Jet Reconstruction at the LC

Junping Tian, Keisuke Fujii (KEK)

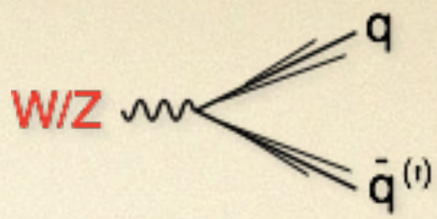
Workshop on Top physics at the LC 2015, June 30 - July 2 @ IFIC, Valencia

outline

- brief introduction to jet clustering
- new challenges at future e+e- colliders
  - ongoing studies on beam jet removal
  - efforts towards new color-singlet clustering

Many thanks to I. Garcia for providing some material of nice study in Valencia group, see his talk @ CLIC Workshop 2015

introduction to jet clustering



motivation: reconstruct parton level information of main hard process which is calculable using perturbative QCD

requirement: infrared safe — well behaved for the soft and collinear limit

sequential type

i) define inter-particle distance (jet resolution variable) y_{ij}

ii) find the smallest y_{ij} and combine particle i and j into one pseudo-particle

iii) iterate the recombination process until $y_{ij} > y_{cut}$ or #jet = #fixed

global type

- Cone algorithm (Sterman-Weinberg jet): maximize the energy inside one fixed cone
- Georgi algorithm (arxiv: 1408.1161 / 1408.3823): maximize Jet function

$$J_\beta(P_\alpha) \equiv E_\alpha - \beta \frac{P_\alpha^2}{E_\alpha} = E_\alpha \left[(1-\beta) + \beta v_\alpha^2\right]$$

introduction to jet clustering: sequential algorithms at e+e-

Algorithm		resolution scale	comment
Jade	J	$2E_iE_j(1-\cos\theta_{ij})$	
Durham	D	$2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})$	
Durham (Luclus $k_\perp$) (Also Durham/Lu)	DL	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	Luclus without reassignment and preclustering
Luclus	L	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	
Geneva	G	$\frac{8}{9}E_{\text{vis}}^2\frac{E_iE_j(1-\cos\theta_{ij})}{(E_i+E_j)^2}$	
Angular-Ordered Durham	A	$2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})$	Cambridge without soft-freezing
Cambridge	C	$2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})$	
Cambridge (Luclus $k_\perp$)	CL	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	
Diclus mode 0	Di0	$\frac{(s_{ji}-(m_i+m_j)^2)(s_{ik}-(m_i+m_k)^2)}{s_{ijk}}$	$3\to 2$ clustering
Diclus mode 1	Di1	$\frac{(s_{ji}-(m_i+m_j)^2)(s_{ik}-(m_i+m_k)^2)}{s_{ijk}}$	largest initial cluster retains its direction
Diclus mode 2	Di2	$\frac{s_{ji}s_{ik}}{s_{ijk}}$	largest initial cluster retains its direction

S. Morreti, et al., JHEP08 (1998) 001

introduction to jet clustering: examples of improvement

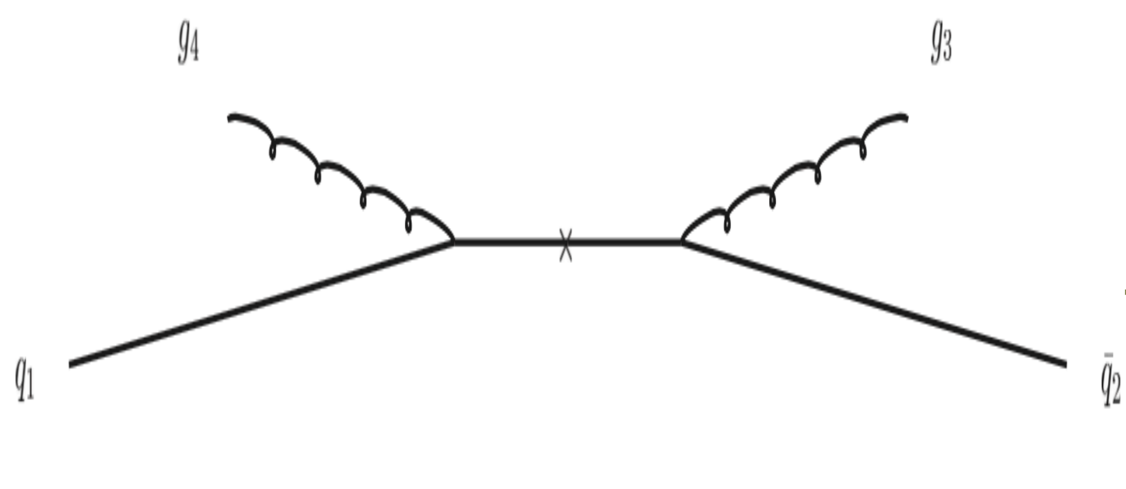


Figure 2: The seagull diagram with $E_3, E_4 \ll E_1, E_2$ and $\theta_{14}, \theta_{23} \ll \theta_{34}$.

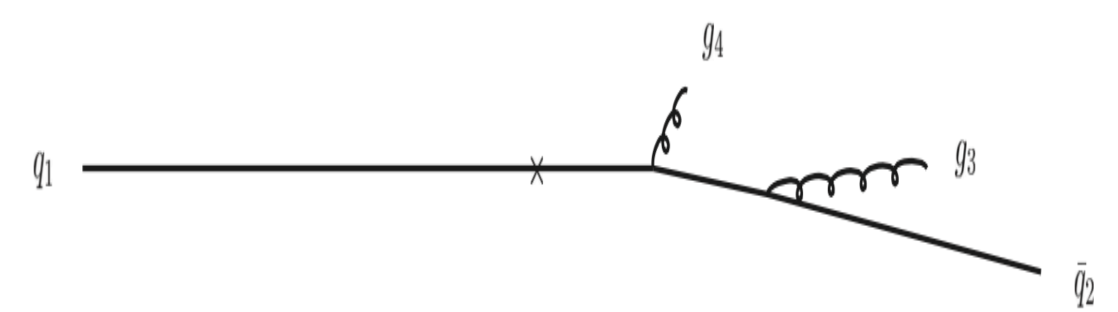


Figure 3: Parton branching with 'unresolved', soft, large-angle gluon emission g_4. Here, one has the following configuration: $E_2 \gg E_3 \gg E_4$, and $\theta_{23} \ll \theta_{24} \approx \theta_{34}$.

check complete reviews by

S. Morreti, et al., JHEP08 (1998) 001
A. Ali, et al. arxiv: 1012.2288

JADE: $$y_{ij} = \frac{2E_iE_j(1-\cos\theta_{ij})}{E_{\mathrm{vis}}^2}$$

remedy unnatural soft jets by $g_3 g_4$

Durham: $$y_{ij} = \frac{2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})}{E_{\mathrm{vis}}^2}$$

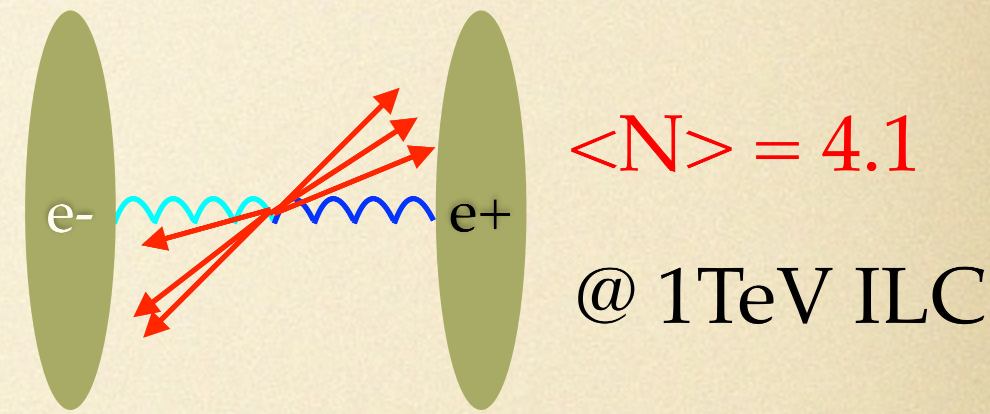
remedy soft large angle g_4

Angular-Ordered Durham:
introduce angular measure v_{ij}, start pair with smallest angle, $v_{ij}=2(1-\cos\theta_{ij})$

