JOINT INSTITUTE FOR NUCLEAR RESEARCH
Dzhelepov Laboratory of Nuclear Problems

FINAL REPORT ON THE START PROGRAMME

Determination of pp-collision time and particle identification for SPD NICA

Supervisor:
Dr. Mikhail Zhabitsky

Student:
Semyon Yurchenko, Russia
Saint Petersburg State University

Participation period:
July 17 – September 03, Summer Session 2022

Dubna, 2022

Contents

1 Abstract

Identification of final particles is a very important task in high energy physics experiments. If final particles are identified, reconstruction of decayed particles($\phi - meson, \Lambda - baryon, D - meson$) becomes easier. In order to identify final particles Time-Of-Flight(TOF) detectors are used. TOF detectors measure only time when particle cross the detector, but time of collision is also needed. In this work an algorithm to reconstruct time of the pp-collision and a PID procedure will be presented.

2 Introduction

Usage of TOF detectors to identify particles is a most reliable technique. To reconstruct particle mass only three parameters are needed: $\vec{p}$ - momentum of the particle, L - arc length of the particle trajectory and TOF - Time-Of-Flight. If those parameters are known, mass of particle is given by the simple formula:

$$m^2 = p^2\left(\frac{c^2 \cdot TOF^2}{L^2} - 1\right) \tag{2.0.1}$$

Where c - is speed of light. TOF can be calculated as a difference between a time mark in TOF detector - t_i and a moment of pp-collision - t_0:

$$TOF = t_i - t_0 \tag{2.0.2}$$

While t_i can be measured directly, t_0 can not. So it is needed to create procedure that will reconstruct t_0 by using information about particles trajectories and time marks in TOF detector. This procedure should fulfill two requirements - it should be precise and fast. In section 4 it will be shown that it is hard to meet those requirements at the same time. And after determination of the pp-collision time, a particle identification(PID) procedure should be developed.

3 Project Goals

During START program two goals should be reached. First is to develop an algorithm for fast and precise t_0 reconstruction and compare it with the Brute Force Algorithm, which should be implemented as well. Second is two perform PID with different methods and compare them.

4 Determination of t_0

4.1 Setup

Initial conditions:

1. Collision energy - $\sqrt{s} = 27\ GeV$
2. PYTHIA8 - Monte Carlo generator where collision will be simulated.
3. SPD experiment parameters [1]:
   (a) Resolution of TOF detector - $\sigma_{TOF} = 70\ ps$,
   (b) Momentum resolution - $\frac{\sigma_p}{p} = 2\%$,
   (c) Detector has cylindric shape with radius - $r = 1\ m$ and half-length - $l = 1,886\ m$,
   (d) Magnetic field - $B = 1\ T$(uniform, solenoidal).

Plan for this project:

1. Simulate collisions in PYTHIA8,
2. Calculate the intersection point of tracks with the TOF detector,
3. Calculate TOF and arc length of trajectory,
4. Smear TOF with $N(TOF,\ \sigma_{TOF})$ and p with $N(p,\ 0.02p)$,
5. Reconstruct t_0 with smeared data,
6. Perform PID.

4.2 PYTHIA8 and intersection point

For simulation PYTHIA8 Monte Carlo generator was chosen. To obtain tracks two types of processes were chosen: $HardQCD : all$ and $SoftQCD : all$. Processes selection had no effect on algorithms performance.
All charged particles were propagated through the magnetic field. Intersection points with TOF detector were analytically calculated:
Trajectory in the magnetic field(in z direction) of point charge is a spiral:

$$x(t) = x_0 + r \cdot sin(\omega t + \alpha) \tag{4.2.1}$$

$$y(t) = y_0 + r \cdot cos(\omega t + \alpha) \tag{4.2.2}$$

$$z(t) = z_0 + v_z \cdot t \tag{4.2.3}$$

with parameters:

$$\omega = \frac{e \cdot c \cdot B}{E} \tag{4.2.4}$$

$$r = \frac{c \cdot p_\perp}{e \cdot B} \tag{4.2.5}$$

Here: e - charge of particle($+1$ for proton, -1 for electron), B - magnetic field, $p_\perp$ - transverse momentum of the particle, $v_z = \frac{c \cdot p_z}{E}$ - particle velocity in z - direction.

