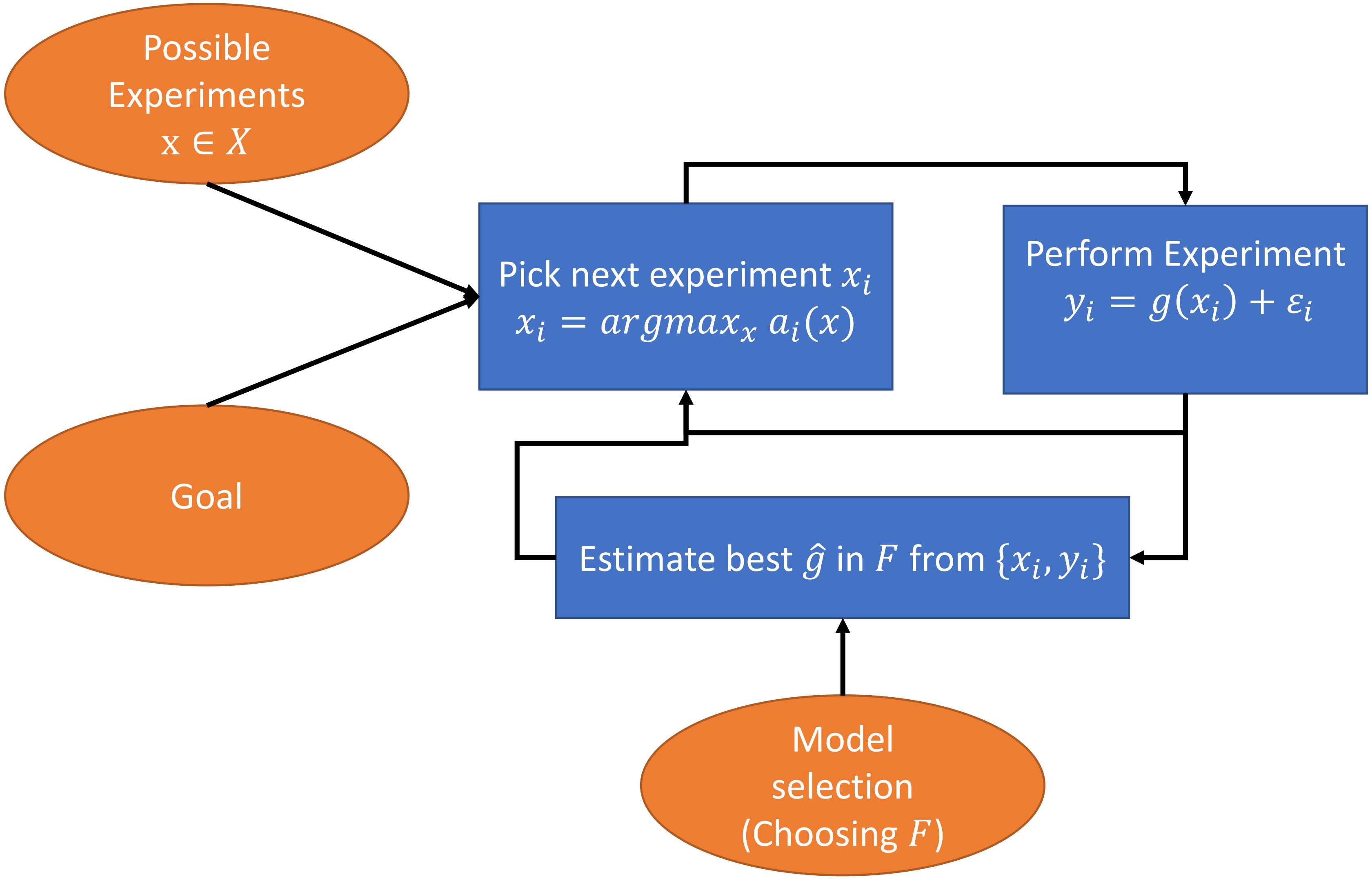
Optimization for Knowledge Discovery

Modern design of statistical experiments

Mojmír Mutný, ETH Zürich

Joint work with Andreas, Johannes and others

Experiment Design



- How many iterations do we need to finish with the goal?
- What is the right goal?
- How to select F?
- How to be robust to noise?
- Can we efficiently pick the experiment?
- How to select the right F?

In one sentence:

"How to make sequential decision making scale to more realistic and modern model assumptions and requirements?"

