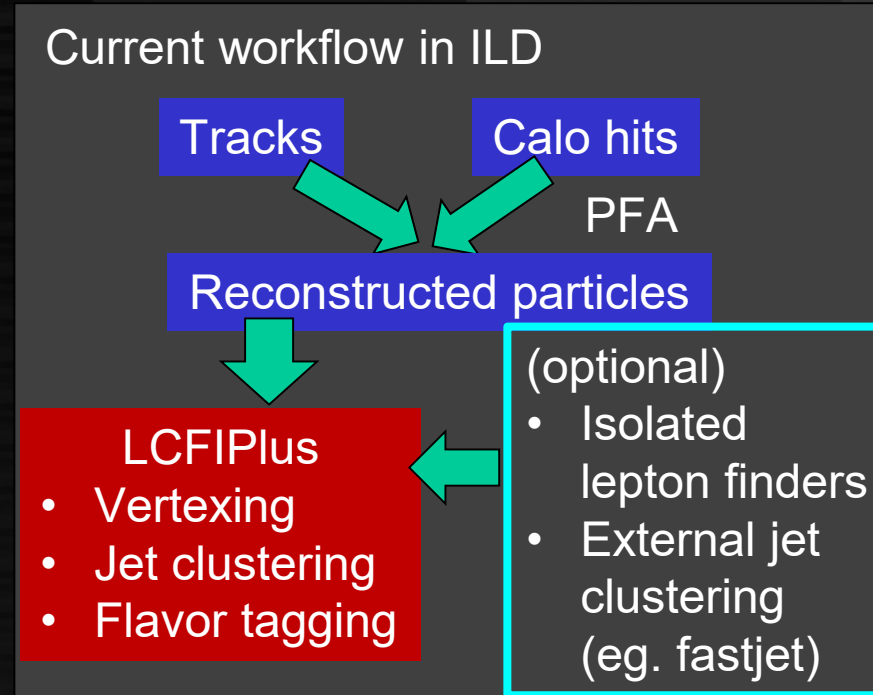


Introduction of LCFIPlus (+ ongoing effort)

Taikan Suehara
(Kyushu University)

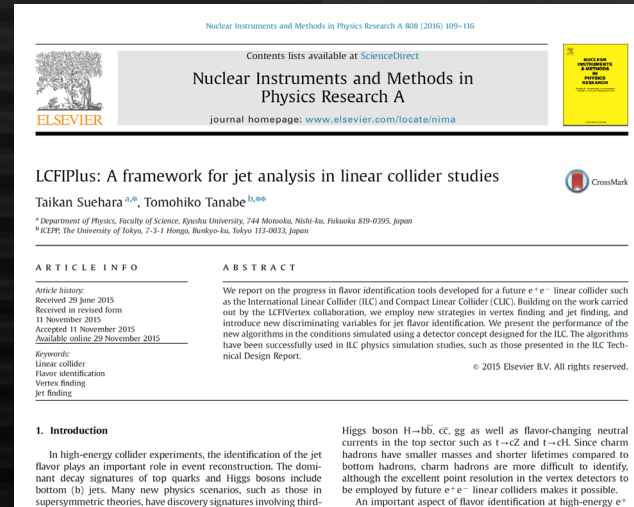
Contents

- LCFIPlus: integrated jet analysis tools since 2013
 - Vertex finding
 - Jet clustering
 - Flavor tagging
- Design consideration of LCFIPlus
 - Focusing difference to modern DNN methods



LCFIPlus: introduction

- LCFIVertex developed by UK group
 - <https://doi.org/10.1016/j.nima.2009.08.059>
 - Based on topological clustering (ZVTOP for SLD)
 - Budget cutoff at UK in ~2010 stopped all activities
- LCFIPlus transfers the effort
 - Target: exceed the performance of LCFIVertex
 - More robust to multi-jet environment ($H \rightarrow bb/cc, ZHH$)
 - Published for **ILC DBD** in 2013, also used for **CLIC CDR**
 - Features
 - Build-up secondary vertex finder
 - Integrated jet clustering with vertex information
 - Beam-jet rejection with Durham/Valencia/KT algorithm
 - BDT (by TMVA) for flavor tagging



