FM-GPS
An Efficient Framework for Global Non-Convex Polynomial Optimization

Pierre-David Letourneau, Principal Engineer
M. Harper Langston, Principal Engineer
Dalton Jones, Senior Engineer
Mitchell Harris, Intern
Matthew Morse, Senior Engineer

Qualcomm Technologies Inc.
pletourn@qti.qualcomm.com

Agenda

1. Very brief recapitulation of Polynomial Optimization (PO)
2. FM-GPS : An Efficient Framework for Global Non-Convex Polynomial Optimization
3. Discussion of Results and Proofs
4. Numerical Algorithms
5. Numerical Results
6. Conclusion

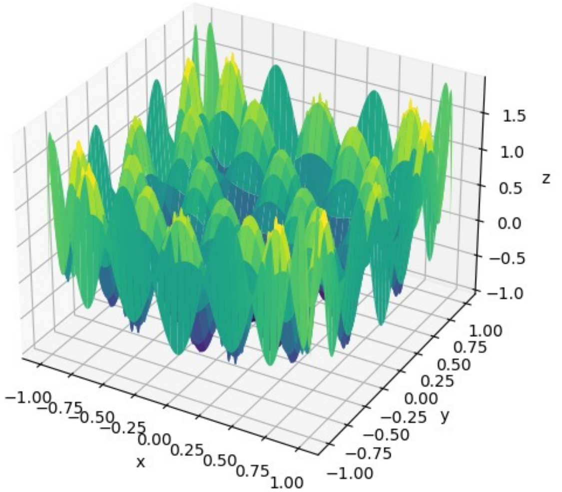
Polynomial Optimization (PO)

Problem Statement

Given a *polynomials* of degree **d** in **D** dimensions

$$p(x) = \sum_{|n|\leq d} p_n\, x^n = \sum_{|n|\leq d} p_n \left(\prod_{i=1}^{D} x_i^{n_i}\right)$$

$$g^{(k)}(x) = \sum_{|n|\leq d} g_n^{(k)}\, x^n = \sum_{|n|\leq d} g_n^{(k)} \left(\prod_{i=1}^{D} x_i^{n_i}\right)$$



Compute the *global minimum value* and find a *global minimum location* of the problem

$$\min_{x\in[-1,1]^D} \quad p(x)$$

$$\text{subject to} \quad g^{(k)}(x) \geq 0$$

Problem Statement

Fundamental Problem of Polynomial Optimization

Source sample text

Problem Statement

Why is PO Difficult?

Challenges:

1. Non-convex objective

- Multiple local minima
- Failure of descent methods

2. Non-convex / disconnected domain

- Multiple connected components
- Non-convex feasible region

3. Scale

- Expensive descriptions (many parameters/coefficients)
- Dimension

FM-GPS

An efficient framework for global non-convex optimization

FM-GPS

General Idea

Traditional Lasserre SDP relaxation

- **Convex relaxation**
  - General measures
  - Semi-definite constraints (approximation/relaxation)
  - Convex (Semi-Definite Program)

$$\int_{\mathbb{R}^D} p(x)\,\mathrm{d}\mu(x)$$

FM-GPS

- **Exact reformulation**
  - Convex combinations of product measures
  - Exact reformulation (no approximation)
  - Non-linear

$$\int_{\mathbb{R}^D} p(x)\,\mathrm{d}\left(\sum_{l=1}^{L}\prod_{i=1}^{D}\mu_i^{(l)}(x_i)\right)$$

FM-GPS

Reformulation

FM-GPS Reformulation

$$\min_{\mu \in \mathbb{R}^{(d+1)\times D \times L}} \sum_{n \in \text{supp}(p)} p_n \phi_n(\mu)$$

$$\begin{aligned}
\text{subject to } & \mathcal{M}_d(\mu_i^{(l)}) \succeq 0, \\
& \mathcal{M}_{d-1}(\mu_i^{(l)}; 1 - x_i^2) \succeq 0, \quad i = 1, ..., D, l = 1, ..., L, k = 1, ..., K \\
& \gamma^{(j)}(y_k) \cdot \phi(\mu) \geq 0 \\
& y_k \geq 0 \\
& \mu_{i,0}^{(l)} \geq 0, \\
& \phi_{(0,\ldots,0)}(\mu) = 1 \\
& \phi_n(\mu) = \sum_{l=1}^{L} \prod_{i=1}^{D} \mu_{i,n_i}^{(l)}.
\end{aligned}$$

Problem:

- Nonlinear Objective
- Semi-definite constraints
- Scalar (linear & nonlinear) constraints

Computations:

- Memory cost *linear* (O(N)) in dimension (D), degree (d) independently
- Memory cost *at most quadratic* ($O(N^2)$) in nonzero constraint coefficients
- Computational cost of each iteration *at most polynomial* in degree and dimension $o(\min(D^d, d^D))$
- *Essentially no spurious local minima*

Results and Proofs

Discussion

FM-GPS

Characterization of product measures

Source sample text

Product measures supported over compact sets can be characterized efficiently using semi-definite constraints

Proposition 1 *Let* $D, d \in \mathbb{N}$ *and* $(\mu_1, \mu_2, ..., \mu_D) \in \mathbb{R}^{(d+1)\times D}$ *be such that for each* $i = 1, ..., D$, $\mu_{i,0} = 1$, *and*

$$\mathcal{M}_d(\mu_i) \succeq 0,$$
$$\mathcal{M}_{d-1}(\mu_i;\ 1 - x_i^2) \succeq 0.$$

Then, there exists a regular Borel product measure,

$$\mu(\cdot) := \prod_{i=1}^{D} \mu_i(\cdot),$$

supported over $[-1,1]^D$ *such that* $\mu\left([-1,1]^D\right) = 1$, *and*

$$\int_{[-1,1]^D} x^n \left(\prod_{i=1}^{D} \mu_i(x_i)\right) \mathrm{d}x = \prod_{i=1}^{D} \mu_{i,n_i}$$

for all multi-index $n \in \mathbb{N}^D$ *such that* $0 \le n_i \le d$ *for all* i.

