

Optimization of the SHiP Spectrometer Tracker geometry using the Bayesian Optimization with Gaussian Processes

Oleg Alenkin¹, Mikhail Hushchyn^{1,2,*}, Andrey Ustyuzhanin^{2,3}, and Alexander Baranov^{2,3}

¹Skolkovo Institute of Science and Technology

²National Research University Higher School of Economics

³Yandex School of Data Analysis

Abstract. One of the most important aspects of data processing at SHiP [1] experiments is tracks pattern recognition. The purpose of the SHiP Spectrometer Tracker (SST) is efficient reconstruction of charged particle tracks originating from decays of neutral New Physics objects. The reconstruction performance strongly depends on the tracker design and should be considered as an objective to define the best SST geometry parameters. In this study the SHiP Spectrometer Tracker geometry optimization using Bayesian optimization with Gaussian processes is considered. The study has been done on MC data. The first results of the optimization are also considered.

1 Introduction

SHiP is a fixed-target experiment at the CERN SPS accelerator proposed recently for searches of the hidden particles predicted by a variety of physics models beyond the Standard Model [1, 2]. The purpose of the SHiP Spectrometer Tracker (SST) is efficient reconstruction of charged particle tracks originating from decays of neutral New Physics objects. The SST geometry parameters have to be optimized to achieve the highest reconstruction efficiency.

One of the most popular methods of optimization is grid search. The disadvantage of this method is exponential grid size growth with increase of the number of optimization parameters. This requires large computational resources to check all grid nodes. In the SST geometry optimization task, evaluation of every node requires large statistics Monte Carlo simulation of SHiP physics events, what makes the grid search approach too expensive. Bayesian optimization method [3] is well-suited to the optimization of expensive objective functions. Using information from all previously evaluated nodes, the method chooses the next node to be checked requiring the objective function to have larger probability for the optimum. This allows to find the optimum with lower number of checks than with the grid search.

Bayesian optimization method was tested in our previous study on simulator tuning [4] and showed promising result. In this study the SHiP Spectrometer Tracker geometry optimization using the Bayesian optimization with Gaussian processes (GP) [3, 5] is considered.

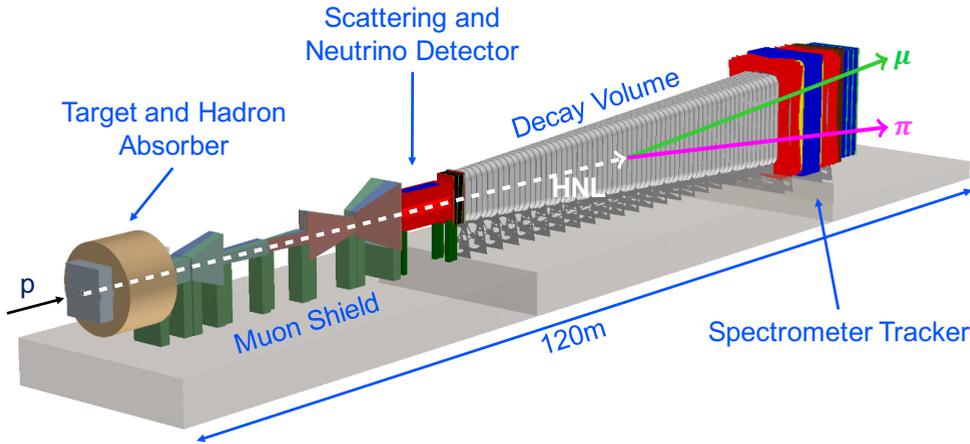


Figure 1: The SHiP detector.

2 SHiP Spectrometer Tracker

SHiP [6, 7] experiment was proposed in 2013 for searches of weakly interacting long-lived particles predicted by the Hidden Sector theoretical models. The general scheme of the detector is demonstrated in Fig. 1. SHiP is a detector with a fixed target. A proton beam from the CERN SPS accelerator with energy of 400 GeV bumps into it. The target [8] is designed to maximize production of charm and beauty hadrons and photons. It is followed by the 5m hadron absorber that absorbs all particles except weakly interacting ones: muons, neutrinos, and particles of new physics. The Muon Shield [9, 10] is used to reduce muon flux in the detector acceptance. The shield is a system of 6 magnets with a total length of 35m. The average magnetic field is 1.7T. The Scattering and Neutrino Detector (SND) [11, 12] is located after the muon shield. The goal of the detector is to study neutrino physics and to detect light dark matter candidates.