<https://doi.org/10.1016/j.nima.2015.11.054>



<https://github.com/lcfiplus/LCFIPlus>

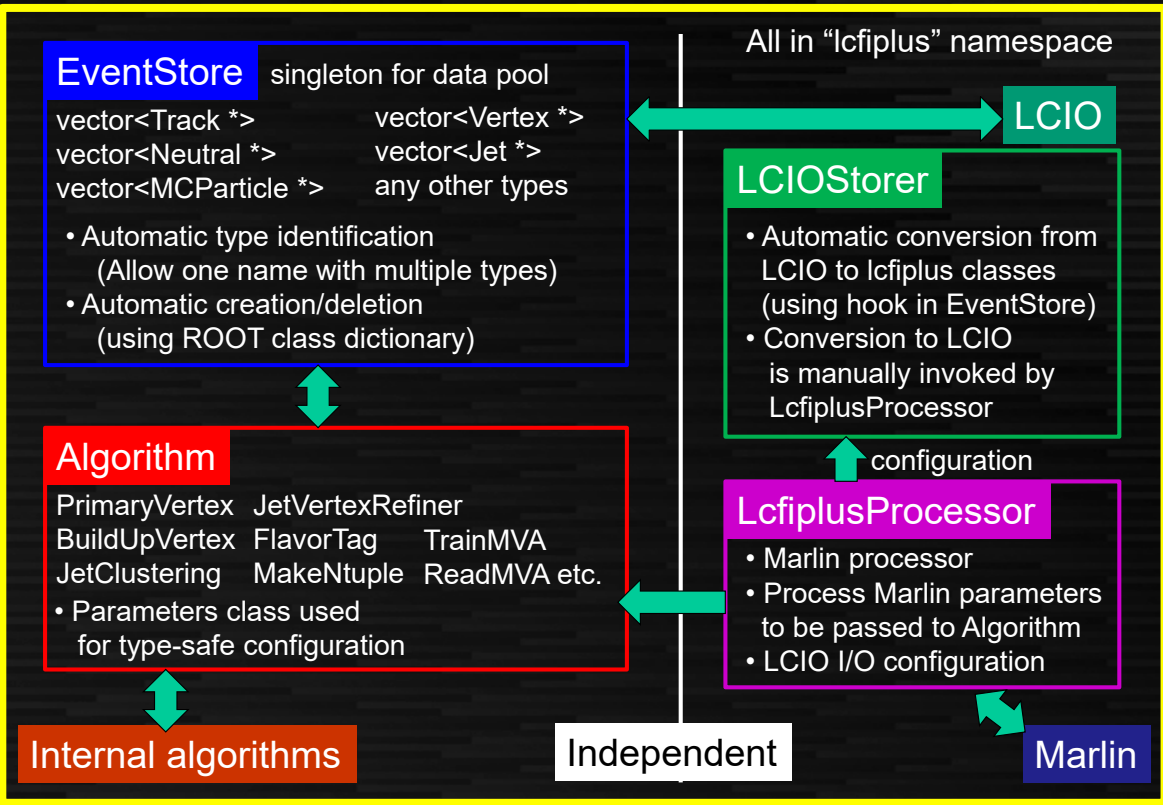
Currently mainly maintained
by R. Yonamine (KEK)

Record of a tutorial:

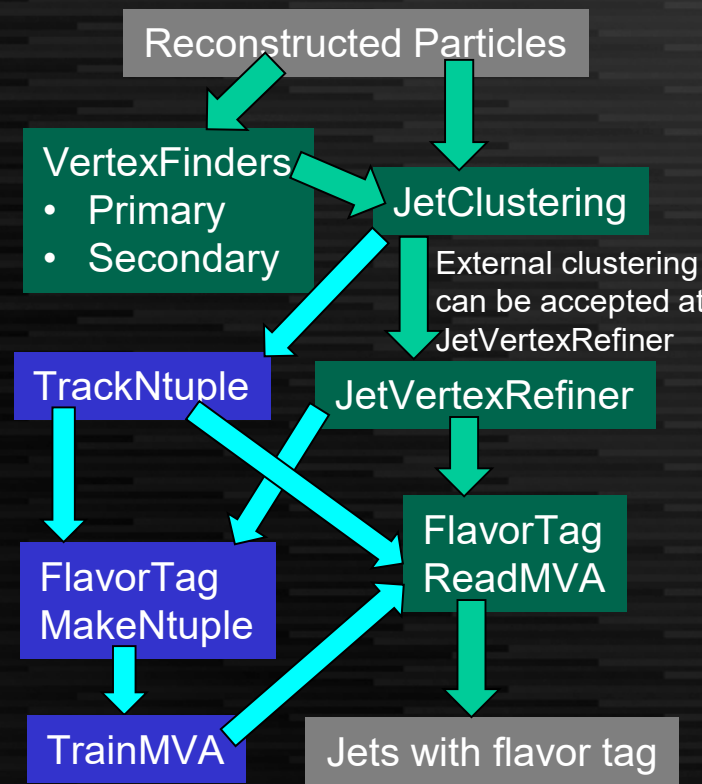
<https://agenda.linearcollider.org/event/9318/>

