Workshop on top physics at the LC
5-6 March 2014 LPNHE Paris

A new jet reconstruction algorithm for lepton colliders

I. García*, M. Vos
IFIC (CSIC/UVEG)
Valencia, Spain

With help from Gavin Salam (CERN)
And Roman Poeschl and François Richard (LAL)

Introduction

Jet algorithms must be:

- IR and collinear safe
  - observables are insensitive to soft or collinear emission
- Simple to use in experiment and calculations
  - describe in a few lines
  - FastJet implementation
- Subject to small hadronization corrections
  - Cambridge/Aachen at LEP
- **Future high-energy lepton colliders** present an environment that **differs** in several important respects from that encountered **at the Z-pole**
  Do we need to rethink jet reconstruction? which algorithms are most suitable?

A brief history of sequential recombination algorithms

JADE 1980s

$$y_{ij} = \frac{E_i^2, E_j^2}{Q^2}(1 - \cos\theta_{ij})$$

Experience on e^+e^- data at Z-pole

Durham or e^+e^- k_t algorithm (LEP and SLC)

$$d_{ij} = 2min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$$

Adapt to hadron colliders

$$d_{ij} = min(p_{Ti}^{2n}, p_{Tj}^{2n})\Delta R_{ij}^{2n}/R^{2n}$$
$$d_{iB} = p_{Ti}^{2n}$$

n=0: Cambridge-Aachen
n=1: Longitudinally invariant k_t
n=-1: Anti-k_t (LHC default)

Feed back into e^+e^- algorithms

Generalised e^+e^- k_t algorithm

$$d_{ij} = min(E_i^2, E_j^2)(1 - \cos\theta_{ij})/(1 - \cos R)$$
$$d_{iB} = E_i^2$$

Moretti, Lonblad, Sjostrand, JHEP9808 (1998)
Catani, Dokshitzer, Webber, Phys.Lett. B285 (1992)
Catani, Dokshitzer, Seymour, Webber, Nucl.Phys. B406 (1994)
Ellis, Soper, Phys.Rev. D48 (1993)
All algorithms available in FastJet

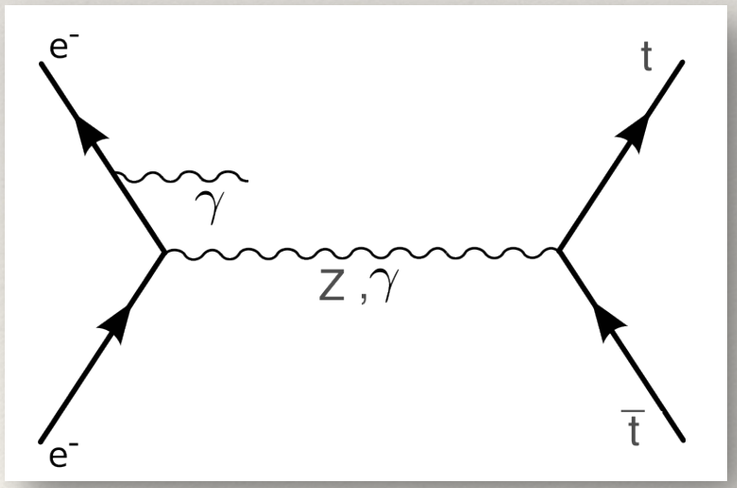
Time to rethink e^+e^- algorithms!!

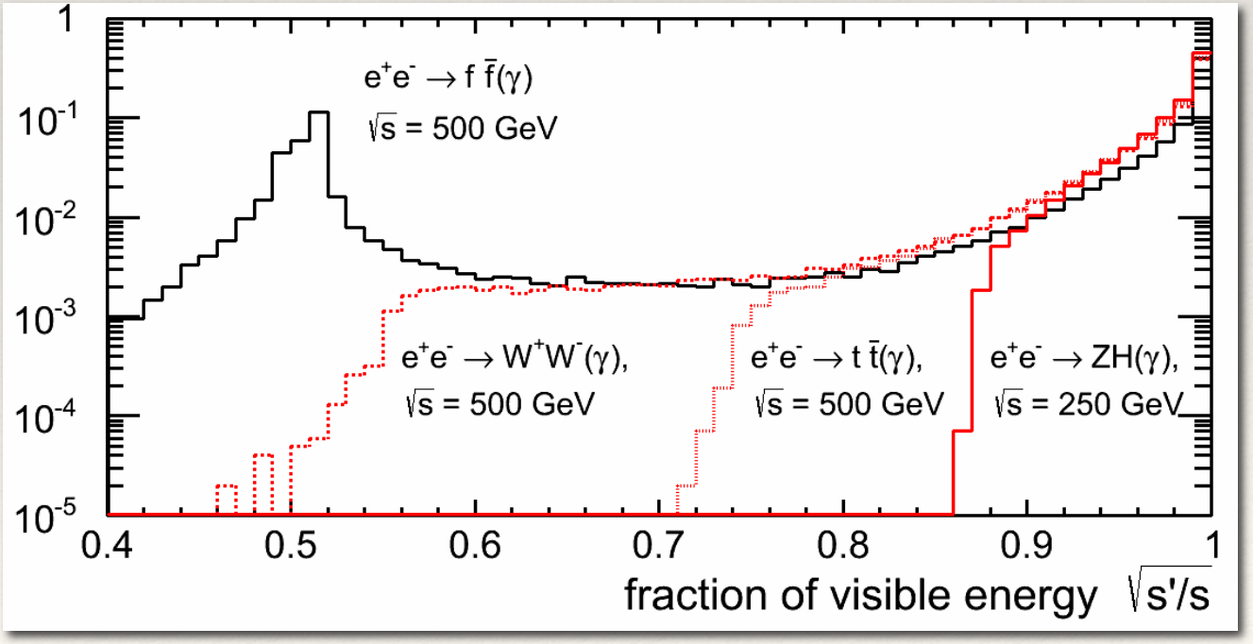
Boost invariance

- At **hadron colliders** **the partons** that participate in the hard process generally **carry different fractions of the initial** hadron **energy.**
- The final state acquires a substantial **Lorentz boost** along the beam axis.
  - LHC di-jets: $\beta_z \sim 1$
  - LHC tt: $\beta_z \sim 0.5$
- Replace the [**energy, polar angle**] basis by [**transverse momentum, rapidity**]

Boost invariance

- Photons emitted by the incoming beam particles (**Initial State Radiation**) can carry away a significant fractions of the nominal center-of-mass energy
- For $e^+e^- \rightarrow Z/\gamma^* \rightarrow f\bar{f}$ process, with $m_f < M_Z/2$ → large fraction of events tends to ***return*** to the **Z-pole**
- However for **most interesting processes** at a future lepton collider ISR plays a much **less important role**
  - At **lepton colliders ISR** leads to **a minor boost**
  - **The basis [E,θ]** is the most natural choice





Background levels at future LC

- The **pile-up** at the **LHC** is a serious challenge that has led to a large body of work on mitigation and correction methods
- **LEP** or **SLC presented effectively negligible background**
- The γγ—> **hadrons** background at **CLIC** has **strong** impact on jet reconstruction performance [CLIC CDR, Marshall & Thomson, arXiv:1308.4537]
- Less pronounced, but **non-negligible** impact on ILC physics [many studies, arXiv:1307.8102]
- Using hadron collider algorithms can reduce these problems [CLIC CDR]

The Valencia jet algorithm

A new clustering jet reconstruction algorithm that combines the good features of lepton collider algorithms, in particular the **Durham-like distance criterion;**

$$d_{ij} = min(E_i^{2\beta}, E_j^{2\beta})(1 - \cos\theta_{ij})/R^2$$

with the **robustness against background of** the longitudinally invariant **k_t algorithm**

