Genetic Algorithm for determination of event collision time and particle identification by Time-Of-Flight at SPD NICA

Semyon Yurchenko

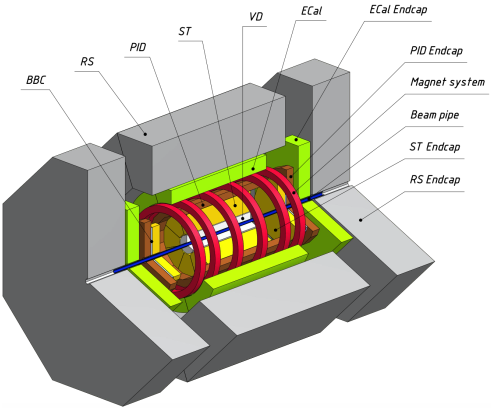
Saint Petersburg State University

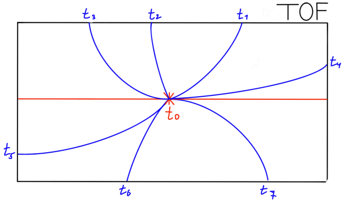
6-th International Conference on Particle Physics and Astrophysics

December 1, 2022

The Spin Physics Detector

- Beams: $p\uparrow, d\uparrow$,
- $\sqrt{s} = 27 GeV$,
- Physics:
- Spin dependent physics
- Structure of proton
- Charmonia, Open charm, prompt-photons production
- Double spin asymmetries





Task and initial conditions

Using information about particles trajectories and hits from TOF detector determine time of *pp*-collision.

1. In this work only tracks with the momentum above 500 MeV will be considered.
2. Resolution of the TOF detector $\sigma_t = 70\ ps$.
3. Momentum resolution: $\frac{\sigma_p}{p} = 2\%$
4. TOF radius is 1 m and length of 3.772 m.

Plan of simulation

1. Get tracks of charged particles with momentum over 500 MeV from Pythia8 events with $\sqrt{s} = 27\ GeV$.
2. Calculate intersection point with TOF detector(t_i).
3. Calculate arc length of trajectory(L_i).
4. Smear t_i with $N(t_i, 70\ ps)$ and p_i with $N(p_i, 0.02 \cdot p_i)$.
5. Using information about arc lengths of trajectories, TOF hits and particle momentum determine time of *pp*-collision(t_0).

Plan of simulation

1. Get tracks of charged particles with momentum over 500 MeV from Pythia8 events with $\sqrt{s} = 27\ GeV$.
2. Calculate intersection point with TOF detector(t_i).
3. Calculate arc length of trajectory(L_i).
4. Smear t_i with $N(t_i, 70\ ps)$ and p_i with $N(p_i, 0.02 \cdot p_i)$.
5. **Using information about arc lengths of trajectories, TOF hits and particle momentum determine time of pp-collision(t_0).**

Brute force

Brute Force - checking all possible solutions.

1. Choose particles types to make tof hypotheses $\rightarrow [\pi^{\pm}, K^{\pm}, p^{\pm}]$.

$$tof_{ik} = \frac{L_i}{c}\sqrt{1 + \frac{m_k^2}{p_i^2}} \tag{1}$$

2. For every event check all tracks hypotheses combinations - 3^N variants.
3. On every step calculate t_0 and χ^2 and find χ^2_{min}.

$$\chi^2 = \sum^N \frac{(t_0 + tof_{ik} - t_i)^2}{\sigma_t^2 + \sigma_{p_i}^2}, \quad t_0 = \frac{1}{\omega}\sum^N \frac{t_i - tof_{ik}}{\sigma_t^2 + \sigma_{p_i}^2}, \quad \omega = \sum^N \frac{1}{\sigma_t^2 + \sigma_{p_i}^2} \tag{2}$$

4. Time complexity $O(N \cdot 3^N)$.

Brute force

Brute Force - checking all possible solutions.

