Techniques for hadronic decays of ultra-heavy resonances decaying to vector-like quark pairs

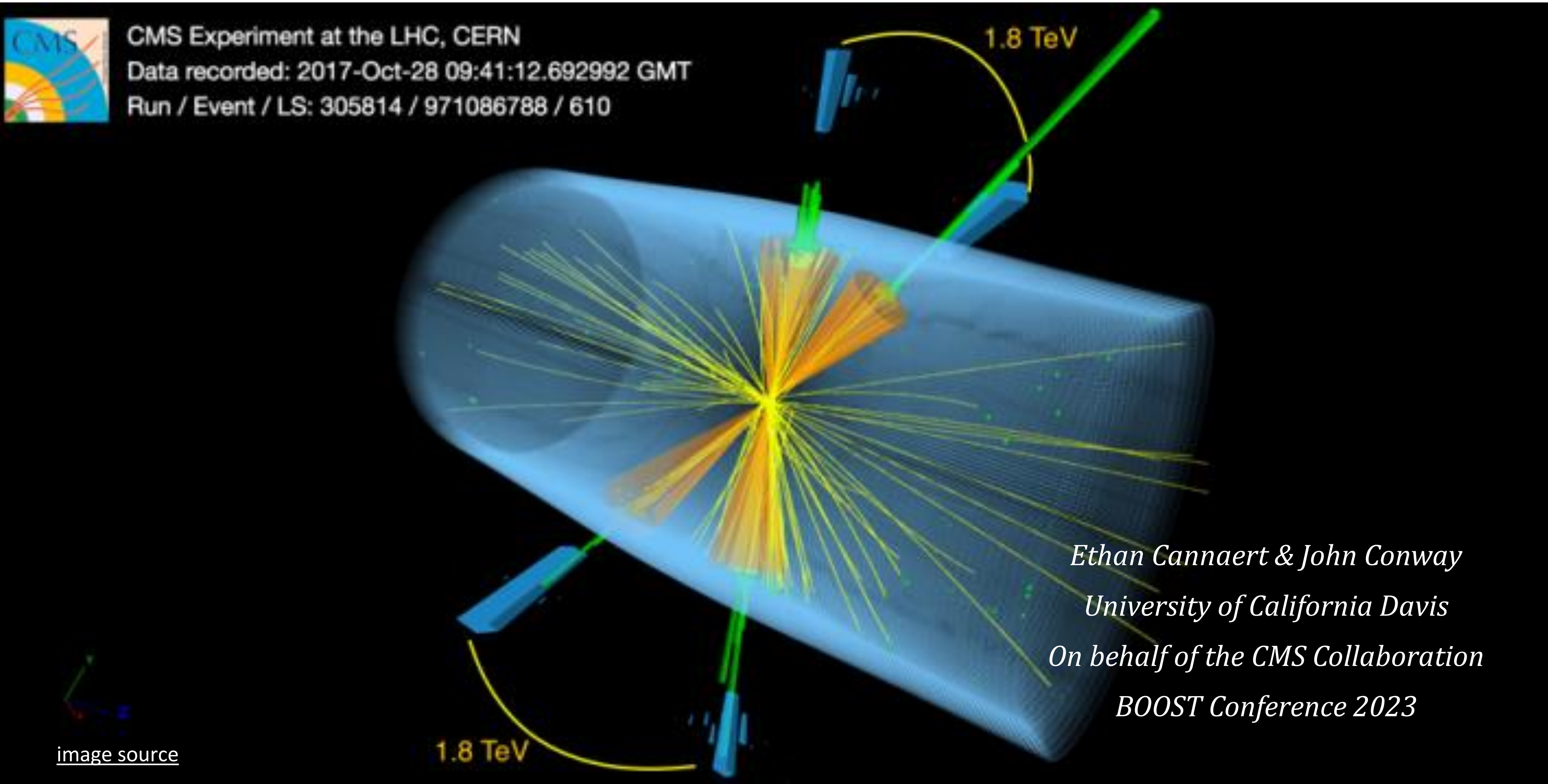


image source

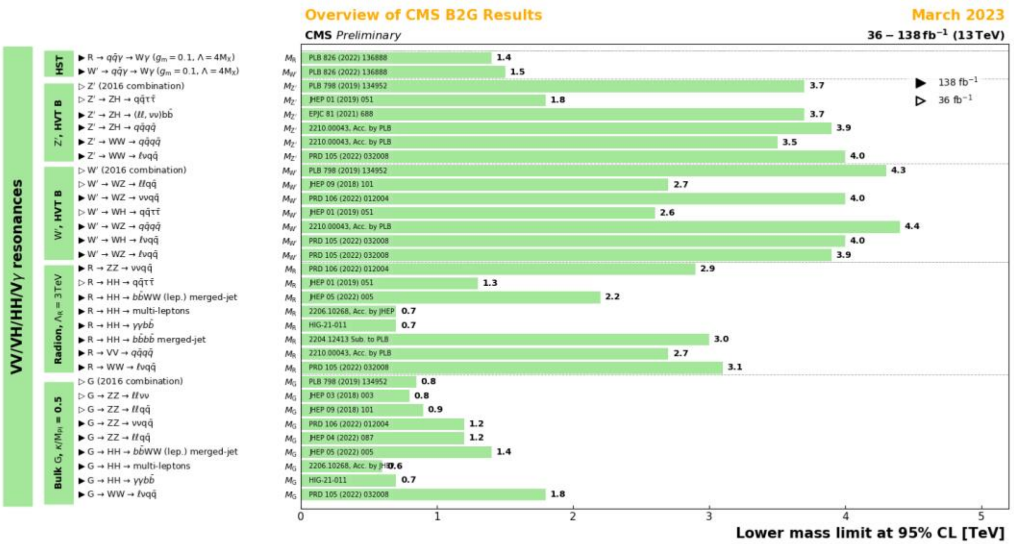
Ethan Cannaert & John Conway

University of California Davis

On behalf of the CMS Collaboration

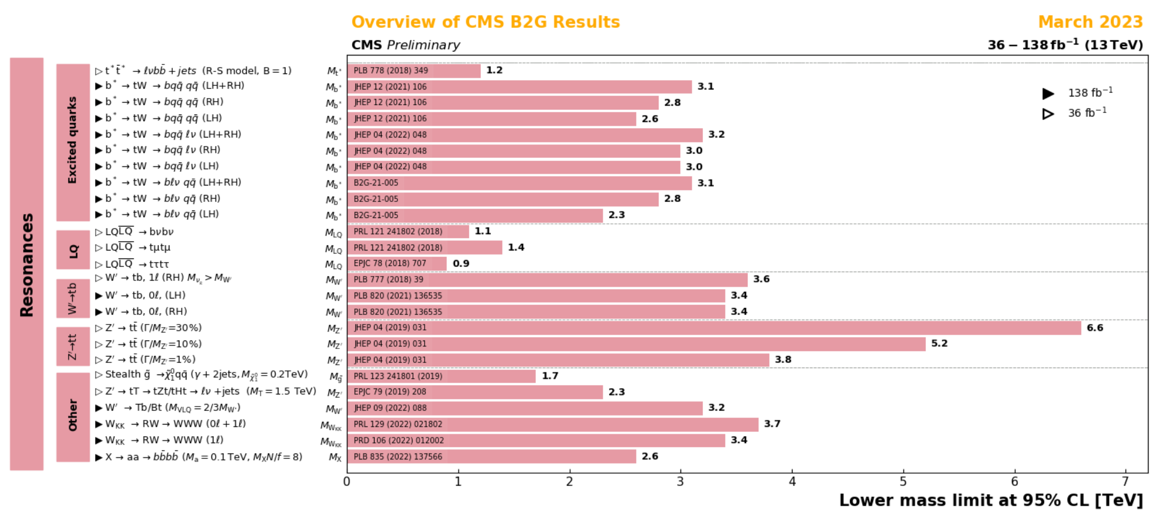
BOOST Conference 2023

Resonance Searches in Run 2



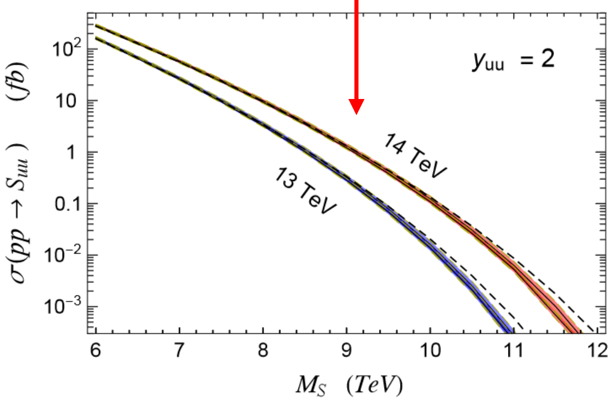
- Run 2 CMS analyses extended our limits on many resonance searches up to several TeV and beyond
- Much of the possible phase space has been covered
- [Check out more CMS B2G analyses](#)!

- Is new physics hiding in plain sight or on the horizon?
