We Still Don't Understand High-Dimensional Bayesian Optimization

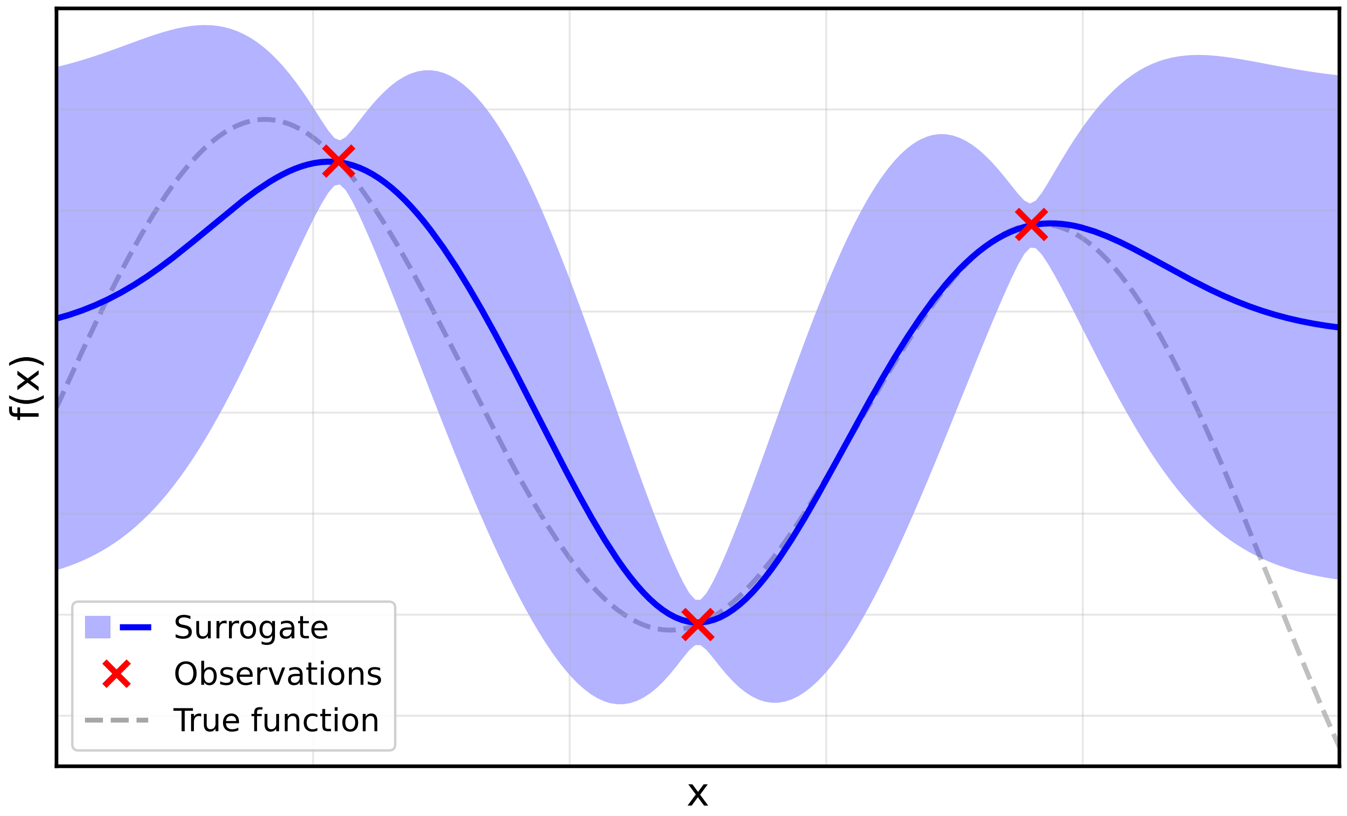
Colin Doumont, **Donney Fan,** Natalie Maus, Jacob R. Gardner, Henry Moss, Geoff Pleiss

