Journal of Physics: Conference Series

Vertex Reconstruction at STAR: Overview and Performance Evaluation

To cite this article: D. Smirnov *et al* 2017 *J. Phys.: Conf. Ser.* **898** 042058

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 042058 doi:10.1088/1742-6596/898/4/042058

Vertex Reconstruction at STAR: Overview and Performance Evaluation

D. Smirnov, J. Lauret, V. Perevoztchikov, G. Van Buren, and J. Webb

Physics Department, Brookhaven National Laboratory, Upton, NY 11973, USA

E-mail: `dsmirnov@bnl.gov`

Abstract.
The STAR experiment at the Relativistic Heavy Ion Collider (RHIC) has a rich physics program ranging from studies of the Quark Gluon Plasma to the exploration of the spin structure of the proton. Many measurements carried out by the STAR collaboration rely on the efficient reconstruction and precise knowledge of the position of the primary-interaction vertex. Throughout the years two main vertex finders have been predominantly utilized in event reconstruction by the experiment: `MinutVF` and `PPV` with their application domains focusing on heavy ion and proton-proton events respectively. In this work we give a brief overview and discuss recent improvements to the vertex finding algorithms implemented in the STAR software library. In our studies we focus on the finding efficiency and the quality of the reconstructed primary vertex. We examine the effect of an additional constraint, imposed by an independent measurement of the beam line position, when it is applied during the fit. We evaluate the significance of the improved primary vertex resolution on identification of the secondary decay vertices occurring inside the beam pipe. Finally, we present a method and its software implementation developed to measure the performance of the primary vertex reconstruction algorithms.

1. Introduction

In 2014 the STAR experiment was upgraded with the Heavy Flavor Tracker (HFT) [1], a group of three high precision silicon-based detectors. The data recorded from the HFT detectors for three consecutive runs after the installation warranted improvements in track reconstruction and benefited the pointing resolution of the existing Time Projection Chamber (TPC) [2]. The significant improvement in the primary vertex resolution resulting from this upgrade prompted us to revisit the vertex reconstruction algorithms currently employed by the STAR collaboration. It was also noted that for physics analyses relying on event-by-event topological event selection a good vertex resolution is crucial to identify displaced vertices from decays of short-lived particles such as heavy-flavor mesons and hyperons.

1.1. Tracking with STAR and Pile-up

The TPC is the principle tracking detector in the STAR experiment. It is immersed in a uniform magnetic field of 0.5 T and has an acceptance of $|\eta| < 1.3$ and $0 < \phi < 2\pi$. A charged particle going through the gas of the TPC leaves an ionization trail that drifts away from the central membrane to the ends of the barrel where it is read out. The particle momenta are then

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 042058 doi:10.1088/1742-6596/898/4/042058

determined from the curvature and pitch of the helix reconstructed from the hits registered in the TPC volume. The challenge with the TPC is that its readout time is much longer than the time between the beam collisions. For this reason, the recorded event, initiated by a collision that fired the trigger, is likely to contain tracks from other beam crossings. The effect of pile-up tracks on primary vertices becomes an increasing problem with higher luminosities delivered by the RHIC.

With respect to the triggered collision the pile-up tracks can originate in either preceding or subsequent beam crossings. Besides such out-of-time tracks, additional collisions in the same beam crossing can produce more pile-up tracks when the instantaneous luminosities are high. These additional collisions can have the same properties as the triggered one and, therefore, their tracks cannot be distinguished by means of the TPC alone. In order to mitigate the effect on the main vertex reconstruction due to tracks coming from other than triggered collisions, STAR makes use of fast non-tracking detectors. These detectors include the Barrel and the Endcap Electromagnetic Calorimeters (BEMC and EEMC) and the Time Of Flight (TOF) subsystem. The BEMC and EEMC have the same 2π acceptance in azimuth while in pseudo-rapidity their coverages are respectively $|\eta| < 1$ and $1.1 < \eta < 2$. The coverage of the TOF detector matches that of the BEMC. The readout time for TPC electronics is on the order of 40 μs whereas the calorimeters and the TOF can be reset in 10's of nanoseconds allowing them to be read out for every beam crossing. The next section discusses in more details how we take advantage of the facts that the out-of-time tracks are usually of lower quality and do not deposit any energy in the fast detectors.

