Strange Jet Tagging

Yuichiro Nakai

T. D. Lee Institute & Shanghai Jiao Tong U.

Based on YN, D. Shih and S. Thomas, in preparation.

Particle Physics in Computing Frontier, December 2019

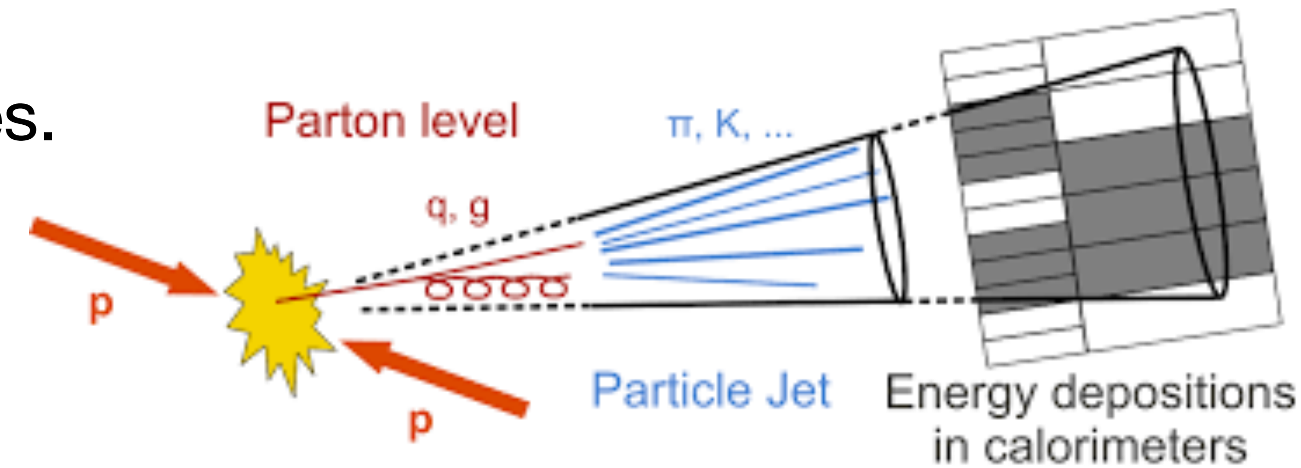
Jets at the LHC

Jet : collimated bunch of hadrons as the signatures of quarks and gluons produced in high-energy collisions

✓ QCD partons are never observed isolated due to confinement.

✓ They give cascades of radiation (parton shower) by QCD processes.

✓ Hadrons are formed at $\sim \Lambda_{\mathrm{QCD}}$

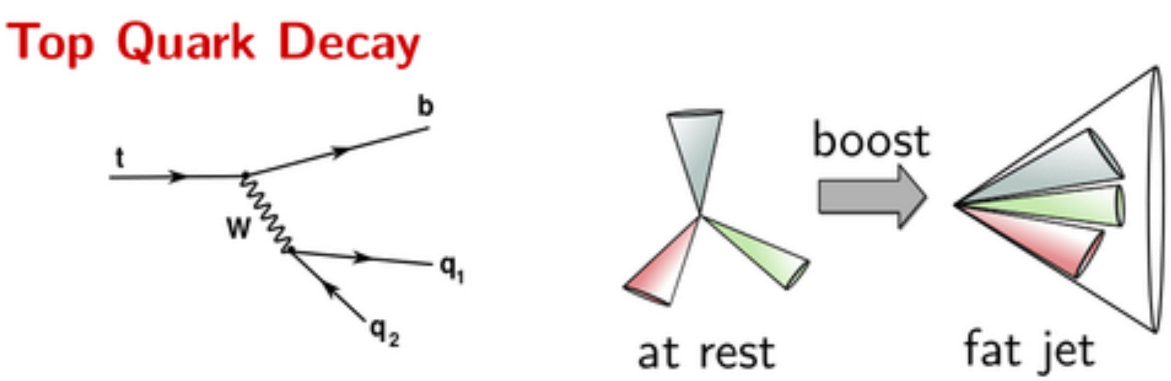


Understanding jets is a key ingredient of physics measurements and new physics searches at the LHC.

What initial parton produces a jet ?

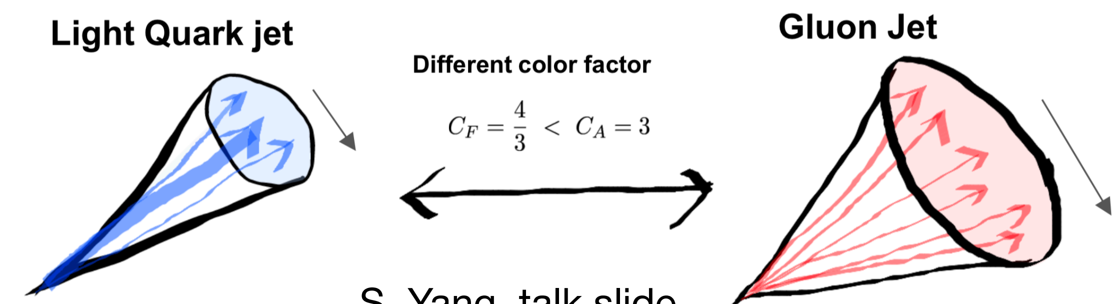
Quark and Gluon Tagging

☑ Top quark



Jet mass, N-subjettiness, …

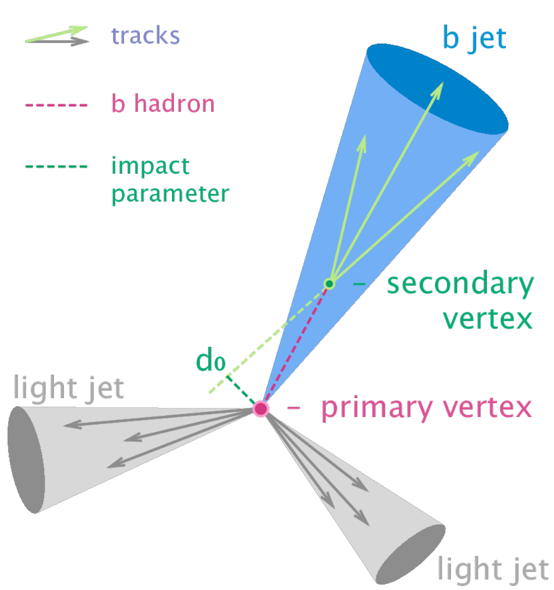
☑ Gluon



S. Yang, talk slide

More constituents with more uniform energy fragmentation and wider.

☑ Bottom/Charm



Look for a displaced (secondary) vertex.

Wikipedia

☑ Up-type vs Down-type

p_T-weighted jet charge

$$Q_\kappa^i = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa$$

The last missing piece :

Strange quark tagging?

Applications of Strange Tagging

- **Strange Yukawa**

The SM predictions of the Higgs couplings to heavy gauge bosons and fermions, $2m^2_{W,Z}/v$ and m_f/v, have been confirmed for the W and Z bosons and for the third-generation fermions.

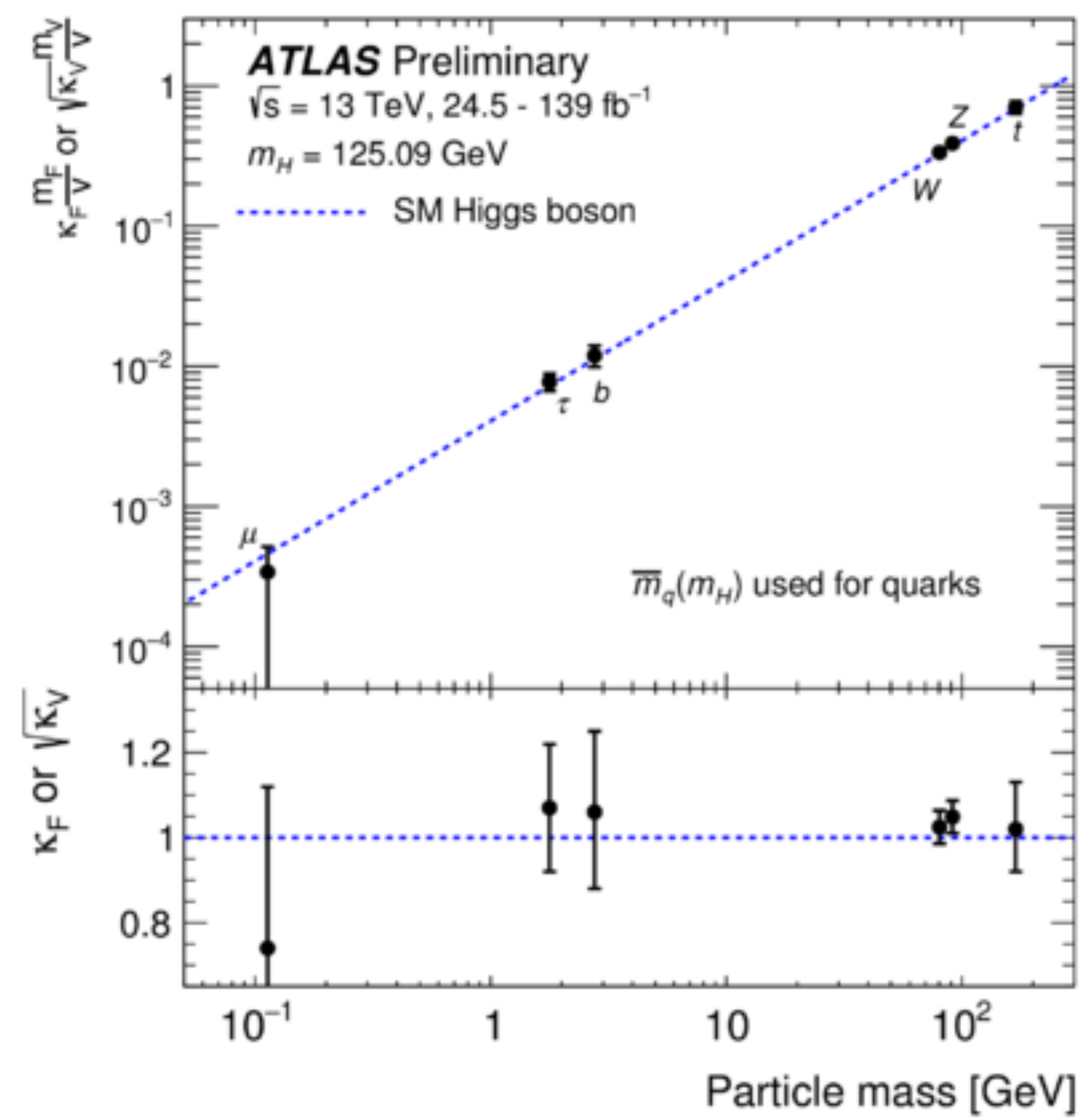
However…

No direct measurements for the first two generation fermions.

We can easily come up with models where these couplings can deviate significantly from the SM predictions.

Probing the Higgs couplings to light fermions is very important !

Strange tagging is essential.



Applications of Strange Tagging

• CKM mixings

The CKM matrix elements are fundamental parameters of the SM and their precise determination is important.

However… **The values for $|V_{cs}|$ and $|V_{cd}|$ are not measured very well.**

Because the charm quark mass is too heavy to be considered light but not heavy enough to treat in the heavy quark limit.

One process to probe $|V_{cs}|$ is through **the semileptonic decays $D \to Ke\nu$**.

Our best effort is to use lattice QCD :

$$V_{cs} = 0.98 \pm 0.01_{exp} \pm 0.10_{th}$$

The experimental error is small but the theoretical error is huge !

W boson decay $W \to cs$ gives the most direct measurement of $|V_{cs}|$ if strange tagging is possible.

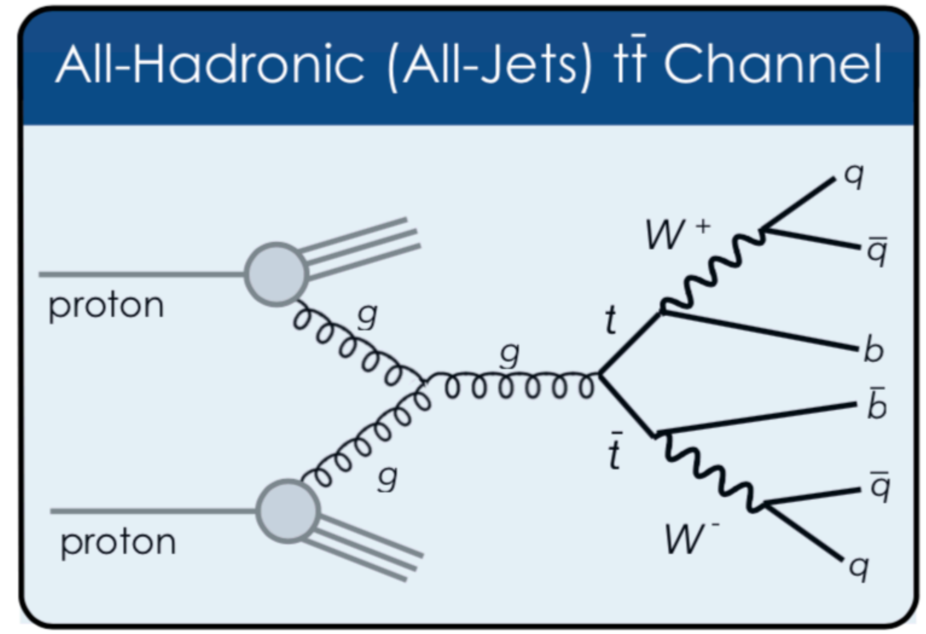
Applications of Strange Tagging

- **Top quark reconstruction**

  ✓ **All-hadronic channel**

  🙂 Full event reconstruction is possible.

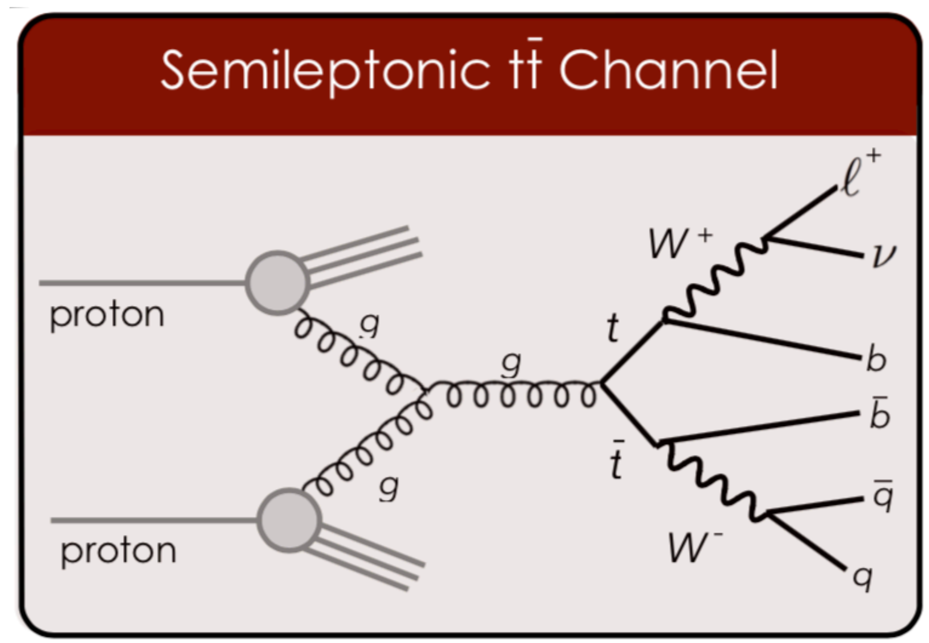
  😱 Jet combinatorics and large multi-jet background are problematic.



  ✓ **Semileptonic channel**

  😉 Leptonic top identifies event and hadronic top can be reconstructed.

  😰 Jet combinatorics and multi-jet background are still issues.



T. McCarthy, talk slide

Which jets are W → cs, us, cd, ud decay products?

Identification of strange jet may give some help.

Applications of Strange Tagging

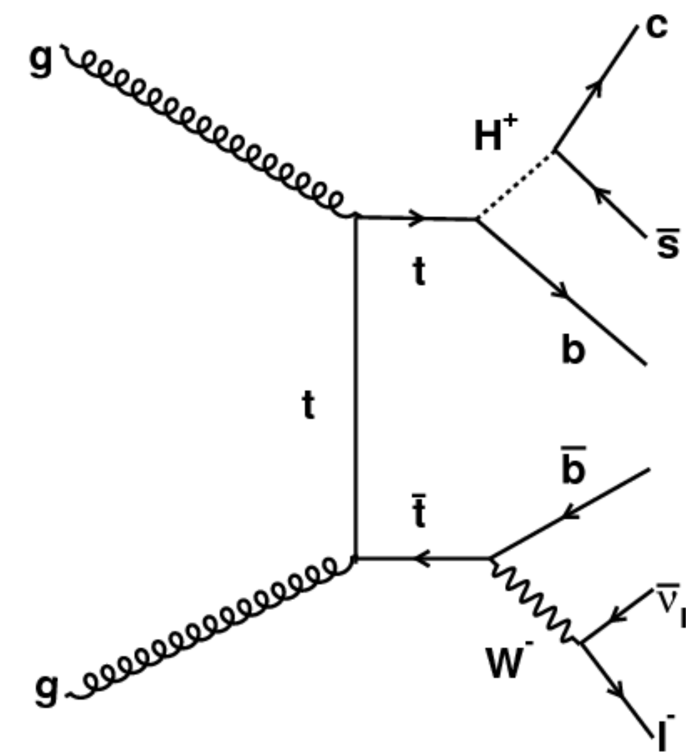
- **Light charged Higgs search**

Production : $t\bar{t} \rightarrow W^{\pm}bH^{\mp}\bar{b}$

Decay : $H^{+} \rightarrow c\bar{s}$

😰 The same issue as top quark reconstruction is applied.

😱 We do not know the charged Higgs mass!



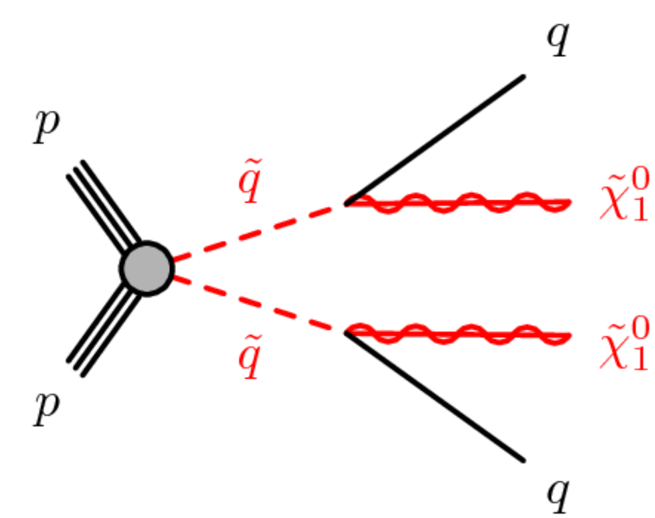
CMS

- **Squark search**

Identification of strange jet can …

✓ reduce the background $Z(\rightarrow \nu\nu) +$ jets

✓ identify squark flavor after the discovery



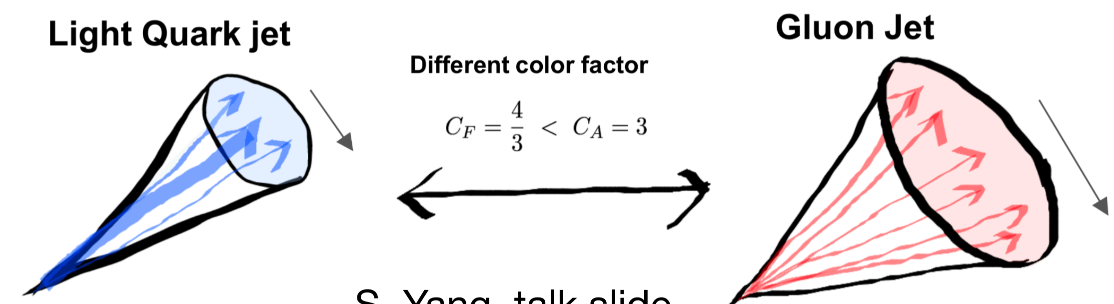
CMS

<u>Strange tagging may be important !</u>

Tagging Strategy

- **Strange vs Gluon**

  We can expect the same thing as quark/gluon discrimination.



S. Yang, talk slide

More constituents with more uniform energy fragmentation and wider.

- **Strange vs Up**

  We can expect the same thing as up/down discrimination.

p_T-weighted jet charge

$$Q_\kappa^i = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{j \in \text{jet}} Q_j (p_T^j)^\kappa$$

- **Strange vs Down**

  Possible ??

  Both are quarks with the same charge.

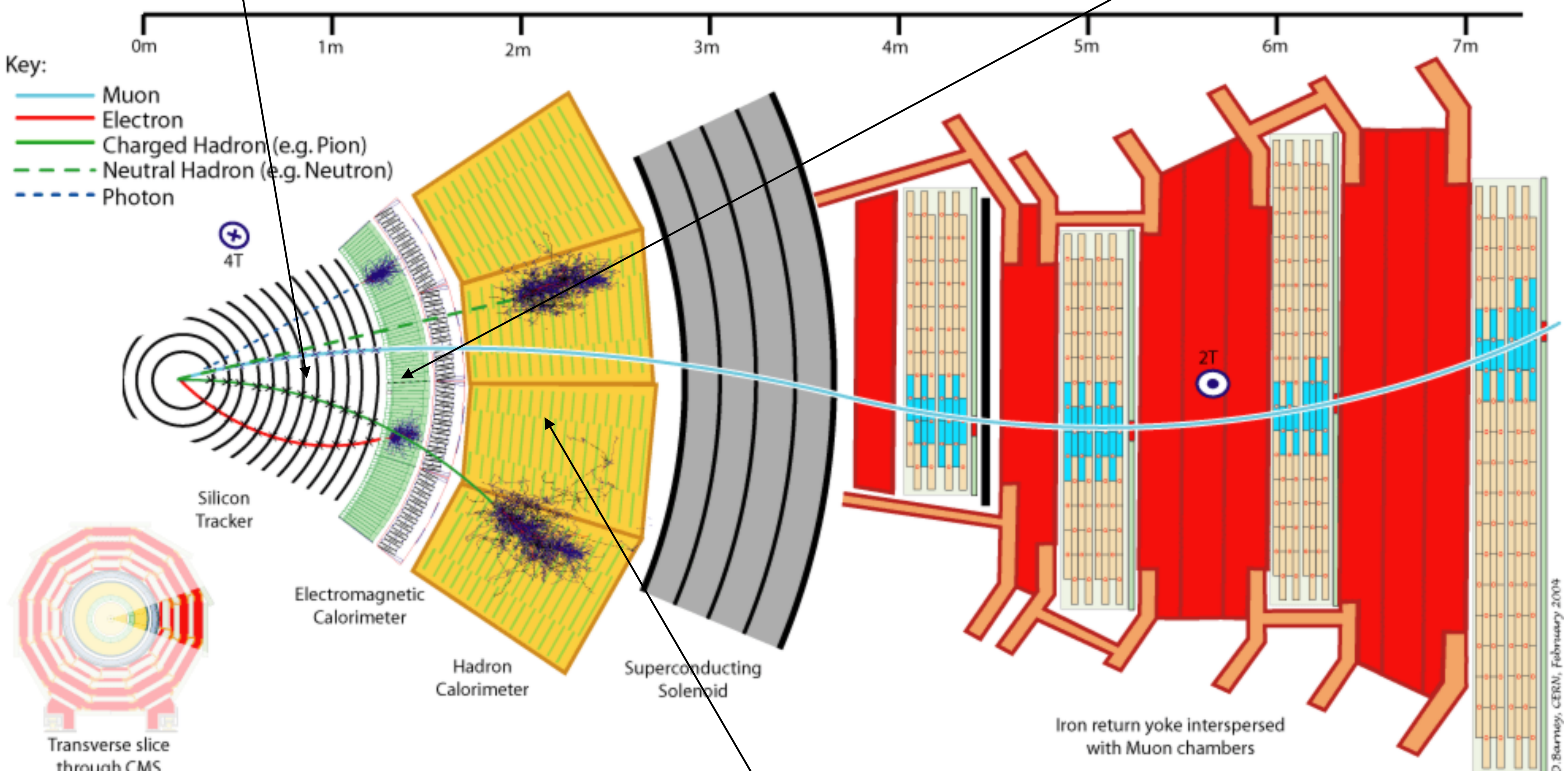
Main theme of this talk

Tagging Strategy

CMS experiment at the LHC

Tracker : trajectories of charged particles

ECAL : energy of electrons and photons



HCAL : energy deposits of hadrons

Tagging Strategy

After hadronization, strange quarks form Kaons :

$$K^- = s\bar{u}, \quad K^+ = \bar{s}u, \quad K_L \approx \frac{s\bar{d} - d\bar{s}}{\sqrt{2}}, \quad K_S \approx \frac{s\bar{d} + d\bar{s}}{\sqrt{2}}$$

No difference

$\boxed{K_L,\ K^\pm}$ $\gamma c\tau \sim 3$ m

No decay inside the detectors

$\boxed{K_S}$ $\gamma c\tau \sim 3$ cm

Decay inside the detectors

$K_S \to \pi^+\pi^-(\sim 70\%),\ \pi^0\pi^0(\sim 30\%)$

Detector responses to hadrons : Down jets

	K_L	K_S	$K^\pm$	π^0	$\pi^\pm$
HN	○	△			
ECAL		△		○	
Tracker		△	○		○

Strange jets

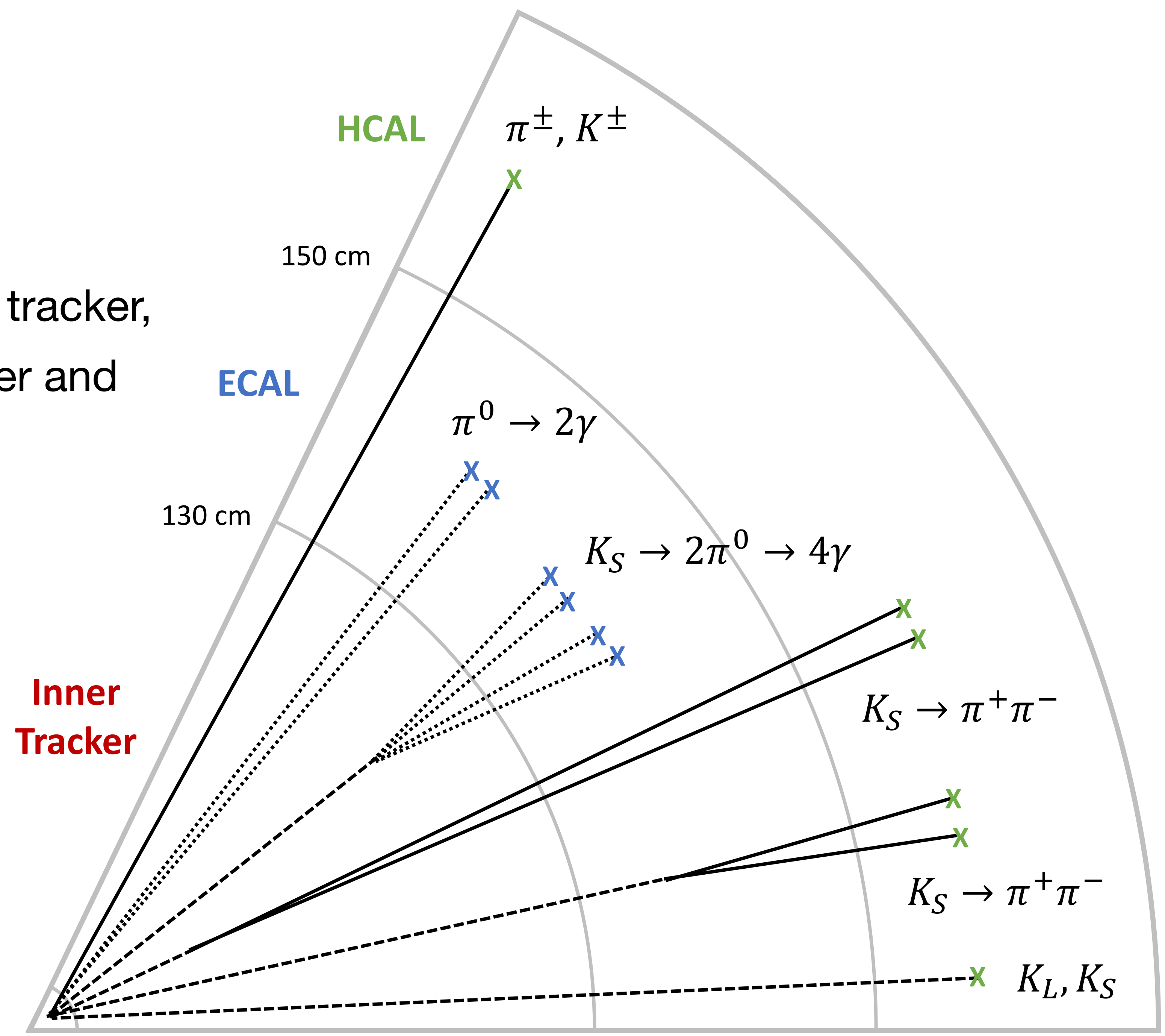
Hadronic Neutral (HN) = HCAL - Tracker

K-long (and K-short) can be used for tagging !

Tagging Strategy

K-short behaves very differently in detectors depending on decay length and decay mode.