Best Student Paper Award, AISTATS 2026

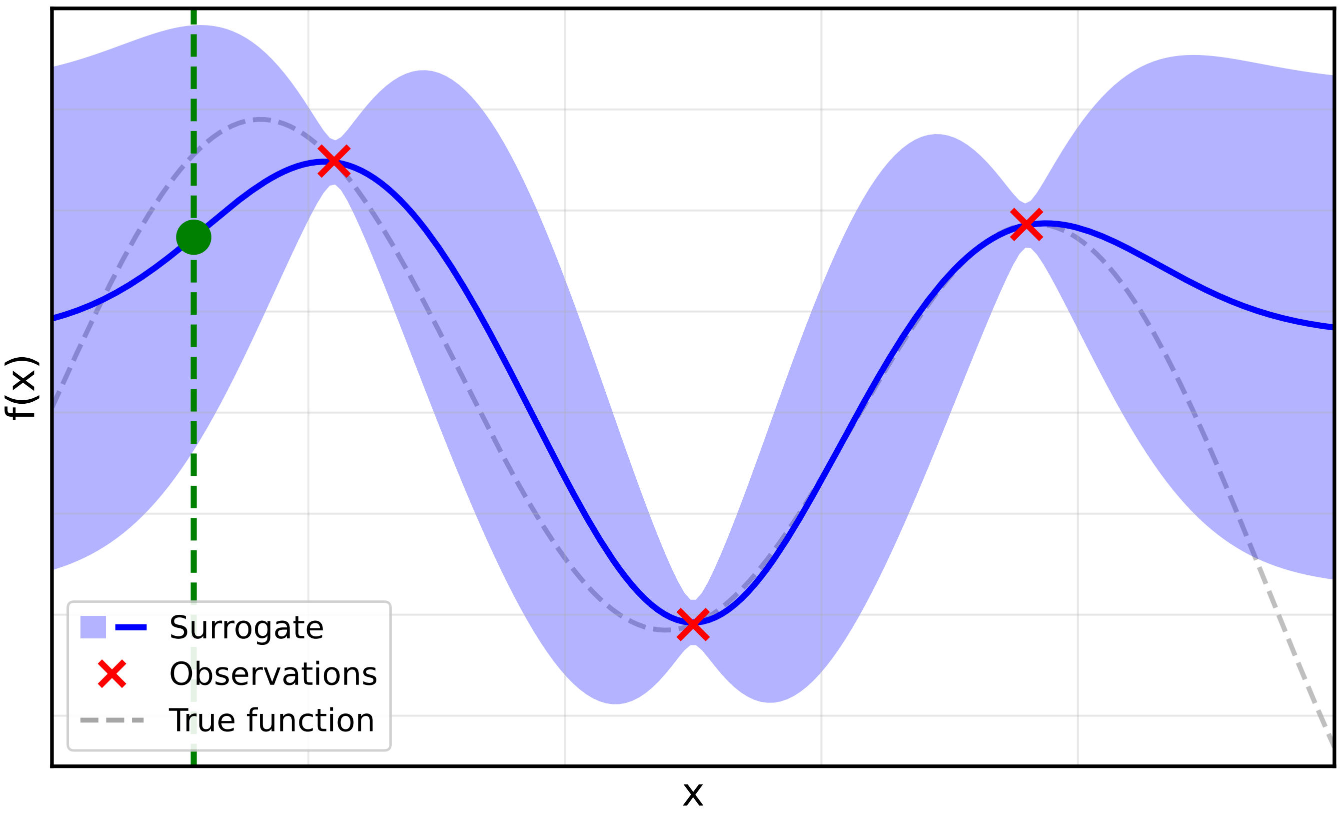
28 May 2026

VECTOR INSTITUTE

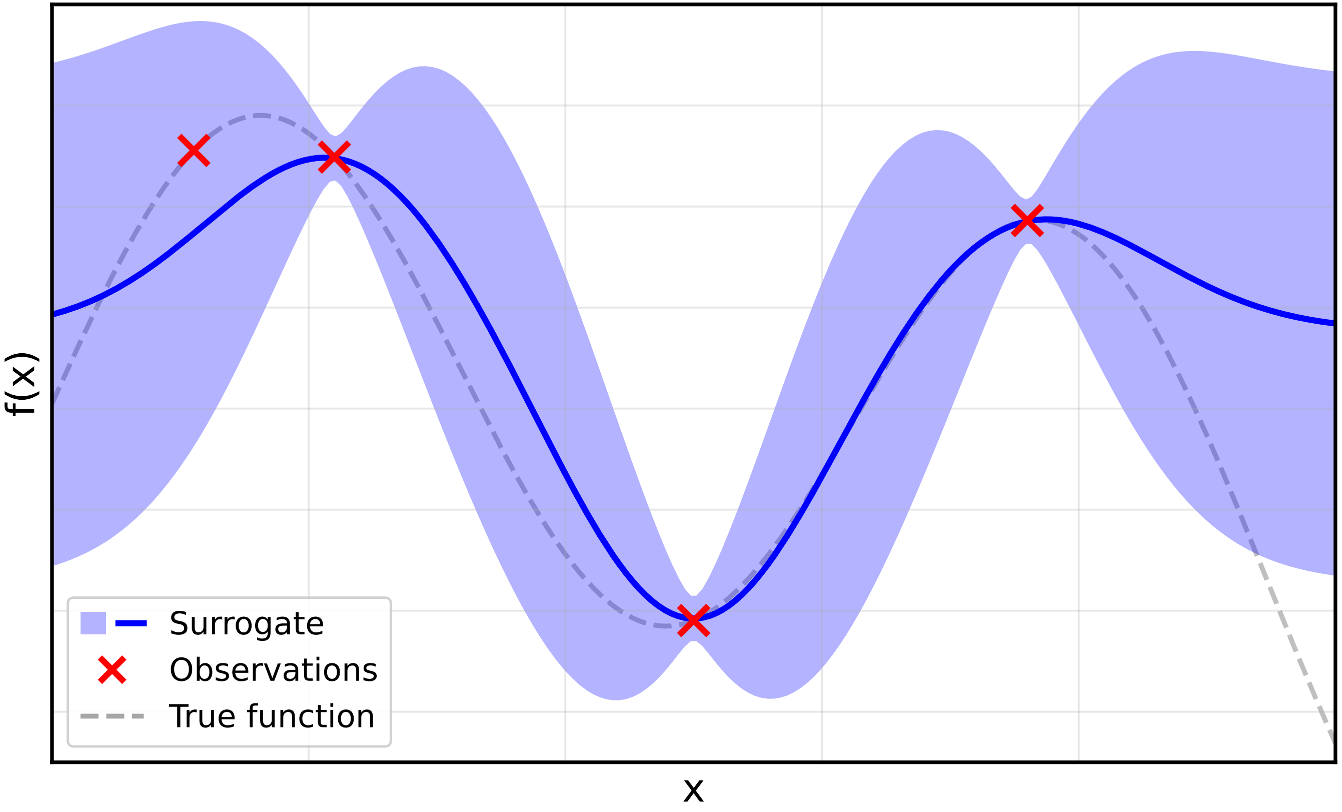
Bayesian optimization



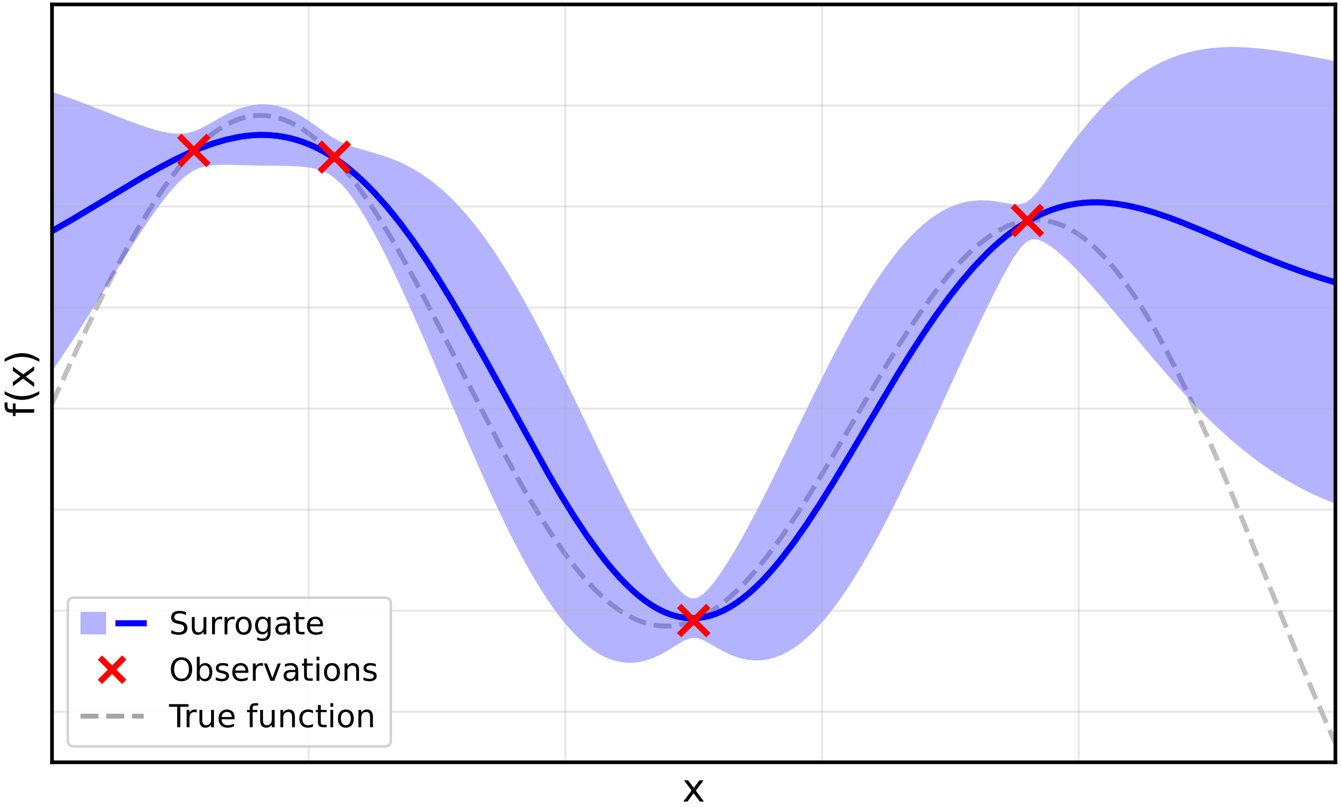
Bayesian optimization