1. Choose particles types to make tof hypotheses$\rightarrow [\pi^{\pm}, K^{\pm}, p^{\pm}]$.

$$tof_{ik} = \frac{L_i}{c}\sqrt{1 + \frac{m_k^2}{p_i^2}} \tag{1}$$

2. For every event check all tracks hypotheses combinations - 3^N variants.
3. On every step calculate t_0 and χ^2 and find χ^2_{min}.

$$\chi^2 = \sum^N \frac{(t_0 + tof_{ik} - t_i)^2}{\sigma_t^2 + \sigma_{p_i}^2}, \quad t_0 = \frac{1}{\omega}\sum^N \frac{t_i - tof_{ik}}{\sigma_t^2 + \sigma_{p_i}^2}, \quad \omega = \sum^N \frac{1}{\sigma_t^2 + \sigma_{p_i}^2} \tag{2}$$

4. Time complexity $O(N \cdot 3^N)$ - **very slow!!!**

Genetic algorithm

How genetic algorithm works:

1. Associate masses of particles with indices: $[m_\pi, m_K, m_p] \to [0, 1, 2]$
2. Create population of random candidate solutions - $v([m_i]_k)$.
3. Create mutant vector from random candidates in population(DE-inspired):

$$v_{mut} = v_r + (v_p - v_q) \quad (3)$$

4. Check if $\chi^2_{mut} < \chi^2_r$ then replace v_r with v_{mut}. If not, population remains unchanged - **Darwinian selection**.
5. Repeat
6. After some number of steps stop and choose χ^2_{min} as an answer.
7. Time complexity - $O(N \cdot N_{population} \cdot N_{steps})$, $800 < N_{steps} < 1000$.

Genetic algorithm: Example of iteration

Three candidate solutions and their χ^2:

1. $v_r = [2, 1, 1, 2, 0, 0]$, $\chi^2 = 300$
2. $v_p = [1, 0, 0, 1, 0, 1]$, $\chi^2 = 200$
3. $v_q = [0, 1, 2, 2, 1, 2]$, $\chi^2 = 100$

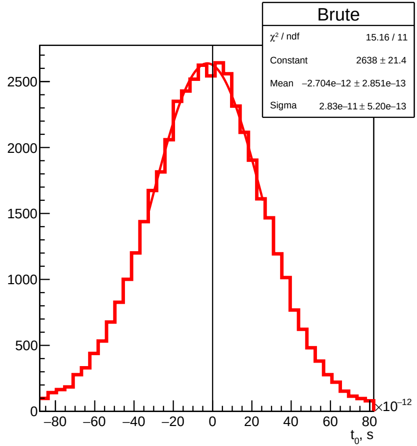
$v_{mut} = v_r + (v_p - v_q) = [2, 0, 0, 1, 0, 0]$, $\chi^2 = 10$
$\chi^2_{mut} < \chi^2_r$, so we change v_r in population for v_{mut}.

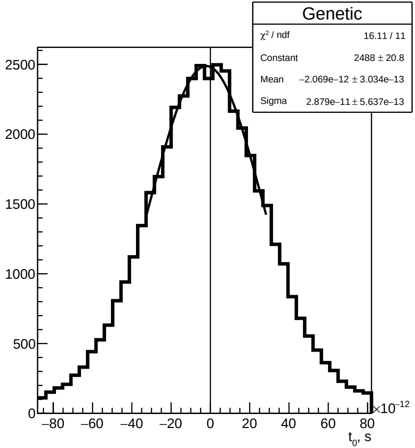
Genetic algorithm vs Brute force

Comparison of 2 algorithms was done on events with number of tracks $5 \leq N \leq 14$.

Brute Force's efficiency$= \frac{Events(-3\sigma,3\sigma)}{Event_{Normal}(-3\sigma,3\sigma)} \sim 98.6\%$.

Genetic Algorithm's efficiency$\sim 97\%$





Genetic algorithm vs Brute Force

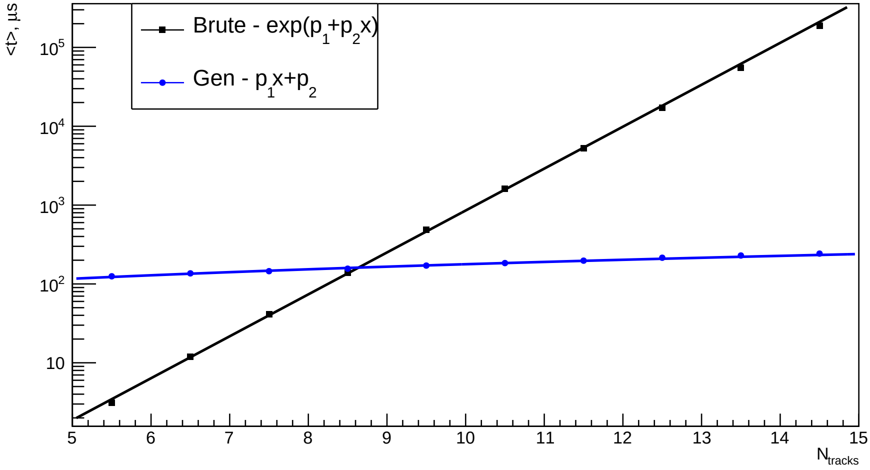
Differences:

1. Brute force gives solution with minimal χ^2, but very slowly.
2. Genetic algorithm is nearly as accurate, but much faster.

