Wedge Sampling: Efficient Tensor Completion with Nearly-Linear Sample Complexity

Hengrui Luo HRLUO@RICE.EDU
Department of Statistics, Rice University

Anna Ma ANNA.MA@UCI.EDU
Department of Mathematics, University of California, Irvine

Ludovic Stephan LUDOVIC.STEPHAN@ENSAI.FR
Univ Rennes, Ensai, CNRS, CREST-UMR 9194, F-35000 Rennes, France

Yizhe Zhu YIZHEZHU@USC.EDU
Department of Mathematics, University of Southern California

Editors: Steve Hanneke and Tor Lattimore

Abstract

We introduce *Wedge Sampling*, a new non-adaptive sampling scheme for low-rank tensor completion. We study the recovery of an order-k low-rank tensor of dimension $n \times \cdots \times n$ from a subset of its entries. Unlike the standard uniform entry model, in which entries are sampled independently from $[n]^k$, wedge sampling allocates observations to structured length-two patterns, or wedges, in an associated bipartite sampling graph obtained from tensor unfolding. By directly promoting these length-two connections, the sampling design strengthens the spectral signal that underlies efficient initialization, particularly in regimes where uniform entry sampling is too sparse to generate sufficiently many informative correlations.

Our main result shows that this change in sampling paradigm enables polynomial-time algorithms to achieve both weak and exact recovery with nearly linear sample complexity in n. At a high level, uniform entry sampling spends its sampling budget on isolated tensor entries and only creates useful second-order correlations indirectly. In contrast, wedge sampling samples pairs of entries sharing a common unfolded index, thereby directly producing the correlations needed for accurate spectral subspace estimation. This leads to reliable spectral initialization using only $\tilde{O}(n)$ observations, substantially below the $\tilde{O}(n^{k/2})$ sample complexity typically required by efficient methods under uniform entry sampling.

The proposed approach is also plug-and-play. A wedge-sampling-based spectral initializer can be combined with existing refinement procedures, including spectral denoising and gradient-based nonconvex optimization, using only an additional $\tilde{O}(n)$ uniformly sampled entries. Thus, rather than replacing the refinement stage of existing tensor completion algorithms, wedge sampling addresses the initialization bottleneck that prevents these methods from operating at nearly linear sample size. Our analysis develops the concentration and perturbation tools needed to handle the dependencies induced by wedge sampling, including concentration for long matrices under wedge sampling, leave-one-out $\ell_{2,\infty}$ singular subspace guarantees, and sparse tensor concentration under an incoherent norm.

Overall, our results suggest that the statistical-to-computational gap highlighted in (Barak and Moitra, 2022) is, to a large extent, tied to the uniform entry sampling model for tensor completion. Alternative non-adaptive measurement designs that guarantee a sufficiently strong initialization, such as wedge sampling, can mitigate this barrier and enable computationally efficient tensor completion at nearly linear sample complexity.

Keywords: Tensor completion, statistical-to-computational gap, wedge sampling, random tensor

1. Extended abstract. Full version appears as `https://arxiv.org/abs/2602.05869v2`.

Acknowledgments

H.L. was partially supported by the National Science Foundation Grant NSF-DMS 2412403. A.M. was partially supported by the Rose Hills Innovators Fellowship and the Simons Grant MPS-TSM-00014026. Y.Z. was partially supported by the Simons Grant MPS-TSM-00013944.

References

Boaz Barak and Ankur Moitra. Noisy tensor completion via the sum-of-squares hierarchy. *Mathematical Programming*, 193(2):513–548, 2022.

Changxiao Cai, Gen Li, Yuejie Chi, H Vincent Poor, and Yuxin Chen. Subspace estimation from unbalanced and incomplete data matrices: $\ell_{2,\infty}$ statistical guarantees. *The Annals of Statistics*, 49(2), 2021.

Changxiao Cai, Gen Li, H Vincent Poor, and Yuxin Chen. Nonconvex low-rank tensor completion from noisy data. *Operations Research*, 70(2):1219–1237, 2022.

Yudong Chen. Incoherence-optimal matrix completion. *IEEE Transactions on Information Theory*, 61(5):2909–2923, 2015.

Yuxin Chen, Yuejie Chi, Jianqing Fan, and Cong Ma. Spectral methods for data science: A statistical perspective. *Foundations and Trends® in Machine Learning*, 14(5):566–806, 2021.

Cullen Haselby, Mark Iwen, Santhosh Karnik, and Rongrong Wang. Tensor deli: Tensor completion for low cp-rank tensors via random sampling. *arXiv preprint arXiv:2403.09932*, 2024.

Andrea Montanari and Nike Sun. Spectral algorithms for tensor completion. *Communications on Pure and Applied Mathematics*, 71(11):2381–2425, 2018.

Aaron Potechin and David Steurer. Exact tensor completion with sum-of-squares. In *Conference on Learning Theory*, pages 1619–1673. PMLR, 2017.

Dong Xia and Ming Yuan. On polynomial time methods for exact low-rank tensor completion. *Foundations of Computational Mathematics*, 19(6):1265–1313, 2019.

Ming Yuan and Cun-Hui Zhang. On tensor completion via nuclear norm minimization. *Foundations of Computational Mathematics*, 16(4):1031–1068, 2016.

Ming Yuan and Cun-Hui Zhang. Incoherent tensor norms and their applications in higher order tensor completion. *IEEE Transactions on Information Theory*, 63(10):6753–6766, 2017.