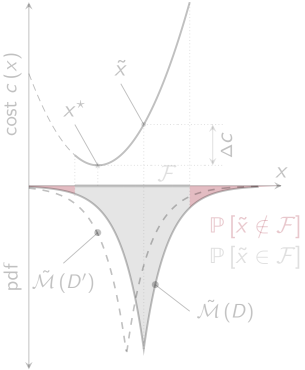
Privacy-Preserving Perturbations of Convex Optimization Programs

Vladimir Dvorkin, MIT Energy Initiative

Stats&LIDS Tea Talks

May 4, 2022



The Learning ORDER project

- **Learning for Operationalizing Data into Energy Management**
- with A. Botterud and D. Mallapragada
- Started in Mar 2022, spans two years
- Three main research thrusts:
  - Marketplace for energy datasets
  - Differential privacy for energy datasets
  - Performance-oriented learning for control

MARIE CURIE ACTIONS

Fundación IBERDROLA ESPAÑA

Convex optimization solves real-world problems

$$
\begin{aligned}
\min_{x} \quad & c^\top x \\
\text{s.t.} \quad & b - Ax \in \mathcal{K}
\end{aligned}
$$

- Conic optimization program
- Optimization dataset $\mathcal{D} = \{c, b, A\}$
- Optimal solution $x^\star$ is dataset-specific
- Often, $x^\star(\mathcal{D}) \neq x^\star(\mathcal{D}')$ for different datasets $\mathcal{D}$ and $\mathcal{D}'$

Healthcare

Logistics

Credit scoring

eCommerce

Energy forecasting

Distribution grids

Formalization of privacy

Privacy definition

The right of data owner to be protected from an unauthorized disclosure of the private information when his/her data is **utilized**.

- Optimization as a mapping $x^\star : \mathbb{D} \mapsto \mathbb{X}$
- Privacy adversary mapping $\mathcal{A} : \mathbb{X} \mapsto \mathbb{D}$
- **Privacy goal** is to mislead the adversary

This is achieved with **ε−differential privacy:**

- Let $\tilde{x}^\star$ be a random counterpart of $x^\star$
- Any dataset pair $\mathcal{D}, \mathcal{D}' \in \mathbb{D}$ is adjacent if

$$\|\mathcal{D} - \mathcal{D}'\| \leqslant \alpha$$

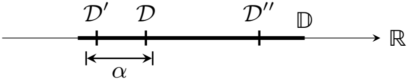
- For two adjacent datasets $\mathcal{D}$ and $\mathcal{D}'$:

$$x^\star(\mathcal{D}) \neq x^\star(\mathcal{D}') \;\; \text{but} \;\; \tilde{x}^\star(\mathcal{D}) \approx \tilde{x}^\star(\mathcal{D}')$$

Formalization of privacy

Privacy definition

The right of data owner to be protected from an unauthorized disclosure of the private information when his/her data is **utilized**.



- Optimization as a mapping $x^\star : \mathbb{D} \mapsto \mathbb{X}$
- Privacy adversary mapping $\mathcal{A} : \mathbb{X} \mapsto \mathbb{D}$
- **Privacy goal** is to mislead the adversary

This is achieved with **ε—differential privacy:**

- Let $\tilde{x}^\star$ be a random counterpart of $x^\star$
- Any dataset pair $\mathcal{D}, \mathcal{D}' \in \mathbb{D}$ is adjacent if

$$\|\mathcal{D} - \mathcal{D}'\| \leqslant \alpha$$

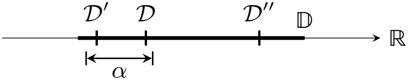
- For two adjacent datasets $\mathcal{D}$ and $\mathcal{D}'$:

$$x^\star(\mathcal{D}) \neq x^\star(\mathcal{D}') \text{ but } \tilde{x}^\star(\mathcal{D}) \approx \tilde{x}^\star(\mathcal{D}')$$

Formalization of privacy

Privacy definition

The right of data owner to be protected from an unauthorized disclosure of the private information when his/her data is **utilized**.