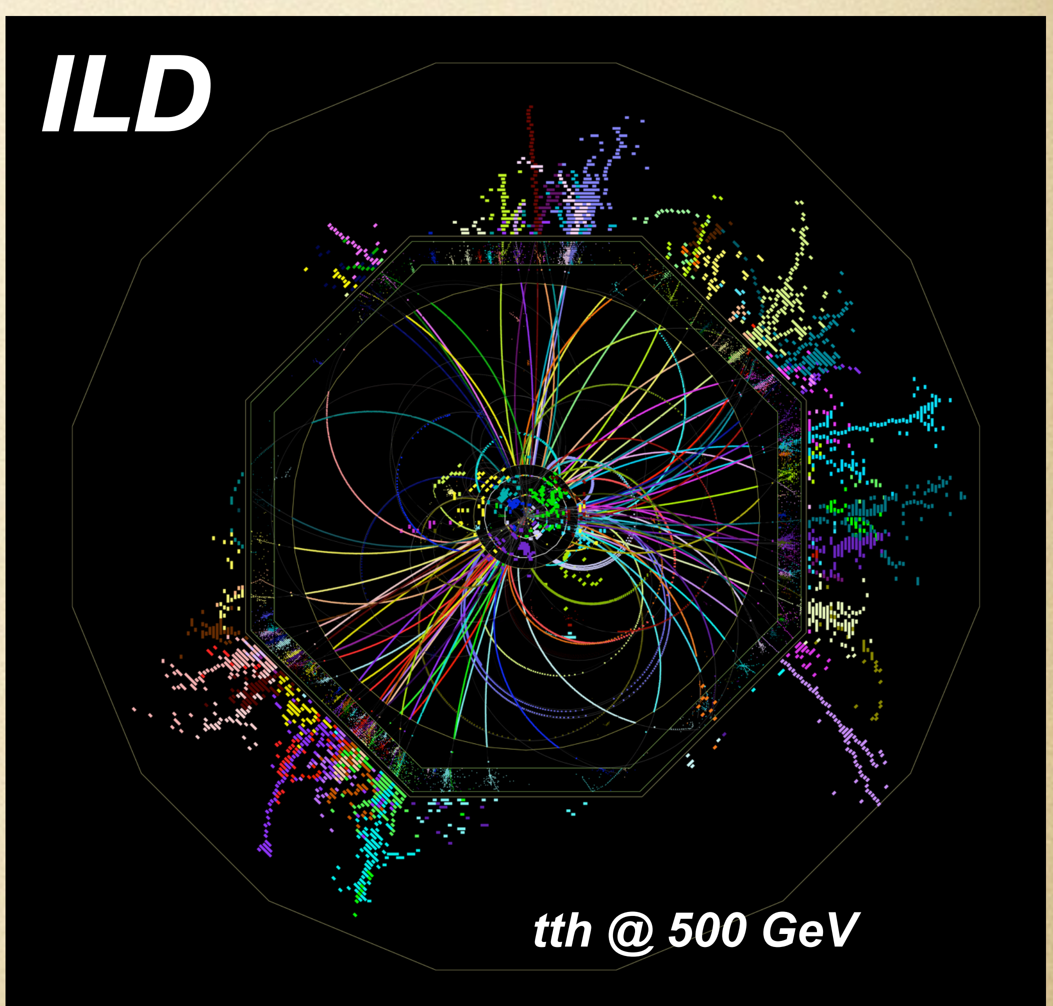
Cambridge: "soft-freezing"

new challenges at next high energy e+e- colliders

- beam jets from underlying events, mainly $\gamma\gamma$—>hadrons: not all particles reconstructed in one event are from interested hard processes

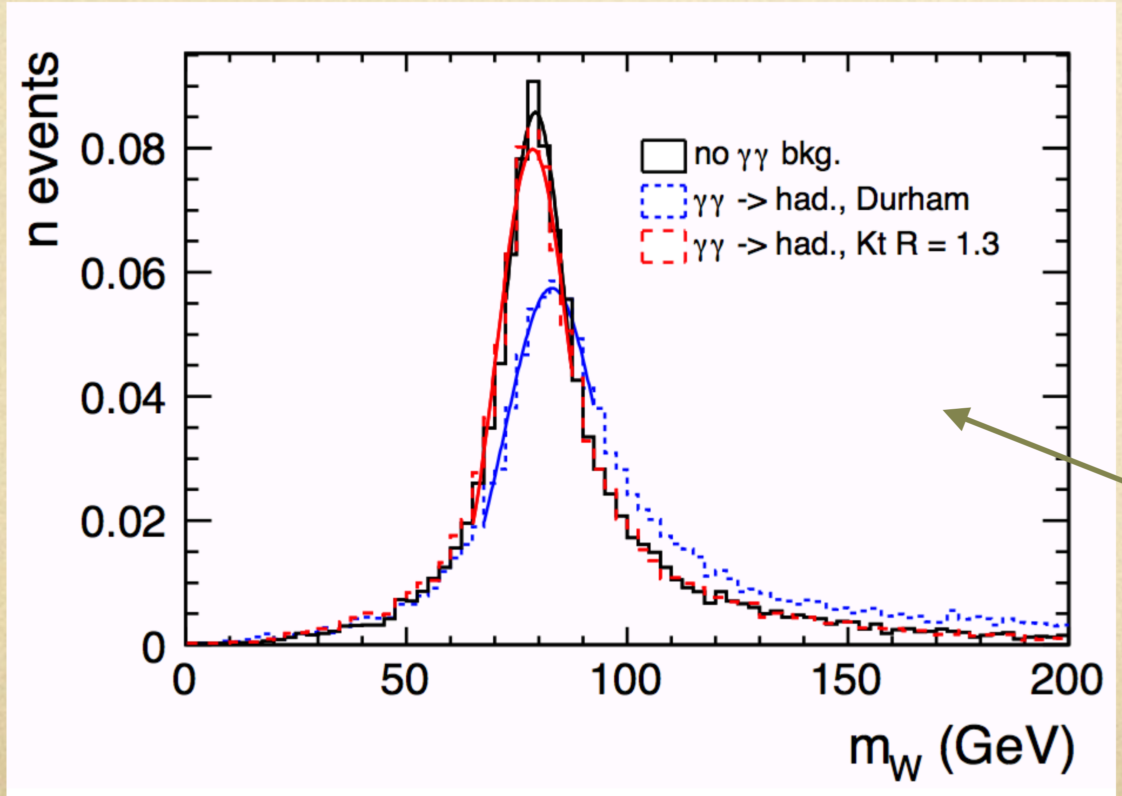


- more jet multiplicities: tt-6jet, ZHH-6jet, ttH-8jet, etc.
- higher available Q^2: deeper parton shower process
- possible y_{ij} inside one jet larger than y_{ij} between two jets —> mis-clustering due to jet overlap



efforts to handle the beam jets — Valencia jet-clustering

- Durham algorithm at e+e- has been adapted to hadron collider where beam remnant is more relevant (quite long time ago): longitudinal invariant k_t algorithm (anti-k_t).
- recently, Valencia algorithm has been proposed to combine good features of lepton colliders (Durham like distance).



