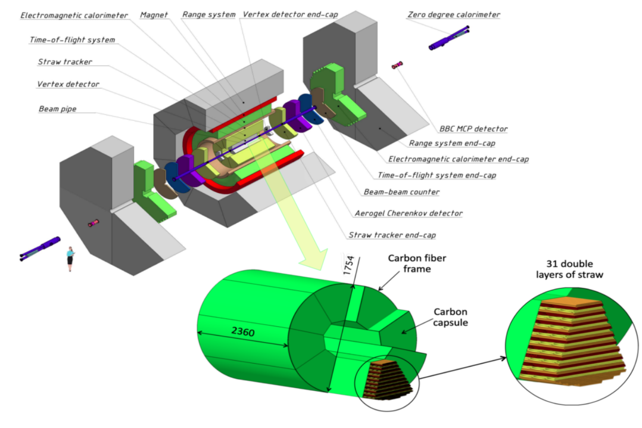
Realistic simulation and hit reconstruction for the Straw Tracker

E. Mosolova, V. Bautin, S. Bulanova, A. Mukhamejanova
Supervisors: Katerina Kuznetsova, Temur Enik, Viktor Kim

November 8, 2024

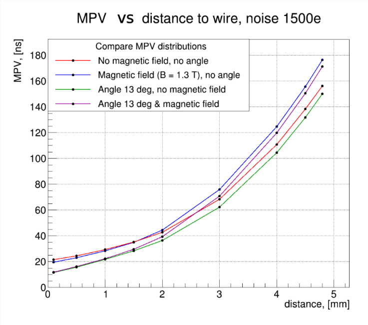
PNPI | VIII SPD collaboration meeting

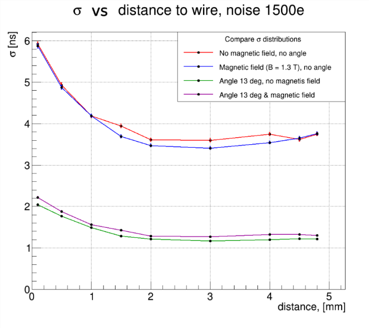
Straw Tracker – the main tracking system of SPD



Straw diameter: 10mm thickness 36μm PET
Barrel is made of 8 modules with up to 31 double-layers, with the ZUV orientation $(0^\circ, +3^\circ, -3^\circ)$

2023| Sonya B. & Vitalii B. parameterized mode and variance of the straw signal registration time distribution by Garfield++/LTSpice





Straw diameter: 10 mm

Anode diameter: 30 mkm

Gas mixture: Ar+CO2 / 70:30 [%]

Gas gain = 4.5E4

Peaking time 25 ns

Signal amplification 3 mV/fC

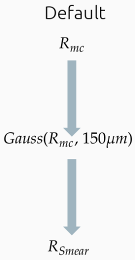
Noise is implemented, Threshold 10 mV

VMM3-based readout

model by Vitalii B.

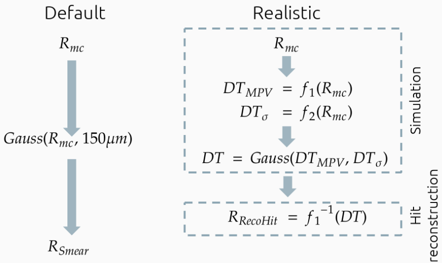
Source: Diploma by Sonya B.

By default SPDROOT accounts for the final straw resolution by smearing the MC hit coordinates



- Monte Carlo Point was smearing in an almost infinite while loop with a fixed variance of 150 μm

We introduced the realistic signal parameterization and hit reconstruction in the SPDROOT

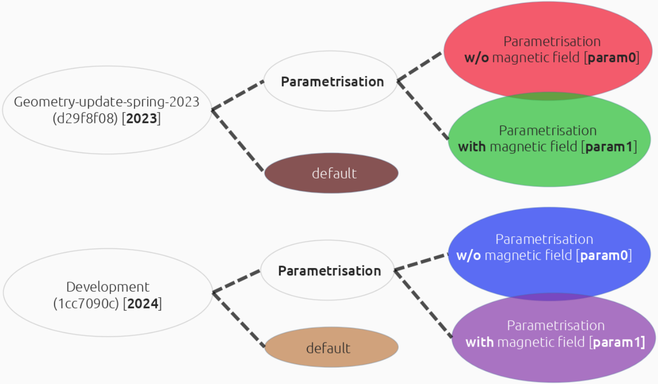


- The distribution of the drift time (DT) is provided by Sonya B. & Vitalii B.
- The DT is calculated for each Monte Carlo point
- Afterward, DT is smeared by $\sigma(DT) = f(R_{MC})$
- Roots of the inverse function (parabola) provide $R_{RecoHit}$

Two branches of SPDROOT were taking for comparison

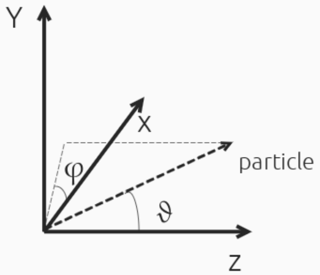
1. Geometry-update-spring **2023**: $\sigma(R_{MC})$ is const $= 150\mu$m
2. Development **2024**: $\sigma(R_{MC})$ is $0.06506 * exp(-3.26 * R_{MC})$
   implemented by R.Akhunzyanov

The comparision includes simuls by two versions of parametrisation not acounting for the magnetic field (param0) and accounting for the magnetic field 1.3T (param1)

