Global Optimization over Compact Domains

Seminaire Pascaline, LIP, Lyon

Georgy Scholten

MPI-CBG Dresden

Joint work with Mohab Safey El Din and Emmanuel Trélat
Ongoing work with Alexander Demin

Partially supported by Grant FA8665-20-1-7029 of the EOARD-AFOSR

October 24, 2025

MAX PLANCK INSTITUTE
OF MOLECULAR CELL BIOLOGY
AND GENETICS

CENTER FOR
SYSTEMS BIOLOGY
DRESDEN

The Global Optimization Problem

The Global Optimization Problem

Problem Statement:

- Let $f : \mathcal{C} \longrightarrow \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$

The Global Optimization Problem

Problem Statement:

- Let $f : \mathcal{C} \longrightarrow \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Goal:** Find all local minimizers of f on $\mathcal{C}$

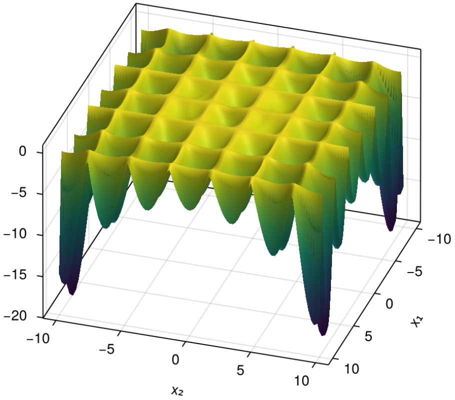
$$\text{Find all } x^\star \in \mathcal{C} \text{ s.t. } f(x^\star) \leq f(x) \text{ for all } x \text{ near } x^\star$$

Motivational Example: The Hölder Table Function

$$f(x, y) = -\left|(\sin(x)\cos(y)\exp(1 - \sqrt{x^2 + y^2}/\pi))\right|$$

Motivational Example: The Hölder Table Function

$$f(x,y) = -|(\sin(x)\cos(y)\exp(1-\sqrt{x^2+y^2}/\pi))|$$

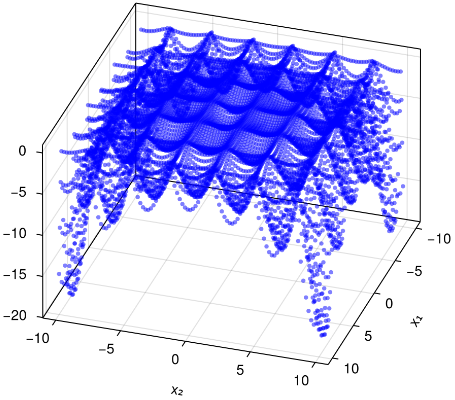


Motivational Example: The Hölder Table Function

$$f(x, y) = -|(\sin(x)\cos(y)\exp(1 - \sqrt{x^2 + y^2}/\pi))|$$

Motivational Example: The Hölder Table Function

$$f(x, y) = -|(\sin(x)\cos(y)\exp(1 - \sqrt{x^2 + y^2}/\pi))|$$



Critical Points Capture of The Hölder Table Function

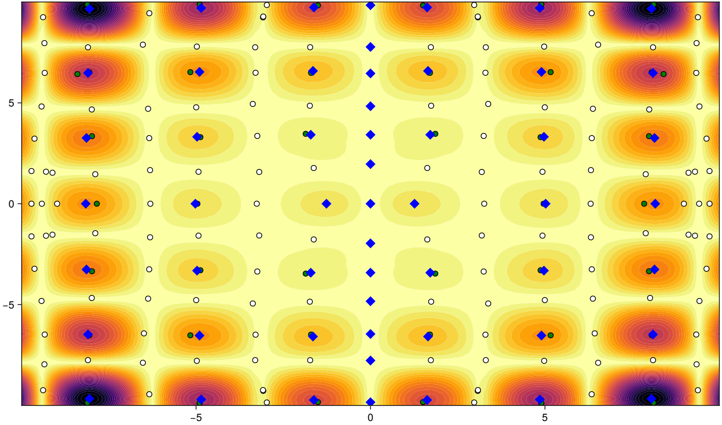


Figure: Level sets and critical points of $w_{18,\mathcal{S}}$

The Global Optimization Problem (Practical Setting)

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available
  - No derivatives can be computed directly

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available
  - No derivatives can be computed directly
- **Goal:** Find all local minimizers of f on $\mathcal{C}$ ×

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available
  - No derivatives can be computed directly
- **Goal:** Find all local minimizers of f on $\mathcal{C}$ ×

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available
  - No derivatives can be computed directly
- **Goal:** Find all local minimizers of f on $\mathcal{C}$ $\times$

We need more assumptions on f:

- Let spec(f) denote eigenvalues of the Hessian of f at all local minimizers in interior of $\mathcal{C}$

The Global Optimization Problem (Practical Setting)

Updated Problem Statement:

- Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function on a compact domain $\mathcal{C} = [-1, 1]^n$
- **Constraint:** The function f is only accessible through an *evaluation program*
  - We can evaluate $f(x)$ at any point $x \in \mathcal{C}$
  - No analytical formula is available
  - No derivatives can be computed directly
- **Goal:** Find all local minimizers of f on $\mathcal{C}$ ×

We need more assumptions on f:

- Let spec(f) denote eigenvalues of the Hessian of f at all local minimizers in interior of $\mathcal{C}$
- For $\lambda > 0$, $\rho > 0$ define

$$\mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C}) = \{f \in \mathscr{M}^m(\mathcal{C}) \mid \min \operatorname{spec}(f) \geq \lambda, \max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho, j \leq m\}$$

where $\mathscr{M}^m(\mathcal{C})$ = m-times diff. Morse functions on $\mathcal{C}$

Overall Methodology: Three-Step Process

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
    - generate a tensorized grid of evaluation points from Chebyshev measure
    - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
    - ✓ Good numerical stability over rectangular domains