longitudinal inv. k_t

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\,\Delta R_{ij}^2/R^2$$

$$d_{iB} = p_{ti}^2$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

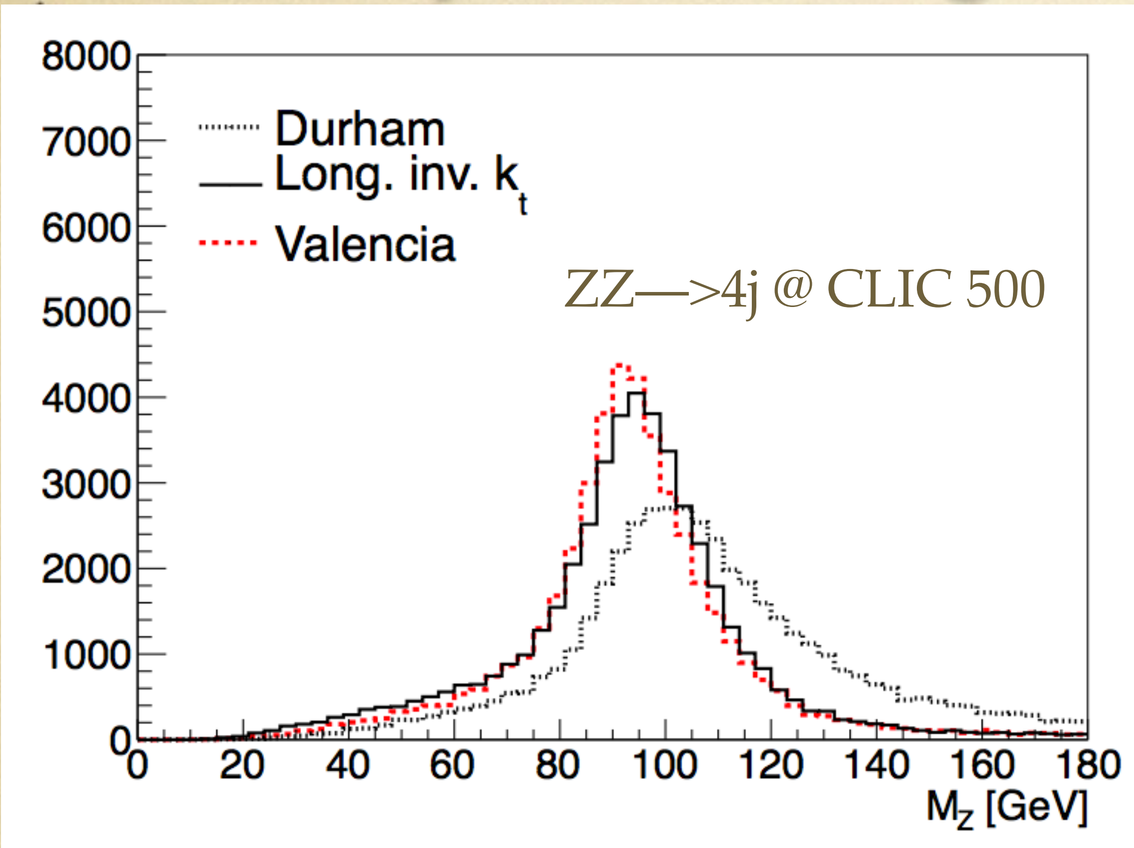
Valencia algorithm

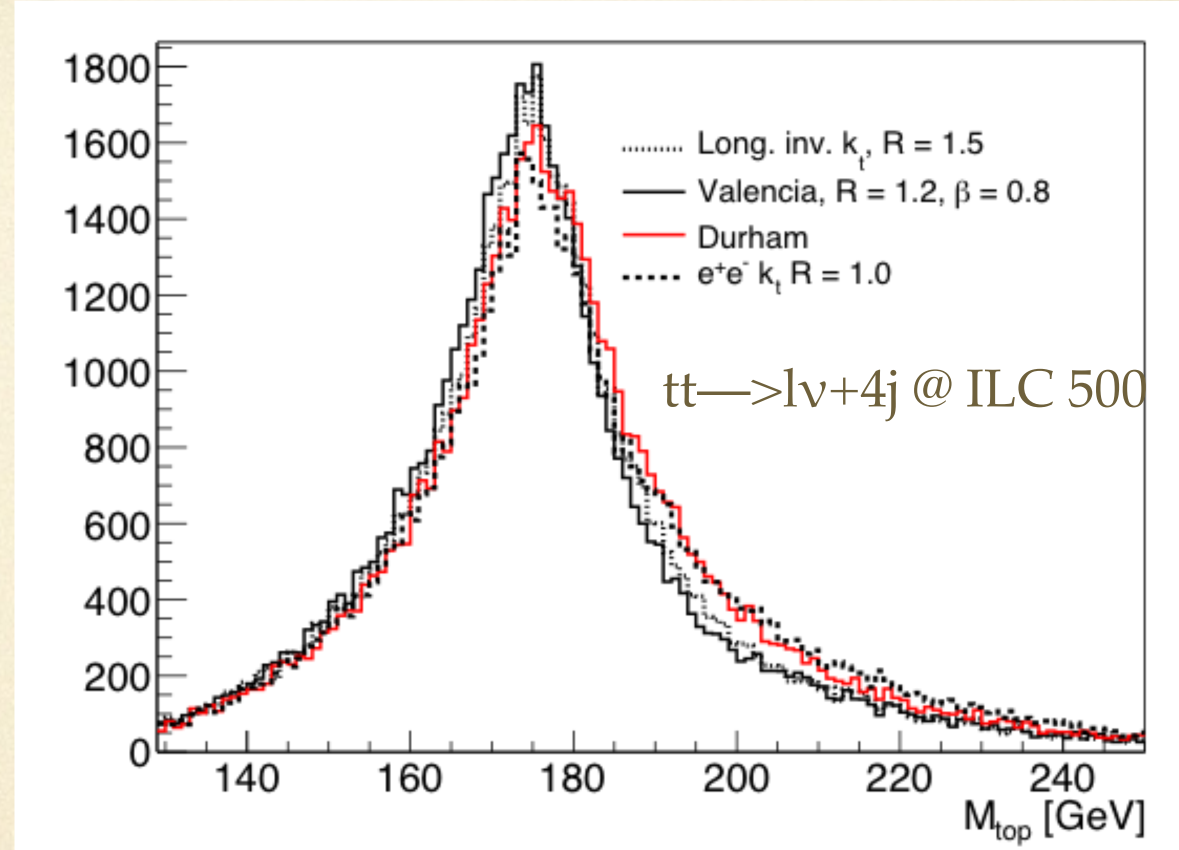
$$d_{ij} = min(E_i^{2\beta}, E_j^{2\beta})(1 - \cos\theta_{ij})/R^2$$

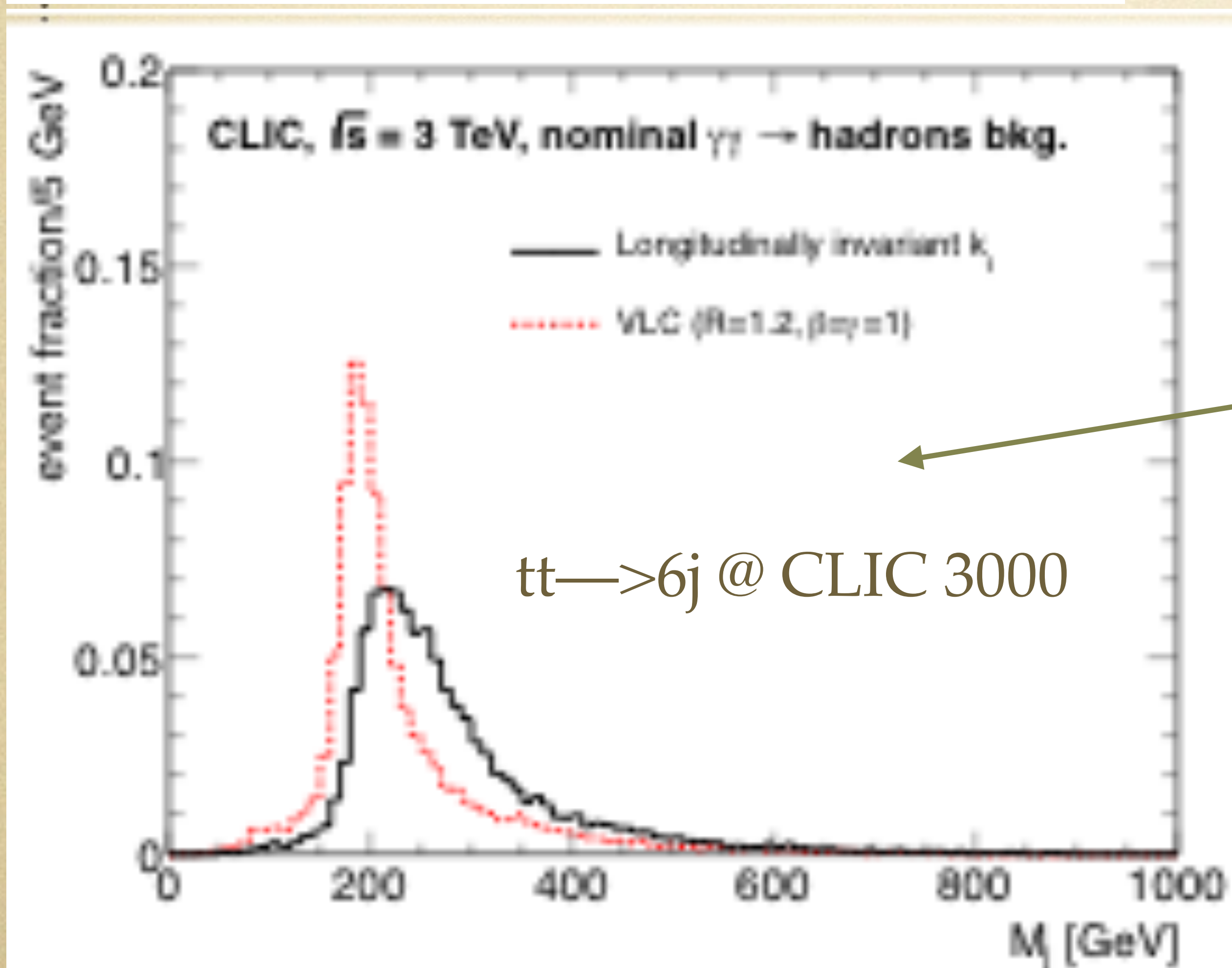
$$d_{iB} = p_{ti}^{2\beta} = E_i^2 \sin^{2\beta}\theta_i$$

