# Planted Random Number Partitioning Problem

Eren C. Kızıldağ (UIUC)

**Algorithmic Learning Theory, 2026**

# Number Partitioning Problem (NPP)

Given $X \in \mathbb{R}^n$, find a balanced partition: $\min_{\sigma \in \{-1,1\}^n} |\langle \sigma, X \rangle|$.

- One of **Karp**'s **21 NP-complete** problems.
  Approximation is as hard as **Shortest Vector**
  [Karp 72, Garey-Johnson 79, Hoberg-Ramadas-Rothvoss-Yang 17]
- Instance of **discrepancy minimization**
  [Spencer 85, Chazelle 01, Bansal 10, Lovett-Meka 15]
- Exhibits **phase transitions. Locally REM-like** [Mertens 98, Borgs-Chayes-Pittel 01, Bauke-Mertens 04, Borgs-Chayes-Mertens-Pittel 09]
- Linked to design of **randomized controlled trials**
  [Krieger-Azriel-Kapelner 19, Harshaw-Sävje-Spielman-Zhang 19]

# Planted Models

- Recover signal hidden in noise:
  stochastic block model, planted clique, sparse PCA
  [Jerrum 92, Alon-Krivelevich-Sudakov 98, Berthet-Rigollet 13, Meka-Potechin-Wigderson 15, Abbe 17]

- Instrumental in studying **random CSPs**
  [Achlioptas & Coja-Oghlan 08, Montanari-Restrepo-Tetali 11, Feldman-Perkins-Vempala 15, Perkins-Xu 21]

A planted variant of the **NPP**?

# Average-Case NPP

## Statistical-Computational Gap

For $X \sim \mathcal{N}(0, I_n)$, the optimal value is $\Theta(\sqrt{n}2^{-n})$, whereas best alg reaches $n^{-\Theta(\log n)}$ [Karmarkar-Karp-Lueker-Odylzko 86, Yakir 96]

## Algorithmic Hardness Results

- Unconditional lower bounds via **multi OGP** [Gamarnik-**K.** 23, Mallarapu-Sellke 25]
- Conditional lower bound, under **worst-case hardness** of **SVP** [Vafa-Vaikuntanathan 25]

## Multi Overlap Gap Property

- Links geometry to algorithmic barriers in **random optimization**
- **Planted:** No multi OGP, rule out only MCMC/simple local search.

# Planted NPP

**Model**

Fix unknown $\boldsymbol{\sigma}^* \in \{-1, 1\}^n$. Sample $X \sim \mathcal{N}(0, I_n)$ conditional on $|\langle \boldsymbol{\sigma}^*, X \rangle| \leq 3^{-n}$

Plant $\boldsymbol{\sigma}^*$. Statistically distinguishable from random model.

**Hamiltonian** $H(\boldsymbol{\sigma}) := |\langle \boldsymbol{\sigma}, X \rangle|$.

**Questions**

Optimizing Hamiltonian, recovery (of $\boldsymbol{\sigma}^*$), hypothesis testing

**This Talk: Optimization**

Landscape results, intricate geometry, algorithmic lower bounds

# Global Optima (other than $\pm\boldsymbol{\sigma}^*$)

All guarantees under planted measure, $\mathbb{P}_{\mathrm{pl}}[\mathcal{E}] := \mathbb{P}\left[\mathcal{E} \,\middle|\, H(\boldsymbol{\sigma}^*) \le 3^{-n}\right]$.

**Theorem**

*It holds that* $\min_{\boldsymbol{\sigma} \neq \pm\boldsymbol{\sigma}^*} H(\boldsymbol{\sigma}) = \widetilde{\Theta}(2^{-n})$ *w.h.p.*

Planting doesn't induce $\boldsymbol{\sigma}$ with $H(\boldsymbol{\sigma}) \ll 2^{-n}$.

**Comparison with Planted Clique**

Planting a $k$-clique with $k \gg \log_2 n$ to $G(n, \frac{1}{2})$ automatically induces $\ell$-cliques for all $\log_2 n \ll \ell \le k$.

