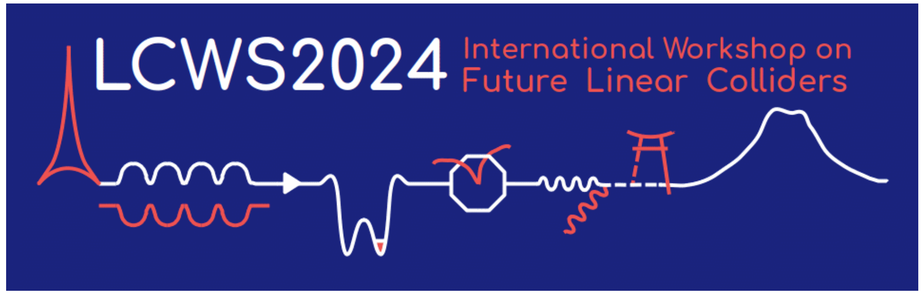
Software, Reconstruction, Computing

Summary

Carsten Hensel on behalf of the SRC Conveners

Andre Sailer, Frank Gaede, Hiroaki Ono, Lindsey Grey, Ryu Sawada



Software, Reconstruction, Computing

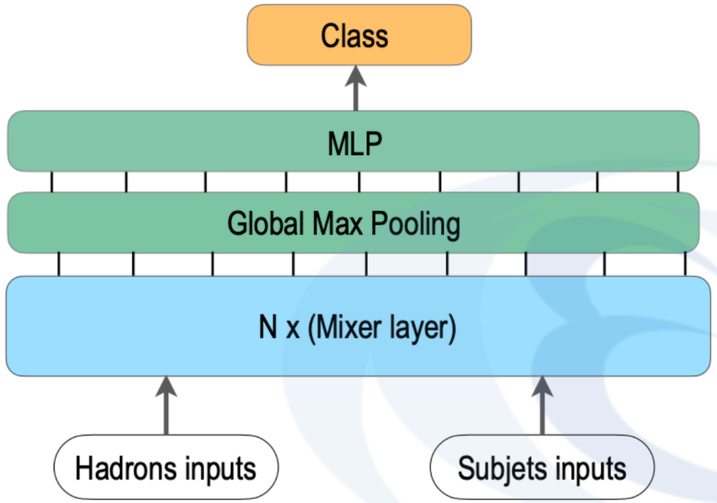
- 3 sessions
- 9 presentations
- Ample time for discussions

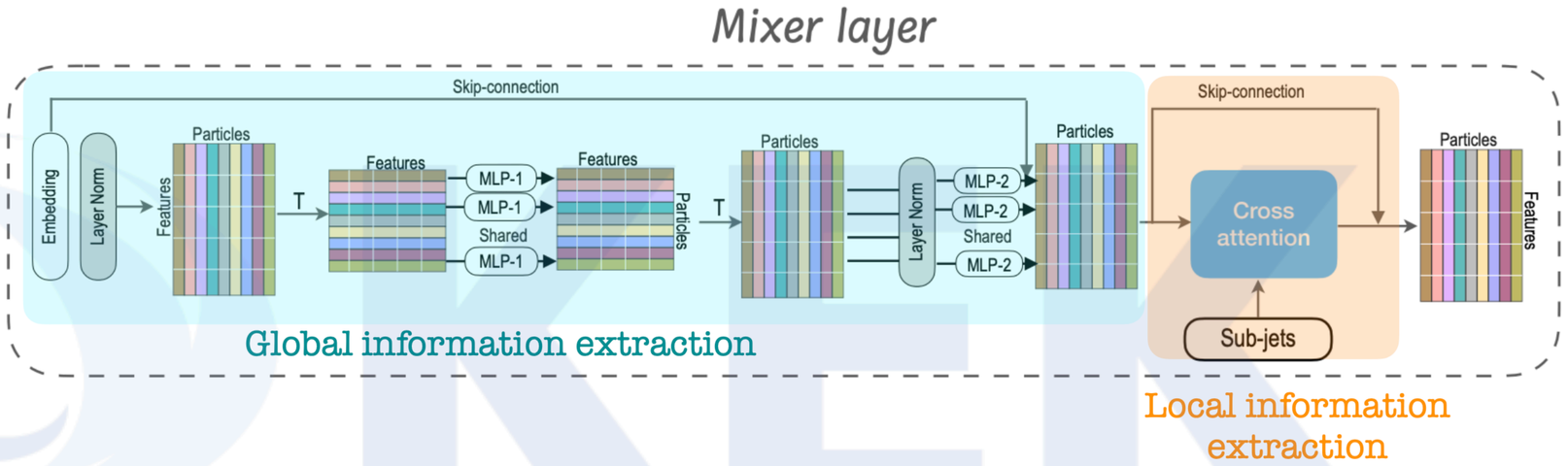
Day 1

- Streamlined jet tagging network assisted by jet prong structure
  (Ahmed Hammad)
- Jet origin identification: AI enhanced reconstruction for Higgs factory
  (Manqi Ruan)
- Application of Particle Transformer to Quark Flavor Tagging in the ILC Project
  (Risako Tagami)

Mixer network

Mixer layer has two MLP that mix both features and Particle tokens (similar to the transformer) which allow for fast extraction of the global features of the event. Local information is extracted from the subjects via Cross-attention layer.