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions
2. **Algebraic Step**

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions
2. **Algebraic Step**
   - Compute $\mathrm{crit}(w)$ using algebraic methods

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions
2. **Algebraic Step**
   - Compute $\mathrm{crit}(w)$ using algebraic methods
   - Ensure all critical points are captured

Overall Methodology: Three-Step Process

To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

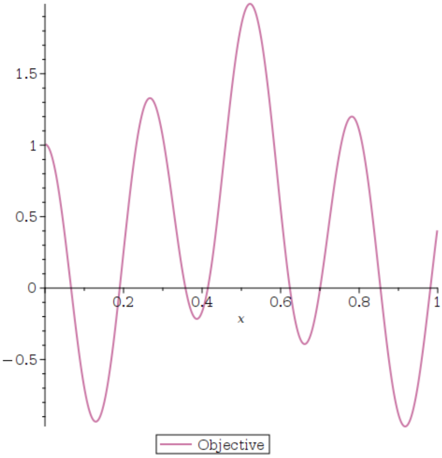
1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions
2. **Algebraic Step**
   - Compute $\mathrm{crit}(w)$ using algebraic methods
   - Ensure all critical points are captured
3. **Local Optimization**

Overall Methodology: Three-Step Process

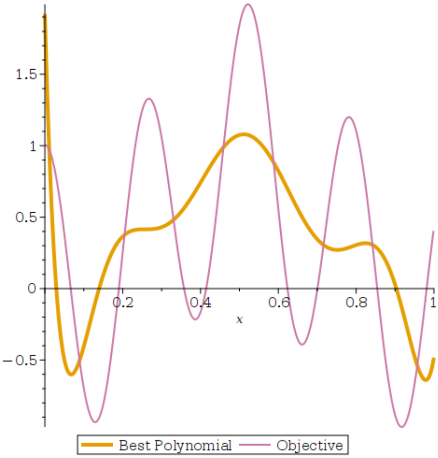
To compute all local minimizers of f on $\mathcal{C}$, we employ a three-step approach:

1. **Polynomial Approximation Step**
   - generate a tensorized grid of evaluation points from Chebyshev measure
   - Create polynomial approximant $w_{d,\mathcal{S}}$ using discrete least-squares
   - ✓ Good numerical stability over rectangular domains
   - ✓ Control quality via regularity assumptions
2. **Algebraic Step**
   - Compute $\mathrm{crit}(w)$ using algebraic methods
   - Ensure all critical points are captured
3. **Local Optimization**
   - Initiate local optimization methods at the computed points

L^∞ vs L^2 Polynomial Approximant

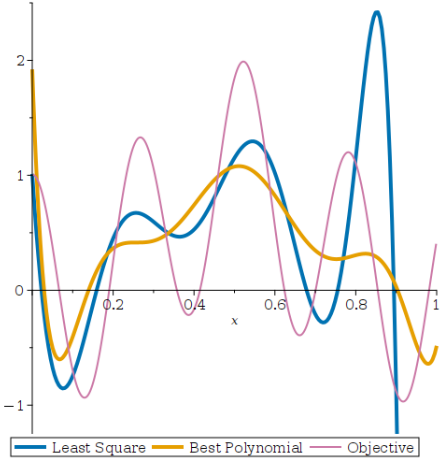


L^∞ vs L^2 Polynomial Approximant



$$\| f \|_{\mathcal{L}^\infty} = \max_{x \in \mathcal{C}} |f(x)|. \quad (1)$$

L^∞ vs L^2 Polynomial Approximant



$$\| f \|_{\mathcal{L}^2} = \left(\int_{\mathcal{C}} f(x)^2 \mathrm{d}\mu \right)^{\frac{1}{2}} . \tag{2}$$

Chebyshev Polynomials

[1] Keaton J. Burns et al. "Dedalus: A flexible framework for numerical simulations with spectral methods". *Phys. Rev. Res.* 2 (2 Apr. 2020), p. 023068. DOI: 10.1103/PhysRevResearch.2.023068. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.023068.

Chebyshev Polynomials

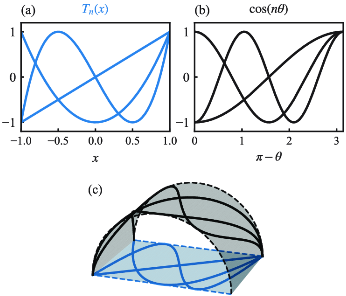


Figure: Dedalus[1]

[1] Keaton J. Burns et al. "Dedalus: A flexible framework for numerical simulations with spectral methods". *Phys. Rev. Res.* 2 (2 Apr. 2020), p. 023068. DOI: 10.1103/PhysRevResearch.2.023068. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.023068.

Chebyshev Polynomials

Chebyshev Polynomials

- Orthogonality:

$$\int_{\mathcal{C}} T_m(x) T_n(x) \mathrm{d}\mu = \gamma_n \delta_{mn}$$

where $\gamma_n > 0$

[a] Zaïneb Bel-Afia, Chiara Meroni, and Simon Telen. *Chebyshev Varieties*. 2024. arXiv: 2401.12140 [math.AG]. URL: https://arxiv.org/abs/2401.12140.

Chebyshev Polynomials

- Orthogonality:

$$\int_{\mathcal{C}} T_m(x)\, T_n(x) \mathrm{d}\mu = \gamma_n \delta_{mn}$$

  where $\gamma_n > 0$
- Well-suited for solving Least Squares

[a] Zaïneb Bel-Afia, Chiara Meroni, and Simon Telen. *Chebyshev Varieties*. 2024. arXiv: 2401.12140 [math.AG]. URL: https://arxiv.org/abs/2401.12140.

