Differential Privacy: Ideas, Concepts and Advances

Saswat Das

School of Mathematical Sciences, NISER, HBNI

September 21, 2022

Abstract

Amidst a host of concerns about privacy, chief of which are concerned with the advent of sophisticated data analysis on huge databases, older ”privatising” and ”anonymising” protocols have been shown to be insufficient for protecting the privacy of various data subjects and for protecting catastrophic privacy violations. So differential privacy (DP), known as the gold standard of privacy, was introduced by Dwork et al with the underlying idea that perturbing responses to queries on databases by noise addition can help enforce privacy. What follows here is a fairly comprehensive yet concise body of various case studies, motivation arguments, definitions regarding differential privacy, the elegant theory surrounding it, and tools and mechanisms for differentially private algorithm design. Also presented are some prominent DP algorithms, industrial implementations, and a brief discussion on some prominent and recent advances in differential privacy.

1 Motivation behind Differential Privacy

In a massively data driven world, concern for the privacy of individuals or entities whose data is being used for data analysis, and the possibility of said data being used or compromised to the detriment of said participants is a natural one; for otherwise it would all but render anyone under the sun unwilling and apprehensive of having their data being used for any sort of analysis. For example, a person recently diagnosed with, say, diabetes would not want their diagnosis being revealed deliberately or inadvertently for fear of their medical insurance premiums rising as a result. That is just one of the many possible ramifications of such privacy breaches, some of which we will discuss shortly to further motivate our study of differential privacy.

Provision of data to analysts has been ridden with these problems and questions about privacy for a while now, and there have been several attempts, the bulk of them not being quite successful at it, to preserve the privacy of individuals in a database, including so-called ”anonymisation” of data, and suppressing statistics, or providing merely summary statistics (viz. averages, medians etc.) in hopes of achieving that. We shall take a look at some attempts like these and the kind of attacks, if any, that have been mounted on them.

1.1 Case Study: Linkage Attack on Massachusetts Group Insurance Commission

This attack[1], which is one of the most famous examples of what is called a data linkage attack, was carried out by Sweeney in the 1990s. Back then, the Massachusetts Group Insurance Commission had begun a programme, with the personal blessing of the then governor of Massachusetts as regards its security, which allowed any researcher or analyst to access hospital visit records for every state employee for free, and since medical records are very sensitive information, an attempt at anonymising that database was made: the database included features like each individual's name, social security number, ZIP code, date of birth, sex and diagnosis, and anonymisation was carried out by simply omitting the name and social security number of each individual from the database.

However, this in itself was not sufficient, for it did not take into consideration the existence of auxiliary sources of information, containing data about individuals present in said database/medical records, which can be used to cross-reference this database with, and this is what Sweeney ended up exploiting. She used the voter rolls she obtained from the city of Cambridge, MA - which contained data on every registered voter's name, address, ZIP code, date of birth, and sex. Notice that both the databases here overlap in containing each individual's ZIP code, date of birth, and sex, which was a unique combination of features for about 87% of the American population back then, and this helped Sweeney to **link** the two databases and find out who most of the people in the "anonymised" medical records were, including the governor himself, and to rub it in, Sweeney ended up mailing the governor his own medical records that she was able to obtain.

This suffices to say that mere omission of certain parameters in a database in a bid to anonymise it is futile, especially if we do not factor in the presence of auxiliary, compromising information, and, in the presence of those, uniquely identifying combinations of information about an individual/entity. This led Sweeney, along with Samarati, to introduce the concept of k-anonymity.

1.1.1 k-anonymity

To neutralise the above vulnerability, Samarati and Sweeney [2] created k-anonymity, achieving which involves taking a database, which includes,

1. **Identifiers**: viz. names, AADHAR numbers, social security numbers, i.e. features that identify an individual/entity directly;
2. **Pseudo-indentifiers**: such as ZIP/PIN codes, date of birth etc., i.e. non-sensitive features that can potentially be used to identify an individual even if they are not identifiers;
3. **Sensitive features**: Features such as medical conditions, or whether a person smokes/is an alcoholic, that needs to be used for analysis but shouldn't be linked to an individual's identity in the database to avoid undesirable consequences.

and,

1. Removing *identifiers*;
2. Considering non-sensitive *pseudo-identifiers*, and manipulating the database, by making necessary omissions or making pseudo-identifying data less precise (e.g. replacing the

date of birth by the year of birth), till we obtain a database that is such that any combination of pseudo-identifying data is common to at least k individuals in the database.

To sum up, a database that is such that for any combination of pseudo-identifying features has at least k many individuals sharing said combination is said to be k-anonymous.

To be fair, k-anonymity would prevent an attack of the kind Sweeney mounted in the Massachusetts Group Insurance Commission's case but it is still vulnerable to other attacks, including some linkage attacks, a deeper discussion and illustration of which we will not go into lest we should digress from the task at hand. Interested readers may go through the critiques of k-anonymity, such as the one by Ganta, Kasiviswanathan, and Smith.[3]

1.2 Case Study: The Netflix Prize

This attack, also a linkage attack, was carried out successfully by Narayanan and Shmatikov [4] on a database shared by Netflix containing anonymised data of almost half a million users, which contained the movie ratings that they had given between December, 1999 and December, 2005 as part of a competition to find better recommendation algorithms for its platform. The company had claimed that the database released for analysis, by virtue of removal of identifying information about customers, was completely private, which going off the previous example, we know is not near enough to make such a bold claim.

Without going into much detail, Narayanan and Shmatikov simply began to cross-reference these Netflix provided movie ratings with movie ratings people had posted publicly on IMDb, even if it was not a straight forward process of matching as in the previous example, and ended up reidentifying several users of the platform by drawing correlations between the combinations of movies they had liked or disliked on Netflix and how well they had rated them on IMDb, and the dates on which said ratings were respectively made.

Now one might argue that a person's IMDb ratings are public anyway and this would not constitute a leak of any additional information that was not public already, this is not true, for the reidentified users were simply identified on the basis of ratings that they had submitted publicly on IMDb and which seemed to concur with their Netflix data, but there are a number of ratings of movies that people might have submitted privately on Netflix but had not rated said movies on IMDb, perhaps because it would reveal their political leanings, sexual orientation, or similarly sensitive information, and now being reidentified in the Netflix database ended up compromising their private information.

This led to Netflix being sued and a follow up to the Netflix prize being cancelled.

This just strengthens the case against adopting an approach of merely anonymising data by suppressing/omitting identifying data, and that auxiliary information can be used quite ingeniously to breach the supposed privacy of a database, even if the auxiliary database might have differences in presentation or content from the one being reidentified, or is relatively "incomplete"; strong correlations suffice to successfully execute a linkage attack.

1.3 Reconstruction Attacks

In a database containing sensitive information about individuals or entities, where said individuals (or entities; we shall simply use the all inclusive term "individual" for the sake of brevity) are represented as rows in said database, we might want to protect or obscure certain *secret bits*

pertaining to each row, or make it difficult to link an individual's identity with their secret bit, which for example might be whether or not they have a certain medical condition, their political affiliation, or something similar that would be sensitive information to share.

This problem has been long studied by statistical agencies and computer scientists since the 70s, and a family of various approaches, called *statistical disclosure limitation* (SDL), have been proposed or adopted to achieve this goal, such as[5]

1. *top-coding*, where values above a certain limit are effectively censored prior to the computation of various statistics;
2. *noise-injection*, where random noise, often from a chosen distribution, is added to some attributes in the database that would enable an analyst to draw useful statistical conclusions without revealing an individual's personal/actual data;
3. *swapping*, in which attributes belonging to different rows are swapped so as to preserve every individual row's privacy while allowing meaningful data analysis even after said swapping.

Then this was taken up in the 1970's by computer scientists who wanted to find ways to securely release summary statistics in response to queries from a database without compromising the privacy of an individual. Some proposals in this regard were

1. *Database Query Auditing:* Where the queries that can be possibly made by an analyst are restricted so as to avoid the possibility of someone making a set of queries that would reveal an individual row's information or the presence of an individual in said database;
2. *Adding noise to the data in the original data in the database itself*;
3. *Adding noise to the results produced due to queries*.

Out of these approaches, adding noise was proven to be the most viable one because as the volume and number of features of a database increase, the complexity of query auditing increases exponentially, and it is hard to decide for a large database what combination of queries might be dangerous, or innocuous.

We shall look into some of those techniques, their utility, and their flaws, if any, as required for our discussion in due time, and build on those to find some very interesting results and techniques.

Now what reconstruction attacks aim to do is to somehow, given the information provided or known to the adversary/analyst, *reconstruct* the database to find the secret bits of each individual.

First we describe a fairly straightforward instance of a reconstruction attack which had some significant effect on something as vital as the conduct of the US Census (starting with the 2020 census).

1.3.1 Case Study: Reconstruction Attack on the 2010 US Census

When a census is conducted, specifically an American census, a data collector collects information such as a person's name, age, sex, race, address, marital status, and more; this collection of data is of course highly identifying and therefore is sensitive. This sensitive raw data is called the *microdata*, and under American law (Title 13, Section 9 of the U.S. Code), anything the Census Bureau publishes must not contain data that can be identified with any particular individual or establishment, and ergo, the Census Bureau cannot publish data such

Statistic	Group	Count	Median Age	Mean Age
1A	total population	7	30	38
2A	female	4	30	33.5
2B	male	3	30	44
2C	black or African American	4	51	48.5
2D	white	3	24	24
3A	single adults	(D)	(D)	(D)
3B	married adults	4	51	54
4A	black or African American female	3	36	36.7
4B	black or African American male	(D)	(D)	(D)
4C	white male	(D)	(D)	(D)
4D	white female	(D)	(D)	(D)
5A	persons under 5 years	(D)	(D)	(D)
5B	persons under 18 years	(D)	(D)	(D)
5C	persons 64 years or older	(D)	(D)	(D)
	Note: Married persons must be 15 or over.			

Table 1: Fictional Block Census Data

as somebody's name, address, or any other identifying information. In short, unprocessed microdata cannot be published in Census reports as it is, for privacy related reasons.

Therefore, what the Census Bureau, and concerned statistical agencies did up to the 2010 census was publish Census data blockwise, or citywise, or statewise, but not on an individual level, by aggregating statistics, i.e. providing values of measures of central tendencies (viz. arithmetic mean and median) of certain values for features like the age of people living in an area, for various demographics.

Of course, a quick observation here would be that if there were only one, or two, or just a handful of individuals belonging to a category, then publishing the mean or the median of, say, their ages would constitute a privacy violation, for example, if two people lived in an area and the mean of the value of the age living in said area was published, one of them, knowing their own age of course, could easily calculate that for the other person, which is a privacy violation, insignificant as it may seem.[5] Or alternatively, if two Asian men and three Asian people lived in a block, then it is a straightforward inference that there is only one Asian woman in that block, and if the report released the mean or median of the age of Asian females in the block, it would effectively disclose the age of the Asian woman.

To counter this, data of this kind, where easy inferences like above could be drawn, was actually omitted from a Census report and marked with a **(D)** instead.

Even after aggregation, we can draw inferences from the aggregate statistics provided in the report, which can provide redundant information about a person's personal information (here we are considering the value of someone's age, as earlier, as the sensitive data to be hidden), and help us determine constraints on the possible values of the microdata and after determining enough constraints, we can deduce conclusively as to what someone's microdata might be.

To make this clearer, let us take the published data of a fictional block in the above prescribed format (provided in [5] for demonstrative purposes) and see how we can compromise the privacy of certain individuals here.

In Table 1, we can easily make certain assumptions: ages are reported as non negative integers, and that the value of anybody's age has to be somewhere between 0 and 125 (both inclusive), with the upper bound thus chosen as the oldest concretely recorded age of any human so far is 122, and setting it at 125 leaves room for unreported cases of ages exceeding even that.

Now let us look at statistic 2B, here we have 3 males in this block with a median age of 30 and mean age of 44. Without the information about the mean and median ages, we would have $\binom{125}{3} = 317750$ many possible age values for the males, but with the information about the values of measures of central tendency of the ages, we can narrow that down to merely 30 combinations, by seeing that one person has the age 30, the median age, trivially, and the other two males have a mean age of $\frac{1}{2}(3 \times 44 - 30) = \frac{1}{2}(102) = 51$, and this suffices to have one of those other two males being able to calculate the other male's age. Apart from that, even if an observer was not one of these two males, a reconstruction attack could still be mounted by taking every constraint into account and creating an appropriate mathematical model with the same, and then (even if solving such a system of constraints is NP hard in general there exist efficient solvers for such problems viz. SAT solvers) we can use an automatic solver to find a suitable assignment for all the unknowns satisfying all of those constraints, given by statistic 2B and other statistics that could help reconstruct personal data pertaining to statistic 2B. We shall not go into details of this process, but an interested reader can take a thorough look at the paper on this by Garfinkel et al[5].

Using a similar process, it was shown that this reconstruction attack succeeded in reconstructing the microdata precisely for 46% of the American population, and allowing for an error to the tune of ± 1 year in the age attribute, for an astounding 71% of the American census respondents.

So as mentioned earlier, this led to drastic changes in the conduct of the 2020 US Census, which will involve use of the eponymous subject of this paper, differential privacy, to avoid such attacks from working.

2 Some Prerequisite Results from Mathematics/Statistics

Here we shall mention some (well known) mathematical/statistical facts that we shall be making use of moving forward.[6]

Note that here for $n \in \mathbb{N}$, we define $[n] := \{i \in \mathbb{N} : i \leq n\}$ (i.e. the set of the first n natural numbers.)

Theorem 2.1. ***Additive Chernoff's Bound***
Let $X_1, \ldots, X_n$ be independent random variables, each of which is bounded as $0 \leq X_i \leq 1, \forall\, i \in [n]$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the mean of all these X_i's and let $\mu = \mathbb{E}(\bar{X})$ denote their expected mean, then

$$\Pr[\bar{X} > \mu + \varepsilon] \leq e^{-2m\varepsilon^2},$$
$$\Pr[\bar{X} < \mu - \varepsilon] \leq e^{-2m\varepsilon^2}.$$

Theorem 2.2. ***Multiplicative Chernoff's Bound***
Let $X_1, \ldots, X_n$ be independent random variables, each of which is bounded as $0 \leq X_i \leq 1, \forall\, i \in [n]$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the mean of all these X_i's and let $\mu = \mathbb{E}(\bar{X})$ denote their expected mean, then

$$\Pr[\bar{X} > \mu(1+\varepsilon)] \leq e^{-2m\varepsilon^2/3},$$
$$\Pr[\bar{X} < \mu(1-\varepsilon)] \leq e^{-2m\varepsilon^2/2}.$$

Theorem 2.3. ***Chernoff-Hoeffding Bound***
For a finite set of random variables $\{X_i : i \in [r]\}$, if for known a_i, b_i, we have that $X_i \in [a_i, b_i]$, let $\Delta_i = b_i - a_i$. Let $M = \sum_{i\in[r]} X_i$, then $\forall\, \alpha \in (0, \frac{1}{2}$,

$$\Pr[|M - \mathbb{E}[M]| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{\sum_{i\in[r]} \Delta_i^2}\right).$$

Theorem 2.4. ***Stirling's Approximation***
For large values of n, $n!$ can be approximated by $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

We also state the following well known inequality below without proof.

Theorem 2.5. ***Azuma's Inequality***
Let f be a function of m random variables $X_1, \ldots, X_m$, with each X_i taking values from a set A_i such that $\mathbb{E}[f]$ is bounded. Let c_i be such that, $\forall\, a_i, a_i' \in A_i$

$$|\mathbb{E}[f|X_1, \ldots, X_{i-1}, X_i = a_i] - \mathbb{E}[f|X_1, \ldots, X_{i-1}, X_i = a_i']| \leq c_i$$

Then

$$\Pr[f(X_1, \ldots, X_m) \geq \mathbb{E}[f] + t] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{m} c_i^2}\right).$$

3 Introduction to Differential Privacy[7]

To begin with, we shall discuss what came to be known as *central differential privacy* first (to be contrasted with *local differential privacy* later); we formalise the model of data analysis as follows: a *data curator*, who is someone who holds the data contributed by various data subjects/individuals in a database, to be analysed and accessed by (potentially malicious/adversarial) analysts (who could possibly compromise the privacy of an individual, let's say viz. by uncovering their secret bits, if any).

In the framework of (central) differential privacy, the trusted curator makes a promise to the data subjects (referred to often hereon as *individuals*) that no single one of them will be affected in any manner, beneficially or adversely, as a result of their participation in a study or data analysis exercise, by virtue of providing their personal data for the same (this is a promise made by differential privacy in general with or without the presence of said trusted curator).

Ideally, databases when accessed via differential privacy *mechanisms* should be confidential and protect any individual's privacy without any need for restricting data access, query auditing, data usage agreements, or any other such regulatory measures; privacy should be enforced in and of itself in while an analyst continues to make queries to the database, but that also means that the accuracy of responses to queries must necessarily degrade with increasing number of queries being asked, lest, according to the Fundamental Law of Information Recovery[1], privacy be compromised; which it will be for exceedingly large number of queries, practically speaking, but finding better mechanisms and techniques to delay when privacy is breached is an ongoing thread of research in differential privacy.

To elaborate, we want to be able to study an entire population without learning anything more about an individual than what is publicly known. Or in other words, we want to be able to draw conclusions from a database that are practically the same whether an individual's data is or is not in it; which also confers the added advantage of deniability to an individual in terms of inclusion in a given database, i.e. a person can deny that they have a certain secret bit or even that they were present in the database plausibly.

Now of course it would be a tall order, if not impossible, to demand the conclusions drawn from a database that contains a particular individual and those from one that does not are absolutely identical, so a parameter, ϵ, also known as the privacy budget is defined to allow for small deviations within reason, dictated by the value of ϵ: smaller the value of epsilon, the more privacy we get, but at the cost of accuracy, ergo utility, of responses. Suffice it to say that designing differentially private algorithms is a delicate game of balance between preserving the privacy of individuals in a database and maximising the accuracy of responses, which sets it apart from something like, say, the cryptographic description of semantic security, where the adversary and intended recipient are distinct entities.

3.1 A Review of Techniques used for and Concerns regarding Privacy in Data Analysis

For a more detailed treatment of this subject, refer to section 1.1 of [7].

1. *Data cannot be fully anonymised and stay useful:*
   This is the pet peeve of researchers working in the domain of differential privacy; as mentioned earlier, privacy degrades with increase in accuracy of responses to queries and vice-versa. A corollary of the above statement is that the richer and more explicit the data, the more shall be its utility, but it would of course grant access to a lot of information to an analyst using which an individual's personal, including secret, data could be compromised with ease.

   Said corollary was used to justify mere anonymisation of data, as in the Netflix Prize case study and other instances of such an approach (viz. the attack on taxi cab data released by the NYC Taxi and Limousine commission.), which almost always ended up in a successful attack (often linkage attacks) reidentifying the anonymised data with fair ease.

[1] **Fundamental Law of Information Recovery[8]:** Accurate answers to too many queries to a database will destroy privacy.

Differential privacy neutralises the risk of linkage attacks as discussed earlier, as will be made evident in time. Besides, we can build off this point to segue into the next.

2. *Reidentification of Anonymised Records can have huge ramifications:*
   This is something that has been talked about at length earlier; reidentification of individual information in sensitive databases can reveal personal secrets like someone's medical history (thus violating legislations like HIPAA, and putting a person at risk of their medical insurance premiums rising).

3. *Queries over large sets are not protective:*
   Adopting a hide-in-a-crowd-in-plain-sight approach does not work, because this approach is vulnerable to fairly straightforward differencing attacks; for example an adversary could query how many people have a particular secret bit (like let's say the number of people who have tested positive for COVID or have diabetes) and how many people not named X have the same secret bit, where X is an individual being targeted by an adversary, and this would end up revealing X's secret bit (i.e. continuing with our earlier example, whether X has said disease or not).

4. *Query Auditing is Problematic:*
   As mentioned earlier, noise addition is much more feasible to implement than query auditing, whose complexity increases exponentially with increase in size of the database in question and the type of queries that can be made. To audit queries huge databases manually would be a labour intensive or practically impossible task, especially when for queries in a feature rich query language in which determining and listing all possible compromising queries is next to impossible, and adopting an algorithmic approach to auditing might be computationally infeasible or downright impossible.

   Also say an auditor, a human or an algorithm, is on the lookout for queries that might be used to execute a differencing attack, refusing to answer those specific queries is disclosive in itself: it could imply the presence of an individual in the database, or worse, help reveal what their secret bit is, based on said refusal.

5. *Purely providing Summary Statistics is not a safe option:*
   If the above mentioned differencing attack by using queries asking for summary statistics (i.e. a count) is any indication, carefully crafted queries asking for summary statistics can be used to successfully mount a differencing attack, and even some reconstruction attacks, such as the one carried out on one of the Human Genome Project's databases that was able to identify individuals based on their data in said database, which was a major privacy violation and exposed individuals' sensitive data in the process. (The discussion of how sensitive the HGP's data is and possible consequences of any unintended disclosure is a whole another discussion.)

6. *Revelation of seemingly mundane and "everyday" facts is perilous:*
   A common example cited, including in [7], to explain this point is that let us say that someone is accustommed to buying white, sliced bread everyday over a year or two, and they suddenly stop buying it from a certain day onwards; this can be suggestive of that person having been diagnosed with something like diabetes. So we should consider protecting an individual row in the database in its entirety just as important as protecting certain sensitive, secret bits.

7. *Compromising a few outliers' information - is it advisable?:*
   Now this is highly subjective, in certain cases a technique that only compromises a few outliers' data and protect that of everyone else by then would be seen as being appropriate, but when in cases where the outliers' data is precisely the most sensitive, there is a problem.

   Suffice it to say that this is not a wholesale dismissal of this approach, but differential privacy allows one to circumvent this whole dilemma with privacy being preserved while not opting to reveal the outliers' data.

3.2 The Model of Computation for Differential Privacy

As mentioned earlier, in the model used to discuss differential privacy, we have a *trustworthy curator* who holds the data of several individuals in a database D with (finitely many) n rows, and the goal here is to permit statistical analysis of the entire database without compromising any individual row in particular simultaneously. The curator can optionally be replaced by a protocol run by the data subjects/individuals using cryptographically secure multiparty protocols.

There are two different models of computation in this context: the *offline* or *non-interactive* model, and the *online* or the *interactive* model.

In the *offline* model, especially but not necessarily when given a set of (type of) queries that an analyst wants to get answered, the curator produces a *synthetic database* which is a collection of summary statistics carefully chosen to answer the queries the analyst(s) might be interested in making while maintaining (differential) privacy (differential privacy will be defined more formally in a moment) and releases it to the analyst(s), post which the original database may be destroyed, because the synthetic database can now answer all of the queries securely and adequately (hopefully), and this in addition eliminates the need for a curator post the release of the synthetic database.

In the *online* model, also known aptly as the *interactive* model, the analyst is permitted to ask queries, which are essentially functions applied to a database, adaptively one after the other.

The offline model is best used when all the queries are known in advance so that noise addition can be correlated accordingly to provide an balance optimum of privacy and the best possible accuracy. Whereas when no queries are not known in advance, the offline model might run into problems as any synthetic database produced will have to be able to give accurate answers to all possible queries.

Now recall that the Fundamental Law of Information Recovery says that the provision of accurate responses to too many queries can cause a privacy catastrophe, ergo to preserve privacy, the accuracy of the responses must necessarily deteriorate with the number of queries made; aliter, it is not possible to simultaneously maintain privacy and provide accurate answers to all possible queries.

So rather than provide responses to a query straight from the database, one provides answers via a *privacy mechanism*.

Definition 3.1. Privacy Mechanism
An algorithm that takes as input a database, a universe χ of datatypes (i.e. the set of all possible database rows), random bits, and (optionally) a set of queries and outputs an *output string* is called a *privacy mechanism*.

The output string can be (but need not be) like a synthetic database; if queries are known beforehand, it should provide relatively accurate answers to them. If not, then we are in an interactive framework.

If the output string is indeed a synthetic database, then this is essentially a multiset drawn from χ, and decoding here is done by applying the query in the synthetic database and applying some specified, simple transformation (viz. multiplying by a scaling factor).

Now we shall take a look at a naïve attempt at defining private data analysis and examine it.

3.3 Attempt at Defining Private Data Analysis

Nothing is Learnt: That is the prior and posterior views of an analyst/adversary should not differ much/be nearly identical post accessing the database.
Remark(s):

1. This is a straightforward adaptation of the idea behind semantic security towards defining private data analysis. However this is not achievable; for example an analyst who has a prior view that is practically a common fallacy (an apt example straight out of [7] is that an analyst who believes that people have two left feet will definitely have a different posterior view after learning from the database that a person has a right foot and a left foot each) will stand corrected after viewing the database;
2. An easy observation to make here would be that the model of analysis of data stands in stark contrast with the model of encrypted communication over a channel as used for semantic security and cryptographic purposes: here the analyst is both legitimately at the receiving end of the responses to their queries, but also can act adversarially by attempting to mount an attack on the privacy of the database;
3. The analyst needs to learn something useful from the database that they didn't know prior to accessing the database, which flies in the face of any comparison to semantic security;
4. An adversary simulator will not be able to learn what an analyst does, as a simulator is not capable of predicting what an analyst might learn, and therefore due to this gap, it is futile to model privacy in data analysis as above.

Also note that differential privacy needs to perform and preserve privacy somehow while taking into account the presence of auxiliary sources of information to a reasonable extent, but defining what said reasonable extent is is problematic, that is putting a reasonable cap on what an analyst might know given auxiliary sources of information.

3.4 Technical Description of Differential Privacy

Differential privacy provides, well, privacy by virtue of introducing randomness into the responses provided in response to a query, which very much resembles noise addition as discussed earlier.

Now let us take a look at what is possibly one of the earliest documented implementations of ensuring privacy by the means of introducing randomness.

3.4.1 Warner's Randomised Response

This was initially described in Warner's paper on the same published as early as 1965[9]. The setup is as follows: we want to test for how many individuals in set of respondents have a certain property, $\mathcal{P}$, which might be a controversial one to possess or lack thereof; and this might ordinarily lead to said subset of respondents becoming what Warner described as a "non-cooperative" group, who might refuse to be surveyed straight away, or provide a dishonest answer which would introduce an oft difficult to assess bias in the survey results.

To simplify our discussion, let $\mathcal{P}$ be an incriminating property to possess, and therefore to simultaneously ensure that respondents answer honestly (and as a result avoid said bias) and that the respondents' privacy is not violated, and that is by giving them the ability to deny the veracity of a positive response, if they give one, is ensured by introducing randomness into the process of surveying as follows.

1. The respondent takes a fair coin and flips it;
2. If *tails* is obtained, then the respondent answers truthfully, and if *heads* is obtained, the respondent flips the coin again and
   (a) Responds affirmatively if the the outcome is heads;
   (b) Responds negatively if the outcome is tails.

It can be observed that this process endows certain privacy properties to the survey, one being the *deniability* of one's response, as only the respondent should be privy to their coin tosses and the number of times that they flipped the coin, and more importantly, it also provides sufficient accuracy for an analyst's benefit, as the results of the survey might not be completely correct at first glance but a set of simple transformations will be able to yield the correct answers to the survey's queries (recall that something similar was discussed in the context of synthetic databases and decoding the responses received to queries by applying a simple transformation to them, earlier), which are described below.

