Private k-Means Clustering with Stability Assumptions

Moshe Shechner
Ben-Gurion University

Or Sheffet
Bar-Ilan University

Uri Stemmer
Ben-Gurion University

Abstract

We study the problem of differentially private clustering under input-stability assumptions. Despite the ever-growing volume of works on differential privacy in general and differentially private clustering in particular, only three works (Nissim et al., 2007; Wang et al., 2015; Huang and Liu, 2018) looked at the problem of privately clustering "nice" k-means instances, all three relying on the sample-and-aggregate framework and all three measuring utility in terms of Wasserstein distance between the true cluster centers and the centers returned by the private algorithm. In this work we improve upon this line of works on multiple axes. We present a simpler algorithm for clustering stable inputs (not relying on the sample-and-aggregate framework), and analyze its utility in both the Wasserstein distance and the k-means cost. Moreover, our algorithm has straightforward analogues for "nice" k-median instances and for the local-model of differential privacy.

1 Introduction

In recent years differential privacy (Dwork et al., 2006b) has been established as the de-facto gold standard of privacy preserving data analysis. The notion of differential privacy guarantees that any single datum has a limited effect on the outcome of the algorithm, and so it is often presented as a formal notion of robustness. Indeed, it is commonly believed that objectives which are sensitive to the change of a single datapoint are hard to approximate in a differentially private manner. One such notorious example is the median, which may shift drastically by a single datapoint.

And yet, the median is easy to approximate on stable instances (Nissim et al., 2007). In fact, the median problem was the first to be studied in the context of the interplay between input-stability notions and the stability enforced by differential privacy. In fact, at the very same paper, Nissim et al. (2007) gave the first differentially private algorithm for clustering *well-separated* k-means instances, a notion first introduced by Ostrovsky et al. (2012). A k-means clustering instance is called ϕ-well separated (or simply ϕ-separated) if the ratio of the optimal k-means cost to the optimal $(k-1)$-means cost is at most ϕ^2. Following the work of Ostrovsky et al., several other works have related other notions of input-stability to clustering (Balcan et al., 2009; Awasthi et al., 2010; Bilu and Linial, 2010; Kumar and Kannan, 2010; Awasthi et al., 2012).

The construction of differentially private k-means clustering algorithms has attracted a lot of attention over the last 14 years.[1] In particular, three works — the work of Nissim et al. (2007) and two followup papers (Wang et al., 2015; Huang and Liu, 2018) — have constructed private k-means algorithms for stable instances. While several interesting concepts arise from these three works, their algorithms — and more importantly, their analysis — can be tighten up, simplified, and at the same time be applied in a broader setting. Our work does precisely this: we simplify the existing constructions for private clustering on stable instances, while improving upon their analysis and relating the well-separability no-

Proceedings of the 23rdInternational Conference on Artificial Intelligence and Statistics (AISTATS) 2020, Palermo, Italy. PMLR: Volume 108.

[1] Blum et al. (2005); Nissim et al. (2007); Feldman et al. (2009); McSherry (2009); Gupta et al. (2010); Mohan et al. (2012); Wang et al. (2015); Nock et al. (2016); Su et al. (2016); Nissim et al. (2016); Feldman et al. (2017); Balcan et al. (2017); Nissim and Stemmer (2018); Huang and Liu (2018); Kaplan and Stemmer (2018); Stemmer (2020)

tion to other notions of stability, applicable for both the k-means and the k-median objectives. Moreover, we give the first local-differentially private algorithm for clustering well-separated instances.

Before formally presenting our results, we describe our setting more precisely. Consider an input database X containing n points in $\mathbb{R}^d$. In k-means clustering, the goal is to identify a set C of k *centers* in $\mathbb{R}^d$, approximately minimizing the sum of squared distances from each input point to its nearest center, a quantity referred to as the *cost* of the centers. That is,

$$\text{cost}_X(C) = \sum_{x \in X} \min_{c \in C} \|x - c\|^2.$$

We denote the lowest possible cost as $\text{OPT}_k(X)$. As minimizing the k-means objective is NP-hard, the literature has focused on approximation algorithms, with the current (non-private) state-of-the-art achieving a multiplicative error of 6.357 (Ahmadian et al., 2017). That is, their algorithm identifies a set of k centers whose cost is no more than 6.357 times the lowest possible cost. Furthermore, for *well-separated instances*, the works mentioned above obtain significantly improved guarantees, with error arbitrarily close to 1 (non-privately).[2]

In our context, every input point $x \in X$ is assumed to be the (private) information of one individual (such as a location or a text file), and we would like to identify a set of centers C with low cost while at the same time providing differential privacy for the points in X.

Definition 1.1 (Dwork et al. (2006c))**.** *A randomized algorithm $\mathcal{A} : \mathcal{X}^n \to Y$ is (ε, δ)* differentially private *if for every two databases $X, X' \in \mathcal{X}^n$ that differ in one row, and every set $T \subseteq Y$, we have*

$$\Pr[\mathcal{A}(X) \in T] \leq e^{\varepsilon} \cdot \Pr[\mathcal{A}(X') \in T] + \delta.$$

Unlike in the non-private literature, it is known that every *private* algorithm for approximating the k-means must have an *additive* error (even computationally unbounded algorithms), which scales with the diameter of the input space. Hence, a standard assumption for private k-means is that the input points come from the d-dimensional ball of radius Λ around the origin $\mathcal{B}(0, \Lambda)$. This is the setting we consider in this work, where we fix $\Lambda = 1$ for the introduction. As private k-means algorithms have both multiplicative and additive errors, different guarantees can easily be incomparable. Typically (though not always), one aims to minimize the multiplicative error while keeping the additive error at most polylogarithmic in the size of the database (note that an additive error of size $|X|$ is meaningless). The current state-of-the-art construction for private k-means by Kaplan and Stemmer (2018) obtains $O(1)$ multiplicative error and $\text{poly}(\log(n), k, d)$ additive error.

Given the success of (non-private) stability-based clustering algorithms, it is not surprising that such stability assumptions were also utilized in the privacy literature, specifically by Nissim et al. (2007); Wang et al. (2015); Huang and Liu (2018). However, the error measure pursued in these three works is different. Instead of aiming to find k centers with low k-means cost, these three works aim to find centers that are close to the optimal centers in terms of the *Wasserstein distance*, defined as follows.

