Search for SUSY in Dijet Events with Novel Data-Driven Background Estimation

Henning Flächer
(CERN)
for the CMS Collaboration

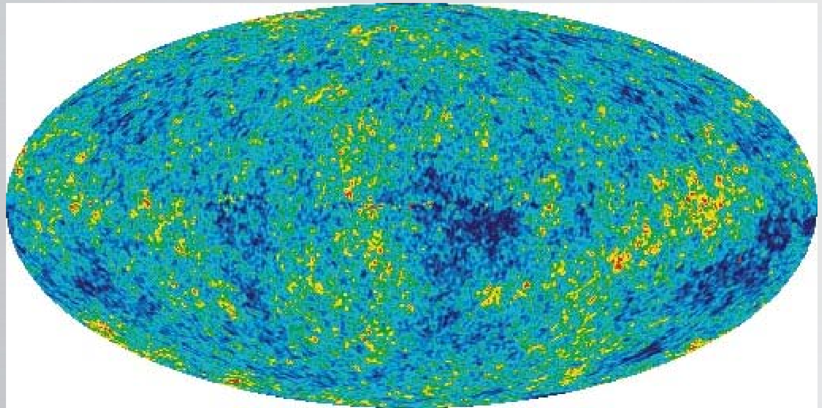
- Motivation for di-jet search
  - Dark Matter at Colliders
- Missing Energy Signatures in Di-Jet Events
  - robust analysis techniques
  - data-driven background estimates
- Conclusions

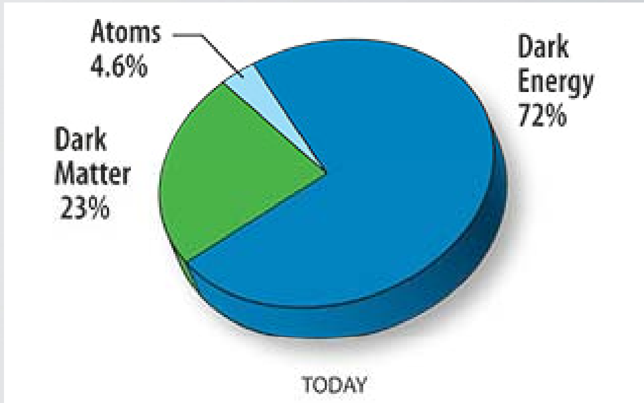
La Thuile 2009 (March 1st -7th, 2009)

A Look at the Energy and Matter Content of the Universe

- Cosmic microwave background gives precise information about dark matter content of the universe

- WMAP 5 year result:





Credit: NASA/WMAP Science Team

  - relic dark matter density of the universe

    $\Omega_{DM}h^2 = 0.110 \pm 0.006$

- Only 5% is made from baryonic matter, 23% from unknown “dark matter”

- Attractive explanation for Dark Matter:
  - new weakly interacting particle

Can we produce Dark Matter particles at Colliders?

- Dark Matter candidate is a weakly interacting massive particle (WIMP)

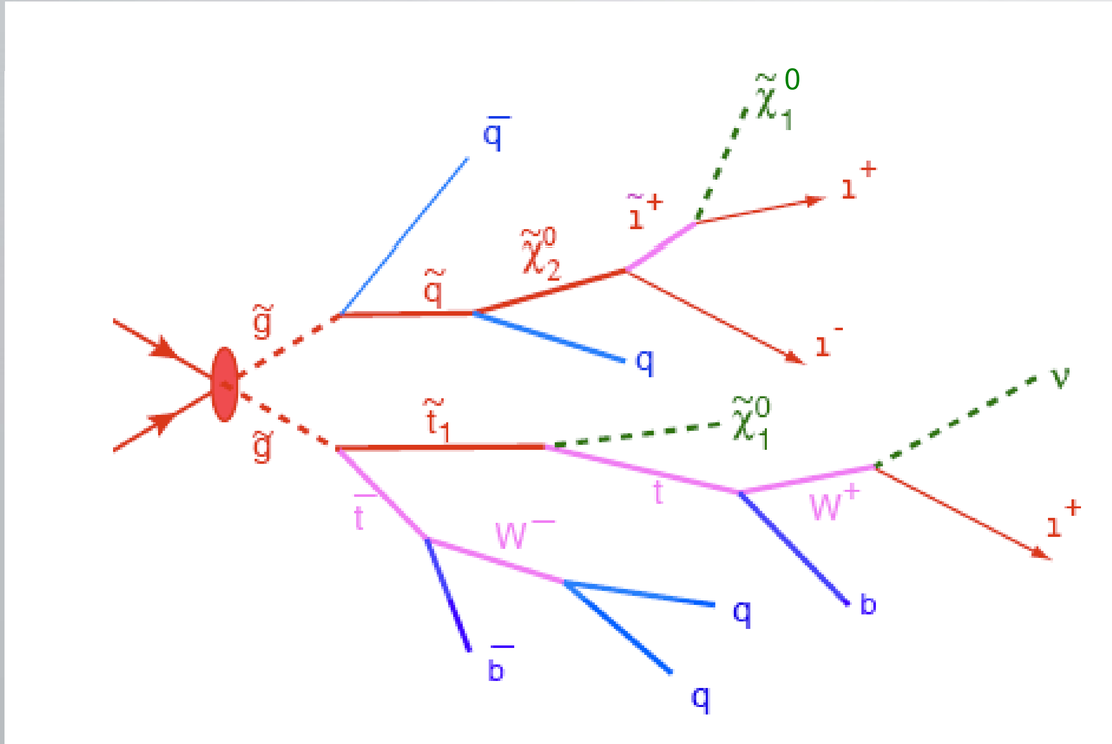
- Many New Physics Models provide viable dark matter candidates, e.g.
  - R-parity conserving Supersymmetry
    - minimal super gravity mSugra ➔ neutralino is WIMP
    - Gauge mediated SUSY ➔ gravitino is WIMP (too light)
  - Universal Extra Dimensions
  - Warped Extra Dimensions
  - Little Higgs Models
  - Technicolor Models

- Production of WIMP's in cascade decays of heavy new particles
  - WIMP's escape the detector and remain undetected
  - Leads to a missing energy signature

An Example from SUSY

e.g. gluino pair-production

lots of missing energy, many jets, and possibly leptons in the final state



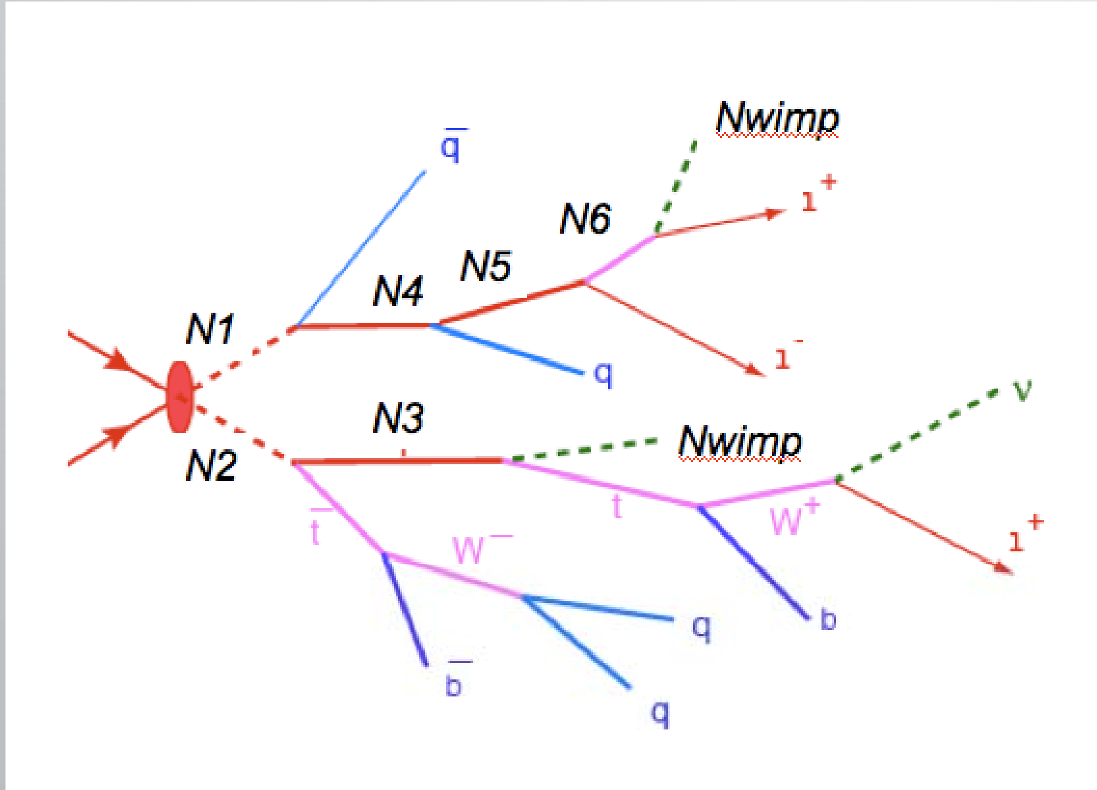
Missing Energy:
- from LSP

Multi-Jets:
- from cascade decay (gaugino)