- What is the limit of what can be discovered at the LHC?
  - Run 2? Run 3?



Resonances in Run 3

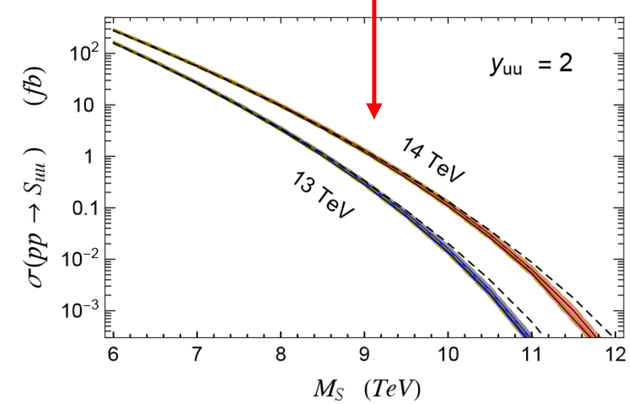
- With Run 3 underway, it's an exciting time to be studying ultra-heavy resonances!
- Increased luminosity and ~4.5% increase in energy that significantly enhances many of these processes



Exploring the laws of nature with CMS, Dobrescu 2020

Resonances in Run 3

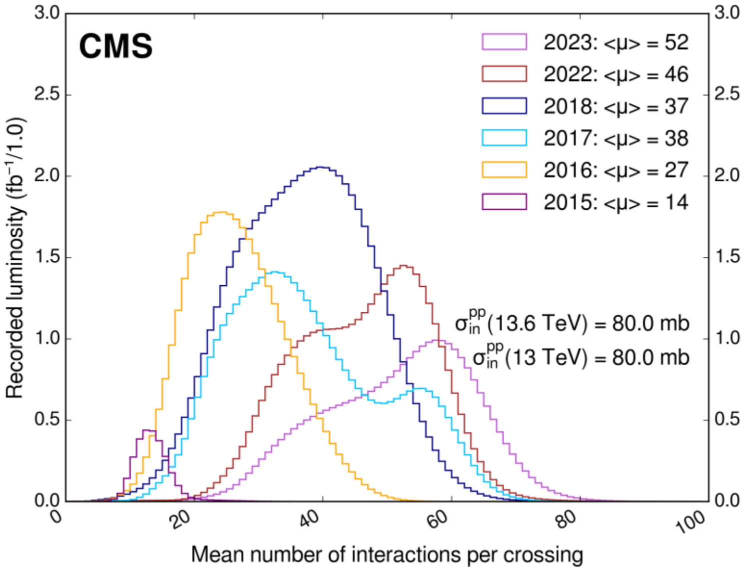
- With Run 3 underway, it's an exciting time to be studying ultra-heavy resonances!
- Increased luminosity and ~4.5% increase in energy that significantly enhances many of these processes



