Tradeoffs in Privacy, Welfare, and Fairness for Facility Location

Sara Fish
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Yannai A. Gonczarowski
Department of Economics and Department of Computer Science, Harvard University, Cambridge, MA, USA

Jason Z. Tang
Department of Computer Science, Harvard University, Cambridge, MA, USA

Salil Vadhan
School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Abstract

The differentially private (DP) facility location problem seeks to determine a socially optimal placement for a public facility while ensuring that each participating agent's location remains private. In order to privatize its input data, a DP mechanism must inject noise into its output distribution, producing a placement that will have lower expected social welfare than the optimal spot for the facility. The privacy-induced welfare loss can be viewed as the "cost of privacy," illustrating a tradeoff between social welfare and privacy that has been the focus of prior work. Yet, the imposition of privacy also induces a third consideration that has not been similarly studied: *fairness* in how the "cost of privacy" is distributed across individuals. For instance, a mechanism may satisfy differential privacy with minimal social welfare loss, yet still be undesirable if that loss falls entirely on one individual. In this paper, we quantify this new notion of unfairness and design mechanisms for facility location that attempt to simultaneously optimize across these three objectives of privacy, social welfare, and fairness.

Under this setup, we first derive an impossibility result, showing that privacy and fairness cannot be simultaneously guaranteed over all possible datasets that could represent the locations of individuals in a population. We then consider a relaxation of the original problem that still requires worst-case differential privacy, but only seeks fairness and appealing social welfare over smaller, more "realistic-looking" families of datasets. For this relaxation, we construct a DP mechanism and demonstrate that it is simultaneously optimal (or, for a harder family of datasets, near-optimal up to small factors) on fairness and social welfare. This suggests that while there is a tradeoff between privacy and each of social welfare and fairness, there is no additional tradeoff when we consider all three objectives simultaneously, provided that the population data is sufficiently natural.

2012 ACM Subject Classification Theory of computation → Design and analysis of algorithms; Security and privacy → Privacy-preserving protocols; Theory of computation → Algorithmic game theory and mechanism design

Keywords and phrases differential privacy, facility location, fairness, mechanism design

Digital Object Identifier 10.4230/LIPIcs.FORC.2026.12

Related Version *Full Version*: `https://arxiv.org/abs/2604.10443`

Funding *Sara Fish*: Supported by an NSF Graduate Research Fellowship and a Kempner Institute Graduate Fellowship.
Yannai A. Gonczarowski: Supported by the National Science Foundation (NSF-BSF grant No. 2343922) and Harvard FAS Inequality in America Initiative.
Salil Vadhan: Supported in part by NSF grant BCS-2218803.

Acknowledgements This paper is based on Tang's undergraduate thesis [32], advised by Fish, Gonczarowski, and Vadhan.

1 Introduction

The *facility location* problem in game theory and economics examines the scenario where a public facility should be placed to maximize the total welfare of its prospective users. While this social welfare objective is the primary focus of traditional literature on facility location [21, 8, 17, 29], differential privacy [9, 12] has recently become an additional important desideratum for applications that rely on personal data. Since the welfare-maximizing location for the facility may depend on sensitive information about the population (e.g., which members of the population would use the facility and where they reside), the facility location mechanism should not inadvertently leak its data inputs.

A facility location mechanism can only satisfy differential privacy (DP) by reducing its dependence on individualized data, which leads to lower social welfare compared to the optimal (but non-private) mechanism. While previous work in DP facility location attempts to minimize the *total* loss in welfare [20, 22, 15], this objective fails to consider the unfairness induced by the *distribution* of this welfare loss across participating agents.

For example, consider the facility location problem depicted by Figure 1, where an agent's welfare is inversely related to their distance from the facility. Upon switching from the optimal facility placement to a potential placement chosen under differential privacy, the *total* distance between agents and the facility only increases moderately. However, Agent 4 bears almost all of that increase in distance while the remaining agents are minimally impacted. If, over the randomness of the DP mechanism, the changes in distance are frequently uneven across agents, then the mechanism might be undesirable as it introduces a new form of unfairness caused directly by the enforcement of privacy.

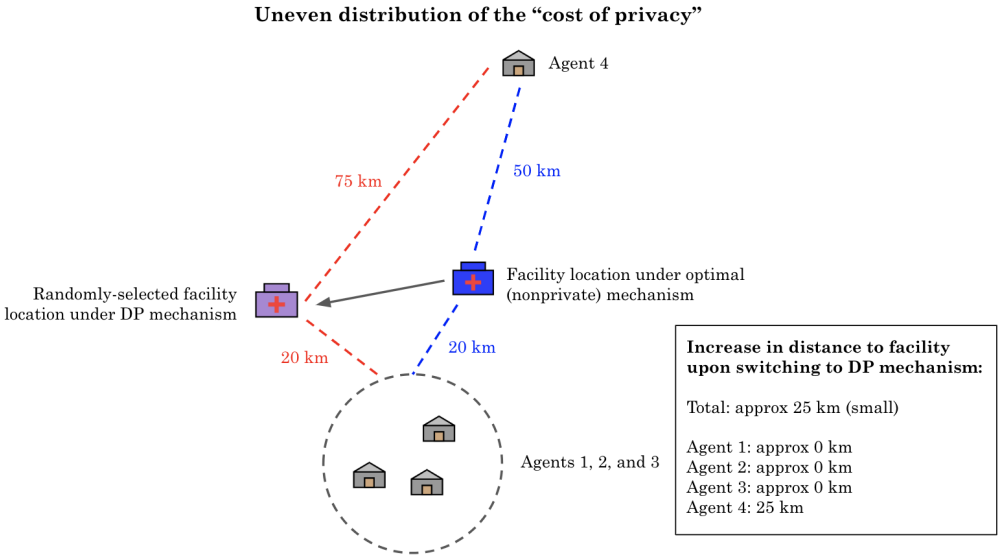


Figure 1 A depiction of how the switch from the optimal mechanism to the differentially private one can disparately impact agent utilities. In this case, the change in distance to the facility differs significantly between agents, suggesting a new form of unfairness.

In this paper, we consider three objectives in facility location mechanism design simultaneously, by characterizing the tradeoff between privacy and its cost in terms of both social welfare and fairness. We require privacy through DP and introduce two metrics (or loss functions) to quantify the remaining objectives of social welfare and fairness. Both metrics

are formulated by comparing the utilities of agents in two worlds: (1) the world where the location is selected via the optimal (but nonprivate) facility location mechanism $\mathcal{T}$, and (2) the world where the facility location is selected via a DP mechanism $\mathcal{M}$.

- The first loss function, SWDIFF, corresponds to social welfare. SWDIFF measures the difference in social welfare under the outputs of $\mathcal{T}$ and $\mathcal{M}$ and is standard in the literature.
- The second loss function, FAIR, is our novel proposal and measures the *maximum* loss in utility experienced by an individual agent when we move from $\mathcal{T}$ to $\mathcal{M}$. FAIR corresponds to fairness in the sense that minimizing it ensures no individual pays significantly for the imposition of privacy (in the spirit of Rawlsian social welfare in economics, which measures the minimum utility of any individual agent).

We analyze the facility location problem under these metrics. We focus specifically on the bounded and one-dimensional setting, under which the non-private optimum for SWDIFF is given by the median. We offer the following contributions.

1. **Quantifying fairness under the imposition of privacy:** We propose a new individual-utility-based notion of fairness, FAIR, as described above. This metric measures the extent to which privacy-induced welfare loss may be distributed unevenly among agents.
2. **Impossibility result on achieving both privacy and fairness:** We show that no meaningfully private (i.e., ϵ-DP for any reasonably small ϵ) facility location mechanism preserves fairness, as quantified by FAIR, for all possible arrangements (henceforth referred to as "datasets") of agents.
3. **Simultaneous privacy, social welfare, and fairness under a relaxation:** The previous impossibility result only arises upon considering datasets that are "pathological" and unrepresentative of many real-world ones. We thus consider a relaxation of the problem where we still require mechanisms to be (pure) DP over all datasets, but only require good fairness and welfare over a smaller family of natural datasets. We propose and motivate two such families – datasets that are "collapsing towards the median" and datasets that are roughly "single-peaked at the median" – on which we demonstrate that there exists a DP mechanism that achieves optimal (or, for the latter harder family, near-optimal) bounds on SWDIFF and FAIR *simultaneously*. To arrive at this result, we establish three sets of technical results for each family:
   a. **Analyzing a strong DP mechanism:** We first show that there exists a DP mechanism, $\texttt{DPExpMed}_{\alpha}$, that performs well on both SWDIFF and FAIR upon tuning its parameter value α. The mechanism is an instantiation of the "widened exponential mechanism" [26] for the median.
   b. **Deriving high probability upper bounds:** We identify an optimal parameter value $\alpha = \alpha^*$ and prove upper bounds on SWDIFF and FAIR for the mechanism $\texttt{DPExpMed}_{\alpha^*}$ over each family of datasets.
   c. **Deriving information-theoretic lower bounds:** We prove information-theoretic lower bounds for SWDIFF and FAIR that must be incurred by *any* DP mechanism over each family of datasets.

   The key results from parts (b) and (c) are detailed in Table 1. All pairs of upper and lower bounds either match or nearly match up to small factors (see Table 1 caption), suggesting our proposed mechanism from (a) is simultaneously optimal on both SWDIFF and FAIR. This then implies that social welfare and fairness can co-exist when designing optimal private mechanisms for facility location.

Table 1 Lower and upper bounds on loss derived in Section 6. n is the number of agents, m is the maximum possible distance between agents, β is the DP failure probability, and λ is a parameter that is typically $O(1/\sqrt{n})$. Our proposed ϵ-DP mechanism achieves all upper bounds simultaneously, and the lower bounds apply to all ϵ-DP mechanisms. The bounds match closely, with only minor discrepancies. Under typical assumptions $\lambda = O(1/\sqrt{n})$ and $\epsilon = \Theta(1)$, the extra $1/\epsilon$ in the $\mathcal{SPM}_\lambda$ upper bound on SWDIFF disappears, and the extra $\ln(n\lambda)$ factors are logarithmic relative to the leading n term in the denominator.

