Qjets

U. Washington: Steve Ellis, Tuhin Roy

Harvard: David Krohn, Matt Schwartz

arXiv:1201.1914
(PRL 108 (2012) 182003)
& work(s) in progress

Andrew Hornig
U. Washington
Particle, Field, and String Theory Group

BOOST 2012
July 26, 2012

Overview of Talk

- (BRIEF!) Review of Standard Approach to Substructure
- Basics of Qjets approach, application to W jets
- 2 main aspects of Qjets:
  - 1: statistical improvement (non-Poissonian)
  - 2: new types of jet variables (example: “Volatility”)
- towards theoretical improvements (e.g., resummation of Qjet observables, “Qthrust”)

Standard Jet Substructure Technique

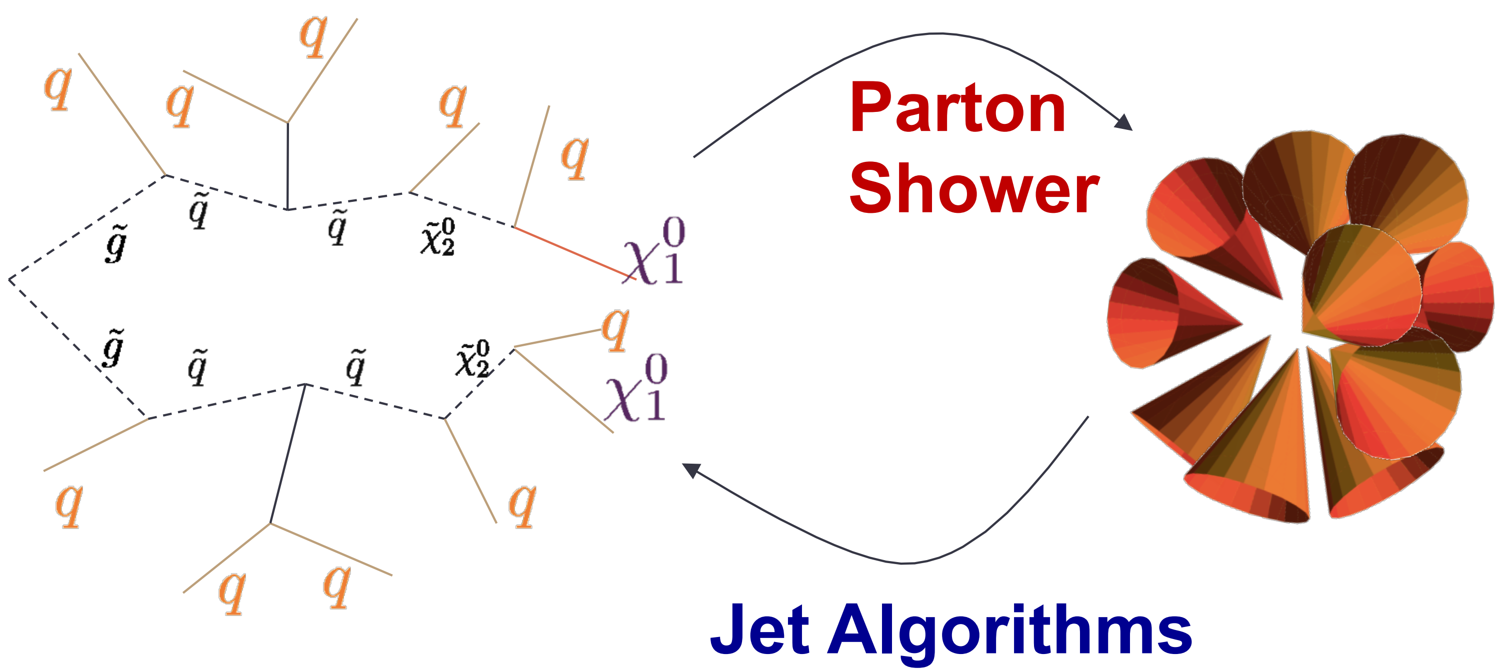
- from Jet (particles or constituents) get Tree
- done using algorithm to find the “best” tree (e.g., CA or kT)
- see if tree has structure
- examples:
  - BDRS (mass-drop + filtering)
  - Grooming (pruning & trimming)
  - top-tagging (JHU, HEP, ...)
  - N-jettiness

Standard Jet Substructure Technique

- from Jet (particles or constituents) get Tree
- done using algorithm to find the “best” tree (e.g., CA or kT)
- see if tree has structure
- examples:
    - BDRS (mass-drop + filtering)
    - Grooming (pruning & trimming)
    - top-tagging (JHU, HEP, ...)
    - N-jettiness

Basics of Qjets

- substructure assumes a shower creates trees, and best tree is good enough



- not *really* one-to-one, invertible
- "structure" can be highly dependent on which tree you take, *especially* for QCD

Basics of Qjets

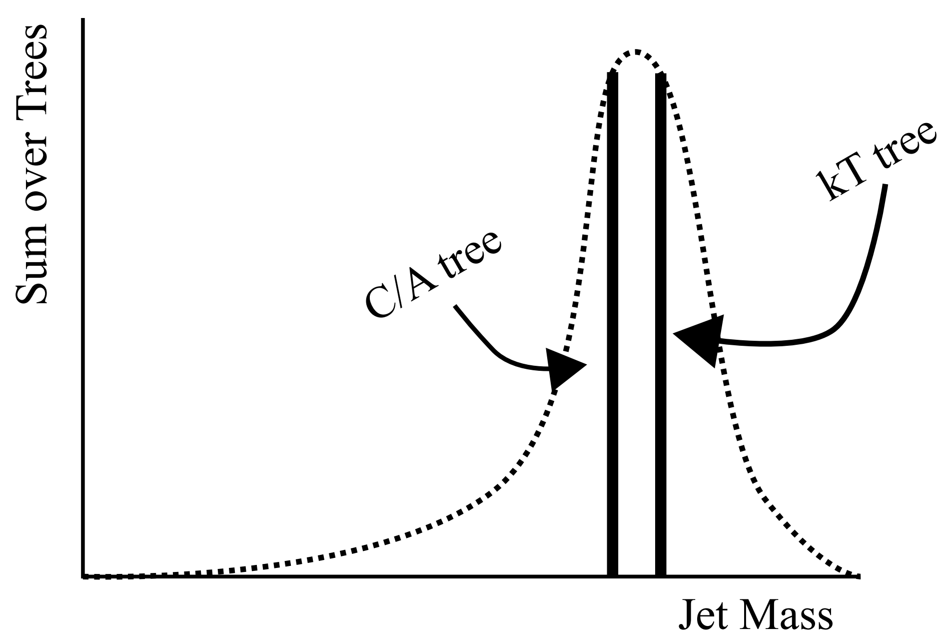
- substructure assumes a shower creates trees, and best tree is good enough

- however, even if we knew “best” tree, many other options (showering itself is a random/markovian process), and interference + UE contamination complicates this even more....

