Supplementary Materials

A Additional details about related research

In Hsu et al.'s paper on private linear programming (Hsu et al., 2014), the authors study two different categories of linear program (LPs) when the constraint vector $\boldsymbol{b}(D)$ depends on private data:

1. *High-sensitivity LPs*: For any two neighboring databases D and D', there is exactly one component $i \in [m]$ where $\boldsymbol{b}(D)_i \neq \boldsymbol{b}(D')_i$, and for every other component $j \neq i$, $\boldsymbol{b}(D)_j = \boldsymbol{b}(D')_j$.
2. *Low-sensitivity LPs*: For any two neighboring databases D and D' of size N, $\|\boldsymbol{b}(D) - \boldsymbol{b}(D')\|_\infty \leq \frac{1}{N}$.

They prove that in general, high-sensitivity LPs cannot be solved privately. Specifically, for any database $D \in \{0,1\}^n$ of size n, they define the following high-sensitivity LP:

$$\begin{array}{ll} \text{find} & \boldsymbol{x} \\ \text{such that} & x_i = D_i \quad \text{for all } i \in [n]. \end{array}$$

They prove that for any (ϵ, δ)-differentially private mechanism with output $\boldsymbol{x}(D)$, there will be at least one component $i \in [n]$ such that $|\boldsymbol{x}(D)_i - D_i| \geq \frac{1}{2}$ (otherwise, the mechanism would be able to reconstruct D exactly). This fact does not contradict our upper bound from Theorem 3.4 since this worst-case problem does not satisfy Assumption 3.1.

For low-sensitivity LPs, they show that a private version of the multiplicative weights algorithm returns a solution that is close to the LP's optimal solution. Their solution is allowed to violate the LP's constraints, so their algorithm does not apply in our setting.

B Omitted proofs from Section 3 about multi-dimensional optimization

Theorem 3.2. *The mapping $D \mapsto \boldsymbol{b}(D) - s + \boldsymbol{\eta}$ preserves (ϵ, δ)-differential privacy.*

Proof. Let D and D' be two neighboring databases. We write the density function of $\boldsymbol{b}(D) - s + \boldsymbol{\eta}$ as $f_D(\boldsymbol{u}) \propto \prod_{i=1}^m \exp\left(-\frac{\epsilon|u_i+s-b(D)_i|}{\Delta}\right)$ when $\boldsymbol{u} \in [\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)]$ and $f_D(\boldsymbol{u}) = 0$ when $\boldsymbol{u} \notin [\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)]$. This proof relies on the following two claims. The first claim shows that in the intersection of the supports $[\boldsymbol{b}(D)-2s, \boldsymbol{b}(D)] \cap [\boldsymbol{b}(D')-2s, \boldsymbol{b}(D')]$, the density functions f_D and $f_{D'}$ are close.

Claim B.1. *Let $\boldsymbol{u}$ be a vector in the intersection of the supports $[\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)] \cap [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$. Then $f_D(\boldsymbol{u}) \leq e^\epsilon f_{D'}(\boldsymbol{u})$.*

Proof of Claim B.1. Since $\boldsymbol{u}$ is a vector in the intersection of the support $[\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)] \cap [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$,

$$\begin{aligned}
\frac{f_D(\boldsymbol{u})}{f_{D'}(\boldsymbol{u})} &= \prod_{i=1}^m \frac{\exp\left(-\epsilon|u_i + s - b(D)_i|/\Delta\right)}{\exp\left(-\epsilon|u_i + s - b(D')_i|/\Delta\right)} \\
&= \prod_{i=1}^m \exp\left(\frac{\epsilon\left(|u_i + s - b(D')_i| - |u_i + s - b(D)_i|\right)}{\Delta}\right) \\
&\leq \prod_{i=1}^m \exp\left(\frac{\epsilon|b(D)_i - b(D')_i|}{\Delta}\right) \\
&= \exp\left(\frac{\epsilon\sum_{i=1}^m |b(D)_i - b(D')_i|}{\Delta}\right) \\
&\leq \exp\left(\frac{\epsilon\Delta}{\Delta}\right) \\
&= e^\epsilon,
\end{aligned}$$

as claimed. □

The second claim shows that the total density of $\boldsymbol{b}(D) - s + \boldsymbol{\eta}$ on vectors not contained in the support of $\boldsymbol{b}(D') - s + \boldsymbol{\eta}$ is at most δ.

