Nonlinear balanced truncation: Computing energy functions and model reduction

Boris Kramer

Workshop on Nonlinear Model Reduction for Control
22-26 May 2023, Blacksburg, VA

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Mechanical and Aerospace Engineering

Acknowledgements

- Thanks for support through the National Science Foundation, division of Dynamics, Control and System Diagnostics for sponsoring this work and workshop under Grant CMMI-2130727
- Part of this work is based upon work supported by the National Science Foundation under Grant No. DMS-1929284 while the authors were in residence at ICERM in Providence, RI, during the Spring 2020 Semester Program "Model and dimension reduction in uncertain and dynamic systems" and Spring 2020 Reunion Event.

Nick Corbin
(UC San Diego)

Linus Balicki
(Virginia Tech)

Jeff Borggaard
(Virginia Tech)

Serkan Gugercin
(Virginia Tech)

Nonlinear control-affine model reduction

We are interested in high-dimensional nonlinear systems:

$$\mathbf{E}\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x})\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)),$$

with states $\mathbf{x} \in \mathbb{R}^n$, controls $\mathbf{u} \in \mathbb{R}^m$, nonlinearity $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$, outputs $\mathbf{y} \in \mathbb{R}^p$.

- For the large-scale systems of interest, i.e., semi-discretized PDEs, and differential algebraic equations (DAEs), $n \gg 1,000$.
- We assume invertible $\mathbf{E}$ matrices, and present the methods for $\mathbf{E} = \mathbf{I}$, which can be obtained with a suitable change of variables (see numerical examples).

Goal of control-affine nonlinear model reduction:

Find a low-dimensional coordinate transformation $\mathbf{x} \approx \Phi(\mathbf{z}_r)$, $\mathbf{z}_r \in \mathbb{R}^r$ and derive a reduced-order model (ROM)

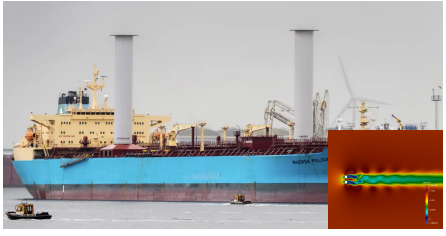
$$\dot{\mathbf{z}}_r(t) = \mathbf{f}_r(\mathbf{z}_r(t)) + \mathbf{g}_r(\mathbf{z}_r)\mathbf{u}(t), \qquad \mathbf{y}_r(t) = \mathbf{h}_r(\mathbf{z}_r(t)),$$

with reduced states $\mathbf{z}_r \in \mathbb{R}^r$ with $r \ll n$, such that $\|\mathbf{x} - \Phi(\mathbf{z}_r)\|_{\mathcal{X}}$ or $\|\mathbf{y} - \mathbf{y}_r\|_{\mathcal{X}}$ are small (in some norm $\|\cdot\|_{\mathcal{X}}$), and where the system has favorable control theoretic properties.

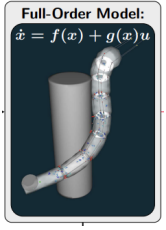
Motivation: control-oriented model reduction

For input-driven and controlled systems, taking into account the effects of the inputs & controls in the model reduction process is paramount.

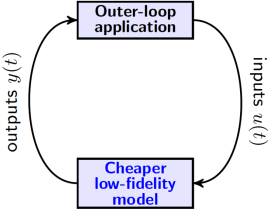
- **Trajectory-based methods** (proper orthogonal decomposition, reduced basis method, dynamic mode decomposition, ...) require carefully choosing representative forcing functions/initial conditions.
- **System-theoretic methods** use the underlying transfer function ($\mathcal{H}_2$, $\mathcal{H}_\infty$, Loewner), moments thereof, or system energies (balanced truncation) to find appropriate subspaces for projection.



`www.norsepower.com/technology`: Flettner rotors



Soft robots



generic control system

In these next two hours, we will discuss:

1. Various energy functions for nonlinear systems

2. Scalable computation of energy functions via polynomial approximations and tensor calculus[1]

3. Simultaneous balance-and-reduce strategy: ROMs on nonlinear balanced manifolds[2]

---

[1] K./Gugercin/Borggaard, *Nonlinear Balanced Truncation: Part 1—Computing Energy Functions, arxiv:2209.07645*

[2] K./Gugercin/Borggaard, *Nonlinear Balanced Truncation: Part 2—Model Reduction on Manifolds, arXiv:2302.02036*

Part 1:
Energy functions for nonlinear systems

Controllers and energy functions for nonlinear systems

Theorem ([Lukes, 1969])

Consider a control-affine nonlinear dynamical system and a quadratic cost (or energy)

$$\widehat{\mathcal{E}}(\mathbf{x}_0, \mathbf{u}) = \frac{1}{2}\int_0^\infty \mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t)\mathrm{d}t, \qquad \mathbf{Q}, \mathbf{R} \succ \mathbf{0}.$$

Let the following assumptions hold: (1) there is a neighborhood Ω of the origin where $\mathbf{f} \in C^2(\Omega)$; $\mathbf{f}(\mathbf{0}) = \mathbf{0}$; (2) the pair $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{0}), \mathbf{g}(\mathbf{0})\right)$ is stabilizable; (3) the nonlinear system is stabilizable on Ω, so there exists a stabilizing controller so that the closed-loop system is asymptotically stable on Ω. Then there exists a unique solution $\mathbf{u}^*(\mathbf{x})$ to the HJB equation

$$0 = \min_{\mathbf{u}} \left\{ \mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t) + \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}} \left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}\right] \right\}$$

where $\mathcal{E}(\mathbf{x}) = \min_{\mathbf{u}} \widehat{\mathcal{E}}(\mathbf{x}, \mathbf{u})$ and the unique continuously differentiable minimizer for the optimal feedback control $\mathbf{u}^*(\mathbf{x})$ is

$$\mathbf{u}^*(\mathbf{x}) = -\mathbf{R}^{-1}\mathbf{g}(\mathbf{x})^\top \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}}.$$

Moreover, if $\mathbf{f}(\mathbf{x})$ is analytic, so are $\mathbf{u}^*(\mathbf{x})$ and $\mathcal{E}(\mathbf{x})$.

Controllers and energy functions for nonlinear systems ctd

Inserting the optimal control $\mathbf{u}^*(\mathbf{x}) = -\mathbf{R}^{-1}\mathbf{g}(\mathbf{x})^\top \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}}$ into the HJB equation we obtain

$$0 = \mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t) + \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}} \left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \left(-\mathbf{R}^{-1}\mathbf{g}(\mathbf{x})^\top \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}}\right)\right]$$

which after reorganizing becomes

$$0 = \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{R}^{-1}\mathbf{g}(\mathbf{x})^\top \frac{\partial \mathcal{E}(\mathbf{x})}{\partial \mathbf{x}} + \mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t)$$

- The HJB equation is therefore a necessary and sufficient condition to the optimal control problem

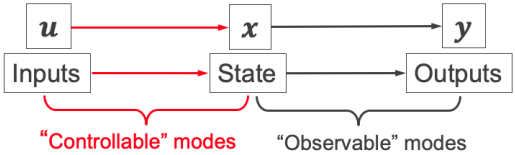
$$\begin{aligned} &\min_u \widehat{\mathcal{E}}(\mathbf{x}_0, \mathbf{u}) \\ &\text{s.t.} \qquad \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x})\mathbf{u}(t), \qquad \mathbf{x}_0 = \mathbf{x}(0) \end{aligned}$$

- Analytic solutions: since we assume polynomial dynamics going forward, $\mathbf{f}(\mathbf{x})$ is analytic, so we know that we can search for Taylor series of $\mathbf{u}^*(\mathbf{x})$ and $\mathcal{E}(\mathbf{x})$.

Energy functions for LTI systems

Consider a linear time-invariant system:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)$$



The energy to reach $\mathbf{x}_0$ from zero, and the observability energy associated with state $\mathbf{x}_0$, can be defined as

$$\mathcal{E}_c(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2(-\infty,0] \\ \mathbf{x}(-\infty)=\mathbf{0} \\ \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2}\int_{-\infty}^{0} \|\mathbf{u}(t)\|^2 \mathrm{d}t, \qquad \mathcal{E}_o(\mathbf{x}_0) := \frac{1}{2}\int_{0}^{\infty} \|\mathbf{y}(t)\|^2 \mathrm{d}t$$

It can be shown that they are **quadratic functions of the state**:

$$\mathcal{E}_c(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}_0^\top \mathbf{P}^{-1}\mathbf{x}_0, \qquad \mathcal{E}_o(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}_0^\top \mathbf{Q}\mathbf{x}_0,$$

where *controllability (observability) Gramians* $\mathbf{P}, \mathbf{Q}$ are solutions to the Lyapunov equations:

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^\top + \mathbf{B}\mathbf{B}^\top = \mathbf{0}, \qquad \mathbf{A}^\top\mathbf{Q} + \mathbf{Q}\mathbf{A} + \mathbf{C}^\top\mathbf{C} = \mathbf{0}.$$

What does this have to do with model reduction?

Let us decompose the (symmetric positive definite) controllability Gramian using the SVD:

$$\mathbf{P} = \mathbf{W}\mathbf{\Sigma}\mathbf{W}^\top, \qquad \mathbf{W}^\top\mathbf{W} = \mathbf{I}_n, \quad \mathbf{\Sigma} = \mathrm{diag}(\sigma_1, ..., \sigma_n).$$

The energy to reach a state $\mathbf{x}_0 = \mathbf{w}_i$ (a column of $\mathbf{W}$) from $\mathbf{x}(-\infty) = \mathbf{0}$ is:

$$\mathcal{E}_c(\mathbf{w}_i) = \mathbf{w}_i^\top \mathbf{P}^{-1}\mathbf{w}_i = \mathbf{w}_i^\top \mathbf{W}\mathbf{\Sigma}^{-1}\mathbf{W}^\top\mathbf{w}_i = \frac{1}{\sigma_i},$$

so the energy to reach $\mathbf{w}_i$ is given by $\frac{1}{\sigma_i}$. This leads us to make two observations:

1. "Easy" to reach states correspond to large σ_i.

2. "Hard" to reach states correspond to small σ_i.

What does this have to do with model reduction?

Similar observations can be made for the observability Gramian. Let us decompose the observability Gramian using its SVD:

$$\mathbf{Q} = \mathbf{V}\mathbf{S}\mathbf{V}^\top, \mathbf{V}^\top\mathbf{V} = \mathbf{I}, \qquad \mathbf{V}^\top\mathbf{V} = \mathbf{I}_n, \quad \mathbf{\Sigma} = \mathrm{diag}(\sigma_1, ..., \sigma_n)$$

The output energy that is generated by $\mathbf{x}_0 = \mathbf{v}_i$ (the observability energy) is

$$\mathcal{E}_o(\mathbf{v}_i) = \mathbf{v}_i^\top \mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top \mathbf{v}_i = \sigma_i.$$

We can make similar observations:

1. The eigenvectors corresponding to large σ_i are easy to observe.
2. The eigenvectors corresponding to small σ_i are hard to observe.

