Nonlinear Balancing for Quadratic-Polynomial Systems

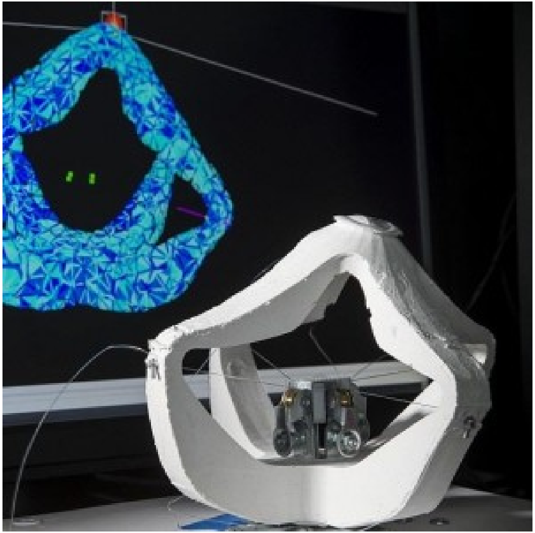
Nicholas A. Corbin

May 24th, 2023, Blacksburg, VA

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.

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Mechanical and Aerospace Engineering

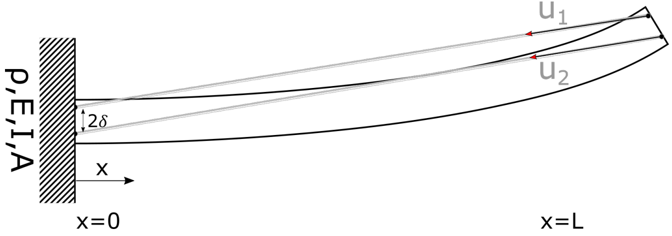
Motivating application: soft robotics

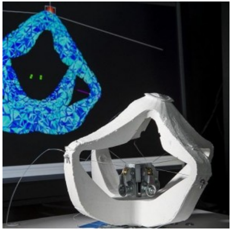


[Zhang et al., 2016]

"Even if different results exist in the literature for linear model reduction, like balanced truncation or iterative tangential interpolation, the only method suitable for non-linear system is the Proper Orthogonal Decomposition (POD)." - [Thieffry et al., 2018]

Challenges with soft robot control





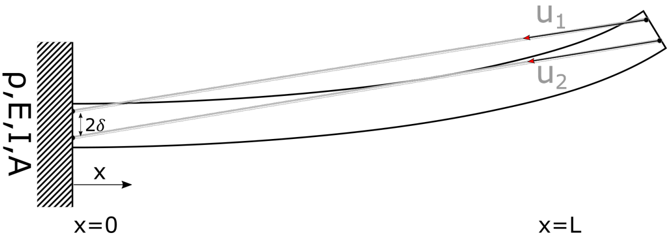
Control-affine dynamical system

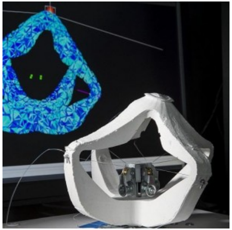
Material and geometric nonlinearities

Large state dimension due to PDE discretization

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

Challenges with soft robot control





Control-affine dynamical system

Material and geometric nonlinearities

Large state dimension due to PDE discretization

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

Geometric nonlinearities, e.g. cable angles

Background on nonlinear balancing and computation

$\mathcal{H}_\infty$ energy function definitions

Definition ([Scherpen, 1996])

Let γ be a positive constant $\gamma > 0, \gamma \neq 1$, and define $\eta := 1 - \gamma^{-2}$. The $\mathcal{H}_\infty$ past energy of the nonlinear system is defined as

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

If $\gamma < 1$, the $\mathcal{H}_\infty$ future energy of the nonlinear system is defined as

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

whereas if $\gamma > 1$, the $\mathcal{H}_\infty$ future energy is defined as

$$\mathcal{E}_\gamma^+(\mathbf{x}_0) := \min_{\substack{\mathbf{u} \in L_2[0, \infty) \\ \mathbf{x}(0) = \mathbf{x}_0,\, \mathbf{x}(\infty) = \mathbf{0}}} \frac{1}{2} \int_{0}^{\infty} \|\mathbf{y}(t)\|^2 + \frac{\|\mathbf{u}(t)\|^2}{\eta} \mathrm{d}t. \tag{3}$$

Energy functions solve Hamilton-Jacobi-Bellman PDEs

Theorem ([Scherpen, 1996])

Assume that the HJB equation

$$0 = \frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) + \frac{1}{2} \frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} \mathbf{g}(\mathbf{x}) \mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} - \frac{\eta}{2} \mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x}) \tag{4}$$

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

$$0 = \frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}) - \frac{\eta}{2} \frac{\partial \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}} \mathbf{g}(\mathbf{x}) \mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^+(\mathbf{x})}{\partial \mathbf{x}} + \frac{1}{2} \mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x}) \tag{5}$$

has a solution with $\mathcal{E}_\gamma^+(\mathbf{0}) = 0$ *such that* $\mathbf{f}(\mathbf{x}) - \eta \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \partial^\top \mathcal{E}_\gamma^+(\mathbf{x})/\partial \mathbf{x}$ *is asymptotically stable. Then this solution is the future energy function* $\mathcal{E}_\gamma^+(\mathbf{x})$.

Analytic dynamics → analytic energy function

If the dynamics are analytic (polynomial):

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} \quad \rightarrow \quad \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \sum_{\xi=2}^{\ell} \mathbf{F}_\xi \mathbf{x}^{\textcircled{\xi}} + \sum_{\xi=1}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{\textcircled{\xi}} \otimes \mathbf{u}\right) + \mathbf{B}\mathbf{u}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) \quad \rightarrow \quad \mathbf{y} = \mathbf{C}\mathbf{x} + \sum_{\xi=2}^{\ell} \mathbf{H}_\xi \mathbf{x}^{\textcircled{\xi}}$$

then the energy functions are also analytic (polynomial) [Al'brekht, 1961, Lukes, 1969]:

$$\mathcal{E}_\gamma^-(\mathbf{x}) = ??? \quad \rightarrow \quad \mathcal{E}_\gamma^-(\mathbf{x}) = \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{\textcircled{2}} + \mathbf{v}_3^\top \mathbf{x}^{\textcircled{3}} + \dots\right)$$

$$\mathcal{E}_\gamma^+(\mathbf{x}) = ??? \quad \rightarrow \quad \mathcal{E}_\gamma^+(\mathbf{x}) = \frac{1}{2}\left(\mathbf{w}_2^\top \mathbf{x}^{\textcircled{2}} + \mathbf{w}_3^\top \mathbf{x}^{\textcircled{3}} + \dots\right)$$

HJB PDEs become a set of algebraic equations for the coefficients $\mathbf{v}_i, \mathbf{w}_i$.

