Benign landscapes for synchronization

Faniriana Tommy RAKOTO ENDOR [1 2]

[1]CNRS CEREMADE
Université Paris-Dauphine PSL

[2]Équipe-Projet MOKAPLAN
Inria Paris

10th June 2026
YRD 2026

1. Orthogonal synchronization
2. Benign landscapes for low-dimensional relaxations
3. A sharp condition-number criterion
4. Applications: $\mathbb{Z}_2$, signed Kuramoto and Kuramoto

What is **single-particle cryo-electron microscopy**?

- We observe noisy 2D projection images of a 3D molecule.

What is **single-particle cryo-electron microscopy**?

- We observe noisy 2D projection images of a 3D molecule.
- We only know relative measurements/angle between projections.

What is **single-particle cryo-electron microscopy**?

- We observe noisy 2D projection images of a 3D molecule.
- We only know relative measurements/angle between projections.
- A central task is to estimate these unknown orientations so that the 2D images can be combined into a 3D reconstruction.

What is **single-particle cryo-electron microscopy**?

- We observe noisy 2D projection images of a 3D molecule.
- We only know relative measurements/angle between projections.
- A central task is to estimate these unknown orientations so that the 2D images can be combined into a 3D reconstruction.

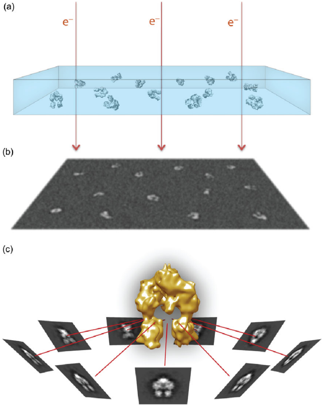


Figure: Single-particle cryo-EM workflow. Adapted from Skiniotis and Southworth [1].

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

- The image I_i is approximately a 2D projection of the same 3D volume, viewed from the direction determined by Z_i.

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

- The image I_i is approximately a 2D projection of the same 3D volume, viewed from the direction determined by Z_i.
- Although the absolute orientations Z_i are unknown, pairs of images can provide noisy information about their relative orientation.

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

- The image I_i is approximately a 2D projection of the same 3D volume, viewed from the direction determined by Z_i.
- Although the absolute orientations Z_i are unknown, pairs of images can provide noisy information about their relative orientation.
- Thus, for some pairs (i, j), we obtain noisy measurements of the form

$$\widehat{R}_{ij} \approx Z_i Z_j^\top,$$

or partial information about $Z_i Z_j^\top$.

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

- The image I_i is approximately a 2D projection of the same 3D volume, viewed from the direction determined by Z_i.
- Although the absolute orientations Z_i are unknown, pairs of images can provide noisy information about their relative orientation.
- Thus, for some pairs (i, j), we obtain noisy measurements of the form

$$\widehat{R}_{ij} \approx Z_i Z_j^\top,$$

or partial information about $Z_i Z_j^\top$.

Synchronization viewpoint

Estimate globally consistent orientations $Z_1, \dots, Z_n \in \mathrm{SO}(3)$ from noisy pairwise relative information $\widehat{R}_{ij}$.

From cryo-EM images to synchronization over $\mathrm{SO}(3)$

Each particle image has an unknown orientation

$$Z_i \in \mathrm{SO}(3).$$

- The image I_i is approximately a 2D projection of the same 3D volume, viewed from the direction determined by Z_i.
- Although the absolute orientations Z_i are unknown, pairs of images can provide noisy information about their relative orientation.
- Thus, for some pairs (i, j), we obtain noisy measurements of the form

$$\widehat{R}_{ij} \approx Z_i Z_j^\top,$$

or partial information about $Z_i Z_j^\top$.

Synchronization viewpoint

Estimate globally consistent orientations $Z_1, \ldots, Z_n \in \mathrm{SO}(3)$ from noisy pairwise relative information $\widehat{R}_{ij}$.

This viewpoint is used in the cryo-EM viewing-direction estimation method of Shkolnisky and Singer [2].

Unknowns

We give ourselves n orthogonal matrices

$$Z_1, \ldots, Z_n \in \mathrm{O}(d), \qquad Z_i Z_i^\top = I_d.$$

And a graph

$$G = (V, E), \qquad V = \{1, \ldots, n\},$$

with adjacency matrix A. Think of Z_i as living on vertex i.

Unknowns

We give ourselves n orthogonal matrices

$$Z_1, \ldots, Z_n \in \mathrm{O}(d), \qquad Z_i Z_i^\top = I_d.$$

And a graph

$$G = (V, E), \qquad V = \{1, \ldots, n\},$$

with adjacency matrix A. Think of Z_i as living on vertex i.