Now one gets each of tails and heads from a fair coin-flip with probability $\frac{1}{2}$, and when having obtained heads, we get affirmative and negative responses with probability$= \frac{1}{2}$ alike. So given the true answer is "Yes" for a given person, if they get tails (with probability $\frac{1}{2}$), they answer "Yes" anyway. If they get heads instead, they answer truthfully (i.e. affirmatively) if they get heads, i.e. with a probability of $P[\text{Heads}] \times P[\text{Heads}] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, and with the same probability, they get to lie and respond negatively if they get tails on the second toss.

$$\therefore \mathbb{E}[\text{Yes}] = \frac{3}{4}n(\text{has } \mathcal{P}) + \frac{1}{4}n(\text{does not have } \mathcal{P})$$

Where $\mathbb{E}[\text{Yes}]$ is the expected number of affirmative responses, and $n(X)$ is the number of people initially reported to have the property X at face value.

Randomisation is a form of noise addition to the results that still allows an analyst to perform statistical analysis on the data in the database. Without noise addition or randomisation, non trivial guarantees of privacy fail, as that would allow an adversary to link two databases and perform straightforward linkage and differencing attacks to know the secret bit(s) of any individual in them/whose data/presence differs across these databases.

We will discuss how this fits into the scheme of differential privacy formally, shortly.

3.4.2 Some Formal Definitions

N.B. For the most part while describing mechanisms or with definitions, we shall be working on discrete distributions; and whenever we opt to sample from continuous distributions, they shall be discretised to a fair degree of finite precision.

Definition 3.2. Probability Simplex
For a given discrete set S, the *probability simplex* of S, denoted by $\Delta(S)$, is given by

$$\Delta(S) = \left\{ x \in \mathbb{R}^{|S|} : x_i \geq 0 \, \forall \, i \in [|S|], \sum_{i=1}^{|S|} x_i = 1 \right\}.$$

Definition 3.3. Randomised Algorithm
A *randomised algorithm* $\mathcal{M}$ with domain R and discrete range S, given the natural map $M : R \to \Delta(S)$, $\mathcal{M}(r) = s$ with probability $(M(r))_s \; \forall \, s \in S$, with the probability space being over the coin flips of the algorithm $\mathcal{M}$.

A database x will be thought of as a collection of records from χ, and can be represented as histograms as $x \in \bar{\mathbb{N}}^{|\chi|}, x = (x_i)_{i \in \bar{\mathbb{N}}}$, where $\bar{\mathbb{N}} = \mathbb{N} \cup \{0\}$ (though often, via notational abuse, we shall often just denote the set of non negative integers by $\mathbb{N}$, and $\bar{\mathbb{N}}^{|\chi|}$ as $\mathbb{N}^{|\chi|}$) and $x_i =$ number of elements of x of type $i \in \chi$. Distances between databases shall be measured using the metric induced by the ℓ_1 norm.

Definition 3.4. ℓ_p Norm
The ℓ_p *norm* of a database x, denoted by $|| \cdot ||_p$ for $p \in \mathbb{N}$, is given by

$$||x||_p = (\sum_{i=1}^{|\chi|} |x_1|^p)^{\frac{1}{p}}.$$

The ℓ_1 norm is just the ℓ_p norm with $p = 1$, i.e. $||x||_1 = \sum_{i=1}^{|\chi|} |x_1|$, and the metric thus induced is known as the *taxicab* or *Manhattan* metric/distance.

The choice of the ℓ_1 norm for this purpose is natural, given that the taxicab distance between two databases would yield the number of records said databases differ in. We shall discuss databases and distances between them using these most of the time.

Now we are in a position to finally define differential privacy.

Definition 3.5. Differential Privacy
A randomised algorithm $\mathcal{M}$ on the domain $\mathbb{N}^{|\chi|}$ is (ε, δ)-differentially private if $\forall \, S \subseteq$

Range($\mathcal{M}$) and $\forall\, x, y \in \mathbb{N}^{|\chi|}$ such that $||x - y||_1 \leq 1$ (i.e. for *neighbouring databases*),

$$\Pr[\mathcal{M}(x) \in S] \leq \exp(\varepsilon) \Pr[\mathcal{M}(y) \in S] + \delta$$

with the probability space being over the coin flips of the mechanism $\mathcal{M}$.

For $\delta = 0$, $\mathcal{M}$ is simply said to be ε-*differentially private* or *purely differentially private* (in contrast to the case where $\delta > 0$, i.e. approximate differntial privacy), and it can be represented as when

$$\ln \left(\frac{\Pr[\mathcal{M}(x) \in S]}{\Pr[\mathcal{M}(y) \in S]} \right) \leq \varepsilon.$$

N.B. Whenever $\delta > 0$, we want δ to be less than the order of inverse of any polynomial in the size of the database. To elaborate on this, if δ is of the order of $\frac{1}{||x||_1}$ for a database x, then it effectively means adopting the just a few philosophy as described earlier, i.e. releasing the records of certain database outliers in their entirety, aliter this is equivalent to releasing the records of δ many data subjects with probability 1, or all the records with probability δ[10], which we would ideally like to avoid. (Though it is worth noting that, in the words of one of the co-authors' of the original paper that introduced approximate differential privacy[11], the Gaussian mechanism, which is something we shall soon encounter as a means of endowing (and actually the motivating example behind) approximate differential privacy, does not suffer a catastrophic failure revealing the data of individuals in this fashion precisely.)

The difference between $\varepsilon-$differential privacy and $(\varepsilon, \delta)-$differential privacy is that in $\varepsilon-$differential privacy, it is evident that the promise is one of that the mechanism's output for every run is almost equally likely to be observed on any neighbouring database; whereas for $(\varepsilon, \delta)-$differential privacy, it makes a slightly weaker promise in the sense that upon observing a given output of $\mathcal{M}(x)$, it is unlikely that the output is substantially more (or less) likely to be that of $\mathcal{M}(x)$ than that of $\mathcal{M}(y)$, where x and y are neighbouring databases.

We define a quantity useful for the sake of discussing the efficacy of mechanisms, and one whose absolute value we would seek to minimise.

Definition 3.6. Privacy Loss
The *privacy loss* incurred due to observing a value ζ in the mechanism $\mathcal{M}$'s range is given by

$$\mathcal{L}^{(\zeta)}_{\mathcal{M}(x)||\mathcal{M}(y)} = \ln \left(\frac{\Pr[\mathcal{M}(x) = \zeta}{\Pr[\mathcal{M}(y) = \zeta} \right).$$

Note that the above quantity is bounded by ε in the case of $\varepsilon-$differential privacy, and is a good measure of differential privacy afforded by a mechanism. A result, which may seem apparent intuitively, that we shall prove later on states that $(\varepsilon, \delta)-$differential privacy, on the other hand, fails to bound privacy loss by ε with probability at most being δ, and for very small values of δ, this is a good guarantee of privacy.

3.4.3 Some Properties of Differential Privacy

The following proposition about differential privacy ensures that an analyst cannot possibly increase the privacy loss of a differentially private mechanism by composing it with a mapping (i.e. by feeding its output to a function to increase privacy loss), i.e. the output of an $(\varepsilon, \delta)-$differentially private mechanism cannot be processed further to reveal any more about the original database

or any individual than it already has. This is called attempting to *post process* the output of a differentially private mechanism.

Proposition 3.7. ***Post Processing***
Let $\mathcal{M} : \mathbb{N}^{|\chi|} \to R$ be a randomised algorithm that is $(\varepsilon, \delta)-$differentially private. Let $f : R \to R'$ be an arbitrary randomised mapping. Then $f \circ \mathcal{M} : \mathbb{N}^{|\chi|} \to R'$ is $(\epsilon, \delta)-$differentially private.

(*Note:* We shall prove this proposition for a deterministic mapping f, and then we can say that the result follows for randomised mapping as every randomised mapping can be seen as a convex combination of deterministic functions[12], and differential privacy is closed under convex combination of mechanisms.)

Proof. [7]
For any neighbouring databases x, y, i.e. $||x - y||_1 \leq 1$, take any event $S \subseteq R'$. Define $\bar{S} = \{r \in R : f(r) \in S\}$. Then

$$\begin{aligned}\Pr[f(\mathcal{M}(x)) \in S] &= \Pr[\mathcal{M}(x) \in \bar{S}] \\ &\leq \exp(\varepsilon) \Pr[\mathcal{M}(y) \in \bar{S}] + \delta \qquad (\because \mathcal{M} \text{ is } (\varepsilon, \delta)\text{-differentially private}) \\ &= \exp(\varepsilon) \Pr[f(\mathcal{M}(y)) \in S] + \delta\end{aligned}$$

$\therefore$ $f \circ \mathcal{M}$ is $(\varepsilon, \delta)-$differentially private. ■

These definitions in time will give way to composition theorems which exhibit differential privacy's scalability, naturally so at that, with respect to the number of queries or the size of the group being queried on.

Theorem 3.8. ***Group Privacy***
Let $\mathcal{M} : \mathbb{N}^{|\chi|} \to R$ (R being an appropriate codomain/range of the mechanism) be an $\varepsilon-$differentially private (i.e. $(\varepsilon, 0)-$differentially private) mechanism; for groups of size k, i.e. $\forall$ databases x, y such that $||x - y||_1 \leq k$ and all $T \subseteq R$,

$$\Pr[\mathcal{M}(x) \subseteq T] \leq \exp(k\varepsilon) \Pr[\mathcal{M}(y) \subseteq T].$$

Proof. Let $\|x - y\|_1 = k$, (as if this holds for x, y that are exactly k positions apart, then this will work for $\|x - y\|_1 \leq k$ as a whole) define $x^{(0)} := x, x^{(k)} = y$, then let $\{x^{(i)}\}_{0 \leq i \leq k}$ be a sequence of databases such that $x^{(i)}, x^{(i+1)}$ are neighbouring for $0 \leq i < k$. Then $\forall\, T \subseteq R$,

$$\begin{aligned}\Pr[\mathcal{M}(x^{(0)}) \in T] &\leq e^{\varepsilon} \Pr[\mathcal{M}(x^{(1)}) \in T] \\ &\leq e^{2\varepsilon} \Pr[\mathcal{M}(x^{(2)}) \in T] \\ &\vdots \qquad\qquad\qquad\qquad \leq e^{k\varepsilon} \Pr[\mathcal{M}(x^{(k)}) \in T] \\ &= e^{k\varepsilon} \Pr[\mathcal{M}(y) \in T]\end{aligned}$$

■

Another elegant property of differential privacy that arises naturally is that of **composition**, and the most elementary form of composition is simply that if in the responses to multiple

queries are each $(\varepsilon_i, \delta_i)-$-differentially private, then the total release is $\left(\sum_i \varepsilon_i, \sum_i \delta_i\right)-$differentially private.

In layman terms, this simply means that differential privacy breaks down slowly and gradually, and not abruptly, in response to multiple queries. This is a pretty neat property to have at hand, as it allows one to define a set *privacy budget*, i.e. a threshold which if the cumulative values of all these ε_i's (and δ_i's at times too) exceeds, then the analyst is not allowed to make any more queries (useful in real life, interactive deployments).

For now, we shall be proving a simpler version of this property that can be extended to the above general case.

Theorem 3.9. ***Simple Composition***
For any two given neighbouring databases $x, y \in \mathbb{N}^{|\chi|}$, and a given sequence of outputs $z = (z_1, z_2, \ldots, z_k)$ via $\varepsilon-$differentially private mechanisms $\mathcal{M}_i : \mathbb{N}^{|\chi|} \to R, 1 \leq i \leq k$, then consider $\mathcal{M}(\cdot) := (\mathcal{M}_1(\cdot), \mathcal{M}_2(\cdot), \ldots, \mathcal{M}_k(\cdot))$. Then

$$\ln\left(\frac{\Pr[\mathcal{M}(x) = z]}{\Pr[\mathcal{M}(y) = z]}\right) \leq k\varepsilon.$$

Proof.

$$\begin{aligned}\frac{\Pr[\mathcal{M}(x) = z]}{\Pr[\mathcal{M}(y) = z]} &= \prod_{i=1}^{k} \frac{\Pr[\mathcal{M}_i(x) = z_i | (M_1(x), \ldots, M_{i-1}(x)) = (z_1, \ldots, z_{i-1})]}{\Pr[\mathcal{M}(y) = z_i | (M_1(y), \ldots, M_{i-1}(y)) = (z_1, \ldots, z_{i-1})]} \\ &\leq \prod_{i=1}^{k} e^{\varepsilon} \\ &= e^{k\varepsilon}.\end{aligned}$$

■

Theorem 3.10. ***General Composition for Two Differentially Private Algorithms***
Let $\mathcal{M}_1 : \mathbb{N}^{|\chi|} \to \mathcal{R}_1$ and $\mathcal{M}_2 : \mathbb{N}^{|\chi|} \to \mathcal{R}_2$ be ε_1- and ε_2-differentially private algorithms respectively, then $\mathcal{M}_{1,2} : \mathbb{N}^{|\chi|} \to \mathcal{R}_1 \times \mathcal{R}_2$, given by $\mathcal{M}_{1,2}(x) = (\mathcal{M}_1(x), \mathcal{M}_2(x))$ is $\varepsilon_1 + \varepsilon_2-$differentially private.

Proof. For any two neighbouring databases $x, y \in \mathbb{N}^{|\chi|}$, and for any $(r_1, r_2) \in \mathcal{R}_1 \times \mathcal{R}_2$. Then,

$$\begin{aligned}\frac{\Pr[\mathcal{M}_{1,2}(x) = (r_1, r_2)]}{\Pr[\mathcal{M}_{1,2}(y) = (r_1, r_2)]} &= \frac{\Pr[\mathcal{M}_1(x) = r_1] \Pr[\mathcal{M}_2(x) = r_2]}{\Pr[\mathcal{M}_1(y) = r_1] \Pr[\mathcal{M}_2(y) = r_2]} \\ &= \frac{\Pr[\mathcal{M}_1(x) = r_1]}{\Pr[\mathcal{M}_1(y) = r_1]} \frac{\Pr[\mathcal{M}_2(x) = r_2]}{\Pr[\mathcal{M}_2(y) = r_2]} \\ &\leq \exp(\varepsilon_1) \exp(\varepsilon_2) \\ &= \exp(\varepsilon_1 + \varepsilon_2)\end{aligned}$$

■

We can iteratively apply this on a (finite) sequence of differentially private algorithms to obtain the following useful corollary.

Corollary 3.11. *Let $\{\mathcal{M}_i\}_{i\in[k]}$, where $\mathcal{M}_i : \mathbb{N}^{|\chi|} \to \mathcal{R}_i$, be a finite sequence of ε_i-differentially private algorithms. If*

$$\mathcal{M}_{[k]} : \mathbb{N}^{|\chi|} \to \prod_{i\in[k]} \mathcal{R}_i, \mathcal{M}_{[k]}(x) = (\mathcal{M}_1(x), \ldots, \mathcal{M}_k(x)),$$

then $\mathcal{M}_{[k]}$ is $\sum_{i=1}^{k} \varepsilon_i-$differentially private.

We can further generalise this to approximate (i.e. $(\varepsilon, \delta)-$) differentially private algorithms. First, we present the following lemma without proof.

Lemma 3.12. *Let $\mathcal{M}_1 : \mathbb{N}^{|\chi|} \to \mathcal{R}_1, \mathcal{M}_2 : \mathbb{N}^{|\chi|} \to \mathcal{R}_2$ be $(\varepsilon_1, \delta_1)-$ and $(\varepsilon_2, \delta_2)$-differentially private algorithms respectively, then $\mathcal{M}_{1,2} : \mathbb{N}^{|\chi|} \to \mathcal{R}_1 \times \mathcal{R}_2, x \mapsto (\mathcal{M}_1(x), \mathcal{M}_2(x))$ is $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-differentially private.*

Now this can be iteratively applied to obtain the following composition theorem.

Theorem 3.13. *Let $\{\mathcal{M}_i\}_{i\in[k]}$, where $\mathcal{M}_i : \mathbb{N}^{|\chi|} \to \mathcal{R}_i$, be a finite sequence of $(\varepsilon_i, \delta_i)-$differentially private algorithms. If*

$$\mathcal{M}_{[k]} : \mathbb{N}^{|\chi|} \to \prod_{i\in[k]} \mathcal{R}_i, \mathcal{M}_{[k]}(x) = (\mathcal{M}_1(x), \ldots, \mathcal{M}_k(x)),$$

then $\mathcal{M}_{[k]}$ is $\left(\sum_{i=1}^{k} \varepsilon_i, \sum_{i=1}^{k} \delta_i\right)-$differentially private.

It is worth noting that more advanced, and better (in terms of minimising privacy loss) forms of composition exist. We shall discuss those in due time.

Denote $\alpha \cong_{(\varepsilon,\delta)} \beta$ if $\left|\frac{\Pr[\mathcal{M}(\alpha)\in S]-\delta}{\Pr[\mathcal{M}(\beta)\in S]}\right| \leq \varepsilon$. Then some fairly intuitive properties of differential privacy are as follows:

Monotonicity: For $\varepsilon' > \varepsilon > 0, \delta' > \delta > 0$, (ε, δ)-differential privacy implies (ε', δ')-differential privacy;

Triangle Inequality: Let $\alpha_1 \cong_{(\varepsilon_1,\delta_1)} \alpha_2$ and $\alpha_2 \cong_{(\varepsilon_2,\delta_2)} \alpha_3 \implies \alpha_1 \cong_{(\varepsilon_1+\varepsilon_2,\delta_1+\delta_2)} \alpha_3$; and

Quasi-Convexity: Let $\alpha_1 \cong_{(\varepsilon,\delta)} \alpha_2$ and $\beta_1 \cong_{(\varepsilon,\delta)} \beta_2$, then for any $p \in [0, 1], (1-p)\alpha_1 + p\beta_1 \cong_{(\varepsilon,\delta)} (1-p)\alpha_2 + p\beta_2$.

3.4.4 When is Differential Privacy not violated?

Differential privacy comes with a promise that an individual will not be additionally harmed in any way owing to their participation in a survey; that is, from an analyst's point of view, the survey will yield practically similar results as when a certain participant is in it as compared to when said participant is not in it.

Say that a smoker took part in a survey collecting statistics on smoking, and analysis based on that survey concluded that smoking causes cancer, and the smoker's insurance provider happens to be aware of his smoking or was informed about the fact by other, external sources, which leads to said smoker's insurance premiums rising.

The question is: is this a violation of differential privacy? The answer is no, as this conclusion that smoking causes cancer would have been reached with or without that smoker's

data being in the database, and is ergo not additional harm directly owing to their participation in the survey itself.

4 Implementation of Differential Privacy

4.1 Some Basic Tools and Mechanisms

Designing differentially private algorithms for any particular task makes use of certain well known method/mechanisms, chief of which are[13]

1. Global Sensitivity Method,
2. Exponential Mechanism,
3. $\varepsilon-$Differentially Private Randomised Response.

Among many others we might go into later.

4.1.1 Global Sensitivity Method

Given databases $x, y \in \mathbb{N}^{|\chi|}$ and a query $f : \mathbb{N}^{|\chi|} \to R$, and here we shall consider numeric queries, so we will take $R = \mathbb{R}^k, k \in \mathbb{N}$ (but it can be any appropriate range, depending on the context), the *global sensitivity of* f, or the ℓ_1 sensitivity of f is given by $\Delta(f) = \max_{\|x-y\|_1 \leq 1} |f(x) - f(y)|$.

The noise added to the raw data in the global sensitivity method is taken from distributions whose parameters are calibrated with respect to the global sensitivity of the query in question.

Laplace Mechanism This is among the most popular global sensitivity based mechanisms in existence, and is used to endow ε-differential privacy to query responses.
The Laplace distribution is given by $\text{Lap}(\mu, b)$, with the pdf $p(z|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|z-\mu|}{b}\right)$.
When $\mu = 0$, we denote $\text{Lap}(b) := \text{Lap}(0, b)$, where the pdf reduces to $p(z|b) := p(z|\mu = 0, b) = \frac{1}{2b} \exp\left(\frac{-|z|}{b}\right)$.

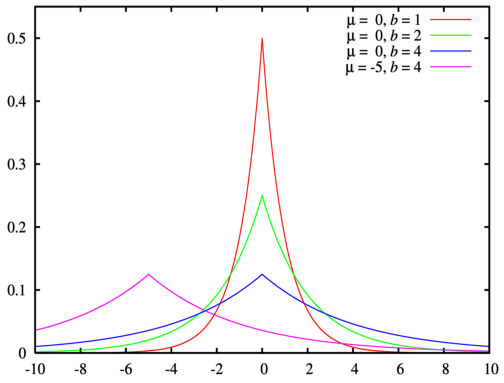

Figure 1: The Laplace Distribution

Or the Laplace mechanism, $\mathcal{M}_L(x, f(\cdot), \varepsilon) = f(x) + (Y_1, Y_2, \ldots, Y_k)$ where $Y_i \sim \frac{\Delta(f)}{\varepsilon}$.

Algorithm 1 Laplace Mechanism

Require: $x \in \mathbb{N}^{|\chi|}, f : \mathbb{N}^{|\chi|} \to \mathbb{R}^k, \epsilon > 0$
1: $\Delta \leftarrow \max_{\|x-y\|_1 \leq 1} |f(x) - f(y)|$
2: **for** $i \in [k]$ **do**
3: $\quad Y_i \leftarrow \text{Lap}(\frac{\Delta}{\varepsilon})$
4: **end for**
5: **Output** $f(x) + (Y_1, Y_2, \ldots, Y_k)$

This is useful for answering numeric queries like counting queries (which essentially ask for the number of elements in a database satisfying a particular set of properties), where $\Delta(f) = 1$, and histogram queries, where number of items of a specified set of datatypes are queried for.

Note that ε is inversely proportional to the loss in accuracy of the data release (note the $\frac{1}{\varepsilon}$ term in the pdf of the distribution from which the noise is drawm; the higher the value of ε, the less the standard deviation is, and the less is the amount of noise added, and thus the value of accuracy increases.

Proposition 4.1. *The Laplace Mechanism, with noise addition from the Laplace distribution with $\mu = 0, b = \frac{\Delta(f)}{\varepsilon}$, i.e. Lap $\left(\frac{\Delta(f)}{\varepsilon}\right)$, endows ε-differential privacy to the data release.*

Proof. Given two neighbouring databases $x, y \in \mathbb{N}^{|\chi|}$, we consider the distributions given by $\mathcal{M}_L(x, f, \varepsilon)$, and $\mathcal{M}_L(y, f, \varepsilon)$ and their respective probability distribution functions p_x, p_y. Then $\forall\, z \in \mathbb{R}^k$,

$$\begin{aligned}
\frac{p_x(z)}{p_y(z)} &= \prod_{i=1}^{k} \left(\frac{\frac{1}{2b}}{\frac{1}{2b}} \times \frac{\exp\left(-\frac{\varepsilon|f(x)_i - z_i|}{\Delta(f)}\right)}{\exp\left(-\frac{\varepsilon|f(y)_i - z_i|}{\Delta(f)}\right)} \right) \\
&= \prod_{i=1}^{k} \exp\left(\frac{\varepsilon(|f(y)_i - z_i| - |f(y)_i - z_i|)}{\Delta(f)}\right) \\
&\leq \prod_{i=1}^{k} \exp\left(\frac{\varepsilon|f(y)_i - f(x)_i|}{\Delta(f)}\right) \\
&= \exp\left(\frac{\varepsilon\|f(y) - f(x)\|_1}{\Delta(f)}\right) \\
&\leq \exp(\varepsilon)
\end{aligned}$$

■

Now we shall present a bound on the accuracy for the Laplace mechanism. For that, along with a union bound, we shall use the following fact,

$$\text{If } Y \sim \text{Lap}(b), \text{ then } \Pr[|Y| \geq t \cdot b] = e^{-t}.$$

Theorem 4.2. ***Bound on Accuracy for Laplace Mechanism***
Let $f : \mathbb{N}^{|\chi|} \to \mathbb{R}^k$, and let $y = \mathcal{M}_L(x, f(\cdot), \varepsilon)$. Then $\forall\, \delta \in (0, 1]$,

$$\Pr\left[\|f(x) - y\|_\infty \geq \ln\left(\frac{k}{\delta}\right) \cdot \left(\frac{\Delta f}{\varepsilon}\right)\right] \leq \delta.$$

Proof.

$$\begin{aligned}
\Pr\left[\|f(x) - y\|_\infty \geq \ln\left(\frac{k}{\delta}\right) \cdot \left(\frac{\Delta f}{\varepsilon}\right)\right] &= \Pr\left[\max_{i\in[k]} |Y_i| \geq \ln\left(\frac{k}{\delta}\right) \cdot \left(\frac{\Delta f}{\varepsilon}\right)\right] \\
&\leq k \cdot \Pr\left[|Y_i| \geq \ln\left(\frac{k}{\delta}\right) \cdot \left(\frac{\Delta f}{\varepsilon}\right)\right] \\
&= k \cdot e^{-\frac{k}{\delta}} && (\because\ Y_i \in \mathbf{Lap}(\Delta f/\varepsilon) \implies \Pr[|Y| \geq t \cdot b] = e^{-t}) \\
&= k \cdot \frac{\delta}{k} \\
&= \delta.
\end{aligned}$$

■

Now consider the problem of which counting query yields the highest value in a differentially private fashion. This can be solved by the application of the Laplace mechanism in what is called *Report Noisy Max*. Denote the Laplace mechanism by $\texttt{LapMech}(x, q, \varepsilon)$, for query q on database x with privacy parameter $\varepsilon > 0$. Note that only the *index* of the highest noisy count

Algorithm 2 Report Noisy Max

Require: $x \in \mathbb{N}^{|\chi|}, q_1, q_2, \ldots, q_k : \mathbb{N}^{|\chi|} \to \mathbb{R}^k, \epsilon > 0$
1: **for** $i \in [k]$ **do**
2: $\quad Y_i \leftarrow \texttt{LapMech}(x, q_i, \varepsilon)$
3: **end for**
4: **Output** $\text{argmax}_i Y_i$

is released. We claim that this is ε-differentially private.

Theorem 4.3. *The Report Noisy Max algorithm is ε-differentially private.*

Proof. Let $x = x' \cup \{a\}$, where $x, x' \in \mathbb{N}^{|\chi|}$. Let for the counting queries $q_i, 1 \leq i \leq k, q := (q_1, q_2, \ldots, q_k), c := q(x)$, and $c' := q(x')$. We use the following properties.

1. $\forall\, i \in [k], c_i \geq c'_i$ (i.e. counts are monotonic); and
2. $\forall\, i \in [k], 1 + c'_i \geq c_i$ (this is called the *Lipschitz property*).

Then for any $i \in [k]$, take r_{-i} as a draw from $\left[\text{Lap}\left(\frac{1}{\varepsilon}\right)\right]^{m-1}$ that shall be used for noise addition to all the counts except for q_i.

