Kink Finder within the Belle II Tracking Software

Denis Bodrov
(Soochow University, HSE University)
on behalf of the Belle II Tracking Group

The 2024 International Workshop on Future Tau Charm Facilities

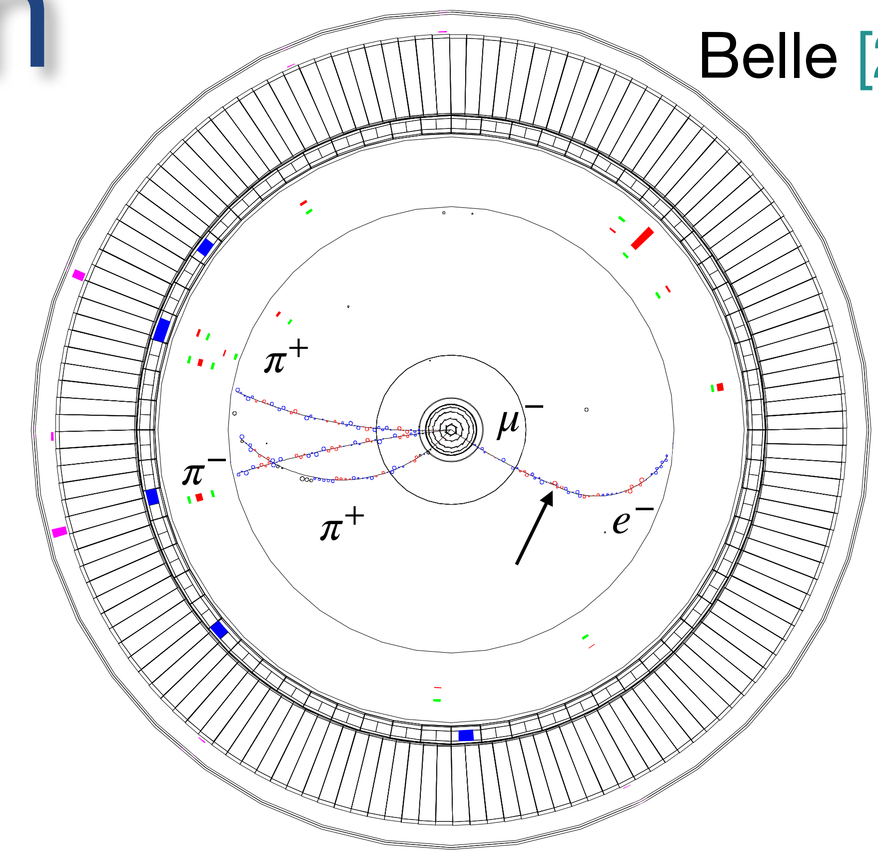
Motivation

- Kinks are tracks with breakpoints inside the experiment tracker
- They can be caused by charged particles decay in flight or scattering in the detector material

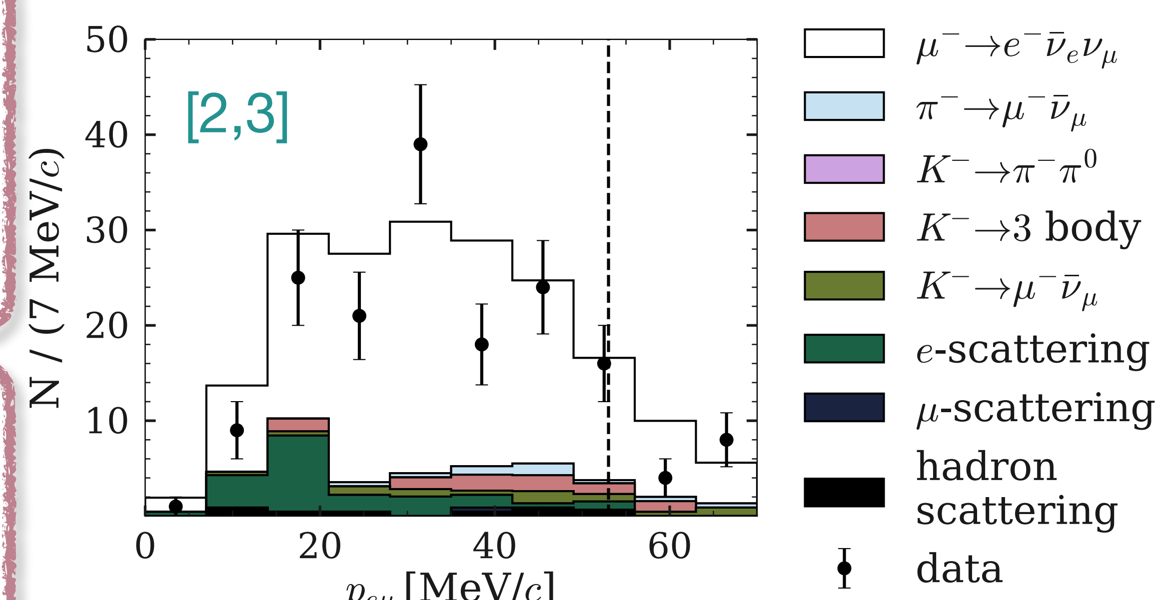
- Their reconstruction is one of the steps to improve the performance:
  - It can reduce the fake rate in the particle identification (PID), e.g., $K^- \to \pi^-\pi^0$ is a natural cause of initial kaons to be identified as pions
  - $K^- \to \pi^-\pi^0$, $K^- \to \mu^-\bar{\nu}_\mu$, and $\pi^- \to \mu^-\bar{\nu}_\mu$ can be used for dE/dx-calibration
  - Clone tracks can imitate the kink (their reconstruction can reduce their rate)

- Kinks can be used in physics studies:
  - E.g., $\mu^- \to e^-\bar{\nu}_e\nu_\mu$ reconstruction provides information about muon polarization [1]

Belle [2,3]



[1] JHEP 10 (2022) 035
[2] Phys.Rev.Lett. 131 (2023) 021801
[3] Phys.Rev.D 108 (2023) 012003

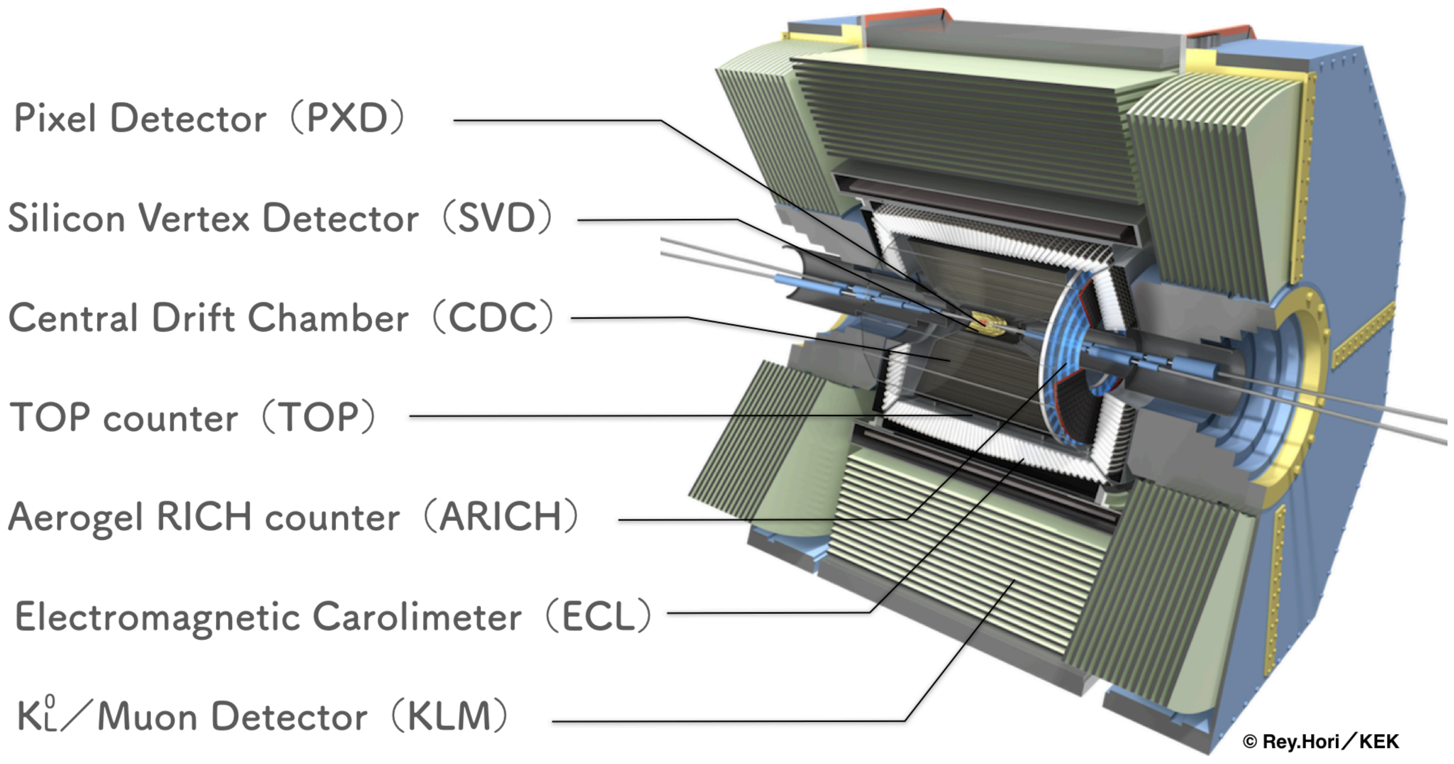


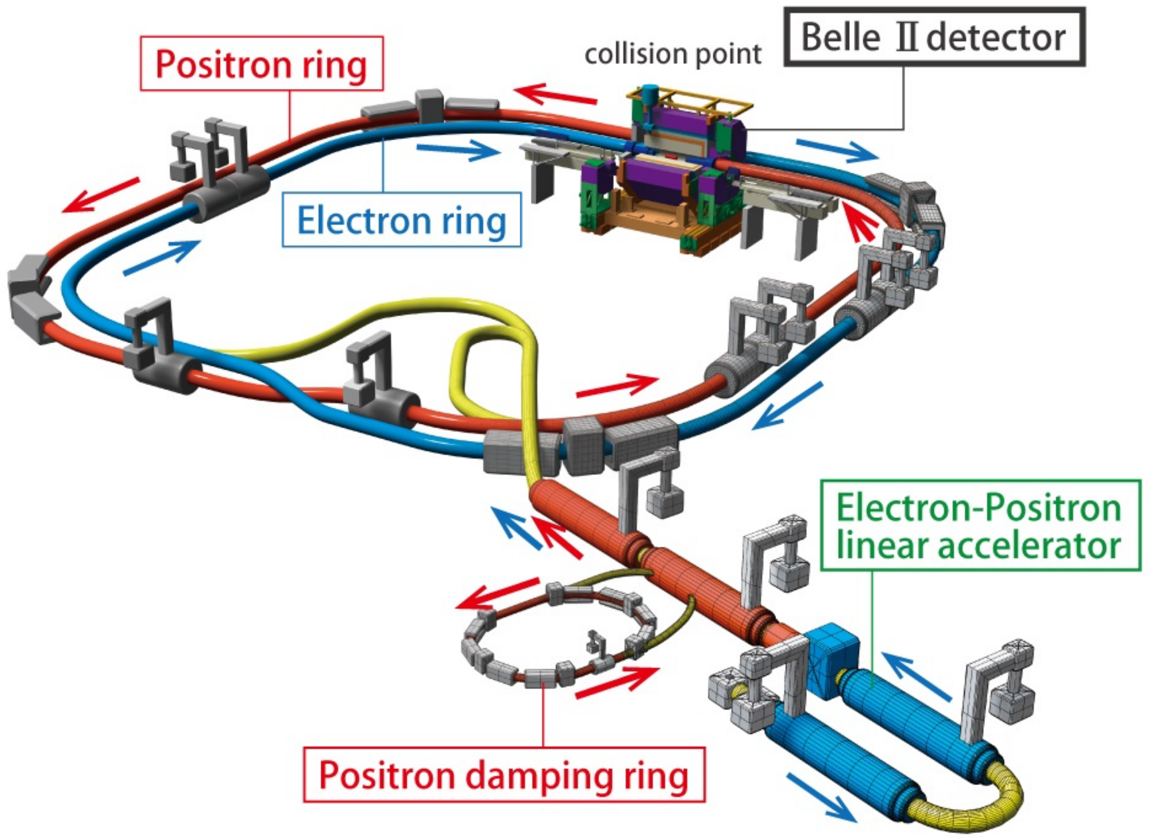
Belle II @SuperKEKB

- The Belle II experiment is located at the asymmetric energy e^+e^- collider SuperKEKB (Tsukuba, Japan)
- Belle II expects integrated luminosity of $\mathscr{L} = 50\,\text{ab}^{-1}$ by 2035 ($\sim 530\,\text{fb}^{-1}$ collected)
- More details in Monday talk by Alexander Kuzmin

