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

Thesis Defense

20.06.2016

About

Mathematical Motivation:

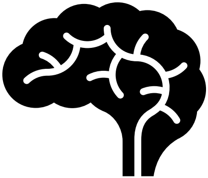
- Large-scale networks and systems
- High-dimensional parametrization
- subject to many-query settings such as optimization

Aim: **Combined State and Parameter Reduction**

- Empirical-Gramian-based [Lall et al.'99]
- Greedy-optimization-based [Liebermann et al.'10]

Neuroscientific Application:

- Dynamic causal modelling [Friston et al.'03]
- Brain region connectivity reconstruction
- from functional neuroimaging data
- by solving an inverse problem



Model Reduction for Control Systems

Control Systems

Time-invariant linear control system:

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

Input: $u : \mathbb{R} \to \mathbb{R}^M$ State: $x : \mathbb{R} \to \mathbb{R}^N$ Output: $y : \mathbb{R} \to \mathbb{R}^Q$

Time-invariant parametric nonlinear control system:

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

Model Order Reduction

Projection-based reduced order model:

$$\begin{aligned}\dot{x}_r(t) &= Vf(Ux_r(t), u(t), \Pi\theta_r) \\ y_r(t) &= g(Ux_r(t), u(t), \Pi\theta_r) \\ x_r(0) &= Vx_0 \\ \theta_r &= \Lambda\theta\end{aligned}$$

State Reduction:

- Trunc. proj. $U \in \mathbb{R}^{N \times n}$
- Trunc. proj. $V \in \mathbb{R}^{n \times N}$
- $VU = \mathbb{1}$

Parameter Reduction:

- Trunc. proj. $\Pi \in \mathbb{R}^{P \times p}$
- Trunc. proj. $\Lambda \in \mathbb{R}^{p \times P}$
- $\Lambda\Pi = \mathbb{1}$

Task: Determine U, V, Π, Λ

Gramian-Based Combined Reduction

Cross Gramian Matrix [Fernando & Nicholson'83, Sorensen & Antoulas'02]

Controllability:

$$\mathcal{C}(u) := \int_{-\infty}^{0} \mathrm{e}^{At} Bu(t)\mathrm{d}t$$

$$\mathcal{C} : L_2^M \to \mathbb{R}^N$$

Observability:

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

$$\mathcal{O} : \mathbb{R}^N \to L_2^Q$$

Hankel operator (maps past inputs to future outputs):

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

Cross Gramian (only applicable to square systems $M = Q$):

$$W_X := \mathcal{C}\mathcal{O} = \int_0^{\infty} \mathrm{e}^{At}\,BC\,\mathrm{e}^{At}\,\mathrm{d}t$$

Cross Gramian's central property:

$$CA^{-1}B = (CA^{-1}B)^{\intercal} \Rightarrow \sigma_i(H) = |\lambda_i(W_X)|$$

From Linear to Parametric Nonlinear Model Reduction

State-space direct truncation:

$$W_X \overset{\text{SVD}}{=} UDV \to U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \to V_1 := U_1^\intercal$$

Extension of the cross Gramian:

- Non-square systems
- Nonlinear (MIMO) systems
- Parametric systems
- Parameter identification

Non-Symmetric Cross Gramian [H. & Ohlberger'15c]

Cross Gramian superposition $B = \big(b_1 \dots b_M\big)$, $C = \big(c_1 \dots c_Q\big)^\intercal$:

$$W_X = \sum_{i=1}^{M=Q} \int_0^\infty e^{At}\, b_i c_i\, e^{At}\, \mathrm{d}t$$

Non-symmetric cross Gramian[a]:

$$W_Z := \sum_{i=1}^{M} \sum_{j=1}^{Q} \int_0^\infty e^{At}\, b_i c_j\, e^{At}\, \mathrm{d}t$$

$$= \int_0^\infty e^{At}\, \big(\sum_{i=1}^{M} b_i\big)\big(\sum_{j=1}^{Q} c_j\big)\, e^{At}\, \mathrm{d}t$$

---

[a] The non-symmetric cross gramian has a connection to tangential interpolation.

Empirical Cross Gramian [H. & Ohlberger'14a,H. & Ohlberger'16]

Empirical linear cross Gramian:

$$W_X = \int_0^\infty \mathrm{e}^{At} BC \, \mathrm{e}^{At} \mathrm{d}t = \int_0^\infty \left(\mathrm{e}^{At} B\right)\left(\mathrm{e}^{A^\intercal t} C^\intercal\right)^\intercal \mathrm{d}t$$

Empirical cross Gramian:

$$\widehat{W}_X := \frac{1}{KLM} \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{m=1}^{M} \frac{1}{c_k d_l} \int_0^\infty \Psi^{klm}(t) \mathrm{d}t \in \mathbb{R}^{N \times N}$$

$$\Psi_{ij}^{klm}(t) = \langle x_i^{km}(t) - \bar{x}_i, y_m^{lj}(t) - \bar{y}_m \rangle$$

- State trajectory $x^{km}(t)$ for the impulse input $u(t) = c_k e_m \delta(t)$.
- Output trajectory $y^{lj}(t)$ for the initial state $x_0 = d_l e_j$.

