Kink Finding in MDC
Using Graph Neural Networks

Xinghui Yang[1] Jin Zhang[1] Xiaorong Zhou[2]

[1]Sun Yat-sen University [2]University of Science and Technology of China

The Super Tau Charm Facility Workshop
July 1, 2026

Outline

Introduction

Kink Finding Using Graph Neural Networks (GNN)

Summary and Outlook

Introduction

A sudden change in direction ('kink') of the charged particle can also cause tracks with breakpoints inside the track system.

Such tracks cannot be reconstructed properly by the existing track finding algorithms, leading to low efficiency in reconstructing particles like $K^{\pm}$, $\pi^{\pm}$, $\mu^{\pm}$, etc.

Some causes of kinked tracks include:

- Decay in flight: e.g. $K^{-}(\pi^{-}) \rightarrow \mu^{-}\bar{\nu}_{\mu}$ and $\mu^{-} \rightarrow e^{-}\bar{\nu}_{e}\nu_{\mu}$.
- Bremsstrahlung: e.g. high-energy electrons emitting photons.
- Multiple scattering: e.g. interactions with detector material.

The kinked tracks lead to several issues in tracking reconstruction:

- Lower track finding efficiency and Poor momentum resolution
- Creating a pair of duplicates known as clone tracks.
- Misidentification(fake rate) in particle identification (PID), like $K^{\pm} \rightarrow \pi^{\pm}X$

Introduction

Multilayer Drift Chamber (MDC)

In experiments like BESIII, Belle II, and STCF, a core component of the tracking system is the multilayer drift chamber, such as the CDC (Central Drift Chamber) in Belle II or the MDC in BESIII and STCF. When charged particles traverse the large volume of the MDC, they are prone to producing kinks through various physical mechanisms, especially via decay in flight.

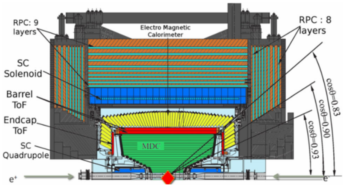


Figure: BESIII Detector

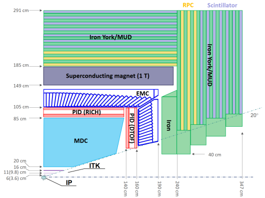


Figure: STCF Detector

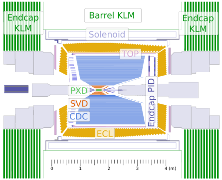


Figure: Belle II Detector

Introdution

Decay in flight in MDC

Generate kaon and pion samples using particle gun at STCF:

- Software Version: OSCAR v2.6.2
- MC samples: 180k K^- and π^-, with the range of momentum $100 \sim 1150\mathrm{MeV}/c$ for kaon and $50 \sim 1100\mathrm{MeV}/c$ for pion.
- Decay modes:

Particle	Decay Mode	Branching Ratio (%)	Decay in MDC(%)	Mass (MeV/c^2)	$c\tau$(m)
K^-	$\mu^-\bar{\nu}_\mu$	63.56	8.19	463.68	3.7
	$\pi^-\pi^0$	20.67	2.66	-	-
	$\pi^-\pi^+\pi^-$	5.58	0.71	-	-
	$\pi^0\mu^-\bar{\nu}_\mu$	5.07	0.63	-	-
	$\pi^0 e^-\bar{\nu}_e$	3.35	0.42	-	-
	$\pi^-\pi^0\pi^0$	1.76	0.24	-	-
	Total	99.99	12.85	-	-
π^-	$\mu^-\bar{\nu}_\mu$	99.99	2.81	139.57	7.8

Introduction

Kink Finding Preformance at Belle II

- Traditional kink finding algorithm at Belle II[1]
- Two type of kink topologies studied:
  - Type I: the daughter and parent tracks are individually reconstructed(10 – 20%).
  - Type II: the daughter and parent tracks are combined one track(40 – 70%).
- MC simulation for $K^-(\pi^-) \to \mu^-\bar{\nu}_\mu$ from $\tau^- \to K^-(\pi^-)\nu_\tau$ decay:
  - PID fake rate: correlated with Type I, but too small at Belle II.
    (Kink Finder has improved 1% for kaon and 5% for pion)
  - Too few hits inside CDC: $\sim$ 10% can not reconstructed for daughter track.
  - Track reconstructed efficiency for Type I: above 85% for parent track and 14 $\sim$ 30% for daughter track(wrong charge about 5 $\sim$ 20%).
  - Overall reconstruction performance: The tracking efficiency with the KinkFinder is 40%, a slight but limited improvement compared to the 11% from standard reconstruction in Belle II.

[1] https://arxiv.org/pdf/2511.14171

Introduction

Challenges and Improvements

- Main challenges faced by traditional kink finding algorithms at present:
  - Low reconstruction efficiency of daughter tracks: Existing algorithms struggle to reconstruct the short track segments of daughter particles accurately.
  - Misidentification of normal tracks: Normal long tracks are easily misidentified as kink events, decreasing efficiency and purity when separating normal tracks.
  - High computational cost: Traditional geometry-based algorithms are computationally intensive, causing long offline data processing times.
- Potential improvements:
  - Enhancing traditional algorithms: Introducing geometric constraints for coupled helices at the vertex, or including often-missed information from stereo hits in reconstruction.
  - Introducing Machine Learning: Utilizing deep learning algorithms, such as Graph Neural Networks (GNNs), which are currently being extensively developed in Belle II.

Kink Finding Using Graph Neural Networks

Construction Strategy of Deep Learning Datasets

The dataset mainly consists of $K^{\pm}$ and $\pi^{\pm}$ mesons undergoing decay in flight.

- MC simulation: OSCAR v2.6.2
- The momentum range is set to $100 \sim 1150\mathrm{MeV}/c$ for K mesons and $50 \sim 1100\mathrm{MeV}/c$ for π mesons.
- The solid angle is within the detector acceptance ($|\cos\theta| < 0.93,\ 0 \le \phi < 2\pi$).

Due to the rarity of kink events, the dataset construction focuses on multi-track events containing tracks from a single kink and a normal track, mixed with beam background. The dataset construction is as follows:

Type	Event Number(has bkg)	Event Number(no bkg)
1 kink track + 1 normal track	100,000	100,000
2 kink tracks	50,000	50,000
2 normal tracks	50,000	50,000
1 kink track	50,000	50,000
1 normal track	50,000	50,000

GNN Construction Strategy

Graph Building from MDC Data

To handle unknown track counts and massive hit data, the network models MDC point clouds as a graph using Object Condensation and GravNet.

- **Nodes**: Individual sense wire hits.
- **Node Features**:
  - **Spatial**: Cartesian coordinates (x, y) at the $z = 0$ plane.
  - **Geometry**: Hierarchical Cell ID, Layer ID, and Super-layer ID.
  - **Electronics**: ADC (energy loss) & TDC (drift time).
- **Edges (*Dynamic*)**: No predefined physical edges. Connections are dynamically built in the latent space via GravNet iterations.