- ✓ $K_S \to \pi^0\pi^0 \to \gamma\gamma\gamma\gamma$ before ECAL, energy deposit in ECAL
- ✓ $K_S \to \pi^+\pi^-$ before or within the inner tracker, momenta measured in the inner tracker and energy deposit in HCAL
- ✓ $K_S \to \pi^+\pi^-$ outside the inner tracker before reaching HCAL, energy deposit in HCAL
- ✓ No decay before reaching HCAL, energy deposit in HCAL



Tagging Strategy

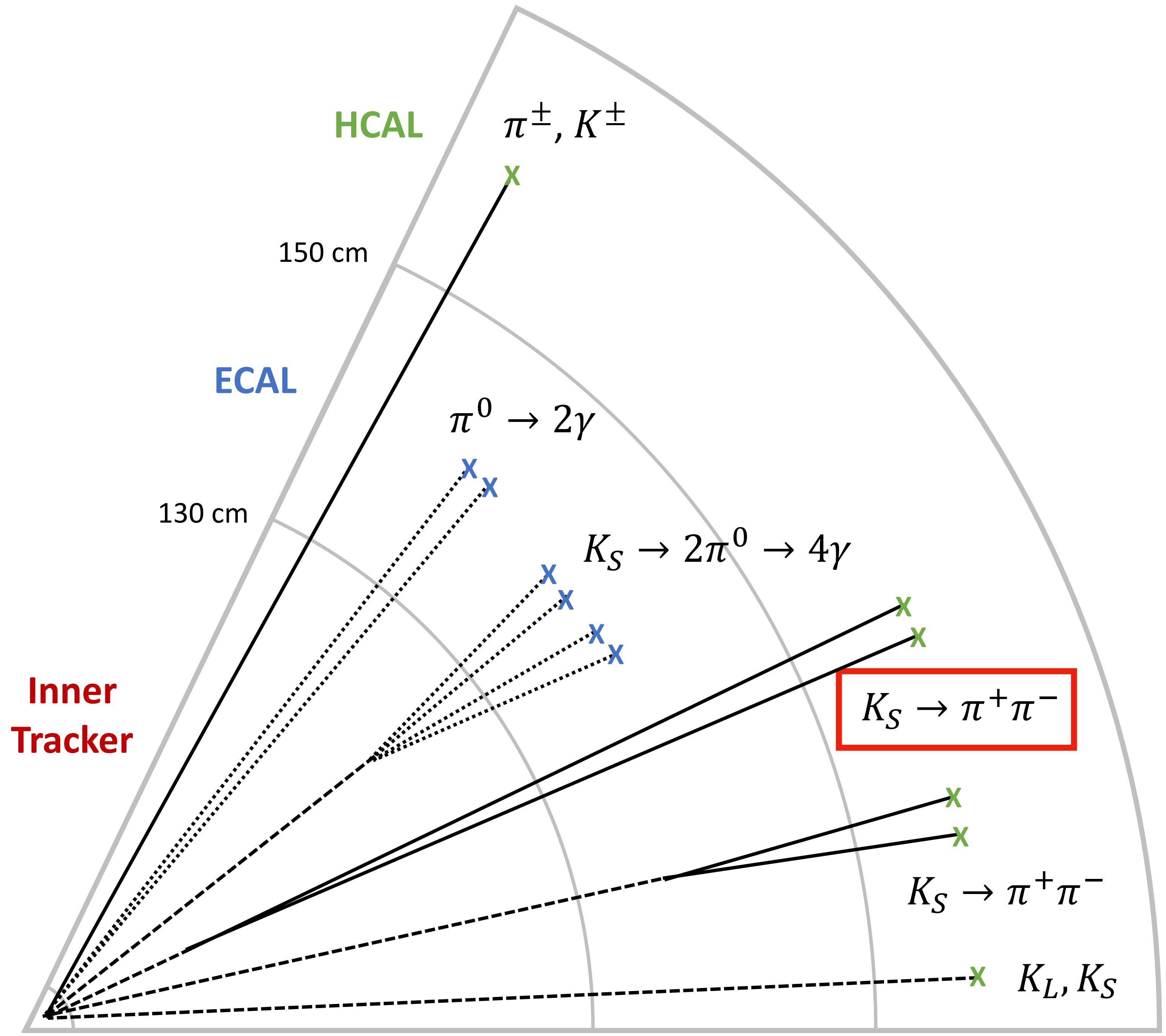
$K_S \to \pi^+\pi^-$ **inside Tracker leaves charge track pairs from the secondary vertex.**

We can identify such K-shorts by reconstructing the invariant mass !

CMS (2010), ATLAS (2011)

✓ $K_S \to \pi^+\pi^-$ outside the inner tracker before reaching HCAL, energy deposit in HCAL

We assume the efficiency of the reconstruction is 40%.



Tagging Strategy

1. K-long (and K-short) can be used for tagging !

No difference

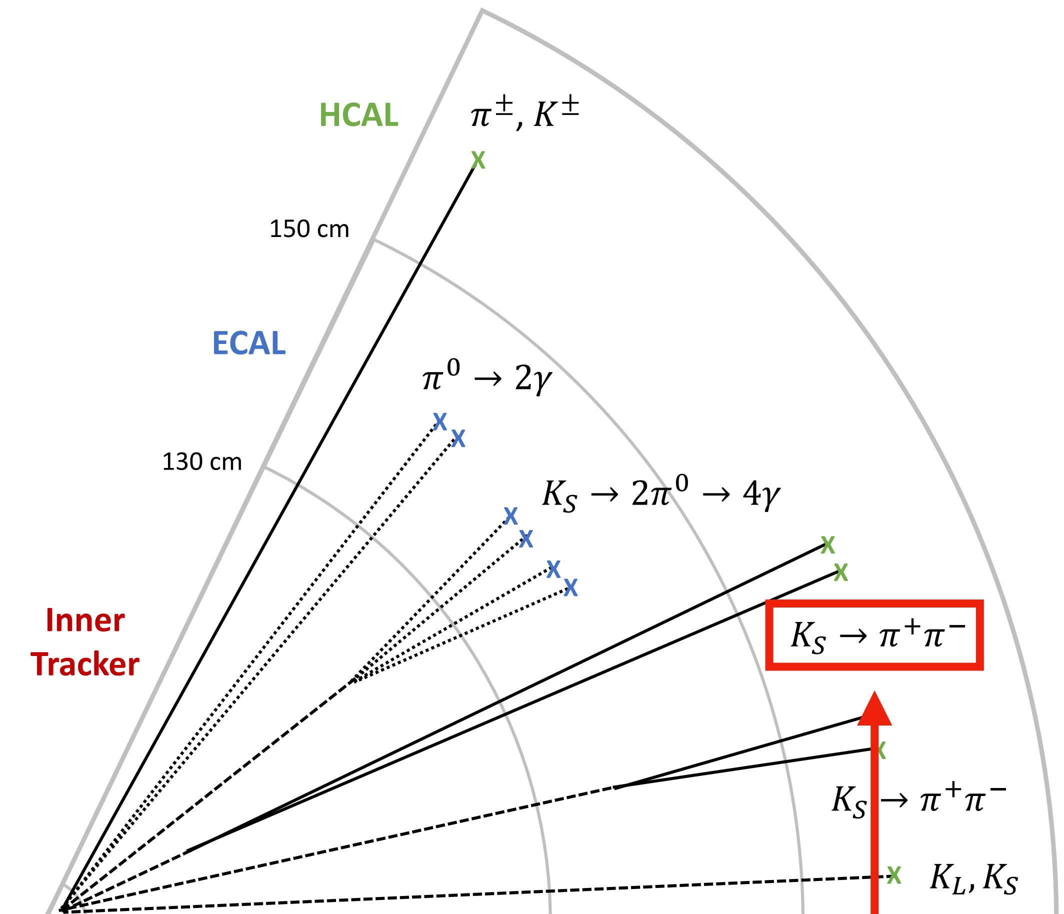
Detector responses to hadrons :

Down jets

	K_L	K_S	$K^{\pm}$	π^0	$\pi^{\pm}$
HN	○	△			
ECAL		△		○	
Tracker		△	○		○

Strange jets

Hadronic Neutral (HN) = HCAL - Tracker



2. We can identify such K-shorts by reconstructing the invariant mass !

Jet Samples

Generate strange/down jet samples by using MadGraph, PYTHIA and Delphes.

100000 events for each case of :

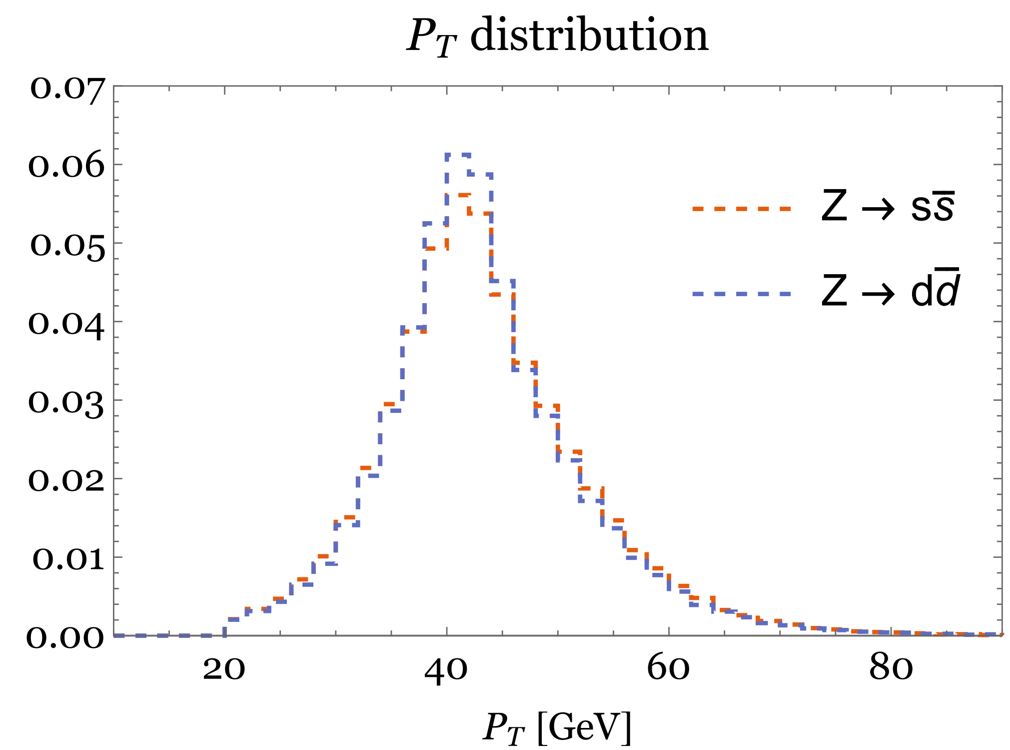
$$Z \to s\bar{s} \quad (p_T > 20\ \mathrm{GeV})$$
$$Z \to d\bar{d} \quad (p_T > 20\ \mathrm{GeV})$$

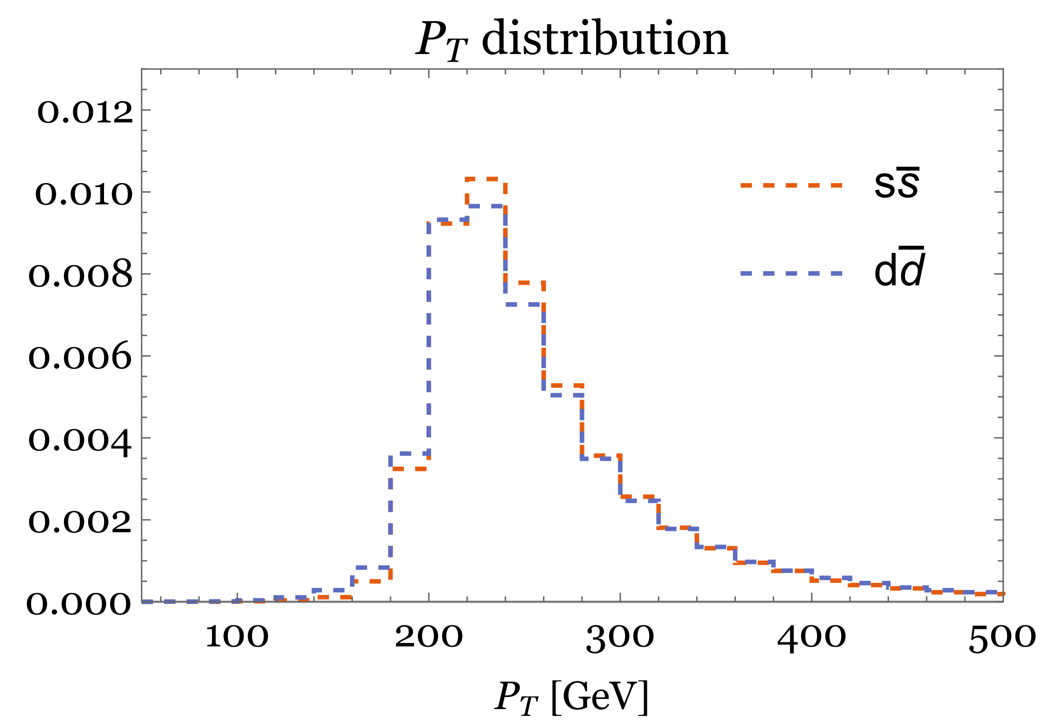
$$s\bar{s} \quad (p_T > 200\ \mathrm{GeV})$$
$$d\bar{d} \quad (p_T > 200\ \mathrm{GeV})$$

$$(|\eta| < 0.05)$$

Initial parton is required to be inside the leading jet : $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$

Herwig gives the similar results.





Ingredients of strange/down jets

Strange jets contain more energetic Kaons than down jets?

Analyze the pT fraction of a detector-stable particle averaged over jet samples.

✓ The detector stable particles: K_L, K^+, π^+, γ, e^-, ν_e, μ^-, ν_μ, p, n

and the corresponding antiparticles

✓ Sometimes long-lived and decay outside ECAL:

K_S, Λ(uds), Σ^+(uus), Σ^-(dds), Ξ^0(uss), Ξ^-(dss)

★ These detector stable and long-lived particles should be all generated before or inside the ECAL.

The pT fraction averaged over our jet samples : $$\epsilon_p \equiv \frac{1}{N_{\mathbf{jets}}} \sum_{j \in \mathbf{jets}} \frac{p^p_{Tj}}{\sum_p p^p_{Tj}}$$

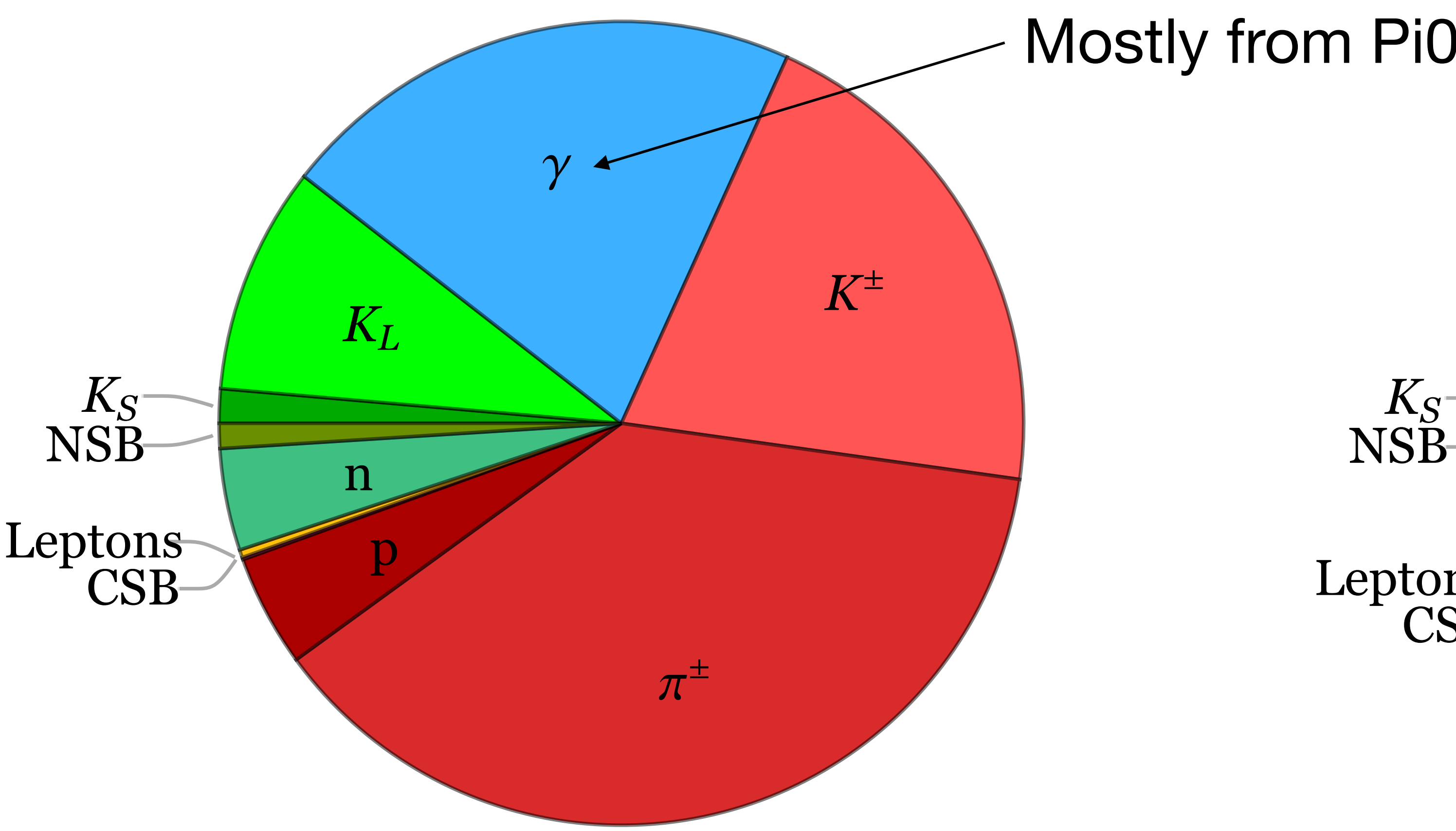
N_{jets}(= 10000): the number of jets, p: all the stable particles

Ingredients of strange/down jets

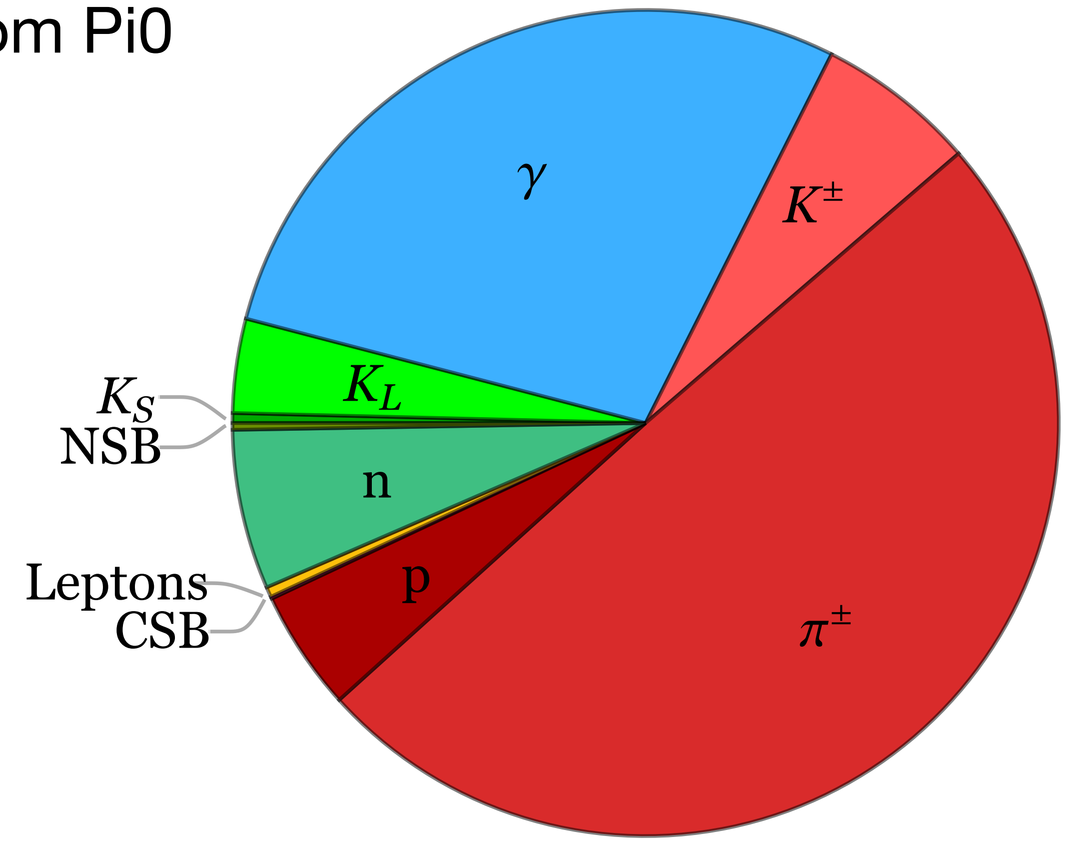
Strange jets contain more energetic Kaons than down jets.

The pT fraction of a detector-stable particle averaged over jet samples :

$Z \to s\bar{s} \quad (p_T > 20\ \mathrm{GeV})$



$Z \to d\bar{d} \quad (p_T > 20\ \mathrm{GeV})$



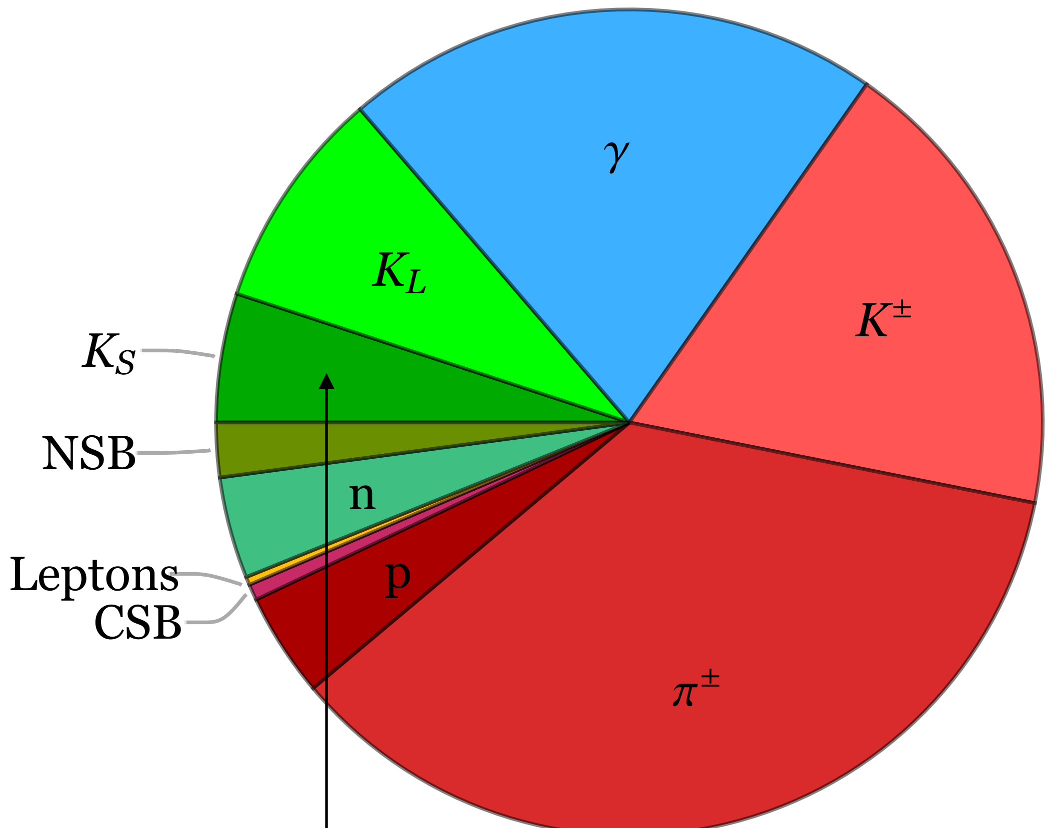
NSB: neutral strange baryons, CSB: charged strange baryons

Ingredients of strange/down jets

Strange jets contain more energetic Kaons than down jets.

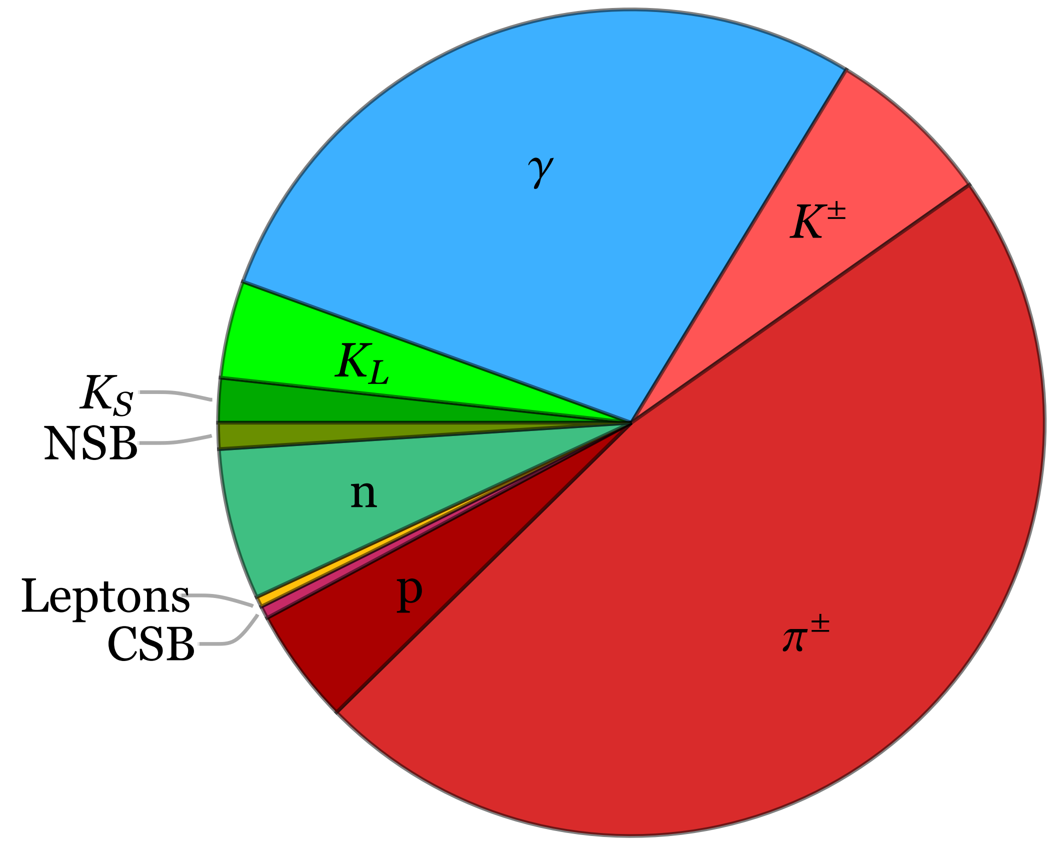
The pT fraction of a detector-stable particle averaged over jet samples :

Strange $P_T > 200$ GeV



More K-shorts due to boost factor

Down $P_T > 200$ GeV



NSB: neutral strange baryons, CSB: charged strange baryons

Various taggers

The list of our taggers and their inputs :

Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	π^0 , K_L , $K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	H_N , E , T	
BDT4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	H_N , E , T	13×13 Jet Image
CNN4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

Truth-level classifier

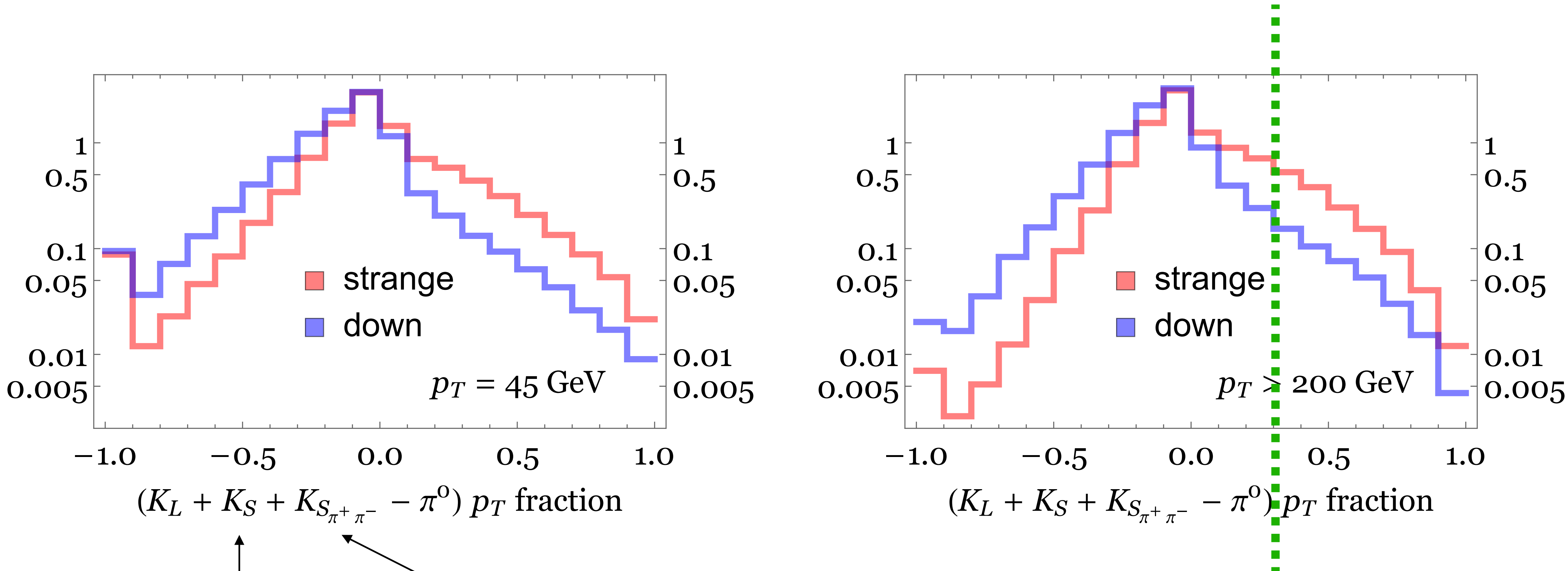
The list of our taggers and their inputs :

Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	π^0 , K_L , $K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	H_N , E , T	
BDT4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	H_N , E , T	13×13 Jet Image
CNN4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

Truth-level classifier

- **Truth Cut1** Input variable : $-\pi^0 + K_L + K_S + K_{S_{\pi+\pi-}}$

★ Use information before going through Delphes



Put a cut in distribution

Long-lived K-short which hits HCAL

Reconstructable K-short decaying to two charged pions

Truth-level classifier

- **Truth BDT3**

Inputs have 3 dimensions: K_L , K_S , π^0 p_T

Long-lived K-short which hits HCAL + Reconstructable K-short decaying to two charged pions

★ Use information before going through Delphes

★ Use **Boosted Decision Tree (BDT)** for classification.