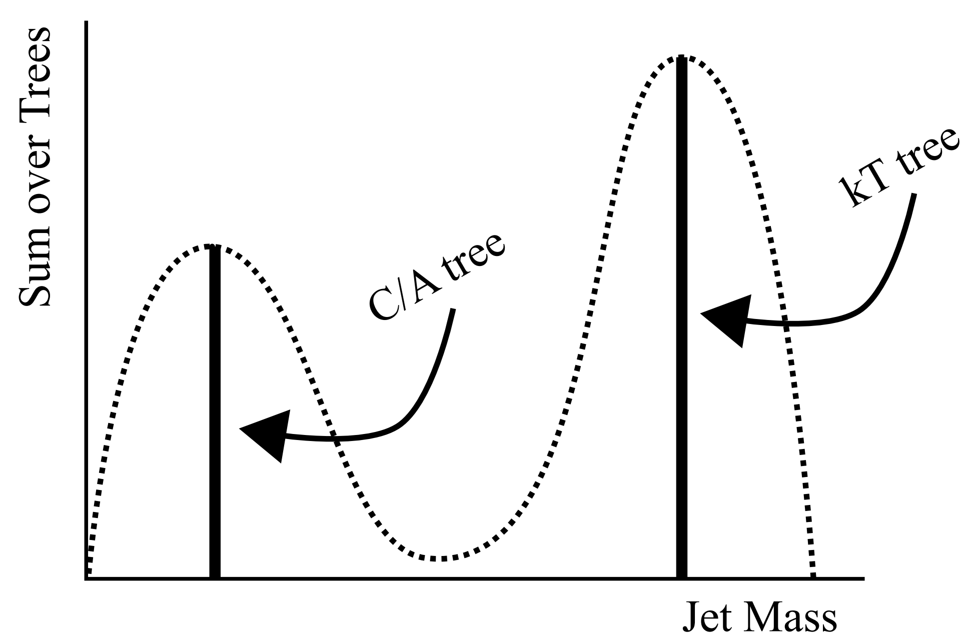
- “structure” can be highly dependent on which tree you take, especially for QCD

Basics of Qjets

- substructure assumes a shower creates trees, and best tree is good enough
- Qjets: take all (or many trees)
- example: apply pruning to the various recombinations allowed within a single jet



VS



Qjets in practice

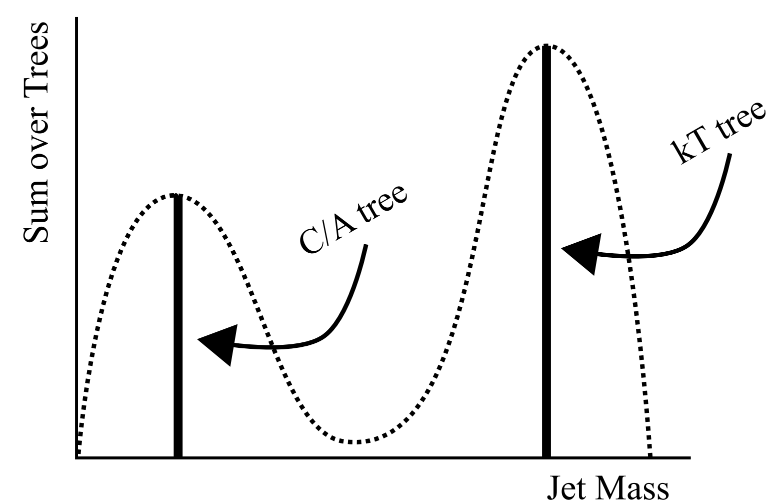
Fastjet Plugin: http://jets.physics.harvard.edu/Qjets

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

$$\omega_{ij}^{(\alpha)} \equiv \exp\left\{-\alpha \frac{(d_{ij} - d^{\min})}{d^{\min}}\right\}$$ where $$d_{ij} = \begin{cases} d_{\mathrm{k_T}} \equiv \min\{p_{Ti}^2, p_{Tj}^2\}\Delta R_{ij}^2 \\ d_{\mathrm{C/A}} \equiv \Delta R_{ij}^2 \end{cases}$$

  - this gives a tree, on which any
- results in a *distribution* for each jet
- (typically) stable after ~100 runs

  (and 100 << 10! to 20!)