Some notes on the Kronecker product[1]

The Kronecker product of $\mathbf{A} \in \mathbb{R}^{p\times q}$ and $\mathbf{B} \in \mathbb{R}^{s\times t}$ is the $ps \times qt$ block matrix:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1q}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{p1}\mathbf{B} & \cdots & a_{pq}\mathbf{B} \end{bmatrix}$$

where a_{ij} denotes the (i,j)th entry of $\mathbf{A}$. We write repeated Kronecker products as

$$\mathbf{x}^{\textcircled{k}} := \underbrace{\mathbf{x} \otimes \cdots \otimes \mathbf{x}}_{k \text{ times}} \in \mathbb{R}^{n^k}.$$

We adopt the following notation to define the *k-way Lyapunov matrix*

$$\mathcal{L}_k(\mathbf{A}) := \underbrace{\mathbf{A} \otimes \cdots \otimes \mathbf{I}}_{k \text{ times}} + \cdots + \underbrace{\mathbf{I} \otimes \cdots \otimes \mathbf{A}}_{k \text{ times}}$$

ID. 1. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{D} \otimes \mathbf{G}) = \mathbf{AD} \otimes \mathbf{BG}$

ID. 2. $\mathbf{A} \otimes \mathbf{B} = \mathbf{S}_{s\times p}(\mathbf{B} \otimes \mathbf{A})\mathbf{S}_{q\times t}$

ID. 3. $(\mathbf{I}_p \otimes \mathbf{x})\mathbf{A} = (\mathbf{A} \otimes \mathbf{x})$

ID. 4. $(\mathbf{x} \otimes \mathbf{I}_p)\mathbf{A} = (\mathbf{x} \otimes \mathbf{A})$

ID. 5. $\mathrm{vec}(\mathbf{ADB}) = (\mathbf{B}^\top \otimes \mathbf{A})\mathrm{vec}(\mathbf{D})$

ID. 6.
$$\begin{aligned} \mathrm{vec}(\mathbf{AD}) &= (\mathbf{I}_s \otimes \mathbf{A})\mathrm{vec}(\mathbf{D}) \\ &= (\mathbf{D}^\top \otimes \mathbf{I}_p)\mathrm{vec}(\mathbf{A}) \\ &= (\mathbf{D}^\top \otimes \mathbf{A})\mathrm{vec}(\mathbf{I}_q) \end{aligned}$$

ID. 7. $\mathbf{u}^\top \mathbf{Z}\mathbf{x} = \mathrm{vec}(\mathbf{Z})^\top(\mathbf{x} \otimes \mathbf{u})$

ID. 8. $\mathrm{vec}\left(\mathbf{x}^\top \otimes \mathbf{I}_m\right) = (\mathbf{x} \otimes \mathrm{vec}(\mathbf{I}_m))$

ID. 9. $\mathrm{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_q \otimes \mathbf{S}_{p\times t} \otimes \mathbf{I}_s)\left(\mathrm{vec}\left(\mathbf{A}\right) \otimes \mathrm{vec}\left(\mathbf{B}\right)\right)$

[1] IDs from various sources, including [Brewer, 1978, Magnus and Neudecker, 2019], etc.

Nonlinear balancing for quadratic systems [Kramer et al., 2022]

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{②} + \mathbf{B}\mathbf{u},\qquad \mathcal{E}_\gamma^-(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{②} + \mathbf{v}_3^\top \mathbf{x}^{③} + \cdots + \mathbf{v}_d^\top \mathbf{x}^{ⓓ}\right)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x},$$

Theorem

$\mathbf{v}_2 = \mathrm{vec}\left(\mathbf{V}_2\right)$ *is the symmetric positive definite solution to the $\mathcal{H}_\infty$ Riccati equation*

$$\mathbf{0} = \mathbf{A}^\top \mathbf{V}_2 + \mathbf{V}_2\mathbf{A} - \eta \mathbf{C}^\top \mathbf{C} + \mathbf{V}_2\mathbf{B}\mathbf{B}^\top \mathbf{V}_2. \tag{6}$$

For $3 \leq k \leq d$, let $\tilde{\mathbf{v}}_k \in \mathbb{R}^{n^k}$ solve the linear system

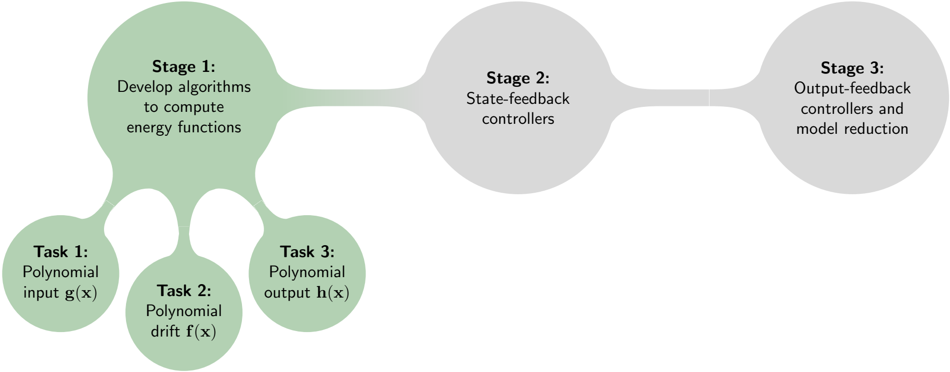
$$\mathcal{L}_k(\mathbf{A}^\top + \mathbf{V}_2\mathbf{B}\mathbf{B}^\top)\tilde{\mathbf{v}}_k = -\mathcal{L}_{k-1}(\mathbf{F})^\top \mathbf{v}_{k-1} - \frac{1}{4}\sum_{\substack{i,j>2 \\ i+j=k+2}} ij\ \mathrm{vec}(\mathbf{V}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{V}_j).$$

Then the coefficient vector $\mathbf{v}_k = \mathrm{vec}\left(\mathbf{V}_k\right)$ is obtained by the symmetrization of $\tilde{\mathbf{v}}_k$.