➡ **Approximately set the maximal performance we can achieve.**

Cut-Based Tagging

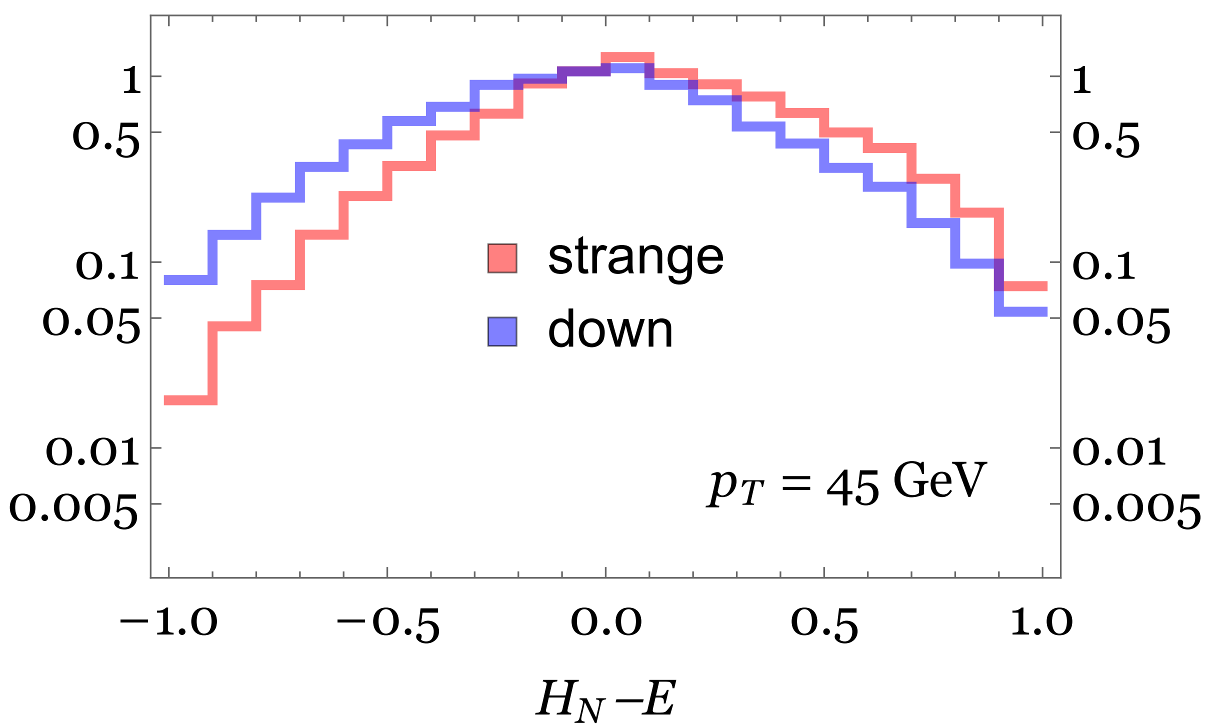
The list of our taggers and their inputs :

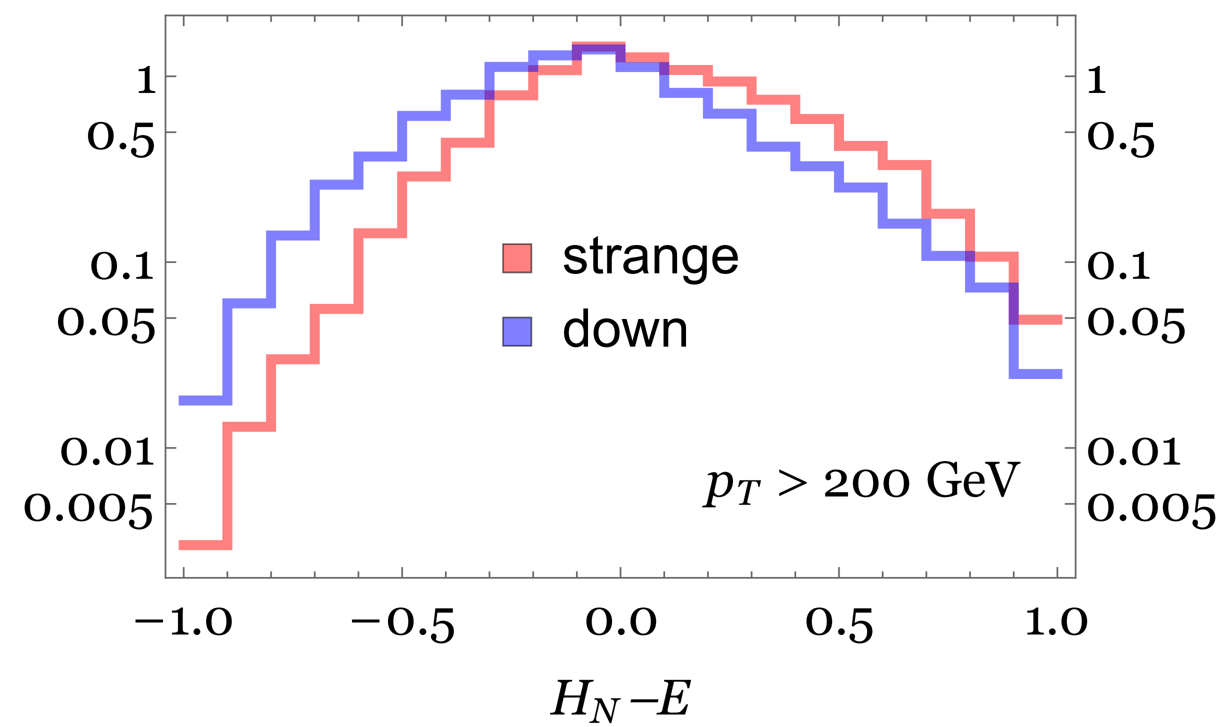
Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	π^0 , K_L , $K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	H_N , E , T	
BDT4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	H_N , E , T	13×13 Jet Image
CNN4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

Cut-Based Tagging

- **Cut1** Input variable : H_N–E

H_N : jet neutral hadronic energy fraction, E : electromagnetic energy fraction, T : track momentum fraction





- **Cut1+** Input variable : H_N –E +$KS_{\pi+\pi-}$

Cut-Based Tagging

Classify each jet into strange jet (signal) or down jet (background).

Put a cut in distribution.

Measures to estimate efficiency and accuracy of taggers

$$\varepsilon_S = \frac{\text{(Correctly classified into signals)}}{\text{(Total number of signal jets)}}$$

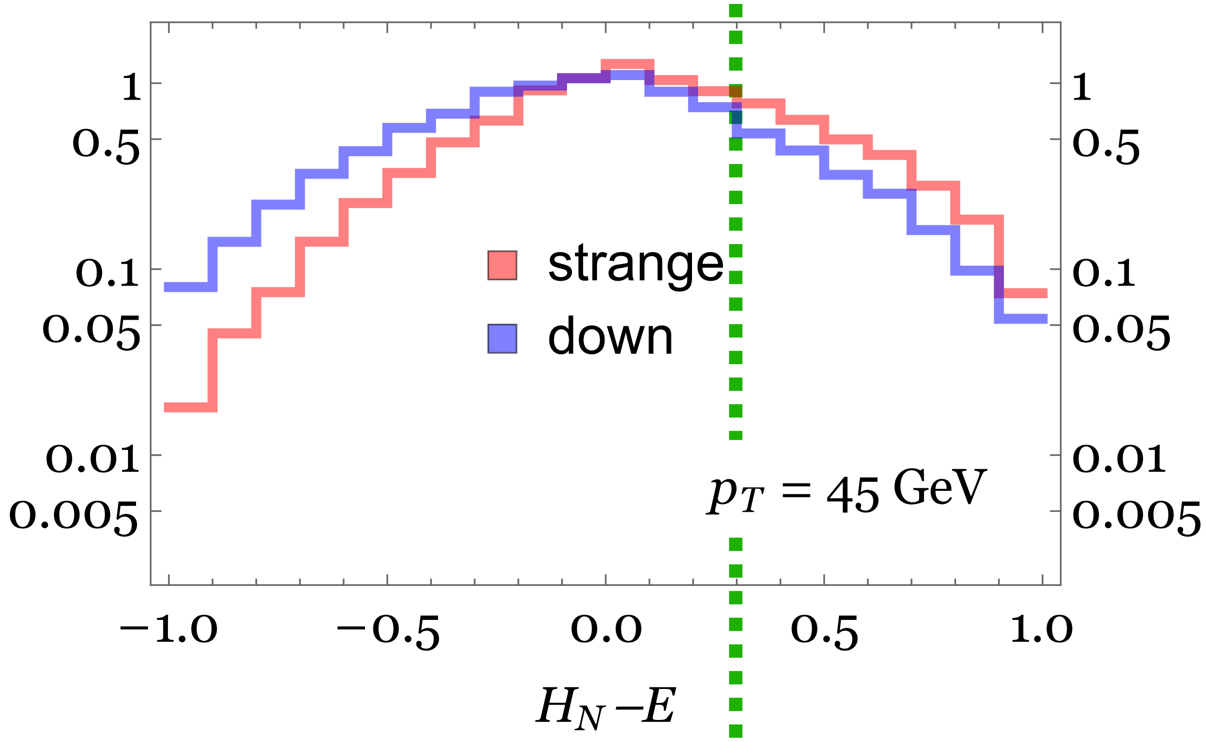
$$\varepsilon_B = \frac{\text{(Misclassified into signals)}}{\text{(Total number of backgrounds)}}$$

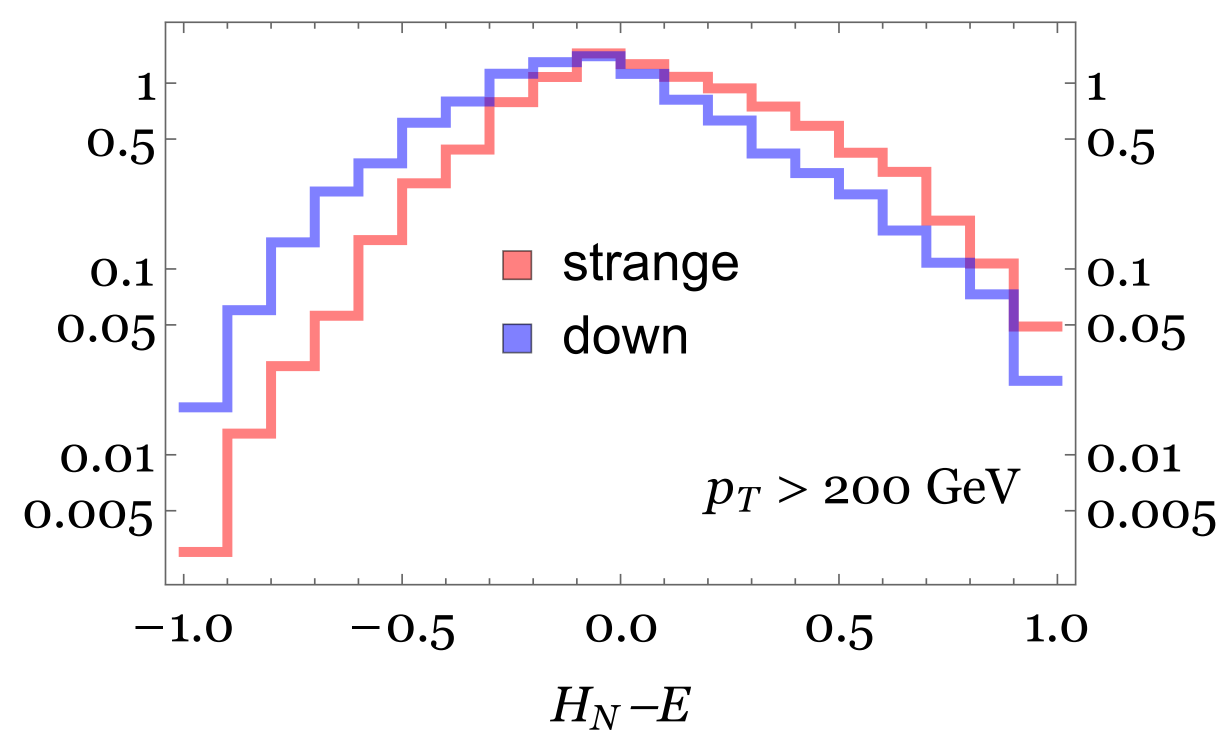
Larger ε_S

Larger ε_B

Smaller ε_S

Smaller ε_B





BDT

The list of our taggers and their inputs :

Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	$\pi^0\,,\ K_L\,,\ K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	$H_N\,,\ E\,,\ T$	
BDT4	Delphes	$H_N\,,\ E\,,\ T\,,\ K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	$H_N\,,\ E\,,\ T$	13×13 Jet Image
CNN4	Delphes	$H_N\,,\ E\,,\ T\,,\ K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

CNN

The list of our taggers and their inputs :

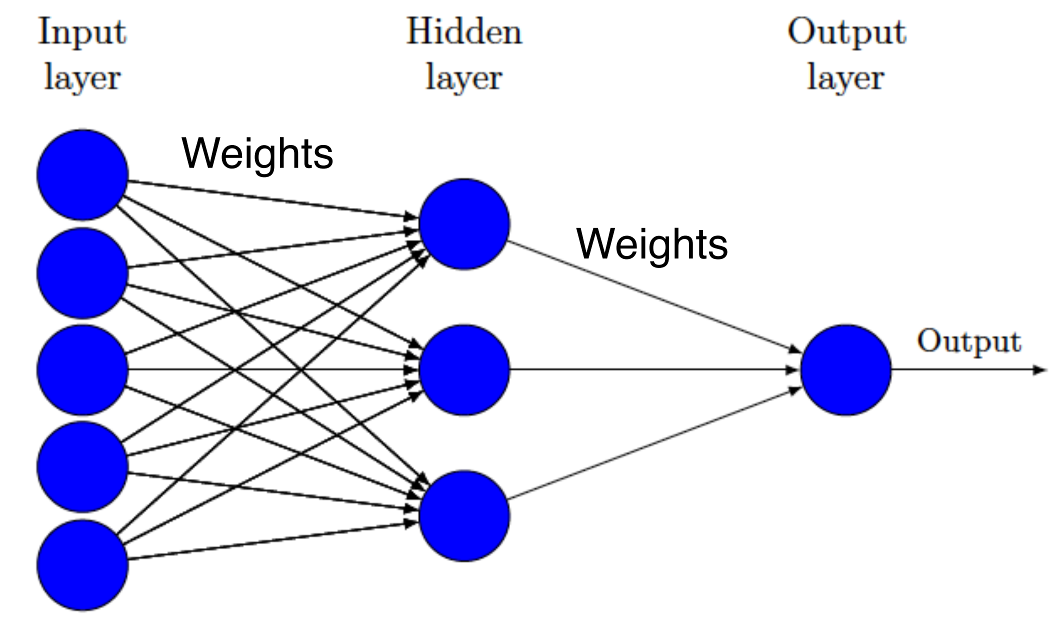
Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	$\pi^0 \, , \, K_L \, , \, K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	$H_N \, , \, E \, , \, T$	
BDT4	Delphes	$H_N \, , \, E \, , \, T \, , \, K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	$H_N \, , \, E \, , \, T$	13×13 Jet Image
CNN4	Delphes	$H_N \, , \, E \, , \, T \, , \, K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

Neural Networks

✓ Powerful machine learning-based techniques used to solve many real-world problems

✓ Modeled loosely after the human brain and designed to recognize patterns

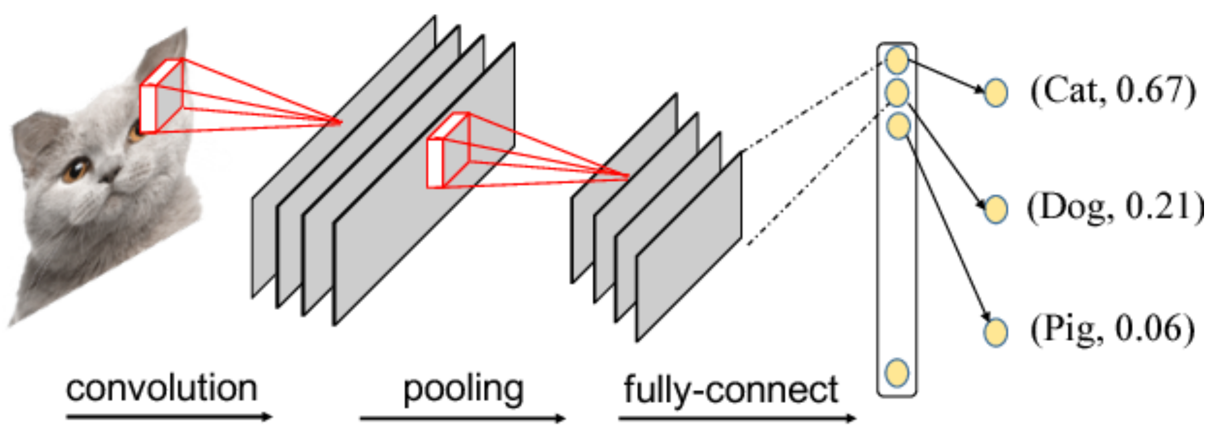
✓ Containing weights between neurons that are tuned by learning from data



Convolutional Neural Network (CNN)

✓ Show high performance for image recognitions

✓ Maintain the spacial information of images



arXiv:1712.01670

Apply a convolution operation to the input, passing the result to the next layer

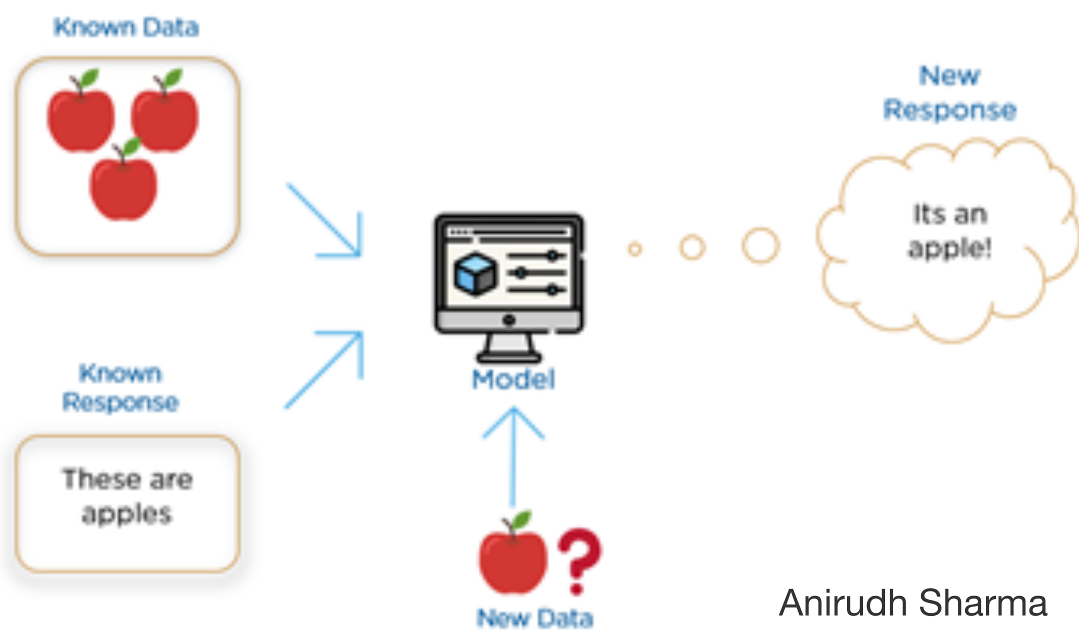
Reduce the image size

Supervised or Unsupervised

Machine learning algorithms can be classified into:

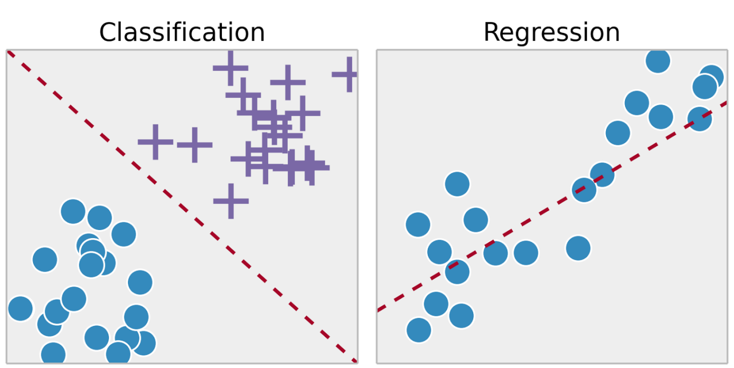
Supervised learning

Learn from labeled data



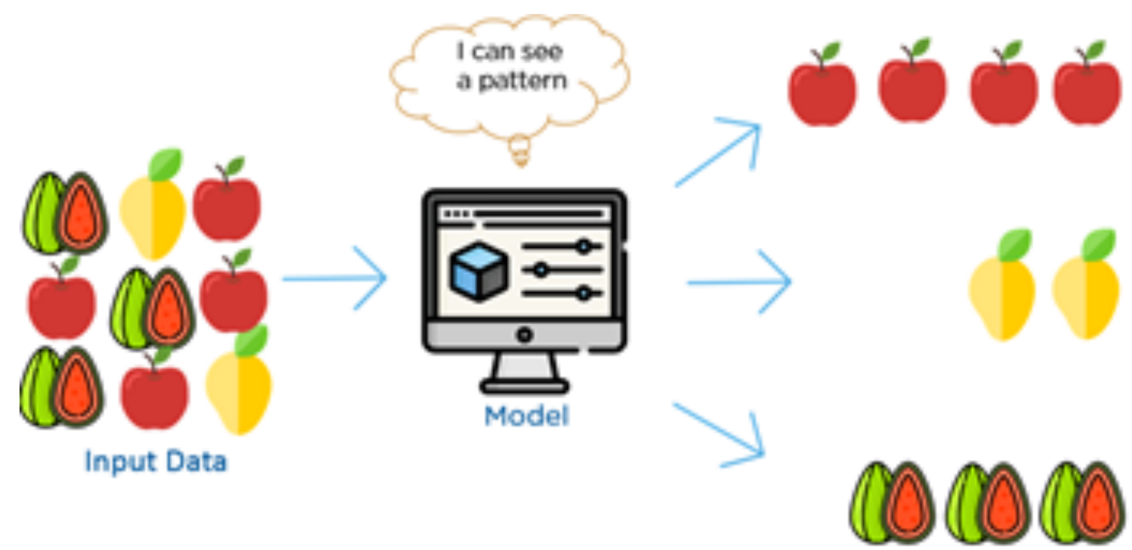
Anirudh Sharma

Applications)



Unsupervised learning

Learn from unlabeled data



The system looks for patterns and extracts features in data.

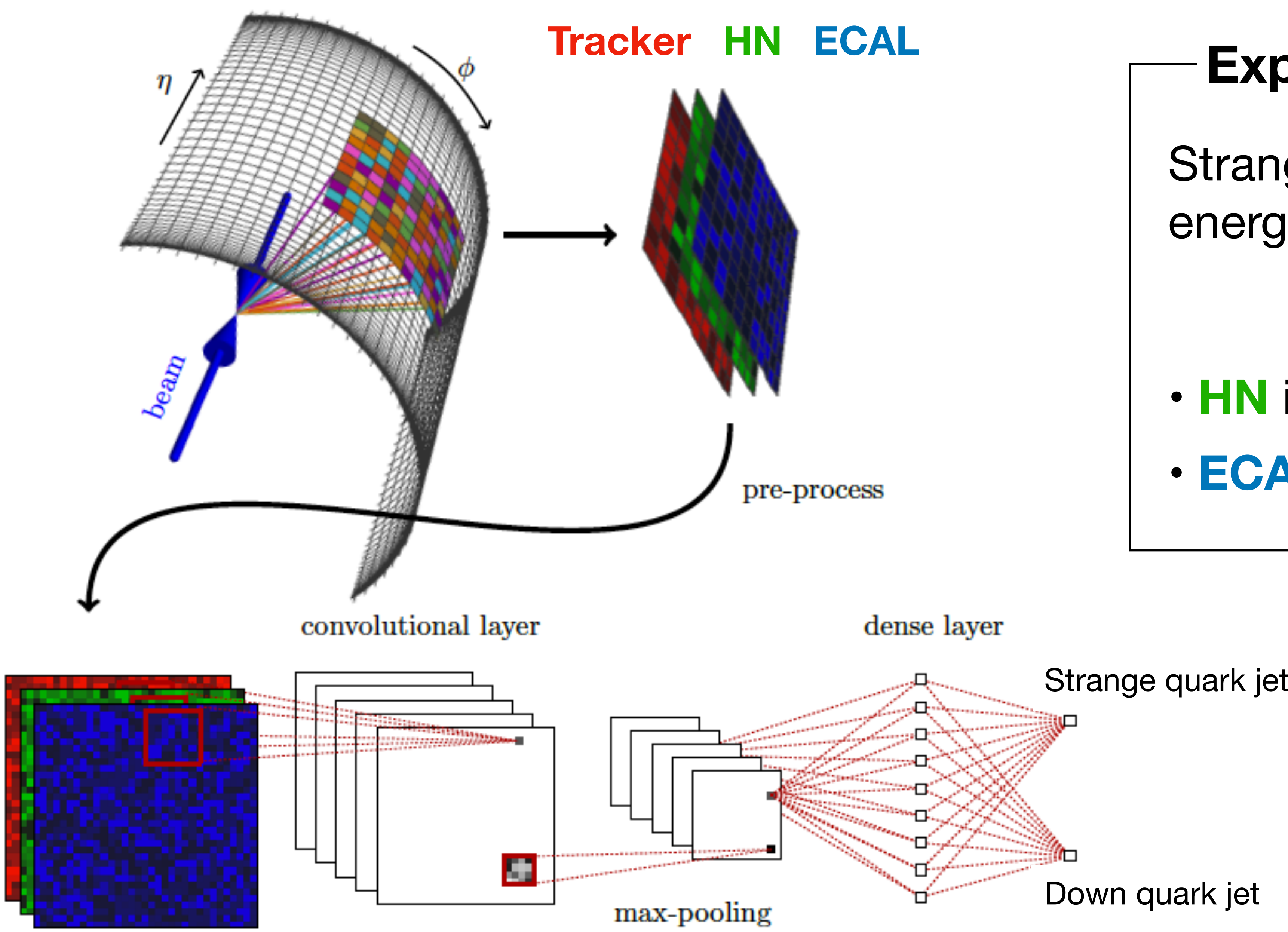
Applications) Clustering

Anomaly detection

Jet Images and CNN

Classification problem : **Strange jet** vs **Down jet**

Create jet images with colors (Tracker, HN, ECAL) and feed them into CNN.



Komiske, Metodiev, Schwartz (2016)

Expectation

Strange jet contains more energetic neutral Kaons.

- **HN** image : brighter
- **ECAL** image : darker

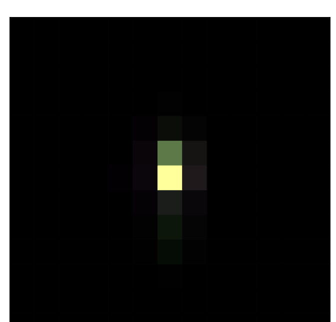
Jet Images

Create jet images with colors (Tracker, HN = HCAL - Tracker, ECAL).

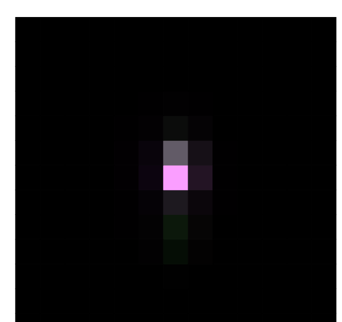
Image pre-processing

1. Shift an image so that the centroid is at the origin
2. Rotate the image so that the major principal axis is vertical
3. Flip the image so that the maximum intensity is in the upper right region
4. Normalize the image to unit total intensity : $\sum_{jet} (\hat{p}_T^{track} + \hat{E}_{had} + \hat{E}_{em}) = 1$
5. Pixelate the image : $\Delta\eta = \Delta\phi = 1.2$ 13 x 13 pixels

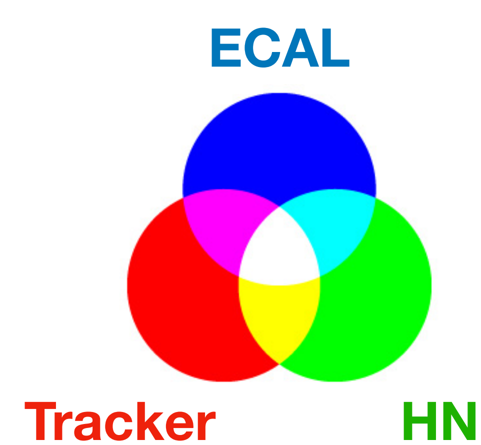
Average images :



$Z \to s\bar{s} \quad (p_T > 20 \text{ GeV})$

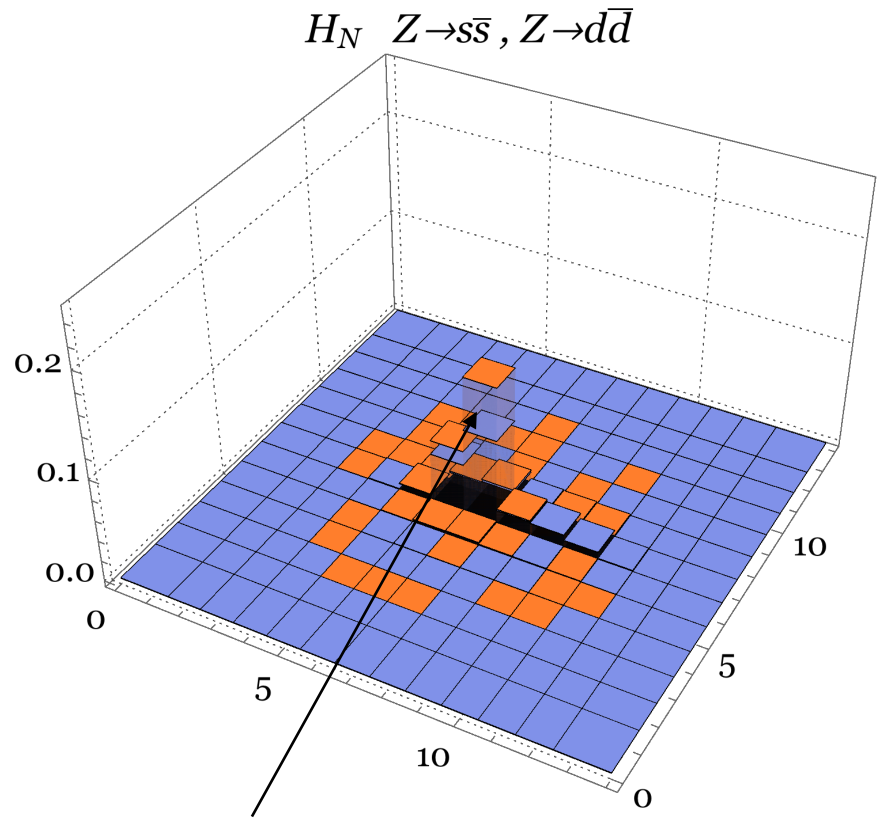


$Z \to d\bar{d} \quad (p_T > 20 \text{ GeV})$



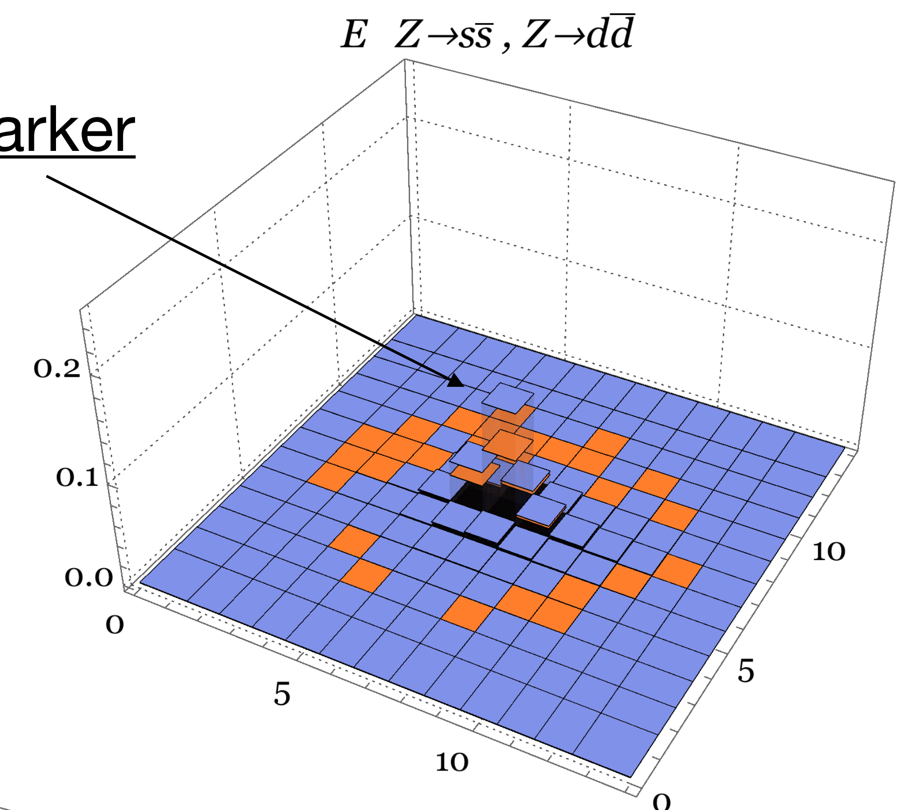
Average Images

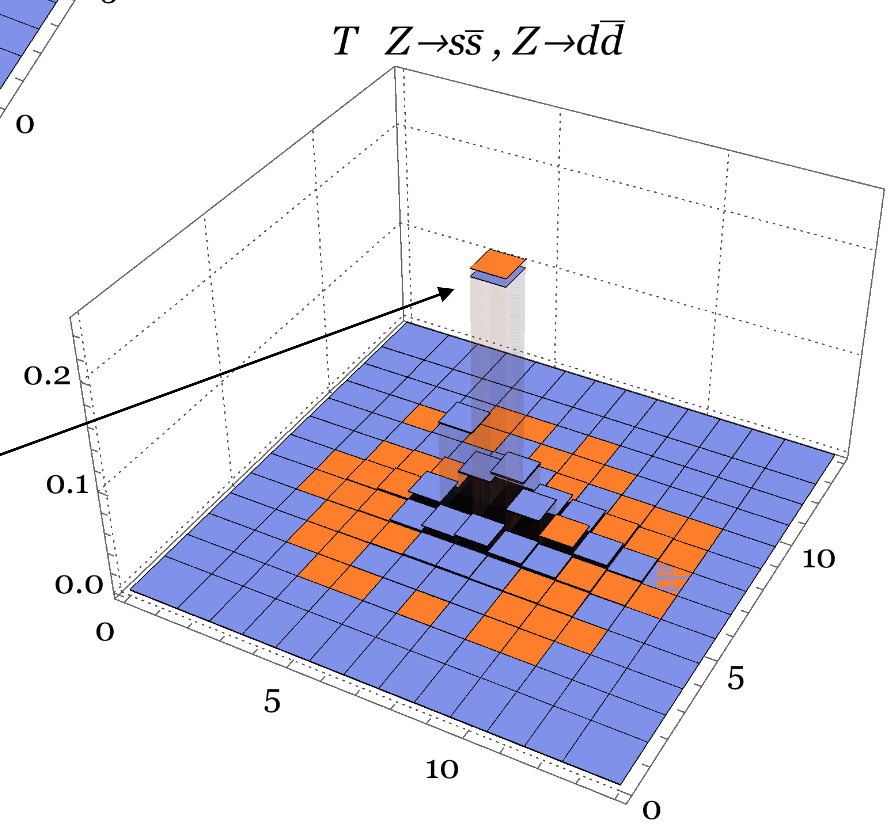
Strange jet (average) image is brighter in HN and darker in ECAL.



