The Planted Number Partitioning Problem

Eren C. Kızıldağ KIZILDAG@ILLINOIS.EDU
Department of Statistics, University of Illinois Urbana-Champaign [*]

Editors: Matus Telgarsky and Jonathan Ullman

Abstract

Given a list of numbers $X \sim \mathcal{N}(0, I_n)$, the random number partitioning problem (NPP) seeks a partition $\boldsymbol{\sigma} \in \{-1,1\}^n$ with a small $H(\boldsymbol{\sigma}) = \frac{1}{\sqrt{n}}|\langle \boldsymbol{\sigma}, X\rangle|$. Being a worst-case NP-hard problem, it appears among the *six 'basic' NP-complete problems* by Garey and Johnson (1990) and exhibits rich features, including a rigorous phase transition (Mertens, 1998; Borgs et al., 2001). It is closely related to certain models in statistical mechanics (Bauke and Mertens, 2004; Borgs et al., 2009a,b), discrepancy minimization (Spencer, 1985; Matousek, 1999; Chazelle, 2000), and the design of randomized controlled trials in statistics (Harshaw et al., 2019). Additionally, the NPP exhibits a *statistical-computational gap* between the optimal value (Karmarkar et al., 1986) and the best known polynomial-time algorithms (Karmarkar and Karp, 1982; Yakir, 1996); recent work by Gamarnik and Kızıldağ (2023) established algorithmic barriers for the NPP.

In this paper, we introduce a planted version of the random NPP: fix a $\boldsymbol{\sigma}^* \in \{-1,1\}^n$ and generate $X \sim \mathcal{N}(0, I_n)$ conditional on the event $H(\boldsymbol{\sigma}^*) \le 3^{-n}$. The random and planted models are statistically distinguishable, since in the former case $\min_{\boldsymbol{\sigma}} H(\boldsymbol{\sigma}) = \Theta(\sqrt{n}2^{-n})$ w.h.p.

We first analyze the values of $H(\boldsymbol{\sigma})$. We show that, perhaps surprisingly, planting does not induce partitions with objective values substantially smaller than 2^{-n}: we have $\min_{\boldsymbol{\sigma} \ne \pm \boldsymbol{\sigma}^*} H(\boldsymbol{\sigma}) = \widetilde{\Theta}(2^{-n})$ w.h.p. Moreover, we precisely characterize the minimal $H(\boldsymbol{\sigma})$ achievable at any fixed distance from $\boldsymbol{\sigma}^*$. Specifically, we establish that for any $\rho \in (0,1)$,

$$\min_{\boldsymbol{\sigma}: d_H(\boldsymbol{\sigma}, \boldsymbol{\sigma}^*) = \rho n} H(\boldsymbol{\sigma}) = \widetilde{\Theta}\big(2^{-n h_b(\rho)}\big),$$

w.h.p. under the planted measure, where $h_b(\cdot)$ is the binary entropy function (logarithm base 2).

Turning to the algorithmic problem, we ask whether one can efficiently find a partition $\boldsymbol{\sigma}$ with small $H(\boldsymbol{\sigma})$. We prove that planted NPP exhibits the multi Overlap Gap Property (m-OGP) at scales $2^{-\Theta(n)}$. Building on this barrier, we show that stable algorithms satisfying a natural anti-concentration property cannot find partitions with $H(\boldsymbol{\sigma}) = 2^{-\Theta(n)}$. Stable algorithms are those for which a small perturbation of its inputs leads to only a small change in its output—namely, the returned partition. Numerous powerful algorithmic frameworks exhibit such an input-stability.

This is the first instance where the m-OGP rules out stable algorithms in a planted setting. Our results demonstrate that the multi OGP framework, previously developed for unplanted models, extends naturally to planted ones when the goal is to recover low-objective solutions. They further point to a statistical–computational gap: although the random and planted NPP are statistically distinguishable, we conjecture that no polynomial-time algorithm can distinguish them with non-trivial advantage. Our results demonstrate that planted NPP harbors intriguing features and it is a particularly promising model for probing algorithmic barriers in planted problems.

Acknowledgments

I would like to thank David Gamarnik and Muhammet Emre Şahinoğlu for valuable discussions during the initial stages of this work, to Cindy Rush for a helpful discussion during which she suggested a different version of planting, and to anonymous ALT 2026 referees and the area chair for their detailed feedback.

[*] Extended abstract. Full version appears as (Kızıldağ, 2023).

References

Heiko Bauke and Stephan Mertens. Universality in the level statistics of disordered systems. *Physical Review E*, 70(2):025102, 2004.

Christian Borgs, Jennifer Chayes, and Boris Pittel. Phase transition and finite-size scaling for the integer partitioning problem. *Random Structures & Algorithms*, 19(3-4):247–288, 2001.

Christian Borgs, Jennifer Chayes, Stephan Mertens, and Chandra Nair. Proof of the local rem conjecture for number partitioning. i: Constant energy scales. *Random Structures & Algorithms*, 34(2):217–240, 2009a.

Christian Borgs, Jennifer Chayes, Stephan Mertens, and Chandra Nair. Proof of the local rem conjecture for number partitioning. ii. growing energy scales. *Random Structures & Algorithms*, 34(2):241–284, 2009b.

Bernard Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, 2000. doi: 10.1017/CBO9780511626371.

David Gamarnik and Eren C Kızıldağ. Algorithmic obstructions in the random number partitioning problem. *The Annals of Applied Probability*, 33(6B):5497–5563, 2023.

Michael R. Garey and David S. Johnson. *Computers and Intractability; A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., USA, 1990. ISBN 0716710455.

Christopher Harshaw, Fredrik Sävje, Daniel Spielman, and Peng Zhang. Balancing covariates in randomized experiments using the gram-schmidt walk. *arXiv preprint arXiv:1911.03071*, 2019.

Narendra Karmarkar and Richard M Karp. *The differencing method of set partitioning*. Computer Science Division (EECS), University of California Berkeley, 1982.

Narendra Karmarkar, Richard M Karp, George S Lueker, and Andrew M Odlyzko. Probabilistic analysis of optimum partitioning. *Journal of Applied probability*, 23(3):626–645, 1986.

Eren C Kızıldağ. Planted random number partitioning problem. *arXiv preprint arXiv:2309.15115*, 2023.

Jiri Matousek. *Geometric discrepancy: An illustrated guide*, volume 18. Springer Science & Business Media, 1999.

Stephan Mertens. Phase transition in the number partitioning problem. *Physical Review Letters*, 81 (20):4281, 1998.

Joel Spencer. Six standard deviations suffice. *Transactions of the American mathematical society*, 289(2):679–706, 1985.

Benjamin Yakir. The differencing algorithm LDM for partitioning: a proof of a conjecture of Karmarkar and Karp. *Mathematics of Operations Research*, 21(1):85–99, 1996.