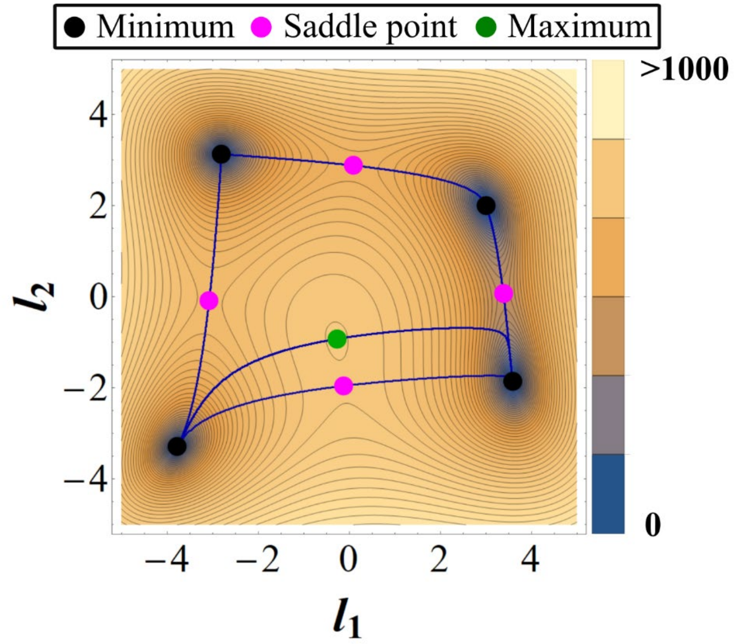
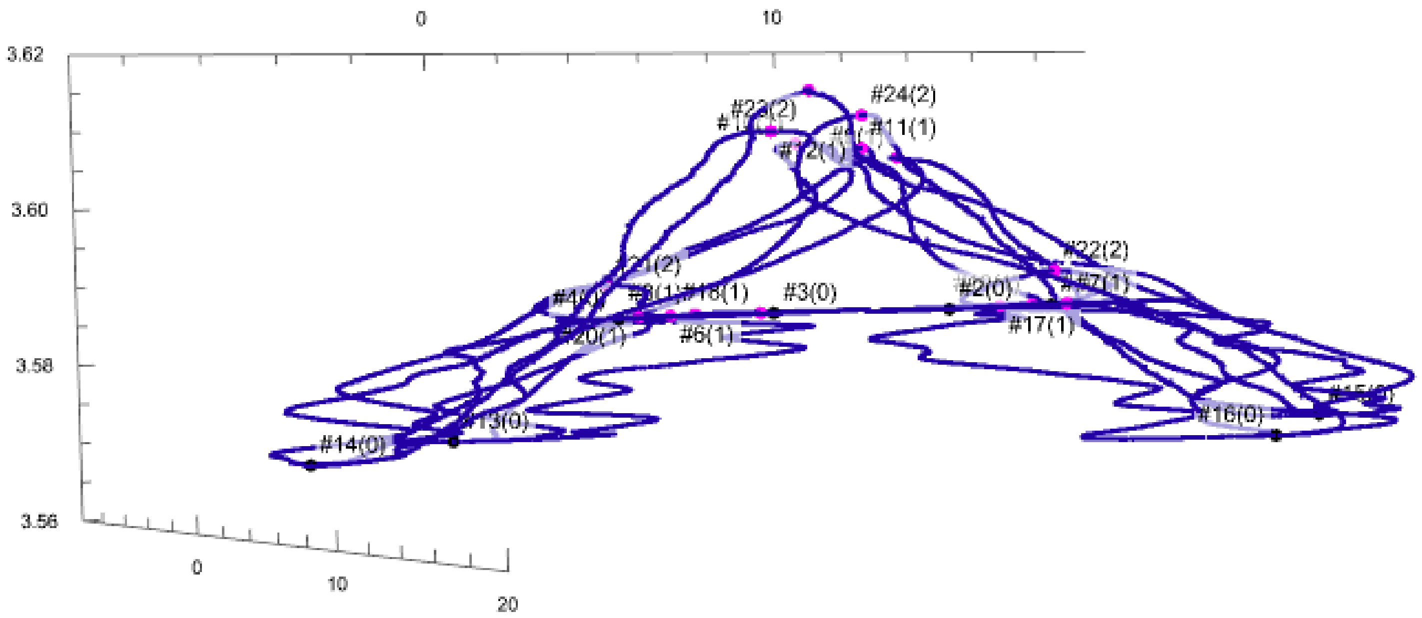
Fixed points, saddle graphs, and numerical algebraic geometry





Jonathan Hauenstein

Dynamical systems and Semi-algebraic geometry: interactions with Optimization and Deep Learning

July 19, 2023

Postdoc position available now

Optimization, Machine Learning, and Algebraic Geometry

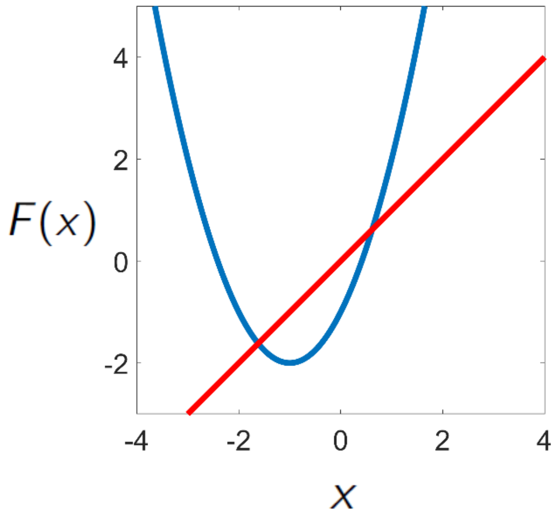
https://www.mathjobs.org/jobs/list/22544

hauenstein@nd.edu

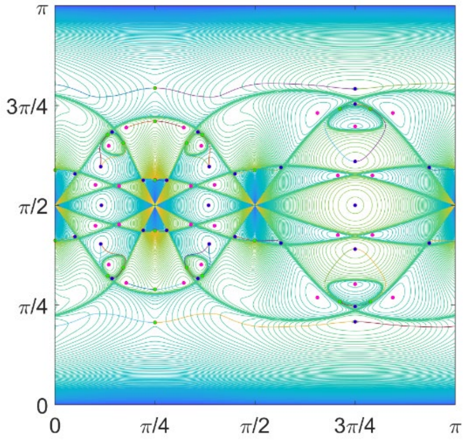
Introduction

Polynomial system $F : \mathbb{R}^n \to \mathbb{R}^n$

map *fixed point* solves $F(x) - x = 0$

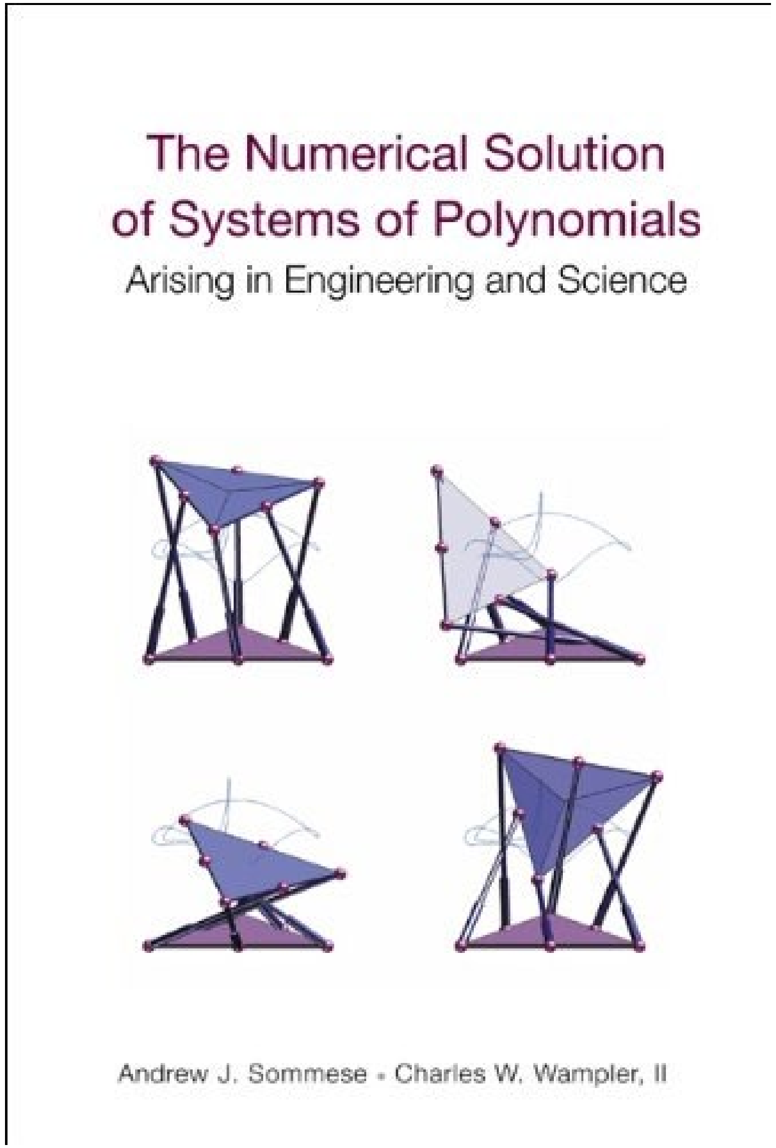


dyn sys *fixed point* of $\dot{x} = F(x)$ solves $F(x) = 0$

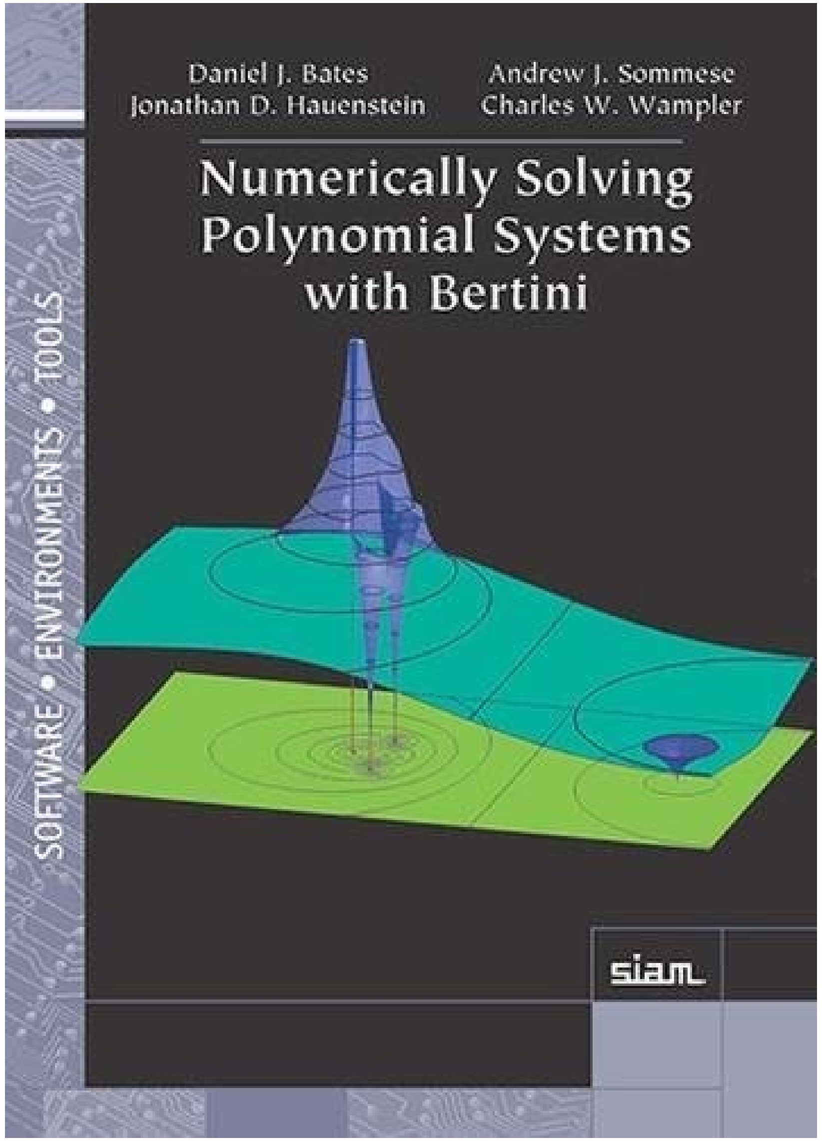


Key: compute fixed points by solving polynomial equations

Numerical Algebraic Geometry



Sommese-Wampler (2005)



Bates-H-Sommese-Wampler (2013)

Homotopy Continuation

Example

Solve

$$f(x,y) = \begin{bmatrix} x^2 + 4xy + 4y^2 - 8x - 9y + 8 \\ 4x^2 - 12xy + 9y^2 - 7x + 14y - 2 \end{bmatrix} = 0$$

Homotopy Continuation

Example

Solve

$$f(x,y) = \begin{bmatrix} x^2 + 4xy + 4y^2 - 8x - 9y + 8 \\ 4x^2 - 12xy + 9y^2 - 7x + 14y - 2 \end{bmatrix} = 0$$