Empirical Cross Gramian [H. & Ohlberger'14a,H. & Ohlberger'15a]

Empirical cross Gramian:

$$\widehat{W}_X := \frac{1}{KLM} \sum_{k=1}^{K} \sum_{l=1}^{L} \sum_{m=1}^{M} \frac{1}{c_k d_l} \int_0^\infty \Psi^{klm}(t)\mathrm{d}t \in \mathbb{R}^{N \times N}$$

$$\Psi_{ij}^{klm}(t) = \langle x_i^{km}(t) - \bar{x}_i, y_m^{lj}(t) - \bar{y}_m \rangle$$

- For asymptotically stable linear systems: $\widehat{W}_X = W_X$
- Applicable to nonlinear systems since purely **data-driven**
- Extended to arbitrary input (empirical covariance matrix)
- Extended to parametric systems
- Extended to an empircal non-symmetric cross gramian

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

Augmented system (treat parameters as constant states):

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

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

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

Joint Gramian (cross Gramian of the augmented system):

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

Cross-Identifiability Gramian (Schur complement of symmetric part of W_J):

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

Cross-Gramian-Based Combined Reduction

State-space direct truncation:

$$W_X \stackrel{\text{SVD}}{=} UDV \rightarrow U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \rightarrow V_1 := U_1^\intercal$$

Parameter-space direct truncation:

$$W_{\ddot{I}} \stackrel{\text{SVD}}{=} \Pi\Delta\Lambda \rightarrow \Pi = \begin{pmatrix} \Pi_1 & \Pi_2 \end{pmatrix} \rightarrow \Lambda_1 := \Pi_1^\intercal$$

Empirical-cross-Gramian-based reduced order model:

$$\begin{aligned} \dot{x}_r(t) &= V_1 f(U_1 x_r(t), u(t), \Pi_1 \theta_r) \\ y_r(t) &= g(U_1 x_r(t), u(t), \Pi_1 \theta_r) \\ x_r(0) &= V_1 x_0 \\ \theta_r &= \Lambda_1 \theta \end{aligned}$$

Optimization-Based Combined Reduction

Iterative Projection Assembly

Idea: Alternatingly extend parameter and state projection

1. Select parameter vector θ_{I+1} for parameter projection:
   $\Pi_{I+1} = \{\Pi_I \cup (\theta_{I+1} - \Pi_I \Pi_I^\intercal \theta_{I+1})\}$

2. Select state vector $\bar{x}(\theta_{I+1})$ based on parameter θ_{I+1}:
   $U_{I+1} = \{U_I \cup (\bar{x}(\theta_{I+1}) - U_I U_I^T \bar{x}(\theta_{I+1}))\}$

- Parameter selection: Greedy method
- State selection: Energy-based

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

Greedy Sampling Strategy:

- Locate currently worst approximated parameter
- Formulation and solution as optimization problem:

$$\theta_{I+1} = \arg\max_{\theta \in \Theta} \|y(\theta) - y(\Pi_I \Pi_I^T \theta)\|_{L_2}^2 - \beta_2 \|\theta\|_2^2$$

subject to:

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

- Adaptive sampling (no pre-defined grid)
- Tikhonov regularization

State Reduction

In each iteration, after selection of θ_{I+1}:

1. Simulate a state trajectory $x(\theta_{I+1})$
2. Select principal component $\bar{x}(\theta_{I+1})$

- Based on input-to-state mapping $u \to x$
- $U_1 = \arg\min_{U \in \mathbb{R}^{N \times n}, U^\intercal U = \mathbb{1}} \|x(\theta_{I+1}) - U x_r(\theta_{I+1})\|_{L_2}^2$
- POD method-of-snapshots
- proven method for nonlinear systems

Optimization-Based Algorithm [H. & Ohlberger'15b]

Optimization-Based Combined Reduction Algorithm[b]:

1. $\theta_0 \leftarrow \bar{\theta}, \;\; \Pi_0 \leftarrow \theta_0, \;\; U_0 \leftarrow \operatorname{pod}_1(x(\theta_0))$
2. for $l = 1 \dots p$
    3. $\theta_l \leftarrow \arg\min \; -\|y(\theta) - y_r(\theta_r)\|_{L_2}^2 + \beta_2\|\theta\|_2^2$
    4. $\Pi_l \leftarrow \operatorname{orth}(\Pi_{l-1}, \theta_l)$
    5. $U_l \leftarrow \operatorname{orth}(U_{l-1}, \operatorname{pod}_1(x(\theta_l)))$

- Practically, maximization is a minimization with negative cost function
- Typically, no analytical (parameter) derivatives are available
- Computational complexity lies in approximation of derivative information

[b] Orthogonalization orth uses re-iterated Gram-Schmidt.

