A new tagger for hadronically decaying heavy particles at the LHC

arxiv:1606.04961

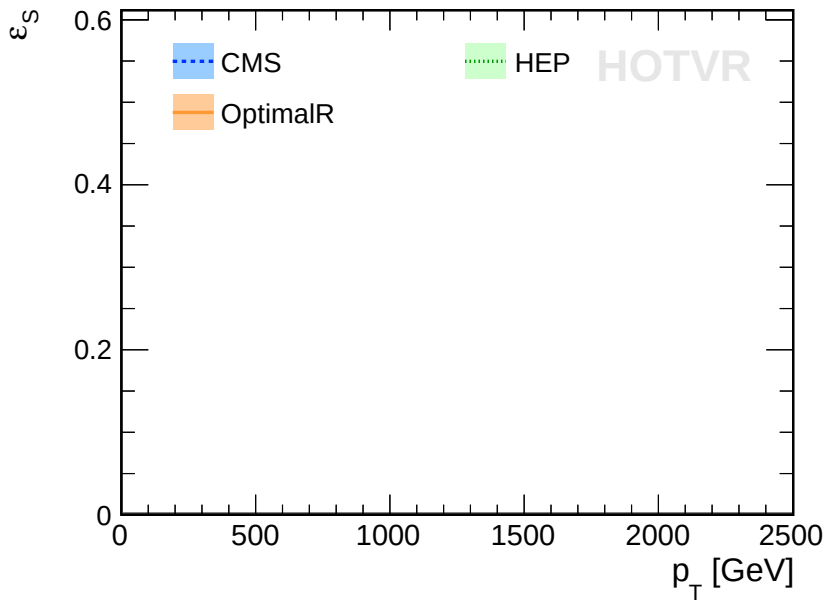
Tobias Lapsien, Roman Kogler, Johannes Haller

Universität Hamburg

BOOST 18.07.16 - 22.07.16

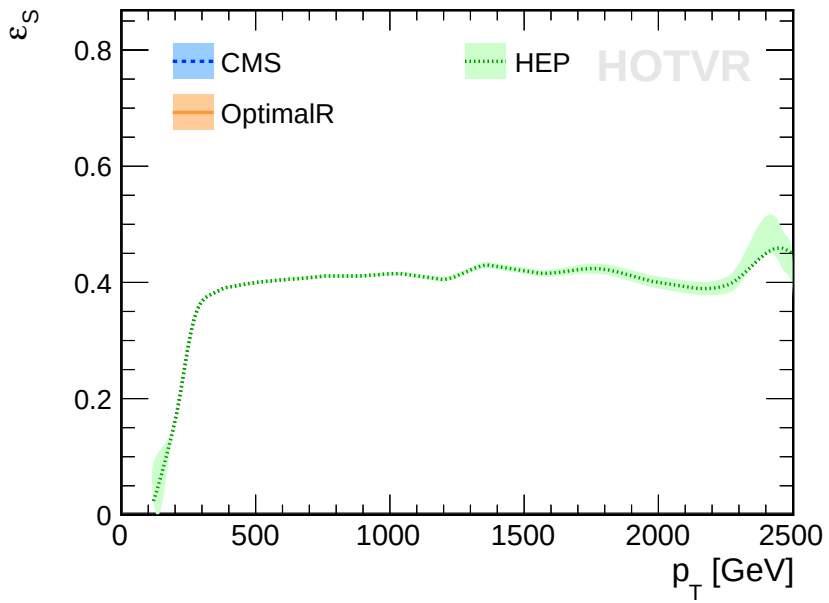
Motivation

Why a new top tagger?



Motivation

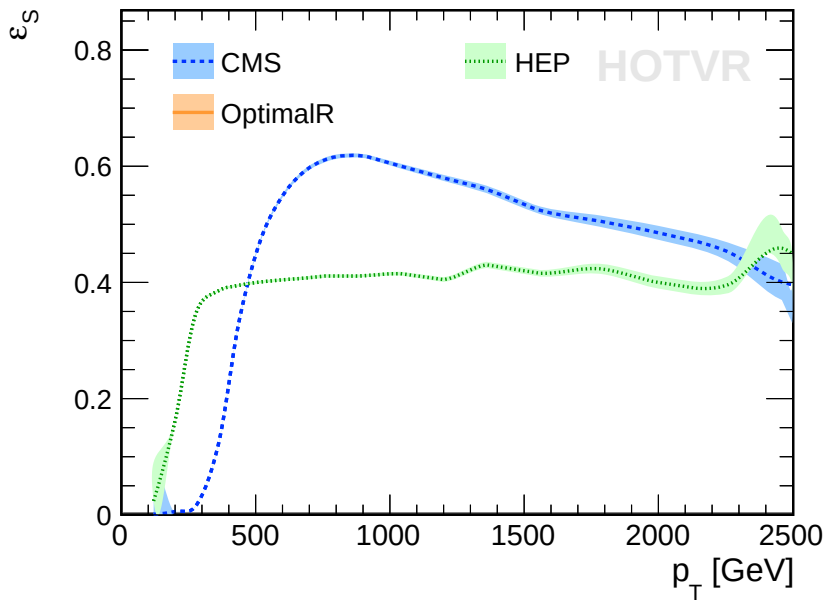
Why a new top tagger?



HEP TT
[JHEP 1010:078 (2010)]

Motivation

Why a new top tagger?



fixed conesize

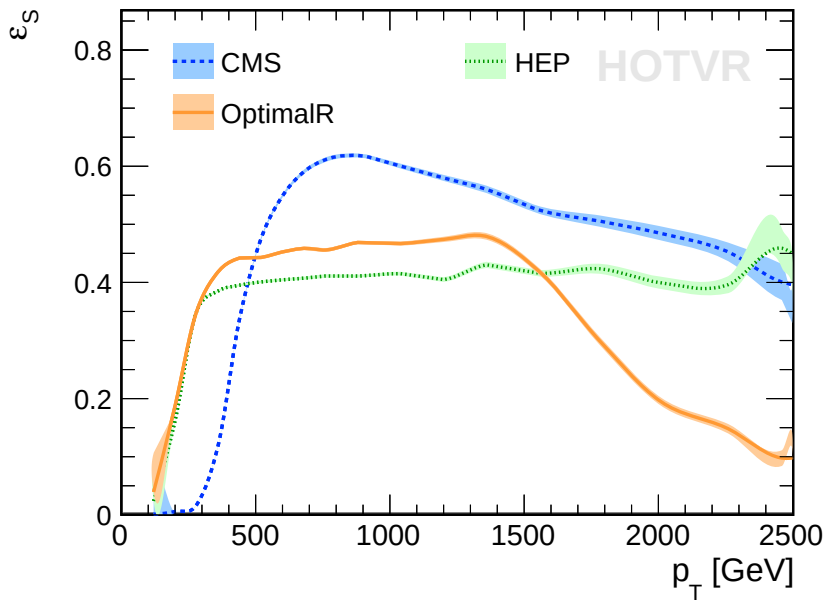
HEP TT
[JHEP 1010:078 (2010)]

CMS TT
[Phys.Rev.Lett.101:142001 (2008),CMS-PAS-JME-09-001 (2009)]

example: CMS Search
Two taggers used in different p_T regions
Phys. Rev. D 93 (2016) 012001

Motivation

Why a new top tagger?



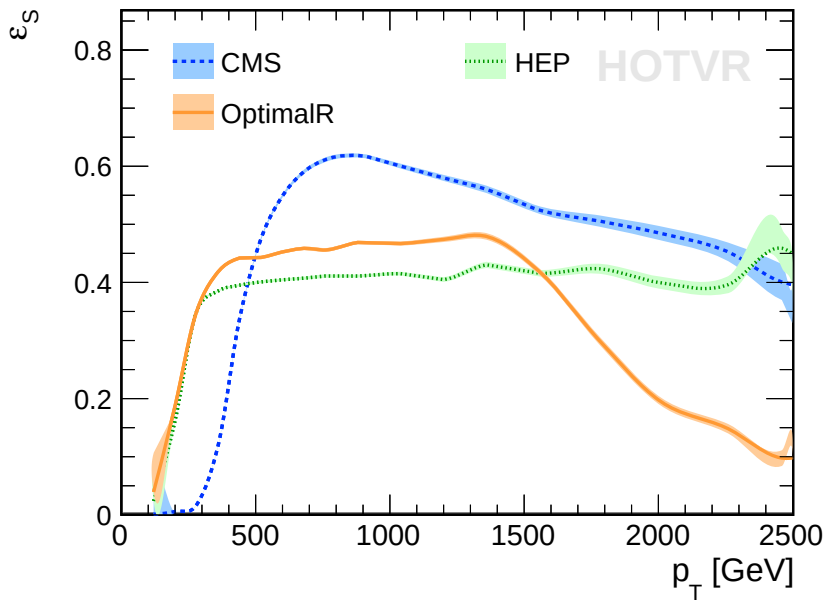
fixed conesize

- HEP TT [JHEP 1010:078 (2010)]
- CMS TT [Phys.Rev.Lett.101:142001 (2008),CMS-PAS-JME-09-001 (2009)]
- OptimalR HEP TT [JHEP 06 (2015) 203]

example: CMS Search
Two taggers used in different p_{T} regions
Phys. Rev. D 93 (2016) 012001

Motivation

Why a new top tagger?



fixed conesize

- HEP TT [JHEP 1010:078 (2010)]
- CMS TT [Phys.Rev.Lett.101:142001 (2008),CMS-PAS-JME-09-001 (2009)]
- OptimalR HEP TT [JHEP 06 (2015) 203]

example: CMS Search
Two taggers used in different p_T regions
Phys. Rev. D 93 (2016) 012001

Aim for new tagger

Top tagging history

Study combination of complex observables [CMS-PAS-JME-15-002]

- N-subjettness

  JHEP 03 (2011) 015

- Subjet b-tagging

  CMS-PAS-BTV-13-001

- Shower deconstruction

  PhysRevD.87.054012

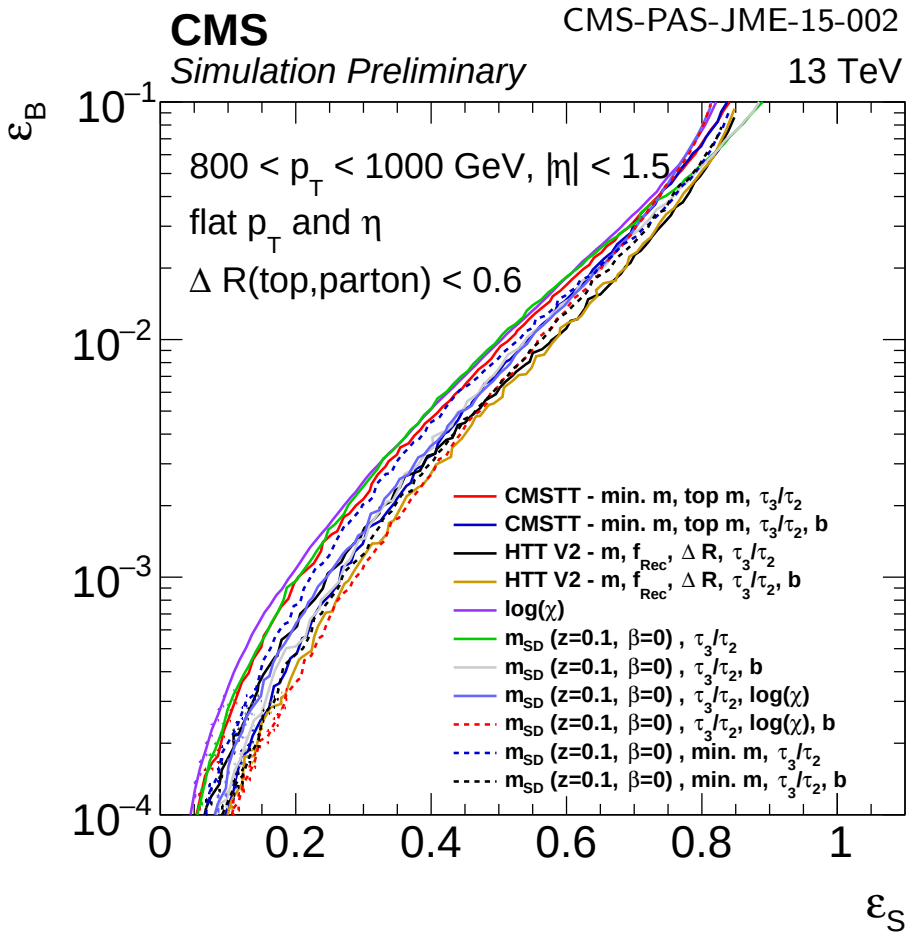
- OptimalR HEP tagger

  [JHEP 06 (2015) 203]

- Soft drop

  JHEP 1405 146 (2014)

Improvement moderate, cost of increased complexity



Top tagging history

Study combination of complex observables [CMS-PAS-JME-15-002]

- N-subjettness

  JHEP 03 (2011) 015

- Subjet b-tagging

  CMS-PAS-BTV-13-001

- Shower deconstruction

  PhysRevD.87.054012

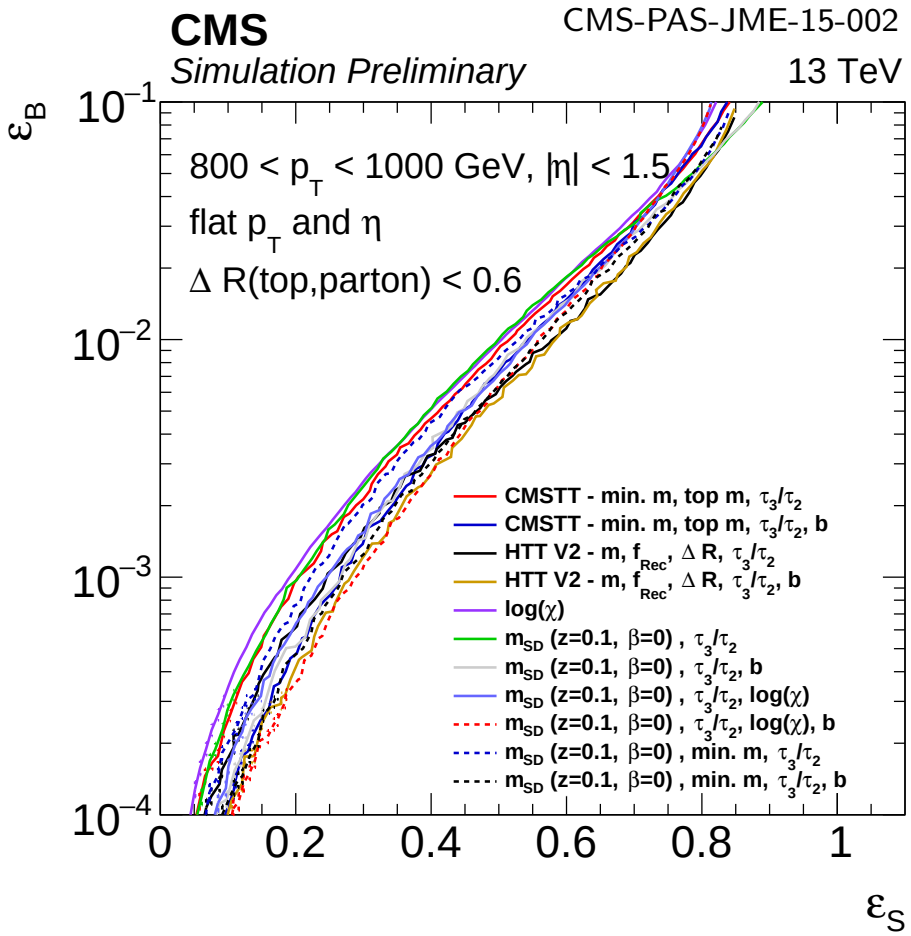
- OptimalR HEP tagger

  [JHEP 06 (2015) 203]

- Soft drop

  JHEP 1405 146 (2014)

Improvement moderate, cost of increased complexity



Goal

→ Top tagger with low complexity and good performance of a large p_{T} range

The Heavy Object Tagger with Variable R (HOTVR)

Main features HOTVR

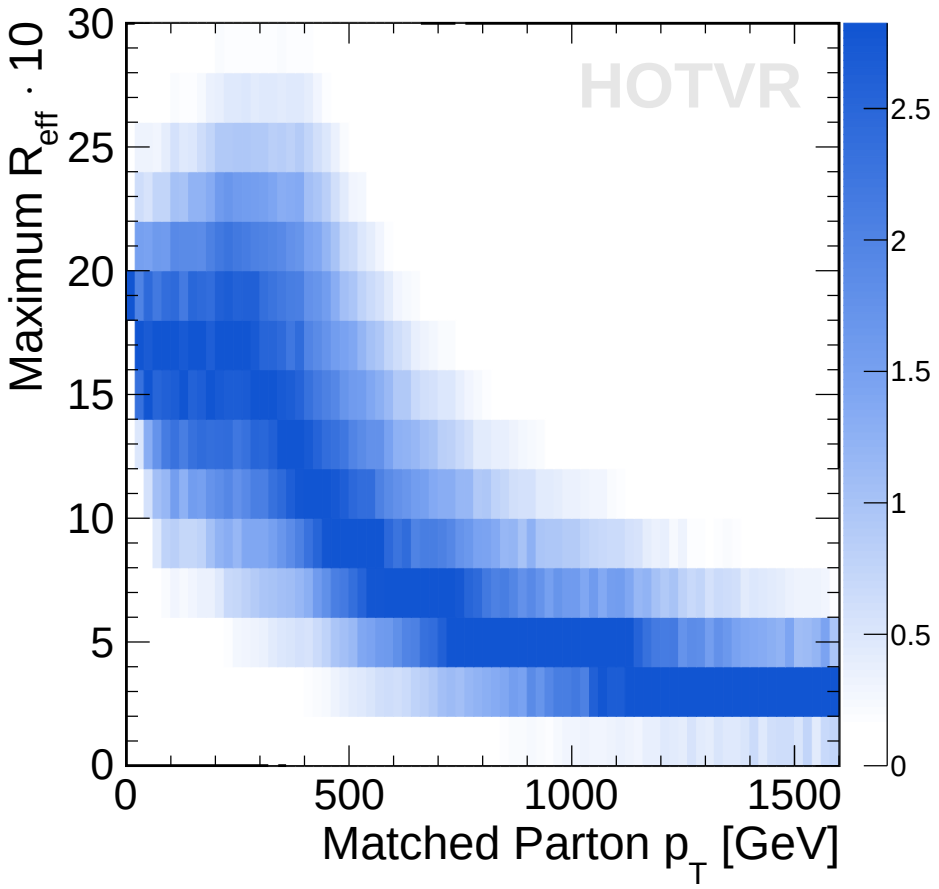
- Clustering with
  - Adaptive jet size
  - Subjet finding
  - Rejection of soft clusters

- Jets with Variable R

  JHEP 0906:059 (2009)

- Vetoed jet clustering: The mass-jump algorithm

  JHEP04 111 (2015)



The Heavy Object Tagger with Variable R (HOTVR)

- Use known distance measures (with Variable R)

$$d_{ij} = \min[p_{\mathrm{T},i}^{2n}, p_{\mathrm{T},j}^{2n}]\Delta R_{ij}^2$$

$$d_{i\mathrm{B}} = p_{\mathrm{T},i}^{2n} R_{\mathrm{eff}}^2 \qquad R_{\mathrm{eff}} = \frac{\rho}{p_{\mathrm{T}}}$$

- Veto condition for clustering step

$$m_{ij} < \mu$$

$$\theta \cdot m_{ij} > \max[m_i, m_j]$$

- Store i and j as subjets if

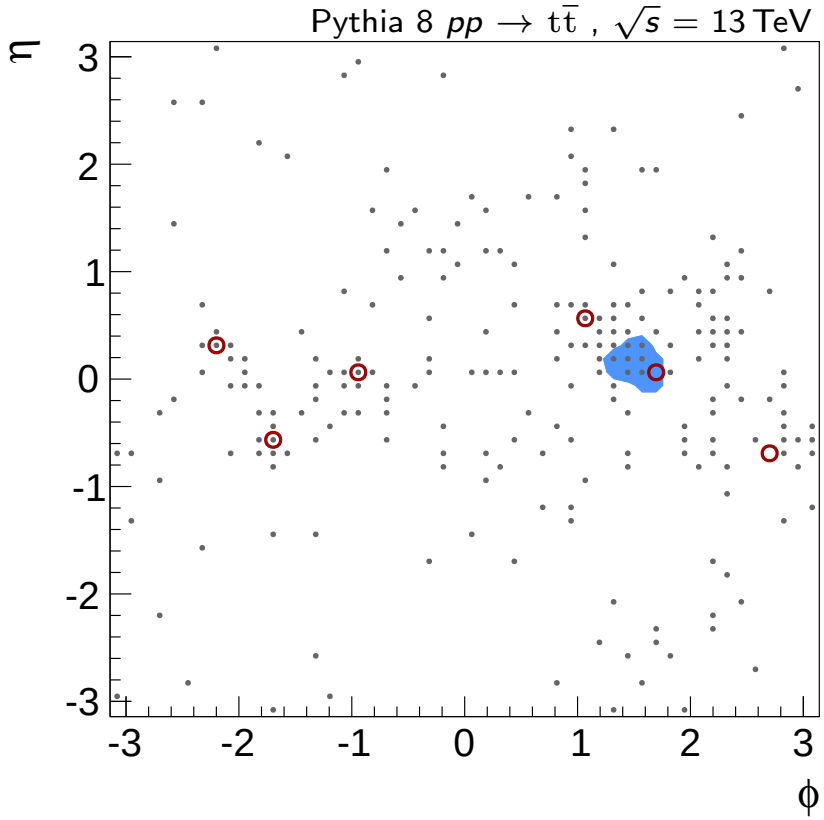
$$p_{\mathrm{T},i}, p_{\mathrm{T},j} > p_{\mathrm{T,sub}}$$

Parameter in top tagging mode

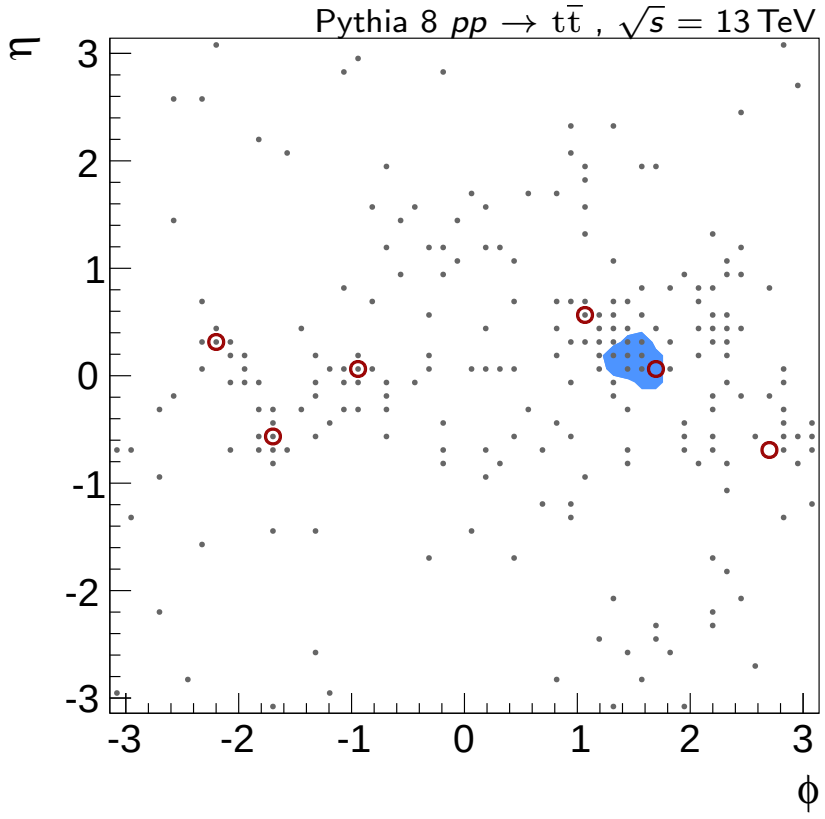
- $R_{\mathrm{min}} = 0.1$: Minimum radius
- $R_{\mathrm{max}} = 1.5$: Maximum radius
- $\rho = 600\,\mathrm{GeV}$: Slope of R_{eff}
- $\mu = 30\,\mathrm{GeV}$: Mass jump threshold
- $\theta = 0.7$: Mass jump strength
- $p_{\mathrm{T,sub}} = 30\,\mathrm{GeV}$: Minimum subjet p_{T}

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



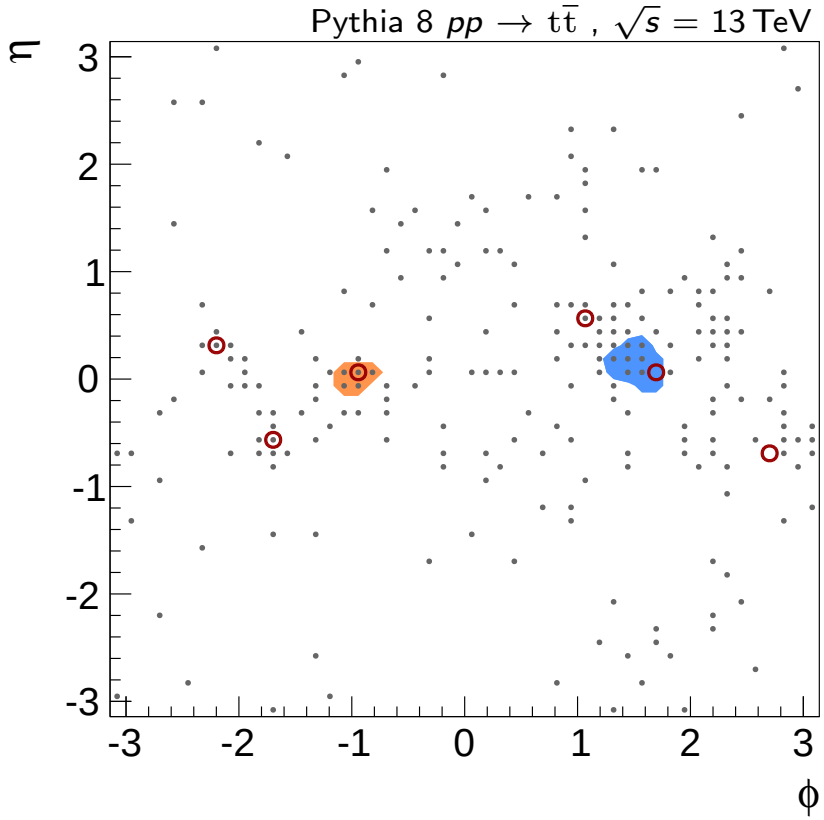
HOTVR clustering



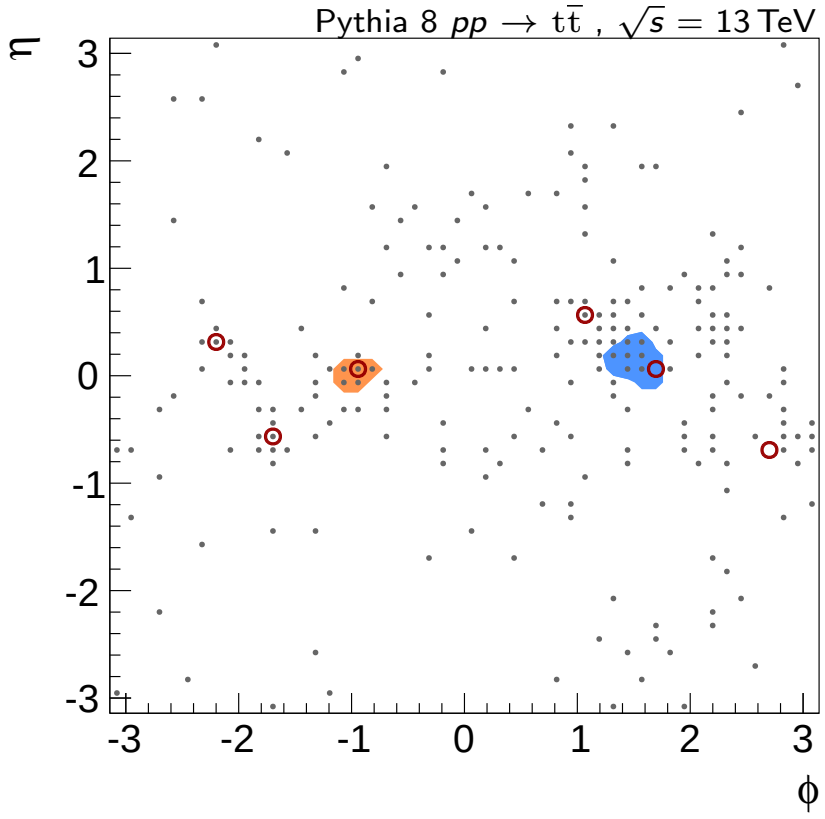
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



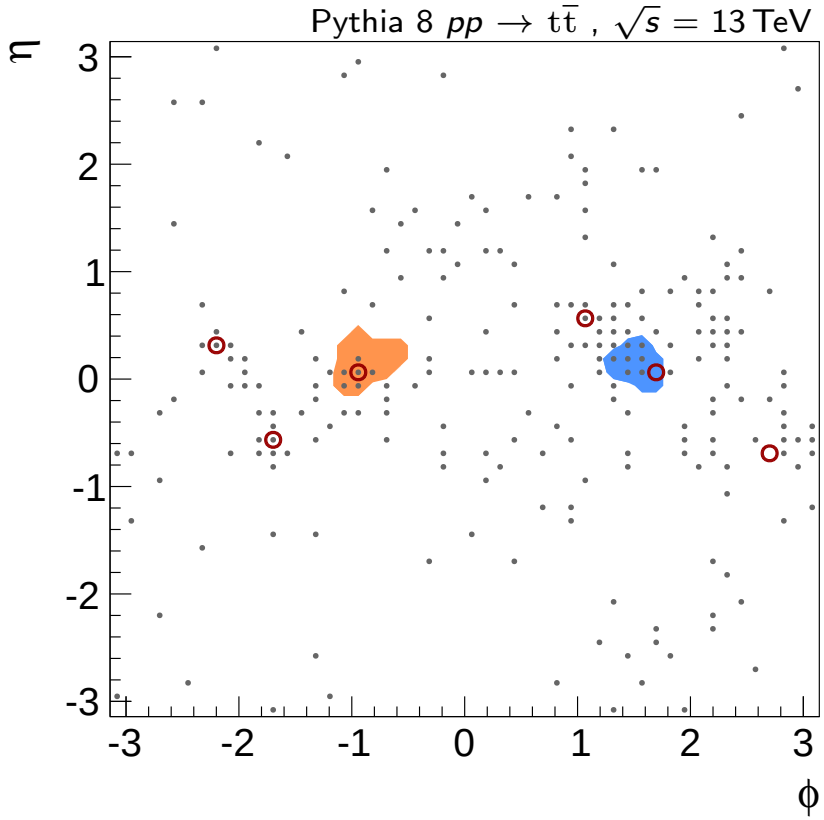
HOTVR clustering



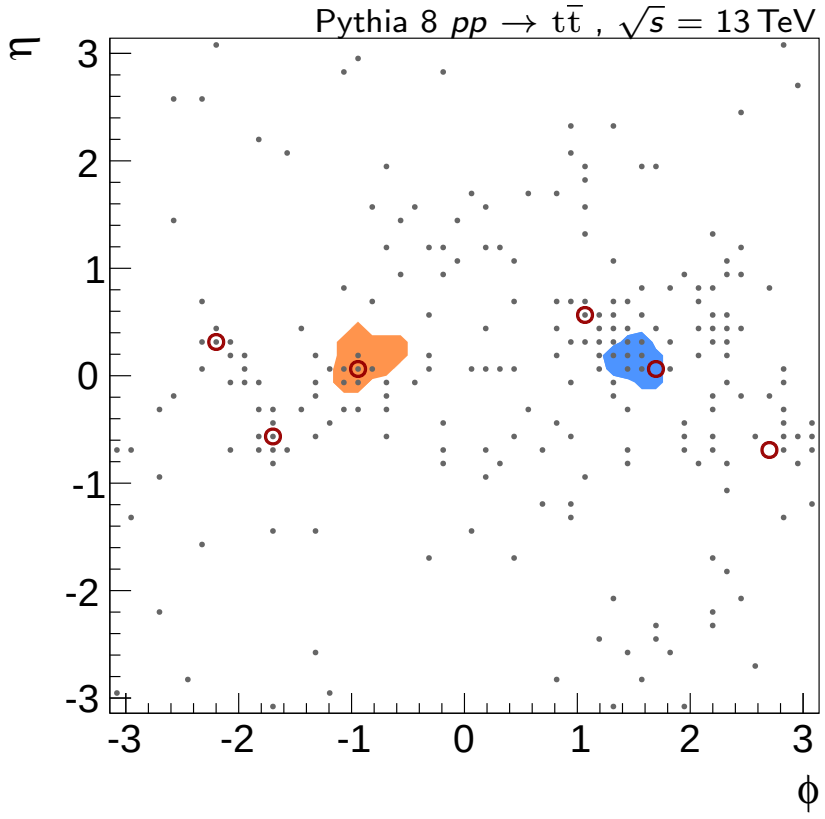
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



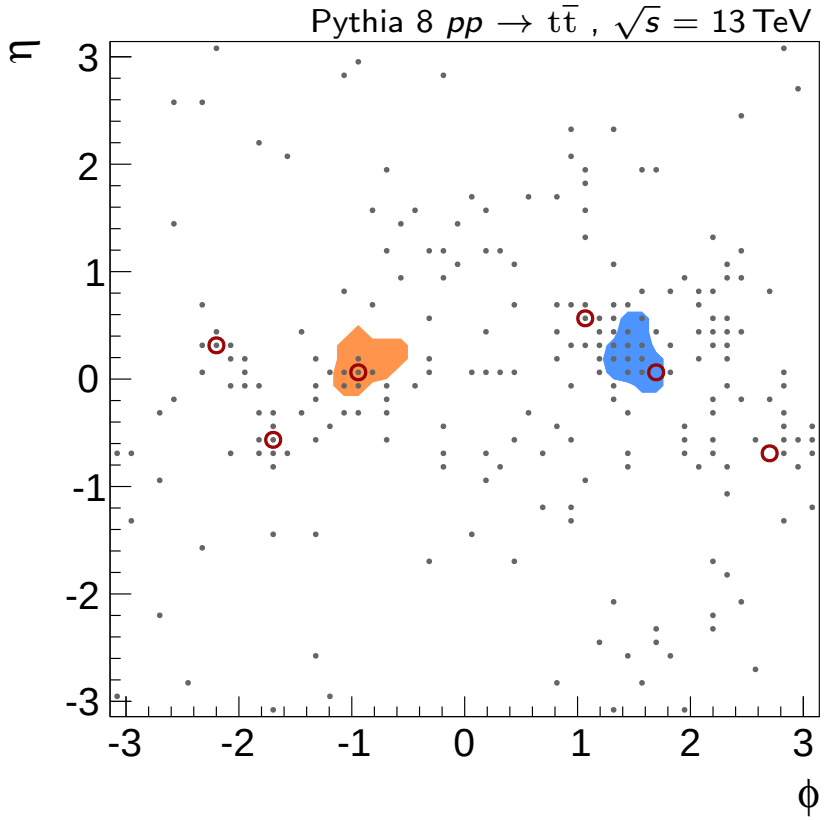
HOTVR clustering



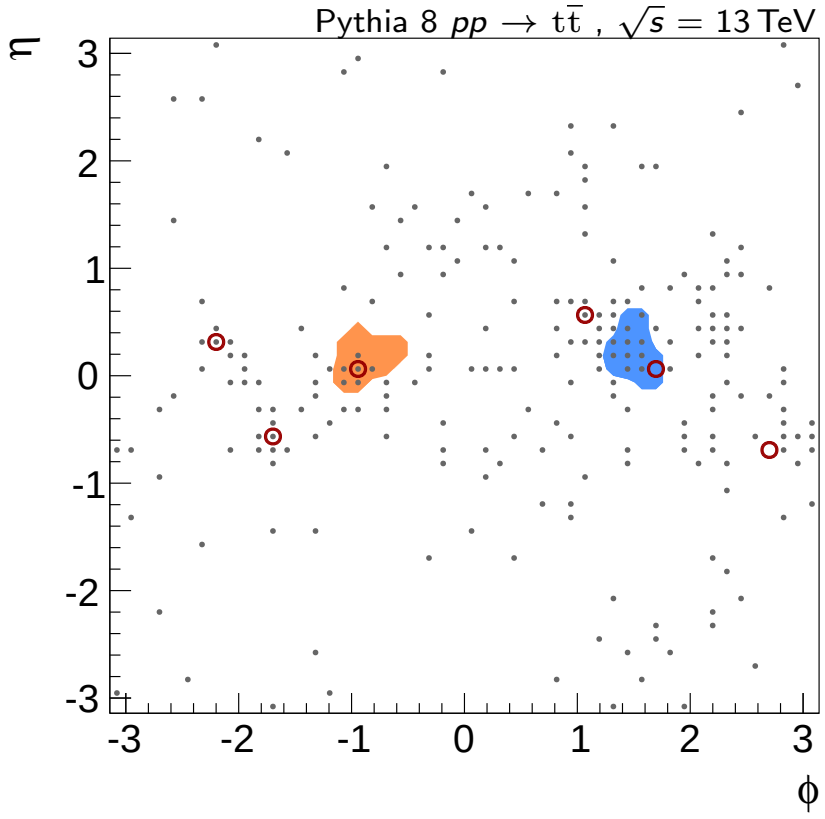
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



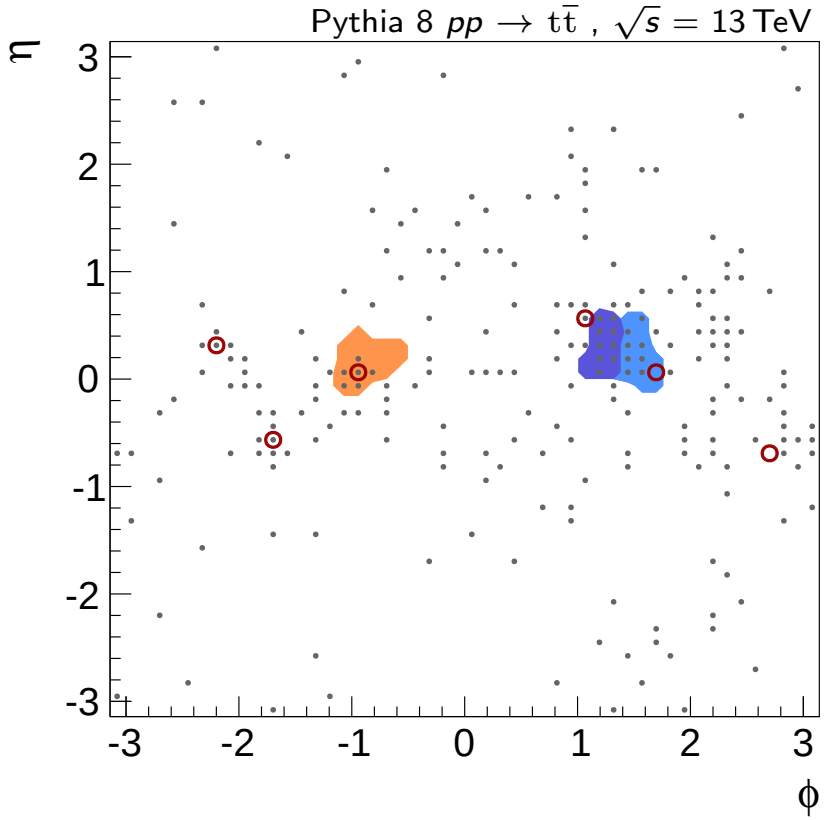
HOTVR clustering



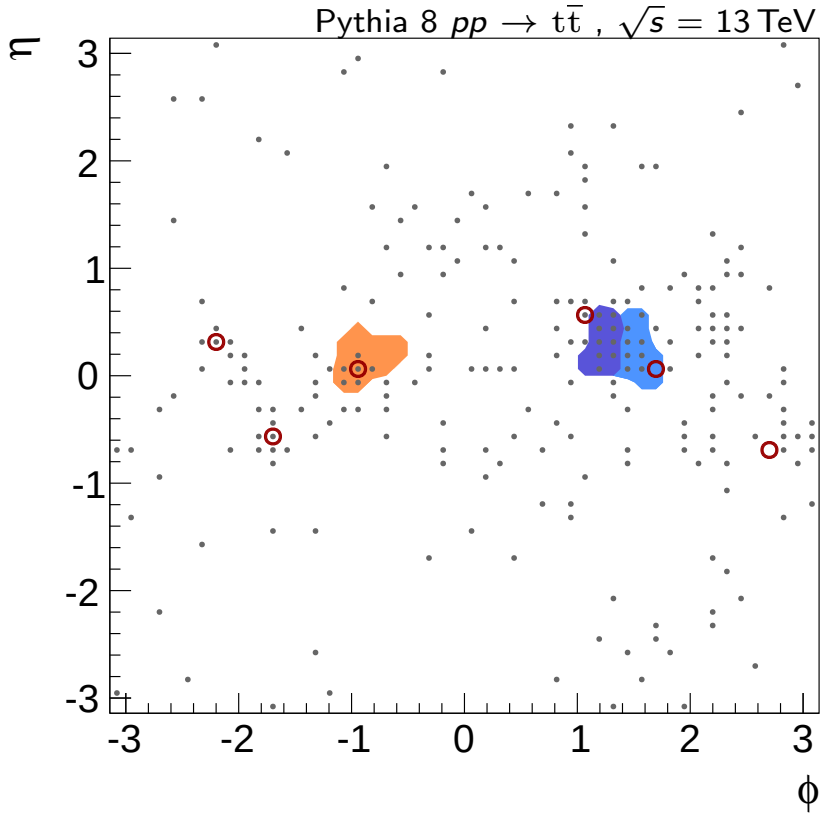
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



