MR4600989 62G05 62C20 62G10

Butucea, Cristina (F-ENSAE-CRE); **Rohde, Angelika** (D-FRBG);
Steinberger, Lukas (A-WIEN-SOR)

Interactive versus noninteractive locally differentially private estimation: two elbows for the quadratic functional. (English. English summary)

Ann. Statist. **51** (2023), *no.* 2, 464–486.

In this article, an estimation problem under α-local differential privacy (LDP) is investigated. Suppose that n individual data providers hold confidential data $\{X_i\}_{i=1}^n$, which are assumed to be drawn independently from an unknown distribution having a density f with support on $[0,1]$. Also suppose that for the privacy protection of data providers, the sanitized data $\{Z_i\}_{i=1}^n$ are generated from the original data $\{X_i\}_{i=1}^n$ so that α-LDP is satisfied.

Under this setup, the authors compare minimax optimal convergence rates of estimators of the quadratic functional of the density $D(f)=\int_0^1 f^2(x)dx$ for two cases: the *noninteractive* (NI) and *sequentially-interactive* (SI) cases. For each case, empirical wavelet coefficients based on X_i are sanitized by adding Laplace noise to generate Z_i.

For NI, the i-th individual generates his/her own Z_i from the original X_i independently of all other individuals. A U-statistic of order 2 is considered as the estimator of $D(f)$ in this case. It is demonstrated that the optimal convergence rate of this estimator improves from a nonparametric one to the parametric α-private one $(n\alpha^2)^{-1/2}$ if $s>3/4$ where s is the degree of smoothness in the density f.

In contrast, for SI, the i-th individual has access to the sanitized data that were generated by other individuals in order to generate his/her own Z_i. For concreteness, the entire sample is split into two subsamples of equal sample sizes, and the first subsample is used to compute a density estimate. An estimator of $D(f)$ is then constructed under the assumption that individuals belonging to the second subsample can generate their own sanitized data after observing the density estimate. The convergence rate of the two-step estimator is shown to improve from a nonparametric one to the parametric α-private one $(n\alpha^2)^{-1/2}$ if $s>1/2$. Even when $s\le 1/2$ so that the estimators for NI and SI both have nonparametric convergence rates, the rate for SI is faster than the one for NI. Such rate improvements of SI over NI have not been observed in many other estimation problems under LDP. Monte Carlo simulations also confirm better finite-sample properties of SI over NI for a small α (and thus under a strong privacy protection). *Masayuki Hirukawa*

[References]

1. Acharya, J., Canonne, C. L., Sun, Z. and Tyagi, H. (2022). The role of interactivity in structured estimation. Preprint. Available at arXiv:2203.06870.
2. Berrett, T. and Butucea, C. (2020). Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms. *NeurIPS* **34**.
3. Bickel, P. J. and Ritov, Y. (1988). Estimating integrated squared density derivatives: Sharp best order of convergence estimates. *Sankhyā Ser. A* **50** 381–393. MR1065550 MR1065550
4. Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11–29. MR1331653 https://doi.org/10.1214/aos/1176324452
5. Butucea, C. (2007). Goodness-of-fit testing and quadratic functional estimation from indirect observations. *Ann. Statist.* **35** 1907–1930. MR2363957 https://doi.org/10.1214/009053607000000118 MR2363957
6. Butucea, C. and Issartel, Y. (2021). Locally differentially private estimation of nonlinear functionals of discrete distributions. *NeurIPS* **34**.

7. Butucea, C., Rohde, A. and Steinberger, L. (2023). Supplement to "Interactive versus noninteractive locally differentially private estimation: Two elbows for the quadratic functional." https://doi.org/10.1214/22-AOS2254SUPP
8. Butucea, C., Dubois, A., Kroll, M. and Saumard, A. (2020). Local differential privacy: Elbow effect in optimal density estimation and adaptation over Besov ellipsoids. *Bernoulli* **26** 1727–1764. MR4091090 https://doi.org/10.3150/19-BEJ1165 MR4091090
9. Cai, T. T. and Low, M. G. (2005). Nonquadratic estimators of a quadratic functional. *Ann. Statist.* **33** 2930–2956. MR2253108 https://doi.org/10.1214/009053605000000147 MR2253108
10. Cai, T. T. and Low, M. G. (2006). Optimal adaptive estimation of a quadratic functional. *Ann. Statist.* **34** 2298–2325. MR2291501 https://doi.org/10.1214/009053606000000849 MR2291501
11. Cai, T. T., Wang, Y. and Zhang, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *Ann. Statist.* **49** 2825–2850. MR4338894 https://doi.org/10.1214/21-aos2058 MR4338894
12. Collier, O., Comminges, L. and Tsybakov, A. B. (2017). Minimax estimation of linear and quadratic functionals on sparsity classes. *Ann. Statist.* **45** 923–958. MR3662444 https://doi.org/10.1214/15-AOS1432 MR3662444
13. DeVore, R. A., Jawerth, B. and Popov, V. (1992). Compression of wavelet decompositions. *Amer. J. Math.* **114** 737–785. MR1175690 https://doi.org/10.2307/2374796 MR1175690
14. de la Peña, V. H. and Montgomery-Smith, S. J. (1995). Decoupling inequalities for the tail probabilities of multivariate U-statistics. *Ann. Probab.* **23** 806–816. MR1334173 MR1334173
15. Donoho, D. L. and Nussbaum, M. (1990). Minimax quadratic estimation of a quadratic functional. *J. Complexity* **6** 290–323. MR1081043 https://doi.org/10.1016/0885-064X(90)90025-9 MR1081043
16. Duchi, J. C., Jordan, M. I. and Wainwright, M. J. (2013a). Local privacy and statistical minimax rates. In 2013 *IEEE* 54*th Annual Symposium on Foundations of Computer Science—FOCS* 2013 429–438. IEEE Computer Soc., Los Alamitos, CA. MR3246246 https://doi.org/10.1109/FOCS.2013.53 MR3246246
17. Duchi, J., Jordan, M. I. and Wainwright, M. J. (2013b). Local privacy and minimax bounds: Sharp rates for probability estimation. In *Adv. Neural Inf. Process. Syst.* 1529–1537.
18. Duchi, J. C., Jordan, M. I. and Wainwright, M. J. (2014). Local privacy, data processing inequalities, and statistical minimax rates. Preprint. Available at arXiv:1302.3203. MR3246246
19. Duchi, J. C., Jordan, M. I. and Wainwright, M. J. (2018). Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* **113** 182–201. MR3803452 https://doi.org/10.1080/01621459.2017.1389735 MR3803452
20. Duchi, J. C. and Ruan, F. (2018). The right complexity measure in locally private estimation: It is not the Fisher information. Preprint. Available at arXiv:1806.05756v1.
21. Duchi, J. C. and Ruan, F. (2020). The right complexity measure in locally private estimation: It is not the Fisher information. Available at https://arxiv.org/abs/1806.05756v3.
22. Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography* (S. Halevi and T. Rabin, eds.). *Lecture Notes in Computer Science* **3876** 265–284. Springer, Berlin. MR2241676 https://doi.org/10.1007/11681878_14 MR2241676