Roadmap

Control-affine dynamical system

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



Extension to polynomial inputs

Quadratic-Polynomial System

Consider the quadratic-polynomial system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{\circled{2}} + \sum_{\xi=1}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{\circled{\xi}} \otimes \mathbf{u}\right) + \mathbf{B}\mathbf{u}, \tag{7}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}. \tag{8}$$

Now we wish to solve the HJB PDE

$$0 = \frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\top \frac{\partial^\top \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} - \frac{\eta}{2}\mathbf{h}(\mathbf{x})^\top \mathbf{h}(\mathbf{x}) \tag{9}$$

with the quadratic-polynomial dynamics given by

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{\circled{2}}, \qquad \mathbf{g}(\mathbf{x}) = \sum_{\xi=1}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{\circled{\xi}} \otimes \mathbf{I}_m\right) + \mathbf{B}, \qquad \mathbf{h}(\mathbf{x}) = \mathbf{C}\mathbf{x}. \tag{10}$$

What do the additional $\mathbf{G}_\xi$ terms change?

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{②} + \sum_{\xi=1}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{ⓧ} \otimes \mathbf{u}\right) + \mathbf{B}\mathbf{u}, \qquad \mathcal{E}_\gamma^-(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{②} + \mathbf{v}_3^\top \mathbf{x}^{③} + \cdots + \mathbf{v}_d^\top \mathbf{x}^{ⓓ}\right)$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

Lemma

The additional $\mathbf{G}_\xi$ terms only add terms to the right-hand side of the linear systems for $\mathbf{v}_k$ for $3 \leq k \leq d$.

So $\mathbf{v}_2 = \mathrm{vec}\left(\mathbf{V}_2\right)$ still solves 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 $3 \leq k \leq d$, $\tilde{\mathbf{v}}_k \in \mathbb{R}^{n^k}$ solves the linear system

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

Main result: additional terms in the linear system for $\mathbf{v}_k$

Theorem

For $3 \leq k \leq d$, let $\tilde{\mathbf{v}}_k \in \mathbb{R}^{n^k}$ solve the linear system

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

$$-\frac{1}{4}\sum_{o=1}^{2\ell}\left(\sum_{\substack{p,q\geq 0 \\ p+q=o}}\left(\sum_{\substack{i,j\geq 2 \\ i+j=k-o+2}} ij \ \mathrm{vec}\left[\left(\mathbf{I}_{n^p} \otimes \mathrm{vec}\left[\mathbf{I}_m\right]^\top\right)\left(\mathrm{vec}\left[\mathbf{G}_q^\top \mathbf{V}_j\right]^\top \otimes \left(\mathbf{G}_p^\top \mathbf{V}_i \otimes \mathbf{I}_m\right)\right) \times \left(\mathbf{I}_{n^{j-1}} \otimes \mathbf{S}_{n^{i-1}\times n^q m} \otimes \mathbf{I}_m\right)\left(\mathbf{I}_{n^{k-p}} \otimes \mathrm{vec}\left[\mathbf{I}_m\right]\right)\right]\right)\right) \tag{11}$$

Then the coefficient vector $\mathbf{v}_k = \mathrm{vec}\left(\mathbf{V}_k\right)$ *is obtained by the symmetrization of* $\tilde{\mathbf{v}}_k$.

Main result proof sketch

The gradient of the energy function $\mathcal{E}_\gamma^-(\mathbf{x}) \approx \frac{1}{2}\left(\mathbf{v}_2^\top \mathbf{x}^{\textcircled{2}} + \mathbf{v}_3^\top \mathbf{x}^{\textcircled{3}} + \cdots + \mathbf{v}_d^\top \mathbf{x}^{\textcircled{d}}\right)$ is

$$\begin{aligned}\frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} = \frac{1}{2}\Big(&\mathbf{v}_2^\top(\mathbf{I}_n \otimes \mathbf{x}) + \mathbf{v}_2^\top(\mathbf{x} \otimes \mathbf{I}_n) \\ &+ \mathbf{v}_3^\top(\mathbf{I}_n \otimes \mathbf{x} \otimes \mathbf{x}) + \mathbf{v}_3^\top(\mathbf{x} \otimes \mathbf{I}_n \otimes \mathbf{x}) + \mathbf{v}_3^\top(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{I}_n) \\ &+ \mathbf{v}_4^\top(\mathbf{I}_n \otimes \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) + \mathbf{v}_4^\top(\mathbf{x} \otimes \mathbf{I}_n \otimes \mathbf{x} \otimes \mathbf{x}) \\ &+ \mathbf{v}_4^\top(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{I}_n \otimes \mathbf{x}) + \mathbf{v}_4^\top(\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{I}_n) + \cdots\Big). \end{aligned} \tag{12}$$

According to the Lemma, the terms we consider only appear on the right-hand side, so we may assume symmetry of the already computed $\mathbf{v}_i$ and write

$$\begin{aligned}\frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} &= \frac{1}{2}\left(2\mathbf{v}_2^\top(\mathbf{I}_n \otimes \mathbf{x}) + 3\mathbf{v}_3^\top(\mathbf{I}_n \otimes \mathbf{x} \otimes \mathbf{x}) + 4\mathbf{v}_4^\top(\mathbf{I}_n \otimes \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) + \cdots\right) \\ &= \frac{1}{2}\sum_{i=2}^{k-1} i\mathbf{v}_i^\top(\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{i-1}}). \end{aligned} \tag{13}$$

Main result proof sketch

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

Gradient of the energy function: $$\frac{\partial \mathcal{E}_\gamma^-(\mathbf{x})}{\partial \mathbf{x}} = \frac{1}{2} \sum_{i=2}^{k-1} i \mathbf{v}_i^\top (\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{\scriptsize i-1}}).$$

Input vector field: $$\mathbf{g}(\mathbf{x}) = \sum_{\xi=1}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{\textcircled{\scriptsize \xi}} \otimes \mathbf{I}_m\right) + \mathbf{B} := \sum_{\xi=0}^{\ell} \mathbf{G}_\xi \left(\mathbf{x}^{\textcircled{\scriptsize \xi}} \otimes \mathbf{I}_m\right).$$

Then we can write an arbitrary kth-order HJB term containing $\mathbf{G}_\xi$ generally as

$$\frac{1}{8} i \mathbf{v}_i^\top (\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{\scriptsize i-1}}) \mathbf{G}_p (\mathbf{x}^{\textcircled{\scriptsize p}} \otimes \mathbf{I}_m)(\mathbf{x}^{\textcircled{\scriptsize q}\,\top} \otimes \mathbf{I}_m) \mathbf{G}_q^\top (\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{\scriptsize j-1}\,\top}) \mathbf{v}_j j \tag{14}$$

$$\text{with} \quad \begin{cases} p \in [0, o], & o \in [1, 2\ell], \\ q = o - p, & i + j + o = k + 2. \end{cases} \tag{15}$$