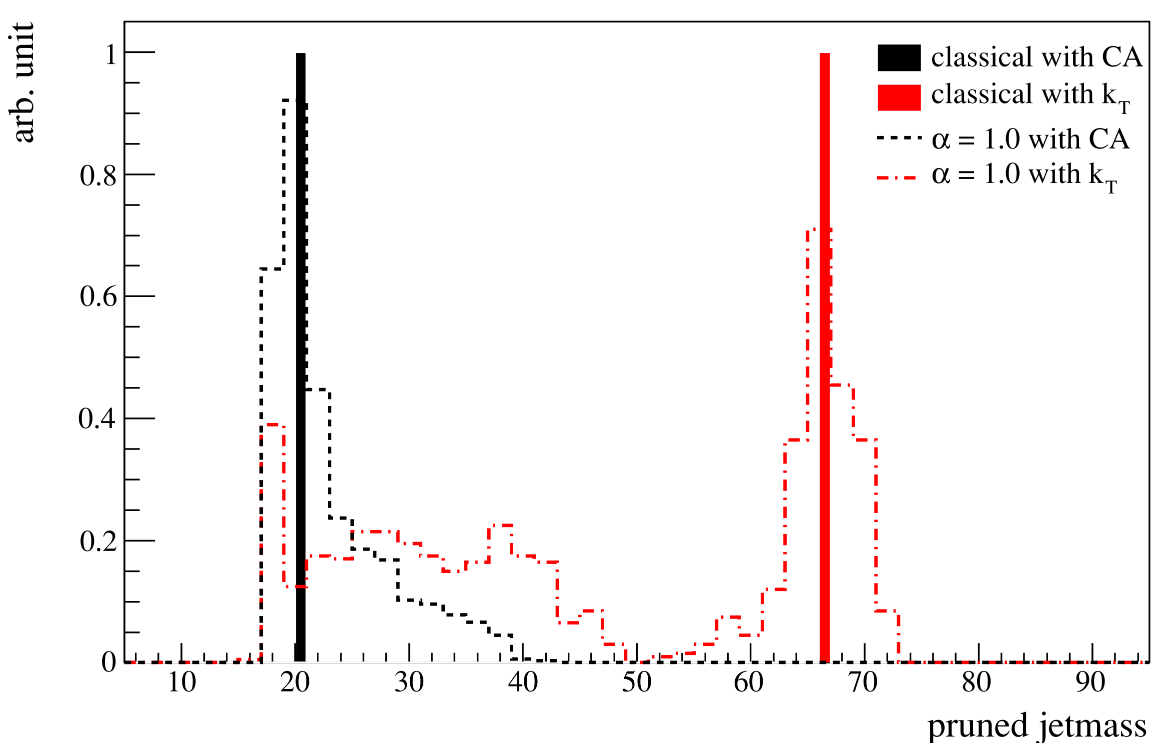


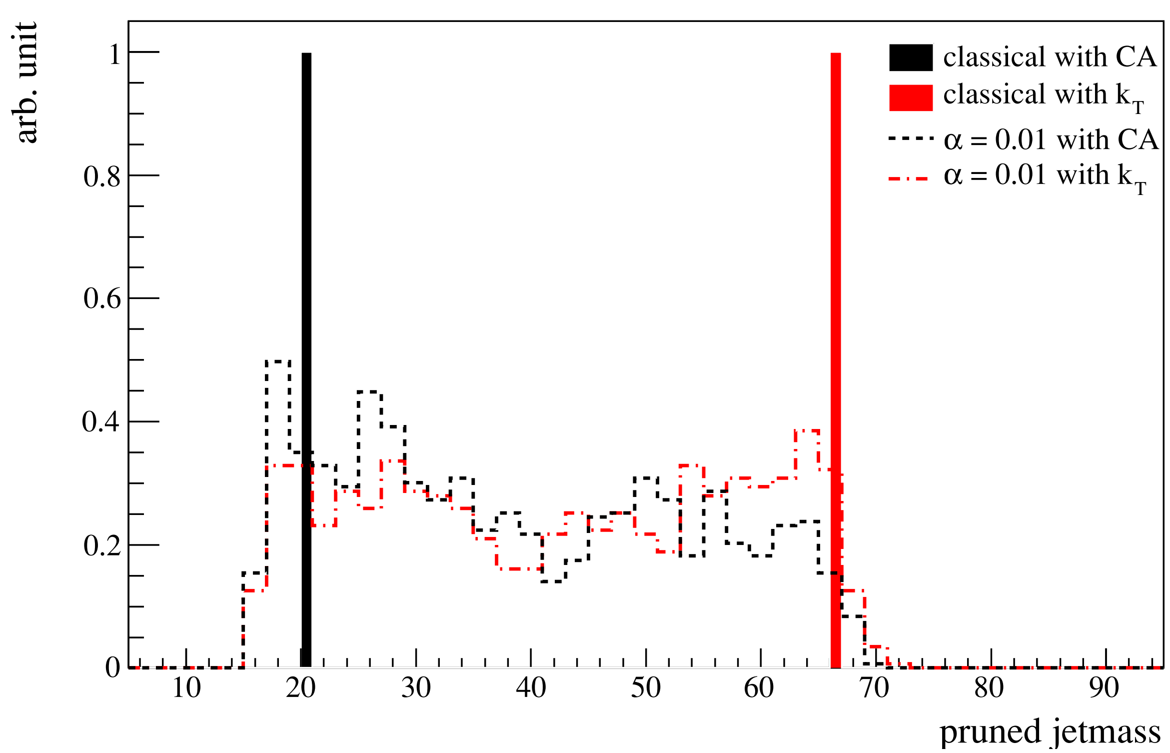
Qjets in Practice: our (ad hoc) metric

- α ≡ "ridigity":
  - α → ∞, exact CA (kT)
  - α → 0, all combos equal

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

- CA and kT are "close" for small enough α:





Application I: Statistics

- Classical:
- assumptions:
  - 1) production is Poisson: $P_N(n) \equiv \frac{e^{-N} N^n}{n!}$
  - 2) if 1 event has prob. $\epsilon_{\rm cl}$ of being tagged (“tagging efficiency”)

  ⇒ tagging (for fixed #n) is binomial: $B_\epsilon(n;r) \equiv {}_nC_r \epsilon^r (1-\epsilon)^{n-r}$

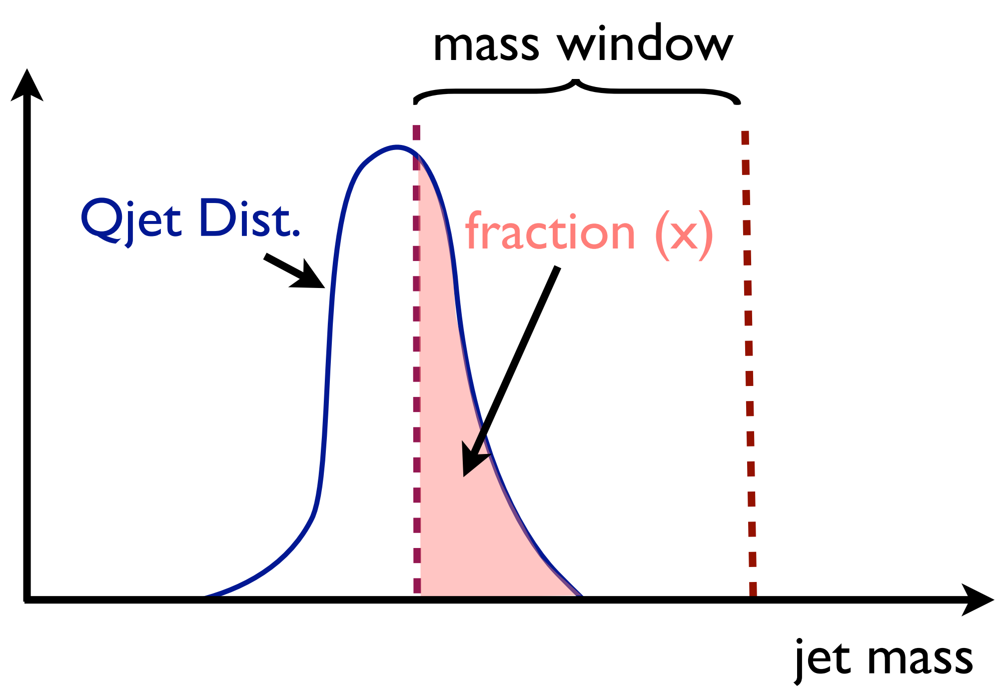
  ⇒ tagging (for any n) is also Poisson:

$$F_{\epsilon,N}(r) \equiv \sum_{n=r}^{\infty} F_{\epsilon,N}(r|n) = \frac{e^{-N\epsilon} N^r \epsilon^r}{r!} \equiv P_{N\epsilon}(r)$$

$$\frac{\delta\sigma_{\rm cl}}{\sigma_{\rm cl}} = \frac{1}{\sqrt{N\epsilon_{\rm cl}}}$$ and $$\frac{\delta\sigma_{cl}^2}{\sigma_c} = 1$$