Similarities

1. Resolution of $t_0 \sim 29$ *ps*,
2. Percentage of events where all tracks have been identified correctly $\sim 63.3\%$,
3. Overall percentage of tracks that were identified correctly $\sim 96.8\%$.

Time complexity



Particle Identification

Some possible strategies:

1. Take particles types from χ^2 minimum(chi2_min),
2. N-sigma criteria,
3. Bayesian approach.

In 2 and 3 strategies we exclude particle from determination of t_0 to avoid correlations.

$$\sigma^2 = \sigma^2_{TOF} + \sigma^2_p + \sigma^2_{t_0}, \quad (4)$$

N-sigma criteria	Bayesian approach
$\mathrm{n}_{\sigma^i_k} = \frac{S_i - \hat{S}_i(m_k)}{\sigma^i_k}$	$\mathrm{P}(\mathrm{H}_i \vert \vec{S}) = \frac{P(\vec{S} \vert H_i) C(H_i)}{\sum_{k=\pi,K,p} P(\vec{S} \vert H_k) C(H_k)}$
If $n_{sigma} < 3$ for certain type,	$\mathrm{P}(\mathrm{S} \vert H_i) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2} n^2_{\sigma_i}\right)$
we associate this type with particle	Probability of being particle
$\rightarrow$more than 1 type	of certain type
Only smearing is considered	Abundance of particles
	is also taken into account

Bayesian approach

Bayes formula:

$$P(H_i|\vec{S}) = \frac{P(\vec{S}|H_i)C(H_i)}{\sum_{k=\pi,K,p} P(\vec{S}|H_k)C(H_k)} \tag{5}$$

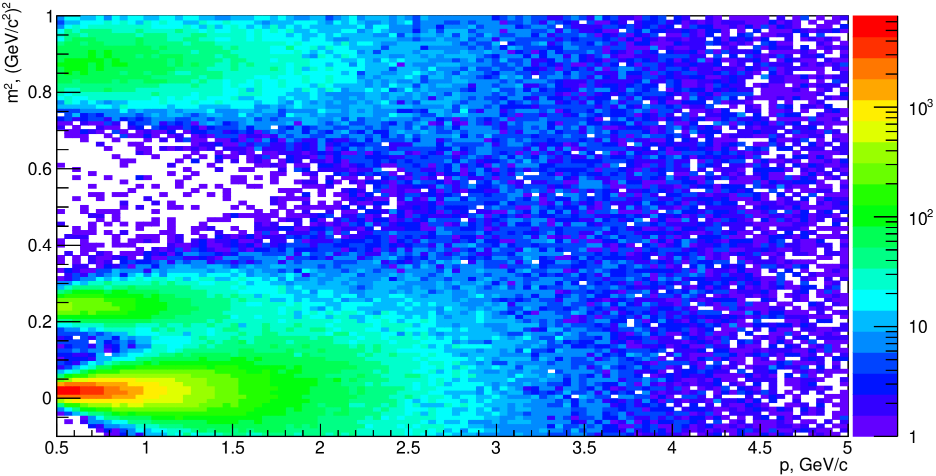
$C(H_i)$-prior probability, $P(H_i|\vec{S})$ - posterior probability, which is calculated iterativly. On first step $C(H_i)$ are chosen flat.
Detector signal:

$$P(S|H_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}n_{\sigma_i}^2\right) \tag{6}$$

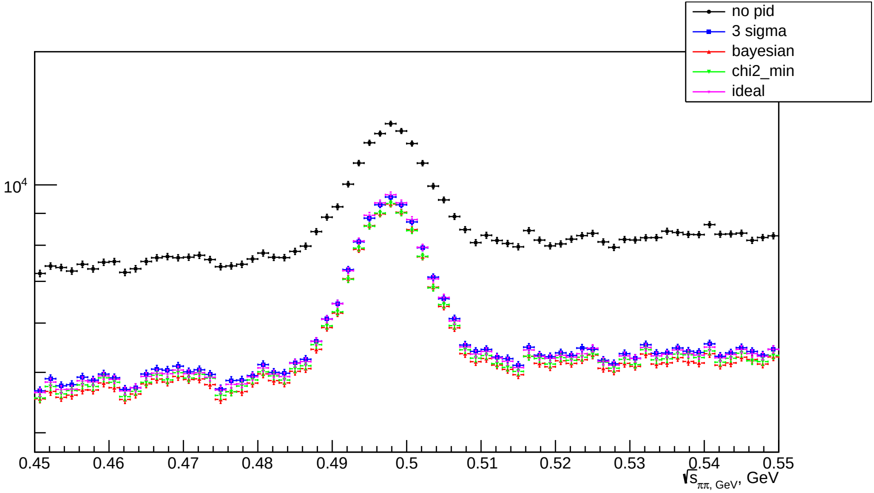
In case of multiple signal(i.e $dE/dx + TOF + Cherenkov$):

