Differential Privacy and Noisy Confidentiality Concepts for European Population Statistics

To allow for blinded review:
do NOT indicate author information or affiliation

Keywords: *Noise injection, population statistics, census, cell key method, differential privacy*

1 Introduction

After Dinur and Nissim published their seminal database reconstruction theorem almost two decades ago [1], it has shaped and accelerated research activities across many domains involved with data protection, data privacy and confidentiality, including disclosure control in official statistics. In its wake, 'differential privacy' was proposed in 2006 [2, 3] initially as a rigorous privacy or risk measure addressing consequences from the database reconstruction theorem. Differentially private noise mechanisms were then picked up and developed further to test and improve its use for (official) statistics; see e.g. [4, 5, 6, 7, 8, 9].

Now a first strict line must be drawn between differential privacy as a *risk measure*, and differentially private (noisy) *output mechanisms* that are engineered to manifestly guarantee a given differential privacy level. However, many other noisy output mechanisms, using bounded or unbounded noise distributions, can be set up to give at least a relaxed differential privacy guarantee too [3, 9]. For instance, the cell key method originally proposed by the Australian Bureau of Statistics [10, 11, 12] can be turned into a (relaxed) differentially private mechanism [13]. On the other hand, strictly differentially private output mechanisms require unbounded noise distributions with infinite tails, which may have particularly negative effects on utility. This paper aims to first address all these different notions separately, and then to present a consolidated discussion from both utility and risk perspectives.

Population and census-like statistics are chosen as a particular topical setting for two distinct motivations: On the one hand, treating only unweighted person counts in contingency tables simplifies many technical discussions without touching key issues of the noise discussion. On the other hand, global efforts on the 2020/2021 census round are peaking right now, with many important (and urgent) contact points to this paper. For instance, the U.S. Census Bureau has adopted a strictly differentially private noise mechanism for the 2020 U.S. census [14, 15, 16], which received mixed reactions down to grave utility concerns [17, 18]. On the other hand, the European Statistical System[1] has developed recommendations for a harmonised protection of 2021 EU census outputs based on the cell key method [19, 20, 21], where sizeable disclosure risks from massive averaging attacks were claimed recently [22]. Also these issues will be put into scope in the further course.

Our goals are to give a comparative overview of the various terms and concepts, and to present some analytic evidence that may contribute to the process of setting up an appropriate noise mechanism for particular output scenarios of official population or census statistics.

[1] The joint body of Eurostat and the national statistical institutes of all EU countries and Iceland, Liechtenstein, Norway and Switzerland. It is responsible for the development and quality assurance of official European statistics.

2 Methods

After recalling the database reconstruction theorem [1], we introduce *differential privacy* [2, 3] (DP) as a risk measure, coming either as strict ε-DP or relaxed (ε, δ)-DP. Then *noise distributions* are probability densities defining the distribution of random noise terms over a given range. While a strict ε-DP guarantee requires unbounded noise (infinite range), relaxed (ε, δ)-DP or non-manifestly DP distributions can be bounded (noise within finite range $\pm E$).
Finally, (noisy) *output mechanisms* are defined as holistic processes adding random noise drawn from a given *noise distribution* to each output statistics and managing the global amount of noise injected on the whole output. Such output mechanisms can be either *flexible* (amount of output unknown a priori) or *static* (all output is fixed before publication). Census-like unweighted contingency tables are discussed as a typical example of (often) static output, where specific risk/utility aspects are to be considered (see section 3) when setting up a suitable noisy output mechanism.

3 Results

3.1 Risk aspects

While traditional statistical disclosure control (SDC) approaches are often aimed at protecting only the small counts at particular risk of disclosure, it is meanwhile known that entire microdata databases (incl. rare or unique records) can be reconstructed accurately from too detailed output statistics [23]. Noise injection can preempt this effectively, but some considerations on the detailed noise setup apply, as argued below.
First, the database reconstruction theorem requires the noise amount to scale with the amount of published output statistics t as $\sim \sqrt{t}$, where it is argued that strictly ε-DP mechanisms generally have an overprotective scaling as $\sim t$. On the other hand, if the output amount is fixed (static output), also the noise amount can be fixed appropriately (i.e. no need for scaling).
Moreover, outputs based on bounded noise may be susceptible to attacks *exploiting output constraints* (e.g. table margins) and the finite noise bound E [22]. However, it is shown that a bounded noise distribution can be set up in a way that factually removes this risk. The given structure of a static output can be analysed to infer a risk-motivated lower limit on E (noise bound) as a function of V (noise variance).
Finally, all noisy outputs are in principle susceptible to *massive averaging* attacks [22], where V and the amount of redundancy in the output are the key risk parameters. Again the structure of a static output motivates a generic lower limit on V. The 2021 EU census programme is analysed as an example to find a V limit, which translates to an upper limit on the DP privacy budget parameter ε. The *same participants-same noise* (SPSN) principle [12, 9] is shown to be effective in reducing output redundancies.

