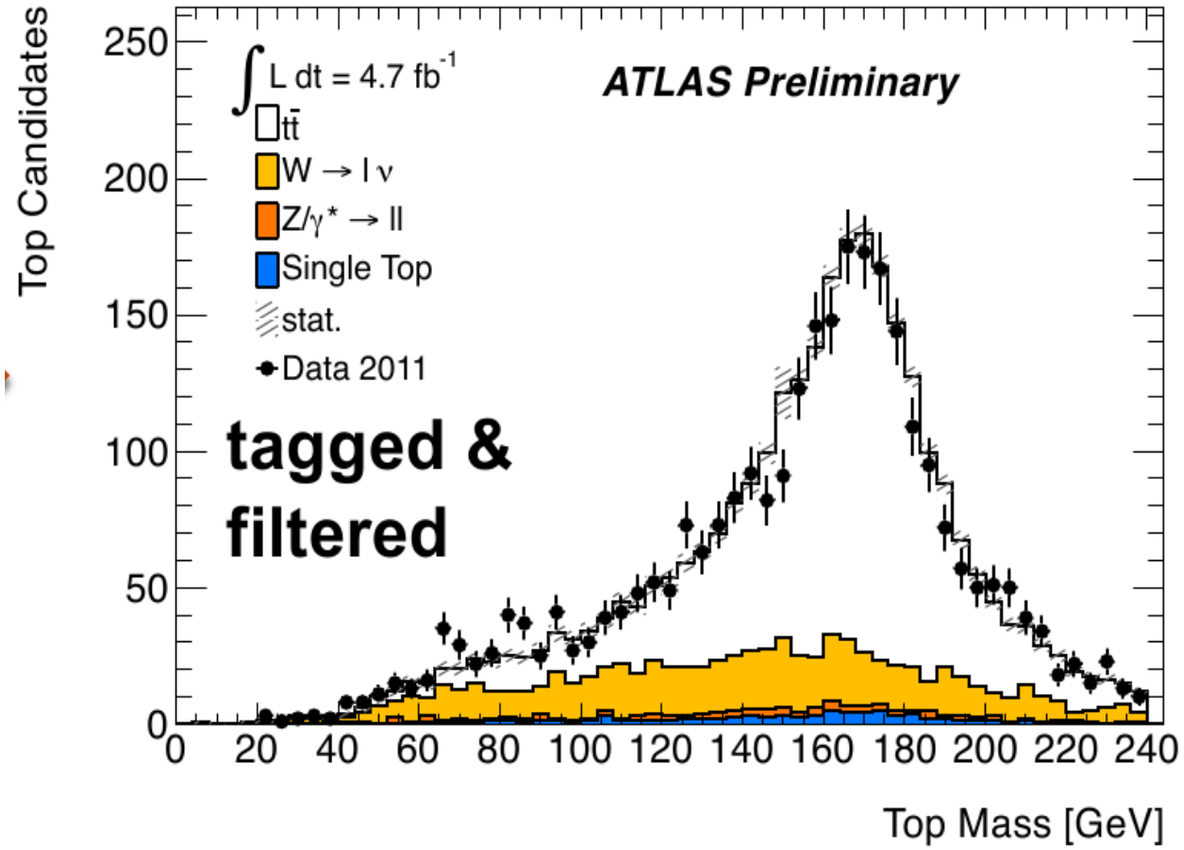
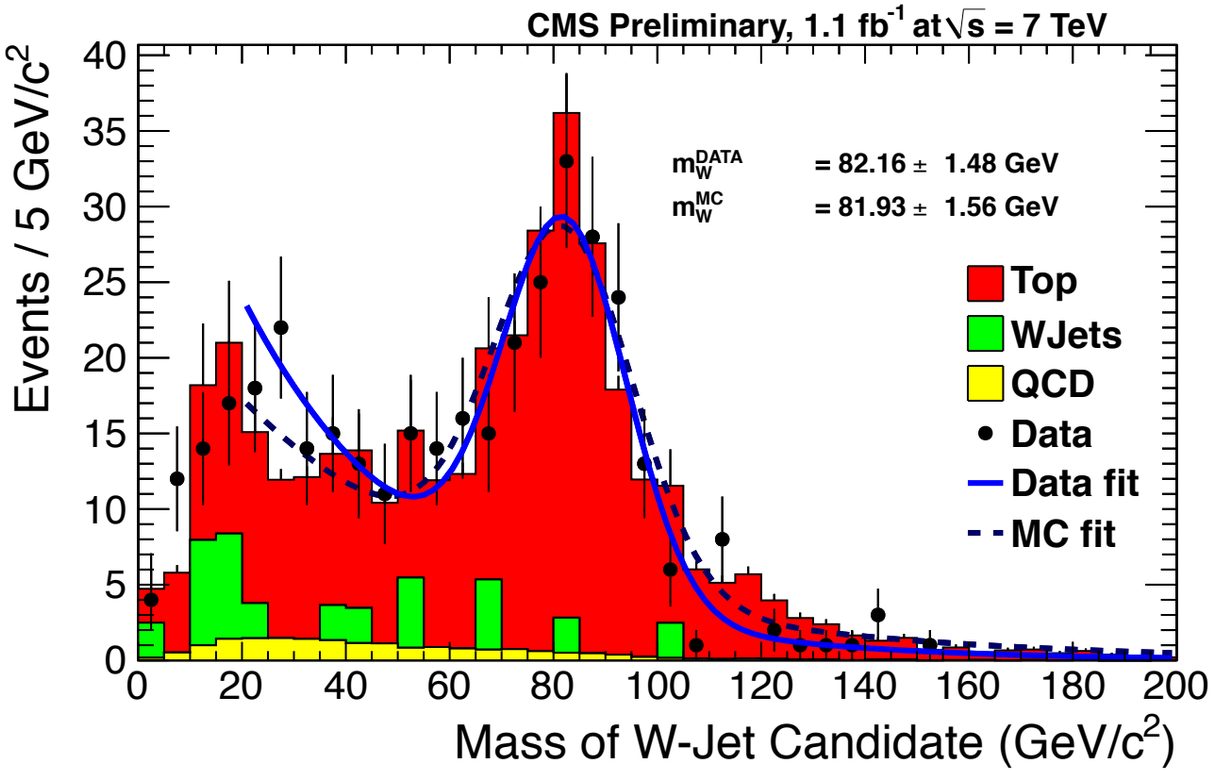
Progress and challenges in jet substructure

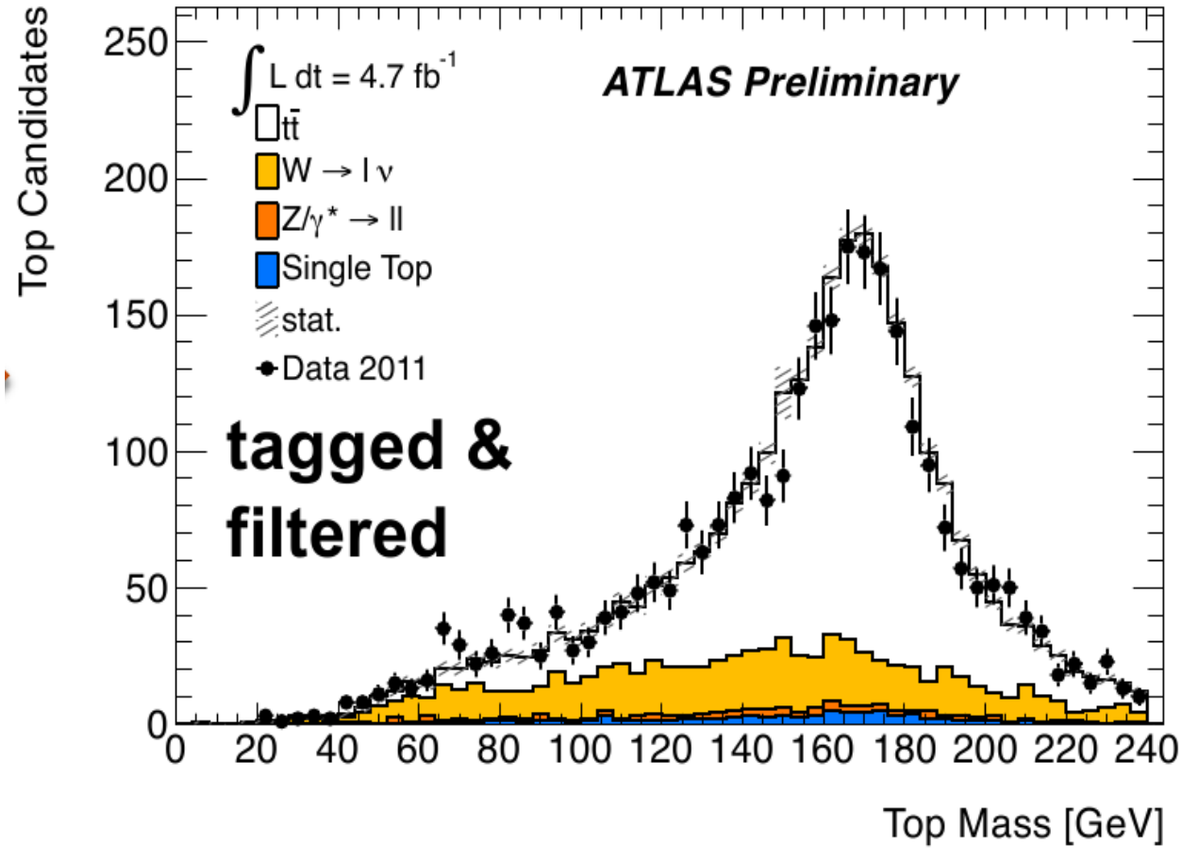




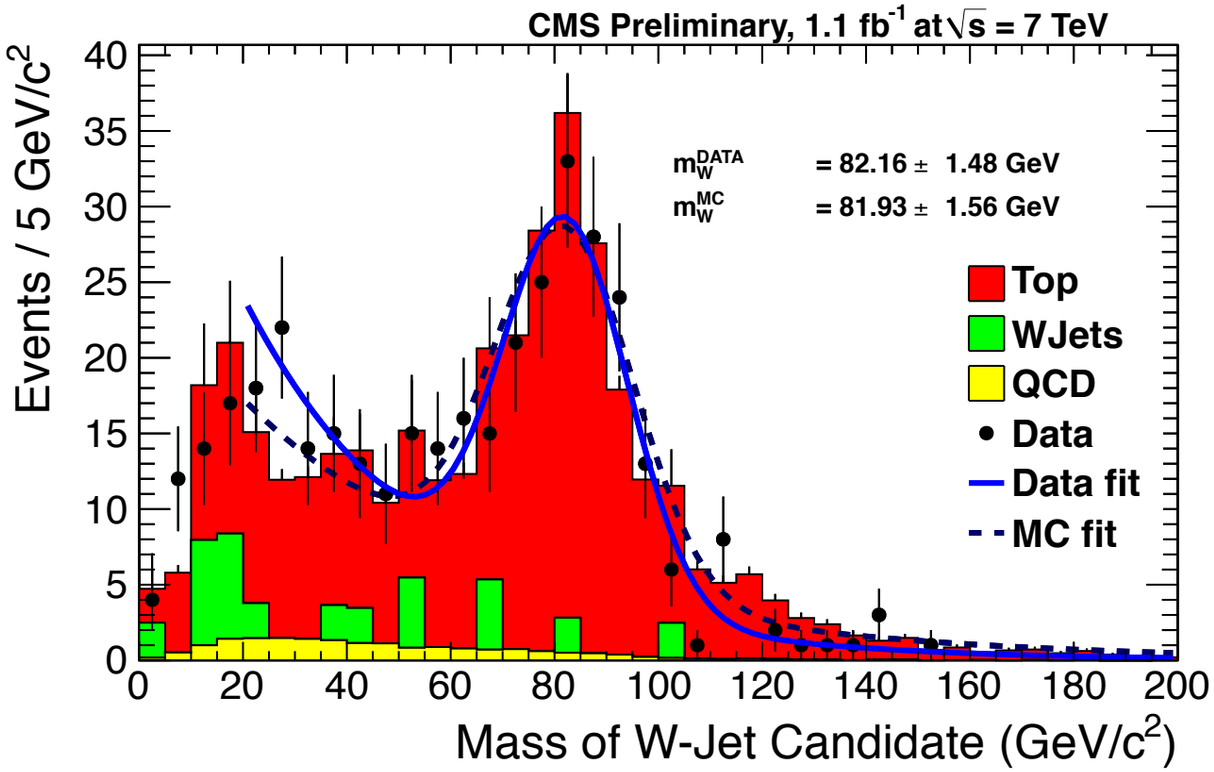
Christopher Vermilion
Michigan State University

8/20/2012

Progress and challenges in jet substructure



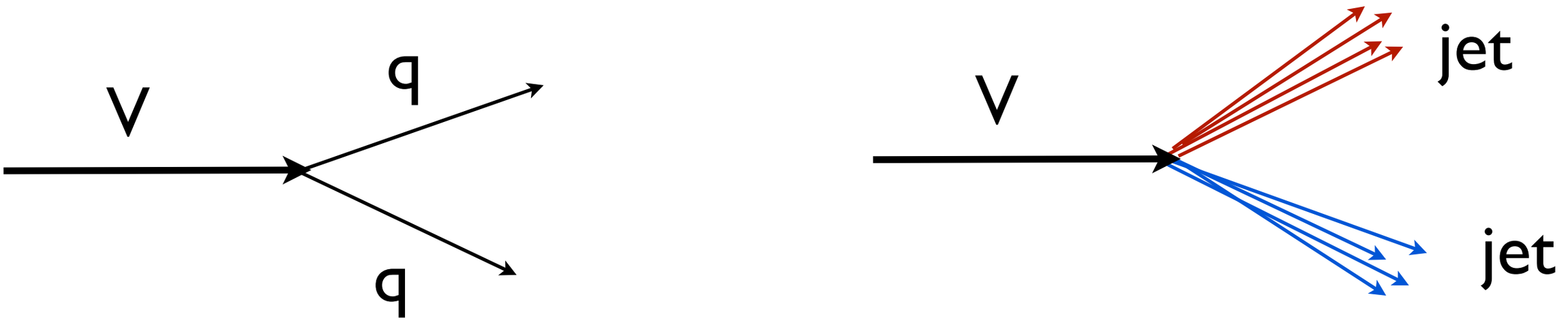
Or, How I Spent My Summer Vacation at BOOST 2012



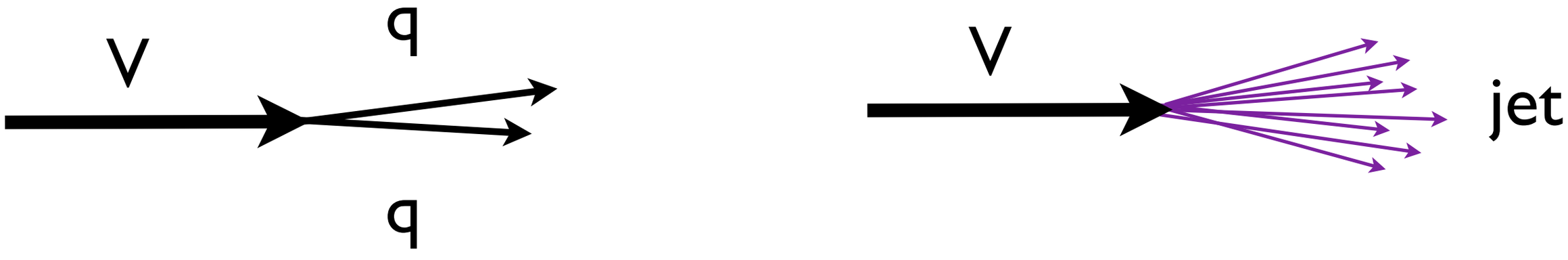
Christopher Vermilion
Michigan State University

8/20/2012

Low-pT boson decay



High-pT boson decay



History: five years of hard work on problem of boosted heavy objects

Lead to BOOST series of workshops, starting in 2009

arXiv:1012.5412

Boosted objects: a probe of beyond the standard model physics ⋆

A. Abdesselam[1], A. Belyaev[2,3], E. Bergeaas Kuutmann[4], U. Bitenc[5], G. Brooijmans[6], J. Butterworth[7], P. Bruckman de Renstrom[8], D. Buarque Franzosi[9], R. Buckingham[1], B. Chapleau[10], M. Dasgupta[11], A. Davison[7], J. Dolen[12], S. Ellis[13], F. Fassi[14], J. Ferrando[1], M.T. Frandsen[15], J. Frost[16], T. Gadfort[17], N. Glover[18], A. Haas[19], E. Halkiadakis[20], K. Hamilton[21], C. Hays[1], C. Hill[22], J. Jackson[3], C. Issever[1], M. Karagoz[1], A. Katz[23], L. Kreczko[24], D. Krohn[25], A. Lewis[1], S. Livermore[1], P. Loch[26], P. Maksimovic[27], J. March-Russell[15], A. Martin[28], N. McCubbin[3], D. Newbold[24], J. Ott[29], G. Perez[30], A. Policchio[13], S. Rappoccio[27], A.R. Raklev[31], P. Richardson[18], G.P. Salam[25,32,33], F. Sannino[34], J. Santiago[35], A. Schwartzman[19], C. Shepherd-Themistocleous[3], P. Sinervo[36], J. Sjoelin[37,38], M. Son[39], M. Spannowsky[40], E. Strauss[19], M. Takeuchi[41], J. Tseng[1], B. Tweedie[27,42], C. Vermilion[43], J. Voigt[29], M. Vos[44], J. Wacker[19], J. Wagner-Kuhr[29], and M.G. Wilson[19]

M. Karagoz, G. P. Salam, M. Spannowsky, M. Vos (editors)

arXiv:1201.0008

Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*

A Altheimer[1], S Arora[2], L Asquith[3], G Brooijmans[1], J Butterworth[4], M Campanelli[4], B Chapleau[5], A E Cholakian[1,6], J P Chou[7], M Dasgupta[8], A Davison[4], J Dolen[9], S D Ellis[10], R Essig[11,12,13], J J Fan[14], R Field[15], A Fregoso[8], J Gallicchio[6], Y Gershtein[2], A Gomes[16], A Haas[11], E Halkiadakis[2], V Halyo[14], S Hoeche[11], A Hook[11,17], A Hornig[10], P Huang[18], E Izaguirre[11,17], M Jankowiak[11,17], G Kribs[19,20], D Krohn[6], A J Larkoski[11], A Lath[2], C Lee[21], S J Lee[22], P Loch[23], P Maksimovic[24], M Martinez[25], D W Miller[11,17], T Plehn[26], K Prokofiev[27], R Rahmat[28], S Rappoccio[24], A Safonov[29], G P Salam[30,14,31], S Schumann[32], M D Schwartz[6], A Schwartzman[11], M Seymour[8], J Shao[33], P Sinervo[34], M Son[35], D E Soper[20], M Spannowsky[36], I W Stewart[21], M Strassler[2], E Strauss[11], M Takeuchi[26], J Thaler[21], S Thomas[2], B Tweedie[37], R Vasquez Sierra[9], C K Vermilion[38,39], M Villaplana[40], M Vos[40], J Wacker[11], D Walker[6], J R Walsh[38,41], L-T Wang[14], S Wilbur[42] and W Zhu[14]

L. Asquith, S. Rappoccio, CV (editors)

The story of the film so far (more details in the backup slides)

Dozens of jet substructure techniques proposed (studied in MC!)