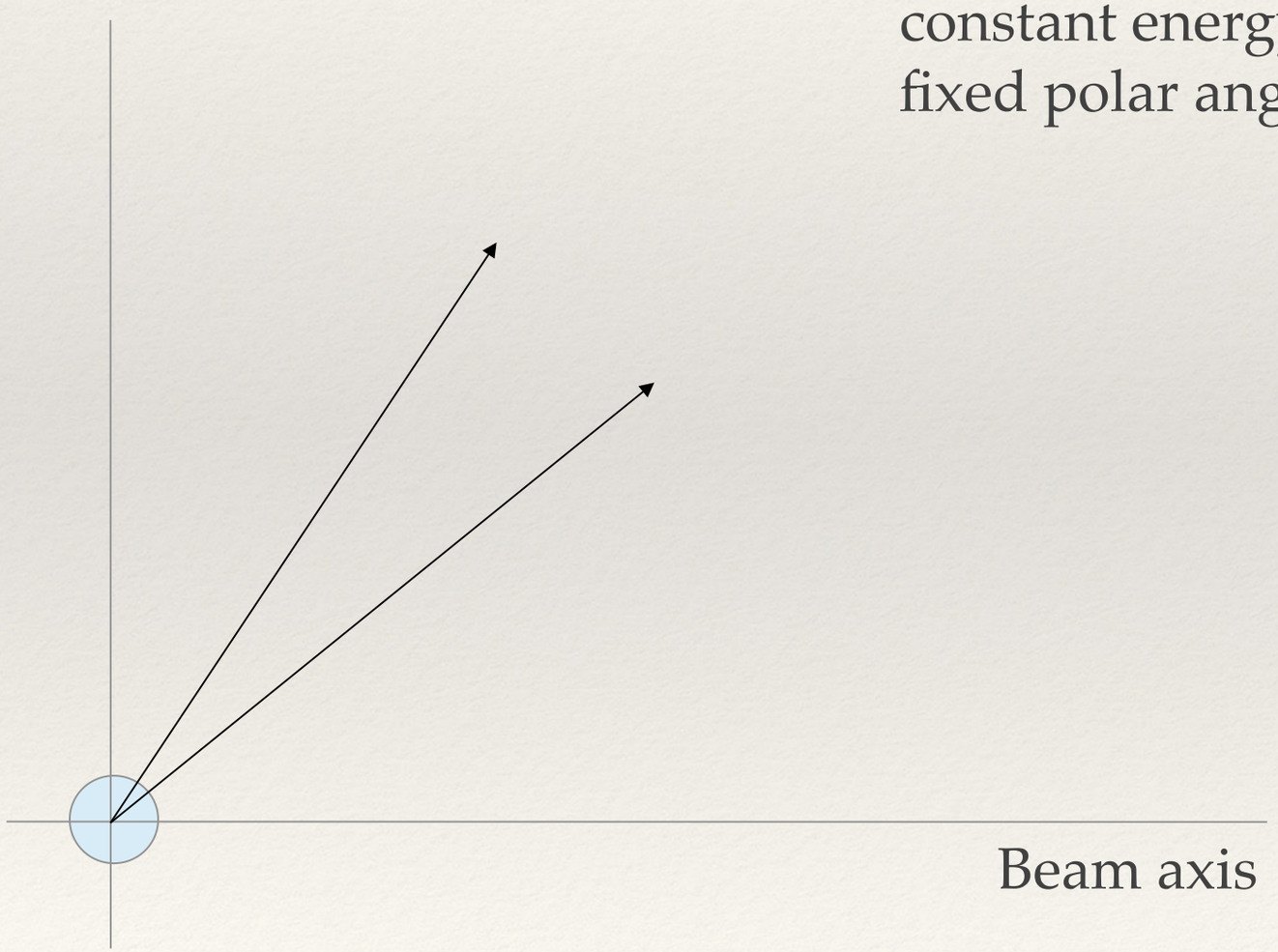
$$d_{iB} = p_T^{2\beta}$$

The exponent β allows to ***tune*** **the background rejection** level

The algorithm has been implemented as a plugin for the *FastJet* package and will be made available in the fjcontrib area

Comparison of the distance criteria

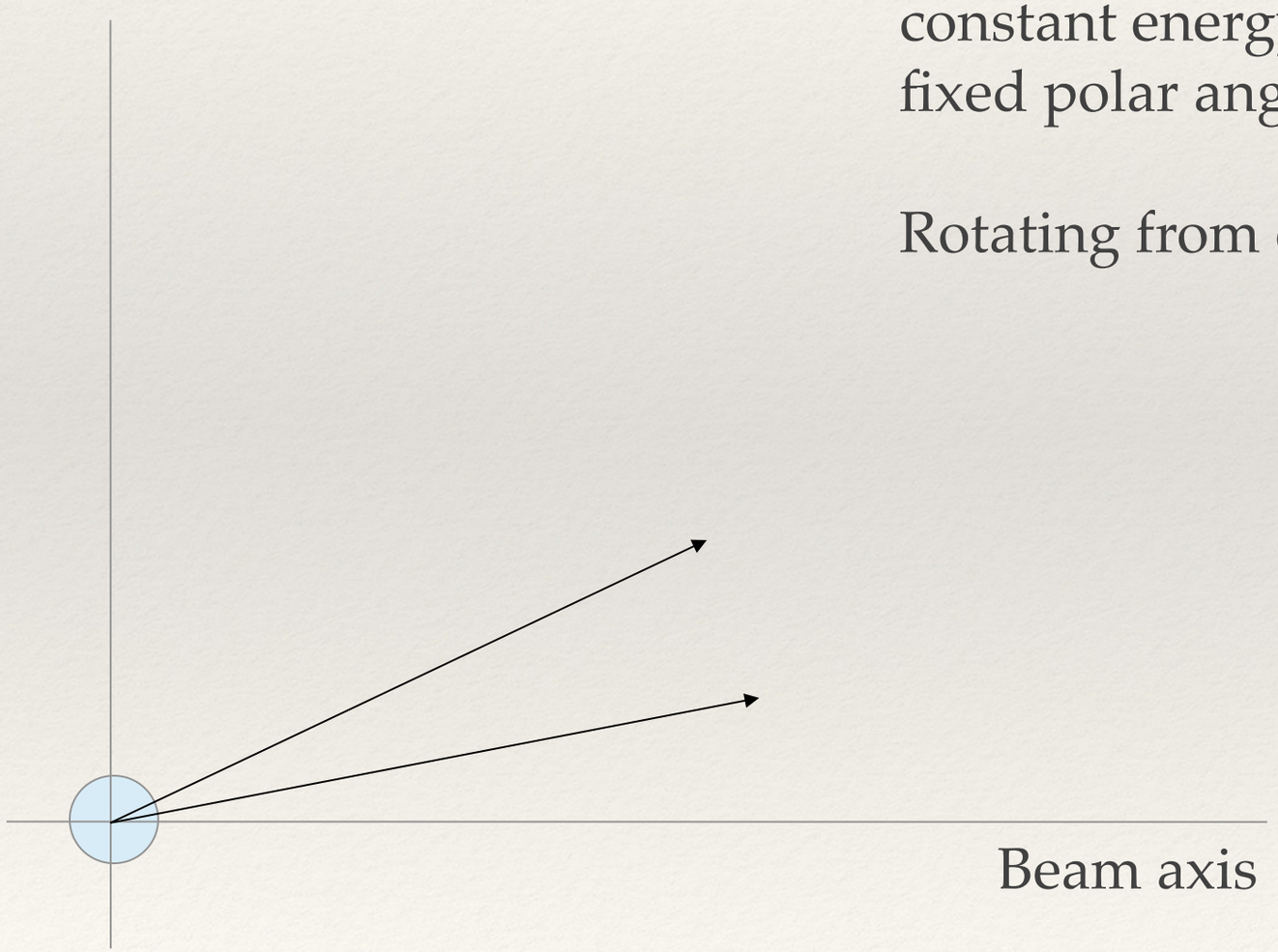
Two test particles with
constant energy (E = 1 GeV) and
fixed polar angle separation (100 mrad)



