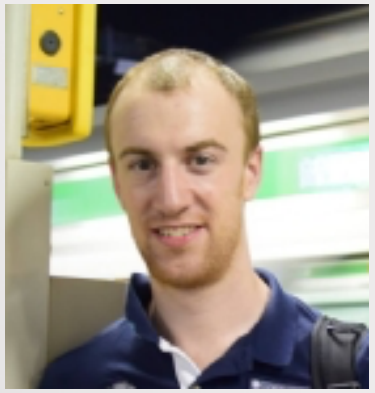
Uber AI

Scalable Global Optimization via Local Bayesian Optimization

David Eriksson
Uber AI
eriksson@uber.com

Michael Pearce



Jake Gardner

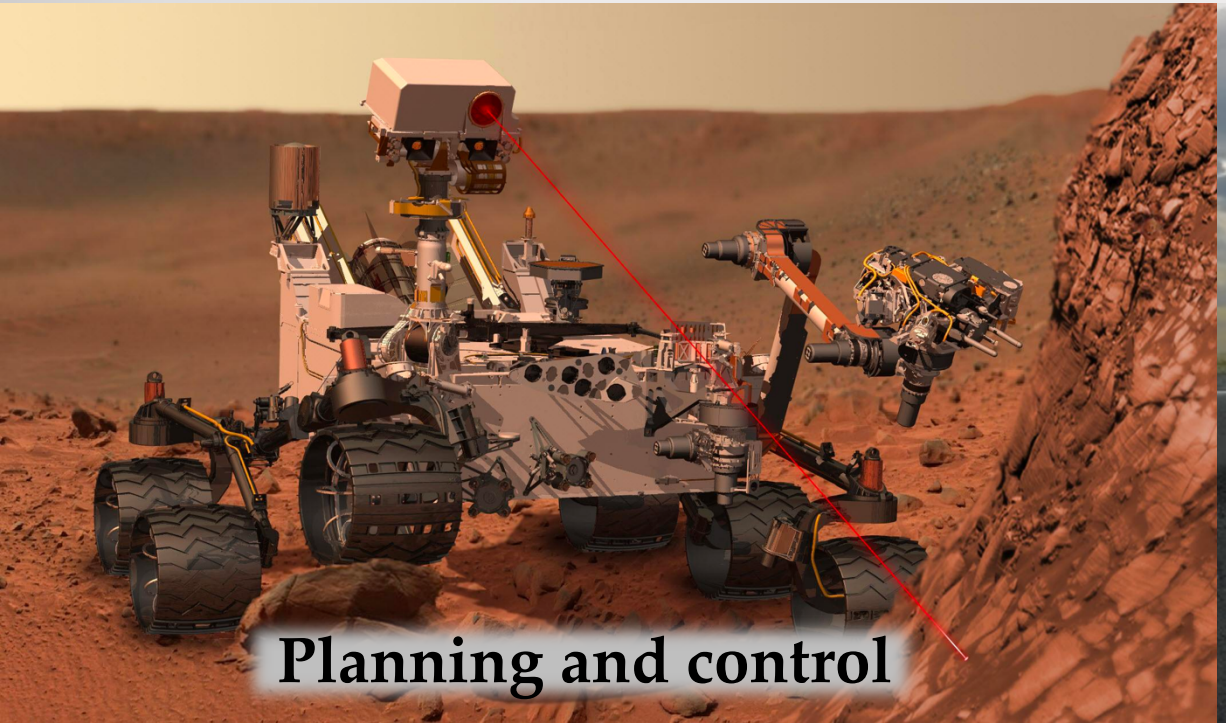
Ryan Turner

Matthias Poloczek

Global Optimization

$$\text{Find } x^* \in \Omega \text{ such that } f(x^*) \leq f(x), \quad \forall x \in \Omega$$

- f is a continuous, computationally expensive, black-box function
- $\Omega \subset \mathbb{R}^d$ is a hyper-rectangle



