



MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG



COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

Réduction de Modèle à Base de Grammaires pour Système Linéaire et Non Linéaire

Christian Himpe

38e École Internationale d'été de Contrôle Automatique
Approximation des Systèmes Dynamiques à Grande Échelle

2017-09-11



MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG



COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

Gramian-Based Model Reduction

for Linear and Nonlinear Systems

Christian Himpe

38th International Summer School of Automatic Control
Approximation of Large-Scale Dynamical Systems

2017-09-11

- Given a differential equation model:
- Algorithmic computation of simple surrogate models,
- preserving essential features of the original model.

- Enable computational simulation
- Reduce numerical complexity
- Multi-query / Many-query settings:
 - Optimization
 - Control
 - Sensitivity Analysis
 - Uncertainty Quantification
 - Computer-Aided Design
- Gain new insights
- Surrogate modelling

- Mechanical problems
- Electrical circuits
- Fluid dynamics
- Inference of brain connectivity
- Scenario analysis for gas networks¹

¹BMW funded MathEnergy project.

- Moment Matching
- Transfer Function Interpolation
- Matrix Interpolation
- Hankel Norm Approximation
- Proper Orthogonal Decomposition
- Balanced Proper Orthogonal Decomposition
- Balanced Truncation
- Approximate Balancing
- Singular Perturbation
- Balanced Gains
- Empirical Gramians
- ...

- Encode **relevant** information about a system
- in operators (matrices) with specific **attributes**,
- which induce **projections**
- to spaces with **sorted** components.
- **Truncated** projections then
- map to a **meaningful** subspace
- inducing a **reduced order model**.



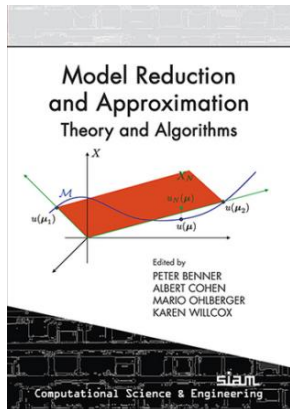
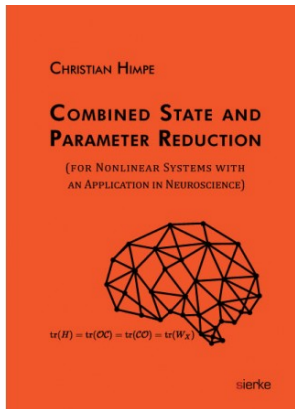
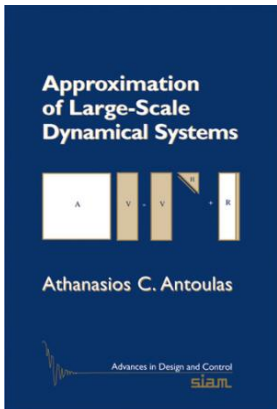
1. System-Theoretic Preliminaries
2. Linear Gramian-Based Model Reduction
3. Nonlinear Gramian-Based Model Reduction
4. Combined Reduction in Action
5. Other Uses for System Gramians

- A.C. Antoulas, D.C. Sorensen, and S. Gugercin. **A survey of model reduction methods for large-scale systems**. In Structured Matrices in Mathematics, Computer Science, and Engineering I, Contemporary Mathematics, vol. 280: 193–219, 2001.
- A.C. Antoulas. **An overview of approximation methods for large-scale dynamical systems**. Annual Reviews in Control, 29(2): 181–190, 2005.
- U. Baur, P. Benner, and L. Feng. **Model Order Reduction for Linear and Nonlinear Systems: A System-Theoretic Perspective**. Archives of Computational Methods in Engineering, 21(4): 331–358, 2014.
- P. Benner, S. Gugercin, and K. Willcox. **A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems**. SIAM Reviews, 57(4): 483–531, 2015.
- C. Himpe. **emgr - The Empirical Gramian Framework**. arXiv e-print, cs.MS: 1611.00675, 2016.



CSC

Useful Books





Model Order Reduction Wiki:

`http://modelreduction.org`

- Methods
- Benchmarks
- Software
- Event Calendar
- Comprehensive bib



Before We Start

- The slides will be online.
- Ask questions.
- There is a hands-on lab session tomorrow!

1. Motivating Model Reduction
2. Dynamical Systems Summary
3. Basic System Theory
4. Generic Model Reduction
5. Principal Axis Transformation

Differential equation:

$$F(s, x, D^\alpha x) = 0$$

- Multi-index differential operator: $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$.
- Analytical solution: $x(s)$.
- Discrete approximate solution: $x_h(s_h)$.
- Large-Scale discretized dimension: $\dim(x_h(s_h)) \gg 1$.
- Example: 1D heat equation: $F(x, \frac{\partial^2 x}{\partial s^2})$.

Continuous dynamical system over \mathbb{R} :

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

- Time: $t \in \mathbb{R}^+$
- State: $x : \mathbb{R} \rightarrow \mathbb{R}^N$
- State-space: \mathbb{R}^N , $N < \infty$
- Vector field: $f : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$
- Ordinary Differential Equation (ODE): $\dot{x} = \frac{\partial x}{\partial t}$
- Initial Value Problem (IVP): $x_0 = x(t_0)$

Linear time invariant homogenous system of ODEs:

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

IVP Solution:

$$x(t) = e^{At} x_0$$

Fundamental Solution:

$$L(\cdot)(t) = e^{At}$$

- System matrix: $A \in \mathbb{R}^{N \times N}$
- Matrix exponential: $e^A = \exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$

Linear time invariant inhomogenous system:

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

Duhamel's principle:

$$x(t) = L(x_0)(t) + (L * b)(t) = e^{At} x_0 + \int_0^t e^{A\tau} b(\tau) d\tau$$

- Continuous inhomogeneity: $b : \mathbb{R} \rightarrow \mathbb{R}^N$, $b \in C^0$.

Stability:

$$\|x(t)\| < \infty, \forall t > 0$$

Lyapunov stability:

$$\forall \epsilon > 0, \exists \delta > 0 : \|x(t) - \bar{x}\| < \epsilon, \forall t \geq 0, \forall x_0 : \|x_0 - \bar{x}\| < \delta$$

Lyapunov stability (asymptotic):

$$\forall \epsilon > 0, \forall \delta > 0 : \|x(t) - \bar{x}\| < \epsilon, \forall t \geq T, \forall x_0 : \|x_0 - \bar{x}\| < \delta$$

Lyapunov stability (exponential):

$$\exists c_1, c_2 \in \mathbb{R} : \|x(t) - \bar{x}\| < c_1 e^{-c_2 t} \|x_0 - \bar{x}\|, \forall t > 0$$

- Steady state \bar{x} : $f(t, \bar{x}) = 0$

Global stability via spectrum:

$$\begin{aligned} & A \in \mathbb{R}^{N \times N} \text{ Hurwitz} \\ \Leftrightarrow & \operatorname{Re}(\lambda_i(A)) < 0, i = 1 \dots N \\ \Leftrightarrow & x \text{ exponentially stable} \\ \Leftrightarrow & x \text{ asymptotically stable} \end{aligned}$$



Local stability via linearization:

$$f \in C^1, \bar{x} \in \mathbb{R}^N : f(\bar{x}) = 0$$

$$\rightarrow A_\ell := \frac{\partial f}{\partial x}(\bar{x})$$

$$\Rightarrow \dot{x}_\ell(t) = A_\ell x_\ell(t) \approx f(x(t)), \|x(t) - \bar{x}\| < \epsilon$$

$\rightarrow x_\ell$ globally stable $\Rightarrow x$ locally stable



Nonlinear parametric control system:

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

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

- Input: $u : \mathbb{R} \rightarrow \mathbb{R}^M$
- State: $x : \mathbb{R} \rightarrow \mathbb{R}^N$
- Output: $y : \mathbb{R} \rightarrow \mathbb{R}^Q$
- Parameter: $\theta \in \mathbb{R}^P$
- Vector field: $f : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^Q \rightarrow \mathbb{R}^N$
- Output functional: $g : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^Q \rightarrow \mathbb{R}^Q$



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

- $M = Q \rightarrow$ Square system
- $M = Q = 1 \rightarrow$ SISO system
- $M = 1, Q > 1 \rightarrow$ SIMO system
- $M > 1, Q = 1 \rightarrow$ MISO system
- $M > 1, Q > 1 \rightarrow$ MIMO system

Linear time-invariant (LTI) control system:

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

$$y(t) = Cx(t) + Du(t)$$

- System matrix: $A \in \mathbb{R}^{N \times N}$
- Input matrix: $B \in \mathbb{R}^{N \times M}$
- Output matrix: $C \in \mathbb{R}^{Q \times N}$
- Feed-forward matrix: $D \in \mathbb{R}^{Q \times M}$
- Here: Assume $D = 0$

Impulse Input:

$$u(t) = \delta(t) = \begin{cases} \Delta^{-1} & t < \Delta \\ 0 & t \geq \Delta \end{cases}$$

LTI Impulse Response:

$$g(t) = \begin{cases} C e^{At} B & t \geq 0 \\ 0 & t < 0 \end{cases}$$

- Fundamental solution: $y(t) = (g * u)(t)$.
- $\int_0^{\infty} \delta(t) dt = 1$
- $g(t < 0) = 0$ ensures causality.