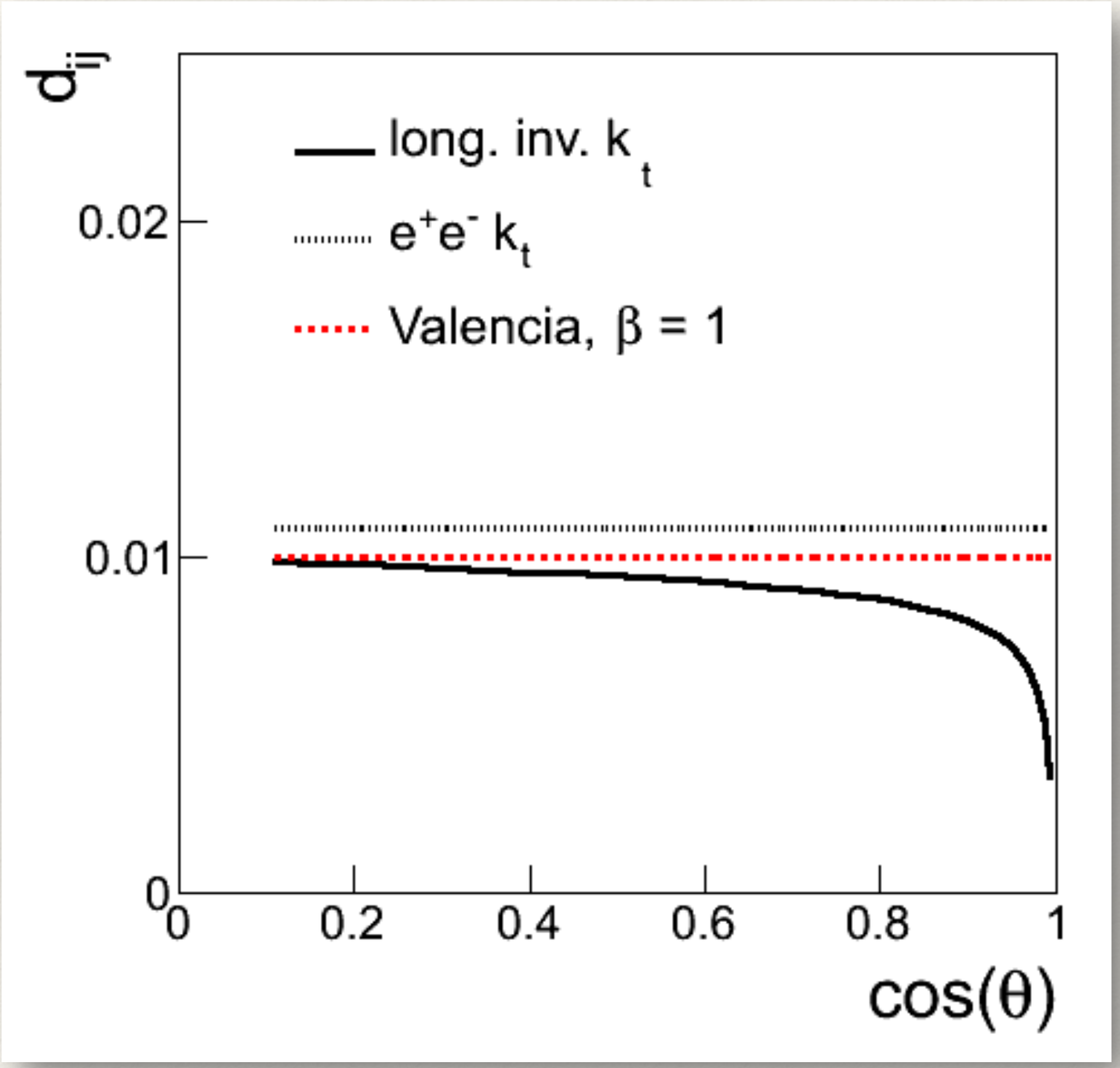
Comparison of the distance criteria

Two test particles with
constant energy (E = 1 GeV) and
fixed polar angle separation (100 mrad)

Rotating from central to forward region



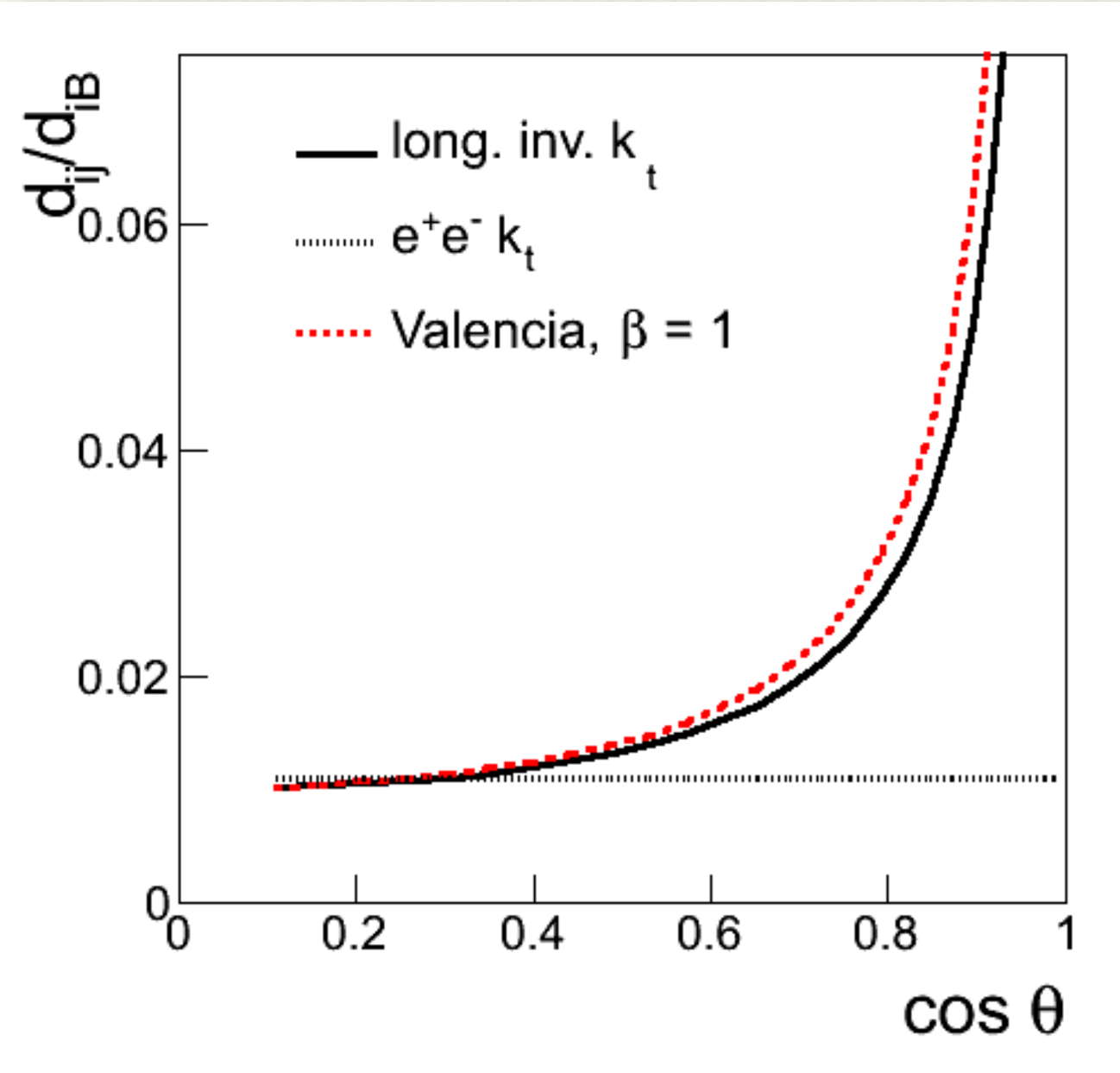
Comparison of the distance criteria



As the two-particle system rotates into the forward region, the distance d_{ij} of longitudinally invariant k_t decreases ($\Delta\eta$ increases, p_T decreases faster)

Traditional e^+e^- algorithms and Valencia have constant d_{ij}

Comparison of the distance criteria



The ratio of the inter-particle distance and the beam distance:

d_{ij}/d_{iB}

drives the robustness to (forward) background: the decision to assign the particle to final-state or beam jets depends on this ratio (and R)