$O(d)$-synchronization

The goal is to recover $Z_1, \ldots, Z_n$ from corrupted pairwise measurements

$$C_{ij} = Z_i Z_j^\top + \Delta_{ij} \in \mathbb{R}^{d \times d}, \ \forall \{i, j\} \in E.$$

Here Δ_{ij} is noise.

A natural estimator is

$$\min_{Y_1,\ldots,Y_n \in \mathbb{R}^{d\times d}} \sum_{(i,j)\in E} \left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 \quad \text{s.t.} \quad Y_i Y_i^\top = I_d.$$

Since Y_i is square, the constraint means $Y_i \in \mathrm{O}(d)$.

A natural estimator is

$$\boxed{\min_{Y_1,\dots,Y_n \in \mathbb{R}^{d\times d}} \sum_{(i,j)\in E} \left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 \quad \text{s.t.} \quad Y_i Y_i^\top = I_d.}$$

Since Y_i is square, the constraint means $Y_i \in \mathrm{O}(d)$.

What the objective measures

For each edge, the fitted relative transformation is

$$Y_i Y_j^\top.$$

The least-squares loss penalizes the mismatch between C_{ij} and $Y_i Y_j^\top$.

A natural estimator is

$$\boxed{\min_{Y_1,\ldots,Y_n \in \mathbb{R}^{d\times d}} \sum_{(i,j)\in E} \left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 \quad \text{s.t.} \quad Y_i Y_i^\top = I_d.}$$

Since Y_i is square, the constraint means $Y_i \in \mathrm{O}(d)$.

What the objective measures

For each edge, the fitted relative transformation is

$$Y_i Y_j^\top.$$

The least-squares loss penalizes the mismatch between C_{ij} and $Y_i Y_j^\top$.

Computational issue

This is a nonconvex optimization problem with many coupled orthogonality constraints. For $d = 1$, it contains binary quadratic optimization problems such as MaxCut-type problems, so it is **NP-hard** in general.

For the square problem $Y_i \in \mathrm{O}(d)$, the least-squares objective is equivalent to a trace maximization. Indeed,

$$\left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 = \|C_{ij}\|_F^2 + d - 2 \left\langle C_{ij}, Y_i Y_j^\top \right\rangle .$$

For the square problem $Y_i \in \mathrm{O}(d)$, the least-squares objective is equivalent to a trace maximization. Indeed,

$$\left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 = \|C_{ij}\|_F^2 + d - 2 \left\langle C_{ij}, Y_i Y_j^\top \right\rangle .$$

The first two terms are constant, hence least squares is equivalent to

$$\boxed{\max_{Y_i \in \mathrm{O}(d)} \sum_{(i,j) \in E} \left\langle C_{ij}, Y_i Y_j^\top \right\rangle = \langle C, YY^\top \rangle.}$$

For the square problem $Y_i \in \mathrm{O}(d)$, the least-squares objective is equivalent to a trace maximization. Indeed,

$$\left\| C_{ij} - Y_i Y_j^\top \right\|_F^2 = \|C_{ij}\|_F^2 + d - 2 \left\langle C_{ij}, Y_i Y_j^\top \right\rangle .$$

The first two terms are constant, hence least squares is equivalent to

$$\boxed{\max_{Y_i \in \mathrm{O}(d)} \sum_{(i,j) \in E} \left\langle C_{ij}, Y_i Y_j^\top \right\rangle = \langle C, YY^\top \rangle.}$$

The original variable is square: $Y_i \in \mathbb{R}^{d \times d}$ with $Y_i Y_i^\top = I_d$.

The original variable is square: $Y_i \in \mathbb{R}^{d \times d}$ with $Y_i Y_i^\top = I_d$. We **relax or overparametrize** it by allowing more columns:

$$Y_i \in \mathbb{R}^{d \times p}, \qquad p \geq d, \qquad Y_i Y_i^\top = I_d.$$

The original variable is square: $Y_i \in \mathbb{R}^{d \times d}$ with $Y_i Y_i^\top = I_d$. We **relax or overparametrize** it by allowing more columns:

$$Y_i \in \mathbb{R}^{d \times p}, \qquad p \geq d, \qquad Y_i Y_i^\top = I_d.$$

Thus

$$Y_i \in \mathrm{St}(d, p) := \{U \in \mathbb{R}^{d \times p} : UU^\top = I_d\}.$$

The original variable is square: $Y_i \in \mathbb{R}^{d \times d}$ with $Y_i Y_i^\top = I_d$. We **relax or overparametrize** it by allowing more columns:

$$Y_i \in \mathbb{R}^{d \times p}, \qquad p \geq d, \qquad Y_i Y_i^\top = I_d.$$

Thus

$$Y_i \in \mathrm{St}(d, p) := \{U \in \mathbb{R}^{d \times p} : UU^\top = I_d\}.$$

$$\boxed{\max_{Y_i \in \mathrm{St}(d,p)} \sum_{(i,j) \in E} \left\langle C_{ij}, Y_i Y_j^\top \right\rangle = F_p(Y).}$$