LCFIPlus framework

 Training chain
 Reconstruction chain



LCFIPlus algorithms



LCIO/Marlin implementation via independent adapters

Vertex finder in LCFIPlus

- PrimaryVertex

- Tear-down method

- Fit all tracks with beam constraint
- Remove a track with worst χ^2
- Repeat until all tracks within acceptable χ^2

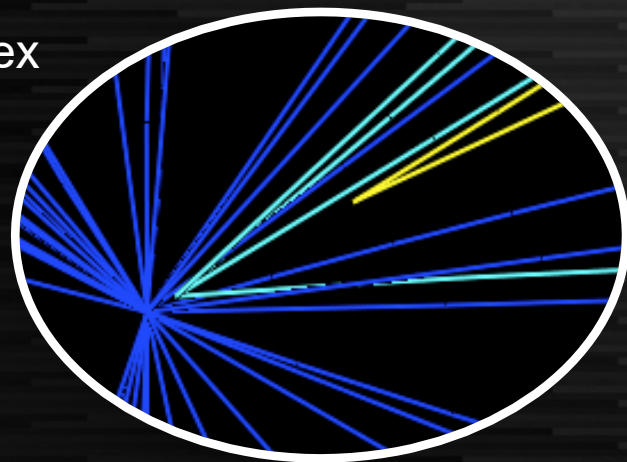
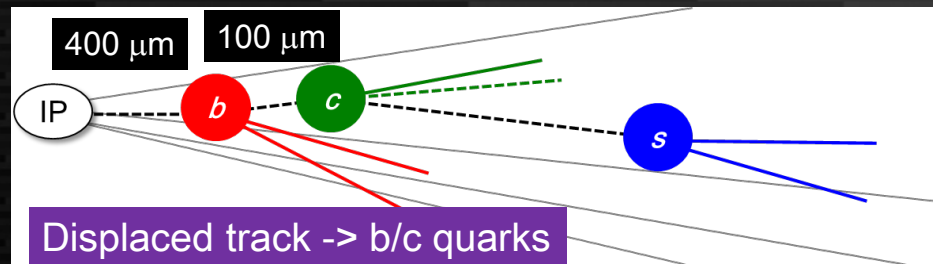
- BuildUpVertex

- Secondary vertex finder to be run after PrimaryVertex

- Build-up method

- Examine every track pair if it is consistent with a vertex (χ^2 etc)
- List “good” vertex candidates
- Try to attach more tracks to the candidates

- “Adaptive vertex fitter” recently added (optional)



Jet clustering / JetVertexRefiner in LCFIPlus

- Durham with beam rejection

$$y_{\text{beam}} = 2E^2\alpha^2(1-\cos\theta)/E_{\text{vis}}^2$$

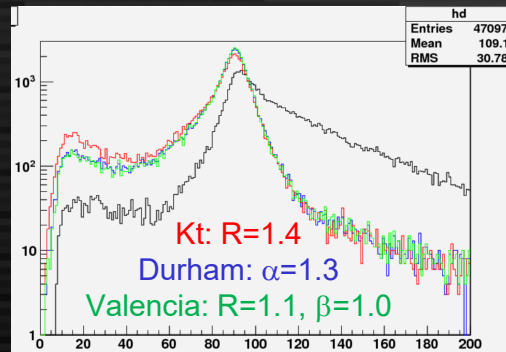
- Plain Durham (still available, of course)
- kT algorithm
 - No need to run it outside any more
- Valencia algorithm