Claim B.2. *Let $V = [\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)] \setminus [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$ be the set of vectors in the support of $\boldsymbol{b}(D) - s + \boldsymbol{\eta}$ but not in the support of $\boldsymbol{b}(D') - s + \boldsymbol{\eta}$. Then $\mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in V] \leq \delta$.*

Proof of Claim B.2. Suppose $\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in V$. Then for some $i \in [m]$, either $b(D)_i - s + \eta_i < b(D')_i - 2s$ or $b(D)_i - s + \eta_i > b(D')_i$. This implies that either $\eta_i < -s + \Delta$ or $\eta_i > s - \Delta$. The density function of the truncated Laplace distribution with support $[-s, s]$ and scale $\frac{\Delta}{\epsilon}$ is

$$f(\eta) = \begin{cases} \frac{1}{Z}\exp\left(-\frac{|\eta|\epsilon}{\Delta}\right) & \text{if } \eta \in [-s, s] \\ 0 & \text{otherwise,} \end{cases}$$

where $Z = \frac{2\Delta\left(1 - e^{-\epsilon s/\Delta}\right)}{\epsilon}$ is a normalizing constant. Therefore, the probability that for some $i \in [m]$, either $\eta_i < -s + \Delta$ or $\eta_i > s - \Delta$ is

$$\begin{aligned}
m\left(\int_{-s}^{-s+\Delta} f(\eta)\, d\eta + \int_{s-\Delta}^{s} f(\eta)\, d\eta\right) &= \frac{m}{Z}\left(\int_{-s}^{-s+\Delta} \exp\left(-\frac{|\eta|\epsilon}{\Delta}\right) d\eta + \int_{s-\Delta}^{s} \exp\left(-\frac{|\eta|\epsilon}{\Delta}\right) d\eta\right) \\
&= \frac{2m\Delta(e^\epsilon - 1)e^{-s\epsilon/\Delta}}{Z\epsilon} \\
&= \frac{m\left(e^\epsilon - 1\right)e^{-s\epsilon/\Delta}}{1 - e^{-\epsilon s/\Delta}} \\
&= \frac{m\left(e^\epsilon - 1\right)}{e^{s\epsilon/\Delta} - 1} \\
&= \delta,
\end{aligned}$$

where the final equality follows from the fact that $s = \frac{\Delta}{\epsilon} \ln\left(\frac{m(e^\epsilon - 1)}{\delta} + 1\right)$. In turn, this implies that $\mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in V] \leq \delta$. □

These two claims imply that the mapping $D \mapsto \boldsymbol{b}(D) - s + \boldsymbol{\eta}$ preserves (ϵ, δ)-differential privacy. To see why, let $W \subseteq [\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)]$ be an arbitrary set of vectors in the support of $\boldsymbol{b}(D) - s + \boldsymbol{\eta}$. Let $W_0 = W \cap [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$ be the set of vectors in W that are also in the support of $\boldsymbol{b}(D') - s + \boldsymbol{\eta}$ and let $W_1 = W \setminus [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$ be the remaining set of vectors in W. As in Claim B.2, let $V = [\boldsymbol{b}(D) - 2s, \boldsymbol{b}(D)] \setminus [\boldsymbol{b}(D') - 2s, \boldsymbol{b}(D')]$ be the set of vectors in the support of $\boldsymbol{b}(D) - s + \boldsymbol{\eta}$ but not in the support of $\boldsymbol{b}(D') - s + \boldsymbol{\eta}$. Clearly, $W_1 \subseteq V$. Therefore,

$$\begin{aligned}
\mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in W] &= \mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in W_0] + \mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in W_1] \\
&\leq \mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in W_0] + \mathbb{P}[\boldsymbol{b}(D) - s + \boldsymbol{\eta} \in V] \\
&= \int_{W_0} f_D(\boldsymbol{u})\, d\boldsymbol{u} + \int_V f_D(\boldsymbol{u})\, d\boldsymbol{u} \\
&\leq \int_{W_0} e^\epsilon f_{D'}(\boldsymbol{u})\, d\boldsymbol{u} + \int_V f_D(\boldsymbol{u})\, d\boldsymbol{u} && \text{(Claim B.1)} \\
&\leq \int_{W_0} e^\epsilon f_{D'}(\boldsymbol{u})\, d\boldsymbol{u} + \delta && \text{(Claim B.2)} \\
&\leq e^\epsilon\, \mathbb{P}[\boldsymbol{b}(D') - s + \boldsymbol{\eta} \in W] + \delta,
\end{aligned}$$

so differential privacy is preserved. □

Lemma B.3. *Suppose D and D' are two neighboring databases with disjoint feasible regions: $\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\} \cap \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D')\} = \emptyset$. There is no (ϵ, δ)-DP mechanism with $\delta < 1$ that satisfies the constraints with probability 1.*