A consequence of the fact that 1D moment problem may be solved efficiently (not true in higher dimensions. See, e.g., Nie-Schweighofer),

Theorem 4 (Powers, Reznick, from Fekete, [25]) *Let* $p \in \mathbb{P}_{1,d}$ *be a 1D polynomial of degree* d *positive on* $[-1,1]$. *Then,*

$$p(x) = (f(x))^2 + (1 - x^2)\,(g(x))^2$$

for some polynomials $f(x)$ *and* $g(x)$ *of degree at most* d *and* $d-1$ *respectively.*

FM-GPS

Theoretical Underpinnings

The FM-GPS reformulation has the *same optimal value* as the original problem and *does not possess spurious local minima*

- "Every local minimum is essentially a global minimum"
- We can use local descent techniques to find a global minimum!
- **Caveat**: high-order stationary points
- Similar re-formulation/conclusions for semi-algebraic sets (*upcoming paper*)

Theorem 2 *The original problem,*

$$\min_{[-1,1]^D} \quad p(x),$$

and the re-formulated problem share the same global minimum value.

Theorem 3 *Let* $\left\{\mu_i^{(0)}, \mu_i^{(1)}\right\}_{i=1}^{D} \in \mathbb{R}^{(2d+1)\times D\times 2}$ *be a feasible point of Problem 14 corresponding to a local minimum, and let* $\alpha^{(j)} = \prod_{i=1}^{D} \mu_{i,0}^{(j)}$ *for* $j \in \{0,1\}$*. Then, such a point is a global minimum if and only if either*

$$\begin{aligned} &\alpha^{(0)} \neq 0 \quad \text{and} \quad \left\{\frac{\prod_{i=1}^{D} \mu_{i,n_i}^{(0)}}{\alpha^{(0)}}\right\}_{|n|\leq d} \quad \text{is a local minimum, or} \\ &\alpha^{(1)} \neq 0 \quad \text{and} \quad \left\{\frac{\prod_{i=1}^{D} \mu_{i,n_i}^{(1)}}{\alpha^{(1)}}\right\}_{|n|\leq d} \quad \text{is a local minimum.} \end{aligned} \tag{20}$$

FM-GPS

Equal optimal value

The original formulation and the FM-GPS re-formulation have the same global optimal value

The feasible set of the reformulated problem *corresponds to convex combinations of product measures*

$$\begin{aligned}
\min_{\mu \in \mathbb{R}^{(d+1)\times D \times L}} \quad & \sum_{n \in \mathrm{supp}(p)} p_n\, \phi_n(\mu) \\
\text{subject to} \quad & \mathcal{M}_d(\mu_i^{(l)}) \succeq 0, \\
& \mathcal{M}_{d-1}(\mu_i^{(l)}; 1 - x_i^2) \succeq 0, \\
& \gamma^{(j)}(y_k) \cdot \phi(\mu) \geq 0 \\
& y_k \geq 0 \\
& \mu_{i,0}^{(l)} \geq 0, \\
& \phi_{(0,\ldots,0)}(\mu) = 1 \\
& \phi_n(\mu) = \sum_{l=1}^{L} \prod_{i=1}^{D} \mu_{i,n_i}^{(l)}.
\end{aligned}$$

The feasible set contains Dirac deltas supported over the optimal set

$$\int_{\mathbb{R}^D} p(x)\, \mathrm{d}\left(\sum_{l=1}^{L} \prod_{i=1}^{D} \mu_i^{(l)}(x_i) \right)$$

Source sample text

FM-GPS

Local minima are global minima

For every non-optimal point of the re-formulation, there exists a non-increasing feasible path

Consider L=2,

Case 1: $\mu^{(0)}\left([-1,1]^D\right) \in \{0,1\}$

$$\pi(t) := \prod_{i=1}^{D} \pi_i^{(0)}(t) + \prod_{i=1}^{D} \pi_i^{(1)}(t)$$

$$\pi_i^{(0)}(t) := (1-t)^{1/D}\,\mu_i^{(0)}(\cdot)$$
$$\pi_i^{(1)}(t) := t^{1/D}\,\delta(\cdot - x_i^*)$$

Only one non-trivial measure. Can reach the global minimum directly

Caveat: high-order stationary point

Case 2: $\mu^{(0)}\left([-1,1]^D\right) \in (0,1)$.

$$\pi(t) := \prod_{i=1}^{D} \pi_i^{(0)}(t) + \prod_{i=1}^{D} \pi_i^{(1)}(t)$$

$$\pi_i^{(0)}(t) = (1-t)^{1/D}\;\mu_i^{(0)}(\cdot)$$
$$\pi_i^{(1)}(t) = \left(1 + t\,\frac{\alpha}{1-\alpha}\right)^{1/D}\;\mu_i^{(1)}(\cdot)$$

$$\frac{1}{1-\alpha}\int p(x)\,\mathrm{d}\mu^{(1)}(x) < \frac{1}{\alpha}\int p(x)\,\mathrm{d}\mu^{(0)}(x).$$

