MPI Higgs Physics Analyses, 07.10.2015, Munich

Status and Plans for VHbb

Felix Müller
Dan Nebe
David Joseph
Ilona Weimer
Sandra Kortner

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Introduction (1)

Why H → bb?

- Test of fermionic coupling
- No evidence in this channel yet

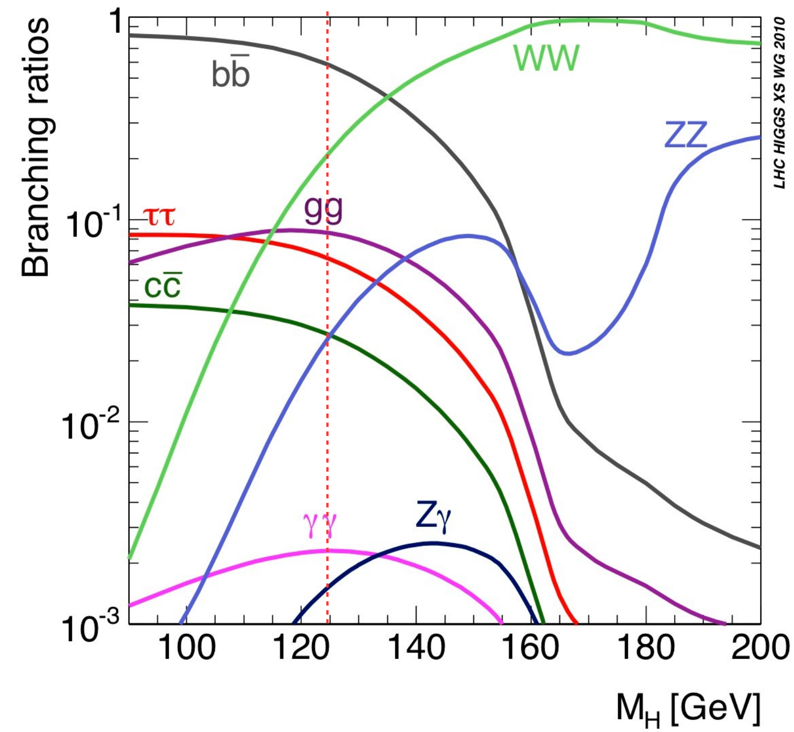
Expected sensitivity of **2.6σ**: observed **1.4σ** deviation from background only hypothesis

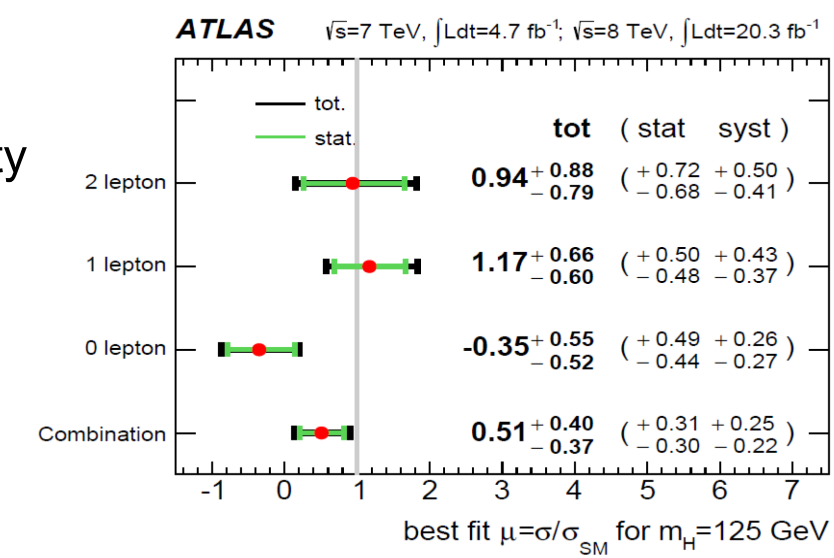
⇒ *VH* signal strength $\mu = 0.52 \pm 0.32(\text{stat.}) \pm 0.24(\text{syst.})$

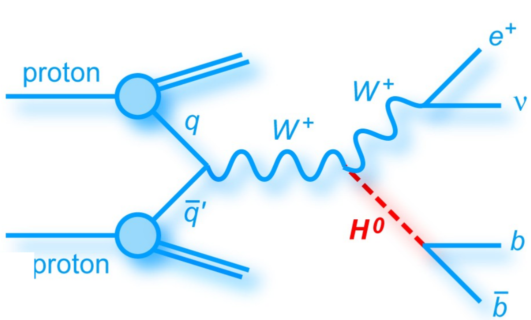
Check of *VZ*,*Z* → $b\bar{b}$ signal strength : $\mu = 0.74 \pm 0.09(\text{stat.}) \pm 0.14(\text{syst.})$

Experimental challenge

- H → bb largest BR [58% at m_H = 125 GeV]
- Very difficult in hadron collider
  - Gluon fusion: gg → H → bb
    [no handle against QCD background]
  - Associate production: VH → Vbb
    [$\sigma_{VH} = 0.1 \times \sigma_{gg\to H}$]
- ZH → ννbb and WH → lνbb with similar sensitivity







Content

- Introduction to WHbb
  - Overview about analysis
  - Optimization of selection criteria at 13 TeV
- Substructure in WHbb
  - Introduction to large-R jets, substructure
  - Potential improvement in WHbb using substructure
  - Current status and next steps
- Technical background
  - CxAOD Framework
  - Xbb tagger
- Summary and Outlook

Introduction to WHbb

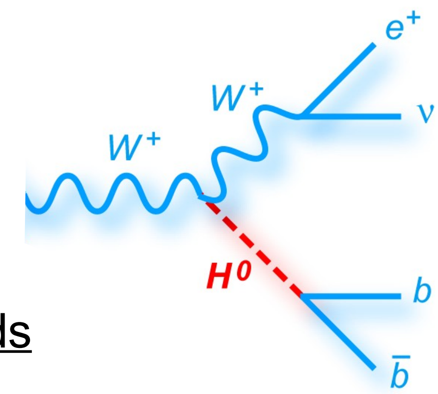
Selection

- 1 lepton [electon, muon]
- 2 b-jets with tight working point
  [optimized for signal efficiency, c-quark rejection]
- Classification using p_T^V
  [most sensitive in p_T^V > 200 GeV]
- Cuts on dR_{bb}, m_T^W, H_T, MET, m_{bb}
  → Final analysis with multi-variate approach [BDT with ~12 variables]

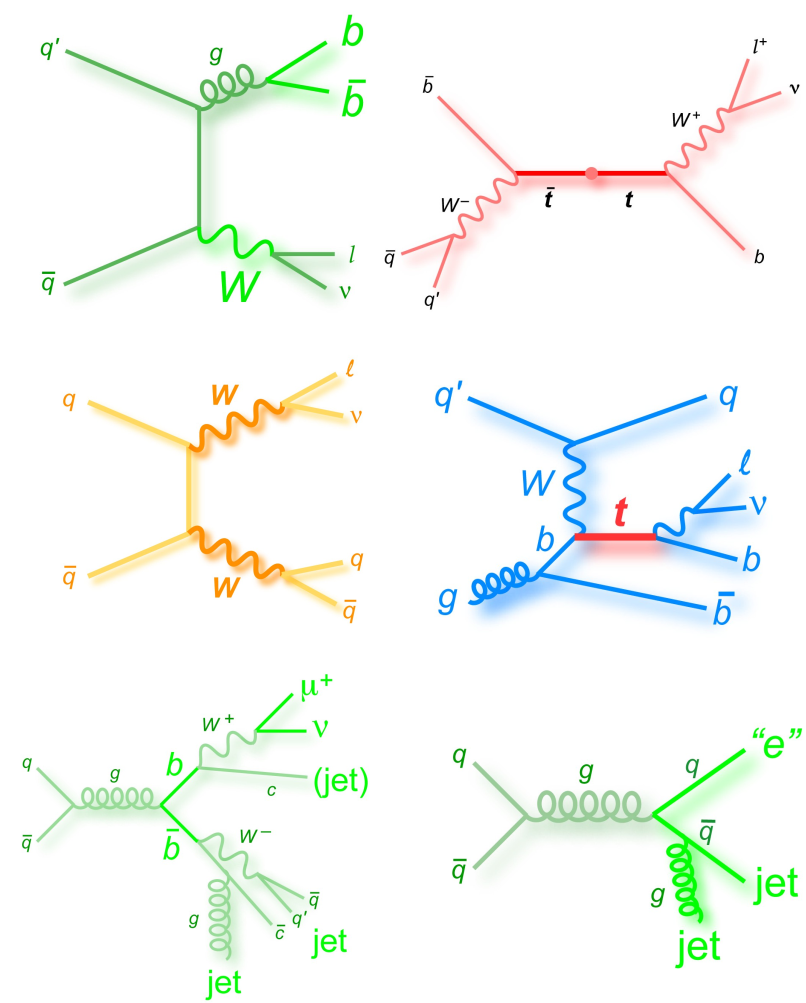
Backgrounds

- W+jets
- Ttbar
- single top
- Di-boson
- Fake-V
- QCD

signal



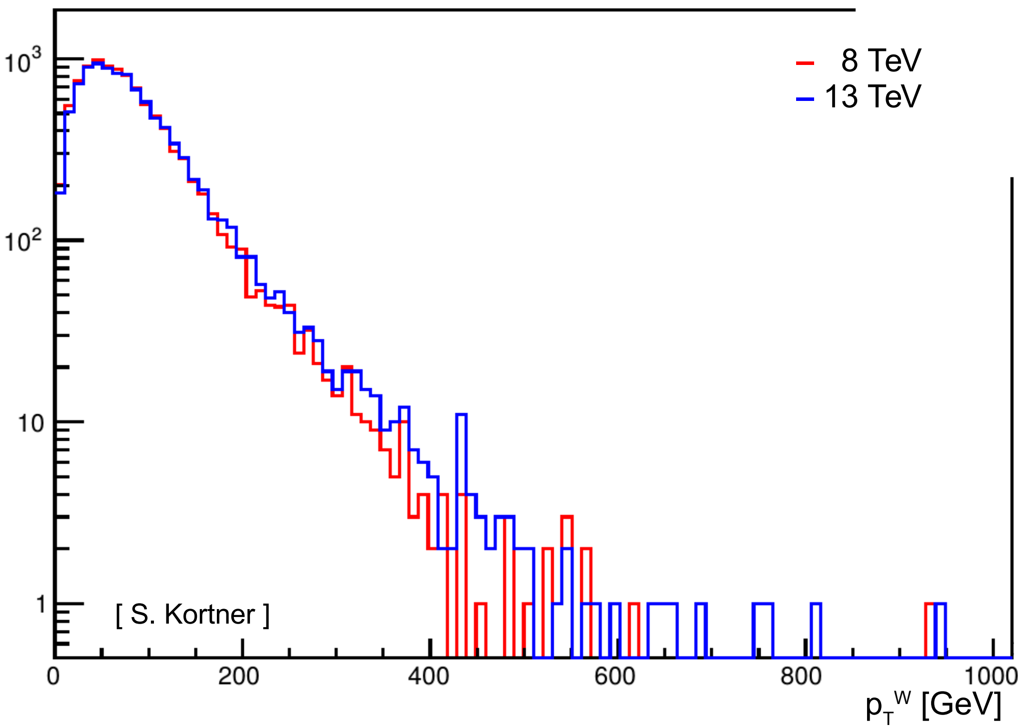
backgrounds

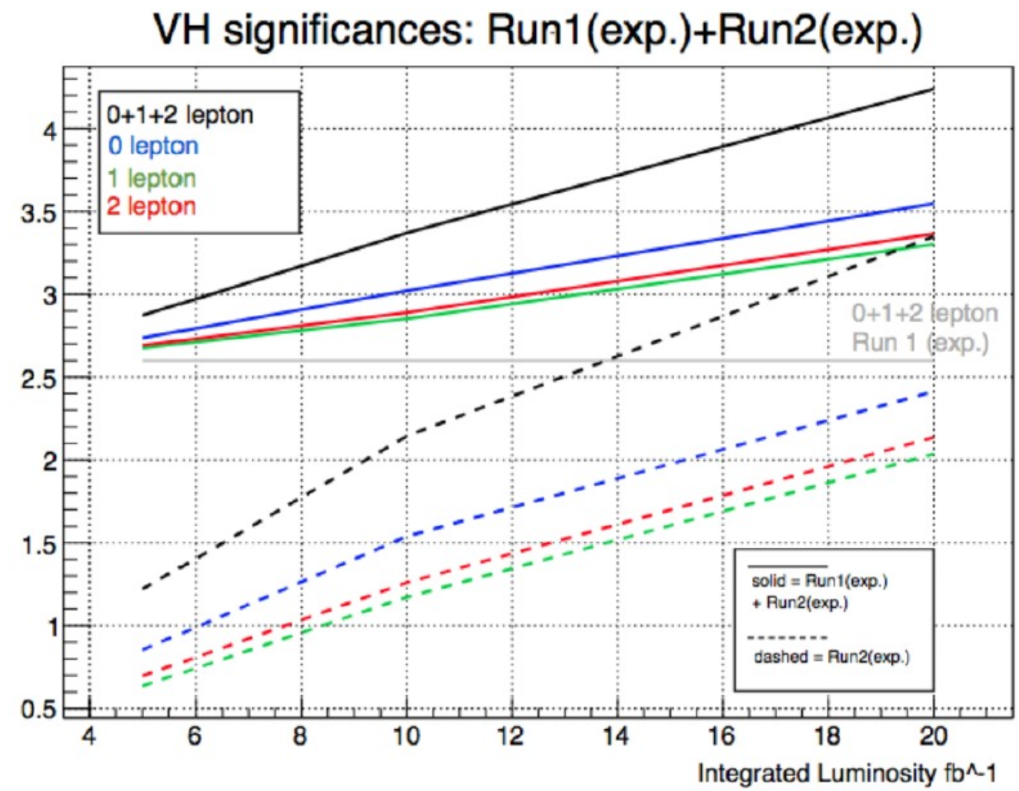


From Run1 to Run2

- Increased signal cross section
  [13 TeV yields: σ_{VH} x 2, σ_{tt} x 3,3]
- Improved detector (b-layer)
- Harder spectrum

- Some backgrounds increase
- High pile-up conditions
- Need at least same CP performance than for Run I





- SM VHbb analysis needs O(10) fb^{-1} for evidence
- Short term (winter conferences): Searches for AZh, HVT, ...

