Maxime Gouzevitch

On jet substructure

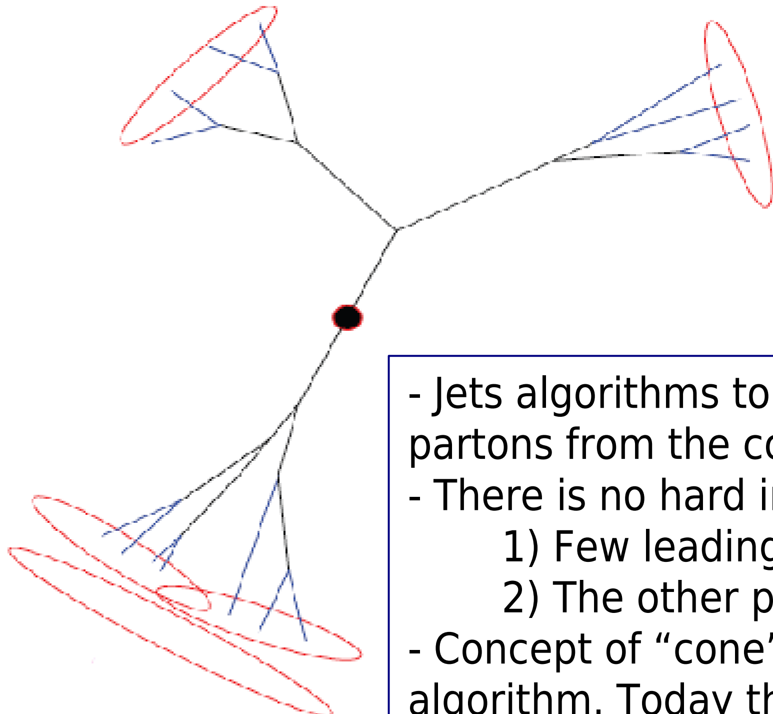
DGR Terascale (Clermont-Ferrand, April 2012)

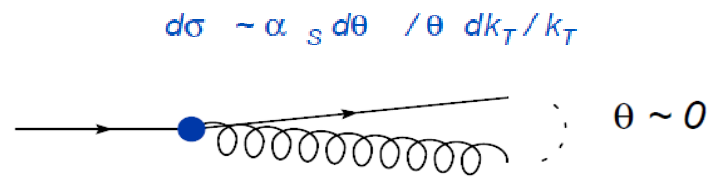
1) On QCD jets substructure.
2) Heavy resonances and boosted topology.
3) Superstructure and substructure.

Help from: A. Hinzmann (CERN - CH), Petar Maksimovic (JHU - US).

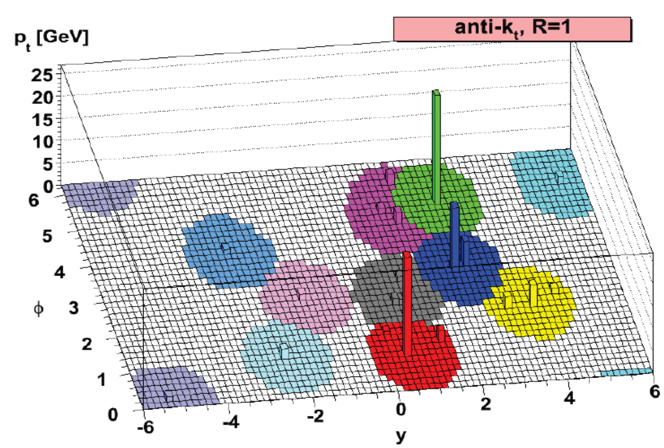
On QCD jets substructure

1.1) How the QCD jets are designed?





- Jets algorithms to extract the information on the “hard” partons from the collinear QCD radiation.
- There is no hard intrinsic scale. Typically:
    1) Few leading particles.
    2) The other particles distributed in log(θ), symmetric in φ.
- Concept of “cone” in θ x φ aimed by a large bunch of algorithm. Today the most successful anti-k_T.

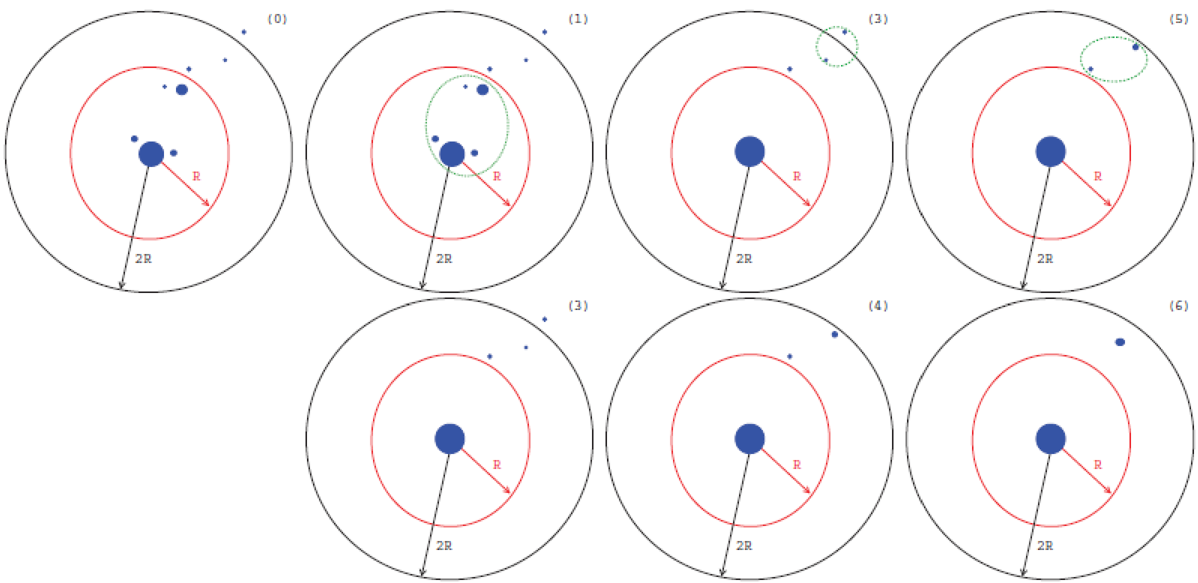


1.2) Sequential jet algorithms

- Iterative clustering using the distances:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} ,$$
$$d_{iB} = k_{ti}^{2p} ,$$

- particles recombined till $d_{ij} < d_{iB}$
- anti-k_T: p = -1, “big ogre eat first”, coherent with idea of 1 leading particle coming from the parton dominating a population of radiative gluons.
- Cambridge-Aachen: p = 0. Purely geometrical. Used for substructure.
- Resulting: n jets with Δ > R ~ 0.5-1



