# Differential privacy, entropy and security in distributed control of cyber-physical systems

YU WANG, ZHENQI HUANG, SAYAN MITRA, GEIR DULLERUD

APRIL 26, 2016

# General Question

For distributed control systems, how expensive is it to preserve privacy? How to optimize?

Navigation

- Routing delays vs location privacy

Smart Grid

- Peak demand vs schedule privacy

## *Section I:*
## On Differential Privacy of Distributed Control System

# Distributed control

Consider a network of vehicles evolving in a shared environment (road congestion)

State of each agent (vehicle) $x_i$

- Evolve with coupled dynamics (delays)

Agents want to share state to estimate delays

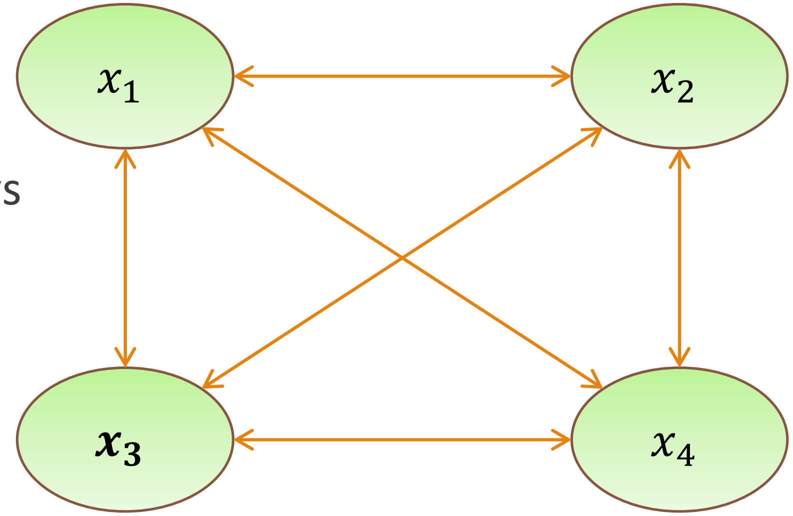
Private preferences $p_i$,

- initial states + sequence of waypoint

Report value $z_i = x_i + noise$

Dynamics of agent:

$$z_i = x_i + w_i$$
$$u_i = g(x_i, p_i, z)$$
$$x_i^+ = f(x_i, x, u_i)$$



# Some notations

$$\boxed{\begin{aligned} z_i &= x_i + w_i \\ u_i &= g(x_i, p_i, z) \\ x_i^+ &= f(x_i, x, u_i) \end{aligned}}$$

- Sensitive data set: $D = \{p_i\}_{i\in[N]}$ collects agent preference
  - Two data set $D, D'$ are adjacent if they differ in one agent's data
- Observation sequence: $O = \{z(t)\}_{t\in[T]} \in \Re^{\{nNT\}}$
- Trajectory: $\xi = \{x(t)\}_{t\in[T]}$,
  - Fully defined by a data set $D$ and observation $O$, $\xi_{D,O}$

# $\epsilon$-differential privacy

Definition: The randomized communication is **ε-**differentially private with $\epsilon > 0$, if for all adjacent datasets $D$ and $D'$ for all subset of observations $S$,

$$\Pr[O_D \in S] \leq e^{\epsilon} \Pr[O_{D'} \in S]$$

- Difference in one agent's data doesn't change the output distribution much
- Small $\epsilon$, high privacy; $\epsilon \to 0$, no communication; $\epsilon \to \infty$, no privacy
- How to design the noise to achieve $\epsilon$-differential privacy?

[Dwork2006], [Ny2014], [Huang2012]

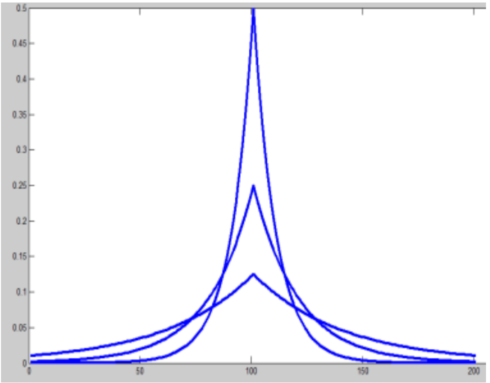
# Laplace mechanism for one-shot queries[Dwork06]

No dynamics involve, just exchanging initial states

- $p_i \in \Re$ is the initial state of agent $i$

Laplace mechanism: $z_i = p_i + Lap\left(\frac{1}{\epsilon}\right)$ gives $\epsilon$-differential privacy for any $\epsilon$