Subjet kinematics, jet shapes, grooming, hybrid taggers

Slow but important progress in *calculations* relevant to these techniques

Mostly focused on improving jet cross sections, jet mass

Data: wait and see (I'll come back to this)

Brings us to Boost 2012

http://ific.uv.es/boost2012/

Boost 2012: A Feast of Data

Recent theory progress: new phenomenology, new calculations

Many slides taken from Gavin Salam's summary talk at BOOST

Qjets in practice

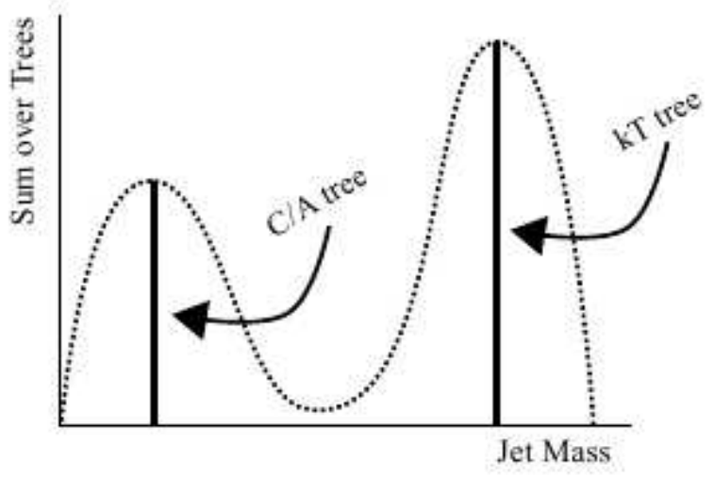
Fastjet Plugin: http://jets.physics.harvard.edu/Qjets

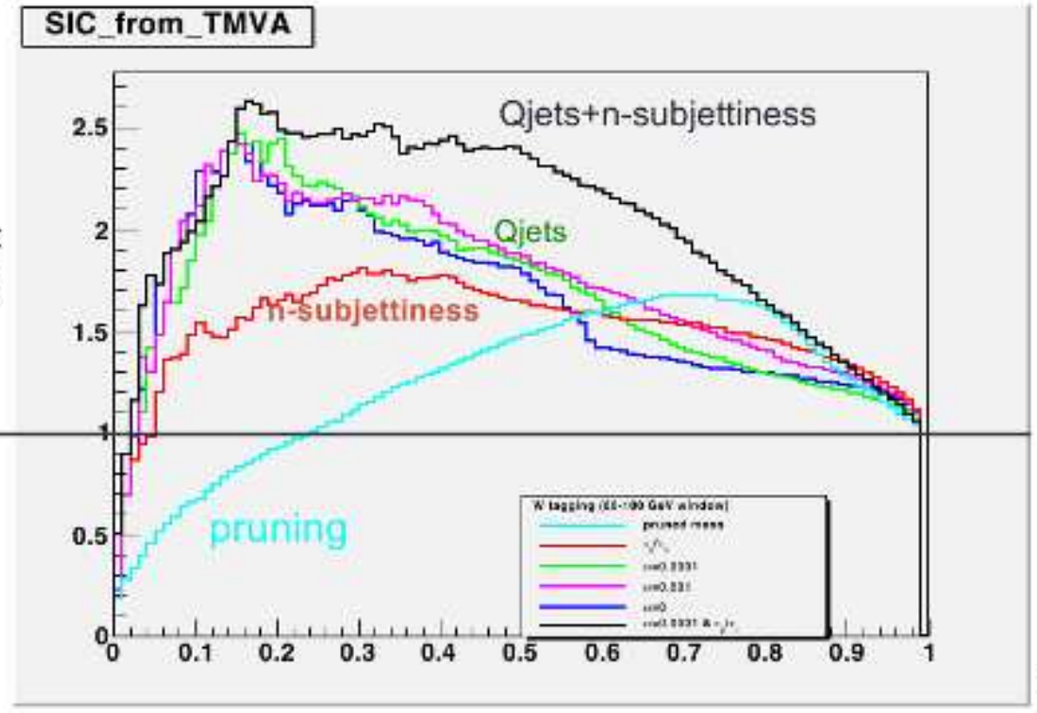
Hornig

- too many trees to consider all
- can sample kT like (or CA like) randomly:
  - at each stage, merge pair w/ prob.

$$\omega_{ij}^{(\alpha)} \equiv \exp\left\{-\alpha \frac{(d_{ij} - d^{\min})}{d^{\min}}\right\}$$

  - this gives a tree,





[Optimal weighting turns out to be flat??]

Larkoski

- Angular Correlation Function (ACF)

$$\mathcal{G}(R) \equiv \sum_{i \neq j} p_{\perp i} p_{\perp j} \Delta R_{ij}^2 \Theta[R - \Delta R_{ij}]$$

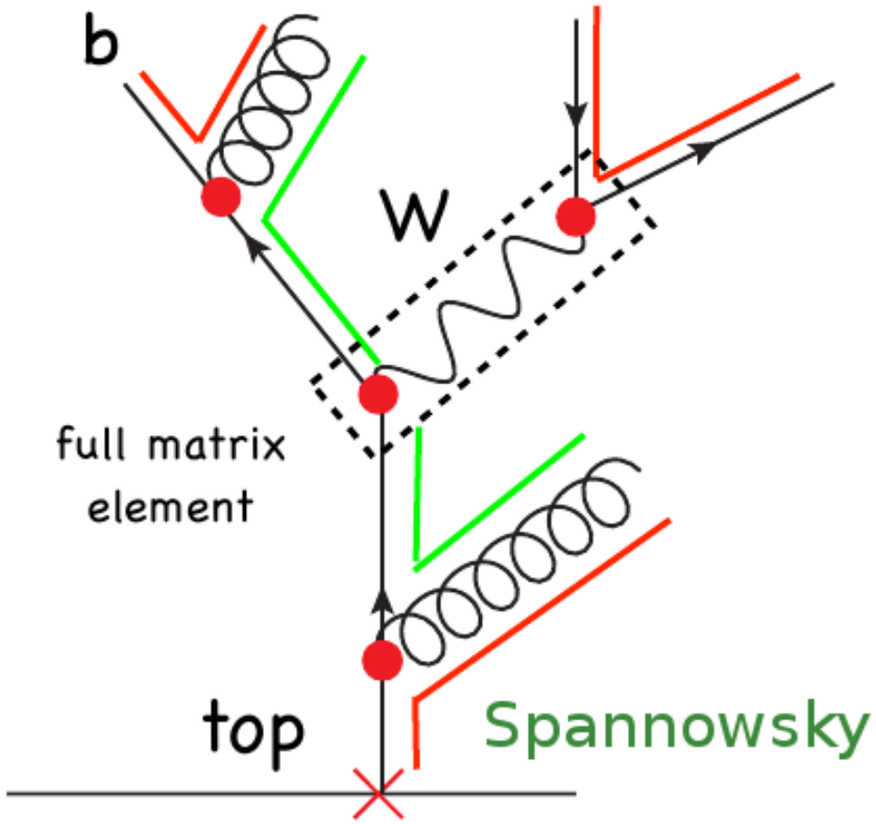
- IRC safe
- If jet is scale invariant, $\mathcal{G}(R) \sim R^D$
- D = IRC-safe definition of the scaling/ correlation dimension of the jet

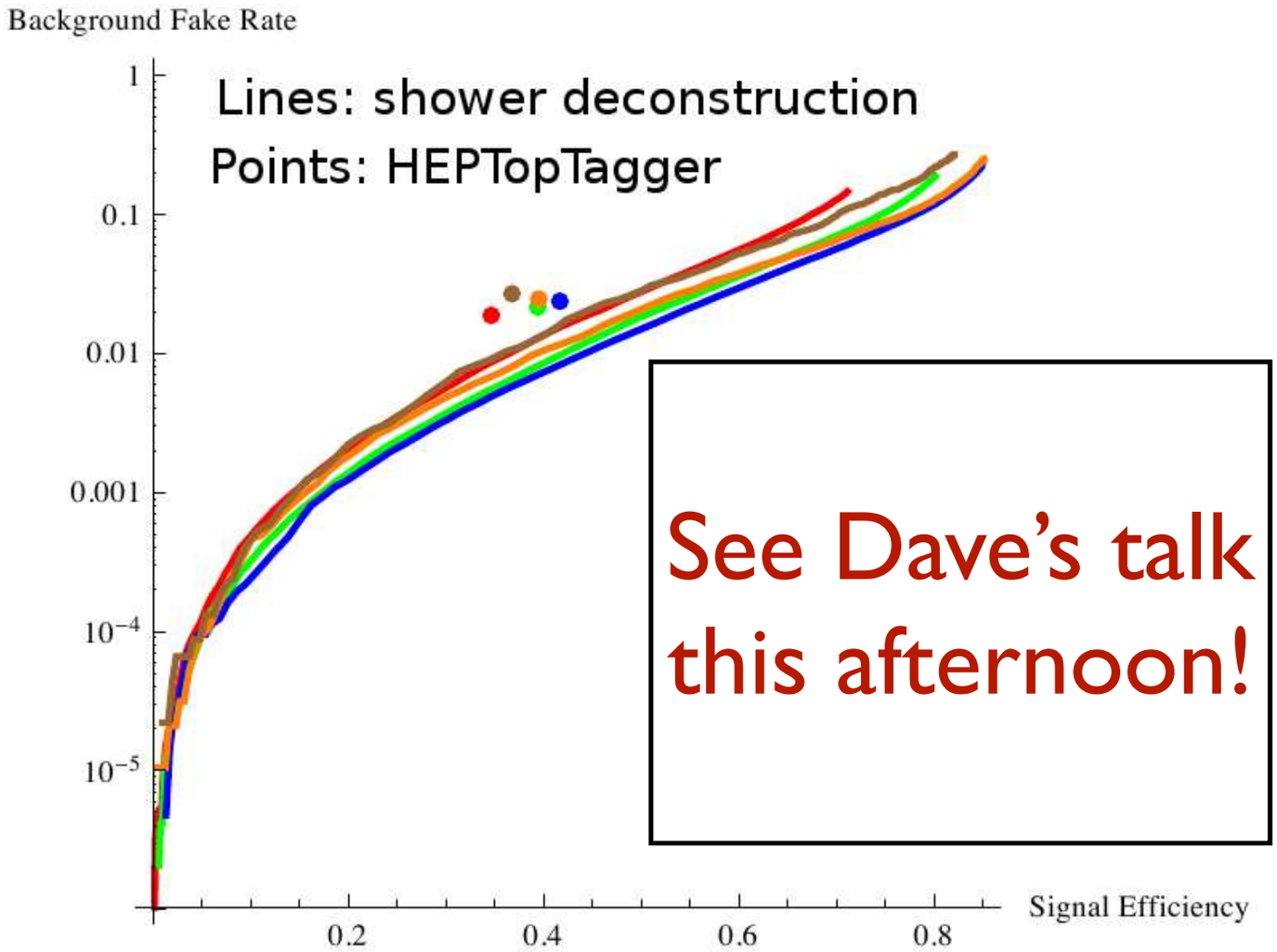
ACF (aka energy-energy-correlation moments) is an observable a bit like angularities, but with very special resummation properties — keep an eye on it for the future

studied for e^+e^- also in Banfi, GPS & Zanderighi '04

Shower deconstruction: matrix element method "on steroids"

Need to convert the shower history into analytic expression





See Dave's talk this afternoon!

Final and distant goal:

Shower deconstruction is related to **Matrix Element Method** like

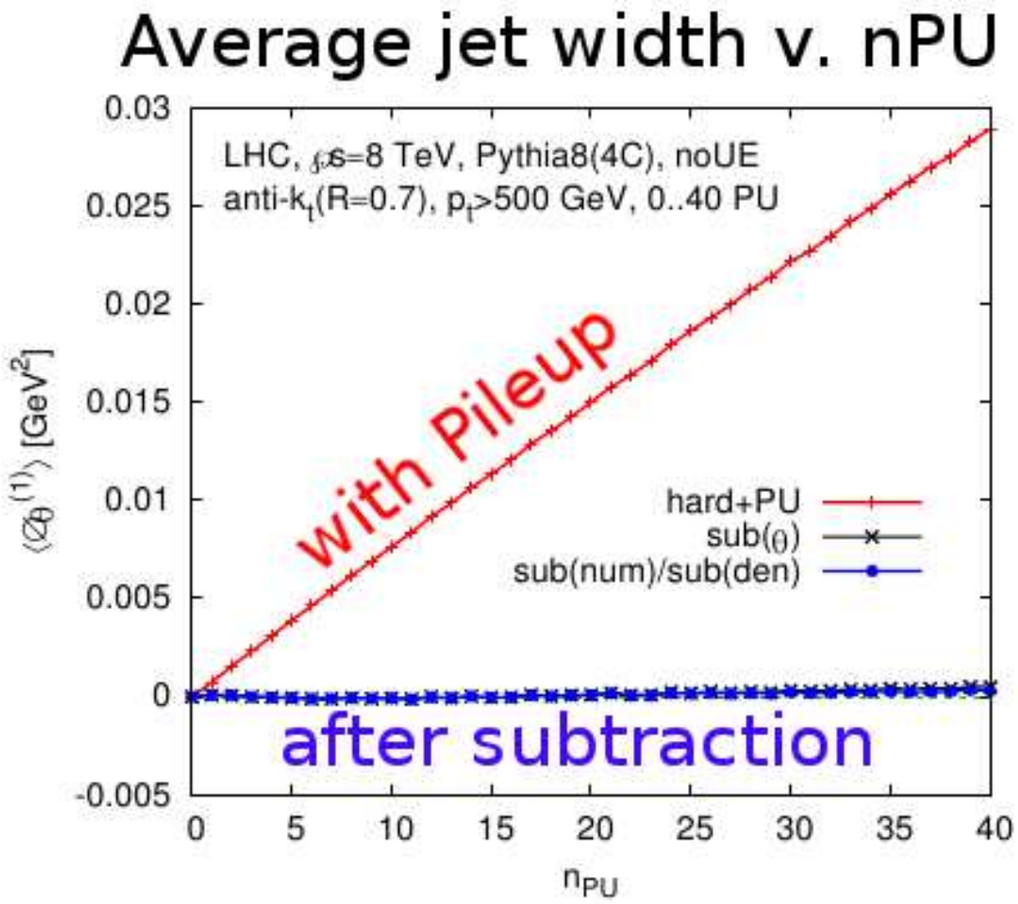
Full Event Generators to **Parton Level simulation**

