Optimization of muon EDM experimental setup using simulations

Ritwika Chakraborty (PSI)

SPS Annual Meeting 2023

5.09.2023

MuEDM Experiment at PSI

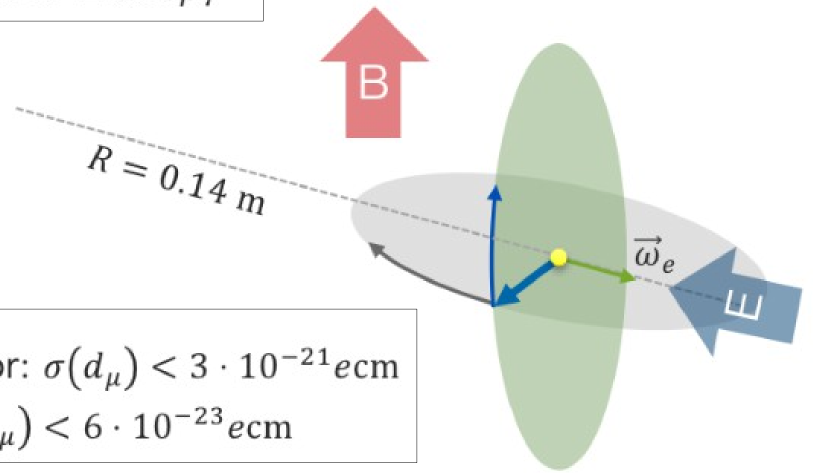
- PSI Muon EDM measurement using Frozen-Spin Technique
- Phase I demonstrates Frozen-Spin technique
- Injecting muons with right experimental design parameters essential for ensuring optimum storage
- Longitudinal asymmetry in positrons decaying from stored muons
  → signal for muon EDM

This talk → simulation studies to optimize initial parameters affecting injection efficiency

$$\vec{\Omega} = \frac{q}{m}\left[a\vec{B} - \left(a + \frac{1}{1-\gamma^2}\right)\frac{\vec{\beta}\times\vec{E}}{c}\right] + \frac{\eta q}{2m}\left[\vec{\beta}\times\vec{B} + \frac{\vec{E}}{c}\right]$$

g-2 term EDM term

set $E \cong aBc\beta\gamma^2$

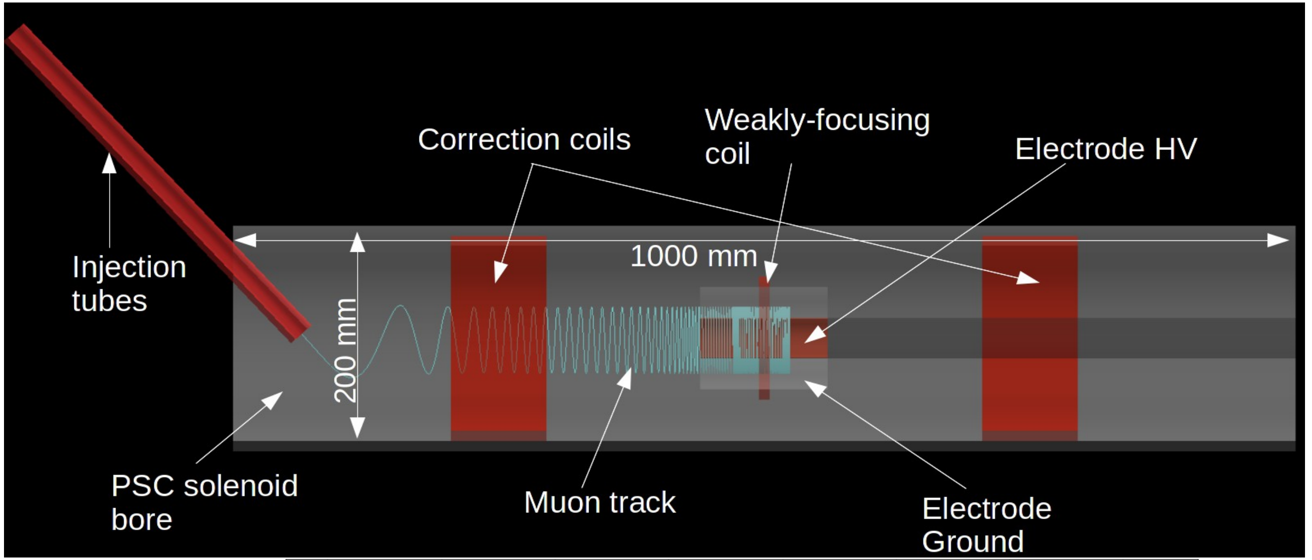


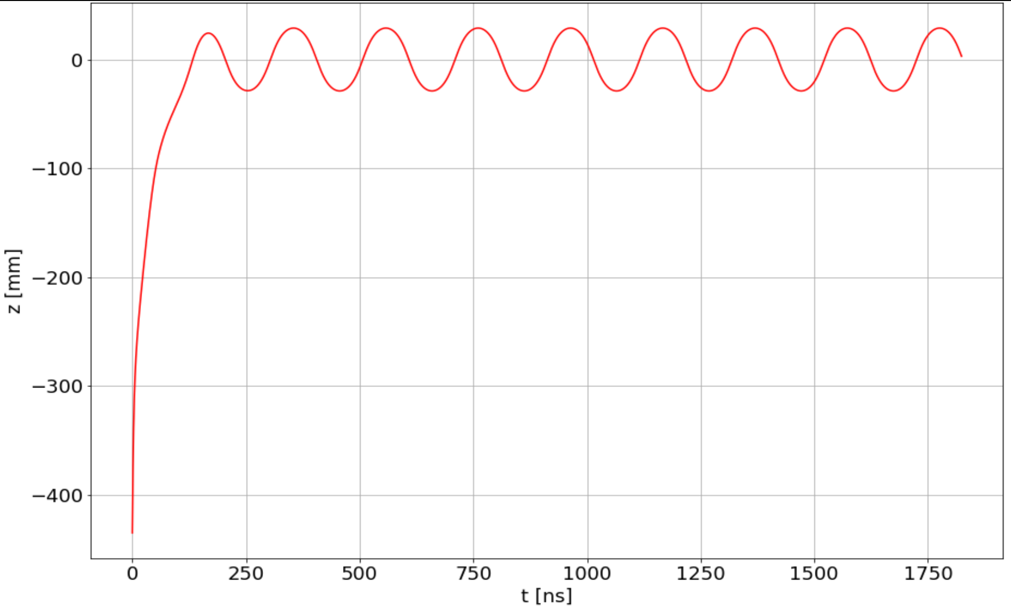
Prescursor: $\sigma(d_\mu) < 3 \cdot 10^{-21} ecm$