- $Lap\left(\frac{1}{\epsilon}\right)$ has p.d.f. : $f(x) = \frac{\epsilon}{2} e^{\epsilon|x|}$
- $\forall x, x'$: $\frac{f(x)}{f(x')} \leq e^{\epsilon|x-x'|}$
- The average reported value is $\sum z_i$ which gives DP with accuracy bounds



# When dynamics come into the picture

Definition: the sensitivity of the system is supremum 1-norm between agent trajectories

$$S(t) = \sup_{\substack{\mathrm{adj}(D,D') \\ O \in Obs}} |\, \xi_{D,O,i}(t) - \xi_{D',O,i}(t)|_1$$

- Sensitivity is a property of dynamics of the network
- It can be computed [HiCoNS2014], [CAV2014]

# Laplace Mechanism for dynamical systems

Theorem: The following distributed control system is $\epsilon$-differentially private:

- at each time $t$, each agent adds an vector of independent Laplace noise $Lap(\frac{S(t)T}{\epsilon})$ to its actual state:

$$z(t) = x_i(t) + Lap(\frac{S(t)T}{\epsilon})$$

- Larger time horizon, higher privacy level, larger sensitivity $\Rightarrow$ more noise $\Rightarrow$ worse accuracy

# Cost of Privacy

Average Cost: $Cost_p = \frac{1}{\mathrm{N}} \sum_{t=0}^{T} \sum_i |x_i(t) - p_i(t)|^2$

Baseline cost $\overline{Cost_p}$: the cost when $z_i(t) = x_i(t)$

- No noise

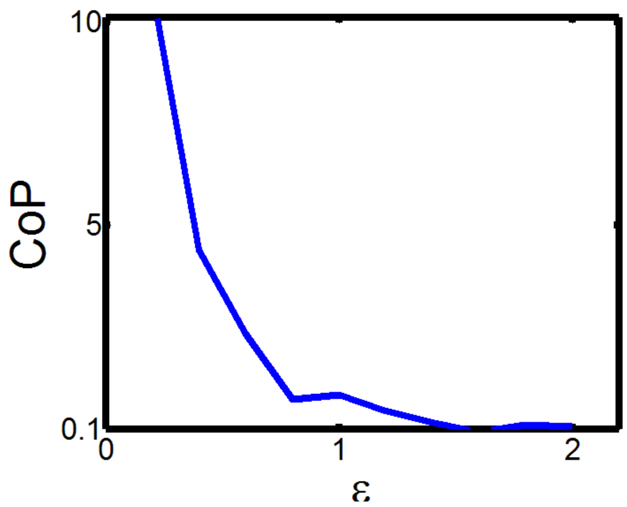
The Cost of Privacy of a DP mechanism $M$ is:

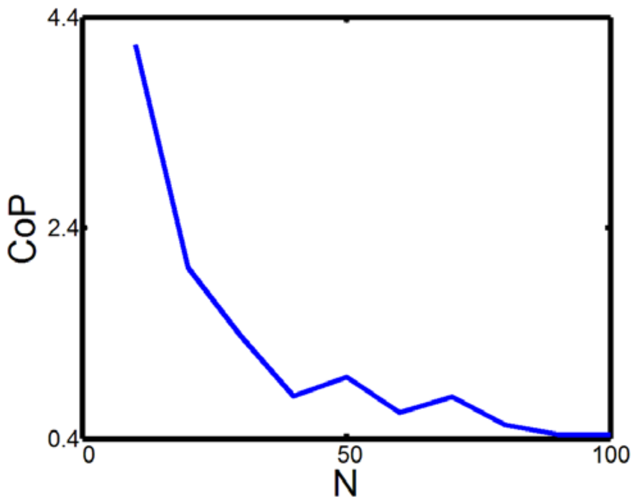
$$\boldsymbol{CoP} = \sup_{p} \mathbf{E}[Cost_p - \overline{Cost_p}]$$

# CoP for linear dynamical system

For stable dynamics: $\mathrm{CoP} \sim O(\frac{T^3}{N^2\epsilon^2})$,

otherwise exponential in $T$



# Summary

Extend the notion of differential privacy to dynamical systems

Generalize Laplace mechanism to dynamical observation using sensitivity of trajectories

For stable dynamics $\mathrm{CoP} \sim O(\frac{T^3}{N^2 \epsilon^2})$, otherwise, exponential in $T$

# *Section II:* Entropy-minimization of Differential Privacy

# Feedback control system

$$\boxed{\begin{aligned} z &= x + w \\ x^+ &= f(x, z) \end{aligned}}$$

- Feedback control of agent:
  - Sensitive data: $x_0$ initial state of agent
    - Protecting the initial state is equivalent to protecting the whole trajectory
  - Observation sequence: $O = \{z(t)\}_{t \in [T]}$
- Question: how much information is lost by adding noise? How to minimize the information loss?