Jet tagging task can be divided into two main parts:

- Global information extraction

The network learns how important each jet constituent to all other constituents via two MLPs.

$$Y_{i,j} = X_{i,j} + [\left(W_2 \sigma W_1(\text{LayerNorm}(\mathbf{X})^T)\right)^T]_{i,j},$$

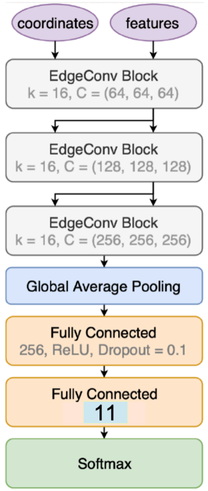
$$\widetilde{X}_{i,j} = Y_{i,j} + \left(W_4 \sigma W_3(\text{LayerNorm}(Y_{i,j}))\right)$$

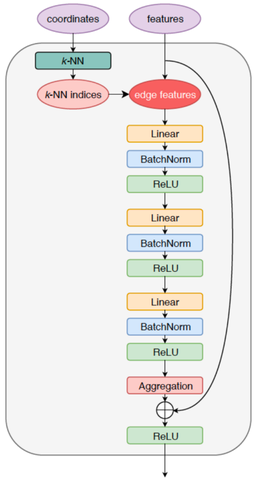
- Local information extraction

The network learns how important each jet constituent to the sub-jet it belongs to

$$P\left(\text{Hadrons inside jet} \mid \text{subjet cluster}\right) = P\left(x_i | y_\alpha\right)$$

Particle Net: IO





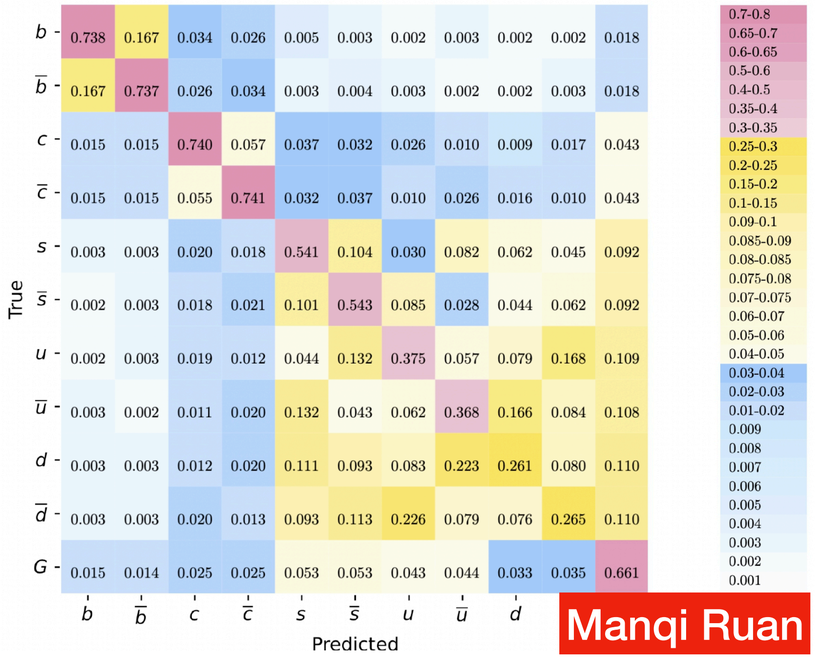
Variable	Definition
$\Delta\eta$	difference in pseudorapidity between the particle and the jet axis
$\Delta\phi$	difference in azimuthal angle between the particle and the jet axis
$logp_T$	logarithm of the particle's p_T
$logE$	logarithm of the particle's energy
$log\frac{p_T}{p_T(jet)}$	logarithm of the particle's p_T relative to the jet p_T
$log\frac{E}{E(jet)}$	logarithm of the particle's energy relative to the jet energy
ΔR	angular separation between the particle and the jet axis ($\sqrt{(\Delta\eta)^2+(\Delta\phi)^2}$)
d0	transverse impact parameter of the track
d0err	uncertainty associated with the measurement of the d0
z0	longitudinal impact parameter of the track
z0err	uncertainty associated with the measurement of the z0
charge	electric charge of the particle
isElectron	if the particle is an electron
isMuon	if the particle is a muon
isChragedKaon	if the particle is a charged Kaon
isChragedPion	if the particle is a charged Pion
isProton	if the particle is a proton
isNeutralHadron	if the particle is a neutral hadron
isPhoton	if the particle is a photon

Table 3. The input variables used in ParticleNet for jet flavor tagging at the CEPC.

