Bayesian Optimization for Distributionally Robust Chance-constrained Problem

Yu Inatsu [1] Shion Takeno [1]

Masayuki Karasuyama [1] Ichiro Takeuchi [1 2]

[1]Nagoya Institute of Technology

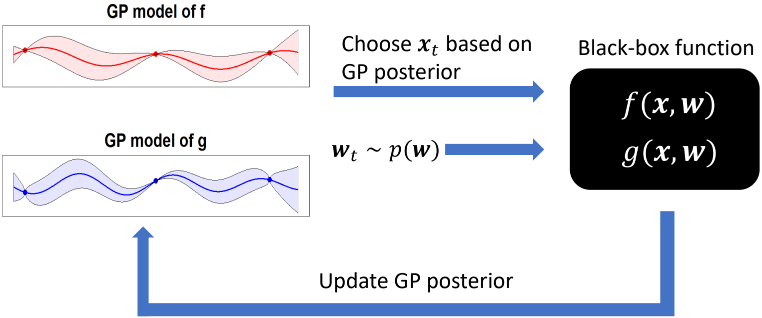
[2]RIKEN Center for Advanced Intelligence Project

Black-box Function under Uncertainty Environment

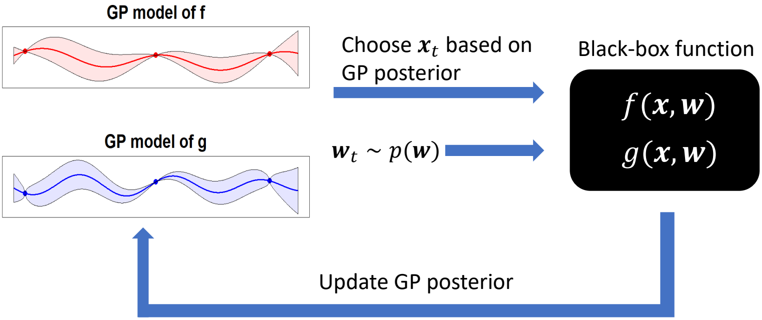
- Consider a constrained optimization for black-box functions under uncertainty environment using Gaussian process (GP).

Black-box Function under Uncertainty Environment

- Consider a constrained optimization for black-box functions under uncertainty environment using Gaussian process (GP).



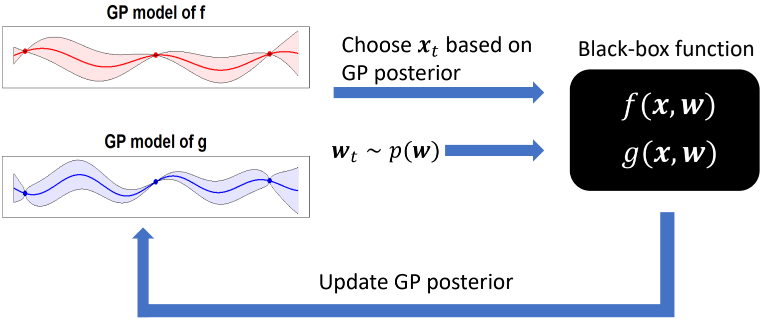
- Consider a constrained optimization for black-box functions under uncertainty environment using Gaussian process (GP).



- Two types of variables exist:
  - **Design variable** x: Fully controllable.
  - **Environmental variable** w: Realized randomly.

Black-box Function under Uncertainty Environment

- Consider a constrained optimization for black-box functions under uncertainty environment using Gaussian process (GP).



- Two types of variables exist:
  - **Design variable** x: Fully controllable.
  - **Environmental variable** w: Realized randomly.

Goal: Identify desired x by taking into account the uncertainty of w.

Robustness Measure

How to define the optimal design variable x?

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

$$
\text{(CC)}: \quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\mathsf{true})}}[f(\boldsymbol{x},\boldsymbol{w})]
$$
$$
\text{subject to} \quad \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\mathsf{true})}}[g(\boldsymbol{x},\boldsymbol{w}) > h] > \alpha.
$$

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

$$\text{(CC):} \quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x}, \boldsymbol{w})]$$
$$\text{subject to} \quad \mathbb{P}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x}, \boldsymbol{w}) > h] > \alpha.$$

  - CC problem is regard as the constrained optimization between the expectation and probability w.r.t. the uncertainty of $\boldsymbol{w}$.