We shall now show that $\Pr[i|x, r_{-i}] \leq e^\varepsilon \Pr[i|x'|r_{-i}]$. Define $r^* := \min\{r_i : c_i + r_i > c_j + r_j, \forall\, j \leq i\}$. Now clearly i shall only be output when the database is x if and only if $r_i > r^*$. Then $\forall\, i, j \in [k], i \neq j$,

$$\begin{aligned}
& c_i + r^* > c_j + r_j \\
\implies & (1 + c'_i) + r^* \geq c_i + r^* > c_j + r_j \geq c'_j + r_j
\end{aligned}$$

$$\implies c'_i + (r^* + 1) > c'_j + r_j.$$

That is, when the database is x' and when r_i is added to c_i and for every other $c_j, i \neq j, r_{-i}$ is added, if $r_i \geq r^* + 1$, then the i^{th} count will be the maximum. Then over the choice of $r_i \sim \text{Lap}\left(\frac{1}{\varepsilon}\right)$

$$\begin{aligned}
&\Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|x, r_{-i}] \\
\implies &\Pr[i|x', r_{-i}] \geq \Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|x, r_{-i}] \\
\implies &\Pr[i|x, r_{-i}] \leq e^{\varepsilon} \Pr[i|x', r_{-i}].
\end{aligned}$$

Now we have to show that $\Pr[i|x', r_{-i}] \leq e^{\varepsilon} \Pr[i|x, r_{-i}]$; for this, define $r^{**} := \min\{r_i : c'_i + r_i > c'_j + r_j, \forall\, j \leq i\}$.

Then for a fixed r_{-i}, i shall be the output with database x' if and only if the noise added to the i^{th} count, $r_i \geq r^{**}$.

Then $\forall\, i, j \in [k], i \neq j$,

$$\begin{aligned}
&c'_i + r^{**} > c'_j + r_j \\
\implies &1 + c'_i + r^{**} > 1 + c'_j + r_j \\
\implies &c'_i + (r^{**} + 1) > (1 + c'_j) + r_j \\
\implies &c_i + (r^{**} + 1) \geq c'_i + (r^{**} + 1) > (1 + c'_j) + r^{**} \geq c_j + r_j
\end{aligned}$$

$\therefore$ if $r_i \geq r^{**} + 1$, then i shall be output by Report Noisy Max on the database x with noise r_i being added to the i^{th} count, and r_{-i} to the rest. $\therefore$ We have the following, over the choice of r_i.

$$\begin{aligned}
&\Pr[i|x, r_{-i}] \geq \Pr[r_i \geq r^{**} + 1] \geq e^{-\varepsilon} \Pr[r_i \geq r^{**}] = e^{-\varepsilon} \Pr[i|x', r_{-i}] \\
\implies &\Pr[i|x', r_{-i}] \leq e^{\varepsilon} \Pr[i|x, r_{-i}].
\end{aligned}$$

So we have that $\ln \left| \frac{\Pr[i|x, r_{-i}]}{\Pr[i|x', r_{-i}]} \right| \leq \varepsilon$. ■

Gaussian Mechanism This helps endow approximate (i.e. (ε, δ)-)differential privacy to a query response.

Here we define global sensitivity of a query $f : \mathbb{N}^{|\chi|} \to R$, for two neighbouring databases $x, y \in \mathbb{N}^{|\chi|}$, to be the ℓ_2- sensitivity of f,

$$\Delta_2(f) = \max_{x, y \in \mathbb{N}^{|\chi|}} \|f(x) - f(y)\|_2.$$

As the name suggests, here the noise is taken from the Gaussian distribution: $N(\mu, \sigma^2)$, which has the pdf $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.

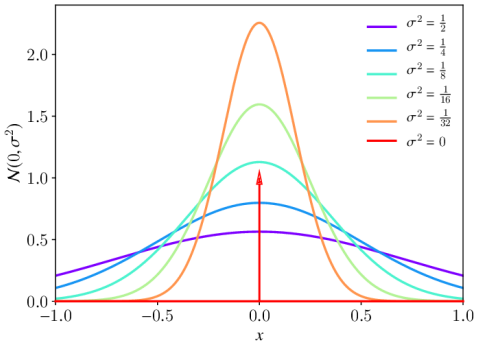

Figure 2: The Gaussian/Normal Distribution

Algorithm 3 Gaussian Mechanism

Require: $x \in \mathbb{N}^{|\chi|}, f : \mathbb{N}^{|\chi|} \to \mathbb{R}^k, \epsilon > 0$

1: $\Delta_2 \leftarrow \max_{x,y \in \mathbb{N}^{|\chi|}; \|x-y\|_1 \leq 1} \|f(x) - f(y)\|_2$
2: **for** $i \in [k]$ **do**
3: $\quad Y_i \leftarrow N(0, 2\ln(\frac{1.25}{\delta})\frac{\Delta_2^2}{\varepsilon^2})$
4: **end for**
5: **Output** $f(x) + (Y_1, Y_2, \ldots, Y_k)$

N.B. We can actually take $Y_i \sim N(0, \sigma^2), 1 \leq i \leq k$, where $\sigma \geq c\frac{\Delta_2(f)}{\varepsilon}$, and $c^2 > 2\ln\left(\frac{1.25}{\delta}\right)$.

This entire process endows $(\varepsilon, \delta)-$differential privacy to the data release. We shall not be producing the rather convoluted proof here for now.

Notice that this has certain advantages, the Gaussian distribution is more "natural" than the Laplace distribution to use, as it arises naturally in many a place. But more importantly, recall that the sum of two Gaussian random variables is again a Gaussian random variable, thus making composition and related calculations easy. All that is in addition to it providing approximate differential privacy which is a slightly weaker guarantee than $\varepsilon-$differential privacy but is strong enough for small values of δ, and thus allows less noise addition to achieve the same privacy budget (i.e. more accuracy).

4.1.2 Exponential Mechanism[14]

Sometimes just simply perturbing data by naïvely adding noise is not suitable for all purposes, for instance, (this is a classic example) we might have a seller who has an ample amount of items of a particular kind to sell to buyers who propose what is the highest amount of money (we shall call it the buyer's valuation of that item, per unit) they are willing to pay for a unit of said item.

Suppose buyer A's valuation is \$ 1, that of B is \$ 1.01, and that of C is \$ 3.01. If the seller decides to sell the item at \$ 1 each, then his total revenue will be (consider A, B, C to be the only buyers in this scenario) \$ 3; if he sells at \$ 1.01 apiece, then it will fetch him \$ 2.02, and if he sells at \$ 3.01 apiece, he will get a revenue of \$ 3.01.

But what if he sells it at \$ 3.02 apiece, perhaps given that maybe each buyer wants to keep their valuation private and thus noise perturbation gave him that value as the highest

valuation? In that case, he will actually end up selling nothing, having overshot all the valuations.

So perturbing raw data as in the global sensitivity method is not the solution here. To deal with situations like these, Sherry and Talwar came up with the *exponential mechanism*, which doesn't explicitly and simply add noise from a distribution but instead chooses a particular object with probability that increases exponentially with respect to what is called its score/utility.

Let us consider a database $x \in \mathbb{N}^{|\chi|}$, a set of objects to choose from ($\mathcal{H}$), and a score/utility function $s : \mathbb{N}^{|\chi|} \times \mathcal{H} \to \mathbb{R}$ (outputs how "good", defined contextually, an object in $\mathcal{H}$ is w.r.t. x).

Define $\Delta s = \max_{h \in \mathcal{H}} \max_{x,y \in \mathbb{N}^{|\chi|}, \|x-y\|_1} |s(x,h) - s(y,h)|$.

Algorithm 4 Exponential Mechanism

Require: $x \in \mathbb{N}^{|\chi|}, \mathcal{H}, s : \mathbb{N}^{|\chi|} \times \mathcal{H} \to \mathbb{R}$

1: $\Delta s \leftarrow \max_{h \in \mathcal{H}} \max_{x,y \in \mathbb{N}^{|\chi|}, \|x-y\|_1} |s(x,h) - s(y,h)|$

2: **Output** $\mathcal{M}_E(x, \mathcal{H}, s, \varepsilon) = h (\in \mathcal{H})$ with probability$= c \exp(\frac{\varepsilon s(x,h)}{2\Delta s})$, where $c \in \mathbb{R}_+$ is a suitable constant.

Theorem 4.4. *The exponential mechanism is ε-differentially private.*

Proof. For any two neighbouring databases x, y and some outcome $h \in \mathcal{H}$,

$$\begin{aligned}
\frac{\Pr[\mathcal{M}_E(x) = h]}{\Pr[\mathcal{M}_E(y) = h]} &= \frac{\left(\frac{\exp\left(\frac{\varepsilon s(x,h)}{2\Delta s}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(x,h')}{2\Delta s}\right)}\right)}{\left(\frac{\exp\left(\frac{\varepsilon s(y,h)}{2\Delta s}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(y,h')}{2\Delta s}\right)}\right)} \\
&= \exp\left(\frac{\varepsilon(s(x,h) - s(y,h))}{2\Delta s}\right) \frac{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(y,h')}{2\Delta s}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(x,h')}{2\Delta s}\right)} \\
&\leq \exp\left(\frac{\varepsilon}{2}\right) \exp\left(\frac{\varepsilon}{2}\right) \frac{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(x,h')}{2\Delta s}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(x,h')}{2\Delta s}\right)} \\
&= \exp(\varepsilon) \frac{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(y,h')}{2\Delta s}\right)}{\sum_{h' \in \mathcal{H}} \exp\left(\frac{\varepsilon s(y,h')}{2\Delta s}\right)} \\
&= \exp(\varepsilon)
\end{aligned}$$

■

Armed with this, we can solve the above sale problem by taking x as being the set of all the buyers' valuations, $\mathcal{H}$ as the set of all possible prices, and the score function s as the revenue earned at a given price.

The exponential mechanism is fairly versatile and can be fitted to suit various contexts. In fact, an interesting observation is that the exponential mechanism can be used to reproduce the Laplace mechanism, aliter, the Laplace mechanism can be seen as an instance of the exponential mechanism; let $f : \mathbb{N}^{|\chi|} \to \mathbb{R}$ be a numeric query with ℓ_1-sensitivity Δ on a dataset x to be answered using the Laplace mechanism. then we take $\mathcal{H} = \mathbb{R}$, and the

score function, $s(x,h) = -|f(x) - h|$, then $h \in \mathbb{R}$ is output with probability proportional to $\exp(-\frac{\varepsilon|f(x)-h}{2\Delta})$, which is precisely what the Laplace mechanism does.

Now as the exponential mechanism effectively adds noise by choosing an object whose score is close to that of the best object (let us define said best score by $OPT(x) = \max_{h\in\mathcal{H}} s(x,h)$), we stand to lose a little amount of score, so to speak. The following theorem and its corollary explain this, and therefore the utility of a response, concretely.[15]

Theorem 4.5. *Let x be a dataset, let $\mathcal{H}^* = \{h \in \mathcal{H} : s(x,h) = OPT(x)\}$ be the set of objects that achieve $OPT(x)$ as their score. Then*

$$\Pr\left[s(M_E(x)) \le OPT(x) - \frac{2\Delta}{\varepsilon}\left(\ln\left(\frac{|\mathcal{H}|}{|\mathcal{H}^*|}\right) + t\right)\right] \le \exp(-t).$$

Proof. Let $m = OPT(x) - \frac{2\Delta}{\varepsilon}\left(\ln\left(\frac{|\mathcal{H}|}{|\mathcal{H}^*|} + t\right)\right)$, then

$$\begin{aligned}
\Pr[s(M_E(x)) \le m] &= \frac{\sum_{h:s(x,h)\le m \wedge h\in\mathcal{H}} \exp\left(\frac{\varepsilon s(x,h)}{2\Delta}\right)}{\sum_{h'\in\mathcal{H}} \exp\left(\frac{\varepsilon s(x,h')}{2\Delta}\right)} \\
&\le \frac{|\mathcal{H}|\exp\left(\frac{\varepsilon m}{2\Delta}\right)}{|\mathcal{H}^*|\exp\left(\frac{\varepsilon OPT(x)}{2\Delta}\right)} \\
&= \frac{|\mathcal{H}|}{|\mathcal{H}^*|}\exp\left(\frac{\varepsilon(m - OPT(x))}{2\Delta}\right) \\
&= \frac{|\mathcal{H}|}{|\mathcal{H}^*|}\exp\left(\frac{\varepsilon(-\frac{2\Delta}{\varepsilon}\left(\ln\left(\frac{|\mathcal{H}|}{|\mathcal{H}^*|}\right) + t\right))}{2\Delta}\right) \\
&= \frac{|\mathcal{H}|}{|\mathcal{H}^*|}\exp\left(-\ln\left(\frac{|\mathcal{H}|}{|\mathcal{H}^*|}\right) + t\right) \\
&= \exp(-t)
\end{aligned}$$

■

The following corollary is simply the case where $|\mathcal{H}^*| = 1$.

Corollary 4.6. $\Pr\left[s(M_E(x)) \le OPT(x) - \frac{2\Delta}{\varepsilon}(\ln(|\mathcal{H}|) + t)\right] \le \exp(-t)$.

Now let us consider the problem of determining which datatype out of two, A, B, is more common. Then WLOG (perhaps by translation), let us consider the count of A to be 0 and $c > 0$ for B. We define score/utility s here as being the actual counts, so $\Delta s = 1$, $s(A) = 0, s(B) = c$. Then by corollary 4.6, we have that the probability of observing the wrong outcome, A, after noise addition via the exponential mechanism, is at most $2e^{-c(\varepsilon/2\Delta u)} = 2e^{-c\varepsilon/2}$. This is in contrast to understanding the error margin, so to speak, when we use algorithm 2 (Report Noisy Max), especially considering the case when positive noise is added to the count for A and negative noise is added to that for B.

So *Report One-Sided Noisy Arg-Max* was introduced to help with this problem, which adds noise drawn from the one-sided exponential distribution (let us denote it by Exp)

$\left(\text{with parameter } \begin{cases} \frac{\varepsilon}{\Delta s} \text{ if } s \text{ is monotonic;} \\ \frac{\varepsilon}{2\Delta s} \text{ otherwise} \end{cases}\right)$ to the utility of each potential output, before reporting the arg-max of the maximum noisy count.

The proof of the following result on the ε-differential privacy of Report One-Sided Noisy Arg-Max is nearly identical to the one for the theorem on ε-differential privacy of algorithm 2 (Report Noisy Max).

Theorem 4.7. *Report One-Sided Noisy Arg-Max with the parameter $\varepsilon/2\Delta s$ for the exponential distribution, is ε-differentially private.*

Proof. Let $x = x' \cup \{a\}$, where $x, x' \in \mathbb{N}^{|\chi|}$. Let for the counting queries $q_i, 1 \leq i \leq k, q := (q_1, q_2, \ldots, q_k), c := q(x)$, and $c' := q(x')$.

Then for any $i \in [k]$, take r_{-i} as a draw from $\left[\text{Exp}\left(\frac{\varepsilon}{\Delta s}\right)\right]^{m-1}$, if s is monotonic, else from $\left[\text{Exp}\left(\frac{\varepsilon}{2\Delta s}\right)\right]^{m-1}$ that shall be used for noise addition to all the counts except for q_i.

We shall now show that $\Pr[i|x, r_{-i}] \leq e^{\varepsilon} \Pr[i|x'|r_{-i}]$. Define $r^* := \min\{r_i : c_i + r_i > c_j + r_j, \forall\, j \leq i\}$. Now clearly i shall only be output when the database is x if and only if $r_i > r^*$. Then $\forall\, i, j \in [k], i \neq j$,

$$\begin{aligned} & c_i + r^* > c_j + r_j \\ \implies & (1 + c'_i) + r^* \geq c_i + r^* > c_j + r_j \geq c'_j + r_j \\ \implies & c'_i + (r^* + 1) > c'_j + r_j. \end{aligned}$$

That is, when the database is x' and when r_i is added to c_i and for every other $c_j, i \neq j, r_{-i}$ is added, if $r_i \geq r^* + 1$, then the i^{th} count will be the maximum. Then over the choice of r_i from the exponential distribution with the appropriate parameter,

$$\begin{aligned} & \Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|x, r_{-i}] \\ \implies & \Pr[i|x', r_{-i}] \geq \Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|x, r_{-i}] \\ \implies & \Pr[i|x, r_{-i}] \leq e^{\varepsilon} \Pr[i|x', r_{-i}]. \end{aligned}$$

Now we have to show that $\Pr[i|x', r_{-i}] \leq e^{\varepsilon} \Pr[i|x, r_{-i}]$; for this, define $r^{**} := \min\{r_i : c'_i + r_i > c'_j + r_j, \forall\, j \leq i\}$.

Then for a fixed r_{-i}, i shall be the output with database x' if and only if the noise added to the i^{th} count, $r_i \geq r^{**}$.

Then $\forall\, i, j \in [k], i \neq j$,

$$\begin{aligned} & c'_i + r^{**} > c'_j + r_j \\ \implies & 1 + c'_i + r^{**} > 1 + c'_j + r_j \\ \implies & c'_i + (r^{**} + 1) > (1 + c'_j) + r_j \\ \implies & c_i + (r^{**} + 1) \geq c'_i + (r^{**} + 1) > (1 + c'_j) + r^{**} \geq c_j + r_j \end{aligned}$$

$\therefore$ if $r_i \geq r^{**} + 1$, then i shall be output by Report One-Sided Noisy Arg-Max on the database x with noise r_i being added to the i^{th} count, and r_{-i} to the rest. $\therefore$ We have the following, over

the choice of r_i.

$$\begin{aligned} &\Pr[i|x, r_{-i}] \geq \Pr[r_i \geq r^{**} + 1] \geq e^{-\varepsilon} \Pr[r_i \geq r^{**}] = e^{-\varepsilon} \Pr[i|x', r_{-i}] \\ \implies &\Pr[i|x', r_{-i}] \leq e^{\varepsilon} \Pr[i|x, r_{-i}]. \end{aligned}$$

So we have that $\ln \left| \frac{\Pr[i|x, r_{-i}]}{\Pr[i|x', r_{-i}]} \right| \leq \varepsilon$. ■

4.1.3 Randomised Response

Now it is worth recalling that randomised response was formulated in its most basic form prior to the advent of differential privacy as a formally defined concept, but it has seen a resurgence post the same, especially as a primitive used to enforce what is called *local differential privacy*, which we shall discuss later.

The ε-differentially private version of randomised response is defined as given below.

In general, given $\varepsilon > 0$, for every private bit X in a piece of data, output

$$\mathcal{M}(X) = \begin{cases} X, \text{ with probability} = \frac{\exp(\varepsilon)}{1+\exp(\varepsilon)}; \\ 1 - X, \text{ with probability} = \frac{1}{1+\exp(\varepsilon)}. \end{cases}$$

For *any*[2] two bits X, Y, and for $z \in \{0, 1\}$ this disclosure is $\ln \left(\frac{\Pr[\mathcal{M}(X)=z]}{\Pr[\mathcal{M}(Y)=z]} \right) \leq \ln \left(\frac{\frac{e^{\varepsilon}}{1+e^{\varepsilon}}}{\frac{1}{1+e^{\varepsilon}}} \right) =$ ε-differentially private.

We can then aggregate and calculate the mean of n bits using $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\exp(\varepsilon)+1}{\exp(\varepsilon)-1} \cdot Y_i - \frac{1}{\exp(\varepsilon)-1} \right)$.

Also it can be seen that

$$\begin{aligned} \mathbb{E}(\hat{\mu}) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} \left(\frac{\exp(\varepsilon)+1}{\exp(\varepsilon)-1} \cdot Y_i - \frac{1}{\exp(\varepsilon)-1} \right) \right) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\exp(\varepsilon)+1}{\exp(\varepsilon)-1} \cdot \mathbb{E}\left(Y_i\right) - \frac{1}{\exp(\varepsilon)-1} \right) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\exp(\varepsilon)+1}{\exp(\varepsilon)-1} \cdot \left(\frac{\exp\left(\varepsilon\right)}{1+\exp\left(\varepsilon\right)} \times X_i + \frac{1}{1+\exp(\varepsilon)} \times (1 - X_i) \right) - \frac{1}{\exp(\varepsilon)-1} \right) \\ &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{\exp(\varepsilon)+1}{\exp(\varepsilon)-1} \cdot \left(\frac{\exp\left(\varepsilon\right)-1}{1+\exp\left(\varepsilon\right)} \times X_i + \frac{1}{1+\exp(\varepsilon)} \right) - \frac{1}{\exp(\varepsilon)-1} \right) \\ &= \frac{1}{n} \sum_{i=1}^{n} X_i. \end{aligned}$$

Which means that $\hat{\mu}$ is an unbiased estimator for the true mean.

5 Advanced Composition

The object of finding advanced composition results was two-fold: not only do we want the privacy budget to get exhausted as slowly as possible via repeated application of a privacy mechanism on a database, we also want the theorem to be able to handle more complicated applications of composition in certain, more general situations, such as

[2]This leads to a stronger guarantee that will come in handy in local differential privacy, which we shall define later.

1. The repeated application of differentially private algorithms on the same database, which is something we have covered in a previous section, and to be assured that said repeated application of the same mechanism/algorithm on the same database does not degrade its privacy guarantees too much; or
2. The combination of known differentially private mechanisms/algorithms to come up with new ones (i.e. modular design of new differentially private algorithms) and study them; and
3. The repeated application of differentially private algorithms on different databases that may have information pertaining to the same individual/set of individuals in common; even in this scenario we seek to assure the data subjects that their privacy is not compromised (or by much).

Before we can begin discussing this more powerful composition theorem[16], we need to be armed with some more theory. As a convention here, if the denominator in any fraction henceforth happens to be 0, given that the numerators shall always be positive, then we shall take the value of said fraction to be $+\infty$.

Definition 5.1. Kullback-Leibler (KL)-Divergence
Also known as *relative entropy* between two random variables Y and Z which have the same domain, this quantity is defined as

$$D(Y\|Z) = \mathbb{E}_{y\sim Y}\left[\ln \frac{\Pr[Y=y]}{\Pr[Z=y]}\right].$$

Note that D here is not a metric on the set of random variables sharing the same domain, as evidently $D(Y\|Z) \geq 0$ (for this we can use the log sum inequality[3], and then see that $\sum_y \Pr[Y=y] \ln \frac{\Pr[Y=y]}{\Pr[Z=y]} \geq \sum_{y\in Y} \log \frac{\sum_y \Pr[Y=y]}{\sum_y \Pr[Z=y]} = 1 \log \frac{1}{1} = 0$) and $D(Y\|Z) = 0 \iff Y$ and Z are identically distributed.

On the other hand, it need not be that $D(Y\|Z) = D(Z\|Y), \forall$ random variables Y, Z taking values from the same domain, and does not satisfy the triangle inequality, and can be infinite when $\text{Supp}(Y) \setminus \text{Supp}(Z) \neq \phi$.

Definition 5.2. Max Divergence and δ-Approximate Max Divergence
For two random variables Y, Z taking values from the same domain, this quantity is defined as

$$D_\infty(Y\|Z) = \max_{S\subseteq \text{Supp}(Y)} \left[\ln \frac{\Pr[Y=y]}{\Pr[Z=y]}\right].$$

The δ-Approximate Max Divergence between Y and Z is given by

$$D_\infty^\delta(Y\|Z) = \max_{S\subseteq \text{Supp}(Y):\Pr[Y\in S]\geq\delta} \left[\ln \frac{\Pr[Y=y]-\delta}{\Pr[Z=y]}\right].$$

In the light of having defined these terms, it is evident that a sufficient and necessary condition for a mechanism $\mathcal{M}$ to be ε- (respectively, (ε,δ)-) differentially private is that $\forall$ neighbouring databases $x, y \in \mathbb{N}^{|\chi|}$, $D_\infty(\mathcal{M}(x)\|\mathcal{M}(y)) \leq \varepsilon$ and $D_\infty(\mathcal{M}(y)\|\mathcal{M}(x)) \leq \varepsilon$ (respectively, $D_\infty^\delta(\mathcal{M}(x)\|\mathcal{M}(y)) \leq \varepsilon$ and $D_\infty^\delta(\mathcal{M}(y)\|\mathcal{M}(x)) \leq \varepsilon$).

[3] *Log Sum Inequality*: Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be non-negative numbers. Define $a := \sum_{i=1}^n a_i$, $b = \sum_{i=1}^n b_i$, then $\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq a \log \frac{a}{b}$., with equality iff $\frac{a_i}{b_i}$ is the same for all i.

Definition 5.3. Statistical Distance between Random Variables
For two random variables Y, Z, we define the *statistical distance* between them to be

$$\Delta(Y, Z) := \max_S |\Pr[Y \in S] - \Pr[Z \in S]|$$

The above distance can easily be seen to be a metric. We term Y and Z to be δ-*close* if $\Delta(Y, Z) \leq \delta$.

Now to bring it all together, we shall reformulate δ-Approximate Max Divergence in terms of Max Divergence and statistical distance.

Lemma 5.4. ***Necessary and Sufficient Conditions*** δ***-Approximate Max Divergence***

1. $D_\infty^\delta(Y\|Z) \leq \varepsilon \iff \exists$ *a random variable* X *such that* $\Delta(Y, X) \leq \delta$ *and* $D_\infty(X\|Z) \leq \varepsilon$*;*
2. $D_\infty^\delta(Y\|Z) \leq \varepsilon$ *and* $D_\infty^\delta(Z\|Y) \leq \varepsilon \iff$ *there exist random variables* Y', Z' *such that* $\Delta(Y, Y') \leq \frac{\delta}{e^\varepsilon+1}$, $\Delta(Z, Z') \leq \frac{\delta}{e^\varepsilon+1}$, *and* $D_\infty(Y'\|Z') \leq \varepsilon$.

Proof.
For 1., the if part Consider the random variables Y, Z and suppose that there exists a random variable Y' such that $\Delta(Y, Y') \leq \delta$ and $D_\infty(Y'\|Z)\varepsilon$, then $\forall$ set S,

$$\Pr[Y \in S] \leq \Pr[Y' \in S] + \delta \leq e^\varepsilon \cdot \Pr[Z \in S] + \delta \implies D_\infty^\delta(Y\|Z) \leq \varepsilon.$$

For the only if part Suppose that $D_\infty^\delta(Y\|Z) \leq \varepsilon$. Define the set $S := \{y : \Pr[Y = y] \geq e^\varepsilon \cdot \Pr[Z = y]\}$., then we have

$$\sum_{y \in S} (\Pr[Y = y] - e^\varepsilon \cdot \Pr[Z = y]) = \Pr[Y \in S] - e^\varepsilon \cdot \Pr[Z \in S] \leq \delta.$$

Also for $T = \{y : \Pr[Y = y] < \Pr[Z = y]\}$, then we have

$$\begin{aligned}
\sum_{y \in T} (\Pr[Z = y] - \Pr[Y = y]) &= \sum_{y \notin T} (\Pr[Y = y] - \Pr[Z = y]) \\
&\geq \sum_{y \in S} (\Pr[Y = y] - \Pr[Z = y]) \\
&\geq \sum_{y \in S} (\Pr[Y = y] - e^\varepsilon \cdot \Pr[Z = y).
\end{aligned}$$

Now we define Y' (with a lower probability on S and hence higher probability on T w.r.t. Y) as follows,

1. $\forall\, Y \in S, \Pr[Y' = y] = e^\varepsilon \cdot \Pr[Z = y] < \Pr[Y = y]$;
2. $\forall\, y \in T, \Pr[Y = y] \leq \Pr[Y' = y] \leq \Pr[Z = y]$;
3. $\forall\, y \notin S \cup T, \Pr[Y' = y] = \Pr[Y = y] \leq e^\varepsilon \cdot \Pr[Z = y]$.