# Estimation & Entropy

Definition. An estimate of the agent's initial state is the expectation of the initial state given the history of the agents' report

$$\widetilde{x_t} = \mathbf{E}[x_0 | z_0, z_1, \dots, z_t]$$

Definition. The entropy of a random variable $x$ with probability distribution function $f(x)$ is defined as

$$H(x) = -\int f(x) \ln x \; dx$$

# Entropy-minimization problem

For minimizing the amount of information loss for achieving differential privacy, we design the noise $w$ to be added :

Minimize $H(\widetilde{x_t})$

Subject to: $\forall a, b$: $\mathrm{P}[\widetilde{x_t} = a] \leq e^{\epsilon|a-b|}\mathrm{P}[\widetilde{x_t} = b]$

# Result for one-shot case

$$z = x + w$$

The estimate $\tilde{x} \in \Re^n$ is computed by the first observation $z \in \Re^n$, no dynamics is involved.

Theorem: The lower-bound of estimate entropy is $n - n \ln \frac{\epsilon}{2}$, which is achieved by adding Laplace noise $w \sim Lap(1/\epsilon)$

# Sketch of proof [CDC14]

- Let $p(x, z)$ be the joint distribution of initial state $x$ and report $z$, we find a symmetric property
- Claim 1: for any $x$, $p(x, z - x)$ is even
  - Since the noise to add is $n = z - x$, the noise is mean-zero
- Claim 2: for any $c$, $p(x, z) = p(2c - x, 2c - z)$
  - The noise added is independent of the state
- We can define $f(w) = f(z - x) = p(x, z)$
- Claim 3: $f(w)$ is non-decreasing

# Extension with dynamics

$$\boxed{\begin{gathered} z = x + w \\ x^+ = f(x, z) \end{gathered}}$$

The estimate $\widetilde{x}_t = \mathrm{E}[x_0 | z_0, z_1, \ldots, z_t]$ is computed by the first $t$ observation $\{z_s\}_{s \in [T]}$

- Theorem: The lower-bound of estimate entropy is still $n - n \ln \frac{\epsilon}{2}$, which is achieved by a Laplace mechanism.

# Optimal Laplace mechanism

$$\boxed{\begin{gathered} z = x + n \\ x^+ = f(x, z) \end{gathered}}$$

- The first noise to add is the same as the one-shot case:

$$w_0 \sim Lap(1/\epsilon)$$

- In the following round $t > 0$, the noise to be added is by evolving the initial noise with the dynamics:

$$w_t = \xi(w_0, t)$$

# Summary

- Formulate a general estimation problem for which we want to minimize the entropy of estimate

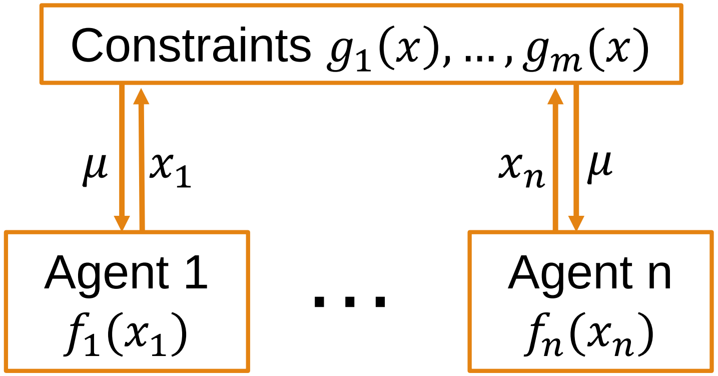
- Prove a lower bound of estimation entropy $n - n \ln \frac{\epsilon}{2}$