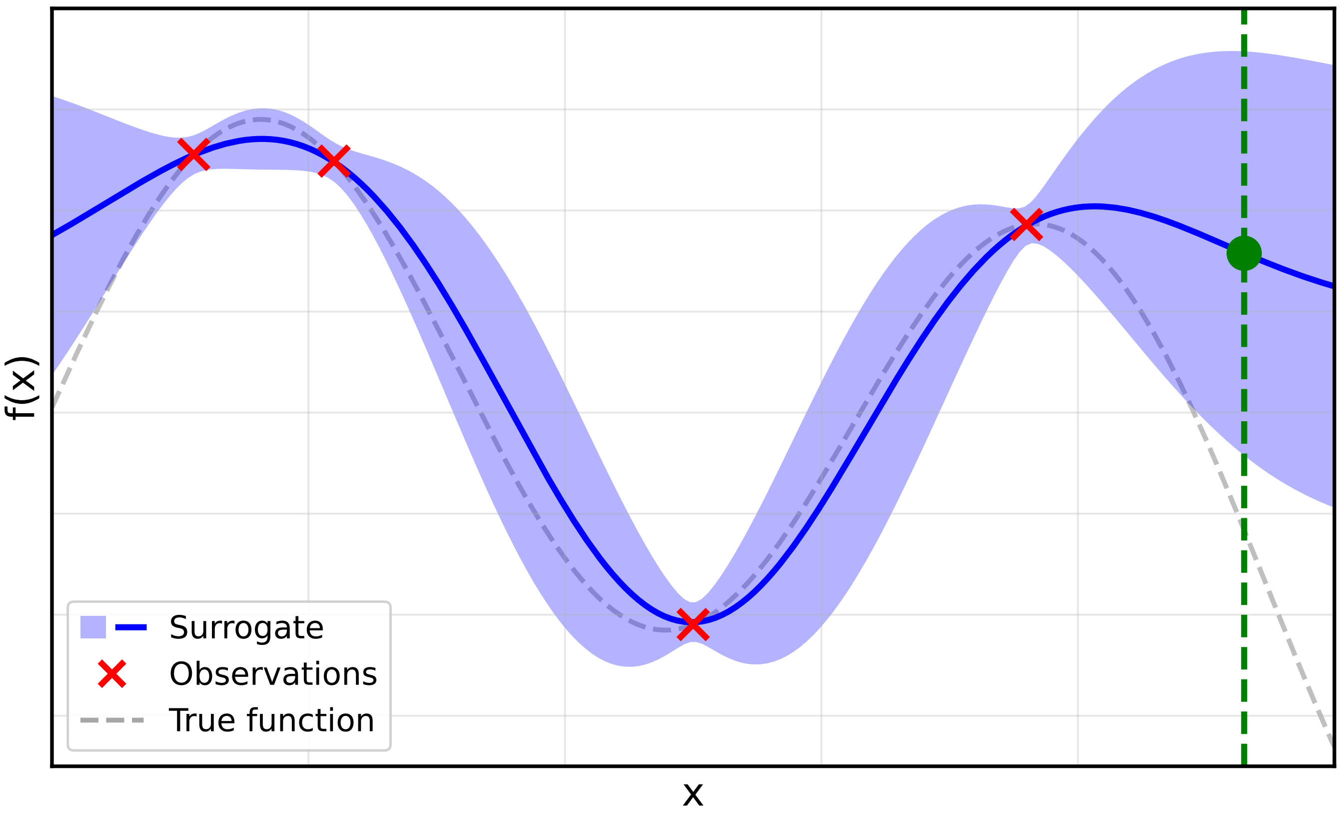
Bayesian optimization



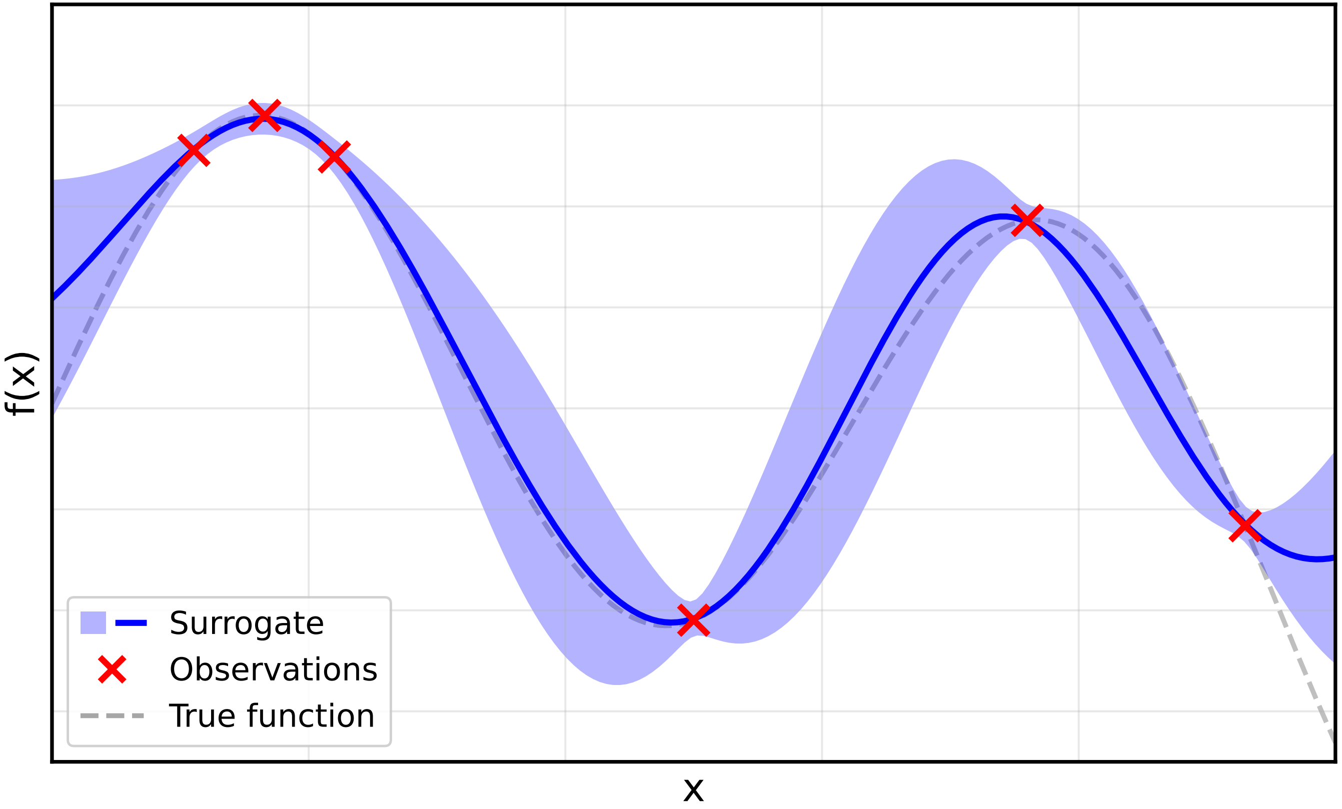
Bayesian optimization



Bayesian optimization



Bayesian optimization



Curse of dimensionality — problem

GP-UCB simple regret (Srinivas et al., 2010):

