Sensor network localization has benign landscape under mild rank relaxation

February 11, 2025

Chris Criscitiello
with
Andrew McRae, Quentin Rebjock, Nicolas Boumal

OPTIM, Chair of Continuous Optimization
Institute of Mathematics, EPFL

EPFL

Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

← Solve with local algorithm (gradient descent, trust regions, …)

Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

← Solve with local algorithm (gradient descent, trust regions, …)

When can we have guarantees for global optimality?

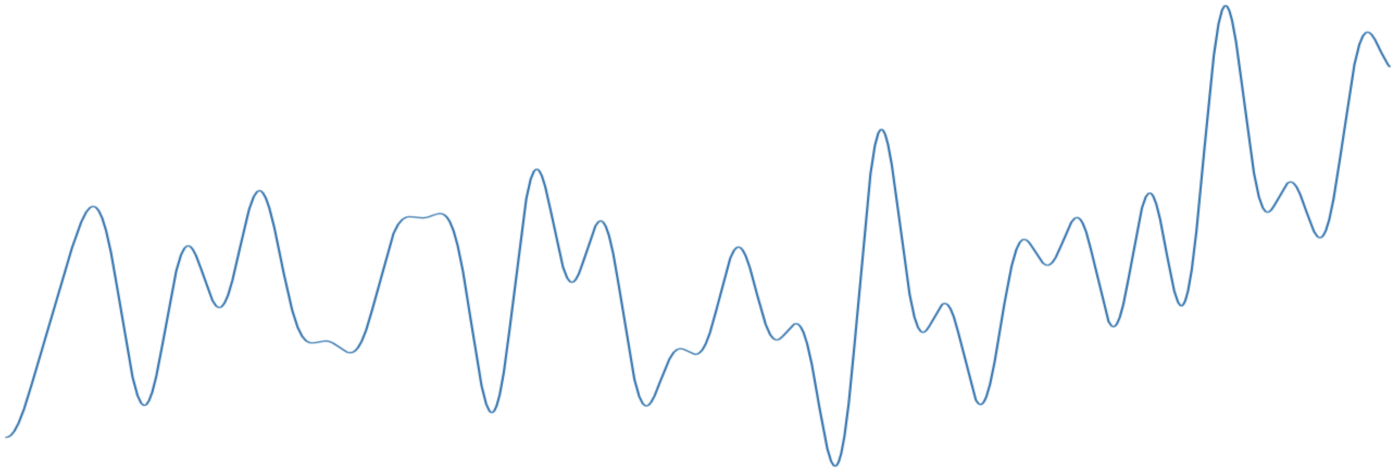
Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

Solve with local algorithm (gradient descent, trust regions, …)

When can we have guarantees for global optimality?

Need **structure!**



Global guarantees for optimization

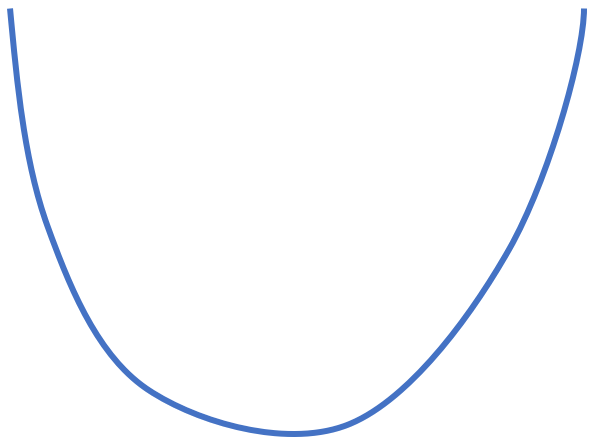
$$\min_{x \in \mathbb{R}^d} f(x)$$

Solve with local algorithm (gradient descent, trust regions, …)

When can we have guarantees for global optimality?

Need **structure!**

Convexity



Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

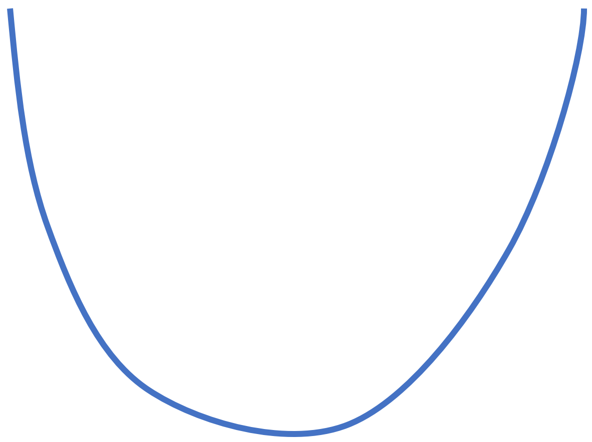
Solve with local algorithm (gradient descent, trust regions, ...)

When can we have guarantees for global optimality?

Need **structure!**

$$\nabla f(x) = 0$$

Convexity (implies all 1-critical points are optimal)



Global guarantees for optimization

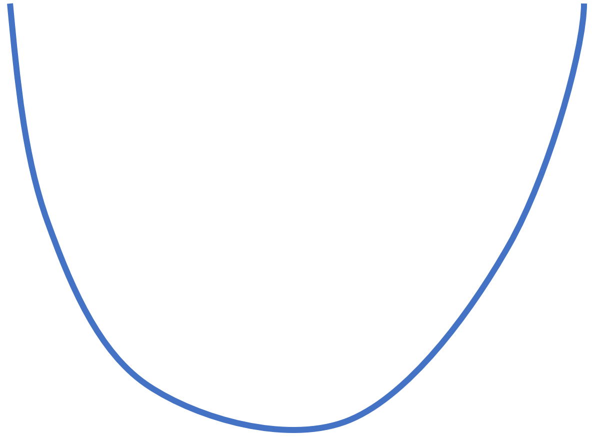
$$\min_{x \in \mathbb{R}^d} f(x)$$ ← Solve with local algorithm (gradient descent, trust regions, …)

When can we have guarantees for global optimality?

Need **structure!**

$$\nabla f(x) = 0$$

~~**Convexity**~~ (implies all 1-critical points are optimal)



Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

Solve with local algorithm
(gradient descent, trust regions, …)

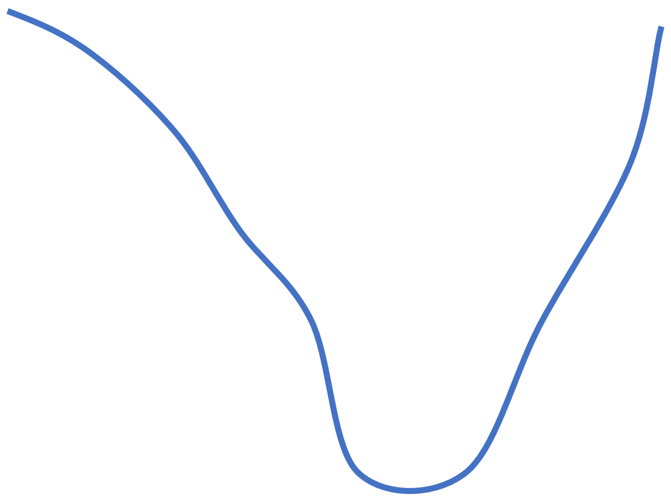
When can we have guarantees for global optimality?

Need **structure!**

$$\nabla f(x) = 0$$

Convexity (implies all 1-critical points are optimal)

Invexity (definition: all 1-critical points are optimal)



Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$ ← Solve with local algorithm (gradient descent, trust regions, …)

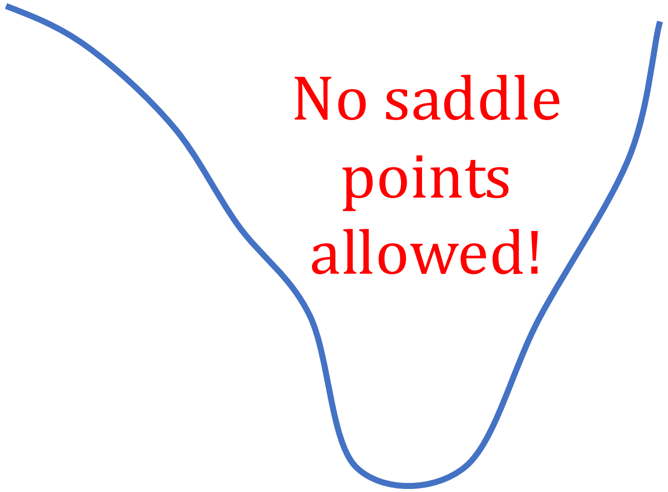
When can we have guarantees for global optimality?

Need **structure!**

$\nabla f(x) = 0$

Convexity (implies all 1-critical points are optimal)

~~**Invexity**~~ (definition: all 1-critical points are optimal)



Global guarantees for optimization

$$\min_{x \in \mathbb{R}^d} f(x)$$

Solve with local algorithm (gradient descent, trust regions, …)

When can we have guarantees for global optimality?

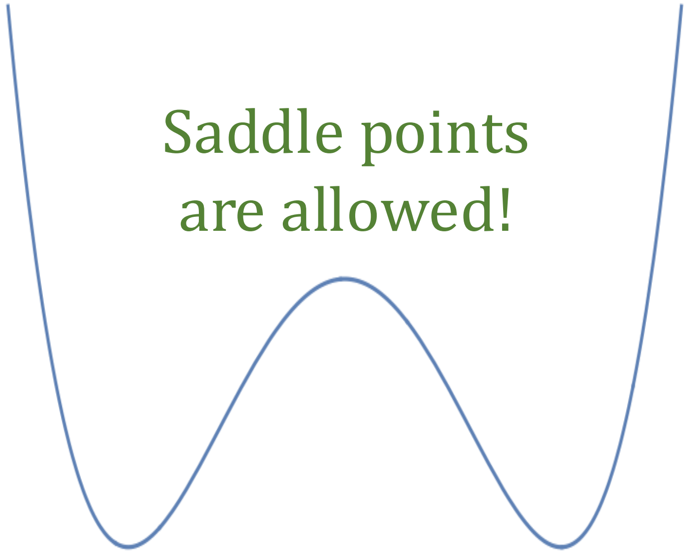
Need **structure!**

$$\nabla f(x) = 0$$

Convexity (implies all 1-critical points are optimal)

Invexity (definition: all 1-critical points are optimal)

Benign landscape (all local minima are optimal)



Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

Definition: f has a **benign landscape** if all 2-critical points are optimal:

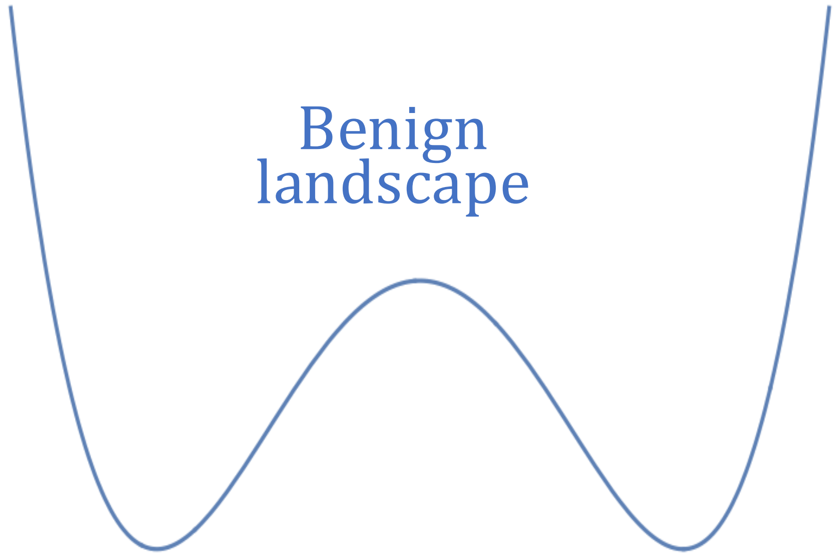
$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min

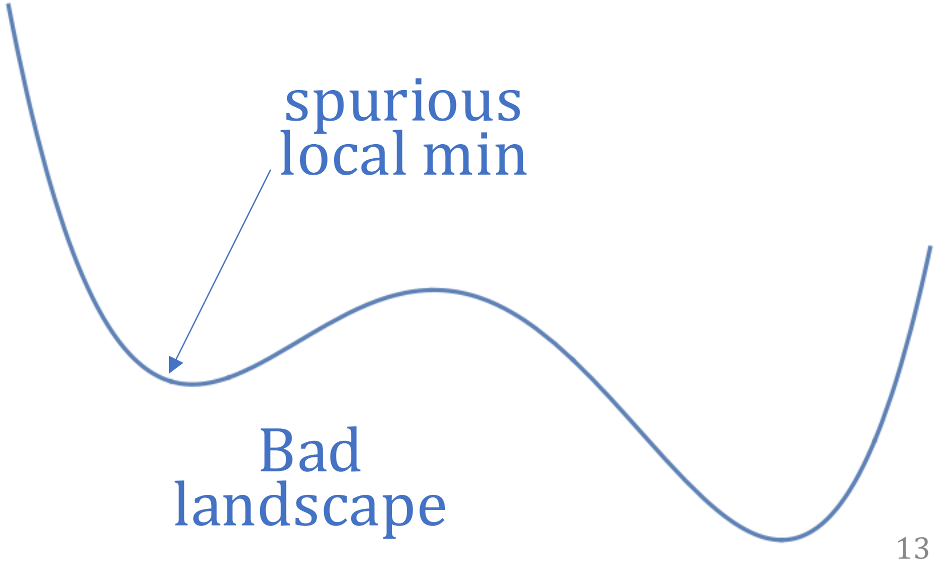
Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

Definition: f has a **benign landscape** if all 2-critical points are optimal:

$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min





Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

Definition: f has a **benign landscape** if all 2-critical points are optimal:

$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min

Why useful?

