Accelerating the Computation of Empirical Gramians

and Related Methods

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

WWU Münster
Institute for Computational and Applied Mathematics

5th IWMRRF
2015-06-29

Motivation

Research Question:

- Is it possible to reconstruct a large network,
- stimulated by low-dimensional stimuli,
- and low-dimensional measurements.
- Let's say: the brain

Mathematical Problem:

- Nonlinear (!)
- Parametric
- Large-Scale ODEs / Discretized PDEs
- Many-Query Setting
- Such as: Model-Constrained Optimization

MRRF?

What does this have to do with
"model reduction in reactive flows"?

Same models if investigating
input-output behaviour of a flow!

See for example: [Bagheri et al'09], [Holmes et al'12], [Nguyen et al'14].

Outline

Control Systems

Linear Control System:

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

$(x(t) \in \mathbb{R}^{\boldsymbol{N}}, u(t) \in \mathbb{R}^{\boldsymbol{M}}, y(t) \in \mathbb{R}^{\boldsymbol{O}}, A \in \mathbb{R}^{\boldsymbol{N} \times \boldsymbol{N}}, B \in \mathbb{R}^{\boldsymbol{N} \times \boldsymbol{M}}, C \in \mathbb{R}^{\boldsymbol{O} \times \boldsymbol{N}})$

General Control System:

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

$(x(t) \in \mathbb{R}^{\boldsymbol{N}}, u(t) \in \mathbb{R}^{\boldsymbol{M}}, y(t) \in \mathbb{R}^{\boldsymbol{O}}, f : \mathbb{R}^{\boldsymbol{N}} \times \mathbb{R}^{\boldsymbol{M}} \to \mathbb{R}^{\boldsymbol{N}}, g : \mathbb{R}^{\boldsymbol{N}} \times \mathbb{R}^{\boldsymbol{M}} \to \mathbb{R}^{\boldsymbol{O}}, \theta \in \mathbb{R}^{\boldsymbol{P}})$

Model Order Reduction

General Control System:

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

(In a many query setting, numerous solutions for varying x_0, u, θ are required)

Common System Dimensions:

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

(Important is a mapping $u \mapsto y$)

Projection-Based Combined Reduction

Reduced Order Model (ROM):

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

$(x_r(t) \in \mathbb{R}^n, \theta_r \in \mathbb{R}^p, \dim(x_r(t)) \ll \dim(x(t)), \dim(\theta_r) \ll \dim(\theta), \|y(\theta) - y_r(\theta_r)\| \ll 1)$

(Galerkin) Projection-Based ROM:

$$\dot{x}_r(t) = U^T f(Ux_r(t), u(t), Q\theta_r)$$
$$y_r(t) = g(Ux_r(t), u(t), Q\theta_r)$$
$$x_r(0) = U^T x_0$$

$(U \in \mathbb{R}^{N \times n}, U^T U = \mathbb{1}, Q \in \mathbb{R}^{P \times p}, \theta_r = Q^T \theta)$

Cross-Gramian-Based Model Reduction

Controllability & Observability:

$$\mathcal{C}(u) := \int_0^\infty e^{At} Bu(t) dt$$
$$\mathcal{O}(x) := Ce^{At} x$$

Cross Gramian (square systems only):

$$W_X := \mathcal{C}\mathcal{O} = \int_0^\infty e^{At} BCe^{At} dt$$

$(AW_{\boldsymbol{X}} + W_{\boldsymbol{X}} A = -BC)$

Direct Truncation (symmetric systems only):

$$W_X \stackrel{TSVD}{=} UDV$$

$(U \in \mathbb{R}^{\boldsymbol{N} \times \boldsymbol{n}}, D \in \mathbb{R}^{\boldsymbol{n} \times \boldsymbol{n}}, V \in \mathbb{R}^{\boldsymbol{n} \times \boldsymbol{N}}, D_{\boldsymbol{ii}} \approx \sigma_{\boldsymbol{i}}(H)$, all SISO systems are symmetric)