$$r_T = \mathcal{O}\left(\sqrt{\frac{\log(T)^{D+1}}{T}}\right)$$

Achieving ε-accuracy requires T growing **exponentially** in D!

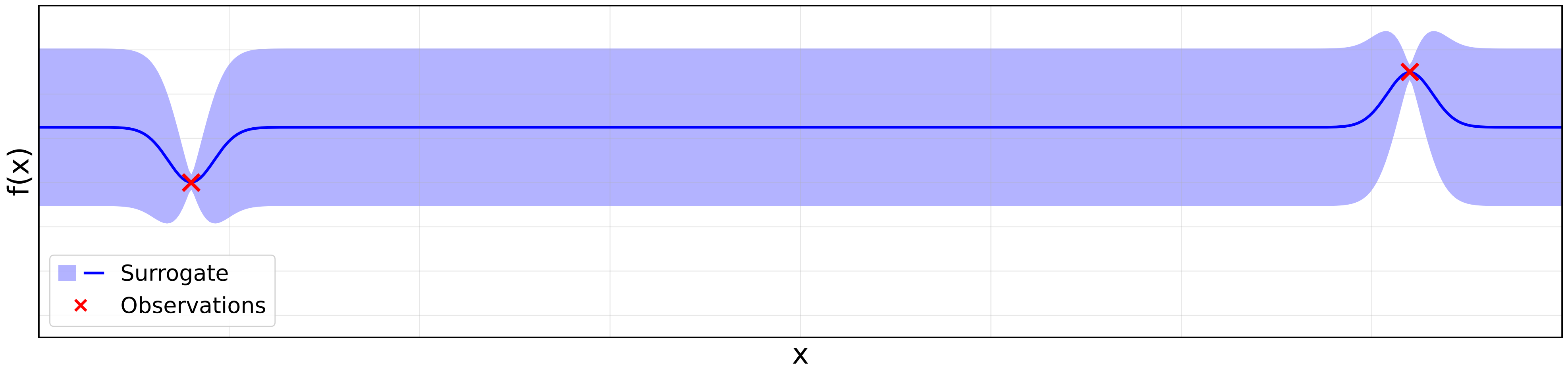
Curse of dimensionality — solutions

1. Structural assumptions:
    - trust regions
    - random embeddings