Exploring the laws of nature with CMS, Dobrescu 2020

Challenges in Run 3

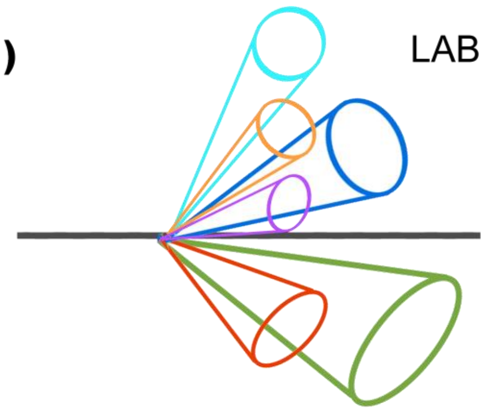
- Increased pileup → increased jet multiplicity that complicates jet sorting
- Increased energy → more highly boosted events with many overlapping jets
- Conventional techniques often break down at high energies
    - e.g. b-jet identification loses efficiency at high p_T
- If you are looking at fully hadronic channels, these are all challenges that need to be considered
- On top of this, there is often a large phase space to cover
- New techniques must be flexible to various kinematics yet robust enough to function at high energy



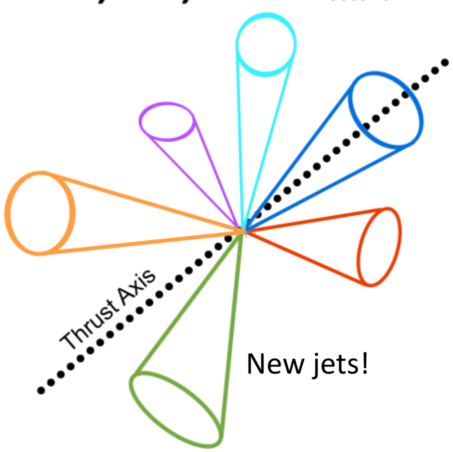
Above: Distribution of the average number of interactions per crossing (pileup) for pp collisions at the LHC for years 2011-2012, 2015-2018 and 2022-2023. Source: Public CMS Luminosity Results, The CMS Collaboration, 2023, link to luminosity results

- We have been working on a new technique to study complex, pair-produced processes in Run 2
- **Goal**
  - Use event geometry to reconstruct hadronic decays of heavy pair produced signals
  - Maintain maximum flexibility to account for many models and levels of jet merging
    - Applicable to any pair-production process
- **Algorithm Steps**
  1. Select high H_T (> 1.5 TeV) events with many jets
     - How loosely will jets be selected?
  2. Boost all jet particles along their axis to the frame that minimizes parallel momentum (MPP frame)
  3. Recluster particles into new AK8 or CA8 MPP jets
     - This resolves jets that were merged in the lab frame and gives the opportunity to shed soft radiation that gets clustered into the original jets

1)



2) + 3)



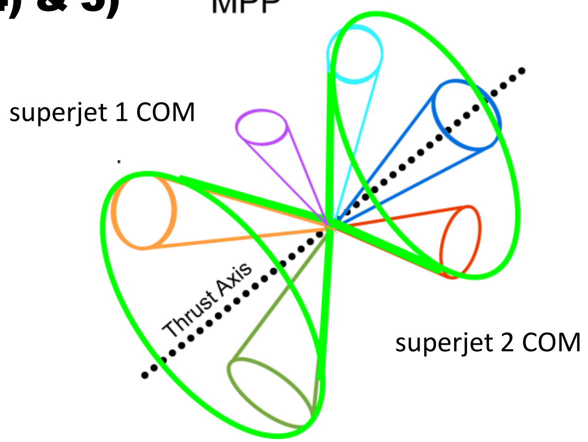
(continued)

Sorting Technique Continued

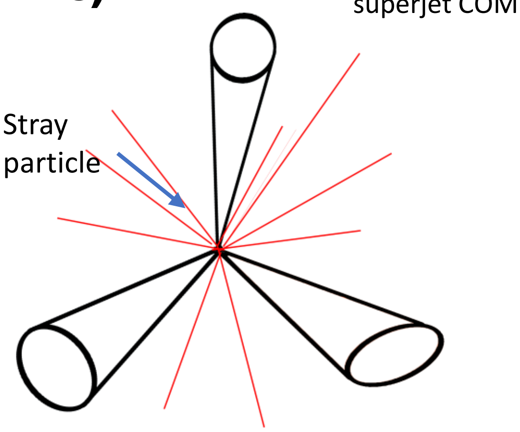
$$\text{*Thrust} = max_{|n|=1}\frac{\sum_i |n \bullet p_i|}{\sum_i |p_i|}, \text{[9]}$$

4. Calculate the MPP thrust* axis.
   - This should split the event into hemispheres
5. Sort jets into two "superjets" from their angles relative to the thrust axis.
   - Nearly perpendicular jets sorted to balance superjet mass
6. Boost superjets to their COM & recluster particles into smaller (AK4 or CA4) jets

4) & 5)