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

Linear Empirical Cross Gramian:

$$W_X = \int_0^\infty (e^{At} B)\,(e^{A^T t} C^T) dt$$

Nonlinear Empirical Cross Gramian[1]:

$$\widehat{W}_X = \frac{1}{|Q_u||R_u|M|Q_x||R_x|} \sum_{h=1}^{|Q_{\boldsymbol{u}}|} \sum_{i=1}^{|R_{\boldsymbol{u}}|} \sum_{j=1}^{M} \sum_{k=1}^{|Q_{\boldsymbol{x}}|} \sum_{l=1}^{|R_{\boldsymbol{x}}|} \frac{1}{c_h d_k} \int_0^\infty T_l \Psi^{hijkl}(t) T_l^T \mathrm{d}t$$

$$\Psi_{ab}^{hijkl}(t) = f_b^T T_k^T \Delta x^{hij}(t) e_i^T S_h^T \Delta y^{kla}(t) \in \mathbb{R},$$

$$\Delta x^{hij}(t) = (x^{hij}(t) - \bar{x}^{hij}),$$

$$\Delta y^{kla}(t) = (y^{kla}(t) - \bar{y}^{kla}).$$

---

[1] [H. & Ohlberger'15] proposes a non-symmetric extension to the cross gramian, which can be efficiently computed for the empirical cross gramian.

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

Parameter augmented system (assume constant parameters):

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

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

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

Cross Gramian parameter augmented system (joint gramian W_J):

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

Schur-complement of symmetric part of W_J (cross-identifiability gramian $W_{\ddot{I}}$):

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

Trajectories for Nonlinear Systems

Single-Step Methods:

- Averages vector field evaluations of intermediate time-steps
- i.e. Runge-Kutta methods

Multi-Step Methods:

- Averages vector field evaluations of past time-steps
- Require s starting values, i.e. by single-step methods
- i.e. Adams-Bashforth methods

General Linear Methods:

- Averages vector field evaluations of past and intermediate time-steps
- i.e. Two-Step Runge-Kutta (TSRK) methods

Improved Runge-Kutta Methods [Rabiei'13]

TSRK Example: IRK3

$$
\begin{aligned}
k_{-1} &= hf(t_{i-1}, y_{i-1}) \\
k_{-2} &= hf(t_{i-1} + \frac{1}{2}h, y_{i-1} + \frac{1}{2}k_{i-1}) \\
k_1 &= hf(t_i, y_i) \\
k_2 &= hf(t_i + \frac{1}{2}h, y_i + \frac{1}{2}k_i) \\
y_{k+1} &= y_k + \frac{2}{3}k_1 - (-\frac{1}{3}k_1) + \frac{5}{6}(k_2 - k_{-2})
\end{aligned}
$$

Starting Values [LeVeque'07]:

- Ensure order of one-step-error
- 2nd order explicit midpoint-rule, because: same sample points
- Two starting values yield more accurate results for TSRK

Numerical Experiment Setup

Hyperbolic Network Model:

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

($K \in \mathbb{R}^{N \times N}, |K_{ii}| \leq 1, K_{ij, i \neq j} = 0$)

System Dimensions:

- $\dim(u(t)) = 1$
- $\dim(x(t)) \in \{4, 16, 256, 1024\}$
- $\dim(y(t)) = 1$
- $\dim(\theta) = \dim(x(t))$
- 100 Time-Steps

(In PDE terms: $\dim(x(t)) = 1024 \Rightarrow 102400$ DOFs)

Numerical Comparison (I)

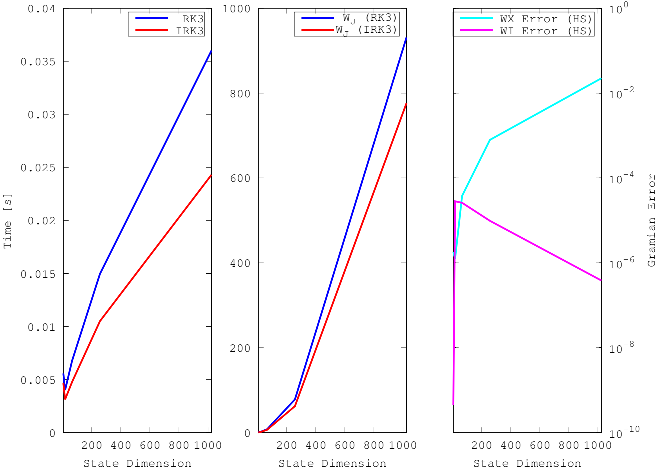

Figure : Comparison of Timings and Frobenius Norm.