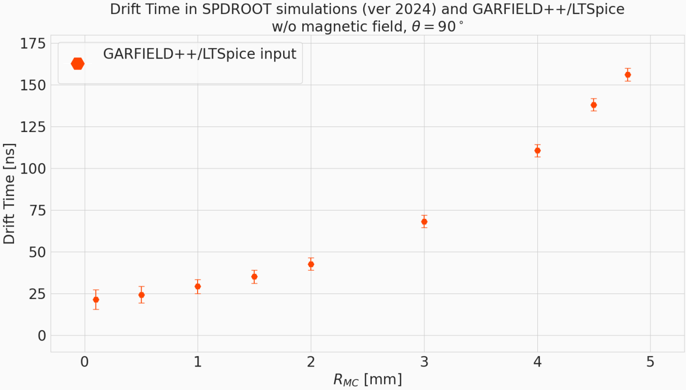


Simulation settings

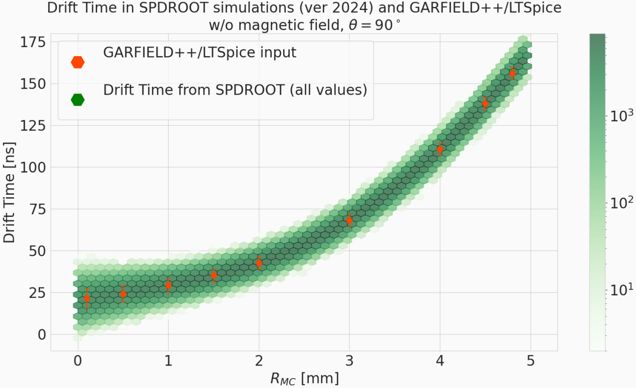
- **Patricle:** muon (μ, pdg = 13)
- **Energy:** 1GeV
- **Generator:** SpdIsotropicGenerator
  - θ**:** is angle between Z-axis and beam (**now we used** $\theta = 90^\circ$)
  - ϕ**:** From 0° to 360°
- **Detectors:** Only Straw Barrel
- **Vertex:** Off
- **Magnet:** field_full1_8.bin
- **Events:** 10k



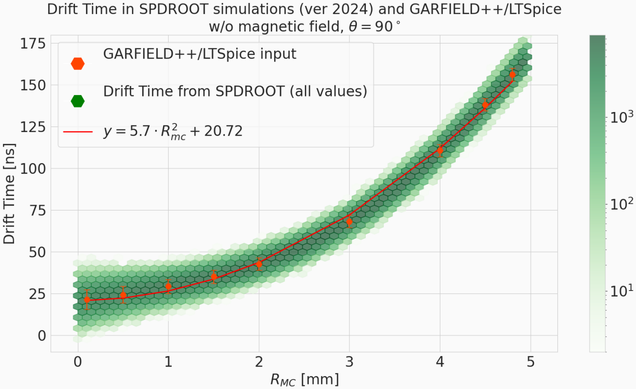
The distribution of the drift time (DT) is provided by Sonya B. Vitalii B.



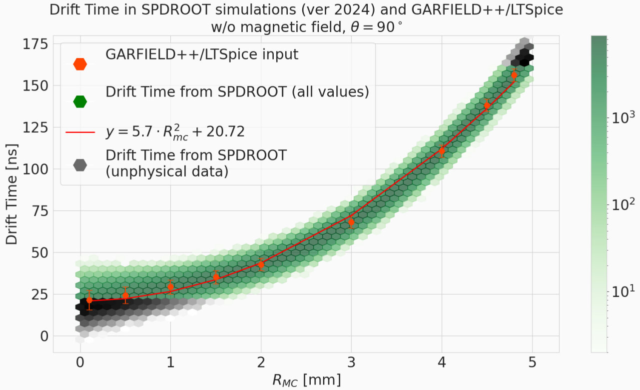
The DT is calculated for each Monte Carlo point and smeared



Calibration curve for hit reconstruction

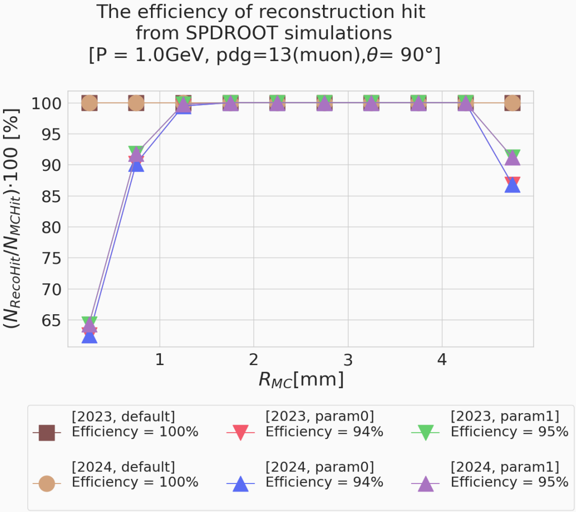


4% of hits are lost near the anode
Less than 1% is reconstructed outside the tube



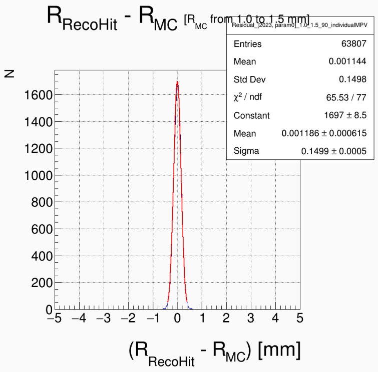
Therefore, the accuracy of hits position estimation is an object of utter importance.

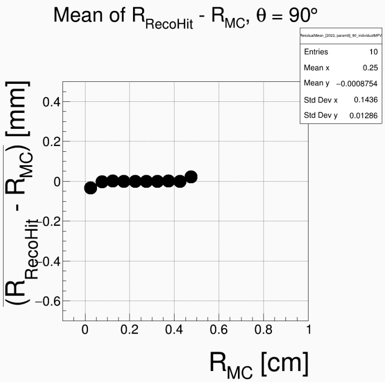
Reconstruction efficiencies for param0/1 difference are the same



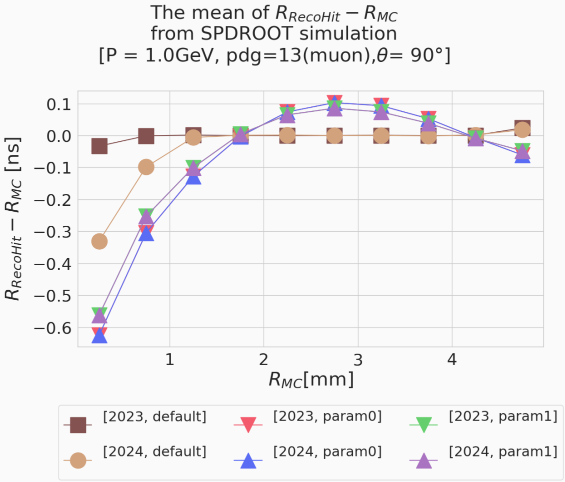
Distributions of residuals over areas R_{mc}
We are considering 10 ranges [mm]: [0.0–0.5), [0.5–1.0), [1.5–2.0), etc.