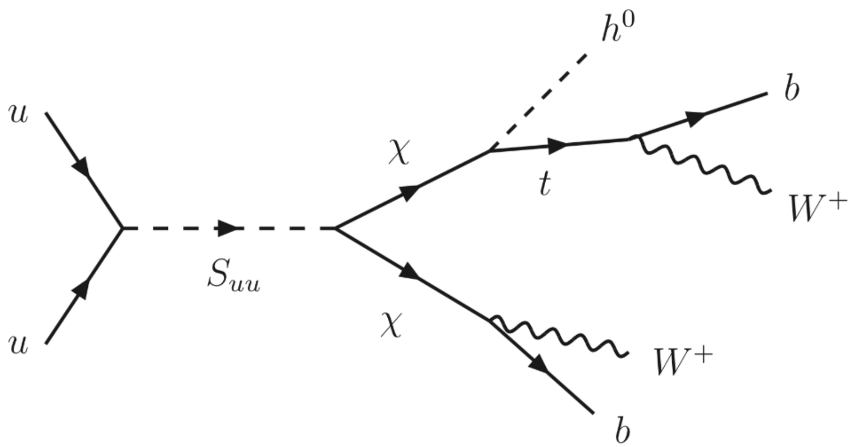
6)

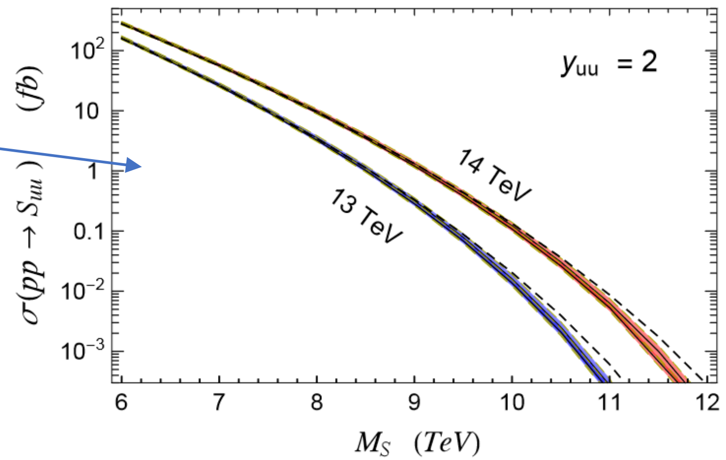


- If everything went well, you'll have two sorted superjets representing your daughter pair
- Looking at substructure is a powerful tool for discriminating against backgrounds
  - MPP frame event shape
  - superjet COM event shapes

Our Pet Process - $S_{uu} \to \chi\ \chi$

- The $\boldsymbol{S_{uu}} \to \boldsymbol{\chi\ \chi}$ process provides a fantastic opportunity to put this technique into practice
- $\boldsymbol{S_{uu}}$ - diquark, ultra-heavy resonance, couples to up quarks
- **χ** – vector-like quark
  - decays to W^+b, Z t, h t
- Relatively large σ at LHC means there is potential for Run 2 discoveries up to $M_{S_{uu}}$= 8 TeV
- Many complex hadronic final states with different levels of jet merging
- Main backgrounds – QCD multijet & top pair production





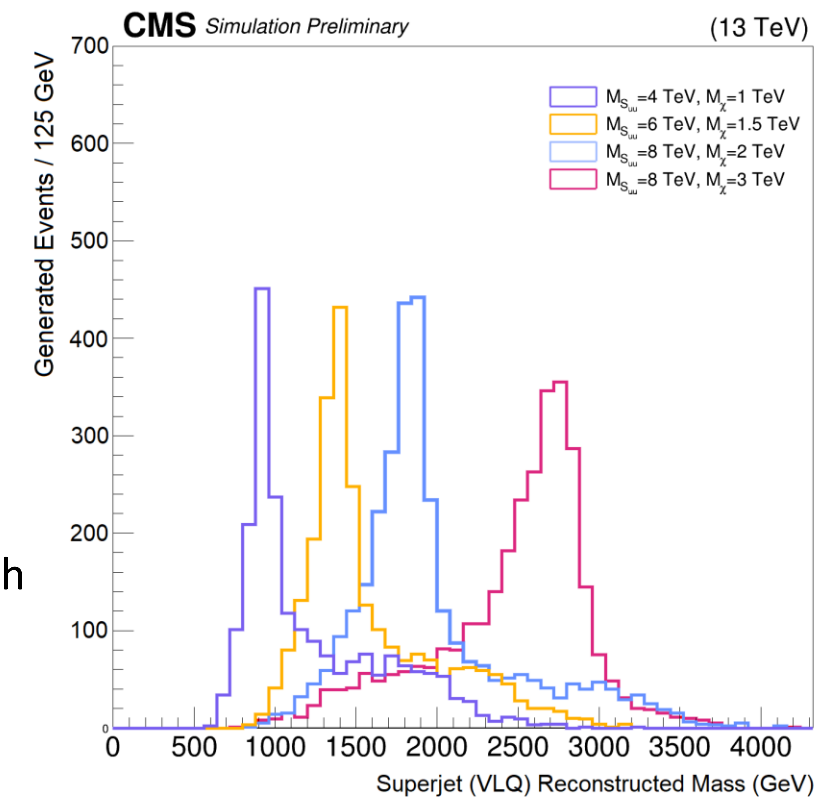
* $S_{uu} \to \chi\ \chi$ model source: Bogdan Dobrescu, Robert Harris and Joshua Isaacson, FNAL

Sorting Algorithm on MC

- How well does this work with reconstructed jets?
- Define some baseline cuts to target the hadronic channel