Dataset family		Lower bound	Upper bound achieved by $\texttt{DPExpMed}_{\alpha^*}$
Smaller family $\mathcal{CTM}$	FAIR	$\Omega\left(\frac{m}{n\epsilon}\ln\frac{1}{\beta}\right)$	$O\left(\frac{m}{n\epsilon}\ln\frac{1}{\beta}\right)$
	SWDIFF	$\Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right)$	$O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right)$
Larger family $\mathcal{SPM}_\lambda$	FAIR	$\Omega\left(\frac{m}{n\epsilon}\ln\frac{1}{\beta} + m\lambda\right)$	$O\left(\frac{m}{n\epsilon}\ln\frac{n\lambda}{\beta} + m\lambda\right)$
	SWDIFF	$\Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2 + m\lambda\right)$	$O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{n\lambda}{\beta}\right)^2 + \frac{m\lambda}{\epsilon}\right)$
General datasets	FAIR	$m/2$	–
	SWDIFF	–	–

Related Work

Our contributions connect to the broader literature on both DP facility location and intersections of fairness with DP.

DP Facility Location and Median-Finding. Gupta, Ligett, McSherry, Roth, and Talwar [20] were among the first to consider the problem of DP facility location across general metric spaces, proposing a solution based on the "exponential mechanism" [27] with a social welfare score function. Subsequent work [22, 15, 14, 23, 24] has offered tighter runtime and utility bounds, including for relaxed versions of facility location originally proposed by Gupta et al. Other papers have focused on developing algorithms for related forms of privacy like local DP [7] and metric DP [13]. Likewise, significant literature has emerged on the closely-related problem of accurate and sample-efficient DP median estimation [11, 19, 34, 31, 2, 28]. However, to our knowledge, we are the first in the domain to examine the *individual-utility loss vector* (which is used in our objective FAIR) in addition to the more-commonly studied social welfare loss function. Studying these two utility-based metrics simultaneously allows us to characterize a tradeoff not previously addressed.

Fairness under DP. The existing literature that analyzes fairness under DP constraints has done so with varying notions of fairness. The vast majority of research in this area has been for computer science contexts such as deep learning [16, 35], synthetic data [18], and federated learning [6], where fairness is typically characterized via standards from the field of algorithmic fairness [10]. We also analyze tradeoffs between privacy and fairness, but we do this within a more economic setting, where we use a comparison of agent utilities to measure fairness.

The two works most similar to ours are Pujol et al. [30] and Tran et al. [33], which study the impact of a DP data release (e.g., the release of noisy Census data) on the fairness of downstream economic problems. These papers both examine the tasks of fund allocation and voting rights decision rules. (Pujol et al. focus on simulations, while Tran et al. is a theoretical paper.) The measure of "bias" or unfairness utilized by both papers is similar in spirit to the one we propose, as it is also based on individual agent utility. We differ from Pujol et al. and Tran et al. in two primary ways. First, we analyze the facility location problem, which

has a distinct nature compared to allocation and classification, thereby presenting unique challenges in enforcing DP. Second, both previous works focus on mechanisms that follow a particular structure: those that first release summary statistics data in a DP manner before using the noisy data to solve the downstream problem. In this sense, those papers measure how adding DP to an upstream data release impacts the fairness of the later allocation or classification problem (where no additional privacy layers are imposed). Meanwhile, we consider mechanisms that *directly* protect the "downstream" facility location problem through DP and assess how much that affects both fairness and social welfare.

2 Model and Preliminaries

2.1 Facility Location Framework

The canonical one-dimensional facility location problem is defined as follows.

▶ **Definition 1** (One-Dimensional Continuous Facility Location Setting). *Let there be n agents located on a continuous one-dimensional metric space $V = [-m/2, m/2]$ of diameter $m > 0$ with metric $d(x, y) = |x - y|$.*

- *For each agent i, let $x_i \in V$ be the agent's location, and denote* a dataset of agent locations *$D \in V^n$ by $D = (x_1, x_2, \ldots, x_n)$. Moreover, assume without loss of generality that the indices are sorted by location, i.e., $x_1 \leq x_2 \leq \ldots \leq x_n$.*
- *Define utility and social welfare functions $u : V^n \times V \to \mathbb{R}^n$ and $s : V^n \times V \to \mathbb{R}$ that take as input a dataset D and a proposed facility location ℓ, as follows:*
  - *The utility function outputs an* individual utility vector *$u(D, \ell) \in \mathbb{R}^n$, where the ith component $u(D, \ell)_i = -d(x_i, \ell)$ is the utility of agent i given the facility placement ℓ and their location x_i.*
  - *The social welfare function outputs the overall* social welfare *$s(D, \ell)$ of the placement, measured as the sum of utilities across agents $\sum_{i=1}^{n} u(D, \ell)_i$.*
- *Assume that the number of agents n is odd. This is solely for ease of exposition; analogous results hold for even n via slight adjustments in theorem statements and proofs.*

While previous work [20, 22, 14] on differentially private facility location has examined more general metric spaces, we primarily consider one-dimensional continuous intervals as this constitutes the most fundamental instance of facility location where we can analyze the privacy, welfare, and fairness implications simultaneously.

▶ **Definition 2** (Facility location mechanisms). *Denote a general* facility location mechanism *by $\mathcal{M} : V^n \to V$, which takes in an input dataset of agent locations D and outputs (possibly in a randomized manner) a facility location $\mathcal{M}(D) \in V$.*

For every dataset D of agent locations, there exists a set $\mathcal{O}_{\mathcal{T}}(D) = \operatorname{argmax}_{\ell \in V} s(D, \ell)$ of optimal facility placements that maximize the resulting social welfare. This set $\mathcal{O}_{\mathcal{T}}(D)$ then induces an optimal facility location mechanism *$\mathcal{T}$, which takes as input a dataset D and outputs a facility location $\mathcal{T}(D)$ selected from $\mathcal{O}_{\mathcal{T}}(D)$.*

In the one-dimensional model, it is well-known that the set $\mathcal{O}_{\mathcal{T}}(D)$ of optimal facility location(s) occur at or between the most central points:

▶ **Proposition 3** (Optimal facility location is the median). *Over $V = [-m/2, m/2]$, the optimal set of facility locations for any dataset D, sorted as $x_1 \leq \cdots \leq x_n$, is given by*

$$\mathcal{O}_{\mathcal{T}}(D) = \begin{cases} \{x_{\lceil n/2 \rceil}\}, & n \text{ odd;} \\ [x_{\lfloor n/2 \rfloor}, x_{\lceil n/2 \rceil}], & n \text{ even.} \end{cases}$$

Proposition 3 clarifies why it is helpful notationally to adopt the assumption that n is odd: since $\mathcal{O}_{\mathcal{T}}(D)$ is a singleton when n is odd, $\mathcal{T}$ is deterministic and uniquely defined.

2.2 Differential Privacy

While there are many mechanisms for facility location, we are interested only in ones that satisfy *differential privacy* (DP) [12]. Roughly speaking, differential privacy requires that a mechanism treat similar datasets similarly, where we use the *change-one* notion of similarity.

▶ **Definition 4** (Change-one distance). *Let $D, D' \in V^n$ be two datasets that each contain n agent locations. Viewing D and D' as size-n multisets over elements in V, the* change-one distance $d_{co} : V^n \times V^n \to \mathbb{Z}_{\geq 0}$ *between D and D' is given by $d_{co}(D, D') = |D - D'|$ where $-$ denotes multiset difference, i.e., the number of elements (counting multiplicity) that lie in D but not D'. Two datasets $D, D' \in V^n$ are* neighboring *if $d_{co}(D, D') = 1$.*

The notion of neighboring datasets helps define the concept of differential privacy.

▶ **Definition 5** (Differentially Private Mechanism [9, 12]). *A mechanism $\mathcal{M} : V^n \to \mathcal{Y}$ that takes in datasets from V^n satisfies ϵ-differential privacy (ϵ-DP) when*

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(D') \in S]$$

for all neighboring datasets D, D' and subsets S of the mechanism's codomain $\mathcal{Y}$.

We typically require $\epsilon = O(1)$ so that DP provides meaningful privacy protection. Additionally, we generally assume $n \gg 1/\epsilon$ for an ϵ-DP mechanism $\mathcal{M}$; this is necessary for $\mathcal{M}$ to have nontrivial performance on the mechanism's primary goal (in our setting, this would be good social welfare). Note that we use the "pure" notion of DP like many related papers [20, 30, 33], but similar median-finding problems can behave quite differently under approximate DP [5].

2.3 Utility-based metrics for fairness and social welfare

We are interested in the tradeoff between privacy, social welfare, and fairness within facility location. We require privacy via differential privacy. We now define our metrics for social welfare and fairness. Both metrics compare a DP mechanism $\mathcal{M}$'s output location $\mathcal{M}(D)$ (or more generally, any proposed facility location ℓ) to the optimal but nonprivate mechanism $\mathcal{T}$'s output location $\mathcal{T}(D)$.

▶ **Definition 6** (SWDIFF and FAIR). *Both of the following functions take as input a dataset of agent locations D and a proposed facility location ℓ. In both, $\mathcal{T}$ is any optimal mechanism.*

1. Social welfare difference:

$$\text{SWDIFF}(D, \ell) = s(D, \mathcal{T}(D)) - s(D, \ell).$$

2. Maximum individual loss in utility:

$$\text{FAIR}(D, \ell, \mathcal{T}) = \max_{i \in \{1,2,\ldots,n\}} \text{LOSS}(D, \ell, \mathcal{T})_i,$$

where $\text{LOSS}(D, \ell, \mathcal{T}) = \mathbb{E}_T[u(D, \mathcal{T}(D))] - u(D, \ell)$ is the individual utility loss vector of switching the facility from $\mathcal{T}(D)$ to ℓ.