2. STAR Vertex Finders

Two different vertex finders, `MinutVF` and `PPV`, have been predominantly utilized by STAR to cope with the distinctive nature of heavy-ion and proton-proton collisions. Historically, `MinutVF` was optimized for high multiplicity $A + A$ events whereas `PPV` was mainly optimized for low multiplicity $p + p$ events with a high pile-up rate. Both vertexing algorithms have been previously reported to demonstrate a good overall performance [3].

Vertex reconstruction is an integral part of the full event reconstruction at STAR. Once the tracking algorithm finishes the reconstruction of tracks from hits left by charged particle in the sensitive detector material, the control flow is passed to a vertex-finding algorithm selected by the user. The tracks that can be associated with a collision vertex are called *primary* tracks. It is customary to identify the following common stages pertaining to most vertexers including the STAR ones:

- **Track selection.** Substantial requirements are applied to select tracks which are allowed to participate in reconstruction of primary vertex candidates.
- **Vertex seeding.** Approximate spatial positions (usually along the z axis, i.e. the nominal beam line) of vertex candidates are estimated from the clusters of tracks believed to originate from the same interaction points.
- **Track fitting.** Using all tracks compatible with a vertex candidate, a precise position and the corresponding covariance matrix are determined. At this stage the tracks consistent with the final position are designated as primary tracks.
- **Ranking of vertices.** While not widely adopted outside of STAR the ranking step is aimed to sort the final vertices such that the highest-rank vertex is the best match for the primary interaction point.

It is evident that distinguishing between the vertex reconstruction steps as just described is not always possible. Some elaborate algorithms, such as adaptive vertex fits, may not have a dedicated seeding step while doing some track rejection during the fit. However, the STAR algorithms are the classical ones with distinct stages as outlined above.

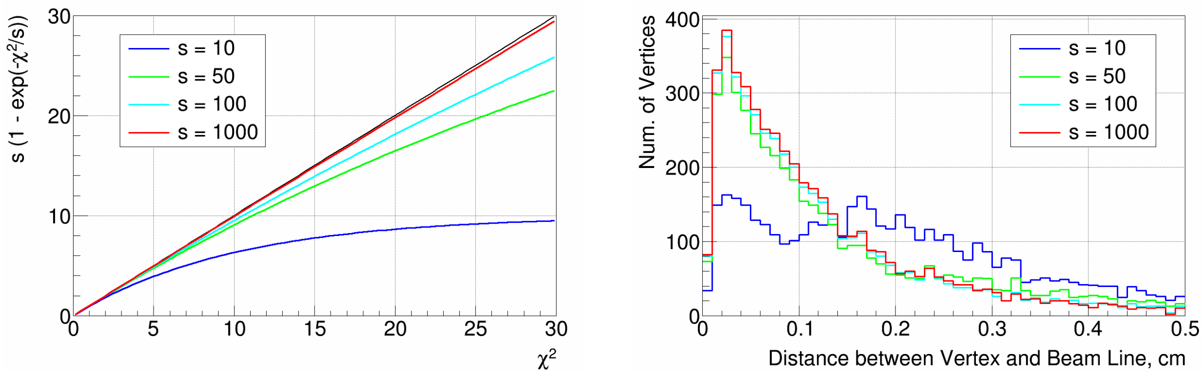
Both STAR vertex finders apply similar track quality cuts to populate the initial collection of tracks used for seeding. The tracks are required to have a distance-of-closest-approach (DCA) to the z-axis less than the beam pipe radius (2 or 3 cm depending on the STAR configuration) and do not project beyond the longitudinal TPC dimension $|\mathrm{DCA}_z| < 200$ cm. The minimum number of TPC hits on track is another principle selection criteria that is applied. These and other parameters can be tuned depending on the type of data being reconstructed.

The seeding procedure in both `MinutVF` and `PPV` is based on finding clusters of tracks along the z-axis. In `MinutVF` a simple distribution of tracks is constructed by extrapolating them to the nominal beam line position. In `PPV` the same distribution is weighted by the uncertainty of the track's projection at the point of closest approach to the beam line. In addition, the `PPV` method increases the track's weight if it can be extrapolated to either an EMC or TOF cell with a signal above a certain threshold. Also, the method boosts the weight of higher quality tracks with an established continuity at the central TPC membrane while those with a discontinuity are rejected. The purpose of such re-weighting is to minimize the contribution from the pile-up tracks while giving preference to the tracks coming from the triggered beam crossing. The vertex candidates are then determined by iteratively finding local maxima in the binned longitudinal DCA distributions. In fact, in the earlier version of `PPV` this was the last step of the algorithm as there was no refitting performed for the tracks associated with the vertex seed. In the revised `PPV`, we added the same fitting procedure as in `MinutVF` which includes a proper calculation of the full covariance matrix.