- Optimization as a mapping $x^\star : \mathbb{D} \mapsto \mathbb{X}$
- Privacy adversary mapping $\mathcal{A} : \mathbb{X} \mapsto \mathbb{D}$
- **Privacy goal** is to mislead the adversary

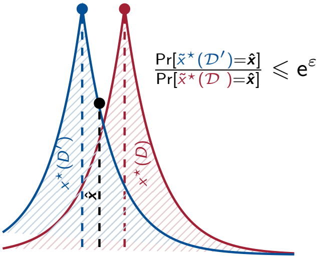
This is achieved with **ε—differential privacy:**

- Let $\tilde{x}^\star$ be a random counterpart of $x^\star$
- Any dataset pair $\mathcal{D}, \mathcal{D}' \in \mathbb{D}$ is adjacent if

$$\|\mathcal{D} - \mathcal{D}'\| \leqslant \alpha$$

- For two adjacent datasets $\mathcal{D}$ and $\mathcal{D}'$:

$$x^\star(\mathcal{D}) \neq x^\star(\mathcal{D}') \;\text{ but }\; \tilde{x}^\star(\mathcal{D}) \approx \tilde{x}^\star(\mathcal{D}')$$



Key principle of data privacy: perturb data but preserve its value

The Starry Night by Vincent Van Gogh

1889 (Museum of Modern Art, NYC)

1888 (Musée d'Orsay's, Paris)

The value of each painting is well over $100 million

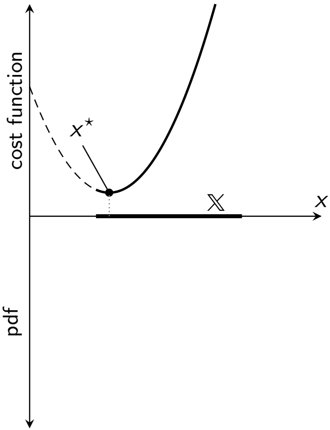
The Starry Night by Vincent Van Gogh

1889 (Museum of Modern Art, NYC)

1888 (Musée d'Orsay's, Paris)

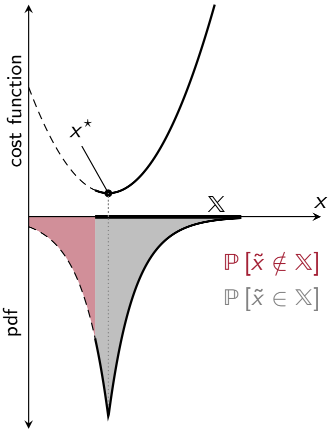
The value of each painting is well over $100 million

- Perturbation of the optimal solution $x^{\star}$ often fails to ensure feasibility
- We seek solution $\overline{x}$ whose perturbation is
  - $\varepsilon-$differentially private
  - Feasible with a high probability
  - Cost-optimal w.r.t. some risk measure



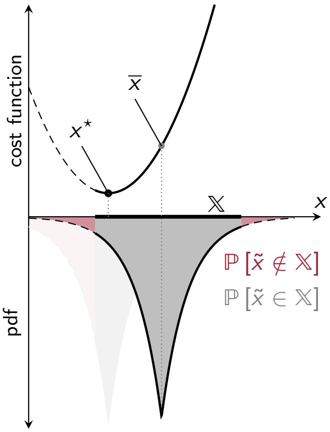
Problem statement

- Perturbation of the optimal solution $x^\star$ often fails to ensure feasibility
- We seek solution $\overline{x}$ whose perturbation is
  - $\varepsilon-$differentially private
  - Feasible with a high probability
  - Cost-optimal w.r.t. some risk measure



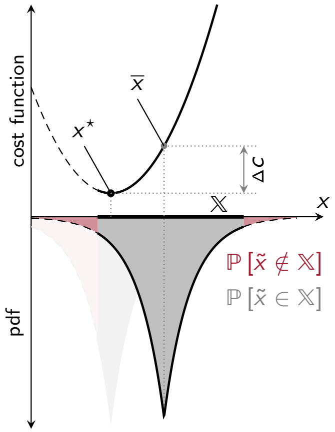
Problem statement

- Perturbation of the optimal solution $x^\star$ often fails to ensure feasibility
- We seek solution $\overline{x}$ whose perturbation is
  - $\varepsilon-$differentially private
  - Feasible with a high probability
  - Cost-optimal w.r.t. some risk measure