Balanced truncation model reduction

We want to find a coordinate system (i.e., a state-space transformation) where states are easy to reach and easy to observe. [Moore, 1981] pioneered balanced truncation for LTI systems:

- A **linear transformation** $\mathbf{x} = \mathbf{T}\mathbf{z}$ simultaneously diagonalizes $\mathbf{P}, \mathbf{Q}$.
- Truncating the balanced high-dimensional model yields a **balanced ROM with states that are easy to reach and easy to control**.

Model reduction for nonlinear systems via energy functions

Energy-function-based approaches

- [Scherpen, 1993] introduced the concept of nonlinear balancing via energy functions for (locally) stable, open-loop nonlinear systems.
- HJB balancing [Scherpen and Van der Schaft, 1994], $\mathcal{H}_\infty$ balancing [Scherpen, 1996]
- [Newman and Krishnaprasad, 2000] : controllability energy function is related to the stationary density p_∞ of a Markov process; suggest to solve Fokker-Planck equations
- Symbolic computing toolbox: [Krener, 2008]
- [Fujimoto and Tsubakino, 2008] use Taylor series for open-loop controllability and observability energy functions ($n = 4$)
- Interpretation from a Hankel singular value perspective: [Fujimoto and Scherpen, 2010]
- Machine-learning for balancing transformation based on RKHS [Bouvrie and Hamzi, 2017]

Gramian-based approaches (linear transformation $\Rightarrow$ quadratic energy function)

- Empirical Gramians for nonlinear systems [Lall et al., 2002]
- Algebraic Gramians for local balancing [Gray and Verriest, 2006, Benner and Goyal, 2017, Kramer and Willcox, 2019].

Nonlinear open-loop observability & controllability energy fcts.

For *stable* nonlinear systems controllability and observability energy functions are fully nonlinear, and can be defined [Scherpen, 1993] as

$$\mathcal{E}_c(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2(-\infty,0] \\ \mathbf{x}(-\infty)=\mathbf{0} \\ \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2}\int_{-\infty}^{0} \|\mathbf{u}(t)\|^2 \mathrm{d}t, \qquad \mathcal{E}_o(\mathbf{x}_0) := \frac{1}{2}\int_{0}^{\infty} \|\mathbf{y}(t)\|^2 \mathrm{d}t$$

- $\mathcal{E}_c(\mathbf{x}_0)$: minimum energy to steer system from $\mathbf{x}(-\infty)=\mathbf{0}$ to $\mathbf{x}(0)=\mathbf{x}_0$.
- $\mathcal{E}_o(\mathbf{x}_0)$: output energy generated by $\mathbf{x}_0 \neq 0$ and $\mathbf{u}(t)\equiv\mathbf{0}$.

Energy functions are solutions to Hamilton-Jacobi equations:

$$\begin{aligned} 0 =& \frac{\partial \mathcal{E}_o(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x}), \\ 0 =& \frac{\partial \mathcal{E}_c(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\frac{\partial \mathcal{E}_c(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_c(\mathbf{x})}{\partial \mathbf{x}}. \end{aligned}$$

- $\mathcal{E}_o$ exists if $\mathbf{f}$ is asymptotically stable in a neighborhood of the origin
- $\mathcal{E}_c$ exists if $-\left(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_c(\mathbf{x})}{\partial \mathbf{x}}\right)$ is asympt. stable in a neighborhood of origin.

HJB-balancing energy functions

HJB balancing [Scherpen and Van der Schaft, 1994] (applicable to unstable systems) defines the *past* and *future energy function* as

$$\mathcal{E}^-(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2(-\infty,0]\\ \mathbf{x}(-\infty)=\mathbf{0}\\ \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2}\int_{-\infty}^{0} \|\mathbf{y}(t)\|^2 + \|\mathbf{u}(t)\|^2 \mathrm{d}t$$

$$\mathcal{E}^+(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2[0,\infty)\\ \mathbf{x}(0)=\mathbf{x}_0\\ \mathbf{x}(\infty)=\mathbf{0}}} \frac{1}{2}\int_{0}^{\infty} \|\mathbf{y}(t)\|^2 + \|\mathbf{u}(t)\|^2 \mathrm{d}t$$

and they are solutions to the Hamilton-Jacobi-Bellman equation

$$0 = \frac{\partial \mathcal{E}^-(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\frac{\partial \mathcal{E}^-(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}^-(\mathbf{x})}{\partial \mathbf{x}} - \frac{1}{2}\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x})$$

$$0 = \frac{\partial \mathcal{E}^+(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{1}{2}\frac{\partial \mathcal{E}^+(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}^+(\mathbf{x})}{\partial \mathbf{x}} + \frac{1}{2}\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x}).$$

Note: LQG-balancing [Verriest, 1981, Jonckheere and Silverman, 1983] and HJB-balancing are identical concepts for linear systems.

$\mathcal{H}_\infty$ energy functions

Definition [Scherpen, 1996]

For a nonlinear system, the $\mathcal{H}_\infty$ past energy in the state $\mathbf{x}_0$ is defined for $0 < \gamma \neq 1$ as

$$\mathcal{E}_\gamma^-(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2(-\infty,0] \\ \mathbf{x}(-\infty)=\mathbf{0},\ \mathbf{x}(0)=\mathbf{x}_0}} \frac{1}{2}\int_{-\infty}^{0} (1-\gamma^{-2})\|\mathbf{y}(t)\|^2 + \|\mathbf{u}(t)\|^2 \mathrm{d}t$$

and the $\mathcal{H}_\infty$ future energy in the state $\mathbf{x}_0$ is defined for $\gamma > 1$ as

$$\mathcal{E}_\gamma^+(\mathbf{x}_0) := \min_{\substack{\mathbf{u}\in L_2[0,\infty) \\ \mathbf{x}(0)=\mathbf{x}_0,\ \mathbf{x}(\infty)=\mathbf{0}}} \frac{1}{2}\int_{0}^{\infty} \|\mathbf{y}(t)\|^2 + \left(\frac{1}{1-\gamma^{-2}}\right)\|\mathbf{u}(t)\|^2 \mathrm{d}t$$

and for $0 < \gamma < 1$ as

$$\mathcal{E}_\gamma^+(\mathbf{x}_0) := \max_{\substack{\mathbf{u}\in L_2[0,\infty) \\ \mathbf{x}(0)=\mathbf{x}_0,\ \mathbf{x}(\infty)=\mathbf{0}}} \frac{1}{2}\int_{0}^{\infty} \|\mathbf{y}(t)\|^2 + \left(\frac{1}{1-\gamma^{-2}}\right)\|\mathbf{u}(t)\|^2 \mathrm{d}t.$$

Hamilton-Jacobi equations for $\mathcal{H}_\infty$ balancing

Theorem [Scherpen, 1996, Thm 5.2]

Assume that the HJB equation

$$0 = \frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\frac{\partial \mathcal{E}_\gamma^-}{\partial \mathbf{x}}(\mathbf{x})\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} - \frac{1}{2}(1-\gamma^{-2})\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x})$$

has a solution with $\mathcal{E}_\gamma^-(\mathbf{0}) = 0$ that also satisfies that $-\left(\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}}\right)$ is asymptotically stable. Then this solution is the past energy function $\mathcal{E}_\gamma^-(\mathbf{x})$. Furthermore, assume that the HJB equation

$$0 = \frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{1}{2}(1-\gamma^{-2})\frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^+}{\partial \mathbf{x}}(\mathbf{x}) + \frac{1}{2}\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x})$$

has a solution with $\mathcal{E}_\gamma^+(\mathbf{0}) = 0$ which satisfies that $\left(\mathbf{f}(\mathbf{x}) - (1-\gamma^{-2})\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^+}{\partial \mathbf{x}}(\mathbf{x})\right)$ is asymptotically stable. Then this solution is the future energy function.

Observations

1. For $\gamma = \sqrt{1/2}$ the past and future energy functions are identical:

$$\mathcal{E}^-_{\gamma=\frac{1}{\sqrt{2}}}(\mathbf{x}) = \mathcal{E}^+_{\gamma=\frac{1}{\sqrt{2}}}(\mathbf{x}).$$

2. The $\mathcal{H}_\infty$ energy functions are related to HJB balancing as [Scherpen, 1996]

$$\lim_{\gamma\to\infty} \mathcal{E}^-_\gamma(\mathbf{x}) = \mathcal{E}^-(\mathbf{x}). \qquad \lim_{\gamma\to\infty} \mathcal{E}^+_\gamma(\mathbf{x}) = \mathcal{E}^+(\mathbf{x}).$$

3. Under certain technical conditions we also have that the $\mathcal{H}_\infty$ energy functions approach the standard open-loop balancing energy functions:

$$\lim_{\gamma\to 1} \mathcal{E}^+_\gamma(\mathbf{x}) = \mathcal{E}_o(\mathbf{x}), \qquad \lim_{\gamma\to 1} \mathcal{E}^-_\gamma(\mathbf{x}) = \mathcal{E}_c(\mathbf{x}).$$

4. For an LTI system, $\mathcal{E}^-_\gamma(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}^\top \mathbf{Y}_\infty^{-1}\mathbf{x}$ and $\mathcal{E}^+_\gamma(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}^\top \mathbf{X}_\infty \mathbf{x}$, where $\mathbf{Y}_\infty,\ \mathbf{X}_\infty$ are the usual stabilizing positive definite solutions to the $\mathcal{H}_\infty$ AREs

$$\mathbf{A}\mathbf{Y}_\infty + \mathbf{Y}_\infty\mathbf{A}^\top + \mathbf{B}\mathbf{B}^\top - (1-\gamma^{-2})\mathbf{Y}_\infty\mathbf{C}^\top\mathbf{C}\mathbf{Y}_\infty = \mathbf{0},$$
$$\mathbf{A}^\top\mathbf{X}_\infty + \mathbf{X}_\infty\mathbf{A} + \mathbf{C}^\top\mathbf{C} - (1-\gamma^{-2})\mathbf{X}_\infty\mathbf{B}\mathbf{B}^\top\mathbf{X}_\infty = \mathbf{0}.$$