Multi-Leptons:
- from decay of charginos and neutralinos

...but signature is more general

- pair production of new heavy particles



Missing Energy:
- Nwimp - end of the cascade

Multi-Jets:
- from decay of the Ns (possibly via heavy SM particles like top, W/Z)

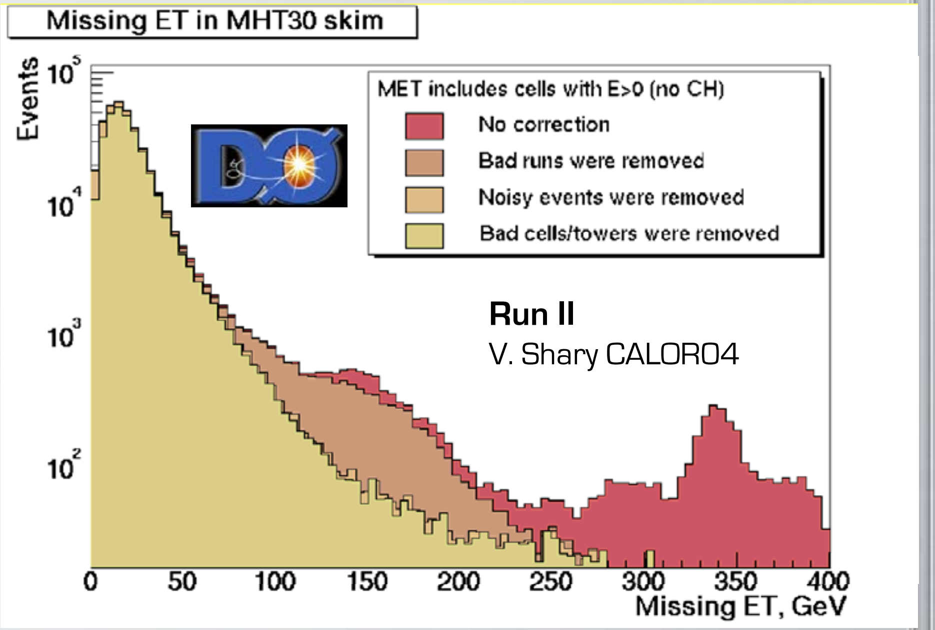
Multi-Leptons:
- from decay of the N's

Model examples are Extra dimensions, Little Higgs, Technicolour, etc

Missing Energy Measurement

- “Traditional” approach:
  - Calculate missing energy as negative vectorial sum of all calorimeter deposits
  - Susceptible to mis-measurements from, e.g.
    - Calorimetric noise (hot cells)
    - Cosmic rays
    - Beam-gas interactions
    - Beam-halo events
  - Difficult to understand in the early days of data taking
- Need for robust measurement techniques

Missing Energy from Tevatron during several cleanup stages:

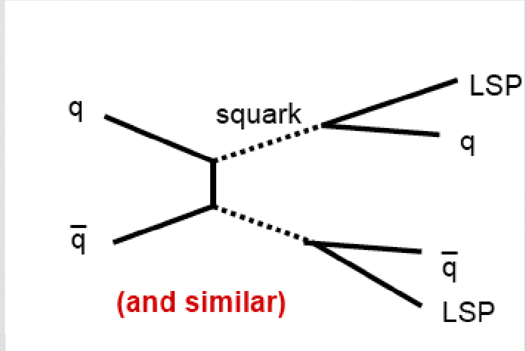


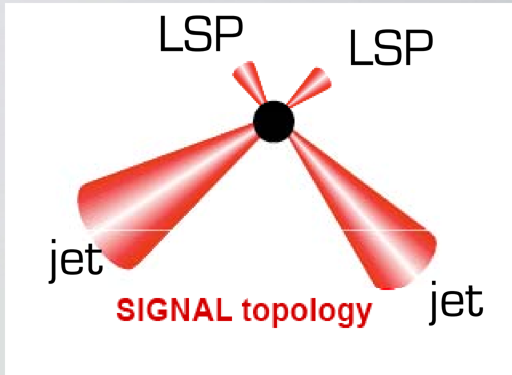
IDEA:
infer missing energy from well measured objects by applying transverse energy/momentum conservation

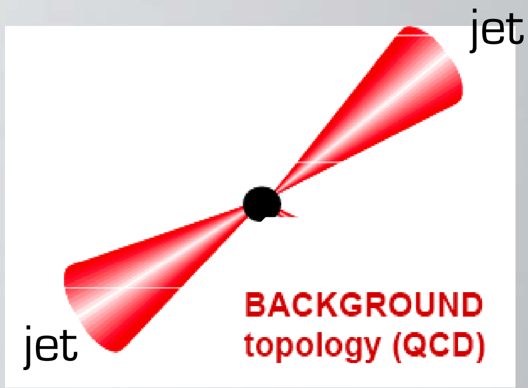
Di-jet Analysis

- New CMS study: PAS-SUS-08/005
  - CMS PTDR II focused on inclusive SUSY searches with ≥ 3 jets
- Motivated in addition by recent paper by
  - L. Randall, D.Tucker-Smith (Phys.Rev.Lett.101:221803,2008)
- Idea:
  - Squarks pair produced and directly decaying to quarks and neutralinos
- Event topology
  - Only two jets + missing energy
- Background:
  - QCD dijet events
    - No real missing momentum
  - $Z \to \nu\nu$ events
    - Irreducible background due to real missing E_T
  - $W \to l\nu$
    - Leads to missing Et when lepton not reconstructed or out of acceptance

Transverse momentum conservation
Jets back-to-back in phi
E_T of jets equal in magnitude







Event Selection

- Main variables of interest
  - HT: Scalar sum of Jet p_T's:
    - $HT = p_T^{Jet1} + p_T^{Jet2}$
  - MHT: Jet based missing E_T
    - $MHT = -(\vec{p}_T^{Jet1} + \vec{p}_T^{Jet2})$
  - MT: Transverse Mass
    - $$M_T = \sqrt{\left(\Sigma_i E_{T_i}\right)^2 - \left(\Sigma_i p_{x_i}\right)^2 - \left(\Sigma_i p_{y_i}\right)^2}$$
      $$= \sqrt{HT^2 - MHT^2}$$
  - but also p_T of a possible 3rd jet
  - $\Delta\varphi$ between the jets
  - α (α_T) from 2 leading jets

- Trigger
  - di-jet trigger (2 jets with $p_T > 150$ GeV)
- Preselection:
  - Jet Selection
    - 2 jets with $p_T > 50$ GeV, $F_{em} < 0.9$
    - 3rd jet veto: $p_T < 50$ GeV
    - $\Delta\varphi(MHT, jet_{1,2,3}) > 0.3$ rad
    - $|\eta_{j1}| < 2.5$
  - Lepton veto's:
    - no e, μ with pt >10 GeV
- Full Selection
  - $HT > 500$ GeV
  - α (α_T) > 0.55
  - ($\Delta\phi < 2\pi/3$)

Accounting for finite resolution (not optimised)

Discriminating Variables

- Exploit kinematics of the event
  - Define new variable α (Randall – Tucker-Smith):

$$\alpha = \frac{E_{T\ j2}}{M_{j1j2}} = \frac{E_{T\ j2}}{\sqrt{2E_1E_2(1-\cos\theta)}}$$

    - Can be at most 0.5 for QCD, $\alpha < 0.5$
    - $\alpha > 0.5$ implies missing momentum

  - And transverse α_T:

$$\alpha_T = \frac{E_{T\ j2}}{M_{T\ j1j2}} = \frac{\sqrt{E_{T\ j2}/E_{T\ j1}}}{\sqrt{2(1-\cos\Delta\varphi)}}$$

    - Exploits that for QCD jets need to be back-to-back and of equal magnitude
    - For QCD dijets $\alpha = 0.5$