Study of event selection at 13 TeV [David Joseph]

Idea

- Check selection of 8 TeV analysis at 13 TeV
  → try to optimize cuts, improve significance

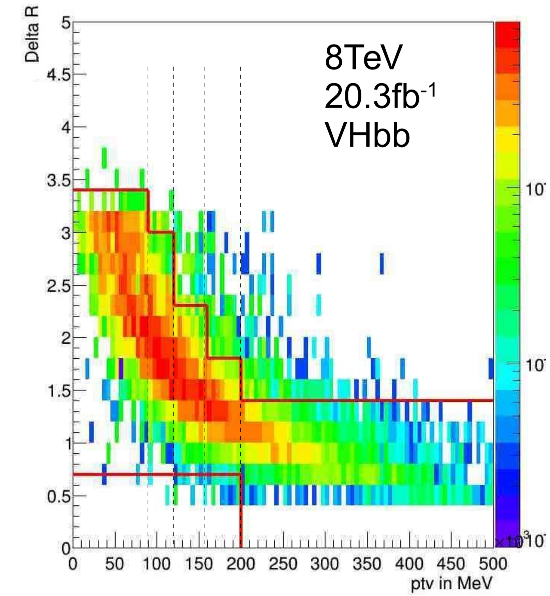
Results

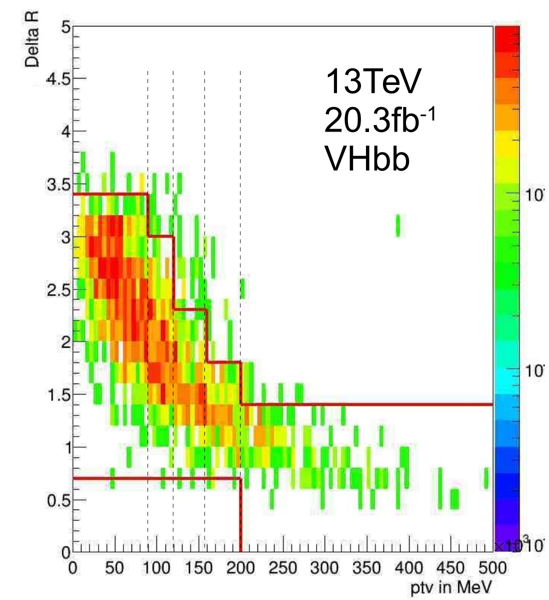
- Event topology not very different
  → improvements only at low p_T^W (dR, H_T)
- Comparison of significance

P_T^W [GeV]	$\sqrt{s} = 8\,TeV$ [σ]	$\sqrt{s} = 13\,TeV$ standard [σ]	$\sqrt{s} = 13\,TeV$ optimized [σ]
0 - 90	0,170 ± 0,003	0,319 ± 0,03	0,344 ± 0,02
90 - 120	0,174 ± 0,004	0,23 ± 0,02	0,23 ± 0,02
120 -160	0,277 ± 0,004	0,32 ± 0,02	0,33 ± 0,02
160 - 200	0,37 ± 0,01	0,4232 ± 0,02	0,4232 ± 0,02
> 200	0,68 ± 0,05	0,79 ± 0,09	0,79 ± 0,09

small improvement

largest improvement, no optimization





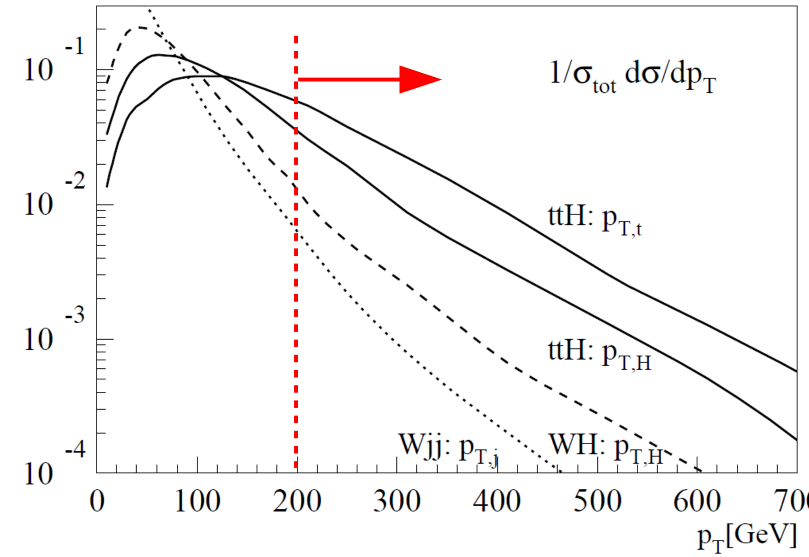
Outlook

- Choice of event selection and signal regions (e.g. 1-2 additional jets) still under study in VHbb group

Large-R jets and substructure

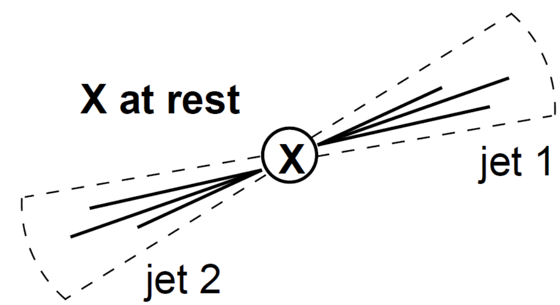
High p_T region is most promising

- Relevant fraction of signal at high p_T
- Signal harder than backgrounds
  [best significance at high p_T]
- Easier jet combinatorics at high p_T
  [*cleaner* events]



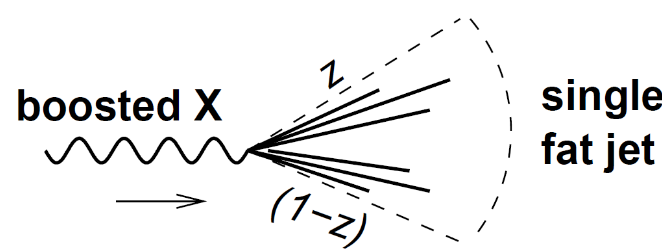
Boosted event topologies!

→ Jets start to merge
$p_T > 2m / R$ [R = 0.4: $p_T^{Higgs} > 600$ GeV]
→ Single fat jet



Large-R jets and substructure

- Use dedicated algorithms to select
  H → bb system at high boost
- Large-R jet to preselect decay particles
- b-tagging and substructure to identify
  H → bb



Large-R jets

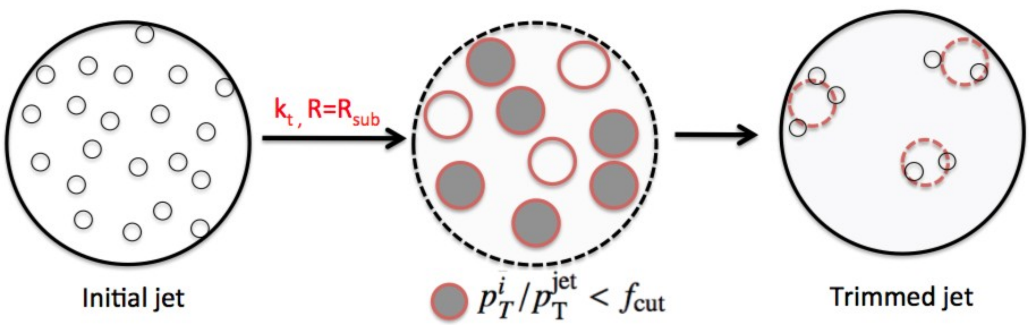
D. Krohn, J. Thaler and L.-T. Wang, “Jet Trimming", arXiv:0912.1342
[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetGrooming]

2 b-jets vs. 1 large-R jet

- m_{bb} from 2 b-jets might miss final-state radiation (FSR)
- Mass of large-R jet includes radiation
- But: contamination from pile-up, initial-state radiation (ISR), ...

Trimming algorithm

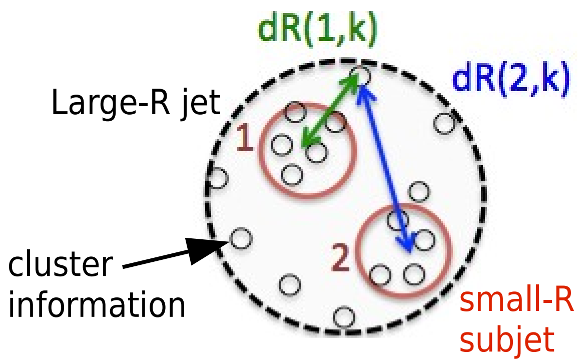
- Contribution from pile-up, ISR usually much softer than FSR
- Trim soft constituents of the large-R jet
- Ratio of p_T of constituents with respect to the jet p_T used as criterion



Method

- Create subjets from large-R jet constituents
  [Large-R jet: anti-k_t, R=1.0; subjet: k_t, R=0.2]
- Remove subjets *i* with p_T^i / p_T^{jet} < f_{cut}
  [typ: f_{cut} = 0.05]

Substructure techniques



Moment	xAOD Jet attribute names
N-subjettiness	Tau1, Tau2, Tau3, Tau1_wta, Tau2_wta, Tau3_wta
N-subjettiness ratios	Tau21, Tau32, Tau21_wta, Tau32_wta
kT splitting scale	Split12, Split23, Split34
zCut	ZCut12, ZCut23, ZCut34
Dipolarity	Dip12, Dip13, Dip23, DipExcl12
Angularity	Angularity
kT Delta R	KtDR
kT Mass drop	Mu12
Planar flow	PlanarFlow
Energy correlations	ECF1, ECF2, ECF3, ECF1_Beta2, ECF2_Beta2, ECF3_Beta2
ECF ratios	C1, C2, D2, C1_Beta2, C2_Beta2, D2_Beta2
Thrust	ThrustMin, ThrustMaj
FoxWolfram	FoxWolfram0, FoxWolfram1, FoxWolfram2, FoxWolfram3, FoxWolfram4
Sphericity	Sphericity, Aplanarity
Jet charge	Charge
Pull	PullMag, PullPhi, Pull_C00_, Pull_C01_, Pull_C10_, Pull_C10_
Shower deconstruction	ShowerDeconstructionW, ShowerDeconstructionTop
Q-jets volatility	Volatility

subjettiness

energy correlation functions

[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetSubstructureTools]

- Substructure of large-R jets contains additional information, e.g. energy distribution inside subjet

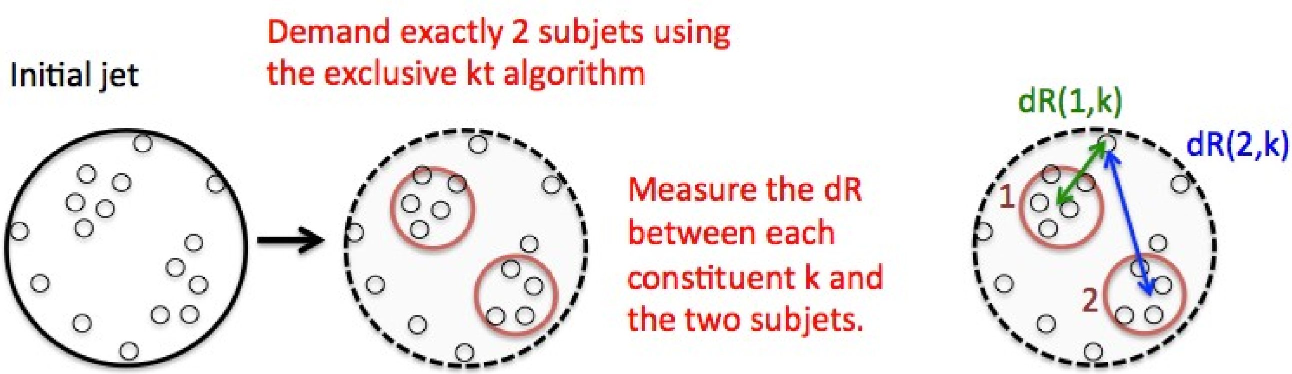
- Many substructure variables
  [not taylored to a particular decay / tagger]
  - Might have separation power for VH → $b\bar{b}$ analysis

- Additional variable for selection
  - Starting point for investigation
    [→ which features can improve S/B]

Promising to study

Example: n-subjettiness

[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetSubstructureVariables]



d_k for cluster k with pT(k):

d_k = pT(k) x min(dR(1,k),dR(2,k))

Normalisation d_0 for jet with radius R:

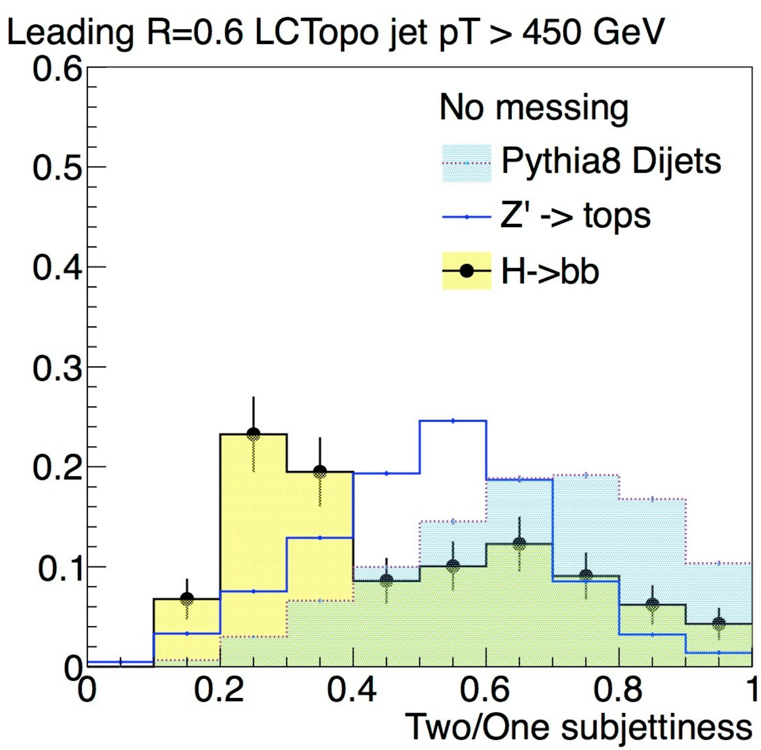
$d_0 = \Sigma(p_T(k) \times R)$

τ_2 for jet with exactly 2 subjets:

$\tau_2 = \Sigma d_k / d_0$

Jet with exactly 2 subjets

- Calculate d_k of all consitutents
  - → if consituent is close / within subjet: d_k small
  - → if constituent is far from all subjets: d_k large
- Calculate normalisation d_0
- τ_2 is the the two-subjettiness
  - → collimated subjets: τ_2 small
- Ratios of 1-, 2- and 3-subjettiness: τ_{32} ,τ_{21}



Substructure in resolved analysis [Dan Nebe]

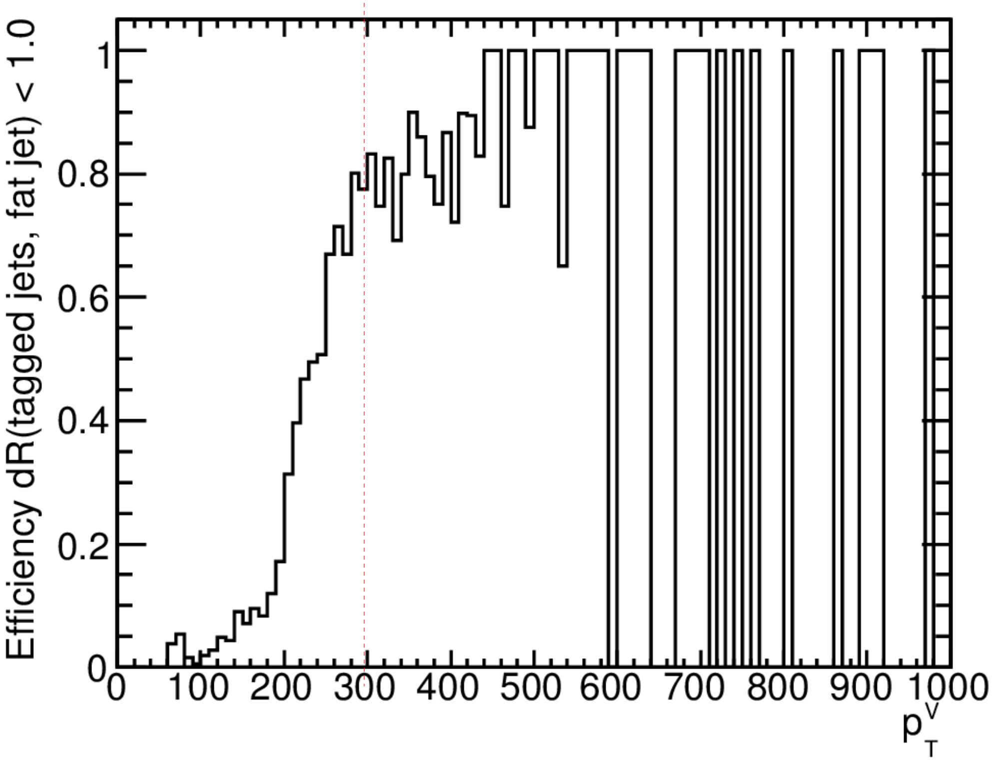
Idea

- Study WH→bb, at 13 TeV
  → event selection of resolved analysis from Run1 as baseline
- Match large-R jets to bb system
  → only use those events in this analysis
- Investigate potential of different large-R jet algorithms and substructure variables on significance

<u>Higgs tagging *light*</u>
- **Large-R jet used for sub-structure informtion only**
- **No mass, no calibration**

Matching

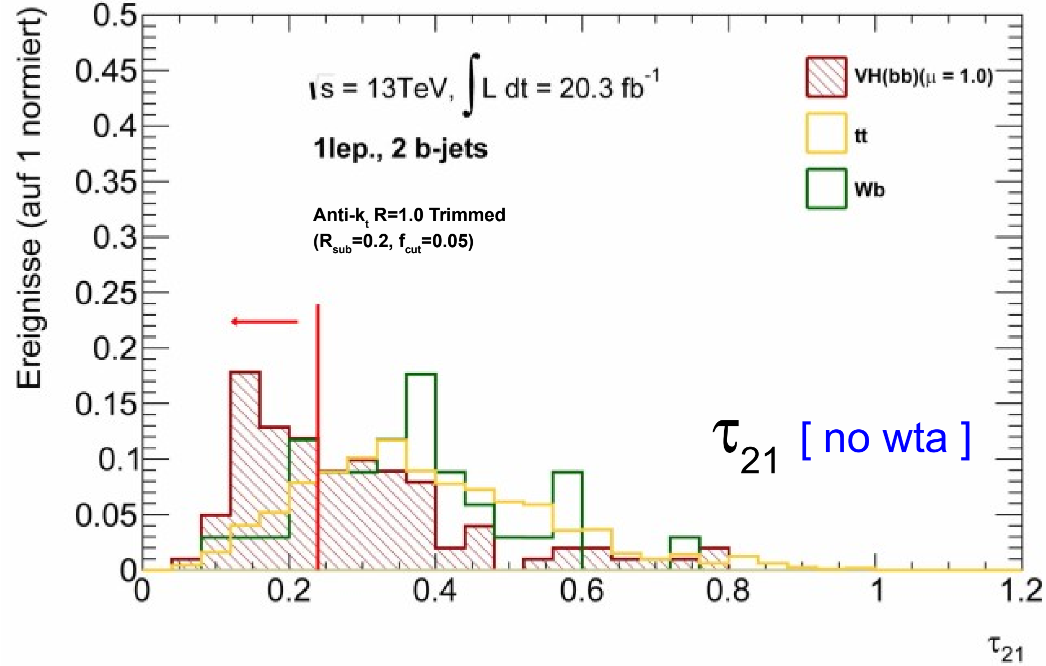
- R = 1.0 for tested algorithms
  → limits matching to high p_T^W
- $p_T^W > 300$ GeV

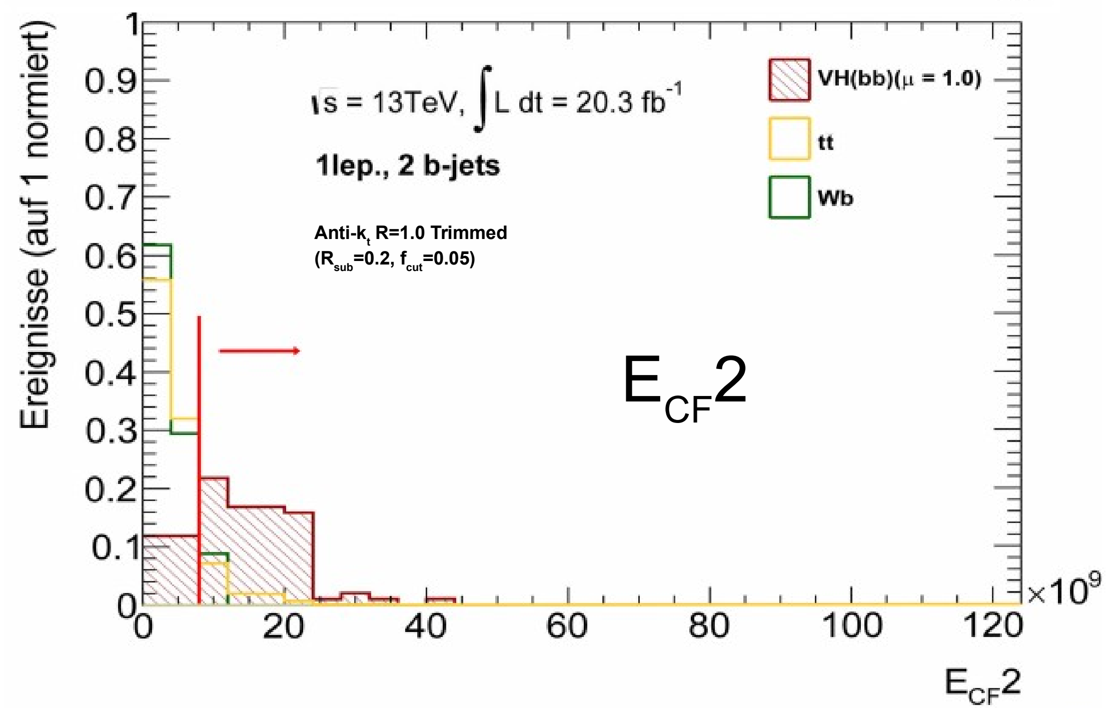


Substructure variables

- Clear separation of signal and background
  - τ_{21} - subjettiness
  - Energy correlation functions $E_{CF}1$, $E_{CF}2$
    [should have tested ratios as well]

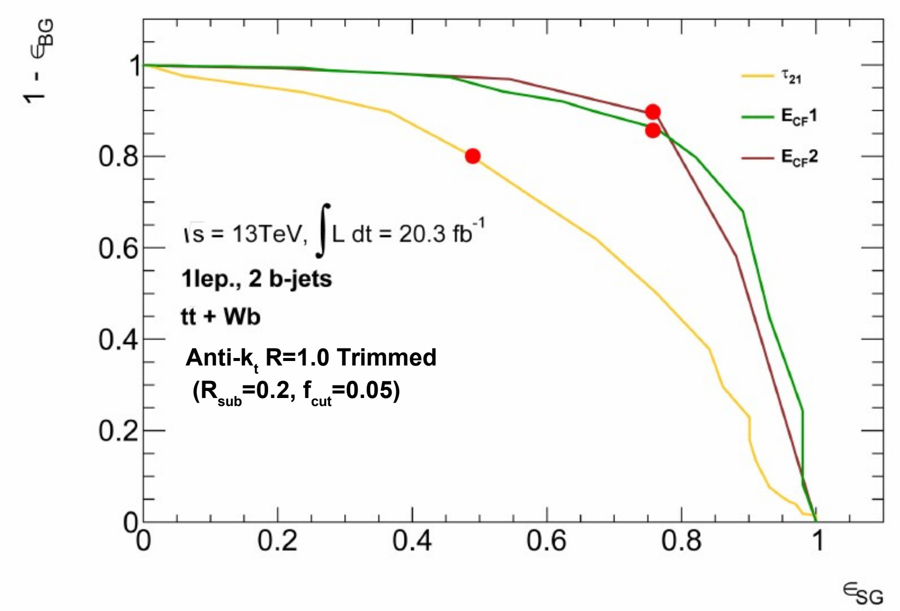
looks very promising





- Non-optimal b-tagging
  [WP not optimized]
  [SV1_IP3D instead of MV2c20]
- No cut on m_{bb}
  [seems to make difference]

Comparison and selection efficiency



ROC curve

- Large improvements using E_{CF}
- Improvements from subjetiness not as large
  → results depend on jet algorithms

suggests large improvement for signal significance

Sample	τ_{21}	$E_{CF}1$	$E_{CF}2$
VH→$b\bar{b}$	49%	76%	76%
$t\bar{t}$	21%	13%	12%
Wb	19%	15%	9%

Sample	τ_{21}	$E_{CF}1$	$E_{CF}2$
VH→$b\bar{b}$	68%	68%	68%
$t\bar{t}$	40%	11%	10%
Wb	40%	11%	6%

Signal and background efficiencies

- Visual cuts:
  76% of signal with rejection of almost 90% for $E_{CF}2$

- Constant signal efficiency (68%):
  Energy correlation function clearly better than subjetiness

Significance using substructure

- Calculation of significance
  → m_{bb} in [100, 150] GeV

$$Z = \sqrt{2 \cdot \left((s+b) \cdot \ln\left(1+\frac{s}{b}\right) - s\right)}$$

- Significance in matched events

	Anti-kt20		
Selektion	VHbb	$t\bar{t}$	Wb
$\Delta R_{max} \geq 1$	$43,8 \pm 1,4$	$7924,8 \pm 57,2$	$1248,7 \pm 98,1$
$\Delta R_{max} < 1$	$5,3 \pm 0,5$	$492,1 \pm 14,7$	$314,4 \pm 53,9$
τ_{21}	$2,6 \pm 0,4$	$103,3 \pm 3,1$	$59,7 \pm 10,1$
$E_{CF}1$	$4 \pm 0,5$	$64 \pm 1,9$	$47,2 \pm 8$
$E_{CF}2$	$4 \pm 0,5$	$59,1 \pm 1,7$	$28,3 \pm 4,8$
Signalsignifikanz	Gesamt		
$Z_{\text{o.Ü.}}$	0,45 } 0,483		
$Z_{\text{Überlapp}}$	0,187		
$Z_{\tau_{21}}$	0,209		
$Z_{E_{CF}1}$	0,379		
$Z_{E_{CF}2}$	0,428		

