Noisy Computing of the Threshold Function

Ziao Wang   ZIAOW@UMICH.EDU
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor

Nadim Ghaddar   NADIM.GHADDAR@UTORONTO.CA
Department of Electrical and Computer Engineering, University of Toronto

Banghua Zhu   BANGHUA@UW.EDU
Department of Electrical and Computer Engineering, University of Washington

Lele Wang   LELEWANG@ECE.UBC.CA
Department of Electrical and Computer Engineering, University of British Columbia

Editors: Gautam Kamath and Po-Ling Loh

Abstract

Coping with noise in computing is an important problem to consider in large systems. With applications in fault tolerance (Hastad et al., 1987; Pease et al., 1980; Pippenger et al., 1991), noisy sorting (Shah and Wainwright, 2018; Agarwal et al., 2017; Falahatgar et al., 2017; Heckel et al., 2019; Wang et al., 2024a; Gu and Xu, 2023; Kunisky et al., 2024), noisy searching (Berlekamp, 1964; Horstein, 1963; Burnashev and Zigangirov, 1974; Pelc, 1989; Karp and Kleinberg, 2007), among many others, the goal is to devise algorithms with the minimum number of queries that are robust enough to detect and correct the errors that can happen during the computation.

In this work, we consider the noisy computing of the threshold-k function. For n Boolean variables $\mathbf{x} = (x_1, \ldots, x_n) \in \{0,1\}^n$, the threshold-$k$ function $\mathsf{TH}_k(\cdot)$ computes whether the number of 1's in $\mathbf{x}$ is at least k or not, i.e.,

$$
\mathsf{TH}_k(\mathbf{x}) \triangleq \begin{cases} 1 & \text{if } \sum_{i=1}^n x_i \geq k; \\ 0 & \text{if } \sum_{i=1}^n x_i < k. \end{cases}
$$

The noisy queries correspond to noisy readings of the bits, where at each time step, the agent queries one of the bits, and with probability p, the wrong value of the bit is returned. It is assumed that the constant $p \in (0, 1/2)$ is known to the agent. Our goal is to characterize the optimal query complexity for computing the TH_k function with error probability at most δ.

This model for noisy computation of the TH_k function has been studied by Feige et al. (1994), where the order of the optimal query complexity is established; however, the exact tight characterization of the optimal number of queries is still open. In this paper, our main contribution is tightening this gap by providing new upper and lower bounds for the computation of the TH_k function, which simultaneously improve the existing upper and lower bounds.

The main result of this paper can be stated as follows: for any $1 \leq k \leq n$, there exists an algorithm that computes the TH_k function with an error probability at most $\delta = o(1)$, and the algorithm uses at most

$$
(1 + o(1)) \frac{n \log \frac{m}{\delta}}{D_{\mathsf{KL}}(p\|1-p)}
$$

---

queries in expectation. Here we define $m \triangleq \min\{k, n - k + 1\}$ and denote the Kullback-Leibler divergence between $\mathsf{Bern}(p)$ and $\mathsf{Bern}(1-p)$ distributions by $D_{\mathsf{KL}}(p\|1-p)$. Conversely, we prove that to achieve an error probability of $\delta = o(1)$, any algorithm must make at least

$$(1 - o(1)) \frac{(n - m) \log \frac{m}{\delta}}{D_{\mathsf{KL}}(p\|1-p)}$$

queries in expectation. When $m = o(n)$, the ratio between these upper and lower bounds is $1+o(1)$, and hence we provide an asymptotically tight characterization for the optimal number of queries. For general m, these bounds are tight within a multiplicative factor of 2. When specialized to the case of $k = 1$, our results recover the optimal bounds for computing the logic OR function found by Zhu et al. (2024).

Keywords: Noisy computing, sample complexity, active learning

Acknowledgement

References

Arpit Agarwal, Shivani Agarwal, Sepehr Assadi, and Sanjeev Khanna. Learning with limited rounds of adaptivity: Coin tossing, multi-armed bandits, and ranking from pairwise comparisons. In *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 39–75. PMLR, 07–10 Jul 2017.

Elwyn R Berlekamp. *Block coding with noiseless feedback*. Ph.D. thesis, MIT, Cambridge, MA, USA, 1964.

Marat Valievich Burnashev and Kamil'Shamil'evich Zigangirov. An interval estimation problem for controlled observations. *Problemy Peredachi Informatsii*, 10(3):51–61, 1974.

Moein Falahatgar, Alon Orlitsky, Venkatadheeraj Pichapati, and Ananda Theertha Suresh. Maximum selection and ranking under noisy comparisons. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1088–1096. PMLR, 06–11 Aug 2017.

Uriel Feige, Prabhakar Raghavan, David Peleg, and Eli Upfal. Computing with noisy information. *SIAM Journal on Computing*, 23(5):1001–1018, 1994.

Yuzhou Gu and Yinzhan Xu. Optimal bounds for noisy sorting. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, page 1502–1515, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450399135. doi: 10.1145/3564246. 3585131.

J. Hastad, T. Leighton, and M. Newman. Reconfiguring a hypercube in the presence of faults. In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, STOC '87, page 274–284, New York, NY, USA, 1987.

Reinhard Heckel, Nihar B. Shah, Kannan Ramchandran, and Martin J. Wainwright. Active ranking from pairwise comparisons and when parametric assumptions do not help. *Ann. Stat.*, 47(6):3099 – 3126, 2019.

M. Horstein. Sequential transmission using noiseless feedback. *IEEE Trans. Inf. Theory*, 9(3): 136–143, 1963. doi: 10.1109/TIT.1963.1057832.

Richard M. Karp and Robert Kleinberg. Noisy binary search and its applications. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, page 881–890, USA, 2007. Society for Industrial and Applied Mathematics. ISBN 9780898716245.

Dmitriy Kunisky, Daniel A. Spielman, and Xifan Yu. Inference of rankings planted in random tournaments, 2024. URL https://arxiv.org/abs/2407.16597.

M. Pease, R. Shostak, and L. Lamport. Reaching agreement in the presence of faults. *J. ACM*, 27 (2):228–234, April 1980.

Andrzej Pelc. Searching with known error probability. *Theoretical Computer Science*, 63(2):185– 202, 1989.

Nicholas Pippenger, George D Stamoulis, and John N Tsitsiklis. On a lower bound for the redundancy of reliable networks with noisy gates. *IEEE Transactions on Information Theory*, 37(3): 639–643, 1991.

Nihar B. Shah and Martin J. Wainwright. Simple, robust and optimal ranking from pairwise comparisons. *J. Mach. Learn. Res.*, 18(199):1–38, 2018.

Ziao Wang, Nadim Ghaddar, Banghua Zhu, and Lele Wang. Noisy sorting capacity. *IEEE Transactions on Information Theory*, 70(9):6121–6138, 2024a. doi: 10.1109/TIT.2024.3425281.

Ziao Wang, Nadim Ghaddar, Banghua Zhu, and Lele Wang. Noisy computing of the threshold function, 2024b. URL https://arxiv.org/abs/2403.07227.

Banghua Zhu, Ziao Wang, Nadim Ghaddar, Jiantao Jiao, and Lele Wang. Noisy computing of the or and max functions. *IEEE Journal on Selected Areas in Information Theory*, 5:302–313, 2024. doi: 10.1109/JSAIT.2024.3396787.