The original variable is square: $Y_i \in \mathbb{R}^{d\times d}$ with $Y_i Y_i^\top = I_d$. We **relax or overparametrize** it by allowing more columns:

$$Y_i \in \mathbb{R}^{d\times p}, \qquad p \geq d, \qquad Y_i Y_i^\top = I_d.$$

Thus

$$Y_i \in \mathrm{St}(d,p) := \{U \in \mathbb{R}^{d\times p} : UU^\top = I_d\}.$$

$$\boxed{\max_{Y_i \in \mathrm{St}(d,p)} \sum_{(i,j)\in E} \left\langle C_{ij}, Y_i Y_j^\top \right\rangle = F_p(Y).}$$

$\implies$ This becomes an optimization problem on a manifold $\to$ more tractable than an NP hard problem.

The original variable is square: $Y_i \in \mathbb{R}^{d\times d}$ with $Y_i Y_i^\top = I_d$. We **relax or overparametrize** it by allowing more columns:

$$Y_i \in \mathbb{R}^{d\times p}, \qquad p \geq d, \qquad Y_i Y_i^\top = I_d.$$

Thus

$$Y_i \in \mathrm{St}(d,p) := \{U \in \mathbb{R}^{d\times p} : UU^\top = I_d\}.$$

$$\boxed{\max_{Y_i \in \mathrm{St}(d,p)} \sum_{(i,j)\in E} \left\langle C_{ij}, Y_i Y_j^\top \right\rangle = F_p(Y).}$$

$\implies$ This becomes an optimization problem on a manifold $\to$ more tractable than an NP hard problem.

The question of this talk

Does increasing p improve the nonconvex landscape?

More precisely: can a small amount of overparameterization remove spurious stable critical points?

We have introduced a family of nonconvex problems indexed by the relaxation dimension

$$\max_{Y \in \mathrm{St}(d,p)^n} F_p(Y), \qquad p \geq d.$$

We have introduced a family of nonconvex problems indexed by the relaxation dimension

$$\max_{Y \in \mathrm{St}(d,p)^n} F_p(Y), \qquad p \geq d.$$

Next question

The original problem corresponds to $p = d$. The relaxation keeps the same objective, but enlarges the search space.

Does increasing p remove bad local optima?

We have introduced a family of nonconvex problems indexed by the relaxation dimension

$$\max_{Y \in \mathrm{St}(d,p)^n} F_p(Y), \qquad p \geq d.$$

Next question

The original problem corresponds to $p = d$. The relaxation keeps the same objective, but enlarges the search space.

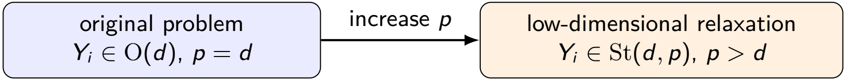
Does increasing p remove bad local optima?

Why second-order points matter

Gradient-based methods naturally settle at stable critical points. A stable local maximum must satisfy the second-order optimality conditions, so we study SOCPs.

The relaxed problem is still nonconvex, but increasing p gives each block Y_i **more tangent directions**.

The relaxed problem is still nonconvex, but increasing p gives each block Y_i **more tangent directions**.



The relaxed problem is still nonconvex, but increasing p gives each block Y_i **more tangent directions**.

original problem	increase p →	low-dimensional relaxation
$Y_i \in \mathrm{O}(d),\ p = d$		$Y_i \in \mathrm{St}(d,p),\ p > d$

Landscape question

For which p is the implication

$$\mathrm{SOCP} \implies \text{global optimum}$$

true, i.e. no spurious local maxima?

The relaxed problem is still nonconvex, but increasing p gives each block Y_i **more tangent directions**.

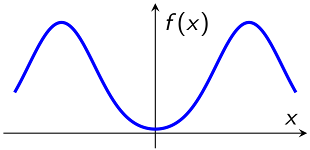
original problem	increase p →	low-dimensional relaxation
$Y_i \in \mathrm{O}(d),\ p = d$		$Y_i \in \mathrm{St}(d,p),\ p > d$

Landscape question

For which p is the implication

$$\mathrm{SOCP} \implies \text{global optimum}$$

true, i.e. no spurious local maxima?



Example: neither convex nor concave on $\mathbb{R}$, but it has no spurious local maximum. This is benign but non-concave.

Let

$$f(Y) = \left\langle C, YY^\top \right\rangle, \qquad Y \in \mathrm{St}(d, p)^n.$$

Let

$$f(Y) = \left\langle C, YY^\top \right\rangle, \qquad Y \in \mathrm{St}(d,p)^n.$$

Second-order critical point

A feasible point Y is a SOCP if

$$\operatorname{grad} f(Y) = 0, \qquad \operatorname{Hess} f(Y) \preceq 0$$

for the maximization problem on the product Stiefel manifold.