Long. inv. k_t's robustness is indeed due to its increasing d_{ij}/d_{iB} ratio

Valencia with β=1 is similar (by design) to long. inv. k_t

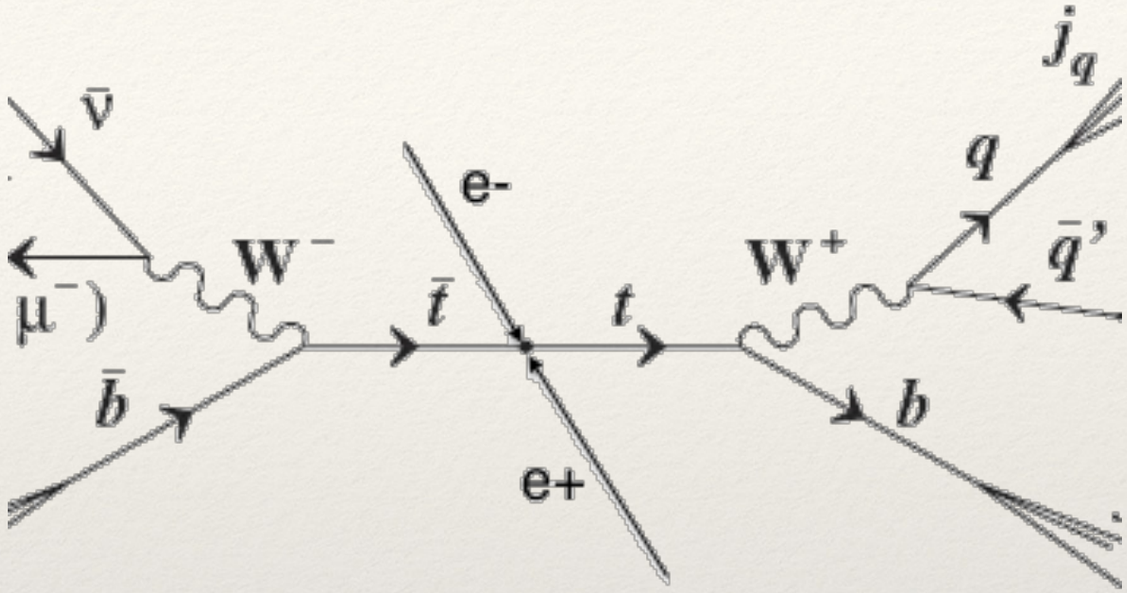
Jet reconstruction performance

IFIC/LAL study of lepton+jets tt @ 500 GeV, [arXiv:1307.8102]

Event Generation
Whizard 1.95

Reconstruct Particle Flow objects using PANDORA

Reconstruct jets (exclusive, n=4)

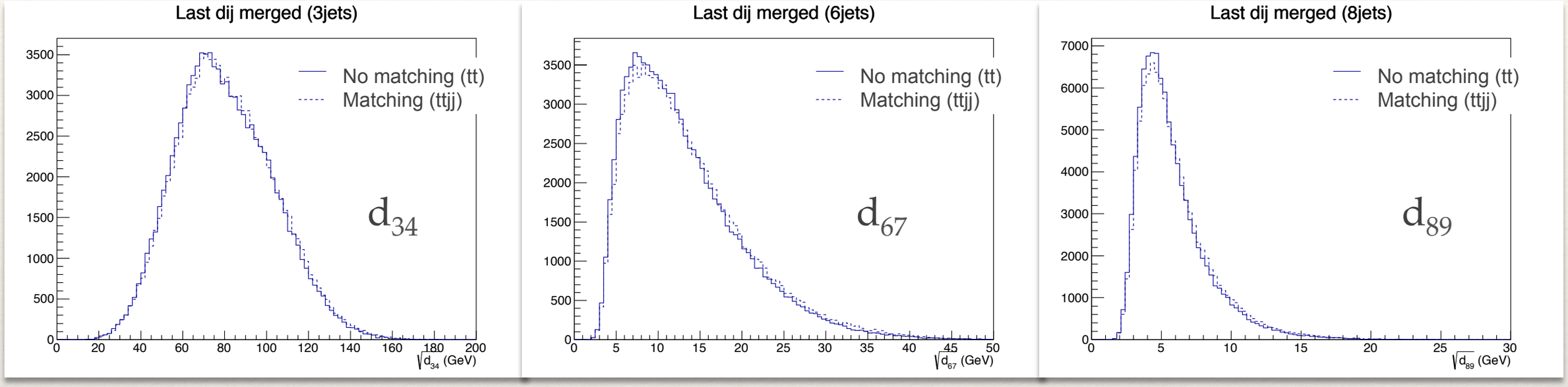


DBD Samples

The signal is reconstructed by choosing the combination of *b* quark jet and W boson that minimises the following equation

$$d^2 = \left(\frac{m_{cand.} - m_t}{\sigma_{m_t}}\right)^2 + \left(\frac{E_{cand.} - E_{beam}}{\sigma_{E_{cand.}}}\right)^2 + \left(\frac{p_b^* - 68}{\sigma_{p_b^*}}\right)^2 + \left(\frac{cos\theta_{bW} - 0.23}{\sigma_{cos\theta_{bW}}}\right)^2$$

Whizard-Pyhtia Matching



The inter-particle distance d_{ij} of the last reconstructed jet

Event generation (Whizard 2.2.0 beta) 10^5events
- ttbar (Matching off)
- ttbar+ttbarj+ttbarjj (Matching on)

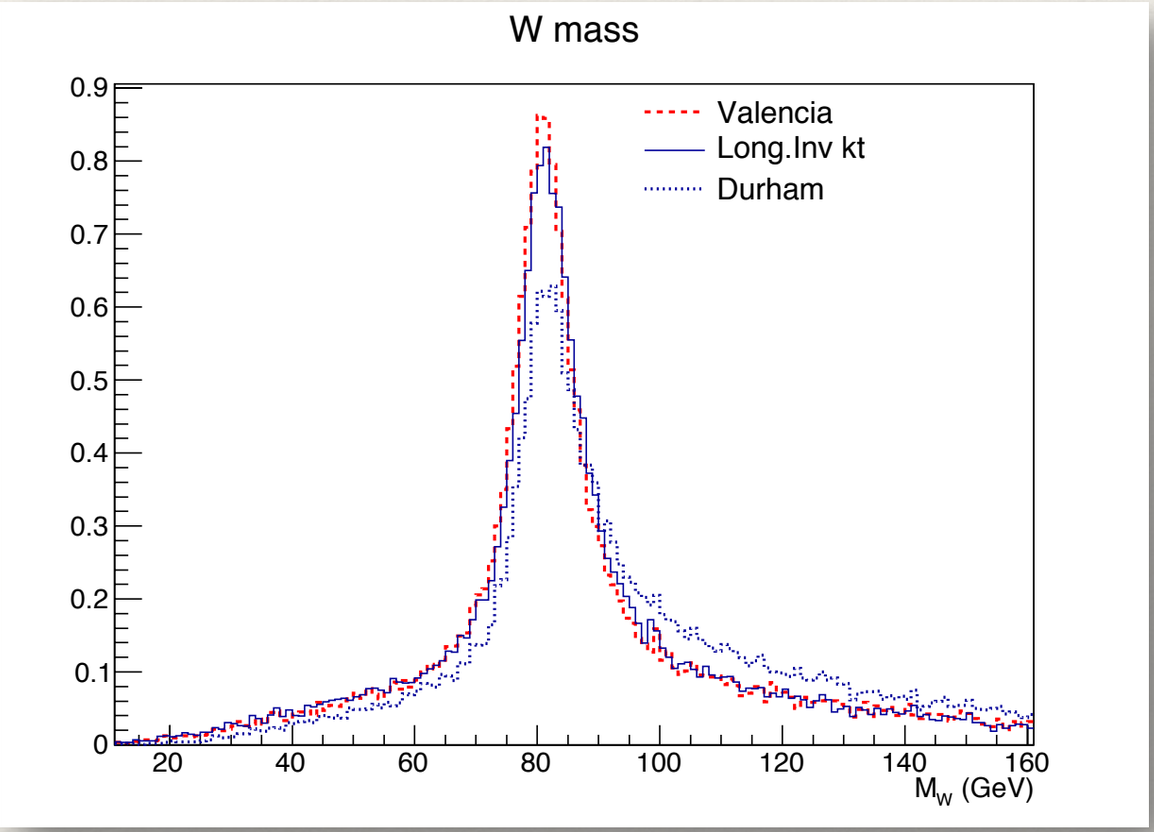
Pythia 8 Hadronisation: ttbar decay into fully hadronic channel (W->jj)