SND is followed by the second detector complex [11] for studying visible decays of the Hidden Sector particles — Heavy Neutral Leptons (HNL). The complex consists of the Decay Volume, Spectrometer Tracker, Electromagnetic Calorimeter and Muon Detector. HNL decays in the Decay Volume into visible particles. Fig. 1 shows an example of the $HNL \rightarrow \mu\pi$ decay. The spectrometer is used to reconstruct the tracks of the visible particles. The Electromagnetic Calorimeter and the Muon Detector are used to identify their types.

The SHiP Spectrometer Tracker has 2 straw tube stations before the magnet and 2 stations after it. Each station has 4 views: 2 Y-views with straw tubes oriented parallel to X-axis and U, V-views rotated by a small angle, $+\alpha$ or $-\alpha$, around the Z axis. A view has 2 planes with 2 straw tubes layers in each plane as it is shown in Fig 2. The view geometry defines the particle track recognition [13] quality. The geometry is defined by the following parameters: straw pitch, Z shift between layers, Z shift between planes, Z shift between views, Y offset between layers, Y offset between planes and angle α between Y and U, V views. The goal of this study is to find optimal values of these parameters.

*e-mail: mikhail.hushchyn@cern.ch

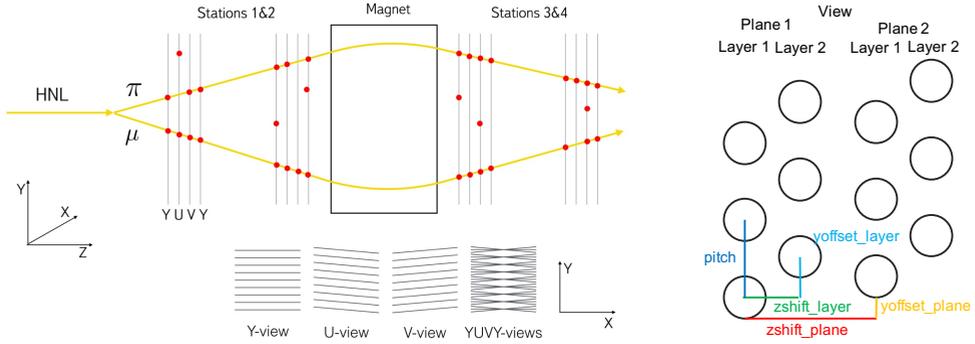


Figure 2: The SHiP Spectrometer Tracker stations and an example of HNL decay (left). The view layout and parameters for optimization (right).

3 Gaussian Processes Regression

GP regression [3, 5] is used for the spectrometer geometry optimization as described in the next section. Consider an objective function $y = f(x)$ and a set of observations $\{x_i, y_i\}_{i=1}^n$. In the GP concept, the sequence of observations can be represented as a sample from a Gaussian distribution:

$$\mathbf{y} = (y_1, y_2, \dots, y_n)^T \sim \mathcal{N}(0, K) \quad (1)$$

where $\mathcal{N}(0, K)$ represents multivariate Gaussian (or normal) distribution with zero mean and covariance matrix K . For a new point (x_{n+1}, y_{n+1}) a similar relation with the new covariance matrix K' can be written:

$$\begin{pmatrix} \mathbf{y} \\ y_{n+1} \end{pmatrix} \sim \mathcal{N}(0, K') \quad (2)$$

where the covariance coefficients k_{ij} characterize the correlation between different points of the process: for nearby points x_i and x_j the observables y_i and y_j are strongly correlated than for distant points. Here squared exponential covariance function is used to define coefficients k_{ij} :

$$K' = \begin{pmatrix} K & \mathbf{k} \\ \mathbf{k}^T & k(x_{n+1}, x_{n+1}) \end{pmatrix}, \quad \mathbf{k} = \begin{pmatrix} k(x_1, x_{n+1}) \\ k(x_2, x_{n+1}) \\ \vdots \\ k(x_n, x_{n+1}) \end{pmatrix}, \quad k(x_i, x_j) = \exp\left(-\frac{1}{2\theta^2} \|x_i - x_j\|^2\right) \quad (3)$$

where θ defines the smoothness of the objective function.