Problem statement

- Perturbation of the optimal solution $x^\star$ often fails to ensure feasibility
- We seek solution $\overline{x}$ whose perturbation is
  - $\varepsilon-$differentially private
  - Feasible with a high probability
  - Cost-optimal w.r.t. some risk measure



Differential privacy meets stochastic programming

Deterministic program

Business as usual

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & b - Ax \in \mathcal{K} \end{aligned}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\begin{aligned} \min_{\tilde{x}(\xi)} \quad & \mathbb{F}[c^\top \tilde{x}(\xi)] \\ \text{s.t.} \quad & \mathbb{P}[b - A\tilde{x}(\xi) \in \mathcal{K}] \geqslant 1 - \eta \end{aligned}$$

Differential privacy meets stochastic programming

Deterministic program

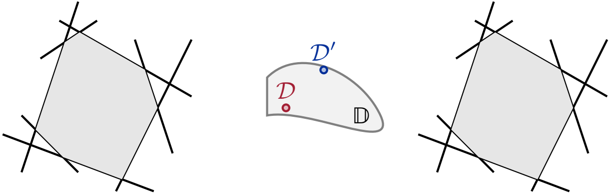
Business as usual

$$\min_{x} \quad c^\top x$$
$$\text{s.t.} \quad b - Ax \in \mathcal{K}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\min_{\tilde{x}(\xi)} \quad \mathbb{F}[c^\top \tilde{x}(\xi)]$$
$$\text{s.t.} \quad \mathbb{P}[b - A\tilde{x}(\xi) \in \mathcal{K}] \geqslant 1 - \eta$$



Differential privacy meets stochastic programming

Deterministic program

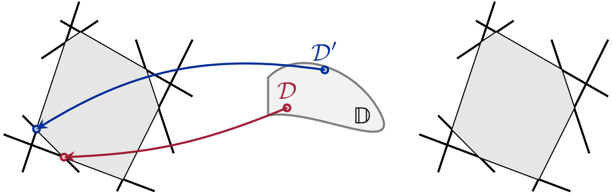
Business as usual

$$\min_{x} \quad c^\top x$$
$$\text{s.t.} \quad b - Ax \in \mathcal{K}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\min_{\tilde{x}(\xi)} \quad \mathbb{F}[c^\top \tilde{x}(\xi)]$$
$$\text{s.t.} \quad \mathbb{P}[b - A\tilde{x}(\xi) \in \mathcal{K}] \geqslant 1 - \eta$$



Differential privacy meets stochastic programming

Deterministic program

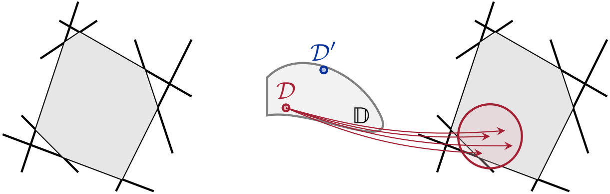
Business as usual

$$\min_{x} \quad c^\top x$$
$$\text{s.t.} \quad b - Ax \in \mathcal{K}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\min_{\tilde{x}(\xi)} \quad \mathbb{F}[c^\top \tilde{x}(\xi)]$$
$$\text{s.t.} \quad \mathbb{P}[b - A\tilde{x}(\xi) \in \mathcal{K}] \geqslant 1 - \eta$$