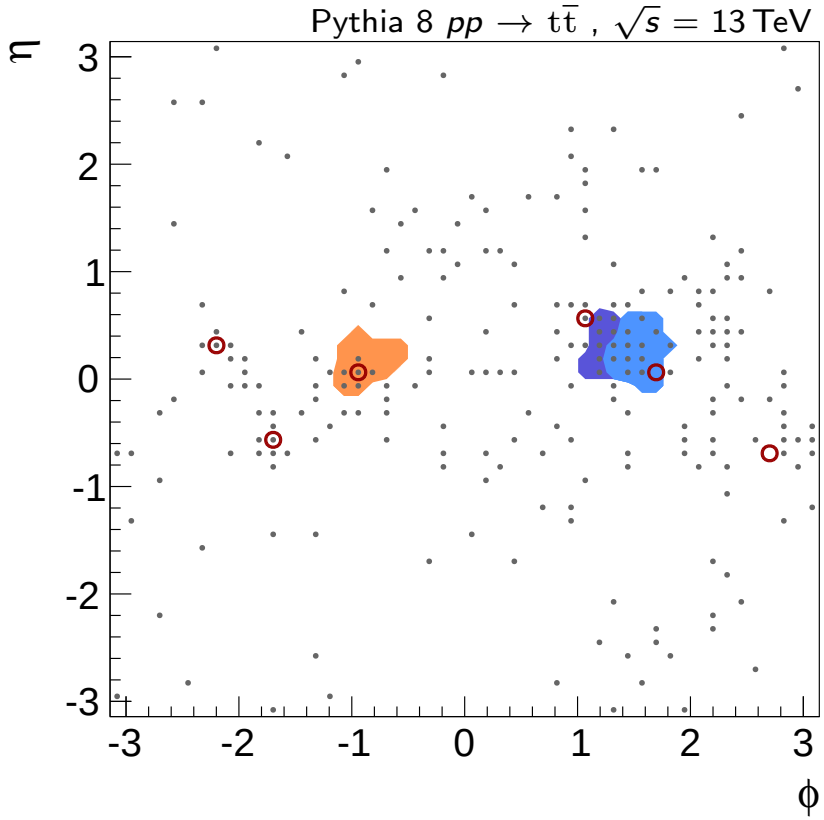
HOTVR clustering



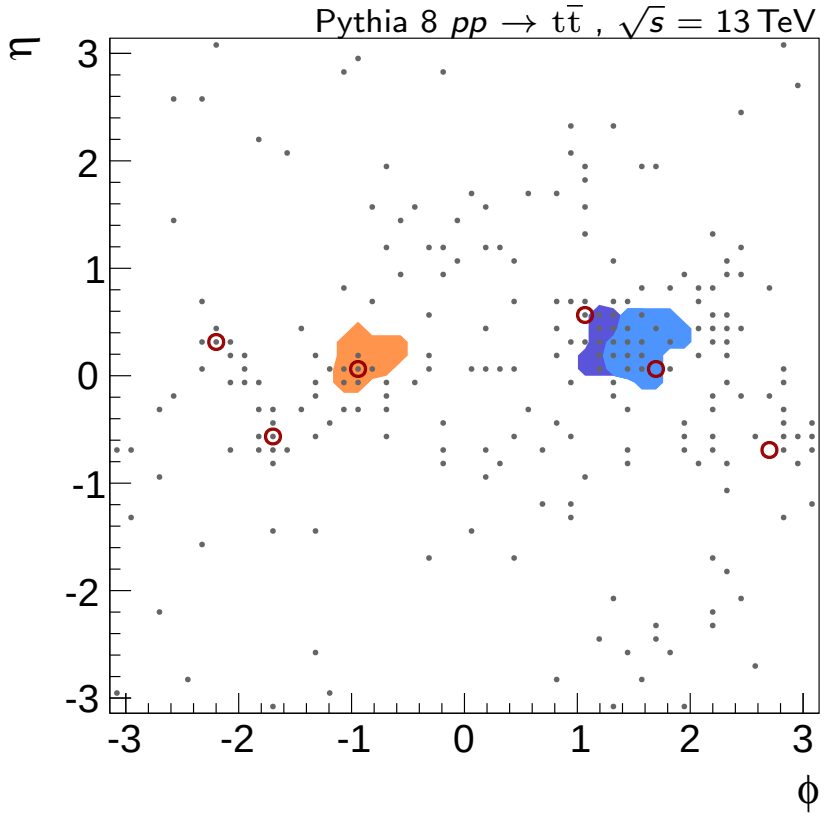
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



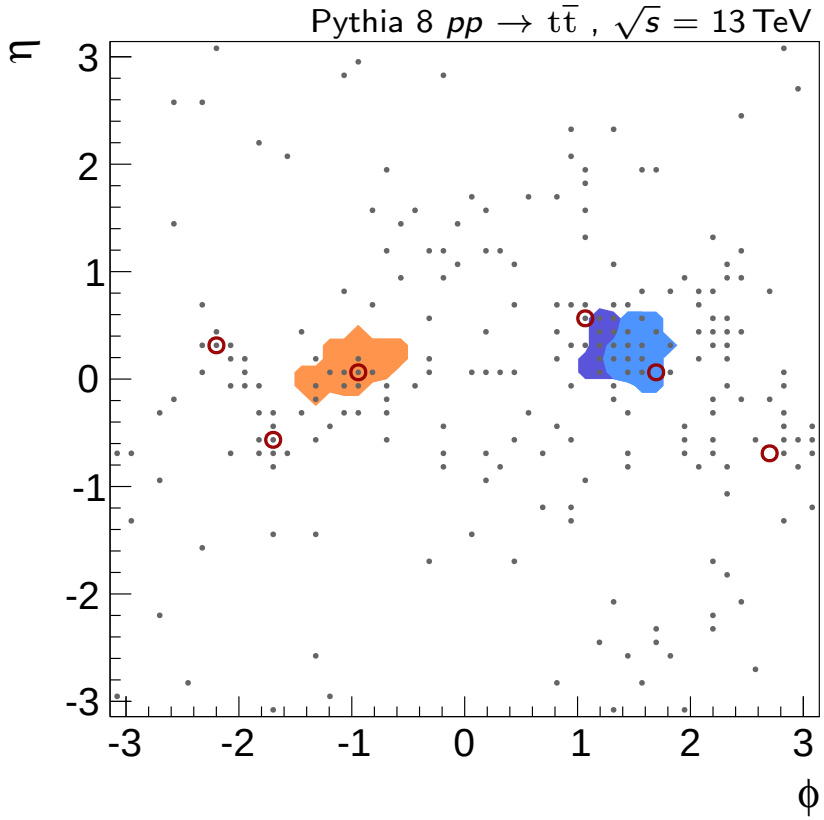
HOTVR clustering



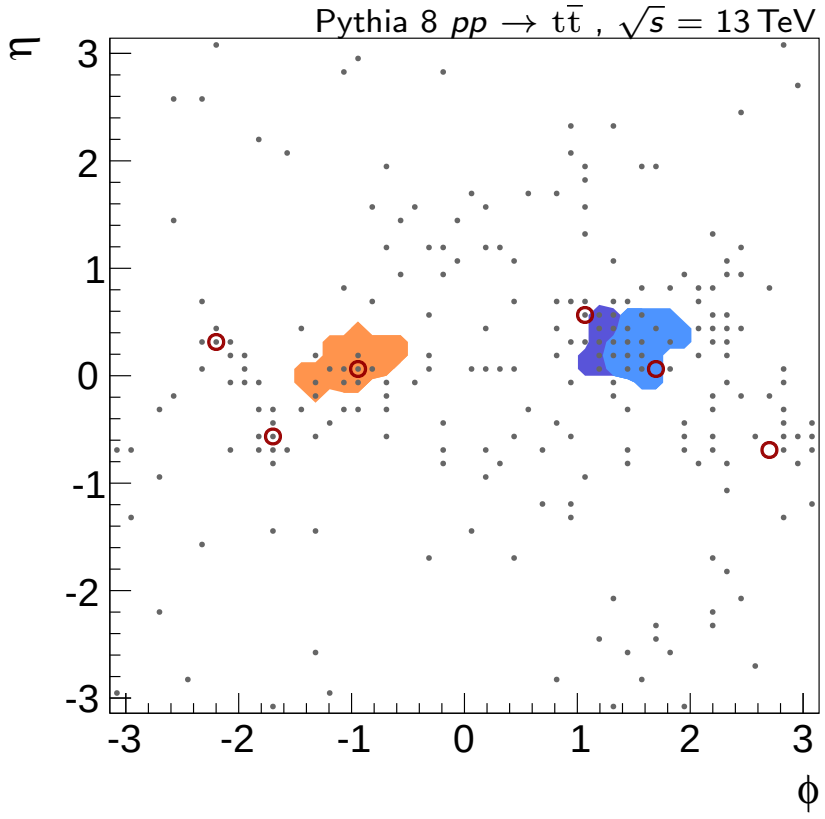
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



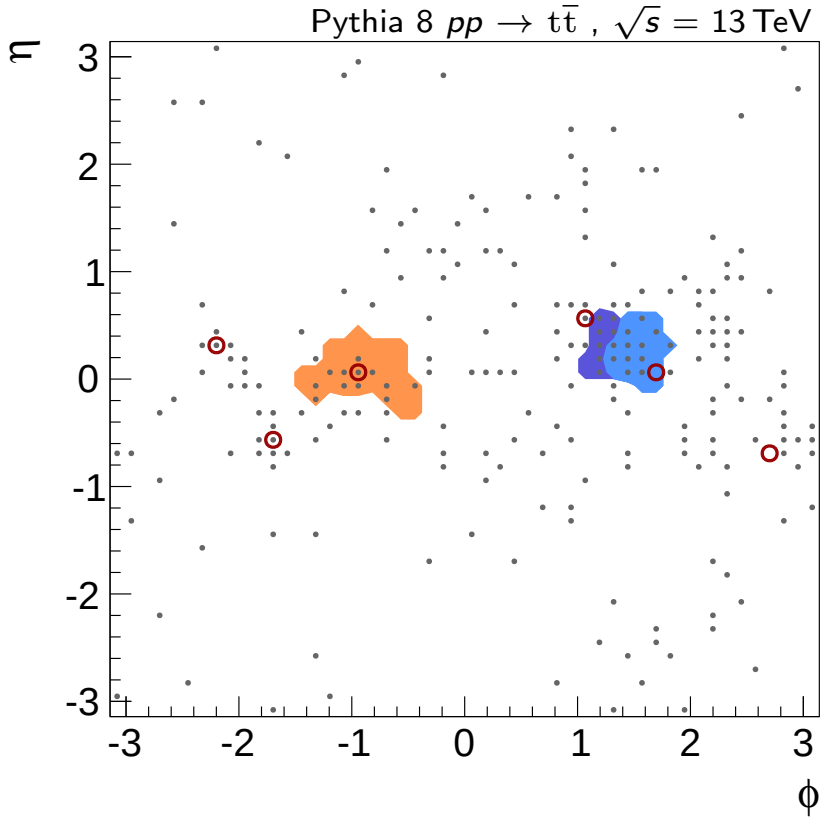
HOTVR clustering



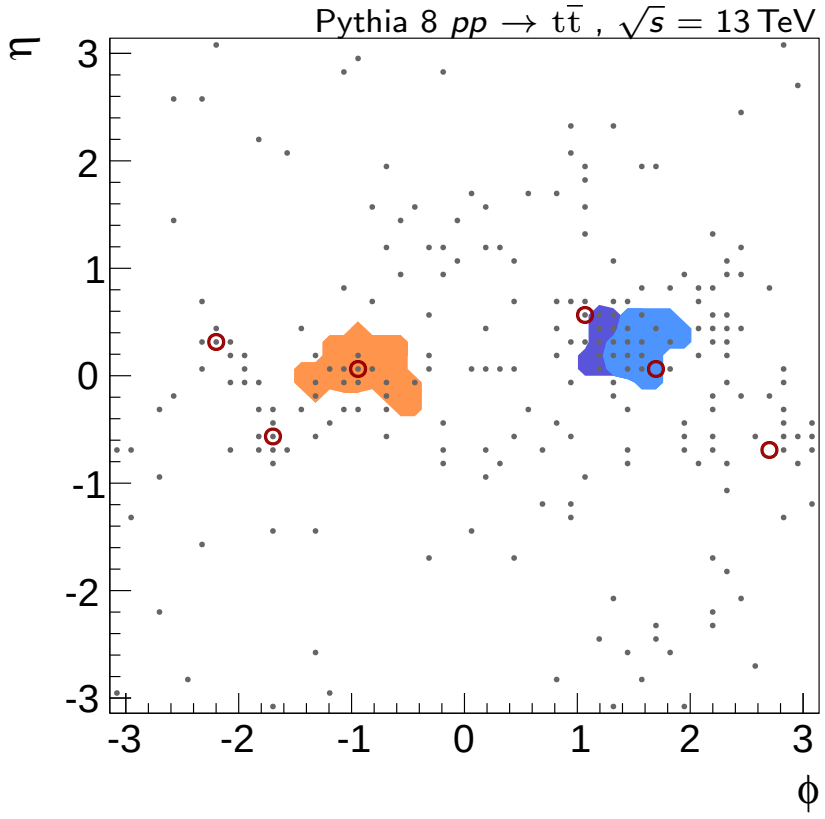
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



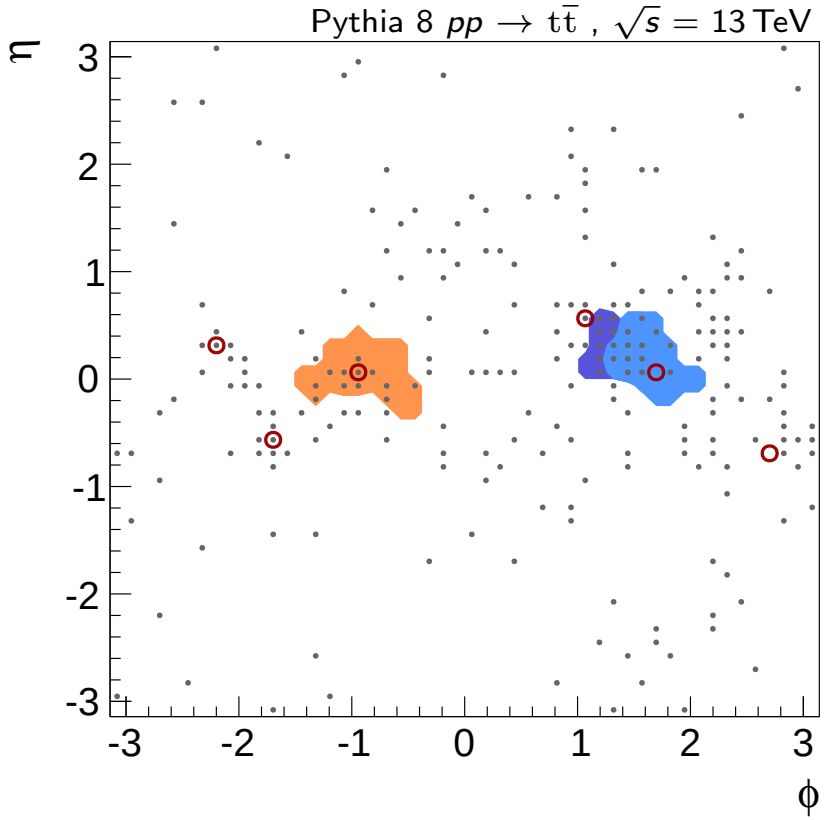
HOTVR clustering



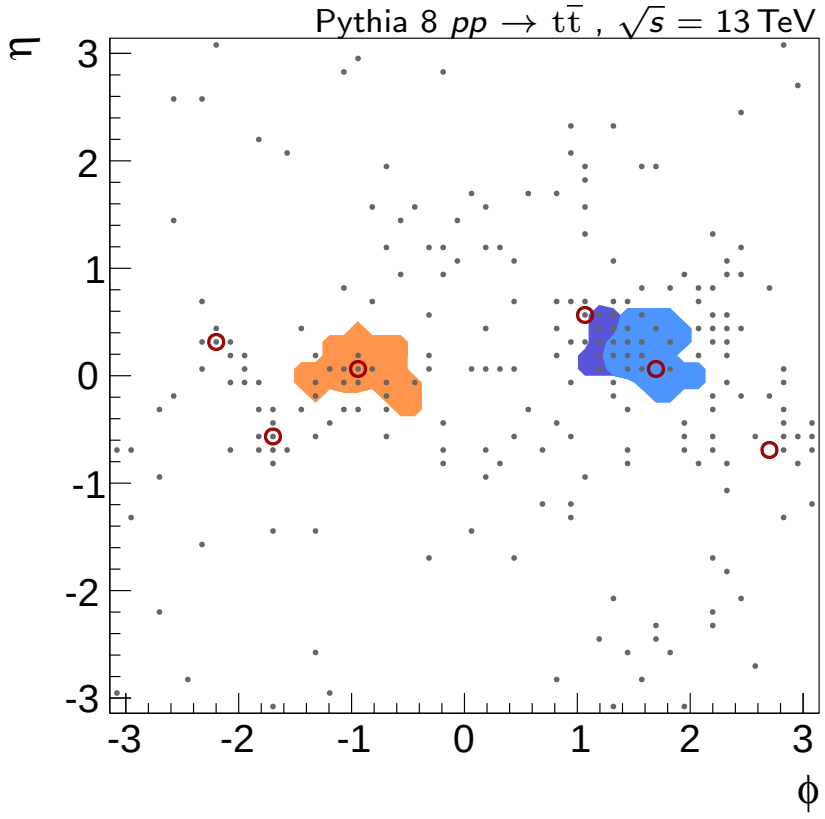
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



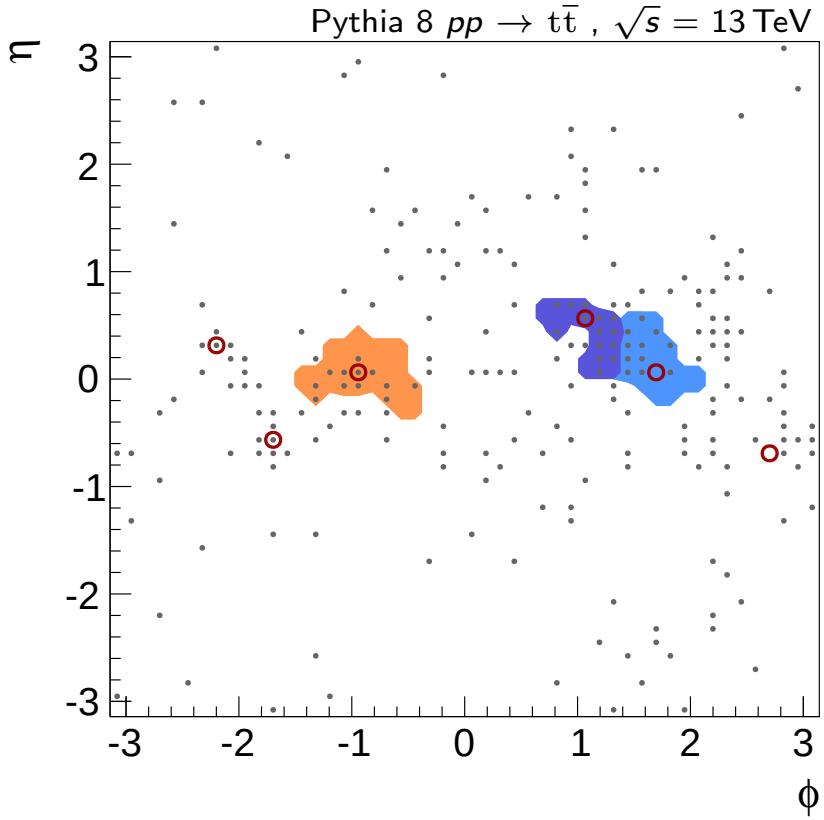
HOTVR clustering



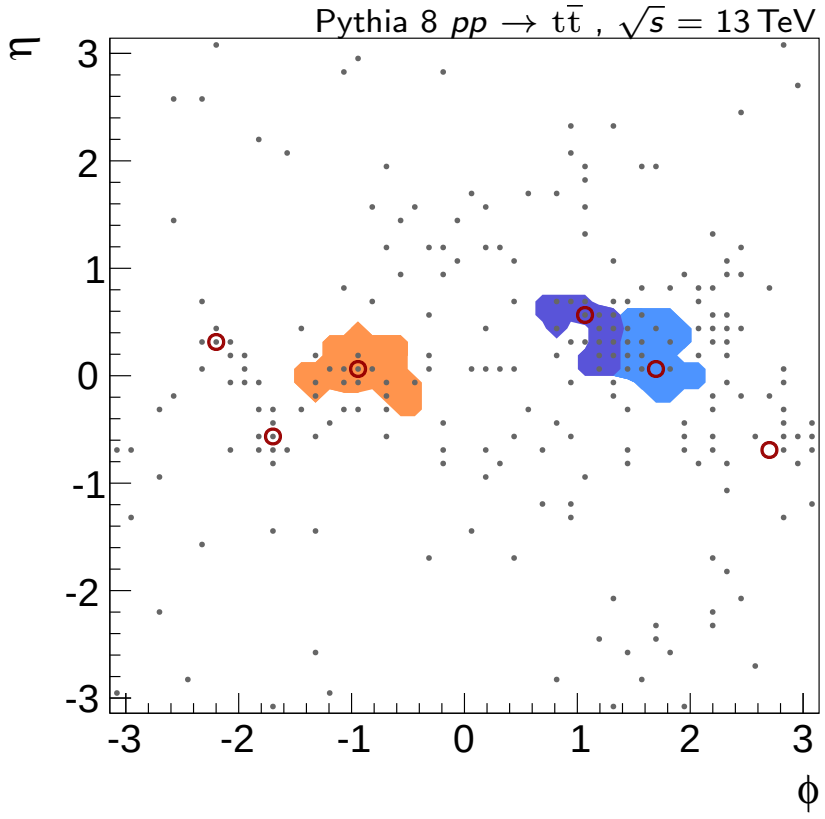
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



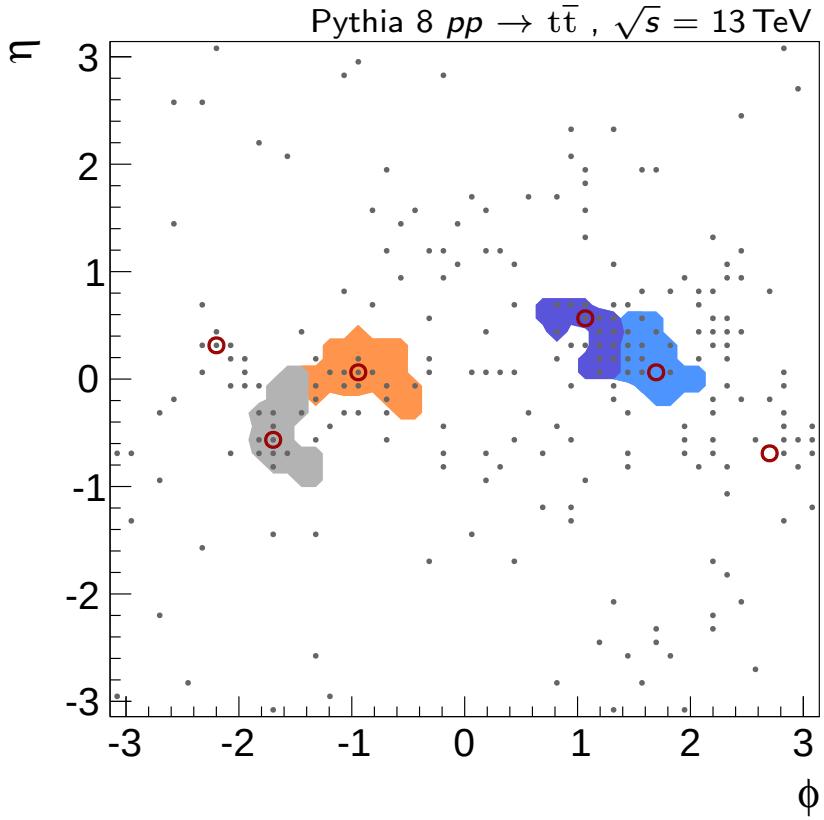
HOTVR clustering



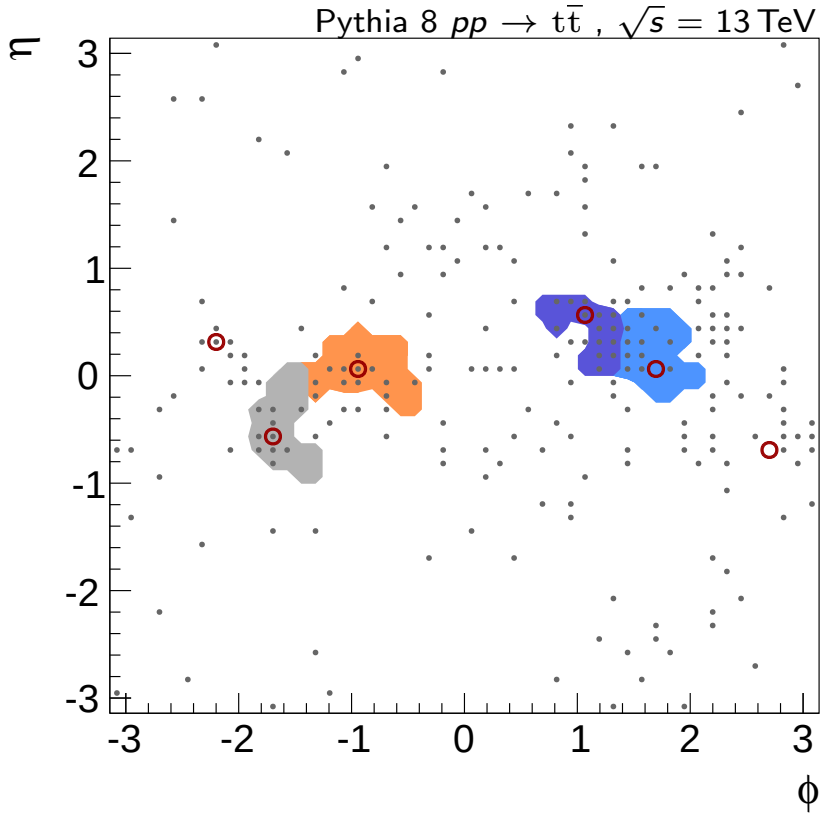
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



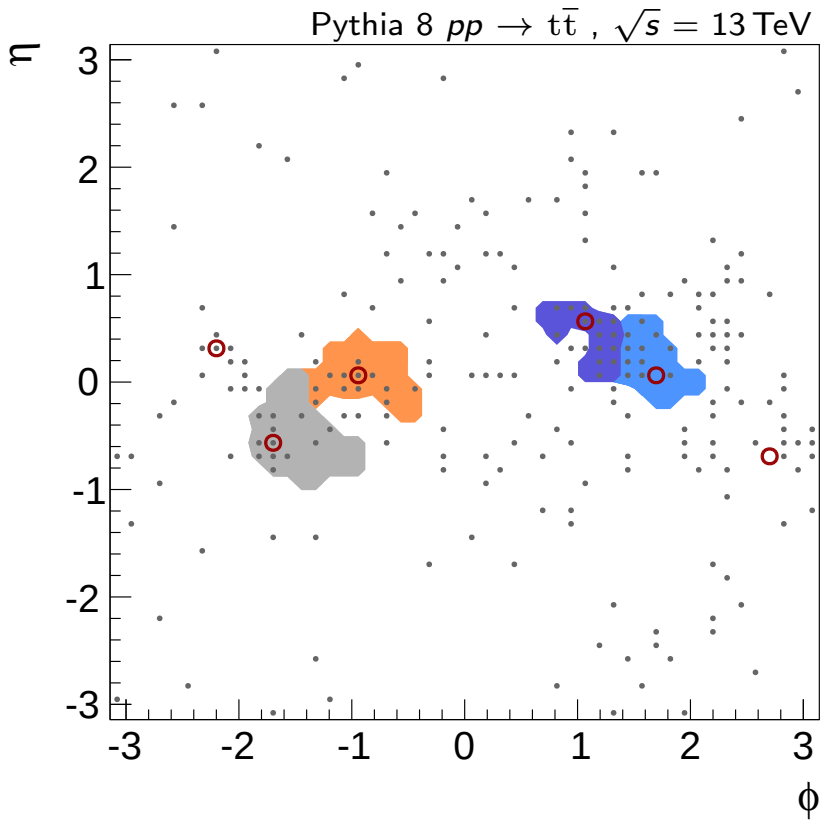
HOTVR clustering



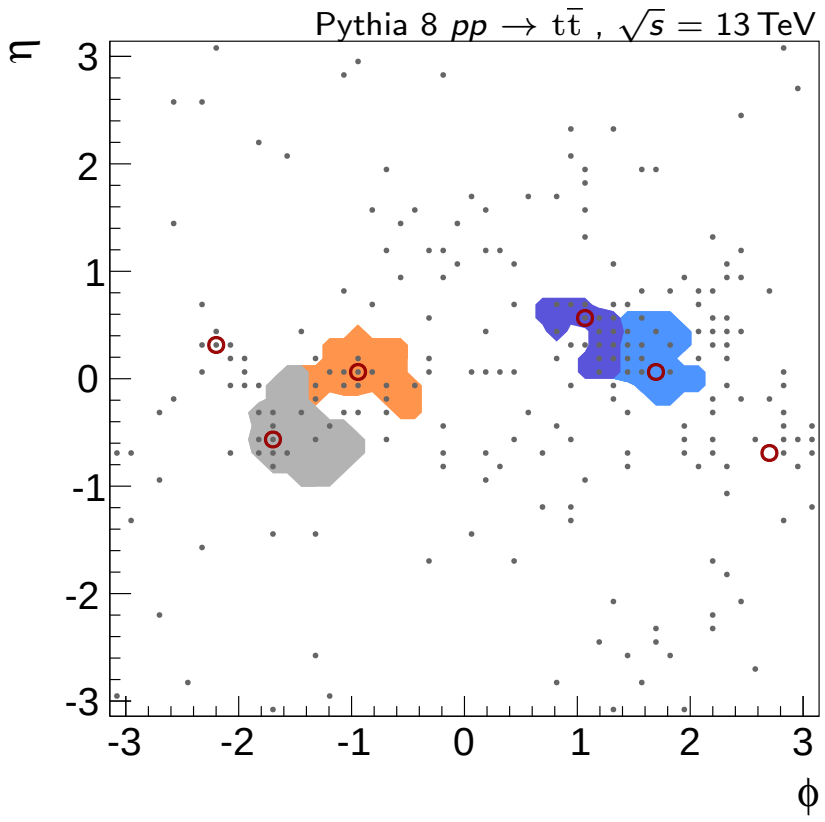
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



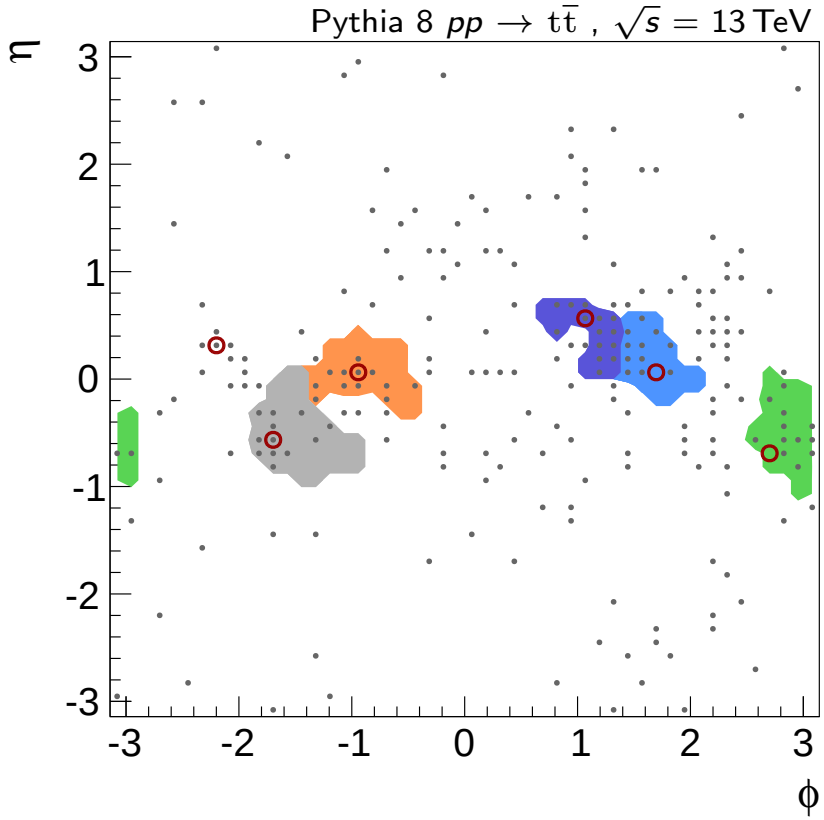
HOTVR clustering



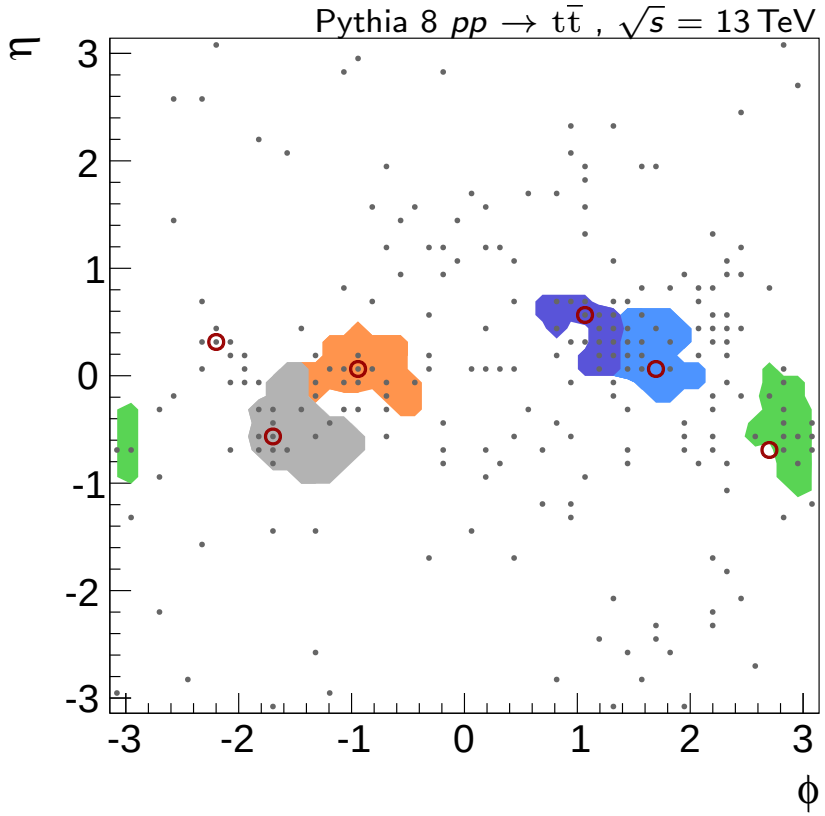
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



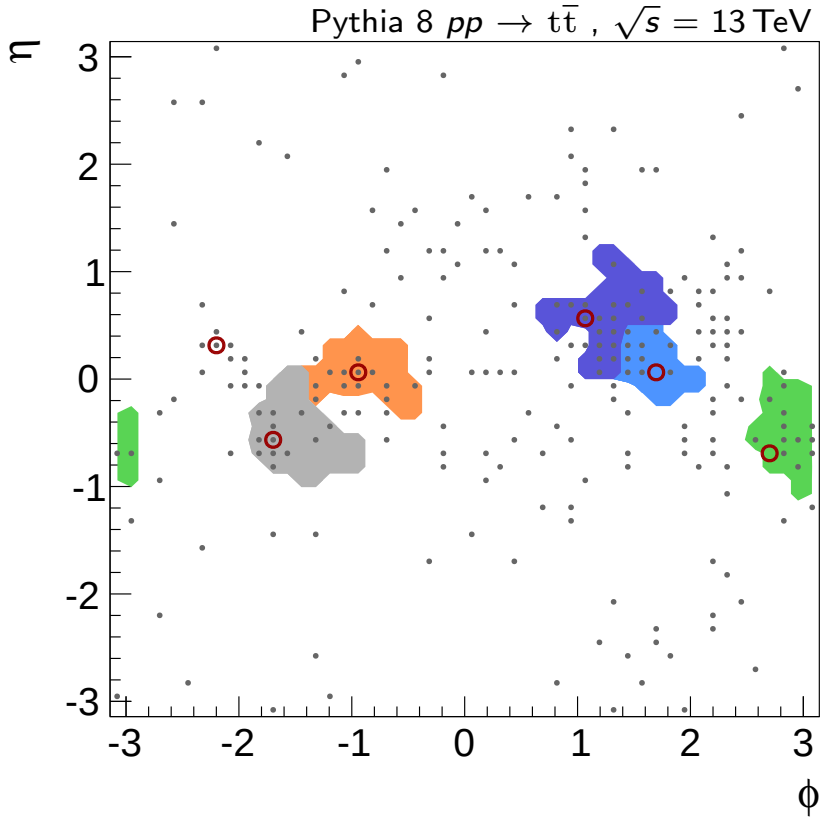
HOTVR clustering



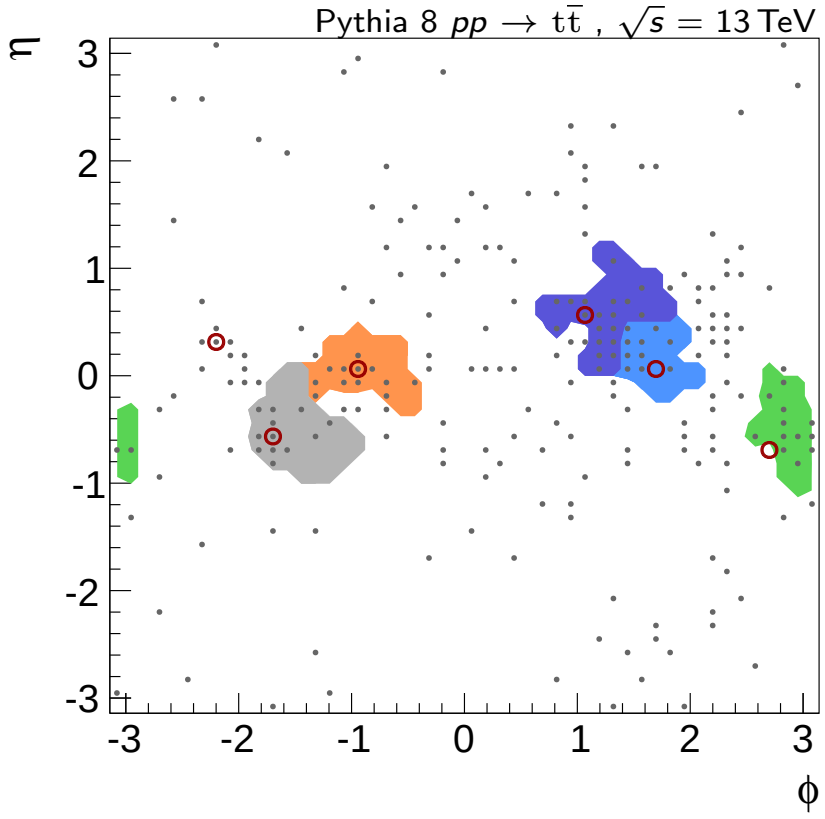
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



