Hom-PGD+: Fast Reparameterized Optimization over Non-convex Ball-Homeomorphic Set

Chenghao Liu[1], Enming Liang[1], Minghua Chen[1,2]

[1]City University of Hong Kong [2]Chinese University of Hong Kong (Shenzhen)

Problem and Motivation

Non-Convex Constrained Optimization.

$$\min_{\mathbf{x}} f_\theta(\mathbf{x}), \quad \text{s.t} \quad \mathbf{x} \in \mathcal{K}_\theta,$$

- $\mathbf{x} \in \mathbb{R}^n$ is the decision variable
- $\theta \in \mathbb{R}^d$ is the input parameter
- f_θ can be non-convex
- $\mathcal{K}_\theta \cong \mathcal{B}$ for any θ and is non-convex

Limitations of Existing Methods.

✓ Non/Low-convergence for classical projection-based methods

✓ High complexity to ensure feasibility

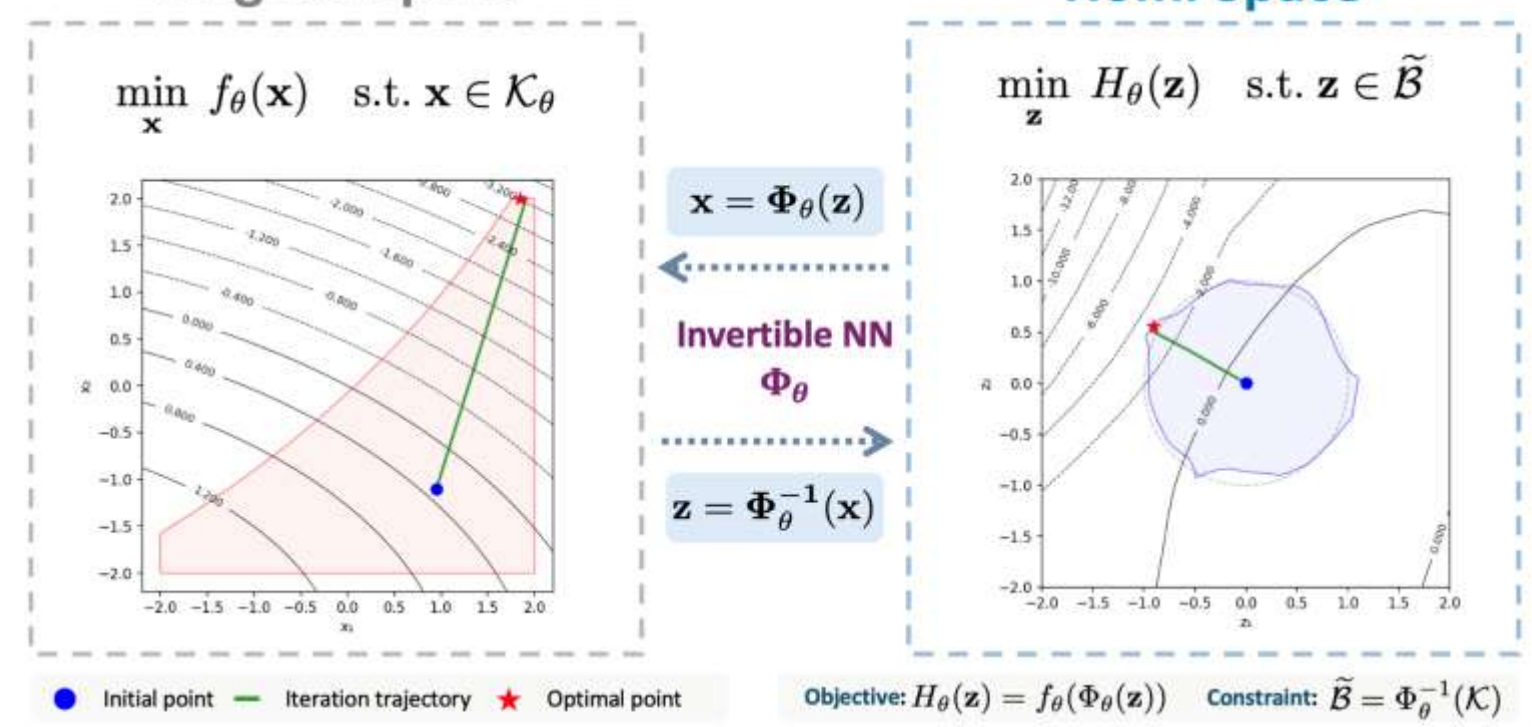
Research Gap: *efficient optimization over a* ***structured non-convex set*** *with* ***fast convergence*** *and* ***low per-iteration cost***