Chebyshev Polynomials

- Orthogonality:

$$\int_{\mathcal{C}} T_m(x) T_n(x) \mathrm{d}\mu = \gamma_n \delta_{mn}$$

  where $\gamma_n > 0$

- Well-suited for solving Least Squares
- Chebyshev Varieties[a]:

$$x_1^{\nu_1} \dots x_n^{\nu_n} \to T_{\nu_1}(x_1) \dots T_{\nu_n}(x_n)$$

[a] Zaïneb Bel-Afia, Chiara Meroni, and Simon Telen. *Chebyshev Varieties*. 2024. arXiv: 2401.12140 [math.AG]. URL: https://arxiv.org/abs/2401.12140.

Chebyshev Polynomials

- Orthogonality:

$$\int_{\mathcal{C}} T_m(x)T_n(x)\mathrm{d}\mu = \gamma_n \delta_{mn}$$

  where $\gamma_n > 0$
- Well-suited for solving Least Squares
- Chebyshev Varieties[a]:

$$x_1^{\nu_1}\dots x_n^{\nu_n} \to T_{\nu_1}(x_1)\dots T_{\nu_n}(x_n)$$

[a] Zaïneb Bel-Afia, Chiara Meroni, and Simon Telen. *Chebyshev Varieties*. 2024. arXiv: 2401.12140 [math.AG]. URL: https://arxiv.org/abs/2401.12140.

Chebyshev Polynomials

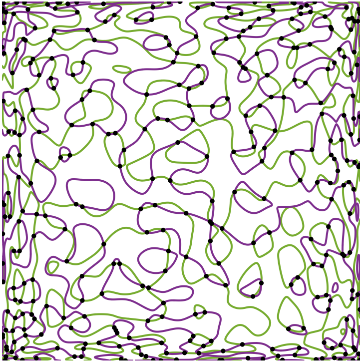
- Orthogonality:

$$\int_{\mathcal{C}} T_m(x)\,T_n(x)\mathrm{d}\mu = \gamma_n \delta_{mn}$$

where $\gamma_n > 0$

- Well-suited for solving Least Squares
- Chebyshev Varieties[a]:

$$x_1^{\nu_1}\dots x_n^{\nu_n} \longrightarrow T_{\nu_1}(x_1)\dots T_{\nu_n}(x_n)$$



[a] Zaïneb Bel-Afia, Chiara Meroni, and Simon Telen. *Chebyshev Varieties*. 2024. arXiv: 2401.12140 [math.AG]. URL: https://arxiv.org/abs/2401.12140.

Jackson's Approximation Theorem

Jackson's Approximation Theorem

ρ-Regularity

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$, meaning:

$$\max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho \quad \text{for all } j \leq m$$

Jackson's Approximation Theorem

ρ-Regularity

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$, meaning:

$$\max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho \quad \text{for all } j \leq m$$

Jackson's Theorem

There exists a constant $C_{n,m} > 0$ such that:

$$\min_{p \in \mathscr{P}_{n,d}} \max_{x \in \mathcal{C}} |f(x) - p(x)| \leq C_{n,m} \frac{\rho}{d^{m-1}}$$

Jackson's Approximation Theorem

ρ-Regularity

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$, meaning:

$$\max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho \quad \text{for all } j \leq m$$

Jackson's Theorem

There exists a constant $C_{n,m} > 0$ such that:

$$\min_{p \in \mathscr{P}_{n,d}} \max_{x \in \mathcal{C}} |f(x) - p(x)| \leq C_{n,m} \frac{\rho}{d^{m-1}}$$

- Error decreases as $\mathcal{O}(d^{-(m-1)})$ with polynomial degree d

Jackson's Approximation Theorem

ρ-Regularity

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$, meaning:

$$\max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho \quad \text{for all } j \leq m$$

Jackson's Theorem

There exists a constant $C_{n,m} > 0$ such that:

$$\min_{p \in \mathscr{P}_{n,d}} \max_{x \in \mathcal{C}} |f(x) - p(x)| \leq C_{n,m} \frac{\rho}{d^{m-1}}$$

- Error decreases as $\mathcal{O}(d^{-(m-1)})$ with polynomial degree d
- Higher smoothness $m \Rightarrow$ faster convergence rate

Jackson's Approximation Theorem

ρ-Regularity

Let $f \in \mathscr{C}^{m,\rho}_{\lambda}(\mathcal{C})$, meaning:

$$\max_{x \in \mathcal{C}} \|d^j f(x)\| \leq \rho \quad \text{for all } j \leq m$$

Jackson's Theorem

There exists a constant $C_{n,m} > 0$ such that:

$$\min_{p \in \mathscr{P}_{n,d}} \max_{x \in \mathcal{C}} |f(x) - p(x)| \leq C_{n,m} \frac{\rho}{d^{m-1}}$$

- Error decreases as $\mathcal{O}(d^{-(m-1)})$ with polynomial degree d
- Higher smoothness $m \Rightarrow$ faster convergence rate
- Constant $C_{n,m}$ depends on dimension n and smoothness m

Main Theorem [Safey El Din, S., Trélat]

Main Theorem [Safey El Din, S., Trélat]

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$ with $m \geq \max(3, \beta n + 1)$, where $\beta = \frac{\ln(3)}{2\ln(2)}$.
Let $r > 0$ satisfy $\rho r \leq 3\lambda$, $\alpha \in (0,1)$ and $C_{n,m}$ the Jackson constant.