Final: $\sigma(d_\mu) < 6 \cdot 10^{-23} ecm$

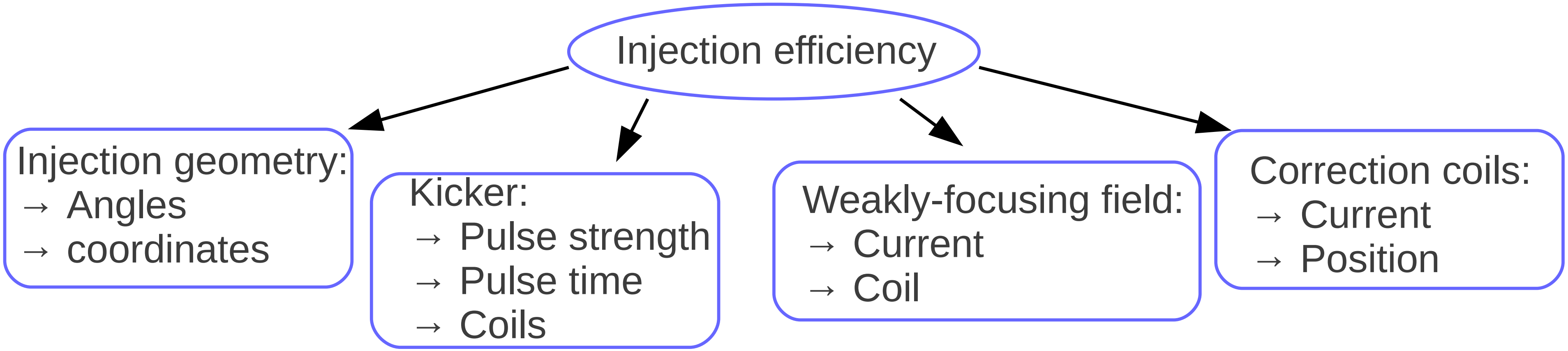
MuEDM Phase I in G4Beamline

- Muons are injected off-axis into 3 T solenoid
- Field gradient at injection corrected by correction coils
- Radial magnetic pulse generated by pair of anti-helmholtz coil
- Muons trapped in weakly-focusing magnetic field at the center generated by thin coils





Multivariate Optimization



- Multivariate optimization → computationally expensive
- G4Beamline simulation run for one configuration of input variables on HPC cluster takes ~6 hrs (for 1M muons)
- High fidelity simulations impossible to do given time constraint

PCE Based Surrogate Modelling

- Replace complex model with approximation → Surrogate Model
- Polynomial Chaos Expansion (PCE) :

$$Y = \sum_{i=0}^{\infty} \alpha_i \Psi_i(\vec{x})$$

Y → Model response (injection efficiency), $\boldsymbol{\Psi_i}$ → Orthogonal polynomials
x → input variables, $\boldsymbol{\alpha_i}$ → expansion coefficients

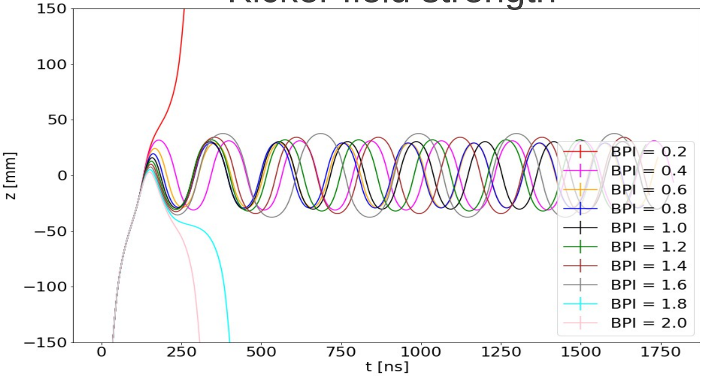
- Polynomial basis based on input variable distribution
- Coefficients determined using regression based methods

$$\vec{\alpha} = \text{Argmin} \frac{1}{N} \sum_{j=1}^{N} \left\{ f(\vec{\xi}^j) - \sum_{i=0}^{P-1} \alpha_i \Psi_i(\vec{x}^j) \right\}^2$$

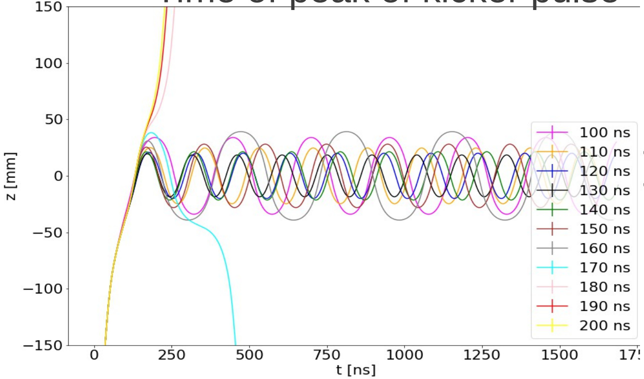
Initial Parameter Bounds

- Preliminary optimization with 8 parameters with range determined using rough scans in G4Beamline with the reference particle