Example: One-dimensional quadratic dynamical system

Consider the equation

$$\dot{x}(t) = ax(t) + \mathrm{n}x(t)^2 + bu(t), \qquad y(t) = cx(t).$$

With $\eta = 1 - \gamma^{-2}$ the HJB equation is

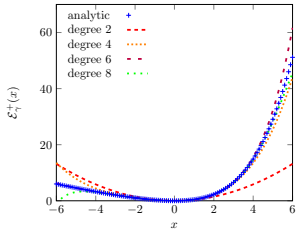
$$0 = \frac{\mathrm{d}\,\mathcal{E}_\gamma^+}{\mathrm{d}\,x}(x)[ax + \mathrm{n}x^2] - \frac{1}{2}b^2\eta\left(\frac{\mathrm{d}\,\mathcal{E}_\gamma^+}{\mathrm{d}\,x}(x)\right)^2 + \frac{1}{2}c^2x^2,$$

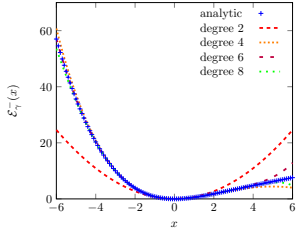
with analytical solution

$$\mathcal{E}_\gamma^+(x) = \frac{1}{b^2\eta}\left(\mp\frac{ab^2c^2\eta\sqrt{x^2((a+\mathrm{n}x)^2+b^2c^2\eta)}\log\left(\sqrt{(a+\mathrm{n}x)^2+b^2c^2\eta}+a+\mathrm{n}x\right)}{2n^2x\sqrt{(a+\mathrm{n}x)^2+b^2c^2\eta}}\right.$$
$$\left.\pm\frac{\sqrt{x^2((a+\mathrm{n}x)^2+b^2c^2\eta)}\left(\frac{(a+\mathrm{n}x)^2}{3\mathrm{n}}-\frac{a(a+\mathrm{n}x)}{2\mathrm{n}}+\frac{b^2c^2\eta}{3n}\right)}{\mathrm{n}x}+\frac{ax^2}{2}+\frac{\mathrm{n}x^3}{3}\right)$$

Figure: energy functions, $a = -2$, $b = 2$, $\mathrm{n} = 1, c = 2, \gamma = \sqrt{2}$.

Need higher degree terms to approximate energy functions!



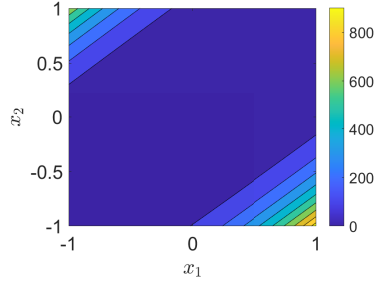


Two-dimensional example

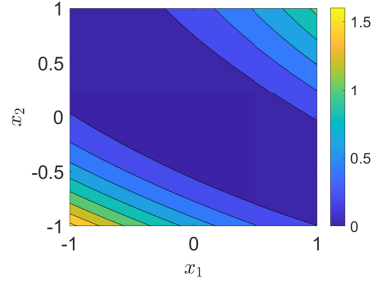
We modify the 2d nonlinear example from [Kawano and Scherpen, 2016, IV.C] :

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} -x_2^2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{u}, \qquad \mathbf{y} = [1\ 1]\mathbf{x}$$

- Plot for polynomial approximations with $d = 4$ for energy functions, and $\eta = 0.1$ ($\gamma \approx 1.054$) and
- Quadratic approximations again would not be sufficient (higher order terms needed)



$\mathcal{E}_\gamma^-(\mathbf{x})$



$\mathcal{E}_\gamma^+(\mathbf{x})$

Worksheet: Assume LTI system and obtain HJB solution

Part 2:
Computing energy functions via polynomial approximations

Notation and setting

- **Polynomial nonlinear systems** for scalability:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \sum_{k=2}^{\ell} \mathbf{F}_k \mathbf{x}^{\textcircled{k}}(t) + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

  For now, assume quadratic nonlinear system ($\ell = 2$ and $\mathbf{F}_2 = \mathbf{F}$);

- Define k-term Kronecker product of $\mathbf{x}$:

$$\mathbf{x}^{\textcircled{k}} := \underbrace{\mathbf{x} \otimes \ldots \otimes \mathbf{x}}_{k \text{ times}}.$$

- Define the *d-way Lyapunov matrix*/special *Kronecker sum*:

$$\mathcal{L}_d(\mathbf{A}) := \underbrace{\mathbf{A} \otimes \ldots \otimes \mathbf{I}}_{d \text{ times}} + \cdots + \underbrace{\mathbf{I} \otimes \ldots \otimes \mathbf{A}}_{d \text{ times}}.$$

- Define $\eta := (1 - \gamma^{-2})$ and note that $\eta \in (-\infty, 1)$ since $\gamma > 0$.

Symmetry considerations

For convenience and to ensure a unique representation of the coefficients, we impose symmetry of our coefficients in all monomial terms in the energy functions.

Definition (Symmetric Coefficients)

A monomial term with real coefficients $\mathbf{w}_d^\top \mathbf{x}^{\textcircled{d}}$ has *symmetric coefficients* if it satisfies

$$\mathbf{w}_d^\top (\mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \cdots \otimes \mathbf{a}_d) = \mathbf{w}_d^\top (\mathbf{a}_{i_1} \otimes \mathbf{a}_{i_2} \otimes \cdots \otimes \mathbf{a}_{i_d}),$$

where the indices $\{i_k\}_{k=1}^d$ are any permutation of $1, \ldots, d$.

- This definition generalizes the definition of symmetry from matrices to tensors. For example,

$$\mathbf{w}_2^\top (\mathbf{a} \otimes \mathbf{b}) = \mathbf{w}_2^\top (\mathbf{b} \otimes \mathbf{a}) \quad \forall \mathbf{a}, \mathbf{b} \quad \Leftrightarrow \quad (\mathbf{a}^\top \otimes \mathbf{b}^\top)\mathbf{w}_2 = (\mathbf{b}^\top \otimes \mathbf{a}^\top)\mathbf{w}_2.$$

  Hence, using $\mathbf{w}_2 = \text{vec}(\mathbf{W}_2)$, we have $\mathbf{b}^\top \mathbf{W}_2 \mathbf{a} = \mathbf{a}^\top \mathbf{W}_2 \mathbf{b}$. Since these are real scalars, this implies $\mathbf{W}_2 = \mathbf{W}_2^\top$.

Symmetry considerations

We also remark that any polynomial can be uniquely written in Kronecker product form with symmetric coefficients. For example,

$$c_1 x_1^2 + c_2 x_1 x_2 + c_3 x_2^2 = [x_1 \ x_2] \begin{bmatrix} c_1 & \frac{1}{2}c_2 \\ \frac{1}{2}c_2 & c_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c_1 & \frac{1}{2}c_2 & \frac{1}{2}c_2 & c_3 \end{bmatrix} (\mathbf{x} \otimes \mathbf{x}) = \mathbf{F}\mathbf{x}^{\textcircled{2}}.$$

The same set of quadratic terms would be realized with the coefficient matrices corresponding to either $[c_1 \ c_2 \ 0 \ c_3]$ or $[c_1 \ 0 \ c_2 \ c_3]$. However, the requirement of symmetry leads to a unique representation.

- We assume that each row of the coefficient matrices $\mathbf{F}_k$ in the FOM is symmetric as defined above, and that the polynomial representations of the energy functions and controls share this symmetric representation.
- Our algorithms are designed to ensure symmetry in the computed coefficients.

Expansion of future energy function

We approximate the future energy function as

$$\mathcal{E}_\gamma^+(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{w}_2^\top \mathbf{x}^{②} + \mathbf{w}_3^\top \mathbf{x}^{③} + \ldots + \mathbf{w}_d^\top \mathbf{x}^{ⓓ}\right) = \frac{1}{2}\left(\mathbf{w}_2^\top + \tilde{\mathbf{w}}_3^\top(\mathbf{x}) + \ldots + \tilde{\mathbf{w}}_d^\top(\mathbf{x})\right)\mathbf{x}^{②}.$$

Theorem (K./Gugercin/Borggaard/Balicki '22)

Let $\gamma > \gamma_0 > 0$, (can be computed), $\eta = 1 - \gamma^{-2}$. Let the future energy $\mathcal{E}_\gamma^+(\mathbf{x})$ for the quadratic nonlinear system ($\ell = 2$ and $\mathbf{F}_2 = \mathbf{F}$) be expanded with coefficients $\mathbf{w}_i, i = 2, \ldots, d$. Then, $\mathbf{w}_2 = vec(\mathbf{W}_2)$ where $\mathbf{W}_2$ is the s.p.d. solution to the $\mathcal{H}_\infty$ Riccati equation

$$\mathbf{0} = \mathbf{A}^\top \mathbf{W}_2 + \mathbf{W}_2 \mathbf{A} + \mathbf{C}^\top \mathbf{C} - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top \mathbf{W}_2.$$

For $2 < k \leq d$, let $\tilde{\mathbf{w}}_k \in \mathbb{R}^{n^k}$ solve the linear system

$$\mathcal{L}_k(\mathbf{A}^\top - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{w}}_k = -\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{w}_{k-1} + \frac{\eta}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \ \mathrm{vec}(\mathbf{W}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{W}_j).$$

Then, the coefficient vector $\mathbf{w}_k = \mathrm{vec}(\mathbf{W}_k) \in \mathbb{R}^{n^k}$ is obtained by symmetrizing $\tilde{\mathbf{w}}_k$.