Subtraction pileup from shapes

pile-up Soyez ghosts

$$\propto \rho \times$$

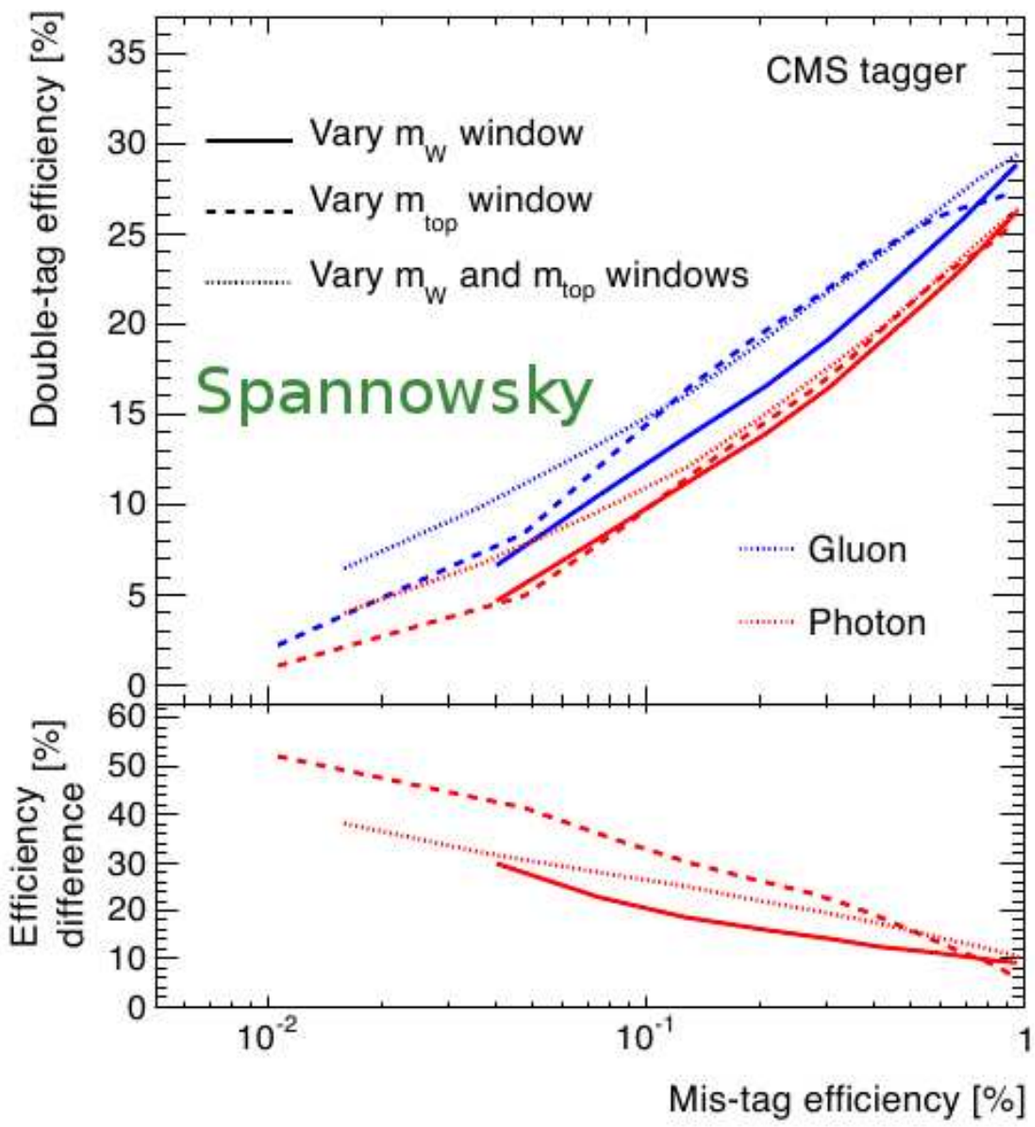
$$f_{\text{sub}}(\text{jet}) = f_{\text{full}}(\text{jet}) - \rho\, a_{\text{ghost}}\, \partial_{\text{ghostscale}} f_{\text{full}}(\text{jet}) + \frac{1}{2}(\rho\, a_{\text{ghost}})^2\, \partial^2_{\text{ghostscale}} f_{\text{full}}(\text{jet}) + \dots$$

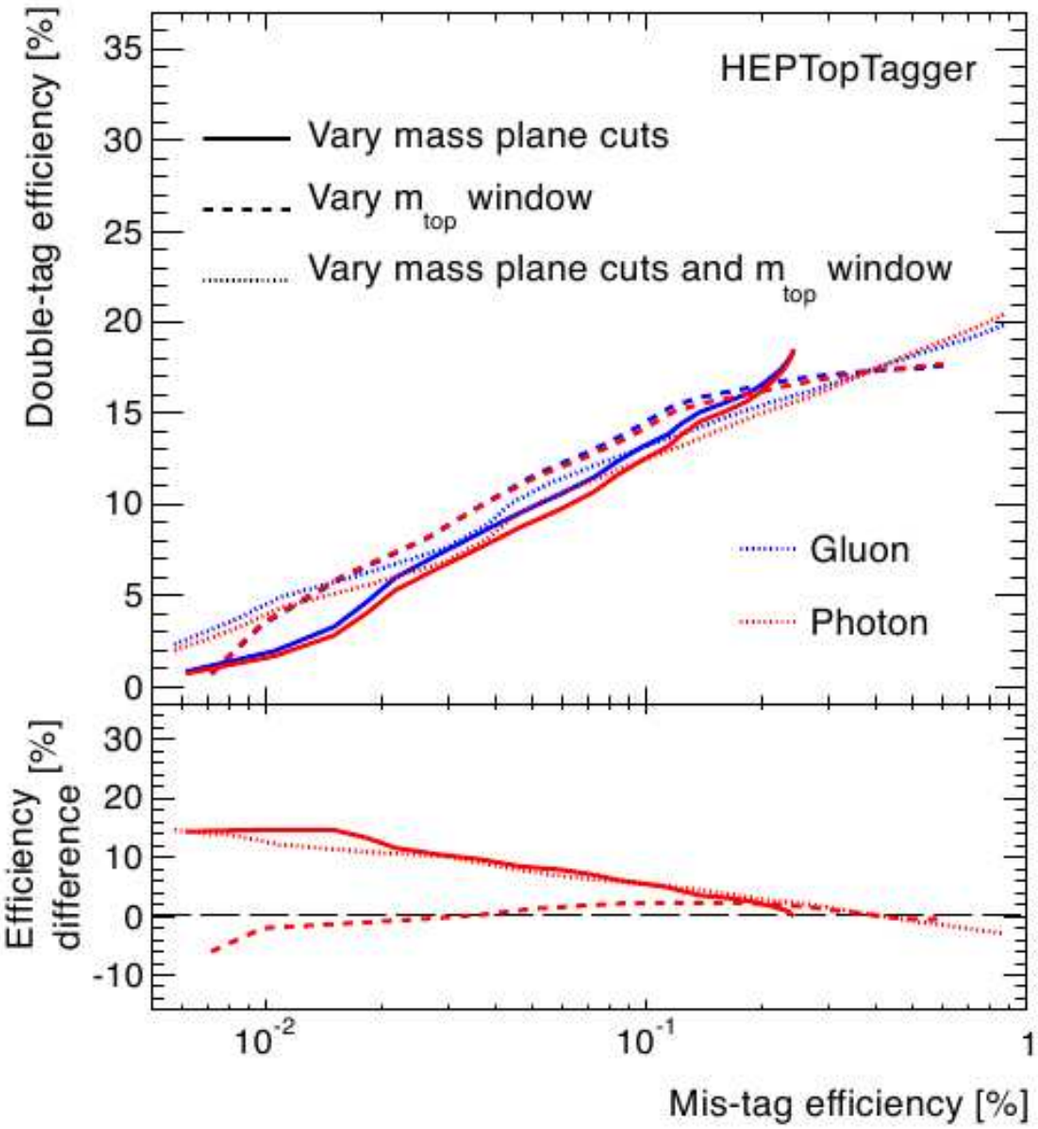


Application to τ_3/τ_2 based top tagging:

without pileup:	35% v. 2%
with 30 pileup:	20% v. 0.5%
after subtraction:	35% v. 2%

Robustness to colour-flow? (Octet v. singlet resonances)





Conclusion: top-tagger with in-built grooming $\ll$ affected

NB: sensitivity appears for aggressive working points (very low mistag), while experiments work with higher mistag rates

More exciting (to me!):

Progress in substructure calculations

First: jet masses

Gavin on why jet mass is difficult:

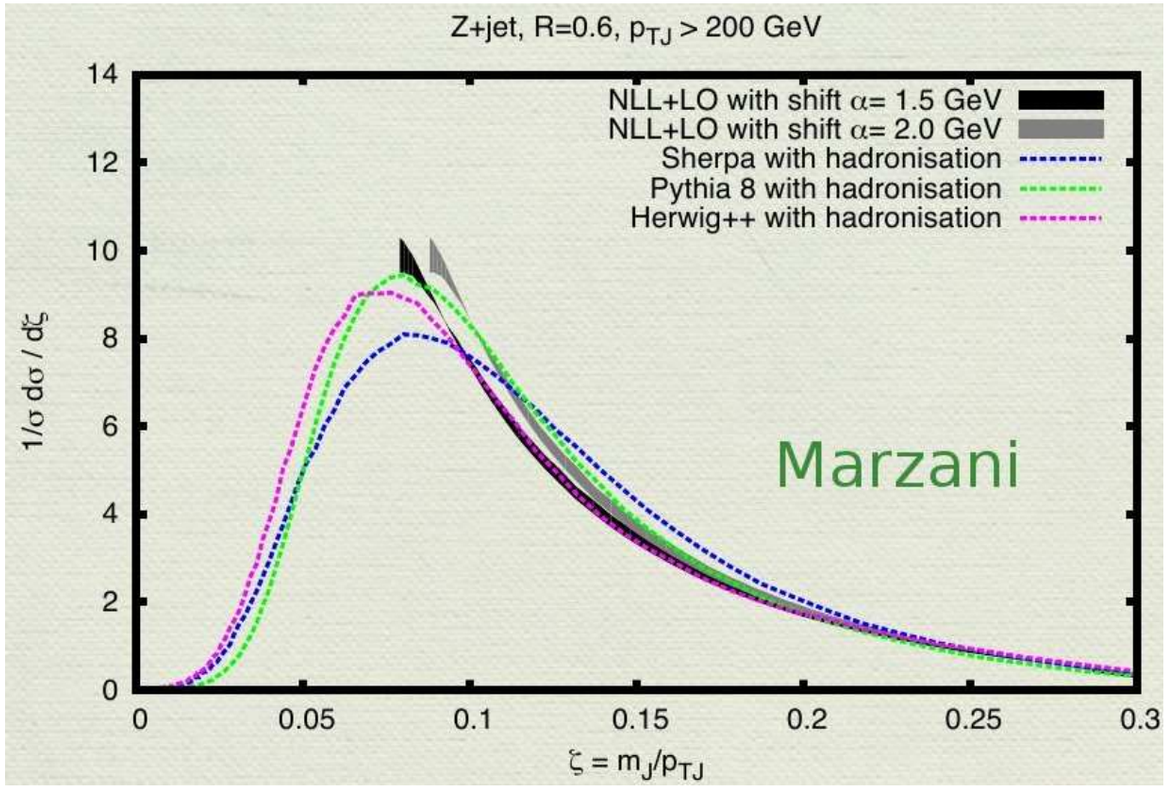
$$m^2 \simeq p_{t1} p_{t2} \Delta R_{12}^2$$

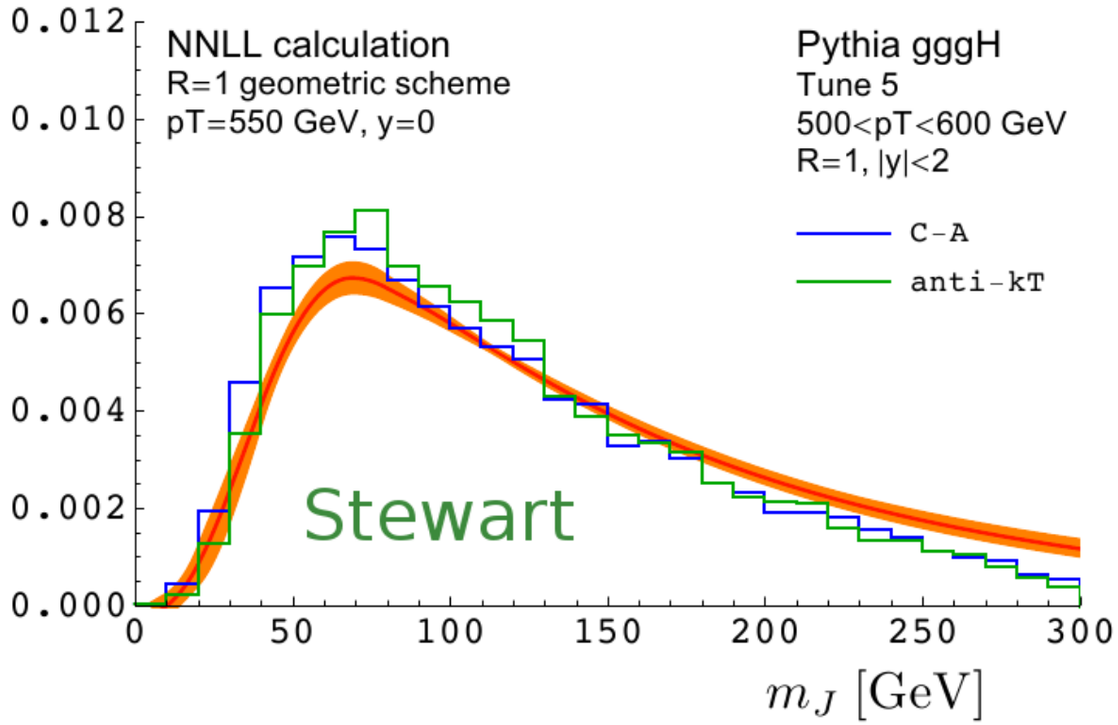
for parton with $p_t = 300$ GeV, jet $R = 0.7$,
30 GeV of jet mass comes from 7 GeV emission

for parton with $p_t = 2$ TeV, jet $R = 0.7$,
30 GeV of jet mass comes from 1 GeV emission

This sensitivity to low momentum scales is what makes masses difficult

Jet masses from two groups





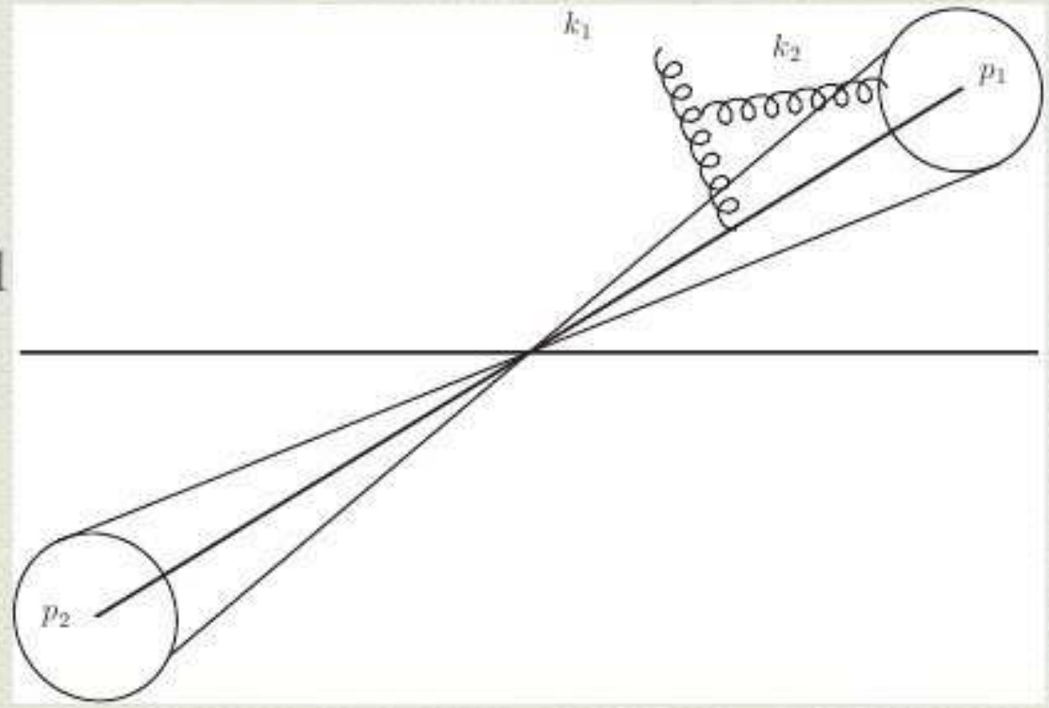
Resummed jet masses compared to Monte Carlo showers

These results are a significant theory development of the workshop

but, be aware of the fine print

Non-global logarithms

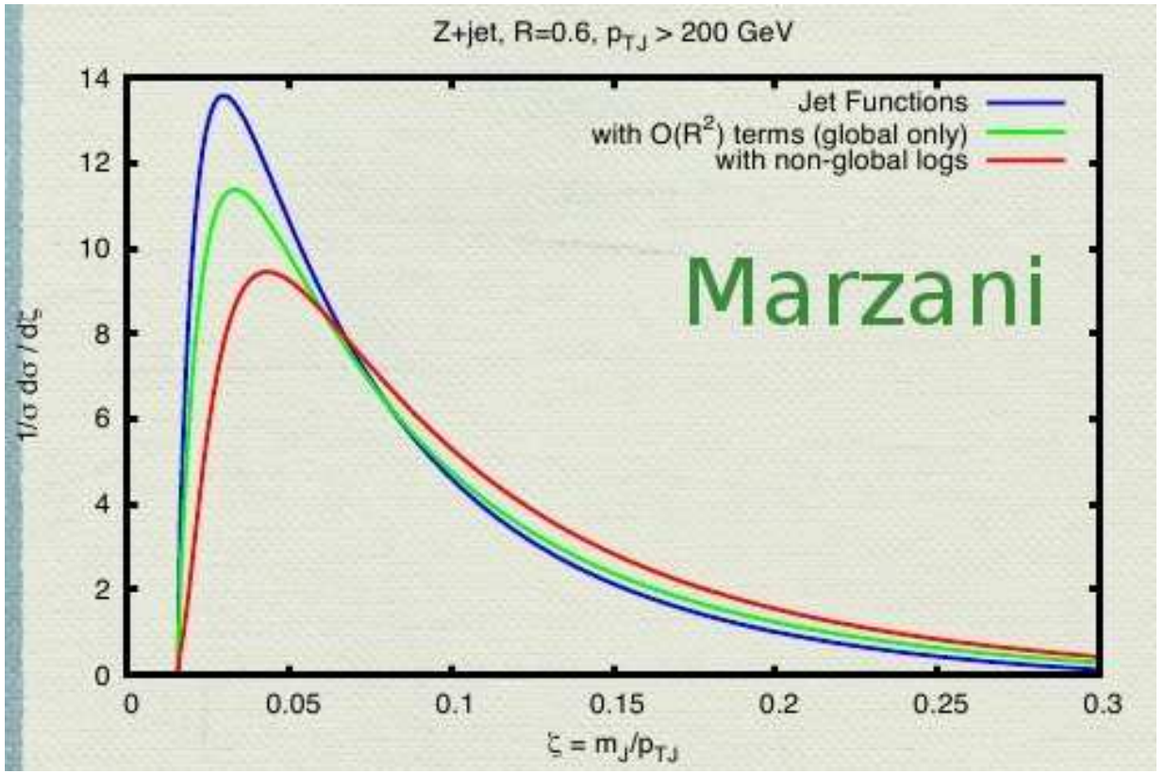
- BUT, even if we use anti-k_t, exponentiation of the independent emission is not the whole story
- The jet-mass is a non-global observable: it receives single log corrections from correlated emission
- This is a $C_F C_A$ term and it's missed by single gluon exponentiation
- In principle we need to consider any number of gluons outside the jet
- Colour structure becomes intractable, so the resummation is performed in the large N_c limit

