The (Recent) History of Jet Substructure & Boosted Physics

Steve Ellis

Big Picture:

The LHC is intended to find new “stuff” (BSM physics)

At the LHC new and old heavy particles will often be boosted → 1 jet

It will be a challenge to find the new stuff!

☞ Experimenters and Theorists will need to work together!!!

Borrowed results from
many sources – thanks!

Department of Physics
University of Washington

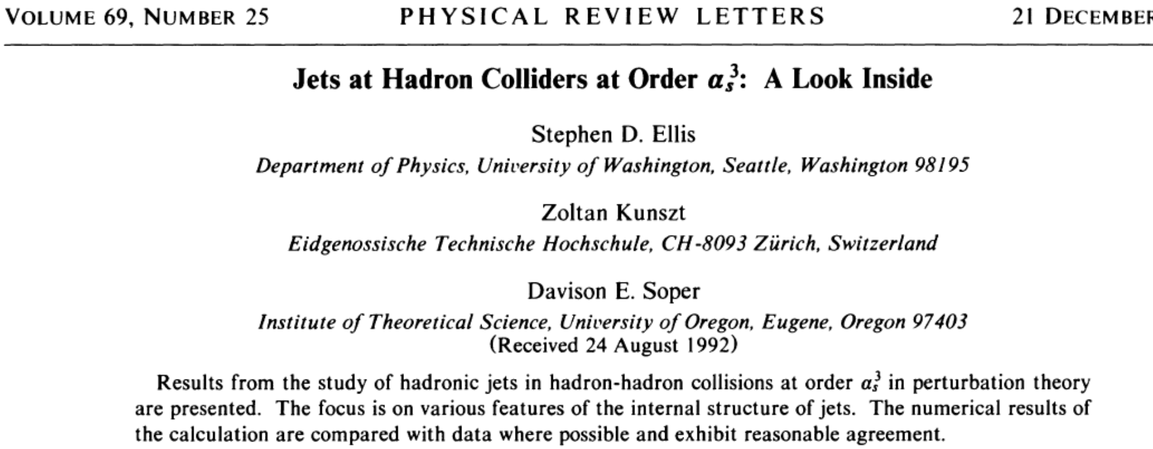
West Coast ATLAS Forum 6/15/11

A Brief History – (as an older person history is my thing)

- In the Beginning (~1970):Jets were thought of as single partons
- Considered to have no relevant substructure
- No unique definition – a colored object showering – details depend on specific algorithm
- Algorithms evolved in 1980’s leading to Snowmass 1990, and the iterative cone algorithm
- Beginning of pQCD era of jet studies

More History

- First jet substructure – the pT profile



VOLUME 69, NUMBER 25 PHYSICAL REVIEW LETTERS 21 DECEMBER 1992

Jets at Hadron Colliders at Order α_s^3: A Look Inside

Stephen D. Ellis
Department of Physics, University of Washington, Seattle, Washington 98195

Zoltan Kunszt
Eidgenossische Technische Hochschule, CH-8093 Zürich, Switzerland

Davison E. Soper
Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403

(Received 24 August 1992)

Results from the study of hadronic jets in hadron-hadron collisions at order α_s^3 in perturbation theory are presented. The focus is on various features of the internal structure of jets. The numerical results of the calculation are compared with data where possible and exhibit reasonable agreement.

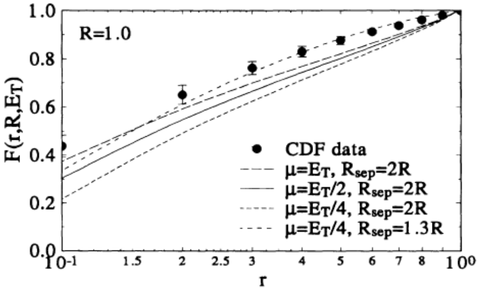


FIG. 2. $F(r,R,E_T)$ vs r for $R=1.0$, $\sqrt{s}=1800$ GeV, $E_T=100$ GeV, and $0.1<|\eta|<0.7$ with $\mu=E_T/4$, $E_T/2$, E_T compared to data from CDF [7]; the dot-dashed curve is explained in the text.

$F(r,R,E_T)$ is the average fraction of the jets' transverse energy that lies inside an inner cone of radius $r<R$ (concentric with the jet defining cone). Stated another way, the quantity $1-F(r,R,E_T)$ describes the fraction of E_T that lies in the annulus between r and R. It is this latter

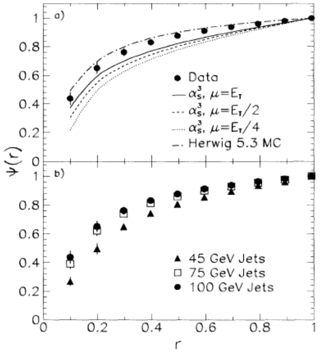
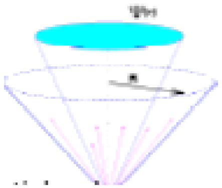
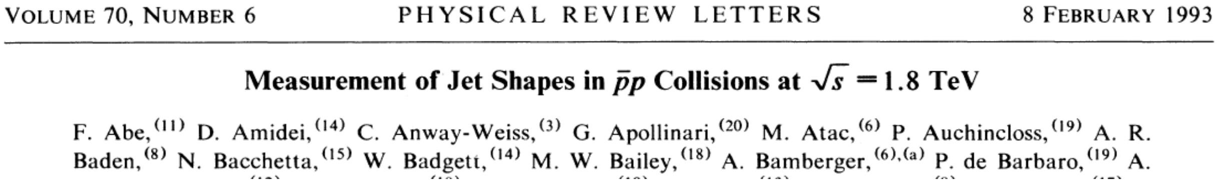


FIG. 1. (a) The distribution of the P_T fraction in a cone for 100 GeV E_T jets and cone size of $R_0=1.0$. The variable plotted, $\Psi(r)$, is the ratio of P_T within a cone of radius r to the P_T within a cone of radius $R_0=1.0$. Systematic uncertainties dominate the errors. Also shown are QCD calculations: α_s^3 theory calculations, using HMRS B structure functions for $\Lambda_{QCD}=122$ MeV and different scales μ; the prediction from the HERWIG Monte Carlo version 5.3. (b) $\Psi(r)$ for 45, 70, and 100 GeV jets.





VOLUME 70, NUMBER 6 PHYSICAL REVIEW LETTERS 8 FEBRUARY 1993

Measurement of Jet Shapes in $\bar{p}p$ Collisions at $\sqrt{s}=1.8$ TeV

F. Abe,(11) D. Amidei,(14) C. Anway-Weiss,(3) G. Apollinari,(20) M. Atac,(6) P. Auchincloss,(19) A. R. Baden,(8) N. Bacchetta,(15) W. Badgett,(14) M. W. Bailey,(18) A. Bamberger,(6),(a) P. de Barbaro,(19) A.

$$\rho(r)=\frac{\xi(r)}{\int_0^{R_0}\xi(r')dr'}, \quad \text{with} \quad \xi(r)\equiv\frac{1}{\mathcal{N}_{\text{jet}}}\sum_{\text{jets}}\int_{P_T>P_T^{\min}}\frac{P_T}{\mathcal{P}_T^{\text{jet}}}\frac{d^2N}{dr\,dP_T}dP_T\,.$$

The integral shape variable $\Psi(r)=\int_0^r\rho(r')dr'$ is used to compare data with theory. Note that r is related to the

In Fig. 1(a)

And at the same time:

- Mike Seymour was already talking about boosted things (but largely unheard)

2) **Tagging a heavy Higgs boson.**
M.H. Seymour, (Cambridge U.) . CAVENDISH-HEP-90-25, Jan 1991. 16pp. Talk presented at the ECFA LHC Workshop, Aachen, Germany, Oct 4-9, 1990.
Published in Aachen ECFA Workshop 1990:0557-569 (QCD183:L25:1990:V.2)
References | LaTeX(US) | LaTeX(EU) | Harvmac | BibTeX | Keywords | Cited 1 time | More Info
CERN Library Record
Scanned Version (KEK Library)
Conference Info
Bookmarkable link to this information

"The W-finder used in this study utilises this cut by running a jet-finder twice, with cone sizes of ΔR=0.75 and ΔR=0.25, and then demands a big jet containing two small jets, with $|m_{jj}-m_W|<10$ GeV."

