# Learning how to count

## A high multiplicity search for the LHC

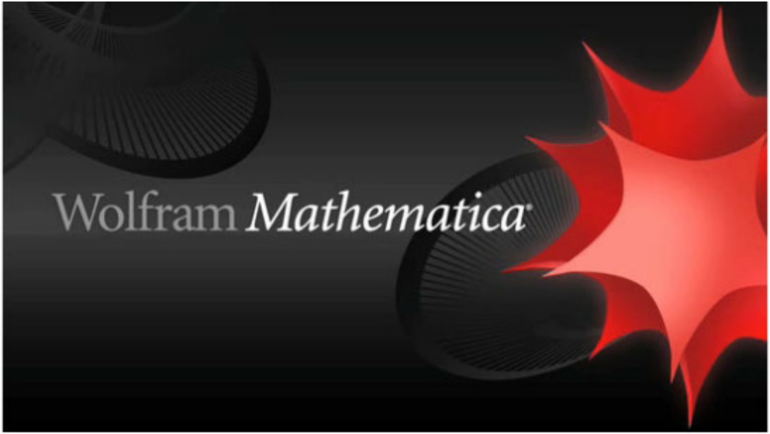
Sonia El Hedri

with

Anson Hook, Martin Jankowiak and Jay Wacker

August 13, 2013

arXiv:1302.1870v1


Wolfram Mathematica

Overview

Counting subjets

Results

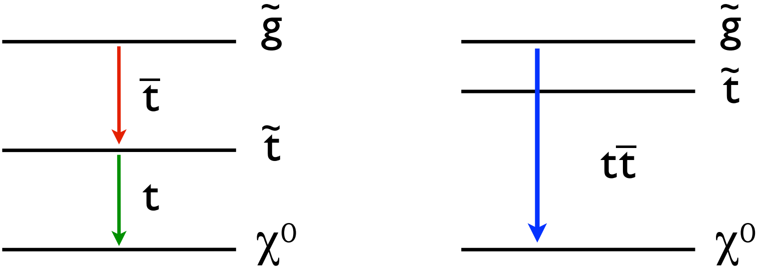
# Traditional searches

Handles to look at new physics signals:

- Leptons
- Heavy flavor jets (b-tagging)
- Kinematic reconstruction ($m_T$, MT2, ...)
- Boosted jets, tagging using jet substructure
- High $p_T$ jets, radius $R = 0.4, 0.5$
- Missing $E_T$

# One target: natural SUSY

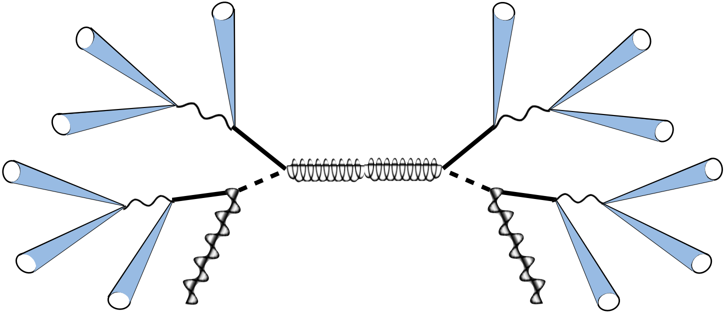
Decouple all particles not cancelling the top quadratic divergences




$$\tilde{g} \rightarrow t\bar{t}\chi^0$$

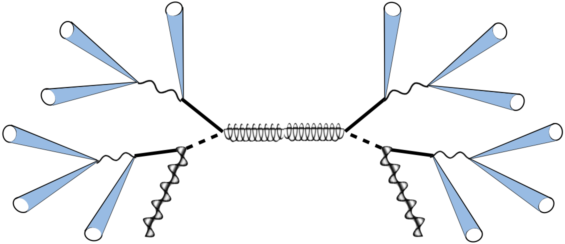
# High multiplicity signals

$>$ 12 jet signals from natural SUSY



Other signals: RPV, strong dynamics, cascade decays, ...

# High multiplicity signals



- Dominating if the light particles are hard to see
- Low production rate
- Signatures distributed across many channels
  - Exclusive searches are low efficiency
  - Inclusive searches are high background

## Traditional approaches

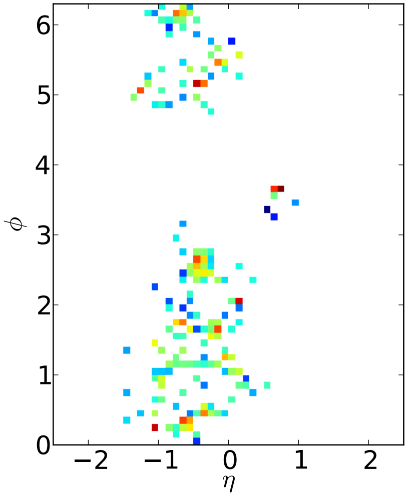
- Cluster thin jets, $R = 0.4 - 0.5$, $p_T > 50\,\mathrm{GeV}$
- Cut on the number of jets
- Cut on $\not{E}_T$