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

$$\text{(CC):} \quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x},\boldsymbol{w})] \quad \text{subject to} \quad \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x},\boldsymbol{w}) > h] > \alpha.$$

  - CC problem is regard as the constrained optimization between the expectation and probability w.r.t. the uncertainty of $\boldsymbol{w}$.
- This is used in many situations such as power system control.

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

$$\text{(CC):}\quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x},\boldsymbol{w})]$$
$$\text{subject to}\ \ \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x},\boldsymbol{w})>h]>\alpha.$$

  - CC problem is regard as the constrained optimization between the expectation and probability w.r.t. the uncertainty of $\boldsymbol{w}$.
- This is used in many situations such as power system control.

If the distribution of w is unknown, CC problem cannot be defined.

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

$$\text{(CC):} \quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x},\boldsymbol{w})] \\ \text{subject to} \quad \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x},\boldsymbol{w}) > h] > \alpha.$$

  - CC problem is regard as the constrained optimization between the expectation and probability w.r.t. the uncertainty of $\boldsymbol{w}$.
- This is used in many situations such as power system control.

If the distribution of w is unknown, CC problem cannot be defined.

- We focus on a distributionally robust CC (DRCC) problem:

How to define the optimal design variable x?

- Chance-constrained (CC) problem is reasonable:

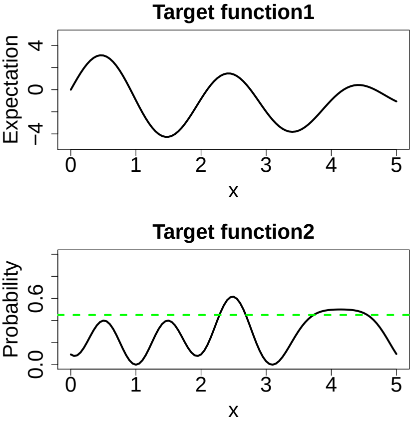
$$\text{(CC):}\quad \arg\max_{\boldsymbol{x}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x},\boldsymbol{w})] \quad \text{subject to} \quad \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x},\boldsymbol{w}) > h] > \alpha.$$

  - CC problem is regard as the constrained optimization between the expectation and probability w.r.t. the uncertainty of $\boldsymbol{w}$.
- This is used in many situations such as power system control.

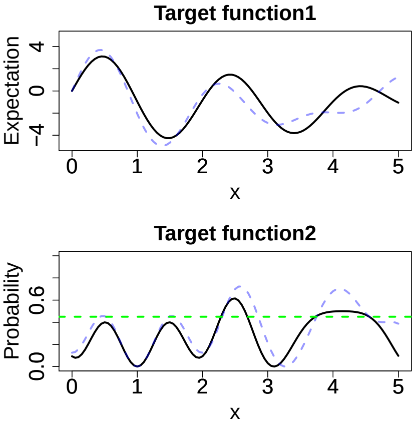
If the distribution of w is unknown, CC problem cannot be defined.

- We focus on a distributionally robust CC (DRCC) problem:

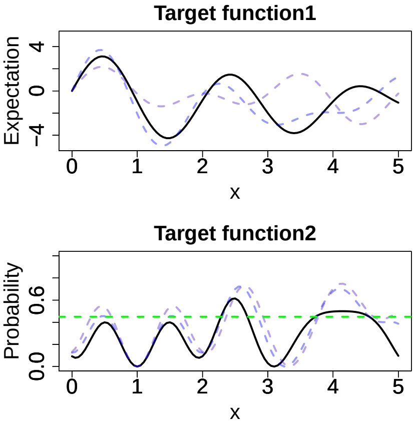
$$\text{(DRCC):}\quad \arg\max_{\boldsymbol{x}} \inf_{\mathbb{P}_{\boldsymbol{w}}\in\mathcal{A}} \mathbb{E}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}}[f(\boldsymbol{x},\boldsymbol{w})] \quad \text{subject to} \quad \inf_{\mathbb{P}_{\boldsymbol{w}}\in\mathcal{A}} \mathbb{P}_{\boldsymbol{w}\sim\mathbb{P}_{\boldsymbol{w}}}[g(\boldsymbol{x},\boldsymbol{w}) > h] > \alpha.$$