- We still we didn't

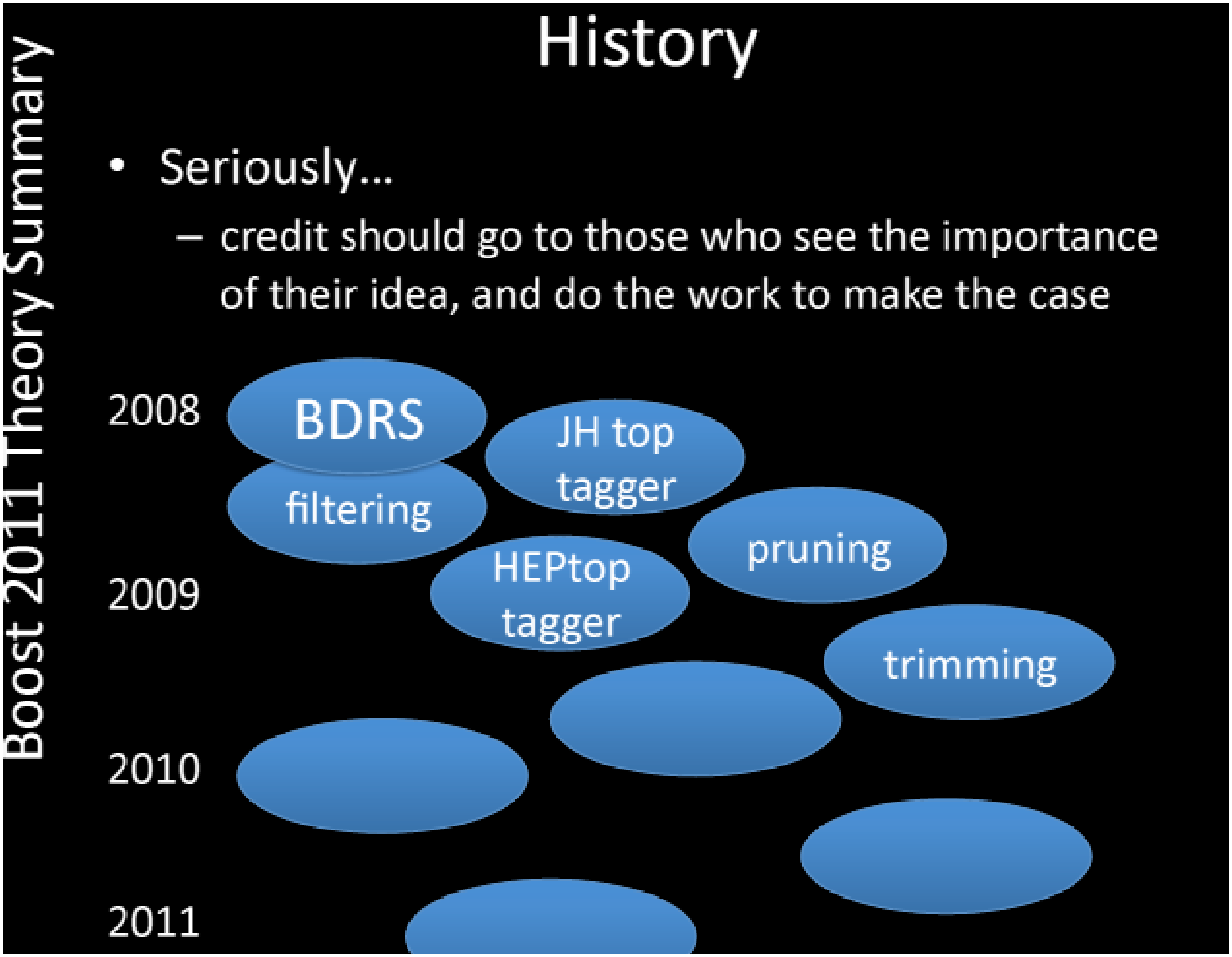
  discuss jet masses!!

- Jet algorithms became more sophisticated in 2000's
  - recombination algorithms including anti-kT

From MS's Boost 2011 Summary Talk

Finally in Late 2000’s things picked up

A variety of jet grooming & tagging tools



Boost 2011 Theory Summary

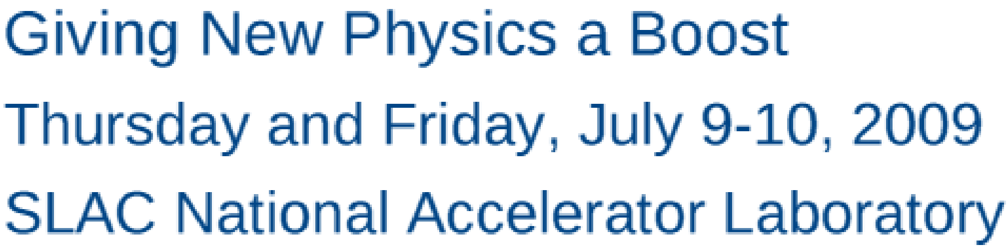
History

- Seriously…
  - credit should go to those who see the importance of their idea, and do the work to make the case

From Mike Seymour’s Summary talk

A real physics success story of rapid transition from theory ideas to experimental tools:

- Aided by the BOOST meetings –
  Princeton, May 2011
  Oxford, June 2010



Giving New Physics a Boost

Thursday and Friday, July 9-10, 2009

SLAC National Accelerator Laboratory

+ U. Washingtop, january 2010
+ Manchester, November 2010
+ Boston, january 2011
+ Oregon, february 2011
+ LPCC, CERN, february 2011

“New” Results at BOOST 2011

- No time to review all, but note the outstanding summary talks:
  ⇒ Mike Seymour – Theory
  ⇒ Jon Butterworth – Experiment

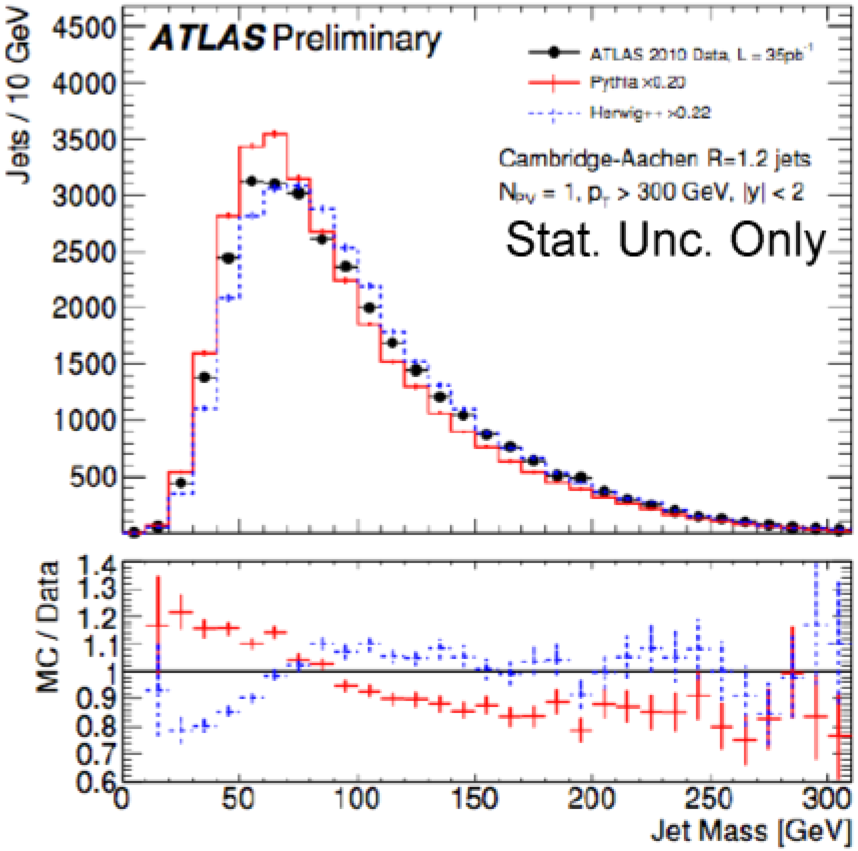
- Results in many areas:
  Tools – SpartyJet, FastJet, various Taggers
  Observables - color flow, N-subjetiness, Gluon-tagging, No Tree substructure,…
  Calculations – improved MCs, Improved SCET
  Applications - top FB asym, unconventional SUSY stops,…