Main Theorem [Safey El Din, S., Trélat]

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$ with $m \geq \max(3, \beta n + 1)$, where $\beta = \frac{\ln(3)}{2\ln(2)}$.
Let $r > 0$ satisfy $\rho r \leq 3\lambda$, $\alpha \in (0,1)$ and $C_{n,m}$ the Jackson constant.
Given any $\delta \in (0,1)$, set

$$A_1 = 16\left(\frac{1}{\pi^n} + \frac{2}{1-\delta}\right)\frac{C_{n,m}^2\rho^2}{(\pi e)^n} \quad \text{and} \quad A_2 = \frac{16(1+\delta)}{(\pi e)^n(1-\delta)^2}\left(8\delta^2(1-\delta)^4 + 4\right).$$

Main Theorem [Safey El Din, S., Trélat]

Let $f \in \mathscr{C}_\lambda^{m,\rho}(\mathcal{C})$ with $m \geq \max(3, \beta n + 1)$, where $\beta = \frac{\ln(3)}{2\ln(2)}$.
Let $r > 0$ satisfy $\rho r \leq 3\lambda$, $\alpha \in (0,1)$ and $C_{n,m}$ the Jackson constant.
Given any $\delta \in (0,1)$, set

$$A_1 = 16\left(\frac{1}{\pi^n} + \frac{2}{1-\delta}\right)\frac{C_{n,m}^2\rho^2}{(\pi e)^n} \quad \text{and} \quad A_2 = \frac{16(1+\delta)}{(\pi e)^n(1-\delta)^2}\left(8\delta^2(1-\delta)^4 + 4\right).$$

If $d \in \mathbb{N}$ satisfies the **degree condition**:

$$\left(\frac{A_1}{d^{2m-2}} + A_2\varepsilon^2\right) d^{2\beta n} \leq \lambda^2 r^4,$$

Main Theorem [Safey El Din, S., Trélat]

Let $f \in \mathscr{C}_{\lambda}^{m,\rho}(\mathcal{C})$ with $m \geq \max(3, \beta n + 1)$, where $\beta = \frac{\ln(3)}{2\ln(2)}$.
Let $r > 0$ satisfy $\rho r \leq 3\lambda$, $\alpha \in (0,1)$ and $C_{n,m}$ the Jackson constant.
Given any $\delta \in (0,1)$, set

$$A_1 = 16\left(\frac{1}{\pi^n} + \frac{2}{1-\delta}\right)\frac{C_{n,m}^2\rho^2}{(\pi e)^n} \quad \text{and} \quad A_2 = \frac{16(1+\delta)}{(\pi e)^n(1-\delta)^2}\left(8\delta^2(1-\delta)^4 + 4\right).$$

If $d \in \mathbb{N}$ satisfies the **degree condition**:

$$\left(\frac{A_1}{d^{2m-2}} + A_2\varepsilon^2\right) d^{2\beta n} \leq \lambda^2 r^4,$$

then with a large enough sample set $\mathcal{S}$ chosen from the Chebyshev measure:

the DLSP approximant $w_{d,\mathcal{S}}$ **captures all local minimizers** at precision r with probability $\geq 1 - \alpha$

Main Theorem [Safey El Din, S., Trélat]

Let $f \in \mathscr{C}^{m,\rho}_{\lambda}(\mathcal{C})$ with $m \geq \max(3, \beta n + 1)$, where $\beta = \frac{\ln(3)}{2\ln(2)}$.
Let $r > 0$ satisfy $\rho r \leq 3\lambda$, $\alpha \in (0,1)$ and $C_{n,m}$ the Jackson constant.
Given any $\delta \in (0,1)$, set

$$A_1 = 16\left(\frac{1}{\pi^n} + \frac{2}{1-\delta}\right)\frac{C_{n,m}^2\rho^2}{(\pi e)^n} \quad \text{and} \quad A_2 = \frac{16(1+\delta)}{(\pi e)^n(1-\delta)^2}\left(8\delta^2(1-\delta)^4 + 4\right).$$

If $d \in \mathbb{N}$ satisfies the **degree condition**:

$$\left(\frac{A_1}{d^{2m-2}} + A_2\varepsilon^2\right) d^{2\beta n} \leq \lambda^2 r^4,$$

then with a large enough sample set $\mathcal{S}$ chosen from the Chebyshev measure:

> the DLSP approximant $w_{d,\mathcal{S}}$ **captures all local minimizers** at precision r with probability $\geq 1 - \alpha$

Constants dependency:

- A_1: dimension n, smoothness m, regularity ρ, parameter δ
- A_2: dimension n, parameter δ

Probabilistic Least Squares Polynomials

Probabilistic Least Squares Polynomials

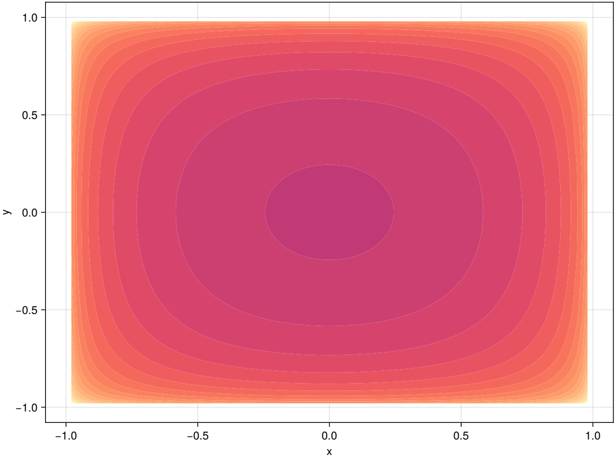


Figure: probability density of μ.

Probabilistic Least Squares Polynomials

Probabilistic Least Squares Polynomials

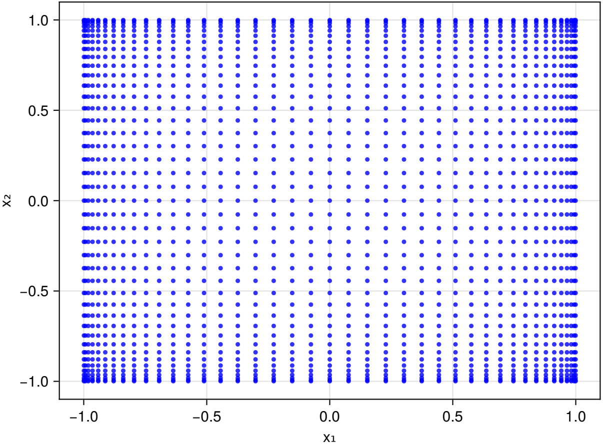


Figure: Discretization of μ.

