Differentially Private Linear Programming with Guaranteed Constraint Satisfaction

Alexander Benvenuti*, Brendan Bialy ‡, Miriam Dennis ‡, Matthew Hale*
*Georgia Institute of Technology, ‡Air Force Research Lab

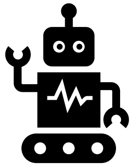
Linear programming is used across disciplines

- Linear programs (LPs) are commonly used in
  - Finance: optimizing portfolios
  - Marketing: pricing advertisements
  - Logistics: building travel itineraries
  - Autonomy: synthesizing policies/controllers
- Typically formulated using
  - Finance: budgets, company valuations
  - Marketing: website traffic, ad effectiveness
  - Logistics: travel costs, destinations
  - Autonomy: system limitations, environment information, mission specs

D

$$\underset{x \in \mathbb{R}^n}{\text{maximize}}\; c^T x$$
$$\text{Subject to } Ax \leq b$$



$$\underset{x \in \mathbb{R}^n}{\text{maximize}}\; c(D)^T x$$
$$\text{Subject to } A(D)x \leq b(D)$$

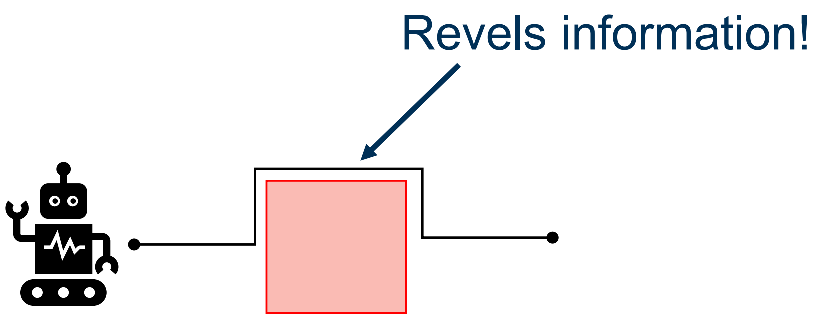
Issue: This information is very sensitive!

Privacy is required to protect LPs

- Solutions of LPs can reveal information about the data used to formulate them
- Hsu et al. [1] attempted to privately solve LPs
  - This work allows for constraint violations

In autonomy, this means systems may crash, operate unsafely, and not meet mission objectives

- Privately solving LPs with constraint satisfaction is an open problem [2]



<u>Munoz</u>

$$\underset{x\geq 0}{\text{maximize}}\; c^T x$$

$$\text{Subject to } Ax \leq b(D)$$

<u>Us</u>

$$\underset{x\geq 0}{\text{maximize}}\; c(D)^T x$$

$$\text{Subject to } A(D)x \leq b(D)$$

In this talk: Solve

$$\underset{x\geq 0}{\text{maximize}}\; c(D)^T x$$

$$\text{Subject to } A(D)x \leq b(D)$$

in a differentially private manner while guaranteeing feasibility in the original constraints

[1] Hsu, J., Roth, A., Roughgarden, T., and Ullman, J. Privately solving linear programs. In Automata, Languages, and Programming: 41st International Colloquium, pp.612–624. Springer, 2014b.
[2] Munoz, A., Syed, U., Vassilvtiskii, S., and Vitercik, E. Private optimization without constraint violations. In International Conference on Artificial Intelligence and Statistics, pp. 2557–2565. PMLR, 2021.

We use differential privacy to formulate private LPs

- Differential privacy goal: Make "similar" data appear "approximately indistinguishable", enforced by a mechanism M
- Similar is defined by Adjacency

Definition (Adjacency): Two databases D, D' are adjacent if they differ in at most one entry

- To be approximately indistinguishable

Definition (Differential Privacy): A mechanism M is (ϵ, δ)-differentially private if

$$\mathbb{P}(M(D) \in S) \leq e^{\epsilon}\mathbb{P}(M(D') \in S) + \delta$$

- Small ϵ, δ = strong privacy,
- Usually, $0.1 \leq \epsilon \leq 10, \quad 0 \leq \delta \leq 0.05$

What guarantees does this give us?

Differential privacy provides useful guarantees

Definition (Differential Privacy): A mechanism M is (ϵ, δ)-differentially private if

$$\mathbb{P}(M(D) \in S) \leq e^{\epsilon}\mathbb{P}(M(D') \in S) + \delta$$

Properties of Differential Privacy:

- Immunity to post-processing
- Robustness to side information
- Compositions remain differentially private

We want these guarantees for D

How do we make a differential privacy mechanism?