Application I: Statistics

- Qjets: distributions have an overlap ($\in[0,1]$), not binomial!
  - 1) production is Poisson
  - ~~2) tagging (for fixed #n) is binomial~~
- tagging now a distribution $f_1(x)$:



Application I: Statistics

- Qjets: distributions have an overlap (∈[0,1]), not binomial!
  - ✓ 1) production is Poisson
  - ✗ ~~2) tagging (for fixed #n) is binomial~~
  - tagging now a distribution $f_1(x)$

$$\epsilon_Q = \langle x \rangle_{f_1} \qquad \sigma_1^2 = \langle (x - \bar{x})^2 \rangle_{f_1}$$

- upshot:

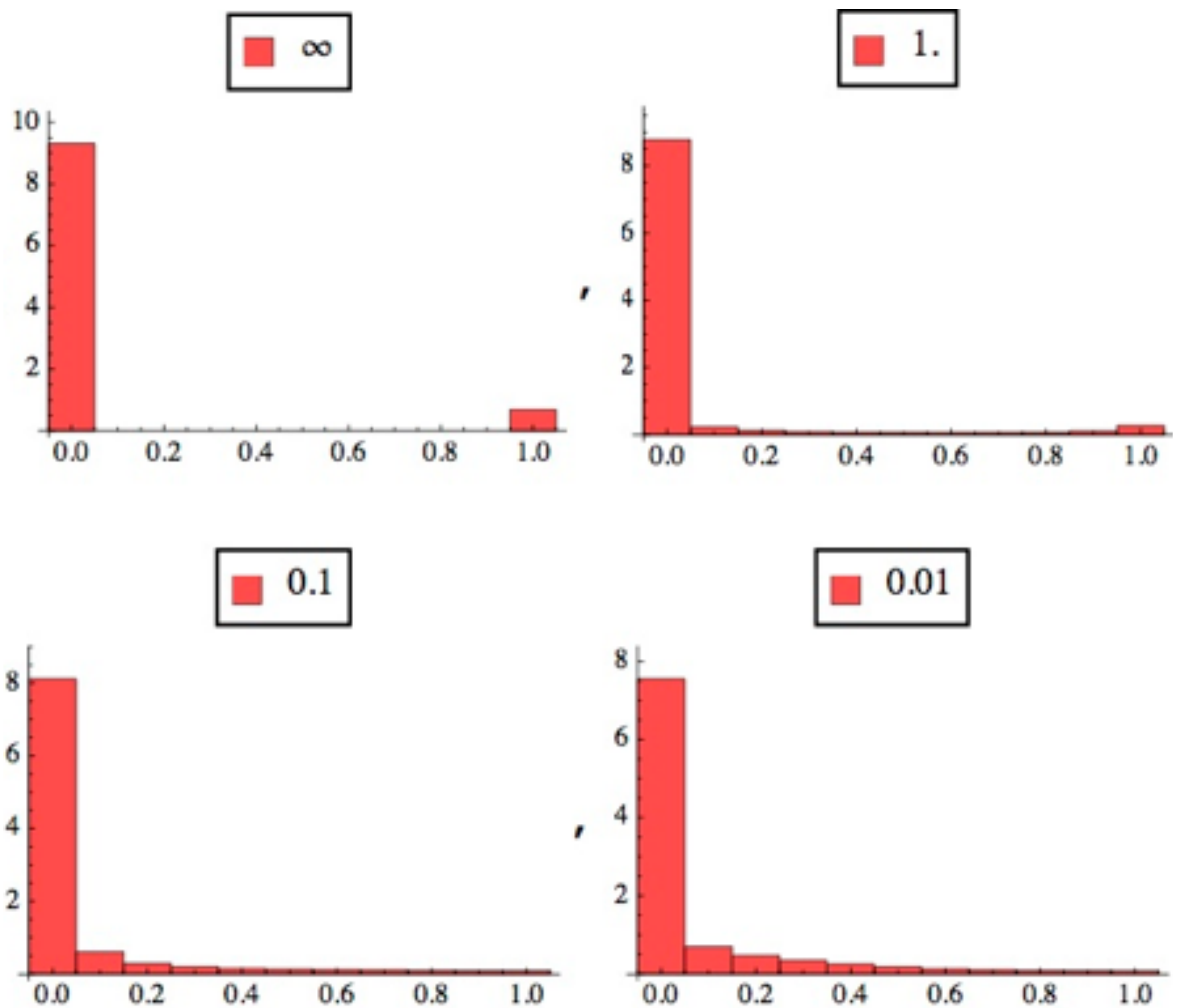
$$\frac{\delta\sigma_Q}{\sigma_Q} = \sqrt{\frac{1 + (\sigma_1/\epsilon_Q)^2}{N}} \quad \text{VS} \quad \frac{\delta\sigma_{\rm cl}}{\sigma_{\rm cl}} = \frac{1}{\sqrt{N\epsilon_{\rm cl}}}$$