- Intermediate algorithm of Durham and kT

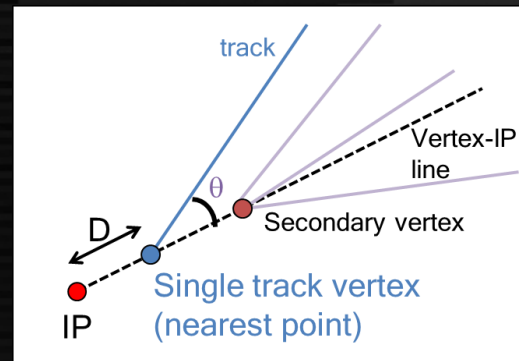
$$d_{ij} = \min(E_i^{2\beta}, E_j^{2\beta})(1 - \cos\theta_{ij})/R^2 \quad d_{iB} = p_T^{2\beta}$$

- JetVertexRefiner

- Assign vertices to jets if jets are created externally
- Looking for “single track vertices”
- Limit vertices/jet to two, by forced combining of existing vertices if # vtx > 3 (based on χ^2)
- “Bness” recently added



Invariant mass of Z
($Z_{VV} \rightarrow qq_{VV}$, 500 GeV, 2-jet clustering)



Flavor tagging in LCFIPlus

- Flavor tagging – BDTG in TMVA

- Training algorithms

- TrackNtuple and TrackProb.C
 - To prepare data files to be used in flavor tagging
 - MakeNtuple ← TrackProb files
 - To prepare training data
 - Train ← MakeNtuple files
 - Run TMVA to train the flavor tagging BDT

- FlavorTag/ReadMVA

- Evaluate RefinedJets with trained BDT data
 - Output “n”-likenesses, categories can be defined (default: b, c, other. Can add eg. bb.)

Flavor tagging variables (selectable in steering file)

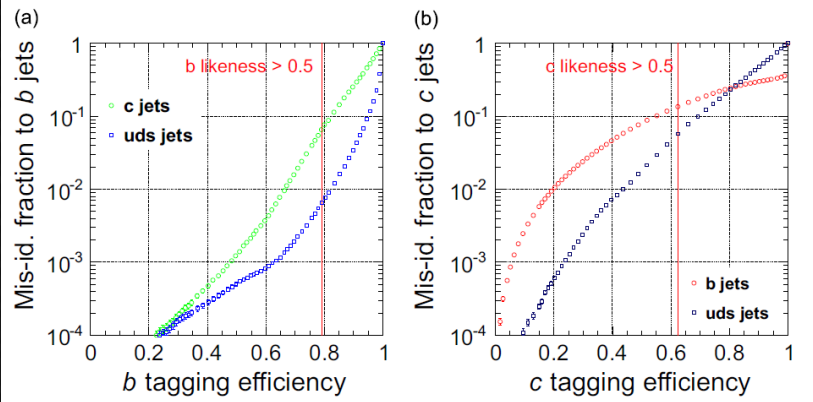
- Categories by # reco vtx / jet (0, 1, 1+1 single track, 2)
- Track variables
 - d0/z0 significance, pt, joint probability, b/c/q prob., # leptons, track mass
- Vertex variables (only 1+ vtx)
 - decay length, angle, mass, pt-corrected mass, momentum, multiplicity, probability, significance
- 2-vertex variables (> 1 vtx)
 - single vertex prob., length and significance of 1st-2nd vertex
- New variables possible