Then we can see that

$$D_\infty(Y'\|Z) = \max_{S \subseteq \text{Supp}(Y')} \left(\ln \frac{\Pr[Y' \in S]}{\Pr[Z \in S]} \right)$$

$$\leq \max_{S \subseteq \text{Supp}(Y')} \left(\ln \frac{e^{\varepsilon} \Pr[Z \in S]}{\Pr[Z \in S]} \right)$$
$$\leq \varepsilon$$

And $\Delta(Y, Y') = \Pr[Y \in S] - \Pr[Y' \in S] = \Pr[Y \in S] - e^{\varepsilon} \cdot \Pr[Z \in S] \leq \delta$.

For 2., the if part Then $\forall$ set S,

$$\begin{aligned}
\Pr[Y \in S] &\leq \Pr[Y' \in S] + \frac{\delta}{e^{\varepsilon} + 1} \\
&\leq e^{\varepsilon} \cdot \Pr[Z' \in S] + \frac{\delta}{e^{\varepsilon} + 1} \\
&\leq e^{\varepsilon} \cdot \left(\Pr[Z \in S] + \frac{\delta}{e^{\varepsilon} + 1} \right) + \frac{\delta}{e^{\varepsilon} + 1} \\
&= e^{\varepsilon} \cdot \Pr[Z \in S] + \delta \\
&\implies D_{\infty}^{\delta}(Y \| Z) \leq \varepsilon.
\end{aligned}$$

By symmetry, $D_{\infty}^{\delta}(Z \| Y) \leq \varepsilon$.

The only if part Define $S := \{y : \Pr[Y = y] \geq e^{\varepsilon} \cdot \Pr[Z = y]\}$. Then we shall take, $\forall\, y \in S$,

$$\begin{aligned}
\Pr[Y' = y] &= e^{\varepsilon} \cdot \Pr[Z' = y] \\
&= \frac{e^{\varepsilon}}{1 + e^{\varepsilon}} \cdot (\Pr[Y = y] + \Pr[Z' = y]) \\
&\in [e^{\varepsilon} \cdot \Pr[Z = y], \Pr[Y = y]].
\end{aligned}$$

Which in turn implies that for $y \in S$,

$$\Pr[Y = y] - \Pr[Y' = y] = \Pr[Z' = y] - \Pr[Z = y] \frac{\Pr[Y = y] - e^{\varepsilon} \cdot \Pr[Z = y]}{e^{\varepsilon} + 1}$$

So

$$\begin{aligned}
\alpha := \sum_{y \in S} (\Pr[Y = y] - \Pr[Y' = y]) = \sum_{y \in S} (\Pr[Z' = y] - \Pr[Z = y]) \\
&= \frac{\Pr[Y \in S] - e^{\varepsilon} \cdot \Pr[Z \in S]}{e^{\varepsilon} + 1} \\
&\leq \frac{\delta}{e^{\varepsilon} + 1}.
\end{aligned}$$

Similarly we can reduce the probability mass of Z and increase that of Y by some $\alpha' \leq \frac{\delta}{e^{\varepsilon}+1}$ to get Z' and Y' respectively, so that $\forall\, y \in S' := \{y : \Pr[Z = y] > e^{\varepsilon} \cdot \Pr[Y = y]\}$, $\Pr[Z' = y] = e^{\varepsilon} \Pr[Y' = y]$.

If $\alpha = \alpha'$, then we can take, $\forall\, y \in S \cup S'$, $\Pr[Z' = y = \Pr[Z = y]$ and $\Pr[Y' = y = \Pr[Y = y] \implies D_{\infty}(Y \| Z) \leq \varepsilon$ and $\Delta(Y, Y') = \Delta(Z, Z') = \alpha$.

If $\alpha \neq \alpha'$, then if WLOG $\alpha ? \alpha'$, we shall need to reduce the probability mass of Y' and increase that of Z' for any $y \notin S \cup S'$ by $\beta = \alpha - \alpha'$ so that the probabilities end up summing to 1, i.e. $\sum_y \Pr[Y' = y] = 1 - \beta$ and $\sum_y \Pr[Z' = y] = 1 + \beta$. (In addition to this, we already have that $\Pr[Y' = y] \leq e^{\varepsilon} \cdot \Pr[Z' = y]$ and $\Pr[Z' = y] \leq e^{\varepsilon} \cdot \Pr[Y' = y]$.)

So if we define $R := \{y : \Pr[Y' = y] < \Pr[Z' = y]\}$, then $\sum_{y \in R} (\Pr[Z' = y] - \Pr[Y' = y]) \geq \sum_y (\Pr[Z' = y] - \Pr[Y' = y]) = 2\beta$.

So upon increasing (respectively, decreasing) the probability mass of Y' (respectively, Z') for points in R by a total of β, we still have that $\forall\, y \in R, \Pr[Y' = y] \leq \Pr[Z' = y]$. Then we have that for these new Y' and Z', $D_\infty(Y', Z') \leq \varepsilon$ and $\Delta(Y, Y'), \Delta(Z, Z') \leq \alpha$. ■

Lemma 5.5. *Given that random variables Y and Z satisfy $D_\infty(Y\|Z) \leq \varepsilon$ and $D_\infty(Z\|Y) \leq \varepsilon$, then $D(Y\|Z) \leq \varepsilon(e^\varepsilon - 1)$.*

Proof. It is known that $\forall$ random variables Y, Z, their KL-divergence, $D(Y\|Z) \geq 0$.

$$
\begin{aligned}
D(Y\|Z) &\leq D(Y\|Z) + D(Z\|Y) \\
&= \sum_y \Pr[Y = y] \cdot \left(\ln \frac{\Pr[Y = y]}{\Pr[Z = y]}\right) + \Pr[Z = y] \left(\ln \frac{\Pr[Z = y]}{\Pr[Y = y]}\right) \\
&= \sum_y \Pr[Y = y] \cdot \left(\ln \frac{\Pr[Y = y]}{\Pr[Z = y]} + \ln \frac{\Pr[Z = y]}{\Pr[Y = y]}\right) + (\Pr[Z = y] - \Pr[Y = y]) \cdot \left(\ln \frac{\Pr[Z = y]}{\Pr[Y = y]}\right) \\
&\leq \sum_y (\Pr[Y = y] \cdot (\ln 1) + |\Pr[Z = y] - \Pr[Y = y]| \cdot \varepsilon) \\
&= \sum_y (|\Pr[Z = y] - \Pr[Y = y]| \cdot \varepsilon) \\
&= \varepsilon \cdot \sum_y (\max\{\Pr[Y = y], \Pr[Z = y]\} - \min\{\Pr[Y = y], \Pr[Z = y]\} \\
&\leq \varepsilon \cdot \sum_y ((e^\varepsilon - 1) \cdot \min\{\Pr[Y = y], \Pr[Z = y]\}) \\
&\leq \varepsilon \cdot (e^\varepsilon - 1)
\end{aligned}
$$

■

Now we consider the situation where the adversary can adaptively act on the databases being fed to mechanisms, and adaptively affect queries to said mechanisms.

Let $\mathcal{F}$ be a family of database access mechanisms, and A be a probabilistic, stateful adversary. Consider the following experiment, called *k-fold adaptive composition*, in this setup.

Definition 5.6. k-Fold Adaptive Composition
For $b \in \{0, 1\}$, family $\mathcal{F}$, and adversary A, for each $i \in [k]$, A produces two neighbouring databases x_i^0 and x_i^1, a mechanism $\mathcal{M}_i \in \mathcal{F}$, and parameters w_i, and is returned a randomly chosen $y \in \mathcal{M}_i(w_i, x_i^b)$.
The choice of b, once made is kept constant throughout the experiment (ergo giving us two different variants of the experiment).

The adversary A here, being stateful, can adaptively choose the databases and parameters to input, and the mechanisms to use adaptively based upon previously observed outputs of previous mechanisms, and can only view their own coin tosses and the mechanism outputs produced as a result of this process.

Let A choose x_i^0 to always be a database containing a data subject, B's data, and x_i^1 to always be x_i^0 with B's data omitted. For $b = 0$, we are considering a real world situation wherein B's data is contained in the database and is allowed to be a part of various query releases, and for $b = 1$, we are in an ideal situation where the query releases (and their

accuracy) are independent of B's data. Now differential privacy promises indistinguishability of releases in both these cases.

Definition 5.7. **ε-Differential Privacy under k-Fold Adaptive Composition**
A family $\mathcal{F}$ of database access mechanisms satisfies *ε-differential privacy under k-fold adaptive composition* if $\forall$ adversary A, $D_\infty(V^0 \| V^1) \leq \varepsilon$, where V^b is the view of A for when the input databases chosen are of the form x_i^b.
(ε, δ)-differential privacy under k-fold adaptive composition is defined as when given the above defined $\mathcal{F}$ and arbitrarily chosen adversary A, $D_\infty^\delta(V^0 \| V^1) \leq \varepsilon$.

Now we are in a position to state and prove the Advanced Composition Theorem for ε-differentially private mechanisms[17], which is a celebrated result due to Dwork, Rothblum, and Vadhan[16].

Theorem 5.8. ***Advanced Composition for Pure Differentially Private Mechanisms***
$\forall \varepsilon > 0, \delta > 0$, and $k \in \mathbb{N}$, the class of ε-differentially private mechanisms is (ε', δ) differentially private under k-fold adaptive composition, where

$$\varepsilon' = \sqrt{2k \ln(1/\delta)} \cdot \varepsilon + k\varepsilon(e^\varepsilon - 1).$$

Proof. The view of the adversary A can be represented as a tuple $v := (r, y_1, \ldots, y_k)$, where r is the coin tosses of A and $y_1, \ldots, y_k$ are the outputs of the mechanisms $\mathcal{M}_1, \ldots, \mathcal{M}_k$. Let

$$B = \{v : \Pr[V^0 = v] > e^{\varepsilon'} \cdot \Pr[V^1 = v].$$

It shall suffice to show that $\Pr[V^0 \in B] \leq \delta$, as for every set S, $\Pr[V^0 \in S] \leq \Pr[V^0 \in B] + \Pr[V^0 \in (S \setminus B)] \leq \delta + \varepsilon^{\varepsilon'} \cdot \Pr[V^1 \in S]$, i.e. $D_\infty^\delta(V^0 \| V^1) \leq \varepsilon'$.

Let $V^0 = (R^0, Y_1^0, \ldots, Y_k^0)$ be the random variable denoting the view of A when $b = 0$, and let $V^1 = (R^1, Y_1^1, \ldots, Y_k^1)$ that when $b = 1$, then for a particular view $v = (r, y_1, \ldots, y_k)$,

$$\begin{aligned}
\ln\left(\frac{\Pr[V^0 = v]}{\Pr[V^1 = v]}\right) &= \ln\left(\frac{\Pr[R^0 = r]}{\Pr[R^1 = r]} \cdot \prod_{i=1}^{k} \frac{\Pr[Y_i^0 = y_i | R^0 = r, Y_1^0 = y_1, \ldots, Y_{i-1}^0 = y_{i-1}]]}{\Pr[Y_i^1 = y_i | R^1 = r, Y_1^1 = y_1, \ldots, Y_{i-1}^1 = y_{i-1}]}\right) \\
&= \sum_{i=1}^{k} \ln\left(\frac{\Pr[Y_i^0 = y_i | R^0 = r, Y_1^0 = y_1, \ldots, Y_{i-1}^0 = y_{i-1}}{\Pr[Y_i^1 = y_i | R^1 = r, Y_1^1 = y_1, \ldots, Y_{i-1}^1 = y_{i-1}]}\right) \\
&\stackrel{\text{def}}{=} \sum_{i=1}^{k} c_i(r, y_1, \ldots, y_i).
\end{aligned}$$

Now given a fixed prefix $(r, y_1, \ldots, y_{i-1})$, we would like to examine the random variables $C_i := c_i(R^0, Y_1^0, \ldots, Y_{i-1}^0, Y_i^0) = c_i(r, y_i, \ldots, y_{i-1}, y_i)$ given that $R^0 = r, Y_1^0 = y_1, \ldots, Y_{i-1}^0 = y_{i-1}$, and given the aforementioned fixed prefix, A produces the databases x_i^0, x_i^1, the mechanism $\mathcal{M}_i$, and the parameter w_i for $b = 0, 1$, and therefore Y_i^0 is distributed according to $\mathcal{M}_i(w_i, x_i^0)$. So for any y_i we have

$$|c_i(r, y_1, \ldots, y_i)| = \left|\ln\left(\frac{\Pr[\mathcal{M}_i(w_i, x_i^0) = y_i]}{\Pr[\mathcal{M}_i(w_i, x_i^1) = y_i]}\right)\right| \leq \varepsilon.$$

$$\begin{aligned}
&\implies \max\{D_\infty(\mathcal{M}_i(w_i, x_i^0) \| \mathcal{M}_i(w_i, w_i^1)), D_\infty(CM_i(w_i, x_i^1) \| \mathcal{M}_i(w_i, w_i^0)))\} = \varepsilon \\
&\implies D_\infty(CM_i(w_i, x_i^0) \| \mathcal{M}_i(w_i, w_i^1)), D_\infty(CM_i(w_i, x_i^1) \| \mathcal{M}_i(w_i, w_i^0)) \leq \varepsilon
\end{aligned}$$

And using lemma 5.5, we can see that

$$\mathbb{E}[c_i(R^0, Y_1^0, \ldots, Y_i^0)|R^0 = r, Y_1^0 = y_1, \ldots, Y_{i-1}^0 = y_{i-1}] = D(\mathcal{M}_i(w_i, x_i^0)\|\mathcal{M}_i(w_i, x_i^1)) \leq \varepsilon(e^\varepsilon - 1).$$

Then taking the random variables C_i, we can apply Azuma's inequality (theorem 2.5) with $\alpha := \varepsilon$, $\beta := \varepsilon \cdot \varepsilon_0$, and $z = \sqrt{2\ln(1/\delta)}$ to get

$$\Pr[V^0 \in B] = \Pr\left[\sum_i C_i > \varepsilon'\right] < e^{-z^2/2} = \delta.$$

This completes our proof. ■

The variant of this for approximate DP is merely stated below.

Theorem 5.9. ***Advanced Composition for Approximate Differentially Private Mechanisms***
$\forall\, \varepsilon > 0, \delta, \delta' > 0$, and $k \in \mathbb{N}$, the class of (ε, δ)-differentially private mechanisms is $(\varepsilon', k\delta + \delta')$ differentially private under k-fold adaptive composition, where

$$\varepsilon' = \sqrt{2k\ln(1/\delta')} \cdot \varepsilon + k\varepsilon(e^\varepsilon - 1).$$

6 Some More Useful Techniques

6.1 The Sparse Vector Technique (SVT)

On a tight privacy budget, one needs to prioritise what information they want to ask for or release, depending on which end of the aisle they are, with minimum noise addition. Often there are a large multitude of queries to respond to, which can exhaust the privacy budget quickly, or in other words, lead to massive amounts of privacy loss in trying to answer them with reasonable amounts of accuracy.[4]

That being said, in certain cases, all analysts care about are values above a certain threshold, and only those need be released with along with added noise whilst discarding values that lie significantly below said threshold, in a form of reasoning similar to, but more general than, say, Report Noisy Max.

Such releases of data would ensure that privacy would only degrade with respect to the number of queries that are actually released as a result of lying above that aforementioned threshold, and this can lead to significant savings if the released answer vector seems to be sparse with respect to the answer vector consisting of (numeric) responses to all the queries made.

In our upcoming analysis of sparse vector techniques, we shall consider the queries made as having sensitivity 1 and possibly being made adaptively. A common threshold can be fixed universally, though queries can have distinct, query specific thresholds with the same results as for the analysis of the common threshold situation.

[4]The bulk of this discussion is based on the sections of [7] pertaining to the sparse vector technique.

Consider the following algorithm which generates a bit vector denoting whether a query's response value is above the specified threshold or not, post noise addition.

Algorithm 5 AboveThreshold

Require: $x \in \mathbb{N}^{|\chi|}$, a sequence of adaptively chosen sensitivity 1 queries $\{q_i\}$ on $\mathbb{N}^{|\chi|}$, a threshold T and $\epsilon > 0$

$\hat{T} \leftarrow T + \text{Lap}\left(\frac{2}{\varepsilon}\right)$.
for each i **do**
  $v_i \leftarrow \text{Lap}\left(\frac{4}{\varepsilon}\right)$
  if $q_i(x) + v_i \geq \hat{T}$ **then**
    $a_i \leftarrow \top$
  else
    $a_i \leftarrow \perp$
  end if
end for
Output $a := \{a_i\}$

Theorem 6.1. *AboveThreshold is ε-differentially private.*

Proof. For any two neighbouring databases, $x, x' \in \mathbb{N}^{|\chi|}$, define A, A' as being the random variables representing the outputs of AboveThreshold$(x, \{q_i\}, T, \varepsilon)$ and AboveThreshold$(x', \{q_i\}, T, \varepsilon)$, and each bit of said outputs is from the set $\{\top, \perp\}$. Also $\hat{T}$ and each v_i in the algorithm are random variables. Now what we shall do is examine the case when the algorithm terminates upon receiving one above response above the threshold, then we can generalise to $c > 1$ above threshold query responses by using known composition theorems. For our analysis, we shall do a few things.

1. Let for some $k > 1, a := \{a_i\}_{i \in [k]}, a_i = \perp$ whenever $i \in [k-1]$ and $a_k = \top$;
2. Fix the arbitrary values of $v_1, \ldots, v_{k-1}$, so all that we need to do now is to consider the probability distributions/randomness of v_k and $\hat{T}$.

Define the maximum noisy value of any query response on x for among the queries $\{q_i\}_{i \in [k-1]}$ as

$$g(x) := \max_{i \in [k-1]} (q_i(x) + v_i).$$

By way of notational abuse, we shall write $\Pr[\hat{T} = t]$ as shorthand for the pdf of T evaluated at t, and $\Pr[v_k = \nu]$ as the pdf of v_k evaluated at ν. We denote the indicator function of event ζ by $1[\zeta]$. As v_i, for $i \in [k-1]$ are fixed, $g(x)$ is a deterministic quantity. So

$$\begin{aligned}
\Pr_{\hat{T}, v_k} &= \Pr_{\hat{T}, v_k} [\hat{T} > g(x) \text{ and } v_k \geq \hat{T}] \\
&= \Pr_{\hat{T}, v_k} [\hat{T} \in (g(x), q_k(x) + v_k]] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pr[v_k = v] \cdot \Pr[\hat{T}] 1[t \in (g(x), f_k(x) + v_k]] dv dt \\
&=: I
\end{aligned}$$

Define $\hat{v} := v + g(x) - g(x') + q_k(x') - q_k(x)$ and $\hat{t} := t + g(x) - g(x')$.
Then $\forall\, x, x'$, we have $|\hat{v} - v| \leq 2$ and $|\hat{t} - t| \leq 1$, as the sensitivity of each query is 1, and

therefore the sensitivity of $g(x)$ is 1 as well. Changing variables v, t to v', t' respectively,

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Pr[v_k = \hat{v}] \cdot \Pr[\hat{T} = \hat{t}] 1[(t + g(x) - g(x')) \in (g(x), q_k(x') + v + g(x) - g(x'))] dv dt \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Pr[v_k = \hat{v}] \cdot \Pr[\hat{T} = \hat{t}] 1[t \in (g(x), q_k(x') + v]] dv dt \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(\varepsilon/2) \Pr[v_k = v] \cdot \exp(\varepsilon/2) \Pr[\hat{T} = t] 1[(t \in (g(x'), q_k(x') + v]] dv dt \\
&\text{(due to the bounds on } |\hat{v} - v|, |\hat{t} - t| \text{ and the nature of the Laplace pdf.)} \\
&= \exp(\varepsilon) \Pr_{\hat{T}, v_k}[\hat{T} > g(x') \text{ and } q_k(x') + v_k \geq \hat{T}] \\
&= \exp(\varepsilon) \Pr_{\hat{T}, v_k}[A' = a].
\end{aligned}
$$

■

Definition 6.2. (α, β)**-Accuracy**
An algorithm that outputs a stream of answer bits $a_1, \ldots \in \{\top, \bot\}^*$ in response to a sequence of k many queries $\{q_i\}_{i \in [k]}$ is said to be (α, β)-*accurate* with respect to a threshold T if the algorithm only halts before the query q_k with probability $\leq \beta$, and on a database x and for all $a_i = \top$,

$$q_i(x) \geq T - \alpha$$

and for all $a_i = \bot$,

$$q_i(x) \leq T + \alpha.$$

Now the problem with algorithm 5 (AboveThreshold) is that the noisy threshold $\hat{T}$ may be too far removed from T, i.e. $|\hat{T} - T| > \alpha$, and also that a small count $q_i(x)$ may end up appearing above the threshold as a result of too much noise being added to it, even if $\hat{T}$ happens to be close to T. However, these events are very unlikely and happen with probability that are exponentially small in α. Now as we can have either error, or even both of them occur, we shall allocate the parameter $\alpha/2$ in the error analysis presented below.

Theorem 6.3. *For any sequence of k many queries $\{q_i\}_{i \in [k]}$, where all the queries except q_k are such that $q_i(x) < T - \alpha$ (i.e. firmly below the threshold), algorithm 5 (AboveThreshold) is (α, β)-accurate where*

$$\alpha = \frac{8\left(\ln\left(\frac{2}{\beta}\right) + \ln k\right)}{\varepsilon}.$$

Proof. It is sufficient to show that with probability β, the sum of the maximum noise added to any count with the difference between the actual and the noised threshold is at most α, i.e.,

$$\max_{i \in [k]} |v_i| + |\hat{T} - T| \leq \alpha.$$

Because if so, then for any $a_i = \top$, $q_i(x) + v_i \geq \hat{T} \geq T - |\hat{T} - T| \implies q_i(x) \geq T - |\hat{T} - T| - |v_i| \geq T - \alpha$.

Now for $a_i = \bot$, similarly, we get that $q_i(x) < \hat{T} \leq T + |\hat{T} - T| + |v_i| \leq T + \alpha$.

Also note that, as stated earlier, the count of any query for $i < k$ does not exceed the threshold, up to noise addition of α on either end of said threshold, i.e.,

$$\begin{aligned} & q_i(x) < T - \alpha < T - |v_i| - |\hat{T} - T| \\ \implies & q_i(x) + v_i \leq q_i(x) + |v_i| \leq T - |\hat{T} - T| \leq \hat{T} \\ \implies & a_i = \perp \end{aligned}$$

$\therefore$ That condition ensures that the algorithm does not halt before the query q_k is made.

Having justified the rationale behind our current objective, we shall proceed to prove the theorem in full.

For a Laplace random variable $Y \sim \text{Lap}(b)$, we know that $\Pr[|Y| \geq t \cdot b] = \exp(-t)$, and as the noise added in the context of AboveThreshold is taken from the Laplace distribution, $|\hat{T} - T|$ is a Laplace random variable.

$$\Pr\left[|\hat{T} - T| \geq \frac{\alpha}{2}\right] = \exp\left(-\frac{\varepsilon\alpha}{4})\right)$$

So to have this quantity bounded above by $\frac{\beta}{2}$, i.e.

$$\begin{aligned} & \exp\left(-\frac{\varepsilon\alpha}{4}\right) \leq \frac{\beta}{2} \\ \implies & -\frac{\varepsilon\alpha}{4} \leq \ln\frac{\beta}{2} \\ \implies & \alpha \leq \frac{4\ln\frac{2}{\beta}}{\varepsilon}. \end{aligned}$$

Similarly, and by applying a union bound, we get

$$\Pr[\max_{i\in[k]} |v_i| \geq \frac{\alpha}{2}] \leq k \cdot \exp\left(-\frac{\varepsilon\alpha}{8}\right)$$

And, similarly, bounding this above by $\frac{\beta}{2}$ requires us to have $\alpha \geq \frac{8\ln\left(\frac{2}{\beta}\right)+\ln k}{\varepsilon}$. ■

Now we simply have described results for a single AboveThreshold release. To generalise this, we have to look at composing multiple AboveThreshold queries. The Spare algorithm was devised to handle that: it simply calls AboveThreshold (with ε-DP guarantees) repeatedly, with it reinitialising its calls to AboveThreshold after receiving an above threshold response on the remainder of the sequence of queries. After it calls AboveThreshold a set number $(c > 0)$ of times, it halts. We shall see shortly that by an application of the advanced composition theorem, Sparse gives us (ε, δ)-differential privacy.

Theorem 6.4. *Sparse is (ε, δ)-differentially private.*

Proof. We shall use the notation used in the pseudocode for Sparse (algorithm 6). We know that AboveThreshold$(x, \{q_i\}, T, \varepsilon)$ gives us ε-differential privacy, and in the algorithm, we have assigned

$$\varepsilon' = \begin{cases} \frac{\varepsilon}{c}, \text{if } \delta = 0; \\ \frac{\varepsilon}{\sqrt{8c\ln\frac{1}{\delta}}} \text{ otherwise.} \end{cases}$$

Algorithm 6 Sparse

Require: $x \in \mathbb{N}^{|\chi|}$, a sequence of adaptively chosen sensitivity 1 queries $\{q_i\}$ on $\mathbb{N}^{|\chi|}$, a threshold T, a cutoff $c \in \mathbb{N}$, $\delta \geq 0$, and $\epsilon > 0$

if $\delta = 0$ **then**
  $\sigma \leftarrow \frac{2c}{\varepsilon}$
else
  $\sigma \leftarrow \frac{\sqrt{32c \ln \frac{1}{\delta}}}{\varepsilon}$
end if
$\hat{T} \leftarrow T + \text{Lap}(\sigma)$.
$j \leftarrow 0$
for each i **do**
  $v_i \leftarrow \text{Lap}(2\sigma)$
  if $q_i(x) + v_i \geq \hat{T}_j$ **then**
    $a_i \leftarrow \top$
    $j \leftarrow j + 1$
    $\hat{T}_j = T + \text{Lap}(\sigma)$
  else
    $a_i \leftarrow \perp$
  end if
  if $j \geq c$ **then**
    Halt.
  end if
end for
Output $a := \{a_i\}$

So we take repeated instances of AboveThreshold$(x, \{q_i\}, T, \varepsilon')$ (at most c times) and when $\delta = 0$, the result that c instances of a $\varepsilon' = \frac{\varepsilon}{c}$-differentially private algorithms are ε-differentially private simply follows from basic composition.

For when $\delta > 0$, we can apply the advanced composition theorem to get that c applications of an $\varepsilon' = \frac{\varepsilon}{\sqrt{8c \ln \frac{1}{\delta}}}$-differentially private algorithm is (ε, δ)-differentially private. ■

Now we shall discuss the accuracy of Sparse.

Theorem 6.5. *For any sequence of k queries $\{q_i\}_{i \in [k]}$ such that $L(T) := |\{i : q_i(x) \geq T - \alpha\}| \leq c$, if $\delta > 0$, then Sparse is (α, β)-accurate, where*

$$\alpha = \frac{\left(\ln k + \ln \frac{2c}{\beta}\right) \sqrt{512c \ln \frac{1}{\delta}}}{\varepsilon},$$

and if $\delta = 0$, Sparse is (α, β)-accurate, where

$$\alpha = \frac{8c\left(\ln\left(\frac{2c}{\beta}\right) + \ln k\right)}{\varepsilon}.$$

Proof. This can be treated as a corollary/application of the accuracy result for AboveThreshold,

by taking the β parameter to be $\frac{\beta}{c}$ and the ε parameter to be

$$\begin{cases} \frac{\varepsilon}{\sqrt{8c \ln \frac{1}{\delta}}} \text{ if } \delta > 0 \\ \frac{\varepsilon}{c} \text{ if } \delta = 0 \end{cases}.$$

The result follows. ■

So Sparse answers queries with noise that scales as $\Theta(\log k)$, which is nice compared to the noise for straightforward application of ε-DP ($\Theta(k)$) and (ε, δ)-DP ($\Theta(\sqrt{k \ln(\frac{1}{\delta})})$).