    - variable selection
    - …

2. Non-linear surrogate

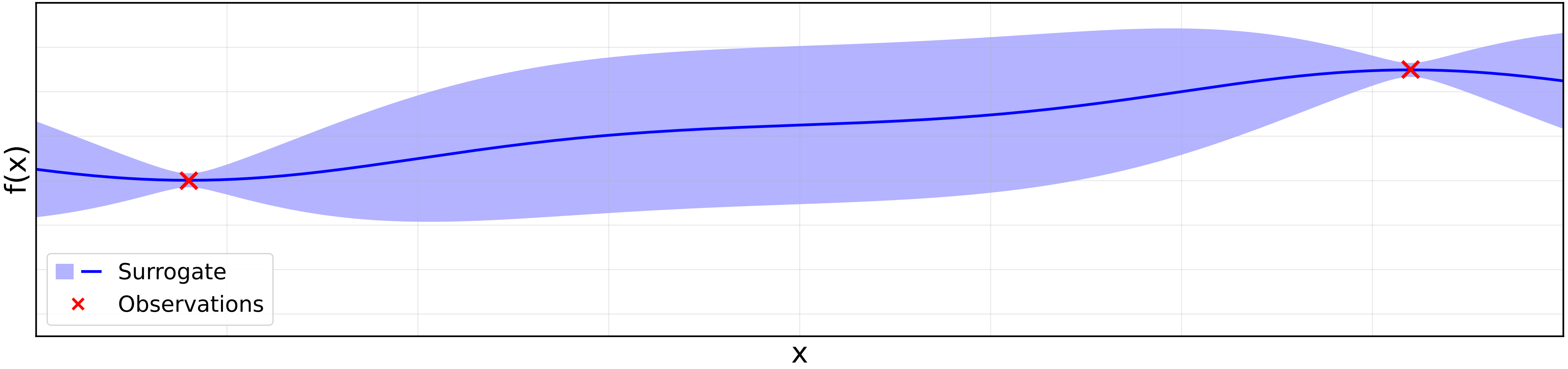
Smoother surrogates ("Vanilla BO")

$$\mathrm{Cov}[f(\mathbf{x}), f(\mathbf{x}')] \propto \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2}\right)$$