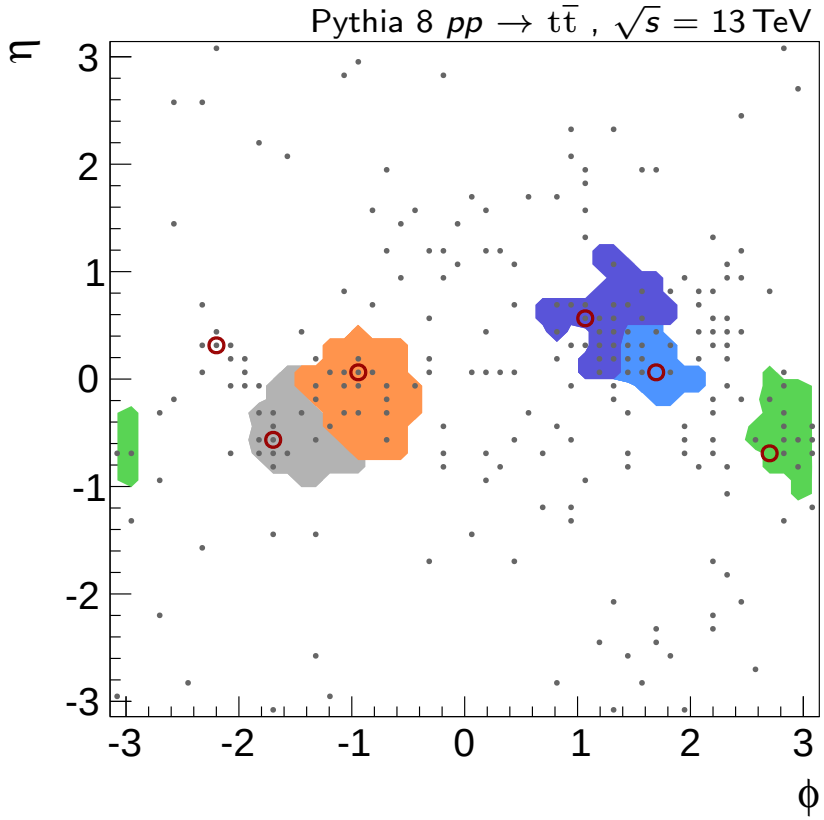
HOTVR clustering



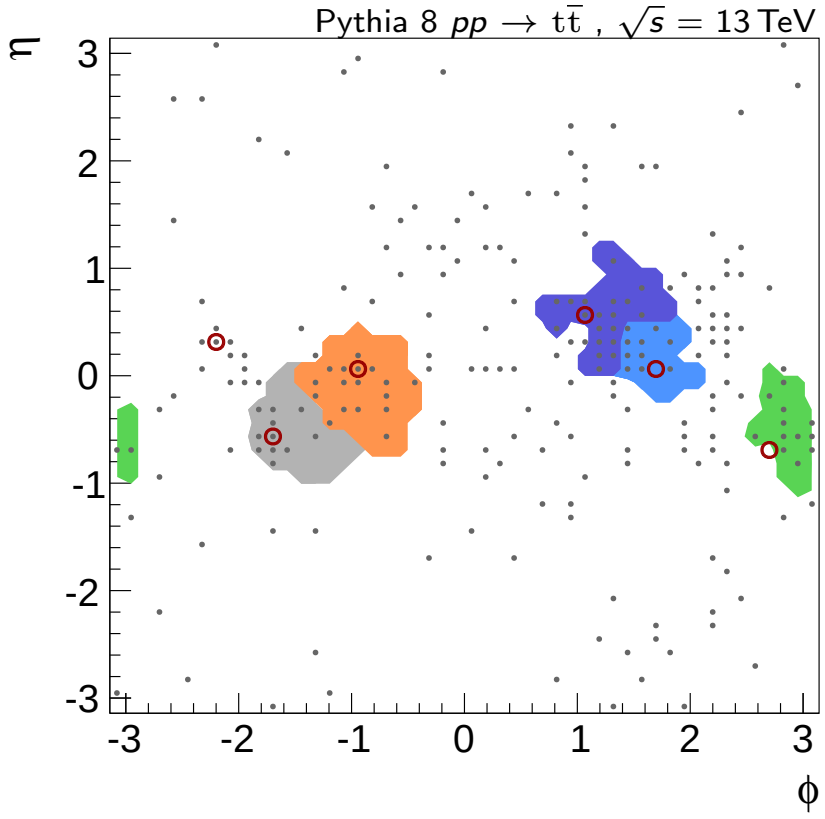
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



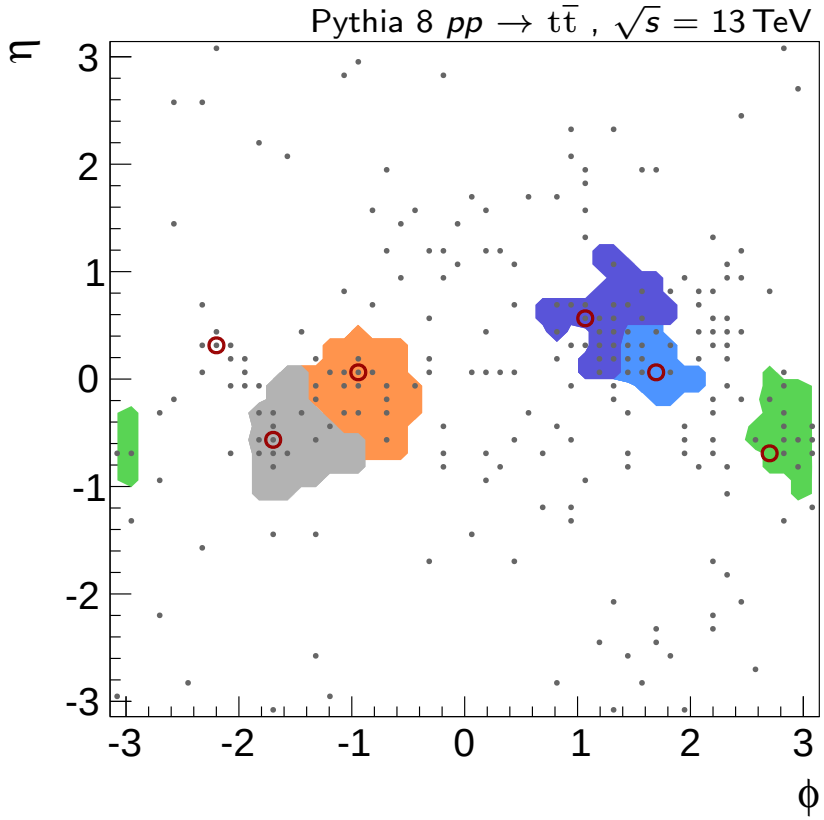
HOTVR clustering



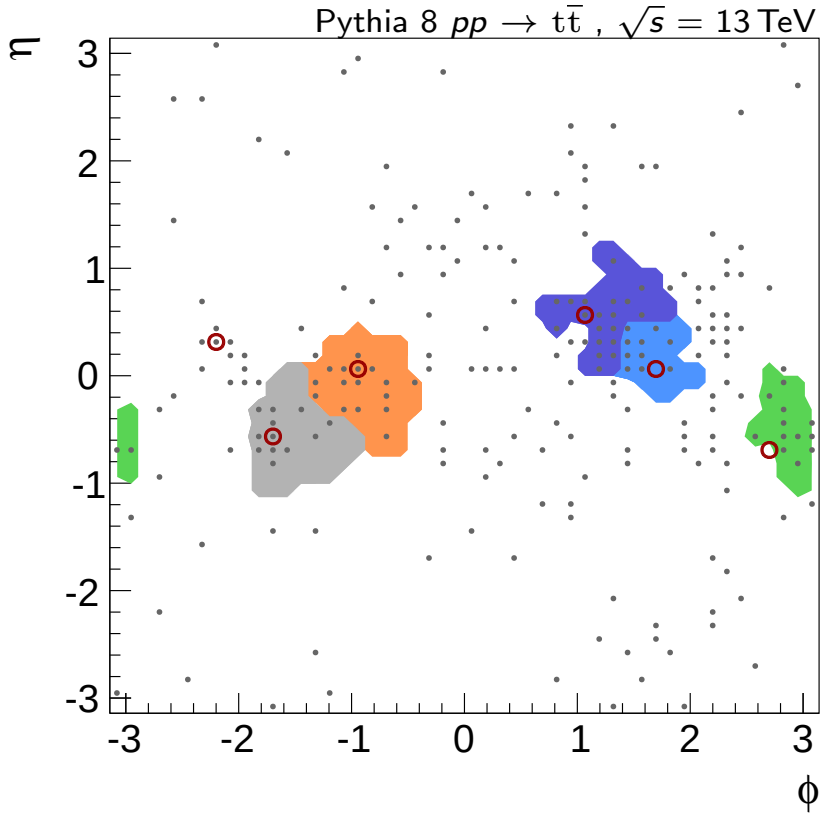
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



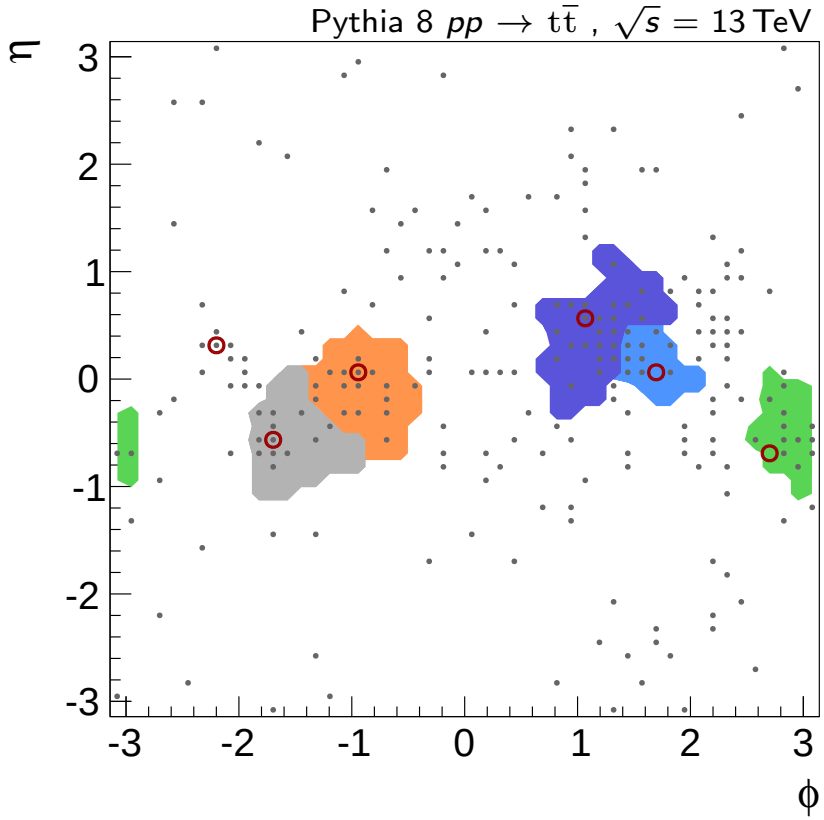
HOTVR clustering



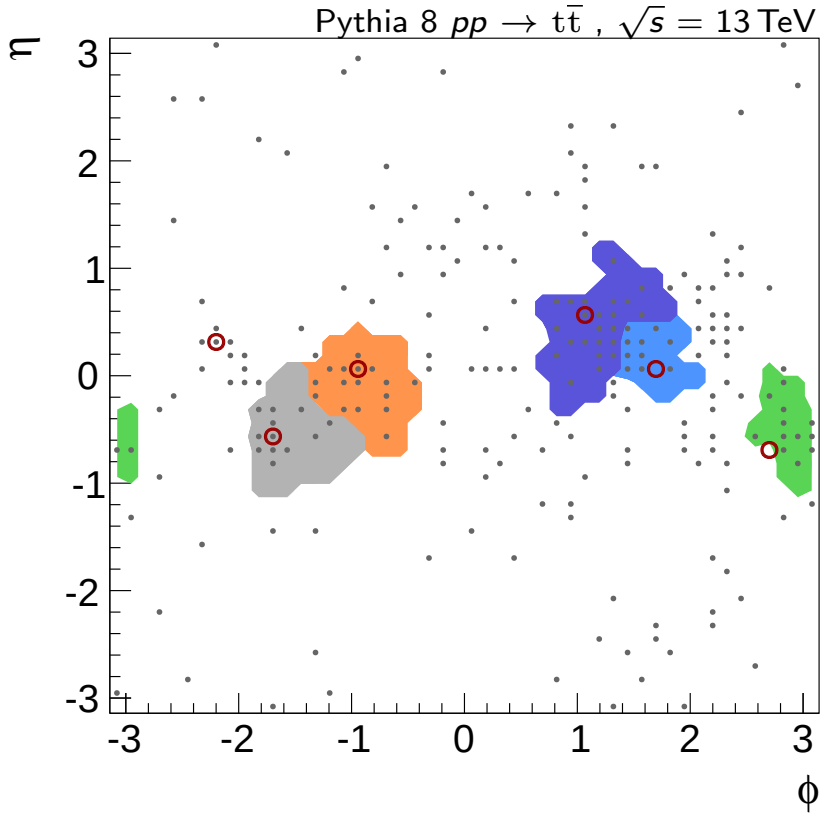
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



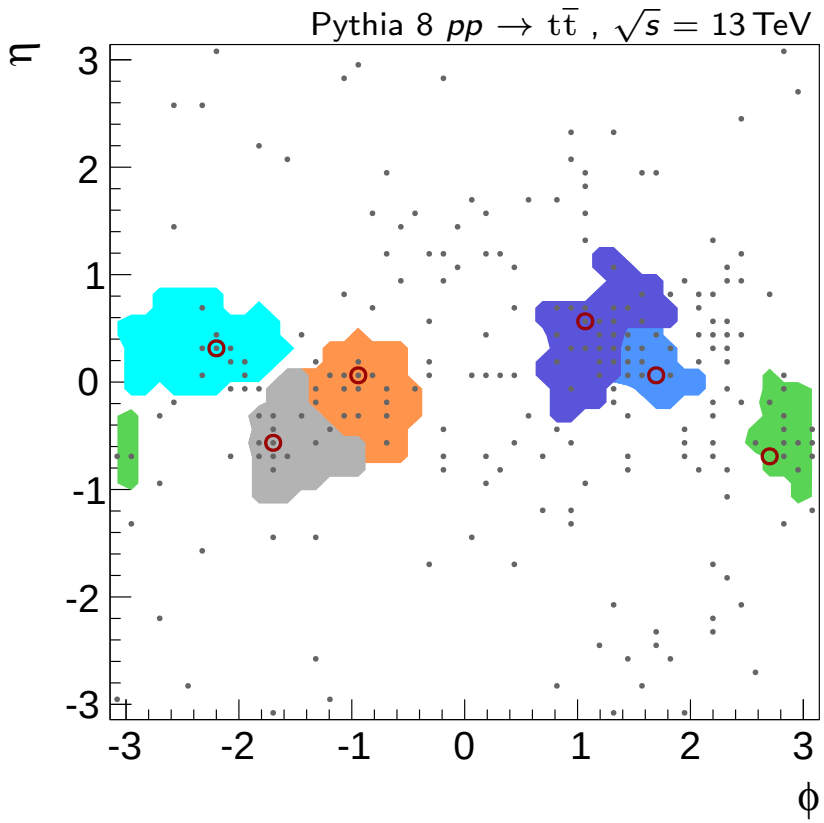
HOTVR clustering



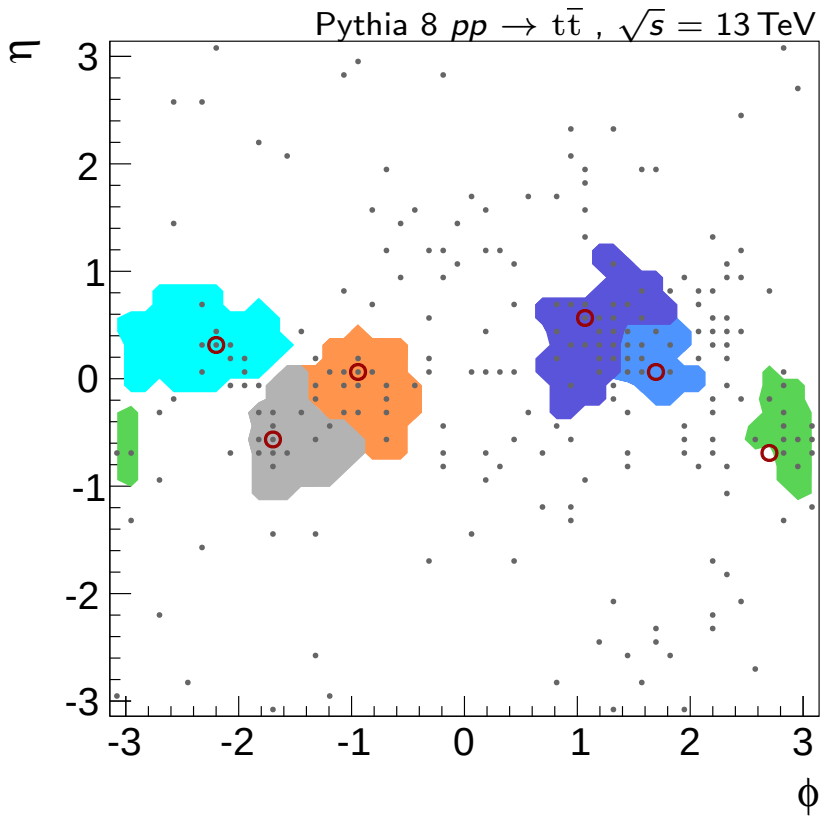
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