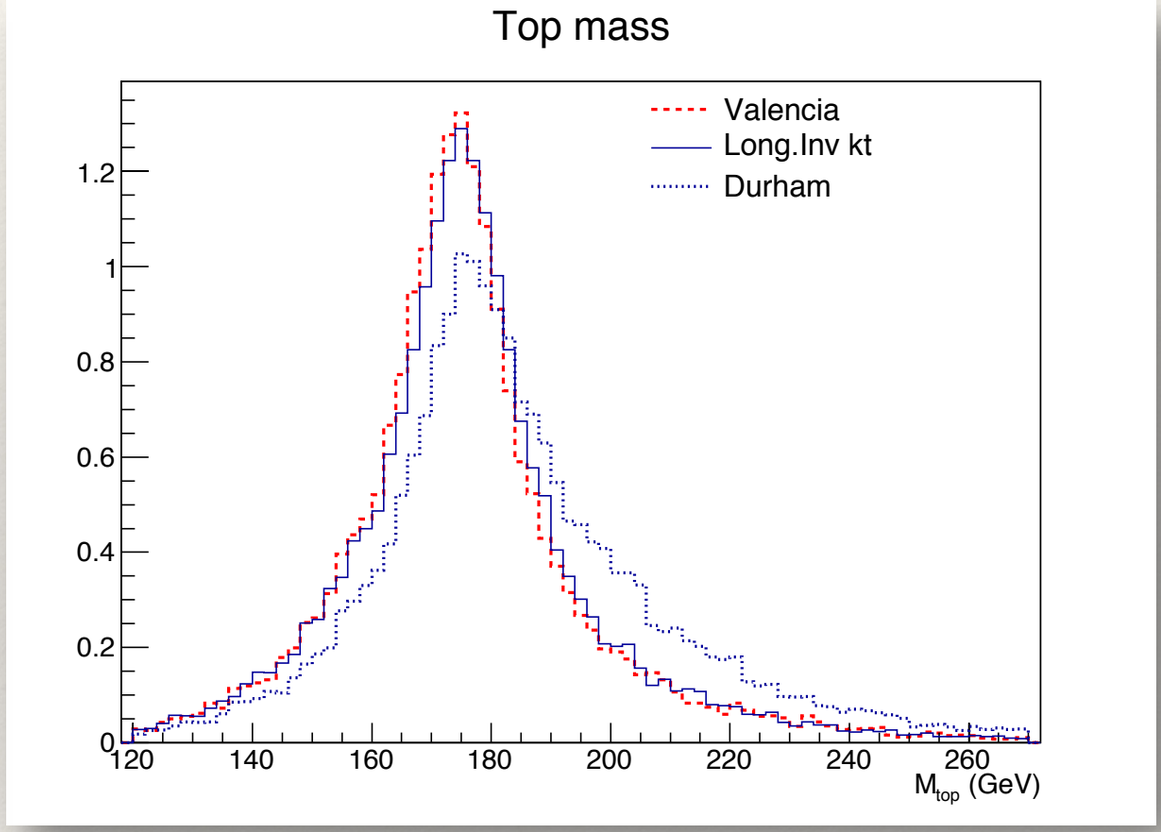
Jet reconstruction with FastJet 3.0.6
- Exclusive mode: 3, 6 and 8 jets

dij is not so affected by gluon emission
- preliminary results
- opened to discussion

tt → (bjj)(blν)

Hadronic top and W candidates





Durham is affected by γγ -> hadrons, longitudinally invariant k_t and Valencia OK

Resolution on jets reconstruction

Degradation of all jet-related measurements due to $\gamma\gamma \rightarrow$ hadrons background

RMS_{90} [GeV]	E_{4j}	E_W	m_W	E_t	m_t
Durham	23.2	19.6	20.3	19.5	21.4
e^+e^- k_t	25.6	20.8	21.6	20.5	22.8
long. inv. k_t	21.7	18.4	18.9	18.4	20.1
Valencia	21.4	18.0	18.8	18.2	20.0

Four-jet system

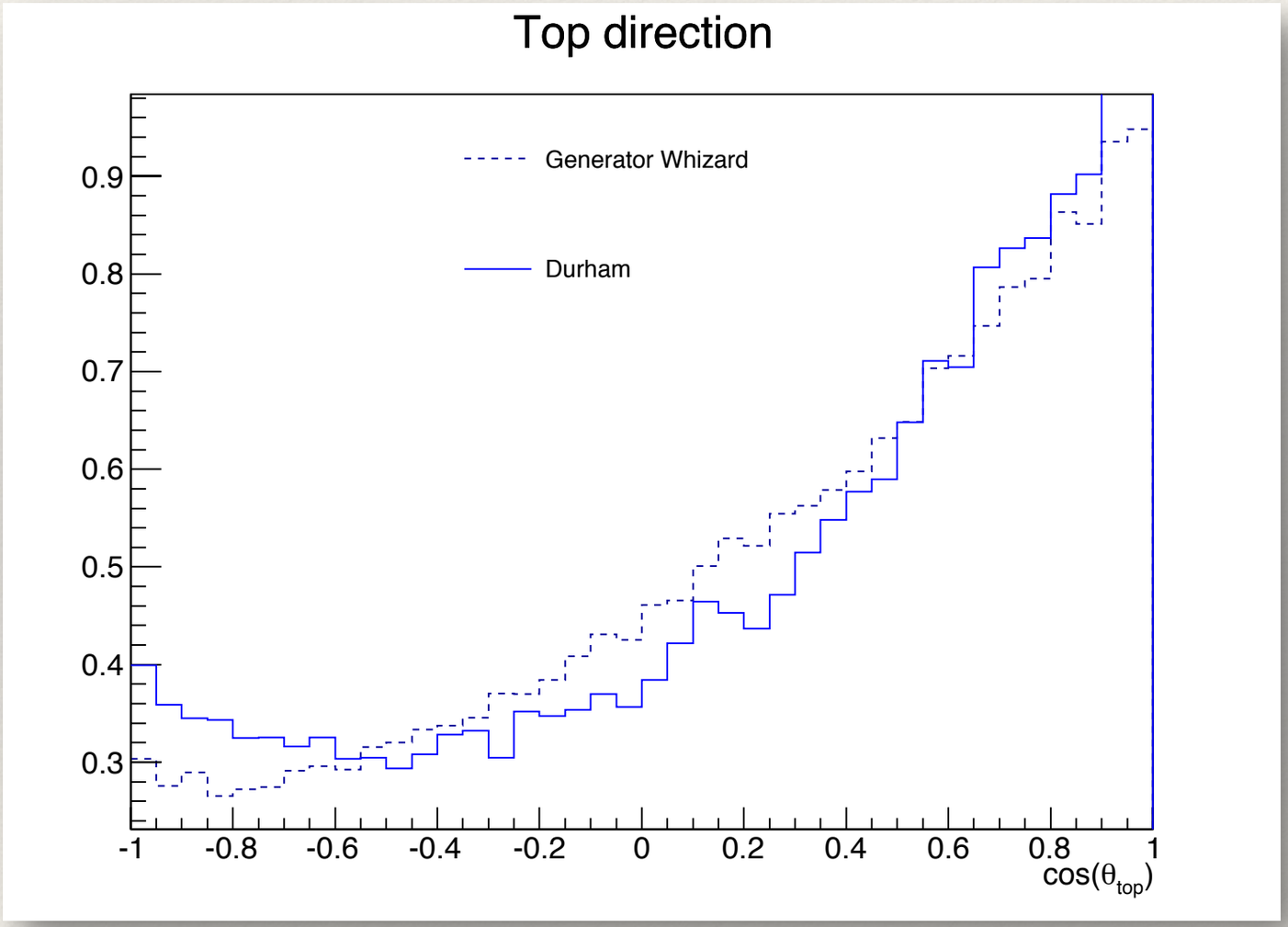
Hadronic W candidate

Hadronic top candidate

Durham significantly degraded.
Hadron collider algorithm and Valencia offer better reconstruction for all hadronic observables

Forward-Backward asymmetry

Durham jet algorithm (first choice)

