Jet definition and event selection for |Vts|

May 13, 2024

Lars Bogner, Sofia Giappichini, Simon Keilbach, Jan Kieseler, Markus Klute, Matteo Presilla, **Xunwu Zuo**

Outline

Context:

- $t \to Ws$ decay is a clean and direct measurement of $|V_{ts}|$
- Requires excellent s-tagging performance and calibration
- Initial study at https://indico.cern.ch/event/1390113/#31-top-couplings-and-vts, with a lot of open questions on jet definition and tagging

This update tries to address the following questions

- How well can we assign “true flavor” to jets in ttbar events?
- Which is the better jet definition/selection for this case?
- What are the background composition and yields?

REMARK: this work is far from optimal. The following results provide ideas for next steps, but do not represent the final sensitivity.

Common labels in following slides

jet reconstruction:

(e or μ with pt > 15 GeV are removed)

- **kt2**: exclusive jet reco with exactly 2 jets in the event
- **R5**: inclusive jet reco with radius R=0.5 (and pt > 1 GeV)
- **R8, R10**: inclusive jet reco with radius R=0.8, or 1.0 (and pt > 1 GeV)

jet selection:

jet sel: default selection is $m(j) < 50$ GeV and $E(j) > 15$ GeV

Event labels:

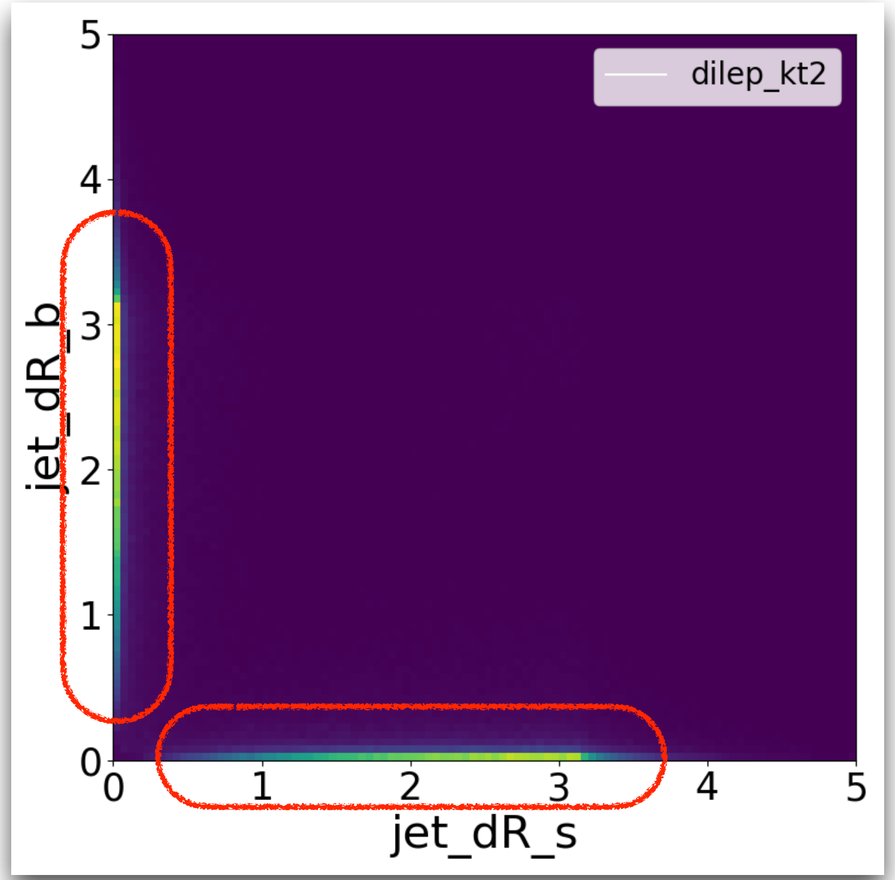
- **dilep**: $t\bar{t} \rightarrow WsWb$ decays, with both $W \rightarrow \ell\nu$ (e or μ)
- **semilep Wud**: $t\bar{t} \rightarrow WsWb$ decays, with one $W \rightarrow ud$
- **semilep Wcs**: $t\bar{t} \rightarrow WsWb$ decays, with one $W \rightarrow cs$
- **dihad**: $t\bar{t} \rightarrow WsWb$ decays, with both $W \rightarrow qq$

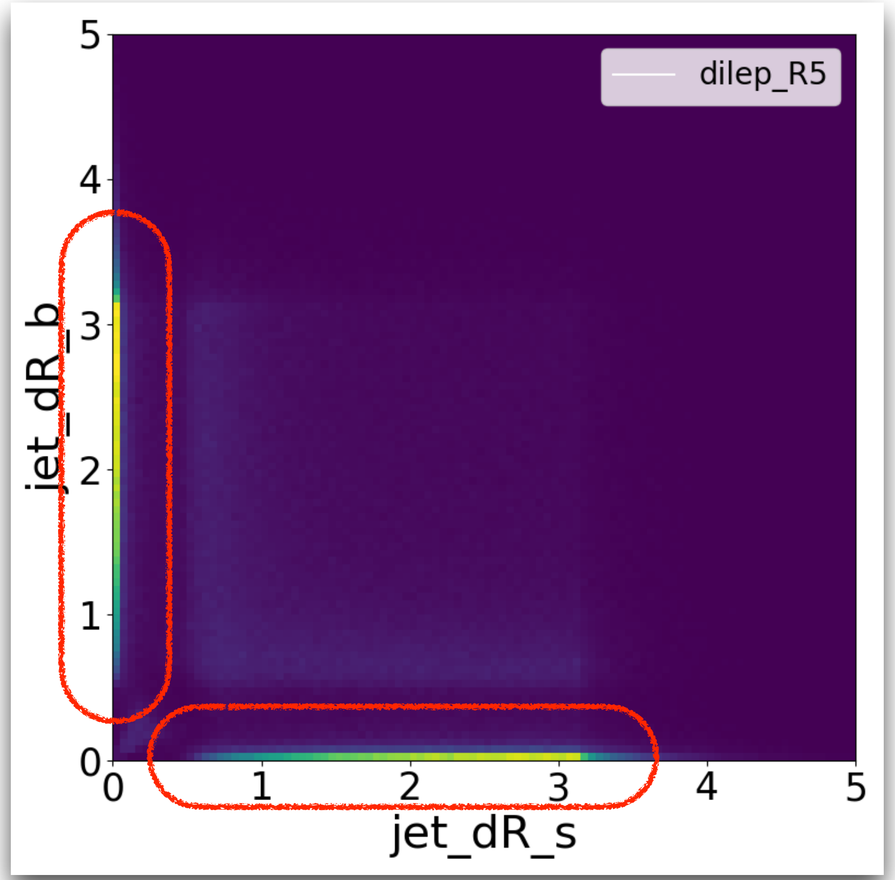
Jet “true flavor” assignment

- Can we get a clean collection of b- and s- jets?
- What about other jets with other flavors?

Spatial separation of flavored jets

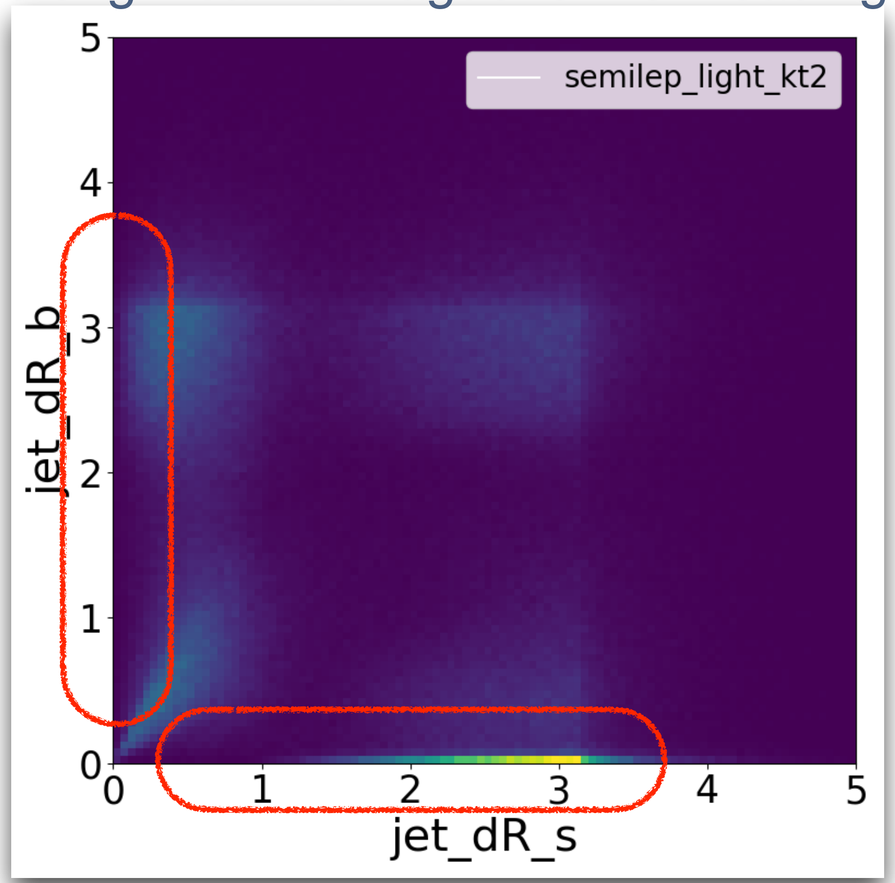
- In **dileptonic** events, compare the proximity between jets and ME quarks (before showering)
  - The true quark origin is well-defined for almost all jets