When $\mathcal{T}(D)$ is uniquely defined (i.e., when $\mathcal{O}_T(D)$ is a singleton), we can drop $\mathcal{T}$ as an input for FAIR *and simply denote* $\text{FAIR}(D, \ell, \mathcal{T})$ *as* $\text{FAIR}(D, \ell)$*. In particular, since $\mathcal{T}(D)$ is deterministic by our assumptions that n is odd and V is one-dimensional, we henceforth only write* $\text{FAIR}(D, \ell)$*.*

The formulation of SWDIFF is standard in the literature for measuring social welfare loss due to DP [20]. Meanwhile, FAIR is an unstudied metric for unfairness, reminiscent of the loss functions used in [30, 33] to analyze DP data releases. FAIR accounts for utility losses on an individual level. By penalizing for the worst-off agent's utility loss upon switching from $\mathcal{T}$ to ℓ, we ensure that no individual is disproportionately harmed by the imposition of privacy.

Observe that the smallest possible SWDIFF and FAIR are both 0, achieved by any optimal facility placement $\mathcal{T}(D)$, and the largest possible SWDIFF and FAIR are mn and m, respectively, where m is the diameter of the space and n is the number of agents.

2.4 Closed forms for FAIR and SWDIFF

It turns out the conceptually intuitive formulations of SWDIFF and FAIR can be transformed into mathematically useful ones. For one, FAIR under an outputted facility location ℓ can be written cleanly as its distance from the optimal location $\mathcal{T}(D)$.

▶ **Theorem 7** (FAIR closed form)**.** *For every dataset D and proposed facility location $\ell \in V = [-m/2, m/2]$, it holds that* $\text{FAIR}(D, \ell) = d(\mathcal{T}(D), \ell)$*.*

Meanwhile, to derive a closed form for SWDIFF, we must first characterize the set of agents most negatively impacted by a suboptimal facility placement.

▶ **Definition 8** (Set of crossed agents)**.** *For a given dataset D and proposed facility location $\ell \in V = [-m/2, m/2]$, define the set of* crossed agents $C(D, \ell)$ *by*

$$C(D,\ell) = \begin{cases} \{i : i < \lceil n/2 \rceil \text{ and } x_i \in [\ell, \mathcal{T}(D)]\}, & \text{if } \ell < \mathcal{T}(D) \\ \{i : i > \lceil n/2 \rceil \text{ and } x_i \in [\mathcal{T}(D), \ell]\}, & \text{if } \ell > \mathcal{T}(D) \\ \emptyset, & \text{if } \ell = \mathcal{T}(D) \end{cases}.$$

In words, $C(D, \ell)$ represent the indices of agents that are "crossed" when the facility location is moved along V from $\mathcal{T}(D)$ to ℓ.

Intuitively, $C(D, \ell)$ is the set of agents we have to "pay" for (in units of SWDIFF) when we move from $\mathcal{T}(D)$ to a suboptimal location ℓ.

▶ **Theorem 9** (SWDIFF closed form)**.** *For every dataset D and proposed facility location $\ell \in V = [-m/2, m/2]$, it holds that*

$$\text{SWDIFF}(D, \ell) = d(\mathcal{T}(D), \ell) + 2 \sum_{j \in C} d(x_j, \ell)$$

where $C = C(D, \ell)$ is the set of crossed agents.

The proofs of Theorems 7 and 9 are in the full version of the paper.

3 General incompatibility of privacy and fairness

We now turn towards the analysis of how DP mechanisms perform on SWDIFF and FAIR. We first demonstrate that a DP mechanism cannot possibly do well on fairness for all inputs.

▶ **Theorem 10** (Incompatibility of DP and fairness over all datasets). *For every ϵ-DP facility location mechanism $\mathcal{M} : V^n \to V$ on the interval $V = [-m/2, m/2]$, there exists a dataset D of agent locations for which*

$$\text{FAIR}(D, \mathcal{M}(D)) \geq m/2$$

with probability at least $\frac{1}{1+e^\epsilon}$.

Proof. We consider a pair of datasets D and D' that are commonly used to prove accuracy lower bounds for DP medians. Let D be a dataset with $\lceil n/2 \rceil$ agents at $-m/2$ and $\lfloor n/2 \rfloor$ agents at $m/2$, and let D' be a dataset with $\lfloor n/2 \rfloor$ agents at $-m/2$ and $\lceil n/2 \rceil$ agents at $m/2$. Note that D and D' can be viewed as neighbors since we can transform D into D' by moving one agent from $-m/2$ to $m/2$.

Since $\mathcal{T}(D) = -m/2$ and $\mathcal{T}(D') = m/2$, the closed form for FAIR from Theorem 7 implies that $\text{FAIR}(D, \mathcal{M}(D))$ is at least $m/2$ when $\mathcal{M}(D) \geq 0$ and $\text{FAIR}(D', \mathcal{M}(D'))$ is at least $m/2$ when $\mathcal{M}(D') \leq 0$. However, it holds by DP that $e^{-\epsilon}\Pr[\mathcal{M}(D) \leq 0] \leq \Pr[\mathcal{M}(D') \leq 0]$ so

$$\begin{aligned}\min(\Pr[\mathcal{M}(D) \geq 0], \Pr[\mathcal{M}(D') \leq 0]) &\geq \min(1 - \Pr[\mathcal{M}(D) \leq 0], e^{-\epsilon}\Pr[\mathcal{M}(D) \leq 0]) \\ &\geq \frac{e^{-\epsilon}}{1 + e^{-\epsilon}}\end{aligned}$$

where the second inequality follows by setting $\Pr[\mathcal{M}(D) \leq 0]$ equal to $\frac{1}{1+e^{-\epsilon}}$ to make the two terms in the min expression coincide. ◀

Under the typical expectation that $\epsilon \leq 1$ for DP mechanisms, the above theorem implies that any DP mechanism is bound to have egregious FAIR with probability at least $1/(1+\epsilon) > 25\%$ on some dataset.

4 Positive results on restricted datasets

Since Theorem 10 says that privacy is inherently incompatible with fairness across general datasets, we next consider a relaxation of the problem commonly adopted in the DP literature [11, 28, 3, 4, 2, 34] when preserving utility over all datasets is infeasible. Specifically, we consider smaller families of datasets where we can feasibly expect good (low) FAIR and SWDIFF while maintaining differential privacy over *all* datasets. Because the datasets that produce impossibility results are more "pathological" than what may we expect of real-world datasets, we might not actually need our DP mechanism to do well over those engineered datasets for it to be generally useful. In particular, we propose two "natural" families of datasets, $\mathcal{CTM}$ and $\mathcal{SPM}$, that guarantee the optimal facility location does not differ drastically across neighboring datasets.

▶ **Definition 11** ($\mathcal{CTM}$ family). *Call a dataset D* collapsing towards the median *if agents are packed increasingly tightly as they near the median (and optimal facility location) $x_{\lceil n/2 \rceil}$:*

- *$|x_{i+1} - x_i| \geq |x_{j+1} - x_j|$ for all $1 \leq i < j \leq \lceil n/2 \rceil - 1$, and*
- *$|x_i - x_{i-1}| \geq |x_j - x_{j-1}|$ for all $\lceil n/2 \rceil + 1 \leq j < i \leq n$.*

Define $\mathcal{CTM}$ as the family of all datasets $D \in V^n$ that are collapsing towards the median.

Datasets that collapse towards the median are appealing because they cleanly ensure concentration around the median agent and the nonexistence of two (or more) peaks of agent density, which help rule out the problematic datasets from the above example. However, the collapsing condition may be too stringent for real-world datasets to fully obey.

Thus, we also consider a larger family of datasets that are close to nice absolutely continuous density distributions P over the space V. For any distribution P, let F_P denote its CDF (so $F_P(\ell) = \Pr_{p\sim P}[p \le \ell]$) and let f_P denote its PDF.

▶ **Definition 12** (Kolmogorov-Smirnov Distance [25]). *Let D, D' be two distributions over the reals, and let $F, F' : \mathbb{R} \to [0,1]$ be their respective CDFs. The* Kolmogorov-Smirnov (K-S) distance *between F and F' is given by*

$$\mathrm{KS}(F, F') = \sup_{x\in\mathbb{R}} |F(x) - F'(x)|.$$

Notationally, we will also refer to the K-S distance between D and D' as $\mathrm{KS}(F, F')$.

▶ **Definition 13** ($\mathcal{SPM}$ family). *Call a density distribution P over V*single-peaked *at $\ell \in V$ if $f_P(x) \le f_P(y)$ for all $x, y \in V$ satisfying either $x \le y \le \ell$ or $x \ge y \ge \ell$. Let $\mathcal{P}$ be the class of all distributions over V that are single-peaked at $F_P^{-1}(0.5)$.*

For any fixed $\lambda \ge 0$, define $\mathcal{SPM}_\lambda$ as the family of all datasets $D \in V^n$ for which there exists a distribution $P \in \mathcal{P}$ such that the K-S distance between D and P is at most λ.

$\mathcal{SPM}$ has a natural distributional interpretation. If one interprets an observed dataset D as a set of n i.i.d. samples from a true single-peaked distribution $P \in \mathcal{P}$ of agent locations, then $\mathcal{SPM}_\lambda$ for $\lambda = \Theta(1/\sqrt{n})$ is the collection of datasets that, accounting for sampling error, could arise with nontrivial probability.

While $\mathcal{CTM}$ and $\mathcal{SPM}$ are similar to common distributional assumptions in the median-finding literature, they come with their limitations. Notably, the demand for single-peakedness excludes distributions with a stable median but multiple local peaks. Other papers [28, 34] have previously proposed alternative distributional requirements that better handle this case, and a natural follow-up would be to extend our work to those formulations.

5 Near-optimal mechanism over the relaxation

We now propose a DP mechanism $\texttt{DPExpMed}_\alpha$ that performs well on both SWDIFF and FAIR for the two dataset families. We build the mechanism on the (continuous) exponential mechanism of McSherry and Talwar [27], which uses a scoring function with low *sensitivity*.