Let

$$f(Y) = \left\langle C, YY^\top \right\rangle, \qquad Y \in \mathrm{St}(d,p)^n.$$

Second-order critical point

A feasible point Y is a SOCP if

$$\text{grad} f(Y) = 0, \qquad \text{Hess} f(Y) \preceq 0$$

for the maximization problem on the product Stiefel manifold.

Benign landscape

The landscape is benign if

$$\mathrm{SOCP} \quad \Longrightarrow \quad Y \text{ is globally optimal.}$$

Equivalently, there are no spurious local maxima.

Theorem 1 (McRae–Boumal, 2023)

Assume:

- $G = (V, E)$ is connected;

Theorem 1 (McRae–Boumal, 2023)

Assume:

- $G = (V, E)$ is connected;
- measurements are exact **no noise** : $C_{ij} = Z_i Z_j^\top$ for all $(i, j) \in E$;

Theorem 1 (McRae–Boumal, 2023)

Assume:

- $G = (V, E)$ is connected;
- measurements are exact **no noise** : $C_{ij} = Z_i Z_j^\top$ for all $(i, j) \in E$;
- the relaxation dimension satisfies $p \geq d + 2$.

Theorem 1 (McRae–Boumal, 2023)

Assume:

- $G = (V, E)$ is connected;
- measurements are exact **no noise** : $C_{ij} = Z_i Z_j^\top$ for all $(i, j) \in E$;
- the relaxation dimension satisfies $p \geq d + 2$.

Then every second-order critical point Y of the low-dimensional relaxation satisfies

$$YY^\top = ZZ^\top.$$

Equivalently,

$$Y = ZU, \qquad U \in \mathbb{R}^{d \times p}, \quad UU^\top = I_d.$$

Theorem 1 (McRae–Boumal, 2023)

Assume:

- $G = (V, E)$ is connected;
- measurements are exact **no noise** : $C_{ij} = Z_i Z_j^\top$ for all $(i, j) \in E$;
- the relaxation dimension satisfies $p \geq d + 2$.

Then every second-order critical point Y of the low-dimensional relaxation satisfies

$$YY^\top = ZZ^\top.$$

Equivalently,

$$Y = ZU, \qquad U \in \mathbb{R}^{d\times p}, \quad UU^\top = I_d.$$

Sharpness

For general connected graphs, $p = d$ and $p = d + 1$ can have spurious local optima because $\mathrm{St}(d, p)$ is connected but **not simply connected**.

Let Δ be the block noise matrix and L_G the graph Laplacian of G :

$$L_G = \mathrm{diag}(A\mathbf{1}) - A, \qquad 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n.$$

For $p > d + 2$, set $C_p = \dfrac{2(p+d-2)}{p-d-2}$.

Let Δ be the block noise matrix and L_G the graph Laplacian of G :

$$L_G = \mathrm{diag}(A\mathbf{1}) - A, \qquad 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n.$$

For $p > d + 2$, set $C_p = \dfrac{2(p+d-2)}{p-d-2}$.

Theorem 3 (McRae–Boumal, 2023)

If

$$\frac{\|\Delta\|_2}{\lambda_2(L_G)} < \frac{1}{\sqrt{5}\, C_p \sqrt{dn}},$$

then any second-order critical point Y of the relaxation satisfies:

Let Δ be the block noise matrix and L_G the graph Laplacian of G :

$$L_G = \operatorname{diag}(A\mathbf{1}) - A, \qquad 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n.$$

For $p > d+2$, set $C_p = \dfrac{2(p+d-2)}{p-d-2}$.

Theorem 3 (McRae–Boumal, 2023)

If

$$\frac{\|\Delta\|_2}{\lambda_2(L_G)} < \frac{1}{\sqrt{5}\, C_p \sqrt{dn}},$$

then any second-order critical point Y of the relaxation satisfies:

- Y is the unique solution of the relaxation, up to a global orthogonal transformation in $\mathrm{O}(p)$;

Let Δ be the block noise matrix and L_G the graph Laplacian of G :

$$L_G = \mathrm{diag}(A\mathbf{1}) - A, \qquad 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n.$$

For $p > d + 2$, set $C_p = \dfrac{2(p+d-2)}{p-d-2}$.

Theorem 3 (McRae–Boumal, 2023)

If

$$\frac{\|\Delta\|_2}{\lambda_2(L_G)} < \frac{1}{\sqrt{5}\, C_p \sqrt{dn}},$$

then any second-order critical point Y of the relaxation satisfies:

- Y is the unique solution of the relaxation, up to a global orthogonal transformation in $\mathrm{O}(p)$;
- Y has rank d, so $Y = \widehat{Z}U$ with $\widehat{Z} \in \mathbb{R}^{dn \times d}$, $U \in \mathbb{R}^{d \times p}$, $UU^\top = I_d$;

Let Δ be the block noise matrix and L_G the graph Laplacian of G :

$$L_G = \mathrm{diag}(A\mathbf{1}) - A, \qquad 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n.$$

For $p > d + 2$, set $C_p = \dfrac{2(p+d-2)}{p-d-2}$.