This is default version 2023

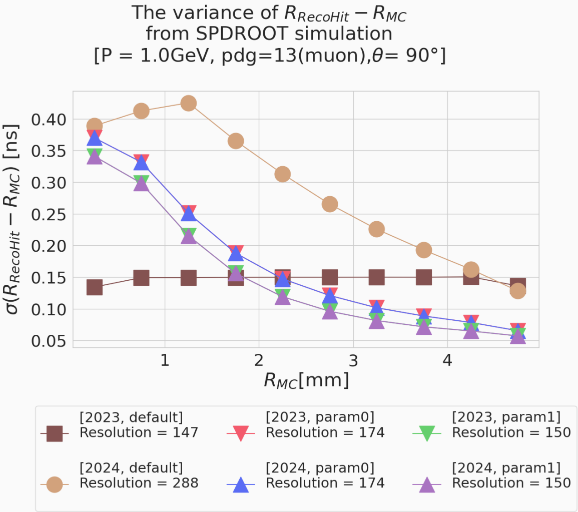




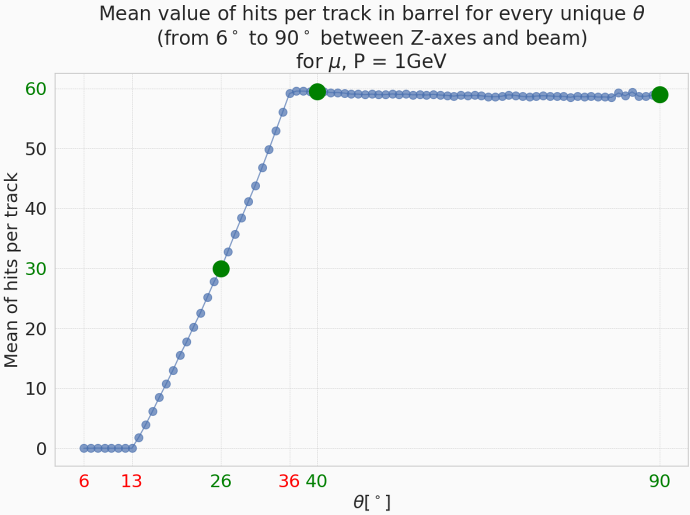
Bias Analysis of hit reconstruction: Default vs. Parametric Versions.



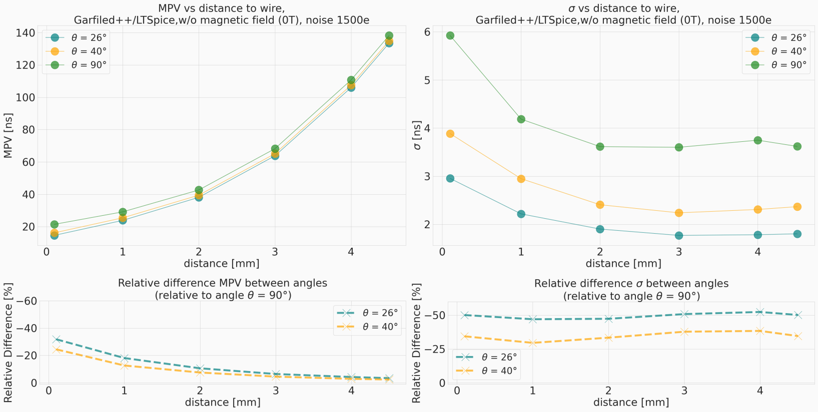
Variance of $R_{MC} - R_{RecoHit}$



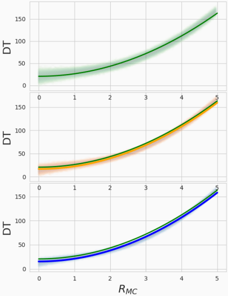
Next, we'll analyze the behavior at an angle with 30 hits per track (26°) and at the left edge of the saturation plateau (40°).

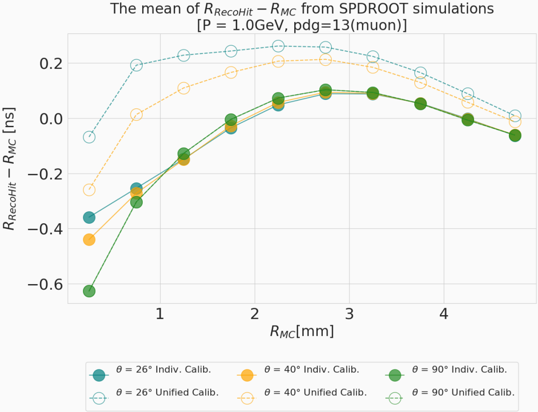


Difference near anode $> 20\% \rightarrow$ switch to individual calibration curves

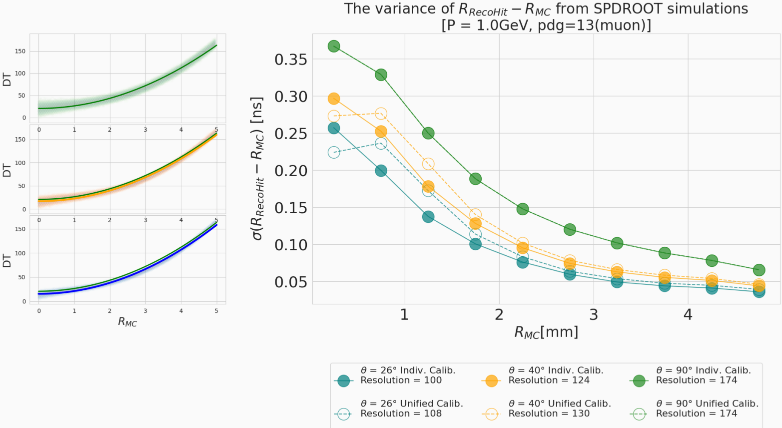


Individual calibration curve: less bias, higher resolution
Bias increases with the unified calibration curve (90°)