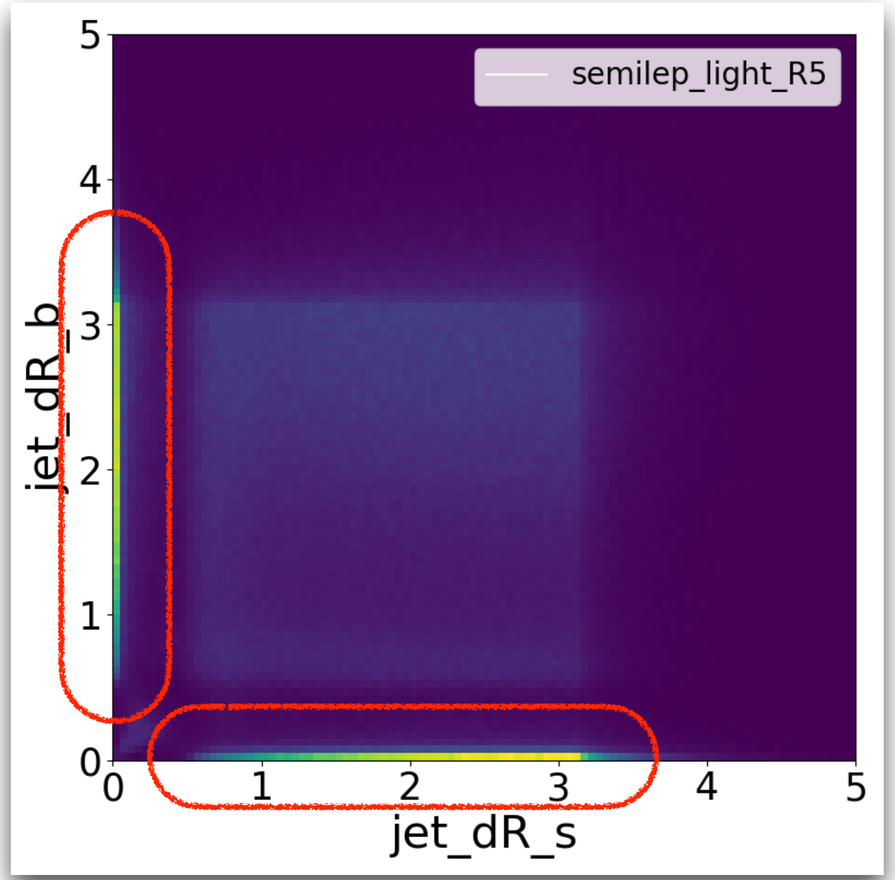




Spatial separation of flavored jets

- In **semilep** $W \to ud$ events, compare the proximity between jets and ME quarks (before showering)
  - kt2 is no longer valid. R5 is rather stable.
  - True origin can be vague for non-leading jets





Jet definition

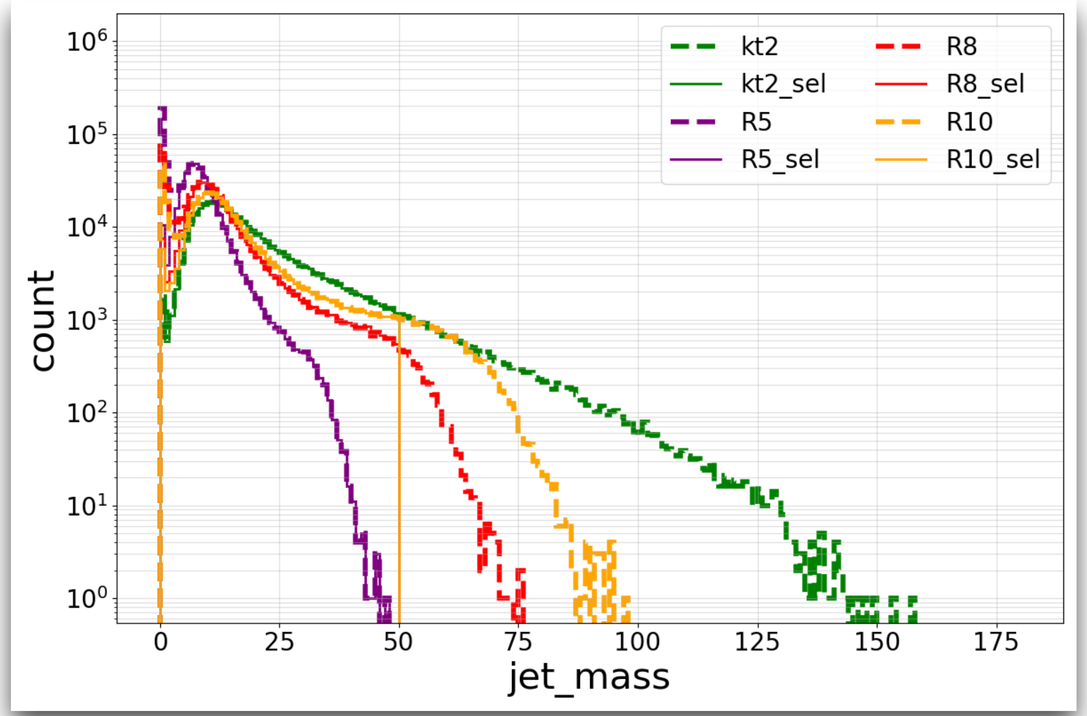
- In this study, flavor tagging is more important than the jet energy calibration —> prefer inclusive jet reco over exclusive
- What jet radius and energy threshold to use?

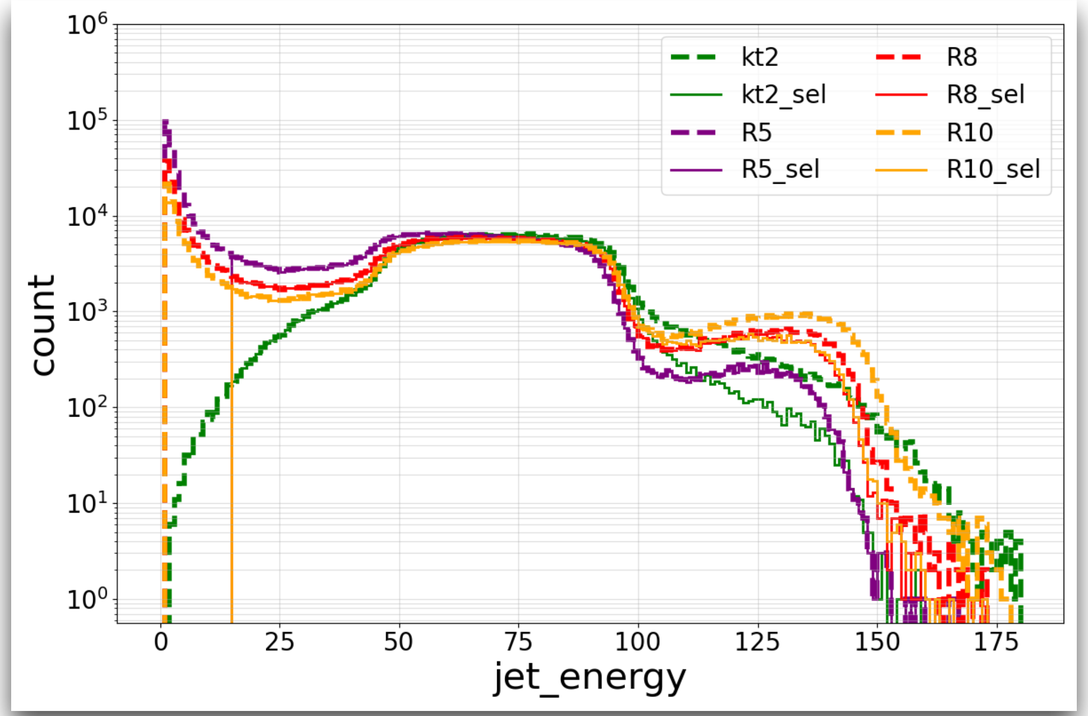
Jet selection

- In **dileptonic** events, the kt2 is the most correct
- A lot of low energy jets in the inclusive jet collections
- After jet selection, profiles of the inclusive jets look similar to the kt2 jets

jet sel:

$m(j) < 50$ GeV and
$E(j) > 15$ GeV



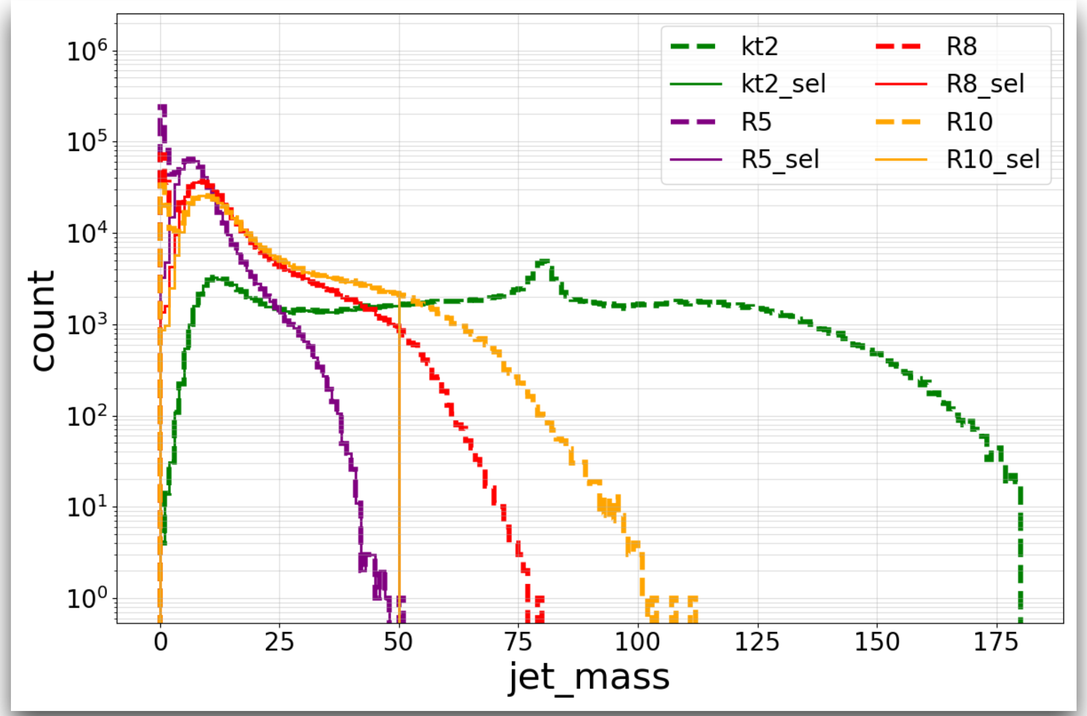


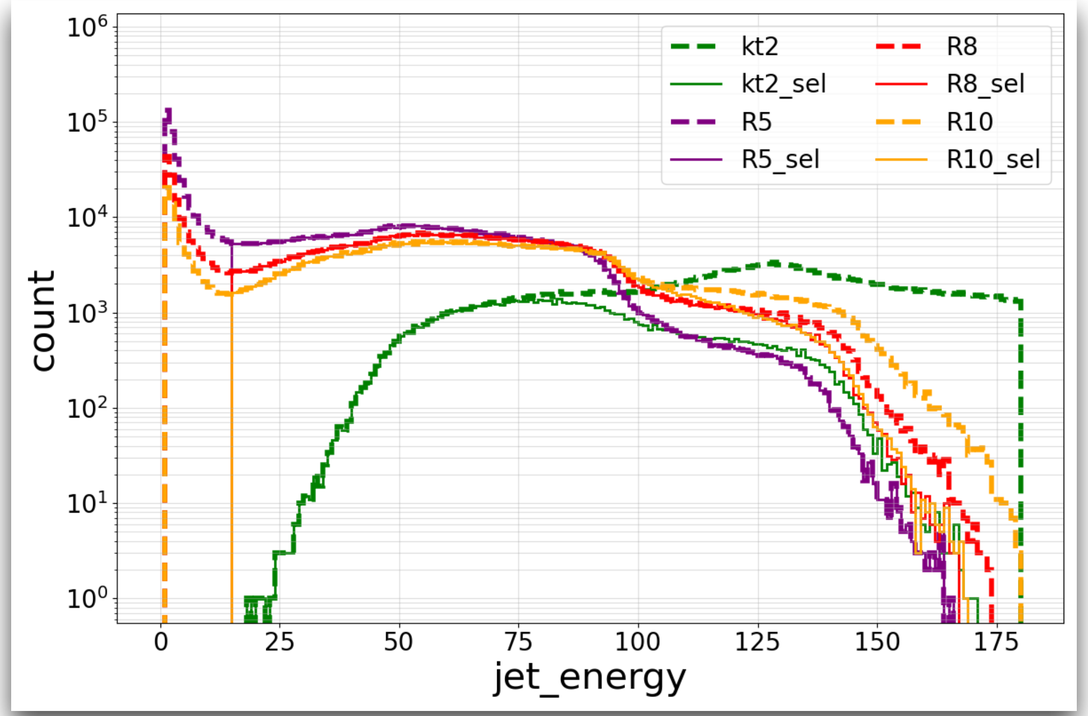
Jet selection

- In **semileptonic Wcs** events, the kt2 biased
- Inclusive clustering gives stable performance
- Kinematic profiles with different jet radius definitions are somewhat consistent

jet sel:

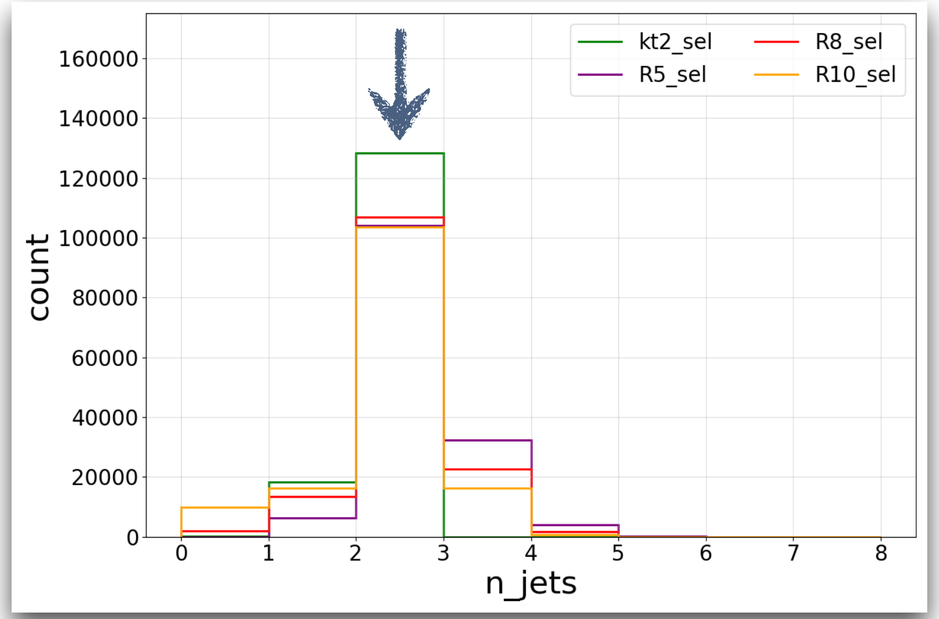
$m(j) < 50$ GeV and
$E(j) > 15$ GeV



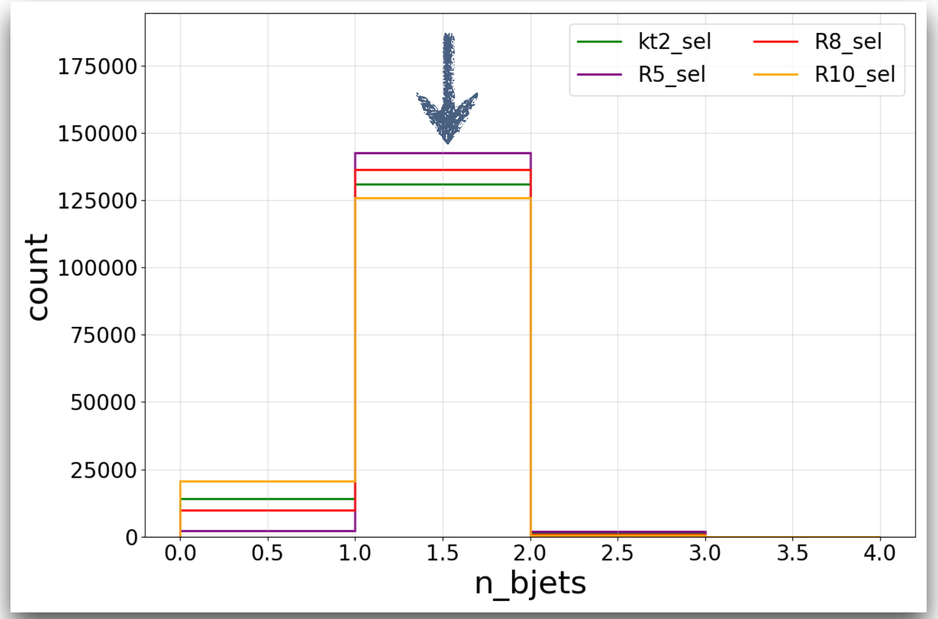


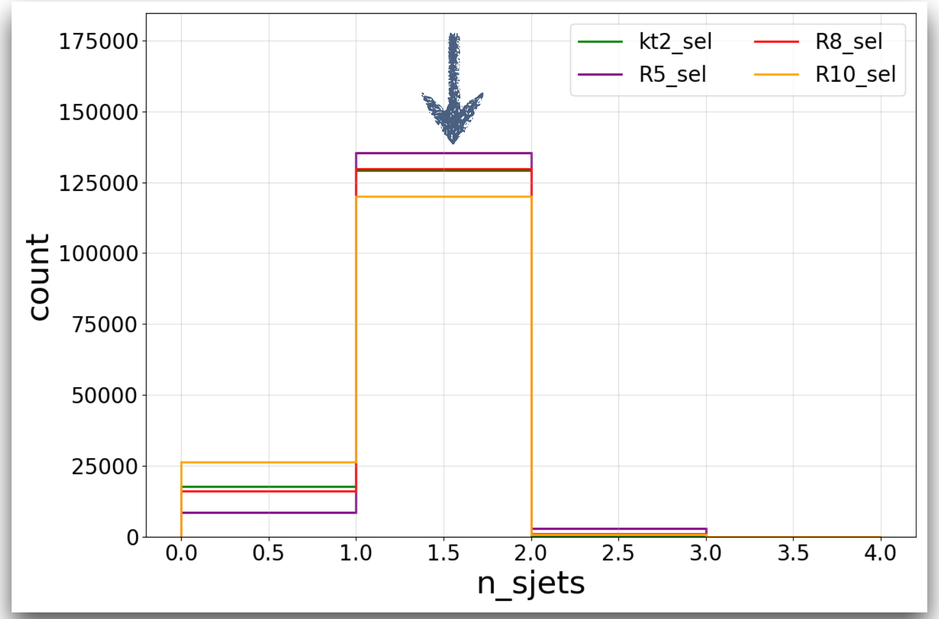
Clustering perf - **dileptonic**

- Check how well the reco jets reflects the true parton number and flavor (prior to any tagging algo)
  - “true flavor” is assigned as the heaviest parton (after showering) flavor within R=0.3 cone of the jet.
- With **dileptonic** events,
  - expect njets = 2, n_bjets=1, n_sjets = 1