We can now introduce NumericSparse, a version of Sparse that does not only produce a bit vector in response to a sequence of queries, but also releases the numeric values of the query responses that are above the threshold; it combines Sparse with the Laplace mechanism and outputs a vector $a \in (\mathbb{R} \cup \{\perp\})^*$.

Algorithm 7 NumericSparse

Require: $x \in \mathbb{N}^{|\chi|}$, a sequence of adaptively chosen sensitivity 1 queries $\{q_i\}$ on $\mathbb{N}^{|\chi|}$, a threshold T, a cutoff $c \in \mathbb{N}$, $\delta \geq 0$, and $\varepsilon > 0$

if $\delta = 0$ **then**
  $\varepsilon_1 \leftarrow \frac{8}{9}\varepsilon, \varepsilon_2 \leftarrow 29\varepsilon$
  $\sigma(\varepsilon) = \frac{2c}{\varepsilon}$
else
  $\varepsilon_1 = \frac{\sqrt{512}}{\sqrt{512}+1}\varepsilon, \varepsilon_2 = \frac{2}{\sqrt{512}+1}$
  $\sigma \leftarrow \frac{\sqrt{32c \ln \frac{1}{\delta}}}{\varepsilon}$
end if
$\hat{T} \leftarrow T + \text{Lap}\,(\sigma(\varepsilon_1))$.
$j \leftarrow 0$
for each i **do**
  $v_i \leftarrow \text{Lap}\,(2\sigma(\varepsilon_1))$
  if $q_i(x) + v_i \geq \hat{T}_j$ **then**
    $v_i \leftarrow \text{Lap}(\sigma(\varepsilon_2))$
    $a_i \leftarrow q_i(x) + v_i$
    $j \leftarrow j + 1$
    $\hat{T}_j = T + \text{Lap}(\sigma(\varepsilon_1)$
  else
    $a_i \leftarrow \perp$
  end if
  if $j \geq c$ **then**
    Halt.
  end if
end for
Output $a := \{a_i\}$

Theorem 6.6. *NumericSparse is (ε, δ)-differentially private.*

Proof. If $\delta = 0$, NumericSparse$(x, \{q_i\}, T, c, \varepsilon, \delta)$ is just the adaptive composition of Sparse$(x, \{q_i\}, T, c, \frac{8}{9}\varepsilon, 0)$ with the Laplace mechanism on the privacy parameters $(\varepsilon', \delta) = \left(\frac{1}{9}\varepsilon, 0\right)$. Applying simple composition, we get that this composition yields (ε, δ)-differential privacy.

If $\delta > 0$, then NumericSparse$(x, \{q_i\}, T, c, \varepsilon, \delta)$ is the adaptive composition of Sparse$(x, \{q_i\}, T, c, \frac{\sqrt{512}}{\sqrt{512}+1}\varepsilon, \frac{\delta}{2})$ with the Laplace mechanism on the privacy parameters $(\varepsilon', \delta) = \left(\frac{1}{\sqrt{512}+1}\varepsilon, \frac{\delta}{2}\right)$. Again applying simple composition, we get that this composition yields (ε, δ)-differential privacy. ■

Definition 6.7. Numeric Accuracy
An algorithm that outputs a sequence/vector of answers $a_1, \ldots \in (\mathbb{R} \cup \{\perp\})^*$ in response to a sequence of k queries $\{q_i\}_{i\in[k]}$ is said to be (α, β)-*accurate* with respect to a threshold T if it only halts before the query q_k is made with probability $\leq \beta$, and $\forall\, a_i \in \mathbb{R}$

$$|q_i(x) - a_i| \leq \alpha$$

and for all $a_i = \perp$

$$q_i(x) \leq T + \alpha.$$

Theorem 6.8. *For any sequence of k queries, $\{q_i\}_{i\in[k]}$, and $L(T) := |\{i : q_i(x) \geq T - \alpha\}| \leq c$, if $\delta > 0$, then NumericSparse is (α, β)-accurate, where*

$$\alpha = \frac{\left(\ln k + \ln \frac{4c}{\beta}\right)\sqrt{c \ln \frac{2}{\delta}}(\sqrt{512}+1)}{\varepsilon},$$

and if $\delta = 0$, NumericSparse is (α, β)-accurate, where

$$\alpha = \frac{9c\left(\ln k + \ln\left(\frac{4c}{\beta}\right)\right)}{\varepsilon}.$$

Proof. For (α, β)-accuracy, we need that $\forall\, a_i = \perp, q_i(x) \leq T + \alpha$, which holds with probability $1 - \frac{\beta}{2}$ by the theorem about the accuracy of Sparse. For all $a_i \in \mathbb{R}, |q_i(x) - a_i| \leq \alpha$, which again holds with probability $\frac{\beta}{2}$ by the result on the bound on the accuracy of the Laplace mechanism that we discussed earlier. ■

Thus, we see that with accuracy comparable to that of the Laplace mechanism (up to difference in constants and a factor of $\ln k$), the sparse vector technique allows us to identify and publish the noised outputs of large queries, so tot speak, with only a logarithmic cost being paid for the queries that yield $\perp$ as a response.

6.2 SmallDB

6.2.1 Linear Queries

Now we shall discuss certain techniques to release query responses to a class of queries called linear queries, which are a generalisation of counting queries, but instead of only taking boolean values, it takes values in $[0, 1]$. More concretely, a linear query $\bar{q}$ is a query that takes $q : \chi :\to [0, 1]$ and returns either the sum or the mean of the query's responses with each $\chi_i \in \chi$ (i.e. each datatype) as input. With some commonly used notational abuse, $\bar{q}$ is often represented by q taken on a database (and hence with the domain $\mathbb{N}^{|\chi|}$, instead of a datatype, by overloading the definition of the query q itself.

When linear queries return the average of the aforementioned query responses, they are called *normalised linear queries*, and is defined as follows, (and thus, as the term "normalised" suggests, takes values in $[0, 1]$)

$$(q(x) =)\bar{q}(x) := \frac{1}{\|x\|_1} \sum_{i=1}^{|\chi|} x_i q(\chi_i).$$

And the linear queries that return the sum of the aforementioned query responses are called *unnormalised linear queries*, which are given by,

$$(q(x) =)\bar{q}(x) := \sum_{i=1}^{|\chi|} x_i q(\chi_i).$$

Which is simply the formula for normalised linear queries without the normalisation factor, $\frac{1}{\|x\|_1}$, and thus take values in $[0, \|x\|_1]$.

Here linear queries have sensitivity ≤ 1, (denoting q applied on the database x as q_x) as

$$\Delta(\bar{q}) = \max_{\|x-y\|_1 \leq 1} |\bar{q}(x) - \bar{q}(y)| \leq |\sum_{i=1}^{|\chi|} x_i q_x(\chi_i) - \sum_{i=1}^{|\chi|} y_i q_y(\chi_i)|.$$

Now x and y differ in only 1 record, let it be at the index i, i.e. $x_i = y_i$ except for when $i = k$, where $|x_k - y_k| \leq 1$. Then

$$\Delta(\bar{q}) \leq |\sum_{i=1}^{|\chi|} x_i q_x(\chi_i) - \sum_{i=1}^{|\chi|} y_i q_y(\chi_i)| \leq \sum_{i=1}^{|\chi|} x_i - \sum_{i=1}^{|\chi|} y_k| = |x_k - y_k| \leq 1.$$

6.2.2 Back to SmallDB

We saw that for Sparse, we answer k queries, releasing their responses (noised) if and only if they exceed a threshold, making the answer vector sparse and thus reducing the amount of privacy loss and ergo total noise added.

What SmallDB, which is an offline algorithm, promises[18] is essentially answering all of those k queries with noise in $\Theta(\log k)$, although with some caveats. As the name of the algorithm suggests, it outputs fairly small databases, that can be used to in a way so as to give a good approximation for a query's responses on a database with high probability. This works by correlating noise that is added across queries. Another salient feature is refusing to answer a query on the same database more than once to avoid an adversary trying to mount an averaging attack. This is similar to memoisation, something we shall cover in future sections. It uses the Exponential Mechanism to add noise and produce said small (synthetic) database.

SmallDB takes as input a database $x \in \mathbb{N}^{|\chi|}$, a collection of queries Q, $\varepsilon > 0$, and an accuracy bound α. It outputs a smaller database (w.r.t. x) y whose size depends on $\log(|Q|)$ and α. Noise addition is done via the use of the Exponential mechanism with the score/utility function being $u(x, y) = -\max_f \in Q|f(x) - f(y)|$.

Theorem 6.9. *SmallDB is ε-differentially private.*

Proof. The noise addition in this algorithm is done via the Exponential Mechanism with privacy parameter $\varepsilon > 0$, hence satisfies ε-differential privacy. ■

Algorithm 8 SmallDB

Require: $x \in \mathbb{N}^{|\chi|}$, a collection of queries Q, $\varepsilon > 0$, and accuracy bound α

1: $\mathcal{R} = \{y \in \mathbb{N}^{|\chi|} : \|y\|_1 = \frac{\log|Q|}{\alpha^2}\}$

2: Let $u : \mathbb{N}^{|\chi|} \times \mathcal{R} \to \mathbb{R}$ be the utility function, where

$$u(x, y) = -\max_f \in Q|f(x) - f(y)|$$

3: **Sample and Output** $y \in \mathcal{R}$ with the Exponential Mechanism $\mathcal{M}_E(x, u, \varepsilon)$.

Now to discuss the accuracy of SmallDB, we look at the following theorem.

Theorem 6.10. *For accuracy bound $\alpha > 0$, finite collection of linear queries Q, if $\mathcal{R} = \{y \in \mathbb{N}^{|\chi|} : \|y\|_1 = \frac{\log|Q|}{\alpha^2}\}$, then for all $x \in \mathbb{N}^{|\chi|}, \exists\, y \in \mathcal{R}$ such that*

$$\max_{q \in Q} |f(x) - f(y)| \leq \alpha.$$

Proof. Construct database y having the elements $(x_1, \ldots, x_m)$ by taking $m = \log|Q|/\alpha^2$ samples uniformly from the elements of x, with X_i being a random variable such that $P(X_i = \chi_j \in \chi) = \frac{x_j}{\|x\|_1}$. $\forall\, q \in Q$,

$$q(y) = \frac{1}{\|y\|_1} \sum_{i=1}^{|\chi|} y_i f(\chi_i) = \frac{1}{m} \sum_{i=1}^{m} f(X_i).$$

With the second equality basically is switching to look at value per entry in y. Then we have

$$\begin{aligned}
\mathbb{E}[q(y)] &= \mathbb{E}[\frac{1}{m} \sum_{i=1}^{m} q(X_i) \\
&= \frac{1}{m} \sum_{i=1}^{m} \mathbb{E}(q(X_i)) \\
&= \frac{1}{m} \sum_{i=1}^{m} \left(\sum_{i=1}^{|\chi|} \frac{x_j}{\|x\|_1} q(\chi_j) \right) \\
&= \frac{1}{m} \sum_{i=1}^{m} q(x) = q(x)
\end{aligned}$$

We take a probabilistic approach to showing the existence of such a good database y. Using the additive Chernoff bound we get that

$$\Pr[|q(y) - q(x)| > \alpha] \leq 2e^{-2m\alpha^2}.$$

And applying a union bound over the linear queries in Q, we get

$$\Pr[\max q \in Q|q(y) - q(x)| > \alpha] \leq 2|Q|e^{-2m\alpha^2}.$$

And as $m = \frac{\log|Q|}{\alpha^2}$,

$$\Pr[|q(y) - q(x)| > \alpha] < 1.$$

Which means that $\exists$ a $y \in \mathbb{N}^{|\chi|}$ of size $\frac{\log|Q|}{\alpha^2}$ which ensures that $\Pr[|q(y) - q(x)| > \alpha \Pr[|q(y) - q(x)| < \alpha$, and clearly $y \in \mathcal{R}$, as $\mathcal{R}$ contains all databases of size $\frac{\log|Q|}{\alpha^2}$. ■

Now all we need to do is show that we can sample such an appropriate, good database y with high probability.

Theorem 6.11. *Let Q be a finite class of linear queries. Let $y = \mathit{SmallDB}(x, Q, \varepsilon, \alpha)$. Then for some $\beta \in [0, 1]$,*

$$\Pr\left[\max_{q \in Q} |q(x) - q(y)| < \alpha + \frac{2\left(\frac{\log|\chi| \cdot \log|Q|}{\alpha^2} + \log\frac{1}{\beta}\right)}{\varepsilon\|x\|_1}\right] \geq 1 - \beta.$$

Proof. By the accuracy theorem for the Exponential Mechanism,

$$\Pr\left[u(\mathcal{M}_E(x, u, \varepsilon)) \leq OPT_u(x) - \frac{2\Delta u(\ln|\mathcal{R}| + t)}{\varepsilon}\right] \leq e^{-t}.$$

Clearly $|\mathcal{R}| = |\chi|^{\frac{\log|Q|}{\alpha^2}} \implies \ln|\mathcal{R}| = \frac{\log|Q| \times \log|\chi}{\alpha^2}$. Δu here is simply $\frac{1}{\|x\|_1}$. By setting $t = \log(\frac{1}{\beta})$, we get

$$\Pr\left[u(\mathcal{M}_E(x, u, \varepsilon)) \leq OPT_u(x) - \frac{2\Delta u(\frac{\log|Q| \times \log|\chi}{\alpha^2} + \log\frac{(}{1}\beta))}{\varepsilon}\right] \leq \beta$$

By the preceding existence theorem, we know that $\exists\, y \in \mathbb{N}^{|\chi|}$ such that $|q(x) - q(y)| = -u(x, y) < \alpha$, which means that $\alpha > -OPT_u(x)$, so

$$\Pr\left[\max_{q \in Q} |q(x) - q(y)| \geq \alpha + \frac{2\Delta u(\frac{\log|Q| \times \log|\chi}{\alpha^2} + \log\frac{(}{1}\beta))}{\varepsilon}\right] \leq \beta$$

$$\implies \Pr\left[\max_{q \in Q} |q(x) - q(y)| < \alpha + \frac{2\Delta u(\frac{\log|Q| \times \log|\chi}{\alpha^2} + \log\frac{(}{1}\beta))}{\varepsilon}\right] \geq 1 - \beta.$$

■

Theorem 6.12. *Let $y = \mathit{SmallDB}(x, Q, \varepsilon, \frac{\alpha}{2})$, then*

$$\Pr\left[\max_{q \in Q} |q(x) - q(y)| < \left(\frac{16\log|\chi|\log|Q| + 4\log(\frac{1}{\beta})}{\varepsilon\|x\|_1}\right)^{\frac{1}{3}}\right] \geq 1 - \beta.$$

Proof. By the preceding theorem, for $y = \text{SmallDB}(x, Q, eps, \frac{\alpha}{2})$,

$$\Pr\left[\max_{q \in Q} |q(x) - q(y)| \geq \frac{\alpha}{2} + \frac{2\left(\frac{4\log|\chi| \cdot \log|Q|}{\alpha^2} + \log(\frac{1}{\beta})\right)}{\varepsilon\|x\|_1}\right] < \beta.$$

Taking $\alpha = \frac{4\left(\frac{4\log|\chi| \cdot \log|Q|}{\alpha^2} + \log(\frac{1}{\beta})\right)}{\varepsilon\|x\|_1}$, we get

$$\Pr\left[\max_{q \in Q} |q(x) - q(y)| \geq \frac{2\left(\frac{4\log|\chi| \cdot \log|Q|}{\alpha^2} + \log(\frac{1}{\beta})\right)}{\varepsilon\|x\|_1} + \frac{2\left(\frac{4\log|\chi| \cdot \log|Q|}{\alpha^2} + \log(\frac{1}{\beta})\right)}{\varepsilon\|x\|_1}\right] < \beta$$

$$\implies \Pr\left[\max_{q \in Q} |q(x) - q(y)| \geq \left(\frac{16\log|\chi|\log|Q| + 4\log(\frac{1}{\beta})}{\varepsilon\|x\|_1}\right)^{\frac{1}{3}}\right] < \beta.$$

$$\implies \Pr\left[\max_{q\in Q}|q(x)-q(y)| < \left(\frac{16\log|\chi|\log|Q|+4\log(\frac{1}{\beta})}{\varepsilon\|x\|_1}\right)^{\frac{1}{3}}\right] \geq 1-\beta.$$

■

7 Local Differential Privacy

Earlier, we discussed what is called central differential privacy, i.e. a model of enforcing differential privacy with the aid of a trusted curator entrusted with the raw data of multiple data subjects, but problems arise with this approach when the existence of a trusted curator is not possible or is inconvenient. Add to that the fact that if the trusted curator is somehow compromised, it would be a major privacy catastrophe.

Consider the situation we used when discussing Warner's randomised response, in that situation, we had a property $\mathcal{P}$ that was potentially too highly incriminating (viz. having used a certain narcotic substance or cheated in an exam) to be disclosed in its raw form to anyone (other than the data subject themselves) at all, and so what randomised response does in general is to provide noise addition from the data subjects' end itself and no third party trusted curator is involved in this process. Recall that we have also defined a general form of randomised response that provides ε-differential privacy for a given $\varepsilon > 0$.

Enter local differential privacy, where we discard the notion of a trusted curator (a curator may still exist but is no longer trusted, if so) and transfer the power to add noise and thus make differentially private disclosures to the data subject themselves.

Formally, local differential privacy is a slightly stronger guarantee than, but in essence similar to, differential privacy.

Definition 7.1. Local Differential Privacy (LDP)
A randomised mechanism $\mathcal{M}$ for $\varepsilon > 0$ is said to be $\varepsilon-$locally differentially private if for all pairs x, y of a user's private data, and for all possible outputs $z \in \text{Range}(\mathcal{M})$,

$$\ln\left(\frac{\Pr[M(x)=z]}{\Pr[M(y)=z]}\right) \leq \varepsilon$$

with the probability space being over the coin flips of $\mathcal{M}$.

As it turns out, randomised response as we defined earlier does provide ε-local differential privacy, and happens to be one of the most prominent primitives used for endowing ε-local differential privacy upon data releases. Recall that in randomised response w.r.t. a fixed $\varepsilon > 0$ (let's call this mechanism/algorithm $\mathcal{M}$ again) we showed that a slightly stronger guarantee to differential privacy was satisfied, i.e. for any pair of bits X_1, X_2 of a user, and for $z \in \{0, 1\}$ we showed that

$$\ln\left(\frac{\Pr[\mathcal{M}(X_1)=z]}{\Pr[\mathcal{M}(X_2)=z]}\right) \leq \varepsilon,$$

which simply means that randomised response is locally differentially private!

7.1 LDP: A Silver Bullet?

Which begs the question: why bother with central differential privacy at all if local differential privacy seems to work?

The answer lies in the subtle differences between the two: local differential privacy is a stronger guarantee than "vanilla' differential privacy, and hence for the same $\varepsilon > 0$, the noise added (and thus the amount of accuracy lost) by the application of an ε-differentially private algorithm is more than that for an application of a central ε-differentially private algorithm.

Also it has been seen that local differential privacy gives decent or even good results for large databases, but for smaller databases, it suffers in terms of providing accuracy, which was ameliorated by the introduction of *condensed local differential privacy*[19], which we shall not discuss much in this paper.

Definition 7.2. Condensed Local Differential Privacy (CLDP)
A randomised mechanism $\mathcal{M}$ for $\varepsilon > 0$ is said to be $\alpha-$condensed locally differentially private if for all pairs x, y of a user's private data, and for all possible outputs $z \in$ Range($\mathcal{M}$),

$$\left(\frac{\Pr[M(x) = z]}{\Pr[M(y) = z]}\right) \leq e^{\alpha d(x,y)}$$

with the probability space being over the coin flips of $\mathcal{M}$, and d is a suitable metric (contextually chosen, viz. the Manhattan distance).

In essence what CLDP does is scale privacy degradation w.r.t. the distance between the pair of the user's data, and confers less privacy degradation for pairs of data with less distance between them. More variants of LDP, just like CLDP, have been proposed, and this is an active field of study. Also it is worth pointing out that there is currently no known separation between the local variants of ε- and (ε, δ)-differential privacy.

Now we can illustrate how the amount of noise differs between central differential privacy and local differential privacy.

7.2 Comparative Study on Noise Addition in Central vis-à-vis Local DP

Let us consider the problem of differentially private estimation of mean of bits $\{X_i\}_{i\in[n]}$. Using central differential privacy, this can be done using the Laplace mechanism as follows,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i + \text{Lap}(\frac{1}{\varepsilon n}).$$

Here it can be easily seen that the magnitude of noise, $\mu - \hat{\mu}| \leq \frac{1}{\varepsilon n}$. So to bound the error by α, we have

$$\mu - \hat{\mu}| \leq \frac{1}{\varepsilon n} \leq \alpha$$
$$\implies n \geq \frac{1}{\alpha\epsilon}.$$

Now for the above described randomised response, if $\varepsilon = O(1)$, then the variance of $\hat{\mu}$ is

$O\left(\frac{1}{\epsilon^2 n}\right)$.

By Chebyshev's inequality,

$$\Pr\left[|\hat{\mu} - \mu| \geq \frac{1}{\varepsilon\sqrt{n}}\right] \leq \frac{\text{Var}(\hat{\mu})}{\frac{1}{\varepsilon^2 n}}.$$

So a good choice would be to bound $|\hat{\mu} - \mu| \leq \frac{1}{\varepsilon\sqrt{n}} \implies$ if we want to bound the error (LHS of the above equation, i.e. $\Pr\left[|\hat{\mu} - \mu| \geq \frac{1}{\varepsilon\sqrt{n}}\right]$) by α, then $n \geq \frac{1}{\alpha^2\epsilon^2}$.

That happens to be the square of the number of bits required as for centrally differentially private Laplace mechanism based mean calculation for the same error bound!

Which is precisely why local differential privacy is mostly used in contexts where there is an ample abundance of data and by companies with large stores of data like Microsoft, Google, and Apple, among others.

7.3 Models of Local Differential Privacy

There are three main models of local differential privacy based on curator-user interactivity.

Non-Interactive: There is only a single round of privatised communication of all users with the curator; that is the curator issues a single call for data and each user sends privatised messaged to the curator.

Sequentially Interactive: Here the curator can sequentially (i.e. one user after another) and adaptively query users, each at most once.

Fully Interactive: Here there is no holds barred, free flow of queries and (privatised) responses between the curator and individuals.

As it happens, the implications of using and the differences between each of these models is still unclear and is a field of research.

7.4 Deployments of LDP

First, we define an important concept/technique that is used while designing a number of LDP deployments, called *memoisation*, which is used in response to the possibility of suffering *noise averaging attacks*, i.e. when an adversary makes the exact same query to a database via a differentially private mechanism in order to degrade its privacy guarantee by observing many perturbed responses of said query and then averaging it after a large number of times of making that query, at which point the noise added to the raw data should more or less cancel out, leaving the adversary with (a close approximation of) the raw data, which is a privacy violation.

Definition 7.3. Memoisation
Memoisation can informally be defined as sending the same noise perturbed response to the same query being made repeatedly.

Memoisation, by making sure the adversary gets the same response for each instance of the same query nullifies noise averaging attacks; but a memoisation is not without its own perils:

an adversary could very well make many equivalent queries which essentially ask for the same thing, which still leaves some room for possible noise averaging attacks anyway. But it still works for instances in which there is a system, which is genuinely focused on the users' privacy, that makes the same query regularly (say a system like a web browser's/operating system's feedback system).

Now we shall discuss some popular deployments of LDP.

7.4.1 RAPPOR (Google; Erlingsson et al, 2014)[20]

RAPPOR (Randomized Aggregatable Privacy-Preserving Ordinal Response) was developed by researchers at Google to collect locally differentially private reports from users of its apps, a primary example being Google Chrome. It makes use of bitwise randomised response and bloom filters to produce reports. We shall just discuss in fair detail how it produces locally differentially private reports from the clients' end, leaving the aggregation of said reports out to keep our discussion focused on differential privacy.

It is implemented as follows.

1. **Signal** Take string x from a known universe χ and pass through Bloom filter of size k with h hash functions to get a k bit string B;
2. **Permanent Randomised Response** Apply bitwise randomised response to B to obtain B' as follows, taking B_i and B'_i to be the i^{th} bit of B and B' respectively,

$$B'_i = \begin{cases} 1, \text{ with probability } \frac{1}{2}f \\ 0, \text{ with probability } \frac{1}{2}f \\ B_i, \text{ with probability } 1-f \end{cases}$$

   where f is a parameter that can be tuned by the user, and determines the degree of longitudinal privacy thus conferred;
3. **Instantaneous Randomised Response** B' can be possibly unique to a user, so apply bitwise randomised response again prior to sending it to the curator as follows; create a k bit array S with all 0s, and denoting the i^{th} bit of S by S_1, assign 1 to S_i with probability

$$\Pr[S_i = 1] = \begin{cases} p, \text{ if } B'_i = 1; \\ q, \text{ if } B'_i = 0 \end{cases}$$

   where $p, q \in [0, 1]$ are locally defined parameters;
4. Curator uses a sophisticated aggregation technique to get results.

RAPPOR can be modified to fit various contexts.Three prominent modifications are listed below.

One-Time RAPPOR In this case, the data is collected only once, so longitudinal attacks are not a concern here, and applying instantaneous randomised response can be skipped in favour of applying direct randomisation to provide ample privacy.

Basic RAPPOR If the known universe of strings, χ, has short and well defined strings that can be simply and deterministically mapped to a single bit on a bit array (e.g. for a single toss of a fair coin, we can map "heads" to 10 and tails to 01, i.e. the first and the second bits correspond to "heads" and "tails" respectively), then it does not make sense to use

a Bloom filter with multiple hash functions. Instead, in the signal step, one can simply replace the Bloom filter with the deterministic mapping of strings to bits in the bit array itself.

Basic One-Time RAPPOR This is simply applied in a melange of the above two situations, so we would simply need to (1.) replace the Bloom filter with the deterministic mapping described above, and (2.) skip applying instantaneous randomised response.

