SHiP Spectrometer Optimization using Bayesian Optimization with Gaussian Processes

23rd International Conference on Computing in High Energy and Nuclear Physics

Oleg Alenkin[1], **Mikhail Hushchyn**[2,3], **Katerina Kuznetsova**[4], **Massimiliano Ferro-Luzzi**[4], **Alexander Baranov**[2,3]

[1]Skolkovo Institute of Science and Technology, [2]Yandex School of Data Analysis
[3]National Research University Higher School of Economics, [4]SHiP

Introduction

SHiP is a new proposed fixed-target experiment at the CERN SPS accelerator. The goal of the experiment is to search for hidden particles predicted by models of Hidden Sectors. The purpose of the SHiP Spectrometer Tracker is to reconstruct tracks of charged particles from the decay of neutral New Physics objects with high efficiency. Efficiency of the track reconstruction depends on the spectrometer geometry. Parameters of the geometry can be optimized to archive the higher efficiency. In this study the SHiP Spectrometer Tracker optimization using Bayesian optimization with Gaussian processes in considered.

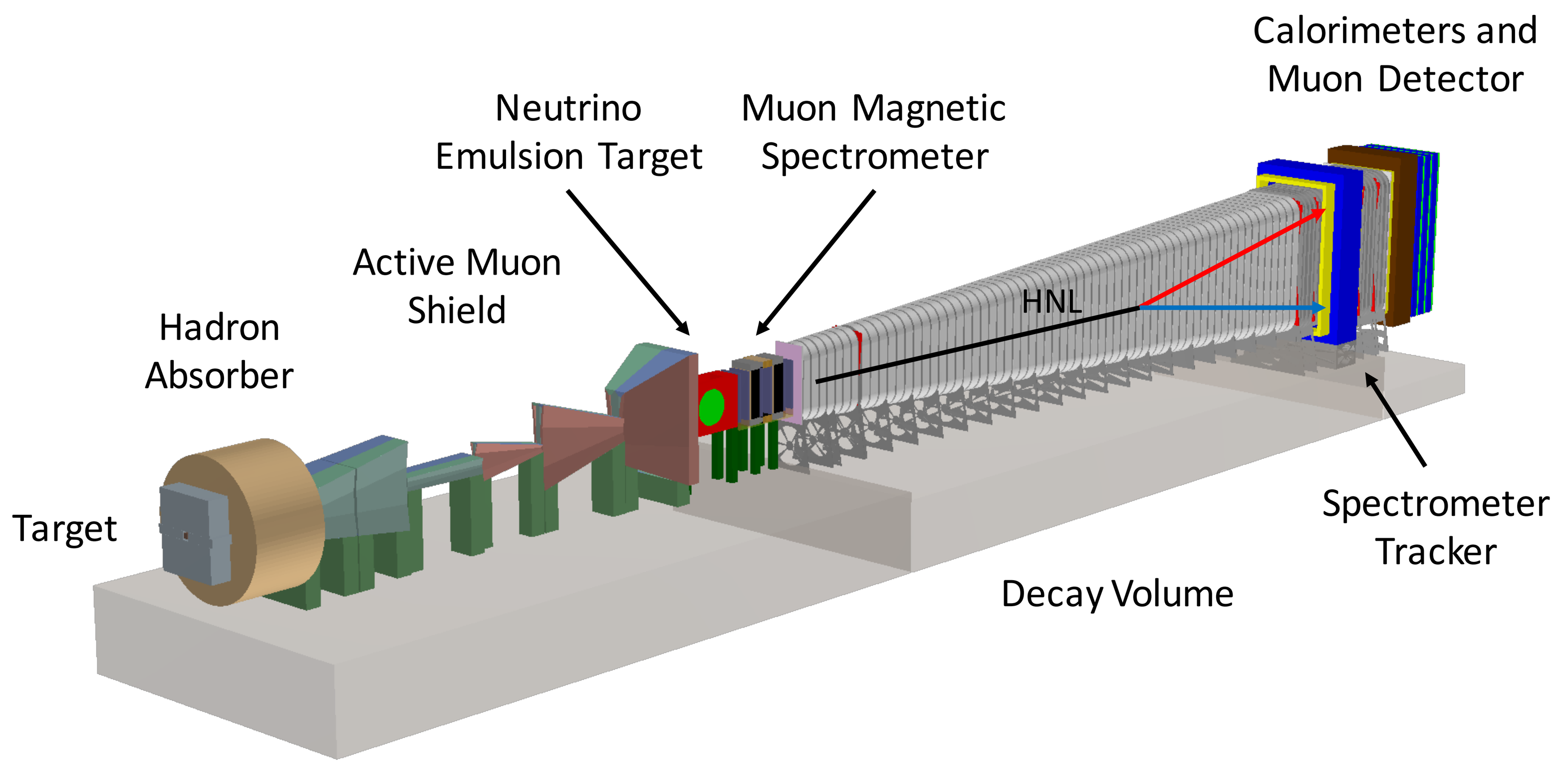

Figure 1: SHiP detector layout

SHiP Spectrometer

The SHiP Spectrometer has 2 straw tube stations before the magnet and 2 stations after it. Each station has 4 views: 2 Y-views with straw tubes along Y axis and U, V-views rotated relative to Y axis. A view has 2 planes with 2 straw tubes layers in each plane as it is shown in Fig 2. In additional to the parameters in the figure, there are z shift between views inside one station and angle between Y and Stereo (U, V) views.

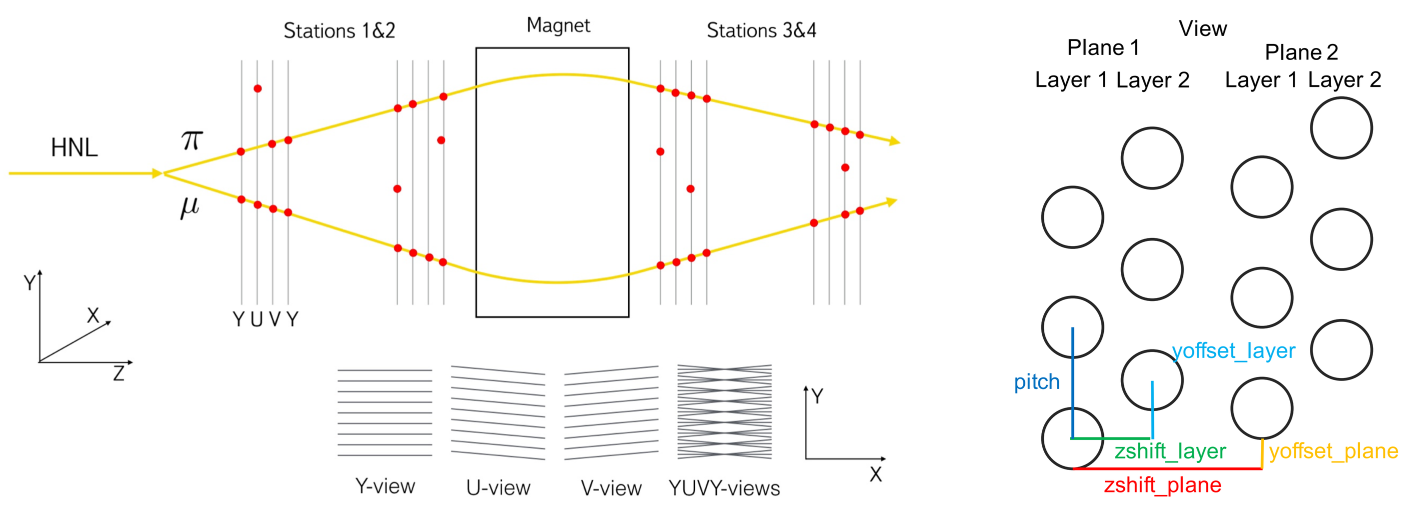

Figure 2: The SHiP Spectrometer (left) and one view layout (right).