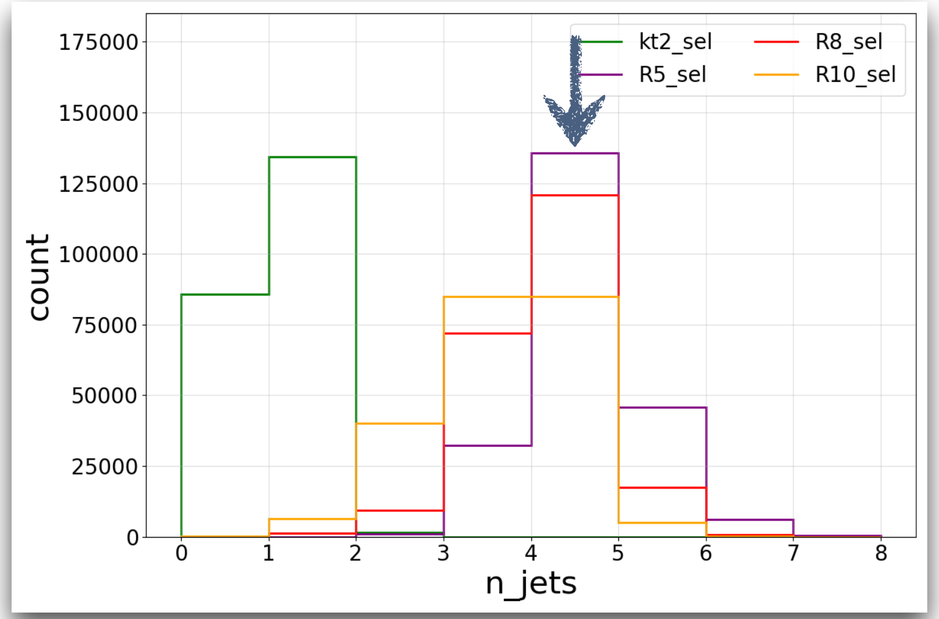
Clustering with **R5** is more correct than **R8** and **R10**



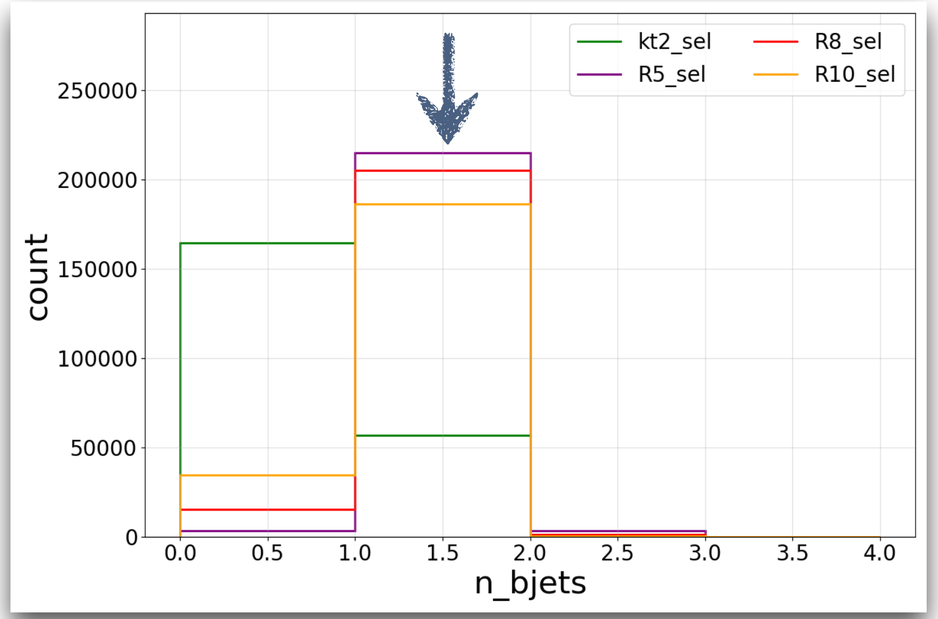


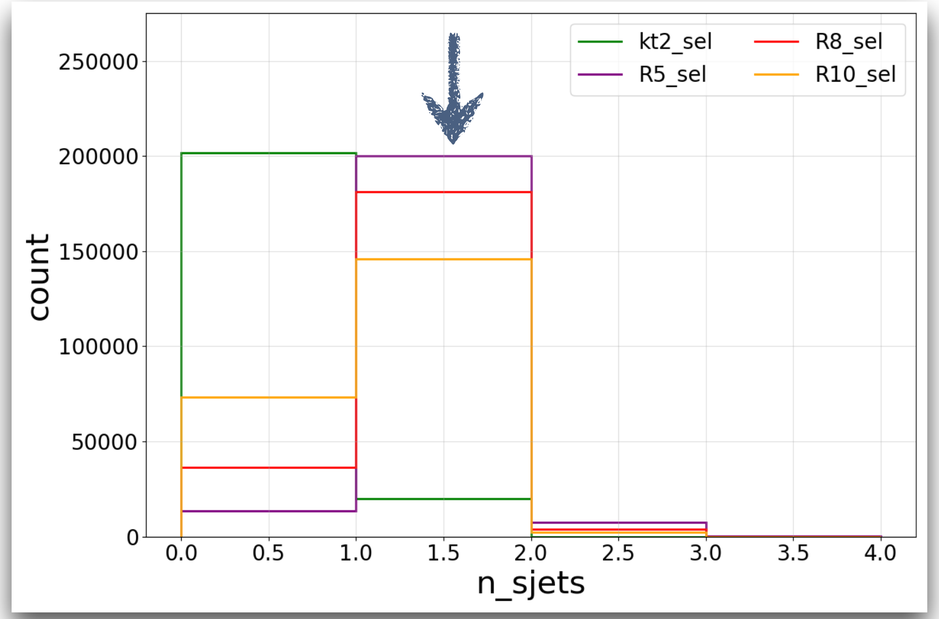
Clustering perf - **semilep Wud**

- Check how well the reco jets reflects the true parton number and flavor (prior to any tagging algo)
  - “true flavor” is assigned as the heaviest parton (after showering) flavor within R=0.3 cone of the jet.
- With **semileptonic Wud** events,
  - expect njets = 4, n_bjets=1, n_sjets = 1



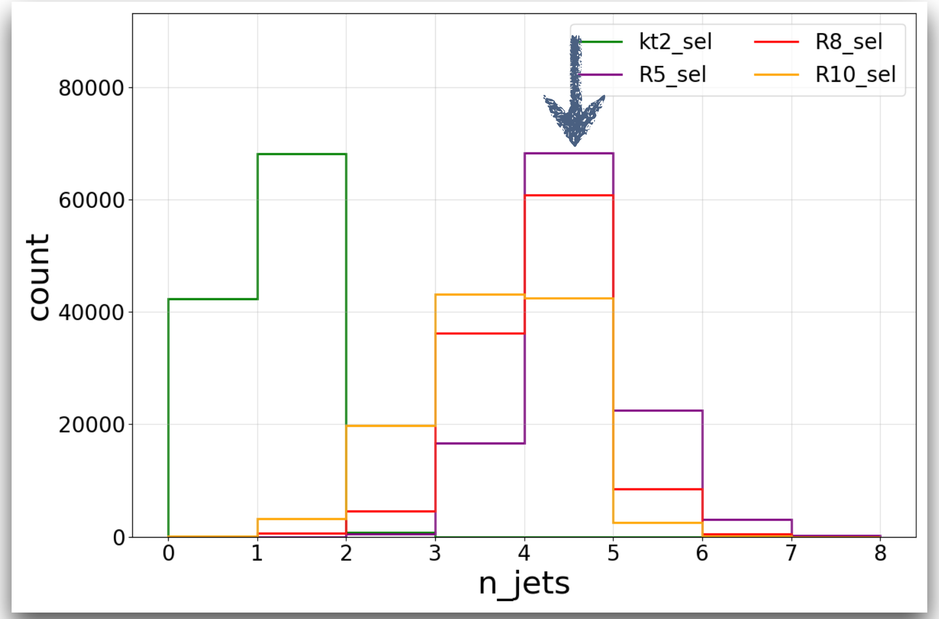
Clustering with **R5** is more correct than **R8** and **R10**



