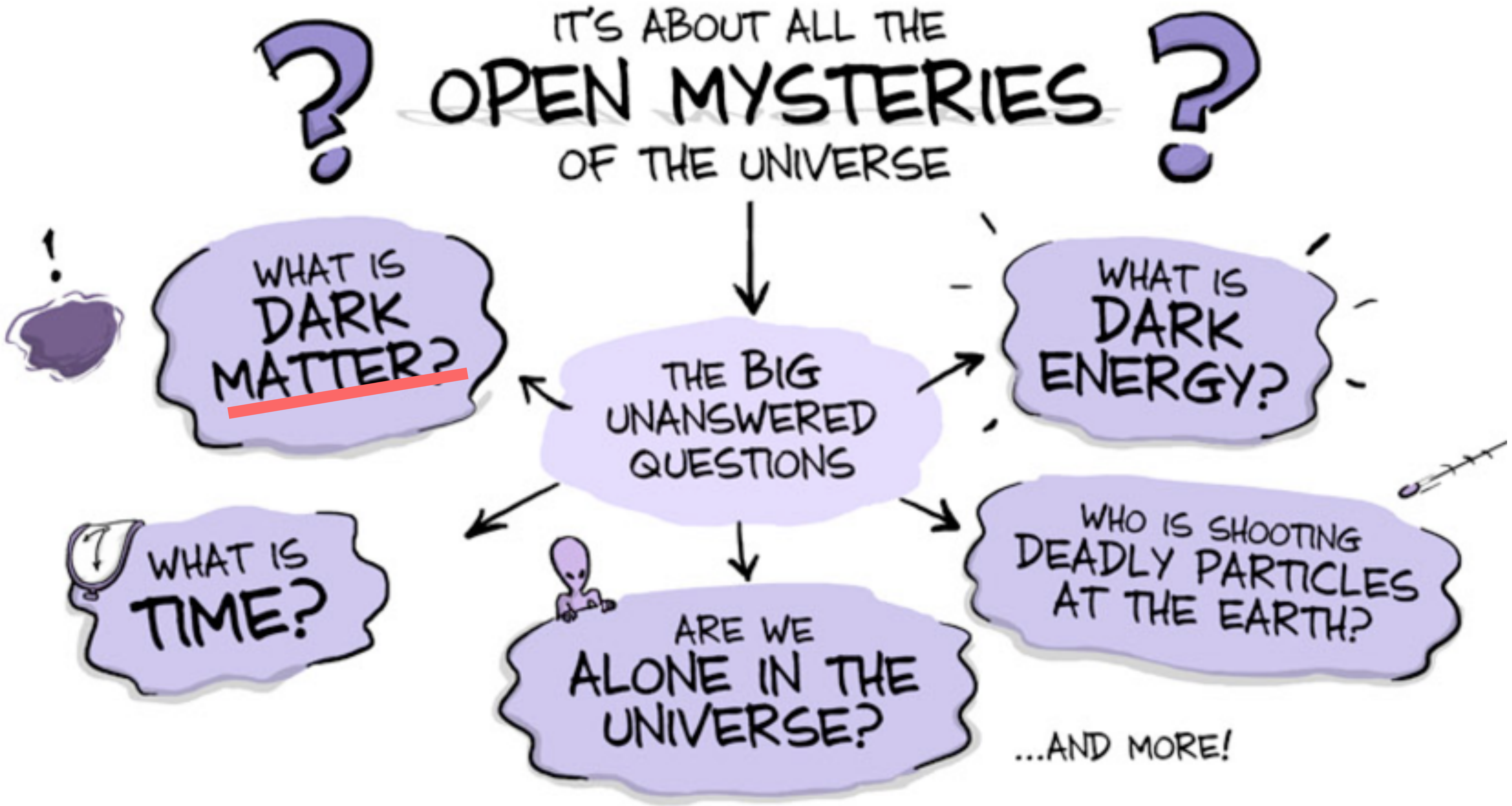
SHiP shield optimization

Artem Filatov
Yandex School of Data Analysis, Higher School of Economics

SHiP experiment

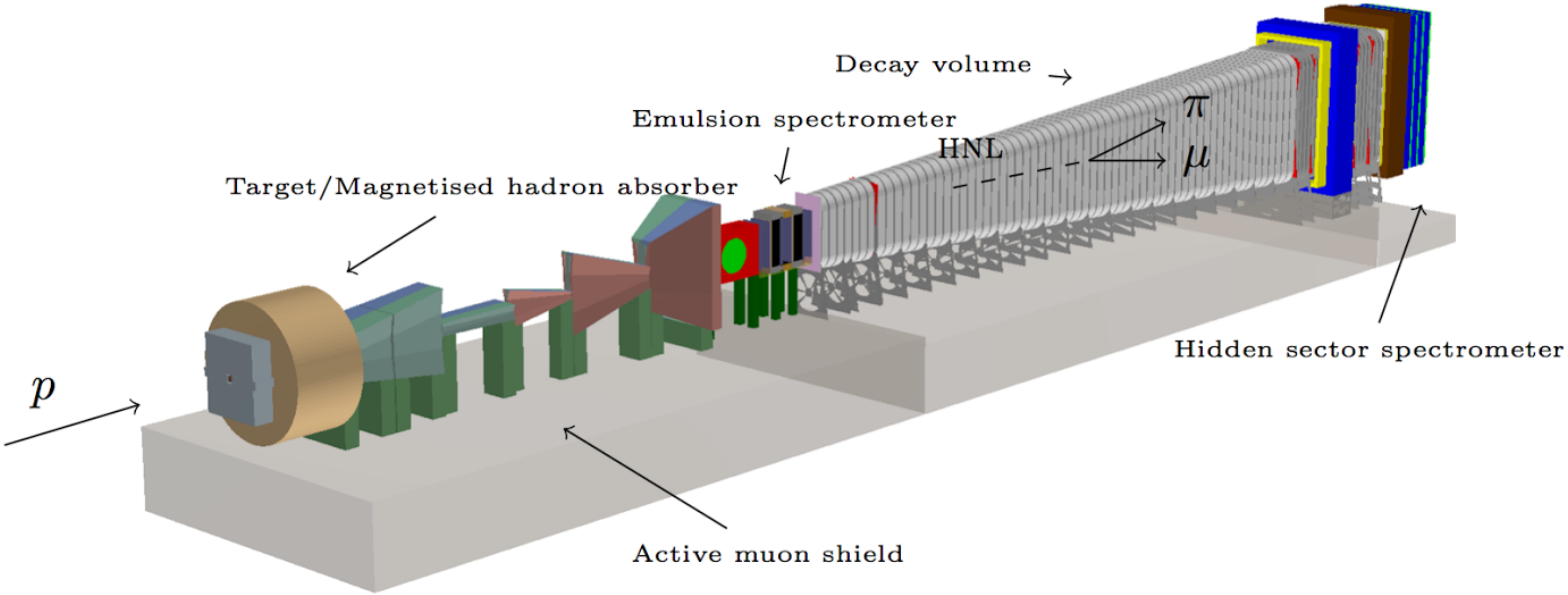


SHiP experiment

> Many theoretical ideas which predict dark matter, and which can be tested experimentally.

SHiP is designed to find a solution for new physics by searching for very weakly interacting particles of the low mass.

SHiP experiment



SHiP shield

SHiP shield

- The experiment needs to minimize backgrounds from all known particles.
- Critical part is the deflection of muons by a magnetic shield.

SHiP shield

- The shield contains 8 magnets and each magnet parametrized by 7 values.
- It cost about 4000 $ per ton.
- We need to find a cheap and efficient solution which minimize backgrounds.