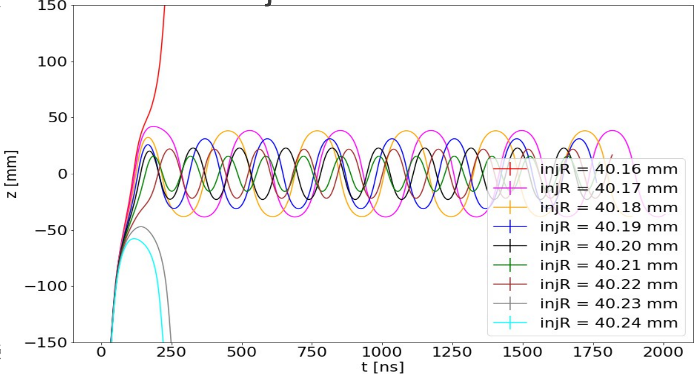
Kicker field strength



Time of peak of kicker pulse



Injection radius

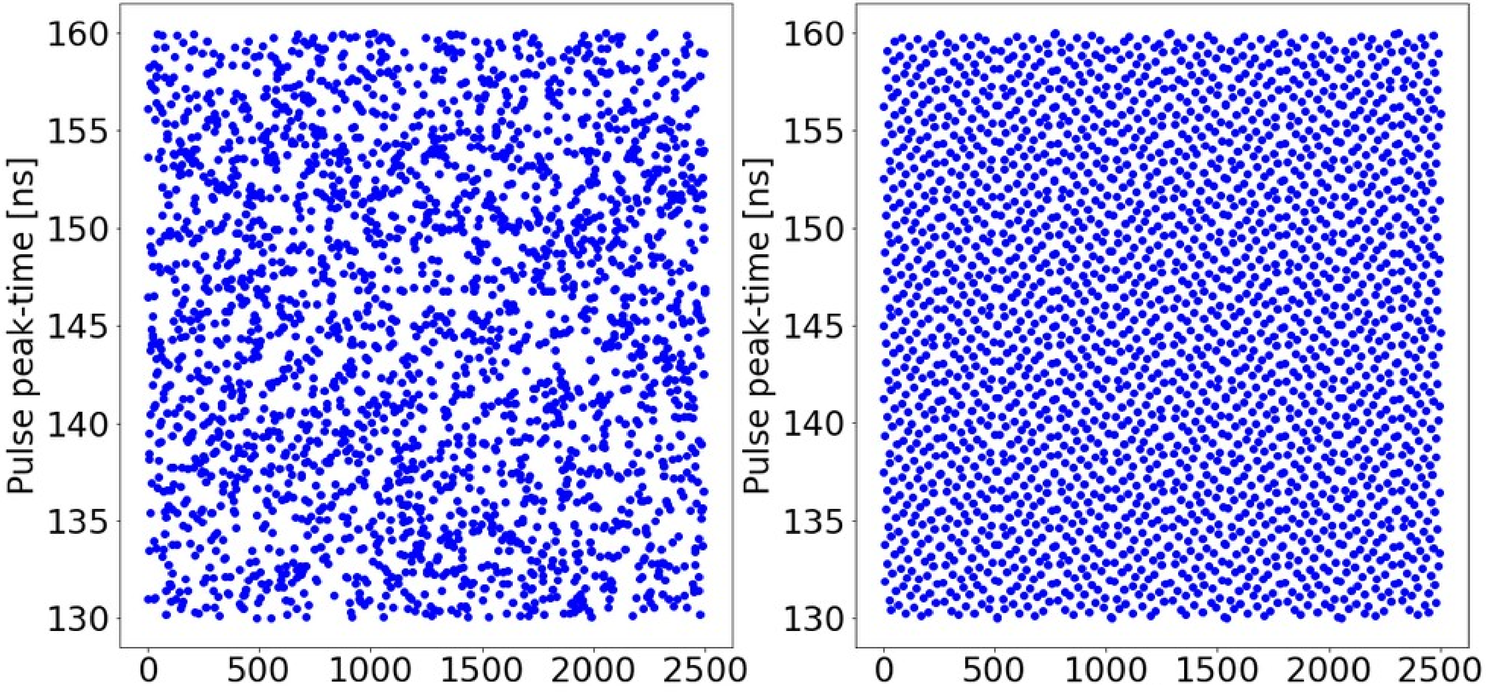


- Best guess input parameter range:

Parameter symbol	Physics Parameter	Lower bound	Upper bound
Theta	Injection angle (degrees)	47.38	47.45
Phi	Transverse angle (degrees)	5.4	5.9
InjR	Injection radius (mm)	40.0	40.4
Z	Longtudinal Injection coordinate (mm)	-437	-433
BPI	Kicker Field Strength (arb. unit)	-1.4	-0.2
KPT	Time of peak of Kicker Pulse (ns)	130	160
W	Width of the Kicker Pulse (ns)	15	55
WC	Weak Current*100 (A)	100	700

Initial Distribution: Sobol Points

- Monte Carlo techniques rely on random distribution of samples
- Prone to clusters and empty spaces, slow convergence, probabilistic error bounds



Random distribution | Sobol distribution

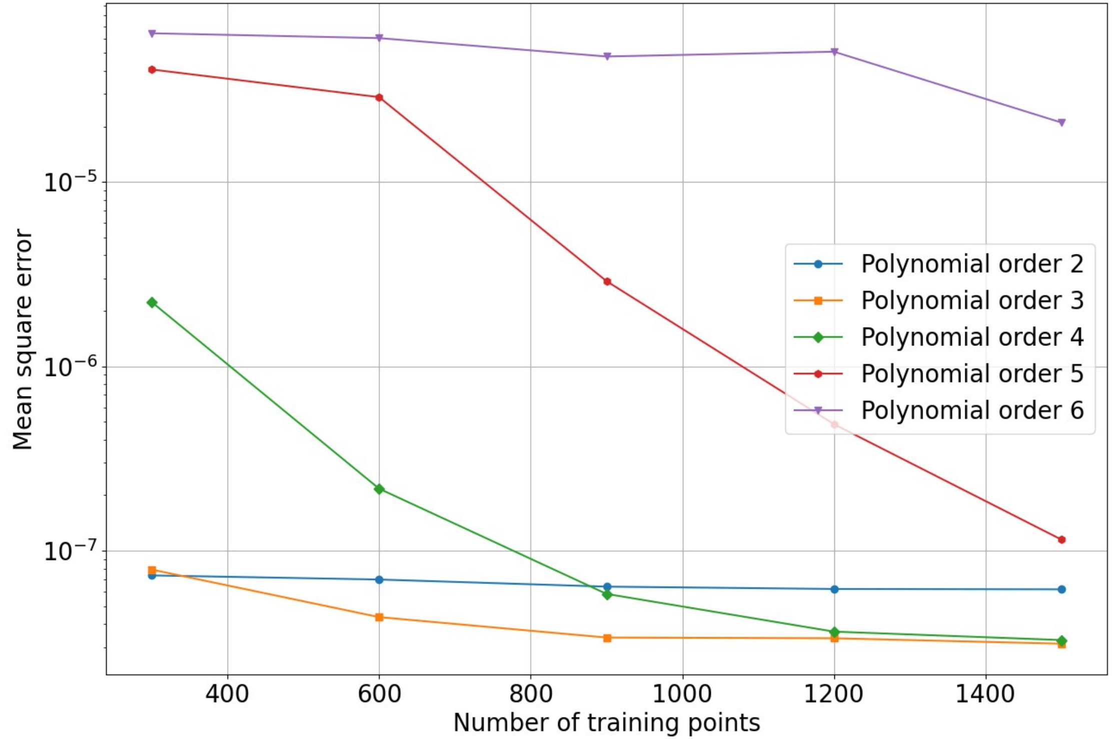
- Low discrepancy distributions like Sobol sequence preferred, deterministic error bounds, uniformly spans the given range