Bounded-input-bounded-output (BIBO) stability:

$$\|u\|_{L_\infty} < \infty \rightarrow \|y\|_{L_\infty} < \infty$$

Young's Inequality²:

$$\|y\|_{L_\infty} = \|g * u\|_{L_\infty} < \|g\|_{L_1} \|u\|_{L_\infty}$$

- $\|g\|_{L_1} < \infty \Rightarrow$ BIBO stable

²W.H. Young. **On the Multiplication of Successions of Fourier Constants.** Proceedings of the Royal Society of London, Series A, 87(596): 331–339, 1912.



Adjoint Operator:

$$\begin{aligned}\langle y, g * u \rangle &= \langle \hat{g} * y, u \rangle \\ \Rightarrow \int_{-\infty}^{\infty} y^T(t) \left(\int_{-\infty}^{\infty} C e^{A(t-\tau)} B u(\tau) d\tau \right) dt \\ &= \int_{-\infty}^{\infty} u^T(t) \left(\int_{-\infty}^{\infty} B e^{A(t-\tau)} C y(t) dt \right) d\tau \\ &= \int_{-\infty}^{\infty} u^T(t) \left(- \int_{\infty}^{-\infty} B e^{-A(\tau-t)} C y(t) dt \right) d\tau\end{aligned}$$

Adjoint System (backward running):

$$\begin{aligned}\dot{z}(\tau) &= -A^T z(\tau) - C^T y(\tau) \\ u(\tau) &= B^T z(\tau)\end{aligned}$$

Transfer function:

$$G(s) = C(\mathbb{1}s - A)^{-1}B$$

Properties:

- Laplace transformation of the impulse response.
- Rational function for LTI systems.
- Frequency response $G(s)$, for frequency $s \in [0, \infty)$.

Symmetric System:

$$G(s) = (G(s))^T$$

$$\Leftrightarrow g(t) = (g(t))^T$$

$$\Leftrightarrow (CA^k B) = (CA^k B)^T$$

$$\Leftrightarrow \exists J = J^T \in \mathbb{R}^{N \times N} : AJ = JA^T, B = JC^T$$

State-Space Symmetric System ($J = \mathbb{1}$):

$$A = A^T, B = C^T$$

- All SISO systems are symmetric!

Abstract Input-Output System:

$$u \mapsto x \mapsto y$$

Typical properties:

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

Is there a shortcut?



Generic Reduced Order Model:

$$\dot{x}_r(t) = f_r(x_r(t), u(t), \theta_r)$$

$$y_r(t) = g_r(x_r(t), u(t), \theta_r)$$

$$x_r(0) = x_{r,0}$$

- Reduced State: $x_r : \mathbb{R} \rightarrow \mathbb{R}^n$, $n \ll N$
- Reduced Parameter: $\theta_r \in \mathbb{R}^p$, $p \ll P$
- Approximate Output: $y_r : \mathbb{R} \rightarrow \mathbb{R}^Q$
- Reduced Vector Field: $f_r : \mathbb{R}^n \times \mathbb{R}^M \times \mathbb{R}^p$
- Reduced Output Functional: $g_r : \mathbb{R}^n \times \mathbb{R}^M \times \mathbb{R}^p$

Linear Projection:

$$V \in \mathbb{R}^{N \times N}, V^2 = V$$

Truncated Projection:

$$\text{rank}(V) < N \Rightarrow V \sim (V_1 \ 0)$$

- Truncated projection: $V_1 \in \mathbb{R}^{n \times N}$
- Petrov-Galerkin: $U_1 \in \mathbb{R}^{N \times n}, V_1 U_1 = \mathbb{1} \in \mathbb{R}^{n \times n}$
- Galerkin: $U_1 = V_1^T, U_1$ orthogonal.

Approximate Trajectory:

$$x_r(t) := V_1 x(t) \Rightarrow x(t) \approx U_1 x_r(t)$$

- Reducing truncated projection: $V_1 \in \mathbb{R}^{n \times N}$
- Reconstructing truncated projection: $U_1 \in \mathbb{R}^{N \times n}$
- **Task:** Find projections U_1 and V_1

State-Space Reduced Model:

$$\dot{x}_r(t) = V_1 f(U_1 x_r(t), u(t))$$

$$y_r(t) = g(U_1 x_r(t), u(t))$$

$$x_r(0) = V_1 x_0$$

- Reduced vector field: $f_r := V_1 \circ f \circ U_1$
- Reduced output functional: $g_r := g \circ U_1$
- $\dim(x_r(t)) \ll \dim(x(t))$
- $\|y - y_r\| \ll 1$
- Requires evaluations of full order f, g !

Linear State-Space Reduced Model:

$$\dot{x}_r(t) = V_1 A U_1 x_r + V_1 B u(t) = A_r x_r(t) + B_r u(t)$$

$$y_r(t) = C U_1 x_r(t) = C_r x_r(t)$$

$$x_r(0) = V_1 x_0$$

Linear Reduced Order Model:

- Reduced system matrix: $A_r \in \mathbb{R}^{n \times n}$
- Reduced input matrix: $B_r \in \mathbb{R}^{n \times M}$
- Reduced output matrix: $C_r \in \mathbb{R}^{Q \times n}$

Affine State-Space Reduced Model:

$$\dot{x}_r(t) = V_1 f(\bar{x} + U_1 x_r(t), u(t))$$

$$y_r(t) = g(\bar{x} + U_1 x_r(t), u(t))$$

$$x_r(0) = V_1(x_0 - \bar{x})$$

- $\bar{x} \in \mathbb{R}^N$, often a steady state or mean.
- x becomes perturbation from \bar{x} .
- Affine subspace usually better than linear subspace.

Parametric State-Space Reduced Model:

$$\dot{x}_r(t) = V_1 f(U_1 x_r(t), u(t), \theta)$$

$$y_r(t) = g(U_1 x_r(t), u(t), \theta)$$

$$x_r(0) = V_1 x_0$$

- State-space reduction preserving θ dependency.
- U_1, V_1 are global projections.
- $\|y(\theta) - y_r(\theta)\| \ll 1$

Parameter Reduced Model:

$$\dot{x}(t) = f(x(t), u(t), \Pi_1 \theta_r)$$

$$y(t) = g(x(t), u(t), \Pi_1 \theta_r)$$

$$x(0) = x_0$$

- $\theta_r := \Lambda_1 \theta \Rightarrow \theta \approx \Pi_1 \theta_r$
- Reducing truncated projection: $\Lambda_1 \in \mathbb{R}^{p \times p}$
- Reconstructing truncated projection: $\Pi_1 \in \mathbb{R}^{p \times p}$
- (Bi-)Orthogonality: $\Lambda_1 \Pi_1 = \mathbb{1} \in \mathbb{R}^{p \times p}$
- $\|y(\theta) - y(\theta_r)\| \ll 1$

Combined State and Parameter Reduced Model:

$$\dot{x}_r(t) = V_1 f(\bar{x} + U_1 x_r(t), u(t), \Pi_1 \theta_r)$$

$$y_r(t) = g(\bar{x} + U_1 x_r(t), u(t), \Pi_1 \theta_r)$$

$$x_r(0) = V_1 x_0$$

$$\theta_r = \Lambda_1 \theta$$

- Affine state-space reduction
- Orthogonal parameter-space reduction
- $\|y(\theta) - y_r(\theta_r)\| \ll 1$



Error System:

$$\dot{x}_e(t) := \begin{pmatrix} \dot{x}(t) \\ \dot{x}_r(t) \end{pmatrix} = \begin{pmatrix} f(x(t), u(t), \theta) \\ f_r(x_r(t), u(t), \theta_r) \end{pmatrix}$$
$$y_e(t) := g(x(t), u(t), \theta) - g_r(x_r(t), u(t), \theta_r)$$

Time-Domain Model Reduction Errors:

- $\|y_e\|_? = \|y - y_r\|_?$
- $\|y_e(\theta)\|_? = \|y(\theta) - y_r(\theta)\|_?$
- $\|y_e(\theta, \theta_r)\|_? = \|y(\theta) - y_r(\theta_r)\|_?$



Lebesgue L_1 -Norm (Action):

$$\|y\|_{L_1} := \int_0^\infty \|y(t)\|_1 dt = \int_0^\infty \sum_{i=1}^Q |y_i(t)| dt$$

Lebesgue L_2 -Norm (Energy):

$$\|y\|_{L_2} := \sqrt{\int_0^\infty \|y(t)\|_2^2 dt} = \sqrt{\int_0^\infty \sum_{i=1}^Q y_i(t)^2 dt}$$

Lebesgue L_∞ -Norm (Peak):

$$\|y\|_{L_\infty} := \sup_{t \in [0, \infty)} \|y(t)\|_\infty = \sup_{t \in [0, \infty)} \max_j y_j(t)$$



Generic (time-domain) joint state and parameter norm:

$$\|y(\theta)\|_{L_p \otimes L_q} := (\|\cdot\|_{L_p} \circ \|\cdot\|_{L_q})(y(\theta))$$

Sample joint norms:

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

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

- Practically joint norms require a discrete parameter space,
- which needs to be sampled sufficiently.
- Originally, joint frequency-parameter norms were introduced³.