Theorem 7.4. ***Differential Privacy of Permanent Randomised Response of RAPPOR*** *Permanent Randomised Response satisfies ε-local differential privacy, where*

$$\varepsilon = 2h \ln \left(\frac{1 - \frac{1}{2}f}{\frac{1}{2}f} \right).$$

Proof. (Given in [20]) Given a RAPPOR generated report, $S = s = s_1, s_2, \ldots, s_n$, and knowing the true client value $V = v$, where S and V are random variables pertaining to the RAPPOR report received and true client value respectively, we have the following.

$$\begin{aligned} \Pr[S = s|V = v] &= \Pr[S = s|B', B, v] \cdot \Pr[B'|B, v] \Pr[B|v] \\ &= \Pr[S = s|B'] \cdot \Pr[B'|B] \cdot \Pr[B|v] \\ &= \Pr[S = s|B'] \cdot \Pr[B'|B]. \end{aligned}$$

Note that the probability of observing $S = s$ is independent of what the value of B is as long as B' is known, and hence $Pr[S = s|B']$ does not provide any information about B, and is not something that can be used to find/deduce about B in a longitudinal attack, but the second term $\Pr[B'|B]$ does, and is the focus for studying longitudinal privacy in this context. Consider the following probabilities.

$$\Pr[b'_i = 1|b_i = 1] = \frac{1}{2}f + (1 - f) = 1 - \frac{1}{2}f,$$

$$\Pr[b'_i = 1|b_i = 0] = \frac{1}{2}f.$$

Now WLOG, let the bits indexed $1 \leq i \leq h$ be set w.r.t. the hash functions, i.e. $b^* = \{b_1 = 1, b_2, \ldots, b_h = 1, b_{h+1} = 1, \ldots, b_n = 0\}$, then,

$$\Pr[B' = b'|B = b^*] = \prod_{i=1}^{n} \left(\frac{1}{2}f \right)^{b'_i} \left(1 - \frac{1}{2}f \right)^{1-b'_i}$$

For any two distinct values of B, B_1, B_2, we have, for any set R,

$$\begin{aligned} \frac{\Pr[B' \in R|B = B_1]}{\Pr[B' \in R|B = B_2]} &= \frac{\sum_{B'_i \in R} \Pr[B' = B'_i|B = B_1]}{\sum_{B'_i \in R} \Pr[B' = B'_i|B = B_2]} \\ &\leq \max_{B'_i \in R} \frac{\Pr[B' = B'_i|B = B_1]}{\Pr[B' = B'_i|B = B_2]} \\ &= \left(\frac{1}{2}f \right)^{2(\sum_{i=1}^{h} b_i - \sum_{j=h+1}^{n} b_j)} \times \left(1 - \frac{1}{2}f \right)^{2(\sum_{j=h+1}^{n} b_j - \sum_{i=1}^{h} b_i)} \end{aligned}$$

The sensitivity is maximised when $b'_i = \begin{cases} 1 \text{ when } i \in [h] \\ 0 \text{ otherwise} \end{cases}$, so,

$$\frac{\Pr[B' \in R | B = B_1]}{\Pr[B' \in R | B = B_2]} \leq \left(\frac{1 - \frac{1}{2}f}{\frac{1}{2}} \right)^{2h}$$

So taking $\varepsilon = \ln \left(\frac{1-\frac{1}{2}f}{\frac{1}{2}} \right)^{2h} = 2h \ln \left(\frac{1-\frac{1}{2}f}{\frac{1}{2}f} \right)$ works and completes the proof. ■

Limitations and Drawbacks of RAPPOR

That being said, RAPPOR ended up being used quite extensively, being the first large scale industrial implementation of differential privacy, but it was not without its flaws.

1. It cannot handle slight, gradual variations in data viz. age in days of a user, leaving it precariously open to longitudinal attacks.
2. It struggles with discovering new, unknown strings not in χ.

7.4.2 Microsoft's Low Communication LDP; Ding, Kulkarni, Yekhanin[21]

This approach is a considerable improvement over RAPPOR, and drastically brings down communication costs (with the cost itself being of the tune of 1 bit per user! That is impressive.).

Before we introduce this consider averaging numbers in $[0, m]$ under $\varepsilon-$local DP. A simple and conventional manner to achieve this would be to use the Laplace mechanism as follows

1. User i has a datapoint $X_i \in [0, m]$, transmits $X_i + \text{Lap}(\frac{m}{\varepsilon})$;
2. Curator averages responses as $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} (X_i + \text{Lap}(\frac{m}{\varepsilon}))$.

The resulting error would be $\hat{\mu} - \frac{1}{n} \sum_{i=1}^{n} X_i = \frac{1}{n} \sum_{i=1}^{n} \text{Lap}(\frac{m}{\varepsilon})$. and the standard deviation of error is of $O\left(\frac{m}{\varepsilon \sqrt{n}} \right)$. The values transmitted by the users will be $\Omega(m)$ (being in $[0, m]$) (corresponding to the magnitude of the noise), and ergo needs $\Omega(\log m)$ bits of communication.

But what Ding et al proposed, called *one bit estimation*, does the job with, as mentioned earlier, just a single bit of communication

In their approach, we use randomised response on X_i to generate a report, i.e. we transmit

$$Y_i = \begin{cases} 1 \text{ with probability } = \frac{1}{e^\varepsilon + 1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon - 1}{e^\varepsilon + 1} \\ 0 \text{ otherwise} \end{cases}$$

The reports are then aggregated by a curator using the formula

$$\hat{\mu} = \frac{m}{n} \sum_{i=1}^{n} \frac{Y_i \cdot (e^\varepsilon + 1) - 1}{e^\varepsilon - 1}.$$

This can be easily seen to be unbiased, i.e. $\mathbb{E}[\hat{\mu}] = \frac{1}{n} \sum_{i=1}^{n} X_i$, as follows,

$$\begin{aligned} \mathbb{E}[\hat{\mu}] &= \mathbb{E}\left[\frac{m}{n} \sum_{i=1}^{n} \frac{Y_i \cdot (e^\varepsilon + 1) - 1}{e^\varepsilon - 1} \right] \\ &= \frac{m}{n} \sum_{i=1}^{n} \frac{\mathbb{E}[Y_i] \cdot (e^\varepsilon + 1) - 1}{e^\varepsilon - 1} \end{aligned}$$

$$
\begin{aligned}
&= \frac{m}{n} \sum_{i=1}^{n} \frac{\left(\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}\right) \cdot (e^\varepsilon + 1) - 1}{e^\varepsilon - 1} \\
&= \frac{m}{n} \sum_{i=1}^{n} \frac{\left(1 + \frac{X_i}{m} \cdot (e^\varepsilon - 1)\right) \cdot -1}{e^\varepsilon - 1} \\
&= \frac{m}{n} \sum_{i=1}^{n} \frac{\left(\frac{X_i}{m} \cdot (e^\varepsilon - 1)\right)}{e^\varepsilon - 1} \\
&= \frac{1}{n} \sum_{i=1}^{n} X_i
\end{aligned}
$$

Lemma 7.5. ***ε-LDP of 1-Bit Mean Estimation***
For each round of data collection, 1-bit mean estimation is ε-locally differentially private for each user.

Proof. Each user sends a single Y_i to the curator, so we have

$$
\begin{aligned}
\ln\left|\frac{\Pr[Y_i = 1]}{\Pr[Y_i = 0]}\right| &= \ln\left|\frac{\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}{1 - \left(\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}\right)}\right| \\
&= \ln\left|\frac{\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}{\frac{e^\varepsilon}{e^\varepsilon+1} - \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}\right| \\
&= \ln\left|\frac{\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}{\frac{e^\varepsilon}{e^\varepsilon+1} - \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}\right| \\
&= \ln\left|\frac{\frac{e^\varepsilon}{e^\varepsilon+1} - \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}{\frac{1}{e^\varepsilon+1} + \frac{X_i}{m} \cdot \frac{e^\varepsilon-1}{e^\varepsilon+1}}\right| \\
&\leq \ln\left|\frac{\frac{e^\varepsilon}{e^\varepsilon+1}}{\frac{1}{e^\varepsilon+1}}\right| \\
&= \varepsilon.
\end{aligned}
$$

■

And furthermore, by a Chernoff bound, we can see that the error is $O\left(\frac{m}{\varepsilon\sqrt{n}}\right)$, the same as the Laplace mechanism approach! But to show that, we need to prove a lemma first.

Lemma 7.6. ***Chernoff-Hoeffding Bound on the Error for 1-Bit Mean Estimation***
For an arbitrary $\theta \in (0, 1)$, $\Pr[|\hat{\mu} - \mu| \geq \theta m] \leq 2 \cdot e^{-2\theta^2 \cdot n\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)^2}$.

Proof. By an application of the Chernoff-Hoeffding bound, and taking $M = \sum_i Y_i$, $1 \leq i \leq n$ (while considering that $Y_i \in \{0, 1\} \subseteq [0, 1]$, hence $\Delta = 1$), we get

$$
\begin{aligned}
&\Pr[|M - \mathbb{E}[M]| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{\sum_{i\in[]} \Delta_i^2}\right) \\
\implies &\Pr[|\sum_{i\in[n]} Y_i - \mathbb{E}\left[\sum_{i\in[n]} Y_i\right]| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{n}\right)
\end{aligned}
$$

$$\implies \Pr[|\sum_{i\in[n]} Y_i - \sum_{i=1}^{n} \mathbb{E}[Y_i]| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{n}\right)$$

$$\implies \Pr[|\sum_{i\in[n]} Y_i - \sum_{i=1}^{n} \left(\frac{1}{e^\varepsilon+1} + \frac{X_i}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{n}\right)$$

$$\implies \Pr[|\sum_{i\in[n]} Y_i - \left(\frac{n}{e^\varepsilon+1} + \frac{n\mu}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)| > \alpha] \leq 2\exp\left(\frac{-2\alpha^2}{n}\right)$$

$$\implies \Pr\left[|\mu - \hat{\mu}| > \alpha\cdot\frac{m}{n}\cdot\frac{e^\varepsilon+1}{e^\varepsilon-1}\right] \leq 2\exp\left(\frac{-2\alpha^2}{n}\right)$$

Let $\alpha = \theta n\dfrac{e^\varepsilon-1}{e^\varepsilon+1}$, then

$$\implies \Pr[|\hat{\mu}-\mu| \geq \theta m] \leq 2\cdot e^{-2\theta^2\cdot n\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)^2}.$$

Hence, proved. ■

Theorem 7.7. *For any $0 \leq \delta \leq 1$, with probability at least $1-\delta$, the error,*

$$|\hat{\mu}-\mu| \leq \frac{m}{\sqrt{2n}}\cdot\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot\sqrt{\log\frac{2}{\delta}}.$$

Proof. $\forall\,\delta \in [0,1]$, set $\delta = 2\cdot e^{-2\theta^2\cdot n\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right)^2}$, then.

$$\theta m \leq \frac{m}{\sqrt{2n}}\cdot\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot\sqrt{\log\frac{2}{\delta}}.$$

Which simply means that

$$\Pr\left[|\hat{\mu}-\mu| \geq \frac{m}{\sqrt{2n}}\cdot\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot\sqrt{\log\frac{2}{\delta}}\right] \leq \delta$$

$$\implies \Pr\left[|\hat{\mu}-\mu| \leq \frac{m}{\sqrt{2n}}\cdot\frac{e^\varepsilon+1}{e^\varepsilon-1}\cdot\sqrt{\log\frac{2}{\delta}}\right] \geq 1-\delta.$$

■

7.4.3 Bonus: Memoisation in Ding et al's Approach

Now let us consider queries for rapidly/realtime updated values like the age of the user (in days). Such queries are difficult to memoise against. So we can possibly we discretise this data into bins and round it to the centre of each in response to an incoming instance of such a query?

Yes; but having too many bins, i.e. overly fine discretisation), naturally implies a small bin width and ergo makes the discretised data sensitive to even small fluctuations, and therefore would involve repeated randomisation/movement across bins; and suffice this to say that this is precarious, for this can lead leakage of some information via the detection and study of said changes.

So what if we play safe and reduce the number of bins to a comfortable low? Well as it happens, in line with the seesaw of tradeoffs between accuracy and privacy, having too few bins (i.e. having coarse discretisation) would lead to a larger magnitude of error post rounding off to

the centres of the bins.
Ding et al's solution to this is their randomised rounding algorithm, which involves coarse discretisation but without the large amount of error.

1. Let $A_i(\cdot)$ be the 1-bit mean estimation function/algorithm;
2. Select offset $\alpha_i \in [m-1]$ uniformly at random;
3. For $X_i \in [0, m]$, output $Y_i = \begin{cases} A_i(0) \text{ if } X_i + \alpha_i \leq m \\ A_i(m) \text{ if } X_i + \alpha_i > m \end{cases}$;
4. Aggregate as for regular one bit estimation.

The output here would appear as a sequence of $A_i(0)$s followed by a sequence of $A_i(m)$s observed over time. Ding et al show that this enjoys the same accuracy guarantees as one bit mean estimation as described earlier, which is outstanding.

Lemma 7.8. *Define $\tilde{\mu} = \frac{1}{n}\sum_i Y_i$. Then $\mathbb{E}[\tilde{\mu}] = \mu$.*

Proof. Let $a := X_i$ and $b = m - X_i$. Let $Z_i = \begin{cases} b \text{ with probability } \frac{a}{a+b} \\ -a \text{ with probability } \frac{b}{a+b} \end{cases}$.
Then $\mathbb{E}[Z_i] = \frac{ba}{a+b} - \frac{ab}{a+b} = 0$. Note that in this case, it is easy to see that we can write $Y_i = X_i + Z_i$. Then,

$$\begin{aligned}
\mathbb{E}[\tilde{\mu}] &= \mathbb{E}\left[\frac{1}{n}\sum_{i\in[n]} Y_i\right] \\
&= \frac{1}{n}\sum_{i\in[n]} \mathbb{E}[Y_i] \\
&= \frac{1}{n}\sum_{i\in[n]} \mathbb{E}[X_i + Z_i] \\
&= \frac{1}{n}\sum_{i\in[n]} (\mathbb{E}[X_i] + \mathbb{E}[Z_i]) \\
&= \frac{1}{n}\sum_{i\in[n]} \mathbb{E}[X_i] \\
&= \mu.
\end{aligned}$$

■

The only revelation that can be made by a query is when the user i's value (say, their age in days) passes $m - \alpha_i$, which is private as the user gets to choose what their value of α_i is and said α_i is private to them.

Theorem 7.9. *This memoisation algorithm enjoys the same accuracy guarantees as that of 1-bit mean estimation.*

Proof. For this, it suffices to show that the output bit is still sampled according to the distribution (Bernoulli with the same parameters) as for 1-bit mean estimation. To see this, consider, with

$a := X_i, b := m - X_i$ (i.e. we use the notation as in the proof of the last lemma),

$$\begin{aligned}\Pr[Y_i = 1] &= \frac{b}{a+b}\left(\frac{1}{e^\varepsilon+1} + \frac{0}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right) + \frac{a}{a+b}\left(\frac{1}{e^\varepsilon+1} + \frac{m}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right) \\ &= \frac{m-X_i}{m}\left(\frac{1}{e^\varepsilon+1}\right) + \frac{X_i}{m}\left(\frac{1}{e^\varepsilon+1} + \cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}\right) \\ &= \frac{1}{e^\varepsilon+1} + \frac{X_i}{m}\left(\frac{e^\varepsilon-1}{e^\varepsilon+1}\right).\end{aligned}$$

And this completes our proof, as now the rest follows from our earlier discussion on 1-bit mean estimation's accuracy. ■

8 Bonus: Different Versions/Definitions of (Approximate) Differential Privacy

8.1 Rényi Differential Privacy (RDP)

Introduced by Ilya Mironov in 2016[10], this is a generalisation of the (ε, δ)- formulation of differential privacy that is equivalent to ε-differential privacy for pure DP, and affords certain guarantees like better behaviour in terms of optimal composition of "approximately" (in a loose sense, not exactly expressed in a (ε, δ) form, which seems to run into problems[22] while trying to compose heterogeneously approximately differentially private algorithms optimally; it is a $\#P$-hard problem) differentially private mechanisms/algorithms, notably for different instances of the Gaussian mechanism, and no catastrophic failure as discussed earlier for (ε, δ)-differential privacy. For Rényi differential privacy, advanced composition seems to behave better on, say, heterogeneous instances of the Gaussian mechanism. We shall discuss Rényi DP briefly here, perhaps without going into much, if any, proofs (all of which are beautifully laid out in [10], but we shall, for the sake of brevity, not go into all of, if any of, those).

The main motivation for RDP, and another concept that we shall soon discuss is that while we have some upper bounds on the privacy cost of a mechanism in terms of (ε, δ)-differential privacy, these bounds are loose, i.e. the actual privacy loss bound is quite less than the aforementioned upper bounds, so we need a better definition of approximate differential privacy (bounds for composition of pure (i.e. ε-) DP given by simple composition are tight) in order to obtain tighter privacy loss bounds. In other words, these new versions of differential privacy result in less values of ε for the same amount of noise added.[23]

We define RDP using the concept of what is called *Rényi divergence*.

Definition 8.1. Rényi Divergence
We define Rényi divergence between databases P and Q with parameter α as

$$D_\alpha(P\|Q) := \frac{1}{\alpha-1}\ln \mathbb{E}_Q\left[\left(\frac{P(x)}{Q(x)}\right)^\alpha\right]$$

where $\mathbb{E}_Q$ denotes expectation over the probability distribution for Q.

We define

$$D_\infty(P\|Q) := \lim_{\alpha\to\infty} D_\alpha(P\|Q) = \ln \max_x \frac{P(x)}{Q(x)}$$

and

$$D_1(P\|Q) := \lim_{\alpha\to 1} D_\alpha(P\|Q) = \mathbb{E}_P\left[\ln \frac{P(x)}{Q(x)}\right].$$

Now one can look at the definition of $D_\infty(\cdot\|\cdot)$ and (almost) immediately think of how it resembles the LHS of the definition of pure differential privacy, and if so, one would not be mistaken. We now define *Rényi Differential Privacy* concretely.

Definition 8.2. Rényi Differential Privacy (RDP)
We say that a mechanism $\mathcal{M}$ on $\mathbb{N}^{|\chi|}$ is (α, ε)- (or $\varepsilon(\alpha)$-)Rényi Differentially Private if $\forall$ neighbouring $x, x' \in \mathbb{N}^{|\chi|}$

$$D_\alpha(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \varepsilon.$$

In other terms, we can express the condition for (α, ε)-Rényi differential privacy as, for any arbitrary set S, and for any neighbouring databases x, x',

$$\Pr[\mathcal{M}(x) \in S] \leq \left(e^\varepsilon \cdot \Pr[CM(x') \in S]^{1-\frac{1}{\alpha}}.\right)$$

Note that for $\alpha \to \infty$, this is equivalent to ε-differential privacy as

$$\begin{aligned} & D_\infty(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \varepsilon \\ \implies & \ln \max \frac{\mathcal{M}(x)}{\mathcal{M}(x')} \leq \varepsilon \end{aligned}$$

And as Rényi DP holds for any pair of adjacent databases, we can repeat the same process for $D_\infty(\mathcal{M}(x')\|\mathcal{M}(x)) \leq \varepsilon$. This essentially gives us the definition of ε-differential privacy. We can summarise this little discussion as follows.

Theorem 8.3. (∞, ε) *Rényi differential privacy is equivalent to* ε*-differential privacy.*

One can perhaps intuitively say at this point that the higher α is, the less "approximate" the guarantee of differential privacy is. More concretely, we have the following result.

Theorem 8.4. (α, ε)*-Rényi differential privacy implies* $\left(\varepsilon + \frac{\log \frac{1}{\delta}}{\alpha - 1}, \delta\right)$ *differential privacy.*

In addition to all that, RDP satisfies all the salient properties of differential privacy, viz. post processing invariance, group privacy, ease of composition, and even advanced composition. We shall not venture into proving any of those results for RDP here, for aforementioned reasons.[5]

Theorem 8.5. ***Monotonicity***
For $\alpha_1 \geq \alpha_2, (\alpha_1, \varepsilon)$*-Rényi differential privacy implies* (α_2, ε)*-Rényi differential privacy.*

[5]Again, interested readers may refer to the brilliantly presented results and proofs in [10] if they are so inclined.

Theorem 8.6. ***Composition***
Composing a (α, ε_1)-Rényi differentially private algorithm with a (α, ε_2)-Rényi differentially private algorithm yields $(\alpha, \varepsilon_1 + \varepsilon_2)$-Rényi differential privacy.

Theorem 8.7. ***Advanced Composition***
Let $\mathcal{M}$ be the adaptive composition of k ε-differentially private mechanisms. Then for any two adjacent databases, x, x', and any arbitrary set S,

$$\Pr[\mathcal{M}(x) \in S] \leq \exp\left(2\varepsilon\sqrt{n \ln \frac{1}{\Pr[\mathcal{M}(x') \in S]}}\right) \cdot \Pr[\mathcal{M}(x') \in S].$$

The above formulation of advanced composition, upon close examination, is equivalent to the one we introduced earlier for (ε, δ)-differentially private algorithms.

Theorem 8.8. ***From RDP to DP***
If a randomised mechanism $\mathcal{M}$ satisfies $(\alpha, \varepsilon(\alpha))$-RDP, then $\mathcal{M}$ also satisfies $(\varepsilon(\alpha) + \frac{\log(\frac{1}{\delta})}{\alpha - 1}, \delta)$-DP for any $\delta \in (0, 1)$.

Rényi differential privacy has lately been a thriving field of research and has been touted as being a convenient mathematical framework, vis-à-vis the formulation of (ε, δ)-differential privacy, to work with in certain contexts[24].

8.2 Zero-Concentrated Differential Privacy (zCDP)

This is a modification of Rényi differential privacy formulated by Bun and Steinke[25] and used to prove better quantitative results, establish lower privacy loss/error bounds. We shall, however, merely dive into the very basics of zCDP, viz. the definition, simple composition, post processing invariance, and group privacy of zCDP, sans proofs, as further discussion on this would be a digression and is beyond the scope of this work. Interested readers may refer to the very comprehensive paper by Bun and Steinke[25] introducing the same for the proofs and an in depth look at zCDP.

Definition 8.9. Zero-Concentrated Differential Privacy (zCDP)
Given a privacy parameter ρ, a randomised mechanism $\mathcal{M}$ satisfies (η, ρ)-zCDP if $\forall$ neighbouring databases x, y and all numbers $\alpha > 1$,

$$D_\alpha(\mathcal{M}(x) \| \mathcal{M}(x')) \leq \eta + \rho\alpha.$$

If $\eta = 0$, we say that we have ρ-zCDP.

Theorem 8.10. ***Composition***
For a set $\{\mathcal{M}_i\}_{i \in [k]}$ of algorithms that are ρ_i-zero concentrated differentially private respectively, then $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k)$ is $\left(\sum_{i=1}^{k} \rho_i\right)$-zero concentrated differentially private.

Theorem 8.11. ***Post-Processing Invariance***
Let $\mathcal{M} : \mathbb{N}^{|\chi|} \to Y$ and $f : Y \to Z$ be randomized algorithms, with $\mathcal{M}$ satisfying ρ-zCDP. Then $f \circ \mathcal{M}$ satisfies $\rho - zCDP$.

Theorem 8.12. ***Group Privacy***
If $\mathcal{M}$ is a randomised algorithm satisfying ρ-zCDP, then for groups of size k, $\mathcal{M}$ guarantees $k^2\rho$-zCDP, i.e. for $x, y \in \mathbb{N}^{|\chi|}$ such that $\|x - y\|_1 \leq k$,

$$D_\alpha(\mathcal{M}(x) \| \mathcal{M}(y')) \leq (k^2\rho)\alpha.$$

Now to tie this into (ε, δ)-differential privacy, we have the following results.

Theorem 8.13. ***ρ-zCDP in terms of (ε, δ)-Differential Privacy***
If $\mathcal{M}$ is a randomised algorithm that satisfies ρ-zCDP, then $\mathcal{M}$ satisfies $(\rho + 2\sqrt{\rho \log(\frac{1}{\delta})}, \delta)$-differential privacy.

Theorem 8.14. ***ε-DP to ρ-zCDP***
If $\mathcal{M}$ is an ε-differentially private randomised algorithm, then for $\rho = \frac{1}{2}\varepsilon^2$, $\mathcal{M}$ satisfies ρ-zCDP.

The Gaussian mechanism can be expressed in terms of ρ-zCDP as follows.

Definition 8.15. Gaussian Mechanism (in terms of ρ-zCDP
Given $\rho > 0$, a set of queries, Q, and an input database x, the Gaussian mechanism $\mathcal{M}_G$ returns noisy answers $\{q(x) + \mathrm{N}(0, \Delta_2(Q)^2/(2\rho))\}_{q \in Q}$.

This endows ρ-zCDP upon the query response(s).

9 Some Recent Papers

Now we shall discuss a couple of recent papers published in/accepted to premier ML conferences like NeurIPS/ICML, and mostly restrict out attention to some important results in and the mathematical aspects of the same.

9.1 NeurIPS 2021; Abowd et al.[26]

This portion will majorly focus on the very recent (accepted to) NeurIPS 2021 paper by Abowd et al[26]. Among the authors are some of the same people at the U.S. Census Bureau who carried out the reconstruction attack on the 2010 US Census (refer to the the pertinent case study in this document for details on that).
The mode of operation in this new differentially private approach, keeping in mind a few factors, viz. the need to release synthetic databases for offline analysis in the absence of a curator to actively answer queries, the fact that noise addition might lead to negative (noised) values being reported and any synthetic database for demographic analysis cannot have negative values in it (as demographic data in censuses is not expressed in negative values), etc., is as follows.

- Given a defined set of queries, the system produces *measurements*, i.e. query responses with differentially private noise;
- Then it produces synthetic database, or privacy-protected microdata such that querying on this microdata matches the respective measurements as closely as possible.

This might intuitively seem very reasonable and one would expect this to work at least decently for demographic analysis.

However an end user survey of an early prototype of this approach showed considerable anomalies in the privacy-protected microdata: query responses from queries made on the privacy-protected microdata produced much larger query error than the measurements.

This is due to the nature of privacy-protected microdata; there exists an uncertainty principle describes a tradeoff between accuracy of queries answered with differential privacy on populations (viz. states) and those on subpopulations (viz. counties within states). Abowd et al in [26] strive to show fairly tight bounds on errors on the populations or subpopulations, given that we prioritise one to have the least error possible, like that comparable to that when Laplace noise is added.

Why this is a concern is apparent: if we have too much error at a "larger" level (viz. states), it will affect the way federal funds are allocated to states, and if it happens at the subpopulation level (viz. for counties/cities in the state), then it will affect, for instance, the layout of electoral constituencies, which are redefined post every census in the US.

9.1.1 Some Mathematical Definitions

Abowd et al use per-query squared error for error analysis; which is given as follows: for a collection Q of queries, raw microdata x, and privacy-protected $\tilde{x}$, per query squared error is given by $PQSE(Q) = \max_{q \in Q} \mathbb{E}_{\tilde{x}}[(q(x) - q(\tilde{x})^2]$. This is in contrast to simultaneous/outlier error, $\mathbb{E}_{\tilde{x}}[\max_{q \in Q}(q(x) - q(\tilde{x}))^2]$, which most other existing literature on this topic uses.

Per-query squared error considers the average behaviour of each query and tells whether $\exists$ bad queries that have large errors on average. Outlier error focuses on all possible $\tilde{x}$ instead of the queries themselves, cannot tell whether it is large errors in some queries or some outliers that occur by chance when there are a lot of random variables.

A useful fact (stated without proof) used in this paper is that $\max_{q \in Q} \mathbb{E}_{\tilde{x}}[(q(x) - q(\tilde{x})^2] \leq \mathbb{E}_{\tilde{x}}[\max_{q \in Q}(q(x) - q(\tilde{x}))^2]$.

We now indulge in defining some pertinent terms.

Definition 9.1. Disjoint Queries
A collection of queries Q is called disjoint if changing one record of the database affects at most one of the queries in Q..

Now we define a dichotomy of queries that will help make the discussion about queries on populations and subpopulations more formal.