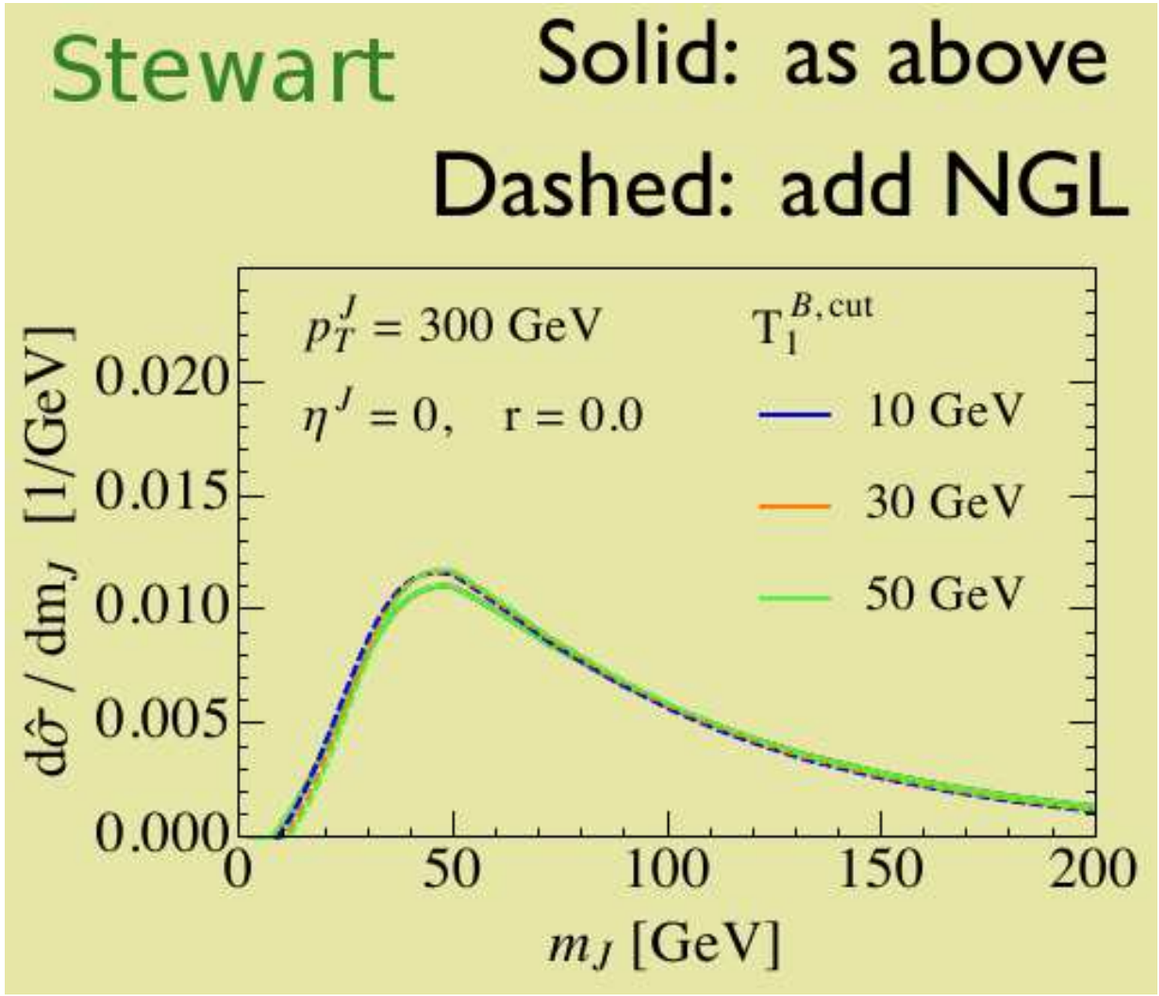


Dasgupta and Salam (2001)
Banfi, Marchesini and Smye (2002)

Marzani

Do (NLL!) non-global logs matter when you do NNLL?

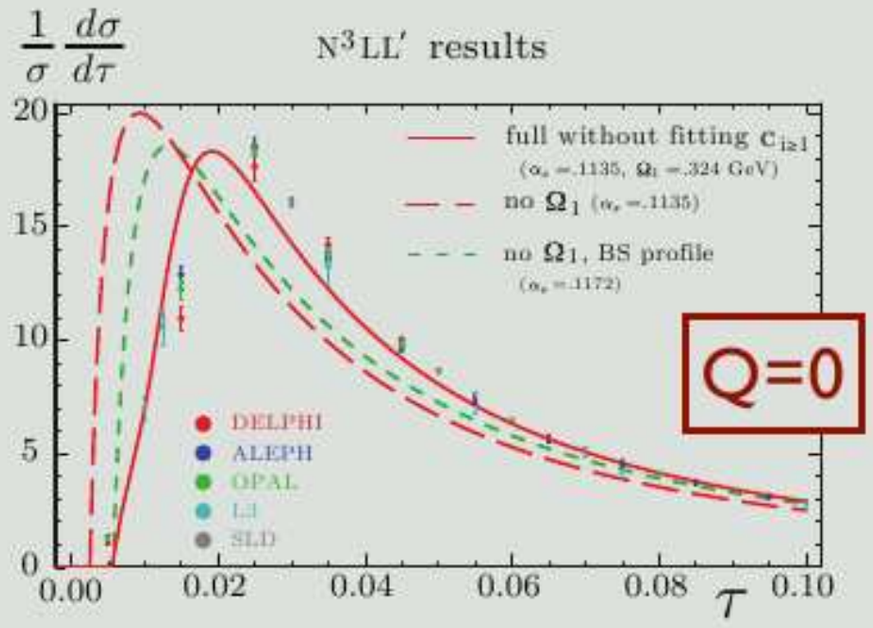


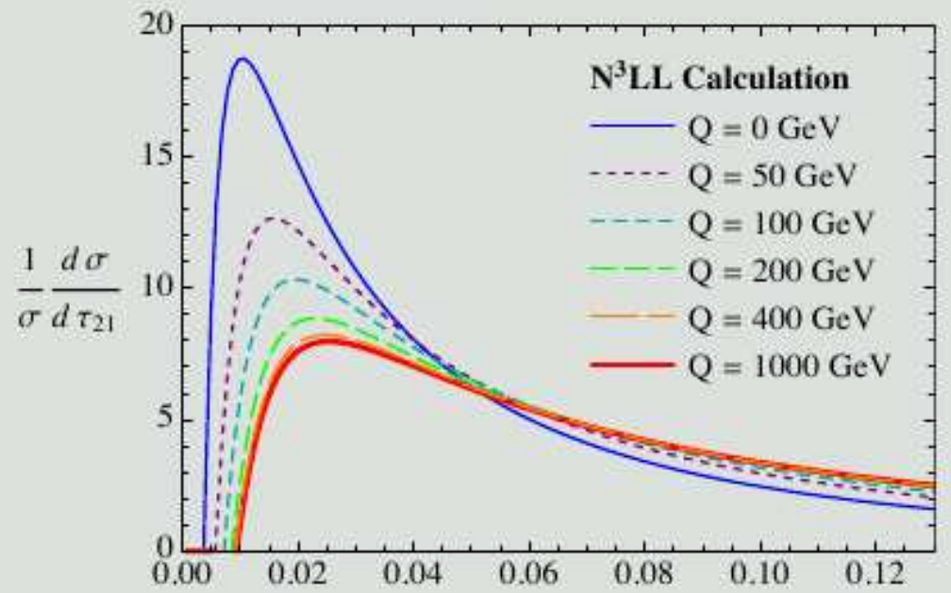


[Opposing conclusions! Still to be shaken out...]

Thaler:

Recycling Thrust Results





[Abbate,Fickinger,Hoang,Mateu,Stewart;
see also Becher, Schwartz]

Hard, Jet, and Soft Functions to $O(\alpha_s^2)$

Resummation to N^{3}LL

Leading Shift from Non-Perturbative Power Correction
[See Vicent Mateu's Talk]

NB: one still needs to select the jet mass (which has NLL ISR distortion)

Meanwhile on the experimental side...

Google

meanwhile

christopher.vermilion@gmail.com 0 + Share

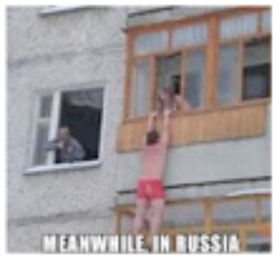
Search

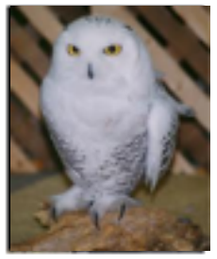
About 249,000,000 results (0.43 seconds)

SafeSearch

Web
Images
Videos
News
Shopping
More

Related searches: meanwhile book





Any time
Past 24 hours
Past week
Custom range...

All results
By subject
Personal

Any size
Large
Medium
Icon
Larger than

MEANWHILE, IN RUSSIA

MEANWHILE IN
ALABAMA
MEANWHILEIN.ORG

MEANWHILE
IN MICHIGAN
quickmeme.com

MEANWHILE

In Canada.....

Experimental progress

(As measured by data talks at Boost)

2009: 1

2010: 5

2011: 14

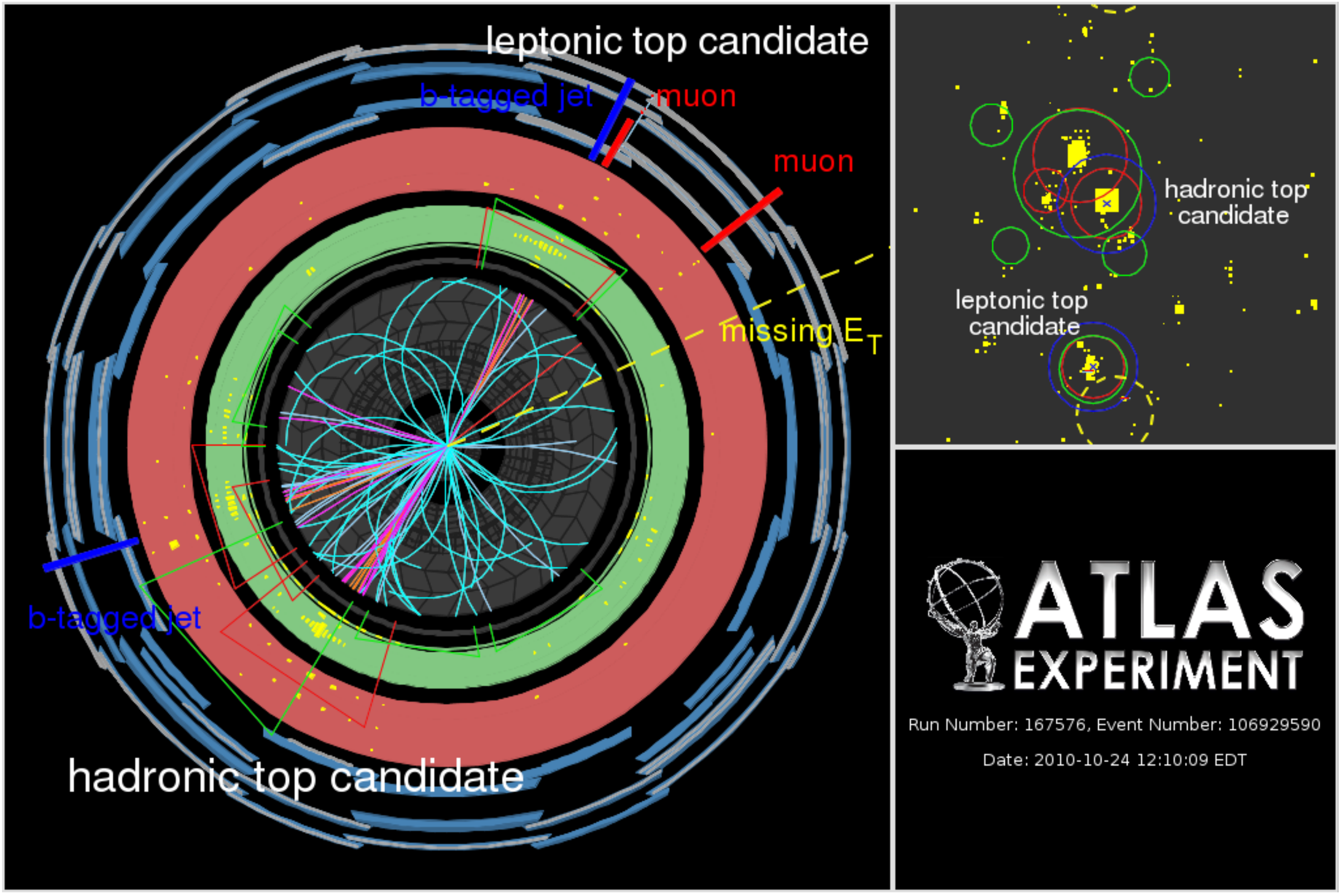
2012: 16

Experimental progress

(As measured by data talks at Boost)

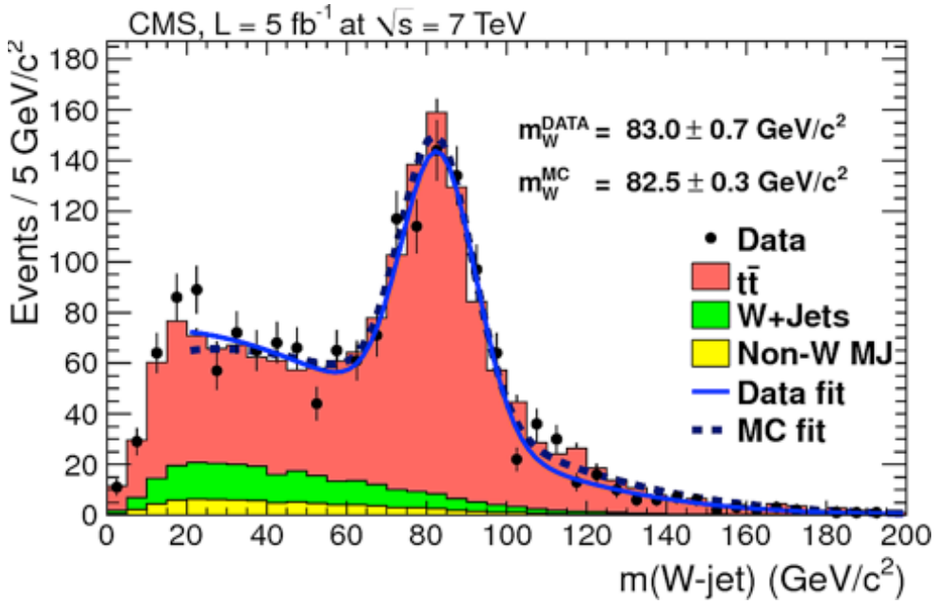
2009: 1	
2010: 5	Calibration
2011: 14	Multijets, lepton jets, background studies
2012: 16	Searches, SM heavy particles

This was the big result at Boost 2011!



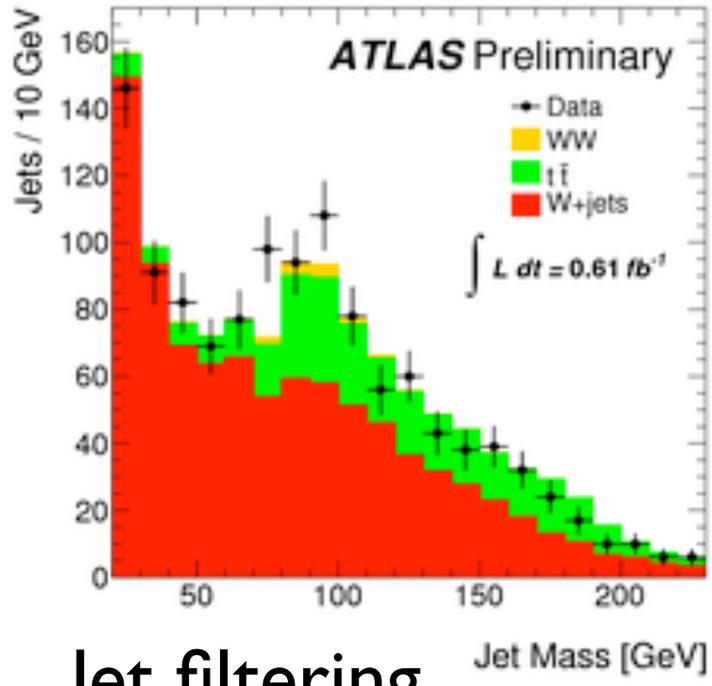
Boost 2012: Tops, Ws and searches (oh my!)