Proof idea

From the polynomial energy function it follows that

$$\begin{aligned}\frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}} =& \frac{1}{2}\left(\mathbf{w}_2^\top(\mathbf{I}\otimes\mathbf{x}) + \mathbf{w}_2^\top(\mathbf{x}\otimes\mathbf{I})\right.\\ &+\mathbf{w}_3^\top(\mathbf{I}\otimes\mathbf{x}\otimes\mathbf{x}) + \mathbf{w}_3^\top(\mathbf{x}\otimes\mathbf{I}\otimes\mathbf{x}) + \mathbf{w}_3^\top(\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{I})\\ &+\mathbf{w}_4^\top(\mathbf{I}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}) + \mathbf{w}_4^\top(\mathbf{x}\otimes\mathbf{I}\otimes\mathbf{x}\otimes\mathbf{x}) + \mathbf{w}_4^\top(\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{I}\otimes\mathbf{x}) + \mathbf{w}_4^\top(\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{I})\\ &\left.+\cdots\right).\end{aligned}$$

Given $\mathbf{g}(\mathbf{x}) = \mathbf{B}$, $\mathbf{h}(\mathbf{x}) = \mathbf{C}\mathbf{x}$, $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{F}(\mathbf{x}\otimes\mathbf{x})$ the HBJ reads as

$$0 = \frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}}(\mathbf{A}\mathbf{x} + \mathbf{F}(\mathbf{x}\otimes\mathbf{x})) - \frac{1}{2}(1-\gamma^{-2})\frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}}\mathbf{B}\mathbf{B}^\top\frac{\partial^\top \mathcal{E}_\gamma^+}{\partial \mathbf{x}}(\mathbf{x}) + \frac{1}{2}\mathbf{x}^\top\mathbf{C}^\top\mathbf{C}\mathbf{x},$$

- We now collect terms by degree of $\mathbf{x}$, starting with quadratic, to cubic, to higher-order.
- We can pull out $\mathbf{x}^{\textcircled{k}}$'s etc and set terms inside to zero (similar to what you did for the LTI into HJB example)

Expansion of past energy function

We approximate the past energy function as

$$\mathcal{E}_\gamma^-(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{②} + \mathbf{v}_3^\top \mathbf{x}^{③} + \cdots \mathbf{v}_d^\top \mathbf{x}^{ⓓ}\right).$$

Theorem (K./Gugercin/Borggaard/Balicki '22)

Let $\gamma > \gamma_0 > 0$, (can be computed), $\eta = 1 - \gamma^{-2}$. Let the past energy function $\mathcal{E}_\gamma^-(\mathbf{x})$ for the quadratic nonlinear system ($\ell = 2$ and $\mathbf{F}_2 = \mathbf{F}$) be expanded as above with the coefficients $\mathbf{v}_i, i = 2, 3, \ldots, d$. Then, $\mathbf{v}_2 = \mathrm{vec}(\mathbf{V}_2)$ where $\mathbf{V}_2$ is the symmetric positive definite solution to the $\mathcal{H}_\infty$ Riccati equation

$$\mathbf{0} = \mathbf{A}^\top \mathbf{V}_2 + \mathbf{V}_2 \mathbf{A} - \eta \mathbf{C}^\top \mathbf{C} + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top \mathbf{V}_2.$$

For $2 < k \leq d$, let let $\tilde{\mathbf{v}}_k \in \mathbb{R}^{n^k}$ solve the linear system

$$\mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{v}}_k = -\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{v}_{k-1} - \frac{1}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \ \mathrm{vec}(\mathbf{V}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{V}_j).$$

Then, the coefficient vector $\mathbf{v}_k = \mathrm{vec}(\mathbf{V}_k) \in \mathbb{R}^{n^k}$ is obtained by the symmetrizing $\tilde{\mathbf{v}}_k$.

Algorithm for Energy Function Approximation

Algorithm 1 Computing HJB energy function approximations: $\mathcal{E}_\gamma^-(\mathbf{x})$ and $\mathcal{E}_\gamma^+(\mathbf{x})$.

Input: System matrices $\mathbf{A}, \mathbf{F}, \mathbf{B}, \mathbf{C}$; polynomial degree d; constant $\gamma > \gamma_0 > 0$, $\gamma \neq 1$.

Output: Coefficients $\{\mathbf{v}_i\}_{i=2}^d$ of the past energy and $\{\mathbf{w}_i\}_{i=2}^d$ of the future energy functions.

1: Set $\eta = (1 - \gamma^{-2})$.

2: Solve the $\mathcal{H}_\infty$ Riccati equations

$$\mathbf{0} = \mathbf{A}^\top \mathbf{V}_2 + \mathbf{V}_2 \mathbf{A} - \eta \mathbf{C}^\top \mathbf{C} + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top \mathbf{V}_2,$$
$$\mathbf{0} = \mathbf{A}^\top \mathbf{W}_2 + \mathbf{W}_2 \mathbf{A} + \mathbf{C}^\top \mathbf{C} - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top \mathbf{W}_2$$

and set $\mathbf{v}_2 = \mathrm{vec}(\mathbf{V}_2)$ and $\mathbf{w}_2 = \mathrm{vec}(\mathbf{W}_2)$.

3: **For** $k = 3, 4, \ldots, d$: Solve the systems for $\tilde{\mathbf{v}}_d$ and $\tilde{\mathbf{w}}_d$:

$$\mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{v}}_k = -\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{v}_{k-1} - \frac{1}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \ \mathrm{vec}(\mathbf{V}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{V}_j)$$

$$\mathcal{L}_k(\mathbf{A}^\top - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{w}}_k = -\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{w}_{k-1} + \frac{\eta}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \ \mathrm{vec}(\mathbf{W}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{W}_j)$$

4: Symmetrize $\tilde{\mathbf{w}}_k$ and $\tilde{\mathbf{v}}_k$ to obtain $\mathbf{w}_k$ and $\mathbf{v}_k$.

Solvability of the coefficient systems

Theorem [Thm 8, K./Gugerin/Borggaard/Balicki]

Let $\gamma > \gamma_0 \geq 0$ hold so that the $\mathcal{H}_\infty$ ARE is solvable and $\eta = 1 - \gamma^{-2}$. Then, for any $k = 1, \ldots, d$ the matrices $\mathcal{L}_k(\mathbf{A}^\top - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top)$ and $\mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top)$ are invertible, thus the coefficients $\mathbf{w}_i$ and $\mathbf{v}_i$ are uniquely determined.

Proof idea: A result from [Horn et al., 1994] states that for any $\mathbf{M} \in \mathbb{R}^{n \times n}$ the spectrum

$$\Lambda\left(\mathcal{L}_k(\mathbf{M})\right) = \left\{ \sum_{i \in \mathcal{P}_k} \lambda_i : \lambda_i \in \Lambda(\mathbf{M}) \right\},$$

where $\mathcal{P}_k$ denotes the set of all possible selection of k-indices from the set $\{1, 2, \ldots, n\}$. Since $\mathbf{W}_2$ is the unique stabilizing solution of the $\mathcal{H}_\infty$ Riccati equation, $\mathbf{M} = \mathbf{A}^\top - \eta \mathbf{W}_2 \mathbf{B}\mathbf{B}^\top$ has all eigenvalues in the open left-half plane. Therefore, all eigenvalues of $\mathcal{L}_k(\mathbf{M})$ are contained in the open left-half plane as well, thus $\mathcal{L}_k(\mathbf{M})$ is invertible and the $\tilde{\mathbf{w}}_k$ can be uniquely determined. For the second statement, we use that

$$\mathbf{A}^\top + \mathbf{V}_2 \mathbf{B}\mathbf{B}^\top = -\mathbf{V}_2 \mathbf{A} \mathbf{V}_2^{-1} + \eta \mathbf{C}^\top \mathbf{C} \mathbf{V}_2^{-1}.$$

and use similar arguments.

Solving Tensor Systems Efficiently

- For some $\mathbf{b}$, the kth-order polynomial terms require solving linear systems of the form

$$\mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2\mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{v}}_k = \mathbf{b},$$

  which grow exponentially in k and polynomially in n.

- We leverage the k-way Bartels-Stewart algorithm in [Borggaard and Zietsman, 2021]. By first performing a Schur factorization[3] of $\mathbf{A}^\top + \mathbf{V}_2\mathbf{B}\mathbf{B}^\top = \mathbf{U}\mathbf{T}\mathbf{U}^*$ and defining a matrix $\mathbf{U}^{\textcircled{k}} = \mathbf{U} \otimes \mathbf{U} \otimes \cdots \otimes \mathbf{U} \in \mathbb{R}^{n^k \times n^k}$, we convert the above linear system to

$$[\mathbf{U}^{\textcircled{k}}]^* \mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2\mathbf{B}\mathbf{B}^\top)[\mathbf{U}^{\textcircled{k}}]\hat{\mathbf{v}}_k = \hat{\mathbf{b}}, \quad \text{where} \quad \hat{\mathbf{v}}_k = [\mathbf{U}^{\textcircled{k}}]^*\tilde{\mathbf{v}}_k, \hat{\mathbf{b}} = [\mathbf{U}^{\textcircled{k}}]^*\mathbf{b}.$$

- Resulting system $\mathcal{L}_k(\mathbf{T})\hat{\mathbf{v}}_k = \hat{\mathbf{b}}$ is upper triangular and can be solved by a block backsubstitution procedure requiring n^{k-1} linear system solutions of size n.
- From here, we can compute the solution $\tilde{\mathbf{v}}_k = \mathbf{U}^{\textcircled{k}}\hat{\mathbf{v}}_k$.
- **Overall computational cost of solving k-th order system is $O(n^{k+1})$.**

[3] Note: Schur decomposition of $\mathbf{A}$ okay for medium-scale problems to avoid performing any operation in n^k-dim. space. For large-scale problems: iterative methods to exploit tensor structure such as the Krylov methods [Kressner and Tobler, 2010] or low-rank ADI type methods [Benner and Saak, 2013].

Total Cost of Solving for Coefficients

- Next, we have a look at the right-hand side of the linear systems:

$$\mathbf{b} = -\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{v}_{k-1} - \frac{1}{4} \sum_{\substack{i,j>2 \\ i+j=k+2}} ij \ \text{vec}(\mathbf{V}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{V}_j)$$

- We efficiently compute products $\mathcal{L}_{k-1}(\mathbf{F}^\top)\mathbf{v}_{k-1}$ in $O(kn^k)$; a direct product of the $n^{2k} \times n^k$ matrix times a vector would require $O(n^{3k})$ operations.
- Cost of forming the summation terms is dominated by multiplying the stored matrices $\mathbf{V}_i^\top \mathbf{B}$ and $\mathbf{B}^\top \mathbf{V}_j$. The cost of forming the summation terms are $O(kmn^k)$.
- We perform a final step to impose symmetry.
- **In sum, the computational complexity of computing a dth order approximation of the energy functions is (for $n > dm$):**

$$O(n^{d+1}) \qquad \text{(vs } O(n^{3d}) \text{ for naive implementation)}$$

Numerical Results: Burgers' equation

We consider the one-dimensional Burgers' equation

$$z_t(x,t) = \epsilon z_{xx}(x,t) - \frac{1}{2}\left(z^2(x,t)\right)_x + \sum_{j=1}^{m} b_j^m(x)u_j(t),$$

$$y_i(t) = \int_{\chi_{[(i-1)/p,\, i/p]}} z(x,t)\mathrm{d}x, \qquad i = 1,\ldots,p,$$

- periodic BCs $z(0,t) = z(1,t)$ and $z_x(0,t) = z_x(1,t)$
- IC: $z(\cdot,0) = z_0(\cdot) \in H_0^1(0,1)$
- $\epsilon = 0.001$ to make the nonlinearity significant.
- $p = 4$ outputs: spatial averages
- $m = 4$ controls/inputs with $b_j^m(x) = \chi_{[(j-1)/m,\, j/m]}(x)$.