Analysis does not rely on calorimetric MET, MHT inferred from 2 jets

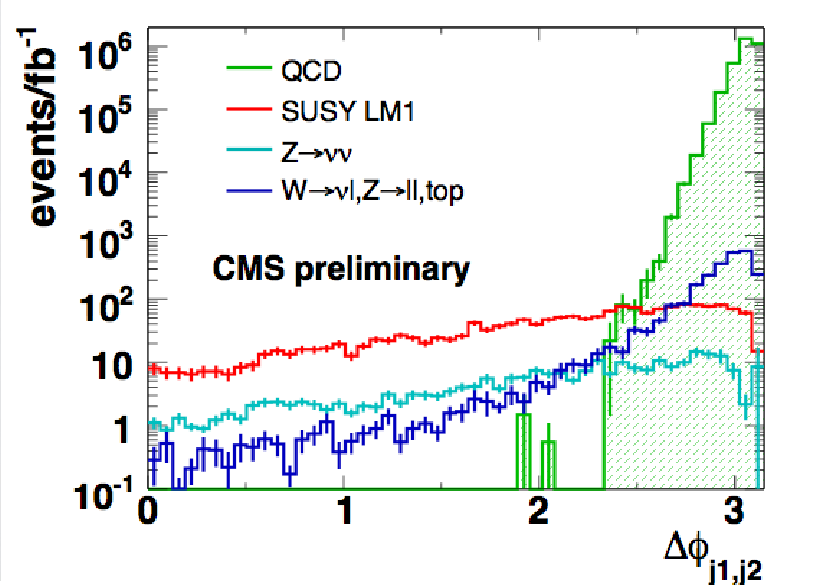
⇒ well suited for early data

Discriminating Variables (II)

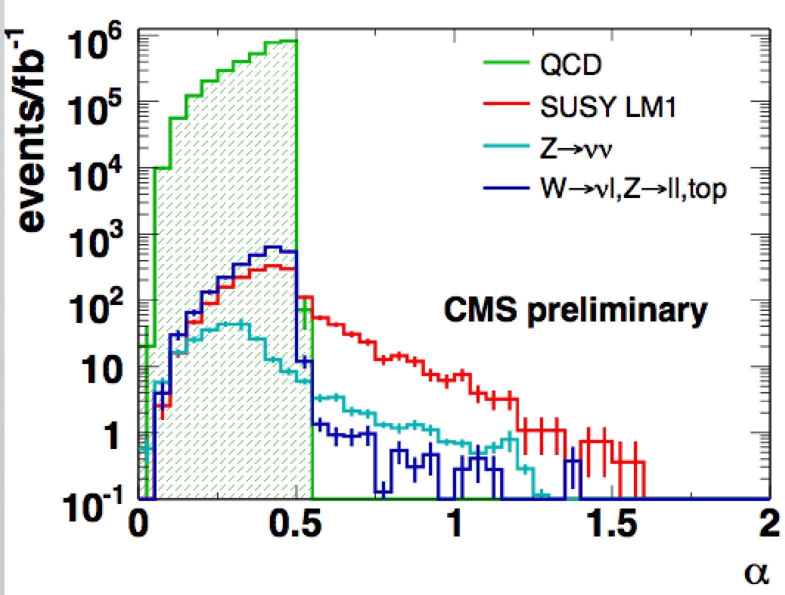
- Cuts applied:
  - Preselection & HT > 500 GeV

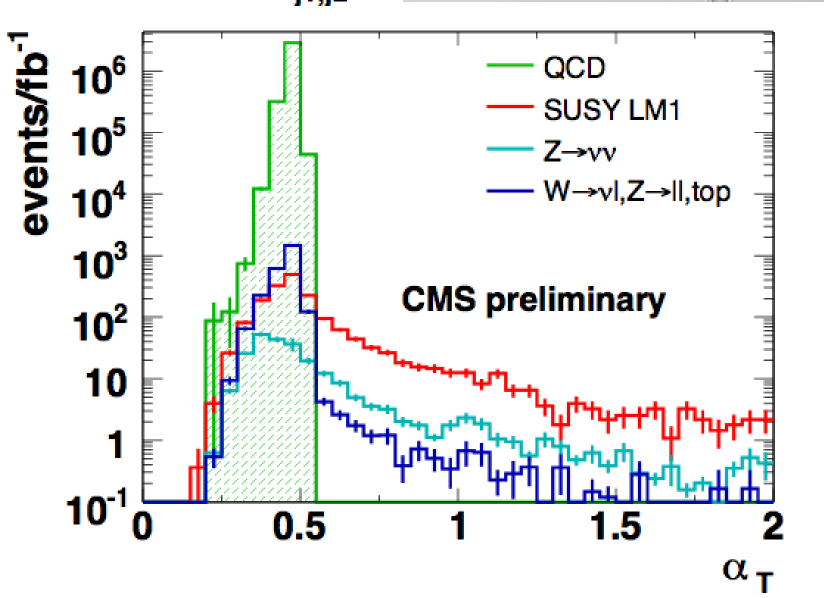
1 fb^{-1}

QCD peaking
at $\Delta\phi = \pi$,
but exhibiting tail



Sharp drop of
QCD background
for α (α_T) > 0.5





- $Z\rightarrow\nu\nu$ is main background
- W and other Z decays small

Signal & Background yields

- Expected event yields for 1 fb^{-1}

Selection cut	QCD	$t\bar{t}$,W,Z	$Z \to \nu\bar{\nu}$	LM1
Trigger	1.1×10^8	147892	1807	25772
Preselection	3.4×10^7	9820	878	2408
HT > 500 GeV	3.2×10^6	2404	243	1784
$\alpha > 0.55$	0	7.2	19.7	227.6
$\alpha_T > 0.55$	0	19.9	58.2	439.6
$\Delta\phi_{j1,j2} < 2\pi/3$	0	18.7	57.2	432.4

=> Signal/Background = 5.6

Reminder: desired topology is 2 squarks decaying to 2 squarks and 2 neutralinos (LSPs)

Sample	Events	$\tilde{q}\,\tilde{q}$ (invisible)	$\tilde{q}\,\tilde{q}$ (other)	$\tilde{q}\,\tilde{g}$	$\tilde{g}\,\tilde{g}$	other
LM1	432	39%	22%	34%	3%	1%
LM2	132	46%	33%	18%	0%	2%
LM3	138	69%	17%	12%	0%	2%
LM4	195	49%	10%	36%	3%	1%

- Variation of jet energy scale and resolution
  - 10% gaussian smearing of jet p_T's and of 0.1 rad of φ measurement
  - Scaling of jet energy by ± 5%
  - Scaling of jet energy by ± 3% for endcap/forward ($|\eta|$>1.4)
- Stable S/B for all variations

Data-Driven Background Studies

- LHC data in explores a new energy regime
  - Monte Carlo simulations should not be taken at face value
  - develop data-driven techniques
  - identify data control samples

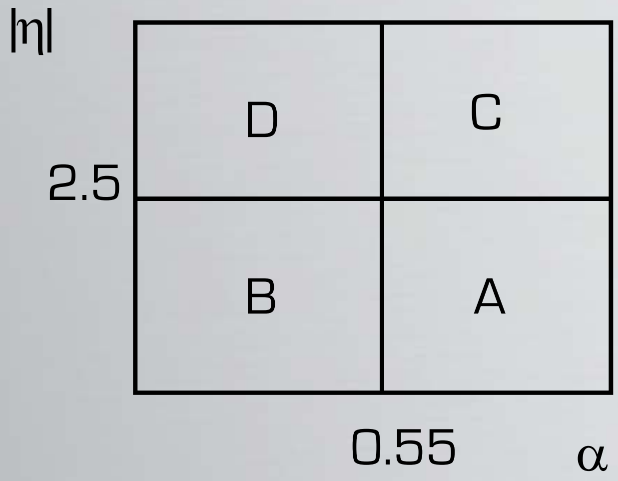
- Two main sources of background:

- QCD
  - Seems to be under control but huge cross-section
  - MC uncertainties due to higher order QCD effects

- $Z \rightarrow \nu\nu$ + Jets
  - represents an irreducible background
  - two jets + real missing E_T
  - Ideally study $Z \rightarrow \mu\mu$ events but not enough statistics in the early days
  - Other control samples:
  - W + Jets
  - Photon + Jets as shown in CMS-AN 36/2008

