Westfälische Wilhelms-Universität Münster

Applied Mathematics Münster

Combined State and Parameter Reduction

for the Inversion of Functional Neuroimaging Data

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

WWU Münster
Institute for Computational and Applied Mathematics

CSE 2015
2015-03-15

Connectivity Inference (A problem in systems neuroscience)

Procedure:

- Record Functional Neuroimaging Data
  - fMRI / fNIRS
  - EEG / MEG
- Set Up Parametrized Network Model
  - Dynamic Causal Modelling[1]
- Solve Inverse Problem

Large Networks $\rightarrow$ Many (homogeneous) connectivity parameters.

[1] K. Friston, L. Harrison, W. Penny. Dynamic Causal Modelling. NeuroImage 19(4): 1273–1303, 2003.

Connectivity Inference (A problem in systems neuroscience)

Procedure:

- Record Functional Neuroimaging Data
  - fMRI / fNIRS
  - EEG / MEG
- Set Up Parametrized Network Model
  - Dynamic Causal Modelling[1]
- Solve Inverse Problem

Large Networks $\rightarrow$ Many (homogeneous) connectivity parameters.

Model Reduction!

[1] K. Friston, L. Harrison, W. Penny. Dynamic Causal Modelling. NeuroImage 19(4): 1273–1303, 2003.

Outline

Linear Connectivity Model I

Dynamical System:

$$\dot{x}(t) = Ax(t)$$

Adjacency Matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{2N} & a_{22} & & \vdots \\ \vdots & & \ddots & \\ a_{N1} & & & a_{NN} \end{pmatrix}$$

External Stimuli:

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

Measurements:

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

Linear Connectivity Model II

Connectivity Parametrization:

$$\theta := \mathrm{vec}(A)$$

Parametrized Linear System:

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

Parameter Space Dimension:

$$\dim(\theta) = \dim(x(t))^2$$

Actual Models

fMRI / fNIRS:

$$\begin{pmatrix} \dot{x}(t) \\ \dot{z}_1(t) \\ \vdots \\ \dot{z}_N(t) \end{pmatrix} = \begin{pmatrix} A(\theta)x(t) + Bu(t) \\ f_1(z_1(t), x_1(t)) \\ \vdots \\ f_N(z_N(t), x_N(t)) \end{pmatrix}$$

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_N(t) \end{pmatrix} = \begin{pmatrix} g_1(z_1(t)) \\ \vdots \\ g_N(z_N(t)) \end{pmatrix}$$

EEG / MEG:

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

Inverse Problem

General Nonlinear System:

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

Data Model:

$$y_d = y(\theta) + \varepsilon$$

Parameter Estimation:

$$\tilde{\theta} = \text{argmin}_\theta \, \|y(\theta) - y_d\|_2^2 + \alpha \|\theta\|_K^2$$

Aim

- Fast Inversion
- of Nonlinear Models
- with High-Dimensional Parameter Space
- and High-Dimensional State-Space

Issues:

- $\dim(\theta) \gg 1 \rightarrow$ Costly Inversion
- $\dim(x(t)) \gg 1 \rightarrow$ Costly Individual Forward Solution

Aim

- Fast Inversion
- of Nonlinear Models
- with High-Dimensional Parameter Space
- and High-Dimensional State-Space

Issues:

- $\dim(\theta) \gg 1 \rightarrow$ Costly Inversion
- $\dim(x(t)) \gg 1 \rightarrow$ Costly Individual Forward Solution

Combined State and Parameter Reduction!

Model Order Reduction (State-Reduction)

Setting:

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

Reduced Order Model:

$$\left.\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= g(x(t), u(t)) \end{aligned}\right\} \overset{MOR}{\rightarrow} \begin{cases} \dot{x}_r(t) = f_r(x_r(t), u(t)) \\ y_r(t) = g_r(x_r(t), u(t)) \end{cases}$$

ROM Quality:

$$\|y - y_r\| \ll 1$$

Parametric Model Order Reduction

Setting:

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

Reduced Order Model:

$$\left.\begin{aligned} \dot{x}(t) &= f(x(t), u(t), \theta) \\ y(t) &= g(x(t), u(t), \theta) \end{aligned}\right\} \overset{MOR}{\rightarrow} \begin{cases} \dot{x}_r(t) = f_r(x_r(t), u(t), \theta) \\ y_r(t) = g_r(x_r(t), u(t), \theta) \end{cases}$$