After the seeding stage, the fit in `MinutVF` is performed for every vertex candidate starting with tracks in some proximity. The precise position is then determined for each vertex seed using a robust χ^2-based fit where contributions from outliers are down-weighted comparing to the normal χ^2 calculation. The modified χ^2 contributes to the objective function with the following term:

$$s\left(1 - \exp\left(\frac{-\chi^2}{s}\right)\right) \xrightarrow[s\to\infty]{} \chi^2, \tag{1}$$

where s is a scale factor to control the level at which a measurement should be qualified as an outlier. With large values of s expression (1) asymptotically approaches the normal χ^2 as



(a) (b)

Figure 1. (a) Comparison of contributions from robust χ^2 given by (1) evaluated for different values of scale s. (b) The effect of small s applied to the beam contributing to the measurement is illustrated by the bias in the vertex position with respect to the beam position. With $s = 10$ the position is primarily defined by the projected tracks rather than by the precise measurement of the beam line.

illustrated in Figure 1(a). By adjusting the scale, the vertex position can be protected from a significant bias caused by accidental high precision tracks, such as those with HFT hits. A possible application of this feature is explained in Figure 1(b) using the beam line with small transverse uncertainties as an example.

In PPV the ranking is an integral part of the seeding stage due to the appropriate re-weighting of the track DCA distribution. The seeds are selected starting with the highest local maxima until all tracks are attributed. This naturally orders the vertex candidates by their probability to match the primary event vertex. In MinutVF, however, the ranking is done as a dedicated last step which implements the same principle of assigning higher rank to vertices with higher fraction of tracks pointing to the fast detectors as in PPV. In addition, the average angle between the daughter tracks and the z-axis is calculated for each vertex. This average angle correlates with the z coordinates of the vertex due to the TPC acceptance. Vertices with larger deviations from the nominal correlation are assumed less likely to represent the primary event vertex and, thus, are given a smaller rank.

3. Primary Vertex Fit with Beam Line Constraint

When the position of the colliding beams is known, it can be used as an additional measurement of the primary interaction point. At STAR the beam line is calculated by finding vertices inside the beam pipe using high-multiplicity events. The tracks are fitted using MinutVF without any additional constraints on the vertex position. The obtained vertex position distributions are projected onto the xz and yz planes (see Figure 2), and a straight line is then fit to the data to extract a parametrization for the beam line as:

$$\begin{cases} x = x_0 + z \tan \phi_{xz} \\ y = y_0 + z \tan \phi_{yz} \end{cases} \tag{2}$$

The beam line calibration outlined above can potentially lead to very small statistical uncertainties on the extracted parameters $x_0, y_0, k_{xz} \equiv \tan \phi_{xz}$, and $k_{yz} \equiv \tan \phi_{yz}$ due to high statistics available for the measurement. Unfortunately, systematic variations on the order of 100 μm have been observed in the vertex x and y distributions along the z-axis. To determine whether changes in the beam shape near the interaction region or some other sources contribute to this effect will require an additional detailed study. For the purposes of this work, we assume that the resolution of the beam line constraint in xy is limited by the nominal width of the beam which is estimated to be about 100 μm.

Both MinutVF and PPV can use the beam line parametrization when estimating the primary vertex position. However, the previous implementation did not take into account the measured uncertainties on the beam line parameters; instead the vertex coordinates were forced to satisfy exactly the beam line equations (2). The fitting stage, thus, would decay to a one dimensional fit with z_v as the only parameter while keeping x_v and y_v fixed by (2).