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

An inverse problem provides data y_d.

Extended cost function with data mismatch as regularization:

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

- Reduced order model is specific to data
- Partial inversion during reduction
- Accelerates projection assembly
- Determining weighting coefficient more complicated

Optimization-Based Combined Reduction

Truncated parameter-space projection:

$$\Pi_{I+1} \to \Lambda_{I+1} := \Pi_{I+1}^{\intercal}$$

Truncated state-space projection:

$$U_{I+1} \to V_{I+1} := U_{I+1}^{\intercal}$$

Greedy-optimization-based reduced order model:

$$\begin{aligned}
\dot{x}_r(t) &= V_{I+1} f(U_{I+1} x_r(t), u(t), \Pi_{I+1}\theta_r) \\
y_r(t) &= g(U_{I+1} x_r(t), u(t), \Pi_{I+1}\theta_r) \\
x_r(0) &= V_{I+1} x_0 \\
\theta_r &= \Lambda_{I+1}\theta
\end{aligned}$$

Implementation

Implementation [H. & Ohlberger'13]

- `emgr` - **Empirical Gramian Framework**

```
W = emgr(f,g,[M,N,Q],[h,T],w,p);
```

- `optmor` - **Optimization-Based Model Order Reduction**

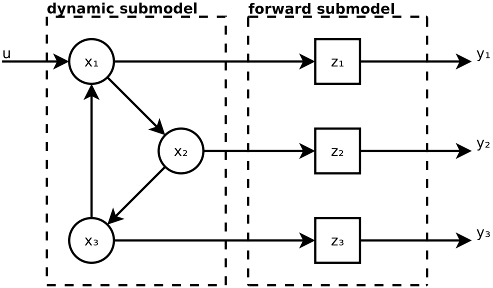
```
[X,P] = optmor(f,g,[M,N,Q],[h,T],p);
```

- Reusable toolboxes
- Released under open-source license
- Compatible with MATLAB and OCTAVE
- Granular configurability
- Modularity: custom solver interface
- Extensively vectorized
- Verification and validation tests

Neuroscientific Application & Numerical Experiments[c]

[c] All experiment source code is available at: doi:10.5281/zenodo.48122 .

fMRI & fNIRS Dynamic Causal Model (fMRI DCM)



Dynamic Submodel:

$$\dot{x}(t) = A(\theta)x(t) + Bu(t)$$

Forward Submodel:

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

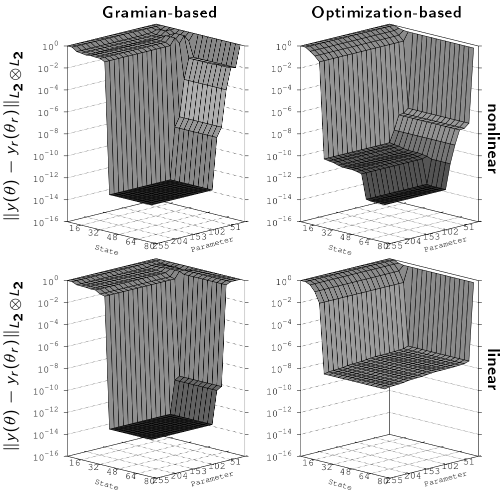
$$\dot{f}_i(t) = s_i(t)$$

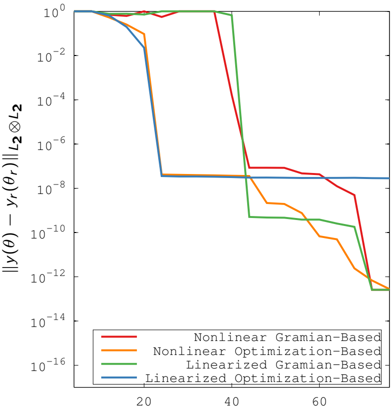
$$\dot{v}_i(t) = \frac{1}{\tau}(f_i(t) - v_i(t)^{\frac{1}{\alpha}})$$

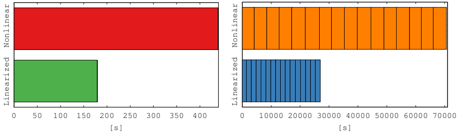
$$\dot{q}_i(t) = \frac{1}{\tau}(\frac{1}{\rho} f_i(t)(1 - ((1 - \rho))^{\frac{1}{f_i(t)}}) - v_i(t)^{\frac{1}{\alpha} - 1} q_i(t))$$