ROM Quality:

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

Combined State and Parameter Reduction

Setting:

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

Reduced Order Model:

$$\left.\begin{aligned} \dot{x}(t) &= f(x(t), u(t), \theta) \\ y(t) &= g(x(t), u(t), \theta) \end{aligned}\right\} \overset{MOR}{\rightarrow} \left\{\begin{aligned} \dot{x}_r(t) &= f_r(x_r(t), u(t), \theta_r) \\ y_r(t) &= g_r(x_r(t), u(t), \theta_r) \end{aligned}\right.$$

ROM Quality:

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

Projection-Based Model Reduction

Galerkin Projections:

$$U \in \mathbb{R}^{N \times n} : U^T U = \mathbb{1}$$
$$\Pi \in \mathbb{R}^{P \times p} : \Pi^T \Pi = \mathbb{1}$$

Reduced Order Model:

$$\begin{aligned} \dot{x}_r(t) &= U^T f(U x_r(t), u(t), \Pi \theta_r) \\ y_r(t) &= \quad\;\; g(U x_r(t), u(t), \Pi \theta_r) \end{aligned}$$

Investigated Methods:

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

Gramian-Based Combined Reduction

- A system-theoretic approach based on system gramians:
  B. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction.
  IEEE Transactions on Automatic Control, 26(1): 17–32, 1981.
- Extended to nonlinear systems through empirical gramians:
  S. Lall, J.E. Marsden and S. Glavaski. Empirical model reduction of controlled nonlinear systems.
  Proceedings of the 14th IFAC Congress, F: 473–478, 1999.
- Focus on the cross gramian:
  K.V. Fernando and H. Nicholson. On the structure of balanced and other principal representations of SISO systems.
  IEEE Transactions on Automatic Control, 28(2): 228–231, 1983.
- and its empirical and parametric variant:
  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.

Controllability & Observability

Linear System:

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

Convolution Operator:

$$y(t) = \int_0^\infty Ce^{A(t-\tau)}Bu(\tau)d\tau = (h * u)(t) =: S(u)(t)$$

Hankel Operator:

$$H(u)(t) := \int_{-\infty}^0 Ce^{A(t-\tau)}Bu(\tau)d\tau = \mathcal{O}\mathcal{C} \to \text{rank}(H) \leq N < \infty$$

- Controllability Operator: $\mathcal{C}(u) = \int_{-\infty}^0 e^{-A\tau}Bu(\tau)dt$
- Observability Operator: $\mathcal{O}(x) = Ce^{At}x$

Cross Gramian

Cross Gramian:

$$W_X := \mathcal{C}\mathcal{O} = \int_0^\infty e^{A\tau} B C e^{A\tau} d\tau$$

Core Property:

$$\mathcal{O}\mathcal{C} = (\mathcal{O}\mathcal{C})^* \Rightarrow \sigma_i(H) = |\lambda_i(W_X)|$$

Direct Truncation:

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

Non-Symmetric Cross Gramian[2]:

$$\widehat{W}_X := \sum_i \sum_j W_{X,SISO}^{i,j}$$

[2] C. Himpe and M. Ohlberger. A Note on the Non-Symmetric Cross Gramian. Submitted, 2015. Preprint at http://arxiv.org/pdf/1501.05519

Nonlinear Systems: Empirical Gramians

Cross Gramian:

$$W_X := \mathcal{C}\mathcal{O} = \int_0^\infty e^{A\tau} B \;\; C e^{A\tau} d\tau$$

- $\mathcal{C}$ - Input-To-State mapping for impulse input.
- $\mathcal{O}$ - State-To-Output mapping for an initial state.