Definition 9.2. Point Queries and their Sum Query
Consider a collection of disjoint counting queries $q_1, q_2, \ldots, q_d$, and define the query q_*, where $\forall\, x \in \mathbb{N}^{|\chi|}, q_*(x) = \sum_{i=1}^{n} q_i(x)$.

Then we call each q_i, for $1 \leq i \leq n$ a *point query*, and q_* as the *sum query* of this set of point queries.

To slightly generalise and to simplify the proofs that we shall encounter, we shall consider databases (alternatively called datasets in the context of discussions regarding DP) which are (positively) weighted, i.e. each record r of the database has a weight $w_r \geq 0$.

For a query q with a predicate ψ, we evaluate, for a database x and every record r of x, $q(x) = \sum_{r \in x, r \text{ satisfies } \psi} w_r$.

Notice that normal microdata is merely a positively weighted database with all its records having weight $= 1$.

9.1.2 Error Bounds and Analysis

Therefore, it is apparent that the sensitivity of each point query (being counting queries) is 1, and as each query q_i works on a disjoint partition of x, we have that change in a record affects only the pertinent point query and the sum query, so for neighbouring databases x, y, they differ only in one place and $\|x - y\|_1$ so $\Delta_1(q_1, \ldots, q_d, q_*) = 2$; and thus adding $\text{Lap}(\frac{2}{\varepsilon})$ noise to the response to each each query in the tuple $(q_1, q_2, \ldots, q_d, q_*)$ gives us ε-DP, which in turn implies that the expected squared error of each query answer is $\frac{8}{\varepsilon^2}$ ($\because X \sim \text{Lap}(\sigma) \implies \text{Var}(X) = 2\sigma^2$).

Similarly, as $\Delta_2(q_1, \ldots, q_d, q_*) = \sqrt{1^2 + 1^2} = \sqrt{2}$, adding $N(0, (\frac{(\sqrt{2})^2}{2\rho}) = N(0, \frac{1}{\rho})$ noise to each query gives us ρ-zCDP.

While this seems to be fine so far, after noise addition to obtain measurements, post-processing (as mentioned earlier) is done to generate privacy-protected microdata (simply called microdata in the paper when the context is clear), so the end-result's accuracy is no longer described by the above analysis.

The paper establishes the following tradeoff in its most prominent results, which we shall soon discuss: Let $\mathbb{E}_{\tilde{x}}[(q_*(x) - q_*(\tilde{x}))^2] \leq D^2, \max \mathbb{E}_{\tilde{x}}[(q_i(x) - q_i(\tilde{x}))^2] \leq C^2$ for constants C, D.

- If we prioritise the accuracy of the sum query, i.e. $D^2 \in O(\frac{1}{\varepsilon^2}) \implies C^2 \in \Omega(\frac{1}{\varepsilon^2} \log^2(d))$, i.e. the bound for expected squared error for point queries incurs a penalty by a factor of $\log^2(d)$; and
- If we prioritise the accuracy of the point queries, i.e. $C^2 \in O(\frac{1}{\varepsilon^2}) \implies D^2 \in \Omega(\frac{d^2}{\varepsilon^2})$, i.e. the bound for expected squared error for point queries incurs a penalty by a factor of d^2

Note that these errors are unavoidable for "difficult" databases, of which a trivial/pathological example would be one where all the query answers (unnoised) equal zero, except for one which is equal to $\frac{\log(d)}{\varepsilon}$. The Laplace ordinarily would incur a per query error of $\frac{8}{\varepsilon^2}$ but a lot of the query responses will be negative after noise addition. So we shall soon see that any algorithm creating privacy-protected microdata from this can perform as well, error-wise, than merely applying Laplace noise.

Whereas for an "easy" database, i.e. if all the unnoised query answers happen to be "large" (i.e. $\geq \frac{\log(d)}{\varepsilon}$) on said database, then after adding Laplace noise we shall get only non negative answers and hence it will mean that we have a low error ($\frac{8}{\varepsilon^2}$).

Now we can come to the main theorems of this paper, which specify lower and upper bounds on the per query squared errors for point queries and their sum query in various contexts. We shall only be proving the results for ε- and (ε, δ)-DP in the interest of keeping this discussion fairly concise, for the rest (i.e. ρ-zCDP), one can refer to the cited paper's appendix.

Theorem 9.3. ***Theorem on Lower Bounds***
$\{q_i : i \in [d]\}$ is a set of disjoint queries and q_* is the respective sum query. Let $\mathcal{M}$ be a randomised algorithm as described above. Let, taking expectations over the randomness in $\mathcal{M}$, $\mathbb{E}[(q_i(x) - q_i(\mathcal{M}(x)))^2] \leq C^2$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}(x)))^2] \leq D^2$ for constants C, D. Then

1. *$\mathcal{M}$ satisfies ε-DP $\implies \forall\, k > 0, e^{2\varepsilon(2C+k)} \geq \frac{k(d-1)}{16C+8D+4k}$, which implies*
   (a) *if $D^2 \leq \lambda/\varepsilon^2$ for some constant λ, then $C^2 \in \Omega(\frac{1}{\varepsilon^2}\log^2(d))$; and*
   (b) *if $C^2 \leq \lambda/\varepsilon^2 \implies D^2 \in \Omega(d^2/\varepsilon^2)$.*
2. *If $\mathcal{M}$ satisfies (ε, δ)-DP $\implies \forall\, k > 0, \left(\frac{\delta}{\varepsilon} + \frac{4C+2D+k}{k(d-1)}\right) e^{4\varepsilon C + 2k\varepsilon} \geq \frac{1}{4}$, which implies*
   (a) *if $D^2 \leq \frac{\lambda}{\varepsilon^2}$ for some λ, then $C^2 \in \Omega(\min(\frac{1}{\varepsilon^2}\log^2(d), \frac{1}{\varepsilon^2}\log^2\frac{\varepsilon}{\delta}))$; and*
   (b) *if $C^2 \leq \frac{\lambda}{\varepsilon^2} \implies$ either $\varepsilon \in O(\delta)$ or $D^2 \in \Omega(d^2/\varepsilon^2)$.*
3. *If $\mathcal{M}$ satisfies ρ-zCDP, then the tradeoff function between C, D (omitted in the statement by Abowd et al) implies*
   (a) *if $D^2 \leq \lambda/\rho \implies C^2 \in \Omega(\log(d)/\rho)$; and*
   (b) *if $C^2 \leq \lambda/\rho$, then $\forall\, \lambda \in (0,1)$, $D^2 \in \Omega(d^{2\gamma}/\rho)$.*

Proof. Let us denote $a[i] = q_i(x)$ and therefore $\sum_i a[i] = q_*(x)$. Similarly, let $\tilde{a}[i] := q_i(\tilde{x})$. Therefore the vectors a and $\tilde{a}$ represent the query answers obtained on the raw and privacy-protected microdata respectively. Notice that a here is a vector of non-negative integers, and $\tilde{a}$ is one of non-negative real numbers. All probabilities in the following proof are taken with respect to only the randomness of $\mathcal{M}$.

For constants α, β, k (which we shall assign appropriate values to later), and for any fixed j, by an application of Markov's inequality, we have,

$$\Pr[|\tilde{a}[j] - a[j]| \geq \alpha C] \leq \frac{\mathbb{E}[(a[j] - \tilde{a}[j])^2]}{C^2\alpha^2} \leq \frac{1}{\alpha^2}$$

$$\Pr\left[|\sum_{i=1}^{d}\tilde{a}[i] - \sum_{i=1}^{d} a[i]| \geq \beta D\right] \leq \frac{1}{\beta^2}$$

$\forall\, n \in \mathbb{N}, k > 0$, and $i = 2, \ldots, d$, define

$$G_{i,n,k} = \{\tilde{a} : \tilde{a}[i] \in [k, k + 2\alpha C], \tilde{a}[1] \in [n - 2\alpha C - k, n - k], \sum_{j=1}^{d}\tilde{a}[j] \in [n - \beta D, n + \beta D]\}$$

Now for any fixed n, k, we see how many $G_{i,n,k}$ a vector $\tilde{a}$ can belong to. If $\tilde{a} \in G_{i,n,k}$, then $\tilde{a}[i] \geq k, \tilde{a}[1] \geq n - 2\alpha C - k \implies \sum_{i=2}^{d}\tilde{a}[i] \leq k + 2\alpha C + \beta D$ and as for each i, $\tilde{a}[i] \geq k$, $\tilde{a}$ can belong to $G_{i,n,k}$ for at most $\frac{k+2\alpha C+\beta D}{k}$ indices i.

Define the vectors $a_i, 1 \leq i \leq d$, where $a_1[i] = \begin{cases} n \text{ for } i = 1 \\ 0 \text{ otherwise} \end{cases}$, and for $2 \leq i \leq d$,

$$a_i[j] = \begin{cases} n - \alpha C - k \text{ for } j = 1; \\ \alpha C + k \text{ for } j = i \\ 0 \text{ otherwise} \end{cases}.$$

Now for each i, let x_i be a database whose point query answer vector is a_i, and x_i exists because the point queries are disjoint, and x_1 differs from all other x_i by at least a difference of $2(\alpha C + k)$ records.

1. **For ε-DP**,

$$
\begin{aligned}
1 &\geq \Pr\left[\mathcal{M}(x_1) \in \bigcup_{i=2}^{d} G_{i,n,k}\right] \geq \frac{k}{2\alpha C + \beta D + k} \sum_{i=2}^{d} \Pr[\mathcal{M}(x_1) \in G_{i,n,k}] \\
&\geq \exp(-\varepsilon 2(\alpha C + k)) \frac{k}{2\alpha C + \beta D + k} \sum_{i=2}^{d} \Pr[\mathcal{M}(x_i) \in G_{i,n,k}] && (\because \text{ of group privacy}) \\
&\geq \exp(-\varepsilon 2(\alpha C + k)) \frac{k}{2\alpha C + \beta D + k} \sum_{i=2}^{d} \left(1 - \frac{2}{\alpha^2} - \frac{1}{\beta^2}\right) && (\text{by Markov inequality and a union bound}) \\
&= \exp(-\varepsilon 2(\alpha C + k)) \frac{k(d-1)}{2\alpha C + \beta D + k} \left(1 - \frac{2}{\alpha^2} - \frac{1}{\beta^2}\right)
\end{aligned}
$$

   Setting $\alpha = 2, \beta = 2$, we get

$$\exp(2\varepsilon(2C + k)) \geq \frac{k(d-1)}{16C + 8D + 4k}.$$

   If D is allowed to be $\leq C$, then let $k = C$, then

$$\exp(6\varepsilon C) \geq d - 128 \implies C \geq \frac{1}{6\varepsilon} \log \frac{d-1}{28}$$

   In general, if $D \in O(C)$ then by similar arguments we see that $C \in \Omega(\frac{1}{\varepsilon} \log(d))$. If D is allowed to be $> C$, then upon setting $k = \frac{1}{\varepsilon}$ we get

$$e^{4\varepsilon C + 2} \geq \frac{(d-1)}{24\varepsilon D + 4} \implies C \geq \frac{1}{4\varepsilon} \left(\log\left(\frac{d-1}{24\varepsilon D + 4}\right)\right).$$

   More generally, if $D \in \Omega(C)$, then similarly, we get that $C \in \Omega(\frac{1}{\varepsilon} \log(\frac{d}{\varepsilon D}))$. $\implies$ $D \in O(\frac{1}{\varepsilon})$ $\implies$ $C \in \Omega(\frac{1}{\varepsilon} \log(d))$.

   And it can be seen similarly that $C \in O(\frac{1}{\varepsilon})$, then $D \in \Omega(d \times \frac{1}{\varepsilon})$.

2. **For (ε, δ)-DP**,

$$
\begin{aligned}
1 &\geq \Pr\left[\mathcal{M}(x_1) \in \bigcup_{i=2}^{d} G_{i,n,k}\right] \geq \frac{k}{2\alpha C + \beta D + k} \sum_{i=2}^{d} \Pr[\mathcal{M}(x_1) \in G_{i,n,k}] \\
&\geq \frac{k}{2\alpha C + \beta D + k} \sum_{k=2}^{d} (\exp(-2\varepsilon(\alpha C + k)) \Pr[\mathcal{M}(x_i) \in G_{i,n,k}] - \delta/\varepsilon) \\
&\geq \frac{k}{2\alpha C + \beta D + k} \sum_{i=2}^{d} \left(\exp(-2\varepsilon(\alpha C + k)) \left(1 - \frac{2}{\alpha^2} - \frac{1}{\beta^2}\right) - \delta/\varepsilon\right) \\
&= \frac{k(d-1)}{2\alpha C + \beta D + k} \left(\exp(-2\varepsilon(\alpha C + k)) \left(1 - \frac{2}{\alpha^2} - \frac{1}{\beta^2}\right) - \delta/\varepsilon\right)
\end{aligned}
$$

   Setting $\alpha = 2, \beta = 2$ gives

$$1 \geq \frac{k(d-1)}{4C + 2D + k} \left(\frac{1}{4} \exp(-2\varepsilon(2C + k)) - \delta/\varepsilon\right) \implies \left(1 + (\delta/\varepsilon) \frac{k(d-1)}{4C + 2D + k}\right) \exp(4\varepsilon C + 2k\varepsilon) \geq \frac{1}{4} \frac{k(d-1)}{4C + 2D + k}.$$

Which implies

$$\left(\frac{\delta}{\varepsilon} + \frac{4C + 2D + k}{k(d-1)}\right) \exp(4\varepsilon C + 2k\varepsilon) \geq \frac{1}{4}.$$

$\forall\, z \in \mathbb{R}, 1 + z \leq 2\max(1, z)$, which implies

$$\exp(4\varepsilon C + 2k\varepsilon) \geq \frac{1}{8} \min\left(\frac{k(d-1)}{4C + 2D + k}, \frac{\varepsilon}{\delta}\right).$$

Proceeding as we did in the ε-DP case, we get that if $D \in O(C) \implies C \in \Omega(\min(\frac{1}{\varepsilon}\log(d), \frac{1}{\varepsilon}\log\frac{\varepsilon}{\delta}))$.
If $D \in \Omega(C) \implies C \in \Omega(\min(\frac{1}{\varepsilon}\log\frac{d}{\varepsilon D}, \frac{1}{\varepsilon}\log\frac{\varepsilon}{\delta}))$.
Which implies that if $D \in O(\frac{1}{\varepsilon})$, then $C \in \Omega(\min(\frac{1}{\varepsilon}\log(d), \frac{1}{\varepsilon}\log\varepsilon\delta))$; and if $C \in O(\frac{1}{\varepsilon})$, then $\varepsilon \in O(\delta)$ or $D \in \Omega(\frac{d}{\varepsilon})$.

■

For the result on upper bounds, we first need to look at some results, including some facts about Gaussian and Laplace random variables (fairly well known results and thus stated without proof) and a lemma.

Lemma 9.4. *Let $y_i, 1 \leq i \leq d$, be i.i.d. random variables from a distribution A*

- *If A is $N(0, \sigma^2)$, then*
  - $\mathbb{E}[y_i^2] = \sigma^2, \forall\, i$;
  - $\mathbb{E}[|y_i|] \leq \sigma, \forall\, i$;
  - $\mathbb{E}[\max_i |y_i|] \in O(\sigma\sqrt{\log(d)})$;
  - $\mathbb{E}[\max_i y_i^2] \in O(\sigma^2 \log(d)$.
- *If A is Lap$(\frac{1}{\varepsilon})$, then*
  - $\mathbb{E}[y_i^2] = \frac{2}{\varepsilon^2}, \forall\, i$;
  - $\mathbb{E}[|y_i|] = \frac{1}{\varepsilon}, \forall\, i$;
  - $\mathbb{E}[\max_i |y_i|] \leq \frac{1}{\varepsilon}(\ln(d) + 1)$;
  - $\mathbb{E}[\max_i y_i^2] \leq \frac{1}{\varepsilon^2}(\ln(d) + 2\ln(d) + 2)$.

Then we state the following theorem, for whose proof one can refer to the appendix of [26].

Lemma 9.5. *Result on the Solution to a Constrained Non-Negative Least Squares Problem*
Let $r_1, \ldots, r_d \in \mathbb{R}, r_* \geq 0$, then the solution to the optimisation problem

$$\arg\min_{t_1,\ldots,t_d} \frac{1}{2}\sum_{i=1}^{d}(t_i - r - i)^2$$

$$\textit{s.t.} \sum_{i=1}^{d} t_i = r_*$$

$$t_i \geq 0, \textit{ for } i = 1, \ldots, d$$

is $t_i = \max\{r_i - \gamma, 0\}, \forall\, i$ where γ is a value chosen such that $\sum_{i=1}^{d}\max\{0, r_i - \gamma\} = r_*$.

Theorem 9.6. ***Theorem on Upper Bounds***
Given privacy parameters $\varepsilon > 0, \rho > 0, \exists$ algorithms $\mathcal{M}_\varepsilon, \mathcal{M}_\rho, \mathcal{M}'_\varepsilon, \mathcal{M}'_\rho, \mathcal{M}'_{\varepsilon,\delta}$ that output positive weighted databases and have the properties:

1. *$\mathcal{M}_\varepsilon$ satisfies ε-DP, and $\forall\, x \in \mathbb{N}^{|\chi|}$ and $\forall\, i$, $\mathbb{E}[(q_i(x) - q_i(\mathcal{M}_\varepsilon(x))^2] \leq \frac{2}{\varepsilon^2}$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}_\rho(x)))^2] \leq \frac{2d^2}{\varepsilon^2}$;*
2. *$\mathcal{M}_\rho$ satisfies ρ-zCDP, and $\forall\, x \in \mathbb{N}^{|\chi|}$ and $\forall\, i$, $\mathbb{E}[(q_i(x) - q_i(\mathcal{M}_\rho(x))^2] \leq \frac{1}{2\rho}$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}_\rho(x)))^2] \leq \frac{d^2}{2\rho}$;*
3. *$\mathcal{M}'_\varepsilon$ satisfies ε-DP, then $\forall\, x, i, \mathbb{E}[(q_i(x) - q_i(\mathcal{M}'_\varepsilon(x))^2] \in O(\log^2(d)/\varepsilon^2)$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}'_\varepsilon(x)))^2] \in O(1/\varepsilon^2)$;*
4. *$\mathcal{M}_\rho$ satisfies ρ-zCDP, then $\forall\, x, i, \mathbb{E}[(q_i(x) - q_i(\mathcal{M}'_\rho(x)))^2] \in O(\log(d)/rho)$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}'_\rho(x)))^2] \in O(1/\rho)$;*
5. *$\mathcal{M}'_{\varepsilon,\delta}$ satisfies (ε, δ)-DP, then $\forall\, x, i, \mathbb{E}[(q_i(x) - q_i(\mathcal{M}'_{\varepsilon,\delta}(x)))^2] \in O(\log^2(1/\delta)/\varepsilon^2 + 1$ and $\mathbb{E}[(q_*(x) - q_*(\mathcal{M}'_{\varepsilon,\delta}(x)))^2] \in O(\frac{1}{\varepsilon^2})$.*

Proof. We first define $DGeo(\varepsilon)$, which is the double-sided geometric distribution and a discrete version of the Laplace distribution, supported over integers and with the pmf, $p(k) = \frac{1-e^{-\varepsilon}}{1+e^{-\varepsilon}} e^{-\varepsilon|k|}$. Some of its useful properties are as follows.

1. It has mean$= 0$;
2. Its variance$= 2\frac{e^{-\varepsilon}}{(1-e^{-\varepsilon})^2} \leq 2/\varepsilon^2$;
3. Given an integer-valued query q, adding noise from $DGeo(\varepsilon/\Delta_1(q)$ to its response yields ε-DP;
4. If $k > 0$, then $\Pr[z \geq k] = \Pr[z \leq -k] = \frac{1}{1+e^{-\varepsilon}} e^{-\varepsilon k}$.

And we define the discrete Gaussian $DGauss(0, \frac{1}{2\rho})$, which is, as the name suggests, a discrete version of the Gaussian distribution. It possesses some salient properties,

1. Its mean is 0;
2. Its variance is less than that of$N(0, \frac{1}{2\rho})$; and
3. Given a query with outputs$\in \mathbb{Z}$, adding $DGauss(0, \Delta_2(q)^2/(2\rho)$ yields ρ-zCDP.

Now we shall prove each part of the theorem separately.

1. Let $r_1, \ldots, r_d$ be records satisfying predicates ψ_i for point queries $q_i, 1 \leq i \leq d$ respectively. Let $\mathcal{M}_\varepsilon$ be an algorithm which first computes non-negative query answers $a_i = \max\{0, q_i(x) + DGeo(\frac{1}{\varepsilon})\}$ for $i \in [d]$ and then outputs a synthetic database $\bar{x}$ that contains a_i copies of record r_i for each i. $\mathcal{M}_\varepsilon$ satisfies ε-DP as it does not compute a noisy answer for q_* and as $\Delta(q_1, \ldots, q_d) = 1$. Since $q_i(x) \geq 0$ for all i, we have

$$\begin{aligned}
\mathbb{E}[(q_i(x) - q_i(\mathcal{M}_\varepsilon(x))]^2 &= \mathbb{E}[(q_i(x) - \max\{0, q_i(x) + DGeo(\varepsilon)\})^2] \\
&\leq \mathbb{E}[(q_i(x) - (q_i(x) + DGeo(\varepsilon)))^2] \\
&\leq \frac{2}{\varepsilon^2}.
\end{aligned}$$

The same for the sum query is given by

$$\mathbb{E}[(q_*(x) - q_8(\mathcal{M}_\varepsilon(x)))^2] = \mathbb{E}\left[\left(\sum_i q_i(x) - \sum_i q_i(\mathcal{M}_\varepsilon(x))\right)^2\right]$$
$$= \sum_i \mathbb{E}[(q_i(x) - \max\{0, q_i(x) + DGeo(\varepsilon)\})^2]$$
$$+ 2\sum_{1,j:i<j} \mathbb{E}[(q_i(x) - \max\{0, q_i(x) + DGeo(\varepsilon)\})]\mathbb{E}[(q_j(x) - \max\{0, q_j(x) - \max\{0, q_j(x) + DGeo(\varepsilon)\})]$$
$$\leq d\frac{2}{\varepsilon^2} + d(d-1)\frac{2}{\varepsilon^2} = d^2\frac{2}{\varepsilon^2}$$

2. We prove this similarly as above, with the exception that $\mathcal{M}_\rho$ synthesises $\tilde{D}$ with the noisy answers $a_i = q_i(x) + \max\{0, DGauss(0, \frac{1}{2\rho})\}$, and then proceeding practically similarly as above, we obtain that the expected squared error of each point query q_i is $\leq \frac{1}{2\rho}$, and for the sum query q_* is $\leq \frac{d^2}{2\rho}$.

3. Let $r_1, \ldots, r_d$ be records satisfying predicates ψ_i for point queries $q_i, 1 \leq i \leq d$ respectively. Let $\mathcal{M}'_\varepsilon$ be an algorithm which first computes noisy answers for each query, i.e. $a_i = q_i(x) + \text{Lap}(\frac{2}{\varepsilon})$ and $a_* = q_*(x) + \text{Lap}(\frac{2}{\varepsilon})$. By our earlier discussion, we know that this satisfies ε-DP. We have $\mathcal{M}$ that solves the following optimisation problem.

$$\arg\min_{w_1,\ldots,w_d} \frac{1}{2}\sum_{i=1}^{d}(w_i - a_i)^2$$
$$\text{such that } \sum_{i=1}^{d} w_i = \max\{0, a_*\}$$
$$w_i \geq 0, \forall\, i \in [d].$$

And thus produces a privacy protected microdata $\tilde{x}$ which contains the records $r_1, \ldots, r_d$ with respective weights $w_1, \ldots, w_d$.

Since the sum query is non-negative and we need to have that $\sum_{i=1}^{d} w_i = \max\{0, a_*\}$, then it follows that $\mathbb{E}[(q_*(\mathcal{M}'_\varepsilon(x)) - q_*(x)))^2] \leq \frac{2}{\varepsilon^2}$.

Now $\forall\, i \in [d]$, define $z_i := a_i - q_i(x), z_* = a_* - q_*(x)$, which are just the value of the actual noises added (and are all i.i.d. $\text{Lap}(\frac{2}{\varepsilon})$.

By *lemma* 9.5, w_i is if the form $\max\{a_i - \gamma, 0\} = \max\{q_i(x) + z_i - \gamma, 0\}$ for some γ such that $\sum_{i=1}^{d} \max\{a_i - \gamma, 0\} = \max\{0, a_*\}$, and the LHS is monotonic in γ.

We now seek to find suitable upper and lower bounds on γ. Define

$$L := -|z_*| + \min_i z_i; U := |z_*| + \max_i z_i.$$

Then,

$$\sum_i \max\{0, a_i - U\} = \sum_{i=1}^{d} \max\{0, q_i(x) + z_i - U\} \leq \sum_{i=1}^{d} \max\{0, q_i(x) - |z_*|\}$$
$$\leq \max\left\{0, \left(\sum_i q_i(x)\right) - |z_*|\right\} \qquad \because\ q_i(x) \geq 0$$
$$= \max\{0, q_*(x) - |z_*|\}$$

$$\leq \max\{0, a_*\} \implies \gamma \leq U.$$

and we have

$$\begin{aligned}
\sum_i \max\{0, a_i - L\} &= \sum_{i=1}^{d} \max\{0, q_i(x) + z_i - L\} \geq \sum_{i=1}^{d} \max\{0, q_i(x) + |z_*|\} \\
&= \sum_i (q_i(x) + |z_*|) && \because\ q_i(x) \geq 0 \\
&\geq \left(\sum_i q_i(x)\right) + |z_*| \\
&= q_*(x) + |z_*| \geq \max\{0, a_*\} \\
\implies \gamma &\geq L.
\end{aligned}$$

Now we find a bound on $\mathbb{E}[(q_i(\mathcal{M}_\varepsilon(x)) - q_i(x))^2]$ in terms of γ.

$$\begin{aligned}
\mathbb{E}[(q_i(\mathcal{M}'_\varepsilon(x)) - q_i(x))^2] &= \mathbb{E}[(\max\{0, q_i(x) + z_i - \gamma\} - q_i(x))^2] && \text{with the r.v.s being } z_i, \gamma \\
&\leq \mathbb{E}[((q_i(x) + z_i - \gamma) - q_i(x))^2] && \because\ q_i(x) \geq 0 \text{ and removing max} \\
& && \text{moves LHS away from } q_i(x) \\
&= \mathbb{E}[(z_i - \gamma)^2] \leq \mathbb{E}[(|z_i| + \max\{|L|, |U|\})^2] \\
&\leq \mathbb{E}\left[\left(|z_i| + |z_*| + \max_j |z_j|\right)^2\right] && \because\ z_j \text{ are symmetric around } 0 \\
&\leq \mathbb{E}\left[\left(|z_*| + 2\max_j |z_j|\right)^2\right] \\
&= \mathbb{E}[z_*^2] + 4\mathbb{E}[|z_*|]\mathbb{E}[\max_j |z_j|] + 4\mathbb{E}[(\max_j |z_j|)^2] \\
&\in O(\frac{1}{\varepsilon^2}\log^2(d)) && \text{by an aforementioned property of} \\
& && \text{Laplace noise.}
\end{aligned}$$

4. For this we proceed similarly as for the previous case, but adding noise from $N(0, \frac{1}{\rho})$ instead of $\text{Lap}\frac{2}{\varepsilon}$, and by a prior discussion, we are aware that this imparts ρ-zCDP to the output. Also the variance of the sum query is $\leq \frac{1}{\rho}$, and for the point queries, the only dissimilarity w.r.t above is is the last step where we use results for Gaussian noise from lemma 9.5, instead of those for Laplace noise, to get that $\mathbb{E}[(q_i(\mathcal{M}(x)) - q_i(x))^2] \in O(\frac{1}{\rho}\log(d)), \forall\, i \in [d]$.
5. We follow a procedure that is similar to that as before, only this time we draw noise from the double geometric distribution.