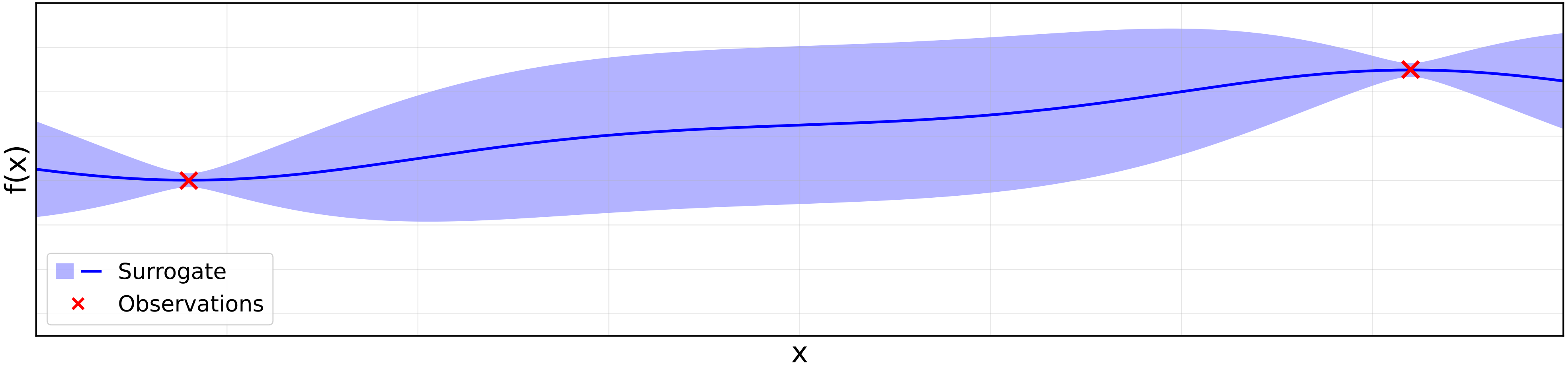
Smoother surrogates ("Vanilla BO")

$$\mathrm{Cov}[f(\mathbf{x}), f(\mathbf{x}')] \propto \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2D}\right)$$



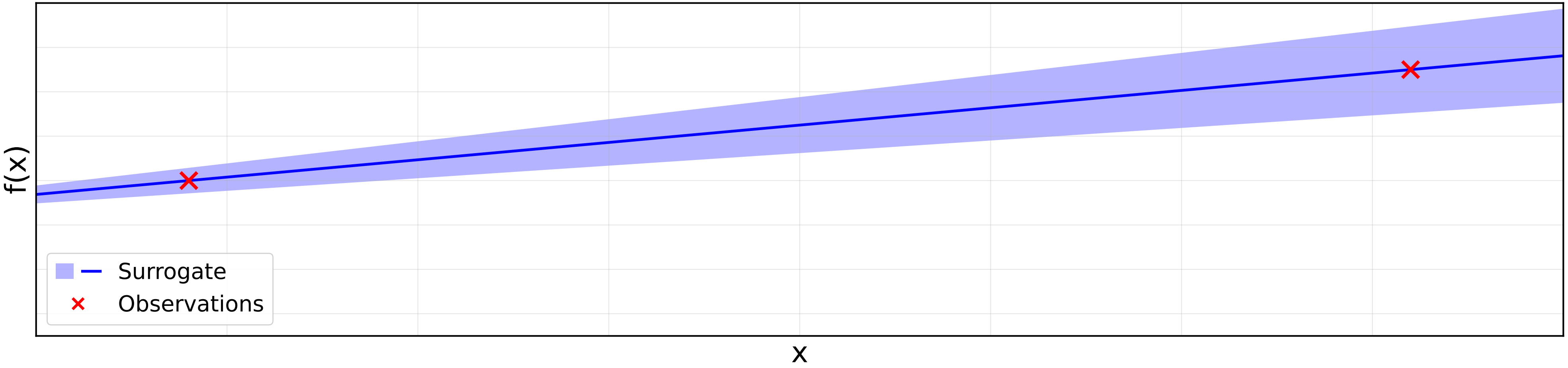
Pushing smoothness to the extreme

What if we replace the non-linear surrogate...



