Convex Relaxation Regression (CoRR)

Eva Dyer

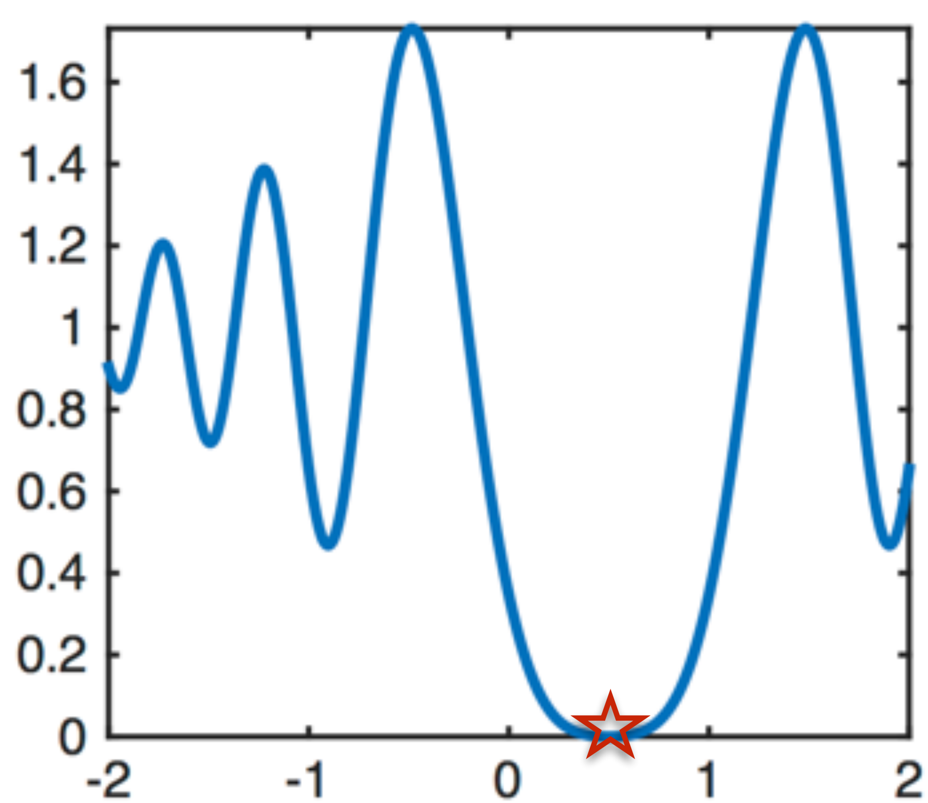
work with: Mohammad Azar
Konrad Körding

NORTHWESTERN UNIVERSITY