Parameters $x_0,\ y_0,\ z_0,\ \alpha$ will be determined from initial conditions for coordinates. For $t = 0$:

$$x(0) = 0 = x_0 + r \cdot sin(\alpha) \tag{4.2.6}$$

$$y(0) = 0 = y_0 + r \cdot cos(\alpha) \tag{4.2.7}$$

$$z(0) = 0 = z_0 \tag{4.2.8}$$

Initial conditions for velocities:

$$v_x(0) = v_\perp \cdot cos(\alpha) = v_x^{Pythia} \tag{4.2.9}$$

$$v_y(0) = -v_\perp \cdot sin(\alpha) = v_y^{Pythia} \tag{4.2.10}$$

$$\alpha = arccos\left(\frac{v_x^{Pythia}}{v_\perp}\right) = arcsin\left(-\frac{v_y^{Pythia}}{v_\perp}\right) \tag{4.2.11}$$

Now we obtain that:

$$x_0 = -r \cdot sin(\alpha) \tag{4.2.12}$$

$$y_0 = -r \cdot cos(\alpha) \tag{4.2.13}$$

Physical meaning of $x_0,\ y_0$ - coordinates of the center of spiral.

Intersection can occur in 2 ways: with thee side surface or with the end of cylinder. First let's consider intersection with side surface of cylinder with radius R.

Condition for intersection is:

$$x^2(t) + y^2(t) = R^2 \tag{4.2.14}$$

$$(x_0 + r \cdot sin(\omega t + \alpha))^2 + (y_0 + r \cdot cos(\omega t + \alpha))^2 = R^2 \tag{4.2.15}$$

$$x_0^2 + 2 \cdot r \cdot x_0 \cdot sin(\omega t + \alpha) + r^2 \cdot sin^2(\omega t + \alpha) + y_0^2 + 2 \cdot r \cdot y_0 \cdot cos(\omega t + \alpha) + r^2 \cdot cos^2(\omega t + \alpha) = R^2 \tag{4.2.16}$$

$$x_0 \cdot sin(\omega t + \alpha) + y_0 \cdot cos(\omega t + \alpha) = \frac{R^2 - 2r^2}{2 \cdot r} \tag{4.2.17}$$

$$\frac{x_0}{r} \cdot sin(\omega t + \alpha) + \frac{y_0}{r} \cdot cos(\omega t + \alpha) = \frac{R^2 - 2r^2}{2 \cdot r^2} \tag{4.2.18}$$

$$\frac{x_0}{r} = sin(\beta), \quad \frac{y_0}{r} = cos(\beta) \tag{4.2.19}$$

$$sin(\beta) \cdot sin(\omega t + \alpha) + cos(\beta) \cdot cos(\omega t + \alpha) = \frac{R^2 - 2r^2}{2 \cdot r^2} \tag{4.2.20}$$

$$cos(\omega t + \alpha - \beta) = \frac{R^2 - 2r^2}{2 \cdot r^2} \tag{4.2.21}$$

$$t = \frac{2 \cdot \pi + arcsin\left(\frac{R^2-2r^2}{2 \cdot r^2}\right) - \alpha - \angle([x_0, y_0],\ x - axis)}{\omega} \tag{4.2.22}$$

For negatively charged particles we obtain:

$$t = \frac{arcsin\left(\frac{R^2-2r^2}{2 \cdot r^2}\right) - \alpha + \angle([x_0, y_0],\ x - axis)}{\omega} \tag{4.2.23}$$

Arc length is:

$$L = v \cdot t \tag{4.2.24}$$

Intersection with the end of cylinder happen when:

$$v_z \cdot t = l \Rightarrow t = \frac{l}{v_z} \tag{4.2.25}$$

4.3 General idea for t_0 determination

After production in pp-collision short-living particles will decay before they reach detector system. At TOF detector possible particle types are: $\pi^{\pm},\ K^{\pm},\ p^{\pm},\ e^{\pm},\ \mu^{\pm}$. So idea is that we can take an event and start inserting those particles in tracks and find the combination that gives the best result. Best result is determined with minimisation of χ^2 function.