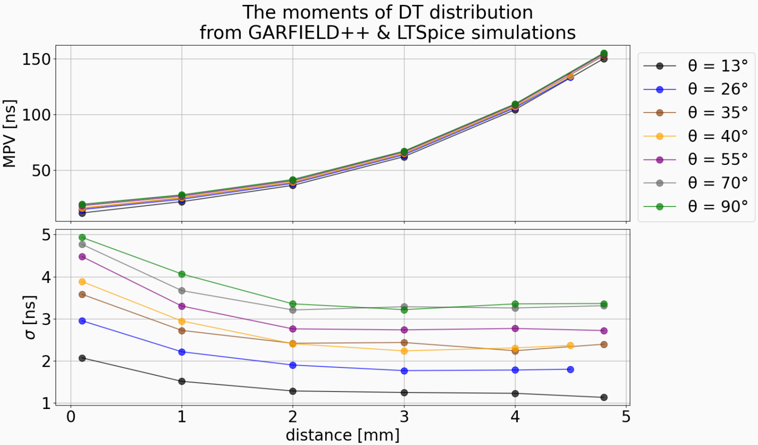
Individual calibration curve: less bias, higher resolution
Bias increases with the unified calibration curve (90°)



Expanding parameterization capabilities

- Next step we implement general parameterization as a function of R_{MC} and θ
- We performing 2D fitting of drift time distribution moments
- This will help calculate mean and variance based on θ and R_{MC}
- We applied a second degree polynomial regression for this fitting

Maximum variance at 90° muon emission angle



Modeling the relationship between mpv and σ based on the variables R_{mc} and θ

The linear regression model is trained using the method of least squares. The goal is to find such coefficients β that minimize the sum of squared differences between the predicted and actual values of the target variable.

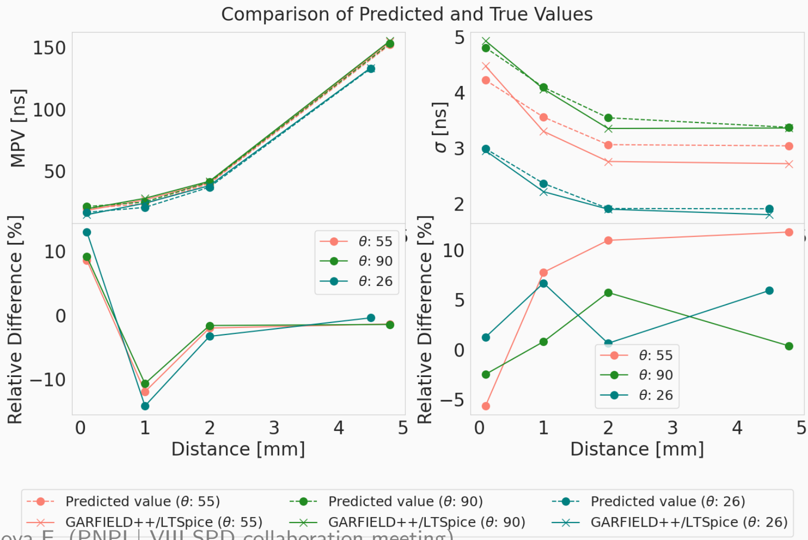
The mathematical formula of the **second degree polynomial regression** model is:

$$data = \beta_0 + \beta_1 \cdot R_{mc} + \beta_2 \cdot \theta + \beta_3 \cdot R_{mc}^2 + \beta_4 \cdot (R_{mc} \cdot \theta) + \beta_5 \cdot \theta^2 \quad (1)$$

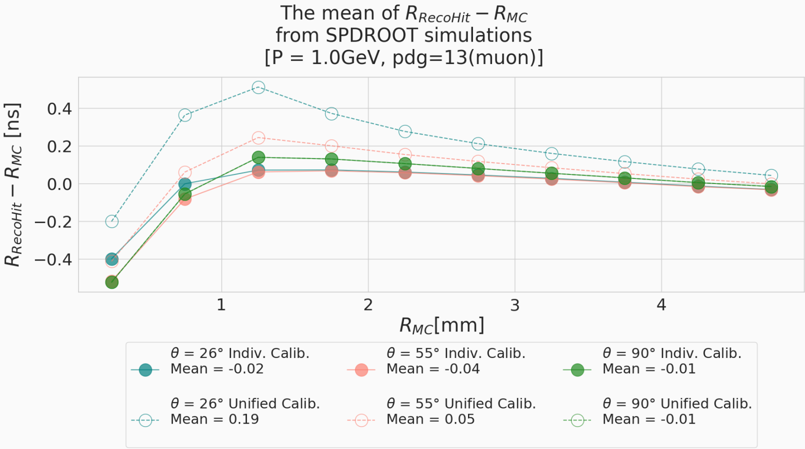
Where:

- (β_0) — intercept (constant term)
- $(\beta_1, \beta_2, ..., \beta_5)$ — model coefficients that are learned from the data