CMS

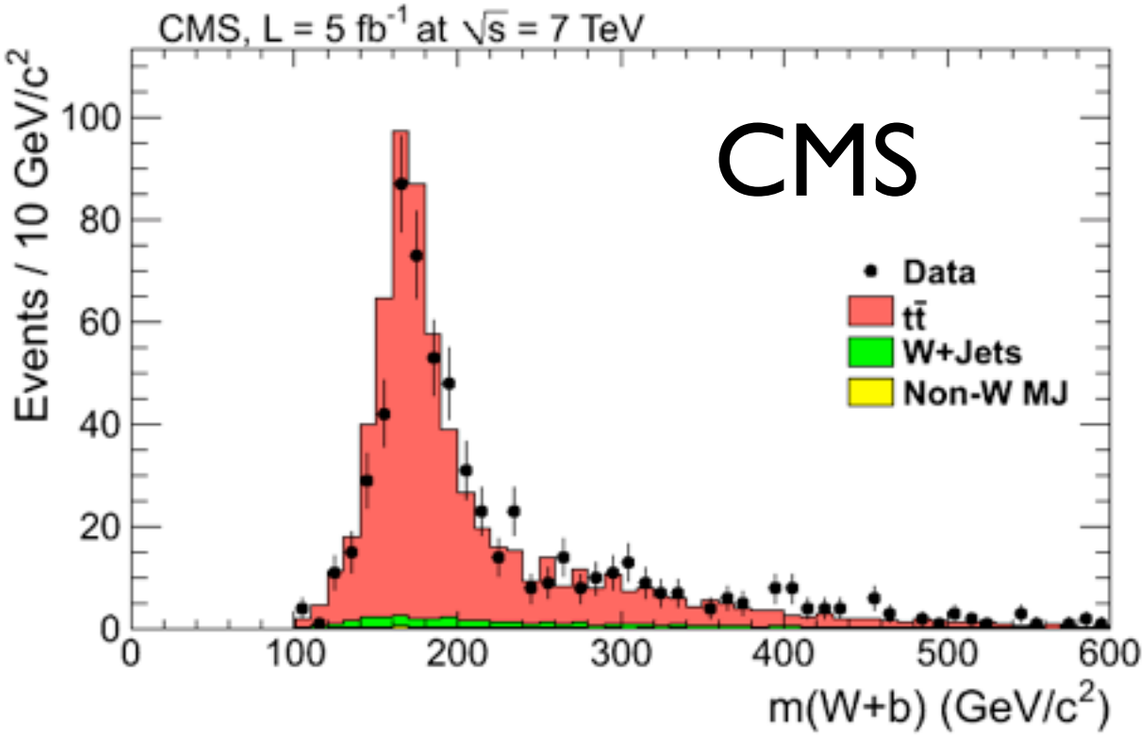


Jet pruning

ATLAS



Jet filtering



CENSORED

See David's talk, next.

2011 theory wish list

Jet mass (in V+j, jj, multijet; multi-differential useful)

Jet shapes: N-subjettiness (N=1,2,3; beta=1,2), planar flow,

subjet/track multiplicities

Groomed observables (all of the above)

Event shapes: 0/1-jettiness, Y_23

Nearly all of these have been measured!

Outlook

~~Outlook~~ To-Do List

It's tough to make predictions,
especially about the future.

Yogi Berra

HUGE experimental output on substructure in the last few years.

Really impressive speed of experimental implementation of new ideas: BDRS, pruning/trimming/N-subjettiness...

Time for the flow to go the other way!
Theorists need to grapple with explosion of data. What have we learned?

How to bring theory/
experiment together?

shared tools ⊗ Monte Carlo = Profit

Tools

FastJet

http://fastjet.fr

Standard library of jet algorithms

Now includes extensive set of jet measurements, manipulations

Coming soon: fastjet/contrib: project to collate user-submitted tools

SpartyJet

http://projects.hepforge.org/spartyjet

Extends FastJet with flexible I/O, Python control, GUI browser

Easy to build up complicated analyses, add in pile-up, etc.

Rivet

http://projects.hepforge.org/rivet

Really nice package for standardized experimental analysis write-up

Major caveat: designed to run on unfolded data or MC

All three of these being used in Boost 2012 follow-up

Calculation wish-list

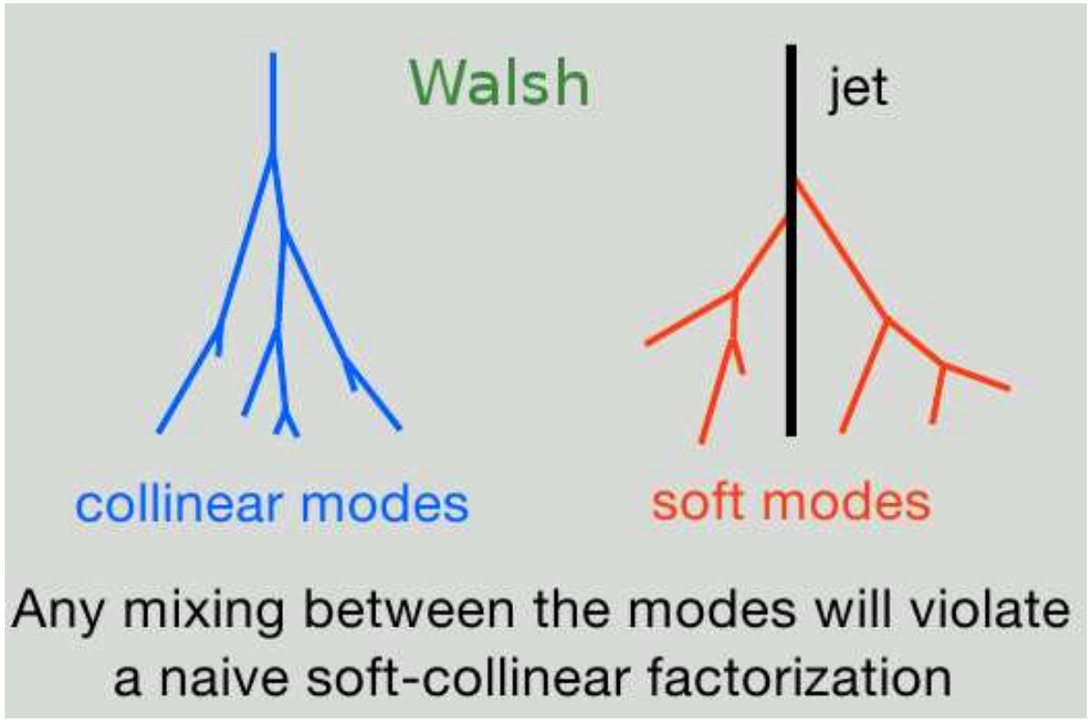
Great progress so far, esp. on jet mass and shapes.

Can we resolve question of "non-global logs"?

Can we synthesize our understanding of various substructure tools?

Groomed jet cross sections?

What about calculating masses with grooming?



arXiv.org > hep-ph > arXiv:1002.4557

Non-Global Logarithms in Filtered Jet Algorithms

Mathieu Rubin

(Submitted on 24 Feb 2010 (v1), last revised 10 May 2010 (this version, v2))

We analytically and numerically study the effect of perturbative gluons emission on the "Filtering analysis", which is part of a subjet analysis procedure proposed two years ago to possibly identify a low-mass Higgs boson decaying into b\bar{b} at the LHC. This leads us to examine the non-global structure of the resulting perturbative series in the leading single-log large-N_c approximation, including all-orders numerical results, simple analytical approximations to them and comments on the structure of their series expansion. We then use these results to semi-analytically optimize the parameters of the Filtering analysis so as to suppress as much as possible the effect of underlying event and pile-up on the Higgs mass peak reconstruction while keeping the major part of the perturbative radiation from the b\bar{b} dipole.

My opinion: with all techniques we have at hand, I see no reason why we can't obtain reasonably precise predictions for groomed techniques

Experimental wish-list

Experimental wish-list

Keep up the good work!

Experimental wish-list

More apples-to-apples comparable measurements between ATLAS/CMS

Continuous theorist involvement from beginning

Techniques are phenomenological, best seen as iterative

Not all of param. space is calculationally sensible [Walsh, Zuberi 1110.5333]

Rest of 2011 report wish list: mostly QCD precision as opposed to searches [see Albert's quotation of Michelangelo this morning]

Parting thoughts

Extraordinary progress in last few years

Should be a really interesting year for substructure:

Precision QCD substructure calculations

Comprehensive understanding of substructure in MC

Substructure as a standard tool in the experimental tool kit

Parting thoughts

Extraordinary progress in last few years

Should be a really interesting year for substructure:

Precision QCD substructure calculations

Comprehensive understanding of substructure in MC

Substructure as a standard tool in the experimental tool kit

New physics with jet substructure?

Thank you!

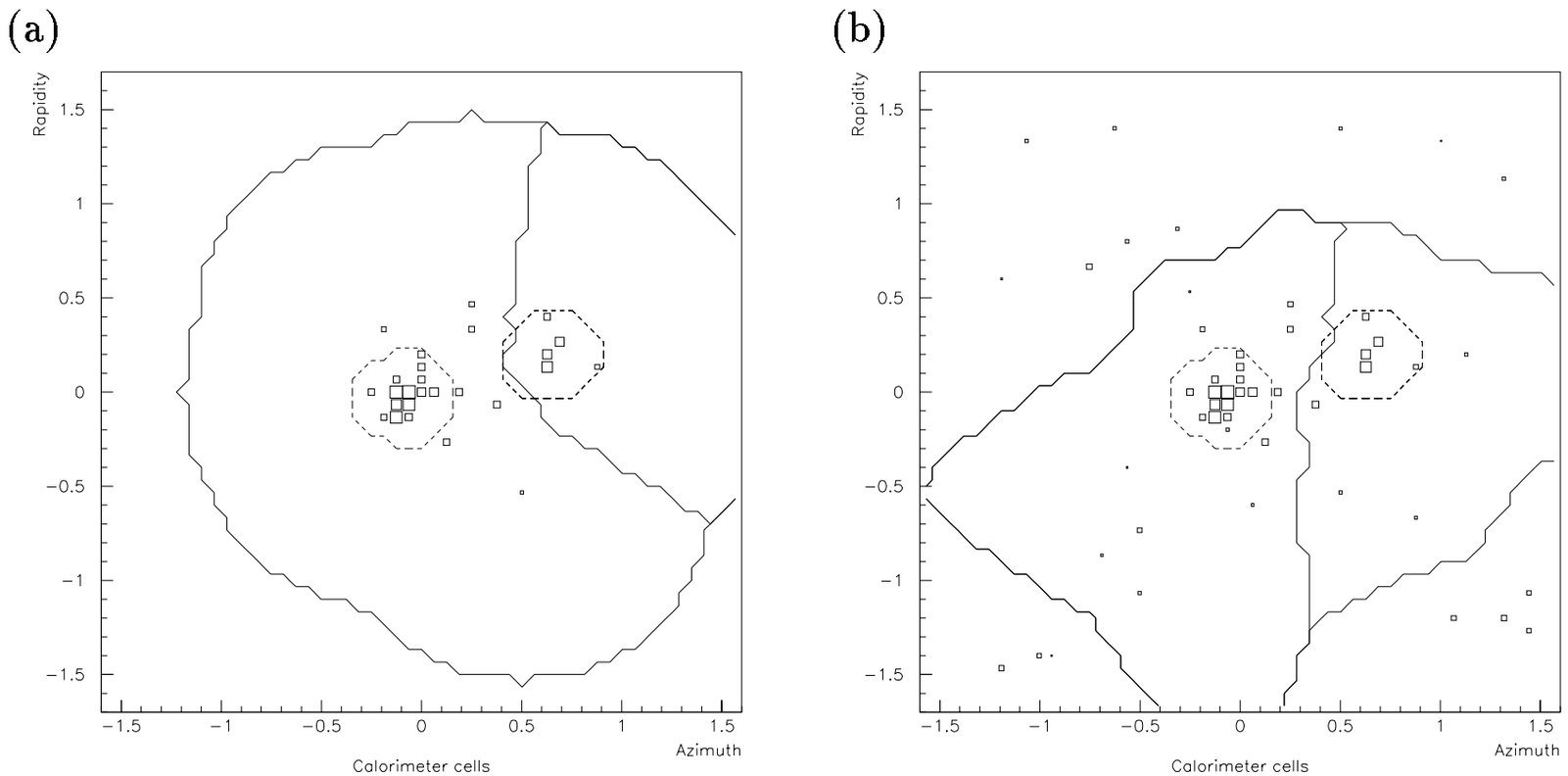


Figure 2: A hadronic W decay, as seen at calorimeter level, (a) without, and (b) with, particles from the underlying event. Box sizes are logarithmic in the cell energy, lines show the borders of the sub-jets for infinitely soft emission according to the cluster (solid) and cone (dashed) algorithms.

LU TP 93-8

Searches for New Particles Using Cone and Cluster Jet Algorithms: A Comparative Study

Michael H. Seymour
Department of Theoretical Physics, University of Lund,
Sölvegatan 14A, S-22362 Lund, Sweden

Mike Seymour

Z. Phys. C62 (1994) 127

Google

superconducting super collider

christopher.vermilion… 0 + Share

Get directions My places

superconducting super collider

Search the web for ***superconducting super collider*** Top web result shown

Superconducting Super Collider - Wikipedia, the free encyclopedia

The **Superconducting Super Collider** (**SSC**) (also nicknamed the Desertron) was a particle accelerator complex under construction in the vicinity of Waxahachie, ...

en.wikipedia.org

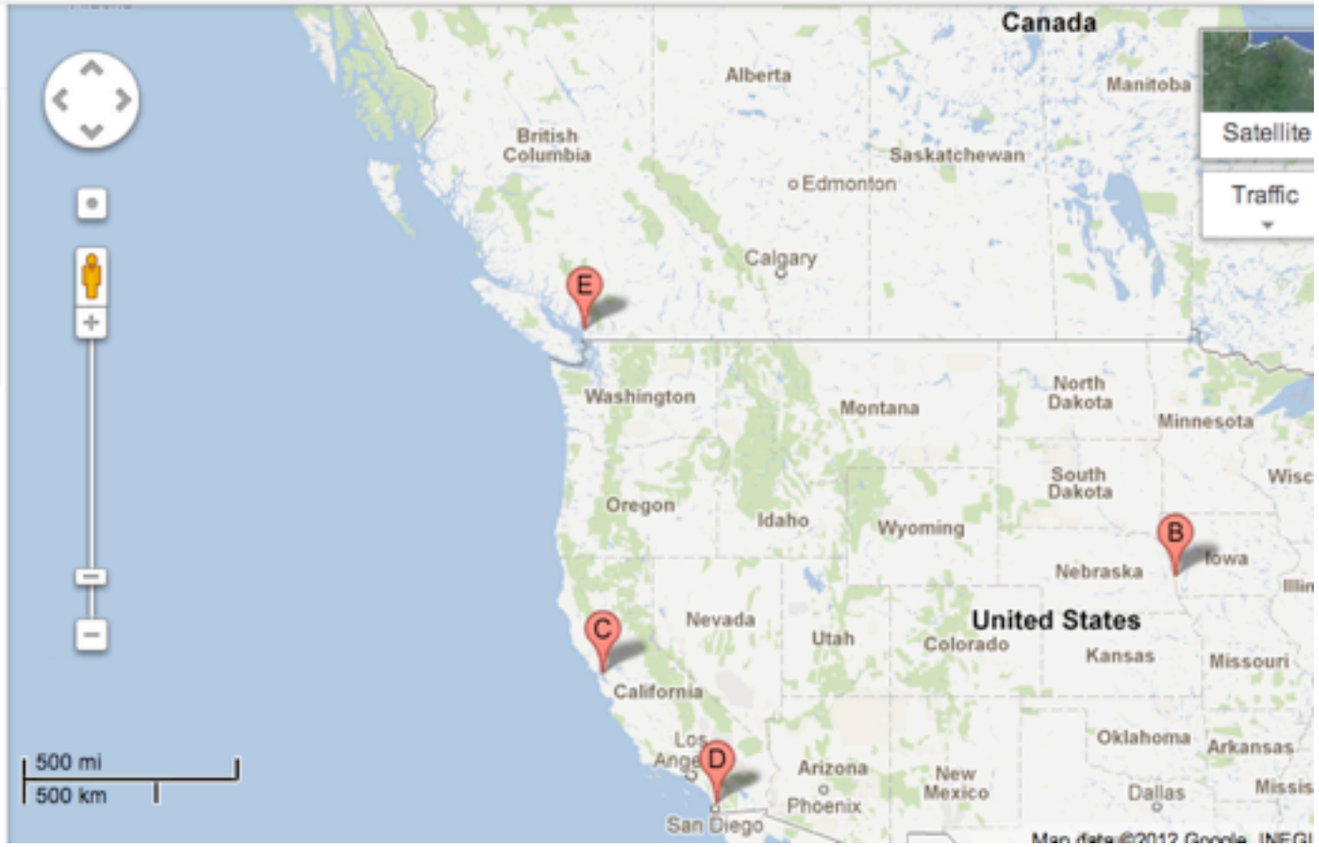
A **Ice Cave Technology**

5600 Lake Otis Parkway, Anchorage, AK

(907) 771-0543 · icecavetechnology.com

"**Super Conducting Super Collider**. PowerPoint presentation. Television commercial. Computer kiosk. Data visualization and animation. Web sites and marketing" - icecavetechnology.com

B **Cryo Nebraska Inc**



Butterworth, Davison, Rubin, Salam (0802.2470)

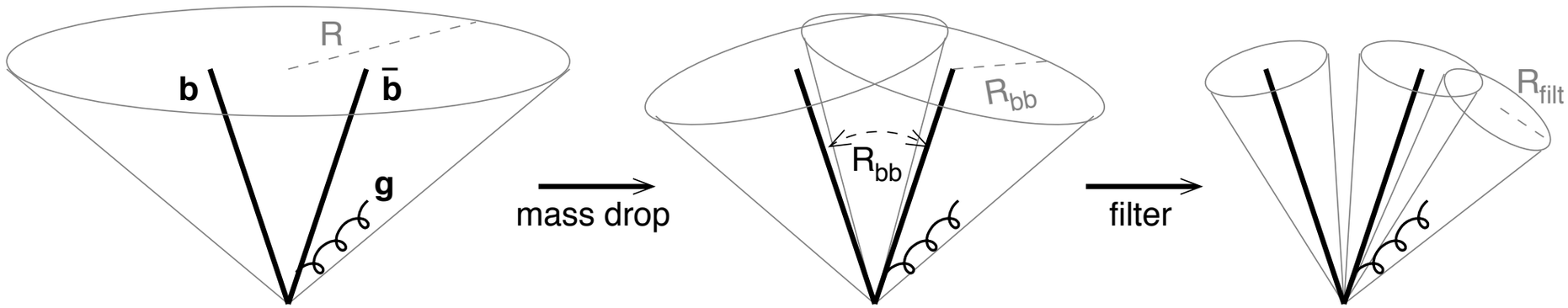


FIG. 1: The three stages of our jet analysis: starting from a hard massive jet on angular scale R, one identifies the Higgs neighbourhood within it by undoing the clustering (effectively shrinking the jet radius) until the jet splits into two subjets each with a significantly lower mass; within this region one then further reduces the radius to $R_{\rm filt}$ and takes the three hardest subjets, so as to filter away UE contamination while retaining hard perturbative radiation from the Higgs decay products.