Definition 1.2 (Wasserstein (1969))**.** *Let $C = (c_1, \ldots, c_k) \in (\mathbb{R}^d)^k$ and $\hat{C} = (\hat{c}_1, \ldots, \hat{c}_k) \in (\mathbb{R}^d)^k$ be two sets of centers. The* Wasserstein distance *between C and $\hat{C}$ is the L_2^{dk} distance under the best possible permutation π of the centers in each set.*

Nissim et al. (2007) presented a private algorithm that, for a ϕ-separated instance, computes k centers of Wasserstein distance at most $O(\frac{k\sqrt{d}}{\varepsilon} \cdot \phi^2)$ from the optimal k-means centers. Wang et al. (2015) extended the results of Nissim et al. to *subspace clustering*[3] with similar error bounds. Finally, Huang and Liu (2018), presented a clever algorithm that reduced the error down to $O(\phi^2)$ – a significant improvement over the previous error bounds of Nissim et al. (2007) and Wang et al. (2015). In addition, Huang and Liu (2018) showed that their error bound is tight, and that Wasserstein distance of $O(\phi^2)$ is the best possible under differential privacy (for ϕ-separated instances).

We comment that even though a set of centers $\hat{C}$ might be close to the optimal centers C in terms of the Wasserstein distance, say $d_{\mathsf{W}}(C, \hat{C}) = \gamma$, the k-means cost of $\text{cost}_X(\hat{C})$ might be as big as $\text{OPT}_k(X) + |X| \cdot \gamma^2$. That is, the additive error obtained by translating a bound on the Wasserstein distance to a bound on the k-means cost scales with $|X|$. In this work we are aiming for an additive error of at most $\text{polylog}\,|X|$, which means that approximation guarantees w.r.t. the Wasserstein distance

[2] Ostrovsky et al. (2012); Balcan et al. (2009); Awasthi et al. (2010); Bilu and Linial (2010); Kumar and Kannan (2010); Awasthi et al. (2012)

[3] In subspace clustering we aim to group the data points into clusters so that data points in a single cluster lie approximately on a low-dimensional linear subspace.

do not imply (in general) satisfactory approximation guarantees w.r.t. the k-means cost.

Our Contribution and Organization. First, we establish equivalence between several notions of input-stability for clustering problems. This result is given in the preliminaries, Section 2, and should come as no surprise considering all of these notions (and others) yield a PTAS for the clustering problem (Ostrovsky et al., 2012; Awasthi et al., 2010). Second, we present our – absurdly simple – private algorithm for clustering well-separated instances in Section 3, which can be summarized as follows: run an arbitrary (private) k-means approximation algorithm and then take a Lloyd-step (averaging only the points with clear preference for one center over all others). We give a short proof arguing that the result of applying an algorithm with a worst-case guarantee of v-approximation[4] to the k-means objective on a ϕ-well separable instance is (effectively) a $(1+O(\phi^2))$-approximation, provided v is small in comparison to ϕ^{-2}. We obtain the following theorem.

Theorem 1.3 (informal)**.** *There exists an (ε,δ)-differentially private algorithm such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0,1)$, and assume that X is ϕ-separated for k-means for $\phi = O(1)$ (sufficiently small). When applied to X, the algorithm returns (w.h.p.) a set of k centers C' satisfying*

$$\mathrm{cost}_X(C') \leq (1+O(\phi^2)) \cdot \mathrm{OPT}_k(X) + \Delta,$$

for $\Delta \lesssim \frac{k(\sqrt{d}+\sqrt{k})}{\varepsilon}$.

We analyze this algorithm's utility also in terms of Wasserstein distance, as follows.

Theorem 1.4 (informal)**.** *There exists an (ε,δ)-differentially private algorithm such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0,1)$. Assume that X is ϕ-separated for k-means for $\phi = O(1)$ (sufficiently small), and assume that $\mathrm{OPT}_{k-1}(X) \gtrsim \frac{k(\sqrt{d}+\sqrt{k})}{\phi^2 \cdot \varepsilon}$.*

When applied to X, the algorithm returns (w.h.p.) a set of k centers C' satisfying

$$d_{\mathsf{W}}(C, C') \leq O(\phi^2),$$

where C are the optimal k-means centers.

The error bound in this theorem matches the state-of-the-art result of Huang and Liu (2018), and offers some improvements in terms of the requirement on $\mathrm{OPT}_{k-1}(X)$.[5]

Due to the simplicity of our algorithm, we give its local-model analogue and k-median analogue in Sections 5 and 4 respectively. This is the first locally-private clustering algorithm for well-separated instances (even w.r.t. the Wasserstein distance). Unlike our algorithm, previous constructions (Nissim et al., 2007; Wang et al., 2015; Huang and Liu, 2018) are based on the *sample-and-aggregate* framework (Nissim et al., 2007), which is inapplicable (in general) in the local-model. It is hence unclear whether these previous constructions have analogues for the local-model.

2 Preliminaries

We require the following two folklore lemmas. The first lemma quantifies the 1-means cost of a center $\hat{c}$ in terms of its distance from the optimal center (for a proof see, e.g., (Awasthi, 2013, Fact 2.3.1)).

Lemma 2.1. *Let $X \in (\mathbb{R}^d)^n$ and let c denote the average of X. For any $\hat{c} \in \mathbb{R}^d$ it holds that*

$$\sum_{x \in X} \|x - \hat{c}\|^2 = n \cdot \|\hat{c} - c\|^2 + \sum_{x \in X} \|x - c\|^2.$$

The next lemma bounds the distance from the average of X to the average of a subset of X.

Lemma 2.2 (Ostrovsky et al. (2012))**.** *Let X be a set of points in $\mathbb{R}^d$ and let $S \subseteq X$ with $S \neq \emptyset$. Let c and s denote the averages of X and S, resp. Then,*

$$\|s - c\|^2 \leq \frac{\mathrm{OPT}_1(X)}{|X|} \cdot \frac{|X \setminus S|}{|S|}.$$

2.1 Clustering Stable Instances

In recent years, many works have studied the notion of clustering under various input-stability assumptions, showing how to rely on input-niceness in order to obtain a good approximation and even a PTAS for clustering problems.[6] The main focus of this work is the stability notion of Ostrovsky et al. (2012) who

[4] With an additive error, as analyzed in Theorem 3.3.