3.2 Utility aspects

From a utility perspective focusing on census-like outputs, the key information is how much noise was added to each single output count, i.e. the variance V and bound E of the noise distribution used (cf. parameters of the cell key method [12, 24]). Utility implications of V in a bounded noise scenario were discussed in detail in [9], so the focus here is on tail effects of unbounded noise (as required for ε-DP output mechanisms). In particular, the U.S. Census Bureau announced that it will apply strictly ε-DP unbounded noise to its 2020 census outputs, with a global privacy budget being tested in the range $[0.25, 8.0]$ [14, 15]. The way this budget is split between

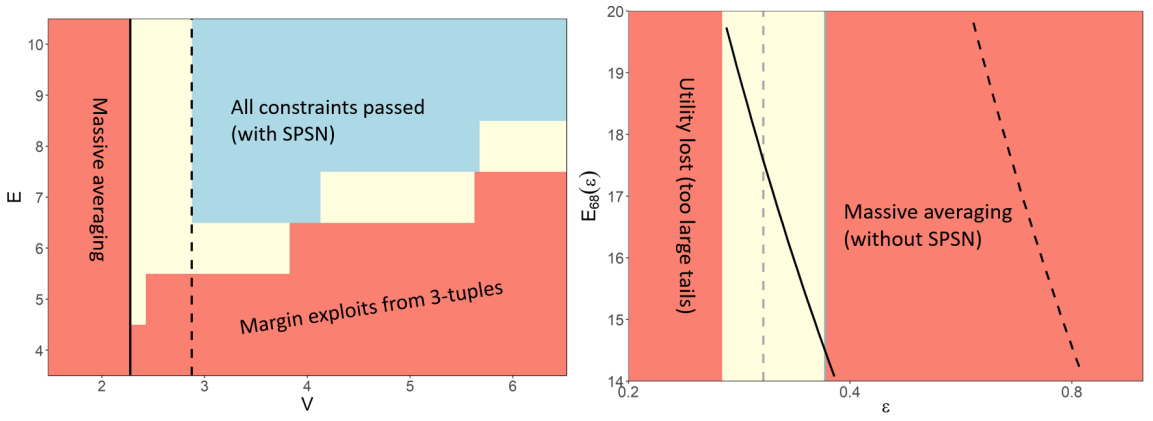

Figure 1: Generic noise parameter space highlighting regions that survive all risk/utility constraints (blue/yellow): the utility-driven generic V–E plane (left) and the risk-driven (table-level) ε range (right). Yellow regions are not excluded; they rather indicate that such setups may work in certain circumstances, or with slightly relaxed constraints. Note that the ε range (right) is a one-parameter space, where the utility constraint is from section 3.2. The SPSN principle is assumed to be invoked on the left, but not on the right (DP default).
On the right, no blue region survives all constraints conservatively: averaging (left of grey dashed line) and $E_\alpha < 20$ at $\alpha = 68\,\%$ (right of black dashed line interception with $E = 20$). When relaxing certain constraints (between black solid line interception with $E = 20$ and grey solid line), a small yellow band $\varepsilon \in [0.27, 0.37]$ remains.

individual tables [16] suggests an individual $\varepsilon \in [0.025, 0.8]$ for each table, which corresponds to noise of $\sqrt{V} \in [1.8, 57]$.
Applying such a noise setup to Local Administrative Unit (LAU) tables in the 2021 EU census output (exemplifying small-area census outputs), it is shown that tail effects of unbounded noise can lead to $> 100\,\%$ distortions for sizeable numbers of LAU units with observed counts > 20 and up to > 200. Accurate information at high geographic detail being a key unique feature of censuses, such distortions on individual LAUs may be unacceptable. This can be turned into a utility-motivated lower bound on the privacy budget ε spent on small-area tables, by requiring that unbounded noise remain within a set limit $\pm E_\alpha$ (e.g. $E_\alpha = 20$) at confidence level α for a given number of output statistics (e.g. all LAU total counts).

3.3 Risk vs. utility for upcoming censuses

In an attempt to integrate the findings of sections 3.1 and 3.2 for the scope of the 2021 EU census, the risk-motivated resp. utility-motivated parameter limits obtained there can be combined into a global picture of the generic noise parameter space: Fig. 1 illustrates that utility-driven parametrisations using individual count-level variance V and noise bound E can be set up within a range that avoids all risk/utility constraints assessed in this paper (e.g. $V \gtrsim 2$ to 3 and $E \gtrsim 5$ to 10). On the other hand, risk-driven approaches such as strictly ε-DP mechanisms with unbounded noise are severely constrained by the simultaneous requirements of risk (massive averaging) and utility (small-area accuracy) considerations. In particular, only a narrow window around individual table-level $\varepsilon \simeq 0.3$ seems to remain with acceptable compromises.
Global constraints as in Fig. 1 do depend on the exact (static) output, but in general such constraints can always be obtained systematically from the static output structure. This is what makes the risks controllable: if no satisfying parameter setup is found, the output can be curated to relax the constraints. While ε-DP as a *risk*

measure may contribute to an assessment of appropriate noise amounts, the flexibility of ε-DP *mechanisms* is heavily limited with just a single parameter (the privacy budget). It is the presence of a second parameter—the noise bound E, or δ in (ε, δ)-DP mechanisms—that adds flexibility to arbitrate between risk and utility constraints.