PCE Regression Analysis

- Run G4bl simulations with parameter sample space given by Sobol distribution
- For regression based estimation of coefficients, number of training samples is **N=(d-1)P**;
- P → terms in polynomial expansion: $P = \frac{(p+d)!}{p!d!}$, d → dimension, p → polynomial degree
- Storage criteria:

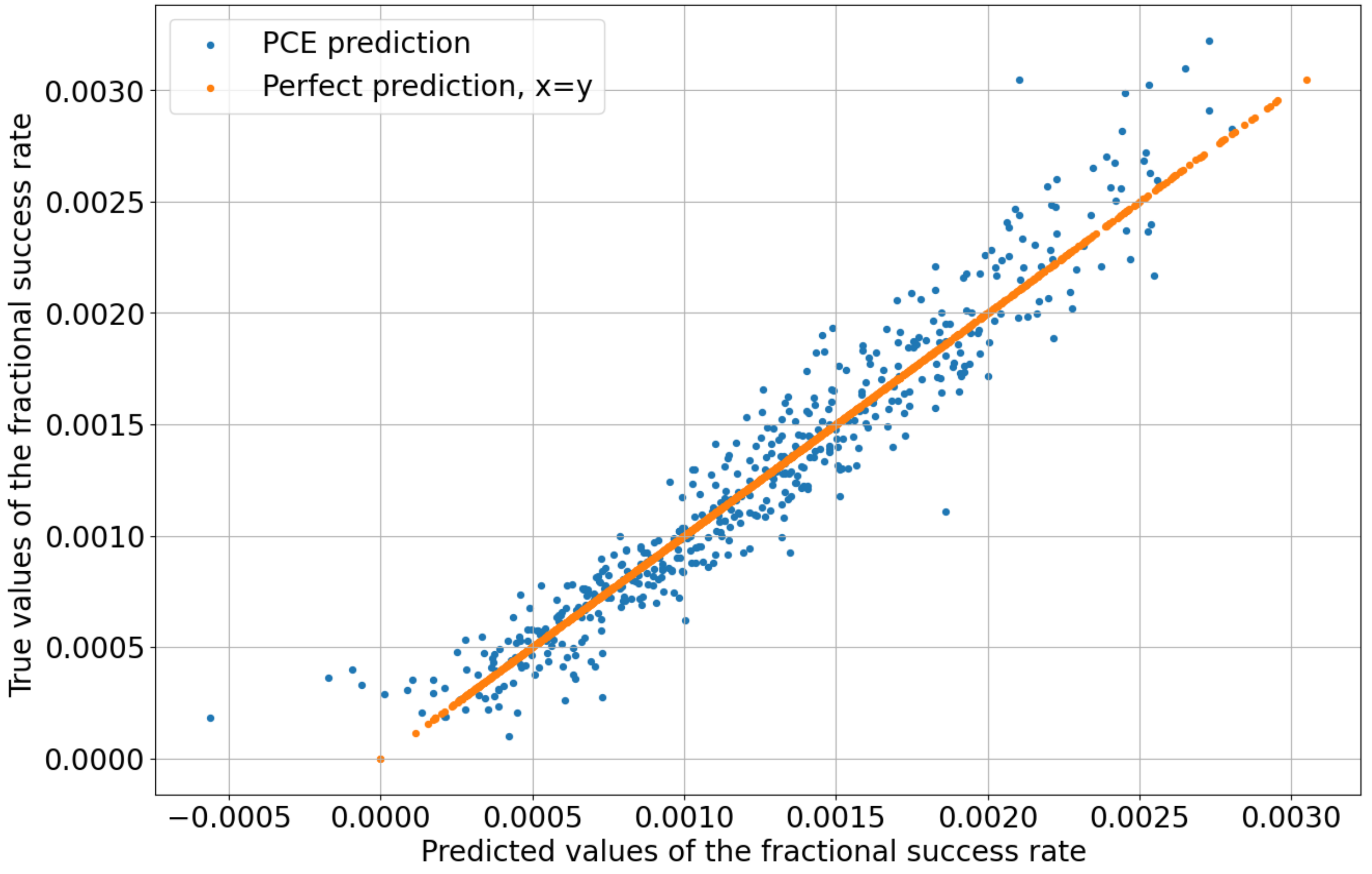
  muon decaying/killed
  with **-40mm<z<40mm**
  and **-0.5MeV<Pz<0.5MeV**

- Python toolbox chaospy for generating expansion and regression fit
- Lowest MSE for poly order 3, poly order 4 converging fast
- ~1100 points needed for poly order 3, ~3500 points needed for poly order 4