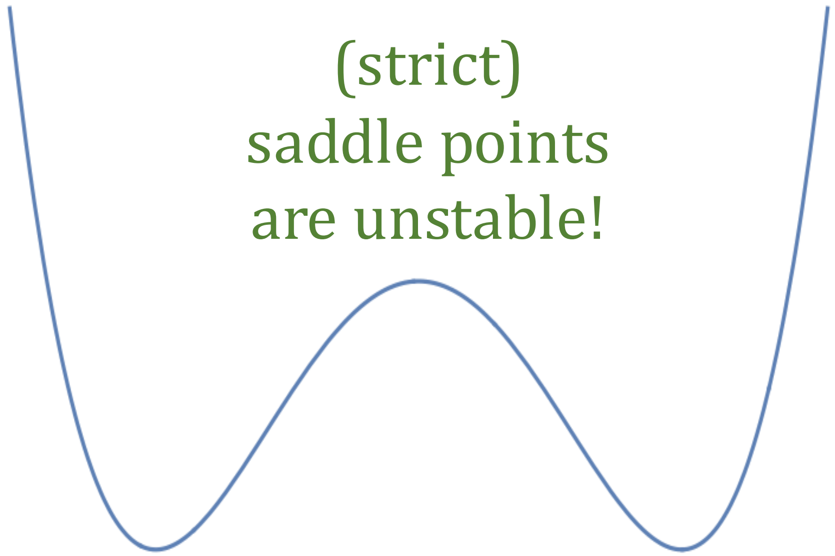
Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

Definition: f has a **benign landscape** if all 2-critical points are optimal:

$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min

Why useful?



Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

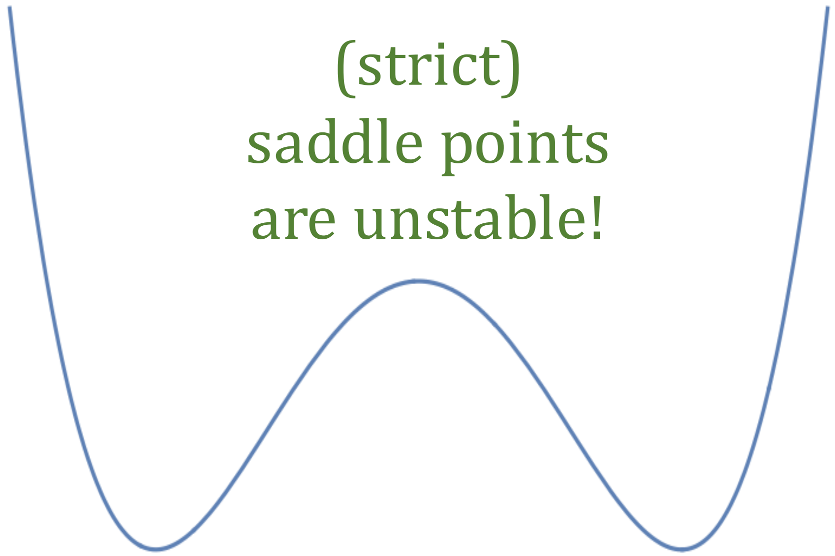
Definition: f has a **benign landscape** if all 2-critical points are optimal:

$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min

Why useful?

Local algorithm (GD, TR, …)

converges to 2-critical point (w/ prob 1)



Benign landscape

$$\min_{x \in \mathbb{R}^d} f(x)$$

Definition: f has a **benign landscape** if all 2-critical points are optimal:

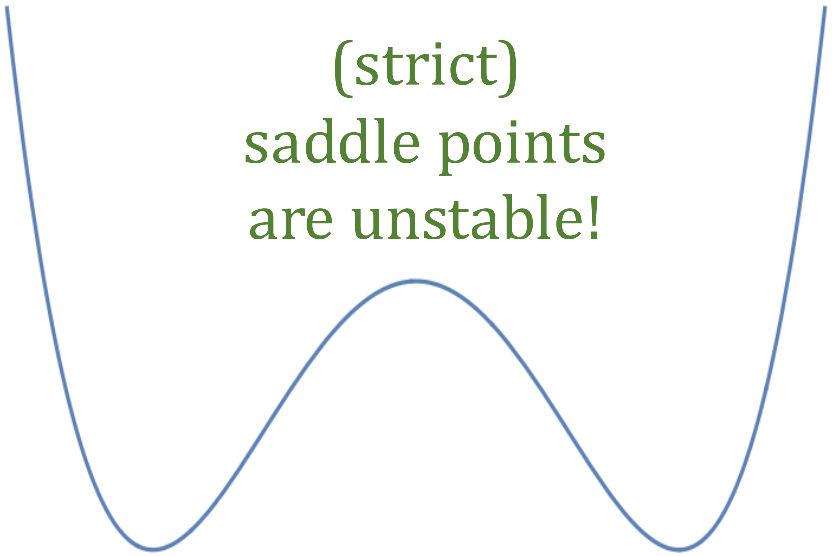
$\nabla f(x) = 0$ and $\nabla^2 f(x) \succcurlyeq 0$ implies x is a global min

Why useful?

Local algorithm (GD, TR, …)

converges to 2-critical point (w/ prob 1)

Stable manifold theorems
+
Łojasiewicz theorem



Some of my previous work …

"Negative curvature obstructs acceleration for g-convex optimization"

C, Boumal, 2022

"Curvature and complexity: Lower bounds for g-convex optimization"

C, Boumal, 2023

Invexity

"Synchronization on circles and spheres with nonlinear interaction"

C, McRae, Rebjock, Boumal, 2024

"Sensor network localization has benign landscape under relaxation"

C, McRae, Rebjock, Boumal, to be public very soon

Benign Landscape

Some of my previous work ...

"Negative curvature obstructs acceleration for g-convex optimization"

C, Boumal, 2022

"Curvature and complexity: Lower bounds for g-convex optimization"

C, Boumal, 2023

Invexity

"Synchronization on circles and spheres with nonlinear interaction"

C, McRae, Rebjock, Boumal, 2024

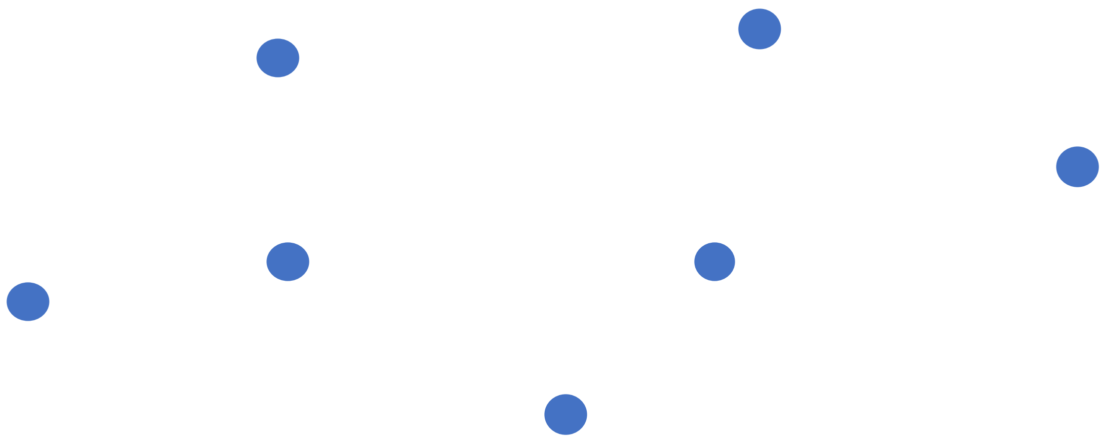
"Sensor network localization has benign landscape under relaxation"

C, McRae, Rebjock, Boumal, to be public very soon

Benign Landscape

The problem: SNL

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

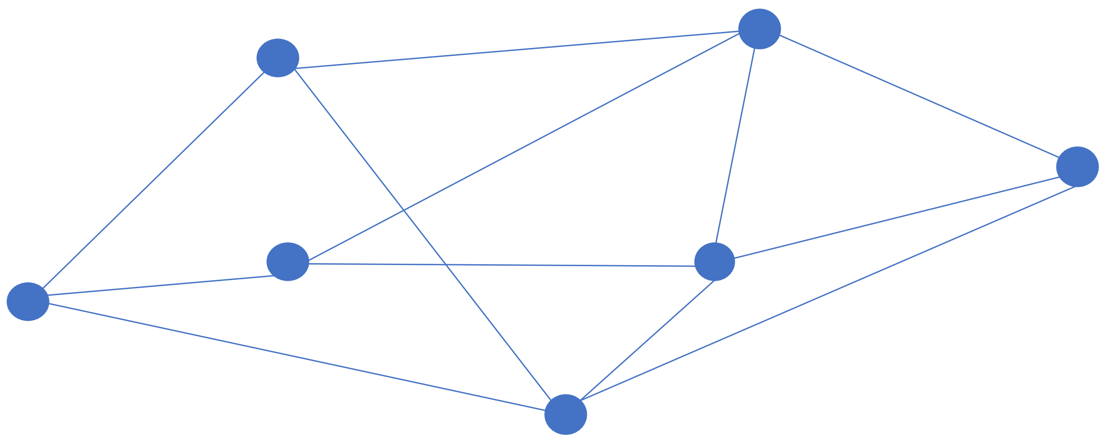


The problem: SNL

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

Know a subset of the pairwise distances (measurements)

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



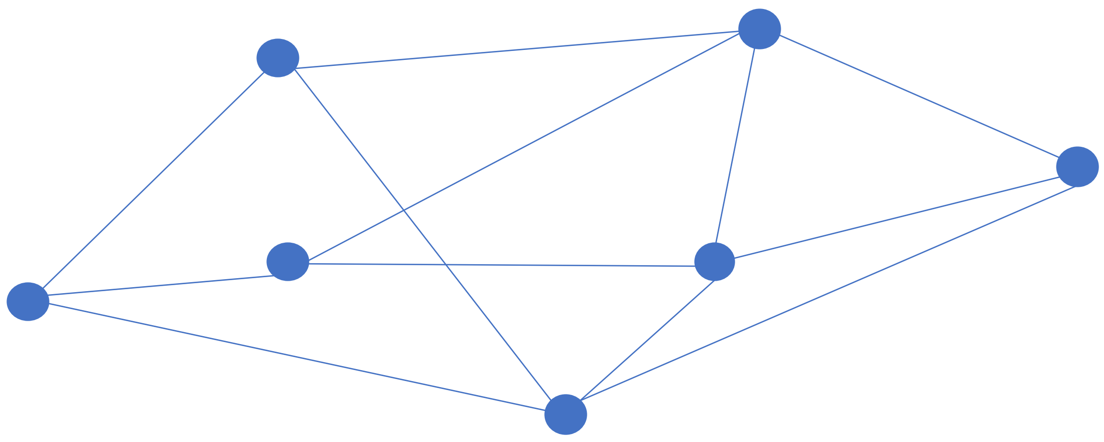
The problem: SNL

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)