Proof. For the sake of a contradiction, suppose $\mu : 2^{\mathcal{X}} \to \mathbb{R}^n$ is an (ϵ, δ)-DP mechanism with $\delta < 1$ that satisfies the constraints with probability 1. Let $V = \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$. Since $V \cap \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D')\} = \emptyset$, it must be that $\mathbb{P}[\mu(D') \in V] = 0$. This means that $1 = \mathbb{P}[\mu(D) \in V] \leq e^\epsilon\, \mathbb{P}[\mu(D') \in V] + \delta = \delta$, which is a contradiction. Therefore, the lemma statement holds. □

Lemma B.4. *With probability 1, the optimization problem in Equation* (3) *is feasible.*

Proof. By definition, the constraint vector $\bar{\boldsymbol{b}}$ is component-wise greater than the vector $\boldsymbol{b}^* = (b_1^*, \dots, b_m^*)$, where $b_i^* = \inf_{D \subseteq \mathcal{X}} b(D)_i$. Therefore, $\left\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \bar{\boldsymbol{b}}(D)\right\} \supseteq \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*\}$. By Lemma B.5, we know that $\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*\} = \bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$ and by Assumption 3.1, we know that $\bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$ is nonempty. Therefore, the feasible set of the linear program in Equation (3), $\left\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \bar{\boldsymbol{b}}(D)\right\}$, is nonempty. □

We now prove Lemma B.5, which we used in the proof of Lemma B.4. Lemma B.5 guarantees that the (nonempty) intersection of the feasible regions across all databases is equal to the set of all $\boldsymbol{x}$ such that $\mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*$.

Lemma B.5. *The set $\bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$ is equal to the set $\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*\}$.*

Proof. Suppose that $\boldsymbol{x} \in \bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$. We claim that $\mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*$. To see why, let $\boldsymbol{a}_i$ be the i^{th} row of the matrix $\mathbf{A}$. We know that for all datasets $D \subseteq \mathcal{X}$, $\boldsymbol{a}_i \cdot \boldsymbol{x} \leq b(D)_i$. By definition of the infimum, this means that $\boldsymbol{a}_i \cdot \boldsymbol{x} \leq \inf_{D \subseteq \mathcal{X}} b(D)_i = b_i^*$. Therefore, $\mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*$.

Next, suppose $\mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*$. Then $\mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)$ for every database D, which means that $\boldsymbol{x} \in \bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$. We conclude that $\bigcap_{D \subseteq \mathcal{X}} \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\} = \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}^*\}$. □

Theorem 3.7. *Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ be an arbitrary diagonal matrix with positive diagonal entries and let $g : \mathbb{R}^m \to \mathbb{R}$ be the function $g(\boldsymbol{x}) = \langle 1, \boldsymbol{x} \rangle$. For any $\Delta > 0$, there exists a mapping from databases $D \subseteq \mathcal{X}$ to vectors $\boldsymbol{b}(D) \in \mathbb{R}^m$ such that:*

1. *The sensitivity of $\boldsymbol{b}(D)$ equals Δ, and*

2. *For any $\epsilon > 0$ and $\delta \in (0, 1/2]$, if $\boldsymbol{\mu}$ is an (ϵ, δ)-differentially private mechanism such that $\mathbf{A}\boldsymbol{\mu}(D) \leq \boldsymbol{b}(D)$ with probability 1, then*

$$\begin{aligned} &\max\left\{g(\boldsymbol{x}) : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\right\} - \mathbb{E}[g(\boldsymbol{\mu}(D))] \\ &\geq \frac{\Delta}{4\epsilon} \cdot \inf_{p \geq 1}\left\{\alpha_{p,1}(\mathbf{A})\sqrt[p]{m}\right\} \cdot \ln\left(\frac{e^{\epsilon} - 1}{2\delta} + 1\right). \end{aligned}$$

Proof. For ease of notation, let $t = \frac{1}{\epsilon}\ln\left(\frac{e^{\epsilon}-1}{2\delta} + 1\right)$. Notice that $\delta \leq \frac{1}{2}$ implies $t \geq 1$. For each vector $\boldsymbol{d} \in \mathbb{Z}^m$, let $D_{\boldsymbol{d}}$ be a database where for any $\boldsymbol{d}, \boldsymbol{d}' \in \mathbb{Z}^m$, if $\left\|\boldsymbol{d} - \boldsymbol{d}'\right\|_1 \leq 1$, then $D_{\boldsymbol{d}}$ and $D_{\boldsymbol{d}'}$ are neighboring. Let $\boldsymbol{b}\left(D_{\boldsymbol{d}}\right) = \Delta\boldsymbol{d}$ and let $a_1, \dots, a_m > 0$ be the diagonal entries of $\mathbf{A}$. Since $\mathbf{A}\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \leq \boldsymbol{b}\left(D_{\boldsymbol{d}}\right)$ with probability 1, $\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)$ must be coordinate-wise smaller than $\Delta\left(\frac{d_1}{a_1}, \dots, \frac{d_m}{a_m}\right)$.