Clustering anti-k_T

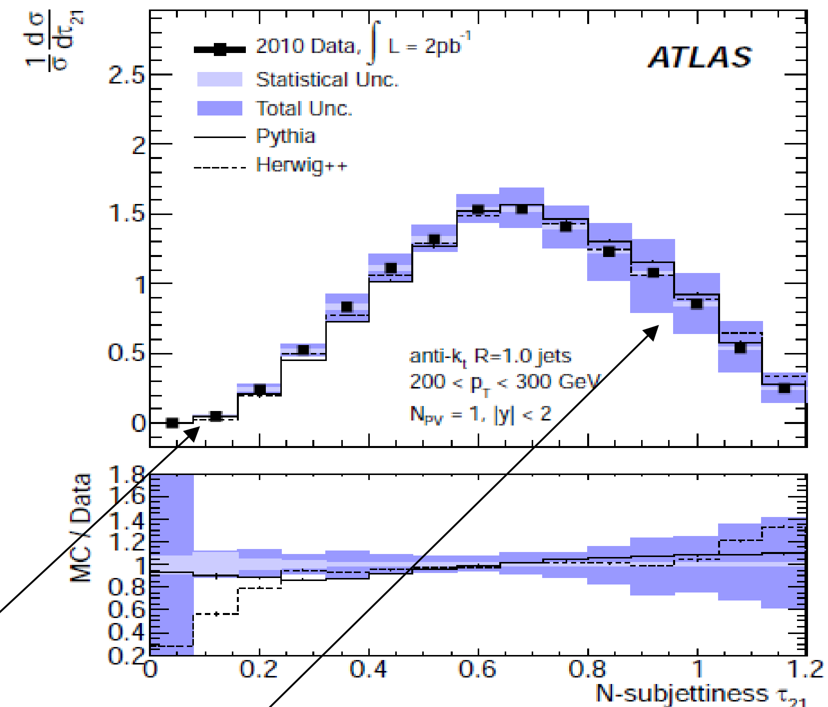
1.3) A measurement of QCD jets substructure

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \times \min(\delta R_{1,k}, \delta R_{2,k}, \ldots, \delta R_{N,k})$$

$$d_0 = \sum_k p_{T,k} R,$$

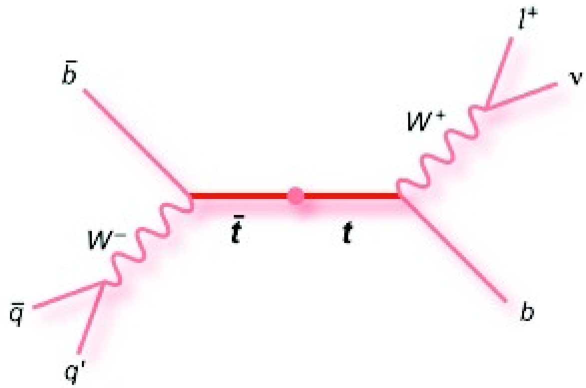
arXiv:1203.4606

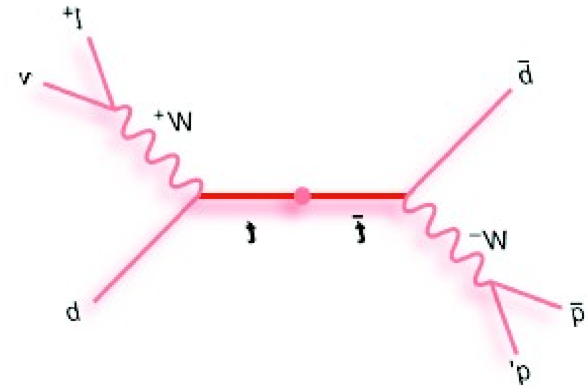
- τ_N: designed to check how well a jet is described by $N_{\text{sub-jets}}$.
- $\tau_{12} = \tau_2/\tau_1$: tells if a jet is better described by 2 sub-jets than by 1 narrow jets.



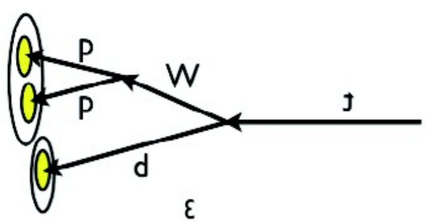
Prefer 2 sub-jets

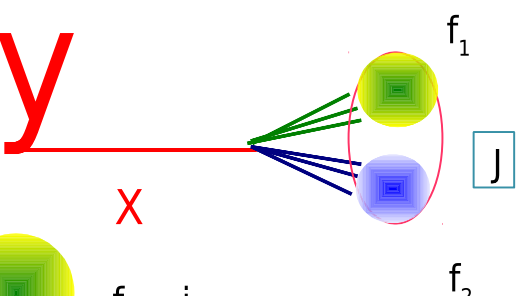
Prefer 1 narrow jet

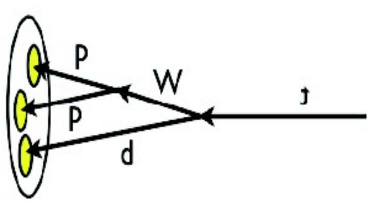


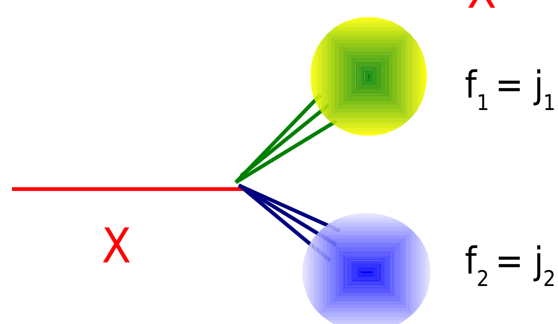


Heavy resonances and boosted topology

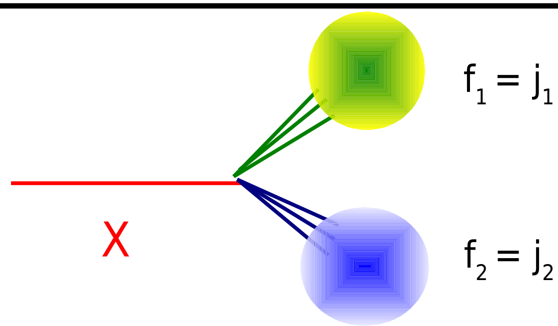


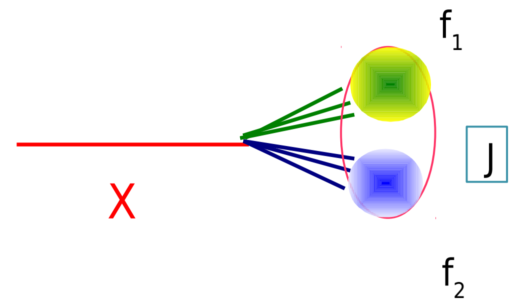






2.1) Topology of a boosted 2 body decay





X: heavy resonances

f_1, f_2 : 2 collimated flux resulting from X decay.