# Hamiltonian at Fixed Distance from $\pm\boldsymbol{\sigma}^*$

**Entropy-Energy Trade-off:**

Smallest $H(\boldsymbol{\sigma})$ achieved furthest from $\pm\boldsymbol{\sigma}^*$.

Can we probe the landscape in finer detail?

**Theorem**

*Denote by $h_b(\rho)$ the binary entropy and let:*

$$\zeta(\rho) := \min_{\boldsymbol{\sigma}: d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}^*) = \rho n} H(\boldsymbol{\sigma}).$$

*Then, for any $\rho \in (0,1)$, $\zeta(\rho) = \widetilde{\Theta}(2^{-n h_b(\rho)})$ w.h.p.*

Note that $d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}^*) + d_H(\boldsymbol{\sigma}, -\boldsymbol{\sigma}^*) = n$ and thus $\zeta(\rho) = \zeta(1-\rho)$.

# Planted Partition and Energy Barriers

**Theorem**

*For any* $E = \Omega(\log n)$*, there exists* $d = \omega(1)$ *such that w.h.p.*

$$\min_{\boldsymbol{\sigma}: 1 \le d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}^*) \le d} H(\boldsymbol{\sigma}) > 2^{-E}.$$

$\boldsymbol{\sigma}^*$ and $\boldsymbol{\sigma}$ with $H(\boldsymbol{\sigma}) \sim 2^{-\epsilon n}$ are separated by steep energy barriers (order $2^{-\Omega(\log n)}$).

$\boldsymbol{\sigma}^*$ is isolated. Recovery is likely computationally hard.

[Mézard-Mora-Zecchina 05, Achlioptas & Ricci-Tersenghi 06, Achlioptas & Coja-Oghlan 08]

# Multi Overlap Gap Property

Intricate geometry of tuples of near optima.
Implies hardness against broad classes of algorithms.

[Gamarnik-Sudan 14, 17, Gamarnik-Jagannath-Wein 20, Gamarnik-**K.** 21, Bresler-Huang 21, Huang-Sellke 22, 23, Gamarnik-**K.**-Perkins-Xu 22,23, Gamarnik-**K.**-Warnke 25]

**Symmetric *m*-OGP**

The model exhibits symmetric *m*-OGP at level $2^{-E}$ if $\exists m \in \mathbb{N}$ and $\beta \in (0,1)$, such that there is no $m$-tuple $\boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_m \in \{-1,1\}^n$ satisfying $\max_i H(\boldsymbol{\sigma}_i) \le 2^{-E}$ and $\frac{1}{n}\langle \boldsymbol{\sigma}_i, \boldsymbol{\sigma}_j \rangle \sim \beta$.

No nearly equidistant *m*-tuples of solutions

# Multi OGP in Planted NPP

**Theorem**

*For any $\epsilon$, planted NPP exhibits m-OGP at level $2^{-n\epsilon}$ with suitable $m = \Theta_\epsilon(1/\epsilon)$.*

**First** *m*-OGP result for a planted model.

**Future Work**

Is $2^{-\Theta(n)}$ sharp? (Unplanted) NPP exhibits *m*-OGP at $2^{-\omega(\sqrt{n \log n})}$.

# Algorithmic Lower Bound

**Algorithmic Framework**

Randomized algorithm $\mathcal{A} : \mathbb{R}^n \times \Omega \to \{-1, 1\}^n$, satisfying:

- **Stability:** If $\|X - Y\|$ is small, then $d_H(\mathcal{A}(X), \mathcal{A}(Y))$ is also small
- **Anti-Concentration:** $\mathbb{P}_{\mathrm{pl}}[\mathcal{A}(X, \omega) \in \{\pm \boldsymbol{\sigma}^*\}] = O(1)$

There are $\omega(1)$ partitions at value $2^{-E}$. So, 'reasonable' $\mathcal{A}$ satisfies anti-concentration.

**Theorem**

*Stable algorithms fail to output $\boldsymbol{\sigma} \in \{-1, 1\}^n$ beyond $H(\boldsymbol{\sigma}) = 2^{-o(n)}$.*

Thank you!