$$\chi^2 = \sum^{N} \frac{(t_0 + tof_{ik} - t_i)^2}{\sigma_t^2 + \sigma_{t(p_ik)})^2} \tag{4.3.1}$$

Where tof_{ik} is Time-Of-Flight of i-th track in assumption that it has type k. Time-Of-Flight uncertainty due to uncertainty in momentum σ_{p_ik} is determined:

$$\sigma_{t(p_ik)} = \frac{L}{c} \cdot \frac{m_k^2}{p_i^2} \left(\sqrt{1 + \frac{m_k^2}{p_i^2}} \right)^{-1} \cdot \frac{\sigma_p}{p} = 0.02 \cdot \frac{L}{c} \cdot \frac{m_k^2}{p_i^2} \left(\sqrt{1 + \frac{m_k^2}{p_i^2}} \right)^{-1} \tag{4.3.2}$$

For a certain mass hypothesis an analytic solution for t_0 reads:

$$t_0 = \frac{1}{\mu} \sum^{N} \frac{t_i - tof_{ik}}{\sigma_t^2 + \sigma_{p_i}^2},\ \ \mu = \sum^{N} \frac{1}{\sigma_t^2 + \sigma_{p_i}^2} \tag{4.3.3}$$

So task is now reduced to minimisation of χ^2 by finding the right mass hypotheses for tracks. Sections 4.4 and 4.5 are deducted to developing algorithm to perform minimisation. Due to low abundance of muons and electrons and their similarity in TOF with pions they were excluded from possible particle types. Results with and without electron are shown in section 4.6.

4.4 Brute Force Algorithm

Most straightforward solution is to check all mass hypotheses and find one that has minimal χ^2. If $N_m = 3$ - number of possible masses and N_{tr} - number of tracks in event then number of combinations is $3^{N_{tr}}$ and time complexity of this algorithm will be $O(N_{tr}3^{N_{tr}})$. Exponential time means that this algorithm is very slow and can not be used even for offline analysis. So it is needed to find a way to speed up this procedure.

4.5 Genetic Algorithm

One way is to use some type of Genetic Algorithm. In this work a Differential evolution-inspired [2] (DE-inspired) genetic algorithm was used. In this algorithm few control parameters are used: N_{pop} - size of the population, C_r - crossover rate, F - mutation scale factor. It means that algorithm will work if we change external parameters like resolutions of detectors or colliding particle types. In this work following parameters were used: $N_{pop} = 15$, $C_r = 1$, $F = 1$.
General procedure is:

1. Create population with N_{pop} random candidates solutions. Candidate solution is random set of masses associated with tracks,
2. Begin mutation process:
   (a) Choose three random solution vectors and create a mutant vector:
   $$v_{mut} = v_r + F \cdot (v_p - v_q), \tag{4.5.1}$$
   (b) Because $C_r = 1$ crossover operation means that we change v_r with v_{mut},
   (c) Calculate t_0^{mut} and χ^2_{mut} (see 4.3.1 and 4.3.2),
   (d) Compare χ^2_r and χ^2_{mut},
   (e) If $\chi^2_{mut} < \chi^2_r$ - new mutant vector is better than parent. Otherwise population remain unchanged. This step is called Darwinian selection.
3. After some number of steps, solution with smallest χ^2 is chosen as an answer.

Time complexity of Genetic Algorithm is $O(N_{tr} \cdot N_{pop} \cdot N_{steps})$, where $800 < N_{steps} < 1000$.