Too difficult! Solve an easier problem:

$$g(x,y) = \begin{bmatrix} x^2 - 4 \\ y^2 - 1 \end{bmatrix} = 0$$

$$\text{Solutions} = \{(2,1),(2,-1),(-2,1),(-2,-1)\}$$

Homotopy Continuation

Example

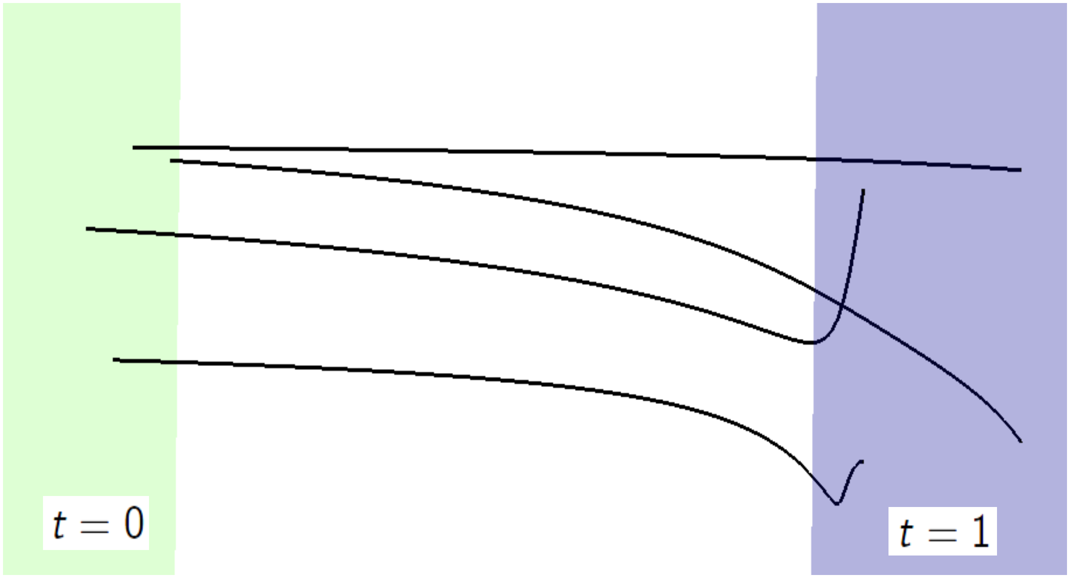
$$f(x,y) = \begin{bmatrix} x^2 + 4xy + 4y^2 - 8x - 9y + 8 \\ 4x^2 - 12xy + 9y^2 - 7x + 14y - 2 \end{bmatrix} = 0 \qquad g(x,y) = \begin{bmatrix} x^2 - 4 \\ y^2 - 1 \end{bmatrix} = 0$$

Deform from simplified (start) system to original (target) system

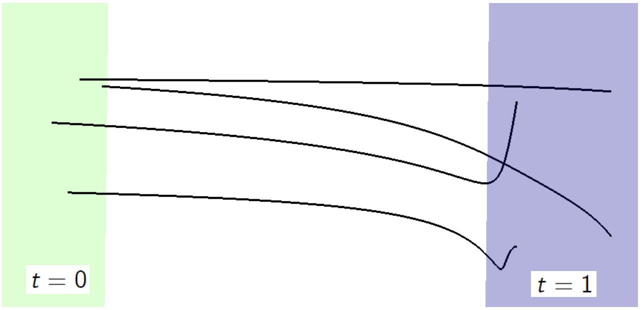
$$H(x,t) = (1-t)f(x) + tg(x) = 0$$

start $H(x,1) = g(x) = 0$ has known solutions

target $H(x,0) = f(x) = 0$ is system want to solve



Homotopy Continuation



Paths are defined by $H(x, t) = 0$:

$$H(x,t) = 0 \implies \frac{d}{dt}H(x,t) = 0$$

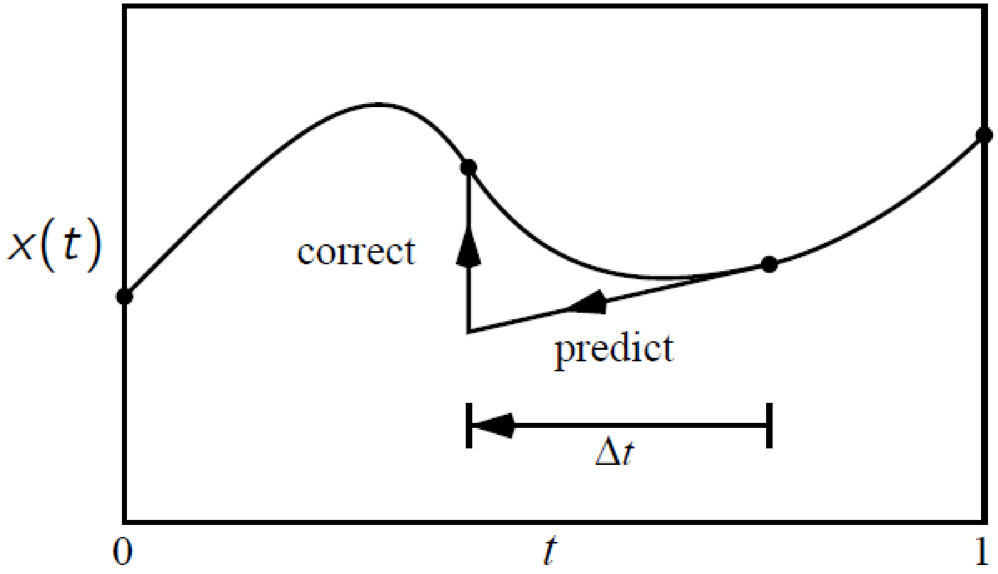
$$\implies \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial t} = 0$$

$$\implies \frac{dx}{dt} = -\left(\frac{\partial H}{\partial x}\right)^{-1}\frac{\partial H}{\partial t}$$

Davidenko differential equation

- solving polynomial systems $\implies$ solving initial value problems

Homotopy Continuation



predict Use $\dfrac{dx}{dt} = -\left(\dfrac{\partial H}{\partial x}\right)^{-1} \dfrac{\partial H}{\partial t}$ to estimate $x(t + \Delta t)$ given $x(t)$.

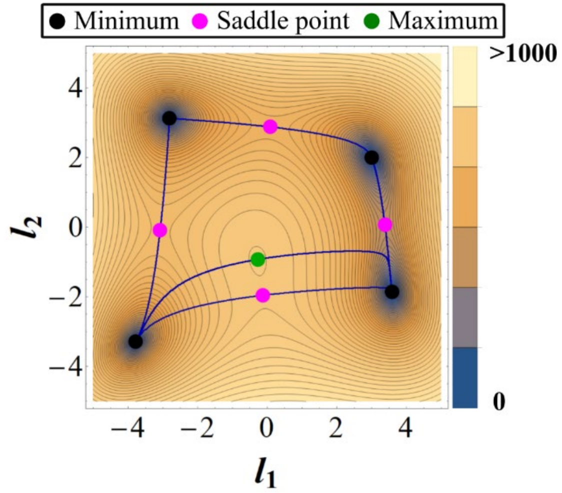
correct Use Newton's method applied to $H(x, t + \Delta t) = 0$.

- error control

Saddle Graphs

Represent landscapes by graphs with

- vertices: fixed points
- edges: gradient descent paths



Joint work with

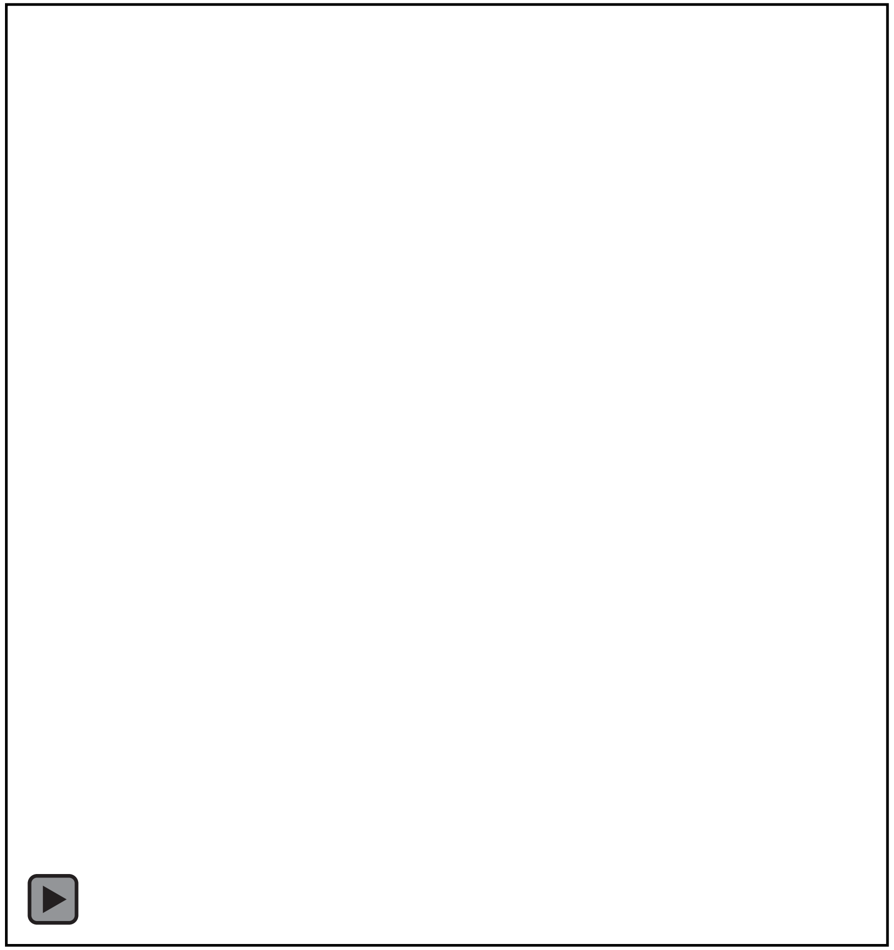
- Aravind Baskar (Notre Dame)
- Mark Plecnik (Notre Dame)

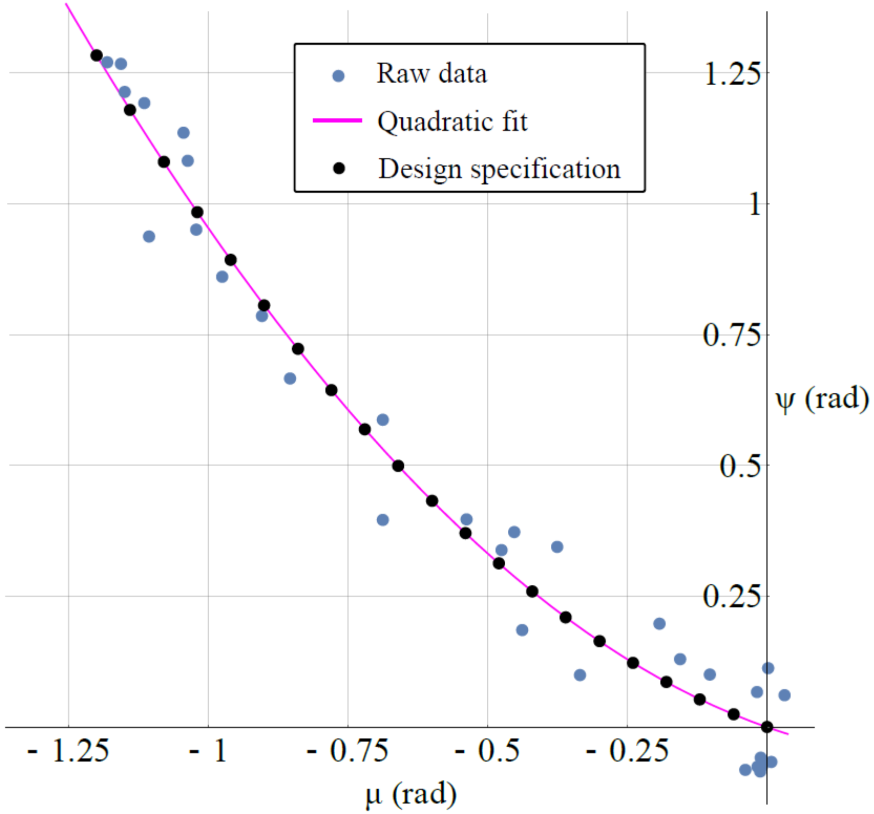
A. Baskar, M. Plecnik, and J.D. Hauenstein.
Computing saddle graphs via homotopy continuation for the approximate synthesis of mechanisms.
Mech. Mach. Theory, 176, 104932, 2022.

Mechanism Synthesis

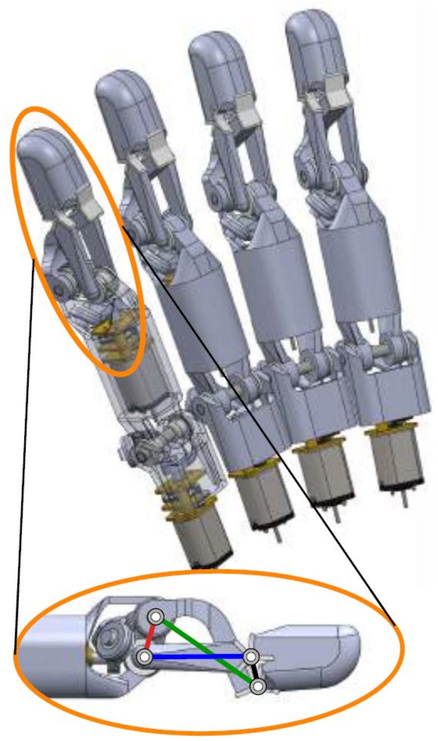
Design a mechanism to accomplish specified tasks.

