Searching For All-Hadronic SUSY With Jet Substructure

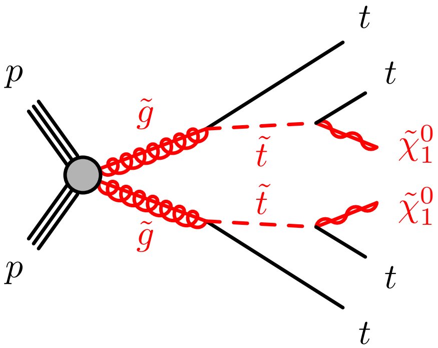
US LUA Lightning Round

Maximilian Swiatlowski
Enrico Fermi Institute, University of Chicago

13 November, 2015

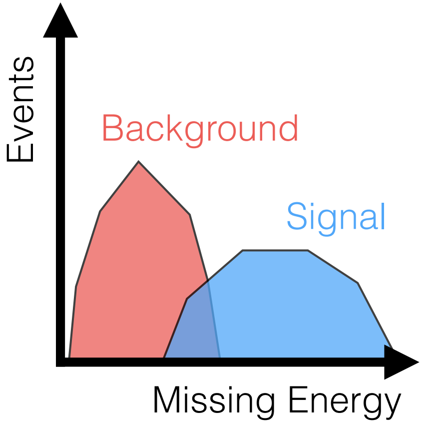
SUSY Signatures

Gluino ($\tilde{g}$) production is potentially easiest way to search for $\tilde{t}$



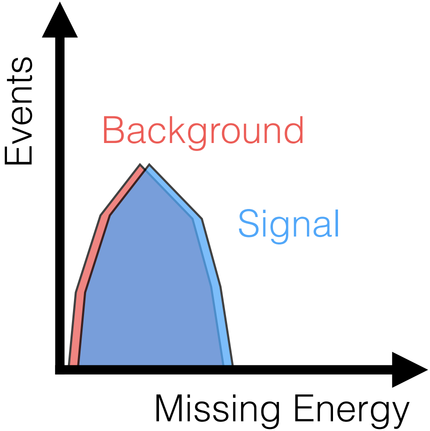
- Many different possible SUSY signatures to search for!
- Most commonly, look for high p_T jets, leptons, and **missing energy**
  - Most models assume *R*-parity, which means the Lightest Supersymmetric Particle ($\tilde{\chi}_1^0$) **is stable**
  - This neutralino **does not interact**: escapes detection, appears as E_T^{miss}
- But what if the LSP decays? Existing searches will not work!

SUSY Signatures



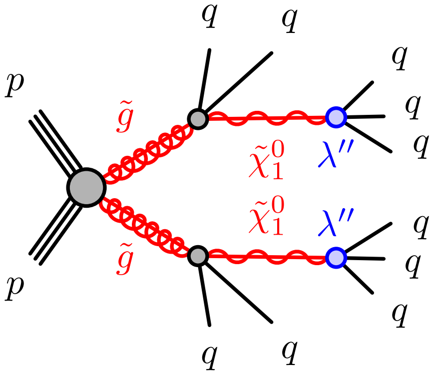
- Many different possible SUSY signatures to search for!
- Most commonly, look for high p_T jets, leptons, and **missing energy**
  - Most models assume R-parity, which means the Lightest Supersymmetric Particle ($\tilde{\chi}_1^0$) **is stable**
  - This neutralino **does not interact**: escapes detection, appears as E_T^{miss}
- But what if the LSP decays? Existing searches will not work!

SUSY Signatures



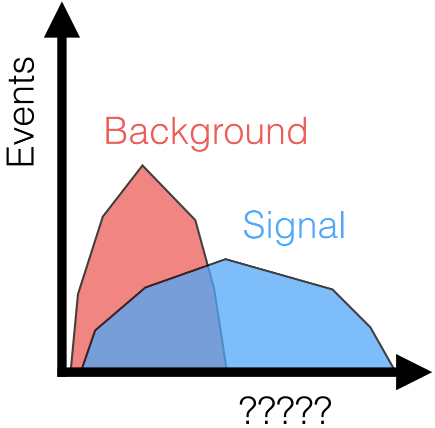
- Many different possible SUSY signatures to search for!
- Most commonly, look for high p_T jets, leptons, and **missing energy**
  - Most models assume R-parity, which means the Lightest Supersymmetric Particle ($\tilde{\chi}_1^0$) **is stable**
  - This neutralino **does not interact**: escapes detection, appears as E_T^{miss}
- But what if the LSP decays? Existing searches will not work!

- Consider $\tilde{g}$ pair production, decaying with $\tilde{g} \to qq\tilde{\chi}_1^0$, $\tilde{\chi}_1^0 \to 3q$: R-parity violating!



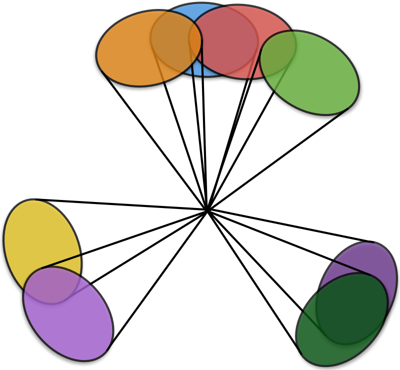
- Final state has huge number of quarks!
  - Between **10** (light quarks only) and **22** (top decays) partons
  - **Extremely difficult background estimation:** high-mass extremes of QCD are difficult to model
  - No source of E_T^{miss}: **need other discrimination handles**

- Consider $\tilde{g}$ pair production, decaying with $\tilde{g} \rightarrow qq\tilde{\chi}_1^0$, $\tilde{\chi}_1^0 \rightarrow 3q$: *R*-parity violating!