GNN Construction Strategy

Multi-Task Learning Heads

The network utilizes multi-task learning to share representations and improve clustering and kinematic predictions:

- **Object Condensation Head**: Outputs the condensation score β_i and latent coordinates.
- **Kinematics Head**: Regresses initial momentum, kink position, and classifies charge.
- **Hit Classification Head**: Categorizes hit types (parent, daughter, normal track, or noise).
- **Edge Prediction Head**: Predicts parent-to-daughter decay relations between hits.

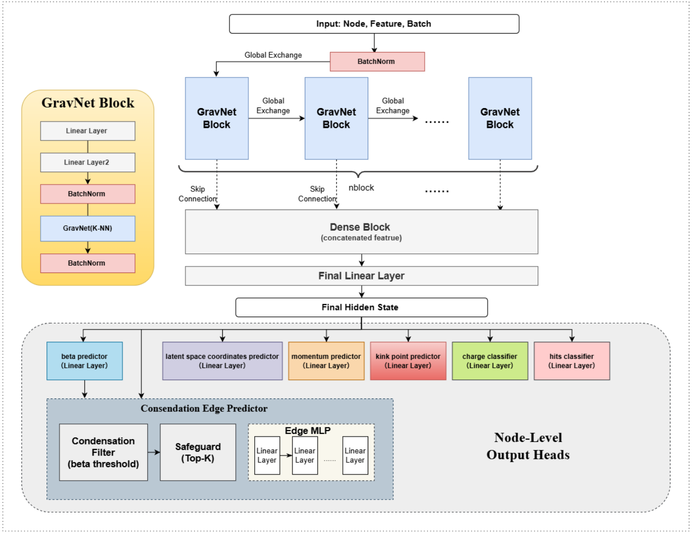
GNN Construction Strategy

Data Pipeline: Labeling & Processing

- **Data Labeling (Based on MC Truth)**:
  - **Hit Categories**: Parent, daughter, normal track, or background.
  - **Graph Node Labels**: 3D momentum, charge, and kink coordinates.
- **Data Quality Checking**:
  - Spatial hit distribution in the detector's XY plane.
  - MC simulated kinematic variable distributions.
  - Electronics signal distributions (ADC & TDC).
- **Data Preprocessing**:
  - **Clipping**: Restricting feature ranges and removing outliers.
  - **Scaling**: Normalizing input features for stable training.

Model Overview

- Object Condensation
- GravNet
- Multi-task learning



Model Overview

Object Condensation

To solve the issue of unknown track counts, Object Condensation [2] dynamically clusters hits using an attractive-repulsive potential field.

- **Network Outputs**: The GNN assigns each hit latent space coordinates and a confidence score $\beta \in [0, 1]$.
- **Node Charge (q):** Defines the condensation strength:

$$q_i = \operatorname{arctanh}^2(\beta_i) + q_{min} \tag{1}$$

  - $\beta_i \to 1$: High charge, strong condensation capability.
  - $\beta_i \to 0$: Low charge, mostly repelled (or background).
- **Clustering Mechanism**:
  - **Center Selection**: The node with the highest β acts as the condensation center α_k.
  - **Attraction**: Hits of the *same* track are pulled toward α_k.
  - **Repulsion**: Hits of *different* tracks are pushed away.

[2] http://arxiv.org/abs/2002.03605

Model Overview

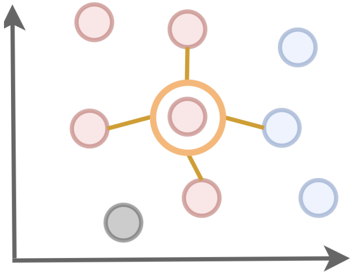
GravNet: Dynamic Graph Construction

- **Dynamic k-NN**: Recompute neighbors in latent space per iteration.
- **Gaussian Weights**: Distance controls information flow:

$$W_{ij} = \exp(-d_{ij}^2) \tag{2}$$

- **Message Aggregation**: Fuse neighbor features (f_j) and update:

$$\mathbf{m}_i = \sum_{j \in \mathcal{N}(i)} W_{ij} f_j \tag{3}$$



Model Overview

Loss Function Construction of Multi-task Learning

The basic loss function consists of:

- **Attractive Potential**: Pulls hits i of the same track toward their representative center α_k in the latent space.
- **Repulsive Potential**: Pushes nodes j of different tracks away from the center α_k.
- **β Penalty Term**: Forces $\beta \to 1$ for true track centers (α_k) and $\beta \to 0$ for background noise.
- **Multi-task Payload**: Weights auxiliary task losses by confidence β, allowing high-confidence nodes to drive the backpropagation.

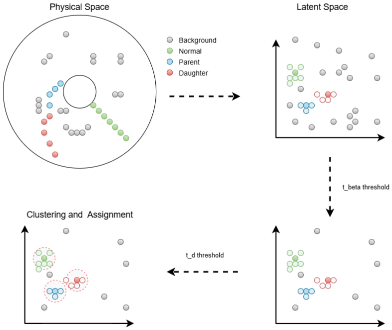
Model Training

Methodology & Hyperparameters

- **Dataset Partition**: Split into Training, Validation, and Test sets.
- **Optimization Target**: Joint optimization of the multi-task loss function (OC potentials + β penalty + payload tasks).
- **Training Strategies**:
  - **Optimizer**: AdamW with an initial learning rate of 0.001, paired with dynamic learning rate decay.
  - **Stage-wise Training**: Adjust the weights of different loss terms across training stages to balance clustering and payload tasks.
  - **Regularization**: Monitor validation loss and apply Early Stopping to prevent overfitting.
  - **Checkpointing**: Periodically save the best-performing models for final evaluation.

Model Post-processing

- **Latent Mapping**: Project hits from physical to clustering space.
- **Noise Filter (t_β)**: Reject background hits below the confidence threshold t_β.
- **NMS Clustering (t_d)**: Find condensation centers via Non-Maximum Suppression with distance threshold t_d; assign hits within radius $2t_d$.
- **Parameter Prediction**: Predict track properties directly from the condensation centers.



Model Post-processing

Track Clustering Metrics

Grid search for optimal t_β and t_d based on validation set performance.

Track Clustering Metrics:

- **Hit Efficiency (ϵ_{hits})**: Hit-count-weighted fraction of truth-track hits assigned to their matched reconstructed tracks
- **Hit Purity ($\mathcal{P}_{\text{hits}}$)**: Hit-count-weighted fraction of hits in matched reconstructed tracks that originate from their corresponding truth tracks
- **Object Recall (R_{obj})**: Fraction of truth track objects successfully reconstructed.
- **Fake Rate (R_{fake})**: Fraction of predicted tracks classified as fake or background-associated.
- **Track Quality Score (S_{track})**:

$$S_{\text{track}} = \frac{2\epsilon_{\text{hits}}\mathcal{P}_{\text{hits}}}{\epsilon_{\text{hits}} + \mathcal{P}_{\text{hits}}} \cdot R_{\text{obj}} \cdot (1 - R_{\text{fake}}). \tag{4}$$

Model Post-processing

Kink Separation Metrics

Kink Separation Metrics:

- **Separation Rate (A_{sep})**: Fraction of truth parent–daughter pairs reconstructed as two distinct clusters.
- **Parent Object Recall (R^{p}_{obj})**: Fraction of truth parent objects successfully reconstructed.
- **Daughter Object Recall (R^{d}_{obj})**: Fraction of truth daughter objects successfully reconstructed.
- **Parent–Daughter Recall Balance (H_{pd})**:

$$H_{pd} = \frac{2R^{p}_{obj}R^{d}_{obj}}{R^{p}_{obj} + R^{d}_{obj}}. \tag{5}$$

- **Kink Separation Score (S_{sep})**:

$$S_{sep} = A_{sep} \cdot H_{pd} \cdot (1 - R_{fake}). \tag{6}$$

Model Post-processing

Daughter Rescue Metrics

Daughter Recovery Metrics:

- **Daughter Hit Efficiency (ϵ_{d})**: Hit-count-weighted fraction of truth daughter hits recovered in the best-matched daughter cluster.
- **Daughter Hit Purity ($\mathcal{P}_{\mathrm{d}}$)**: Hit-count-weighted fraction of hits in the matched cluster that originate from the truth daughter.
- **Daughter Hit F1 (F_{d})**: Harmonic mean of daughter hit efficiency and purity:

$$F_{\mathrm{d}} = \frac{2\epsilon_{\mathrm{d}}\mathcal{P}_{\mathrm{d}}}{\epsilon_{\mathrm{d}} + \mathcal{P}_{\mathrm{d}}}. \tag{7}$$

- **Daughter Object Recall ($R^{\mathrm{d}}_{\mathrm{obj}}$)**: Fraction of unique truth daughter objects successfully reconstructed.
- **Separation Rate (A_{sep})**: Requires the reconstructed daughter to remain distinct from its parent.

Model Post-processing

Daughter Rescue Score

Daughter Contamination Penalties:

- **Fake Daughter Association ($R^{\mathrm{d}}_{\mathrm{fake}}$)**: Fraction of predicted daughter associations not matched to a truth daughter.
- **Background Confusion ($C^{\mathrm{d}}_{\mathrm{bg}}$)**: Rate at which daughter reconstruction is confused with background hits.
- **Other-object Confusion ($C^{\mathrm{d}}_{\mathrm{other}}$)**: Rate at which daughter hits are assigned to another truth object.

$$P^{\mathrm{d}}_{\mathrm{fake}} = 1 - 0.5R^{\mathrm{d}}_{\mathrm{fake}}, \tag{8}$$

$$C_{\mathrm{d}} = \max\left(0,\ 1 - 0.25C^{\mathrm{d}}_{\mathrm{bg}} - 0.25C^{\mathrm{d}}_{\mathrm{other}}\right). \tag{9}$$

Daughter Rescue Score ($S_{\mathrm{d,rescue}}$):

$$S_{\mathrm{d,rescue}} = F_{\mathrm{d}} \cdot R^{\mathrm{d}}_{\mathrm{obj}} \cdot A_{\mathrm{sep}} \cdot P^{\mathrm{d}}_{\mathrm{fake}} \cdot C_{\mathrm{d}}. \tag{10}$$

Model Post-processing

Low-fake Score

Low-fake Control Metrics:

- **Hit Purity ($\mathcal{P}_{\text{hits}}$)**: Controls the cleanliness of reconstructed tracks.
- **Object Recall (R_{obj})**: Prevents the low-fake working point from becoming excessively inefficient.
- **Fake Rate (R_{fake})**: Directly penalizes fake reconstructed tracks.
- **Low-fake Score (S_{lowfake})**:

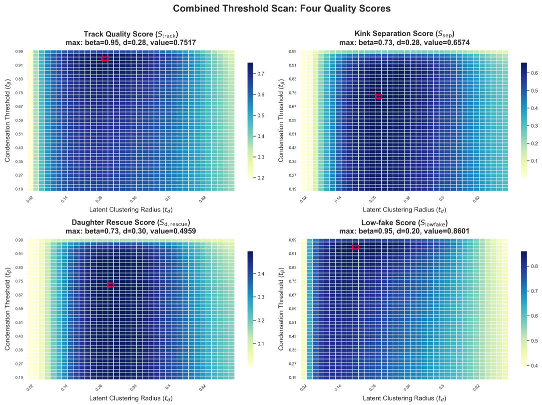
$$\widetilde{R}_{\text{obj}} = \max(R_{\text{obj}}, 0.05), \tag{11}$$

$$S_{\text{lowfake}} = (1 - R_{\text{fake}}) \cdot \frac{2\mathcal{P}_{\text{hits}}\widetilde{R}_{\text{obj}}}{\mathcal{P}_{\text{hits}} + \widetilde{R}_{\text{obj}}}. \tag{12}$$

Post-processing Parameter Scan Results

Four motivated working points are selected on the validation set. The scores have different definitions and should not be compared directly.

Working point	t_β	t_d	Maximum score
Track quality S_{track}	0.95	0.28	0.7517
Kink separation S_{sep}	0.73	0.28	0.6574
Daughter rescue $S_{\text{d,rescue}}$	0.73	0.30	0.4959
Low-fake control S_{lowfake}	0.95	0.20	0.8601



Model Evaluation

Test-Set Evaluation of Four Working Points

Four validation-selected working points are evaluated on the same held-out test set.

Per-metric cell shading: green = favorable, red = unfavorable.

Global Clustering Metrics

Working point	t_β	t_d	ϵ_{hits}	$\mathcal{P}_{\text{hits}}$	R_{obj}	R_{fake}
Track quality	0.95	0.28	**90.8%**	91.1%	88.3%	2.0%
Kink separation	0.73	0.28	86.4%	88.6%	93.0%	7.1%
Daughter rescue	0.73	0.30	87.2%	87.9%	**93.1%**	7.4%
Low-fake	0.95	0.20	87.6%	**94.1%**	86.2%	**1.6%**

Daughter Kink Metrics

Working point	$R^{\text{d}}_{\text{obj}}$	A_{sep}	$R^{\text{d}}_{\text{fake}}$
Track quality	77.0%	70.1%	11.2%
Kink separation	86.0%	79.3%	24.3%
Daughter rescue	**86.1%**	**79.5%**	24.8%
Low-fake	73.1%	66.1%	**10.1%**

Working-point choice. **Track quality** is the balanced default ($\epsilon_{\text{hits}} = 90.8\%$, $R_{\text{fake}} = 2.0\%$). **Daughter rescue** maximizes daughter recovery and separation ($R^{\text{d}}_{\text{obj}} = 86.1\%$, $A_{\text{sep}} = 79.5\%$). **Low-fake** is the purity-first option ($\mathcal{P}_{\text{hits}} = 94.1\%$, $R_{\text{fake}} = 1.6\%$).

Model Evaluation

Kinematic Regression: Component-wise Scan

Four working points are evaluated on the held-out test set; lower residuals indicate better regression accuracy.

Momentum ($\Delta p = p_{\mathrm{pred}} - p_{\mathrm{truth}}$)		Δp_x	Δp_y	Δp_z	Δp
Working point	**Metric**				
Track quality	MAE	37.8	36.2	55.3	88.9
	RMSE	78.8	78.1	103.8	151.9
Kink separation	MAE	46.6	45.3	62.3	105.6
	RMSE	98.5	98.2	119.7	183.5
Daughter rescue	MAE	46.7	45.4	62.5	105.8
	RMSE	98.8	98.6	119.9	184.0
Low-fake	MAE	**36.7**	**35.0**	**54.0**	**86.5**
	RMSE	**76.2**	**74.8**	**101.3**	**147.2**

Momentum residuals in MeV/c.