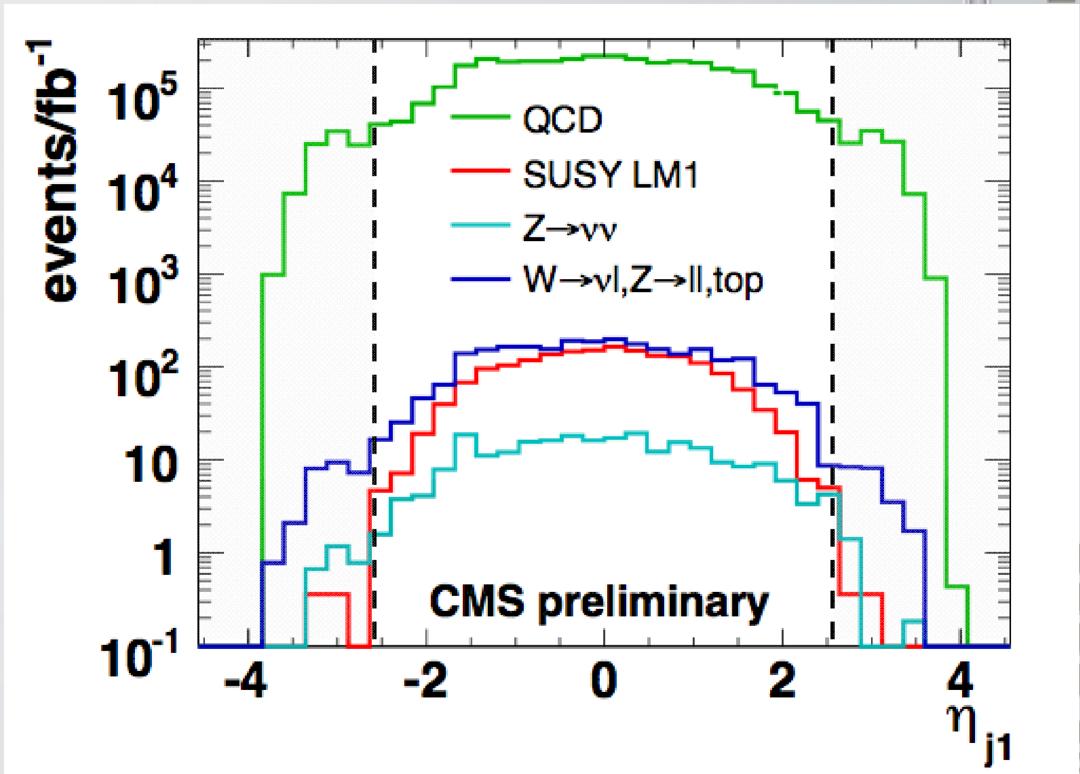
Central Production of Heavy Objects

- Idea: define signal enriched and depleted regions by splitting data sample in events with first jet in barrel and forward region
  - SUSY jets are more central
  - Use ratio of events $R_a = \alpha_T > 0.55 / \alpha_T < 0.55$ in (signal depleted) forward η region to predict background in (signal enriched) barrel region.



See also: Background Modeling in New Physics Searches Using Forward Events at LHC.
V. Pavlunin, D. Stuart, Phys.Rev.D78:035012,2008.

Pre-selection (no η cut) + HT > 500 GeV

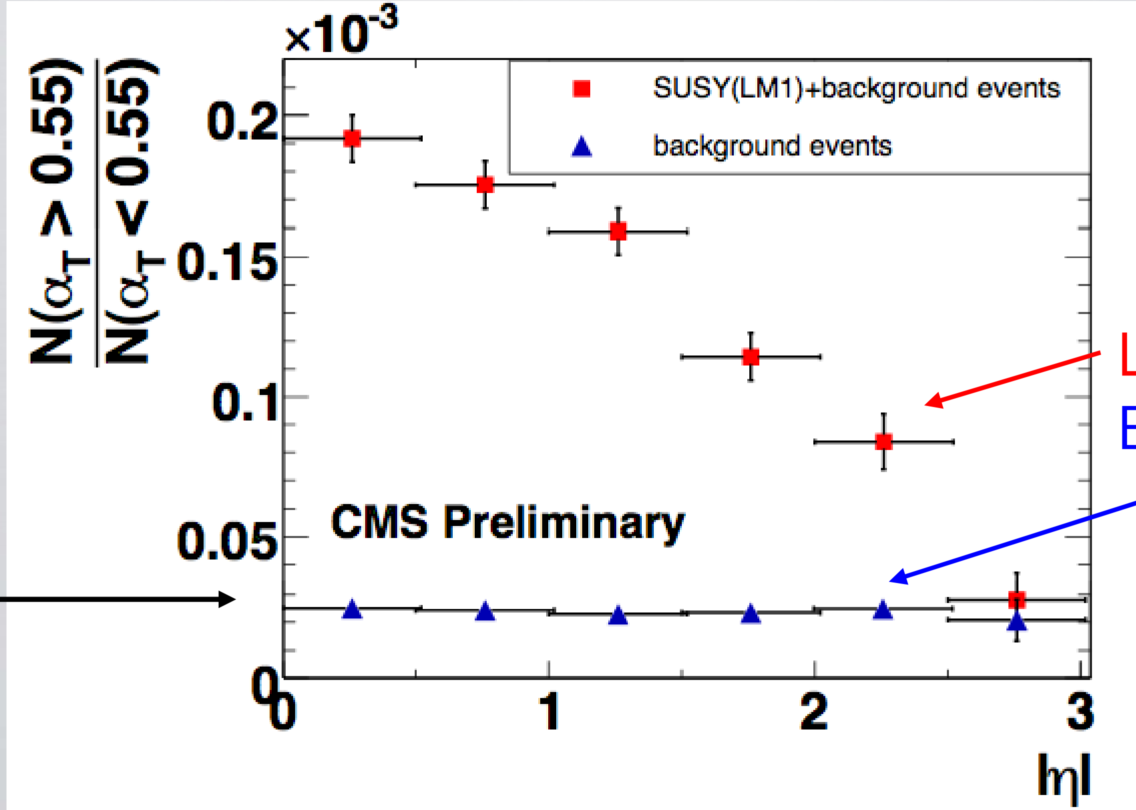


R = C/D: assumed to be constant over η and nearly signal free
also: constant for all background contributions individually

Then, background in A can be obtained as:
A = B * R

η Dependence of Matrix Method

Pre-selection (no |η| cut) + HT > 500 GeV



LM1+BK

Background only

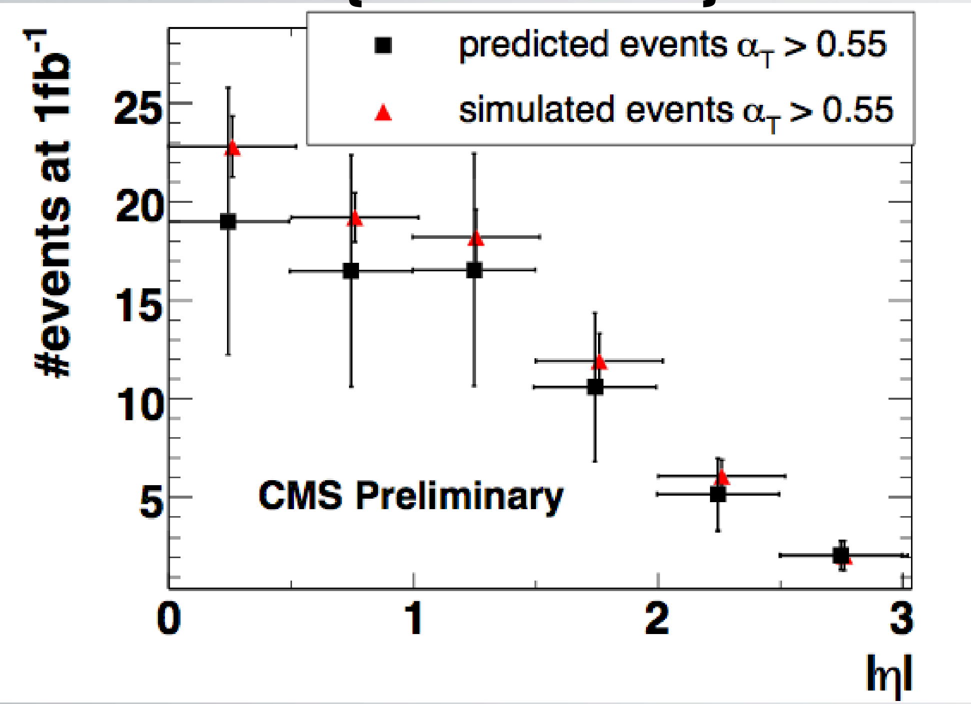
Verified that flat for all Bkgd contributions (see AN-08-071)

- R_α flat for background as function of $|\eta_{j1}|$
- α_T and $|\eta_{j1}|$ can be used for ABCD-matrix method

✸ Measure $R\alpha_T$ in $2.5 < |\eta| < 3.0$ region.