Algebraic manipulations using Kronecker product identities

$$\frac{1}{8} i \mathbf{v}_i^\top (\mathbf{I}_n \otimes \mathbf{x}^{(i-1)}) \mathbf{G}_p (\mathbf{x}^{(p)} \otimes \mathbf{I}_m)(\mathbf{x}^{(q)^\top} \otimes \mathbf{I}_m) \mathbf{G}_q^\top (\mathbf{I}_n \otimes \mathbf{x}^{(j-1)^\top}) \mathbf{v}_j j,$$

Table: Dimensions of matrices used in identities

$\mathbf{A}(p \times q)$	$\mathbf{D}(q \times s)$	$\mathbf{u}(s \times 1)$
$\mathbf{B}(s \times t)$	$\mathbf{G}(t \times u)$	$\mathbf{w}(t \times 1)$

ID. 1. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{D} \otimes \mathbf{G}) = \mathbf{AD} \otimes \mathbf{BG}$

ID. 2. $\mathbf{A} \otimes \mathbf{B} = \mathbf{S}_{s \times p}(\mathbf{B} \otimes \mathbf{A})\mathbf{S}_{q \times t}$

ID. 3. $(\mathbf{I}_p \otimes \mathbf{w})\mathbf{A} = (\mathbf{A} \otimes \mathbf{w})$

ID. 4. $(\mathbf{w} \otimes \mathbf{I}_p)\mathbf{A} = (\mathbf{w} \otimes \mathbf{A})$

ID. 5. $\mathrm{vec}(\mathbf{ADB}) = (\mathbf{B}^\top \otimes \mathbf{A})\mathrm{vec}(\mathbf{D})$

ID. 6. $\mathbf{u}^\top \mathbf{B} \mathbf{w} = \mathrm{vec}(\mathbf{B})^\top (\mathbf{w} \otimes \mathbf{u})$

ID. 7. $\mathrm{vec}\left(\mathbf{w}^\top \otimes \mathbf{I}_m\right) = (\mathbf{w} \otimes \mathrm{vec}(\mathbf{I}_m))$

ID. 8. $\mathrm{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_q \otimes \mathbf{S}_{p \times t} \otimes \mathbf{I}_s)\left(\mathrm{vec}\left(\mathbf{A}\right) \otimes \mathrm{vec}\left(\mathbf{B}\right)\right)$

Algebraic manipulations using Kronecker product identities

$$\frac{1}{8}i \; \underbrace{\mathbf{v}_i^\top(\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{\scriptsize i-1}})}_{=\mathbf{x}^{\textcircled{\scriptsize i-1}^\top}\mathbf{V}_i^\top \text{ by ID. 5}} \; \mathbf{G}_p(\mathbf{x}^{\textcircled{\scriptsize p}} \otimes \mathbf{I}_m)(\mathbf{x}^{\textcircled{\scriptsize q}^\top} \otimes \mathbf{I}_m)\mathbf{G}_q^\top \underbrace{(\mathbf{I}_n \otimes \mathbf{x}^{\textcircled{\scriptsize j-1}^\top})\mathbf{v}_j}_{=\mathbf{V}_j\mathbf{x}^{\textcircled{\scriptsize j-1}} \text{ by ID. 5}} j,$$

Table: Dimensions of matrices used in identities

$\mathbf{A}(p \times q)$	$\mathbf{D}(q \times s)$	$\mathbf{u}(s \times 1)$
$\mathbf{B}(s \times t)$	$\mathbf{G}(t \times u)$	$\mathbf{w}(t \times 1)$

ID. 1. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{D} \otimes \mathbf{G}) = \mathbf{AD} \otimes \mathbf{BG}$

ID. 2. $\mathbf{A} \otimes \mathbf{B} = \mathbf{S}_{s\times p}(\mathbf{B} \otimes \mathbf{A})\mathbf{S}_{q\times t}$

ID. 3. $(\mathbf{I}_p \otimes \mathbf{w})\mathbf{A} = (\mathbf{A} \otimes \mathbf{w})$

ID. 4. $(\mathbf{w} \otimes \mathbf{I}_p)\mathbf{A} = (\mathbf{w} \otimes \mathbf{A})$

ID. 5. $\text{vec}(\mathbf{ADB}) = (\mathbf{B}^\top \otimes \mathbf{A})\text{vec}(\mathbf{D})$

ID. 6. $\mathbf{u}^\top\mathbf{Bw} = \text{vec}(\mathbf{B})^\top(\mathbf{w} \otimes \mathbf{u})$

ID. 7. $\text{vec}\left(\mathbf{w}^\top \otimes \mathbf{I}_m\right) = (\mathbf{w} \otimes \text{vec}(\mathbf{I}_m))$

ID. 8. $\text{vec}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{I}_q \otimes \mathbf{S}_{p\times t} \otimes \mathbf{I}_s)\left(\text{vec}\left(\mathbf{A}\right) \otimes \text{vec}\left(\mathbf{B}\right)\right)$

Algebraic manipulations using Kronecker product identities

$$\dots \frac{1}{8} ij \ \mathrm{vec}\left[\left(\mathbf{I}_{n^p} \otimes \mathrm{vec}\left[\mathbf{I}_m\right]^\top\right)\left(\mathrm{vec}\left[\mathbf{G}_q^\top \mathbf{V}_j\right]^\top \otimes \left(\mathbf{G}_p^\top \mathbf{V}_i \otimes \mathbf{I}_m\right)\right) \times \right.$$

$$\left. \left(\mathbf{I}_{n^{j-1}} \otimes \mathbf{S}_{n^{i-1} \times n^q m} \otimes \mathbf{I}_m\right)\left(\mathbf{I}_{n^{k-p}} \otimes \mathrm{vec}\left[\mathbf{I}_m\right]^\top\right)\right] \mathbf{x}^{\textcircled{k}}$$

Table: Dimensions of matrices used in identities

$\mathbf{A}(p \times q)$	$\mathbf{D}(q \times s)$	$\mathbf{u}(s \times 1)$
$\mathbf{B}(s \times t)$	$\mathbf{G}(t \times u)$	$\mathbf{w}(t \times 1)$

ID. 1. $(\mathbf{A} \otimes \mathbf{B})(\mathbf{D} \otimes \mathbf{G}) = \mathbf{AD} \otimes \mathbf{BG}$

ID. 2. $\mathbf{A} \otimes \mathbf{B} = \mathbf{S}_{s \times p}(\mathbf{B} \otimes \mathbf{A})\mathbf{S}_{q \times t}$

ID. 3. $(\mathbf{I}_p \otimes \mathbf{w})\mathbf{A} = (\mathbf{A} \otimes \mathbf{w})$

ID. 4. $(\mathbf{w} \otimes \mathbf{I}_p)\mathbf{A} = (\mathbf{w} \otimes \mathbf{A})$