benchmark using k_t for WW—>lvqq in ILD-DBD

Valencia jet-clustering: more applications





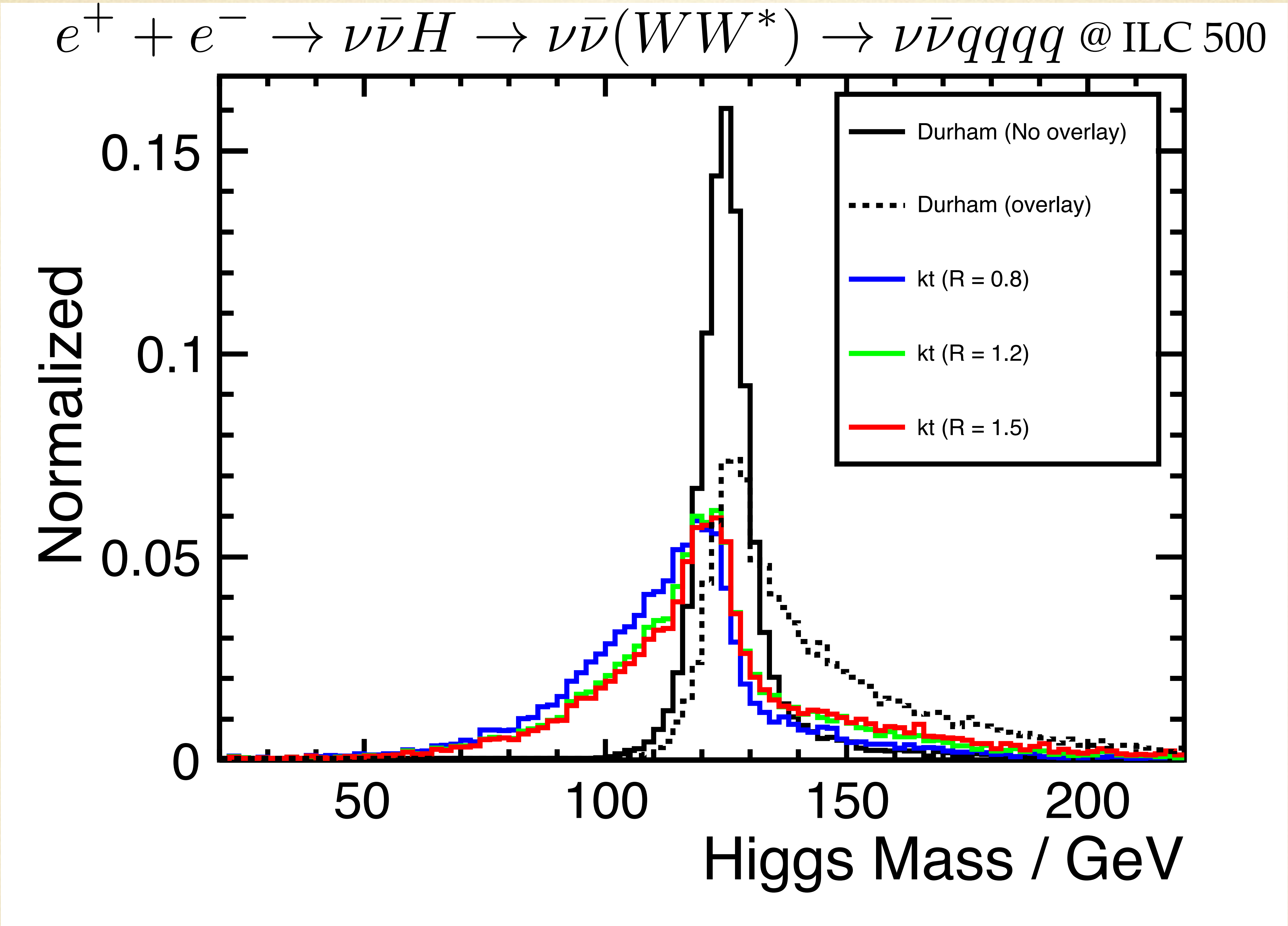


VLC is significantly better

I. Garcia @ CLIC Workshop 2015

M. Vos, et al., arxiv: 1404:4294

impact of γγ—>hadrons overlay: one more example

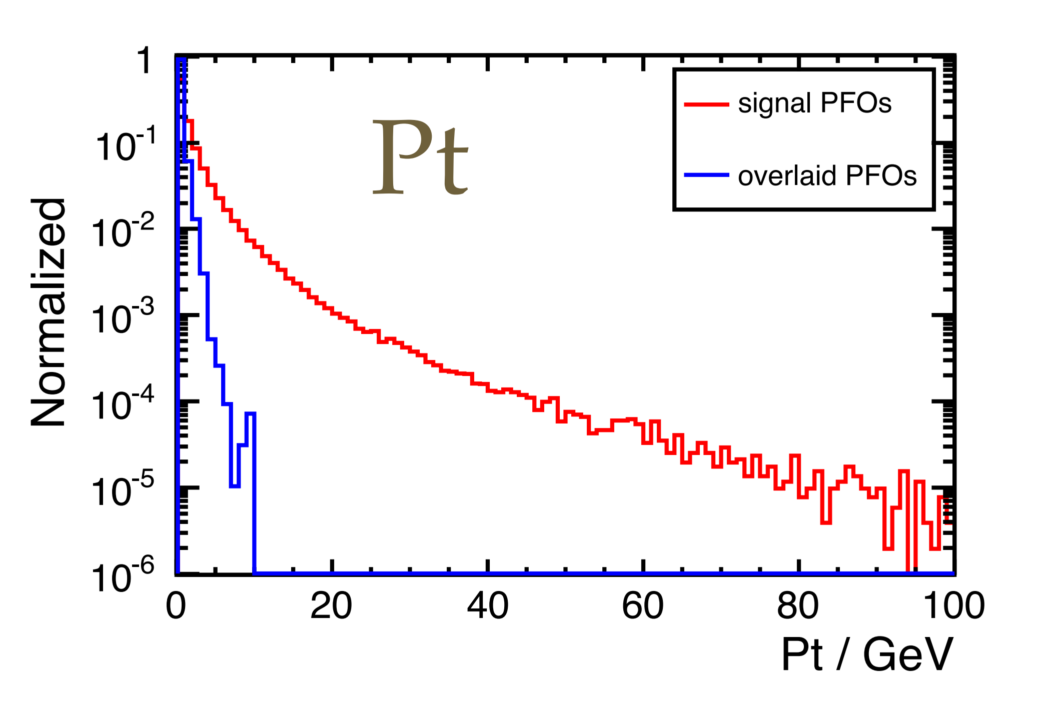


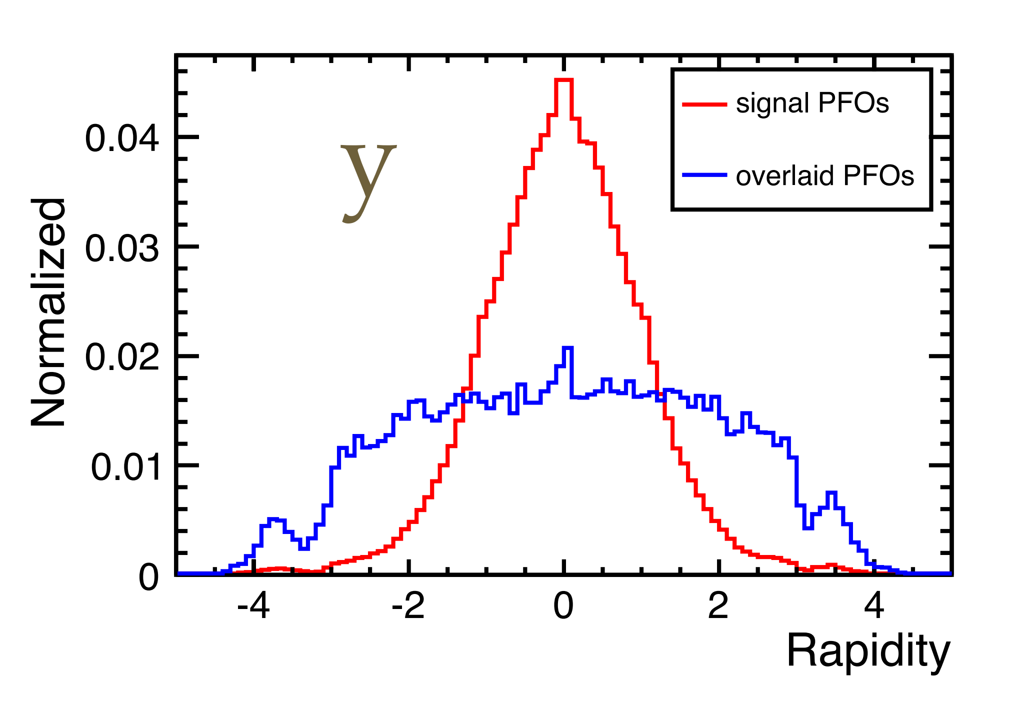
longitudinal inv. kt is barely working in
WW-fusion channel H—>WW*—>4j

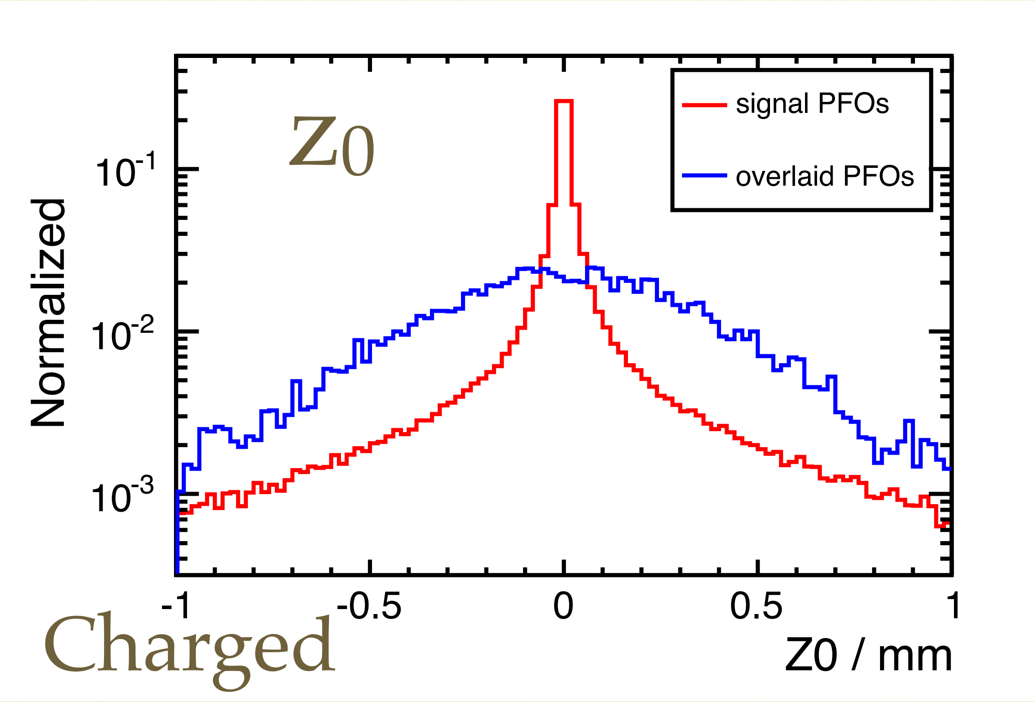
efforts to handle the beam jets at LC— alternative approach

