Tagging with the Lund Jet Plane

Charanjit Kaur Khosa

MANCHESTER 1824

First Lund Jet Plane Institute - CERN
6th July, 2023

In Collaboration With
L.Cavallini, A.Coccaro, O.Fedkevych, F. Giuli, G.Manco, S.Marzani, F.Parodi, D.Rebuzzi, A.Rescia, F. Sforza And G.Stagnitto

Outline

- W tagging — F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064
- Higgs tagging
  - Lund jet plane and jet color ring — CKK and S.Marzani, Phys. Rev. D 104 (2021), 055043
  - Lund jet plane and color sensitive observables — Cavallini *et al.*, Eur. Phys. J. C 82 (2022) no.5, 493
- B-tagging — O.Fedkevych, CKK, S. Marzani and F. Sforza, Phys. Rev. D 107(2023) no.3, 034032
- Top tagging — F.A.Dreyer, and H.Qu, JHEP 03 (2021), 052
- Quark-Gluon tagging — F.A.Dreyer, G. Soyez and A.Takacs, JHEP 08 (2022), 177

Boosted Jets

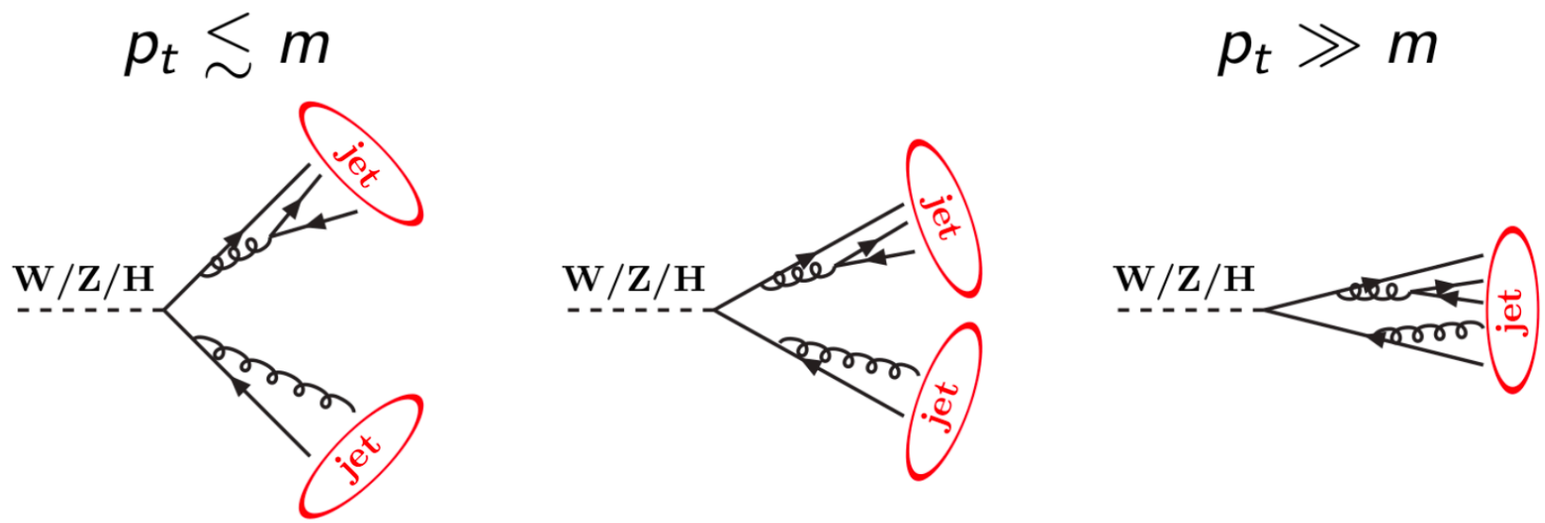
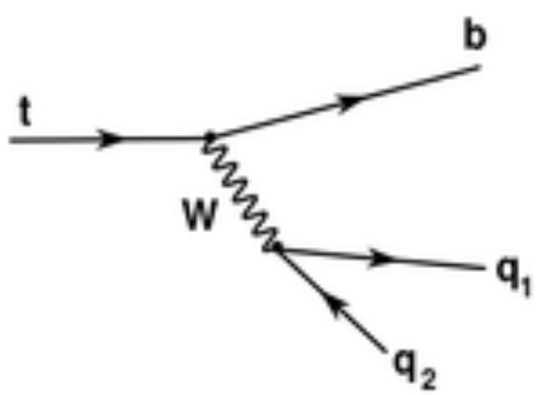


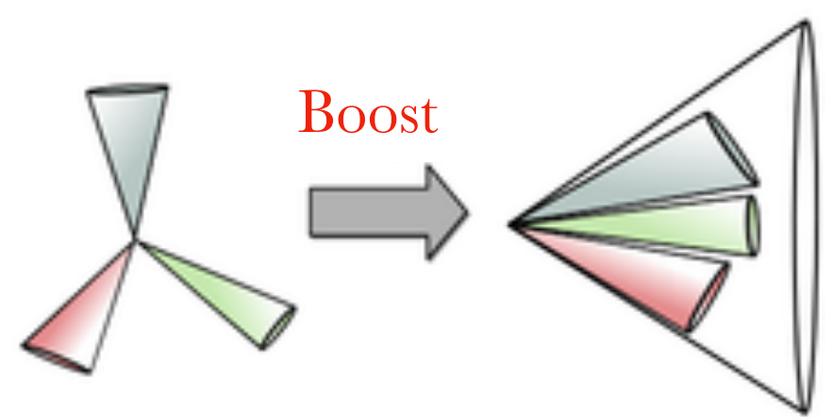
Figure taken form F.Dreyer's slide, original figure credit: G.Soyez

2-jets

Adapted from: https://www.physi.uni-heidelberg.de



Top Quark Decay

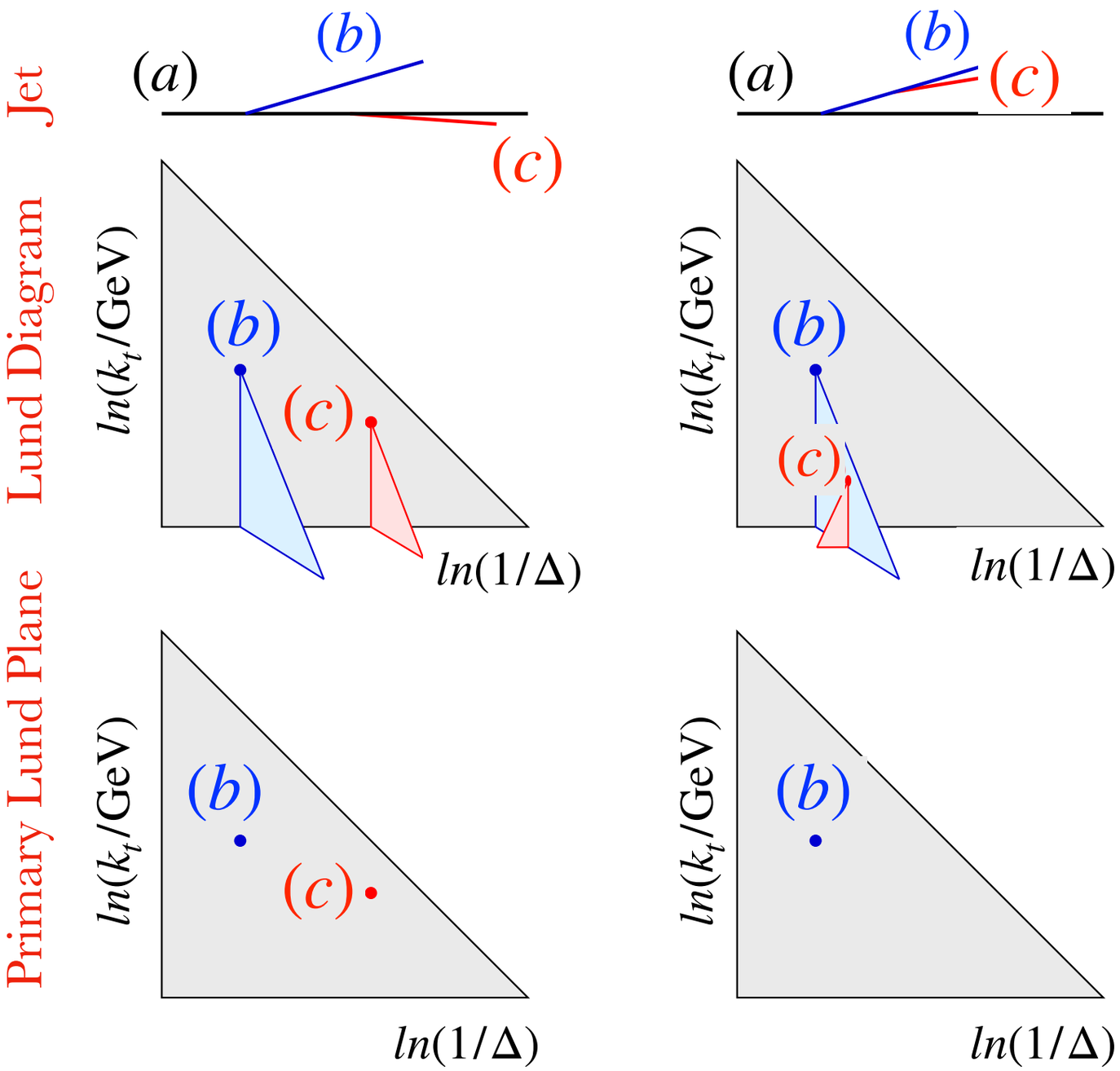


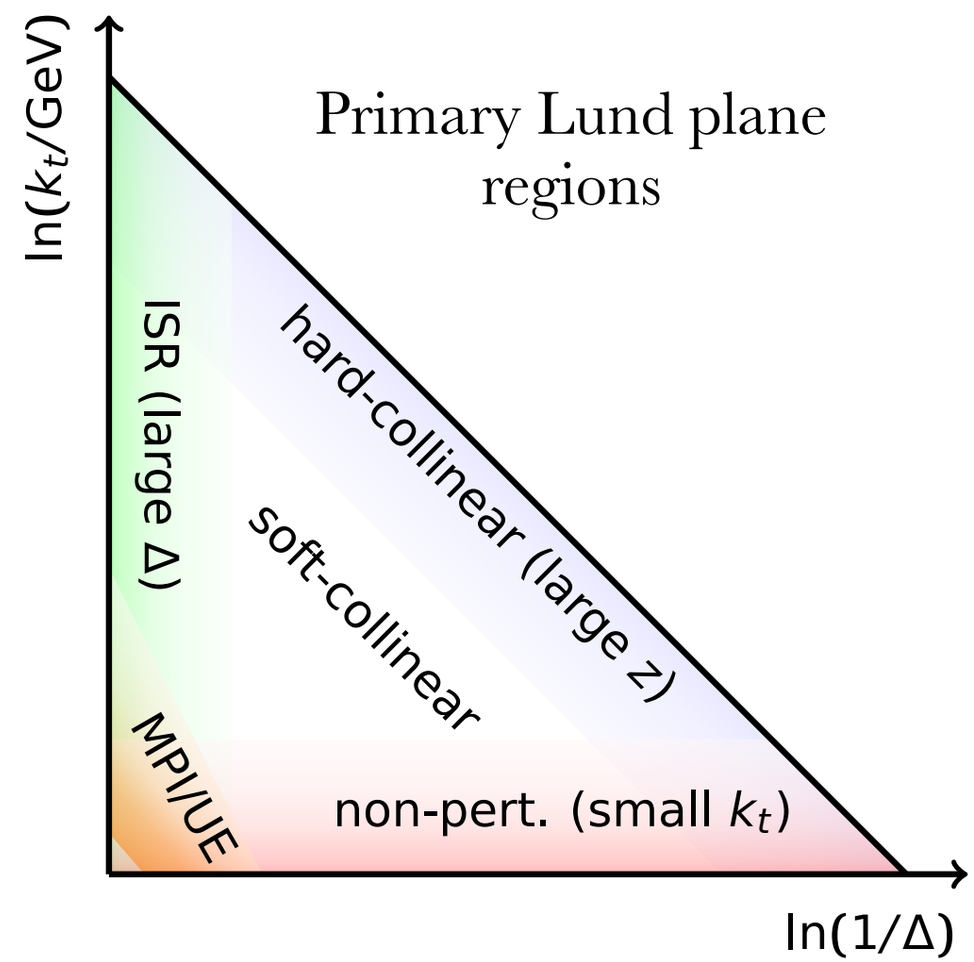
At rest

Large Radius Jet

Jet substructure

Lund Jet Plane





Δ is the opening angle of the emission.

k_T is the transverse momentum w.r.t the jet axis.

F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064

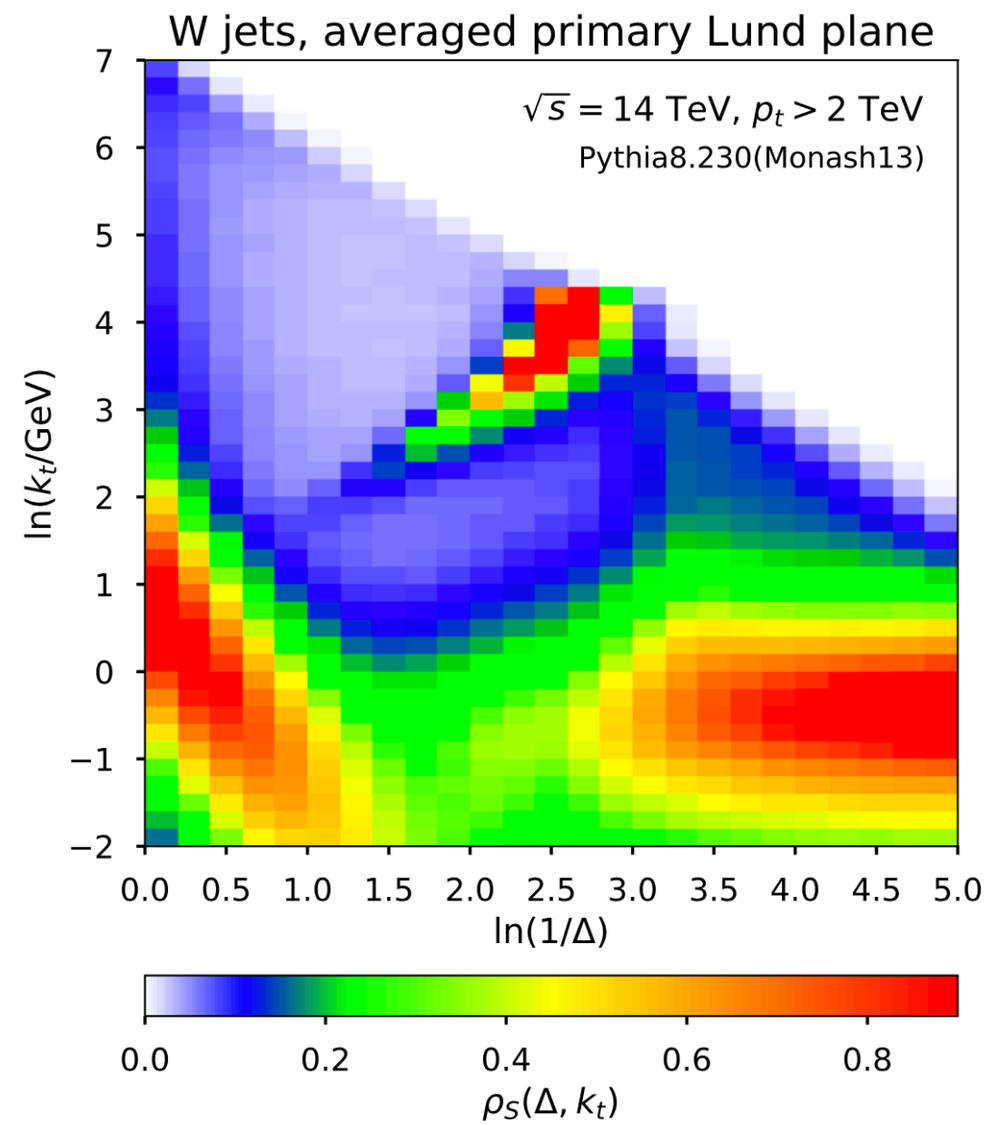
(Primary) Lund Jet Plane

- Re-cluster the constituents using C/A algorithm.
- Start with declustering C/A jet (undo last clustering step) we get two subjects (ordered by momentum).
- Save the kinematic of each declustering step as a tuple $\mathcal{T} = \{k_t, \Delta, z, m, \psi\}$

$$\Delta^2 \equiv (y_1 - y_2)^2 + (\phi_1 - \phi_2)^2 \qquad k_t \equiv p_{t2}\Delta$$

$$z \equiv \frac{p_{t2}}{p_{t1} + p_{t2}} \qquad m^2 \equiv (p_1 + p_2)^2 \qquad \psi \equiv \tan^{-1}\frac{y_2 - y_1}{\phi_2 - \phi_1}$$