ID. 5. $\mathrm{vec}(\mathbf{ADB}) = (\mathbf{B}^\top \otimes \mathbf{A})\mathrm{vec}(\mathbf{D})$

ID. 6. $\mathbf{u}^\top \mathbf{B} \mathbf{w} = \mathrm{vec}(\mathbf{B})^\top (\mathbf{w} \otimes \mathbf{u})$

ID. 7. $\mathrm{vec}\left(\mathbf{w}^\top \otimes \mathbf{I}_m\right) = (\mathbf{w} \otimes \mathrm{vec}(\mathbf{I}_m))$

ID. 8. $\mathrm{vec}(\mathbf{A} \otimes \mathbf{B}) = \left(\mathbf{I}_q \otimes \mathbf{S}_{p \times t} \otimes \mathbf{I}_s\right)\left(\mathrm{vec}\left(\mathbf{A}\right) \otimes \mathrm{vec}\left(\mathbf{B}\right)\right)$

Recap before numerical results

- Interested in soft robot control with control-affine nonlinear models
- Use Taylor expansions to get polynomial models
- Polynomial model yields polynomial energy functions
- **Contribution: Kronecker product form of the algebraic equations for the energy coefficients when you have polynomial $\mathbf{g}(\mathbf{x})$**
- Solve Riccati equation and then a series of linear systems instead of a PDE
- Resulting energy function can be used for
  - control,
  - model reduction,
  - etc.

Numerical results

1D Example: exact solution

Consider the 1D quadratic-polynomial model

$$\dot{x} = ax + nx^2 + bu + g_1 xu + g_2 x^2 u,$$
$$y = cx.$$

The HJB PDE is actually an ODE for this 1D model:

$$0 = \frac{\mathrm{d}\mathcal{E}_\gamma^-(x)}{\mathrm{d}x} f(x) + \frac{1}{2}\left(\frac{\mathrm{d}\mathcal{E}_\gamma^-(x)}{\mathrm{d}x}\right)^2 g(x)^2 - \frac{\eta}{2} h(x)^2.$$

Let $q(x) = \mathrm{d}\mathcal{E}_\gamma^-(x)/\mathrm{d}x$, $\mathrm{a}(x) = g(x)^2/2$, $\mathrm{b}(x) = f(x)$, and $\mathrm{c}(x) = -\eta h(x)^2/2$. The HJB PDE takes the form of a standard quadratic equation in $q(x)$

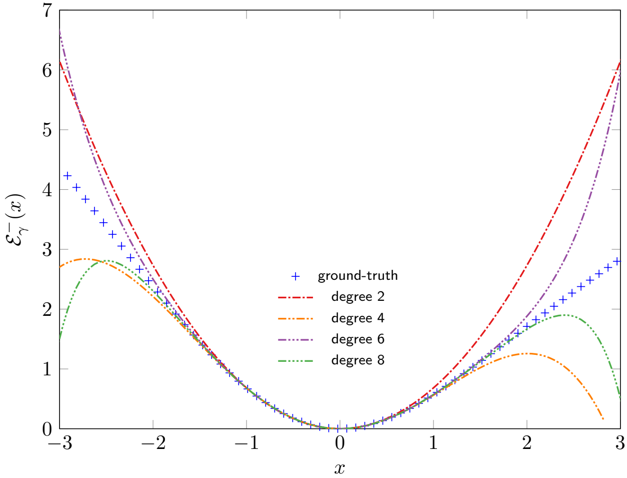
$$0 = \mathrm{a}(x) q(x)^2 + \mathrm{b}(x) q(x) + \mathrm{c}(x)$$

whose roots are given by

$$q(x) = \frac{-\mathrm{b}(x) \pm \sqrt{\mathrm{b}(x)^2 - 4\mathrm{a}(x)\mathrm{c}(x)}}{2\mathrm{a}(x)} \quad \rightarrow \quad \mathcal{E}_\gamma^-(x) = \int \frac{-\mathrm{b}(x) \pm \sqrt{\mathrm{b}(x)^2 - 4\mathrm{a}(x)\mathrm{c}(x)}}{2\mathrm{a}(x)} \mathrm{d}x.$$

1D Example: energy function plots

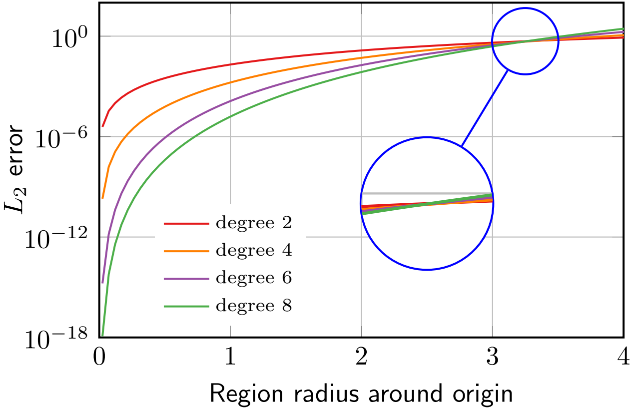
Results for $a = -2$, $n = 1$, $b = 2$, $g_1 = -0.2$, $g_2 = 0.2$, $c = 2$, and $\eta = 0.5$.



$$\begin{aligned}\mathcal{E}_\gamma^-(x) \approx\ & 1.3660x^2 - 1.1925 \times 10^{-1}x^3 \\ & - 1.2475 \times 10^{-1}x^4 + 1.0509 \times 10^{-2}x^5 \\ & + 1.4282 \times 10^{-2}x^6 - 1.5279 \times 10^{-4}x^7 \\ & - 1.6133 \times 10^{-3}x^8\end{aligned}$$

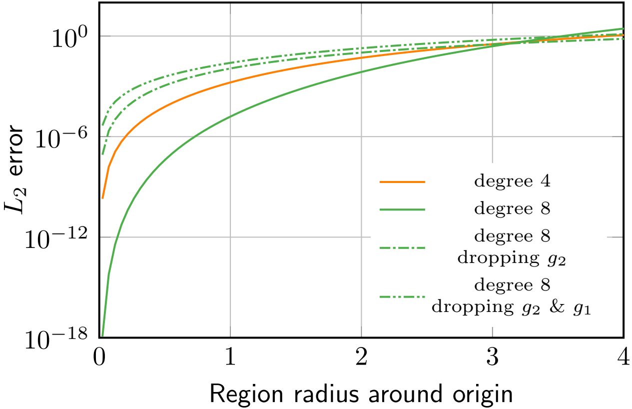
- The energy functions are accurate locally, and adding degrees aids accuracy locally
- As is often the case with polynomial approximation, higher-order approximations also diverge more severely beyond the region of convergence

1D Example: energy function L_2 error



- Adding higher-order terms increases accuracy locally.
- It does not necessarily increase the radius of convergence!