Summary and Contribution

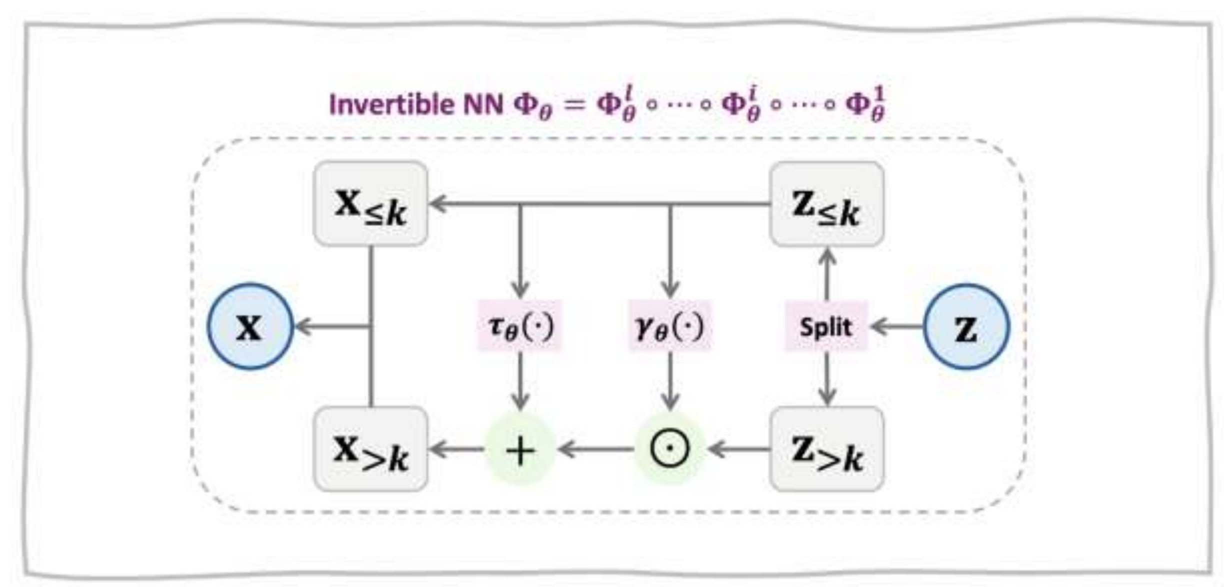
Reference	Settings Obj.	Settings Ctr.	Key Assumption	Parameterization Techniques	Algorithm	Convergence Rate
[LMY23]	NC	Simplex	–	*Hadamard Transformation*	Perturbed RGD	$\mathcal{O}(\epsilon^{-2})$
[CV25]	C	Simplex	–		Cauchy-simplex	$\mathcal{O}(\epsilon^{-1})$
[TT24]	(N)C	Polyhedra	Full-rank constraints.		RGD + PGD	N/A
[LLC25]	C SC NC	Convex	Non-degeneracy. – –	*Gauge Mapping*	PGD over ball	$\mathcal{O}(\epsilon^{-1})$ $\mathcal{O}(\log \epsilon^{-1})$ $\mathcal{O}(\epsilon^{-2})$
[BH18]	SC	NC	Small local concavity coefficients of constraints.	–	PGD	$\mathcal{O}(\log \epsilon^{-1})$
[LMX22]	WC	WC	Certain non-singularity. Initial feasible points.	–	Proximal-point penalty method	$\tilde{\mathcal{O}}(\epsilon^{-3})$ $\tilde{\mathcal{O}}(\epsilon^{-4})$
[BSH23]	IV SIV	IV	Contraction and triangle inequality w.r.t. invexity.	–	Invex PGD	$\mathcal{O}(\epsilon^{-1})$ $\mathcal{O}(\log \epsilon^{-1})$
Theorem 1	NC	NC	Ball-homeomorphic.	*Invertible Neural Network*	Bisected-PGD	$\mathcal{O}(\epsilon^{-2})$