- Final state has huge number of quarks!
  - Between **10** (light quarks only) and **22** (top decays) partons
  - **Extremely difficult background estimation:** high-mass extremes of QCD are difficult to model
  - No source of E_T^{miss}: **need other discrimination handles**

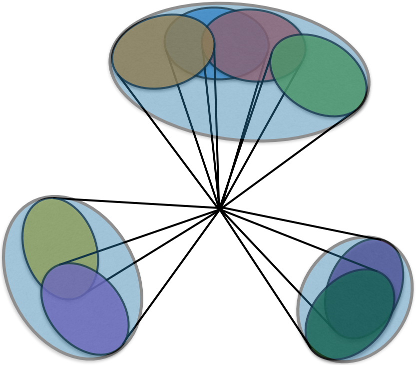
Accidental Substructure



- High multiplicity leads to **significant "accidental" overlaps**

→ We can use **large jets** to capture these overlaps: jets have **mass**

- We can use this structure to search for new physics!

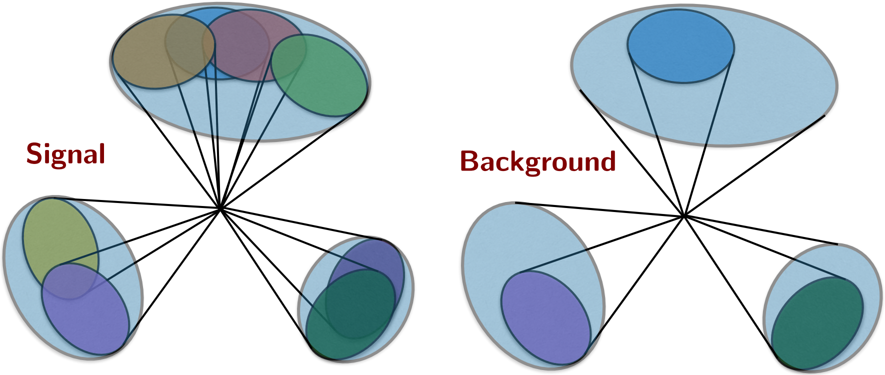


- High multiplicity leads to **significant "accidental" overlaps**

$\rightarrow$ We can use **large jets** to capture these overlaps: jets have **mass**

- We can use this structure to search for new physics!

Accidental Substructure



- High multiplicity leads to **significant "accidental" overlaps**

→ We can use **large jets** to capture these overlaps: jets have **mass**

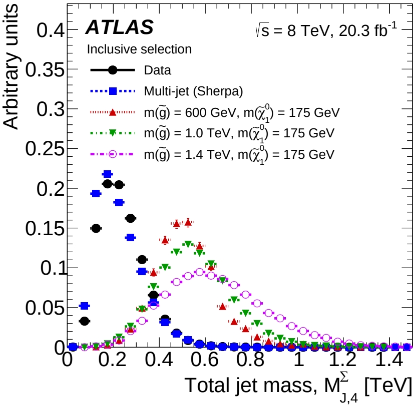
- We can use this structure to search for new physics!

- First proposed by Wacker et al., arXiv:1202.0558
- **Mass** comes when combining **widely spaced** particles
  - → Jets with substructure have **high mass**!
- Define our sensitive variable as:

$$M_J^\Sigma = \sum_{i=1}^{4} M_J^i$$

- **Many different signals** look very different from **background**
  - → But there are some challenges to using large jet masses!

SUSY-2013-07



NB: only consider *trimmed* $R = 1.0$ jets with $p_T > 100$ GeV, $|\eta| < 2.5$

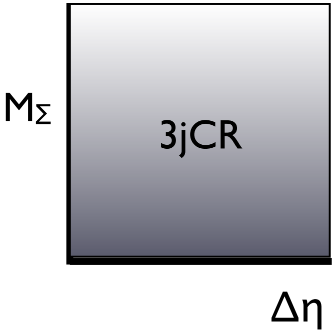
Background Estimation Strategy

Search for new physics in 4-jet regions

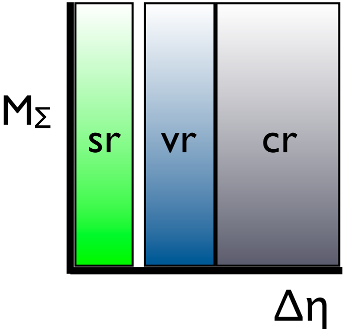
Estimate background in 3-jet control regions: QCD dominates

$|\Delta\eta(j_1, j_2)|$ allows for splitting 4-jet regions for validation

3-jet Region



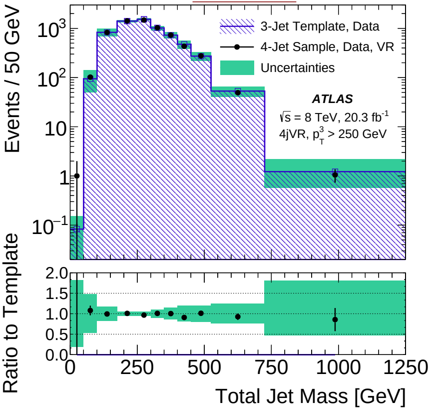
4-jet Regions



- Build *jet mass templates* here: measure $M(p_T, \eta)$
- Various sub-regions for validation and search