Differential privacy meets stochastic programming

Deterministic program

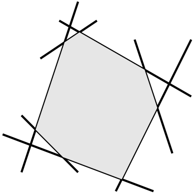
Business as usual

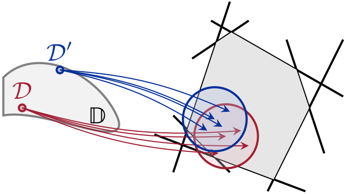
$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & b - Ax \in \mathcal{K} \end{aligned}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\begin{aligned} \min_{\tilde{x}(\xi)} \quad & \mathbb{F}[c^\top \tilde{x}(\xi)] \\ \text{s.t.} \quad & \mathbb{P}[b - A\tilde{x}(\xi) \in \mathcal{K}] \geqslant 1 - \eta \end{aligned}$$





Differential privacy meets stochastic programming

Deterministic program

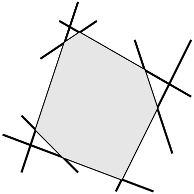
Business as usual

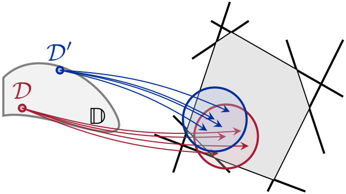
$$\min_{x} \quad c^\top x$$
$$\text{s.t.} \quad b - Ax \in \mathcal{K}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\min_{\overline{x}, X \in \mathbb{X}} \quad \mathbb{F}[c^\top(\overline{x} + X\xi)]$$
$$\text{s.t.} \quad \mathbb{P}[b - A(\overline{x} + X\xi) \in \mathcal{K}] \geqslant 1 - \eta$$





We achieve such a randomization using linear decision rules:

$$\tilde{x}(\xi) = \overline{x} + X\xi$$

- $\overline{x}$−mean value, data dependent $\overline{x} = \overline{x}(\mathcal{D})$
- $X\xi$−recourse, must be made data independent

Differential privacy meets stochastic programming

Deterministic program

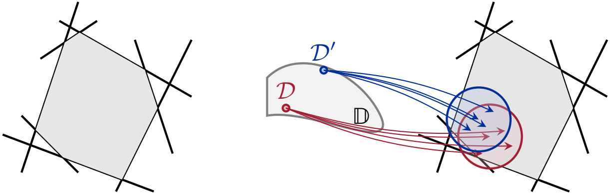
Business as usual

$$\min_{x} \quad c^\top x$$
$$\text{s.t.} \quad b - Ax \in \mathcal{K}$$

Stochastic program

Perturbation, risk-min. s.t. chance constraint

$$\min_{\overline{x}, X \in \mathbb{X}} \quad \mathbb{F}[c^\top(\overline{x} + X\xi)]$$
$$\text{s.t.} \quad \mathbb{P}[b - A(\overline{x} + X\xi) \in \mathcal{K}] \geqslant 1 - \eta$$



Our key result is to prove $\varepsilon-$differential privacy of $\tilde{x}(\mathcal{D})$, i.e.,

- Any pair of $\alpha-$adjacent datasets $\mathcal{D}', \mathcal{D} \in \mathbb{D}$
- Perturbation $\xi \sim \text{Lap}(0, \frac{\Delta_\alpha}{\varepsilon})$
- Worst-case sensitivity of $x^\star$ to $\alpha-$adjacent datasets

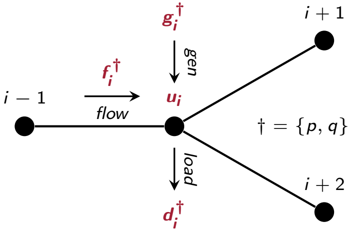
$$\frac{\Pr[\overline{x}(\mathcal{D}') + X(\mathcal{D}')\xi = \hat{x}]}{\Pr[\overline{x}(\mathcal{D}) + X(\mathcal{D})\xi = \hat{x}]} \leqslant e^{\varepsilon}$$

Two applications of private convex optimization:

- Private distribution grid control
- Private monotone wind power curve fitting

Private distribution optimal power flow (OPF)

- Distribution grid topology:



- Distribution AC optimal power flow:
  - Minimize total dispatch cost
  - Subject to OPF equations:

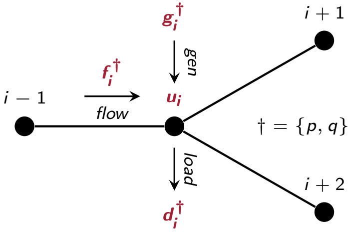
$$f_i^\dagger = d_i^\dagger - g_i^\dagger + \sum_{\ell \in \mathcal{D}_i} f_\ell^\dagger, \quad \forall \ell \in \mathcal{L}$$

$$u_i = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \quad \forall i \in \mathcal{N}$$

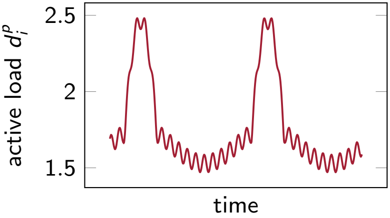
and flow, generation, and voltage limits

Private distribution optimal power flow (OPF)

- Distribution grid topology:



- Dynamic load profile ...



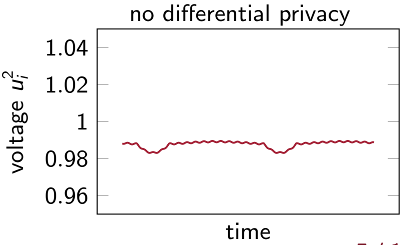
- Distribution AC optimal power flow:
  - Minimize total dispatch cost
  - Subject to OPF equations:

$$\boldsymbol{f_i^\dagger} = \boldsymbol{d_i^\dagger} - \boldsymbol{g_i^\dagger} + \sum_{\ell \in \mathcal{D}_i} f_\ell^\dagger, \quad \forall \ell \in \mathcal{L}$$

$$\boldsymbol{u_i} = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \quad \forall i \in \mathcal{N}$$

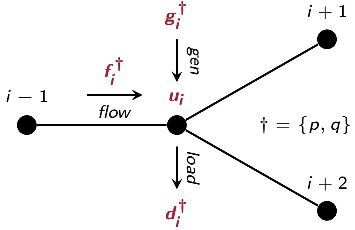
and flow, generation, and voltage limits

- ... leaks through voltage measurements

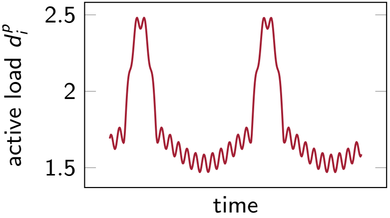


Private distribution optimal power flow (OPF)

- Distribution grid topology:



- Dynamic load profile ...



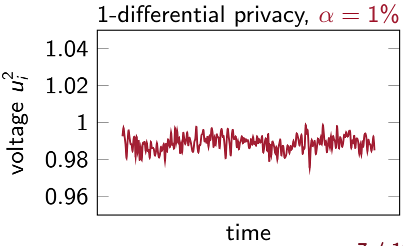
- Distribution AC optimal power flow:
  - Minimize total dispatch cost
  - Subject to OPF equations:

$$\boldsymbol{f_i^\dagger} = \boldsymbol{d_i^\dagger} - \boldsymbol{g_i^\dagger} + \sum_{\ell \in \mathcal{D}_i} f_\ell^\dagger, \quad \forall \ell \in \mathcal{L}$$

$$\boldsymbol{u_i} = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \quad \forall i \in \mathcal{N}$$

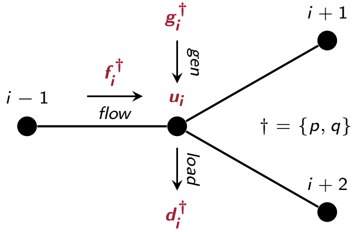
and flow, generation, and voltage limits

- ... leaks through voltage measurements

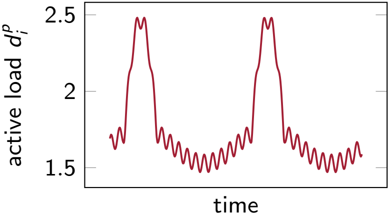


Private distribution optimal power flow (OPF)

- Distribution grid topology:



- Dynamic load profile ...