Probabilistic Least Squares Polynomials

Probabilistic Least Squares Polynomials

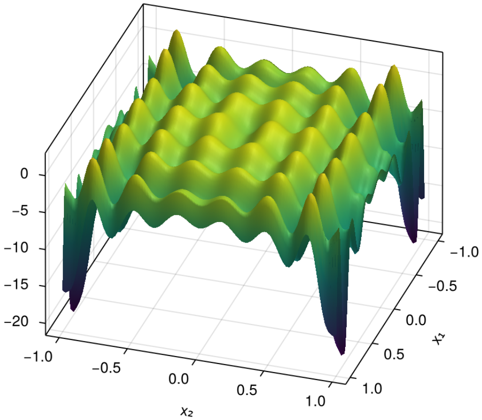


Figure: w_{18}

Probabilistic Least Squares Polynomials

Probabilistic Least Squares Polynomials

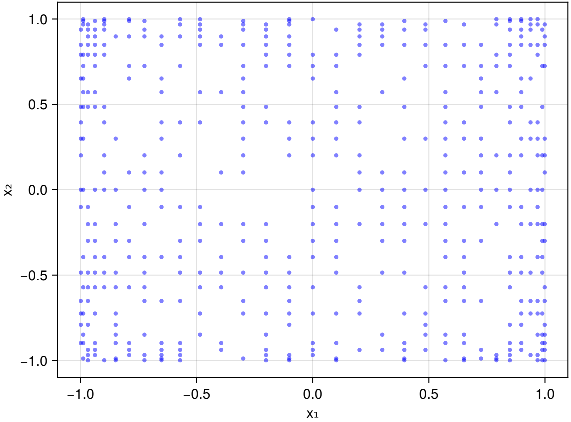


Figure: reduced sample set $\mathcal{S}$.

Probabilistic Least Squares Polynomials

Probabilistic Least Squares Polynomials

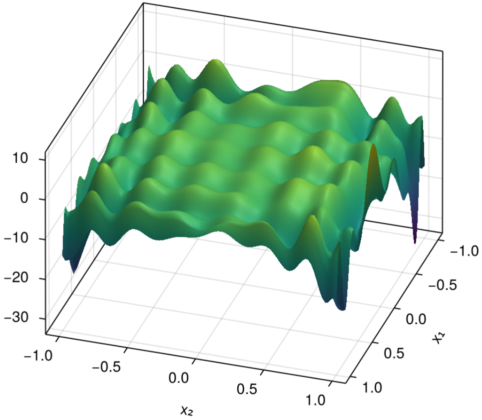


Figure: $w_{18,\mathcal{S}}$.

Probabilistic Bounds

[2] Giovanni Migliorati, Fabio Nobile, and Raúl Tempone. "Convergence estimates in probability and in expectation for discrete least-squares with noisy evaluations at random points". *Journal of Multivariate Analysis* 142 (2015), pp. 167–182. ISSN: 0047-259X. DOI: https://doi.org/10.1016/j.jmva.2015.08.009. URL: https://www.sciencedirect.com/science/article/pii/S0047259X15001931.

Probabilistic Bounds

Theorem:[2] if the number of samples k satisfies

$$\frac{\binom{n+d}{n}^{\frac{\ln(3)}{2\ln(2)}}}{\delta + (1-\delta)\ln(1-\delta)} \leq \frac{k}{\ln(k) + \ln(6\alpha^{-1})}, \tag{3}$$

then with probability at least $1-\alpha$, we have

$$\| w_{d,\mathcal{S}} - f \|^2_{\mathcal{L}^2} \leq \left(1 + \frac{C(\delta)}{\alpha \ln(6k\alpha^{-1})}\right) e_d(f)^2, \tag{4}$$

where

$$C(\delta) = \frac{4\left(\delta + (1-\delta)\ln(1-\delta)\right)}{(1-\delta)^2}, \tag{5}$$

[2] Giovanni Migliorati, Fabio Nobile, and Raúl Tempone. "Convergence estimates in probability and in expectation for discrete least-squares with noisy evaluations at random points". *Journal of Multivariate Analysis* 142 (2015), pp. 167–182. ISSN: 0047-259X. DOI: https://doi.org/10.1016/j.jmva.2015.08.009. URL: https://www.sciencedirect.com/science/article/pii/S0047259X15001931.

Practical Considerations

- Converting to standard monomial basis
- Computing partial derivatives
- Computing critical points:
  1. `Msolve`

Practical Considerations

- Converting to standard monomial basis
- Computing partial derivatives
- Computing critical points:
    1. `Msolve`
    2. `HomotopyContinuation.jl`

Practical Considerations

- Converting to standard monomial basis
- Computing partial derivatives
- Computing critical points:
  1. `Msolve`
  2. `HomotopyContinuation.jl`
  3. Bertini Real

Practical Considerations

- Converting to standard monomial basis
- Computing partial derivatives
- Computing critical points:
  1. `Msolve`
  2. `HomotopyContinuation.jl`
  3. Bertini Real
  4. `HypersurfaceRegions.jl` (?)