- Input: measurable information of all reconstructed jet particles
- Output: 10(11)-likelihoods to different categories

11-dim migration behavior

- Let the jet be identified as the category with highest likelihood:
- Pid: ideal Pid – three categories
  - Lepton identification
  - **Charged Kaon identification**
  - Neutral Kaon identification
- Patterns:
  - ~ Diagonal at quark sector...
  - P(g→q) < P(q→g)...
  - Light jet id...



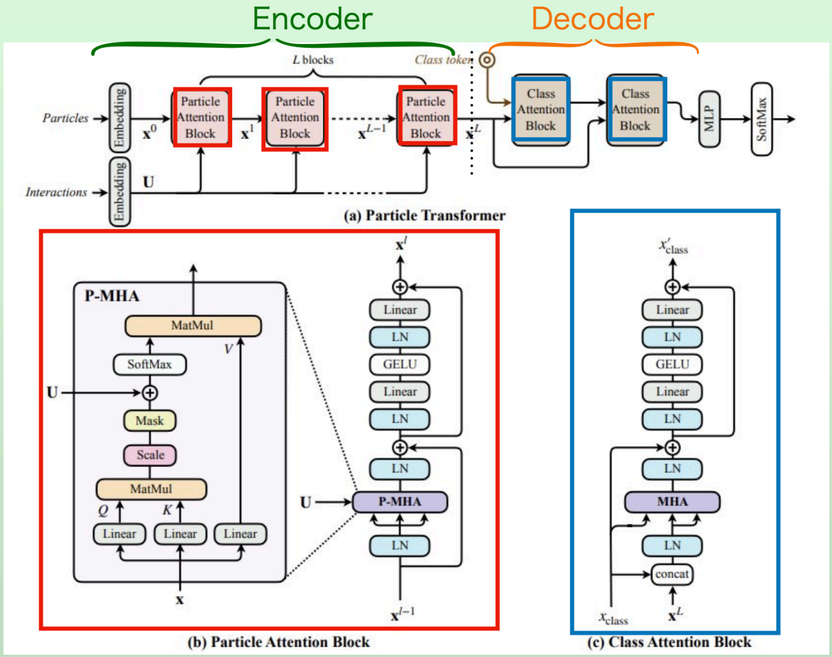
Manqi Ruan

Particle Transformer (ParT)

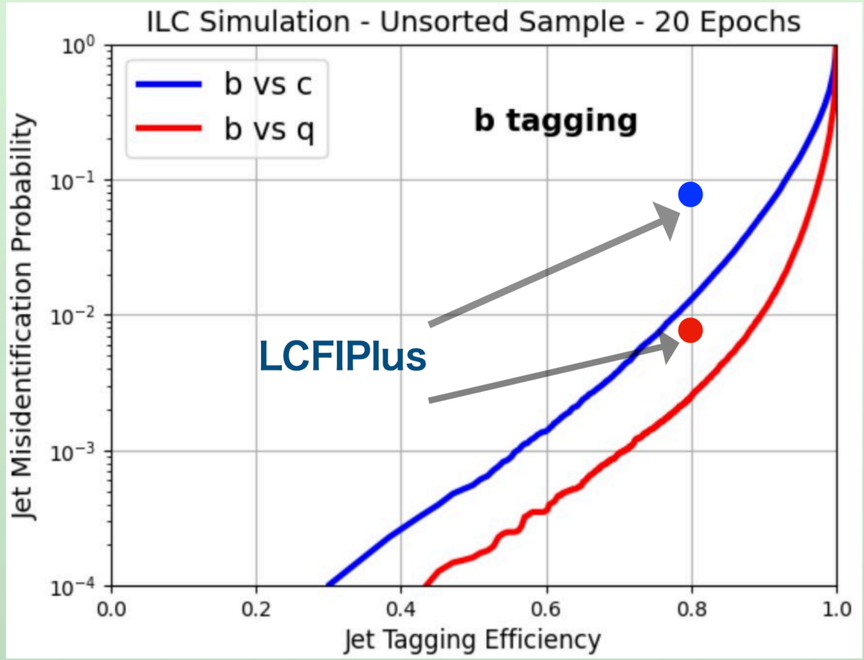
- **ParT** is a modified Transformer model for Jet research（published in 2022.）
  - Considering the nature of Jet, input the physical quantity calculated from the quaternion momentum of two particles to Multihead attention.
- ParT has surpassed the performance of ParticleNet, which has been the highest-performing（arXiv: 2202.03772）。