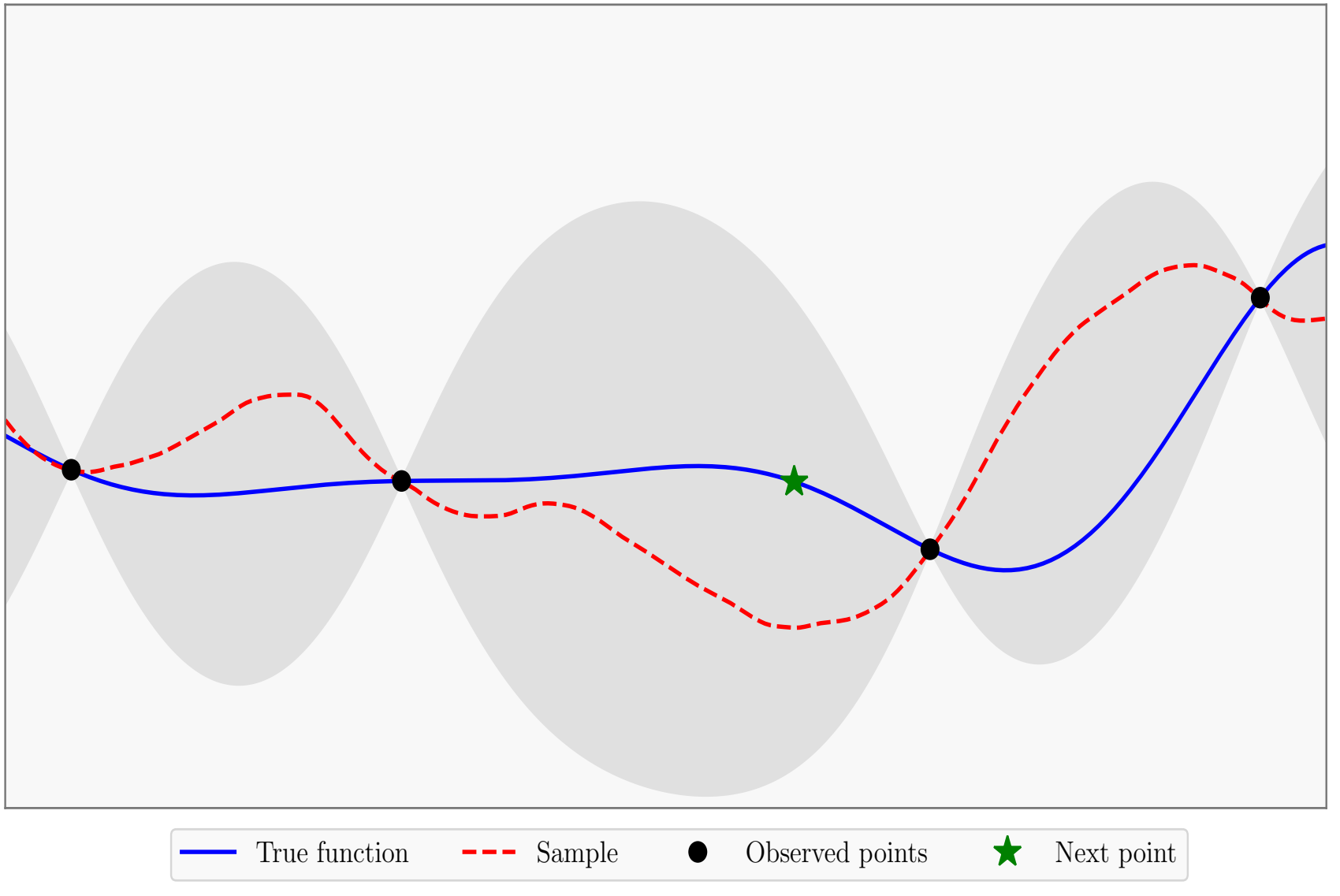
Planning and control

Design of aerodynamic structures

Bayesian Optimization (BO)

Common restrictions:

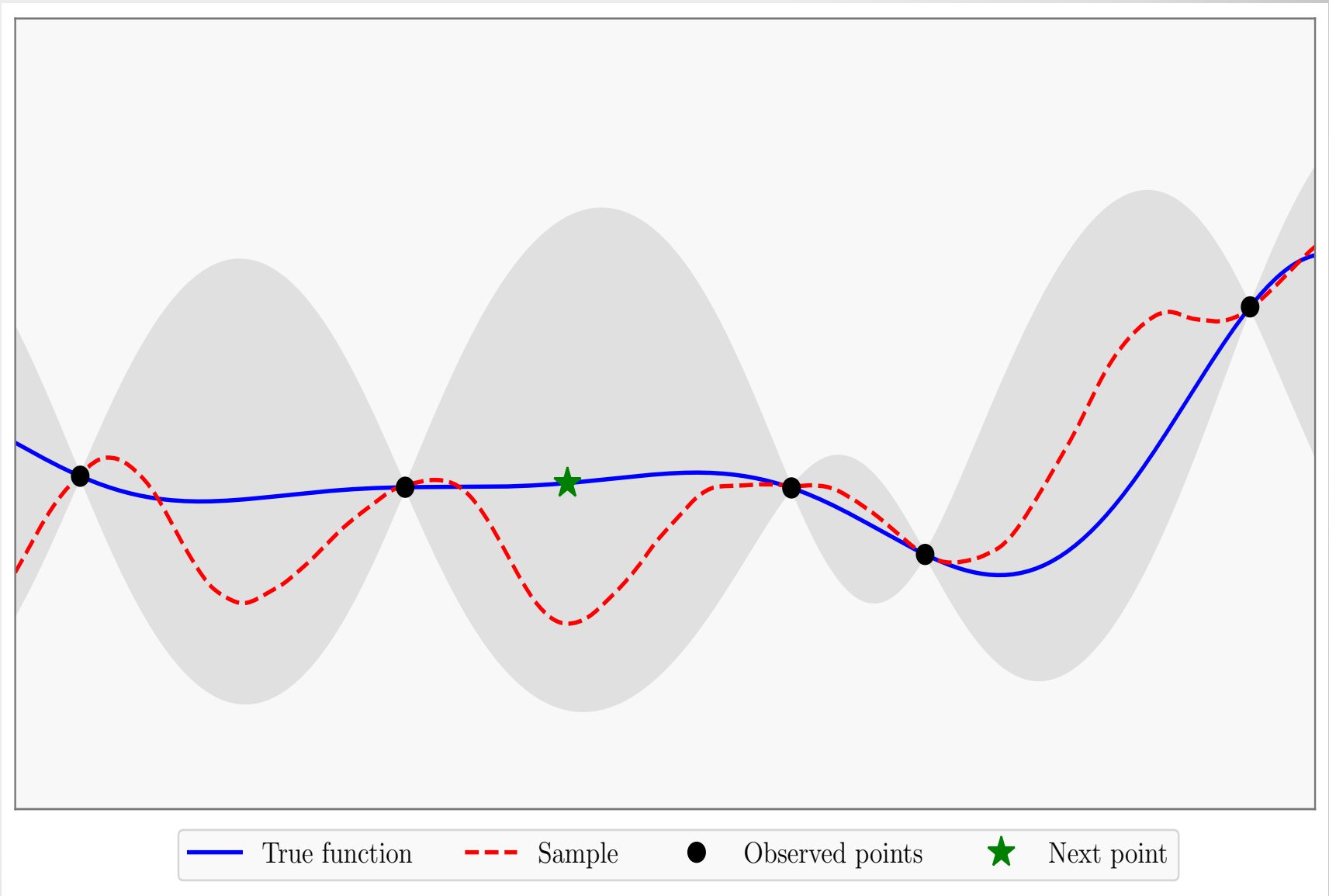
- A few hundred evaluations
- Less than 10 tunable parameters



Bayesian Optimization (BO)

Common restrictions:

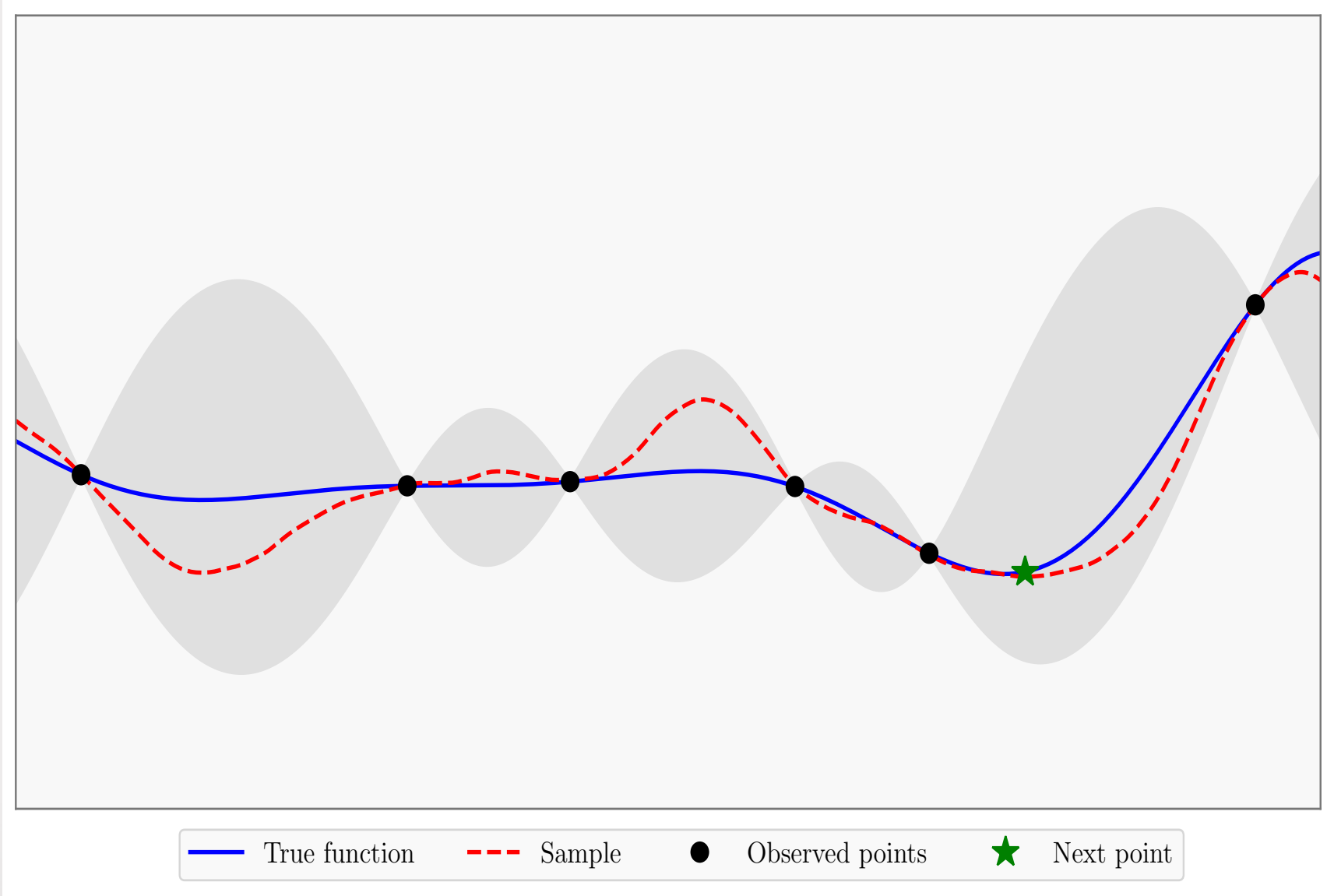
- A few hundred evaluations
- Less than 10 tunable parameters



Bayesian Optimization (BO)

Common restrictions:

- A few hundred evaluations
- Less than 10 tunable parameters



High-dimensional BO is challenging

Challenges:

1. The search space grows exponentially with dimensionality
2. A global GP model may not fit the data everywhere
3. Large areas of uncertainty leads to over-exploration