Validation region is an area we expect no signal (due to extra cuts on $\Delta\eta$), but can check performance of the templates

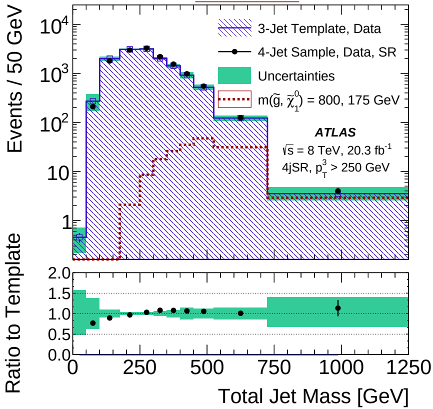
SUSY-2013-07



- **Templates**– background prediction– agree very well with **data**
- **Uncertainties** come from template method, and cross-checks in control regions

Opening the Box

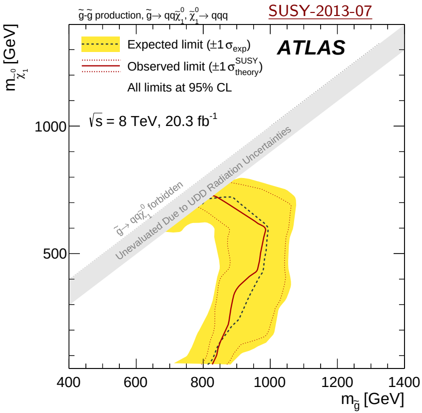
SUSY-2013-07



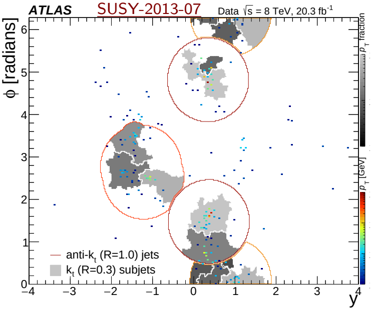
- Open the box: compare signal region prediction to data
- **Templates** agree very well with **data**: no sign of **BSM**
- Different bins provide different $S/\sqrt{B}$: multi-bin fit improves final limits

Limits

- Interpret results in the $m_{\tilde{g}}$, $m_{\tilde{\chi}_1^0}$ plane: can we exclude a signal with various properties?
- **Strong limits**: out to $m_{\tilde{g}} \sim 1$ TeV
- An entirely new final SUSY state has been explored
  - Light gluinos could have been hiding here— but no evidence for them
- **Jet substructure is the key to the analysis**



- We have performed a new analysis searching for gluino pair production with hadronically decaying LSP's
- New **jet substructure reconstruction** techniques are at the heart of the analysis:
  - **Accidental substructure** provides a new variable, M_J^Σ, with powerful background suppression
  - **Jet substructure templates** provide a mechanism for measuring QCD backgrounds

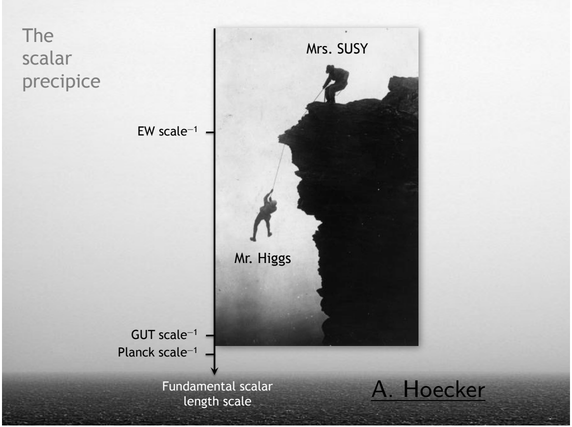


Thank You For Your Attention!

Backup

Naturalness with Supersymmetry

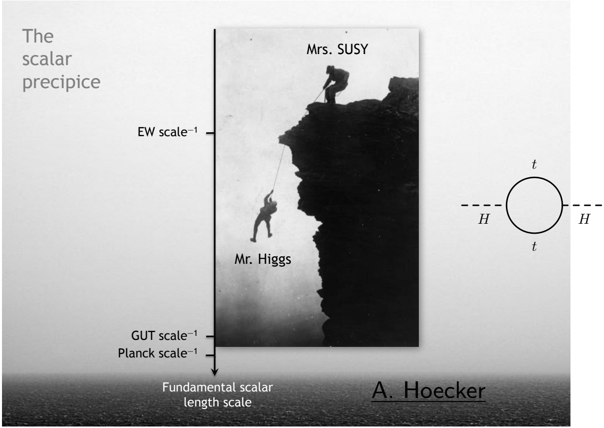
One way to think about the Hierarchy problem: what is preventing the Higgs mass from "falling" to very high mass?



SUSY is a framework which introduces partners for every SM particle

Stop quark ($\tilde{t}$), gluino ($\tilde{g}$), and neutralino ($\tilde{\chi}_1^0$) are the most important

One way to think about the Hierarchy problem: what is preventing the Higgs mass from "falling" to very high mass?

