Top quark generation, full simulation and reconstruction with the CLD detector

Jeremy Andrea, Auguste Besson, Ziad El Bitar, Gaelle Sadowski

Introduction

- The top-quark program at FCCee is very relevant and rich, and the runs at $t\bar{t}$ threshold (and beyond), might require to perform specific detector optimisations.

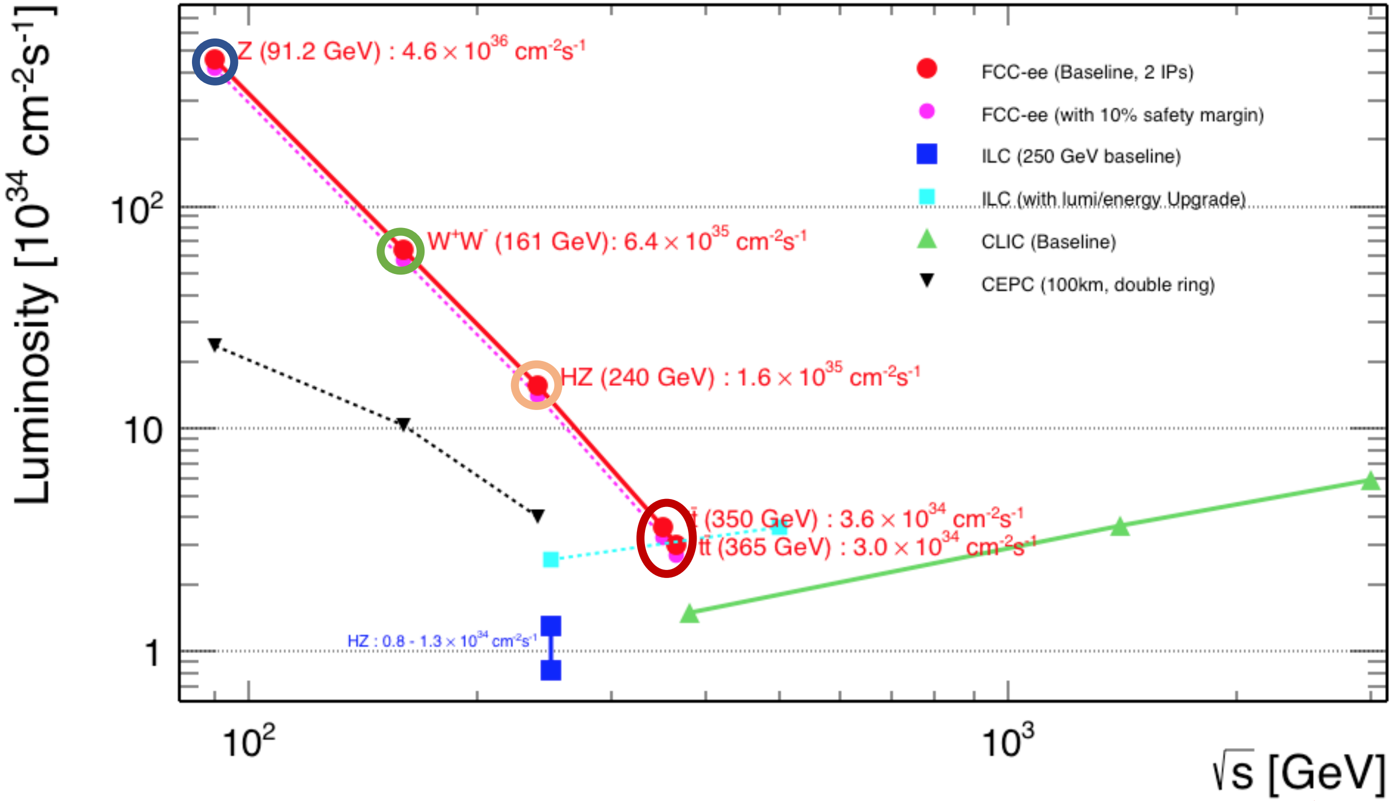
- At the FCC, sensitivity/precision studies are mainly based on fast-simulation or extrapolations from CLIC/ILC. Precise estimation and study of detector requirements and designs for physics requires full-simulation.
  - State of the art generation at highest possible orders (QCD+QED),
  - Accounting for beam-related effects : ISR, Beam Energy Spread, beam backgrounds,
  - Full detector simulation,
  - Complete event reconstruction with Particle Flow.

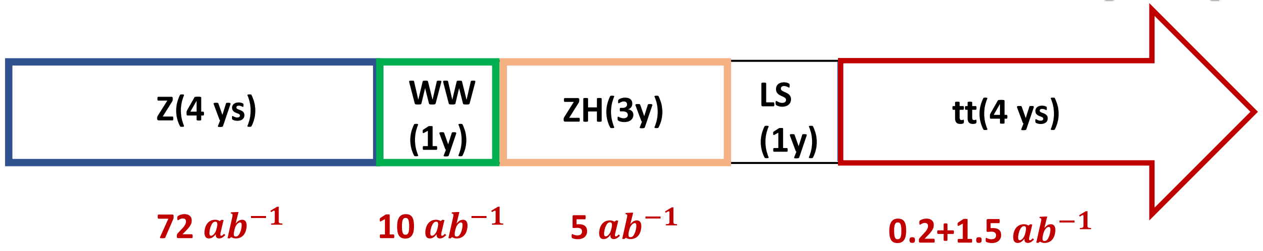
- $t\bar{t}$ events constitute a perfect playground for testing full-sim and object reconstruction/identification performances.

- Outline:
  - Quick reminder of top quark physics and production plan at FCCee,
  - Reminder of previous studies of MC generators,
  - CLD, full sim and reconstruction,
  - Results from $t\bar{t}$ fullsim : first studies and validation.

$t\bar{t}$ program at FCCee

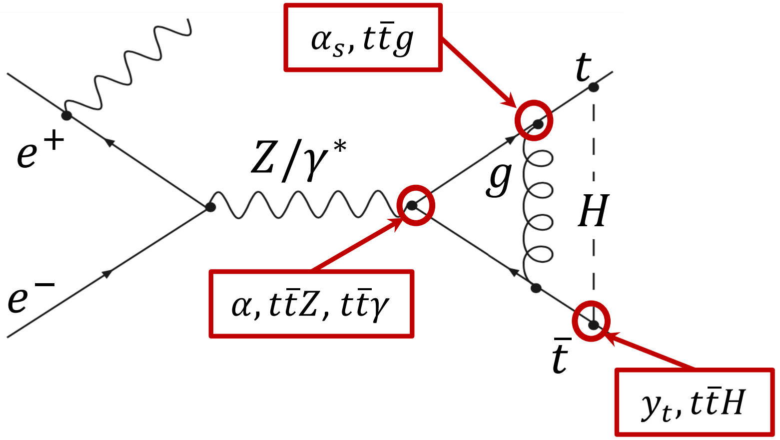
- FCCee schedule : run at different collision energies.
- Top quark physics program in two steps :
  - Scan energies from ~340 to ~350 GeV, 0.2 ab^{-1} in total,
  - Large statistic run at 365 GeV, 1.5 ab^{-1}.
- Cross section at 365 GeV ~1000 fb => ~2 Mevts in total.
- 1-2 order of magnitude lower during the scan.
- Large statistics => important to identify (and reduce) dominating systematics ➔ improve detector design, analysis and run plan.