→ significance for unmatched events
→ significance for matched events
→ significance for matched events including cut on $E_{CF}2$

significance improves by 229%in matched events

- Significance in all events

	Anti-kt20	
Selektion	Signalsignifikanz	Verbesserung
τ_{21}	0,496	$2,7\%$
$E_{CF}1$	0,588	$21,8\%$
$E_{CF}2$	0,621	$28,6\%$

→ total improvement of significance using $E_{CF}2$

significance improves by 28.6% in total

Status and plans

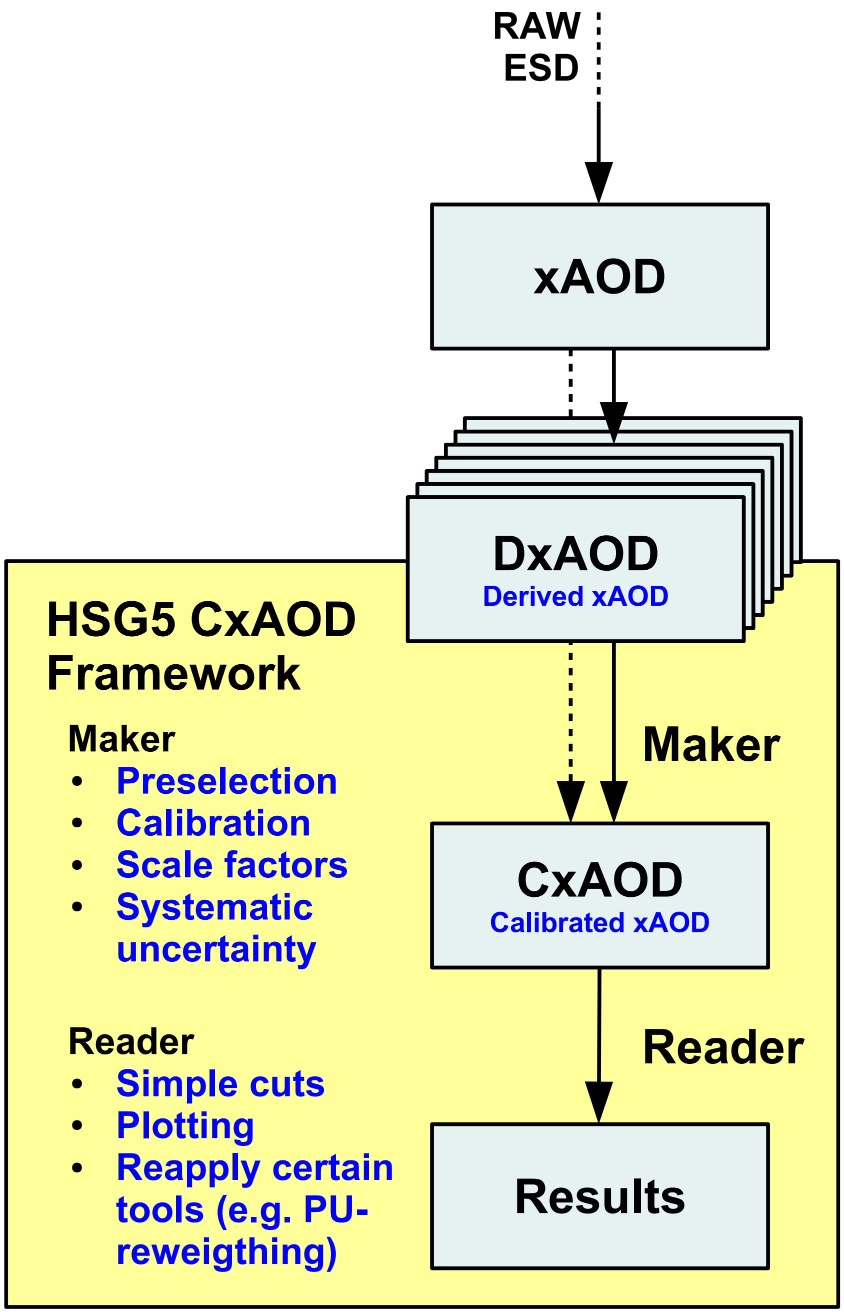
Cross-checks of substructure analysis ongoing

- Latest release [R19 → R20]
- Latest CxAOD version [00-09-03 → 00-14-00]
- Update b-tagging [MV2c20, track jets]
- Validate against Xbb tagger
  → use official tool
  → check against Xbb results [no large improvements from substructure]

Plans

- Study Hbb tagging at *low* p_T for SM Higgs analysis
- Validate Xbb tagger in VHbb analysis
  [Xbb recommendations from generic background estimates]
- Optimize / extend Xbb tagger to low p_T
  [*light* tagger since no calibration for R>1.0; question of mass resolution]
- Study combination of b-tagging and substructure
  - 1 tight b-tag + substructure
  - 2 loose b-tags + substructure

CxAOD and HSG5Framework



- New Event Data Model (EDM)
  - Improve turn-around times
  - Reduce disc space
- XAOD: large production campaigns
- DxAOD: regular production
  [dedicated DxAODs shared among CP / physics groups]
- CxAOD: analysis specific

HSG5Framework

- Common framework
  - Started from VHbb group
  - Used by many different analyses
  -
- Common production
  - CxAOD (calibrated xAOD)
  - All tasks done at Maker level
    [only minimal information in CxAOD]

[switch from PyRoot to Reader?]

Xbb tagger: introduction

Idea

- Tagger for $H \rightarrow b\bar{b}$ decays
  [analogous to top tagger or W/Z tagger]
- Most scenarios deal with highly boosted Higgs decays
- Higgs identification and mass using large-R jet

Incredients

- Large-R jet
- 2 b-tags
- Mass window around 125 GeV
- Substructure variable for additional background rejection

Working points

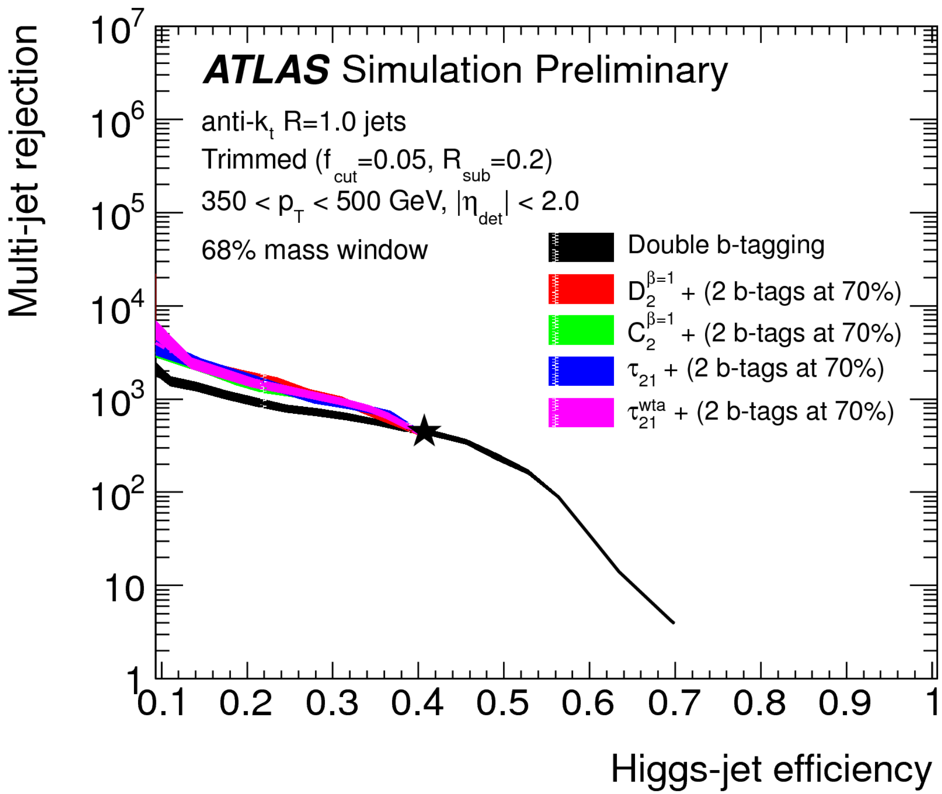
WP	jet algorithm	#bjets	b-tag	mass	substructure
loose	AK10LCTRIMF5R20	2	70%	90%	-
medium	AK10LCTRIMF5R20	2	70%	68%	-
tight	AK10LCTRIMF5R20	2	70%	68%	D2

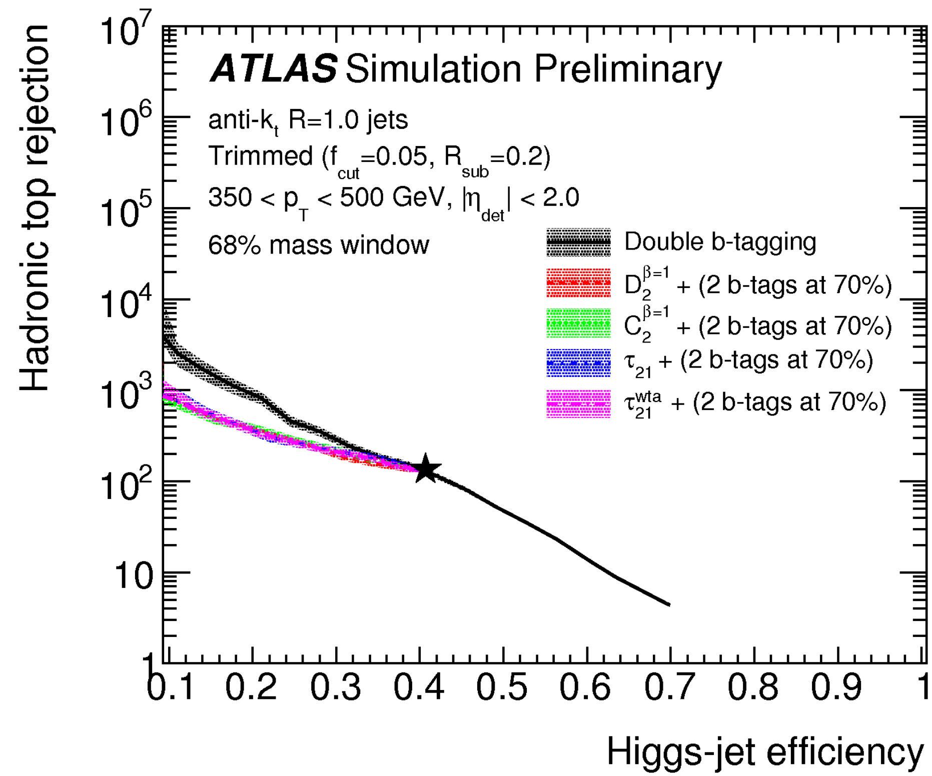
Xbb tagger: details

- Jet algorithm: Anti-k_t R=1.0 LCTopo Trimmed f=0.05, r=0.2
- Match track jets to fat jet
  - GhostAntiKt2TrackJet with p_T > 10 GeV, $|\eta|<2.5$, ≥ 2 constituents
  - ≥ 2 matches
- B-tagging from track jets
  - Improved angular resolution with respect to b-hadron
  - MVc20 70% (medium WP)
- Muon correction
  - Combined muons with medium quality, p_T > 10 GeV, $|\eta|<2.5$
  - dR < 0.2 [closest match]
  - Correct 4-vector by adding muon [subtracting muon energy loss]
- Mass cut
  - $m_{min} < m_{corr} < m_{max}$ [90% or 68% window]
- Substructure cut
  - D2 [ratio of $E_{CF}1$ and $E_{CF}2$, similar to τ_{21}]
  - Cut p_T dependent (4^{th} order plynomial)

Xbb tagger: performance

Performance tested using RSG → HH → 4b, multijet, hadronic $t\bar{t}$





Multijet rejection
for 350 < p_T < 500 GeV

- Substructure variables bring improvements

Hadronic top rejection
for 350 < p_T < 500 GeV

- b-tagging much more powerful
- Improvements seen in VHbb study
  → better supp for semi-hadronic top
  → topology dependence?