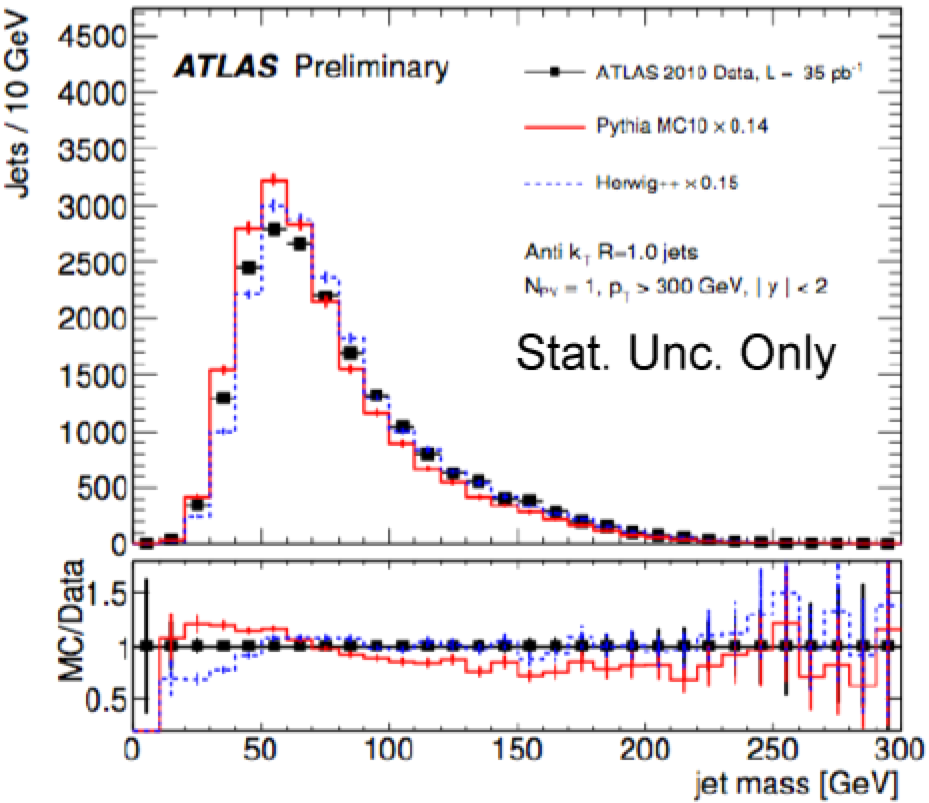
- Here just mention some (personally) favorite highlights

- Main Message - Boosted techniques are being tested/certified at the LHC (although few public results as yet)

Certifying Jet Substructure

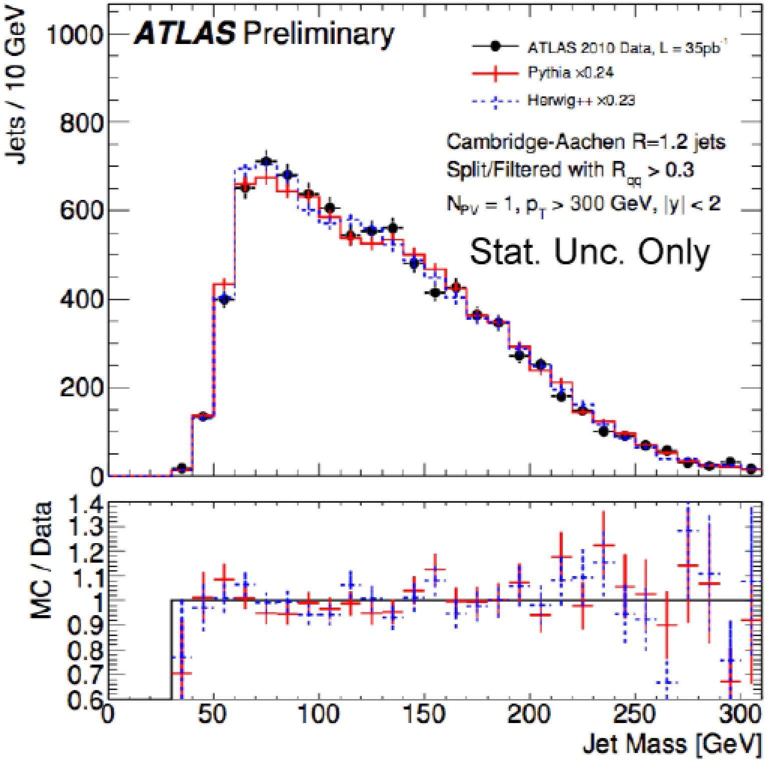
- Start with jet masses – Adam Davison - ATLAS
- Description of mass shape is reasonable in MC samples