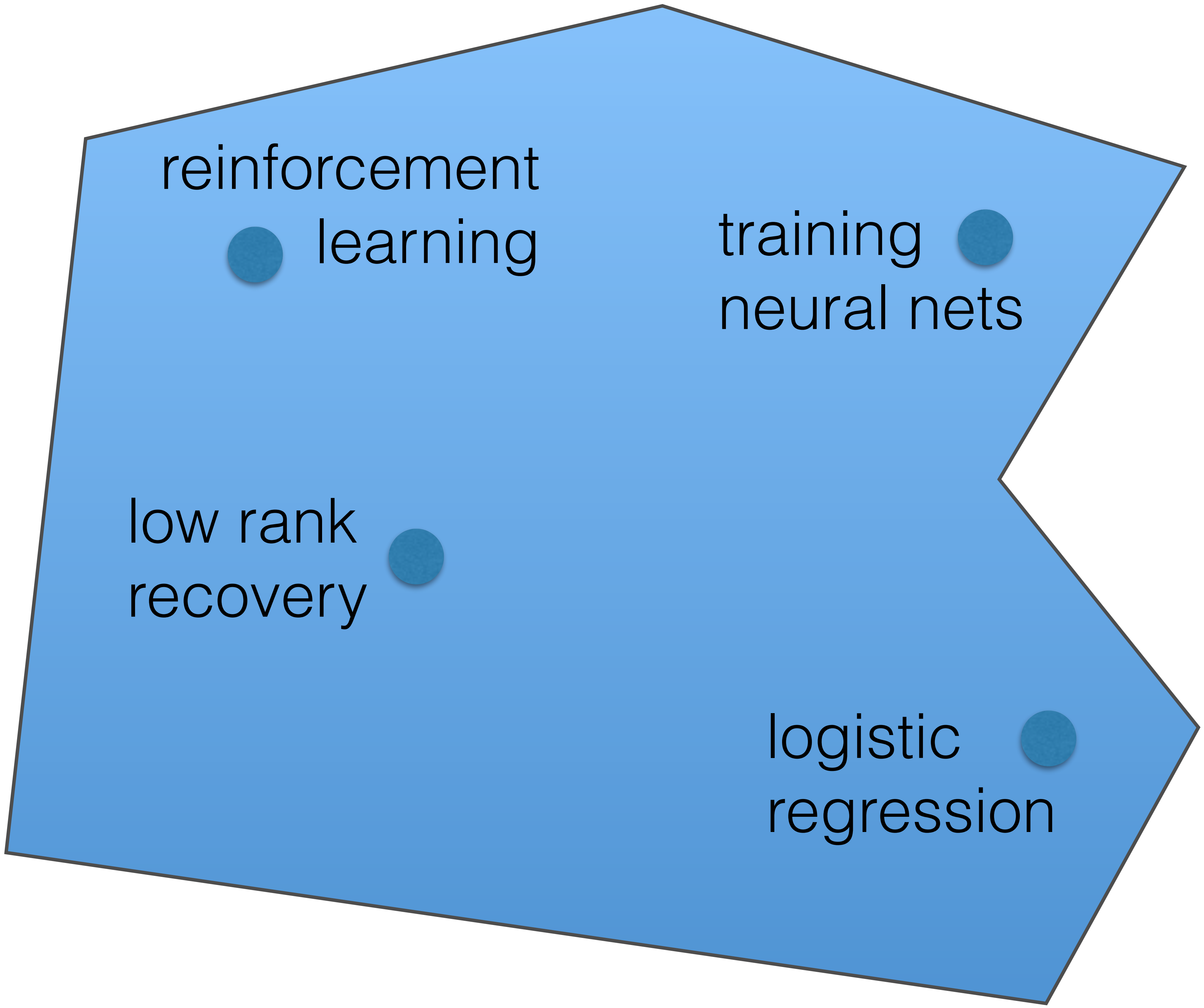
RIC
Rehabilitation Institute of Chicago

global optimization

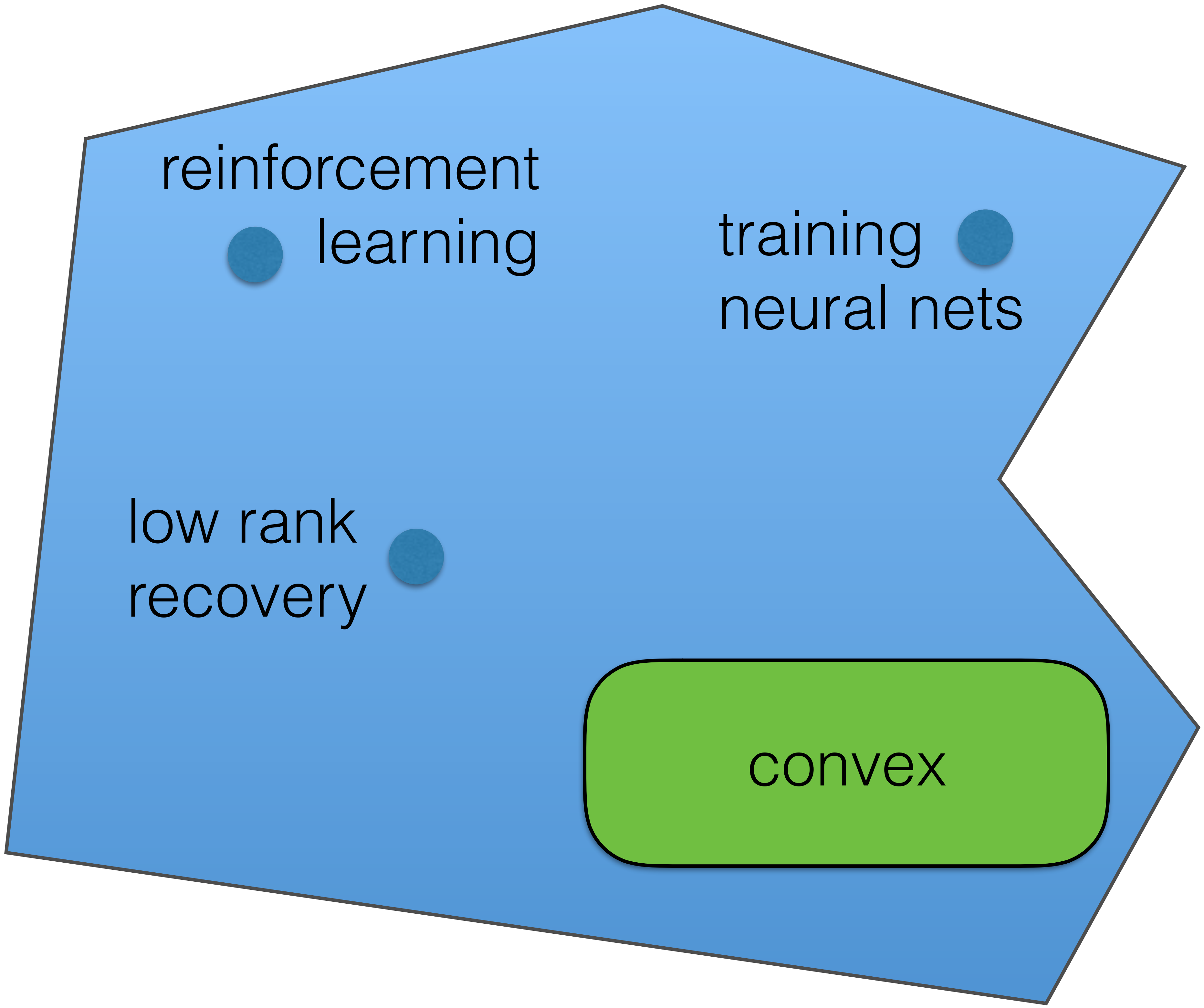
$$f^* := \min_{x \in \mathcal{X}} f(x)$$