$E(e^-) = 7$ GeV, $E(e^+) = 4$ GeV
Target luminosity: 6×10^{35} cm^{-2}s^{-1}
Achieved luminosity: 4.7×10^{34} cm^{-2}s^{-1}

PTEP 2019 (2019) 12, 123C01





Tracking detectors

Pixel Detector (PXD)

- Silicon pixel detector
- 2 layers
- Radii: 14 and 22 mm

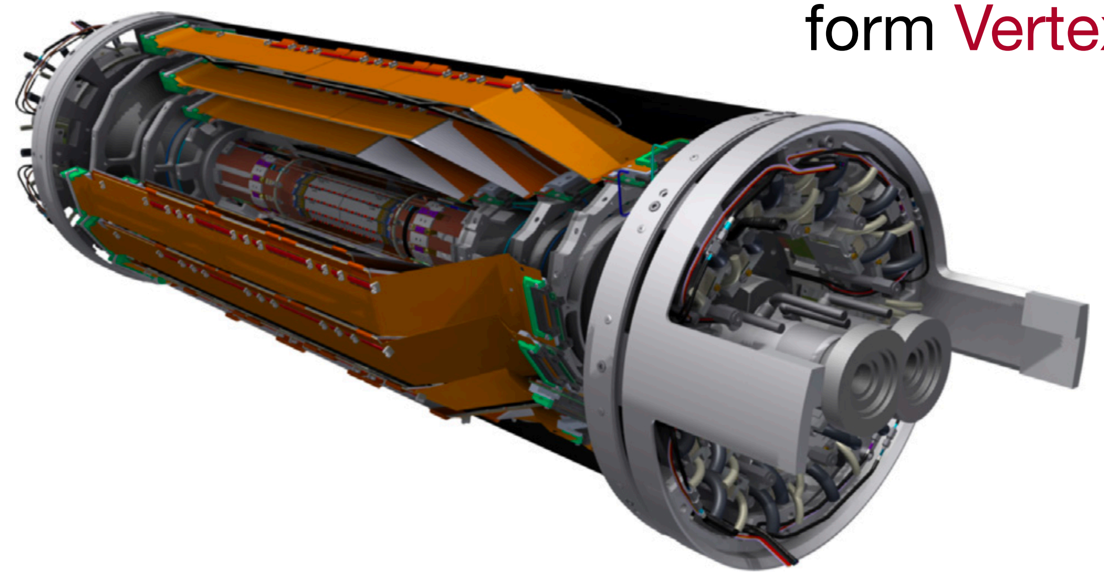
Silicon Vertex Detector (SVD)

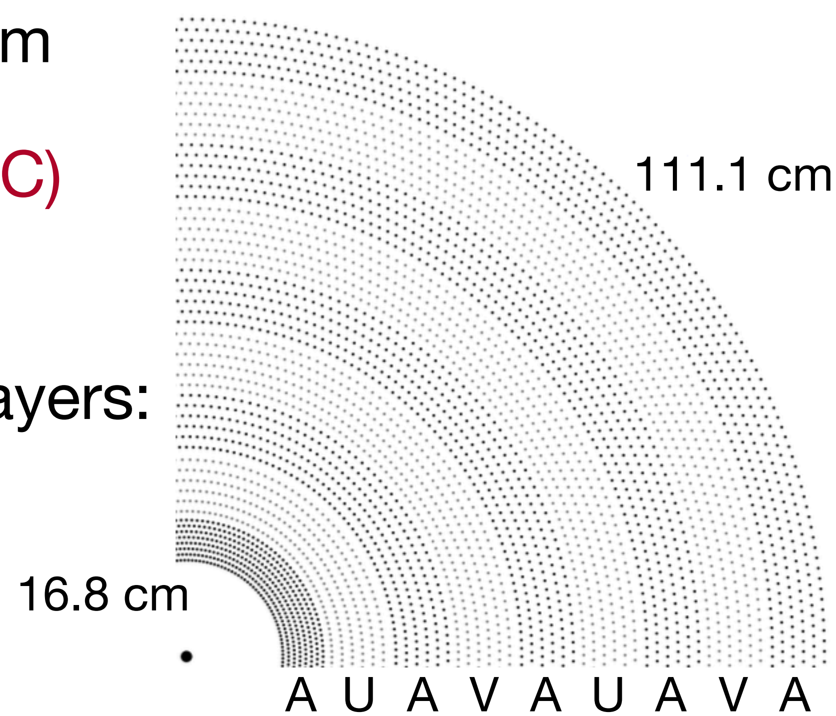
- Double-sided silicon strip detector
- 4 layers
- Radii: from 39 to 135 mm

Central Drift Chamber (CDC)

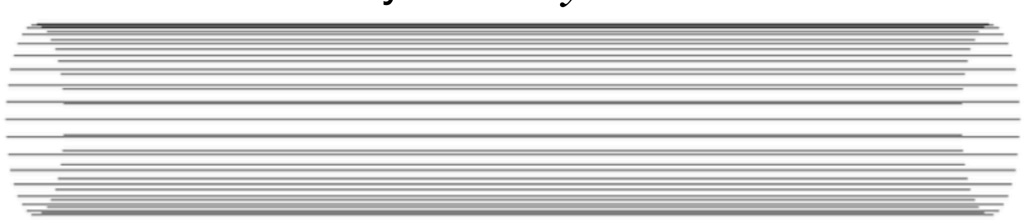
- ≈ 14000 sence wires
- 56 layers form 9 superlayers:
  - Axial (A)
  - Stereo (U, V)

PXD and SVD together form Vertex Detector (VXD)

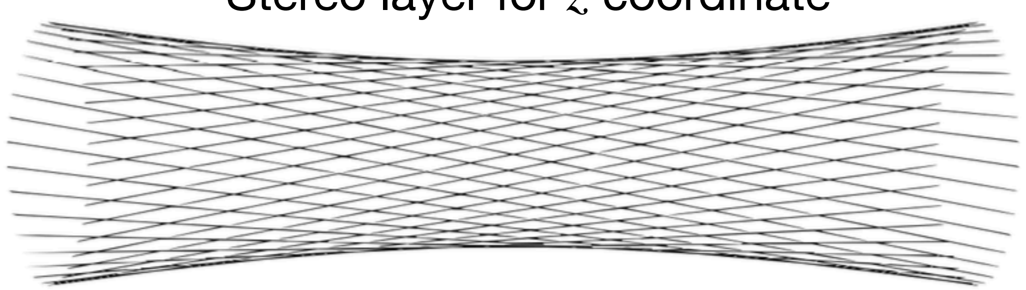




Axial layer for xy coordinate



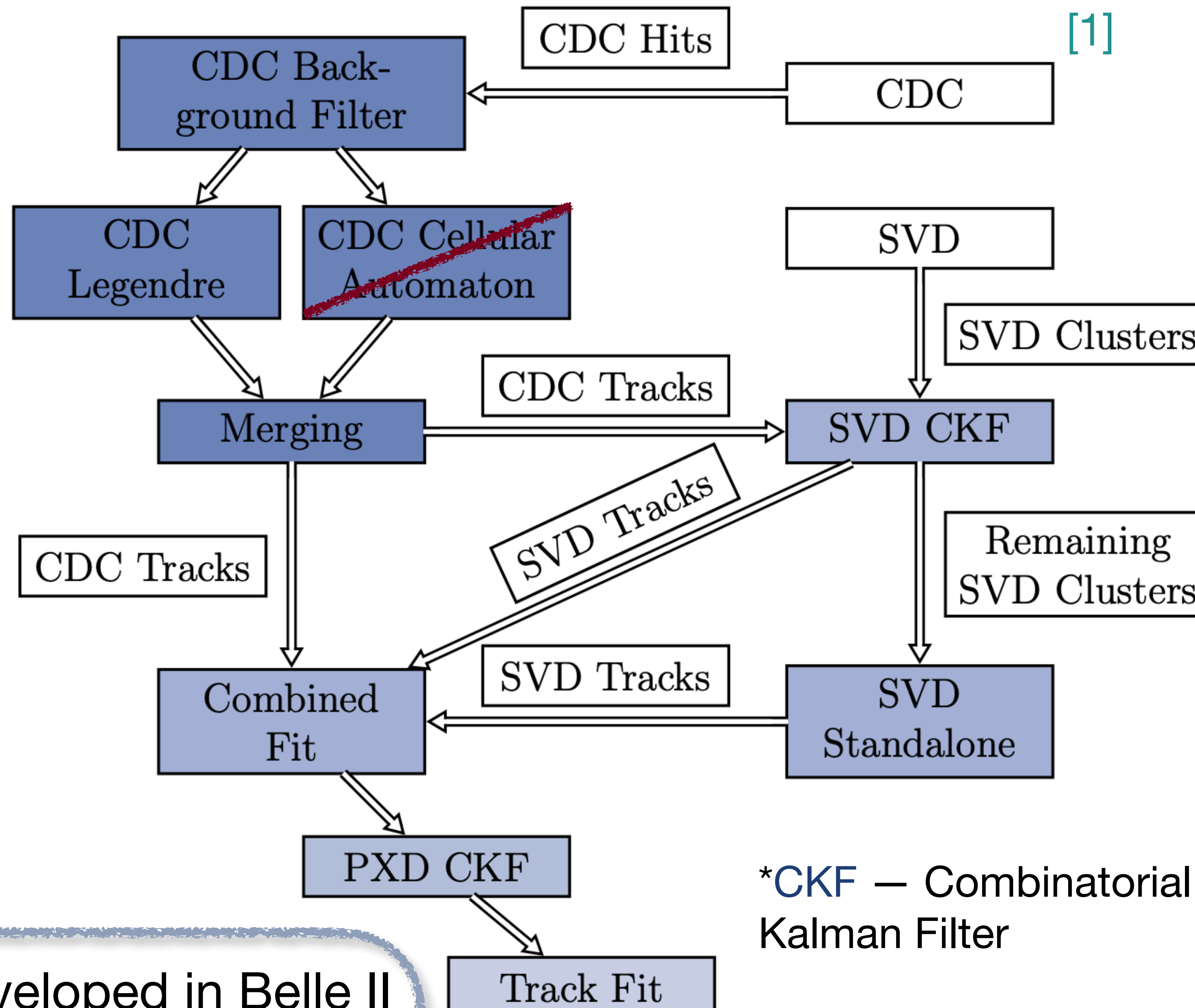
Stereo layer for z coordinate



Track Finding @Belle II

- Local track finder based on Cellular Automaton
  - Has high efficiency for displaced tracks
  - Currently is partially switched off due to high fake rate