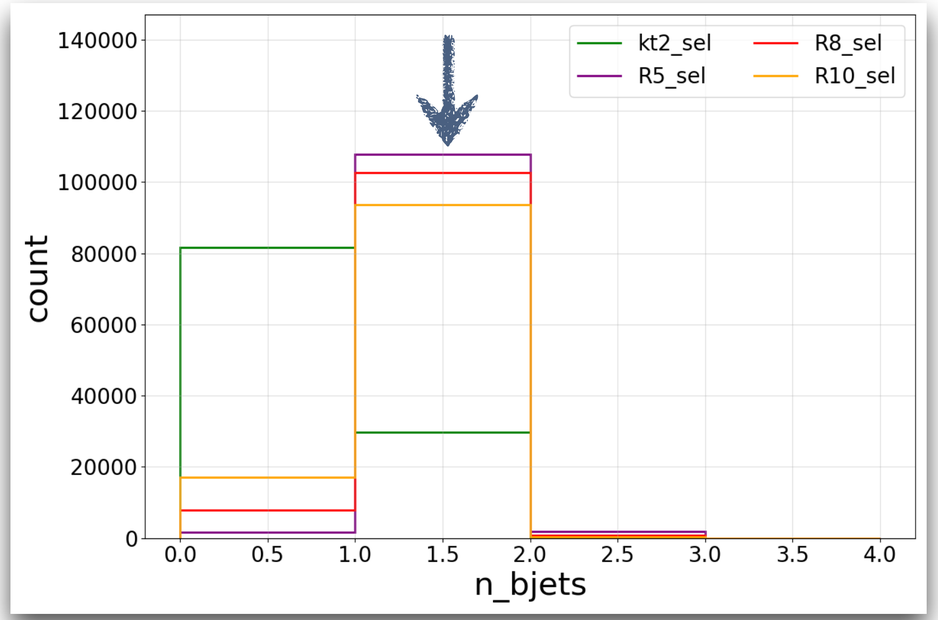


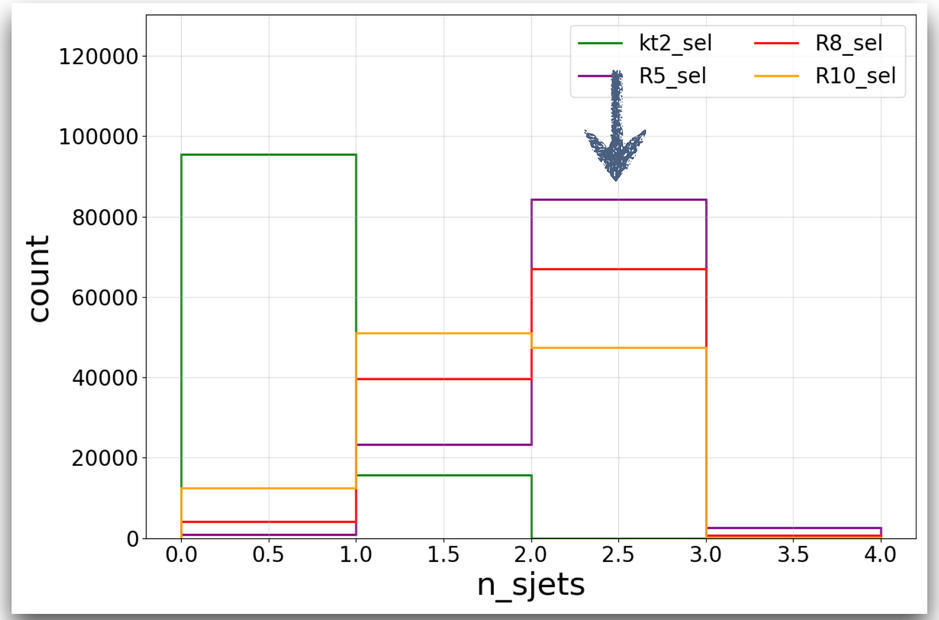
Clustering perf - **semilep Wcs**

- Check how well the reco jets reflects the true parton number and flavor (prior to any tagging algo)
  - “true flavor” is assigned as the heaviest parton (after showering) flavor within R=0.3 cone of the jet.
- With **semileptonic Wcs** events,
  - expect njets = 4, n_bjets=1, n_sjets = 2



Clustering with **R5** is more correct than **R8** and **R10**





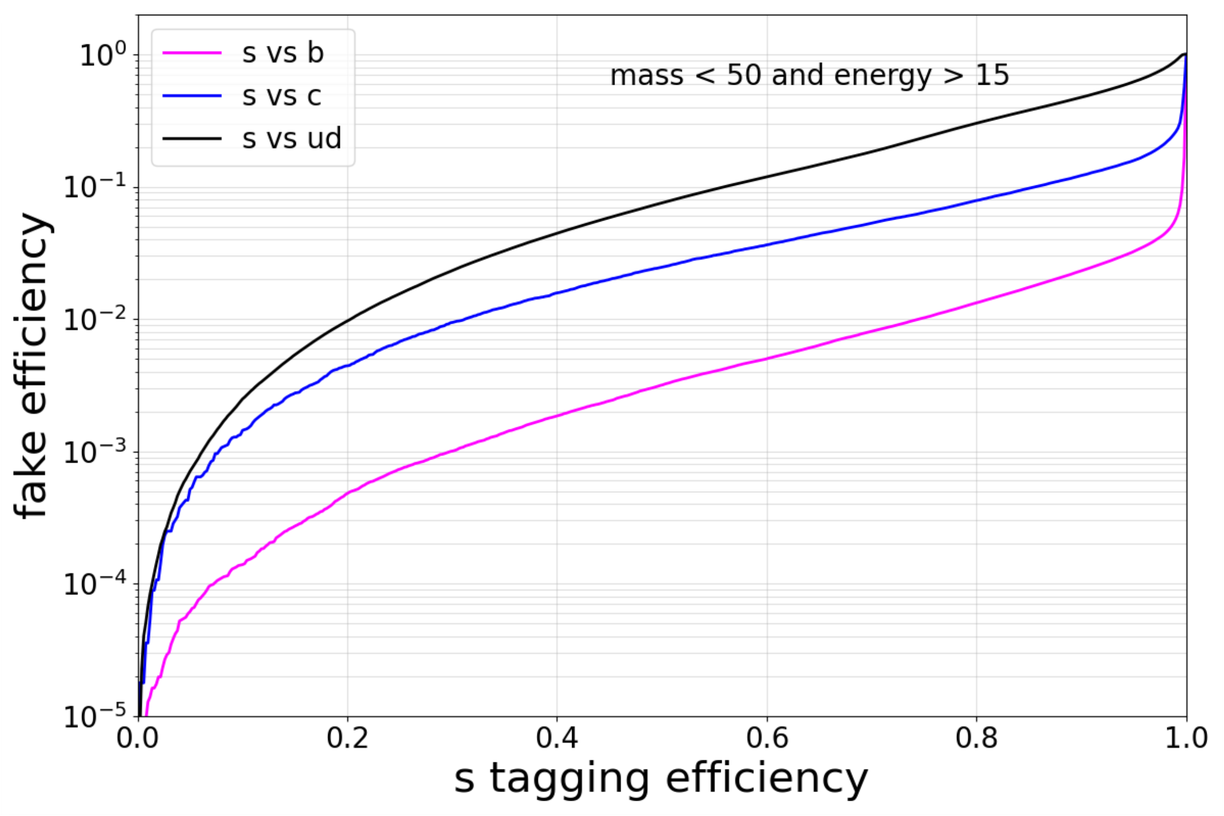
tWs event selection

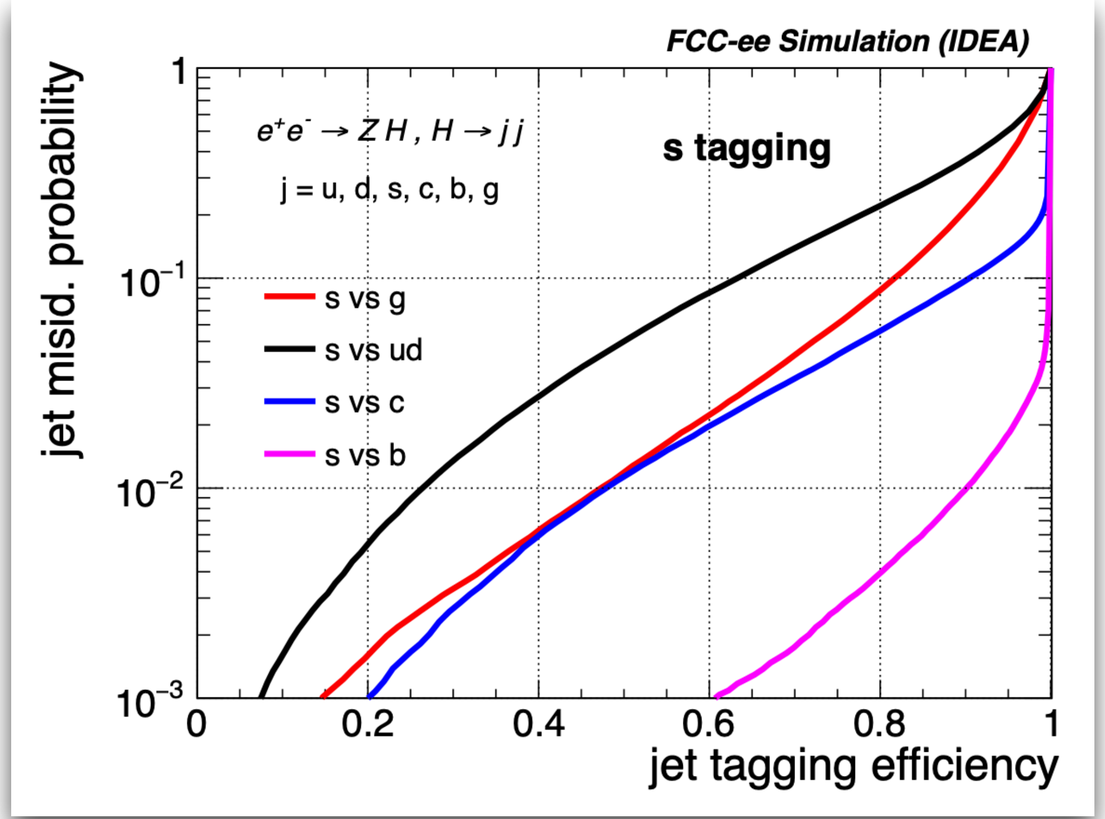
- Proceed with R5 jets (mass < 50 GeV, energy > 15 GeV) for the analysis
- Simple selections to get a sense of the signal yield and background composition

s-tagging performance

- As a reference, below is the tagging performance by applying the standard training to the R5 jets defined and selected as described in previous slides