<u>Initial Selection</u>

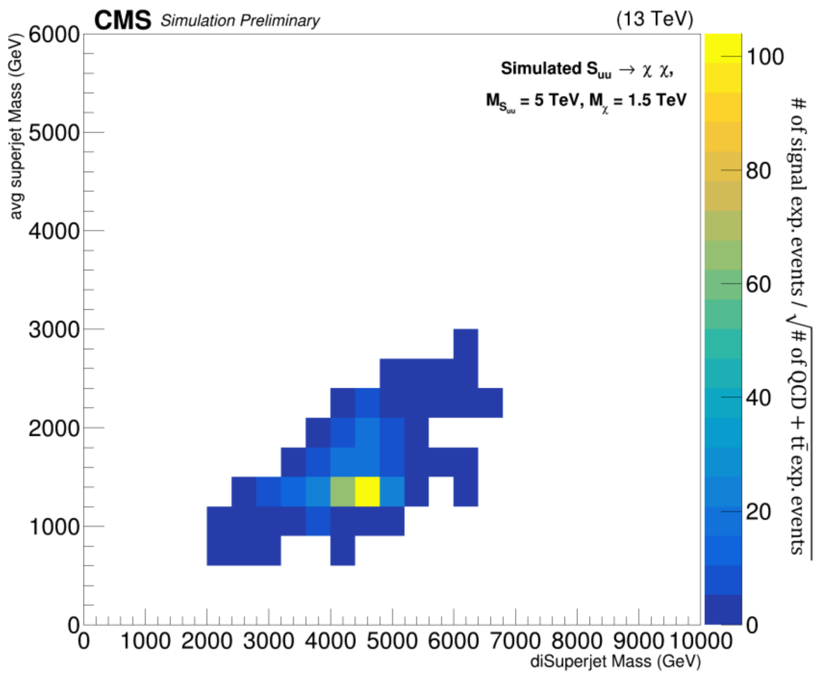
- Lepton veto
- Event H_T > 1500 GeV
- 3+ AK8 jets
- 2+ heavy AK8 jets ($M_{PUPPI,SoftDrop}$> 45 GeV) or two dijet pairs w/ M_{dijet} > 1 TeV
- 1+ tight b-tagged AK4 jet

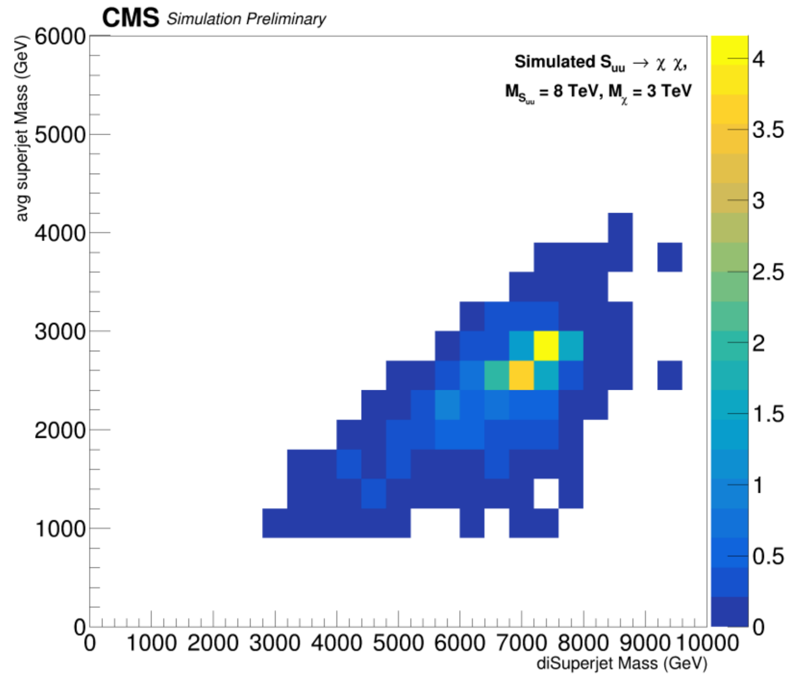


Above: the average reconstructed superjet mass from this technique obtained from tests on various simulated $S_{uu}/\ \chi$ mass combinations.

- Tests on $S_{uu} \rightarrow \chi\ \chi$ simulation samples for various $M_{S_{uu}}$ and M_χ combinations show high resolution in M_χ reconstruction
  - $M_{S_{uu}}$: 4-8 TeV, M_χ: 1-3 TeV
- Backgrounds are not resonant, so superjet mass vs diSuperjet distribution can give extra signal sensitivity