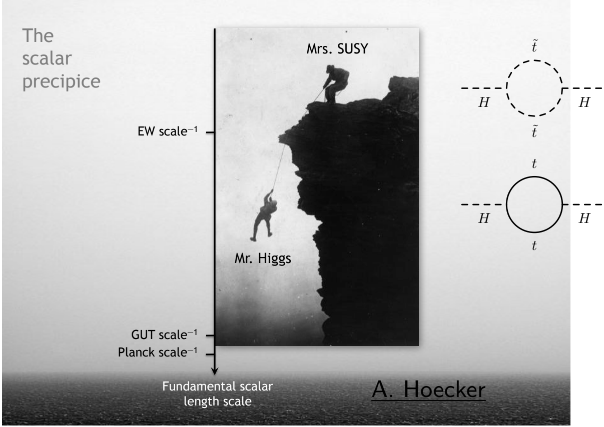


SUSY is a framework which introduces partners for every SM particle

Stop quark ($\tilde{t}$), gluino ($\tilde{g}$), and neutralino ($\tilde{\chi}_1^0$) are the most important

Naturalness with Supersymmetry

One way to think about the Hierarchy problem: what is preventing the Higgs mass from "falling" to very high mass?

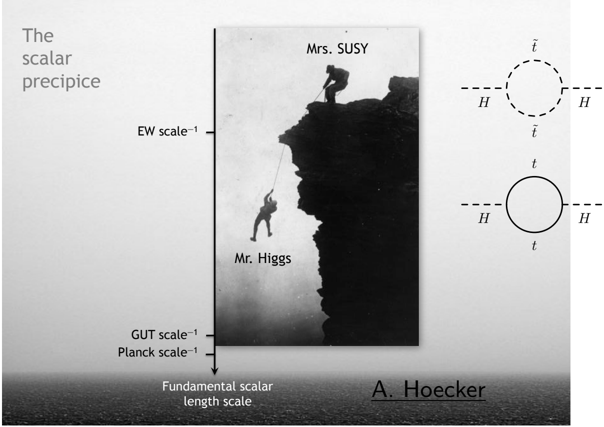


SUSY is a framework which introduces partners for every SM particle

Stop quark ($\tilde{t}$), gluino ($\tilde{g}$), and neutralino ($\tilde{\chi}^0_1$) are the most important

Naturalness with Supersymmetry

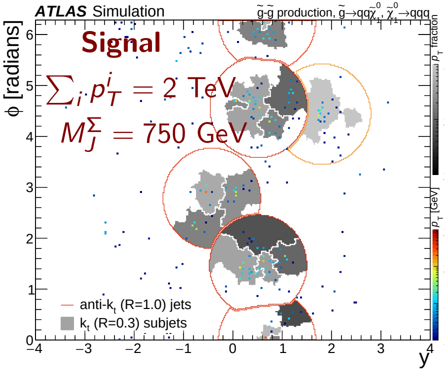
One way to think about the Hierarchy problem: what is preventing the Higgs mass from "falling" to very high mass?

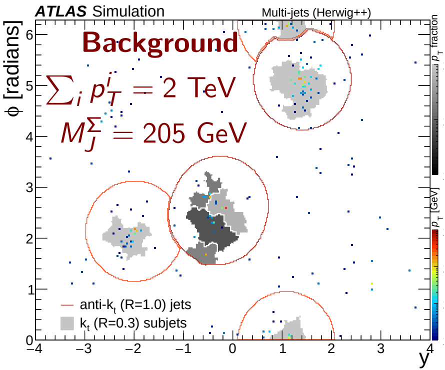


SUSY is a framework which introduces partners for every SM particle

Stop quark ($\tilde{t}$), gluino ($\tilde{g}$), and neutralino ($\tilde{\chi}_1^0$) are the most important

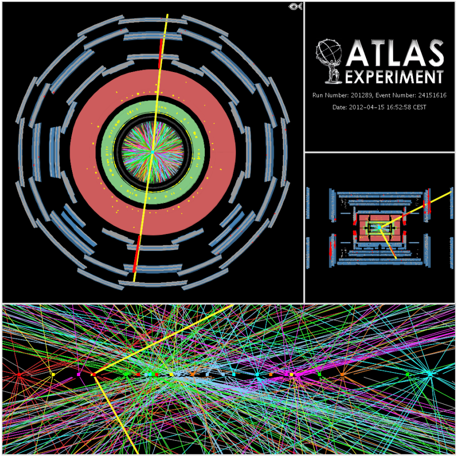
Total Jet Mass: an Example





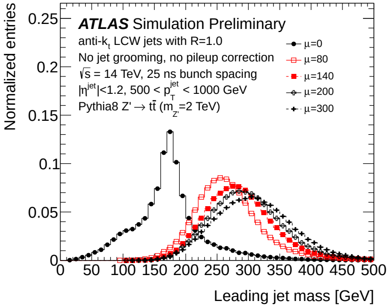
- Typical kinematic variable used for analyses is $H_T = \sum_i p_T^i$
- **Structure** in these events is different, even if energies are similar