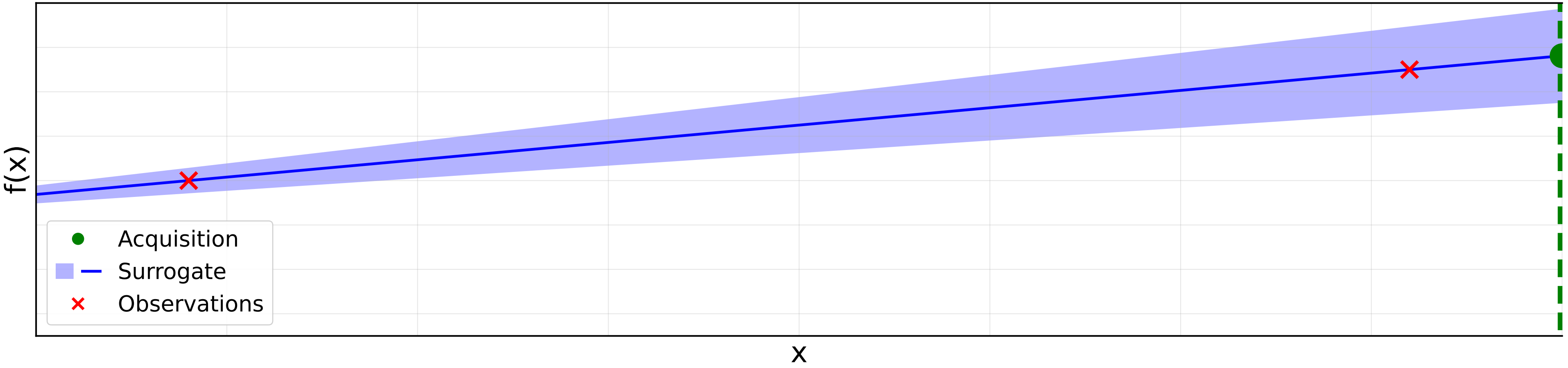
Pushing smoothness to the extreme

…by a (Bayesian) linear regression?

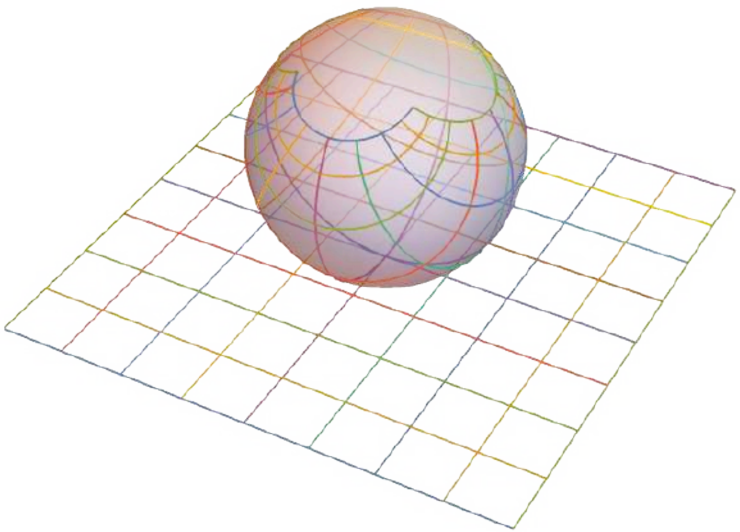


Pushing smoothness to the extreme

We always obtain acquisitions on the boundary!



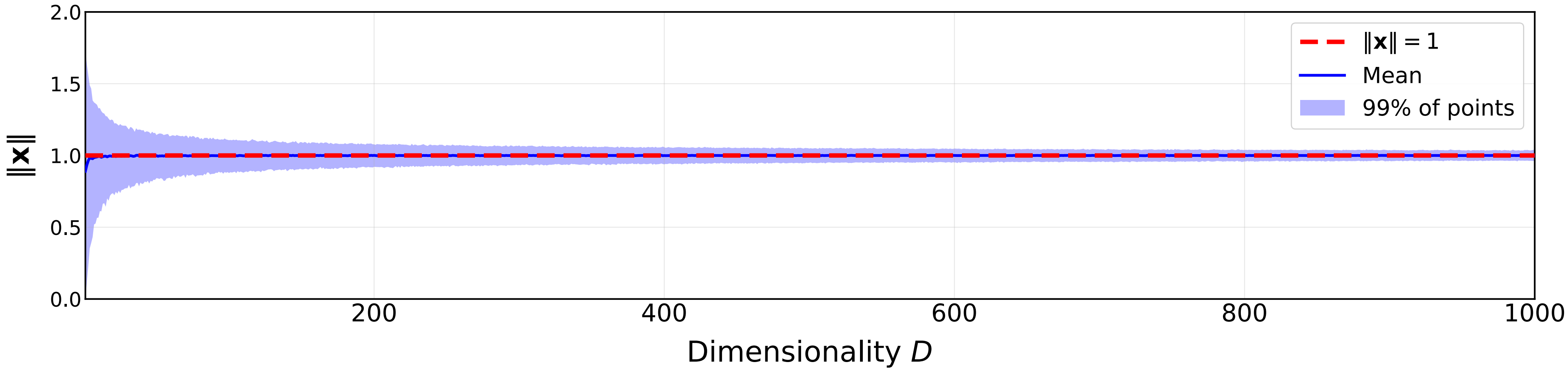
Spherical projection prevents boundary acquisitions