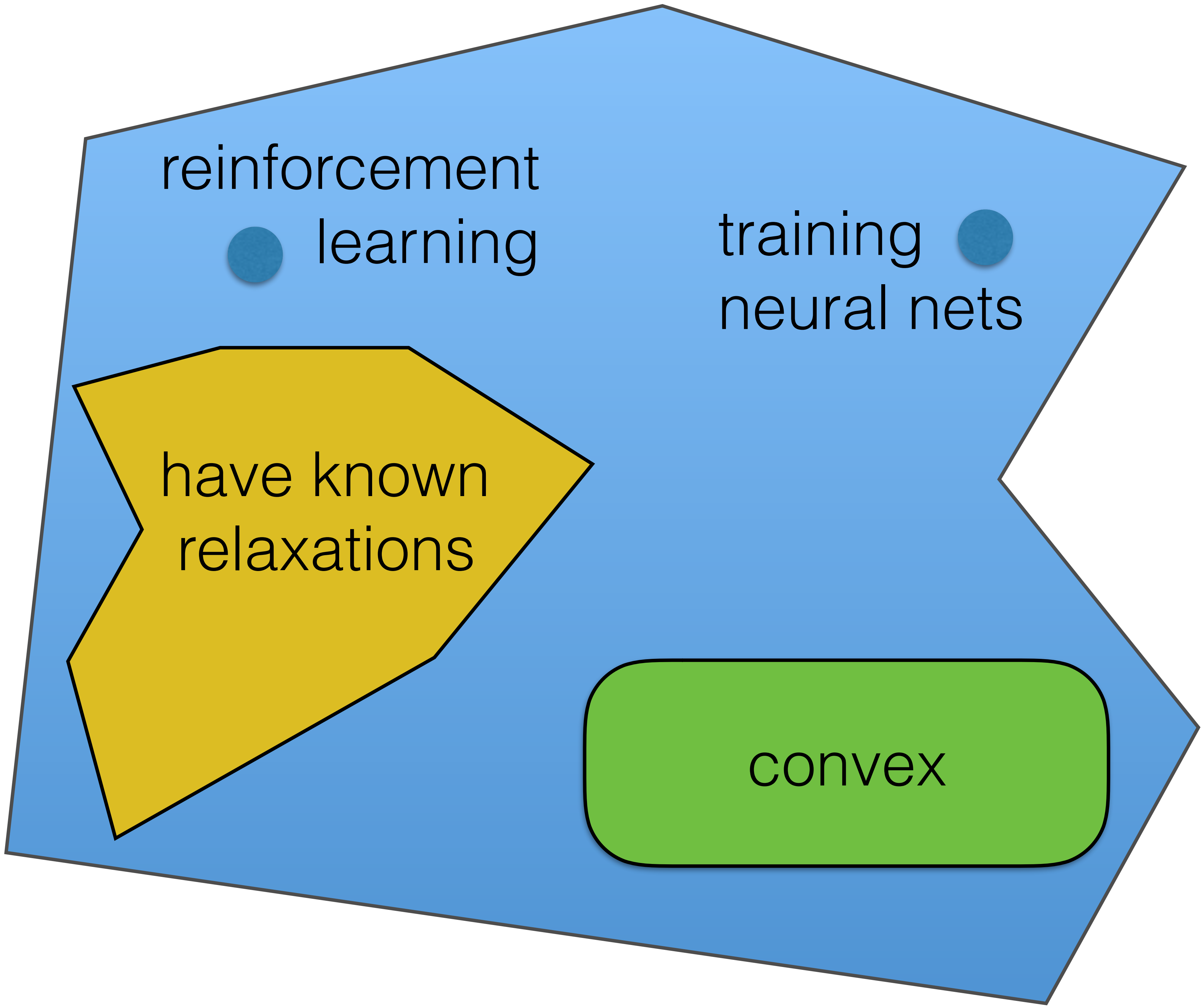
optimization problems



optimization problems

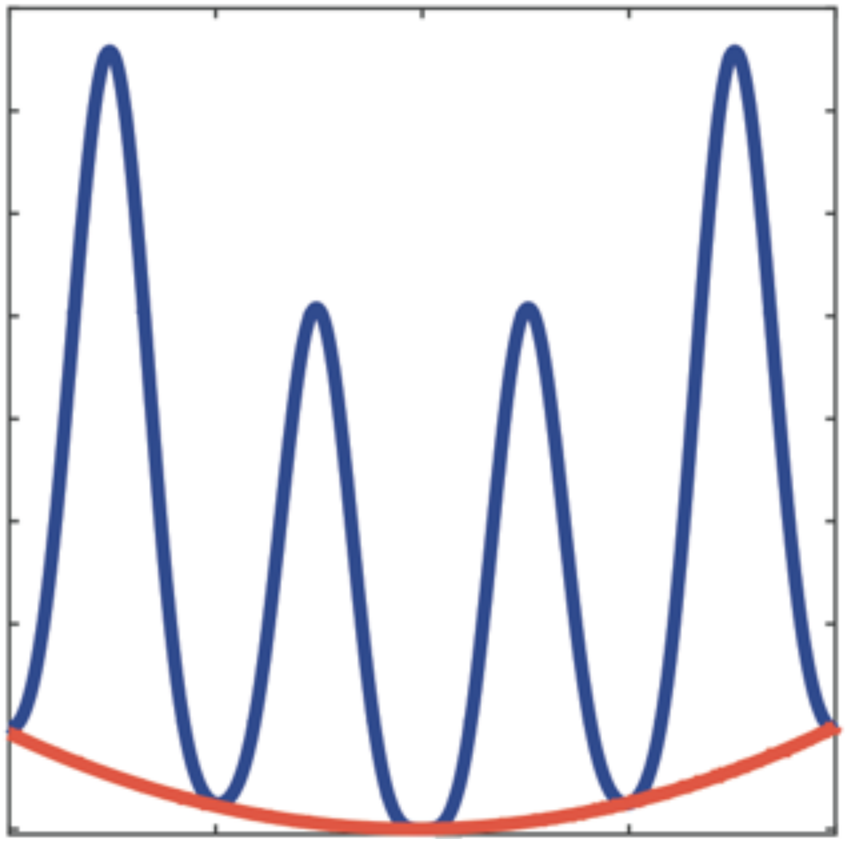


optimization problems



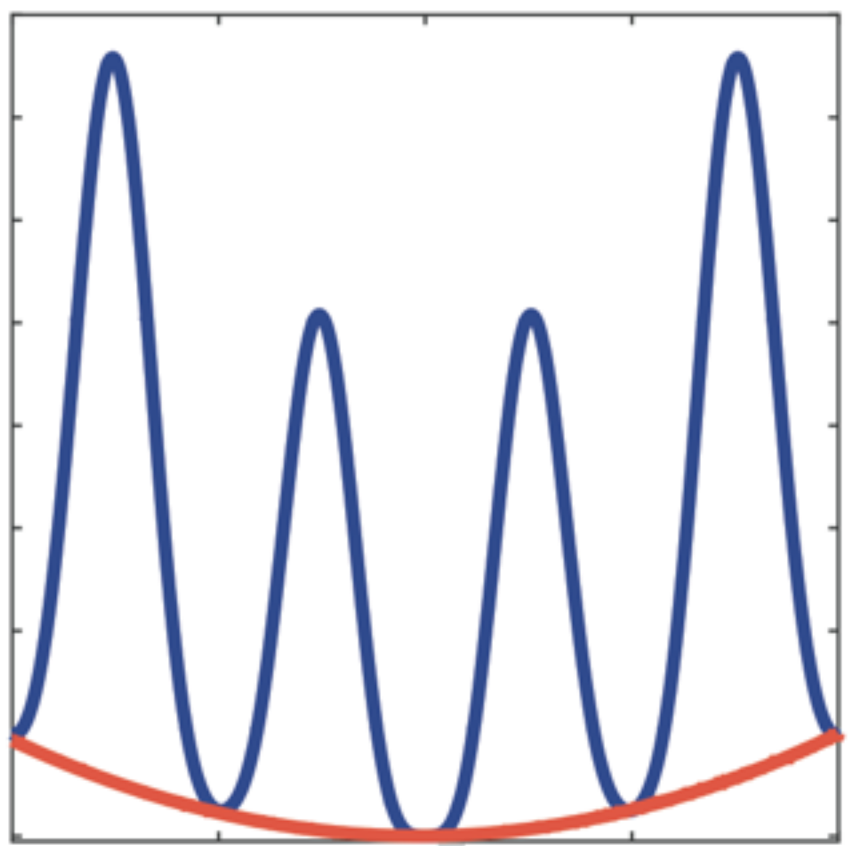
convex relaxation

Step 1: find a tight lower bound to f(x)



convex relaxation

Step 2: optimize the convex surrogate instead of f(x)



convex envelope

convex envelope = tightest convex lower bound

Kleibohm, 1967

Let f_c be the convex envelope of $f : \mathcal{X} \to \mathbb{R}$.
Then (a) $\min_{x \in \mathcal{X}} f_c(x) = f^*$ and (b) $\mathcal{X}_f^* \subseteq \mathcal{X}_{f_c}^*$.

$\mathcal{X}_f^*$: set of the optimizers of f

convex envelope

convex envelope = tightest convex lower bound

Kleibohm, 1967

Let f_c be the convex envelope of $f : \mathcal{X} \to \mathbb{R}$.
Then (a) $\min_{x \in \mathcal{X}} f_c(x) = f^*$ and (b) $\mathcal{X}_f^* \subseteq \mathcal{X}_{f_c}^*$.

$\mathcal{X}_f^*$: set of the optimizers of f