Target function1
Expectation
4
0
−4
0
1
2
3
4
5
x
Target function2
Probability
0.0
0.6
0
1
2
3
4
5
x

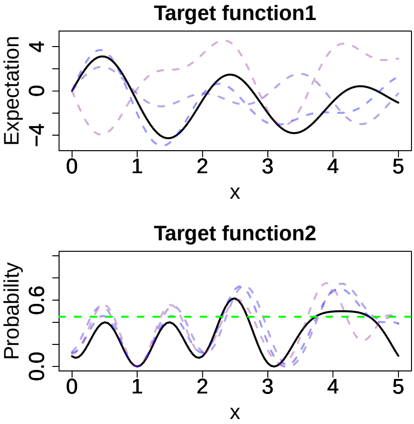
Problem Setup (DRCC Problem)



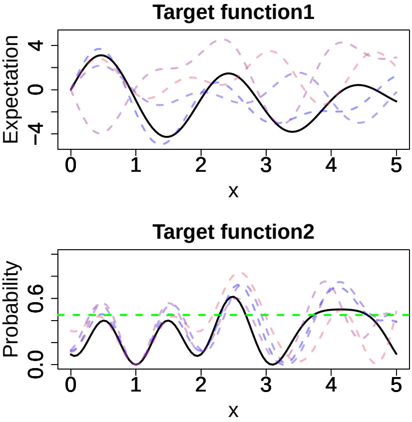
Problem Setup (DRCC Problem)



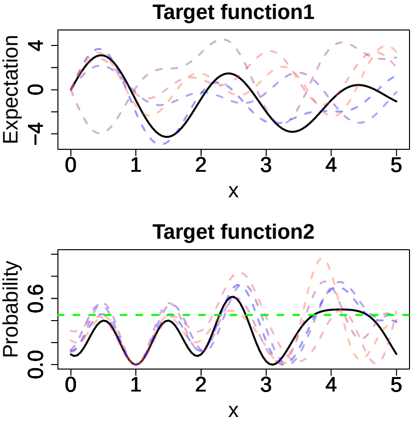
Problem Setup (DRCC Problem)



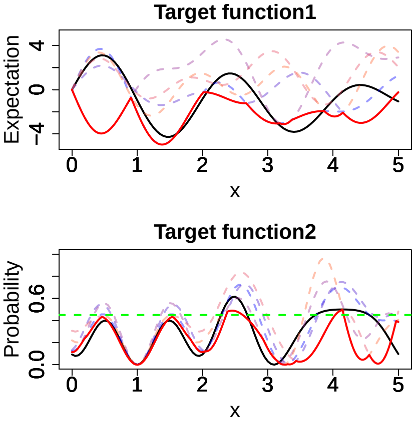
Problem Setup (DRCC Problem)



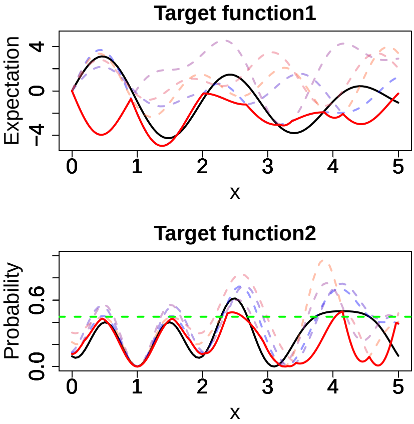
Problem Setup (DRCC Problem)



Problem Setup (DRCC Problem)



Problem Setup (DRCC Problem)



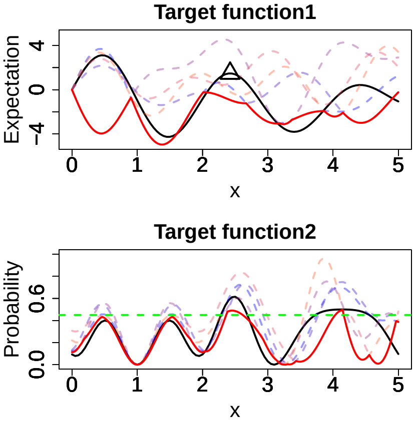
- **Target functions**:

$$F(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[f(\boldsymbol{x}, \boldsymbol{w})],$$

$$G(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{P}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[g(\boldsymbol{x}, \boldsymbol{w}) > h].$$

  - $\mathcal{A} \equiv \{\mathbb{P}_{\boldsymbol{w}} \mid d(\mathbb{P}_{\boldsymbol{w}}, \mathbb{P}_{\text{ref}}) < \epsilon\}$: Family of candidate distributions.
  - h: Given threshold.
  - $\mathbb{P}_{\text{ref}}$: Given reference distribution.