We begin by partitioning the support of $\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)$ so that we can analyze $\mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right]$ using the law of total expectation. We organize this partition using axis-aligned rectangles. Specifically, for each index $i \in [m]$, let S_i^0 be the set of vectors $\boldsymbol{x} \in \mathbb{R}^m$ whose i^{th} components are smaller than $\frac{\Delta}{a_i}\left(d_i - \lfloor t \rfloor\right)$:

$$S_i^0 = \left\{\boldsymbol{x} \in \mathbb{R}^m : x_i \leq \frac{\Delta}{a_i}\left(d_i - \lfloor t \rfloor\right)\right\}.$$

Similarly, let

$$S_i^1 = \left\{\boldsymbol{x} \in \mathbb{R}^m : \frac{\Delta}{a_i}\left(d_i - \lfloor t \rfloor\right) < x_i \leq \frac{\Delta d_i}{a_i}\right\}.$$

For any vector $\boldsymbol{I} \in \{0,1\}^m$, let $S_{\boldsymbol{I}} = \cap_{i=1}^m S_i^{I_i}$. The sets $S_{\boldsymbol{I}}$ partition the support of $\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)$ into rectangles. Therefore, by the law of total expectation,

$$\mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right] = \sum_{\boldsymbol{I} \in \{0,1\}^m} \mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right) \mid \boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right] \mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right]. \tag{10}$$

Conditioning the vector $\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)$ to lie within a rectangle $S_{\boldsymbol{I}}$ makes it much easier to analyze the expected value of $g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)$. Suppose that $\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}$ for some $\boldsymbol{I} \in \{0,1\}^m$. If $I_i = 0$, then we know that $\mu\left(D_{\boldsymbol{d}}\right)_i \leq \frac{\Delta}{a_i}\left(d_i - \lfloor t \rfloor\right)$. Meanwhile, if $I_i = 1$, then $\mu\left(D_{\boldsymbol{d}}\right)_i \leq \frac{\Delta d_i}{a_i}$ since $\mathbf{A}\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \leq \boldsymbol{b}\left(D_{\boldsymbol{d}}\right)$ with probability 1. Since $g(\boldsymbol{x}) = \langle \mathbf{1}, \boldsymbol{x} \rangle$, we have that for each $\boldsymbol{I} \in \{0,1\}^m$,

$$E\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right) \mid \boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right] \leq \sum_{i=1}^m \frac{\Delta\left(d_i - \lfloor t \rfloor\right)}{a_i}\mathbf{1}_{\{I_i = 0\}} + \frac{\Delta d_i}{a_i}\mathbf{1}_{\{I_i = 1\}} = \sum_{i=1}^m \frac{\Delta d_i}{a_i} - \frac{\Delta \lfloor t \rfloor}{a_i}\mathbf{1}_{\{I_i = 0\}}.$$

Combining this inequality with Equation (10) and rearranging terms, we have that

$$\begin{aligned} \mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right] &\leq \Delta\sum_{i=1}^m \frac{d_i}{a_i} - \sum_{\boldsymbol{I} \in \{0,1\}^m}\sum_{i=1}^m \frac{\Delta\lfloor t \rfloor}{a_i}\mathbf{1}_{\{I_i = 0\}}\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right] \\ &= \Delta\sum_{i=1}^m \frac{d_i}{a_i} - \Delta\lfloor t \rfloor\sum_{i=1}^m \frac{1}{a_i}\sum_{\boldsymbol{I} \in \{0,1\}^m}\mathbf{1}_{\{I_i = 0\}}\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right]. \end{aligned}$$

For any $i \in [m]$, $\sum_{\boldsymbol{I} \in \{0,1\}^m}\mathbf{1}_{\{I_i = 0\}}\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_{\boldsymbol{I}}\right] = \mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_i^0\right]$. Therefore,

$$\mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right] \leq \Delta\sum_{i=1}^m \frac{d_i}{a_i} - \Delta\lfloor t \rfloor\sum_{i=1}^m \frac{1}{a_i}\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_i^0\right]. \tag{11}$$

We now prove that for every index $i \in [m]$, $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_i^0\right] > \frac{1}{2}$. This proof relies on the following claim.