Evaluation of the shield

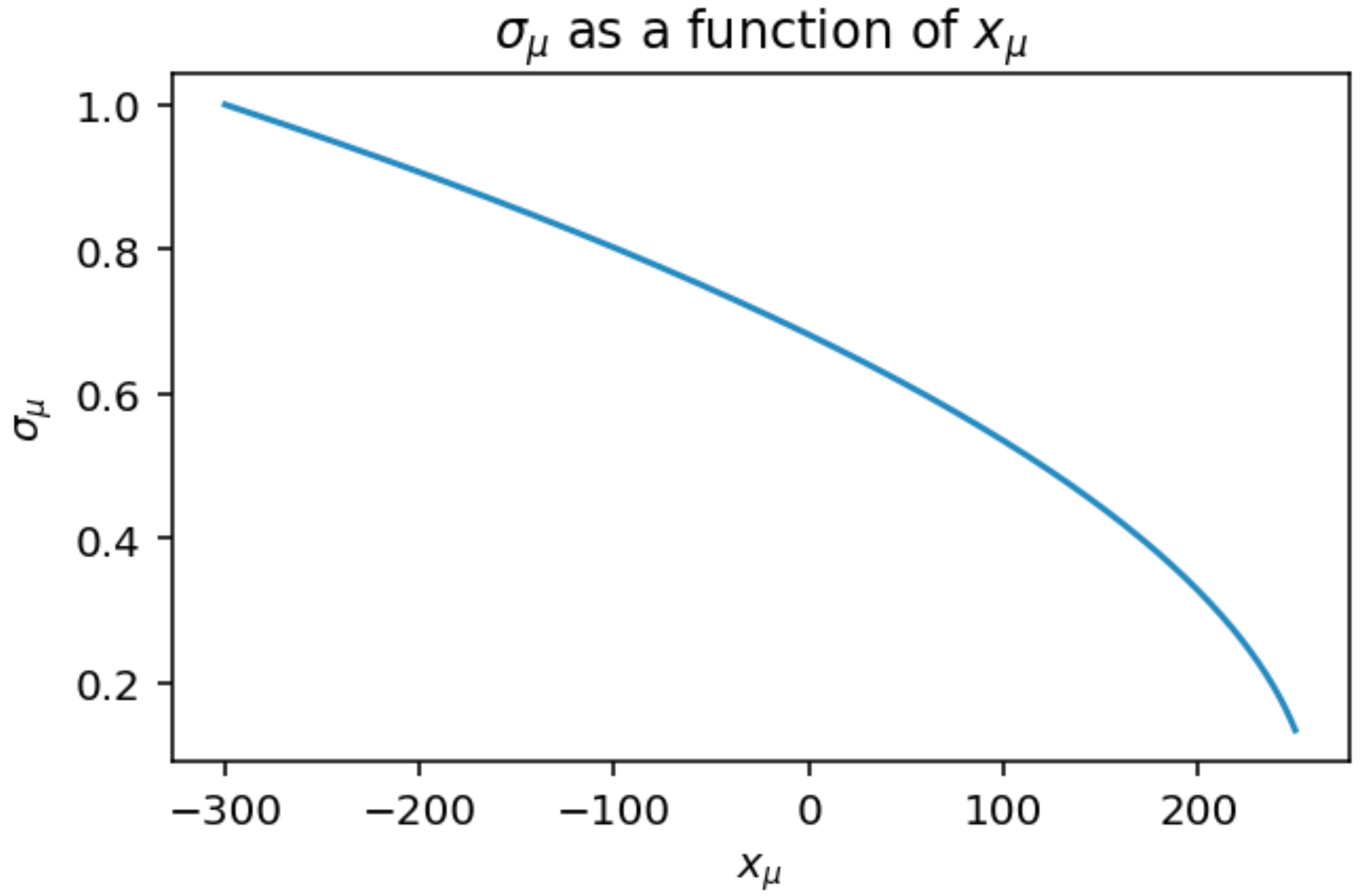
- For given configuration we can make MC simulations.
- 17.8M muons pass thought the shield.
- For every muon we can compute the following value

$$\sigma_\mu = \sqrt{1 - (x_\mu + 300)/560}$$

where x_μ is the coordinate of the muon with respect to the center of the scoring plate

σ_μ

- σ_μ represents the effectiveness of the magnet to the given muon.
- We are trying to put all the muons to the right part of the scoring plate.



Loss function

- We need to formalize the notion of a good configuration in some function.
- Let design loss function which reflects our views about best solution.

Loss function

We can measure performance of the shield by Σ value, which is computed over all muons μ.

$$\Sigma = \sum_{\mu} \sigma_{\mu}$$

We design loss function that depends on the Σ, weight W and some fixed weight W_{bl}.