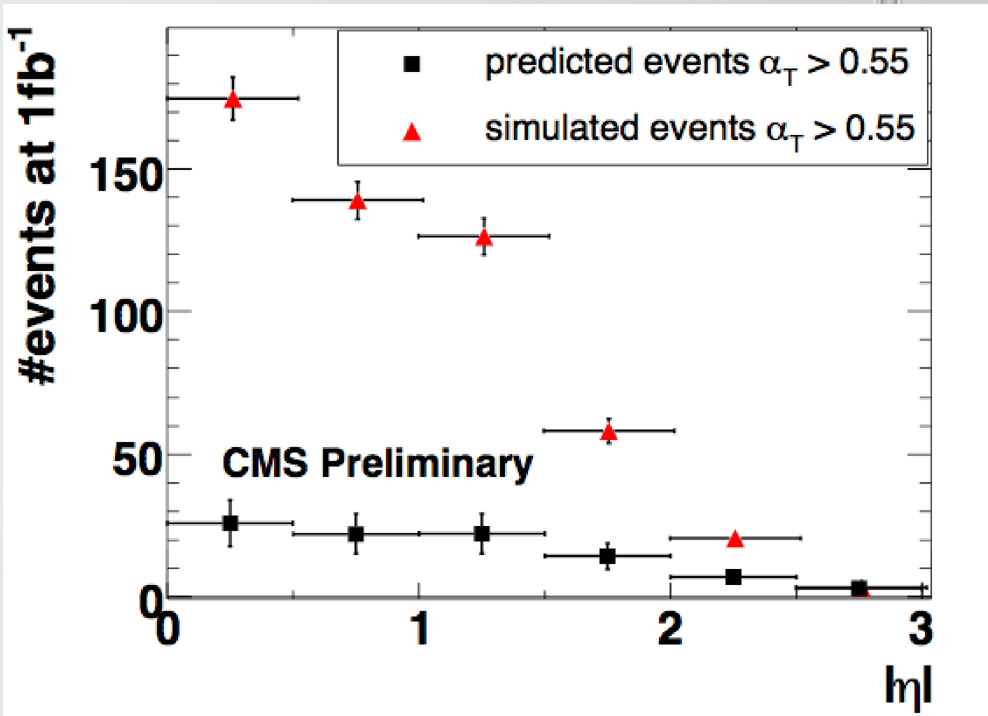
η Dependence of Matrix Method

without SUSY (closure test)



simulated: 77 ± 3 (8 @ 1fb^{-1})
predicted: 68 ±24 (42 @ 1fb^{-1})

with LM1 SUSY "contamination"



simulated: 517 ± 13 (22 @ 1 fb^{-1})
predicted: 91 ± 30 (51 @ 1 fb^{-1})

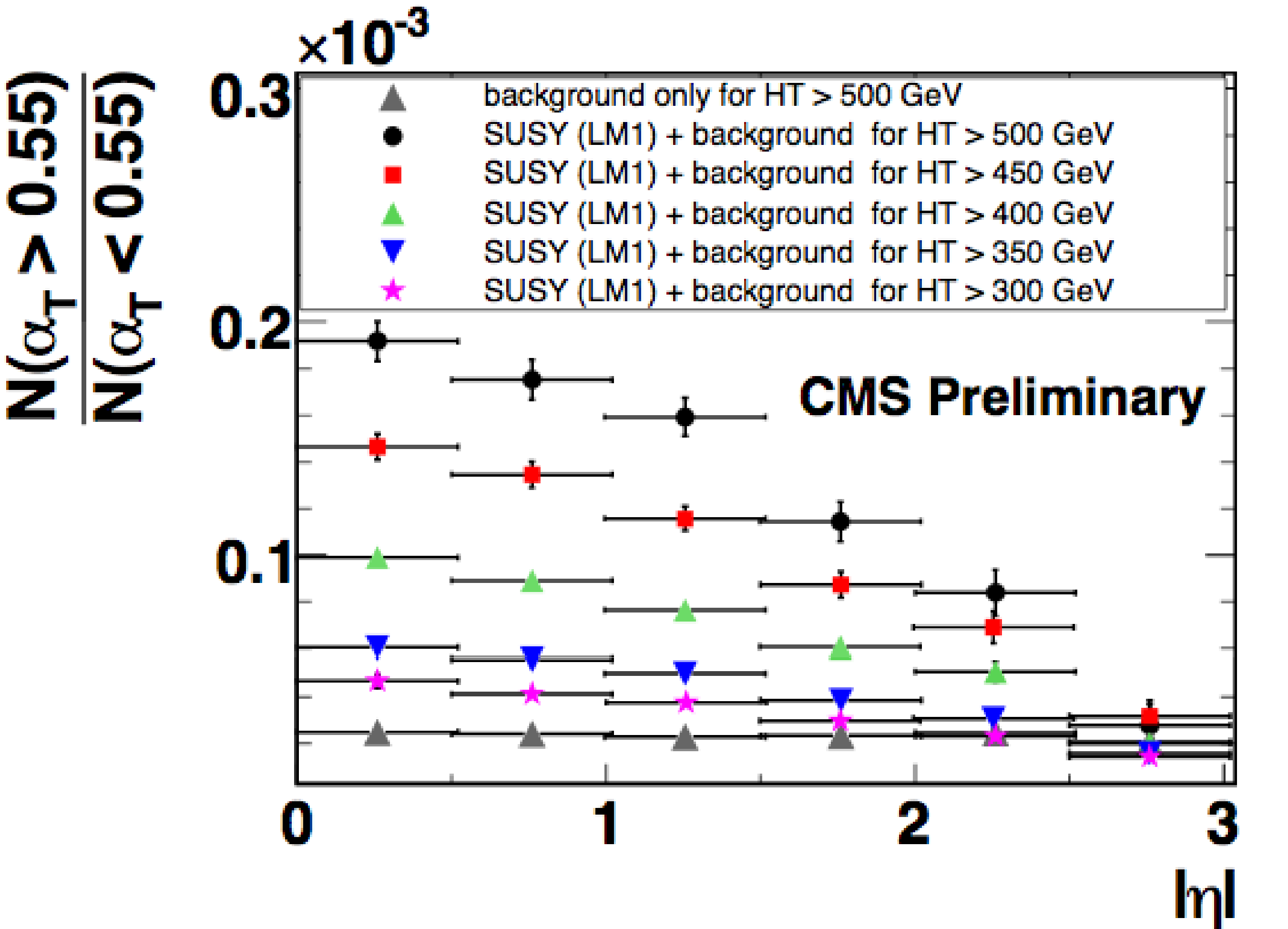
- Predicted BKGD agrees well with simulated BKGD
- SUSY LM1 leads to **significant** excess in signal region
- Method also verified with systematic variations

Test Background Estimation from Data

Variation of HT cut

Idea:
Increase background to check that $R\alpha_T$ is flat in $|\eta|_{j1}$ when signal sufficiently diluted

- Loosen HT cut to decrease signal to background ratio.



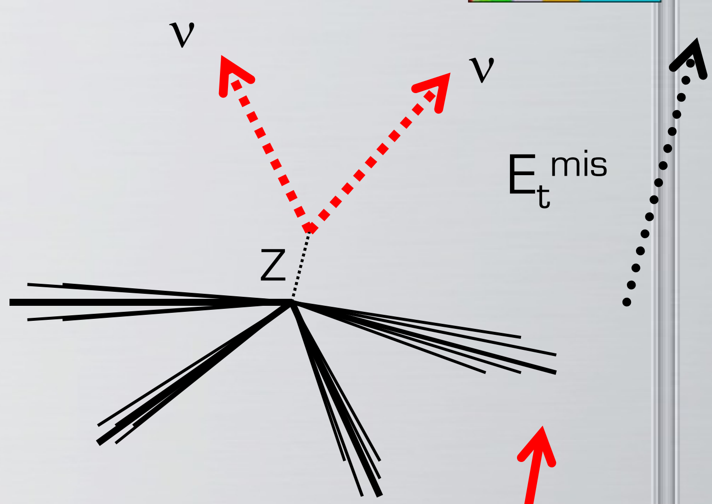
- As HT loosened $|\eta|_{j1}$ dependence gets flatter