Theorem 3 (McRae–Boumal, 2023)

If

$$\frac{\|\Delta\|_2}{\lambda_2(L_G)} < \frac{1}{\sqrt{5}\, C_p \sqrt{dn}},$$

then any second-order critical point Y of the relaxation satisfies:

- Y is the unique solution of the relaxation, up to a global orthogonal transformation in $\mathrm{O}(p)$;
- Y has rank d, so $Y = \widehat{Z}U$ with $\widehat{Z} \in \mathbb{R}^{dn\times d}$, $U \in \mathbb{R}^{d\times p}$, $UU^\top = I_d$;
- $\widehat{Z}$ is the unique solution of the original $\mathrm{O}(d)$ problem, up to a global $\mathrm{O}(d)$ transformation.

$\Longrightarrow$ If the SNR is small, the landscape is benign.

In the noiseless gauge $Z_i = I_d$, the synchronization energy can be written as the quantity we want to **minimize**

$$E(Y) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left\| Y_i - Y_j \right\|_F^2 , \qquad Y_i \in \mathrm{St}(d, p).$$

In the noiseless gauge $Z_i = I_d$, the synchronization energy can be written as the quantity we want to **minimize**

$$E(Y) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left\| Y_i - Y_j \right\|_F^2 , \qquad Y_i \in \mathrm{St}(d, p).$$

Since $Y_i Y_i^\top = I_d$, minimizing E is equivalent to maximizing

$$F(Y) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \langle Y_i, Y_j \rangle.$$

In the noiseless gauge $Z_i = I_d$, the synchronization energy can be written as the quantity we want to **minimize**

$$E(Y) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left\| Y_i - Y_j \right\|_F^2, \qquad Y_i \in \mathrm{St}(d, p).$$

Since $Y_i Y_i^\top = I_d$, minimizing E is equivalent to maximizing

$$F(Y) = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \langle Y_i, Y_j \rangle.$$

Synchronized configurations

$$Y_1 = Y_2 = \cdots = Y_n$$

are exactly the zero-energy configurations when G is connected.

The Euclidean gradient of the energy with respect to Y_i is

$$\nabla_{Y_i} E(Y) = \sum_{j=1}^{n} A_{ij}(Y_i - Y_j).$$

The Euclidean gradient of the energy with respect to Y_i is

$$\nabla_{Y_i} E(Y) = \sum_{j=1}^{n} A_{ij}(Y_i - Y_j).$$

Projecting onto the tangent space of $\mathrm{St}(d, p)$ gives the Riemannian gradient flow

$$\boxed{\dot{Y}_i = -P_{T_{Y_i}\mathrm{St}(d,p)}\Big(\sum_j A_{ij}(Y_i - Y_j)\Big) = P_{T_{Y_i}\mathrm{St}(d,p)}\Big(\sum_j A_{ij} Y_j\Big).}$$

This is the $\mathrm{St}(d, p)$-valued **high-dimensional Kuramoto model**.

The Euclidean gradient of the energy with respect to Y_i is

$$\nabla_{Y_i} E(Y) = \sum_{j=1}^{n} A_{ij}(Y_i - Y_j).$$

Projecting onto the tangent space of $\mathrm{St}(d, p)$ gives the Riemannian gradient flow

$$\boxed{\dot{Y}_i = -P_{T_{Y_i}\mathrm{St}(d,p)}\Big(\sum_j A_{ij}(Y_i - Y_j)\Big) = P_{T_{Y_i}\mathrm{St}(d,p)}\Big(\sum_j A_{ij} Y_j\Big).}$$

This is the $\mathrm{St}(d, p)$-valued **high-dimensional Kuramoto model**.

Gradient flow selects stable critical points

Along the flow,

$$\frac{d}{dt} E(Y(t)) = -\left\|\mathrm{grad}\, E(Y(t))\right\|^2 \leq 0.$$

Since $\mathrm{St}(d, p)^n$ is compact and E is analytic, trajectories converge to critical points; stable equilibria satisfy the second-order necessary conditions, hence are SOCPs for $\min E$, equivalently for $\max F$.

Corollary 1 (McRae–Boumal, 2023)

For any connected graph G, the $\mathrm{St}(d, p)$-valued Kuramoto network synchronizes when

$$p \geq d + 2.$$

Equivalently, every stable equilibrium is synchronized.

Corollary 1 (McRae–Boumal, 2023)

For any connected graph G, the $\mathrm{St}(d,p)$-valued Kuramoto network synchronizes when

$$p \geq d + 2.$$

Equivalently, every stable equilibrium is synchronized.