[5] Specifically, the bound of Huang and Liu (2018) is guaranteed to hold whenever $\mathrm{OPT}_{k-1}(X) \gtrsim n^{\frac{11}{20}} k^{\frac{7}{4}} d^{\frac{3}{4}} \varepsilon^{-\frac{1}{2}} \phi^{-4}$, whereas our bound holds also for smaller values of $\mathrm{OPT}_{k-1}(X)$.

[6] Ostrovsky et al. (2012); Ackerman and Ben-David (2009); Balcan et al. (2009); Bilu and Linial (2010); Awasthi et al. (2010); Kumar and Kannan (2010); Awasthi et al. (2012); Cohen-Addad and Schwiegelshohn (2017)

defined a clustering instance to be *well-separated* for k clusters if the optimal partitioning of the data into k clusters has cost noticeably smaller than the cost of any partitioning of the data into $k-1$ clusters.

Definition 2.3 (Ostrovsky et al. (2012)). *A clustering instance X is ϕ-*well-separated *(or simply ϕ-separated) for k-means if $\mathrm{OPT}_k(X) \leq \phi^2 \cdot \mathrm{OPT}_{k-1}(X)$.*

The following theorem relates the task of approximating the k-means cost and the task of approximating the true k-means centers in Wasserstein distance.

Theorem 2.4 (Ostrovsky et al. (2012)). *Let α and ϕ be such that $\frac{\alpha+\phi^2}{1-\phi^2} < \frac{1}{16}$. Suppose that $X \subseteq \mathbb{R}^d$ is ϕ-separated for k-means, and let $C = (c_1, \ldots, c_k)$ be a set of optimal centers for X. For $i \in [k]$ let D_i denote the distance from c_i to its nearest optimal center, that is $D_i = \min_{j \neq i} \|c_j - c_i\|$. Let $\hat{C} = (\hat{c}_1 \ldots, \hat{c}_k)$ be centers such that $\mathrm{cost}_X(\hat{C}) \leq \alpha \cdot \mathrm{OPT}_{k-1}(X)$. Then for each $\hat{c}_i$ there is a distinct optimal center, call it c_i, such that $\|\hat{c}_i - c_i\| \leq 2\sqrt{\frac{\alpha+\phi^2}{1-\phi^2}} \cdot D_i$.*

While Ostrovsky et al. were the first to define formally a certain input-stability notion, other works quickly followed. Next, we show equivalence between several such input-stability notions.

Lemma 2.5. *The following notions of stability are all equivalent up to a constant factor.*

1. **ϕ-well separability (Ostrovsky et al., 2012):** $\mathrm{OPT}_k \leq \phi^2 \cdot \mathrm{OPT}_{k-1}$.
2. **β-center deletion (Awasthi et al., 2010):** *For every cluster i and $j \neq i$, delete center c_i and assign all of its points to center c_j. The result is a $(k-1)$-clustering of cost $\geq \beta\,\mathrm{OPT}_k$.*
3. **γ-center separation (Awasthi et al., 2010):** *For every cluster i, denote its size by $|X_i|$ and let $D_i^p = \min_{j \neq i} \|c_i - c_j\|^p$. Then $D_i^p \geq \frac{\gamma}{|X_i|}\,\mathrm{OPT}_k$.*
4. **$(\delta, \frac{1}{4})$-approximation stability (Ostrovsky et al., 2012):** *For any k-tuple $\hat{c}_1, .., \hat{c}_k$ of cost at most $\delta\,\mathrm{OPT}_k$, we have a matching φ such that $\|c_i - \hat{c}_{\varphi(i)}\|^p < \frac{1}{4} \cdot D_i^p$.*[7]

As an immediate corollary, it follows that our algorithms are applicable to any instance satisfying one of the above mentioned stability notions (with suitable stability parameters). The proof of Lemma 2.5 appears in the full version of this paper.

[7] We comment that we can replace the constant $\frac{1}{4}$ with any constant $< \frac{1}{2}$.

3 Stability Improves Accuracy for Private Clustering Algorithms

In this section we show that applying a private clustering algorithm with a worst-case guarantee of v-approximation on a ϕ-well separable instance results in (effectively) a $(1 + O(\phi^2))$-approximation for the k-means, provided that v is small in comparison to ϕ^{-2}. In other words, we show that when running a private clustering algorithm $\mathcal{A}$ on *stable instances*, then $\mathcal{A}$ actually performs much better than its worst case bounds. We focus here on the k-means cost objective, and present an analogues result for k-median in Section 4. Our construction appears in algorithm `Private-Stable-`k`-Means`.

The privacy properties of algorithm `Private-Stable-`k`-Means` are immediate from composition properties of differential privacy (see Dwork et al. (2010)). We proceed with its utility analysis. Let X be ϕ-separated for k-means with optimal centers $C^* = \{c_1^*, \ldots, c_k^*\}$, and let $X_1^*, \ldots, X_k^* \subseteq X$ be the clusters induced by C^*. For $i \in [k]$ we denote $n_i = |X_i^*|$ and $r_i^* = \sqrt{\frac{1}{n_i} \sum_{x \in X_i^*} \|x - c_i^*\|^2}$. Consider the execution of `Private-Stable-`k`-Means` on X, and let $B = \{b_1, \ldots, b_k\}$ and $\overline{C} = \{\bar{c}_1, \ldots, \bar{c}_k\}$ denote the centers obtained in Steps 1 and 4. We assume for simplicity (and without loss of generality) that the set of optimal centers $C^* = \{c_1^*, \ldots, c_k^*\}$ is sorted s.t. c_i^* is the closest to b_i. We note that such a matching exists provided that the requirements of Theorem 2.4 are met. The next lemma shows that, provided that B has a low enough cost, then the distance from each $\bar{c}_i$ to its corresponding optimal center is low.

