physics

MDPI

*Article*

# Genetic Algorithm for Determination of the Event Collision Time and Particle Identification by Time-of-Flight at NICA SPD

**Semyon Yurchenko [1] and Mikhail Zhabitsky [2,*]**

[1] Laboratory of Ultra-High Energy Physics, Saint Petersburg State University, St. Petersburg 199034, Russia; semyon.yurchenko@luhep.ru
[2] Dzhelepov Laboratory of Nuclear Problems, Joint Institute for Nuclear Research, Dubna 141980, Russia
* Correspondence: mikhail.zhabitskiy@cern.ch

**Abstract:** Particle identification is an important feature of the future SPD (Spin Physics Detector) experiment at the NICA (Nuclotron-based Ion Collider fAcility) collider. In particular, the identification of particles with momenta up to a few GeV/$c$ (with $c$ the speed of light) by their time-of-flight facilitates the reconstruction of events of interest. The high time resolution of modern TOF (Time-Of-Flight) detectors demands the need to obtain the event collision time, $t_0$, with comparable accuracy. While the determination of the collision time is feasible through the use of TOF signals supplemented by track reconstruction, it proves to be computationally expensive. In the presented study, a dedicated Genetic Algorithm is developed as a fast and accurate method to determine the proton–proton collision time by the measurements of the TOF detector at the SPD experiment. By using this reliable method for the $t_0$ determination we compare different approaches for the particle identification procedure based on TOF signals.

**Keywords:** genetic algorithm; time-of-flight; particle identification; spin physics; NICA project; SPD; TOF detector



**Citation:** Yurchenko, S.; Zhabitsky, M. Genetic Algorithm for Determination of the Event Collision Time and Particle Identification by Time-of-Flight at NICA SPD. *Physics* **2023**, *5*, 423–435. https://doi.org/10.3390/physics5020030

Received: 27 February 2023
Revised: 22 March 2023
Accepted: 29 March 2023
Published: 17 April 2023



## 1. Introduction

The Spin Physics Detector (SPD) is a future experiment that is planned to be placed in one of the two interaction points of the NICA (Nuclotron-based Ion Collider fAcility) collider in the Joint Institute for Nuclear Research (JINR, Dubna, Russia). By studying collisions of polarized proton and deuteron beams, the SPD Collaboration is going to perform a comprehensive study of the unpolarized and polarized gluon content of nucleons and other spin-related QCD (Quantum Chromodynamics) phenomena [1]. With polarized proton–proton (pp) collision energies, $\sqrt{s}$, up to 27 GeV, and a luminosity up to $10^{32}$ cm$^{-2}$s$^{-1}$, SPD is designed to cover a kinematic range between the low-energy measurements at ANKE-COSY [2] and SATURNE [3] experiments and the high-energy measurements at RHIC (Relativistic Heavy Ion Collider, Brookhaven, USA) [4] and the LHC (Large Hadron Collider, Geneva, Switzerland) [5].

The SPD experimental setup is planned as a general-purpose $4\pi$ detector with advanced tracking and particle identification capabilities. The particle identification will be performed by means of the measurement of their ionization losses, Time-Of-Flight (TOF) detector, electromagnetic calorimetry, and muon-filtering techniques. The experiment is considered to use a system of Multigap Resistive Plate Chambers (MRPC) [6,7] as the TOF detector. A main aim of the TOF detector is to provide $\pi/K/$p-identification of charged particles with momenta up to a few GeV/$c$, where $c$ denotes the speed of light.

Identification of particle types by their time-of-flight is an established technique in high energy physics collider experiments [8–12]. It requires just three ingredients: $p$—momentum of the particle, $L$—arc length of its trajectory between the primary collision point and the TOF detector, and $\tau$—the corresponding time-of-flight. The latter is calculated as a time-difference between the stop and start signals. While the stop signal is measured with

high precision by the TOF detector, the collision time, $t_0$, cannot be obtained directly with the same accuracy. The collision time can be estimated from the timing of the accelerator or deduced from signals of a corresponding T0-detector, which is supposed to detect the appearance of secondary particles scattered on small angles with respect to the beams axis, but, in this case, the uncertainty of the collision time dominates the uncertainty of the difference between stop and start signals. Fortunately, $t_0$ can be determined with sufficient precision from the TOF measurements by means of $\chi^2$-minimization procedure [8–10].

As a significant number of secondary particles, originating from the proton–proton collision, subsequently enter the acceptance and are detected by the TOF detector, one can reconstruct $t_0$ as a common value for all detected particles through minimization of the sum of the squares of residuals, $\chi^2$. A residual is defined as a difference between the measured TOF signal and its expected arrival time, assuming a given mass hypothesis. Thus, the $\chi^2$ is minimized with respect not only to $t_0$ but to all mass hypotheses, which proved to be quite a complicated computational task [9].

As the bulk of secondary particles are pions, kaons or protons, it is natural to try different combinations of their masses in order to minimize $\chi^2$, thus the minimization is performed over a discrete set of particle types. The global minimum can be found using the Brute Force Algorithm (BFA), which is characterized by a relatively long run-time. In this study, an Asynchronous Differential Evolution-inspired [13] Genetic Algorithm (ADE-GA) is developed which solves the $\chi^2$-minimization problem within significantly reduced computational time.

This reliable method for $t_0$ determination facilitates the identification of particles by their time of flight. Several approaches can be used for the PID (particle identification) procedure [14]. In this paper, we compare the performance of the Bayesian approach, the so-called "$n$-sigma" criteria, and the direct solution of the $\chi^2$-minimization problem.

## 2. Time-of-Flight Detector and Event Selection

The TOF system is designed to consist of a barrel and two end-cap parts (see Figure 1) with a radius of about 105 cm and a length of 370 cm. It will have an overall active area of 27 m$^2$ and cover polar angles greater than 100 mrad. The short distance from the TOF detector to the interaction point dictates the TOF resolution to be within 50–60 ps, which can be achieved with the MRPC technologies [7]; in the current study, we use a conservative estimation, $\sigma_t = 70$ ps. The TOF detector is considered to be located within the solenoidal magnetic field ($B = 1$ T, parallel to the beam axis) outside the inner tracker, which is aimed to measure particles momenta with relative precision, $\sigma_p/p = 2\%$ [1].