## But

- Soft jets, $p_T \sim 50\,\mathrm{GeV}$
- Low $\not{E}_T$
- Parton shower adds jets
- Complicated phase space $(3^{N_j})$
- No top tagging

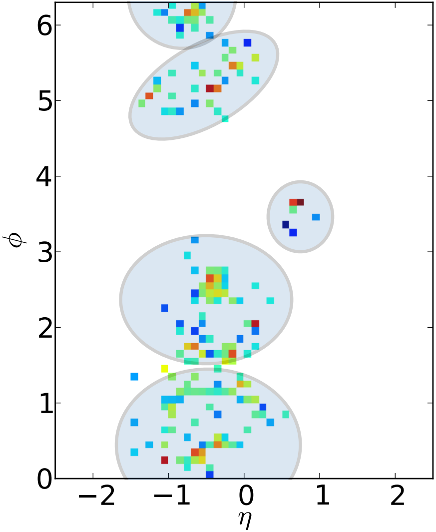
# High multiplicity signals

Jets hard to resolve individually...



# High multiplicity signals

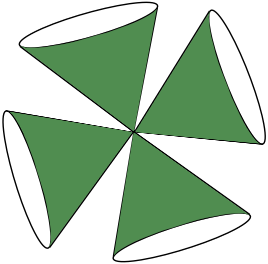
...or accidental boost!



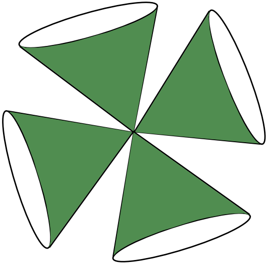
# Using fat jet techniques on high multiplicity events

$>$12 low $p_T$ thin jets $\rightarrow$ four high $p_T$ fat jets

Signal



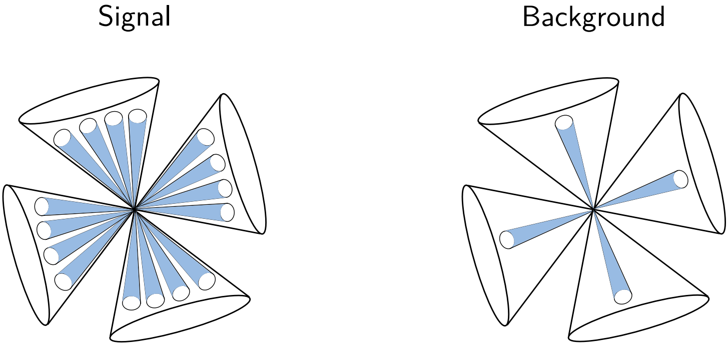
Background



QCD still dominates, even with a $\not{E}_T$ cut

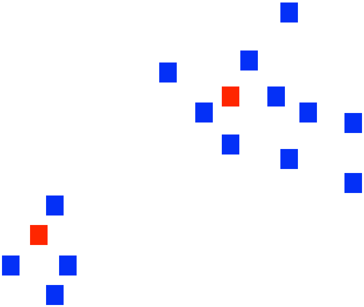
# Using fat jet techniques on high multiplicity events

"Count" the number of subjets using jet substructure techniques



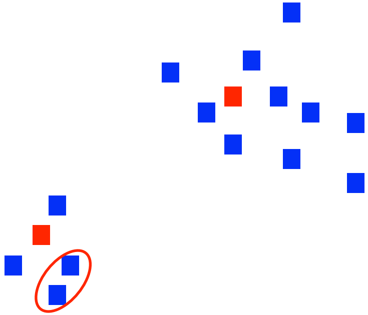

$$N = \sum_i N_i^{subjets}$$

# The exclusive $k_t$ algorithm



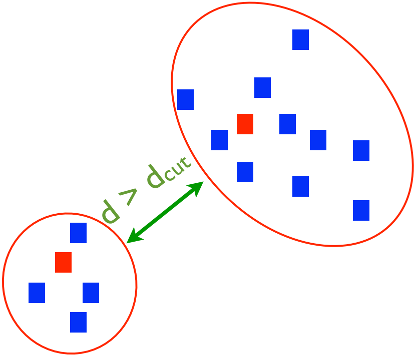
- $d_{ij} = \min(p_{Ti}^2, p_{Tj}^2)\Delta R_{ij}^2$