Kink Vertex ($\Delta r = r_{\mathrm{pred}} - r_{\mathrm{truth}}$)		Δr_x	Δr_y	Δr_z	Δr
Working point	**Metric**				
Track quality	MAE	**58.2**	**58.5**	**100.9**	**152.6**
	RMSE	**110.1**	**109.1**	**183.1**	**239.9**
Kink separation	MAE	74.2	74.2	119.2	187.1
	RMSE	137.0	134.7	221.5	293.2
Daughter rescue	MAE	73.6	73.6	118.5	185.8
	RMSE	136.2	134.0	220.0	291.4
Low-fake	MAE	59.0	59.5	101.3	154.0
	RMSE	111.0	110.5	184.3	241.8

Vertex residuals in mm.

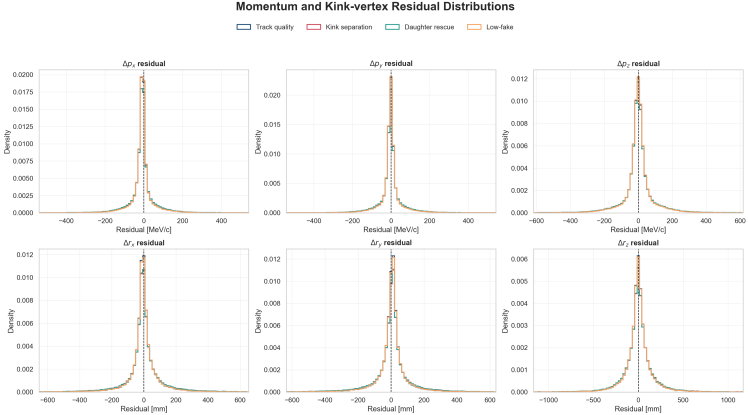
Metric guide. MAE = $\langle |e| \rangle$: typical absolute error; RMSE = $\sqrt{\langle e^2 \rangle}$: more sensitive to large residuals. RMSE $\gg$ MAE indicates broad tails.

Key result. **Low-fake** gives the best momentum regression; **Track quality** gives the best kink-vertex regression; the z component is consistently the least precise.

Model Evaluation

Residual Distributions

Held-out test set: $\Delta p_i = p_i^{\text{pred}} - p_i^{\text{truth}}, \qquad \Delta r_i = r_i^{\text{pred}} - r_i^{\text{truth}}.$

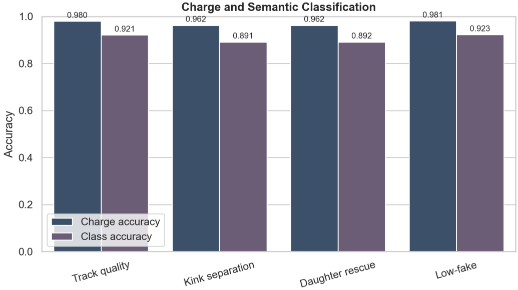


Takeaway. All residuals peak near zero and the four working points largely overlap in the core. The z components are broader, especially Δr_z; visible tails motivate reporting both MAE and RMSE.

Model Evaluation

Charge and Semantic Classification

Charge assignment is robust for all working points; semantic labels provide the more discriminating comparison.



Semantic accuracy is evaluated for the Normal, Parent, and Daughter labels.

Takeaway. Charge accuracy remains high for all points (96.2%–98.1%). Semantic accuracy distinguishes the working points more clearly: **Low-fake** (92.3%) and **Track quality** (92.1%) outperform the two $t_\beta = 0.73$ configurations (89.1%–89.2%).

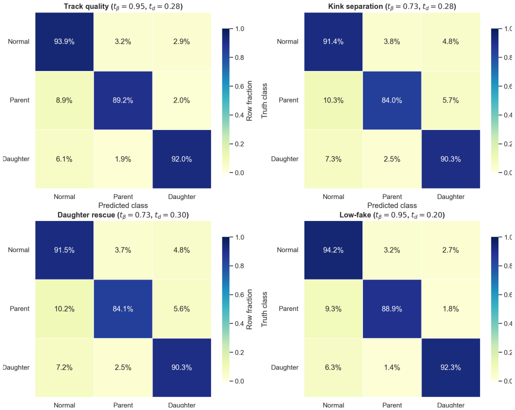
Model Evaluation

Semantic Classification: Confusion Patterns

Reading guide
Rows: truth labels.
Columns: predicted labels.
Diagonal cells: correct assignments.

Main confusion
Parent → Normal: 8.9%–10.3%.
Daughter → Normal: 6.1%–7.3%.
Parent–Daughter interchange remains small ($\lesssim$ 5.7%).

Best diagonal entries
Low-fake: Normal 94.2%, Daughter 92.3%.
Track quality: Parent 89.2%.



Implementation in the STCF Offline Software

Deployment Strategy

Goal: deploy the Python GNN in the STCF C++ offline framework with preserved reconstruction performance.

Route A — ONNX Runtime

Deploy: export an ONNX-compatible graph and run it with `Ort::Session`.

Issue: dynamic kNN and scatter are not native ONNX operators.

Adapt: fixed-size kNN + deterministic sorting + gather aggregation.

Pro: lightweight deployment. **Risk:** graph mismatch.

Route B — Native C++ / LibTorch

Deploy: rebuild the forward pass with `LibTorch` and load Python weights.

Implement: C++ kernels for dynamic kNN, message passing, and scatter reduction.

Match: reproduce Python neighbor construction and ordering.

Pro: highest GravNet fidelity. **Cost:** higher development effort.

Deployment rule. Start with ONNX. Accept only if Δ clustering/daughter metrics < 1 percentage point and Δ MAE/RMSE $< 3\%$. Otherwise move to the native C++ implementation.

Summary and Outlook

Key Results

- **Proof of concept**: A GNN with Object Condensation and GravNet jointly performs hit clustering, kink separation, kinematic regression, and semantic classification.
- **Balanced default**: At $(t_\beta, t_d) = (0.95, 0.28)$, the held-out test set reaches $\epsilon_{\text{hits}} = 90.8\%$, $\mathcal{P}_{\text{hits}} = 91.1\%$, $R_{\text{obj}} = 88.3\%$, and $R_{\text{fake}} = 2.0\%$.
- **Configurable operation**: The daughter-rescue point achieves $R^d_{\text{obj}} = 86.1\%$ and $A_{\text{sep}} = 79.5\%$, while the low-fake point gives 94.1% hit purity with a 1.6% fake rate.
- **Auxiliary predictions**: Charge accuracy remains 96.2–98.1%, with semantic classification accuracy up to 92.3%.

Summary and Outlook

Limitations and Next Steps

- **Vertex localization**: The balanced point has Δr MAE = 152.6 mm, with the z component least precise. Study stereo-hit and full 3D geometry information to strengthen longitudinal constraints.
- **Sparse daughter tracks**: Report efficiency versus daughter hit count, decay radius, and momentum; then introduce short-track reweighting and targeted data augmentation.
- **Physics validation**: Evaluate realistic full-event MC and real-data control samples, and compare with the existing STCF reconstruction at matched fake rate and runtime.
- **Software deployment**: Start with ONNX Runtime and require clustering/daughter metric shifts below 1 percentage point and regression shifts below 3% before C++ integration.