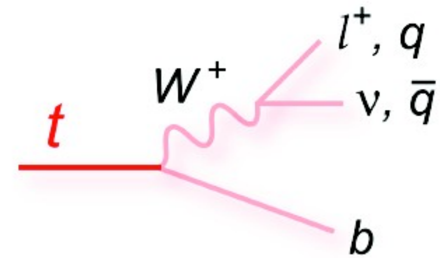
J, J_1, J_2 : reconstructed jets.

- If p(X) >~ M(X) the f_1 and f_2 appears as 2 different jets.
- If p(X) >> M(X) >> M(f_1,f_2) then the flux f_1 and f_2 merge into 1 jets.

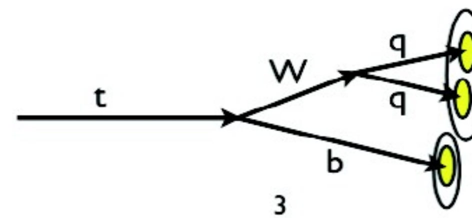
  Substructure ~ presence of an intrinsic mass scale.
- Till recently (late 90th) there was no such resonances.
  - B-tagging is already a “ kind of “ jets substructure, but heavy mesons have a mass close to the confinement.
  - W, Z mesons, even top quark produced nearly at rest.

2.2) Top physics cases

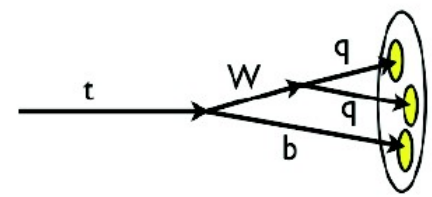
- Low mass regime
  - isotropic event topology
  - standard top selection
  - combinatorial event reconstruction
  - b-tagging

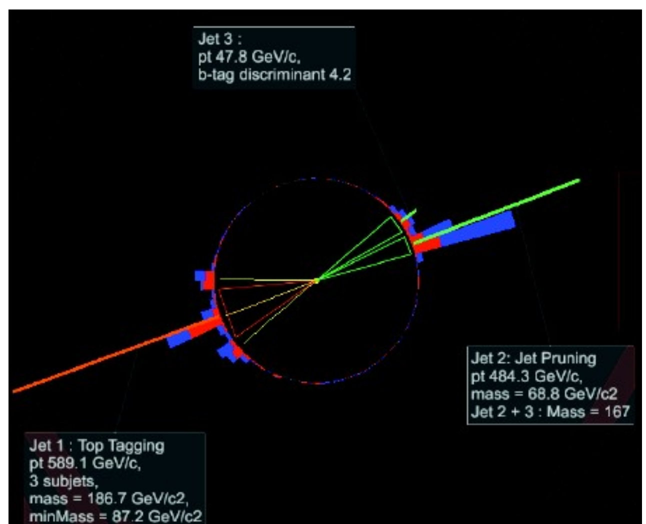


- Intermediate mass regime
  - partially merged hadronic top decays
  - neither high nor low mass work well
  - b-tagging works



- High mass regime
  - relativistic (boosted) top quarks
  - hadronic decay products (jets) merge
  - need to break them into sub-jets
  - no b-tagging



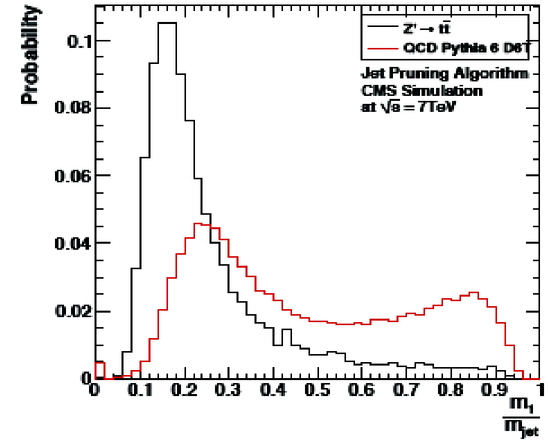


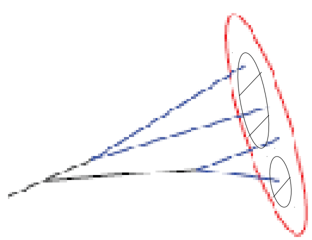
2.3) Detecting the substructure: mass drop

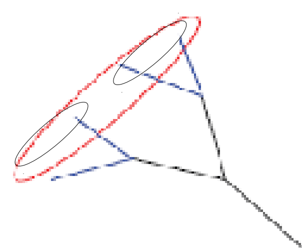
$$\langle M_J^2 \rangle_{NLO} \simeq \overline{C}\left(\frac{p_J}{\sqrt{s}}\right)\alpha_s\left(\frac{p_J}{2}\right) p_J^2 R^2,$$

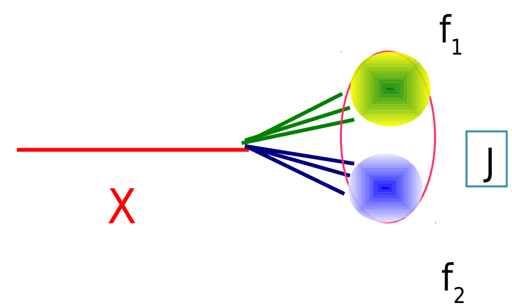
QCD: mass comes from large angles. Let's play a game:

1) Take the Cambridge-Aachen algorithm (geometrical)
2) Cluster jets.
3) Take jet J. Uncluster the last step:
   2 p_T ordered sub-jet J_1, J_2.
4) Compare $M(J)$ to $M(J_1)$, $M(J_2)$.