arxiv 2202.03285





Event selection

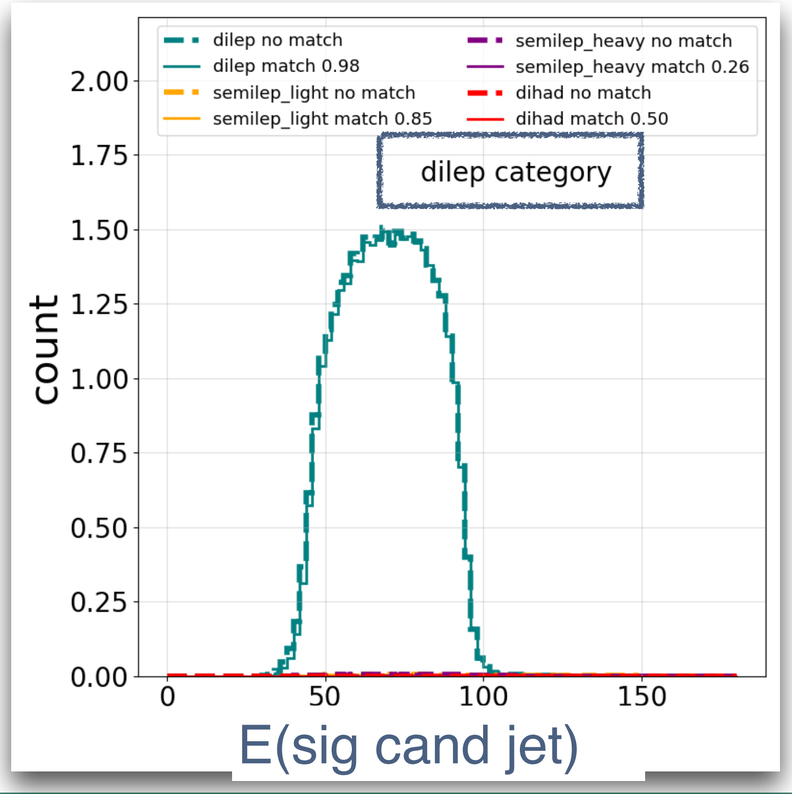
- Exactly 1 b-tagged jet (b score > 0.5)
- At least 1 “tightly” s-tagged jet (s score > 0.9)
  - roughly, ϵ(s-jet) ~ 0.2, ϵ(b-jet) ~ 5E-4, ϵ(c-jet) ~ 4E-3, ϵ(ud-jet) ~ 1E-2
- Events further classified into “dilep” ($n_J = 2$), “semilep” ($n_J = 4$), “dihad” ($n_J = 6$) categories purely based on the number of selected jets.
  - Even without lepton information, the category assignment is quite accurate.

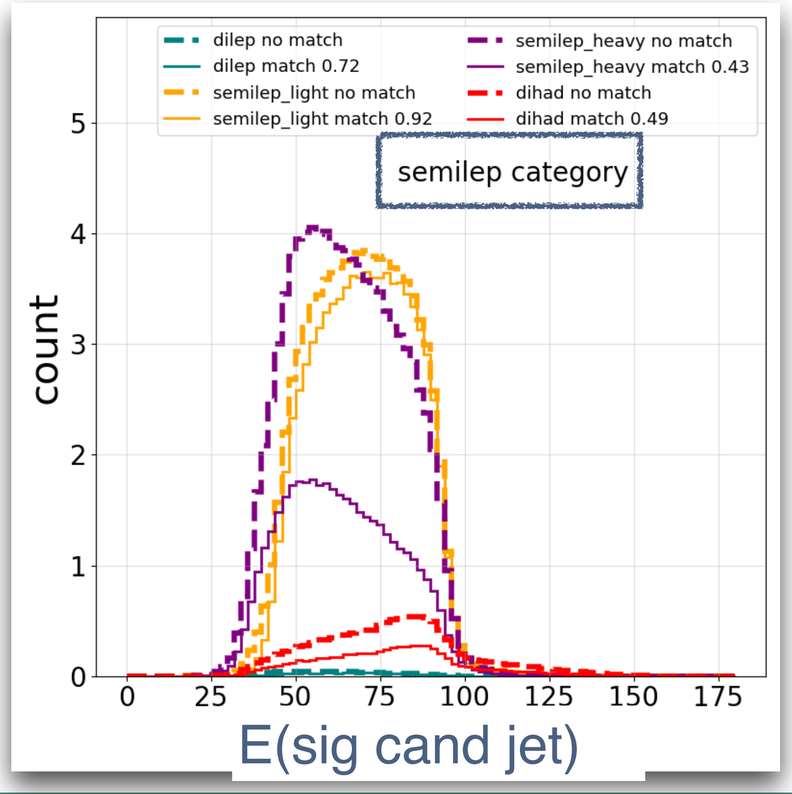
- Backgrounds: consider all decay channels of $t\bar{t} \rightarrow WbWb$
- Signals: consider all decay channels of $t\bar{t} \rightarrow WsWb$, except $W \rightarrow \tau\nu$

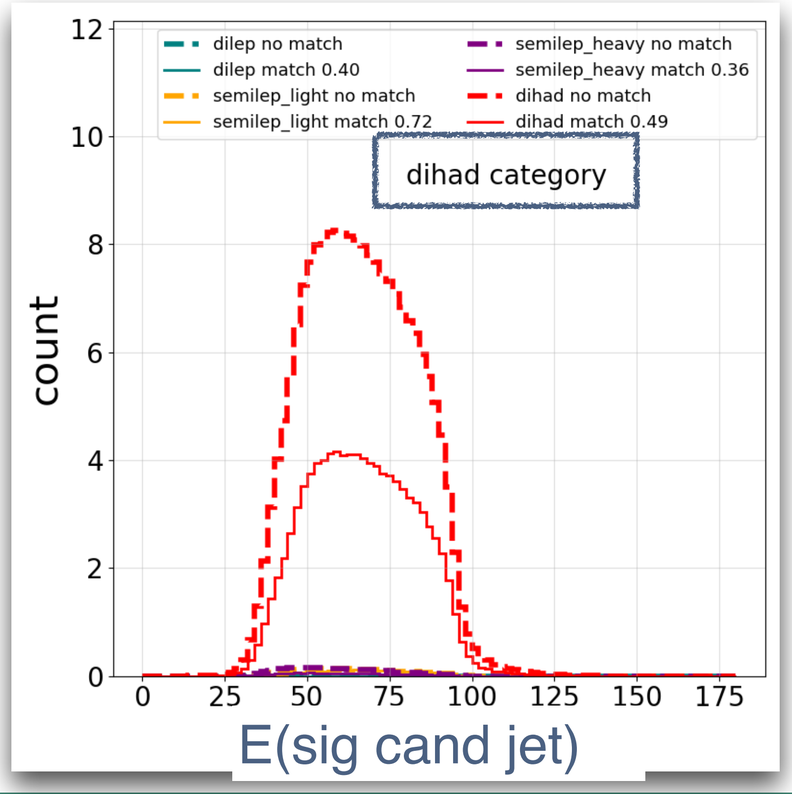
Signal candidate efficiency

Check whether the signal candidate can be matched to the s quark from $t \rightarrow Ws$ decay

- In **dilep** and **semilep Wud** events, signal is correctly chosen most of the time
- In **semilep Wcs** and **dihad** events, the s quark from W decay is a serious contamination