Event classification for JetClass

Event	H→bb Rej. 50%	H→cc Rej. 50%
Particle Net	0.013 %	0.04 %
ParT	0.0094%	0.024%



Performance of ParT



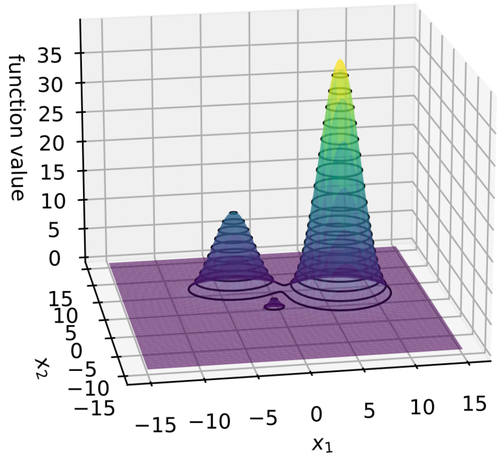
Risako Tagami

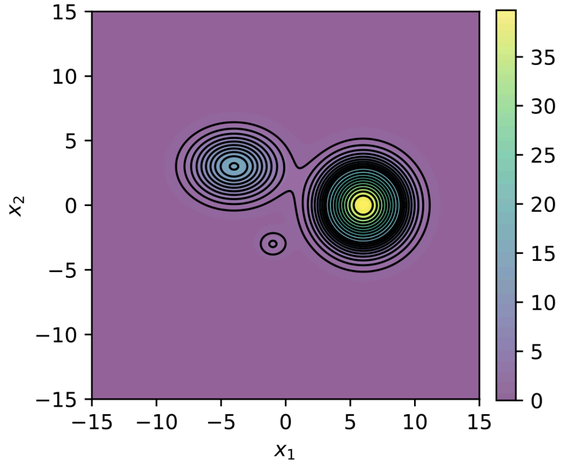
Day 2
Part 1

- Classifying Importance Regions in Monte Carlo Simulations with Machine Learning
(Raymundo Ramos)
- Development of Particle Flow Algorithm with GNN for Higgs Factories
(Tatsuki Murata)
- Impact of NLO QCD on Key Physics Processes at Future Higgs Factories
(Zhijie Zhao)

Learn divisions of a function with multiple peaks

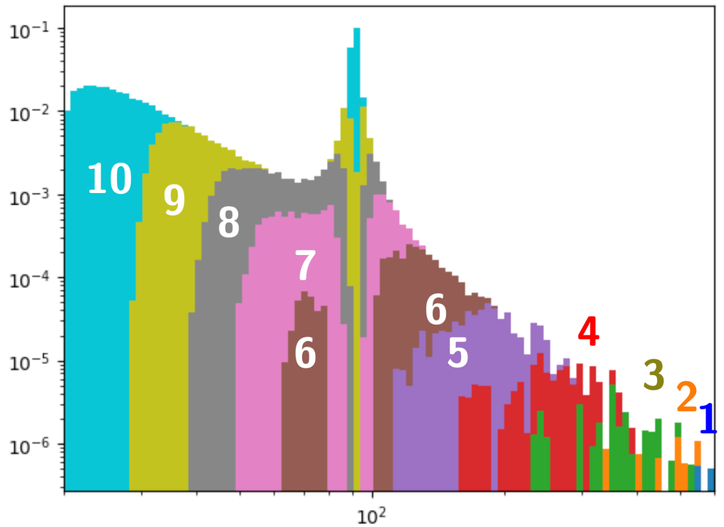
20 regions with similar contribution to value of integral





Generate events: 10 usable regions

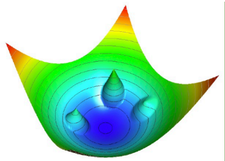
e^-e^+ invariant mass projection



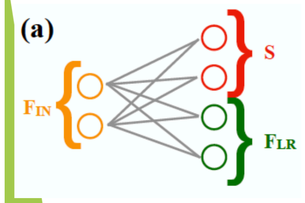
- Sample each region until enough events are accumulated.
  NN can tell which regions points belong to.
- Select points using correct result.