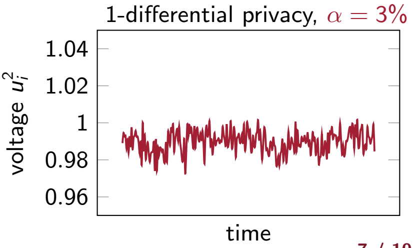
- Distribution AC optimal power flow:
  - Minimize total dispatch cost
  - Subject to OPF equations:

$$\boldsymbol{f_i^\dagger} = \boldsymbol{d_i^\dagger} - \boldsymbol{g_i^\dagger} + \sum_{\ell \in \mathcal{D}_i} f_\ell^\dagger, \quad \forall \ell \in \mathcal{L}$$

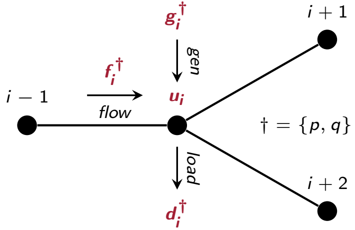
$$\boldsymbol{u_i} = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \quad \forall i \in \mathcal{N}$$

and flow, generation, and voltage limits

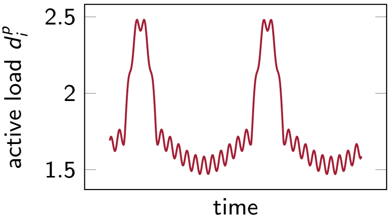
- ... leaks through voltage measurements



- Distribution grid topology:



- Dynamic load profile ...



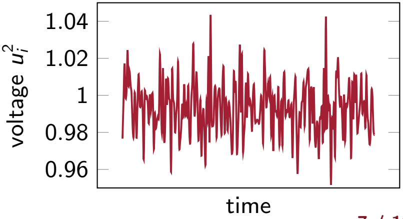
- Distribution AC optimal power flow:
  - Minimize total dispatch cost
  - Subject to OPF equations:

$$\boldsymbol{f_i^\dagger} = \boldsymbol{d_i^\dagger} - \boldsymbol{g_i^\dagger} + \sum_{\ell \in \mathcal{D}_i} f_\ell^\dagger, \quad \forall \ell \in \mathcal{L}$$

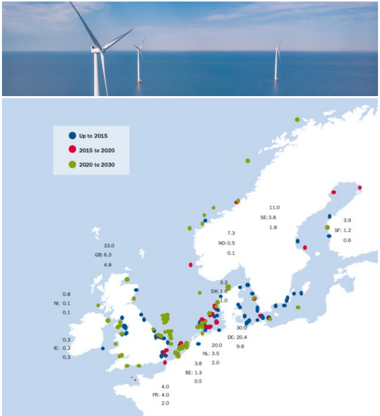
$$\boldsymbol{u_i} = u_0 - 2 \sum_{\ell \in \mathcal{R}_i} (f_\ell^p r_\ell + f_\ell^q x_\ell), \quad \forall i \in \mathcal{N}$$

and flow, generation, and voltage limits

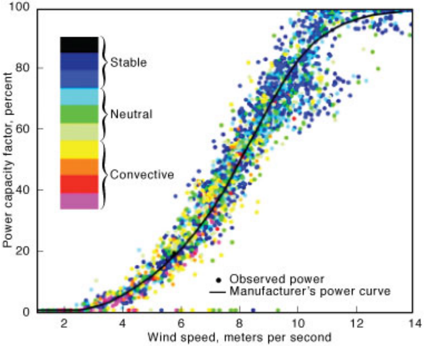
- ... leaks through voltage measurements



Private monotone wind power curve fitting – Motivation



Wind farm locations in N. Europe. Wind Europe (c)



S&TR (April–May 2014)

- Wind farms benefit from coll. learning
- How to release the model while privatizing individual farm data?