In the new version of MinutVF and PPV, we added a routine calculating the χ^2 for the vertex and the beam line, i.e. the point of closest approach, in three dimensions. The vector connecting the two points is shown on Figure 2 and can be expressed in our notation as:

$$\mathbf{d} = \frac{1}{1 + k_{xz}^2 + k_{yz}^2} \begin{pmatrix} (k_{yz}^2 + 1)(x_v - x_0) & - & k_{xz}k_{yz}(y_v - y_0) & - & k_{xz}z_v \\ k_{xz}k_{yz}(x_v - x_0) & + & (k_{xz}^2 + 1)(y_v - y_0) & - & k_{yz}z_v \\ -k_{xz}(x_v - x_0) & - & k_{yz}(y_v - y_0) & + & (k_{xz}^2 + k_{yz}^2)z_v \end{pmatrix}$$

The mapping from the beam-line parameter space onto the direction $\mathbf{d}$ is done by means of a Jacobian matrix defined as:

$$J = \left(\frac{\partial d}{\partial x_0}, \frac{\partial d}{\partial y_0}, \frac{\partial d}{\partial k_{xz}}, \frac{\partial d}{\partial k_{yz}} \right).$$

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 042058 doi:10.1088/1742-6596/898/4/042058

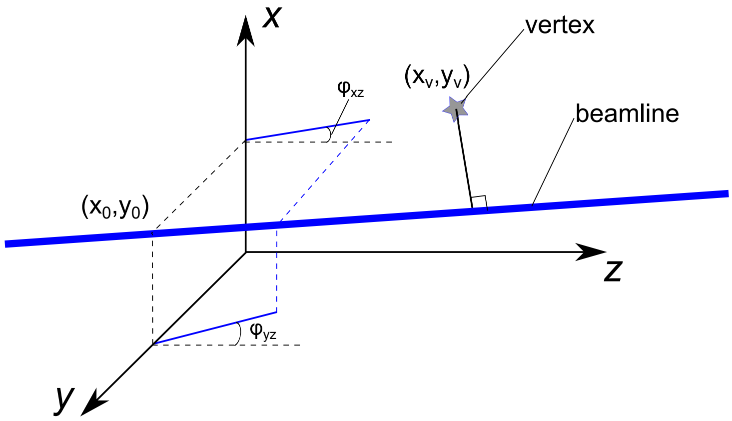


Figure 2. Schematics illustrating the beam line parametrization in 3D by two projections on the xz and yz planes. In newly released STAR vertex finders the vertex position can be additionally constrained by minimizing the χ^2 w.r.t. the shortest distance to the beam line during the fit.

Using this transformation one can easily calculate the variance for $\mathbf{d}$ by propagating the measured uncertainties as $C_d = JCJ^T$ where C is the covariance matrix for the beam-line parameters.

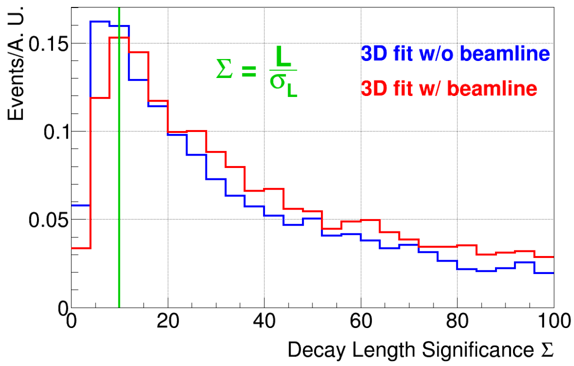
4. Impact on Secondary Decay Vertex

With the new vertex fitting procedure implemented in our software we studied its effect on reconstruction of secondary vertices associated with decays of short-lived particles. The PYTHIA generator [4] was used to produce a sample of events enriched with Λ^0 baryons originating in proton-proton interactions at $\sqrt{s} = 200$ GeV. These signal events were then passed through the full STAR detector simulation and embedded into real data minimally biased events in order to model a realistic background. The primary interaction point was normally distributed along the beam line with a standard deviation of 25 cm and the Λ^0 particles were generated with the minimum transverse momentum P_T of 2 GeV. At this energies a significant fraction of Λ^0's is expected to decay inside the beam pipe that has a radius of 3 cm.

