Crowd-Blending Privacy*

Johannes Gehrke†, Michael Hay‡, Edward Lui, and Rafael Pass§
Department of Computer Science, Cornell University
{johannes,mhay,luied,rafael}@cs.cornell.edu

August 2, 2012

Abstract

We introduce a new definition of privacy called *crowd-blending privacy* that strictly relaxes the notion of differential privacy. Roughly speaking, k-crowd blending private sanitization of a database requires that each individual i in the database "blends" with k other individuals j in the database, in the sense that the output of the sanitizer is "indistinguishable" if i's data is replaced by j's.

We demonstrate crowd-blending private mechanisms for histograms and for releasing synthetic data points, achieving strictly better utility than what is possible using differentially private mechanisms. Additionally, we demonstrate that if a crowd-blending private mechanism is combined with a "pre-sampling" step, where the individuals in the database are randomly drawn from some underlying population (as is often the case during data collection), then the combined mechanism satisfies not only differential privacy, but also the stronger notion of zero-knowledge privacy. This holds even if the pre-sampling is slightly biased and an adversary knows whether certain individuals were sampled or not. Taken together, our results yield a practical approach for collecting and privately releasing data while ensuring higher utility than previous approaches.

1 Introduction

Data privacy is a fundamental problem in today's information age. Large amounts of data are collected from people by government agencies, hospitals, social networking systems, and other organizations, and are stored in databases. There are huge social benefits in analyzing this data, and in many situations, these organizations would like to release the data in some form for people to analyze. However, it is important to protect the privacy of the people that contributed their data; organizations need to make sure that sensitive information about individuals is not leaked to the people analyzing the data.

*A preliminary version of this paper appeared in the 32nd International Cryptology Conference (CRYPTO 2012).

†Gehrke's work on this material was supported by the NSF under Grant IIS-1012593. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

‡Hay's work was supported by the Computing Innovation Fellows Project (http://cifellows.org/), funded by the Computing Research Association/Computing Community Consortium through NSF Grant 1019343.

§Pass is supported in part by a Alfred P. Sloan Fellowship, Microsoft New Faculty Fellowship, NSF CAREER Award CCF-0746990, AFOSR YIP Award FA9550-10-1-0093, and DARPA and AFRL under contract FA8750-11-2-0211. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the US government.

Many privacy definitions and schemes for releasing data have been proposed in the past (see [CKLM09] and [FWCY10] for surveys). However, many of them have been shown to be insufficient due to realistic attacks on such schemes (e.g., see [Kif09]). The notion of *differential privacy* [DMNS06, Dwo06], however, has remained strong and resilient to these attacks. Differential privacy requires that when one person's data is added or removed from the database, the output distribution of the database access mechanism changes very little (by at most an ϵ amount, where a specific notion of closeness of distributions is used). Differential privacy has quickly become the standard definition of privacy, and mechanisms for releasing a variety of functions (including histogram queries, principal component analysis, learning, and many more; see [Dwo09, Dwo08] for a survey) have been developed.

One way to interpret the notion of differential privacy is that an attacker does not learn more about an individual i than what can be deduced from the data of everyone else in the database (see the appendix of [DMNS06]). In the context of e.g., social networks, where the data of an individual may be strongly correlated with the data of his/her friends, such a notion may not always provide sufficient privacy guarantees. To address this issue, an even stronger privacy definition, *zero-knowledge privacy*, was introduced in [GLP11]. Roughly speaking, zero-knowledge privacy requires that whatever an adversary learns about an individual i can be "simulated" given just some "aggregate" information about the *remaining* individuals in the database; for instance, this aggregate information could be k random samples of the remaining individuals in the database. If the aggregate information contains all individuals (excluding i), zero-knowledge privacy collapses down to differential privacy, but for more restrictive classes of aggregate information (such as k random samples, where k is smaller than the number of individual in the database) zero-knowledge privacy is strictly stronger, and provides stronger privacy guarantees in contexts where there is correlation between individuals.

Privacy from Random Sampling of Data. Both differential privacy and zero-knowledge privacy provide strong privacy guarantees. However, for certain tasks, mechanisms satisfying these privacy definitions have to add a lot of "noise", thus lowering the utility of the released data. Also, many of these mechanisms run in exponential time (e.g., [DRV10, BLR08]), so efficiency is also an issue. This leaves open the question of whether there exists a practical approach to sanitizing data, without harming utility too much.

One approach for circumventing the above-mentioned issues is to rely on the fact that in many cases of interest, the data to be sanitized has been collected via *random sampling* from some underlying population. Intuitively, this initial random sampling already provides some basic privacy guarantees, and may thus help us in decreasing the amount of noise added during sanitization. Indeed, there are several results in the literature indicating that random sampling helps in providing privacy: In [CM06] the authors quantify the level of the privacy that may be obtained from just random sampling of data (without any further sanitization); in [NRS07] the authors consider a certain type of "sample-and-aggregate" mechanism for achieving differential privacy (but the sampling technique here is more elaborate than just random sampling from a population); a result in [KLN$^+$08] shows that random pre-sampling can be used to amplify the privacy level of a differentially private mechanism; finally, in a manuscript [LQS11], the authors demonstrate that a random pre-sampling step applied to a particular mechanism leads to a differentially private mechanism.

In this paper, we continue the investigation of using random sampling as a means to achieve privacy. In particular, our goal is to provide a *general* definition of privacy that allows us to achieve both differential and zero-knowledge privacy in situations where the data is collected using random sampling from some population. In order to be realistic, we allow the random sampling during

data collection to be *biased*, and an adversary may even know whether certain individuals were sampled or not. (Although the mechanisms in the earlier papers rely on random sampling, the random sampling is usually thought of as being part of the sanitization procedure and thus the mechanisms are only analyzed under the assumption that the sampling has been done "ideally".) Additionally, we will require that the privacy notion is meaningful in its own right, also without any pre-sampling; we believe this requirement is crucial for guaranteeing a strong fall-back guarantee even in case the result of the pre-sampling is leaked (and thus the attacker knows exactly who was sampled).

1.1 Towards a Weaker Notion of Privacy

We aim to develop a new privacy definition that allows us to design mechanisms that have greater utility or efficiency than differentially private mechanisms, but still provide a meaningful notion of privacy; furthermore, we want mechanisms satisfying the new definition to achieve differential and zero-knowledge privacy when the underlying data was collected via biased random sampling from some population. To this end, we begin by reconsidering some older notions of privacy.

k-Anonymity and Blending in a Crowd. k-anonymity [Swe02] is a privacy definition specifically for releasing data tables, where a data table is simply a table of records (rows), each of which has values for the attributes (columns) of the table. Roughly speaking, a released data table satisfies *k-anonymity* if every record in the table is the same as $k-1$ other records in the table with respect to certain "identifying" attributes (chosen beforehand). k-anonymity imposes constraints on the syntax of the released data table, but does not consider the way the released data table was computed from the underlying database; this issue has led to several practical attacks against the notion of k-anonymity (e.g., see [WFWP07, ZJB07]). k-anonymity can be viewed as being based on the intuition of *"blending in a crowd"*, since the records in the released output are required to "blend" with other records. Intuitively, in many cases, if an individual blends in a crowd of many people in the database, then the individual's privacy is sufficiently protected. However, as demonstrated by known attacks, k-anonymity does not properly capture this intuition as it does not impose any restrictions on the algorithm/mechanism used to generate the released output. Indeed, one of the key insights behind the notion of differential privacy was that privacy should be a property of the sanitization mechanism and not just the output of it.

Relying on this insight, we aim to develop a privacy notion that captures what it means for a mechanism to guarantee that individuals "blend in a crowd". (Another definition partly based on the intuition of blending in a crowd is (c,t)-*isolation* [CDM$^+$05], which requires adversaries to be unable to isolate an individual, represented by a data point in $\mathbb{R}^d$, by roughly determining the individual's location in $\mathbb{R}^d$; we formalize the intuition of blending in a crowd in a very different way.)

Crowd-Blending Privacy – A New Privacy Definition. Let us now turn to describing our new privacy definition, which we call *crowd-blending privacy*. We say that an individual *blends* with another individual with respect to a mechanism *San* if the two individuals are *indistinguishable by the mechanism San*, i.e., whenever we have a database containing either one or both of the individuals, we can replace one of the individual's data with the other individual's data, and the mechanism's output distribution remains essentially the same. We say that an individual t *blends in a crowd of k people in the database D with respect to the mechanism San* if there exist at least $k-1$ other individuals in the database D that blend with individual t with respect to *San*. The intuition behind this notion is that if an individual t blends in a crowd of k people in the database,

then the mechanism essentially does not release any information about individual t beyond the general characteristics of the crowd of k people; in particular, the mechanism does not release any personal information that is specific to individual t and no one else.

Roughly speaking, we say that a mechanism *San* is *crowd-blending private* if the following property holds: For every database and every individual in the database, either the individual *blends in a crowd of k people in the database with respect to San*, or the mechanism *San essentially ignores the individual's data*.

We do not claim that crowd-blending privacy provides sufficiently strong privacy protection in *all* scenarios: the key weakening with respect to differential privacy is that an attacker who knows the data of everyone in an individual i's crowd (except i) may learn information about individual i, as long as this information is "general" in the sense that it applies to the entire crowd. For instance, if the attacker knows everyone in the crowd of individual i, it may deduce that i has, say, three children, as long as everyone in i's crowd has three children. Although to some extent, this may be viewed as a privacy violation (that would not be allowed by the notion of differential privacy), we would argue that the attribute leaked about individual i is "non-sensitive" as it is shared by a sufficiently large crowd. Thus, we view this weakening as desirable in many contexts as it allows us to trade privacy of "non-sensitive information" for improved utility.

A potentially more serious deficiency of the definition is that (in contrast to differential and zero-knowledge privacy) crowd-blending privacy is not closed under composition: San_1 and San_2 may both be crowd-blending private, but the crowds for an individual with respect to San_1 and San_2 could be essentially disjoint, making the individual's crowd for the combination of San_1 and San_2 very small. Although we view composition as an important property of a privacy definition, our goal here is to study the weakest possible "meaningful" definition of "stand-alone" privacy that when combined with pre-sampling leads to strong privacy notions (such as differential and zero-knowledge privacy) that themselves are closed under composition.

1.2 New Database Mechanisms

As it turns out, achieving crowd-blending privacy is significantly easier than achieving differential privacy, and crowd-blending private mechanisms may yield significantly higher utility than differentially private ones.

Privately Releasing Histograms with No Noise for Sufficiently Large Counts. We show that we can release histograms with crowd-blending privacy where no noise is added to bins with a sufficiently large count (and only a small amount of noise is added to bins with a small count). Intuitively, individuals in the same bin blend with each other; thus, the individuals that belong to a bin with a sufficiently large count already blend in a crowd, so no noise needs to be added to the bin. It is easy to see that it is impossible to release the exact count of a bin in a histogram while satisfying differential privacy or zero-knowledge privacy. Using crowd-blending privacy, we can overcome this limitation (for bins with a sufficiently large count) and achieve better utility. These results can be found in Section 3.1.

Privately Releasing Synthetic Data Points in $\mathbb{R}^d$ for Computing Smooth Functions. Given a class $\mathcal{C}$ of counting queries whose size is not too large, it is shown in [BLR08] how to release a synthetic database for approximating all the queries in $\mathcal{C}$ simultaneously while satisfying differential privacy; however, the mechanism is not necessarily efficient. It is known that it is impossible (assuming the existence of one-way functions) to *efficiently* and privately release a

synthetic database for approximating certain classes of counting queries, such as the class of all 2-way marginals (see [UV11, DNR+09]). However, these query functions are non-smooth in the sense that even slightly changing one row of the input database can affect the output of the query functions quite a lot. Here, we focus on efficiently and privately releasing synthetic data for approximating *all* "smooth" functions $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$.

Roughly speaking, a function $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$ is *smooth* if the value of g does not change much when we perturb the data points of the input slightly. We show that we can *efficiently* release synthetic data points in $\mathbb{R}^d$ for approximating *all* smooth functions simultaneously while satisfying crowd-blending privacy. On the other hand, we show that there are smooth functions that cannot even be approximated with non-trivial utility from any synthetic data that has been released with differential privacy (even if the differentially private mechanism is *inefficient*). These results can be found in Section 4.

1.3 From Crowd-Blending Privacy to Zero-Knowledge Privacy

Our main technical result shows that if we combine a crowd-blending private mechanism with a natural pre-sampling step, then the combined algorithm satisfies zero-knowledge privacy (and thus differential privacy as well). We envision the pre-sampling step as being part of the data collection process, where individuals in some population are sampled and asked for their data. Thus, if data is collected using random sampling of individuals from some population, and next sanitized using a crowd-blending private mechanism, then the resulting process ensures zero-knowledge privacy.

We first prove our main theorem for the case where the pre-sampling step samples each individual in the population with probability p independently. In reality, the sampling performed during data collection may be slightly biased or done slightly incorrectly, and an adversary may know whether certain individuals were sampled or not. Thus, we next extend our main theorem to also handle the case where the sampling probability is not necessarily the same for everybody, but the sampling is still "robust" in the sense that most individuals are sampled independently with probability in between p and p' (this probability can even depend on the individual's data), where p and p' are relatively close to one another, while the remaining individuals are sampled independently with arbitrary probability. As a result, we have that in scenarios where data has been collected using any robust sampling, we may release data which both ensures strong utility guarantees and satisfies very strong notions of privacy (i.e., zero-knowledge privacy and differential privacy). In particular, this methodology can allow us to achieve zero-knowledge privacy and differential privacy while guaranteeing utility that is better than that of previous methods (such as for releasing histograms or synthetic data points as described above). Our main theorems can be found in Section 5.

It is worthwhile to note that the particular mechanism considered in [LQS11] (which in fact is a particular mechanism for achieving k-anonymity) can easily be shown to satisfy crowd-blending privacy; as a result, their main result can be derived (and significantly strengthened) as a corollary of our main theorem.[1] (See Section 3.1 and 5 for more details.)

[1] As mentioned, none of the earlier work using random pre-sampling focus on the case when the sampling is biased; furthermore, even for the case of perfect random sampling, the authors of [LQS11] were not able to provide a closed form expression of the level of differential privacy achieved by their mechanism, whereas a closed form expression can be directly obtained by applying our main theorem.