³U. Baur, C. Beattie, P. Benner, and S. Gugercin. **Interpolatory Projection Methods for Parameterized Model Reduction**. *SIAM Journal on Scientific Computing*, 33(5): 2489–2518, 2011.



Hardy H_2 -Norm:

$$\|G\|_{H_2} := \sqrt{\frac{1}{2\pi} \int_0^\infty \text{tr}(G^*(-i\omega)G(i\omega)) d\omega}$$

Hardy H_∞ -Norm:

$$\|G\|_{H_\infty} := \sup_{\omega>0} \sigma_{\max}(G(i\omega))$$

System L_1 -Norm:

$$\|g\|_{L_1} := \left\| \int_0^\infty |g(t)| dt \right\|_\infty$$

H_2 / L_∞ Relation (Young's Inequality):

$$\|y\|_{L_\infty} = \|g * u\|_{L_\infty} \leq \|g\|_{L_2} \|u\|_{L_2} = \|G\|_{H_2} \|u\|_{L_2}$$

H_∞ / L_2 Relation (Paley-Wiener Theorem / Parseval's Equation):

$$\|y\|_{L_2} \leq \|G\|_{H_\infty} \|u\|_{L_2}$$

L_1 / L_1 Relation (Young's Inequality):

$$\|y\|_{L_1} \leq \|g\|_{L_1} \|u\|_{L_1}$$

Principal Axis Theorem:

For every real symmetric matrix exists an orthogonal base of eigenvectors.

Singular Value Decomposition:

For a matrix $X \in \mathbb{R}^{N \times M}$ there exist orthogonal matrices $U \in \mathbb{R}^{N \times N}$, $V \in \mathbb{R}^{M \times M}$ and a diagonal matrix $D \in \mathbb{R}^{N \times M}$ such that:

$$X = UDV.$$

With the diagonal entries $D_{ii} = \sqrt{\lambda_i(XX^T)}$ being the singular values of X , this is a singular value decomposition (SVD).

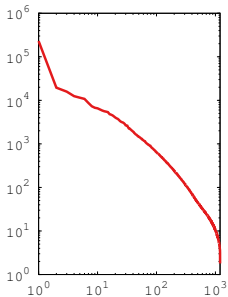


CSC

Dimension Reduction



0 600 1200



0 600 1200



0 600 1200



1. Hankel Operator
2. Controllability and Observability
3. System Gramians
4. Balanced Truncation
5. Singular Perturbation

Evolution Operator:

$$S(x_0, u)(t) := C e^{At} x_0 + \int_0^t C e^{A\tau} B u(\tau) d\tau$$

- Convolution of impulse response with input.
- Maps inputs to outputs ($x_0 \equiv 0$): $S : L_2^M \rightarrow L_2^Q$.
- Spectrum is not finite!



Hankel Operator:

$$\begin{aligned} H(u)(t) &:= S \circ F(u)(t) = \int_{-\infty}^0 C e^{A(t-\tau)} B u(\tau) d\tau \\ &= \left(C e^{At} \right) \left(\int_0^{\infty} e^{A\tau} B u(-\tau) d\tau \right) \\ &= \mathcal{O} \quad \circ \quad \mathcal{C} \end{aligned}$$

- Time-flip operator: $F(u)(t) = u(-t)$.
- Maps past inputs to future outputs.
- Spectrum is finite!

Hankel Norm:

$$\|G\|_H := \sigma_{\max}(H)$$

Lower Model Reduction Bound⁴:

$$\varepsilon_H \geq \sigma_{n+1}(H) \geq 0$$

- Compactness \Rightarrow SVD
- Hilbert-Schmidt operator: $\|H\|_F < \infty$
- Nuclear operator: $\|H\|_* < \infty$

⁴K. Glover and J.R. Partington. **Bounds on the Achievable Accuracy in Model Reduction**. In Modelling, Robustness and Sensitivity Reduction in Control Systems, vol. 34: 95–118. Springer, 1987.



Gramian aka Grammian aka Gram matrix:

$$W_{ij} := \langle v_i, v_j \rangle \rightarrow W = \int_0^\infty V(t)V^T(t) dt \in \mathbb{R}^{N \times N}$$

- $v_{i=1\dots M} \in L_2^N[0, \infty)$
- $V(t) = (v_1(t) \ \dots \ v_M(t))$
- W symmetric, positive semi-definite

1. Controllability Gramian
2. Observability Gramian
3. Cross Gramian (not an actual Gramian matrix)



Controllability:

A system is controllable if for any state $\tilde{x} \in \mathbb{R}^N$ there exists an input function $u : [0, \mathcal{T}] \rightarrow \mathbb{R}^M$, $\mathcal{T} < \infty$, such that $x(0) = \tilde{x}$ and $x(\mathcal{T}) = \tilde{x}$.

Reachability:

A system is reachable if for any state $\tilde{x} \in \mathbb{R}^N$ there exists an input function $u : [0, \mathcal{T}] \rightarrow \mathbb{R}^M$, $\mathcal{T} < \infty$, such that $x(0) = \bar{x}$ and $x(\mathcal{T}) = \tilde{x}$.

Stabilizability:

A system is stabilizable if all uncontrollable subsystems are (asymptotically) stable.

Controllability Operator:

$$\mathcal{C}(u) := \int_0^{\infty} e^{At} B u(-t) dt$$

Adjoint Controllability Operator:

$$\mathcal{C}^*(z_0) = B^T e^{-A^T t} z_0$$

Controllability Gramian Matrix:

$$\begin{aligned} W_C &:= \mathcal{C}\mathcal{C}^* = \int_0^{\infty} e^{At} B B^T e^{A^T t} dt \in \mathbb{R}^{N \times N} \\ &\Leftrightarrow A W_C + W_C A^T = -B B^T \end{aligned}$$

Reconstructibility:

A system is reconstructible if an initial state x_0 is uniquely determined by the output $y(t) \in \mathbb{R}^O$ on a finite time interval $[0, \mathcal{T}]$.

Observability:

A system is observable if any state $x(\mathcal{T})$ is uniquely determined by the output $y(t) \in \mathbb{R}^O$ on a finite time interval $[0, \mathcal{T}]$.

Detectability:

A system is detectable if all unobservable subsystems are (asymptotically) stable.

Observability Operator:

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

Adjoint Observability Operator:

$$\mathcal{O}^*(y) = \int_0^{\infty} e^{A^T t} C^T y(t) dt$$

Observability Gramian Matrix:

$$\begin{aligned} W_O := \mathcal{O}^* \mathcal{O} &= \int_0^{\infty} e^{A^T t} C^T C e^{At} dt \in \mathbb{R}^{N \times N} \\ &\Leftrightarrow W_O A^T + A W_O = -C^T C \end{aligned}$$



Cross Gramian Matrix⁵:

$$W_X := \mathcal{CO} = \int_0^{\infty} e^{At} BC e^{At} dt$$
$$\Leftrightarrow AW_X + W_X A = -BC$$

- Combines controllability and observability.
- Requires square system: $M = Q$.
- Cross Gramian trace: $\text{tr}(W_X) = \text{tr}(H)$.

⁵K.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.