Thanks for your attention!

Questions & Answers

Backup

Momentum Distributions

- Momentum of parent and daughter tracks(e.g. $\pi^{\pm}$, μ^{-} and e^{-}) for K^{-} and π^{-} decay:

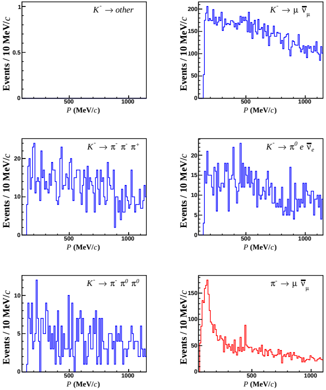


Figure: parent track.

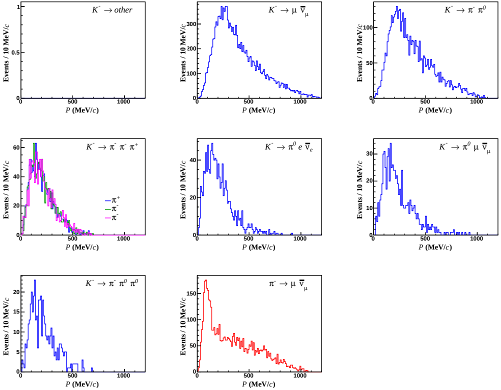
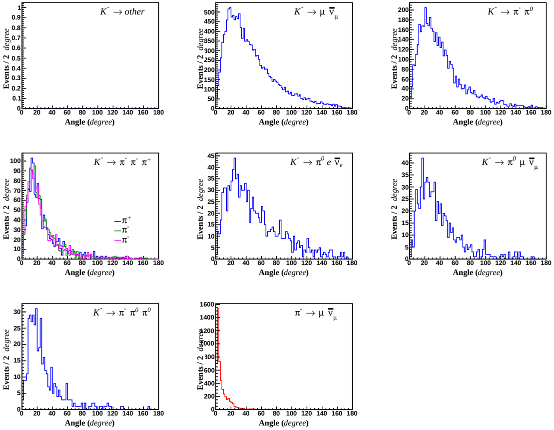


Figure: daughter track.

Backup

Kink angle distribution

- Kink angle distribution between mother and daughter tracks for K^- and π^- decay:



Backup

- Kink angle and momentum 2D distribution for K^- and π^- decay:

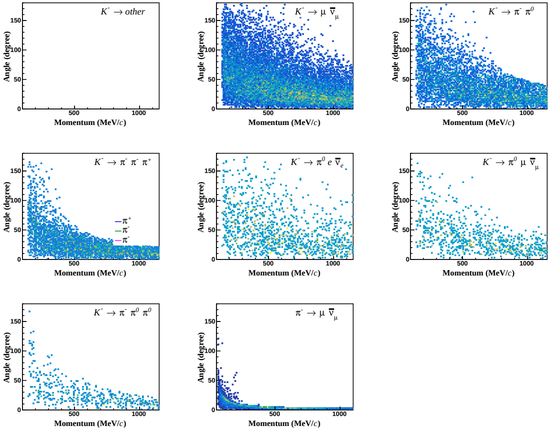


Figure: mother track.

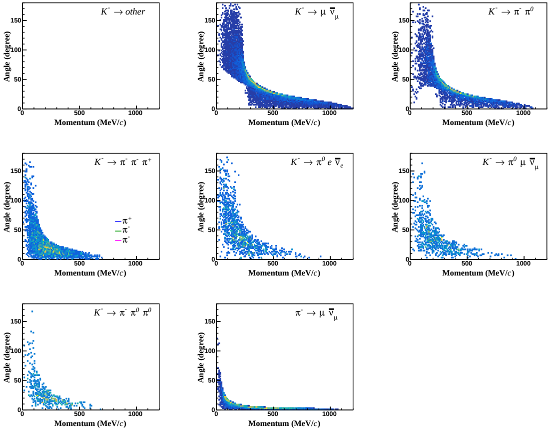
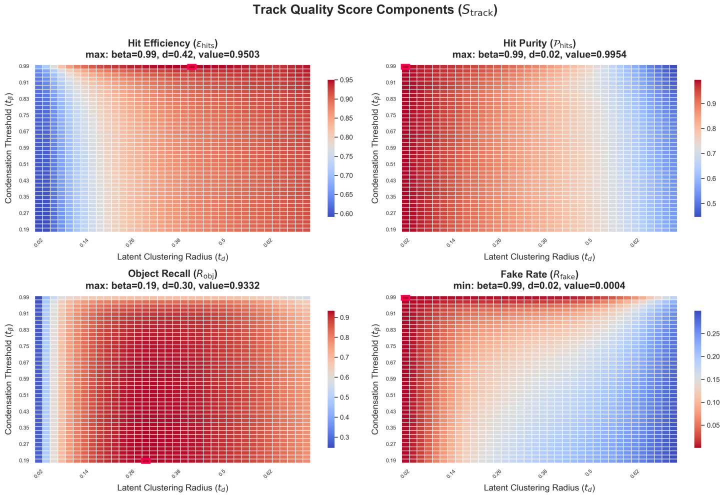


Figure: daughter track.

Backup

Track-quality score components: efficiency, purity, recall, and fake rate

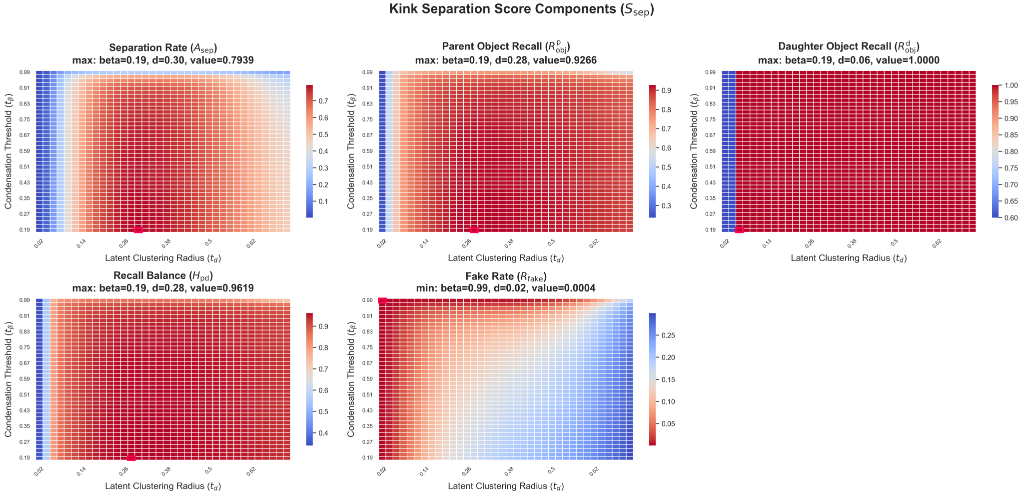
Component response across the (t_β, t_d) scan; magenta boxes mark the optimum of each component.