Unknown Model

Sequential hypothesis testing

Complex models beyond GPs

MKL, neural networks

New Tasks

Subspace estimation, level sets

New noise models

Poisson sensing, heavy tail bandits, likelihood

Scalable Assumptions

Additive models, subspaces, Projection pursuit

Efficient Algorithms

LineBO, Batching, Sketching

In one sentence:

"How to make sequential decision making scale to more realistic and modern model assumptions and requirements?"

Unknown Model

Sequential hypothesis testing

Complex models beyond GPs

MKL, neural networks

New Tasks

Subspace estimation, level sets

New noise models

Poisson sensing, heavy tail bandits, likelihood

Scalable Assumptions

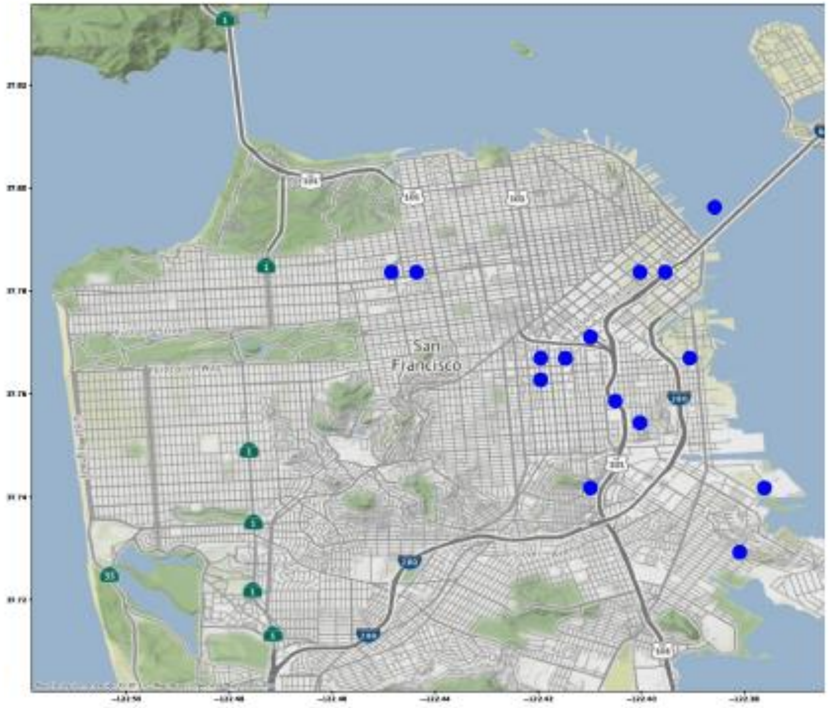
Additive models, subspaces, Projection pursuit