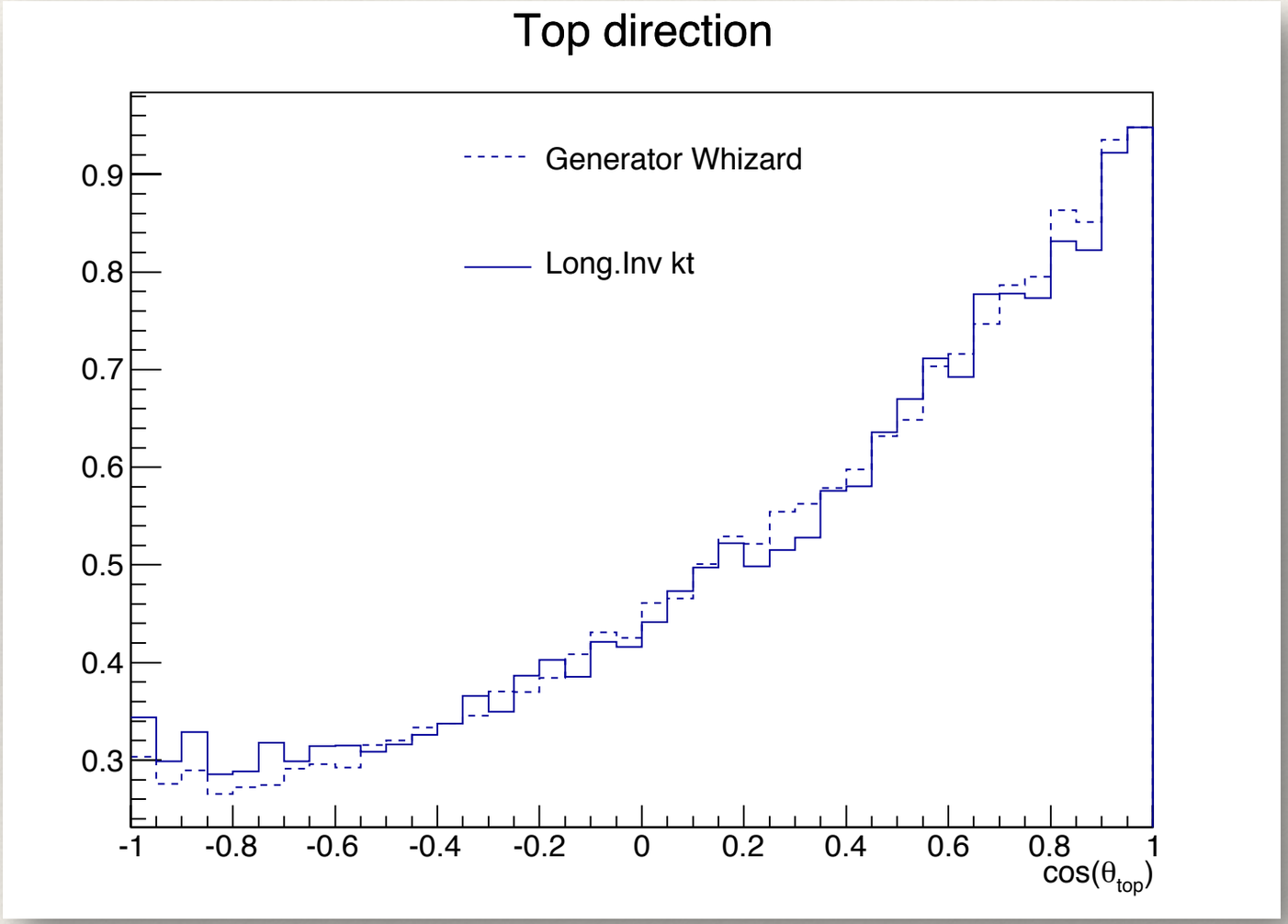


$$AFB_{Whizard} = 33.98\%$$
$$AFB_{reco} = 33.08\%$$

AFB value is ok but Distribution is degraded

Forward-Backward asymmetry

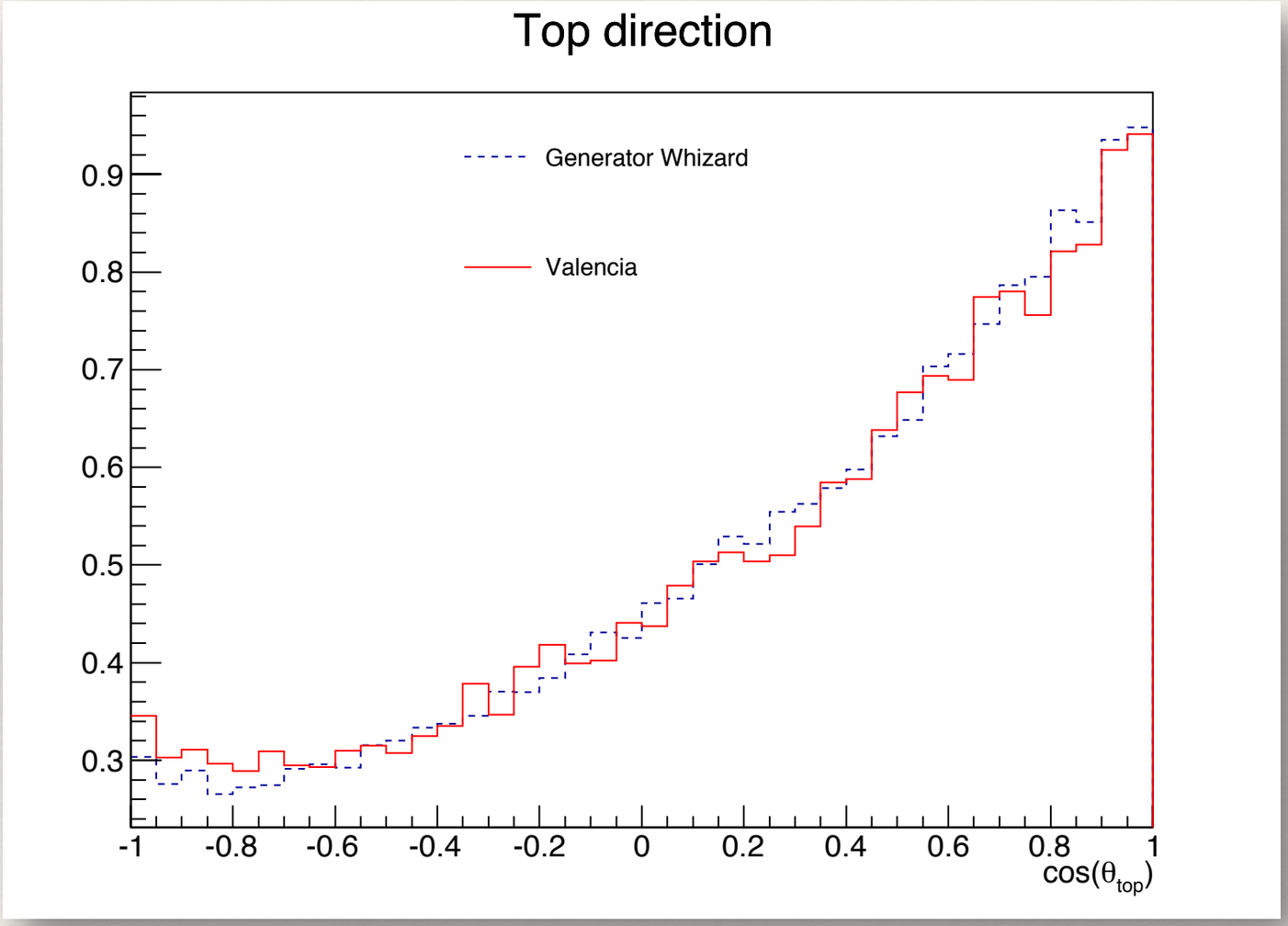
Longitudinally invariant kt (algorithm for hadron colliders – removes γγ->hadrons)



$AFB_{whizard} = 33.98\%$
$AFB_{reco} = 31.90\%$

Forward-Backward asymmetry

Valencia (algorithm for lepton colliders – removes γγ->hadrons)