Backup

Kink-separation score components: parent–daughter separation and balance

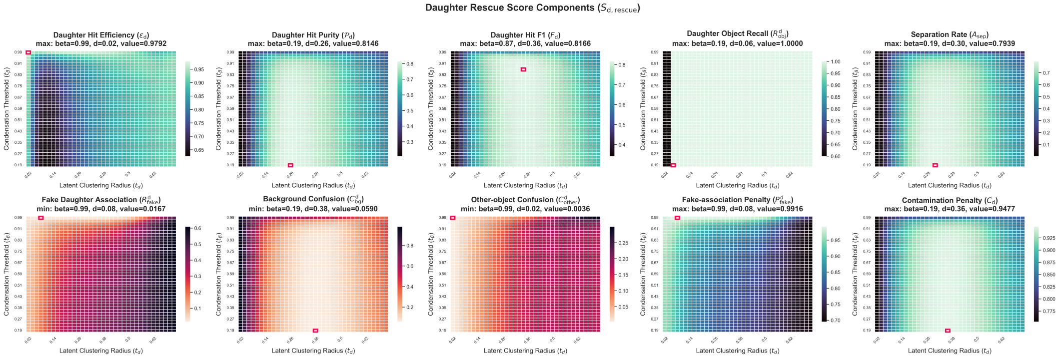
Separation rate, parent/daughter recall, recall balance, and fake-rate response over the parameter plane.



Backup

Daughter-rescue score components: recovery, contamination, and association

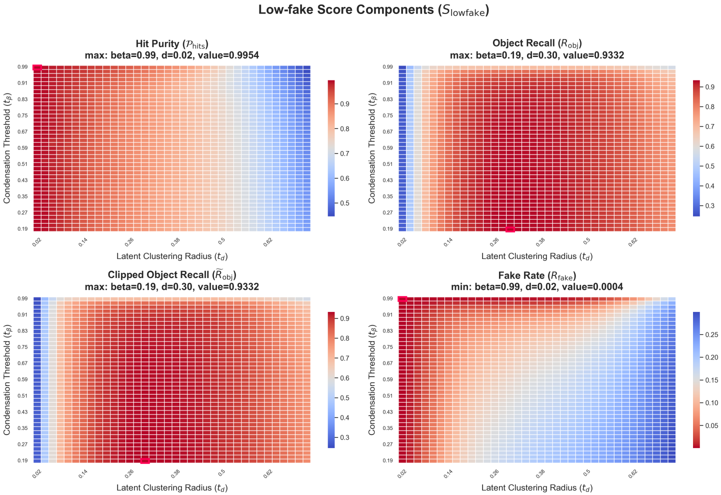
Daughter-track efficiency, purity, recall, and the penalties from background, other objects, and fake associations.



Backup

Low-fake score components: purity–recall trade-off and fake suppression

Hit purity, object recall, clipped recall, and fake-rate response across the (t_β, t_d) scan.

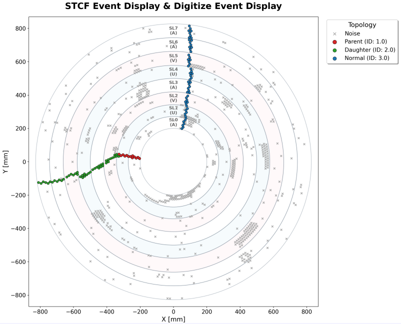


Backup

Dataset Topology: 1 Kink + 1 Normal Track

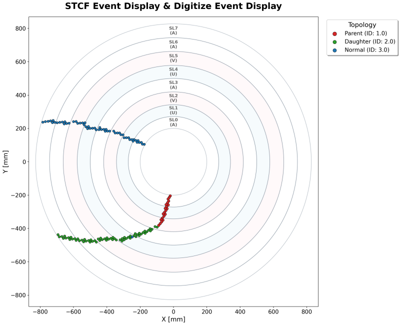
With Beam Background

C1: 1 Kink + 1 Normal + Bkg



Without Beam Background

C2: 1 Kink + 1 Normal



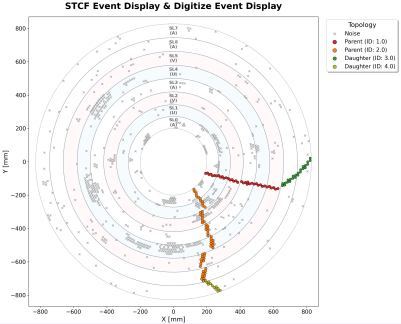
Use: mixed-object baseline for testing parent–daughter separation in the presence of a normal track.

Backup

Dataset Topology: Two Kink Tracks

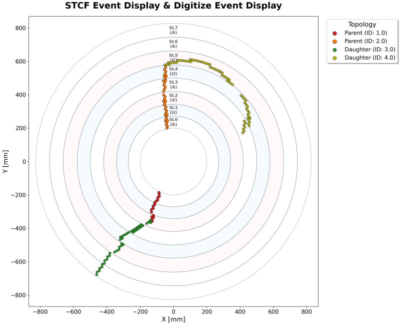
With Beam Background

C3: 2 Kinks + Bkg



Without Beam Background

C4: 2 Kinks



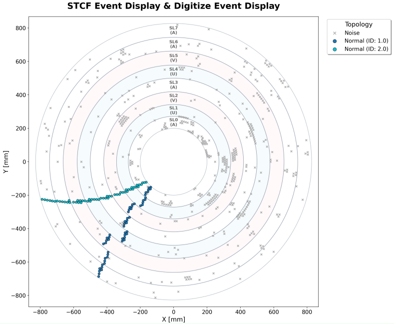
Use: tests competition between multiple parent–daughter structures and possible cluster overlap.

Backup

Dataset Topology: Two Normal Tracks

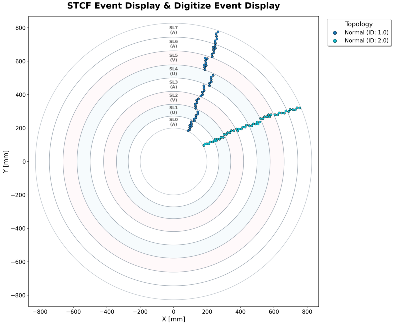
With Beam Background

C5: 2 Normals + Bkg



Without Beam Background

C6: 2 Normals



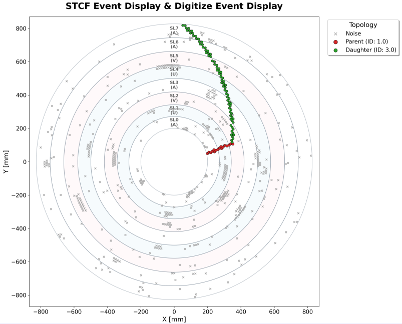
Use: normal-track control sample for diagnosing false kink reconstruction and background-induced fake clusters.