Strange is brighter

Strange is darker





No difference

$Z \to s\bar{s}$ ($p_T > 20$ GeV)

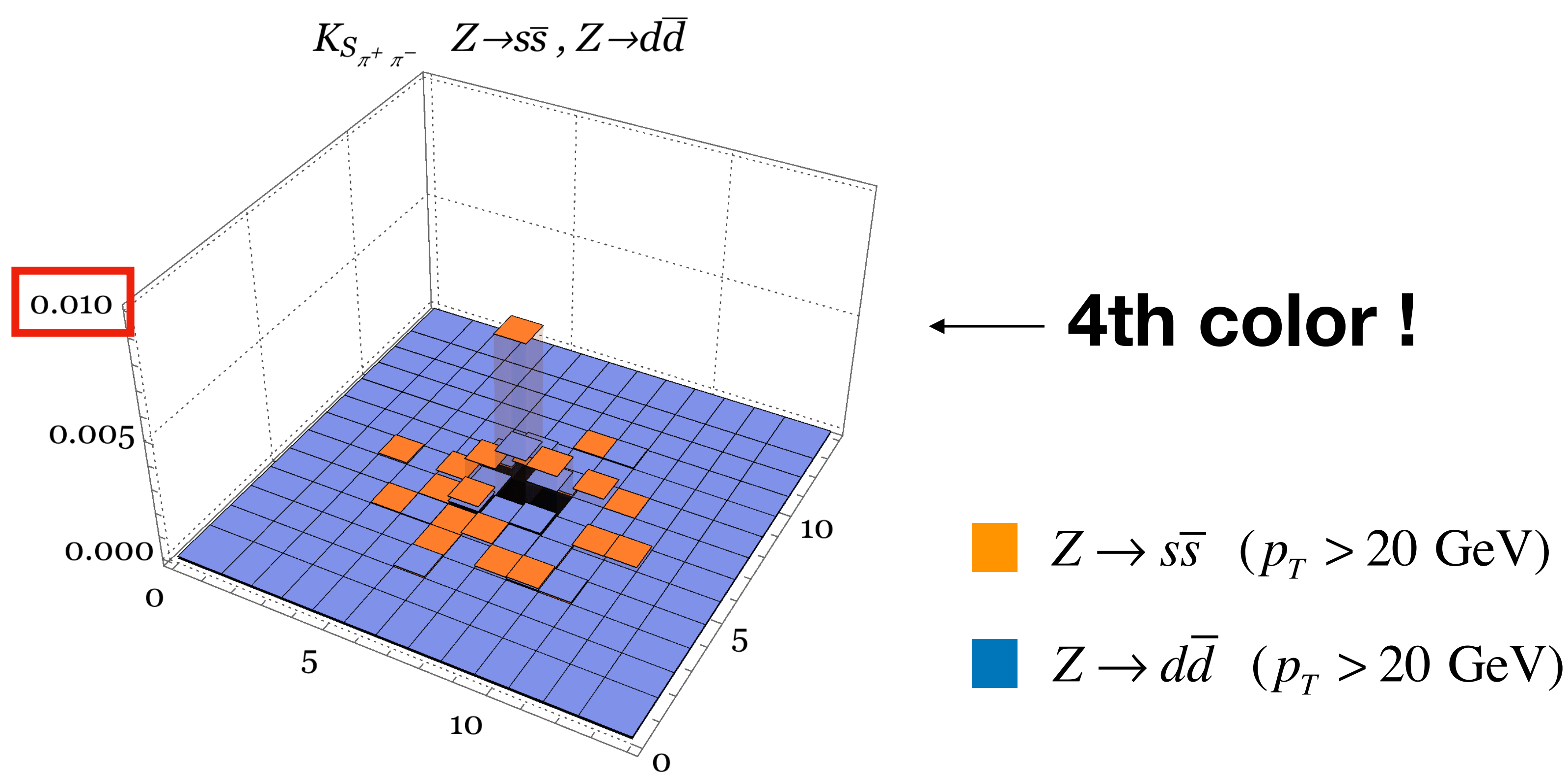
$Z \to d\bar{d}$ ($p_T > 20$ GeV)

Average Images

We add **the fourth color of the reconstructable K_S p_T**.

The intensity is normalized by the sum of the track p_T , ECAL and HN in the whole image.

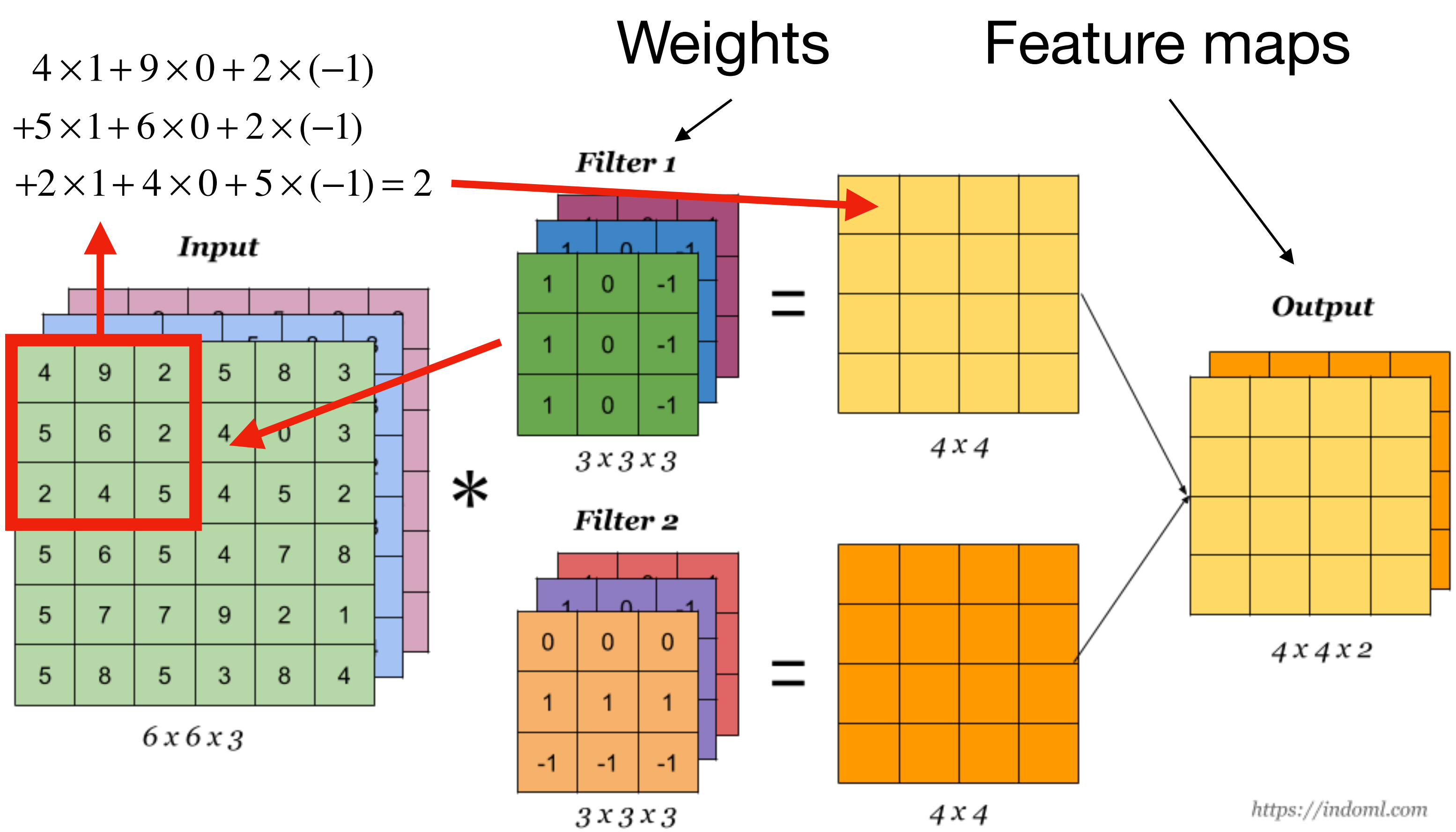
Strange jets contain more energetic K-shorts than down jets.



★ The intensity is much small compared to the other colors because the number of images including the reconstructable K_S is less than 8% (5%) of the total number of images for strange (down) jets.

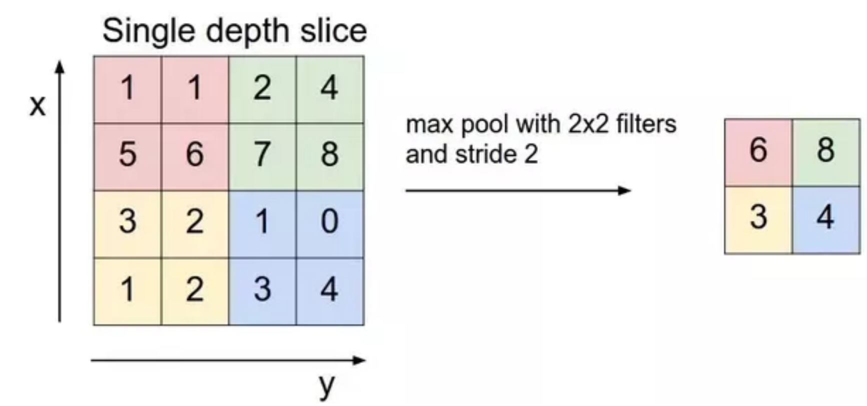
Network Architecture

Convolutional layer

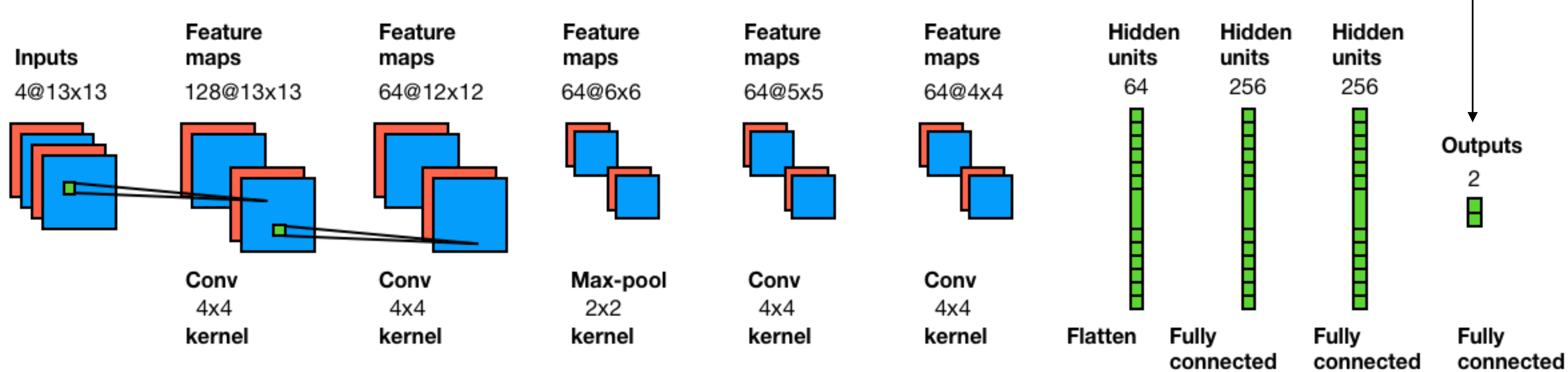


Max pooling

Reduce the image size



Probabilities of signal and background



Training

The goal of training is to minimize loss function :

Network prediction

$$L = \sum_i f(p(\theta, x_i), y_i)$$

$p(\theta, x_i)$: Signal probability

θ : Weights

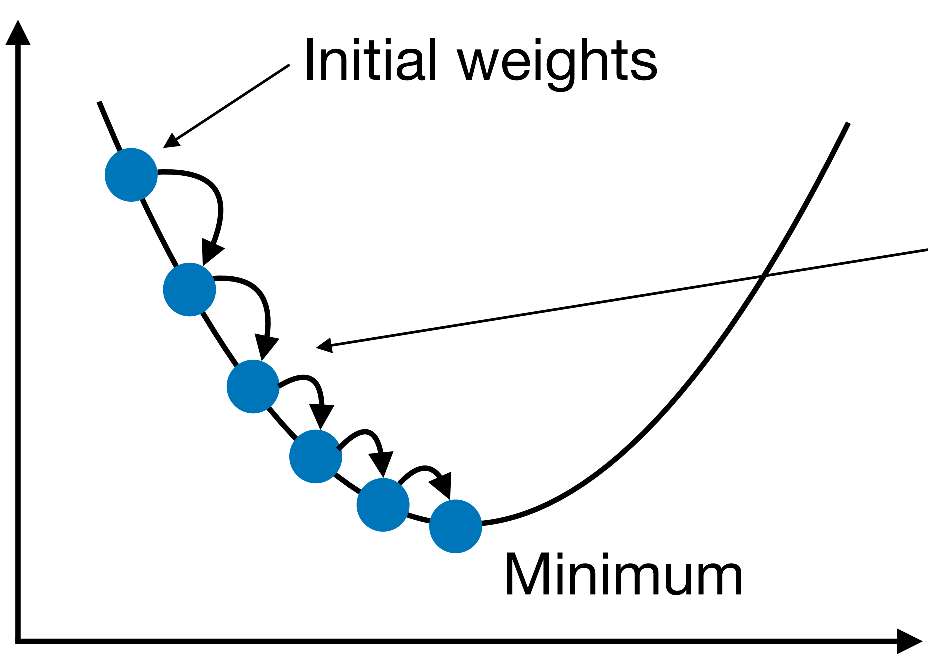
x_i : Input

y_i : Truth label of example i

$$\left(\begin{array}{ll} y_i = 0 & \text{: Signal} \\ y_i = 1 & \text{: Background} \end{array} \right)$$

We use cross entropy : $f(p, y) = -(y \log(1-p) + (1-y)\log p)$

Loss function L



Weights θ

Weights are updated according to derivative of loss function :

$$\Delta\theta = -\eta \nabla L$$

Learning rate

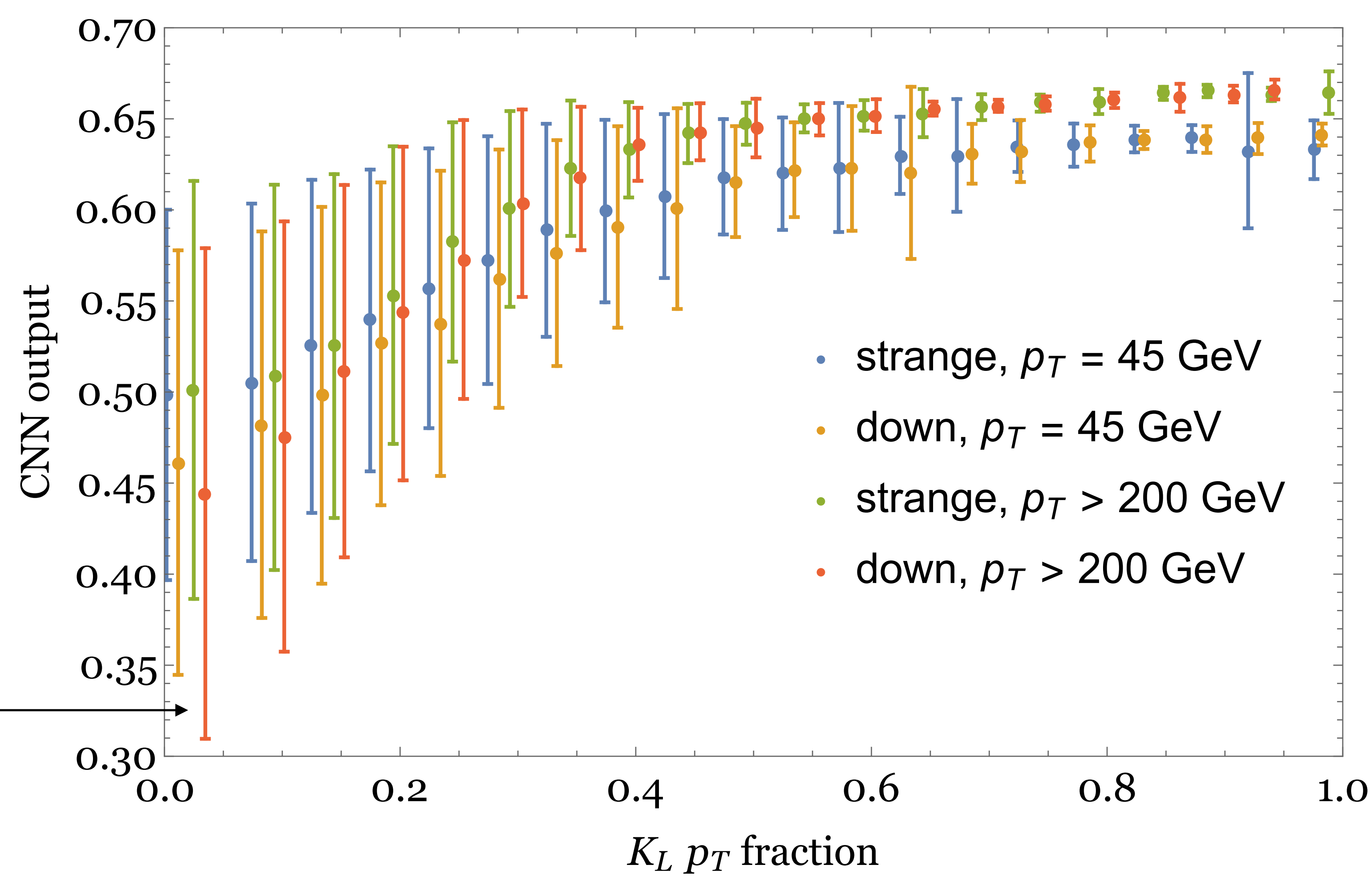
Neural Network Output

The correlation between the K_L pT ratio of input images and the CNN (with 3 colors of tracker, HN and EN) outputs.

The K_L pT ratio : p^{KL} /pT j p^{KL} : the sum of K_L pT in a jet, pTj : jet pT

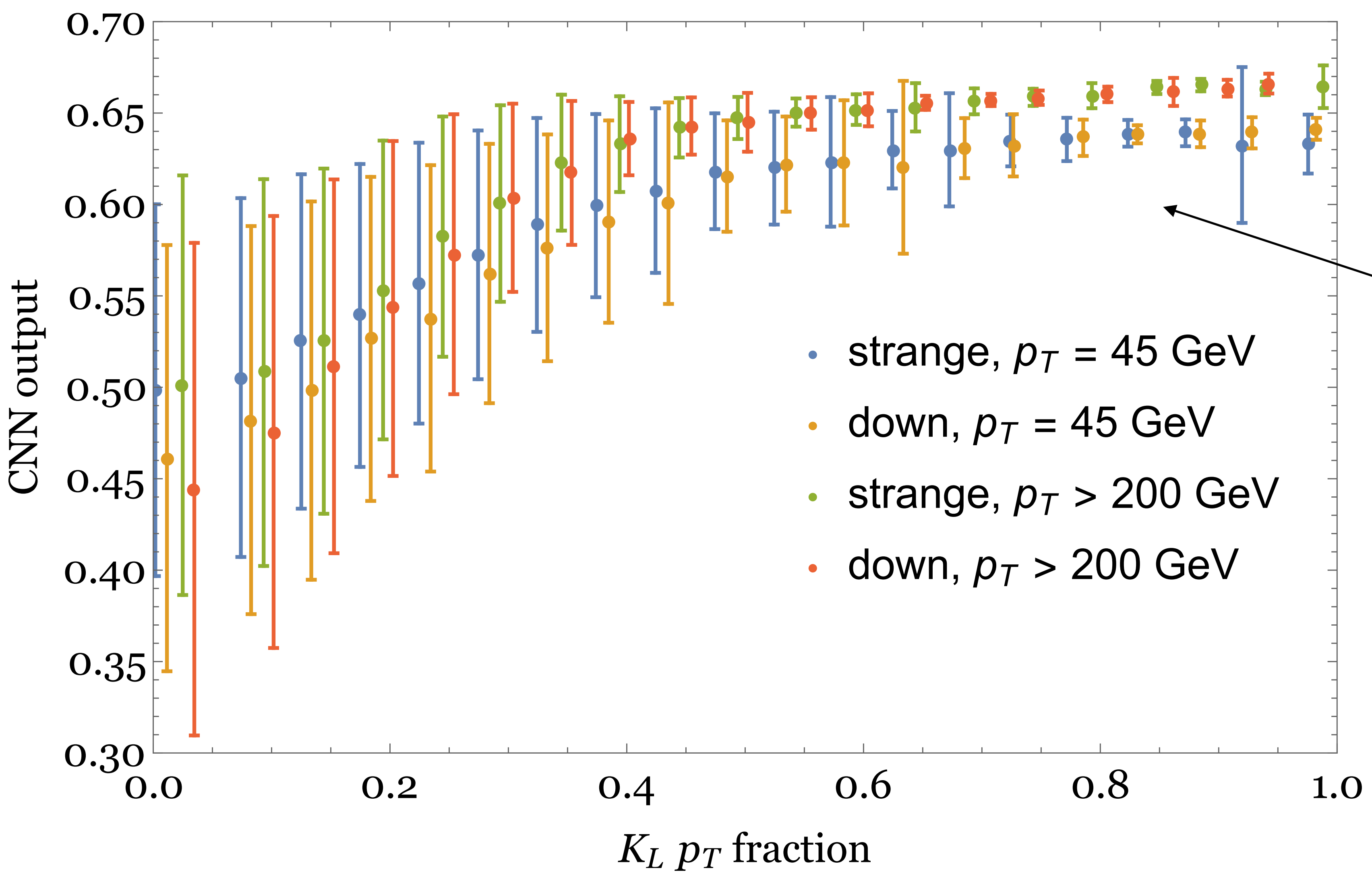
Take the average in each bin.

Standard deviation



Neural Network Output

A clear correlation: **Signal probability increases as K-long pT ratio increases**



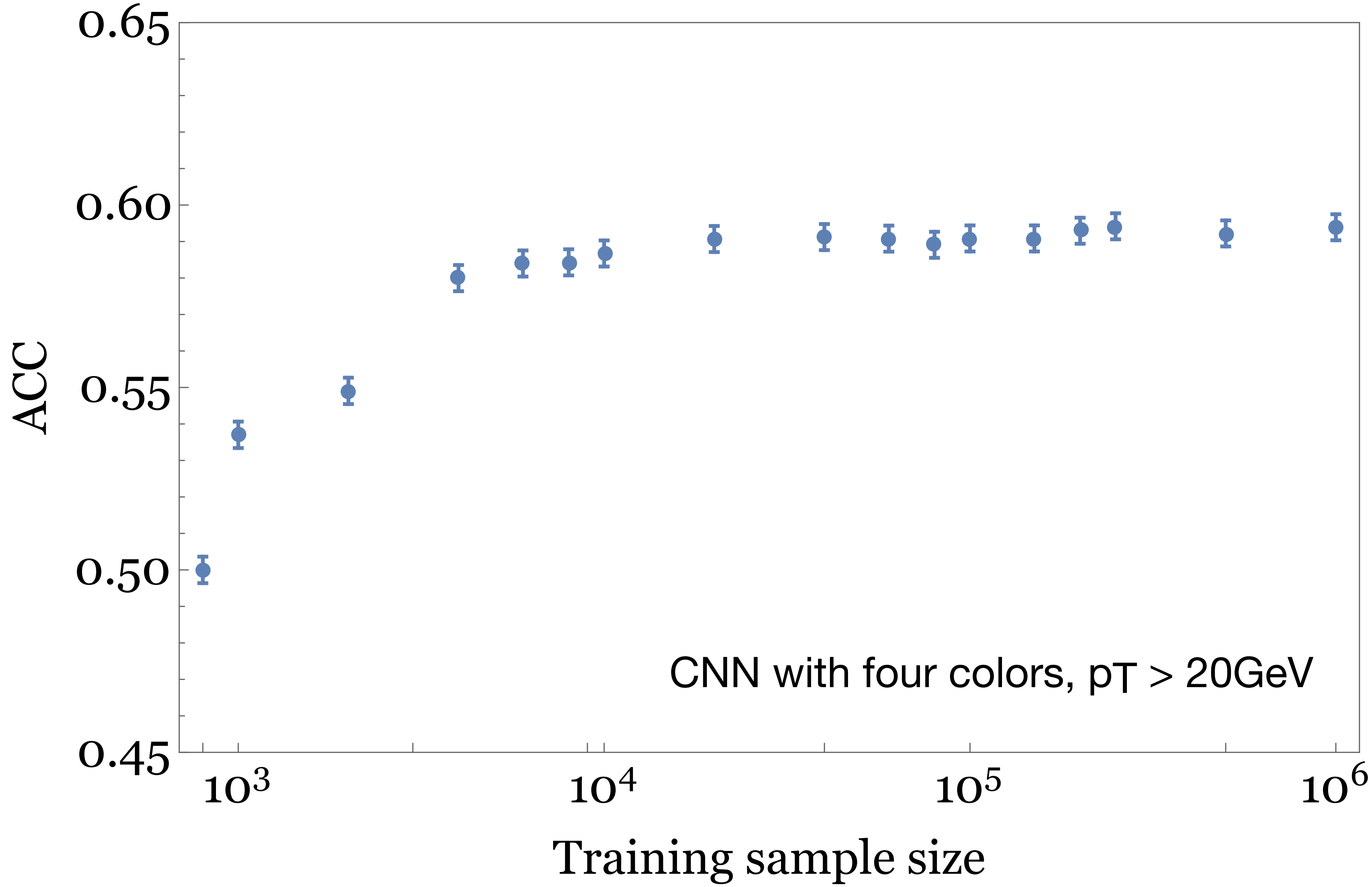
Standard deviation decreases.

In the low K_L pT ratio region, the signal probability of strange jets is larger than that of down jets which can be understood by taking into account of the difference between strange and down jets in terms of the K_S component.

Training Curve

How the performance of the CNN is affected by the number of training samples.

The best accuracy



The performance saturates immediately for more than 10000 training samples.