Two non-trivial measures. Can reach the one with lowest value

The descent path is explicit and constructed as part of the proof

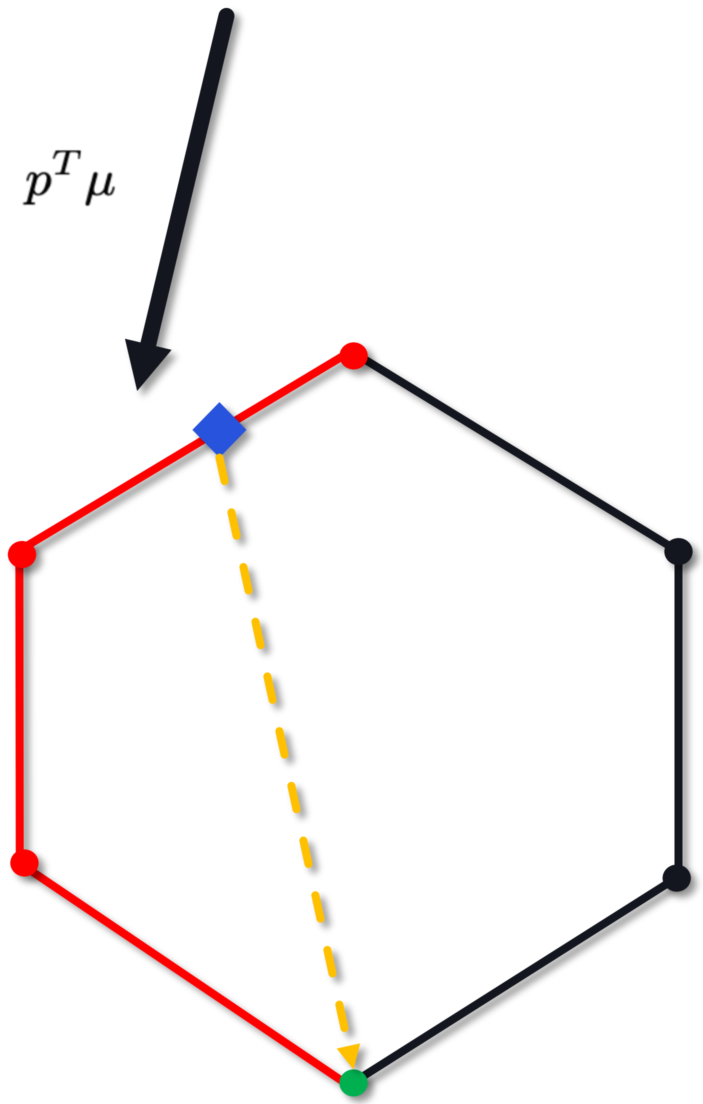
Source sample text

FM-GPS

Geometrical intuition

Descent on Boundary

- For L=2, may only descend along 1D faces.
- Non-trivial descent direction if not-orthogonal to the linear objective vector
- Can always reach a minimum by following path composed of L-D faces
- Similar to the simplex method

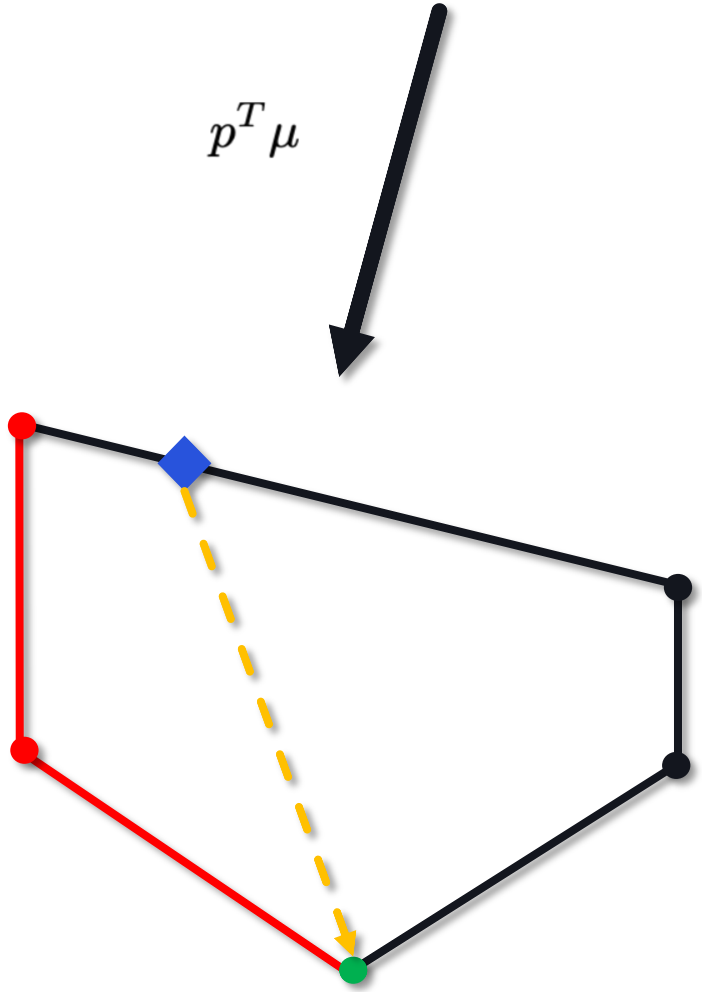


FM-GPS

Geometrical intuition

Descent on Boundary (Rare pathology)

- May only occur once
- Can be easily resolved by moving to the boundary
- Never encountered in practice



Reference

More details on Mathematical Underpinnings and Theory

Mathematics > Optimization and Control

[Submitted on 31 Aug 2023]

An Efficient Framework for Global Non-Convex Polynomial Optimization over the Hypercube

Pierre-David Letourneau, Dalton Jones, Matthew Morse, M. Harper Langston

We present a novel efficient theoretical and numerical framework for solving global non-convex polynomial optimization problems. We analytically demonstrate that such problems can be efficiently reformulated using a non-linear objective over a convex set; further, these reformulated problems possess no spurious local minima (i.e., every local minimum is a global minimum). We introduce an algorithm for solving these resulting problems using the augmented Lagrangian and the method of Burer and Monteiro. We show through numerical experiments that polynomial scaling in dimension and degree is achievable for computing the optimal value and location of previously intractable global polynomial optimization problems in high dimension.