4.6 Comparison

Brute Force Algorithm finds exact solution of χ^2 minimisation and is used as reference to check performance of Genetic Algorithm. Due to high time complexity for Brute Force Algorithm we can use it as a reference only in events with low multiplicities ($4 < N_{tr} < 15$).
On figure 4.6.1 distributions of $t_0 - t_0^{true}$ are presented. Those distributions are unbiased, and

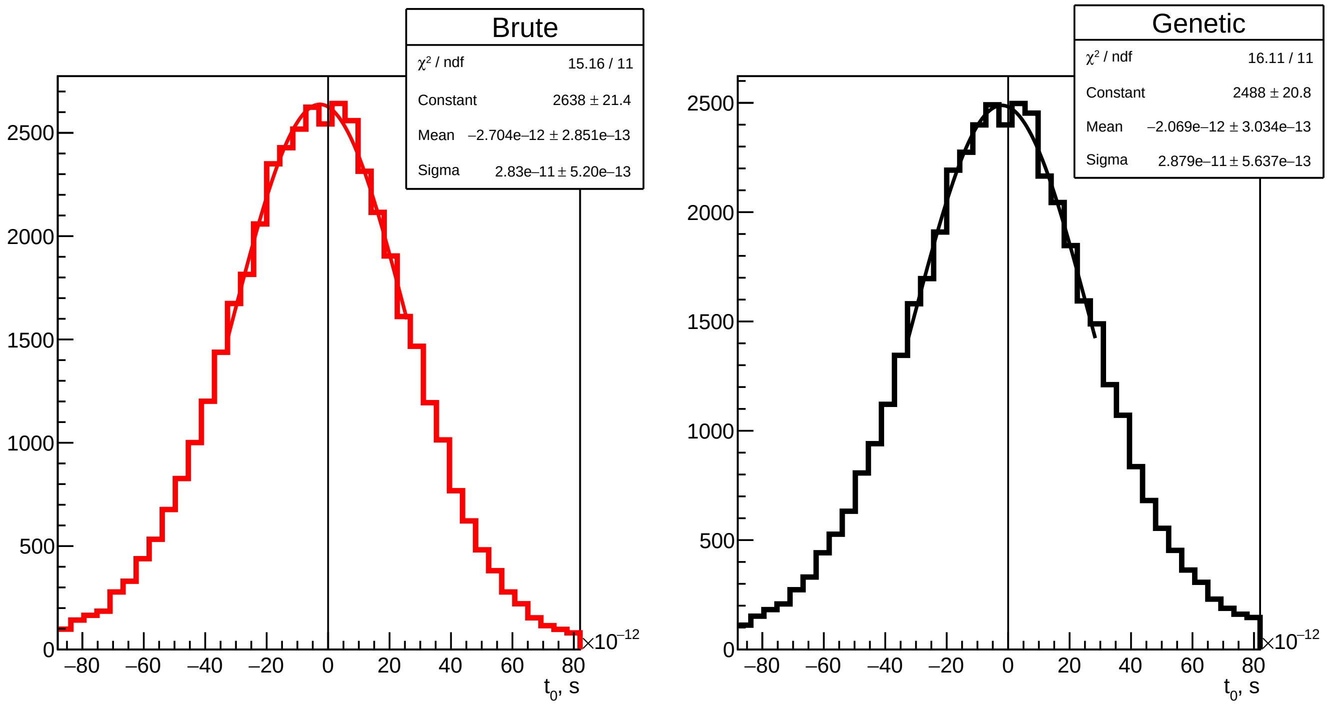

Figure 4.6.1: $t_0 - t_0^{true}$ distributions

have resolutions of 28 ps for Brute Force Algorithm and 29 ps for Genetic Algorithm. Also two additional metrics were calculated. PID event efficiency - percentage of event where all tracks were reconstructed correctly and PID track efficiency - percentage of tracks that were reconstructed correctly. PID event efficiency for Brute Force Algorithm is 66.7% and for Genetic Algorithm it is 63.3%. PID track efficiency is 97.2% for Brute Force Algorithm and for Genetinc Algorithm it is 96.8%.

Time complexity comparison is presented on figure 4.6.3. In events with multiplicity smaller than 8 tracks Brute Force Algorithm has shorter run time. Due to exponential time complexity it becomes unusable when multiplicity is higher. Average run time of Brute Force Algorithm on events with $4 < N_{tr} < 15$ is 5 ms and for Genetic Algorithm this time is 160 μs. On figure 4.6.4 comparison of two possible particle types hypotheses are shown. First plot obtained using mass set of π, K, p and second using mass set of e, π, K, p. When electron is added distribution becomes biased. This happens due to smearing of TOF detector signal (fig. 4.6.2). If electron is added, new minimum for χ^2 is obtained, where pions that look faster due to smearing switched to electron hypothesis.