Two steps:

Mass drop identifies hard splitting within clustering history

Filtering "zooms" in and only keeps three hardest subjets

[Johns Hopkins] Top Tagging (Kaplan, Rehermann, Schwartz, Tweedie; 0806.0848)

3-body kinematic variables (Thaler, Wang; 0806.0023)

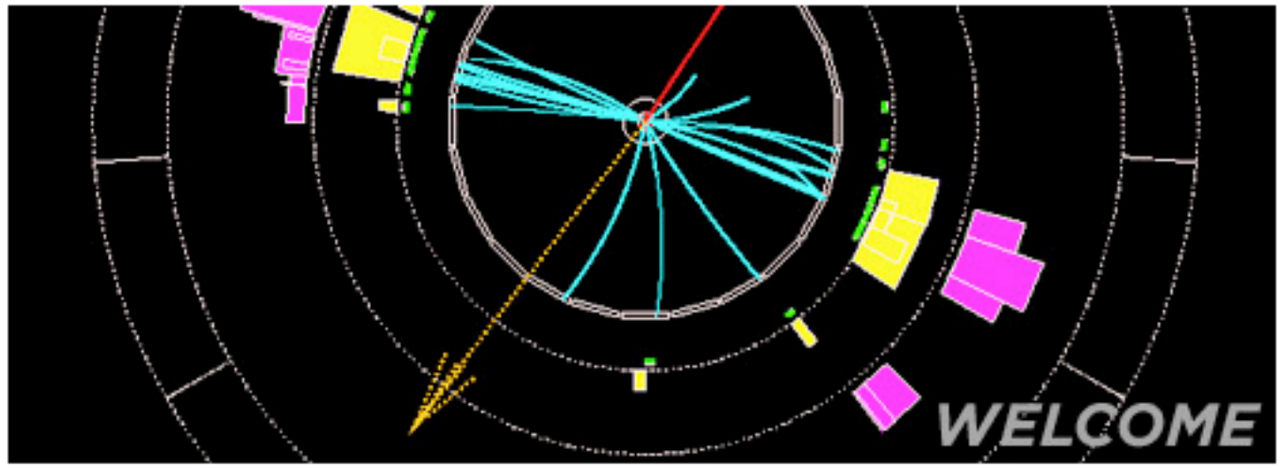
New jet shapes (Almeida, et al.; 0807.0234, 0810.0934)

Pruning (Ellis, Vermilion, Walsh; 0903.5081)

BOOST 2009

GIVING NEW PHYSICS A BOOST SLAC NATIONA

- Home
- Registration
- Participant List
- Agenda
- Accommodations
- General Information
- Travel and Directions
- Visa Information
- Social Event
- Contact



Giving New Physics a Boost

Thursday and Friday, July 9-10, 2009 from 8:00 am to 5:00 pm.
Kavli Auditorium
SLAC National Accelerator Laboratory
Menlo Park, California

Check out the indico page for a fascinating look at boosted physics in 2009...

Jet Trimming (Krohn, Thaler, Wang; 1912.1342)

CMS Top Tagging (0909.4894)

More jet shapes (Chekanov, Proudfoot, Levy, Yoshida; 1002.3982, 1009.2749)

Template overlap (Almeida et al.; 1006.2035)

Jet Pull (Gallicchio, Schwartz; 1001.5027)

HEP Top Tagger (Plehn, Salam, Spannowsky; 0910.5472)

http://www.physics.ox.ac.uk/boost2010/index.asp

N-Subjettiness (Thaler, Van Tilburg 1011.2268; Kim 1011.1493)

Dipolarity (Hook, Jankowiak, Wacker; 1102.1012)

"Jet substructure without trees" (Jankowiak, Larkoski; 1104.1646)

Shower deconstruction (Spannowsky, Soper 1102.3480)

HEP Top Tagger++ (Plehn, Spannowsky, Takeuchi, Zerwas; 1006.2833...)

Multivariate quark/gluon discrimination (Gallicchio, Schwartz; 1106.3076)

ISR tagging (Krohn, Randall, Wang 1101.0810)

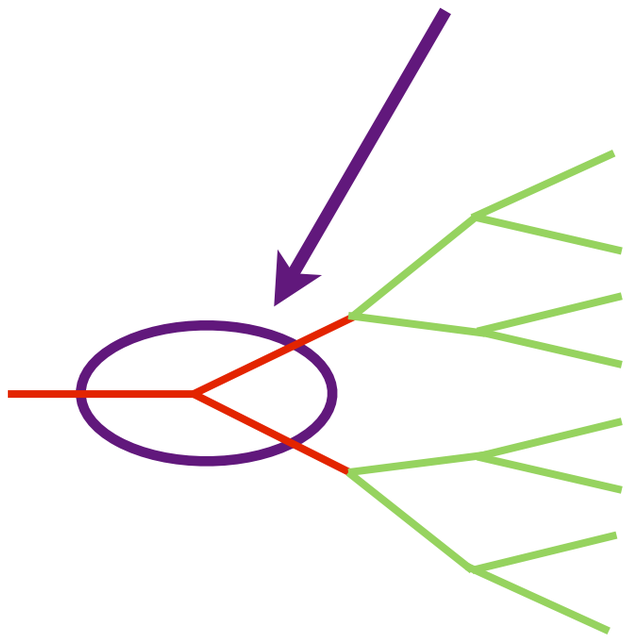
New physics multi-tagging (Kribs, Martin, Roy, Spannowsky; 0912.4731, 1006.1656)

http://www.boost2011.org/

hep-ph/0201098, hep-ph/0702150

ATLAS Y-Splitter (ATL-PHYS-CONF-2008-008): Same variable, applied to tops

Look at kT merging scale here



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

Really great idea: follow-up report!

arXiv:1012.5412

Boosted objects: a probe of beyond the standard model physics ⋆

A. Abdesselam[1], A. Belyaev[2,3], E. Bergeaas Kuutmann[4], U. Bitenc[5], G. Brooijmans[6], J. Butterworth[7], P. Bruckman de Renstrom[8], D. Buarque Franzosi[9], R. Buckingham[1], B. Chapleau[10], M. Dasgupta[11], A. Davison[7], J. Dolen[12], S. Ellis[13], F. Fassi[14], J. Ferrando[1], M.T. Frandsen[15], J. Frost[16], T. Gadfort[17], N. Glover[18], A. Haas[19], E. Halkiadakis[20], K. Hamilton[21], C. Hays[1], C. Hill[22], J. Jackson[3], C. Issever[1], M. Karagoz[1], A. Katz[23], L. Kreczko[24], D. Krohn[25], A. Lewis[1], S. Livermore[1], P. Loch[26], P. Maksimovic[27], J. March-Russell[15], A. Martin[28], N. McCubbin[3], D. Newbold[24], J. Ott[29], G. Perez[30], A. Policchio[13], S. Rappoccio[27], A.R. Raklev[31], P. Richardson[18], G.P. Salam[25,32,33], F. Sannino[34], J. Santiago[35], A. Schwartzman[19], C. Shepherd-Themistocleous[3], P. Sinervo[36], J. Sjoelin[37,38], M. Son[39], M. Spannowsky[40], E. Strauss[19], M. Takeuchi[41], J. Tseng[1], B. Tweedie[27,42], C. Vermilion[43], J. Voigt[29], M. Vos[44], J. Wacker[19], J. Wagner-Kuhr[29], and M.G. Wilson[19]

M. Karagoz, G. P. Salam, M. Spannowsky, M. Vos (editors)

The Eagle and Chi

Outcome I: Survey of motivation

What signals involve boosted heavy particles?

Which techniques have been suggested for each of these signals?

Outcome II: Survey of techniques

This time organized by strategy, not target

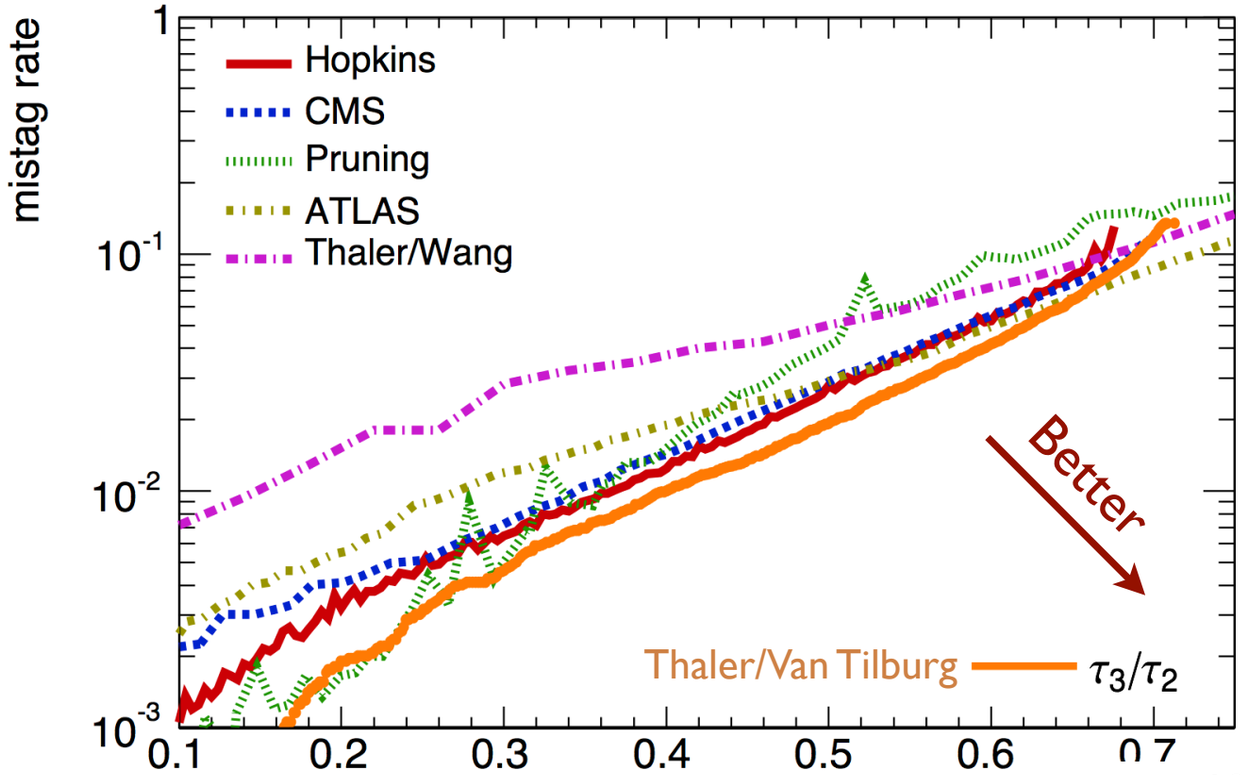
Allowed some degree of comparison between techniques.

Outcome III: Survey of experimental results

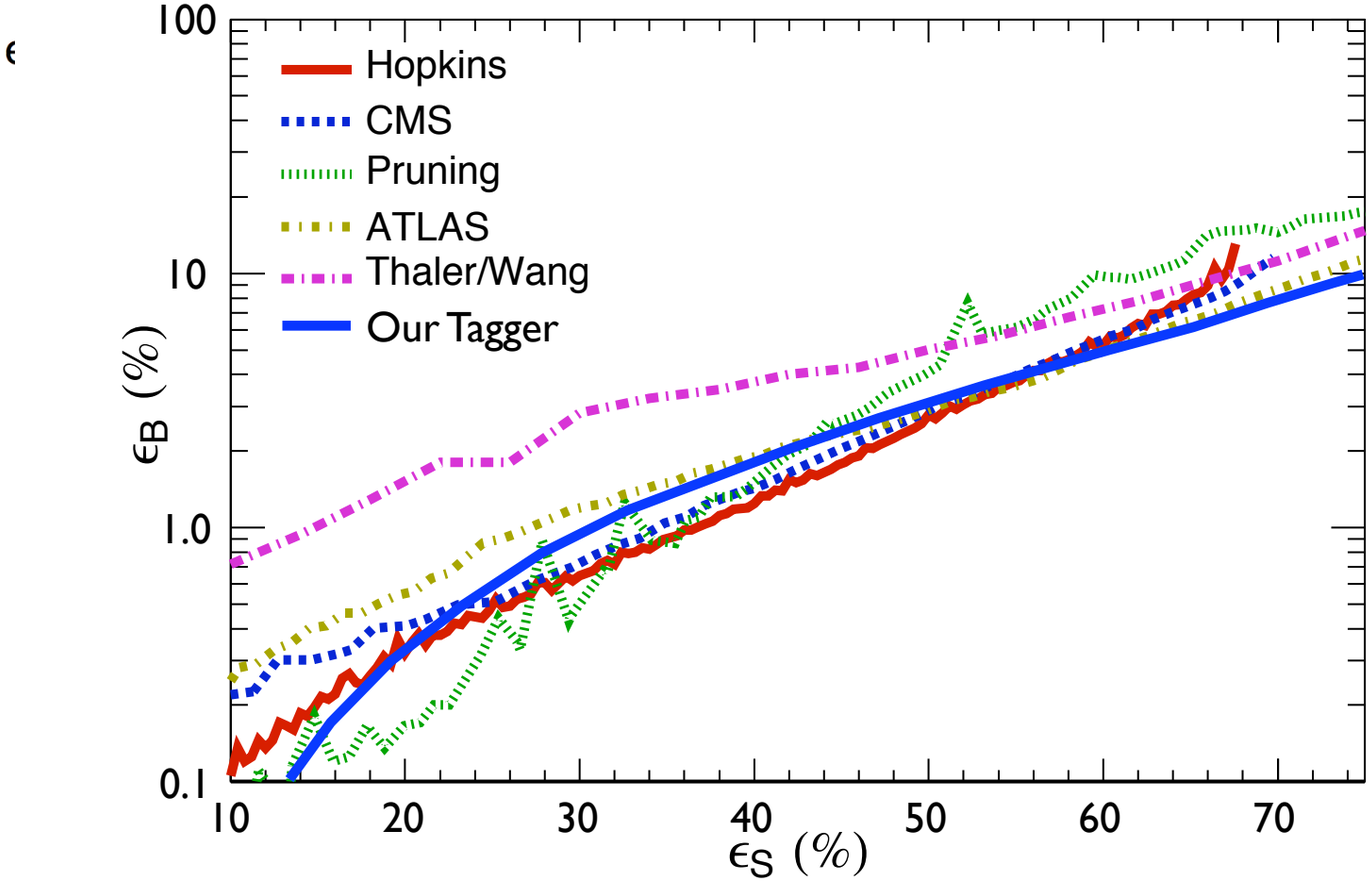
First survey of its kind -- starts from the beginning!

Outcome IV: Comparisons on public samples

NSubjettiness

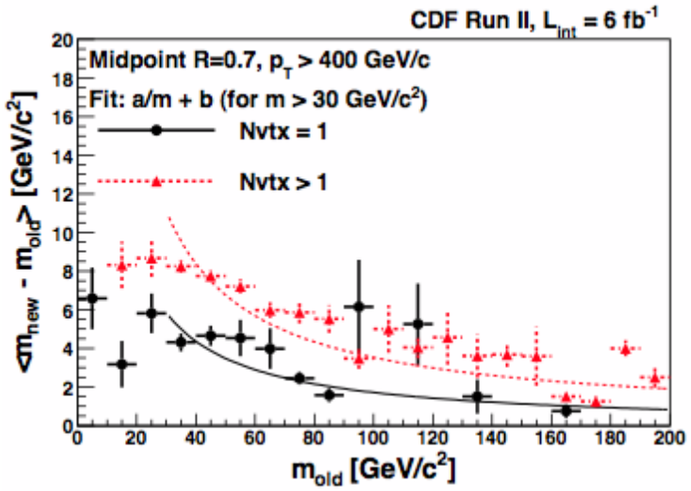


"Substructure without trees" / angular correlation functions

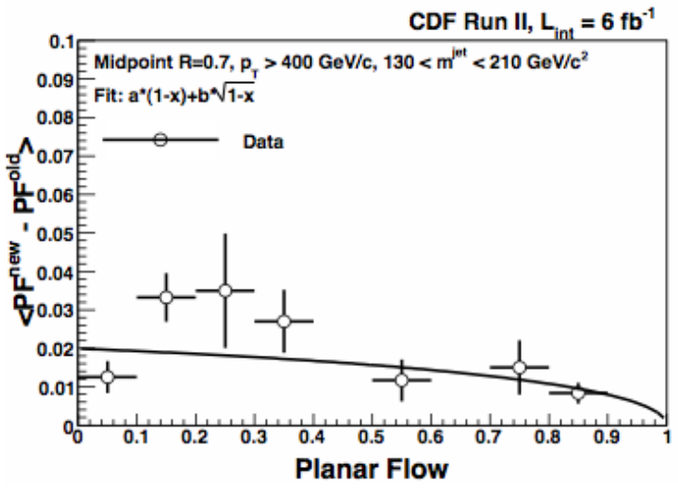


anti-kT, R=1.0, 500 GeV < pT < 600 GeV

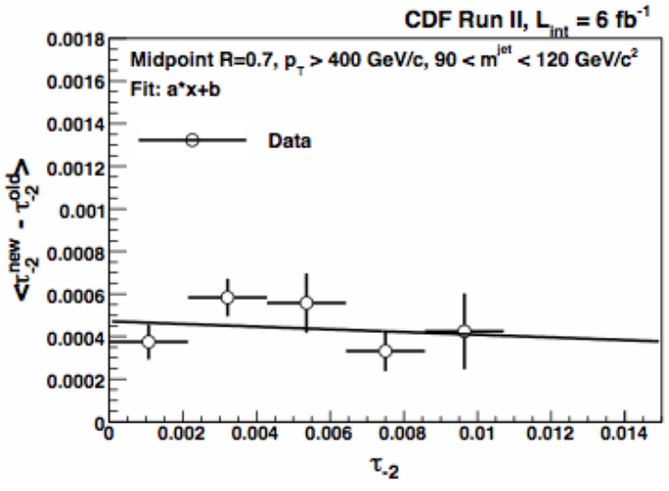
BOOST 2011 Datastravaganza!