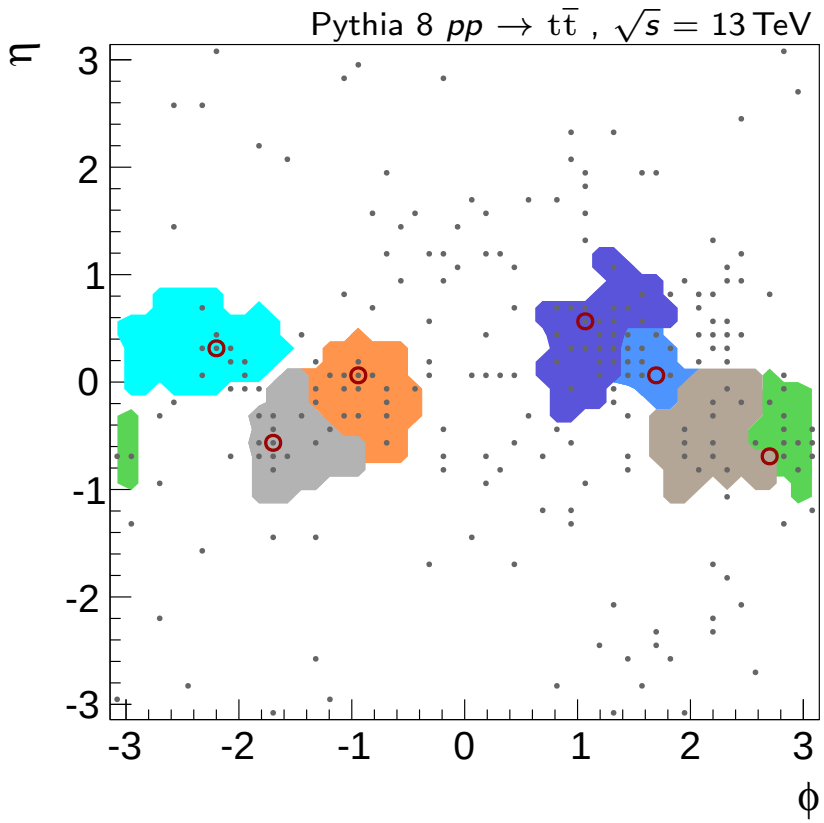
HOTVR clustering



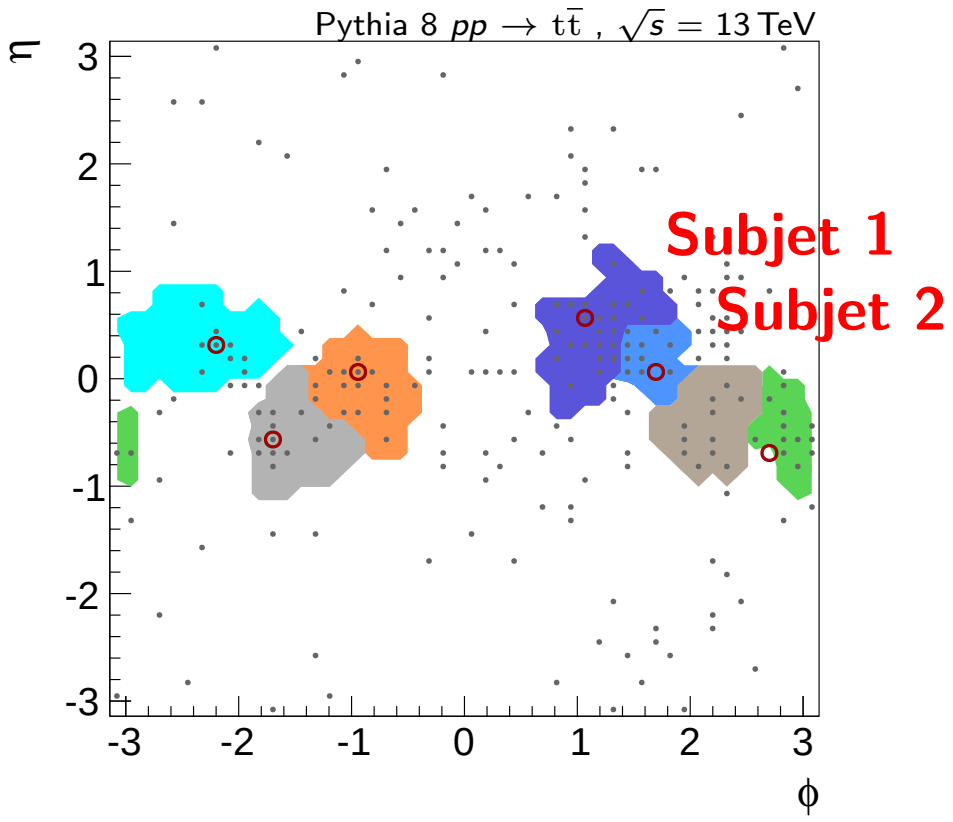
- $m_{ij} < \mu$

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



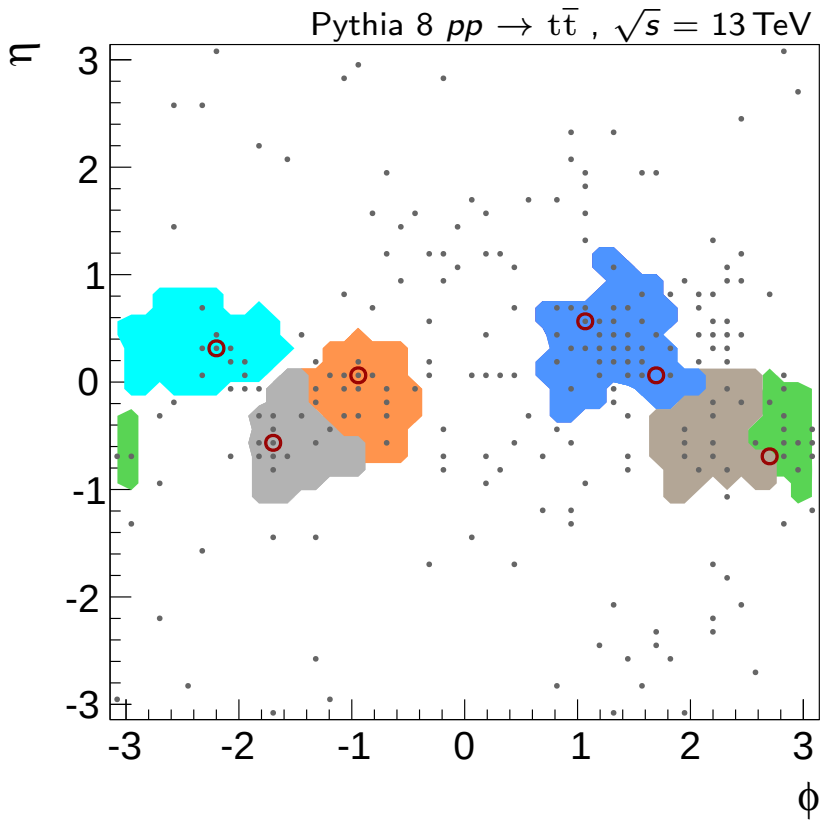
HOTVR clustering



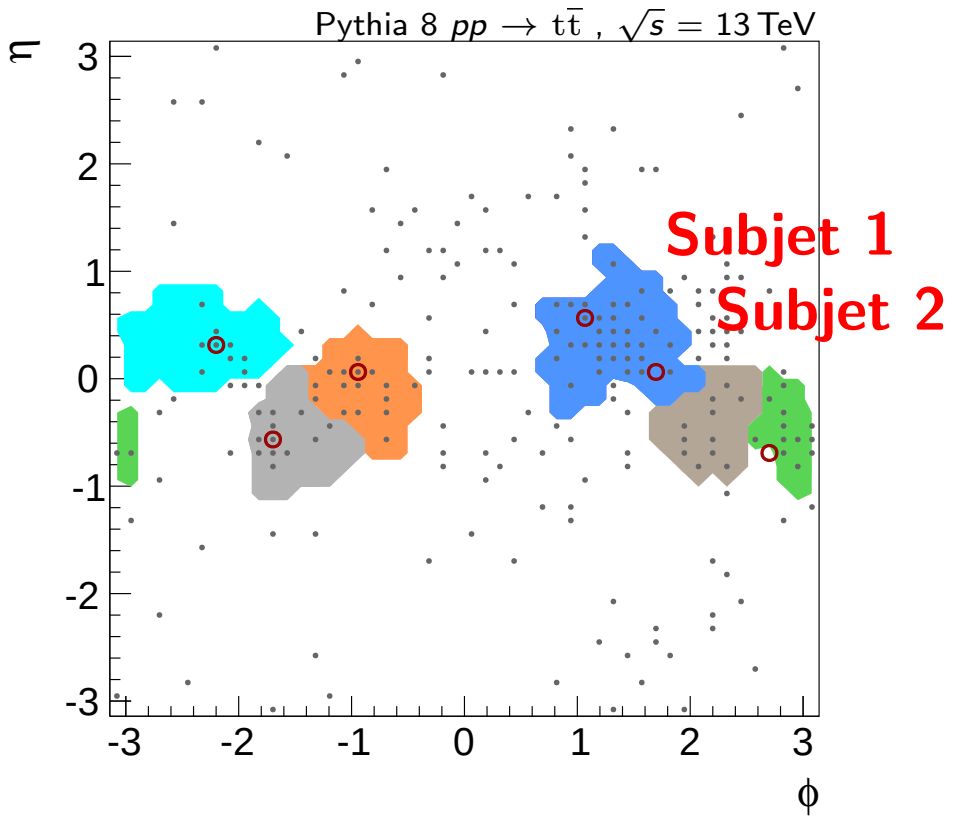
- Massjump found: save subjets

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



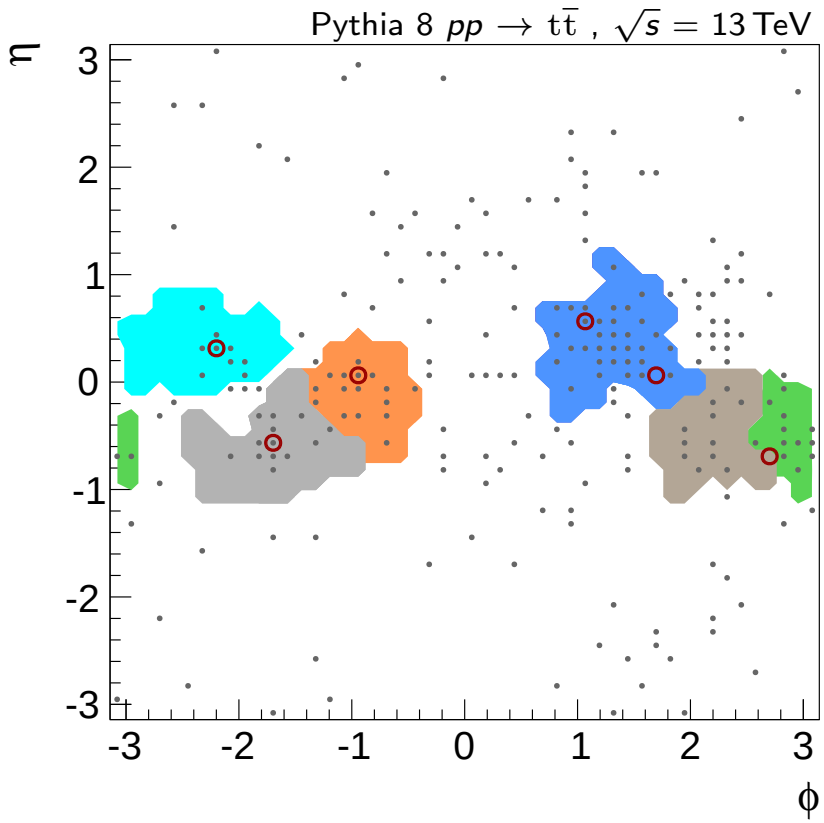
HOTVR clustering



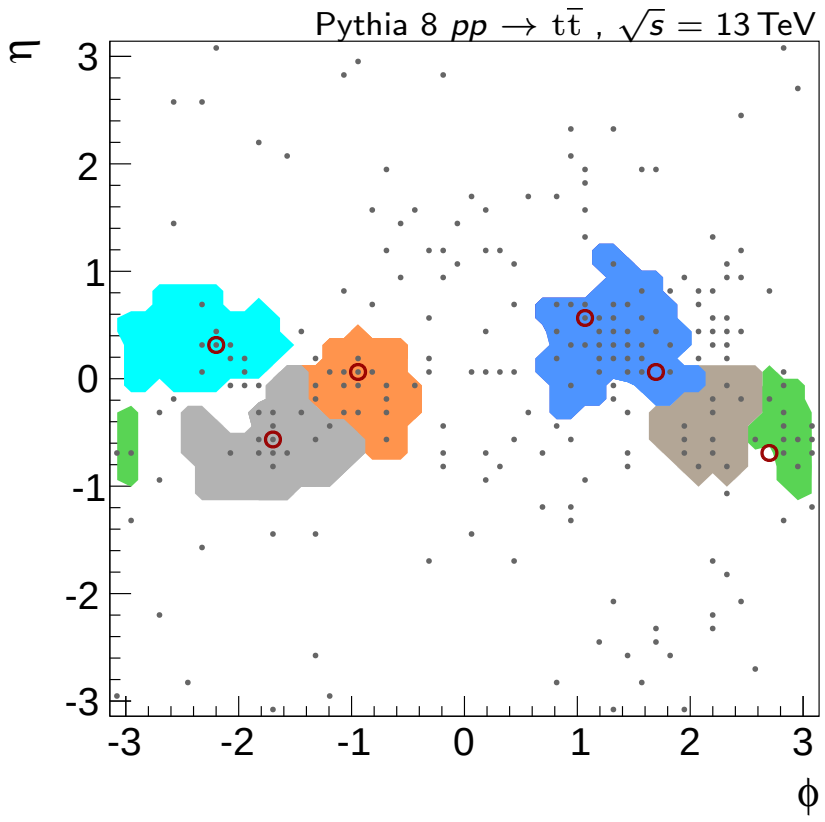
- Massjump found: save subjets

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

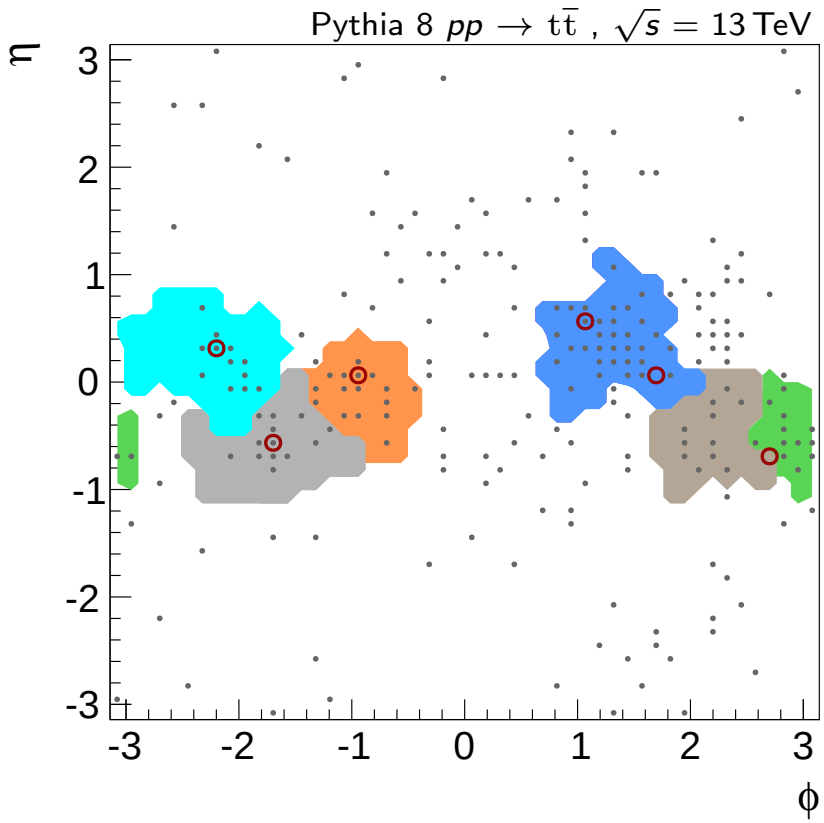


HOTVR clustering

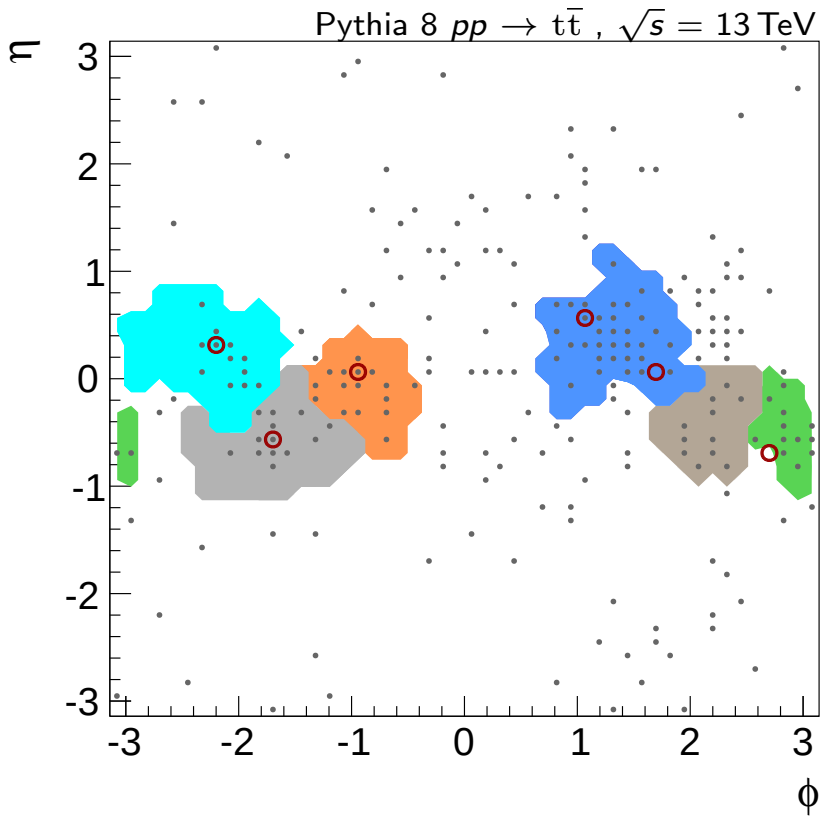


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

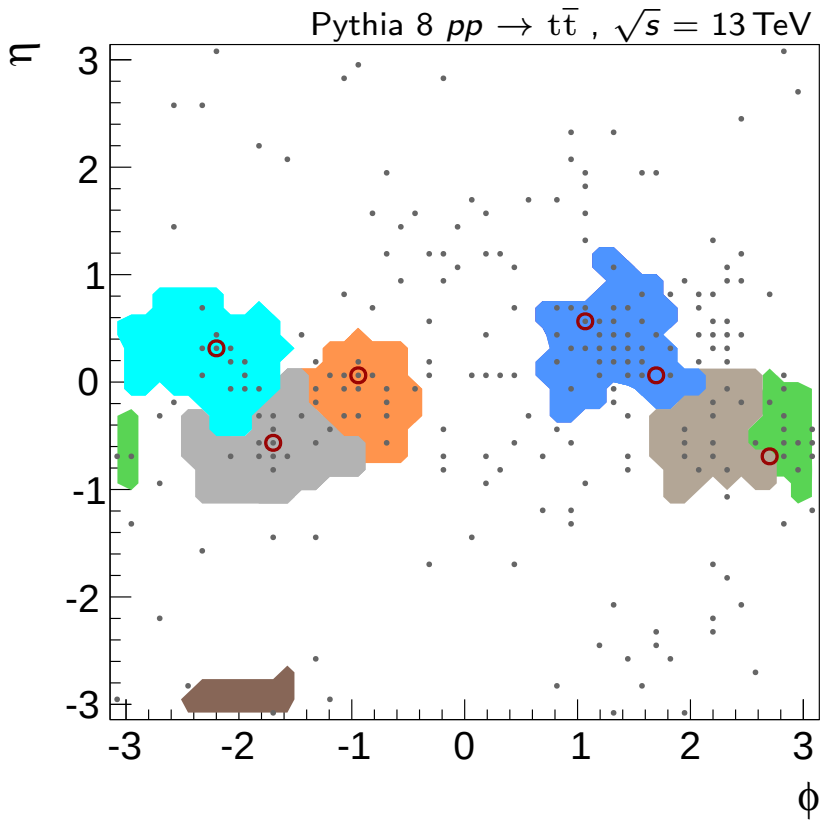


HOTVR clustering

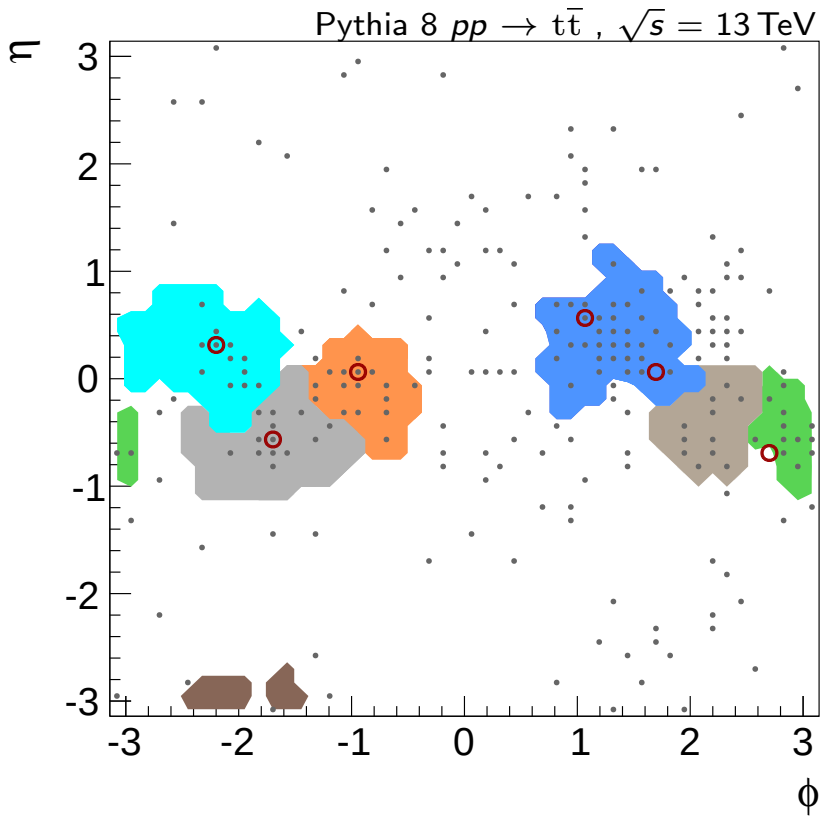


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

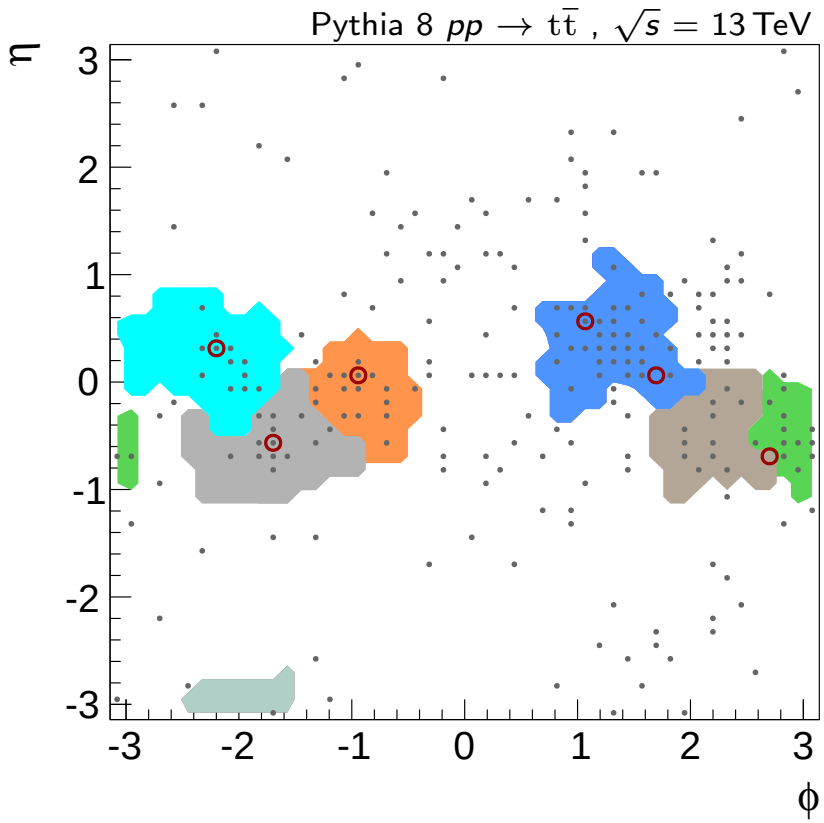


HOTVR clustering

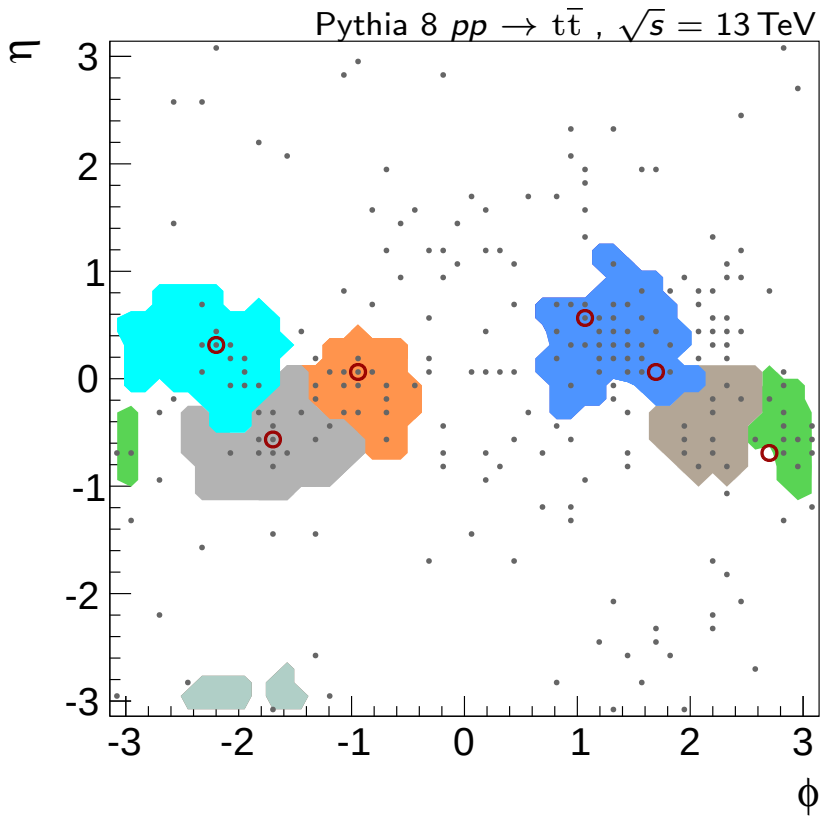


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

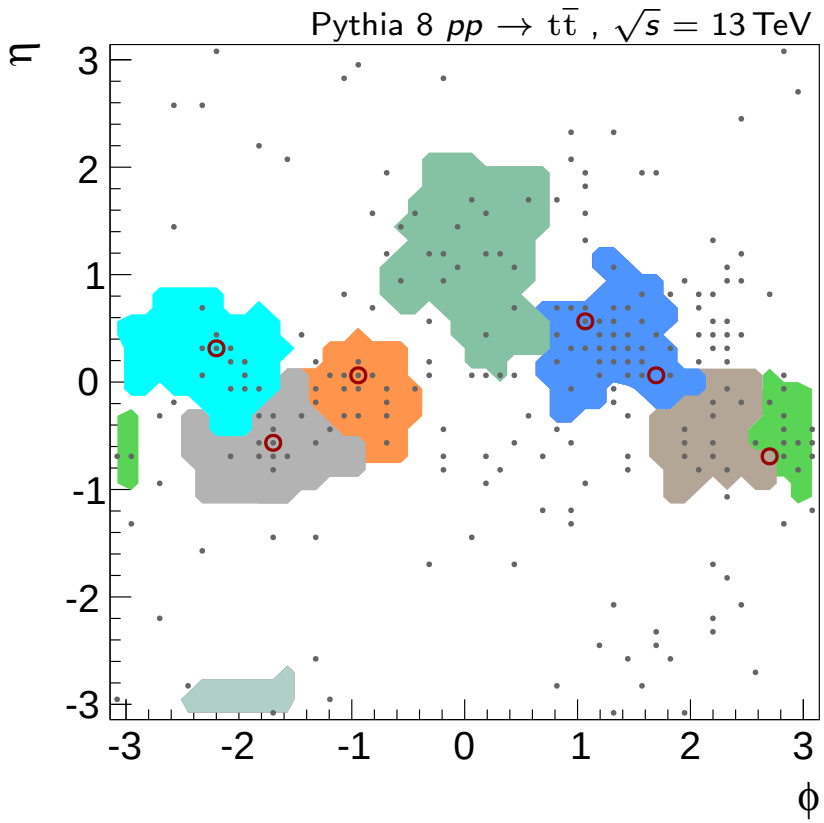


HOTVR clustering

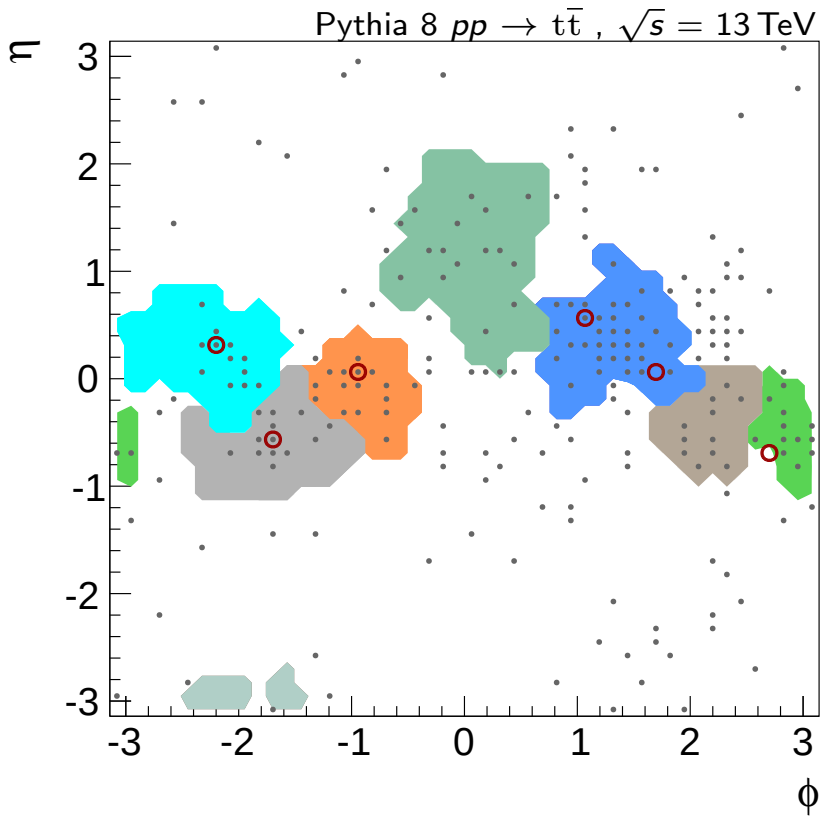


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

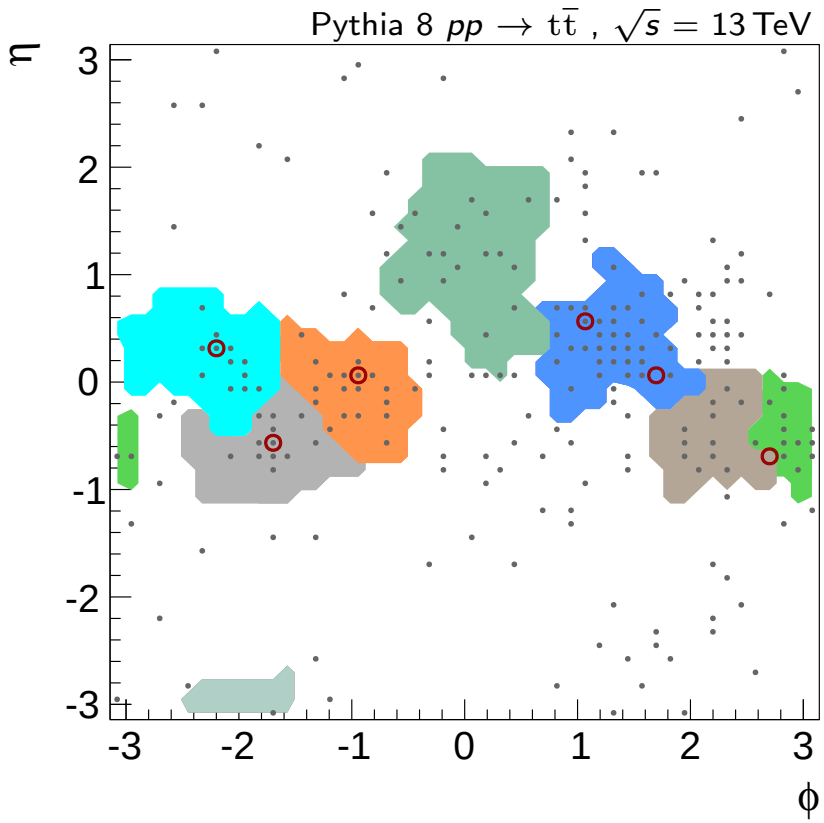


HOTVR clustering

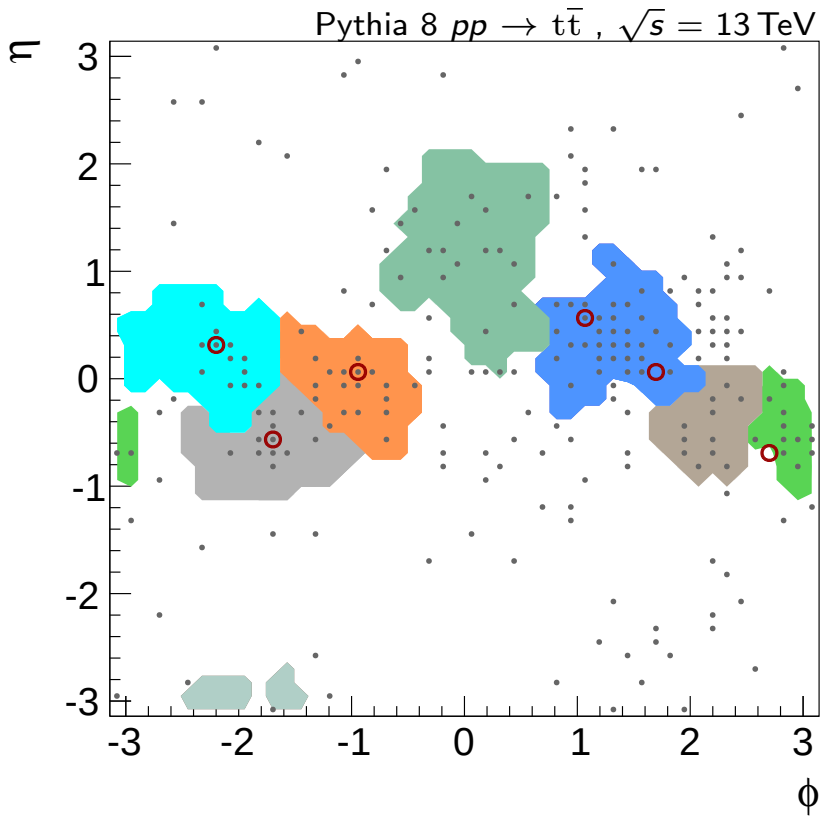


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

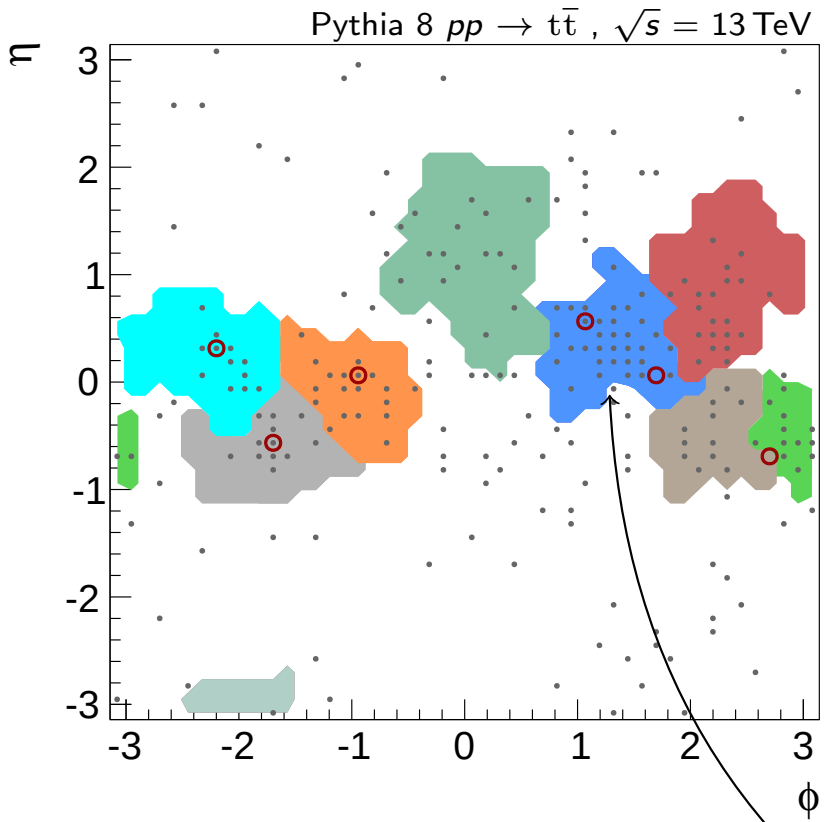


HOTVR clustering

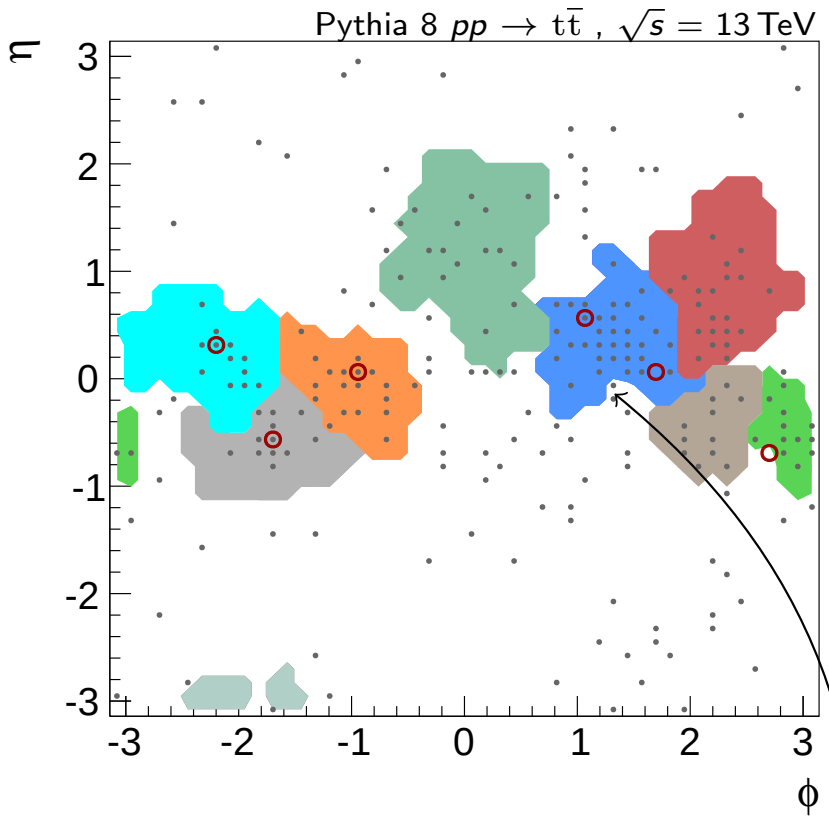


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



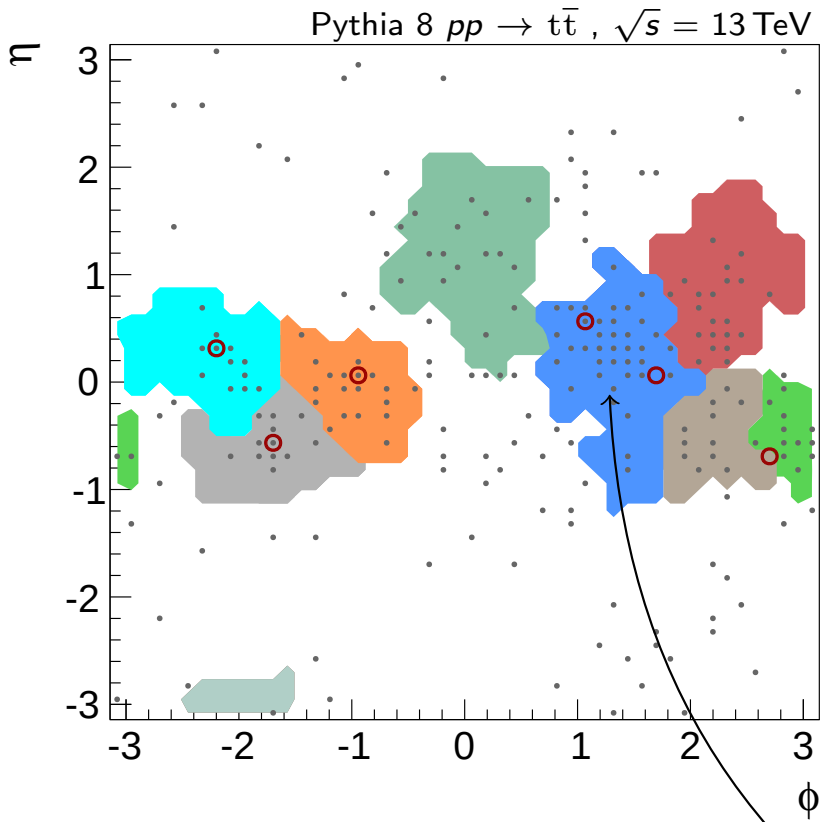
HOTVR clustering



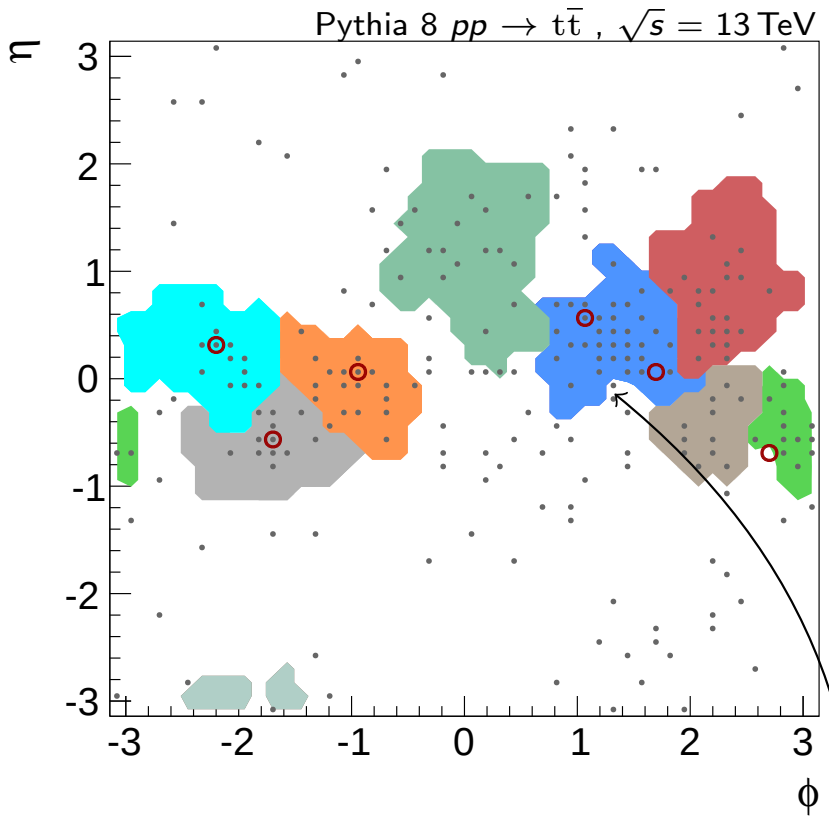
Do not cluster soft particles

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



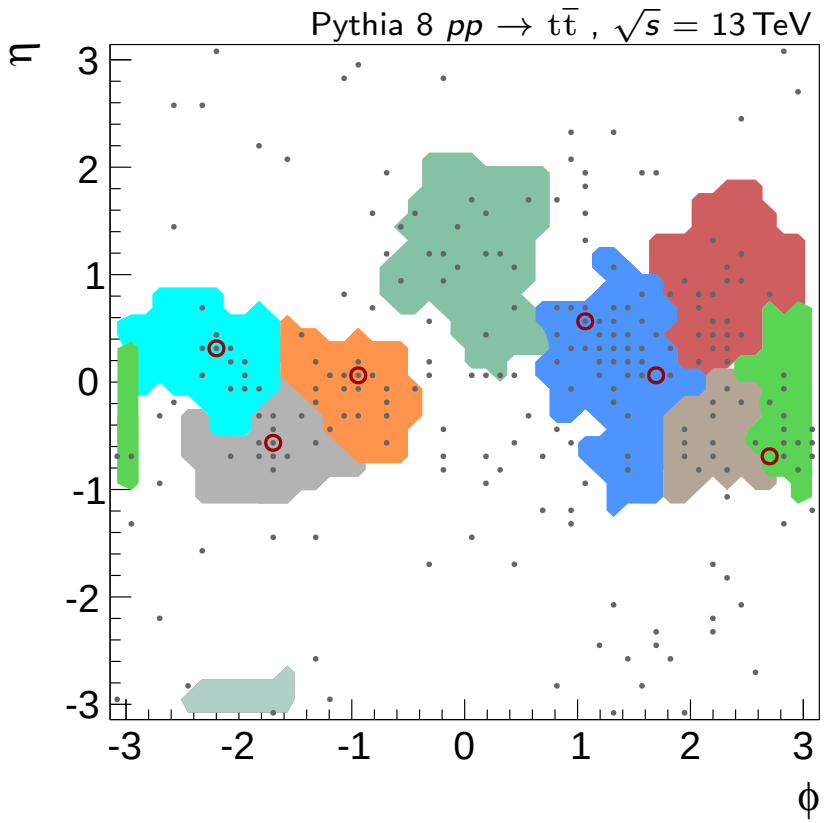
HOTVR clustering



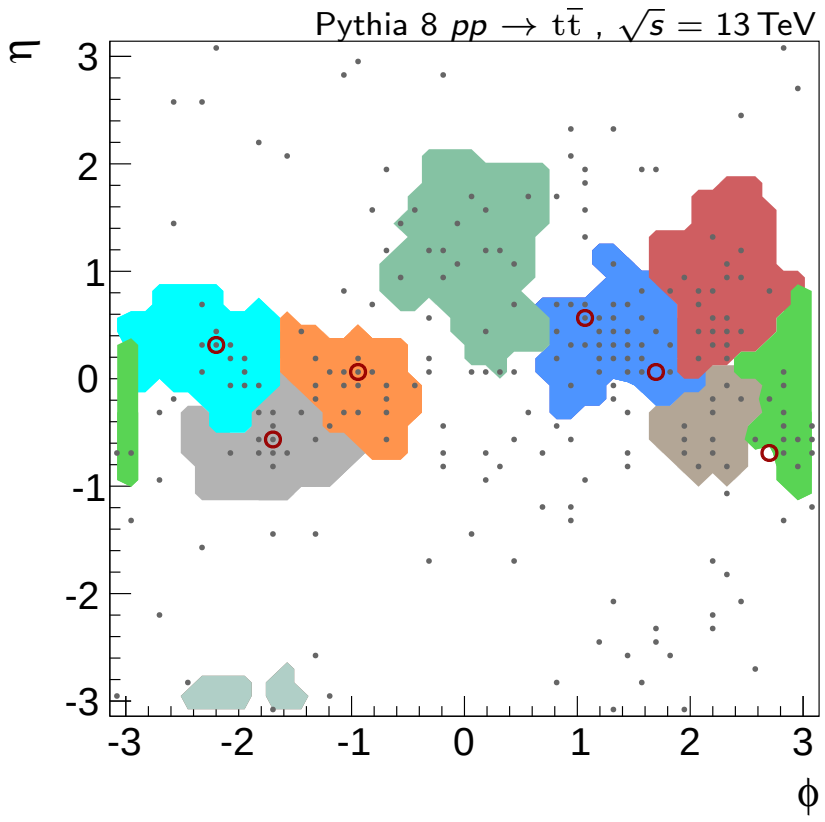
Do not cluster soft particles

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

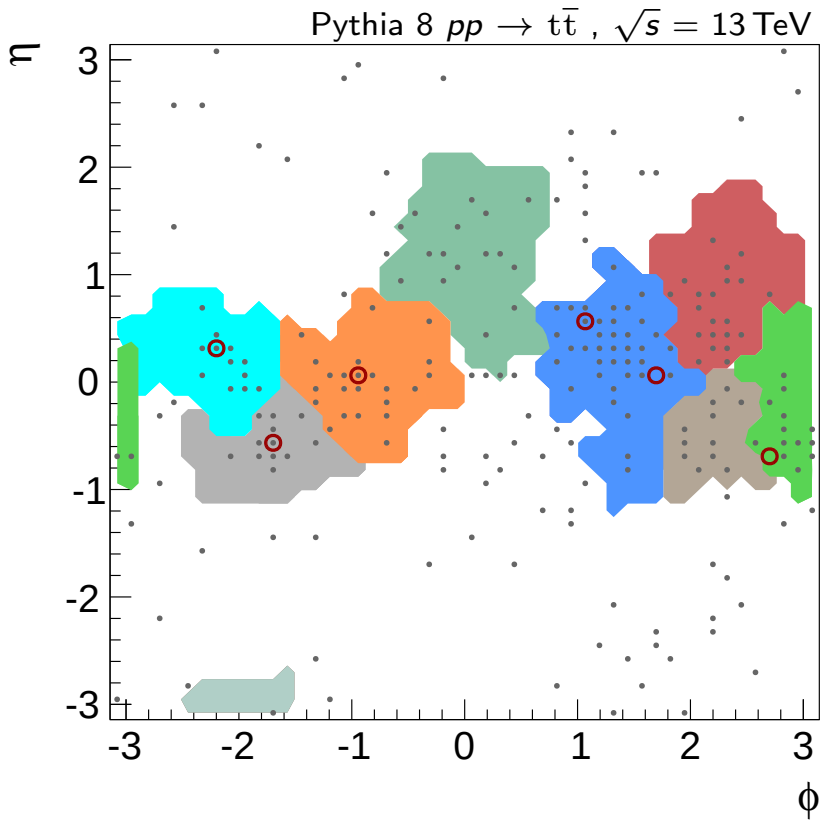


HOTVR clustering

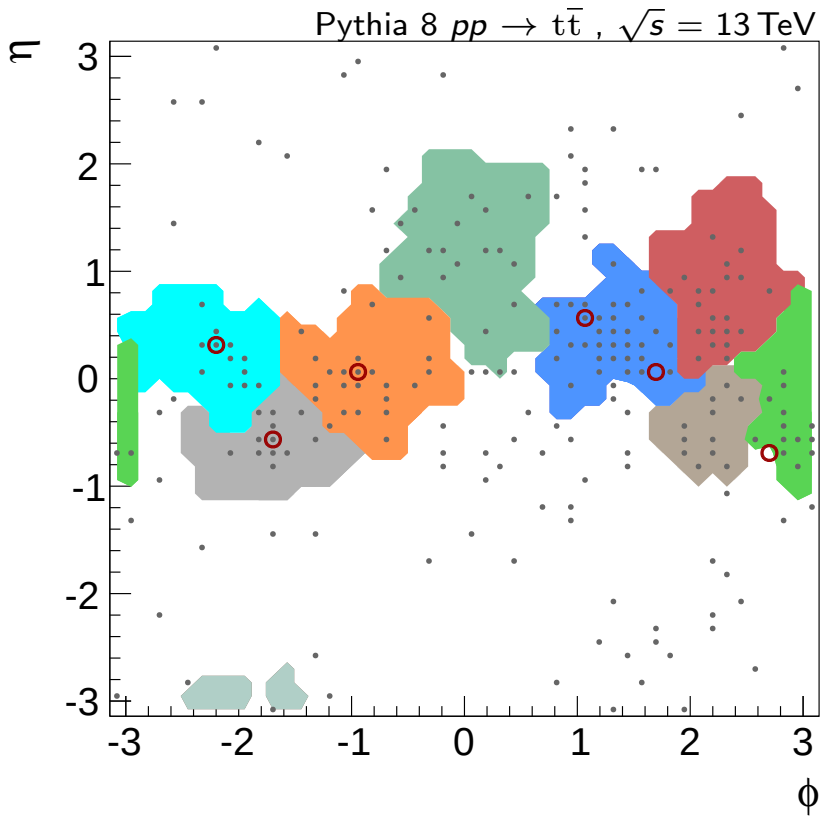