# The exclusive $k_t$ algorithm



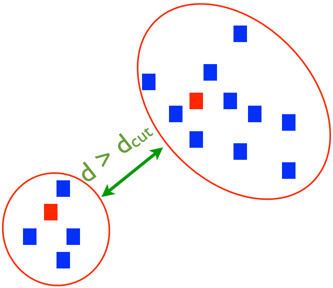
- $d_{ij} = \min(p_{Ti}^2, p_{Tj}^2)\Delta R_{ij}^2$
- Cluster soft components first

# The exclusive $k_t$ algorithm




- $d_{ij} = \min(p_{Ti}^2, p_{Tj}^2)\Delta R_{ij}^2$
- Cluster soft components first
- Stops when $d_{ij} > d_{cut}$

# Counting with $k_T$




- Find $d_{cut}$ which maximizes S/B

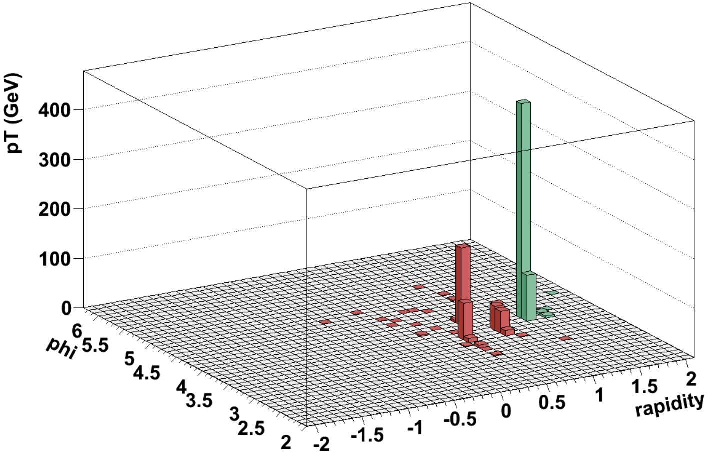
$$\sqrt{d_{cut}} = 0.065\, p_{TJ}$$

- Run the exclusive $k_T$ algorithm
- Select jets with $p_T > 40\,\mathrm{GeV}$

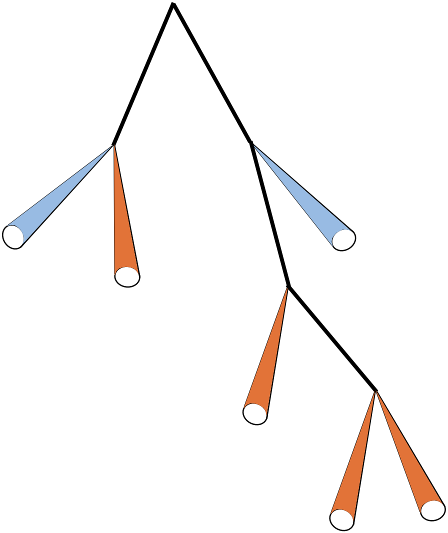
# Advantages of $k_T$

Soft wide angle radiation is clustered with the hard jets

QCD jets have a low $n_{k_T}$



# Counting with the Cambridge-Aachen algorithm

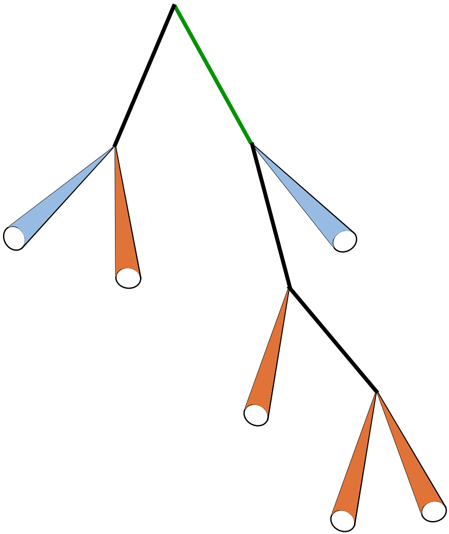


$$d_{ij} = \Delta R_{ij}^2 \qquad (1)$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$

# Counting with the Cambridge-Aachen algorithm

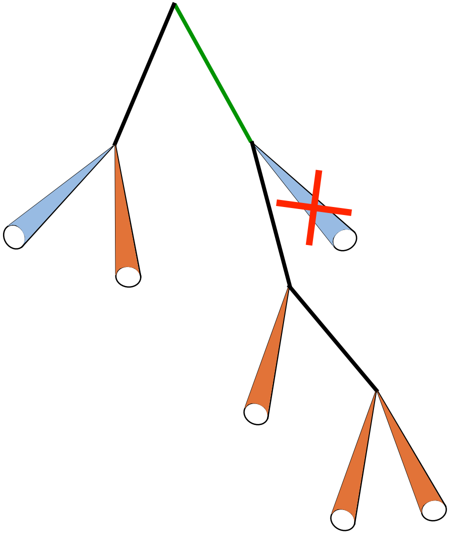


$$d_{ij} = \Delta R_{ij}^2 \qquad (2)$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$
- If $p_T$s are imbalanced, remove soft jet