Xbb tagger in HSG5 framework

- Xbb tagger implemented in the framework [FatJetHandler]
- Running on Maker level [DxAOD → CxAOD]
  - Flags each jet (loose, medium, tight WP)
  - Work ongoing to make this possible on Reader level
    - Issue with trimmed instead of untrimmed jet collection:
      [untrimmed jets should be used to match track jets]
- Systematic uncertainty
  - Jet mass, jet p_T, D2 scale uncertainty
  - Using CP tools mechanism
    - Not sure if D2 scale uncertainty is included in jet variation
    - No variation of muon correction missing
- HSG5Framework needs volunteers
  - Checking CP recommendations for large-R jets [Felix, Rainer]

Summary and Outlook

Results

- Optimization of selection criteria at 13 TeV
  - No large improvements from cut adjustments
  - Event selection and signal regions still under study in VHbb group
- Substructure in WHbb
  - Potential improvement in WHbb using substructure
  - Several cross-checks necessary
    [not ready for this meeting, sorry]

Plans

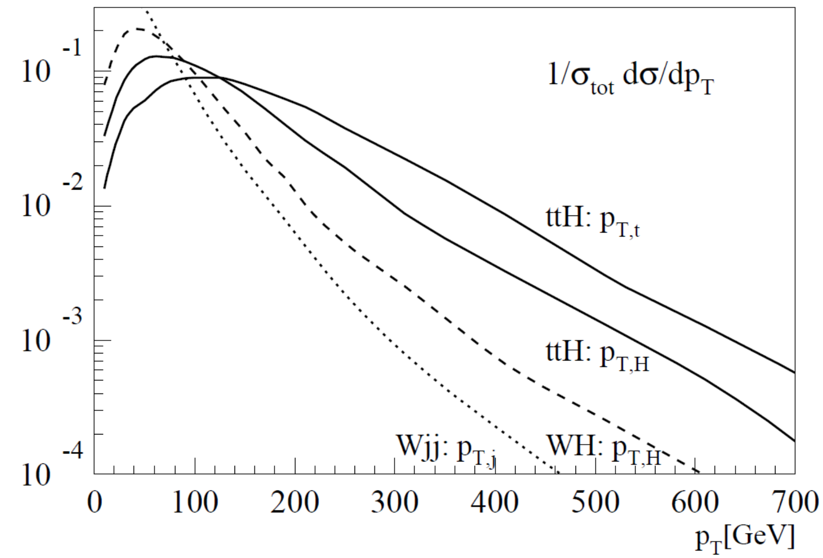
- Study Xbb tagging performance in WHbb
- Try to optimize Xbb tagging
  - Focus on low p_T regime for SM Higgs analysis
  - combination of b-tagging and substructure
    - 1 tight b-tag + substructure
    - 2 loose b-tags + substructure

Backup

Fat jets in VH → $b\bar{b}$

High p_T region is most promising

- Relevant fraction of signal at high p_T
- Signal harder than backgrounds
  [S/B improves at high p_T]
- Easier jet combinatorics at high p_T
  [*cleaner* events]

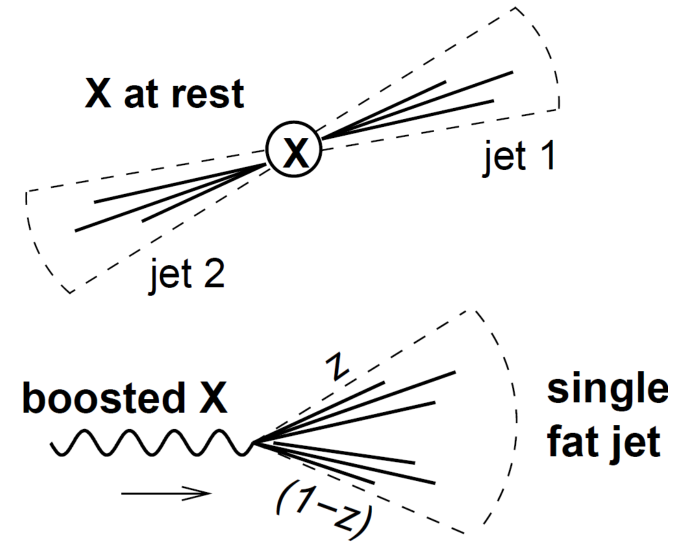


Boosted event topologies

→ Jets start to merge

$p_T > 2m / R$ [$R = 0.4$: $p_T^{Higgs} > 600$ GeV]

→ Single fat jet



Idea of fat jets / jet substructure

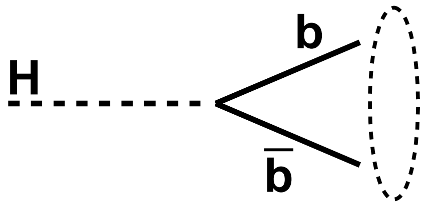
- *Preselection* of decay products
  [reduce combinatorics]
- Decompose decay
  → Tag heavy states like b-jet
  → High pile-up environment
  Filtering, Pruning, Trimming [see backup]

Not used in VH → $b\bar{b}$ yet

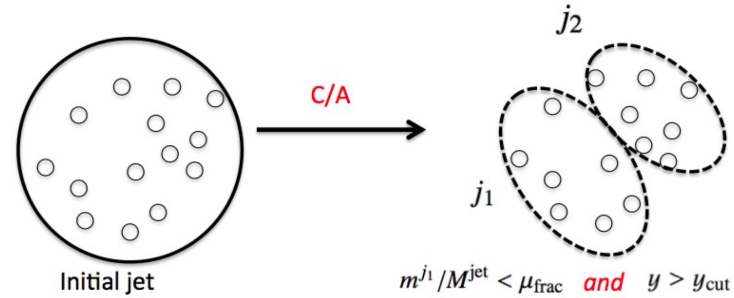
BDRS algorithm

J.M. Butterworth, A.R. Davison, R. Mathieu and G.P. Salam:
Jet substructure as a new Higgs search channel at the LHC
Phys. Rev. Lett., 100, 242001, 2008, arXiv:0802.2470 [BDRS algorithm or mass drop filter]

- Initial jet using C/A algorithm with R ~ 1.2 [fat jet]
- Use cluster sequence of jet algorithm
  - Undo last step of C/A clustering
    [jet splitting into two subjets]



fat jet



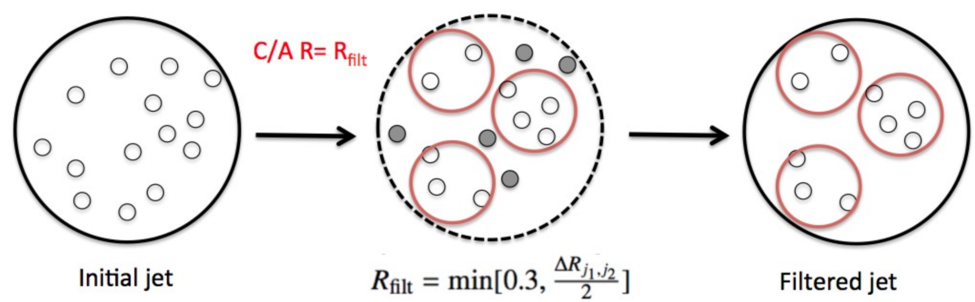
  - Mass drop criterion:
    [require large mass difference between decaying particle and decay products]

$$m_1/m < \mu_{\rm frac}$$ [0.67]

  - Symmetry criterion:
    [decaying particles carry similar p_T]

$$\frac{\min\left(p_{T1}^2, p_{T2}^2\right)}{m^2} \times \Delta R_{12}^2 > y_{cut}$$ [0.09]

  - Recluster subjets:
    [jet filtering]

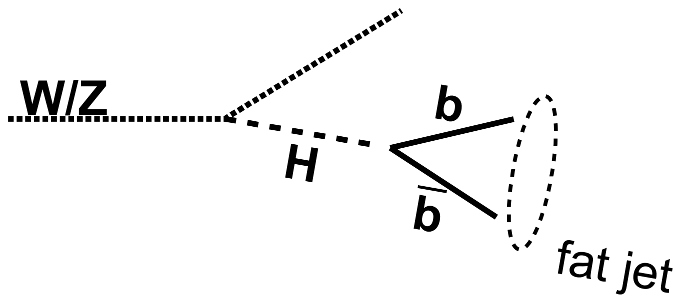


  - B-tagging: require b-tag for leading two subjets

BDRS algorithm: examples

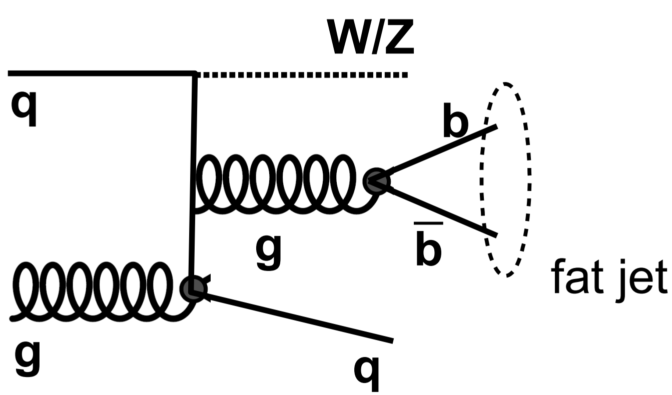
- **Large-R jet**
- **Mass drop criterion**
- **Symmetry criterion**
- **B-tagging**

Signal



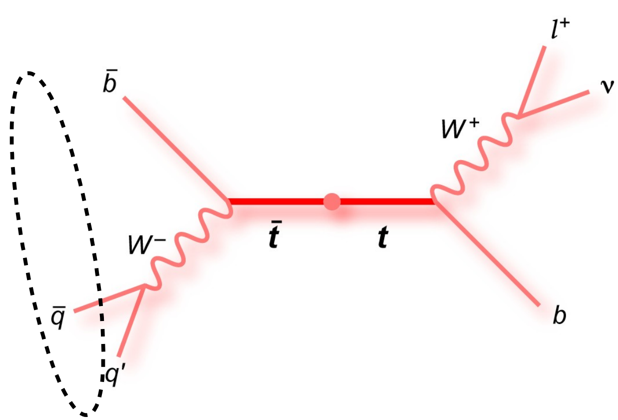
Backgrounds

- V + jets



- Top

fat jet



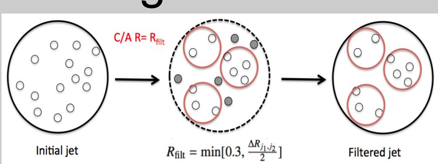
- Mass drop criterion
  [mass drop small for gluon splitting
- B-tagging
  [most effective cut]
- Jet Radius

Jet definition	σ_S/fb	σ_B/fb	$S/\sqrt{B \cdot \text{fb}}$
C/A, $R = 1.2$, MD-F	0.57	0.51	0.80
$K_\perp$, $R = 1.0$, y_{cut}	0.19	0.74	0.22
SISCone, $R = 0.8$	0.49	1.33	0.42

[S: signal, B: Z+jets, $200 < p_T^Z < 600$, $110 < m_J < 125$, b-tag]

To be studied

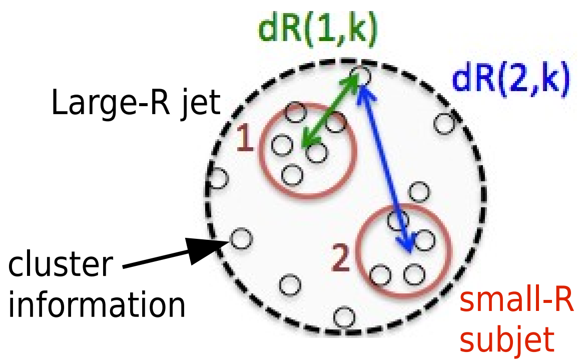
Potential improvements for Run I analysis

Fat jet	Standard analysis
C/A, R=1.5 anti-k_T, R > 0.3	dR (b_1, b_2) anti-k_T R=0.4 → similar jets as subjets
Mass drop criterion Symmetry criterion	$p_T^{b_1}$ $p_T^{b_2}$ m_{bb} $\Delta R(b_1, b_2)$ $\lvert\Delta\eta(b_1, b_2)\rvert$ exploited in BDT?
Filtering  → better energy and angular resolution → hard to calibrate	No equivalent → but anti-k_T R=0.4 already close
B-tagging → might profit from filtering	B-tagging → provides very good background suppression already

Variable	0-Lepton	1-Lepton	2-Lepton
p_T^V		×	×
E_T^{miss}	×	×	×
$p_T^{b_1}$	×	×	×
$p_T^{b_2}$	×	×	×
m_{bb}	×	×	×
$\Delta R(b_1, b_2)$	×	×	×
$\lvert\Delta\eta(b_1, b_2)\rvert$	×		×
$\Delta\phi(V, bb)$	×	×	×
$\lvert\Delta\eta(V, bb)\rvert$			×
H_T	×		
$\min[\Delta\phi(\ell, b)]$		×	
m_T^W		×	
$m_{\ell\ell}$			×
$MV1c(b_1)$	×	×	×
$MV1c(b_2)$	×	×	×
	Only in 3-jet events		
$p_T^{jet_3}$	×	×	×
m_{bbj}	×	×	×

Can we improve?