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

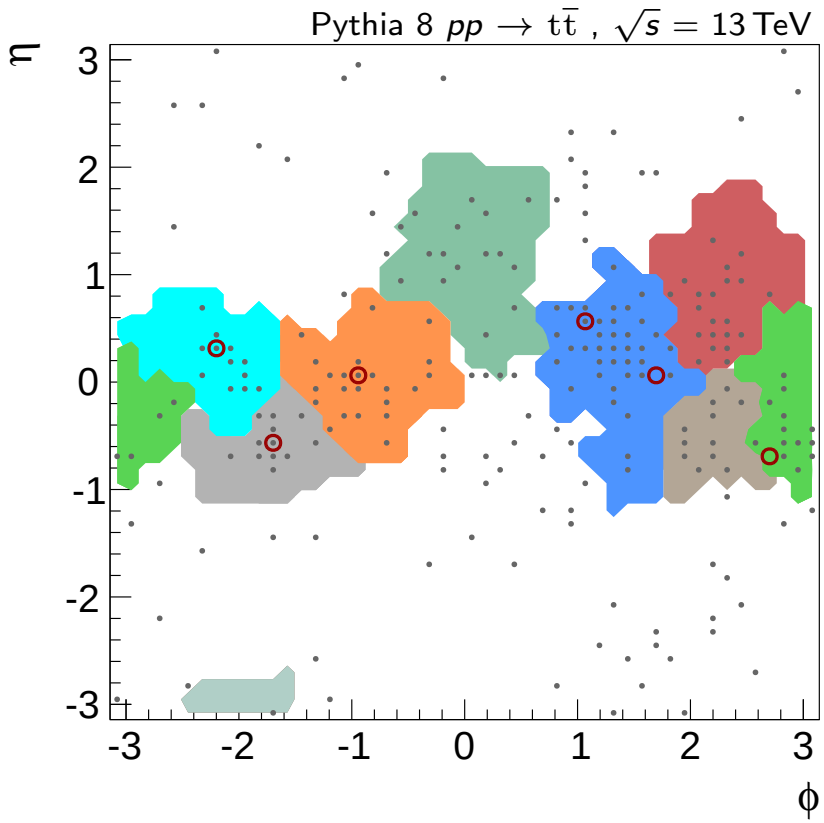


HOTVR clustering

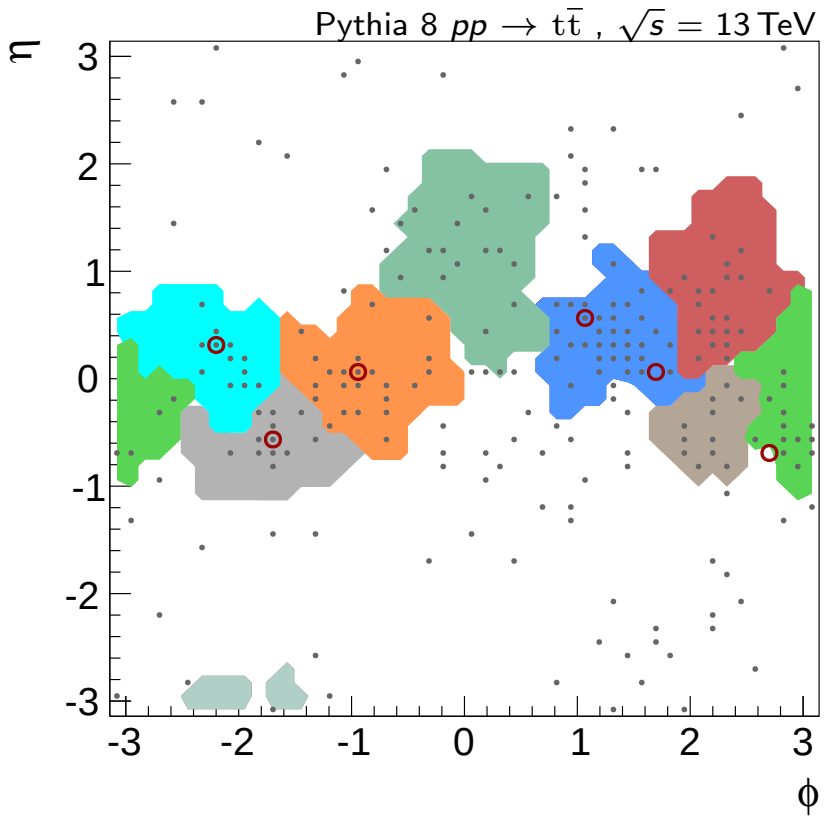


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

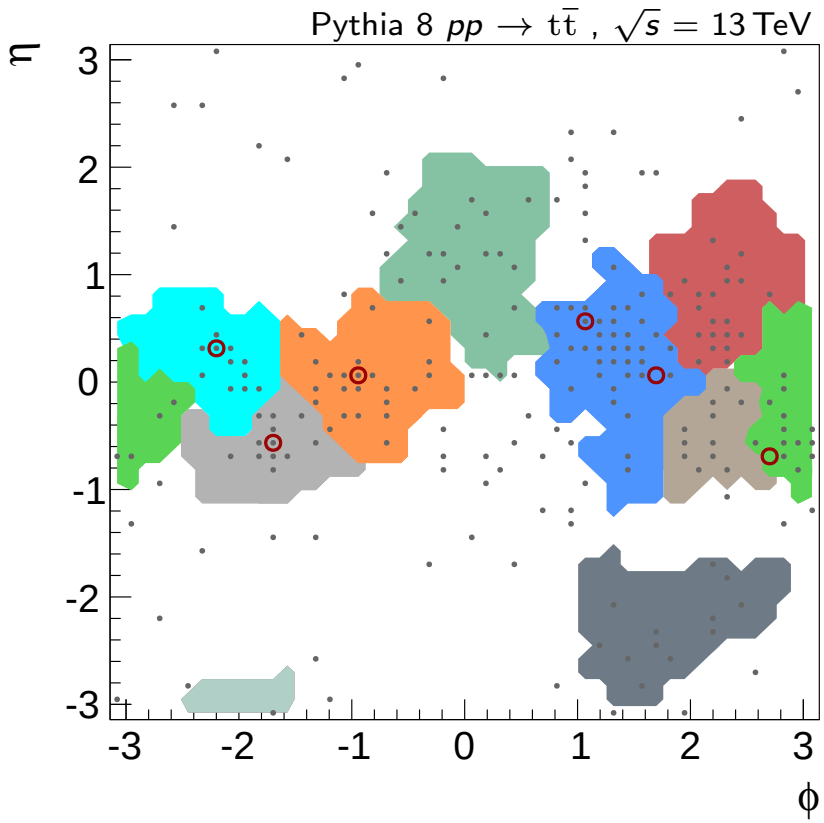


HOTVR clustering

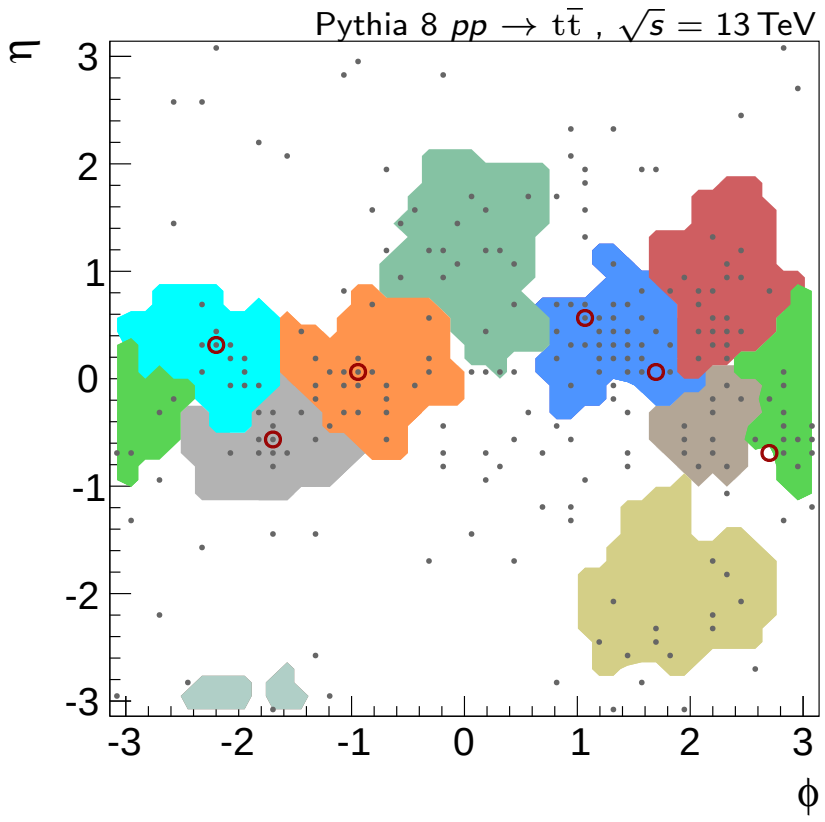


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

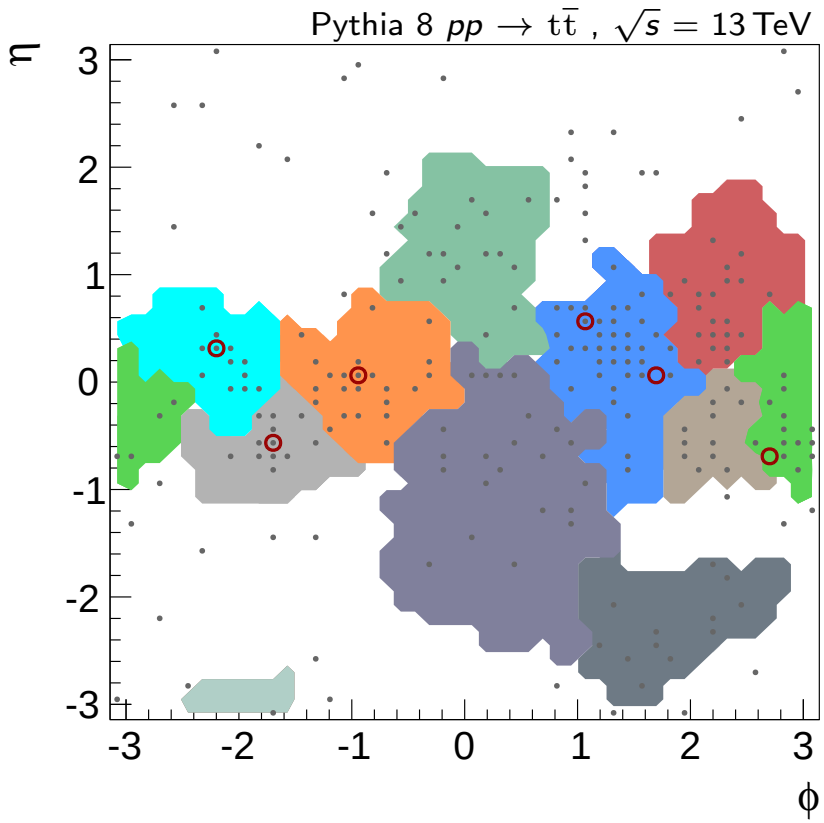


HOTVR clustering

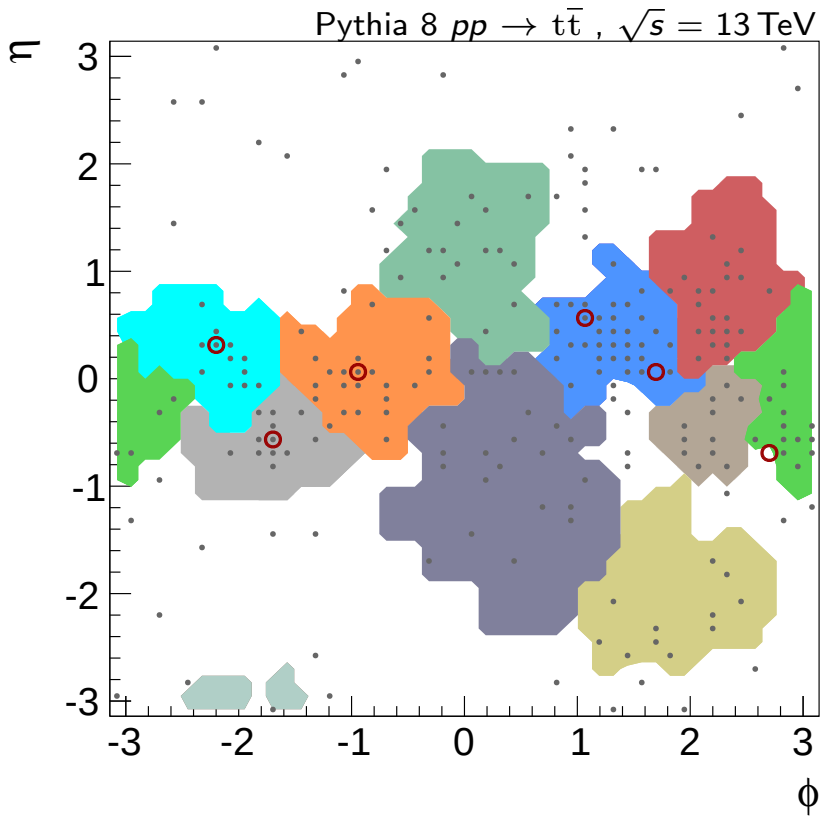


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering

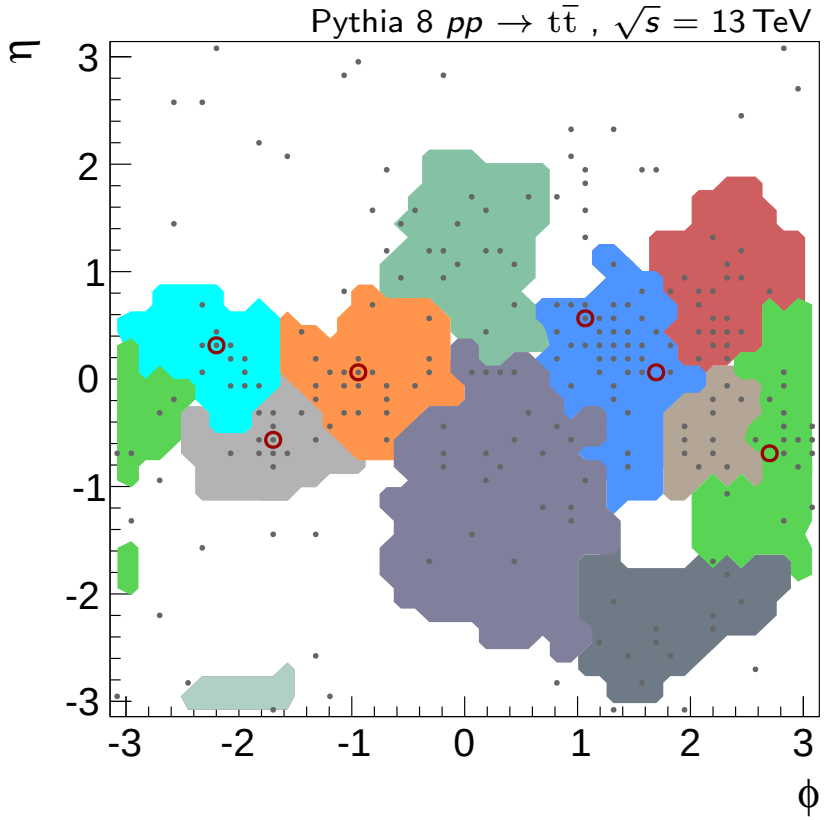


HOTVR clustering

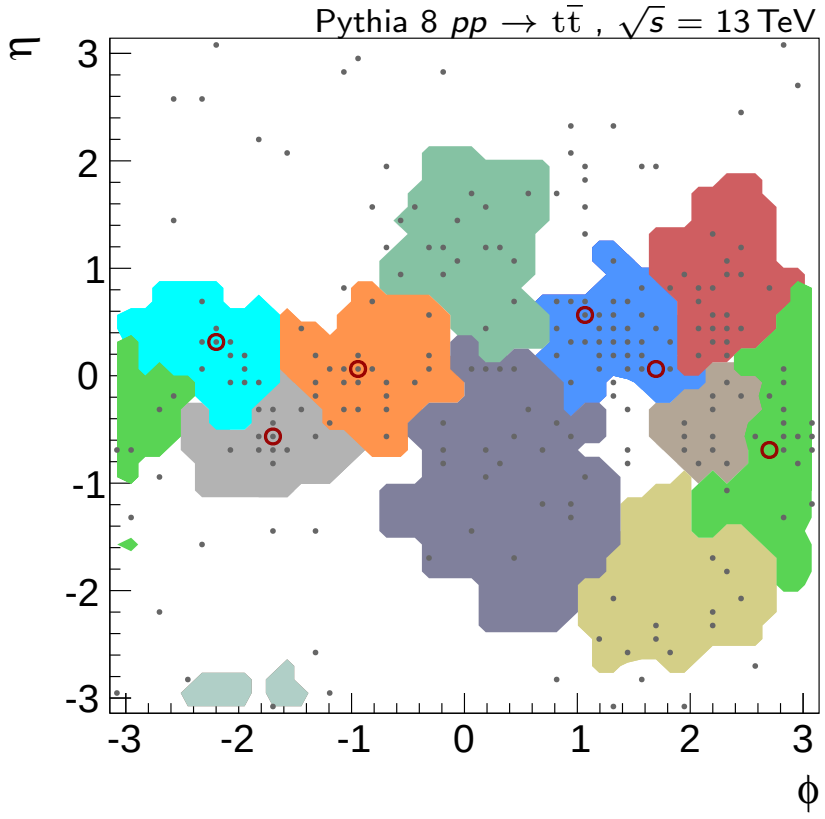


Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



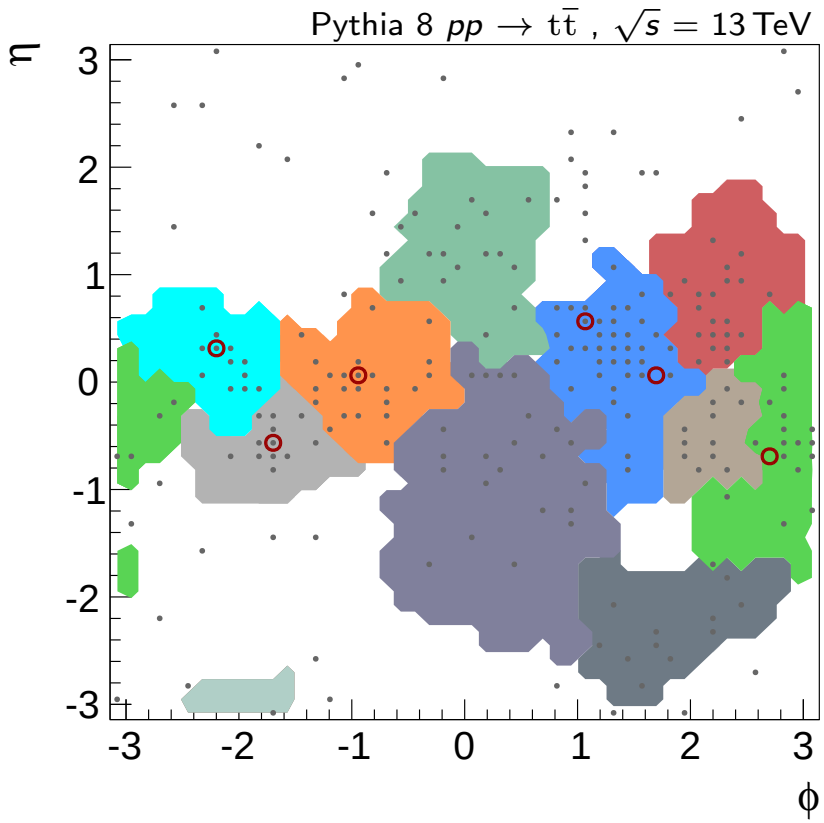
HOTVR clustering



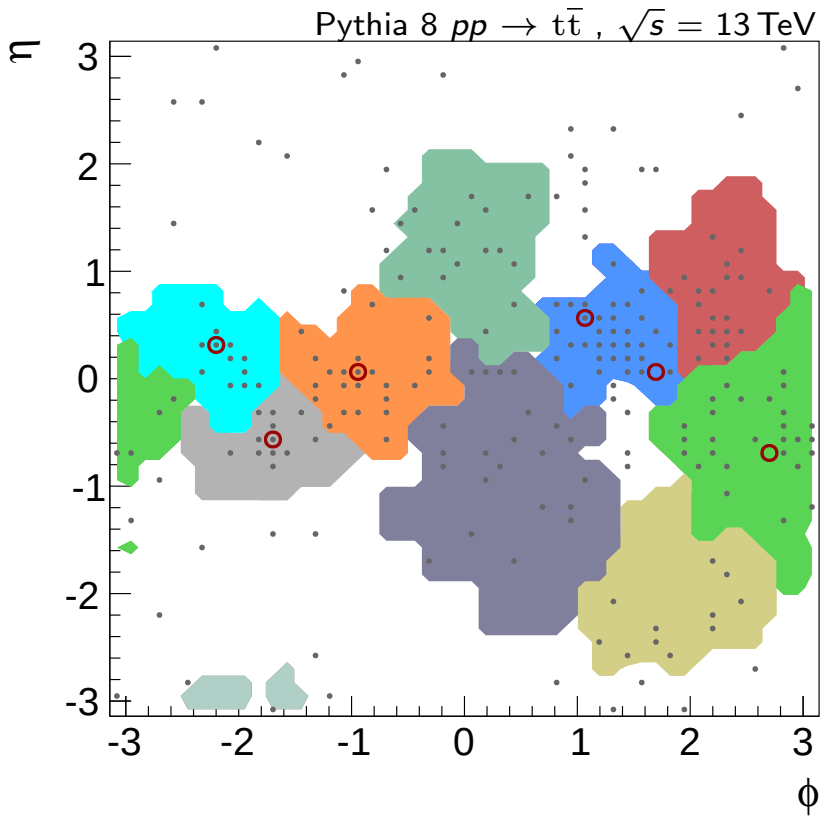
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



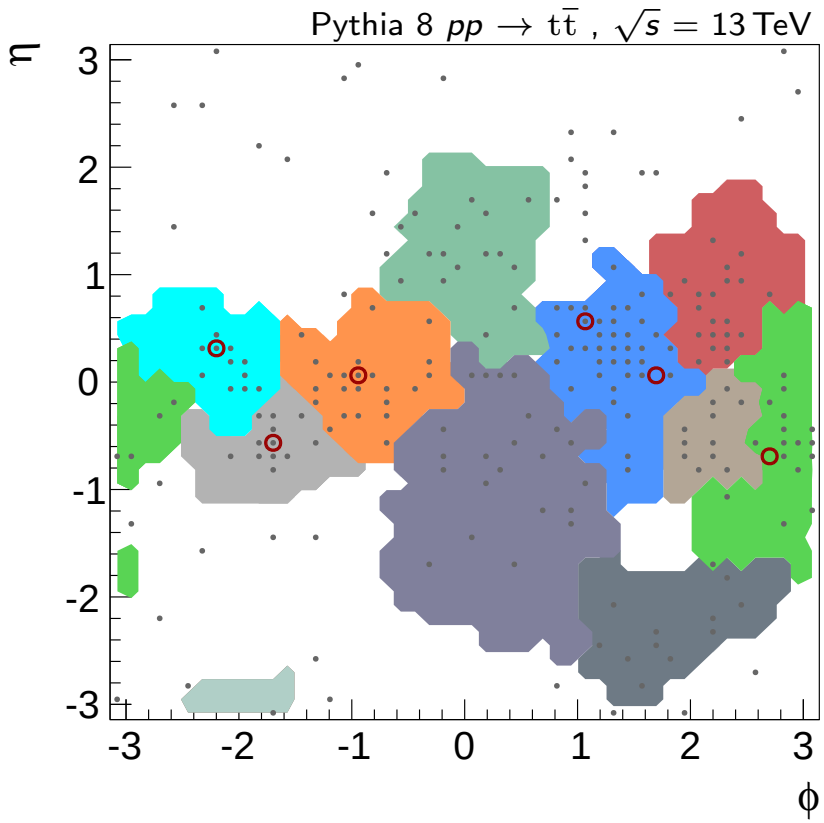
HOTVR clustering



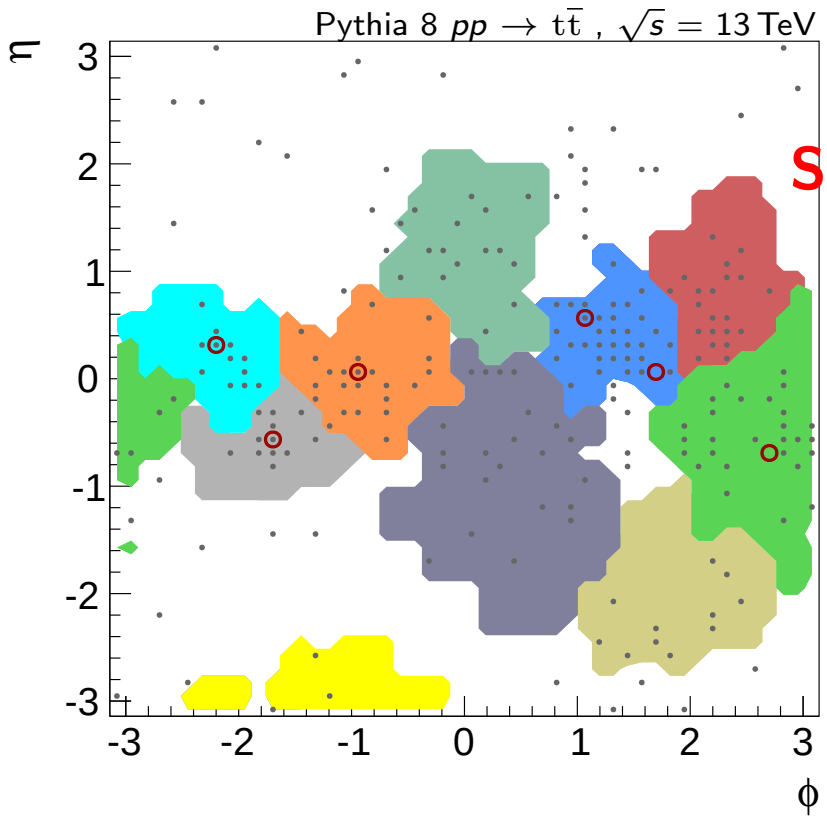
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



HOTVR clustering

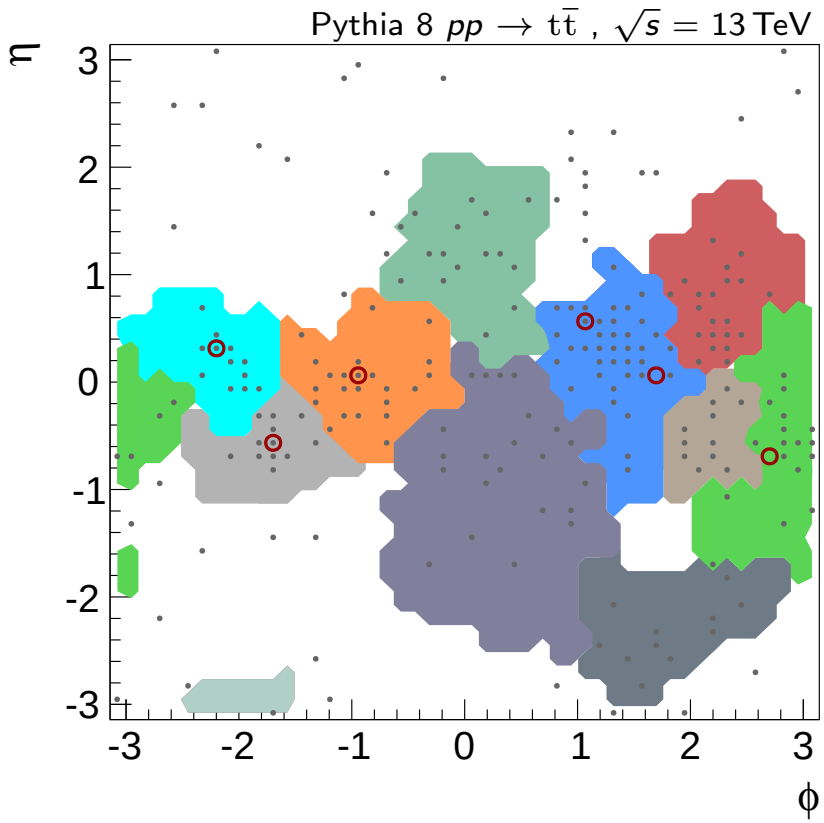


Subjet 3?

- Stop clustering
- Reject soft subjets

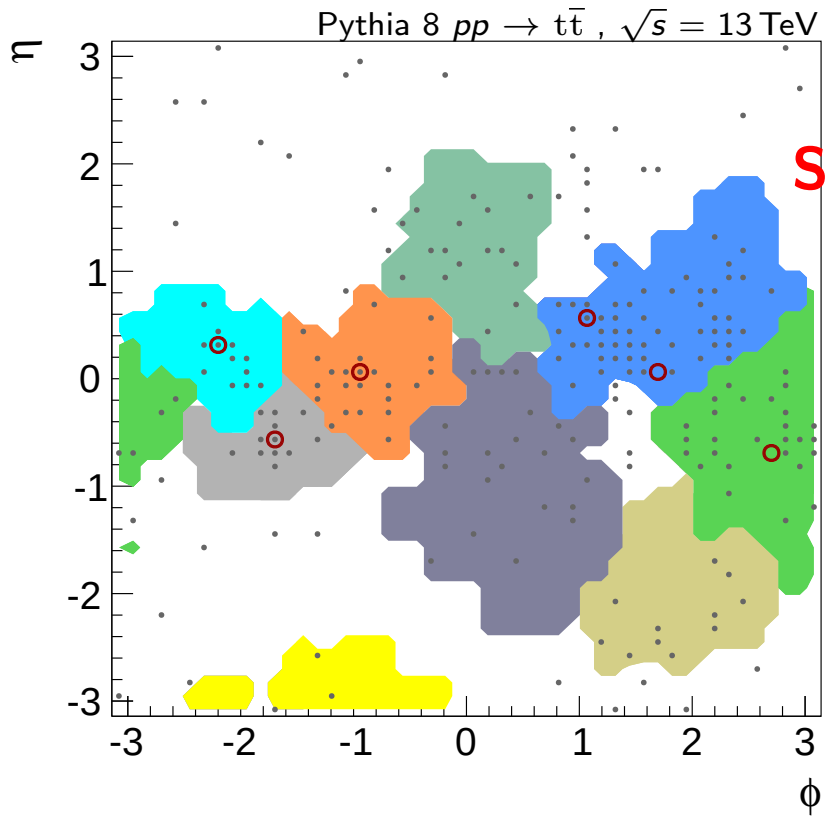
Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



- Stop clustering

HOTVR clustering

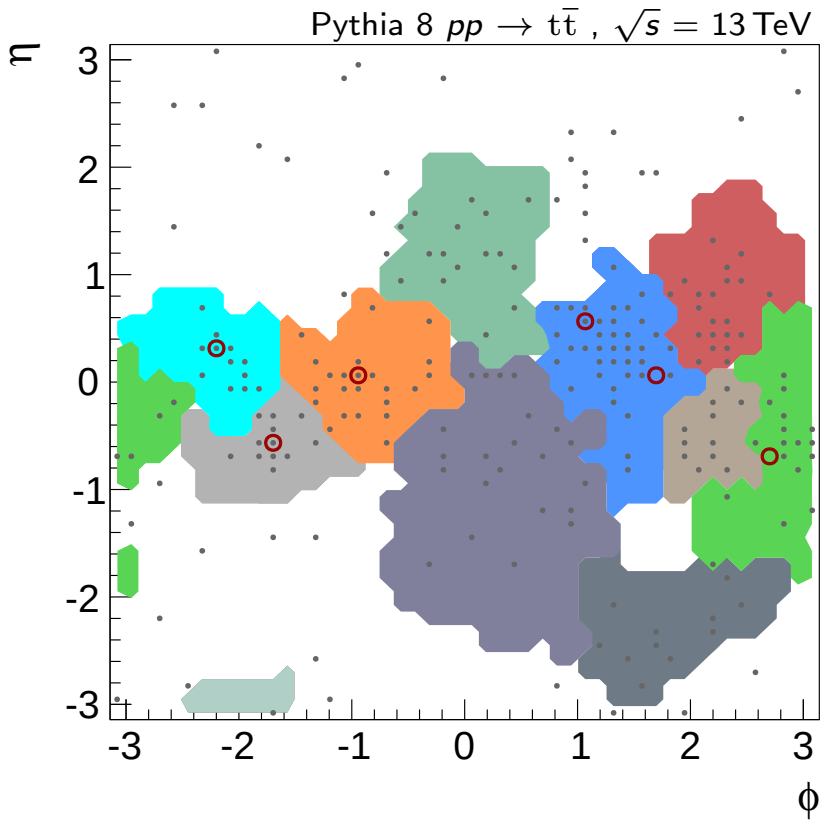


Subjet 3?

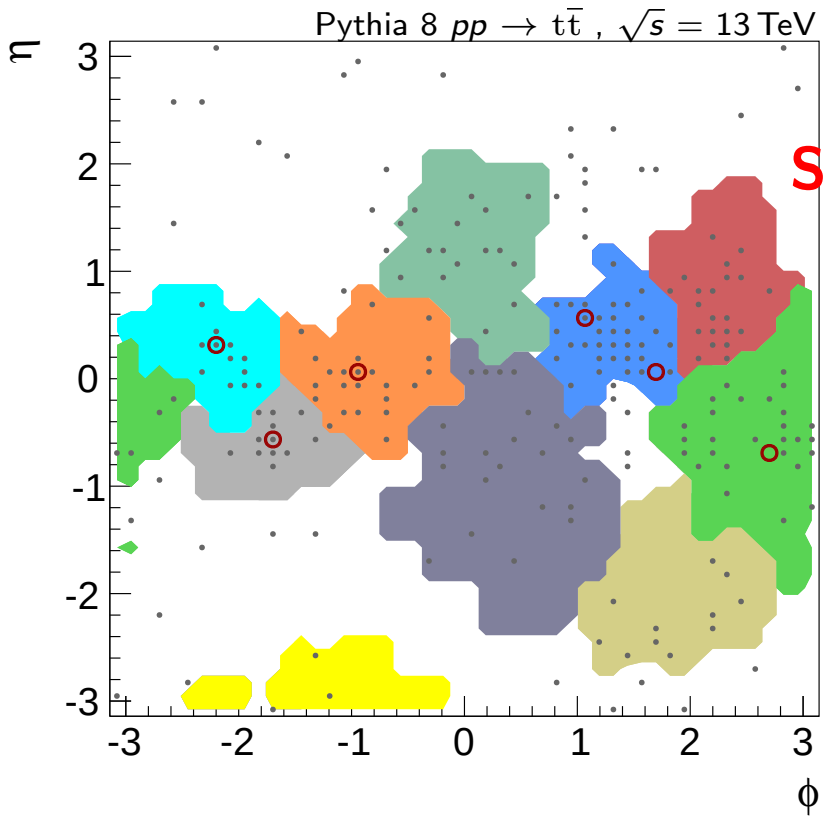
- Reject soft subjets

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



HOTVR clustering

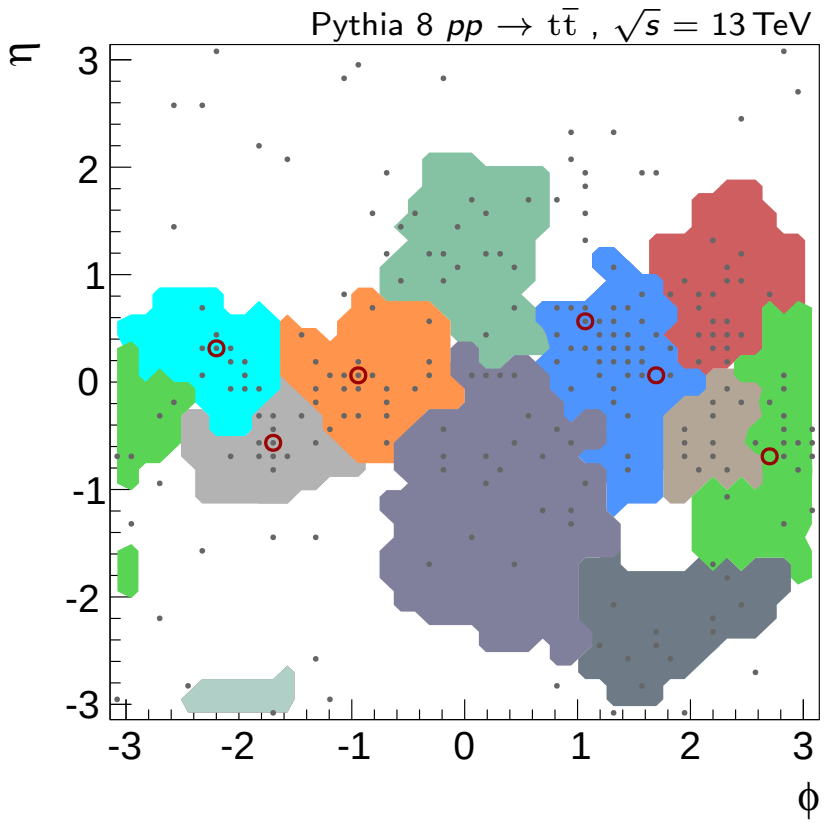


Subjet 3?

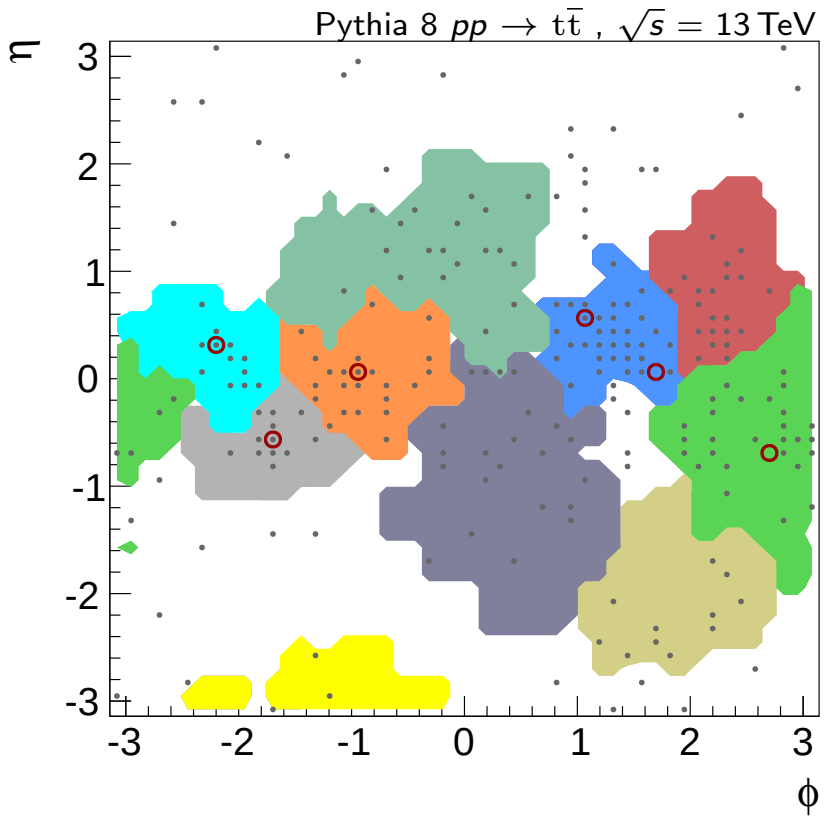
- Stop clustering
- Reject soft subjets

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



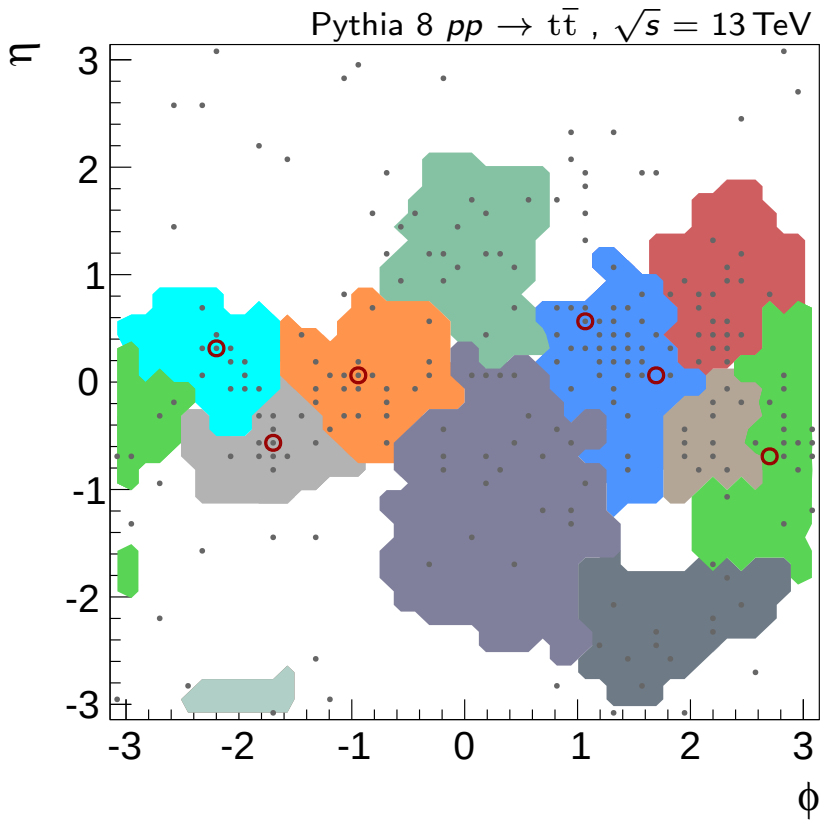
HOTVR clustering



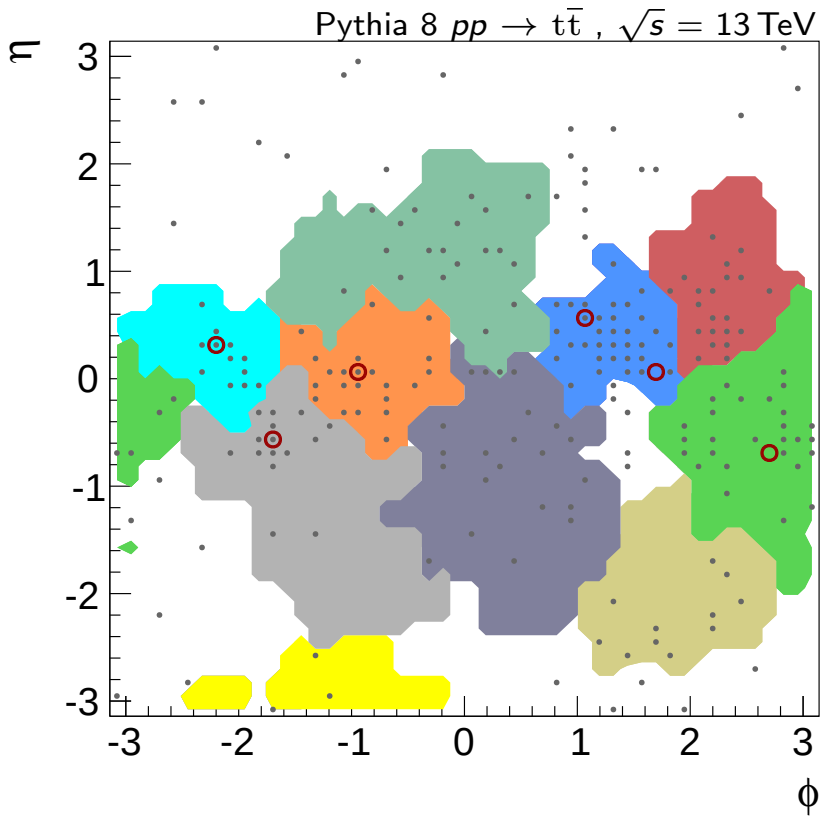
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



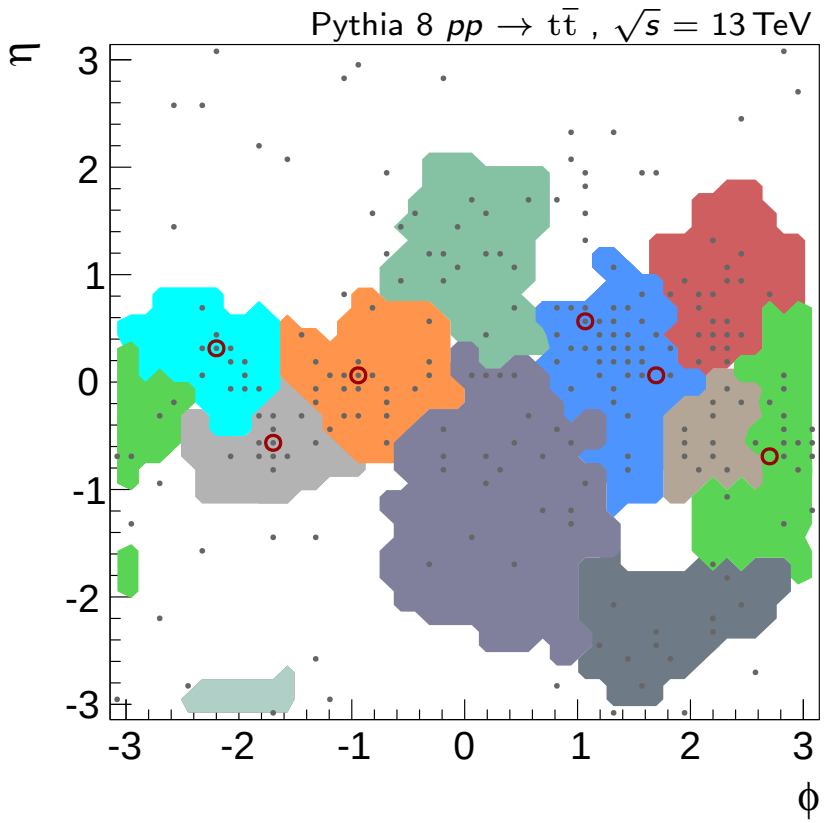
HOTVR clustering



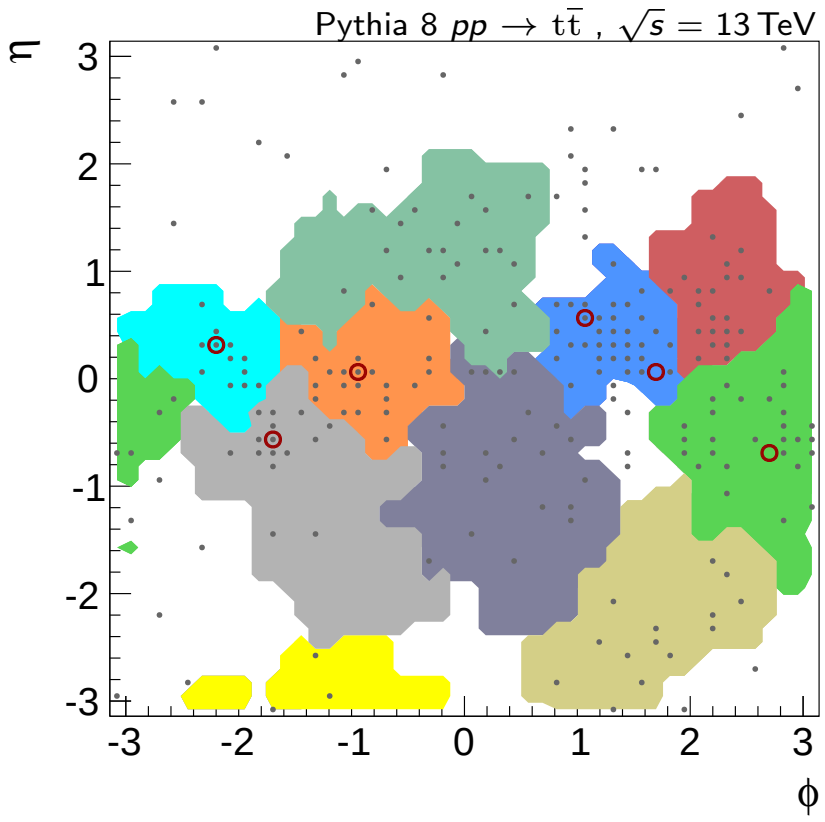
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