# Counting with the Cambridge-Aachen algorithm

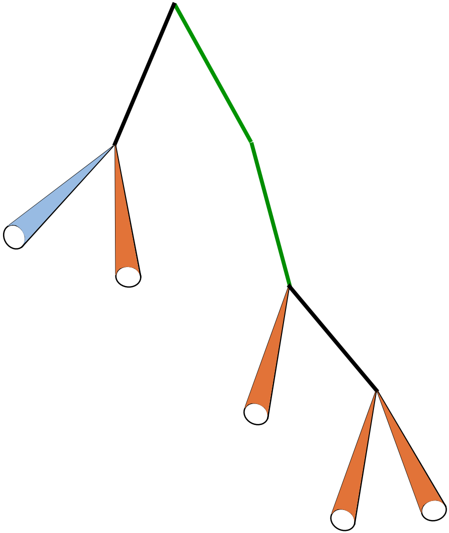


$$d_{ij} = \Delta R_{ij}^2 \qquad (3)$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$
- If $p_T$s are imbalanced, remove soft jet
- If $m_j < m_{cut}$ or $d_{12} < R_{min}$, $j$ is a subjet
- Keep subjets with $p_T > p_{Tcut}$

# Counting with the Cambridge-Aachen algorithm

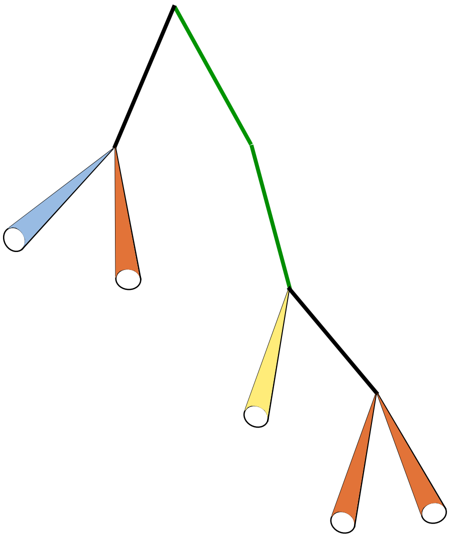


$$d_{ij} = \Delta R_{ij}^2 \qquad (4)$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$
- If $p_T$s are imbalanced, remove soft jet
- If $m_j < m_{cut}$ or $d_{12} < R_{min}$, $j$ is a subjet
- Keep subjets with $p_T > p_{Tcut}$

# Counting with the Cambridge-Aachen algorithm

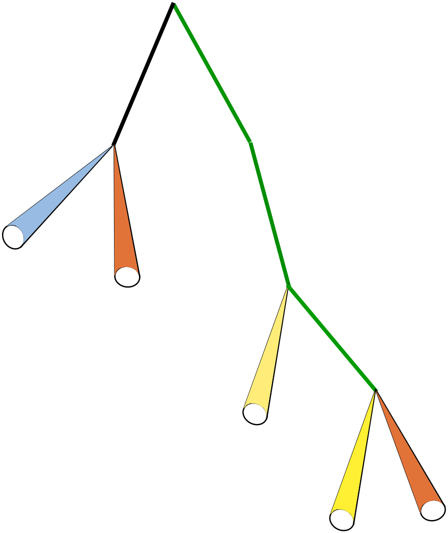


$$d_{ij} = \Delta R_{ij}^2 \qquad (5)$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$
- If $p_T$s are imbalanced, remove soft jet
- If $m_j < m_{cut}$ or $d_{12} < R_{min}$, $j$ is a subjet
- Keep subjets with $p_T > p_{Tcut}$

