Sensor network localization has benign landscape under mild rank relaxation

November 29, 2024

Chris Criscitiello

with

Andrew McRae, Quentin Rebjock, Nicolas Boumal

OPTIM, Chair of Continuous Optimization

Institute of Mathematics, EPFL

EPFL

Landscape analysis

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, …)

Landscape analysis

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, …)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$

Stable manifold theorems
+
Łojasiewicz theorem

Landscape analysis

$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, …)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$

Stable manifold theorems
+
Łojasiewicz theorem

Goal: Show all 2-critical points x are global minima.

Landscape analysis

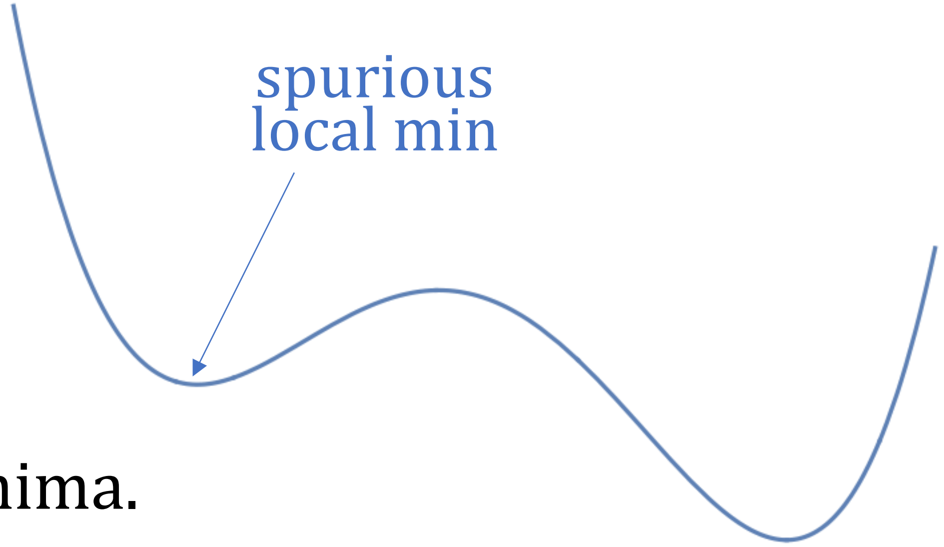
$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, …)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$



Goal: Show all 2-critical points x are global minima.

Landscape analysis

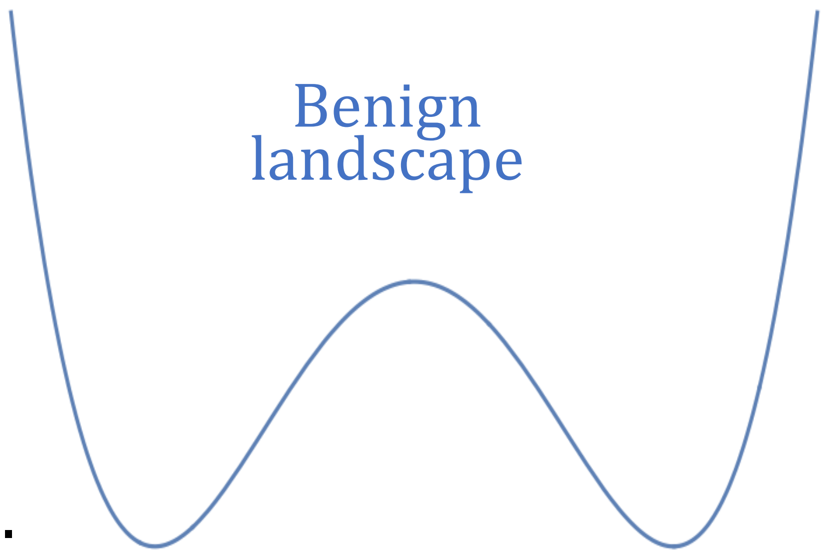
$$\min_{x \in \mathcal{M}} f(x)$$

Run local algorithm (gradient descent, trust regions, …)

Will converge to 2-critical point (w/ prob 1)

[1-critical] $\nabla f(x) = 0$

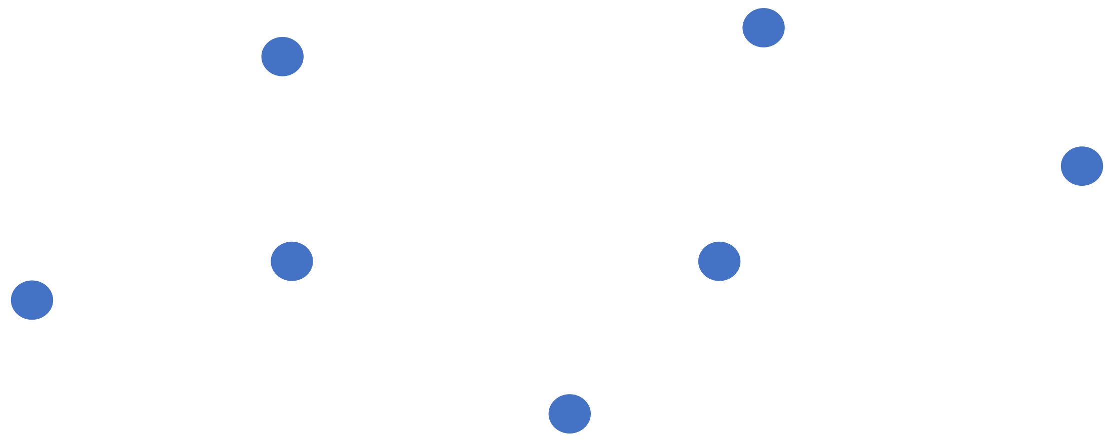
[2-critical] $\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$



Goal: Show all 2-critical points x are global minima.

The problem

n unknown points $z_1^*, z_2^*, \ldots, z_n^*$ in $\mathbb{R}^\ell$.

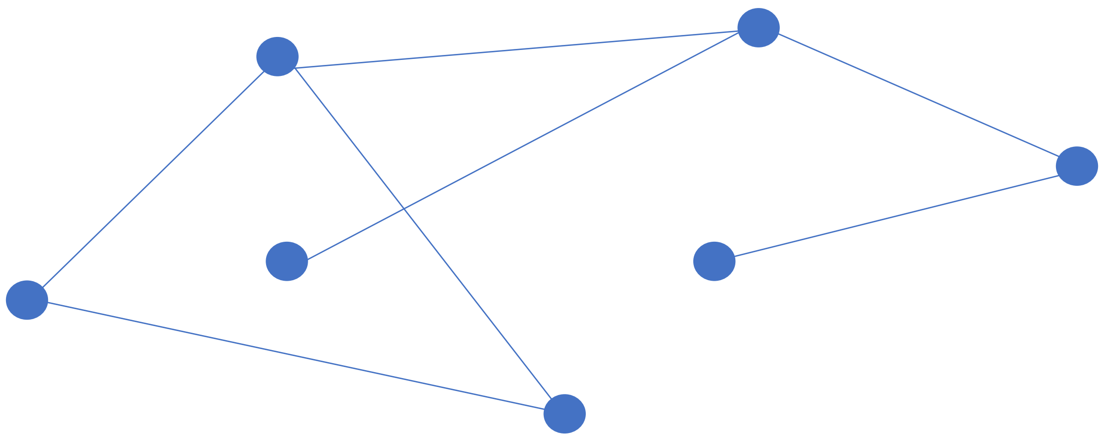


The problem

n unknown points $z_1^*, z_2^*, \ldots, z_n^*$ in $\mathbb{R}^\ell$.