Claim B.6. *For any index $i \in [m]$, vector $\bar{\boldsymbol{d}} \in \mathbb{Z}^m$, and integer $j \geq 1$, let $S_{\bar{\boldsymbol{d}},i,j}$ be the set of all vectors $\boldsymbol{x} \in \mathbb{R}^m$ whose i^{th} component is in the interval $\left(\frac{\Delta(\bar{d}_i - j)}{a_i}, \frac{\Delta\bar{d}_i}{a_i}\right]$:*

$$S_{\bar{\boldsymbol{d}},i,j} = \left\{\boldsymbol{x} \in \mathbb{R}^m : \frac{\Delta(\bar{d}_i - j)}{a_i} < x_i \leq \frac{\Delta\bar{d}_i}{a_i}\right\}.$$

Then $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j}\right] \leq \delta\sum_{\ell=0}^{j-1} e^{\epsilon\ell}$.

Notice that $S_{\boldsymbol{d},i,\lfloor t\rfloor} = S_i^1$, a fact that will allow us to prove that $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_i^0\right] > \frac{1}{2}$.

Proof of Claim B.6. We prove this claim by induction on j.

Base case ($j = 1$)**.** Fix an arbitrary index $i \in [m]$ and vector $\bar{\boldsymbol{d}} \in \mathbb{Z}^m$. Let $\boldsymbol{e}_i \in \{0,1\}^m$ be the standard basis vector with a 1 in the i^{th} component and 0 in every other component. Since $\boldsymbol{b}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) = \Delta\left(\bar{\boldsymbol{d}} - \boldsymbol{e}_i\right)$, we know the probability that $\mu\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right)_i > \frac{\Delta(\bar{d}_i - 1)}{a_i}$ is zero. In other words,

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}},i,1}\right] = 0. \tag{12}$$

Since $D_{\bar{\boldsymbol{d}}}$ and $D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}$ are neighboring, this means that

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,1}\right] \leq e^{\epsilon}\,\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}},i,1}\right] + \delta = \delta.$$

Inductive step. Fix an arbitrary $j \geq 1$ and suppose that for all indices $i \in [m]$ and vectors $\bar{\boldsymbol{d}} \in \mathbb{Z}^m$, $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j}\right] \leq \delta\sum_{\ell=0}^{j-1} e^{\epsilon\ell}$. We want to prove that for all indices $i \in [m]$ and vectors $\bar{\boldsymbol{d}} \in \mathbb{Z}^m$, $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] \leq \delta\sum_{\ell=0}^{j} e^{\epsilon\ell}$. To this end, fix an arbitrary index $i \in [m]$ and vector $\bar{\boldsymbol{d}} \in \mathbb{Z}^m$. By the inductive hypothesis, we know that

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i,i,j}\right] \leq \delta\sum_{\ell=0}^{j-1} e^{\epsilon\ell}. \tag{13}$$

Note that

$$\begin{aligned}
S_{\bar{\boldsymbol{d}},i,j+1} &= \left\{\boldsymbol{x} : \frac{\Delta\left(\bar{d}_i - j - 1\right)}{a_i} < x_i \leq \frac{\Delta\bar{d}_i}{a_i}\right\}\\
&= \left\{\boldsymbol{x} : \frac{\Delta\left(\bar{d}_i - j - 1\right)}{a_i} < x_i \leq \frac{\Delta(\bar{d}_i - 1)}{a_i}\right\} \cup \left\{\boldsymbol{x} : \frac{\Delta\left(\bar{d}_i - 1\right)}{a_i} < x_i \leq \frac{\Delta\bar{d}_i}{a_i}\right\}\\
&= S_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i,i,j} \cup S_{\bar{\boldsymbol{d}},i,1}.
\end{aligned}$$

We can now use this fact, the definition of differential privacy, and Equation (13), to prove the inductive hypothesis holds. By the definition of differential privacy,

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] \leq e^{\epsilon}\,\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] + \delta.$$

Since $S_{\bar{\boldsymbol{d}},i,j+1} = S_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i,i,j} \cup S_{\bar{\boldsymbol{d}},i,1}$, we have that

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] \leq e^{\epsilon}\left(\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i,i,j}\right] + \mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}},i,1}\right]\right) + \delta.$$

By Equation (12), we know that $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}},i,1}\right] = 0$, so

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] \leq e^{\epsilon}\,\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i}\right) \in S_{\bar{\boldsymbol{d}}-\boldsymbol{e}_i,i,j}\right] + \delta.$$

Finally, by the inductive hypothesis (Equation (13)),

$$\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\bar{\boldsymbol{d}}}\right) \in S_{\bar{\boldsymbol{d}},i,j+1}\right] \leq e^{\epsilon}\delta\sum_{\ell=0}^{j-1} e^{\epsilon\ell} + \delta = \delta\sum_{\ell=0}^{j} e^{\epsilon\ell},$$

so the inductive hypothesis holds. □

Since $S_{\boldsymbol{d},i,\lfloor t\rfloor} = S_i^1$, our careful choice of the value t allows us to prove that for every index i, $\mathbb{P}\left[\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right) \in S_i^0\right] > \frac{1}{2}$ (see Claim B.7 in Appendix B). This inequality, Equation (11), and the fact that $t \geq 1$ together imply that

$$\mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right] < \Delta\sum_{i=1}^{m}\frac{d_i}{a_i} - \frac{\Delta\lfloor t\rfloor}{2}\sum_{i=1}^{m}\frac{1}{a_i} \leq \Delta\sum_{i=1}^{m}\frac{d_i}{a_i} - \frac{\Delta t}{4}\sum_{i=1}^{m}\frac{1}{a_i}.$$