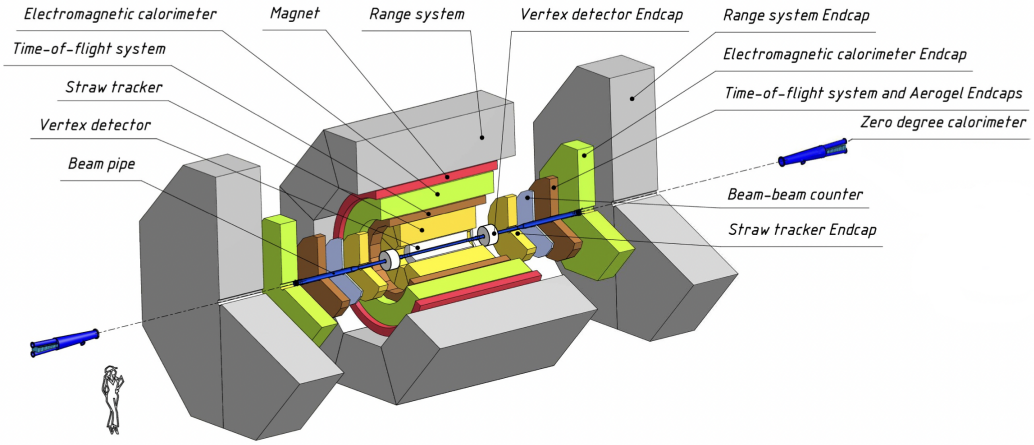


**Figure 1.** General layout of the Spin Physics Detector (SPD) [1].

PYTHIA 8 Monte Carlo generator [15] has been used for the simulation of pp collisions at $\sqrt{s} = 27$ GeV, "SoftQCD:all" settings have been selected to simulate minimum-bias events. All charged particles have been propagated through the uniform magnetic field.

Intersection points of helix trajectories with the TOF detector have been calculated through the closed-form expressions. Only charged tracks with momenta greater than 0.5 GeV/$c$ have been used in the analysis, as relativistic particles are characterized by negligible track distortions due to multiple scattering and ionization losses in the walls of the beam pipe and material of the inner tracker. For the open-charm production, which is of interest to the SPD, more than 10 charged tracks are supposed to be detected by the setup in 90% of events. In the following, for each track $i$ the generated momentum value, $p_{i,0}$, and the arrival time, $t_{i,0}$, to the TOF detector are smeared according to Normal distributions, $p_i = N(p_{i,0}, \sigma_p)$ and $t_i = N(t_{i,0}, \sigma_t)$, respectively.

## 3. Event Collision Time Measurement Performed by the TOF Detector

### *3.1. Method to Reconstruct the Event Collision Time*

A large fraction of particles produced in pp collisions at $\sqrt{s} = 27$ GeV has momenta below 2 GeV/$c$, which suggests a time-of-flight as a powerful technique for particles identification. Statistically, more than half of secondary charged particles detected by the TOF detector are pions. Other major contributions are protons and charged kaons. Admixtures of electrons and muons do not exceed a few percent and to be identified by the Electromagnetic Calorimeter and the Muon Range System [1], see Figure 1. One can calculate all possible times-of-flight,

$$\tau_{ij} = \frac{L_{ij}}{c}\sqrt{1 + \frac{m_j^2}{p_i^2}}, \tag{1}$$

by assigning independently for each track $i$ a certain particle type $j$. Equation (1) is correct only for relativistic particles; for low momenta, Equation (1) has to be replaced by a piecewise summation along the particle's trajectory intersecting coordinate detectors. The arc length, $L_{ij}$, should take into account details of a type $j$ particle interaction with matter. In this analysis, we select only the tracks with momentum above 0.5 GeV/$c$ so that the matter effects can be neglected. For the event with $N$ reconstructed tracks, the event collision time can be found by a $\chi^2$-minimization procedure:

$$\chi^2 = \sum_i^N \frac{(t_0 + \tau_{ij} - t_i)^2}{\sigma_t^2 + \sigma_{\tau(j,p_i)}^2} = \sum_i^N \frac{(t_0 + \tau_{ij} - t_i)^2}{\sigma_{ij}^2}, \tag{2}$$

where the time-of-flight uncertainty, $\sigma_{\tau(j,p_i)}$, due to the uncertainty in momentum is given by

$$\sigma_{\tau(j,p_i)} = \frac{L_{ij}}{c} \cdot \frac{m_j^2}{p_i^2}\left(\sqrt{1 + \frac{m_j^2}{p_i^2}}\right)^{-1} \cdot \frac{\sigma_p}{p} = 0.02 \cdot \frac{L_{ij}}{c} \cdot \frac{m_j^2}{p_i^2}\left(\sqrt{1 + \frac{m_j^2}{p_i^2}}\right)^{-1}. \tag{3}$$

For pions with momenta 0.5 GeV/$c$, $\sigma_{\tau(j,p_i)}$ is as large as 100 ps and is taken into account. Another contribution to $\sigma_{\tau(j,p_i)}$ is the uncertainty in the reconstructed track length. In the SPD experiment, tracking detectors are supposed to provide us with up to 30 to 40 hits per track in spatially separated detector planes [1], thus, for 0.5 GeV/$c$ pions, the time-of-flight uncertainty, due to the uncertainty in the reconstructed track length, is less than 10 ps and is neglected in this study.

For a certain choice of mass hypotheses, an analytic solution for $t_0$ reads:

$$t_0 = \sigma_0^2 \sum_i^N \frac{t_i - \tau_{ij}}{\sigma_{ij}^2}, \quad \text{where} \quad \frac{1}{\sigma_0^2} = \sum_i^N \frac{1}{\sigma_{ij}^2}. \tag{4}$$

So, the task is reduced to a minimisation of $\chi^2$ (2)–(4) by finding the proper mass hypothesis—vector of masses $(m_1, m_2, \ldots, m_N)$ for tracks in the event-by-event way. The emphasis is paid to deducing an accurate and unbiased estimation of the collision time $t_0$.

Sections 3.2 and 3.3 are dedicated to developing algorithms to perform the minimization step.

### *3.2. Brute Force Algorithm*

The most straightforward solution is to check all mass hypotheses and thus locate the global minimum—a combination of masses which has minimal $\chi^2$, so-called exhaustive search or BFA. If $N_m$ is the number of possible masses (possible particle types), then the total number of combinations is $N_m^N$ and the time complexity of this algorithm is $\mathcal{O}(N \cdot N_m^N)$. Exponential running time means that this algorithm is computationally expensive if the number of reconstructed tracks $N$ exceeds 10. To keep BFA execution time reasonable, possible particle types are restricted to $\pi^{\pm}$, $K^{\pm}$ and (anti)protons ($N_m = 3$).