Generalized Transpositions

The Tensor Transpose:

- Given an n-Tensor A, a n-Tensor B is called its transpose if
- $B_{i_{\sigma(1)},\ldots,i_{\sigma(n)}} = A_{i_1,\ldots,i_n}$
- for a permutation $\sigma \neq \mathbb{1}$

For example:

- Given a 3-tensor
- it has $n! - 1 = 5$ transposes
- which corresponds to "rotations"

Empirical Cross Gramian

Essence of the Empirical Cross Gramian:

$$W_X = \sum_k \omega_X^k, \quad \omega_{X,ij}^k = \langle x_i^k(t), y_k^j(t) \rangle$$

$$\begin{array}{ll}
\begin{pmatrix} x^3(t_1) & \dots & x^3(t_T) \end{pmatrix} & \begin{pmatrix} y^6(t_1) & \dots & y^6(t_T) \end{pmatrix} \\
\begin{pmatrix} x^2(t_1) & \dots & x^2(t_T) \end{pmatrix} & \begin{pmatrix} y^5(t_1) & \dots & y^5(t_T) \end{pmatrix} \\
\begin{pmatrix} x^1(t_1) & \dots & x^1(t_T) \end{pmatrix} & \begin{pmatrix} y^4(t_1) & \dots & y^4(t_T) \end{pmatrix} \\
 & \begin{pmatrix} y^3(t_1) & \dots & y^3(t_T) \end{pmatrix} \\
 & \begin{pmatrix} y^2(t_1) & \dots & y^2(t_T) \end{pmatrix} \\
 & \begin{pmatrix} y^1(t_1) & \dots & y^1(t_T) \end{pmatrix}
\end{array}$$

Empirical Cross Gramian

Essence of the Empirical Cross Gramian:

$$W_X = \sum_k \omega_X^k, \quad \omega_{X,ij}^k = \langle x_i^k(t), y_k^j(t) \rangle$$

$\omega_{X,11}^1$:

$$\begin{pmatrix} x^3(t_1) & \dots & x^3(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x^2(t_1) & \dots & x^2(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x_1^1(t_1) & \dots & x_1^1(t_T) \end{pmatrix}$$

$$\begin{pmatrix} y^6(t_1) & \dots & y^6(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y^5(t_1) & \dots & y^5(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y^4(t_1) & \dots & y^4(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y^3(t_1) & \dots & y^3(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y^2(t_1) & \dots & y^2(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y_1^1(t_1) & \dots & y_1^1(t_T) \end{pmatrix}$$

Empirical Cross Gramian

Essence of the Empirical Cross Gramian:

$$W_X = \sum_k \omega_X^k, \quad \omega_{X,ij}^k = \langle x_i^k(t), y_k^j(t) \rangle$$

$$\begin{pmatrix} y_3^1(t_1) \dots y_3^1(t_T) \\ \vdots \\ y_3^3(t_1) \dots y_3^3(t_T) \end{pmatrix}$$

$$\begin{pmatrix} x^3(t_1) & \dots & x^3(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x^2(t_1) & \dots & x^2(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x^1(t_1) & \dots & x^1(t_T) \end{pmatrix}$$

$$\begin{pmatrix} y_2^1(t_1) \dots y_2^1(t_T) \\ \vdots \\ y_2^3(t_1) \dots y_2^3(t_T) \end{pmatrix}$$

$$\begin{pmatrix} y_1^1(t_1) \dots y_1^1(t_T) \\ \vdots \\ y_1^3(t_1) \dots y_1^3(t_T) \end{pmatrix}$$