Pileup at High Luminosity

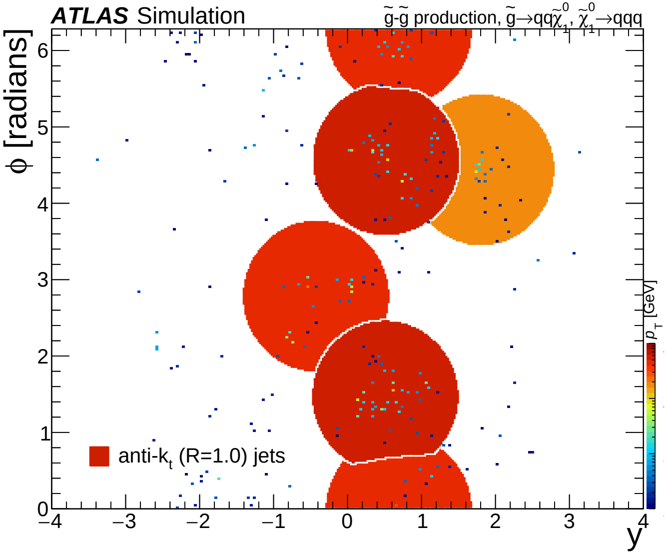


- **Pileup** are the extra un-interesting simultaneous collisions which **contaminate** the interesting collisions we are trying to study
- Compare **no pileup** ($\mu = 0$), to pileup conditions
- Same events– but contamination **washes out the mass peak**
- How can we remove this contamination?

Pileup at High Luminosity



- **Pileup** are the extra un-interesting simultaneous collisions which **contaminate** the interesting collisions we are trying to study
- Compare **no pileup** ($\mu = 0$), to pileup conditions
- Same events– but contamination **washes out the mass peak**
- How can we remove this contamination?

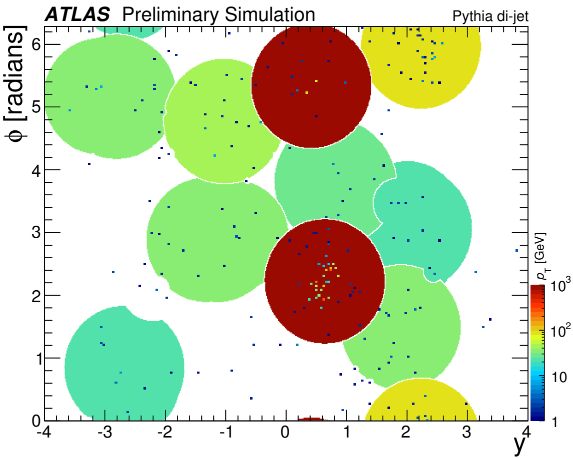


Using large-R jets reduces multiplicity: **increases overlaps**

But how do you control for pileup contamination in these jets?

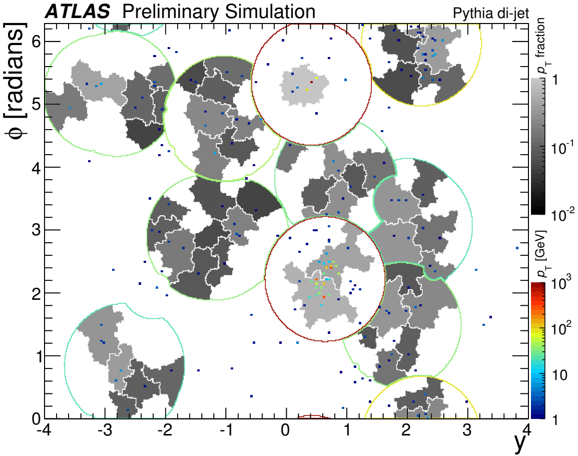
Jet Trimming

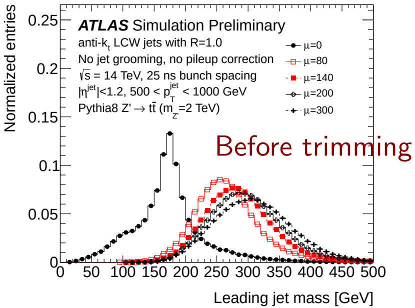
To clean up large-R jets, use *jet trimming*

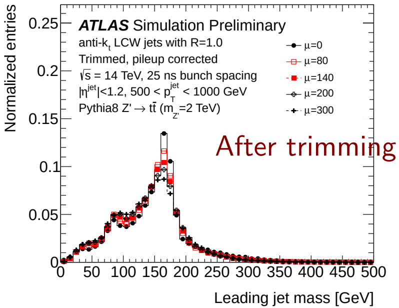


Jet Trimming

To clean up large-R jets, use *jet trimming*







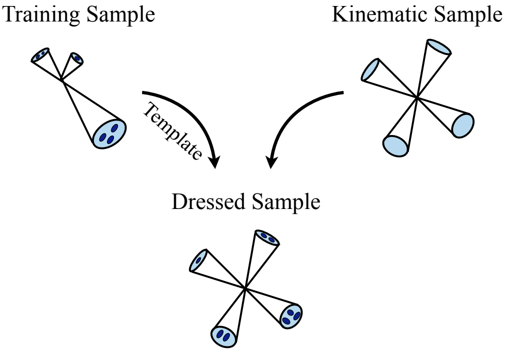
- Before trimming, **large dependence of mass on pileup**
- Trimming mostly **removes pileup dependence** of jet mass!