Know a subset of the pairwise distances (measurements)

$$d_{ij} = \|z_i^* - z_j^*\| \text{ for } ij \in E.$$



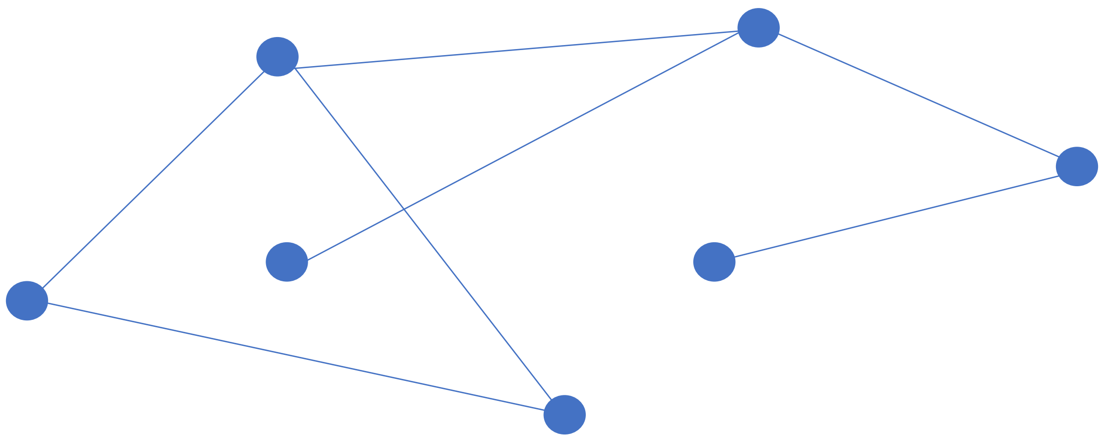
The problem

n unknown points $z_1^*, z_2^*, \ldots, z_n^*$ in $\mathbb{R}^\ell$.

Know a subset of the pairwise distances (measurements)

$$d_{ij} = \|z_i^* - z_j^*\| \text{ for } ij \in E.$$

Goal: recover the n points (up to translation & rotation)



Applications

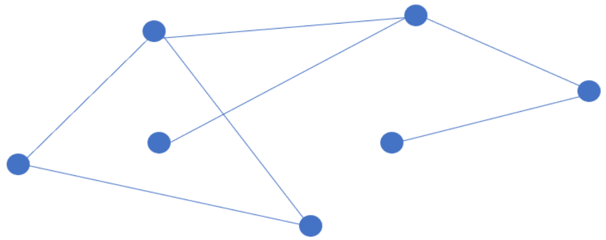
Robotics (**sensor network localization**), $\ell = 2,3$

Molecular conformation

Data analysis (metric **multidimensional scaling**)

Graph theory (rigidity)

When is recovery possible?

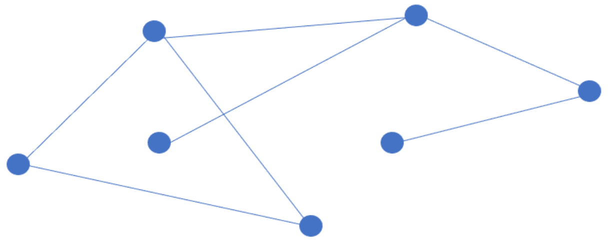


Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^\ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

When is recovery possible?



Global rigidity: Configuration space

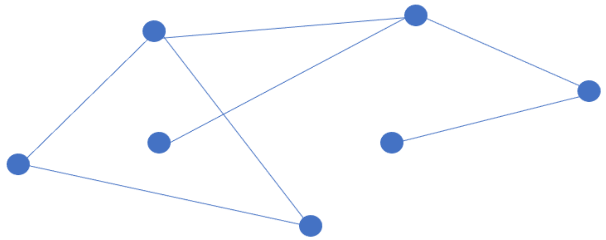
$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^\ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

NP-hard!

“Euclidean distance geometry and applications” -- Liberti, et al

When is recovery possible?



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^\ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

NP-hard!

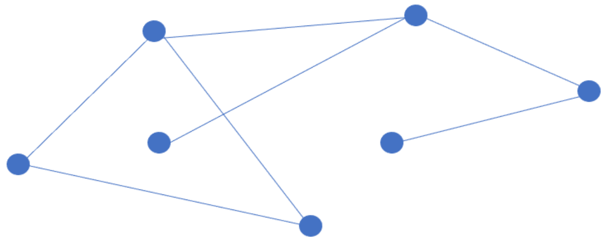
"Euclidean distance geometry and applications" -- Liberti, et al

Universal rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^k \text{ for all } k \geq \ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

When is recovery possible?



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^\ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

NP-hard!

"Euclidean distance geometry and applications" -- Liberti, et al

Universal rigidity: Configuration space

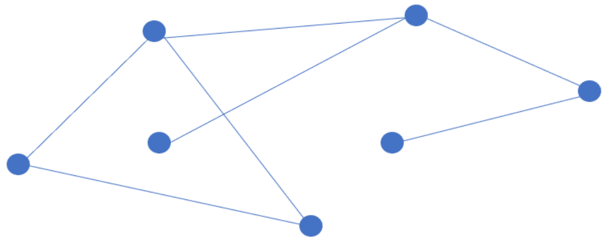
$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^k \text{ for all } k \geq \ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

Polynomial time by SDPs

"Theory of semidefinite programming for Sensor Network Localization" -- So, Ye

When is recovery possible?



Global rigidity: Configuration space

$$\{z_1, z_2, \dots, z_n \in \mathbb{R}^\ell : d_{ij} = \|z_i - z_j\|\}$$

should be a singleton (after quotienting).

NP-hard!

“Euclidean distance geometry and applications” -- Liberti, et al

Universal rigidity: Configuration space

Drawback: SDP involves $(n + \ell)\times(n + \ell)$ matrices

should be a singleton (after quotienting).

Polynomial time by SDPs

“Theory of semidefinite programming for Sensor Network Localization” -- So, Ye

Optimization problem

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^\ell$

"s-stress"

Optimization problem

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^\ell$

"s-stress"

Solved via local algorithms. Guarantees?

Nonconvex! How bad?

Optimization problem

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^\ell$

"s-stress"

Solved via local algorithms. Guarantees?

Nonconvex! How bad?

Possible variations: Noisy measurements, landmarks, …

Our focus: (nearly) complete graphs, no noise

Synthetic experiments, complete graph

Recipe (all distances known):

(1) Choose ground truths $z_1^*, z_2^*, \ldots, z_n^*$ at random (normal iid)

(2) Run gradient descent/trust regions/etc.

(3) Find global min?

(4) Repeat

Synthetic experiments, complete graph

Recipe (all distances known):

(1) Choose ground truths $z_1^*, z_2^*, \ldots, z_n^*$ at random (normal iid)

(2) Run gradient descent/trust regions/etc.

(3) Find global min?

(4) Repeat

Always finds global min!

Synthetic experiments, complete graph

Recipe (all distances known):

(1) Choose ground truths $z_1^*, z_2^*, \ldots, z_n^*$ at random (normal iid)

(2) Run gradient descent/trust regions/etc.

(3) Find global min?

(4) Repeat

Always finds global min!