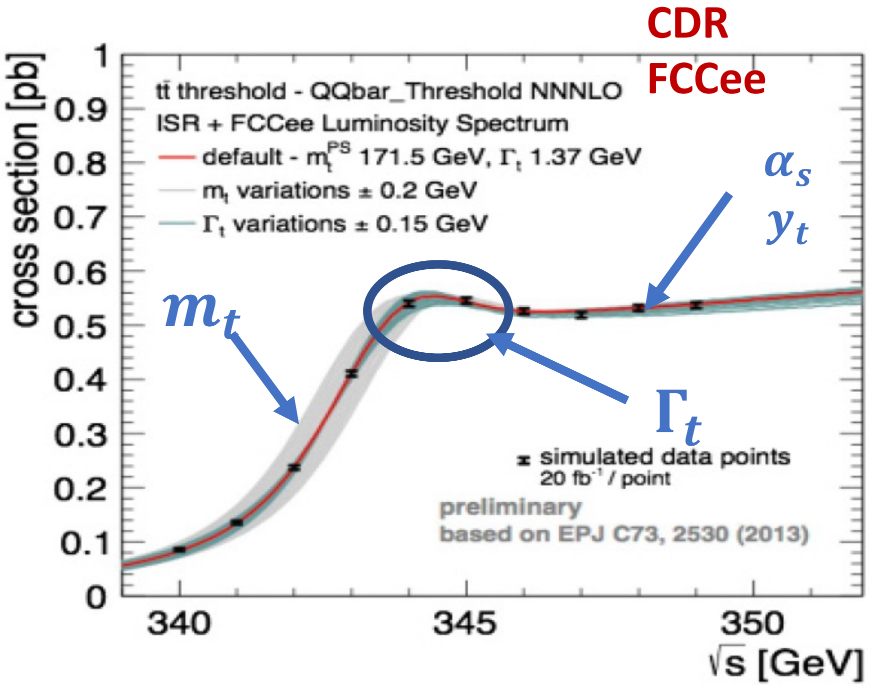




$t\bar{t}$ measurements in a nutshell

- Inclusive and differential => probe $\boldsymbol{t\bar{t}}Z$ and $\boldsymbol{t\bar{t}}\gamma$, also sensitive to $\boldsymbol{t\bar{t}}H$ and $\boldsymbol{t\bar{t}}g$.
- Top mass measurement from cross sections => resolving top mass "ambiguities" : MC mass vs mass in various renorm. scheme. Help to understand EWK vacuum stability.
- Mass extracted from various cross section measurements while scanning $\sqrt{s}$, and then compared to theoretical predictions.
- Cross section measurement precision : 1-2% to reach <200 MeV.
  - Physics backgrounds small (diboson),
  - High selection efficiency : related to detector performance (lepton/jets selection, flavour tagging) => impact on acceptance and modelling uncertainties.
  - Excellent control of selection efficiencies (from data).
- Several other topics of interests : EFT (!), top quark width, Yukawa, FCNC, direct mass measurement, CKM.
- Sensitivity of top quark measurements through ILC/CLIC extrapolations or fast-simulation. FCCee fullsim studies required.



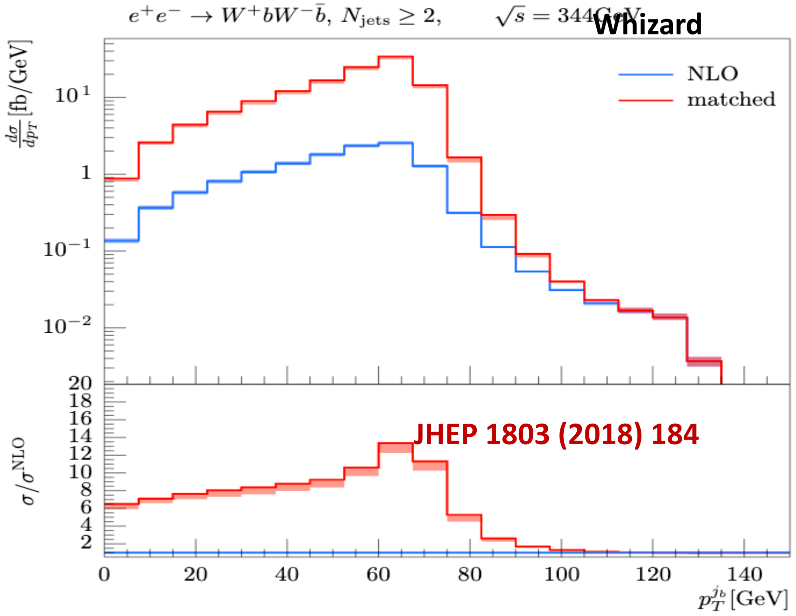


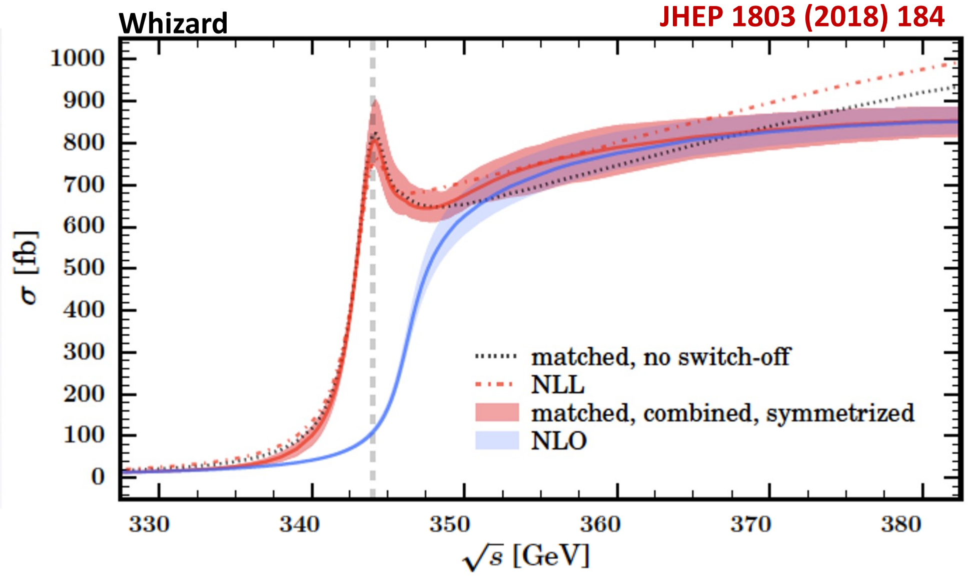
$t\bar{t}$ event generation and simulation

Generators : aMC@NLO and Whizard

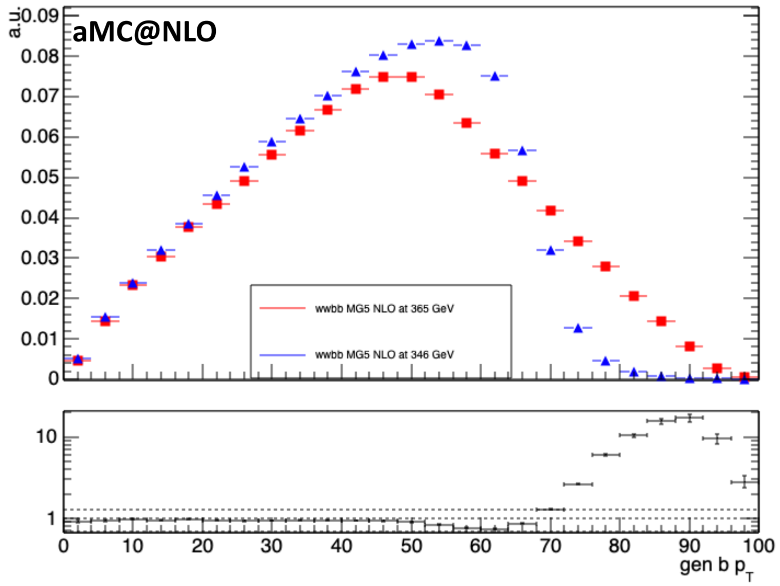
From 2020

- Having "state-of-the-art" generators is a key element for precisions
  - Maximum possible accuracy : NLO QCD+QED,
  - NLL+NLO matching : differential cross sections at threshold, effects of $\sqrt{s}$ on kinematics,
  - We need at least 2 generators, **under investigations** : Whizard and aMC@NLO.
- Status of generators (was in a not-yet public release for aMC@NLO, current status to be checked) :
  - NLO accuracy, **Whizard** : QCD , **MadGraph** :QCD (QED under developments for both generators),
  - Initial State (QED) Radiation, **both,**
  - **Proper description of the threshold : Whizard** only, but work needed.