The mean and standard deviation of the objective function approximation, $\mu(x_{n+1})$ and $\sigma(x_{n+1})$, are defined by the conditional distribution:

$$(y_{n+1} | \mathbf{y}) \sim \mathcal{N}(\mu(x_{n+1}), \sigma(x_{n+1})) \quad (4)$$

$$\mu(x_{n+1}) = \mathbf{k}^T K^{-1} \mathbf{y} \quad (5)$$

$$\sigma(x_{n+1}) = k(x_{n+1}, x_{n+1}) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k} \quad (6)$$

76 The optimal θ value in the covariance matrix \mathbf{K} is found based on the likelihood function
 77 optimization:

$$\theta^{opt} = \arg \max_{\theta} \log p(\mathbf{y}|\mathbf{X}, \theta) \quad (7)$$

$$\log p(\mathbf{y}|\mathbf{X}, \theta) = -\frac{1}{2} \mathbf{y}^T \mathbf{K}^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K}| - \frac{n}{2} \log 2\pi \quad (8)$$

78 Equations for $\mu(x_{n+1})$ and $\sigma(x_{n+1})$ are used during Bayesian optimization to approximate
 79 the objective function based on known observations.

80 4 Bayesian Optimization

81 Bayesian optimization [3] is a method of finding the optimum of an expensive objective
 82 function. The goal of the Bayesian optimization is to find the optimum using an as small as
 83 possible number of function calculations. The Bayesian optimization loop has the following
 84 steps which are repeated until the optimum is found:

- 85 • Given observations $\{x_i, y_i = f(x_i)\}_{i=1}^n$, fit a GP regression model to get $\mu(x_{n+1})$ and $\sigma(x_{n+1})$
 86 of the objective function approximation.
- Optimize the Expected Improvement (EI) acquisition function [14] based on the regression
 model for sampling the next point:

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

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

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

88 5 SHiP Spectrometer Tracker Optimization

89 The goal of this study is to find the optimal geometry of the SHiP Spectrometer Tracker us-
 90 ing Bayesian optimization with GP. The parameter values of the SHiP Spectrometer Tracker
 91 geometry used during the optimization are listed in Tab. 1.

Parameter	Range
Straw Tube Diameter	2.0 cm
Pitch	3.6 cm
Y Offset Layer	1.8 - 3.6 cm
Y Offset Plane	0.9 - 4.5 cm
Z Shift Layer	1.8 - 3.6 cm
Z Shift Plane	1 - 12 cm
Z Shift View	10 - 12 cm
Angle	5° - 15°

Table 1: Geometry parameter ranges for the optimization.

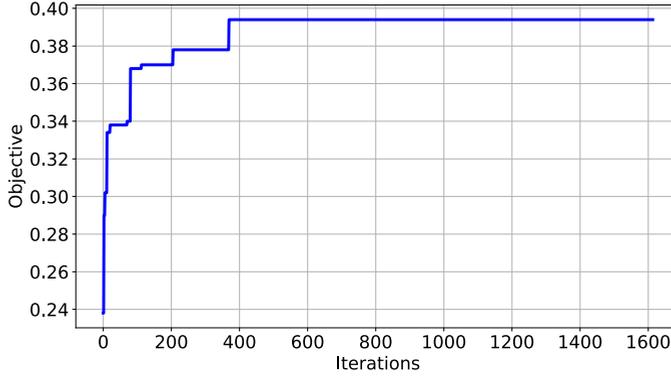


Figure 3: The maximal value of the objective function defined as the number of reconstructed track per a generated event, as a function of the number of iterations.