=> Clear indication that at HT > 500 GeV signal is present

Data Driven Background Estimations

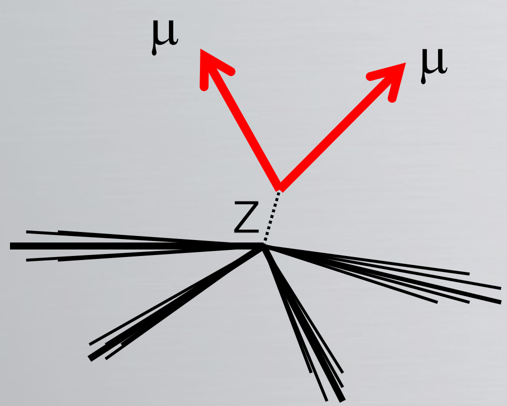
An illustrative example: $Z \to \nu\nu$+jets

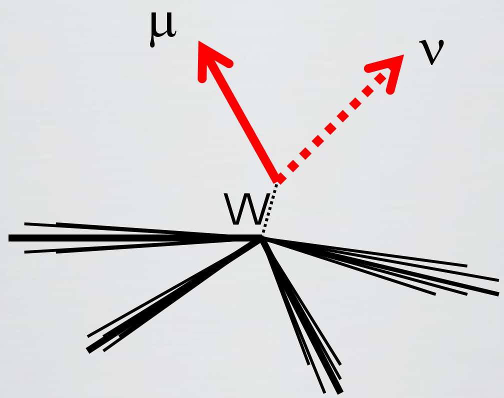
Irreducible background for Jets+E_t^{mis} search

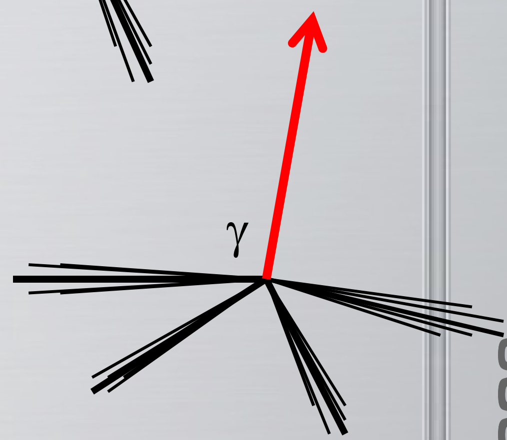


Data driven strategy:

- define control samples and understand their strength and weaknesses:







$Z \to ll$+jets

Strength:
- very clean, easy to select

Weakness:
- low statistic: factor 6 suppressed wrt. to $Z \to \nu\nu$

$W \to l\nu$+jets

Strength:
- larger statistic

Weakness:
- not so clean, SM and signal contamination

γ+jets

Strength:
- large stat, clean for high E_γ

Weakness:
- not clean for E_γ<100 GeV,
- possible theo. issues for normalization (u. investigation)

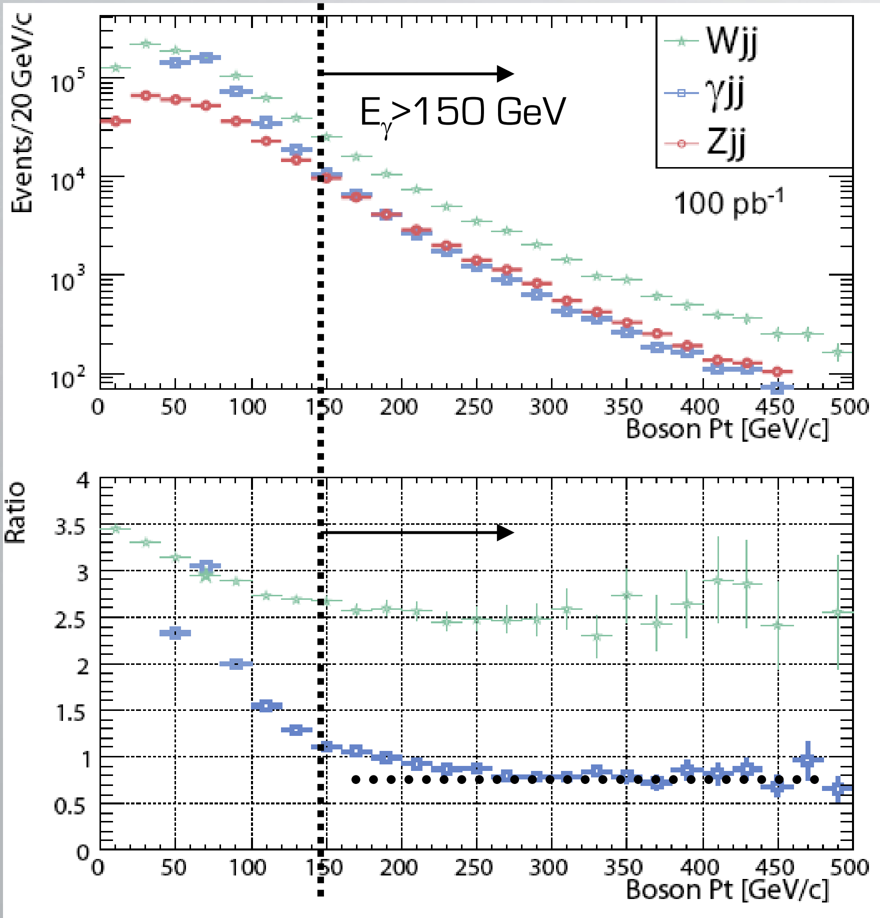
γ+jets: Estimate Z to invisible

γ+jets selection & properties:

- E_γ>150 GeV

→ clean sample: S/B>20

→ ratio σ(Z+jet)/σ(γ+jet) constant

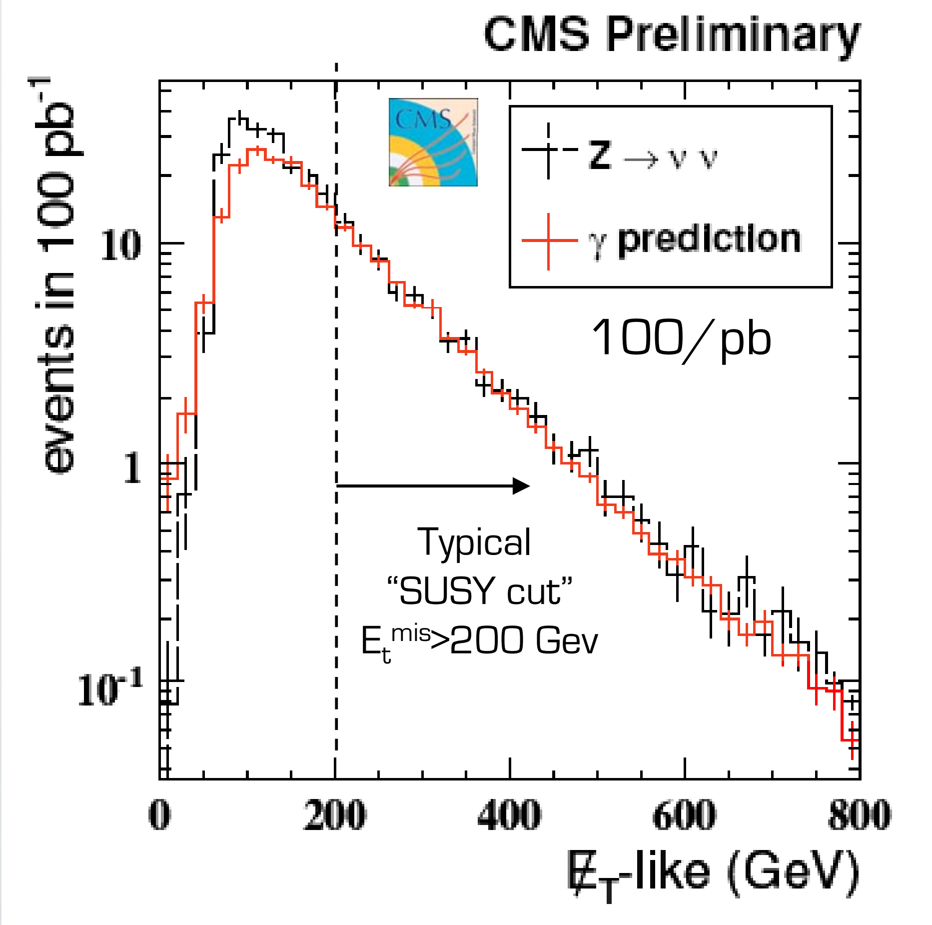


γ+jets: Strategy:

- remove γ from the event:

→ γ becomes E_T^{mis}

- take σ(Z+jet)/σ(γ+jet) for E_γ>200 GeV from MC or measure in data



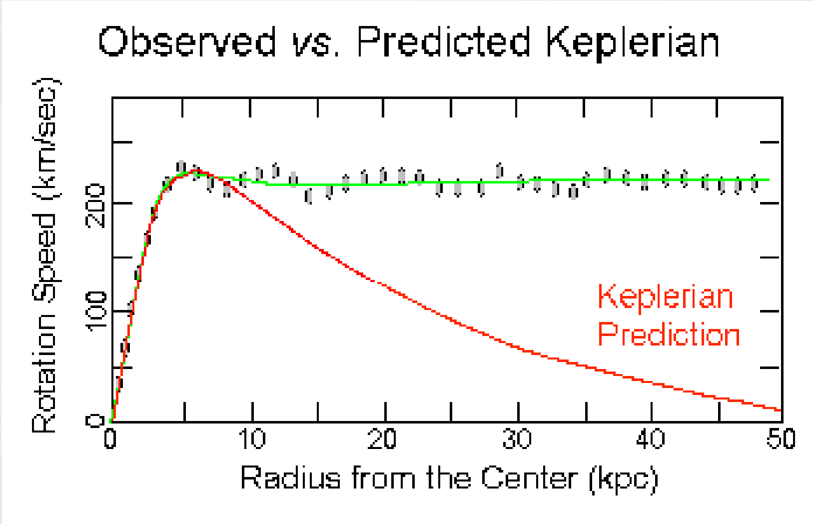
Conclusions

- Good prospects to find Dark Matter Candidate at LHC
- Di-jet analysis promising, exploiting particular event topology
  - α (α_T) and $\Delta\phi$ very powerful
  - Shown results do not rely on calorimetric MET
- Data-driven background determinations have been developed
  - Subtraction of all backgrounds using matrix method — CMS PAS SUS-08-005
    - checks on real data in place
  - $Z \to \nu\nu$ can be obtained from γ + jets — CMS PAS SUS-08-002
- Extension to calorimetric MET independent multi-jet analyses under study
- Benchmark points (e.g. LM1) could be observed in dataset of ~100pb-1
  - Assuming detector performance is understood
- Eagerly awaiting first collision data in fall of this year
  - Exciting times are ahead!

BACKUP

Experimental Evidence for Dark Matter

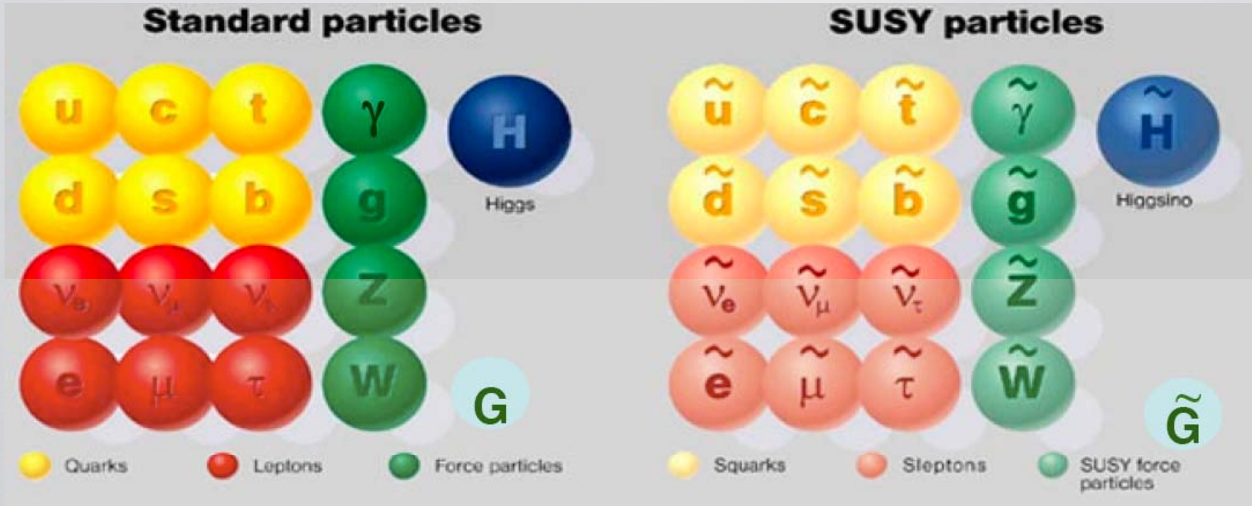
- Zwicky1933
  - rotation frequencies of galaxies
  - high rotation speed at large radii suggests matter far from the center of the galaxy that is not emitting light
  - Dark matter within the galactic halo



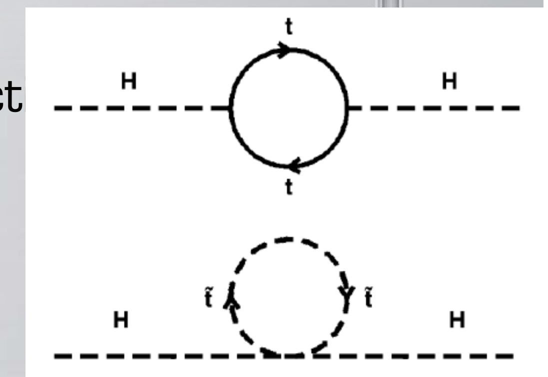
- Bullet cluster
  - collision of two galaxy clusters
  - mass distribution shown in blue
    - determined with gravitational lensing
  - hot gas distribution in red
  - Most of the mass does not interact, only visible matter (gas) is slowed down

Credit: X-ray: NASA/CXC/CfA/M.Markevitch et al.; Optical: NASA/STScl; Magellan/U.Arizona/D.Clowe et al.; Lensing Map: NASA/STScl; ESO WFI; Magellan/U.Arizona/D.Clowe et al.

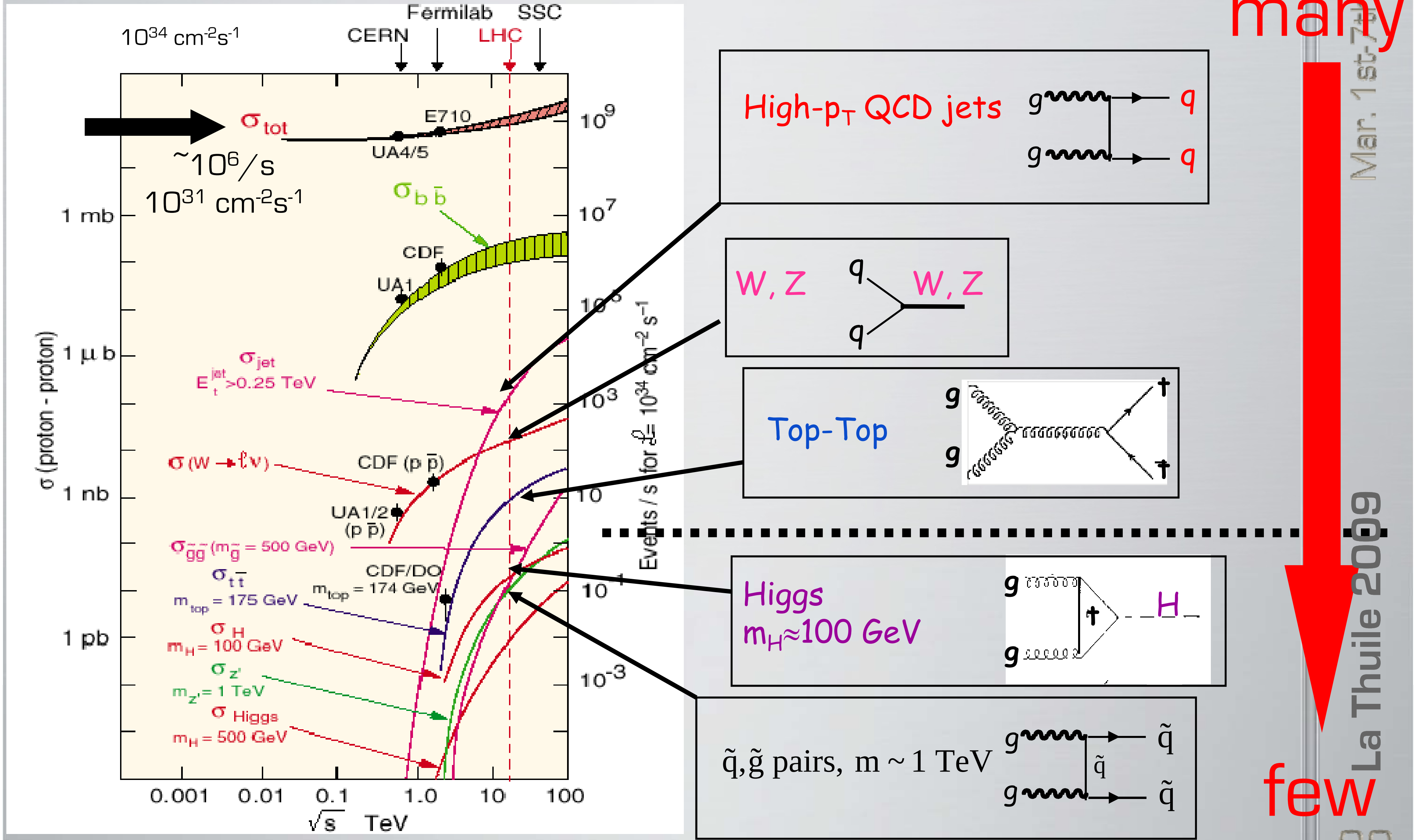
Brief intro to SUSY



- SUSY partner for every SM particle (with ½ unit of spin different)
  - spin 0 Sfermions (squark, sleptons)
  - spin ½ Gauginos (chargino, neutralino)
- SUSY mass scale expected to be ~1TeV in order to:
  - Solve hierarchy problem (stabilize Higgs mass to radiative correct
  - Allow unification of strong and electroweak forces
  - Provide sensible dark matter candidate (R-parity)
  - Naturalises scalar (Higgs) sector of SM
- Downside of SUSY
  - Large parts of parameter space ruled out already
  - Many parameters



