XIth Meeting of the Spanish Network for Future Linear Colliders

Jet reconstruction at Linear Colliders

M.Boronat, J.Fuster, **I. García***, E.Ros, M.Vos

IFIC
Valencia, Spain

With thanks to Gavin Salam, André Sailer, Jesse Thaler

Introduction

- **Jet reconstruction at the ILC is not simply an extension of the LEP/SLC experience**

  - higher energy, higher jet multiplicity, more background, better detectors

- **After introduction of $\gamma\gamma \rightarrow$ hadrons in full simulation, most LC physics studies now use hadron collider algorithms**

  - is this the best we can do?

- **Time for a critical evaluation…**

  - understand impact of jet reconstruction on physics performance

  -which algorithms are most suitable?

Jet algorithms

Lepton colliders

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)

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

Generalised e^+e^- k_t algorithm

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

$$d_{iB} = E_i^2$$

Hadron colliders

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)

Include beam distance in e^+e^- algorithms

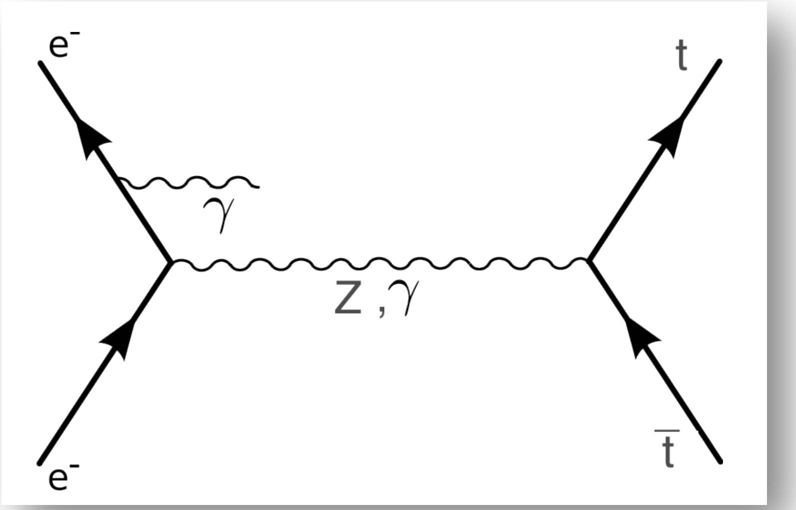
Time to rethink e^+e^- algorithms!!

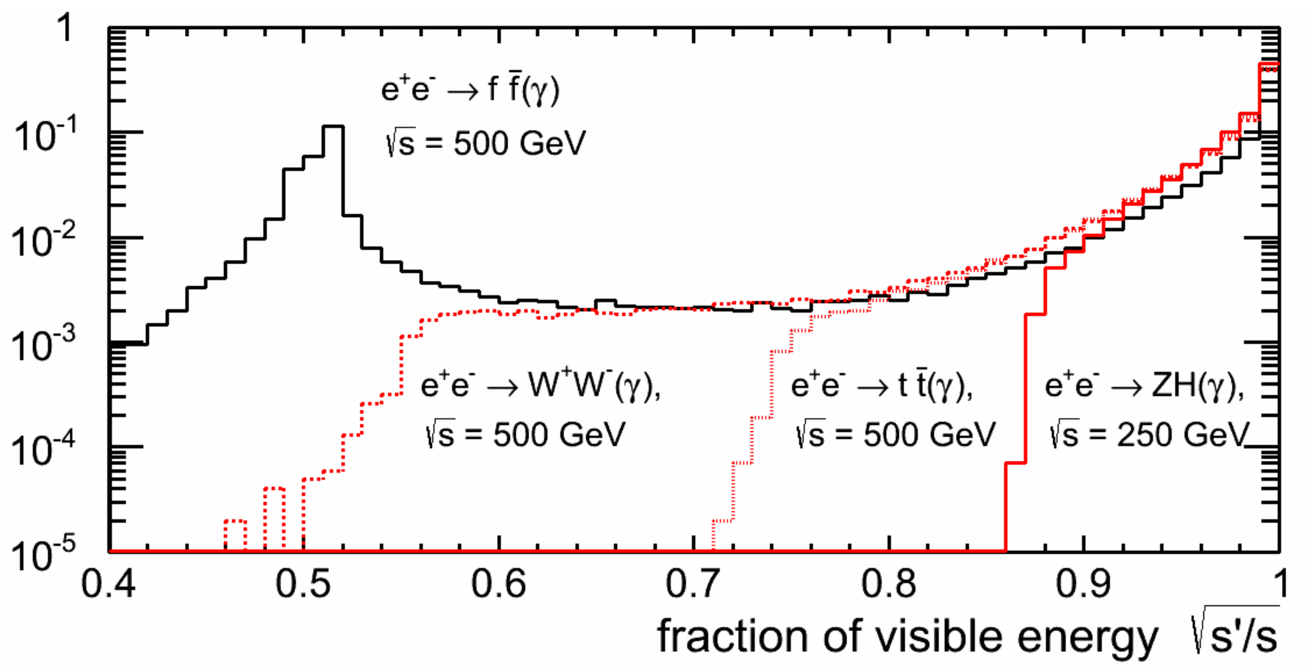
Boost invariance at hadron colliders

- 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 at lepton colliders

- Photons emitted by the incoming beam particles (**Initial State Radiation**) can carry away a significant fractions of the nominal center-of-mass energy
- 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



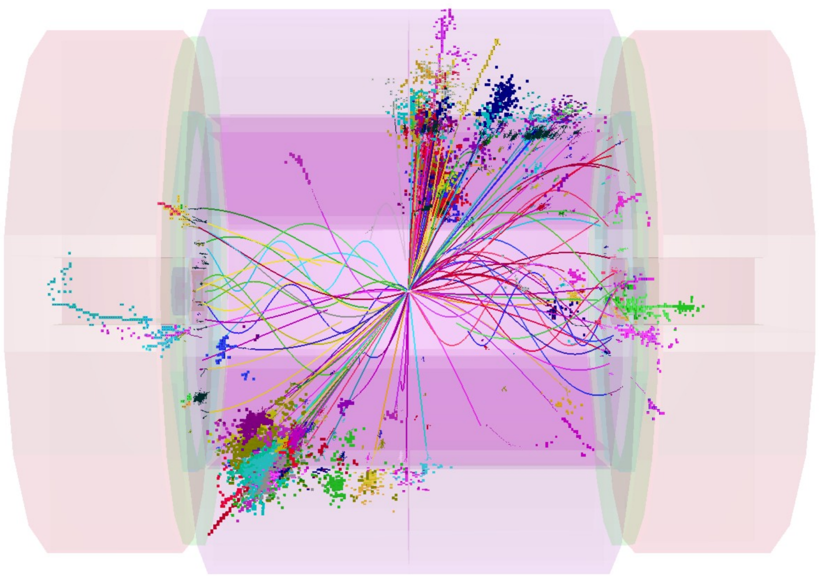


LC backgrounds

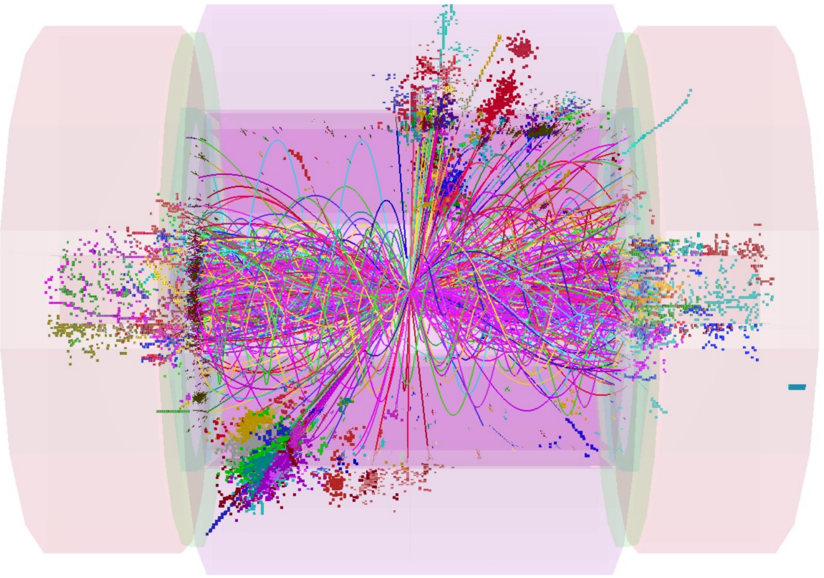
The $\gamma\gamma$ —> **hadrons** background at **CLIC** has **strong** impact on jet reconstruction performance [CLIC CDR, Marshall & Thomson, arXiv:1308.4537]

Less pronounced, but **non-negligible** impact on **ILC** physics [many studies, arXiv:1307.8102]

CLIC 500GeV



CLIC 3TeV



$\gamma\gamma \rightarrow$ hadrons:

1. Strongly peaked in the **forward region**
2. Background **scales with instantaneous luminosity** -> Much larger at 3TeV than at 500GeV
3. Its impact depends on the **bunch structure** and **detector read-out speed**
   → *ILC, 1300 bunches spaced by 500 ns*
   → *CLIC, 312 bunches spaced by 0.5 ns*

Use CLIC case to take jet reconstruction to the limit; if it works there, it's good for ILC too.