   For any $B \in \mathbb{Z}, B > 0$, the truncated double geometric distribution is given by $TDGeo(\varepsilon, B)$ is merely the result of clipping $DGeo(\varepsilon)$ at $\pm B$. More precisely, if $z' \sim TDGeo(\varepsilon, B)$, then

$$\Pr[z' = k] = \begin{cases} \frac{1}{1+e^{-\varepsilon}} e^{-\varepsilon B} \text{ if } k = \pm B \\ \frac{1-e^{-\varepsilon}}{1+e^{-\varepsilon}} e^{-\varepsilon|k|} \text{ for } k = -(B-1), \ldots, B-1 \end{cases}$$

   Now we proceed similarly as for the proof of the 3^{rd} case, with the difference that we draw

noise from $RDGeo(\frac{\varepsilon}{2}, B$ to answer each (point/sum) query. We seek to satisfy $(\frac{\varepsilon}{2}, \frac{\delta}{2})$-DP for each of the point queries and the sum query so upon composition of each point query with the sum query, we get (ε, δ)-DP (recall that point queries are disjoint).

Now we consider the points not in the boundary of $v + TDGeo$ or that of $v - 1 + TDGeo$: $\forall\, v \in \mathbb{Z}$ and $\forall\, k \in [v - B + 1, v + B - 2]$

$$e^{-\varepsilon/2} \leq \frac{\Pr[v + TDGeo(\frac{\varepsilon}{2}, B) = k]}{\Pr[v - 1 + TDGeo(\frac{\varepsilon}{2}, B) = k]} \leq e^{\varepsilon/2}.$$

Now for the boundary points,

$$\begin{aligned}\Pr[v + TDGeo(\frac{\varepsilon}{2}) \in \{v - B, v + B - 1, v + B\}] &= \Pr[DGeo(\varepsilon/2) \geq B - 1] + \Pr[DGeo(\varepsilon/2) \leq -B] \\ &= \frac{1}{1 + e^{-\varepsilon/2}} e^{-\varepsilon B/2} + \frac{1}{1 + e^{-\varepsilon/2}} e^{-\varepsilon(B-1)/2} \\ &=\leq 2e^{-\varepsilon(B-1)/2}.\end{aligned}$$

Setting this equal to $\frac{\delta}{2}$ and performing similar calculations for $\frac{\Pr[v-1+TDGeo(\frac{\varepsilon}{2},B)=k]}{\Pr[v+TDGeo(\frac{\varepsilon}{2},B)=k]}$, we see that if $B \geq \frac{2}{\varepsilon}\log(4/\delta) + 1$, then noise addition from $TDGeo(\frac{\varepsilon}{2}, B)$ satisfies (ε, δ)-DP.

Then utilising the same post-processing as in the proof of the third case, we get that the expected square error of the sum query's response on the privacy-protected microdata is $\leq Var(TDGeo(\varepsilon/2, B) \leq Var(DGeo(\varepsilon/2) \leq \frac{8}{\varepsilon^2}$.

As for point queries, we do follow the same process as in the third case, but since our noise distribution is different now, the last step changes; $\because$ the absolute value of the noise added is bounded by B, the expected squared error of the point queries is $\in O(B^2) = O(\frac{1}{\varepsilon^2}\log^2(\frac{1}{\delta}) + 1)$.

■

One easy conclusion here would be that the lower bounds are nearly tight in comparison to these upper bounds (except for a slight deviation for ρ-zCDP, where we have d^2 instead of $d^{2\gamma}$, with γ approaching 1).

The authors then proceed to describe some post-processing algorithms, a couple of baselines and a couple that they came up with as improvements to said baselines, however, they only publish experimental results based on the same and not the analysis of algorithms. That is auxiliary to the main point of the theoretical/mathematical discussion of this paper.

9.2 NeurIPS 2020; Zhu et al[27]

In their paper titled "Improving Sparse Vector Technique with Rényi Differential Privacy", Zhu and Wang seek to define tighter bounds on SVT and run it with lesser noise by using a variant of it that uses Gaussian noise instead of noise drawn from the Laplace distribution, thus making it an attractive option to implement in a practical sense. We shall discuss some salient results from this paper, whose proofs are present in the supplementary appendix to the same.

What they suggest is taking advantage of the more concentrated nature of Gaussian noise and the fact that as Gaussians occur naturally very often, it is convenient to use noise drawn from the Gaussian distribution for practical purposes. They admit that while their approach

is novel, in certain regimes, their improvement is only by a constant factor, but argue that when talking about practical, real life implementations, even an improvement by a constant factor matters.

They borrow some results from earlier work, including from [25] (which is a non-trivial result that the authors state wholly without proof), one of which is as follows.

Lemma 9.7. ***From DP to RDP***
Let $\mathcal{M}$ (a randomised algorithm) satisfy ε-DP, then $\mathcal{M}$ satisfies $\varepsilon(\alpha), \alpha$)-RDP with $\varepsilon(\alpha) = \frac{1}{\alpha-1}\log\left(\frac{\sinh(\alpha\varepsilon)-\sinh((\alpha-1)\varepsilon)}{\sinh(\varepsilon)}\right) \leq \frac{\alpha\varepsilon^2}{2}$.

Definition 9.8. Low Sensitivity Queries
Define the set of queries, $Q(\Delta) := \{q : \mathbb{N}^{|\chi|} \to \mathbb{R} : |q(x) - q(x') \leq \Delta| \, \forall \, x, x' \in \mathbb{N}^{|\chi|}, \|x - x'\|_1 \leq 1\}$. Then every query in $Q(\Delta)$ is called a low-sensitivity query.

They propose a generalised version of the SVT algorithm (which defines a family of SVT algorithms) which goes as follows.

Algorithm 9 Generalised SVT

Require: $D \in \mathbb{N}^{|\chi|}$, a sequence of adaptively chosen sensitivity Δ queries $\{q_i\}$ on $\mathbb{N}^{|\chi|}$, noise adding mechanisms $\mathcal{M}_\rho, \mathcal{M}_\nu$, a threshold T, a cutoff $c \in \mathbb{N}$, max-length $k_{\max} \in (0, \infty]$, option RESAMPLE

Sample $\bar{T} \sim \mathcal{M}_\rho(D, T)$, count=0
For $i = 1, 2, \ldots, k_{\max}$
  Sample $\tilde{q}_i \sum \mathcal{M}_\nu(D, q_i)$
  if $\tilde{q}_i \geq \hat{T}$ **then**
    Output $a_i = \top$, count = count + 1
    if RESAMPLE, $\hat{T} \sim \mathcal{M}_\rho(D, T)$
    if count$\geq c$, **abort.**
  else
    Output: $a_i = \perp$
  end if

The technical reason behind preferring the Gaussian mechanism over the Laplace mechanism is that the Laplace distribution is a heavy tailed distribution which needs the threshold to be $O(\log \frac{1}{\beta})$ to minimise the chance of false positives (w.r.t. the threshold) occurring, but (sub-)Gaussian tailed distributions merely require the threshold to be $(\sqrt{\log(\frac{1}{\beta})})$, and can enjoy tighter error/noise addition bounds with composition done using RDP, and as it is not as heavy tailed, it adds more concentrated noise to query responses.

Now we move on to analysing various variants of SVT with different choices of noise adding mechanisms $\mathcal{M}_\rho, \mathcal{M}_\nu$.

9.2.1 RDP for $c = 1$

Recall that the more general case of using SVT with $c > 1$ is treated as a composition of multiple SVTs with cutoff $c = 1$. x, x' are any two neighbouring databases.

Theorem 9.9. *Let K be a random variable indicating the stopping time, i.e. the number of outputs $\perp$ plus 1. Let $\mathcal{M}_\rho, \mathcal{M}_\nu$ be noise adding mechanisms. Assume that $\mathcal{M}_\rho$ satisfies $\varepsilon_\rho(\alpha)$-($(\alpha, \varepsilon_\rho(\alpha))$-)RDP for queries with sensitivity Δ and $\mathcal{M}_\nu$ satisfies $\varepsilon_\nu(\alpha)$-RDP for queries with sensitivity 2Δ. Then algorithm 9, denoted by $\mathcal{M}$ with $c = 1$ satisfies*

$$D_\alpha(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \varepsilon_\rho(\alpha) + \varepsilon_\nu(\alpha) + \frac{\log \sup_z \mathbb{E}[K|\rho = z]}{\alpha - 1}$$

$$[D_\alpha(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \frac{\alpha - \frac{\gamma-1}{\gamma}}{\alpha - 1}\varepsilon_\rho\left(\frac{\gamma}{\gamma - 1}\alpha\right) + \varepsilon_\nu(\alpha) + \frac{\log(\mathbb{E}_{z\sim p_\rho}[\mathbb{E}[K|\rho = z]^\gamma])}{\gamma(\alpha - 1)}$$

for all $\gamma > 1, 1 < \alpha < \infty$. More specifically when $\varepsilon_\rho(\infty) \leq \infty$, we get

$$D_\alpha(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \varepsilon_\rho(\alpha) + \varepsilon_\nu(\infty).$$

This is essentially a transfer theorem that bounds the RDP of the generalised SVT with respect to the RDP of its noise addition subroutines, $\mathcal{M}_\rho$ and $\mathcal{M}_\nu$. The authors discuss this in the context of some cases.

- **For pure DP:** Recall that when $\alpha \to \infty$, then RDP is essentially ε-DP, and in that case we can use any noise adding mechanism that is ε-differentially private.
- **Hybrid SVT:** Where $\mathcal{M}_\rho$ is the Gaussian mechanism but $\mathcal{M}_\nu$ are Laplace mechanisms.
- **Bounded-Length SVT:** i.e. when $k_{\max} < \infty$, which implies an RDP bound of the form $\varepsilon_\rho(\alpha) + \varepsilon_\nu(\alpha) + \log \frac{1+k_{\max}}{\alpha - 1}$, which can be expressed in terms of (ε, δ)-differential privacy using the conversion result from RDP to DP(theorem 8.8); if $\delta \leq \frac{1}{1+k_{\max}}$, then we get (ε, δ)-DP with $\varepsilon = \min_{\alpha>1} \varepsilon_\rho(\alpha) + \varepsilon_\nu(\alpha) + 2\log(1/\delta)/(\alpha - 1)$.
- **When $k_{\max} = \infty$:** This does not imply RDP in this case, because SVT can have unbounded length (in fact, even the expected length can be unbounded). If we use Gaussian-mechanism as a subroutine for SVT, we shall have a dependence on the sequence length is unavoidable. Also the second bound in the theorem above, which is only dependent on the moments of the conditional expectation seems to suggest that we can potentially obtain some meaningful RDP bounds for generalized SVT even if $k_{\max} = \infty$ in some cases.

Keeping this in mind, we define a family of queries that allow us to work with RDP even with unbounded sequence length for approximate DP.

Definition 9.10. Non-Negative, Low-Sensitivity Queries Model
The adversary is allowed to choose non-negative low-sensitivity queries $q_1, q_2, \ldots \in Q_+(\Delta)$ where

$$Q_+(\Delta) = \{q : \mathbb{N}^{|\chi|} \to \mathbb{R} : q \text{ is low sensitivity w.r.t. } \Delta, q(x) \geq 0, \forall\, x \in \mathbb{N}^{|\chi|}\}.$$

This class covers some important use cases of SVT, viz. Guess and Check, which is a subroutine of Private Multiplicative Weights, and for model-agnostic private learning.

Theorem 9.11. ***Gaussian SVT with non-negative queries***
Let algorithm 9 be initialised with the query set $Q_+(x), \mathcal{M}_\rho, \mathcal{M}_\nu$ be instances of the Gaussian mechanism with parameters σ_1, σ_2 respectively. Then $\forall\, T < \infty$, and $\gamma > 1$ such that

$\sigma_2 > \sqrt{\gamma+1}\sigma_1$, and with $c = 1$, algorithm 9 halts with K rounds satisfying

$$\mathbb{E}[\mathbb{E}[K|\rho = z]^\gamma] \leq 1 + (c_\gamma \sqrt{2\pi}\max\{\frac{T(1+\gamma}{\sigma_1}, 1\}^\gamma (1+\gamma)^{\frac{1}{2}} \exp(\frac{\gamma T^2}{2\sigma_1^2}).$$

And if $\sigma_2 \geq \sqrt{3}\sigma_1$, then it satisfies RDP with the parameter

$$\frac{\alpha\Delta^2}{\sigma_1^2} + \frac{2\alpha\Delta^2}{\sigma_2^2} + \frac{\log(1 + 2\sqrt{3}\pi(1+\frac{9T^2}{\sigma_1^2})\exp(\frac{\gamma T^2}{\sigma_1^2})}{2(\alpha-1)}.$$

If we choose $T = \sqrt{2(\sigma_1^2+\sigma_2^2)\log 1/\rho} = \sqrt{8\sigma_1^2\log(1/\rho)}$ where we intend to bound the Type I error (false positive rate) by ρ, which is often seen when working with SVT for statistical applications. Then simplifying the aforementioned bound using the fact that $\log(1+x) \leq x$, given that ρ is sufficiently small, we have an RDP bound for Gaussian SVT as follows

$$\frac{5\alpha\Delta^2}{3\sigma_1^2} + \frac{8\log(1/\rho) + \log(4\sqrt{3}\pi(1+72\log(1/\rho)))}{2(\alpha-1)} \leq \frac{5\alpha\Delta^2}{3\sigma_1^2} + \frac{5\log(1/(\rho)}{\alpha-1}$$

Which allows us to get nearly the same approximate/(ε,δ)-DP bound for Gaussian SVT as if working with a Gaussian mechanism's RDP bound given that δ is such that $\log(1/\delta)$ is greater than either $\log(k_{\max})$ when $k_{\max} < \infty$ or is $O(\log(1/\rho))$ in the context of non-negative queries.

Hereon, the authors denote a data-independent upper bound of $(\mathbb{E}[K|\rho]^\gamma])^{\frac{1}{\gamma}}$ by k_γ. Then we have that $k_\infty = k_{\max} = k_1$, and therefore the first two bounds in the above theorem are unified.

We define γ^* such that for a given γ, $\frac{1}{\gamma^*} + \frac{1}{\gamma} = 1$.

9.2.2 Generalised SVT with $c > 1$

Theorem 9.12. ***RDP for Generalised SVT with length cap k_{max} and $c > 1$***
Generalised SVT (algorithm 9) with cutoff $c > 1$ and maximum length k_{max} satisfies

$$D_\alpha(\mathcal{M}(x)\|\mathcal{M}(x')) \leq \varepsilon_\rho(\alpha) + c\varepsilon_v(\alpha) + \frac{1 + \log\sum_{k=0}^{c}\binom{k_{max}}{k}}{\alpha-1}.$$

Note that $\sum_{k=0}^{c}\binom{k_{\max}}{k}$ is simply the cardinality of the set of possible outputs, which is the set of vectors with boolean values of length $k_{\max}$ with at most c many $\perp$s.

When $\mathcal{M}_\nu$ and $\mathcal{M}_\rho$ are Gaussian mechanisms, the above theorem and theorem 8.8 imply for the Gaussian SVT an (ε,δ)-DP bound with

$$\varepsilon(\delta) \leq \frac{\Delta^2}{2\sigma_1^2} + \frac{2c\Delta^2}{\sigma_2^2} + \sqrt{2\left(\frac{\Delta^2}{2\sigma_1^2} + \frac{2c\Delta^2}{\sigma_2^2}\right)\left(\log(\frac{1}{\delta}) + \log c\binom{k_{\max}}{c}\right)}.$$

Which means that noise scales as $O(\sqrt{c})$ when $\delta \leq c^{-1}\binom{k_{\max}}{c}^{-1}$, which is a bound similar to that for vanilla SVT. Though needing to take $\delta < k_{\max}^{-c}$ is limiting.

For (ε,δ)-DP composition: Here the authors introduce a meta-algorithm to Generalised SVT based on the premise that if we use advanced composition on (ε,δ)-DP, we get a bound with $\sqrt{c}$ scaling for a broader set of parameters than those that we are already considering for SVT.

This stage-wise meta-algorithm runs Generalised SVT as a subroutine and resamples the noise on the threshold, ρ, after every c' rounds, with each round having a length cap (on the number of queries per round) $k'_{\max}$ chosen in each round. Here we have to choose $c', k'_{\max}$ in each round in such a way that for the parameters c, δ, we end up with that $\log\left(c'\binom{k'_{\max}}{c'}\right)$ is comparable to $\log\frac{1}{\delta}$.

Algorithm 10 Stage-wise Generalised SVT

Require: $D \in \mathbb{N}^{|\chi|}$, a sequence of adaptively chosen sensitivity Δ queries $\{q_i\}$ on $\mathbb{N}^{|\chi|}$, noise adding mechanisms $\mathcal{M}_\rho, \mathcal{M}_\nu$, a threshold T, total cutoff $c \in \mathbb{N}$, per-stage cutoff c', per-stage max-length $k'_{\max} \in (0, \infty]$, option RESAMPLE

1: Initialise output vector $v \leftarrow \phi$
2: **for** $\ell = 1, 2, 3, \ldots, \lceil c/c' \rceil$ **do**
3: **if** $\ell = \lceil c/c' \rceil$ **then**
4: $\tilde{c} \leftarrow c - c'(\lceil c/c' \rceil - 1)$
5: **else**
6: $\tilde{c} = c'$
7: **end if**
8: Invoke algorithm 9 with parameters $D, T, \mathcal{M}_\rho, \mathcal{M}_\nu, \tilde{c}, k'_{\max}$, RESAMPLE and the query which is currently at the front of the adaptively chosen stream of queries.
9: Append the new output vector to v
10: **end for**
11: **Output:** v.

Theorem 9.13. ***Stage-wise Length-capped Gaussian SVT for*** $Q(\Delta)$
Let $\delta' \in (0, 1)$, and let $\mathcal{M}$ be an instance of algorithm 10 invoked with cutoff c', length cap k'_{max}, with no resampling of noise (i.e. RESAMPLE=False); and let $\mathcal{M}_\rho, \mathcal{M}_\nu$ be chosen as Gaussian mechanisms with parameters σ_1, σ_2 respectively such that $\sigma_2 = 2\sigma_1$ and $\sigma_1 \geq 8\Delta\sqrt{c\log(1/\delta')}$. If we choose $c' \leq c$ such that $c'\binom{k'_{max}}{c'} \leq \frac{1}{\delta'}$, then $\mathcal{M}$ satisfies, for some $\tilde{\delta} > 0$, $(\varepsilon, \tilde{\delta} + \frac{c}{c'}\delta')$-DP with $\varepsilon \in O\left(\sqrt{\frac{c\Delta^2}{\sigma_1^2}\log(\frac{1}{\delta'})\log(\frac{1}{\tilde{\delta}})}\right)$.

Note that any appropriate positive real value may be assigned to $\tilde{\delta}$

Theorem 9.14. ***Adaptive Stage-wise Gaussian SVT for*** $Q_+(\Delta)$
Let $\mathcal{M}$ be an instance of algorithm 10 invoked with the same parameters as for Stage-wise Length-capped Gaussian SVT, but this time RESAMPLE is set to be True and $k_{max} = \infty$. Then $\forall\, c', \gamma$ such that $k_\gamma^{c'} \leq \frac{1}{\delta'}$, and for some $\tilde{\delta} > 0$, $\mathcal{M}$ is $(\varepsilon, \tilde{\delta} + \frac{c}{c'}\delta')$-differentially private with $\varepsilon = O\left(\sqrt{\frac{c\Delta^2}{\sigma_1^2}\log(1/\delta')\log(1/\tilde{\delta})}\right)$ for all adaptively chosen sequences of queries in Q_+.

Note that enabling RESAMPLE (i.e. setting it to True) makes algorithm 10 practically identical to algorithm 9 for all $c' \geq 1$, and then we can optimise privacy bounds by tweaking the values of c', δ', and to determine which part of the composition process relies on the composition of RDP, and which part relies on (ε, δ)-DP, as c' in a way dictates the length of a run of each round, and δ' is involved in (ε, δ)-DP style composition amongst rounds. The best choice, as the authors state, would be to pick c' to be as large as possible to maximise the privacy savings that one would get due to RDP composition, while not compromising $O(\sqrt{C})$ advanced composition when c is large.

To see that this bound is pretty good, consider the probably unachievable situation in which Generalised SVT ends up satisfying RDP with parameter $\varepsilon_\rho(\alpha) + c\varepsilon_\nu(\alpha)$, which would give us $(\varepsilon, \tilde{\delta} + \frac{c}{c'}\delta)$-DP with $\varepsilon = O\left(\sqrt{\frac{c\Delta^2}{\sigma_1^2}\log(\frac{1}{\delta})}\right)$, so we see that the stage-wise and adaptive algorithms only have a bound which is worse than this best case scenario by a factor of $\sqrt{\log(1/\delta)}$, and not to forget, with some restrictions on δ.

References

[1] D. Desfontaines, *K-anonymity, the parent of all privacy definitions - ted is writing things*, Aug. 2017. [Online]. Available: `https://desfontain.es/privacy/k-anonymity.html`.

[2] P. Samarati and L. Sweeney, "Protecting privacy when disclosing information: K-anonymity and its enforcement through generalization and suppression," Tech. Rep., 1998.

[3] S. R. Ganta, S. P. Kasiviswanathan, and A. Smith, *Composition attacks and auxiliary information in data privacy*, 2008. arXiv: `0803.0032 [cs.DB]`.

[4] A. Narayanan and V. Shmatikov, "How to break anonymity of the netflix prize dataset," *CoRR*, vol. abs/cs/0610105, 2006. arXiv: `cs/0610105`. [Online]. Available: `http://arxiv.org/abs/cs/0610105`.

[5] S. Garfinkel, J. M. Abowd, and C. Martindale, "Understanding database reconstruction attacks on public data: These attacks on statistical databases are no longer a theoretical danger.," *Queue*, vol. 16, no. 5, pp. 28–53, Oct. 2018, ISSN: 1542-7730. DOI: `10.1145/3291276.3295691`. [Online]. Available: `https://doi.org/10.1145/3291276.3295691`.

[6] P. Sengupta, S. Paul, and S. Mishra, "Learning with differential privacy," *CoRR*, vol. abs/2006.05609, 2020. arXiv: `2006.05609`. [Online]. Available: `https://arxiv.org/abs/2006.05609`.

[7] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014, ISSN: 1551-305X. DOI: `10.1561/0400000042`. [Online]. Available: `http://dx.doi.org/10.1561/0400000042`.

[8] J. M. Abowd, L. Alvisi, C. Dwork, S. Kannan, A. Machanavajjhala, and J. P. Reiter, "Privacy-preserving data analysis for the federal statistical agencies," *CoRR*, vol. abs/1701.00752, 2017. arXiv: `1701.00752`. [Online]. Available: `http://arxiv.org/abs/1701.00752`.

[9] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965, ISSN: 01621459. [Online]. Available: `http://www.jstor.org/stable/2283137`.

[10] I. Mironov, "Renyi differential privacy," *CoRR*, vol. abs/1702.07476, 2017. arXiv: 1702.07476. [Online]. Available: http://arxiv.org/abs/1702.07476.

[11] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advances in Cryptology - EUROCRYPT 2006, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28 - June 1, 2006, Proceedings*, S. Vaudenay, Ed., ser. Lecture Notes in Computer Science, vol. 4004, Springer, 2006, pp. 486–503. DOI: 10.1007/11761679_29. [Online]. Available: https://doi.org/10.1007/11761679%5C_29.

[12] M. Pycia and M. U. Ünver, "Decomposing random mechanisms," *Journal of Mathematical Economics*, vol. 61, pp. 21–33, 2015, ISSN: 0304-4068. DOI: https://doi.org/10.1016/j.jmateco.2015.06.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304406815000579.

[13] K. Chaudhuri, *Differentially Private Machine Learning: Theory, Algorithms, and Applications*, Sep. 2020. [Online]. Available: https://www.youtube.com/watch?v=kwP4EsFMsu8.

[14] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, 2007, pp. 94–103. DOI: 10.1109/FOCS.2007.66.

[15] G. Kamath, *Lecture Notes for CS 860 - Algorithms for Private Data Analysis, Fall 2020, University of Waterloo*, 2020.

[16] C. Dwork, G. N. Rothblum, and S. Vadhan, "Boosting and differential privacy," in *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, 2010, pp. 51–60. DOI: 10.1109/FOCS.2010.12.

[17] A. Nikolov, *Lecture Notes for CSC 2412 Algorithms for Private Data Analysis, Fall 2019, University of Toronto*, 2020.

[18] R. Cummings, *Lecture Notes for ISYE/CS 8803 Foundations of Data Privacy, Georgia Institute of Technology*, 2017.

[19] M. E. Gursoy, A. Tamersoy, S. Truex, W. Wei, and L. Liu, "Secure and utility-aware data collection with condensed local differential privacy," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 5, pp. 2365–2378, 2021. DOI: 10.1109/TDSC.2019.2949041.

[20] Ú. Erlingsson, A. Korolova, and V. Pihur, "RAPPOR: randomized aggregatable privacy-preserving ordinal response," *CoRR*, vol. abs/1407.6981, 2014. arXiv: 1407.6981. [Online]. Available: http://arxiv.org/abs/1407.6981.

[21] B. Ding, J. Kulkarni, and S. Yekhanin, "Collecting telemetry data privately," *CoRR*, vol. abs/1712.01524, 2017. arXiv: 1712.01524. [Online]. Available: http://arxiv.org/abs/1712.01524.

[22] J. Murtagh and S. P. Vadhan, "The complexity of computing the optimal composition of differential privacy," *CoRR*, vol. abs/1507.03113, 2015. arXiv: 1507.03113. [Online]. Available: http://arxiv.org/abs/1507.03113.

[23] J. P. Near and C. Abuah, *Programming Differential Privacy*. 2021, vol. 1. [Online]. Available: https://uvm-plaid.github.io/programming-dp/.

[24] J. Geumlek, S. Song, and K. Chaudhuri, "Rényi differential privacy mechanisms for posterior sampling," *CoRR*, vol. abs/1710.00892, 2017. arXiv: 1710.00892. [Online]. Available: http://arxiv.org/abs/1710.00892.

[25] M. Bun and T. Steinke, *Concentrated differential privacy: Simplifications, extensions, and lower bounds*, 2016. arXiv: 1605.02065 [cs.CR].

[26] J. Abowd, R. Ashmead, R. Cumings-Menon, S. Garfinkel, D. Kifer, P. Leclerc, W. Sexton, A. Simpson, C. Task, and P. Zhuravlev, *An uncertainty principle is a price of privacy-preserving microdata*, 2021. arXiv: 2110.13239 [cs.CR].

[27] Y. Zhu and Y.-X. Wang, "Improving sparse vector technique with renyi differential privacy," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 20 249–20 258. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/e9bf14a419d77534105016f5ec122d62-Paper.pdf.