[1] Comput. Phys. Commun. 259 (2021) 107610
[2] arXiv:1902.04405 [physics.data-an]



*CKF — Combinatorial Kalman Filter

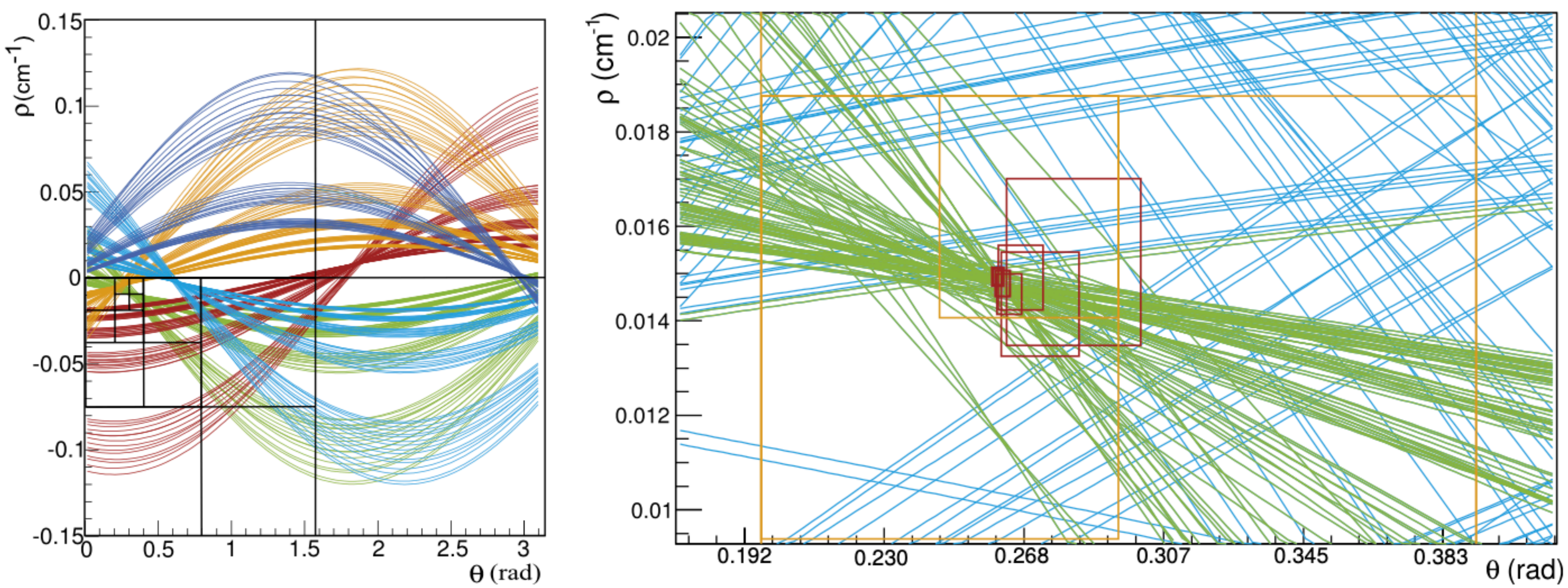
- To measure track Helix parameters, Deterministic Annealing Filter (DAF) from the genfit2 package is used [2]
- It uses iterative Kalman Filter with additional weights calculation to remove outliers
- Three hypotheses are used for fit: pion, kaon, and proton

New GNN Track Finding algorithm is developed in Belle II (incredible performance for displaced tracks)

See a talk by Lea Reuter at Hemholtz AI Conference 2024

Global CDC Track Finding

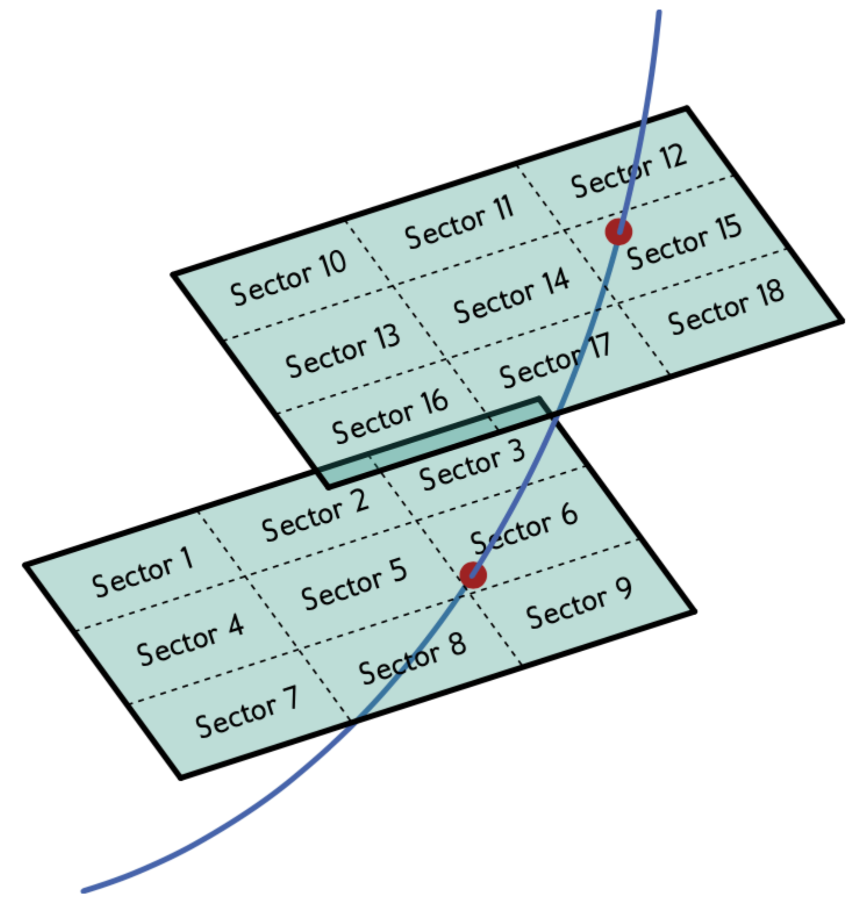
- Conformal transformation: circles crossing IP to straight lines
- Legendre transformation: to find common tangents to the drift circles
- Good for tracks from IP
    1. In $r\phi$ using axial layers only tangent to the drift circle:
       $\rho = x_0 \cos\theta + y_0 \sin\theta \pm R_{dr}$
    2. Extrapolate in 3D by attaching hits in stereo layers, in $z_0 \tan\lambda$:
       $z_0 = z_{\text{rec}} - s_{\text{rec}} \tan\lambda$
- Hit candidates shared by several tracks are not added to any track (on average $19\,\%$ of hits)

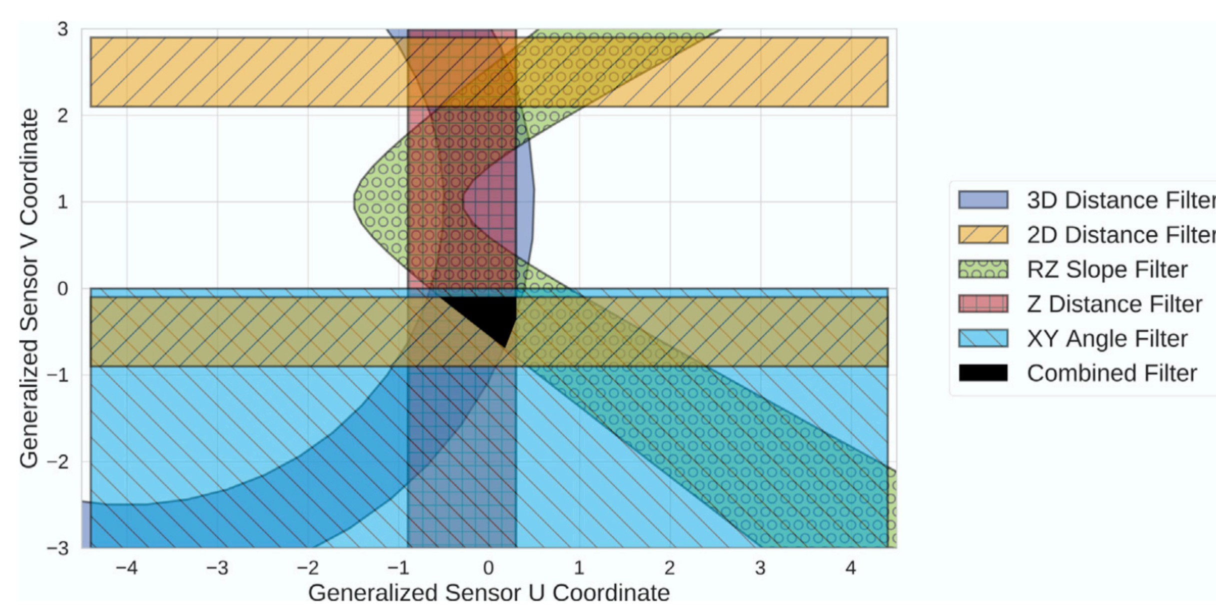


(a) Two-dim. binary search (b) Sliding bins example

SVD Standalone

- Found CDC tracks are extrapolated to SVD to attach hits
- To the remaining hits, SVD Standalone reconstruction is applied:
  - Each SVD sensor is divided in 4×4 sectors
  - Sector map trained on simulated tracks to learn geometrical relations between sectors
  - Filters reject bad space point combinations (angle, timing, distance)
  - Cellular automaton yields a set of paths
- New found tracks are extrapolated to CDC





Kinks at Belle II

- Based on MC simulation for $K^- \to \mu^- \bar{\nu}_\mu$ from $\tau^- \to K^- \nu_\tau$
- $10\,\%$ of decays-in-flight inside CDC cannot be reconstructed in principle
- Mother track:
  - $84\,\%$ of mother tracks reconstructed
- Daughter track:
  - $31\,\%$ of daughter tracks reconstructed
- Reconstructed as a combined track from daughter and mother hits:
  - $17\,\%$ as a mother (more than $66\,\%$ hits from mother)
  - $2\,\%$ as a daughter (more than $66\,\%$ hits from daughter)
  - $13\,\%$ with $\approx 50/50\,\%$ ratio

- More general case for decays-in-flight from $\tau^+\tau^-$, $B\bar{B}$, and $c\bar{c}$ events
- Mother track reconstruction efficiencies:
  - Above $80\,\%$ for kaon and muon decays
  - Around $60\,\%$ for pion decays (hits are assigned to daughter track more often)
  - Around $25\,\%$ of mother tracks are reconstructed by VXD only (great improvement compared to Belle)
- Daughter track reconstruction efficiencies:
  - Around $31\,\%$ for kaon decays
  - Around $23\,\%$ for pion decays
  - Around $15\,\%$ for muon decays