Optimality

For connected graphs in general, this threshold is optimal: it is equivalent to simple connectedness of $\mathrm{St}(d,p)$.

McRae–Boumal (2023) gives a clean answer in the noiseless case:

$$p \geq d + 2 \quad \Longrightarrow \quad \text{no spurious stable points.}$$

McRae–Boumal (2023) gives a clean answer in the noiseless case:

$$p \geq d + 2 \quad \Longrightarrow \quad \text{no spurious stable points.}$$

Next specialization

We now focus on $d = 1$, where

$$\mathrm{O}(1) = \{\pm 1\}.$$

The geometry becomes simpler, and the landscape can be controlled by a deterministic matrix attached to the ground truth: a certificate Laplacian.

When $d = 1$,

$$\mathrm{O}(1) = \{\pm 1\}.$$

When $d = 1$,

$$\mathrm{O}(1) = \{\pm 1\}.$$

The original problem is a binary quadratic (**and NP hard !**) problem:

$$\max_{x \in \{\pm 1\}^n} x^\top C x = \sum_{i,j=1}^{n} C_{ij} x_i x_j.$$

When $d = 1$,

$$\mathrm{O}(1) = \{\pm 1\}.$$

The original problem is a binary quadratic (**and NP hard !**) problem:

$$\max_{x \in \{\pm 1\}^n} x^\top C x = \sum_{i,j=1}^{n} C_{ij} x_i x_j.$$

The low-dimensional sphere relaxation $\{\pm 1\} \to S^{p-1} = \mathrm{St}(1, p)$ is

$$\boxed{\max_{y_i \in S^{p-1}} \sum_{i,j} C_{ij} \langle y_i, y_j \rangle .}$$

When $d = 1$,

$$\mathrm{O}(1) = \{\pm 1\}.$$

The original problem is a binary quadratic (**and NP hard !**) problem:

$$\max_{x \in \{\pm 1\}^n} x^\top C x = \sum_{i,j=1}^{n} C_{ij} x_i x_j.$$

The low-dimensional sphere relaxation $\{\pm 1\} \to S^{p-1} = \mathrm{St}(1, p)$ is

$$\boxed{\max_{y_i \in S^{p-1}} \sum_{i,j} C_{ij} \langle y_i, y_j \rangle .}$$

Same question, simpler geometry

Increasing p replaces signs by vectors on the sphere. We ask whether this extra dimension removes non-global stable critical points.

Fix a candidate ground truth $x \in \{\pm 1\}^n$. Define

$$\boxed{L = L(x) := \mathrm{ddiag}(Cxx^\top) - C.}$$

Then

$$Lx = 0.$$

Fix a candidate ground truth $x \in \{\pm 1\}^n$. Define

$$\boxed{L = L(x) := \operatorname{ddiag}(Cxx^\top) - C.}$$

Then

$$Lx = 0.$$

Ground-truth certificate

The conditions

$$L \succeq 0, \qquad \lambda_2(L) > 0,$$

certify that $xx^\top$ is the unique recoverable rank-one solution.

Fix a candidate ground truth $x \in \{\pm 1\}^n$. Define

$$\boxed{L = L(x) := \mathrm{ddiag}(Cxx^\top) - C.}$$

Then

$$Lx = 0.$$

Ground-truth certificate

The conditions

$$L \succeq 0, \qquad \lambda_2(L) > 0,$$

certify that $xx^\top$ is the unique recoverable rank-one solution.

Nonconvex question

Assuming this certificate holds, how large must p be to ensure

$$Y \text{ SOCP} \quad \Longrightarrow \quad YY^\top = xx^\top?$$

Theorem 2.2 (Rakoto Endor–Waldspurger, 2024)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

Theorem 2.2 (Rakoto Endor–Waldspurger, 2024)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

If

$$\boxed{p > \frac{\lambda_n(L)}{\lambda_2(L)},}$$

then every second-order critical point is globally optimal:

$$YY^\top = xx^\top.$$

Theorem 2.2 (Rakoto Endor–Waldspurger, 2024)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

If

$$\boxed{p > \frac{\lambda_n(L)}{\lambda_2(L)},}$$

then every second-order critical point is globally optimal:

$$YY^\top = xx^\top.$$

Sharpness

The threshold cannot be improved in general: there are instances with

$$\frac{\lambda_n(L)}{\lambda_2(L)} = p$$

and a non-global second-order critical point.