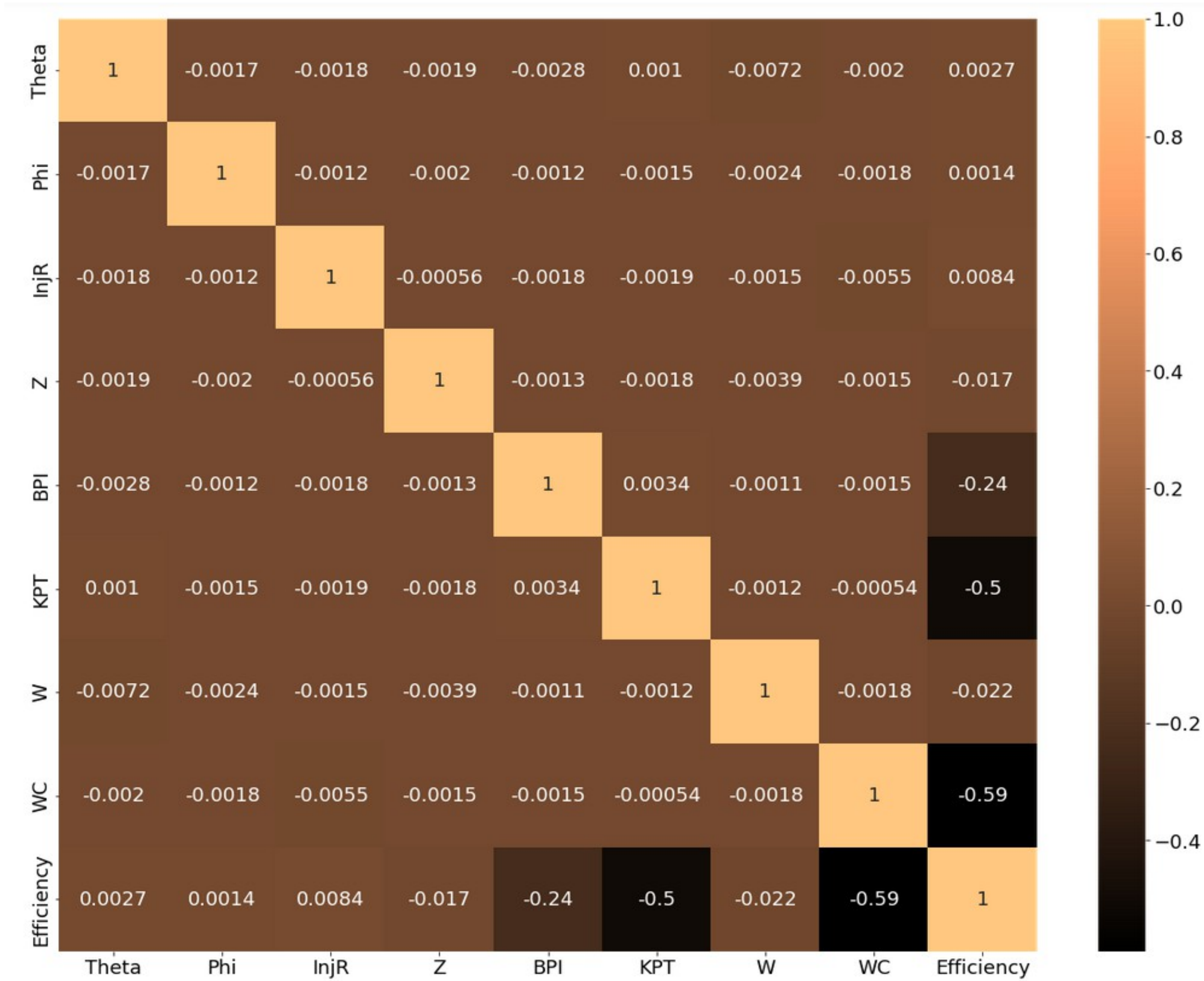


PCE Regression Analysis

- Comparison of true efficiency (orange) and model predicted efficiency (blue) for polynomial order 3