Kink Finding

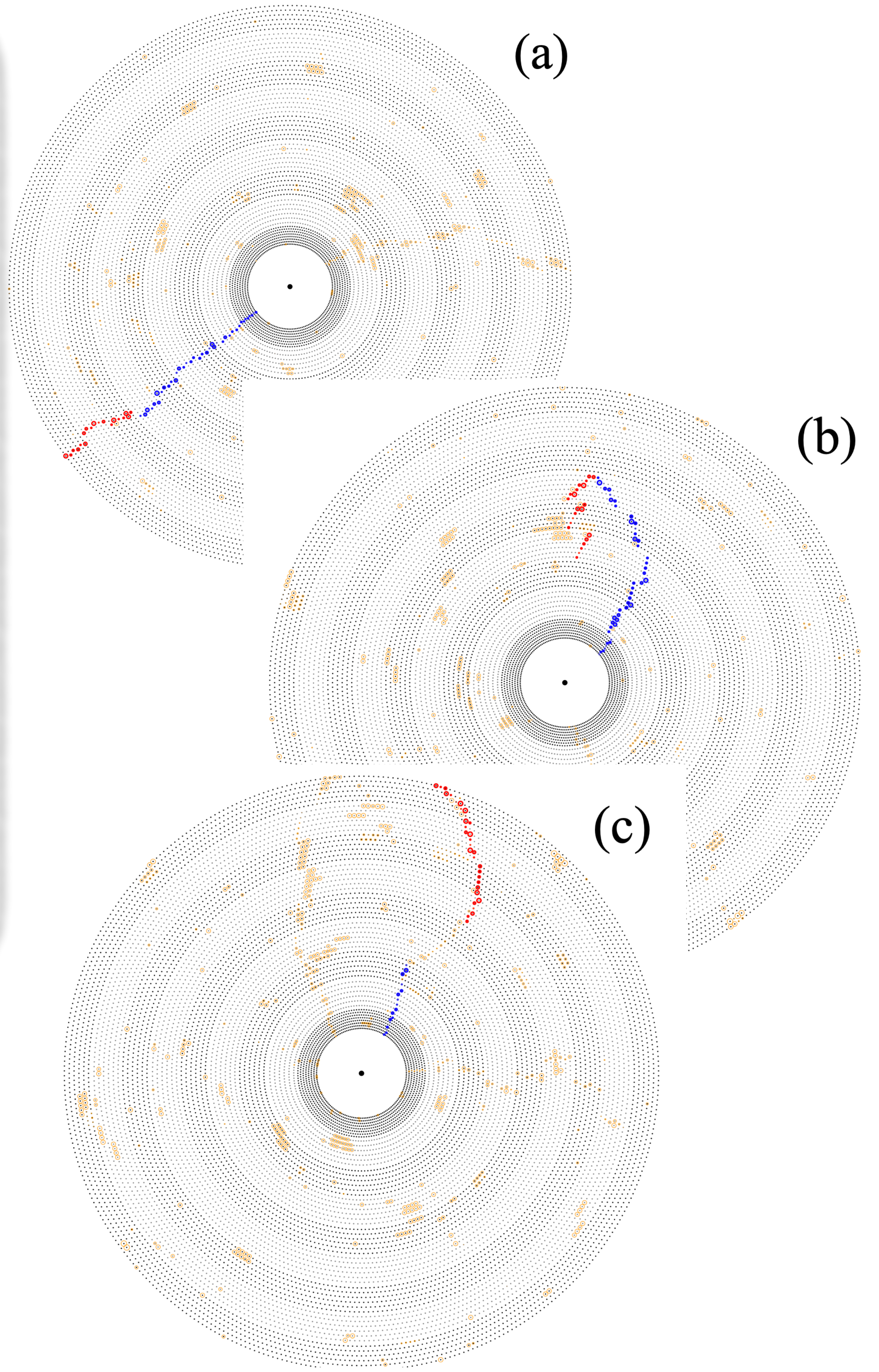
- In Belle II, currently we consider two types of Kinks:
  - Both daughter and mother tracks are individually reconstructed by the default track finding (depending on the decayed particle, it is around $10-20\,\%$ of all decays in flight)
  - Hits from daughter and mother track are combined in one track reconstructed by the default track finding (depending on the decayed particle, it is around $40-70\,\%$ of all decays in flight)

- The last case is the most crucial for the PID fake rate

- Kink Finder is a newly developed algorithm:
  - Shown numbers and overall performance are preliminary and not official
  - It has not been used for the data processing yet

Combination of track pairs

- The default track finding is not optimized for kinks, so we have to consider 6 cases:
  - The daughter's FIRST hit is close to the mother's last hit in 3D (a)
  - The daughter's LAST hit is close to the mother's last hit in 3D (wrong charge reconstruction for daughters coming back to IP as the default track finding is optimized for the tracks from IP) (b)
  - The daughter track helix extrapolation passes close to the mother's last hit in 3D (track finding can miss some superlayers) (c)
  - The other 3 cases are similar, but the z coordinate has poor resolution so check in 2D
- Depending on the case, the processing is different
- In 2D cases, we usually can improve the z coordinate resolution by refitting the daughter with better seeds estimated from the mother last hit information

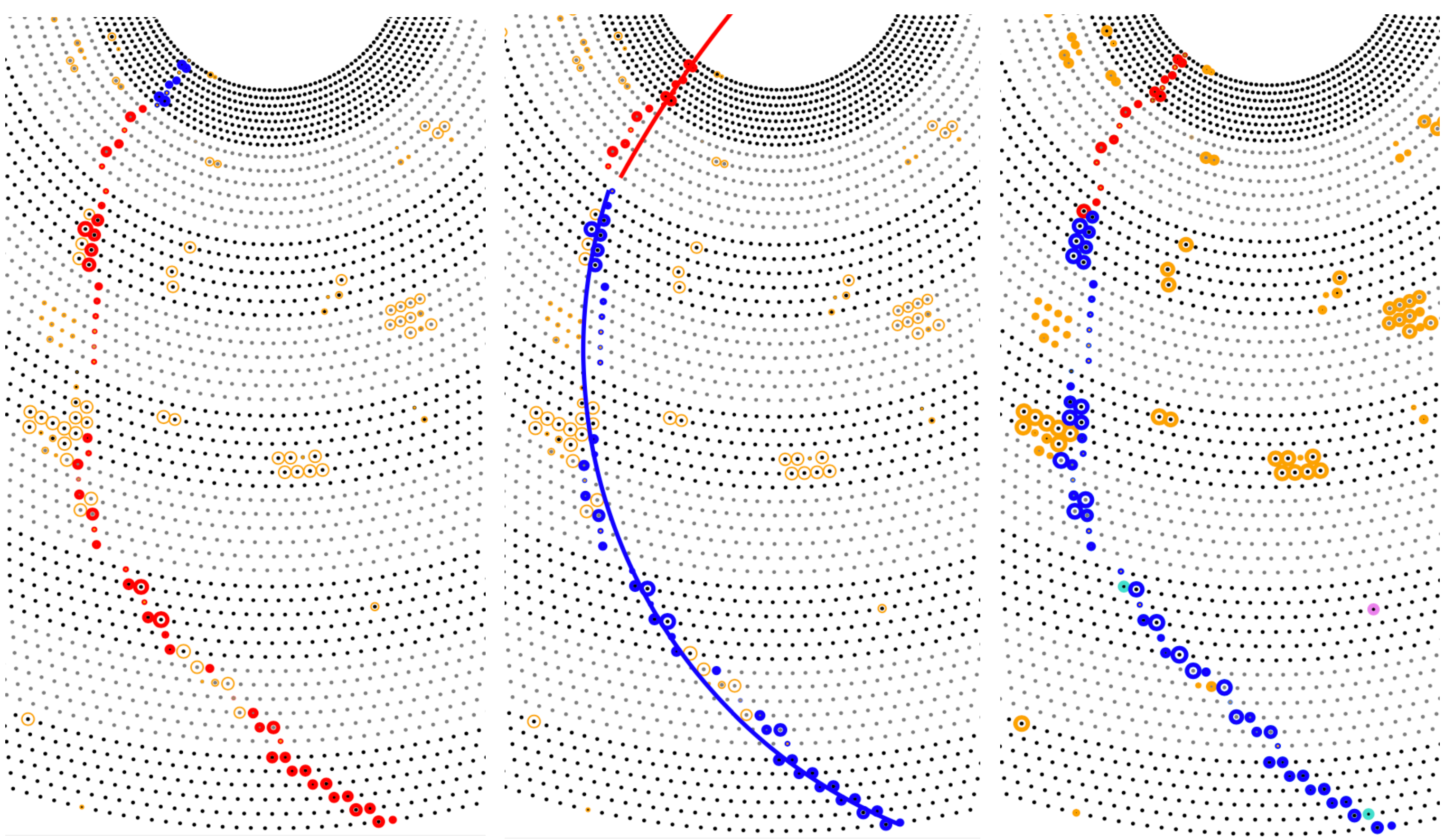


Hit reassignment

- The track finding can assign hits from one of the kink particles to another, leading to a worse resolution of decay vertex and track parameters
- To solve this problem, a hit reassignment is performed between tracks after the first vertex fit

- If the fitted vertex is inside of one of the tracks:
  - Reassign the extra hits to the complementary track
  - Refit both tracks and check if the following value is improved
    $$\frac{\chi^2_{\mathrm{comb}}}{\mathrm{n.d.f.}_{\mathrm{comb}}} = \frac{\chi^2_{\mathrm{m}} + \chi^2_{\mathrm{d}}}{\mathrm{n.d.f.}_{\mathrm{m}} + \mathrm{n.d.f.}_{\mathrm{d}}}$$
  - Refit vertex
  - Repeat until converges or reaches maximum (3 to limit resource usage)

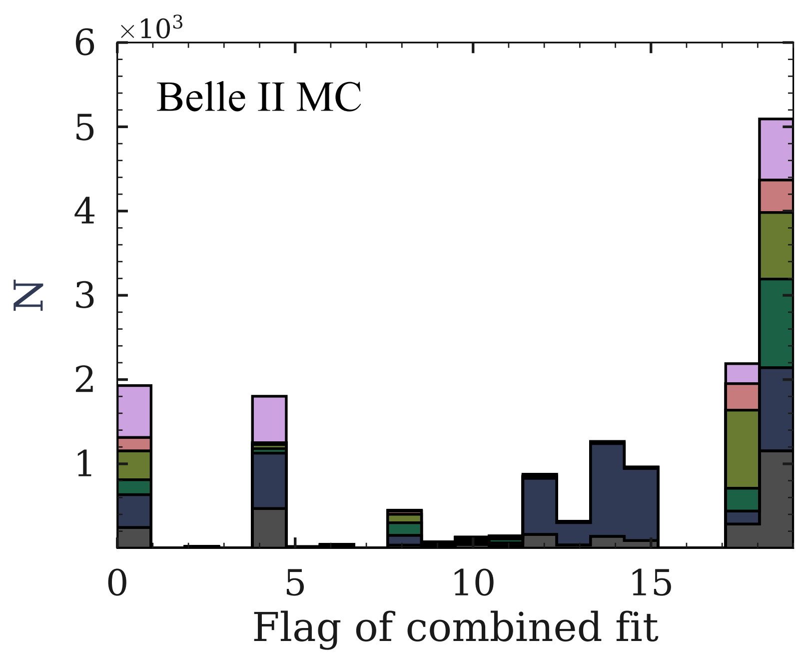
Example of hit reassignment
A whole stereo super layer was reassigned