Let $D \succ 0$ be any positive diagonal matrix and define

$$L_D = D^{-1/2} L D^{-1/2}.$$

Let $D \succ 0$ be any positive diagonal matrix and define

$$L_D = D^{-1/2} L D^{-1/2}.$$

Theorem 2.1 (McRae, 2025)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

Let $D \succ 0$ be any positive diagonal matrix and define

$$L_D = D^{-1/2} L D^{-1/2}.$$

Theorem 2.1 (McRae, 2025)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

If

$$\boxed{p > \frac{\lambda_n(L_D)}{\lambda_2(L_D)},}$$

then every SOCP of the sphere relaxation is globally optimal:

$$YY^\top = xx^\top.$$

Let $D \succ 0$ be any positive diagonal matrix and define

$$L_D = D^{-1/2} L D^{-1/2}.$$

Theorem 2.1 (McRae, 2025)

Assume

$$L \succeq 0, \qquad \lambda_2(L) > 0.$$

If

$$\boxed{p > \frac{\lambda_n(L_D)}{\lambda_2(L_D)},}$$

then every SOCP of the sphere relaxation is globally optimal:

$$YY^\top = xx^\top.$$

Why diagonal scaling helps

A good choice of D can reduce the effective condition number. In graph applications, this often corresponds to using a normalized Laplacian.

Setting	Criterion	Object controlling the landscape
McRae–Boumal	$p \geq d + 2$	connected noiseless graph
Rakoto Endor–Waldspurger	$p > \kappa(L)$	sharp certificate condition
McRae	$p > \kappa(L_D)$	normalized certificate condition

Setting	Criterion	Object controlling the landscape
McRae–Boumal	$p \geq d + 2$	connected noiseless graph
Rakoto Endor–Waldspurger	$p > \kappa(L)$	sharp certificate condition
McRae	$p > \kappa(L_D)$	normalized certificate condition

$$\kappa(L) = \frac{\lambda_n(L)}{\lambda_2(L)}, \qquad \kappa(L_D) = \frac{\lambda_n(D^{-1/2} L D^{-1/2})}{\lambda_2(D^{-1/2} L D^{-1/2})}.$$

The condition-number theorems reduce the landscape question to estimating

$$\kappa(L) \quad \text{or} \quad \kappa(D^{-1/2} L D^{-1/2}).$$

The condition-number theorems reduce the landscape question to estimating

$$\kappa(L) \quad \text{or} \quad \kappa(D^{-1/2} L D^{-1/2}).$$

Applications

In probabilistic and graph models, concentration estimates show that these condition numbers are small in the regimes of interest. This gives benign landscapes for:

$\mathbb{Z}_2$-synchronization, signed Kuramoto networks.

Consequence: $\mathbb{Z}_2$-synchronization

We want to retrieve n numbers $x_1, \ldots, x_n \in O(1) = \{\pm 1\}^n$ from corrupted measurments

$$x_i x_j + \sigma W_{ij}, \; W_{ij} \sim N(0,1).$$

We want to retrieve n numbers $x_1, \ldots, x_n \in O(1) = \{\pm 1\}^n$ from corrupted measurments

$$x_i x_j + \sigma W_{ij}, \;\; W_{ij} \sim N(0,1).$$

Gaussian observation model

$$C = zz^\top + \sigma W, \qquad z \in \{\pm 1\}^n.$$

We want to retrieve n numbers $x_1, \ldots, x_n \in O(1) = \{\pm 1\}^n$ from corrupted measurments

$$x_i x_j + \sigma W_{ij}, \;\; W_{ij} \sim N(0,1).$$

Gaussian observation model

$$C = zz^\top + \sigma W, \qquad z \in \{\pm 1\}^n.$$

Sphere relaxation

$$\max_{y_i \in S^{p-1}} \sum_{i,j} C_{ij} \langle y_i, y_j \rangle .$$

Consequence: $\mathbb{Z}_2$-synchronization

We want to retrieve n numbers $x_1, \ldots, x_n \in O(1) = \{\pm 1\}^n$ from corrupted measurments

$$x_i x_j + \sigma W_{ij}, \;\; W_{ij} \sim N(0,1).$$

Gaussian observation model

$$C = zz^\top + \sigma W, \qquad z \in \{\pm 1\}^n.$$

Sphere relaxation

$$\max_{y_i \in S^{p-1}} \sum_{i,j} C_{ij} \langle y_i, y_j \rangle .$$

Consequence (McRae, 2025)

If for some $\varepsilon > 0$,

$$\sigma \leq \sqrt{\frac{n}{(2+\varepsilon)\log n}},$$

then, with probability tending to 1 as $n \to \infty$, for every fixed $p \geq 2$, every SOCP satisfies

$$YY^\top = zz^\top.$$

We want to retrieve n numbers $x_1, \ldots, x_n \in O(1) = \{\pm 1\}^n$ from corrupted measurments

$$x_i x_j + \sigma W_{ij}, \;\; W_{ij} \sim N(0,1).$$

Gaussian observation model

$$C = zz^\top + \sigma W, \qquad z \in \{\pm 1\}^n.$$

Sphere relaxation

$$\max_{y_i \in S^{p-1}} \sum_{i,j} C_{ij} \langle y_i, y_j \rangle .$$

Consequence (McRae, 2025)

If for some $\varepsilon > 0$,

$$\sigma \leq \sqrt{\frac{n}{(2+\varepsilon)\log n}},$$

then, with probability tending to 1 as $n \to \infty$, for every fixed $p \geq 2$, every SOCP satisfies

$$YY^\top = zz^\top.$$

Sharpness of the noise level

The threshold $\sigma \asymp \sqrt{n/(2\log n)}$ is sharp for exact recovery in this model, as stated in Bandeira (2018).