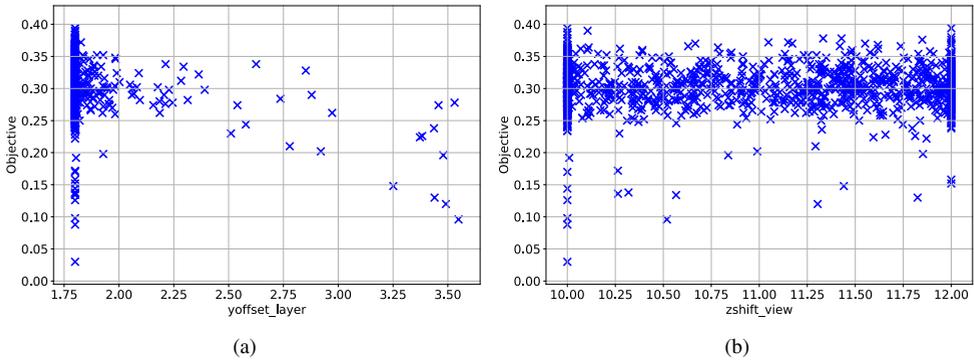


Figure 4: Dependencies of the objective function from (a) Y offset between layers and (b) Z shift between views for observations of the optimization.

92 For each combination of the parameters a set of 500 events of $HNL \rightarrow \mu\pi$ decay is
 93 generated using FairShip [15]. Then, a track pattern recognition algorithm [11] is applied
 94 to find tracks based on hits in the spectrometer. The simulation and recognition for one
 95 combination of the parameters takes about 1000 seconds. For the optimization the following
 96 objective function is calculated:

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

97 where $N_{\text{recognized tracks}}$ is number of recognized tracks in 500 events; $N_{\text{generated events}} = 500$
 98 is number of events. Bayesian optimization with GP realized in GPy [16] python library is
 99 used to maximize the objective function. The found objective maximum after each iteration
 100 of the optimization is shown in Fig. 3. The optimization shows that Y offsets between straw
 101 tubes layers and planes of the spectrometer have the greatest impact on the objective function
 102 as it is shown in Fig. 4a. They have optimal values of $0.5 \times \text{Pitch} = 1.8$ cm and $0.25 \times \text{Pitch} =$
 103 0.9 cm respectively. The smaller Z shift between layers the better, but there are no significant

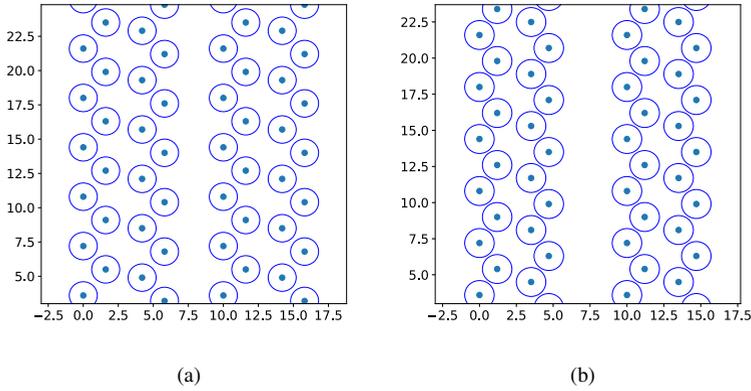


Figure 5: Straw tubes of two views for baseline (a) and optimized (b) geometries.

	Y Offset Layer	Y Offset Plane	Z Shift Layer	Z Shift Plane	Z Shift View	Angle
Baseline	1.9 cm	1.3 cm	1.6 cm	4.4 cm	10 cm	5°
Optimized geometry	1.8 cm	0.9 cm	1.2 cm	3.5 cm	10 cm	5°

Table 2: Values of the spectrometer geometry parameters.

104 differences in the range of 1-2 cm. These three parameters define size of holes between
 105 the straw tubes where a particle can pass through without leaving any hit. So, the smaller
 106 holes the better for the tracks pattern recognition. Z shifts between planes and views, and
 107 angle α do not demonstrate significant influence on the objective function as it is shown in
 108 Fig. 4b. The parameter values for the baseline and one of the optimized geometries are listed
 109 in Tab. 2 and shown in Fig. 5. Check on a set of 50000 events shows the objective function
 110 improvement is within 1% for the new geometry. This means that the baseline is close to the
 111 optimal geometry in terms of the current objective function. However, the optimization using
 112 other quality metrics is needed to be done.