Efficient Algorithms

LineBO, Batching, Sketching

Tolerance to Noise: Poisson Sensing

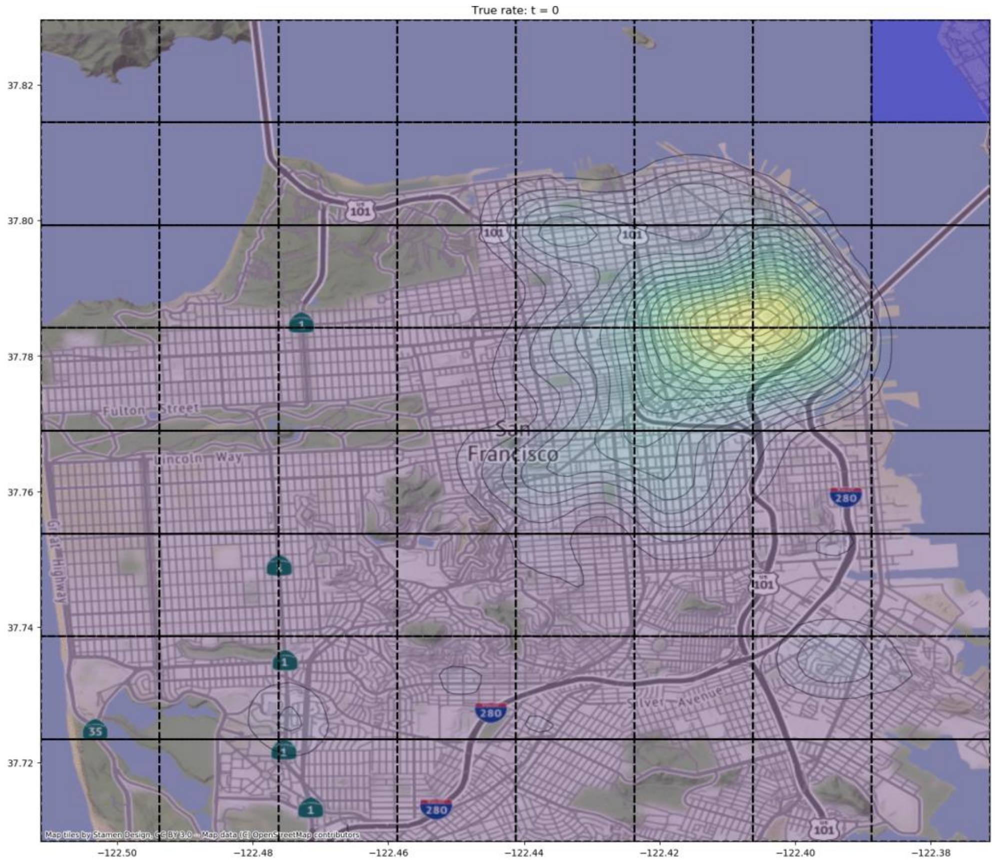
New noise models

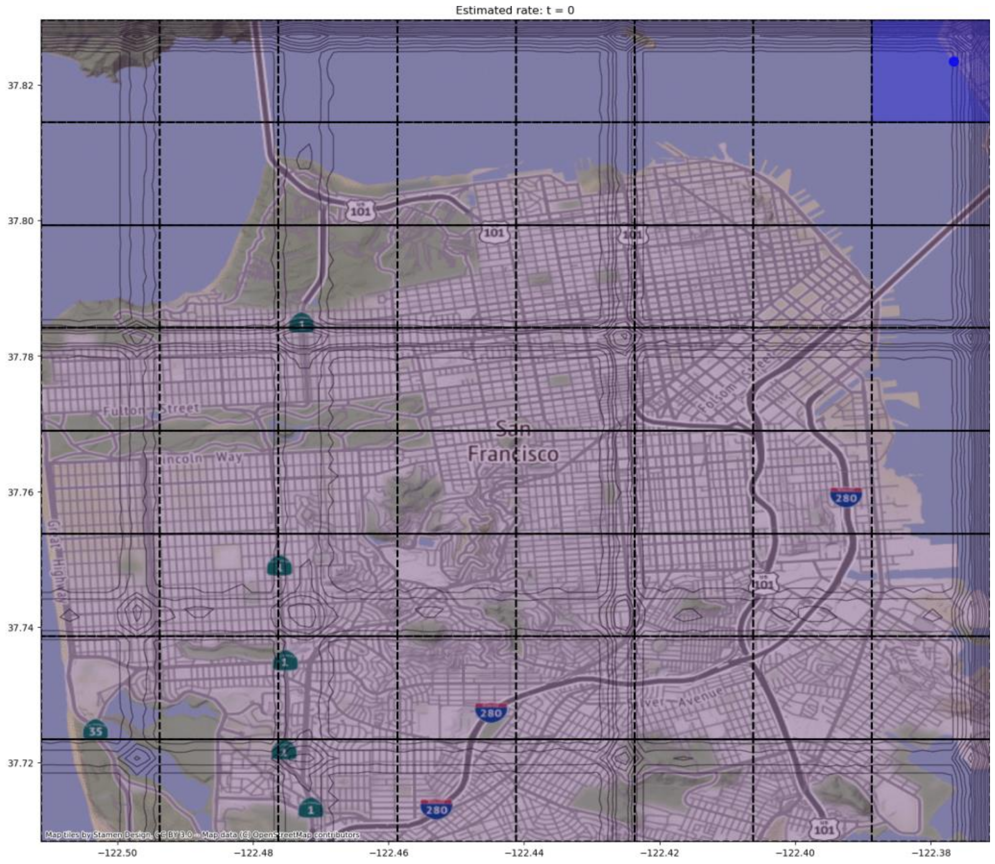


$$N(A) \sim \text{Poisson}\left(\Delta \int_A \lambda(x)dx\right).$$

Tolerance to Noise: Poisson Sensing

New noise models





Scalable models: additive models

Scalable Assumptions

- **Statistical** $\rightarrow$ needs assumptions
  Additive functions

$$g(x_1\, x_2\, x_3\, x_4) = g_1(x_1\, x_2) + g_2(x_2\, x_3) + g_3(x_4)$$