Jet Substructure Templates

J. Wacker et al

Measure jet mass in **training sample**: $M(p_T, \eta)$

Exactly 3-jet region

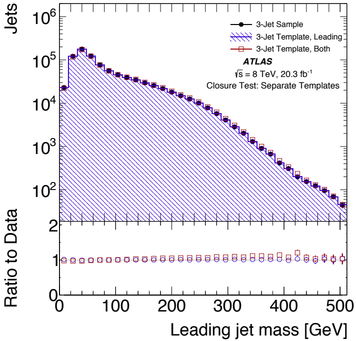


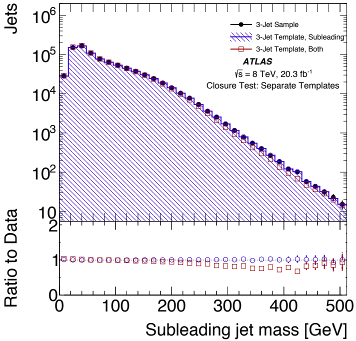
Measure kinematics from **signal region**: p_T, η

$\geq$ 4-jet region

Use kinematics from signal region to **predict mass** in signal region: M

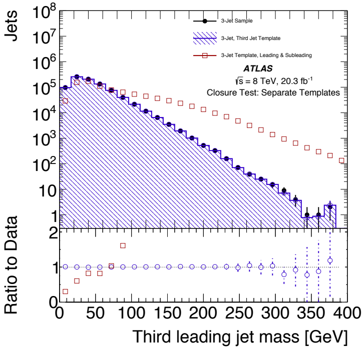
Closure Tests in 3jCR





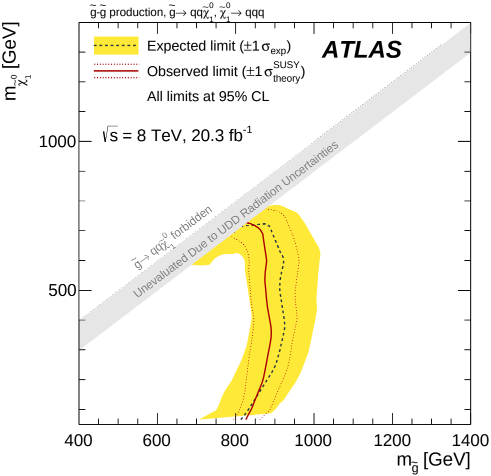
- Train templates separately or inclusively?
- Better agreement between template and data when trained separately
  - Quark/gluon composition different between leading and subleading
  - Averaging over differences with template leads to small bias

Closure Tests in 3jCR

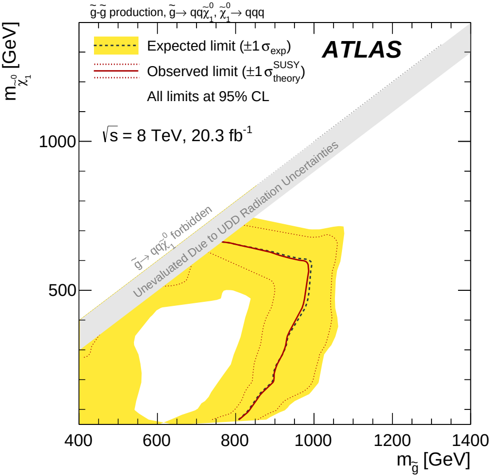


- Third jet is very different from the leading two
- Template trained with third jet only shows large differences from leading two jets
  - 4j regions have four jets: use the template from the third jet to predict fourth jet mass

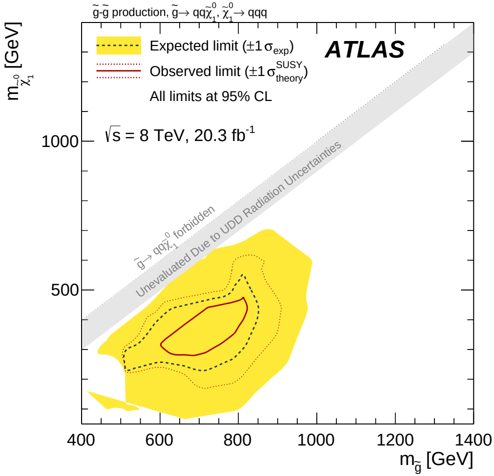
SR100 Limits



SR250 Limits



SR1 Limits

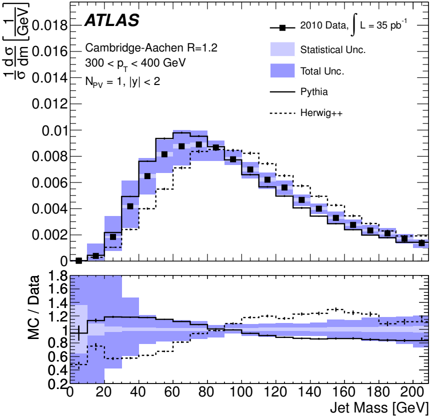

$\tilde{g}$-$\tilde{g}$ production, $\tilde{g} \to qq\tilde{\chi}^0_1$, $\tilde{\chi}^0_1 \to qqq$
Expected limit ($\pm 1\sigma_{exp}$)
Observed limit ($\pm 1\sigma^{SUSY}_{theory}$)
All limits at 95% CL
ATLAS
$\sqrt{s}$ = 8 TeV, 20.3 fb^{-1}
$\tilde{g} \to qq\tilde{\chi}^0_1$ forbidden
Unevaluated Due to UDD Radiation Uncertainties
$m_{\tilde{\chi}^0_1}$ [GeV]
$m_{\tilde{g}}$ [GeV]
1000
500
400
600
800
1000
1200
1400