Sorting Algorithm on MC

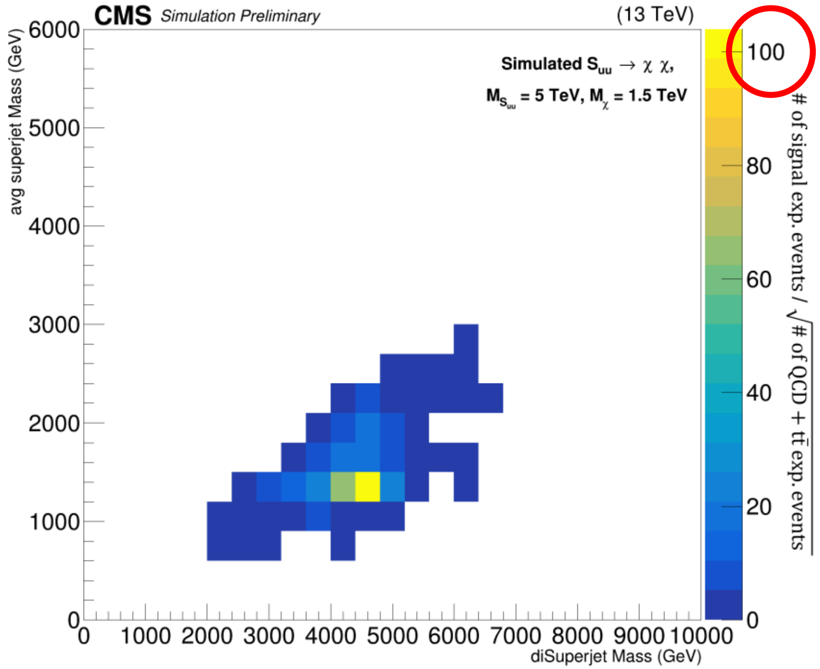


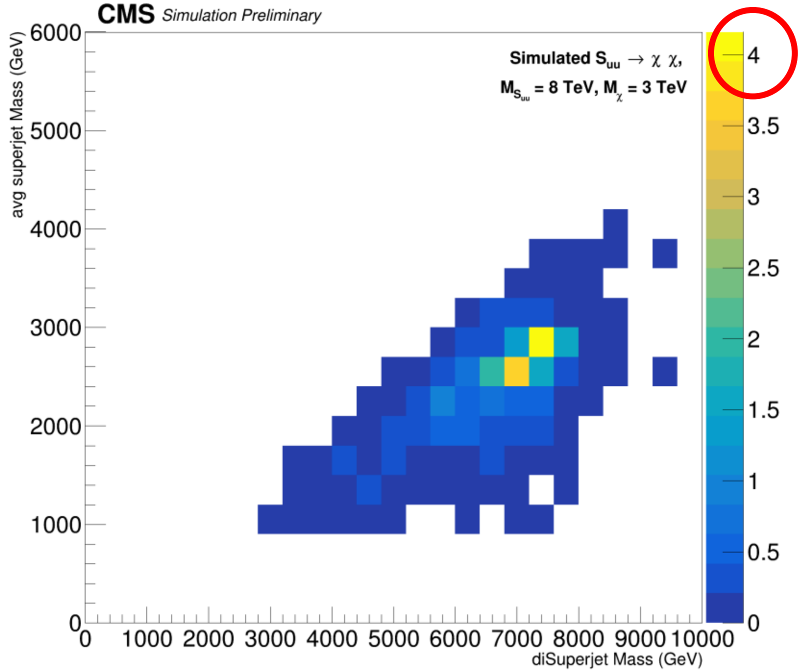


Above: Expected signal sensitivity, given by $\frac{N_{signal\ events}}{\sqrt{N_{QCD+t\bar{t}\ events}}}$, from simulation for this technique as a function of the average superjet mass and diSuperjet mass. The distributions are shown for two different $S_{uu}/\ \chi$ mass combinations

- Comparison of scaled signal to backgrounds (QCD & $t\bar{t}$) shows sensitivity greater than 4 in signal region up to 8 TeV in average SJ vs diSJ masses
- $\sigma_{S_{uu}\rightarrow\chi\ \chi}$ is only an estimate, but this gives us a lot of breathing room to float