Subjects: **Optimization and Control (math.OC)**; Mathematical Software (cs.MS); Numerical Analysis (math.NA)

Cite as: arXiv:2308.16731 [math.OC]

(or arXiv:2308.16731v1 [math.OC] for this version)

https://doi.org/10.48550/arXiv.2308.16731

https://arxiv.org/abs/2308.16731

Reference

More details on Mathematical Underpinnings and Theory

Mathematics > Optimization and Control

[Submitted on 3 Nov 2023]

An Efficient Framework for Global Non-Convex Polynomial Optimization with Nonlinear Polynomial Constraints

Mitchell Tong Harris, Pierre-David Letourneau, Dalton Jones, M. Harper Langston

We present an efficient framework for solving constrained global non-convex polynomial optimization problems. We prove the existence of an equivalent nonlinear reformulation of such problems that possesses essentially no spurious local minima. We show through numerical experiments that polynomial scaling in dimension and degree is achievable for computing the optimal value and location of previously intractable global constrained polynomial optimization problems in high dimension.

https://arxiv.org/abs/2311.02037

Numerical Algorithms

Algorithm

Overview

The FM-GPS reformulation can be solved using *any* appropriate nonlinear solver backend

We have two (2) implementation:

C/C++

- Built from scratch
- Uses only Eigen package
- Higher-performance
- More difficult to use
- Text-based API

Numerical Algorithm:

- Burer-Monteiro for semi-definite constraints
- L-BFGS/GD with line search for descent
- Explicit expressions for functions and gradients

Julia

- Uses packages:
  - `Nonconvex.jl`
  - `DynamicPolynomials.jl`
  - `Forwarddiff.jl`
- Lower-performance
- Easier to use

Numerical Algorithm:

- Burer-Monteiro for semi-definite constraints
- IPOPT (Interior Point) solver as backend
- Automatic differentiation for gradient computations

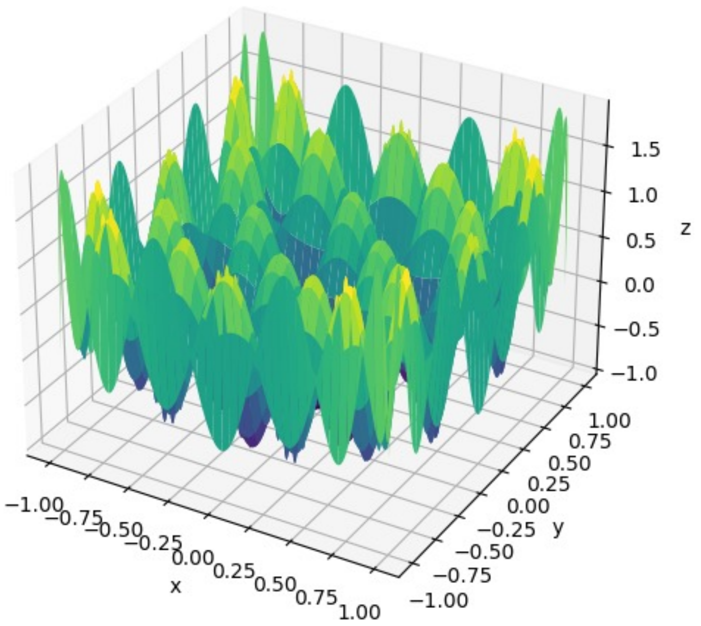
The FM-GPS reformulation can be solved using *any* appropriate nonlinear solver backend

Numerical Results

Numerical Optimization

Hypercube

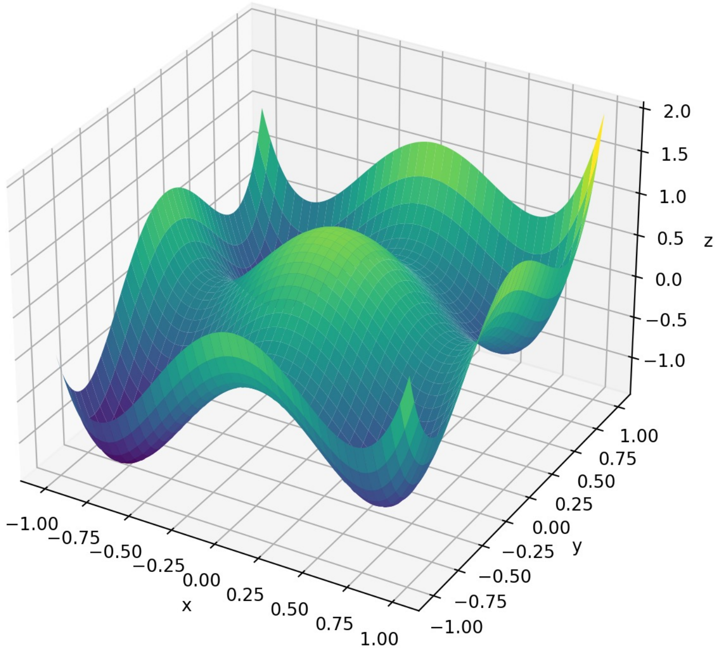
Sparse Polynomial



$$f_D(x) = \frac{1}{D}\sum_{i=1}^{D} T_2(x_i) - \prod_{i=1}^{D} T_8(x_i)$$
$$= \frac{1}{D}\sum_{i=1}^{D}(2x_i^2 - 1) - \prod_{i=1}^{D} \cos(8\,\cos^{-1}(x_i))$$

- Oscillatory
- Symmetric
- Sparse coefficients (Chebyshev)
- Multimodal (4^D local minima)

Dense Polynomial



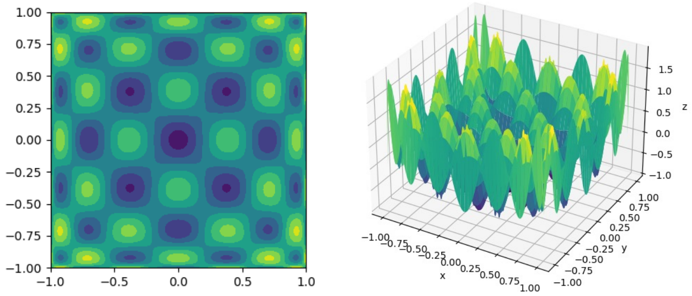
$$g_D(x) = \frac{1}{D}\sum_{i=1}^{D} T_4(x_i) + \left(\frac{1}{D}\sum_{i=1}^{D} T_1(x_i)\right)^3$$
$$= \frac{1}{D}\sum_{i=1}^{D}(8x_i^4 - 8x_i^2 + 1) + \left(\frac{1}{D}\sum_{i=1}^{D} x_i\right)^3$$