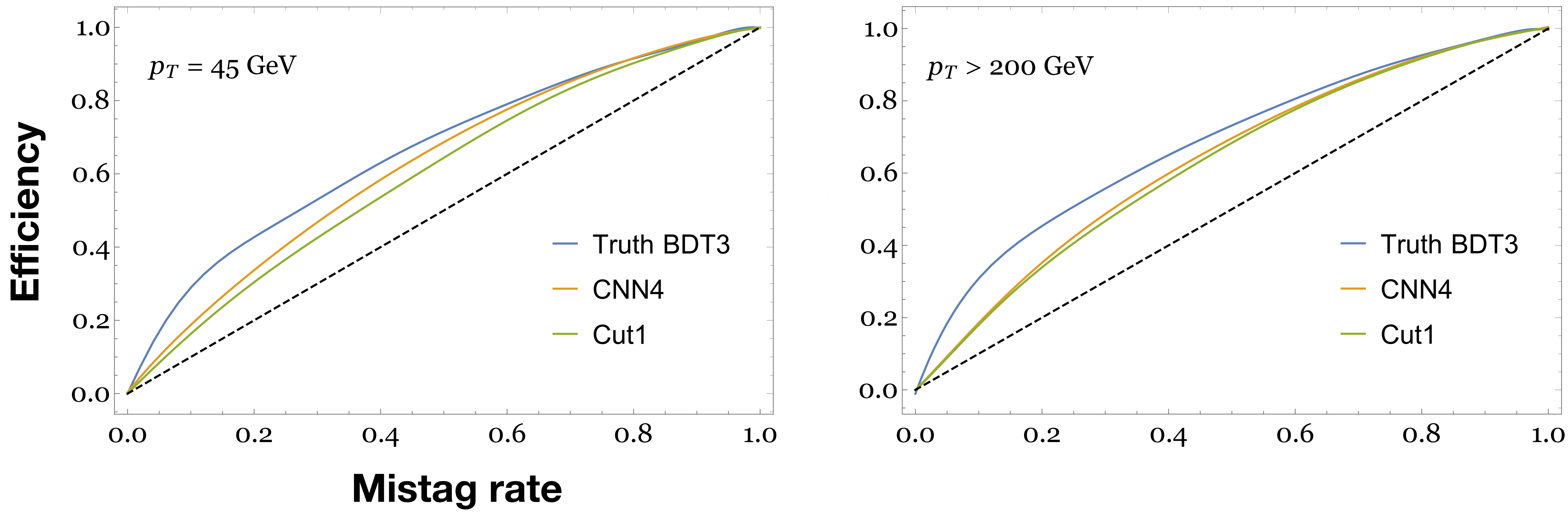
Various taggers

The list of our taggers and their inputs :

Algorithm	Input Source	Input Variable(s)	
Truth Cut1	Pythia 8	$-\pi^0 + K_L + K_S + K_{S_{\pi^+\pi^-}}$	
Truth BDT3	Pythia 8	π^0 , K_L , $K_S + K_{S_{\pi^+\pi^-}}$	
Cut1	Delphes	$H_N - E$	
Cut1+	Delphes	$H_N - E + K_{S_{\pi^+\pi^-}}$	
BDT3	Delphes	H_N , E , T	
BDT4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	
CNN3	Delphes	H_N , E , T	13×13 Jet Image
CNN4	Delphes	H_N , E , T , $K_{S_{\pi^+\pi^-}}$	13×13 Jet Image

Results

ROC curves :



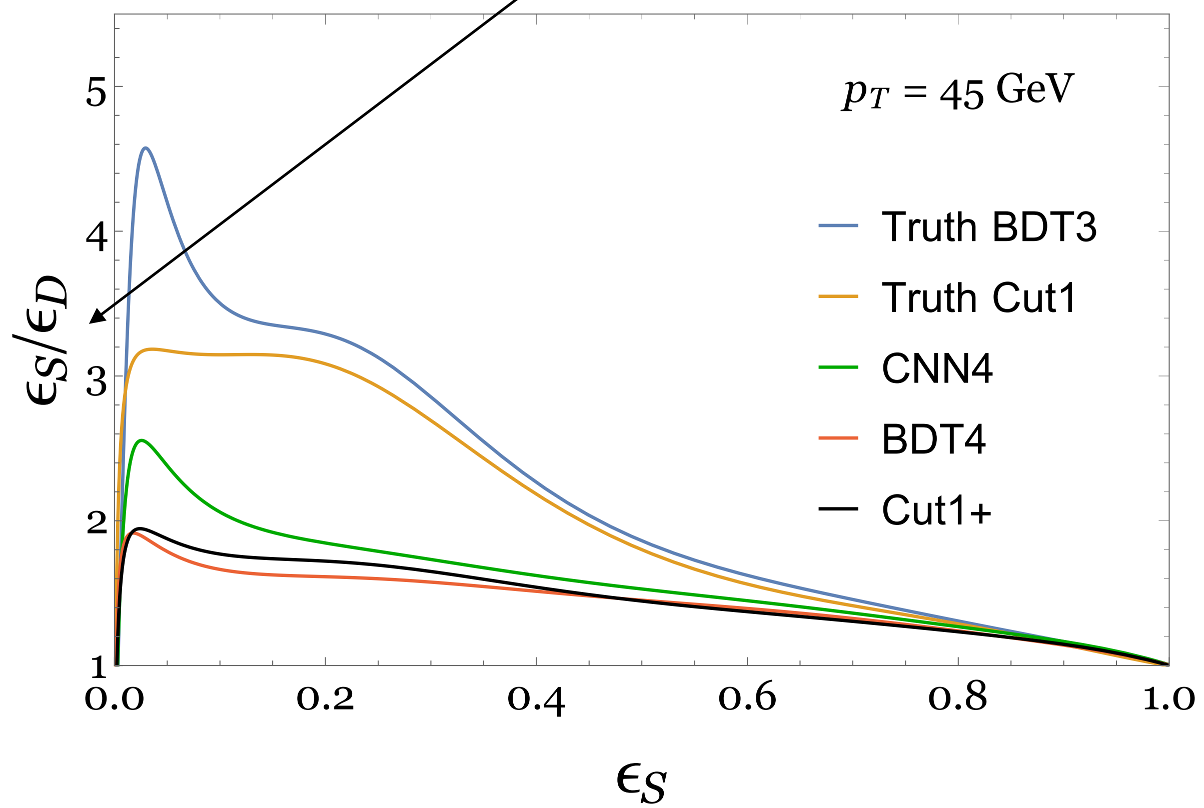
✓ <u>K-short color is less important for pT > 200 GeV</u> because the number of long-lived K-shorts is larger due to boost factor.

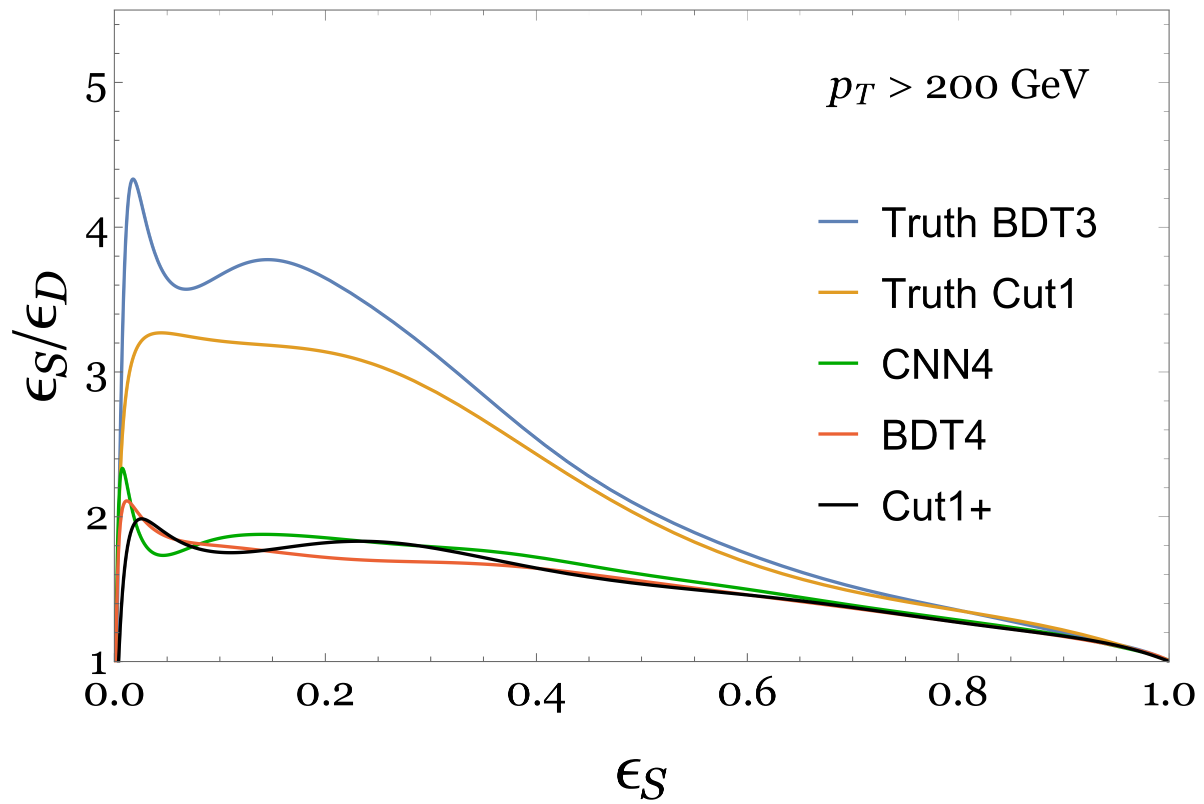
✓ <u>CNN curves are similar for pT = 45 GeV and pT > 200 GeV.</u>

Results

Comparison of various taggers :

(Efficiency) / (Mistag rate)



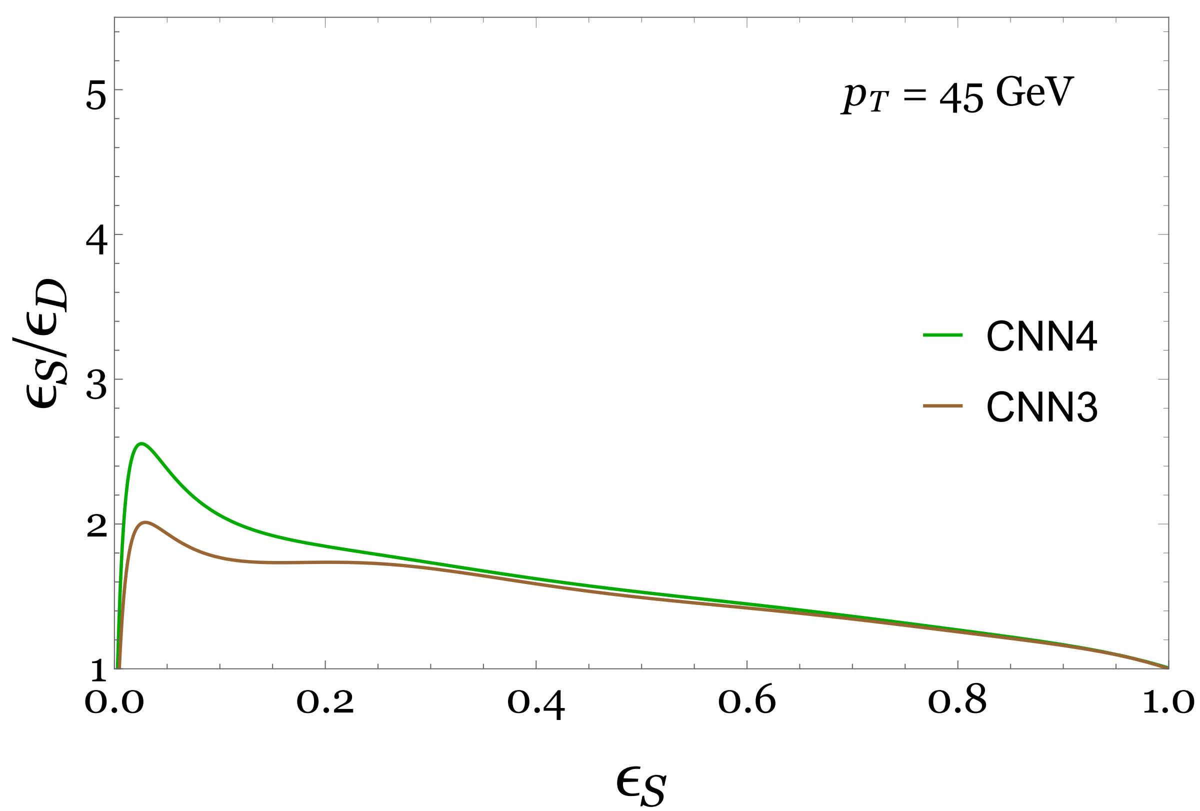


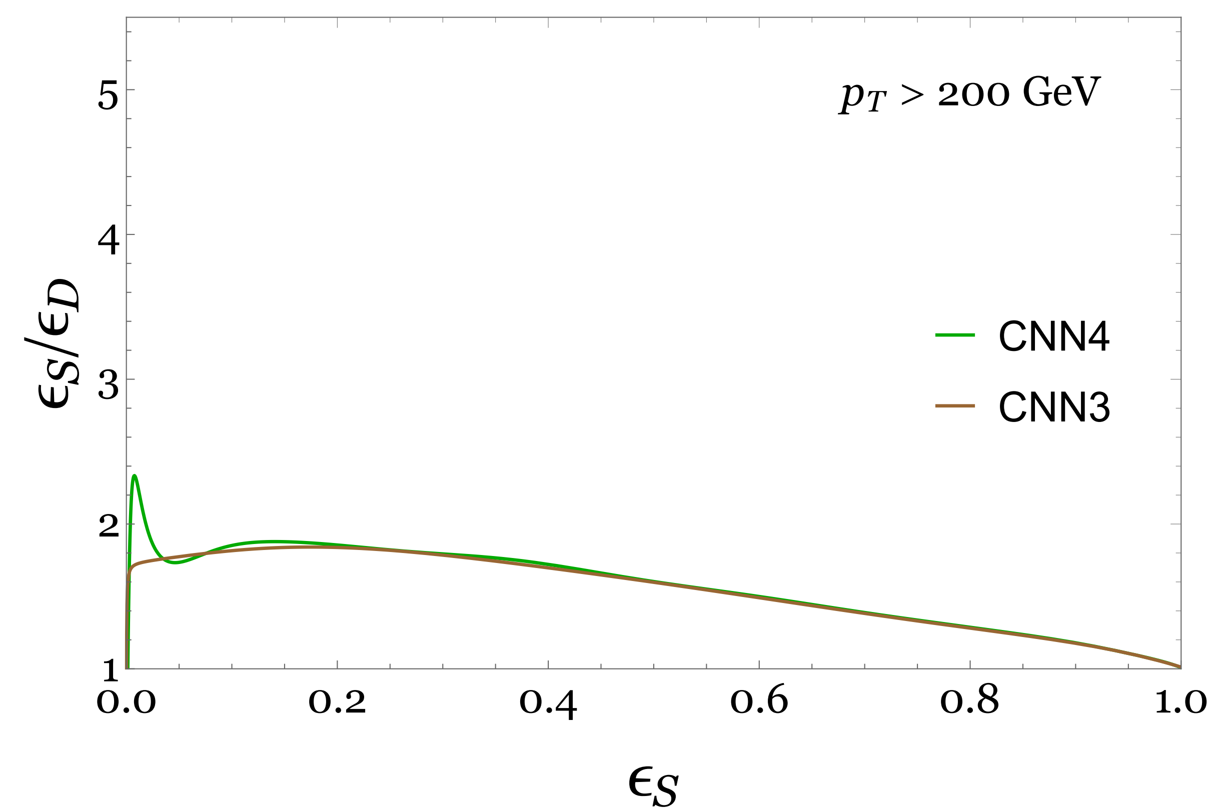
✓ In the case of pT > 20GeV, <u>the CNN tagger outperforms the other taggers</u> and approaches the curve of the truth-level classifiers.

✓ In the case of pT > 200 GeV, <u>the curves of various taggers are more degenerate</u>.

Results

Comparison of the CNN applied to jet images with 3 colors (tracker, HN and EN) and 4 colors (tracker, HN, EN and K_S p_T) :





✓ In the case of p_T > 20 GeV, the CNN tagger with <u>four colors outperforms that with three colors</u> in the low efficiency region.

✓ In the case of p_T > 200 GeV, the two curves are more degenerate because the K_S is more long-lived and <u>the fourth color is not effective</u>.

Results

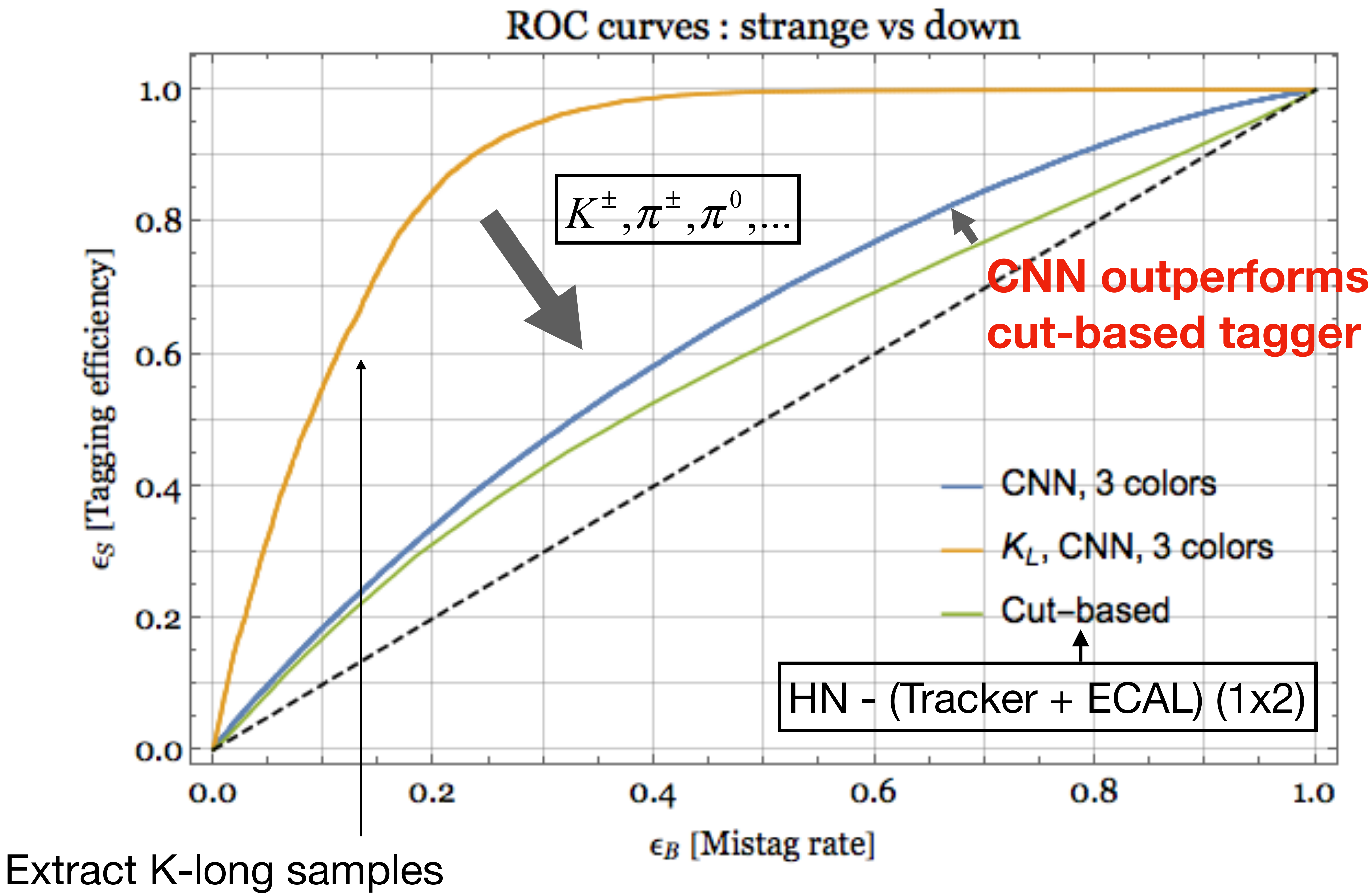
The CNN with four colors shows the best performance.

pT = 45 GeV   pT > 200 GeV

	AUC	ACC	R10	R50
Truth Cut1	0.65 (0.68)	0.61 (0.62)	31.9 (32.3)	3.6 (3.9)
Truth BDT3	0.66 (0.68)	0.62 (0.63)	36.4 (37.0)	3.7 (4.2)
Cut1	0.60 (0.63)	0.57 (0.59)	16.8 (17.9)	2.7 (3.0)
Cut1+	0.62 (0.63)	0.58 (0.60)	17.8 (17.8)	2.9 (3.1)
BDT3	0.61 (0.63)	0.58 (0.60)	16.0 (18.0)	2.8 (3.1)
BDT4	0.61 (0.63)	0.59 (0.60)	17.0 (18.0)	2.9 (3.1)
CNN3	0.62 (0.63)	0.59 (0.60)	17.7 (18.2)	3.0 (3.2)
CNN4	0.63 (0.64)	0.59 (0.60)	20.9 (18.3)	3.1 (3.2)

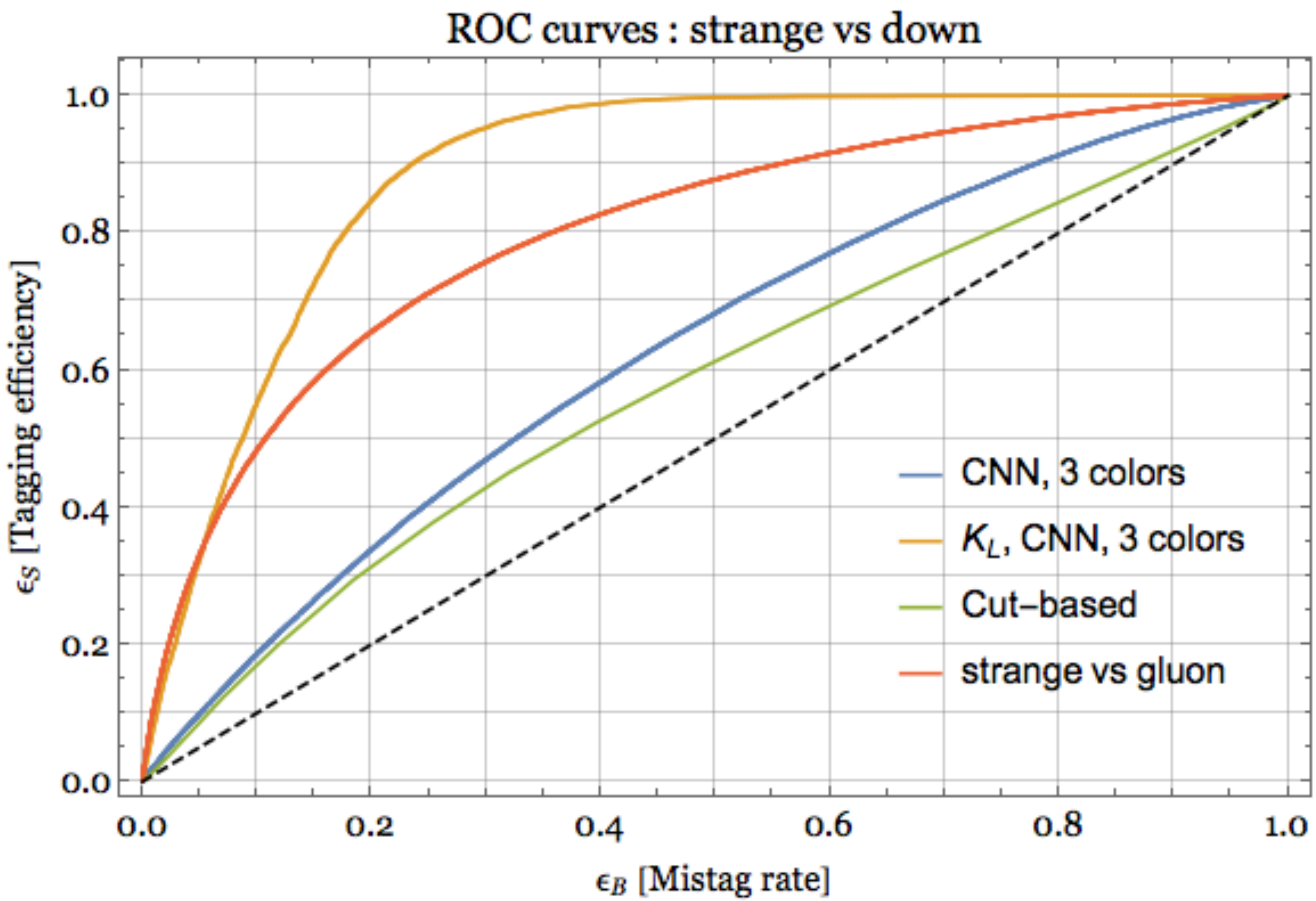
$\mathrm{R10} = 1/\epsilon_D$ for $\epsilon_S = 0.1$ $\mathrm{R50} = 1/\epsilon_D$ for $\epsilon_S = 0.5$

K-long Jets



K-long ROC curve is very good and contamination with other hadrons lowers the performance.

Gluon Jets



Quark and gluon jets are more different than strange and down jets and quark/gluon tagger has higher performance than strange/down tagger.

Results

The CNN with four colors shows the best performance.

pT = 45 GeV   pT > 200 GeV

	AUC	ACC	R10	R50
Truth Cut1	0.65 (0.68)	0.61 (0.62)	31.9 (32.3)	3.6 (3.9)
BDT3	0.61 (0.63)	0.58 (0.60)	16.0 (18.0)	2.8 (3.1)
BDT4	0.61 (0.63)	0.59 (0.60)	17.0 (18.0)	2.9 (3.1)
CNN3	0.62 (0.63)	0.59 (0.60)	17.7 (18.2)	3.0 (3.2)
CNN4	0.63 (0.64)	0.59 (0.60)	20.9 (18.3)	3.1 (3.2)

Can we use such a weak classifier ??

$$R10 = 1/\epsilon_D \text{ for } \epsilon_S = 0.1 \qquad R50 = 1/\epsilon_D \text{ for } \epsilon_S = 0.5$$

Significance Improvement

Consider a binary classifier with efficiency $\boldsymbol{\varepsilon_S}$ and mistag rate $\boldsymbol{\varepsilon_B}$.

Before a cut on the classifier…

Statistical significance of the signal : $S/\sqrt{B}$ (S « B)

After a cut on the classifier…

If we throw away the events that fail the cut…

Statistical significance of the signal : $q = \boxed{\frac{\epsilon_S}{\sqrt{\epsilon_B}}} \frac{S}{\sqrt{B}}$

Significance improvement factor

If a weak classifier gives a significance improvement factor smaller than 1, the classifier reduces our significance ??

Significance Improvement

If we view the classifier as defining two categories (pass vs. fail)…

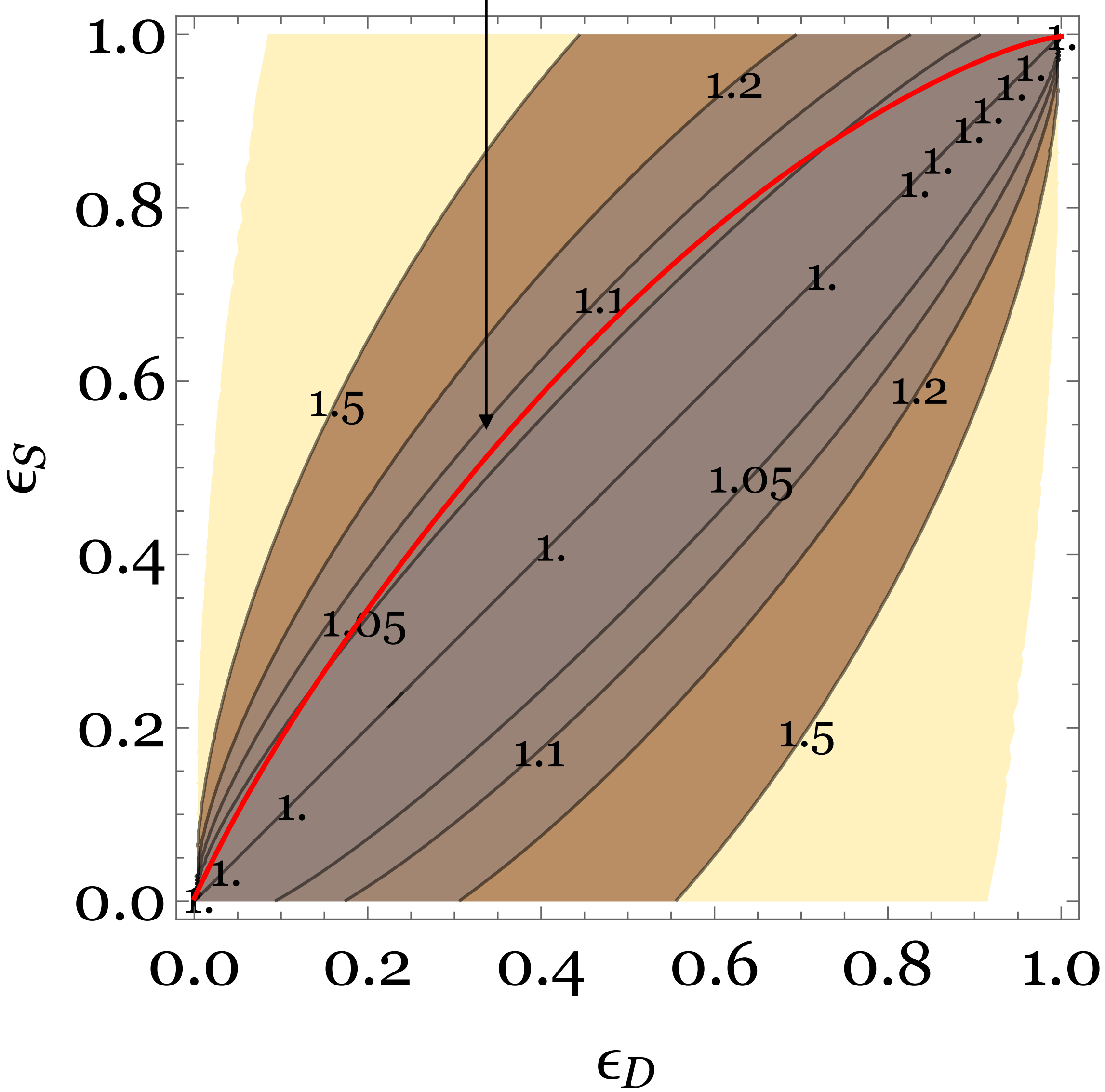
Combined significance of two categories :

$$q = \sqrt{\left(\frac{\epsilon_S}{\sqrt{\epsilon_B}}\frac{S}{\sqrt{B}}\right)^2 + \left(\frac{(1-\epsilon_S)}{\sqrt{(1-\epsilon_B)}}\frac{S}{\sqrt{B}}\right)^2}$$

$$= \sqrt{1 + \frac{(\epsilon_S - \epsilon_B)^2}{\epsilon_B(1-\epsilon_B)}}\frac{S}{\sqrt{B}}$$

The significance can only increase.

Our best strange jet tagger (CNN)



Our best strange jet tagger (CNN)

Significance improvement is only 5-10%.

The importance of our strange tagger is limited…

CKM Mixings

A better use for the strange tagger is to measure the ratio of strange to down jets in some setting.

In this case…

Any amount of discrimination power will make the measurement possible with enough data.

Remember…

The values for $|V_{cs}|$ and $|V_{cd}|$ are not measured very well.

- ✓ Because the charm quark mass is too heavy to be considered light but not heavy enough to treat in the heavy quark limit.
- ✓ W boson decay $W \to cs$ gives the most direct measurement if strange tagging is possible.

Let's consider the measurement of the ratio : $\dfrac{|V_{cs}|^2}{|V_{cd}|^2}$

CKM Mixings

A simple estimate

# of data

A fraction of strange to down

# of events passing the cut : $N_{pass}(f_S) = N(f_S \epsilon_S + (1 - f_S)\epsilon_B)$

# of events failing the cut : $N_{fail}(f_S) = N(f_S(1 - \epsilon_S) + (1 - f_S)(1 - \epsilon_B))$

$$\chi^2 = \frac{(N_{pass}(f_S) - N_{pass}(\hat{f}_S))^2}{N_{pass}(\hat{f}_S)} + \frac{(N_{fail}(f_S) - N_{fail}(\hat{f}_S))^2}{N_{fail}(\hat{f}_S)}$$
$$= \frac{N(\epsilon_B - \epsilon_S)^2}{f_{eff}(1 - f_{eff})}(f_S - \hat{f}_S)^2$$

$$f_{eff} = \hat{f}_S \epsilon_S + (1 - \hat{f}_S)\epsilon_B$$

True fraction

$$\boxed{f_S = \hat{f}_S \pm \delta f_S \qquad \delta f_S = \frac{1}{|\epsilon_B - \epsilon_S|}\sqrt{\frac{f_{eff}(1 - f_{eff})}{N}}}$$

As long as $\varepsilon_B \neq \varepsilon_S$, a sufficiently large N gives an accurate measurement.