▶ **Definition 14** (Global sensitivity with respect to data). *For any function $f : V^n \times \mathcal{P} \to \mathbb{R}$ that takes a dataset in V^n and a vector of parameters $p \in \mathcal{P}$, define the* global sensitivity *of f with respect to the data as*

$$\Delta_f = \max_{p\in\mathcal{P}} \max_{\text{neighboring } D, D'\in V^n} |f(D,p) - f(D',p)|$$

The sensitivity captures the worst-case pair of neighboring datasets and parameter configuration, where f exhibits the biggest change. The sensitivity of the function f we care about dictates the amount of noise that the DP mechanism $\mathcal{M}$ must have.

▶ **Definition 15** (Continuous Exponential Mechanism). *Suppose $s : V^n \times \mathcal{Y} \to \mathbb{R}$ is a scoring function with global sensitivity Δ_s over datasets D in V^n and outcomes y from a continuous outcome space $\mathcal{Y}$. Define the* continuous exponential mechanism $\mathcal{M} : V^n \to \mathcal{Y}$ *where the PDF $f_{\mathcal{M},D}$ of $\mathcal{M}(D)$ satisfies*

$$f_{\mathcal{M},D}(y) = \frac{\exp(\frac{\epsilon}{2\Delta_s} s(D,y))}{\int_{\mathcal{Y}} \exp(\frac{\epsilon}{2\Delta_s} s(x,z))dz}.$$

Intuitively, the exponential mechanism seeks to preserve the selected output's performance under the scoring function. To maintain privacy, it does not always take the best-scoring outcome, instead performing a smoothing such that the probability an outcome is selected increases exponentially with its score. In our case, we base the score off of how far the proposed facility placement is from the optimal placement, as is often done in DP mechanisms for the median.

▶ **Definition 16** (Percentile loss function). *Recall that we denote the ranked ordering of agent locations in D by $x_1 \leq x_2 \leq \cdots \leq x_n$. Define a* percentile loss function $q : V^n \times V \to \{0, 1, \ldots, \lceil n/2 \rceil\}$ *given by*

$$q(D,a) = \begin{cases} \min_{i \in \{1,2,\ldots n\}: a \in [x_i, \mathcal{T}(D)]} |\lceil n/2 \rceil - i|, & \text{if } a \in [x_1, \mathcal{T}(D)]; \\ \min_{i \in \{1,2,\ldots,n\}: a \in [\mathcal{T}(D), x_i]} |\lceil n/2 \rceil - i|, & \text{if } a \in (\mathcal{T}(D), x_n]; \\ \lceil n/2 \rceil, & \text{otherwise.} \end{cases}$$

Roughly speaking, q ranks the proposed location a among $x_1, x_2, \ldots, x_n$ and measures the number of points x_i between a and $x_{\lceil n/2 \rceil}$. We call q a percentile loss function because $\frac{1}{n} q(D,a)$ measures approximately how many percentile points away a is from $\mathcal{T}(D)$. Among all outputs $\ell \in V$, $\mathcal{T}(D)$ has the *lowest* value with respect to q, which makes q a loss function. We thus use $-q$ as the score in an exponential mechanism. We first consider a widened variant of our loss function to account for the continuity of V (see, e.g., [1]).

▶ **Definition 17** (Widened percentile loss function). *For any fixed $\alpha \in [0,1]$, define the* widened percentile loss function $p_\alpha(D, \ell) = \min_{a: |a-\ell| \leq \alpha m} q(D,a)$.

This widening technique is roughly equivalent to binning the continuous space V into bins of width αm. It smoothens the loss function over the output range V, which is necessary to ensure acceptable performance for datasets D with sharp concentration around $\mathcal{T}(D)$. Consider, for example, a dataset D^* where all n points are stacked at 0. Note that all outputs $\ell \neq 0$ look indistinguishable under q, but the widened p_α would credit a small band α around 0 as "good." This allows the following DP mechanism to have sufficient probability mass around the high quality outputs y close to $\mathcal{T}(D)$, even when the set of such y has small measure.

We are now ready to formulate the optimal DP mechanism, which is reminiscent of other DP mechanisms for the median [26].

▶ **Definition 18** (Widened percentile mechanism). *For any fixed α, define the facility location mechanism* $\texttt{DPExpMed}_\alpha$ *as an exponential mechanism with score $-p_\alpha$ where the PDF $f_p(D,\ell)$ of* $\texttt{DPExpMed}_\alpha(D)$ *is given by* $f_p(D,\ell) \propto \exp\left(-\frac{\epsilon}{2} p_\alpha(D,\ell)\right)$.

▶ **Lemma 19** ($\texttt{DPExpMed}_\alpha$ is ϵ-DP). *For every choice of $\alpha \in [0,1]$, the score p_α has sensitivity 1 and thus the mechanism* $\texttt{DPExpMed}_\alpha$ *is ϵ-DP.*

This proof is straightforward and is in the full paper version. As an aside, also note that $\texttt{DPExpMed}_\alpha$ has polynomial runtime $O(n \log n)$, which is not always the case for exponential mechanisms. Sorting the agents is the most expensive operation. Afterwards, scoring and sampling (with the associated probabilities) each only take a linear pass: because $p_\alpha(D, \cdot)$ is piecewise constant with $n+1$ pieces $P_0 \cup P_1 \cup \ldots \cup P_n = V$, the output of the exponential mechanism can be generated by first sampling a piece P_i with the correct probability (proportional to $|P_i| \cdot \exp(-\epsilon p/2)$, where p is the value of $p_\alpha(D,\ell)$ for all $\ell \in P_i$) and then sampling uniformly from P_i.

6 Performance bounds for $\texttt{DPExpMed}_\alpha$

We now benchmark $\texttt{DPExpMed}_\alpha$ by deriving high probability upper and lower bounds for $\texttt{DPExpMed}_\alpha$ on our two loss metrics FAIR and SWDIFF, restricted to datasets in $\mathcal{CTM}$ and $\mathcal{SPM}_\lambda$. The general framework for doing so is to first bound p_α, the loss directly controlling $\texttt{DPExpMed}_\alpha$'s outputs, and then translate this into bounds for FAIR and SWDIFF.

▶ **Theorem 20** (High probability upper bound on p_α for $\texttt{DPExpMed}_\alpha$ over $\mathcal{CTM}$)**.** *There exists a universal constant C such that for every $\alpha, \beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/(\alpha\beta))$, it holds for every $D \in \mathcal{CTM}$ that*

$$p_\alpha(D, \texttt{DPExpMed}_\alpha(D)) = O\left(\frac{1}{\epsilon} \max\left[1, \ln\left(\frac{1}{\alpha\beta n\epsilon}\right)\right]\right)$$

with probability at least $1 - \beta$.

▶ **Theorem 21** (High probability upper bound on p_α for $\texttt{DPExpMed}_\alpha$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/(\alpha\beta))$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$p_\alpha(D, \texttt{DPExpMed}_\alpha(D)) = O\left(\frac{1}{\epsilon} \max\left[1, \ln\left(\frac{\lambda}{\alpha\beta n}\right)\right]\right)$$

with probability at least $1 - \beta$.

These theorems provide, under usual DP size assumptions (small ϵ and large $n\epsilon$), a guarantee on the magnitude of p_α for $\mathcal{CTM}$ and $\mathcal{SPM}_\lambda$ with failure probability at most β.

Proof sketch for Theorem 21. The proof stems from (1) first constructing the dataset with the highest likelihood of producing a large p_α under $\texttt{DPExpMed}_\alpha$, and (2) demonstrating that even on the worse-case dataset, the value of p_α is bounded. More concretely, for step (1) we can show the following lemma.

▶ **Lemma 22.** *Fix $\lambda \geq 0$ and let $s = \lceil \lambda n \rceil - 1$. For every fixed $k \in \{0, 1, \ldots, \lfloor n/2 \rfloor\}$, the probability $\Pr[p_\alpha(D, \texttt{DPExpMed}_\alpha(D) \leq k]$ obtains its minimum across $D \in \mathcal{SPM}_\lambda$ on the dataset $D_k^{\mathcal{SPM}}$ with $\lceil n/2 \rceil + k$ points at $m/2$, s points at $-m/2$, and the remaining $\lfloor n/2 \rfloor - k - s$ points spaced at $-m/2 + i\frac{m}{\lfloor n/2 \rfloor - k + \lambda n}$ for $i = \{1, 2, \ldots, \lfloor n/2 \rfloor - k - s\}$. (See Figure 2.)*

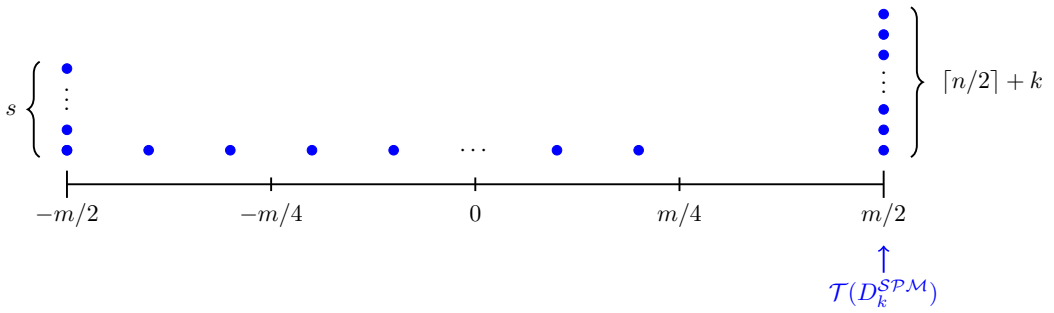


Figure 2 Depiction of Dataset $D_k^{\mathcal{SPM}}$, a worst-case dataset in $\mathcal{SPM}_\lambda$ for Lemma 22. Blue points indicate agent locations along V.