Gaussian Processes Regression

Consider an objective function $y = f(x)$ that is needed to approximate using set of observations $\{x, y\}_{i=1}^{n}$. In the Gaussian processes assumption the sequence of the observations has a Gaussian distribution:

$$\mathbf{y} = (y_1, y_2, ..., y_n) \sim \mathcal{N}(\mu, K), \quad \mu = (\mu(x_1), \mu(x_2), ..., \mu(x_n))$$

Then, for a new point (x_{n+1}, y_{n+1}).

$$\begin{pmatrix} \mathbf{y} \\ y_{n+1} \end{pmatrix} \sim \mathcal{N}(\mu', K')$$

$$K' = \begin{pmatrix} K & \mathbf{k} \\ \mathbf{k^T} & k(x_{n+1}, x_{n+1}) \end{pmatrix}, \quad \mathbf{k^T} = \begin{pmatrix} k(x_1, x_{n+1}) \\ k(x_2, x_{n+1}) \\ \vdots \\ k(x_n, x_{n+1}) \end{pmatrix}, \quad k(x_i, x_j) = exp(-\frac{1}{2\theta^2}\|x_i - x_j\|^2)$$

The objective function approximation $\mu(x_{n+1})$ is defined from the conditional distributions:

$$\mu(x_{n+1}) = \mathbf{k^T} K^{-1} \mathbf{y}$$
$$\sigma(x_{n+1}) = k(x_{n+1}, x_{n+1}) - \mathbf{k^T} K^{-1} \mathbf{k}$$

Bayesian Optimization

Bayesian optimization loop:

- Given observation $\{x_i, y_i = f(x_i)\}_{i=1}^{n}$, fit a Gaussian processes regression model.
- Optimize Expected Improvement (EI) acquisition function based on the regression model for sampling the next point:

$$x_{n+1} = \arg\min_x EI(x)$$

$$EI(x) = \mathbb{E}[\max\{0, \hat{f}(x) - f(x^-)\}], \quad \hat{f}(x) \sim \mathcal{N}(\mu(x), \sigma(x)), \quad x^- = \arg\min_{x_k \in \{x_i\}_{i=1}^{n}} f(x_k)$$

- Sample the next observation $(x_{n+1}, y_{n+1} = f(x_{n+1}))$.