CKM Mixings

W boson decay

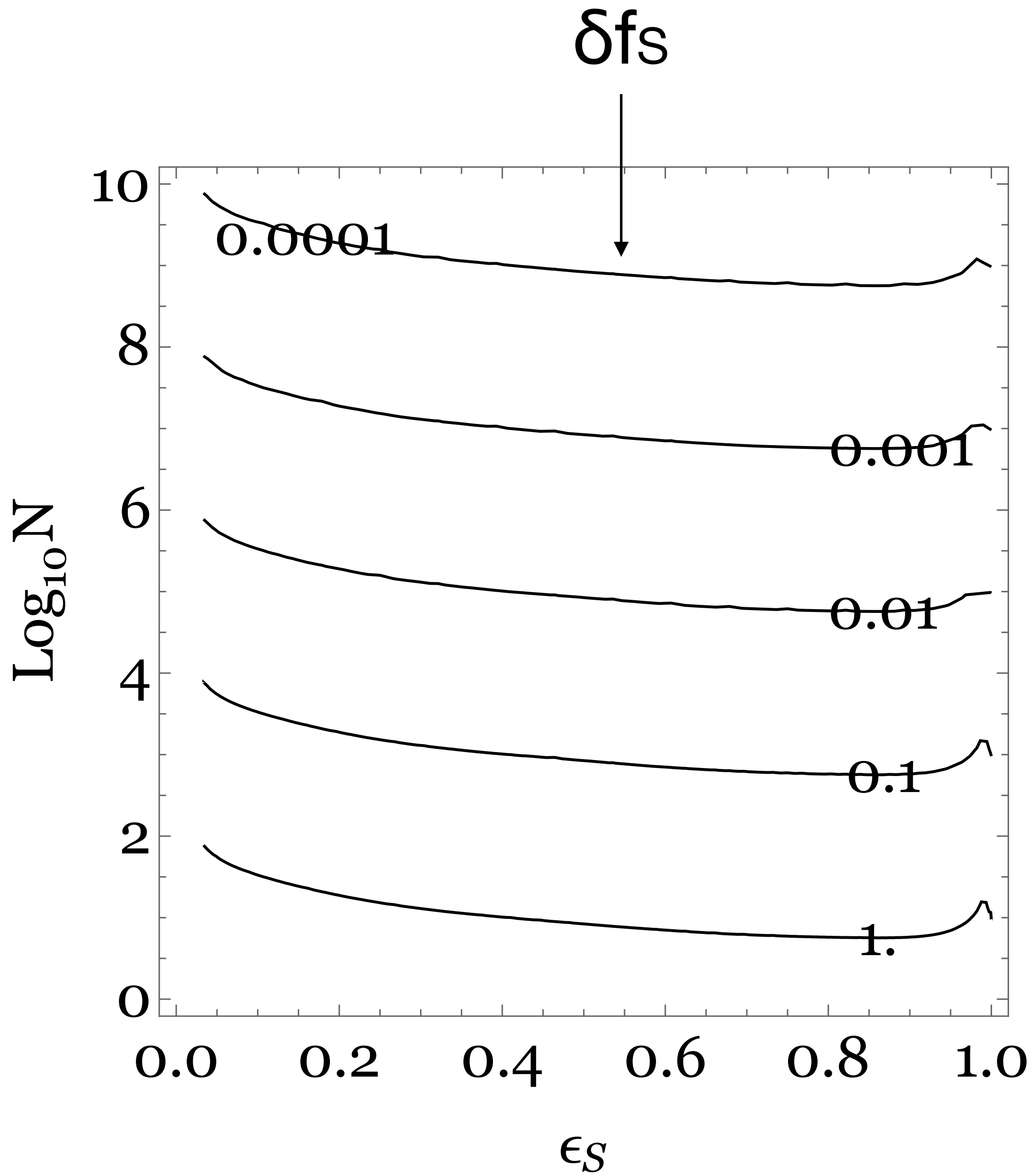
$$\Gamma(W^- \to s\bar{c}) = \Gamma(W^- \to e^-\bar{\nu}) \times 3|V_{cs}|^2$$
$$\Gamma(W^- \to d\bar{c}) = \Gamma(W^- \to e^-\bar{\nu}) \times 3|V_{cd}|^2$$

$$f_S = \hat{f}_S \pm \delta f_S$$

$$\delta f_S = \frac{1}{|\epsilon_B - \epsilon_S|}\sqrt{\frac{f_{eff}(1-f_{eff})}{N}}$$

$$f_{eff} = \hat{f}_S \epsilon_S + (1-\hat{f}_S)\epsilon_B$$

$$\boxed{\hat{f}_s = \frac{\frac{|V_{cs}|^2}{|V_{cd}|^2}}{1+\frac{|V_{cs}|^2}{|V_{cd}|^2}}} \qquad \frac{|V_{cs}|^2}{|V_{cd}|^2} \sim 20$$



Our best strange jet tagger (CNN)

Since the LHC generates a lot of W bosons, a precise measurement is possible !

Summary

✓ Strange tagging is the last missing piece of quark/gluon tagging.

✓ Neutral Kaons can be used for strange tagging.

✓ We create jet images with colors (**Tracker**, **Hadronic Neutral**, **ECAL**, **Ks pT**).
(= HCAL - Tracker)

✓ Average images of strange jets can be distinguished from down images.

✓ Convolutional Neural Network outperforms cut-based tagger.

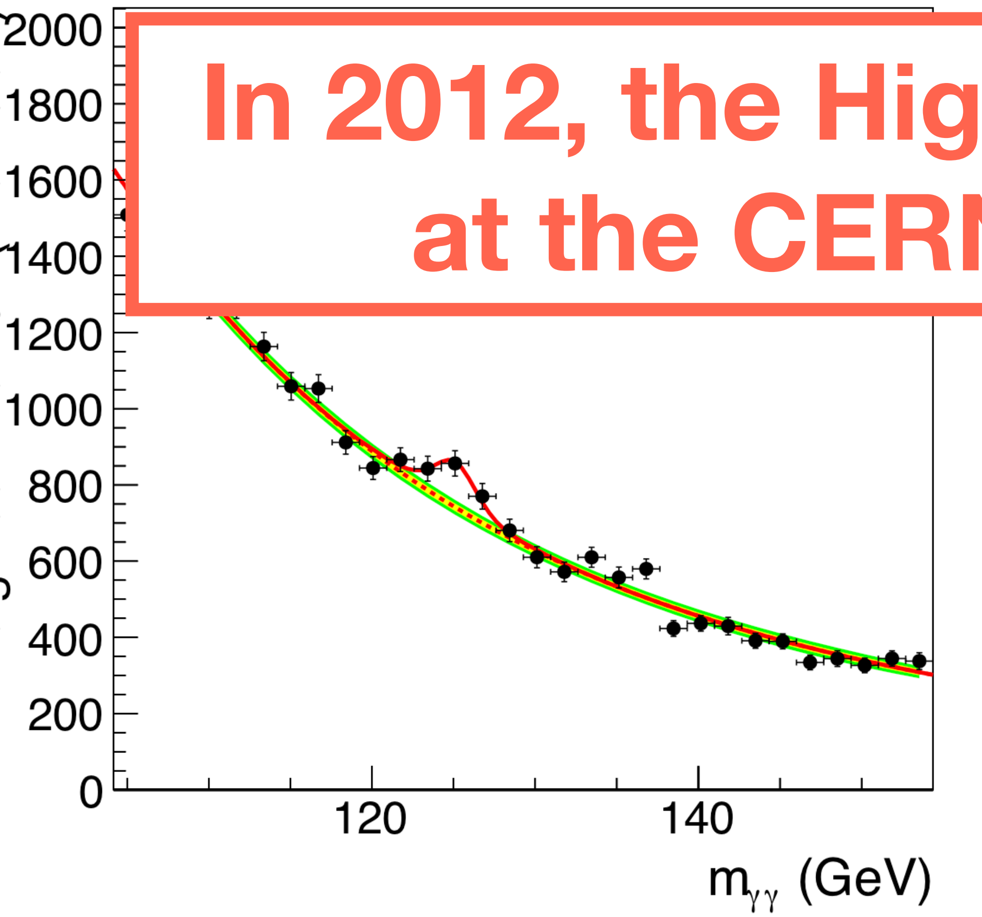
✓ Strange jet tagger may be important for a measurement of CKM mixings.

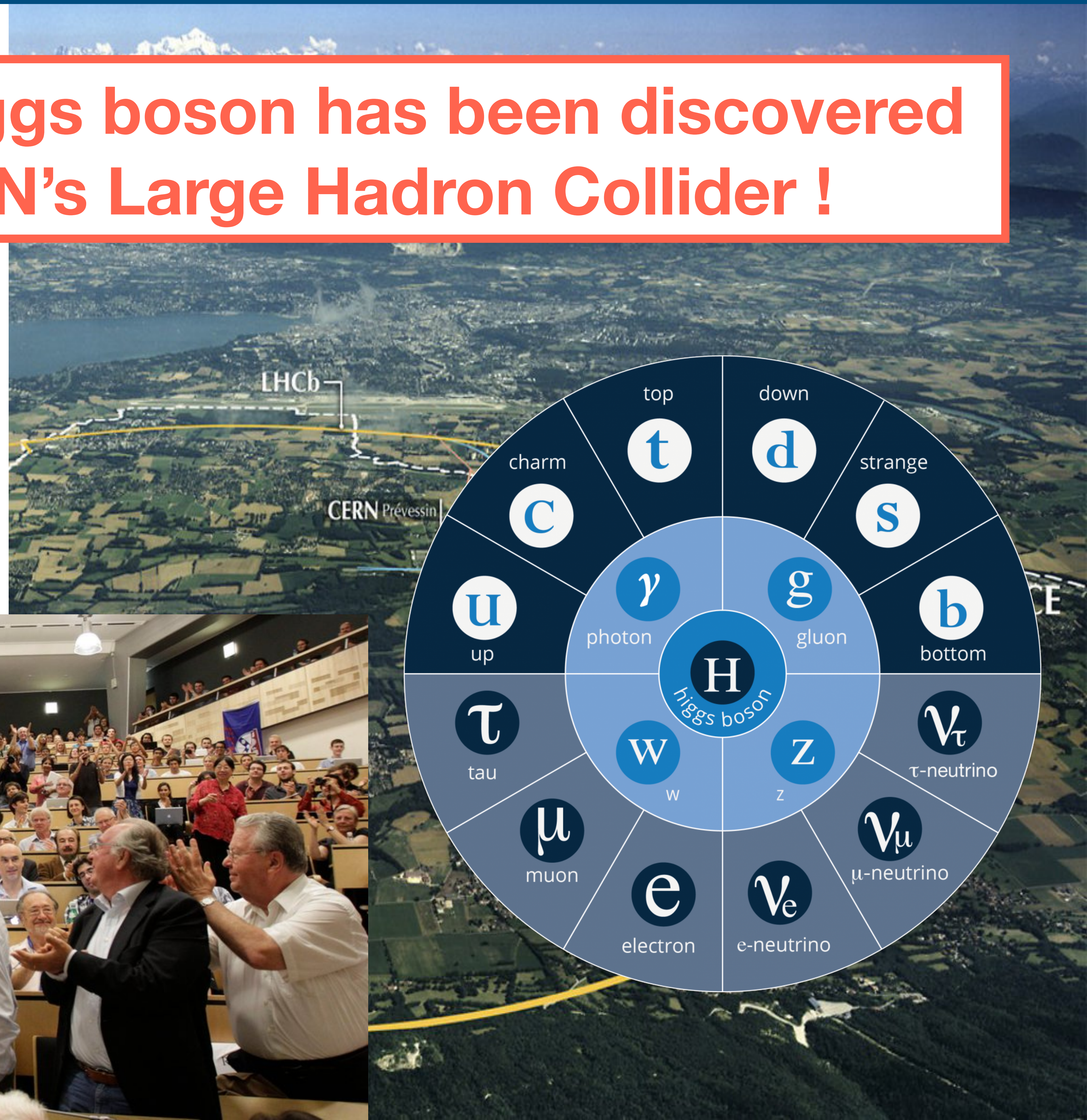
Thank you.

Discussion

The Great Achievement

In 2012, the Higgs boson has been discovered at the CERN's Large Hadron Collider !





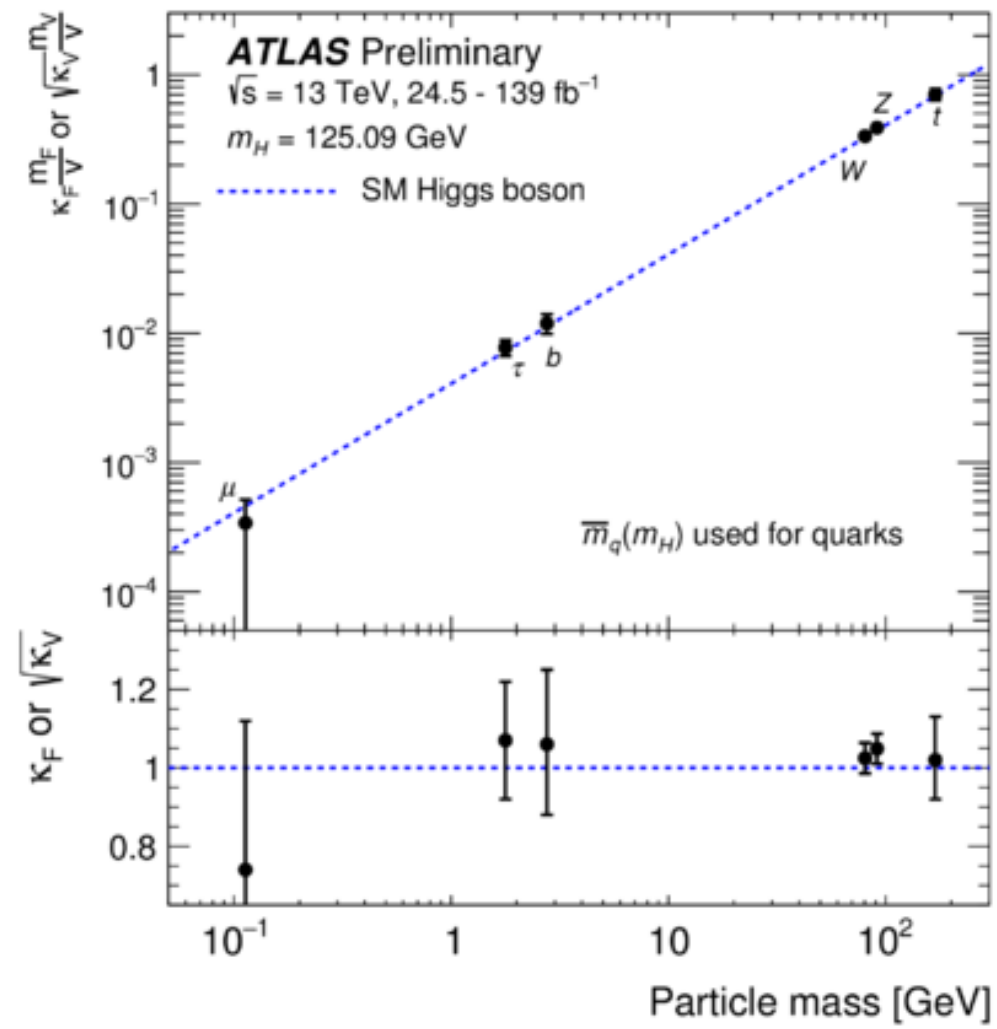
Higgs Precision Era

The SM predictions of the Higgs couplings to heavy gauge bosons and fermions, $2m^2_{W,Z}/v$ and m_f/v, have been confirmed for the W and Z bosons and for the third-generation fermions.

A key aspect of the experimental program of post-LHC colliders includes precision studies of…

- Higgs couplings
- Self-couplings (HH production)
- Total width
- Exotic / Invisible decays

  …



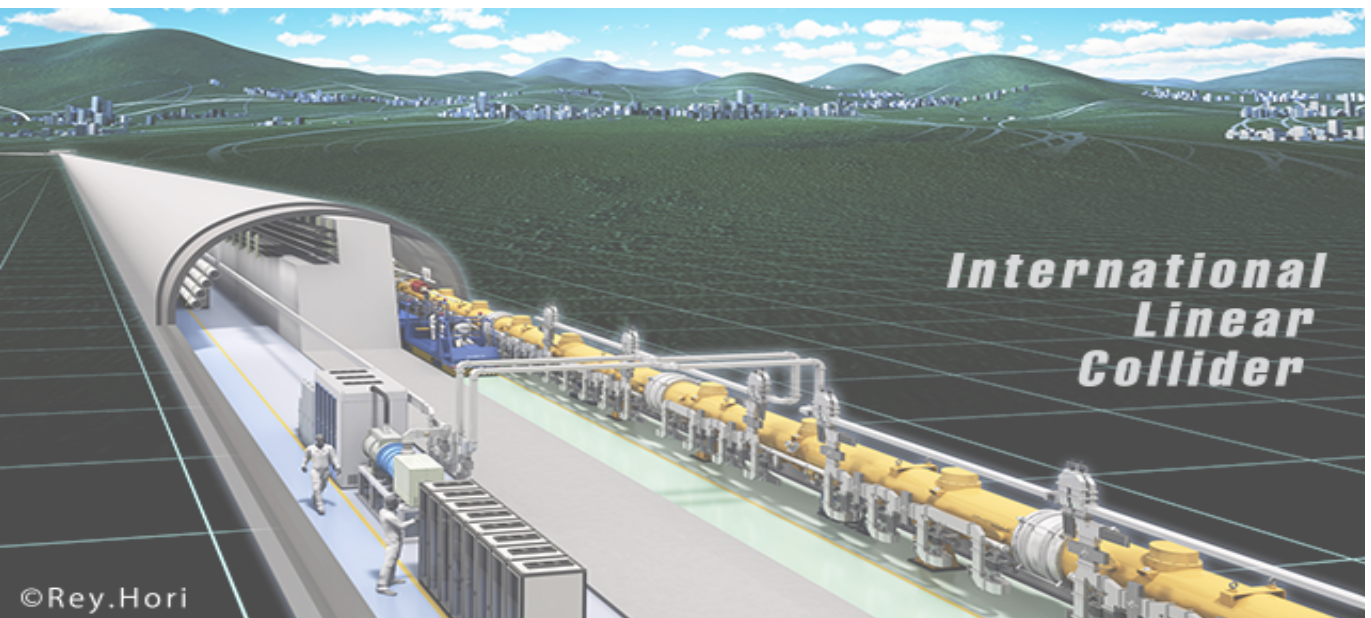
We really need further studies on the newly observed Higgs sector.

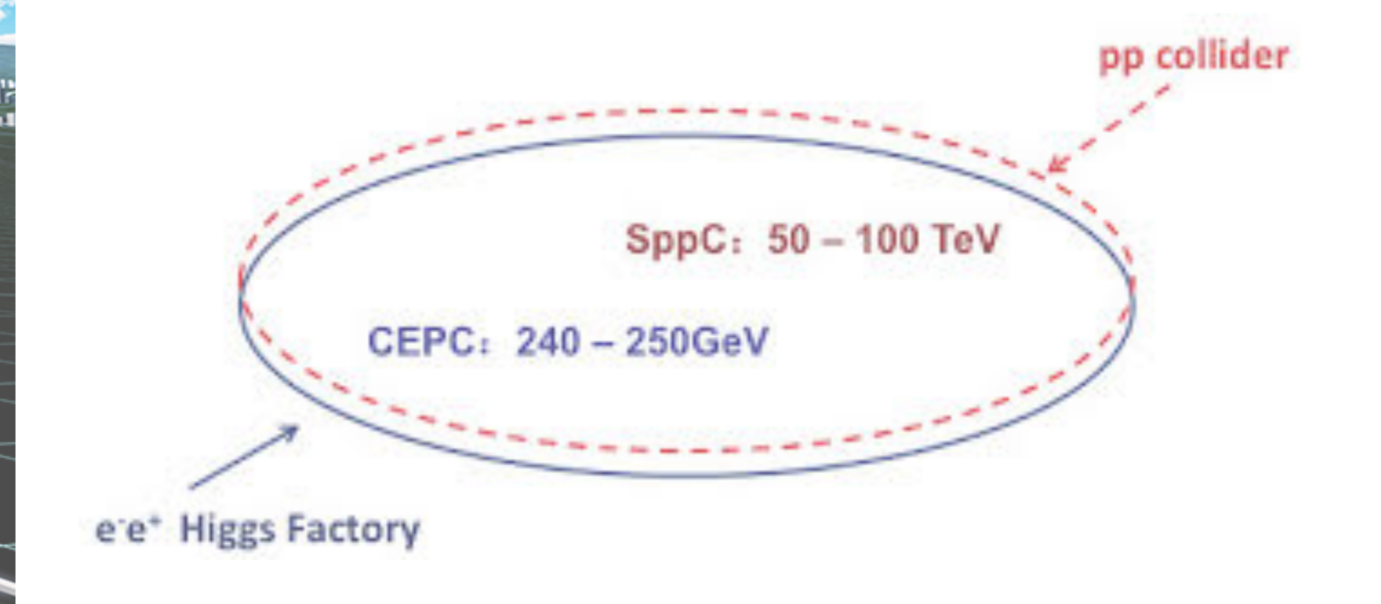
Any small deviations could be a sign of new physics !

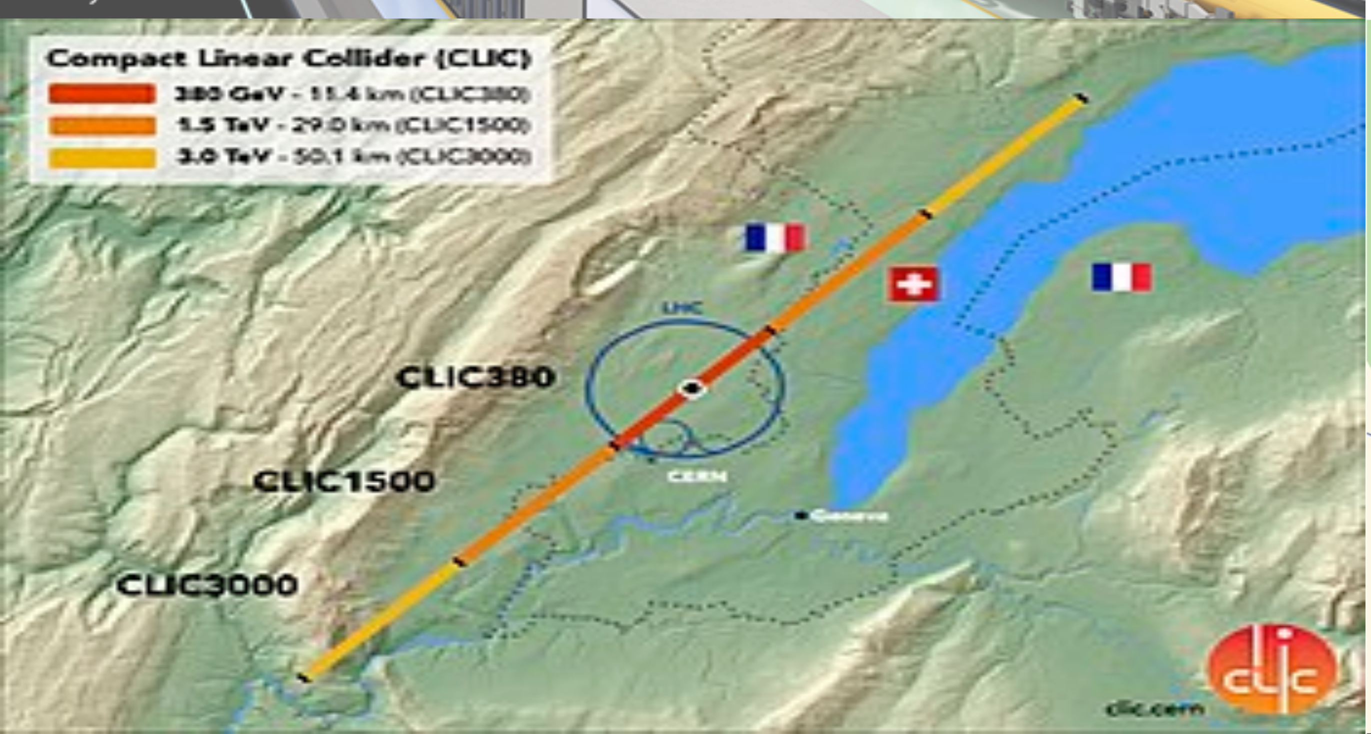
Higgs Precision Era

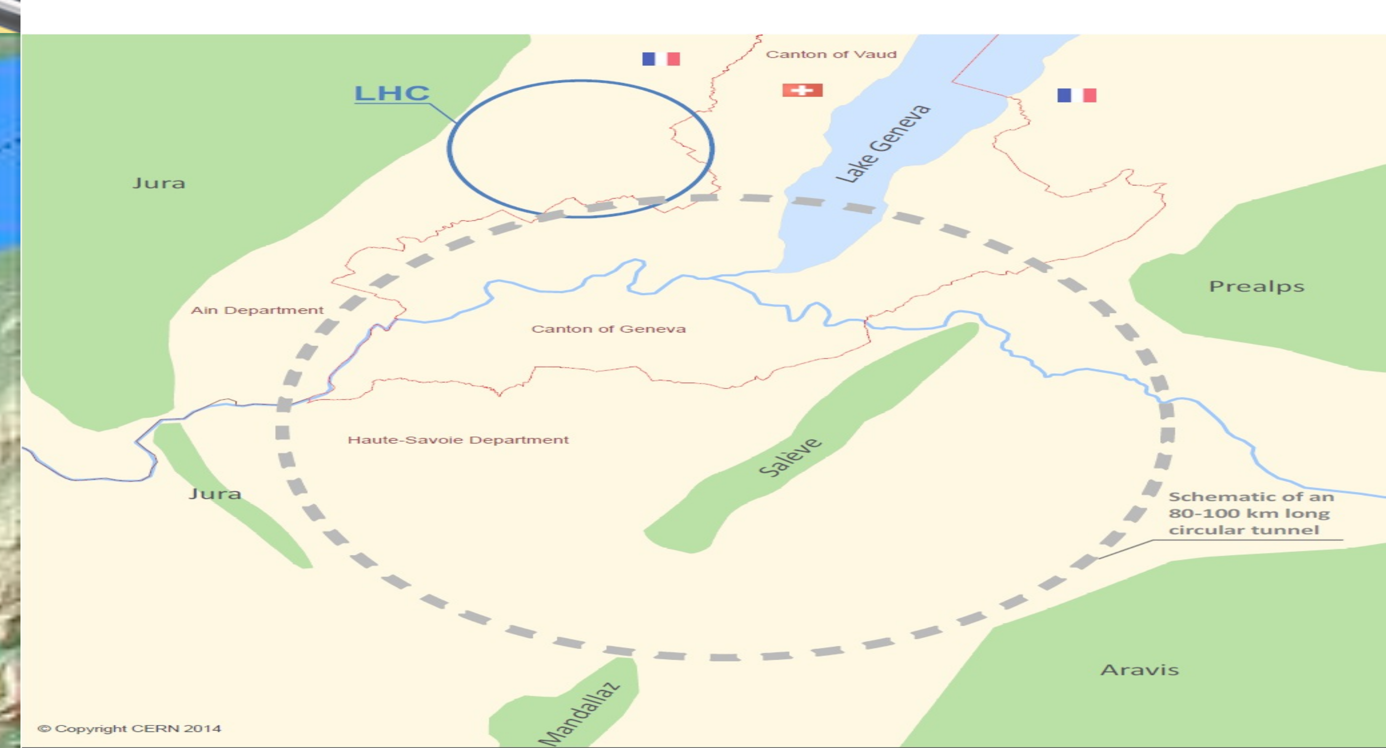
An upgrade of the LHC : **High Luminosity LHC**