### *3.3. Genetic Algorithm*

The minimization of $\chi^2$ (2)–(4) is performed over a discrete set of particle species, thus represents a typical problem in the domain of the discrete optimization. To solve the problem, an Asynchronous Differential Evolution-inspired [13] Genetic Algorithm (ADE-GA) is developed here.

All possible types of particles are represented as a mass-ordered set (genetic representation), e.g., $[m_\pi, m_K, m_\mathrm{p}] \to [0, 1, 2]$. The algorithm maintains a set of candidate solutions called a population. It optimizes a problem by iteratively improving the population through the generation of new candidate solutions, which can replace inferior population members by means of natural (Darwinian) selection. Opposite to the exhaustive search, the algorithm does not check all possible mass combinations but identifies better solutions and performs further searches around them.

The algorithm's workflow for an event with $N$ tracks is as follows:

1. $[m_\pi, m_K, m_\mathrm{p}] \to [0, 1, 2]$
2. Create an initial population of $N_\mathrm{pop}$ random candidate solutions. Each candidate solution is a (pseudo)random set of $N$ masses associated with corresponding tracks, each species has equal probability, $1/N_m$, to be assigned to a given track. The initialization procedure enforces that all population members are unique and for each track, there are at least two different masses within the population.
   At initialization, the population should be as much as possible to cover the search space. The requirement of at least two different masses per track within the population diverts the risk of a degenerated search—the search in a subspace of the search domain. Furthermore, at initialization, the expected abundance of particle species is not taken into account, but each species has an equal probability to be assigned to a given track—this approach facilitates the exploration ability of the algorithm. Example of a population in event with 6 tracks and size of population $N_\mathrm{pop} = 5$:

$$\begin{array}{l|l} v_1 & (0,1,1,2,0,0) \leftrightarrow (m_\pi, m_K, m_K, m_\mathrm{p}, m_\pi, m_\pi) \\ v_2 & (2,2,0,1,0,0) \leftrightarrow (m_\mathrm{p}, m_\mathrm{p}, m_\pi, m_K, m_\pi, m_\pi) \\ v_3 & (1,1,1,0,0,1) \leftrightarrow (m_K, m_K, m_K, m_\pi, m_\pi, m_K) \\ v_4 & (0,0,0,2,1,0) \leftrightarrow (m_\pi, m_\pi, m_\pi, m_\mathrm{p}, m_K, m_\pi) \\ v_5 & (2,0,1,1,2,2) \leftrightarrow (m_\mathrm{p}, m_\pi, m_K, m_K, m_\mathrm{p}, m_\mathrm{p}) \ . \end{array}$$

3. Create a new candidate solution (offspring generation):
   (a) Choose three distinct random solution vectors from the current population and create a mutant vector:

$$v_\mathrm{mut} = v_\mathrm{par} + \Delta v. \tag{5}$$

   Vector $v_\mathrm{par}$ is called a parent vector. Two other vectors form a difference vector, $\Delta v$. If any coordinate falls outside the range $[0, N_m - 1]$, it is projected back to the corresponding boundary. The mutant vector has to be different from any population member, otherwise the generation is repeated.

Example:

$$v_{\text{mut}} = (0,1,1,2,0,0) + (2,2,0,1,0,0) - (1,1,1,0,0,1) = (1,2,0,2,0,0). \quad (6)$$

(b) Calculate fitness of the offspring: $t_0^{\text{mut}}$ and $\chi^2_{\text{mut}}$ (see Equations (2)–(4)).
(c) Compare $\chi^2_{\text{par}}$ and $\chi^2_{\text{mut}}$.
(d) If $\chi^2_{\text{mut}} < \chi^2_{\text{par}}$—the new mutant vector is better than the parent, then the offspring supersedes the parent vector in the population. Otherwise, the population remains unchanged. This step is called natural (Darwinian) selection.

4. Steps 3(a)–3(d) are repeated until a terminating criterion is reached. After a predefined number of iterations, $N_{\text{steps}}$, the solution with the smallest $\chi^2$ is chosen as the best combination.

The size of the population, $N_{\text{pop}}$, is an important parameter in ADE. If one uses too small population size, the algorithm can fail to locate the global minimum. Search with a large $N_{\text{pop}}$ is characterized by better convergence, but consumes longer computer time [13]. Numerical simulations proved that $N_{\text{pop}} = 15$ is sufficient to solve the $\chi^2$-problem (2) even for the largest possible number of tracks. While low-multiplicity events can be resolved with somewhat smaller population sizes, usage of $N_{\text{pop}} < 10$ should be turned down due to a lower convergence to the global minimum. In this paper, a fixed population size $N_{\text{pop}} = 15$ is used for all events.

Time complexity of GA is $\mathcal{O}(N \cdot N_{\text{pop}} \cdot N_{\text{steps}})$, where $800 < N_{\text{steps}} < 1000$. Time complexity increases only linearly as a function of track number, which makes the algorithm suitable for high-multiplicity events. Furthermore, GA is not limited to $N_m = 3$, but can perform the global search for a wider range of possible particle types without loss of performance.

Besides the population size $N_{\text{pop}}$, the canonical DE has two other control parameters: the crossover rate, $C_r$. and the scale factor, $F$. In this study, $C_r$ is fixed to unity as minimization variables are correlated, $F = 1$ is chosen due to the granularity of the mass spectra. The developed ADE-GA algorithm represents the asynchronous evolutionary algorithm: it updates randomly selected population members by the DE/rand/1 strategy [16]. Here, the general convention is used to identify different mutation variants (5) by DE/$x$/$y$/$z$ notation, where $x$ stands for a choice of a parent vector (a randomly chosen vector in the current study), $y$ is the number of difference vectors, and $z$ denotes the crossover scheme (not used here). In this approach, a fitness of many candidate solutions can be calculated in parallel which further speeds up the calculations.

The algorithm can be further accelerated as soon as one monitors the convergence speed and carefully chooses termination (stop) criteria. The results of the current study are obtained using the algorithm performing a predefined fixed number of iterations, $N_{\text{steps}}$. In this approach, the maximal allowed $N_{\text{steps}}$ is chosen to guarantee a high convergence rate to the global minimum. Analysis of the convergence showed that for most events the minimum is found by a much earlier iteration and further iterations waste computing time. In what follows, several approaches toward early detection of global convergence are discussed.