System Gramian relation for **symmetric** systems:

$$W_C W_O = W_X^2$$

System Gramian relation for **state-space symmetric** systems:

$$W_C = W_O = W_X$$

Controllability-based cross Gramian⁶:

$$\begin{pmatrix} \dot{x}(t) \\ \dot{z}(t) \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A^T \end{pmatrix} \begin{pmatrix} x(t) \\ z(t) \end{pmatrix} + \begin{pmatrix} B \\ C^T \end{pmatrix} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} \rightarrow W_C = \begin{pmatrix} W_C & W_X \\ W_X^T & W_O \end{pmatrix}$$

⁶K.V. Fernando and H. Nicholson. **On the Cross-Gramian for Symmetric MIMO Systems**. IEEE Transactions on Circuits and Systems, 32(5): 487–489, 1985.



Orthogonally symmetric systems⁷:

$$\begin{aligned} \exists P = P^T, \exists U, UU^T = \mathbb{1} : AP = PA^T, B = PCU^T, C = PBU \\ \Rightarrow W_X = \int_0^\infty e^{At} BUC e^{At} dt \Rightarrow W_X^2 = W_C W_O \end{aligned}$$

Symmetric embedding⁸:

$$\exists J = J^T \in \mathbb{R}^{N \times N} : AJ = JA^T \rightarrow \bar{A} := A, \bar{B} := (JC^T \ B), \bar{C} := \begin{pmatrix} C \\ B^T J^{-1} \end{pmatrix}$$

⁷J.A. De Abreu-Garcia and F.W. Fairman. **A Note on Cross Grammians for Orthogonally Symmetric Realizations**. IEEE Transactions on Automatic Control, 31(9): 866–868, 1986.

⁸D.C. Sorensen and A.C. Antoulas. **The Sylvester equation and approximate balanced reduction**. Linear Algebra and its Applications, 351–352: 671–700, 2002.



Cross Gramian of a square MIMO as sum of SISOs:

$$W_X = \sum_{i=1}^M \int_0^{\infty} e^{At} B_{:,i} C_{i,:} e^{At} dt$$

Non-Symmetric Cross Gramian⁹ (Cross Gramian of average system):

$$\begin{aligned} W_Z &:= \sum_{i=1}^M \sum_{j=1}^Q \int_0^{\infty} e^{At} B_{:,i} C_{j,:} e^{At} dt \\ &= \int_0^{\infty} e^{At} \left(\sum_{i=1}^M B_{:,i} \right) \left(\sum_{j=1}^Q C_{j,:} \right) e^{At} dt \end{aligned}$$

■ Motivated by decentralized control.

⁹C. Himpe and M. Ohlberger. **A note on the cross Gramian for non-symmetric systems.** System Science and Control Engineering, 4(1): 199–208, 2016.

Hankel Operator:

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

Hankel Singular Values (HSV):

$$\sigma_i(H) = \sqrt{\lambda_i(W_C W_O)}$$

Hankel Singular Values (Symmetric Systems):

$$\sigma_i(H) = |\lambda_i(W_X)|$$



Balancing Transformation (Simultaneous Diagonalization):

$$TW_C T^\top = T^{-\top} W_O T^{-1} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_N \end{pmatrix}$$

- Balancing of controllability and observability¹⁰.
- HSV quantify importance of balanced states.
- Truncating T , T^{-1} yields U_1 , V_1 .

¹⁰B. Moore. **Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction**. IEEE Transactions on Automatic Control, 26(1): 17–32, 1981.

Square-Root Algorithm¹¹:

1. $W_C \stackrel{\text{Cholesky}}{=} L_C L_C^T$
2. $W_O \stackrel{\text{Cholesky}}{=} L_O L_O^T$
3. $L_O L_C^T \stackrel{\text{SVD}}{=} U D V$

SVD-Based Algorithm¹²:

1. $W_C \stackrel{\text{SVD}}{=} U_C D_C U_C^T$
2. $W_O \stackrel{\text{SVD}}{=} U_O D_O U_O^T$
3. $U_C D_C^{\frac{1}{2}} U_C^T U_O D_O^{\frac{1}{2}} U_O^T \stackrel{\text{SVD}}{=} U D V$

¹¹A.J. Laub, M.T. Heath, C. Paige, and R. Ward. **Computation of System Balancing Transformations and Other Applications of Simultaneous Diagonalization Algorithms**. IEEE Transactions on Automatic Control, 32(2): 115–122, 1987.

¹²J.S. Garcia and J.C. Basilio. **Computation of reduced-order models of multivariable systems by balanced truncation**. International Journal of Systems Science, 33(10): 847–854, 2002.

EVD-Based Balancing¹³:

1. $W_X \stackrel{\text{EVD}}{=} UDV$

SVD-Based Approximate Balancing⁷:

1. $W_X \stackrel{\text{SVD}}{=} UDV$

■ Direct Truncation: $V \leftarrow U^T$

¹³R.W. Aldhaheri. **Model order reduction via real Schur-form decomposition.** International Journal of Control, 53(3): 709–716, 1991.

Truncation Operator:

$$S = \begin{pmatrix} \mathbb{1}_n \\ 0 \end{pmatrix} \in \mathbb{R}^{N \times n}$$

- Let U and V be a balancing transformation.
- Then: $U_1 := U \circ S$, $V_1 := S^T \circ V$.
- This masks out the least important (balanced) states.
- Zero vector field components are constant,
- hence, can be discarded.

Computing Lyapunov and Sylvester equations:

- Bartels-Stewart¹⁴
- Sign function¹⁵
- Alternating Direction Implicit Iteration (ADI)¹⁶
- ...

¹⁴R. Bartels and G. Stewart. **Solution of the Matrix Equation $AX + XB = C$** . Communication of the ACM, 15(9): 820–826, 1972.

¹⁵U. Baur and P. Benner. **Cross-Gramian Based Model Reduction for Data-Sparse Systems**. Electronic Transactions on Numerical Analysis, 31: 256–270, 2008.

¹⁶J. Saak, P. Benner and P. Kürschner. **A Goal-Oriented Dual LRCF-ADI for Balanced Truncation**. IFAC Proceedings Volumes, 45(2): 752–757, 2012.

- Balanced truncation is stability preserving¹⁷.
- Approximate balancing generally not, but often is.
- Non-symmetric cross Gramian preserves stability⁸.

¹⁷L. Pernebo and L. Silverman. **Model Reduction via Balanced State Space Representations**. IEEE Transactions on Automatic Control, 27(2): 382–387, 1982.



H_∞ Error Bound^{18,19}:

$$\|G - G_r\|_{H_\infty} \leq 2 \sum_{i=n+1}^N \sigma_i(H)$$

L_1 Error Bound²⁰:

$$\|g - g_r\|_{L_1} \leq 4(N + n) \sum_{i=n+1}^N \sigma_i(H)$$

¹⁸D.F. Enns. **Model Reduction with Balanced Realizations: An Error Bound and a Frequency Weighted Generalization.** In IEEE Conference on Decision and Control, vol. 23: 127–132, 1984.

¹⁹K. Glover. **All optimal Hankel-norm approximations of linear multivariable systems and their L_∞ -error bounds.** International Journal of Control, 39(6): 1115–1193, 1984.

²⁰J. Lam and B.D.O Anderson. **L_1 Impulse Response Error Bound for Balanced Truncation.** System & Control Letters, 18(2): 129–137, 1992.



H_2 Error Indicator⁷:

$$\|G - G_r\|_{H_2} \approx \sqrt{\text{tr}(\tilde{C}_2 W_{22} \tilde{B}_2)}$$

- H_2 norm: $\|G\|_{H_2} = \sqrt{\text{tr}(C W_C C^\top)} = \sqrt{\text{tr}(B^\top W_O B)}$
- Balanced system matrices: \tilde{A} , \tilde{B} , \tilde{C}
- Balanced Gramian: $W = \begin{pmatrix} \sigma_1(H) & & \\ & \ddots & \\ & & \sigma_N(H) \end{pmatrix}$

Second-Order System:

$$M\ddot{q}(t) + G\dot{q}(t) + Kq(t) = B_V u(t)$$
$$y(t) = C_P q(t) + C_V \dot{q}(t)$$

First-Order Representation:

$$\begin{pmatrix} \dot{x}_P(t) \\ \dot{x}_V(t) \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1} \\ -M^{-1}K & -M^{-1}G \end{pmatrix} \begin{pmatrix} x_P(t) \\ x_V(t) \end{pmatrix} + \begin{pmatrix} 0 \\ M^{-1}B_V \end{pmatrix} u(t)$$
$$y(t) = (C_P \quad C_V) \begin{pmatrix} x_P(t) \\ x_V(t) \end{pmatrix}$$



Second-Order System Gramians²¹:

$$W_C = \begin{pmatrix} W_{C,P} & W_{C,12} \\ W_{C,21} & W_{C,V} \end{pmatrix}, \quad W_O = \begin{pmatrix} W_{O,P} & W_{O,12} \\ W_{O,21} & W_{O,V} \end{pmatrix}, \quad W_X = \begin{pmatrix} W_{X,P} & W_{X,12} \\ W_{X,21} & W_{X,V} \end{pmatrix}$$

- Position and velocity projections: $U_{P,1}, V_{P,1}, U_{V,1}, V_{V,1}$
- Generally not stability preserving.
- Preserves second-order structure.