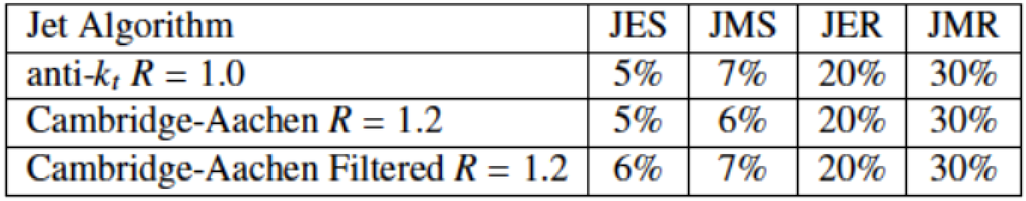
Still understood after Grooming

- Basic description appears to be very good here



- Getting handle on uncertainties

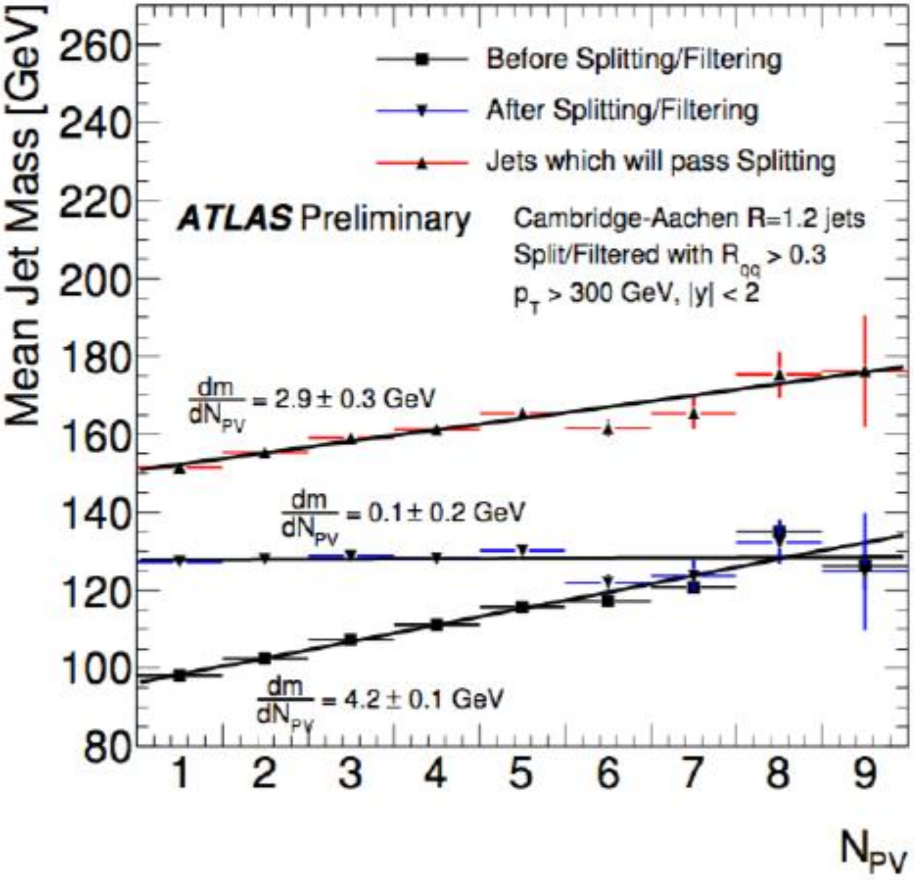
- Have the following (correlated) systematics:



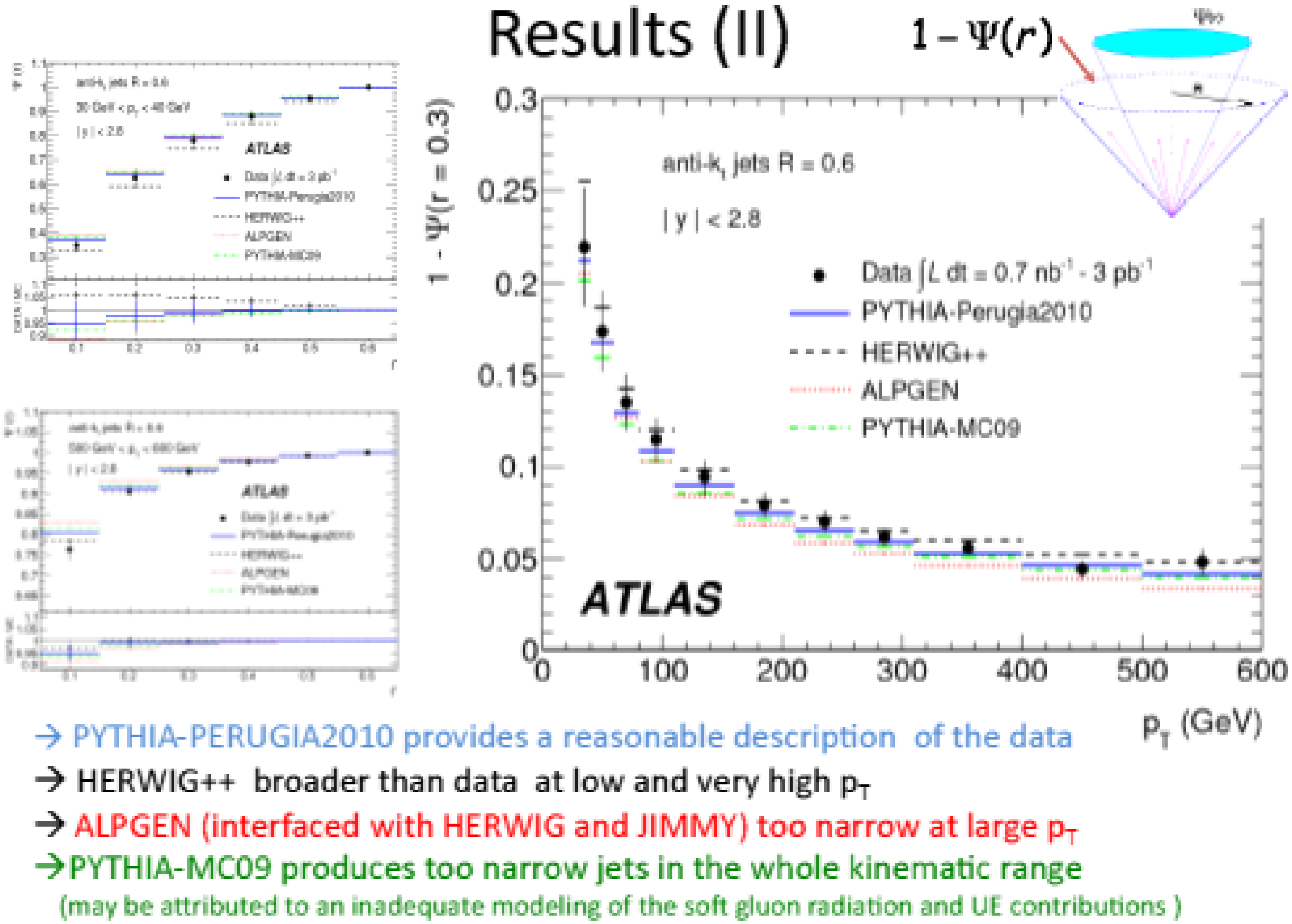
Jet Algorithm	JES	JMS	JER	JMR
anti-k_t $R = 1.0$	5%	7%	20%	30%
Cambridge-Aachen $R = 1.2$	5%	6%	20%	30%
Cambridge-Aachen Filtered $R = 1.2$	6%	7%	20%	30%

As expected Grooming reduces sensitivity to Pile-Up

- Filtering reduces effective jet area
- Should therefore reduce pile-up dependence
- Slope in fact consistent with zero after filtering



From Mario Martinez – good start but precision will require tuning of the MCs



Results (II)

→ PYTHIA-PERUGIA2010 provides a reasonable description of the data

→ HERWIG++ broader than data at low and very high p_T

→ ALPGEN (interfaced with HERWIG and JIMMY) too narrow at large p_T

→ PYTHIA-MC09 produces too narrow jets in the whole kinematic range
(may be attributed to an inadequate modeling of the soft gluon radiation and UE contributions)

A new observable -

Jet Substructure Without Trees

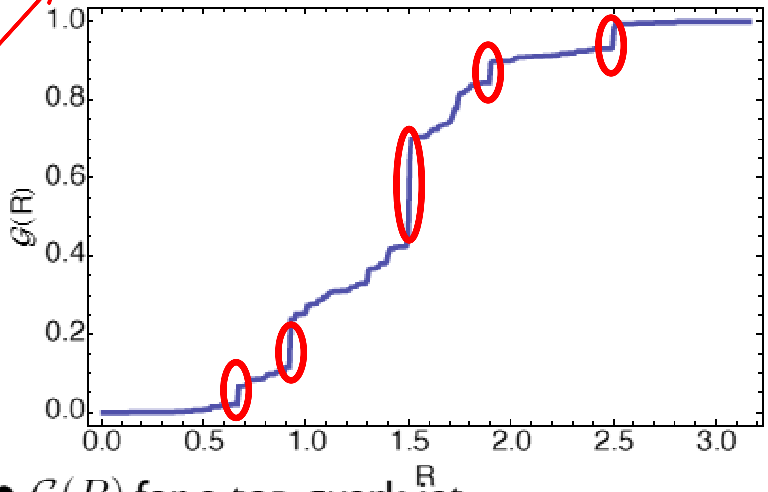
Andrew Larkoski
SLAC, Stanford University
with Martin Jankowiak
arXiv:1104.1646