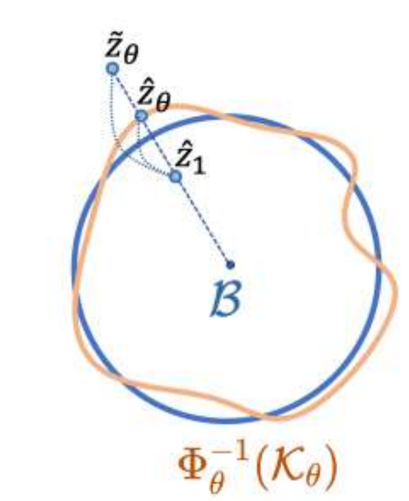
We propose Hom-PGD+ as the *first projection-efficient, learning-based first-order* framework capable of solving optimization **over a broad non-convex set** while achieving the **fastest convergence rate** in the SOTA methods

Homeomorphic Constrained Optimization



Learned Homeomorphism in Practice: Invertible Neural Network





Properties of INNs.

- Universal approximation
- **Bisected projection**
- Bi-Lipschitz continuous
- Unsupervised training

Algorithm 1 Hom-PGD$^+$

Input: initial point $\mathbf{z}_0$, valid INN Φ_θ, reformulated problem $(\mathrm{H}_{\mathrm{inn}})$, and maximum iterations K

for $k = 0$ to K **do**

  Compute stepsize α_k and gradient $\nabla H_\theta(\mathbf{z}_k)$

  Gradient descent: $\mathbf{z}_{k+1} = \mathbf{z}_k - \alpha_k \nabla H_\theta(\mathbf{z}_k)$

  Projection: $\mathbf{z}_{k+1} \leftarrow \mathrm{BP}_{\tilde{\mathcal{B}}}(\mathbf{z}_{k+1})$ **if** $\Phi_\theta(\mathbf{z}_{k+1}) \notin \mathcal{K}_\theta$

end for

Output: $\mathbf{x}_K = \Phi_\theta(\mathbf{z}_K)$

Performance Analysis

Assumptions.

- Smooth objective and constrained functions
- ϵ_{inn} is the INN approximation error (including function and Jacobian approximation)

Landscape Analysis.

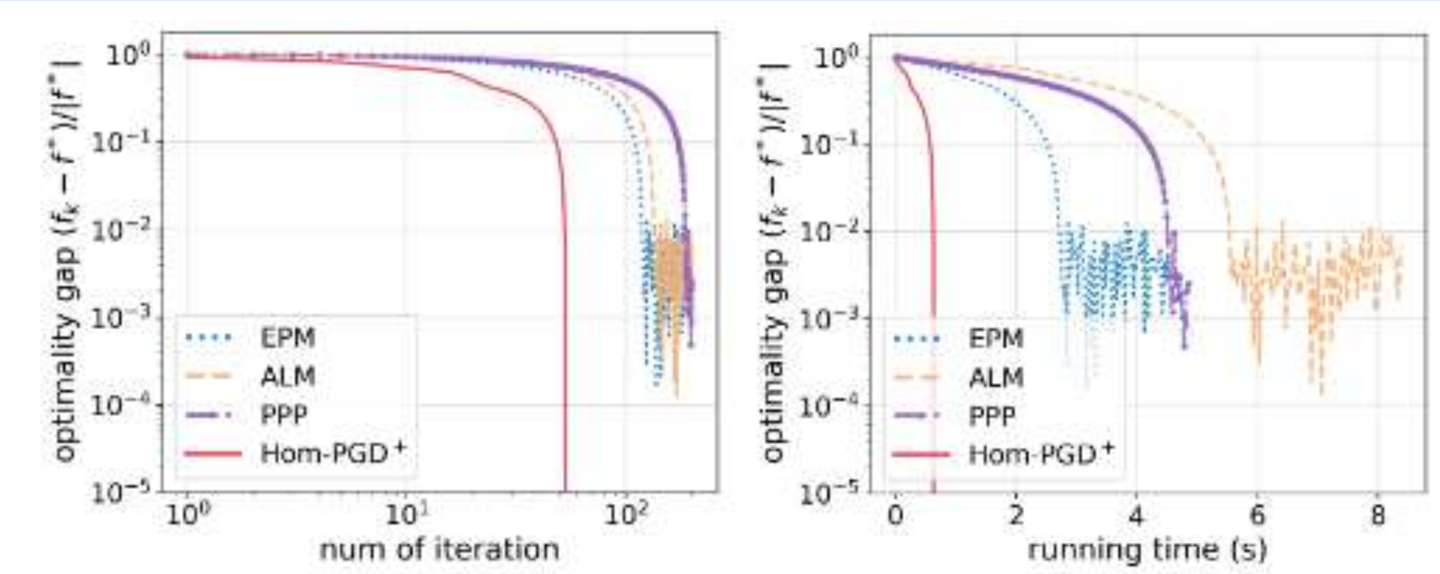
- KKT stationary point equivalence. A point is a KKT stationary point of (P) if and only if it's a KKT stationary point of (H) under mild conditions.