The result of successive iterations of the ADE-GA algorithm is a gradual improvement of the population: naturally selected candidate solutions have smaller $\chi^2$-values than their respective parent vectors. Not only the best vector, but all population members converge to the minimum. Thus, the spread in fitness function values within the population is gradually reduced, the small spread can indicate either convergence or stagnation of the algorithm [17]. To monitor convergence one can sort all population members by their fitness values: $\chi^2_{\text{best}}, \ldots, \chi^2_{\text{m}}, \ldots, \chi^2_{\text{worst}}$, where $\chi^2_{\text{m}}$ denotes the median fitness. As the fitness of the global minimum is expected to be of the order of $N$, where $N$ is the number of tracks, one can end iterations as soon as $(\chi^2_{\text{m}} - \chi^2_{\text{best}})/N < \Delta_{\text{m}}$, where $\Delta_{\text{m}}$ is a predefined small value. Alternatively, the algorithm can monitor the typical number of iterations between successive improvements of the $(\chi^2_{\text{m}} - \chi^2_{\text{best}})$-difference—$N_{\text{progress}}$, which can be

achieved through learning in the process. If there is no progress after $kN_{\text{progress}}$ iterations, where typically $k = 3, \ldots, 5$, then the algorithm is terminated. Monitoring the $(\chi^2_{\text{m}} - \chi^2_{\text{best}})$-difference has major advantages with respect to the terminating criteria based only upon the $\chi^2_{\text{best}}$-value. The improvement steps by the algorithm can be characterized by exploration or exploitation chances of the steps. Exploration is the ability of the algorithm to locate a new region in the search domain with better fitness values. Exploitation is a gradual improvement of the population through testing of potentially interesting candidate vectors around an already found local minimum. DE is known for its exploration abilities. As soon as a new region of interest is located, the algorithm quickly populates the neighborhood of the local minimum, thanks to the mutation operator (5). Improvement steps by exploration are a much rarer case, while improvements through exploitation are common. If one monitors only the $\chi^2_{\text{best}}$-value as a stop criterion, then after a successful exploration step one can cause premature termination of iterations by preventing further fast exploitation around a new minimum. Typically, at this stage exploitation leads not to only a general improvement of the population but also to a better best-so-far solution.

## 4. Results and Discussion

### *4.1. Comparison of the Genetic Algorithm with the Brute Force Algorithm*

The BFA finds the global minimum of $\chi^2$-minimization and is used as a reference to check the performance of the GA. Due to high time complexity, one can use BFA as a reference only in events with low multiplicities ($5 \leq N \leq 14$). The distributions of errors, $\Delta t_0 = t_0 - t_0^{\text{true}}$, for such events are presented in Figure 2. Only $\pi^{\pm}, K^{\pm}, \text{p}^{\pm}$ are used as allowed types of particles. Both BFA and GA provide an unbiased estimation of the reconstructed event collision time with resolutions of 29 ps for BFA and 30 ps for GA.

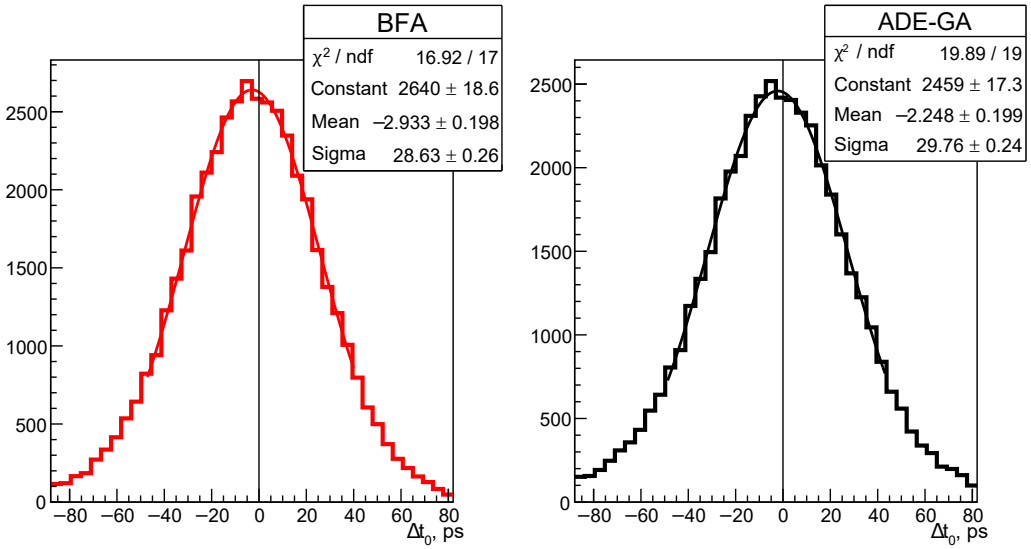


**Figure 2.** Distributions of errors, $\Delta t_0 = t_0 - t_0^{\text{true}}$, for the event collision time reconstructed by (**left**) the Brute-Force Algorithm (BFA) and (**right**) the Asynchronous Differential Evolution-inspired Genetic Algorithm (ADE-GA). The curve represents the Gaussian fit with the parameters shown; "ndf" stays for the numbers of degrees of freedom.

Another important metric is the overall percentage of tracks that were identified correctly: 97.2% for BFA and 96.8% for GA. Non-zero PID inefficiency by BFA looks counter-intuitive, but such an inefficiency appears due to the finite resolution of the TOF detector when the uncertainty of its measurement exceeds the typical time-of-flight difference between two different particle types at a given momentum. In that case, a particle will be misidentified if the global minimum of $\chi^2$-minimization is deeper than the $\chi^2$ of the actual particle configuration. One should note that fast tracks, whose particle type can not be

resolved by the time-of-flight technique, contribute to above cited PID inefficiencies, but are useful for the correct determination of the event collision time in case the difference $(t_i - \tau_{ij})$ in Equation (4) is insensitive to particle species.

The GA performance to solve the $\chi^2$-minimization problem is on par with the exhaustive search, but it demonstrates a different time complexity for high-multiplicity events; see Figure 3. While for events with less than 8 tracks, BFA has a shorter run time, it exponentially slows down as multiplicity grows. The average run time of BFA on events with $5 \leq N \leq 14$ is 5 ms, while GA runs much faster—160 µs. Both BFA and ADE-GA are intrinsically parallel algorithms. The run-times cited have been measured in a single-thread calculation mode to simplify the comparison. Faster execution time can be achieved with multithreading.