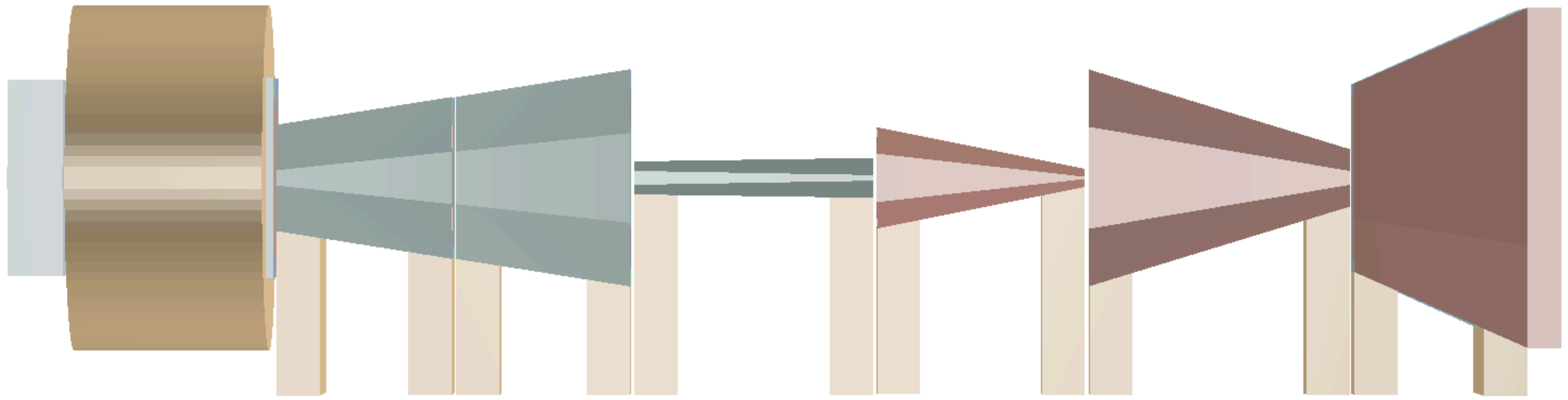
$$L = (1 + \Sigma)(1 + \exp(10(W - W_{bl})/W_{bl}))$$

Our goal is to minimize L.

Baseline configuration

Baseline

- We had a baseline which was derived manually.
- The weight is about 1900 tons and Σ is equal to 32.
- But we would like to find cheaper and more efficient solution.