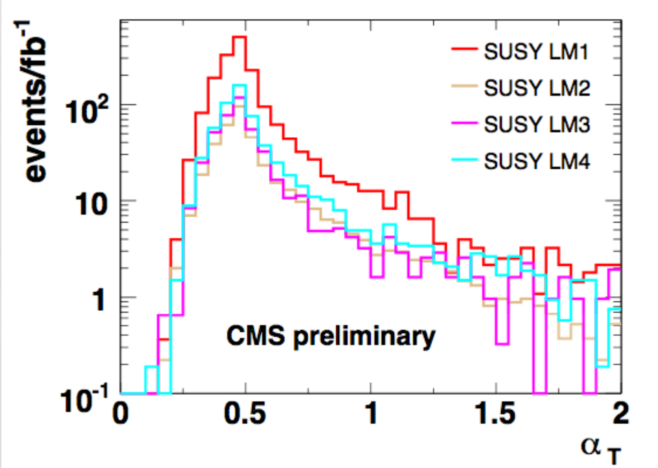
What do we expect to see at LHC?



A closer look at SUSY yields

- CMS SUSY benchmark points

Sample	m_0 (GeV)	$m_{1/2}$ (GeV)	A_0	$\tan\beta$	sign(μ)	σ NLO (pb)	(LO) (pb)	lightest $\tilde{q}$ (GeV)	$\tilde{\chi}_1^0$ (GeV)
LM1	60	250	0	10	+	54.86	(43.28)	410 ($\tilde{t}_1$)	97
LM2	185	350	0	35	+	9.41	(7.27)	582 ($\tilde{t}_1$)	141
LM3	330	240	0	20	+	45.47	(34.20)	446 ($\tilde{t}_1$)	94
LM4	210	285	0	10	+	25.11	(19.43)	483 ($\tilde{t}_1$)	112



- Reminder: desired topology is 2 squarks decaying to squarks and 2 neutralinos (LSPs)

Sample	Events	$\tilde{q}\,\tilde{q}$ (invisible)	$\tilde{q}\,\tilde{q}$ (other)	$\tilde{q}\,\tilde{g}$	$\tilde{g}\,\tilde{g}$	other
LM1	432	39%	22%	34%	3%	1%
LM2	132	46%	33%	18%	0%	2%
LM3	138	69%	17%	12%	0%	2%
LM4	195	49%	10%	36%	3%	1%

For comparison:
QCD: 0
$Z \to \nu\nu$: 57
W/Z: 19
Total: 76

- Dominated by squark-squark, but not only:
  - Squark - gluino contribution, where gluino decays to squark+quark
  - In LM1: small mass difference between gluino and squark => low p_T 3rd jet

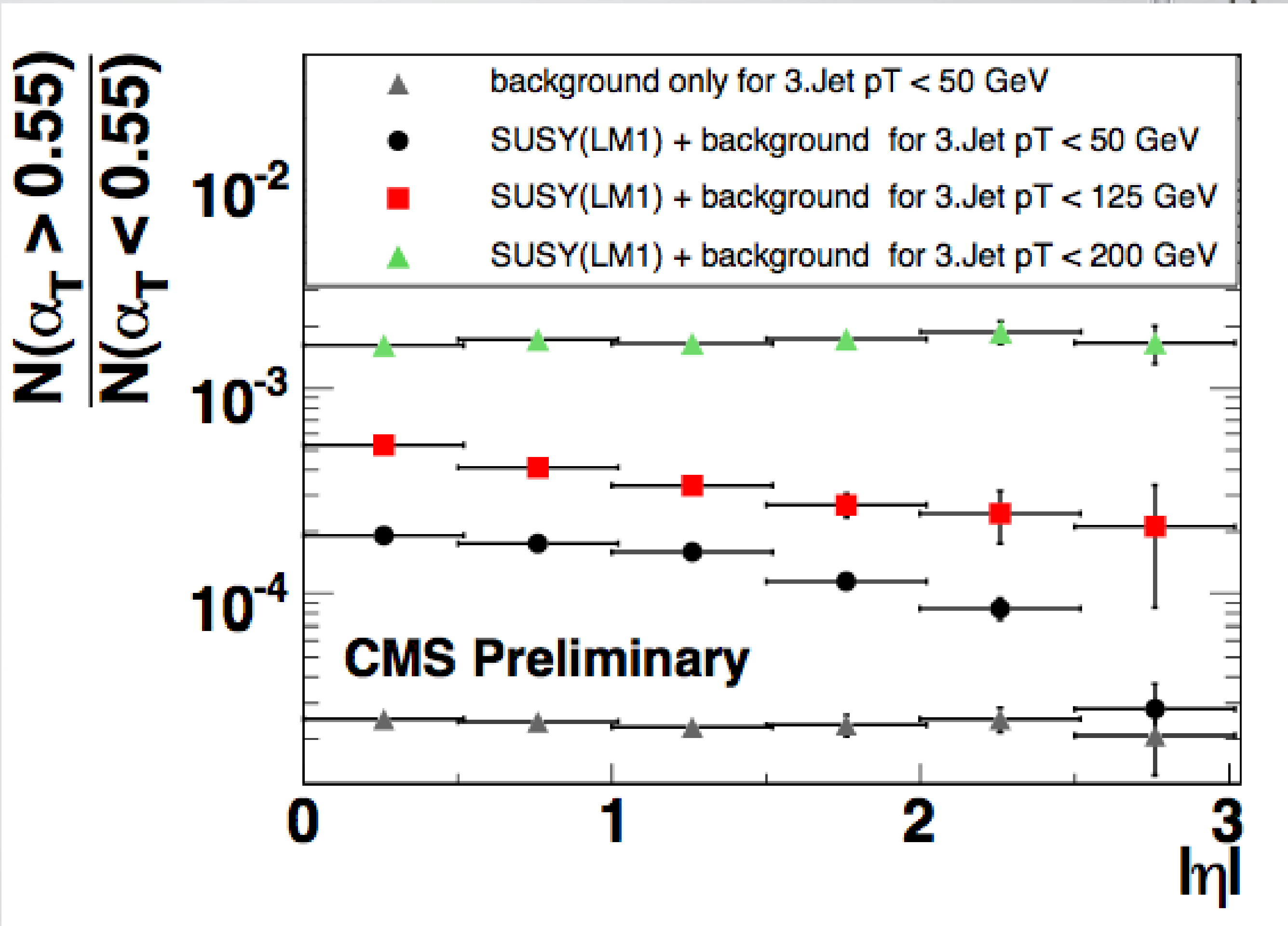
Production process	$p_T^{j3} < 30\,\mathrm{GeV}$	$p_T^{j3} < 50\,\mathrm{GeV}$	$p_T^{j3} < 70\,\mathrm{GeV}$
$\tilde{q}\,\tilde{q}$	80%	61%	51%
$\tilde{q}\,\tilde{g}$	18%	34%	44%
$\tilde{g}\,\tilde{g}$	1%	3%	5%

  - Indeed observe increase in squark-gluino contribution when relaxing 3rd jet veto

Background estimation from data (II)

Variation of 3rd jet p_T

Idea:
dilute signal by increasing background contribution
Loosen cut on 3rd jet p_T to create missing E_T
=> tail in $\alpha(\alpha_T)$



Test if $R\alpha_T$ is stable
Slope should be observed when signal contribution becomes sizable

⇒ Slope is observed for hard enough jet veto