4 Conclusions

Recent results suggest that random noise methods are the most effective countermeasure to systematic database reconstruction [1] attacks (which can reveal all rare and unique microdata records), where the amount of noise should scale with output detail. This scaling rule implies a first important notion: flexible output mechanisms (where the output detail or'complexity' is not fixed a priori) require some kind of noise scaling and are thus much harder to realise within reasonable risk and utility constraints. On the other hand, static output mechanisms (pre-fixed output complexity) allow for a diligent curation, including controlling risks and assessing risk/utility trade-off to fix a static noise amount. Unless imposed by external constraints, a move from static to flexible output mechanisms should be considered very carefully.

Differential privacy (DP) is a useful concept to quantify risk irrespective of a particular output scenario, and hence to compare risk levels consistently between various SDC approaches [2, 7]. DP risk measures may thus contribute to a broadly based SDC assessment. Moreover, DP provides for automatic noise scaling with output complexity, as required by flexible output mechanisms. However, this paper suggests that the complexity scaling of DP noise levels is over-protective for increasingly complex outputs, so DP inferences on absolute noise levels should be handled with care, especially with complex static outputs.

The paper makes a clear separation between DP risk measures and DP output mechanisms, where the latter may give strict ε-DP or relaxed (ε, δ)-DP guarantees with ε the total privacy budget spent on the entire output. However, strictly ε-DP mechanisms must employ unbounded random noise distributions, while relaxed (ε, δ)-DP or not manifestly DP mechanisms can have bounded distributions. It is shown that in static output scenarios, typical generic risks such as margin exploits and massive averaging are controllable with bounded noise, (ε, δ)-DP or not. Conversely, the unbounded noise of strictly ε-DP mechanisms may lead to severe utility damage when the noise amount is tuned up to evade averaging risks. More generally, the fact that ε-DP mechanisms only have a single parameter costs a lot of flexibility.

Censuses are big national investments for comparably narrow purposes, not necessarily to answer any question any user may have on any characteristics of any subpopulation. This suggests a static output mechanism with a utility-driven parametrisation, which allows to maximise utility within purpose scope while controlling risks carefully. Finally, if particular SDC mechanisms jeopardise unique census features, they are bluntly unfit for the purpose. For the scope of the 2021 EU census round, this paper finds noise methods recommended by the European Statistical System [19], including bounded noise from the cell key method, suitable to protect outputs in a controlled way. The generic parameter space (noise variance and noise bound) is constrained by different risk or utility requirements, but various setups remain feasible. Such setups can obtain a relaxed (ε, δ)-DP guarantee, if needed. On the other hand, strictly ε-DP mechanisms are severely constrained, with only a small parameter window remaining for a possibly acceptable compromise. It seems strict ε-DP guarantees are overpriced (in utility) at least for census-like scenarios.

References

[1] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the Twenty-second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 202–210, 01 2003. DOI: 10.1145/773153.773173. URL http://www.cse.psu.edu/~ads22/privacy598/papers/dn03.pdf.

[2] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *Theory of Cryptography*, pages 265–284, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. ISBN 978-3-540-32732-5.

[3] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 486–503. Springer, 2006.

[4] Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *IEEE 24th International Conference on Data Engineering (ICDE)*, pages 277–286, 04 2008. DOI: 10.1109/ICDE.2008.4497436. URL http://www.cse.psu.edu/~duk17/papers/PrivacyOnTheMap.pdf.

[5] Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 705–714, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450300506. DOI: 10.1145/1806689.1806786. URL https://doi.org/10.1145/1806689.1806786.

[6] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. *SIAM Journal on Computing*, 41(6): 1673–1693, 2012.

[7] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014. ISSN 1551-305X. DOI: 10.1561/0400000042. URL http://dx.doi.org/10.1561/0400000042.

[8] Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy, 2016. URL https://arxiv.org/abs/1603.01887.

[9] Yosef Rinott, Christine O'Keefe, Natalie Shlomo, and C.J. Skinner. Confidentiality and differential privacy in the dissemination of frequency tables. *Statistical Science*, 33:358–385, 08 2018. DOI: 10.1214/17-STS641. URL http://eprints.lse.ac.uk/86504/7/Skinner_Dissemination-of-frequency-tables_Published.pdf.