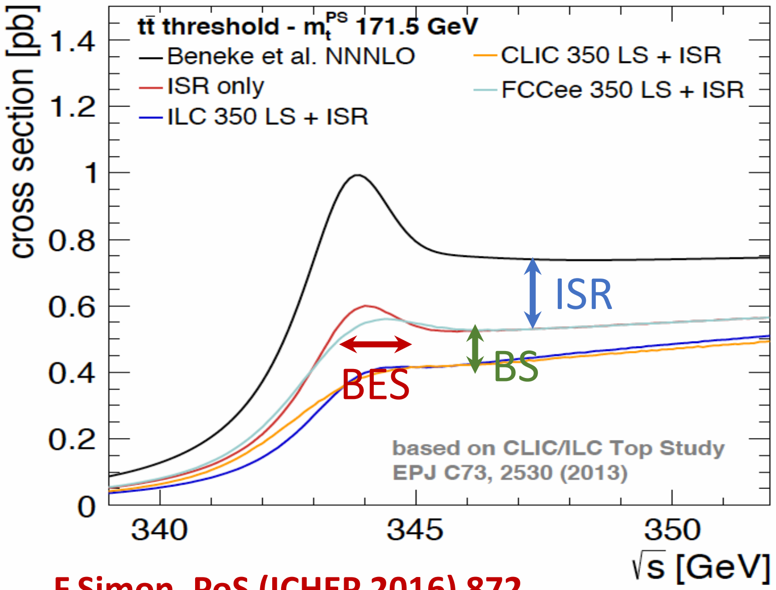
Need MC events which reproduce the kinematic at threshold



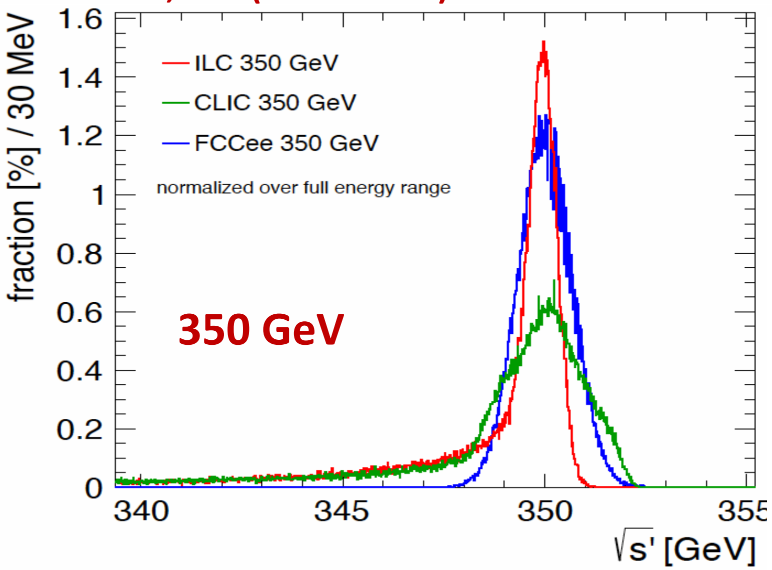
From 2020

Beam effects/ISR

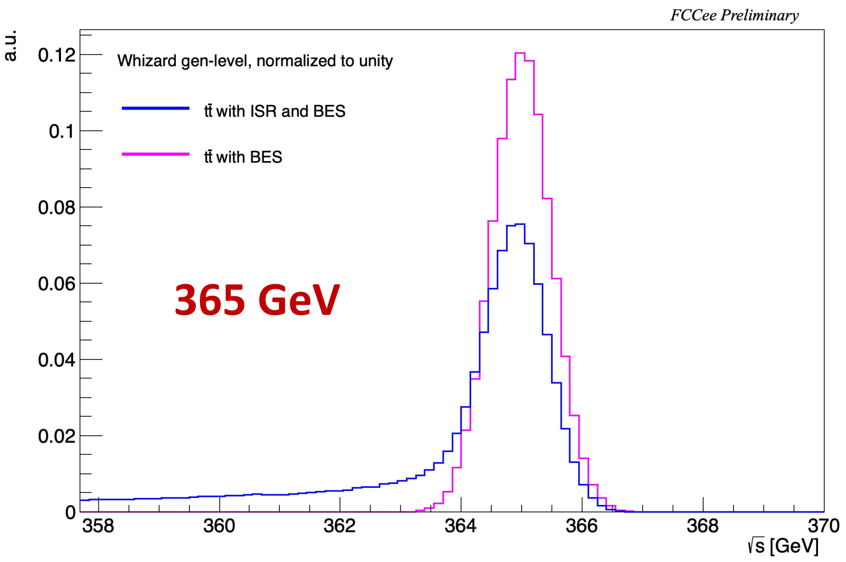
F.Simon, PoS (ICHEP 2016) 872

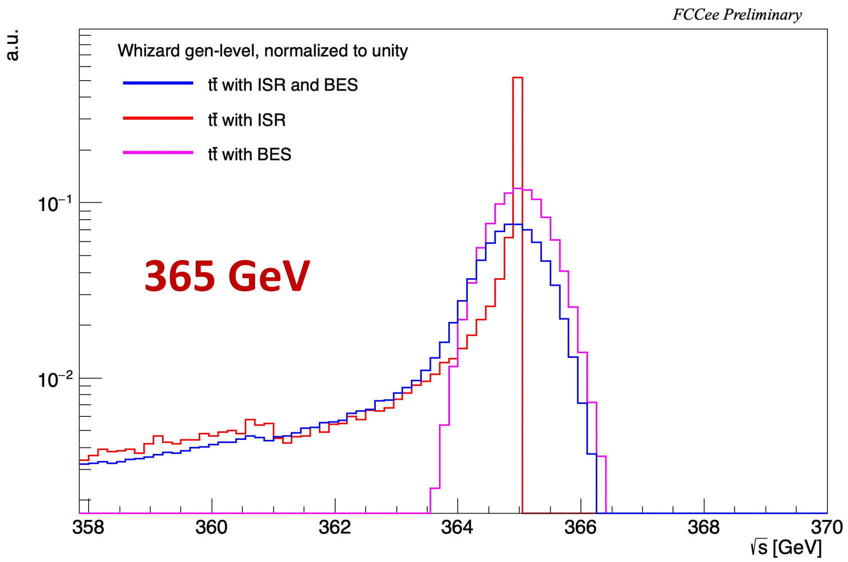


F.Simon, PoS (ICHEP 2016) 872



- ISR and beam effects on the threshold measurement :
  - ISR and Beam Backgrounds : reduces the energy in the e^+e^- centre of mass => tails toward lower energies.
  - Beam Energy Spread (BES) : enlarge the $\sqrt{s}$ distribution. BES ~0.19% per beam.
- At FCCee BES : $\sqrt{s}$ distribution symmetric and gaussian with very good approximation.
- Comparing Whizard and Madgraph (private version)
  - Beamstrahlung : **Whizard** : interface with GuineaPig/CIRCE. **MadGraph** : parametrization fitted to GuineaPig++.
  - Beam Energy Spread : **Whizard** : Gaussian smearing in case of FCCee, **Madgraph** : not available yet.