# Counting with the Cambridge-Aachen algorithm



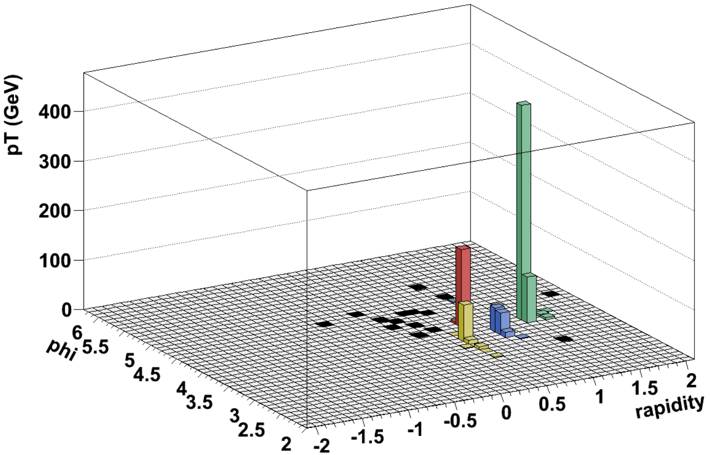
$$d_{ij} = \Delta R_{ij}^2 \tag{6}$$

Cluster the jet with CA and go down the clustering tree

- Uncluster $j$ into $j_1$ and $j_2$
- If $p_T$s are imbalanced, remove soft jet
- If $m_j < m_{cut}$ or $d_{12} < R_{min}$, $j$ is a subjet
- Keep subjets with $p_T > p_{Tcut}$

# Counting with CA

- Subjets consistent with the decay of a massive particle
- Soft radiation discarded
- $m_{cut} = 30\,\mathrm{GeV}$, $y_{cut} = 0.10$, $R_{min} = 0.15$, $p_{Tcut} = 30\,\mathrm{GeV}$

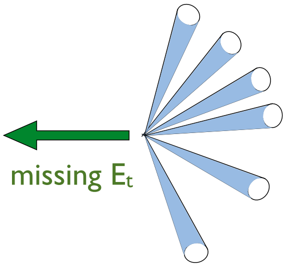


# Fastjet plugin

**fastjet.hepforge.org/trac/browser/contrib/contribs#SubjetCounting**

# ATLAS high multiplicity search

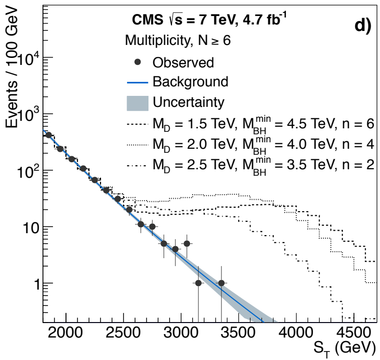
ATLAS-CONF-2012-103




- 8 TeV, $5.8\mathrm{fb}^{-1}$
- Anti-$k_t$ algorithm with $R = 0.4$
- 7, 8 or 9 jets with $p_T > 55\,\mathrm{GeV}$
- 6, 7 or 8 jets with $p_T > 80\,\mathrm{GeV}$
- $\frac{\not{E}_t}{\sqrt{H_t}} > 4\,\mathrm{GeV}^{1/2}$

# CMS black hole search

CERN PH-EP/ 2012 045



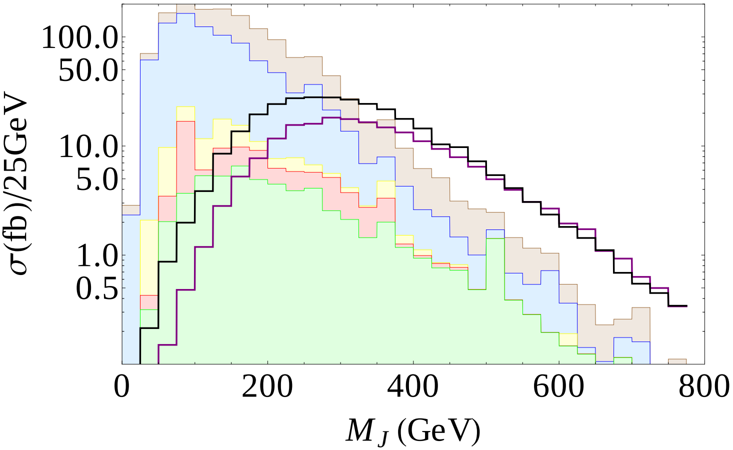

- 7 TeV, $4.7\mathrm{fb}^{-1}$
- Anti-$k_t$ algorithm with $R = 0.5$
- cut on number of objects with $E_T > 50\,\mathrm{GeV}$

  $n \in \{2, 4, 6\}$
- cut on $S_T = \sum E_{Tobj} + \not{E}_t$

  $S_{Tmin} \in [1.9, 4.1]\,\mathrm{GeV}$

# Fat jet mass

$$M_J = \sum_{j \in \text{jets}} m_j$$



$$N_{\mathrm{jets}}\ \text{cut} + \not{E}_T\ \text{cut (ATLAS)}$$

vs

$$N_{\mathrm{objects}}\ \text{cut} + S_T\ \text{cut (CMS)}$$

vs

$$M_J\ \text{cut} + \not{E}_T\ \text{cut}$$

vs

$$M_J\ \text{cut} + \not{E}_T\ \text{cut} + N_{\mathrm{subjets}}\ \text{cut}$$

# Benchmark models

| | | Tops jets? | Cascade decay? | RPV? |
|---|---|---|---|---|
| $\tilde{g}$ | $\rightarrow$ | $t\bar{t}\chi_i^0$ | $\chi_i^0 \rightarrow VV\chi_1^0$ | $\chi_1^0 \rightarrow jjj$ |
| | | $+12$ jets | $+8$ jets | $+6$ jets |