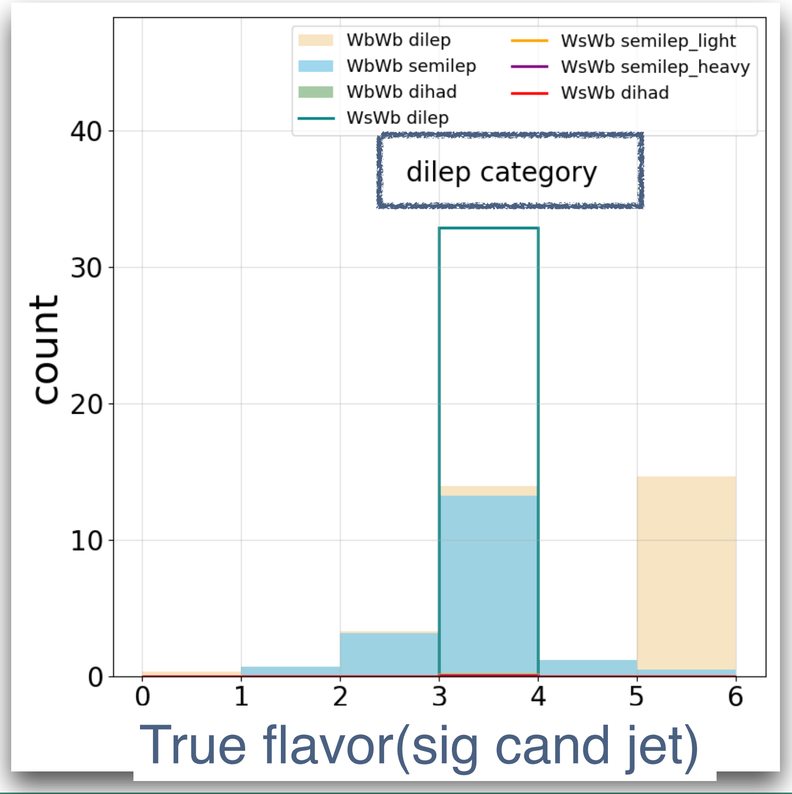


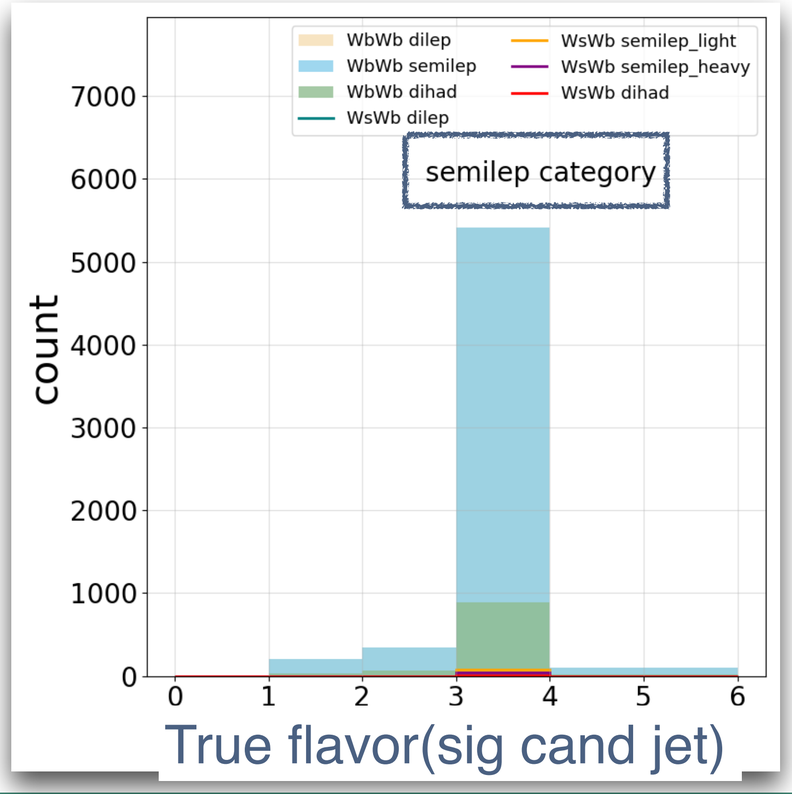


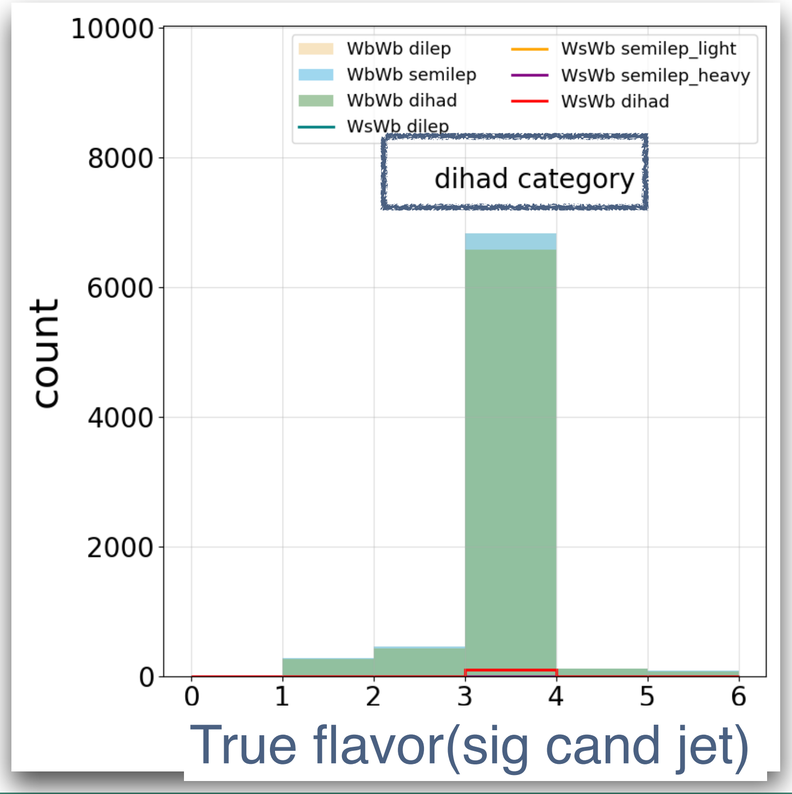


Event yields

- It is possible to reduce the b-jet background yield to the same level as the signal
- The main source of background in hadronic events is $W \rightarrow cs$ decay.
  - Handles in c-tagging and W reconstruction. Need further studies.







Summary

- $t\bar{t} \rightarrow WsWb$ dileptonic events can be a good candle for profiling jet clustering and tagging performance (or for training)
  - Not so straightforward to define a clean collection of u,d,c jets in $t\bar{t}$ events
- For the case of $|V_{ts}|$ measurement, inclusive jet clustering with small jet radius gives better jet flavor assignment
- With simple selections, can achieve S/B ~ 1 against b-jet background
  - Dilep channel is clean —> potential for measurement
  - $W \rightarrow cs$ is the dominant background —> More studies to reduce it

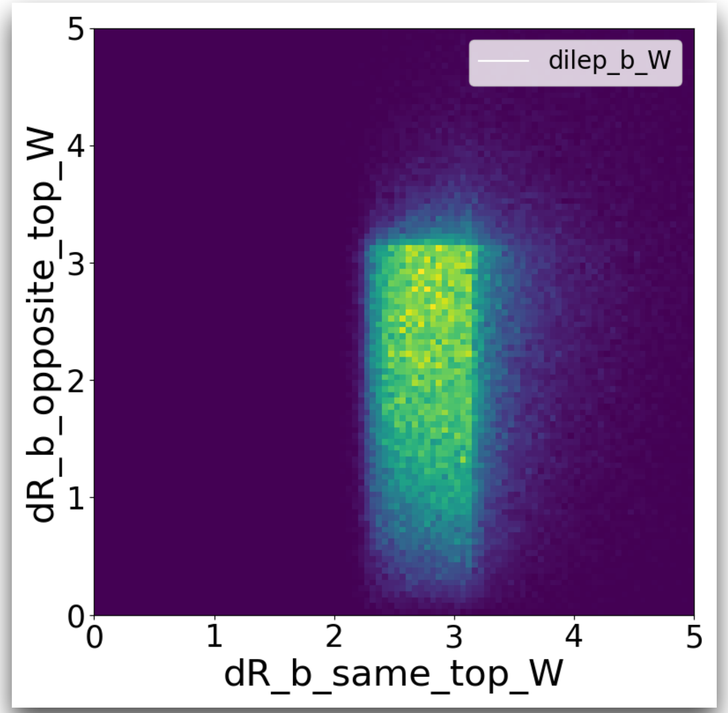
- Caveats: jet tagging uncertainty, and WW and ZZ background
- Another open question: from theory side, what is the potential range of modifications? How much precision is needed for this study to be relevant/interesting?

Backups

Spatial separation of quarks

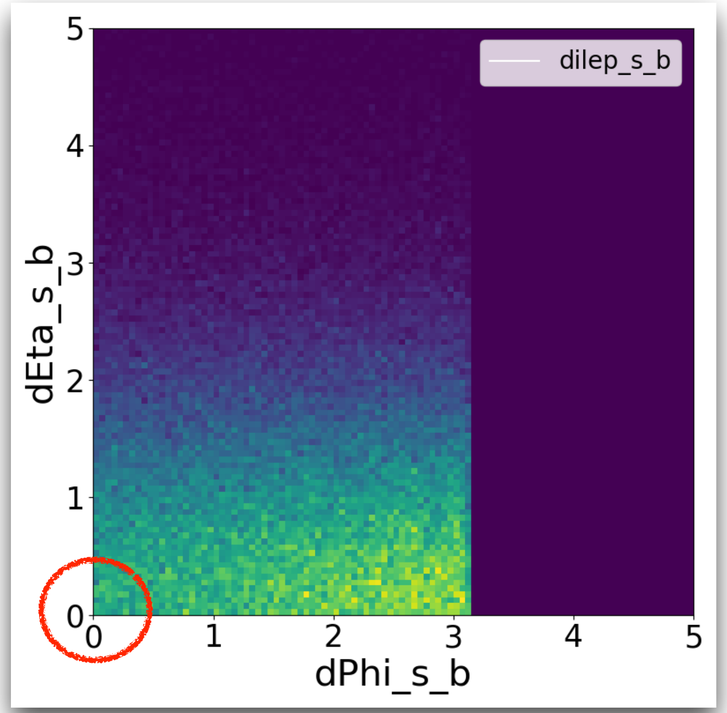
- Using **dileptonic** events, with exactly 1 b quark and 1 s quark and no other quark (or tau).
- Looking at distances between quark momenta at ME level (b and s from top decays, before showering)
- In a small fraction of events, the b and s quarks may be too close and be mixing in one jet

$\Delta R(b, W)$ from different top decays



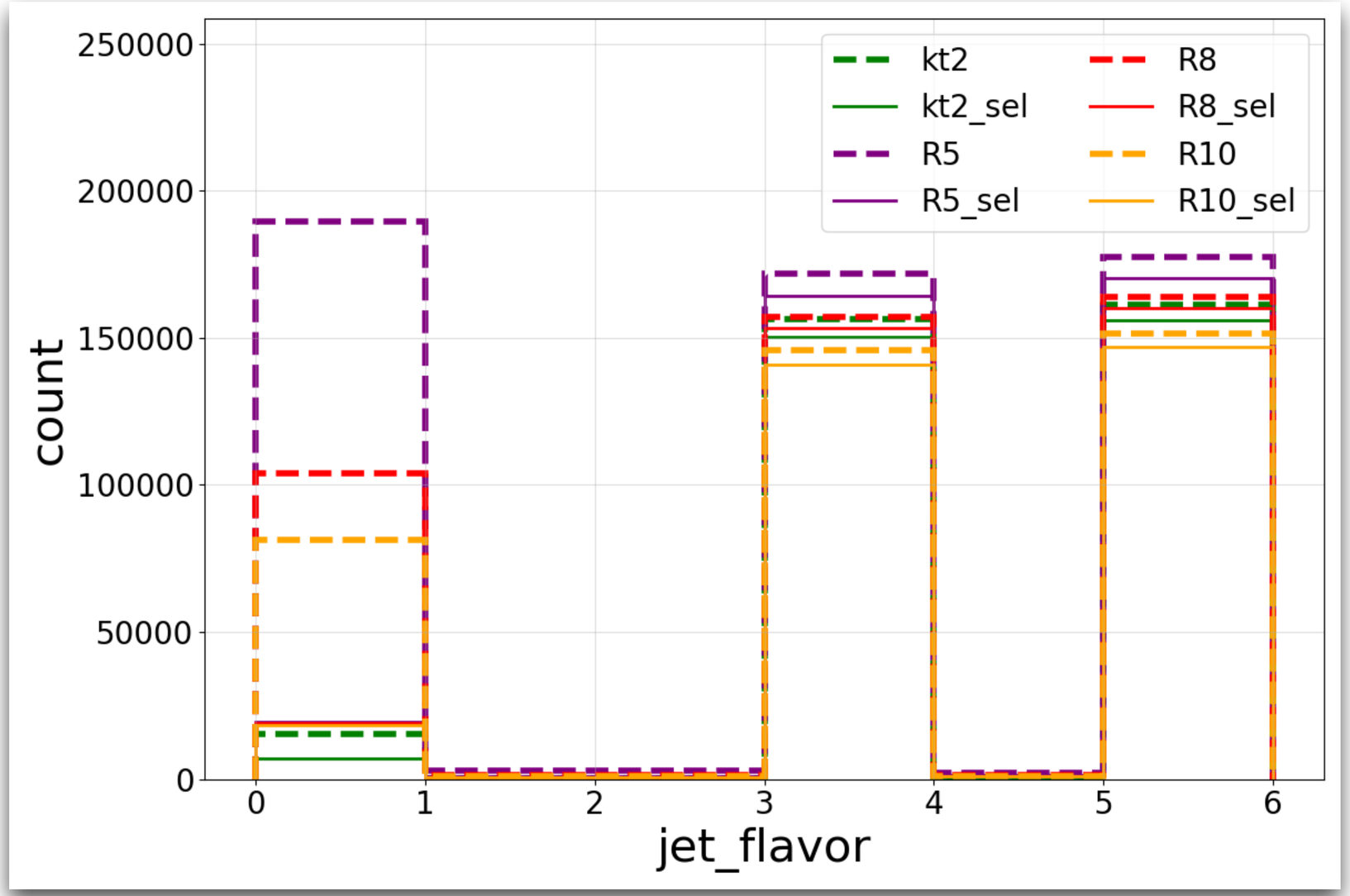
$\Delta R(b, W)$ from the same top decay