Flavor tagging: performance at ILD

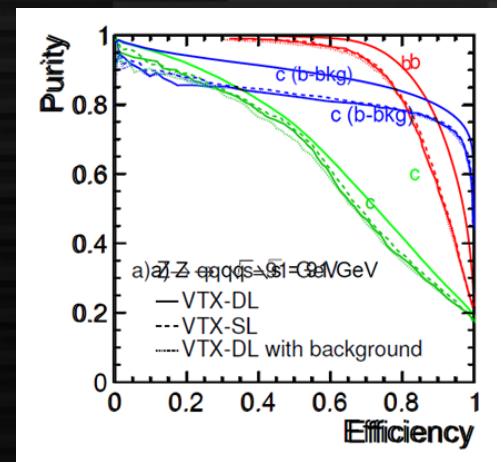
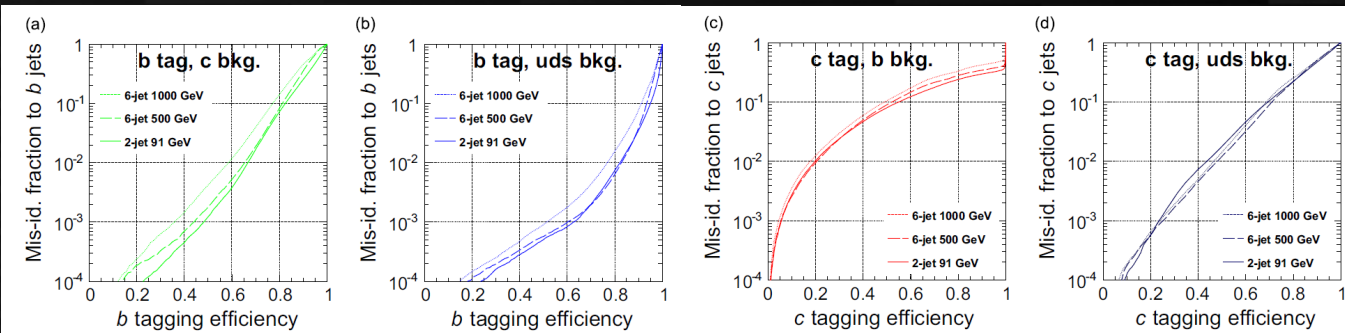
ILD DBD

$Z \rightarrow qq$, 91 GeV

Performance of vertex finder



(#vtx, #pseudo-vtx)	b jet (%)	c jet (%)	uds jet (%)
(0, 0)	21.3	59.3	98.1
(0, 1)	1.61	0.17	0.01
(1, 0)	39.7	39.8	1.80
(1, 1)	13.5	0.54	0.02
(2, 0)	23.8	0.19	0.04



Comparison with LCFIVertex

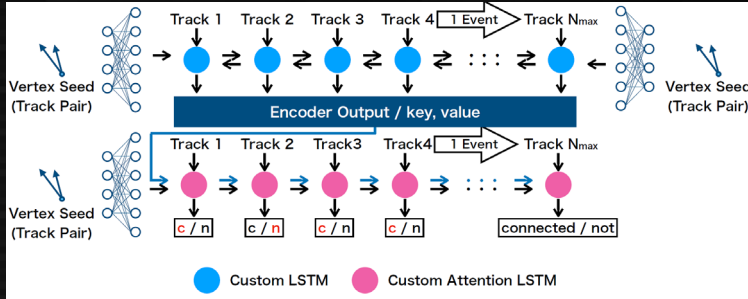
Design consideration of LCFIPlus

- Vertex finder
 - Fitting based – not easy to make equivalent with pure DNN
 - Focused on **high-purity vertices**,
throwing low-quality tracks/vertices away
 - Mandatory to be the input of jet clustering
 - Bit room for improvement with DNN
 - V0 rejection also important (for flavor tagging)
 - Dedicated V0 fitting considering momenta at decay point (not at IP)
- Flavor tagging
 - Input variables are designed as virtually Lorentz invariant
 - Robust to the change of jet energy
 - DNN method usually heavily depends on kinematics – energy dependence

Our trials on DNN flavor tagging – a glance

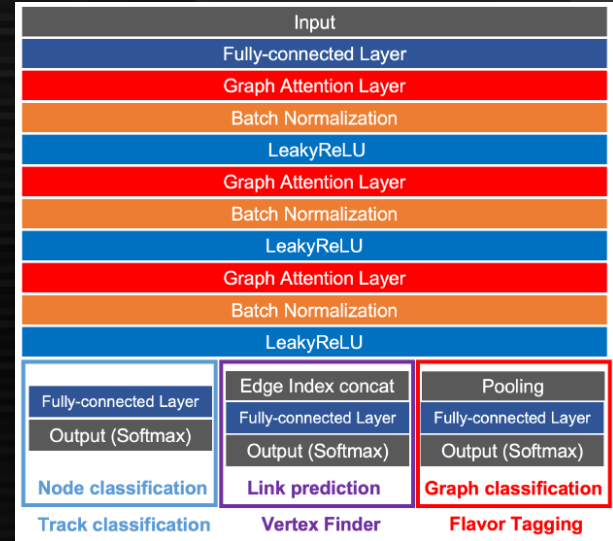
Vertex finder based on recurrent NN

- Modified LSTM with attention [NIMA 1047 \(2023\) 167836](#)