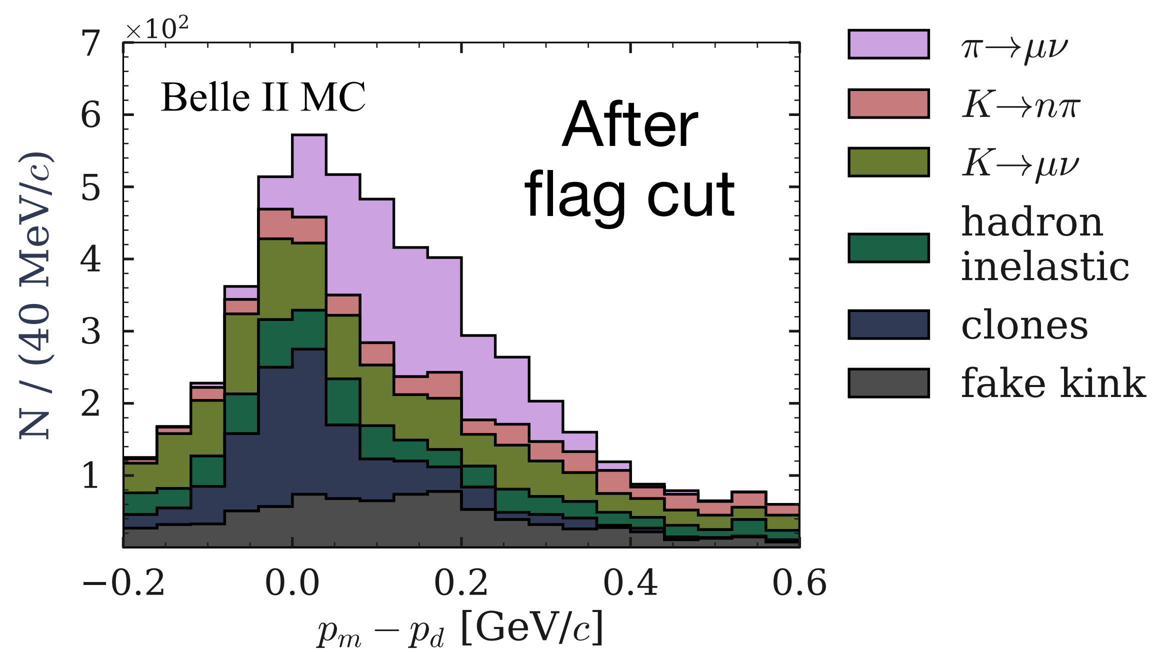


Default Reassigned MCTruth

Combined fit

- Clones of the tracks (one real track is reconstructed as two separate tracks) are inevitable in track finding
- They can imitate kinks
- It will be required to suppress their contribution in analysis
- Combine two tracks into one and fit it
- Do not save the result and only use a flag, based on the improvement of the fit result
- Around $50\ \%$ of clones can be suppressed with this flag, while having more than $90\ \%$ retention for real kinks
- Further suppression can be applied at analysis level with ML using other characteristic variables





Number of kinks for general case

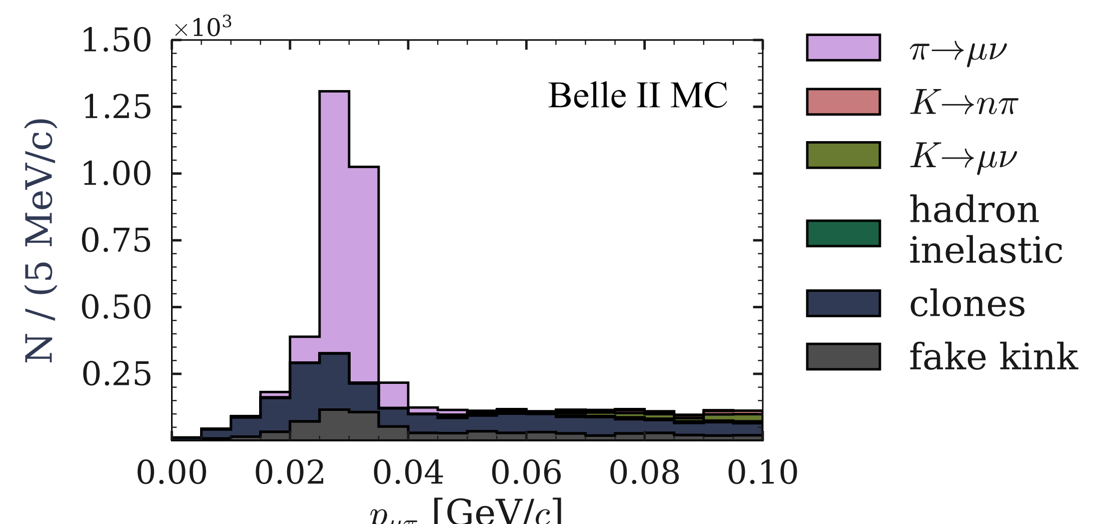
MC study with $\tau^+\tau^-$ and $B\bar{B}$ samples, 10^5 events in each

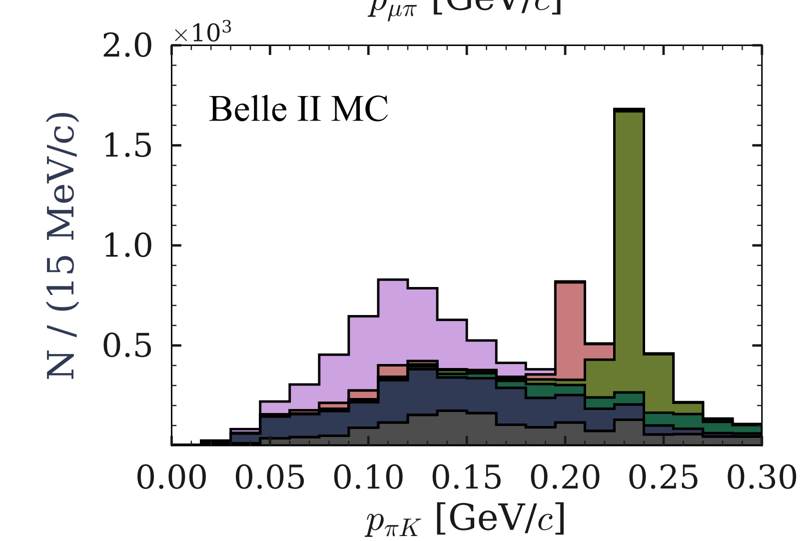
Total number of found kinks: $\approx 2.4 \times 10^3$ for $\tau^+\tau^-$ and $\approx 1.7 \times 10^4$ for $B\bar{B}$.

Among them:

- Combination of fake mother (daughter) track: $2\,\%$ ($8\,\%$) for $\tau^+\tau^-$ and $2\,\%$ ($9\,\%$) for $B\bar{B}$
- Fake combination of real tracks: $4\,\%$ for $\tau^+\tau^-$ and $8\,\%$ for $B\bar{B}$
- True kinks: $55\,\%$ for $\tau^+\tau^-$ and $46\,\%$ for $B\bar{B}$
  - Decay: $30\,\%$ for $\tau^+\tau^-$ and $74\,\%$ for $B\bar{B}$
  - Hadron inelastic: $69\,\%$ for $\tau^+\tau^-$ and $24\,\%$ for $B\bar{B}$
- Clones: $32\,\%$ for $\tau^+\tau^-$ and $36\,\%$ for $B\bar{B}$

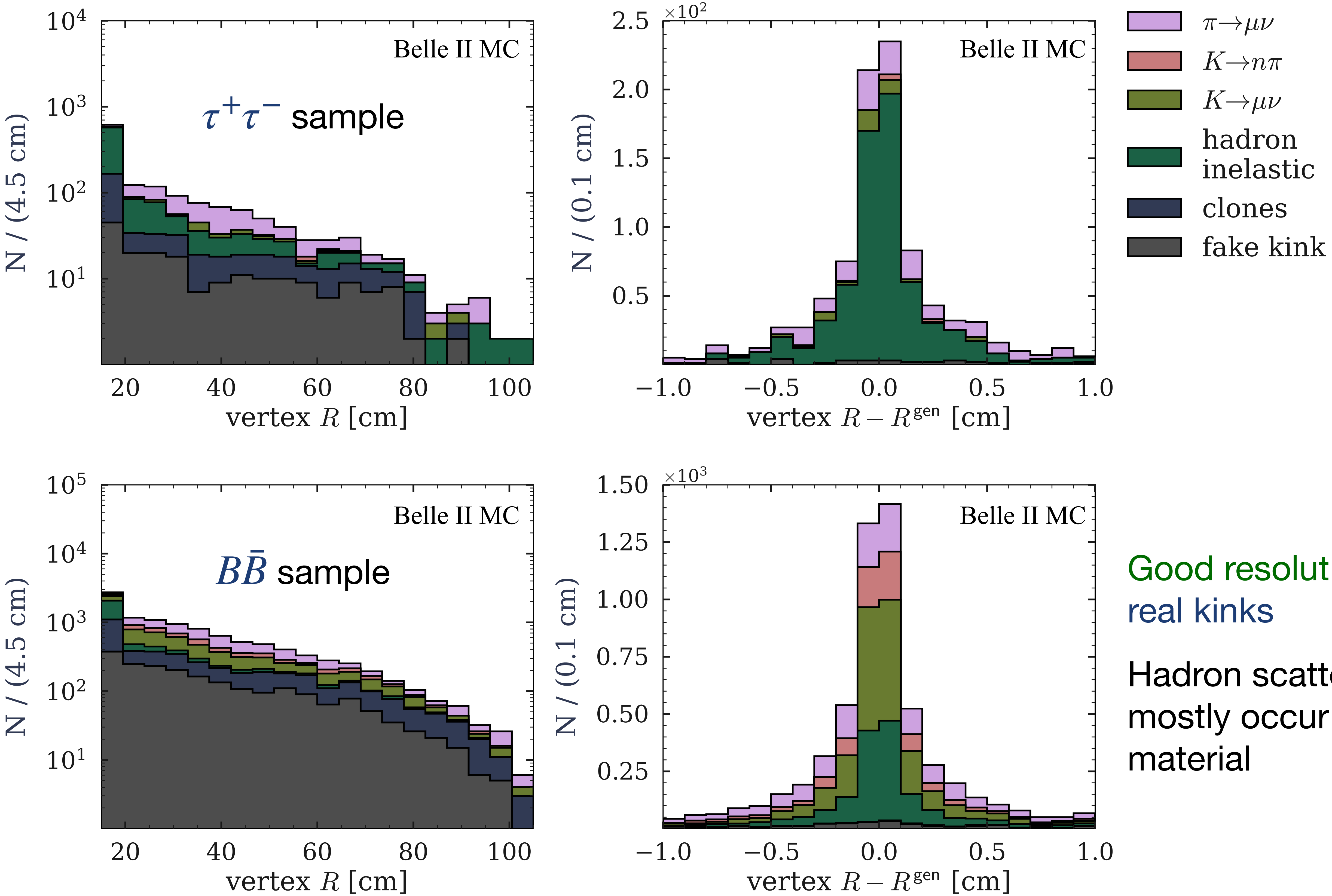
	N^{MC}	$N^{\mathrm{MC}}_{\mathrm{reco}}$	N_{kinks}	$N_{\mathrm{kinks}}/N^{\mathrm{MC}}_{\mathrm{reco}}$
τ sample				
All decays	6179	563	391	69%
Muon decays	48	1	0	–
Pion decays	5774	493	319	65%
Kaon decays	357	69	61	88%
$B\bar{B}$ sample				
Decay-in-flight	95830	8598	5597	65%
Muon decays	576	4	2	–
Pion decays	63964	4348	2204	51%
Kaon decays	31260	4246	3390	80%