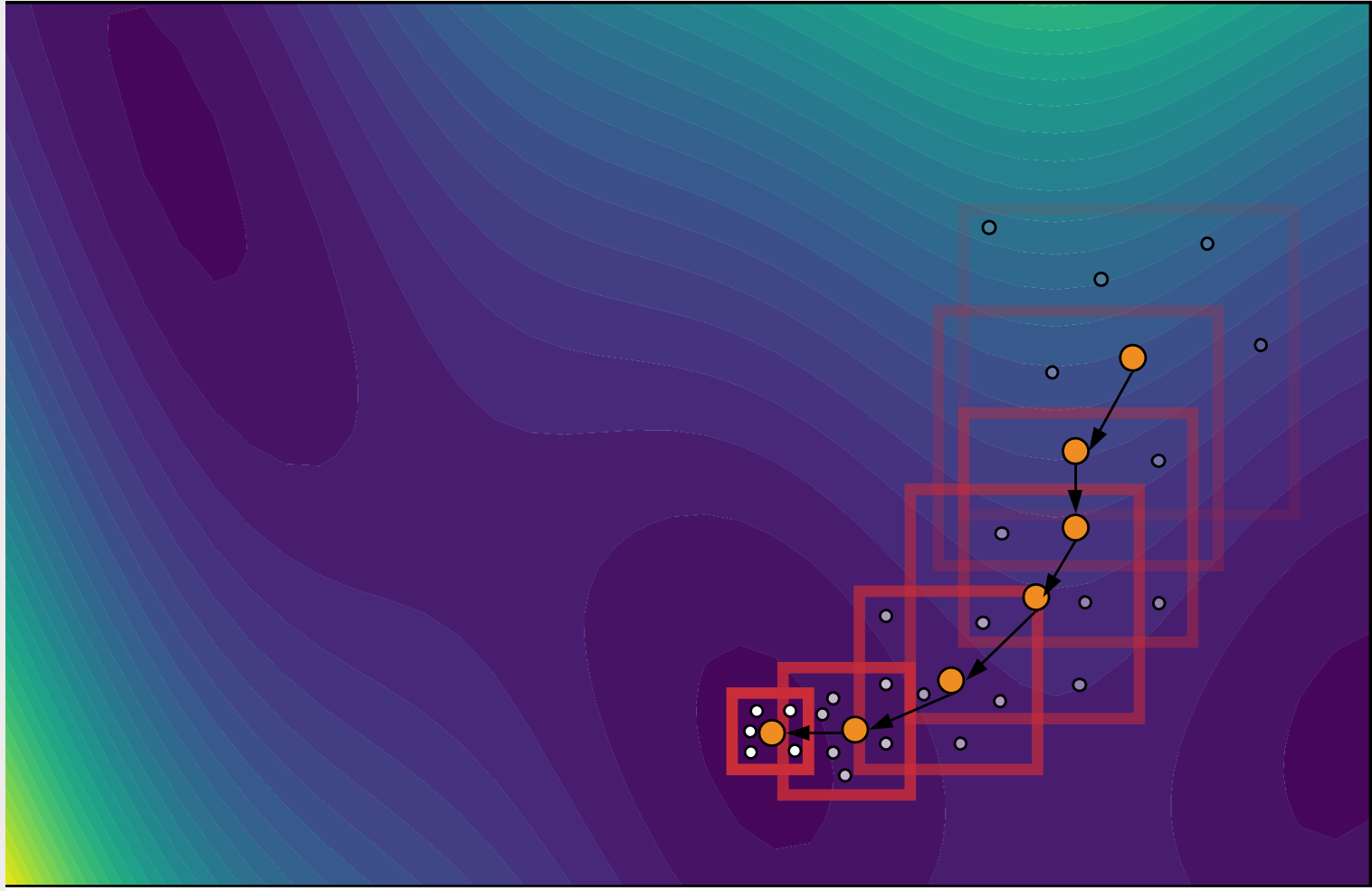
Previous work makes **strong assumptions**:

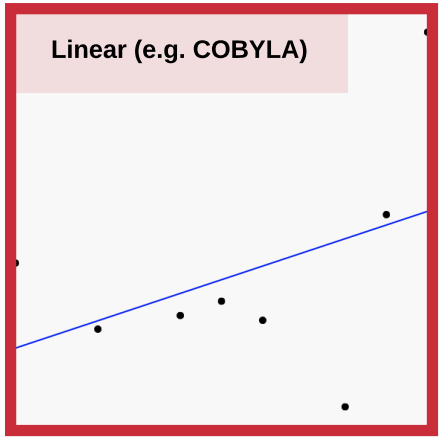
- Additive structure
- Low-dimensional structure