Problem Setup (DRCC Problem)



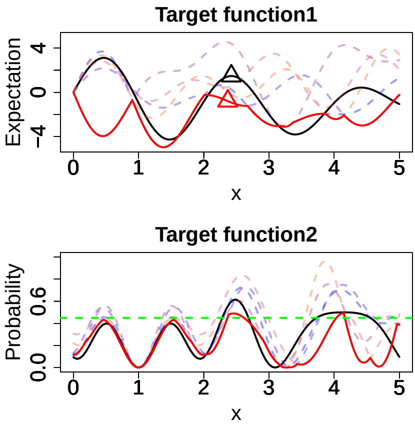
- **Target functions**:

$$F(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[f(\boldsymbol{x}, \boldsymbol{w})],$$

$$G(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{P}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[g(\boldsymbol{x}, \boldsymbol{w}) > h].$$

  - $\mathcal{A} \equiv \{\mathbb{P}_{\boldsymbol{w}} \mid d(\mathbb{P}_{\boldsymbol{w}}, \mathbb{P}_{\text{ref}}) < \epsilon\}$: Family of candidate distributions.
  - h: Given threshold.
  - $\mathbb{P}_{\text{ref}}$: Given reference distribution.

- $\triangle$ is the optimal value of CC problem.



- **Target functions**:

$$F(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[f(\boldsymbol{x}, \boldsymbol{w})],$$

$$G(\boldsymbol{x}) = \inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \mathbb{P}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}}[g(\boldsymbol{x}, \boldsymbol{w}) > h].$$

  - $\mathcal{A} \equiv \{\mathbb{P}_{\boldsymbol{w}} \mid d(\mathbb{P}_{\boldsymbol{w}}, \mathbb{P}_{\text{ref}}) < \epsilon\}$: Family of candidate distributions.
  - h: Given threshold.
  - $\mathbb{P}_{\text{ref}}$: Given reference distribution.
- $\triangle$ is the optimal value of CC problem.

Goal: Find the optimal value of DRCC $\triangle$ efficiently.

- We propose a Bayesian optimization method for DRCC problem by constructing credible intervals that contain target functions w.h.p.

- We propose a Bayesian optimization method for DRCC problem by constructing credible intervals that contain target functions w.h.p.

$$\text{Credible interval of } F(\boldsymbol{x}) : \quad \underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \underbrace{f(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}},$$

$$\text{Credible interval of } G(\boldsymbol{x}) : \quad \underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \overbrace{\mathbb{1}[\underbrace{g(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} > h]}^{\text{(ii)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}}.$$

Basic idea of our proposed method

- We propose a Bayesian optimization method for DRCC problem by constructing credible intervals that contain target functions w.h.p.

$$\text{Credible interval of } F(\boldsymbol{x}) : \quad \underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \underbrace{f(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}},$$

$$\text{Credible interval of } G(\boldsymbol{x}) : \quad \underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \overbrace{\mathbb{1}[\underbrace{g(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} > h]}^{\text{(ii)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}}.$$

Basic idea of our proposed method

(i) Assume GP for f and g, and calculate their credible intervals.

- We propose a Bayesian optimization method for DRCC problem by constructing credible intervals that contain target functions w.h.p.

Credible interval of $F(\boldsymbol{x})$: $$\underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \underbrace{f(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}},$$

Credible interval of $G(\boldsymbol{x})$: $$\underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \overbrace{\mathbb{1}[\underbrace{g(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} > h]}^{\text{(ii)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}}.$$

Basic idea of our proposed method

(i) Assume GP for f and g, and calculate their credible intervals.

(ii) Construct a credible interval of $\mathbb{1}[\cdot]$ using the upper and lower ends of the credible interval of g.