Substructure techniques



Moment	xAOD Jet attribute names
N-subjettiness	Tau1, Tau2, Tau3, Tau1_wta, Tau2_wta, Tau3_wta
kT splitting scale	Split12, Split23, Split34
zCut	ZCut12, ZCut23, ZCut34
Dipolarity	Dip12, Dip13, Dip23, DipExcl12
Angularity	Angularity
kT Delta R	KtDR
kT Mass drop	Mu12
Planar flow	PlanarFlow
Energy correlations	ECF1, ECF2, ECF3
Thrust	ThrustMin, ThrustMaj
FoxWolfram	FoxWolfram0, FoxWolfram1, FoxWolfram2, FoxWolfram3, FoxWolfram4
Sphericity	Sphericity, Aplanarity
Jet charge	Charge
Pull	PullMag, PullPhi, Pull_C00_, Pull_C01_, Pull_C10_, Pull_C10_
Shower deconstruction	ShowerDeconstructionW, ShowerDeconstructionTop
Q-jets volatility	Volatility

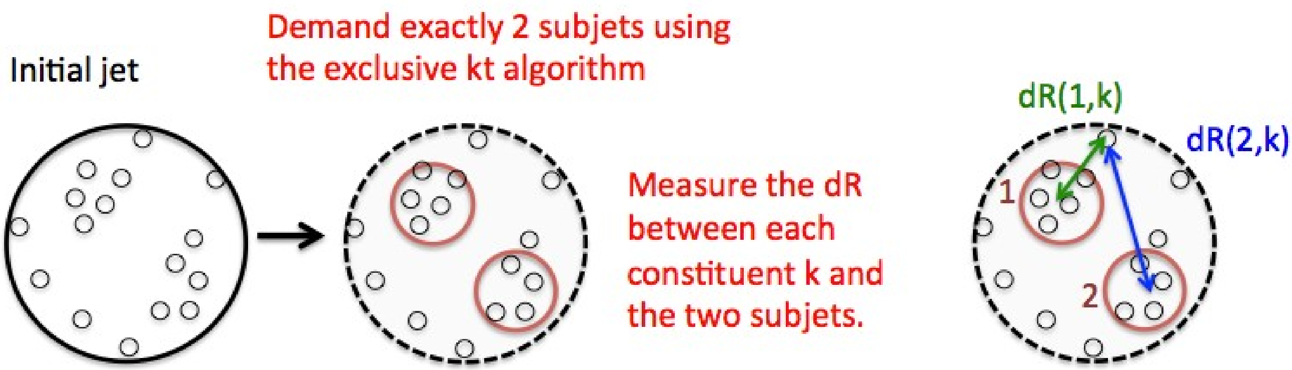
[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetSubstructureTools]

- Fat jets / substructure contain additional information, e.g. shape of (sub)jets
- Generic substructure variables
  [not taylored to a particular decay / tagger]
  - Might have separation power for VH → $b\bar{b}$ analysis
- Additional variable in BDT
  - Simplest approach
    [but for sure not best]
  - Starting point for investigation
    [→ which features can improve S/B]

Promising to study

Example: n-subjettiness

[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetSubstructureVariables]



d_k for cluster k with pT(k):

d_k = pT(k) x min(dR(1,k),dR(2,k))

Normalisation d_0 for jet with radius R:

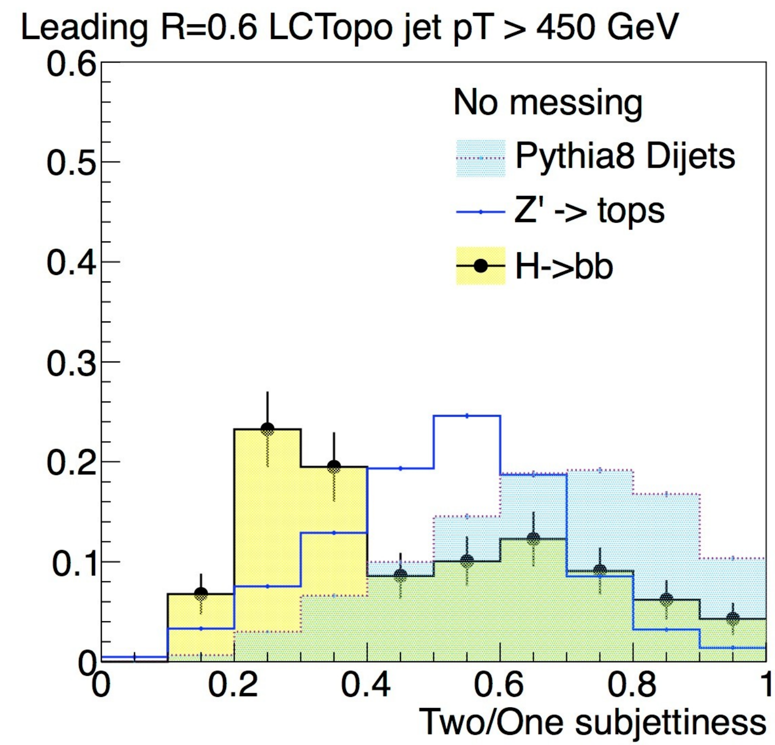
$d_0 = \Sigma(p_T(k) \times R)$

τ_2 for jet with exactly 2 subjets:

$\tau_2 = \Sigma d_k / d_0$

Jet with exactly 2 subjets

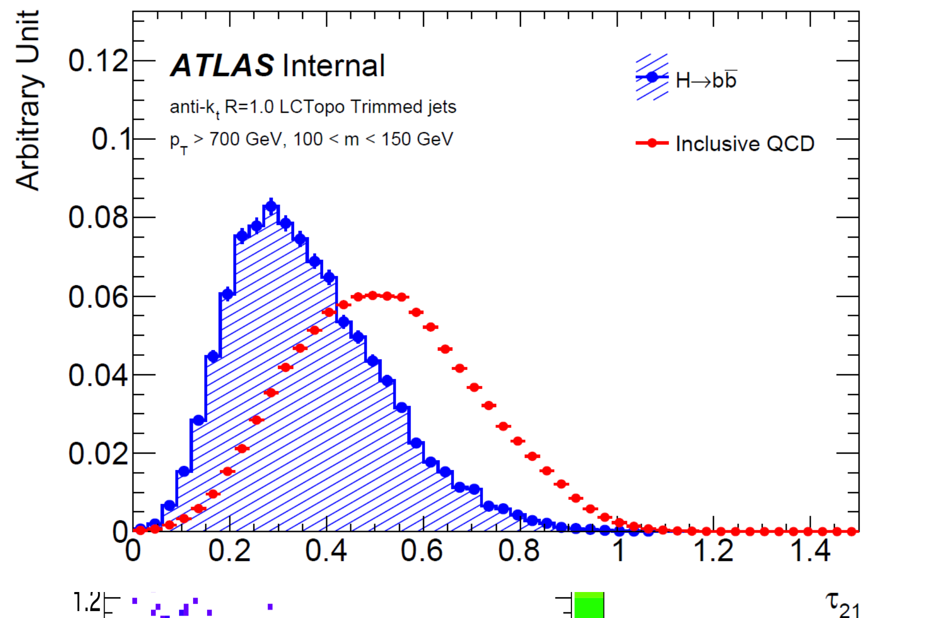
- Calculate d_k of all consitutents
  - → if consituent is close / within subjet: d_k small
  - → if constituent is far from all subjets: d_k large
- Calculate normalisation d_0
- τ_2 is the the two-subjettiness
  - → collimated subjets: τ_2 small
- Ratios of 1-, 2- and 3-subjettiness: τ_{32} ,τ_{21}



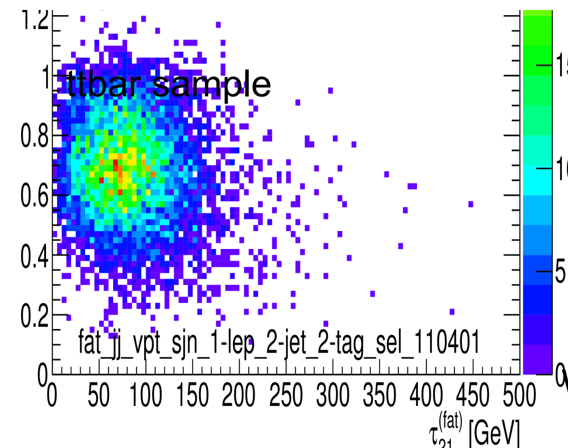
τ_{21}-subjettiness

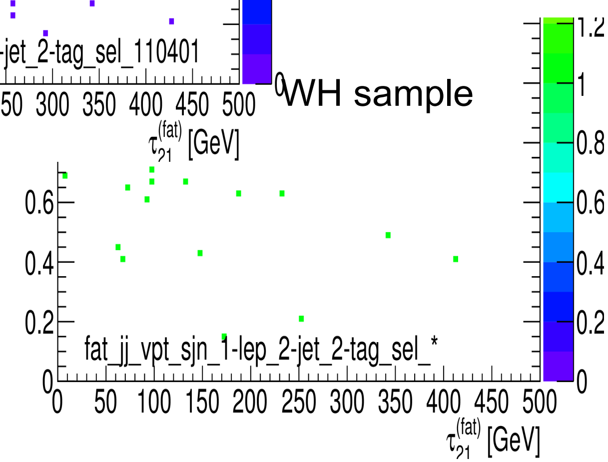
[https://indico.cern.ch/event/346667/]

- Recently presented study
  [Qi Zeng, Michael Kagan, Ariel Schwartzmann]
- p_T > 700 GeV,
  b-tagging on track jets
- τ_{21} gives good rejection against QCD ($g \rightarrow b\bar{b}$) on top of double b-tag
  [yet minor background in VH → $b\bar{b}$]



- Started to study in 13 TeV
  - QCD small background in 1-lep channel
  - No statistics at large p_T^V
    [waiting for full samples]



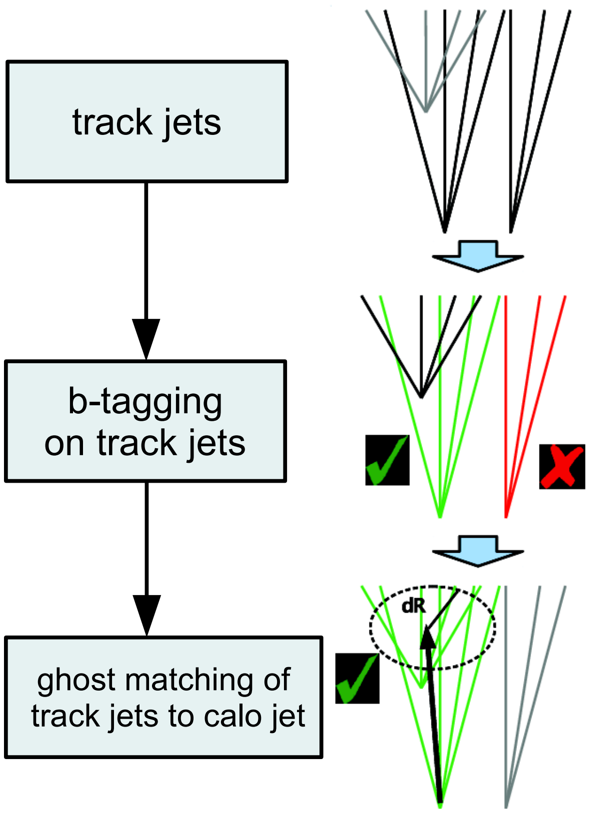


Ongoing work

[proper plot not ready in time (sorry)]

B-tagging

- B-Tagging fundamental part in H→ bb
  - Essential part of standard analysis
    [invariant mass from two tight b-jets]
  - Essential part of Higgs tagging using fat jets
    [identify Higgs decay products]
- Optimized for Run I [c-jet rejection]

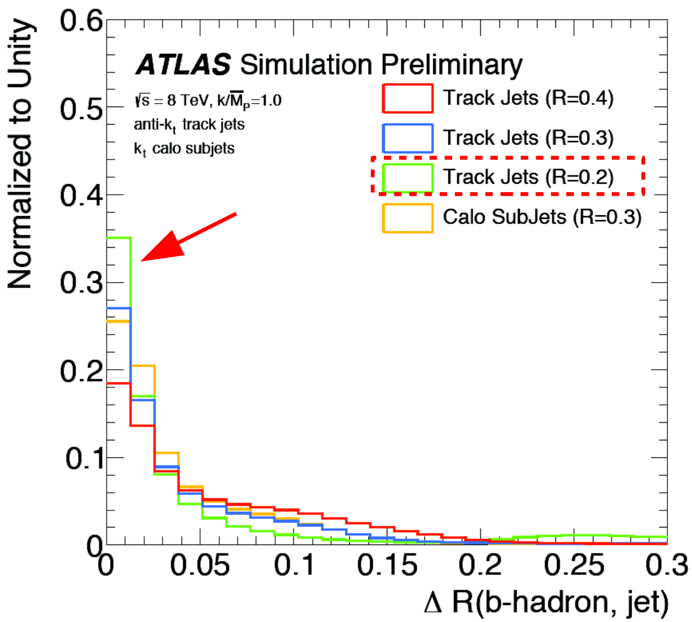


Improvements for Run II

- Track-jet vs. calo-jet
  - Less sensitive to pile-up
  - Better angular resolution
    [useful in boosted topologies]
  - No need for precise jet calibration

No decision for default algorithm yet

  - Both will probably be supported (efficiency points, uncertainties)



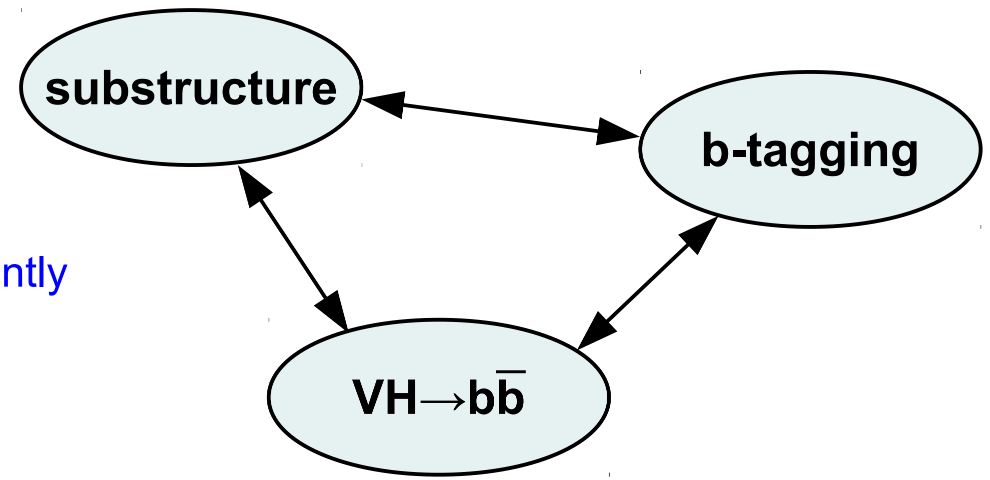
Fields of activity

B-tagging

- Track-jet b-tagging
  → b-tagging only recently available in xAOD
  → Track-jet b-tagging still not available,
  need to re-run (athena)