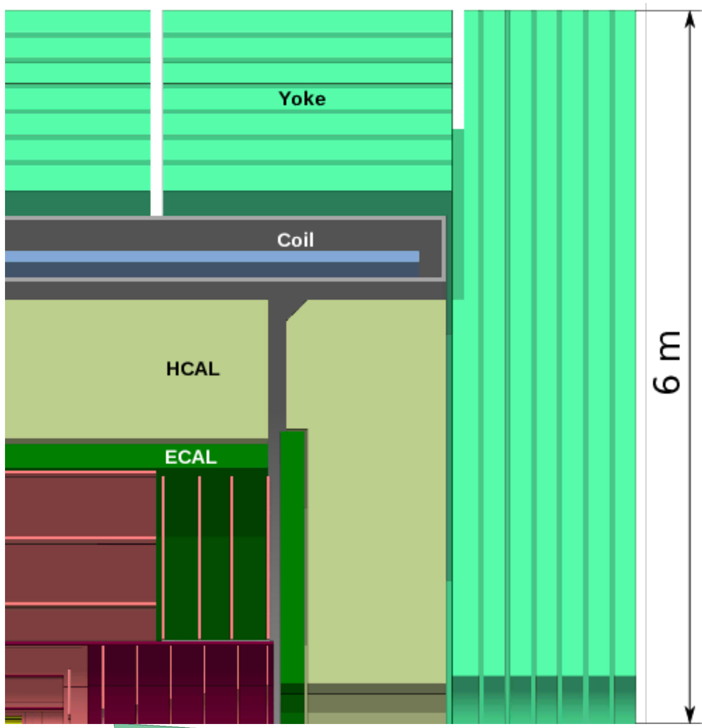


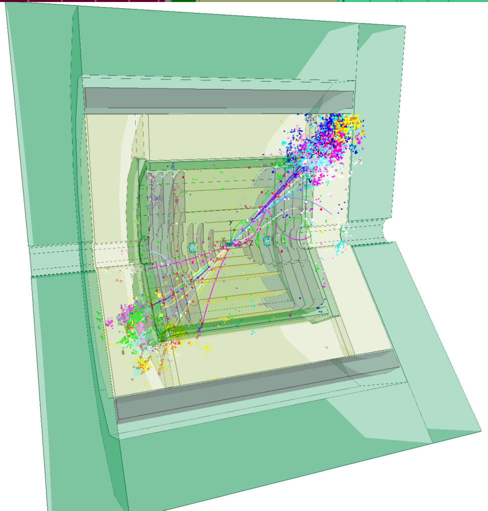
The CLD detector

- CLIC detector revisited for FCCee:
  - Full silicon tracker,
  - High granularity calorimeters for Particle Flow (ECAL silicone tungsten and HCAL scintillator-steel),
  - Tracking+calorimeter inside a solenoid (2T).
  - Muon chambers (RPC) inserted within the return yoke.

- CLD detector, as described in CLD paper (link) :
  - Fully implemented and accessible for full simulation and reconstruction,
  - Reconstruction based on Pandora particle flow algorithm,
  - Benefit a lot from the efforts made by the ILC/CLIC communities,
  - Some variants of the detector (beam pipe/VTX) exists and can be studied (**see talk G.Sadowski**).

- Thanks to edm4hep, the simulation files are compatible with FCCAnalysis.





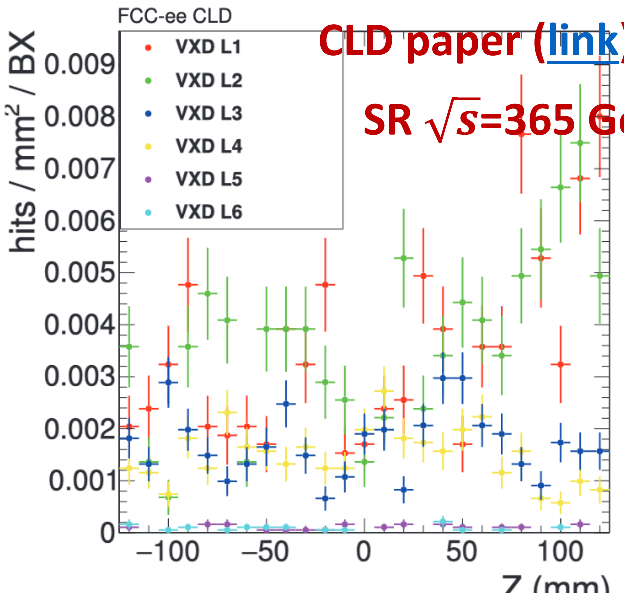
Beam backgrounds at $t\bar{t}$ threshold

- Beam Backgrounds are important, especially for VXD design:
  - $\gamma\gamma \rightarrow hadrons$,
  - Synchrotron Radiation (SR) from last bending magnet,
  - Incoherent Pair Creation (IPC, e^+e^- pair via interaction with beamstrahlung).

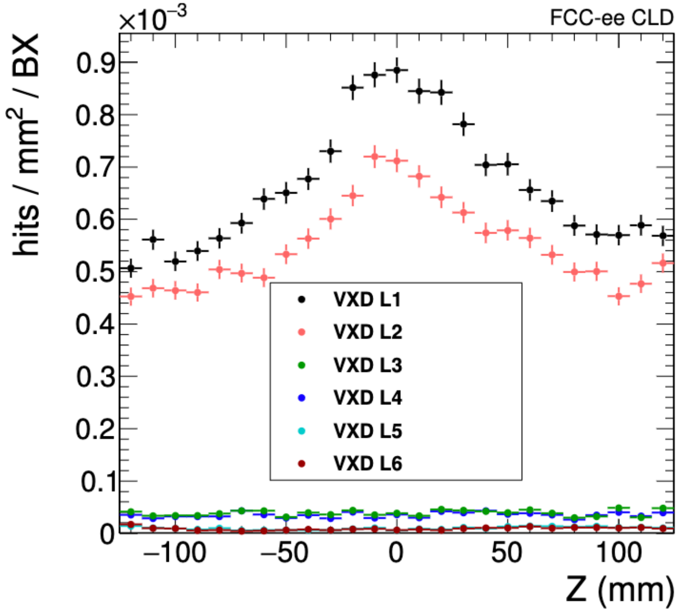
- Effects can be (re)estimated from full simulation, impact on the CLD vertex detector shown.

CLD paper (link)