For Step (2), we can then reference the worst-case datasets from Lemma 22. For any fixed k, the probability $p_\alpha(D_k^{\mathcal{SPM}}, \texttt{DPExpMed}_\alpha(D_k^{\mathcal{SPM}}))$ exceeds k is an upper bound on the probability that $p_\alpha(D, \texttt{DPExpMed}_\alpha(D))$ exceeds k for every $D \in \mathcal{SPM}_\lambda$. We can then bound the former probability via the structure of the exponential mechanism. Using the geometric formula to collapse similar exponential terms, we can demonstrate that

$$\Pr[p_\alpha(D_k^{\mathcal{SPM}}, \texttt{DPExpMed}_\alpha(D_k^{\mathcal{SPM}})) > k] \leq \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \cdot \frac{1}{\alpha(\lfloor n/2 \rfloor - k)} \exp\left(-\frac{\epsilon}{2}(k+1)\right).$$

Upper bounding the right hand side by a desired failure probability β and inverting the inequality demonstrates that for every $D \in V^n$, the mechanism output $\texttt{DPExpMed}_\alpha(D)$ will satisfy

$$p_\alpha(D, \texttt{DPExpMed}_\alpha(D)) = O\left(\frac{1}{\epsilon}\max\left[1, \ln\frac{1}{\alpha\beta n\epsilon}\right]\right)$$

with probability at least $1-\beta$. ◀

Optimizing over α and relating p_α, FAIR, and SWDIFF via Section 2.4, we can transfer Theorems 20 and 21 into bounds for FAIR and SWDIFF seen in Table 1.

▶ **Theorem 23** (High probability bound on FAIR for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{CTM}$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0,1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$, it holds for every $D \in \mathcal{CTM}$ that*

$$\text{FAIR}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon}\ln\frac{1}{\beta}\right)$$

with probability at least $1-\beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

▶ **Theorem 24** (High probability bound on FAIR for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0,1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$ and $n \geq \beta/\lambda$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$\text{FAIR}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon}\ln\frac{n\lambda}{\beta} + m\lambda\right)$$

with probability at least $1-\beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

▶ **Theorem 25** (High probability bound on SWDIFF for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{CTM}$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0,1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$, it holds for every $D \in \mathcal{CTM}$ that*

$$\text{SWDIFF}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right)$$

with probability at least $1-\beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

▶ **Theorem 26** (High probability bound on SWDIFF for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0,1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$ and $n \geq \beta/\lambda$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$\text{SWDIFF}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{n\lambda}{\beta}\right)^2 + \frac{m\lambda}{\epsilon}\ln\frac{n\lambda}{\beta}\right)$$

with probability at least $1-\beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

For the full proof details over $\mathcal{SPM}_\lambda$, see the appendix; meanwhile, the proofs for $\mathcal{CTM}$ are analogous and available in the full paper version. Note that these bounds are considerable improvements over the one in Theorem 10. Moreover, these upper bounds are tight, which we show by proving matching lower bounds.

We generally construct lower bounds via the same intuition used in Theorem 10. To find the most adversarial pair of datasets D, D' within the dataset family, we roughly seek to maximize the difference between $\mathcal{T}(D)$ and $\mathcal{T}(D')$ while keeping $d_{co}(D, D')$ small. The fact that similar datasets must have similar outputs under $\mathcal{M}$ then implies unavoidably large FAIR under one of D or D'. Specifically, we use the following result.

▶ **Theorem 27** (Direct Lower Bound). *Let $D_1, D_2 \in V^n$ be two distinct datasets and let Y_1, Y_2 be disjoint subsets of the outcome space $\mathcal{Y}$. For every ϵ-DP mechanism $\mathcal{M} : V^n \to \mathcal{Y}$, it must hold that*

$$\min_{i\in\{1,2\}} \Pr[\mathcal{M}(D_i) \in Y_i] \leq \frac{e^{d_{co}(D_1,D_2)\epsilon}}{e^{d_{co}(D_1,D_2)\epsilon} + 1}$$

where $d_{co}(D_1, D_2)$ is the change-one distance between the two datasets.

The sets Y_i here represent the range of outputs that are "good" (as measured by amount of loss with respect to a loss metric) for a particular dataset D_i. Therefore, for two datasets with different "good output ranges," Theorem 27 demonstrates that the similarity in their output distributions under DP will limit the likelihood that the actual output falls into the desirable ranges.

▶ **Theorem 28** (Lower Bound for FAIR over $\mathcal{CTM}$). *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0, 1)$ satisfying $n\epsilon \geq C$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{CTM}$ that*

$$\mathrm{FAIR}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon} \ln \frac{1}{\beta}\right)$$

with probability at least β.

▶ **Theorem 29** (Lower Bound for FAIR over $\mathcal{SPM}_\lambda$). *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0, 1)$ satisfying $n\epsilon \geq C$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{SPM}_\lambda$ that*

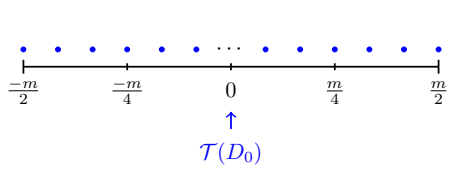
$$\mathrm{FAIR}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon} \ln \frac{1}{\beta} + m\lambda\right)$$

with probability at least β.

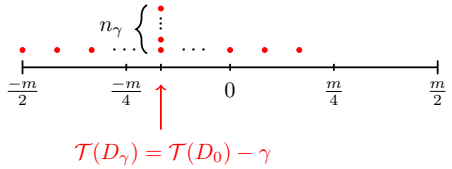
The proof for Theorem 28 is essentially a subset of the proof for Theorem 29. We provide a sketch of the latter and a full version can be found in the appendix.

Proof sketch for Theorem 29. We split the lower bound into two. We first show that FAIR is often at least $\Omega(m/(n\epsilon)\ln(1/\beta))$ on some dataset by considering $D_0, D_\gamma \in \mathcal{SPM}_\lambda$ of the visual structure depicted in Figure 3. We then show that FAIR is often at least $\Omega(m\lambda)$ on some dataset by considering neighboring $D_1, D_2 \in \mathcal{SPM}_\lambda$ with distant optimal facility locations; see Figure 4. Combining the two subresults gives the desired final lower bound.

The same datasets used to prove Theorems 28 and 29 also directly yield bounds over SWDIFF.

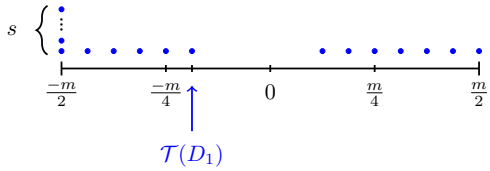


(a) Depiction of Dataset D_0. Blue points indicate agent locations along V.

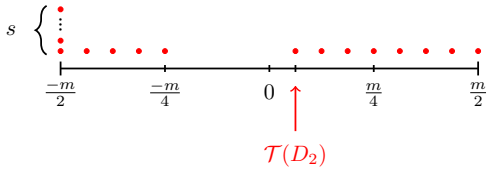


(b) Depiction of Dataset D_γ. Red points indicate agent locations along V.

Figure 3 The datasets D_0 and D_γ that imply a lower bound of $\Omega(m/(n\varepsilon)\ln(1/\beta))$.



(a) Depiction of Dataset D_1. Blue points indicate agent locations along V.



(b) Depiction of Dataset D_2. Red points indicate agent locations along V.

Figure 4 The datasets D_1 and D_2 that imply a lower bound of $\Omega(m\lambda)$. Observe that the two datasets share $n-1$ points and only differ in the median agent.

▶ **Theorem 30** (Lower Bound for SWDIFF over $\mathcal{CTM}$)**.** *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0,1)$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{CTM}$ that*

$$\mathrm{SWDIFF}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right)$$

with probability at least β.

▶ **Theorem 31** (Lower Bound for SWDIFF over $\mathcal{SPM}_\lambda$)**.** *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0,1)$ satisfying $n\epsilon \geq C$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{SPM}_\lambda$ that*

$$\mathrm{SWDIFF}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2 + m\lambda\right)$$

with probability at least β. ◀

7 Discussion

The lower bounds for $\mathcal{CTM}$ in Theorems 28 and 30 match exactly with their counterparts in Theorems 23 and 25, so $\texttt{DPExpMed}_\alpha$ is fully optimal on both FAIR and SWDIFF for $\mathcal{CTM}$. Meanwhile, the lower bounds on $\mathcal{SPM}_\lambda$ also match closely to the upper bounds from Theorems 24 and 26. The only discrepancies are the presence of additional $\max(1, \ln(n\lambda))$ factors in the upper bounds and an extra $1/\epsilon$ factor on $m\lambda$ in the SWDIFF upper bound. However, these extra terms are small enough that we would consider them largely unsubstantial. Indeed, under the common assumptions that $\epsilon = \Theta(1)$ and $\lambda = O(1/\sqrt{n})$, the max is $O(\ln n)$ and thus logarithmic relative to the factor of n in the overall term's denominator, and the $1/\epsilon$ factor is $\Theta(1)$. Both differences are minor enough that we consider the differentially private mechanism $\texttt{DPExpMed}_\alpha$ to be virtually optimal on both FAIR and SWDIFF for $\mathcal{SPM}_\lambda$.

Thus, $\texttt{DPExpMed}_\alpha$'s performance relative to the lower bounds on $\mathcal{CTM}$ and $\mathcal{SPM}_\lambda$ suggests that on each of our two families, FAIR and SWDIFF can be simultaneously optimized over DP mechanisms. Therefore, while there is a tradeoff between privacy and each of social welfare and fairness in facility location mechanism design, there is no additional tradeoff when we consider all three objectives simultaneously, provided that the population data is sufficiently natural.