Definition (Sensitivity): Given adjacent databases D, D' the sensitivity of a function $f: \mathcal{D} \to \mathbb{R}^{m \times n}$ is

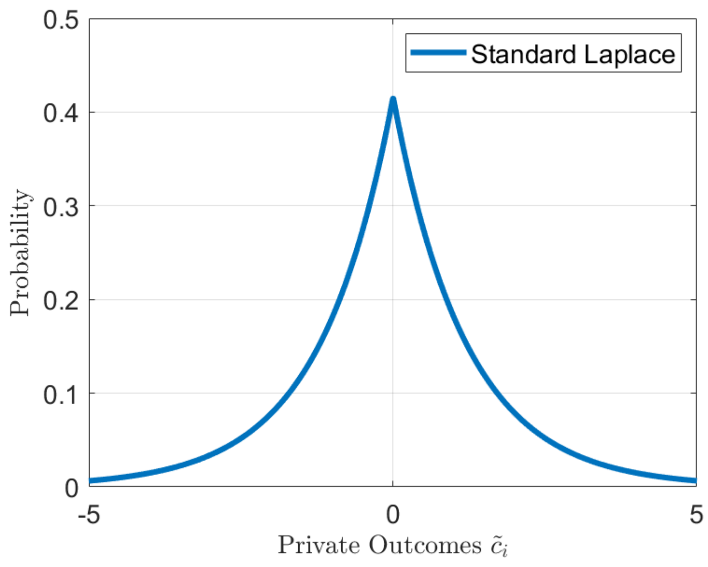
$$\Delta_{1,1} f = \sup_{D, D'} ||f(D) - f(D')||_{1,1}$$

- "The most f can change on adjacent D, D'"

We can add calibrated noise using the sensitivity to attain differential privacy

We privatize each component of an LP

- Fix an LP with components $A(D), b(D), c(D)$
- Fix $\epsilon > 0, \delta \in \left(0, \frac{1}{2}\right)$
- Set of all possible database realizations $\mathcal{D}$



Mechanism for $c(D)$:

- Generate Laplace noise $z_c \sim \mathcal{L}(\sigma_c), \sigma_c \leq \frac{\Delta_1 c}{\epsilon}$
- $\tilde{c} = c(D) + z_c$
- $\tilde{c}$ is $(\epsilon, 0)$- differentially private

We can control how much information is leaked if the private cost function is learned

The cost was easy, what about constraints and feasibility?

We privatize each component of an LP

- Fix an LP with components $A(D), b(D), c(D)$
- Fix $\epsilon > 0, \delta \in \left(0, \frac{1}{2}\right)$
- Set of all possible database realizations $\mathcal{D}$

Mechanism for $b(D)$:

- Compute bounds on Laplace noise

$$s_b = \frac{\Delta_1 b}{\epsilon} \log\left(\frac{m(e^{\epsilon}-1)}{\delta} + 1\right), S_b = [-s_b, s_b]$$

- Generate bounded noise

$$z_{b_i} \sim \mathcal{L}_T(\sigma_b, S_b), \sigma_c \leq \frac{\Delta_1 b}{\epsilon}$$

- $\bar{b}_i = b(D)_i - s_b + z_{b_i}$
- $\tilde{b}_i = \max\{\bar{b}_i, \inf_{d \in \mathcal{D}} b(d)_i\}$

Mechanism for $A(D)$:

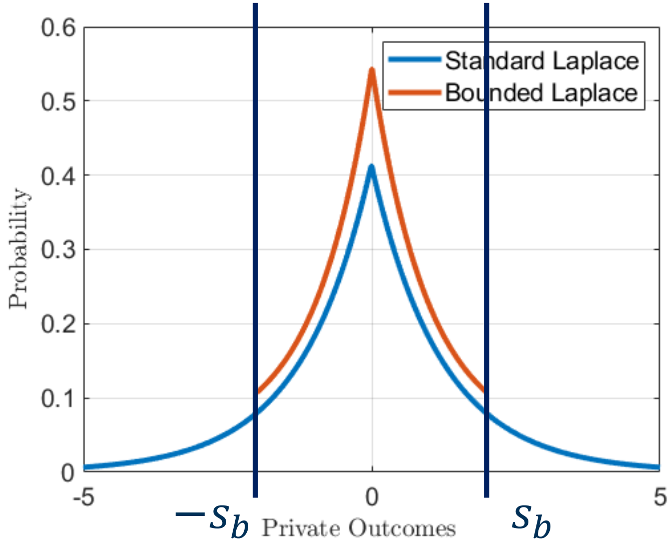
- Compute bounds on Laplace noise

$$s_A = \frac{\Delta_{1,1} A}{\epsilon} \log\left(\frac{m(e^{\epsilon}-1)}{\delta} + 1\right), S_A = [-s_A, s_A]$$

- Generate bounded noise

$$z_{A_{i,j}} \sim \mathcal{L}_T(\sigma_A, S_A), \sigma_c \leq \frac{\Delta_{1,1} A}{\epsilon}$$

- $\bar{A}_{i,j} = A(D)_{i,j} + s_A + z_{A_{i,j}}$
- $\tilde{A}_{i,j} = \min\{\bar{A}_{i,j}, \sup_{d \in \mathcal{D}} A(d)_{i,j}\}$



We can keep constraints private by only making them tighter, ensuring feasibility

Our mechanisms enforce differential privacy

Theorem 1:

- Fix $\epsilon > 0, \delta \in \left(0, \frac{1}{2}\right)$
- Forming and solving

$$\underset{x \geq 0}{\text{maximize}}\ \tilde{c}^T x$$
$$\text{Subject to } \tilde{A}x \leq \tilde{b}$$

is (ϵ, δ)-differentially private

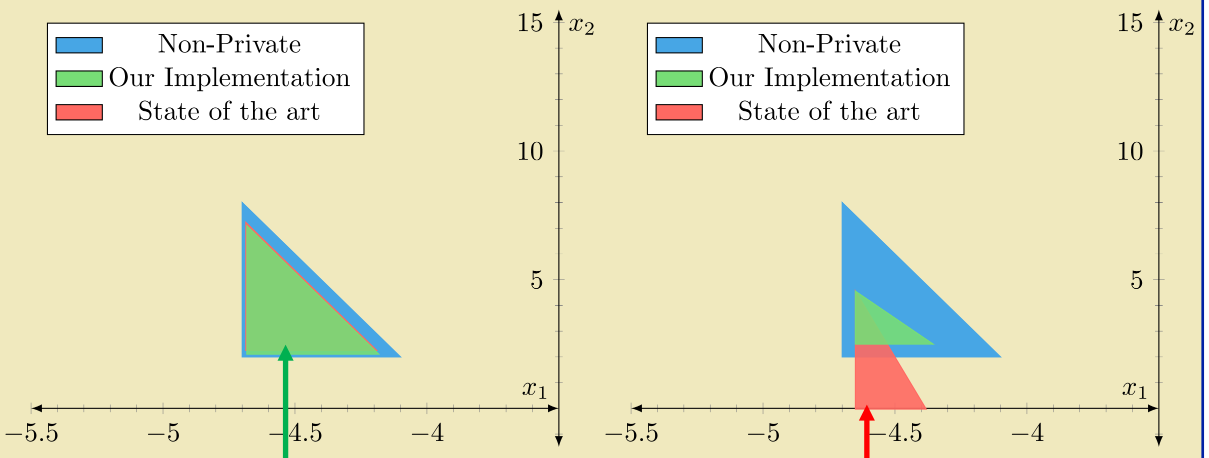
Our mechanism produces a solution which is always feasible in the original constraints

How well does that solution perform?

Example:

$$\underset{x \in \mathbb{R}^2}{\text{maximize}}\ 2x_1 - 3x_2$$
$$\text{Subject to } \quad x_1 + x_2 \leq 12,$$
$$x_1, x_2 \geq 3$$

$\epsilon = 1$ $\qquad$ $\epsilon = 0.1$



Solutions inside the blue: no crashing, completes mission, no unsafe behavior

Solutions outside the blue: a controller which lead to mission failure

We analyze the quality of solutions

$$\underset{x \geq 0}{\text{maximize}}\; c(D)^T x$$
$$\text{Subject to } A(D)x \leq b(D)$$

Our Privacy Implementation

$$\underset{x \geq 0}{\text{maximize}}\; \tilde{c}^T x$$
$$\text{Subject to } \tilde{A}x \leq \tilde{b}$$

No privacy

$$c(D)^T x^*$$

How far apart are these on average?