The problem: SNL

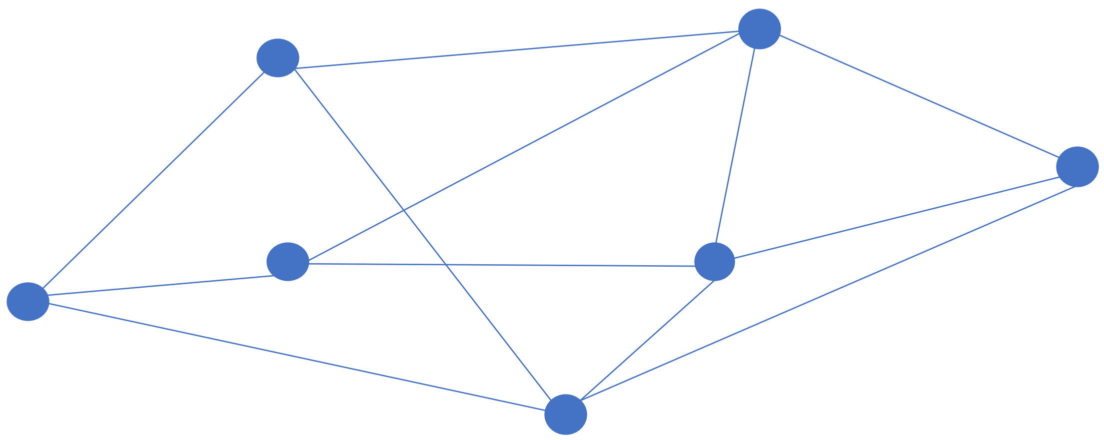
n unknown points $z_1^*, z_2^*, \ldots, z_n^*$ in $\mathbb{R}^\ell$.

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)

Sensor network localization (SNL) – Torgerson ’58, Shepard ‘62



Applications

Robotics (**sensor network localization**), ℓ = dimension = 2,3

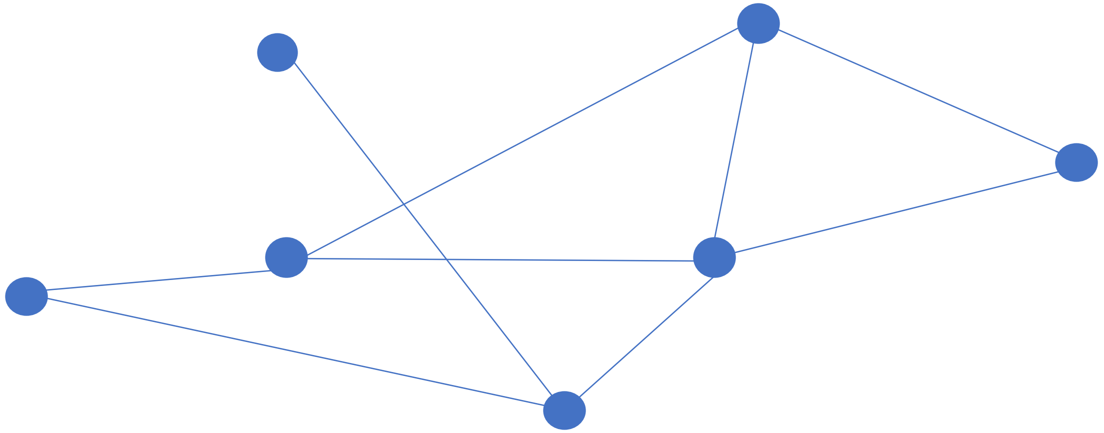
Molecular conformation

Data analysis (metric **multidimensional scaling**)

Graph theory (rigidity)

The problem: SNL

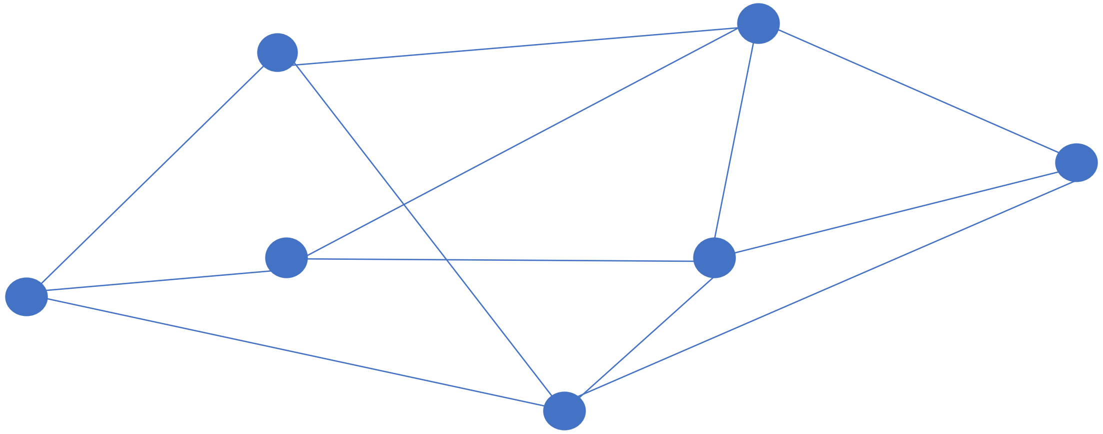
Need enough distance measurements for well-posedness (*global rigidity*)



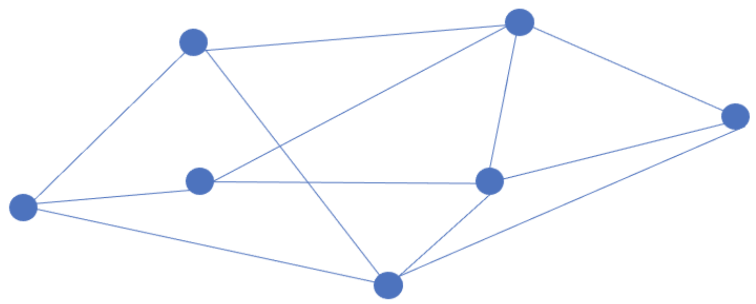
The problem: SNL

Need enough distance measurements for well-posedness (*global rigidity*)

NP-hard! (“Euclidean distance geometry and applications” -- Liberti, et al)



The problem: SNL



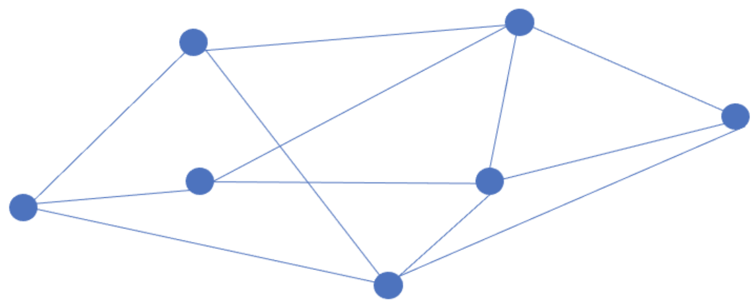
Need enough distance measurements for well-posedness (*global rigidity*)

NP-hard! (“Euclidean distance geometry and applications” -- Liberti, et al)

Common approach: solve with **SDP** (semidefinite program)

Polynomial time, under additional assumptions (*universal rigidity*)

The problem: SNL



Need enough distance measurements for well-posedness (*global rigidity*)

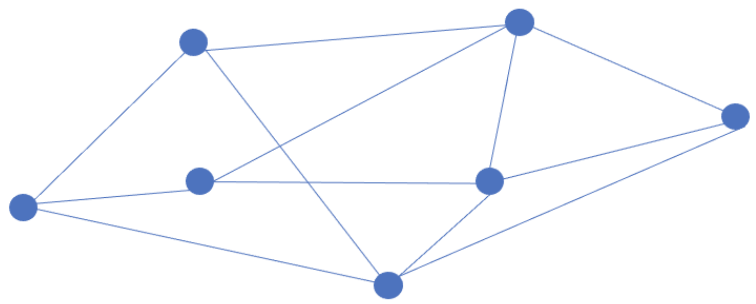
NP-hard! (“Euclidean distance geometry and applications” -- Liberti, et al)

Common approach: solve with **SDP** (semidefinite program)

Polynomial time, under additional assumptions (*universal rigidity*)

“Theory of semidefinite programming for Sensor Network Localization” -- So, Ye ‘06

The problem: SNL



Need enough distance measurements for well-posedness (*global rigidity*)

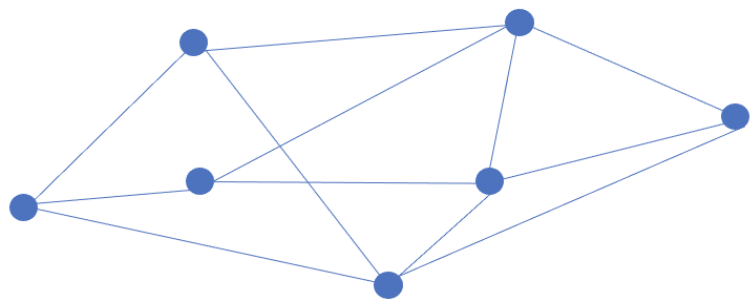
NP-hard! ("Euclidean distance geometry and applications" -- Liberti, et al)

Common approach: solve with **SDP** (semidefinite program)

Polynomial time, under additional assumptions (*universal rigidity*)

Big drawback: **SDP** involves dense $(n + \ell) \times (n + \ell)$ matrices

The problem: SNL



Need enough distance measurements for well-posedness (*global rigidity*)

NP-hard! (“Euclidean distance geometry and applications” -- Liberti, et al)

Common approach: solve with **SDP** (semidefinite program)

Polynomial time, under additional assumptions (*universal rigidity*)

Big drawback: **SDP** involves dense $(n + \ell) \times (n + \ell)$ matrices

quadratic in # of points

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"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

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"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

But 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"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

But nonconvex! How bad?

Little is known!

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

"s-stress"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

But nonconvex! How bad?

Little is known! Even simplest possible question was unanswered:

Open Q: Is landscape benign for complete graph (all distances known)?

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"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

But nonconvex! How bad?

Little is known! Even simplest possible question was unanswered:

Open Q: Is landscape benign for complete graph (all distances known)?

Asked by Malone & Trosset 2000, Parhizkar 2013, …

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

"s-stress"

Kruskal '64

Variable size = $n\ell$ = **linear in # of points**

But nonconvex! How bad?

Little is known! Even simplest possible question was unanswered:

Open Q: Is landscape benign for complete graph (all distances known)?

Computationally easy, via Eigenvalue decomposition

Asked by Malone & Trosset 2000, Parhizkar 2013, …

Synthetic experiment, complete graph

Open Q: Is s-stress landscape always benign for complete graph?

Synthetic experiment, complete graph

Open Q: Is s-stress landscape always benign for complete graph?

(1) Choose ground truths $z_1^*, z_2^*, \dots, z_n^*$ at random (normal iid)
(2) Run gradient descent/trust regions/etc.
(3) Find global min?
(4) Repeat

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

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

"s-stress"

Synthetic experiment, complete graph

Open Q: Is s-stress landscape always benign for complete graph?

(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

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

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

"s-stress"

Always finds global min!

Synthetic experiment, complete graph

Open Q: Is s-stress landscape always benign for complete graph?

(1) Choose ground truths $z_1^*, z_2^*, \dots, z_n^*$ at random (normal iid)
(2) Run gradient descent/trust regions/etc.
(3) Find global min?
(4) Repeat

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

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

"s-stress"

Always finds global min!

Maybe landscape is always benign?

(for all ground truth configurations $z_1^*, z_2^*, \dots, z_n^*$, all 2-critical points of s-stress are global minima?)

Synthetic experiment, complete graph

Open Q: Is s-stress landscape always benign for complete graph?