and

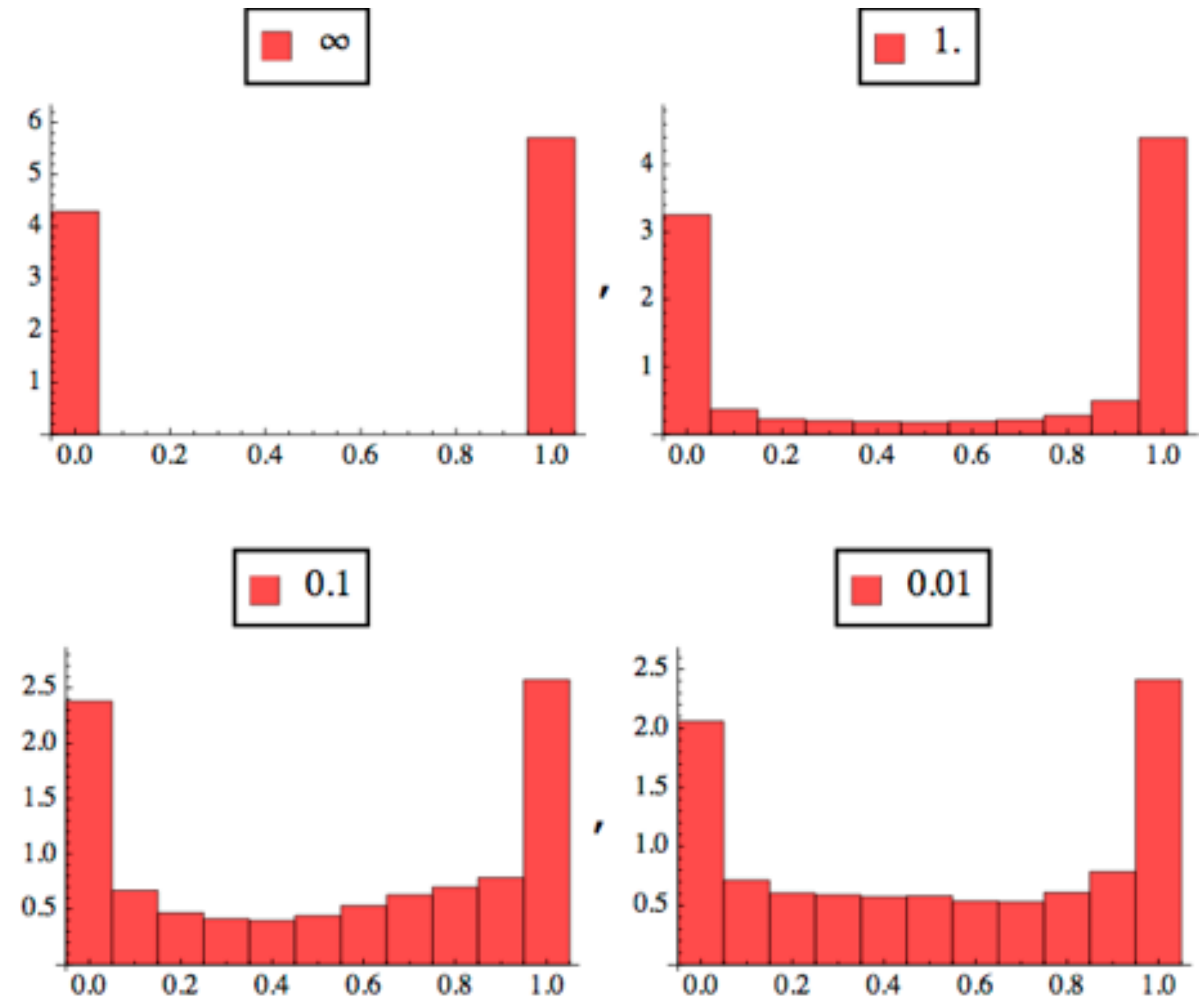
$$\frac{\delta\sigma_Q^2}{\sigma_Q} = \frac{\langle x^2 \rangle}{\langle x \rangle} = \epsilon_Q + \frac{\sigma_1^2}{\epsilon_Q} \quad \text{VS} \quad \frac{\delta\sigma_{cl}^2}{\sigma_c} = 1$$

$f_1(x)$ function for the α-weight

- background (QCD):



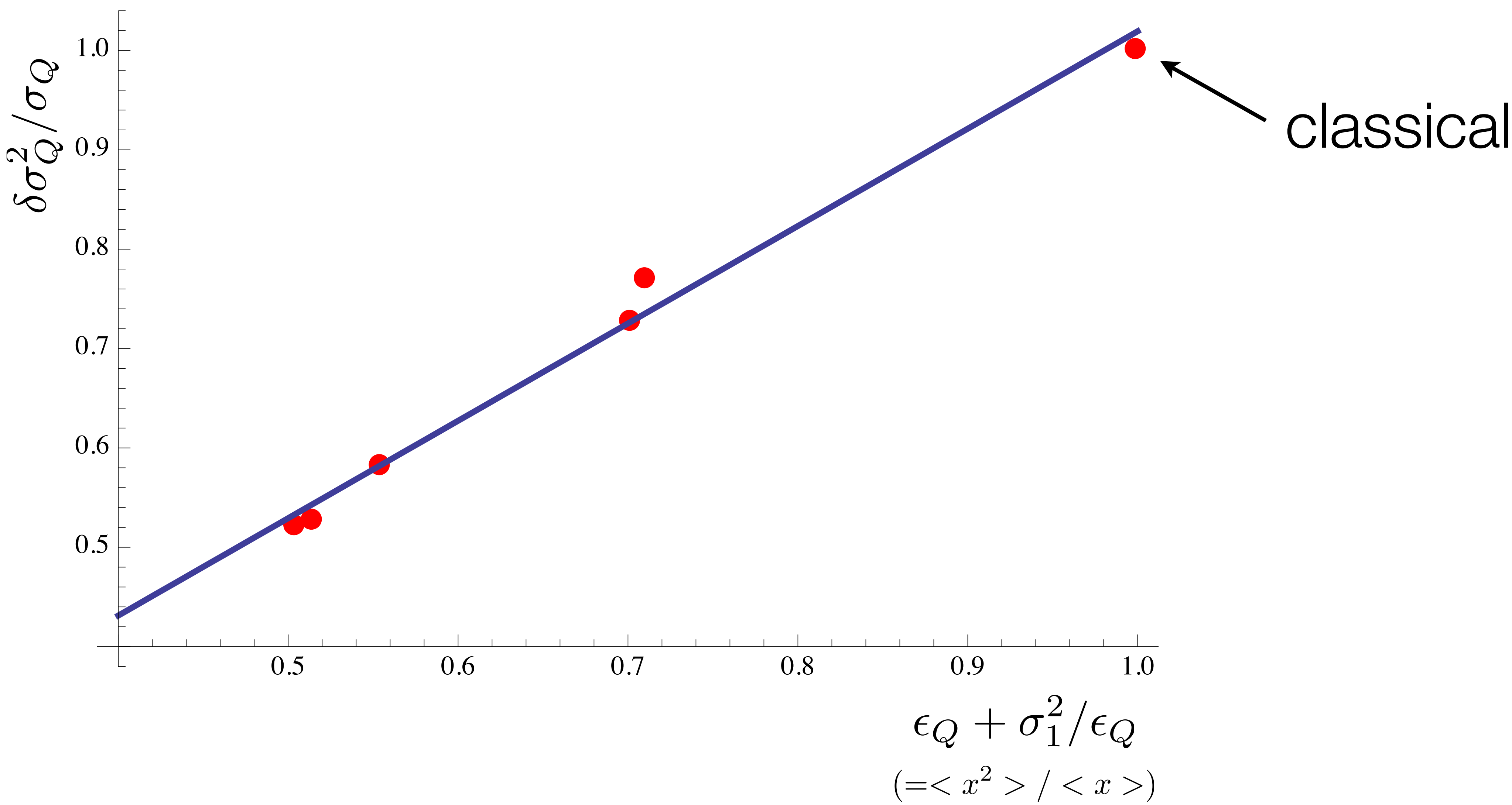
- signal (W):