- For any IRC safe set of particles {*i*}:

$$\mathcal{G}(R) \equiv \frac{\sum_{i\neq j} p_{Ti}p_{Tj}\Delta R_{ij}^2\Theta(R-\Delta R_{ij})}{\sum_{i\neq j} p_{Ti}p_{Tj}\Delta R_{ij}^2} \approx \frac{\sum_{i\neq j} p_i\cdot p_j\Theta(R-\Delta R_{ij})}{\sum_{i\neq j} p_i\cdot p_j}$$

- *R* is **not** measured wrt jet center
  - Distinct from angular profile
- Quantifies jet scaling in an IRC safe way
- Ledges in $\mathcal{G}(R)$ = separation of hard subjets

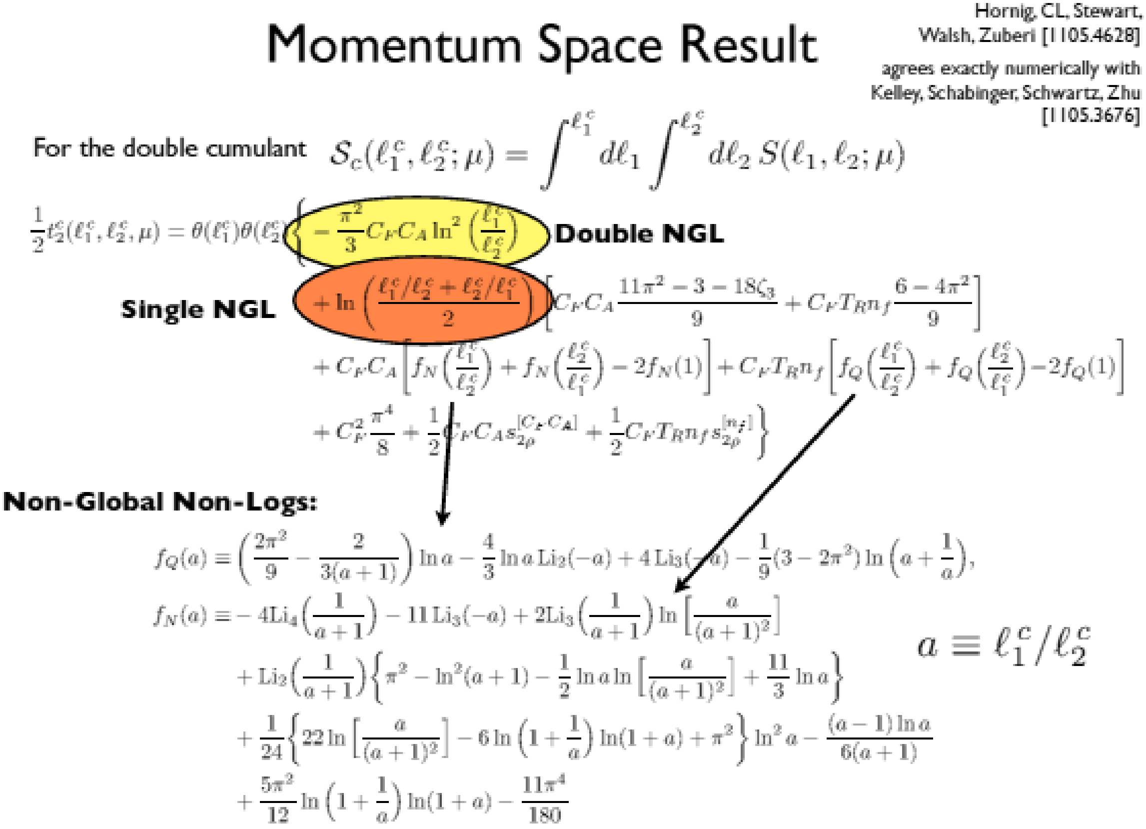
Really CLIFFs



- $\mathcal{G}(R)$ for a top quark jet

SCET applications :

- Jet Substructure is intrinsically multi-scale (jet pT > jet mass > subjet mass) with potentially large logarithms, a natural application of SCET
- New progress with soft function – Chris Lee



Momentum Space Result

Hornig, CL, Stewart, Walsh, Zuberi [1105.4628]

agrees exactly numerically with Kelley, Schabinger, Schwartz, Zhu [1105.3676]

For the double cumulant $S_c(\ell_1^c,\ell_2^c;\mu) = \int^{\ell_1^c} d\ell_1 \int^{\ell_2^c} d\ell_2\, S(\ell_1,\ell_2;\mu)$

$$\frac{1}{2}t_2^c(\ell_1^c,\ell_2^c,\mu) = \theta(\ell_1^c)\theta(\ell_2^c)\Big\{ -\frac{\pi^2}{3}C_F C_A \ln^2\left(\frac{\ell_1^c}{\ell_2^c}\right)$$

Double NGL

Single NGL

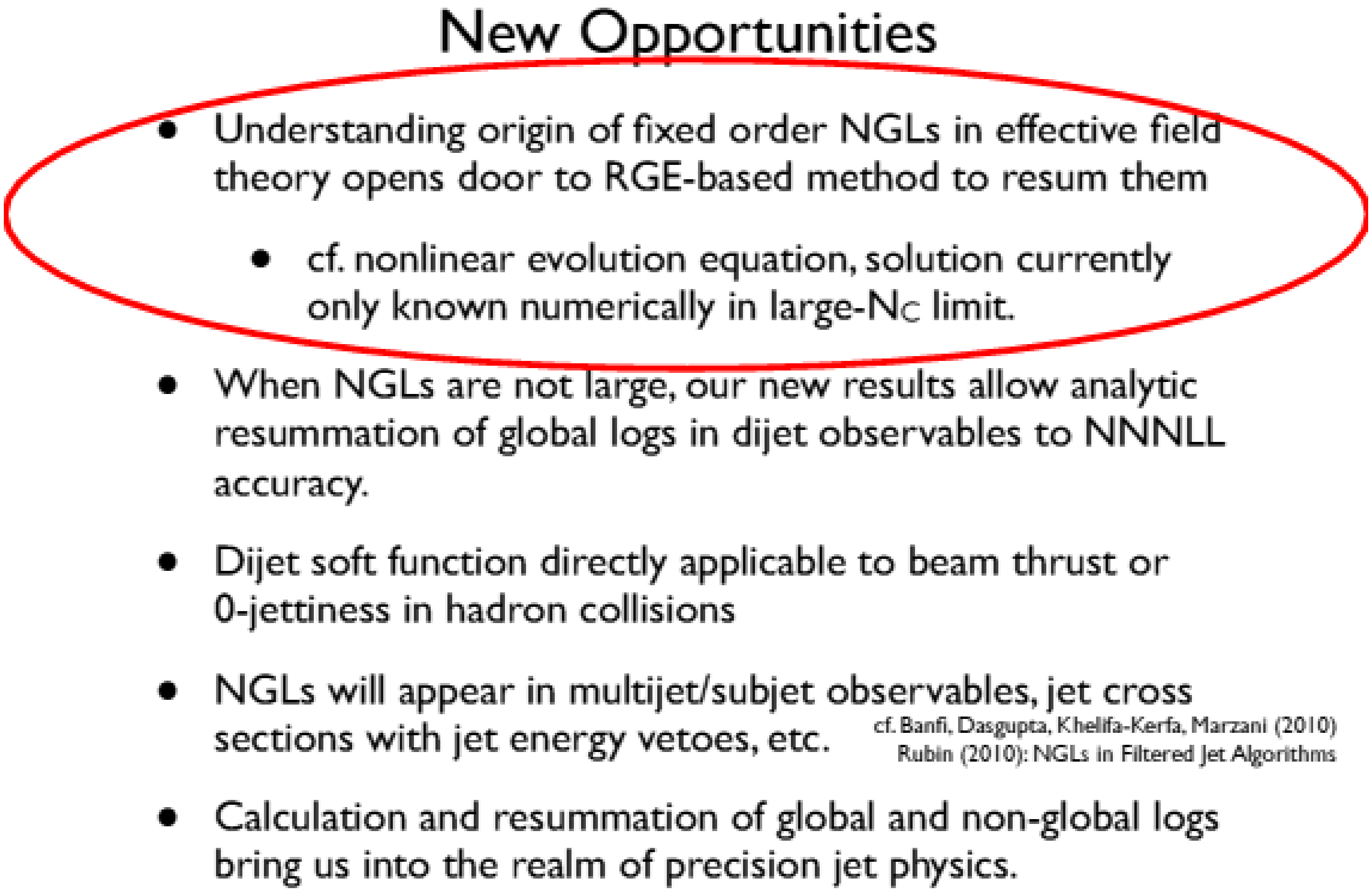
$$+\ln\left(\frac{\ell_1^c/\ell_2^c+\ell_2^c/\ell_1^c}{2}\right)\left[C_F C_A\frac{11\pi^2-3-18\zeta_3}{9}+C_F T_R n_f\frac{6-4\pi^2}{9}\right]$$
$$+C_F C_A\left[f_N\left(\frac{\ell_1^c}{\ell_2^c}\right)+f_N\left(\frac{\ell_2^c}{\ell_1^c}\right)-2f_N(1)\right]+C_F T_R n_f\left[f_Q\left(\frac{\ell_1^c}{\ell_2^c}\right)+f_Q\left(\frac{\ell_2^c}{\ell_1^c}\right)-2f_Q(1)\right]$$
$$+C_F^2\frac{\pi^4}{8}+\frac{1}{2}C_F C_A s_{2\rho}^{[C_F C_A]}+\frac{1}{2}C_F T_R n_f s_{2\rho}^{[n_f]}\Big\}$$