- the beam remnant at e+e- should be much simpler than those at hadron collider —> in principle we should be able to reconstruct those remnant exclusively without generic clustering algorithm.
- the overlay events, though in the same bunch, have a shifted IP in z-direction to the primary IP of signal events.
- the bunch spread in z at ILC ~ 300μm; track impact parameter resolution ~ 10μm; if the low momentum tracking and vertex finder work well, we would be able to reconstructed the IP of those overlay events when the shift is more than 10s of μm.
- as a first step approach, try to check how well we can separate signal and overly particle by particle

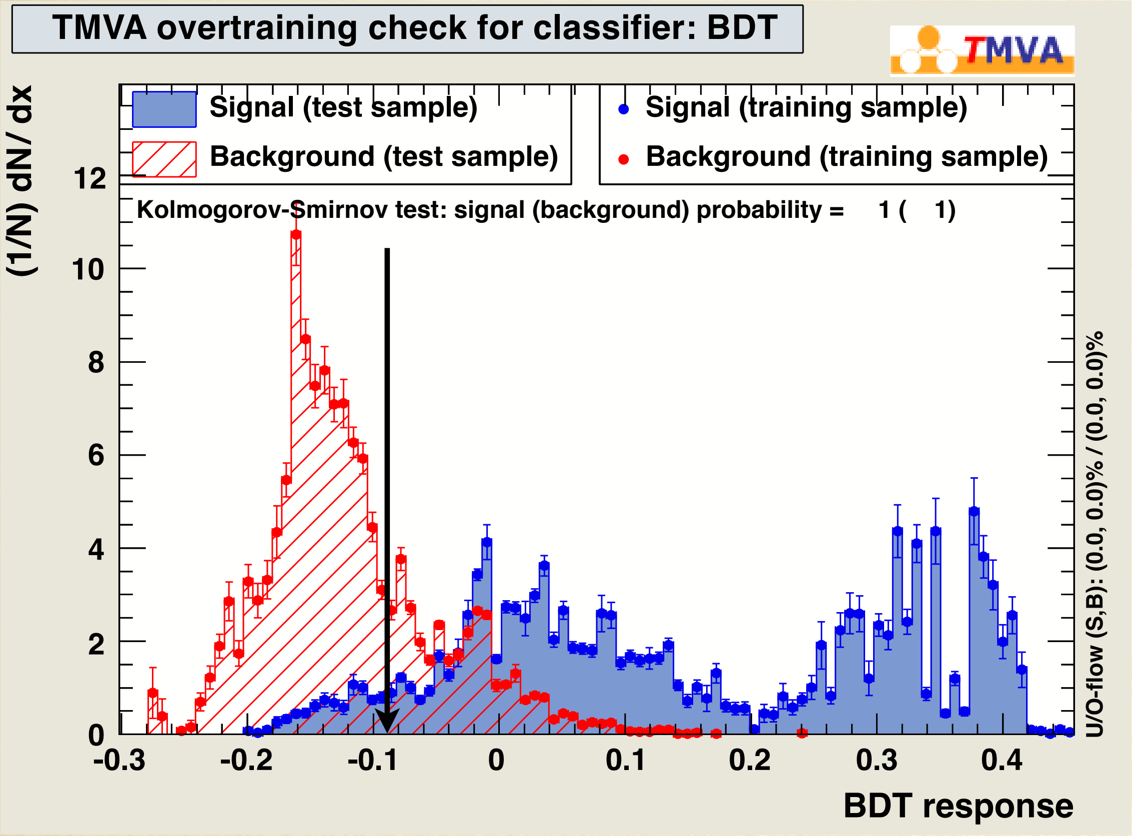
characteristics of particles from overlay events