→ $\langle x^2\rangle/\langle x\rangle$ decr. w/ α (away from “classical”)

Application I: Statistics: W-jet Example

- reminder: $$\frac{\delta\sigma_Q^2}{\sigma_Q} = \frac{< x^2 >}{< x >} = \epsilon_Q + \frac{\sigma_1^2}{\epsilon_Q}$$



Application I: Statistics: W-jet Example

- Signal = boosted W-jets, pT > 500
- BG = light QCD jets, pT > 500
- Measure the signal size in a bin (here 70-90 GeV) and compare it to the size of the BG fluctuations (Poisson stats included)

Algorithm	Mass uncertainty $\delta\langle m\rangle$	Relative Luminosity required
k_T ("classical" pruning)	3.15 GeV	1.00
Qjets α=0	2.20 GeV	0.50
Qjets α=0.001	2.04 GeV	0.45

$\Rightarrow$ ~ ***factor of 2 in luminosity*** needed for given significance

Application II: New Observables

- since Qjets gives distributions for **each jet**, can now cut on these distributions (or more complicated analysis)
- example: "Volatility," a measure of how ambiguous/"fuzzy" jets are

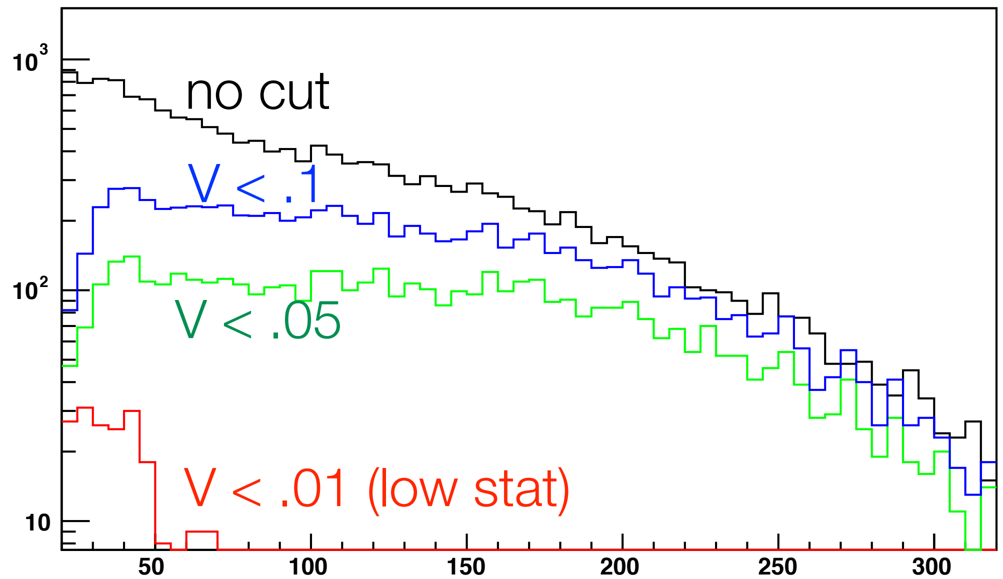
$$\mathcal{V} = \Gamma/\langle m \rangle \qquad \Gamma \equiv \sqrt{\langle m^2 \rangle - \langle m \rangle^2}$$

- QCD jets often have ambiguities, making V larger
- ambiguity larger for smaller m/p_T ($m \sim p_T$ QCD jets have "real" structure)
- Note: Poisson stats for $\mathcal{V}$-cut jets (when $m_{window} >> m_{cut}$)