Private monotone wind power curve fitting – Example

$$\min_{\beta} \quad \mathbb{E}\left[\sum_{i=1}^{n}\left(\underbrace{y_i - \varphi(x_i)^\top\beta}_{\text{business as usual}} \underbrace{-\varphi(x_i)^\top\xi}_{\text{perturbation}}\right)^2\right]$$

$$\text{s.t.} \quad \mathbb{P}\left[C(\beta+\xi) \geqslant \mathbb{0}\right] \geqslant 1-\eta,$$

- Dataset $\{(y_1, x_1), \dots, (y_n, x_n)\}$
- Minimize regression loss function
- By finding optimal weights $\beta^\star$ ...
- ... of basis functions in vector $\varphi(x)$

- Deterministic curve fitting results in the loss of 1,513.4
- We want to make datasets indistinguishable in model weights $\beta^\star$
- The direct weight perturbation, i.e., $\beta^\star + \xi$:
  - Does not effect the goodness of fit, loss
  - Infeasible with a high probability of 13.4%
- Perturbation of the optimal chance-constrained weights:
  - Reduces the empirical infeasibility to 4.0% ($\eta = 5\%$)
  - At the expense of an increasing loss of 2,003.2 (+32%)

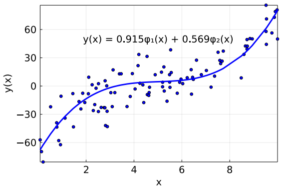
Private monotone wind power curve fitting – Example

$$\min_{\beta} \quad \mathbb{E}\left[\sum_{i=1}^{n}\Big(\underbrace{y_i - \varphi(x_i)^\top\beta}_{\text{business as usual}} \underbrace{-\varphi(x_i)^\top\xi}_{\text{perturbation}}\Big)^2\right]$$

$$\text{s.t.} \quad \mathbb{P}\left[C(\beta+\xi) \geqslant \mathbb{0}\right] \geqslant 1-\eta,$$

- Dataset $\{(y_1, x_1), \ldots, (y_n, x_n)\}$
- Minimize regression loss function
- By finding optimal weights $\beta^\star$ ...
- ... of basis functions in vector $\varphi(x)$

- Deterministic curve fitting results in the loss of 1,513.4
- We want to make datasets indistinguishable in model weights $\beta^\star$
- The direct weight perturbation, i.e., $\beta^\star + \xi$:
  - Does not effect the goodness of fit, loss
  - Infeasible with a high probability of 13.4%
- Perturbation of the optimal chance-constrained weights:
  - Reduces the empirical infeasibility to 4.0% ($\eta = 5\%$)
  - At the expense of an increasing loss of 2,003.2 (+32%)



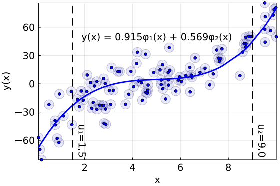
Private monotone wind power curve fitting – Example

$$\min_{\beta} \quad \mathbb{E}\left[\sum_{i=1}^{n}\left(\underbrace{y_i - \varphi(x_i)^\top\beta}_{\text{business as usual}} \underbrace{-\varphi(x_i)^\top\xi}_{\text{perturbation}}\right)^2\right]$$

$$\text{s.t.} \quad \mathbb{P}\left[C(\beta+\xi) \geqslant \mathbb{0}\right] \geqslant 1-\eta,$$

- Dataset $\{(y_1, x_1), \ldots, (y_n, x_n)\}$
- Minimize regression loss function
- By finding optimal weights $\beta^\star$ ...
- ... of basis functions in vector $\varphi(x)$

- Deterministic curve fitting results in the loss of 1,513.4
- We want to make datasets indistinguishable in model weights $\beta^\star$
- The direct weight perturbation, i.e., $\beta^\star + \xi$:
  - Does not effect the goodness of fit, loss
  - Infeasible with a high probability of 13.4%
- Perturbation of the optimal chance-constrained weights:
  - Reduces the empirical infeasibility to 4.0% ($\eta = 5\%$)
  - At the expense of an increasing loss of 2,003.2 (+32%)



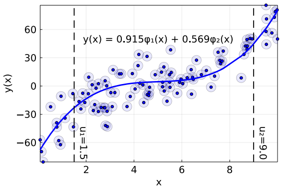
Private monotone wind power curve fitting – Example