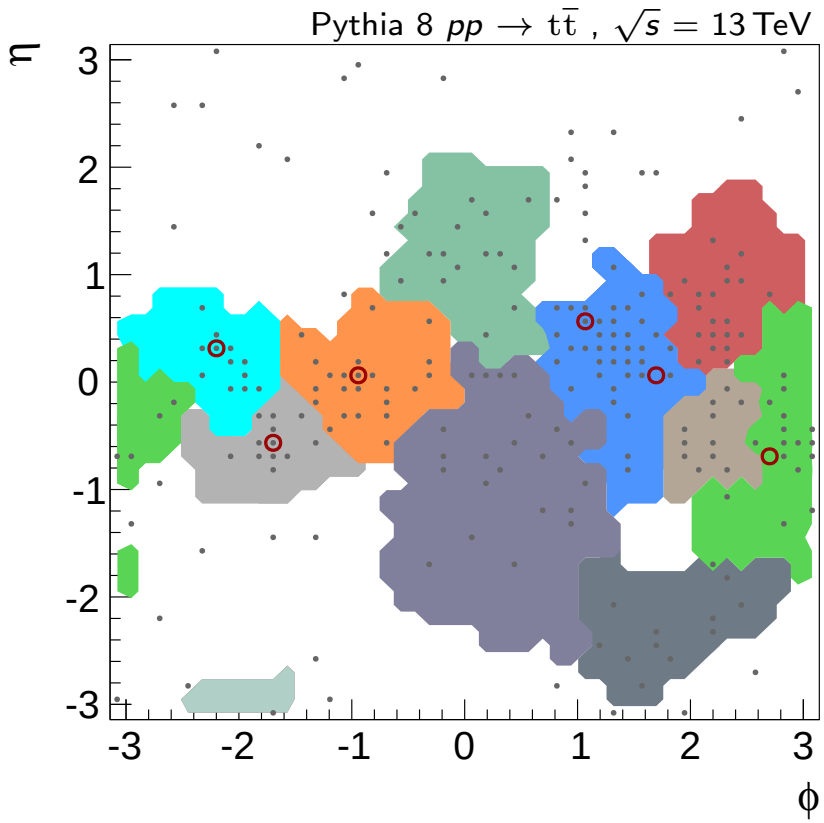
HOTVR clustering



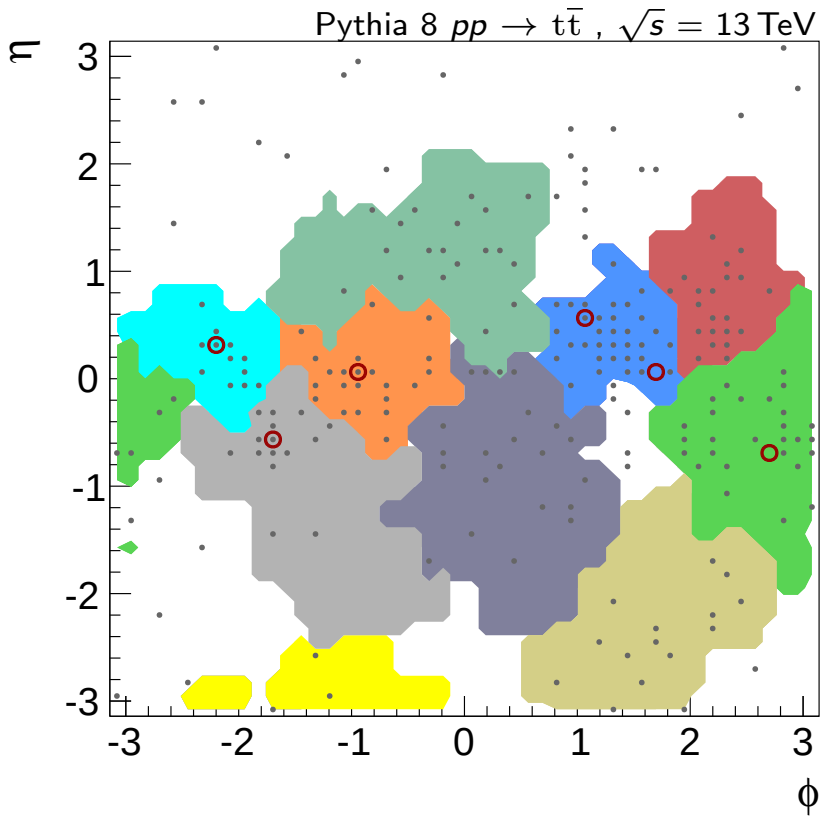
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



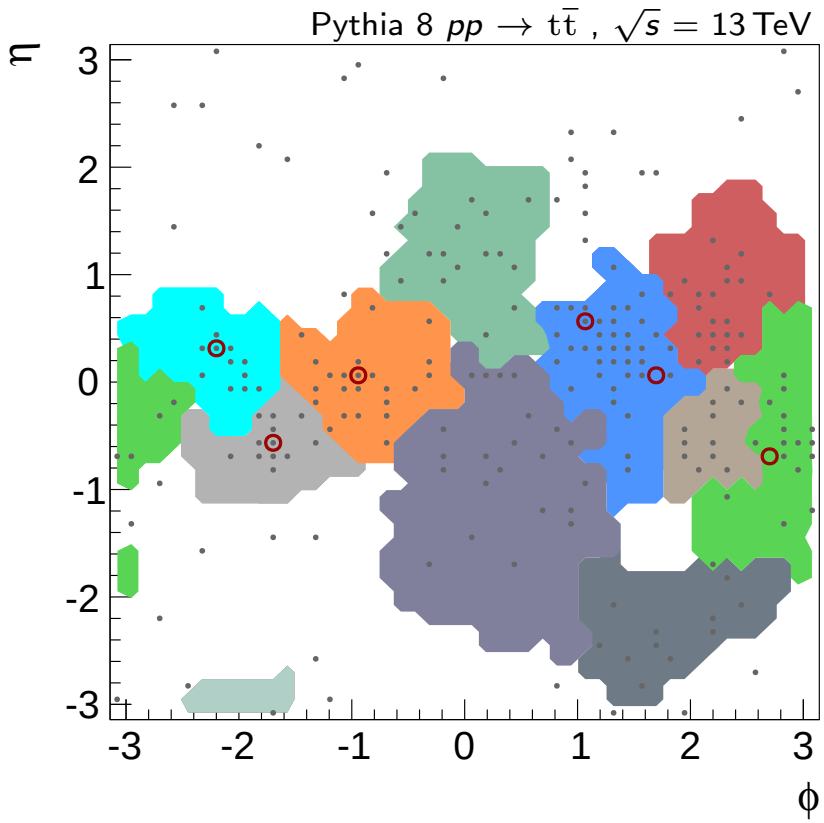
HOTVR clustering



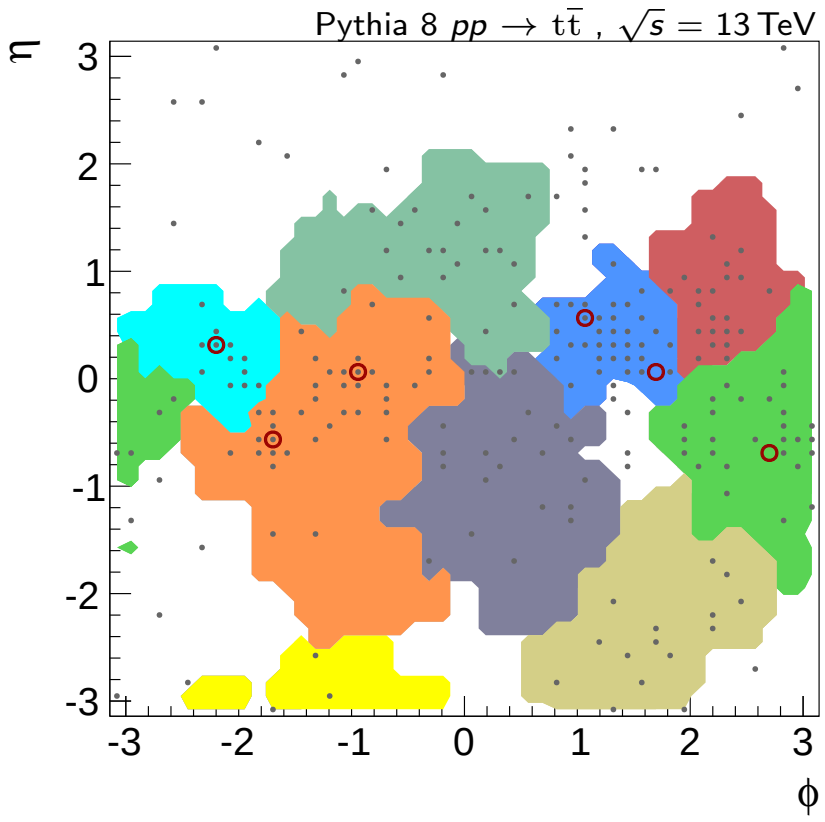
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



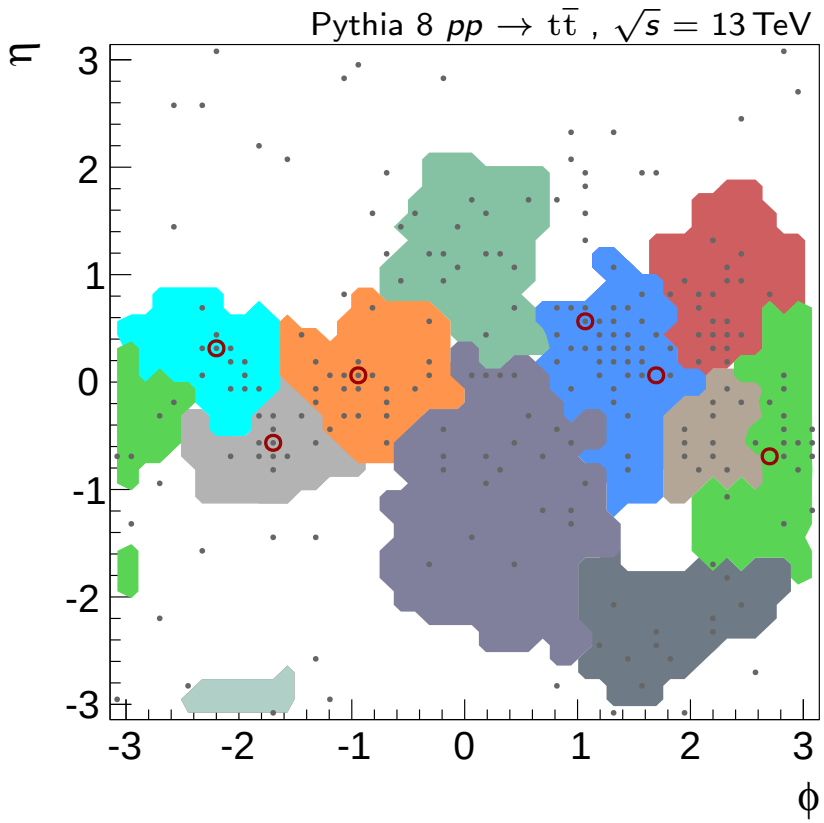
HOTVR clustering



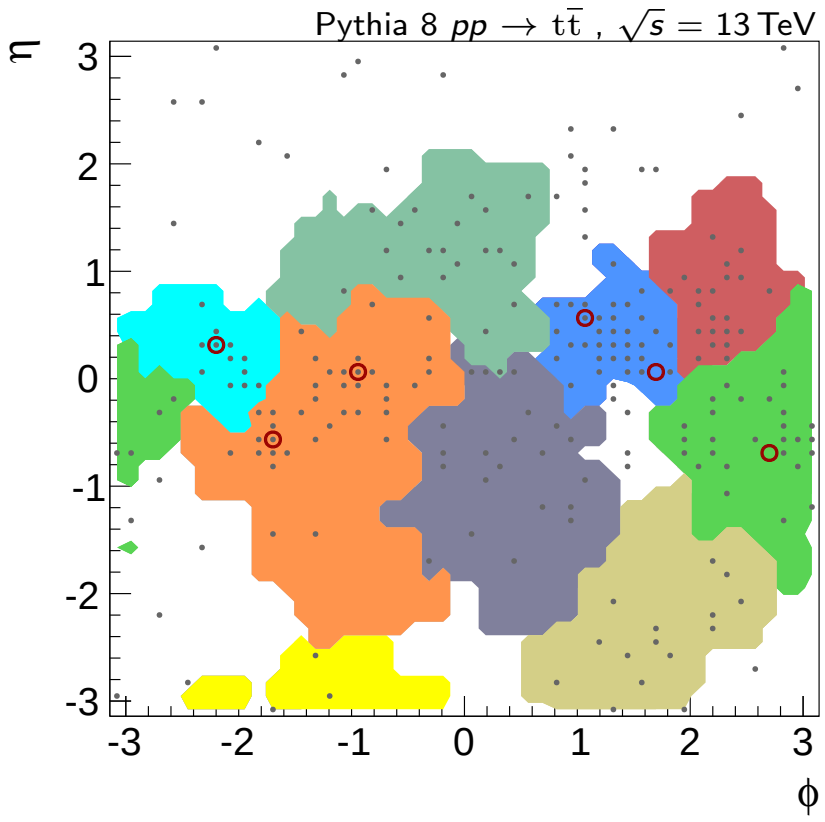
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



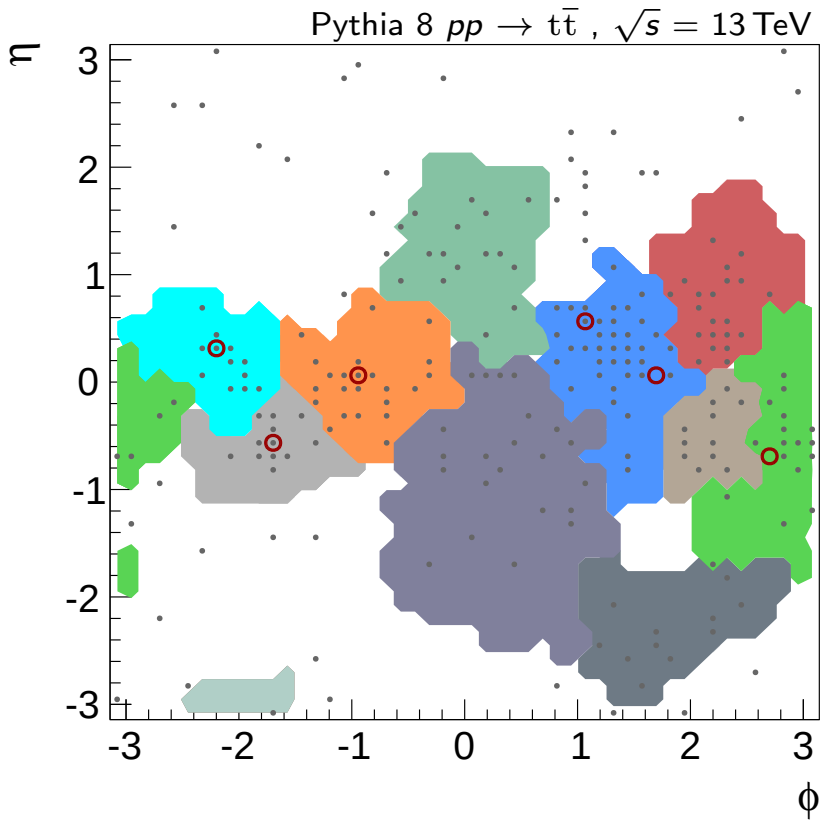
HOTVR clustering



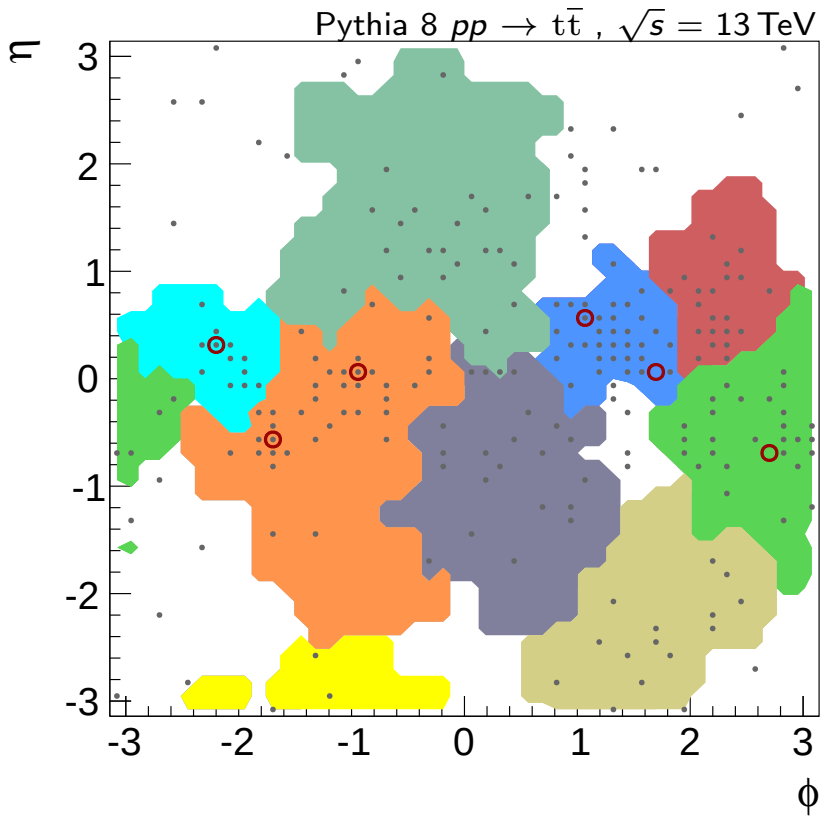
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



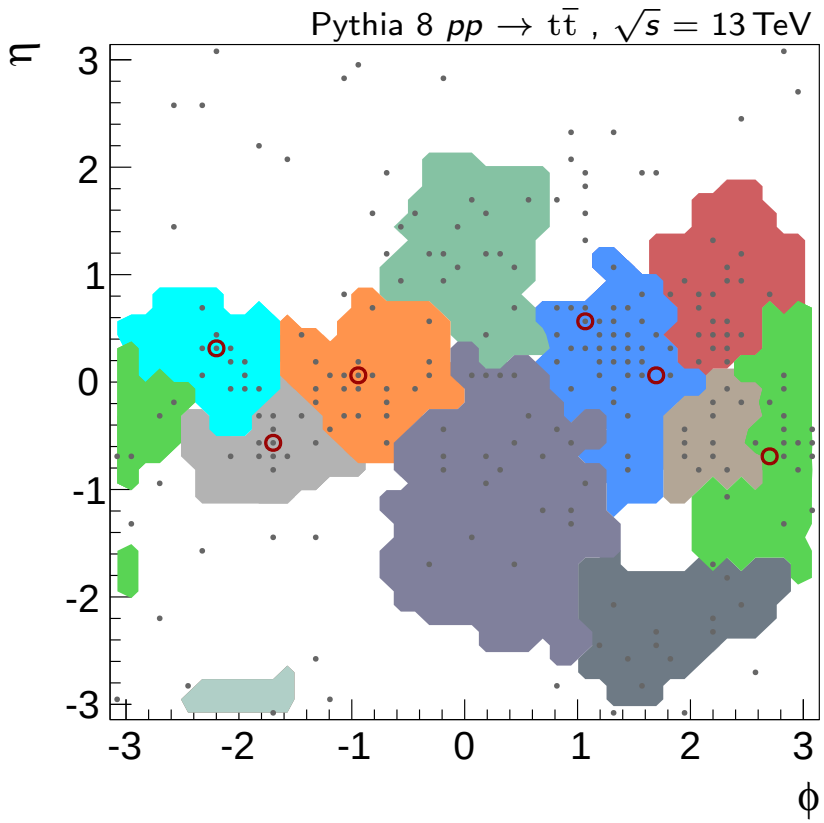
HOTVR clustering



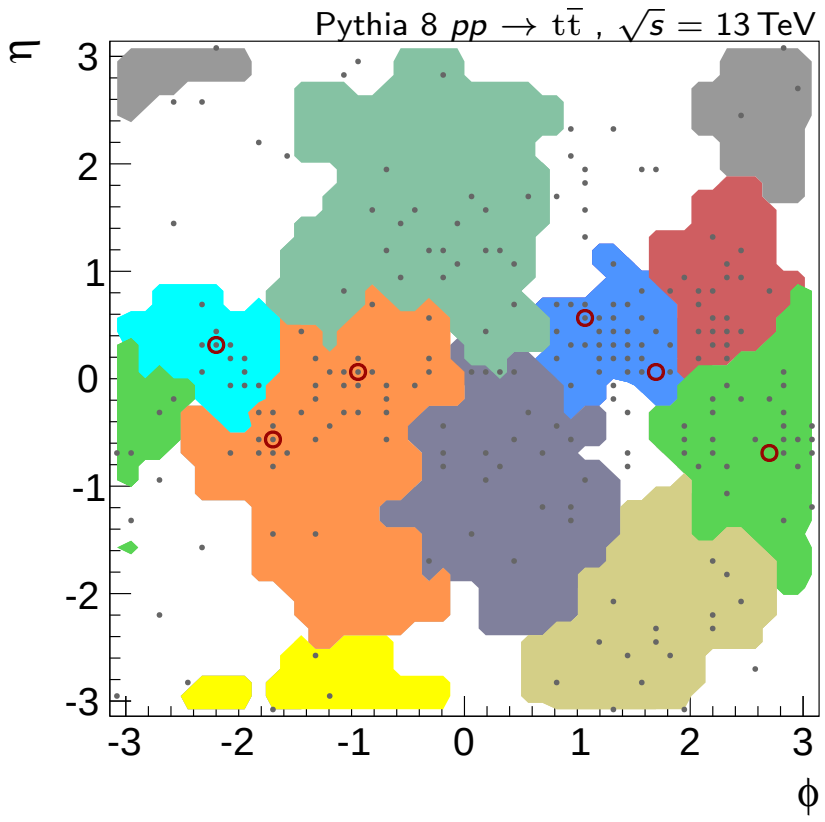
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



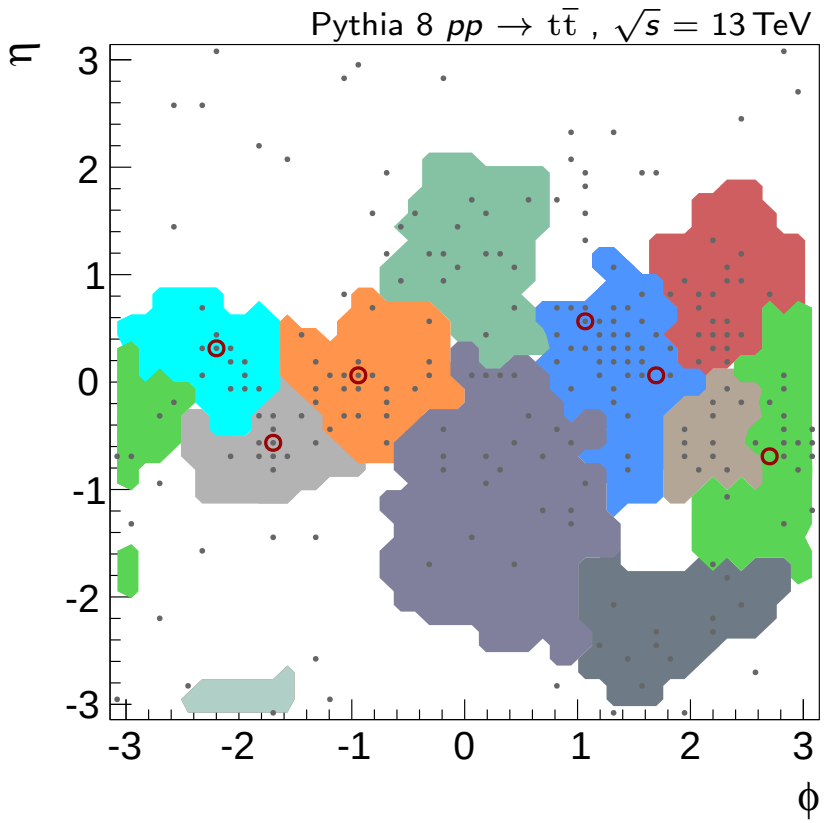
HOTVR clustering



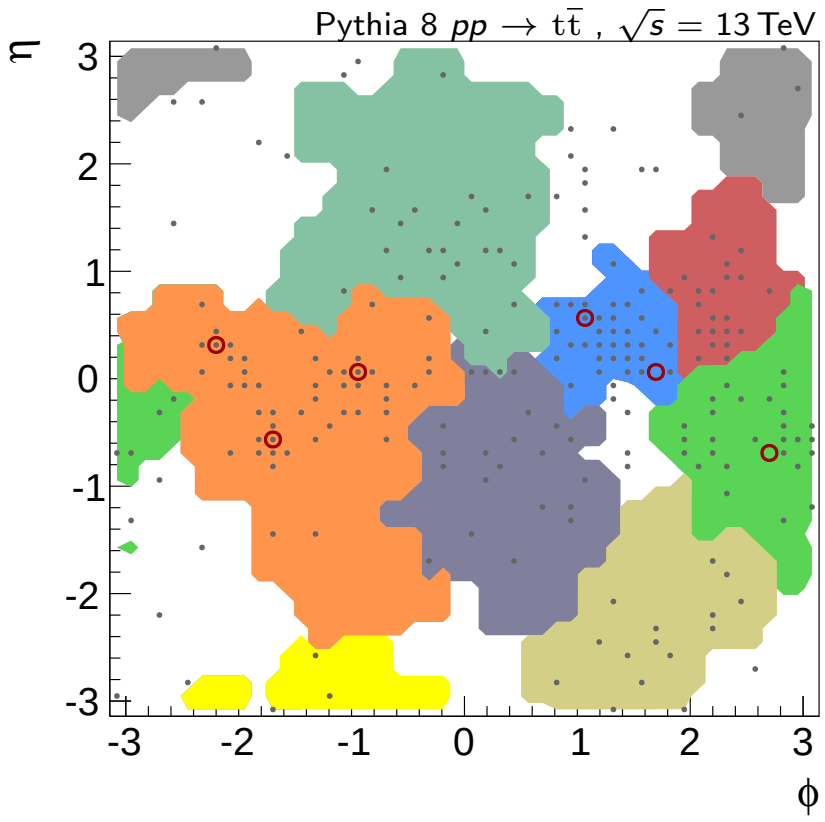
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



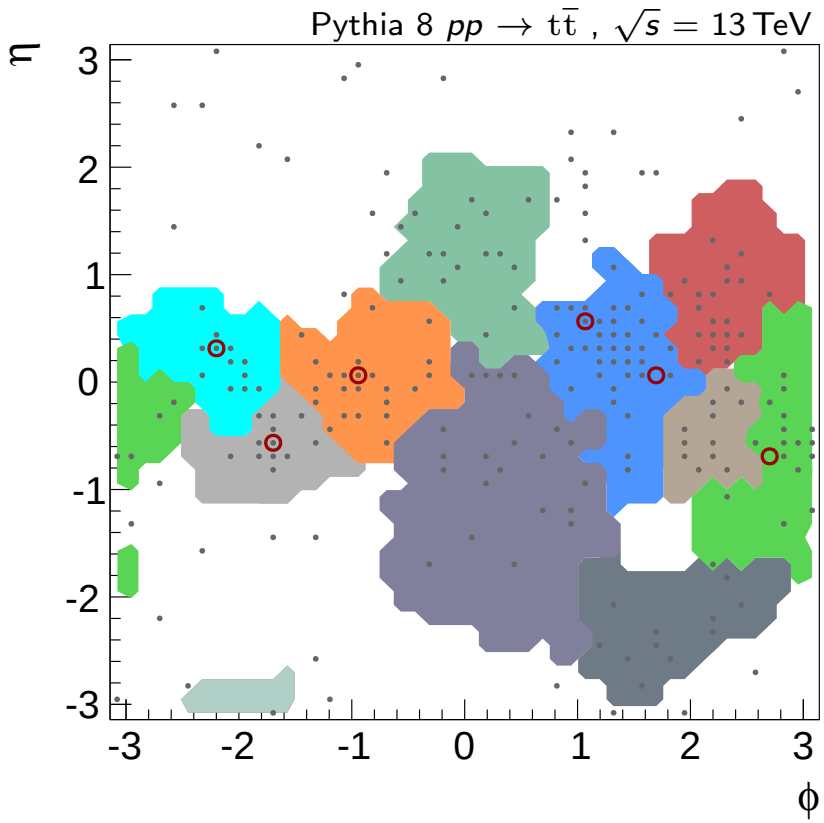
HOTVR clustering



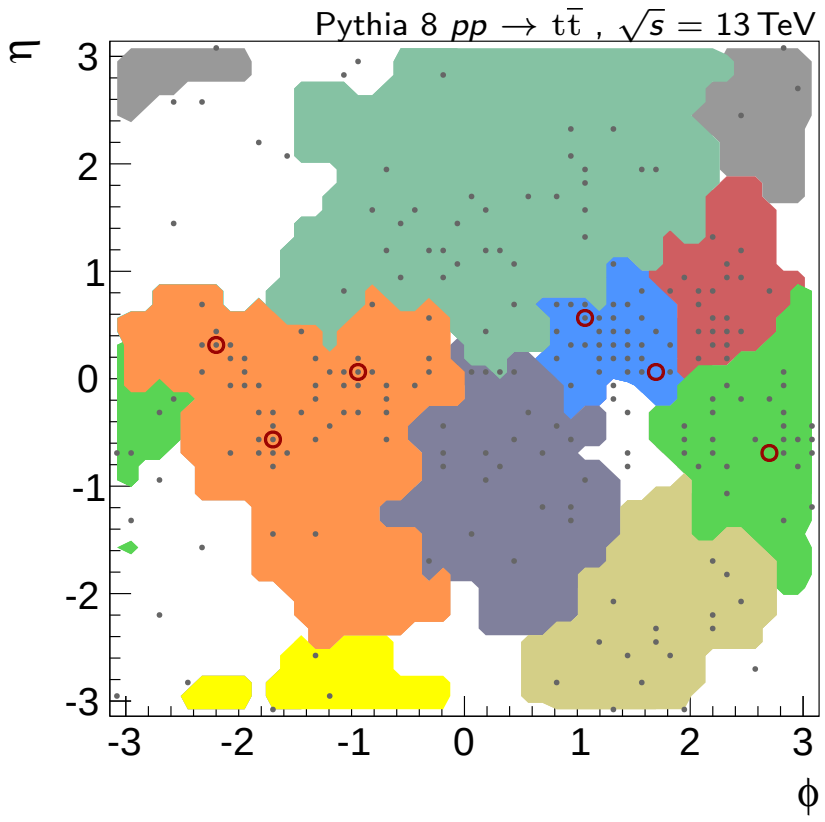
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



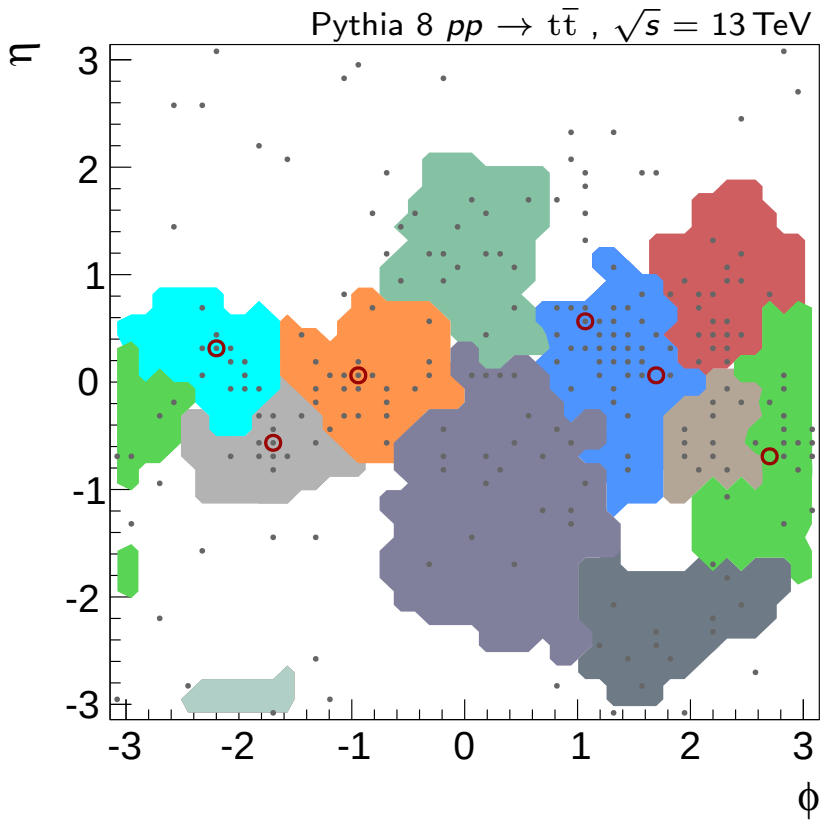
HOTVR clustering



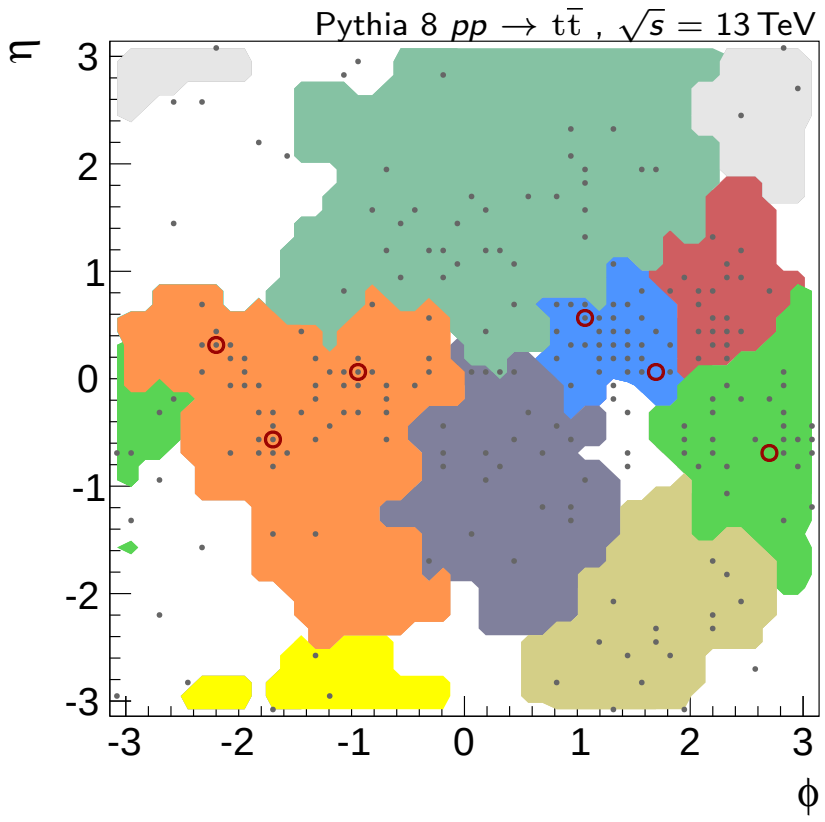
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



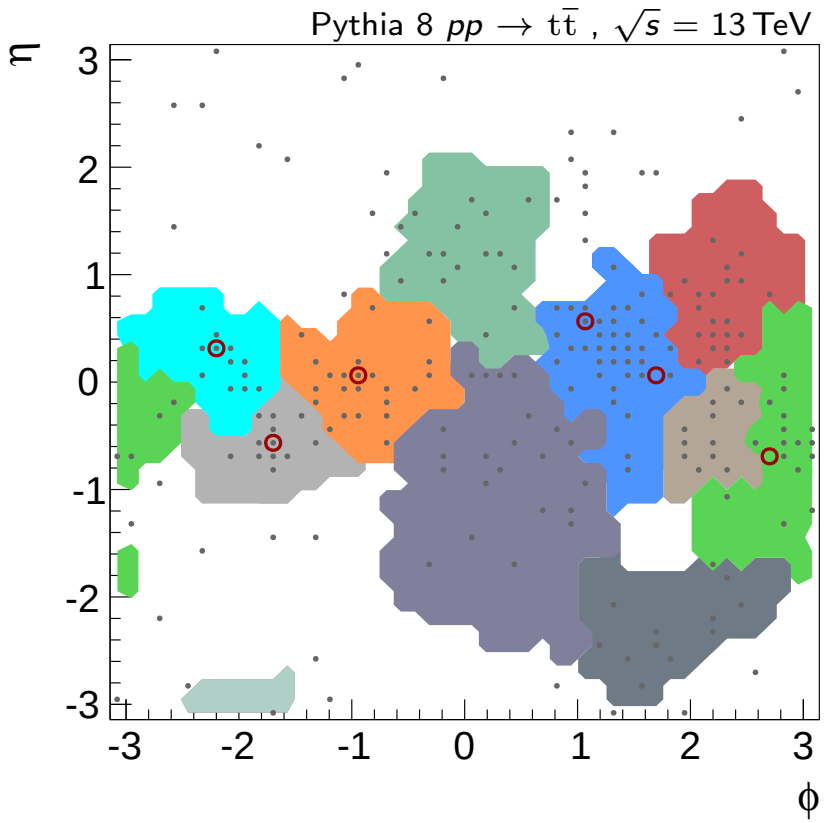
HOTVR clustering



- Stop clustering

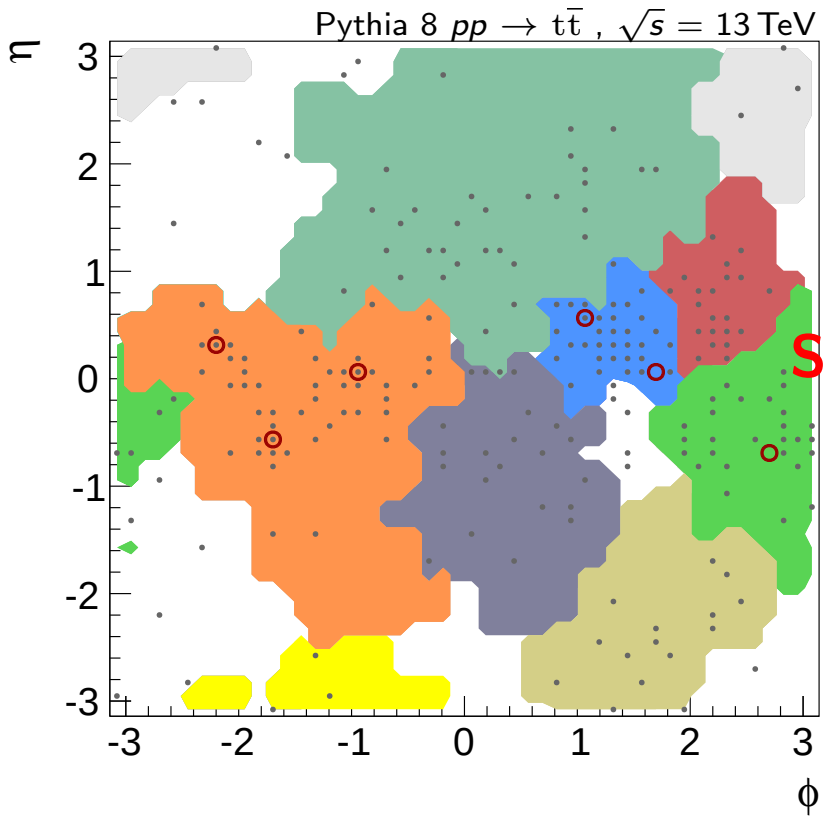
Clustering example 1, top quark $p_T \approx 240\,\text{GeV}$