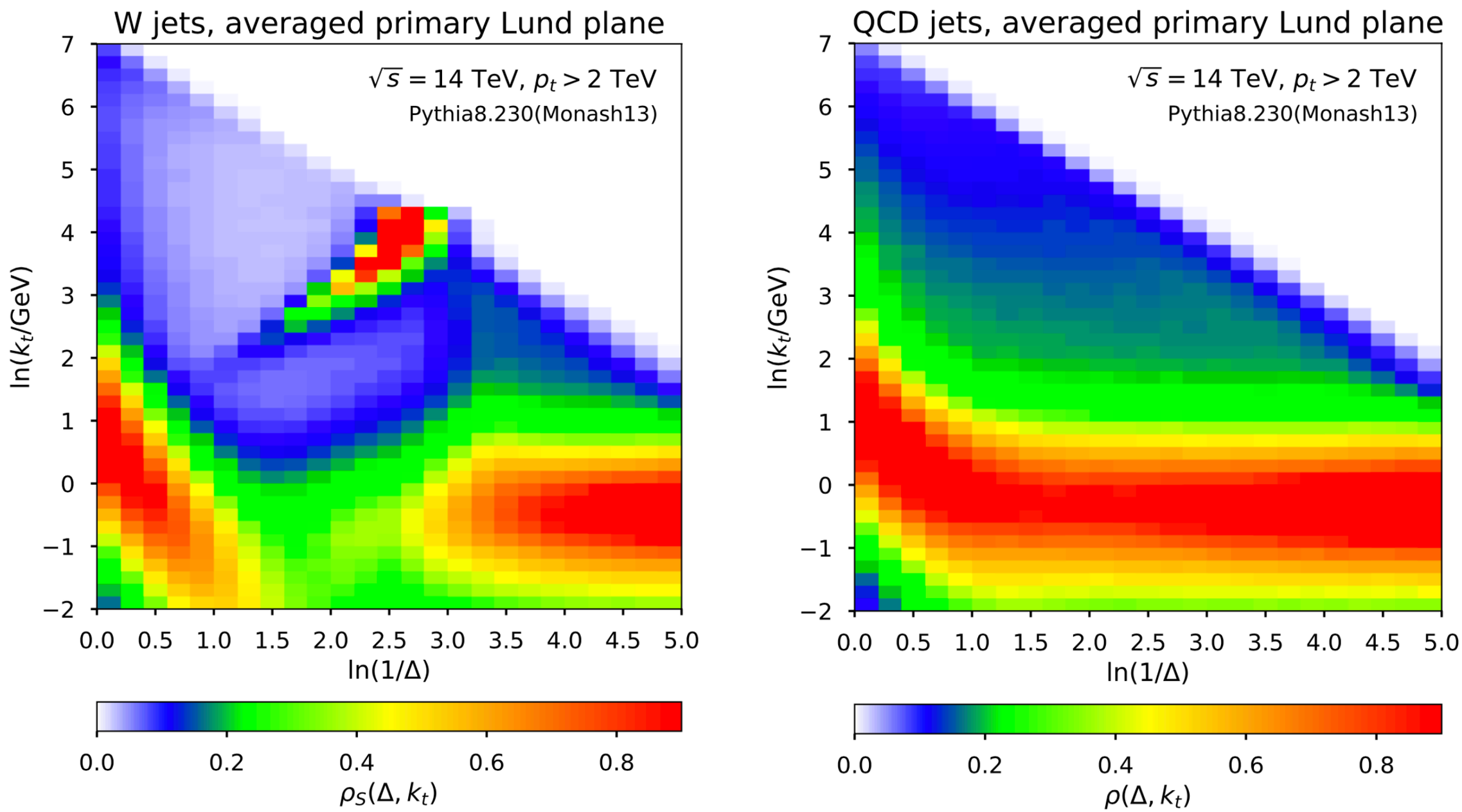
- Following a history of the hardest branch, we get double logarithmic plane $\left(\ln\frac{1}{\Delta}, \ln\frac{k_t}{\text{GeV}}\right)$.
- Pixels corresponding to the splitting are turned on.



LundPlane library is publicly available as a part of FastJet contrib package.

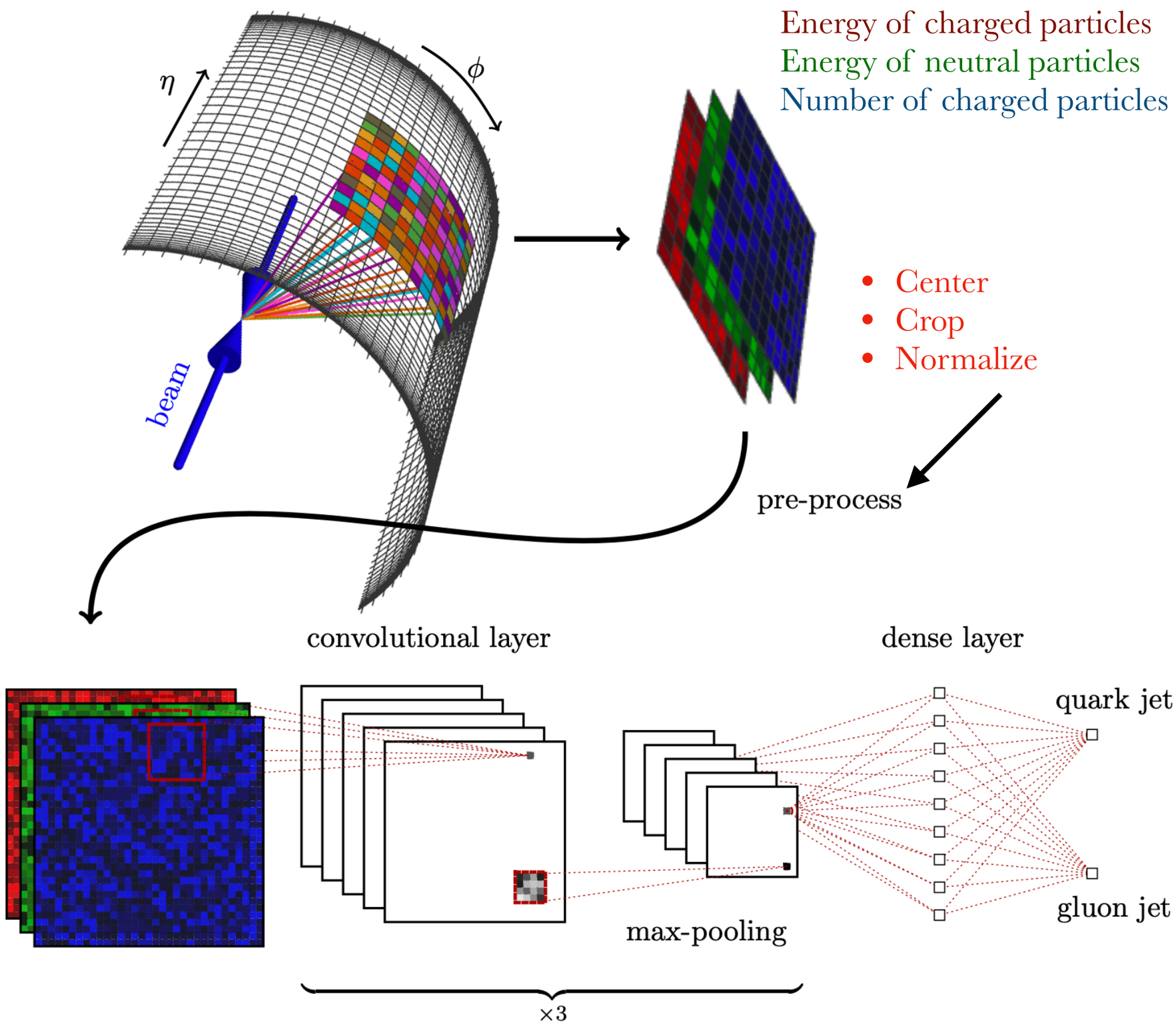
F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064

W and QCD Jets



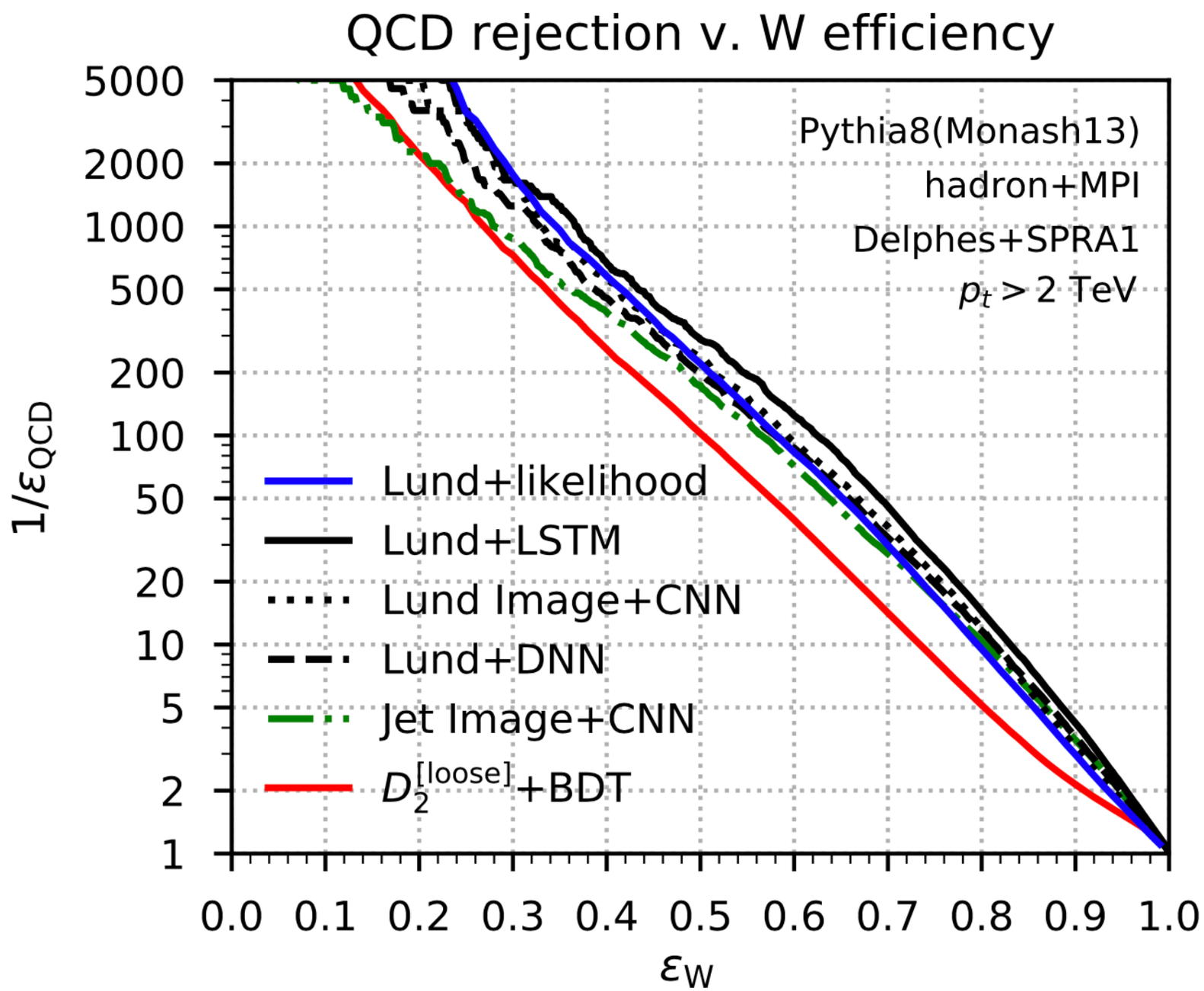
F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064

Example: Jet Images as an input to CNN



P. T. Komiske, E. M. Metodiev and M. D. Schwartz, JHEP 01(2017), 110

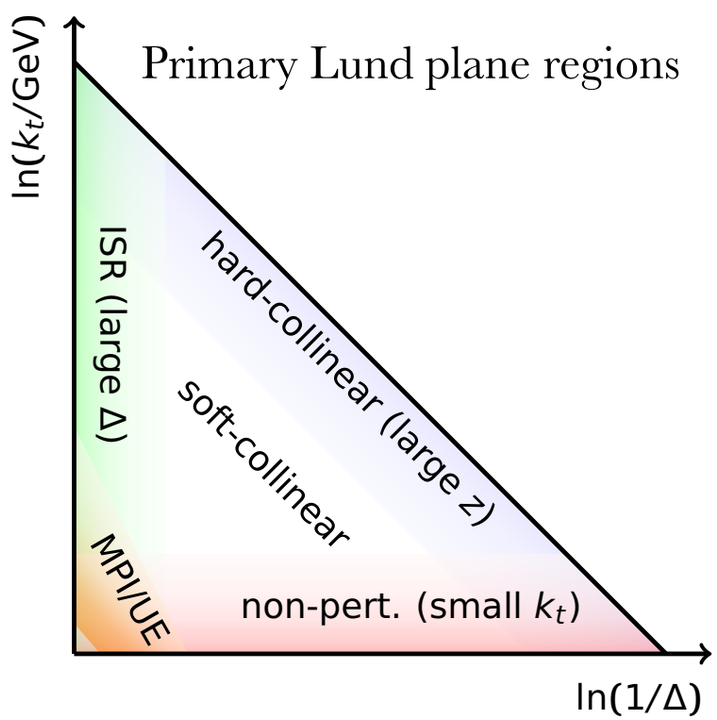
W Tagging using Lund Jet Plane



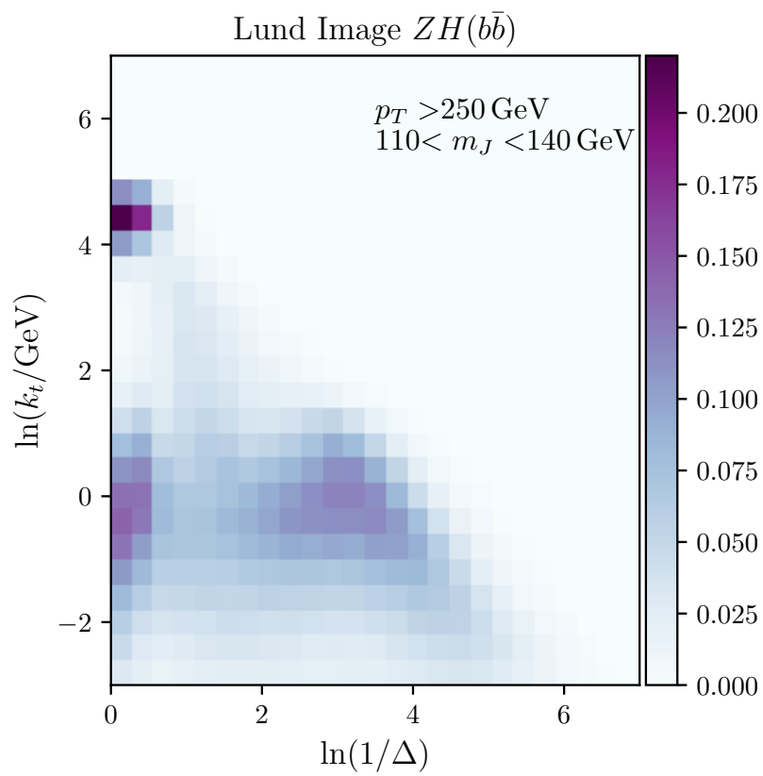
F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064

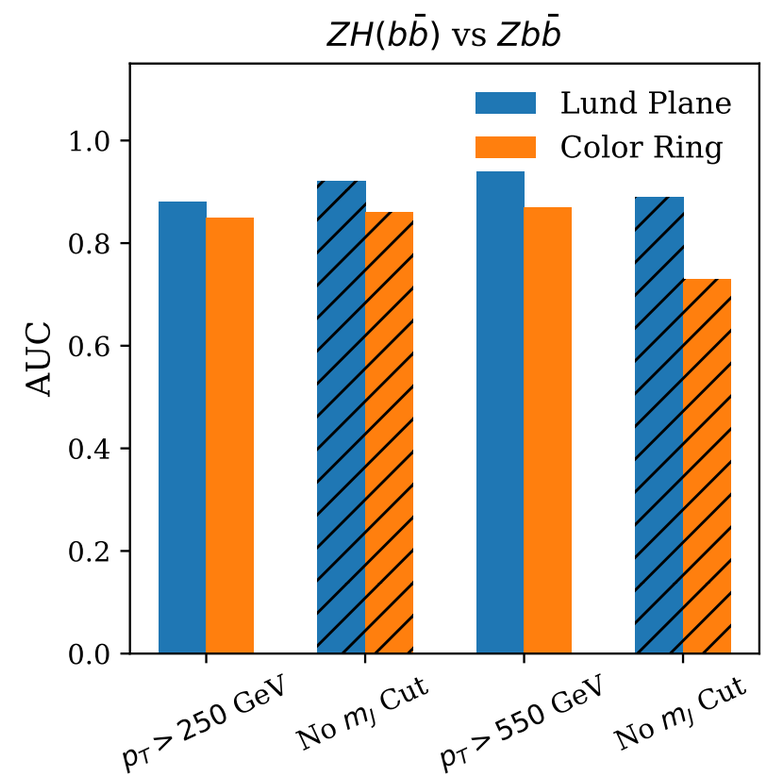
Higgs Tagging using the Lund Jet Plane

CKK and Simone Marzani, Phys. Rev. D 104 (2021) no.5, 055043, arXiv:2105.03989 [hep-ph]



F.A.Dreyer, G.P.Salam and G.Soyez, JHEP 12 (2018), 064





Jet color ring $\mathcal{O} = \dfrac{\theta_{ka}^2 + \theta_{kb}^2}{\theta_{ab}^2}$

Simulation Set-up

Tools: Madgraph 2.7.2, Pythia 8 and Fastjet 3.0.