$$P(S|H_i) = \prod_{k=dE/dx,TOF,Cherenkov} P(S|H_i)_k \tag{7}$$

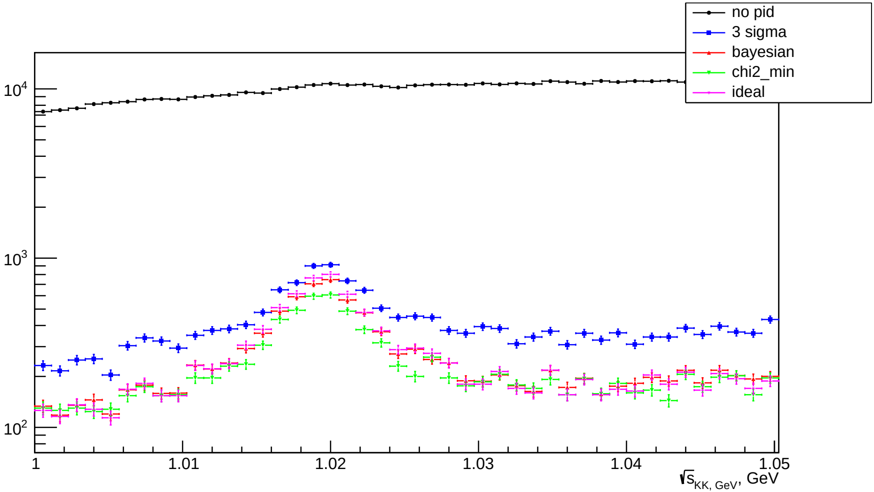
PID masses calculated with exclusion



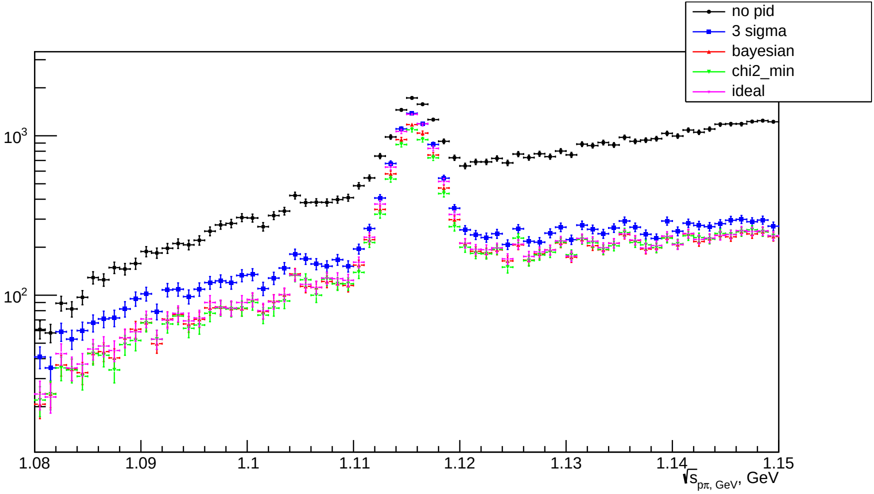
Signals: $K_s \to \pi^+\pi^-$



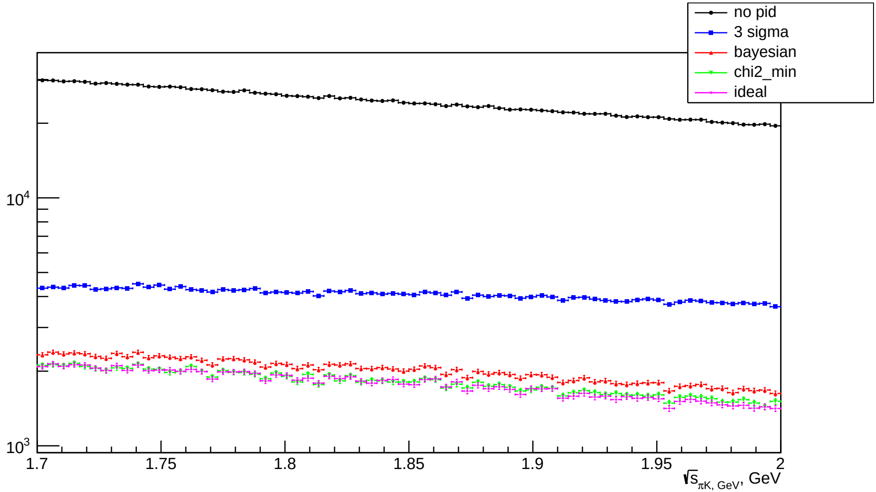
Signals: $\phi \to K^+K^-$



Signals: $\Lambda \to p^+\pi^-$



Signals: $D^0 \to \pi^+ K^-$ $(and\ cc)$



Summary

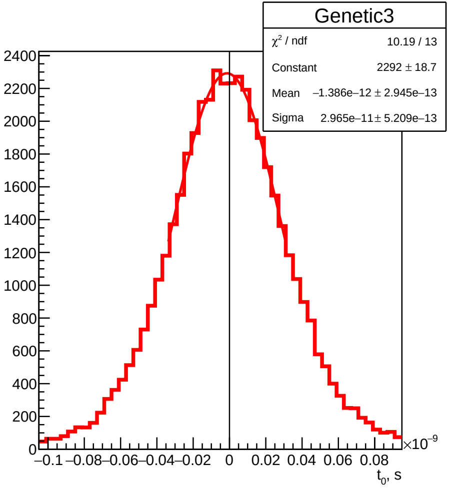
1. Procedure for determination of t_0 has been developed.
2. Genetic Algorithm efficiency $\sim 97\%$,
3. For Genetic Algorithm run time grows slowly as function of multiplicity. In events with high multiplicity it is algorithm of choice.