The VLC 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**

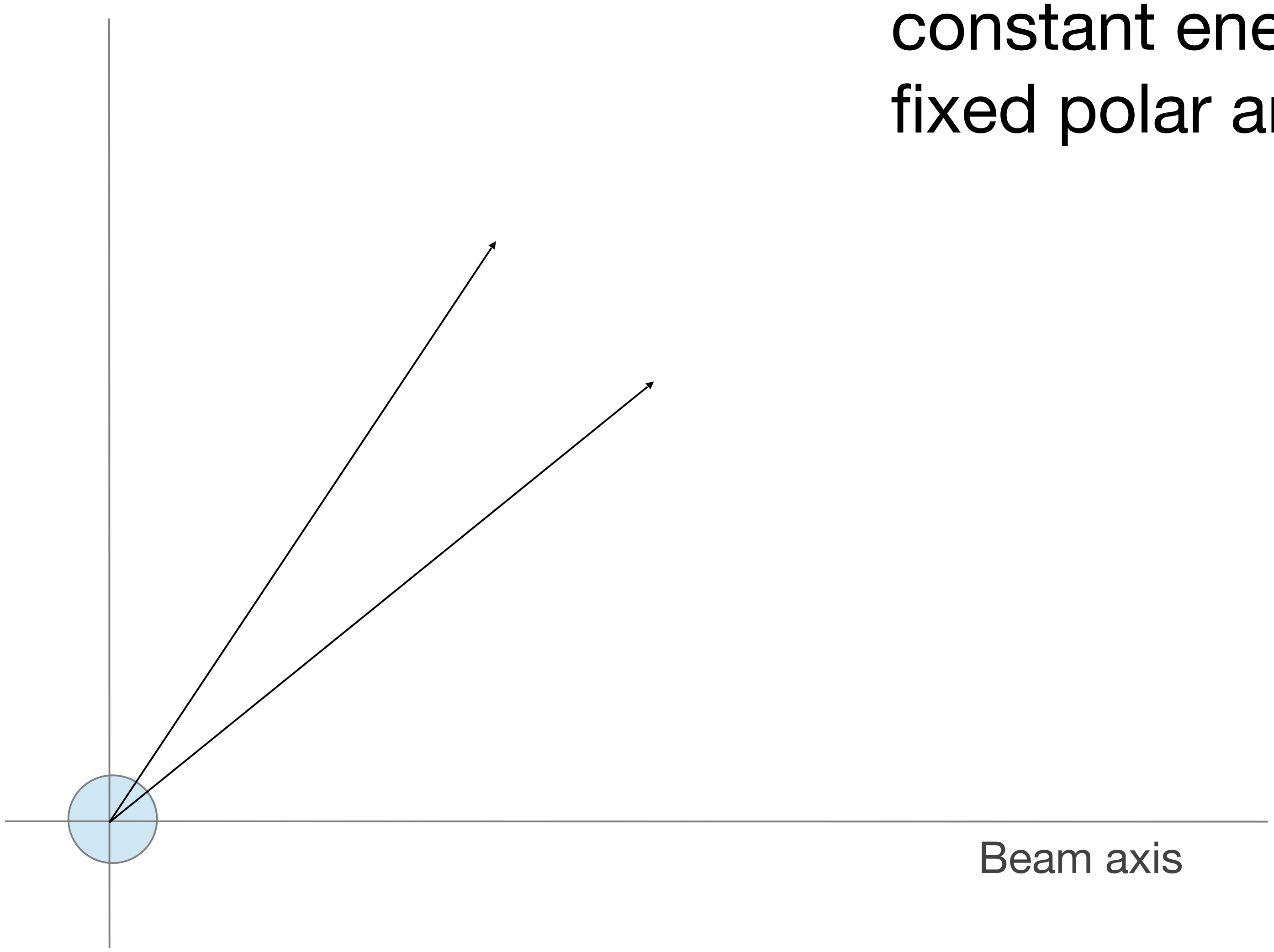
$$d_{iB} = E^{2\beta} \sin^{2\gamma} \theta_{iB}$$

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

*In the default settings the two exponents β and γ are equal. For **β=γ=1** the expression simplifies to $\boldsymbol{d_{iB} = E^2 \sin^2 \theta_{iB} = p^2_{ti}}$

Comparison of the distance criteria

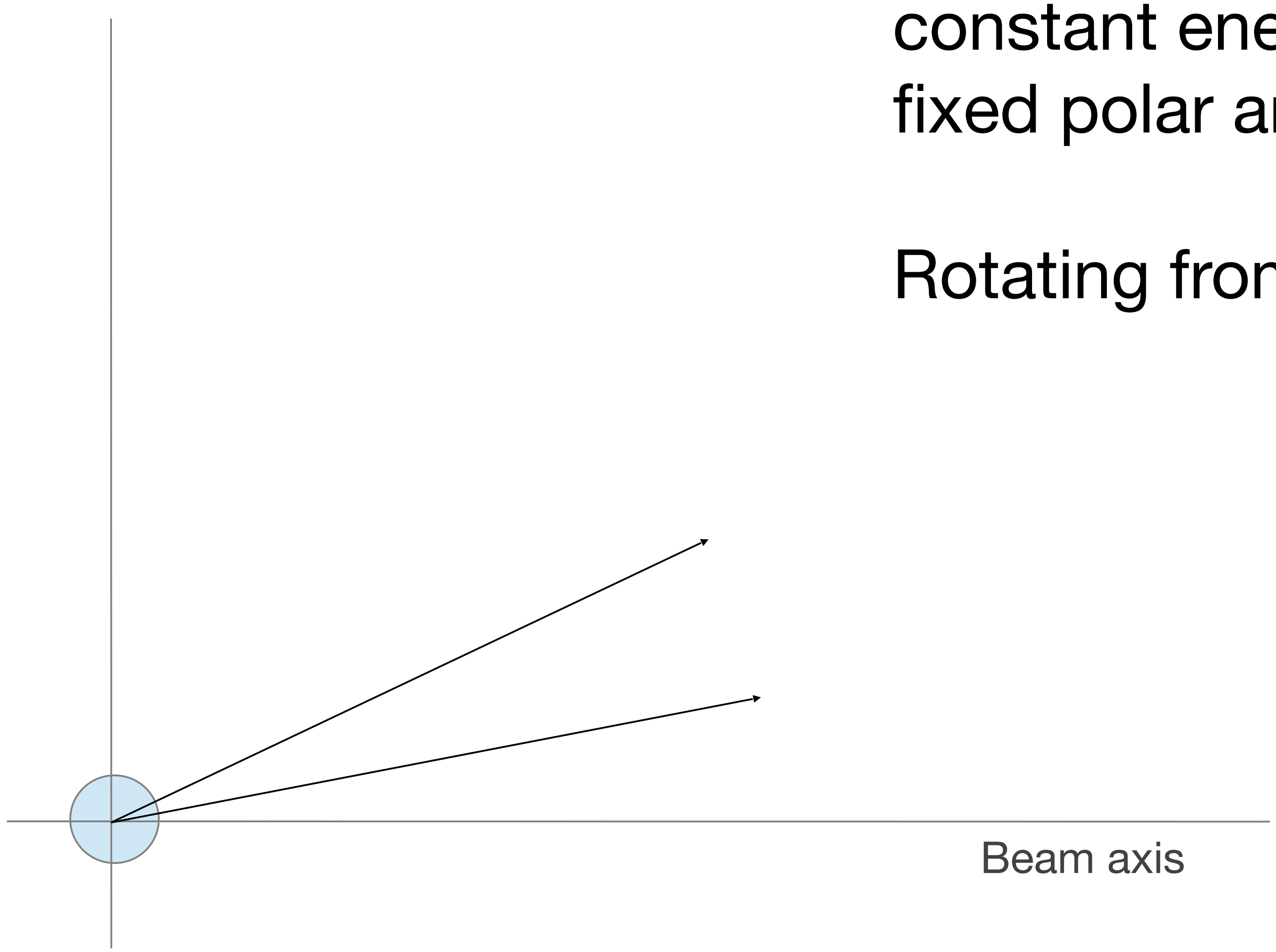
Two test particles with
constant energy (E = 1 GeV) and
fixed polar angle separation (100 mrad)