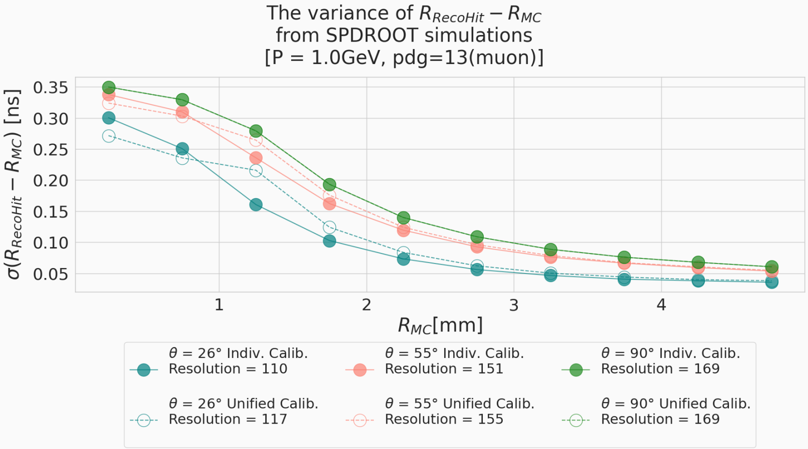
Fit quality control $\theta = 55^\circ$: Unused in training, over 10% variance difference



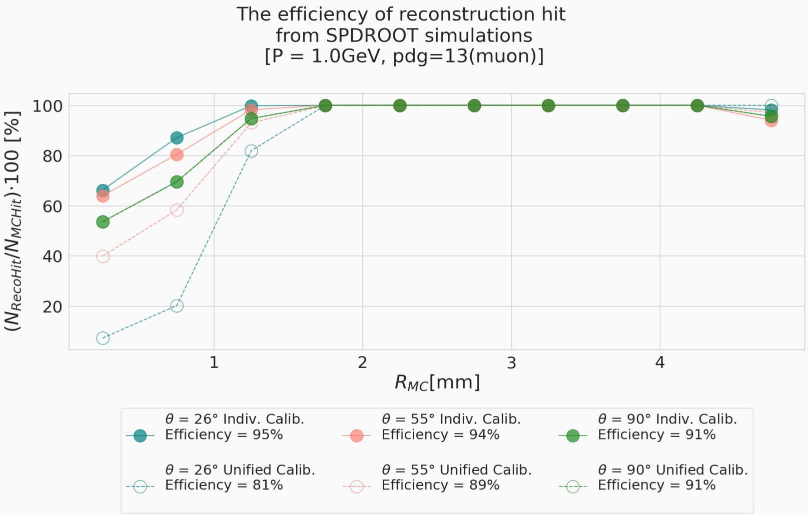
Unified calibration curves markedly influence MVP diminishing effects at the periphery



Unified calibration curves have a minor effect on sigma 6.5% resolution difference at 26°



Individualized calibration curves near the anode boost reconstruction efficiency by 17% across the radius



Summary

- **Signal Parameterization:** Straw signal parameterization by Sonya B., Assel M. and Vitalii B. was implemented in SPDROOT for several angles.
- **SPDROOT:**
  - **Parameterization:** A comparison of different parameterization approaches, considering with and without the magnetic field, was conducted.
  - **Hit Reconstruction:** A straw hit reconstruction procedure has been introduced into SPDROOT.
- `https://git.jinr.ru/nica/spdroot/-/tree/Straw-Signal-Parameterisation`

Result:

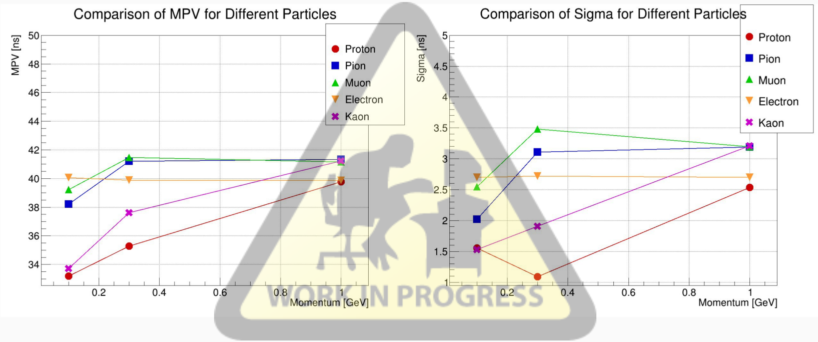
- Both parameterization approaches seem to work similarly, so we'll use the parameterization without considering the magnetic field for future work
- A current simple approach of the hit reconstruction gives a bias up to 100μm and changes with the radius
- The work on improving the hit reconstruction processing is ongoing
- For the current realistic simulation of VMM3-based readout model, the average resolution for 90° is 150μm
- Use of unified calibration curve results in high bias

Result:

- We introduce parametrization based on muon sample with a current model of VMM3-readout
- Work on more realistic readout model and studies with different particles types are ongoing
  - We do not account for the finite TDC (Time-to-Digital Converter) resolution in the electronics. Evaluation of the TDC resolution influence is ongoing
  - Continuing to refine the tube description in LTSpice

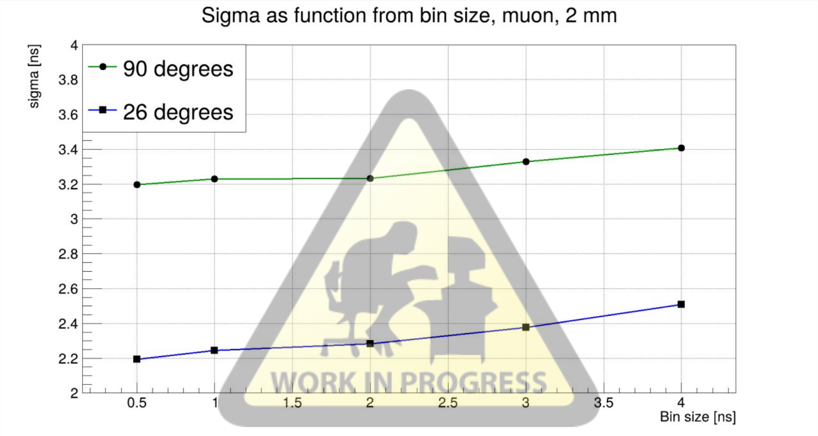
For the 2 mm point

1. There may be a difference in MPV for low energy particles
2. It seems that other particles will give better resolution



Thank you for your attention!

bckp


Sigma as function from bin size, muon, 2 mm
sigma [ns]
4
3.8
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
0.5
1
1.5
2
2.5
3
3.5
4
Bin size [ns]
90 degrees
26 degrees
WORK IN PROGRESS

Transforming Features into Polynomial Features

The class `PolynomialFeatures` (from Python package `sklearn`) is used with the parameter `degree=2` to create new features that are polynomial combinations of the original features up to the second degree.