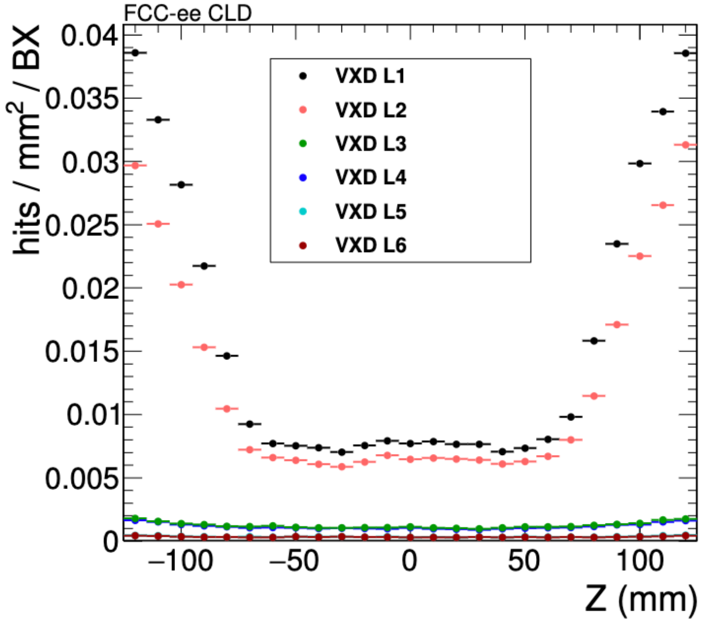
SR $\sqrt{s}$=365 GeV

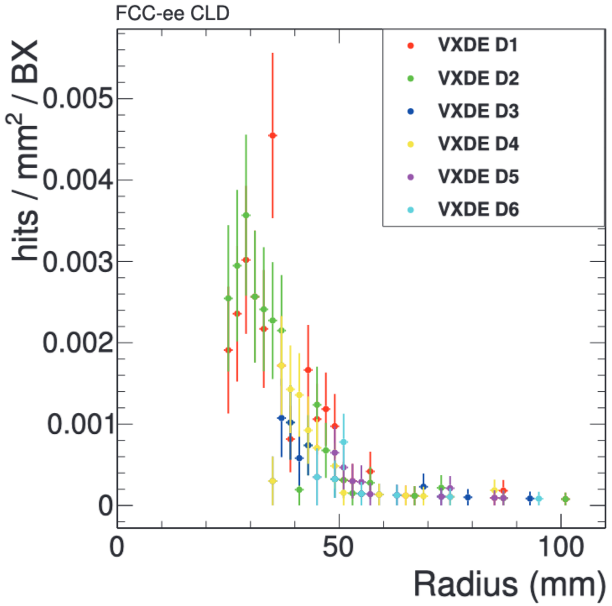


IPC $\sqrt{s}$=91.2 GeV



IPC $\sqrt{s}$=365 GeV



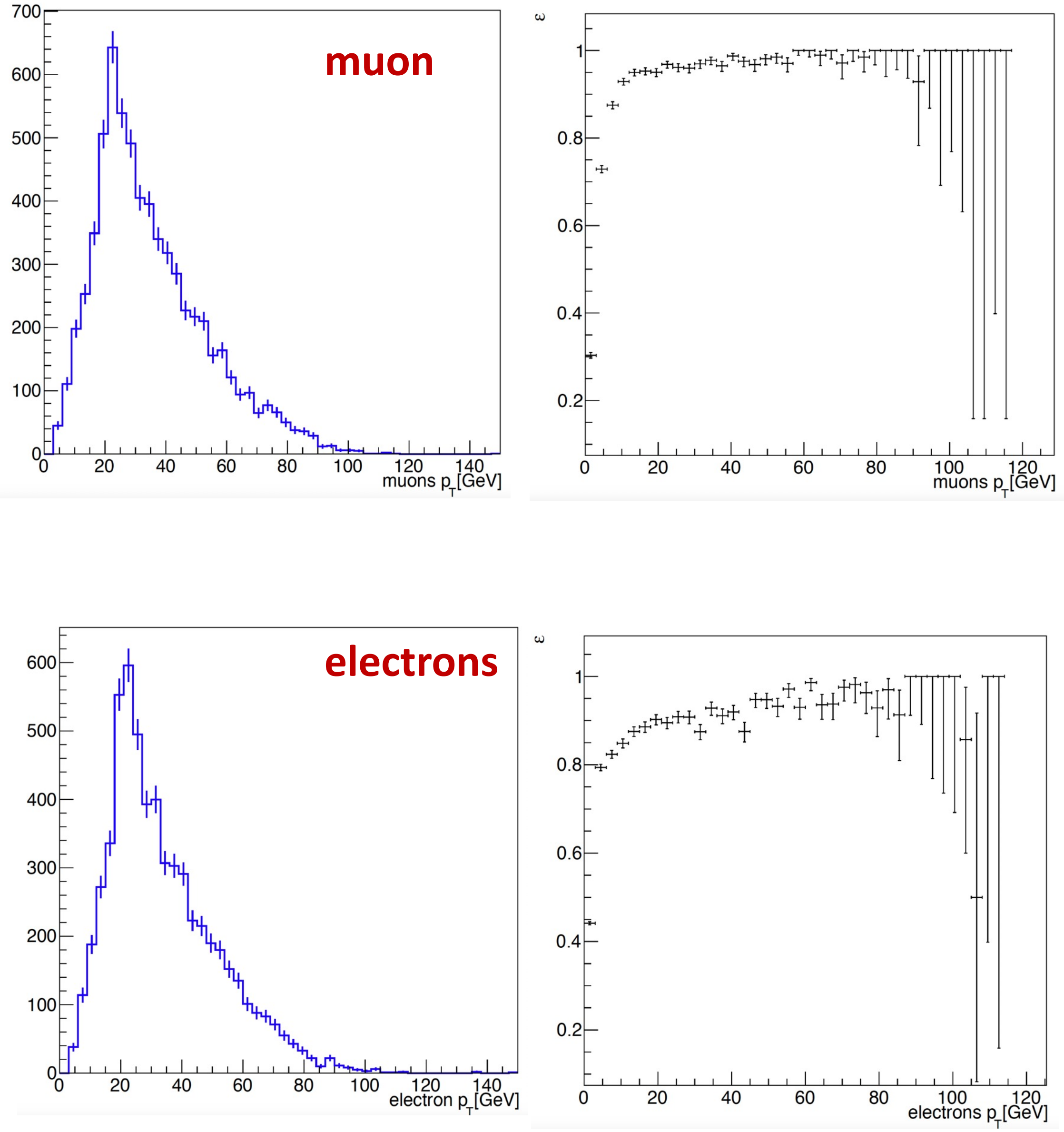


FCC-Analysis on ttbar fullsim

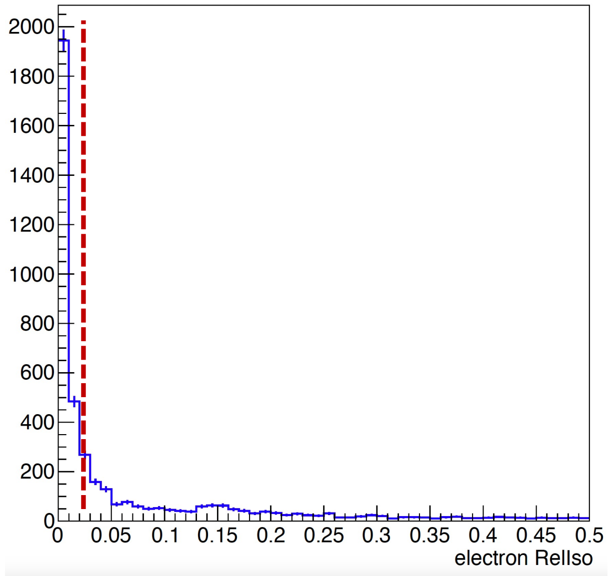
very first (preliminary) results

Event generation and selection

- First attempt (for me at least) to produce fullsim $t\bar{t}$ events with CLD :
  - Madgraph $t\bar{t}$ events, including decays, at $\sqrt{s} = 365$ GeV,
  - Perform ISR and hadronization with pythia8 "standalone" to have hepmc3 files.
  - ddsim fed with hepmc3, then k4run for the reconstruction.
  - ~25kevents generated with condor (limited by eos disk quota).
- Reconstruction with Pandora Particle Flow.
  - Lepton reconstruction efficiencies calculated with simple ΔR matching with MC leptons.
  - No cuts on the leptons.
- Events selection (lepton+jets) :
  - Exactly one isolated leptons (electrons or muons) with p_T>20 GeV,
  - Use of exclusive jet reconstruction (no selection) or inclusive jet reconstruction (p_T > 15 GeV, >3 jets).