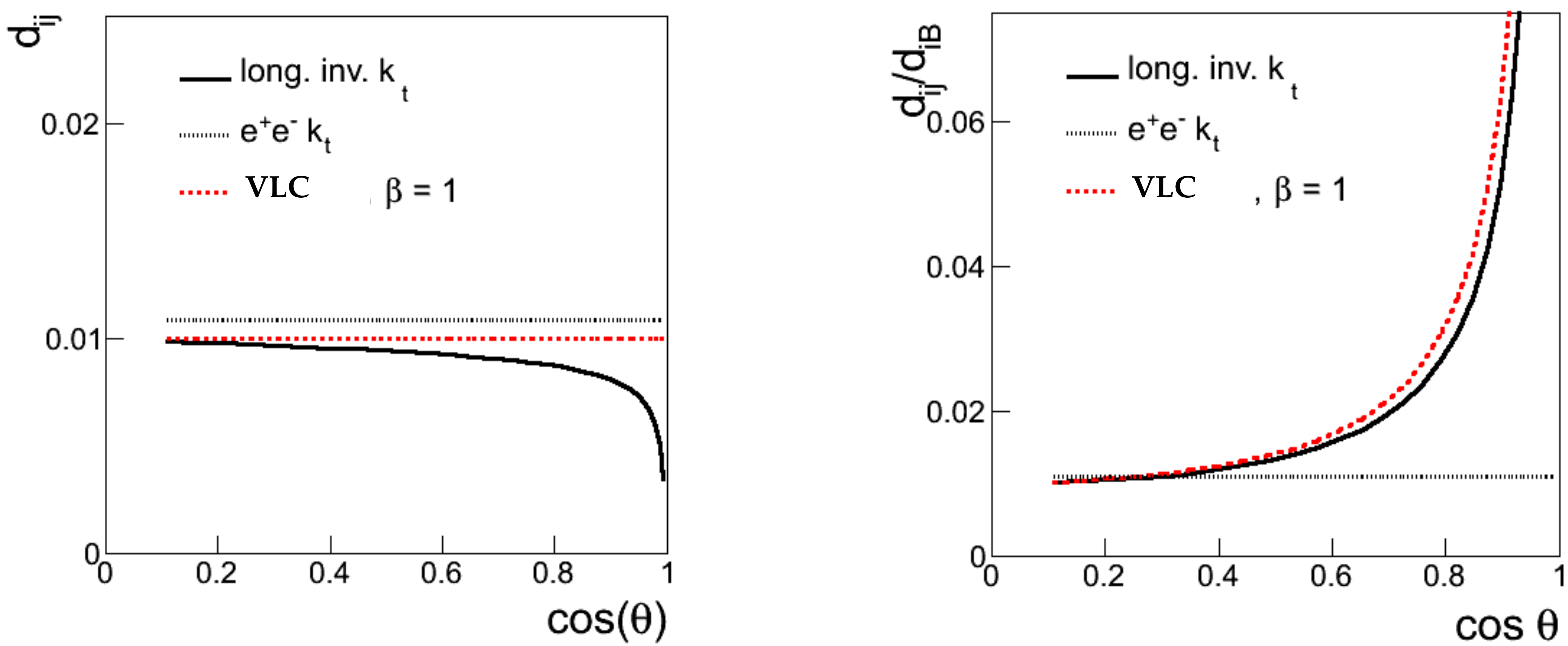
Comparison of the distance criteria

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

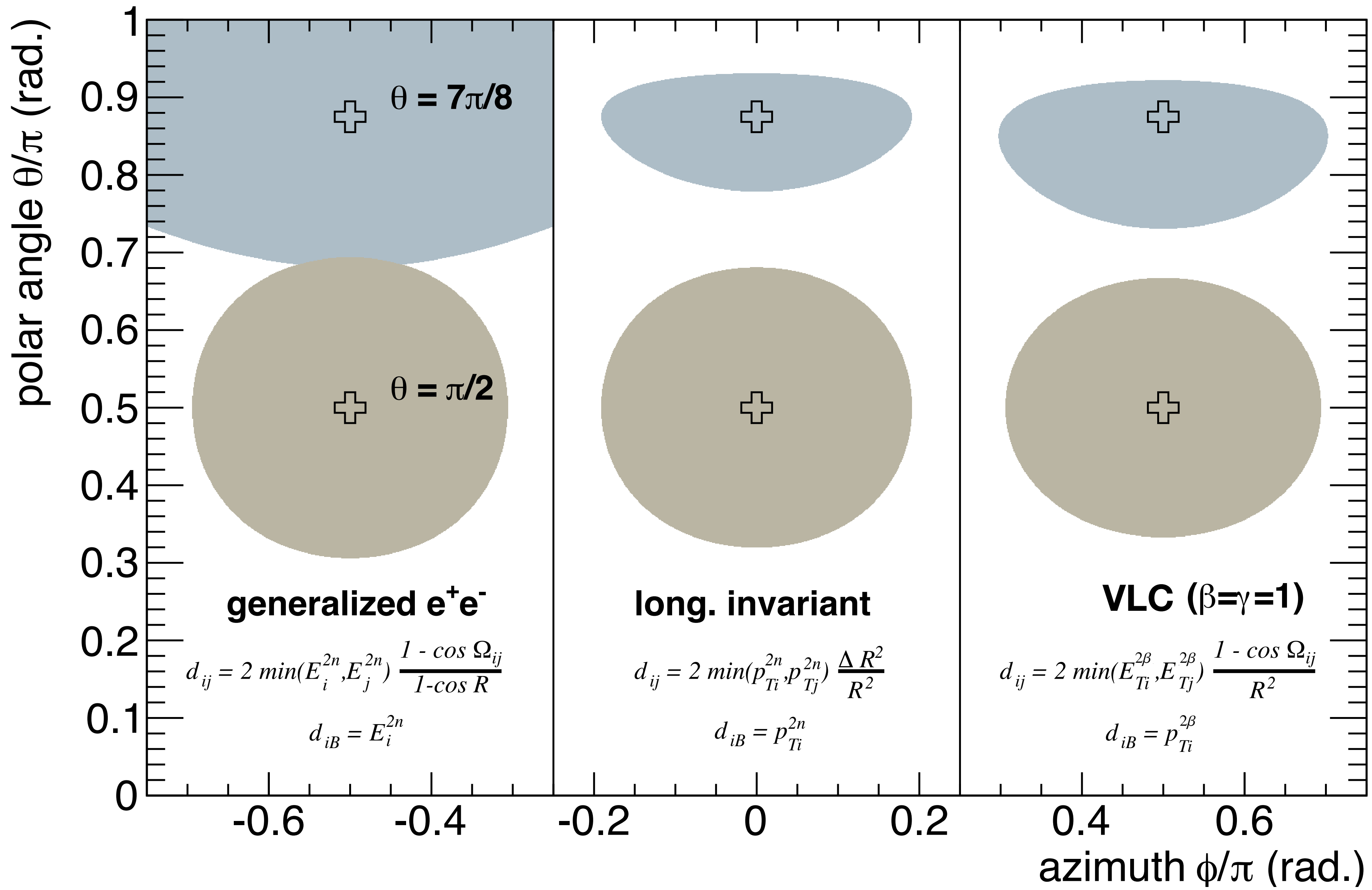


The ratio of the inter-particle distance and the beam distance: d_{ij}/d_{iB} drives the robustness to (forward) background: the decision to assign the particle to final-state or beam jets depends on this ratio (and R)

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

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

Comparison of the jet sizes



The footprint or area of jets depends on the jet algorithm

Three algorithms that yield a similar, circular area in the central detector produce very different jets in the forward region

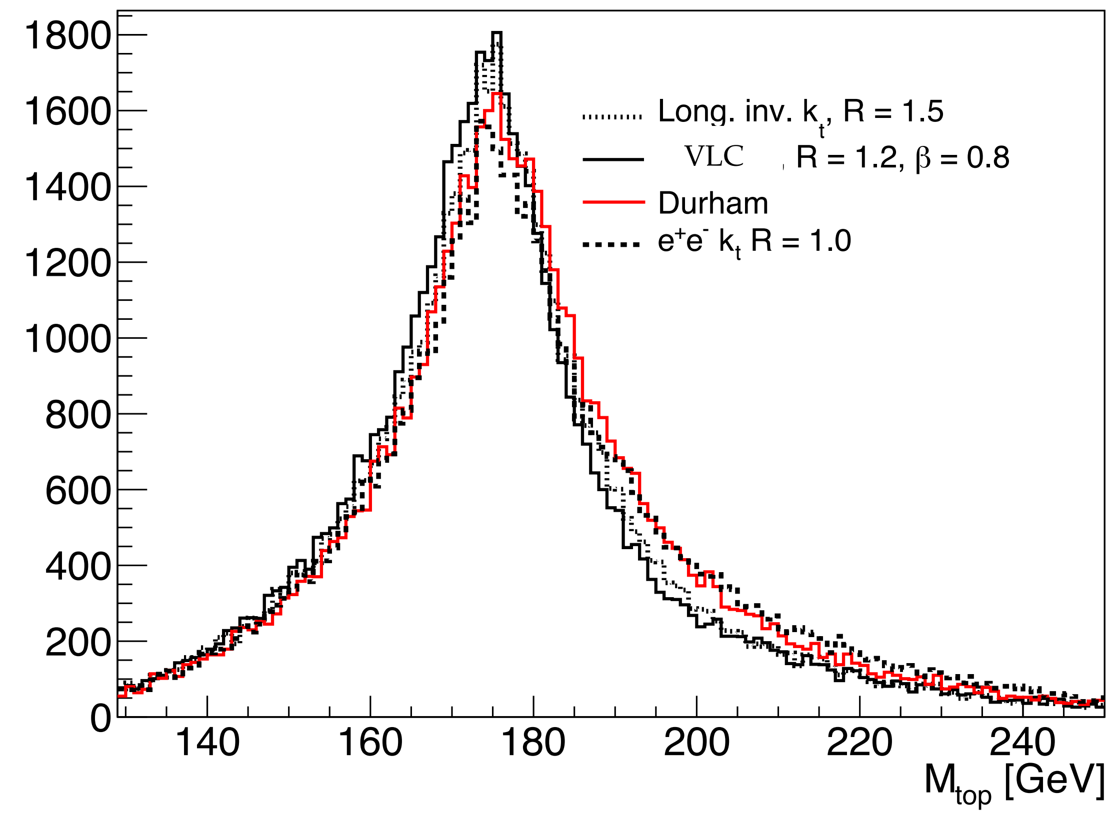
ILC realistic benchmark

IFIC/LAL study of ILC lepton+jets $t\bar{t}$ @ 500 GeV, [arXiv:1307.8102]