$AFB_{whizard} = 33.98\%$
$AFB_{reco} = 31.80\%$

e^+e^- style algorithm can compete with hadron collider algorithm

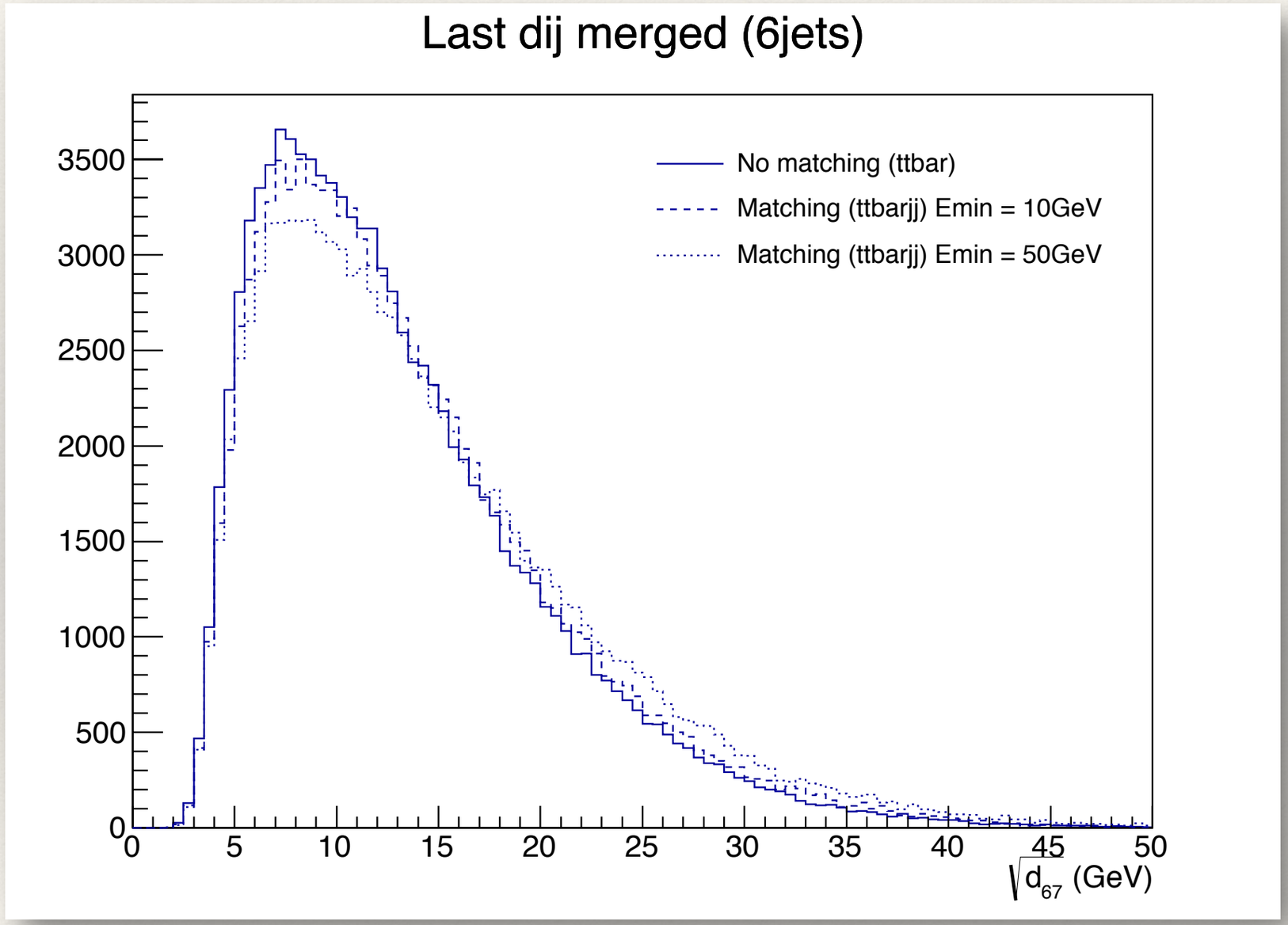
Conclusions

- **$\gamma\gamma \rightarrow$ hadrons bkg. forces us to rethink e^+e^- algorithms**
- **The Valencia** jet algorithm retains the natural inter-particle **distance criterion for e^+e^-** collisions **and** offers **robust performance** in the presence of the (mild) **background** levels expected at lepton colliders
- We used the tt FB asymmetry analysis as the first benchmark

THANK YOU FOR YOUR ATTENTION

BACKUP SLIDES

Matching: Energy cut



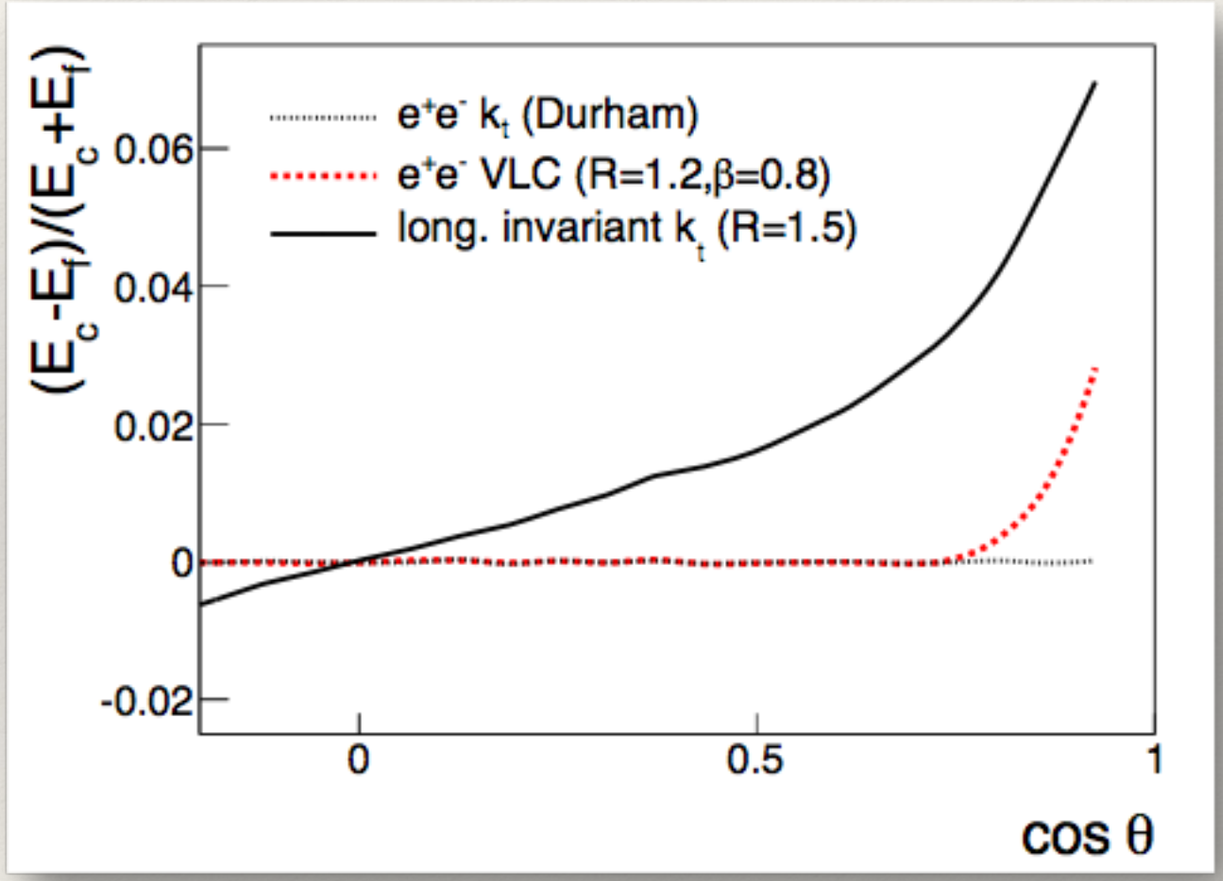
ttbar+ttbarj+ttbarjj

Matching on

$E_{min} = 10$ GeV

$E_{min} = 50$ GeV

Comparison of the distance criteria



Decreasing distance in forward region
→ **bias in energy sharing**

Toy experiment with two jets with typical lateral development, separated by 1...