shared

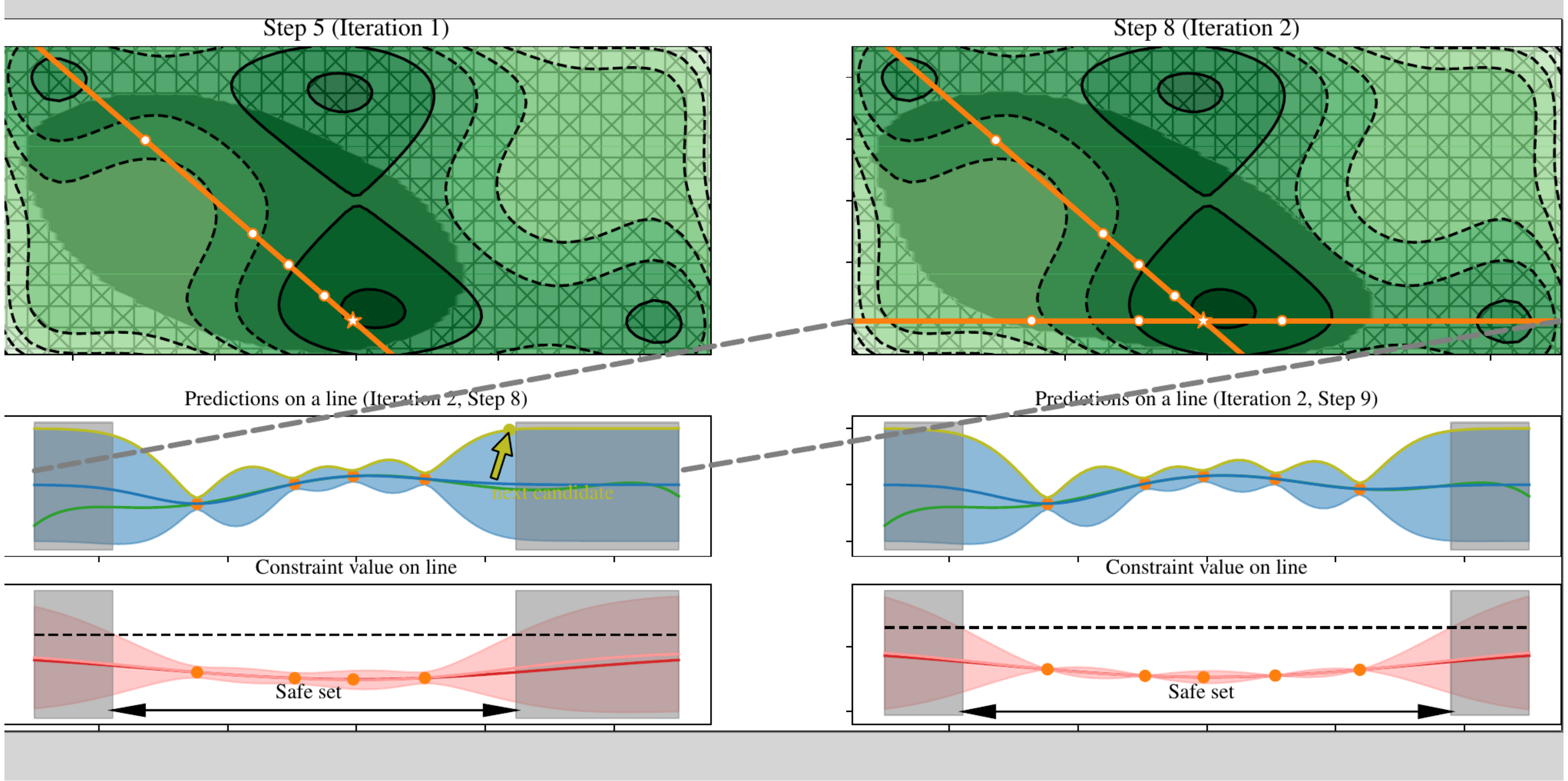
$\bar{d} :=$ *size of largest group*

- **Computational**
  - Kernel inversion $\mathcal{O}(T^3) \rightarrow \mathcal{O}(T \log T)$
  - Optimization of the acquisition function $\rightarrow$ coordinate optimization

Scalable Algorithms - LineBO

Efficient Algorithms

- Prior joint work with PSI, **Johannes,** Nicole, Jochem, Rasmus and Andreas



Pick next experiment x_i
$x_i = argmax_x \, a_i(x)$