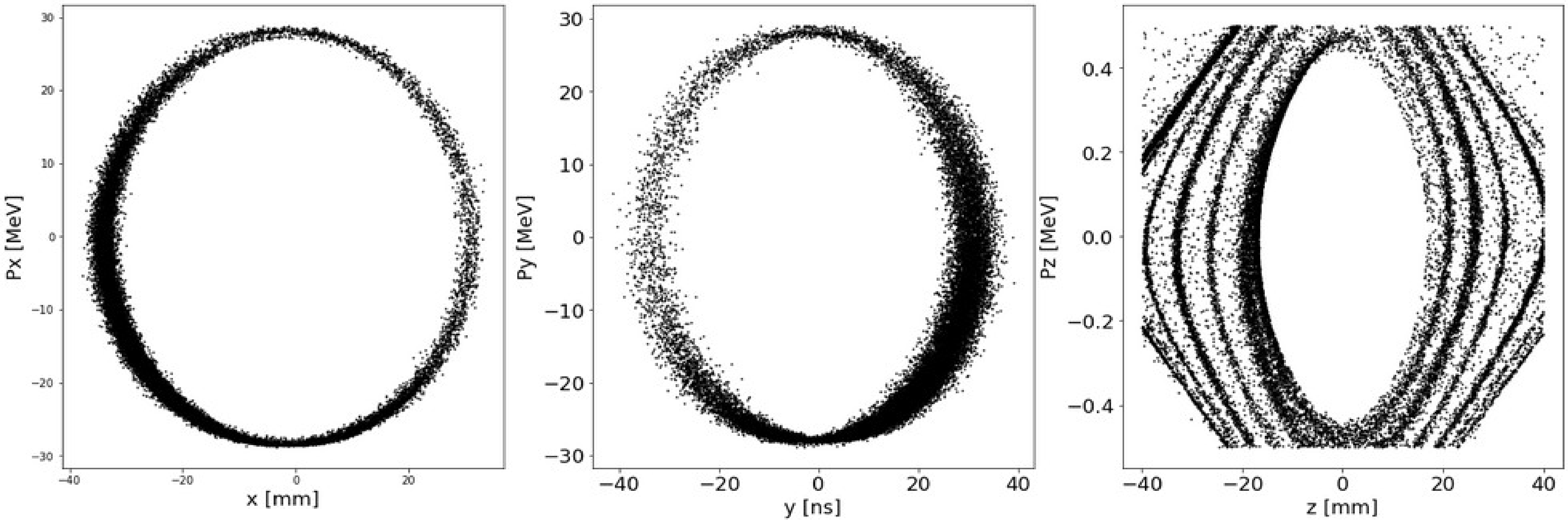


PCE Analysis Correlation Matrix



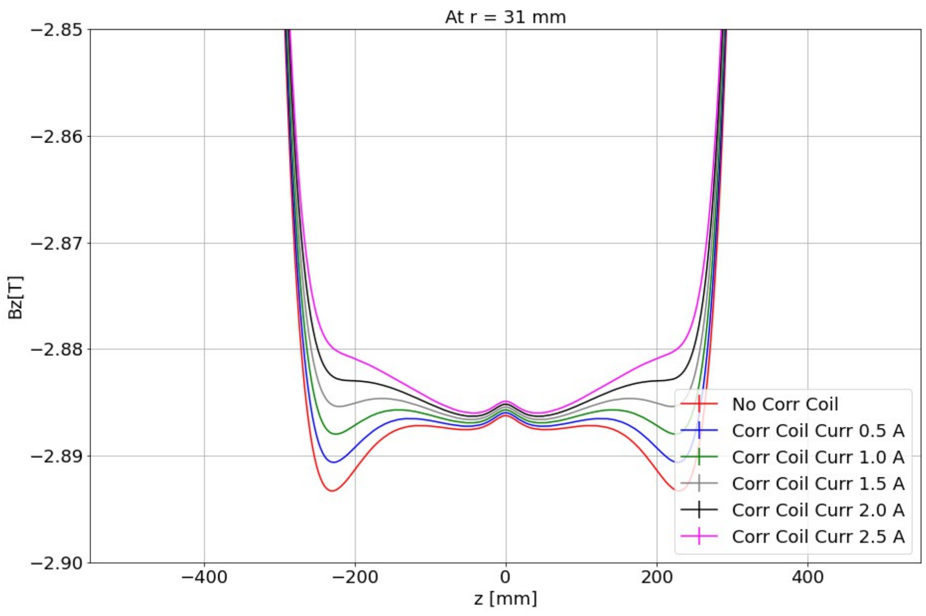
Current Status

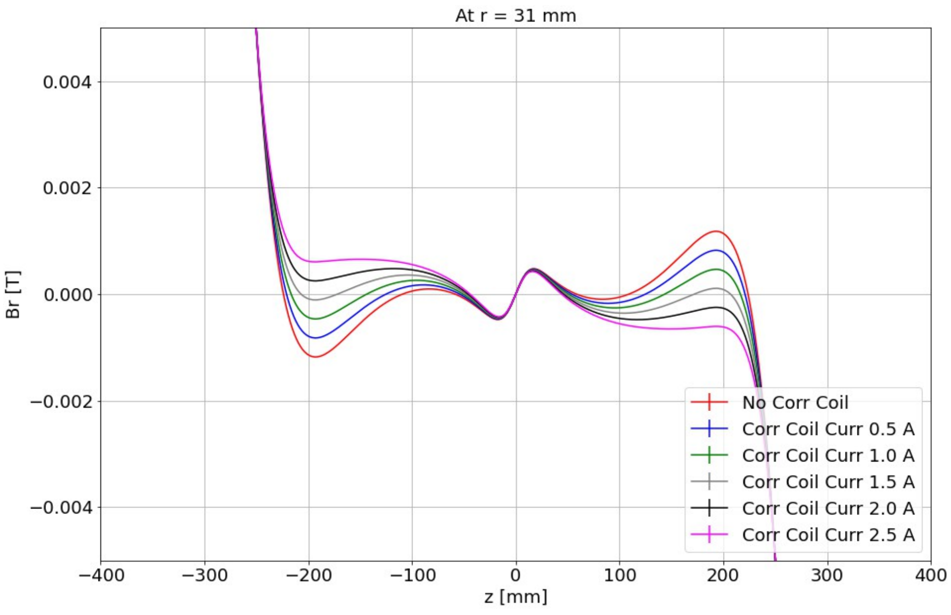
- Need to generate muon distribution in storage region for detector design
- Currently trying to use the maximum efficiency parameters to generate distribution
- 3000 stored from 1M injected muons → Extrapolating to generate larger distributions



Current Status

- Currently running the next iteration of optimization → Excluded geometric parameters
- New parameter → distance between the pair of kicker coils, current in the correction coil
- Correction coils needed to reduce field gradient at injection





- Reduction in CC current desirable to control heat dissipation

Summary

- Injection efficiency depends on a range of design parameter
- Off-axis beam injection simulation in G4Beamline
- Maximize injection efficiency, determine acceptable input variable range
- Multivariate optimization problem → approximated by PCE based surrogate modeling
- Best modeled by polynomial degree 3, less sensitivity to geometric parameters
  → increase in training samples needed to reduce MSE
- Generate muon distribution in storage region based on the maximum efficiency parameters
- Optimize other input variables affecting injection efficiency

* * *

Thank You for Your Attention!

MuEDM Collaboration

ETH · PAUL SCHERRER INSTITUT PSI · UNIVERSITY OF LIVERPOOL · JG|U · UNIVERSITÉ DE GENÈVE · University of Zurich[UZH] · SAPIENZA UNIVERSITÀ DI ROMA · INFN Istituto Nazionale di Fisica Nucleare · UCL · CERN · Cluster of Excellence PRISMA+ · US University of Sussex · tu technische universität dortmund · MANCHESTER 1824 The University of Manchester · 上海交通大学 SHANGHAI JIAO TONG UNIVERSITY · 李政道研究所 Tsung-Dao Lee Institute · TUM · BROOKHAVEN NATIONAL LABORATORY "a passion for discovery" · UNIVERSITÀ DI PISA

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
State Secretariat for Education,
Research and Innovation SERI

Extra

Sensitivity indices: Example, when d=3

Response function representation:
$Y = f(x_1,x_2,x_3) = f_0 + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3) + f_{123}(x_1,x_2,x_3)$

Total varaince:
$D = \int_{Id} f^2(x_1,x_2,x_3)dx_1\, dx_2\, dx_3 - f_0^{\ 2}$

$= D_1 + D_2 + D_3 + D_{12} + D_{13} + D_{23} + D_{123}$

with
$D_1 = \int f_1^{\ 2}(x_1)\, dx_1$, $D_2 = \int f_2^{\ 2}(x_2)\, dx_2$ …....