1D Example: is it better to discard G_ξ terms?



$$\dot{x} = ax + nx^2 + bu + \cancel{g_1 xu} + \cancel{g_2 x^2 u},$$
$$y = cx.$$

- Probably not.
- Better to approximate the *right* energy function to lower order than the *wrong* energy function to higher order.

2D Example: energy functions and residuals

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

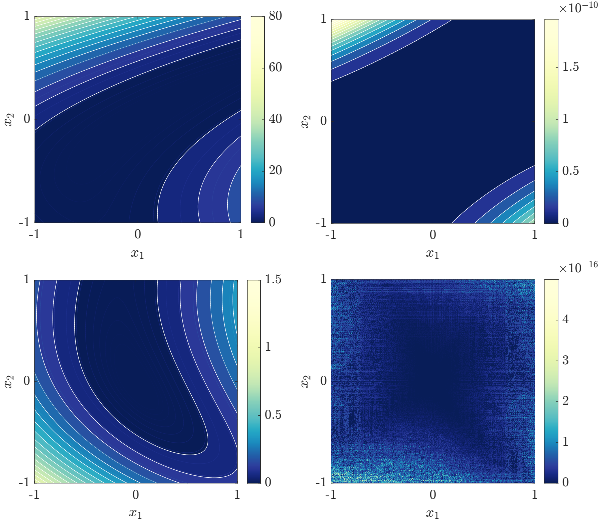
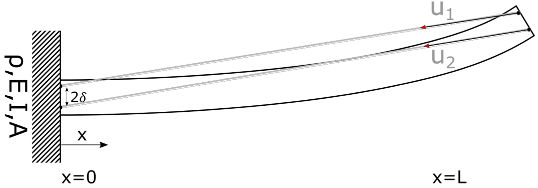


Figure: Left: nonlinear balancing energy functions. Right: HJB residuals.

Nonlinear beam

Consider a cable-actuated cantilever beam. We model this as a nonlinear Euler-Bernoulli beam subject to Von Kármán strains but linear elastic material response [Reddy, 2004].

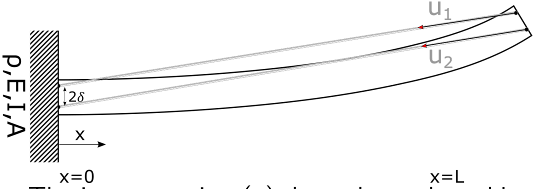


$$
\begin{aligned}
0 &= \rho A \frac{\partial^2 w}{\partial t^2} - \frac{\partial N_{xx}}{\partial x}, \\
0 &= \rho A \frac{\partial^2 v}{\partial t^2} - \frac{\partial}{\partial x}\left(N_{xx}\frac{\partial v}{\partial x}\right) - \rho I \frac{\partial^4 v}{\partial t^2 \partial x^2} + \frac{\partial^2 M_{xx}}{\partial x^2},
\end{aligned}
$$

where

$$
N_{xx} = EA\left[\frac{\partial w}{\partial x} + \frac{1}{2}\left(\frac{\partial v}{\partial x}\right)^2\right] \qquad \text{and} \qquad M_{xx} = EI\frac{\partial^2 v}{\partial x^2}.
$$

Nonlinear beam: finite element discretization



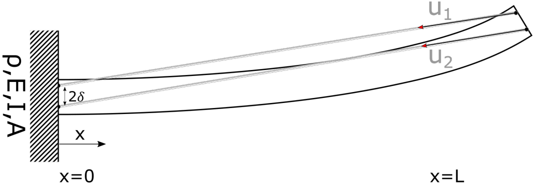
The semi-discretized truncated system can be written in state-space form as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}_2\mathbf{x}^{②} + \mathbf{F}_3\mathbf{x}^{③} + \mathbf{g}(\mathbf{x})\mathbf{u},$$
$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

The input matrix $\mathbf{g}(\mathbf{x})$ depends on the cable angle $\theta(\mathbf{x})$, which changes with the state. We can model the angular dependence using simple geometry and Taylor series expansion as

$$\cos(\theta(v,w)) = \frac{L + w(L,t)}{\sqrt{(L+w(L,t))^2 + v(L,t)^2}}, \quad \rightarrow \quad \approx \left(1 - \frac{v(L,t)^2}{2L^2} + \frac{w(L,t)v(L,t)^2}{L^3}\right),$$

$$\sin(\theta(v,w)) = \frac{v(L,t)}{\sqrt{(L+w(L,t))^2 + v(L,t)^2}}, \quad \rightarrow \quad \approx \left(\frac{v(L,t)}{L} - \frac{w(L,t)v(L,t)}{L^2} + \frac{(2w(L,t)^2 v(L,t) - v(L,t)^3)}{2L^3}\right).$$



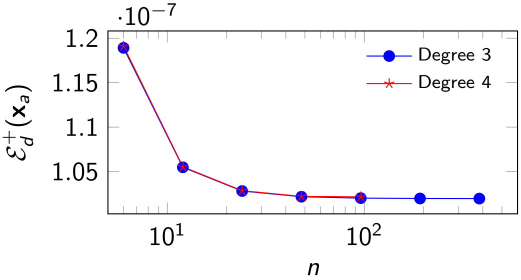


Figure: Convergence with respect to mesh size

With the degree 3 Taylor approximation to $\mathbf{g}(\mathbf{x})$ and discarding the cubic drift term, the model is

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{F}\mathbf{x}^{②} + \sum_{\xi=1}^{3} \mathbf{G}_\xi \left(\mathbf{x}^{ⓧ} \otimes \mathbf{u}\right) + \mathbf{B}\mathbf{u},$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}$$

where $\mathbf{x} \in \mathbb{R}^{6N}$ and N is the number of elements.

By varying the number of elements, we can investigate convergence and scaling performance.