- approximate synthesis: find “best fit” linkage



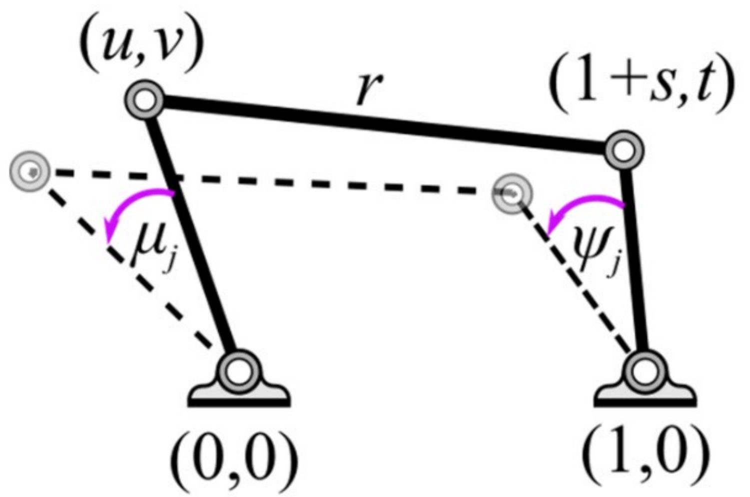


Mechanism Synthesis

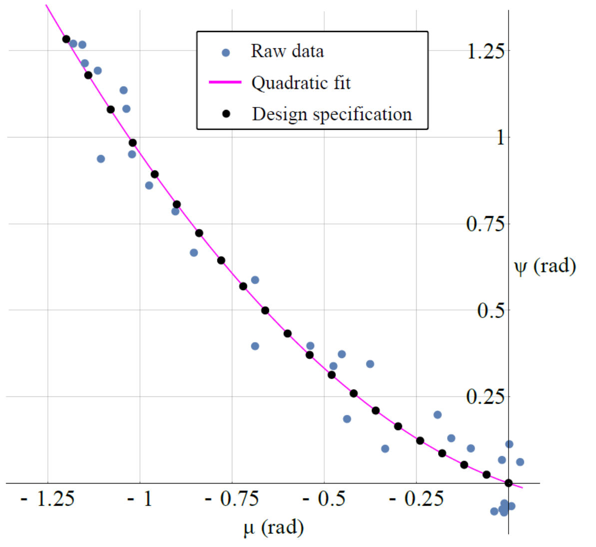


Function generation using a four-bar linkage:

- Model parameters $p = (r, s, t, u, v)$
- Output angle Ψ is a function of input angle μ
  - $\Psi = m(\mu; p)$
- Compute four-bar linkage that "best" fits finger motion data

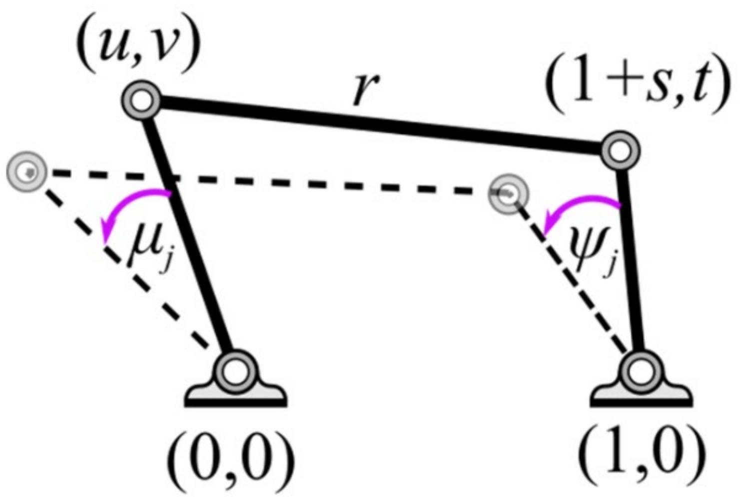


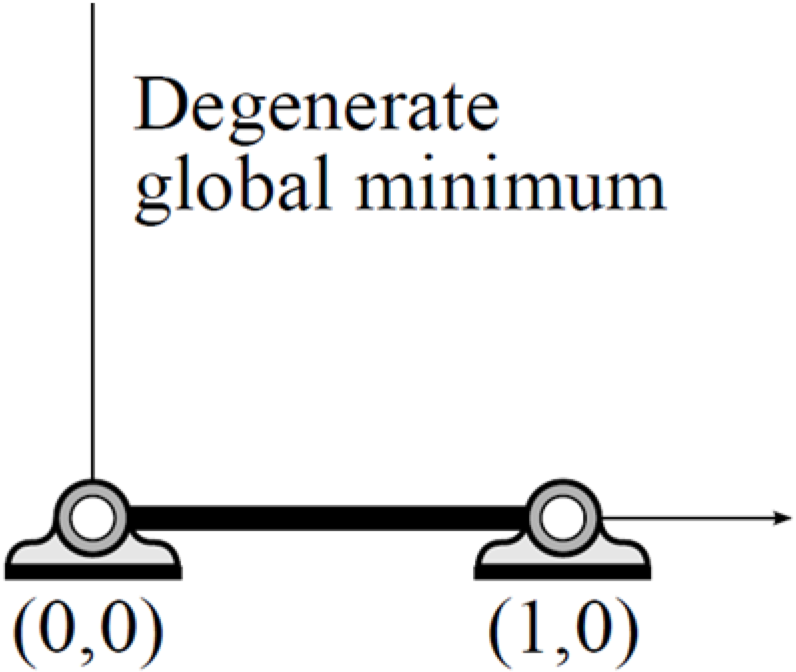
$$\min_{p} \sum_{j} (m(\mu_j; p) - \Psi_j)^2$$



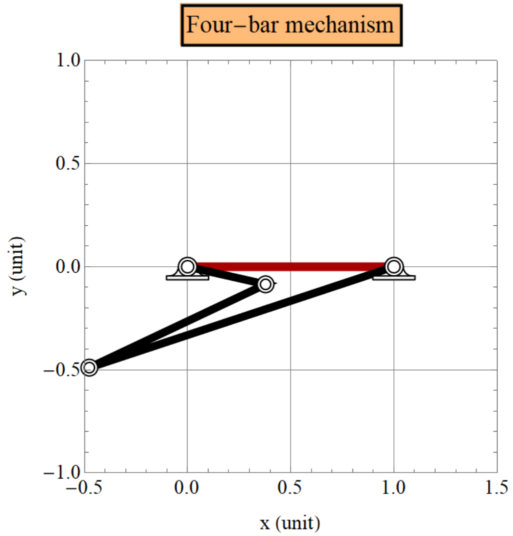
Mechanism Synthesis

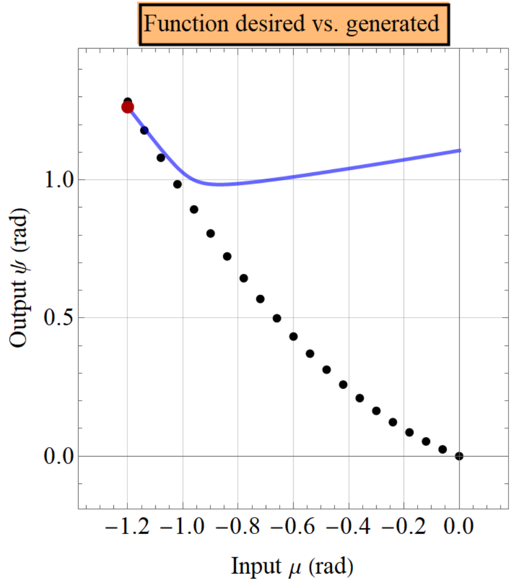
$$\min_{p} \sum_{j} (m(\mu_j; p) - \Psi_j)^2$$