Convergence Rate of Hom-PGD+.

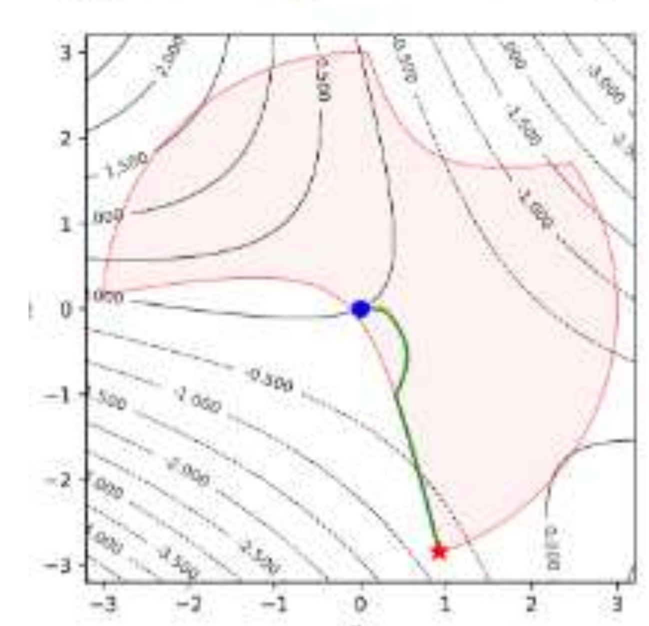
Theorem. Hom-PGD+ with a constant step-size can find an $\epsilon + O(\sqrt{\epsilon_{\mathrm{inn}}})$ approximate KKT stationary point in $O(\epsilon^{-2})$ iterations.

Per-Iteration Complexity: $\mathcal{O}(W)$ where W is the INN size. Usually, $W = O(n^2)$ make excellent performance in practice.

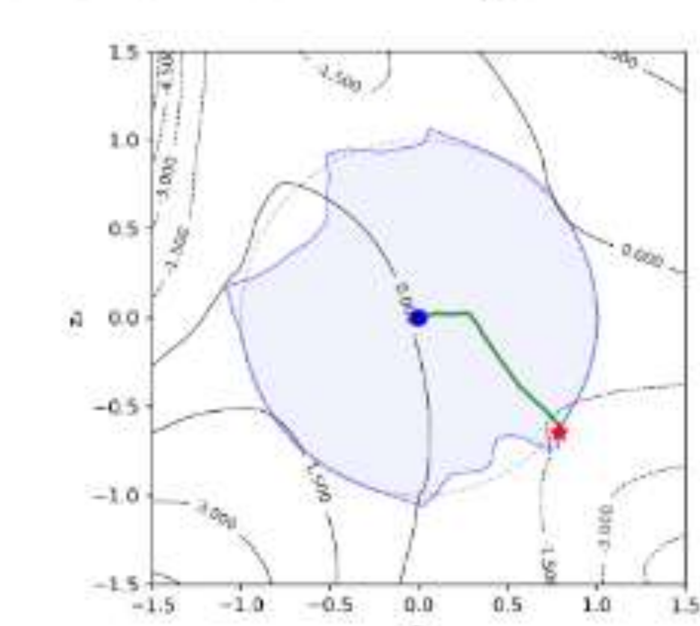
Illustration Example: QCQP



(a) Convergence with respect to iterations and running time.



(b) ALM iter.



(c) Hom-PGD$^+$ iter.