113 6 Conclusions

114 Bayesian optimization with Gaussian processes is successfully applied for the SHiP Spec-
 115 trometer Tracker optimization. The first preliminary results show that the found geometry
 116 configuration provides slightly better quality of the tracks recognition. On the next steps
 117 more quality metrics will be used and the optimization results will be compared with conven-
 118 tional optimization techniques like grid search.

7 Acknowledgments

We wish to thank Katerina Kuznetsova and Massimiliano Ferro-Luzzi for his help and guidance in this work.

The research was carried out with the financial support of the Ministry of Science and Higher Education of Russian Federation within the framework of the Federal Target Program “Research and Development in Priority Areas of the Development of the Scientific and Technological Complex of Russia for 2014-2020”. Unique identifier – RFMEFI58117X0023, agreement 14.581.21.0023 on 03.10.2017.

References

- [1] M. Anelli, S. Aoki, G. Arduini, J. Back, A. Bagulya, W. Baldini, A. Baranov, G. Barker, S. Barsuk, M. Battistin et al. (SHiP Collaboration), Tech. Rep. CERN-SPSC-2015-016. SPSC-P-350. arXiv:1504.04956, CERN, Geneva (2015), technical Proposal, <https://cds.cern.ch/record/2007512>
- [2] S. Alekhin, W. Altmannshofer, T. Asaka, B. Batell, F. Bezrukov, K. Bondarenko, A. Boyarsky, K.Y. Choi, C. Corral, N. Craig et al., Reports on Progress in Physics **79**, 124201 (2016)
- [3] C. Williams, C. Rasmussen, **2** (2005)
- [4] M. Karpov, K. Arzymatov, V. Belavin, A. Sapronov, A. Ustyuzhanin, A. Nevolin, International Journal of Civil Engineering and Technology **9**, 220 (2018)
- [5] E.V. Burnaev, M.E. Panov, A.A. Zaytsev, Journal of Communications Technology and Electronics **61**, 661 (2016)
- [6] C. Ahdida, R. Albanese, A. Alexandrov, A. Anokhina, S. Aoki, G. Arduini, E. Atkin, N. Azorskiy, J. Back, A. Bagulya et al., Journal of Instrumentation **14**, P03025 (2019)
- [7] E. Graverini, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment (2018)
- [8] K. Kershaw, J.L. Grenard, M. Calviani, C. Ahdida, M. Casolino, S. Delavalle, D. Hounsome, R. Jacobsson, M. Lamont, E.L. Sola et al., Journal of Instrumentation **13**, P10011 (2018)
- [9] A. Akmete, A. Alexandrov, A. Anokhina, S. Aoki, E. Atkin, N. Azorskiy, J. Back, A. Bagulya, A. Baranov, G. Barker et al., Journal of Instrumentation **12**, P05011 (2017)
- [10] A. Baranov, E. Burnaev, D. Derkach, A. Filatov, N. Klyuchnikov, O. Lantwin, F. Ratnikov, A. Ustyuzhanin, A. Zaitsev, Journal of Physics: Conference Series **934**, 012050 (2017)
- [11] C. Ahdida, R. Albanese, A. Alexandrov, A. Anokhina, S. Aoki, G. Arduini, E. Atkin, N. Azorskiy, J.J. Back, A. Bagulya et al. (SHiP Collaboration), Tech. Rep. CERN-SPSC-2019-010. SPSC-SR-248, CERN, Geneva (2019), <https://cds.cern.ch/record/2654870>
- [12] A. Buonauro (SHiP Collaboration) (2016)
- [13] M. Hushchyn, A. Ustyuzhanin, O. Alenkin, E. van Herwijnen, Journal of Physics: Conference Series **898**, 042027 (2017)
- [14] T.V. Mockus J, Z. A, *Towards Global Optimization* (1978)
- [15] *Fairship* (2018), <https://github.com/ShipSoft/FairShip>
- [16] *Gpy library* (2018), <http://sheffieldml.github.io/GPy/>