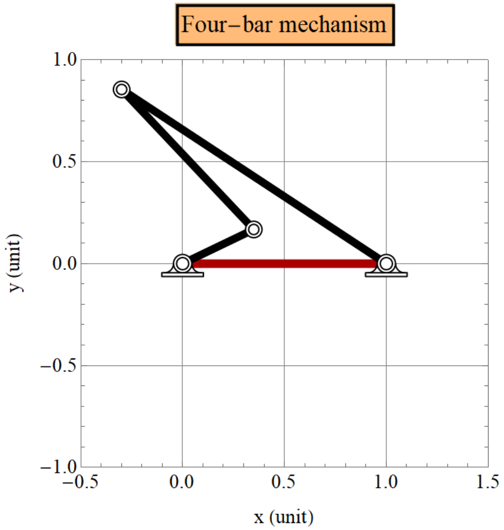


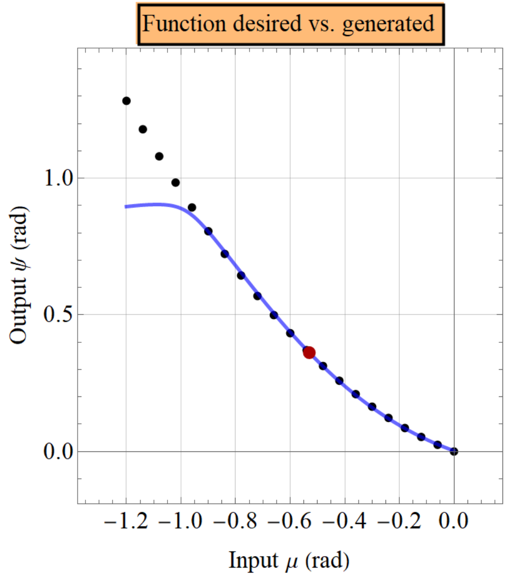


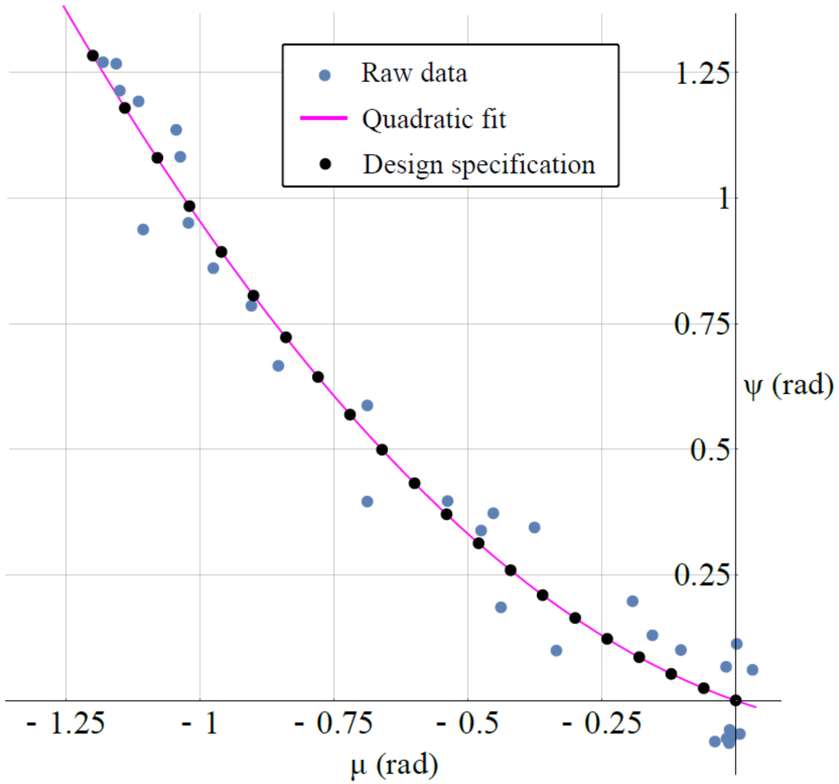
▶ Local minimum with branch defect



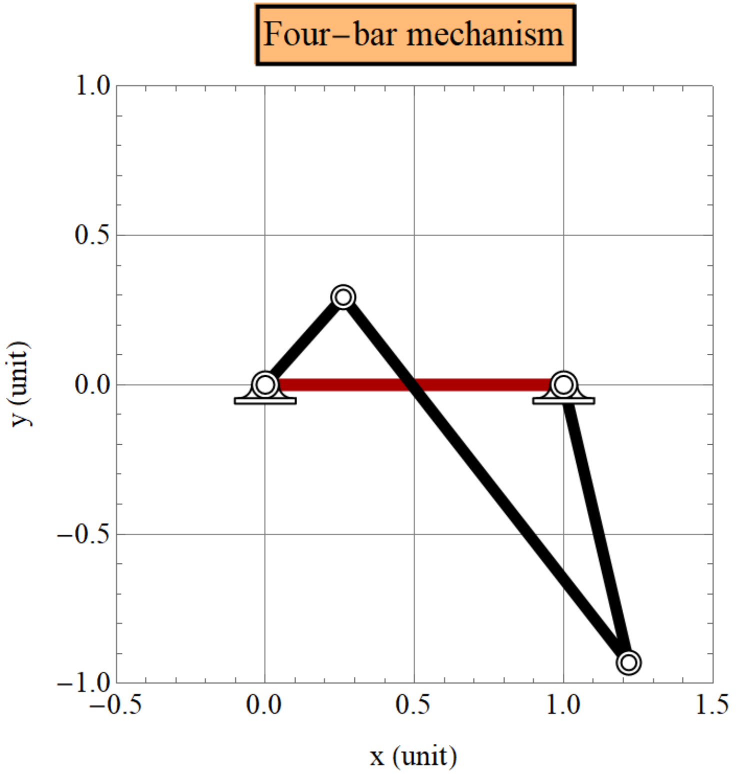


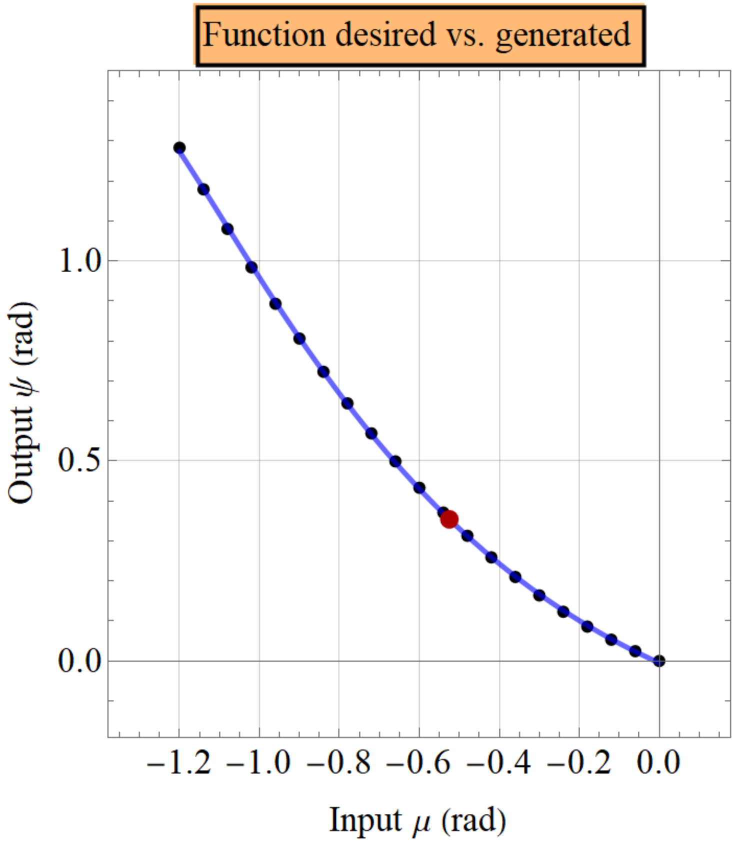


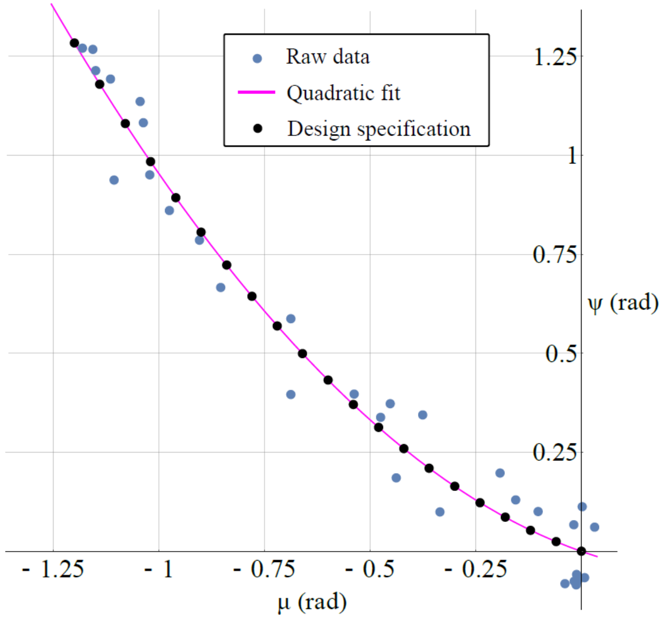


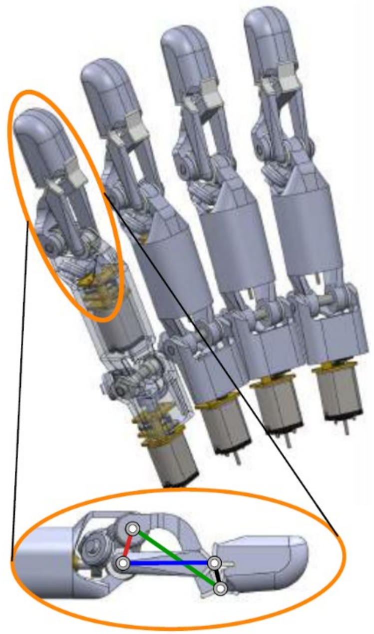


- Local minimum with legs of quite different lengths
  - difficult to package into finger linkage
  - potential problems with torque required for motion









Mechanism Synthesis

$$\min_{p} \sum_{j} (m(\mu_j; p) - \Psi_j)^2$$

Optimization problem does not include all constraints.

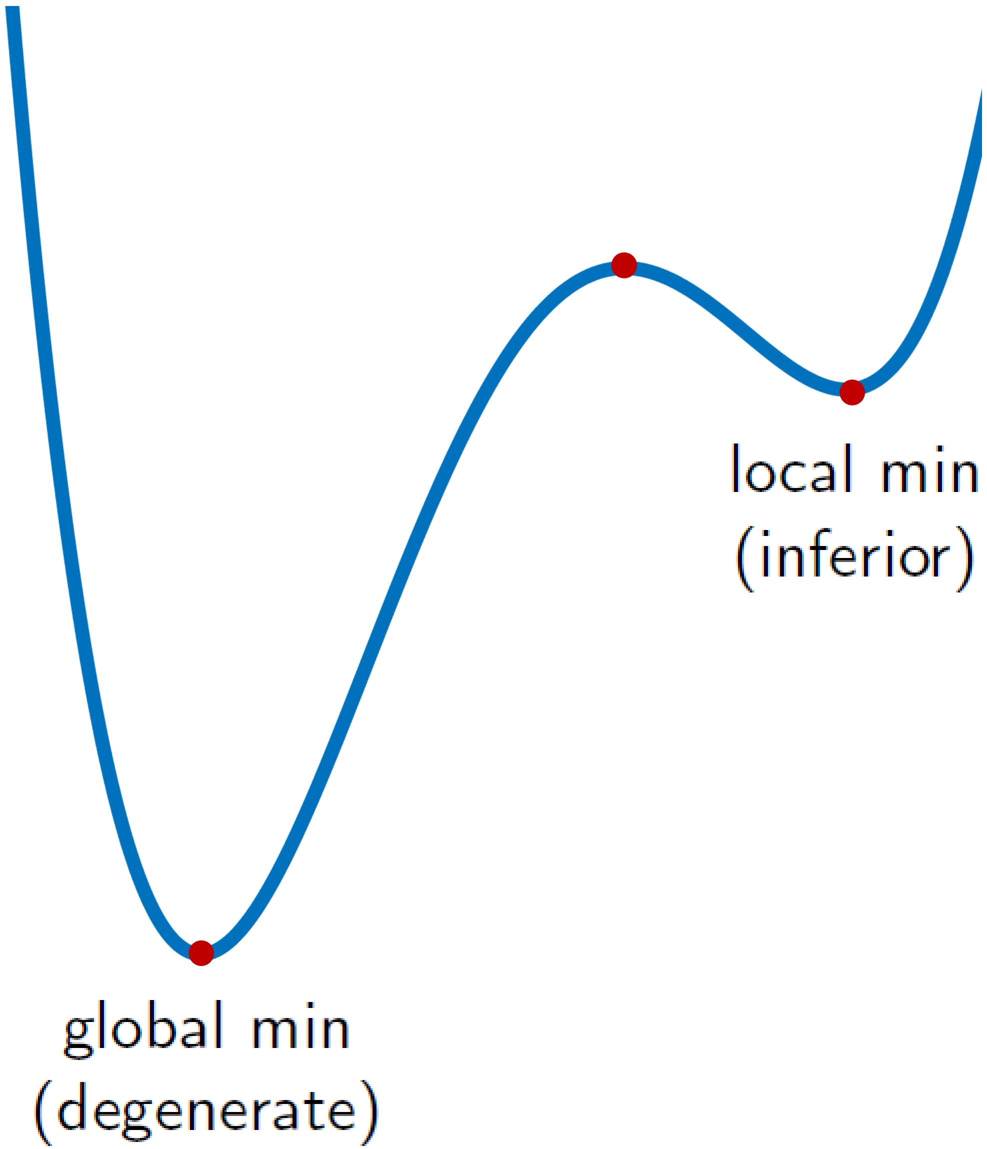
- Hard to formulate

Baskar-Plecnik-H.: compute a 1-dim'l view of landscape

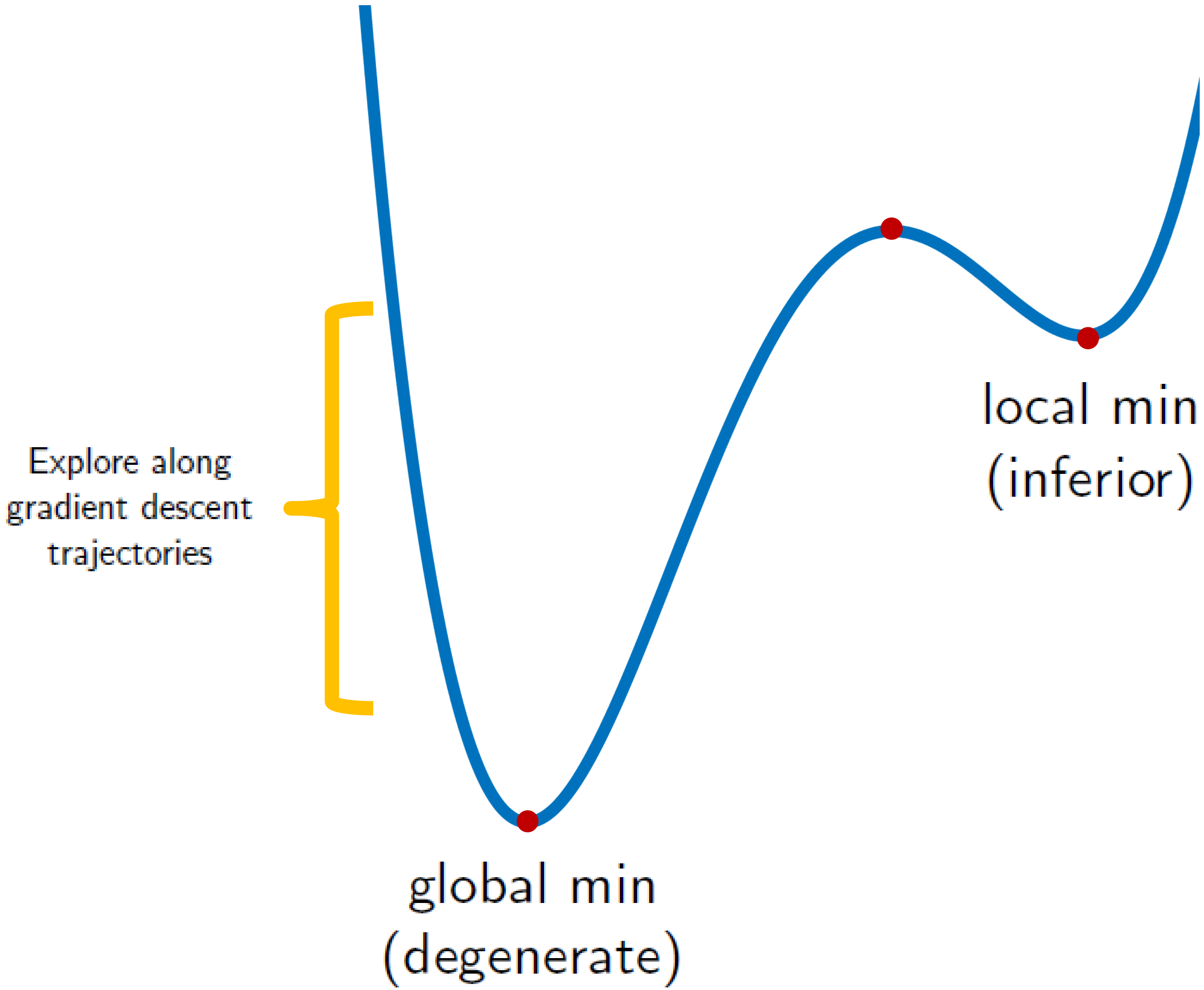
- gives designer freedom to find their "best fit"
- apply their own constraints *a posteriori*