$\Delta \eta(b, s)$

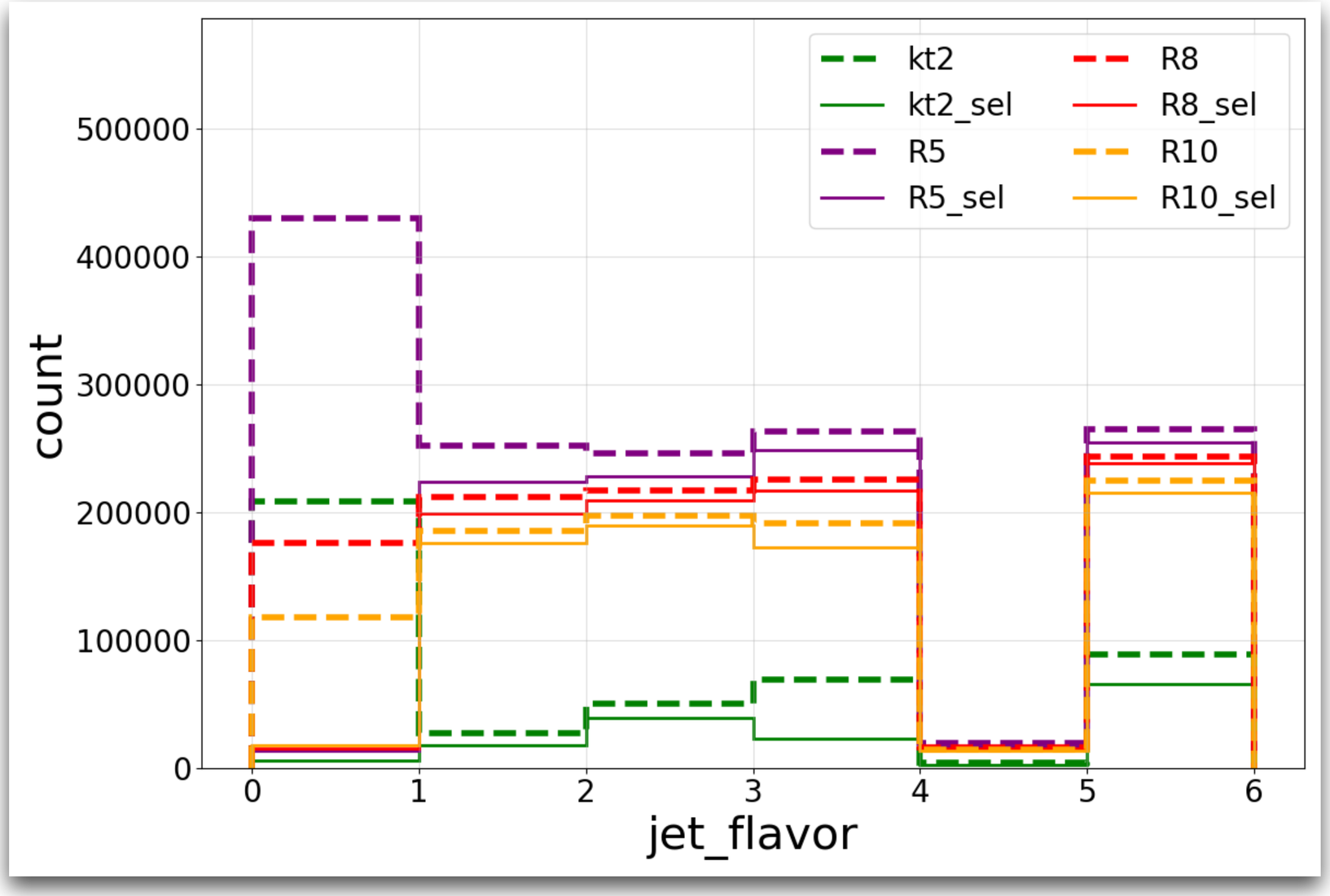


$\Delta \phi(b, s)$

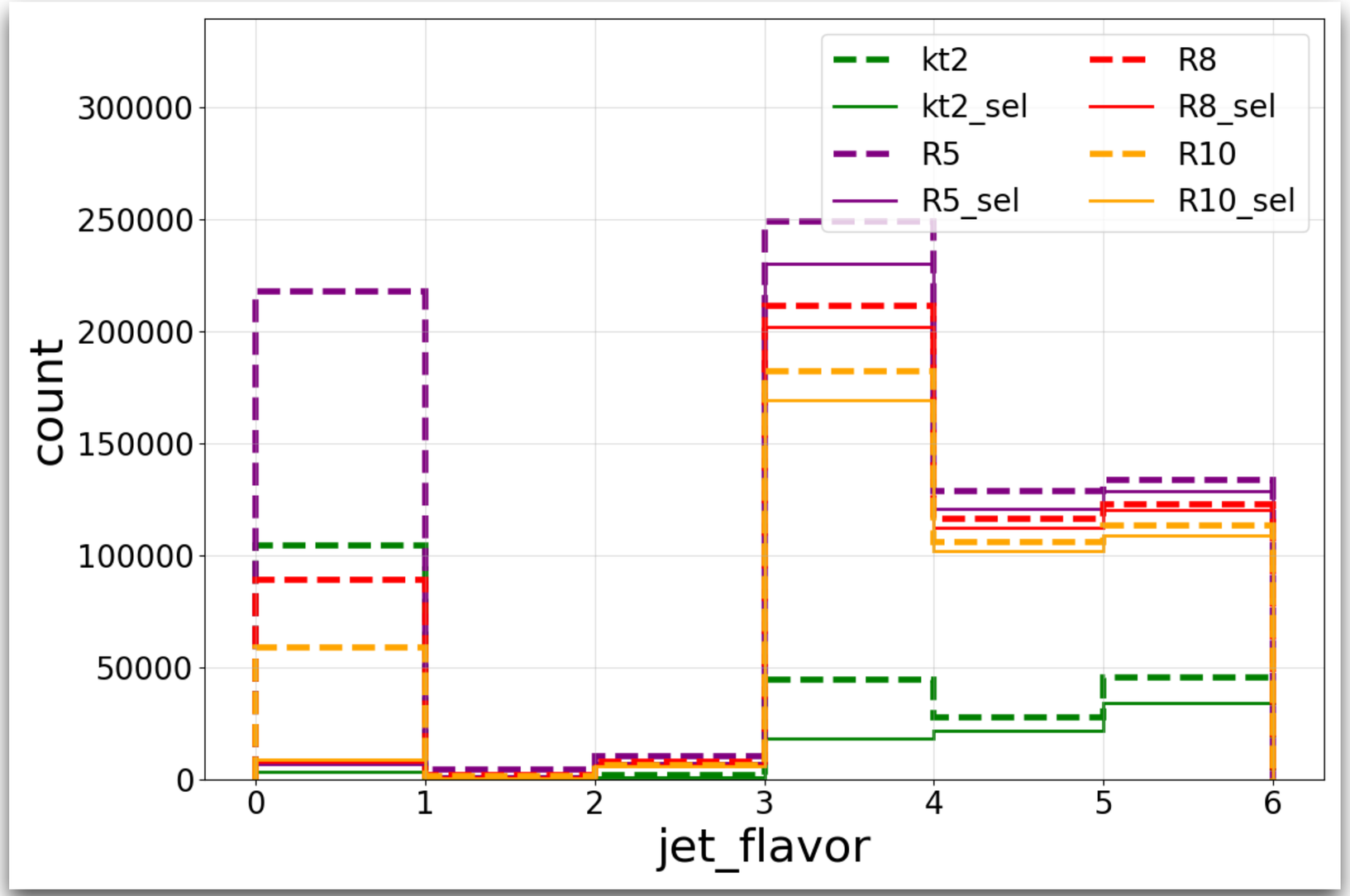
Clustering perf - **dileptonic**



Clustering perf - **semilep Wud**



Clustering perf - **semilep Wcs**



Event yields

- It is possible to reduce the b-jet background yield to the same level as the signal
- The main source of background in hadronic events is $W \to cs$ decay.
  - Handles in c-tagging and W reconstruction. Need further studies.