2 Preliminaries and Existing Privacy Definitions

A *database* is a finite *multiset* of data values, where a data value is simply an element of some fixed set X, which we refer to as the *data universe*. Each data value in a database belongs to an individual, so we also refer to a data value in a database as an *individual* in the database. For convenience, we will sometimes order the individuals in a database in an arbitrary way and think of the database as an element of X^*, i.e., a vector with components in X (the components are referred to as the *rows* of the database). Given a database D and a data value $v \in X$, let (D, v) denote the database $D \cup \{v\}$. A (database) *mechanism* is simply an algorithm that operates on databases.

Given $\epsilon, \delta \geq 0$ and two random variables (or distributions) Z and Z', we shall write $Z \approx_{\epsilon,\delta} Z'$ to mean that for every $Y \subseteq \mathrm{Supp}(Z) \cup \mathrm{Supp}(Z')$ we have

$$\Pr[Z \in Y] \leq e^{\epsilon} \Pr[Z' \in Y] + \delta$$

and

$$\Pr[Z' \in Y] \leq e^{\epsilon} \Pr[Z \in Y] + \delta.$$

We shall also write $Z \approx_{\epsilon} Z'$ to mean $Z \approx_{\epsilon,0} Z'$. Differential privacy (see [DMNS06, Dwo06]) can now be defined in the following manner:

Definition 1 ([DMNS06, Dwo06])**.** A mechanism San is said to be **ϵ-differentially private** if for every pair of databases D and D' differing in only one data value, we have $San(D) \approx_{\epsilon} San(D')$.

There are two definitions in the literature for "a pair of databases D and D' differing in only one data value", leading to two slightly different definitions of differential privacy. In one definition, it is required that D contains D' and has exactly one more data value than D'. In the other definition, it is required that $|D| = |D'|$, $|D \setminus D'| = 1$, and $|D' \setminus D| = 1$. Intuitively, differential privacy protects the privacy of an individual t by requiring the output distribution of the mechanism to be essentially the same regardless of whether individual t's data is included in the database or not (or regardless of what data value individual t has).

We now begin describing zero-knowledge privacy, which is a privacy definition introduced in [GLP11] that is strictly stronger than differential privacy. In the definition of zero-knowledge privacy, *adversaries* and *simulators* are simply randomized algorithms that play certain roles in the definition. Let San be any mechanism. For any database D, any adversary A, and any auxiliary information $z \in \{0,1\}^*$, let $Out_A(A(z) \leftrightarrow San(D))$ denote the output of A on input z after interacting with the mechanism San operating on the database D. San can be interactive or non-interactive. If San is non-interactive, then $San(D)$ simply sends its output (e.g., sanitized data) to A and then halts immediately.

Let agg be any class of randomized algorithms. agg is normally a class of randomized aggregation functions that provide aggregate information to simulators, as described in the introduction.

Definition 2 ([GLP11])**.** A mechanism San is said to be **(ϵ, δ)-zero-knowledge private with respect to** agg if there exists a $T \in agg$ such that for every adversary A, there exists a simulator S such that for every database D, every individual $t \in D$, and every auxiliary information $z \in \{0,1\}^*$, we have

$$Out_A(A(z) \leftrightarrow San(D)) \approx_{\epsilon,\delta} S(z, T(D \setminus \{t\}), |D|).$$

Intuitively, zero-knowledge privacy requires that whatever an adversary can compute about individual t by accessing (i.e., interacting with) the mechanism can also be essentially computed without accessing the mechanism but with certain aggregate information about the remaining individuals; this aggregate information is provided by an algorithm in agg. The adversary in the latter scenario is represented by the simulator S. This ensures that the adversary essentially does not learn any additional information about individual t beyond the aggregate information provided by an algorithm in agg on the remaining individuals.

agg is normally some class of randomized aggregation functions, such as the class of all functions T that draws r random samples from the input database and performs any computation (e.g., computes the average or simply outputs the samples) on the r random samples (note that in the definition, T is applied to $D \setminus \{t\}$ instead of D so that the aggregate information from T does not depend directly on individual t's data). Zero-knowledge privacy with respect to this class of aggregation functions ensures that an adversary essentially does not learn anything more about an individual beyond some "r random sample aggregate information" of the other individuals. One can also consider zero-knowledge privacy with respect to other classes of aggregation functions, such as the class of (randomized) functions that first sample each row of the input database with probability p (or in between p and p') independently and then performs any computation on the samples. We will actually use such classes of aggregation functions when we prove our main theorems later. It can be easily shown that zero-knowledge privacy (with respect to any class agg) implies differential privacy (see [GLP11]).

In the original definition of zero-knowledge privacy in [GLP11], T operates on $(D \setminus \{t\}, \perp)$ instead of $D \setminus \{t\}$, where $\perp$ is any arbitrary element of the data universe X. The main point is that the database that T is applied to does not include individual t's data value (otherwise, a lot of information about individual t could possibly be leaked). Thus, using $T(D \setminus \{t\})$ in the definition also makes sense, and we choose to use $T(D \setminus \{t\})$ in this paper for convenience.[2] This version of zero-knowledge privacy still implies differential privacy (essentially the same "hybrid/transitivity" proof from [GLP11] works).

3 Crowd-Blending Privacy – A New Privacy Definition

We now begin to formally define our new privacy definition. Given $t, t' \in X$, $\epsilon \geq 0$, and a mechanism San, we say that t and t' are *ϵ-indistinguishable* by San, denoted $t \approx_{\epsilon,San} t'$, if $San(D, t) \approx_\epsilon San(D, t')$ for every database D. Intuitively, t and t' are indistinguishable by San if for any database containing t, we can replace the t by t' and the output distribution of San remains essentially the same. Usually, t and t' are the data values of two individuals, and if t and t' are indistinguishable by San, then this roughly means that San cannot distinguish these two individuals regardless of who else is in the database. If t and t' are ϵ-indistinguishable by San, we also loosely say that t *blends* with t' (with respect to San). We now describe what it means for an individual to blend in a crowd of people in the database (with respect to a mechanism).

Definition 3. Let D be any database. An individual $t \in D$ **ϵ-blends in a crowd of k people in D with respect to the mechanism** San if $|\{t' \in D : t' \approx_{\epsilon,San} t\}| \geq k$.

In the above definition, $\{t' \in D : t' \approx_{\epsilon,San} t\}$ should be regarded as a multiset. When the mechanism San is clear from context, we shall simply omit the "with respect to the mechanism San". Intuitively, an individual $t \in D$ blends in a crowd of k people in D if t is indistinguishable

[2] Even if we used $T(D \setminus \{t\}, \perp)$ instead of $T(D \setminus \{t\})$, our results would still hold with only minor modifications and slight differences in privacy parameters. Recall that differential privacy also has two versions of its definition.

by San from at least $k-1$ other individuals in D. Note that by the definition of two individuals being indistinguishable by San, $t \in D$ must be indistinguishable by San from each of these $k-1$ other individuals *regardless of what the database is*, as opposed to only when the database is D. (A weaker requirement would be that for each of these $k-1$ other individuals t', t and t' only need to be "indistinguishable by San with respect to D", i.e., if we take D and replace t by t' or vice versa, the output distributions of San on D and the modified D are essentially the same; we leave investigating this and other possible weaker requirements for future work.) We are now ready to state our new privacy definition.

Definition 4 (Crowd-blending privacy)**.** A mechanism San is (k, ϵ)**-crowd-blending private** if for every database D and every individual $t \in D$, either t ϵ-blends in a crowd of k people in D, or $San(D) \approx_\epsilon San(D \setminus \{t\})$ (or both).

Crowd-blending privacy requires that for every individual t in the database, either t blends in a crowd of k people in the database, or the mechanism essentially ignores individual t's data (the latter case is captured by $San(D) \approx_\epsilon San(D \setminus \{t\})$ in the definition). When an individual t blends in a crowd of k people in the database, the mechanism essentially does not release any information about individual t beyond the general characteristics of the crowd of k people. This is because the mechanism cannot distinguish individual t from the people in the crowd of k people, i.e., individual t's data can be changed to the data of another person in the crowd of k people and the output distribution of the mechanism remains essentially the same. A consequence is that the mechanism does not release any personally identifying information about individual t.

As mentioned in the introduction, crowd-blending privacy is not closed under composition (we later give an example in Section 3.2); however, we note that the privacy guarantee of blending in a crowd of k people in the database (described above) holds regardless of the amount of auxiliary information the adversary has (i.e., the definition is agnostic to the adversary's auxiliary information). Additionally, as mentioned previously, we show in Section 5 that when crowd-blending privacy is combined with "robust pre-sampling", we get zero-knowledge privacy and thus differential privacy as well, both of which satisfy composition in a natural way. Thus, as long as robust sampling is used during data collection before running a crowd-blending private mechanism on the collected data, independent releases from crowd-blending private mechanisms do compose and satisfy zero-knowledge privacy and differential privacy. (We also mention that one can compose a crowd-blending private mechanism with a differentially private mechanism to obtain a crowd-blending private mechanism; see Section 3.2 for details.)

Relationship with Differential Privacy. Differential privacy implies crowd-blending privacy.

Proposition 5 (Differential privacy $\Longrightarrow$ Crowd-blending privacy)**.** *Let San be any ϵ-differentially private mechanism. Then, San is (k, ϵ)-crowd-blending private for every integer $k \geq 1$.*

Proof. This immediately follows from the two privacy definitions. □

(k, ϵ)-crowd-blending privacy for some integer k does not imply differential privacy in general; this will be clear from the examples of crowd-blending private mechanisms that we give later. Crowd-blending privacy requires that for every database D and every individual $t \in D$, at least one of two conditions hold. The second condition $San(D) \approx_\epsilon San(D \setminus \{t\})$ is similar to the condition required in differential privacy. Thus, we can view crowd-blending privacy as a relaxation of differential privacy. If we remove the first condition "t ϵ-blends in a crowd of k people in D" from crowd-blending privacy, we clearly get the same definition as differential privacy. If we remove the second condition instead, it turns out that we also get differential privacy. (When we remove the

second condition $San(D) \approx_\epsilon San(D \setminus \{t\})$, we also change the definition to only consider databases of size at least k, since otherwise it would be impossible for individual t to blend in a crowd of k people in the database.)

Proposition 6 (Removing the condition $San(D) \approx_\epsilon San(D \setminus \{t\})$ in crowd-blending privacy results in differential privacy)**.** *Let San be any mechanism, let $\epsilon \geq 0$, and let k be any integer ≥ 2. Then, San is ϵ-differentially private*[3] *if and only if San satisfies the property that for every database D of size at least k and every individual $t \in D$, t ϵ-blends in a crowd of k people in D with respect to San.*

Proof. If *San* is ϵ-differentially private, then for every database D of size at least k and every individual $t \in D$, t is ϵ-indistinguishable by *San* from every individual in D, so t ϵ-blends in a crowd of k people in D.

Now, suppose *San* is not ϵ-differentially private. Then, there exist a database D and a pair of data values $t, t' \in X$ such that $San(D, t) \not\approx_\epsilon San(D, t')$. Now, consider a database D' consisting of an individual with data value t and $k-1$ individuals with data value t'. Since $San(D, t) \not\approx_\epsilon San(D, t')$, t and t' are not ϵ-indistinguishable by *San*, so the individual $t \in D'$ does not ϵ-blend in a crowd of k people in D'. □

3.1 Examples of Crowd-Blending Private Mechanisms

Given a partition P of the data universe X, and given a database D, one can compute the histogram with respect to the partition P using the database D; the histogram specifies for each block of the partition (which we refer to as a "bin") the number of individuals in D that belong to the block (which we refer to as the "count" of the bin). We first give an example of a crowd-blending private mechanism that computes a histogram and suppresses (i.e., sets to 0) bin counts that are considered too small.

Example (Histogram with suppression of small counts)**.** Let P be any partition of X. Fix $k \in \mathbb{Z}_{\geq 0}$. Let *San* be a mechanism that, on input a database D, computes the histogram with respect to the partition P using the database D, suppresses each bin count that is $< k$ (by setting the count to 0), and then releases the resulting histogram.

Then, *San* is $(k, 0)$-crowd-blending private. To see this, we note that an individual t in a database D is 0-indistinguishable by *San* from all the individuals in D that belong to the same bin as t. If there are at least k such people, then individual t blends with k people in D; otherwise, we have $San(D) \approx_0 San(D \setminus \{t\})$ since *San* suppresses each bin count that is $< k$.

It is easy to see that it is impossible to release the exact count of a bin while satisfying differential privacy. Thus, crowd-blending privacy is indeed weaker than differential privacy. For crowd-blending privacy, we can actually get better utility by adding a bit of noise to bins with low counts instead of completely suppressing them.

Example (Histogram with noise for small counts and no noise for large counts)**.** Let P be any partition of X. Fix $\epsilon > 0$ and $k \in \mathbb{Z}_{\geq 0}$. Let *San* be a mechanism that, on input a database D, computes the histogram with respect to the partition P using the database D. Then, *San* replaces each bin count $i < k$ with $A(i)$, where A is any (randomized) algorithm that satisfies $A(j) \approx_\epsilon A(j-1)$ for every $0 < j < k$ ($A(i)$ is normally a noisy version of i). *San* then releases the noisy histogram.

[3] Here, we are using the version of differential privacy that considers a pair of databases of equal size.

Then, San is (k, ϵ)-crowd-blending private. To see this, we note that an individual t in a database D is ϵ-indistinguishable (in fact, 0-indistinguishable) by San from all the individuals in D that belong to the same bin as t. If there are at least k such people, then individual t blends with k people in D, as required. If not, then we have $San(D) \approx_\epsilon San(D \setminus \{t\})$, since the histogram when using the database D is the same as the histogram when using the database $D \setminus \{t\}$ except for individual t's bin, which differs by one; however, San replaces the count i for individual t's bin with $A(i)$, and the algorithm A satisfies $A(i) \approx_\epsilon A(i-1)$, so $San(D) \approx_\epsilon San(D \setminus \{t\})$, as required.

We can choose the algorithm A to be $A(j) = j + Lap(\frac{1}{\epsilon})$, where $Lap(\lambda)$ is (a random variable with) the Laplace distribution with probability density function $f_\lambda(x) = \frac{1}{2\lambda} e^{|x|/\lambda}$. The proof that $A(j) \approx_\epsilon A(j-1)$ for every $0 < j < k$ is simple and can be implicitly found in [DMNS06].

The differentially private mechanism in [DMNS06] for computing histograms has to add noise to every bin, while our mechanism here only adds noise to the bins that have a count that is $< k$.