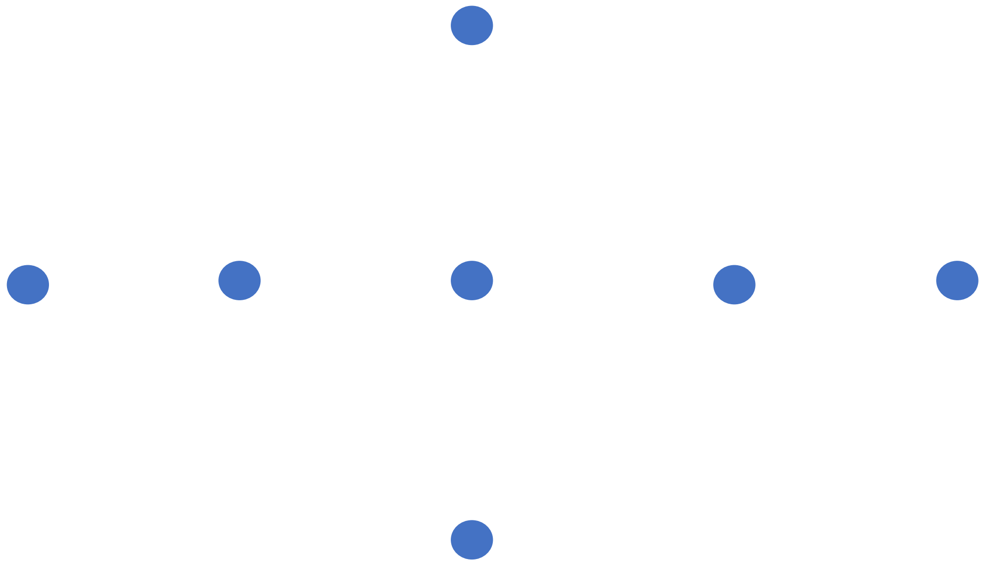
Open Question: Does s-stress have spurious local minima? Are all 2-critical points global minima?

* Malone & Trosset 2000, Parhizkar 2013, etc.

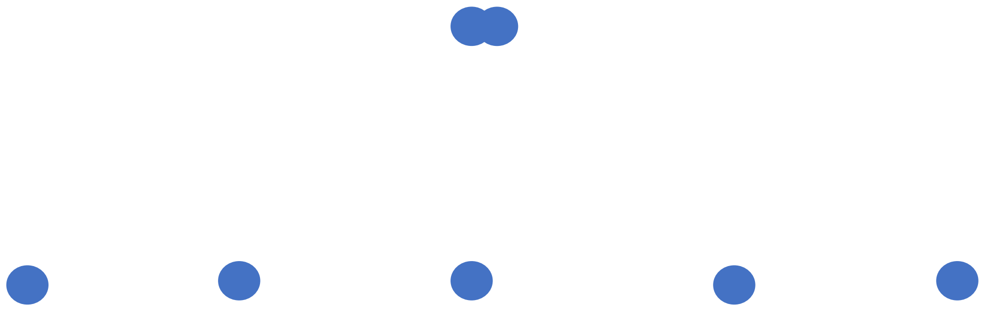
Counterexamples

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, \ldots$



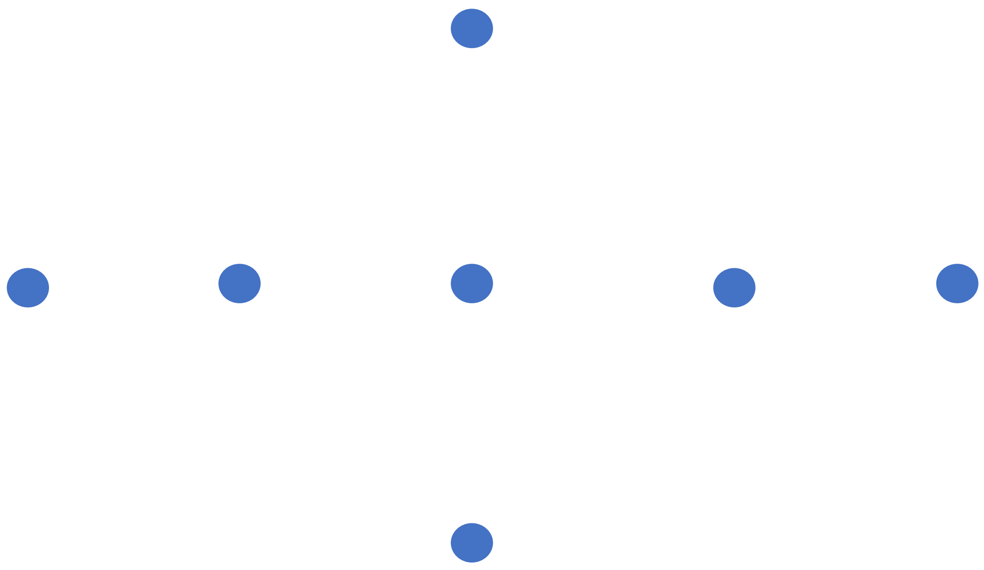
Spurious configuration $z_1, z_2, \ldots$