More about the **journey** rather than the destination

Mechanism Design

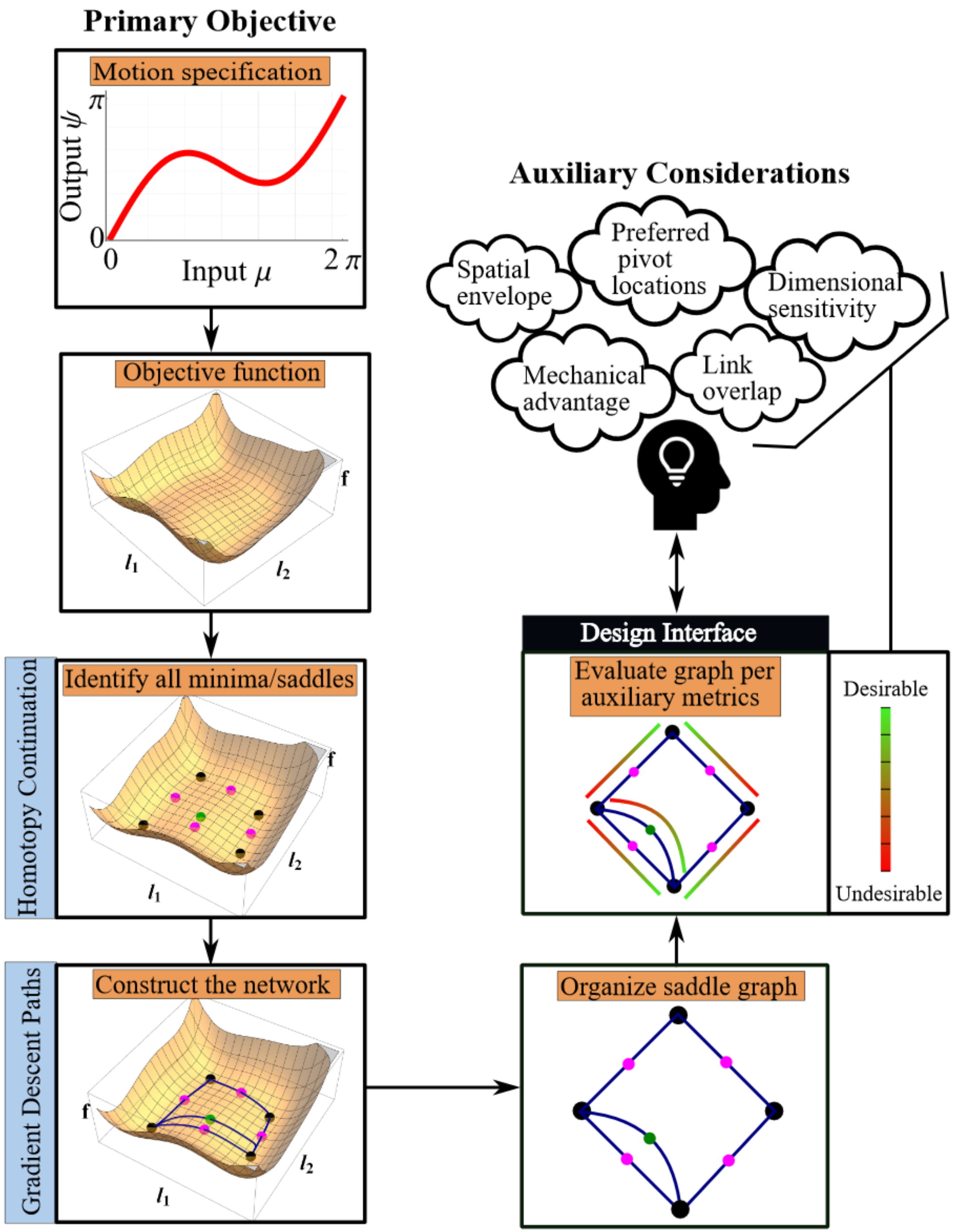


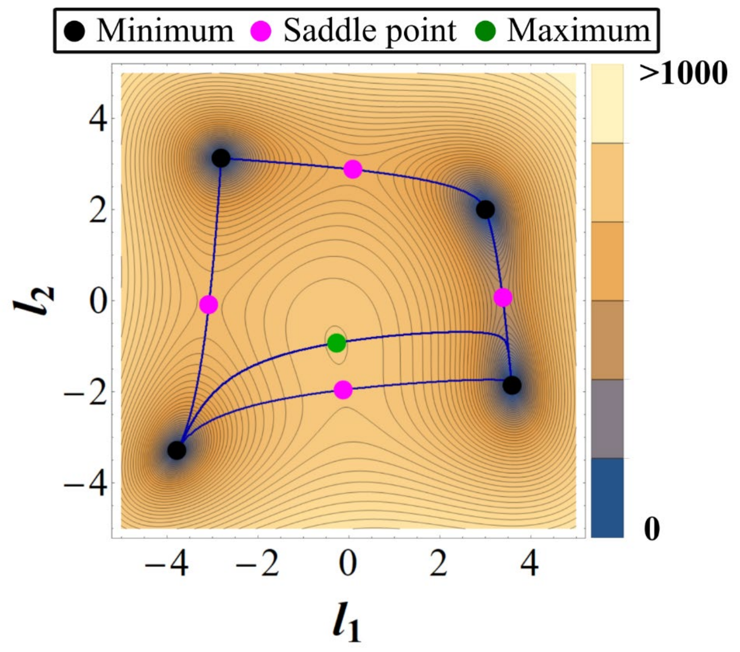
Mechanism Design



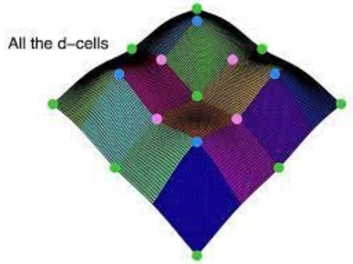
solve $\nabla f(x) = 0$

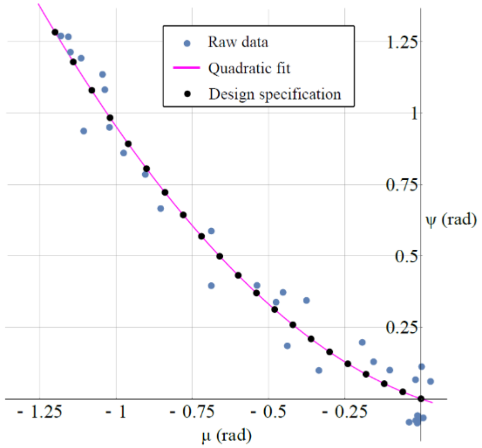
track $\dot{x} = -\nabla f$

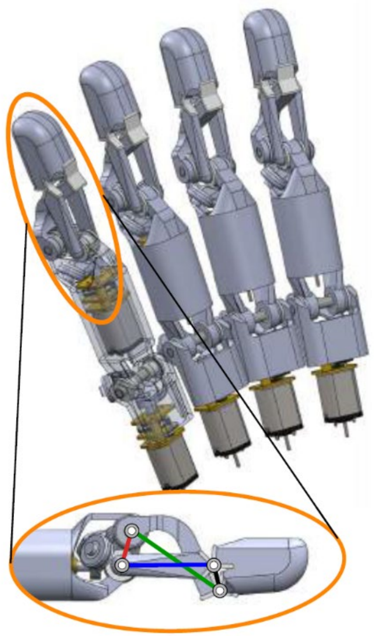




motivated by Morse and Morse-Smale complexes

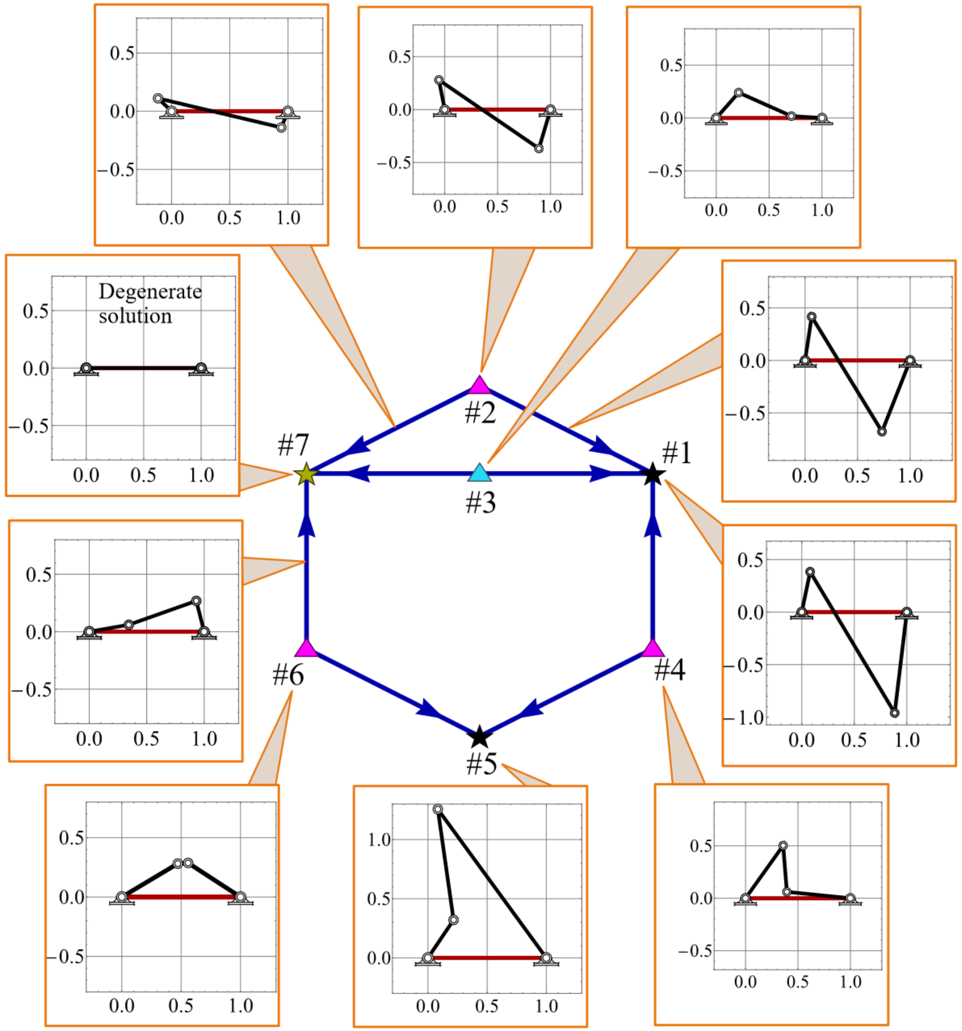




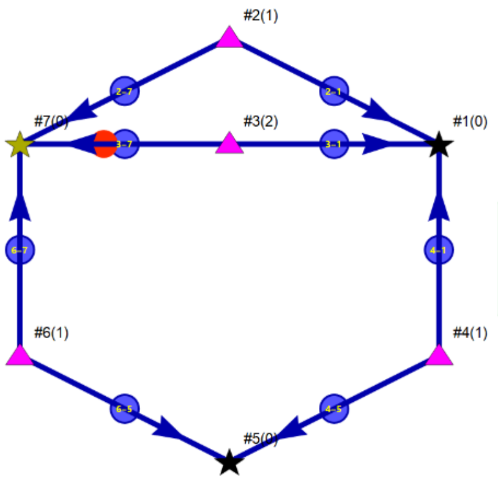


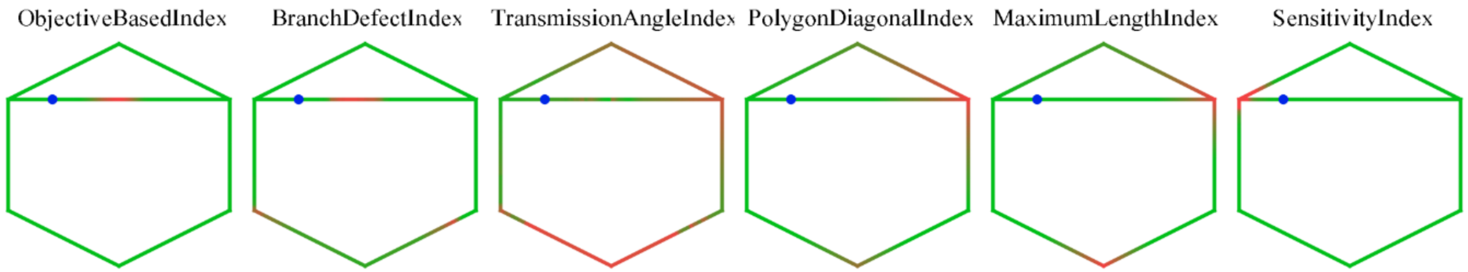
25 critical points over $\mathbb{C}$

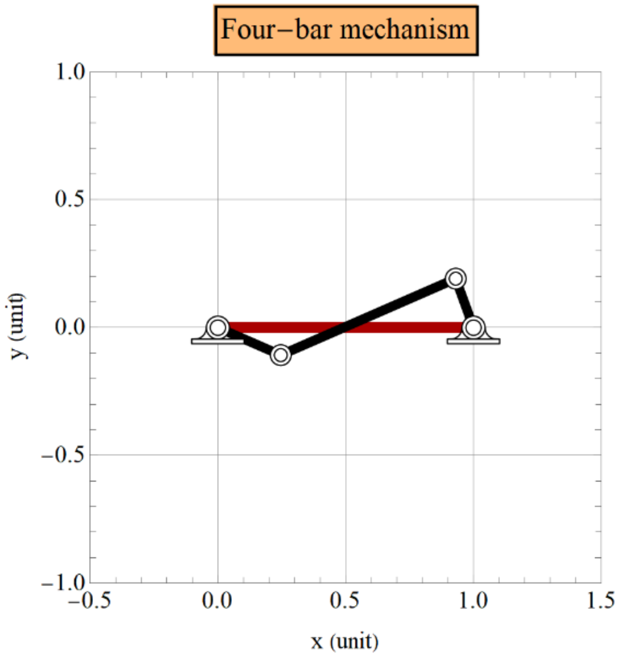
- 7 real with 3 being local minima
- "best" designs: trajectory leading to degenerate global minima