- Oscillatory
- Asymmetric
- Dense coefficients ($O(D^4)$)
- Multimodal (2^D local minima)

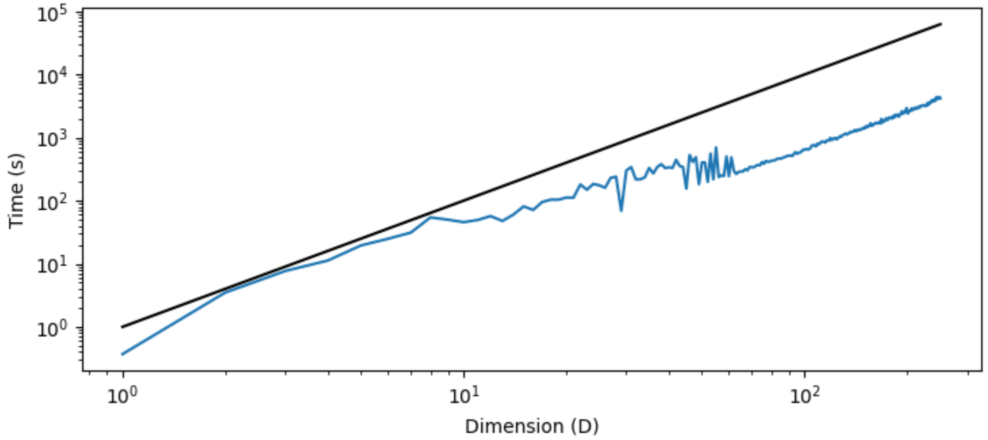
Numerical Optimization

Hypercube

Scaling & Accuracy

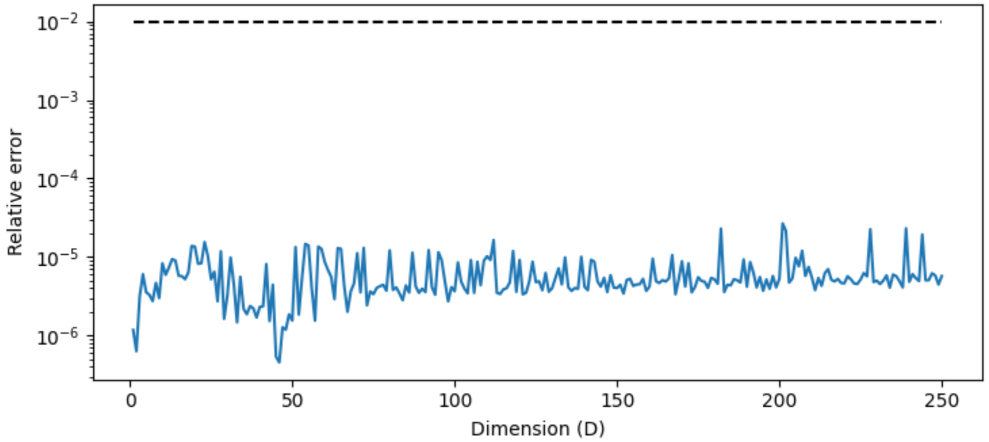


Scaling



Scaling is polynomial in dimension (O(D^2))

Accuracy (optimal value)



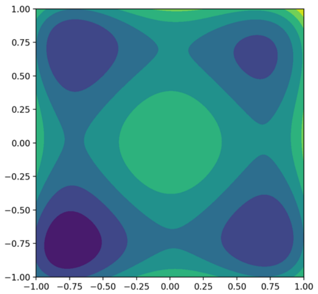
Computed minimum value is accurate to 5 digits of accuracy (solver threshold: 10^{-2})