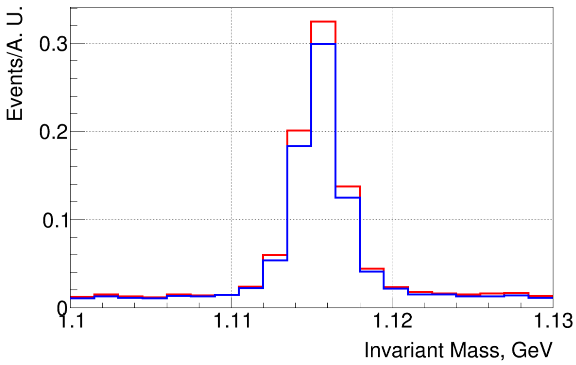
The created sample was fully reconstructed twice with the only difference in how the primary vertex finding was performed. In both passes we used the updated `PPV` algorithm with the 3D fit for the vertex fit, however, in the first scenario the fit was constrained by the beam line, whereas in the second one the fit was unconstrained. In each event thus reconstructed, we search for Λ^0 decays by identifying proton and pion tracks with opposite curvature and a short distance of closest approach. Similar to other STAR analyses we rely on the particle identification based on the track's dE/dx measured in the TPC. From the selected track pairs the momentum of Λ^0 candidates is calculated and the mid-point between the tracks is chosen as the decay vertex position. In addition to a few loose topological selections reducing the background, we require the reconstructed momentum to point back to the primary vertex.

Finally, to compare the results of the constrained and unconstrained primary vertex 3D fits, we plot the significance, $\Sigma = L/\sigma_L$, of the decay length L and its total error σ_L, for the two cases. Figure 3(a) demonstrates that events move towards the tail of the distribution as expected due to the inclusion of the high precision measurement in the fit. This is a positive effect if one chooses to cut on this variable to further reduce the background. In our trial study we observed a 10% increase in the signal comparing to the unconstrained vertex fit when select events with $\Sigma > 10$.

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 042058 doi:10.1088/1742-6596/898/4/042058



(a)



(b)

Figure 3. The decay length significance $\Sigma = L/\sigma_L$ is defined as the ratio of the distance L between the primary and a secondary vertex formed by the $\Lambda^0 \to p\pi^-$ decay products and the total uncertainty σ_L associated with the reconstructed vertices (a) and the invariant mass of all proton-pion pairs in the event (b). The red and blue curves correspond to the vertex fits with and without the beam line constraint.

5. Performance Evaluation Metric

The vertex reconstruction efficiency is one of the commonly used metrics representing the performance of a vertex finding algorithm. The calculation of the efficiency usually has to rely on the information available in the simulated data. In our evaluation we utilize a conventional reconstruction efficiency defined as a ratio of the number of reconstructed vertices to the number of reconstructible vertices. The latter is experiment-specific and, in case of STAR, a simulated vertex is regarded as reconstructible if it has at least one charged particle associated with it that can produce a reconstructible track with at least 15 TPC hits. In turn, the reconstructed vertices are required to match one of the simulated vertices in order to be accounted for in the efficiency calculation. For this purpose we employ a track-based truth matching algorithm which unfolds as follows:

- For each track assigned to a reconstructed vertex we determine whether it can be matched with a simulated Monte-Carlo particle. The matching MC particle is the one that contributes the most number of hits to the reconstructed track.[1] It conveys its unique identifier to the track as well as that of its original parent vertex. The fraction of the winning hits on the track is also recorded and used as a weight in the next step.
- The association of a reconstructed vertex to its simulated counterpart proceeds in a manner similar to that of individual tracks and their MC hits. More specifically, each reconstructed track votes for its presumed MC vertex according to its own weight determined in the previous step. The simulated vertex with the highest cumulative weight then wins.

An alternative way to associate a vertex with a true one can be done by finding the pair with the minimal distance between them. However, this method is more likely to overestimate the efficiency due to possible mismatch of ghost vertices found too close to the true vertex.

The reconstruction efficiency can depend on many parameters of the event therefore, having a single number representing the overall efficiency is not always advisable. We choose to plot the efficiency as a function of reconstructible track multiplicity. As evident from Figure 4 the efficiency has a steep turn-on shape growing rapidly with track multiplicity. As defined above, the truth matching efficiency can be used to tune various parameters of the vertex finding

[1] Thus identified MC particle is commonly referred to as the dominant contributor.

algorithms or vertex selection requirements applied by a specific analysis. We illustrate this with four different windows (1, 2, 3, and 5σ wide) within which a track is considered to match a reconstructed vertex. The subplots in Figure 4 correspond to the cases where the true vertex is found among the entire set of reconstructed vertices (left) and the best one with the maximum rank (right). While the wider window can recover more low multiplicity vertices the quality of the highest-rank primary vertex may suffer from incorrectly associated nearby tracks. Therefore, the balance between the overall efficiency and the purity of the highest-rank vertex relied on by many STAR analyses can be achieved.