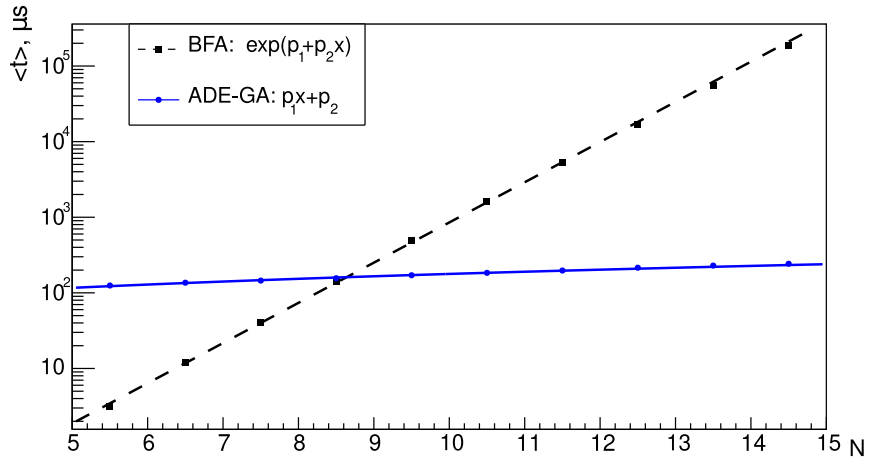


**Figure 3.** Time complexity comparison of the BFA and ADE-GA: the median run-time, $\langle t \rangle$, as a function of event multiplicity, $N$. The functions shown are the parametrizations of the $\langle t \rangle(N)$.

The ADE-GA analysis of events with any number of reconstructed tracks confirms that the uncertainty in the collision time, $\sigma_0$, decreases from about 32 ps for 5-track events down to about 20 ps for high-multiplicity events and scales as $1/\sqrt{N}$. The achieved uncertainty in $t_0$ is much better than the resolution of the TOF detector, $\sigma_t = 70$ ps, thus the latter will dominate the uncertainty in the time-of-flight between the collision point and the TOF detector. The efficiency of the ADE-GA to accurately measure the collision time for events with more than 4 reconstructed tracks is estimated to be about 97%.

The ability of the ADE-GA algorithm to efficiently solve the global minimization problem (2) defined in discrete space is based on the following fundamental principles of the DE [16,18]. First, the algorithm is derivative-free, thus it can perform optimization over discrete variables. Second, DE does not use any assumption about a particular shape of the minimized fitness function, e.g., relying on the linear or quadratic approximation of its shape. Instead, DE adapts its population to a particular landscape through natural (Darwinian) selection. Better candidate solutions have higher chances to stay in a population for a longer time thus more often playing the parent (central) role in the mutation operator ($v_{\text{par}}$ in Equation (5)). In this way, the population's center of gravity is gradually shifted to the deeper minimum in case of multimodal problems. Last but not least, due to the common convergence of population members to a minimum, the algorithm automatically adapts the difference vector $\Delta v$ in Equation (5)) to a typical size of the search region around the minimum. The latter feature enables the comparatively fast convergence of DE.

Thanks to the swiftness of the GA, a wider than $N_m = 3$ range of possible mass types can be taken into account. If electrons/positrons are added into possible particle types then the reconstructed event collision time becomes biased (Figure 4). Due to short flight paths, the expected arrival time of pions with momenta above 1 GeV/$c$ to the TOF detector is delayed to electrons less than the TOF time resolution. In this case, some pions are

misidentified as electrons whenever such a mass hypothesis provides a deeper minimum for $\chi^2$. As pions are much more abundant than electrons/positrons such misidentification results in a biased estimation of the collision time.

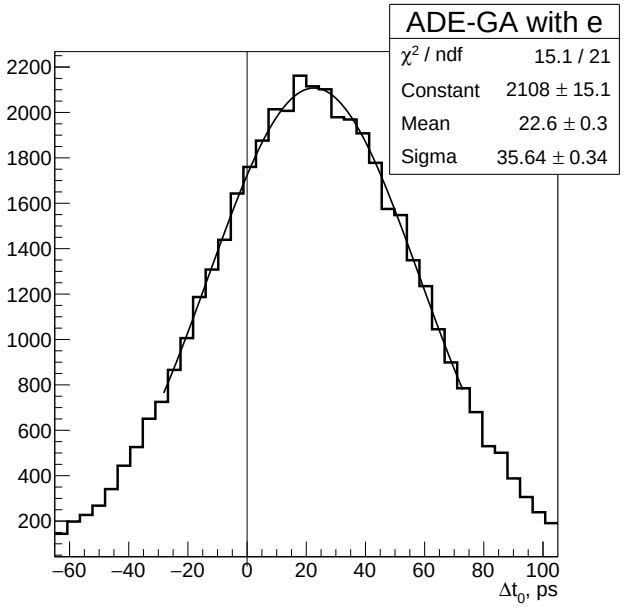


**Figure 4.** Distributions of errors, $\Delta t_0 = t_0 - t_0^{\text{true}}$, for the event collision time reconstructed by ADE-GA with $e^{\pm}, \pi^{\pm}, K^{\pm}, p^{\pm}$ species. The curve represents the Gaussian fit with the parameters shown.

To remedy the probability of misidentifications, one can remove from consideration any track whose type is not identified in a unique way for sure (for example, by $n$-sigma criteria, see Section 4.3 below). This approach was derived in Ref. [19]: using a priori knowledge about the dominant prevalence of pions in the sample of registered particles, one can consider all low-momenta tracks as pions; for each track, estimate the collision time, identify the most probable one and reject all heavier-than-pion particles. In this way, the collision time can be found with uncertainty about 32 ps for events with a fairly high number of tracks, but not in events with less than 3 pions below 1.5 GeV/$c$.

As the accurate and unbiased estimation of the event collision time is the main goal of this study, the correctness of the obtained $t_0$ value is further verified by iteratively removing major addends from the $\chi^2$-sum (2) followed by the $\chi^2$-minimization over the rest of tracks in the event. A statistically significant shift of the $t_0$ value indicates a possible outlier due to noise or misidentification.

### *4.2. Alternative Ways to Measure the Event Collision Time*

Alternatively, the event collision time can be measured by dedicated detectors installed close to the beam tube, so-called T0 detectors [20]. Such detectors typically have fine granularity to cope with a high load of secondary particles and protons scattered on small angles.

In SPD, the intersection region of two colliding beams will cover a few tens centimeters along the beam axis. This demands installing a pair of T0 detectors, located from both beam directions around a collision point, to be used in combination. The Monte-Carlo simulation shows that, if T0 detectors cover polar angles between 60 and 500 mrad, then only about half of pp collisions at $\sqrt{s} = 27$ GeV can produce charged tracks in both forward and backward T0 detectors [1]. This limits the ability of T0 detectors to measure the event collision time in the SPD experimental conditions. Meantime, T0 detectors can determine $t_0$ for events, where other detectors can not be used, e.g., in the case of elastic scattering.