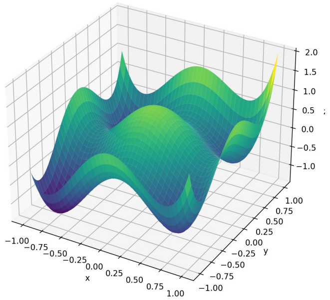
Source sample text

Numerical Optimization

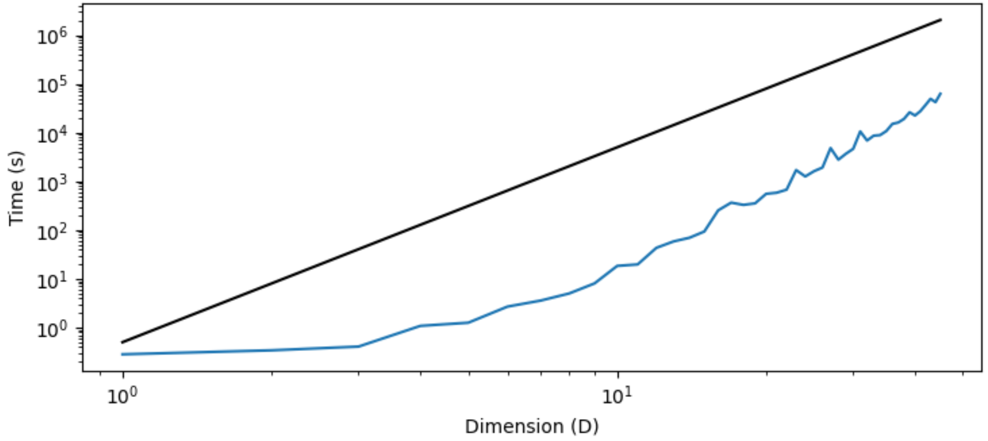
Hypercube

Scaling & Accuracy



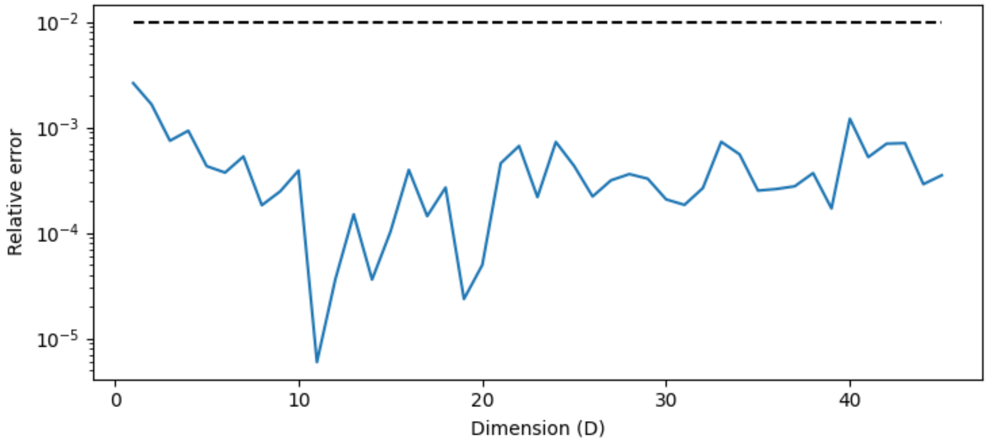


Scaling



Scaling is polynomial in dimension (O(D^4))

Accuracy (optimal value)



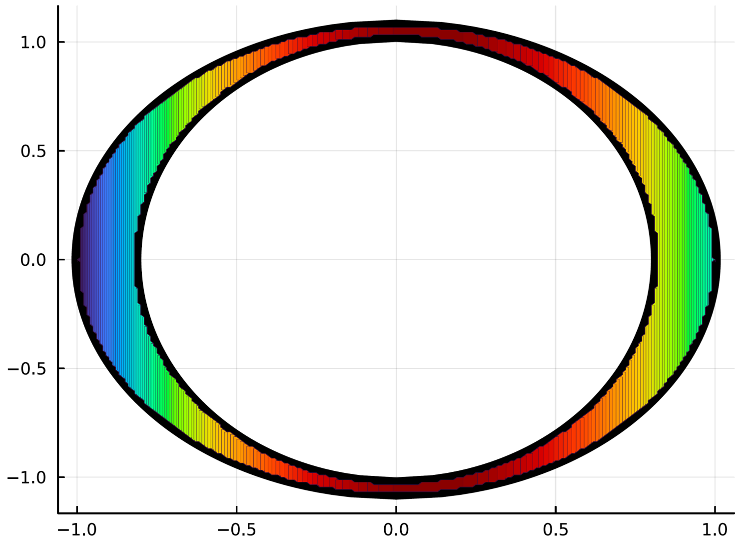
Computed minimum value is accurate to 3 digits of accuracy (threshold: 10^{-2})

Source sample text

Numerical Optimization

Constrained optimization

Connected Nonconvex Domain



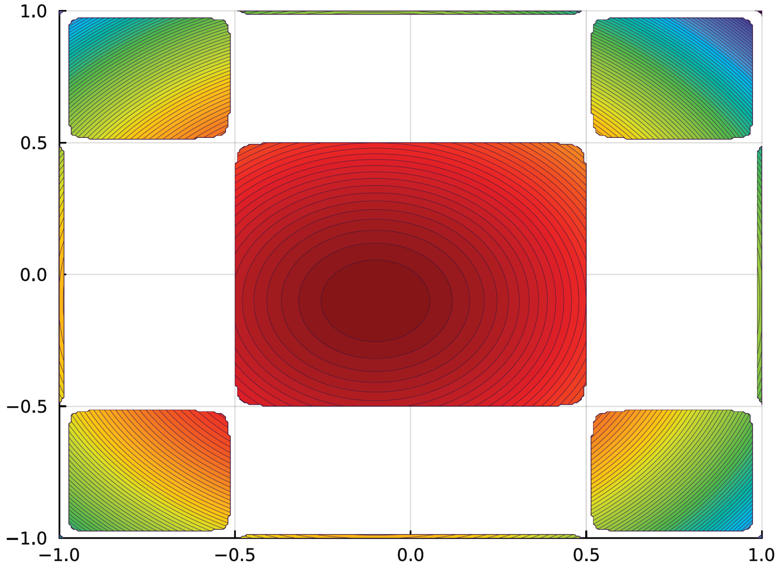
$$\min_{x\in\mathbb{R}^D} \quad p(x) := -(x_1-0.1)^2$$

$$\text{s.t.} \quad c \le x^T \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & M_D \end{pmatrix} x \le 1$$

$$-1 \le x_i \le 1 \ \ \forall i.$$

- Concave objective
- Non-convex feasible region
- Sparse coefficients
- 2 local minima, 1 global minimum

Disconnected Domain



$$\min_{x\in\mathbb{R}^D} \quad p(x) := -\sum_{i=1}^{D}(x_i+0.1)^2$$

$$\text{s.t.} \quad c(\pi x_i)\cdot c(\pi x_j) \ge 10^{-2} \ \forall i \ne j$$

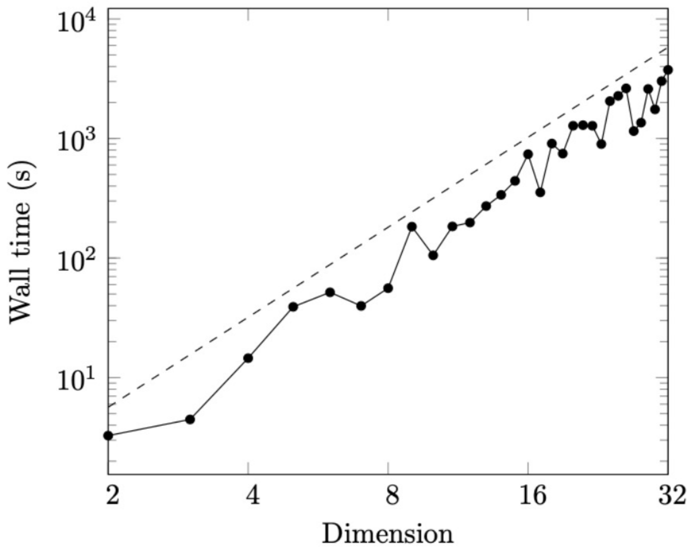
$$c(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$