(1) Choose ground truths $z_1^*, z_2^*, \dots, z_n^*$ at random (normal iid)
(2) Run gradient descent/trust regions/etc.
(3) Find global min?
(4) Repeat

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

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

"s-stress"

Always finds global min!

Maybe landscape is always benign?

(for all ground truth configurations $z_1^*, z_2^*, \dots, z_n^*$, all 2-critical points of s-stress are global minima?)

NO!

Counterexample

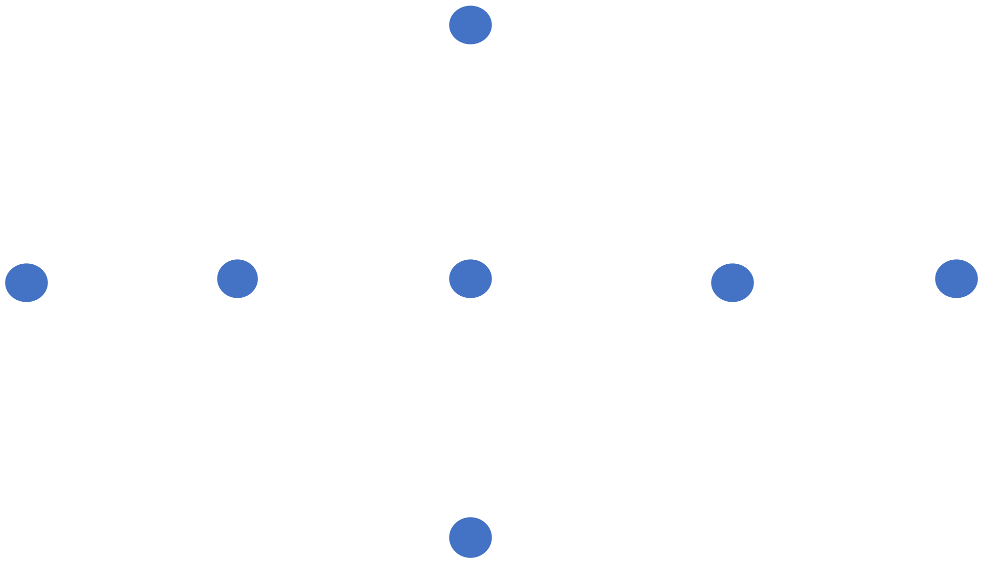
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

"s-stress"

s-stress can have spurious strict local minima!

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



Spurious configuration $z_1, z_2, \ldots$

Counterexample

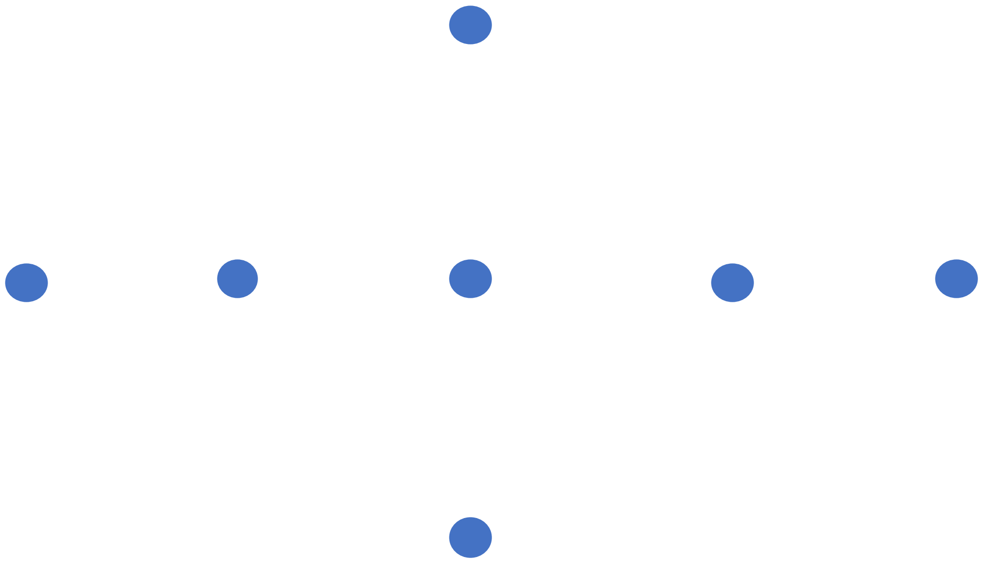
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

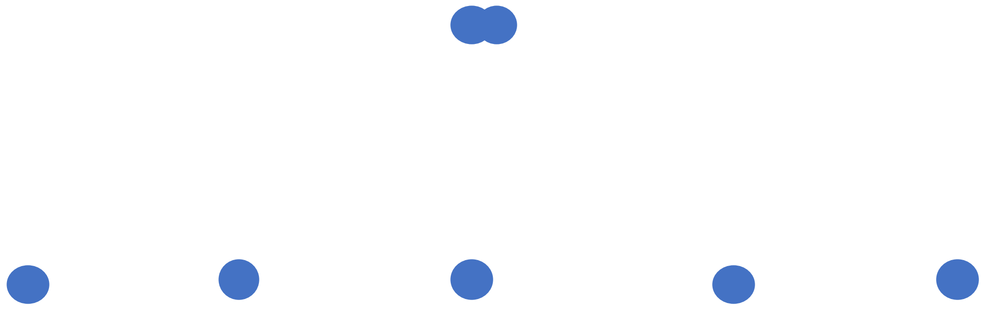
"s-stress"

s-stress can have spurious strict local minima!

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



Spurious configuration $z_1, z_2, \dots$



Counterexample

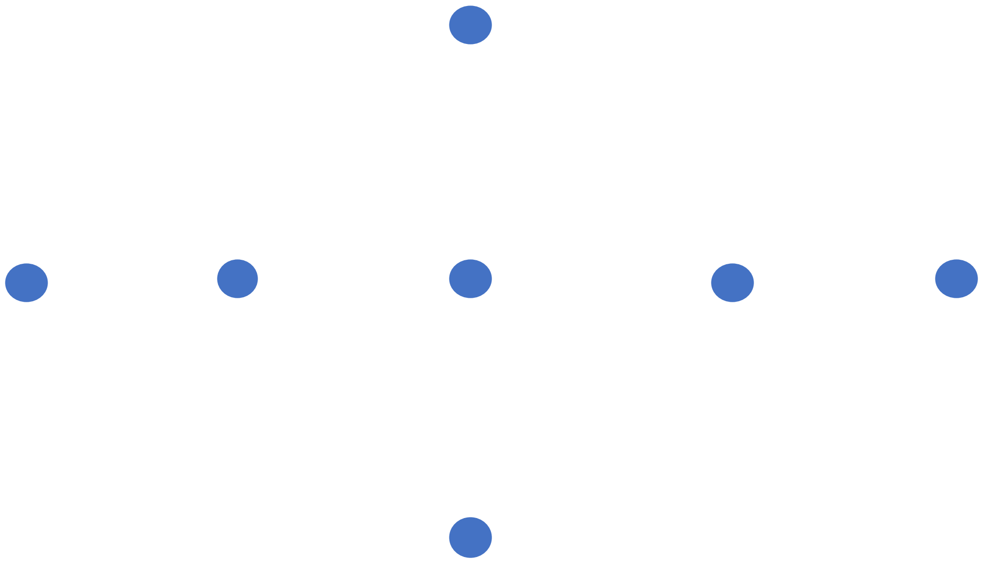
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

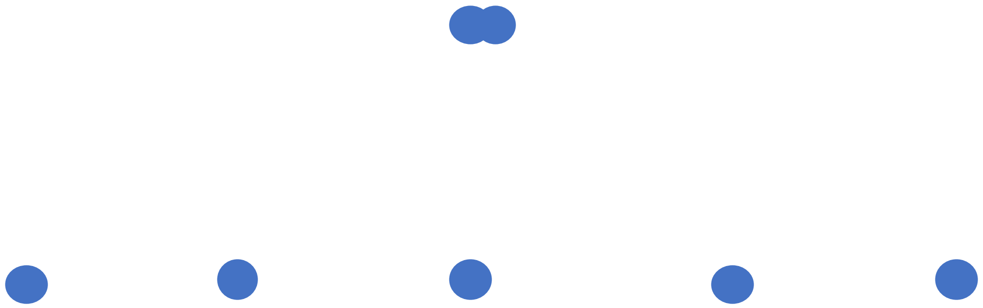
"s-stress"

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

Counterexample

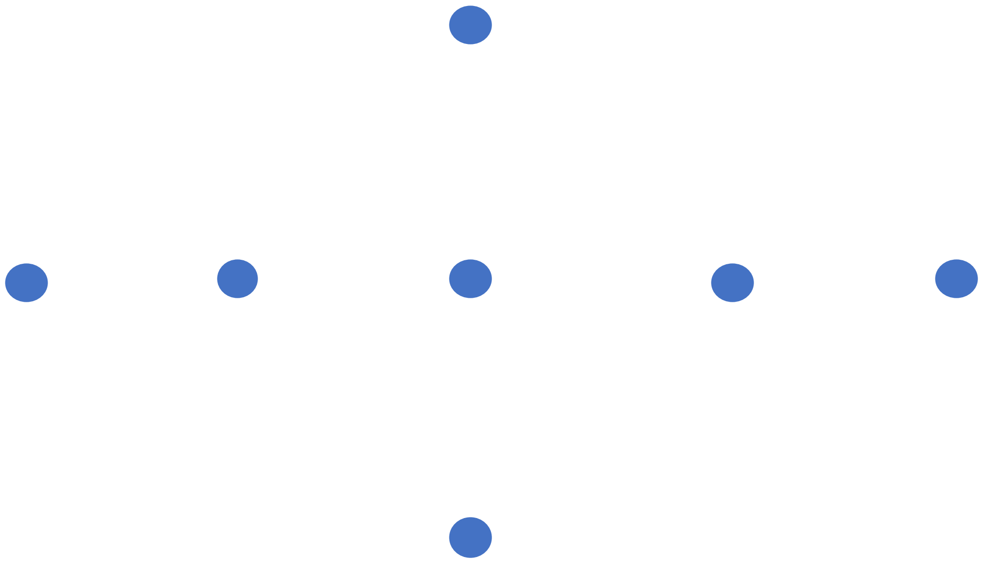
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

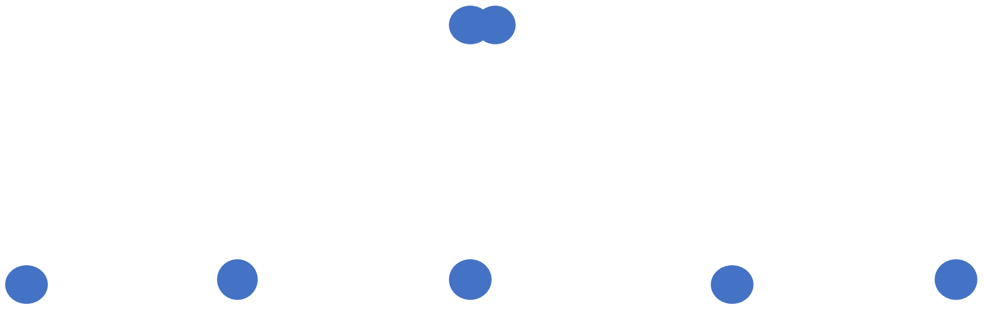
"s-stress"