With the T0 detectors, only a few tracks are involved in the event collision time determination, while $t_0$ measurement by the TOF detector uses many tracks and improves as $1/\sqrt{N}$ for high-multiplicity events. Moreover, measurements by the TOF detector are accompanied by reconstructed tracks that further reduce uncertainties due to uncertainties in time-of-flight distance and particle momentum.

Another approach to analyze high-multiplicity events was developed in the ALICE experiment [10]. Particles originated from pp-collisions at $\sqrt{s} = 7$ TeV (or from p-Pb, Pb-Pb collisions) and detected by the TOF detector were divided into ten momentum intervals. The time of the event was calculated in subsamples and the weighted average of the results was then taken. This strategy reduces the computational time but mathematically is not equivalent to the direct minimization of the $\chi^2$-function (2). The reported results show the high efficiency of the ALICE's mitigation strategy for high-multiplicity events ($N > 15$), but efficiency decreases towards lower multiplicities, which are typical at SPD running conditions.

### *4.3. Particle Identification by Time-of-Flight*

Using the reliable method to reconstruct the event collision time $t_0$, one can perform particle identification through a comparison of track timing by the TOF detector to the expected time of the particle's arrival at the detector. There are several strategies for PID by the time of flight:

1. One can assign particle type for each track from the result of $\chi^2$-minimization: the track type is accepted as the most likely species (maximal probability).
2. On the other hand, for every track $i$ in event, one can exclude this track from the determination of the collision time $t_0$ to avoid correlations. Let us denote as $t_{i0}$ the event collision time calculated over the rest of the tracks in the event. Then there are two common strategies to perform PID by time-of-flight [14]:
   (a) $n$-sigma selection—the most simple threshold discriminator:

$$n_{ij} = \frac{t_i - (t_{i0} + \tau_{ij})}{\sigma_{ij}} = \frac{S_i - \hat{S}_i(m_j)}{\sigma_{ij}}. \tag{7}$$

   Here $S_i$ is a signal obtained for track $i$, $\hat{S}_i(m_j)$ is the expected signal for a particle of species $j$ with momenta $p_i$. If the signal belongs to the range $\pm 2\sigma$ (standard deviation) or $\pm 3\sigma$ of a certain species, this track is accepted as the particle of this species. The track can be accepted as multiple species.
   (b) Bayesian method: takes into account yield of particle species. The conditional probability for track $i$ to be a particle of species $j$ reads:

$$P(H_j|S_i) = \frac{P(S_i|H_j)C(H_j)}{\sum_{\alpha=\pi,K,\mathrm{p}} P(S_i|H_\alpha)C(H_\alpha)}. \tag{8}$$

   Here $C(H_j)$ is a prior probability that is calculated iteratively. It takes into account the relative abundance of species $j$, which depends on particle momenta and emission angle. The likelihood function, $P(S_i|H_j)$, is given by

$$P(S_i|H_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left(-\frac{1}{2}n_{ij}^2\right). \tag{9}$$

Separation power, $n\sigma_{\pi K} = (\tau_{iK} - \tau_{i\pi})/\sigma_{iK}$, can be used as a measure of the PID performance [10]. In the SPD, identification of particles by their time-of-flight can be performed up to 1.7 GeV/$c$ for $\pi/K$ separation and up to 3 GeV/$c$ for $K$/p at $3\sigma$-level, see Figure 5.

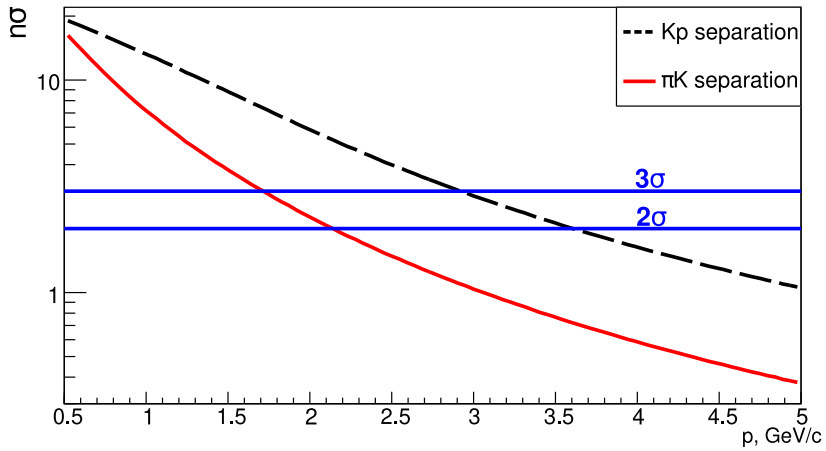


**Figure 5.** $\pi/K$ and $K/\mathrm{p}$ separation powers as functions of momenta. See text for details.

### *4.4. PID Benchmarks for Two-Prong Decay Channels*

Different PID methods have been compared by reconstructing several decay channels with two oppositely charged particles in the final state: $\phi \to K^+K^-$, $\Lambda^0 \to \mathrm{p}^+\pi^-$ and $K^0_s \to \pi^+\pi^-$. For this study, two-prong decay channels have been chosen due to the smaller combinatorial background with respect to multi-prong decays. Only PID by time-of-flight is used in this section. One should note that in real data analysis, it will be accompanied by other methods to reduce background: secondary vertex reconstruction of intermediate particles, particle identification by ionization losses, etc.

In Figures 6–8 (left), the invariant mass of all pairs of oppositely charged tracks is shown. The same combination of masses has been assigned to each pair as for the channel of interest (indicated "no pid"). In the $n$-sigma approach only tracks with TOF signals within $\pm 3\sigma_{ij}$ of a certain species $j$ are selected ("3 sigma"). In the case of the weighted Bayesian PID, all combinations are included with the pair's weight as a product of conditional probabilities (8) for each prong ("bayesian"). The possible pair combinations corresponding to the global minimum of the $\chi^2$-minimization are marked as "chi2_min". Finally, the "ideal" corresponds to Monte-Carlo combinations with known particle types.

The weighted Bayesian approach, which exploits both PID-by-TOF capabilities and the abundance of particle species, provides the best suppression of background while preserving particles of interest; see Figures 6–8 (right). Kaons are less abundant than pions and protons, therefore the advantage of the Bayesian approach is more pronounced if kaons are the decay products (Figure 6 (right)). By applying the $n$-sigma approach one can preserve more signal events, but at the same time, the suppression of the combinatorial background is reduced. Benchmarks analyses, shown in Figures 6–8, demonstrate the power of the identification of particles by the time-of-flight method in the SPD experimental conditions.