There are a few directions for future work. First, one could focus on reconciling the small mismatches in bounds that still remain. Additionally, it may be worth analyzing this same tradeoff under different distributional assumptions used for similar relaxations in the DP median-finding literature [11, 28, 3, 4, 2, 34]. While we offer motivation for the families $\mathcal{CTM}$ and $\mathcal{SPM}_\lambda$, one weakness of our conditions is that the single-peakedness property may still be too strong and not strictly necessary for a DP facility location mechanism to achieve reasonable utility. Moreover, it would be insightful to expand this framework to more general facility location settings in higher dimensional metric spaces, as the social welfare and fairness scoring functions are likely less well-behaved in those more complex settings. Lastly, it could be interesting to consider how the performance and fairness guarantees change under looser forms of differential privacy, such as approximate DP or other variants.

References

1 Daniel Alabi, Audra McMillan, Jayshree Sarathy, Adam Smith, and Salil Vadhan. Differentially private simple linear regression. *Proceedings on Privacy Enhancing Technologies*, 2022. URL: https://petsymposium.org/popets/2022/popets-2022-0041.pdf.

2 Maryam Aliakbarpour, Rose Silver, Thomas Steinke, and Jonathan Ullman. Differentially private medians and interior points for non-pathological data. *arXiv preprint*, 2023. doi:10.48550/arXiv.2305.13440.

3 Christian Borgs, Jennifer Chayes, Adam Smith, and Ilias Zadik. Private algorithms can always be extended. *arXiv preprint*, 2018. arXiv:1810.12518.

4 Christian Borgs, Jennifer Chayes, Adam Smith, and Ilias Zadik. Revealing network structure, confidentially: Improved rates for node-private graphon estimation. In *IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, 2018. URL: https://www.computer.org/csdl/proceedings-article/focs/2018/423000a533/17D45XreC6C.

5 Mark Bun, Kobbi Nissim, Uri Stemmer, and Salil Vadhan. Differentially private release and learning of threshold functions. *arXiv preprint*, 2015. arXiv:1504.07553.

6 Huiqiang Chen, Tianqing Zhu Zhu, Tao Zhang, Wanlei Zhou, and Philip S. Yu. Privacy and fairness in federated learning: On the perspective of tradeoff. *ACM Computing Surveys*, 2023. URL: https://dl.acm.org/doi/pdf/10.1145/3606017.

7 Vincent Cohen-Addad, Yunus Esencayi, Chenglin Fan, Marco Gaboradi, Shi Li, and Di Wang. On facility location problem in the local differential privacy model. *Proceedings of Machine Learning Research*, 2022. URL: https://proceedings.mlr.press/v151/cohen-addad22a.html.

8 Z. Drezner and G. O. Wesolowsky. Facility location on a sphere. *Journal of the Operational Research Society*, 1978. URL: https://www.jstor.org/stable/3009474.

9 Cynthia Dwork. Differential privacy. In *Automata, Languages and Programming*. Springer, 2006. doi:10.1007/11787006_1.

10 Cynthia Dwork, Mortiz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 2012. URL: https://dl.acm.org/doi/10.1145/2090236.2090255.

11 Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. *arXiv preprint*, 2009. URL: https://www.stat.cmu.edu/~jinglei/dl09.pdf.

12 Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. Springer, 2006. doi:10.1007/11681878_14.

13 Allesandro Epasto, Vahab Mirrokni, Shyam Narayana, and Peilin Zhong. k-means clustering with distance-based privacy. *Advances in Neural Information Processing Systems*, 2023. URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/3e8d9bf1dd1eb9d3d9d500fb3543c87b-Paper-Conference.pdf.

14 Yunus Esencayi, Marco Gaboardi, Shi Li, and Di Wang. Facility location problem in differential privacy model revisited. *Advances in Neural Information Processing Systems*, 2019. URL: https://papers.nips.cc/paper_files/paper/2019/hash/41c576a3bac4220845f9427b002a2a9d-Abstract.html.

15 Chenglin Fan, Ping Li, and Xiaoyun Li. k-median clustering via metric embedding: Towards better initialization with differential privacy. *Neural Information Processing Systems*, 2023. URL: https://proceedings.neurips.cc/paper_files/paper/2023/file/e9a612969b4df241ff0d8273656bd5a4-Paper-Conference.pdf.

16 Tom Farrand, Fatemehsadat Mireshghallah, Sahib Singh, and Andrew Trask. Neither private nor fair: Impact of data imbalance on utility and fairness in differential privacy. In *Privacy-Preserving Machine Learning in Practice*, 2020. URL: https://dl.acm.org/doi/10.1145/3411501.3419419.

17 Itai Feigenbaum, Jay Sethuraman, and Chun Ye. Approximately optimal mechanisms for strategyproof facility location: Minimizing L_p norm of costs. *Mathematics of Operations Research*, 2016. URL: https://www.columbia.edu/~js1353/pubs/fsy.pdf.

18 Georgi Ganev, Bristena Oprisanu, and Emiliano De Cristofaro. Robin hood and matthew effects: Differential privacy has disparate impact on synthetic data. *arXiv preprint*, 2021. arXiv:2109.11429.

19 Jennifer Gillenwater, Matthew Joseph, and Alex Kulesza. Differentially private quantiles. *Proceedings of Machine Learning Research*, 2021. URL: https://proceedings.mlr.press/v139/gillenwater21a/gillenwater21a.pdf.

20 Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, and Kunal Talwar. Differentially private combinatorial optimization. *Society for Industrial and Applied Mathematics*, 2009. URL: https://www.cis.upenn.edu/~aaroth/Papers/privateapprox.pdf.

21 Harold Hotelling. Stability in competition. *The Economic Journal*, 1929. URL: https://www.jstor.org/stable/2224214.

22 Matthew Jones, Huy Lê Nguyễn, and Thy Nguyen. Differentially private clustering via maximum coverage. *arXiv preprint*, 2020. URL: https://arxiv.org/pdf/2008.12388.

23 George Z. Li, Dung Nguyen, and Anil Vullikanti. Differentially private partial set cover with applications to facility location. *International Joint Conference on Artificial Intelligence*, 2023. URL: https://www.ijcai.org/proceedings/2023/534.

24 Pasin Manurangsi. Improved lower bound for differentially private facility location. *Information Processing Letters*, 2025. URL: https://www.sciencedirect.com/science/article/pii/S0020019024000322.

25 Jr. Massey, Frank J. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 1951. URL: https://www.jstor.org/stable/2280095.

26 Frank McSherry. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In *ACM SIGMOD International Conference on Management of Data*, 2009. URL: https://www.microsoft.com/en-us/research/wp-content/uploads/2009/06/sigmod115-mcsherry.pdf.

27 Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *IEEE Symposium on Foundations of Computer Science*, 2007. URL: https://ieeexplore.ieee.org/document/4389483.

28 Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. *arXiv preprint*, 2011. URL: https://cs-people.bu.edu/ads22/pubs/NRS07/NRS07-full-draft-v1.pdf.

29 Camilo Ortiz-Astorquiza, Ivan Contreras, and Gilbert Laporte. Multi-level facility location problems. *European Journal of Operational Research*, 2018. URL: https://www.sciencedirect.com/science/article/abs/pii/S0377221717309323.

30 David Pujol, Ryan McKenna, Satya Kuppam, Michael Hay, Ashwin Machanavajjhala, and Gerome Miklau. Fair decision making using privacy-protected data. In *Conference on Fairness, Accountability, and Transparency*, 2020. URL: https://dl.acm.org/doi/abs/10.1145/3351095.3372872.

31 Kelly Ramsay, Aukosh Jagannath, and Shoja'eddin Chenouri. Differentially private multivariate medians. *arXiv preprint*, 2022. URL: https://arxiv.org/abs/2210.06459.

32 Jason Tang. The Best of Three Worlds? Privacy, Welfare, and Fairness for Facility Location Mechanism Design, 2025. Bachelor's thesis, Harvard College.

33 Cuong Tran, Ferdinando Fioretto, Pascal Van Hentenryck, and Zhiyan Yao. Decision making with differential privacy under a fairness lens. In *International Joint Conference on Artificial Intelligence*, 2021. URL: https://www.ijcai.org/proceedings/2021/0078.pdf.

34 Christos Tzamos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, and Ilias Zadik. Optimal private median estimation under minimal distributional assumptions. *Advances in Neural Information Processing Systems*, 2020. URL: https://proceedings.neurips.cc/paper/2020/file/21d144c75af2c3a1cb90441bbb7d8b40-Paper.pdf.

35 Archit Uniyal, Rakshit Naidu, Sasikanth Kotti, Sahib Singh, Patrik Joslin Kenfack, Fatemehsadat Mireshghallah, and Andrew Trask. DP-SGD vs PATE: Which has less disparate impact on model accuracy? *arXiv preprint*, 2021. arXiv:2106.12576.

A Performance upper bound proofs

We offer a more complete proof of Theorems 24 and 26 for $\mathcal{SPM}_\lambda$. The proofs for $\mathcal{CTM}$ are analogous and available in the full version of this paper.

We first characterize the "worst-case" datasets in $\mathcal{SPM}_\lambda$ under which it is most likely that $p_\alpha(D, \texttt{DPExpMed}_\alpha(D))$ is large. We take the following lemma for granted, but it can be proven via a greedy construction (details in full paper version).

▶ **Lemma 32.** *Fix $\lambda \geq 0$ and let $s = \lceil \lambda n \rceil - 1$. For every fixed $k \in \{0, 1, \ldots, \lfloor n/2 \rfloor\}$, the probability $\Pr[p_\alpha(D, \texttt{DPExpMed}_\alpha(D) \leq k]$ obtains its minimum across $D \in \mathcal{SPM}_\lambda$ on the dataset $D_k^{\mathcal{SPM}}$ with $\lceil n/2 \rceil + k$ points at $m/2$, s points at $-m/2$, and the remaining $\lfloor n/2 \rfloor - k - s$ points spaced at $-m/2 + i\frac{m}{\lfloor n/2 \rfloor - k + \lambda n}$ for $i = \{1, 2, \ldots, \lfloor n/2 \rfloor - k - s\}$.*

▶ **Theorem 33** (High probability upper bound on p_α for $\texttt{DPExpMed}_\alpha$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/(\alpha\beta))$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$p_\alpha(D, \texttt{DPExpMed}_\alpha(D)) = O\left(\frac{1}{\epsilon} \max\left[1, \ln\left(\frac{\lambda}{\alpha\beta n}\right)\right]\right)$$

with probability at least $1 - \beta$.