- 8 possible topologies
- from 4 to 26 jets
- signals with and without $\not{E}_T$

# Benchmark models and searches

Optimal cuts depend on :

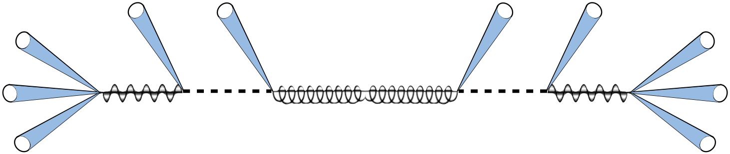
- Jet multiplicity
- $\not{E}_T$
- Presence of leptons
- Mass of the initial particle $m_{\tilde{g}}$

**Inclusive** search:

- Leptons clustered with jets
- Find **minimal** number of cuts on $M_J + \not{E}_T + \ldots$ so that the bounds are close to optimal
  - For each signal
  - For each mass

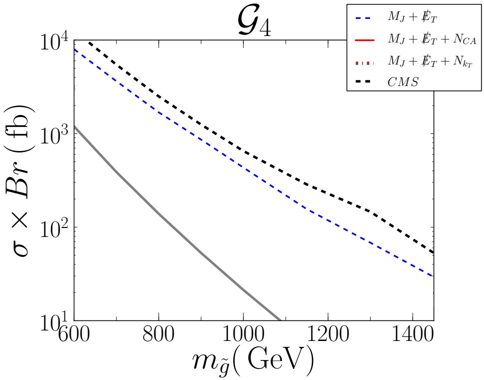
# Gluino decay to light quarks, RPV

$\tilde{g} \rightarrow jj\chi_1^0,\ \chi_1^0 \rightarrow jjj$



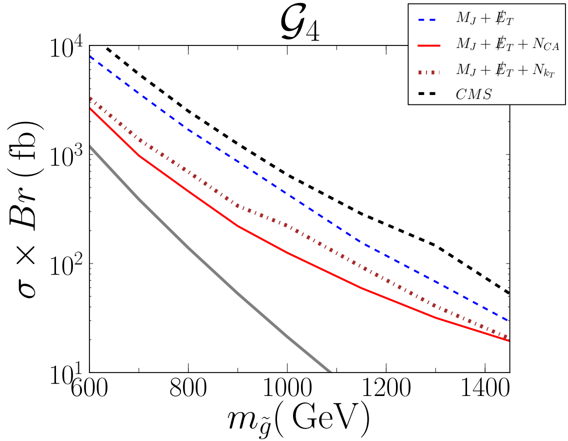
10 jets, no $\not{E}_T$

# Gluino decay to light quarks, RPV – 8 TeV, 30 $\mathrm{fb}^{-1}$




- $M_J > 1.3$ TeV
  low signal efficiency + high background
- CMS search and $M_J + \not{E}_T$ search comparable