$B\bar{B}$ sample after clone flag cut

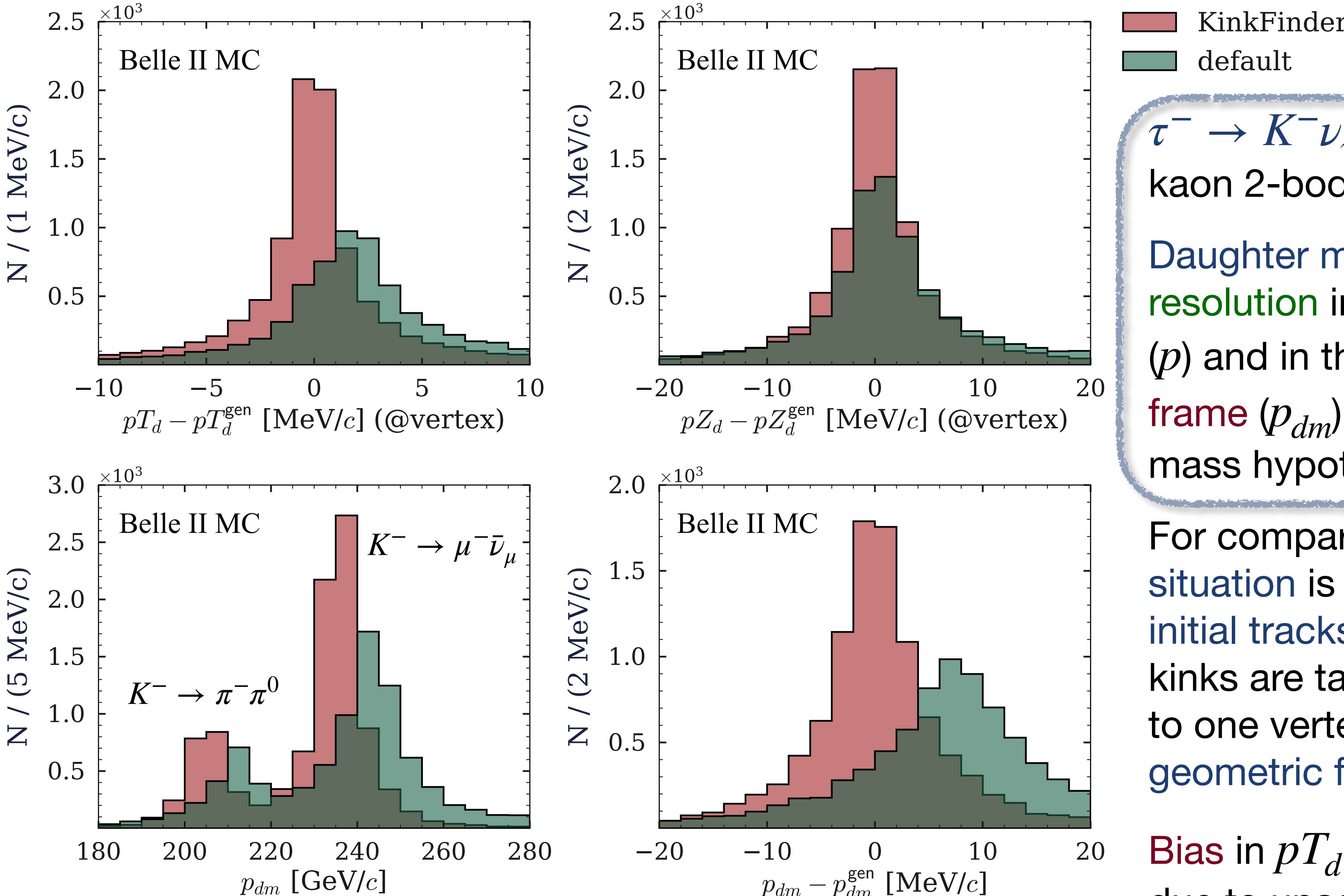
Kink Vertex reconstruction



Good resolution for real kinks

Hadron scattering mostly occur on VXD material

Daughter momentum resolution



$\tau^- \to K^- \nu_\tau$ sample with kaon 2-body decays

Daughter momentum resolution in the lab frame (p) and in the mother rest frame (p_{dm}) with correct mass hypotheses applied

For comparison, a default situation is shown: the initial tracks from found kinks are taken and fitted to one vertex with geometric fitter

Bias in pT_d for default is due to unaccounted dE/dx losses

Track splitting

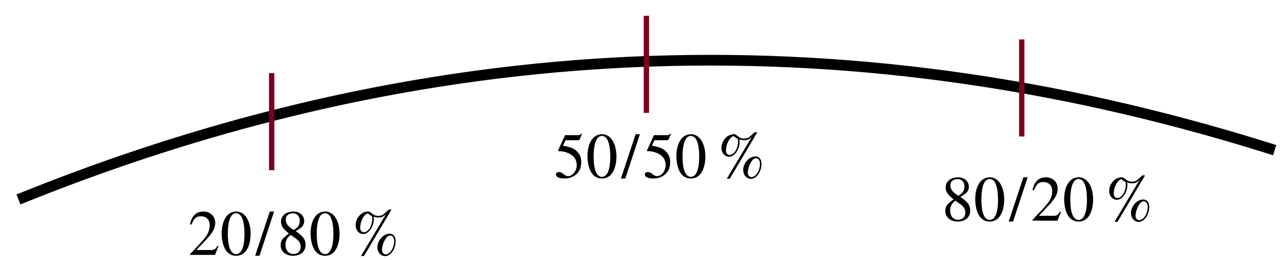
- Around $40\,\%$ of decays in flight of kaons from $\tau^- \to K^-\nu_\tau$ are reconstructed as one (combined) track
- Around $70\,\%$ of decays in flight of pions from $\tau^- \to \pi^-\nu_\tau$ are reconstructed as one (combined) track
- Combined tracks can be distinguished by low p-value and less fitted hits
  - Several times more ordinary tracks with low p-value and a small amount of hits
  - Not all tracks have bad p-value

- Split the track in 5 iterations (usually enough to converge) and do a binary search for the best splitting position based on the following criteria

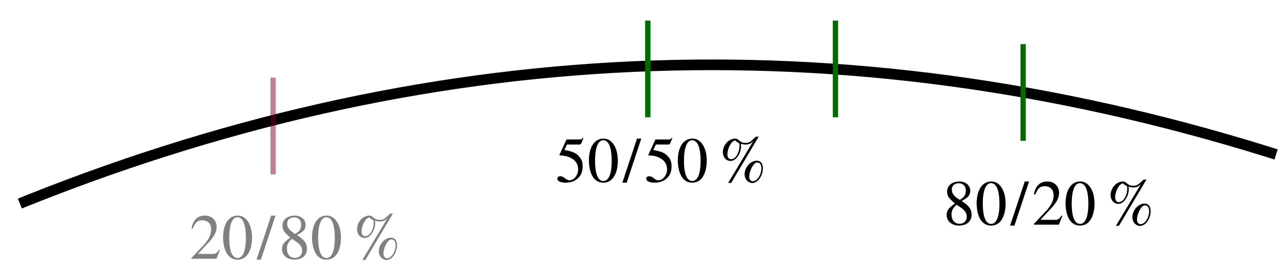
$$\left| \frac{\chi^2_{\text{comb}}}{\text{n.d.f.}_{\text{comb}}} - 1 \right|$$

- Fit new track pairs

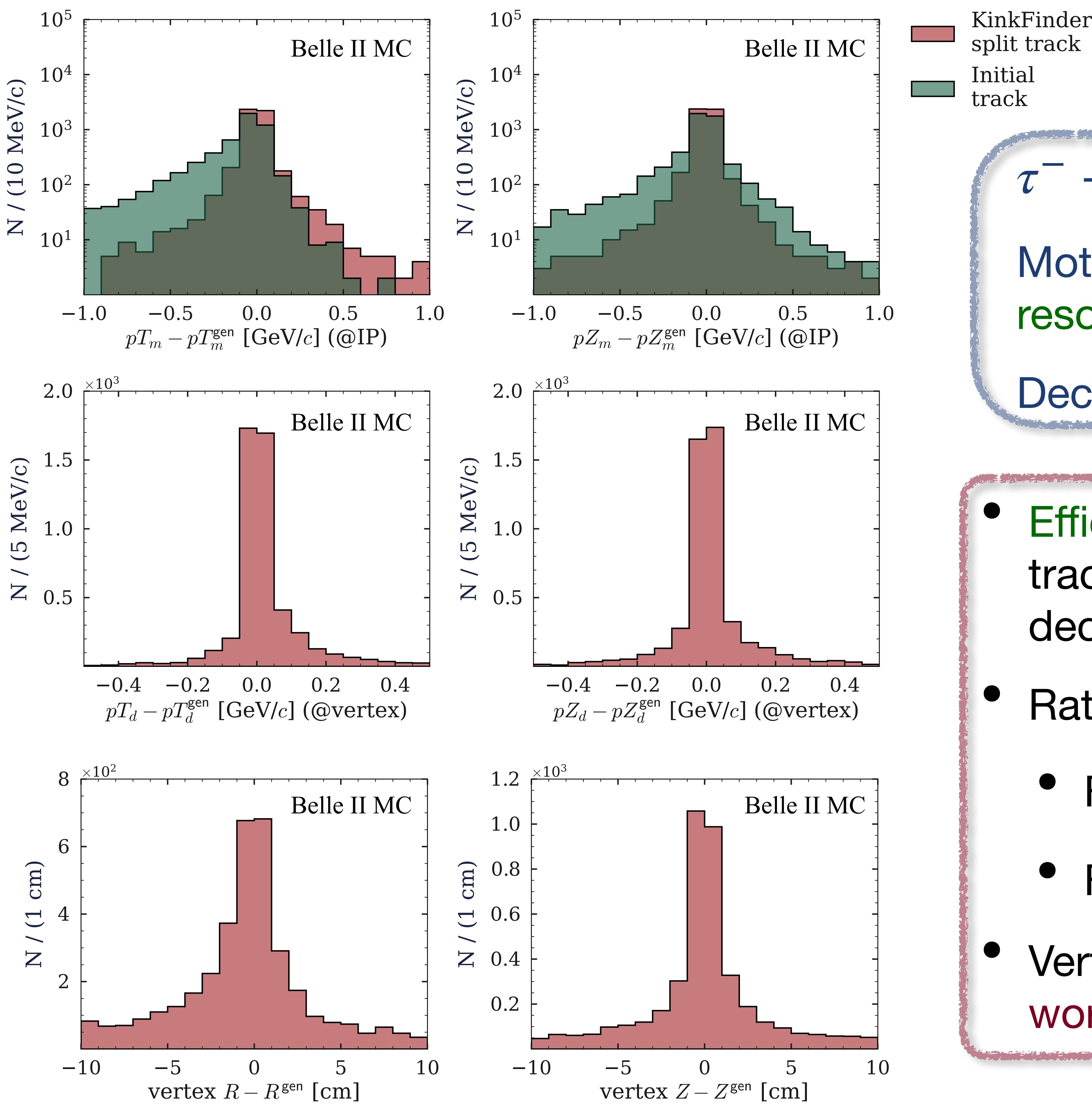
Initial track is split 3 times in 3 points