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

Set of ground truths with spurious local minima has positive measure

Counterexample

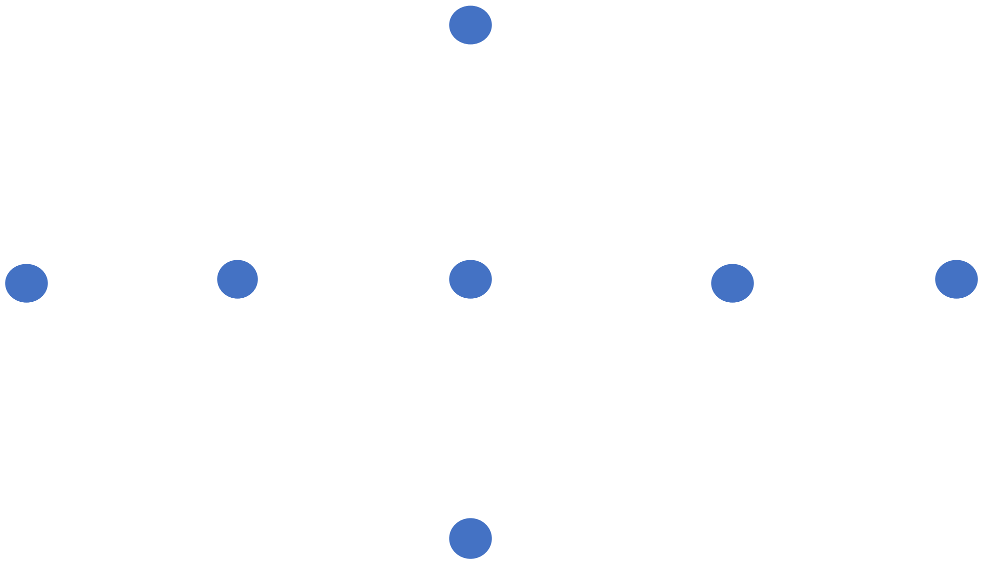
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

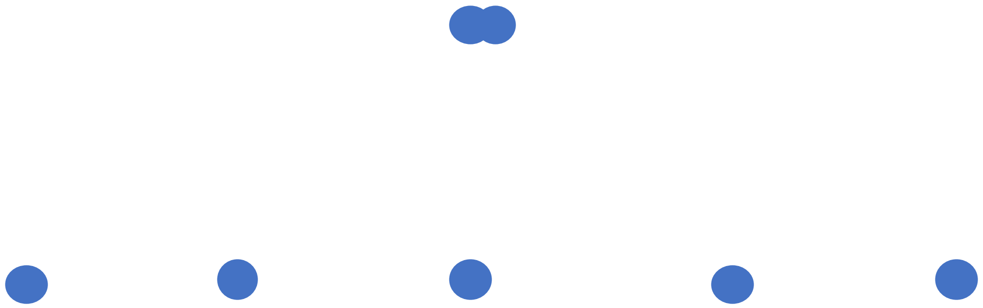
"s-stress"

s-stress can have spurious strict local minima!

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



Spurious configuration $z_1, z_2, \ldots$



Hmm … what should we do?

Counterexample

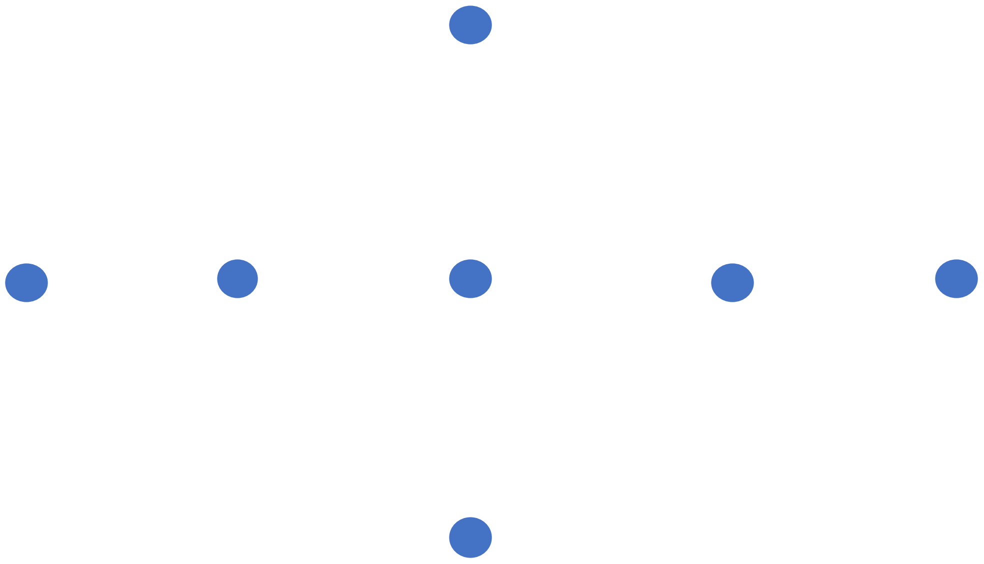
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

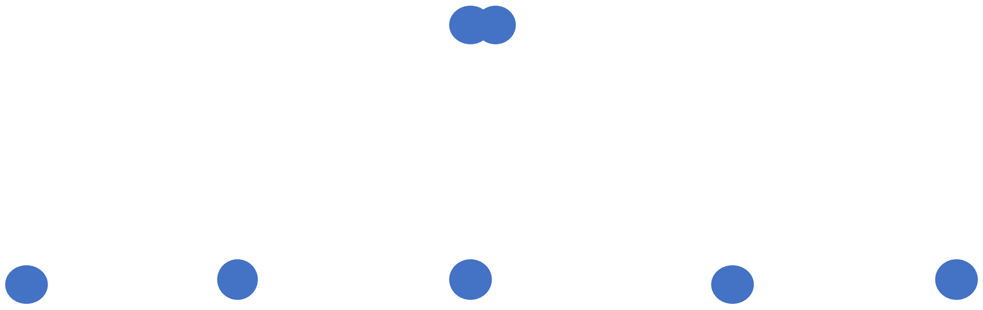
"s-stress"

s-stress can have spurious strict local minima!

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



Spurious configuration $z_1, z_2, \ldots$



Optimize over points in $\mathbb{R}^3$

Counterexample

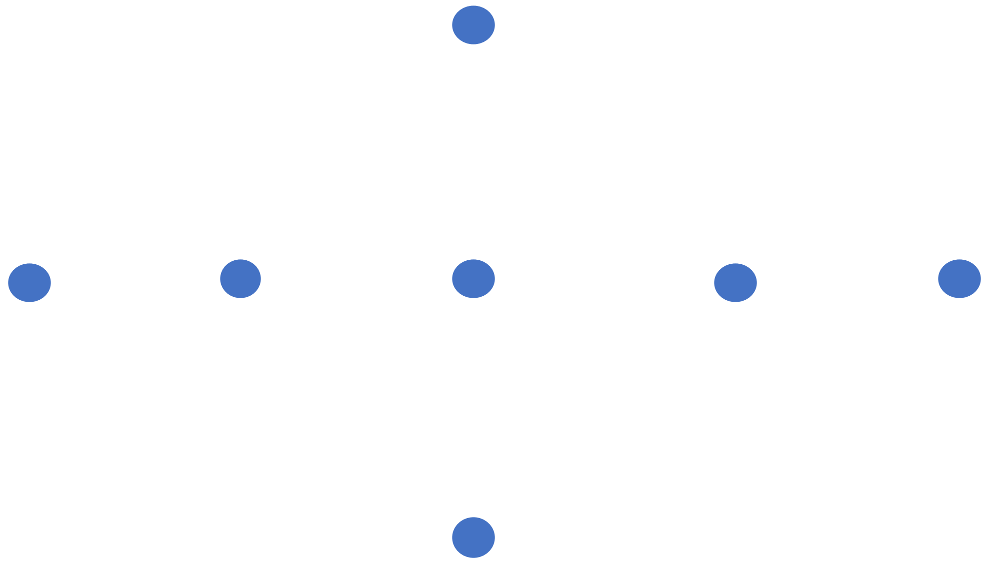
$$\min \sum_{ij\in E} \left(\|z_i - z_j\|^2 - d_{ij}^2\right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

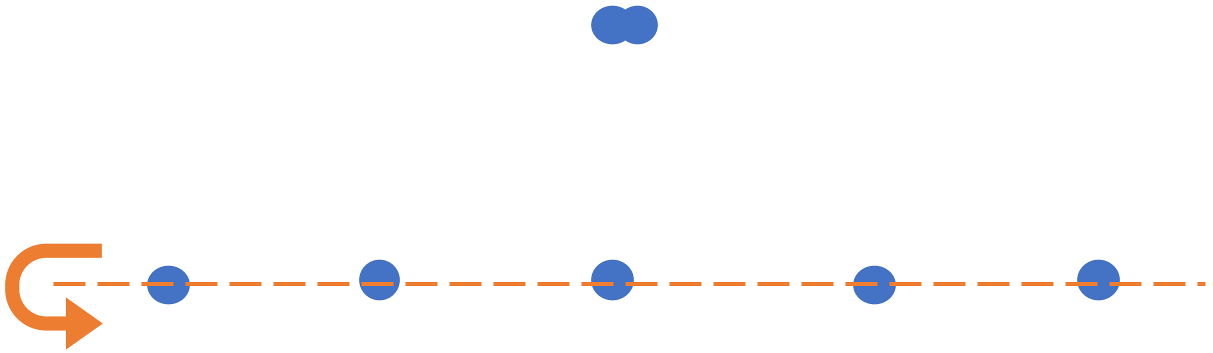
"s-stress"

s-stress can have spurious strict local minima!

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



Spurious configuration $z_1, z_2, \ldots$



Optimize over points in $\mathbb{R}^3$

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

Relax to dimension $k > \ell$

Nonconvex relaxation

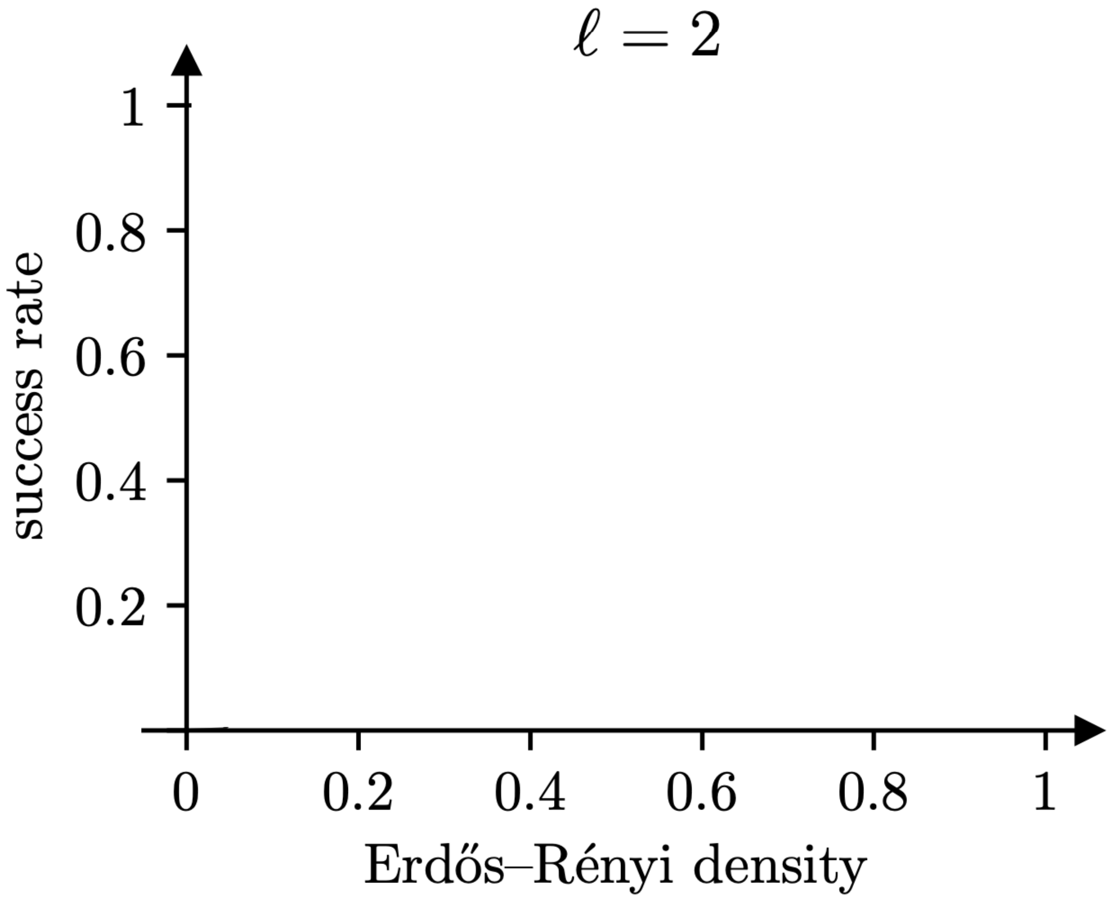
$$\min \sum_{ij\in E} \left(\|z_i - z_j\|^2 - d_{ij}^2\right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