$t\bar{t} \to b\bar{b}j_1j_2l\nu$

We consider four jet reconstruction algorithms

- **Durham** algorithm
- **Generic e+e− k_t** algorithm with beam jets with R = 1
- **Longitudinally invariant k_t** algorithm with R = 1.5
- **VLC** algorithm with R =1.2 and β = 0.8.



The choice of parameters corresponds to the optimal setting determined in a scan over a broad range of parameters.

Durham is affected by $\gamma\gamma$ -> hadrons, longitudinally invariant k_t and VLC 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
VLC	21.4	18.0	18.8	18.2	20.0

Four-jet system (E_{4j})

Hadronic W candidate (E_W, m_W)

Hadronic top candidate (E_t, m_t)

Durham and e^+e^- k_t are degraded

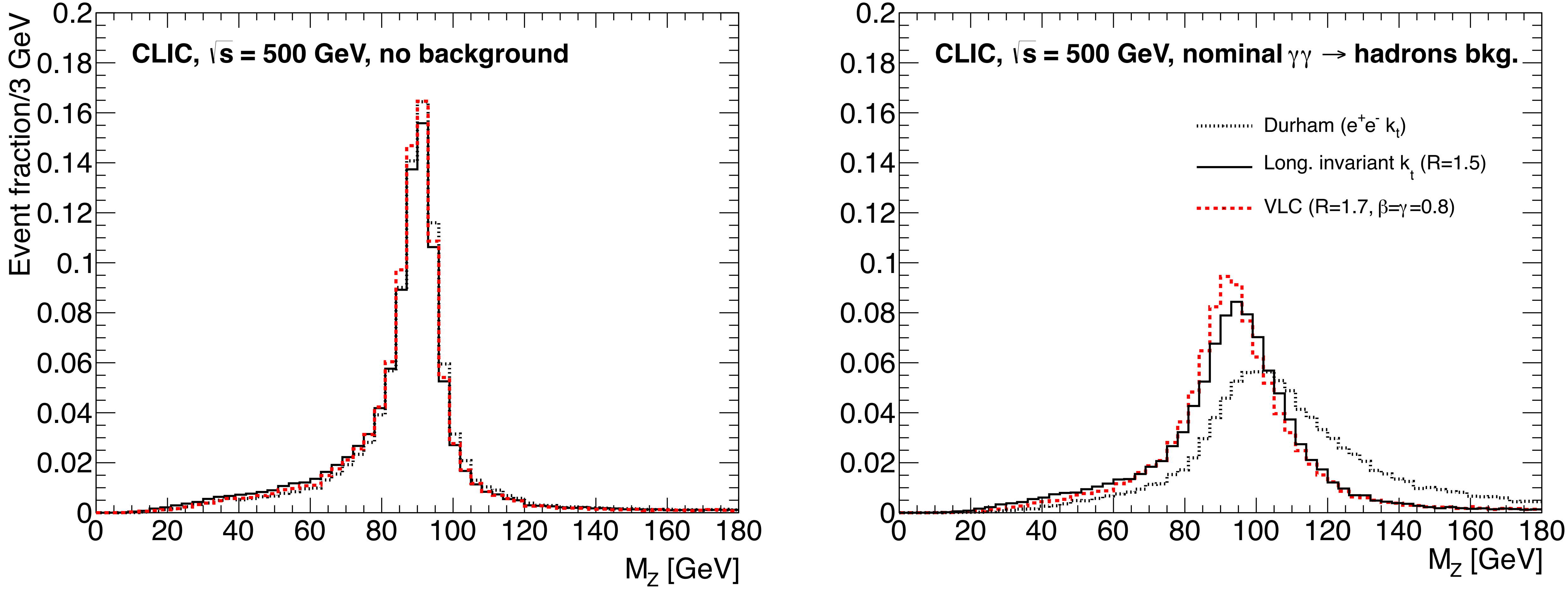
Long. inv. k_t algorithm and VLC offer better reconstruction for all hadronic observables

CLIC realistic benchmark including background

CLIC di-boson (ZZ) production @ 500 GeV

Reconstruct Particle Flow objects using PANDORA

Reconstruct jets (exclusive, n=4) and form Z boson candidates, selecting best jet pairs



Jet energy reconstruction with nominal background much less degraded with algorithms with shrinking footprint (long. Invariant algorithms, VLC) than e^+e^- algorithms (CLIC, high energy)

Jet reconstruction performance

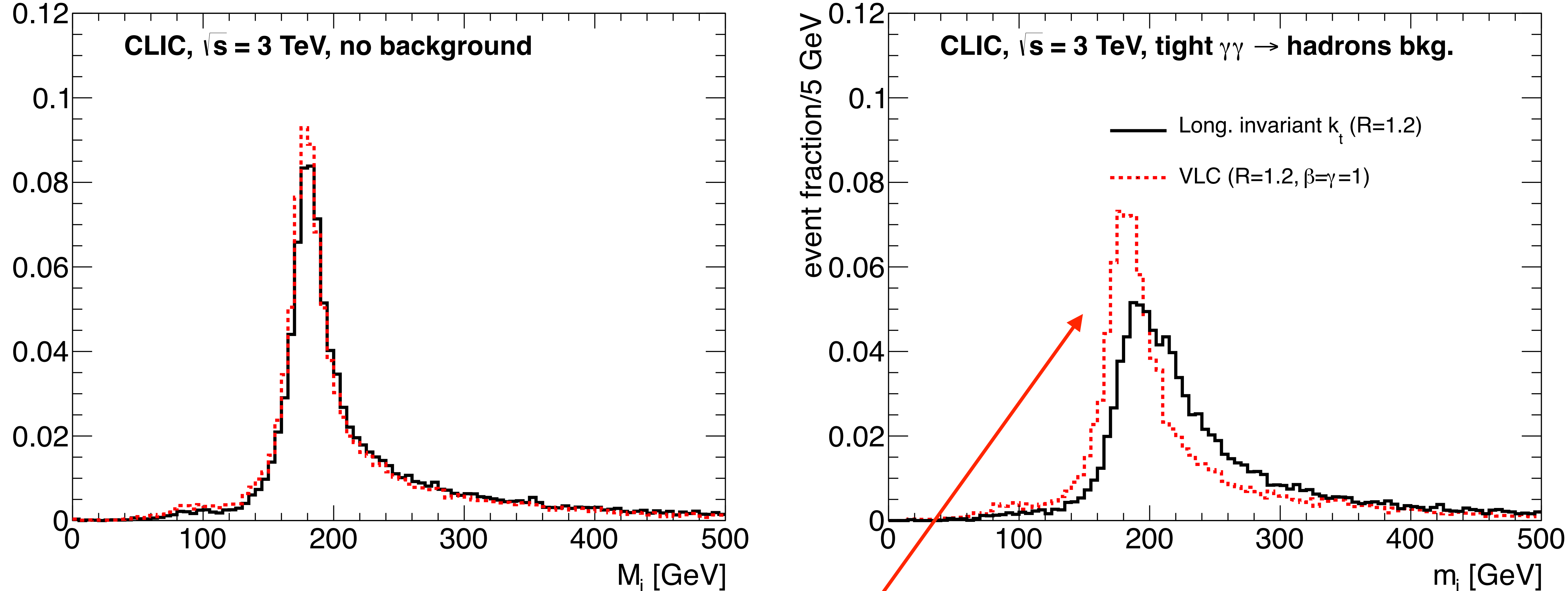
The previous results in numbers: central value, width of the Z-boson mass peak and RMS_{90}

CLIC, $\sqrt{s}$ = 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
VLC ($\beta = \gamma = 1$)	90.3	5.2	12.5
CLIC, $\sqrt{s}$ = 500 GeV, nominal PFO selection			
[GeV]	m_Z	σ_Z	RMS_{90}
Durham	101.1	13.6	28.8
long. inv. k_t	92.0	9.0	17.2
VLC ($\beta = \gamma = 1$)	92.5	9.2	16.2

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

Boosted tops at CLIC 3TeV

$e^+e^- \rightarrow t\bar{t} \rightarrow b\bar{b}q\bar{q}'q''\bar{q}'''$ (fully hadronic decay)



Tight PFO selection is applied

The VLC algorithm performs significantly better than the classical algorithms, including longitudinally invariant k_t .