23. Evfimievski, A., Gehrke, J. and Srikant, R. (2003). Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 211–222. ACM, New York.
24. Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge Series in Statistical and Probabilistic Mathematics* **40**. Cambridge Univ. Press, New York. MR3588285 https://doi.org/10.1017/CBO9781107337862 MR3588285
25. Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications. Lecture Notes in Statistics* **129**. Springer, New York. MR1618204 https://doi.org/10.1007/978-1-4612-2222-4 MR1618204
26. Houdré, C. and Reynaud-Bouret, P. (2003). Exponential inequalities, with constants, for U-statistics of order two. In *Stochastic Inequalities and Applications. Progress in Probability* **56** 55–69. Birkhäuser, Basel. MR2073426 MR2073426
27. Joseph, M., Mao, J. and Roth, A. (2020). Exponential separations in local differential privacy. In *Proceedings of the* 2020 *ACM-SIAM Symposium on Discrete Algorithms* 515–527. SIAM, Philadelphia, PA. MR4141213 MR4141213
28. Joseph, M., Mao, J., Neel, S. and Roth, A. (2019). The role of interactivity in local differential privacy. In 2019 *IEEE* 60*th Annual Symposium on Foundations of Computer Science* 94–105. IEEE Comput. Soc. Press, Los Alamitos, CA. MR4228162 https://doi.org/10.1109/FOCS.2019.00015 MR4228162
29. Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S. and Smith, A. (2011). What can we learn privately? *SIAM J. Comput.* **40** 793–826. MR2823508 https://doi.org/10.1137/090756090 MR2823508
30. Klemelä, J. (2006). Sharp adaptive estimation of quadratic functionals. *Probab. Theory Related Fields* **134** 539–564. MR2214904 https://doi.org/10.1007/s00440-005-0447-2 MR2214904
31. Lam-Weil, J., Laurent, B. and Loubes, J.-M. (2022). Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint. *Bernoulli* **28** 579–600. MR4337717 https://doi.org/10.3150/21-bej1358 MR4337717
32. Laurent, B. (2005). Adaptive estimation of a quadratic functional of a density by model selection. *ESAIM Probab. Stat.* **9** 1–18. MR2148958 https://doi.org/10.1051/ps:2005001 MR2148958
33. Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** 1302–1338. MR1805785 https://doi.org/10.1214/aos/1015957395 MR1805785
34. Ritov, Y. and Bickel, P. J. (1990). Achieving information bounds in non and semiparametric models. *Ann. Statist.* **18** 925–938. MR1056344 https://doi.org/10.1214/aos/1176347633 MR1056344
35. Rohde, A. and Steinberger, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.* **48** 2646–2670. MR4152116 https://doi.org/10.1214/19-AOS1901 MR4152116
36. Smith, A. (2008). Efficient, differentially private point estimators. Preprint. Available at arXiv:0809.4794.
37. Smith, A. (2011). Privacy-preserving statistical estimation with optimal convergence rates. In *STOC'*11*—Proceedings of the* 43*rd ACM Symposium on Theory of Computing* 813–821. ACM, New York. MR2932032 https://doi.org/10.1145/1993636.1993743 MR2932032
38. Steinberger, L. (2023). Efficiency in local differential privacy. Preprint. Available

at arXiv:2301.10600.

39. Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* **60** 63–69. https://doi.org/10.1080/01621459.1965.10480775
40. Wasserman, L. and Zhou, S. (2010). A statistical framework for differential privacy. *J. Amer. Statist. Assoc.* **105** 375–389. MR2656057 https://doi.org/10.1198/jasa.2009.tm08651 MR2656057
41. Ye, M. and Barg, A. (2017). Asymptotically optimal private estimation under mean square loss. Preprint. Available at arXiv:1708.00059.

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.