Sampled State and Output Trajectories:

$$X = \begin{pmatrix} x_1 & \dots & x_M \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 & \dots & y_N \end{pmatrix}$$

Empirical Cross Gramian:

$$\widetilde{W}_X := XY^T = W_X \text{ (for linear systems)}$$

$$\rightarrow \widetilde{W}_X = \sum_i X_i Y_i^T \text{ (more accurate for nonlinear systems)}$$

$$\rightarrow \widetilde{W}_X = \sum_j \sum_i X_i(\theta_j) Y_i(\theta_j)^T \text{ (for parametrized systems)}$$

Joint Gramian

Augmented System:

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

Joint Gramian (Cross Gramian of Augmented System):

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

Cross-Identifiability Gramian:

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

Direct Truncation:

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

Optimization-Based Combined Reduction

- Using greedy sampling[3]:
  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.
- *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.*
- in an inverse problem setting:
  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. Himpe and M. Ohlberger. Data-Driven Combined State and Parameter Reduction for Inverse Problems. Submitted 2014.*

---

[3] K. Veroy and A. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: Rigorous reduced-basis a posteriori error bounds. International Journal for Numerical Methods in Fluids, 47:773–788, 2005.

Parameter-Space Projection

Greedy Sampling:

$$\tilde{\theta} = \text{argmax}_{\theta} \, \|y(\theta) - y_r(\theta)\|_2^2$$

Tikhonov Regularization (Prior Knowledge):

$$\tilde{\theta} = \text{argmax}_{\theta} \, \|y(\theta) - y_r(\theta_r)\|_2^2 - \beta \|\theta\|_K^2$$

1. Initialize Parameter Base with $\Pi_0 = \theta_0$
2. Greedily Select: $\theta_I = \text{argmin}_{\theta} \; -\|y(\theta) - y_r(\theta_r)\|_2^2 + \beta \|\theta\|_K^2$
3. Incorporate θ_I into Π_{I-1} by orthogonalization.

State-Space Projection

POD:

$$\bar{x} = \mathrm{pod}_1(x(\theta))$$

POD-Greedy[4]:

$$\bar{x} = \mathrm{pod}_1(x(\theta) - x_r(\theta))$$

1. Initialize State Base with $U_0 = \mathrm{pod}_1(x(\theta_0))$
2. Greedily-Select: $\bar{x}_I = \mathrm{pod}_1(x(\theta_I) - x_r(\theta_I))$
3. Include $\bar{x}_I$ into U_{I-1}

[4] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. M2AN, Math. Model. Numer. Anal., 42(2):277–302, 2008

Extensions:

Data-Driven Regularization:

$$\theta_I = \text{argmax}_\theta \ \ \|y(\theta) - y_r(\theta_r)\|_2^2 - \beta\|\theta\|_K^2 - \gamma\|y_d - y_r(\theta_r)\|_2^2$$

Monte-Carlo Parameter Basis:

$$\theta \approx \sum_{i=1}^{q \ll \dim(\theta)} \hat{\theta}_i M_i, \quad \hat{\theta}_i \in \mathbb{R}, \quad M_i \in \mathcal{U}_{[0,1]}^{\dim(\theta)}, \quad \langle M_i, M_{j \neq i} \rangle = 0$$

$$\hat{\theta}_I = \text{argmax}_{\hat{\theta}} \, \|y(\hat{\theta}) - y_r(\hat{\theta}_r)\|_2^2 - \beta\|\hat{\theta}\|_K^2$$

Algorithm

Algorithm 1: Optimization-Based Combined Reduction[5]

$\theta_0 \longleftarrow \bar{\theta}$
$\Pi_0 \longleftarrow \theta_0$
$U_0 \longleftarrow \mathrm{pod}_1(x(\theta_0))$
for $I = 1 : R$ **do**
$\quad \hat{\theta}_I \longleftarrow \mathrm{argmin}_{\hat{\theta}} \ -\|y(\hat{\theta}) - y_r(\hat{\theta}_r)\|_2^2 + \beta\|\hat{\theta}\|_K^2 + \gamma\|y_d - y_r(\hat{\theta}_r)\|_2^2$
$\quad \Pi_I \longleftarrow \mathrm{orth}(\Pi_{I-1}, [M_1 \dots M_q]\theta_I)$
$\quad U_I \longleftarrow (U_{I-1}, \mathrm{pod}_1(x(\theta_I) - x_r(\theta_I))$

[5] Also applicable to nonlinear systems, since purely snapshot based!