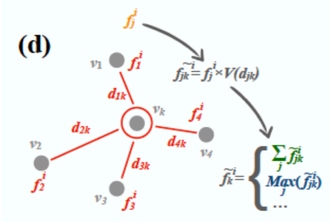
Raymundo Ramos

GravNet and object condensation



- GravNet arXiv:1902.07987
- The virtual coordinate (S) is derived from inputs with simple multilayer-perceptron(MLP)
- Convolution using “distance” at S (bigger convolution with nearer hits)
- Concatenate the output with MLP



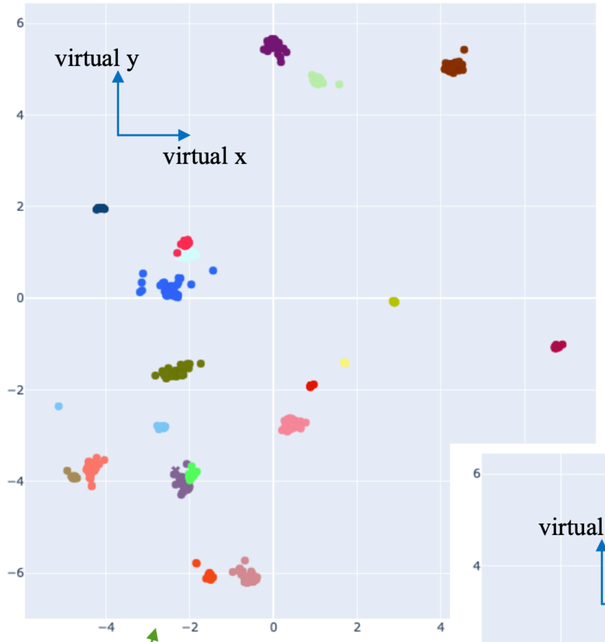


arXiv:2002.03605

Object condensation (loss function)

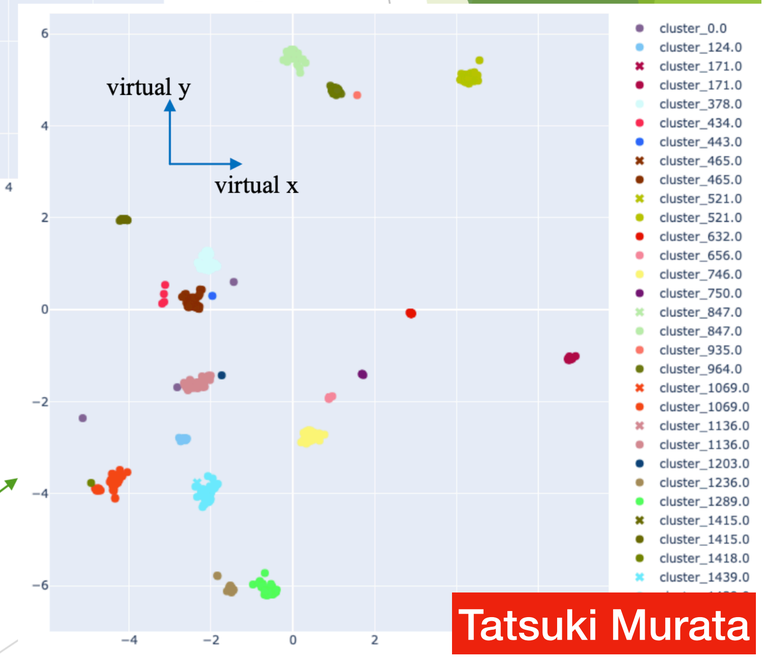
$$L = L_p + s_C(L_\beta + L_V)$$

- Condensation point : the hit with largest β at each MC cluster
- L_V: attractive potential to the condensation point of the same cluster and repulsive potential to the condensation point of different clusters
- L_β: pulling up β of the condensation point (up to 1)
- (L_p: regression to output features)



Output features
Virtual coordinate

Colored by true clusters

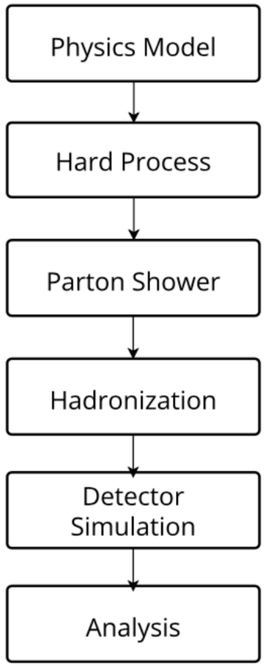


Colored by reconstructed clusters

Tatsuki Murata

Status and Goals

- Present events for analysis of e^+e^- colliders:
  - Leading order matrix elements are calculated by Whizard 1.95.
  - Parton shower and hadronization are performed by Pythia6.
  - OPAL tune for LEP is used.
- Our goals:
  - Upgrade the simulation chain to Whizard3+Pythia8.
  - Get agreement with LEP data, especially the neutral hadrons.
  - Include NLO matching because of the requirement of high precision.



Summary

- The MC simulation chain is necessary to upgrade to modern generators with NLO precision.
- We test the NLO mode of Whizard.
- We get good agreement between LO and NLO events at reconstruction level for $e^+e^- \to q\bar{q}$.
- The NLO corrections play an important role in $e^+e^- \to \mu^+\mu^- b\bar{b}$.
- Further checks are necessary.