Example (Sanitizing a database by generalizing records safely)**.** Many mechanisms for achieving k-anonymity involve "generalizing" the records in the input table by replacing specific values with more general values, such as replacing a specific age with an age range. If this is not done carefully, the privacy of individuals can be breached, as shown by many attacks in the past (e.g., see [WFWP07, ZJB07]). Most of these mechanisms do not satisfy crowd-blending privacy. However, if the generalization of records is done carefully, achieving crowd-blending privacy may be possible.

One example is the mechanism of [LQS11]: Let Y be any set, and let $f : X \to Y$ be any function. We think of Y as a set of possible "generalized records", and f is a function that maps a record to its generalized version. Let San be a mechanism that, on input a database D, applies the function f to each individual in D; let $f(D)$ be the multi-set of images in Y. San then removes each record in $f(D)$ that appears fewer than k times in $f(D)$, and then outputs the result. It is easy to see that San is $(k, 0)$-crowd-blending private. To see this, we note that an individual t in a database D is 0-indistinguishable by San from all the individuals in D that also get mapped to $f(t)$. If there are at least k such people, then individual t blends with k people in D; otherwise, we have $San(D) \approx_0 San(D \setminus \{t\})$ since San removes each record in $f(D)$ that appears fewer than k times in $f(D)$.

3.2 Discussion of Composition

Unfortunately, crowd-blending private mechanisms do not necessarily compose, as we now show:

Proposition 7. *Let $X = \{1, 2, 3\}$ be the data universe, and let $k \in \mathbb{Z}^+$ and $\epsilon > 0$. Let San_1 and San_2 be the histogram mechanism in the "Histogram with noise for small counts and no noise for large counts" example with partitions $P_1 = \{\{1, 2\}, \{3\}\}$ and $P_2 = \{\{1\}, \{2, 3\}\}$, respectively. As shown in the example, San_1 and San_2 are both (k, ϵ)-crowd-blending private.*

Let San be the composition of San_1 and San_2, i.e., $San(D) = (San_1(D), San_2(D))$ for every database D. Then, for every $k' > 1$ and every $\epsilon' \geq 0$, San is not (k', ϵ')-crowd-blending private.

Proof. Fix $k' > 1$ and $\epsilon' \geq 0$. Let D be the database containing exactly $k-1$ individuals with data value 1, exactly 1 individual with data value 2, and exactly $k-1$ individuals with data value 3. Let t be the individual in D with data value 2.

We claim that individual t is not ϵ'-indistinguishable by San from any individual in D other than himself/herself. To see this, we note that if t changes his/her data value to 1, then the number of individuals in the database that belong to the block $\{2, 3\}$ of the partition P_2 decreases from k to $k-1$; since San_2 adds noise to counts that are $< k$ but does not add noise to counts that are $\geq k$, the output distribution of San_2 changes completely and $San(D) \approx_{\epsilon'} San(D \setminus \{t\}, 1)$ clearly does

not hold. If t changes his/her data value to 3, then the number of individuals in the database that belong to the block $\{1,2\}$ of the partition P_1 decreases from k to $k-1$; since San_1 adds noise to counts that are $< k$ but does not add noise to counts that are $\geq k$, the output distribution of San_1 changes completely and $San(D) \approx_{\epsilon'} San(D \setminus \{t\}, 3)$ clearly does not hold. Thus, individual t is not ϵ'-indistinguishable by San from any individual in D other than himself/herself, so individual t does not ϵ'-blend in a crowd of k' people in the database D.

We now claim that $San(D) \not\approx_{\epsilon'} San(D \setminus \{t\})$. To see this, we note that when the database is D, San_1 does not add noise to the bin $\{1,2\}$ of the histogram it computes, since the count of the bin is k. However, when the database is $D \setminus \{t\}$, San_1 does add noise to the bin $\{1,2\}$, since the count of the bin is $k-1$. Thus, $San(D) \not\approx_{\epsilon'} San(D \setminus \{t\})$ clearly does not hold.

It follows that San is not (k', ϵ')-crowd-blending private. □

Although crowd-blending private mechanisms do not necessarily compose, one can compose via concatenation a crowd-blending private mechanism with a differentially private mechanism to obtain a crowd-blending private mechanism.

Proposition 8. *Let San_1 be any (k, ϵ_1)-crowd-blending private mechanism, and let San_2 be any ϵ_2-differentially private mechanism. Then, the mechanism $San(D) = (San_1(D), San_2(D))$ is $(k, \epsilon_1 + 2\epsilon_2)$-crowd-blending private.*

Proof. Let D be any database and t be any individual in D. Since San_1 is (k, ϵ_1)-crowd-blending private, either t ϵ_1-blends in a crowd of k people in D with respect to San_1, or $San_1(D) \approx_{\epsilon_1} San_1(D \setminus \{t\})$.

In the former case, we have $|t' \in D : t' \approx_{\epsilon_1, San_1} t| \geq k$; now, we note that if $t' \in D$ satisfies $t' \approx_{\epsilon, San_1} t$, then t' also satisfies $t' \approx_{\epsilon_1+2\epsilon_2, San} t$ since for every database D', we have

$$San(D', t') = (San_1(D', t'), San_2(D', t')) \approx_{\epsilon_1+2\epsilon_2} (San_1(D', t), San_2(D', t)) = San(D', t).$$

(The factor of 2 in $2\epsilon_2$ appears when we use a "hybrid/transitivity" argument: Since San_2 is ϵ_2-differentially private, we have $San_2(D', t') \approx_{\epsilon_2} San_2(D') \approx_{\epsilon_2} San(D', t)$, so $San_2(D', t') \approx_{2\epsilon_2} San(D', t)$.) Thus, individual t $(\epsilon_1 + 2\epsilon_2)$-blends in a crowd of k people in D with respect to San, as required.

In the latter case, we have $San_1(D) \approx_{\epsilon_1} San_1(D \setminus \{t\})$, so

$$San(D) = (San_1(D), San_2(D)) \approx_{\epsilon_1+\epsilon_2} (San_1(D \setminus \{t\}), San_2(D \setminus \{t\})) = San(D \setminus \{t\}),$$

as required. □

4 Privately Releasing Synthetic Data Points in $\mathbb{R}^d$ for Computing Smooth Functions

Roughly speaking, a function $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$ is smooth if the value of g does not change much when we perturb the data points of the input slightly. In this section, we show that we can *efficiently* release synthetic data points in $\mathbb{R}^d$ for approximating *all* smooth functions simultaneously while satisfying crowd-blending privacy. On the other hand, we show that there are smooth functions that cannot even be approximated with non-trivial utility from synthetic data that has been released with differential privacy (even if the differentially private mechanism is *inefficient*).

In this section, the data universe X is any bounded subset of $\mathbb{R}^d$ for some positive integer d, and the input databases of mechanisms are elements of X^*. We consider mechanisms that always output

a synthetic database where each row is a data point in $\mathbb{R}^d$. We loosely use the term "synthetic data/database" to mean that the data/database was outputted by a mechanism but still has the same format as the original input data/database. Given a database/vector D, let D_i denote the i^{th} row/component of D. We now state the definition of smoothness of a function $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$.

Definition 9. Let $M : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$ and $K : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$ be functions. A function $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$ is said to be $(M(\cdot), K(\cdot))$**-smooth** if for every pair of databases $D, D' \in X^*$ of equal size n such that $||D_i - D'_i||_1 \leq M(n)$ for every $i \in [n]$, we have $||g(D) - g(D')||_1 \leq K(n)$.

Roughly speaking, a function is $(M(\cdot), K(\cdot))$-smooth if the value of the function changes by at most a distance of $K(n)$ when the data points in a database of size n are perturbed by at most a distance of $M(n)$. For example, the function that computes the mean of the data points is $(M(\cdot), M(\cdot))$-smooth for every function $M : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$. In practice, *outliers* are often removed before computing certain statistics on the data points, since outliers often cause the statistics to be less meaningful. Thus, when we consider the utility of a mechanism, we will consider how well the synthetic database released by the mechanism can be used to accurately approximate smooth functions with an outlier removal preprocessing step.

We now discuss how we decide whether a data point is an outlier or not. For the rest of the section, we fix a bounded data universe $X \subseteq \mathbb{R}^d$, a partition P of X, and an integer $k \geq 1$. Given a database D, an individual t in D is said to be an *outlier in* D (with respect to the partition P and the threshold k) if the block of P containing t contains fewer than k data points from D. We now describe what it means for a mechanism to be useful for a class of functions with outlier removal preprocessing.

Definition 10. Let San be any mechanism that always outputs a database whose rows are data points in $\mathbb{R}^d$. Let $\mathcal{C}$ be any class of functions of the form $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$. San is said to be $(\alpha(\cdot), \beta(\cdot))$**-useful for** $\mathcal{C}$ **with outlier removal preprocessing** if for every database $D \in X^*$, if we let $\widehat{D}$ be the database D with all outliers removed and $\widehat{n} = |\widehat{D}|$, then with probability at least $1 - \beta(\widehat{n})$, $San(D)$ outputs a synthetic database $\widetilde{D}$ such that

$$||g(\widetilde{D}) - g(\widehat{D})||_1 \leq \alpha(\widehat{n}) \quad \text{for every } g \in \mathcal{C}.$$

We now give an example of a crowd-blending private mechanism that releases synthetic data points in $\mathbb{R}^d$ for approximating all smooth functions with outlier removal preprocessing. Given a subset $A \subseteq \mathbb{R}^d$, let the diameter of A, denoted $diam(A)$, be defined by $diam(A) = \sup_{x,y \in A} ||x-y||_1$.

Example (Releasing noisy data points in $\mathbb{R}^d$ for approximating all smooth functions with outlier removal preprocessing)**.** Let $\epsilon > 0$. Let San be a mechanism that, on input a database D, looks at each data point $\vec{x}$ in D and does the following: If $\vec{x}$ is an outlier in D, San simply deletes $\vec{x}$. Otherwise, San replaces $\vec{x}$ with $A_B(\vec{x})$, where B is the block of the partition P that contains $\vec{x}$, and A_B is any (randomized) algorithm that satisfies $A_B(\vec{y}) \approx_\epsilon A_B(\vec{z})$ for every pair of vectors $\vec{y}, \vec{z} \in B$ ($A_B(\vec{x})$ is normally a noisy version of $\vec{x}$). San then releases all the noisy data points.

Then, San is (k, ϵ)-crowd-blending private. To see this, let D be any database and let t be any individual in D. If t is an outlier in D, then we have $San(D) = San(D \setminus \{t\})$, since San simply deletes all outliers and the removal of t from D does not change whether the other individuals are outliers or not; thus, San is (k, ϵ)-crowd-blending private, as required. Thus, we now assume t is not an outlier in D. Then, let B be the block of the partition P that contains t. We note that individual t is ϵ-indistinguishable by San from each individual $t' \in D$ in the block B, since for every database D', we have $San(D', t) \approx_\epsilon San(D', t')$ since $A_B(t) \approx_\epsilon A_B(t')$. Since t is not an outlier in

D, there are at least k people in D that belong to the block B, so t ϵ-blends in a crowd of k people in D, as required.

For each block B of the partition P, we can choose the algorithm A_B to be $A_B(\vec{y}) = \vec{y} + Lap(\frac{diam(B)}{\epsilon})^d$, where $Lap(\frac{diam(B)}{\epsilon})^d$ is a random vector with d components, each of which is independently distributed as $Lap(\frac{diam(B)}{\epsilon})$. Using techniques/results found in [DMNS06], it is easy to show that $A_B(\vec{y}) \approx_\epsilon A_B(\vec{z})$ for every pair of vectors $\vec{y}, \vec{z} \in B$.

Remark. Even though *San* essentially runs a differentially private mechanism within each block (that does not contain too few data points), it is not the case that the only information that remains for a block is the number of data points that belong to the block. This is because there can be many data points within a block, and if *San* adds Laplacian noise to each data point as above, the general distribution of data points and many statistics are preserved in expectation and would also be reasonably accurate with high probability. Outputting just the number of data points within a block does not tell us such distributional and statistical information. Thus, we do not get the same result if *San* simply outputs the number of data points within each block like a histogram.

We now show that the above crowd-blending private mechanism with $A_B(\vec{y}) = \vec{y} + Lap(\frac{diam(B)}{\epsilon})^d$ is useful for all smooth functions with outlier removal preprocessing.

Proposition 11. *Let $\epsilon > 0$ and $L > 0$, and let $M : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$ and $K : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$ be arbitrary functions. Suppose $diam(B) \leq L$ for every block B of the partition P. Let San be the mechanism in the above example with $A_B(\vec{y}) = \vec{y} + Lap(\frac{diam(B)}{\epsilon})^d$. Then, San is $(K(\cdot), \beta(\cdot))$-useful for the class $\mathcal{C}$ of all $(M(\cdot), K(\cdot))$-smooth functions with outlier removal preprocessing, where $\beta(\widehat{n}) = d\widehat{n}e^{-\frac{\epsilon M(\widehat{n})}{dL}}$.*

Proof. Let $D \in X^*$, let $\widehat{D}$ be the database D with all outliers removed, and let $\widehat{n} = |\widehat{D}|$. Let $\widetilde{D} = San(D)$. Since $San(D)$ simply removes all outliers in D, we have $\widetilde{D} = San(\widehat{D})$. Now, we note that $|\widetilde{D}| = \widehat{n}$ and for every $i \in [\widehat{n}]$, we have $\widetilde{D}_i = \widehat{D}_i + Lap(\frac{diam(B_i)}{\epsilon})^d$, where B_i is the block of P that contains $\widehat{D}_i$. Let $\lambda = \frac{L}{\epsilon}$ so that $\frac{diam(B_i)}{\epsilon} \leq \lambda$ for every $i \in [\widehat{n}]$. From the p.d.f. or c.d.f. of $Lap(\lambda)$, it is easy to verify that for every $\delta \geq 0$, we have $\Pr_{X \sim Lap(\lambda)}[|X| \leq \delta] = 1 - e^{-\frac{\delta}{\lambda}}$, so $\Pr_{X \sim Lap(\lambda)^d}[||X||_1 \leq \delta] \geq 1 - de^{-\frac{\delta}{d\lambda}}$ by a union bound. Then, for every $i \in [\widehat{n}]$, we have

$$\begin{aligned}
\Pr\left[||\widetilde{D}_i - \widehat{D}_i||_1 \leq M(\widehat{n})\right] &= \Pr_{X \sim Lap(\frac{diam(B_i)}{\epsilon})^d}[||X||_1 \leq M(\widehat{n})] \\
&\geq \Pr_{X \sim Lap(\lambda)^d}[||X||_1 \leq M(\widehat{n})] \\
&\geq 1 - de^{-\frac{\epsilon M(\widehat{n})}{dL}}.
\end{aligned}$$

Then, by a union bound, with probability at least $1 - \widehat{n}de^{-\frac{\epsilon M(\widehat{n})}{dL}}$, we have $||\widetilde{D}_i - \widehat{D}_i||_1 \leq M(\widehat{n})$ for every $i \in [\widehat{n}]$. Then, with probability at least $1 - \widehat{n}de^{-\frac{\epsilon M(\widehat{n})}{dL}}$, we have $||g(\widetilde{D}) - g(\widehat{D})||_1 \leq K(\widehat{n})$ for every $(M(\cdot), K(\cdot))$-smooth function $g : (\mathbb{R}^d)^* \to \mathbb{R}^m$ by definition of $(M(\cdot), K(\cdot))$-smooth. Thus, *San* is $(K(\cdot), \beta(\cdot))$-useful for the class $\mathcal{C}$ of all $(M(\cdot), K(\cdot))$-smooth functions with outlier removal preprocessing. □

We note that $\beta(\widehat{n}) = d\widehat{n}e^{-\frac{\epsilon M(\widehat{n})}{dL}}$ can be made to be negligible by choosing $M(\widehat{n}) = \Omega(\widehat{n}^\kappa)$ for any $\kappa > 0$. We also note that the mechanism in the proposition can clearly be implemented efficiently. We now show that there exist $(M(\cdot), K(\cdot))$-smooth functions that cannot be computed with non-trivial utility from synthetic data released by a differentially private mechanism, regardless of the running time of the mechanism.