Proof of Theorem 21. For brevity, denote the $\texttt{DPExpMed}_\alpha$ mechanism by $\mathcal{M}_\alpha^{med}$ throughout. Let s be the largest integer for which $\lambda > s/n$. For every fixed $k \in \{0, 1 \ldots, \lfloor n/2 \rfloor\}$, applying the exponential mechanism weighting to the dataset $D_k^{\mathcal{SPM}}$ from Lemma 32 will result in an unnormalized mass of αm for the event $p_\alpha(D_k^{\mathcal{SPM}}, \mathcal{M}_\alpha^{med}(D_k^{\mathcal{SPM}})) \leq k$. Meanwhile, an overestimate on the unnormalized mass M for the entire outcome space is given by

$$M = \alpha m + \frac{m}{\lfloor n/2 \rfloor - k + n\lambda} \sum_{i=k+1}^{\lfloor n/2 \rfloor - s + 1} \exp\left(-\frac{\epsilon}{2} i\right) + \left(m - \frac{m(\lfloor n/2 \rfloor - s - k)}{\lfloor n/2 \rfloor - k + n\lambda}\right) \exp\left(-\frac{\epsilon}{2}(k+1)\right),$$

which simplifies via the geometric series formula to be

$$M = \alpha m + \left(\frac{m(1-\exp(-\frac{\epsilon}{2}(\lfloor n/2 \rfloor - s - k + 1)))}{(\lfloor n/2 \rfloor - k + n\lambda)(1-\exp(-\frac{\epsilon}{2}))} + \frac{m(n\lambda + s)}{\lfloor n/2 \rfloor - k + n\lambda} \right) \exp\left(-\frac{\epsilon}{2}(k+1)\right)$$
$$\leq \alpha m + m \cdot \frac{2n\lambda}{\lfloor n/2 \rfloor - k} \cdot \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \exp\left(-\frac{\epsilon}{2}(k+1)\right).$$

Therefore, we can establish the bound

$$\begin{aligned}
\Pr[p_\alpha(D_k^{\mathcal{SPM}}, \mathcal{M}_\alpha^{med}(D_k^{\mathcal{SPM}})) > k] &= 1 - \Pr[p_\alpha(D_k^{\mathcal{SPM}}, \mathcal{M}_\alpha^{med}(D_k^{\mathcal{SPM}})) \leq k] \\
&\leq 1 - \frac{\alpha m}{M} \\
&\leq \frac{M - \alpha m}{\alpha m} \\
&\leq \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \cdot \frac{1}{\alpha(\lfloor n/2 \rfloor - k)} \exp\left(-\frac{\epsilon}{2}(k+1)\right)
\end{aligned}$$

from which we are assured

$$\Pr[p_\alpha(D, \mathcal{M}_\alpha^{med}(D)) > k] \leq \beta$$

for all $k \in \{0, 1, \ldots, \lfloor n/2 \rfloor\}$ that satisfy

$$\frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \cdot \frac{1}{\alpha(\lfloor n/2 \rfloor - k)} \exp\left(-\frac{\epsilon}{2}(k+1)\right) \leq \beta,$$

or equivalently,

$$k \geq -1 + \frac{2}{\epsilon} \ln\left(\frac{2n\lambda}{\alpha\beta(\lfloor n/2 \rfloor - k)} \cdot \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \right). \tag{1}$$

If $k = 0$ makes the right hand side negative, then we immediately obtain the desired result. Otherwise, if this is not the case, we can consider the nonnegative quantity

$$k^* = \left\lceil -1 + \frac{2}{\epsilon} \ln\left(\frac{4n\lambda}{\alpha\beta\lfloor n/2 \rfloor} \cdot \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \right) \right\rceil,$$

which we claim satisfies Inequality (1). Again noting that

$$\frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \leq \lfloor n/2 \rfloor$$

implies

$$k^* \leq \left\lceil -1 + \frac{2}{\epsilon} \ln\left(\frac{4n\lambda}{\alpha\beta} \right) \right\rceil.$$

Therefore, there must exist a constant $C' \geq 2$ for which $\lfloor n/2 \rfloor \geq C' \cdot k^*$ if we choose a sufficiently large constant C such that $n\epsilon \geq C \cdot \ln(1/(\alpha\beta))$. Substituting into the previous inequality demonstrates that k^* fulfills Inequality (1). It thus holds that with probability at least $1 - \beta$, we have

$$\begin{aligned}
p_\alpha(D, \mathcal{M}_\alpha^{med}(D)) \leq \max(0, k^*) &= O\left(\frac{1}{\epsilon} \max\left[1, \ln\left(\frac{\lambda}{\alpha\beta} \cdot \frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} \right) \right] \right) \\
&= O\left(\frac{1}{\epsilon} \max\left[1, \ln\left(\frac{\lambda}{\alpha\beta n} \right) \right] \right)
\end{aligned}$$

by noting that $\frac{1-\exp(-\frac{\epsilon}{2}\lfloor n/2 \rfloor)}{1-\exp(-\frac{\epsilon}{2})} = O(1/\epsilon)$ as before. ◀

This theorem holds for general α but we will later optimize over α to produce the tightest possible result for FAIR and SWDIFF. It turns out we will select $\alpha = 1/(n\epsilon)$, upon which one can verify that Theorem 21 simplifies to the following corollary.

▶ **Corollary 34** (Theorem 20 with $\alpha = 1/(n\epsilon)$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$ and $n \geq \beta/\lambda$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$p_\alpha(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{1}{\epsilon} \ln \frac{n\lambda}{\beta}\right)$$

with probability at least $1 - \beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

Observe that the $n \geq \beta/\lambda$ condition is necessary for technical precision, but does not impose any "new" constraint since we generally assume $\lambda = \Theta(1/\sqrt{n})$.

We now transfer the bound from p_α over to FAIR by relating the two loss functions. We first utilize a lemma which formalizes the idea that the "most spread out" (with respect to its median) distribution $P \in \mathcal{P}$ is the uniform distribution.

▶ **Lemma 35.** *Let F be the CDF for some distribution $P \in \mathcal{P}$. Then, for every $\ell \in V$, it holds that*

$$|F^{-1}(0.5) - \ell| \leq 2m \cdot |0.5 - F(\ell)|.$$

This result now allows us to relate density of agents, represented by expressions involving F, to distances over $\mathcal{P}$, represented by expressions involving F^{-1}. This gives us a mapping between the bound on p_α to one on FAIR.

▶ **Theorem 36** (High probability bound on FAIR for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$ and $n \geq \beta/\lambda$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$\text{FAIR}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon} \ln \frac{n\lambda}{\beta} + m\lambda\right)$$

with probability at least $1 - \beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

Proof. Let P be a distribution in $\mathcal{P}$ that D is λ away from in K-S distance, and let F, F_n be the CDFs of P and D respectively. Set $F^{-1}(0.5) = t$ and $F_n^{-1}(0.5) = t_n$ with $\eta = |t_n - t|$. Finally, let $a = \text{argmin}_{a:|a-\ell| \leq \alpha m} q(D, a)$. We now have

$$\begin{aligned}
\text{FAIR}(D, \ell) = |t_n - \ell| &\leq |t_n - a| + \alpha m \\
&\leq |t - a| + \eta + \alpha m \\
&\leq m|0.5 - F(a)| + \eta + \alpha m \\
&\leq m\left(|F_n(t_n) - F(a)| + \frac{1}{n}\right) + \eta + \alpha m \\
&\leq m\left(|F_n(t_n) - F_n(a)| + \frac{1}{n} + \lambda\right) + \eta + \alpha m \\
&\leq \frac{m}{n}(p(D, \ell) + 1) + m\lambda + \eta + \alpha m
\end{aligned}$$

where we utilize Lemma 35. We can bound η via another application of Lemma 35. We have that

$$\begin{aligned}\eta = |t - t_n| &\leq 2m|F(t) - F(t_n)| \\ &\leq 2m(|F(t) - F_n(t_n)| + |F_n(t_n) - F(t_n)|) \leq 2m\left(\frac{1}{n} + \lambda\right)\end{aligned}$$

so the previous inequality becomes

$$\text{FAIR}(D, \ell) \leq \frac{m}{n}(p(D, \ell) + 1) + \frac{2m}{n} + 3m\lambda + \alpha m.$$

Then, via the result of Theorem 33, we have that $\text{FAIR}(D, \ell)$ is at most

$$\frac{m}{n} \cdot O\left(\frac{1}{\epsilon}\max\left[1, \ln\left(\frac{\lambda}{\alpha\beta n}\right)\right]\right) + 3m\lambda + \alpha m \tag{2}$$

with probability $1 - \beta$ for a general widening parameter α if $n\epsilon$ is sufficiently larger than $\ln(1/\alpha)$. As discussed in Corollary 34, we can select $\alpha = \alpha^* = 1/(n\epsilon)$ and apply Theorem 33 if $n\epsilon$ is bigger than a sufficiently large constant. Combining Corollary 34 with (2) immediately gives for any $D \in D_\lambda$ that

$$\text{FAIR}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon}\ln\frac{n\lambda}{\beta} + m\lambda\right)$$

with probability at least $1 - \beta$. ◀

We can build off Theorems 33 and 36 to obtain a bound on SWDIFF.