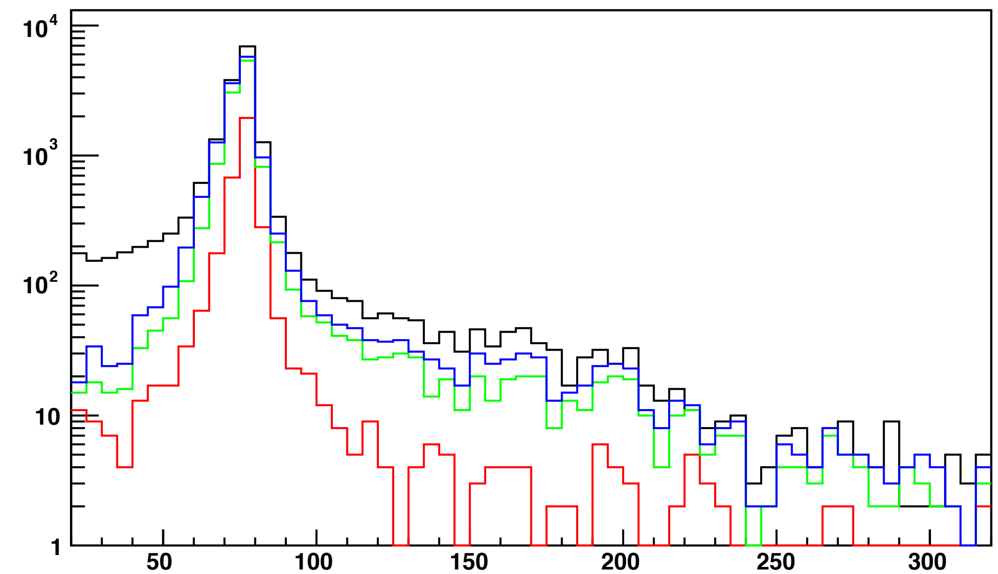
Application II: New Observables

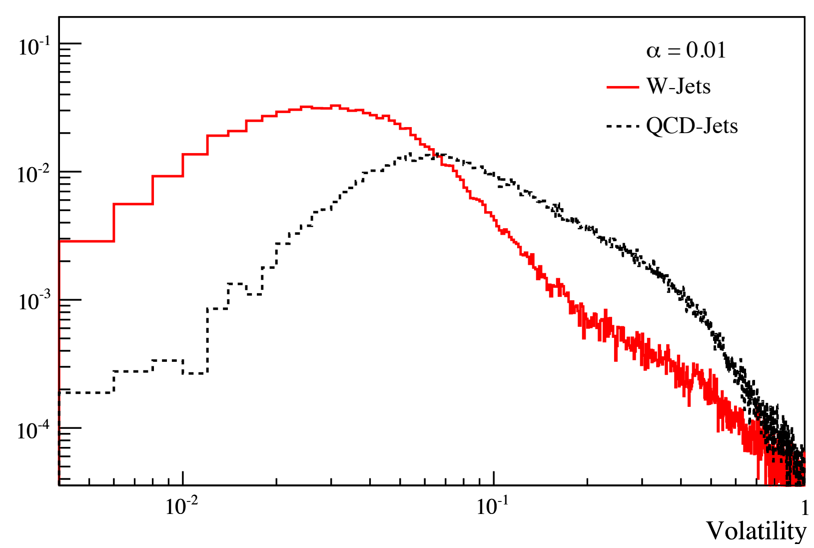
- Volatility for example of pruned jets $\mathcal{V} = \Gamma/\langle m \rangle \qquad \Gamma \equiv \sqrt{\langle m^2 \rangle - \langle m \rangle^2}$

QCD:

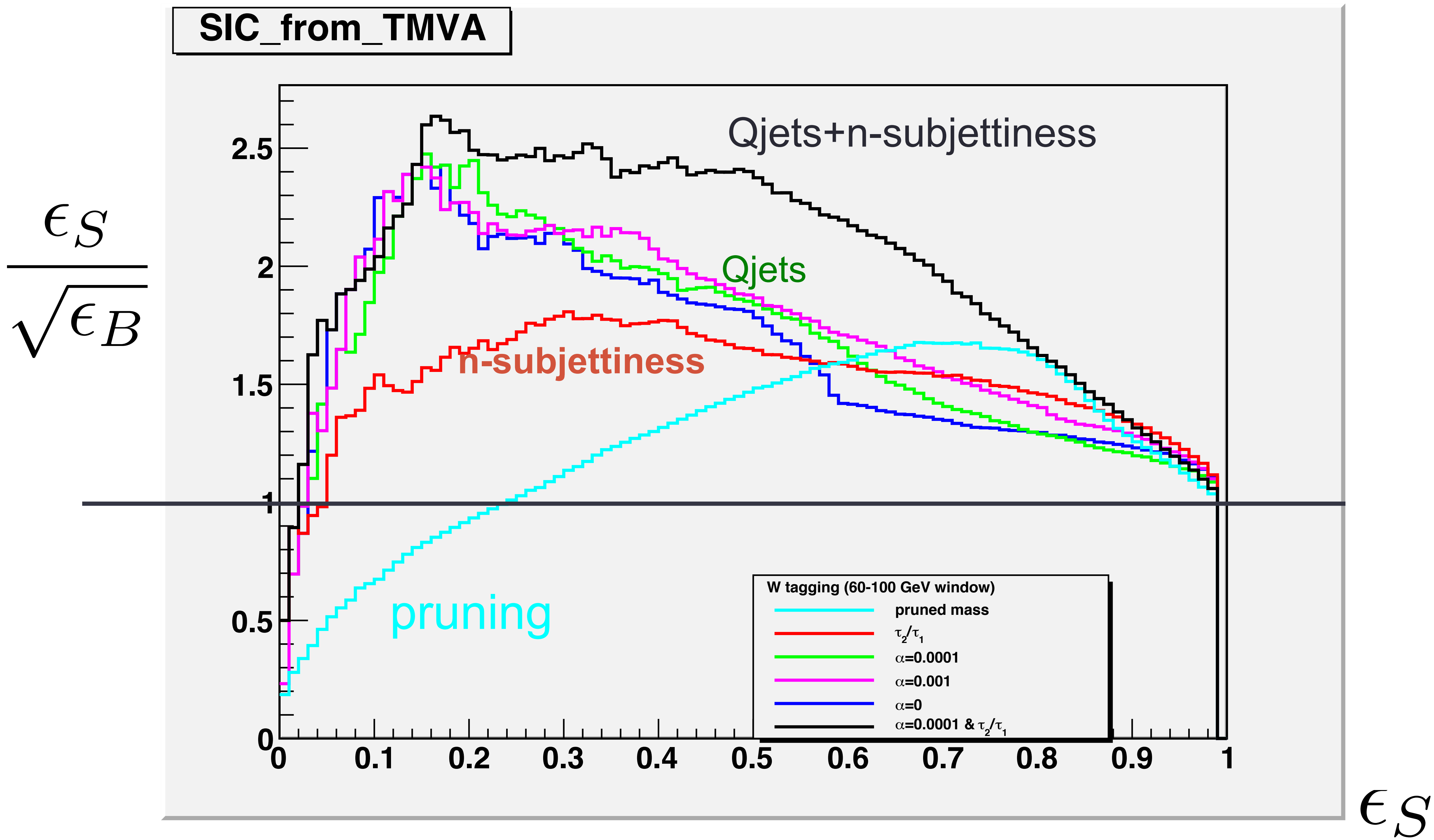


Signal (W):





Volatility vs. N-Subjettiness vs. Combined



Other Weights

- binomial CA-kT: choose CA or kT at each clustering
- generalized kT (q-axis, q=1 is kT, q=0 is CA)
- q-axis -> p,q plane (p=q=1 is JADE)

1. doesn't span space of α weight
2. not nearly as efficient α weight (>> 100 Qjets per jet)

- sudakov/shower inspired weights

1. "no" free parameters
2. only QCD radiation (no separate QCD/signal weights for Qjets, unlike Template & Shower Deconstruction)