Cambridge/Aachen clustering



- Stop clustering

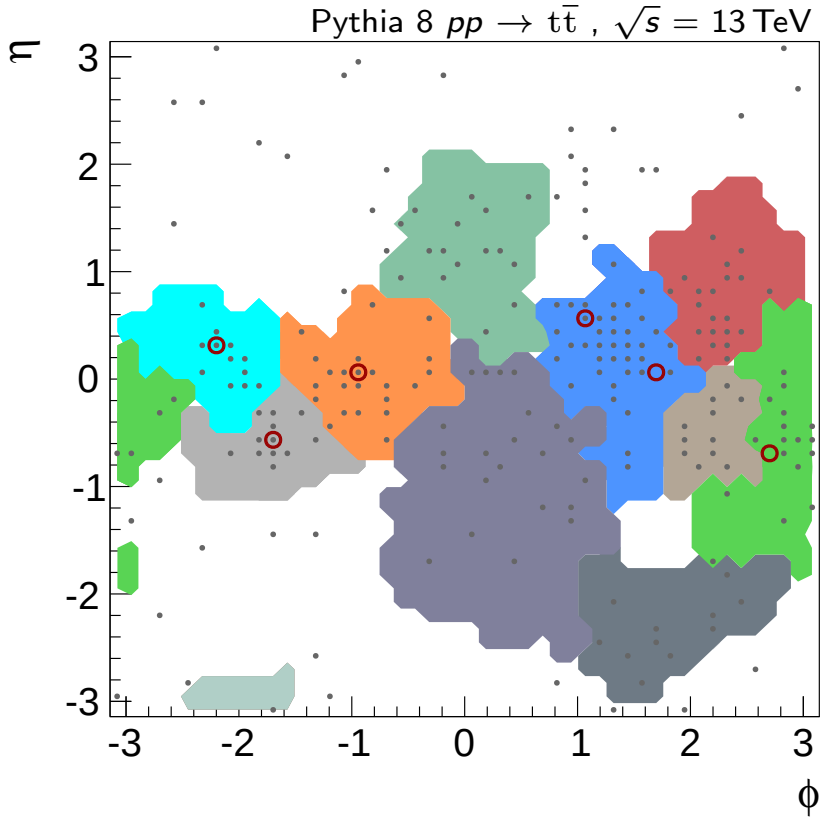
HOTVR clustering



- Massjump found: save subjets

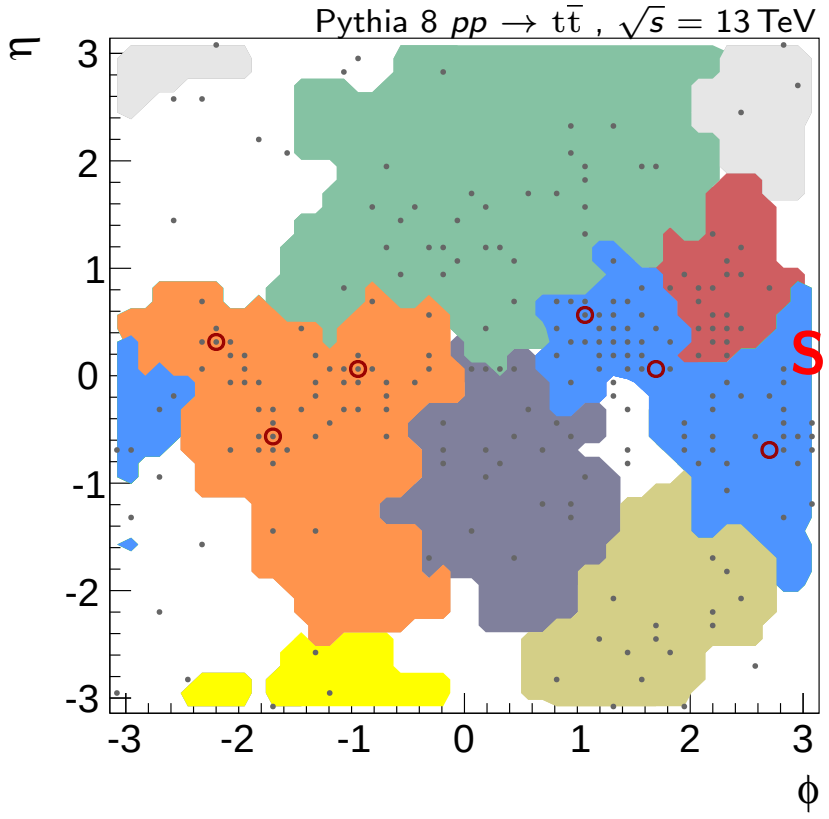
Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



- Stop clustering

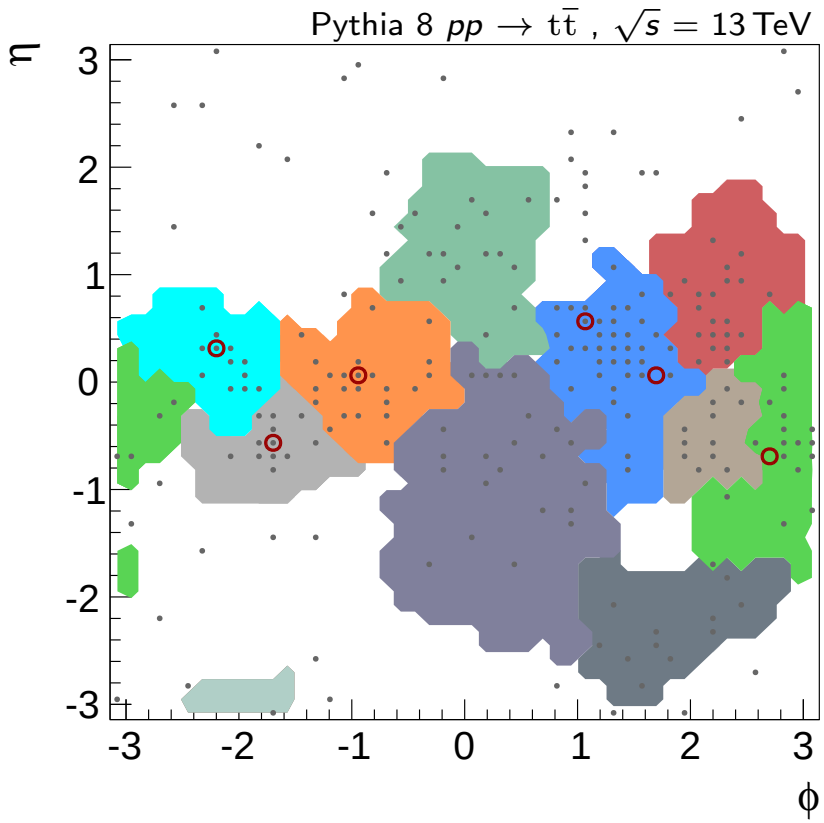
HOTVR clustering



- Massjump found: save subjets

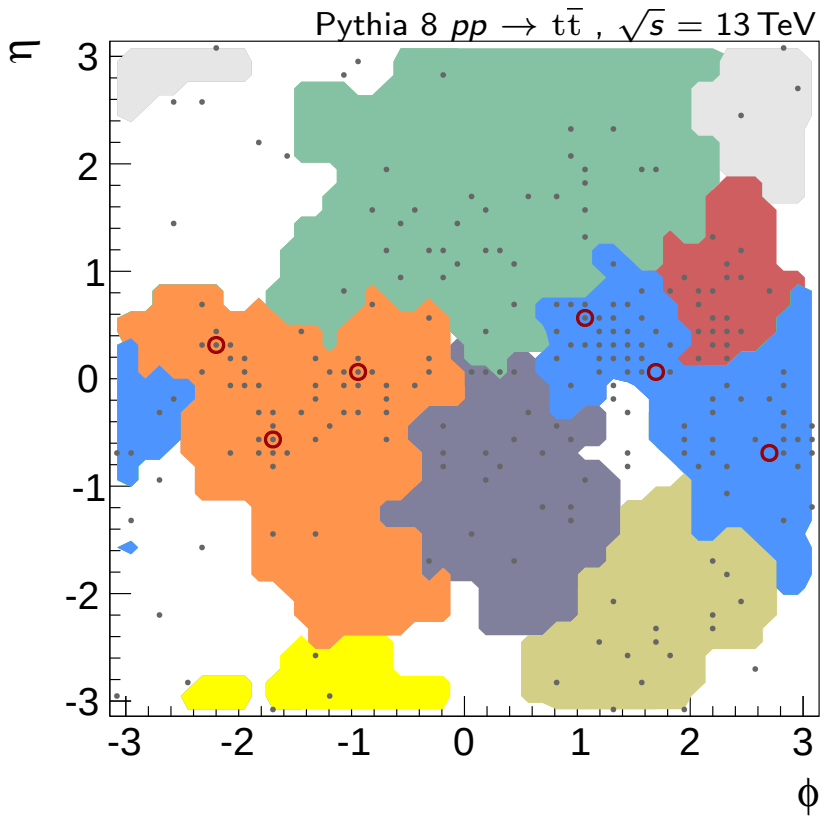
Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



- Stop clustering

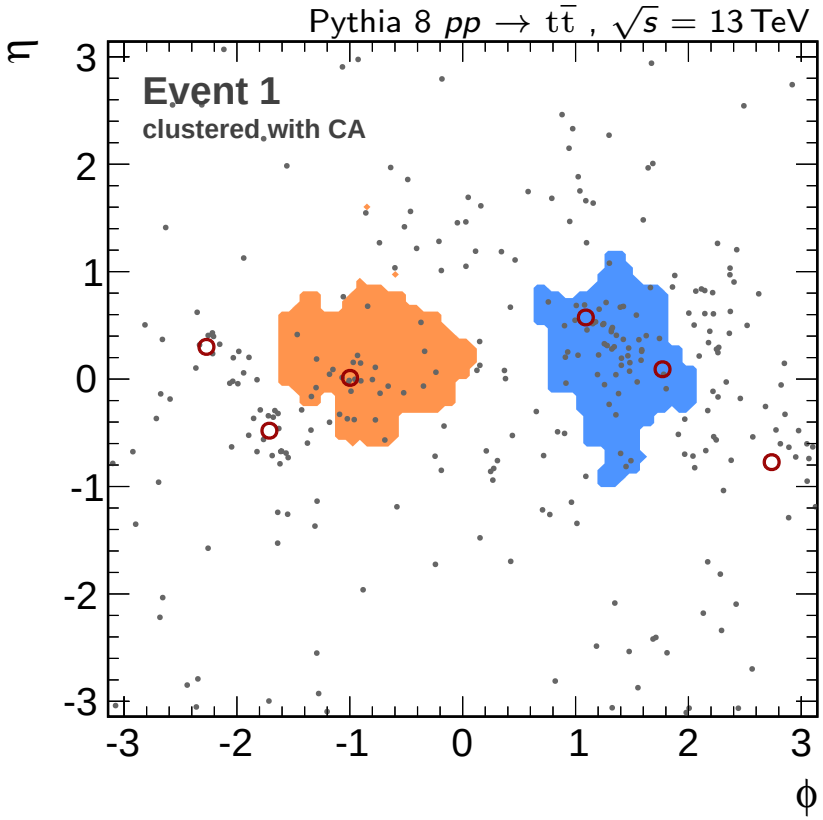
HOTVR clustering



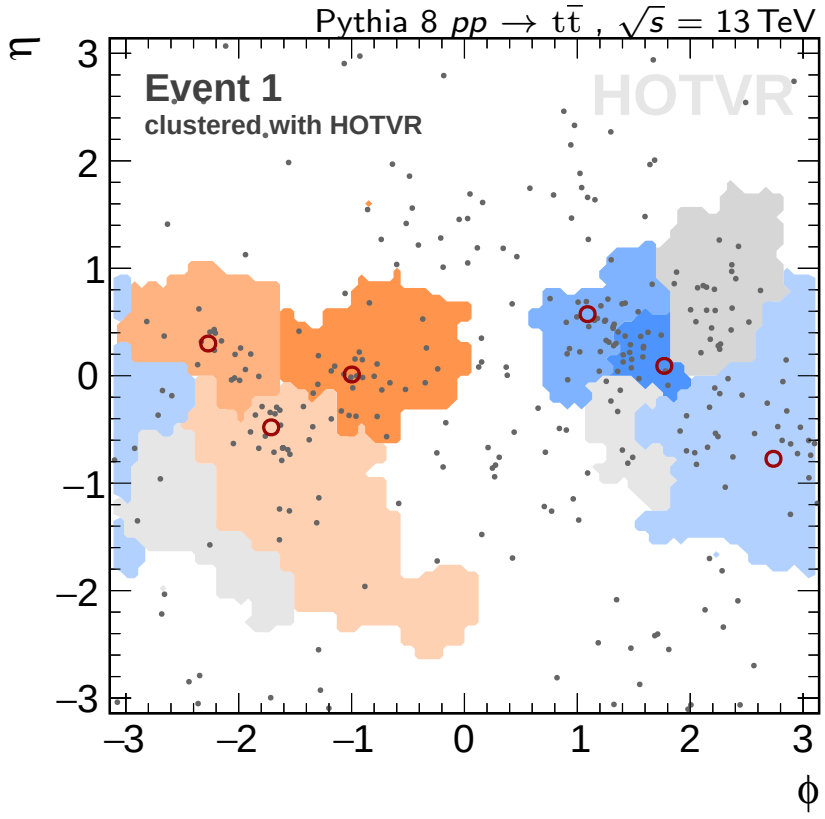
- Stop clustering

Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



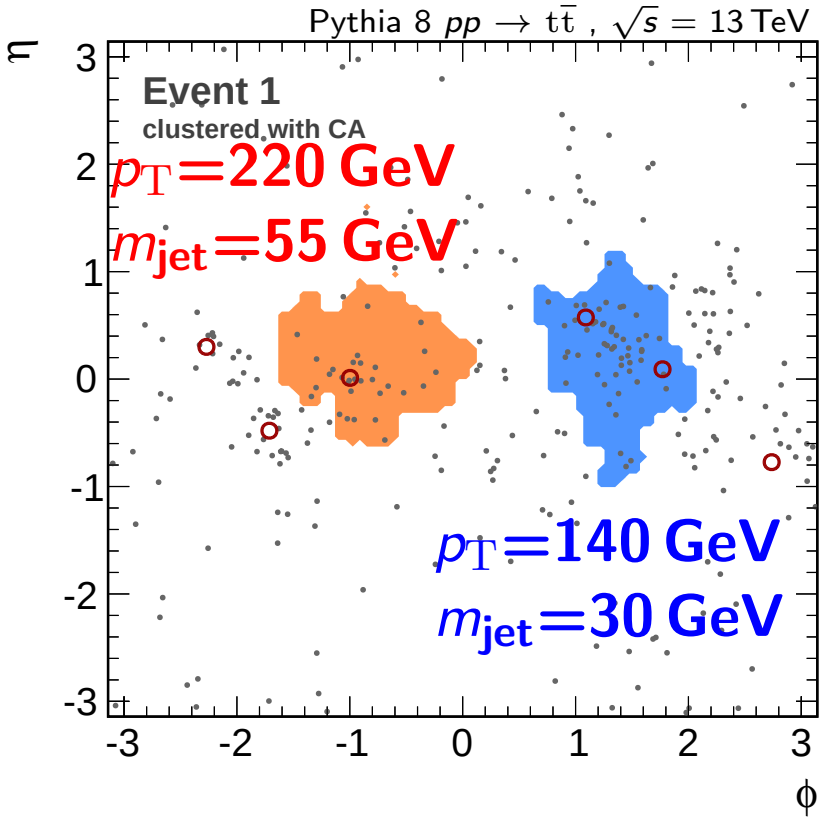
HOTVR clustering



- CA8 jets too small
- All decay products clustered

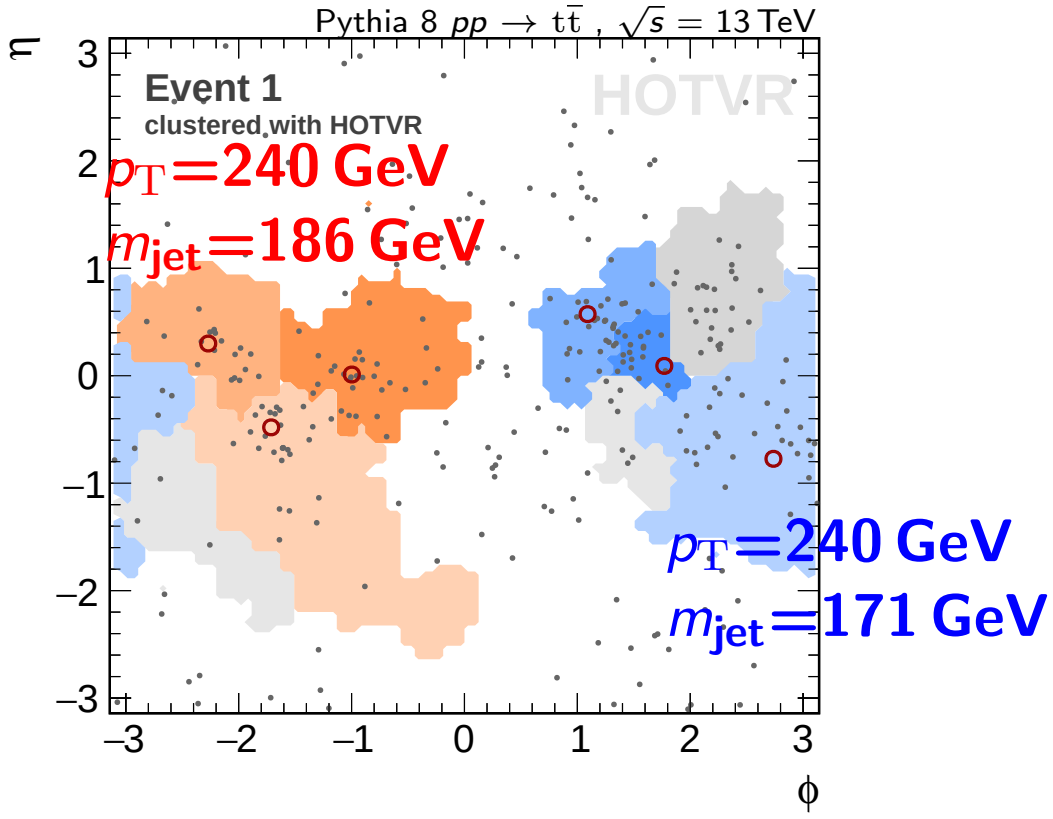
Clustering example 1, top quark $p_{\mathrm{T}} \approx 240\,\mathrm{GeV}$

Cambridge/Aachen clustering



- CA8 jets too small

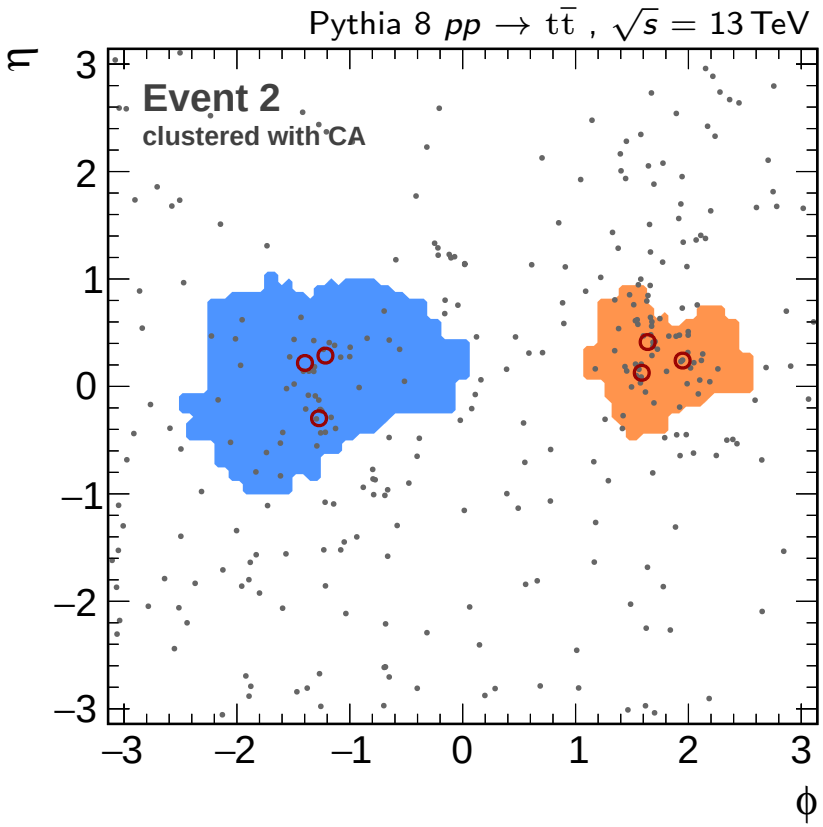
HOTVR clustering



- All decay products clustered

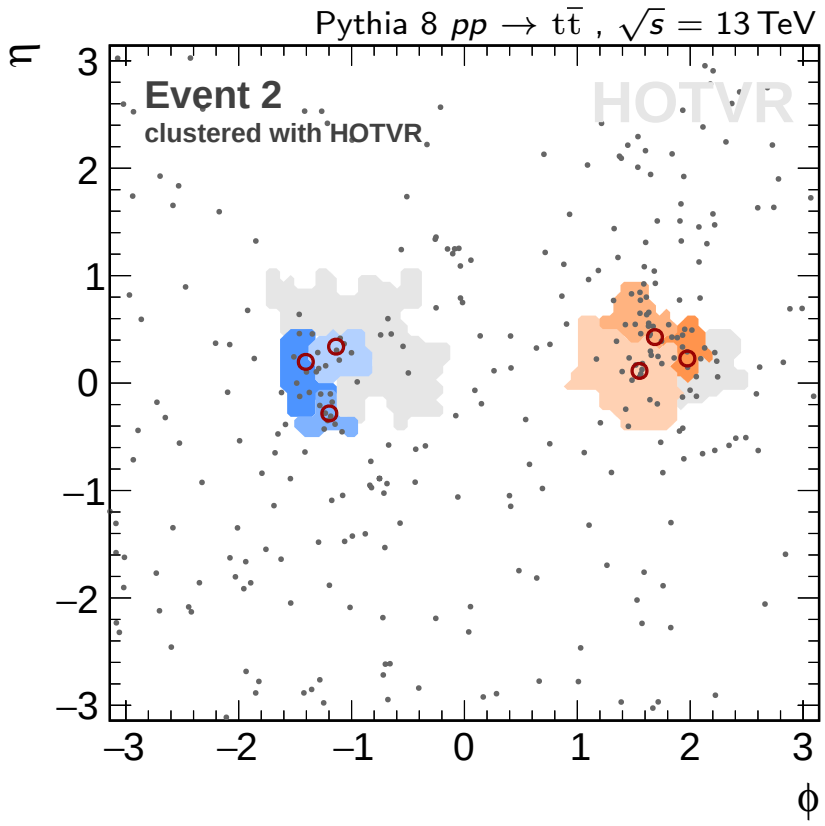
Clustering example 2, top quark $p_{\mathrm{T}} \approx 850\,\mathrm{GeV}$

Cambridge/Aachen clustering



- CA8 jets too large

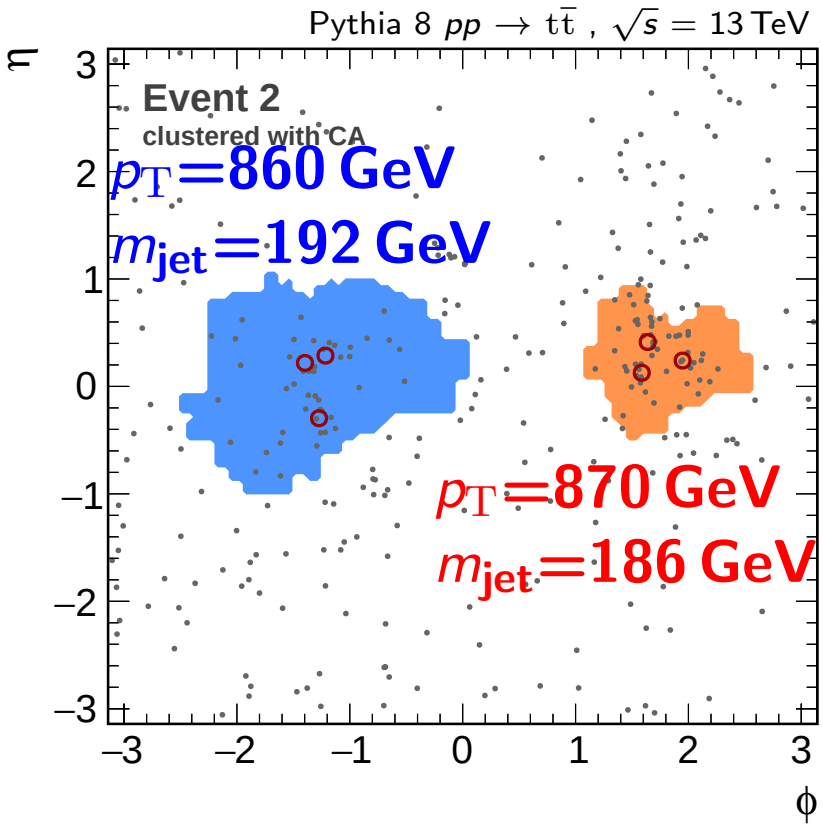
HOTVR clustering



- All decay products clustered

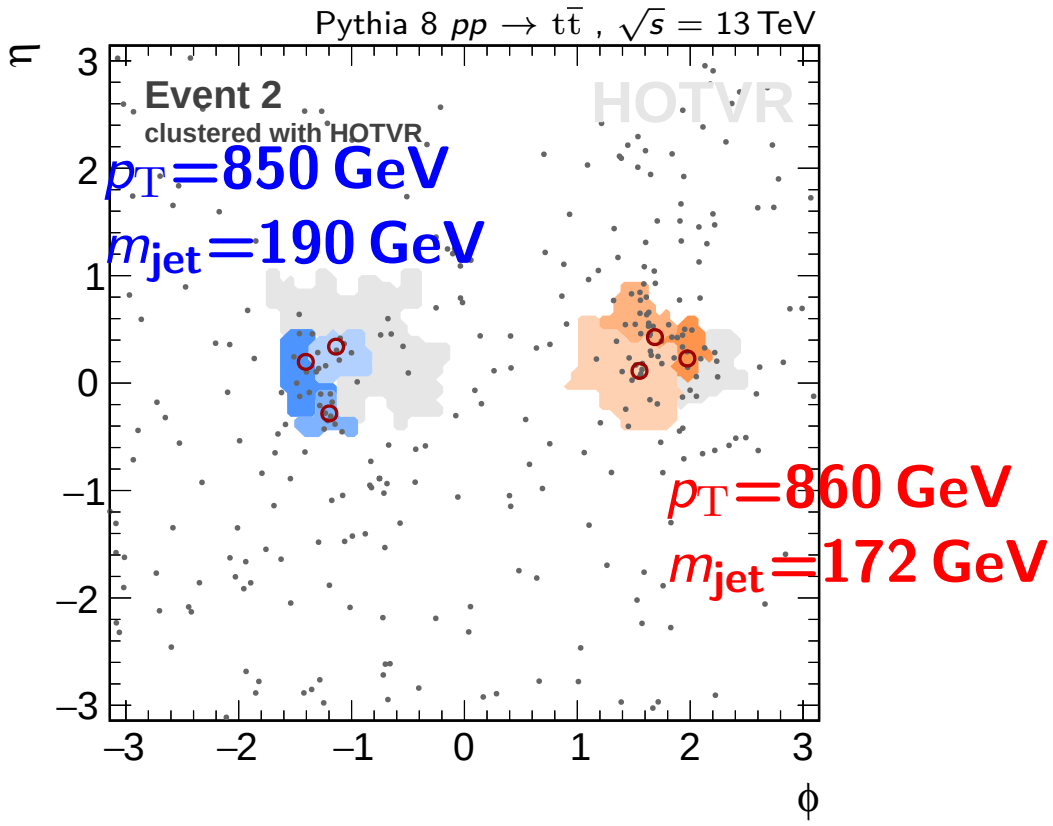
Clustering example 2, top quark $p_{\mathrm{T}} \approx 850\,\mathrm{GeV}$

Cambridge/Aachen clustering



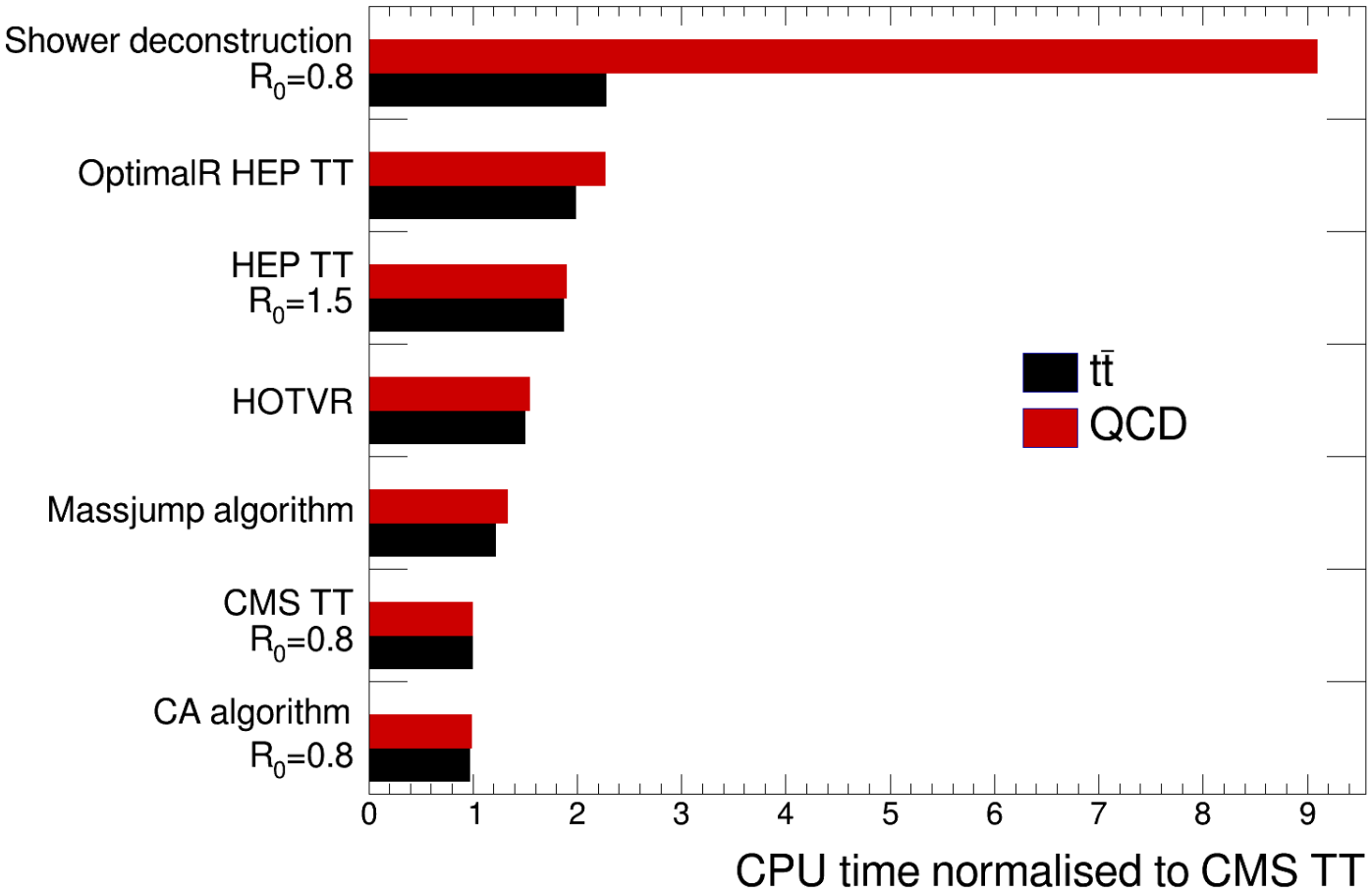
- CA8 jets too large

HOTVR clustering



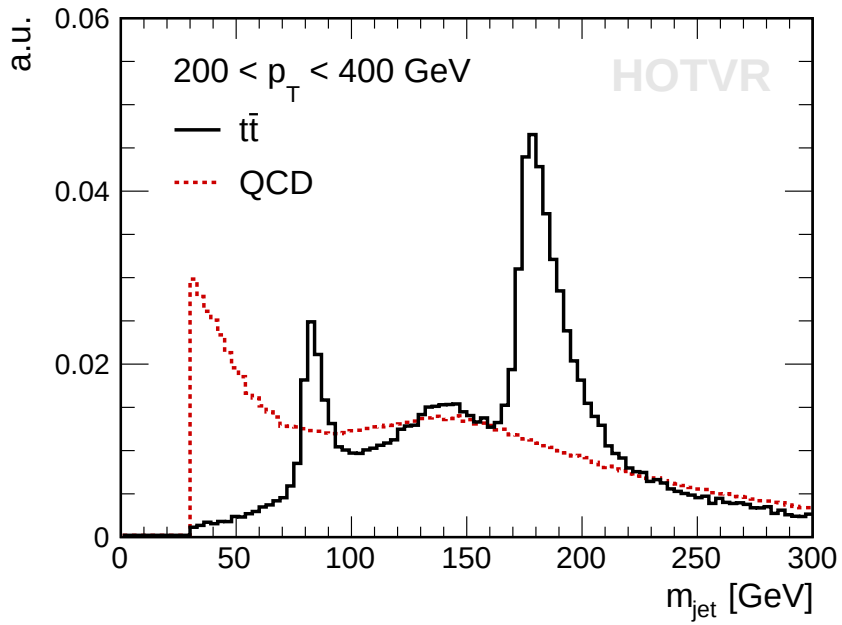
- All decay products clustered

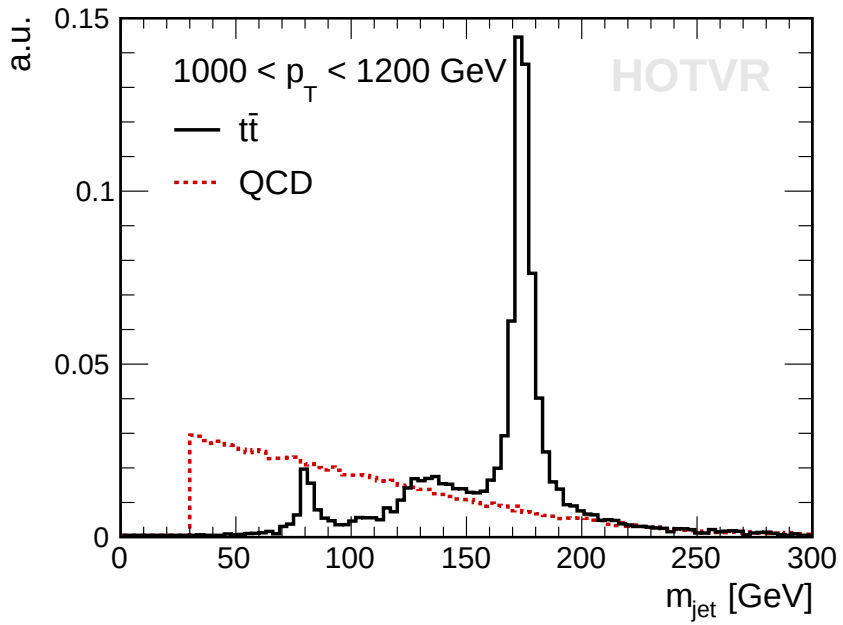
Computational time



- Test performed on 3M events
- Fastjet 3.0.5, Fastjet Contribs 1.017

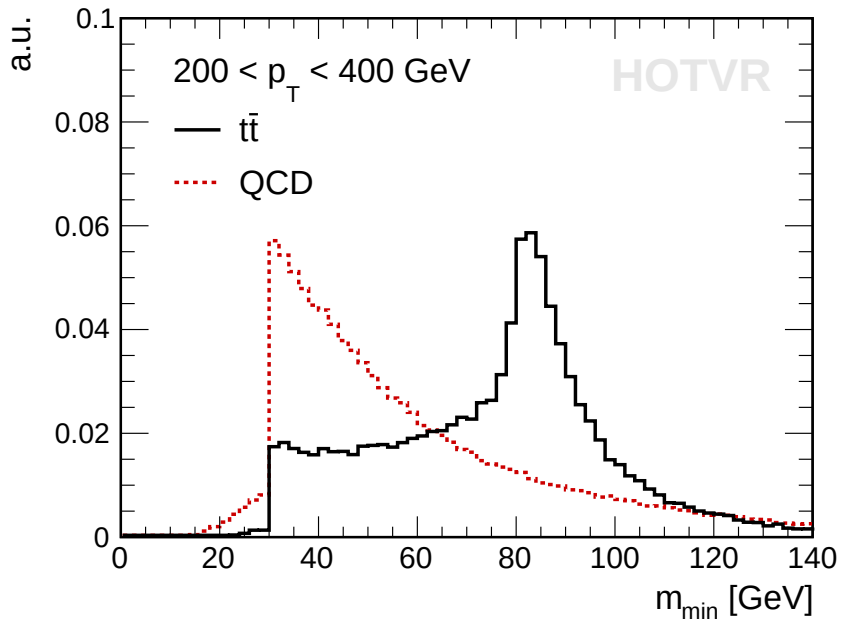
Substructure variables

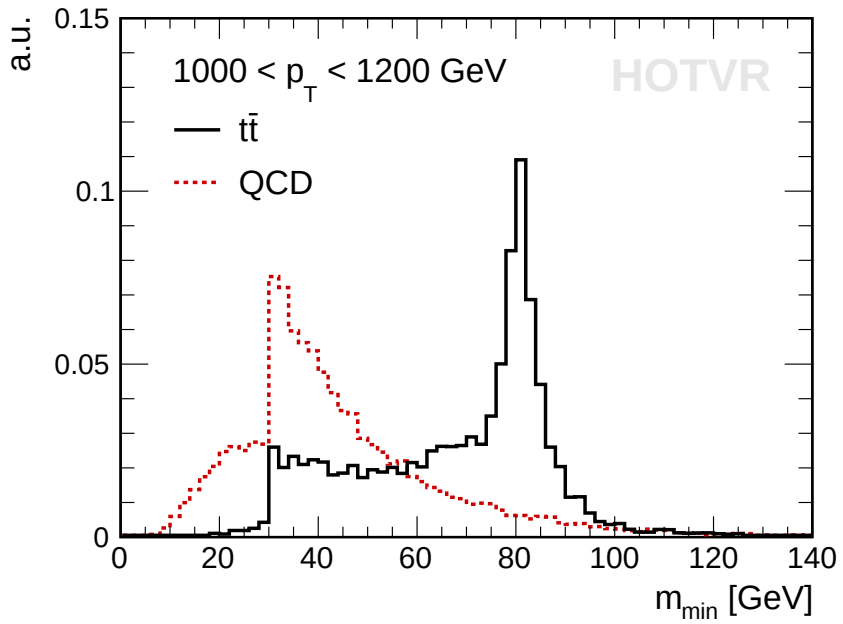




- Jet mass peaks at the top quark mass and is stable over a large p_T range
- Separation power increases with increasing p_T
- Other peaks caused by unmerged jets
- Standard selection: $140 < m_{\text{jet}} < 220\,\text{GeV}$