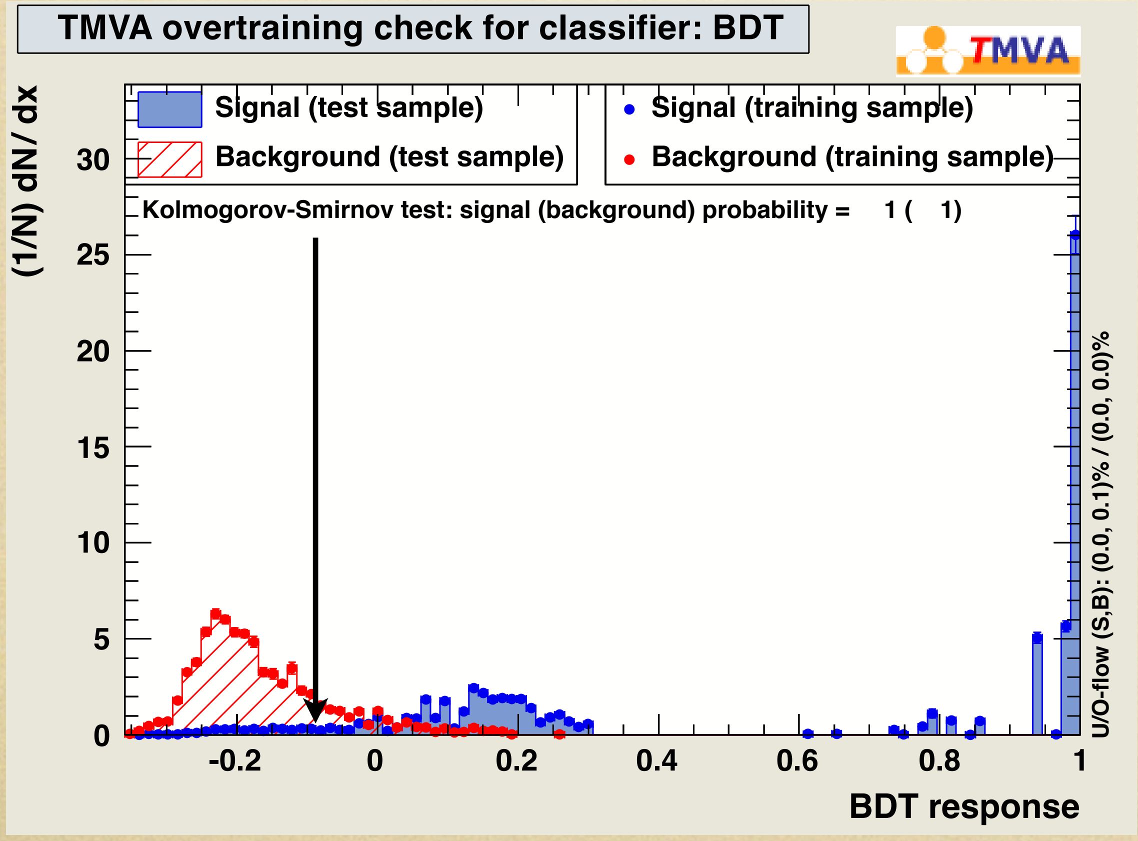


MVA to separate signal particle and overlay particle



category 1: neural or large z0

input: Pt, Rapidity

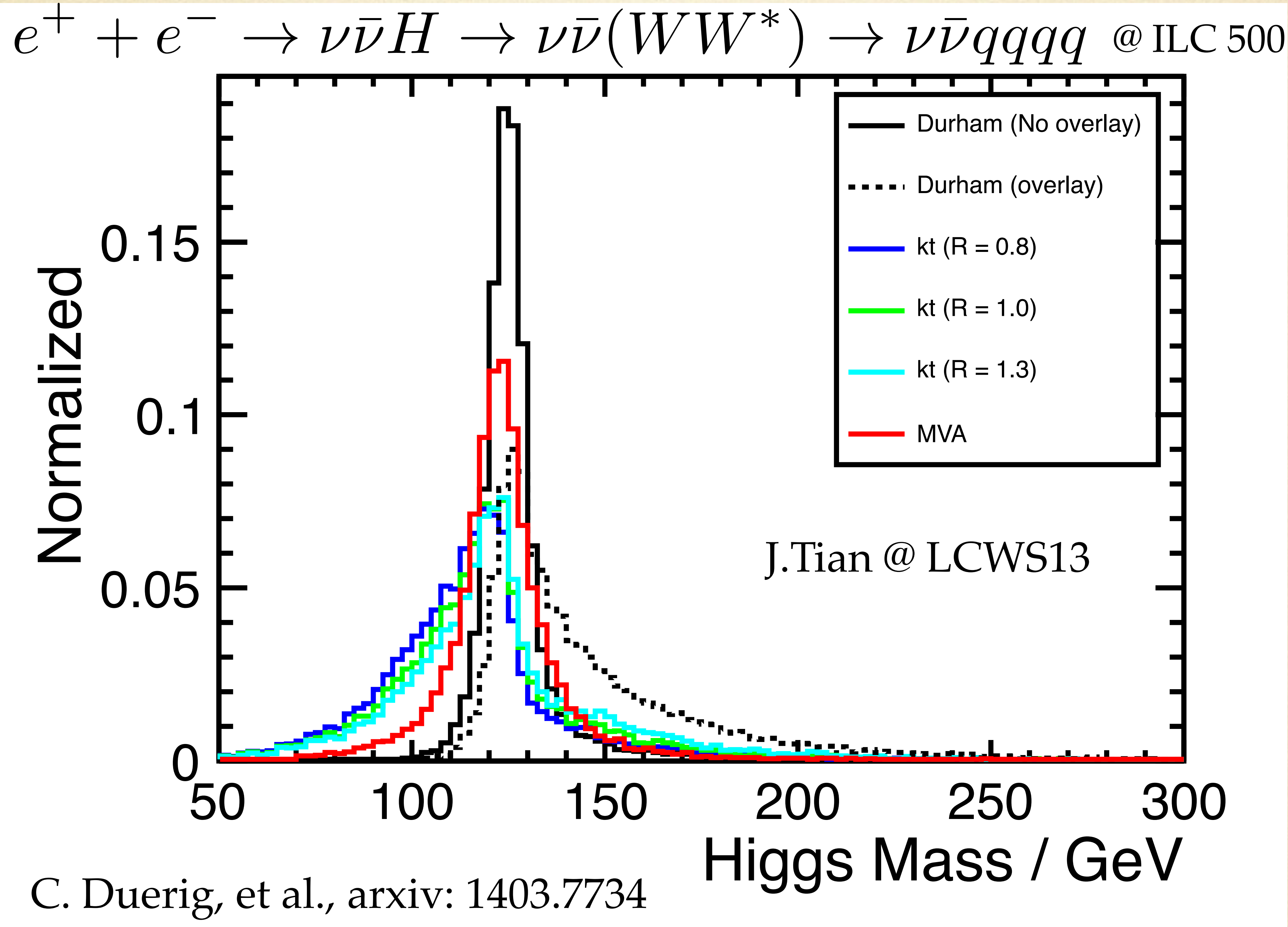


category 2: charged and small z0

input: Pt, Rapidity, z0

each PFO weighted by energy in both cases

efforts to handle the beam jets at LC— MVA approach



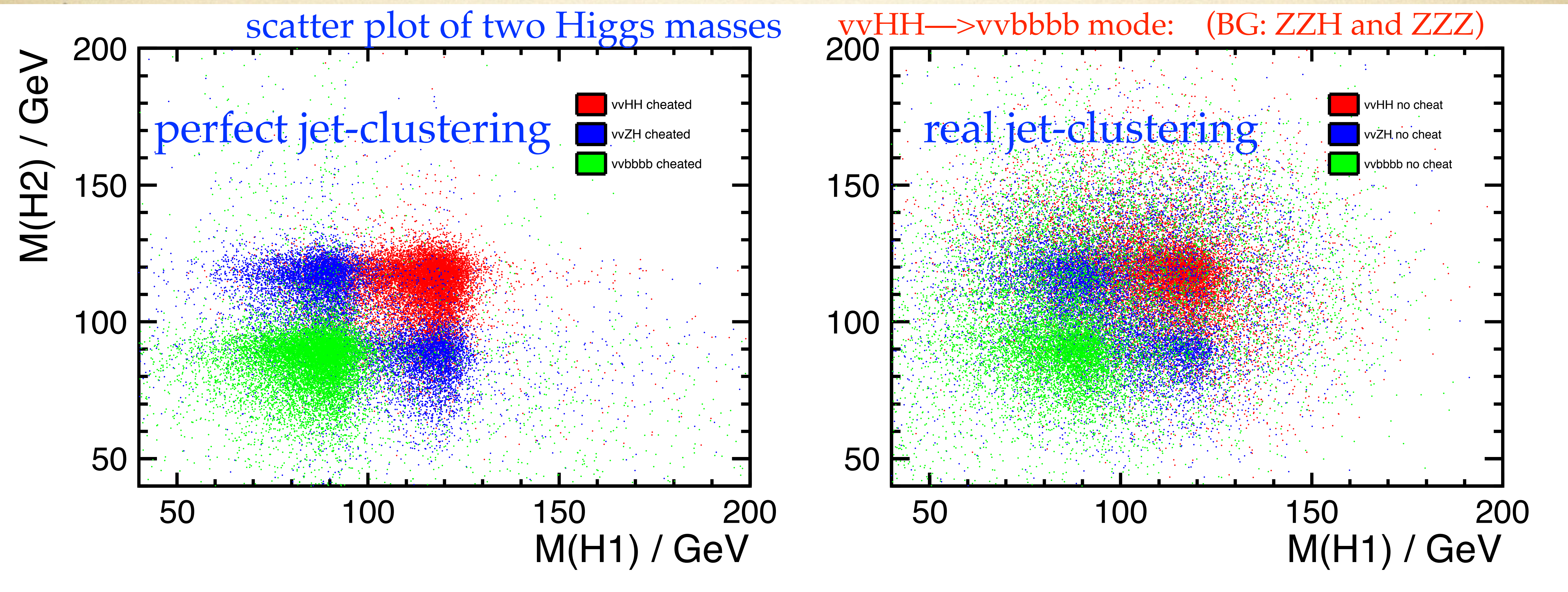
as a first step, MVA seed particle tagging already gives better performance in H—>WW* process; need generalize to clustering algorithm and add vertex seeds, and check whether it's better in other channels; ongoing

towards to a generic color-singlet jet clustering

- first of all this idea is in a conceptual level, feel free to start your coffee break now,
- and I'm an experimentalist, feel free to ignore what's I'm going to propose towards a very ambitious new clustering algorithm.
- but, we do need help from you, in particular QCD experts.
- starting point: why do need a much better color-singlet clustering

impact of jet-clustering in Higgs self-coupling measurement

(without beam overlay now)



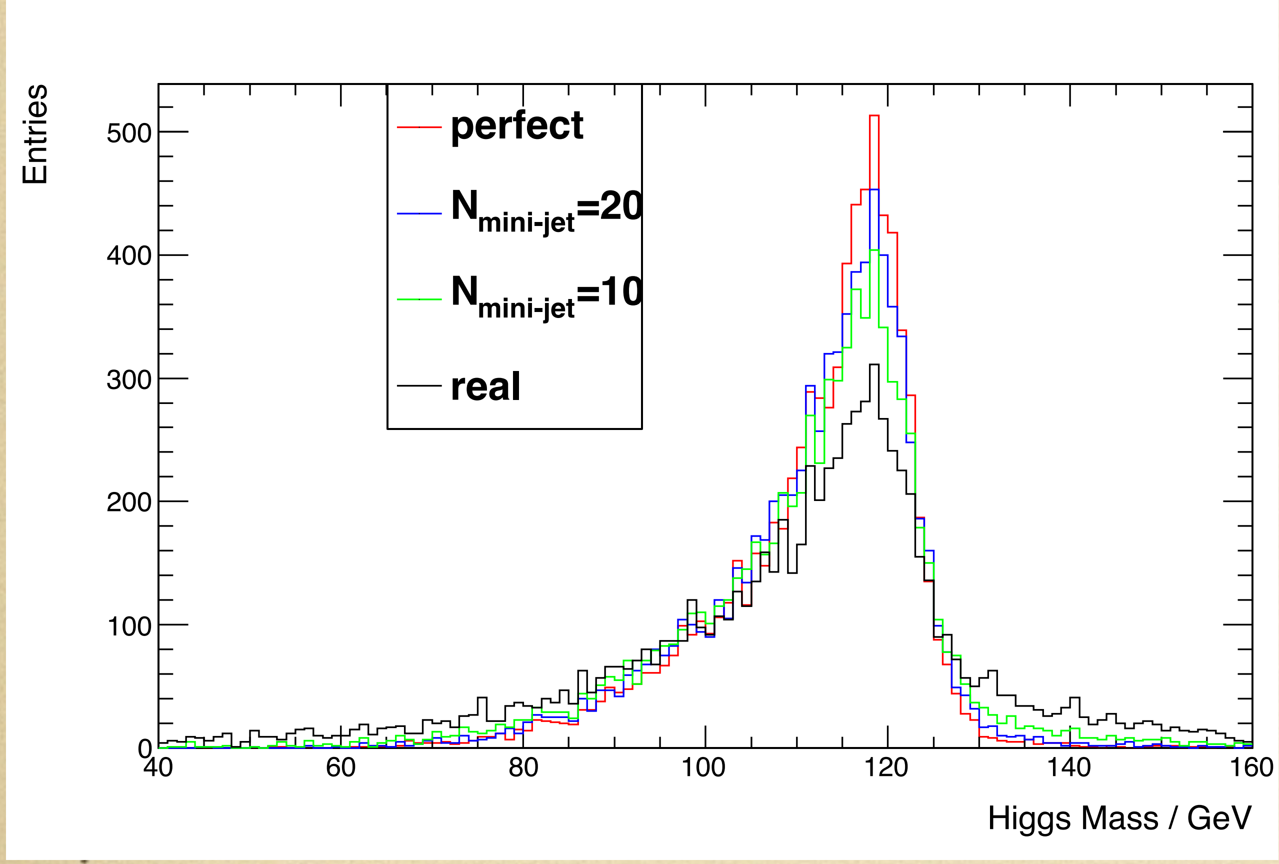
it has been studied if a color singlet jet clustering can be implemented for both signal and BG, λ_{HHH} measurement can be improved by 40%, which means 20% $\delta\lambda_{HHH}/\lambda$ (5σ) would already be possible at 500 GeV ILC with the H20 scenario.