Our approach: Bayesian Optimization

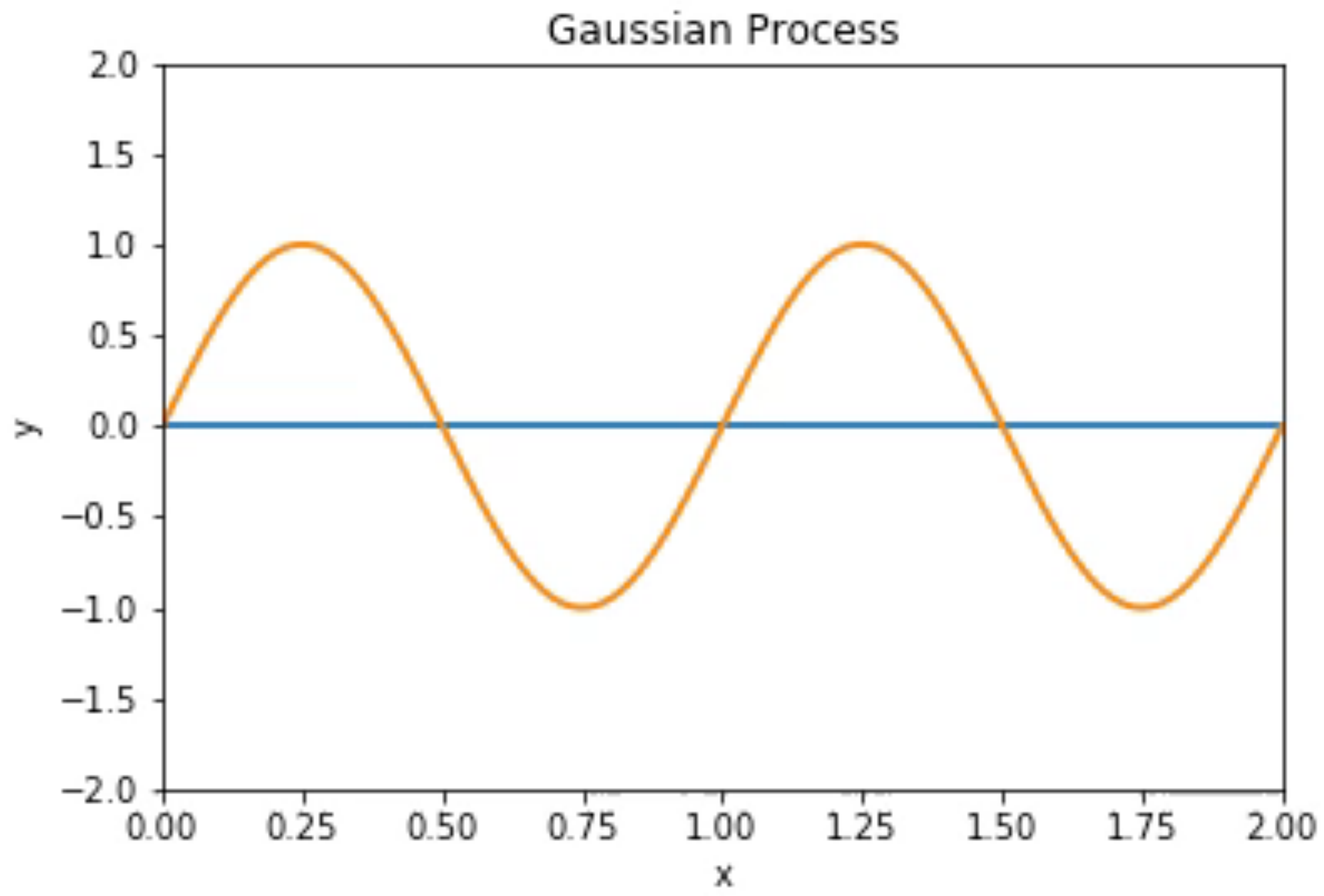
Optimization cycle

- Build a surrogate model over loss function.
- Choose next point according to surrogate model via probabilistic methods.
- Compute next point.