High and moderate p_T benchmarks: $p_T > 250$ GeV and $p_T > 550$ GeV.

Generation level cuts: $p_T > 200$ GeV and $p_T > 500$ GeV, $|\eta_j| < 5.0, |\eta_l| < 2.5$.

Jet mass cut: $110 < m_J < 140\,\text{GeV}$.

$H \to b\bar{b}$

$pp \to ZH\ (Z \to \mu^+\mu^-, H \to b\bar{b})$

$pp \to Zb\bar{b}\ (Z \to \mu^+\mu^-)$

$H \to gg$

$pp \to ZH\ (Z \to \mu^+\mu^-, H \to gg)$

$pp \to Zjj\ (Z \to \mu^+\mu^-)$

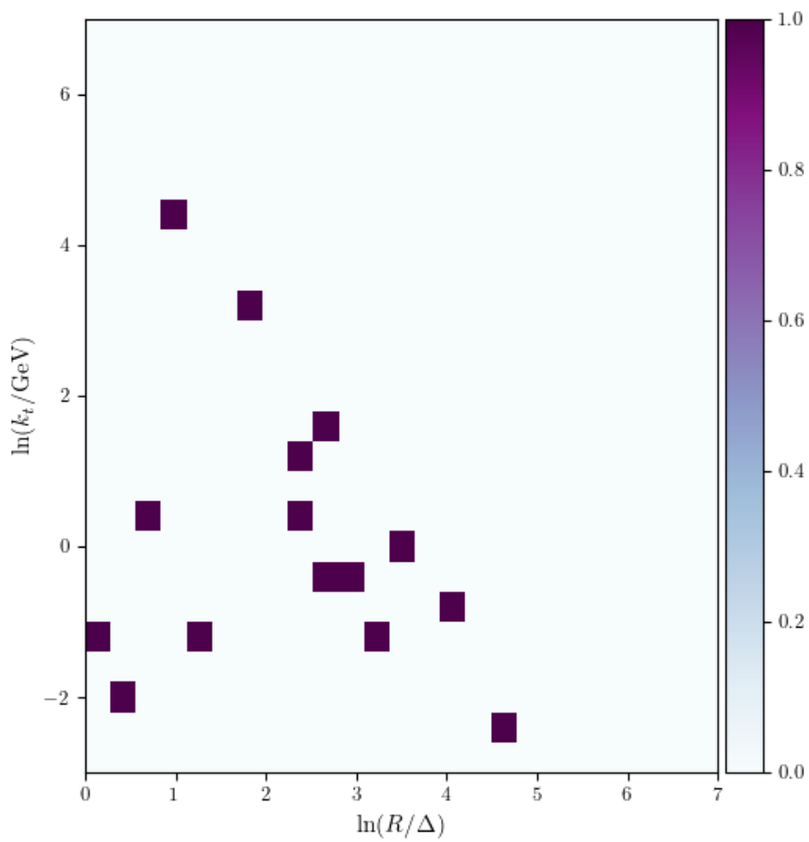
CKK and Simone Marzani, Phys. Rev. D 104 (2021) no.5, 055043, arXiv:2105.03989 [hep-ph]

Mapping Events to the (primary) Lund Jet Plane

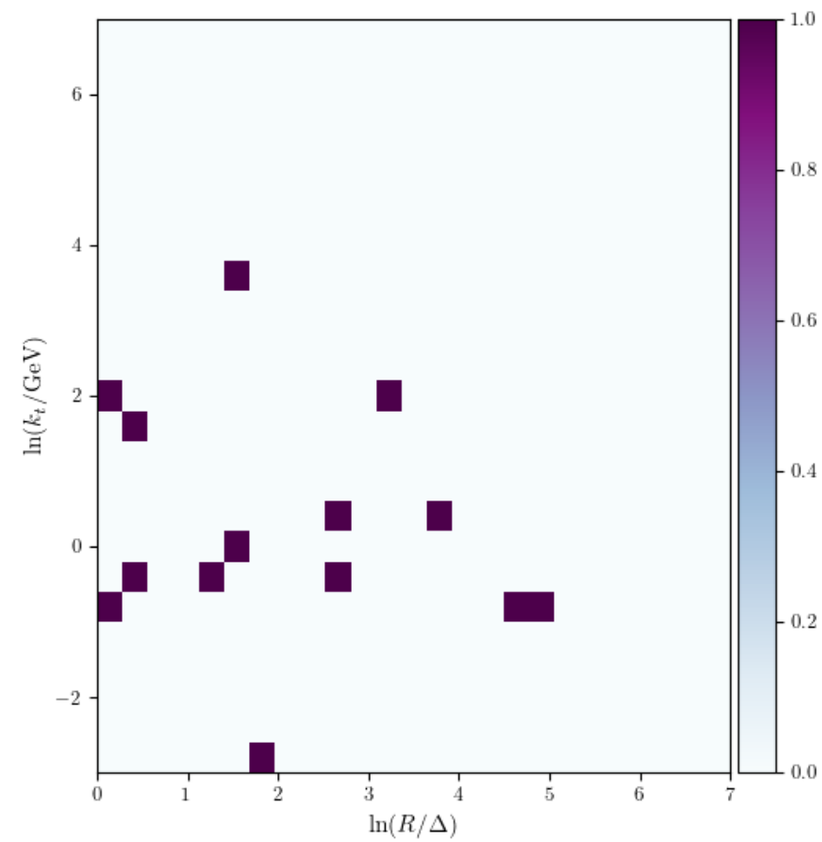
- Consider a list of particles (except the muons from hard scattering).
- Form anti-k_t $R = 1$ jets using Fastjet.
- Cluster the constituents of the leading jet using C/A algorithm with a maximum allowed jet radius.
- Save the Lund coordinates history of the hardest branch.
- Use a double logarithmic plane $\left(\ln \frac{1}{\Delta}, \ln \frac{k_t}{\mathrm{GeV}}\right)$.
- Pixels (25×25) corresponding to the splitting are turned on.

Lund Jet Plane

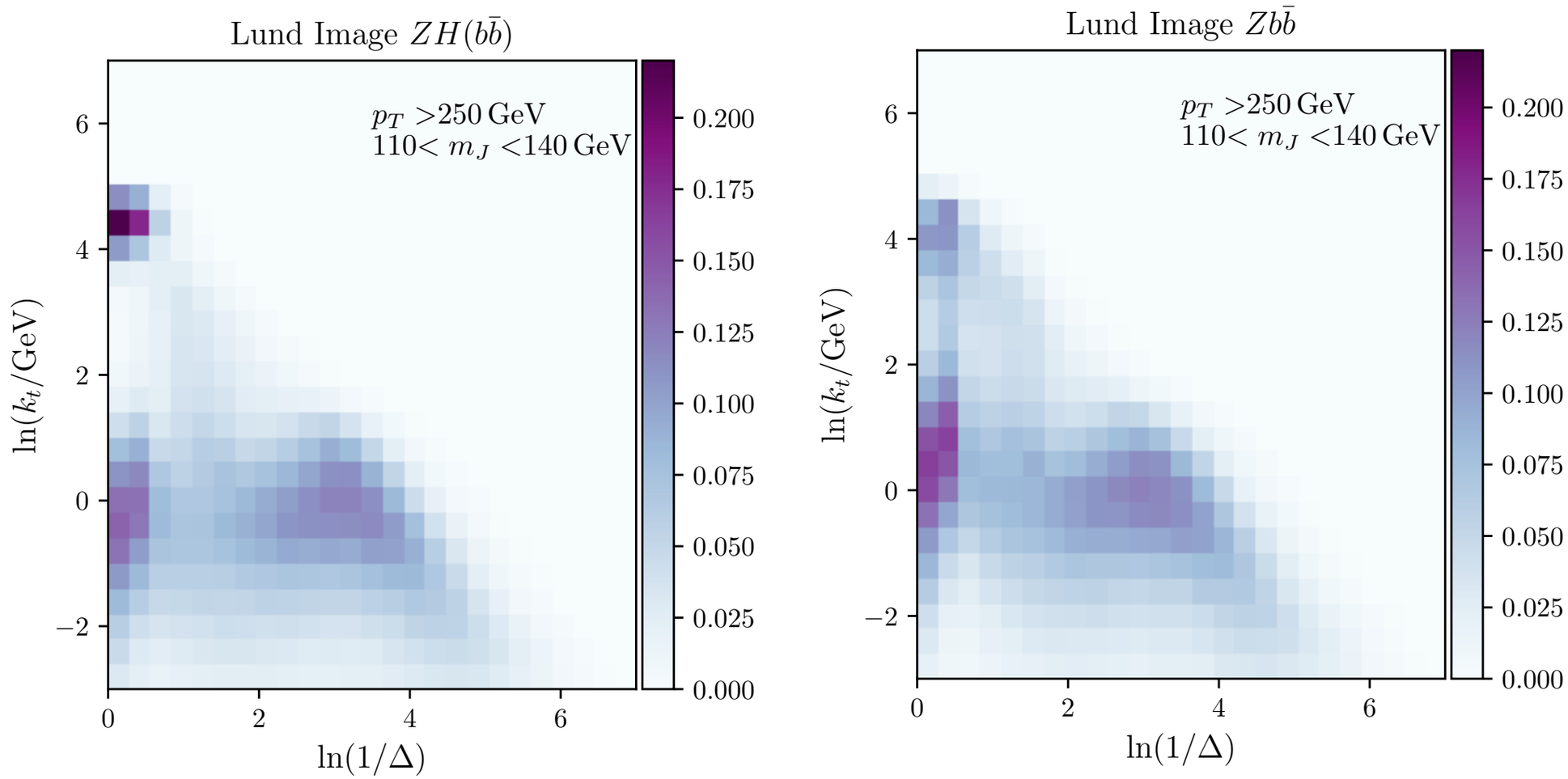
Signal



Background



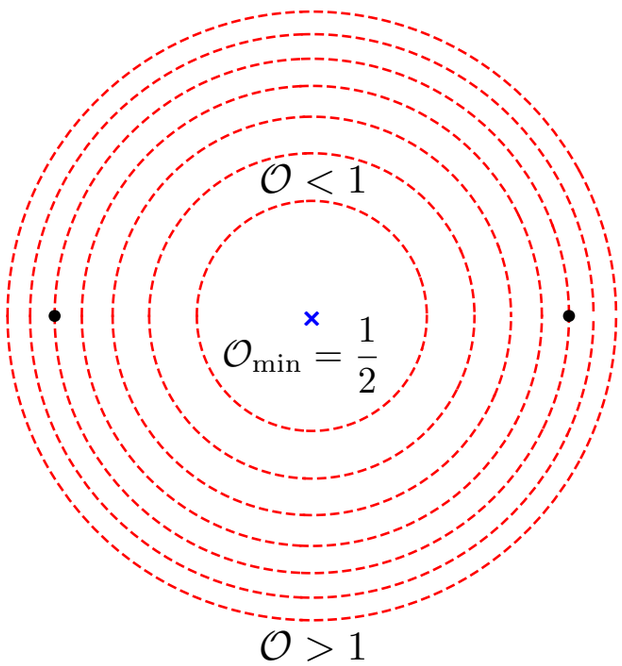
(Averaged) Lund Jet Images



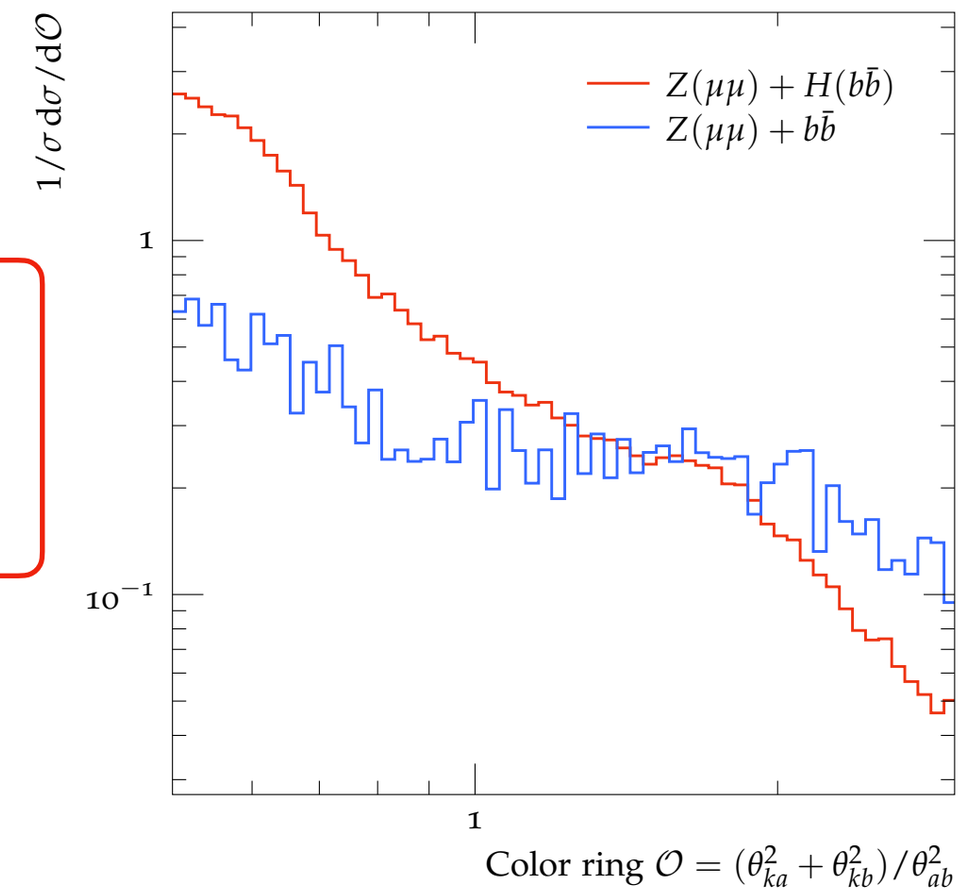
anti-k_T R=1 jets, reclustered the constituents using C/A algorithm

CKK and S.Marzani, Phys. Rev. D 104 (2021) no.5, 055043

Color Singlet Tagger: Jet Color Ring



$$\frac{|\mathcal{M}_S|^2}{|\mathcal{M}_B|^2} \simeq \frac{\theta_{ak}^2 + \theta_{bk}^2}{\theta_{ab}^2}$$