On figure 4.6.5 distribution of t_0 errors is presented. This distribution have peaks in points where $\sigma_{t_0} = \sigma_{TOF}/\sqrt{N_{tr}}$. And between those peaks there are areas with event where some particles have big $\sigma_{t(p)}$. Errors of t_0 are less than the resolution of TOF detector, which means that our procedure do not add significant uncertainties.

From results presented above some conclusions can be made. For events with high multiplicity Genetic Algorithm has significantly better run time. Resolution, PID event and PID track efficiencies have negligible difference.

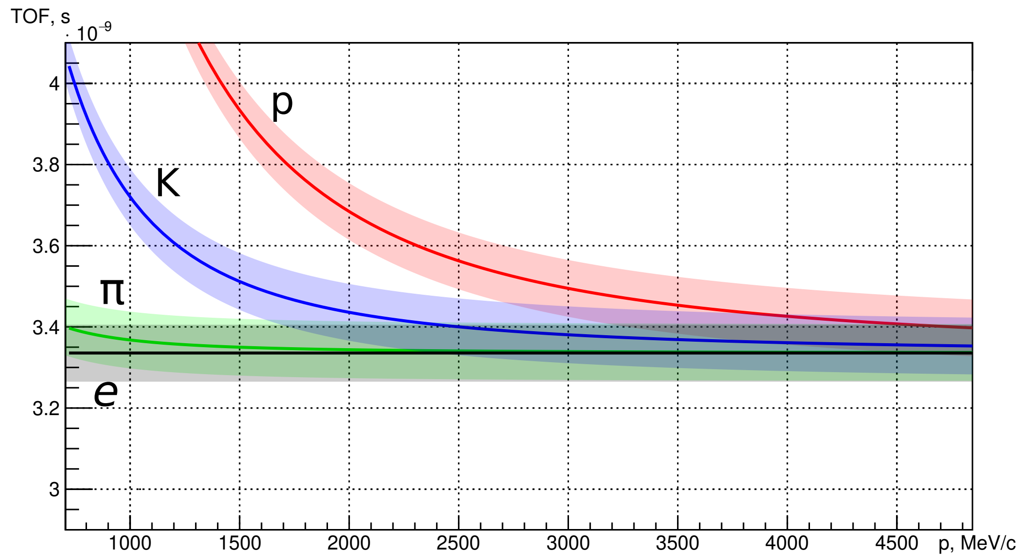

Figure 4.6.2: TOF dependance on momentum for different particles. Band around lines is $\pm 70\ ps$. Only resolution of TOF-detector is taken into account