Relax to dimension $k > \ell$

Experiment:

- $n = 50, \ell =$ dimension $= 2$
- Ground truth = iid Gaussian points
- Graph = ER
- Run TR from random initialization



Nonconvex relaxation

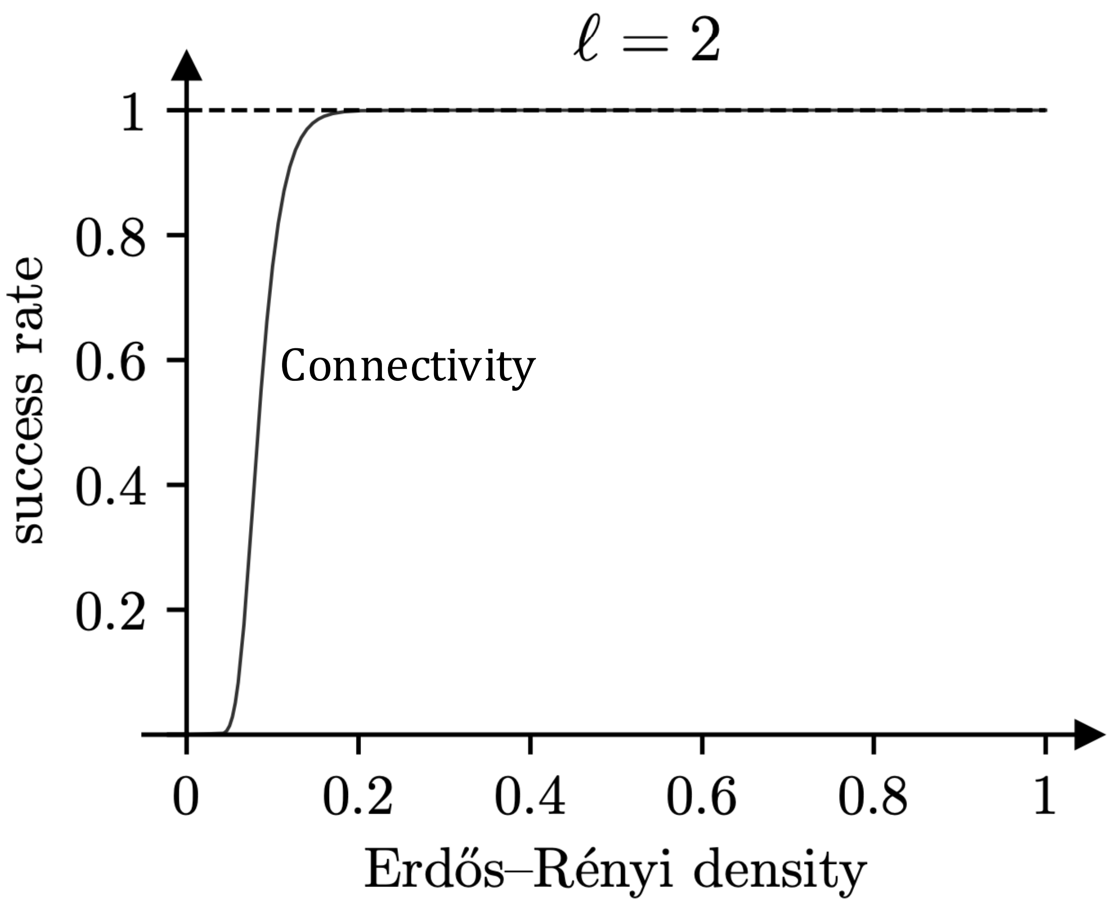
$$\min \sum_{ij\in E} \left(\|z_i - z_j\|^2 - d_{ij}^2\right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

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

Relax to dimension $k > \ell$

Experiment:

- $n = 50, \ell =$ dimension $= 2$
- Ground truth = iid Gaussian points
- Graph = ER
- Run TR from random initialization

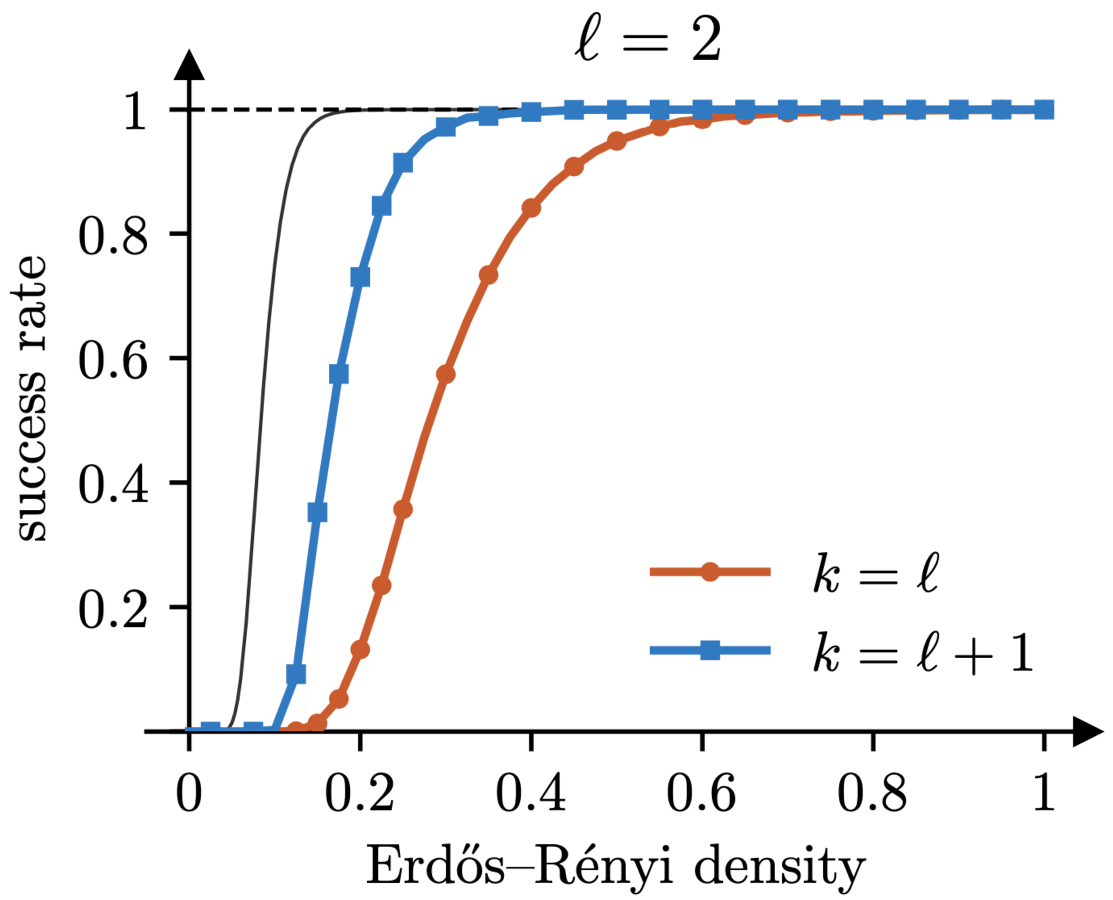


Nonconvex relaxation

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

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

Relax to dimension $k > \ell$



Experiment:

- $n = 50, \ell =$ dimension $= 2$
- Ground truth = iid Gaussian points
- Graph = ER
- Run TR from random initialization

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

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

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

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

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

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.

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

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.

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

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.

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

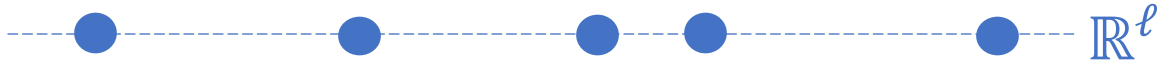
$$k \approx \ell + \log(n),$$

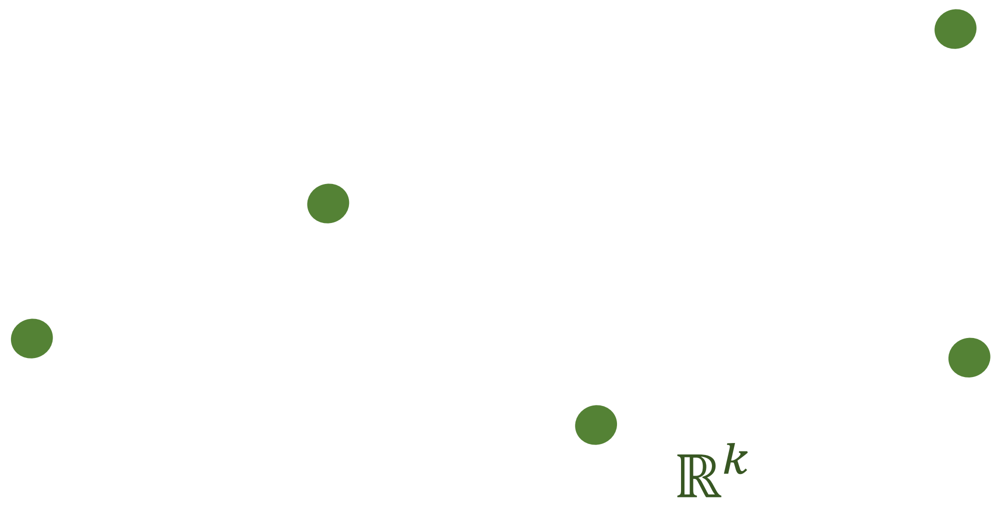
then every 2-critical point is the ground truth, w.h.p.

Proof sketch: isotropic

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

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





Proof sketch: isotropic

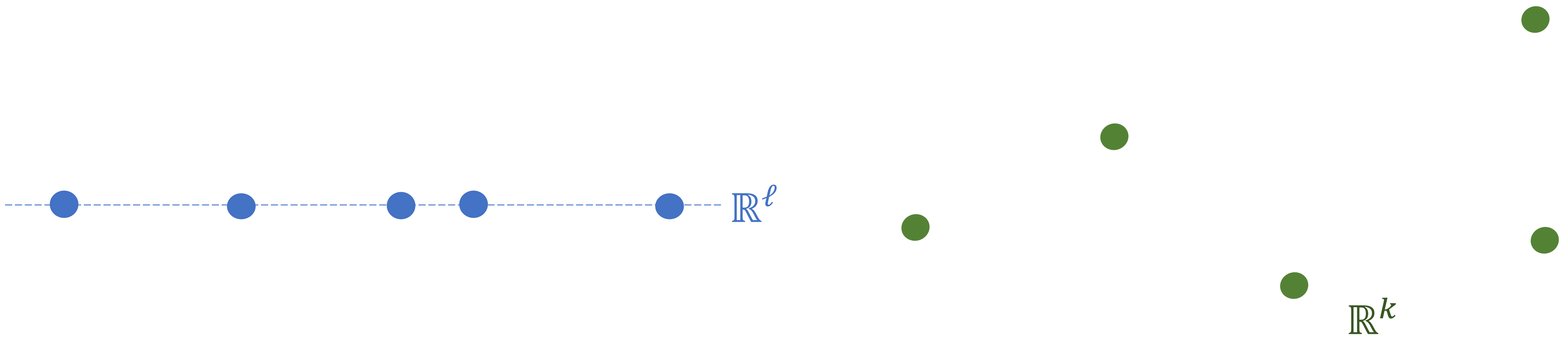
Ground truth $z_1^*, z_2^*, \ldots$ in dimension ℓ

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

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

Goal: **perturb** 1-critical configuration to decrease cost



Proof sketch: isotropic

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

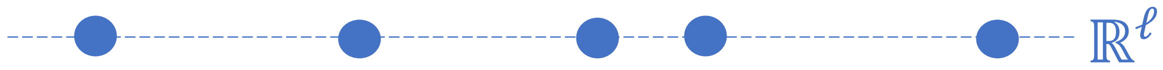
1-critical configuration in dimension $k > \ell$

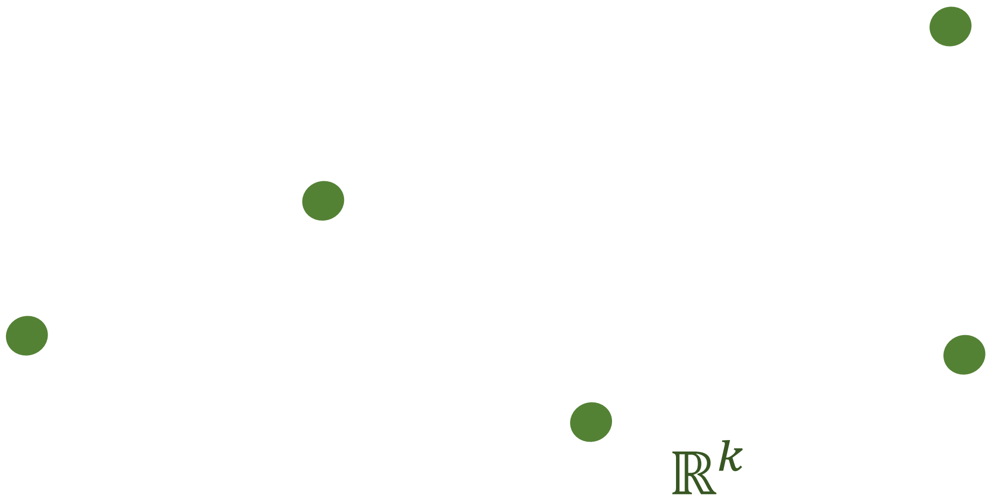
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How?