Pronounced bias for long. invariant k_t

Effect of beam jets visible for very forward jets in Valencia algorithm

Jet reconstruction performance

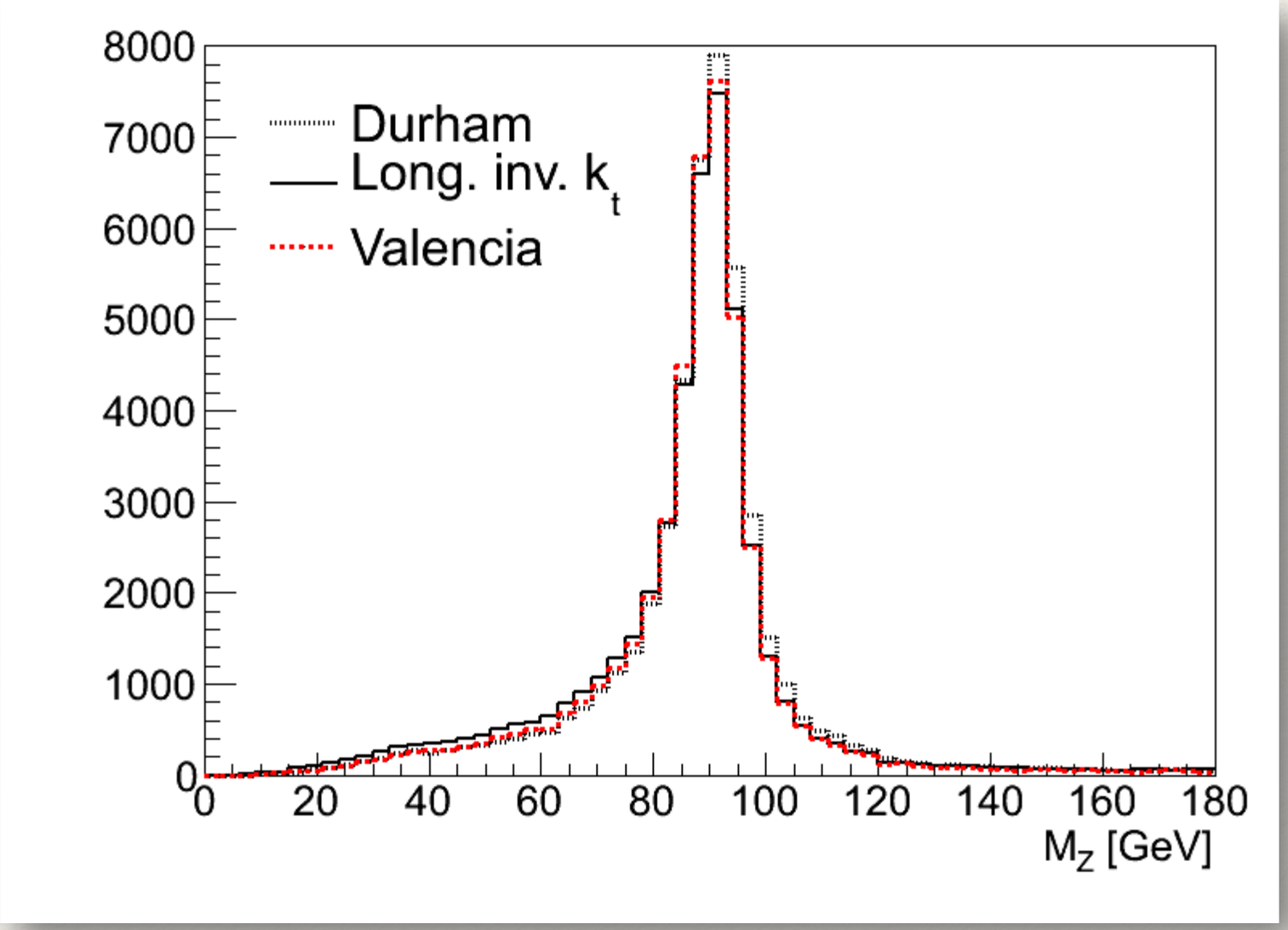
CLIC di-boson (ZZ) production @ 500 GeV

Reconstruct Particle Flow objects using PANDORA

Reconstruct jets (exclusive, n=4)

Form Z boson candidates, selecting best jet pairs

Chosen to facilitate comparison with Marshall&Thomson, CLIC CDR



No background: it doesn't really matter which algorithm you pick

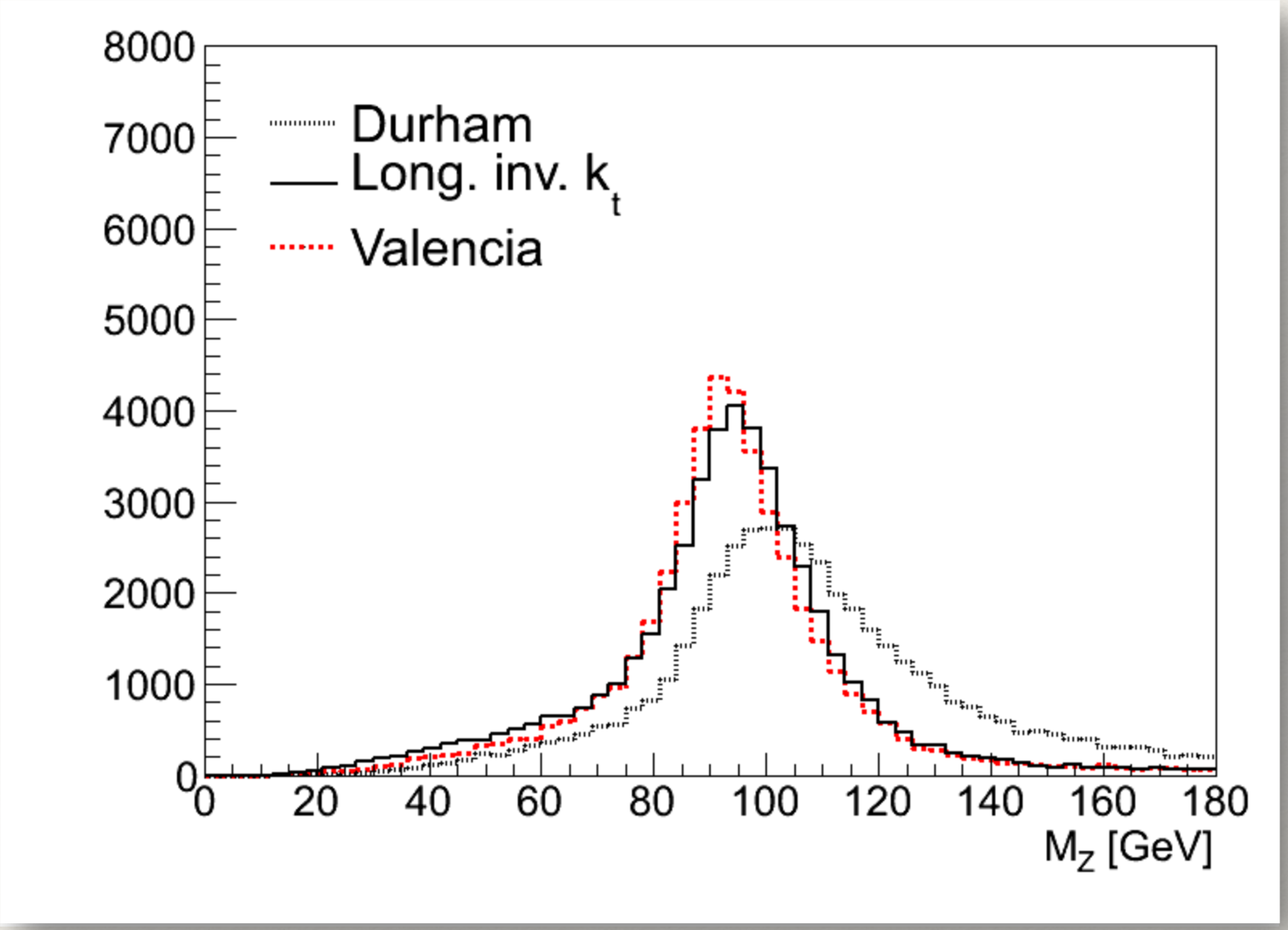
Jet reconstruction performance

CLIC di-boson (ZZ) production @ 500 GeV
+ 300 BX of γγ → hadrons

Reconstruct Particle Flow objects using PANDORA
+ quality and timing cuts

Reconstruct jets (exclusive, n=4)

Form Z boson candidates, selecting best jet pairs



Nominal background: Durham is severely affected, longitudinally invariant k_t and Valencia OK

Jet reconstruction performance

The previous results in numbers: central value and width of the Z-boson mass peak

$\sqrt{s}$ = 500 GeV, no background overlay			
[GeV]	m_Z	σ_Z	RMS_{90}
Durham	90.6	5.4	13.8
long. inv. k_t	90.4	5.3	14.3
Valencia	90.3	5.2	12.5
$\sqrt{s}$ = 500 GeV, 0.3 $\gamma\gamma \rightarrow$ *hadrons* events/BX			
[GeV]	m_Z	σ_Z	RMS_{90}
Durham	101.1	13.6	28.8
long. inv. k_t	95.1	10.9	17.9
Valencia	93.1	10.2	17.1

e^+e^- style algorithm can compete with hadron collider algorithm