problem: finding the **convex envelope** is hard!

idea behind CoRR

solution: approximate the convex envelope!

↓

$$h(x;\theta) = \langle \theta, \phi(x) \rangle$$

$$h(\cdot;\theta) \in \mathcal{H}, x \in \mathcal{X}, \theta \in \Theta$$

idea behind CoRR

solution: **approximate** the convex envelope!

↓

$$\min_{\theta} \; \mathbb{E}[d(h(x;\theta), f(x))]$$

idea behind CoRR

solution: **approximate** the convex envelope!

↓

$$\min_{\theta} \; \mathbb{E}[d(h(x;\theta), f(x))]$$

what is the right objective function??

the key to CoRR

Lemma 1

Assume that the convex envelope $f_c(x) \in \mathcal{H}$ and that $\mu = \mathbb{E}[f_c(x)]$

Then the convex approximation h(x) returned by (**P1**) will coincide with the convex envelope.

P1

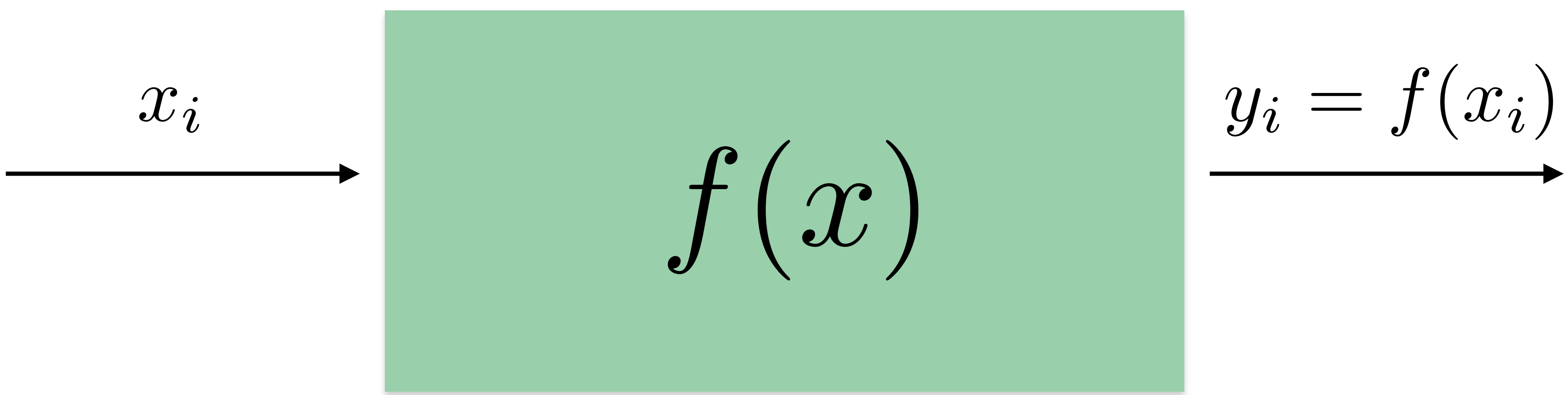
$$\theta_\mu = \arg\min_{\theta \in \Theta} \; \mathbb{E}[|h(x;\theta) - f(x)|] \;\; \text{s.t.} \;\; \mathbb{E}[h(x;\theta)] = \mu.$$