The discretized system has the form

$$\widetilde{\mathbf{E}}\dot{\mathbf{z}} = \widetilde{\mathbf{A}}\mathbf{z} + \widetilde{\mathbf{N}}_2\left(\mathbf{z}\otimes\mathbf{z}\right) + \widetilde{\mathbf{B}}\mathbf{u}$$
$$\mathbf{y} = \widetilde{\mathbf{C}}\mathbf{z},$$

A change of variables $\mathbf{x} = \widetilde{\mathbf{E}}^{1/2}\mathbf{z}$ and redefining $\mathbf{A} = \mathbf{S}^{-1}\widetilde{\mathbf{A}}\mathbf{S}^{-1}$, $\mathbf{B} = \mathbf{S}^{-1}\widetilde{\mathbf{B}}$, $\mathbf{C} = \widetilde{\mathbf{C}}\mathbf{S}^{-1}$, $\widetilde{\mathbf{N}}_2 = \mathbf{N}_2(\mathbf{S}^{-1}\otimes\mathbf{S}^{-1})$ leads to a system with $\mathbf{E} = \mathbf{I}$.

Burgers' equation: Computing the energy functions

$d = 3$, **convergence w.r.t** n.

n	n^3	CPU sec	$\mathcal{E}_3^+(\mathbf{z}_0)$
8	5.1200e+02	2.96e-02	1.144557e-06
16	4.0960e+03	1.08e-02	1.116244e-06
32	3.2768e+04	5.96e-02	1.093503e-06
64	2.6214e+05	4.40e-01	1.099870e-06
128	2.0972e+06	4.29e+00	1.097715e-06
256	1.6777e+07	5.48e+01	1.095300e-06
512	1.3422e+08	6.63e+02	1.096322e-06
1024	1.0737e+09	7.93e+03	1.096093e-06

$n = 8$ **approximation w.r.t.** d.

d	$\mathcal{E}_d^-(\mathbf{z}_0)$	$\mathcal{E}_d^+(\mathbf{z}_0)$
2	3.161325e-05	1.146135e-06
3	2.731740e-05	1.144557e-06
4	2.370917e-05	1.144783e-06
5	2.593642e-05	1.144792e-06
6	2.662942e-05	1.144791e-06
7	2.519892e-05	1.144791e-06
8	2.538956e-05	1.144791e-06

Observations

- Convergence of the energy function as n increases (set gain $\eta = 0.9$ for HJB equation)
- Flop-count analysis predicts computational cost with growth of $O(n^4)$ (since $d = 3$), but CPU times scale as $O(n^{2.84})$. For $d = 4$ case, we find growth of $O(n^{3.57})$. This suggests that CPU time scales more like $O(n^d)$ for our problem sizes.
- First time where a high-resolution approximation of the cubic term in the energy function
- For $n = 1024$, this requires solving linear systems of size 10^9, which, through an efficient BLAS-3 level implementation can be performed in less than 5h CPU time.

Numerical Results: Kuramoto-Sivashinsky equation

Consider the domain $x \in (0, 1)$ and $t > 0$, and

$$z_t(x,t) = -\epsilon z_{xx}(x,t) - \epsilon^2 z_{xxxx}(x,t) - \epsilon(z(x,t)^2)_x + \sum_{j=1}^{m} b_j^m(x) u_j(t)$$

- periodic BCs $z(0,t) = z(1,t)$ and $z_x(0,t) = z_x(1,t)$
- same control input functions b_j^m and observation as for Burgers's equation
- $m = 5$ (five controls) and $p = 2$ (two outputs), and here choosing $\eta = 0.1$
- parameter $\epsilon = 1/13.0291^2$, which is known to exhibit heteroclinic cycles in the open-loop system
- IC: $z(x,0) = z_0(x) = \frac{0.01}{\sqrt{\epsilon}} \sin(4\pi x)$

Table: $d = 3$, convergence w.r.t n.

n	n^3	CPU sec	$\mathcal{E}_3^+(\mathbf{z}_0)$
16	4.0960e+03	1.20e-02	4.369195e+00
32	3.2768e+04	8.44e-02	5.099752e+00
64	2.6214e+05	5.54e-01	4.793412e+00
128	2.0972e+06	9.14e+00	4.732940e+00
256	1.6777e+07	1.37e+02	4.811878e+00
512	1.3422e+08	1.70e+03	4.827930e+00
1024	1.0737e+09	2.04e+04	4.807904e+00

Table: $n = 16$ approximation w.r.t. d.

d	$\mathcal{E}_d^+(\mathbf{z}_0)$	CPU sec
2	4.3690773e+00	6.81e-03
3	4.3691951e+00	9.88e-03
4	4.3469410e+00	1.37e-01
5	4.3467633e+00	2.40e+00
6	4.3467610e+00	4.39e+01

Part 3:
Simultaneous Balance-and-Reduce Model Reduction on Manifolds

Balancing an LTI system

Recall, that to get the quadratic energy functions, we had to solve the Lyapunov equations:

$$\mathbf{AP} + \mathbf{PA}^\top + \mathbf{BB}^\top = \mathbf{0}, \qquad \mathbf{A}^\top\mathbf{Q} + \mathbf{QA} + \mathbf{C}^\top\mathbf{C} = \mathbf{0}.$$

Definition (Balanced system)

An asymptotically stable LTI system is *balanced* if $\mathbf{P} = \mathbf{Q} = \mathbf{\Sigma} = \text{diag}(\sigma_1, ..., \sigma_n)$.

Theorem (Balancing transformation)

Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}]$ be asymptotically stable, controllable and observable. Let $\mathbf{P} = \mathbf{RR}^\top$, $\mathbf{Q} = \mathbf{LL}^\top$ be the Cholesky factorizations and $\mathbf{L}^\top\mathbf{R} = \mathbf{U\Sigma V}^\top$ and let

$$\mathbf{T} = \mathbf{RV\Sigma}^{-\frac{1}{2}}, \qquad \mathbf{T}^{-1} = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{U}^\top\mathbf{L}^\top.$$

Then $[\widetilde{\mathbf{A}}, \widetilde{\mathbf{B}}, \widetilde{\mathbf{C}}, \widetilde{\mathbf{D}}] = [\mathbf{T}^{-1}\mathbf{AT}, \mathbf{T}^{-1}\mathbf{B}, \mathbf{CT}, \mathbf{D}]$ is a balanced LTI system.

The controllability and observability energy functions of the balanced systems are then:

$$\mathcal{E}_c(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}_0^\top\mathbf{\Sigma}^{-1}\mathbf{x}_0, \qquad \mathcal{E}_o(\mathbf{x}_0) = \frac{1}{2}\mathbf{x}_0^\top\mathbf{\Sigma}\mathbf{x}_0.$$

Claim to fame: A few key results

Since the states of the balanced system are now ordered by observability/controllability properties, we delete the states that are not relevant, i.e., $\widetilde{\mathbf{x}} = [x_1, x_2, \ldots, x_r]$, with $r \ll n$ is the reduced state.

Theorem (Stability and minimality)

Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}]$ be an asymptotially stable and minimal system. Let the balanced ROM be $[\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r, \mathbf{D}_r]$ where $\sigma_r > \sigma_{r+1}$ for the Hankel singular values $\sigma_i, i = 1, \ldots, n$. Then, the ROM is asymptotically stable, minimal and balanced with Gramians $\mathbf{P}_r = \mathbf{Q}_r = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) =: \mathbf{\Sigma}_r$.

Theorem (Error bound)

Let $[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}]$ be asymptotically stable and balanced with controllability Gramian and observability Gramian $\mathbf{P} = \mathbf{Q} = \mathrm{diag}(\sigma_1 \mathbf{I}_{s_1}, \sigma_2 \mathbf{I}_{s_2}, \ldots, \sigma_k \mathbf{I}_{s_k})$($\sigma$ could be repeated), where $\sigma_1 > \sigma_2 > \ldots > \sigma_k \geqslant \mathbf{0}$. Let $[\mathbf{A}_r, \mathbf{B}_r, \mathbf{C}_r, \mathbf{D}_r]$ be the balanced ROM with $r = s_1 + s_2 + \ldots + s_l$ for some $l \leq k$. Then, we have:

$$||\mathbf{G} - \mathbf{G}_r||_{\mathcal{H}_\infty} \leqslant \sum_{j=l+1}^{k} 2\sigma_j.$$

Key "Ingredients" for balancing of nonlinear systems

Let's now consider a **quadratic nonlinear systems** again:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{F}\mathbf{x}^{②}(t) + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t).$$

The key ingredients for nonlinear balancing are:

1. Energy functions: controllability/observability; past/future; HJB energy functions ($\Rightarrow$ Part 1 & 2)
2. A nonlinear transformation $\mathbf{x} = \Phi(\mathbf{z})$ (instead of $\mathbf{x} = \mathbf{T}\mathbf{z}$) that "diagonalizes" the energy
3. Singular value (functions) $\sigma_i(z_i)$ to decide on which states to truncate (instead of constant SVs for LTI)
4. A definition of the nonlinearly balanced ROM

Input-normal/output-diagonal balancing

We have computed polynomial expansions of the past and future energy functions of the form

$$\mathcal{E}_\gamma^-(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{②} + \mathbf{v}_3^\top \mathbf{x}^{③} + \cdots \mathbf{v}_d^\top \mathbf{x}^{ⓓ}\right) \quad \text{and} \quad \mathcal{E}_\gamma^+(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{w}_2^\top \mathbf{x}^{②} + \mathbf{w}_3^\top \mathbf{x}^{③} + \ldots + \mathbf{w}_d^\top \mathbf{x}^{ⓓ}\right)$$

Theorem [Fujimoto and Scherpen, 2010, Thm. 2]

Suppose the Jacobian linearization of the nonlinear system is controllable, observable, and asymptotically stable. Then there is a neighborhood $\mathcal{W}$ of the origin and a coordinate transformation $\mathbf{x} = \Phi(\mathbf{z})$ on $\mathcal{W}$ with $\mathbf{z} = [z_1, z_2, \ldots, z_n]$ such that the energy functions have **input-normal** form:

$$\mathcal{E}_\gamma^-(\Phi(\mathbf{z})) = \frac{1}{2}\sum_{i=1}^{n} z_i^2, \qquad \mathcal{E}_\gamma^+(\Phi(\mathbf{z})) = \frac{1}{2}\sum_{i=1}^{n} \xi_i^2(z_i) z_i^2.$$

We assume that the state transformation is analytic, so

$$\mathbf{x} = \Phi(\mathbf{z}) = \mathbf{T}_1\mathbf{z} + \mathbf{T}_2\mathbf{z}^{②} + \ldots + \mathbf{T}_k\mathbf{z}^{ⓚ}$$

where $\mathbf{T}_k \in \mathbb{R}^{n \times n^k}$ are the polynomial coefficients and $\mathbf{T}_1$ is nonsingular.