Substructure

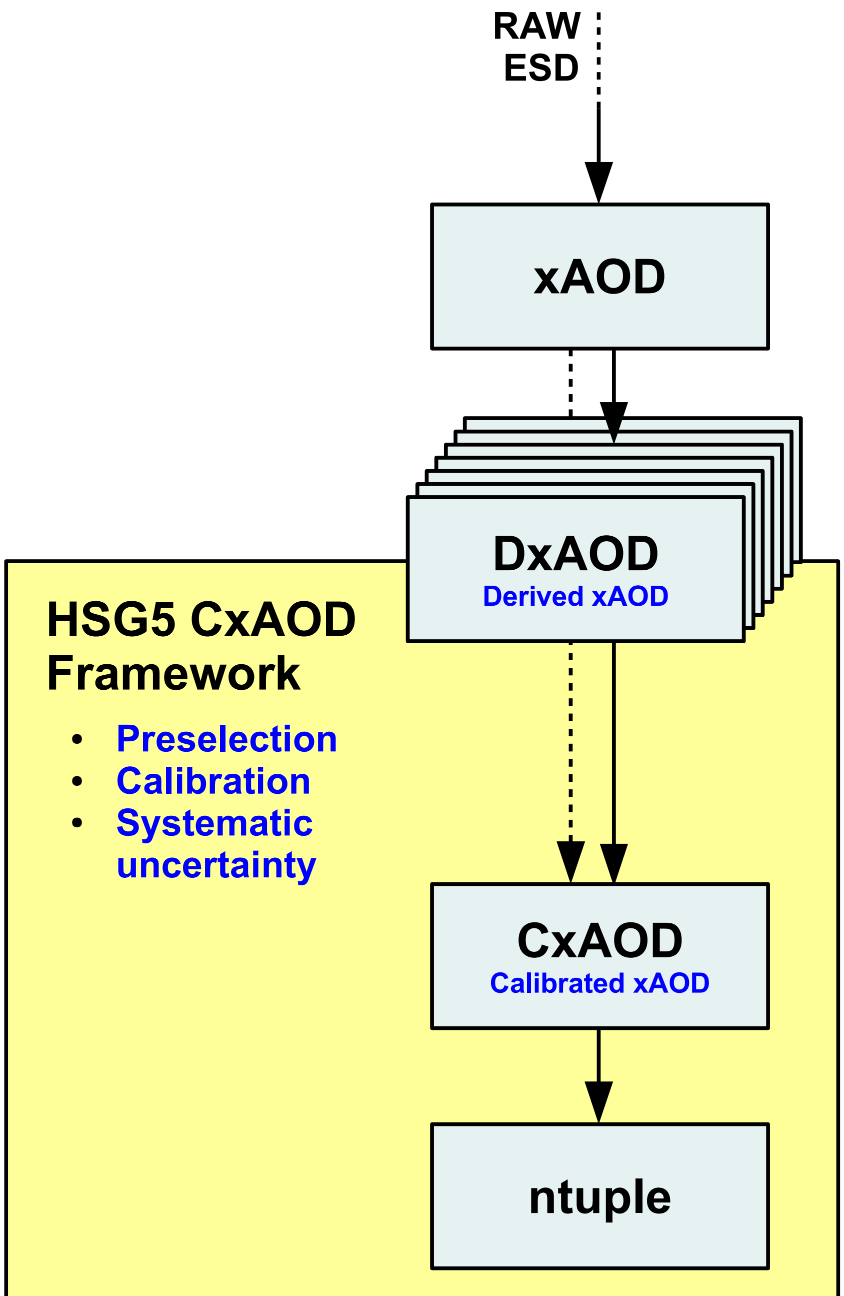
- BDRS algorithm
  Technically most difficult approach
  → Re-run jet algorithm, dual-use tools just recently
- Substructure variables
  → Exist for all jet collections
  → Signal samples only available recently



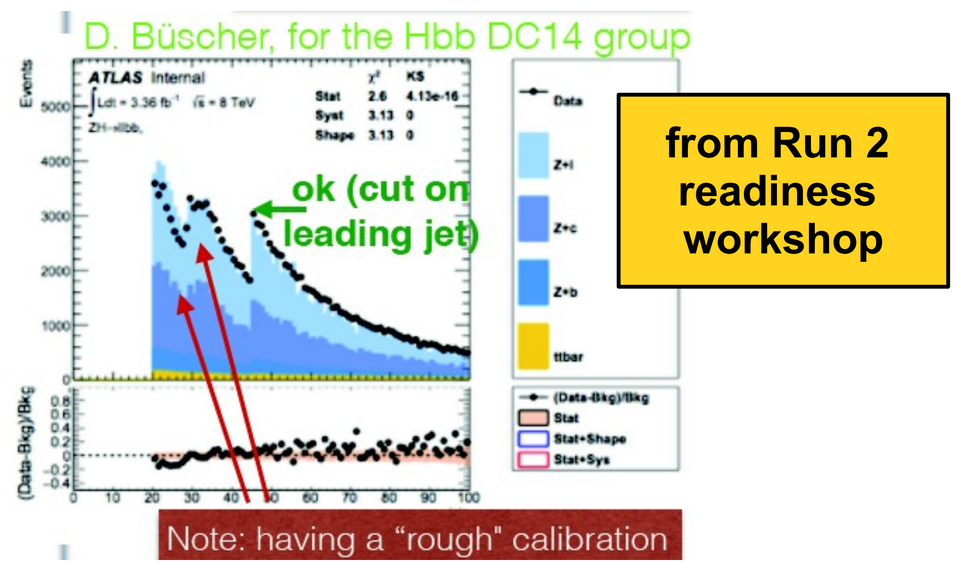
VH → $b\bar{b}$ analysis

- CxAOD Framework
  → Preparation for Run II analysis
  - Implement fat jet collections and substructure variables
  - Reproduce Run I dijet mass analysis
    → serve as reference
  - Study substructure for Run II
- Filter for $t\bar{t}$ production
  → Request from convener, limited work

New analysis model and Higgs framework



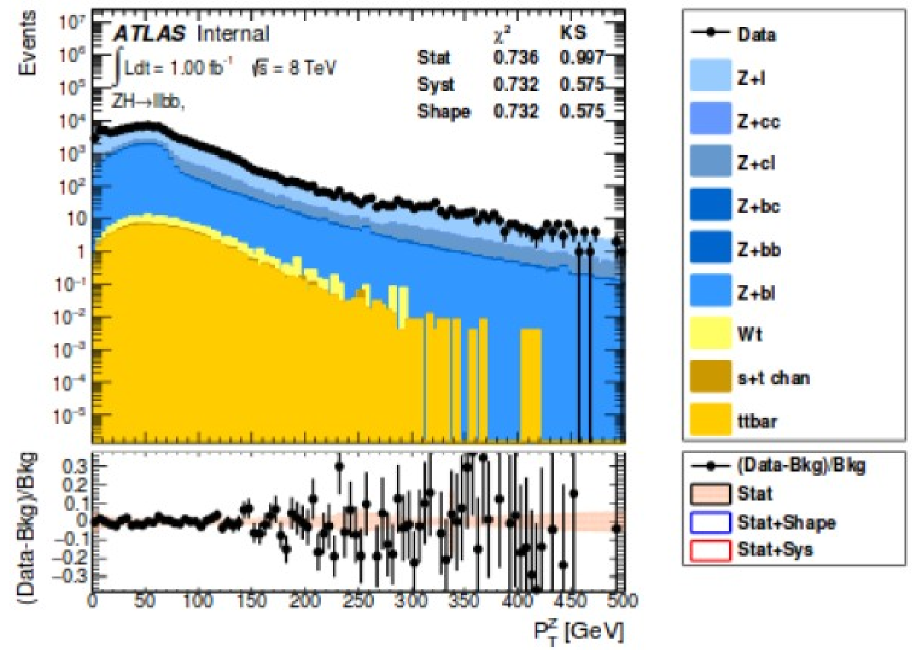
- New Event Data Model (EDM)
  → new file format
  → all tools need(ed) to be adapted
- DxAOD produced centrally
  [derivation framework, train production]
  - Dedicated DxAODs for CP groups
  - DxAOD for 1-lep channel
- CxAOD produced for group
- Data Challenge 14
  - Test of new analysis model
  - HSG5 CxAODFramework



Framework and samples

CxAOD Framework

- Majority of objects / variables available
  - B-Tagging operation points from 8 TeV [not optimised]
  - Trigger [not in xAOD, tools just ready]
  - p_T^{miss} [no tool available?]
  - MC truth [very limited information]
- Preselection, calibration, systematic uncertainty
  [structure created, most CP tools available]
- Validation ongoing
  [almost not been used for studies, yet]



Samples [just finished recently]

Sample	8 TeV			13 TeV		
	xAOD	DxAOD	CxAOD	xAOD	DxAOD	CxAOD
VHbb 0-lep	300k	yes	yes	50k	-	yes
VHbb 1-lep	300k	yes	yes	200k	-	yes
VHbb 2-lep	600k	yes	yes	200k	-	yes
ttbar	30M	yes	yes	10M	-	yes
Single top	10M	yes	yes	2M	-	yes
Z+jet	~ 9M	yes	yes	~ 15M	-	yes
W+jet	~ 16M	yes	yes	~ 15M	-	yes
QCD multijet		*From data*			*From data*	
Data	yes	yes	yes	-	-	-

[production of full statistics about to start]

Splitting and Filtering

[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetGrooming]

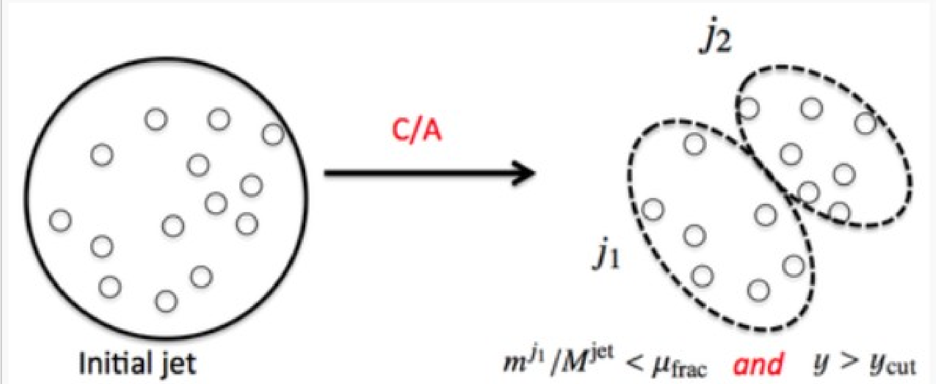
The mass-drop filtering procedure seeks to isolate concentrations of energy within a jet by identifying relatively symmetric subjets, each with a significantly smaller mass than that of their sum.

This technique was developed and optimized using C/A jets in the search for a Higgs boson decaying to two b-quarks.

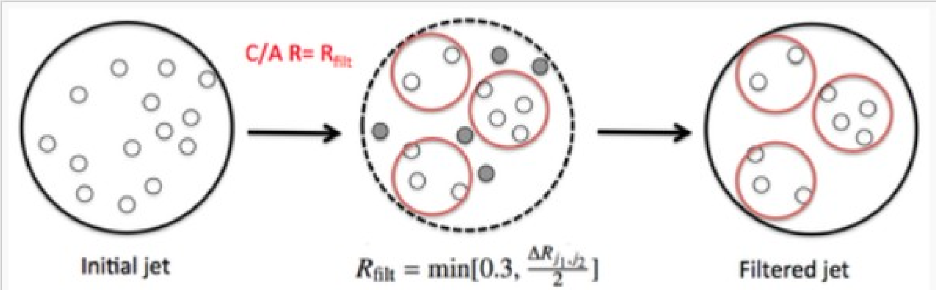
The procedure is applied only to C/A jets since each clustering step of the algorithm combines the two widest angle proto-jets at that point in the shower history.

- **Splitting and Filtering:** J.M. Butterworth, A.R. Davison, R. Mathieu and G.P. Salam (BDRS), "Jet substructure as a new Higgs search channel at the LHC", Phys. Rev. Lett., 100, 242001, 2008, arXiv:0802.2470

- **Step 1:** Undo the last stage of the C/A clustering so that the jet "splits" into two subjets, j_1 and j_{2} , ordered such that the mass of j_1 is larger.
- **Step 2:** Require that $m^{j1}/m^{jet} < \mu_{frac}$ (the **"mass drop"** criterion)
- **Step 3:** Require that $frac{min}[(p_{T}^{j1})^2,(p_{T}^{j2})^2]}{(m^{jet})^2}\times \Delta R(j1,j2)^{2} > y_{cut}$ (the **"symmetry"** criterion)
- **Step 4:** The constituents of $j1$ and $j2$ are reclustered using the C/A algorithm with $R_{filt} < \Delta R(j1,j2)$, where $R_{filt} = min [0.3,\frac{\Delta R(j1,j2)}{2}]$.



Step 1 of the splitting and filtering algorithm.



Step 2 of the splitting and filtering algorithm.

Trimming

[https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetGrooming]

The trimming algorithm takes advantage of the fact that contamination from pile-up, multiple parton interactions (MPI), and initial-state radiation (ISR) in the reconstructed jet is often much softer than the outgoing partons associated with the hard-scatter and their final-state radiation (FSR).