Boosted tops at CLIC 3TeV

CLIC, $\sqrt{s} = 3$ TeV, no background overlay			
RMS_{90} [%]	E_j (top)	E_j (truth)	m_j
Durham	5.8	3.7	12
generic $e^+e^- k_t$	6.2	2.7	4.5
long. inv. k_t	6.1	2.4	3.4
VLC	5.9	2.4	3.4
CLIC, $\sqrt{s} = 3$ GeV, tight PFO selection			
RMS_{90} [%]	E_j (top)	E_j (truth)	m_j
Durham	7.2	5.6	44
generic $e^+e^- k_t$	6.8	3.4	15
long. inv. k_t	6.1	2.6	9.9
VLC	6.0	2.6	6.8

At higher energy including the $\gamma\gamma$->hadrons background, VLC algorithm offers even better resolution than the hadron collider algorithm long. inv. k_t

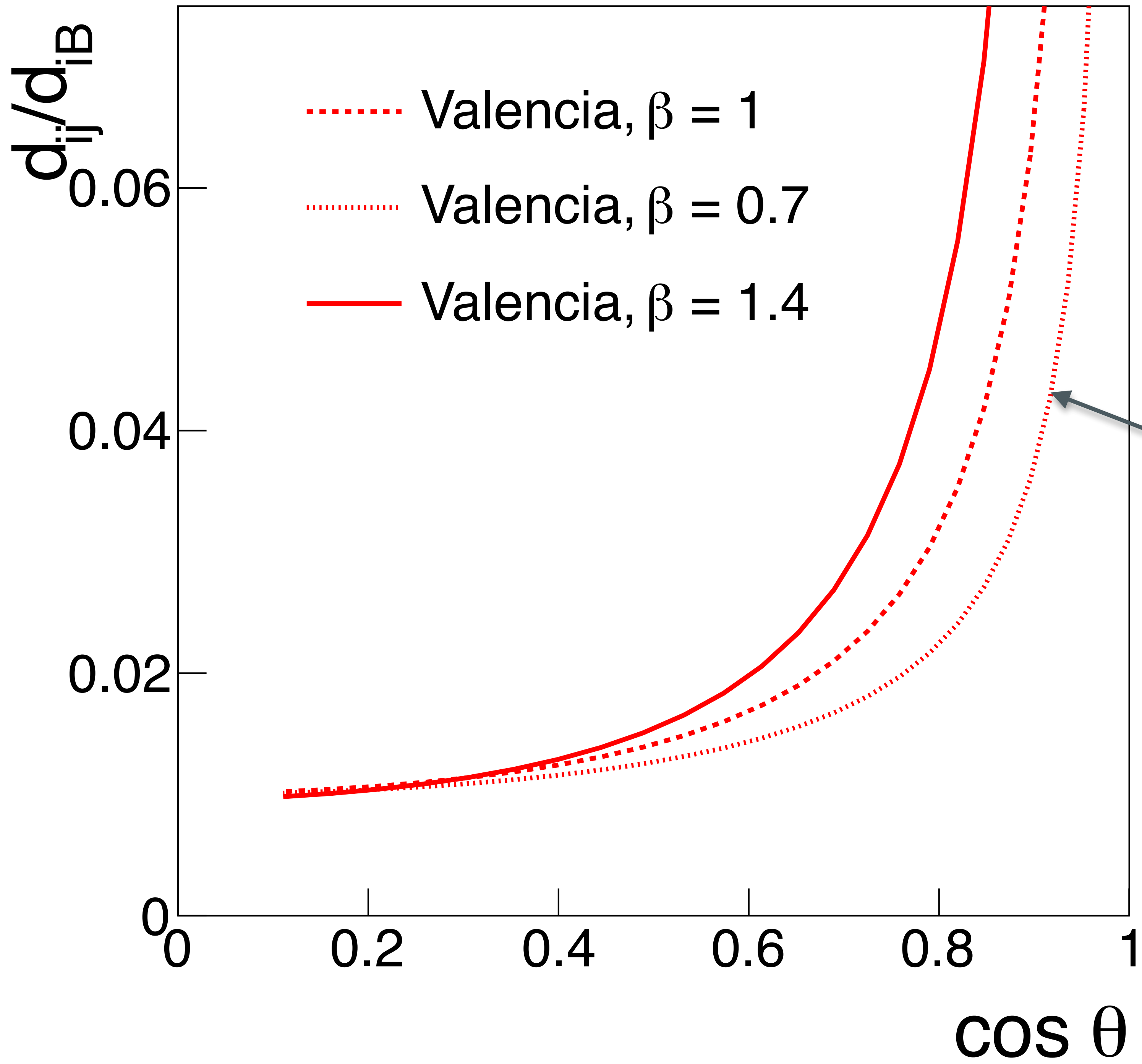
Conclusions

- $\gamma\gamma \rightarrow$ **hadrons bkg. forces us to rethink e^+e^- algorithms** because **old e+e- algorithms** are severally **degraded**
- **The VLC** jet algorithm retains the natural inter-particle **distance criterion for e^+e^-** collisions **and** offers **robust performance** in the presence of the $\gamma\gamma \rightarrow$ **hadrons background** levels expected at lepton colliders
- Shown to work on several benchmark analyses.
- In the most challenging environment the VLC algorithm has significantly better background rejection than the longitudinally invariant k_t algorithm.
- Pre-print out on the arXiv:
  - *Boronat, Fuster, Garcia, Ros, Vos*, **A robust jet reconstruction algorithm for high-energy lepton colliders, arXiv:1404.4294**