Surrogate modeling

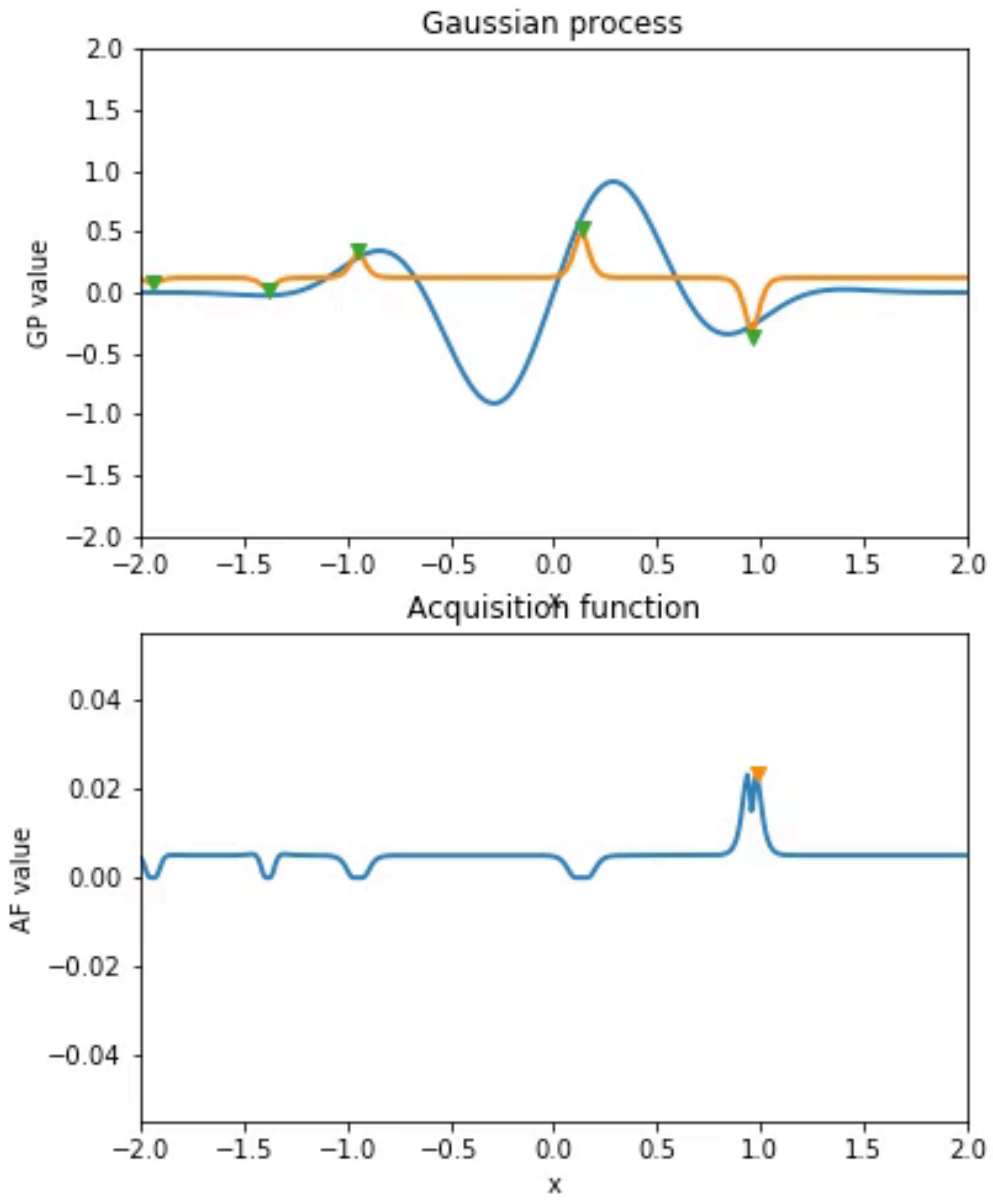
Surrogate models

- Gaussian Processes
- Random Forest
- Gradient Boosting



Expected improvement

- New points suggested by Expected Improvement algorithm.
- Find point which maximize $E(y^* - \hat{f}(x))^+$.