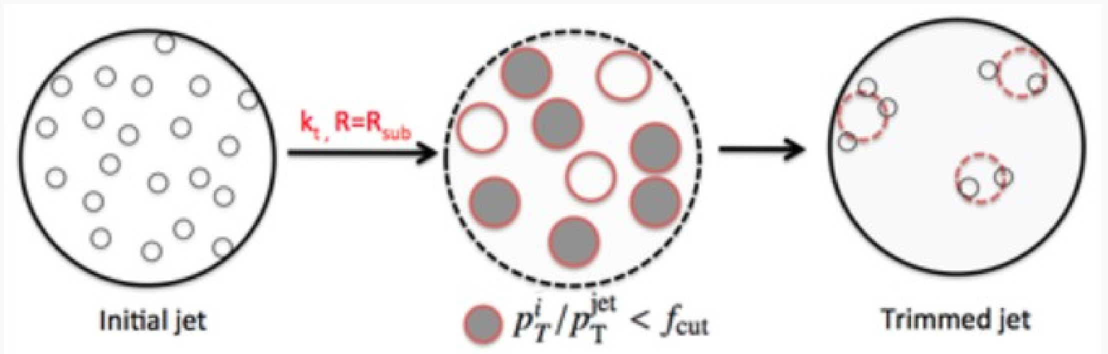
The ratio of the pT of the constituents to that of the jet is used as a selection criterion.

Completely removing the softer components from the final jet is possible as there is generally minimal spatial overlap of the soft additional radiation from pile-up, MPI, and ISR with the hard-scatter decay products.

As the primary effect of pile-up, for example, is additional low-energy topo-clusters as opposed to additional energy being added to topo-clusters from hard-scatter particles, this allows a relatively simple jet energy offset correction for smaller radius jets (R=0.4, 0.6) as a function of the number of primary reconstructed vertices.

- **Trimming:** D. Krohn, J. Thaler and L.-T. Wang, "Jet Trimming", arXiv:0912.1342

- **Step 1:** Uses the inclusive k_{t} algorithm to create subjets of size R_{sub} from the constituents of a jet.
- **Step 2:** Any subjets with $\pti/\ptjet < \fcut$ are removed, where pti is the transverse momentum of the i^{th} subjet, and f_{cut} is a parameter of the method



Schematic of the trimming algorithm.

Pruning

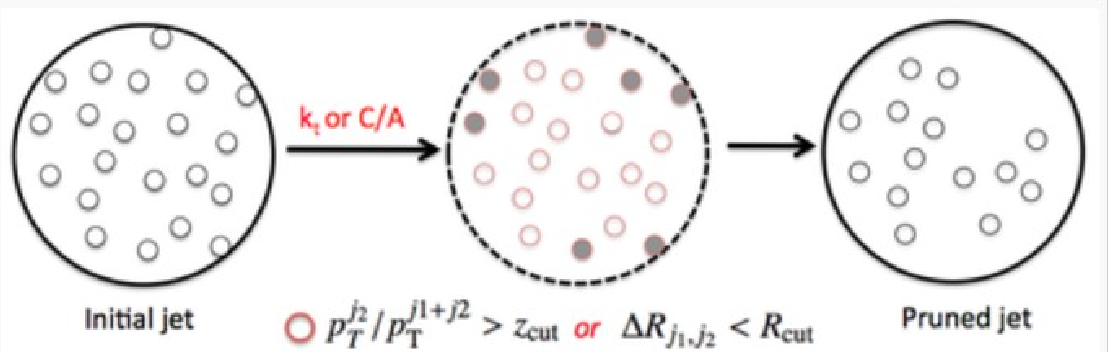
The pruning algorithm is similar to trimming in that it removes constituents with a small relative pT, but additionally utilizes a wide-angle radiation veto.

The design of the procedure is such that the selections are based on each successive recombination of the C/A or k_{t} algorithm, depending on which is used to implement the pruning.

The pruning procedure is invoked at each successive recombination of the jet algorithm used (either C/A or k_{t}), based on the branching at each point in the jet reconstruction, and as such does not require the reconstruction of subjets. This results in definitions of the terms "wide-angle" or "soft" that are not directly related to the original jet but rather to the proto-jets formed in the process of rebuilding the pruned jet.

- **Pruning:** S.D. Ellis, C.K. Vermilion and J.R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches", arXiv:0912.0033

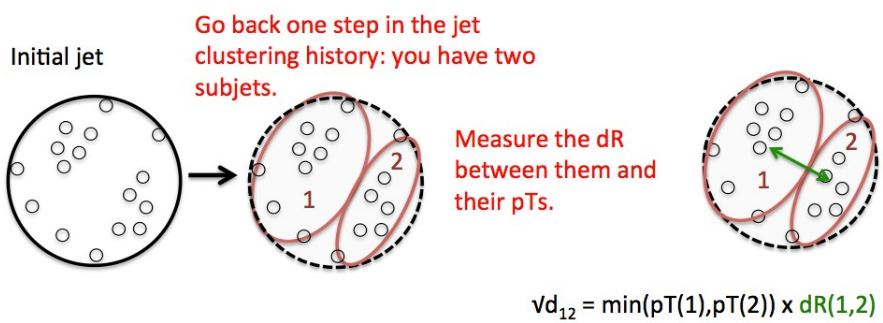
- **Step 1:** Run either the C/A or k_{t} recombination jet algorithm on the constituents found by any jet finding algorithm.
- **Step 2:** At each recombination step with constituents j_1 and j_{2} (where $p_T^{j_1} > p_T^{j_2}$), require that $p_{T}^{j_{2}}/p_{T}^{j1+j2} > z_{cut}$ or $\Delta R_{j_{1}, j_{2}} < R_{cut}\times\frac{2 m^{jet}}{p_{T}}$.
- **Step 3:** Merge j_{2} with j_1 if the above criteria are met, otherwise, discard j_{2} and continue with the algorithm.



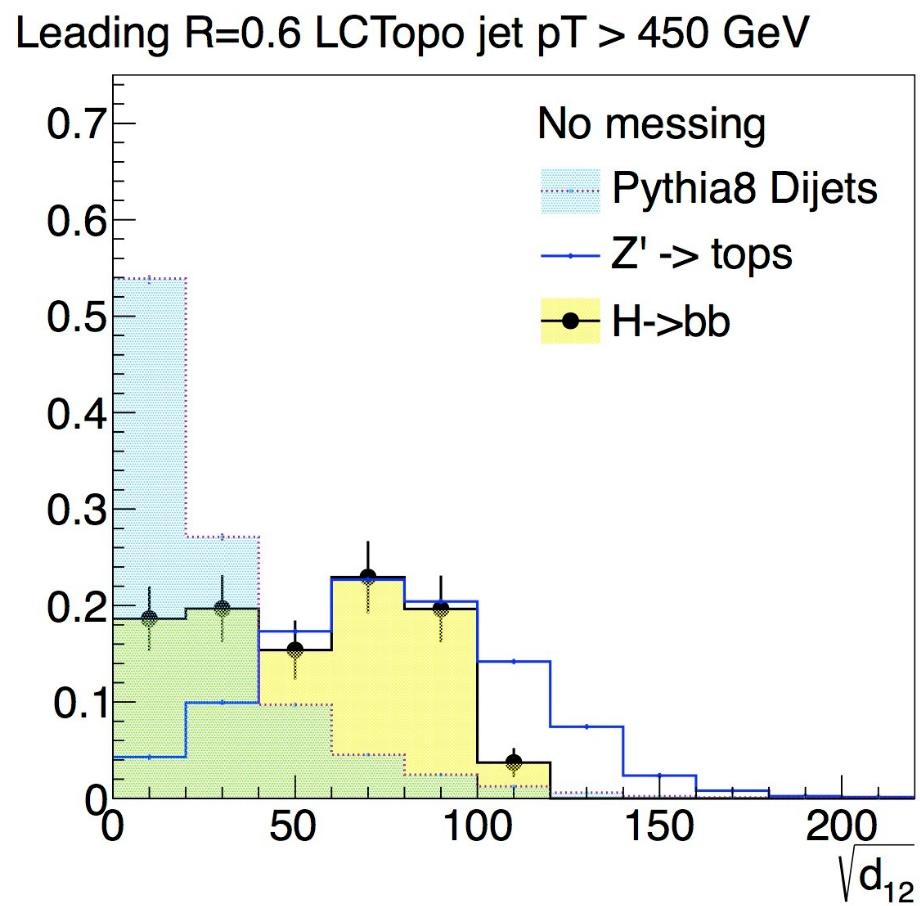
Schematic of of the pruning algorithm.

k_T splitting scale

- also know as sqrt(d_ij)



- If the distance between the subjets is large, $\sqrt{d_{12}}$ is large.
- If the softer of the two subjets in the last clustering has high pT, then $\sqrt{d_{12}}$ is large.
- Both these things indicate large $\sqrt{d_{12}}$ in symmetric two body decays.



BDRS algorithm

J.M. Butterworth, A.R. Davison, R. Mathieu and G.P. Salam:
Jet substructure as a new Higgs search channel at the LHC
Phys. Rev. Lett., 100, 242001, 2008, arXiv:0802.2470 [BDRS algorithm or mass drop filter]

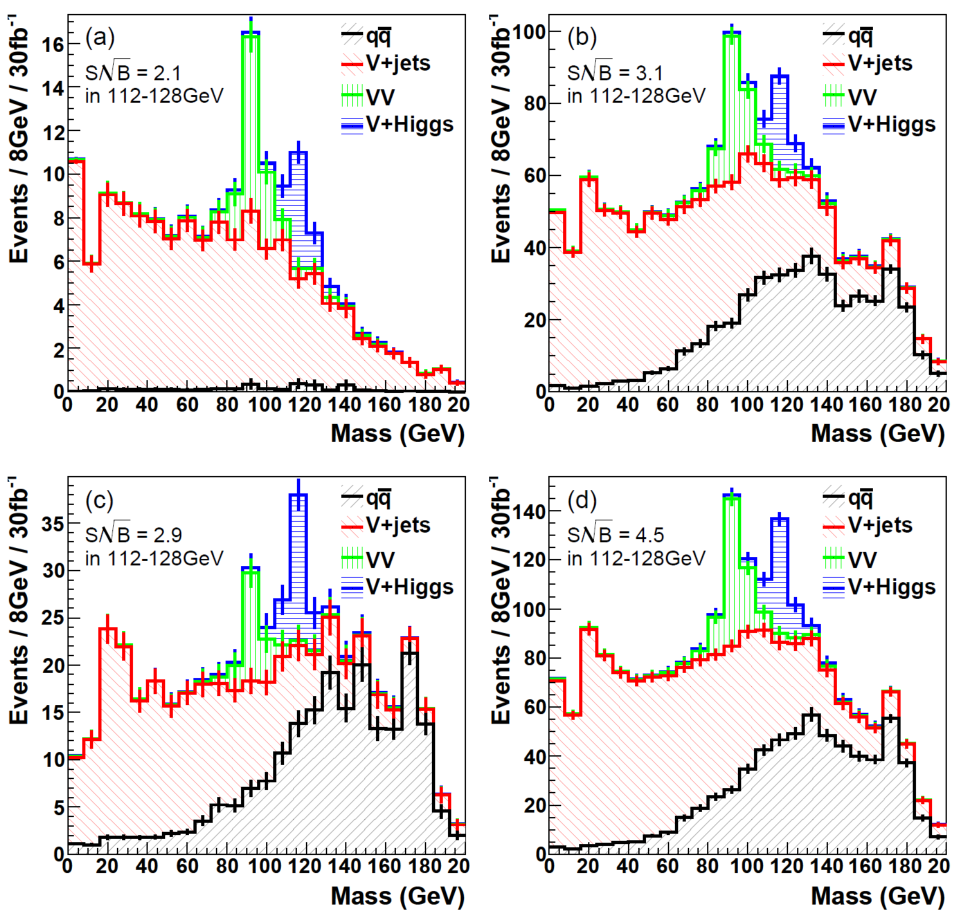


FIG. 2: Signal and background for a 115 GeV SM Higgs simulated using HERWIG, C/A MD-F with $R = 1.2$ and $p_T > 200$ GeV, for 30 fb^{-1}. The b tag efficiency is assumed to be 60% and a mistag probability of 2% is used. The $q\bar{q}$ sample includes dijets and $t\bar{t}$. The vector boson selections for (a), (b) and (c) are described in the text, and (d) shows the sum of all three channels. The errors reflect the statistical uncertainty on the simulated samples, and correspond to integrated luminosities > 30 fb^{-1}.

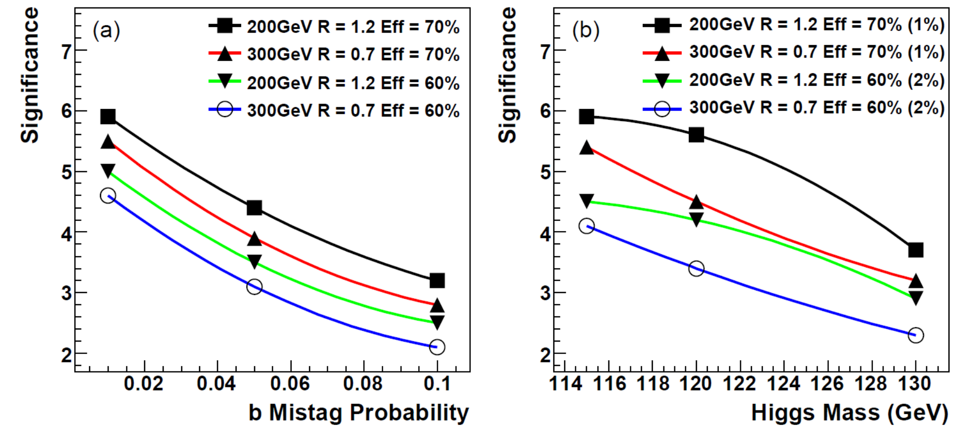


FIG. 3: Estimated sensitivity for 30 fb^{-1} under various different sets of cuts and assumptions (a) for $m_H = 115$ GeV as a function of the mistag probability for b-subjets and (b) as a function of Higgs mass for the b-tag efficiency (mistag rates) shown in the legend. Significance is estimated as signal/$\sqrt{\text{background}}$ in the peak region.