Backup

Dataset Topology: Single Kink Track

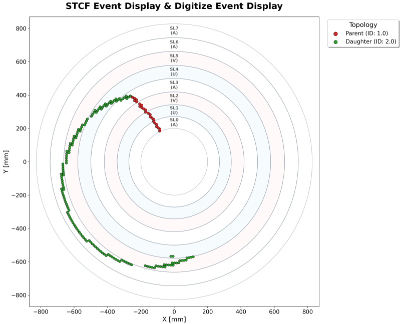
With Beam Background

C7: 1 Kink + Bkg



Without Beam Background

C8: 1 Kink



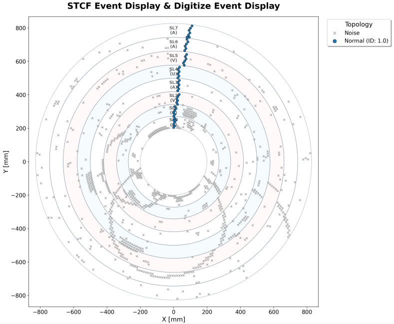
Use: isolates daughter reconstruction and parent–daughter separation without competition from other signal tracks.

Backup

Dataset Topology: Single Normal Track

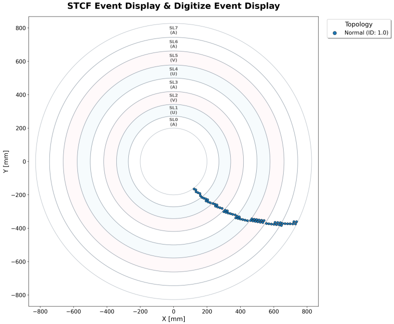
With Beam Background

C9: 1 Normal + Bkg



Without Beam Background

C10: 1 Normal



Use: pure control topology for evaluating whether beam background creates fake clusters or false kink-like structures.

Backup

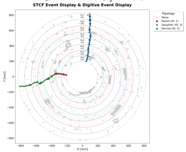
Track Quality Working Point: Predicted Clustering Examples

Prediction examples for the Track quality working point:

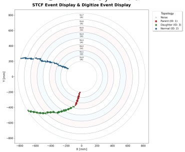
$$t_\beta = 0.95, \qquad t_d = 0.28.$$

The event IDs are identical to the truth gallery; only the display is changed from `Truth` to `Pred`.

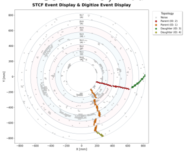
C1: Evt78, Pred



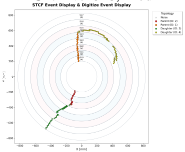
C2: Evt6, Pred



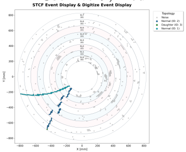
C3: Evt23, Pred



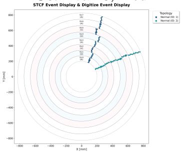
C4: Evt4648, Pred



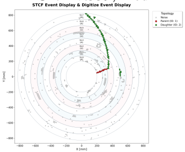
C5: Evt29, Pred



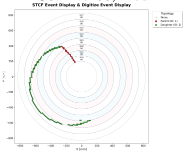
C6: Evt228, Pred



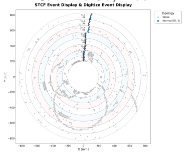
C7: Evt339, Pred



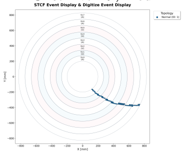
C8: Evt406, Pred



C9: Evt237, Pred



C10: Evt119, Pred

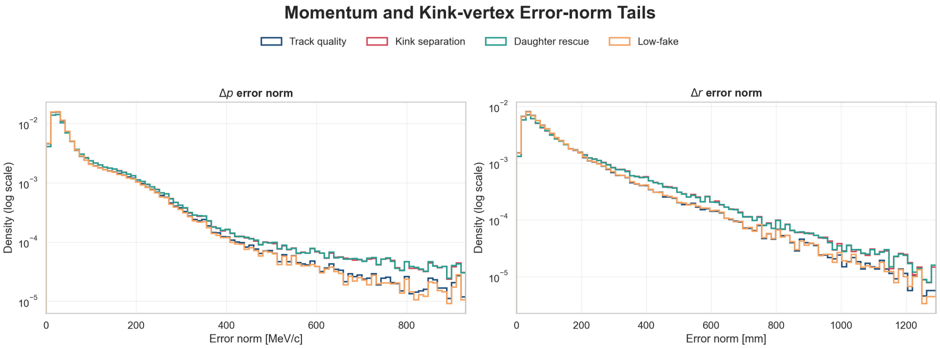


Backup

Kinematic Regression: Error-Norm Tails

Log-scale distributions of the three-dimensional residual norms:

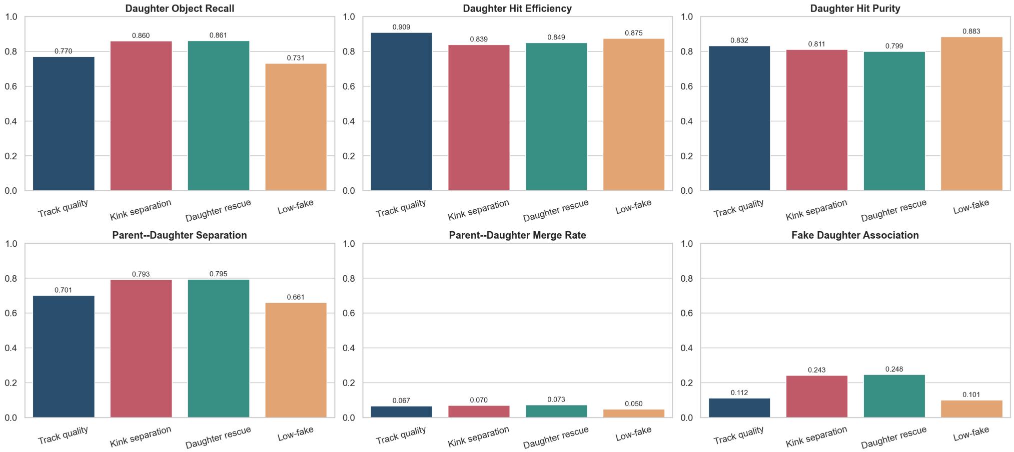
$$\Delta p = \left\|\vec{p}_{\text{pred}} - \vec{p}_{\text{truth}}\right\|, \qquad \Delta r = \left\|\vec{r}_{\text{pred}} - \vec{r}_{\text{truth}}\right\|.$$



Tail behaviour. The logarithmic scale makes tail differences visible. Kink-oriented working points retain larger momentum and vertex tails, especially for Δr, whereas **Low-fake** and **Track quality** suppress the large-error region. This is why RMSE changes more strongly than MAE.

Backup

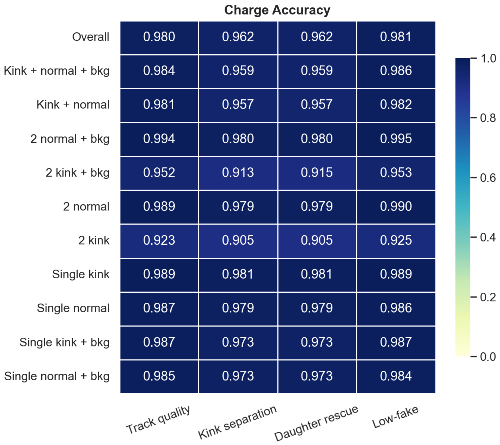
Daughter Reconstruction and Association Diagnostics

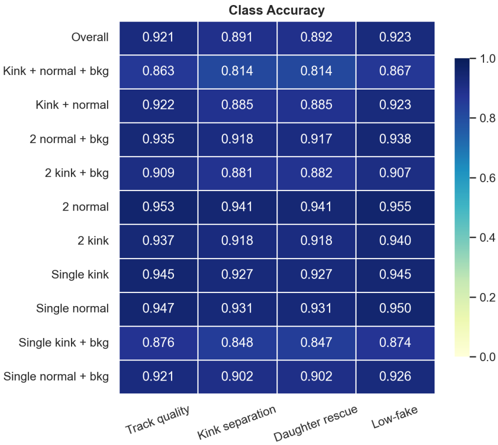


detailed comparison of daughter recall, separation, merge rate, and fake daughter association.