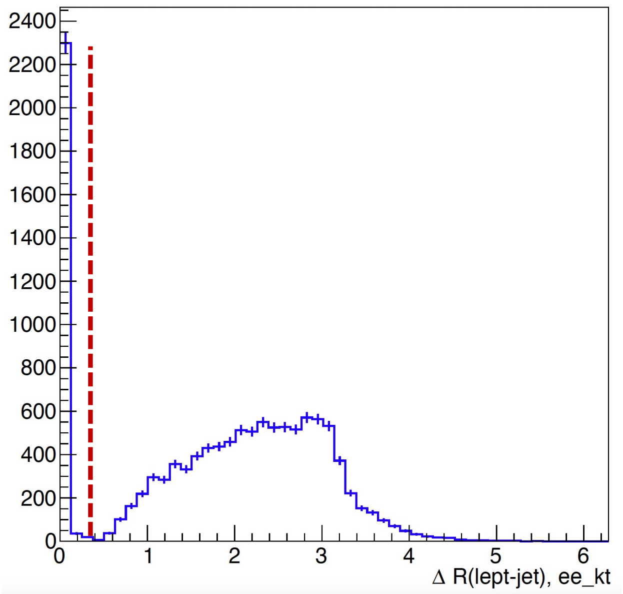


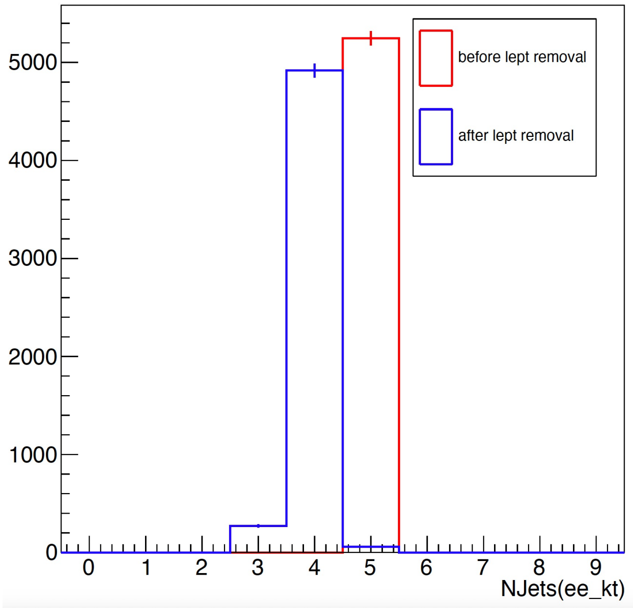
Lepton isolation and jet/lepton removal

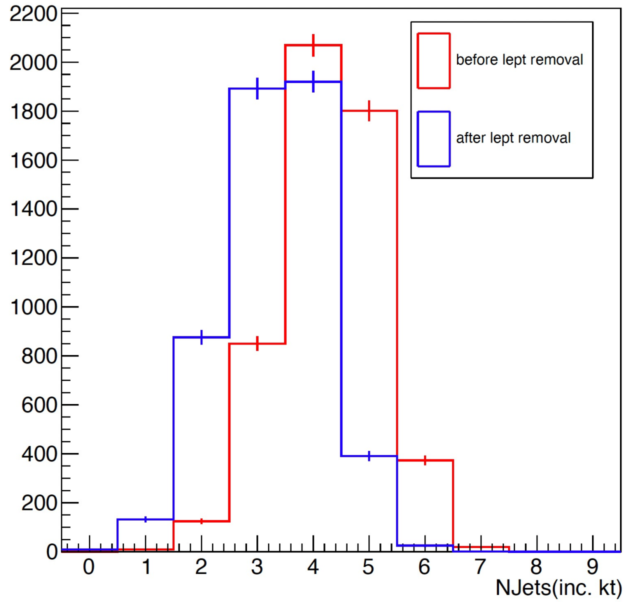


- Lepton isolation requirement :
  - Sum of charged and neutral momenta in a cone around the lepton $0.01<\Delta R<0.5$,
  - Relative isolation calculated : $\mathrm{RelIso} = Sum/p_l$,
  - Lepton isolated if $\mathrm{RelIso}<0.025$.

- Leptons also reconstructed as jets.
  - Remove isolated leptons from jet collection,
  - Exclusive jet reconstruction to be configured accordingly (+1 extra jet to account for the lepton).

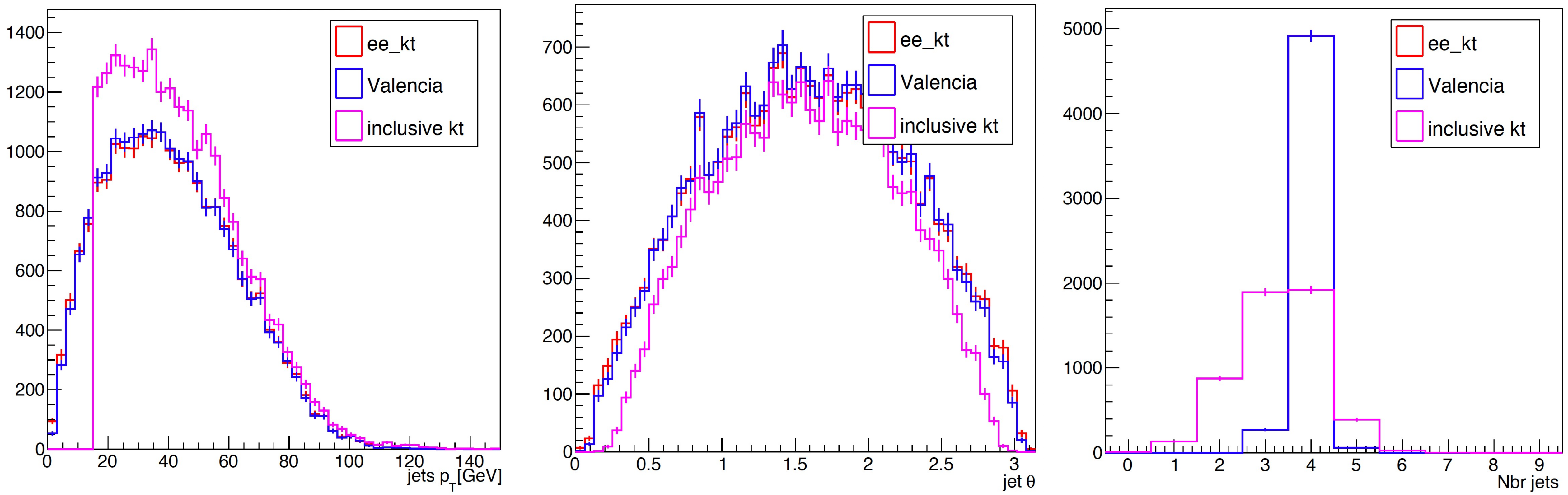






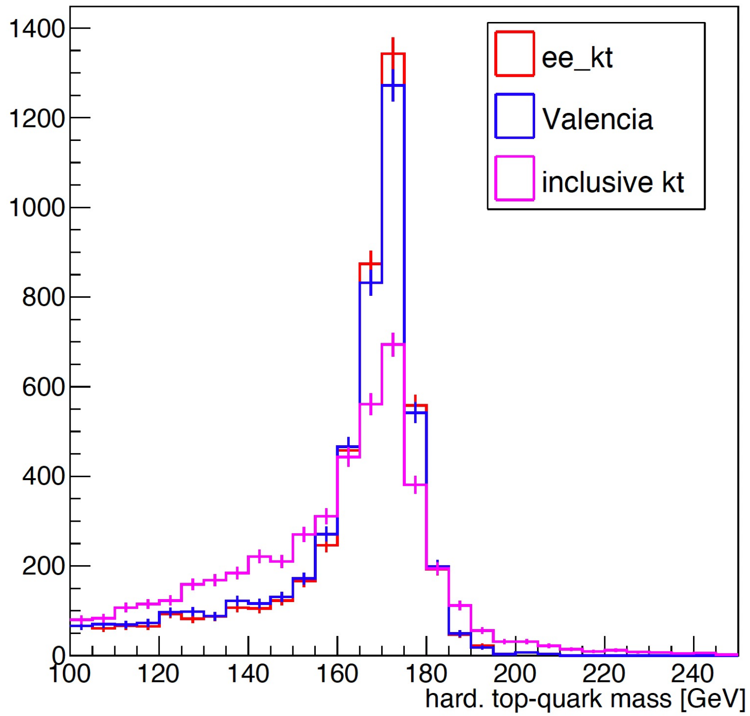
Jet reconstruction comparisons

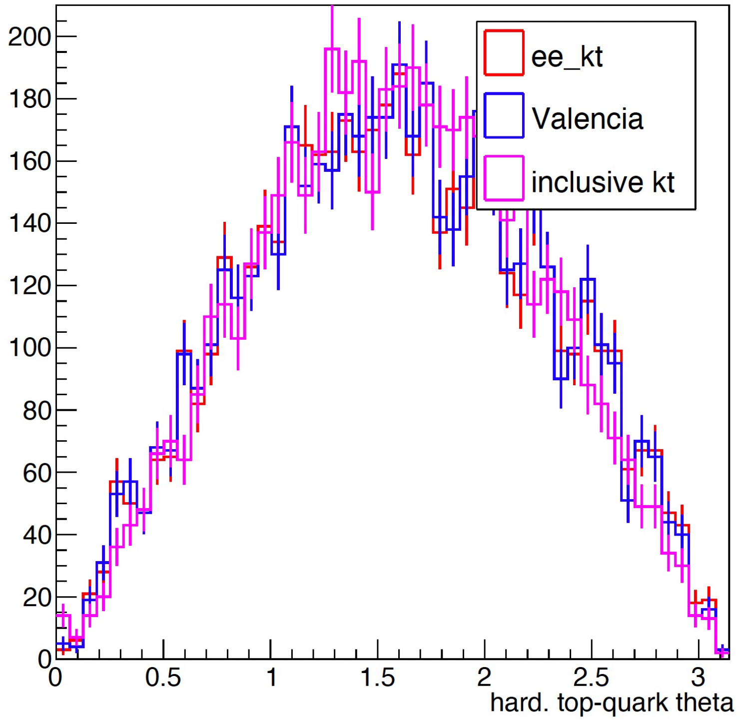
- Example of a simple study :
  - Compare jet reconstruction for various jet algorithms,
  - Here tested with exclusive jets algo (Valencia, ee_kt) and inclusive (kt 0.4).
- Very first look at these distributions. Further studies required.



Simplified top reconstruction

- Top quark mass sensitive to jet resolution.
- Tested with simple method :
  - Compute the invariant mass of all triplet of jets,
  - Selected the triplet with mass the closest to 172.5 as the hadronic top-quark of the lepton+jets channel.





Conclusion

- Top quark physics is major topic at FCCee colliders. Given its rich final states, it also constitutes a perfect playground for testing full simulation.

- With the CLD detector, it is possible to test the complete simulation chain : from event generation to full simulation and full reconstruction. A very simple example of an analysis is shown in this presentation.