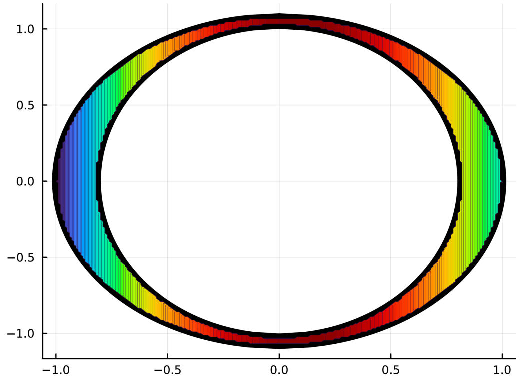
$$-1 \le x_i \le 1 \ \ \forall i.$$

- Concave objective
- Disconnected feasible region
- >2^D local minima, 1 global minimum

Source sample text

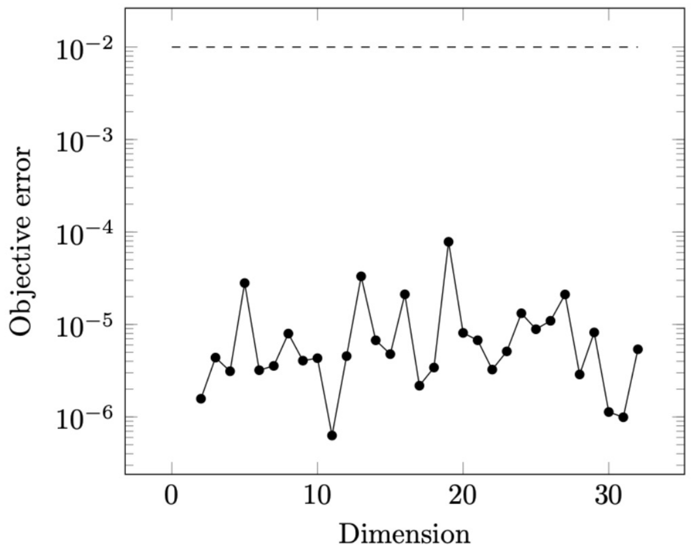
Scaling



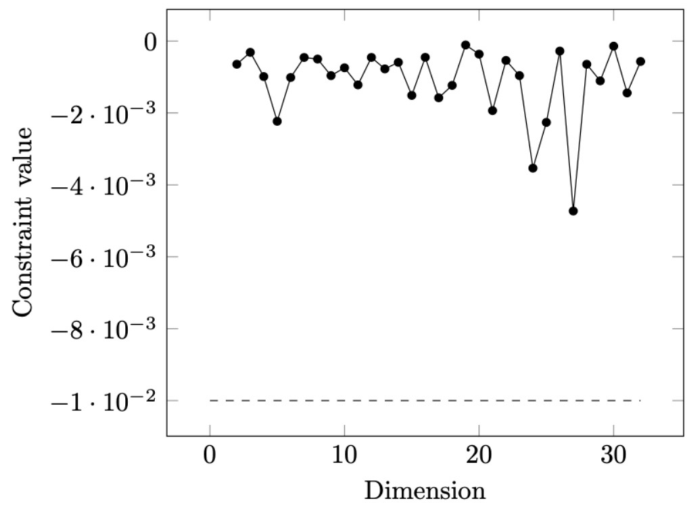


- Scaling is polynomial in dimension ($O(D^2)$)
- Computed minimum value is accurate to 5 digits of accuracy (threshold: 10^{-2})
- Solution on boundary. Feasible to 3 digits (threshold: 10^{-2})

Accuracy (optimal value)



Accuracy (constraint value)