QCD :

Soft large angle radiation
$M(J_1) \sim M(J)$

QCD :

Hard radiation
$M(J_1) \sim M(J)/4$

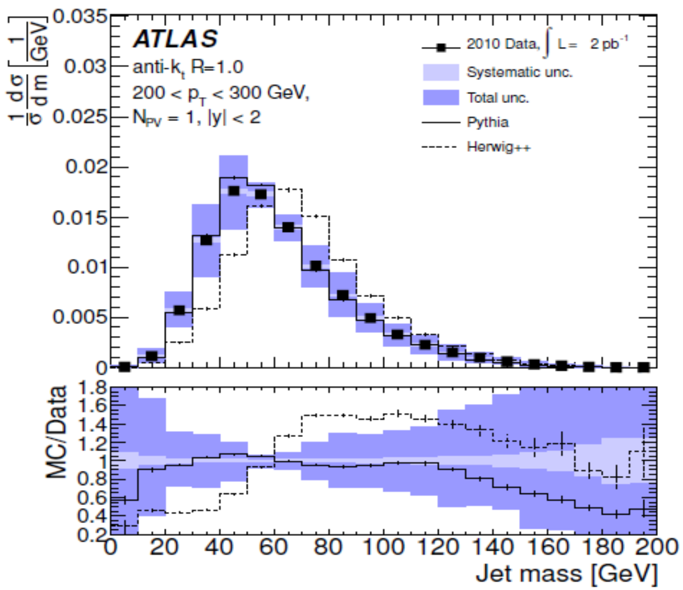
Resonance decay :

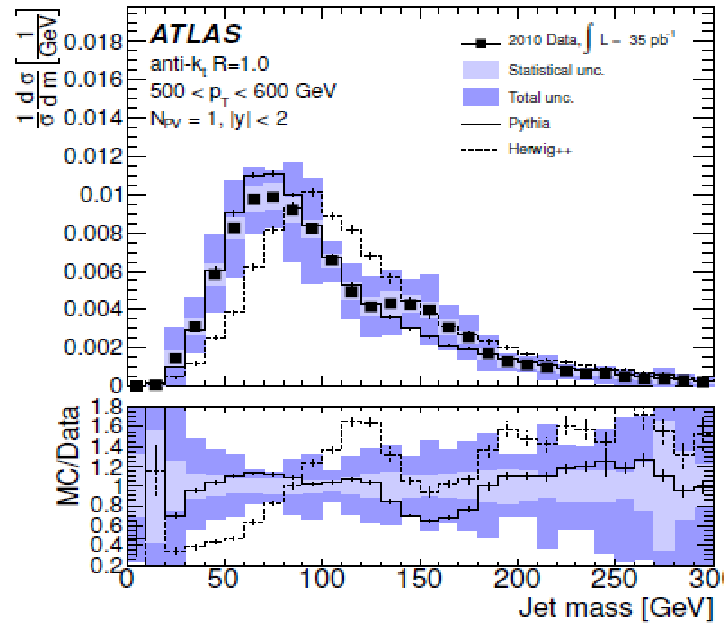
$M(J_1) << M(J)$

2.4) Model the mass in QCD

$$\langle M_J^2\rangle_{NLO} \simeq \overline{C}\left(\frac{p_J}{\sqrt{s}}\right)\alpha_s\left(\frac{p_J}{2}\right)p_J^2R^2,$$

arXiv:1203.4606

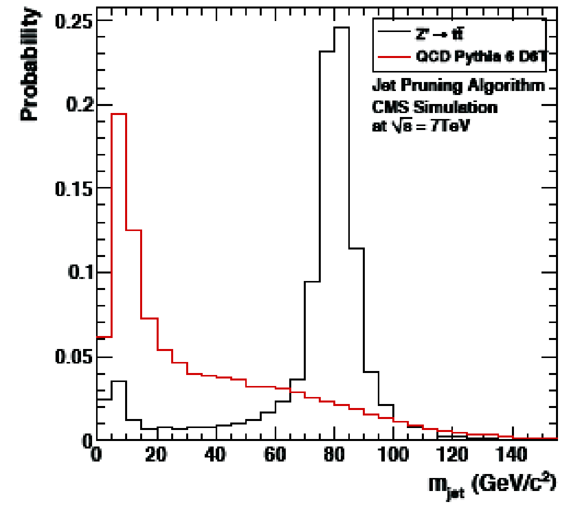


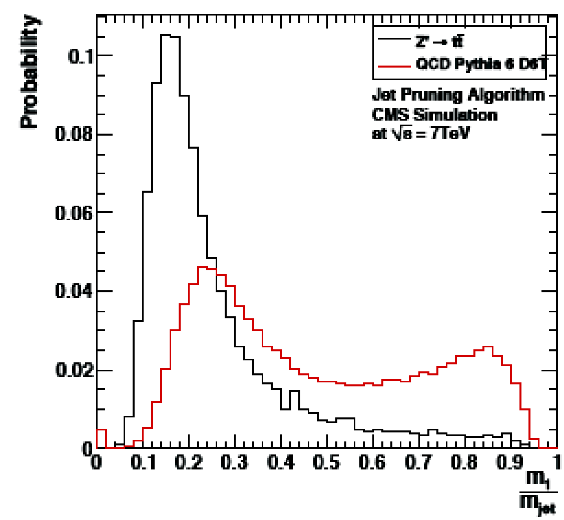


- Jet mass is well described by the LO+PS MC in shape.
- The exact position of the peak is lower in PYTHIA than in data (less FSR radiation usually); too high in Herwig.
- Need to explore full 2011 statistics to cover the large mass tail and jets with p_T ~ 1 TeV.
- CMS measurements coming soon.