Numerical Experiments

1. Generate Synthetic Data Timeseries
   - Random but Stable Network
   - Add Gaussian Noise
2. Combined Reduction
   - Non-Symmetric Empirical Joint Gramian
   - Data-Driven and Monte-Carlo Parameter Basis
3. Inversion of ROMs

$$\tilde{\theta}_r = \text{argmin}_{\theta_r} \|y_d - y_r(\theta_r)\|_2^2 + \alpha \|\theta_r\|_{K_r}^2$$

Hyperbolic Network Model[6]

Model:

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

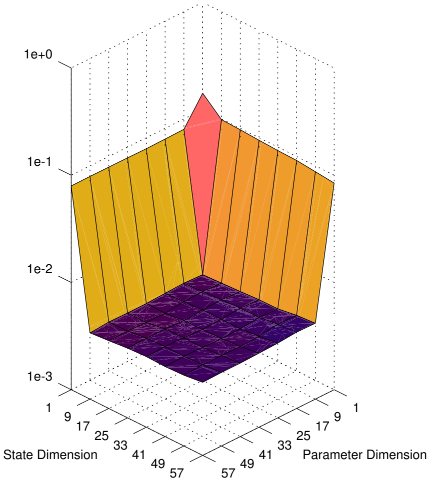
Parametrization:

$$A(\theta) := -\mathbb{1}_N + \text{vec}^{-1}(\theta)$$

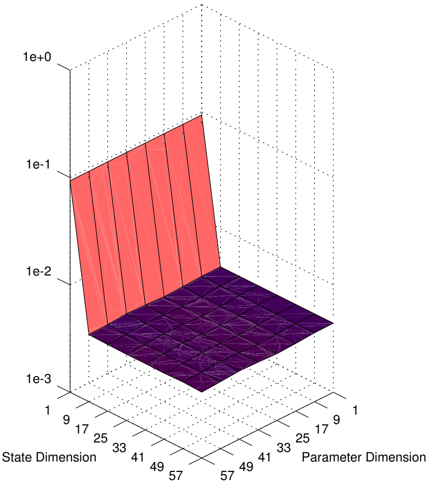
- Nonlinear SIMO System
- $\dim(x(t)) = 64 \Rightarrow \dim(\theta) = 4096$

[6] Y. Quan, H. Zhang and L. Cai. "Modeling and control based on a new neural network model". Proceedings of the American Control Conference 3: 1928–1929, 2001.

Relative L_2 Output Error (Forward Problem)

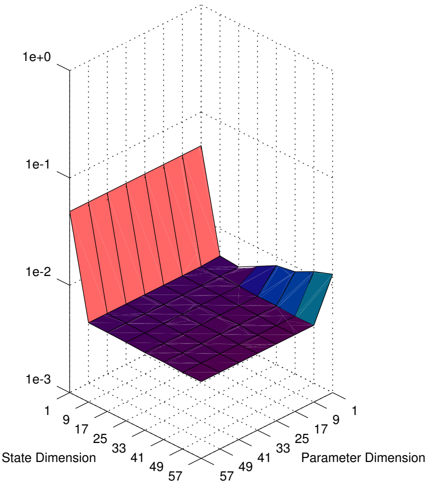


Gramian-Based

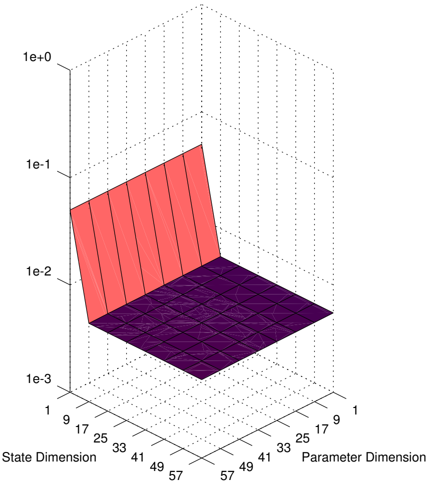


Optimization-Based

Relative L_2 Output Error (Inverse Problem)



Gramian-Based



Optimization-Based

Outlook

- Actual Nonlinear Models (Both)
- Adaptive Parameter Sampling Strategy (Gramian-Based)
- Systematic Regularization Heuristics (Optimization-Based)

tl;dl

- Combined State and Parameter Reduction by
- Gramian-Based and
- Optimization-Based Methods
- for the Inversion of Functional Neuroimaging Data[7].

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

Thanks!

[7] **Get the companion code**: `http://j.mp/siamcse15`

Shameless Plug

Let's meet at my poster:

emgr - Empirical Gramian Framework