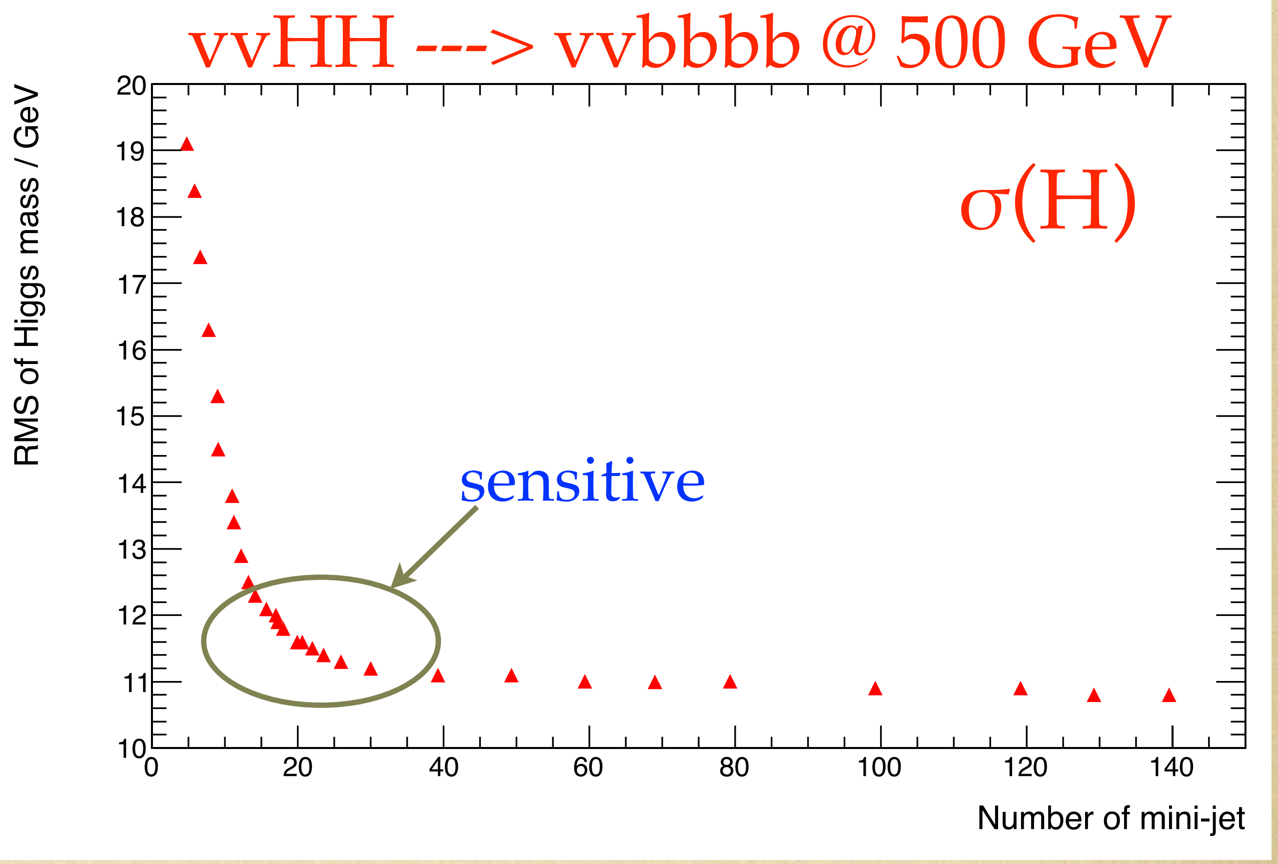
investigation so far: when does mis-clustering start?

i) mini-jet clustering (pre-clustering) with realistic Durham algorithm, stop at certain fixed number of jet

ii) combine those mini-jets using MC truth information to two Higgs (color-singlet)

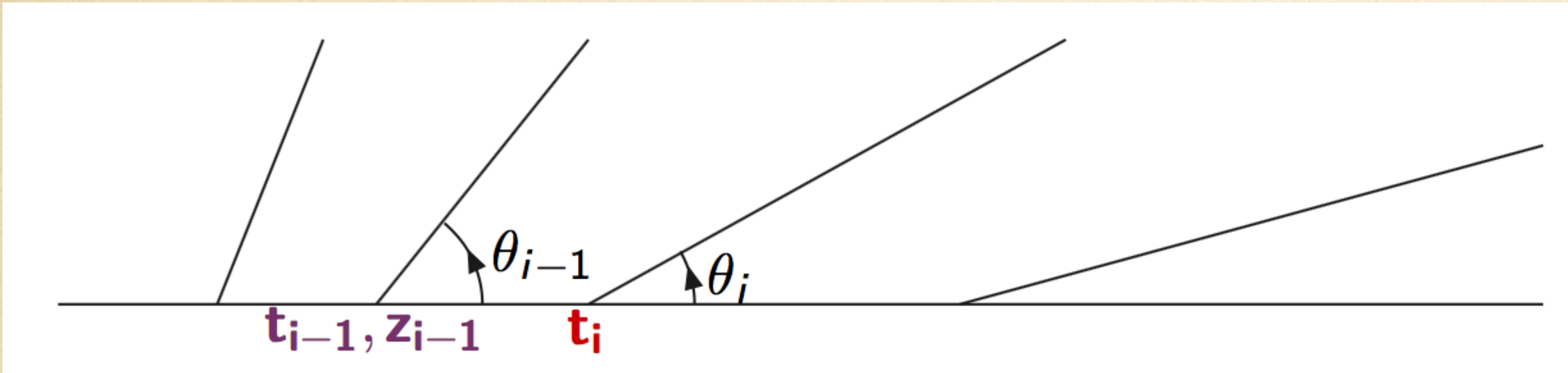
iii) check the dependence of Higgs mass with the number of mini-jets





Durham jet clustering starts to have major mis-clustering when remaining #mini-jet ~ 20

proposal of new clustering: reconstruct the parton shower likelihood



$$
\begin{aligned}
P_{q\to qg}(z) &= C_F\,\frac{1+z^2}{1-z}\,, \\
P_{g\to gg}(z) &= N_C\,\frac{(1-z(1-z))^2}{z(1-z)}\,, \\
P_{g\to q\bar{q}}(z) &= T_R\,(z^2+(1-z)^2)\,, \\
P_{q\to q\gamma}(z) &= e_q^2\,\frac{1+z^2}{1-z}\,, \\
P_{\ell\to\ell\gamma}(z) &= e_\ell^2\,\frac{1+z^2}{1-z}\,,
\end{aligned}
$$

$$
\mathrm{d}\mathcal{P}_a = \sum_{b,c} \frac{\alpha_{abc}}{2\pi} P_{a\to bc}(z)\,\mathrm{d}t\,\mathrm{d}z\,.
$$

- do mini-jet clustering until #mini-jet ~ 20, during which Durham algorithm works well —> which also helps retain infrared safe
- new algorithm comes in to combine the mini-jets: if these mini-jets are actually the relic of parton shower, we should aim for the reconstruction of parton show history as the maximum information
- generic feature of parton shower can be used to help: angular ordering, $\theta_{i-1}<\theta_i<\theta_{i+1}$, $t_{i-1}<t_i<t_{i+1}$.
- assign each branching with a probability $P_{q\to qg}$, etc.
- above is intraJet parton shower; we can use some feature of color correlation interJet, such as rapidity gap, etc. (not really sure).