Energy function convergence for the beam

Table: $n = 18$, scaling and convergence w.r.t. d

d	CPU sec	$\mathcal{E}_d^+(\mathbf{x}_a)$	$\mathcal{E}_d^+(\mathbf{x}_b)$
2	$2.18 \cdot 10^{-2}$	$1.012561 \cdot 10^{-7}$	$1.012561 \cdot 10^{-5}$
3	$9.04 \cdot 10^{-3}$	$1.034814 \cdot 10^{-7}$	$1.235090 \cdot 10^{-5}$
4	$1.89 \cdot 10^{-1}$	$1.035239 \cdot 10^{-7}$	$1.277524 \cdot 10^{-5}$
5	$4.19 \cdot 10^{0}$	$1.035240 \cdot 10^{-7}$	$1.278660 \cdot 10^{-5}$
6	$9.82 \cdot 10^{1}$	$1.035240 \cdot 10^{-7}$	$1.276495 \cdot 10^{-5}$

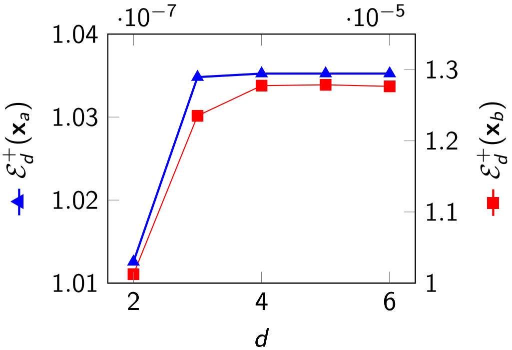


Figure: Energy function convergence as d increases for initial conditions $\mathbf{x}_a, \mathbf{x}_b$

Algorithm scaling for beam example

The solution implementation in [Kramer et al., 2022] for these types of linear systems scales as roughly $O(n^{d+1})$, though their results perform more like $O(n^d)$.

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

No. of Elements	n	n^3	CPU sec	$\mathcal{E}_3^+(\mathbf{x}_a)$
1	6	$2.1600 \cdot 10^2$	$1.37 \cdot 10^{-3}$	$1.189137 \cdot 10^{-7}$
2	12	$1.7280 \cdot 10^3$	$2.80 \cdot 10^{-3}$	$1.054830 \cdot 10^{-7}$
4	24	$1.3824 \cdot 10^4$	$1.30 \cdot 10^{-2}$	$1.028107 \cdot 10^{-7}$
8	48	$1.1059 \cdot 10^5$	$9.57 \cdot 10^{-2}$	$1.021709 \cdot 10^{-7}$
16	96	$8.8474 \cdot 10^5$	$9.34 \cdot 10^{-1}$	$1.020124 \cdot 10^{-7}$
32	192	$7.0779 \cdot 10^6$	$1.05 \cdot 10^1$	$1.019728 \cdot 10^{-7}$
64	384	$5.6623 \cdot 10^7$	$3.89 \cdot 10^2$	$1.019625 \cdot 10^{-7}$

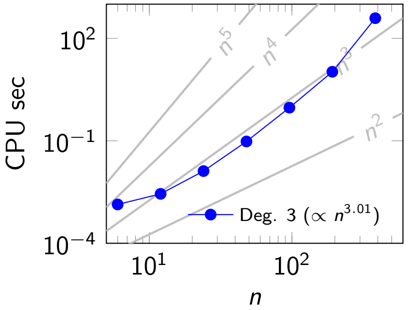


Figure: Scaling of CPU time as n increases

Algorithm scaling for beam example

The solution implementation in [Kramer et al., 2022] for these types of linear systems scales as roughly $O(n^{d+1})$, though their results perform more like $O(n^d)$.

Table: $d = 4$, scaling and convergence w.r.t. n

No. of Elements	n	n^4	CPU sec	$\mathcal{E}_4^+(\mathbf{x}_a)$
1	6	$1.2960{\cdot}10^3$	$2.00{\cdot}10^{-2}$	$1.191000{\cdot}10^{-7}$
2	12	$2.0736{\cdot}10^4$	$3.94{\cdot}10^{-2}$	$1.055358{\cdot}10^{-7}$
4	24	$3.3178{\cdot}10^5$	$5.08{\cdot}10^{-1}$	$1.028498{\cdot}10^{-7}$
8	48	$5.3084{\cdot}10^6$	$8.32{\cdot}10^0$	$1.022069{\cdot}10^{-7}$
16	96	$8.4935{\cdot}10^7$	$1.83{\cdot}10^2$	$1.021582{\cdot}10^{-7}$

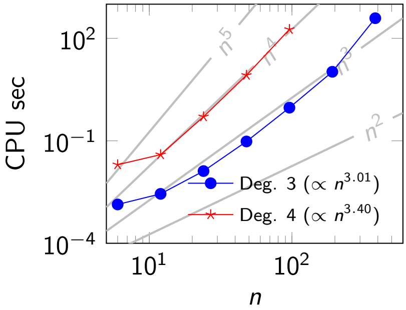


Figure: Scaling of CPU time as n increases

Conclusions

- Presented generalization of [Kramer et al., 2022] to polynomial inputs $\mathbf{g}(\mathbf{x})$
- Linear systems remain similar, only different right-hand sides (allows to inherit solvability, solution algorithms, scaling, etc.)

Future work:

- Extend to polynomial $\mathbf{f}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$ for full polynomial systems
- State feedback control (using these energy functions and others)
- Soft robot examples, model reduction, and output feedback control
- Optimize implementation (current speed OK but also RAM usage), maybe tensor_toolbox [2]

[2] [Bader et al., 2023]

Al'brekht, E. G. (1961).
On the optimal stabilization of nonlinear systems.
Journal of Applied Mathematics and Mechanics, 25(5):1254–1266.

Bader, B. W., Kolda, T. G., et al. (2023).
Tensor Toolbox for MATLAB, Version 3.5.

Brewer, J. (1978).
Kronecker products and matrix calculus in system theory.
IEEE Transactions on Circuits and Systems, 25(9):772–781.

Harrower, M. and Brewer, C. A. (2003).
ColorBrewer.org: an online tool for selecting colour schemes for maps.
The Cartographic Journal, 40(1):27–37.

Kawano, Y. and Scherpen, J. M. A. (2017).
Model reduction by differential balancing based on nonlinear hankel operators.
IEEE Transactions on Automatic Control, 62(7):3293–3308.

Kramer, B., Gugercin, S., Borggaard, J., and Balicki, L. (2022).
Nonlinear balanced truncation: part 1-computing energy functions.

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

Magnus, J. R. and Neudecker, H. (2019).
Matrix differential calculus with applications in statistics and econometrics.
Wiley, third edition edition.

Reddy, J. N. (2004).
An introduction to nonlinear finite element analysis.
Oxford University Press.

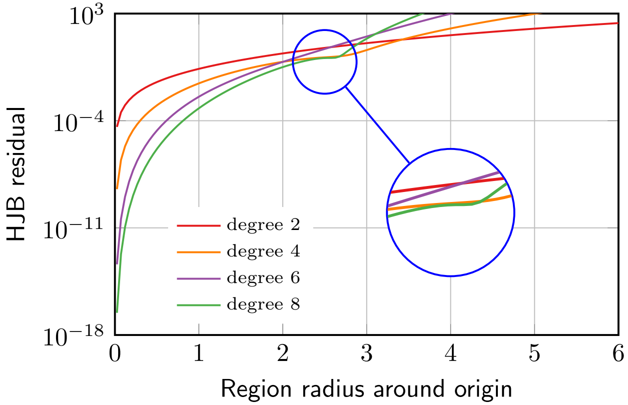
Scherpen, J. M. A. (1996).
H∞ balancing for nonlinear systems.
International Journal of Robust and Nonlinear Control, 6(7):645–668.