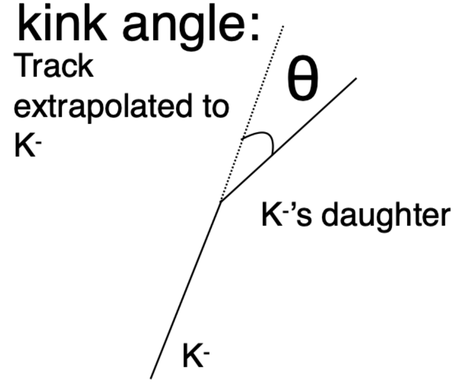
Day 2
Part 2

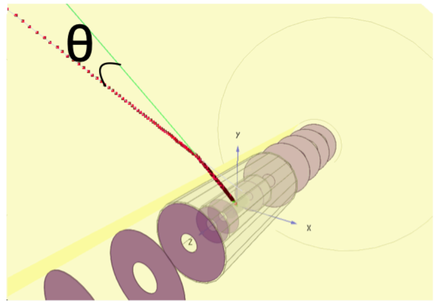
- Finding Kink Signatures of LLPs in TPC at ILC
  (Jurina Nakajima)
- Towards Production Readiness with th Key4HEP Software Stack for Future Colliders
  (Thomas Madlener)
- Fast Timing for Particle ID
  (Jenny List)

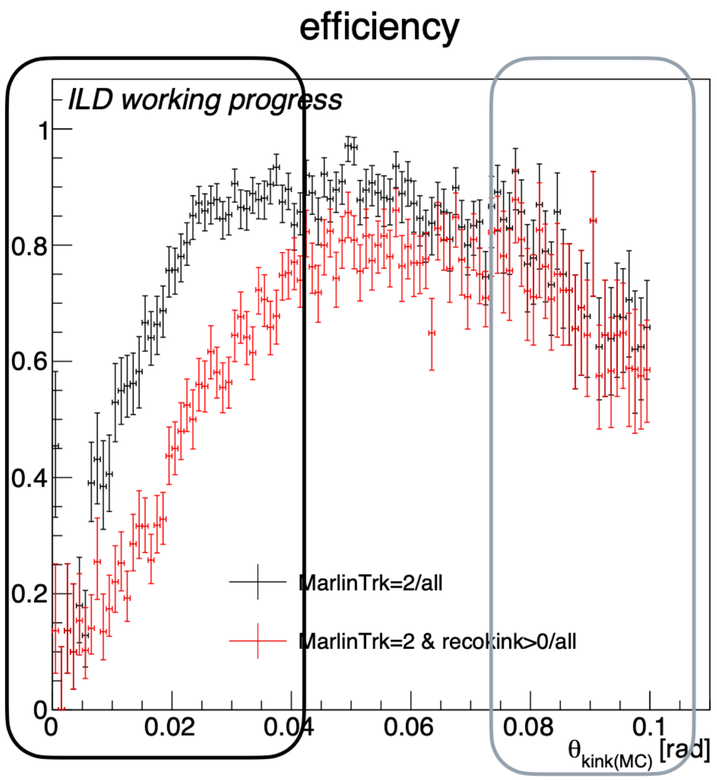
Efficiency dependence on the kink angle

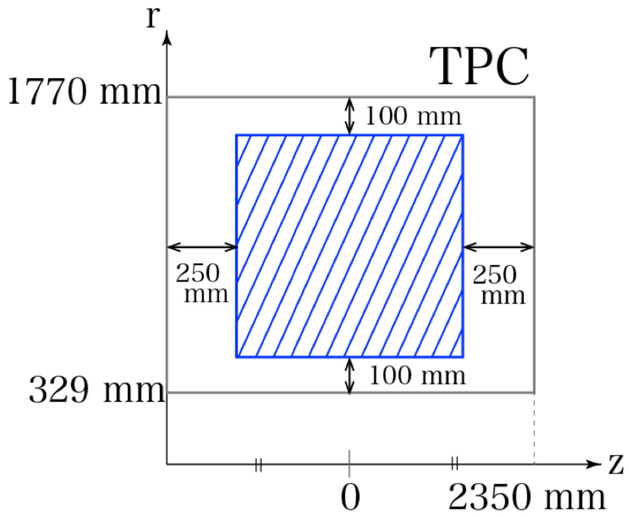
Precut made from **MCparticles**

- Endpoint(MC) inside TPC:
  $329 + 100 < r < 1770 - 100\,[\mathrm{mm}]\ ,\ |z_{max}| < 2350 - 250[\mathrm{mm}]$
- (MC) # of charged daughter of Kaon = 1