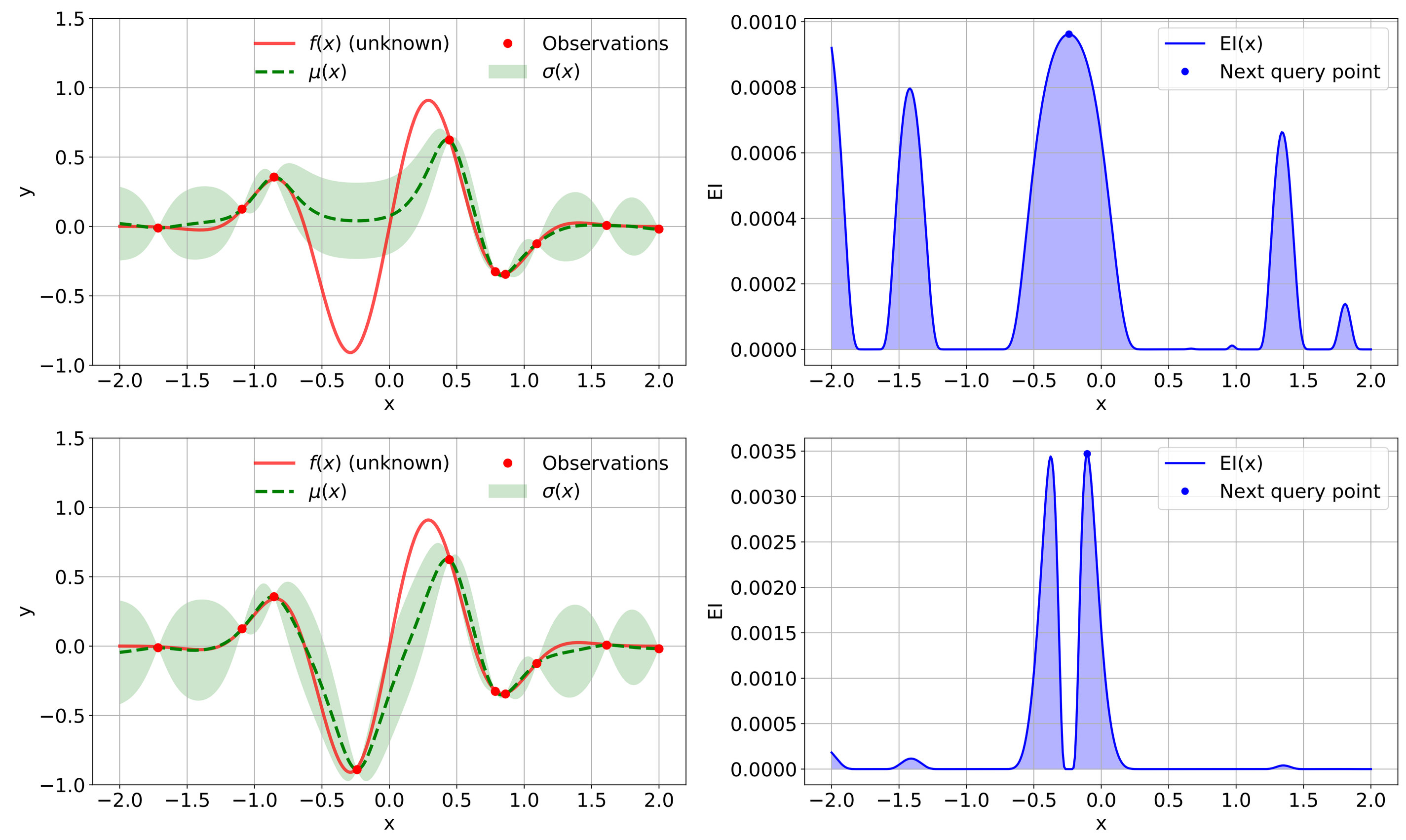

Figure 3: Examples of Gaussian processes regression and Bayesian optimization (minimization).

SHiP Spectrometer Optimization

The following parameters of the SHiP Spectrometer geometry are used during the optimization:

- Straw tube diameter: 2.0 cm
- Pitch: 3.6 cm
- Z shift between layers: 1 - 12 cm
- Z shift between planes: 1 - 12 cm
- Z shift between views: 10 - 12 cm
- Y offset between layers: 1.8 - 3.6 cm
- Y offset between planes: 0.9 - 4.5 cm
- Angle α between Y and U, V views: 5 - 15 degrees

For each combination of the parameters a set of tracks and their hits in the spectrometer are generated. Then, a track pattern recognition algorithm is applied and the following objective function is calculated:

$$Objective = \frac{N_{\text{recognized tracks}}}{N_{\text{generated events}}}$$

Bayesian optimization with Gaussian processes is used to maximize the objective function. The found objective maximum after each iteration of the optimization is shown in Fig 4.