Computation of tensors for transformation

Theorem (K./Gugercin/Borggaard/ '23)

Let $\mathbf{W}_2 = \mathbf{L}\mathbf{L}^\top$ and $\mathbf{V}_2 = \mathbf{R}\mathbf{R}^\top$. Compute the singular value decomposition of $\mathbf{L}^\top\mathbf{R}^{-\top} = \mathcal{U}\mathbf{\Xi}\mathcal{V}^\top$. The linear transformation $\mathbf{T}_1$ and its inverse $\mathbf{T}_1^{-1}$ are given by

$$\mathbf{T}_1 = \mathbf{R}^{-\top}\mathcal{V}, \qquad \mathbf{T}_1^{-1} = \mathbf{\Xi}^{-1}\mathcal{U}^\top\mathbf{L}^\top$$

and they satisfy $\mathbf{T}_1^{-1}\mathbf{V}_2^{-1}\mathbf{W}_2\mathbf{T}_1 = \mathbf{\Xi}^2 = \mathrm{diag}(\xi_1^2(0), \ldots, \xi_n^2(0))$. The higher-order tensors are

$$\mathbf{T}_2 = -\frac{1}{2}\mathbf{T}_1 \ \mathrm{unvec}([\mathbf{T}_1^{③}]^\top \mathbf{v}_3)^\top$$

$$\mathbf{T}_k = -\frac{1}{2}\mathbf{T}_1 \mathrm{unvec}\left(\mathbf{M}_k\right)^\top, \ \text{where} \quad \mathbf{M}_k = \sum_{\substack{i,j>1 \\ i+j=k+1}} \mathrm{vec}\left(\mathbf{T}_j^\top \mathbf{V}_2 \mathbf{T}_i\right) + \sum_{i=3}^{k+1} \mathcal{T}_{i,k+1}^\top \mathbf{v}_i$$

Here, unique tensor products with m terms and and n^l columns are denoted as

$$\mathcal{T}_{m,l} := \sum_{\sum i_j = l} \mathbf{T}_{i_1} \otimes \cdots \otimes \mathbf{T}_{i_m} \in \mathbb{R}^{n^m \times n^l}, \quad i_j \geq 1 \text{ for each } j = 1, \ldots, m,$$

Computation of singular value functions

The state-dependent singular value functions are approximated as

$$\xi_i(z_i) = \xi_i(0) + c_i^{(1)} z_i + c_i^{(2)} z_i^2 + \ldots + c_i^{(\ell)} z_i^\ell, \qquad i = 1, 2, \ldots, n.$$

Define the coefficients of the kth order terms as $\mathbf{c}_k := [c_1^{(k)},\ c_2^{(k)}, \ldots, c_n^{(k)}]^\top$, so that the vector of singular value functions of the input-normal form is

$$\xi(\mathbf{z}) = \boldsymbol{\Xi} \cdot \mathbf{1} + \operatorname{diag}(\mathbf{c}_1)\mathbf{z} + \ldots + \operatorname{diag}(\mathbf{c}_\ell)\mathbf{z}^\ell,$$

Since the transformation matrices $\mathbf{T}_1, \ldots, \mathbf{T}_k$ are already computed, we use the equation

$$\mathcal{E}_\gamma^+(\Phi(\mathbf{z})) = \frac{1}{2} \sum_{i=1}^{n} \xi_i^2(z_i) z_i^2$$

and insert the approximations to obtain

$$\mathbf{z}^\top \mathbf{T}_1^\top \mathbf{W}_2 \mathbf{T}_1 \mathbf{z} + 2\mathbf{z}^\top \mathbf{T}_1^\top \mathbf{W}_2 \Phi^h(\mathbf{z}) + \Phi^h(\mathbf{z})^\top \mathbf{W}_2 \Phi^h(\mathbf{z}) + 2\mathcal{E}_o^h(\Phi(\mathbf{z})) = \sum_{i=1}^{n} z_i^2 \left(\xi_i^2(0) + 2\xi_i(0)\xi_i^h(z_i) + \xi_i^h(z_i)^2 \right).$$

Coefficients of Singular Value Functions

Theorem (K./Gugercin/Borggaard/ '23)

Let $\mathbf{z} = [z_1, z_2, \ldots, z_n]^\top$ be the transformed state and $\mathbf{c}_k = [c_1^{(k)},\ c_2^{(k)}, \ldots, c_n^{(k)}]^\top$ be the vector of n coefficients of the kth order terms. Then,

$$\mathbf{c}_1 = \mathbf{\Xi}^{-1} \left(\text{vec}(\mathbf{T}_2^\top \mathbf{W}_2 \mathbf{T}_1)^\top + \frac{1}{2} \mathbf{w}_3^\top \mathbf{T}_1^{③} \right)_{\mathcal{I}_1}$$

for the indices $\mathcal{I}_1 = \{j \mid j = (i-1)(n^2+n) + i,\ i = 1, \ldots, n\}$. For $k \geq 1$ we obtain

$$\mathbf{c}_k = \frac{1}{2} \mathbf{\Xi}^{-1} \left[\left(\sum_{\substack{i,j \geq 1 \\ i+j=k+2}} \text{vec}\left(\mathbf{T}_j^\top \mathbf{W}_2^\top \mathbf{T}_i \right)^\top + \sum_{i=3}^{k+2} \mathbf{w}_i^\top \mathcal{T}_{i,k+2} \right)_{\mathcal{I}_k} - \sum_{i+j=k} \mathbf{c}_i \odot \mathbf{c}_j \right]$$

where $\mathcal{I}_k$ is the index set $\mathcal{I}_k = \{j \mid j = (i-1) \sum_{l=1}^{k+1} n^l + i,\ i = 1, \ldots, n\}$, and $\odot$ denotes the Hadamard product.

Comparison to the linear case

For LTI systems:

- $\mathbf{T}_i = \mathbf{0}$ for $i \geq 2$ so we recover the usual linear state transformation $\Phi(\mathbf{z}) = \mathbf{T}_1 \mathbf{z}$
- The energy functions are quadratic: $\mathcal{E}_\gamma^-(\mathbf{x}) = \frac{1}{2}\mathbf{v}_2^\top \mathbf{z}^{②}$, and hence $\mathbf{v}_i = \mathbf{0}$ for $i \geq 3$.
- The singular value functions are constant; our algorithm indeed produces $\mathbf{c}_i = \mathbf{0}$ for $i \geq 1$.
- In sum, for LTI, **the energy functions are quadratic, the transformation linear, and the singular value functions constant**.

However, this cascade of degrees does not hold for the general nonlinear case.

- Assume the energy function is exactly cubic, i.e., $\mathcal{E}_\gamma^-(\mathbf{x}) = \frac{1}{2}(\mathbf{v}_2^\top \mathbf{z}^{②} + \mathbf{v}_3^\top \mathbf{z}^{③})$.
- We can still compute $\mathbf{T}_k, k \geq 3$ as $\mathbf{T}_3 \neq \mathbf{0}$ and consequently $\mathbf{T}_k$ is nonzero.
- Similarly the $\mathbf{c}_i$ coefficients can be nonzero.

Thus the degree of the energy function has, in general, no direct impact on the degree of the transformation and singular value functions.

Fully balanced system

- Transformation $\Phi(\mathbf{z})$ brought system in input-normal/output-diagonal form
- Want input-output balanced form, where the singular values appear in both the controllability and observability energy functions.

Theorem [Fujimoto and Scherpen, 2010, Thm. 9]

Suppose that the Jacobian linearization of the nonlinear system is controllable, observable, and asymptotically stable. Then there is a neighborhood $\mathcal{W}$ of the origin and a coordinate transformation $\mathbf{x} = \bar{\Phi}(\bar{\mathbf{z}})$ on $\mathcal{W}$ converting the energy functions into the form

$$\mathcal{E}_c(\bar{\Phi}(\bar{\mathbf{z}})) = \frac{1}{2}\sum_{i=1}^{n} \frac{\bar{z}_i^2}{\sigma_i(\bar{z}_i)}, \qquad \mathcal{E}_o(\bar{\Phi}(\bar{\mathbf{z}})) = \frac{1}{2}\sum_{i=1}^{n} \sigma_i(\bar{z}_i)\bar{z}_i^2.$$

Moreover, if $\mathcal{W} = \mathbb{R}^n$, then the Hankel norm of the nonlinear system is given by

$$\|\Sigma\|_{\mathrm{H}} := \sup_{\mathbf{u}\in L_2(0,\infty), \mathbf{u}\neq\mathbf{0}} \frac{\|\mathcal{H}(\mathbf{u})\|}{\|\mathbf{u}\|} = \sup_{\bar{z}_1} \sigma_1(\bar{z}_1),$$

where $\mathcal{H}$ is the Hankel operator for the nonlinear system.

Balanced high-dimensional model

The nonlinear transformation that brings the dynamical system into a fully balanced coordinate system is

$$\mathbf{x} = \bar{\Phi}(\bar{\mathbf{z}}) = \mathbf{T}_1\mathbf{z} + \mathbf{T}_2\mathbf{z}^{\textcircled{2}} + \cdots + \mathbf{T}_k\mathbf{z}^{\textcircled{k}}$$
$$z_i = \bar{z}_i / \sqrt{\sigma_i(\bar{z}_i)}.$$

The dynamical system when transformed with the input-output balancing transformation $\mathbf{x} = \Phi(\bar{\mathbf{z}})$ (or alternatively the input-normal transform) is

$$\bar{\mathbf{J}}(\bar{\mathbf{z}})\dot{\bar{\mathbf{z}}} = \mathbf{f}(\bar{\Phi}(\bar{\mathbf{z}})) + \mathbf{g}(\bar{\Phi}(\bar{\mathbf{z}}))\mathbf{u},$$

where the Jacobian $\mathbf{J}(\bar{\mathbf{z}}) \in \mathbb{R}^{n\times n}$ of the state-space transformation is given by

$$\bar{\mathbf{J}}(\bar{\mathbf{z}}) := \frac{\mathrm{d}\bar{\Phi}(\bar{\mathbf{z}})}{\mathrm{d}\bar{\mathbf{z}}} = \mathbf{T}_1 + 2\mathbf{T}_2(\bar{\mathbf{z}} \otimes \mathbf{I}) + 3\mathbf{T}_3(\bar{\mathbf{z}} \otimes \bar{\mathbf{z}} \otimes \mathbf{I}) + \ldots$$

which can be computed explicitly without numerical approximation.

How to determine the ROM dimension?

- To determine the reduced dimension r of the ROM, we look for a significant gap in the $\mathcal{H}_\infty$ singular value functions, i.e., we look for the reduced dimension r such that

$$\max_{\bar{z}_r} \sigma_r(\bar{z}_r) \gg \max_{\bar{z}_{r+1}} \sigma_{r+1}(\bar{z}_{r+1})$$

  at a minimum we require that '$>$' holds in a neighborhood of the origin.