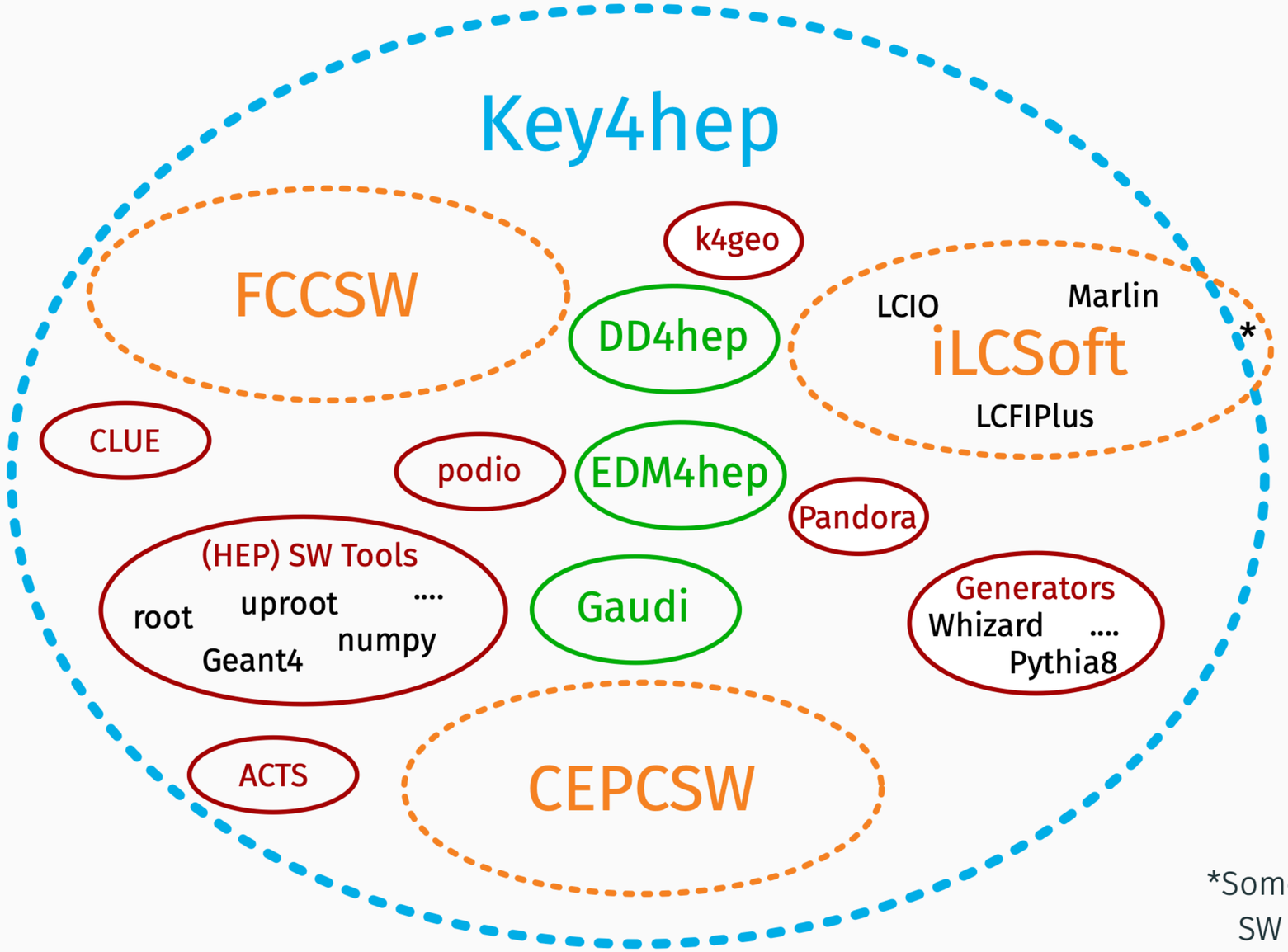
all: imposing only precut

Black: nTrack = 2 (tracking efficiency)
Red: nTrack =2 and Kink > 0 (tracking and KinkFinding efficiency)

Merged into **single** track?

Efficiency for very displaced 2nd track?