Empirical Cross Gramian

Essence of the Empirical Cross Gramian:

$$W_X = \sum_k \omega_X^k, \quad \omega_{X,ij}^k = \langle x_i^k(t), y_k^j(t) \rangle$$

ω_X^1 :

$$\begin{pmatrix} x^3(t_1) & \dots & x^3(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x^2(t_1) & \dots & x^2(t_T) \end{pmatrix}$$
$$\begin{pmatrix} x^1(t_1) & \dots & x^1(t_T) \end{pmatrix}$$

$$\begin{pmatrix} y_3^1(t_1) \dots y_3^1(t_T) \\ \vdots \\ y_3^N(t_1) \dots y_3^N(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y_2^1(t_1) \dots y_2^1(t_T) \\ \vdots \\ y_2^N(t_1) \dots y_2^N(t_T) \end{pmatrix}$$
$$\begin{pmatrix} y_1^1(t_1) \dots y_1^1(t_T) \\ \vdots \\ y_1^N(t_1) \dots y_1^N(t_T) \end{pmatrix}^T$$

Numerical Comparison (II)

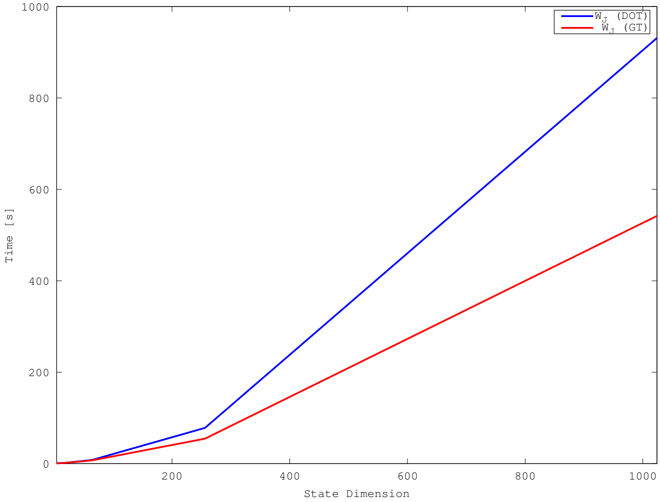

Figure : Comparison of Timings.

Approximate Schur-Complement [Wu et al.'13]

Schur-complement of symmetric part of W_J:

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

(Cross-identifiability gramian $W_{\ddot{I}}$ - observability of parameters)

Neumann Series Represenation of Matrix Inverse:

$$A^{-1} = \sum_{k=0}^{\infty}(\mathbb{1} - A)^k$$

(Truncate for an approximate inverse)

Approximate Inverse in $\mathcal{O}(N^2)$ flops:

$$A^{-1} \approx D^{-1} - D^{-1}ED^{-1}$$

$(A = D + E, D_{ij} = \delta_{ij}A_{ij}, E = A - D)$

Numerical Comparison (III)

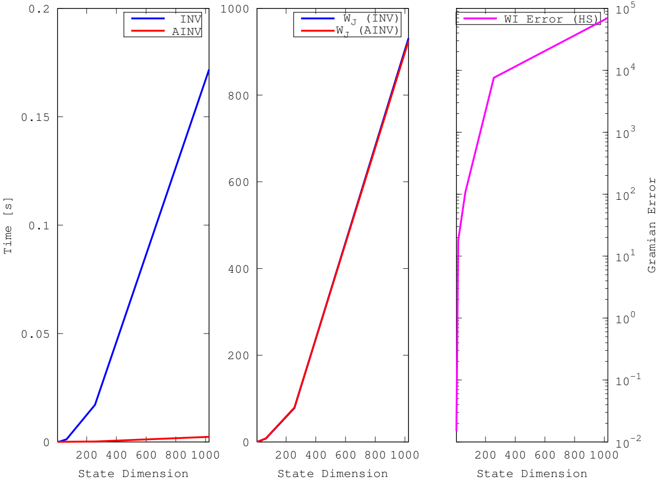

Figure : Comparison of Timings and Frobenius Norm.