Convergence in L^2

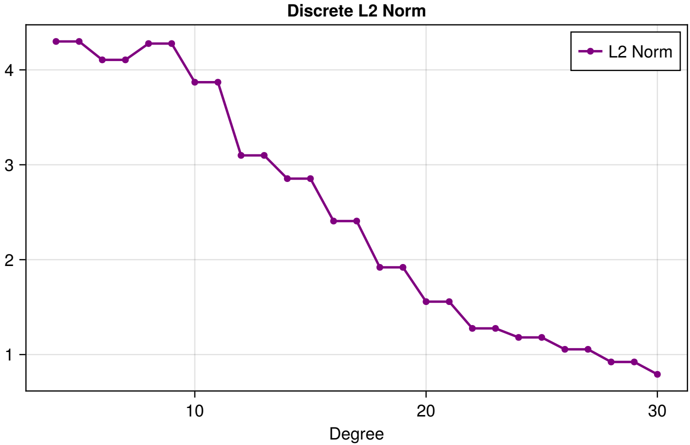


Figure: Error in $\mathcal{L}^2$-norm of the approximant

Convergence of Local Minimizers

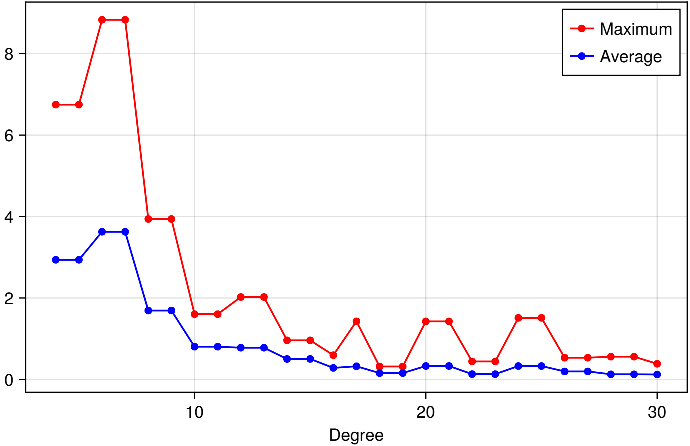


Figure: Distance from local minimizers to nearest critical point.

Tolerance to Noise

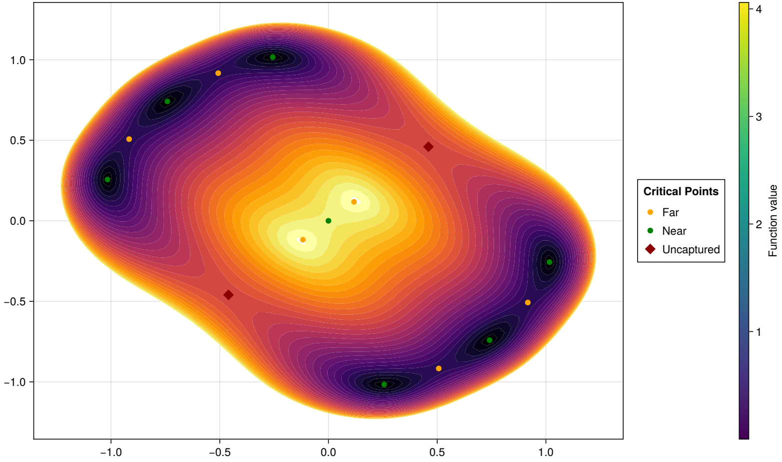


Figure: Deuflhard function

Tolerance to Noise

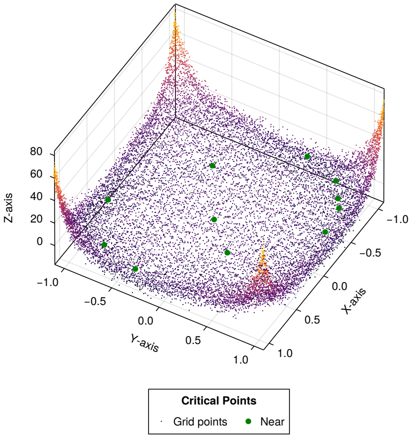


Figure: Noisy Deuflhard function

Tolerance to Noise

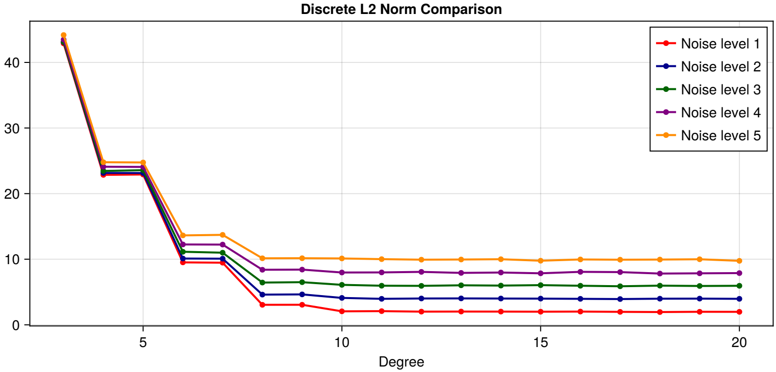


Figure: Noisy Deuflhard L2-norm approximation

Tolerance to Noise

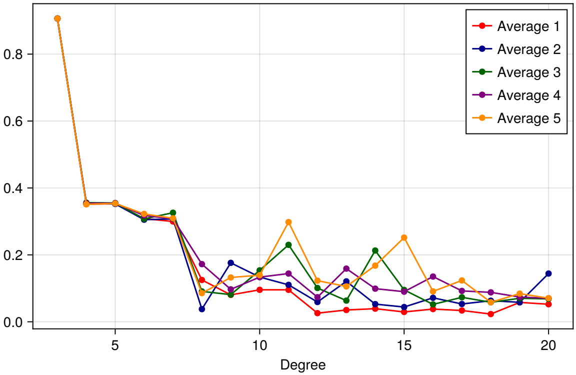


Figure: Noisy Deuflhard convergence of critical points

3D Function Example

$$f(x, y, z) = e^{\sin(50x_1)} + \sin(60e^{x_2})\sin(60x_3) + \sin(70\sin(x_1))\cos(10x_3) + \sin(\sin(80x_2)) - \sin(10(x_1 + x_3)) + \frac{x_1^2 + x_2^2 + x_3^2}{4} \quad (6)$$

ODE System: Parameter Estimation Problem

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) \ = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) \ = \boldsymbol{x_0}, \end{cases}$$

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) \ = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) = \boldsymbol{x_0}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) \ = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) \ = \boldsymbol{x_0}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) \;\; = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) = \boldsymbol{x_0}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)
- $\boldsymbol{\mu}$: scalar parameters (to be estimated)

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) = \boldsymbol{x}_{\boldsymbol{0}}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)
- $\boldsymbol{\mu}$: scalar parameters (to be estimated)
- $\boldsymbol{u}(t)$: input variables (control)