²¹T. Reis and T. Stykel. **Balanced truncation model reduction of second-order systems**. *Mathematical and Computer Modelling of Dynamical Systems: Methods, Tools and Applications in Engineering and Related Sciences*, 14(5): 391–406, 2008.



Structure Preserving Reduced Order Model²²:

$$\begin{pmatrix} \dot{x}_\alpha(t) \\ \dot{x}_\beta(t) \end{pmatrix} = \begin{pmatrix} V_{\alpha,1} & 0 \\ 0 & V_{\beta,1} \end{pmatrix} \begin{pmatrix} A_\alpha & A_{\alpha\beta} \\ A_{\beta\alpha} & A_\beta \end{pmatrix} \begin{pmatrix} U_\alpha & 0 \\ 0 & U_\beta \end{pmatrix} \begin{pmatrix} x_\alpha(t) \\ x_\beta(t) \end{pmatrix} + \begin{pmatrix} V_\alpha & 0 \\ 0 & V_\beta \end{pmatrix} \begin{pmatrix} B_\alpha \\ B_\beta \end{pmatrix} u(t)$$
$$y(t) = \begin{pmatrix} C_\alpha & C_\beta \end{pmatrix} \begin{pmatrix} U_\alpha & 0 \\ 0 & U_\beta \end{pmatrix} \begin{pmatrix} x_\alpha(t) \\ x_\beta(t) \end{pmatrix}$$

- Generalization of second-order balanced truncation.
- Per component projections.
- An H_∞ error bound exists.

²²H. Sandberg and R.M. Murray. **Model reduction of interconnected linear systems**. Optimal Control Applications and Methods, 30(3): 225–245, 2009.



General Structured Model:

$$\begin{pmatrix} \dot{x}_\alpha(t) \\ \dot{x}_\beta(t) \end{pmatrix} = \begin{pmatrix} f_\alpha(x_\alpha(t), x_\beta(t), u(t), \theta) \\ f_\beta(x_\alpha(t), x_\beta(t), u(t), \theta) \end{pmatrix}$$
$$y(t) = g(x_\alpha(t), x_\beta(t), u(t), \theta)$$

Structured Reduced Order Model:

$$\begin{pmatrix} \dot{x}_{\alpha,r}(t) \\ \dot{x}_{\beta,r}(t) \end{pmatrix} = \begin{pmatrix} V_{\alpha,1} f_\alpha(U_{\alpha,1}x_{\alpha,r}(t), U_{\beta,1}x_{\beta,r}(t), u(t), \theta) \\ V_{\beta,1} f_\beta(U_{\alpha,1}x_{\alpha,r}(t), U_{\beta,1}x_{\beta,r}(t), u(t), \theta) \end{pmatrix}$$
$$y_r(t) = g(U_{\alpha,1}x_{\alpha,r}(t), U_{\beta,1}x_{\beta,r}(t), u(t), \theta)$$

Singular Perturbation Reduced Order System Matrices²³:

$$A_r := A_{11} - A_{12}A_{22}^{-1}A_{21}$$

$$B_r := B_1 - A_{12}A_{22}^{-1}B_2$$

$$C_r := C_1 - C_2A_{22}^{-1}A_{21}$$

- Balanced Truncation better for $s \rightarrow \infty$
- Singular Perturbation better for $s \rightarrow 0$
- Instead of truncating (BT), keep at steady state (SP).
- Practically, a Schur complement.
- H_∞ Error bound and stability preservation hold.

²³G. Obinata and B.D.O. Anderson. **Model Reduction for Control System Design.** Communications and Control Engineering. Springer, 2001.

Generalized Singular Perturbation²⁴:

$$A_r := A_{11} + A_{12}(s_0\mathbb{1} - A_{22})^{-1}A_{21}$$

$$B_r := B_1 + A_{12}(s_0\mathbb{1} - A_{22})^{-1}B_2$$

$$C_r := C_1 + C_2(s_0\mathbb{1} - A_{22})^{-1}A_{21}$$

- Balanced Truncation: $s \rightarrow \infty$
- Singular Perturbation: $s \rightarrow 0$
- Generalized Singular Perturbation: $s = s_0$

²⁴Y. Liu and B.D.O. Anderson. **Singular Perturbation Approximation of Balanced Systems**. International Journal of Control, 50(4): 1379–1405, 1989.



Norm identity:

$$\|G\|_{H_2} = \text{tr}\left(\int_0^\infty y^\top(t)y(t) dt\right) = \text{tr}(C^\top CW_C) = \text{tr}(BB^\top W_O)$$

- Balanced system: $\tilde{A}, \tilde{B}, \tilde{C}$
- $\tilde{B} = (\tilde{b}_1 \ \dots \ \tilde{b}_M), \tilde{C} = (\tilde{c}_1^\top \ \dots \ \tilde{c}_Q^\top)^\top$
- In balanced form: $\|G\|_{H_2} = \sum_{i=1}^N \tilde{c}_i^2 \sigma_i(H) = \sum_{i=1}^N \tilde{b}_i^2 \sigma_i(H)$
- Balanced gains²⁵: $d_i := \tilde{c}_i^2 \sigma_i(H) = \tilde{b}_i^2 \sigma_i(H)$
- BG sometimes better, sometimes worse than BT.

²⁵A. Davidson. **Balanced Systems and Model Reduction**. Electronics Letters, 22(10): 531–532, 1986.



1. Empirical Gramians
2. Empirical Balanced Truncation
3. The Averaging Principle
4. Gramian-Based Parameter Identification
5. Combined State and Parameter Reduction



Linearization:

$$A_\ell := \frac{\partial}{\partial x} f(\bar{x}, \bar{u}), \quad B_\ell := \frac{\partial}{\partial u} f(\bar{x}, \bar{u}), \quad C_\ell := \frac{\partial}{\partial x} g(\bar{x}, \bar{u})$$

- Linearize at steady state.
- Compute balanced truncation projections.
- Use these on nonlinear model²⁶.

²⁶X. 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, vol 2: 1165–1166, 1988.



Controllability energy and observability energy:

$$L_C := \min_{u \in L_2} \frac{1}{2} \int_{-\infty}^0 \|u(t)\|^2 dt = \frac{1}{2} x_0^T W_C^{-1} x_0$$

$$L_O := \frac{1}{2} \int_0^{\infty} \|y(t)\|^2 dt = \frac{1}{2} x_0^T W_O x_0$$

- Control-affine system: $\dot{x}(t) = f(x(t)) + h(x(t))u(t)$, $y(t) = g(x(t))$
- Nonlinear controllability Gramian: $\frac{\partial L_C}{\partial x} f(x) + \frac{1}{2} \frac{\partial L_C}{\partial x} h(x) h^T(x) \frac{\partial L_C}{\partial x} = 0$
- Nonlinear observability Gramian: $\frac{\partial L_O}{\partial x} f(x) + \frac{1}{2} g^T(x) g(x) = 0$
- Nonlinear balancing truncation²⁷
- Numerically infeasible for large systems.

²⁷J.M.A. Scherpen. **Balancing for nonlinear systems**. Systems & Control Letters, 21(2): 143–153, 1993.



System Gramians:

$$W_C = \int_0^{\infty} (e^{At} B)(e^{At} B)^T dt$$

$$W_O = \int_0^{\infty} (e^{A^T t} C)(e^{A^T t} C^T)^T dt$$

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

- Integrand is impulse response.
- Data-driven computation.
- Applicable for nonlinear systems.