Has privacy

$$c(D)^T \tilde{x}^*$$

Theorem 2:

- Fix $\epsilon > 0, \delta \in \left(0, \frac{1}{2}\right)$
- Fix an LP with components $A(D), b(D), c(D)$ with $\Delta_1 c, \Delta_1 b, \Delta_{1,1} A$

$$\mathbb{E}[c(D)^T x^* - c(D)^T \tilde{x}^*] \leq \underbrace{||c(D)||_2 H(A, b, c)}_{\text{Problem parameters}} \underbrace{\rho(\epsilon, \delta, \Delta_1 c, \Delta_1 b, \Delta_{1,1} A)}_{\text{Privacy Implementation}}$$

- $H(A, b, c)$: Hoffman constant of the LP, always exist/can be efficiently approximated

Georgia Tech

We empirically trade off privacy and performance

- Consider the following optimization problem

$$\underset{x\geq 0}{\text{maximize}} \sum_{i\in[N]} \sum_{j\in[M]} p(D)_{ij} x_{ij}$$

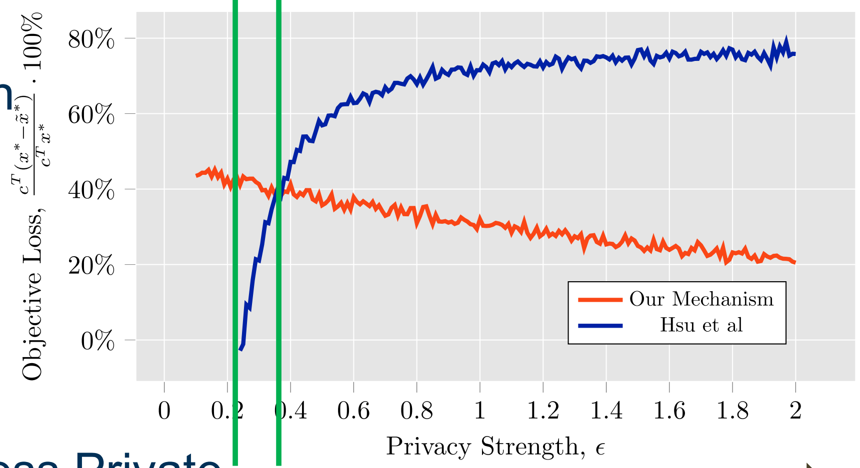
$$\text{Subject to: } \sum_{j\in[M]} x_{ij} \leq n_i \text{ for } i \in [N],$$

$$\sum_{i\in[N]} p(D)_{ij} x_{ij} \leq b(D)_j \text{ for } j \in [M]$$

- We consider $p(D)$ and $b(D)$ sensitive

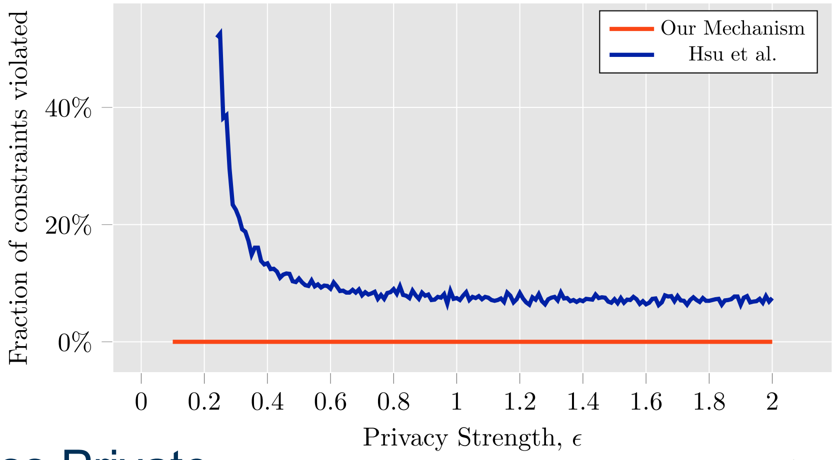
We always satisfy constraints while producing a solution with 65% lower suboptimality than the state of the art