Lemma 3.1. *Let $\rho = \frac{100\phi^2}{1-\phi^2}$. If $\mathrm{cost}_X(B) \leq w \cdot \phi^2 \cdot \mathrm{OPT}_{k-1}(X)$ and if $\frac{\phi^2(w+1)}{1-\phi^2}$ is sufficiently small, then $\|\bar{c}_i - c_i^*\|^2 \leq r_i^2 \cdot \frac{\rho}{1-\rho}$*

Proof. For $i \in [k]$ define $X_i^{\mathrm{cor}} = \{x \in X_i^* : \|x - c_i^*\| \leq \frac{r_i}{\sqrt{\rho}}\}$. Standard Markovian argument shows that $|X_i^{\mathrm{cor}}| \geq (1-\rho) \cdot n_i$. We first show that for every $i \in [k]$ we have $X_i^{\mathrm{cor}} \subseteq \hat{X}_i \subseteq X_i^*$. To that end, fix $i \in [k]$ and recall that $\mathrm{cost}_X(B) \leq w \cdot \phi^2 \cdot \mathrm{OPT}_{k-1}(X)$. Denote $\gamma = 2\sqrt{\frac{\phi^2(w+1)}{1-\phi^2}}$, and $D_i = \min_{j \neq i} \|c_i^* - c_j^*\|$. By Theorem 2.4 we have that $\|c_i^* - b_i\| \leq \gamma \cdot D_i$.

Algorithm `Private-Stable-`k`-Means`

Input: Database X containing n points in the d-dimensional ball $\mathcal{B}(0,\Lambda)$, failure probability β, privacy parameters ε, δ.

Tool used: An (ε,δ)-differentially private algorithm $\mathcal{A}$ for approximating the k-means.

1. Run $\mathcal{A}$ on X to obtain k centers: $B = \{b_1,\ldots,b_k\}$.
2. For $i \in [k]$ let $\hat{D}_i = \min_{j\neq i} \|b_i - b_j\|$.
3. For $i \in [k]$ let $\hat{X}_i = \{x \in X : \|x - b_i\| \leq \hat{D}_i/3\}$.
4. Let $\overline{C} = \{\overline{c}_1,\ldots,\overline{c}_k\}$ denote the average of the points in $\hat{X}_1,\ldots,\hat{X}_k$, respectively. For $i \in [k]$ use the Gaussian mechanism (see (Dwork et al., 2006a)) with privacy parameters (ε,δ) to compute a noisy estimation $\hat{c}_i$ of $\overline{c}_i$. Denote $\hat{C} = \{\hat{c}_1,\ldots,\hat{c}_k\}$.
5. Use the Gaussian mechanism with privacy parameters (ε,δ) to estimate $\mathrm{cost}_X(\hat{C})$ and $\mathrm{cost}_X(B)$. Output the set of centers (either $\hat{C}$ or B) with the lower (estimated) cost.

Now, $\hat{X}_i$ contains every point $x \in X$ whose within a distance from b_i of

$$\begin{aligned}
\frac{1}{3}\hat{D}_i &= \frac{1}{3}\min_{j\neq i}\|b_i - b_j\| \\
&\geq \frac{1}{3}\min_{j\neq i}\left(\|c_i^* - c_j^*\| - \|b_i - c_i^*\| - \|b_j - c_j^*\|\right) \\
&\geq \frac{1}{3}\min_{j\neq i}\left(\|c_i^* - c_j^*\| - \gamma\cdot D_i - \gamma\cdot\min_{\ell\neq j}\|c_\ell^* - c_j^*\|\right) \\
&\geq \frac{1}{3}\min_{j\neq i}\left(\|c_i^* - c_j^*\| - \gamma\cdot D_i - \gamma\cdot\|c_i^* - c_j^*\|\right) \\
&= \frac{1}{3}\min_{j\neq i}\left(\|c_i^* - c_j^*\| - 2\gamma\cdot D_i\right) \\
&= \frac{1}{3}\left[D_i - 2\gamma\cdot D_i\right] = \frac{1-2\gamma}{3}\cdot D_i \qquad (1)
\end{aligned}$$

In particular, $\hat{X}_i$ contains every point $x \in X$ whose within a distance from c_i^* of

$$\frac{1-2\gamma}{3}\cdot D_i - \gamma\cdot D_i \geq \frac{1-5\gamma}{3}\cdot\sqrt{\frac{1-\phi^2}{\phi^2}}\cdot r_i^* \geq \frac{r_i^*}{\sqrt{\rho}}$$

where the first inequality is from Theorem 2.4 and the second inequality holds for sufficiently small ϕ and w. Therefore, $X_i^{\mathrm{cor}} \subseteq \hat{X}_i$, because X_i^{cor} contains points within distance $r_i^*/\sqrt{\rho}$ from c_i^*. Similar arguments (appears in the full version of this paper) show that $\hat{X}_i \subseteq X_i^*$. So, $X_i^{\mathrm{cor}} \subseteq \hat{X}_i \subseteq X_i^*$. Recall that $\overline{c}_i$ denotes the average of the points in $\hat{X}_i$. By Lemma 2.2 we have that

$$\begin{aligned}
\|\overline{c}_i - c_i^*\|^2 &\leq \frac{\mathrm{OPT}_1(X_i^*)}{|X_i^*|}\cdot\frac{|X_i^*\setminus\hat{X}_i|}{|\hat{X}_i|} \\
&\leq \frac{\mathrm{OPT}_1(X_i^*)}{n_i}\cdot\frac{|X_i^*\setminus X_i^{\mathrm{cor}}|}{|X_i^{\mathrm{cor}}|} \\
&\leq \frac{\mathrm{OPT}_1(X_i^*)}{n_i}\cdot\frac{\rho}{1-\rho}
\end{aligned}$$

□

Let $\hat{C}$ be the centers obtained in Step 4 of the execution, and recall that each $\hat{c}_i \in \hat{C}$ is a noisy estimation of $\overline{c}_i$, where $\overline{c}_i$ is the average of the points in $\hat{X}_i$ (all the input points whose distance to b_i is significantly smaller than their distance to any other b_j). The next lemma shows that the k-means cost of $\hat{C}$ is low. This is done by relating the cost of $\hat{C}$ to that of $\overline{C}$, which we then relate to the cost of the optimal centers using Lemma 3.1.

Lemma 3.2. *If* $\mathrm{cost}_X(B) \leq w\cdot\phi^2\cdot\mathrm{OPT}_{k-1}(X)$ *and if* $\frac{\phi^2(w+1)}{1-\phi^2}$ *is sufficiently small, then* $\mathrm{cost}_X(\hat{C}) \leq \left(1+O(\phi^2)\right)\cdot\mathrm{OPT}_k(X) + O(1)\cdot\frac{k\Lambda^2\sqrt{d}}{\varepsilon}\cdot\ln\left(\frac{dk}{\beta\delta}\right)$.