Thank you for your attention

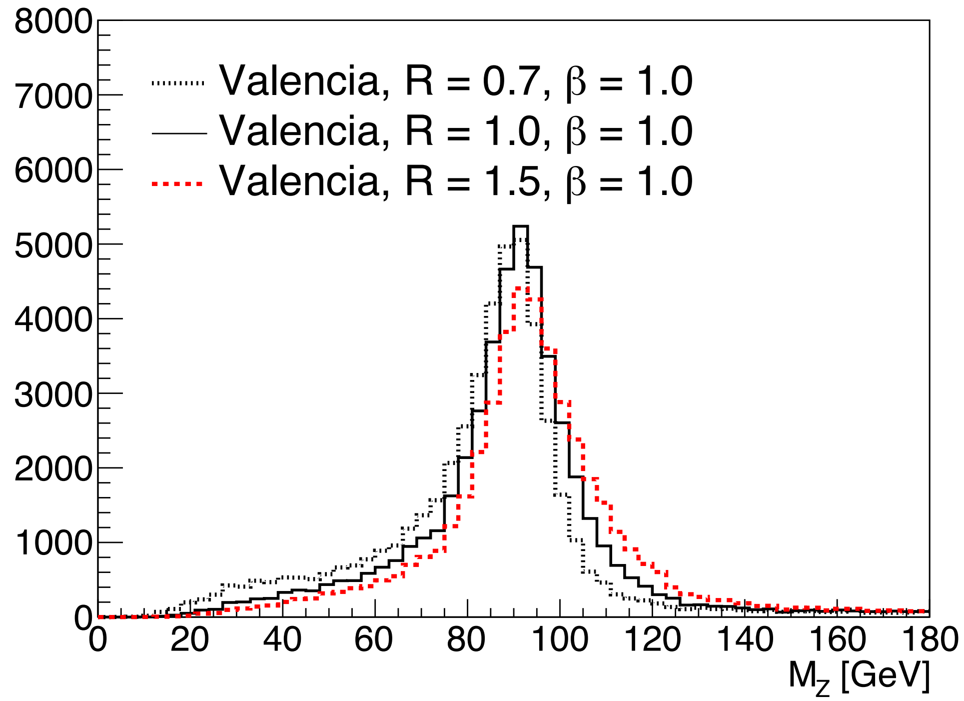
BACK-UP SLIDES

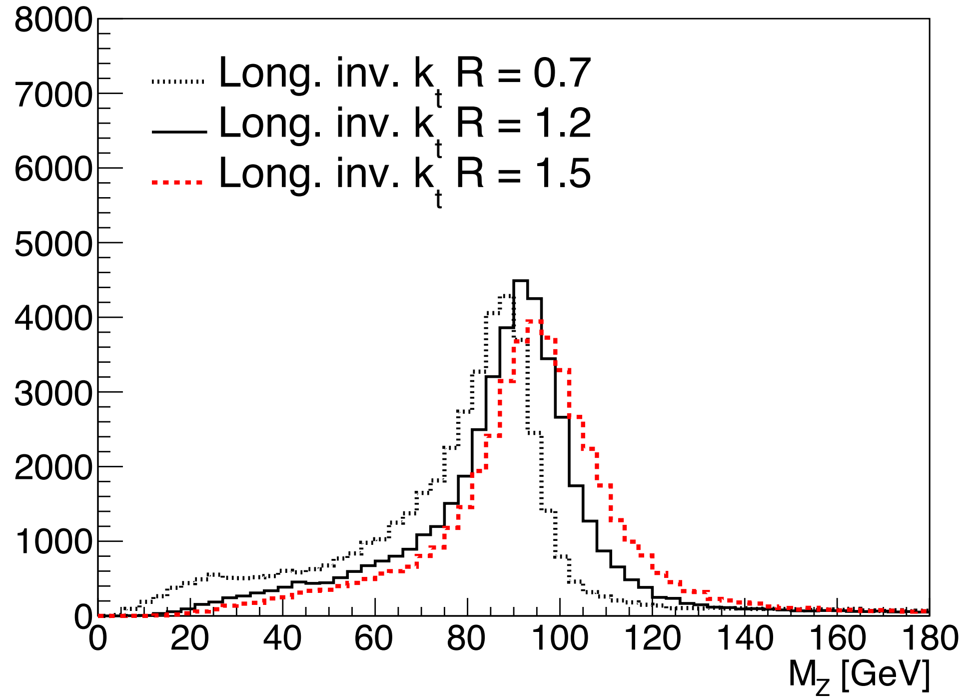
Background rejection



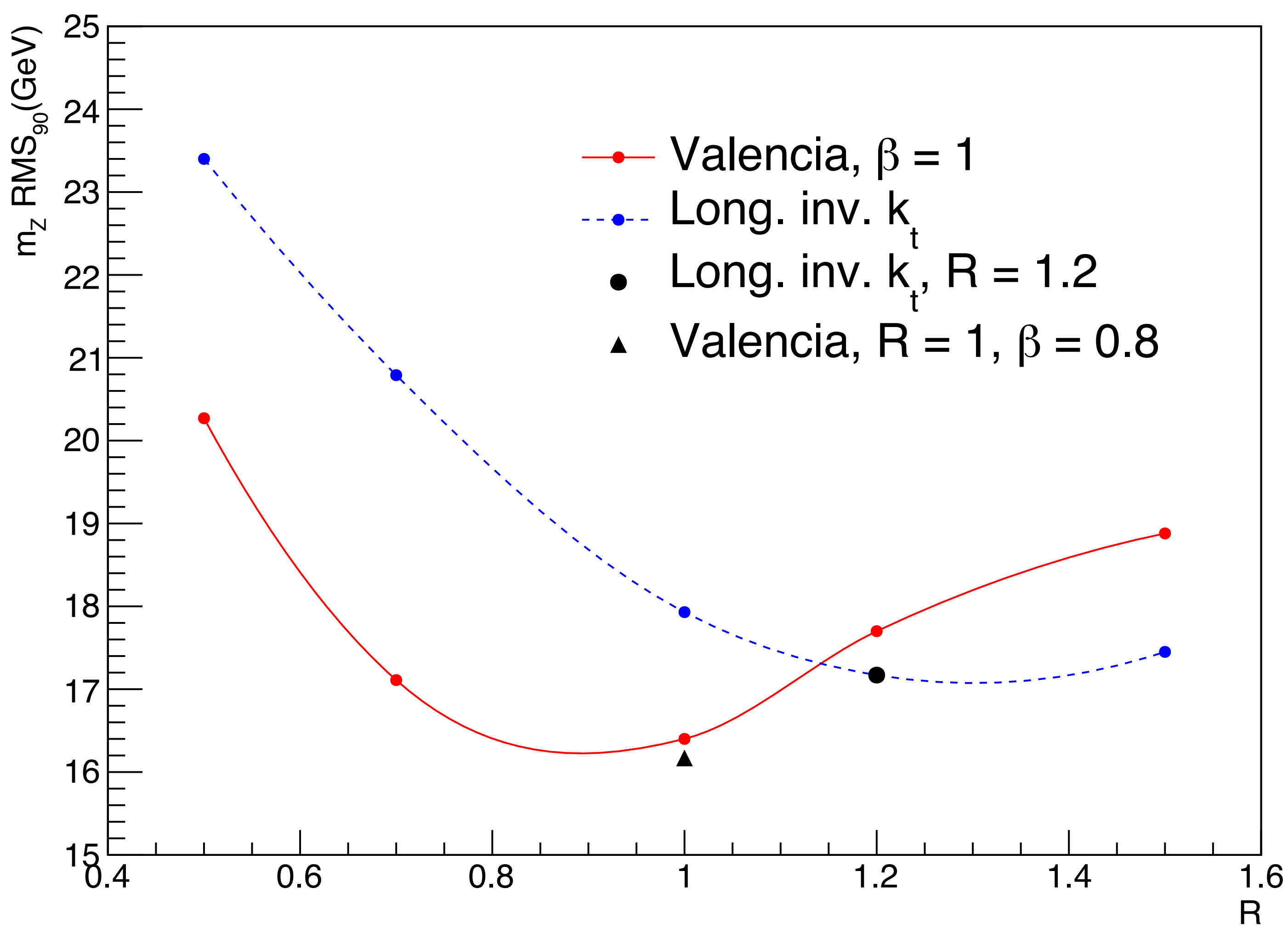
Smaller β larger background rejection

Algorithm parameters optimisation: R scan

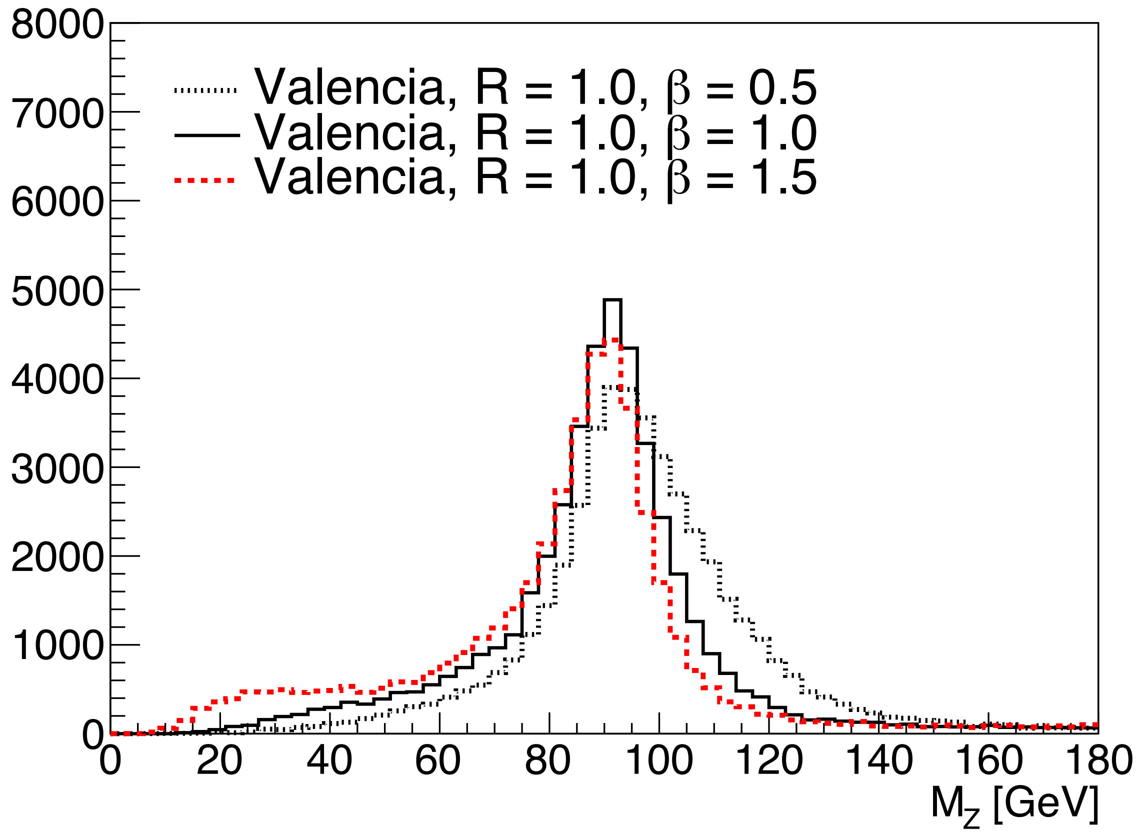


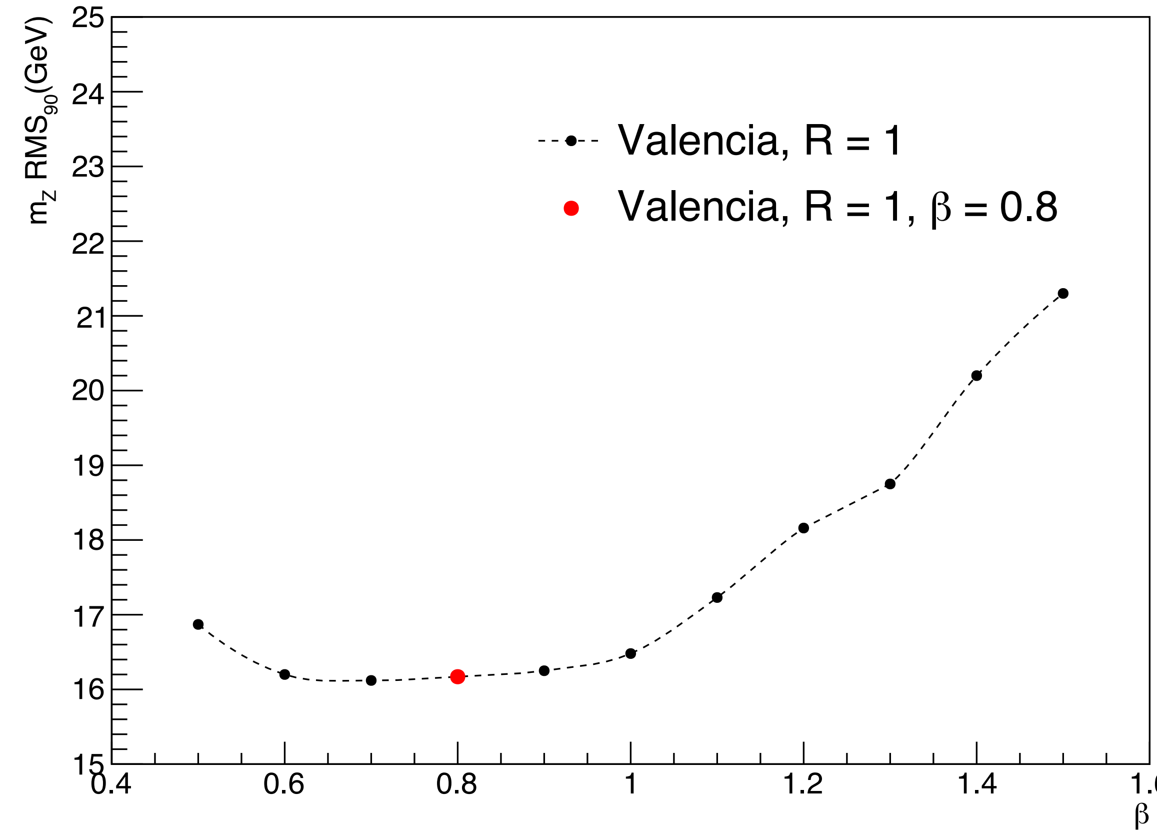


The choice of parameters corresponds to the optimal setting determined in a scan over a broad range of parameters.