A Few Words on Uncertainties

- Data-driven technique eliminates most typical sources of background uncertainty (JES, JMS, etc. apply only to signal)
- Three main data-driven sources of uncertainty instead:
  1. **Variance**: statistical uncertainty propagated from 3jCR
     - Statistical uncertainty from training sample, assessed via bootstrapping
     - 5-10% at low M_J^{Σ}, 20-40% at high M_J^{Σ}
  2. **Bias**: size of oversmoothing of templates
     - Controlled by size of kernel– chosen to be much smaller than variance
     - Assessed by making template out of templates, accurate to first order
     - Typically $< 5\%$
  3. **Non-closure**: disagreement in 4jCR
     - For SR100 regions, use full size of reweighting
     - For SR250 regions, use data/prediction disagreement
     - 5% at low M_J^{Σ}, 15% at high M_J^{Σ}

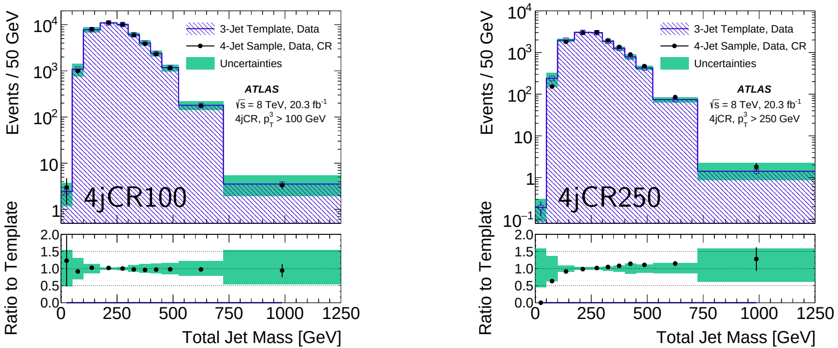
Background Estimation: The Challenge

- Background estimation is **extremely challenging**
  - Need prediction for masses of leading 4 jets— but MC is notoriously unreliable for this
- Data-driven technique is critical!



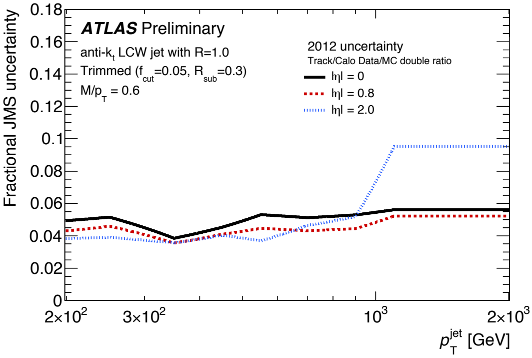
Poor data/MC agreement: difficult to model these effects!

Tests (and uncertainties) in 4jCR ($\Delta\eta > 1.4$)



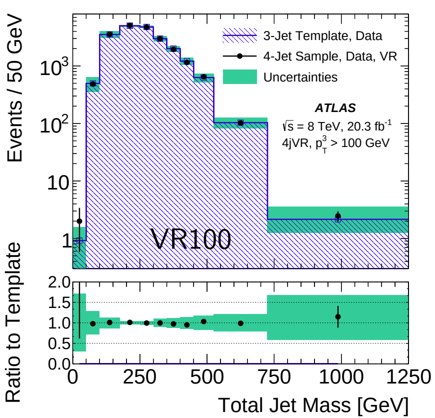
- Templates agree very well with data, within uncertainties
- 4jCR100 has **reweighting** applied: slight disagreement between data/prediction is corrected using 4jCR
  - Plot here **after** correction: looks very good!
- 4jCR250 requires no correction: good agreement to begin

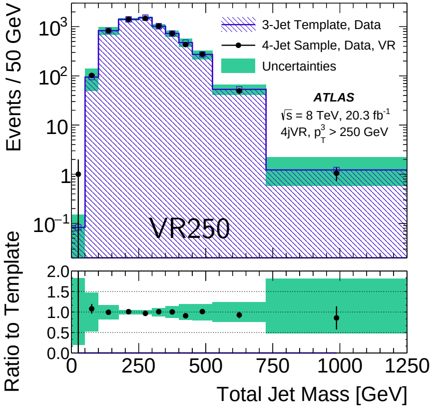
A Few More Words on Uncertainties



- Largest uncertainties on the signal are from **Jet Mass Scale**
  - Assessed using track-to-calorimeter ratios: accurate to 5%
  - Leads to up to 50% acceptance effects at high M_J^Σ
- Other uncertainties– JES, JER, JMR– are subdominant