Numerical Optimization

Constrained Optimization

Scaling & Accuracy

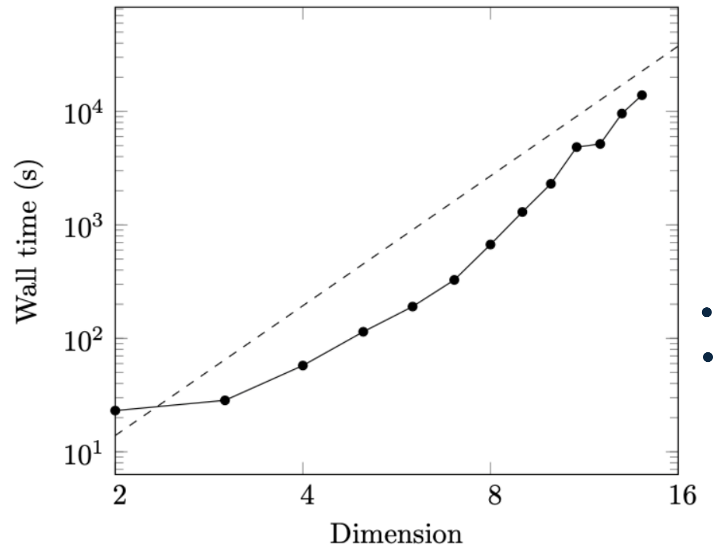
Source sample text

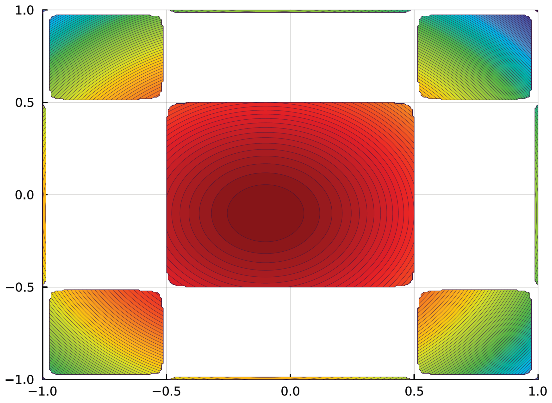
Numerical Optimization

Constrained Optimization

Scaling & Accuracy

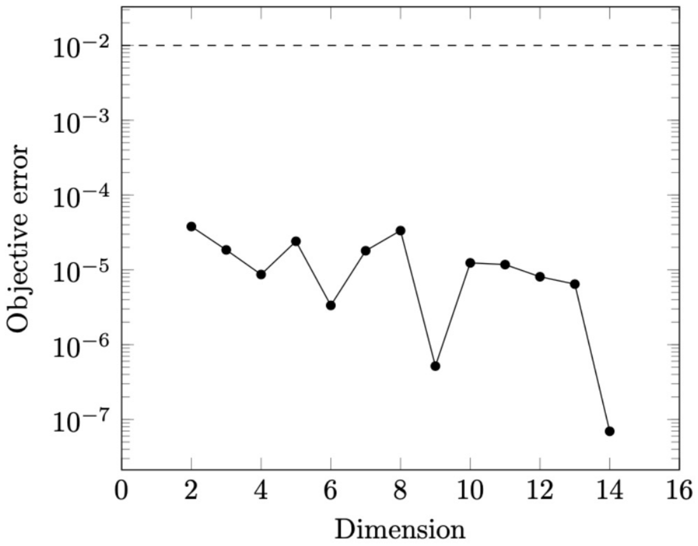
Scaling



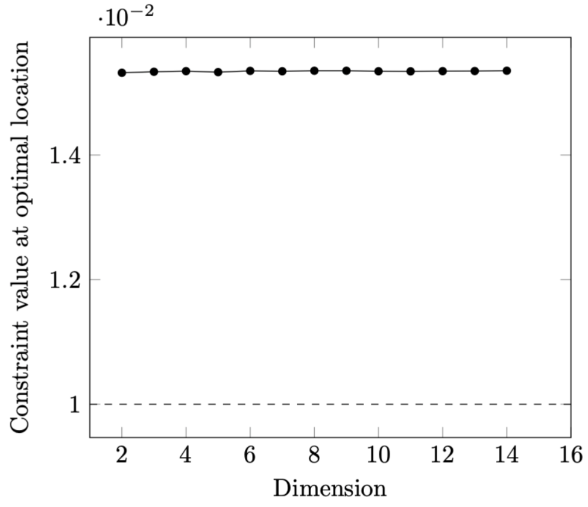


- Scaling is polynomial in dimension ($O(D^4)$)
- Computed minimum value is accurate to 4 digits of accuracy (threshold: 10^{-2})
- Solution on boundary. All feasible.

Accuracy (optimal value)



Accuracy (constraint value)



Source sample text

Conclusion

Conclusion

What we can now do

FM-GPS is a novel framework for PO with the potential to be a game-changing optimization tool

- Efficient reformulation
  - Exact reformulation
  - No spurious local minima
  - Low memory and computational complexity
  - Efficient and scalable
- Numerical demonstrations
  - Solved previously intractable problems
  - Unconstrained (hypercube) and constrained optimization
  - Tractable empirical scaling

Q-OPT References:

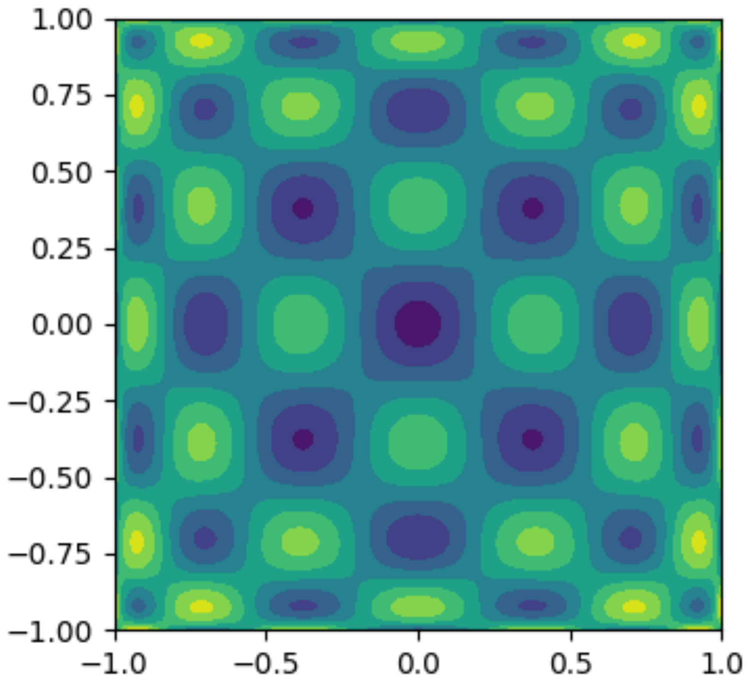
- Letourneau, Pierre-David, et al. "An Efficient Framework for Global Non-Convex Polynomial Optimization over the Hypercube." *arXiv preprint arXiv:2308.16731* (2023). (Submitted Springer Global Optimization)
- Harris, Mitchell, et al. "An Efficient Framework for Global Non-Convex Polynomial Optimization with Nonlinear Polynomial Constraints." *arXiv preprint* arXiv:2311.02037 (2023)

Conclusion

What remains to be done

Future Work

- Stationary points
- Combinatorial problems
- More efficient implementation
  - Parallelism
  - Backend solver
  - Leveraging sparsity & hierarchy
- Integration within Qualcomm's teams



Thank you

Qualcomm

Follow us on:

For more information, visit us at:

qualcomm.com & qualcomm.com/blog