Algorithm parameters optimisation: β scan





Boosted top quarks

CLIC 3 TeV ($e^+e^- \rightarrow t\bar{t}\, e^+e^- \rightarrow t\bar{t} \rightarrow b\bar{b}q\bar{q}'q''\bar{q}'''$)

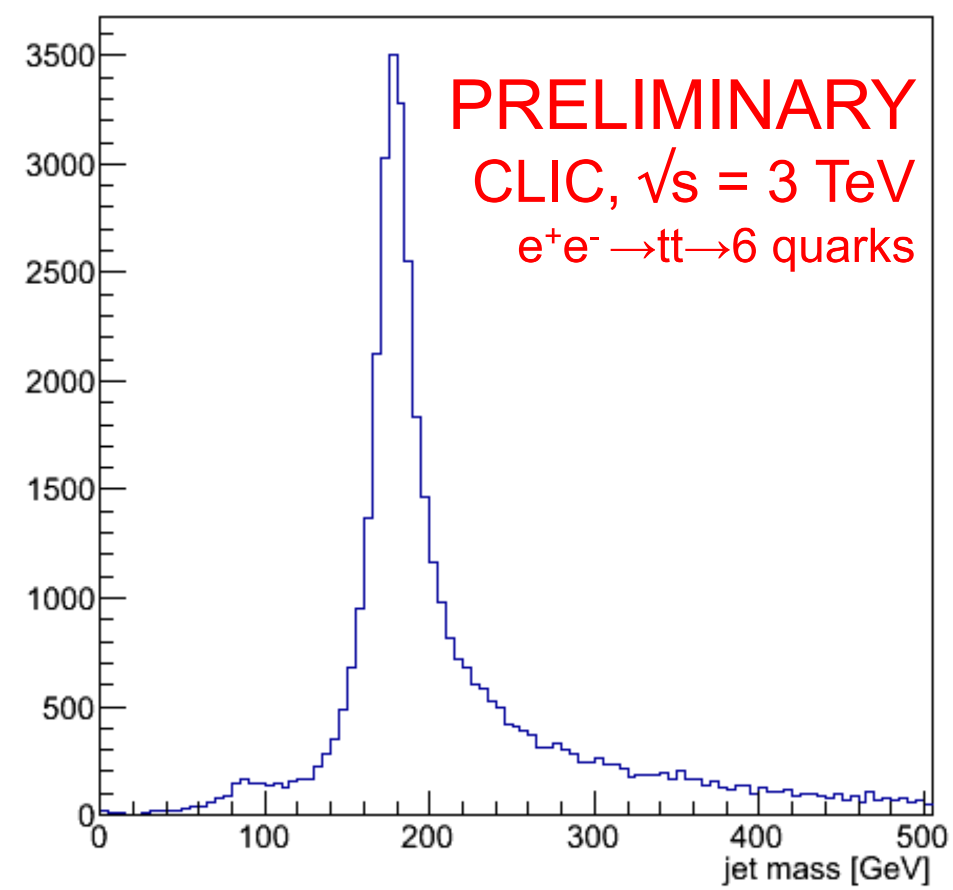
Without $\gamma\gamma \rightarrow$ hadrons background

CLIC-ILD detector simulation
PANDORA PFA

Valencia e^+e^- jet algorithm (N_j =2, R=1, b=1)
Could have picked long. inv. k_t with R=0.8-1.2

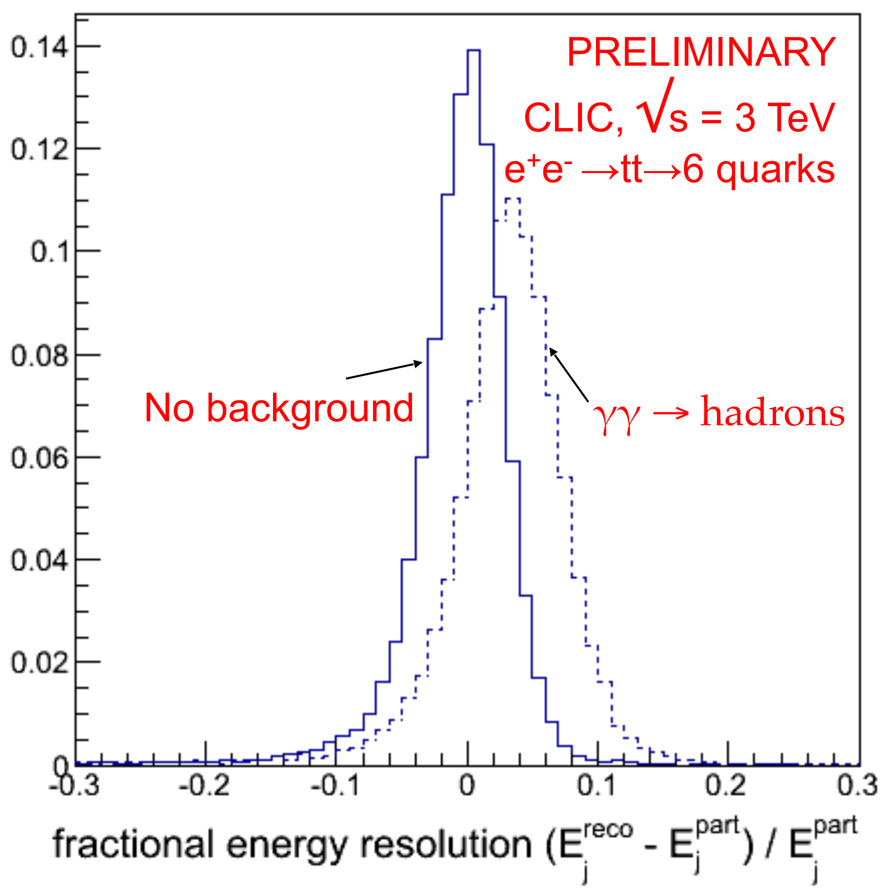
Detector performance for boosted hadronic top jets (E~1200 GeV)