Proposition 12. *Let $g : (\mathbb{R}^d)^* \to \mathbb{R}^d$ be the function defined by $g(D) = D_1$, which is clearly $(M(\cdot), M(\cdot))$-smooth for every function $M : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$. Let $\epsilon \geq 0$, and let San be any (possibly inefficient) ϵ-differentially private mechanism that always outputs a database where each row is a data point in $\mathbb{R}^d$. Then, for every $\delta > 0$, San is not even $(\frac{diam(X)}{4}, \frac{1}{1+e^\epsilon} - \delta)$-useful for the function g with outlier removal preprocessing.*

Proof. Let $\vec{x}$ and $\vec{y}$ be any pair of data points in X such that $||\vec{x} - \vec{y}||_1 \geq \frac{3}{4} diam(X)$. Let D be the database consisting of exactly $k+1$ copies of $\vec{x}$ followed by exactly k copies of $\vec{y}$, and let D' be the same as D except that the first row is changed from $\vec{x}$ to $\vec{y}$. Then, both D and D' do not contain any outliers.

Let $\delta > 0$. To obtain a contradiction, suppose San is $(\frac{diam(X)}{4}, \frac{1}{1+e^\epsilon} - \delta)$-useful for the function g with outlier removal preprocessing. Then, San is also $(\frac{diam(X)}{4}, \frac{1}{1+e^\epsilon})$-useful for g with outlier removal preprocessing. Then, we have

$$\Pr[||San(D)_1 - \vec{x}||_1 \leq diam(X)/4] \geq 1 - \frac{1}{1+e^\epsilon} = \frac{e^\epsilon}{1+e^\epsilon}.$$

Since San is ϵ-differentially private and D and D' differ by only one row, we have $San(D) \approx_\epsilon San(D')$, so

$$\begin{aligned}\Pr[||San(D')_1 - \vec{x}||_1 \leq diam(X)/4] &\geq e^{-\epsilon} \cdot \Pr[||San(D)_1 - \vec{x}||_1 \leq diam(X)/4] \\ &\geq \frac{1}{1+e^\epsilon}.\end{aligned}$$

Since $||\vec{x} - \vec{y}||_1 \geq \frac{3}{4} diam(X)$, if $||San(D')_1 - \vec{x}||_1 \leq diam(X)/4$ holds, then $||San(D')_1 - \vec{y}||_1 \leq diam(X)/4$ does not hold. It follows that

$$\begin{aligned}\Pr[||g(San(D')) - g(D')||_1 \leq diam(X)/4] &= \Pr[||San(D')_1 - \vec{y}||_1 \leq diam(X)/4] \\ &\leq 1 - \Pr[||San(D')_1 - \vec{x}||_1 \leq diam(X)/4] \\ &\leq 1 - \frac{1}{1+e^\epsilon} \\ &< 1 - \left(\frac{1}{1+e^\epsilon} - \delta\right).\end{aligned}$$

This contradicts our assumption that San is $(\frac{diam(X)}{4}, \frac{1}{1+e^\epsilon} - \delta)$-useful for g with outlier removal preprocessing. □

In Proposition 12, we note that the image of X^* under g is X (recall that the databases that we run mechanisms on are elements of X^*) and $\frac{1}{1+e^\epsilon} \approx \frac{1}{2}$ when ϵ is small, and so requiring San to be $(\frac{diam(X)}{4}, \frac{1}{1+e^\epsilon} - \delta)$-useful for g is only requiring San to possibly provide non-trivial utility; however, the proposition says that San cannot even satisfy this non-triviality requirement. If we apply Proposition 11 to the same function g, we see that for every function $M : \mathbb{Z}_{\geq 0} \to \mathbb{R}^+$, the crowd-blending private mechanism is $(M(\cdot), \beta(\cdot))$-useful for g with outlier removal preprocessing, where $\beta(\widehat{n}) = d\widehat{n}e^{-\frac{\epsilon M(\widehat{n})}{dL}}$ and L is a bound on the diameter of every block of the partition P. The utility guarantee of this result is non-trivial in many situations. Thus, it is possible to release synthetic data points for approximating smooth functions while satisfying crowd-blending privacy, but doing this while satisfying differential privacy is impossible in general.

5 Our Main Theorem

In this section, we prove our main theorem that says that when we combine a crowd-blending private mechanism with a natural *pre-sampling* step, the combined algorithm is zero-knowledge private (and thus differentially private as well). The pre-sampling step should be thought of as being part of the data collection process, where individuals in some population are sampled and asked for their data. A crowd-blending private mechanism is then run on the samples to release useful information while preserving privacy.

We first prove our main theorem for the case where the pre-sampling step samples each individual in the population with probability p independently. In reality, the sampling performed during data collection may be slightly *biased* or done slightly incorrectly, and an adversary may know whether certain individuals were sampled or not. Thus, we later extend our main theorem to the case where the sampling probability is not necessarily the same for everybody, but the sampling is still *robust* in the sense that most individuals are sampled independently with probability in between p and p' (this probability can even depend on the individual's data), where p and p' are relatively close to one another, while the remaining individuals are sampled independently with arbitrary probability.

We begin with some necessary terminology and notation. A *population* is a collection of *individuals*, where an individual is simply represented by a data value in the data universe X. Thus, a population is actually a multiset of data values, which is the same as a database. (If we want individuals to have unique data values, we can easily modify X to include personal/unique identifiers.) Given a population $\mathcal{P}$ and a real number $p \in [0,1]$, let $Sam(\mathcal{P}, p)$ be the outcome of sampling each individual in $\mathcal{P}$ with probability p independently.

Although zero-knowledge privacy was originally defined for mechanisms operating on *databases*, one can also consider mechanisms operating on *populations*, since there is essentially no difference between the way we model populations and databases. (In the definition of zero-knowledge privacy, we simply change "database" to "population" and D to $\mathcal{P}$.) We now describe a class of (randomized) aggregation functions that we will use in the definition of zero-knowledge privacy.

- $iidRS(p)$ = i.i.d. random sampling with probability p : the class of algorithms T such that on input a population $\mathcal{P}$, T chooses each individual in $\mathcal{P}$ with probability p independently, and then performs any computation on the data of the chosen individuals.[4]

We now state and prove the basic version of our main theorem.

Theorem 13 (Sampling + Crowd-Blending Privacy ⇒ Zero-Knowledge Privacy)**.** *Let San be any (k,ϵ)-crowd-blending private mechanism with $k \geq 2$, and let $p \in (0,1)$. Then, the algorithm San_{zk} defined by $San_{zk}(\mathcal{P}) = San(Sam(\mathcal{P}, p))$ for any population $\mathcal{P}$ is $(\epsilon_{zk}, \delta_{zk})$-zero-knowledge private*[5] *with respect to $iidRS(p)$, where*

$$\epsilon_{zk} = \ln\left(p \cdot \left(\frac{2-p}{1-p}e^{\epsilon}\right) + (1-p)\right) \quad \text{and} \quad \delta_{zk} = e^{-\Omega(k\cdot(1-p)^2)}.$$

To prove Theorem 13, we will first prove two supporting lemmas. The first lemma essentially says that if an individual t blends with (i.e., is indistinguishable by San from) many people in the

[4] To make zero-knowledge privacy compose naturally for this type of aggregate information, we can extend $iidRS(p)$ to $iidRS(p,r)$, where T is now allowed to perform r rounds of sampling before performing any computation on the sampled data. It is not hard to see that zero-knowledge privacy with respect to $iidRS(p,r)$ composes in a natural way.

[5] The constant hidden by the $\Omega(\cdot)$ in δ_{zk} can be easily computed; however, we did not try to optimize the constant in any way.

population, then t's privacy is protected when we sample from the population and run San on the samples:

Lemma 14 (Protection of individuals that blend with many people in the population)**.** *Let San be any mechanism, $\mathcal{P}$ be any population, $p \in (0,1)$, and $\epsilon \geq 0$. Let t be any individual in $\mathcal{P}$, and let A be any non-empty subset of $\mathcal{P} \setminus \{t\}$ such that $t' \approx_{\epsilon, San} t$ for every individual $t' \in A$. Let $n = |A|$. Then, we have*

$$San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{final}, \delta_{final}} San(Sam(\mathcal{P} \setminus \{t\}, p)),$$

where $\epsilon_{final} = \ln(p \cdot (\frac{2-p}{1-p} e^{\epsilon}) + (1-p))$ and $\delta_{final} = e^{-\Omega((n+1)p(1-p)^2)}$.

In the lemma, A is any non-empty set of individuals in $\mathcal{P} \setminus \{t\}$ that blend with individual t. (We could set A to be the set of *all* individuals in $\mathcal{P} \setminus \{t\}$ that blend with individual t, but leaving A more general allows us to more easily extend the lemma to the case of "robust" sampling later.) We note that δ_{final} is smaller when $n = |A|$ is larger, i.e., when t blends with more people. Intuitively, if an individual t is indistinguishable by San from many other people in the population, then t's presence or absence in the population does not affect the output of $San(Sam(\cdot, p))$ much, since the people indistinguishable from t can essentially take the place of t in almost any situation (and the output of San would essentially be the same). Since it does not matter much whether individual t is in the population or not, it follows that t's privacy is protected.

The proof of the lemma *roughly* works as follows: Consider two scenarios, one where individual t is in the population (i.e., $San(Sam(\mathcal{P}, p))$ in the lemma), and one where individual t has been removed from the population (i.e., $San(Sam(\mathcal{P} \setminus \{t\}, p))$ in the lemma). Our goal is to show that the output of San is essentially the same in the two scenarios, i.e., $San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{final}, \delta_{final}} San(Sam(\mathcal{P} \setminus \{t\}, p))$. Conditional on individual t not being sampled in the first scenario, the two scenarios are exactly the same, as desired. Thus, we now always condition on individual t being sampled in the first scenario. In the lemma, A is a set of individuals in the population (excluding t) that are indistinguishable from t by San. Let $\widetilde{m}$ denote the number of people in A that are sampled. The proof involves showing the following two properties:

1. $\widetilde{m}$ is relatively smooth near its expectation: For every integer m near the expectation of $\widetilde{m}$, $\Pr[\widetilde{m} = m]$ is relatively close to $\Pr[\widetilde{m} = m+1]$.
2. For every integer $m \in \{0, \ldots, n-1\}$, the output of San in the first scenario conditioned on $\widetilde{m} = m$ (and t being sampled) is essentially the same as the output of San in the second scenario conditioned on $\widetilde{m} = m+1$.

For the first property, we note that $\widetilde{m}$ follows a binomial distribution, which can be shown to be relatively smooth near its expectation. To show the second property, we note that when we condition on $\widetilde{m} = m$ (and t being sampled) in the first scenario, m random samples are drawn uniformly from A (one at a time) without replacement, and also $t \notin A$ is sampled for sure (and the remaining individuals are sampled independently with probability p). This is very similar to the second scenario conditioned on $\widetilde{m} = m+1$, where $m+1$ random samples are drawn uniformly from A without replacement, since if we replace the $(m+1)^{th}$ sample by t, we get back the first scenario conditioned on $\widetilde{m} = m$ (and t being sampled). Since the $(m+1)^{th}$ sample is indistinguishable from t by San, the output of San is essentially the same in both scenarios.

Using the two properties above, one can show that when $\widetilde{m}$ is close to its expectation, the output of San is essentially the same in both scenarios. δ_{final} in the lemma captures the probability of the bad event where $\widetilde{m}$ is not close to its expectation, which we bound by essentially using a Chernoff bound. We now give the formal proof of Lemma 14.