Proof. First let us assume that for every $i \in [k]$ it holds that $|\hat{X}_i| \geq \frac{16}{\varepsilon}\log(\frac{4k}{\beta\delta})$, which is true if $n_i = |X_i^*| \geq \frac{16}{\varepsilon(1-\rho)}\log(\frac{4k}{\beta\delta})$. Fix $i \in [k]$. By the properties of the Gaussian mechanism (see (Dwork et al., 2006a)), with probability at least $(1-\frac{\beta}{k})$ we have that $\|\hat{c}_i - \overline{c}_i\| \leq \frac{64\Lambda\sqrt{d}}{\varepsilon\cdot|\hat{X}_i|}\cdot\ln\left(\frac{8dk}{\beta\delta}\right) \leq \frac{64\Lambda\sqrt{d}}{\varepsilon\cdot(1-\rho)n_i}\cdot\ln\left(\frac{8dk}{\beta\delta}\right)$. Thus, by Lemma 2.1 we have that

$$\begin{aligned}
\mathrm{cost}_X(\hat{C}) &\leq \sum_i\sum_{x\in X_i^*}\|x - \hat{c}_i\|^2 \\
&= \sum_i\left(\mathrm{OPT}_1(X_i^*) + n_i\cdot\|\hat{c}_i - c_i^*\|^2\right) \\
&\leq \sum_i\left(\mathrm{OPT}_1(X_i^*) + 3n_i\cdot\|\overline{c}_i - c_i^*\|^2 + 3n_i\cdot\|\hat{c}_i - \overline{c}_i\|^2\right) \\
&\leq \sum_i\left(\mathrm{OPT}_1(X_i^*) + 3\,\mathrm{OPT}_1(X_i^*)\cdot\frac{\rho}{1-\rho}\right. \\
&\quad \left. + O(1)\cdot\min\left\{n_i\cdot\Lambda^2,\ \frac{\Lambda^2 d}{\varepsilon^2(1-\rho)^2\cdot n_i}\cdot\ln^2\left(\frac{dk}{\beta\delta}\right)\right\}\right) \\
&\overset{(*)}{\leq} \sum_i\left(\mathrm{OPT}_1(X_i^*) + 3\,\mathrm{OPT}_1(X_i^*)\cdot\frac{\rho}{1-\rho}\right.
\end{aligned}$$

$$+ O(\Lambda^2) \cdot \frac{\sqrt{d}}{\varepsilon(1-\rho)} \cdot \ln\left(\frac{dk}{\beta\delta}\right)\Bigg)$$
$$= \left(1 + \frac{3\rho}{1-\rho}\right) \mathrm{OPT}_k(X) + O(1) \cdot \frac{k\Lambda^2\sqrt{d}}{\varepsilon(1-\rho)} \ln\left(\frac{dk}{\beta\delta}\right)$$
$$= \left(1 + O(\phi^2)\right) \mathrm{OPT}_k(X) + O(1) \cdot \frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right),$$

where the inequality $(*)$ follows from the fact that $\min\left\{n_i,\ \frac{d}{\varepsilon^2(1-\rho)^2 \cdot n_i} \cdot \ln^2\left(\frac{dk}{\beta\delta}\right)\right\} \leq 2\frac{\sqrt{d}\ln\left(\frac{dk}{\beta\delta}\right)}{\varepsilon(1-\rho)}$. Now, small clusters of size $n_i < \frac{16}{\varepsilon(1-\rho)}\log(\frac{4k}{\beta\delta})$ can increase the cost of $\mathrm{cost}_X(\hat{C})$ by at most $\Lambda^2 \cdot \frac{16}{\varepsilon(1-\rho)}\log(\frac{4k}{\beta\delta})$ additively, and hence, overall we have that $\mathrm{cost}_X(\hat{C})$ is upper bounded by

$$\left(1 + O(\phi^2)\right) \cdot \mathrm{OPT}_k(X) + O\left(\frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right)\right) \quad \square$$

Lemma 3.2 shows that whenever the set of centers B (computed in Step 1) is "good enough" then the resulting set of centers $\hat{C}$ has a low k-means cost (obtaining better guarantees than B). However, the set of centers B is computed using a *private* approximation algorithm, which has both multiplicative and additive errors. In the next theorem we argue that, taking B's additive error into account, either B itself is already a good approximation for the k-means, or its additive error is small enough so that it has only a small effect on the error of $\hat{C}$.

Theorem 3.3. *Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0,\Lambda)$, and assume that X is ϕ-separated for k-means. Let algorithm* `Private-Stable-k-Means` *be executed on X with a subroutine $\mathcal{A}$ that returns, with probability at least $(1-\beta_1)$, a set of centers B satisfying $\mathrm{cost}_X(B) \leq v \cdot \mathrm{OPT}_k(X) + t$. If $\phi^2 \leq O\left(\frac{1}{v}\right)$, then with probability at least $(1-\beta-\beta_1)$, algorithm* `Private-Stable-k-Means` *returns a set of centers C' of cost $\mathrm{cost}_X(C')$ of at most*

$$(1+O(\phi^2)) \cdot \mathrm{OPT}_k(X) + O\left(vt + \frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right)\right).$$

If furthermore $\mathrm{OPT}_{k-1}(X) \geq \frac{t}{\phi^2}$, then $\mathrm{cost}_X(C') \leq (1 + O(\phi^2)) \cdot \mathrm{OPT}_k(X) + O\left(\frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right)\right)$.