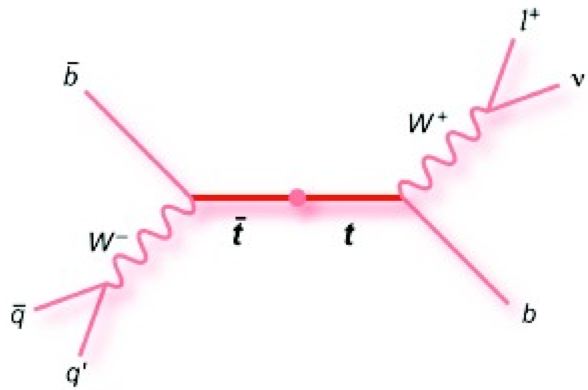
2.5) W tagging algorithm

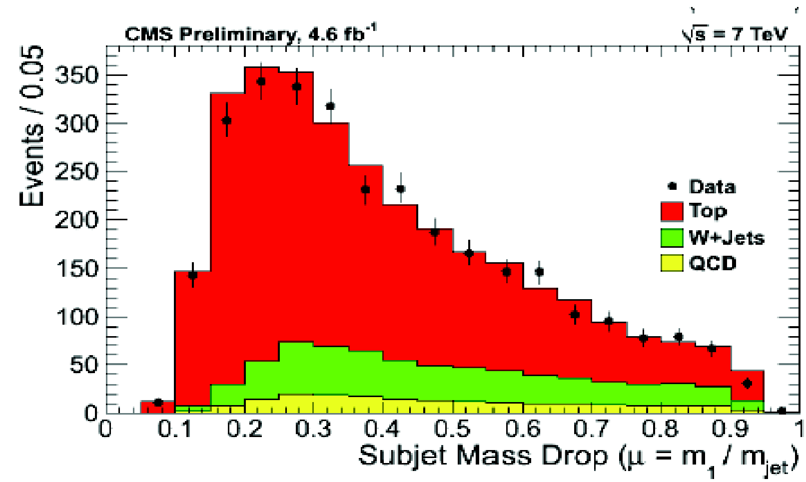
- Jet pruning
  - Ellis, Vermillion, Walsh (arXiv:0903.5081)
  - Improves mass resolution by removing soft, large angle particles
- Cluster jets with Cambridge Aachen (CA) with R=0.8
- Undo last step of jet clustering to find two subjets
- Tag Ws by
  - Butterworth et al. (arXiv:0802.2470)
  - Pruned jet mass $60 < m_{\mathrm{jet}} < 130 GeV/c^2$
  - Mass drop $\mu \equiv \dfrac{m_{\mathrm{leading\ subjet}}}{m_{\mathrm{jet}}} < 0.4$



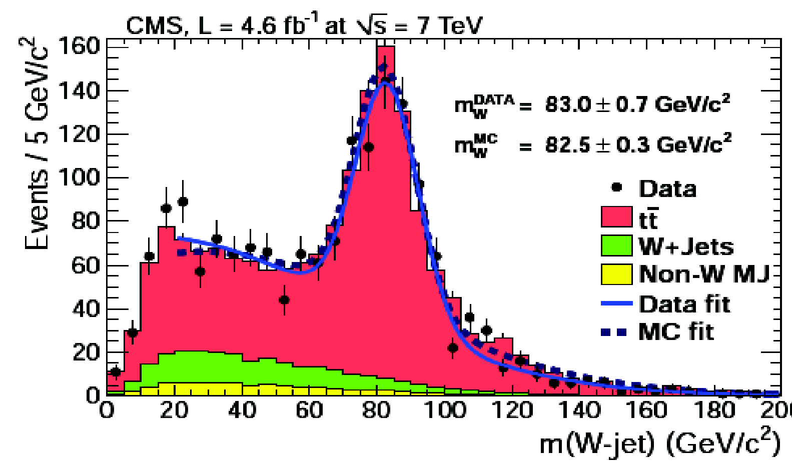


2.6) W tagging algorithm: validation





arXiv:1204.2488

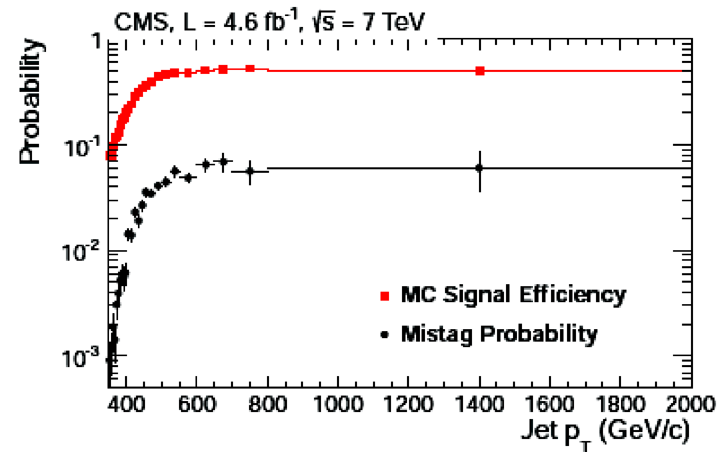


- Look for merged Ws in low mass semileptonic $t\bar{t}$ events
- Use W peak from W-tagged jets to detemine
  - substructure energy correction for MC =1.02 +- 0.01
  - W-tagging efficiency correction for MC = 0.97 +- 0.03
- Madgraph+Pythia Z2 works well

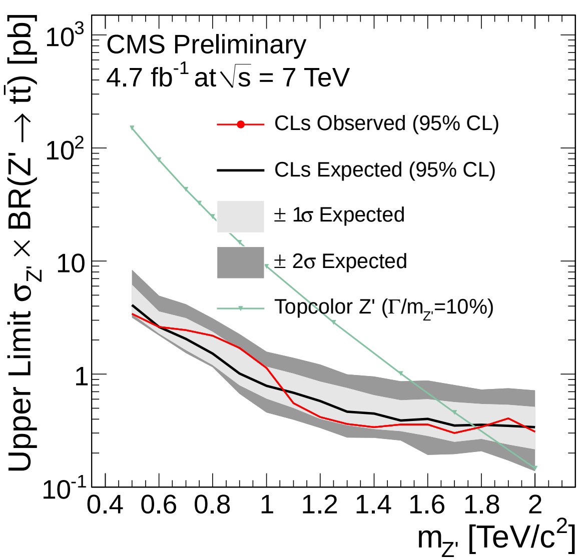
CMS Experiment at the LHC, CERN

2.7) Limits on exotic models

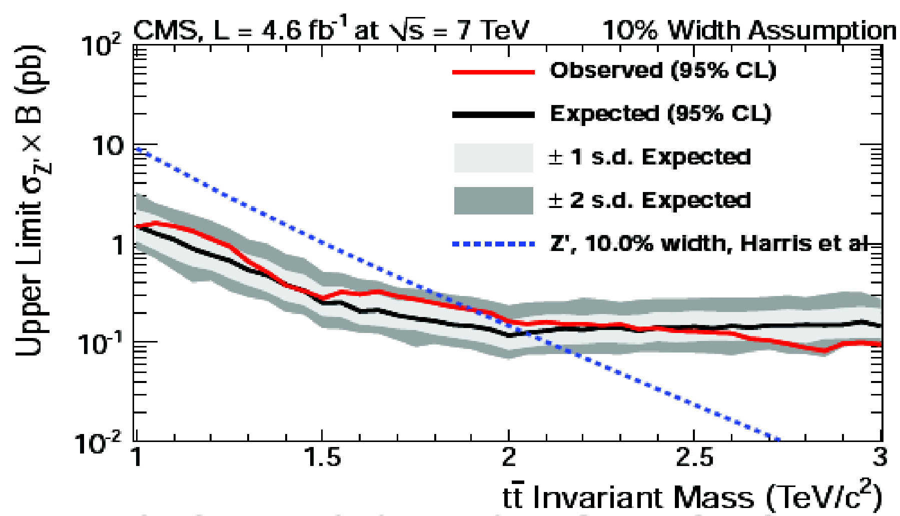
- Fully hadronic ttbar measurement.
- Larger branching fraction than semi-leptonic.
- 2 configurations used in the analysis:
  - W-jet (2 sub-jets) + b jet
  - Top jet: (3 sub-jets)
- Efficient above 1 TeV where discovery potential like semi-leptonic.