Proof of Lemma 14. Let $\epsilon_{Sam} > 0$, $\widehat{D} = Sam(\mathcal{P}, p)$, $\widetilde{D} = Sam(\mathcal{P} \setminus \{t\}, p)$, $\widetilde{m} = |\widetilde{D} \cap A|$, and $Y \subseteq \{0,1\}^*$. Let E be the event that t is sampled when $\widehat{D}$ is chosen. We first observe that

$$\begin{aligned}\Pr[San(\widehat{D}) \in Y] &= \Pr[San(\widehat{D}) \in Y \mid E] \cdot \Pr[E] + \Pr[San(\widehat{D}) \in Y \mid \overline{E}] \cdot \Pr[\overline{E}] \\ &= \Pr[San(\widetilde{D} \cup \{t\}) \in Y] \cdot p + \Pr[San(\widetilde{D}) \in Y] \cdot (1-p). \end{aligned} \tag{1}$$

We will now show that for every $m \in \{0, \ldots, n-1\}$, we have

$$\left| \ln \left(\frac{\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m]}{\Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1]} \right) \right| \leq \epsilon. \tag{2}$$

Fix $m \in \{0, \ldots, n-1\}$. Let $\mathcal{P}_{-t,-A} = (\mathcal{P} \setminus \{t\}) \setminus A$. We note that for $j \in \{0, \ldots, n\}$, the conditional distribution of $\widetilde{D}$ given $\widetilde{m} = j$ is equal to $Sam(\mathcal{P}_{-t,-A}, p) \cup A_j$, where A_j is the outcome of choosing j random samples uniformly without replacement from A. Then, using the fact that $t' \approx_{\epsilon, San} t$ for every individual $t' \in A$, we have

$$\begin{aligned}&\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m] = \Pr[San(Sam(\mathcal{P}_{-t,-A}, p) \cup A_m \cup \{t\}) \in Y] \\ &\leq e^{\epsilon} \Pr[San(Sam(\mathcal{P}_{-t,-A}, p) \cup A_{m+1}) \in Y] = e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1].\end{aligned}$$

Similarly, we also have

$$\begin{aligned}&\Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1] = \Pr[San(Sam(\mathcal{P}_{-t,-A}, p) \cup A_{m+1}) \in Y] \\ &\leq e^{\epsilon} \Pr[San(Sam(\mathcal{P}_{-t,-A}, p) \cup A_m \cup \{t\}) \in Y] = e^{\epsilon} \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m].\end{aligned}$$

Thus, we have shown (2).

Now, we observe that for every $m \in \{0, \ldots, n-1\}$, if $m+1 \leq (n+1)p \cdot \frac{e^{\epsilon_{Sam}}}{pe^{\epsilon_{Sam}} + (1-p)}$, then

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m+1]} = \frac{\binom{n}{m} p^m (1-p)^{n-m}}{\binom{n}{m+1} p^{m+1} (1-p)^{n-(m+1)}} = \frac{m+1}{n-m} \frac{1-p}{p} \leq e^{\epsilon_{Sam}}. \tag{3}$$

Let $\alpha = \frac{e^{\epsilon_{Sam}}}{pe^{\epsilon_{Sam}} + (1-p)}$ and $\delta_{Sam} = \Pr[\widetilde{m} + 1 > (n+1)p \cdot \alpha]$. Now, using (3) and (2) (and the fact that $m = n$ does not satisfy $m + 1 \leq (n+1)p \cdot \alpha$), we have

$$\begin{aligned}&\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \\ &\leq \sum_{\substack{m \in \{0,\ldots,n\} \\ m+1 \leq (n+1)p \cdot \alpha}} \Pr[\widetilde{m} = m] \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m] + \Pr[\widetilde{m} + 1 > (n+1)p \cdot \alpha] \\ &\leq \sum_{\substack{m \in \{0,\ldots,n\} \\ m+1 \leq (n+1)p \cdot \alpha}} e^{\epsilon_{Sam}} \Pr[\widetilde{m} = m+1] \cdot e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1] + \delta_{Sam} \\ &\leq e^{\epsilon + \epsilon_{Sam}} \Pr[San(\widetilde{D}) \in Y] + \delta_{Sam}. \end{aligned} \tag{4}$$

Let $\epsilon_{total} = \max\{\ln(pe^{\epsilon + \epsilon_{Sam}} + (1-p)), \ln(\frac{1}{1-p})\}$. Combining (1) and (4), we have

$$\begin{aligned}\Pr[San(\widehat{D}) \in Y] &\leq (e^{\epsilon + \epsilon_{Sam}} \Pr[San(\widetilde{D}) \in Y] + \delta_{Sam}) \cdot p + \Pr[San(\widetilde{D}) \in Y] \cdot (1-p) \\ &\leq e^{\epsilon_{total}} \Pr[San(\widetilde{D}) \in Y] + p \cdot \delta_{Sam}.\end{aligned}$$

By (1), we also have

$$\Pr[San(\widehat{D}) \in Y] \geq (1-p)\Pr[San(\widetilde{D}) \in Y] \geq e^{-\epsilon_{total}}\Pr[San(\widetilde{D}) \in Y].$$

Thus, we have $San(\widehat{D}) \approx_{\epsilon_{total}, p\cdot\delta_{Sam}} San(\widetilde{D})$.

Now, we set $\epsilon_{Sam} = \ln(\frac{2-p}{1-p})$. Then, we have

$$\begin{aligned}
\epsilon_{total} &= \max\{\ln(pe^{\epsilon+\epsilon_{Sam}} + (1-p)), \ln(\frac{1}{1-p})\} \\
&= \max\{\ln(p\cdot(\frac{2-p}{1-p}e^{\epsilon}) + (1-p)), \ln(\frac{1}{1-p})\} \\
&= \ln(p\cdot(\frac{2-p}{1-p}e^{\epsilon}) + (1-p))
\end{aligned}$$

and

$$\begin{aligned}
p\cdot\delta_{Sam} &= p\cdot\Pr[\widetilde{m}+1 > (n+1)p\cdot(2-p)] \\
&\leq \Pr[\widetilde{m} + Bin(1,p) > (n+1)p\cdot(2-p)] \\
&\leq e^{-\Omega((n+1)p(1-p)^2)},
\end{aligned}$$

where $Bin(1,p)$ is a binomial random variable with 1 trial and success probability p, and the last inequality follows from a multiplicative Chernoff bound. □

We now show how pre-sampling combined with a crowd-blending private mechanism can protect the privacy of individuals who blend with (i.e., are indistinguishable by *San* from) few people in the population.

Lemma 15 (Protection of individuals that blend with few people in the population)**.** *Let San be any (k,ϵ)-crowd-blending private mechanism with $k \geq 2$, let $\mathcal{P}$ be any population, and let $p \in (0,1)$. Let t be any individual in $\mathcal{P}$, and let $n = |\{t' \in \mathcal{P}\setminus\{t\} : t' \approx_{\epsilon,San} t\}|$. Then, if $n \leq \frac{k-1}{p(2-p)}$, we have*

$$San(Sam(\mathcal{P},p)) \approx_{\epsilon_{final},\delta_{final}} San(Sam(\mathcal{P}\setminus\{t\},p)),$$

where $\epsilon_{final} = \ln(pe^{\epsilon}+(1-p))$ and $\delta_{final} = pe^{-\Omega(k\cdot(1-p)^2)}$.

The proof of the lemma *roughly* works as follows: In the lemma, n is the number of people in the population that individual t blends with, and is assumed to be small. We will show that when we remove individual t from the population, the output of *San* does not change much.

Consider two scenarios, one where individual t is in the population, and one where individual t has been removed from the population. Conditional on individual t not being sampled in the first scenario, the two scenarios are exactly the same, as desired. Thus, we now always condition on individual t being sampled in the first scenario. Since individual t blends with few people in the population, we have that with very high probability, the database obtained from sampling from the population would contain fewer than k people that blend with individual t; since *San* is (k,ϵ)-crowd-blending private and individual t does not blend in a crowd of k people in the database, *San* must essentially ignore individual t's data; thus, the first scenario is essentially the same as the second scenario, since individual t's data is essentially ignored anyway. δ_{final} in the lemma captures the probability of the bad event where the database obtained from sampling actually contains k people that blend with individual t. We now give the formal proof of Lemma 15.

Proof of Lemma 15. Suppose $n \leq \frac{k-1}{p(2-p)}$. Let $\widehat{D} = Sam(\mathcal{P}, p)$, $\widetilde{D} = Sam(\mathcal{P} \setminus \{t\}, p)$, and $Y \subseteq \{0,1\}^*$. Let $A = \{t' \in \mathcal{P} \setminus \{t\} : t' \approx_{\epsilon, San} t\}$, so $n = |A|$. Let $\widetilde{m} = |\widetilde{D} \cap A|$, and let E be the event that individual t is in $\widehat{D}$ when $\widehat{D}$ is chosen. We first note that

$$\begin{aligned}\Pr[San(\widehat{D}) \in Y] &= \Pr[San(\widehat{D}) \in Y \mid \overline{E}] \cdot \Pr[\overline{E}] + \Pr[San(\widehat{D}) \in Y \mid E] \cdot \Pr[E] \\ &= \Pr[San(\widetilde{D}) \in Y] \cdot (1-p) + \Pr[San(\widetilde{D} \cup \{t\}) \in Y] \cdot p. \end{aligned} \tag{1}$$

Since San is (k, ϵ)-crowd-blending private, we have

$$\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-1] \leq e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-1]$$

and

$$\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-1] \geq e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-1].$$

Then, we have

$$\begin{aligned}\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \leq\ & \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-1] \Pr[\widetilde{m} < k-1] + \Pr[\widetilde{m} \geq k-1] \\ &\leq e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-1] \Pr[\widetilde{m} < k-1] + \Pr[\widetilde{m} \geq k-1] \\ &\leq e^{\epsilon} \Pr[San(\widetilde{D}) \in Y] + \Pr[\widetilde{m} \geq k-1], \end{aligned} \tag{2}$$

and

$$\begin{aligned}\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \geq\ & \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-1] \Pr[\widetilde{m} < k-1] \\ &\geq e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-1] \Pr[\widetilde{m} < k-1] \\ &\geq e^{-\epsilon} (\Pr[San(\widetilde{D}) \in Y] - \Pr[\widetilde{m} \geq k-1]) \\ &= e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y] - e^{-\epsilon} \Pr[\widetilde{m} \geq k-1]. \end{aligned} \tag{3}$$

Now, combining (1) and (2), we have

$$\Pr[San(\widehat{D}) \in Y] \leq (pe^{\epsilon} + (1-p)) \Pr[San(\widetilde{D}) \in Y] + p \Pr[\widetilde{m} \geq k-1]. \tag{4}$$

Also, combining (1) and (3), we have

$$\Pr[San(\widehat{D}) \in Y] \geq (pe^{-\epsilon} + (1-p)) \Pr[San(\widetilde{D}) \in Y] - e^{-\epsilon} p \Pr[\widetilde{m} \geq k-1].$$

Rearranging this inequality, we get

$$\begin{aligned}\Pr[San(\widetilde{D}) \in Y] &\leq \frac{1}{pe^{-\epsilon} + (1-p)} \Pr[San(\widehat{D}) \in Y] + \frac{e^{-\epsilon}}{pe^{-\epsilon} + (1-p)} p \Pr[\widetilde{m} \geq k-1] \\ &\leq (pe^{\epsilon} + (1-p)) \Pr[San(\widehat{D}) \in Y] + p \Pr[\widetilde{m} \geq k-1], \end{aligned} \tag{5}$$

where the last inequality follows from the fact that the function $f(x) = \frac{1}{x}$ is convex for $x > 0$, so $\frac{1}{pe^{-\epsilon} + (1-p)} \leq pe^{\epsilon} + (1-p)$.

Let $\tau = \frac{k-1}{p(2-p)}$. Then, we have $n \leq \tau$. The lemma now follows from (4), (5), and the inequality

$$\begin{aligned} p \Pr[\widetilde{m} \geq k-1] &= p \Pr[\widetilde{m} \geq \tau p \cdot (2-p)] \\ &\leq p \Pr[\widetilde{m} + Bin(\lfloor \tau \rfloor - n, p) + Bin(1, (\tau - \lfloor \tau \rfloor)p) \geq \tau p \cdot (2-p)] \\ &\leq p e^{-\Omega(\tau p (1-p)^2)} \\ &\leq p e^{-\Omega(k \cdot (1-p)^2)}, \end{aligned}$$

where $Bin(j,p)$ denotes a binomial random variable with j trials and success probability p, and the second inequality follows from a multiplicative chernoff bound (note that the expectation of $\widetilde{m} + Bin(\lfloor\tau\rfloor - n, p) + Bin(1, (\tau - \lfloor\tau\rfloor)p)$ is τp). □

We are now ready to prove Theorem 13. The proof roughly works as follows: By definition of $iidRS(p)$, a simulator in the definition of zero-knowledge privacy is able to obtain the aggregate information $Sam(\mathcal{P}\setminus\{t\}, p)$. With $Sam(\mathcal{P}\setminus\{t\}, p)$, the simulator can easily compute $San(Sam(\mathcal{P}\setminus\{t\}, p))$, which it can then use to simulate the computation of the given adversary. It is not hard to see that the simulation works if $San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P}\setminus\{t\}, p))$ holds. Thus, consider any population $\mathcal{P}$ and any individual $t \in \mathcal{P}$. Recall that Lemma 14 protects the privacy of individuals that blend with many people in $\mathcal{P}$, while Lemma 15 protects the privacy of individuals that blend with few people in $\mathcal{P}$. Thus, if individual t blends with many people in $\mathcal{P}$, we use Lemma 14; otherwise, we use Lemma 15. It then follows that $San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P}\setminus\{t\}, p))$, as required. We now give the formal proof of Theorem 13.

Proof of Theorem 13. We first note that $Sam(\cdot, p) \in iidRS(p)$. Thus, we can let $T = Sam(\cdot, p)$ in the definition of zero-knowledge privacy with respect to $iidRS(p)$. Let A be any adversary. We will describe how to construct a simulator S for A. Let $\mathcal{P}$ be any population, t be any individual in $\mathcal{P}$, and $z \in \{0,1\}^*$. Since the simulator S is given $T(\mathcal{P}\setminus\{t\}) = Sam(\mathcal{P}\setminus\{t\}, p)$ and z as part of its input, S can easily compute $San_{zk}(\mathcal{P}\setminus\{t\}) = San(Sam(\mathcal{P}\setminus\{t\}, p))$ and then simulate the computation of the adversary A that is given $San_{zk}(\mathcal{P}\setminus\{t\})$ and the auxiliary information z; the simulator S then outputs whatever A outputs.

Now, we note that if $San_{zk}(\mathcal{P}) \approx_{\epsilon_{zk},\delta_{zk}} San_{zk}(\mathcal{P}\setminus\{t\})$, then $Out_A(A(z) \leftrightarrow San_{zk}(\mathcal{P})) \approx_{\epsilon_{zk},\delta_{zk}} S(z, T(\mathcal{P}\setminus\{t\}), |\mathcal{P}|)$. Thus, to show that San_{zk} is $(\epsilon_{zk}, \delta_{zk})$-zero-knowledge private with respect to $iidRS(p)$, it suffices to show that $San_{zk}(\mathcal{P}) \approx_{\epsilon_{zk},\delta_{zk}} San_{zk}(\mathcal{P}\setminus\{t\})$, i.e.,

$$San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P}\setminus\{t\}, p)).$$

To this end, let $A = \{t' \in \mathcal{P}\setminus\{t\} : t' \approx_{\epsilon,San} t\}$ and $n = |A|$. Let $\tau = \frac{k-1}{p(2-p)}$. We will consider two cases: $n > \tau$ and $n \leq \tau$.