Choose the best edge and repeat



Splitting performance



$\tau^- \to \pi^- \nu_\tau$ sample with pion decay

Mother and daughter momenta resolution

Decay vertex resolution

- Efficiency to find kink from combined track is 37% (46%) for kaon (pion) decays
- Ratio of fake split to real:
  - For $\tau^+\tau^-$, around 4.2-to-1
  - For $B\bar{B}$, around 2.1-to-1
- Vertex and momenta resolutions are worse than ones for track pairs

Conclusions

- First version of the Kink Finder was developed within the Belle II Tracking Software
- Two major cases were considered: combining two reconstructed tracks and splitting one track, combined from hits of both mother and daughter, to create a kink
- The source code can be found in the Belle II GitHub repository: https://github.com/belle2/basf2/tree/main/tracking/modules/kinkFinder
- Both cases do not change the default track finding
- The obtained finding efficiency is good while the reasons of the inefficiency are understood
- The resolution of the decay vertex and momenta at it are improved significantly compared to the default tracking
- Effect on PID is still to be studied

- There is still room for further improvements

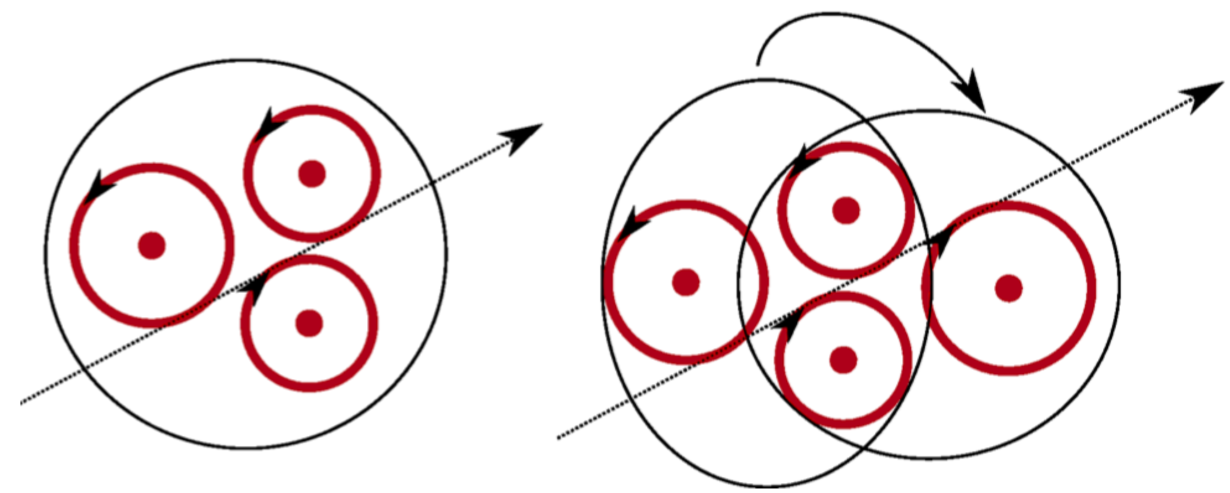
Thank you for attention!

Backup

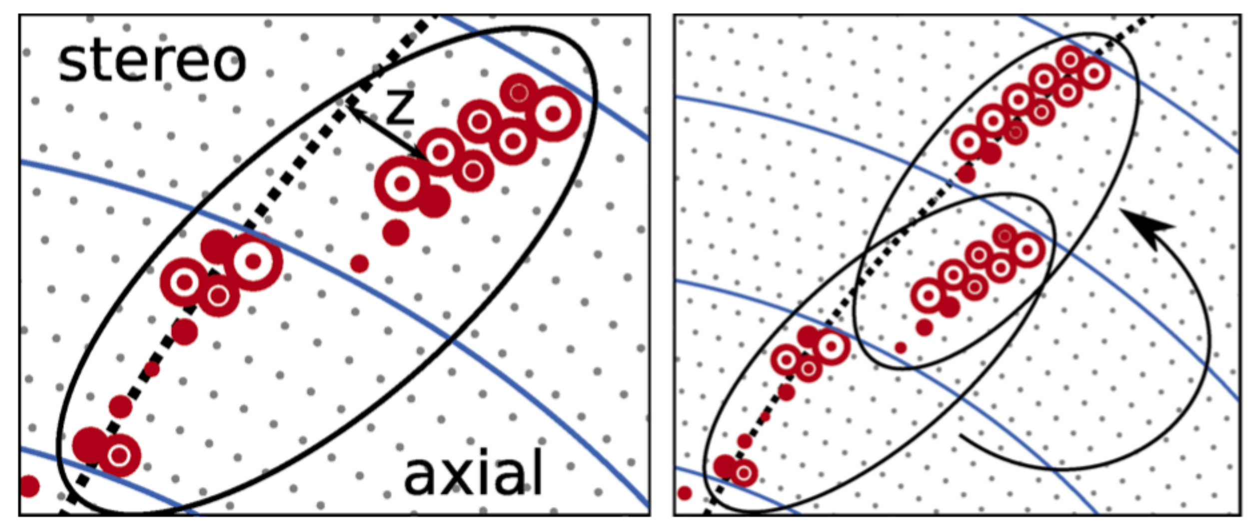
Local CDC Track Finding

- Cellular Automaton based local search for tracks
- $E_i = \sum w_{ij} + \sum \Theta_j$
    1. Combine hits in segments located in one superlayer
    2. Combine segments into tracks
- Segments are first combined with tracks reconstructed by global track finding algorithm
- The remaining segments are used to reconstruct additional tracks
    - High efficiency for displaced tracks
    - Currently is switched off due to high fake rate

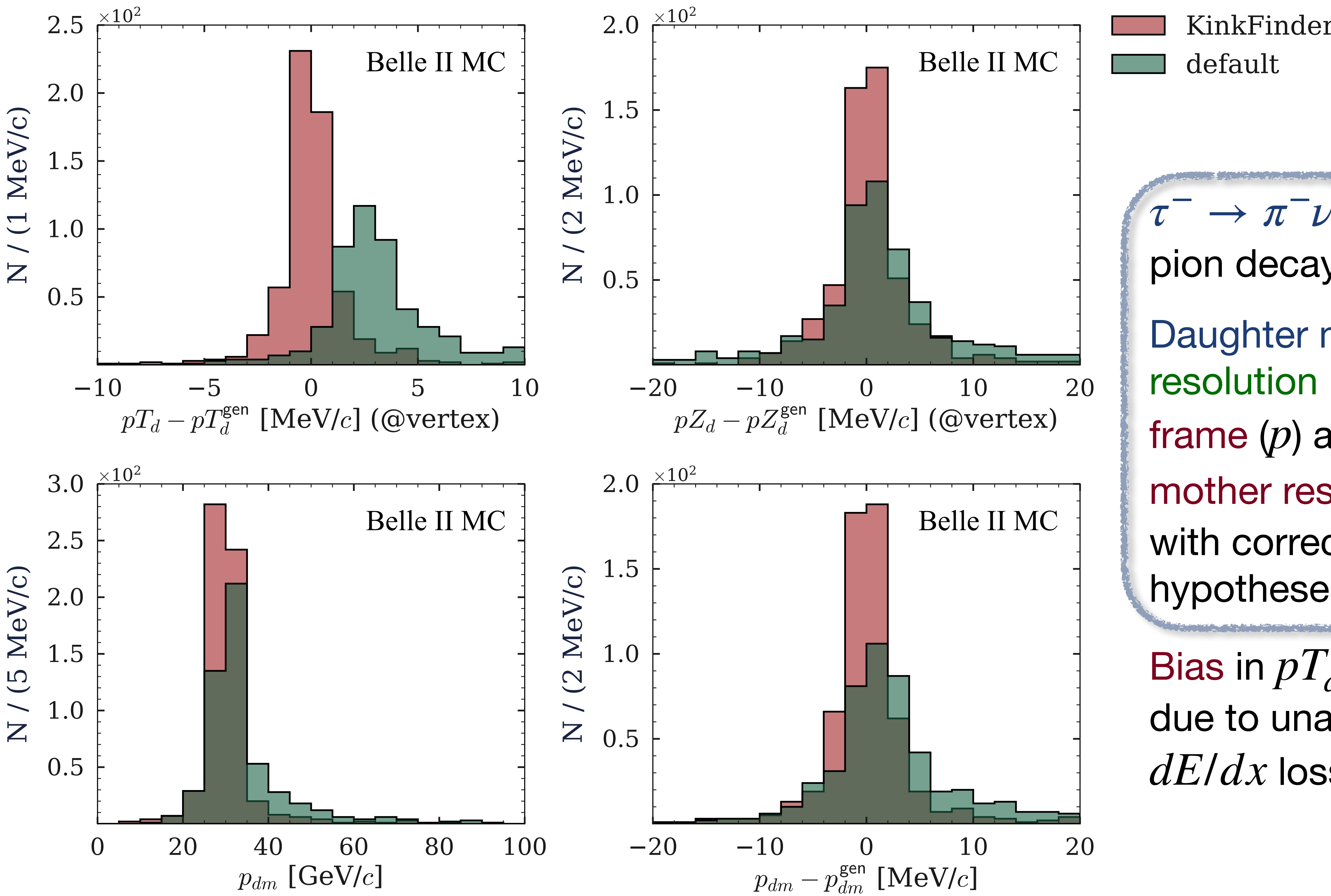
Build segments



Build tracks



Daughter momentum resolution (2)