A.Buckley, G.Callea, A.J.Larkoski and S.Marzani, SciPost Phys. 9 (2020), 026

Constructing the Jet Color Ring

Identify sub-jets within the large R jet

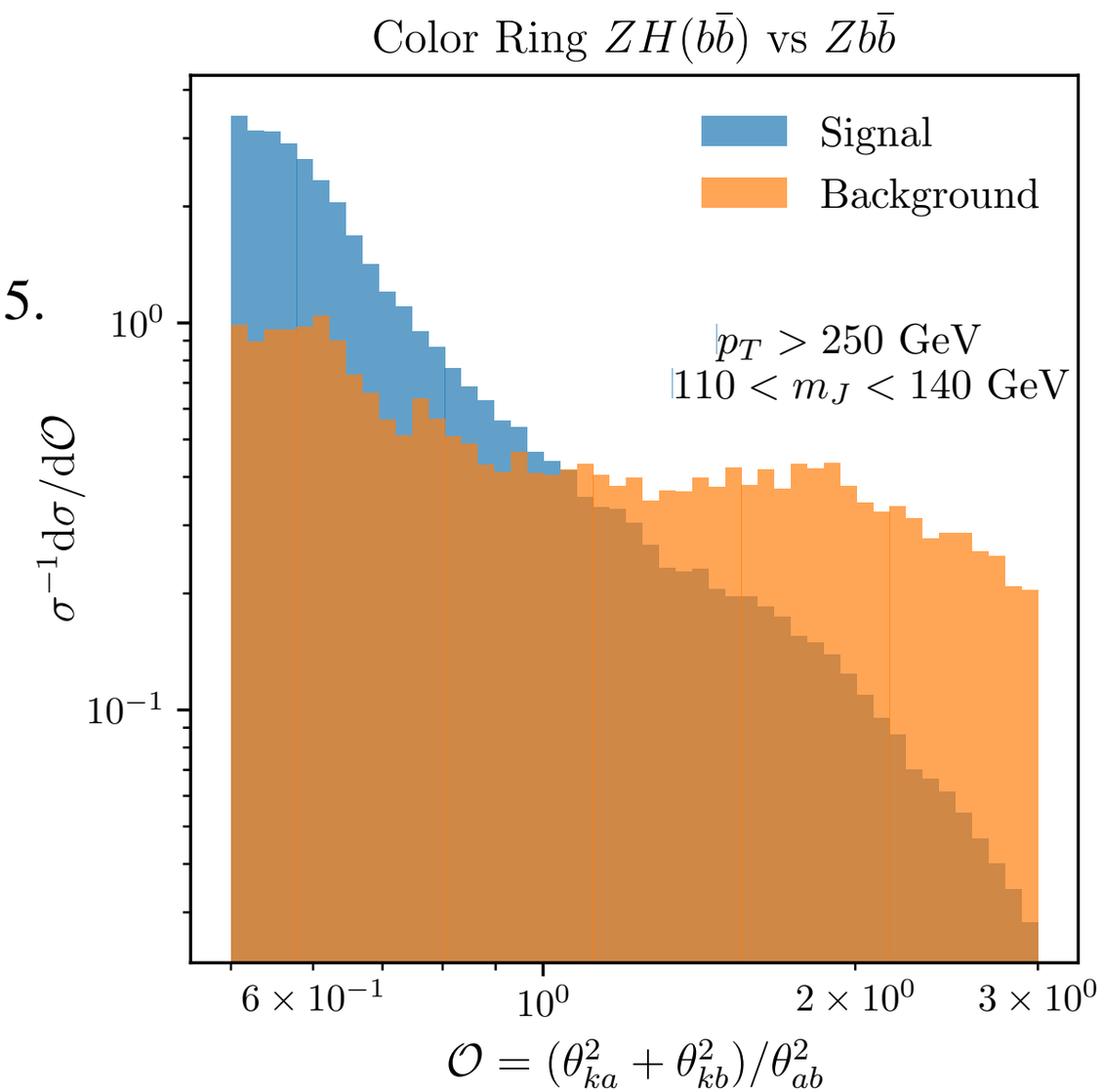
- Charged particles with $p_T > 500$ MeV, $|\eta| < 5$.
- Construct $R = 0.2$ track-jets using anti-k_t algorithm.
- Consider track jets with $p_T > 5$ GeV and $\Delta_{Jj} < 0.8$.

$\Delta_{jb/j\bar{b}} < 0.2$

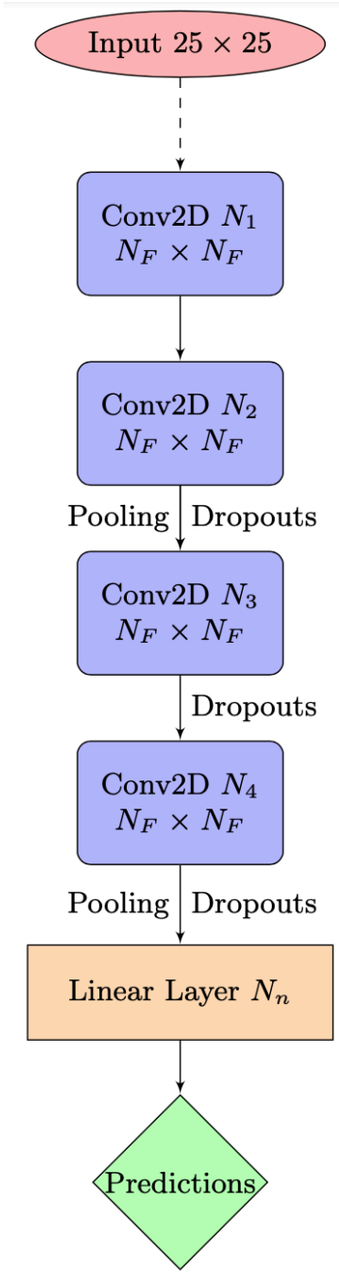
b-jets

light-jets

$$\mathcal{O} = \frac{\theta_{ka}^2 + \theta_{kb}^2}{\theta_{ab}^2}$$



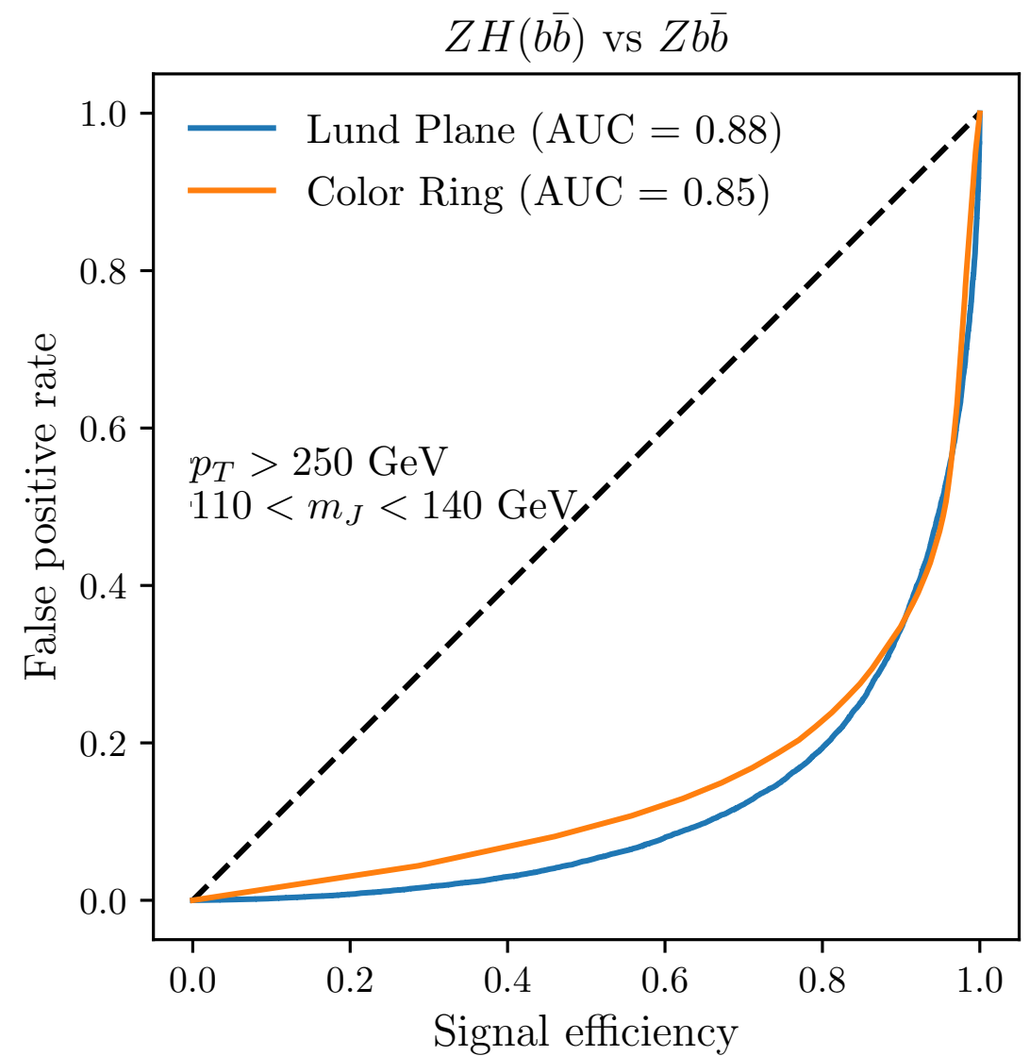
CNN for Binary Classification



100K Events (balanced data)

Training: Validation:Test data= 60:20:20%

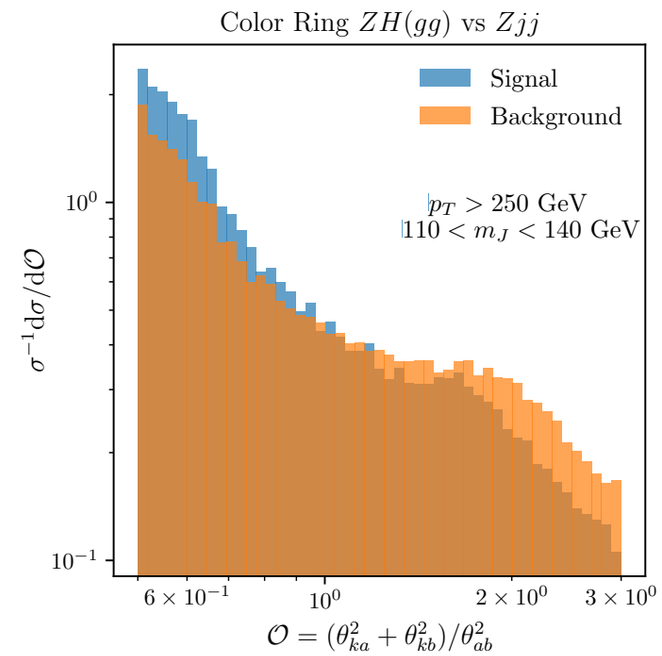
Cross-Entropy Loss function

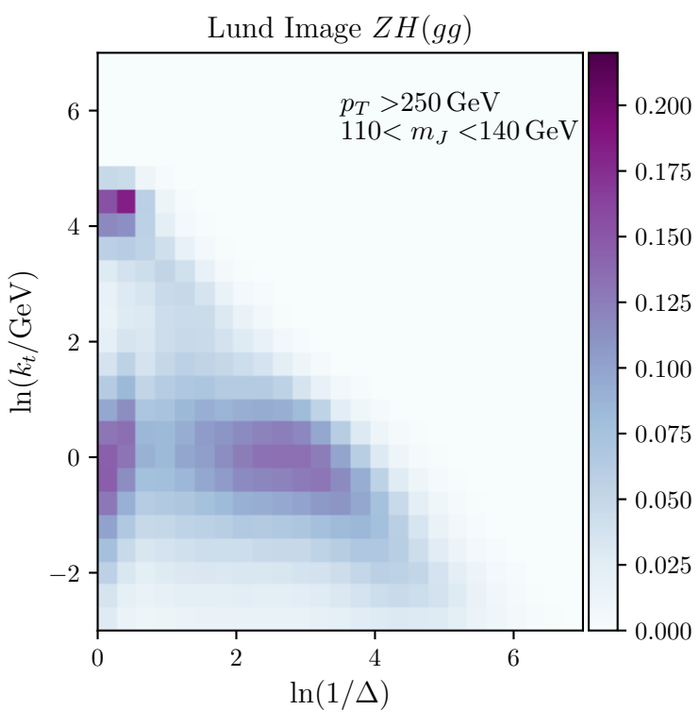


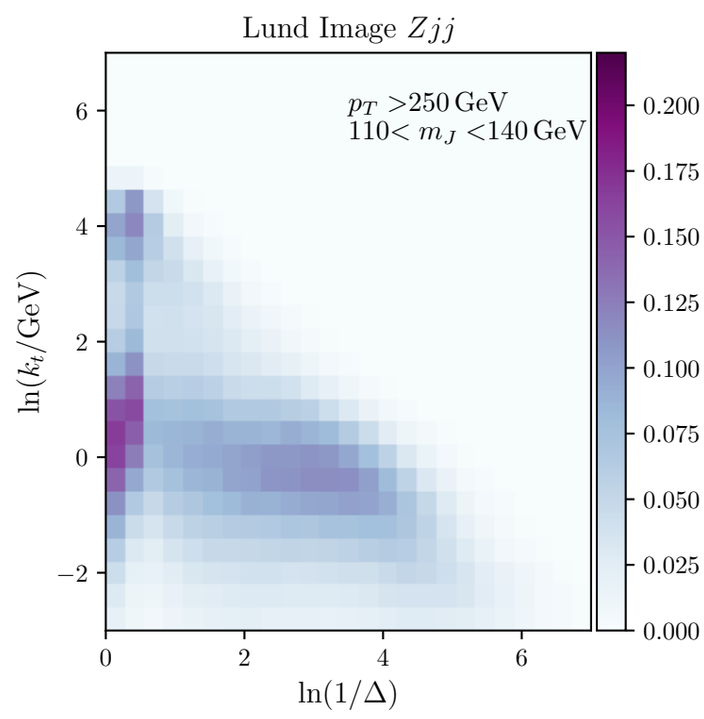
$H \longrightarrow gg$ Analysis

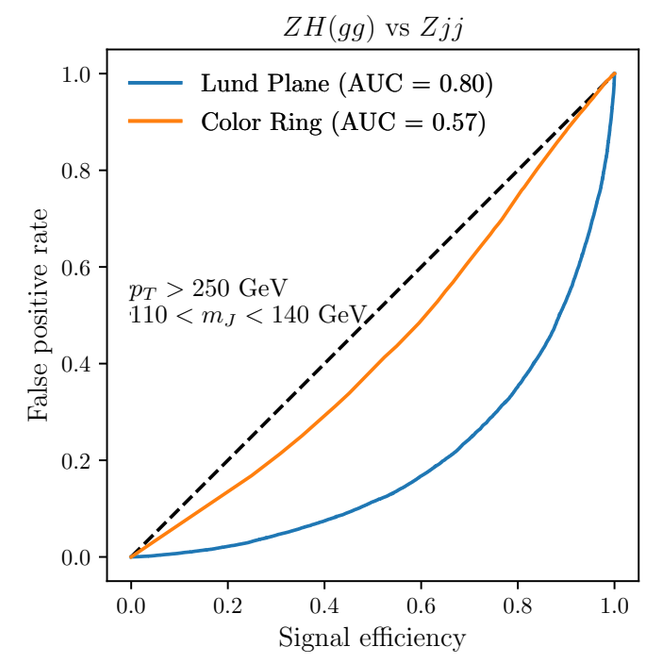
Jet Color Ring:

- Signal distribution is the same as $H \longrightarrow b\bar{b}$ case.
- Several color configurations are contributing to the background.