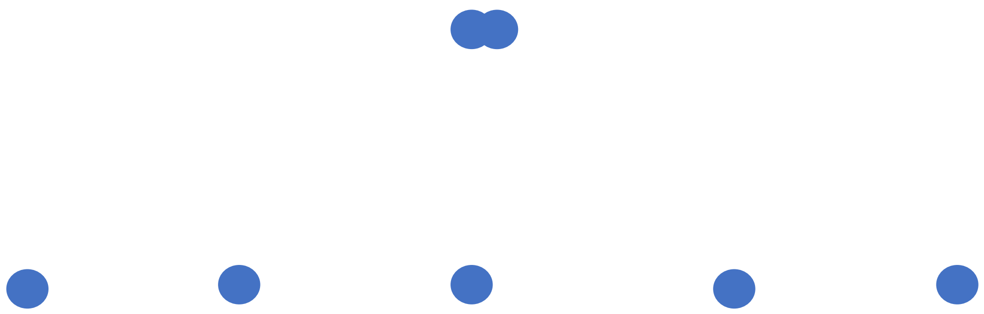
Counterexamples

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, \ldots$



Spurious configuration $z_1, z_2, \ldots$

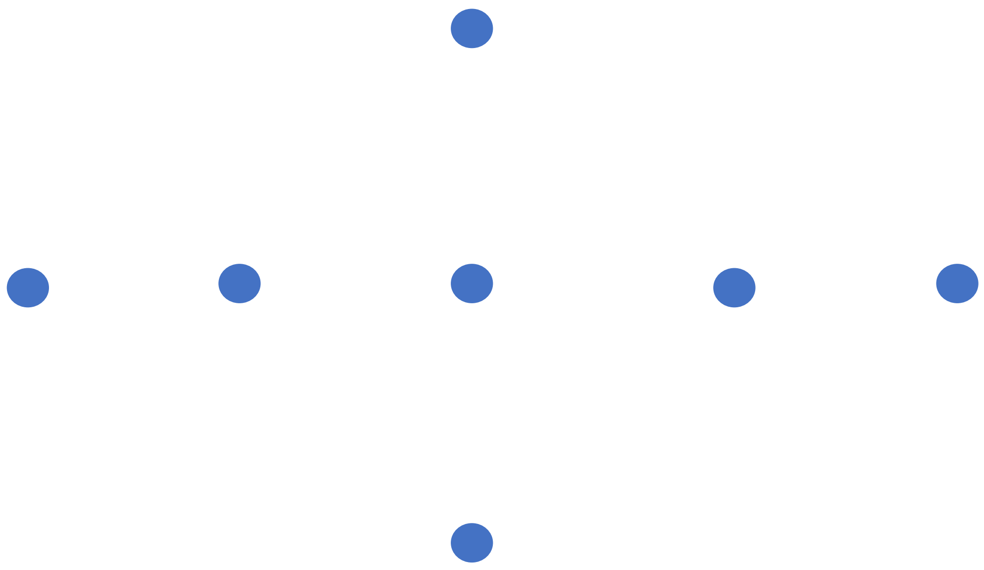


Also see:
Song, Goncalves, Jung, Lavor, Mucherino, Wolkowicz, 2024

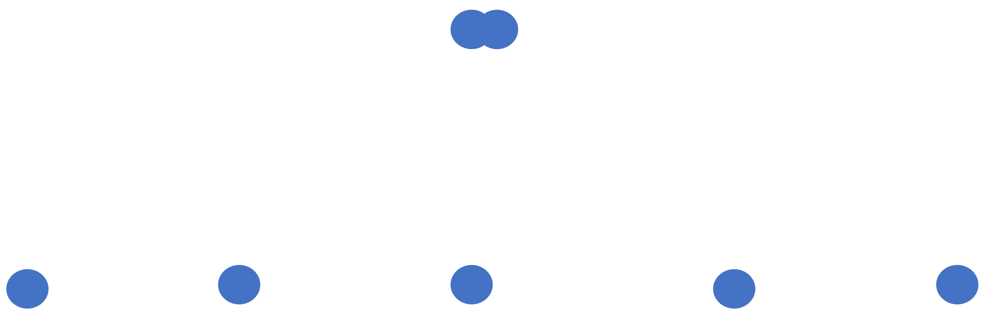
Counterexamples

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, \ldots$



Spurious configuration $z_1, z_2, \ldots$



Set of ground truths with spurious local minima has positive measure

Counterexamples

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, \ldots$

Spurious configuration $z_1, z_2, \ldots$

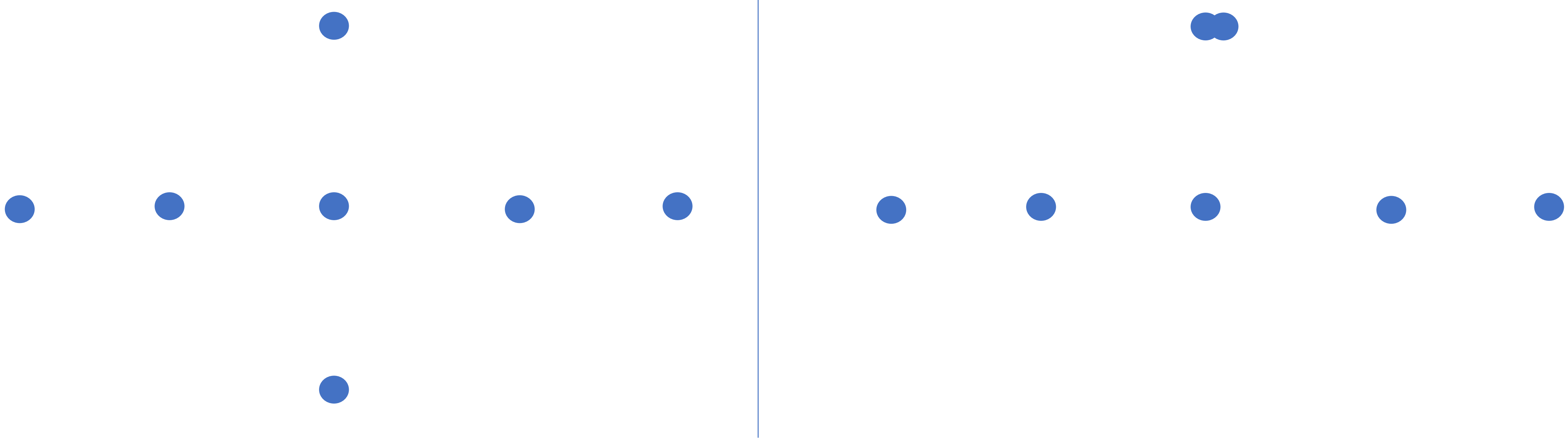
Landscape is not benign, so we have to do something!
What?

Set of ground truths with spurious local minima has positive measure

Counterexamples

s-stress can have spurious strict local minima!

Ground truth $z_1^*, z_2^*, \ldots$ | Spurious configuration $z_1, z_2, \ldots$



Set of ground truths with spurious local minima has positive measure

Nonconvex relaxation

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^{\ell}$

Nonconvex relaxation

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$

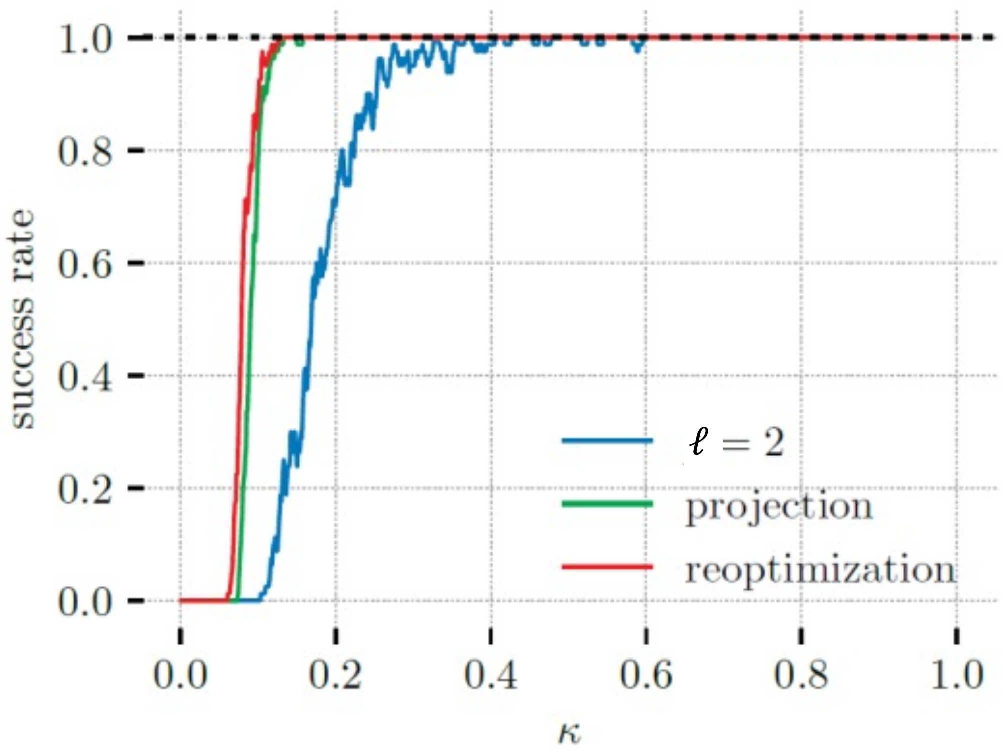
Relax to dimension $k > \ell$

Nonconvex relaxation

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \dots, z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$



$n = 100$
$\ell = 2$
$k = 4$

$\kappa =$ edge density (Erdos-Renyi)

Nonconvex relaxation

$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original?
Yes if graph is complete (or more generally if it is *universally rigid*)

Nonconvex relaxation

$$\min \sum_{ij\in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original?
Yes if graph is complete (or more generally if it is *universally rigid*)

Want k small; new problem has kn variables

Nonconvex relaxation

$$\min \sum_{ij\in E} \left(\|z_i - z_j\|^2 - d_{ij}^2\right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$

Relax to dimension $k > \ell$

Minimizer of relaxed problem same as original?
Yes if graph is complete (or more generally if it is *universally rigid*)

Want k small; new problem has kn variables
If $k = n - 1$, easy to see landscape is benign (Song, Goncalves, Jung, Lavor, Mucherino, Wolkowicz, 2024)
Can we do better?

Results

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell},$$

then every 2-critical point is the ground truth.

Results

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell},$$

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic* and iid, and relax to

$$k \approx \ell + \log(n),$$

then every 2-critical point is the ground truth, w.h.p.

Results

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell},$$

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic* and iid, and relax to

$$k \approx \ell + \log(n)\,,$$

then every 2-critical point is the ground truth, w.h.p.

Conjecture [**arbitrary GT**]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [**isotropic GT**]: Relaxing is not necessary.

Results

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell},$$

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic* and iid, and relax to

$$k \approx \ell + \log(n)\,,$$

then every 2-critical point is the ground truth, w.h.p.

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

- True if $n \leq \ell + 3$
- Numerical optim to explicitly search for counterexamples.

Results

Theorem [arbitrary GT]: If graph is complete and relax to

$$k \approx \ell + \sqrt{n\ell},$$

then every 2-critical point is the ground truth.

Theorem [isotropic GT]: If graph is nearly complete*, ground truth points are isotropic* and iid, and relax to

$$k \approx \ell + \log(n),$$

then every 2-critical point is the ground truth, w.h.p.