New features include:

- Constant term (1)
- Linear features: ($R_{mc},\ theta$)
- Quadratic features: (R_{mc}^2, θ^2)
- Interaction of features: ($R_{mc} \cdot \theta$)

In total, for two features and a polynomial degree of 2, we obtain 6 features.

Predicting Values

After training, the model uses the learned coefficients to predict the target variable based on new values of (R_{mc}) and (θ).

Model: Second Degree Polynomial Regression.

Goal: Modeling the relationship between (mpv) and (sigma) based on the variables (R_{mc}) and (θ).

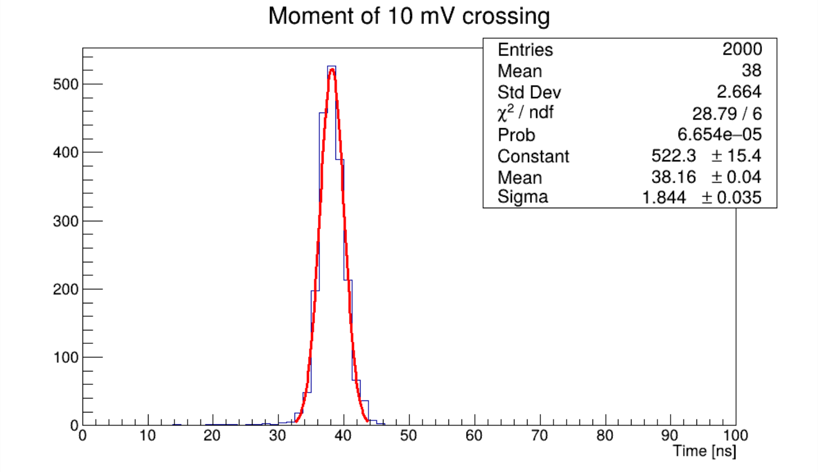
Mathematical methods:

- Least squares method for determining model coefficients.
- Polynomial features to account for nonlinear relationships.
- Multiple linear regression with an extended set of features.

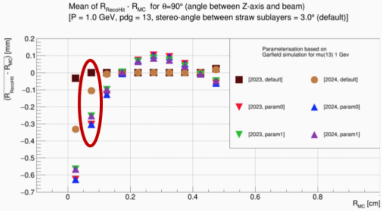
Predicting Values

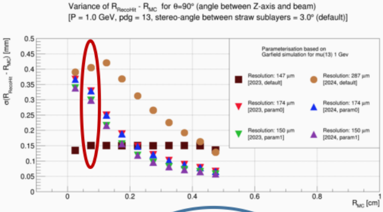
Process:

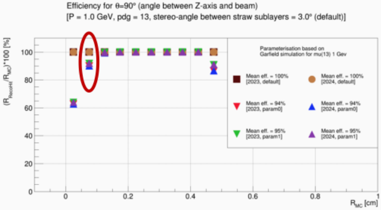
1. Feature transformation: creating polynomial features of degree 2.
2. Model training: using linear regression to train on polynomial features.
3. Model evaluation: using metrics (MSE, (R^2)) to assess model quality.
4. Formula extraction: extracting model coefficients and writing out the explicit formula for the relationship.
5. Prediction: using the model to predict (mpv) and (sigma) on new data.

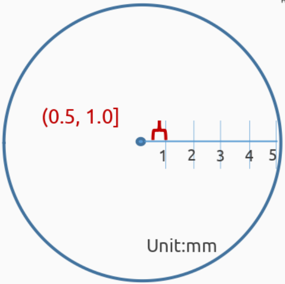

Moment of 10 mV crossing
Entries 2000
Mean 38
Std Dev 2.664
χ2 / ndf 28.79 / 6
Prob 6.654e−05
Constant 522.3 ± 15.4
Mean 38.16 ± 0.04
Sigma 1.844 ± 0.035
0
100
200
300
400
500
0
10
20
30
40
50
60
70
80
90
100
Time [ns]

Let's look at interesting areas: from 0.5 to 1.0;

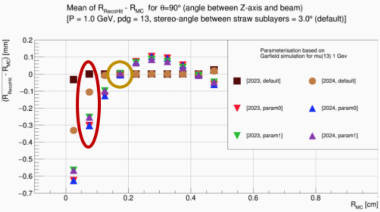


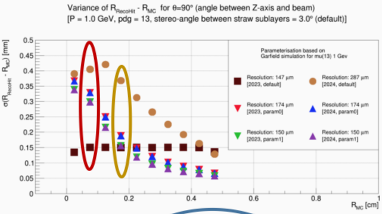


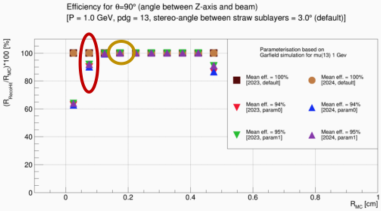


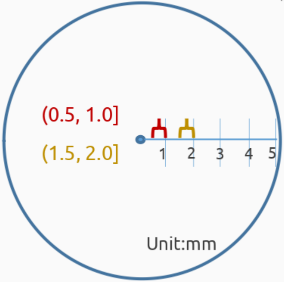


Let's look at interesting areas: from 0.5 to 1.0; from 1.5 to 2.0

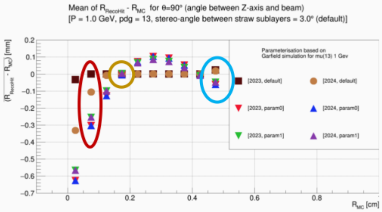


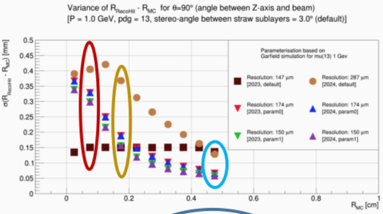


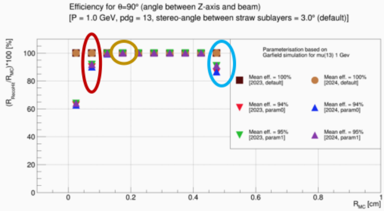


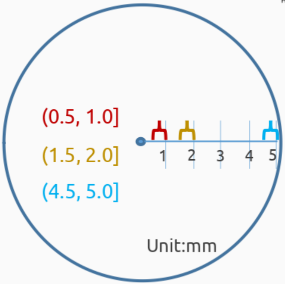


Let’s look at interesting areas: from 0.5 to 1.0; from 1.5 to 2.0 and from 4.5 to 5.0 mm