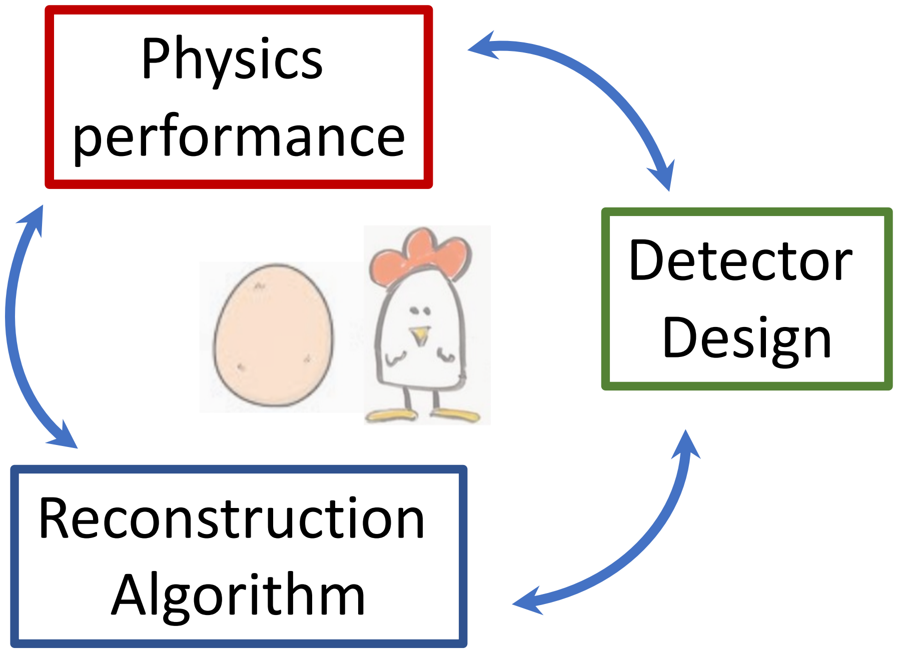
- There are still several topics to work on (toward a centralized generation ”recipe” ?):
  - Test fullsim also with whizard ! (other generator ?),
  - Include “beam” aspects,
  - Performed in-depth validation of the reconstructed events,
  - Algorithmic aspects (like flavour tagging re-training).

- There are only very few CLD “developers/users” !

Backups

General (naïve) comments on detector design “optimisation”

- Needs (resolutions, efficiencies etc…) for top quark physics are probably very similar to the Higgs physics, at first order.
- We need to verify this assumption at $t\bar{t}$ threshold **(different beam conditions and backgrounds)!**



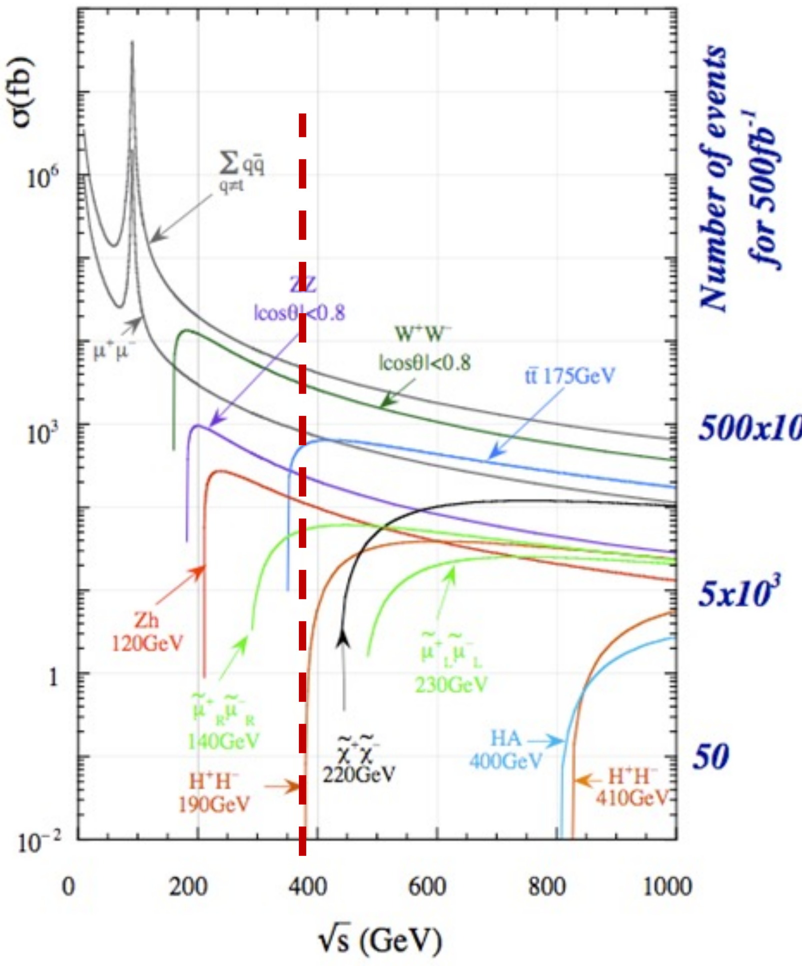
- Tools needed for Physics performance studies :
  - FastSim => interesting to test sensitivity on detector performance, but rapidly limited,
  - FullSim => ultimately needed, but takes time, need (flexible) reconstruction,
  - Intermediate approach with some modelling ? Partial fullsim (not entire detector) to feed fastsim?
- Developments need to proceed in parallel.
- Enough work on all topics to keep us busy for years.

- Some of this work already done for CLD/IDEA : do we want to join effort there, or create our own design? A lot to learn from ILC/CLIC here as well !

High involvement required !

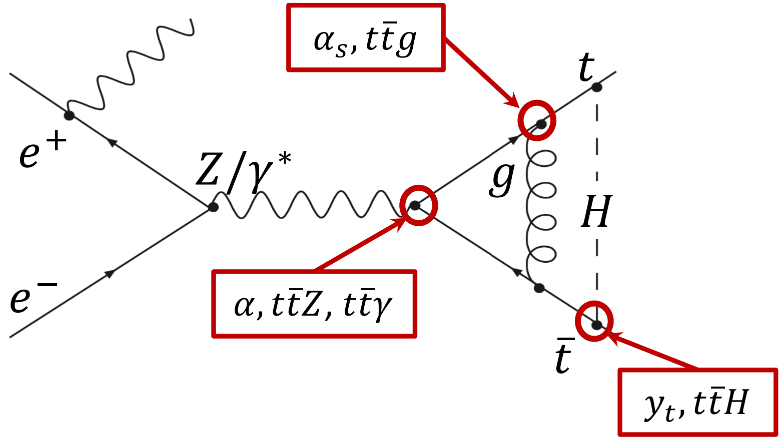
Do we need a trigger at $t\bar{t}$ threshold ?

- Trigger (at least software) might be foreseen for the Z run.
- Effects of trigger selection on analysis (my LHC bias) :
  - Could cause lower signal efficiencies ?
  - Systematics on the trigger efficiency ?
- At FCCee : mainly to reject beam-backgrounds, we want to keep all physic backgrounds (physics, alignment, calibrations and efficiencies measurements etc…).
- Rate of bunch crossing at $t\bar{t}$ (back of the envelop) : ~3000 ns of bunch spacing => ~300kHz, that is ~3 times the actual CMS/ATLAS L1 trigger rate, but half of the HL rates.
- Can/should we avoid L1 and/or HLT triggers ?
- (Naïve) questions to answer :
  - What is the rate of beam backgrounds ?
  - What is a typical size of an event ?
  - What is the needed readout speed and disk throughput ?
- At minima : low trigger requirements to detect a collision (a la LEP). Trigger systematics should be small !