L1 error term

convex constraint

CoRR algorithm

Step 1: draw T samples from the function



black-box setting = **no gradient** information

CoRR algorithm

Step 2: find a convex approximation to the function

$$\widehat{\theta}_c = \arg\min_{\theta \in \Theta} \ \widehat{\mathbb{E}}_1\left[\left|h(x;\theta) - f(x)\right|\right] \ \text{ s.t. } \ \widehat{\mathbb{E}}_2\left[h(x;\theta)\right] = \mu$$

empirical expected loss

empirical constraint

CoRR algorithm

Step 2: fit a convex envelope to the function

$$\widehat{\theta}_c = \arg\min_{\theta \in \Theta} \ \widehat{\mathbb{E}}_1 \left[\left| h(x;\theta) - f(x) \right| \right] \ \text{ s.t. } \ \widehat{\mathbb{E}}_2 \left[h(x;\theta) \right] = \mu$$

Lemma 1 tells us how we should regularize this problem…

finding μ

1. Solve **Step 2** for fixed value of μ $\longrightarrow$ $h(x;\theta_\mu)$
2. Optimize the convex function h(x) $\longrightarrow$ $\widehat{x}_\mu$

$$\widehat{\mu} = \arg\min_{\mu\in[-R,R]} f(\widehat{x}_\mu)$$

guarantees

(Thm. 1) After T function evaluations, CoRR returns an estimate $\widehat{x}$ such that with probability $1 - \delta$

$$f(\widehat{x}) - f^* = \mathcal{O}\left[\left(\frac{\log(1/\delta)}{T}\right)^{\alpha}\right]$$

guarantees

(Thm. 1) After T function evaluations, CoRR returns an estimate $\widehat{x}$ such that with probability $1 - \delta$

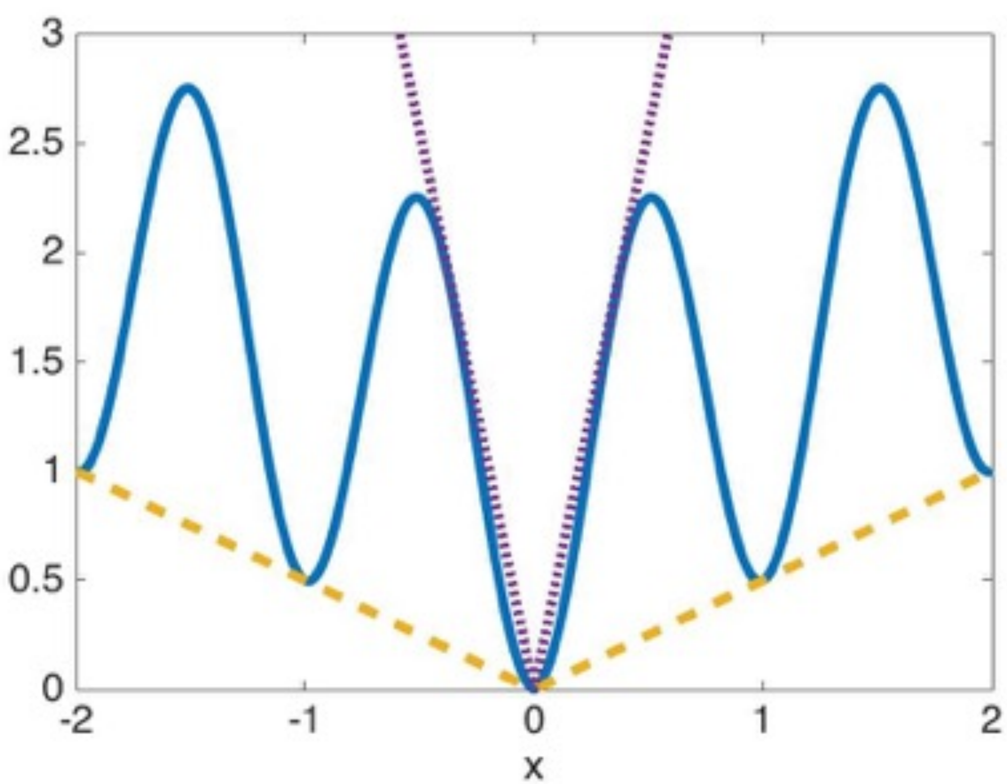
$$f(\widehat{x}) - f^* = \mathcal{O}\left[\left(\frac{\log(1/\delta)}{T}\right)^{\alpha}\right]$$

characterizes difficulty of problem

guarantees

(Thm. 1) After T function evaluations, CoRR returns an estimate $\widehat{x}$ such that with probability $1 - \delta$

$$f(\widehat{x}) - f^* = \mathcal{O}\left[\left(\frac{\log(1/\delta)}{T}\right)^{\alpha}\right]$$



what makes a problem easy:

1. smoothness around its minimum
2. upper and lower bound are matched

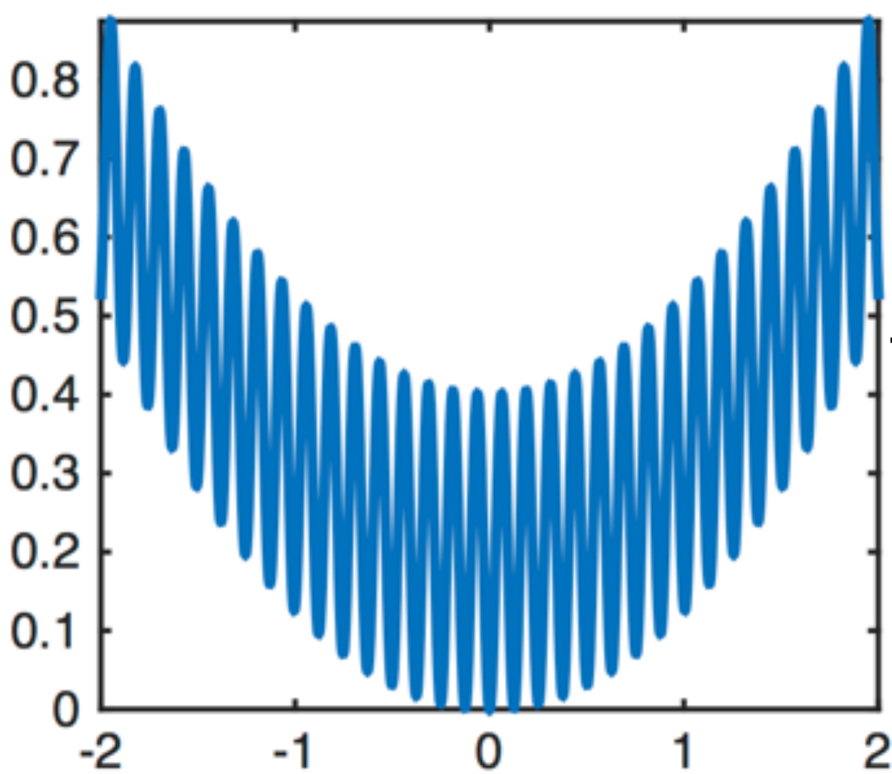
$\Downarrow$

$\alpha = 1/2$

guarantees

(Thm. 1) After T function evaluations, CoRR returns an estimate $\widehat{x}$ such that with probability $1 - \delta$

$$f(\widehat{x}) - f^* = \mathcal{O}\left[\left(\frac{\log(1/\delta)}{T}\right)^{\alpha}\right]$$



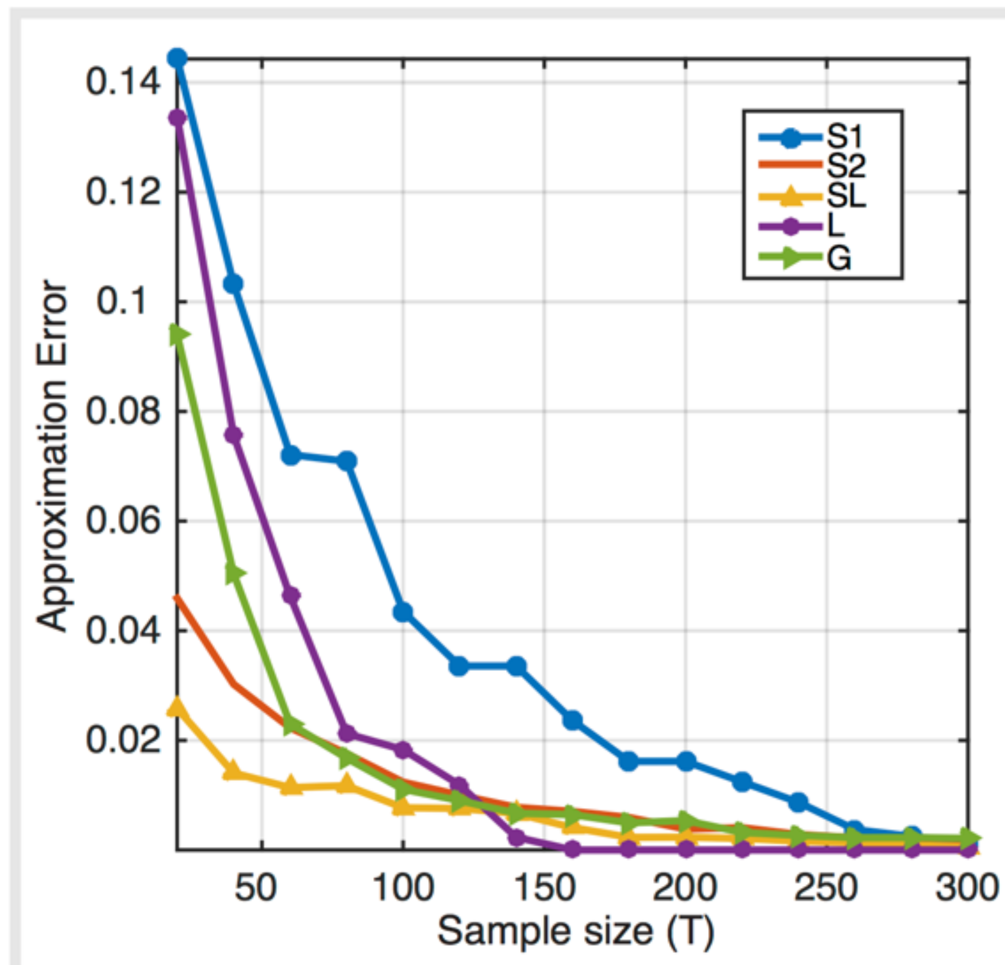
what makes a problem difficult:

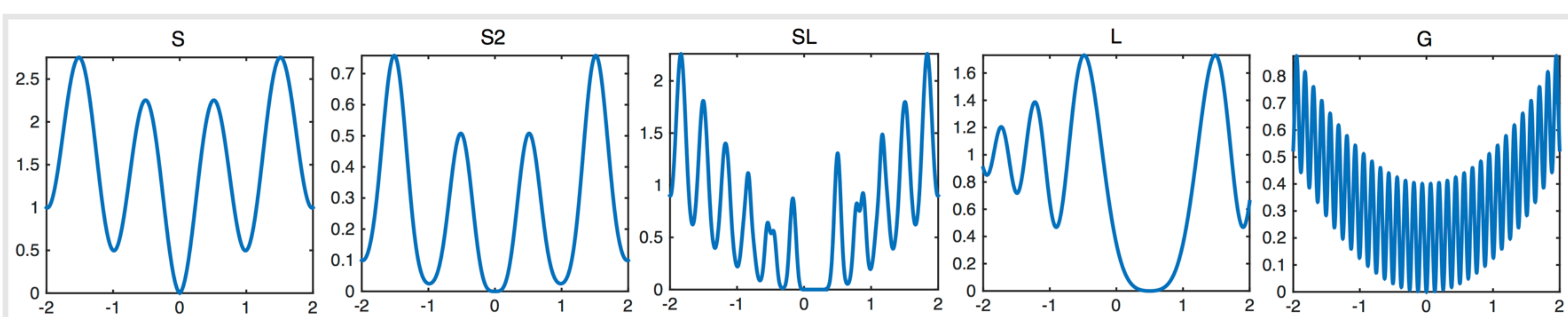
1. fast growth around its minimum
2. needle in the haystack!