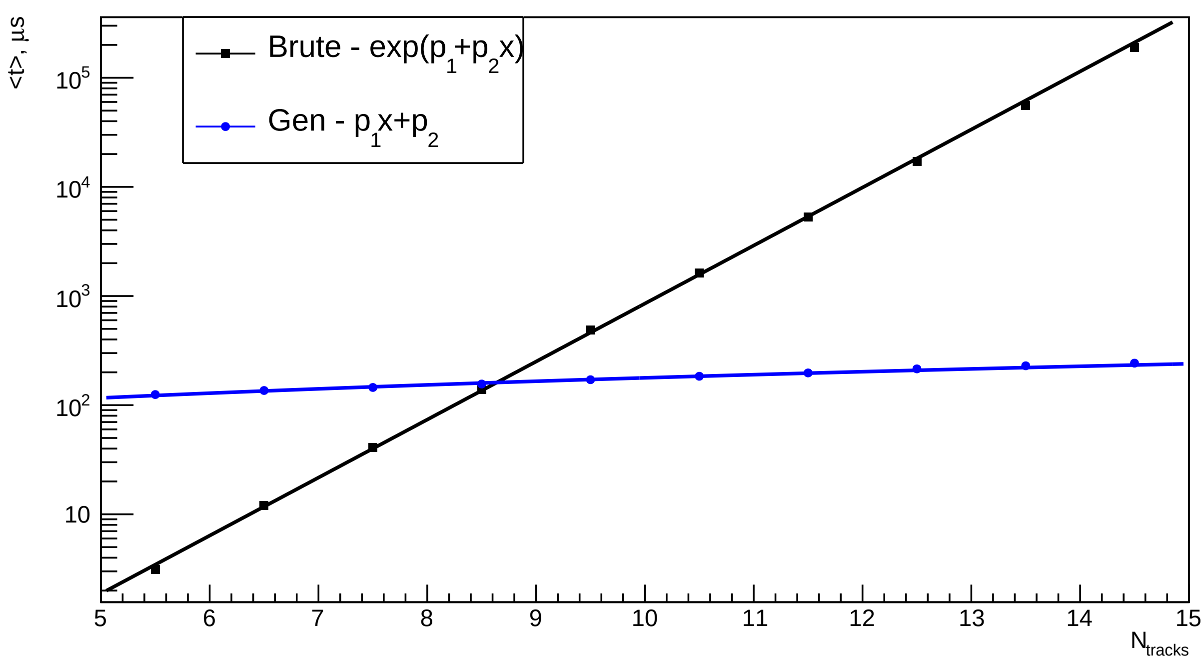

Figure 4.6.3: Time complexity comparison

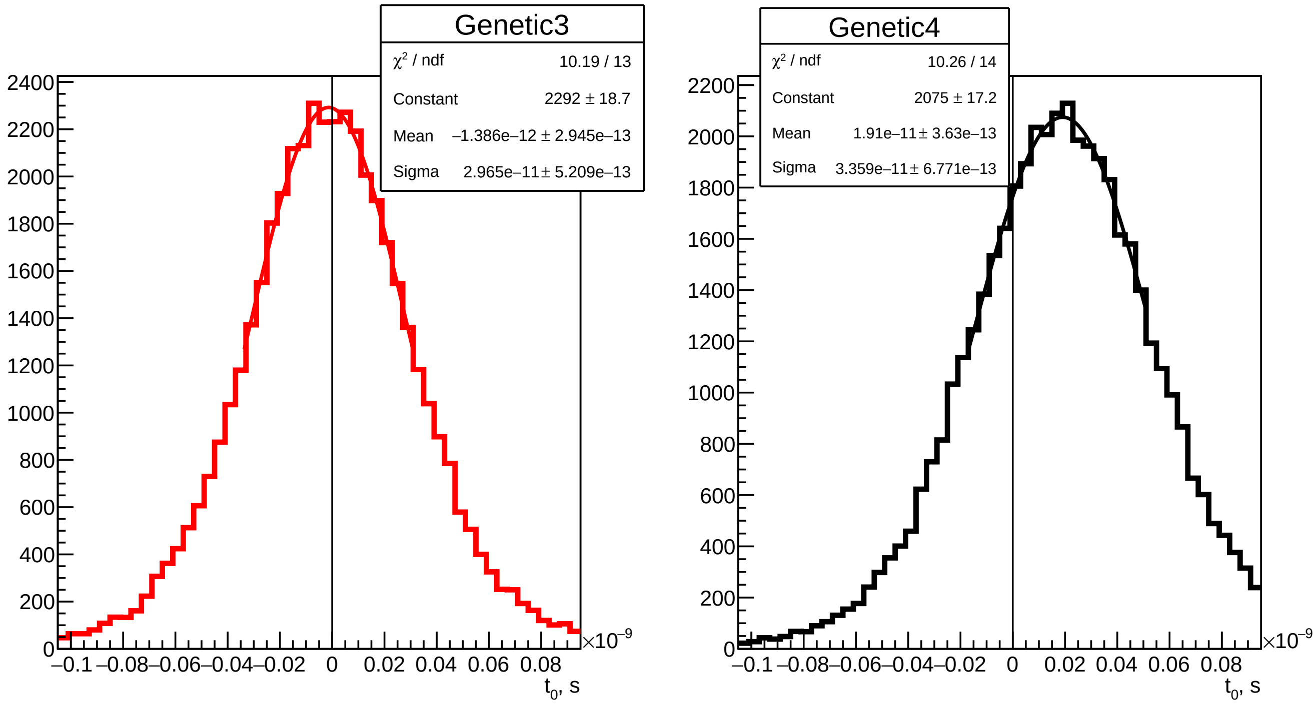

Figure 4.6.4: Distribution of $t_0 - t_0^{true}$ for Genetic Algorithm without (left) and with (right) electron