Suppose $n > \tau$. By Lemma 14, we have

$$San(Sam(\mathcal{P}, p)) \approx_{\epsilon_1,\delta_1} San(Sam(\mathcal{P}\setminus\{t\}, p)),$$

where $\epsilon_1 = \ln(p \cdot (\frac{2-p}{1-p}e^{\epsilon}) + (1-p)) = \epsilon_{zk}$ and $\delta_1 = e^{-\Omega((n+1)p(1-p)^2)} \leq e^{-\Omega(k\cdot(1-p)^2)} = \delta_{zk}$.

Now, suppose $n \leq \tau$. By Lemma 15, we have

$$San(Sam(\mathcal{P}, p)) \approx_{\epsilon_2,\delta_2} San(Sam(\mathcal{P}\setminus\{t\}, p)),$$

where $\epsilon_2 = \ln(pe^{\epsilon} + (1-p)) \leq \epsilon_1 = \epsilon_{zk}$ and $\delta_2 = pe^{-\Omega(k\cdot(1-p)^2)} \leq \delta_{zk}$.

It follows that

$$San(Sam(\mathcal{P}, p)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P}\setminus\{t\}, p)),$$

as required. □

5.1 Our Main Theorem Extended to Robust Sampling

We now extend our main theorem to the case where the sampling probability is not necessarily the same for everybody, but the sampling is still "robust" in the sense that most individuals are sampled independently with probability in between p and p' (this probability can even depend on the individual's data), where p and p' are relatively close to one another (i.e., $\frac{p'}{p}$ is not too large), while the remaining individuals are sampled independently with arbitrary probability.

We begin with some more notation. Given a population $\mathcal{P}$ and a function $\pi : X \to [0,1]$, let $Sam(\mathcal{P}, \pi)$ be the outcome of sampling each individual t in $\mathcal{P}$ with probability $\pi(t)$ independently. We note that for $Sam(\mathcal{P}, \pi)$, two individuals in $\mathcal{P}$ with the same data value in X will have the same probability of being sampled. However, we can easily modify the data universe X to include personal/unique identifiers so that we can represent an individual by a unique data value in X. Thus, for convenience, we now define a population to be a subset of the data universe X instead of being a multiset of data values in X. Then, each individual in a population would have a unique data value in X, so π does not have to assign the same sampling probability to two different individuals. We now describe a class of aggregation functions that we will use in the definition of zero-knowledge privacy.

- $iRS(p, p', \ell)$ = independent random sampling with probability in between p and p' except for ℓ individuals: the class of algorithms T such that on input a population $\mathcal{P}$, T independently chooses each individual $t \in \mathcal{P}$ with some probability $p_t \in [0,1]$ (possibly dependent on t's data), but all except for at most ℓ individuals in $\mathcal{P}$ must be chosen with probability in $\{0\} \cup [p, p']$; T then performs any computation on the chosen individuals' data.

We now state the extended version of our main theorem.

Theorem 16 (Robust Sampling + Crowd-Blending Privacy ⇒ Zero-Knowledge Privacy)**.** *Let San be any (k, ϵ)-crowd-blending private mechanism with $k \geq 2$, let $0 < p \leq p' < 1$, let $\pi : X \to [0,1]$ be any function, let $\ell = |\{x \in X : \pi(x) \notin \{0\} \cup [p, p']\}|$, and let $p_{\max} = \sup_{x \in X} \pi(x)$. Suppose $\ell < k - 1$.*

Then, the algorithm San_{zk} defined by $San_{zk}(\mathcal{P}) = San(Sam(\mathcal{P}, \pi))$ for any population $\mathcal{P}$ is $(\epsilon_{zk}, \delta_{zk})$-zero-knowledge private with respect to $iRS(p, p', \ell)$, where

$$\epsilon_{zk} = \ln\left(p_{\max} \cdot \left(\frac{p'}{p}\frac{(1-p)(2-p)}{(1-p')^2}e^{\epsilon}\right) + (1 - p_{\max})\right) \text{ and}$$
$$\delta_{zk} = \max\left\{\frac{p_{\max}}{p}, \frac{p_{\max}}{1-p'}\right\} e^{-\Omega((k-\ell)\cdot(1-p')^2)}.$$

In the theorem, ℓ represents the number of individuals that are sampled with probability outside of $\{0\} \cup [p, p']$. We prove the theorem by extending Lemmas 14 and 15 to the case of "robust" sampling. We first describe *some* of the main changes to the lemmas and their proofs, and then we give the formal proof of Theorem 16.

Let us first consider Lemma 14, which protects the privacy of individuals that blend with many people in the population. Like before, consider two scenarios, one where individual t is in the population, and one where individual t has been removed. Let $\widetilde{m}$ denote the number of people in A that are sampled (recall that A is a set of individuals that blend with individual t). Recall that in the proof of Lemma 14, we had to show two properties: (1) $\widetilde{m}$ is relatively smooth near its expectation, and (2) the output of San in the first scenario conditioned on $\widetilde{m} = m$ (and t being sampled) is essentially the same as the output of San in the second scenario conditioned on $\widetilde{m} = m + 1$.

For the first property, we used the fact that the binomial distribution is relatively smooth near its expectation. Here, since the sampling is no longer i.i.d. but is still robust, we need the Poisson binomial distribution (the sum of independent Bernoulli trials, where the success probabilities are not necessarily the same) to be relatively smooth near its expectation. This can be shown as long as the success probabilities are all relatively close to one another; this is ensured by changing the lemma so that everyone in the set A is required to have a sampling probability in $[p, p']$.

For the second property, we used the fact that when we condition on $\widetilde{m} = m + 1$ in the second scenario, we are drawing $m + 1$ random samples from A (one at a time) uniformly without replacement, and if we replace the $(m + 1)^{th}$ sample by t, we get the first scenario conditioned on $\widetilde{m} = m$ and t being sampled. This idea still works in the new setting where the sampling probabilities are no longer the same, since there is still a "draw-by-draw" selection procedure for drawing samples from A (one at a time) in a way so that right after drawing the j^{th} sample, the distribution of samples we currently have is the same as if we have conditioned on $\widetilde{m} = j$ (e.g., see Section 3 in [CDL94]).

We now consider Lemma 15, which protects the privacy of individuals that blend with few people in the population. The extension of Lemma 15 to robust sampling redefines what is meant by "few people", since even if an individual blends with few people, many of them could be sampled with probability 1. With this modification, the proof of the extended lemma is similar to the proof of the original lemma.

When we prove the extended theorem using the extended lemmas, when we are trying to show that privacy holds for individual t, we look at how many people blend with t that are sampled with probability in $[p, p']$ (in particular, we exclude the ℓ people that are sampled with probability outside of $\{0\} \cup [p, p']$); similar to before, if this number is large, we use the extended version of Lemma 14; otherwise, we use the extended version of Lemma 15.

We now give the formal proof of Theorem 16. We begin by proving a lemma about the smoothness of the Poisson binomial distribution[6] near its expectation, which will be used later in the proof of Lemma 18.

Lemma 17 (Smoothness of the Poisson binomial distribution near its expectation)**.** *Let $\mathcal{P}$ be any population, $0 < p \leq p' < 1$, $\pi : X \to [0, 1]$ be any function, and $\epsilon_{Sam} > 0$. Let A be any non-empty subset of $\mathcal{P}$ such that $\pi(a) \in [p, p']$ for every $a \in A$. Let $\widetilde{D} = Sam(\mathcal{P}, \pi)$, $\widetilde{m} = |\widetilde{D} \cap A|$, $n = |A|$, and $\bar{p} = \frac{1}{n}\sum_{a \in A} \pi(a)$. Then, for every integer $m \in \{0, \ldots, n-1\}$, we have the following:*

- *If $m + 1 \leq (n + 1)\bar{p} \cdot \frac{e^{\epsilon_{Sam}}}{\bar{p}e^{\epsilon_{Sam}} + (1-\bar{p})}$, then $\Pr[\widetilde{m} = m] \leq \frac{p'}{p}\frac{1-p}{1-p'}e^{\epsilon_{Sam}} \Pr[\widetilde{m} = m + 1]$.*
- *If $m + 1 \geq (n + 1)\bar{p} \cdot \frac{1}{\bar{p} + (1-\bar{p})e^{\epsilon_{Sam}}}$, then $\Pr[\widetilde{m} = m] \geq \frac{p}{p'}\frac{1-p'}{1-p}e^{-\epsilon_{Sam}} \Pr[\widetilde{m} = m + 1]$.*

Proof. Fix $m \in \{0, \ldots, n-1\}$. Given an individual i in $\mathcal{P}$, let $p_i = \pi(i)$, and let $w_i = \frac{p_i}{1-p_i}$. Given any set A' of individuals in $\mathcal{P}$ and any integer m', let $Q(A', m') = \sum_{B \subseteq A', |B| = m'} \prod_{i \in B} w_i$. Then, we have

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m + 1]} = \frac{\sum_{B \subseteq A, |B| = m}(\prod_{i \in B} p_i)(\prod_{i \in A \setminus B}(1 - p_i))}{\sum_{B \subseteq A, |B| = m+1}(\prod_{i \in B} p_i)(\prod_{i \in A \setminus B}(1 - p_i))} = \frac{Q(A, m)}{Q(A, m + 1)}. \tag{1}$$

We will show that for every $j \in A$,

$$\frac{\partial}{\partial w_j}\left(\frac{Q(A, m)}{Q(A, m + 1)}\right) \leq 0. \tag{2}$$

[6]The Poisson binomial distribution is the distribution of the sum of independent Bernoulli random variables, where the success probabilities in the Bernoulli random variables are not necessarily the same.

Fix $j \in A$. We note that for any integer m', we have $\frac{\partial}{\partial w_j} Q(A, m') = Q(A \setminus \{j\}, m' - 1)$ and $Q(A, m') = Q(A \setminus \{j\}, m') + w_j Q(A \setminus \{j\}, m' - 1)$. Then, we observe that

$$\frac{\partial}{\partial w_j}\left(\frac{Q(A,m)}{Q(A,m+1)}\right) = \frac{Q(A,m+1)\cdot Q(A\setminus\{j\},m-1) - Q(A,m)\cdot Q(A\setminus\{j\},m)}{Q(A,m+1)^2}.$$

Now, using the equalities $Q(A, m+1) = Q(A \setminus \{j\}, m+1) + w_j Q(A \setminus \{j\}, m)$ and $Q(A, m) = Q(A \setminus \{j\}, m) + w_j Q(A \setminus \{j\}, m-1)$, we get

$$\frac{\partial}{\partial w_j}\left(\frac{Q(A,m)}{Q(A,m+1)}\right) = \frac{Q(A\setminus\{j\},m+1)\cdot Q(A\setminus\{j\},m-1) - Q(A\setminus\{j\},m)\cdot Q(A\setminus\{j\},m)}{Q(A,m+1)^2}. \tag{3}$$

We will show that this expression is at most 0 by showing that the numerator $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1) - Q(A \setminus \{j\}, m) \cdot Q(A \setminus \{j\}, m)$ is at most 0. If $m = 0$, then $Q(A \setminus \{j\}, m-1) = 0$, so the numerator is clearly at most 0. Thus, we now assume $m \geq 1$. Consider the full expansion of $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1)$ and $Q(A \setminus \{j\}, m) \cdot Q(A \setminus \{j\}, m)$. Each term of both expansions is of the form $w_{i_1}^2 \cdots w_{i_r}^2 w_{j_1} \cdots w_{j_s}$, where the indices $i_1, \ldots, i_r, j_1, \ldots, j_s$ are all distinct, and $2r + s = 2m$. For example, a term $w_{i_1}^2 \cdots w_{i_r}^2 w_{j_1} \cdots w_{j_s}$ that appears in the expansion of $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1)$ is obtained if both $Q(A \setminus \{j\}, m+1)$ and $Q(A \setminus \{j\}, m-1)$ choose $w_{i_1}, \ldots, w_{i_r}$, $Q(A \setminus \{j\}, m+1)$ chooses $m+1-r$ of the factors $w_{j_1}, \ldots, w_{j_s}$, and $Q(A \setminus \{j\}, m-1)$ chooses the remaining factors in $w_{j_1}, \ldots, w_{j_s}$.

Now, consider a term of the form $w_{i_1}^2 \cdots w_{i_r}^2 w_{j_1} \cdots w_{j_s}$, where the indices $i_1, \ldots, i_r, j_1, \ldots, j_s$ are all distinct, and $2r + s = 2m$. It suffices to show that the number of times this term appears in (the full expansion of) $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1)$ is at most the number of times it appears in $Q(A \setminus \{j\}, m) \cdot Q(A \setminus \{j\}, m)$. If $r > m-1$, then this term appears 0 times in $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1)$, since $Q(A \setminus \{j\}, m-1)$ needs to choose more than $m-1$ factors in $w_{i_1}, \ldots, w_{i_r}$ but it can only choose at most $m-1$; thus, the numerator in (3) is at most 0, as required. If $r \leq m-1$, then this term appears $\binom{s}{m+1-r}$ times in $Q(A \setminus \{j\}, m+1) \cdot Q(A \setminus \{j\}, m-1)$ and $\binom{s}{m-r}$ times in $Q(A \setminus \{j\}, m) \cdot Q(A \setminus \{j\}, m)$. Now, we note that $\binom{s}{m+1-r} \leq \binom{s}{m-r}$, since $s = 2(m-r)$ and $\binom{2(m-r)}{m+1-r} \leq \binom{2(m-r)}{m-r}$, as required.