Since $\max\left\{g(\boldsymbol{x}) : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}\left(D_{\boldsymbol{d}}\right)\right\} = \max\left\{\langle \mathbf{1}, \boldsymbol{x}\rangle : \boldsymbol{x} \leq \Delta\left(\frac{d_1}{a_1}, \dots, \frac{d_m}{a_m}\right)\right\} = \Delta \sum_{i=1}^m \frac{d_i}{a_i}$, we have that

$$\max\left\{g(\boldsymbol{x}) : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}\left(D_{\boldsymbol{d}}\right)\right\} - \mathbb{E}\left[g\left(\boldsymbol{\mu}\left(D_{\boldsymbol{d}}\right)\right)\right] \geq \frac{\Delta}{4\epsilon}\left(\sum_{i=1}^m \frac{1}{a_i}\right)\ln\left(\frac{e^\epsilon - 1}{2\delta} + 1\right).$$

We now prove that $\inf_{p\geq 1} \alpha_{p,1}(\mathbf{A})\sqrt[p]{m} \leq \alpha_{\infty,1}(\mathbf{A}) = \sum_{i=1}^m \frac{1}{a_i}$, which proves the theorem statement. Since $\mathbf{A}$ is diagonal, $\alpha_{\infty,1}(\mathbf{A}) = \sup_{\boldsymbol{u}\geq \mathbf{0}}\left\{\|\boldsymbol{u}\|_1 : u_i a_i \leq 1, \forall i \in [m]\right\} = \sum_{i=1}^m \frac{1}{a_i}$. Moreover, since $\sqrt[\infty]{m} = 1$, $\alpha_{\infty,1}(\mathbf{A}) \in \left\{\alpha_{p,1}(\mathbf{A})\sqrt[p]{m} : p \geq 1\right\}$, which implies that $\inf_{p\geq 1}\alpha_{p,1}(\mathbf{A})\sqrt[p]{m} \leq \alpha_{\infty,1}(\mathbf{A})$. Therefore,

$$\max\left\{g(\boldsymbol{x}) : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D_i)\right\} - \mathbb{E}[g(\boldsymbol{\mu}(D))] \geq \frac{\Delta}{4\epsilon} \cdot \inf_{p \geq 1}\left\{\alpha_{p,1}(\mathbf{A})\sqrt[p]{m}\right\} \cdot \ln\left(\frac{e^\epsilon - 1}{2\delta} + 1\right),$$

as claimed. □

Claim B.7. *Let $D_{\boldsymbol{d}}$ and S_i^0 be defined as in Theorem 3.7. Then $\mathbb{P}\left[\boldsymbol{\mu}(D_{\boldsymbol{d}}) \in S_i^0\right] > \frac{1}{2}$.*

Proof. Let S_i^1 and $S_{\boldsymbol{d},i,\lfloor t\rfloor}$ be defined as in Theorem 3.7. We prove that $\mathbb{P}\left[\mu(D_{\boldsymbol{d}}) \in S_{\boldsymbol{d},i,\lfloor t\rfloor}\right] \leq \frac{1}{2}$. Since $S_{\boldsymbol{d},i,\lfloor t\rfloor} = S_i^1$, this implies that $\mathbb{P}\left[\mu(D_{\boldsymbol{d}}) \in S_i^0\right] > \frac{1}{2}$. This claim follows from the following chain of inequalities, which themselves follow from Claim B.6 and the fact that $t = \frac{1}{\epsilon}\log\left(\frac{e^\epsilon - 1}{2\delta} + 1\right)$:

$$\begin{aligned}
\mathbb{P}\left[\mu(D_{\boldsymbol{d}}) \in S_{\boldsymbol{d},i,\lfloor t\rfloor}\right] &\leq \delta \sum_{j=0}^{\lfloor t\rfloor - 1} e^{\epsilon j}\\
&= \frac{\delta\left(e^{\epsilon\lfloor t\rfloor} - 1\right)}{e^\epsilon - 1}\\
&\leq \frac{\delta\left(\exp\left(\epsilon t\right) - 1\right)}{\exp(\epsilon) - 1}\\
&= \frac{\delta\left(\exp\left(\epsilon\left(\frac{1}{\epsilon}\log\left(\frac{e^\epsilon - 1}{2\delta} + 1\right)\right)\right) - 1\right)}{\exp(\epsilon) - 1}\\
&= \frac{\delta\left(\exp\left(\log\left(\frac{e^\epsilon - 1}{2\delta} + 1\right)\right) - 1\right)}{\exp(\epsilon) - 1}\\
&= \frac{\delta\left(\left(\frac{e^\epsilon - 1}{2\delta} + 1\right) - 1\right)}{\exp(\epsilon) - 1}\\
&= \frac{\delta\left(\frac{e^\epsilon - 1}{2\delta}\right)}{\exp(\epsilon) - 1}\\
&= \frac{1}{2}.
\end{aligned}$$

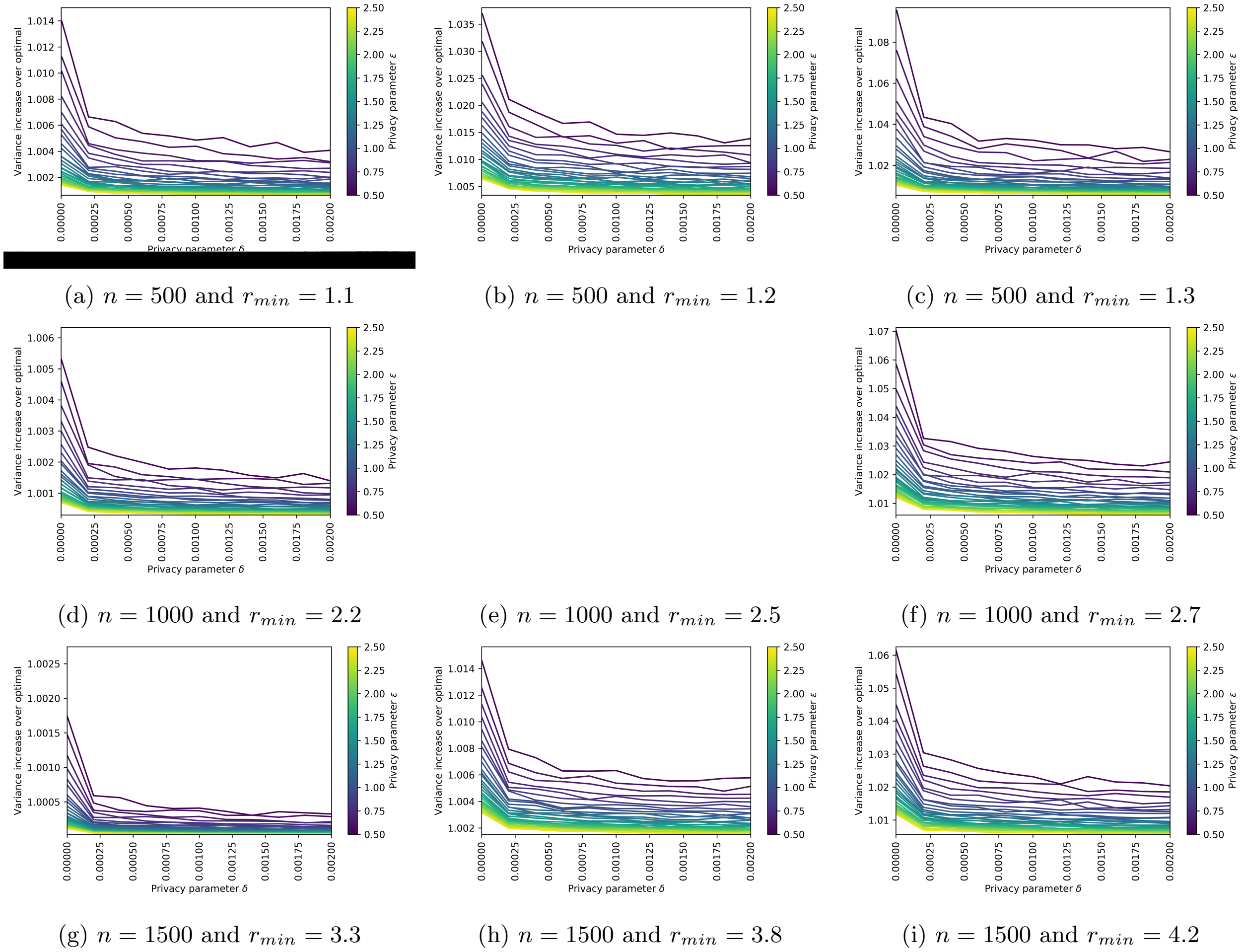
Therefore, $\mathbb{P}\left[\boldsymbol{\mu}(D_{\boldsymbol{d}}) \in S_i^0\right] > \frac{1}{2}$. □

C Characterizing $(\epsilon, 0)$-differentially private mechanisms

In Section 3 we presented a nearly optimal (ϵ, δ)-DP mechanism for private optimization. The optimal $(\epsilon, 0)$-DP mechanism for this problem is considerably easier to characterize.