$\Downarrow$

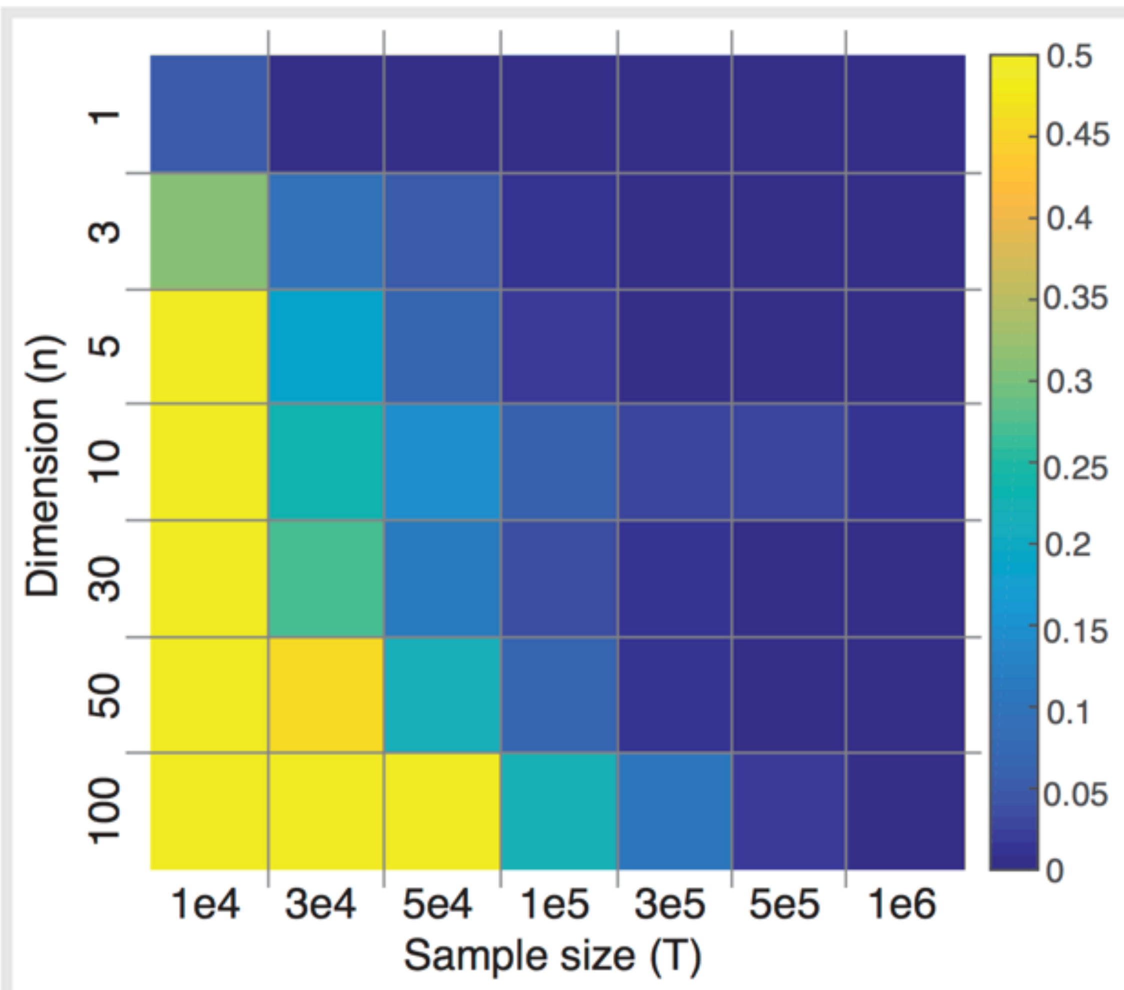
$\alpha \to 0$

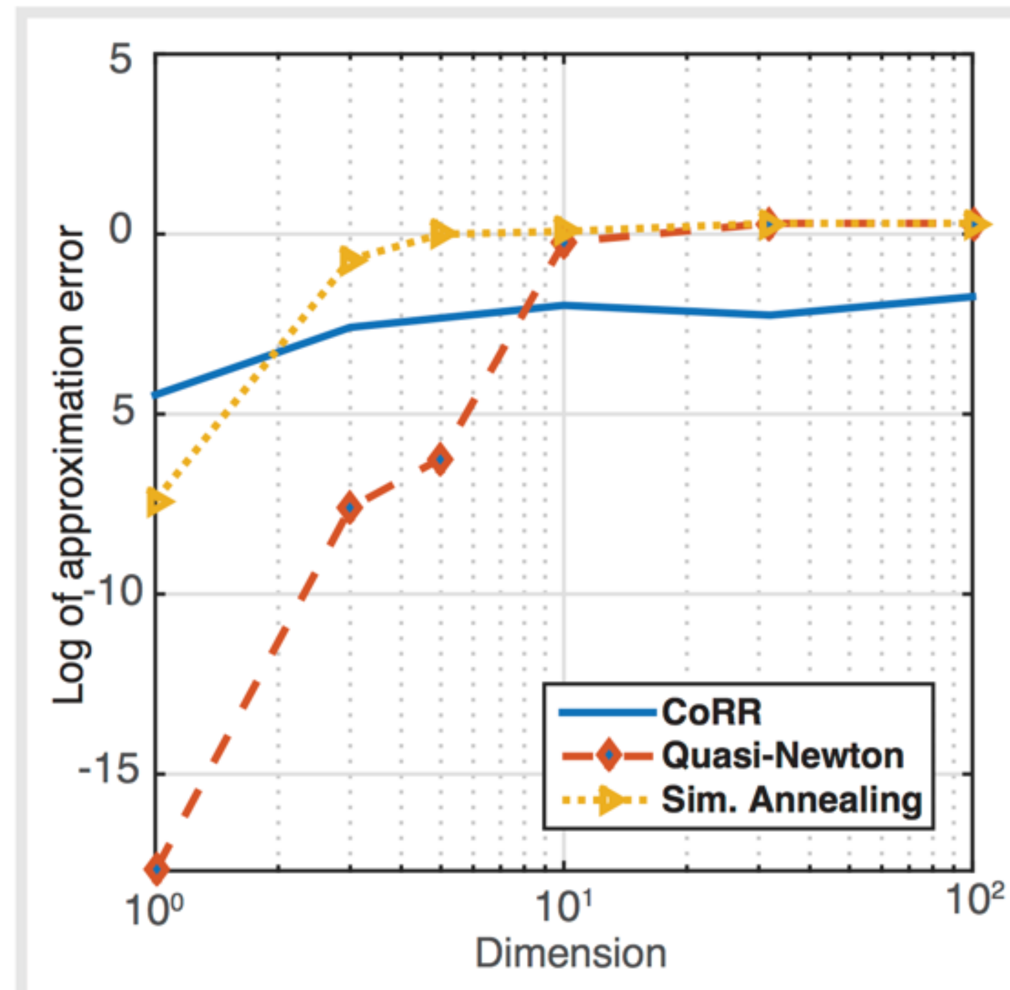
numerical results



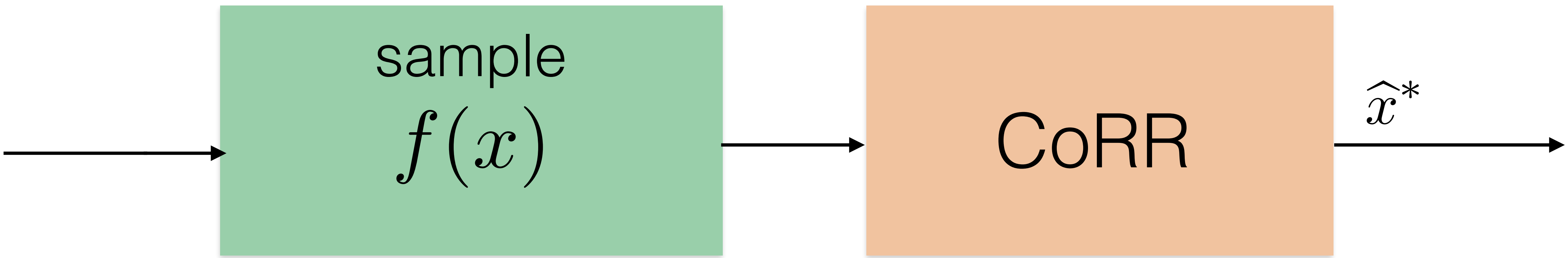


numerical results



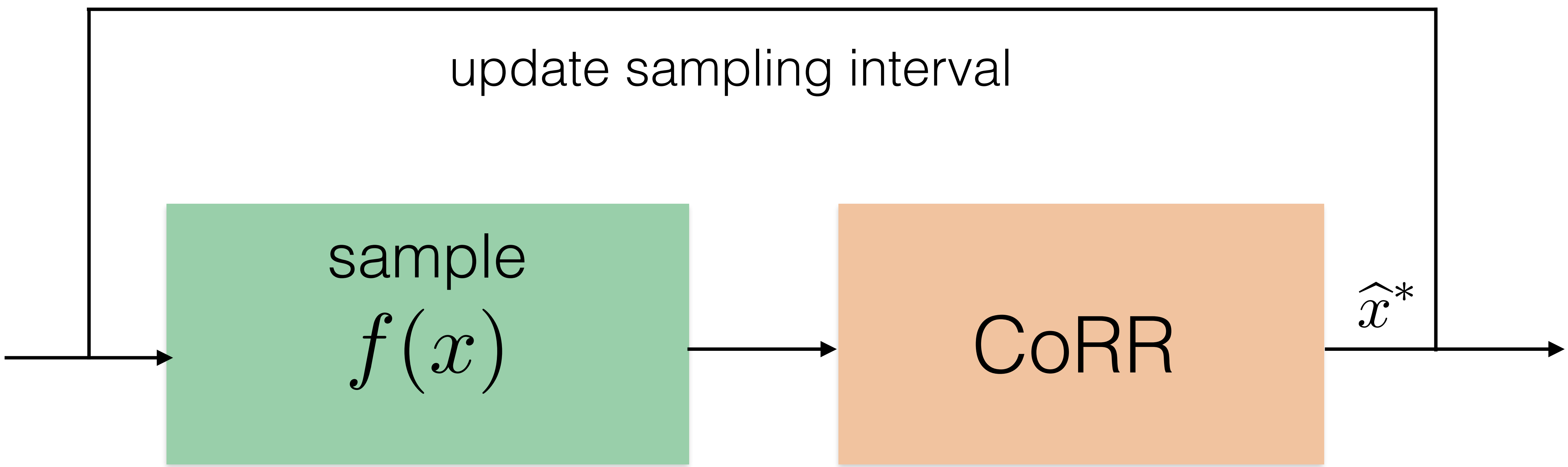


extensions to CoRR



CoRR = fixed set of samples (one-shot)

extensions to CoRR



adaCoRR = iteratively select samples around $\widehat{x}^*$

take-home message…

new approach for solving non-convex problems

CoRR = Convex Relaxation Regression

1. learn a convex approximation h(x) from black-box samples of f(x)
2. optimize a convex surrogate instead of f(x)

the team

Mohammad Azar
(Google DeepMind)

Konrad Körding
(RIC, Northwestern)

thank you!

questions?