Now, from (1),(2), and the fact that $\pi(a) \in [p, p']$ for every $a \in A$, it follows that

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m+1]} = \frac{Q(A,m)}{Q(A,m+1)} \leq \frac{\sum_{B \subseteq A, |B|=m} \prod_{i \in B} \frac{p}{1-p}}{\sum_{B \subseteq A, |B|=m+1} \prod_{i \in B} \frac{p}{1-p}} = \frac{1-p}{p}\frac{m+1}{n-m} \tag{4}$$

and

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m+1]} = \frac{Q(A,m)}{Q(A,m+1)} \geq \frac{\sum_{B \subseteq A, |B|=m} \prod_{i \in B} \frac{p'}{1-p'}}{\sum_{B \subseteq A, |B|=m+1} \prod_{i \in B} \frac{p'}{1-p'}} = \frac{1-p'}{p'}\frac{m+1}{n-m}. \tag{5}$$

If $m+1 \leq (n+1)\bar{p} \cdot \frac{e^{\epsilon_{Sam}}}{\bar{p}e^{\epsilon_{Sam}} + (1-\bar{p})}$, then from (4) we have

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m+1]} \leq \frac{1-p}{p}\frac{m+1}{n-m} = \frac{1-p}{p}\frac{\bar{p}}{1-\bar{p}}\left(\frac{1-\bar{p}}{\bar{p}}\frac{m+1}{n-m}\right) \leq \frac{\bar{p}}{p}\frac{1-p}{1-\bar{p}}e^{\epsilon_{Sam}} \leq \frac{p'}{p}\frac{1-p}{1-p'}e^{\epsilon_{Sam}}.$$

If $m+1 \geq (n+1)\bar{p} \cdot \frac{1}{\bar{p} + (1-\bar{p})e^{\epsilon_{Sam}}}$, then from (5) we have

$$\frac{\Pr[\widetilde{m} = m]}{\Pr[\widetilde{m} = m+1]} \geq \frac{1-p'}{p'}\frac{m+1}{n-m} = \frac{1-p'}{p'}\frac{\bar{p}}{1-\bar{p}}\left(\frac{1-\bar{p}}{\bar{p}}\frac{m+1}{n-m}\right) \geq \frac{\bar{p}}{p'}\frac{1-p'}{1-\bar{p}}e^{-\epsilon_{Sam}} \geq \frac{p}{p'}\frac{1-p'}{1-p}e^{-\epsilon_{Sam}}.$$

□

We now prove a lemma that essentially says that if an individual blends with many people in the population, then the individual's privacy is protected when we robustly sample from the population and run San on the samples. This lemma is essentially the extension of Lemma 14 to robust sampling.

Lemma 18 (Protection of individuals that blend with many people in the population that have a good sampling probability)**.** *Let San be any mechanism, $\mathcal{P}$ be any population, $0 < p \le p' < 1$, $\pi : X \to [0,1]$ be any function, and $\epsilon \ge 0$. Let t be any individual in $\mathcal{P}$, and let A be any non-empty subset of $\mathcal{P} \setminus \{t\}$ such that for every individual $t' \in A$, we have $t' \approx_{\epsilon, San} t$ and $\pi(t') \in [p, p']$. Let $n = |A|$, $p_t = \pi(t)$, and $\bar{p} = \frac{1}{n}\sum_{t' \in A} \pi(t')$. Then, we have*

$$San(Sam(\mathcal{P}, \pi)) \approx_{\epsilon_{final}, \delta_{final}} San(Sam(\mathcal{P} \setminus \{t\}, \pi)),$$

where $\epsilon_{final} = \ln(p_t \cdot (\frac{p'}{p}\frac{(1-p)(2-p)}{(1-p')^2}e^{\epsilon}) + (1-p_t))$ and $\delta_{final} = \max\{\frac{p_t}{p}, \frac{p_t}{1-p'}\} \cdot e^{-\Omega((n+1)\bar{p}(1-\bar{p})^2)}$.

Proof. Let $\epsilon_{Sam} > 0$, $\widehat{D} = Sam(\mathcal{P}, \pi)$, $\widetilde{D} = Sam(\mathcal{P} \setminus \{t\}, \pi)$, $\widetilde{m} = |\widetilde{D} \cap A|$, and $Y \subseteq \{0,1\}^*$. Let E be the event that t is sampled when $\widehat{D}$ is chosen.

We first show that for every $m \in \{0, \ldots, n-1\}$, we have

$$\left| \ln \left(\frac{\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m]}{\Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1]} \right) \right| \le \epsilon. \tag{1}$$

It is known that there exists a "draw-by-draw" selection procedure for drawing samples from A (one at a time) such that right after drawing the j^{th} sample, the samples chosen so far has the same distribution as the conditional distribution of $Sam(A, \pi)$ given $|Sam(A, \pi)| = j$ (e.g., see Section 3 in [CDL94]). More formally, there exists a vector of random variables $(X_1, \ldots, X_n)$ jointly distributed over A^n such that for every $j \in [n]$, $\{X_1, \ldots, X_j\}$ has the same distribution as the conditional distribution of $Sam(A, \pi)$ given $|Sam(A, \pi)| = j$.

Now, fix $m \in \{0, \ldots, n-1\}$. Let $\mathcal{D}_m = Sam(\mathcal{P} \setminus (A \cup \{t\}), \pi) \cup \{X_1, \ldots, X_m\}$. Then, for every $D \subseteq \mathcal{P}$, we have $\Pr[\widetilde{D} \cup \{t\} = D \mid \widetilde{m} = m] = \Pr[\mathcal{D}_m \cup \{t\} = D]$ and $\Pr[\widetilde{D} = D \mid \widetilde{m} = m+1] = \Pr[\mathcal{D}_m \cup \{X_{m+1}\} = D]$. Then, using the fact that $t' \approx_{\epsilon, San} t$ for every individual $t' \in A$, we have

$$\begin{aligned} &\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m] = \Pr[San(\mathcal{D}_m \cup \{t\}) \in Y] \\ &\le e^{\epsilon} \Pr[San(\mathcal{D}_m \cup \{X_{m+1}\}) \in Y] = e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1]. \end{aligned}$$

Similarly, we also have

$$\begin{aligned} &\Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1] = \Pr[San(\mathcal{D}_m \cup \{X_{m+1}\}) \in Y] \\ &\le e^{\epsilon} \Pr[San(\mathcal{D}_m \cup \{t\}) \in Y] = e^{\epsilon} \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m]. \end{aligned}$$

Thus, we have shown (1).

Now, we observe that

$$\begin{aligned} \Pr[San(\widehat{D}) \in Y] &= \Pr[San(\widehat{D}) \in Y \mid E] \cdot \Pr[E] + \Pr[San(\widehat{D}) \in Y \mid \overline{E}] \cdot \Pr[\overline{E}] \\ &= \Pr[San(\widetilde{D} \cup \{t\}) \in Y] \cdot p_t + \Pr[San(\widetilde{D}) \in Y] \cdot (1 - p_t). \end{aligned} \tag{2}$$

Let $\alpha = \frac{e^{\epsilon_{Sam}}}{\bar{p}e^{\epsilon_{Sam}} + (1-\bar{p})}$ and $\beta = \frac{1}{\bar{p} + (1-\bar{p})e^{\epsilon_{Sam}}}$, and let $\delta_{Sam} = \max\{\Pr[\widetilde{m} + 1 > (n+1)\bar{p} \cdot \alpha], \Pr[\widetilde{m} < (n+1)\bar{p} \cdot \beta]\}$. By Lemma 17 and (1) (and the fact that $m = n$ does not satisfy $m+1 \le (n+1)\bar{p} \cdot \alpha$),

we have

$$\begin{aligned}
&\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \\
\leq & \sum_{\substack{m \in \{0,\ldots,n\} \\ m+1 \leq (n+1)\bar{p}\cdot\alpha}} \Pr[\widetilde{m} = m] \cdot \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m] + \Pr[\widetilde{m} + 1 > (n+1)\bar{p} \cdot \alpha] \\
\leq & \sum_{\substack{m \in \{0,\ldots,n\} \\ m+1 \leq (n+1)\bar{p}\cdot\alpha}} \frac{p'}{p}\frac{1-p}{1-p'} e^{\epsilon_{Sam}} \Pr[\widetilde{m} = m+1] \cdot e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1] + \delta_{Sam} \\
\leq & \frac{p'}{p}\frac{1-p}{1-p'} e^{\epsilon+\epsilon_{Sam}} \Pr[San(\widetilde{D}) \in Y] + \delta_{Sam} \qquad (3)
\end{aligned}$$

and

$$\begin{aligned}
&\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \\
\geq & \sum_{\substack{m \in \{0,\ldots,n-1\} \\ m+1 \geq (n+1)\bar{p}\cdot\beta}} \Pr[\widetilde{m} = m] \cdot \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} = m] \\
\geq & \sum_{\substack{m \in \{0,\ldots,n-1\} \\ m+1 \geq (n+1)\bar{p}\cdot\beta}} \frac{p}{p'}\frac{1-p'}{1-p} e^{-\epsilon_{Sam}} \Pr[\widetilde{m} = m+1] \cdot e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} = m+1] \\
\geq & (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) \cdot (\Pr[San(\widetilde{D}) \in Y] - \Pr[\widetilde{m} < (n+1)\bar{p} \cdot \beta]) \\
\geq & (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) \cdot \Pr[San(\widetilde{D}) \in Y] - (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) \cdot \delta_{Sam}. \qquad (4)
\end{aligned}$$

Let $\epsilon_{total} = \ln(p_t \cdot (\frac{p'}{p}\frac{1-p}{1-p'} e^{\epsilon+\epsilon_{Sam}}) + (1-p_t))$. Now, combining (2) and (3), we have

$$\begin{aligned}
\Pr[San(\widehat{D}) \in Y] &\leq (p_t \cdot (\frac{p'}{p}\frac{1-p}{1-p'} e^{\epsilon+\epsilon_{Sam}}) + (1-p_t)) \Pr[San(\widetilde{D}) \in Y] + p_t \cdot \delta_{Sam} \\
&= e^{\epsilon_{total}} \Pr[San(\widetilde{D}) \in Y] + p_t \cdot \delta_{Sam}.
\end{aligned}$$

Combining (2) and (4), we also have

$$\begin{aligned}
&\Pr[San(\widehat{D}) \in Y] \geq (p_t \cdot (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) + (1-p_t)) \Pr[San(\widetilde{D}) \in Y] - p_t \cdot (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) \cdot \delta_{Sam} \\
&\implies \Pr[San(\widetilde{D}) \in Y] \\
&\leq \frac{1}{p_t \cdot (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) + (1-p_t)} \Pr[San(\widehat{D}) \in Y] + \frac{\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}}{p_t \cdot (\frac{p}{p'}\frac{1-p'}{1-p} e^{-(\epsilon+\epsilon_{Sam})}) + (1-p_t)} p_t \cdot \delta_{Sam} \\
&\leq (p_t \cdot (\frac{p'}{p} \cdot \frac{1-p}{1-p'} \cdot e^{\epsilon+\epsilon_{Sam}}) + (1-p_t)) \Pr[San(\widehat{D}) \in Y] + p_t \cdot \delta_{Sam} \\
&= e^{\epsilon_{total}} \Pr[San(\widehat{D}) \in Y] + p_t \cdot \delta_{Sam},
\end{aligned}$$

where the last inequality follows from the fact that the function $f(x) = \frac{1}{x}$ is convex for $x > 0$. Thus, we have $San(\widehat{D}) \approx_{\epsilon_{total}, p_t \cdot \delta_{Sam}} San(\widetilde{D})$.

Now, we set $\epsilon_{Sam} = \ln(\frac{2-\bar{p}}{1-\bar{p}})$. Then, we have

$$\epsilon_{total} = \ln(p_t \cdot (\frac{p'}{p}\frac{(1-p)(2-\bar{p})}{(1-p')(1-\bar{p})}e^{\epsilon}) + (1-p_t)) \leq \ln(p_t \cdot (\frac{p'}{p}\frac{(1-p)(2-p)}{(1-p')^2}e^{\epsilon}) + (1-p_t)) = \epsilon_{final}$$

and

$$\begin{aligned}
p_t \cdot \delta_{Sam} &= p_t \cdot \max\{\Pr[\widetilde{m}+1 > (n+1)\bar{p} \cdot \frac{e^{\epsilon_{Sam}}}{\bar{p}e^{\epsilon_{Sam}} + (1-\bar{p})}], \Pr[\widetilde{m} < (n+1)\bar{p} \cdot \frac{1}{\bar{p} + (1-\bar{p})e^{\epsilon_{Sam}}}]\} \\
&= p_t \cdot \max\{\Pr[\widetilde{m}+1 > (n+1)\bar{p} \cdot (2-\bar{p})], \Pr[\widetilde{m} < (n+1)\bar{p} \cdot \frac{1}{2}]\} \\
&\leq p_t \cdot \max\{\frac{1}{\bar{p}}\Pr[\widetilde{m} + Bin(1,\bar{p}) > (n+1)\bar{p} \cdot (2-\bar{p})], \frac{1}{1-\bar{p}}\Pr[\widetilde{m} + Bin(1,\bar{p}) < (n+1)\bar{p} \cdot \frac{1}{2}]\} \\
&\leq p_t \cdot \max\{\frac{1}{\bar{p}}e^{-\Omega((n+1)\bar{p}(1-\bar{p})^2)}, \frac{1}{1-\bar{p}}e^{-\Omega((n+1)\bar{p})}\} \\
&\leq \max\{\frac{p_t}{p}, \frac{p_t}{1-p'}\} \cdot e^{-\Omega((n+1)\bar{p}(1-\bar{p})^2)} \\
&= \delta_{final},
\end{aligned}$$

where $Bin(1,\bar{p})$ is a binomial random variable with 1 trial and success probability $\bar{p}$, and the second last inequality follows from multiplicative Chernoff bounds. □

We now show how pre-sampling combined with a crowd-blending private mechanism can protect the privacy of individuals who blend with few people in the population. The following lemma is essentially the extension of Lemma 15 to robust sampling. This lemma is stated in a somewhat more general form that allows us to use it to prove Theorem 16 later.

Lemma 19 (Protection of individuals that blend with few people in the population)**.** *Let San be any (k,ϵ)-crowd-blending private mechanism with $k \geq 2$, let $\mathcal{P}$ be any population, and let $\pi : X \to [0,1]$ be any function. Let t be any individual in $\mathcal{P}$, and let A be any non-empty subset of $\mathcal{P} \setminus \{t\}$ such that for every individual $t' \in A$, we have $t' \approx_{\epsilon,San} t$. Let $n = |A|$, $s = |\{t' \in \mathcal{P} \setminus \{t\} : t' \approx_{\epsilon,San} t \text{ and } t' \notin A\}|$, $p_t = \pi(t)$, and $\bar{p} = \frac{1}{n}\sum_{t' \in A}\pi(t')$. Then, if $s < k-1$, $\bar{p} > 0$, and $n \leq \frac{k-s-1}{\bar{p}(2-\bar{p})}$, then we have*

$$San(Sam(\mathcal{P},\pi)) \approx_{\epsilon_{final},\delta_{final}} San(Sam(\mathcal{P} \setminus \{t\},\pi))$$

where $\epsilon_{final} = \ln(p_t e^{\epsilon} + (1-p_t))$ and $\delta_{final} = p_t e^{-\Omega((k-s)\cdot(1-\bar{p})^2)}$.