# Gluino decay to light quarks, RPV – 8 TeV, 30 $\mathrm{fb}^{-1}$



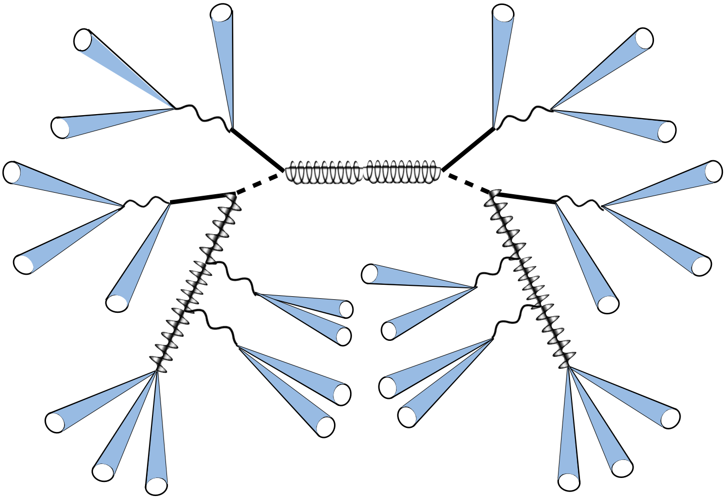

$M_J \geq 1$ TeV

$N_{CA} \geq 13$

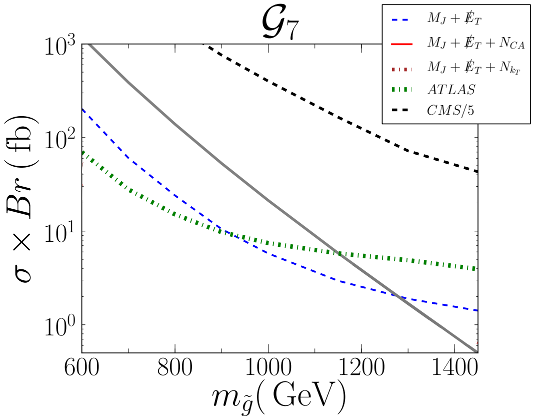
- Factor of 2 to 4 improvement over $M_J + \not{E}_T$ and CMS
- $M_J$ cut loosened
- CA slightly better than $k_T$

# Gluino 2 step decay, RPV

$\tilde{g} \to t\bar{t}\chi_2^0,\ \chi_2^0 \to VV'\chi_1^0,\ \chi_1^0 \to jjj$

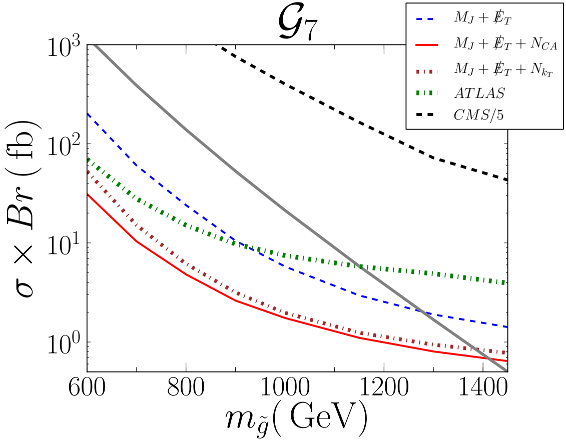


# Gluino 2 step decay, RPV – 8 TeV, $30\mathrm{fb}^{-1}$




- $S_T$ cut too large for CMS search
- $M_J > 725\,\mathrm{GeV}$, $\not{E}_T > 175\,\mathrm{GeV}$
- $M_J + \not{E}_T$ search better at high mass

# Gluino 2 step decay, RPV – 8 TeV, $30\mathrm{fb}^{-1}$




$M_J \geq 425\,\mathrm{GeV}$

$\not{E}_T \geq 125\,\mathrm{GeV}$

$N_{CA} \geq 14$

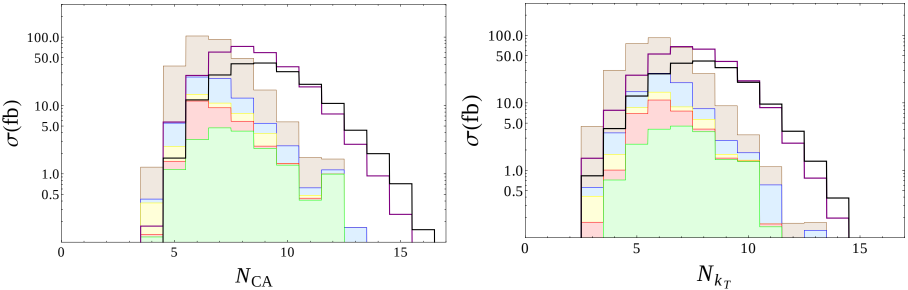
- Factor of $\sim$ 4 improvement over $M_J$ + MET
- Factor of $\sim$ 5 improvement over ATLAS at high mass
- CA slightly better than $k_T$

# Summary

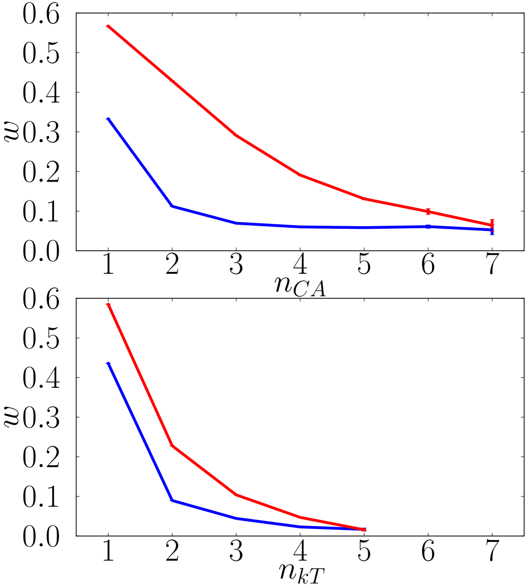
- Common new physics scenarios predict events with very high multiplicity
- Standard handles not appropriate (not boosted, complicated kinematics, low energy)
- Fat jet techniques are more robust but requires finding new jet substructure variables
- Counting subjets in an event provides good discriminating power
- $M_J$ and $\not{E}_T$ cuts loosened, could be used to probe $\not{E}_T$-less signals
- Allows to make data driven estimates of the QCD background