▶ **Theorem 37** (High probability bound on SWDIFF for $\texttt{DPExpMed}_{\alpha^*}$ over $\mathcal{SPM}_\lambda$)**.** *There exists a universal constant C such that for every $\beta \in (0, 1/3), \epsilon \in (0, 1)$, and $n \in \mathbb{N}$ satisfying $n\epsilon \geq C \cdot \ln(1/\beta)$ and $n \geq \beta/\lambda$, it holds for every $D \in \mathcal{SPM}_\lambda$ that*

$$\text{SWDIFF}(D, \texttt{DPExpMed}_{\alpha^*}(D)) = O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{n\lambda}{\beta}\right)^2 + \frac{m\lambda}{\epsilon}\ln\frac{n\lambda}{\beta}\right)$$

with probability at least $1 - \beta$ for widening parameter $\alpha^* = 1/(n\epsilon)$.

Proof. Note that $\text{SWDIFF}(D, \ell) \leq (2p_\alpha(D, \ell) - 1) \cdot \text{FAIR}(D, \ell)$. Like in Theorem 36, we set $\alpha = 1/(n\epsilon)$. Combining the high probability bounds for p_{α^*} and FAIR from Corollary 34 and Theorem 36, we have that $\text{SWDIFF}(D, \texttt{DPExpMed}_{\alpha^*}(D))$ is at most

$$O\left(\frac{1}{\epsilon}\ln\frac{n\lambda}{\beta}\right) \cdot O\left(\frac{m}{n\epsilon}\max\frac{n\lambda}{\beta} + m\lambda\right) = O\left(\frac{m}{n\epsilon^2}\left(\ln\frac{n\lambda}{\beta}\right)^2 + \frac{m\lambda}{\epsilon}\max\ln\frac{n\lambda}{\beta}\right)$$

as required. ◀

B Performance lower bound proofs

We also provide more complete proofs of Theorems 28 and 30 for $\mathcal{CTM}$; the $\mathcal{SPM}_\lambda$ proofs are extensions of similar flavor and are available in the full paper version.

▶ **Theorem 38** (Lower Bound for FAIR over $\mathcal{CTM}$). *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0,1)$ satisfying $n\epsilon \geq C$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{CTM}$ that*

$$\mathrm{FAIR}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon} \ln \frac{1}{\beta}\right)$$

with probability at least β.

Proof. Let D_0 be the dataset which evenly spaces n agents over V so that $x_1 = -m/2$ and $x_n = m/2$, with $|x_{i+1} - x_i| = m/(n-1)$ for all $i \in \{1, \ldots, n-1\}$. Meanwhile, consider values of $\gamma < m/3$ that are integer multiples of $m/(n-1)$, noting that at least one such γ always exists when $n \geq 5$. For any fixed such γ, let D_γ be the dataset with $n_\gamma = 1 + 2\gamma(n-1)/m$ agents stacked at $\mathcal{T}(D_0) - \gamma$ and one agent at each of

$$\mathcal{T}(D_0) - \gamma + j\frac{m}{n-1}$$

for all

$$j \in \left\{-\frac{n-n_\gamma}{2}, -\frac{n-n_\gamma}{2}+1, \ldots, \frac{n-n_\gamma}{2}-1, \frac{n-n_\gamma}{2}\right\} - \{0\}.$$

Reference Figure 3. Note that D_γ is symmetric about $\mathcal{T}(D_0) - \gamma$ so $\mathcal{T}(D_\gamma) = \mathcal{T}(D_0) - \gamma$. Additionally, since D_γ is peaked at its median and otherwise uniform, and D_0 is fully uniform, both datasets lie in $\mathcal{CTM}$ as needed. Moreover, because γ is a multiple of $m/(n-1)$, many agent locations overlap between D_0 and D_γ. In particular, one can construct D_γ from D_0 by simply moving the rightmost $n_\gamma - 1 = 2\gamma(n-1)/m$ agents in D_0 and putting them all at $\mathcal{T}(D_\gamma)$ in D_γ, as visualized in Figure 5.

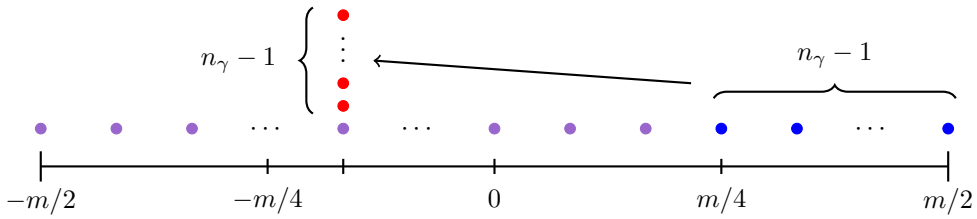


Figure 5 A depiction of how to transform D_0 into D_γ by moving $n_\gamma - 1$ agents. The purple points denote agent locations shared across D_0 and D_γ, whereas the blue points denote agent locations in D_0 that get moved to the red points in D_γ.

Therefore, the change-one distance between D_0 and D_γ is $h = \gamma(n-1)/m$. Let $Y_0 = (\mathcal{T}(D_0) - \gamma/2, \mathcal{T}(D_0) + \gamma/2)$ and $Y_\gamma = (\mathcal{T}(D_\gamma) - \gamma/2, \mathcal{T}(D_\gamma) + \gamma/2)$ be intervals along V. Note that Y_γ's left endpoint is well-defined by the condition $\gamma < m/3$. Then, by Theorem 27,

$$\min_{i \in \{0,\gamma\}} \Pr[\mathcal{M}(D_i) \in Y_i] \leq \frac{e^{h\epsilon}}{1 + e^{h\epsilon}}$$

which means

$$\max_{i \in \{0,\gamma\}} \Pr[\mathrm{FAIR}(D_i, \mathcal{M}(D_i)) > \gamma/2] = \max_{i \in \{0,\gamma\}} \Pr[\mathcal{M}(D_i) \notin Y_i] > \frac{1}{1 + e^{h\epsilon}} \geq \frac{1}{2e^{h\epsilon}}.$$

There then exists a $\gamma = \Theta\left(\frac{m}{n\epsilon} \ln \frac{1}{\beta}\right)$ for which $\max_{i \in \{0,\gamma\}} \Pr[\mathrm{FAIR}(D_i, \mathcal{M}(D_i)) > \gamma/2]$ exceeds β, implying the result for n and ϵ where $n\epsilon$ is sufficiently large (so that $\gamma < m/3$). ◀

▶ **Theorem 39** (Lower Bound for SWDIFF over $\mathcal{CTM}$). *There exist $\beta > 0$ and C such that for every $n \geq 5$, every $\epsilon \in (0,1)$, and every ϵ-DP mechanism $\mathcal{M}$, it holds for some $D \in \mathcal{CTM}$ that*

$$\text{SWDIFF}(D, \mathcal{M}(D)) = \Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right)$$

with probability at least β.

Proof. Define datasets D_0 and D_γ as in Theorem 38. Utilizing the closed form for SWDIFF given in Theorem 9 and the structure of D_0 and D_γ, we can write $\text{SWDIFF}(D, y)$ as a function of $d(\mathcal{T}(D), y) = \text{FAIR}(D, y)$ for these two datasets. First, consider D_0 and suppose $\text{FAIR}(D_0, \mathcal{M}(D_0)) = r$ where r is an integer multiple of $m/(n-1)$. Assume without loss of generality that $\mathcal{M}(D_0) < \mathcal{T}(D_0)$. Then, the set of crossed agents (from Definition 8) corresponds to the agents located at $\mathcal{T}(D_0) - j\frac{m}{n-1}$ for $j = \{1, 2, \ldots, r(n-1)/m\}$. Therefore, by Theorem 9, we have that

$$\begin{aligned}\text{SWDIFF}(D_0, \mathcal{M}(D_0)) &= r + 2\sum_{j=1}^{\frac{r(n-1)}{m}}\left(j \cdot \frac{m}{n-1}\right) \\ &= \Theta\left(\sum_{j=1}^{\frac{r(n-1)}{m}}\left(j \cdot \frac{m}{n-1}\right)\right) = \Theta\left(\frac{r^2 n}{m}\right).\end{aligned}$$

Now, consider D_γ and again suppose $\text{FAIR}(D_\gamma, \mathcal{M}(D_\gamma)) = r$ for r that is an integer multiple of $m/(n-1)$. The set of crossed agents would now be comprised of (1) the agents located between $\mathcal{M}(D_\gamma)$ and $\mathcal{T}(D_\gamma)$ that are $jm/(n-1)$ away from $\mathcal{T}(D_\gamma)$ for $j = \{1, 2, \ldots, r(n-1)/m\}$ (as in the case of D_0), and (2) half of the non-median agents located at $\mathcal{T}(D_\gamma)$, of which there are $2\gamma(n-1)/m$ total. Consequently, by Theorem 9 again,

$$\text{SWDIFF}(D_\gamma, \mathcal{M}(D_\gamma)) = \Theta\left(\frac{\gamma(n-1)}{m}r + \sum_{j=1}^{\frac{r(n-1)}{m}}\left(j \cdot \frac{m}{n-1}\right)\right) = \Theta\left(\frac{n}{m}r(r+\gamma)\right).$$

When $r = \Omega(\gamma)$, then both $\text{SWDIFF}(D_0, \mathcal{M}(D_0))$ and $\text{SWDIFF}(D_\gamma, \mathcal{M}(D_\gamma))$ are $\Theta\left(nr^2/m\right)$. By Theorem 28, there is an at least β probability that one of $\text{FAIR}(D_0, \mathcal{M}(D_0))$ and $\text{FAIR}(D_\gamma, \mathcal{M}(D_\gamma))$ is $\Omega(m\ln(1/\beta)/(n\epsilon))$ when $n\epsilon$ is assumed to be sufficiently large, which translates to an at least β probability that one of $\text{SWDIFF}(D_0, \mathcal{M}(D_0))$ and $\text{SWDIFF}(D_\gamma, \mathcal{M}(D_\gamma))$ is

$$\Omega\left(\frac{n}{m}\left(\frac{m}{n\epsilon}\ln\frac{1}{\beta}\right)^2\right) = \Omega\left(\frac{m}{n\epsilon^2}\left(\ln\frac{1}{\beta}\right)^2\right),$$

as claimed. ◀