Non-Global Non-Logs:

$$f_Q(a) \equiv \left(\frac{2\pi^2}{9}-\frac{2}{3(a+1)}\right)\ln a-\frac{4}{3}\ln a\,\mathrm{Li}_2(-a)+4\,\mathrm{Li}_3(-a)-\frac{1}{9}(3-2\pi^2)\ln\left(a+\frac{1}{a}\right),$$
$$f_N(a) \equiv -4\mathrm{Li}_4\left(\frac{1}{a+1}\right)-11\,\mathrm{Li}_3(-a)+2\mathrm{Li}_3\left(\frac{1}{a+1}\right)\ln\left[\frac{a}{(a+1)^2}\right]$$
$$+\mathrm{Li}_2\left(\frac{1}{a+1}\right)\left\{\pi^2-\ln^2(a+1)-\frac{1}{2}\ln a\ln\left[\frac{a}{(a+1)^2}\right]+\frac{11}{3}\ln a\right\}$$
$$+\frac{1}{24}\left\{22\ln\left[\frac{a}{(a+1)^2}\right]-6\ln\left(1+\frac{1}{a}\right)\ln(1+a)+\pi^2\right\}\ln^2 a-\frac{(a-1)\ln a}{6(a+1)}$$
$$+\frac{5\pi^2}{12}\ln\left(1+\frac{1}{a}\right)\ln(1+a)-\frac{11\pi^4}{180}$$

$$a \equiv \ell_1^c/\ell_2^c$$

A new window on the infamous non-global logs = NGLs



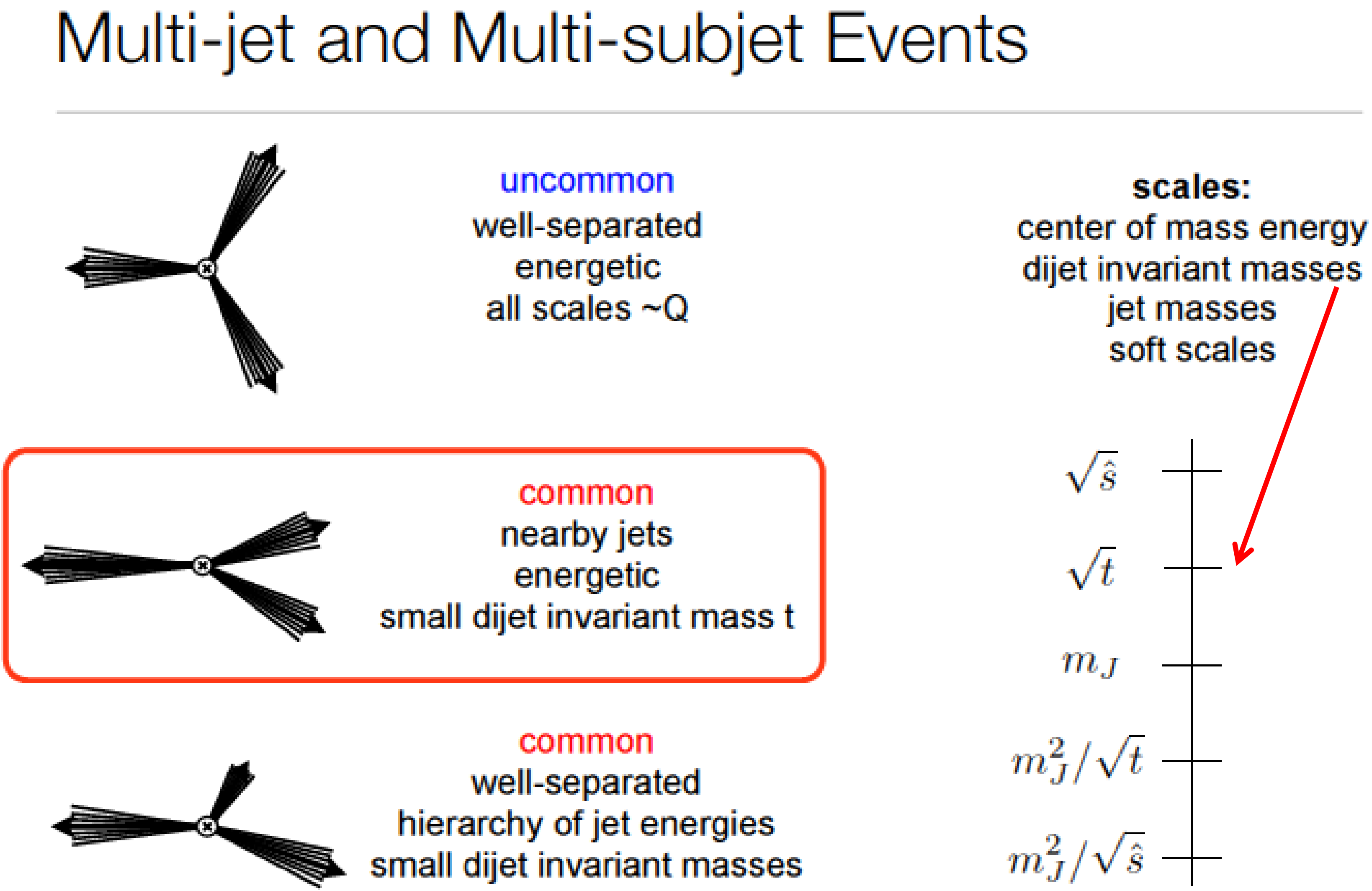
New Opportunities

- Understanding origin of fixed order NGLs in effective field theory opens door to RGE-based method to resum them
    - cf. nonlinear evolution equation, solution currently only known numerically in large-N_C limit.
- When NGLs are not large, our new results allow analytic resummation of global logs in dijet observables to NNNLL accuracy.
- Dijet soft function directly applicable to beam thrust or 0-jettiness in hadron collisions
- NGLs will appear in multijet/subjet observables, jet cross sections with jet energy vetoes, etc.

cf. Banfi, Dasgupta, Khelifa-Kerfa, Marzani (2010)
Rubin (2010): NGLs in Filtered Jet Algorithms

- Calculation and resummation of global and non-global logs bring us into the realm of precision jet physics.