- We propose a Bayesian optimization method for DRCC problem by constructing credible intervals that contain target functions w.h.p.

Credible interval of $F(\boldsymbol{x})$: $$\underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \underbrace{f(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}},$$

Credible interval of $G(\boldsymbol{x})$: $$\underbrace{\inf_{\mathbb{P}_{\boldsymbol{w}} \in \mathcal{A}} \int \overbrace{\mathbb{1}[\underbrace{g(\boldsymbol{x}, \boldsymbol{w})}_{\text{(i)}} > h]}^{\text{(ii)}} \mathrm{d}\mathbb{P}_{\boldsymbol{w}}}_{\text{(iii)}}.$$

Basic idea of our proposed method

(i) Assume GP for f and g, and calculate their credible intervals.

(ii) Construct a credible interval of $\mathbb{1}[\cdot]$ using the upper and lower ends of the credible interval of g.

(iii) Compute credible intervals of target functions by using them.

Theoretical results

- Our proposed method satisfies the following w.h.p.

Theoretical results

- Our proposed method satisfies the following w.h.p.

  (Accuracy for DRCC): Estimated solution $\boldsymbol{x}_{\text{est,DRCC}}$ satisfies

$$F(\boldsymbol{x}_{\text{true,DRCC}}) - F(\boldsymbol{x}_{\text{est,DRCC}}) < \eta,$$
$$G(\boldsymbol{x}_{\text{est,DRCC}}) > \alpha - \eta, \quad \eta > 0.$$

- Our proposed method satisfies the following w.h.p.

  (Accuracy for DRCC): Estimated solution $\boldsymbol{x}_{\text{est,DRCC}}$ satisfies

$$F(\boldsymbol{x}_{\text{true,DRCC}}) - F(\boldsymbol{x}_{\text{est,DRCC}}) < \eta,$$
$$G(\boldsymbol{x}_{\text{est,DRCC}}) > \alpha - \eta, \quad \eta > 0.$$

  (Convergence): Algorithm terminates with sub-linear rate.

Theoretical results

- Our proposed method satisfies the following w.h.p.

(Accuracy for DRCC): Estimated solution $\boldsymbol{x}_{\text{est,DRCC}}$ satisfies

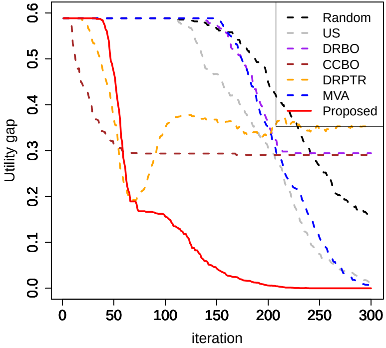
$$F(\boldsymbol{x}_{\text{true,DRCC}}) - F(\boldsymbol{x}_{\text{est,DRCC}}) < \eta,$$
$$G(\boldsymbol{x}_{\text{est,DRCC}}) > \alpha - \eta, \quad \eta > 0.$$

(Convergence): Algorithm terminates with sub-linear rate.

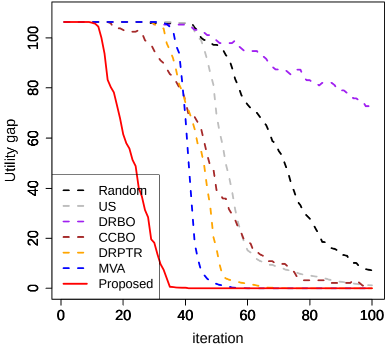
(Accuracy for CC): $\boldsymbol{x}_{\text{est,DRCC}}$ is a good solution of the CC problem:

$$\mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x}_{\text{true,CC}}, \boldsymbol{w})] - \mathbb{E}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[f(\boldsymbol{x}_{\text{est,DRCC}}, \boldsymbol{w})] < 1.5\eta,$$
$$\mathbb{P}_{\boldsymbol{w} \sim \mathbb{P}_{\boldsymbol{w}}^{(\text{true})}}[g(\boldsymbol{x}_{\text{est,DRCC}}, \boldsymbol{w}) > h] > \alpha - 1.5\eta.$$

- Evaluate the performance of our proposed method synthetic and SIR simulation data.
- Measure the performance based on the utility gap.



Synthetic



SIR simulation

- **Solid lines** represent our proposed methods.