Proceedings of Machine Learning Research vol 134:1–2, 2021 34th Annual Conference on Learning Theory

It was "all" for "nothing": sharp phase transitions for noiseless discrete channels

Jonathan Niles-Weed JNW@CIMS.NYU.COM
Courant Institute of Mathematical Sciencess & Center for Data Science, NYU

Ilias Zadik ZADIK@NYU.COM
Center for Data Science, NYU

Editors: Mikhail Belkin and Samory Kpotufe

Abstract

We prove a phase transition known as the "all-or-nothing" phenomenon for noiseless discrete channels. This class of models includes the Bernoulli group testing model and the planted Gaussian perceptron model. Previously, the existence of the all-or-nothing phenomenon for such models was only known in a limited range of parameters. Our work extends the results to all signals with sublinear sparsity.[1]

Keywords: Phase transitions, group testing, teacher-student, perceptron, all-or-nothing

1. Extended Abstract

A surprising feature of high-dimensional inference problems is the presence of *phase transitions* (see, e.g. Mezard and Montanari, 2009). A particularly striking phase transition is known as the *all-or-nothing* phenomenon (Gamarnik and Zadik, 2017; Niles-Weed and Zadik, 2020; Reeves et al., 2019; Barbier et al., 2020; Luneau et al., 2020; Truong et al., 2020; Reeves et al., 2019; Zadik, 2019). In problems evincing this phenomenon, there is a sharp break: for some critical number of samples n^*, with $(1-\epsilon)n^*$ samples it is impossible to infer *almost any* information about a parameter of interest, but with $(1+\epsilon)n^*$ samples it is possible to infer the parameter *almost perfectly*.

In this work, we establish the all-or-nothing phenomenon for a general class of models we call "noiseless discrete channels." We highlight two important special cases where our results significantly generalize previous work.

- *Bernoulli Group testing* (Dorfman, 1943): Assume a population of N individuals, on which we perform group tests to learn the identities of $k = o(N)$ infected individuals. Each individual participates in each test independently w.p. ν/k, and the test outcome is whether some group member is infected or not. Assuming $q = (1-\nu/k)^k \leq 1/2$, we establish the all-or-nothing phenomenon for this task takes place at $n^* = \lfloor k \log \frac{N}{k} / h(q) \rfloor$ tests. This generalizes the work of Truong et al. (2020) which proves the result when $\log k = o(\log N)$ and $q = 1/2$.

- *Planted Gaussian perceptron* (Zdeborová and Krzakala, 2016): Here we assume hidden k-sparse weights $\theta \in \{0,1\}^N$ and we observe i.i.d. samples of the form $(X_i, \mathbb{1}(\langle X_i, \theta\rangle \geq 0))$, where $X_i \sim N(0, I_N)$. Assuming $k = o(N)$, we establishing the all-or-nothing phenomenon takes places at $n^* = \lfloor k \log \frac{N}{k} \rfloor$ samples. Our result generalizes the work of (Luneau et al., 2020) which proves the result when $\omega(N^{\frac{8}{9}}) = k = o(N)$.

1. Extended abstract. Full version appears as [arXiv:2102.12422, v1]

References

Jean Barbier, Nicolas Macris, and Cynthia Rush. All-or-nothing statistical and computational phase transitions in sparse spiked matrix estimation. In Larochelle et al. (2020). URL https://proceedings.neurips.cc/paper/2020/hash/aaa5ebec57257fa776a1990c2bd025c1-Abstract.html.

Robert Dorfman. The detection of defective members of large populations. *The Annals of Mathematical Statistics*, 14(4):436–440, 1943. ISSN 00034851. URL http://www.jstor.org/stable/2235930.

David Gamarnik and Ilias Zadik. High dimensional linear regression with binary coefficients: Mean squared error and a phase transition. *Conference on Learning Theory (COLT)*, 2017. URL https://arxiv.org/abs/1701.04455.

Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors. *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. URL https://proceedings.neurips.cc/paper/2020.

Clément Luneau, Jean Barbier, and Nicolas Macris. Information theoretic limits of learning a sparse rule. In Larochelle et al. (2020). URL https://proceedings.neurips.cc/paper/2020/hash/713fd63d76c8a57b16fc433fb4ae718a-Abstract.html.

Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

Jonathan Niles-Weed and Ilias Zadik. The all-or-nothing phenomenon in sparse tensor PCA. In Larochelle et al. (2020). URL https://proceedings.neurips.cc/paper/2020/hash/cd0b43eac0392accf3624b7372dec36e-Abstract.html.

Galen Reeves, Jiaming Xu, and Ilias Zadik. All-or-nothing phenomena: From single-letter to high dimensions. In *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 654–658, 2019.

Galen Reeves, Jiaming Xu, and Ilias Zadik. The all-or-nothing phenomenon in sparse linear regression. In Alina Beygelzimer and Daniel Hsu, editors, *Proceedings of the Thirty-Second Conference on Learning Theory*, volume 99 of *Proceedings of Machine Learning Research*, pages 2652–2663, Phoenix, USA, 25–28 Jun 2019. PMLR.

Lan V. Truong, Matthew Aldridge, and Jonathan Scarlett. On the all-or-nothing behavior of Bernoulli group testing. arXiv: 2001.10137, 2020.

Ilias Zadik. *Computational and statistical challenges in high dimensional statistical models.* PhD thesis, Massachusetts Institute of Technology; Cambridge MA, 2019.

Lenka Zdeborová and Florent Krzakala. Statistical physics of inference: thresholds and algorithms. *Advances in Physics*, 65(5):453–552, aug 2016. doi: 10.1080/00018732.2016.1211393. URL https://doi.org/10.1080%2F00018732.2016.1211393.