Controlling Jets with SCET – Jon Walsh



Multi-jet and Multi-subjet Events

uncommon
well-separated
energetic
all scales ~Q

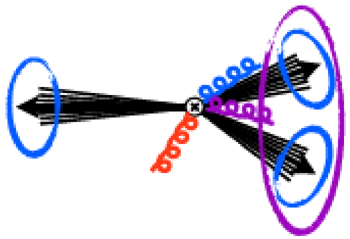
scales:
center of mass energy
dijet invariant masses
jet masses
soft scales

common
nearby jets
energetic
small dijet invariant mass t

common
well-separated
hierarchy of jet energies
small dijet invariant masses

Add new scale & mode

Modes with Nearby Jets: Collinear and Soft Modes



collinear: $p_c \sim E_J(1, \lambda^2, \lambda)$

csoft: $p_{cs} \sim E_J \frac{\lambda^2}{\lambda_t^2}(1, \lambda_t^2, \lambda_t)$

soft: $p_s \sim E_J(\lambda^2, \lambda^2, \lambda^2)$

$$\lambda = \frac{m_J^2}{Q^2}$$

$$\lambda_t = \frac{t}{Q^2}$$

$$\mathcal{T}_j(p) = n_j \cdot p$$

$$\mathcal{T}_j(p_c) = n_j \cdot p_c \sim E_J \lambda^2 \;\; \& \;\; p_c^2 \sim E_J^2 \lambda^2 \;\Rightarrow\; p_c \sim E_J(1, \lambda^2, \lambda)$$

$$\mathcal{T}_j(p_{cs}) = n_j \cdot p_{cs} \sim E_J \lambda^2 \;\; \& \;\; \frac{p_{cs}^+}{p_{cs}^-} \sim \lambda_t^2 \;\Rightarrow\; p_{cs} \sim E_J \frac{\lambda^2}{\lambda_t^2}(1, \lambda_t^2, \lambda_t)$$

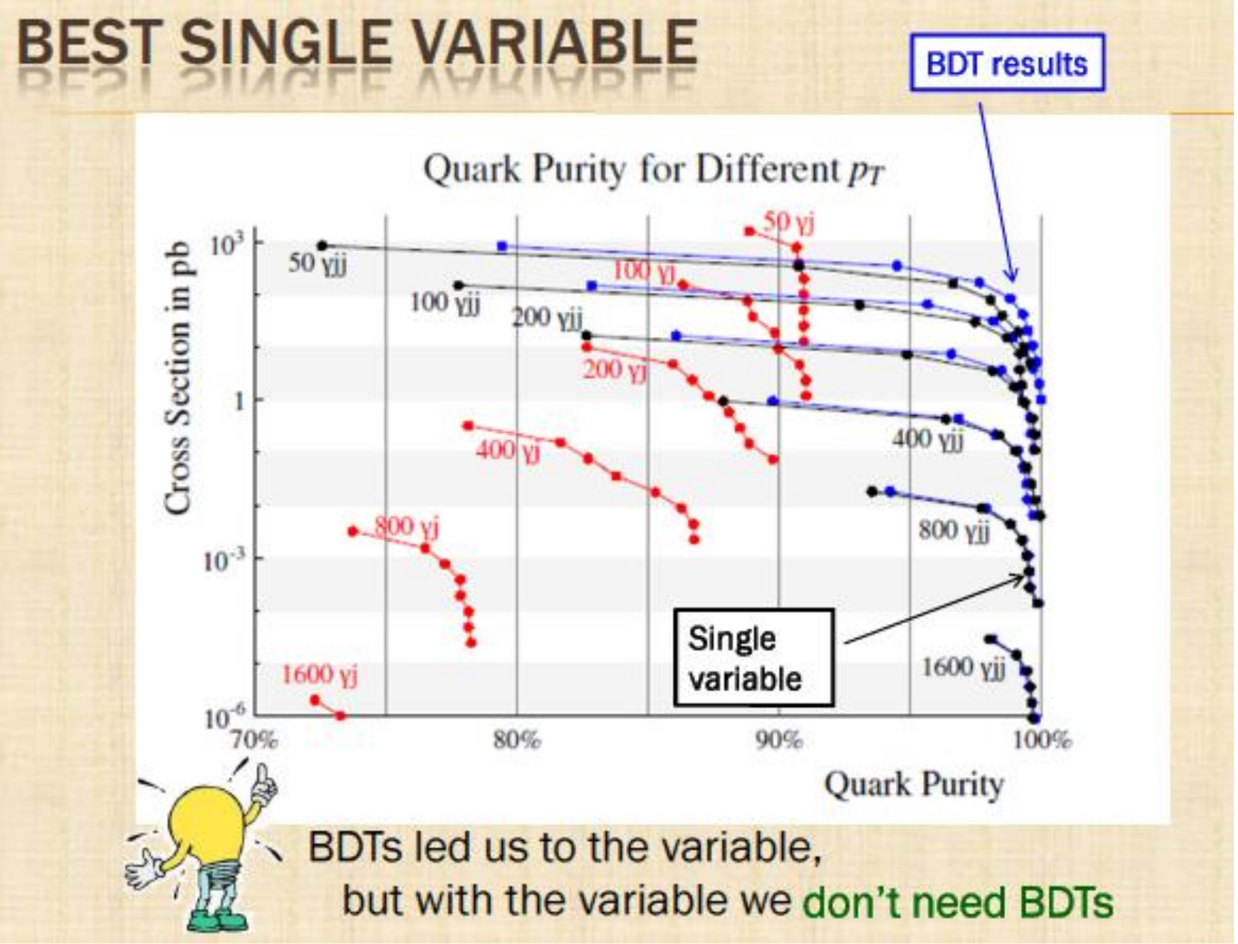
$$\mathcal{T}_j(p_s) = n_j \cdot p_s \sim E_J \lambda^2 \;\; \& \;\; p_s^2 \sim E_J^2 \lambda^4 \;\Rightarrow\; p_s \sim E_J(\lambda^2, \lambda^2, \lambda^2)$$

- Goal: NLO/LL for many multiplicities

GENEVA

Christian Bauer, Calvin Berggren, Nicholas Dunn, Andrew Hornig, Frank Tackmann, Jesse Thaler, Christopher Vermilion, Jonathan Walsh, Saba Zuberi

Really fun was lecture by theorist (Matt Schwartz) on Boosted Decision Trees – an efficient path to optimal observables!