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

Combined Reduction for fMRI DCM

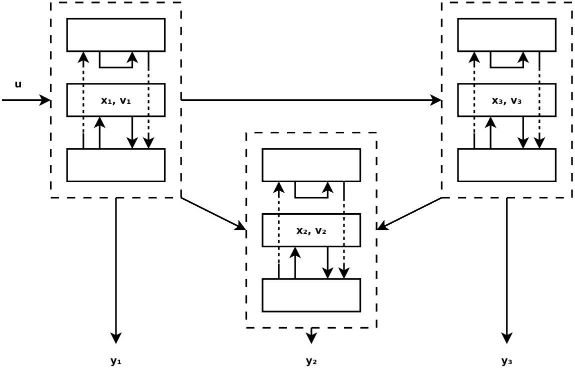






$$\|y(\theta) - y_r(\theta_r)\|_{L_2 \otimes L_2} = \sqrt{\int_\Theta \|y(\theta) - y_r(\theta_r)\|_{L_2}^2 \, \mathrm{d}\theta}$$

EEG & MEG Dynamic Causal Model (EEG DCM)

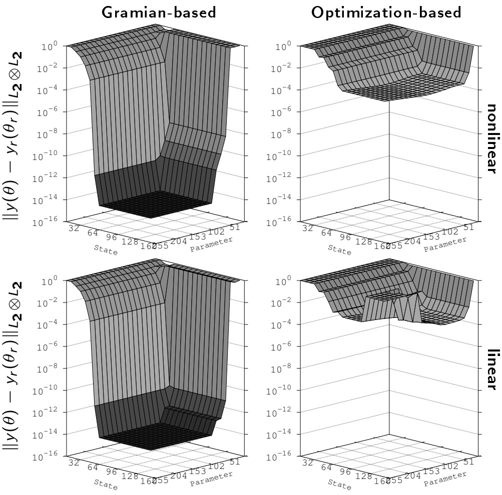


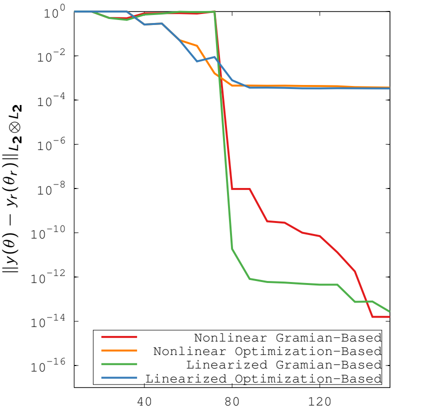
Neural Mass Model:

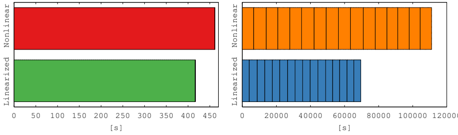
$$\begin{pmatrix} \dot{v}(t) \\ \dot{x}(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1} \\ -T & -T^2 \end{pmatrix} \begin{pmatrix} v(t) \\ x(t) \end{pmatrix} + \begin{pmatrix} 0 \\ A(\theta) \end{pmatrix} \varsigma(Kx(t)) + \begin{pmatrix} 0 \\ B \end{pmatrix} u(t)$$

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

Combined Reduction for EEG DCM

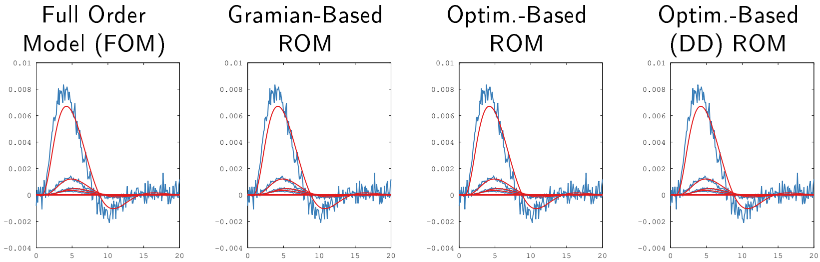






$$\|y(\theta) - y_r(\theta_r)\|_{L_2 \otimes L_2} = \sqrt{\int_\Theta \|y(\theta) - y_r(\theta_r)\|_{L_2}^2 \, \mathrm{d}\theta}$$

Inverse Problem Reduction for fMRI DCM



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
Multiple Breakeven	-	1	10	8

Summary
&
Conclusion

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	$\|\cdot\|_{L_2 \otimes L_2}$	$\|\cdot\|_{L_2 \otimes L_\infty}$
Minimal Complexity	$M + N + P$	$(p-1)(P+1)+1$
Compute Properties	Parallelizable	Sequential
Memory Requirements	Intensive	Economical

Results

Combined state and parameter reduction works

- Both methods are applicable to nonlinear systems
- with high-dimensional parameter-spaces,
- and merely require numerical solutions.
- Suitable to inverse problem reduction

References

- 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)