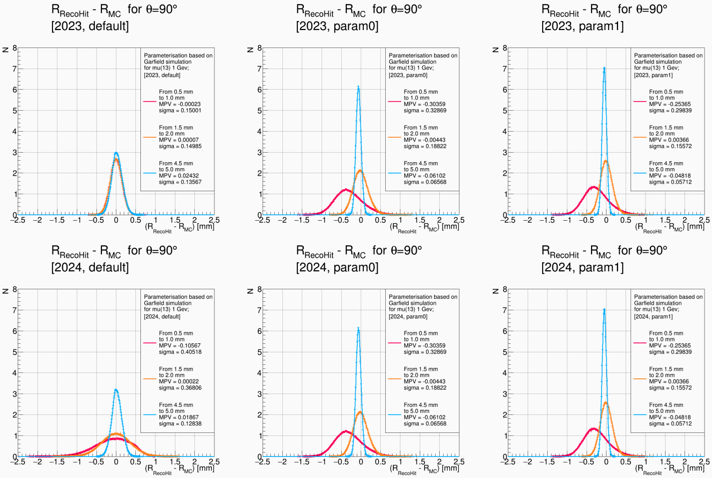




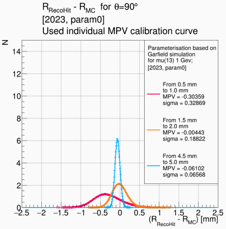


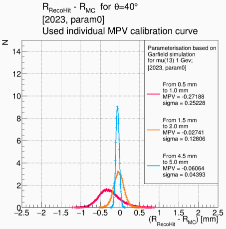


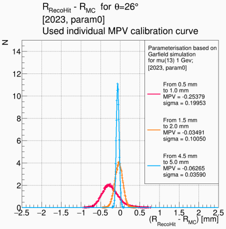
Residual for three selected point

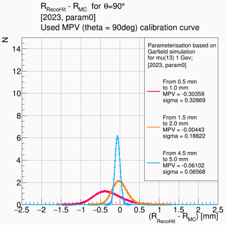


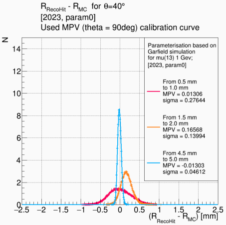
2023 param0 | Residual for three selected point | Bias issues with a single calibration curve

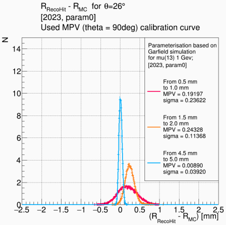




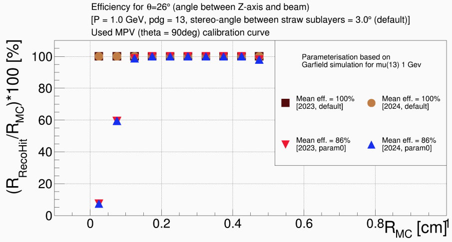


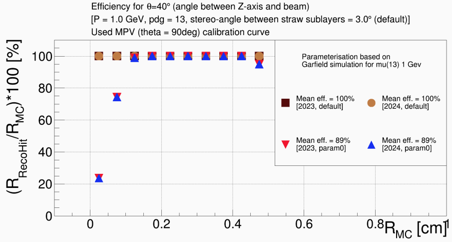


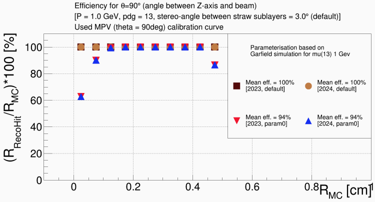




Efficiency


Efficiency for θ=26° (angle between Z-axis and beam)
[P = 1.0 GeV, pdg = 13, stereo-angle between straw sublayers = 3.0° (default)]
Used MPV (theta = 90deg) calibration curve
(R_RecoHit/R_MC)*100 [%]
R_MC [cm]
Parameterisation based on
Garfield simulation for mu(13) 1 Gev
Mean eff. = 100% [2023, default]
Mean eff. = 100% [2024, default]
Mean eff. = 86% [2023, param0]
Mean eff. = 86% [2024, param0]


Efficiency for θ=40° (angle between Z-axis and beam)
[P = 1.0 GeV, pdg = 13, stereo-angle between straw sublayers = 3.0° (default)]
Used MPV (theta = 90deg) calibration curve
(R_RecoHit/R_MC)*100 [%]
R_MC [cm]
Parameterisation based on
Garfield simulation for mu(13) 1 Gev
Mean eff. = 100% [2023, default]
Mean eff. = 100% [2024, default]
Mean eff. = 89% [2023, param0]
Mean eff. = 89% [2024, param0]


Efficiency for θ=90° (angle between Z-axis and beam)
[P = 1.0 GeV, pdg = 13, stereo-angle between straw sublayers = 3.0° (default)]
Used MPV (theta = 90deg) calibration curve
(R_RecoHit/R_MC)*100 [%]
R_MC [cm]
Parameterisation based on
Garfield simulation for mu(13) 1 Gev
Mean eff. = 100% [2023, default]
Mean eff. = 100% [2024, default]
Mean eff. = 94% [2023, param0]
Mean eff. = 94% [2024, param0]