Backup

Classification Accuracy by Event Topology

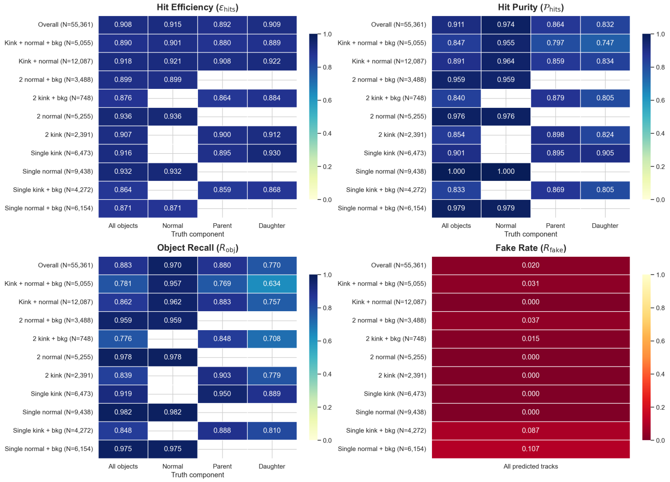




Left: charge accuracy. Right: Normal/Parent/Daughter semantic classification accuracy.

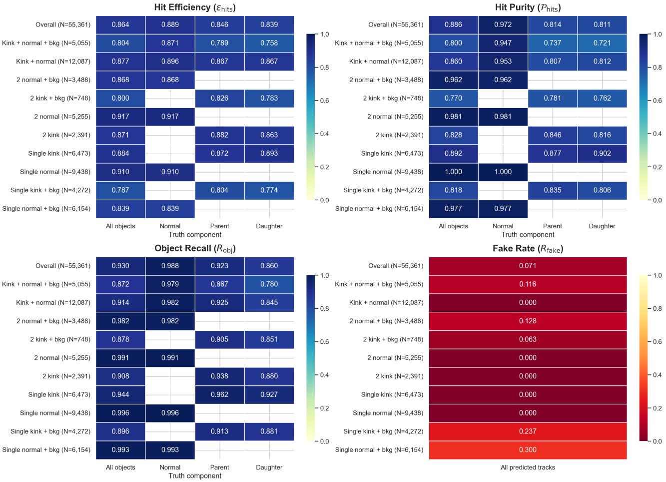
Backup

Track Quality by Topology: Track Quality Working Point



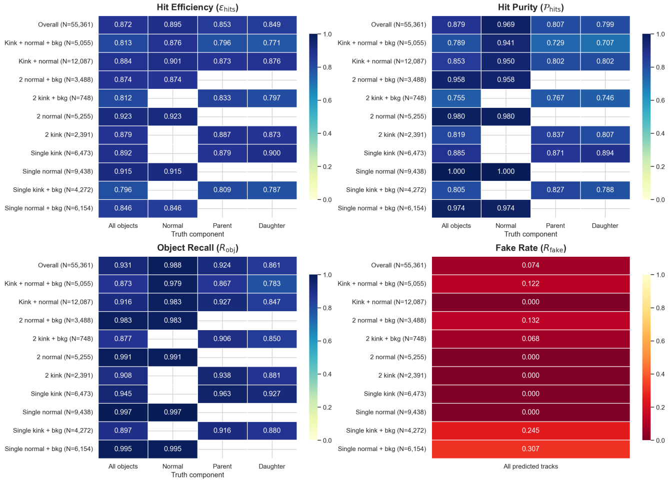
Backup

Track Quality by Topology: Kink-Separation Working Point



Backup

Track Quality by Topology: Daughter-Rescue Working Point



Backup

Track Quality by Topology: Low-Fake Working Point

Hit Efficiency (ε_{hits})

Truth component	All objects	Normal	Parent	Daughter
Overall (N=55,361)	0.876	0.884	0.862	0.875
Kink + normal + bkg (N=5,055)	0.861	0.869	0.857	0.859
Kink + normal (N=12,087)	0.890	0.894	0.881	0.893
2 normal + bkg (N=3,488)	0.869	0.869		
2 kink + bkg (N=748)	0.842		0.831	0.849
2 normal (N=5,255)	0.908	0.908		
2 kink (N=2,391)	0.871		0.865	0.876
Single kink (N=6,473)	0.879		0.858	0.893
Single normal (N=9,438)	0.896	0.896		
Single kink + bkg (N=4,272)	0.822		0.827	0.818
Single normal + bkg (N=6,154)	0.836	0.836		

Hit Purity ($\mathcal{P}_{hits}$)

Truth component	All objects	Normal	Parent	Daughter
Overall (N=55,361)	0.941	0.985	0.902	0.883
Kink + normal + bkg (N=5,055)	0.899	0.975	0.854	0.816
Kink + normal (N=12,087)	0.928	0.979	0.903	0.886
2 normal + bkg (N=3,488)	0.976	0.976		
2 kink + bkg (N=748)	0.896		0.918	0.873
2 normal (N=5,255)	0.987	0.987		
2 kink (N=2,391)	0.892		0.924	0.867
Single kink (N=6,473)	0.929		0.918	0.936
Single normal (N=9,438)	1.000	1.000		
Single kink + bkg (N=4,272)	0.882		0.903	0.863
Single normal + bkg (N=6,154)	0.989	0.989		

Object Recall (R_{obj})

Truth component	All objects	Normal	Parent	Daughter
Overall (N=55,361)	0.862	0.957	0.867	0.731
Kink + normal + bkg (N=5,055)	0.750	0.943	0.751	0.575
Kink + normal (N=12,087)	0.843	0.953	0.870	0.722
2 normal + bkg (N=3,488)	0.943	0.943		
2 kink + bkg (N=748)	0.742		0.836	0.653
2 normal (N=5,255)	0.970	0.970		
2 kink (N=2,391)	0.819		0.895	0.748
Single kink (N=6,473)	0.896		0.934	0.860
Single normal (N=9,438)	0.968	0.968		
Single kink + bkg (N=4,272)	0.821		0.876	0.768
Single normal + bkg (N=6,154)	0.955	0.955		

Fake Rate (R_{fake})

	All predicted tracks
Overall (N=55,361)	0.016
Kink + normal + bkg (N=5,055)	0.023
Kink + normal (N=12,087)	0.000
2 normal + bkg (N=3,488)	0.030
2 kink + bkg (N=748)	0.012
2 normal (N=5,255)	0.000
2 kink (N=2,391)	0.000
Single kink (N=6,473)	0.000
Single normal (N=9,438)	0.000
Single kink + bkg (N=4,272)	0.069
Single normal + bkg (N=6,154)	0.091