Proceedings of Machine Learning Research vol 291:1–1, 2025 38th Annual Conference on Learning Theory

Computational Equivalence of Spiked Covariance and Spiked Wigner Models via Gram-Schmidt Perturbation

Extended abstract. Full version appears as [arXiv:2503.02802v2].

Guy Bresler[*] GUY@MIT.EDU
Alina Harbuzova HADOUGHT@MIT.EDU
MIT EECS

Editors: Nika Haghtalab and Ankur Moitra

For a majority of structured statistical problems of modern interest, the best computationally efficient algorithms require significantly more data than what is theoretically sufficient in the absence of computational constraints, a phenomenon known as a "computational-statistical gap". Understanding the trade-off between computational and statistical resource is a central challenge.

This paper is motivated by the eventual goal of classification of statistical problems into equivalence classes with the implication that problems in the same class are in a precise sense fundamentally the same, with corresponding statistical-computational trade-offs, rather than coincidentally similar. In this work, we show the first average-case reduction transforming the sparse Spiked Covariance Model into the sparse Spiked Wigner Model and as a consequence obtain the first computational equivalence result between two well-studied high-dimensional statistics models.

In the Spiked Covariance Model (SpCov) of Johnstone and Lu (2004), one observes n i.i.d. samples from

$$\mathcal{N}(0, I_d + \theta u u^\top)\,,$$

where $u \in \mathbb{R}^d$ represents the signal vector and $\theta \in \mathbb{R}$ is a signal strength parameter. In the Spiked Wigner Model (SpWig) one observes a $d \times d$ matrix

$$\lambda u u^\top + W\,,$$

where W is symmetric with $\mathcal{N}(0,1)$ off-diagonal and $\mathcal{N}(0,2)$ diagonal entries, $u \in \mathbb{R}^d$ is a signal vector, and $\lambda \in \mathbb{R}$ is a signal strength parameter. For both models, it is often assumed that the signal vector u is *sparse*, having few nonzero entries Johnstone and Lu (2009).

Theorem 1 (Informal Statement of Main Theorem) *For a big part of their parameter ranges, the sparse versions of* SpCov *and* SpWig *are computationally equivalent.*

We develop algorithmic techniques that substantially change the dependence structure of the noise while leaving the signal intact. This is in contrast to the vast majority of reductions in the literature that change the planted structure but for the most part leave the noise i.i.d. Ma and Wu (2015); Hajek et al. (2015); Brennan et al. (2018); Brennan and Bresler (2020). Our main analysis tool is a novel *perturbation equivariance* result for Gram-Schmidt orthogonalization, enabling removal of dependence in the noise while preserving the signal.

[*] Supported by NSF award 2428619.

References

Matthew Brennan and Guy Bresler. Reducibility and statistical-computational gaps from secret leakage. In *Conference on Learning Theory*, pages 648–847. PMLR, 2020.

Matthew Brennan, Guy Bresler, and Wasim Huleihel. Reducibility and computational lower bounds for problems with planted sparse structure. In *Conference on Learning Theory (COLT)*, pages 48–166, 2018.

Bruce E Hajek, Yihong Wu, and Jiaming Xu. Computational lower bounds for community detection on random graphs. In *Conference on Learning Theory (COLT)*, pages 899–928, 2015.

Iain M Johnstone and Arthur Yu Lu. Sparse principal components analysis. *Unpublished manuscript*, 2004.

Iain M Johnstone and Arthur Yu Lu. On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, 104(486):682–693, 2009.

Zongming Ma and Yihong Wu. Computational barriers in minimax submatrix detection. *The Annals of Statistics*, 43(3):1089–1116, 2015.