Sorting Algorithm on MC



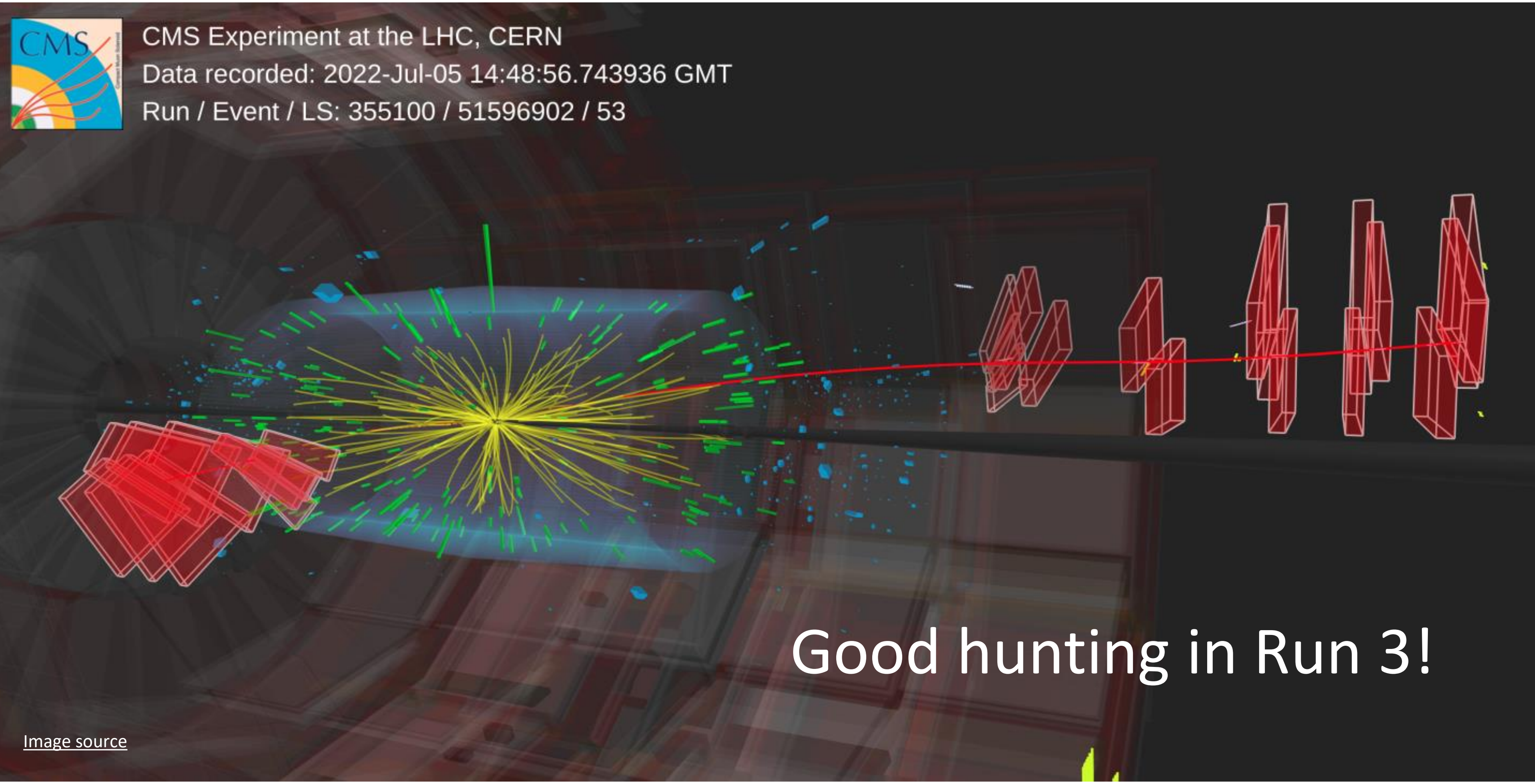


Above: Expected signal sensitivity, given by $\frac{N_{signal\ events}}{\sqrt{N_{QCD+t\bar{t}\ events}}}$, from simulation for this technique as a function of the average superjet mass and diSuperjet mass. The distributions are shown for two different $S_{uu}/\ \chi$ mass combinations

- Comparison of scaled signal to backgrounds (QCD & $t\bar{t}$) shows sensitivity greater than 4 in signal region up to 8 TeV in average SJ vs diSJ masses
- $\sigma_{S_{uu}\to\chi\ \chi}$ is only an estimate, but this gives us a lot of breathing room to float

Future Additions & Summary

- **This technique has been designed with maximum adaptability in mind to accommodate any pair-produced model**
  1. Any number of jets are accepted as inputs
  2. Steps are in place to counteract the effects of highly-boosted topologies
  3. A new class of superjet substructure variables become available for study
- Results on $S_{uu} \rightarrow \chi\,\chi$ simulation show the ability to consistently reconstruct VLQ masses with reasonable resolution
  - The resulting signal sensitivities in the superjet vs diSuperjet plane and expected S_{uu} cross-section facilitate Run 2 studies up to $M_{S_{uu}}$ = 8 TeV
- Room remains to add sophistication
  - Iterative rounds of jet reclustering to remove unwanted particles
  - Machine learning for sorting jets and recognizing superjet substructure
  - Optimization of initial boost value of all jets
- We hope the ideas we have worked on will be expanded and carried over to Run 3
- In the meantime, keep an eye out for our $S_{uu} \rightarrow \chi\,\chi$ Run 2 analysis!


CMS Experiment at the LHC, CERN
Data recorded: 2022-Jul-05 14:48:56.743936 GMT
Run / Event / LS: 355100 / 51596902 / 53
Good hunting in Run 3!
Image source

Thanks!

Glossary

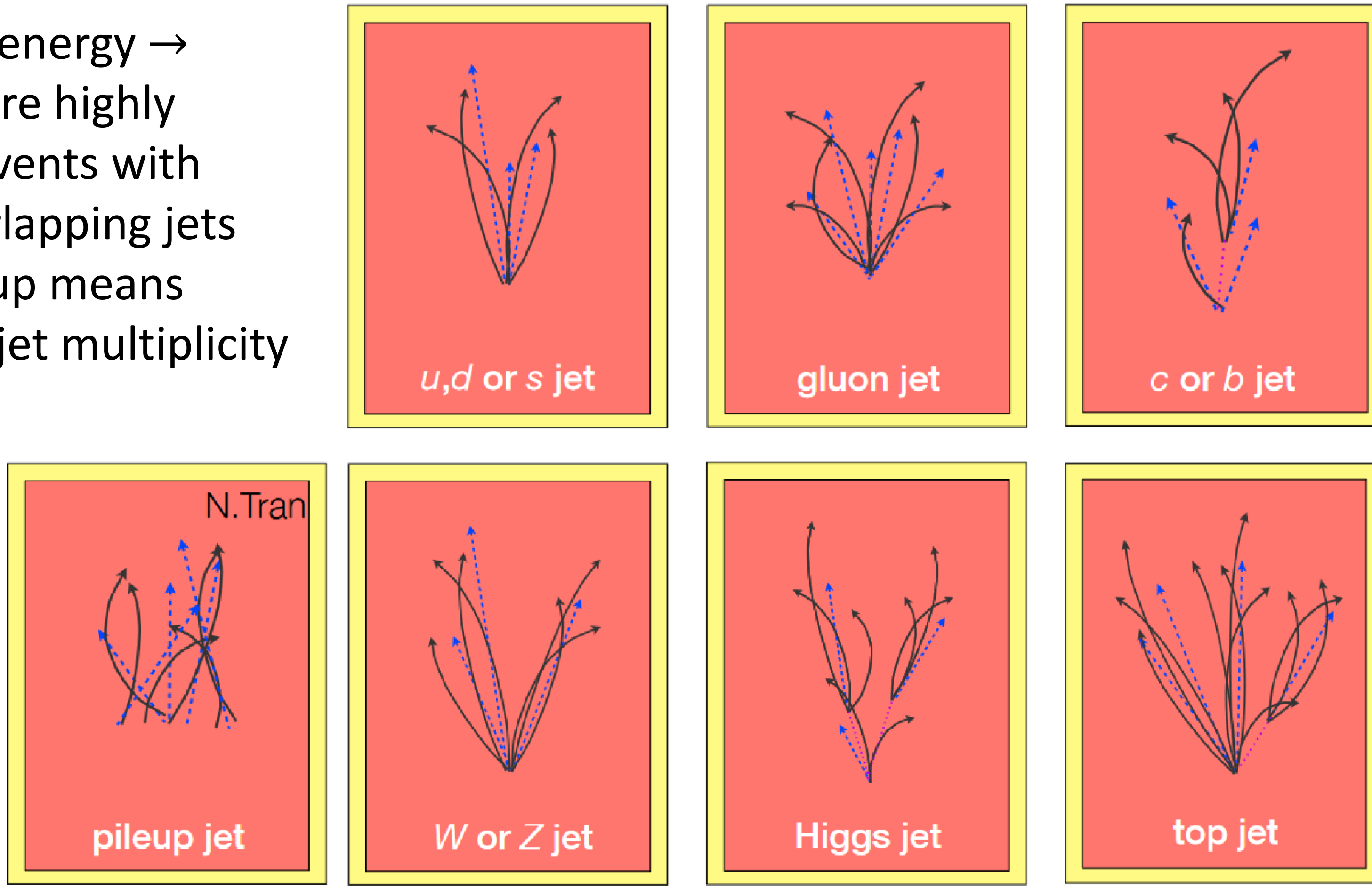
- **η**: pseudorapidity, defined as $\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right]$ for polar coordinate θ.
- **AK8/AK4 jets**: jets clustered using the anti-k_T algorithm [1] with R parameters 0.8 and 0.4 respectively. All jet reclustering in this technique is done using the FastJet package [2].
- $\mathbf{H_T}$: the scalar sum of p_T of AK4 jets with p_T >30 GeV and |η| < 2.5.
- **Vector-like quarks** (VLQ): hypothetical particles that appear in many extensions of the Standard Model.
- $\boldsymbol{S_{uu}}$**/diquark**: a spin-0, ultra-heavy resonance coupling to up quarks [3].
- $\boldsymbol{\chi}$: a vector-like quark with decays to W^+b, h t, and Z t with branching fractions 0.5, 0.25, 0.25 respectively [3].
- **Jet energy corrections** (JECs): corrections applied to jets to account for the detector effects.
- **Jet energy resolution** (JER) corrections: corrections applied simulation jet energy distributions to account for lower energy resolutions in simulation than data.
- **pileup** (PU): residual physics objects from events that are not of interest.
- **PUPPI:** PileUp Per Particle Identification, a method for lessening the effect of pileup in the interactions of interest by scaling out particles that are likely from pileup [4],[5].
- **Softdrop** (SD): a method of pileup mitigation that involves dropping soft, wide-angle radiation from jets [6].
- $\boldsymbol{M_{PUPPI,SoftDrop}}$: the invariant mass of a jet with PUPPI weights scaling each particle's 4-vector and the Softdrop algorithm applied.
- **b-tagging:** the process of identifying jets that came from b quarks. DeepJet/DeepFlavour is the b-tagging algorithm used in this note [7].
- **DeepTau**: a deep neural network trained to identify tau leptons from other physics objects [8].