- True if $n \leq \ell + 3$
- Numerical optim to explicitly search for counterexamples.

Conjecture [arbitrary GT]: Relaxing to $k = \ell + 1$ is enough.

Conjecture [isotropic GT]: Relaxing is not necessary.

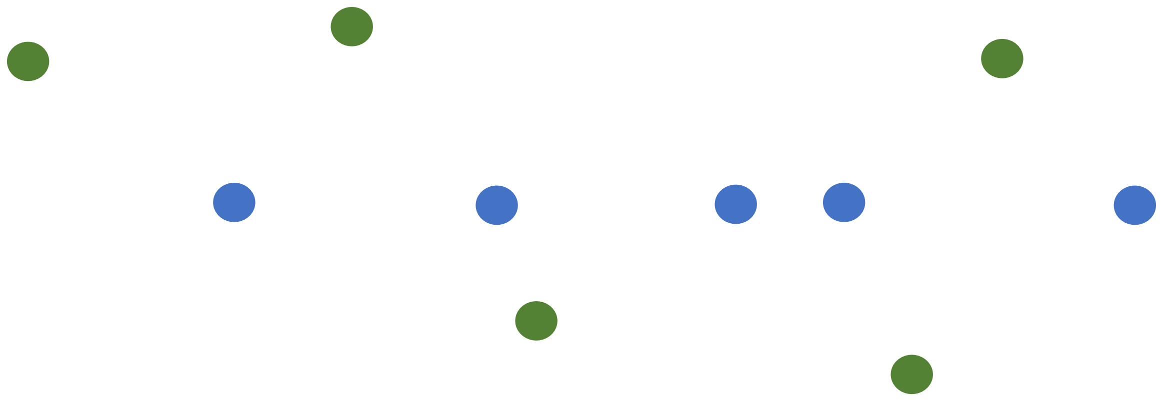
Some ideas from the proof

Ground truth $z_1^*, z_2^*, \ldots$ in dimension ℓ

Some ideas from the proof

Ground truth $z_1^*, z_2^*, \dots$ in dimension ℓ

1-critical configuration in dimension $k > \ell$

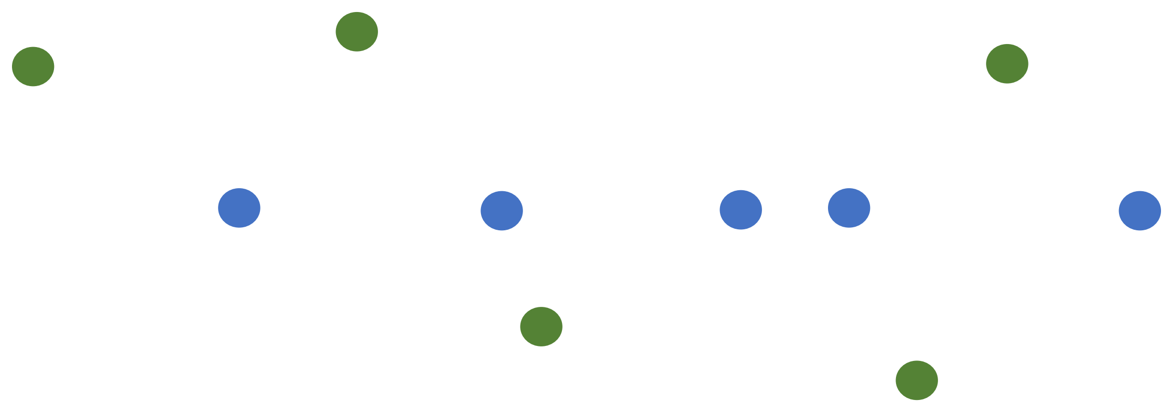


Some ideas from the proof

Ground truth $z_1^*, z_2^*, \ldots$ in dimension ℓ

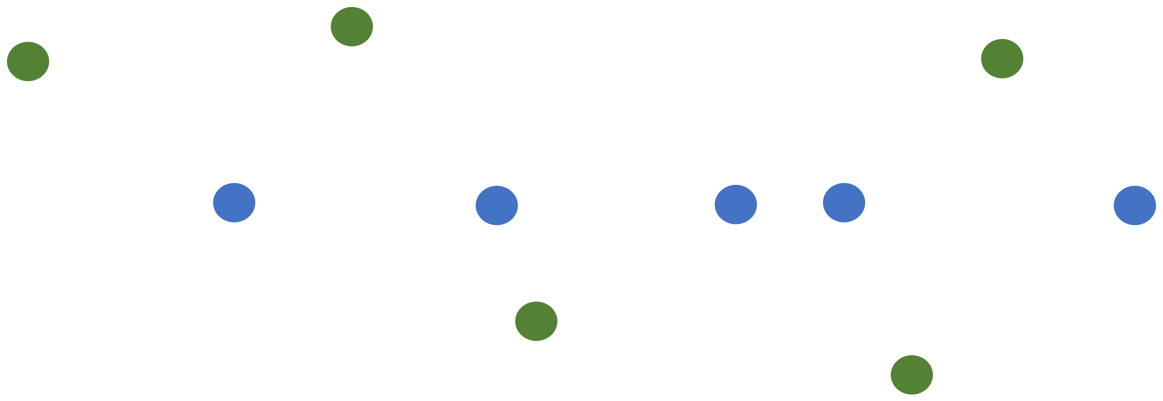
1-critical configuration in dimension $k > \ell$

Goal: perturb 1-critical configuration to decrease cost



Some ideas from the proof

Goal: perturb 1-critical configuration to decrease cost

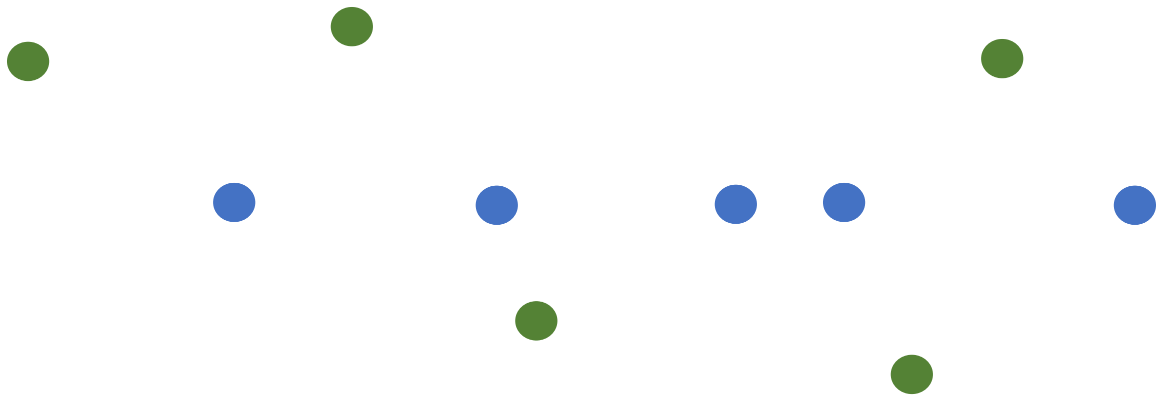


Some ideas from the proof

Goal: perturb 1-critical configuration to decrease cost

Best linear transformation $z_i \mapsto Rz_i$ mapping 1-critical config to ground truth

$$R = \mathrm{argmin}_R \sum_i \|z_i^* - Rz_i\|^2, \qquad R \in \mathbb{R}^{\ell \times k}$$