- This indicates that the state components $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_r$ are more important in terms of the past and future energy functions $\mathcal{E}_\gamma^-$ and $\mathcal{E}_\gamma^+$ than the states $\bar{z}_{r+1}, \bar{z}_{r+2}, \ldots, \bar{z}_n$. We therefore set

$$\bar{z}_{r+1} = \bar{z}_{r+2} = \ldots = \bar{z}_n = 0$$

  in the balanced coordinates.

- Define the reduced state vector as

$$\bar{\mathbf{z}}_r = \mathbf{\Psi}_r^\top \; \bar{\mathbf{z}} = [\bar{z}_1, \bar{z}_2, \ldots \bar{z}_r]^\top, \quad \mathbf{\Psi}_r = [\mathbf{I}_r, \; \mathbf{0}]^\top \in \mathbb{R}^{n \times r}.$$

Balanced ROM

The balance-then-reduce strategy suggested in [Scherpen, 1993, Scherpen, 1996] first computes the full balancing transformation, and then truncates the resulting fully balanced system. Applying this to the FOM yields

$$\begin{aligned}\dot{\bar{\mathbf{z}}}_r &= \underbrace{\boldsymbol{\Psi}_r^\top [\bar{\mathbf{J}}([\bar{\mathbf{z}}_r, \mathbf{0}])]^{-1} \mathbf{f}(\bar{\Phi}([\bar{\mathbf{z}}_r, \mathbf{0}]))}_{=:\mathbf{f}_r(\bar{\mathbf{z}}_r)} + \underbrace{\boldsymbol{\Psi}_r^\top [\bar{\mathbf{J}}([\bar{\mathbf{z}}_r, \mathbf{0}])]^{-1} \mathbf{g}(\bar{\Phi}([\bar{\mathbf{z}}_r, \mathbf{0}]))}_{=:\mathbf{g}_r(\bar{\mathbf{z}}_r)} \mathbf{u}, \\ \mathbf{y}_r &= \underbrace{\mathbf{h}(\bar{\Phi}([\bar{\mathbf{z}}_r, \mathbf{0}]))}_{=:\mathbf{h}_r(\bar{\mathbf{z}}_r)}.\end{aligned}$$

The high-dimensional state is reconstructed as $\mathbf{x} \approx \bar{\Phi}([\bar{\mathbf{z}}_r, \mathbf{0}])$.

Two problems with this approach:

1. Simulating the ROM is computationally expensive
2. The transformation is ill-conditioned due to the need to invert **all** Hankel singular values (in analogy to the linear case)

Simultaneous balancing and reduction

Goal: compute the truncated versions of the linear transformations and higher-order tensors $\mathbf{T}_i$ directly without computing the full-order quantities.

Proposition (K./Gugercin/Borggaard/ '23)

Consider a nonlinear dynamical system and define the embedding $\Phi_r : \mathbb{R}^r \mapsto \mathbb{R}^n$ via

$$\mathbf{x} \approx \Phi_r(\bar{\mathbf{z}}_r) := \mathbf{T}_{1,r}\bar{\mathbf{z}}_r + \mathbf{T}_{2,r}\bar{\mathbf{z}}_r^{②} + \cdots + \mathbf{T}_{k,r}\bar{\mathbf{z}}_r^{ⓚ},$$

with $\mathbf{T}_{k,r} \in \mathbb{R}^{n \times r^k}$ and where $\bar{\mathbf{z}}_r \in \mathbb{R}^r$ is the reduced state. Then, the reduced Jacobian can be computed analytically via

$$\mathbf{J}_r(\bar{\mathbf{z}}_r) := \frac{\mathrm{d}\Phi_r(\bar{\mathbf{z}}_r)}{\mathrm{d}\bar{\mathbf{z}}_r} = \mathbf{T}_{1,r} + 2\mathbf{T}_{2,r}(\bar{\mathbf{z}}_r \otimes \mathbf{I}) + 3\mathbf{T}_{3,r}(\bar{\mathbf{z}}_r \otimes \bar{\mathbf{z}}_r \otimes \mathbf{I}) + \cdots \in \mathbb{R}^{n \times r}.$$

so that the nonlinear ROM with $\mathbf{z}_r \in \mathbb{R}^r$ is

$$\dot{\bar{\mathbf{z}}}_r = \underbrace{\mathbf{J}_r(\bar{\mathbf{z}}_r)^\dagger \mathbf{f}(\Phi_r(\bar{\mathbf{z}}_r))}_{=:\mathbf{f}_r(\bar{\mathbf{z}}_r)} + \underbrace{\mathbf{J}_r(\bar{\mathbf{z}}_r)^\dagger \mathbf{g}(\Phi_r(\bar{\mathbf{z}}_r))}_{=:\mathbf{g}_r(\bar{\mathbf{z}}_r)}\mathbf{u} \qquad \mathbf{y}_r = \underbrace{\mathbf{h}(\Phi_r(\bar{\mathbf{z}}_r))}_{=:\mathbf{h}_r(\bar{\mathbf{z}}_r)}.$$

How to compute the reduced coefficient matrices $\mathbf{T}_{i,r}$?

Truncated (approximate) balanced transformation (K./Gugercin/Borggaard/ '22)

Let $\mathbf{v}_i, \mathbf{w}_i$ be the polynomial coefficients for the energy functions. Let $\mathbf{R}, \mathbf{L}$ be their Cholesky factors, i.e., $\mathbf{V}_2 = \mathbf{R}\mathbf{R}^\top$ and $\mathbf{W}_2 = \mathbf{L}\mathbf{L}^\top$. Let $\mathbf{L}^\top\mathbf{R}^{-\top} = \mathcal{U}\Xi\mathcal{V}^\top$ be the SVD and define

$$\mathcal{U}_r = \mathcal{U}(:, 1:r), \quad \Xi_r = \Xi(1:r, 1:r), \quad \mathcal{V}_r = \mathcal{V}(:, 1:r).$$

Then, the coefficient matrices of the nonlinear embedding $\Phi_r : \mathbb{R}^r \mapsto \mathbb{R}^n$ are

$$\begin{aligned}
\mathbf{T}_{1,r} &= \mathbf{R}^{-\top}\mathcal{V}_r \in \mathbb{R}^{n\times r},\\
\mathbf{T}_{1,r}^\dagger &= \Xi_r^{-1}\mathcal{U}_r^\top\mathbf{L}^\top \in \mathbb{R}^{r\times n}, \text{ (left inverse)}\\
\mathbf{T}_{2,r} &= -\frac{1}{2}\mathbf{T}_{1,r}\ \text{unvec}\left([\mathbf{T}_{1,r}^{③}]^\top\mathbf{v}_3\right)^\top \in \mathbb{R}^{n\times r^2},\\
\mathbf{T}_{k,r} &= -\frac{1}{2}\mathbf{T}_{1,r}\text{unvec}\left(\sum_{\substack{i,j>1\\ i+j=k+1}} \text{vec}\left(\mathbf{T}_{j,r}^\top\mathbf{V}_2\mathbf{T}_{i,r}\right) + \sum_{i=3}^{k+1}\mathcal{T}_{i,k+1}^\top\mathbf{v}_i\right)^\top \in \mathbb{R}^{n\times r^k}.
\end{aligned}$$

Complete balancing algorithm

Algorithm 2 Computation of nonlinear input-output $\mathcal{H}_\infty$-balanced ROM.

Input: Constant $\gamma > \gamma_0 \geq 0$, $\gamma \neq 1$; polynomial degrees $d > k > \ell$; reduced model order r

Output: Input-output nonlinear $\mathcal{H}_\infty$-balanced ROM

1: Obtain a polynomial representation (or approximation) of the past and future energy functions $\mathcal{E}_\gamma^-(\mathbf{x})$ and $\mathcal{E}_\gamma^+(\mathbf{x})$, i.e., coefficients $\{\mathbf{v}_i\}_{i=2}^d$ and $\{\mathbf{w}_i\}_{i=2}^d$.

2: Compute the truncated polynomial coefficient matrices $\{\mathbf{T}_{i,r}\}_{i=1}^k$ for $\mathbf{x} \approx \Phi_r(\bar{\mathbf{z}}_r)$ from Algorithm 1.

3: Symmetrize the coefficients $\{\mathbf{T}_{i,r}\}_{i=1}^r$

4: Assemble the nonlinear ROM functions $\mathbf{f}_r(\bar{\mathbf{z}}_r), \mathbf{g}_r(\bar{\mathbf{z}}_r), \mathbf{h}_r(\bar{\mathbf{z}}_r)$ with the explicit Jacobian.

Nonlinear Manifold ROM approximation

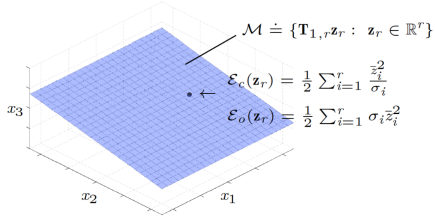
The described nonlinear balanced truncation approach is in essence a model reduction approach on the r-dimensional polynomially nonlinear manifold

$$\mathcal{M} = \left\{ \Phi(\bar{\mathbf{z}}_r) = \mathbf{T}_{1,r}\bar{\mathbf{z}}_r + \mathbf{T}_{2,r}\bar{\mathbf{z}}_r^{②} + \cdots + \mathbf{T}_{k,r}\bar{\mathbf{z}}_r^{ⓚ} \right\}.$$

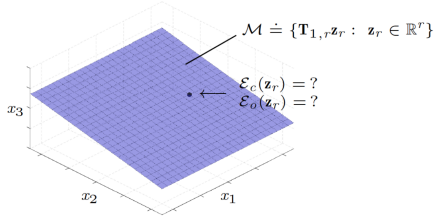
Recent work in NL-ROM on manifolds:

- Autoencoder ROM (fully nonlinear) in [Lee and Carlberg, 2020]
- Quadratic manifolds: use $\mathbf{x} \approx \mathbf{V}\mathbf{z}_r + \bar{\mathbf{V}}(\mathbf{z}_r \otimes \mathbf{z}_r)$ for intrusive [Jain et al., 2017, Barnett and Farhat, 2022] and nonintrusive [Geelen et al., 2022] ROMS.
- Reduced manifold ROM via autoencoder and propagation via feed-forward network, which approximates the ROM, [Fresca et al., 2021]
- Symplectic manifolds for Hamiltonian sysems: [Buchfink et al., 2021]
- Survey of methods to break Kolmogorov n-width problem [Peherstorfer, 2022]
- ...