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) \;= \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) = \boldsymbol{x}_{\boldsymbol{0}}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)
- $\boldsymbol{\mu}$: scalar parameters (to be estimated)
- $\boldsymbol{u}(t)$: input variables (control)
- $\boldsymbol{f}$ and $\boldsymbol{g}$: rational functions in $\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{u}(t)$

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) = \boldsymbol{x}_{\boldsymbol{0}}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)
- $\boldsymbol{\mu}$: scalar parameters (to be estimated)
- $\boldsymbol{u}(t)$: input variables (control)
- $\boldsymbol{f}$ and $\boldsymbol{g}$: rational functions in $\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{u}(t)$

ODE System: Parameter Estimation Problem

Consider a system of ODEs in the state-space form:

$$\Sigma := \begin{cases} \boldsymbol{x}'(t) &= \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{y}(t) &= \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\mu}), \\ \boldsymbol{x}(0) &= \boldsymbol{x_0}, \end{cases}$$

where:

- $\boldsymbol{x}(t)$: state variables
- $\boldsymbol{y}(t)$: output variables (observables)
- $\boldsymbol{\mu}$: scalar parameters (to be estimated)
- $\boldsymbol{u}(t)$: input variables (control)
- $\boldsymbol{f}$ and $\boldsymbol{g}$: rational functions in $\boldsymbol{x}(t), \boldsymbol{\mu}, \boldsymbol{u}(t)$

Given data: $D = ((t_1, \boldsymbol{y}_1), \ldots, (t_n, \boldsymbol{y}_n))$, where $\boldsymbol{y}_i$ is measured at time t_i

Constructing the Objective Function

Constructing the Objective Function

Let $\boldsymbol{\mu}$ be the parameters to be estimated.

Constructing the Objective Function

Let $\boldsymbol{\mu}$ be the parameters to be estimated.
Take time points $\{t_1, t_2, \ldots, t_N\}$ and let $\boldsymbol{\mu} \in \mathbb{R}^n$

$$f(\boldsymbol{\mu}) = \sum_{i=1}^{N} \left(\boldsymbol{y}(t_i; \boldsymbol{\mu}) - \boldsymbol{y}_{\text{target}}(t_i) \right)^2$$

where $\boldsymbol{y}(t_i; \boldsymbol{\mu})$ is the computed response at time t_i for parameters $\boldsymbol{\mu}$, and $\boldsymbol{y}_{\text{target}}(t_i)$ is the target response.

Constructing the Objective Function

Let $\boldsymbol{\mu}$ be the parameters to be estimated.
Take time points $\{t_1, t_2, \ldots, t_N\}$ and let $\boldsymbol{\mu} \in \mathbb{R}^n$

$$f(\boldsymbol{\mu}) = \sum_{i=1}^{N} \Big(\boldsymbol{y}(t_i; \boldsymbol{\mu}) - \boldsymbol{y}_{\text{target}}(t_i) \Big)^2$$

where $\boldsymbol{y}(t_i; \boldsymbol{\mu})$ is the computed response at time t_i for parameters $\boldsymbol{\mu}$, and $\boldsymbol{y}_{\text{target}}(t_i)$ is the target response.
Take a sample set $\mathcal{S} \in \mathbb{R}^n$ and construct

$$w_d := \operatorname*{argmin}_{p \in \mathscr{P}_{n,d}} \sum_{s \in \mathcal{S}} (p(s) - f(s))^2$$

Constructing the Objective Function

Let $\boldsymbol{\mu}$ be the parameters to be estimated.
Take time points $\{t_1, t_2, \ldots, t_N\}$ and let $\boldsymbol{\mu} \in \mathbb{R}^n$

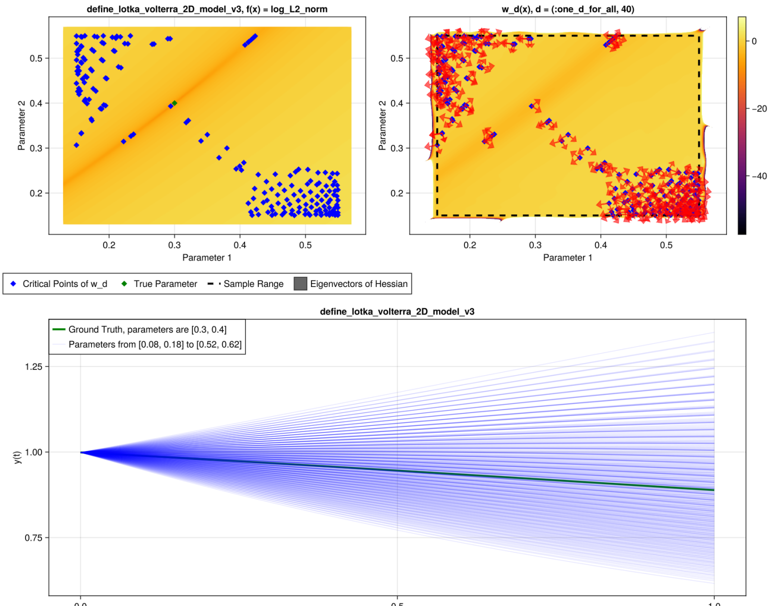
$$f(\boldsymbol{\mu}) = \sum_{i=1}^{N} \Big(\boldsymbol{y}(t_i; \boldsymbol{\mu}) - \boldsymbol{y}_{\text{target}}(t_i) \Big)^2$$

where $\boldsymbol{y}(t_i; \boldsymbol{\mu})$ is the computed response at time t_i for parameters $\boldsymbol{\mu}$, and $\boldsymbol{y}_{\text{target}}(t_i)$ is the target response.
Take a sample set $\mathcal{S} \in \mathbb{R}^n$ and construct