Proof sketch: isotropic

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

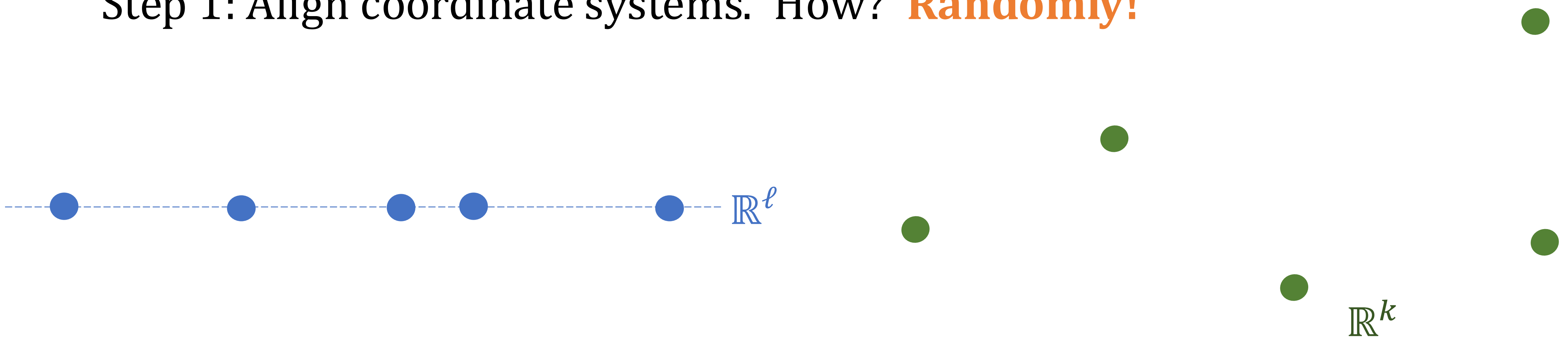
1-critical configuration in dimension $k > \ell$

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**



Proof sketch: isotropic

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

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

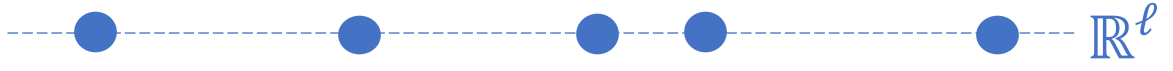
$$\min \sum_{ij \in E} \left(\|z_i - z_j\|^2 - d_{ij}^2 \right)^2, \qquad d_{ij} = \|z_i^* - z_j^*\|$$

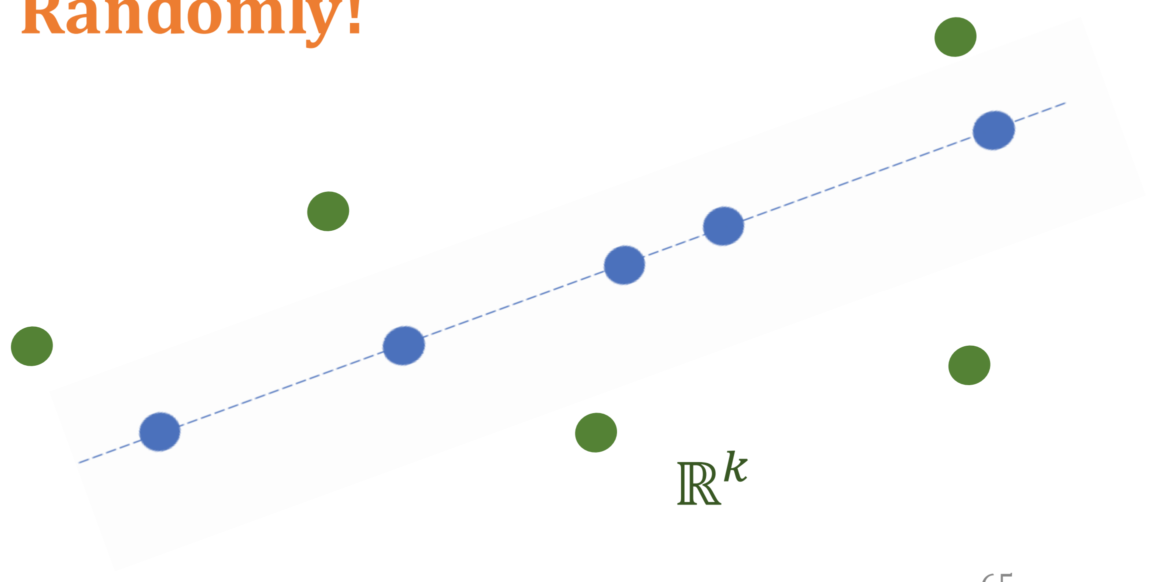
over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$

"s-stress"

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**





Proof sketch: isotropic

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

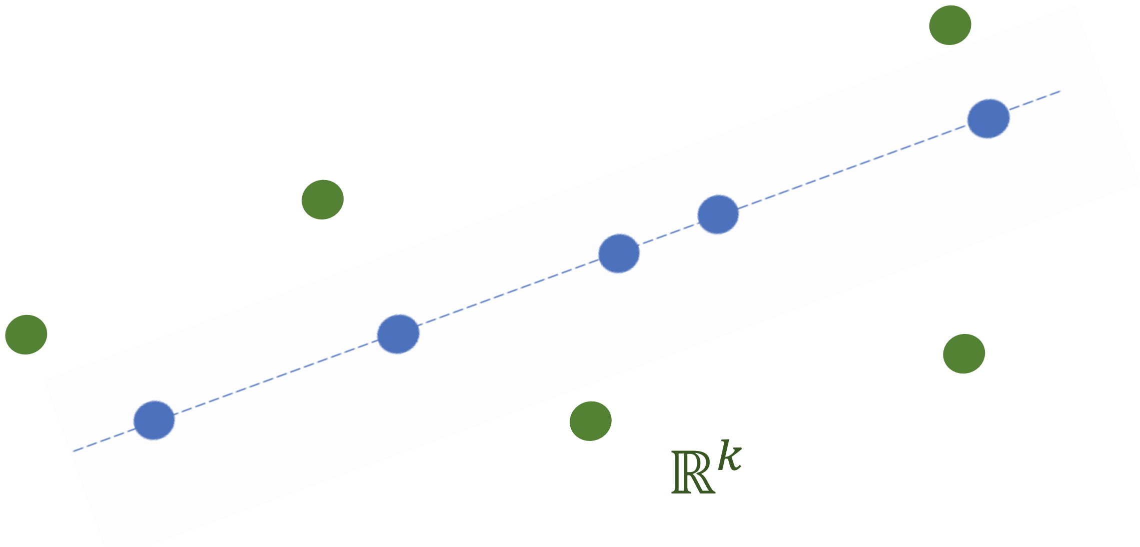
1-critical configuration in dimension $k > \ell$

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

over $z_1, z_2, \ldots, z_n \in \mathbb{R}^k$ "s-stress"

Goal: **perturb** 1-critical configuration to decrease cost

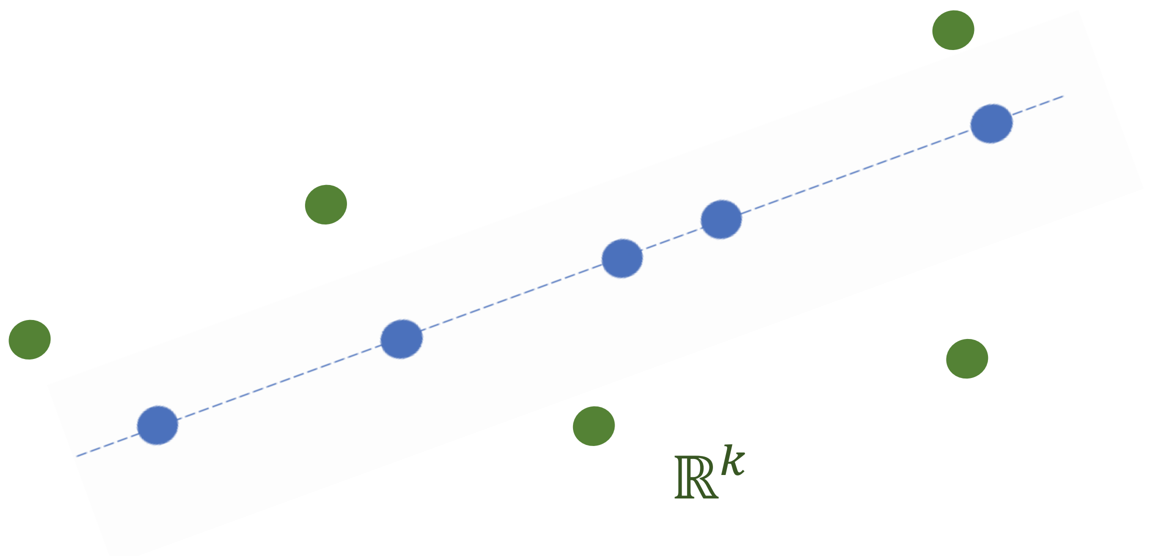
Step 1: Align coordinate systems. How? **Randomly!**



Proof sketch: isotropic

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**

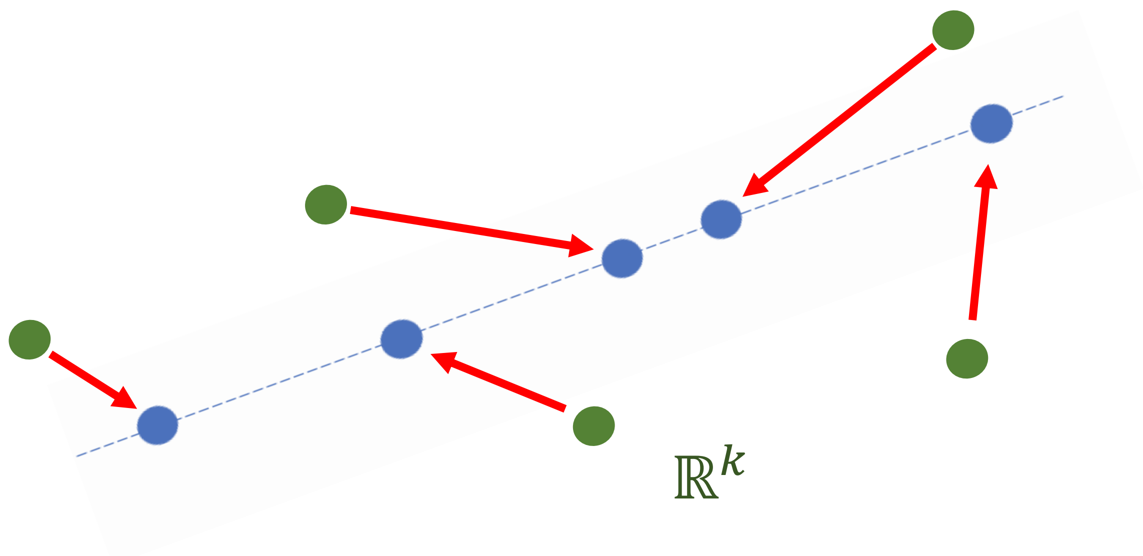


Proof sketch: isotropic

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**

Step 2: Move 1-critical configuration “toward” ground truth.