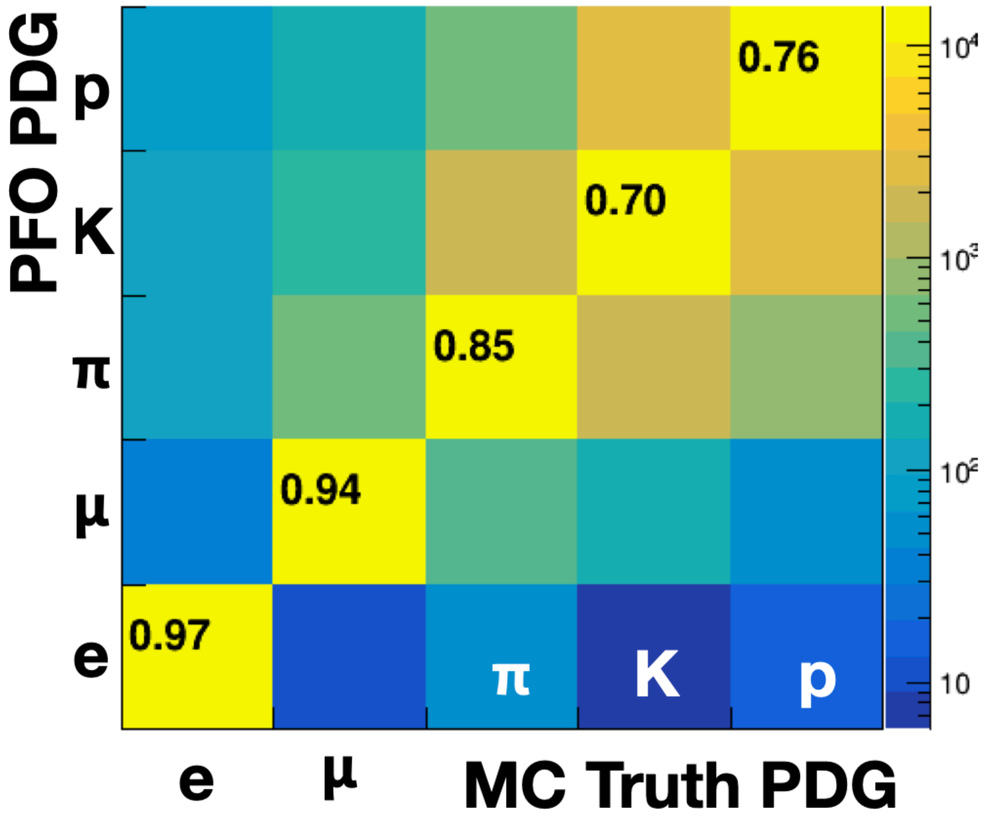
Key4hep (simplified) overview



*Some testbeam related SW not yet included

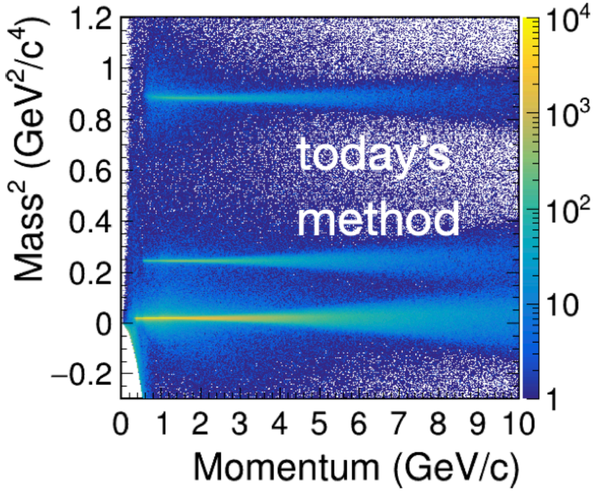
Towards application in analysis

CPID



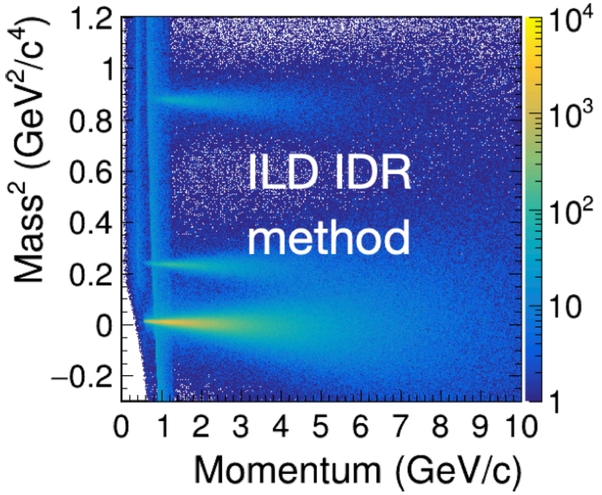
er hits!
ECal hits!
ed for

All this not available on mc-2020 250GeV ILD DST mass production — can only use IDR ToF



CPID on single particles, 1-100 GeV, with

- **dE/dx 4.5%**
- **IDR ToF, 50ps /hit**
- **Pandora PID**
- **LeptonID in jets**



- outlook:
  - new track length in master, could be used in a next ILD MC production
  - hit -> PFO time algorithms not yet committed, but could live with effective smearing of true Geant time

Jenny List

Conclusion/Summary

- Diverse set of contributions
- Many fruitful discussions
- Many thanks to all participants
- Many thanks to the students helping with the setup
- Many thanks to LCWS organizers