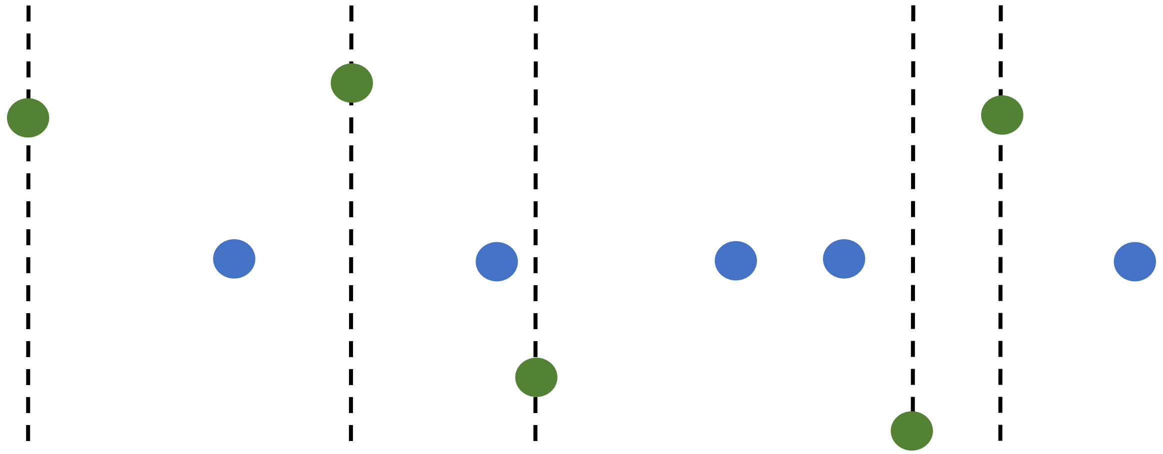


Some ideas from the proof

Goal: perturb 1-critical configuration to decrease cost

Best linear transformation $z_i \mapsto Rz_i$ mapping 1-critical config to ground truth

$$R = \text{argmin}_R \sum_i \|z_i^* - Rz_i\|^2, \quad R \in \mathbb{R}^{\ell \times k}$$

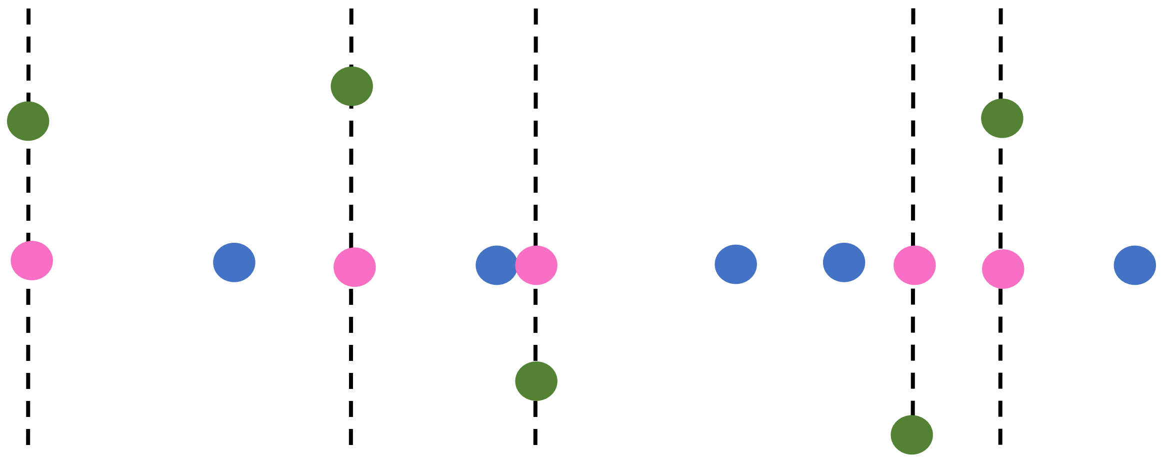


Some ideas from the proof

Goal: perturb 1-critical configuration to decrease cost

Best linear transformation $z_i \mapsto Rz_i$ mapping 1-critical config to ground truth

$$R = \mathrm{argmin}_R \sum_i \|z_i^* - Rz_i\|^2, \quad R \in \mathbb{R}^{\ell \times k}$$

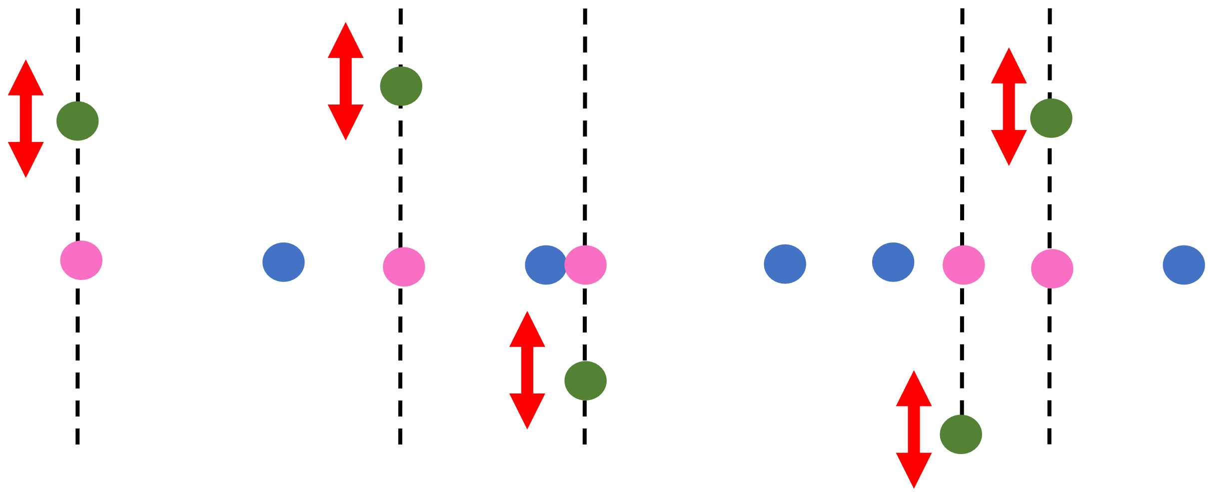


Some ideas from the proof

Goal: perturb 1-critical configuration to decrease cost

Best linear transformation $z_i \mapsto Rz_i$ mapping 1-critical config to ground truth

$$R = \text{argmin}_R \sum_i \|z_i^* - Rz_i\|^2, \qquad R \in \mathbb{R}^{\ell \times k}$$



Alternative perspective:
Low-Rank Optimization

Notation and reformulation

$$Z = \begin{pmatrix} z_1^\top \\ \vdots \\ z_n^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}, \qquad Z_* = \begin{pmatrix} {z_1^*}^\top \\ \vdots \\ {z_n^*}^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}$$

Notation and reformulation

$$Z = \begin{pmatrix} z_1^\top \\ \vdots \\ z_n^\top \end{pmatrix} \in \mathbb{R}^{n\times \ell}, \qquad Z_* = \begin{pmatrix} {z_1^*}^\top \\ \vdots \\ {z_n^*}^\top \end{pmatrix} \in \mathbb{R}^{n\times \ell}$$

Gram matrices $Y = ZZ^\top, Y_* = Z_* Z_*^\top$

Notation and reformulation

$$Z = \begin{pmatrix} z_1^\top \\ \vdots \\ z_n^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}, \qquad Z_* = \begin{pmatrix} {z_1^*}^\top \\ \vdots \\ {z_n^*}^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}$$

Gram matrices $Y = ZZ^\top, Y_* = Z_* Z_*^\top$

MDS map $\Delta : \mathrm{Sym}(n) \to \mathrm{Hollow}(n)$

Gram $\to$ EDM (euclidean distance matrix)

ij-entry $= \langle z_i, z_j \rangle$ $\qquad$ ij-entry $= \|z_i - z_j\|^2$

Notation and reformulation

$$Z = \begin{pmatrix} z_1^\top \\ \vdots \\ z_n^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}, \qquad Z_* = \begin{pmatrix} {z_1^*}^\top \\ \vdots \\ {z_n^*}^\top \end{pmatrix} \in \mathbb{R}^{n\times\ell}$$

Gram matrices $Y = ZZ^\top, Y_* = Z_* Z_*^\top$

MDS map $\Delta : \mathrm{Sym}(n) \to \mathrm{Hollow}(n)$

$\qquad\qquad\qquad \mathrm{Gram} \to$ EDM (euclidean distance matrix)

ij-entry $= \langle z_i, z_j \rangle$ $\qquad$ ij-entry $= \|z_i - z_j\|^2$

$$[\Delta(Y)]_{ij} := Y_{ii} + Y_{jj} - 2Y_{ij}$$

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_* Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

"s-stress"

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_* Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

$\Updownarrow$

Burer-Monteiro factorization!