GNN-based flavor tagging without fitting

- Graph attention, 3 separate outputs



Vertex reconstruction eff. with DNN

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	2.2%	63.3%	68.4%	9.5%
– of same decay chain		62.3%	67.2%	
– of same parent		38.1%	36.2%	6.4%

Vertex reconstruction eff. with LCFIPlus for comparison

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	0.2%	57.9%	60.3%	0.5%
– of same decay chain		57.5%	59.9%	
– of same parent		34.0%	37.2%	0.3%

Tagging efficiency = 0.8	background	Mis-id fraction	
		LCFIPlus	GNN
<i>b</i> jet	<i>c</i> jet	0.073	0.021
	<i>uds</i> jet	0.007	0.015
<i>c</i> jet	<i>b</i> jet	0.22	0.40
	<i>uds</i> jet	0.24	0.14

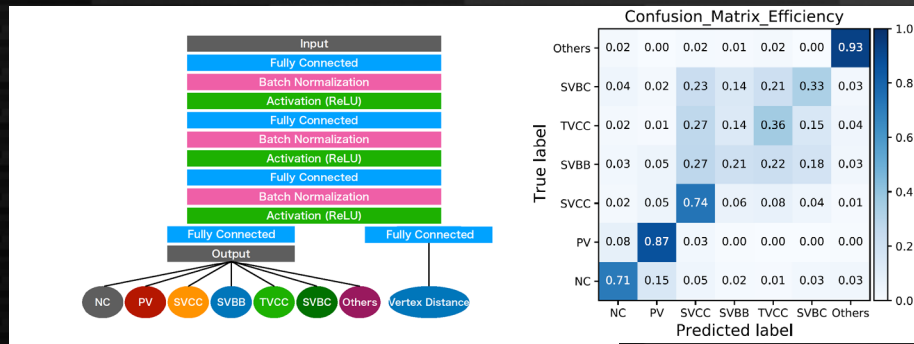
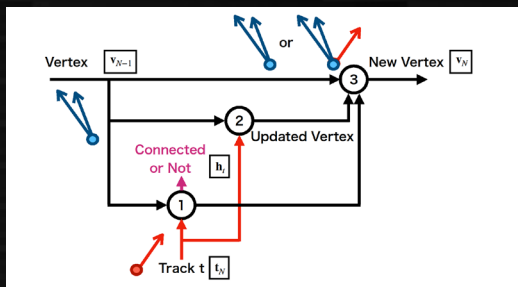
Partially better than LCFIPlus to be investigated

Summary

- LCFIPlus is a standard jet reconstruction tool since 2013
 - Vertex finder (based on tear down / build up)
 - Jet clustering (Durham, Valencia, Kt, beam rejection)
 - Flavor tagging (BDT, Lorentz invariant input variables)
- Reasonable performance
 - b-tag: 80% eff., 10% c / 1% uds acceptance
 - c-tag: 50% eff., 10% b / 2% uds acceptance
- Improvements with DNN?
- Adding another variables like particle ID, picosec timing, neutral particles etc. may open further possibility

Vertex finder with DNN (modified LSTM) arXiv:2101.11906 Under review at NIMA

- Replacing build-up vertex with DNN
 - Track-pair classification with NN
 - Track association with LSTM-based RNN with attention
 - Similar perf. to LCFIPlus obtained w/o vertex fitting



Vertex reconstruction eff. with DNN

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	2.2%	63.3%	68.4%	9.5%
- of same decay chain		62.3%	67.2%	
- of same parent		38.1%	36.2%	6.4%

Vertex reconstruction eff. with LCFIPlus for comparison

Criteria / True label	Primary	Bottom	Charm	Others
All tracks	307 657	187 283	180 143	42 888
In secondary vertex	0.2%	57.9%	60.3%	0.5%
- of same decay chain		57.5%	59.9%	
- of same parent		34.0%	37.2%	0.3%