Future lepton colliders : **ILC, FCC-ee, CEPC, CLIC**



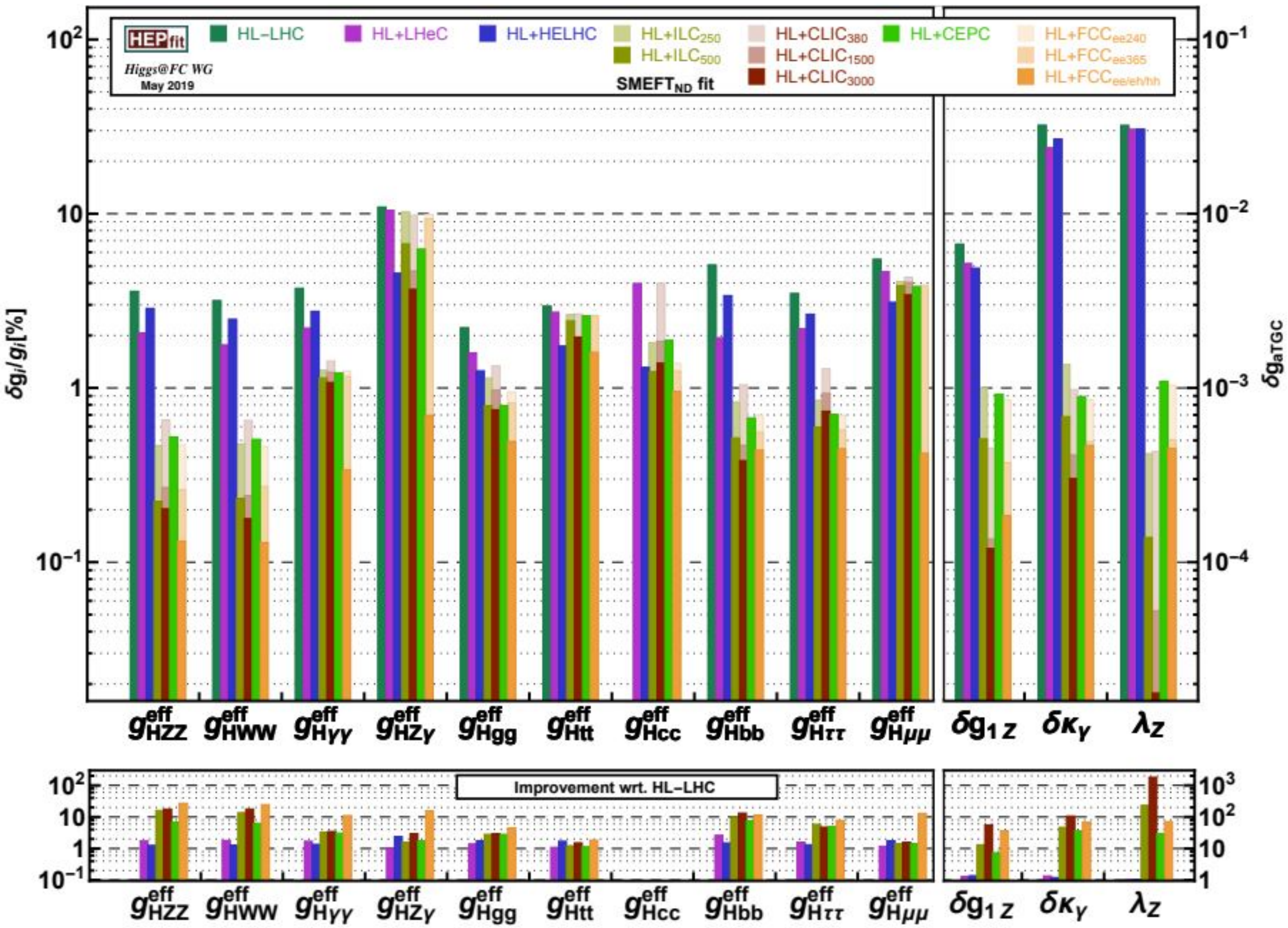






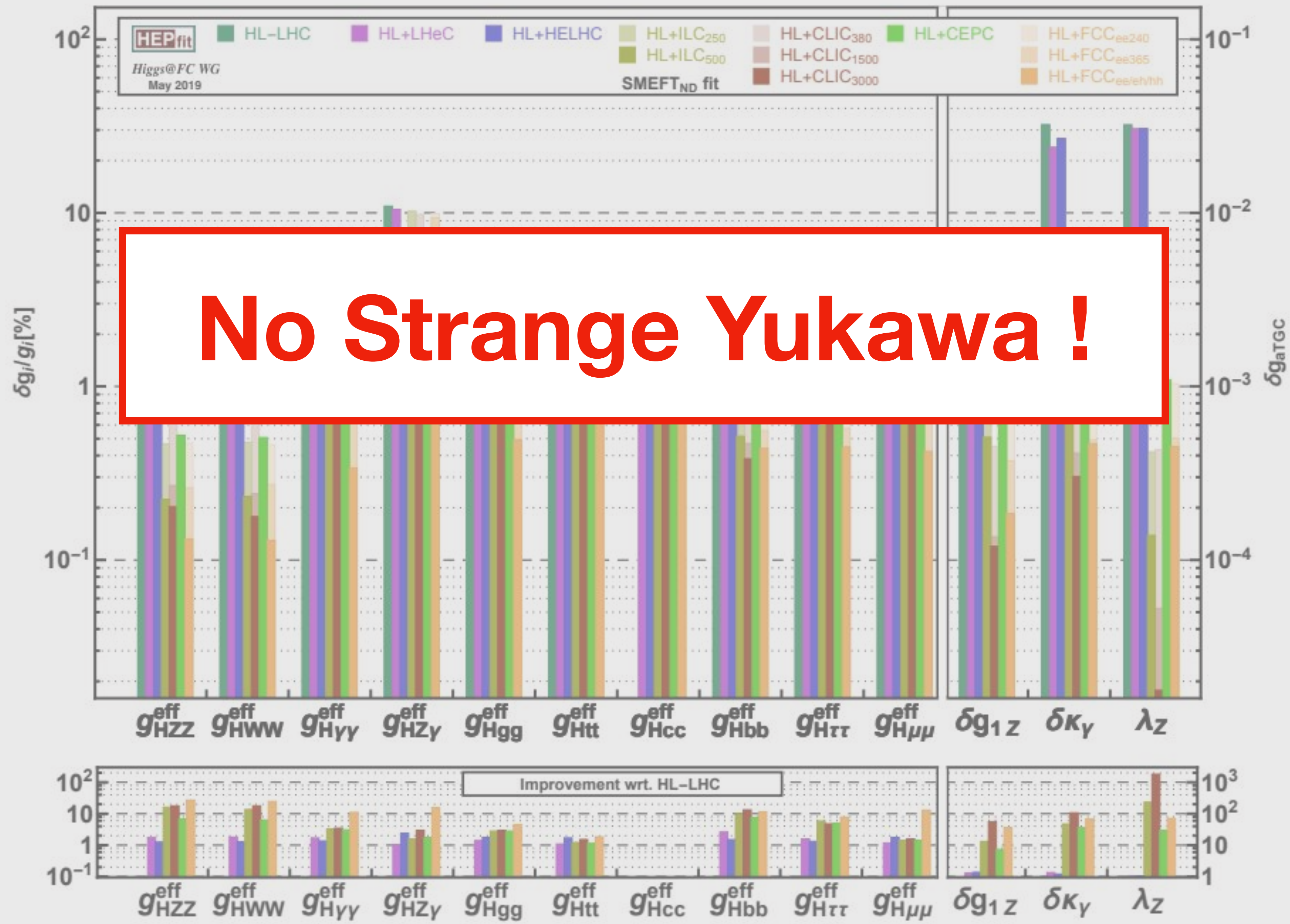
Higgs Precision Era

Sensitivity at 68% probability to deviations in the different effective Higgs couplings



Higgs Precision Era

Sensitivity at 68% probability to deviations in the different effective Higgs couplings

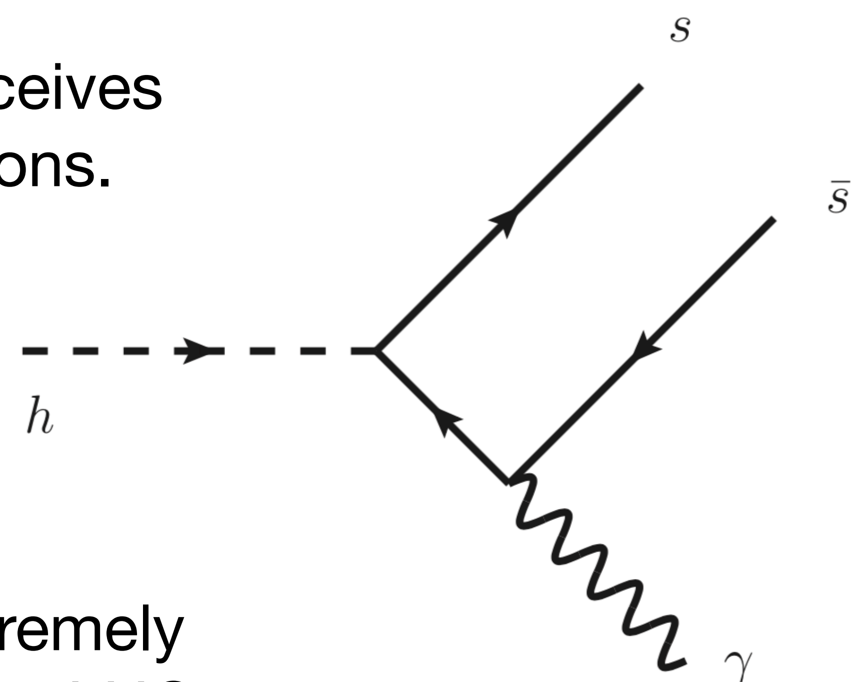


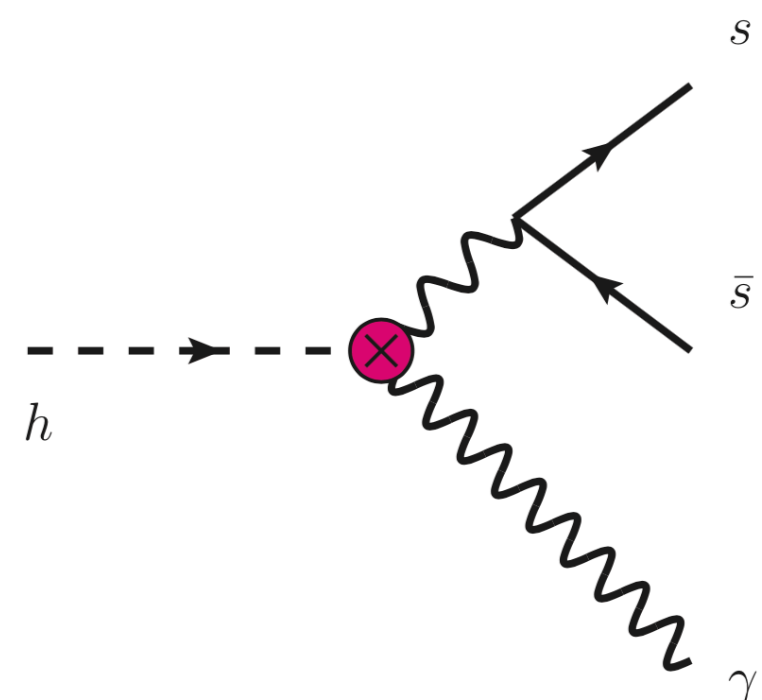
Rare Higgs decay

One way to get access to the strange Yukawa is to focus on the rare decay $h \to \phi\gamma$ (ϕ: a vector meson).

Kagan, Perez, Petriello, Soreq, Stoynev, Zupan (2014)   Konig, Neubert (2015)

The decay amplitude receives two dominant contributions.





The measurement is extremely challenging at the present LHC.

(The branching fraction is in the range of few times 10^{-6}.)

How about HL-LHC ?

HL-LHC can probe O(30) modifications of the strange Yukawa.

Global χ^2 Fit

Another way is a global χ^2 fit to the measured Higgs rates.

Kagan, Perez, Petriello, Soreq, Stoynev, Zupan (2014) Perez, Soreq, Stamou, Tobioka (2015)

The effective Lagrangian

$$\mathcal{L}_{\rm eff}^{\rm Higgs} = \kappa_W \frac{2m_W^2}{v} h W_\mu^+ W^{-\mu} + \kappa_Z \frac{m_Z^2}{v} h Z_\mu Z^\mu - \sum_f \frac{m_f}{v} h \bar{f} (\kappa_f + i\tilde{\kappa}_f \gamma_5) f$$
$$+ \frac{\alpha}{4\pi v}\left(\kappa_{\gamma\gamma} h F_{\mu\nu}F^{\mu\nu} - \tilde{\kappa}_{\gamma\gamma} h F_{\mu\nu}\tilde{F}^{\mu\nu} + \frac{2\kappa_{\gamma Z}}{s_W c_W} h F_{\mu\nu}Z^{\mu\nu} - \frac{2\tilde{\kappa}_{\gamma Z}}{s_W c_W} h F_{\mu\nu}\tilde{Z}^{\mu\nu}\right) + \ldots ,$$

All of the Higgs couplings are allowed to vary from their SM values…

$$\sqrt{|\kappa_u|^2 + |\tilde{\kappa}_u|^2} < 3000 \qquad \sqrt{|\kappa_d|^2 + |\tilde{\kappa}_d|^2} < 1500$$

$$\sqrt{|\kappa_c|^2 + |\tilde{\kappa}_c|^2} < 6.2 \qquad \sqrt{|\kappa_s|^2 + |\tilde{\kappa}_s|^2} < 75$$

(95% CL)

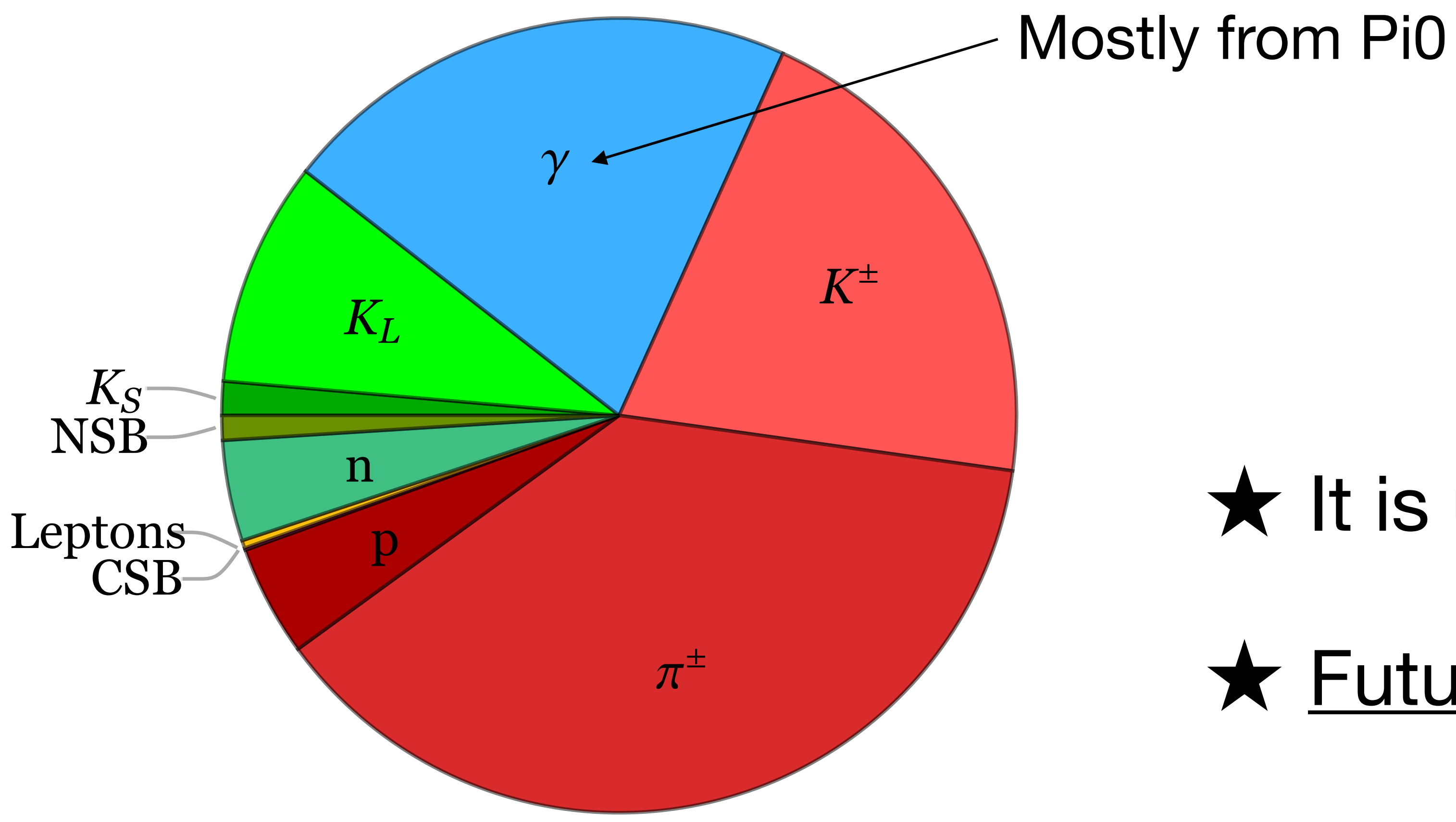
The present LHC data are largely insensitive to the light quark Yukawa.

Can we test the SM strange Yukawa?

Strange tagging is essential.

The pT fraction of a detector-stable particle averaged over jet samples :

$$Z \to s\bar{s} \quad (p_T > 20 \text{ GeV})$$



★ It is unlikely at the LHC.

★ Future lepton colliders can do it?

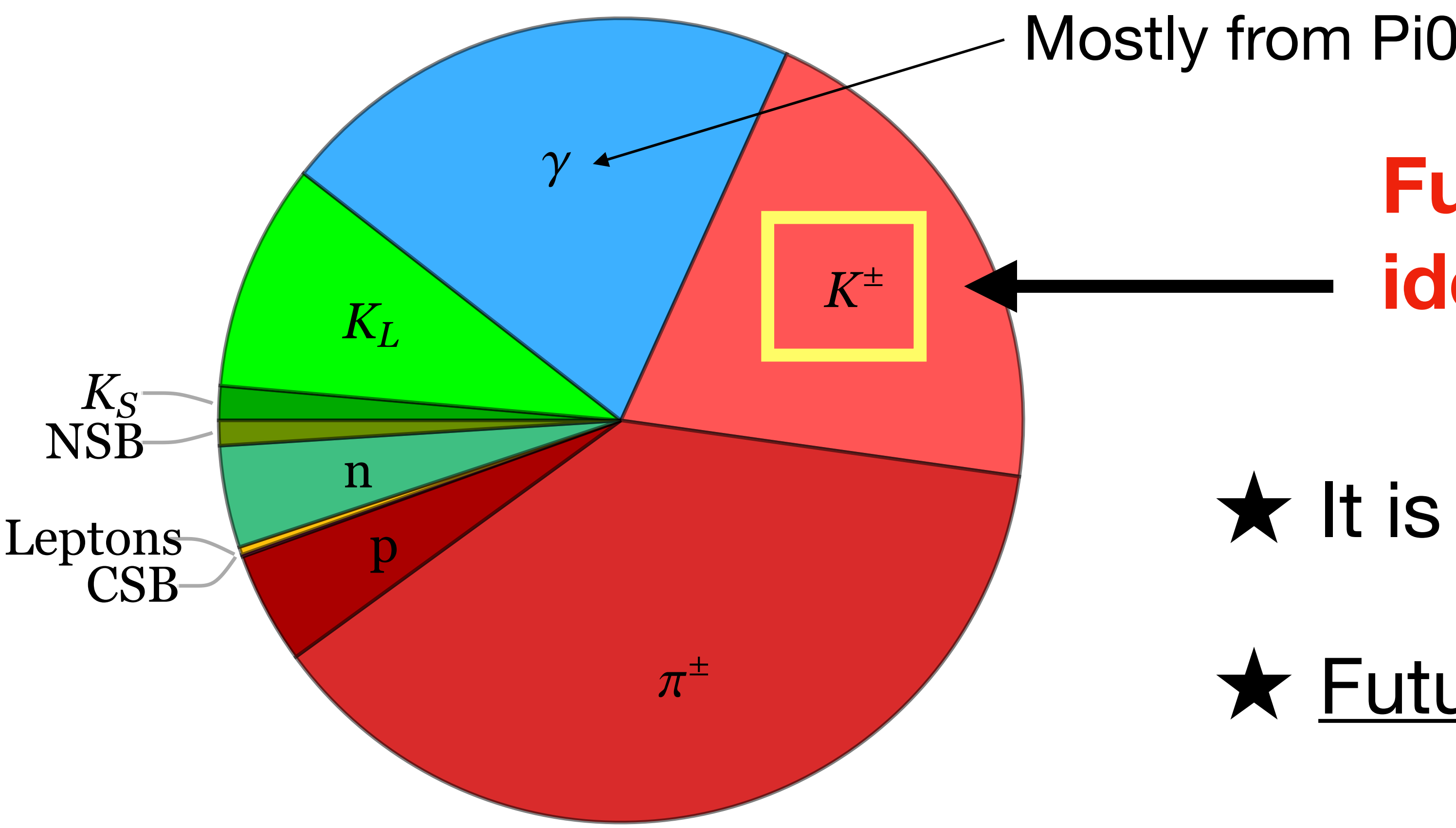
NSB: neutral strange baryons, CSB: charged strange baryons

Can we test the SM strange Yukawa?

Strange tagging is essential.

The pT fraction of a detector-stable particle averaged over jet samples :

$$Z \to s\bar{s} \quad (p_T > 20 \text{ GeV})$$



Future lepton colliders can identify charged Kaons !

★ It is unlikely at the LHC.

★ Future lepton colliders can do it?

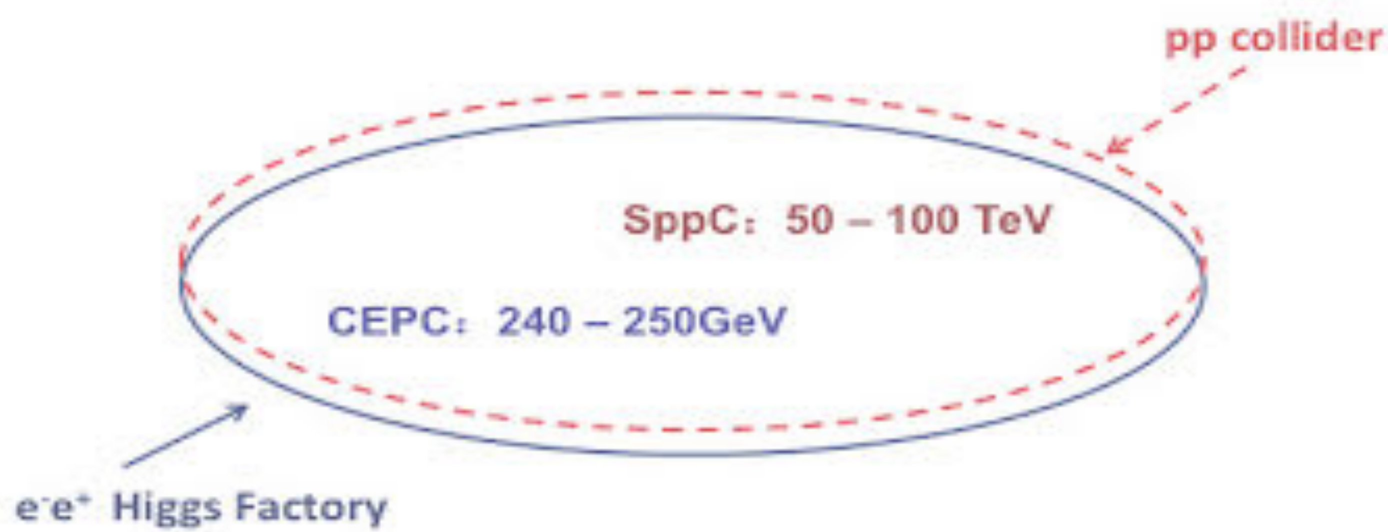
NSB: neutral strange baryons, CSB: charged strange baryons

CEPC

✓ Circular Electron-Positron Collider (CEPC) proposed to be built in China.

✓ CEPC will operate as a Higgs boson factory at center-of-mass energy of around 240 GeV.

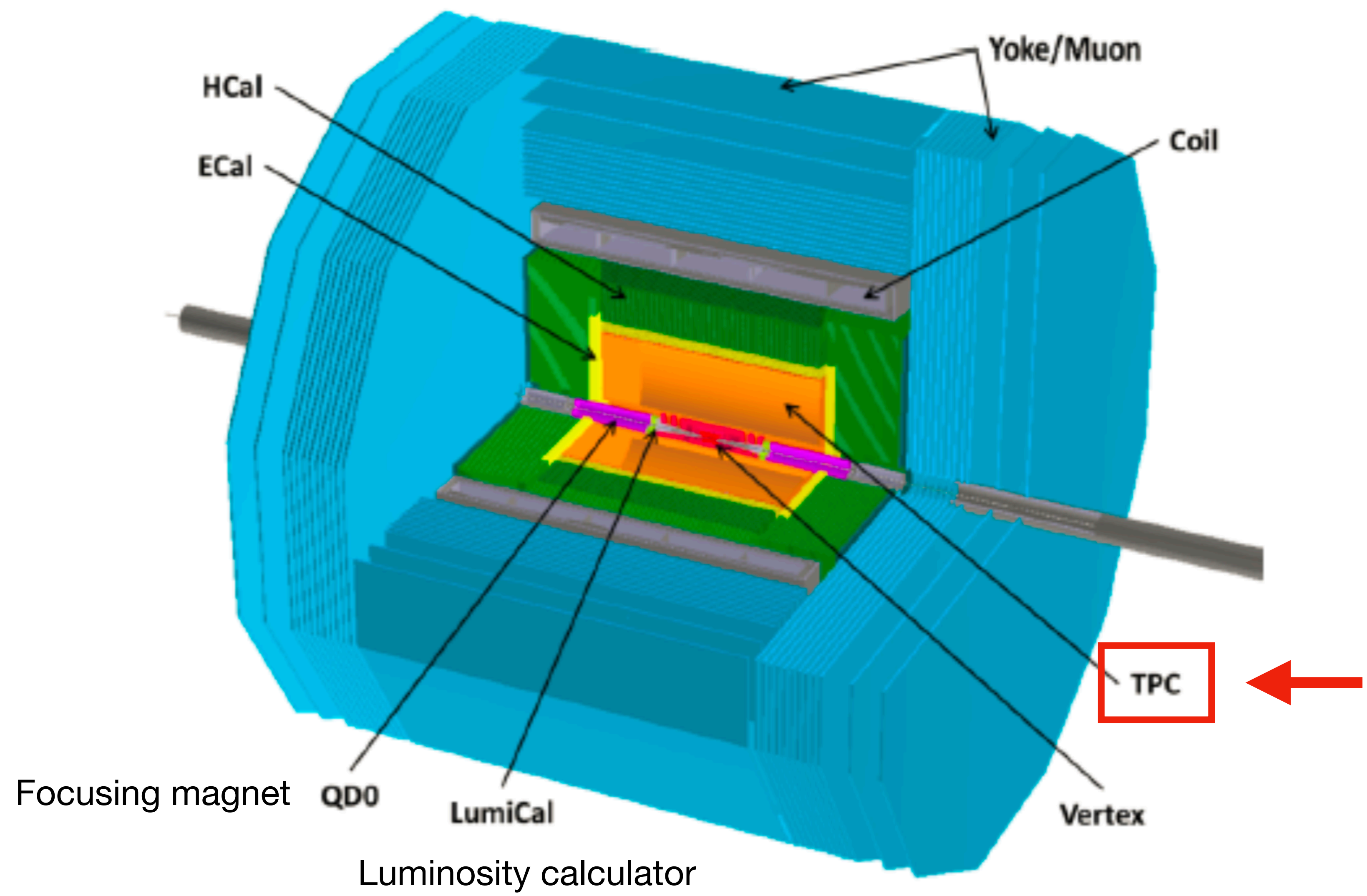
✓ During its lifetime, one million Higgs bosons are expected to be produced, allowing precision measurements of the Higgs boson properties.



✓ The same tunnel could also host a Super Proton Proton Collider (SppC) to reach energies beyond the LHC.

CEPC Detector Design

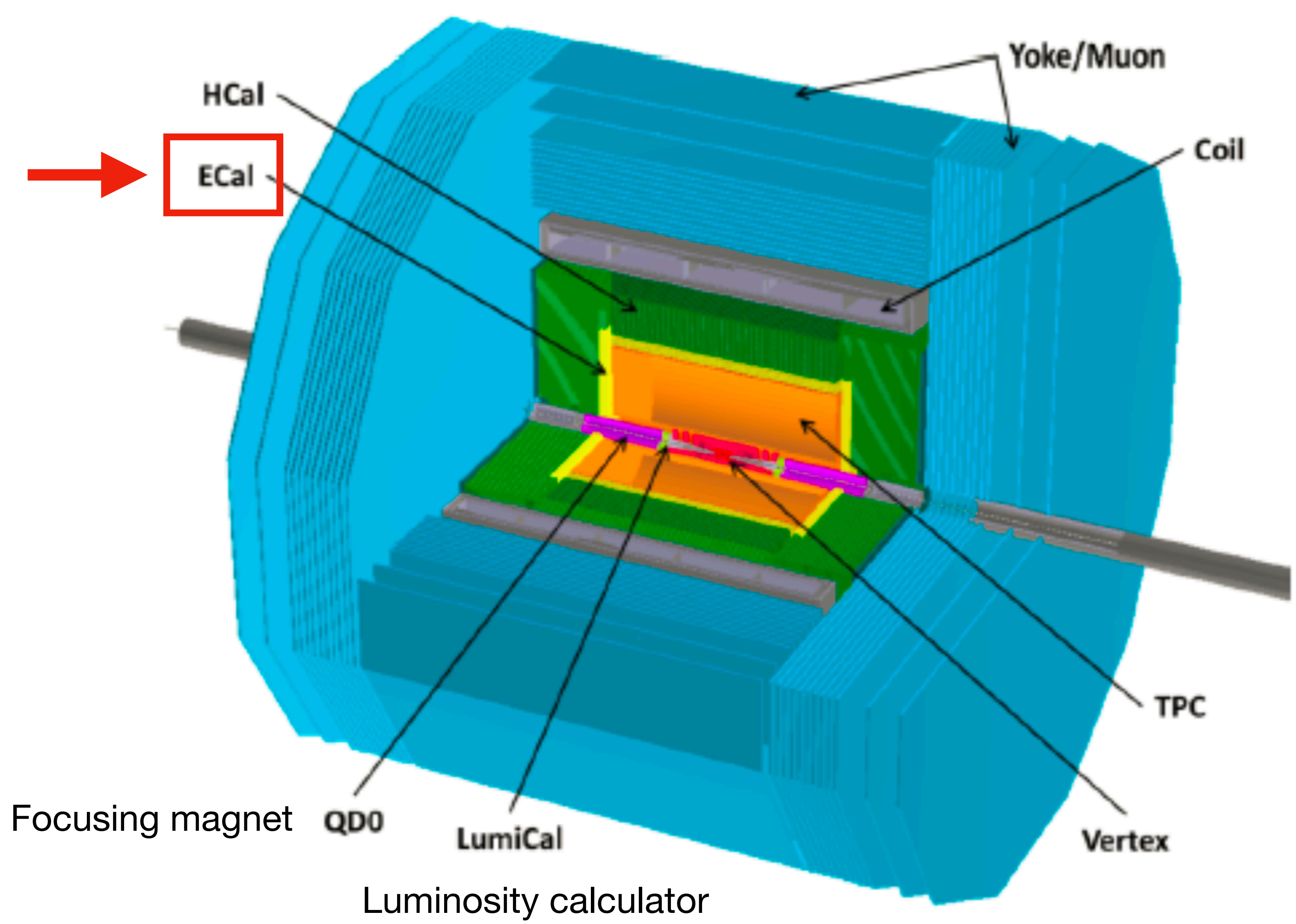
- ✓ Time Projection Chamber (TPC) is proposed as <u>a charged particle tracking device</u>.
- ✓ TPC provides <u>precise momentum and position measurements</u> and a good <u>particle identification (PID)</u> over a wide range of momentum.
- ✓ PID is based on measurements of <u>*dE/dx* (energy deposit per unit path length)</u>.



CEPC Detector Design

In addition…

✓ Electromagnetic Calorimeter (ECAL) provides <u>time-of-flight (TOF) information</u>.