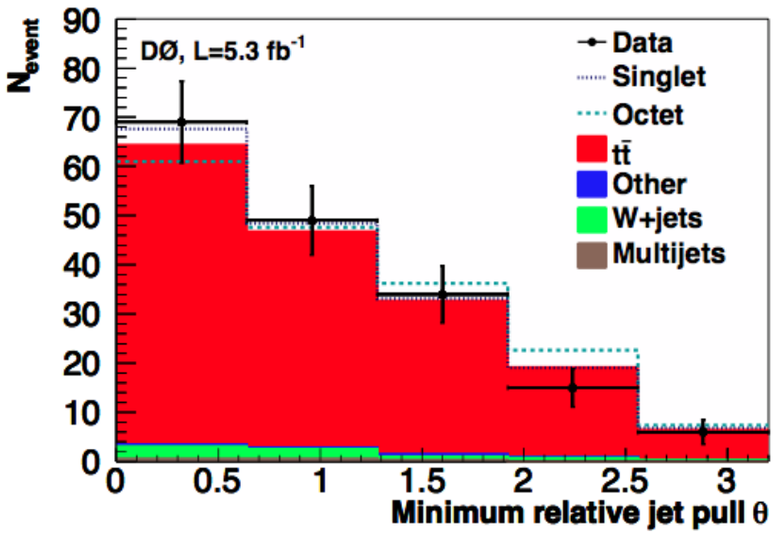
(a) Mass shift.



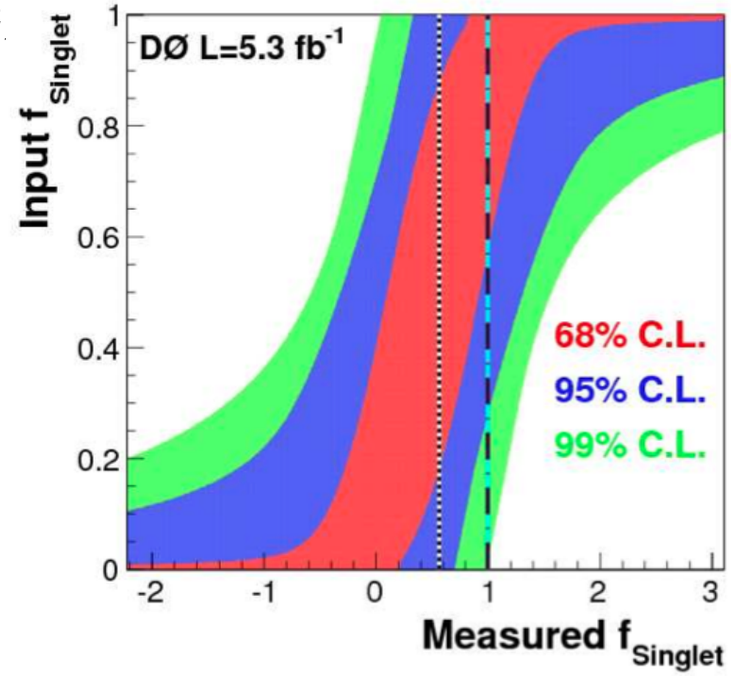
(b) Planar flow shift.



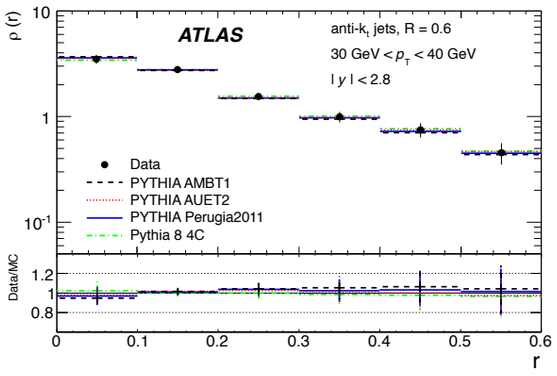
(c



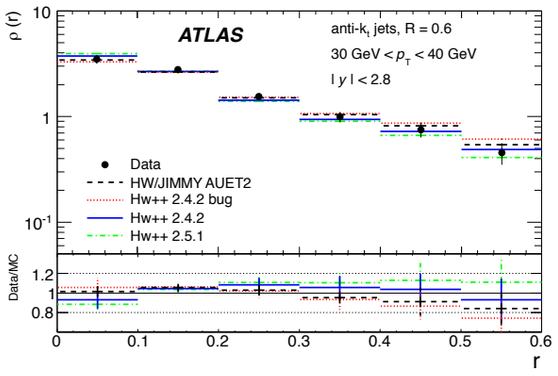
(a)



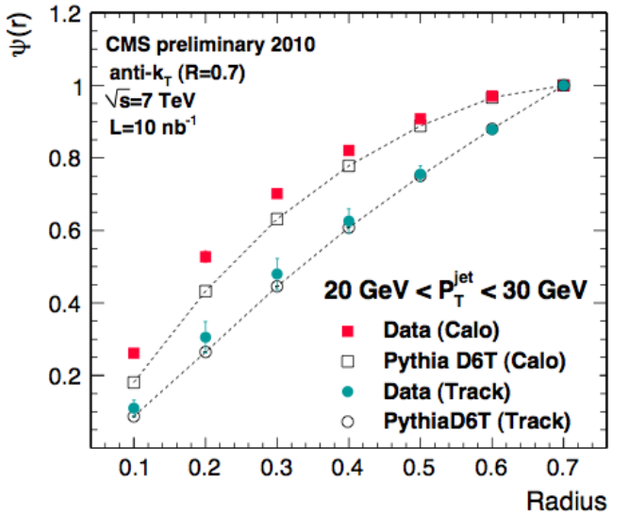
(b)



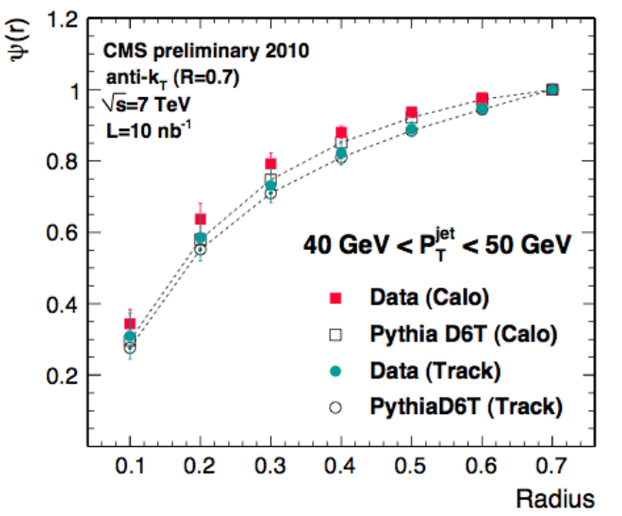
(a) PYTHIA tunes



(b) HERWIG++ tunes



(a) $20 < p_T < 30$ GeV



(b) $40 < p_T < 50$ GeV

Goals for 2011 report

Summarize progress

Establish theory/experiment goals

Extend comparisons

Publish tools

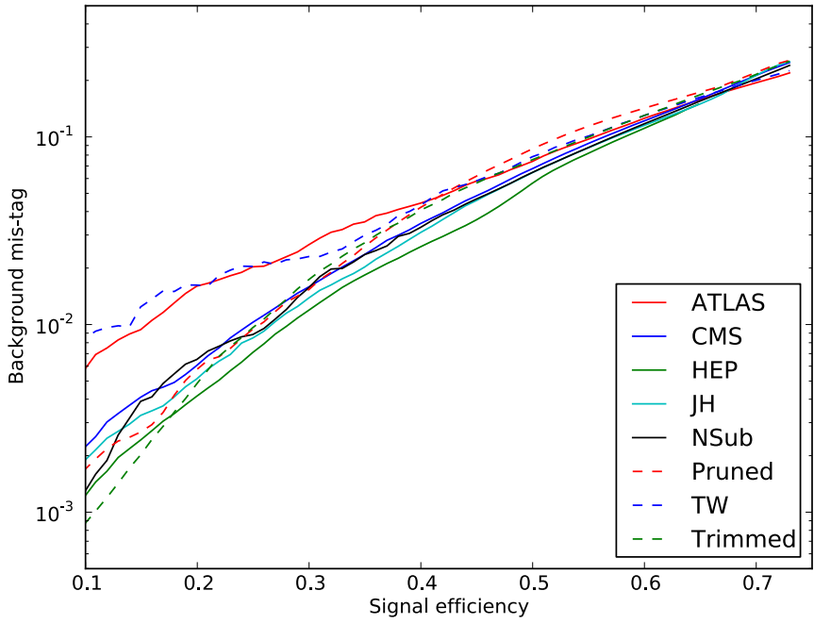
Comparisons

Extend last yet with:

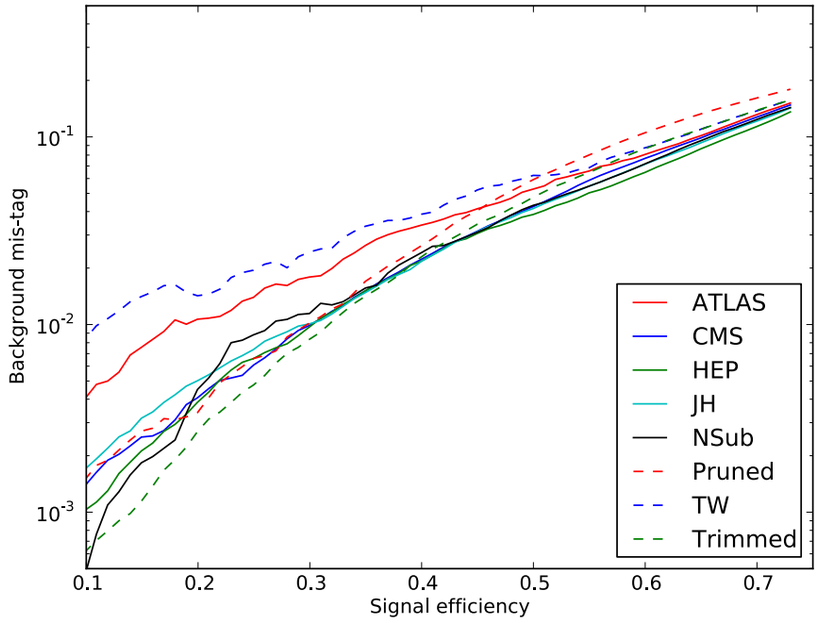
more methods

more MCs

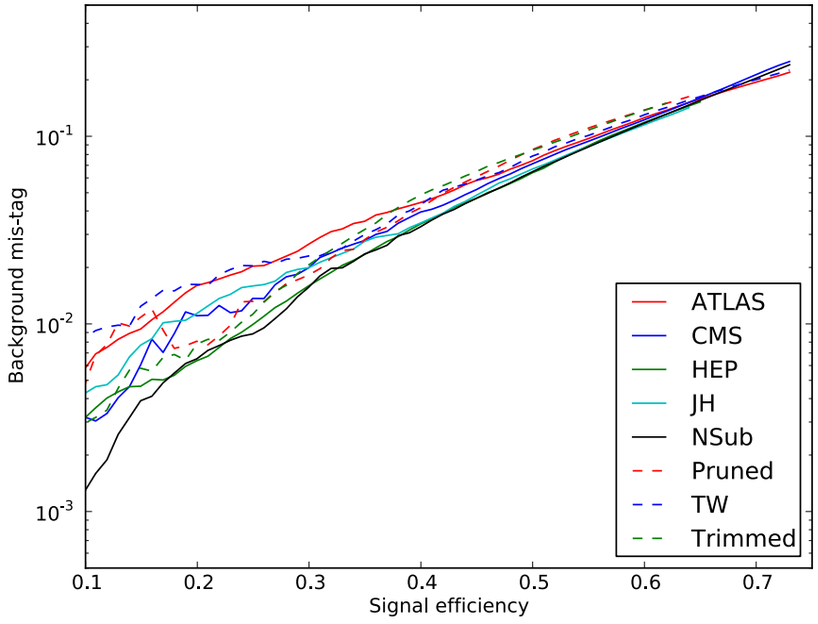
standardized, published tools



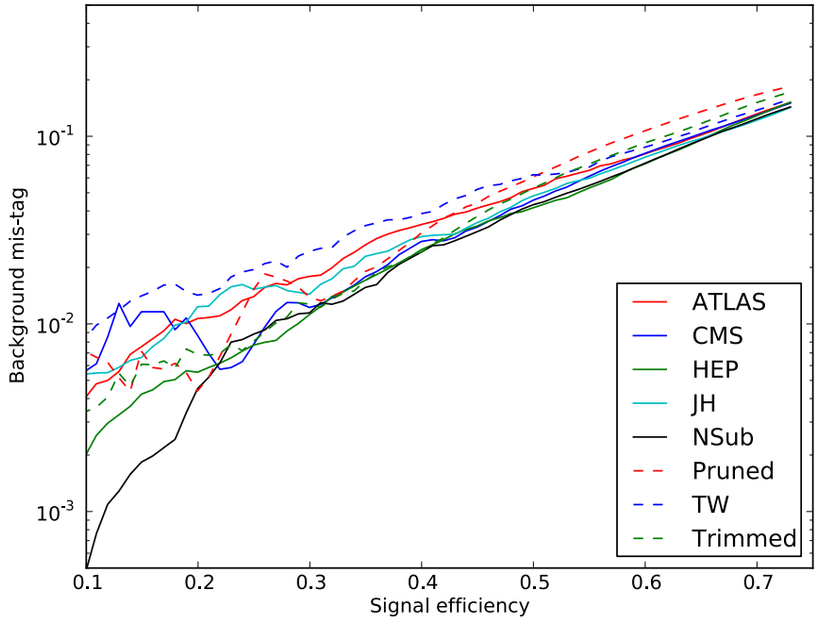
(a) all p_T, optimised



(b) p_T 500–600 GeV, optimised



(c) all p_T

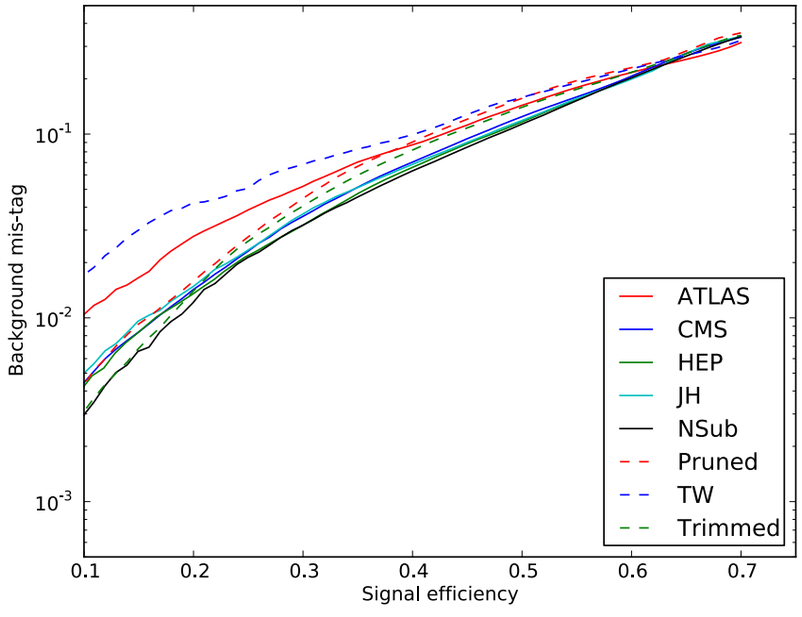


(d) p_T 500–600 GeV

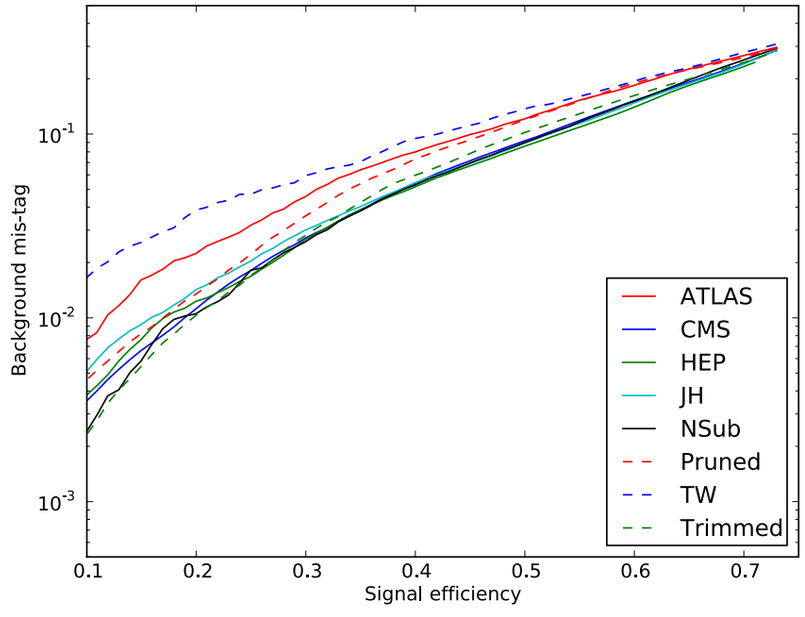
Figure 15. Mis-tag vs. efficiency for several top tagging methods, as tested on HERWIG++ $t\bar{t}$ and dijet samples. For Figures (a) and (b), the input parameters are optimised for each efficiency point. The input parameters for the unoptimised scans are taken from the 35% efficiency point in Figure (b).