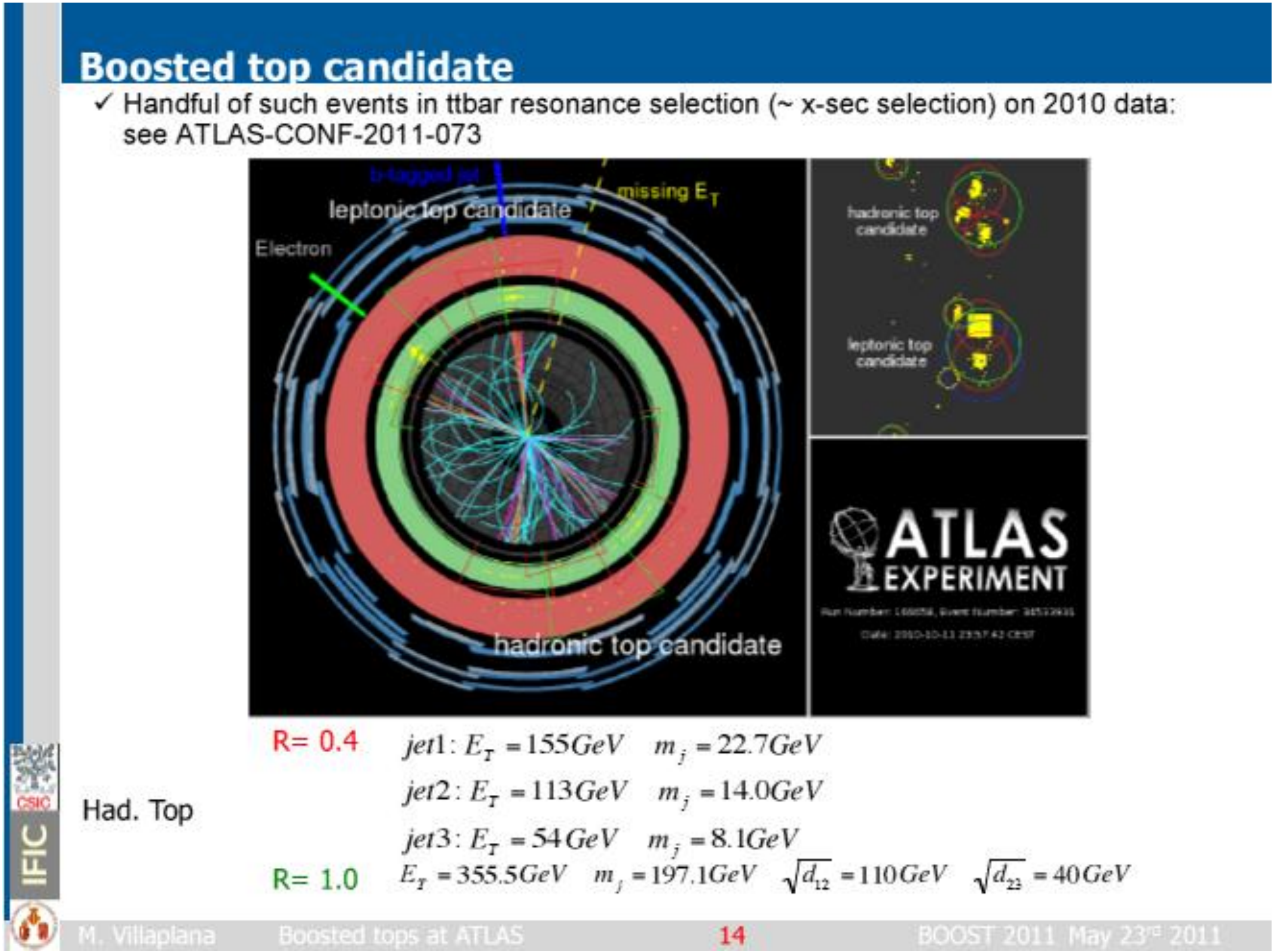
BEST SINGLE VARIABLE

BDT results

Single variable

BDTs led us to the variable, but with the variable we don't need BDTs

A Boosted Top Quark – from Miguel Villaplana



Boosted top candidate

- ✓ Handful of such events in ttbar resonance selection (~ x-sec selection) on 2010 data: see ATLAS-CONF-2011-073

Had. Top

R= 0.4 $jet1: E_T = 155 GeV \quad m_j = 22.7 GeV$

$jet2: E_T = 113 GeV \quad m_j = 14.0 GeV$

$jet3: E_T = 54 GeV \quad m_j = 8.1 GeV$

R= 1.0 $E_T = 355.5 GeV \quad m_j = 197.1 GeV \quad \sqrt{d_{12}} = 110 GeV \quad \sqrt{d_{23}} = 40 GeV$

M. Villaplana Boosted tops at ATLAS 14 BOOST 2011 May 23rd 2011

Summary:

- Exciting new theoretical ideas, tools and observables continue to appear!

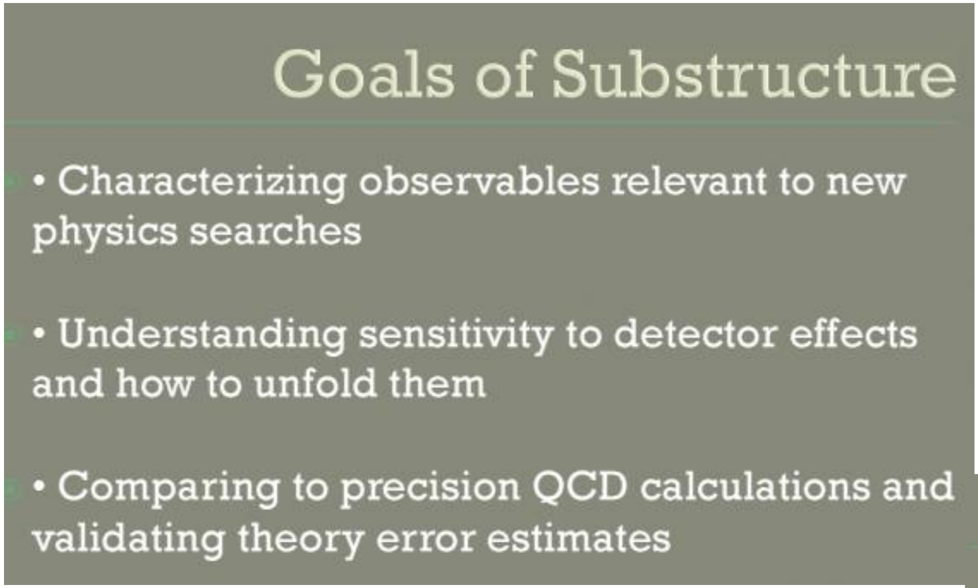
- The certification of these ideas/tools is occurring in real time – maybe public data based results on efficacy later this summer

- Still many challenges to determine and understand the uncertainties for jet substructure, both the theoretical and experimental –

Sum those logs
Tune those Monte Carlos
Understand those detectors
…….

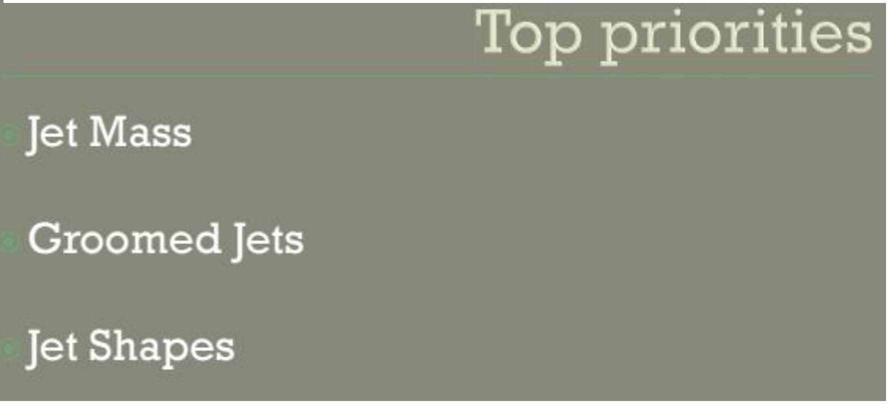
On to BOOST 2012 in Valencia

- See next talk by David Miller & the 2011 Working Group for priorities -



Goals of Substructure

- Characterizing observables relevant to new physics searches
- Understanding sensitivity to detector effects and how to unfold them
- Comparing to precision QCD calculations and validating theory error estimates



Top priorities

- Jet Mass
- Groomed Jets
- Jet Shapes

A Final Highlight – Rick and Steve as the LHC's answer to Laurel and Hardy