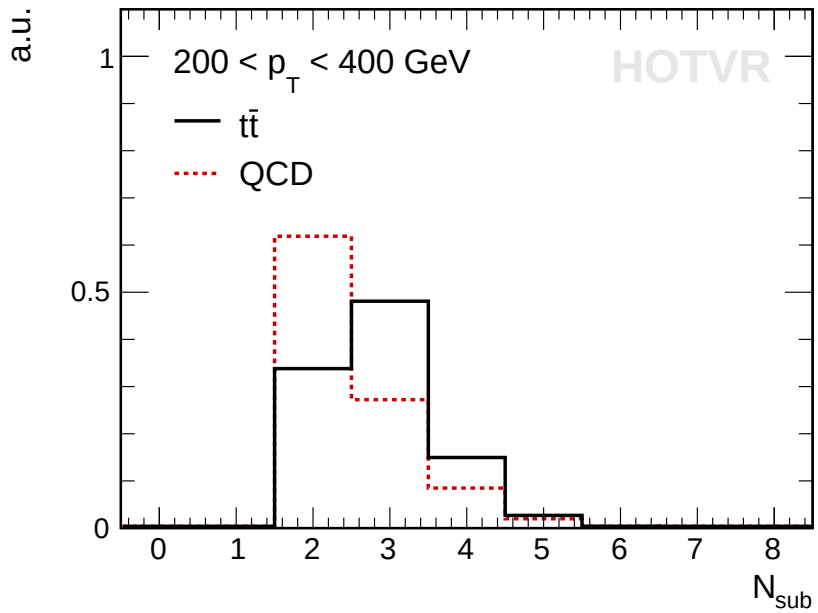
Substructure variables

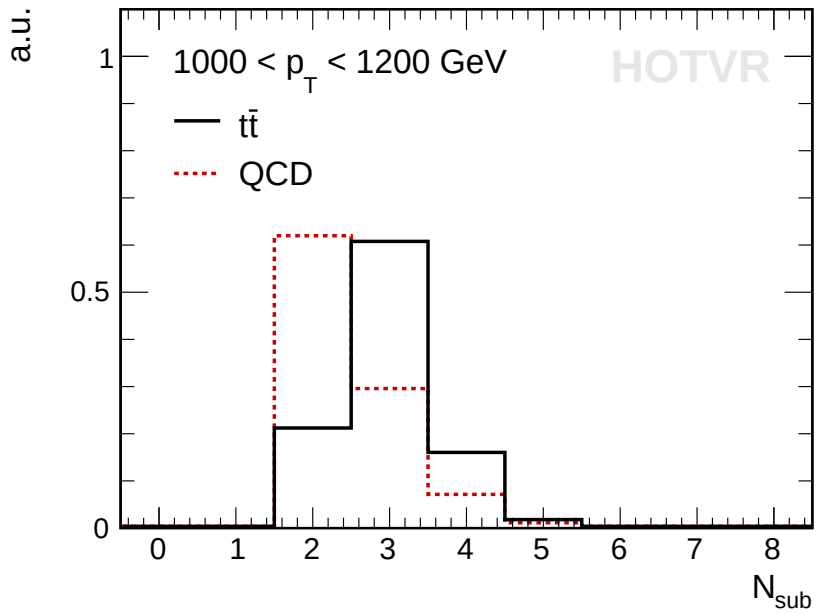




- Minimum pairwise mass: $m_{\text{min}} = \min[m_{12}, m_{13}, m_{23}]$
- Signal distribution peaks at the W mass
- Standard selection: $m_{\text{min}} > 50\,\text{GeV}$

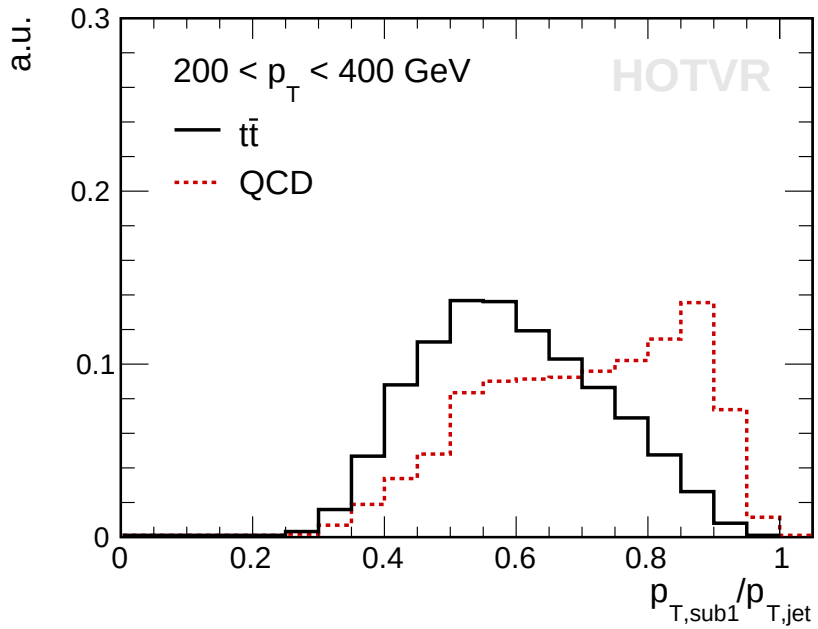
Substructure variables

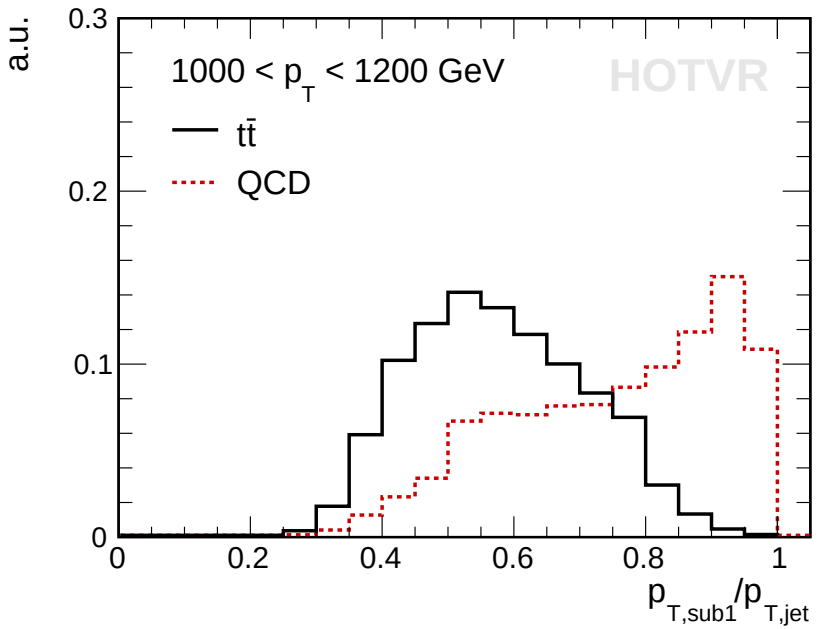




- Signal distribution peaks at 3 (3 decay products of the top quark)
- Separation power increases with increasing p_{T}
- Jets with only one subjet are already rejected
- Standard selection: $N_{\mathrm{sub}} \geq 3$

Substructure variables



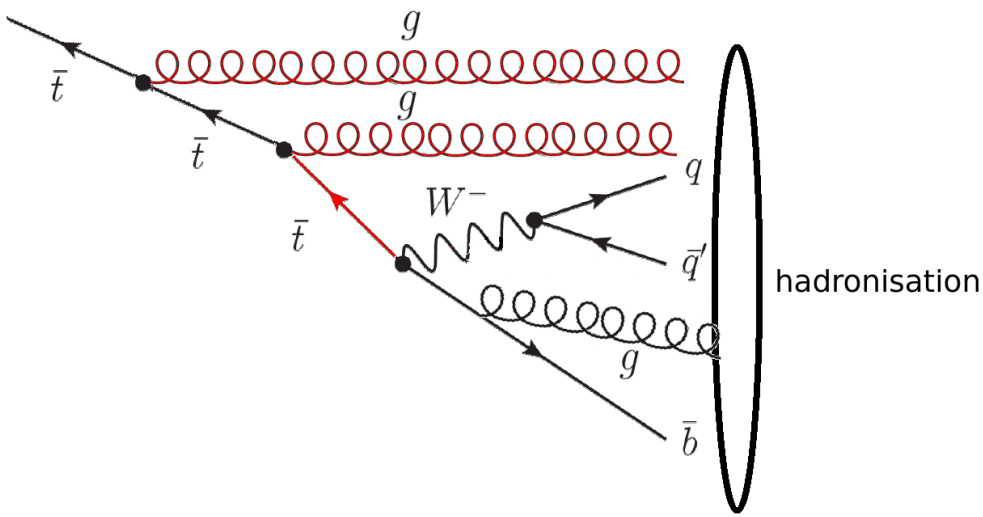


- Transverse momentum more distributed among all subjets for signal events
- Standard selection: $p_{\mathrm{T,sub1}}/p_{\mathrm{T,jet}} < 0.8$

Efficiency & Matching

Parton jet

- Cluster all final state partons, use top quark instead of its decay partons
- Cluster with anti-$\mathrm{k_T}$ with $R_0 = 0.4$ ($p_\mathrm{T} > 100\,\mathrm{GeV}$)



Matching

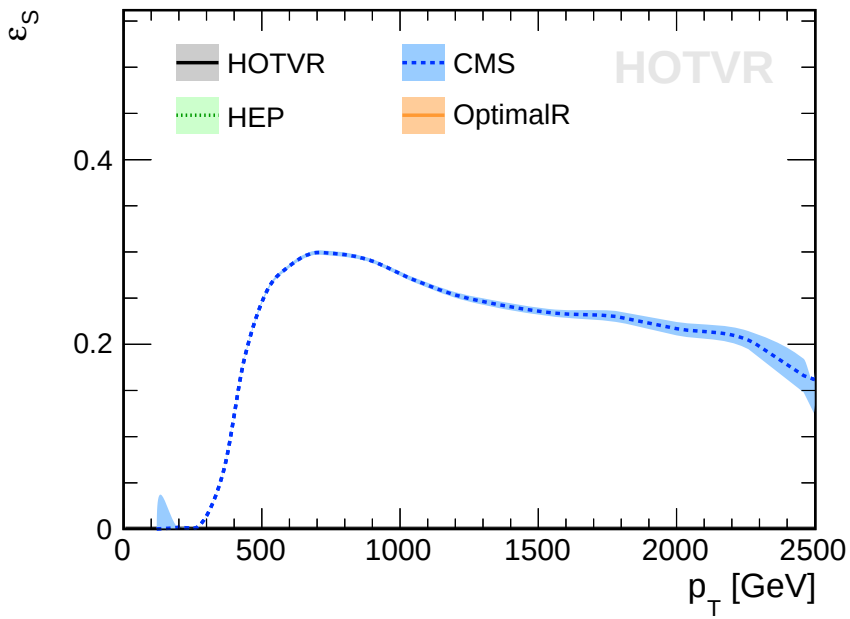
- Matching to parton jet with $\Delta R <$ jet size

Efficiency and mistag-rate

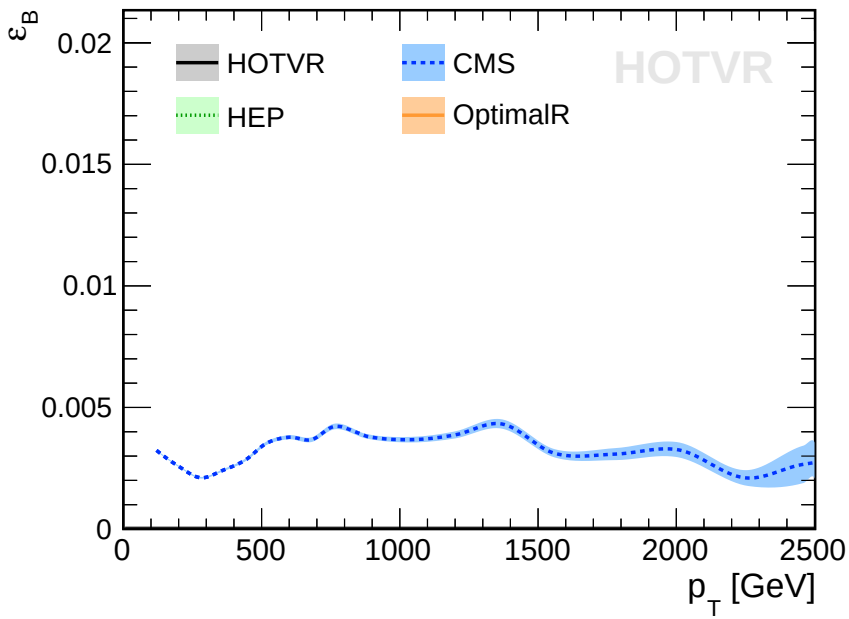
$$\epsilon_{\mathrm{S/B}} = \frac{\text{matched \& tagged parton jets}}{\text{all parton jets (containing a top quark)}}$$

Efficiency & mistag-rate

Efficiency:



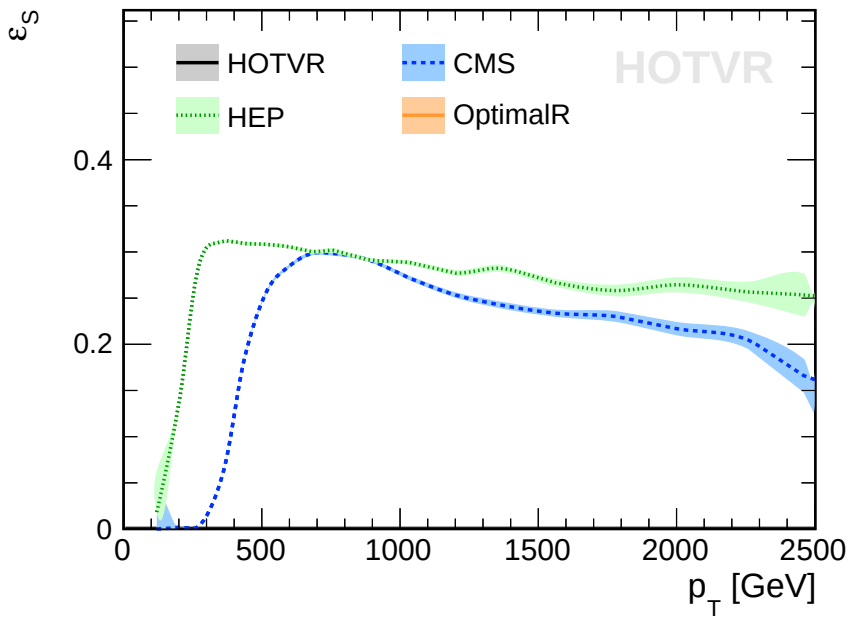
Mistag-rate:



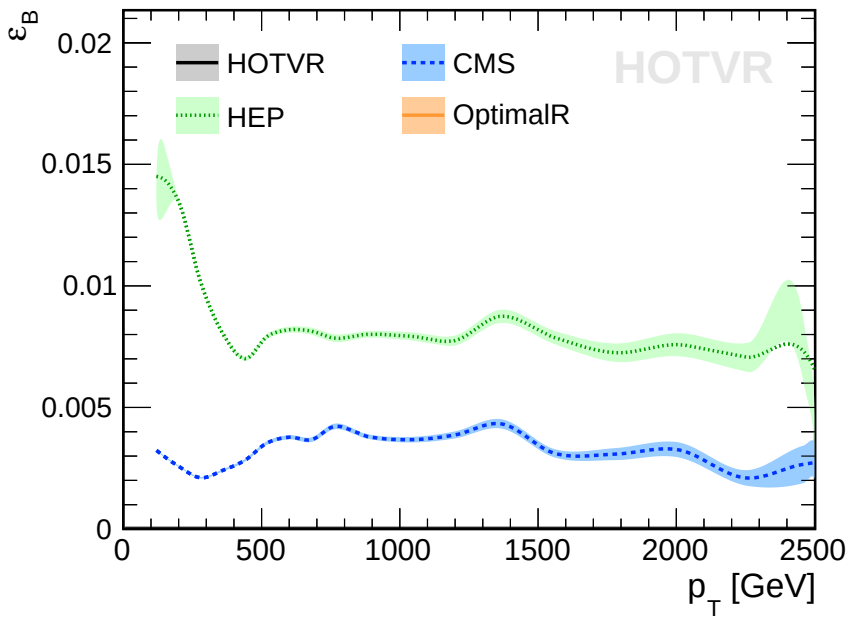
- Normalised to average efficiency of 30% in $p_{\mathrm{T}} = 600 - 1000\,\mathrm{GeV}$ using cut on τ_3/τ_2

Efficiency & mistag-rate

Efficiency:



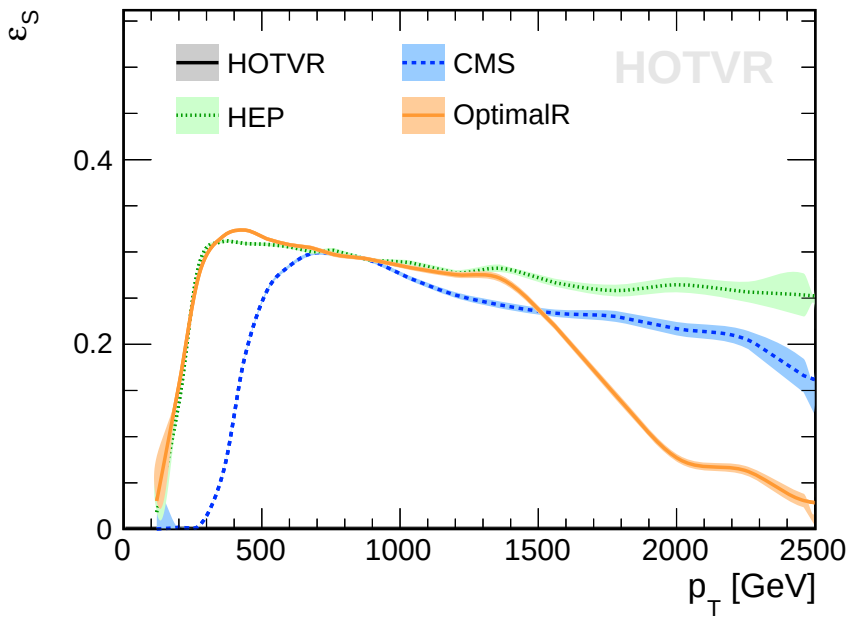
Mistag-rate:



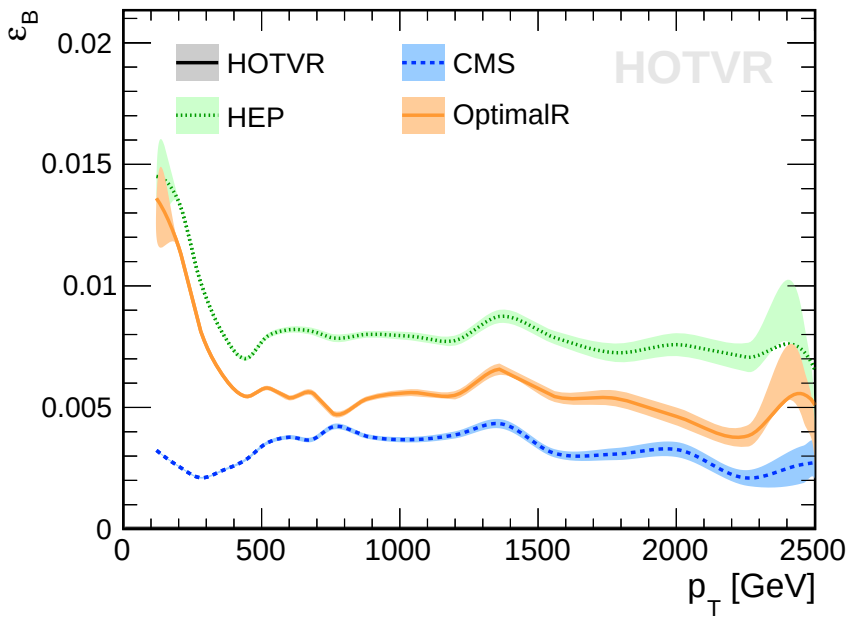
- Normalised to average efficiency of 30% in $p_{\mathrm{T}} = 600 - 1000\,\mathrm{GeV}$ using cut on τ_3/τ_2

Efficiency & mistag-rate

Efficiency:



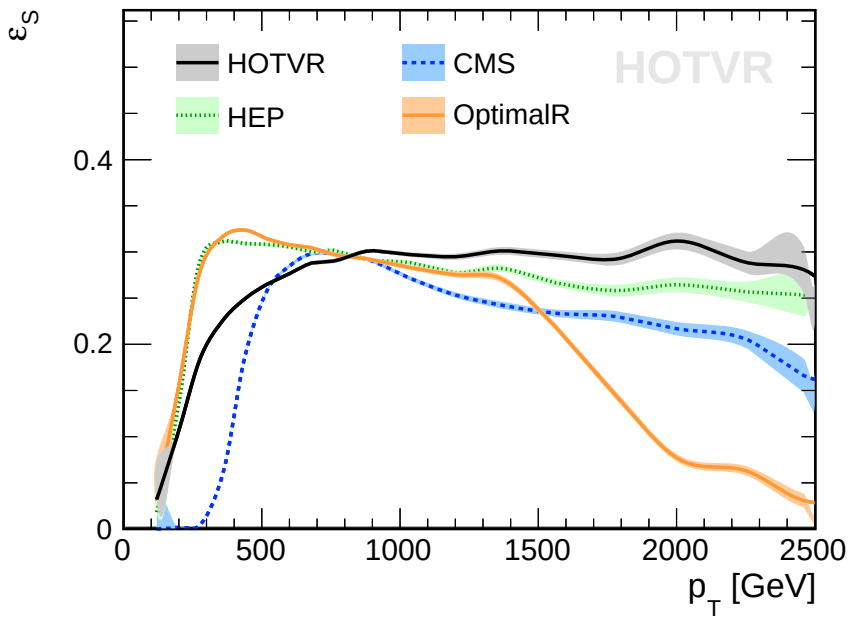
Mistag-rate:



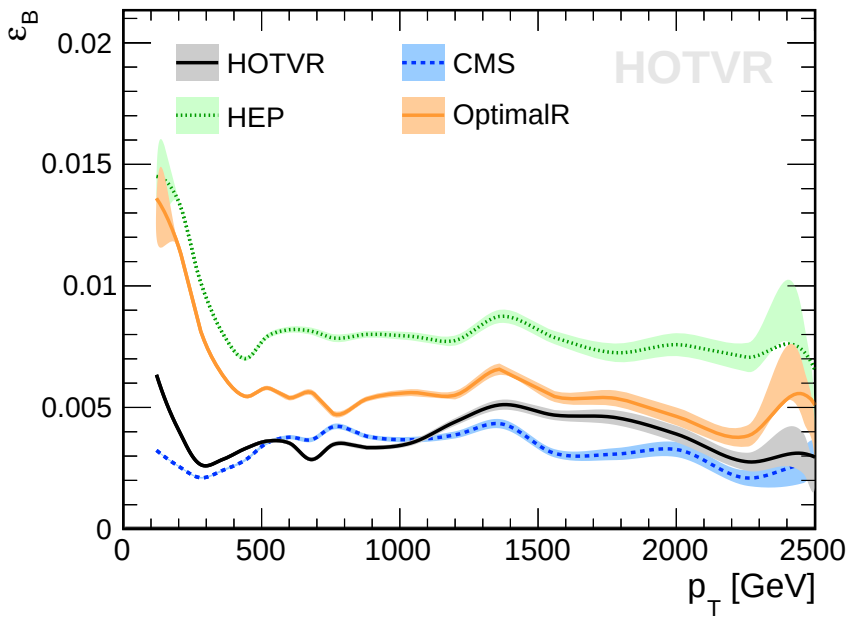
- Normalised to average efficiency of 30% in $p_{\mathrm{T}} = 600 - 1000\,\mathrm{GeV}$ using cut on τ_3/τ_2

Efficiency & mistag-rate

Efficiency:

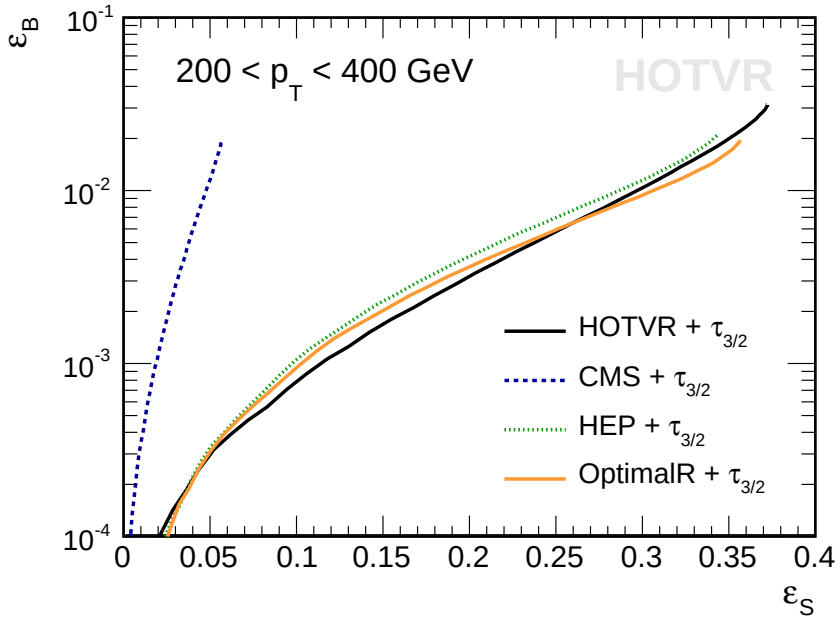


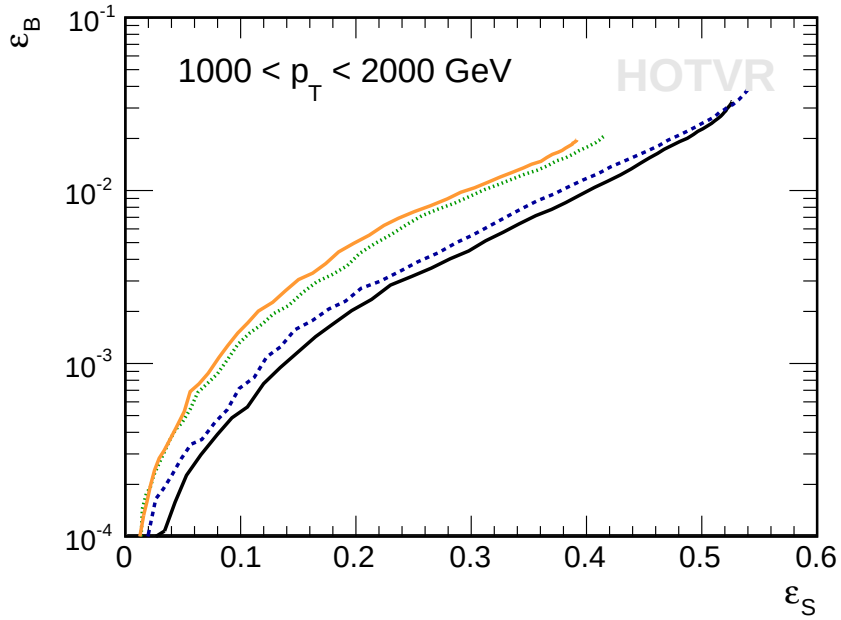
Mistag-rate:



- Normalised to average efficiency of 30% in $p_{\mathrm{T}} = 600 - 1000\,\mathrm{GeV}$ using cut on τ_3/τ_2
- HOTVR efficient already below 300 GeV
- HOTVR reaches plateau around 500 GeV, constant above
- Very small mistag rate (similar to CMS TT)

Performance



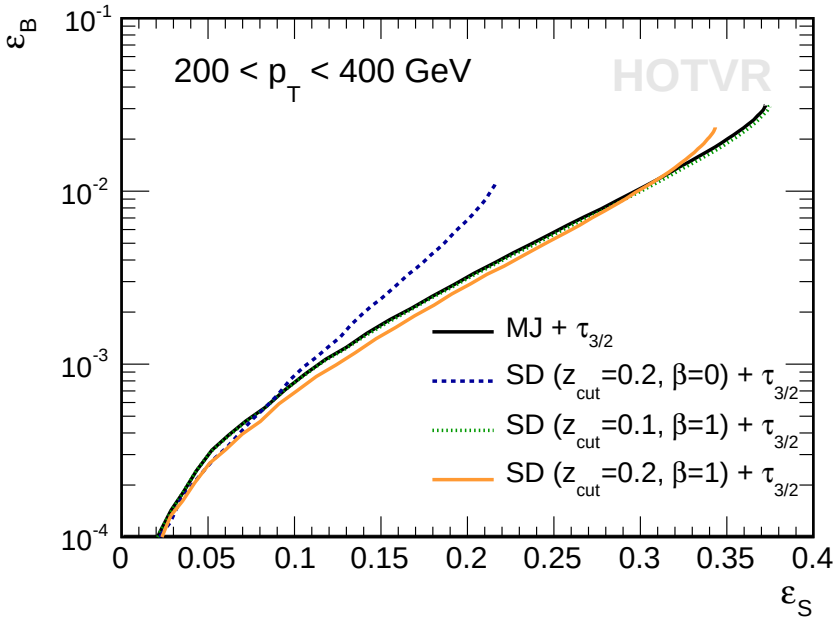


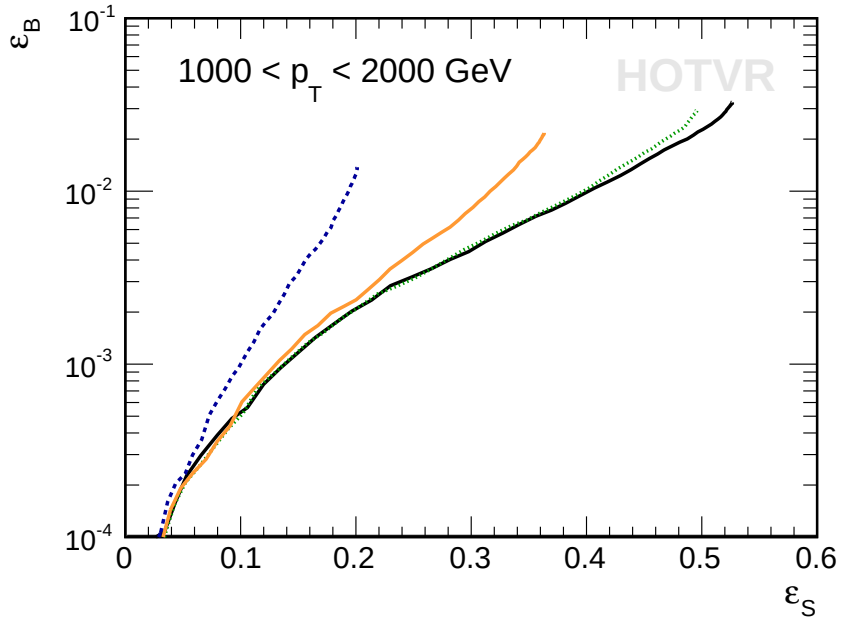
- Performant in low and high p_{T} regions

comparison study: HOTVR & softdrop

Replace mass jump criterion with softdrop criterion

$$\theta \cdot m_{ij} > \max[m_i, m_j] \quad \rightarrow \quad \frac{\min[p_{T,i}, p_{T,j}]}{p_{T,i} + p_{T,j}} > z_{\text{cut}} \left(\frac{\Delta R_{ij}}{R_{\text{eff}}} \right)^{\beta}$$

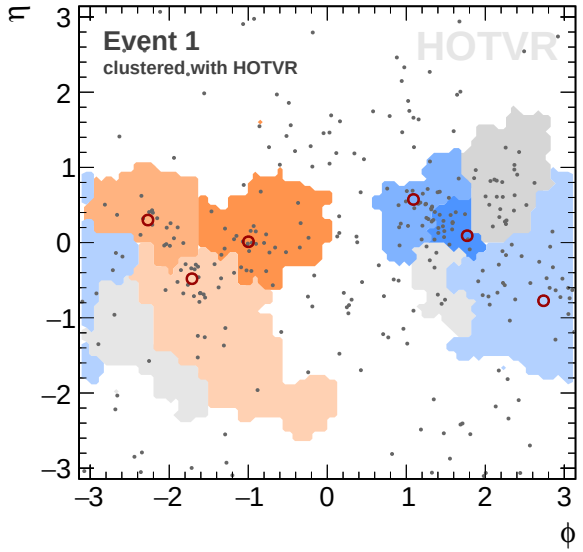


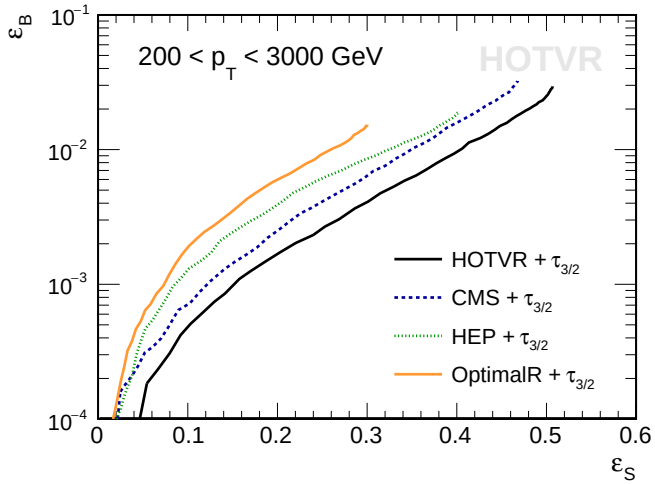


Summary

The HOTVR algorithm

- Low complexity
  - in one sequence:
    adapting jet radius (variable R),
    subjet finding (massjump criterion)
    and rejecting soft clusters
- Good computational performance
- Infrared and collinear safe
- Performant in low and high p_{T}
- Also usable for W/Z/H tagging
  (currently in progress)
- HOTVR in Fastjet contribs:
  "contribs/HOTVR"





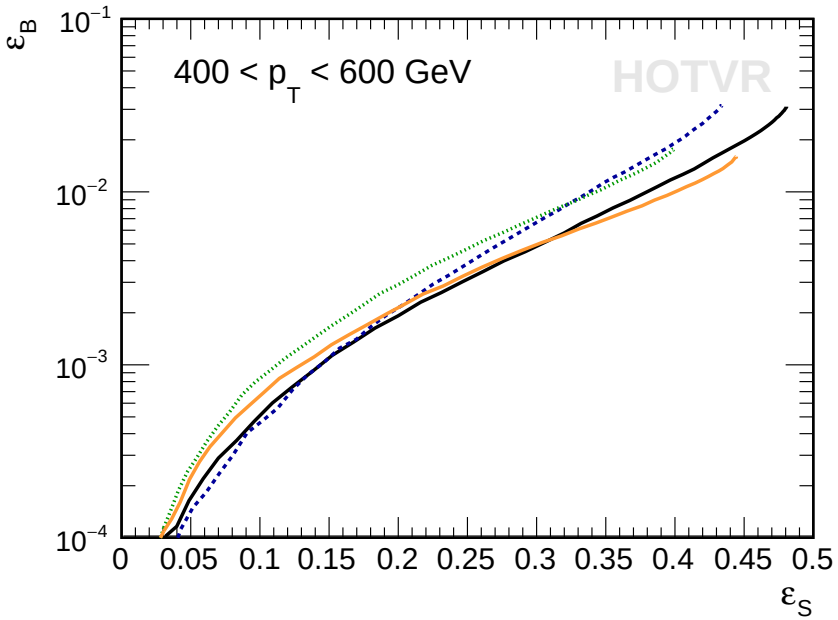
Backup

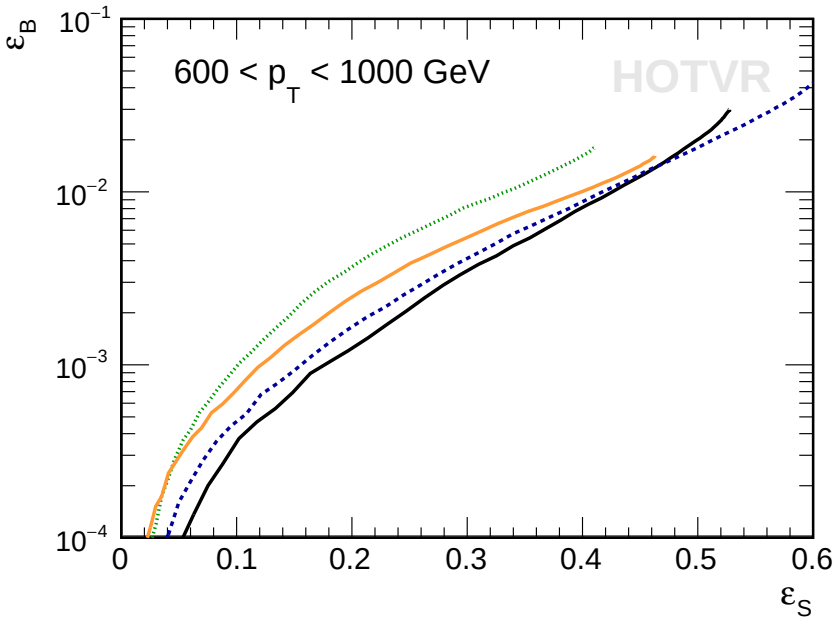
Standard working points

CMS top tagger $R_0 = 0.8$	HEP TT $R_0 = 1.5$		OptimalR $R_0 = 0.5$–1.5
$\delta_p > 0.05$	$f_{\text{drop}} = 0.8$	$m_{23}/m_{123} > 0.35$	same as HEP TT
$A = 0.0004$	$m_{\text{cut}} = 30\,\text{GeV}$	$0.2 < \arctan \frac{m_{13}}{m_{12}} < 1.3$	$\Delta R = 0.1$
$N_{\text{sub}} \geq 3$	$R_{\text{filt}} = 0.3$	$f_W = 0.15$	$m_{\text{rec}}^{1.5} - m_{\text{rec}} > 0.2 m_{\text{rec}}^{1.5}$
$m_{\text{min}} > 50\,\text{GeV}$	$N_{\text{filt}} = 5$	$140 < m_{123} < 250\,\text{GeV}$	$\Delta R_{\text{opt}} < 0.5$
$140 < m_{\text{jet}} < 250\,\text{GeV}$	$p_{\text{T,sub}} > 30\,\text{GeV}$		

Performance

Intermediate p_{T} range





- Best performance in both p_{T} regions