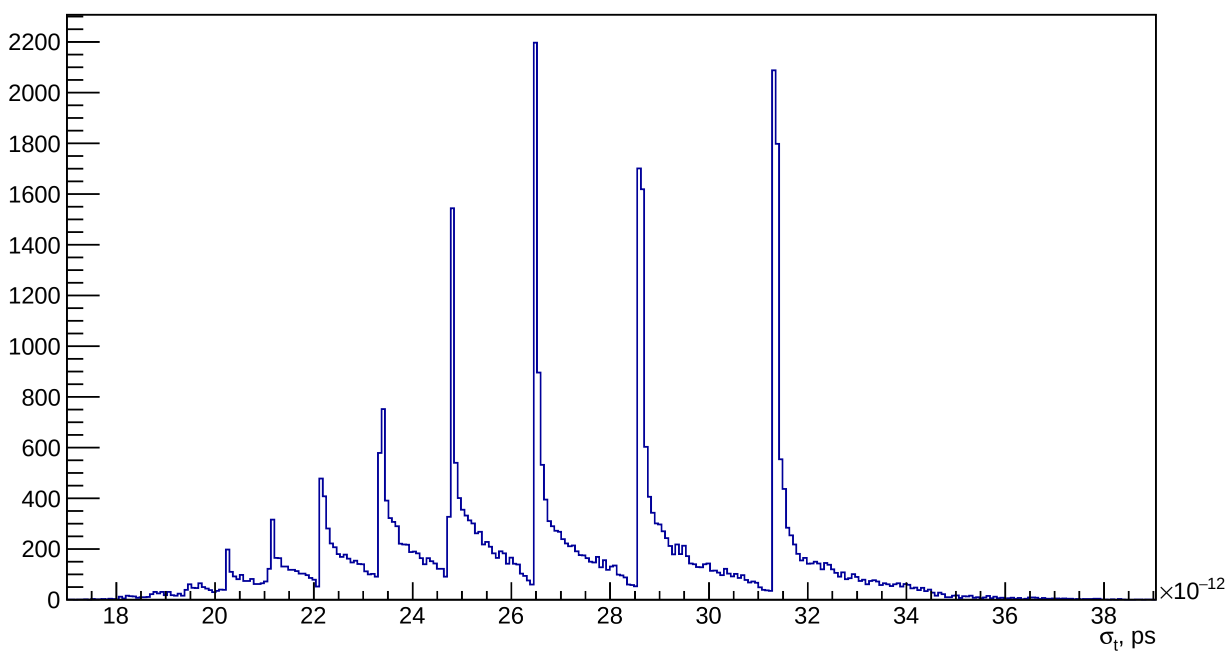

Figure 4.6.5: Distribution of t_0 determination errors

5 Particle identification

5.1 General idea

Now a reliable method for t_0 determination is developed and particle identification should be performed. There are several ways and strategies for PID:

1. Take particle types from χ^2 minimum determined by Genetic Algorithm
2. For every particle in event we should exclude it from determination of t_0 and after that reconstruct it type with [3]:
   (a) n-sigma criteria: discriminating variable $n_{\sigma_k^i}$ is used:

$$n_{\sigma_k^i} = \frac{S_i - \hat{S}_i(m_k)}{\sigma_k^i} \tag{5.1.1}$$

   Where S_i is a signal obtained for particle and $\hat{S}_i(m_k)$ is expected average signal for particle of species k. After that if particle lies in range of 2 or 3 σ of certain species this particle is accepted as particle of this species. Particle can be accepted as multiple species.