- The lower bound is achieved by Laplace mechanism

# *Section III:* Differential Privacy of Distributed Optimization

# Architecture

- Local objective functions
- Global constraints
- Communication via the cloud

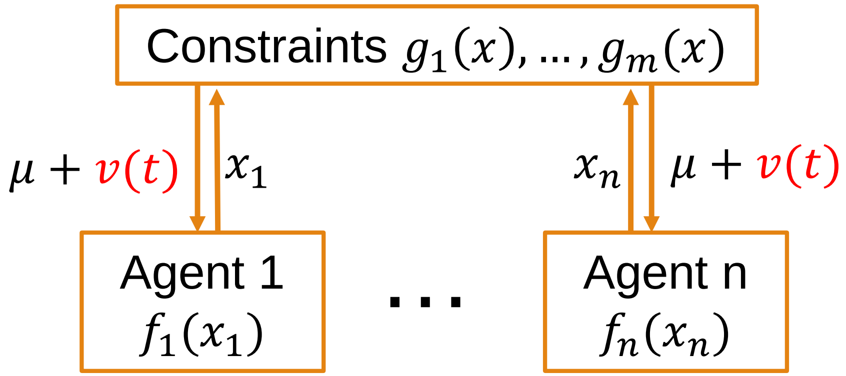



How to keep objective functions differentially private in communication?

# Algorithm

$$x_i \leftarrow \Pi_{X_i}\left[x_i - \gamma_t\left(\frac{\partial f_i}{\partial x_i} + \mu^T\frac{\partial g}{\partial x_i} + \alpha_t x_i\right)\right]$$
$$\mu \leftarrow \Pi_M[\mu + \gamma_t(g(x) - \alpha_t\mu)]$$
$$\mu \leftarrow \mu + v(t)$$

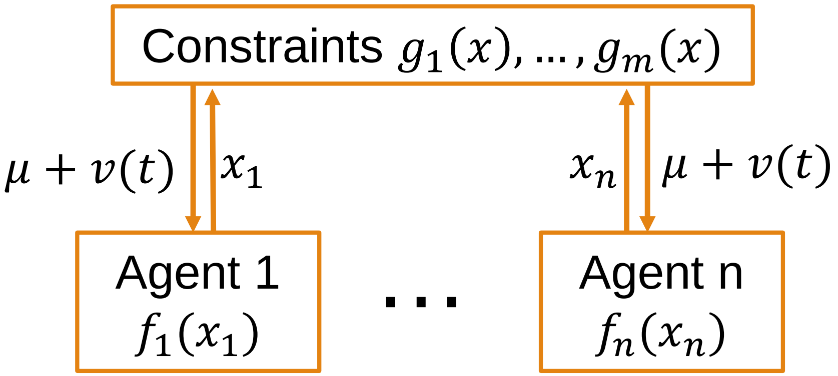
$$\gamma_t = \gamma_1 t^{-c_1}$$
$$\alpha_t = \alpha_1 t^{-c_2}$$
$$c_1 > c_2, c_1 + c_2 < 1$$




*For $v(t) = 0$, the algorithm converges to optima.*

# Assumptions

- Linear objective functions $f_i(x_i) = a_i x_i$
- Lipschitz Constraints $\left\|\frac{\partial g_j}{\partial x_k}\right\| \leq l_{j,k}$
- Completely correlated noise $v(t)$



# Privacy

Two sensitive data $D = \{a_1, \ldots, a_n\}$ and $D' = \{a_1{}', \ldots, a_n{}'\}$ are adjacent if they differ only in the $i$th element. The distance between them is $\|D - D'\| = \|a_i - a_i'\|$.

The algorithm is $\varepsilon$-differentially private if given initial state $x(0), \mu(0)$, the sequence of public multiplier generated by two adjacent sensitive data satisfies

$$Pr\left[\mu_D^{x(0),\mu(0)} \in O\right] \\ \leq e^{\varepsilon\|D-D'\|} Pr\left[\mu_{D'}^{x(0),\mu(0)} \in O\right]$$

# Accuracy

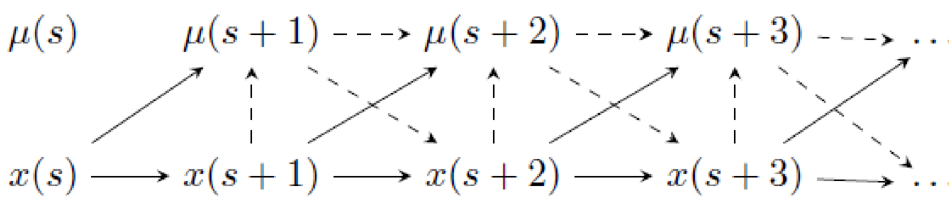
$$x_i \leftarrow \Pi_{X_i}\left[x_i - \gamma_t\left(\frac{\partial f_i}{\partial x_i} + \mu^T \frac{\partial g}{\partial x_i} + \alpha_t x_i\right)\right]$$
$$\mu \leftarrow \Pi_M[\mu + \gamma_t(g(x) - \alpha_t \mu)]$$
$$\mu \leftarrow \mu + v(t)$$