$D_{12} = \int f_{12}^{\ 2}(x_1,x_2)\, dx_1 dx_2$, $D_{23} = \int f_{23}^{\ 2}(x_2,x_3)\, dx_2 dx_3$,........ and so on

Sensitivity indices: $S_1=D_1/D$, $S_{12}=D_{12}/D$, $S_{123}=D_{123}/D$, …. etc

and

$S_1 + S_2 + S_3 + \ldots + S_{123} = 1$

Sensitivity Analysis in Surrogate Model

PCE expansion: $Y = \sum_{i=0}^{\infty} \alpha_i \Psi_i(\vec{x})$, $(\vec{x}) = (x_1, x_2, \dots, x_d)$, d-dimension

For input variables

$$(x_1, x_2, \dots, x_d) \in I^d := [0,1]^d$$

Model response Y can be represented by

$$Y = f(x_1, x_2, .., x_d) = f_0 + \sum_{i=1}^{d} f_i(x_i) + \sum_{1 \le i < j \le n} f_{ij}(x_i, x_j) + \dots + f_{12\dots d}(x_1, x_2, \dots, x_d),$$

if:

1. $f_0 = \int_{I^d} f(\vec{x}) d\vec{x}$

2. $\int_0^1 f_{i_1, i_2, \dots, i_s}(x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_k = 0$

where $k = i_i, \dots, i_d$ and $s = 1, 2, \dots, d$

The variance of the function is:

$$D = \int_{I^d} f^2(\vec{x}) d\vec{x} - f_0^2$$

Alternatively $D = \sum_{i=1}^{d} D_i + \sum_{1 \le i < j \le n} D_{ij} + \dots + D_{12\dots d}$

where

$$D_{i_1, i_2, \dots, i_s} = \int_{I^s} f_{i_1, i_2, \dots, i_s}^2 (x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_{i_i} dx_{i_2} \dots dx_{i_s}$$

with $1 \le i_1 < i_s \le d$

Sobol sensitivity indices are defined as:

$$S_{i_1, i_2, \dots, i_s} = \frac{D_{i_1, i_2, \dots, i_s}}{D}$$

and satisfy the condition

$$\sum_{i=1}^{d} S_i + \sum_{1 \le i < j \le n} S_{ij} + \dots + S_{12\dots d} = 1$$

Sensitivity indices: Results

- Python Chaospy toolbox has built-in functions to calculate sensitivity indices
- First order, second order and total sensitivity indices can be calculated
- The input parameter distribution should be scaled to range [0,1]
- Main Sensitivity indices for poly order 3:

Parameter symbol	Physics Parameter	Main sensitivity index, Polynomial degree 3
Theta	Injection angle (degrees)	3.39230879e-08
Phi	Transverse angle (degrees)	3.86979007e-07
InjR	Injection radius (mm)	1.88544294e-06
Z	Longtudinal Injection coordinate (mm)	4.16966789e-08
BPI	Kicker Field Strength (arb. unit)	1.40754537e-04
KPT	Time of peak of Kicker Pulse (ns)	9.03986321e-04
W	Width of the Kicker Pulse (ns)	4.89511202e-02
WC	Weak Current*100 (A)	9.12374997e-01

Sum = 0.9624

- Relative sensitivities, dependent on initial range provided

Sensitivity indices: From PCE coefficients

$$\hat{S}_{i_1,i_2,\ldots,i_s} = \frac{1}{\hat{D}} \sum_{\boldsymbol{i} \in I_{i_1,i_2,\ldots,i_s}} \alpha_{\boldsymbol{i}}^2 \langle \Psi_{\boldsymbol{i}}^2 \rangle$$

where

$$I_{i_1,i_2,..,i_s} = \left\{ \boldsymbol{i} : \begin{array}{l} i_k > 0 \forall k = 1, \ldots, n, k \in (i_1, \ldots, i_s) \\ i_k = 0 \forall k = 1, \ldots, n, k \notin (i_1, \ldots, i_s) \end{array} \right\}$$

The variance, $\hat{D}$ is given by

$$\hat{D} = \sum_{j=1}^{P-1} \alpha_{\boldsymbol{i}_j}^2 \langle \Psi_{\boldsymbol{i}_j}^2 \rangle.$$

From this, the main and total sensitivity is computed as

$$\hat{S}_j = \frac{1}{\hat{D}} \sum_{\boldsymbol{i} \in I_j} \alpha_{\boldsymbol{i}}^2 \langle \Psi_{\boldsymbol{i}}^2 \rangle$$

where $I_j = \{\boldsymbol{i} = (i_1, i_2, \ldots, i_d) : i_k > 0 \forall k = j \wedge i_k = 0 \forall k \neq j\}$ and

$$\hat{S}_j^T = \frac{1}{\hat{D}} \sum_{\boldsymbol{i} \in I_j} \alpha_{\boldsymbol{i}}^2 \langle \Psi_{\boldsymbol{i}}^2 \rangle$$

where $I_j = \{\boldsymbol{i} = (i_1, i_2, \ldots, i_d) : i_j > 0\}$, respectively.