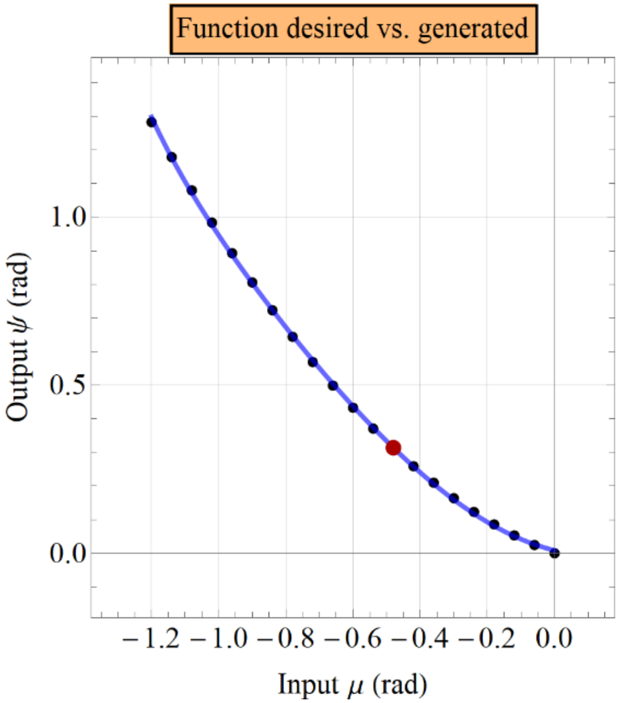


Mathematica interface using Bertini as back end solver







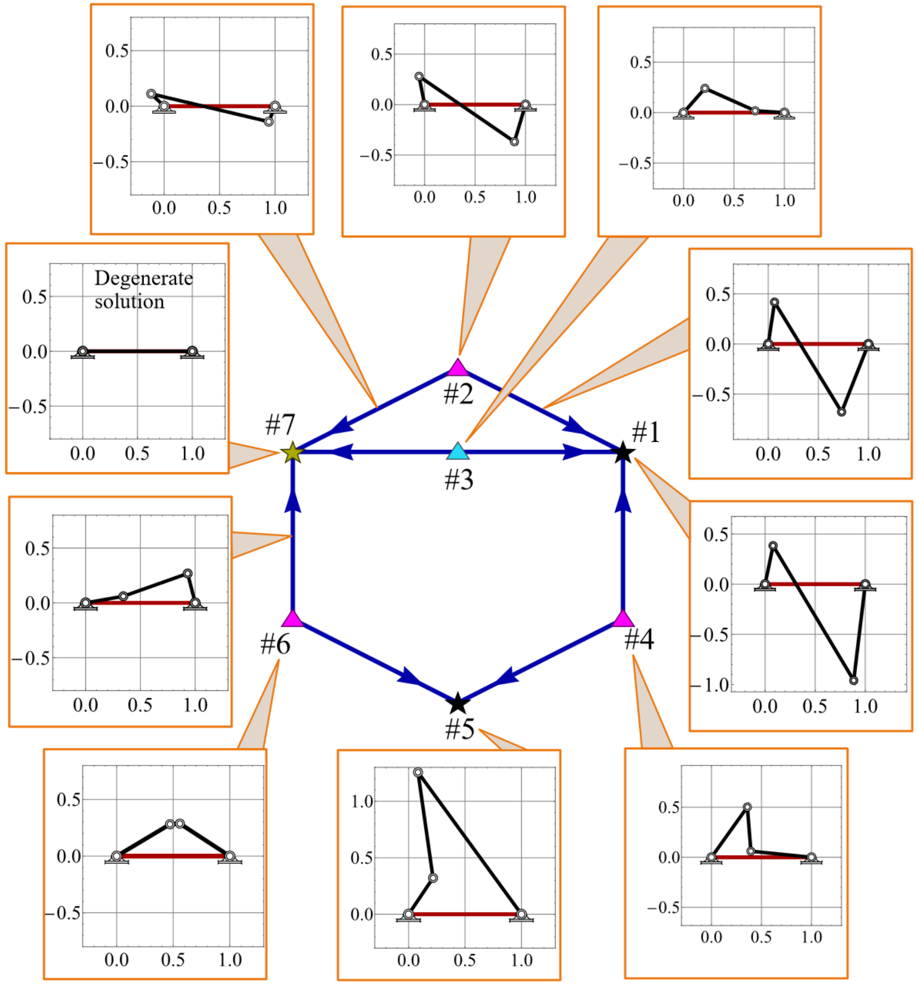


Mechanism Synthesis

Mechanism Synthesis

Saddle graph provides pictorial representation of relationship

- lacks proximity context



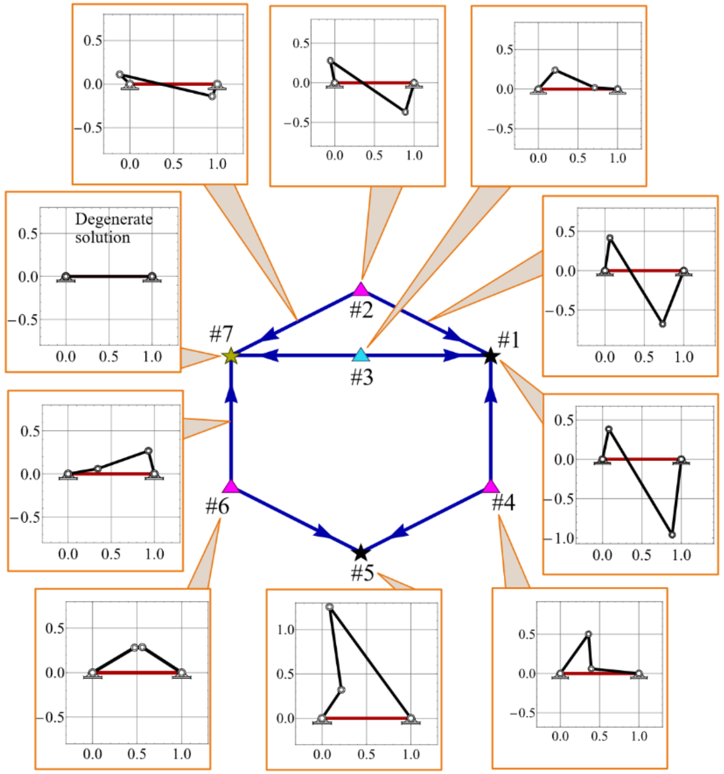
Mechanism Synthesis

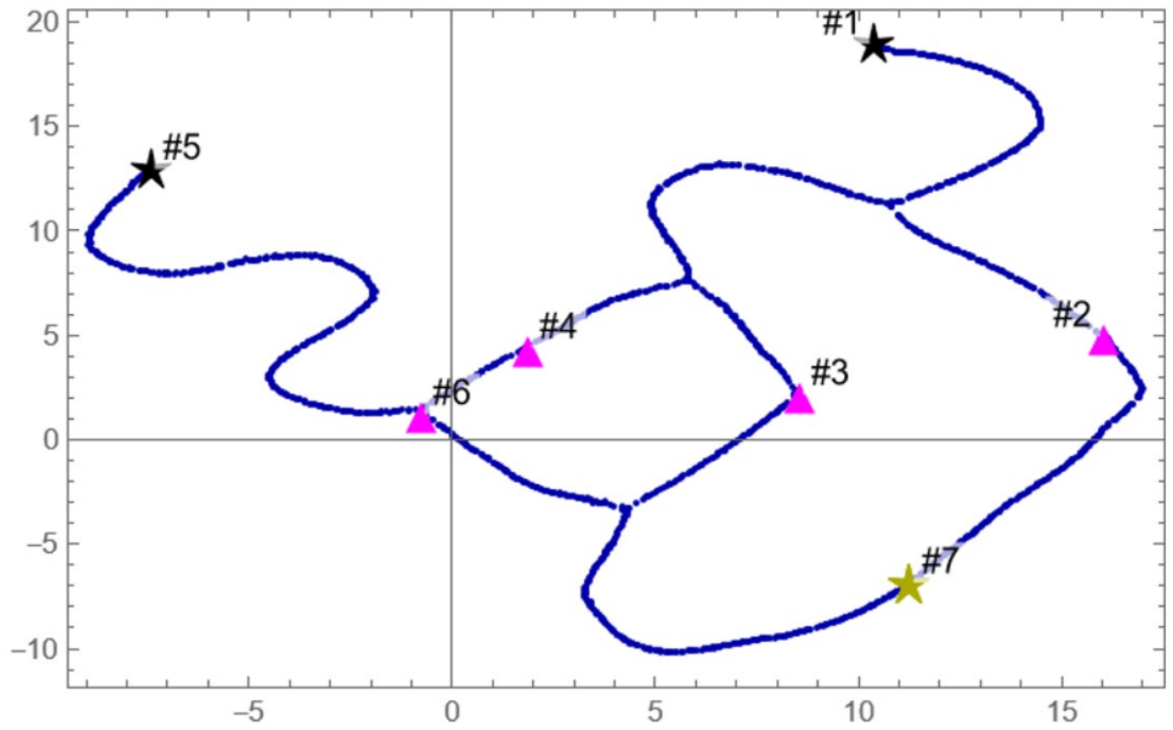
Saddle graph provides pictorial representation of relationship

- lacks proximity context

Project 1-dim'l curves into $\mathbb{R}^2$ using UMAP

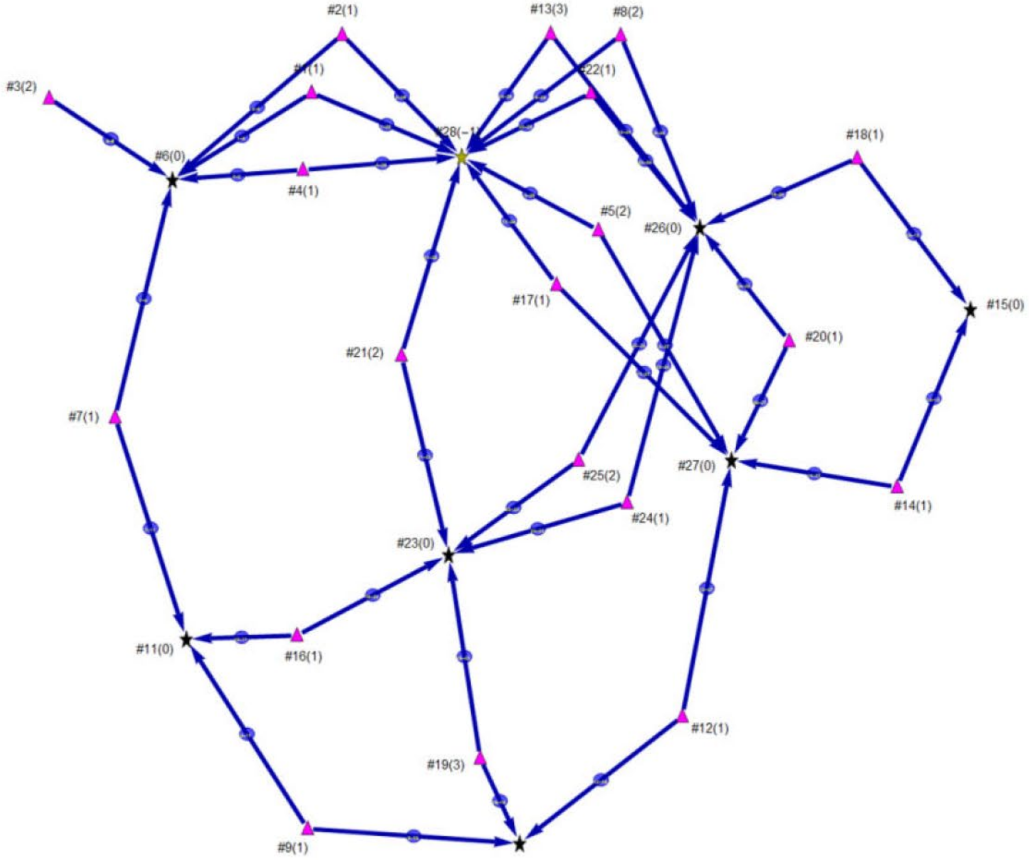
- Uniform Manifold Approximation and Projection

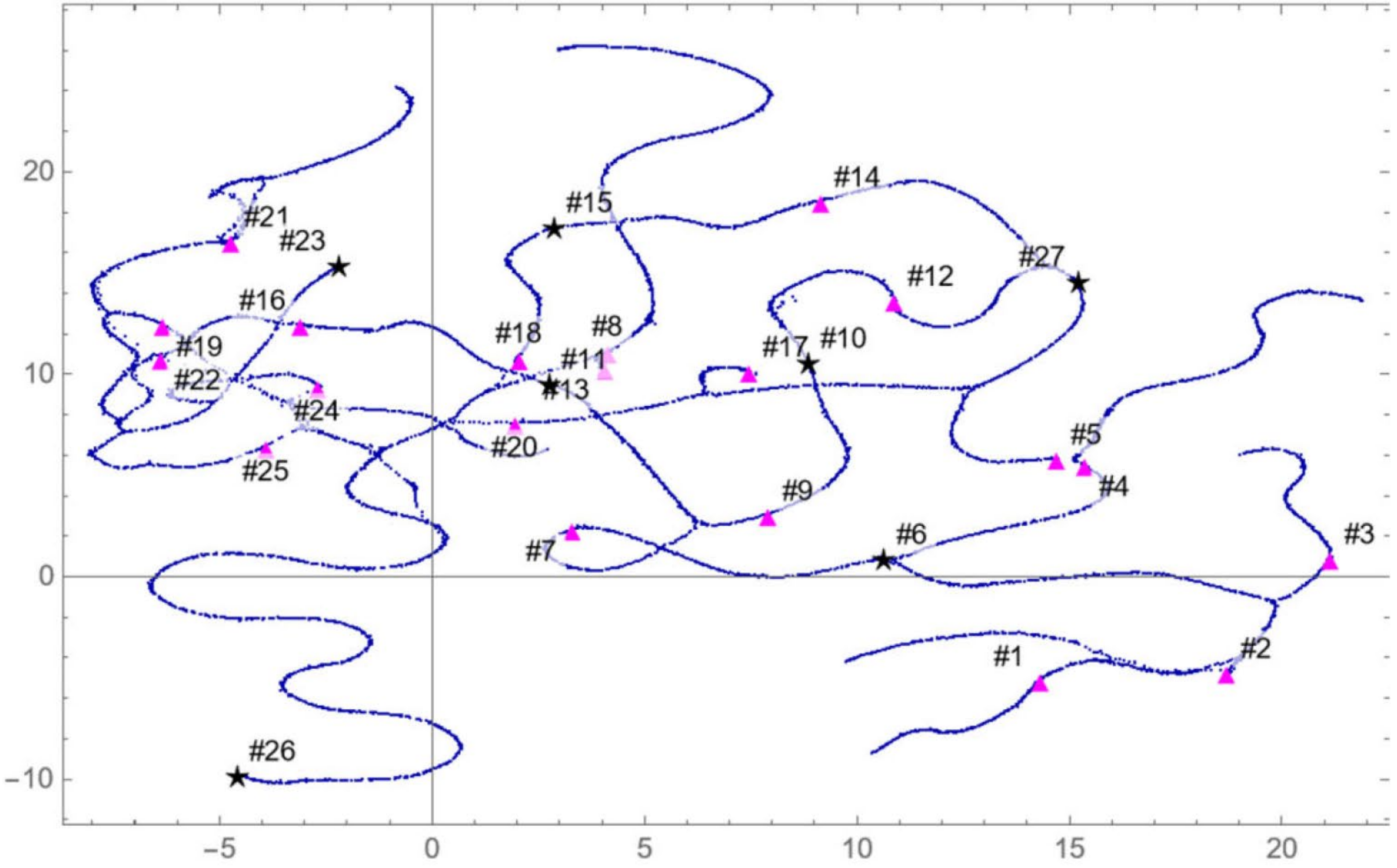




ynamical systems and Semi-algebraic
eometry: interactions with Optimization and
Deep Learning, Conference 2023