Proof sketch. Recall that in Step 5, algorithm `Private-Stable-k-Means` chooses between B and $\hat{C}$ using the Gaussian mechanism. We analyze two cases and show that at least one of these options has small error (the additional error introduces by the Gaussian mechanism is also small). If $\mathrm{OPT}_{k-1}(X) \leq \frac{t}{\phi^2}$ then

$$\begin{aligned}\mathrm{cost}_X(B) &\leq v \cdot \mathrm{OPT}_k(X) + t \\ &\leq v \cdot \phi^2 \cdot \mathrm{OPT}_{k-1}(X) + t \\ &\leq v \cdot t + t \leq O(vt),\end{aligned}$$

and hence, B is a good output. On the other hand, if $\mathrm{OPT}_{k-1}(X) > \frac{t}{\phi^2}$ then

$$\begin{aligned}\mathrm{cost}_X(B) &\leq v \cdot \mathrm{OPT}_k(X) + t \\ &\leq v \cdot \phi^2 \cdot \mathrm{OPT}_{k-1}(X) + \phi^2 \cdot \mathrm{OPT}_{k-1}(X) \\ &= (v+1)\,\phi^2 \cdot \mathrm{OPT}_k(X).\end{aligned}$$

Therefore, for $\phi^2 \leq O(\frac{1}{v})$, we have that the conditions of Lemma 3.2 are met, and so $\mathrm{cost}_X(\hat{C}) \leq$

$$\left(1 + O(\phi^2)\right) \cdot \mathrm{OPT}_k(X) + O(1) \cdot \frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right)$$

and $\hat{C}$ is a good output. $\square$

Combining Theorem 3.3 with the private algorithm of Kaplan and Stemmer (2018) achieving $O(1)$-approximation for the k-means, we get the following corollary.[8]

Corollary 3.4. *There exists an (ε,δ)-differentially private algorithm such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0,\Lambda)$, and assume that X is ϕ-separated for k-means for $\phi = O(1)$ (sufficiently small). When applied to X, the algorithm returns, with probability at least $(1-\beta)$, a set of k centers C' where $\mathrm{cost}_X(C')$ is at most*

$$(1 + O(\phi^2))\,\mathrm{OPT}_k(X) + \tilde{O}\left(\frac{k^{1.01}d^{0.51}\Lambda^2}{\varepsilon^{1.01}} + \frac{k^{1.5}\Lambda^2}{\varepsilon}\right)$$

Furthermore, if $\mathrm{OPT}_{k-1}(X) \geq \tilde{O}\left(\frac{k^{1.01}\cdot d^{0.51}\cdot\Lambda^2}{\varepsilon^{1.01}\phi^2} + \frac{k^{1.5}\cdot\Lambda^2}{\varepsilon\phi^2}\right)$, then $\mathrm{cost}_X(C')$ is upper bounded by

$$(1 + O(\phi^2)) \cdot \mathrm{OPT}_k(X) + O\left(\frac{k\Lambda^2\sqrt{d}}{\varepsilon} \cdot \ln\left(\frac{dk}{\beta\delta}\right)\right)$$

Our algorithm from Corollary 3.4 also results in a new construction for privately approximating the k-means in terms of the Wasserstein distance to the optimal centers. This follows from the fact that, for well-separated instances, centers with near optimal k-means cost must be close to the optimal centers in terms of the Wasserstein distance. Specifically, we get the following result.

[8] For simplicity, throughout the paper we use the $\tilde{O}$ notation to hide logarithmic factors in $k, n, d, 1/\beta, 1/\delta$.

Theorem 3.5. *There exists an (ε, δ)-differentially private algorithm such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0, \Lambda)$. Assume that X is ϕ-separated for k-means for $\phi = O(1)$ (sufficiently small), and assume that $\mathrm{OPT}_{k-1}(X)$ is at least*

$$\tilde{O}\left(\left(\Lambda^2 + \frac{1}{\phi^4}\right) \cdot \left(\frac{k^{1.01} \cdot d^{0.51}}{\varepsilon^{1.01}} + \frac{k^{1.5}}{\varepsilon}\right)\right).$$

When applied to X, the algorithm returns, with probability at least $(1-\beta)$, a set of k centers C' satisfying $d_{\mathsf{W}}(C^*, C') \leq O(\phi^2 \cdot \Lambda)$, where C^* are the optimal centers.

The error bound in this theorem matches the state-of-the-art result of Huang and Liu (2018), and offers some improvements in terms of the requirement on $\mathrm{OPT}_{k-1}(X)$. Specifically, the bound of Huang and Liu (2018) is guaranteed to hold whenever $\mathrm{OPT}_{k-1}(X) \gtrsim n^{\frac{11}{20}} k^{\frac{7}{4}} d^{\frac{3}{4}} \varepsilon^{-\frac{1}{2}} \phi^{-4}$, whereas our bound holds also for smaller values of $\mathrm{OPT}_{k-1}(X)$.

4 Private k-Median Clustering with Stability Assumptions

Our construction for the k-median is conceptually similar to our construction for the k-means. The main difference is that for k-median we cannot use the average to approximate the center of a cluster, as the average can be far from the optimal median of the cluster. We instead use a tool of Bassily et al. (2014) for privately solving convex optimization problems. We obtain the following theorem (the details are given in the full version of this paper).

Theorem 4.1. *There exists an (ε, δ)-differentially private algorithm such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0, \Lambda)$, and assume that X is ϕ-separated for k-median for $\phi = O(1)$ (sufficiently small). When applied to X, the algorithm returns, with probability at least $(1-\beta)$, a set of k centers C' whose k-median cost is upper bounded by*

$$(1 + O(\phi)) \cdot \mathrm{OPT}_k^1(X) + \tilde{O}\left(\frac{k^{1.01} d^{0.51} \Lambda}{\varepsilon^{1.01}} + \frac{k^{1.5} \Lambda}{\varepsilon}\right).$$

Here $\mathrm{OPT}_k^1(X)$ is the lowest possible k-median cost. We remark that for k-median we use the convention that X is ϕ-separable if $\mathrm{OPT}_k^1(X) \leq \phi \cdot \mathrm{OPT}_{k-1}^1(X)$ (that is, the ratio between the optimal costs is at most ϕ and not ϕ^2 like for k-means). The approximation ratio $(1 + O(\phi))$ in Theorem 4.1 is merely a result of this difference in notation.

5 Clustering with Stability Assumptions in the Local Model

In the local model of differential privacy (LDP), there are n users and an untrusted server. Each user i is holding a private input item x_i (a point in $\mathbb{R}^d$ in our case), and the server's goal is to compute some function of the inputs (approximate the k-means in our case). However, in this model, the users do not send their data as is to the server. Instead, every user randomizes her data locally, and sends a differentially private report to the server, who aggregates all the reports. Informally, the privacy requirement is that the input of user i has almost no effect on the distribution on the messages that user i broadcasts.