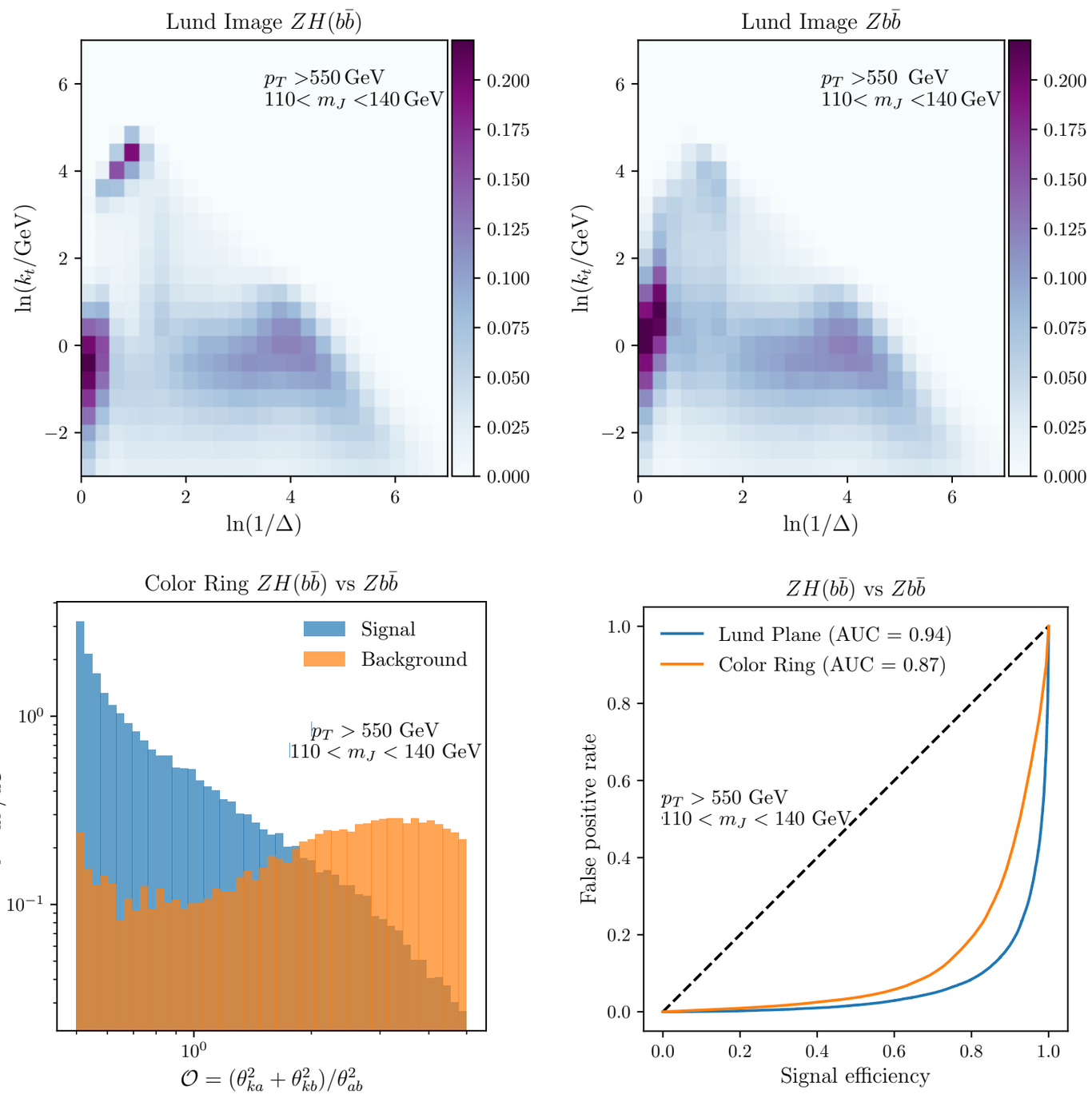








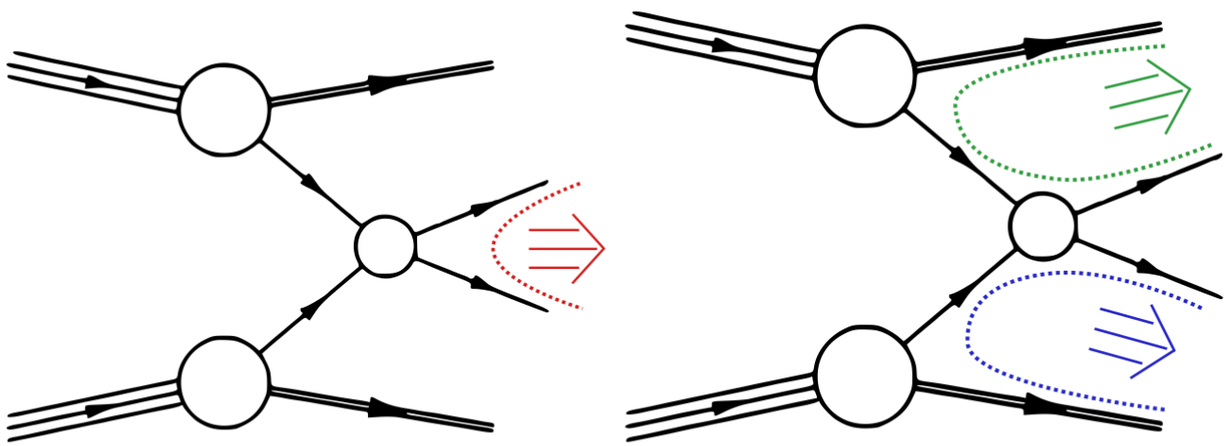
High p_T Scenario

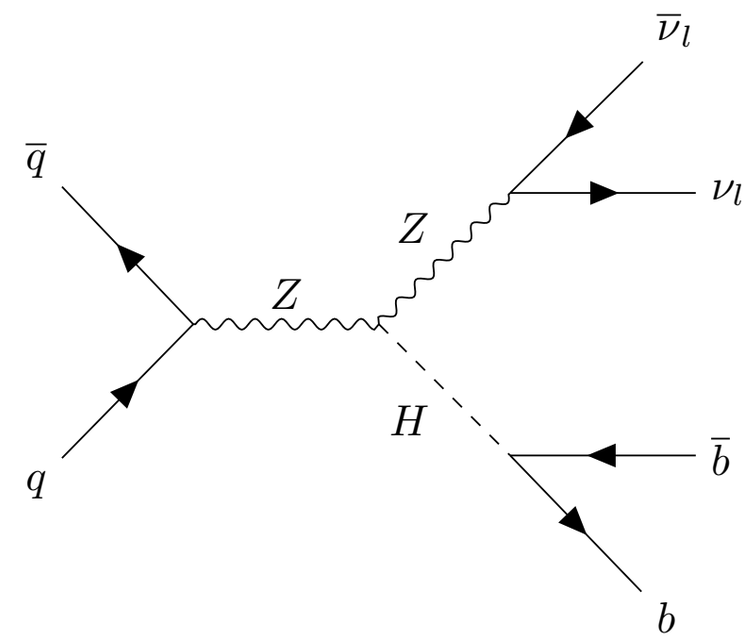


Color sensitive observable and the Lund Jet Plane

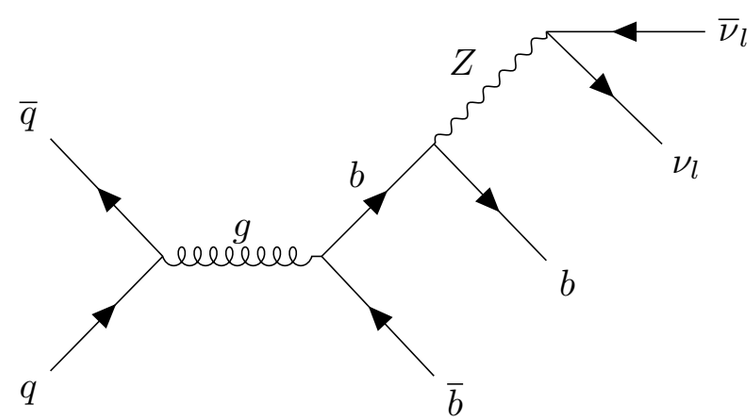
L.Cavallini, A.Coccaro, CKK, G.Manco, S.Marzani, F.Parodi, D.Rebuzzi, A.Rescia and G.Stagnitto,
Eur. Phys. J. C 82 (2022) no.5, 493, [arXiv:2112.09650 [hep-ph]]

Jet pull and its projections, D_2, and color ring, Lund jet plane
(Including detector effects)





Possible color connections for signal ($pp \to H \to b\bar{b}$) and for the background ($pp \to g \to b\bar{b}$)



J. Gallicchio and M. D. Schwartz, Phys. Rev. Lett. 105 (2010), 022001.

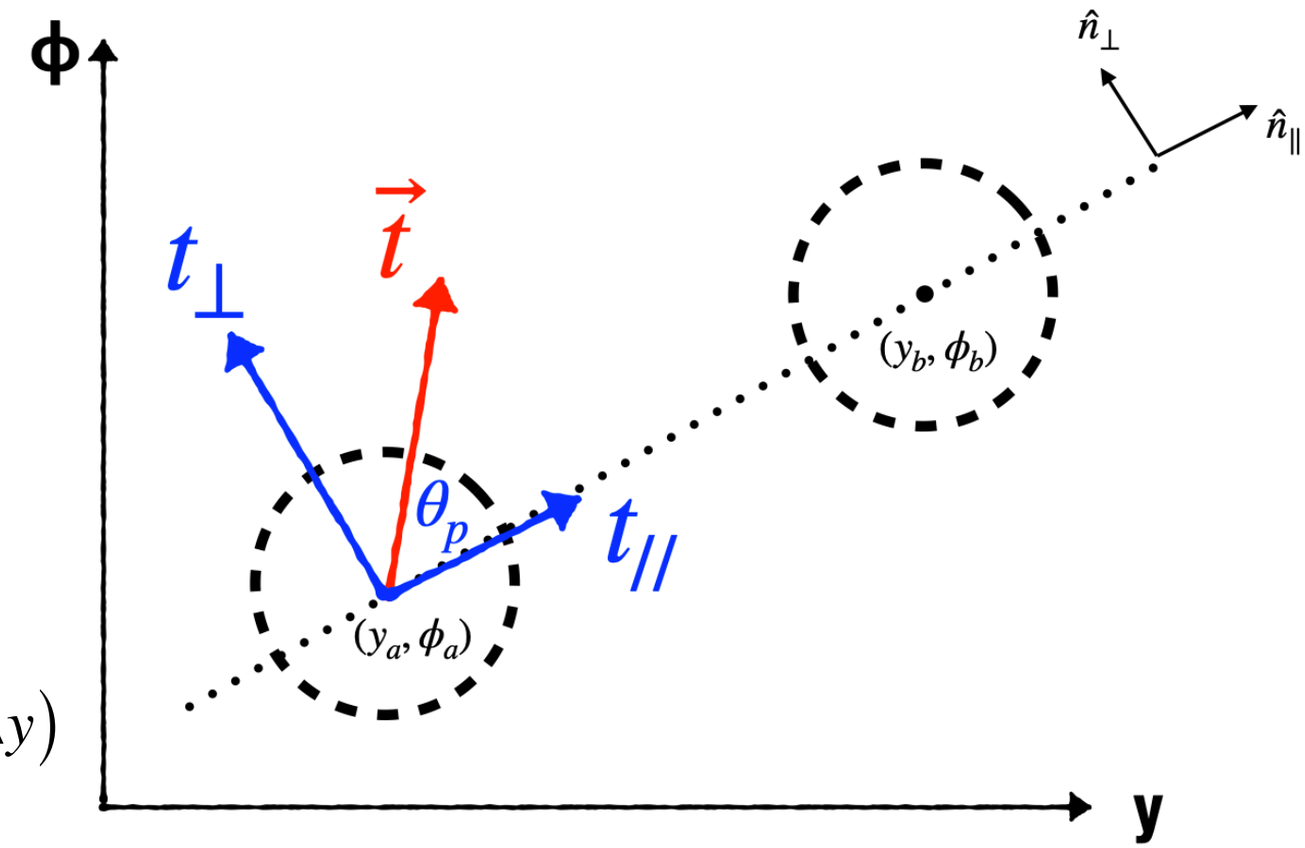
Jet Pull

$$\vec{t} = \frac{1}{p_{Ta}} \sum_{i \in J_a} p_{Ti} |\vec{r}_i|^2 \hat{r}_i$$

$$\vec{r}_i = (y_i - y_a, \phi_i - \phi_a)$$

$$t_{\parallel} = \vec{t} \cdot \hat{n}_{\parallel}, \qquad \hat{n}_{\parallel} = \frac{1}{\sqrt{\Delta y^2 + \Delta\phi^2}} (\Delta y, \Delta\phi)$$

$$t_{\perp} = \vec{t} \cdot \hat{n}_{\perp}, \qquad \hat{n}_{\perp} = \frac{1}{\sqrt{\Delta y^2 + \Delta\phi^2}} (-\Delta\phi, \Delta y)$$