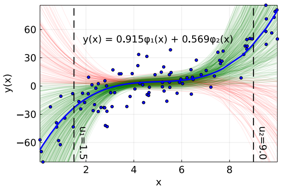
$$\min_{\beta} \quad \mathbb{E}\left[\sum_{i=1}^{n}\left(\underbrace{y_i - \varphi(x_i)^\top\beta}_{\text{business as usual}} \underbrace{-\varphi(x_i)^\top\xi}_{\text{perturbation}}\right)^2\right]$$

$$\text{s.t.} \quad \mathbb{P}\left[C(\beta+\xi) \geqslant \mathbb{0}\right] \geqslant 1-\eta,$$

- Dataset $\{(y_1, x_1), \dots, (y_n, x_n)\}$
- Minimize regression loss function
- By finding optimal weights $\beta^\star$ ...
- ... of basis functions in vector $\varphi(x)$

- Deterministic curve fitting results in the loss of 1,513.4
- We want to make datasets indistinguishable in model weights $\beta^\star$
- The direct weight perturbation, i.e., $\beta^\star + \xi$:
  - Does not effect the goodness of fit, loss
  - Infeasible with a high probability of 13.4%
- Perturbation of the optimal chance-constrained weights:
  - Reduces the empirical infeasibility to 4.0% ($\eta = 5\%$)
  - At the expense of an increasing loss of 2,003.2 (+32%)





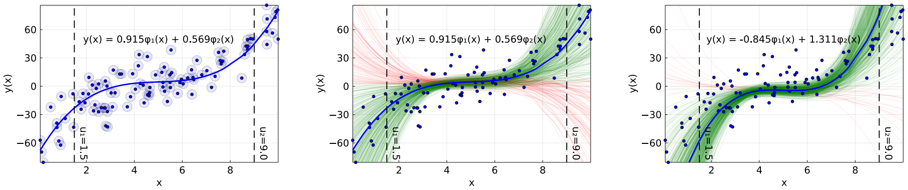
Private monotone wind power curve fitting – Example

$$\min_{\beta} \quad \mathbb{E}\left[\sum_{i=1}^{n}\left(\underbrace{y_i - \varphi(x_i)^\top \beta}_{\text{business as usual}} \underbrace{-\varphi(x_i)^\top \xi}_{\text{perturbation}}\right)^2\right]$$

$$\text{s.t.} \quad \mathbb{P}\left[C(\beta + \xi) \geqslant \mathbb{0}\right] \geqslant 1 - \eta,$$

- Dataset $\{(y_1, x_1), \ldots, (y_n, x_n)\}$
- Minimize regression loss function
- By finding optimal weights $\beta^\star$ …
- … of basis functions in vector $\varphi(x)$

- Deterministic curve fitting results in the loss of 1,513.4
- We want to make datasets indistinguishable in model weights $\beta^\star$
- The direct weight perturbation, i.e., $\beta^\star + \xi$:
  - Does not effect the goodness of fit, loss
  - Infeasible with a high probability of 13.4%
- Perturbation of the optimal chance-constrained weights:
  - Reduces the empirical infeasibility to 4.0% ($\eta = 5\%$)
  - At the expense of an increasing loss of 2,003.2 (+32%)



Thank you for your attention!

Contributions:

1. Dvorkin, V., Fioretto, F., Van Hentenryck, P., Kazempour, J. and Pinson, P.
   Differentially private convex optimization with feasibility guarantees
   Priprint, arXiv preprint arXiv:2006.12338.
2. Dvorkin, V., Fioretto, F., Van Hentenryck, P., Pinson, P. and Kazempour J.
   Differentially private optimal power flow for distribution grids
   IEEE Transactions on Power Systems, 2021
   Best 2019–2021 Paper Award
3. Dvorkin, V., Van Hentenryck, P., Kazempour, J. and Pinson P.
   Differentially private distributed optimal power flow
   2020 Conference on Decision and Control

Let's stay in touch:

DvorkinVladimir   Vladimir-Dvorkin   dvorkin@mit.edu