Take the complete measurement graph (that is, the edge-observation probability is fixed equal to one), and let $\delta \in [0, 1]$ be the signal strength. For each pair $i < j$, set

$$C_{ij} = \begin{cases} z_i z_j, & \text{with probability } \frac{1+\delta}{2}, \\ -z_i z_j, & \text{with probability } \frac{1-\delta}{2}. \end{cases}$$

Take the complete measurement graph (that is, the edge-observation probability is fixed equal to one), and let $\delta \in [0, 1]$ be the signal strength. For each pair $i < j$, set

$$C_{ij} = \begin{cases} z_i z_j, & \text{with probability } \frac{1+\delta}{2}, \\ -z_i z_j, & \text{with probability } \frac{1-\delta}{2}. \end{cases}$$

In the gauge $z = \mathbf{1}$, taking $A = C$ gives a complete signed Kuramoto network:

$$\dot{\theta}_i = K \sum_j A_{ij} \sin(\theta_j - \theta_i),$$

with positive attractive edges and negative repulsive edges.

Theorem 3.2 (McRae, 2025), with complete measurements

Suppose that, for some $\varepsilon > 0$,

$$\frac{n}{\log n}\left(1 - \sqrt{1 - (1 - \varepsilon)\delta^2}\right) \geq 1.$$

Theorem 3.2 (McRae, 2025), with complete measurements

Suppose that, for some $\varepsilon > 0$,

$$\frac{n}{\log n}\left(1 - \sqrt{1 - (1 - \varepsilon)\delta^2}\right) \geq 1.$$

Then, with probability tending to 1 as $n \to \infty$:

- for every $p \geq 2$, every SOCP of the sphere relaxation satisfies

$$YY^\top = zz^\top;$$

Theorem 3.2 (McRae, 2025), with complete measurements

Suppose that, for some $\varepsilon > 0$,

$$\frac{n}{\log n}\left(1 - \sqrt{1 - (1-\varepsilon)\delta^2}\right) \geq 1.$$

Then, with probability tending to 1 as $n \to \infty$:

- for every $p \geq 2$, every SOCP of the sphere relaxation satisfies

$$YY^\top = zz^\top;$$

- equivalently, for $p = 2$, the complete random signed Kuramoto oscillator network globally synchronizes.

Theorem 3.2 (McRae, 2025), with complete measurements

Suppose that, for some $\varepsilon > 0$,

$$\frac{n}{\log n}\left(1 - \sqrt{1 - (1-\varepsilon)\delta^2}\right) \geq 1.$$

Then, with probability tending to 1 as $n \to \infty$:

- for every $p \geq 2$, every SOCP of the sphere relaxation satisfies

$$YY^\top = zz^\top;$$

- equivalently, for $p = 2$, the complete random signed Kuramoto oscillator network globally synchronizes.

Simpler sufficient condition

For small δ, it is enough to have

$$\delta \gtrsim \sqrt{\frac{2\log n}{n}}.$$

Thus the complete signed network tolerates a vanishing average sign bias of order $\sqrt{\log n / n}$.

Thank you for your attention !

References

G. Skiniotis and D. R. Southworth.
"Single-particle cryo-electron microscopy of macromolecular complexes."
Microscopy, 65(1):9–22, 2016. DOI: 10.1093/jmicro/dfv366.

Y. Shkolnisky and A. Singer.
"Viewing direction estimation in cryo-EM using synchronization."
SIAM J. Imaging Sci., 5(3):1088–1110, 2012. DOI: 10.1137/120863642.

A. D. McRae and N. Boumal.
"Benign landscapes of low-dimensional relaxations for orthogonal synchronization on general graphs."
SIAM J. Optim., 34(2):1427–1454, 2024. DOI: 10.1137/23M1584642.

F. Rakoto Endor and I. Waldspurger.
"Benign landscape for Burer–Monteiro factorizations of MaxCut-type semidefinite programs."
arXiv:2411.03103, 2024.

A. D. McRae.
"Benign landscapes for synchronization on spheres via normalized Laplacian matrices."
arXiv:2503.18801, 2025.

A. S. Bandeira.
"Random Laplacian matrices and convex relaxations."
Found. Comput. Math., 18(2):345–379, 2018. DOI: 10.1007/s10208-016-9341-9.