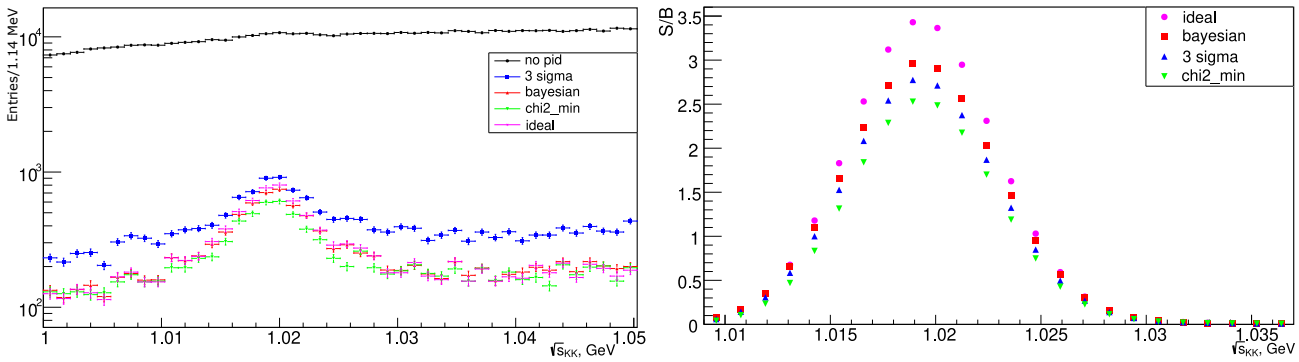


**Figure 6.** Invariant mass (**left**) and signal-to-background ratios (**right**) as a function of the center-of-mass energy, $\sqrt{s_{KK}}$, of pairs of oppositely charged tracks assumed to be $K^+K^-$ under different PID strategies as indicated. Signal of $\phi \to K^+K^-$ decays stands out of the combinatorial background, which is reduced by a factor greater than 20. See text for details.

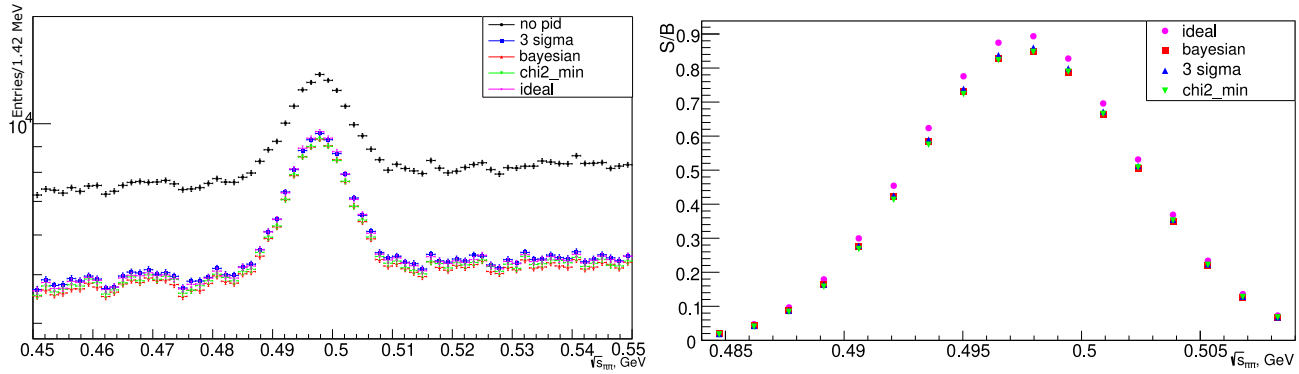


**Figure 7.** Invariant mass (**left**) and signal-to-background ratios (**right**) as a function of the center-of-mass energy, $\sqrt{s_{\pi\pi}}$, of pairs of oppositely charged tracks assumed to be $\pi^+\pi^-$ under different PID strategies as indicated. Combinatorial background to $K^0_S \to \pi^+\pi^-$ decays is suppressed by a factor of $\sim$2. See text for details.

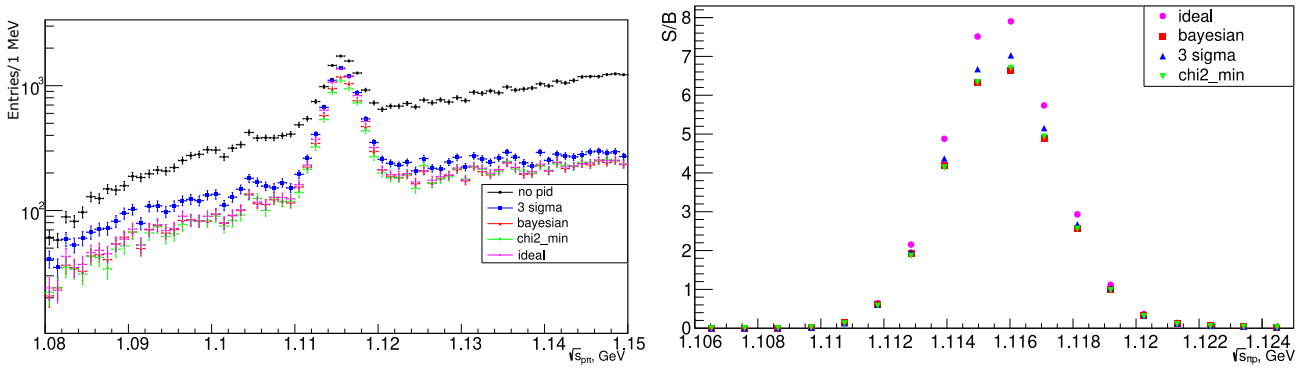


**Figure 8.** Invariant mass (**left**) and signal-to-background ratios (**right**) as a function of the center-of-mass energy, $\sqrt{s_{p\pi}}$, of pairs of oppositely charged tracks assumed to be $p\pi^-$ under different PID strategies as indicated. Combinatorial background to $\Lambda^0 \to p\pi^-$ decays is suppressed by a factor of $\sim$3. See text for details.

## 5. Conclusions

In the SPD experiment, the accurate determination of the event collision time is required to perform $\pi/K/p$ identification in the low-momenta range (0.5–3 GeV/$c$) by the time-of-flight method. The collision time can be reconstructed on an event-by-event basis by a minimization procedure, which uses measurements of the TOF detector combined

with track reconstruction, but the corresponding solution is computationally expensive. In this paper, we present a dedicated Asynchronous Differential Evolution-inspired Genetic Algorithm (ADE-GA), which is shown to solve the optimization problem without any simplifications, thus providing a fast and reliable measurement of the event collision time throughout the range of track multiplicities. Finally, different strategies of particle identification by time-of-flight are tested to prove the power of the PID-by-TOF method in the SPD experimental conditions.