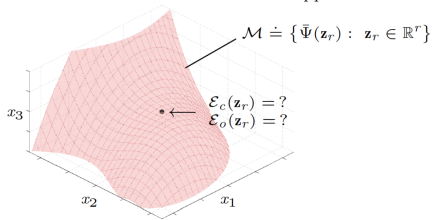
Linear subspace approximation (LTI)



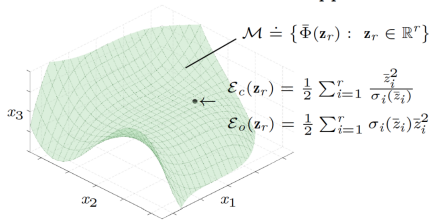
Linear subspace approximation (nonlinear model)



General nonlinear manifold approximation



Nonlinear balanced manifold approximation



Numerical Results: Burgers' equation

We consider the one-dimensional Burgers' equation

$$z_t(x,t) = \epsilon z_{xx}(x,t) - \frac{1}{2}\left(z^2(x,t)\right)_x + \sum_{j=1}^{m} b_j^m(x)u_j(t),$$

$$y_i(t) = \int_{\chi_{[(i-1)/p,\, i/p]}} z(x,t)\mathrm{d}x, \qquad i = 1,\ldots,p,$$

- periodic BCs $z(0,t) = z(1,t)$ and $z_x(0,t) = z_x(1,t)$
- IC: $z(\cdot,0) = z_0(\cdot) \in H_0^1(0,1)$
- $\epsilon = 0.001$ to make the nonlinearity significant.
- $p = 4$ outputs: spatial averages
- $m = 4$ controls/inputs with $b_j^m(x) = \chi_{[(j-1)/m,\, j/m]}(x)$.

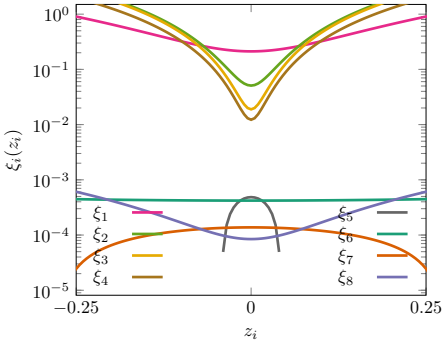
The discretized system has the form

$$\widetilde{\mathbf{E}}\dot{\mathbf{z}} = \widetilde{\mathbf{A}}\mathbf{z} + \widetilde{\mathbf{N}}_2\left(\mathbf{z}\otimes\mathbf{z}\right) + \widetilde{\mathbf{B}}\mathbf{u}$$
$$\mathbf{y} = \widetilde{\mathbf{C}}\mathbf{z},$$

A change of variables $\mathbf{x} = \widetilde{\mathbf{E}}^{1/2}\mathbf{z}$ and redefining $\mathbf{A} = \mathbf{S}^{-1}\widetilde{\mathbf{A}}\mathbf{S}^{-1}$, $\mathbf{B} = \mathbf{S}^{-1}\widetilde{\mathbf{B}}$, $\mathbf{C} = \widetilde{\mathbf{C}}\mathbf{S}^{-1}$, $\widetilde{\mathbf{N}}_2 = \mathbf{N}_2(\mathbf{S}^{-1}\otimes\mathbf{S}^{-1})$ leads to a system with $\mathbf{E} = \mathbf{I}$.

Numerical Results

Singular value functions



- $n = 16$ for FOM model
- Quartic energy functions
- Cubic transformation tensors $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$
- Quadratic singular value functions

Relative output error

$$e(r) = \frac{\left(\int_0^{10} |y(t) - y_r(t)|^2 \ \mathrm{d}t\right)^{1/2}}{\left(\int_0^{10} |y(t)|^2 \ \mathrm{d}t\right)^{1/2}}$$

r	$k = 1$	$k = 3$	$k = 5$
1	0.0714831	0.0714814	0.0713882
2	0.0036861	0.0036778	0.0031076
3	0.0026888	0.0026784	0.0026665
4	0.0024333	0.0024288	0.0024238
5	0.0024095	0.0024032	0.0023853

- Errors decay monotonely w.r.t r and k.
- Linear model transformation, however, already very good in this example.

Review and conclusion

We suggested several new computational and modeling choices for balanced nonlinear ROMs:

1. Scalable computation ($n = 1,024$) of a family ($\mathcal{H}_\infty$, HJB, open-loop) energy function approximations $\mathcal{E}_\gamma^-(\mathbf{x}), \mathcal{E}_\gamma^+(\mathbf{x})$
2. Scalable computation of singular value functions $\sigma_i(z_i)$
3. Nonlinear simultaneous balance-and-reduce state transformation $\mathbf{x} \approx \Phi_r(\bar{\mathbf{z}}_r)$
4. Projection of nonlinear model with nonlinear basis (speed up still needed)
5. Two semi-discretized PDE examples (first time use for PDEs)

Outlook and ongoing work:

1. Nonlinear ROMs still have to be made efficient ((D)EIM, other approximations)
2. Addition of polynomial drift/input/output terms in dynamical system (See Linus Balicki's and Nick Corbin's talk)
3. Other approximation techniques: Sum-of-squares (Hamza Adjerid's talk)
4. More efficient solvers: Low-rank, iterative, ...
5. Controllers based on these energy functions
6. Structured systems (DAEs, (port-) Hamiltonians, etc.

[Barnett and Farhat, 2022] Barnett, J. and Farhat, C. (2022).
Quadratic approximation manifold for mitigating the kolmogorov barrier in nonlinear projection-based model order reduction.
Journal of Computational Physics, page 111348.

[Benner and Goyal, 2017] Benner, P. and Goyal, P. (2017).
Balanced truncation model order reduction for quadratic-bilinear control systems.
arXiv:1705.00160.

[Benner and Saak, 2013] Benner, P. and Saak, J. (2013).
Numerical solution of large and sparse continuous time algebraic matrix Riccati and Lyapunov equations: a state of the art survey.
GAMM-Mitteilungen, 36(1):32–52.

[Borggaard and Zietsman, 2021] Borggaard, J. and Zietsman, L. (2021).
On approximating polynomial-quadratic regulator problems.
IFAC PaersOnLine, 54(9):329–334.

[Bouvrie and Hamzi, 2017] Bouvrie, J. and Hamzi, B. (2017).
Kernel methods for the approximation of nonlinear systems.
SIAM Journal on Control and Optimization, 55(4):2460–2492.

[Breiten et al., 2018] Breiten, T., Kunisch, K., and Pfeiffer, L. (2018).
Numerical study of polynomial feedback laws for a bilinear control problem.
Mathematical Control & Related Fields, 8(3&4):557.

[Buchfink et al., 2021] Buchfink, P., Glas, S., and Haasdonk, B. (2021).
Symplectic model reduction of hamiltonian systems on nonlinear manifolds.
arXiv preprint arXiv:2112.10815.

[Fresca et al., 2021] Fresca, S., Dede, L., and Manzoni, A. (2021).
A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized pdes.
Journal of Scientific Computing, 87(2):1–36.

[Fujimoto and Scherpen, 2010] Fujimoto, K. and Scherpen, J. M. (2010).
Balanced realization and model order reduction for nonlinear systems based on singular value analysis.
SIAM Journal on Control and Optimization, 48(7):4591–4623.

[Fujimoto and Tsubakino, 2008] Fujimoto, K. and Tsubakino, D. (2008).
Computation of nonlinear balanced realization and model reduction based on Taylor series expansion.
Systems & Control Letters, 57(4):283–289.

[Geelen et al., 2022] Geelen, R., Wright, S., and Willcox, K. (2022).
Operator inference for non-intrusive model reduction with nonlinear manifolds.
arXiv preprint arXiv:2205.02304.

[Gray and Verriest, 2006] Gray, W. S. and Verriest, E. I. (2006).
Algebraically defined Gramians for nonlinear systems.
In *Decision and Control, 2006 45th IEEE Conference on*, pages 3730–3735. IEEE.

[Horn et al., 1994] Horn, R. A., Horn, R. A., and Johnson, C. R. (1994).
Topics in matrix analysis.
Cambridge university press.

[Jain et al., 2017] Jain, S., Tiso, P., Rutzmoser, J. B., and Rixen, D. J. (2017).
A quadratic manifold for model order reduction of nonlinear structural dynamics.
Computers & Structures, 188:80–94.

[Jonckheere and Silverman, 1983] Jonckheere, E. and Silverman, L. (1983).
A new set of invariants for linear systems–application to reduced order compensator design.
IEEE Transactions on Automatic Control, 28(10):953–964.

[Kawano and Scherpen, 2016] Kawano, Y. and Scherpen, J. M. (2016).
Model reduction by differential balancing based on nonlinear hankel operators.
IEEE Transactions on Automatic Control, 62(7):3293–3308.

[Kramer and Willcox, 2019] Kramer, B. and Willcox, K. (2019).
Balanced truncation model reduction for lifted nonlinear systems.
arXiv Preprint arXiv:1907.12084.

[Krener, 2008] Krener, A. J. (2008).
Reduced order modeling of nonlinear control systems.
In *Analysis and Design of Nonlinear Control Systems*, pages 41–62. Springer.

[Kressner and Tobler, 2010] Kressner, D. and Tobler, C. (2010).
Krylov subspace methods for linear systems with tensor product structure.
SIAM journal on matrix analysis and applications, 31(4):1688–1714.

[Lall et al., 2002] Lall, S., Marsden, J. E., and Glavaski, S. (2002).
A subspace approach to balanced truncation for model reduction of nonlinear control systems.
International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal, 12(6):519–535.

[Lee and Carlberg, 2020] Lee, K. and Carlberg, K. T. (2020).
Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders.
Journal of Computational Physics, 404:108973.

[Lukes, 1969] Lukes, D. L. (1969).
Optimal regulation of nonlinear dynamical systems.
SIAM Journal on Control, 7(1):75–100.

[Moore, 1981] Moore, B. (1981).
Principal component analysis in linear systems: Controllability, observability, and model reduction.
IEEE Transactions on Automatic Control, 26(1):17–32.

[Newman and Krishnaprasad, 2000] Newman, A. J. and Krishnaprasad, P. S. (2000).
Computing balanced realizations for nonlinear systems.
Technical report, Center for Dynamics and Control of Smart Structures.

[Peherstorfer, 2022] Peherstorfer, B. (2022).
Breaking the Kolmogorov barrier with nonlinear model reduction.
Notices of the American Mathematical Society, 69(5).

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

[Scherpen, 1996] Scherpen, J. M. (1996).
$\mathcal{H}_\infty$ balancing for nonlinear systems.
International Journal of Robust and Nonlinear Control, 6(7):645–668.

[Scherpen and Van der Schaft, 1994] Scherpen, J. M. A. and Van der Schaft, A. (1994).
Normalized coprime factorizations and balancing for unstable nonlinear systems.
International Journal of Control, 60(6):1193–1222.

[Verriest, 1981] Verriest, E. I. (1981).
Suboptimal lqg-design via balanced realizations.
In *1981 20th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pages 686–687. IEEE.