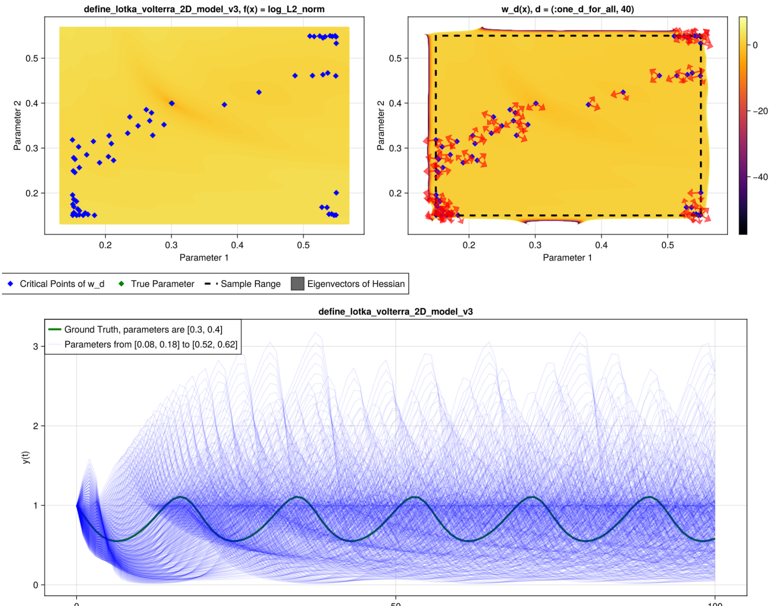
$$w_d := \underset{p \in \mathscr{P}_{n,d}}{\operatorname{argmin}} \sum_{s \in \mathcal{S}} (p(s) - f(s))^2$$

The recovered critical points serve as initial guesses for local optimization methods.

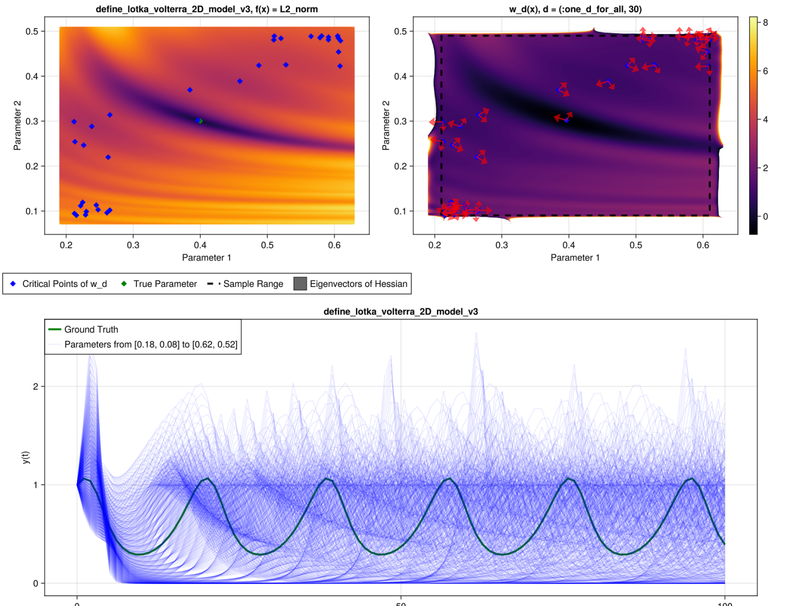
Parameter Estimation Lotka-Volterra: (t=1)



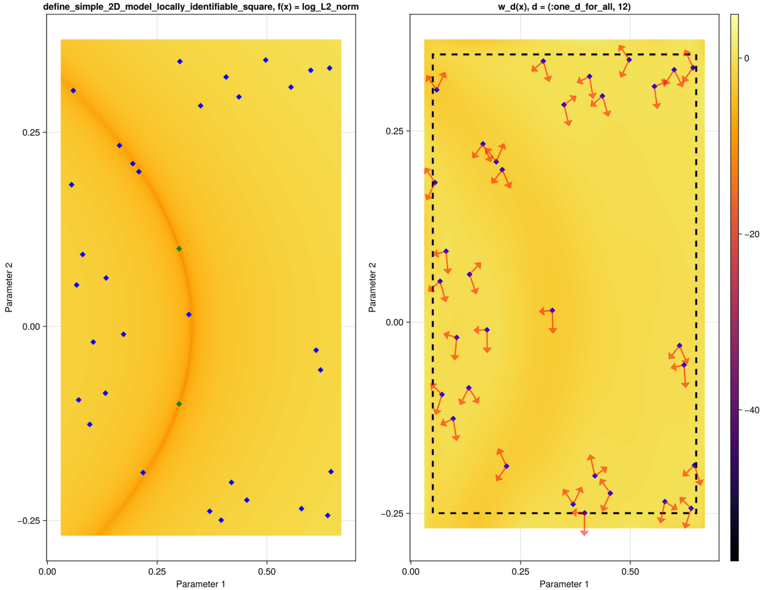
Parameter Estimation Lotka-Volterra (t=100)



Parameter Estimation Lotka-Volterra (t=100)



Locally Identifiable Model



Thank You! I

- Matthieu Dolbeault, David Krieg, and Mario Ullrich. "A sharp upper bound for sampling numbers in $\mathcal{L}^2$". *Applied and Computational Harmonic Analysis* 63 (2023), pp. 113–134. ISSN: 1063-5203. DOI: https://doi.org/10.1016/j.acha.2022.12.001. URL: https://www.sciencedirect.com/science/article/pii/S1063520322000999
- Albert Cohen and Giovanni Migliorati. "Optimal weighted least-squares methods". en. *The SMAI Journal of computational mathematics* 3 (2017), pp. 181–203. DOI: 10.5802/smai-jcm.24. URL: https://smai-jcm.centre-mersenne.org/articles/10.5802/smai-jcm.24/
- Mohab Safey El Din, Georgy Scholten, and Emmanuel Trélat. "Probabilistic algorithm for computing all local minimizers of Morse functions on a compact domain". working paper or preprint. 2025. URL: https://hal.sorbonne-universite.fr/hal-05160251