Our locally-private protocols for k-means and for k-median are obtained from our constructions for the centralized model by instantiating existing LDP tools for computing averages (in the case of k-means) and for solving convex optimization problems (in the case of k-median). Here we present the result for k-means. The full construction appears in algorithm `LDP-Stable-`k`-Means`. A similar analysis to that of Section 3 shows the following theorem.

Theorem 5.1. *There exists an (ε, δ)-LDP protocol such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0, \Lambda)$, and assume that X is ϕ-separated for k-means for $\phi = O(1)$ (sufficiently small). When applied to X, the protocol returns, with probability at least $(1 - \beta)$, a set of k centers C' satisfying*

$$\mathrm{cost}_X(C') \leq (1+O(\phi^2)) \cdot \mathrm{OPT}_k(X) + \tilde{O}\left(\frac{k\sqrt{d} n^{0.51} \Lambda^2}{\varepsilon}\right).$$

As before, using the fact that (for well-separated instances) centers with near optimal k-means cost are close to the optimal centers in terms of the Wasserstein distance, we get the following theorem.

Theorem 5.2. *There exists an (ε, δ)-LDP protocol such that the following holds. Let X be a database containing n points in the d-dimensional ball $\mathcal{B}(0, \Lambda)$. Assume that X is ϕ-separated for k-means for sufficiently small ϕ, and assume that*

$$\mathrm{OPT}_{k-1}(X) \geq \tilde{O}\left(\frac{k\sqrt{d} \cdot n^{0.51} \cdot \Lambda^2}{\varepsilon \cdot \phi^4}\right).$$

Then on X, the protocol returns with probability $\geq (1-\beta)$ a set of k centers C' satisfying $d_{\mathsf{W}}(C^*, C') \leq O(\phi^2 \cdot \Lambda)$, with C^* denoting the optimal centers.

Algorithm `LDP-Stable-`k`-Means`

Input: Failure probability β, privacy parameters ε, δ.

Setting: Each player $i \in [n]$ holds a point x_i in the d-dimensional ball $\mathcal{B}(0, \Lambda)$. Define $X = (x_1, \ldots, x_n)$.

Tool used: An (ε, δ)-LDP protocol $\mathcal{A}$ for approximating the k-means.

1. Run $\mathcal{A}$ on X to obtain k centers: $B = \{b_1, \ldots, b_k\}$.
2. For $i \in [k]$ let $\hat{D}_i = \min_{j \neq i} \|b_i - b_j\|$.
3. For $i \in [k]$ let $R_i = \{x \in \mathcal{B}(0, \Lambda) : \|x - b_i\| \leq \hat{D}_i/3\}$, and denote $\hat{X}_i = X \cap R_i$.
4. Let $\overline{C} = \{\overline{c}_1, \ldots, \overline{c}_k\}$ denote the average of the points in $\hat{X}_1, \ldots, \hat{X}_k$, respectively. Use an (ε, δ)-LDP averaging tool (see, e.g., (Nissim and Stemmer, 2018)) to obtain for every $i \in [k]$ a noisy estimation $\hat{c}_i$ of the average of $X \cap R_i$, i.e., an estimation of $\overline{c}_i$.
5. Estimate $\mathrm{cost}_X(\hat{C})$ and $\mathrm{cost}_X(B)$ using an (ε, δ)-LDP counting tool (see, e.g., (Kasiviswanathan et al., 2011)). Output the set of centers (either $\hat{C}$ or B) with the lower estimated cost.

6 Discussion and Open Problems

This work establishes a new baseline for privately clustering stable instances. More importantly, our work emphasizes the importance of "simplicity" in the design of DP clustering algorithms. In particular, due to its simplicity, our algorithm has straightforward analogues for clustering "nice" k-median instances and for the local-model of DP.

Naturally, several important open problems arise from our work. First, we pose the problem of finding a PTAS for k-means under stability assumptions. Non-privately, there are several papers proposing such clustering algorithms (Awasthi et al., 2010; Cohen-Addad and Schwiegelshohn, 2017) and other works that approximate the target clustering pointwise (Balcan et al., 2009); whereas privately we are only able to derive a $(1 + O(\phi^2))$-approximation for the k-means cost of ϕ-well separated instances. In other words, in the non-private settings the quality of the approximation is independent of the input's stability guarantee, whereas in the private setting a high-quality approximation requires a very strong separation guarantee on the input. What prevents us from deriving private analogues of the abovementioned PTASs which get a $(1+\alpha)$-approximation for any arbitrarily small α? The reason lies in designing a private analogue to one of the most classical approaches for k-means approximation — sampling (Inaba et al., 1994). It is a well-known fact that the centroid obtained by randomly sampling $O(1/\alpha)$ datapoints from a cluster yields a $(1 + \alpha)$-approximation to the cluster's cost, and the abovementioned PTASs rely on this fact. On a high-level, a PTAS for stable inputs works by partitioning the clusters into two types: "cheap" clusters that cost at most $O(\alpha\phi^2 \cdot \mathrm{OPT})$ vs "expensive" (non-cheap) clusters. Approximating the center of a cheap cluster relies on the notion of a core and can be made private using the 1-cluster algorithm, but the difficulty lies in approximating the centers of the expensive clusters. In the non-private setting expensive clusters are simple to handle — since there are at most $O(1/\alpha\phi^2)$ such clusters, one just brute-force tries all possible centers for all expensive clusters. Alas, we have *no private analogue for this approach.* More specifically, should we wish to handle expensive clusters similarly, then we first need to devise a differentially-private analogue of the PTAS of Inaba et al (1994) which runs in $n^{O(k)}$-time. Alternatively, one could potentially derive additional properties of expensive clusters which would allow us to approximate their centers privately; or potentially try a different approach, one that doesn't rely on the separation into cheap vs. heavy clusters.

We also re-pose the question of a definition of clustering which is generalizable. Despite the fact that the k-means and k-median problems are part of the "CS-canon", it is possible these two problems are the "wrong" problems to approximate with a differentially private algorithm. The reason lies in the sensitivity of the optimal k centers, even for stable instances. However, if instead of outputting the "true" k-means centers we shift our focus to outputting some notion of "core centers" or centers that best represent the fraction of the instance with clear preference among centers[9], then such objectives might be less sensitive to a change to a single datapoint and could therefore be better suited for DP.

[9] Note how this proposed "definition" is recursive and thus ill-defined.