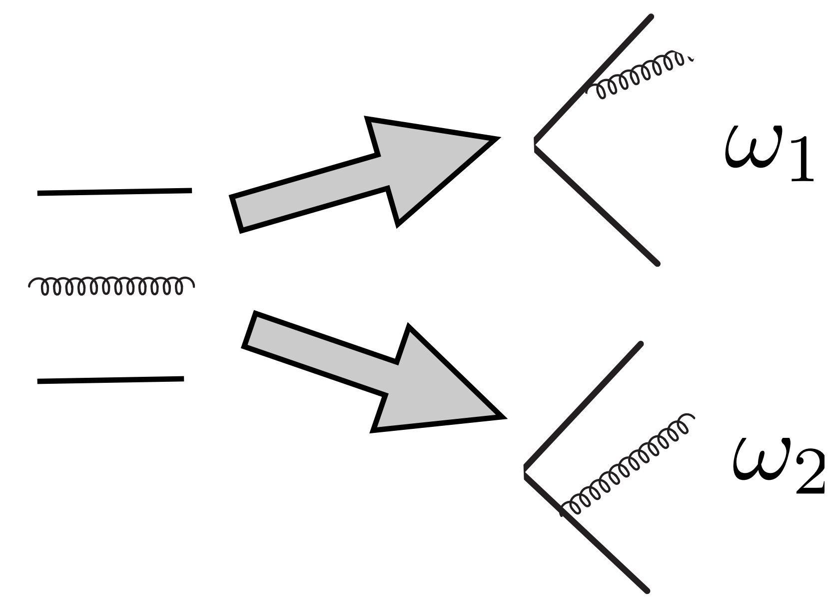
Going forward

- work in progress:
  - (width and var. of) $f_1(x)$ from 1st principles (QCD, SCET, ...)
  - resummation of Qjet obs. (e.g. Qthrust) (S. Ellis, AH, M. Schwartz, in progress)
  - Qanti-kT *events* (D. Krohn, D. Kahawala, M. Schwartz, in progress)
  - top tagging, new-physics searches/measurements, etc

A Simple Qjet Observable Example: “Qthrust”

- Simple example (that factorizes): “Qthrust”
- normal thrust (for $e^+e^- \rightarrow$ 3 partons) :

$$\tau = \frac{1}{Q} \min\{s, t, u\}$$



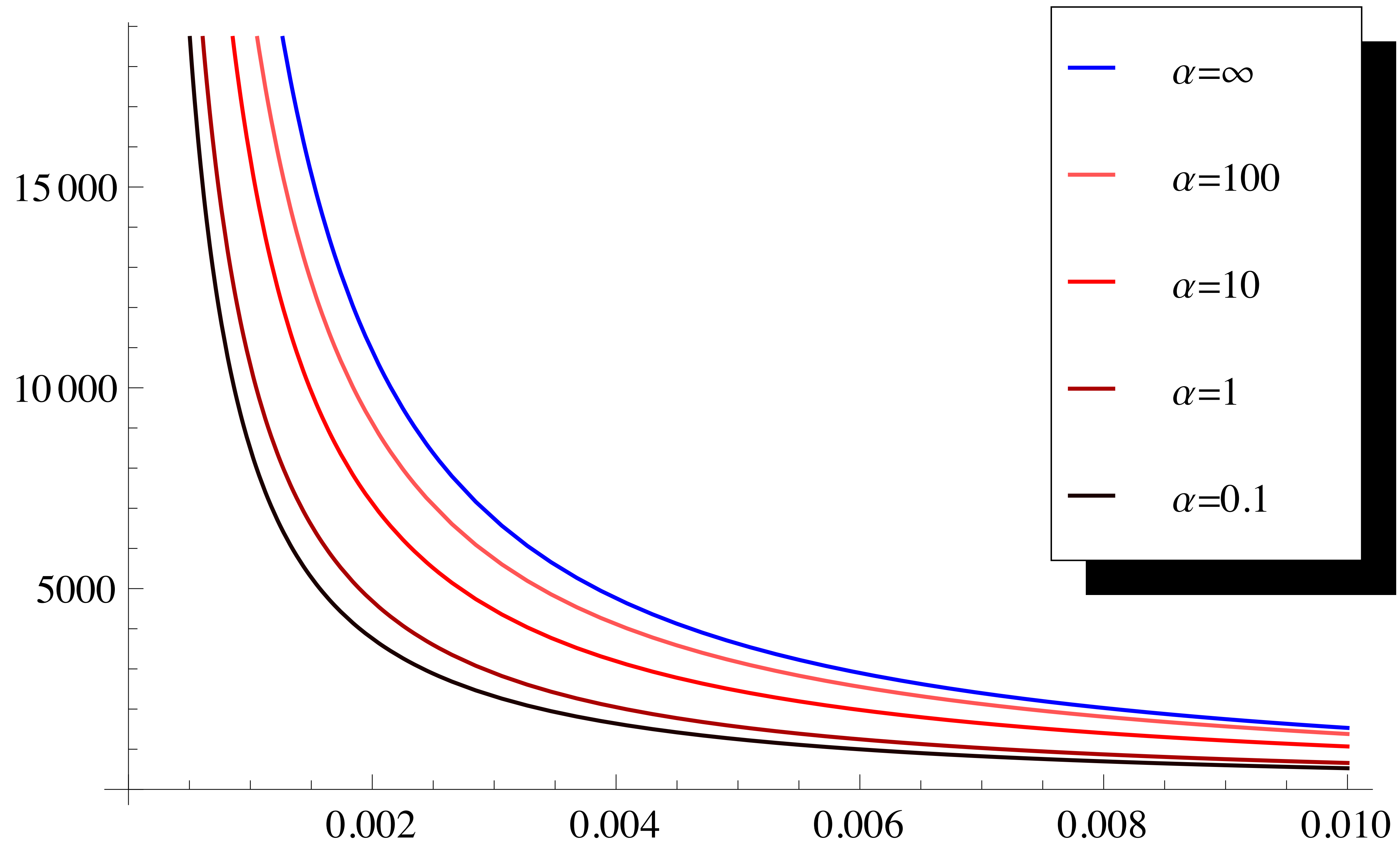
- probabilistic thrust :

$$\tau = \frac{1}{Q}(\omega_s s + \omega_t t + \omega_s s)$$ with $\omega_i \rightarrow 1$ as $i \rightarrow 0$

- example: the α weight...

Towards Calculating Volatility (V): Part I

- fixed-order results for $e^+e^- \to$ jets with weighted clusterings:



Summary

- basis of substructure: trees
- typically, best tree is chosen as CA or kT
- there is no “best” tree and should take many into account
- this improves by:
  - 1) reducing statistical uncertainty (less variability)
  - 2) giving distributions for each jet → new observables
- hopefully, its clear that we’ve only begun to scratch the surface of potential applications....