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \le \ell$$

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_* Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n \times \ell}$$

$\Updownarrow$

Burer-Monteiro factorization!

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \le \ell$$

$\Downarrow$ (relax)

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \le k$$

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_*Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

$\Updownarrow$

Burer-Monteiro factorization!

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \le \ell$$

$\Downarrow$ (relax)

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \le k$$

- If $k = n$, problem is convex (1-critical points are global mins)

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_*Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

$$\Updownarrow$$

Burer-Monteiro factorization!

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq \ell$$

$\Downarrow$ (relax)

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

- If $k = n$, problem is convex (1-critical points are global mins)
- Map $Z \mapsto ZZ^\top$ is $2 \Rightarrow 1$, i.e., 2-critical points map to 1-critical points

[Levin , Kileel, Boumal 2022; Ha, Liu, Barber 2018]

Notation and reformulation

$$\min \|\Delta(ZZ^\top - Z_* Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

$$\Updownarrow$$

Burer-Monteiro factorization!

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \le \ell$$

$\Downarrow$ (relax)

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \le k$$

- If $k = n$, problem is convex (1-critical points are global mins)
- Map $Z \mapsto ZZ^\top$ is $2 \Longrightarrow 1$, i.e., 2-critical points map to 1-critical points

[Levin , Kileel, Boumal 2022; Ha, Liu, Barber 2018]

- Conclusion: Landscape benign if $k = n$

Restricted Isometry Property?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

Restricted Isometry Property?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \leq k$$

Restricted Isometry Property (RIP):

$$\|Y\|_F^2 \leq \|\Delta(Y)\|_F^2 \leq 3\|Y\|_F^2 \qquad \text{for all } Y \text{ s.t. } \mathrm{rank}(Y) \leq 2k.$$

If RIP, then benign landscape [Bhojanapalli et al., 2016; Ge et al., 2017; Zhang et al., 2019]

Restricted Isometry Property?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

Restricted Isometry Property (RIP):

$$\|Y\|_F^2 \leq \|\Delta(Y)\|_F^2 \leq 3\|Y\|_F^2 \qquad \text{for all } Y \text{ s.t.} \operatorname{rank}(Y) \leq 2k.$$

If RIP, then benign landscape [Bhojanapalli et al., 2016; Ge et al., 2017; Zhang et al., 2019]

Δ does not satisfy RIP! Δ has RIP-condition-number n

Special properties of MDS map?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

Special properties of MDS map?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \leq k$$

Special "perturbation" of the identity

$$(\Delta^* \circ \Delta)(Y) = Y + \Theta(Y)$$

$$(\Delta^* \circ \Delta)^{-1}(Y) = Y - \Gamma(Y)$$

Special properties of MDS map?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \leq k$$

Special "perturbation" of the identity

$$(\Delta^* \circ \Delta)(Y) = Y + \Theta(Y)$$

$$(\Delta^* \circ \Delta)^{-1}(Y) = Y - \Gamma(Y)$$

New "general" theorem: If Γ is completely positive, contractive, and satisfies

- $a^\top \Gamma(ab^\top + ba^\top)b \leq 2a^\top \Gamma(bb^\top)a \quad \forall a, b \in \mathbb{R}^n$
- $\langle Y, \Theta(Y)\rangle \leq c\langle Y, \Gamma(Y)\rangle \quad \forall Y$

then landscape is benign when relax to $k \approx \ell + \sqrt{c\ell}$.

Special properties of MDS map?

$$\min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \operatorname{rank}(Y) \le k$$

$$(\Delta^* \circ \Delta)(Y) = Y + \Theta(Y)$$

$$(\Delta^* \circ \Delta)^{-1}(Y) = Y - \Gamma(Y)$$

New "general" theorem: If Γ is completely positive, contractive, and satisfies

- $a^\top \Gamma(ab^\top + ba^\top)b \le 2a^\top \Gamma(bb^\top)a \quad \forall a, b \in \mathbb{R}^n$
- $\langle Y, \Theta(Y)\rangle \le c\langle Y, \Gamma(Y)\rangle \quad \forall Y$

then landscape is benign when relax to $k \approx \ell + \sqrt{c\ell}$.

E.g., $\Gamma(Y) = \sum_{i=1}^{N} a_i a_i^\top \left(a_i^\top Y a_i\right)$ with $a_i \in \mathbb{R}^n$

Takeaways

Summary:

- s-stress can have spurious local mins (even for complete graph)
- If relax mildly ($\sqrt{n}$ or $\log n$), s-stress landscape becomes benign

Takeaways

Summary:

- s-stress can have spurious local mins (even for complete graph)
- If relax mildly ($\sqrt{n}\ \ or \log n$), s-stress landscape becomes benign

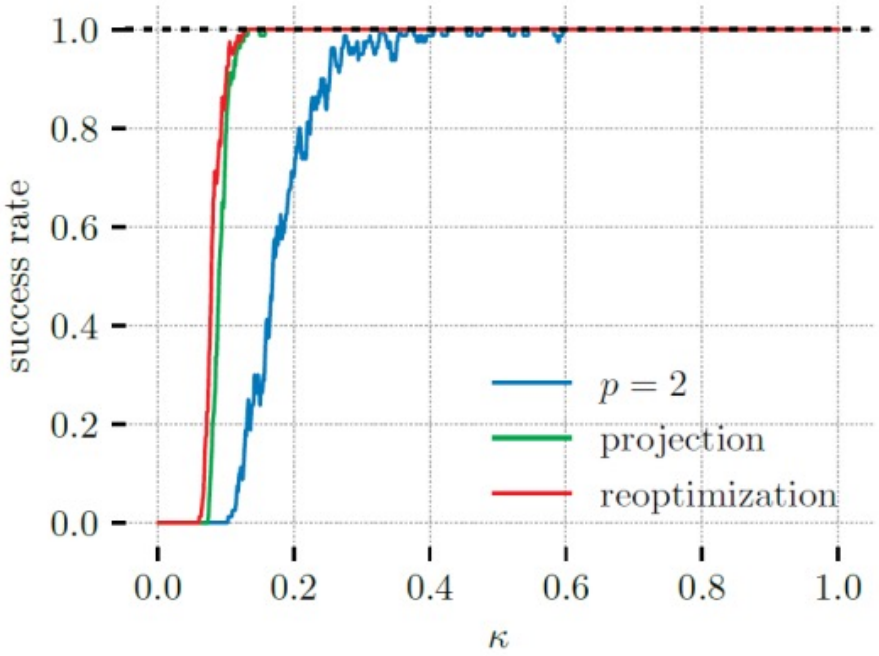
Conceptual takeaways:

- Low-dimensional nonconvex relaxations (cheap and often work!)
- Going beyond RIP: structured “perturbations”

Open questions

- **Conjecture [arbitrary GT]**: Relaxing to $k = \ell + 1$ is enough.
- **Conjecture [isotropic GT]**: Relaxing is not necessary.

- Many other localization problems (trajectory localization, inverse kinemetics, …)
- Incomplete graphs (random, expanders, …)



Conclusion

Conceptual takeaways:

- Low-dimensional nonconvex relaxations (cheap and often work!)
- Going beyond RIP: structured “perturbations”

Open Questions:

- **Conjecture** [**arbitrary GT**]: Relaxing to $k = \ell + 1$ is enough.
- **Conjecture** [**isotropic GT**]: Relaxing is not necessary.