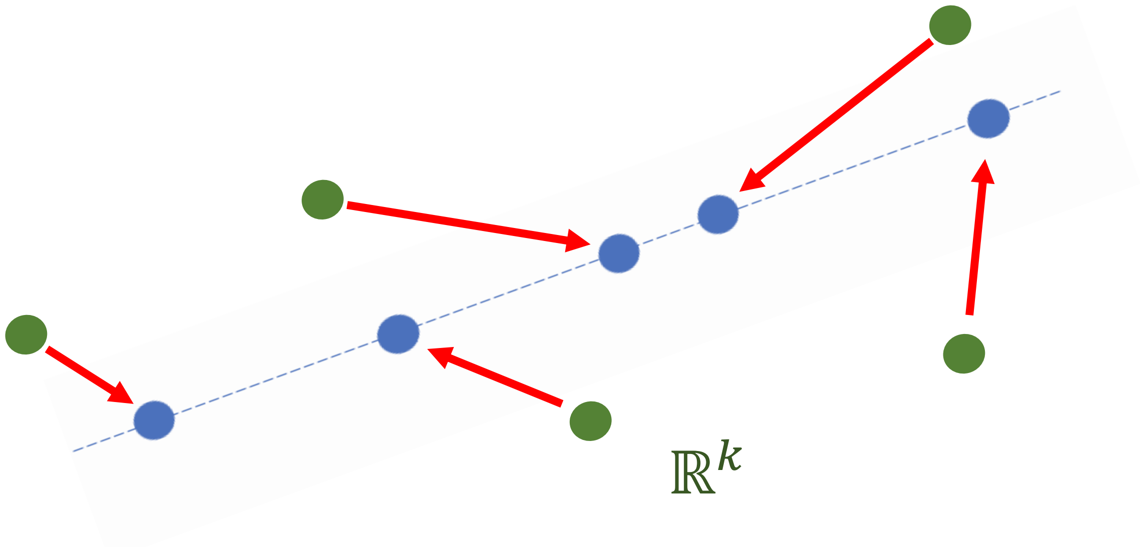
Proof sketch: isotropic

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**

Step 2: Move 1-critical configuration “toward” ground truth.

On average, this *random* perturbation will decrease the cost



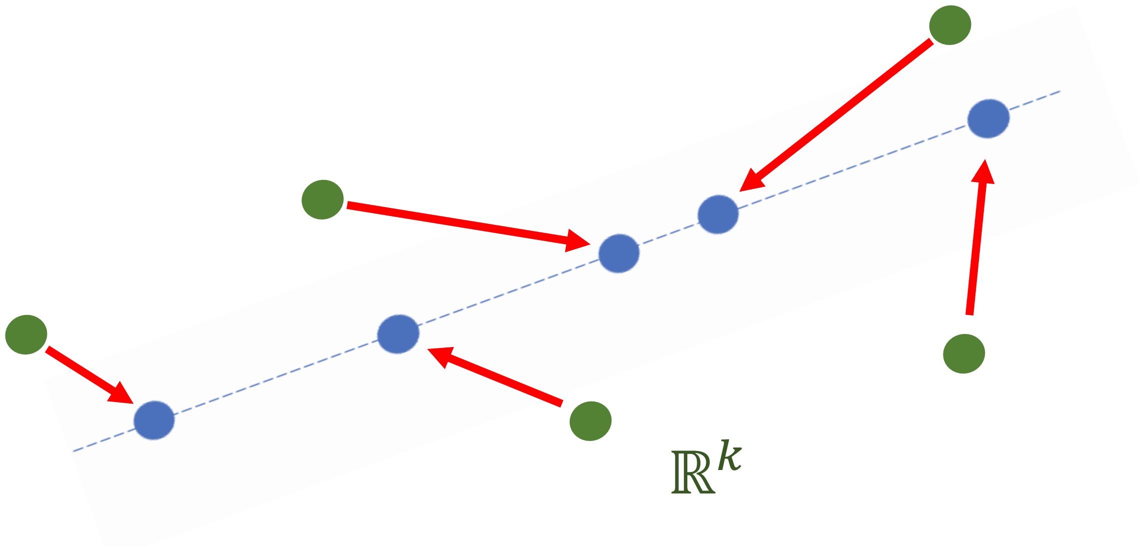
Proof sketch: isotropic

Goal: **perturb** 1-critical configuration to decrease cost

Step 1: Align coordinate systems. How? **Randomly!**

Step 2: Move 1-critical configuration "toward" ground truth.

On average, this *random* perturbation will decrease the cost



Mei, Misiakiewicz, Montanari, Oliveira '17

McRae, Boumal '23

McRae, Abdalla, Bandeira, Boumal '24

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.

Alternative perspective:

Low-Rank Optimization

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

Conjecture [isotropic GT]: Relaxing is not necessary.

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

SNL 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 $Y = ZZ^\top, Y_* = Z_*Z_*^\top$

SNL 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 $= \|z_i - z_j\|^2$

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

Notation and reformulation

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

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

over $z_1, z_2, \dots, z_n \in \mathbb{R}^k$ "s-stress"

Notation and reformulation

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

"s-stress"

$$\Updownarrow$$

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

Notation and reformulation

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

"s-stress"

$$\Updownarrow$$

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

$\Downarrow$ (relax)

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

Notation and reformulation

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

$\Updownarrow$

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

$\Downarrow$ (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)

Notation and reformulation

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

$\Updownarrow$

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

$\Downarrow$ (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 $2 \Rightarrow 1$, i.e., 2-critical points map to 1-critical points

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

Notation and reformulation

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

$$\Updownarrow$$

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

$\Downarrow$ (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?

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

Restricted Isometry Property (RIP):

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

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 SNL map?

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

Special properties of SNL map?

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

Special “perturbation” of the identity

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

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

Special properties of SNL map?

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

Special "perturbation" of the identity

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

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

New "general" theorem: If $\Gamma(Y) = \sum_{i=1}^{N} a_i a_i^\top (a_i^\top Y a_i)$ with $a_i \in \mathbb{R}^n$,

- is contractive (trace and operator norm),
- and satisfies $\langle Y, \Psi(Y)\rangle \leq c\langle Y, \Gamma(Y)\rangle \quad \forall Y$

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

Takeaways for SNL

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 for SNL

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!)
    - Other applications?
- Randomized directions for proving benign landscapes
- Going beyond RIP: structured "perturbations"

Taking a step back

- Most landscape results proved on case-by-case basis

New “general” theorem: If $\Gamma(Y) = \sum_{i=1}^{N} a_i a_i^\top \left(a_i^\top Y a_i\right)$ with $a_i \in \mathbb{R}^n$,
- is contractive (trace and operator norm),
- and satisfies $\langle Y, \Theta(Y)\rangle \leq c\langle Y, \Gamma(Y)\rangle \quad \forall Y$

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

Taking a step back

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - **Reparameterization**

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

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

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

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - **Reparameterization**

$$\min_y f(y) \qquad\qquad \min \|\Delta(Y - Y_*)\|^2 \text{ over } Y \succcurlyeq 0 \text{ with } \mathrm{rank}(Y) \leq \ell$$

$$\downarrow \qquad\qquad \downarrow$$

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

Relationship between landscapes of f and $f \circ \phi$?

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

    Ex: Rayleigh quotient, some low-rank matrix and tensor factorization problems,
    Horn's problem, Paulsen problem, quantum marginals problem, …

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

    Ex: Rayleigh quotient, some low-rank matrix and tensor factorization problems,
    Horn's problem, Paulsen problem, quantum marginals problem, …

$\min z^\top A z$ over z in sphere

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

    Ex: Rayleigh quotient, some low-rank matrix and tensor factorization problems, Horn's problem, Paulsen problem, quantum marginals problem, …

$\min y^\top a$ over y in simplex $\xrightarrow{y = \phi(z) = z \odot z}$ $\min z^\top A z$ over z in sphere

$A = \mathrm{Diag}(a)$

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

    Ex: Rayleigh quotient, some low-rank matrix and tensor factorization problems,
    Horn's problem, Paulsen problem, quantum marginals problem, …

    Kostant and Kirwan convexity theorems (tools from symplectic geometry)

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f\big(\phi(z)\big)$

    Relationship between landscapes of f and $f \circ \phi$?

  - Some problems with enough **symmetry** are nice reparameterizations of convex problems!

    Ex: Rayleigh quotient, some low-rank matrix and tensor factorization problems,
    Horn's problem, Paulsen problem, quantum marginals problem, …

    Kostant and Kirwan convexity theorems (tools from symplectic geometry)

    Can we use these tools to understand other optimization problems?

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?
  - Kostant and Kirwan convexity theorems (tools from symplectic geometry)

    Can we use these tools to understand other optimization problems?

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$

    Relationship between landscapes of f and $f \circ \phi$?
  - Kostant and Kirwan convexity theorems (tools from symplectic geometry)

    Can we use these tools to understand other optimization problems?
  - Morse theory:

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$
    Relationship between landscapes of f and $f \circ \phi$?
  - Kostant and Kirwan convexity theorems (tools from symplectic geometry)
    Can we use these tools to understand other optimization problems?
  - Morse theory:
    Fact: If f is a Morse function defined on a compact manifold, and has no critical points of index 1, then the landscape of f is benign!

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$
    Relationship between landscapes of f and $f \circ \phi$?
  - Kostant and Kirwan convexity theorems (tools from symplectic geometry)
    Can we use these tools to understand other optimization problems?
  - Morse theory:
    Fact: If f is a Morse function defined on a compact manifold, and has no critical points of index 1, then the landscape of f is benign!

    Index of point x = number of negative eigenvalues of $\nabla^2 f(x)$

Tools for landscapes?

- Most landscape results proved on case-by-case basis

- **Tools** for benign landscapes? When does low-dimensional relaxation help?
  - Reparameterization $\min_y f(y) \longrightarrow \min_z f(\phi(z))$
    Relationship between landscapes of f and $f \circ \phi$?
  - Kostant and Kirwan convexity theorems (tools from symplectic geometry)
    Can we use these tools to understand other optimization problems?
  - Morse theory:
    Fact: If f is a Morse function defined on a compact manifold, and has no critical points of index 1, then the landscape of f is benign!

    Index of point x = number of negative eigenvalues of $\nabla^2 f(x)$

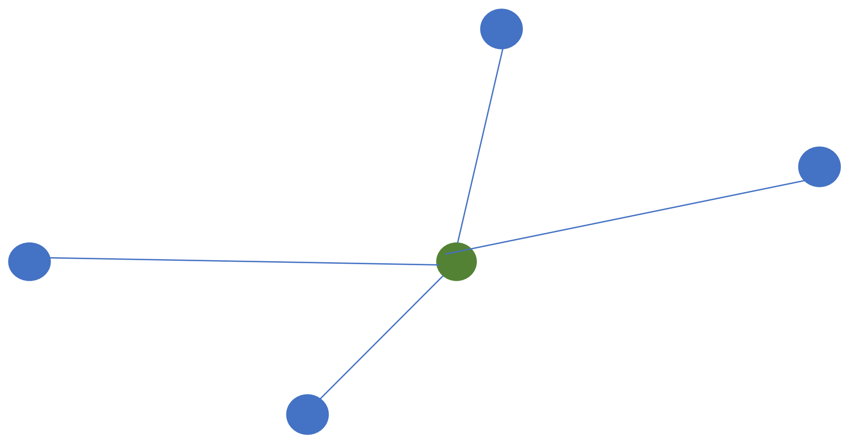
Empirically, relaxing removes critical points of index 1.

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

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

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

Landscape is not benign in general.

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?

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