- Can deal with exploration and exploitation.
- Works with a big class of non-differentiable functions.



Optimization results

Difficulties

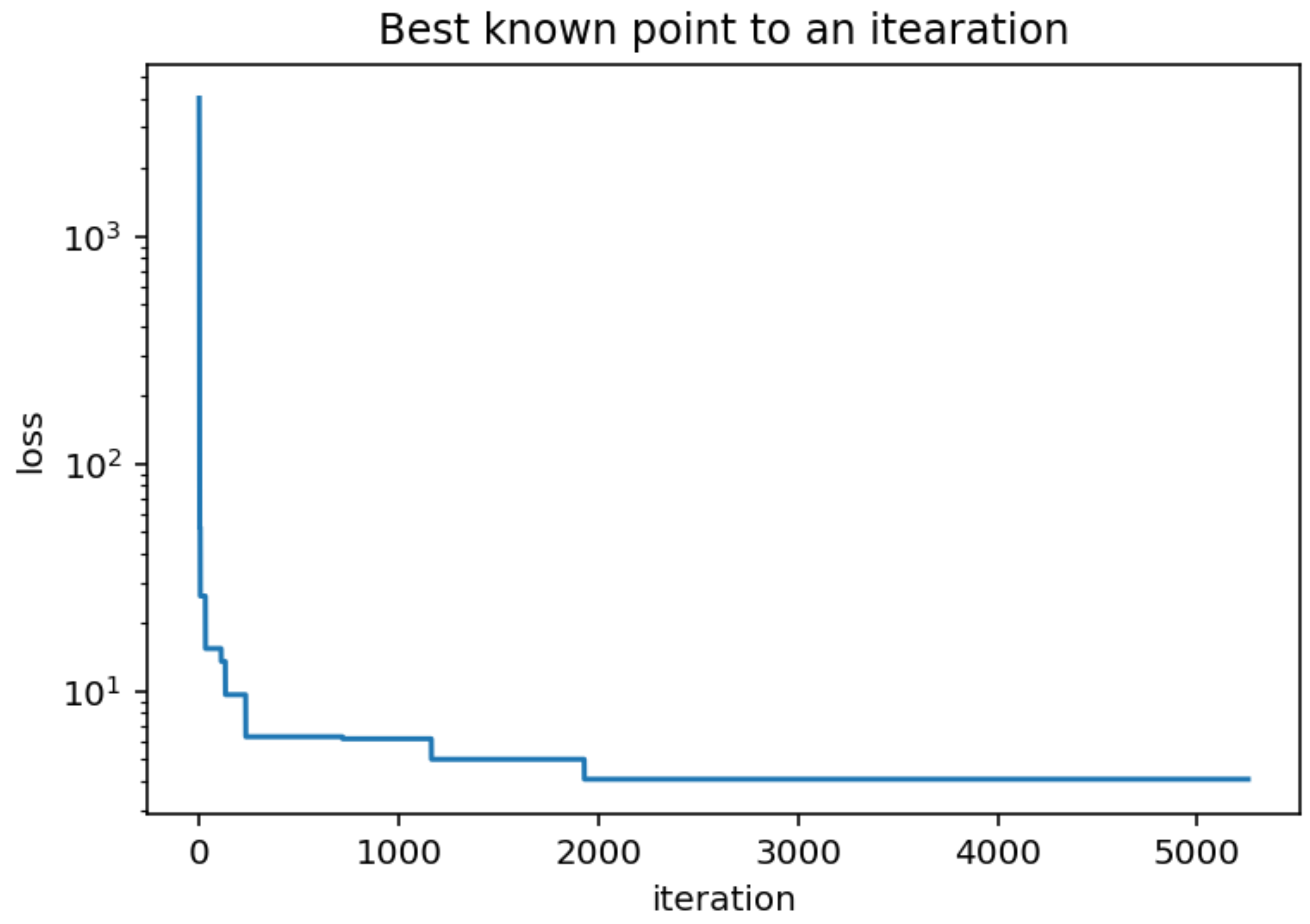
- High dimensional space
- Computation of the Σ is time consuming
- Computation of the Σ is noisy