arXiv:1204.2488



Non boosted semi-leptonic



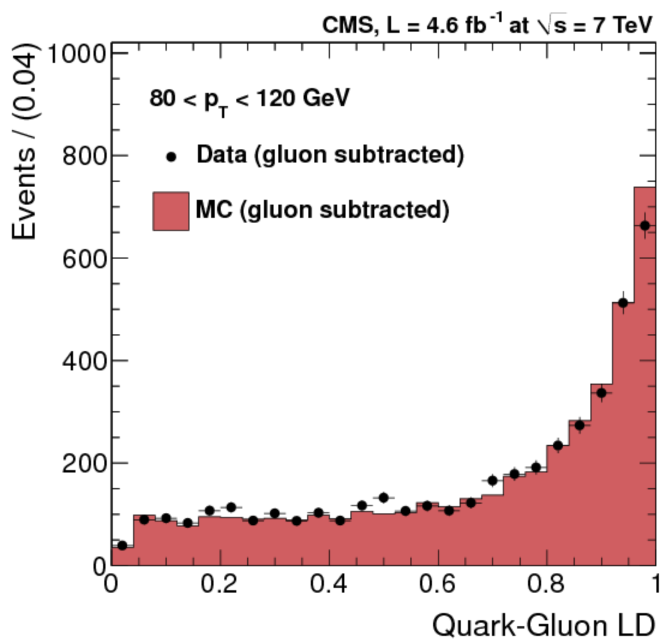
Boosted full hadronic

2.8) Going further

- Mass drop, pruning, filtering ~ “cut based” algorithms. We start to understand them and pushing to the limits.

- Next step: One can imagine to have a generalized MVA like analysis based on jet shapes : dM/dR, dp_T/dR, dN_{part}/dR.
See 10.1103/PhysRevD.83.074023

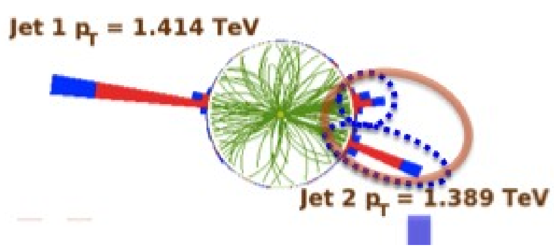
- Already used in CMS/Atlas for Quark-gluon id.

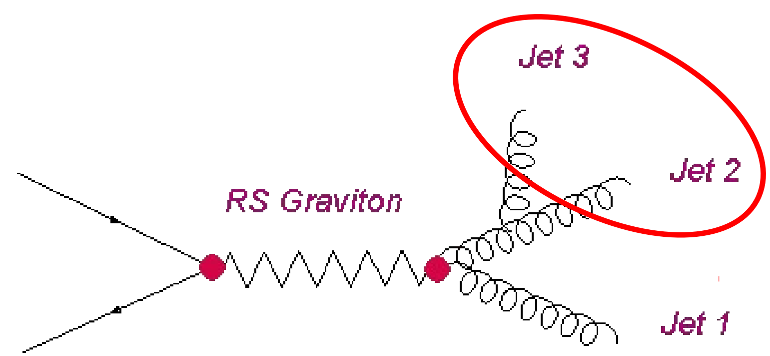


arXiv:1202.1416
$H \to ZZ \to 2l2jets$

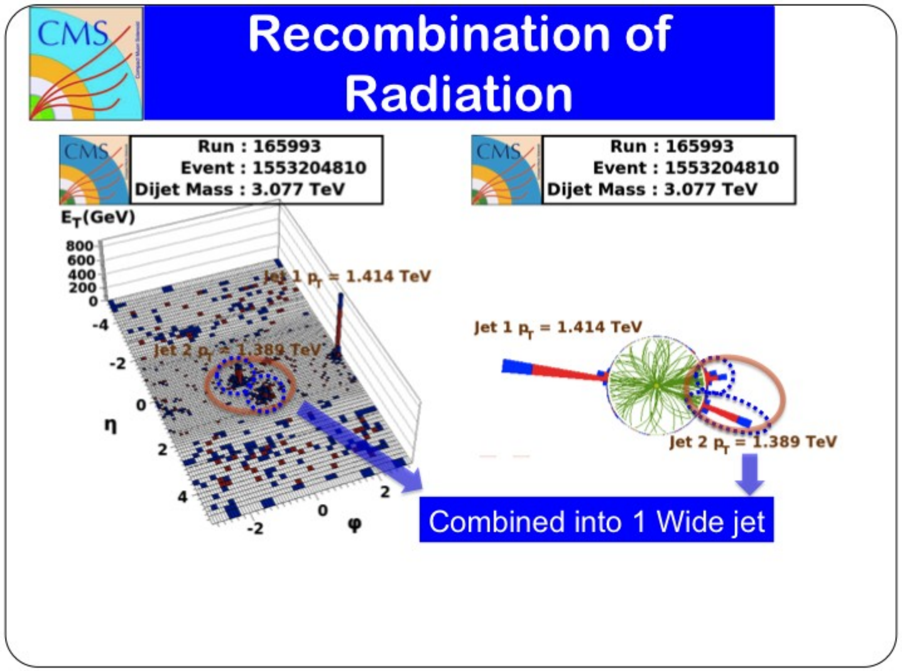
- Expected distributions of QG LD for quark jets.
- The gluon contribution has been subtracted by accessing the MC truth.

Superstructure and Substructure





3.1) Wide jets: Collecting Final state radiation

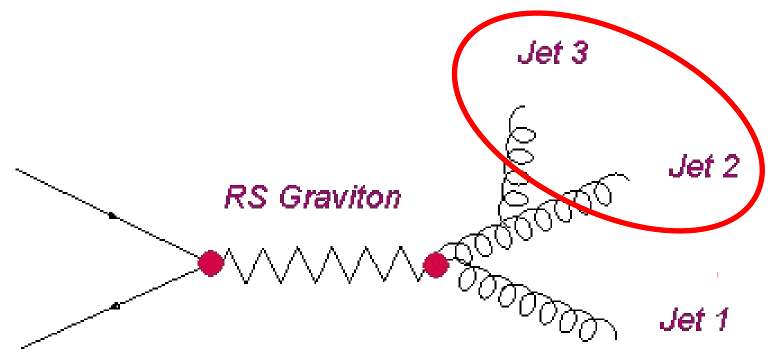


- $X \rightarrow 2$ jets (qq, qg, gg)
- Partons may radiate gluons.
- To better reconstruct M(X) need to collect all FSR:

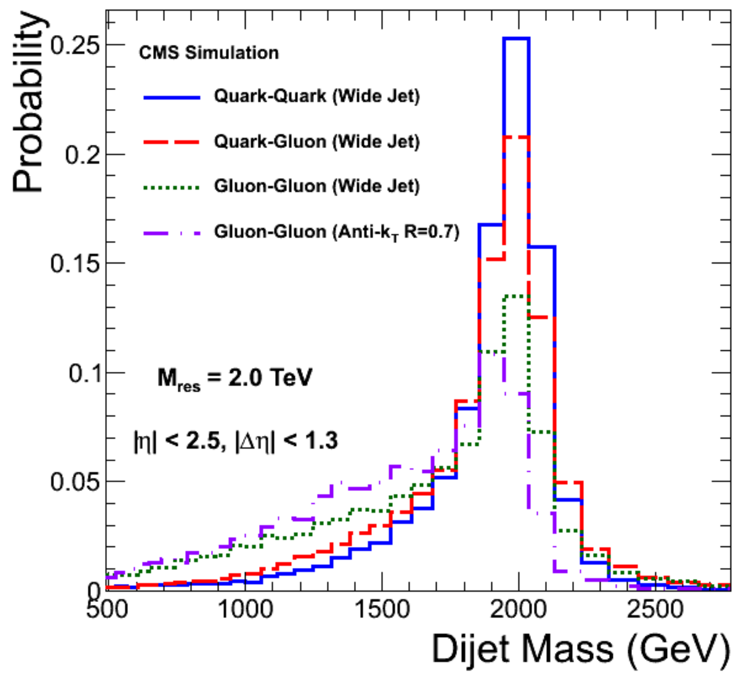
1) Need large radius jets: R > 1.0
2) Need to clean them for pile-up.
3) Need to calibrate.

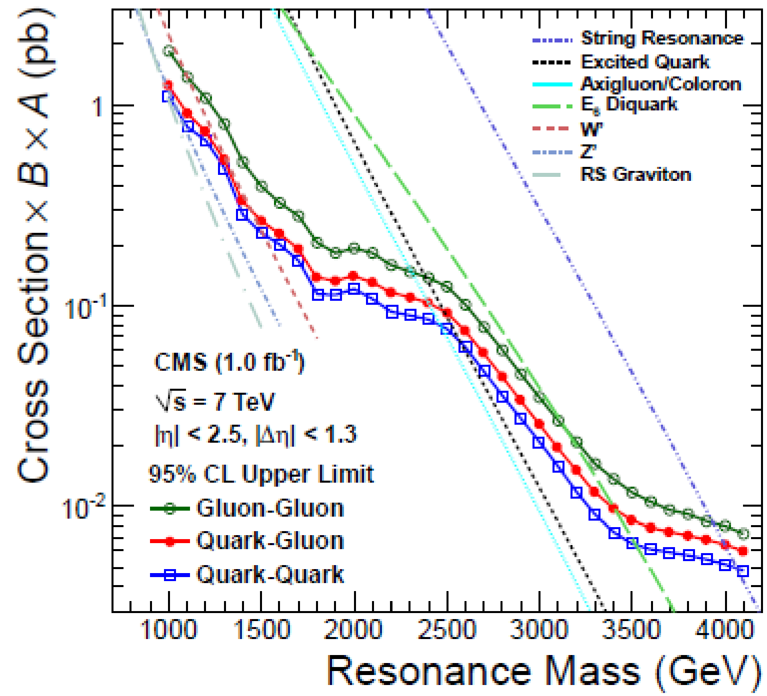
But Atlas and CMS usually maintain only jets with “small” R: CMS: 0.5, 0.7; Atlas: 0.4, 0.6

Idea - make jets with jets :
- Take smallest calibrated jets (R_{low}).
- Make Wide jets (R_{Wide}) with those jets.
- Get pile-up cleaning and large radius “pruning” for free.



3.2) Wide jets



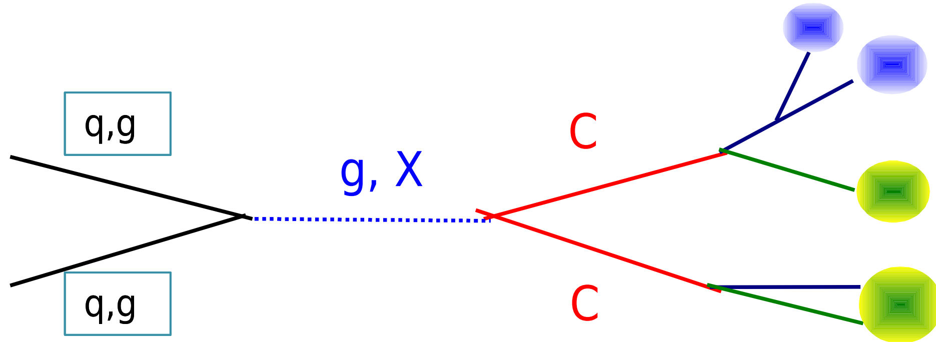


arXiv:1107.4771

- Optimized for discovery potential : R = 1.1
  - Too large radius : catch ISR, pile-up, UE ; large QCD background.
  - Too low radius : loose too much FSR, undershoot M(X).
- Usually M_X ↑, R_{Wide} ↑.
- In gg final state 20% improvement in excluded cross section.

3.3) From substructure to superstructure

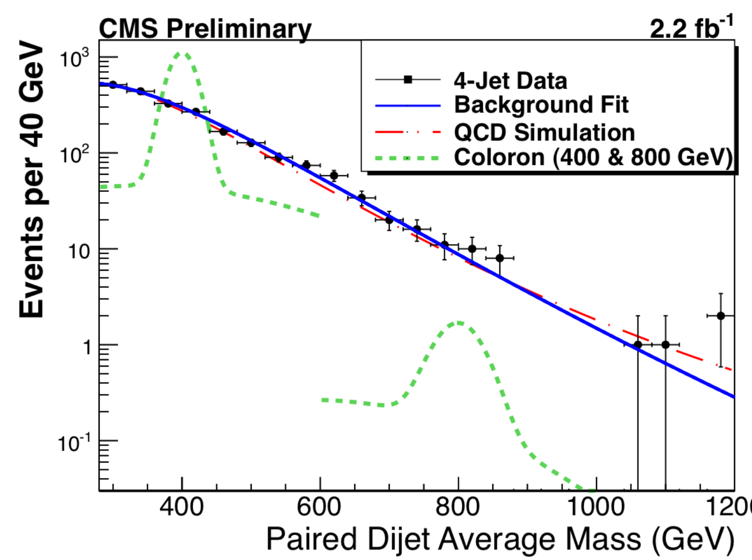
Paired production



- Typical problematic : pair production of heavy resonances.
  – Mass of cascade particles not known.
  – Boosted and unboosted topologies possible.