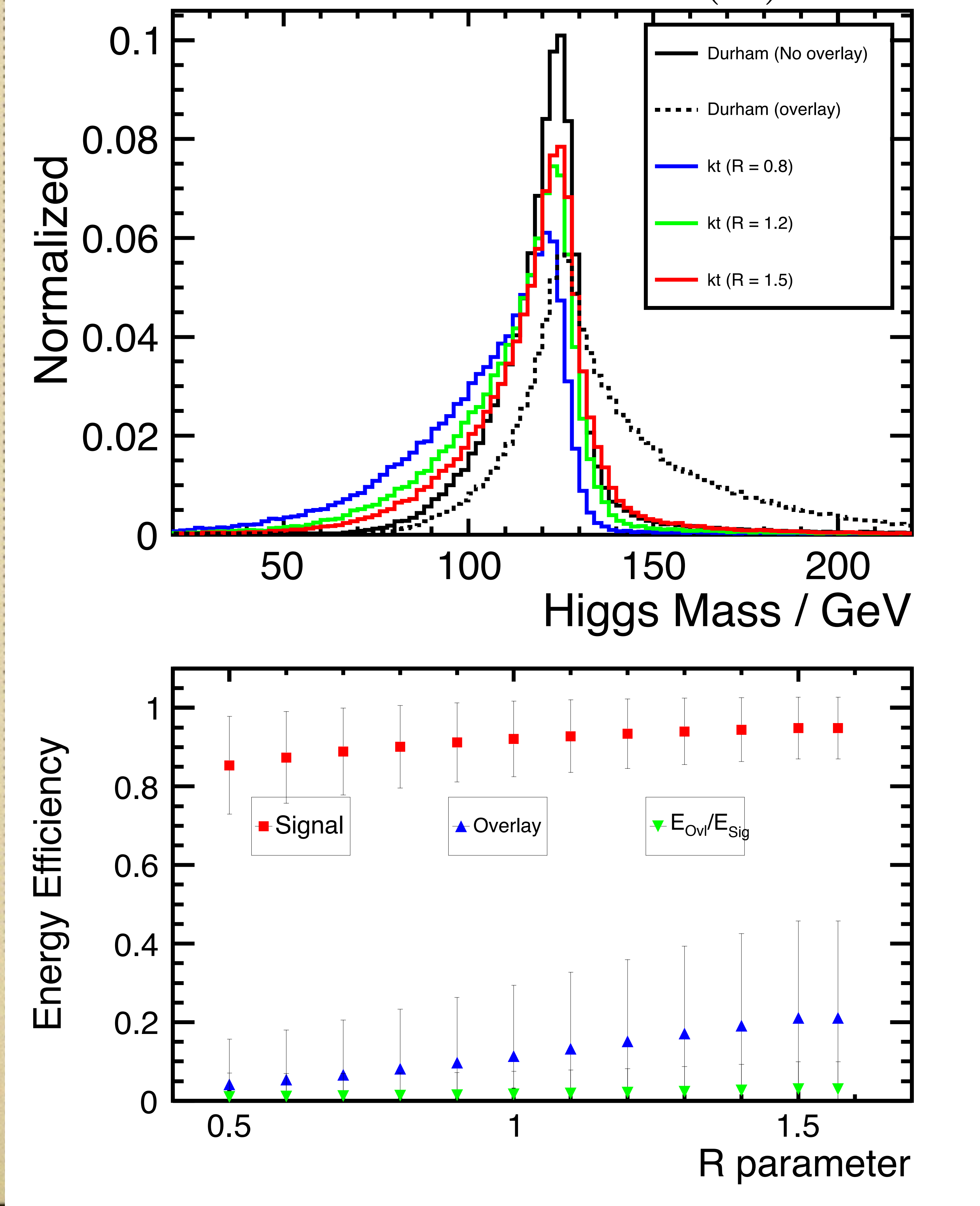
summary

- there have been many jet algorithms for e+e- colliders studied in the past decades, it's worth understanding the advantages in each of them, which could give us some useful hints.
- new challenges are expected at the next high energy e+e- collider, to handle beam jets and mis-clustering.
- improvements on removing beam jets by robust and generic Valencia clustering algorithm, and specific particle/vertex based removal method, have been observed. I would be interesting to take a look at the comparison and more applications. The vertex based method would benefit a lot from improved low momentum tracking and vertex finder.
- more general color-singlet jet clustering algorithm would help a lot for all Higgs self-coupling measurement and possibly most of others with multi-jet final states, a conceptual proposal to maximal use parton shower information is just started.

backup

typical method: kt jet-clustering

$e^+ + e^- \to \nu\bar{\nu}H \to \nu\bar{\nu}(b\bar{b})$ @ 500 GeV



$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\,\Delta R_{ij}^2/R^2$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

overlaid particles usually very forward —> large y —> far from physics jet

paras opt in kt jet clustering

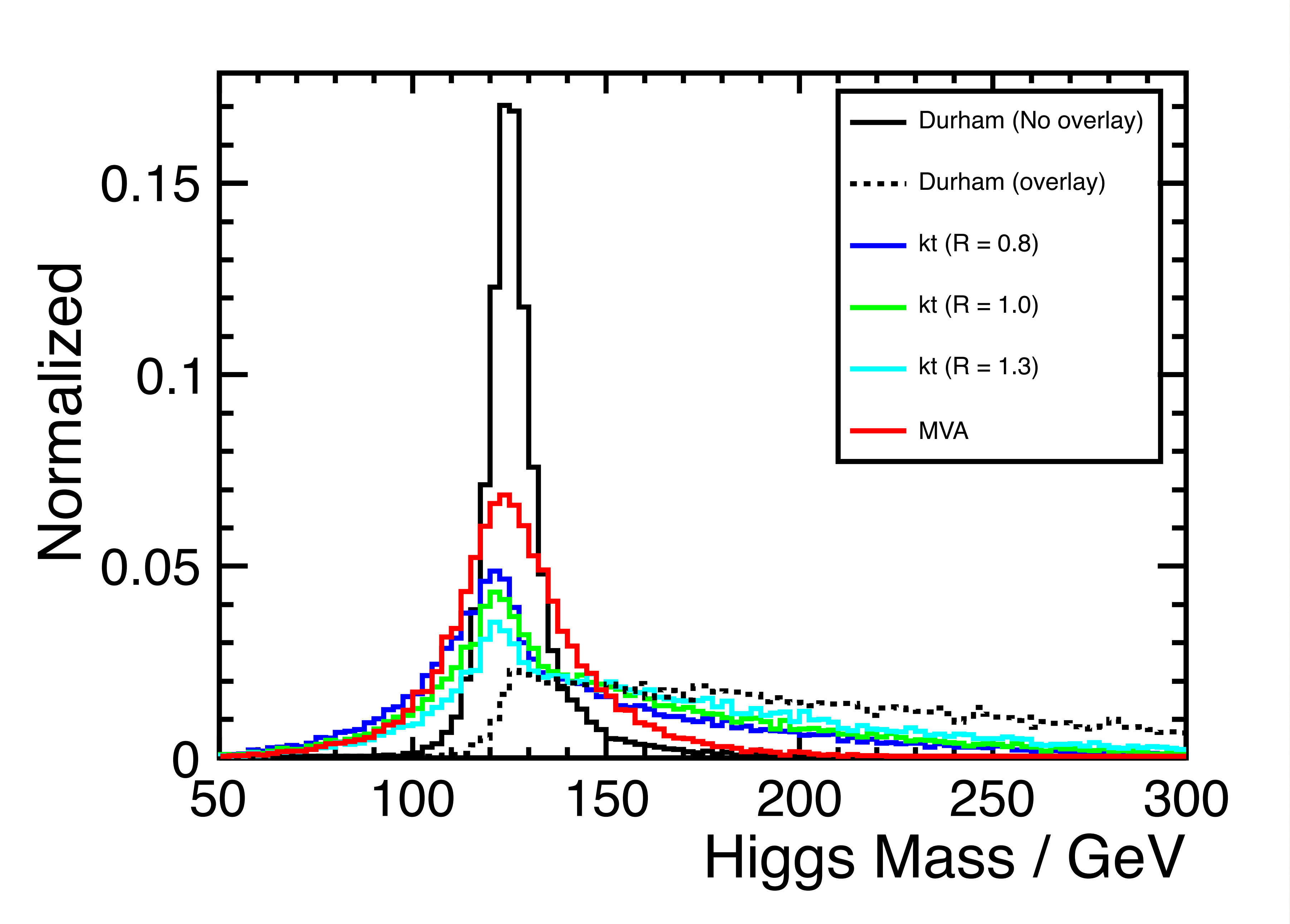
Max No. of Jets = 2

R = 1.5

overlay is removed efficiently

Eff(sig) ~ 95%
Eff(ovl) ~ 21%
purity ~ 97%

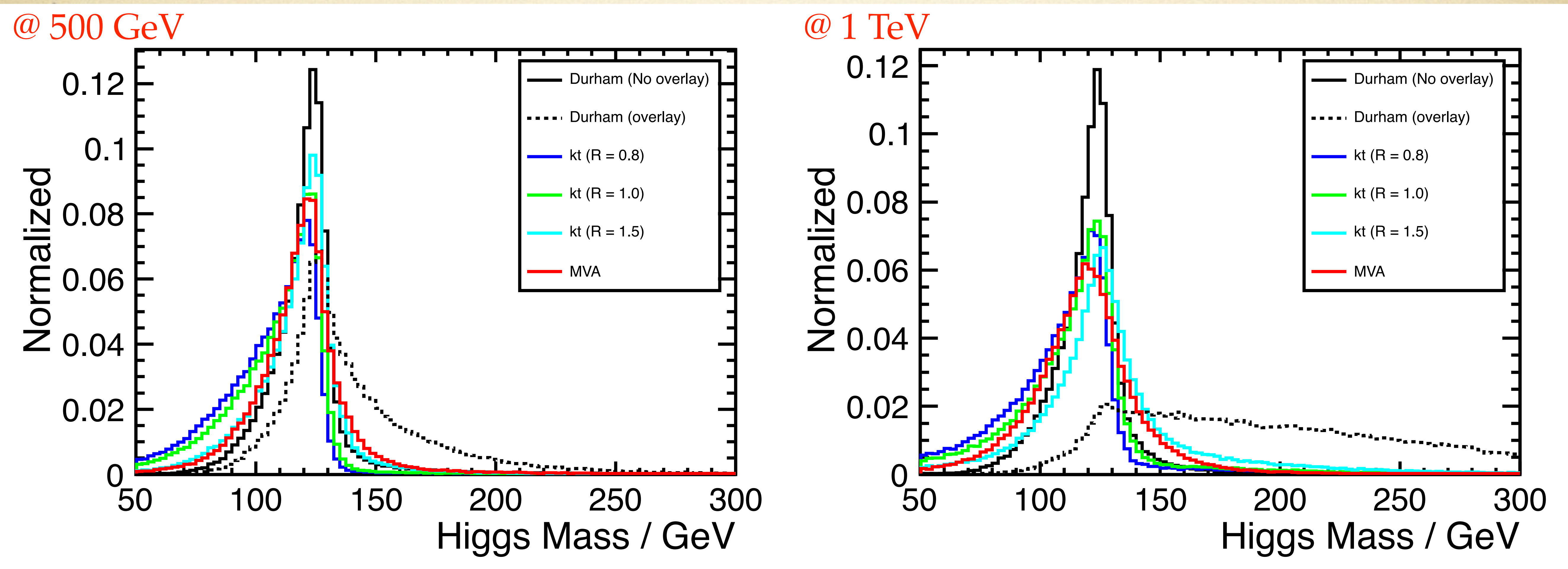
$$e^+ + e^- \to \nu\bar{\nu}H \to \nu\bar{\nu}(WW^*) \to \nu\bar{\nu}qqqq$$ @ 1 TeV



Eff(sig) ~ 89%
Eff(ovl) ~ 16%
purity ~ 95%

better than kt, but obviously not satisfactory

$$e^+ + e^- \to \nu\bar{\nu}H \to \nu\bar{\nu}(b\bar{b})$$



in this channel, MVA is not better than kt…