Better Solution



Less Private

Better Solution



Less Private

Georgia Tech

We empirically trade off privacy and problem size

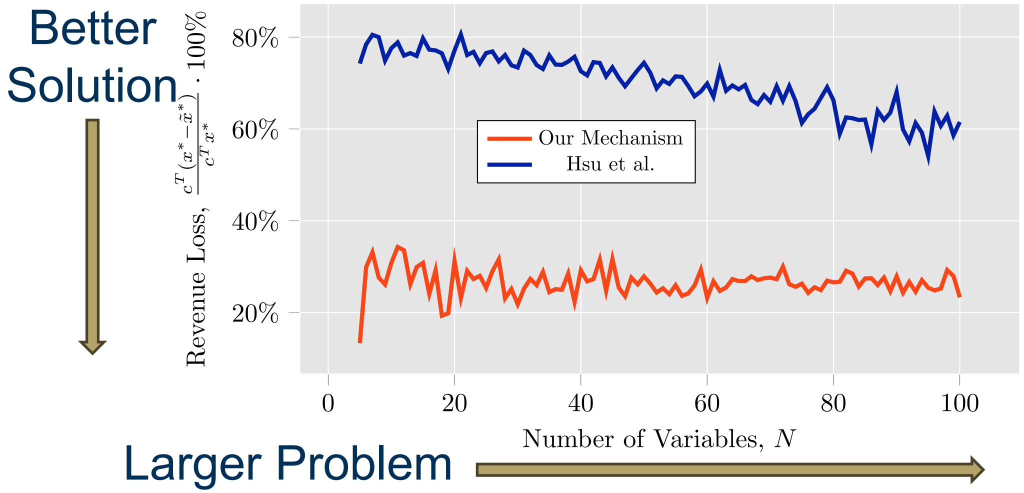
- Consider the following optimization problem

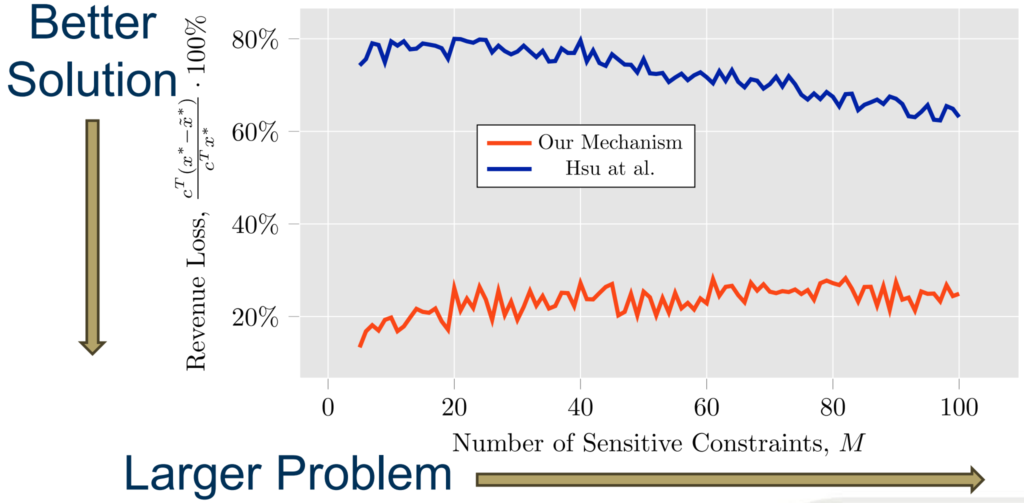
$$\underset{x \geq 0}{\text{maximize}} \sum_{i \in [N]} \sum_{j \in [M]} p(D)_{ij} x_{ij}$$

$$\text{Subject to: } \sum_{j \in [M]} x_{ij} \leq n_i \text{ for } i \in [N],$$

$$\sum_{i \in [N]} p(D)_{ij} x_{ij} \leq b(D)_j \text{ for } j \in [M]$$

- We consider $p(D)$ and $b(D)$ sensitive





Solution quality is unaffected by size

Our mechanism is provably private with strong performance

Takeaways

- Provably conceal mission specs
- Concealment is future-proofed
  - Other methods (i.e., encryption) cannot be
- Maintained good performance
  - Simulation shows strong performance with large systems and strong privacy

Hardware Implementations

- Currently: deploying on ground robots at Georgia Tech's Robotarium platform
- This summer: deploying on drones Eglin AFB's Aviary with AFRL RW

Thank you!
Email: abenvenuti3@gatech.edu