(Auto-)Correlation:

$$W := \int_0^{\infty} (x(t) - \bar{x})(x(t) - \bar{x})^T dt$$

Temporal Mean:

$$\bar{x} = \lim_{T \rightarrow \infty} \int_0^T x(t) dt$$

Averaging:

$$W = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} W_i$$

Input Perturbations:

- Direction: $E_u = \{e_i \in \mathbb{R}^M; \|e_i\| = 1; \langle e_i, e_{j \neq i} \rangle = 0; i, j = 1, \dots, M\}$
- Rotation: $R_u = \{S_i \in \mathbb{R}^{M \times M}; S_i^T S_i = \mathbb{1}_M; i = 1, \dots, s\}$
- Scale: $Q_u = \{c_i \in \mathbb{R}; c_i > 0; i = 1, \dots, q\}$

Initial State Perturbations:

- Direction: $E_x = \{f_i \in \mathbb{R}^N; \|f_i\| = 1; \langle f_i, f_{j \neq i} \rangle = 0; i, j = 1, \dots, N\}$
- Rotation: $R_x = \{T_i \in \mathbb{R}^{N \times N}; T_i^T T_i = \mathbb{1}_N; i = 1, \dots, t\}$
- Scale: $Q_x = \{d_i \in \mathbb{R}; d_i > 0; i = 1, \dots, r\}$



Empirical Controllability Gramian²⁸:

$$\widehat{W}_C = \frac{1}{|Q_u||R_u|} \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^M \frac{1}{c_h^2} \int_0^\infty \Psi^{hij}(t) dt \in \mathbb{R}^{N \times N}$$

$$\Psi^{hij}(t) = (x^{hij}(t) - \bar{x}^{hij})(x^{hij}(t) - \bar{x}^{hij})^T \in \mathbb{R}^{N \times N}$$

- Non-trivial sets: E_u, R_u, Q_u
- $x^{hij}(t)$ is trajectory for input: $u^{hij}(t) = c_h S_i e_j \delta(t)$
- Temporal mean state: \bar{x}^{hij}

²⁸S. Lall, J.E. Marsden, and S. Glavaski. **Empirical Model Reduction of Controlled Nonlinear Systems**. In Proceedings of the 14th IFAC Congress, vol. F: 473–478, 1999.



Empirical Controllability Covariance Matrix²⁹:

$$\widetilde{W}_C = \frac{1}{|Q_u||R_u|} \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^M \frac{1}{c_h^2} \int_0^\infty \Psi^{hij}(t) dt \in \mathbb{R}^{N \times N}$$

$$\Psi^{hij}(t) = (x^{hij}(t) - \bar{x}^{hij})(x^{hij}(t) - \bar{x}^{hij})^T \in \mathbb{R}^{N \times N}$$

- Non-trivial sets: E_u, R_u, Q_u
- $x^{hij}(t)$ is trajectory for input: $u^{hij}(t) = c_h S_i e_j \odot u(t) + \bar{u}$
- Steady-state: \bar{x}
- Steady-state input: \bar{u}

²⁹J. Hahn and T.F. Edgar. **Balancing Approach to Minimal Realization and Model Reduction of Stable Nonlinear Systems**. Industrial & Engineering Chemistry Research, 41(9): 2204–2212, 2002.



Empirical Observability Gramian²⁵:

$$\widehat{W}_O = \frac{1}{|Q_x| |R_x|} \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{1}{d_k^2} \int_0^\infty T_l \Psi^{kl}(t) T_l^T dt \in \mathbb{R}^{N \times N}$$

$$\Psi_{ab}^{kl}(t) = (y^{kla}(t) - \bar{y}^{kla})^T (y^{klb}(t) - \bar{y}^{klb}) \in \mathbb{R}$$

- Non-trivial sets: E_x , R_x , Q_x
- $y^{kla}(t)$ is output trajectory for initial state: $x_0^{kla} = d_k T_l f_a$
- Temporal mean output: \bar{y}^{kla}



Empirical Observability Covariance Matrix²⁶:

$$\widetilde{W}_O = \frac{1}{|Q_x||R_x|} \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{1}{d_k^2} \int_0^\infty T_l \Psi^{kl}(t) T_l^T dt \in \mathbb{R}^{N \times N}$$

$$\Psi_{ab}^{kl}(t) = (y^{kla}(t) - \bar{y}^{kla})^T (y^{klb}(t) - \bar{y}^{klb}) \in \mathbb{R}$$

- Non-trivial sets: E_x , R_x , Q_x
- $y^{kla}(t)$ is output trajectory for initial state: $x_0^{kla} = d_k T_l f_a + \bar{x}$
- Steady-state: \bar{x}
- Steady state output: \bar{y}



Empirical Linear Cross Gramian³⁰:

$$\widehat{W}_Y = \frac{1}{|Q_u||R_u|} \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^M \frac{1}{c_h^2} \int_0^\infty \Psi^{hij}(t) dt \in \mathbb{R}^{N \times N}$$

$$\Psi^{hij}(t) = (x^{hij}(t) - \bar{x}^{hij})(z^{hij}(t) - \bar{z}^{hij})^T \in \mathbb{R}^{N \times N}$$

- Non-trivial sets: E_u, R_u, Q_u
- x^{hij} is trajectory for input: $u^{hij}(t) = c_h S_i e_j \delta(t)$
- z^{hij} is adjoint trajectory for input $v^{hij}(t) = c_h S_i e_j \delta(t)$
- Temporal mean: \bar{x}^{hij}
- Adjoint temporal mean: \bar{z}^{hij}

³⁰U. Baur, P. Benner, B. Haasdonk, C. Himpe, I. Martini and M. Ohlberger.

Comparison of Methods for Parametric Model Order Reduction of Time-Dependent Problems. In: Model Reduction and Approximation: Theory and Algorithms, Editors: P. Benner, A. Cohen, M. Ohlberger and K. Willcox, SIAM: 377–407, 2017.



Empirical Cross Gramian³¹:

$$\widehat{W}_X = \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^M \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{1}{c_h d_k} \int_0^\infty T_l \Psi^{hijkl}(t) T_l^T dt \in \mathbb{R}^{N \times N}$$

$$\Psi_{ab}^{hijkl}(t) = f_b^T T_l^T (x^{hij}(t) - \bar{x}^{hij}) e_j^T S_i^T (y^{kla}(t) - \bar{y}^{kla}) \in \mathbb{R}$$

- Non-trivial sets: $E_u, E_x, R_u, R_x, Q_u, Q_x$
- x^{hij} is trajectory for input: $u^{hij}(t) = c_h S_i e_j \delta(t)$
- y^{kla} is output trajectory for initial state: $x_0^{kla} = d_k T_l f_a$
- Temporal mean state: \bar{x}^{hij}
- Temporal mean output: \bar{y}^{kla}
- Efficient computation of empirical non-symmetric cross Gramian⁸.

³¹C. 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.



Empirical Cross Covariance Matrix³¹:

$$\widetilde{W}_X = \sum_{h=1}^{|Q_u|} \sum_{i=1}^{|R_u|} \sum_{j=1}^M \sum_{k=1}^{|Q_x|} \sum_{l=1}^{|R_x|} \frac{1}{c_h d_k} \int_0^\infty T_l \Psi^{hijkl}(t) T_l^T dt \in \mathbb{R}^{N \times N}$$

$$\Psi_{ab}^{ijkl}(t) = f_b^T T_l^T (x^{hij}(t) - \bar{x}^{hij}) e_j^T S_i^T (y^{kla}(t) - \bar{y}^{kla}) \in \mathbb{R}$$

- Non-trivial sets: $E_u, E_x, R_u, R_x, Q_u, Q_x$
- x^{hij} is trajectory for input $u^{hij}(t) = c_h S_i e_j \odot u(t) + \bar{u}$
- y^{kla} is output trajectory for initial state $x_0^{kla} = d_k T_l f_a + \bar{x}$
- Steady-state: \bar{x}
- Steady-state input: \bar{u}
- Steady-state output: \bar{y}



Columnwise computability of empirical cross Gramian³²:

$$W_X = (w_{X,1} \quad \dots \quad w_{X,N})$$
$$w_{X,j} = \sum_{m=1}^M \int_0^\infty \psi^{jm}(t) dt \in \mathbb{R}^N$$
$$\psi_i^{jm}(t) = (x_i^m(t) - \bar{x}_i^m)(y_m^j(t) - \bar{y}_m^j)$$

- Empirical cross Gramian is a dense $N \times N$ matrix.
- Compute singular vectors (projection) directly,
- without assembly of the empirical cross Gramian matrix.

³²C. Himpe, T. Leibner and S. Rave. **Hierarchical Approximate Proper Orthogonal Decomposition**. arXiv e-prints, math.NA: 1607.05210, 2017.



Relative Information Content:

$$\varepsilon_1 := \frac{\|H_r\|_*}{\|H\|_*} = \frac{\sum_{i=1}^n \sigma_i(H)}{\sum_{i=1}^N \sigma_i(H)}$$

Linear Error Indicators:

- $\|y - y_r\|_{H_\infty} \lesssim 2 \sum_{i=n+1}^N \sigma_i(H)$
- $\|y - y_r\|_{H_2} \approx \sqrt{\text{tr}(\tilde{C}_{\ell,2} W_{22} \tilde{B}_{\ell,2})}$
- $\|y - y_r\|_{L_1} \lesssim 4(N + n) \sum_{i=n+1}^N \sigma_i(H)$

Time-Varying System:

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

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

- Compute average empirical Gramians³³.
- Average over time.
- These are time-invariant Gramians.