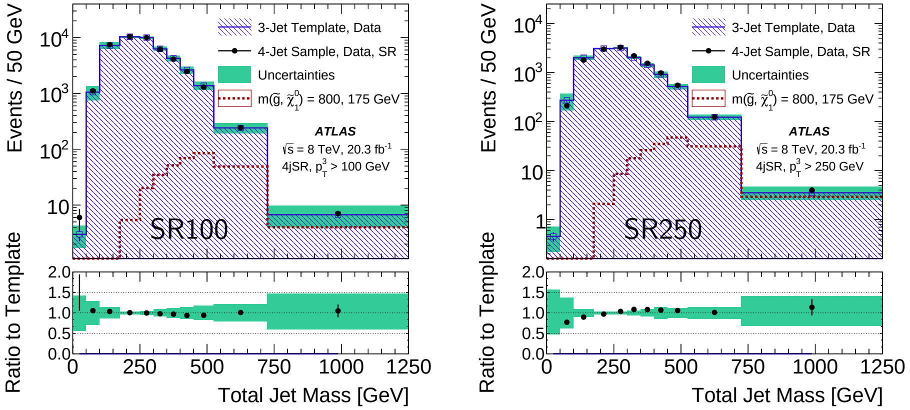
Testing in the Validation Region





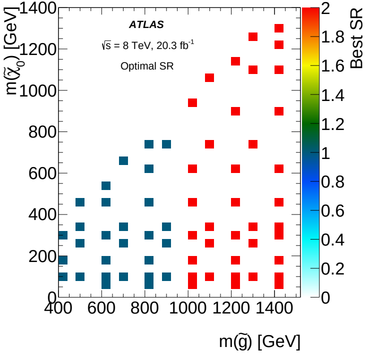
- **Templates**– background prediction– agree very well with **data**
- Validation region allows us to check analysis before looking in sensitive region
- **Uncertainties** come from template method, and cross-checks in

Opening the Box



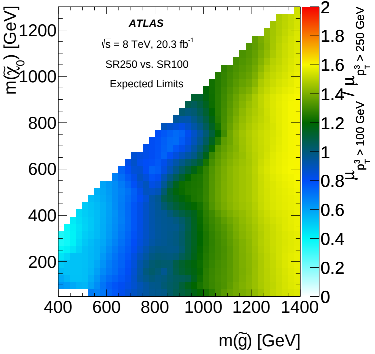
- Open the box: compare signal region prediction to data
- **Templates** agree very well with **data**: no sign of **BSM**
- Different bins provide different $S/\sqrt{B}$: multi-bin fit improves limits

Optimal Signal Regions



- Binning optimized via best expected limits: many bin number and divisions scanned
  - Fit each bin shown previously with $M_J^{\Sigma} > 350$ GeV
- Optimal SR in $m_{\tilde{g}}$, $m_{\tilde{\chi}}$ mass plane is shown for SR100 and SR250
  - High mass points gain from SR250: third jet has higher p_T
- Large improvement in limits from multiple signal regions: 60% gains at high/low $m_{\tilde{g}}$ compared to one SR

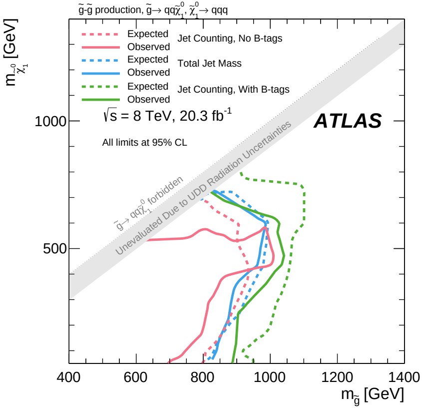
Optimal Signal Regions



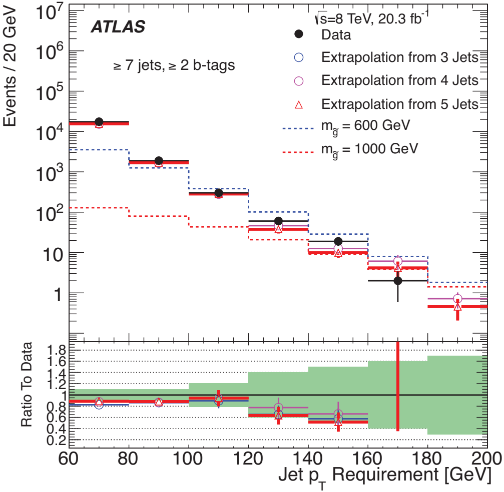
- Binning optimized via best expected limits: many bin number and divisions scanned
  - Fit each bin shown previously with $M_J^{\Sigma} > 350$ GeV
- Optimal SR in $m_{\tilde{g}}$, $m_{\tilde{\chi}}$ mass plane is shown for SR100 and SR250
  - High mass points gain from SR250: third jet has higher p_T
- Large improvement in limits from multiple signal regions: 60% gains at high/low $m_{\tilde{g}}$ compared to one SR

Comparison to Other Analyses

- Compare results to **jet-counting** analysis: simply counts jets above some p_T threshold (i.e., look for events with 7 jets with $p_T > 50$ GeV)
- Total jet mass analysis outperforms nominal jet counting
  - M_J^Σ incorporates *angular structure* of event: should be doing better!
- *b*-tagging jet-counting outperforms total jet mass
  - Squark and λ'' structure inclusive in flavor: often contains at least one *b*



Jet Counting



Future Prospects

- Add *b*-tagging to the analysis: orthogonal information from displaced vertices can clearly help!
- Focus on stops: increases jet multiplicity even higher
- Show results for different assumptions on λ'': extra top quarks, or only light flavor, etc.
- Large challenge for Run II: $\tilde{g}$ cross-section grows rapidly, but so does QCD background!
    - May need to add more variables, or rely on *b*-tagged analyses
    - Many tools left in the toolbox... should be an exciting search to continue!