4. Both methods have resolution ~ 29 *ps*, which is much better than the TOF resolution(70 *ps*).
5. Obtained value of t_0 faciliates the particle identification at SPD.

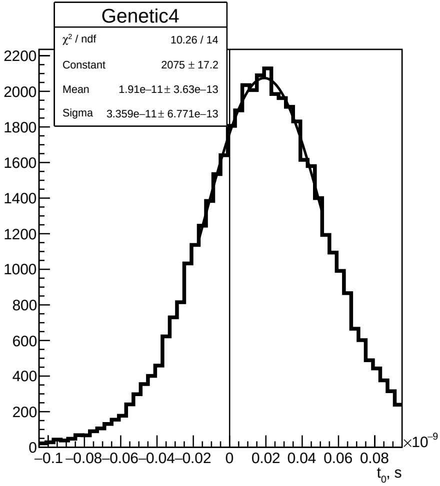
Thank you for your attention!

Backup slides

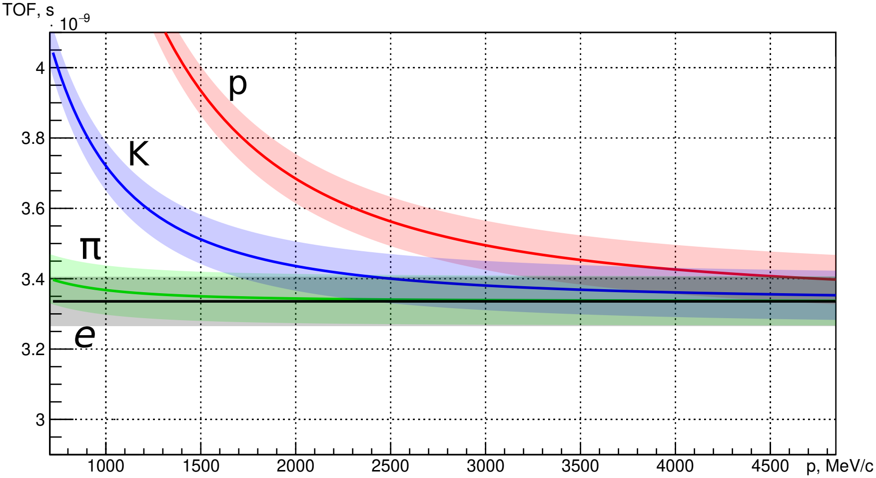
Different hypotheses

Genetic3 hypotheses is $[\pi^{\pm}, K^{\pm}, p^{\pm}]$
Genetic4 hypotheses is $[e^{\pm}, \pi^{\pm}, K^{\pm}, p^{\pm}]$





TOFs



TOF, s
· 10^-9
4
3.8
3.6
3.4
3.2
3
1000
1500
2000
2500
3000
3500
4000
4500
p, MeV/c
p
K
π
e