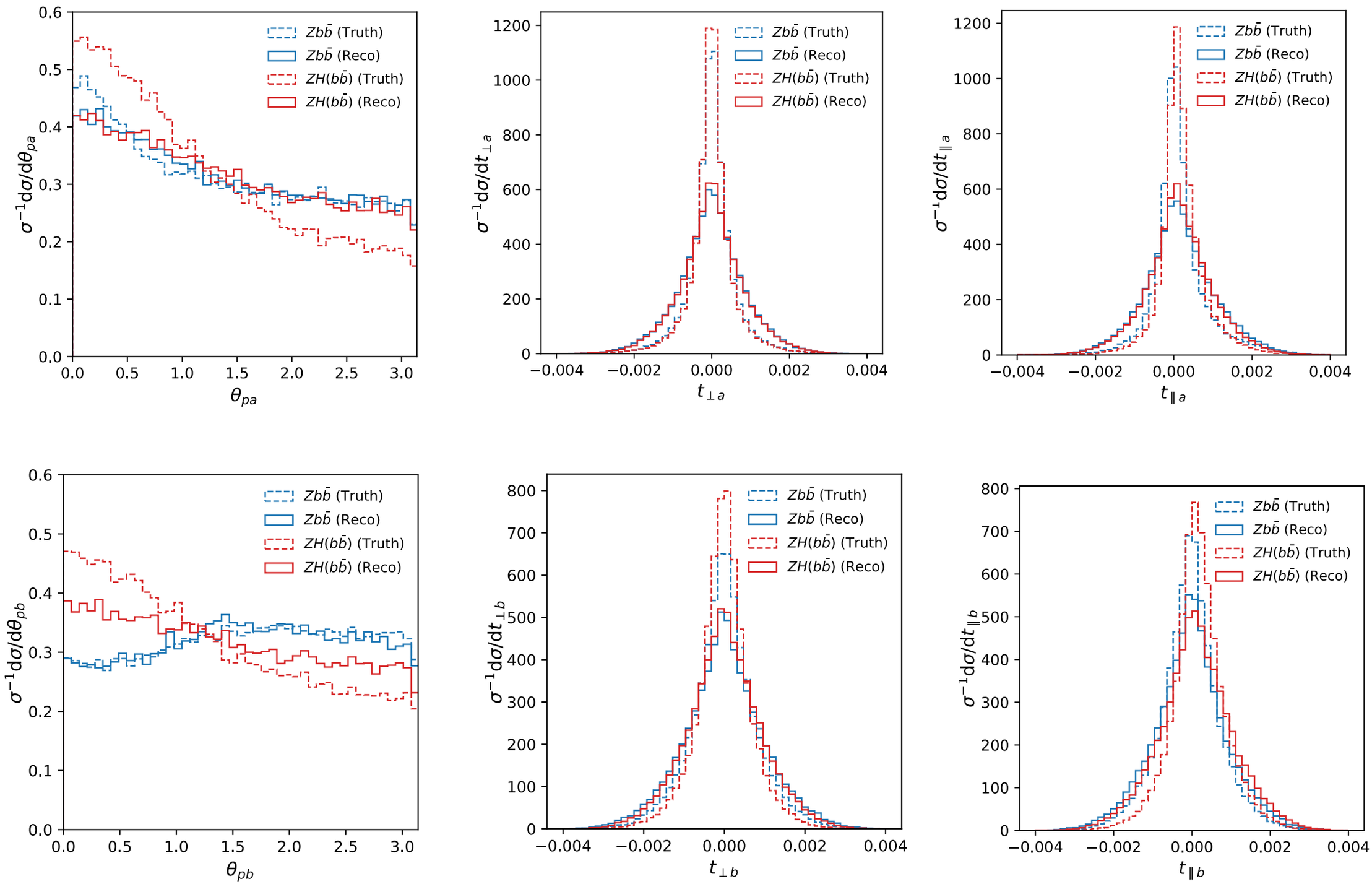


Pull vector components of jet J_a

$$\theta_p = \arccos \frac{t_{\parallel}}{|\vec{t}|}$$

A. Larkoski, S.Marzani and C.Wu, JHEP 01 (2020), 104)

Jet Pull and its projections

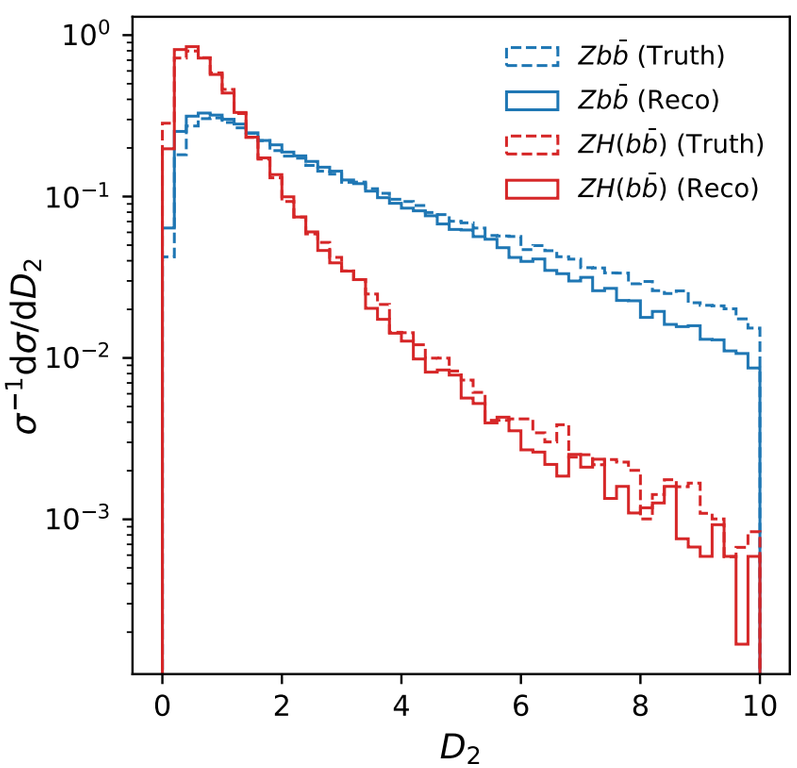


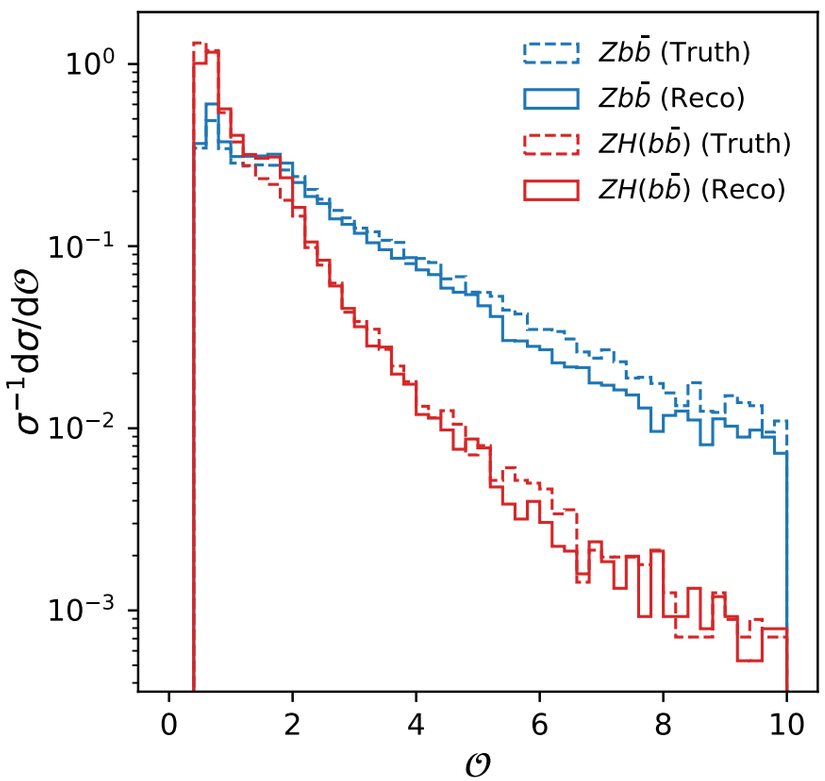
L.Cavallini, A.Coccaro, CKK, G.Manco, S.Marzani, F.Parodi, D.Rebuzzi, A.Rescia and G.Stagnitto, Eur. Phys. J. C 82 (2022) no.5, 493, [arXiv:2112.09650 [hep-ph]]

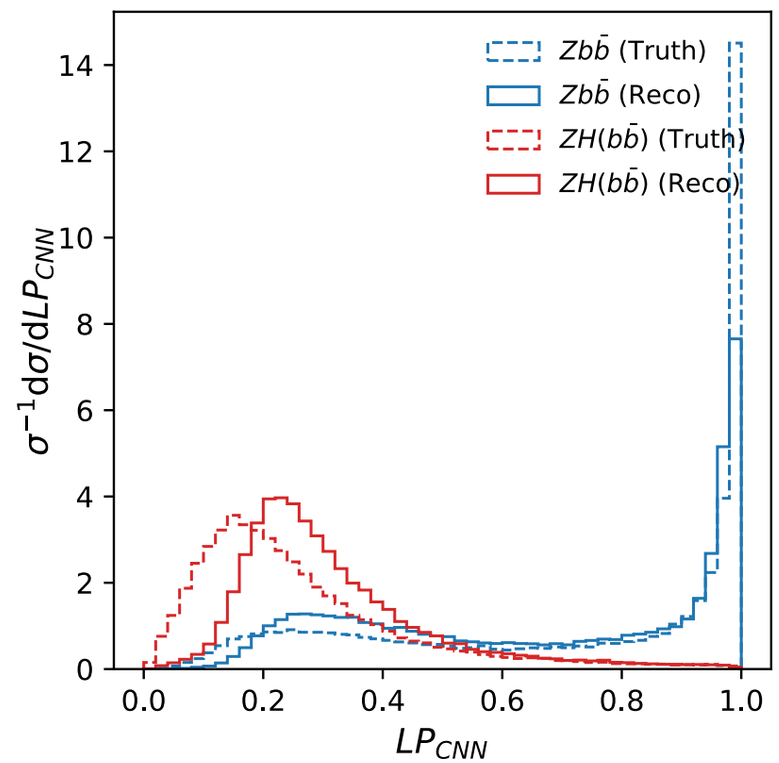
D_2 and Jet Color Ring

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$$

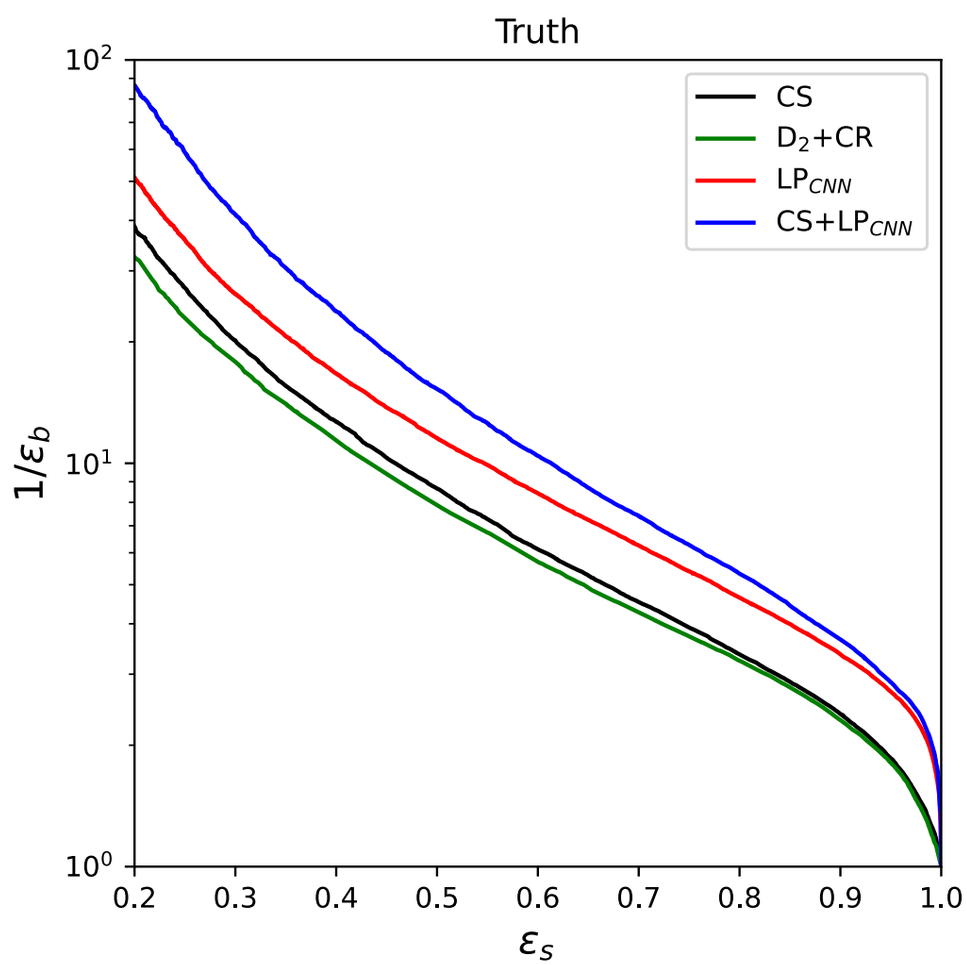
$e_n^{(\beta)}$ is the normalized n-point energy correlation functions