Thieffry, M., Kruszewski, A., Duriez, C., and Guerra, T.-m. (2018).
Control design for soft robots based on reduced order model.
IEEE Robotics and Automation Letters, pages 1–1.

Zhang, Z., Dequidt, J., Kruszewski, A., Largilliere, F., and Duriez, C. (2016).

Kinematic modeling and observer based control of soft robot using real-time finite element method.
In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE.

1D Example: HJB residual as a metric



- Maybe a better measure of a 'good' energy function, since it involves a) the HJB equation and b) gradients of the energy

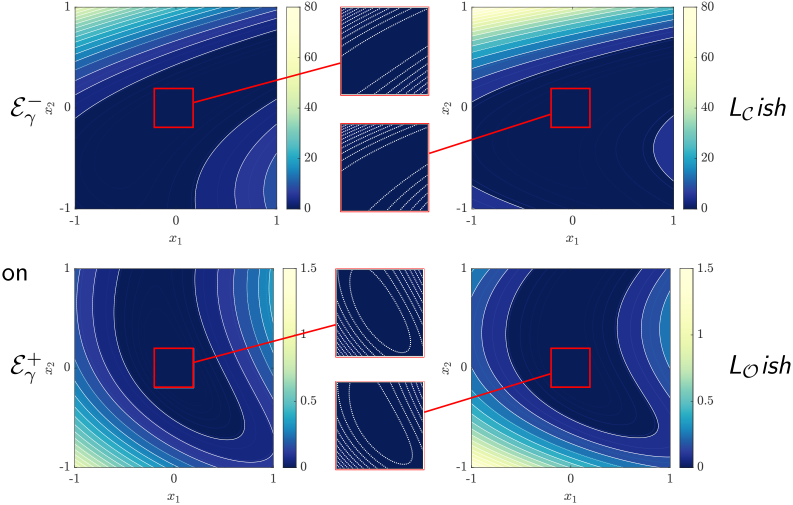
2D Example: nonlinear balancing vs. differential balancing

Next we take the 2D quadratic-bilinear system from [Kawano and Scherpen, 2017]:

$$\dot{x}_1 = -x_1 + x_2 - x_2^2 + u + 2x_2 u$$
$$\dot{x}_2 = -x_2 + u$$
$$y = x_1$$

We use a degree 4 approximation with $\eta = 0$ for open-loop balancing.

$$L_{\mathcal{C}} = \frac{1}{2}\delta \mathbf{x}_0^\top \mathbf{Q}(\mathbf{x}_0)\delta \mathbf{x}_0$$
$$L_{\mathcal{C}} ish = \frac{1}{2}\mathbf{x}_0^\top \mathbf{Q}(\mathbf{x}_0)\mathbf{x}_0$$



Nonlinear balancing for polynomial systems

Theorem (Past energy polynomial coefficients)

For $3 \leq k \leq d$, let $\tilde{\mathbf{v}}_k \in \mathbb{R}^{n^k}$ solve the linear system

$$\mathcal{L}_k\left(\mathbf{A}+\mathbf{B}\mathbf{B}^\top\mathbf{V}_2\right)^\top \tilde{\mathbf{v}}_k = -\sum_{\substack{i,\xi\geq 2\\ \xi+i=k+1}} \mathcal{L}_i(\mathbf{F}_\xi)^\top \mathbf{v}_i - \frac{1}{4}\sum_{\substack{i,j>2\\ i+j=k+2}} ij \operatorname{vec}\left[\mathbf{V}_i^\top \mathbf{B}\mathbf{B}^\top \mathbf{V}_j\right] + \eta \sum_{\substack{p,q\geq 1\\ p+q=k}} \operatorname{vec}\left[\mathbf{H}_p^\top \mathbf{H}_q\right]$$
$$-\frac{1}{4}\sum_{o=1}^{2\ell}\left(\sum_{\substack{p,q\geq 0\\ p+q=o}}\left(\sum_{\substack{i,j\geq 2\\ i+j=k-o+2}} ij \operatorname{vec}\left[\left(\mathbf{I}_{n^p}\otimes \operatorname{vec}\left[\mathbf{I}_m\right]^\top\right)\left(\operatorname{vec}\left[\mathbf{G}_q^\top \mathbf{V}_j\right]^\top \otimes \left(\mathbf{G}_p^\top \mathbf{V}_i \otimes \mathbf{I}_m\right)\right)\times\right.\right.\right.$$
$$\left.\left.\left.\left(\mathbf{I}_{n^{j-1}}\otimes \mathbf{S}_{n^{i-1}\times n^q m}\otimes \mathbf{I}_m\right)\left(\mathbf{I}_{n^{k-p}}\otimes \operatorname{vec}\left[\mathbf{I}_m\right]\right)\right]\right)\right) \tag{16}$$

Then the $\mathbf{v}_k = \operatorname{vec}(\mathbf{V}_k) \in \mathbb{R}^{n^k}$ for $3 \leq k \leq d$ is obtained by symmetrization of $\tilde{\mathbf{v}}_k$.

2D Example: approximate input normal transformations

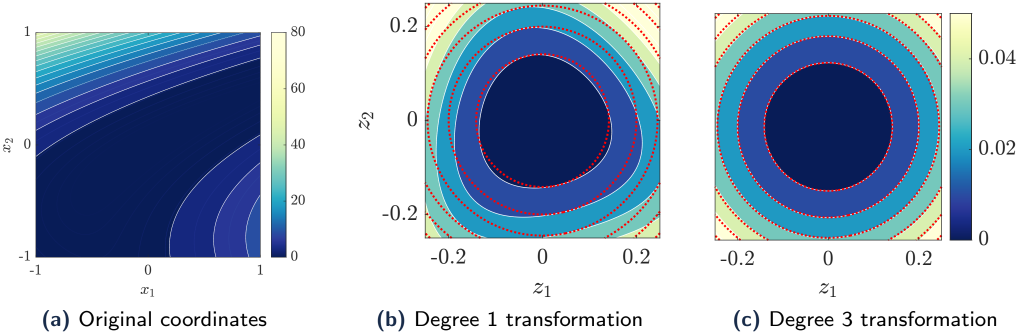


Figure: Past energy function under approximate input normal transformation