³³O. Nilsson and A. Rantzer. **A novel nonlinear model reduction method applied to automotive controller software**. Proceedings of the American Control Conference: 4587–4592, 2009.

Inner Product³⁴:

- Galerkin projections are not stability preserving.
- Adapting the inner product can fix this.
- For example: energy-stable inner products.

Kernel³⁵:

- Think of kernels as generalized inner products.
- Kernel trick: Only evaluate inner product in nonlinear space.
- Mapping back to the original space is more expensive.

³⁴I. Kalashnikova, M.F. Barone, S. Arunajatesan, B.G. Van Bloemen Waanders. **Construction of energy-stable projection-based reduced order models**. Applied Mathematics and Computation, 249: 569–596, 2014.

³⁵J. Bouvrie and B. Hamzi. **Kernel Methods for the Approximation of Nonlinear Systems**. SIAM Journal on Control and Optimization, 55(4): 2460–2492, 2017.

Linearly Parametric Linear System:

$$\dot{x}(t) = Ax(t) + Bu(t) + F\theta = Ax(t) + (B \quad F) \begin{pmatrix} u(t) \\ u_\theta(t) \end{pmatrix}$$

Parametric Empirical Gramians³⁰:

$$\widehat{W}_{C,\theta} := \sum_{i=1}^{\mathcal{P}} \widehat{W}_C(\theta_i)$$

- Parameter input: $u_\theta(t) = \theta$
- Observability is adjoint controllability: $\widehat{W}_{O,\theta} := \sum_i \widehat{W}_O(\theta_i)$
- Parametric empirical cross Gramian: $\widehat{W}_{X,\theta} := \sum_i \widehat{W}_X(\theta_i)$

Discrete Parameter Space:

$$\Theta_h = (\theta^1 \quad \dots \quad \theta^P)$$

- Discretize according to operating region.
- Sampling: i.e. sparse grids, greedy, etc.
- Parametric operating region: $E_u \times R_u \times Q_u \times \Theta_h$

Parameter Covariance:

$$\omega \stackrel{\text{SVD}}{=} \Pi \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_P \end{pmatrix} \Pi^\top$$

- Use same machinery as for states.
- Parameter Galerkin projections via SVD.
- How to compute parameter covariances?



Linearly Parametric Linear System:

$$\dot{x}(t) = Ax(t) + Bu(t) + F\theta = Ax(t) + Bu(t) + \sum_{i=1}^P F_i \theta_i$$

$$\rightarrow W_C = W_{C,0} + \sum_{i=1}^P W_{C,i}$$

$$W_{C,0} = \int_0^{\infty} e^{At} BB^T e^{A^T t} dt, \quad W_{C,i} = \int_0^{\infty} e^{At} F_i F_i^T e^{A^T t} dt$$

Empirical Sensitivity Gramian:

$$W_S := \begin{pmatrix} \text{tr}(W_{C,1}) & & \\ & \ddots & \\ & & \text{tr}(W_{C,P}) \end{pmatrix}$$

Augmented System:

$$\begin{pmatrix} \dot{x}(t) \\ \dot{\theta}(t) \end{pmatrix} = \begin{pmatrix} f(x(t), u(t), \theta(t)) \\ \theta_0 \end{pmatrix}$$
$$y(t) = g(x(t), u(t), \theta(t))$$

Augmented Observability Gramian:

$$\overline{W}_O = \begin{pmatrix} W_O & W_M \\ W_M^T & W_P \end{pmatrix}$$

Empirical Identifiability Gramian:

$$W_I := W_P - W_M^T W_O^{-1} W_M \approx W_P$$

- Related to Fischer information matrix.



Empirical Joint Gramian³¹:

$$W_J := \overline{W}_X = \begin{pmatrix} W_X & W_M \\ W_m & W_P \end{pmatrix} = \begin{pmatrix} W_X & W_M \\ 0 & 0 \end{pmatrix}$$

Empirical Cross-Identifiability Gramian³¹:

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

- Lower blocks are zero as parameters are uncontrollable.
- W_i encodes observability of parameters.
- Schur complement can be approximated efficiently.

Combined Reductions:

- State-space reduction via empirical Gramians
- Parameter-space reduction via empirical Gramians
- Compound computation

Abbreviations:

- BT = Balanced Truncation
- DT = Direct Truncation

1. Compute W_S (includes W_C)
2. Compute W_O
3. State-Space Reduction: $BT(W_C, W_O)$
4. Parameter-Space Reduction: $DT(W_S)$

1. Compute W_C
2. Compute W_I (includes W_O)
3. State-Space Reduction: $BT(W_C, W_O)$
4. Parameter-Space Reduction: $DT(W_I)$

1. Compute W_J (includes $W_X, W_{\ddot{j}}$)
2. State-Space Reduction: $DT(W_X)$
3. Parameter-Space Reduction: $DT(W_{\ddot{j}})$

- Works for high-dimensional parameter-spaces!
- Combined computation of state and parameter Gramians.
- Preferably for control-affine systems.
- Definition of operating region (around a steady state) paramount.
- Simulation trajectory quality determines reduced model quality.



1. Linear Benchmark: Inverse Lyapunov Procedure
2. Nonlinear Benchmark: RC Ladder
3. Hyperbolic Network Model
4. EEG Dynamic Causal Model
5. fMRI Dynamic Causal Model



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:

- Interfaces for: Solver, inner product kernels & distributed memory
- Non-Symmetric option for all cross Gramians
- Functional design
- Compatible with OCTAVE and MATLAB
- Vectorized and parallelizable
- Open-source licensed

More info: <http://gramian.de>

Linear-Linear System:

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

Inverse Lyapunov Procedure³⁶

1. Sample eigenvalues of W_C and W_O .
2. Sample eigenvectors of W_C and W_O .
3. Normalize B and C .
4. Compute A via Lyapunov equation.
5. Unbalance system.

³⁶S.C. Smith and J. Fisher. **On generating random systems: a gramian approach.** In Proceedings of the American Control Conference, vol. 3: 2743–2748, 2003.



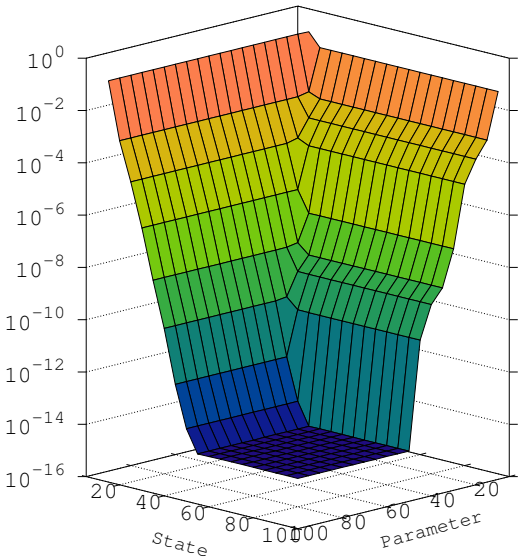
Controllability-based, observability-based, and cross-Gramian-based combined reduction.

SISO Nonlinear RC Cascade³⁷:

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

- Linear Parametrization.
- Dimensionality: $P = N$
- Good-natured nonlinearity.

³⁷Y. Chen. **Model Order Reduction for Nonlinear Systems**. Master's thesis, Massachusetts Institute of Technology, 1999.



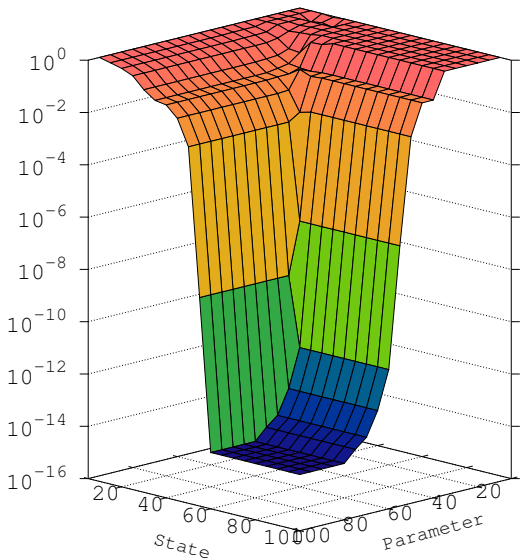


Hyperbolic Network Model³⁸:

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

- Parametrization: $K_{ij}(\theta) = \begin{cases} \theta_i & i = j \\ 0 & i \neq j \end{cases}$
- Dimensionality: $P = N$
- Sigmoid nonlinearity.