- Energy resolution (RMS90) = 2.4%
- Jet mass resolution (RMS90) = 3.2%



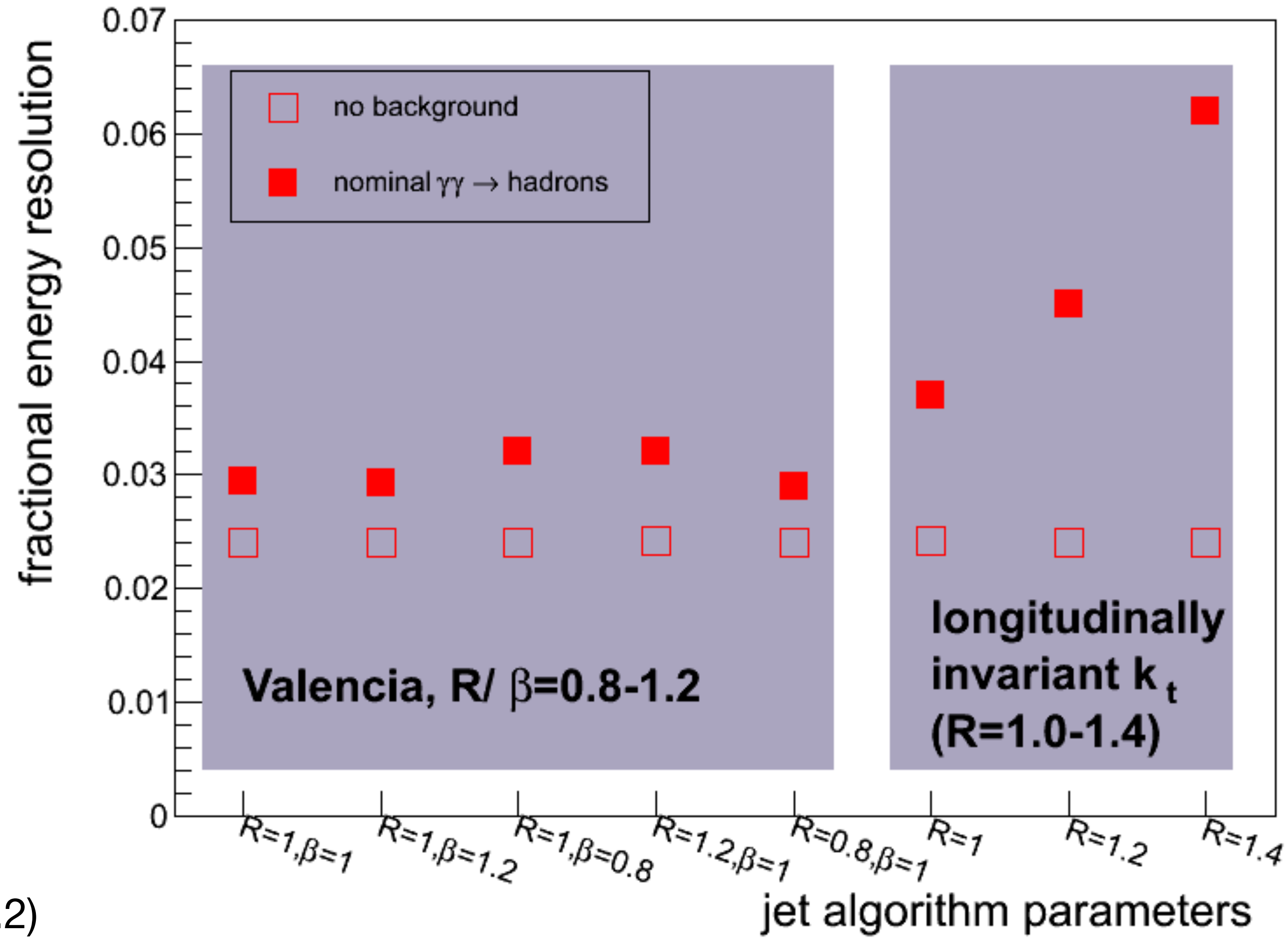
Note: resolution considers reconstructed energy versus stable particle jets; relative to the actual top parton the energy resolution is 5% and the width of the mass peak ~7%

Boosted top quarks



CLIC 3 TeV $e^+e^- \to tt$
Adding $\gamma\gamma \to$ hadrons background
CLIC-ILD detector simulation
PANDORA PFA + quality and timing cuts
Valencia e^+e^- jet algorithm (N_j =2, R=1, b=1.2)
Significantly better now than long. inv. k_t with R=0.8-1.2

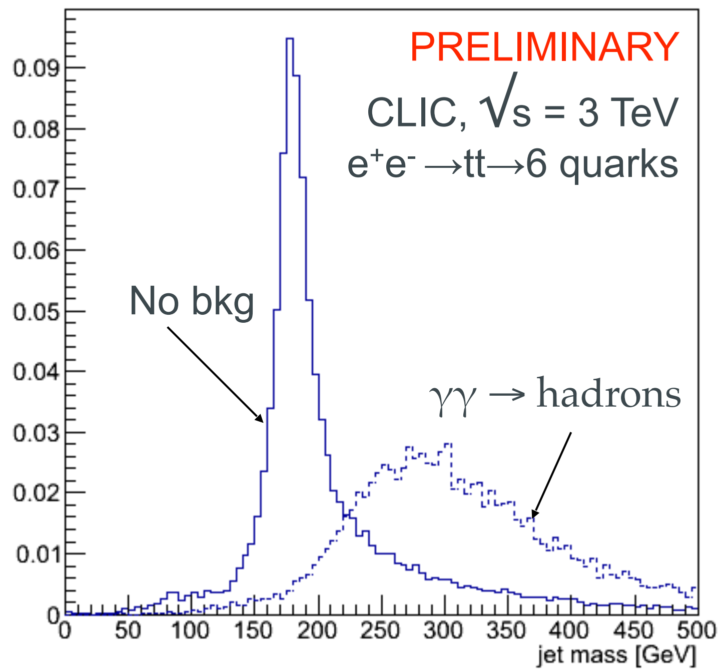
Background has impact on fat jets:
Energy resolution degraded 2.4% → 2.9%



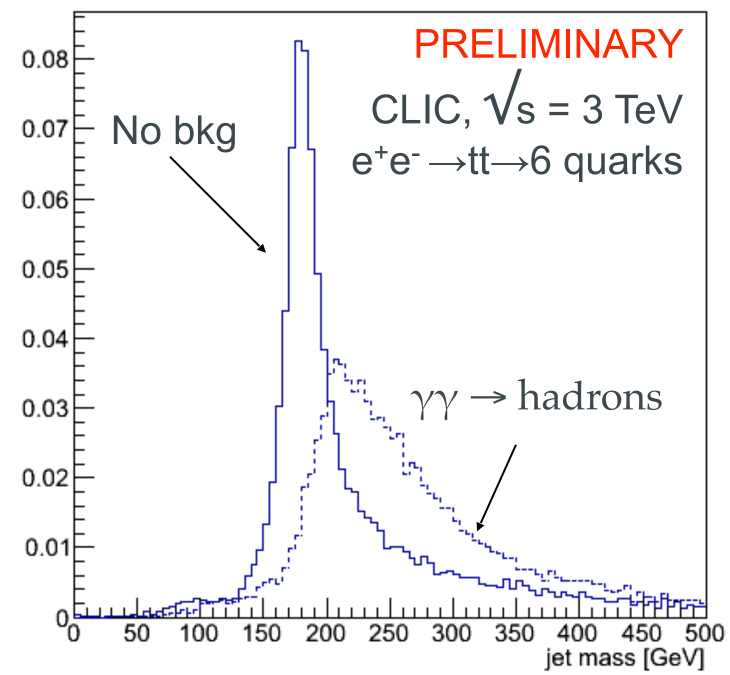
Note: particle jets used to determine resolution do not contain particles from $\gamma\gamma \to$ hadrons

Boosted top quarks

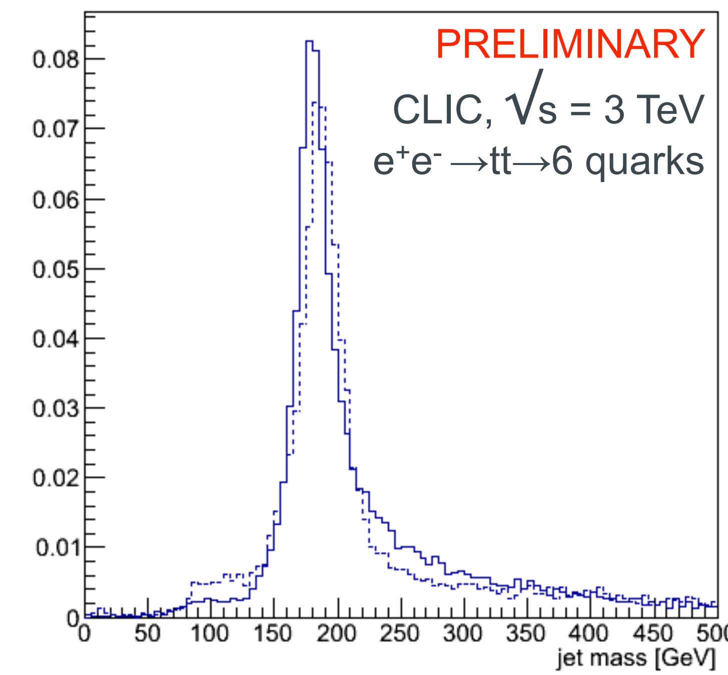
Longitudinally invariant k_t (R=1)



Valencia (R=1, β=1)



Valencia trimming



With $\gamma\gamma \rightarrow$ hadrons background

Background has a profound impact on fat jet substructure:
Raw jet mass resolution badly degraded (from dream 3.2% to nightmare 16%)

Preliminary: grooming jets restores jet mass resolution to ~4%
Results correspond to a primitive e^+e^- variant of trimming based on 3+3 Valencia R=0.2 jets → optimisation needed