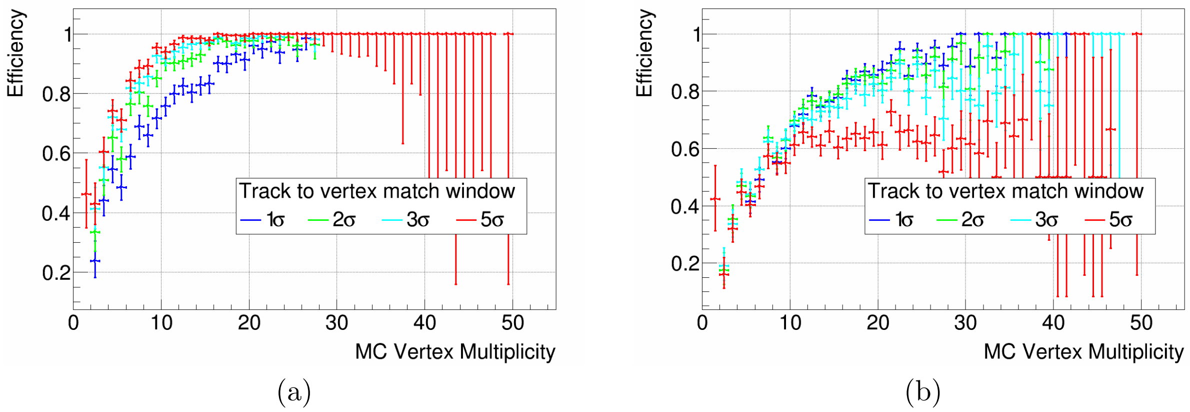


Figure 4. Truth-matching vertex reconstruction efficiency as a function of reconstructible track multiplicity, i.e. the number of reconstructible tracks associated with the simulated vertex. The efficiencies are shown for 1, 2, 3, and 5σ windows within which a track is associated with a found vertex. The overall (a) and the highest-rank (b) vertex efficiencies can be monitored to find a balance between the efficiency and purity of the reconstructed vertices.

While the reconstruction efficiency is an important quantity to monitor, it does not always fully represent the quality of the reconstructed vertex. In order to get a better picture of the vertex finder performance, we also look at such distributions as the distance between the simulated and the reconstructed vertex, uncertainties on the reconstructed position, and pull distributions. These and other quantities of interest can be plotted with our tool developed for STAR [6], however, a dedicated effort is ongoing to make it independent of any experiment as much as possible.

6. Conclusions

A new version of the STAR vertex reconstruction library has been released [5]. The added features include the ability of both `PPV` and `MinutVF` vertex finders to estimate the position and the corresponding covariance matrix of the reconstructed vertices in a three-dimensional fit. In addition, the user can optionally incorporate in the fit an independent high precision measurement of the beam line in order to further constrain the position of the primary interaction. This also expands the available modes for `PPV` by adding the possibility to run without a prior knowledge of the beam line. The use of the algorithm is thus extended to cases when beam line calibration itself is desired, a feature previously available only in `MinutVF`.

We have demonstrated that a proper accounting for the beam line uncertainties can lead to an improved secondary vertex identification due to increased significance of the decay length. In a pilot test with simulated Λ^0 particles decaying to a proton and a pion we observed a 10% increase in the signal comparing to the unconstrained vertex fit.

7. Acknowledgments
This work was supported in part by the Office of Nuclear Physics within the U.S. DOE Office of Science.

References
[1] D. Beavis *et al.*, "The STAR Heavy Flavor Tracker Technical Design Report," drupal.star.bnl.gov/STAR/starnotes/public/sn0600 (2011).
[2] M. Anderson *et al.*, "The Star time projection chamber: A Unique tool for studying high multiplicity events at RHIC," Nucl. Instrum. Meth. A **499**, 659 (2003).
[3] R. Reed, J. Balewski, L. S. Barnby, A. Ogawa, J. Lauret and M. van Leeuwen, "Vertex finding in pile-up rich events for p+p and d+Au collisions at STAR," J. Phys. Conf. Ser. **219**, 032020 (2010).
[4] T. Sjostrand, S. Mrenna and P. Skands, JHEP05 (2006) **026**, Comput. Phys. Comm. **178** (2008) 852.
[5] The source code for the STAR vertex reconstruction library is available at https://github.com/star-bnl/star-vertex. The results presented herein are based on version `v2.0`
[6] The source code for the tools evaluating the performance of the STAR vertex finding algorithms is available at https://github.com/star-bnl/star-travex