*The loss of accuracy is defined by*

$$\Lambda_D(T) = max_{x(0)\in X,\mu(0)\in M}\, Var\left[\mu_{D,v(T)}^{x(0),\mu(0)}(T) - \mu_{D,0}^{x(0),\mu(0)}(T)\right]$$

# Sensitivity

Sensitivity: influence of perturbing the sensitive data on observation




For temporary perturbation on $a(s)$, the noise should be

$$\Delta_s(t) = \begin{cases} 0, & 1 \le t \le s \\ \gamma_s \gamma_{s+1} l, & t = s+1 \\ \gamma_s \gamma_t \Pi_{k=s}^{t-1} (1 - \alpha_k \gamma_k) l, & t \ge s+2 \end{cases}$$

# Noise-adding Mechanism

Mechanism: Add noise

$$v\,(t) = \begin{cases} 0, & t = 1 \\ \gamma_1\gamma_2 lw, & t = 2 \\ \gamma_t\left(\gamma_{t-1} + \Sigma_{s=1}^{t-1}\gamma_s \Pi_{k=s+1}^{t-1}(1-\alpha_k\gamma_k)\right)lw, & t \geq 3 \end{cases}$$

$$w \sim Lap(\frac{1}{\varepsilon})$$

Asymptotics

$$v(t) \preccurlyeq \frac{\gamma_1 lwt^{-(c_1-c_2)}}{\alpha_1},$$

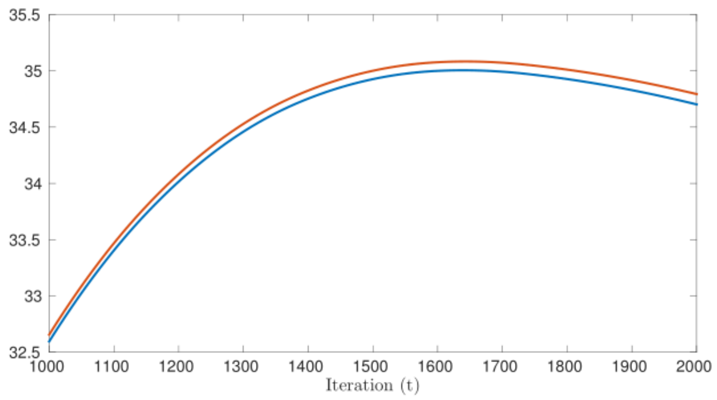
# Trade-off

The loss of accuracy is bounded asymptotically by

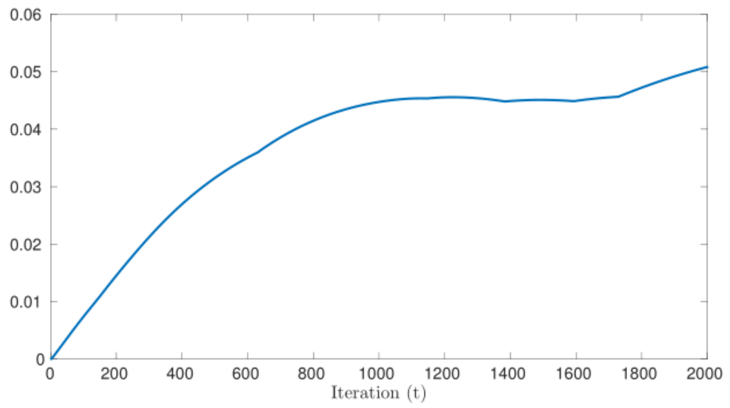
$$\Lambda_D(T) \leq \frac{2T^{2c_2}l}{\alpha_1^2\varepsilon^2}$$

higher privacy level $\leftrightarrow$ smaller $\varepsilon$ $\leftrightarrow$ larger $\Lambda_D$ $\leftrightarrow$ larger error

# Simulations




The dual trajectory $\mu_{D,v(T)}^{x(0),\mu(0)}(T)$ and $\mu_{D,0}^{x(0),\mu(0)}(T)$




$$\frac{\left|\mu_{D,v(T)}^{x(0),\mu(0)}(T) - \mu_{D,0}^{x(0),\mu(0)}(T)\right|}{\Lambda_D(T)}$$

# Summary

- Privacy in distributed optimization
- Trade-off between privacy and accuracy