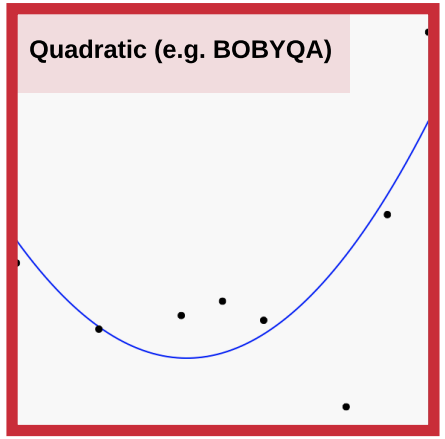
Trust-region methods

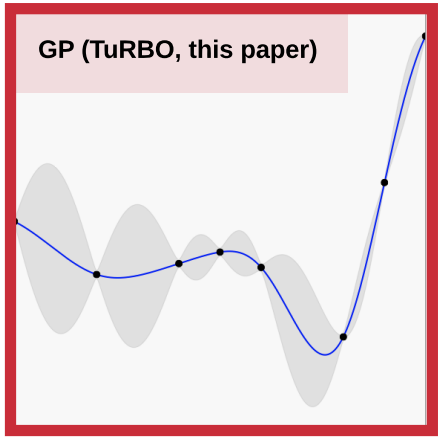
Main idea:

- Optimize a (simple) model in a local region
- Expand/shrink this region based on progress
- Only requires a locally accurate model



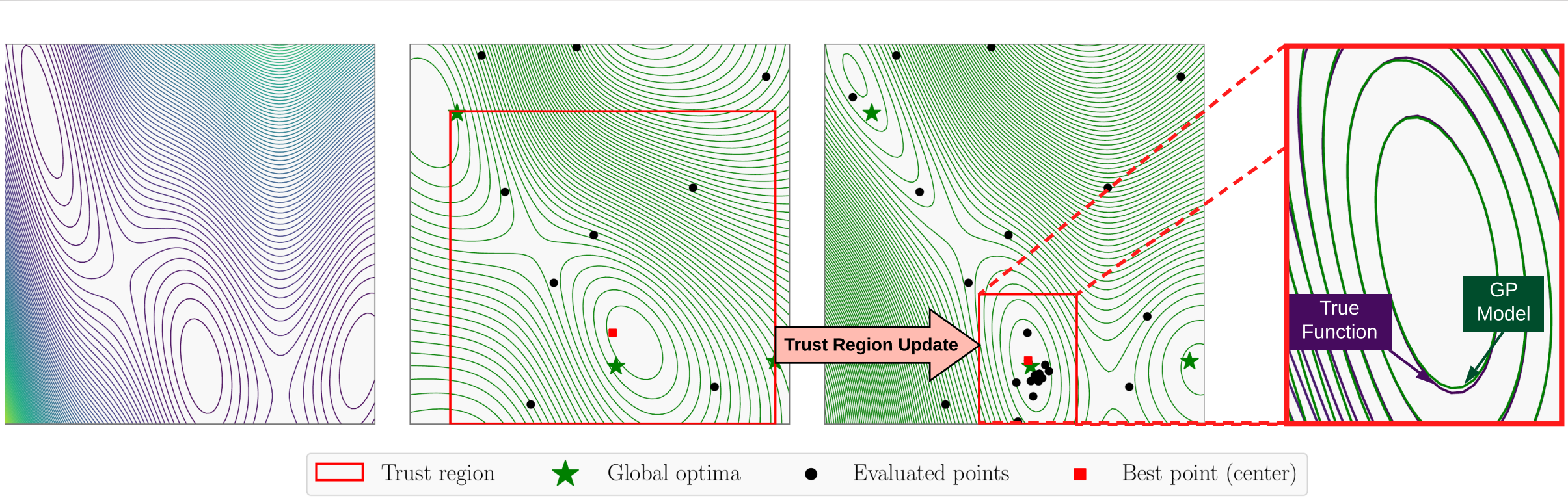






Trust-region BO (TuRBO)