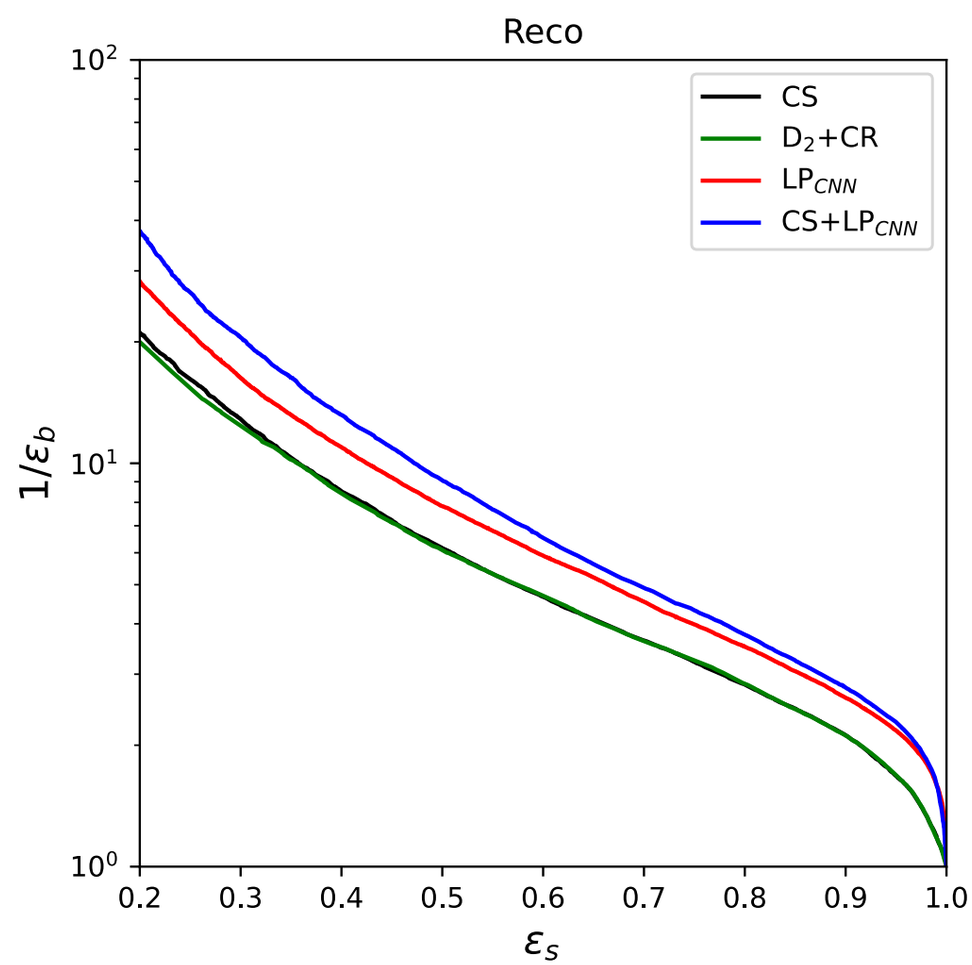






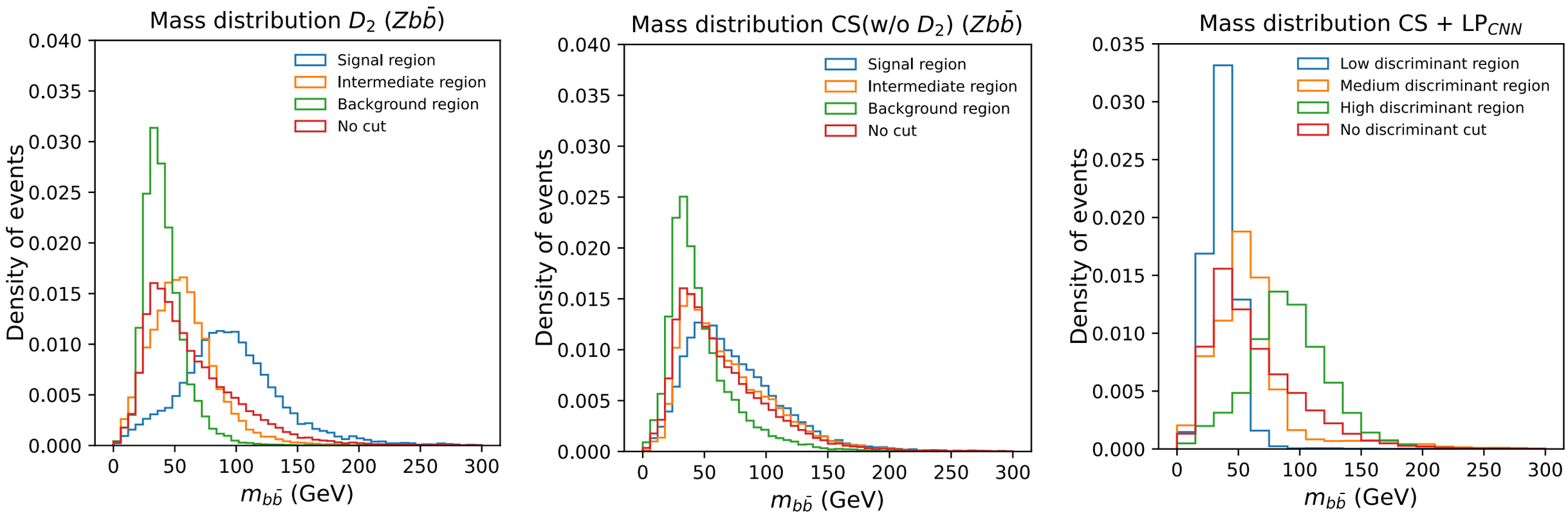
Discrimination Performance





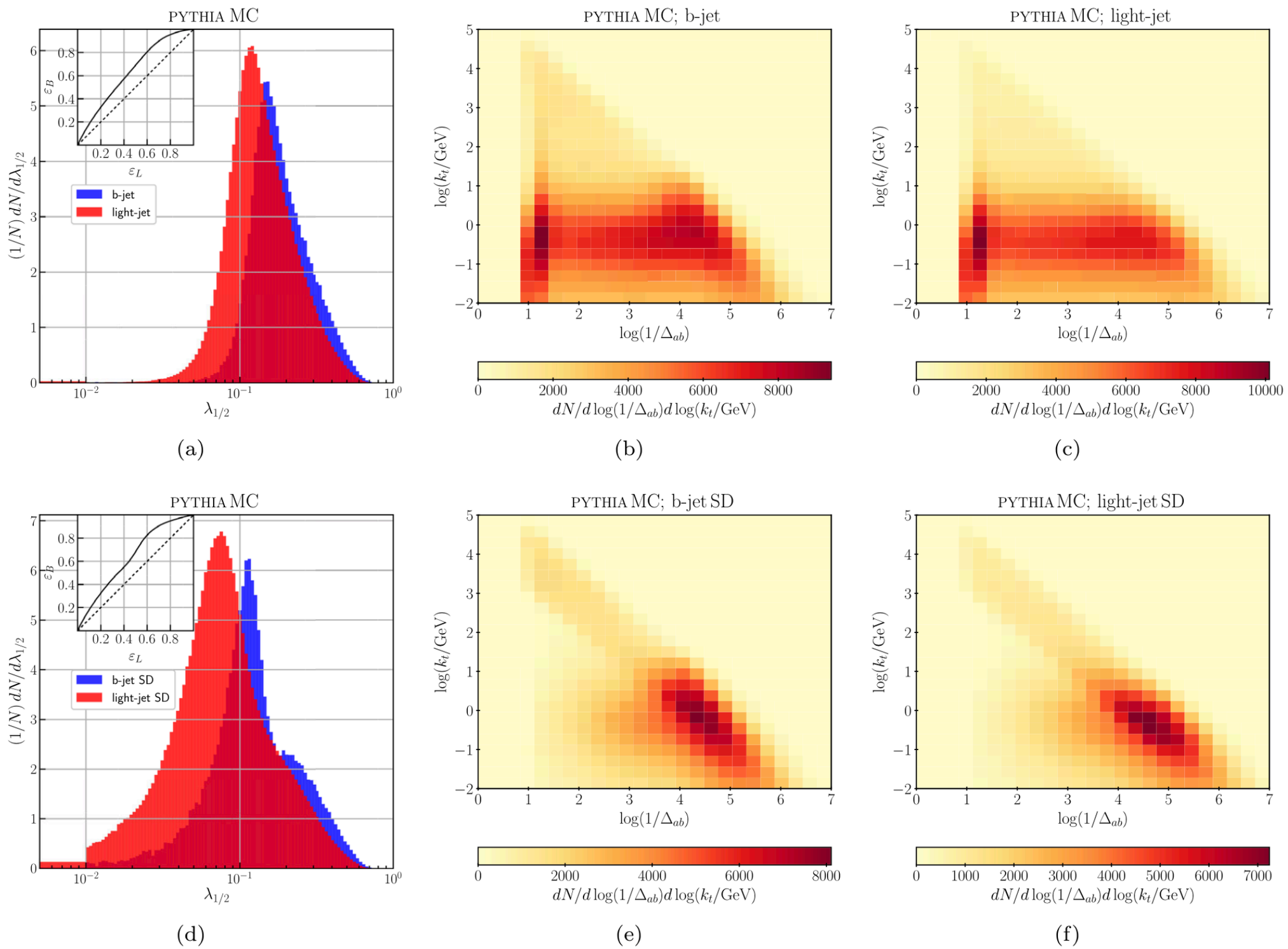
L.Cavallini, A.Coccaro, CKK, G.Manco, S.Marzani, F.Parodi, D.Rebuzzi, A.Rescia and G.Stagnitto, Eur. Phys. J. C 82 (2022) no.5, 493, [arXiv:2112.09650 [hep-ph]]

Invariant Mass Dependence



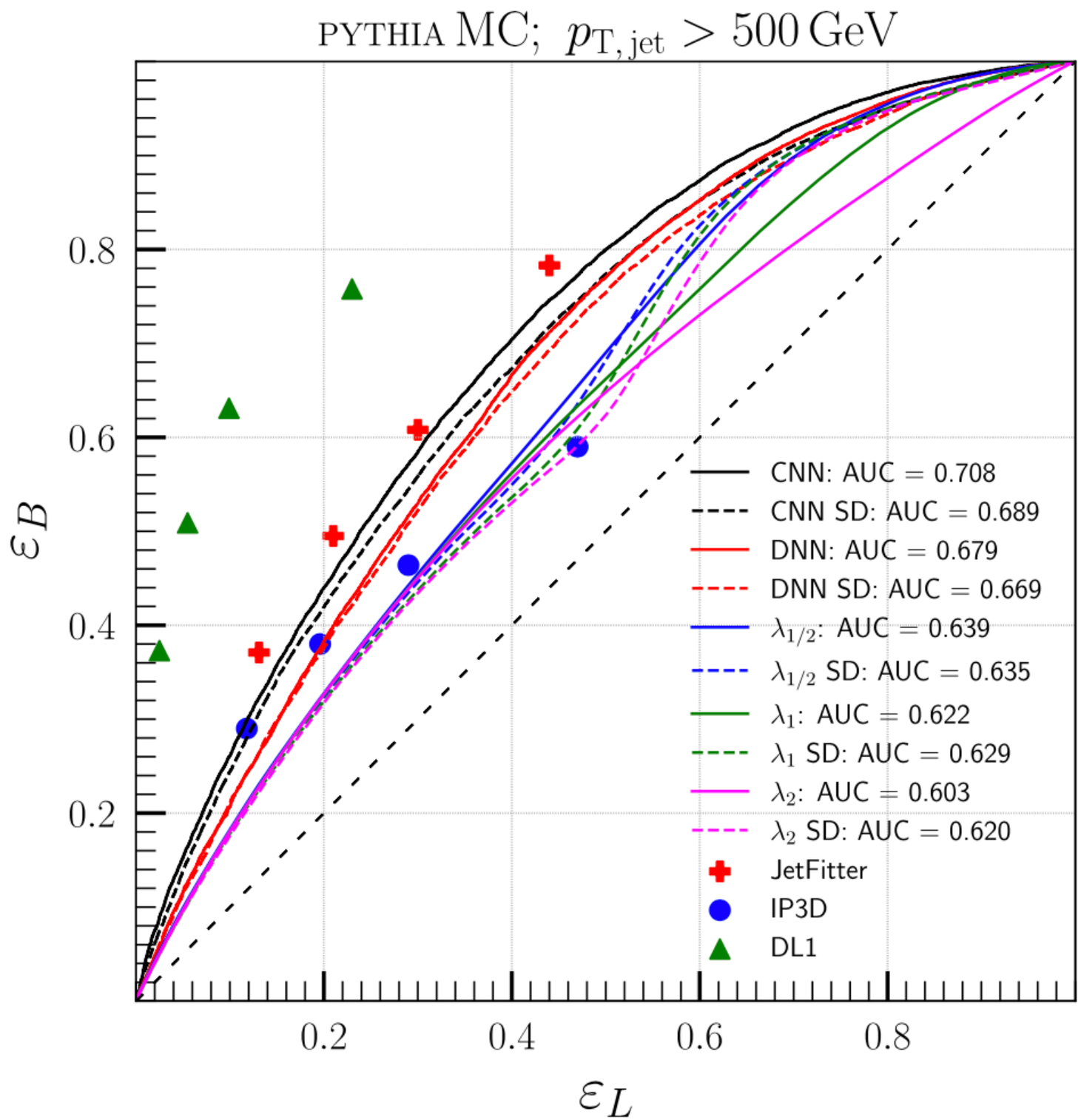
Background distribution of different ranges of classifier output.

Identification of b-jets



O.Fedkevych, CKK, S. Marzani and F. Sforza, Phys. Rev. D 107(2023) no.3, 034032

Identification of b-jets



O.Fedkevych, CKK, S. Marzani and F. Sforza, Phys. Rev. D 107(2023) no.3, 034032

Effect of b-mass in the Lund Jet Plane

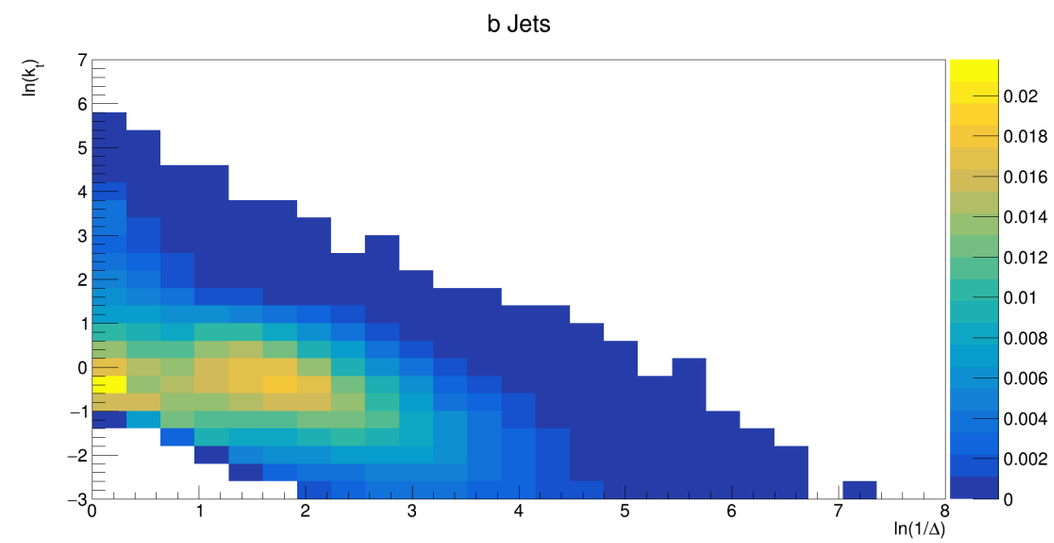
Ongoing work by

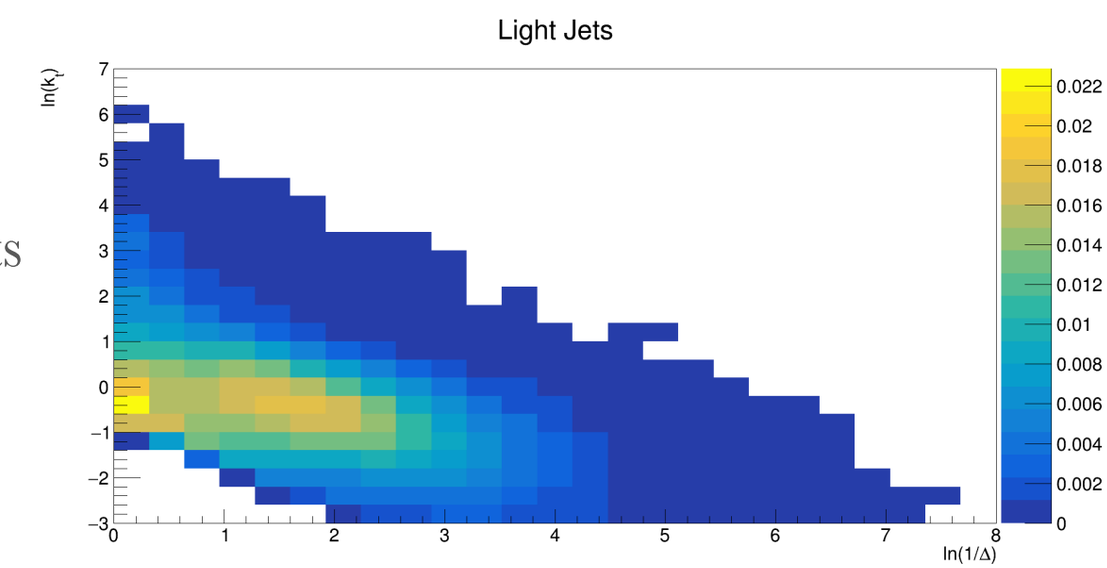
Francesco Giuli, Alberto Rescia, Federico Sforza

- Generate b & light jet events w/ MG5@NLO-Delta (v3.5.0)
  - $Z(l^+l^-) + j$ final state at NLO in 5FS for light events
  - $Z(l^+l^-) + bb$ with NLO in 4FS for b events
  - No kinematic cuts at generator level
- Shower with Pythia8 v8.309

Analyse at truth-level with a dedicated Rivet routine

- **Lepton cuts:**
  - $p_T > 27$ GeV, $|\eta| < 2.5$ and dilepton mass between 76-106 GeV
  - Isolation criterion: discard if $\Delta R < 0.4$ from jets
- **Jets:**
  Cluster final state hadrons into R = 0.4 jets with $p_T > 20$ GeV and $|\eta| < 2.5$
  - Tag jets as b or light via ghost-matching
  - Require 2 tagged jets in final state