[10] Bruce Fraser and Janice Wooton. A proposed method for confidentialising tabular output to protect against differencing. In *Monographs of Official Statistics: Work Session on Statistical Data Confidentiality*, pages 299–302, 11 2005. URL https://op.europa.eu/en/publication-detail/-/publication/bfea1d7d-bc66-4590-933d-90e6f33738c1.

[11] Jennifer K. Marley and Victoria L. Leaver. A method for confidentialising user-defined tables: Statistical properties and a risk-utility analysis. In *Int. Statistical*

Inst.: Proc. 58th World Statistical Congress (Session IPS060), pages 1072–1081, 08 2011. URL http://2011.isiproceedings.org/papers/450007.pdf.

[12] Gwenda Thompson, Stephen Broadfoot, and Daniel Elazar. Methodology for the automatic confidentialisation of statistical outputs from remote servers at the Australian Bureau of Statistics. In *Joint UNECE/Eurostat work session on statistical data confidentiality*, 10 2013. URL https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_1_ABS.pdf.

[13] James Bailie and Chien-Hung Chien. ABS perturbation methodology through the lens of differential privacy. In *Joint UNECE/Eurostat work session on statistical data confidentiality*, 10 2019. URL http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2019/mtg1/SDC2019_S2_ABS_Bailie_D.pdf.

[14] John M. Abowd. The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, page 2867, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450355520. DOI: 10.1145/3219819.3226070. URL https://doi.org/10.1145/3219819.3226070.

[15] Simson L. Garfinkel. Deploying differential privacy for the 2020 census of population and housing. In *JSM 2019 Session: Formal Privacy - Making an Impact at Large Organizations*, 07 2019. URL https://ecommons.cornell.edu/handle/1813/69529.

[16] Samantha Petti and Abraham Flaxman. Differential privacy in the 2020 US census: what will it do? Quantifying the accuracy/privacy tradeoff. *Gates Open Research*, 3, 2019. DOI: 10.12688/gatesopenres.13089.2. URL https://gatesopenresearch.org/articles/3-1722/v2.

[17] Steven Ruggles, Catherine Fitch, Diana Magnuson, and Jonathan Schroeder. Differential privacy and census data: Implications for social and economic research. *AEA Papers and Proceedings*, 109:403–08, May 2019. DOI: 10.1257/pandp.20191107. URL https://www.aeaweb.org/articles?id=10.1257/pandp.20191107.

[18] Alexis R. Santos-Lozada, Jeffrey T. Howard, and Ashton M. Verdery. How differential privacy will affect our understanding of health disparities in the United States. *Proceedings of the National Academy of Sciences*, 117(24): 13405–13412, 2020. ISSN 0027-8424. DOI: 10.1073/pnas.2003714117. URL https://www.pnas.org/content/117/24/13405.

[19] Laszlo Antal, Maël-Luc Buron, Annu Cabrera, Tobias Enderle, Sarah Giessing, Junoš Lukan, Eric Schulte Nordholt, and Andreja Smukavec. Harmonised protection of census data. https://ec.europa.eu/eurostat/cros/content/harmonised-protection-census-data_en, 2017. Accessed on 26 Aug 2020.

[20] Peter-Paul De Wolf, Tobias Enderle, Alexander Kowarik, and Bernhard Meindl. Perturbative confidentiality methods. https://ec.europa.eu/eurostat/cros/content/perturbative-confidentiality-methods_en, 2019. Accessed on 26 Aug 2020.

[21] Peter-Paul De Wolf, Tobias Enderle, Alexander Kowarik, and Bernhard Meindl. SDC Tools - user support and sources of tools for statistical disclosure control. https://github.com/sdcTools, 2019. Accessed on 26 Aug 2020.

[22] Hassan Jameel Asghar and Dali Kaafar. Averaging attacks on bounded noise-based disclosure control algorithms. *Proceedings on Privacy Enhancing Technologies*, 2020(2):358 – 378, 2020. DOI: https://doi.org/10.2478/popets-2020-0031. URL https://content.sciendo.com/view/journals/popets/2020/2/article-p358.xml.

[23] Simson L. Garfinkel, John M. Abowd, and Christian Martindale. Understanding database reconstruction attacks on public data. *Queue*, 16(5):28–53, October 2018. ISSN 1542-7730. DOI: 10.1145/3291276.3295691. URL https://doi.org/10.1145/3291276.3295691.

[24] Bernhard Meindl and Tobias Enderle. cellKey - consistent perturbation of statistical tables. In *Joint UNECE/Eurostat work session on statistical data confidentiality*, 10 2019. URL http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2019/mtg1/SDC2019_S7_Austria_and_Germany_cellKey_Meindl_AD.pdf.