Idea - a transition algorithm for 2 body decay between boosted and unboosted topology :
  – Check if a jet may be broken into 2 subjets.
  – If no look for an extra jet symmetric in p_T.
  – Add soft FSR jets.

CMS EXO-11-016



Summary and prospects

1) Jets substructure very active field since few years:

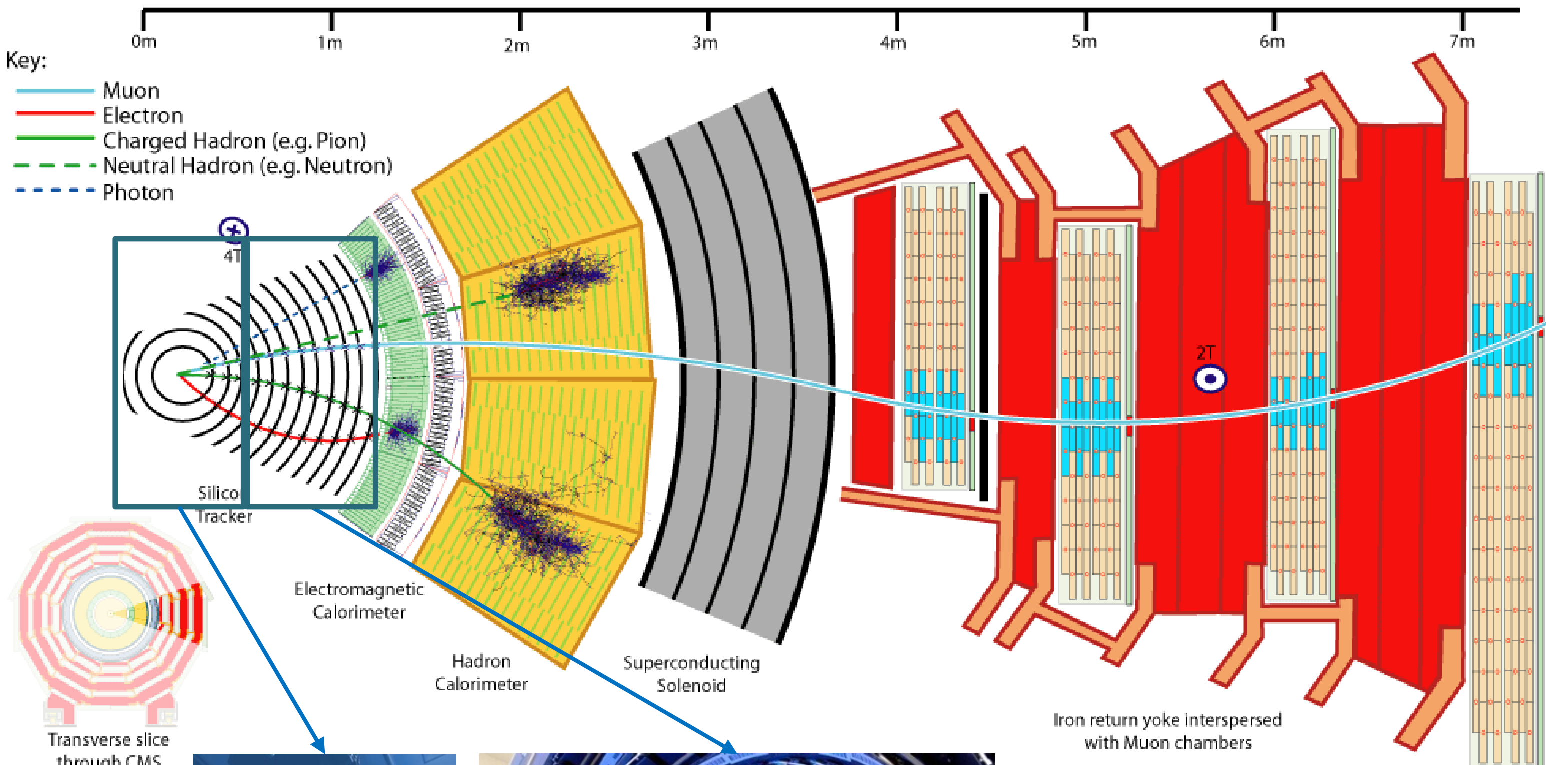
- Is not confined anymore to the specialized MC tuning or QCD study.
- A campaign of systematic QCD measurement just started by CMS and Atlas.
- Look for heavy boosted SM resonances: W, Z, top.
- Search for new boosted resonances in jets final state:
  - $H \rightarrow$ bb, pair produced BSM.
- Systematic Measure

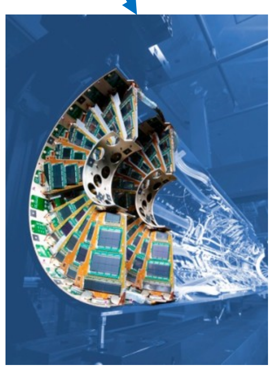
2) Bright future:

- Would have more and more importance while the resonances we are searching for becomes heavier and heavier.
- MVA techniques are invading the area.
- Stay tuned...

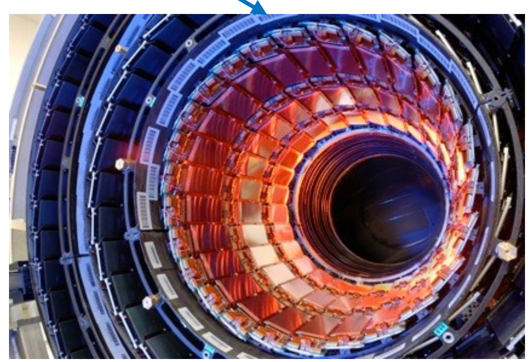
Backup

2) Tracker performance





Pixel



Strips

- *Full silicium technology.*
- *~75.000.000 channels.*

Performance in collisions arXiv:1007.1988

3) Detecting the substructure: mass drop

- JHU Top Tagger (Kaplan, Rehermann, Schwartz, and Tweedie, arXiv: 0806.0848), tweaked by CMS
- Cluster jets with CA R=0.8
- Retrace two steps of clustering sequence back to find subjets
- Variables:
  - Jet mass (~m_{top})
  - Number of subjets
  - Min pairwise mass (~m_W)

$$\begin{aligned} 140 < m_{\text{jet}} &< 250 GeV/c^2 \\ N_{\text{subjets}} &\geq 3 \\ m_{\min} &> 50 GeV/c^2 \end{aligned}$$