**Author Contributions:** Conceptualization, supervision: M.Z.; formal analysis, methodology, software, investigations: S.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was in part supported by the JINR START 2022 program.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors would like to express their deepest appreciation to members of the SPD Collaboration for their valuable feedback.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| ADE-GA | Asynchronous Differential Evolution-inspired Genetic Algorithm |
| ALICE | A Large Ion Collider Experiment |
| ANKE-COSY | The Apparatus for studies of Nucleon and Kaon Ejectiles at the COoler SYnchrotron |
| BFA | Brute Force Algorithm |
| DE | Differential Evolution |
| GA | Genetic Algorithm |
| JINR | Joint Institute for Nuclear Research |
| LHC | Large Hadron Collider |
| MRPC | Multigap Resistive Plate Chamber |
| NICA | Nuclotron-based Ion Collider fAcility |
| PID | Particle IDentification |
| QCD | Quantum Chromodynamics |
| RHIC | Relativistic Heavy Ion Collider |
| SPD | Spin Physics Detector |
| TOF | Time-Of-Flight |

## References

1. Abazov, V.M. et al. [SPD Collaboration]. Conceptual design of the Spin Physics Detector. *arXiv* **2021**, arXiv:2102.00442. [CrossRef]
2. Dymov, S. [ANKE Collaboration] Recent results from the NN-interaction studies with polarized beams and targets at ANKE-COSY. *J. Phys. Conf. Ser.* **2016**, *678*, 012014. [CrossRef]
3. Maggiora, M. New results from DISTO for spin observables in exclusive hyperon production. *Nucl. Phys. A* **2001**, *691*, 329c–335c. [CrossRef]
4. Abdallah, M.S. et al. [STAR Collaboration]. Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 200$ GeV. *Phys. Rev. D* **2021**, *103*, L091103. [CrossRef]
5. Hadjidakis, C.; Kikoła, D.; Lansberg, J.P.; Massacrier, L.; Echevarria, M.G.; Kusina, A.; Schienbein, I.; Seixas, J.; Shao, H.S.; Signori, A.; et al. A fixed-target programme at the LHC: Physics case and projected performances for heavy-ion, hadron, spin and astroparticle studies. *Phys. Rep.* **2021**, *911*, 1–83. [CrossRef]
6. Cerron Zeballos, E.; Crotty, I.; Hatzifotiadou, D.; Lamas Valverde, J.; Neupane, S.; Williams, M.C.S.; Zichichi, A. A new type of resistive plate chamber: The multigap RPC. *Nucl. Instrum. Meth. A* **1996**, *374*, 132–135. [CrossRef]
7. Wang, Y.; Yu, Y. Multigap resistive plate chambers for time of flight applications. *Appl. Sci.* **2021**, *11*, 111. [CrossRef]
8. Basile, M.; Cara Romeo, G.; Cifarelli, L.; D'Ali, G.; Di Cesare, P.; Giusti, P.; Massam, T.; Palmonari, F.; Sartorelli, G.; Valenti, G.; et al. A large-area time-of-flight system for a colliding beam machine. *Nucl. Instrum. Meth. A* **1981**, *179*, 477–485. [CrossRef]
9. Akindinov, A.; Alici, A.; Agostinelli, A.; Antonioli, P.; Arcelli, S.; Basile, M.; Bellini, F.; Cara Romeo, G.; Cifarelli, L.; Cindolo, F.; et al. Performance of the ALICE Time-Of-Flight detector at the LHC. *Eur. Phys. J. Plus* **2013**, *128*, 44. [CrossRef]

10. Adam, J. et al. [ALICE Collaboration]. Determination of the event collision time with the ALICE detector at the LHC. *Eur. Phys. J. Plus* **2017**, *132*, 99. [CrossRef]
11. Shao, M.; Barannikova, O.; Dong, X.; Fisyak, Y.; Ruan, L.; Sorensen, P.; Xu, Z. Extensive particle identification with TPC and TOF at the STAR experiment. *Nucl. Instrum. Meth. A* **2006**, *558*, 419–429. [CrossRef]
12. Zimmermann, S.; Suzuki, K.; Steinschaden, D.; Kratochwil, N.; Nalti, W.; Orth, H.; Schwarz, C.; Lehmann, A.; Böhm, M.; Brinkmann, K.-T. The P̄ANDA Barrel Time-of-Flight detector. *Nucl. Instrum. Meth. A* **2020**, *952*, 161635. [CrossRef]
13. Zhabitskaya, E.; Zhabitsky, M. Asynchronous differential evolution with adaptive correlation matrix. In *GECCO'13: Proceedings of the 15th Genetic and Evolutionary Computation Conference, Amsterdam, The Netherlands, 6–10 July 2013*; Blum, C., Ed.; Association for Computing Machinery: New York, NY, USA; pp. 455–462. [CrossRef]
14. Adam, J. et al. [ALICE Collaboration]. Particle identification in ALICE: A Bayesian approach. *Eur. Phys. J. Plus* **2016**, *131*, 168. [CrossRef]
15. Bierlich, C.; Chakraborty, S.; Desai, N.; Gellersen, L.; Helenius, I.; Ilten, P.; Lönnblad, L.; Mrenna, S.; Prestel, S.; Preuss, C.T.; et al. A comprehensive guide to the physics and usage of PYTHIA 8.3. *arXiv* **2022**, arXiv:2203.11601. [CrossRef]
16. Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optimiz.* **1995**, *11*, 341–359. [CrossRef]
17. Zhabitskaya, E.; Zhabitsky, M. Asynchronous Differential Evolution with restart. In *Numerical Analysis and Its Applications*; Dimov, I., Faragó, I., Vulkov, L., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 555–561. [CrossRef]
18. Das, S.; Mullick, S.S.; Suganthan, P.N. Recent advances in differential evolution—An updated survey. *Swarm Evolut. Comput.* **2016**, *27*, 1–30. [CrossRef]
19. Filonchik, P.G.; Zhabitsky, M.V. Fast way to determine pp-collision time at the SPD experiment. *arXiv* **2022**, arXiv:2212.07887. [CrossRef]
20. Bondila, M.; Grigorev, V.A.; Guber, F.F.; Kaplin, V.A.; Karakash, A.I.; Karavichev, O.V.; Karavicheva, T.L.; Klimov, A.I.; Kondratieva, N.; Kozlov, K.N.; et al. ALICE T0 detector. *IEEE Trans. Nucl. Sci.* **2005**, *52*, 1705–1711. [CrossRef]