Proof. Suppose $s < k-1$, $\bar{p} > 0$, and $n \leq \frac{k-s-1}{\bar{p}(2-\bar{p})}$. Let $\widehat{D} = Sam(\mathcal{P},\pi)$, $\widetilde{D} = Sam(\mathcal{P} \setminus \{t\},\pi)$, and $Y \subseteq \{0,1\}^*$. Let $\widetilde{m} = |\widetilde{D} \cap A|$, and let E be the event that individual t is in $\widehat{D}$ when $\widehat{D}$ is chosen. We first note that

$$\begin{aligned}
\Pr[San(\widehat{D}) \in Y] &= \Pr[San(\widehat{D}) \in Y \mid \overline{E}] \cdot \Pr[\overline{E}] + \Pr[San(\widehat{D}) \in Y \mid E] \cdot \Pr[E] \\
&= \Pr[San(\widetilde{D}) \in Y] \cdot (1-p_t) + \Pr[San(\widetilde{D} \cup \{t\}) \in Y] \cdot p_t. \qquad (1)
\end{aligned}$$

We note that if $\widetilde{m} < k-s-1$, then t ϵ-blends with fewer than k people in $\widetilde{D} \cup \{t\}$, and since San is (k,ϵ)-crowd-blending private, we have

$$\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-s-1] \leq e^{\epsilon}\Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-s-1]$$

and

$$\Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-s-1] \geq e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-s-1].$$

Then, we have

$$\begin{aligned}
\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \leq\ & \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-s-1] \Pr[\widetilde{m} < k-s-1] + \Pr[\widetilde{m} \geq k-s-1] \\
&\leq e^{\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-s-1] \Pr[\widetilde{m} < k-s-1] + \Pr[\widetilde{m} \geq k-s-1] \\
&\leq e^{\epsilon} \Pr[San(\widetilde{D}) \in Y] + \Pr[\widetilde{m} \geq k-s-1], \qquad (2)
\end{aligned}$$

and

$$\begin{aligned}
\Pr[San(\widetilde{D} \cup \{t\}) \in Y] \geq\ & \Pr[San(\widetilde{D} \cup \{t\}) \in Y \mid \widetilde{m} < k-s-1] \Pr[\widetilde{m} < k-s-1] \\
&\geq e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y \mid \widetilde{m} < k-s-1] \Pr[\widetilde{m} < k-s-1] \\
&\geq e^{-\epsilon} (\Pr[San(\widetilde{D}) \in Y] - \Pr[\widetilde{m} \geq k-s-1]) \\
&= e^{-\epsilon} \Pr[San(\widetilde{D}) \in Y] - e^{-\epsilon} \Pr[\widetilde{m} \geq k-s-1]. \qquad (3)
\end{aligned}$$

Now, combining (1) and (2), we have

$$\Pr[San(\widehat{D}) \in Y] \leq (p_t e^{\epsilon} + (1-p_t)) \Pr[San(\widetilde{D}) \in Y] + p_t \Pr[\widetilde{m} \geq k-s-1]. \qquad (4)$$

Also, combining (1) and (3), we have

$$\Pr[San(\widehat{D}) \in Y] \geq (p_t e^{-\epsilon} + (1-p_t)) \Pr[San(\widetilde{D}) \in Y] - e^{-\epsilon} p_t \Pr[\widetilde{m} \geq k-s-1].$$

Rearranging this inequality, we get

$$\begin{aligned}
\Pr[San(\widetilde{D}) \in Y] &\leq \frac{1}{p_t e^{-\epsilon} + (1-p_t)} \Pr[San(\widehat{D}) \in Y] + \frac{e^{-\epsilon}}{p_t e^{-\epsilon} + (1-p_t)} p_t \Pr[\widetilde{m} \geq k-s-1] \\
&\leq (p_t e^{\epsilon} + (1-p_t)) \Pr[San(\widehat{D}) \in Y] + p_t \Pr[\widetilde{m} \geq k-s-1], \qquad (5)
\end{aligned}$$

where the last inequality follows from the fact that the function $f(x) = \frac{1}{x}$ is convex for $x > 0$, so $\frac{1}{p_t e^{-\epsilon} + (1-p_t)} \leq p_t e^{\epsilon} + (1-p_t)$.

Let $\tau = \frac{k-s-1}{\bar{p}(2-\bar{p})}$. Then, we have $n \leq \tau$. The lemma now follows from (4), (5), and the inequality

$$\begin{aligned}
p_t \Pr[\widetilde{m} \geq k-s-1] &= p_t \Pr[\widetilde{m} \geq \tau\bar{p} \cdot (2-\bar{p})] \\
&\leq p_t \Pr[\widetilde{m} + Bin(\lfloor\tau\rfloor - n, \bar{p}) + Bin(1, (\tau - \lfloor\tau\rfloor)\bar{p}) \geq \tau\bar{p} \cdot (2-\bar{p})] \\
&\leq p_t e^{-\Omega(\tau\bar{p}(1-\bar{p})^2)} \\
&\leq p_t e^{-\Omega((k-s)(1-\bar{p})^2)},
\end{aligned}$$

where $Bin(j, q)$ denotes a binomial random variable with j trials and success probability q, and the second inequality follows from a multiplicative Chernoff bound (note that the expectation of $\widetilde{m} + B(\lfloor\tau\rfloor - n, \bar{p}) + B(1, (\tau - \lfloor\tau\rfloor)\bar{p})$ is $\tau\bar{p}$). □

Using the new lemmas (Lemmas 18 and 19), we can now prove Theorem 16 in a way similar to Theorem 13.

Proof of Theorem 16. We first note that $Sam(\cdot,\pi) \in iRS(p,p',l)$. Thus, we can let $T = Sam(\cdot,\pi)$ in the definition of zero-knowledge privacy with respect to $iRS(p,p',l)$. Let A be any adversary. We will describe how to construct a simulator S for A. Let $\mathcal{P}$ be any population, t be any individual in $\mathcal{P}$, and $z \in \{0,1\}^*$. Since the simulator S is given $T(\mathcal{P} \setminus \{t\}) = Sam(\mathcal{P} \setminus \{t\}, \pi)$ and z as part of its input, S can easily compute $San_{zk}(\mathcal{P} \setminus \{t\}) = San(Sam(\mathcal{P} \setminus \{t\}, \pi))$ and then simulate the computation of the adversary A that is given $San_{zk}(\mathcal{P} \setminus \{t\})$ and the auxiliary input z; the simulator S then outputs whatever A outputs.

Now, we note that if $San_{zk}(\mathcal{P}) \approx_{\epsilon_{zk},\delta_{zk}} San_{zk}(\mathcal{P} \setminus \{t\})$, then $Out_A(A(z) \leftrightarrow San_{zk}(\mathcal{P})) \approx_{\epsilon_{zk},\delta_{zk}} S(z, T(\mathcal{P} \setminus \{t\}), |\mathcal{P}|)$. Thus, to show that San_{zk} is $(\epsilon_{zk},\delta_{zk})$-zero-knowledge private with respect to $iRS(p,p',l)$, it suffices to show that $San_{zk}(\mathcal{P}) \approx_{\epsilon_{zk},\delta_{zk}} San_{zk}(\mathcal{P} \setminus \{t\})$, i.e.,

$$San(Sam(\mathcal{P},\pi)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P} \setminus \{t\}, \pi)).$$

To this end, let $A = \{t' \in \mathcal{P} \setminus \{t\} : t' \approx_{\epsilon,San} t \text{ and } \pi(t') \in [p,p']\}$, $n = |A|$, $p_t = \pi(t)$, $\bar{p} = \frac{1}{n}\sum_{t' \in A} \pi(t')$, and $s = |\{t' \in \mathcal{P} \setminus \{t\} : t' \approx_{\epsilon,San} t \text{ and } t' \notin A\}|$. It is easy to see that without loss of generality, we can assume that $\mathcal{P}$ satisfies the property that $\pi(t') \neq 0$ for every $t' \in \mathcal{P}$. We note that $s \leq l$, which we use later in some of the inequalities below. Let $\tau = \frac{k-s-1}{\bar{p}(2-\bar{p})}$. We will consider two cases: $n > \tau$ and $n \leq \tau$.

Suppose $n > \tau$. By Lemma 18, we have

$$San(Sam(\mathcal{P},\pi)) \approx_{\epsilon_1,\delta_1} San(Sam(\mathcal{P} \setminus \{t\}, \pi)),$$

where $\epsilon_1 = \ln(p_t \cdot (\frac{p'}{p}\frac{(1-p)(2-p)}{(1-p')^2}e^{\epsilon}) + (1-p_t)) \leq \ln(p_{\max} \cdot (\frac{p'}{p}\frac{(1-p)(2-p)}{(1-p')^2}e^{\epsilon}) + (1-p_{\max})) = \epsilon_{zk}$ and $\delta_1 = \max\{\frac{p_t}{p}, \frac{p_t}{1-p'}\} \cdot e^{-\Omega((n+1)\bar{p}(1-\bar{p})^2)} \leq \max\{\frac{p_{\max}}{p}, \frac{p_{\max}}{1-p'}\} \cdot e^{-\Omega((k-l)(1-p')^2)} = \delta_{zk}$.

Now, suppose $n \leq \tau$. By Lemma 19, we have

$$San(Sam(\mathcal{P},\pi)) \approx_{\epsilon_2,\delta_2} San(Sam(\mathcal{P} \setminus \{t\}, \pi)),$$

where $\epsilon_2 = \ln(p_t e^{\epsilon} + (1-p_t)) \leq \epsilon_1 \leq \epsilon_{zk}$ and $\delta_2 = p_t e^{-\Omega((k-s)\cdot(1-\bar{p})^2)} \leq p_{\max} e^{-\Omega((k-l)\cdot(1-p')^2)} \leq \delta_{zk}$.

It follows that

$$San(Sam(\mathcal{P},\pi)) \approx_{\epsilon_{zk},\delta_{zk}} San(Sam(\mathcal{P} \setminus \{t\}, \pi)),$$

as required. □

References

[BLR08] Avrim Blum, Katrina Ligett, and Aaron Roth, *A learning theory approach to non-interactive database privacy*, STOC '08: Proc. of the 40th annual ACM symposium on Theory of computing, 2008, pp. 609–618.

[CDL94] Xiang-Hui Chen, Arthur P. Dempster, and Jun S. Liu, *Weighted finite population sampling to maximize entropy*, Biometrika **81** (1994), no. 3, pp. 457–469.

[CDM+05] Shuchi Chawla, Cynthia Dwork, Frank McSherry, Adam Smith, and Hoeteck Wee, *Toward privacy in public databases.*, Second Theory of Cryptography Conference (TCC 2005), 2005, pp. 363–385.

[CKLM09] Bee-Chung Chen, Daniel Kifer, Kristen LeFevre, and Ashwin Machanavajjhala, *Privacy-preserving data publishing*, Foundations and Trends in Databases **2** (2009), no. 1-2, 1–167.

[CM06] Kamalika Chaudhuri and Nina Mishra, *When random sampling preserves privacy*, CRYPTO'06, 2006, pp. 198–213.

[DMNS06] Cynthia Dwork, Frank Mcsherry, Kobbi Nissim, and Adam Smith, *Calibrating noise to sensitivity in private data analysis*, Proc. of the 3rd Theory of Cryptography Conference, 2006, pp. 265–284.

[DNR+09] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil Vadhan, *On the complexity of differentially private data release: efficient algorithms and hardness results*, Proceedings of the 41st annual ACM symposium on Theory of computing, STOC '09, 2009, pp. 381–390.

[DRV10] Cynthia Dwork, Guy Rothblum, and Salil Vadhan, *Boosting and differential privacy*, Proc. of the 51st Annual IEEE Symposium on Foundations of Computer Science, 2010.

[Dwo06] Cynthia Dwork, *Differential privacy*, ICALP, 2006, pp. 1–12.

[Dwo08] Cynthia Dwork, *Differential privacy: A survey of results*, Theory and Applications of Models of Computation, Lecture Notes in Computer Science, vol. 4978, Springer Berlin / Heidelberg, 2008, pp. 1–19.

[Dwo09] C. Dwork, *The differential privacy frontier*, Proc. of the 6th Theory of Cryptography Conference (TCC), 2009.

[FWCY10] Benjamin C. M. Fung, Ke Wang, Rui Chen, and Philip S. Yu, *Privacy-preserving data publishing: A survey of recent developments*, ACM Comput. Surv. **42** (2010), no. 4, 1–53.

[GLP11] Johannes Gehrke, Edward Lui, and Rafael Pass, *Towards privacy for social networks: a zero-knowledge based definition of privacy*, Proceedings of the 8th conference on Theory of cryptography, TCC'11, 2011, pp. 432–449.

[Kif09] Daniel Kifer, *Attacks on privacy and definetti's theorem*, SIGMOD Conference, 2009, pp. 127–138.

[KLN+08] S.P. Kasiviswanathan, H.K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith, *What can we learn privately?*, Foundations of Computer Science, 2008, 2008, pp. 531–540.

[LQS11] Ninghui Li, Wahbeh H. Qardaji, and Dong Su, *Provably private data anonymization: Or, k-anonymity meets differential privacy*, Manuscript, 2011.

[NRS07] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith, *Smooth sensitivity and sampling in private data analysis*, STOC 2007, 2007, pp. 75–84.

[Swe02] Latanya Sweeney, *k-anonymity: a model for protecting privacy*, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **10** (2002), 557–570.

[UV11] Jonathan Ullman and Salil Vadhan, *Pcps and the hardness of generating private synthetic data*, Proceedings of the 8th conference on Theory of cryptography, TCC'11, 2011, pp. 400–416.

[WFWP07] Raymond Chi-Wing Wong, Ada Wai-Chee Fu, Ke Wang, and Jian Pei, *Minimality attack in privacy preserving data publishing*, Proceedings of the 33rd international conference on Very large data bases, VLDB '07, VLDB Endowment, 2007, pp. 543–554.

[ZJB07] Lei Zhang, Sushil Jajodia, and Alexander Brodsky, *Information disclosure under realistic assumptions: privacy versus optimality*, Proceedings of the 14th ACM conference on Computer and communications security, CCS '07, ACM, 2007, pp. 573–583.