Efficiency: details

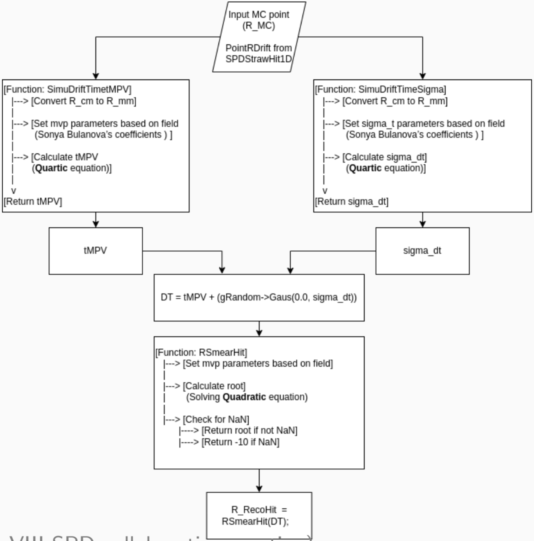
To calculate the efficiency in the range of R_{mc} from 0.0 to 0.5 cm, the total number of R_{mc} was counted, then it was calculated how many of these R_{mc} were reconstructed:

$Eff = \frac{N_{RecoHit}}{N_{totalOfHits}}$

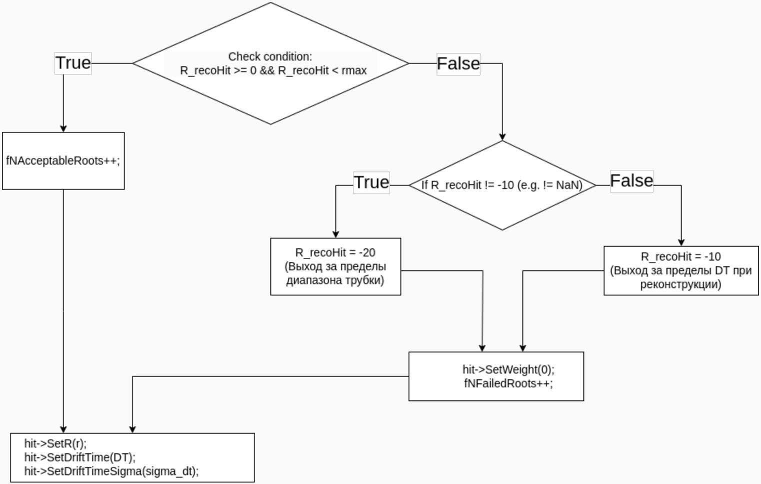
The efficiency of the parameterized version is lower than in the default version.

No reco in default.

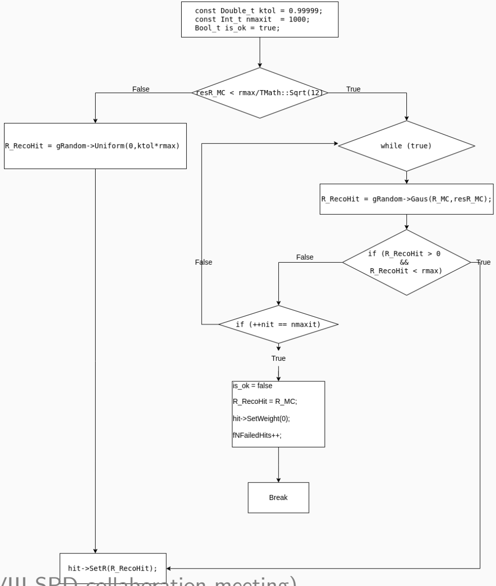
1. Drift time (DT) from R_{mc} and Garfield's simulations
2. $R_{RecoHit}$ from DT



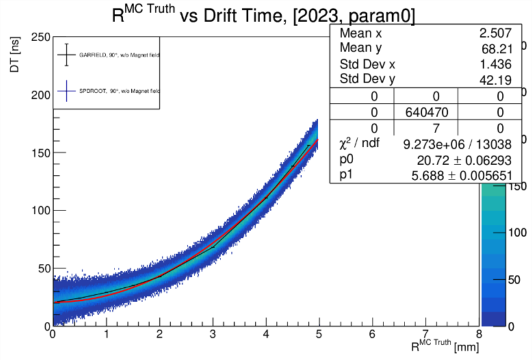
Create hit position in param0/1

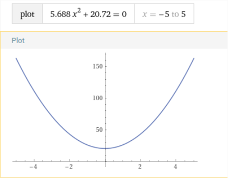


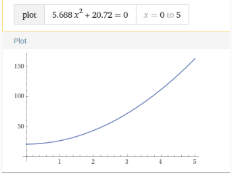
Create hit position in default



Coeff for quadratic equation (param0)



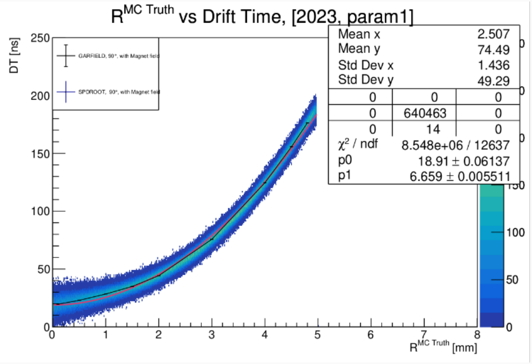


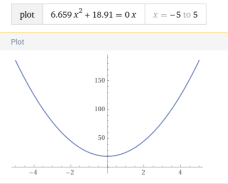


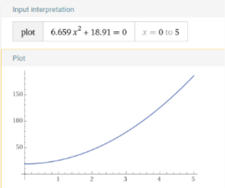
Derivative

$$\frac{d}{dx}\left(5.688\,x^2 + 20.72\right) = 11.376\,x$$

Coeff for quadratic equation (param1)







Derivative

$$\frac{d}{dx}(6.659\,x^2 + 18.91) = 13.318\,x$$