Neural Network Regression

- No. of neurons: **512**

  Hidden layers: **8**

  Activation function: **Leaky ReLU**

  Scheduler: **ReduceLROnPlateau**

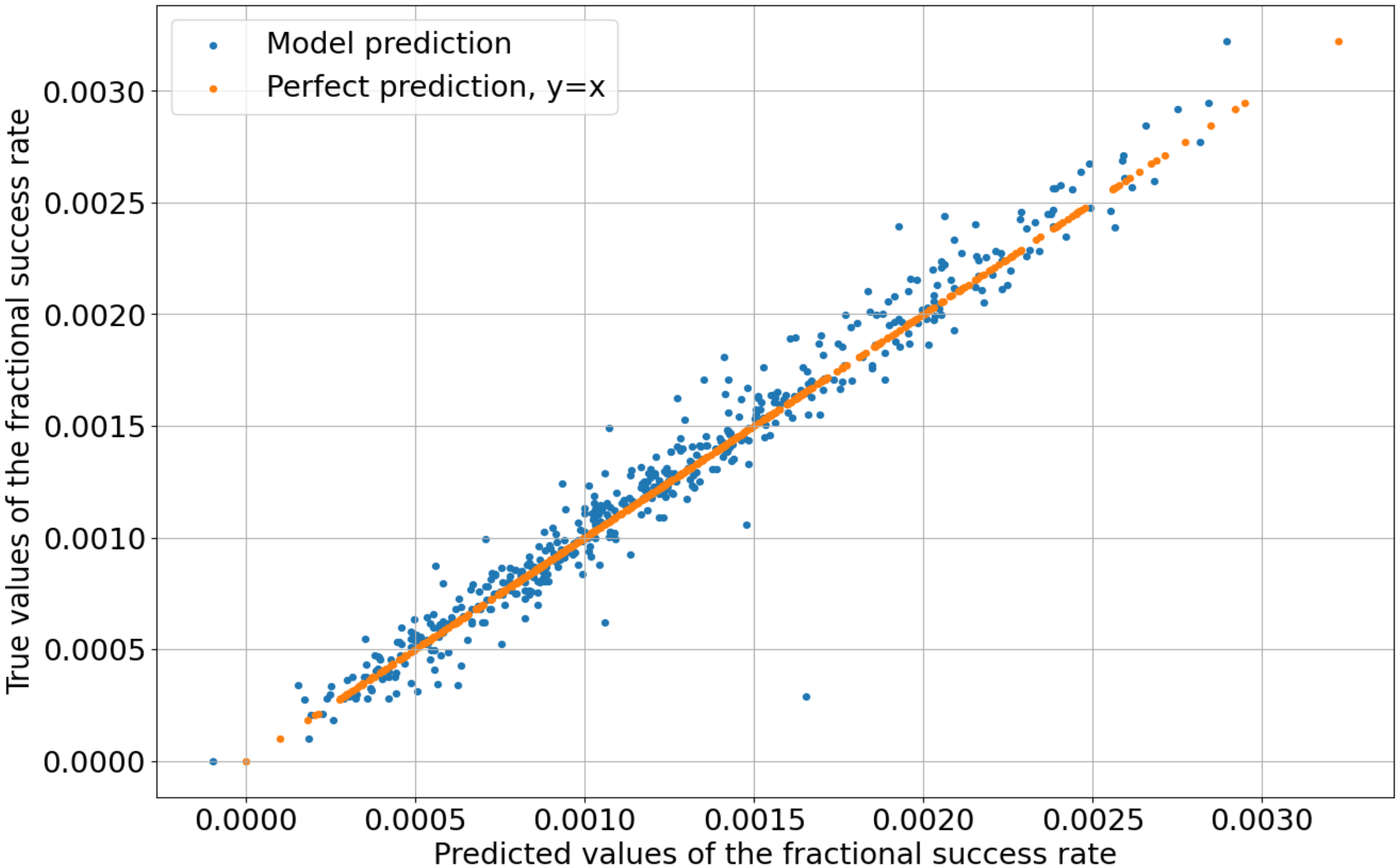
  Batch size: **500**

  Epoch: **400**

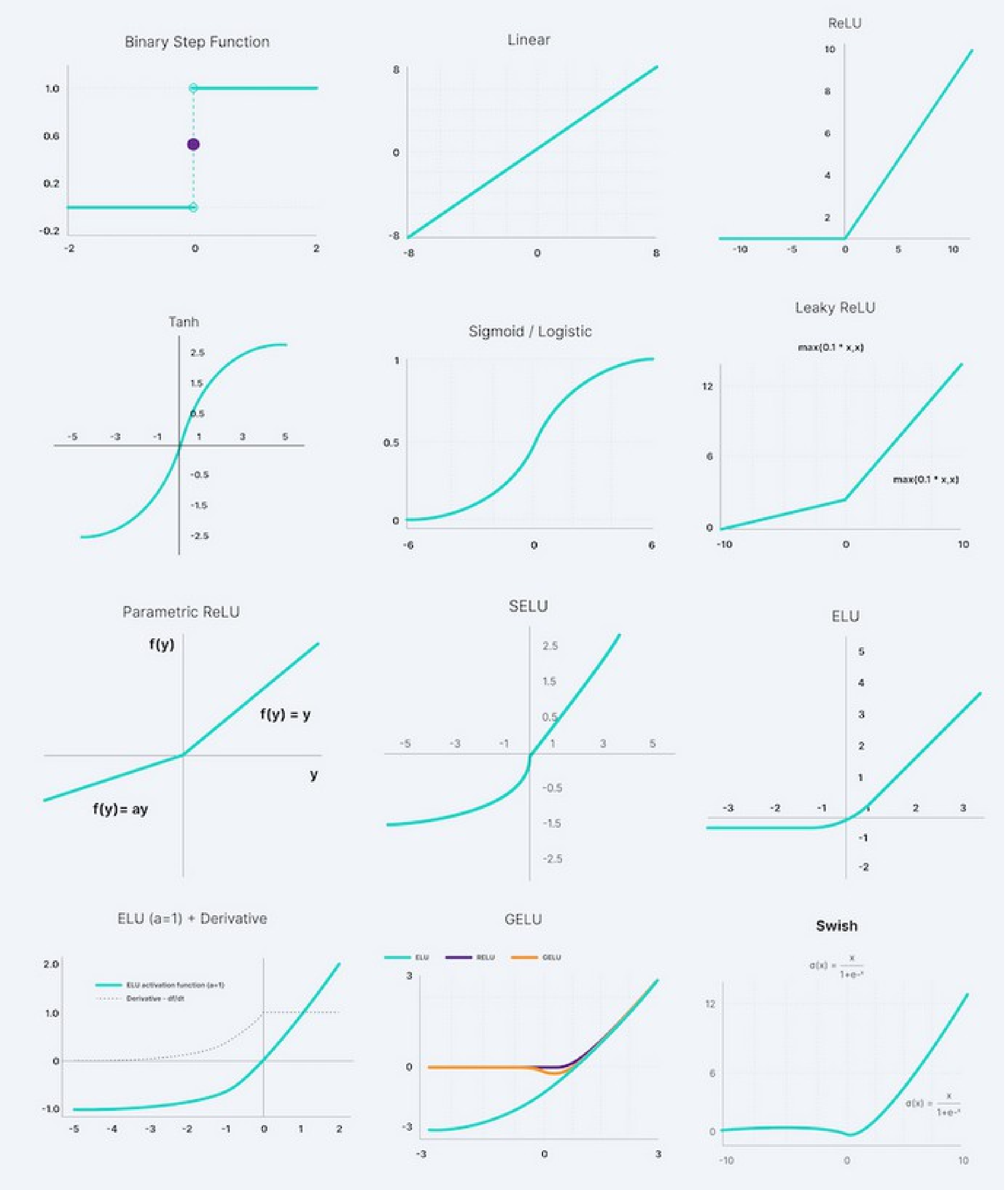
  Optimizer: **Adam**

  Dropout rate: **10%**

- Better agreement between true and predicted



Neural Network activation functions



Number of expansion coefficients, P and Integration points (sample size), N

d	p	P	N
6	3	84	420
6	4	210	1050
6	5	462	2310
7	3	120	720
7	4	330	1980
7	5	792	4752
8	3	165	1155
8	4	495	3465
8	5	1287	9009