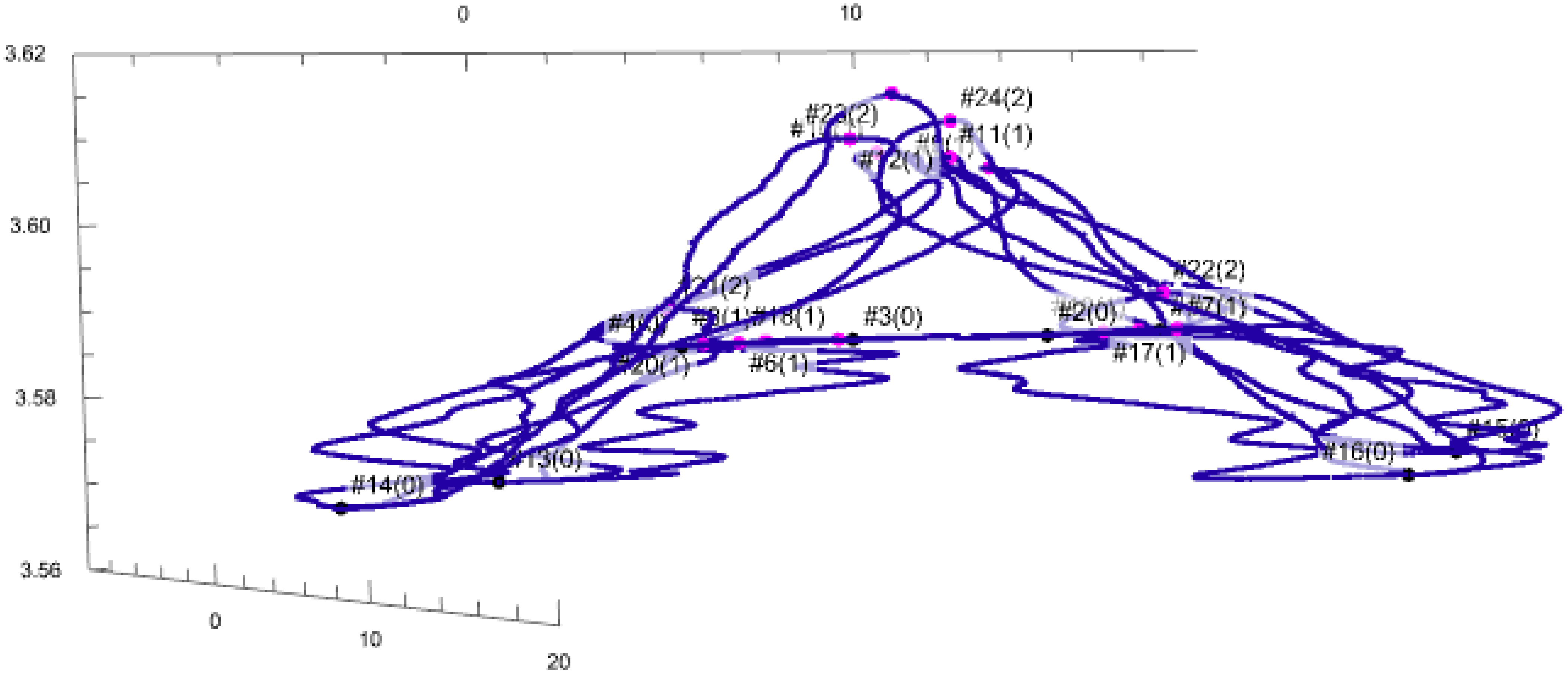
Mechanism Synthesis





Mechanism Synthesis

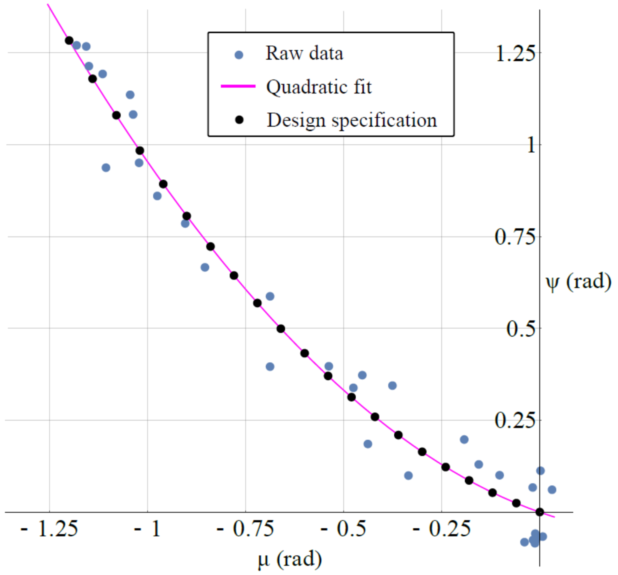
Plot in 3D using objective value to "see" landscape.

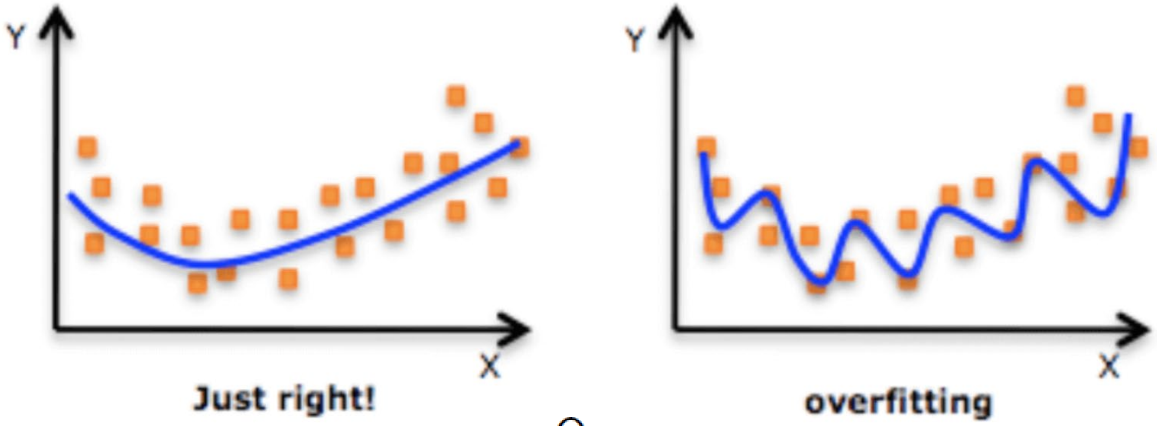


Regularization

Common approach in presence of noisy data is to regularize

- avoid "overfitting" of data





Quora

Regularization

Common approach in presence of noisy data is to regularize

- avoid "overfitting" of data
- avoid ill-conditioning in optimization (full rank Hessian)

$$\min_{x \in \mathbb{R}^n} \left(f(x) + \lambda \cdot R(x) \right)$$

- $f(x)$: original objective function
- $R(x)$: regularization function
- λ: regularization parameter

Tikhonov regularization (ridge regression): $R(x) = \|x\|_2^2$

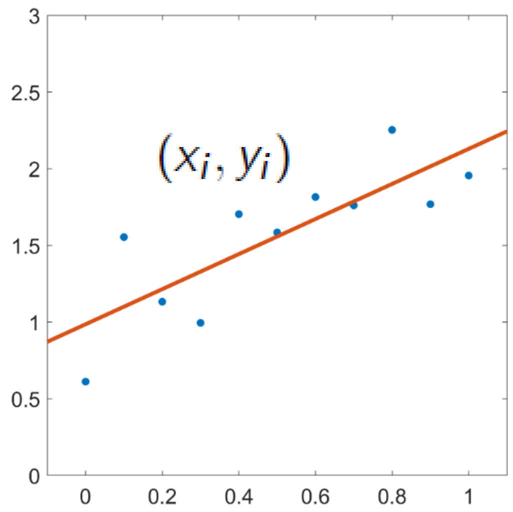
Regularization

What does regularization due to the energy landscape?

- Deep linear networks: Mehta-Chen-Tang-H. (2021)

$$L = \frac{1}{2} \|W_{n+1} \cdot W_n \cdots W_1 X - Y\|_F^2 + \lambda \cdot R(W_1, \ldots, W_{n+1})$$

- Known data matrices X and Y



$$X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_k \end{bmatrix}$$

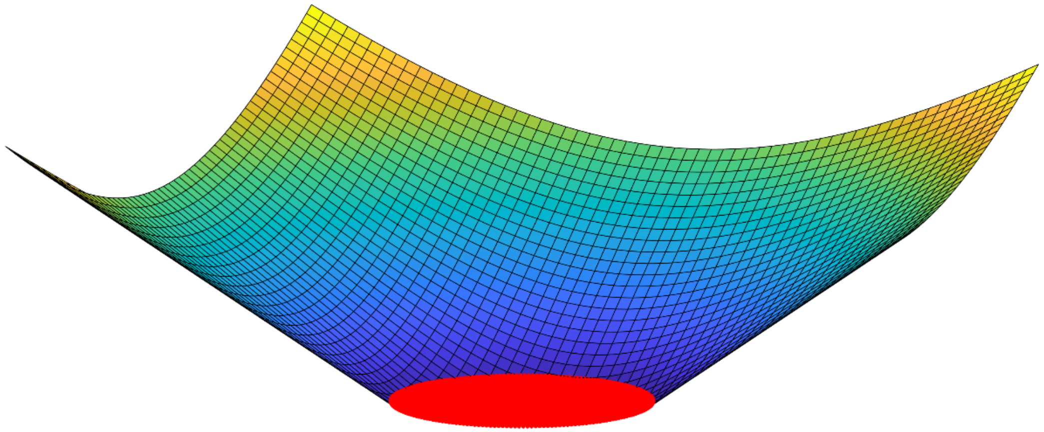
$$Y = \begin{bmatrix} y_1 & \cdots & y_k \end{bmatrix}$$

Regularization

$$L = \frac{1}{2} \|W_{n+1} \cdot W_n \cdots W_1 X - Y\|_F^2 + \lambda \cdot R(W_1, \ldots, W_{n+1})$$

No regularization: $\lambda = 0$

- Analytic expressions of critical points (Zhou-Liang, 2017)
- "No bad minima" (every local minimum is a global minimum)
- "Lakes" of minima from overparameterization (flat minima)
  - replace (W_1, W_2) by (AW_1, W_2A^{-1}) for any invertible A



Regularization

$$L = \frac{1}{2}\|W_{n+1} \cdot W_n \cdots W_1 X - Y\|_F^2 + \lambda \cdot R(W_1, \dots, W_{n+1})$$

With Tikhonov regularization: $R(W_1, \dots, W_{n+1}) = \sum_i \|W_i\|_F^2$

- flat minima can still exist

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & 2 & 3 \\ 1 & -3 & 2 \end{bmatrix}$$

$$W_1 = \begin{bmatrix} a & a \\ \gamma(a,\lambda) & \gamma(a,\lambda) \end{bmatrix} \quad W_2 = \sqrt{\tfrac{2}{197}} \begin{bmatrix} 14a & 14\gamma(a,\lambda) \\ a & \gamma(a,\lambda) \end{bmatrix}$$

$$\gamma(a,\lambda) = \sqrt{\sqrt{394}/56 - a^2 - \lambda/28}$$

Regularization

$$L = \frac{1}{2}\|W_{n+1}\cdot W_n\cdots W_1X - Y\|_F^2 + \lambda\cdot R(W_1,\ldots,W_{n+1})$$

Generalized Tikhonov regularization:

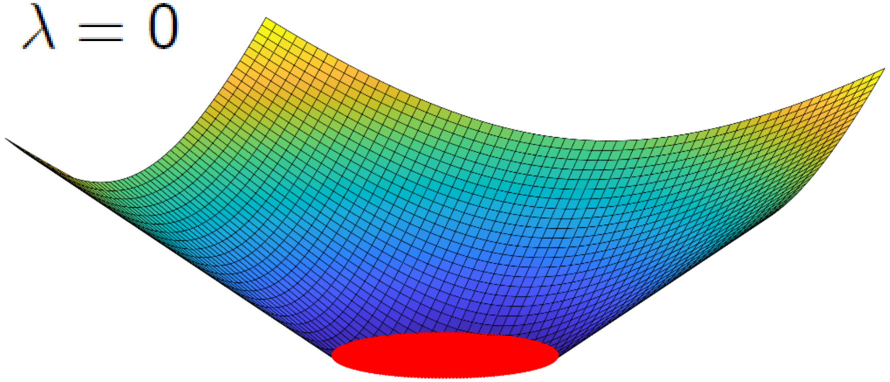
$$R(W_1,\ldots,W_{n+1}) = \sum_i \|B_i\star W_i\|_F^2$$

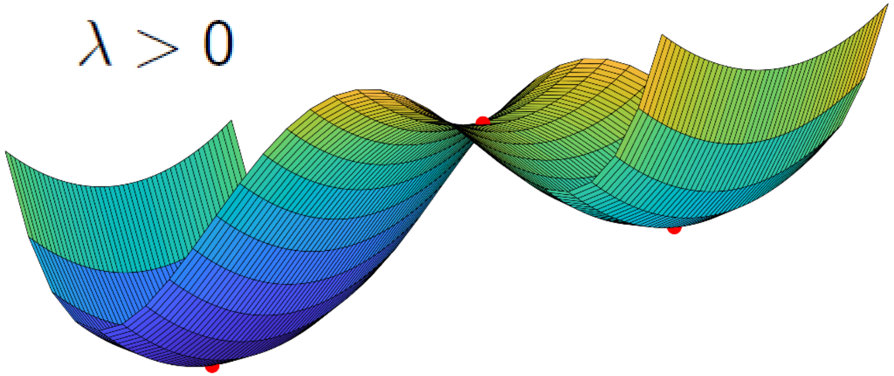
- Weight matrices B_i
- $\star$: Hadamard (entrywise) product

Theorem

All critical points where the matrices W_i have all nonzero entries are isolated and nondegenerate (full rank Hessian) for general weight matrices using a generalized Tikhonov regularization.