³⁸Y. Quan, H. Zhang, and L. Cai. **Modeling and Control Based on a New Neural Network Model**. In Proceedings of the American Control Conference, vol. 3: 1928–1929, 2001.



Practically:

- Inference of brain region connectivity
- from functional imaging data (EEG, fMRI).
- Combined reduction for inverse problem.

Mathematically:

- Two component model
- Network dynamics sub-model
- Measurement sub-model

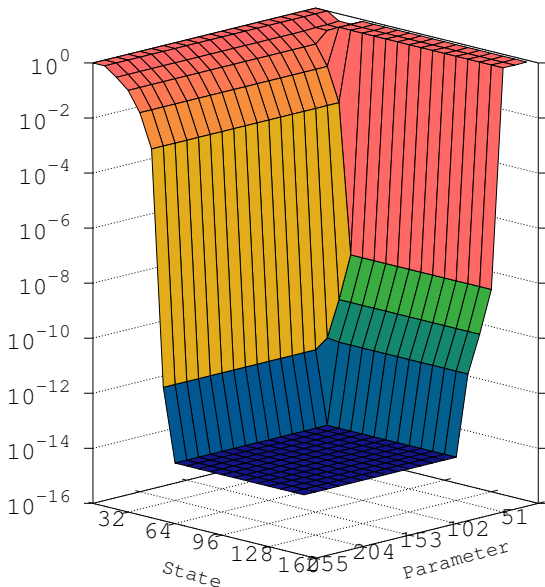


EEG & MEG Dynamic Causal Model³⁹:

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

- Parametrization: $A_{ij}(\theta) = \theta_{i*K+j}$
- Dimensionality: $P = \frac{N^2}{100}$
- Second-order system with sigmoid nonlinearity.

³⁹O. David, S.J. Kiebel, L.M. Harrison, J. Mattout, J.M. Kilner, and K.J. Friston. **Dynamic causal modeling of evoked responses in EEG and MEG**. *NeuroImage*, 4: 1255–1272, 2006.





fMRI Dynamic Causal Model⁴⁰:

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

$$\dot{s}(t) = x(t) - \kappa s(t) - \gamma(f(t) - 1)$$

$$\dot{f}(t) = s(t)$$

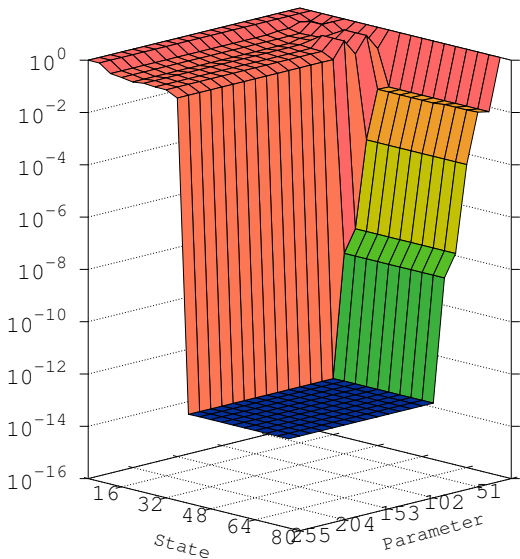
$$\dot{v}(t) = \frac{1}{\tau}(f(t) - v(t)^{\frac{1}{\alpha}})$$

$$\dot{q}(t) = \frac{1}{\tau}(\frac{1}{\rho}f(t)(1 - ((1 - \rho))^{\frac{1}{f(t)}}) - v(t)^{\frac{1}{\alpha}-1}q(t))$$

$$y(t) = V_0(k_1(1 - q(t)) + k_2(1 - v(t)))$$

- Parametrization: $A_{ij}(\theta) = \theta_{i*K+j}$
- Dimensionality: $P = \frac{N^2}{25}$
- Highly nonlinear system.

⁴⁰K.J. Friston, L.M. Harrison, and W. Penny. **Dynamic causal modelling.** *NeuroImage*, 19(4): 1273–1302, 2003.





1. Parameter Hessians
2. Decentralized Control
3. Nonlinearity Quantification
4. Sensitivity Analysis
5. System Indices

Optimization problem subject to control system:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \|y_d - y(\theta)\|_2^2$$

s.t.:

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

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

Newton-Type Algorithm:

- Derivative-based optimization.
- Requires Hessian or an approximation.
- Identifiability Gramian can be used as Hessian⁴¹ (approximation).

⁴¹C. Lieberman and B. Van Bloemen Waanders. **Hessian-Based Model Reduction Approach to Solving Large-Scale Source Inversion Problems**. In CSRI Summer Proceedings: 37–48, 2007.

MIMO System:

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

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

- Which SISO subsystems are dominant?
- Participation matrix: $P_{ij} = \frac{\text{tr}(W_{ij})^2}{\text{tr}(W_{ij}^2)}$, $i = 1 \dots M$, $j = 1 \dots Q$
- Decentralization: Max elements per row or column of P .

Nonlinear System:

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

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

Gramian-based nonlinearity indices⁴²:

- Input nonlinearity: $\sum_{i=1}^N \sum_{j=1}^N |W_C - \widehat{W}_C|_{ij} / \text{tr}(W_C)$
- State nonlinearity: $\sum_{i=1}^N \sum_{j=1}^N |W_X - \widehat{W}_X|_{ij} / \text{tr}(W_X)$
- Output nonlinearity: $\sum_{i=1}^N \sum_{j=1}^N |W_O - \widehat{W}_O|_{ij} / \text{tr}(W_O)$

⁴²J. Hahn and T.F. Edgar. **A Gramian Based Approach to Nonlinearity Quantification and Model Classification**. Industrial & Engineering Chemistry Research, 40(24): 5724–5731, 2001.



System gain aka DC gain aka L_2 gain:

$$S := \text{tr}(G(0)) = \text{tr}(CA^{-1}B)$$

- System gain via cross Gramian: $S(G) = -\frac{1}{2} \text{tr}(W_X)$
- Frequency response for $s = 0$.
- Model reduction based on first order moments (HEV).

Input-Parametric System:

$$\dot{x}(t) = f(x(t), u_{\theta}(t))$$

$$y(t) = g(x(t), u_{\theta}(t))$$

- Treat parameters as inputs: $u_{\theta}(t) = \theta$.
- Parameter sensitivity: $S(\theta) := \frac{y - \bar{y}}{\theta - \bar{\theta}} = S(G(0))$.
- Gain computation via empirical cross Gramian⁴³.

⁴³S. 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.

Cauchy-Index via cross Gramian⁴⁴:

$$C = \sum_{i=1}^N \text{sign}(\lambda_i(W_X))$$

- $C := p_+ - p_-$
- p_+ are transfer function poles w. positive residue
- p_- are transfer function poles w. negative residue

⁴⁴K.V. Fernando and H. Nicholson. **On the Cauchy Index of Linear Systems**. IEEE Transactions on Automatic Control , 28(2):222–224, 1983.

Information entropy via cross Gramian:

$$I := \frac{N}{2} \ln(2\pi e) + \frac{1}{2} \ln(\det(W_X))$$

- Minimizing information loss⁴⁵.
- Meaning minimizing steady-state information entropy.
- Basically Kullback-Leibler divergence minimization.

⁴⁵J. Fu, C. Zhong, Y. Ding, J. Zhou and C. Zhong. **An information theoretic approach to model reduction based on frequency-domain Cross-Gramian information**. In 8th World Congress on Intelligent Control and Automation: 3679–3683, 2010

- Linear Gramian-Based Model Reduction
 - Balanced Truncation
 - Singular Perturbation
 - Approximate Balancing
 - Direct Truncation
- Nonlinear Gramian-Based Model Reduction
 - Empirical Balanced Truncation
 - Empirical Direct Truncation
- Parameter Identification and Parameter Reduction
- Gramian-Based Combined State and Parameter Reduction

- Balanced truncation is the gold standard.
- Use algebraic methods for plain linear systems.
- Empirical Gramians are not always the best,
■ but are almost always computable.
- Empirical Gramians are all about averaging.
- The more you know about the operating region the better.
- The quality of simulations determines the quality of the ROM.

... but also interesting:

- Descriptor Systems
- Unstable Systems
- Bilinear Systems
- Quadratic Systems
- Switched Systems
- Active Subspaces
- ...

- Slides: <http://himpe.science/talks/himpe17-issac.pdf>
- Lab Session: 2017-09-12, 1400-1730
- Questions?

<http://himpe.science>

Acknowledgment:

Supported by the German Federal Ministry for Economic Affairs and Energy, in the joint project: “**MathEnergy** – Mathematical Key Technologies for Evolving Energy Grids”, sub-project: Model Order Reduction (Grant number: 0324019B).