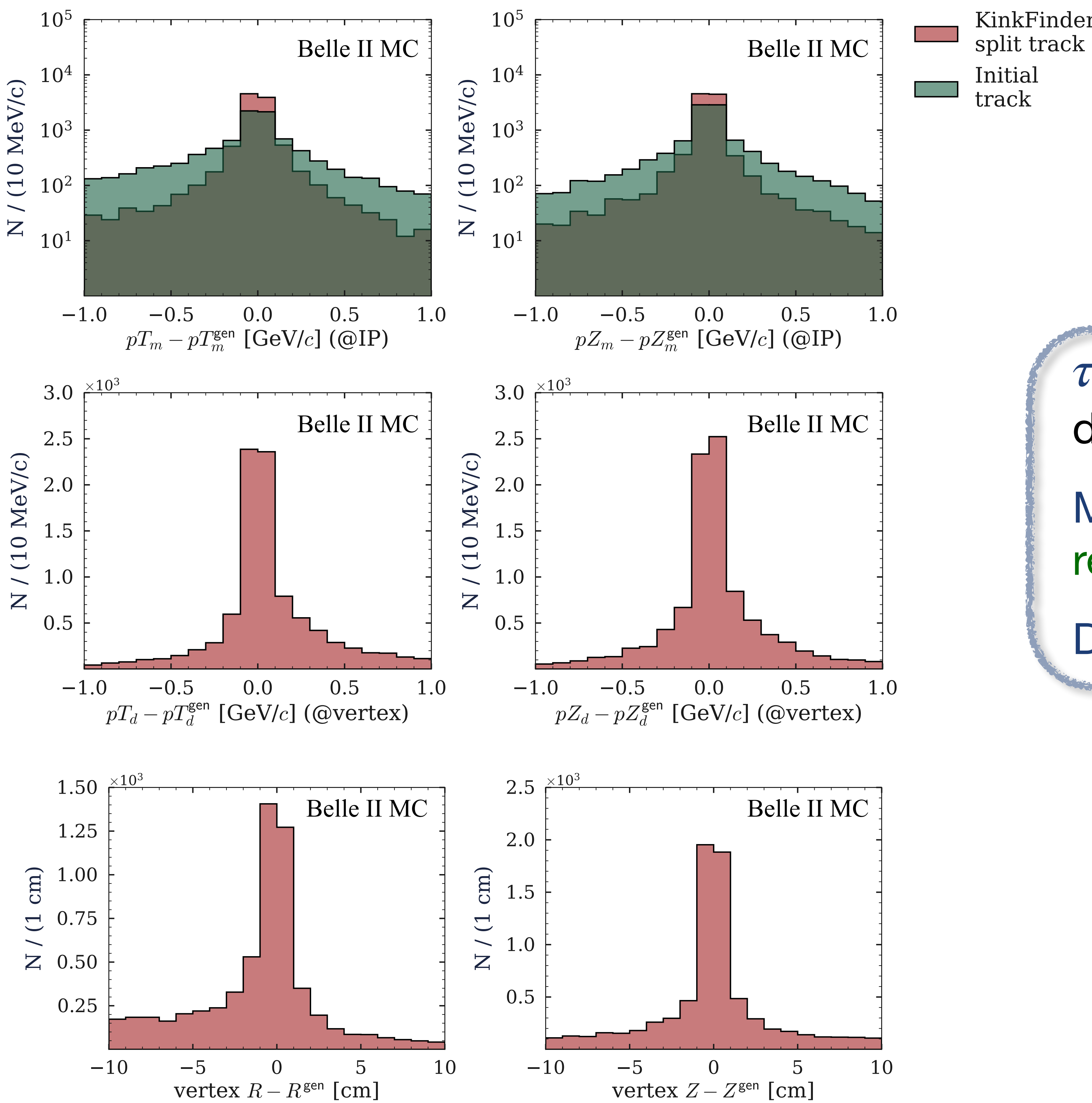


$\tau^- \to \pi^- \nu_\tau$ sample with pion decay

Daughter momentum resolution in the lab frame (p) and in the mother rest frame (p_{dm}) with correct mass hypotheses applied

Bias in pT_d for default is due to unaccounted dE/dx losses

Splitting performance (2)



$\tau^- \to K^- \nu_\tau$ sample with kaon decays

Mother and daughter momenta resolution

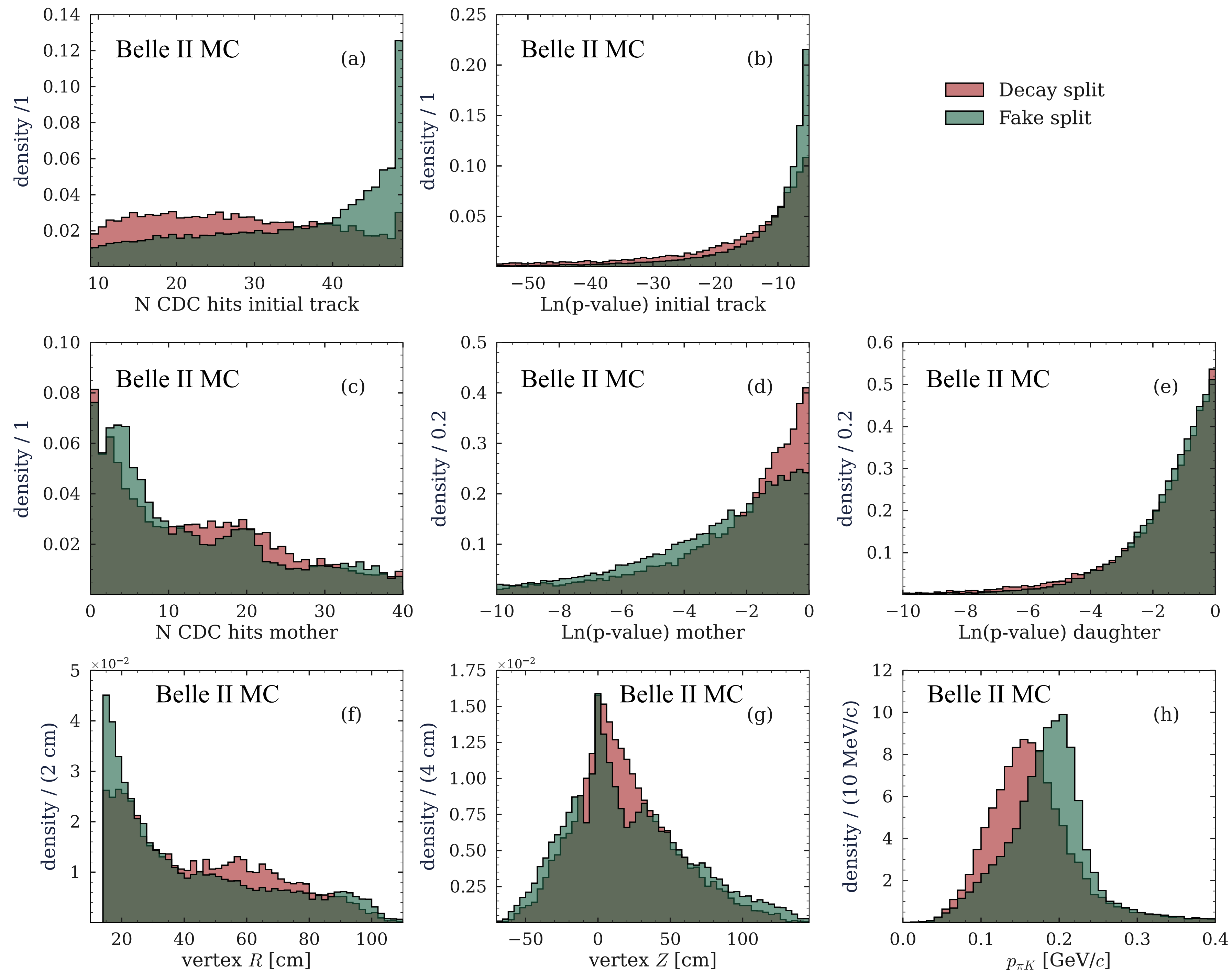
Decay vertex resolution

Fake splitting suppression

To suppress fake track splitting, an ML algorithm can be trained at the analysis level using the following variables as an example

Distributions for $B\bar{B}$ sample are shown

They are slightly different with $\tau^+\tau^-$ sample, thus dependence on multiplicity exists

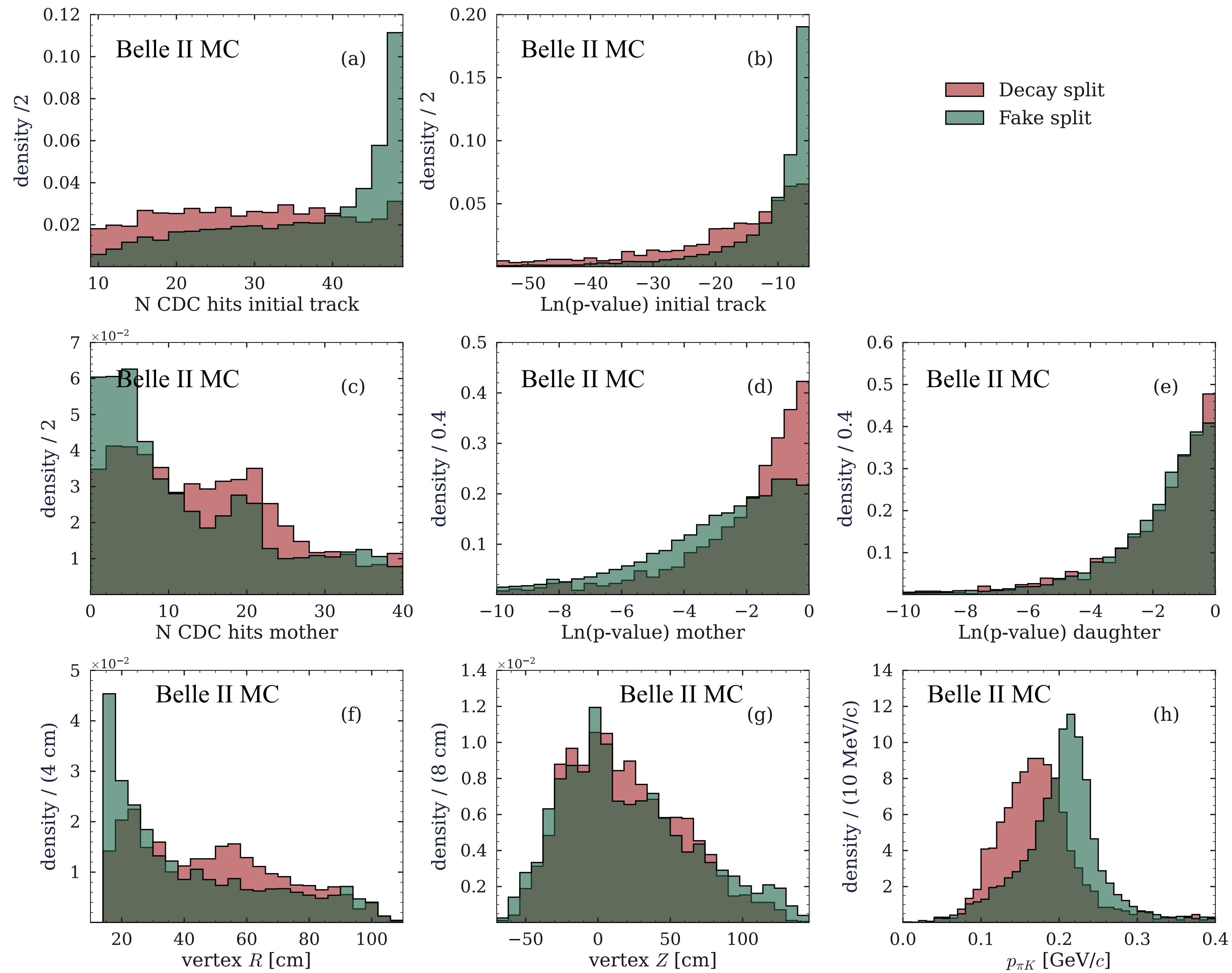


Fake splitting suppression (2)

To suppress fake track splitting, an ML algorithm can be trained at the analysis level using the following variables as an example

Distributions for $\tau^+\tau^-$ sample are shown

They are slightly different with $B\bar{B}$ sample, thus dependence on multiplicity exists



Problems and solutions

- Inefficiency in finding kinks from reconstructed track pairs
    - Finding hits in missing superlayers can improve z resolution
- Inefficiency in reconstruction of the daughter particles
    - Local track finding can improve the efficiency at least by $50\,\%$
    - Short tracks passing one-two superlayers can be reconstructed since usually their segments are found (may generate extra background)
- Processing time due to multiple refitting
    - Prepare a special fitter for the kinks, e.g., saving results of each iteration of the KalmanFilter can reduce the number of fitting to two (forward and backward)
    - Better optimization of the fitter's parameters (number of general iterations to find the optimum)
- Development of special fitter which will consider kink as one continuous object can improve both momentum and vertex resolutions
- New GNN algorithm developed in Belle II shows incredible performance for displaced tracks. It may also increase kink reconstruction efficiency if trained properly
    - See for example, a talk by Lea Reuter at Hemholtz AI Conference 2024