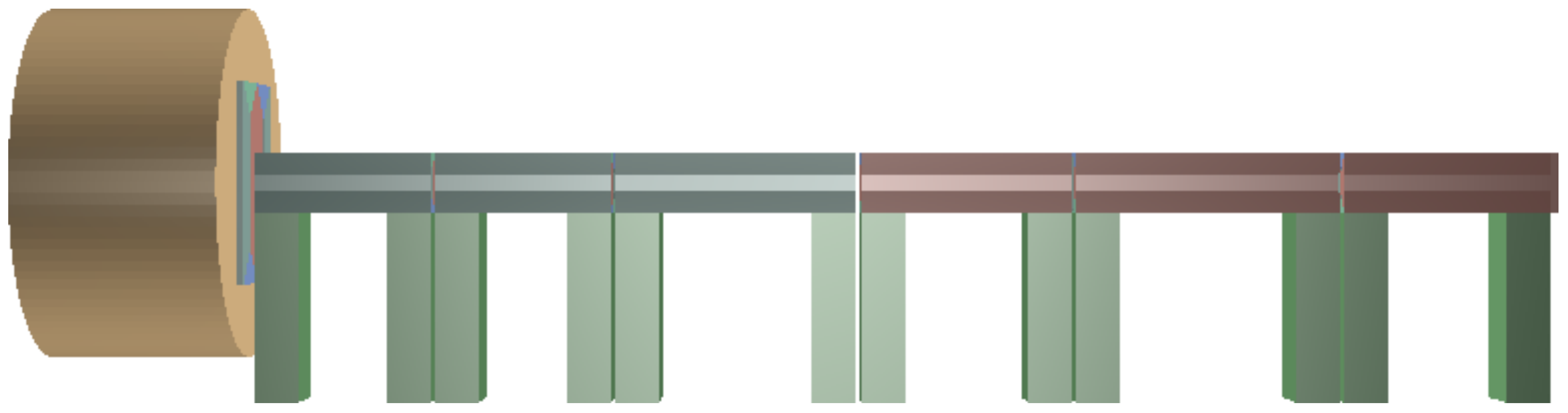
Initial setup

- To increase the speed of computation we made simulations only with ‘bad’ muons.
- Discard many low-momentum muons. Finally we left only 485K muons.
- For computations we have used a large distributed system, as task is well-parallelized.