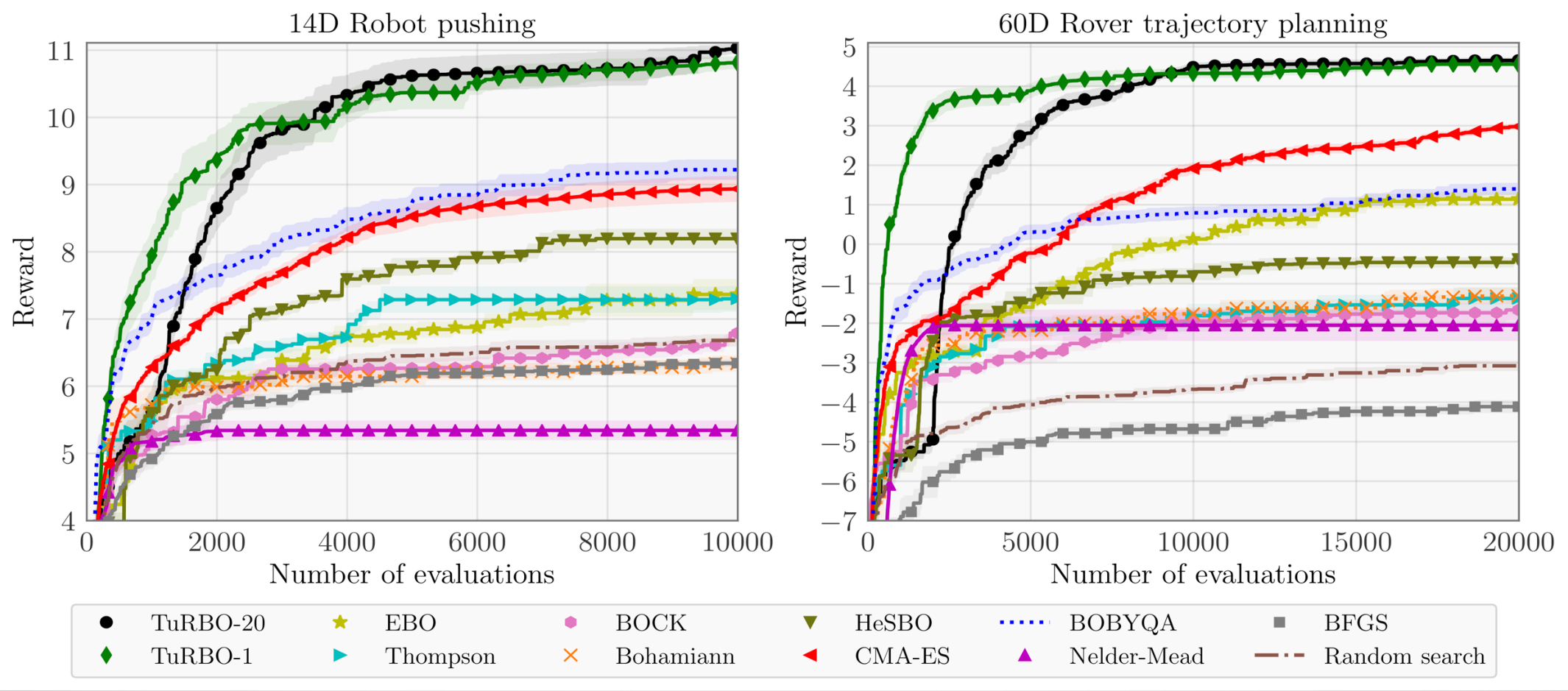
1. Avoids over-exploration by using a trust-region framework
2. Balances exploration/exploitation by using BO inside the trust-region
3. Uses Thompson sampling to scale to large batch sizes