Theorem C.1. *Let $S^* = \bigcap_{D\subseteq\mathcal{X}}\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}(D)\}$ be the intersection of all feasible sets across all databases D. If S^* is nonempty, then the optimal $(\epsilon, 0)$-differentially private mechanism outputs* $\operatorname{argmax}_{\boldsymbol{x}\in S^*} g(\boldsymbol{x})$ *with probability 1. If S^* is empty, then no $(\epsilon, 0)$-differentially private mechanism exists.*

Proof. Fix a mechanism, and let $P(D)$ be the set of vectors $\boldsymbol{x}$ in the support of the mechanism's output given as input the database D. We claim that if the mechanism is $(\epsilon, 0)$-differentially private, then there exists a set P^* such that $P(D) = P^*$ for all databases D. Suppose, for the sake of a contradiction, that there exist databases D and D' such that $P(D) \neq P(D')$. Let $D_1, \dots, D_n$ be a sequence of databases such that $D_1 = D$, $D_n = D'$,



[redacted]

(a) $n = 500$ and $r_{min} = 1.1$

(b) $n = 500$ and $r_{min} = 1.2$

(c) $n = 500$ and $r_{min} = 1.3$

(d) $n = 1000$ and $r_{min} = 2.2$

(e) $n = 1000$ and $r_{min} = 2.5$

(f) $n = 1000$ and $r_{min} = 2.7$

(g) $n = 1500$ and $r_{min} = 3.3$

(h) $n = 1500$ and $r_{min} = 3.8$

(i) $n = 1500$ and $r_{min} = 4.2$

Figure 5: Quality in the portfolio optimization application for various choices of the number n of investors and the minimum return r_{min}. These plots show the multiplicative increase in the objective function value of our algorithm's solution—for various choices of ϵ and δ—over the objective function value of the optimal solution to the original optimization problem (Equation (8)). Darker shading corresponds to lower values of ϵ and therefore stronger privacy.

and each pair of databases D_i and D_{i+1} are neighbors. Then there must exist a pair of neighboring databases D_i and D_{i+1} such that $P(D_i) \neq P(D_{i+1})$, which contradicts the fact that the mechanism is $(\epsilon, 0)$-differentially private. Therefore, if the mechanism is $(\epsilon, 0)$-differentially private, then to satisfy the feasibility requirement, we must have that $P^* \subseteq S^*$. If S^* is empty, then no such mechanism exists. If S^* is nonempty, then the optimal $(\epsilon, 0)$-differentially private mechanism outputs $\operatorname{argmax}_{\boldsymbol{x} \in S^*} g(\boldsymbol{x})$ with probability 1. □

D Additional information about experiments

In Figure 5, we analyze the quality of our algorithm for several different parameter settings. First, we select the number of individuals n to be a value in $\{500, 1000, 1500\}$ (the number of investors is $n = 500$ in Figures 5a-5c, $n = 1000$ in Figures 5d-5f, and $n = 1500$ in Figures 5g-5i). Then, we define each element of the database (money given by individuals to an investor) as a draw from the uniform distribution between 0 and 1, so $b(D)$ equals the sum of these n random variables. The sensitivity of $b(D)$ is therefore $\Delta = 1$. We set the minimum return r_{min} to be a value in $[1, 5]$. We calculate the objective value $v^* \in \mathbb{R}$ of the optimal solution to Equation (8). Then, for $\delta \in \left[\frac{1}{n^2}, 0.002\right]$ and $\epsilon \in [0.5, 2.5]$, we run our algorithm 50 times and calculate the average objective value $\hat{v}_{\epsilon,\delta} \in \mathbb{R}$ of the optimal solutions.

We find that there is a sweet spot for the parameter choices n and r_{min}. If r_{min} is too small, the budget constraint is non-binding with or without privacy, so the variance increase over optimal is always 1. We find this is true when $n = 500$ and $r_{min} \leq 1$, when $n = 1000$ and $r_{min} \leq 2.1$, and when $n = 1500$ and $r_{min} \leq 3.2$. This also explains why the variance increase over optimal improves as r_{min} shrinks, as we can observe from Figure 5. Meanwhile, if r_{min} is too large, then the original quadratic program (Equation (8)) is infeasible. We find this is true when $n = 500$ and $r_{min} \geq 1.4$, when $n = 1000$ and $r_{min} \geq 2.8$, and when $n = 1500$ and $r_{min} \geq 4.3$.