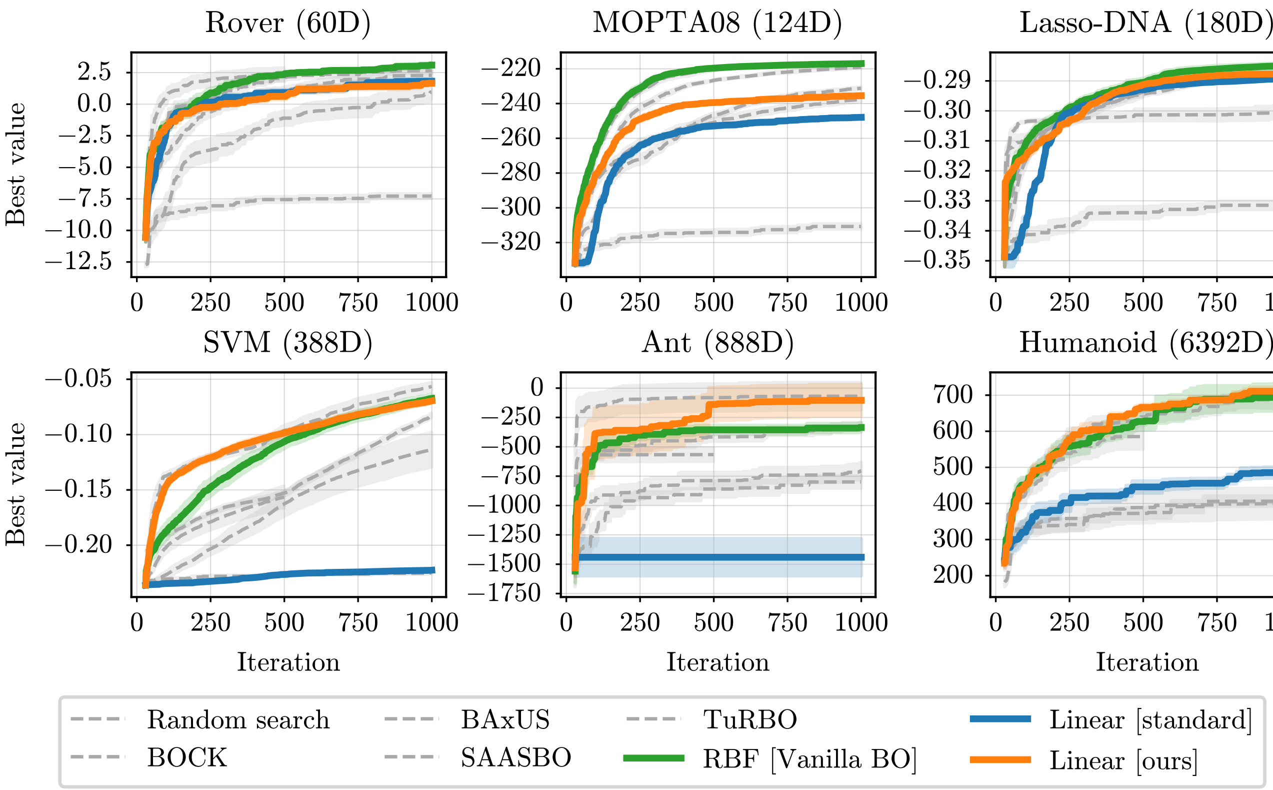
$$\underbrace{\mathbf{x}}_{[-1,1]^D} \rightarrow \underbrace{P(\mathbf{x})}_{\mathbb{S}^D} = \frac{1}{\|\mathbf{x}\|^2 + 1}\left[2x_1,\ \cdots,\ 2x_D,\ \|\mathbf{x}\|^2 - 1\right]$$

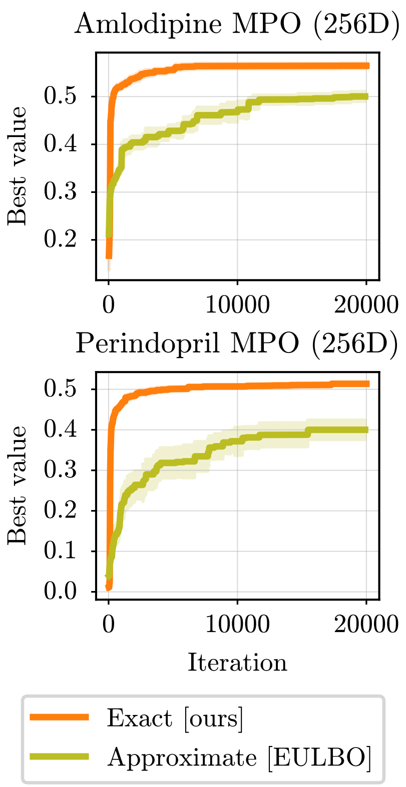
Spherical projection in high dimensions

In high D, $\|\mathbf{x}\| \approx 1 \Longrightarrow P(\mathbf{x}) \approx [\mathbf{x}; 0]$ (projection is nearly a no-op)



Empirical results





Why are these results so surprising?

To achieve state-of-the-art in high-dimensional BO, we need:

- ~~structural assumption + non-linear surrogate~~
- ~~non-linear surrogate*~~
- linear surrogate* (this talk)

What's next?

Less:

- inventing new (complicated) structural-assumption methods

More:

- understanding fundamental mechanisms driving HDBO

Thank you!

Colin Doumont

Donney Fan

Natalie Maus

Jacob R. Gardner

Henry Moss

Geoff Pleiss