tops vs. QCD, pt. 2

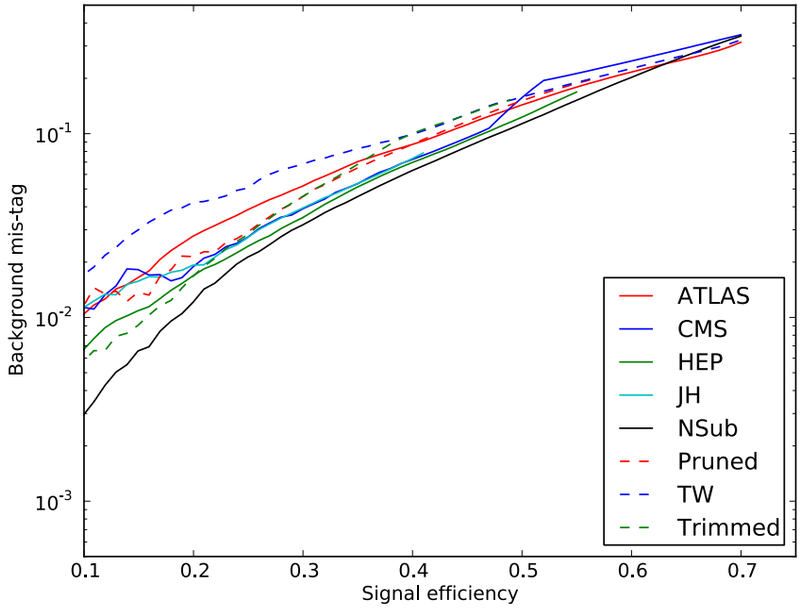
Herwig++



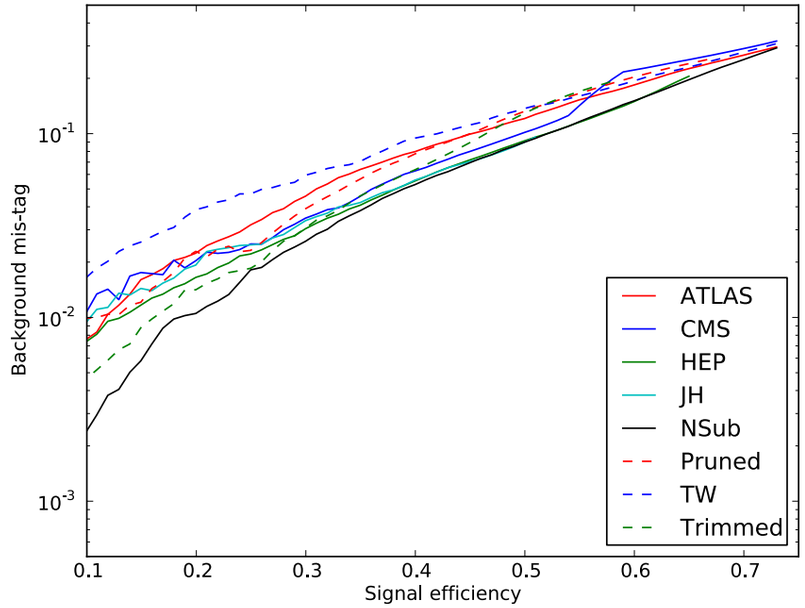
(a) all p_T, optimised



(b) p_T 500–600 GeV, optimised



(c) all p_T

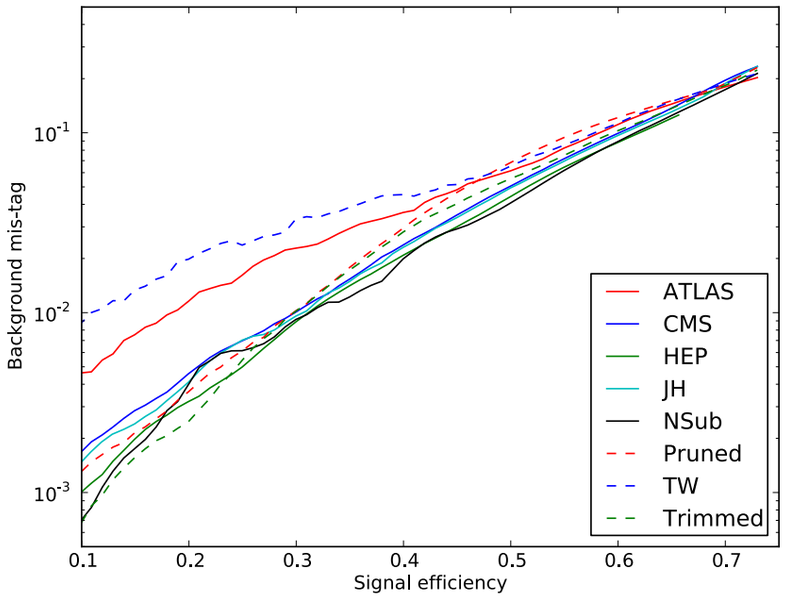


(d) p_T 500–600 GeV

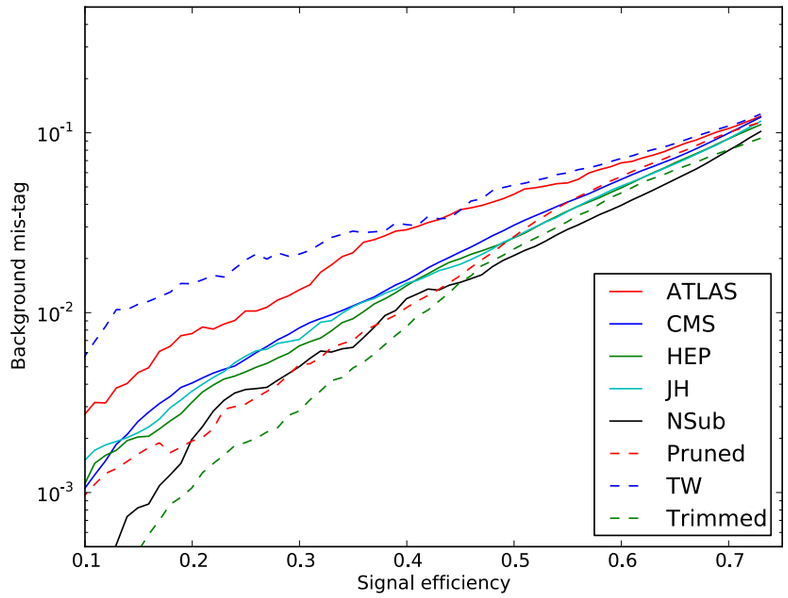
Figure 16. Mis-tag vs. efficiency for several top tagging methods, as tested on SHERPA matched $t\bar{t}$+ jets and multijet samples. For Figures (a) and (b), the input parameters are optimised for each efficiency point. The input parameters for the unoptimised scans are taken from the 35% efficiency point in Figure (b).

tops vs. QCD, pt. 2

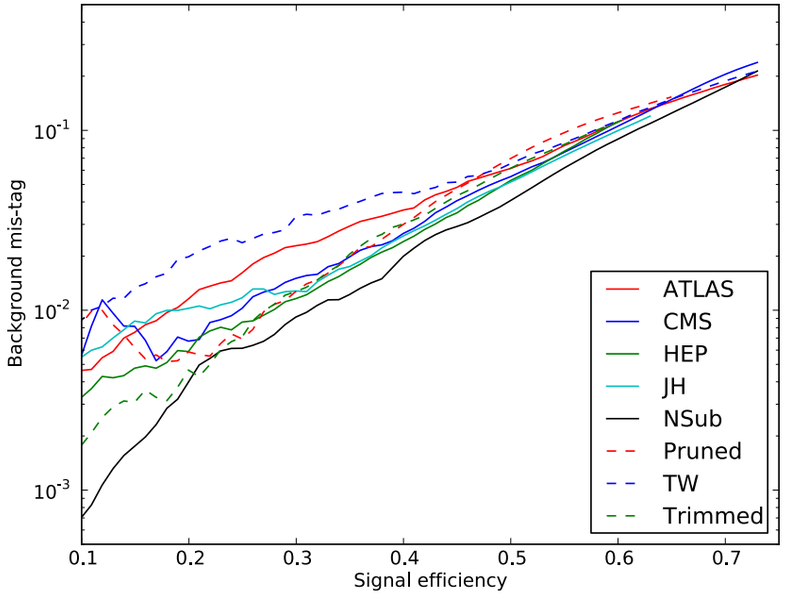
Sherpa



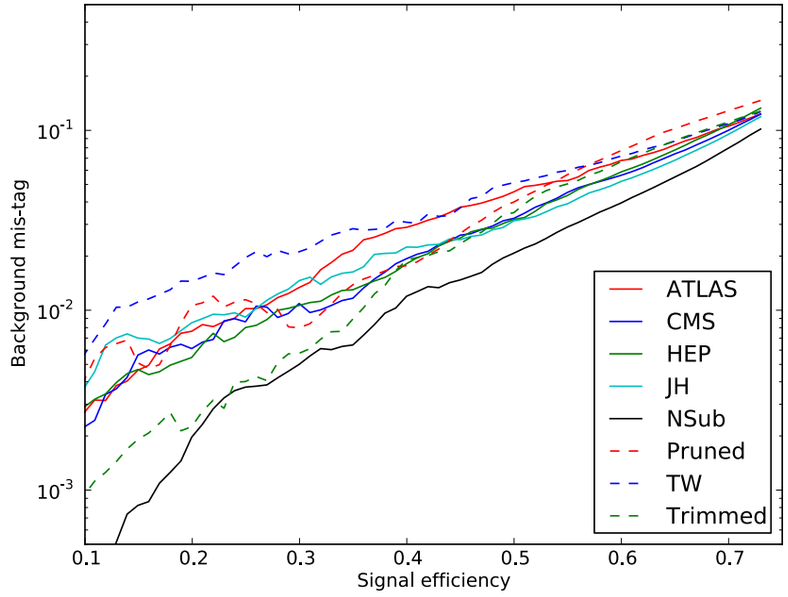
(a) all p_T, optimised



(b) p_T 500–600 GeV, optimised



(c) all p_T

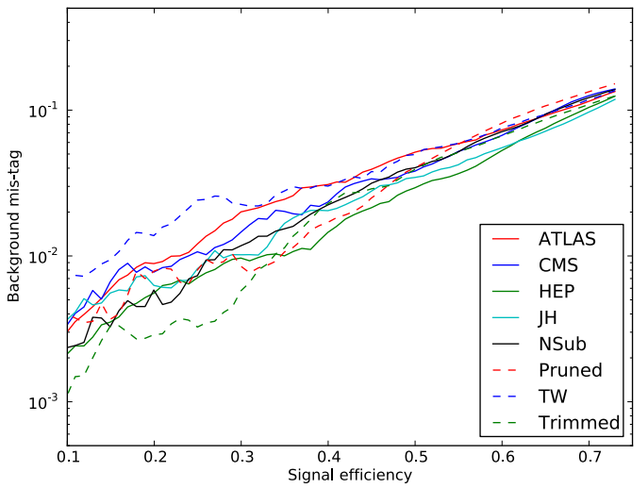


(d) p_T 500–600 GeV

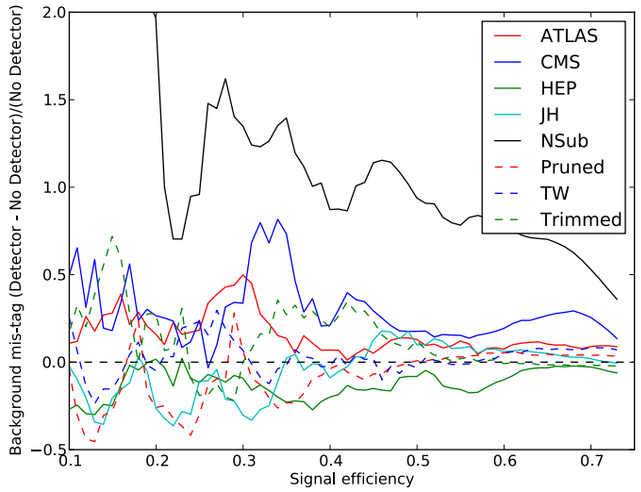
Figure 14. Mis-tag vs. efficiency for several top tagging methods, as tested on HERWIG 6.5 $t\bar{t}$ and dijet samples. For Figures (a) and (b), the input parameters are optimised for each efficiency point. The input parameters for the unoptimised scans are taken from the 35% efficiency point in Figure (b).

tops vs. QCD, pt. 2

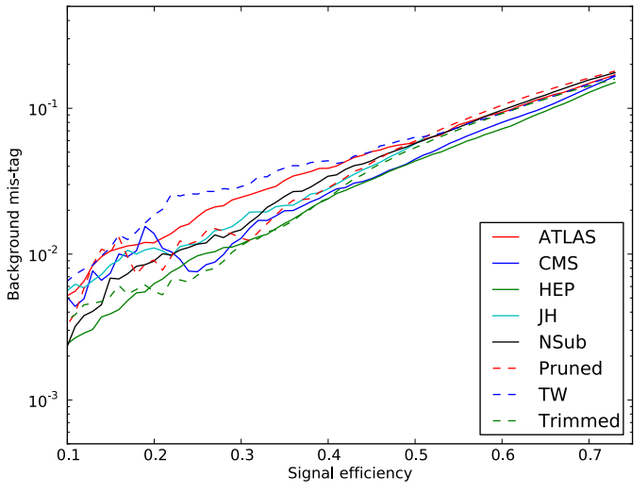
Herwig 6.5



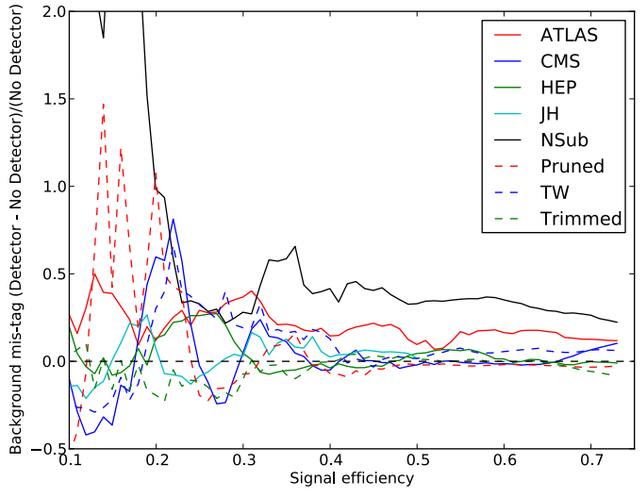
(a) HERWIG



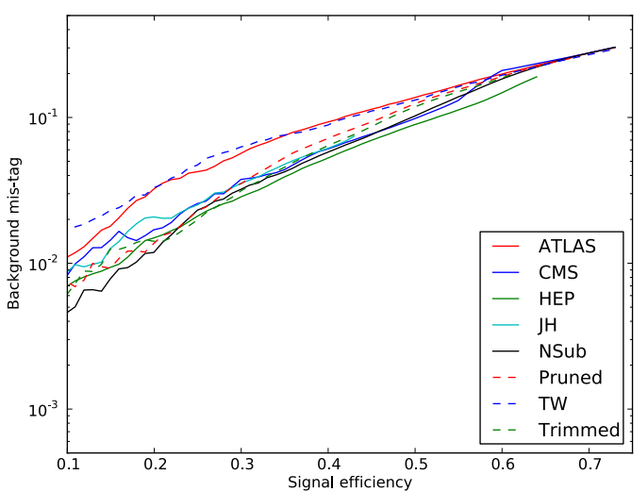
(b) HERWIG, fractional difference

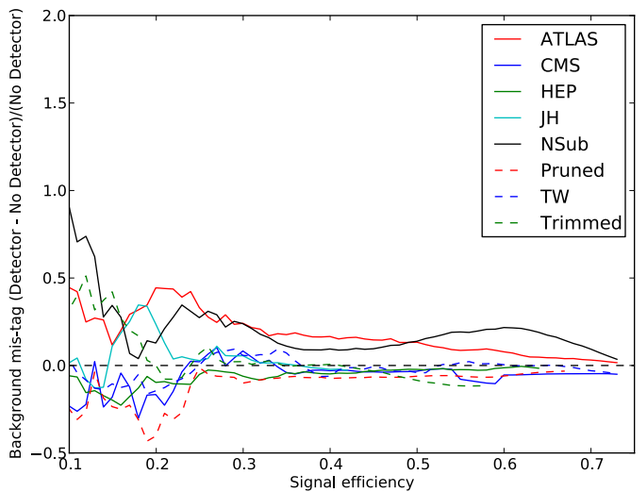


(c) HERWIG++



(d) HERWIG++, fractional difference





tops vs. QCD, pt. 2

Detector effects

Lots of questions left unanswered!

Other signals?

Pile up?

Theory synthesis?

Complicated topologies?

More interesting p_T ranges?

Assignment: Answer these questions!