Optimization run

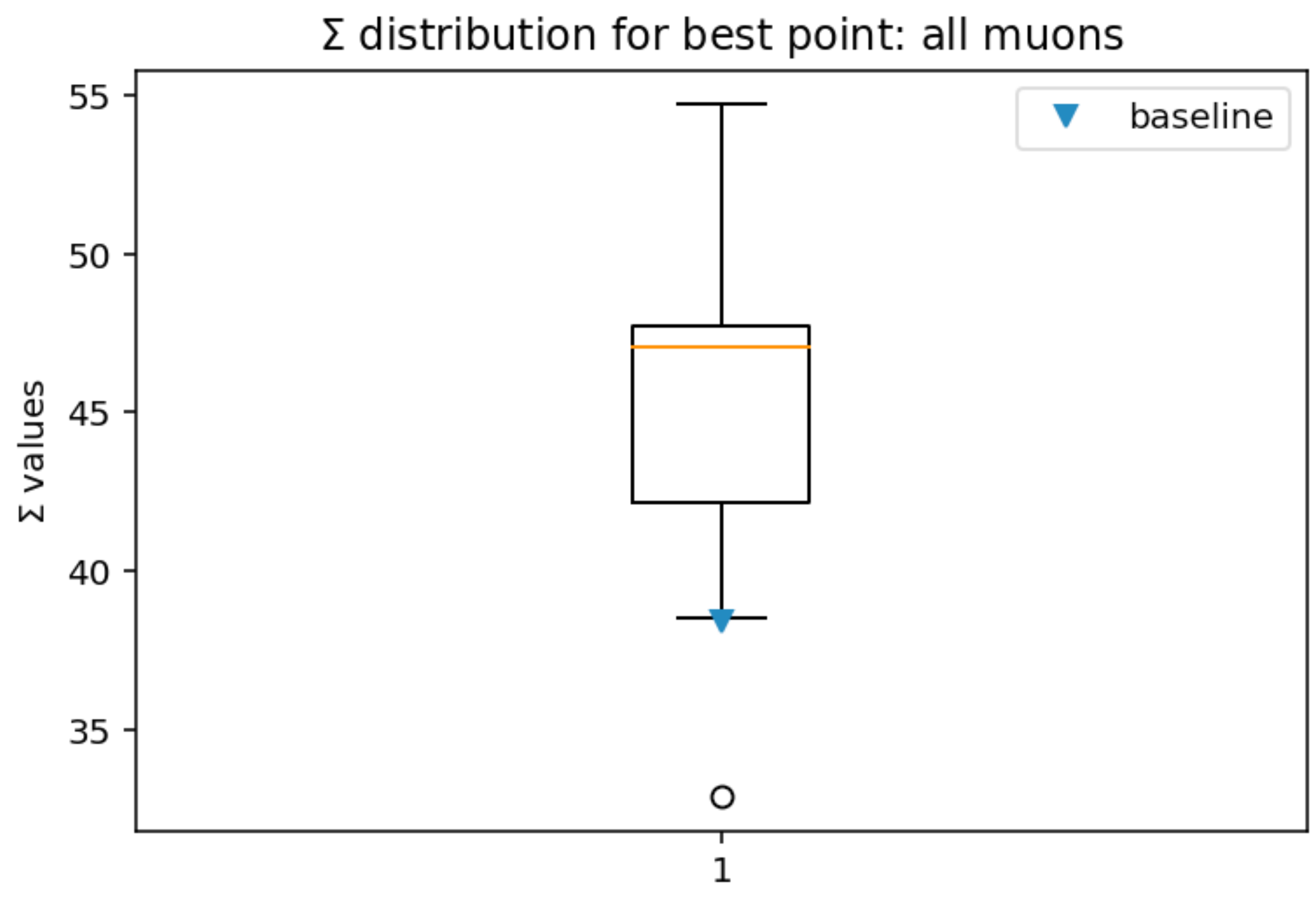
- Optimization started from light rectangular configuration.
- Points was computed in batches.
- After 5000 points we stated a result.

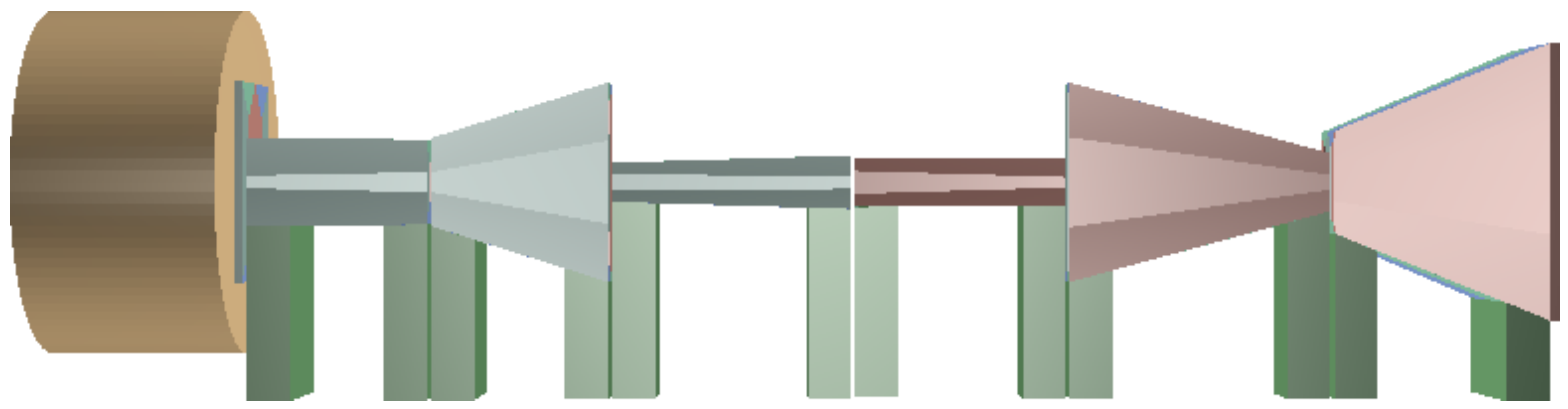




Results

- A new solution is lighter by 25 percent.
- It has a similar performance in terms of Σ value.
- It is significantly cheaper!





Conclusion

- Bayesian Optimization works and solution was found.
- We can optimize a lot of non-differentiable tasks, e.g. physical experiments.