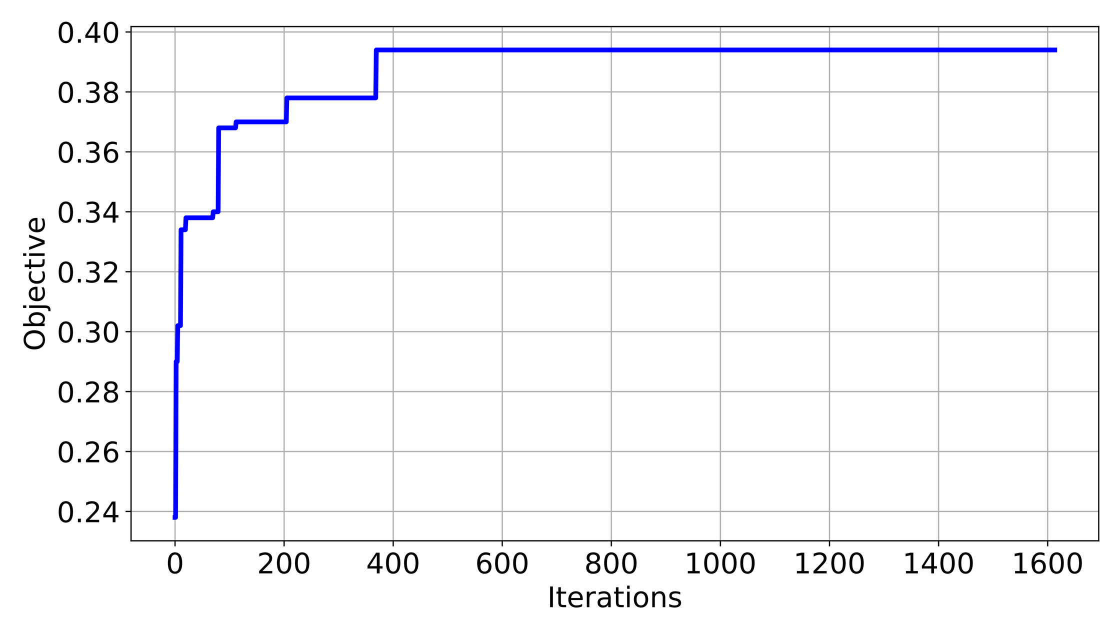

Figure 4: Dependency of the found objective maximum from iterations of the optimization.

Each set of the geometry parameters is checked for agreement with engineering constraints. The found optimal geometry is shown in Fig 5.

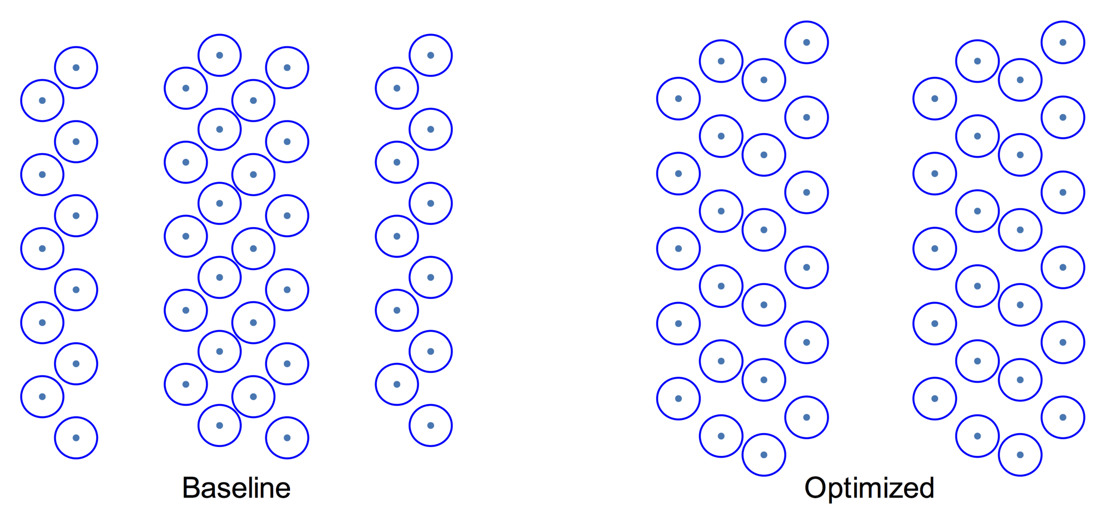

Figure 5: Straw tubes of two views for baseline (left) and optimized (right) geometries.

Conclusions

Bayesian optimization with Gaussian processes is successfully applied for the SHiP Spectrometer optimization. The found geometry configuration provides better track recognition efficiency.