# Backup

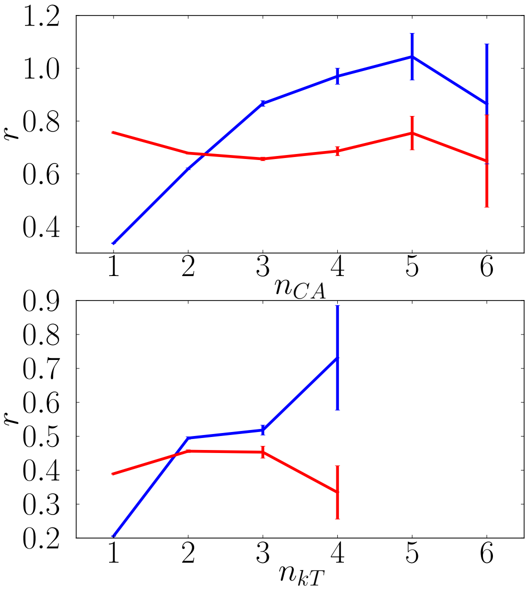
# Signal and background distributions

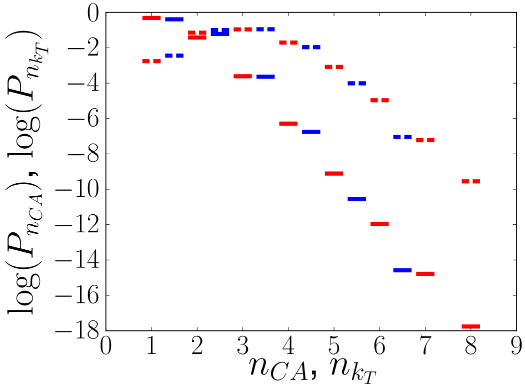


# Scaling patterns

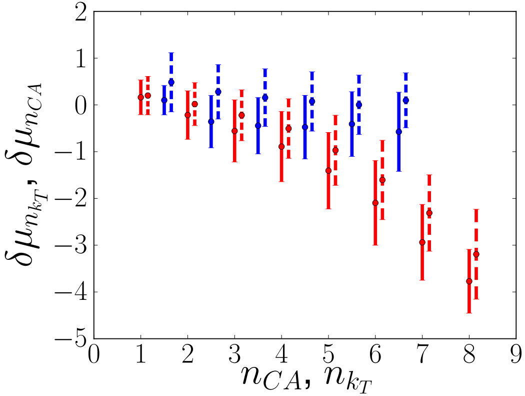


# Scaling patterns

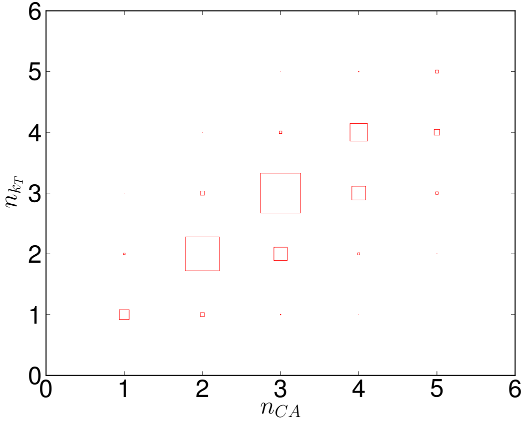



$\log(P_{n_{CA}})$, $\log(P_{n_{k_T}})$
0
−2
−4
−6
−8
−10
−12
−14
−16
−18
0 1 2 3 4 5 6 7 8 9
$n_{CA}$, $n_{k_T}$

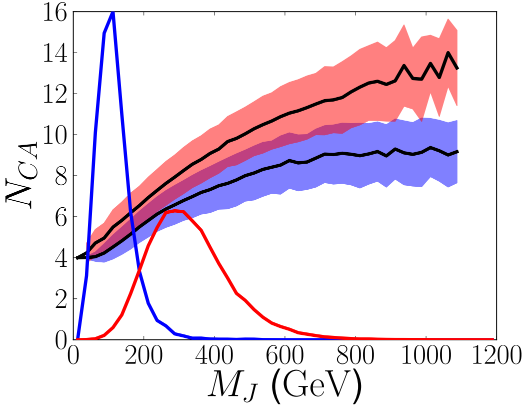
# Correlations between $N_{CA}$ and $N_{k_T}$



# Correlations between $N_{CA}$ and $N_{k_T}$



# $N_{CA}$ vs $M_J$



# $N_{CA}$ vs $M_J$