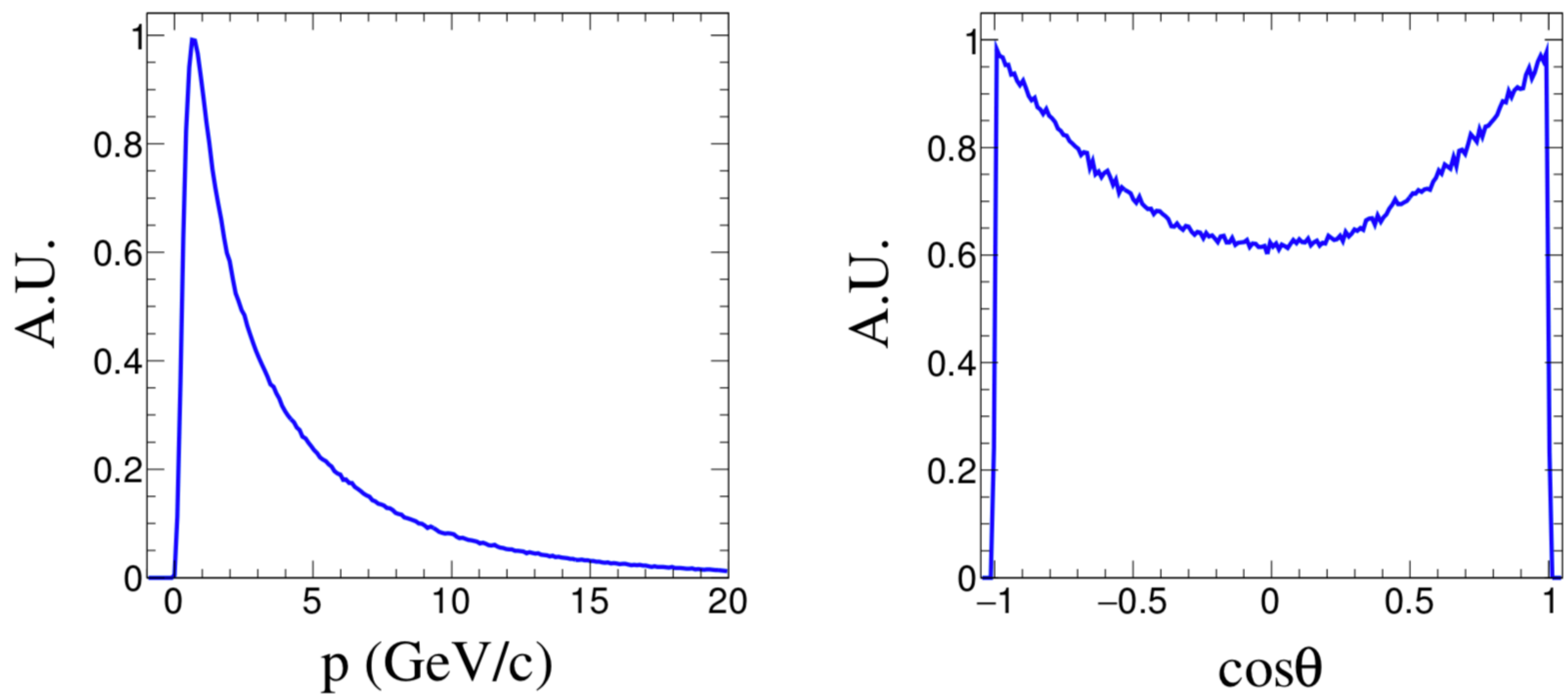
Charged Kaon ID

✓ **Charge kaons can be identified** by combining the information of TPC with TOF of ECAL.

✓ PID of kaons, pions and protons in hadronic decays at the *Z* pole has been studied.

An, Prell, Chen, Cochran, Lou, Ruan (2018)

Kinematic distribution of kaons in $e^+e^- \to Z \to qq$



Polar angle of the tracks with respect to the beam direction

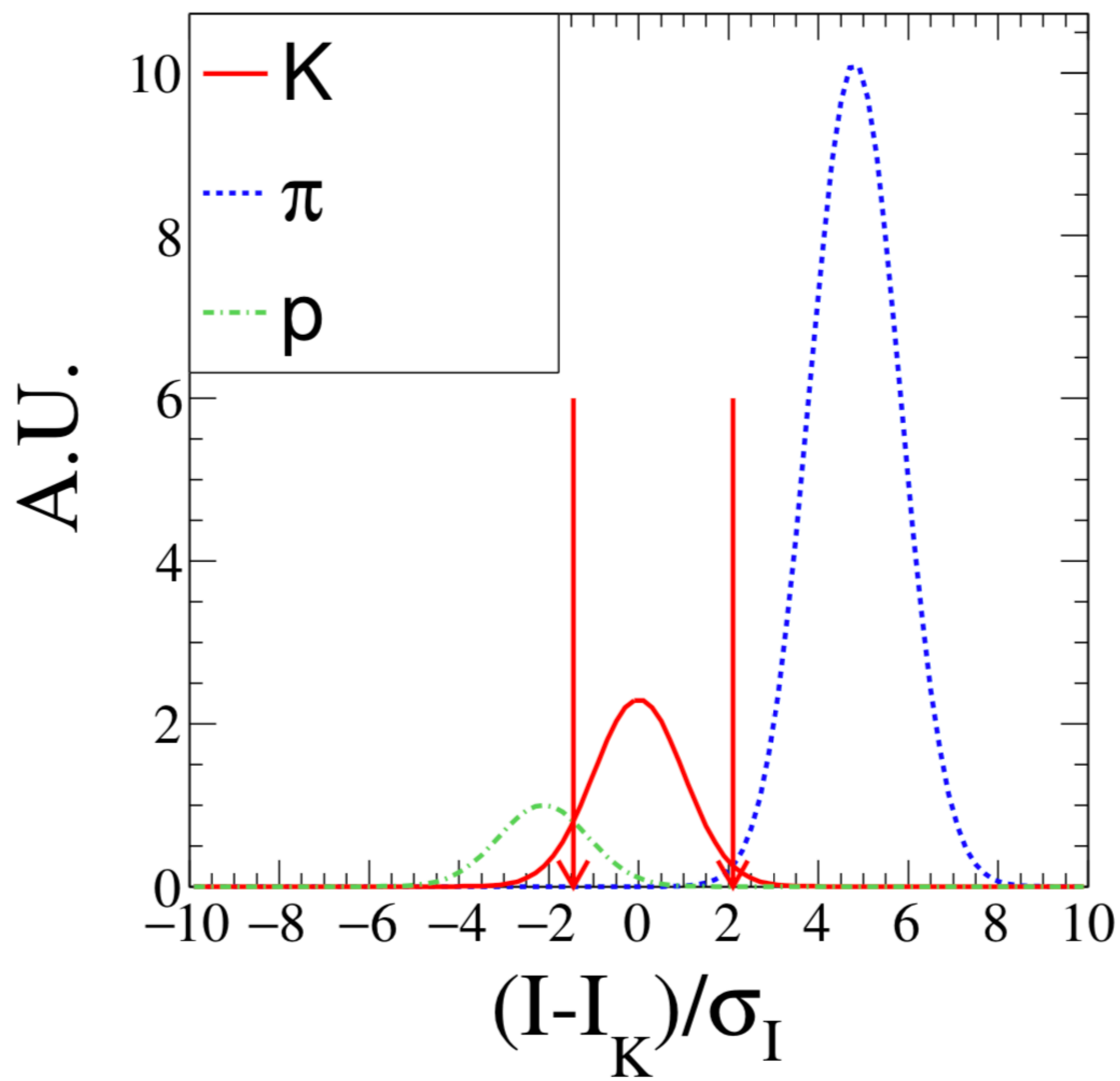
Charged Kaon ID

Measure of separation power between particles A and B :

$$S_{AB} = \frac{|I_A - I_B|}{\sqrt{\sigma_{I_A}^2 + \sigma_{I_B}^2}}$$

I_A (I_B) : average *dE/dx* measurement of particle A (B)

σI_A (σI_B) : the corresponding resolution



An, Prell, Chen, Cochran, Lou, Ruan (2018)

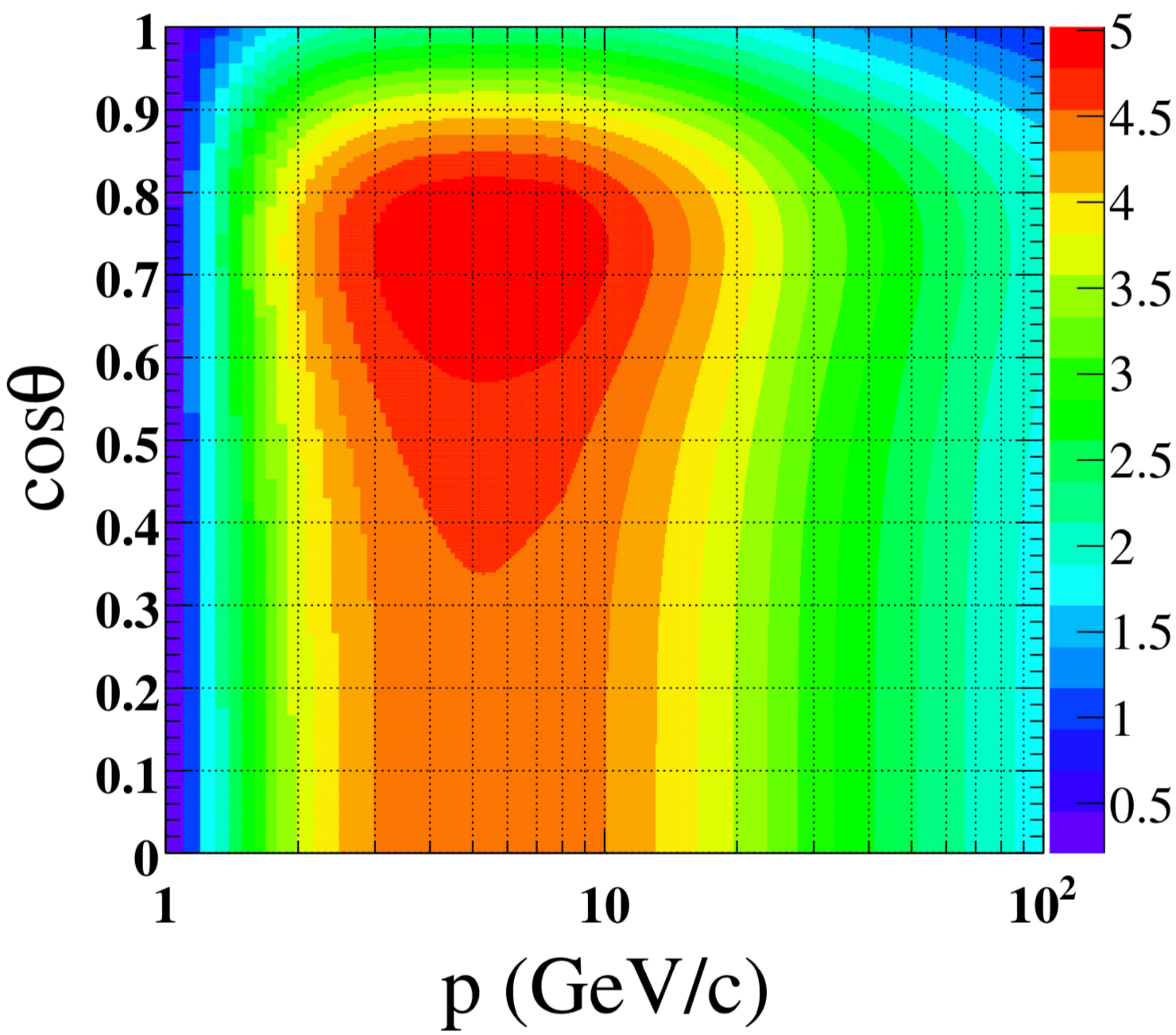
Particles with a momentum of 5 GeV

The relative populations :

$N_\pi = 4.4\, N_K, \quad N_K = 2.3\, N_p$

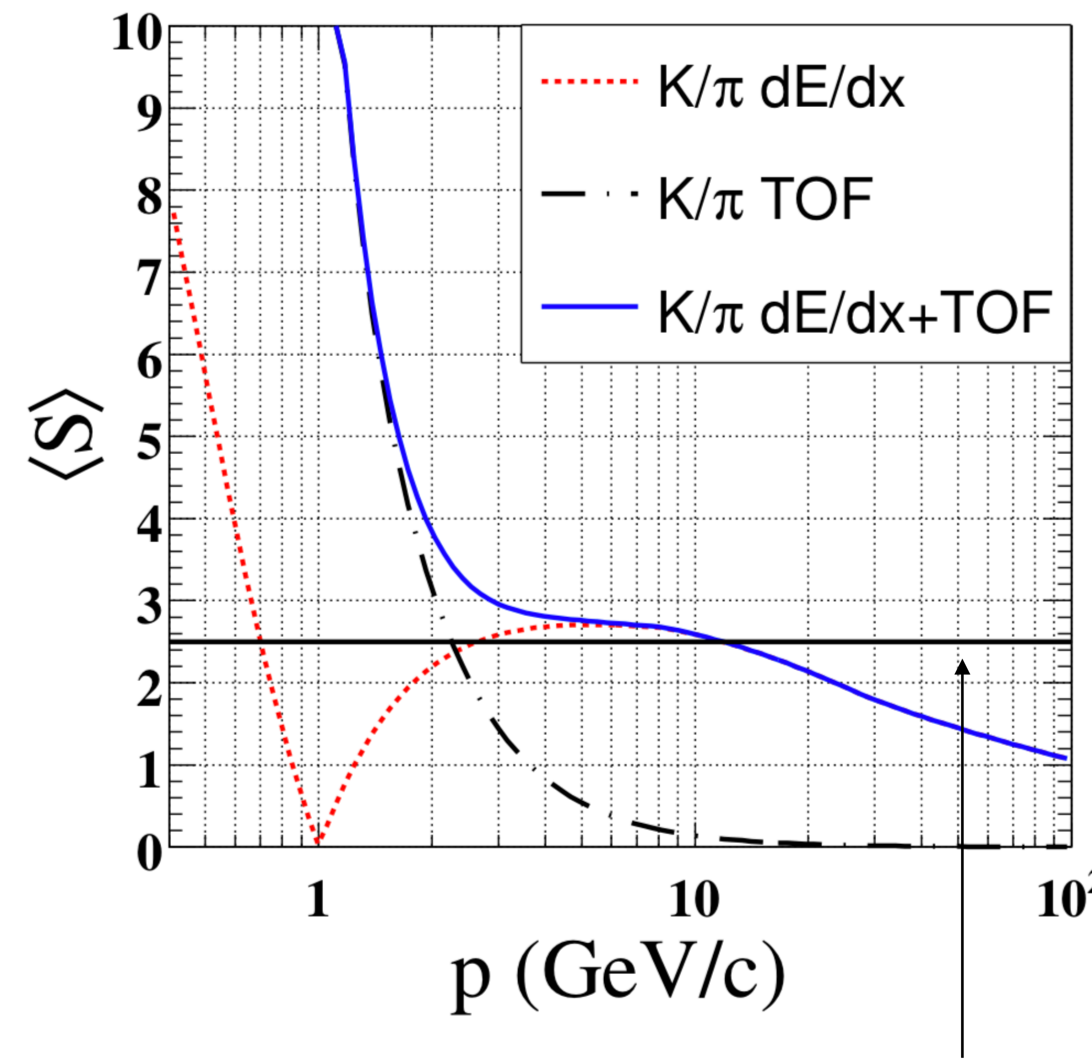
Charged Kaon ID

Separation power S between kaons and pions in the *p*-cosθ plane :



An, Prell, Chen, Cochran, Lou, Ruan (2018)

K/π separation using *dE/dx* and/or TOF :



CEPC can identify charged Kaons with momenta p < 20 GeV !

Prospect of Strange Yukawa

Probe the strange Yukawa by tagging strange jets in future lepton colliders.

Duarte-Campderros, Perez, Schlaffer, Soffer (2018)

$e^+ e^- \to Z h$

1. Separate $h \to jj$ from all non-$h \to jj$ events (preselection).

2. Apply a flavor tag on the selected signal-rich sample.

$Z \to \nu\nu$ has a large branching fraction of 20% and a clean, missing-energy signature that provides good rejection of non-Higgs background and Higgs decays into non-jj final states. (→ preselection)

Non-h → jj background events and their percentages after preselection :

$e^+e^- \to$ Final state	WW $(\tau\nu)(qq')$	$Z(Z+\gamma^*)$ $(\nu\nu)(dd,\ ss,\ bb)$	$Zh+\nu\nu h$ $(\nu\nu)(\text{non-}jj)$	$Z(Z+\gamma^*)$ $(\nu\nu)(uu,\ cc)$	Zh $(\tau\tau)(bb)$	Zh $(qq)(\text{non-}jj)$	WW $(\mu\nu)(qq')$
Fraction [%]	47.1	18.0	13.7	12.2	2.7	2.5	2.0

Prospect of Strange Yukawa

Probe the strange Yukawa by tagging strange jets in future lepton colliders.

Duarte-Campderros, Perez, Schlaffer, Soffer (2018)

$e^+ e^- \rightarrow Z h$

1. Separate $h \rightarrow jj$ from all non-$h \rightarrow jj$ events (preselection).

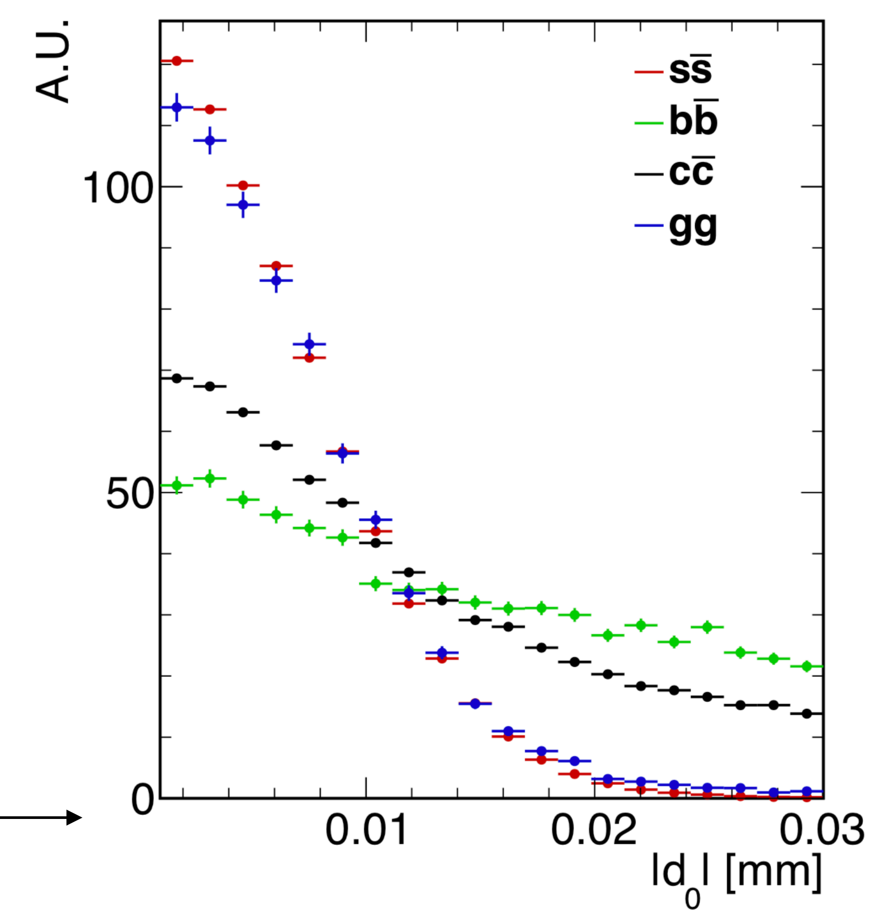
2. Apply a flavor tag on the selected signal-rich sample.

Relevant $h \rightarrow jj$ background decays :

$h \rightarrow bb$, $h \rightarrow cc$, $h \rightarrow gg$

Look for a displaced (secondary) vertex.

d0 : impact parameter



Prospect of Strange Yukawa

Duarte-Campderros, Perez, Schlaffer, Soffer (2018)

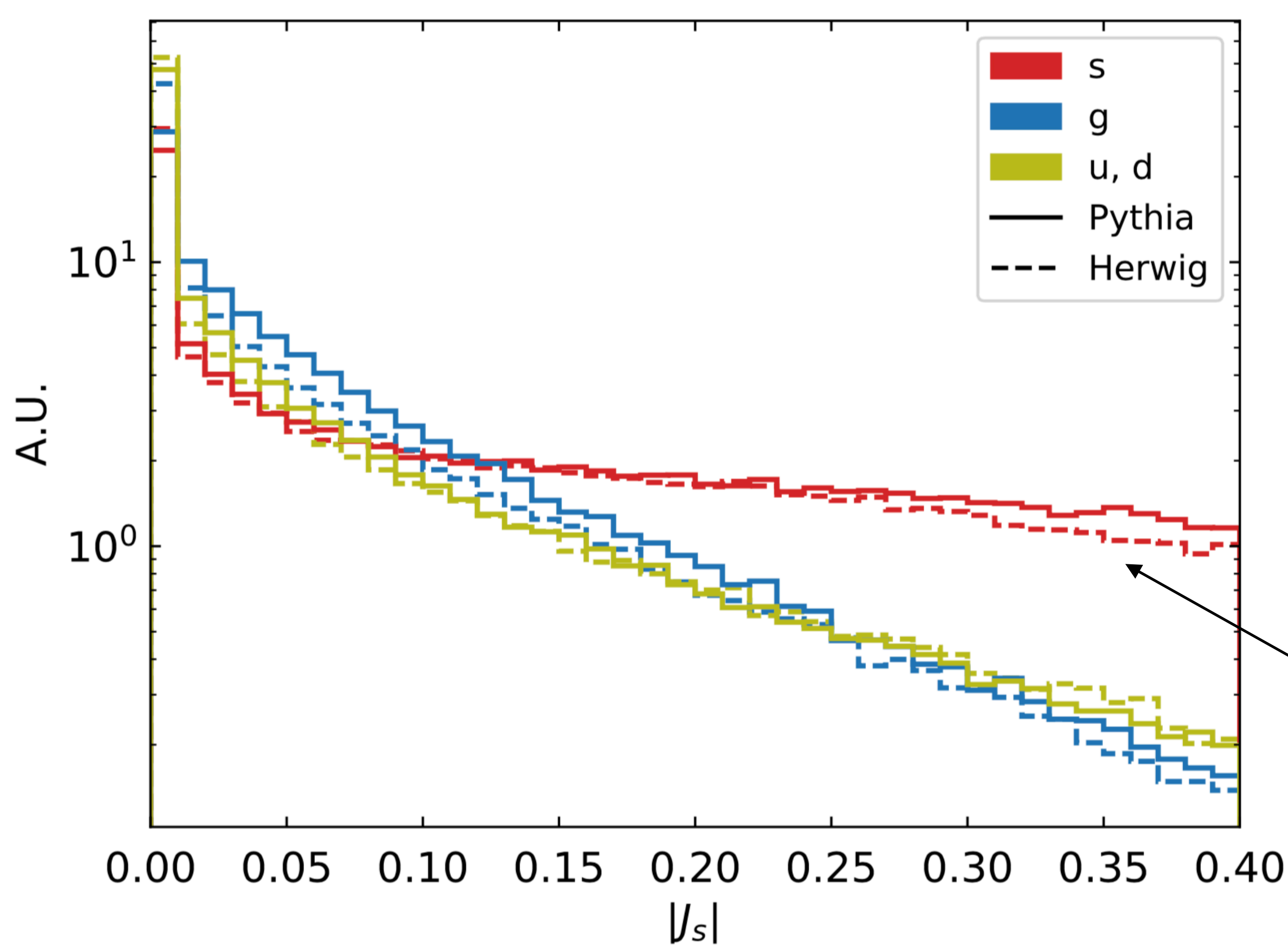
A new jet-flavor variable :

$$J_F = \frac{\sum_H \vec{p}_H \cdot \hat{s}\, R_H}{\sum_H \vec{p}_H \cdot \hat{s}}$$

$R_{K^\pm} = \mp 1$, $R_H = 0$ for $H = \pi^\pm, \pi^0$

Normalized jet axis

All hadrons inside the jet

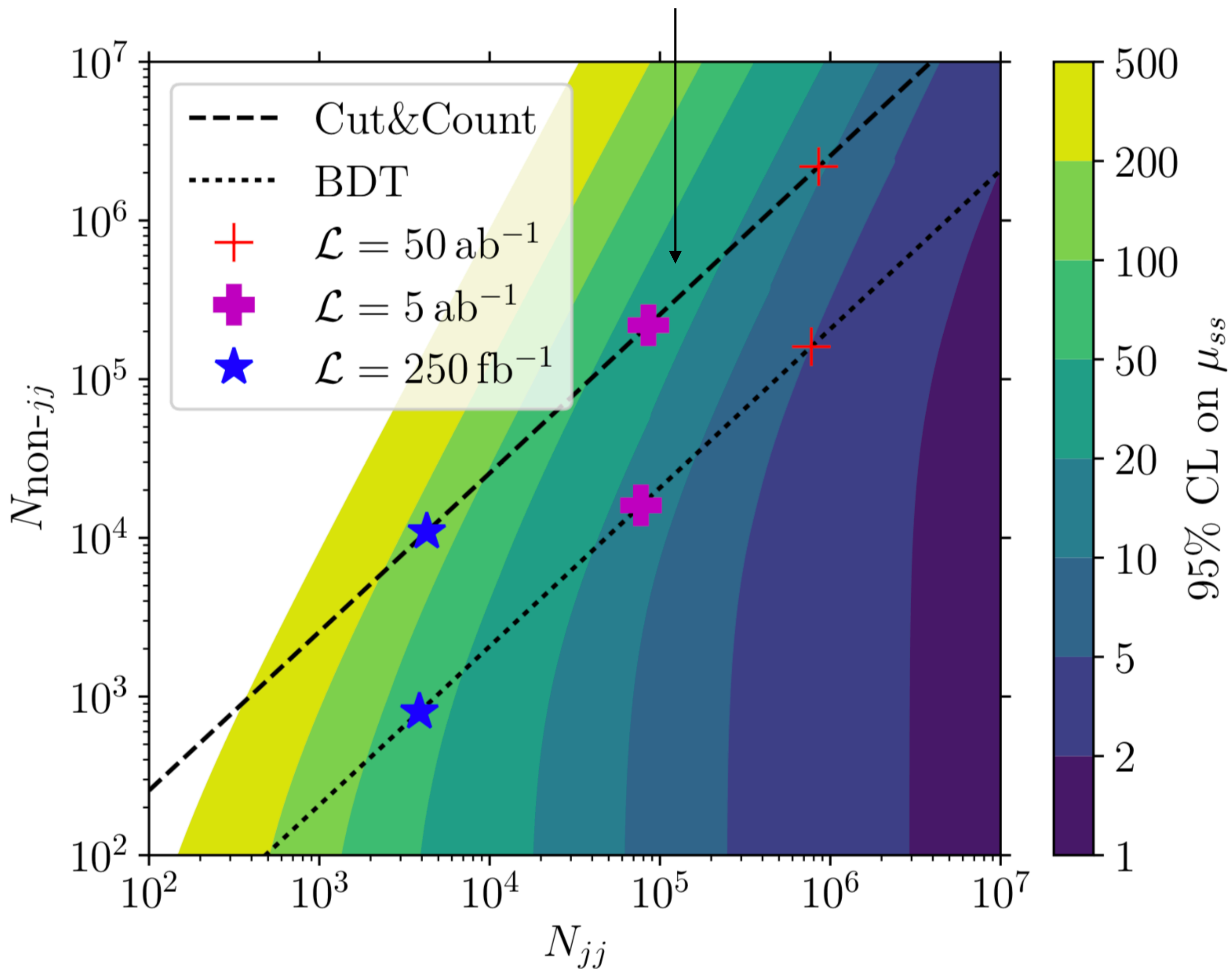


If a K^- (K^+) carries all of the momentum, the $K^\mp$ would have $J_s = \pm 1$.

Distribution of h → ssbar events is broad.

Prospect of Strange Yukawa

The number of non-h → jj events (Nnon−jj) vs. h → jj events (Njj) after preselection but before the s-tagger



Duarte-Campderros, Perez, Schlaffer, Soffer (2018)

The limit on the signal strength after the s-tagger

$$\mu_i \equiv (\sigma \times \mathrm{BR})_i / (\sigma \times \mathrm{BR})_i^{\mathrm{SM}}$$

μss < O(15) and O(5) for integrated luminosities of 5 and 50 ab^{-1}

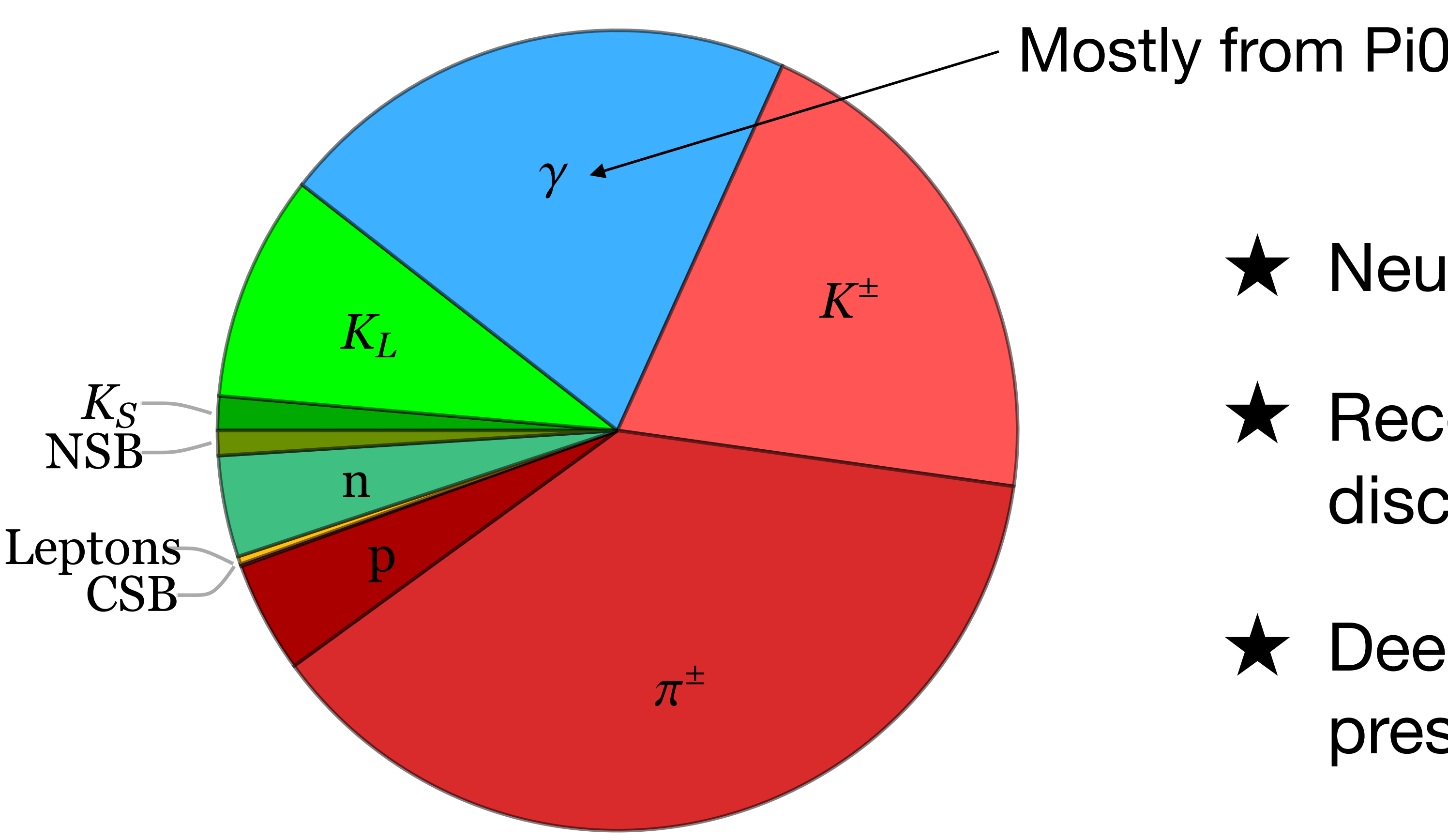
μss < O(75) for an integrated luminosity of 250 fb^{-1}

Can we test the SM strange Yukawa?

Machine learning can help to improve the limit?

The pT fraction of a detector-stable particle averaged over jet samples :

$$Z \to s\bar{s} \quad (p_T > 20 \text{ GeV})$$



★ Neutral Kaon is useful?

★ Recent development of quark/gluon discrimination is useful?

★ Deep learning can improve preselection?

...

NSB: neutral strange baryons, CSB: charged strange baryons