   (b) Bayesian method: calculating probability for particle to be particle of species j by formula:

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=\pi,K,p} P(\vec{S}|H_k)C(H_k)} \tag{5.1.2}$$

   Where C_i is prior probabilities that can be calculated iteratively and $P_{S|H_i}$ is given by:

$$P_{S|H_i} = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}n_{\sigma_i}^2\right) \tag{5.1.3}$$

5.2 Comparison of different methods

Comparison of different methods was done by reconstructing ϕ - meson. For ϕ - meson K^+K^- decay channel was chosen. From figure 5.2.6 one can be conclude that taking particle types from χ^2 and weighted bayesian approaches have better precision than others.

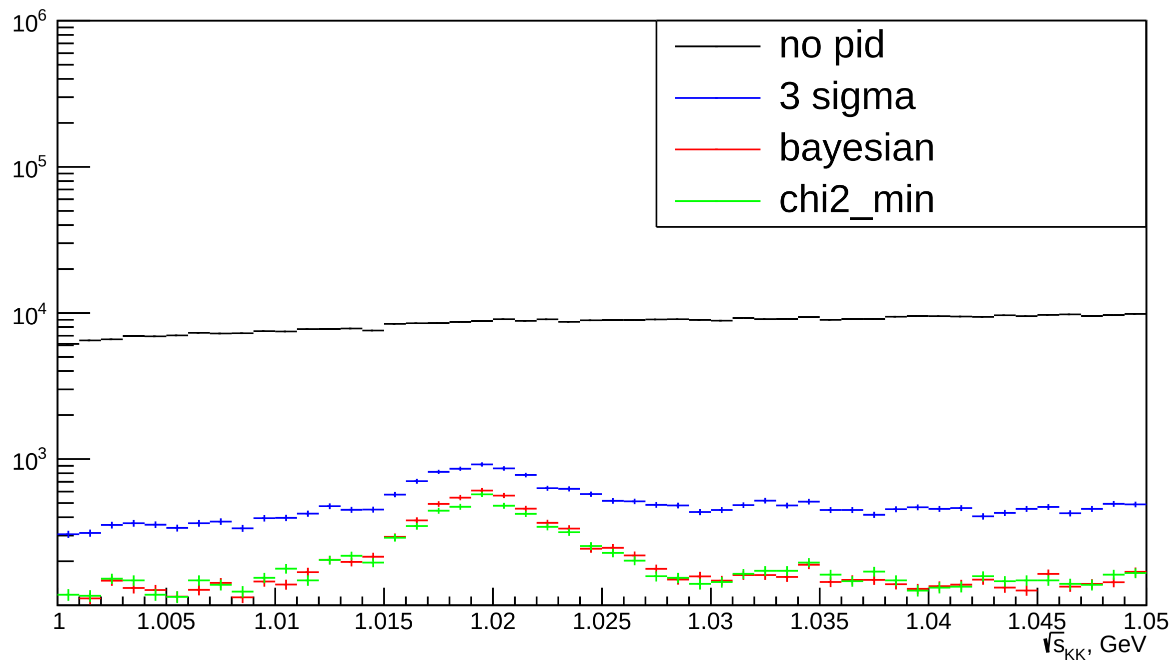

Figure 5.2.6: Invariant mass for kaon pairs

6 Conclusions

Algorithm for t_0 determination was developed. Obtained t_0 distribution is unbiased and have resolution of 29 ps. Particle identification procedure have been performed with different methods. Results of this work have been presented twice on SPD-collaboration meetings and will be presented on the SPD-Physics meeting.

In future Genetic Algorithm should be optimised to decrease run time. Particle identification procedure needs further studies.

7 Acknowledgements

I would like to express my deepest appreciation to Zhabitsky Mikhail for constant support and guidance. Special thanks to SPD collaboration for valuable feedback. Lastly, I'd like to mention START organisers for providing opportunity to take part in this program.

References

[1] Conceptual design of the Spin Physics Detector. arXiv:2102.00442

[2] E. Zhabitskaya and M. Zhabitsky, Proceedings of the 15th annual conference on Genetic and evolutionary computation, GECCO'13, ACM, 455–462 (2013)

[3] The ALICE Collaboration., Adam, J., Adamová, D. et al. Particle identification in ALICE: a Bayesian approach. Eur. Phys. J. Plus 131, 168 (2016). https://doi.org/10.1140/epjp/i2016-16168-5