TSVD via Lanczos

Algorithmic phases:

1. Tridiagonalization
2. Small eigenvalue problem

Properties:

- Iterative (Common strategy for eigenvalue problems)
- Requires only operator evaluations (Similar to the power method)
- Determines dominant singular values (Or eigenvalues respectively)
- Adaptive and bounded variants available (Useful for MOR)

(Double) Re-Orthogonalization

Re-Orthogonalization I [Chen & Saad'09]:

- During tridiagonalization,
- using Gram-Schmidt.

Re-Orthogonalization II:

- Post-processing of (left) singular vectors,
- by an economic (low-dimensional) QR-decomposition

Numerical Comparison (IV)

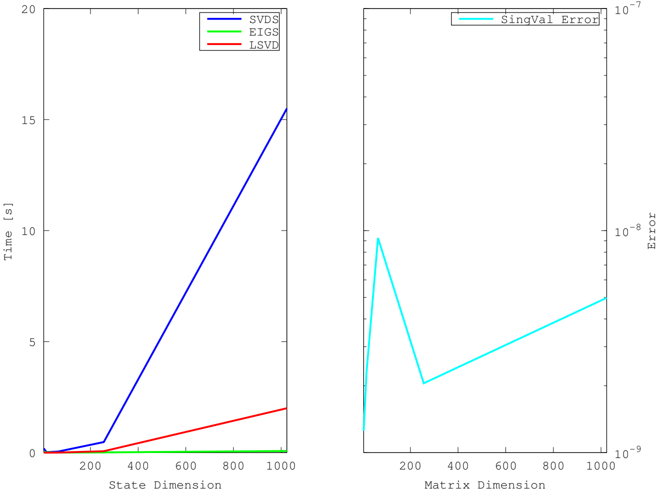

Figure : Comparison of Timings and Euclidean Norm.

Alltogether

"Legacy" Empirical Joint Gramian:

- Ralston's 3rd Order Runge-Kutta Method
- Component-Wise Dot-Products
- $\mathcal{O}(N^3)$ Inverse
- SVD via eigendecomposition of Gramian

vs.

Improved Empirical Joint Gramian:

- 3rd Order Improved Runge Kutta Method
- Generalized Transpositions
- $\mathcal{O}(N^2)$ Approximate Inverse
- Double Re-Orthgonalized Lanczos SVD

Numerical Comparison (V)

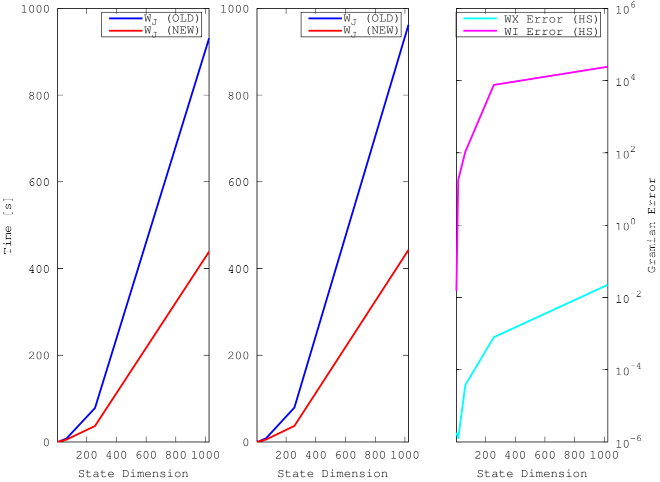

Figure : Comparison of Timings and Frobenius Norm.

tl;dl

- Empirical Cross Gramian for
- Combined State and Parameter Reduction
- Accelerated by
- `irk`, `permute`, `ainv` and `lsvd`.

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

Thanks[2,3]!

[2] Get the companion code: `http://j.mp/iwmrrf5`

[3] If you have large test problems consider contributing them as a benchmark to the MORwiki: `http://modelreduction.org`