Regularization





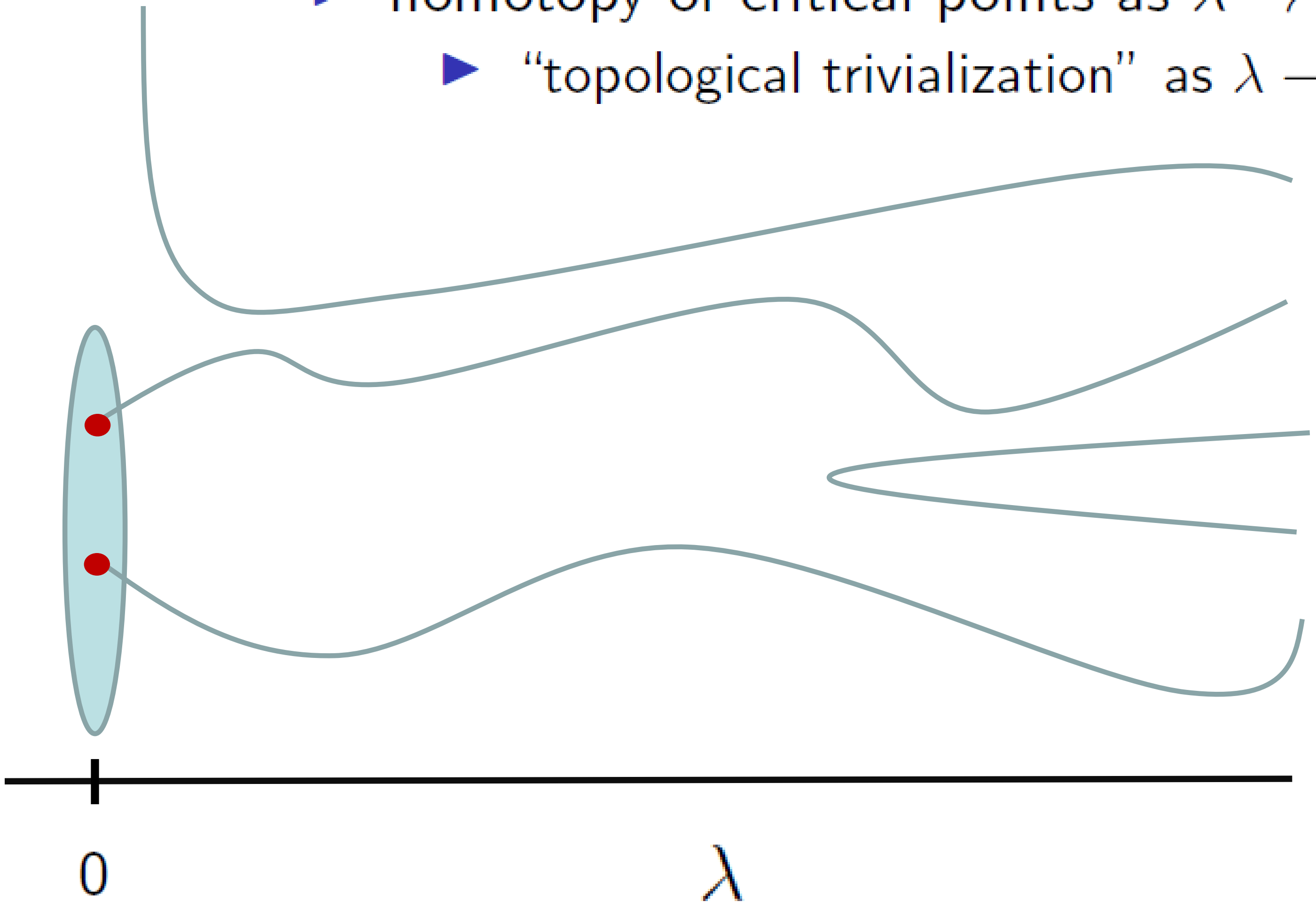
Regularization is a double-edged sword:

+ avoid “overfitting” of data
+ avoid ill-conditioning in optimization (full rank Hessian)
– energy landscape becomes more complicated
  - topological trivialization as $\lambda \to 0^+$
  - non-global local minimizers can (and do!) appear

Regularization

$$L = \frac{1}{2}\|W_{n+1}\cdot W_n\cdots W_1X - Y\|_F^2 + \lambda\cdot R(W_1,\ldots,W_{n+1})$$

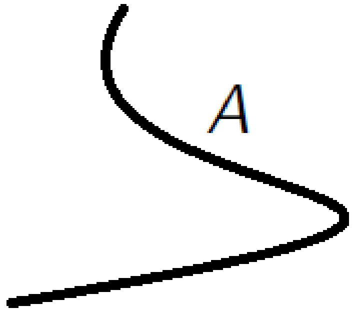
- homotopy of critical points as $\lambda \to 0^+$
  - "topological trivialization" as $\lambda \to 0^+$



Witness Set

How to represent a positive-dimensional variety on a computer?

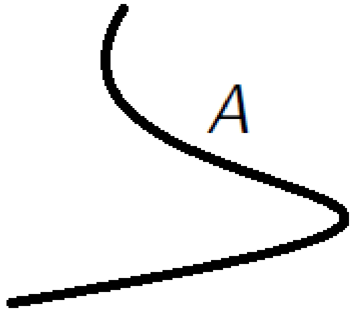
- algebraic: basis for defining ideal
- geometric: witness set



Witness Set

How to represent a positive-dimensional variety on a computer?

- algebraic: basis for defining ideal
- geometric: witness set



Defining ideal: $I(A) = \{g \ : \ g(a) = 0 \text{ for all } a \in A\}$

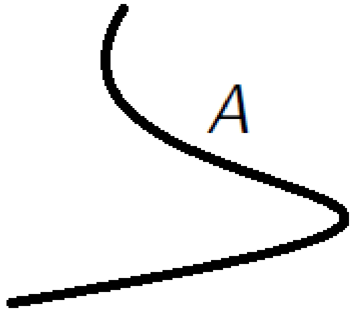
- Hilbert Basis Theorem (1890): there exists $f_1, \ldots, f_k$ such that

$$I(A) = \langle f_1, \ldots, f_k \rangle$$

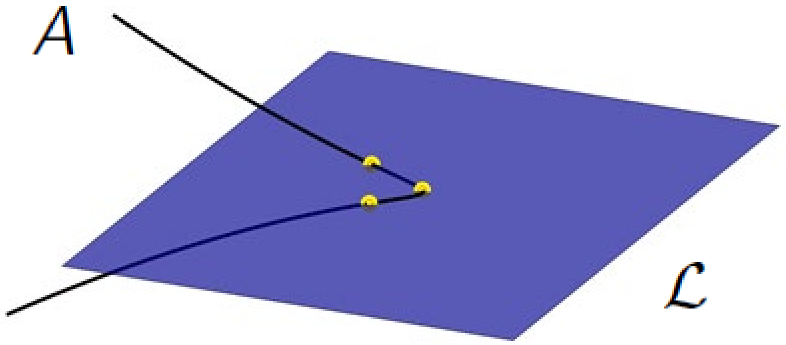
Witness Set

How to represent a positive-dimensional variety on a computer?

- algebraic: basis for defining ideal
- geometric: witness set

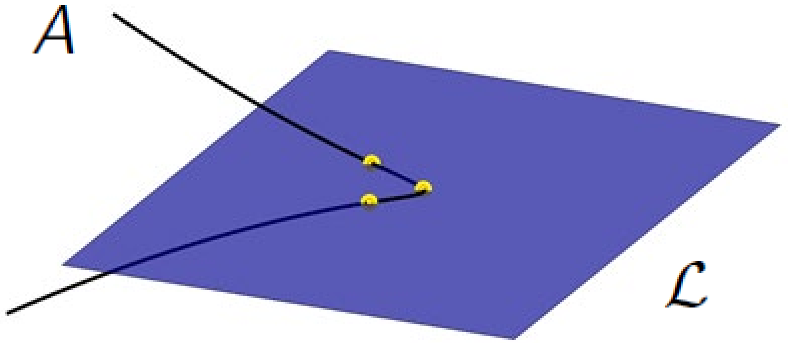


Intersect with complimentary dimensional linear space



Witness Set

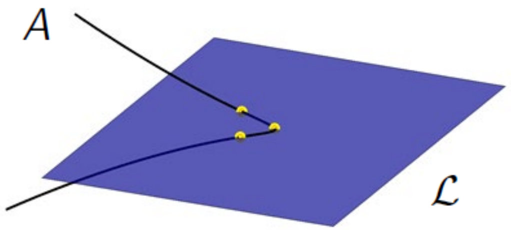
Intersect with complimentary dimensional linear space



- geometric: witness set $\{f, \mathcal{L}, W\}$ where
  - f is polynomial system where A is component of variety of f
  - $\mathcal{L}$ is a general linear space with $\text{codim } \mathcal{L} = \dim A$
  - $W = A \cap \mathcal{L}$ is witness point set with $\#W = \deg A$

Witness Set

Witness sets "localize" computations to component.



Some applications of witness sets:

- compute sample points on A by moving $\mathcal{L}$
- monodromy to compute $W = A \cap \mathcal{L}$ given one $w \in W$
- membership testing in A
- compute images $\overline{\pi(A)}$
- compute intersections $A \cap B$
  - regeneration
- determine arithmetically Cohen-Macaulayness of A
- compute all irreducible components of a variety
  - numerical irreducible decomposition

Fixed Points

Analyze fixed points of a map using numerical algebraic geometry

$$\underbrace{F \circ \cdots \circ F}_{N\text{-times}}(x) = x \qquad F(x_1, \ldots, x_4) = \begin{bmatrix} x_2 \\ -x_4 \\ x_1 - x_1 x_2^2 \\ -x_3 + x_1 x_2 x_4 \end{bmatrix}$$

Joint work with

- Cinzia Bisi (University of Ferrara, Italy)
- Tuyen Trung Truong (University of Oslo, Norway)

Fixed Points

$$\underbrace{F \circ \cdots \circ F}_{k\text{-times}}(x) = x$$

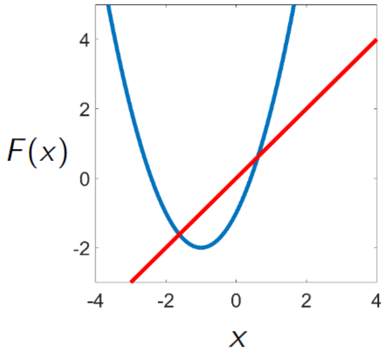
$$F(x_1, \ldots, x_4) = \begin{bmatrix} x_2 \\ -x_4 \\ x_1 - x_1 x_2^2 \\ -x_3 + x_1 x_2 x_4 \end{bmatrix}$$

$$G_N = \{(x, y_1, \ldots, y_N) \mid F(x) = y_1, F(y_1) = y_2, \ldots, F(y_{N-1}) = y_N\}$$

- graph $\rightarrow$ irreducible of dim 4
- $G_{N+1} = (G_N \times \mathbb{C}^4) \cap \{(x, y_1, \ldots, y_N, y_{N+1}) \mid F(y_N) = y_{N+1}\}$

Compute fixed points by intersecting with linear space:

$$G_N \cap \{(x, y_1, \ldots, y_N) \mid x = y_N\}$$



Fixed Points

N	fixed points on general fiber	$[\sharp IsoFix_N(f_c)]^{1/N}$
1	4	**4**
2	C (occurring with multiplicity 1)	0
3	10	**2.15443469003**
4	D_1 (multiplicity 1) & D_2 (multiplicity 2)	0
5	44	**2.13152551327**
6	C (multiplicity 1) AND 12 points	1.51308574942
7	186	**2.10967780991**
8	D_1 (multiplicity 1) & D_2 (multiplicity 2) AND 128 points	1.83400808641
9	820	**2.10744910267**
10	C (multiplicity 1) AND 1440 points	2.06936094886
11	3634	**2.10703309279**
12	D_1 (multiplicity 1) & D_2 (multiplicity 2) AND 6908 points	2.08903649661

$C \subset Z_c$: the curve defined by the ideal $< x_2 - x_1^2 x_2 - x_3, x_1 + x_4, x_1x_4 - x_2x_3 - c >$.

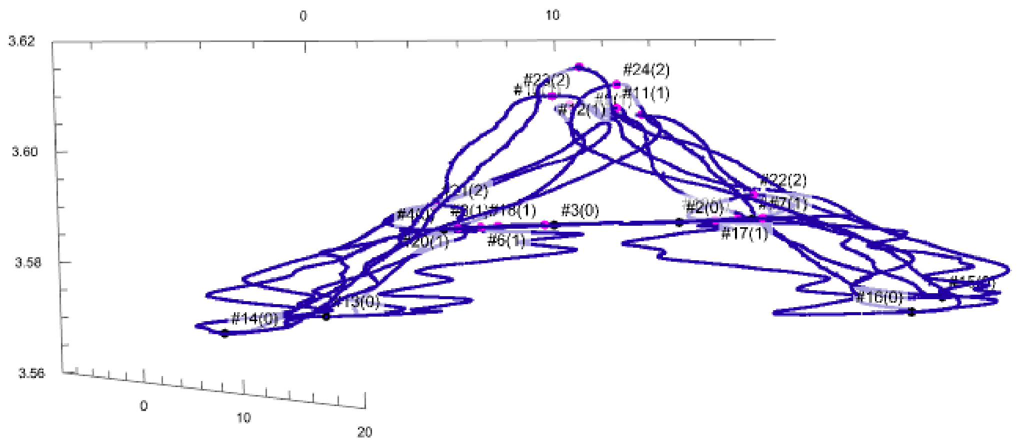
$D_1 \subset Z_c$: the curve with 2 components defined by the ideals $< x_2 - x_1^2x_2 - x_3, x_1 + x_4, x_1x_4 - x_2x_3 - c >$ and $< -x_2 + x_1^2x_2 - x_3, x_1 - x_4, x_1x_4 - x_2x_3 - c >$;

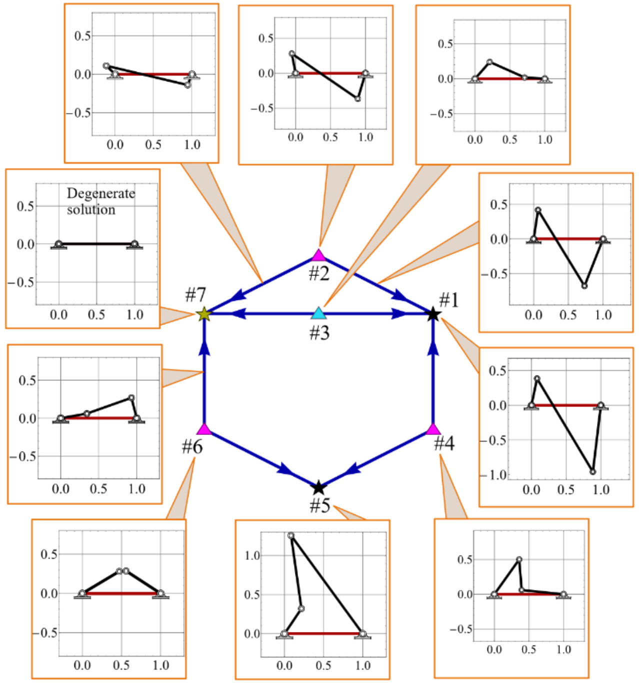
$D_2 \subset Z_c$: the curve with 2 components defined by the ideals $< x_2, x_3, x_1x_4 - x_2x_3 - c >$ and $< x_1, x_4, x_1x_4 - x_2x_3 - c >$.

Conclusion

Solve polynomial systems using numerical algebraic geometry:

- analyze landscapes in optimization and machine learning
  - saddle graphs connect fixed points via gradient descent paths





- analyze maps
  - compute fixed points
  - compute dynamical degrees

Dynamical systems and Semi-algebraic geometry: interactions with Optimization and Deep Learning, Conference 2023

Postdoc position available now

Optimization, Machine Learning, and Algebraic Geometry

https://www.mathjobs.org/jobs/list/22544

hauenstein@nd.edu

Thank You!