Experimental results

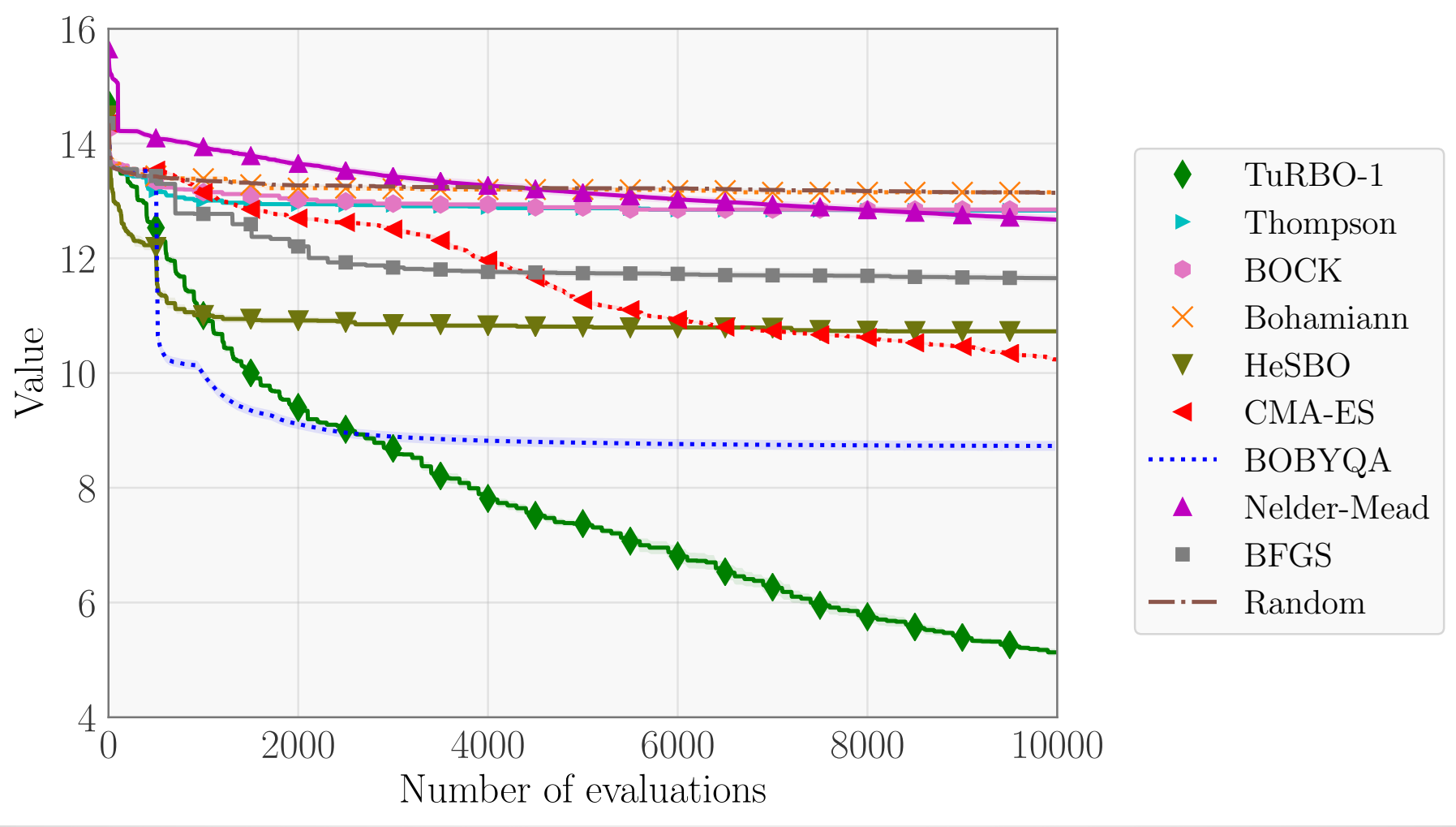
Robot pushing: 10,000 evaluations, batch size 50

Rover trajectory planning: 20,000 evaluations, batch size 100



Experimental results

200D Ackley function: 10,000 evaluations, batch size 100



Summary

TuRBO:

- Achieves excellent results for high-dimensional problems
- Combines BO with trust-regions to avoid over-exploration
- Makes no assumptions about low-dimensional structure

Paper: https://arxiv.org/abs/1910.01739

Code: https://github.com/uber-research/TuRBO

Poster #9