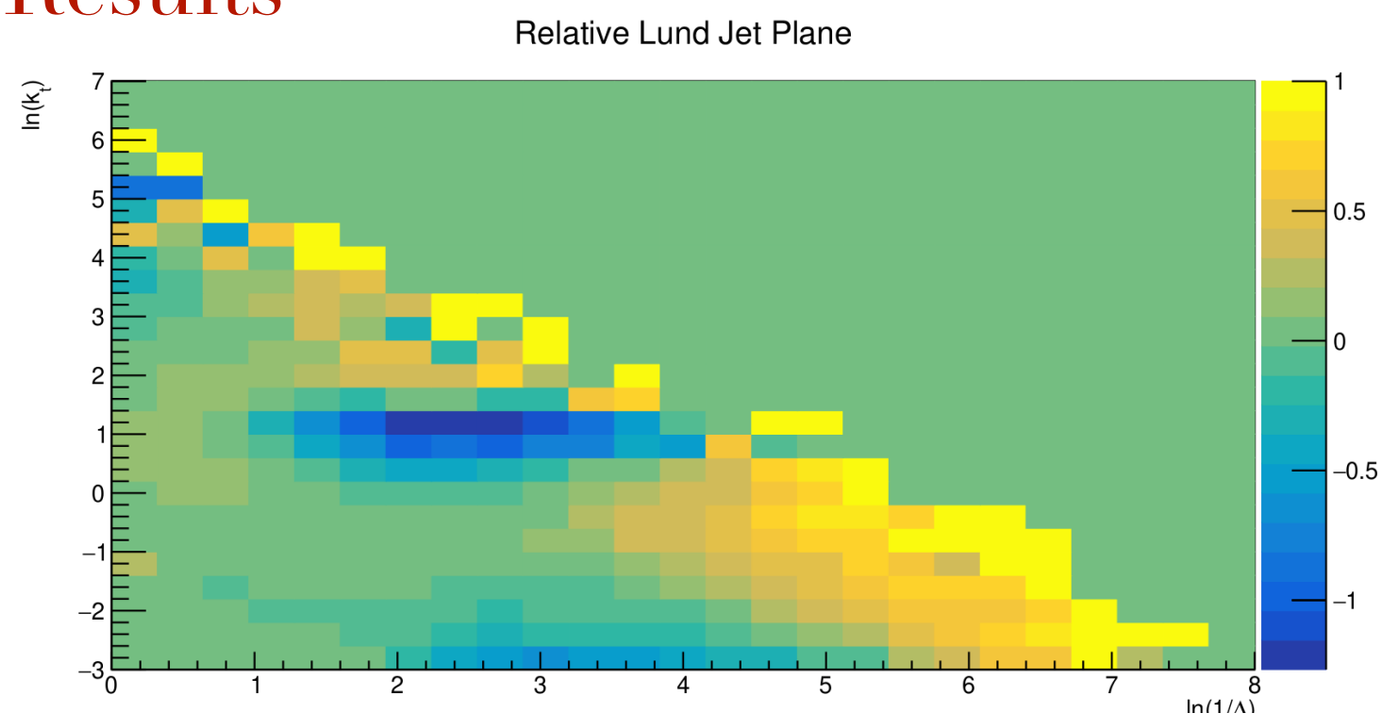


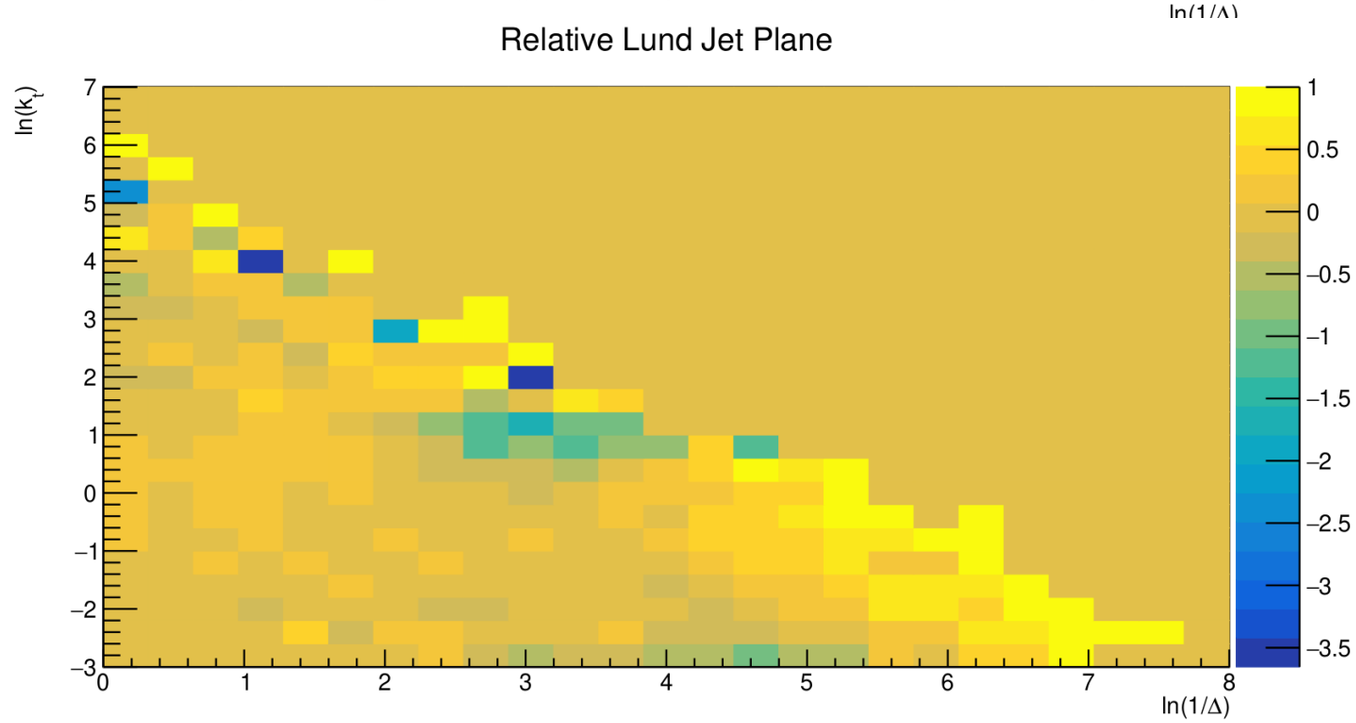
Effect of b-mass in the Lund Jet Plane

Ongoing work by

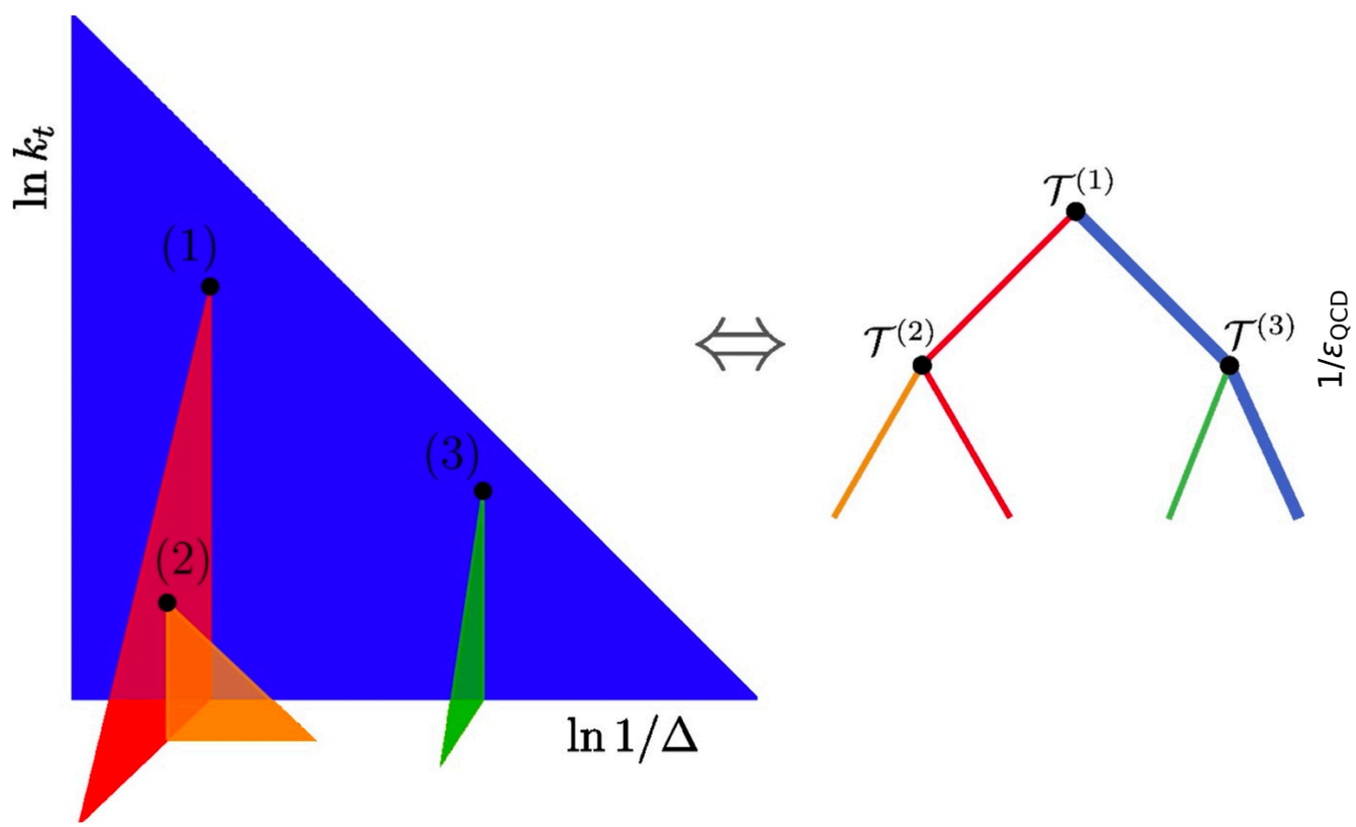
Francesco Giuli, Alberto Rescia, Federico Sforza

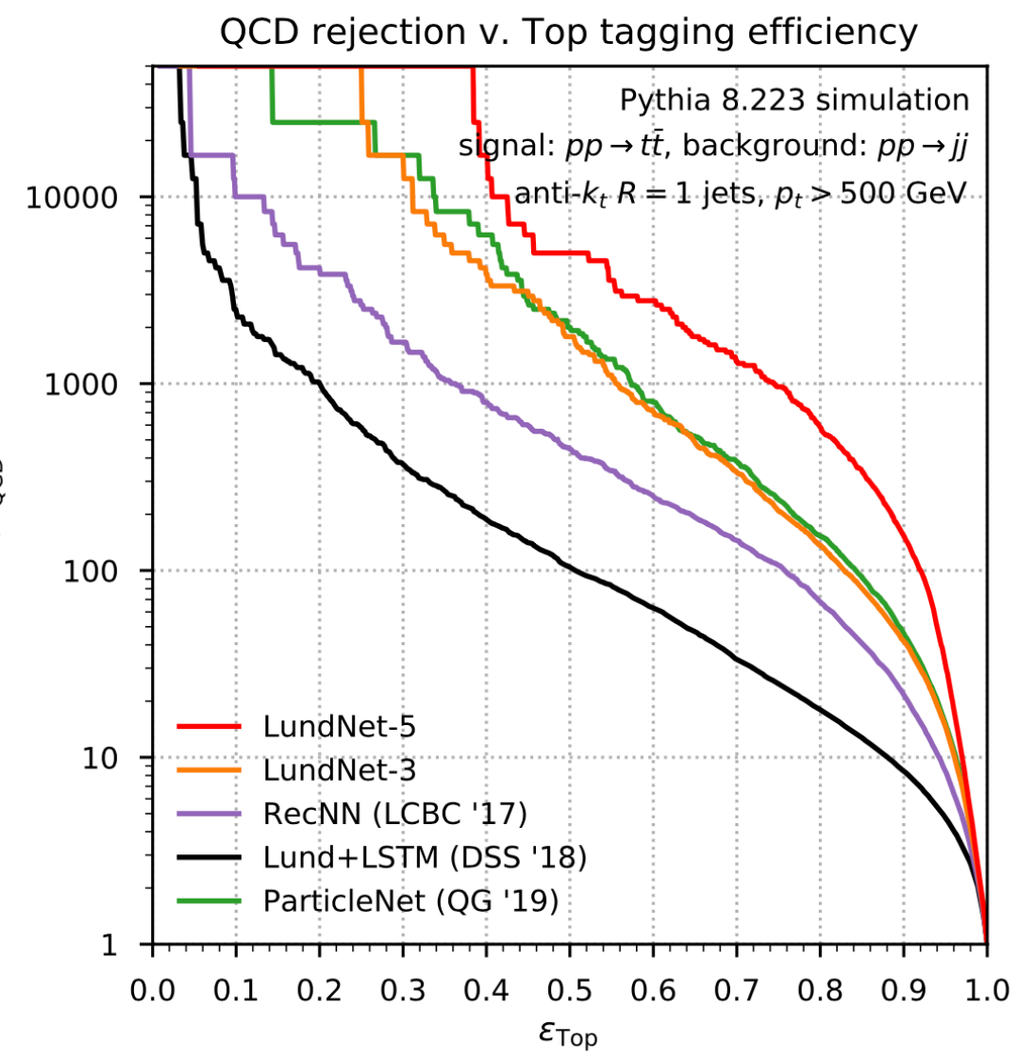
Preliminary Results





LundNet for Top Tagging



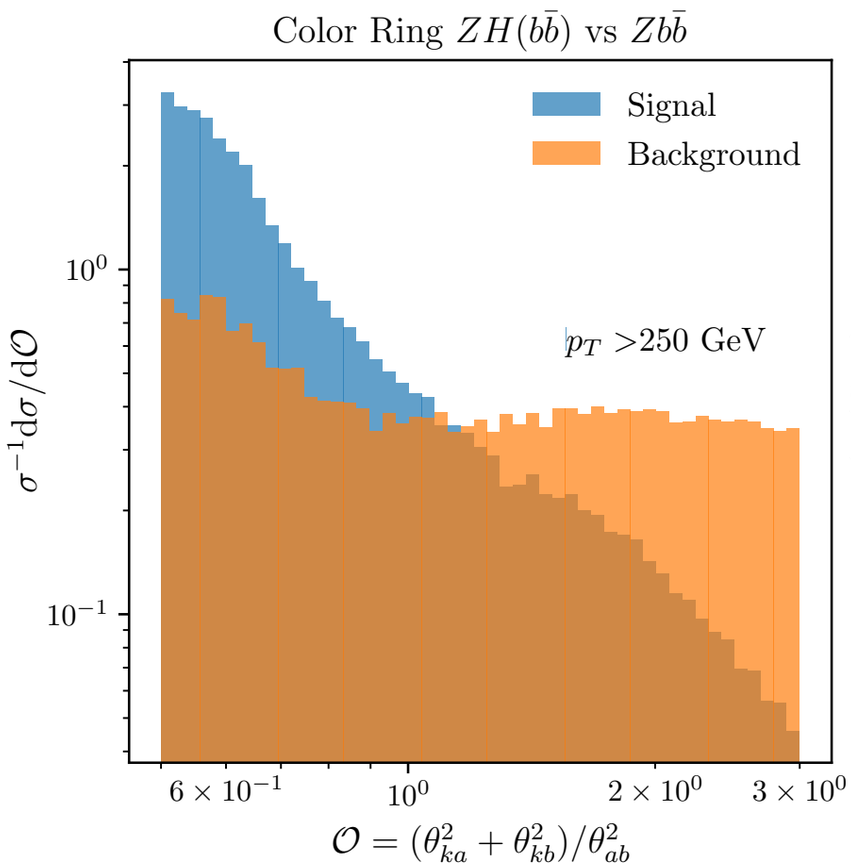


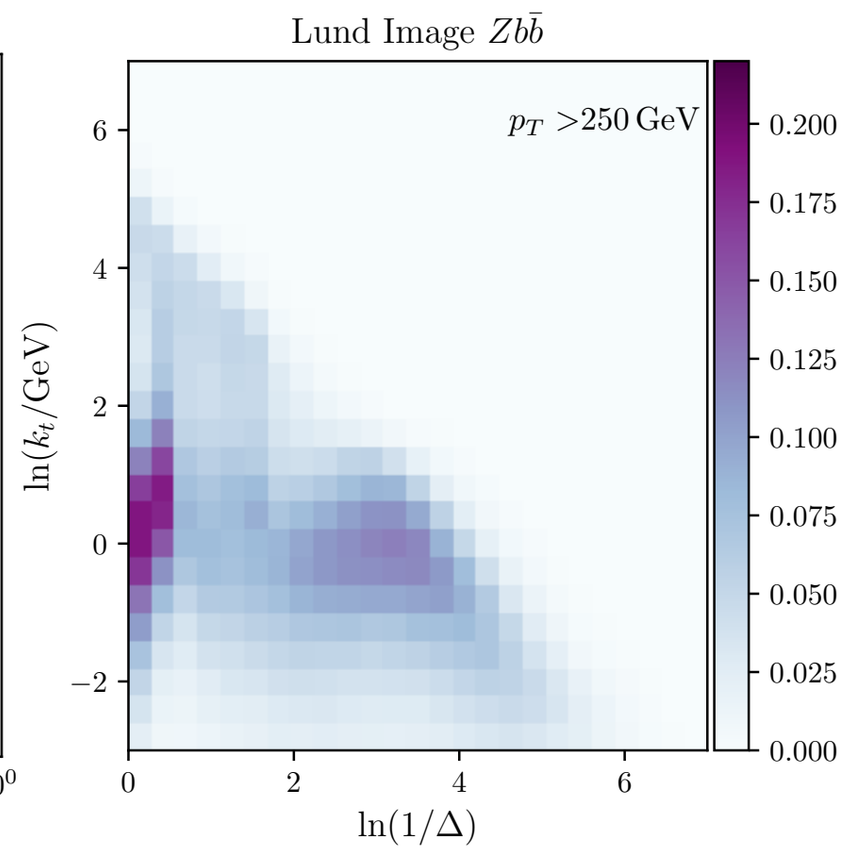
F.A.Dreyer, and H.Qu, JHEP 03 (2021), 052

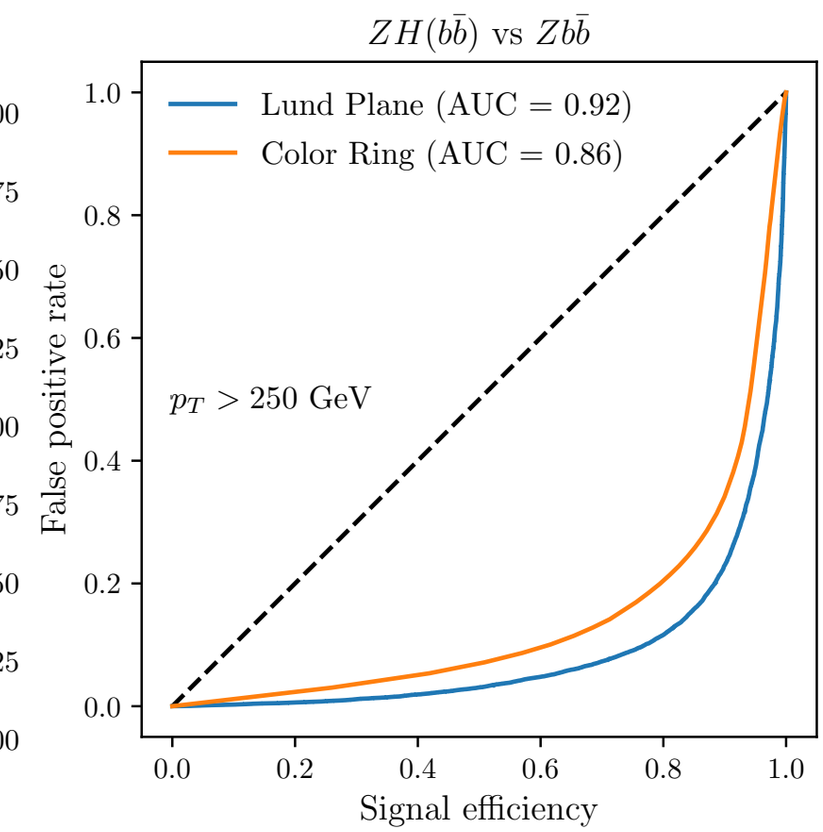
Summary and Outlook

- In this talk, we discussed how Lund Jet Plane and machine learning techniques can be used for jet tagging.
- We are using color-sensitive observables and a primary Lund jet plane for the Higgs tagging.
- Other studies also considered more complex architectures like graph neural networks for boosted jet tagging.
- A comparison of these techniques for the same benchmark will be useful for their implementation in the experimental analysis.

Tagger Without Jet Mass Cut







Mass unspecified tagger using color information rather than prong multiplicities