References

- [1] M. Cacciari, G. P. Salam and G. Soyez, “The anti-kt jet clustering algorithm”, 2008, arXiv:0802.1189
- [2] Matteo Cacciari, Gavin P. Salam, Gregory Soyez, "FastJet user manual", 2011, arXiv:1111.6097
- [3] Bogdan A. Dobrescu, Robert M. Harris, Joshua Isaacson, “Ultraheavy resonances at the LHC: beyond the QCD background”, arXiv:1810.09429
- [4] Daniele Bertolini, Philip Harris, Matthew Low, Nhan Tran, “Pileup Per Particle Identification”, 2014, arXiv:1407.6013
- [5] CMS Collaboration, “Pileup-per-particle identification: optimisation for Run 2 Legacy and beyond”, 2020, CMS-DP-2021-001
- [6] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft drop”, 2014, arXiv:1402.2657
- [7] Emil Bols, Jan Kieseler, Mauro Verzetti, Markus Stoye, Anna Stakia, “Jet Flavour Classification Using DeepJet”, 2020, arXiv:2008.10519
- [8] CMS Collaboration, “Identification of hadronic tau lepton decays using a deep neural network”, 2022, arXiv:2201.08458
- [9] V. D. Barger, R. J. N. Phillips, “Collider Physics” Frontier in Physics, p. 284, 1997, Addison-Wesley Publishing Company, Inc.

Challenges in Run 3

- Increased energy → busier, more highly boosted events with many overlapping jets
- More pileup means increased jet multiplicity



Jet substructure, The CMS Collaboration, 2022

Imagine these jets all superimposed, overlapping, etc.

Analysis Regions

Signal Region
- Lepton veto (medium IDs + isolations)
- PFHT_1050 trigger pass + further 1500 HT cut
- 3+ AK8 jets (pT > 300 GeV, medium ID)
- 2+ heavy AK8 jets (M > 45 GeV) or two dijet pairs w/ $M_{dijet} > 1$ TeV
- 1+ tight b-tagged AK4 jet
- Require each SJ have 2+ recluster AK4 jets with E>300 GeV or be tagged by NN

QCD Control Region
- Lepton veto (medium IDs + isolations)
- PFHT_1050 trigger pass + further 1500 HT cut
- 3+ AK8 jets (pT > 300 GeV, medium ID)
- 2+ heavy AK8 jets (M > 45 GeV) or two dijet pairs w/ $M_{dijet} > 1$ TeV
- 0+ tight b-tagged AK4 jet
- Require each SJ have 2+ recluster AK4 jets with E>300 GeV or be tagged by NN

Anti-tag Region
- Lepton veto (medium IDs + isolations)
- PFHT_1050 trigger pass + further 1500 HT cut
- 3+ AK8 jets (pT > 300 GeV, medium ID)
- 2+ heavy AK8 jets (M > 45 GeV) or two dijet pairs w/ $M_{dijet} > 1$ TeV
- 0+ tight b-tagged AK4 jet
- One (randomly selected) SJ does not pass substructure cuts or is not tagged by NN

TTbar Control region - pending