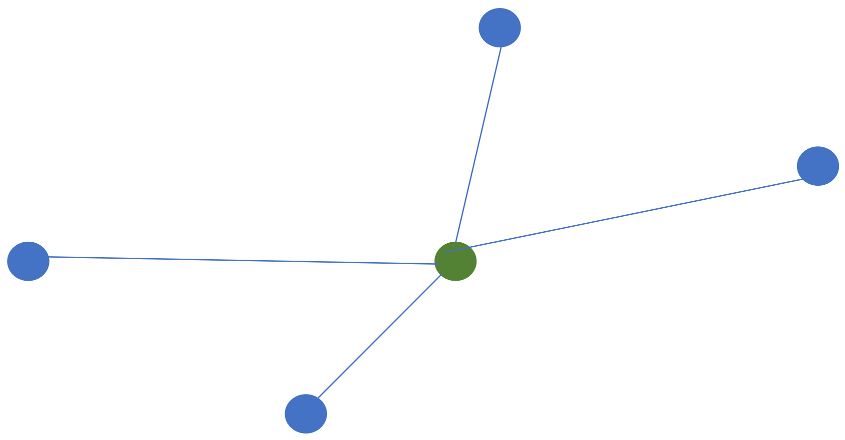
- Incomplete graphs (random, expanders, …)
- Many other localization problems (trajectory localization, inverse kinemetics, …)
- More general theory to analyze landscapes?

Appendix

SNL with landmarks

$$\min \sum_i \left(\|z - z_i\|^2 - d_i^2\right)^2, \qquad d_i = \|z^* - z_i^*\|$$

over $z \in \mathbb{R}^\ell$



SNL with landmarks

$$\min \sum_i \left(\|z - z_i\|^2 - d_i^2\right)^2, \qquad d_i = \|z^* - z_i^*\|$$

over $z \in \mathbb{R}^\ell$

Landscape is not benign in general.

SNL with landmarks

$$\min_{\text{over } z \in \mathbb{R}^\ell} \sum_i \left(\|z - z_i\|^2 - d_i^2 \right)^2, \qquad d_i = \|z^* - z_i^*\|$$

Landscape is not benign in general.

Proposition: If relax to $k = \ell + 1$, the landscape is benign.

Hubs

Theorem [isotropic GT]: If graph is **nearly complete**, ground truth points are isotropic and iid, and relax to

$$k \approx \ell \log(n)\,,$$

then every 2-critical point is the ground truth.

The **hub** of a graph is the set of vertices which are connected to all other vertices.

$$H = \text{size of hub}$$

Theorem [isotropic GT]: If ground truth points are isotropic and iid, and relax to

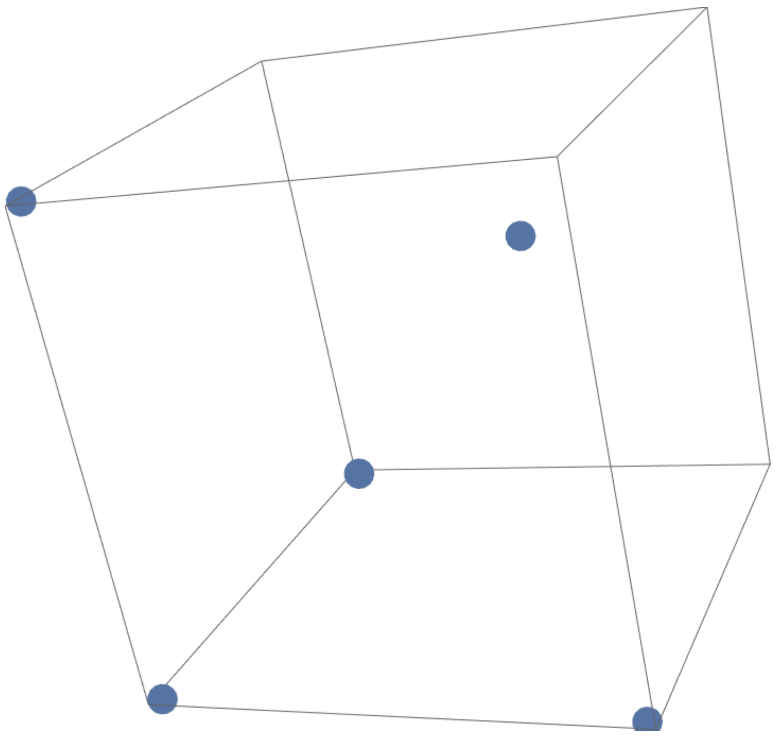
$$k \approx \text{poly}(n - H)\ell \log(n)\,,$$

then every 2-critical point is the ground truth.

Counterexamples

Minima number of points to have spurious local minima?

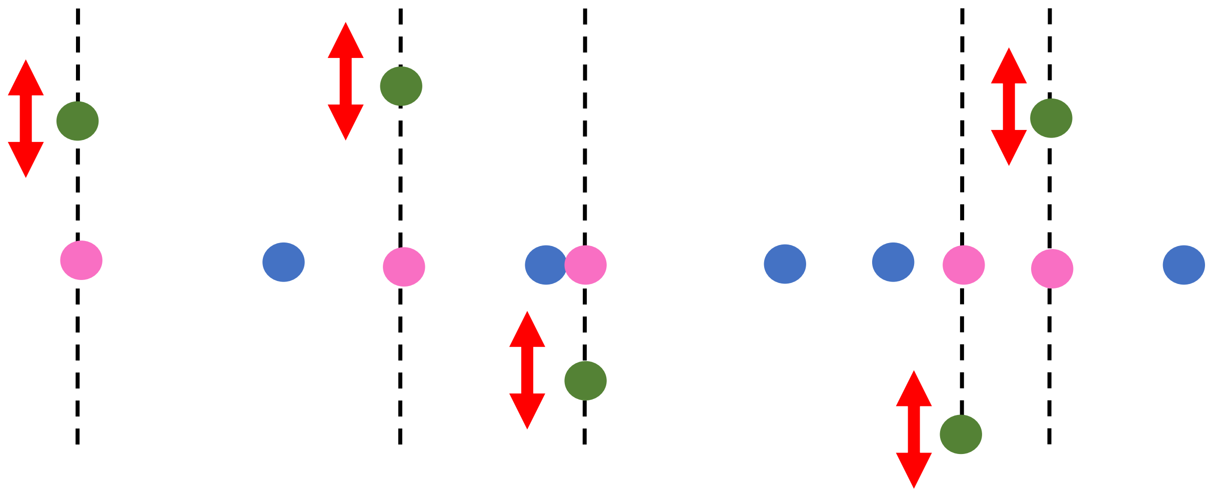
$$n = \ell + 2 \text{ (for } \ell \geq 5)$$



Some ideas from the proof

If relax enough, many ways to perturb this way

Use **eigenvalue interlacing** to argue that a good one exists, if relax enough



Some ideas from the proof

If relax enough, many ways to perturb this way

Use **eigenvalue interlacing** to argue that a good one exists, if relax enough

For **isotropic GT,** $k \approx \ell \log(n)$, similar descent direction

Randomize over descent directions (instead of eigenvalue interlacing)

Can we apply Kirwan convexity, or similar?

$$\min \|\Delta(ZZ^\top - Z_* Z_*^\top)\|^2 \text{ over } Z \in \mathbb{R}^{n\times\ell}$$

$\min \|ZZ^\top - Z_* Z_*^\top\|^2$ over $Z \in \mathbb{R}^{n\times\ell}$ (with $\operatorname{trace}(ZZ^\top) = 1$)

- Kirwan: $K = U(n)$ acts on projective space $\mathbb{P}\left(\mathbb{C}^{n\times\ell}\right)$

No index-1 critical points if relax to $k = \ell + 2$?

- Seems to be a common phenomenon when relaxing dimension

[Index = number of negative eigenvalues of Hessian]