Acknowledgements

We thank Zhiyi Huang and Jinyan Liu for helpful discussions. M.S. and U.S. were supported in part by the Israel Science Foundation (grant No. 1871/19). M.S. was also supported by the Frankel Center for Computer Science. O.S. was supported by grant #201706701 of the Natural Sciences and Engineering Research Council of Canada (NSERC). The bulk of this work was done when O.S. was affiliated with the University of Alberta, Canada.

References

Ackerman, M. and Ben-David, S. (2009). Clusterability: A theoretical study. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, AISTATS 2009, Clearwater Beach, Florida, USA, April 16-18, 2009*, pages 1–8.

Ahmadian, S., Norouzi-Fard, A., Svensson, O., and Ward, J. (2017). Better guarantees for k-means and euclidean k-median by primal-dual algorithms. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 61–72.

Awasthi, P. (2013). *Approximation Algorithms and New Models for Clustering and Learning.* PhD thesis, Carnegie Mellon University. Supervisor-Avrim Blum.

Awasthi, P., Blum, A., and Sheffet, O. (2010). Stability yields a PTAS for k-median and k-means clustering. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 309–318.

Awasthi, P., Blum, A., and Sheffet, O. (2012). Center-based clustering under perturbation stability. *Inf. Process. Lett.*, 112(1-2):49–54.

Balcan, M., Blum, A., and Gupta, A. (2009). Approximate clustering without the approximation. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2009, New York, NY, USA, January 4-6, 2009*, pages 1068–1077.

Balcan, M.-F., Dick, T., Liang, Y., Mou, W., and Zhang, H. (2017). Differentially private clustering in high-dimensional Euclidean spaces. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 322–331, International Convention Centre, Sydney, Australia. PMLR.

Bassily, R., Smith, A., and Thakurta, A. (2014). Private empirical risk minimization: Efficient algorithms and tight error bounds. In *FOCS*, pages 464–473. IEEE.

Bilu, Y. and Linial, N. (2010). Are stable instances easy? In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 5-7, 2010. Proceedings*, pages 332–341.

Blum, A., Dwork, C., McSherry, F., and Nissim, K. (2005). Practical privacy: The SuLQ framework. In Li, C., editor, *PODS*, pages 128–138. ACM.

Cohen-Addad, V. and Schwiegelshohn, C. (2017). On the local structure of stable clustering instances. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 49–60.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. (2006a). Our data, ourselves: Privacy via distributed noise generation. In Vaudenay, S., editor, *EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 486–503. Springer.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b). Calibrating noise to sensitivity in private data analysis. In *TCC*, pages 265–284. Springer.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006c). Calibrating noise to sensitivity in private data analysis. In *TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer.

Dwork, C., Rothblum, G. N., and Vadhan, S. P. (2010). Boosting and differential privacy. In *FOCS*, pages 51–60. IEEE Computer Society.

Feldman, D., Fiat, A., Kaplan, H., and Nissim, K. (2009). Private coresets. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 361–370.

Feldman, D., Xiang, C., Zhu, R., and Rus, D. (2017). Coresets for differentially private k-means clustering and applications to privacy in mobile sensor networks. In *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks*, IPSN '17, pages 3–15, New York, NY, USA. ACM.

Gupta, A., Ligett, K., McSherry, F., Roth, A., and Talwar, K. (2010). Differentially private combinatorial optimization. In *Proceedings of the*

Twenty-first Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '10, pages 1106–1125, Philadelphia, PA, USA. Society for Industrial and Applied Mathematics.

Huang, Z. and Liu, J. (2018). Optimal differentially private algorithms for k-means clustering. In *Proceedings of the 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, Houston, TX, USA, June 10-15, 2018*, pages 395–408.

Inaba, M., Katoh, N., and Imai, H. (1994). Applications of weighted voronoi diagrams and randomization to variance-based k-clustering: (extended abstract). In *Proc. 10th Symp. Comp. Geom.*, pages 332–339.

Kaplan, H. and Stemmer, U. (2018). Differentially private k-means with constant multiplicative error. In *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada.*, pages 5436–5446.

Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S., and Smith, A. (2011). What can we learn privately? *SIAM J. Comput.*, 40(3):793–826.

Kumar, A. and Kannan, R. (2010). Clustering with spectral norm and the k-means algorithm. In *FOCS*.

McSherry, F. (2009). Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2009, Providence, Rhode Island, USA, June 29 - July 2, 2009*, pages 19–30.

Mohan, P., Thakurta, A., Shi, E., Song, D., and Culler, D. (2012). Gupt: Privacy preserving data analysis made easy. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, pages 349–360, New York, NY, USA. ACM.

Nissim, K., Raskhodnikova, S., and Smith, A. (2007). Smooth sensitivity and sampling in private data analysis. In *STOC*, pages 75–84. ACM.

Nissim, K. and Stemmer, U. (2018). Clustering algorithms for the centralized and local models. In Janoos, F., Mohri, M., and Sridharan, K., editors, *Proceedings of Algorithmic Learning Theory*, volume 83 of *Proceedings of Machine Learning Research*, pages 619–653. PMLR.

Nissim, K., Stemmer, U., and Vadhan, S. P. (2016). Locating a small cluster privately. In *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2016, San Francisco, CA, USA, June 26 - July 01, 2016*, pages 413–427.

Nock, R., Canyasse, R., Boreli, R., and Nielsen, F. (2016). k-variates++: more pluses in the k-means++. In *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, pages 145–154.

Ostrovsky, R., Rabani, Y., Schulman, L. J., and Swamy, C. (2012). The effectiveness of lloyd-type methods for the k-means problem. *J. ACM*, 59(6):28:1–28:22.

Stemmer, U. (2020). Locally private k-means clustering. In *SODA*. SIAM.

Su, D., Cao, J., Li, N., Bertino, E., and Jin, H. (2016). Differentially private k-means clustering. In *Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy*, CODASPY '16, pages 26–37, New York, NY, USA. ACM.

Wang, Y., Wang, Y.-X., and Singh, A. (2015). Differentially private subspace clustering. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'15, pages 1000–1008, Cambridge, MA, USA. MIT Press.

Wasserstein, L. N. (1969). Markov processes over denumerable products of spaces describing large systems of automata. *Problems of Information Transmission*, 5(3):47–52.