Precision of $t\bar{t}$ cross section measurements

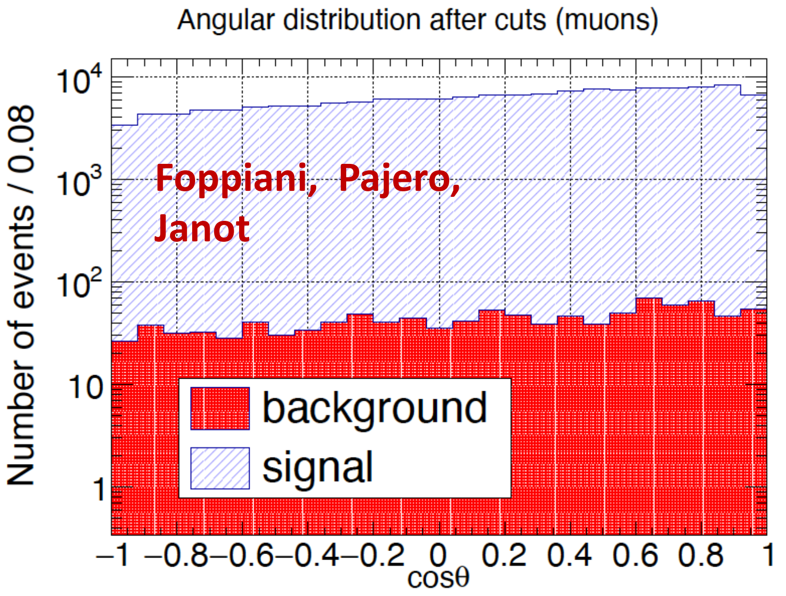
- Inclusive and differential => probe of $t\bar{t}Z$ and $t\bar{t}\gamma$ couplings (EFT related) at LO. Sensitive on $t\bar{t}g$ and $t\bar{t}H$ (EFT) couplings at NLO !



- Dominant backgrounds (lepton+jets):
  - WW(dominant)/ZZ => b-tagging !
  - WWZ, ZH => more difficult to reject, but much lower cross section (/20).

- Events selection :
  - one (relatively loose) isolated leptons with p_T>10 GeV, 80-90% efficiency,
  - no jet reconstruction in the LHC-sense : event reconstruction with expected number of jets as inputs (VTL algo),
  - cut on the "compatibility" of the jet multiplicity hypothesis,
  - b-tagging requirements,
  - jets and lepton association to top, with a kinfit (W and top mass).

- Overall efficiency ~60-70% can be achieved (JHEP 11 (2019) 003), very high purity (>99%).



- Target **systematics ~few %** (even below ?)
  - physics background should not be a problem,
  - highest possible selection efficiency : flavour tagging (!) but lepton sel/jet reco not negligible => impact also acceptance and modelling uncertainties !
  - Excellent control of selection efficiencies (from data).

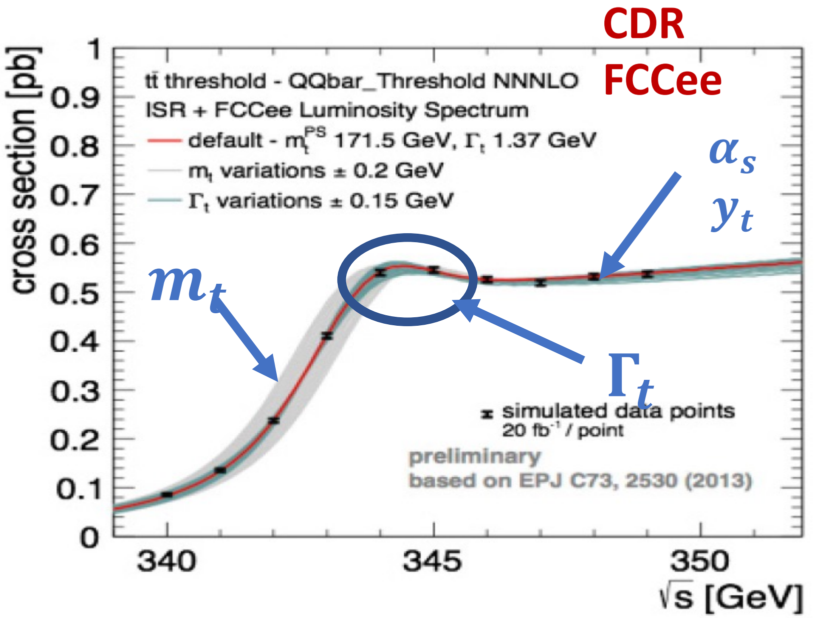
Top mass measurement

- Top mass measurement from cross sections => resolving top mass "ambiguities" : MC mass vs mass in various renorm. scheme. Also important to keep the universe safe against vacuum instability !

- Typical mass difference in the various renorm. schemes ~200 MeV.

- Mass extracted from various cross section measurements while scanning $\sqrt{s}$, and then compared to theoretical predictions.

- Cross section measurement precision : 1-2% to reach <200 MeV.



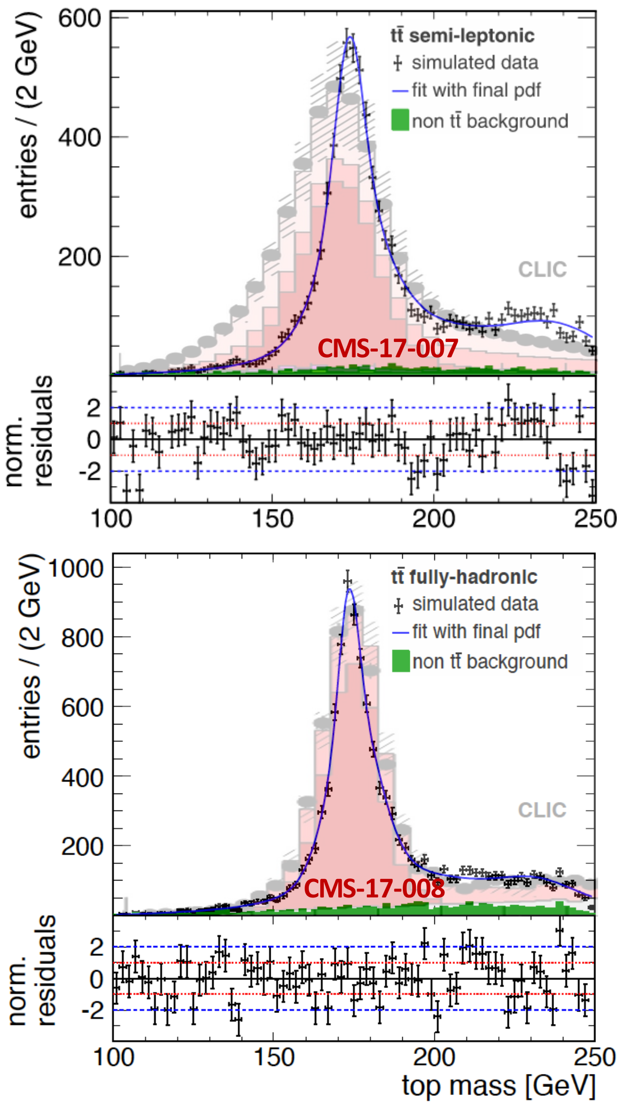
- Expected precisions (CLIC analysis revisited for FCCee):
  - Stat uncertainty at ~15 MeV,
  - Beam energy, reconstruction efficiency and background contamination ~50 MeV ,
  - And luminosity ... ~10 MeV,
  - **Total uncertainty below 100 MeV**, previous measurements of α_s => **reduction to < 50 MeV could be achievable**!

- Experimental uncertainties (close to be) dominated by statistics is possible at the FCCee !

- Still significant impact of theory uncertainties : requires a significant effort from theory community.

Direct measurement of top mass from decay products (above threshold)

CLIC, EPJC 73 (2013) 2530



- Direct mass measurement from top quark decay products (in a nutshell):
  - reconstruct and identify decay products,
  - reconstruct top quarks candidates using a kin fit (determine jets-lepton associations),
  - fit the reconstructed top mass with templates issued from MC generation. Simultaneous fit with JES reduces systematics,
  - requires "calibration" : input $m_t^{MC} \neq m_t^{reco}$.

- Comparisons with CMS top reconstruction at 13 TeV, 35.9 fb^{-1}.

- Estimations of the uncertainties (CLIC@380 GeV) :
  - stat: 30-40 MeV for $1ab^{-1}$,
  - moderate impact of JES : 2% variation of light and b jets = 200 and 350 MeV,
  - JES related uncertainties can be greatly reduced by including the perfect knowledge of the initial stat into the events reconstruction,
  - =>statistically dominated measurement?

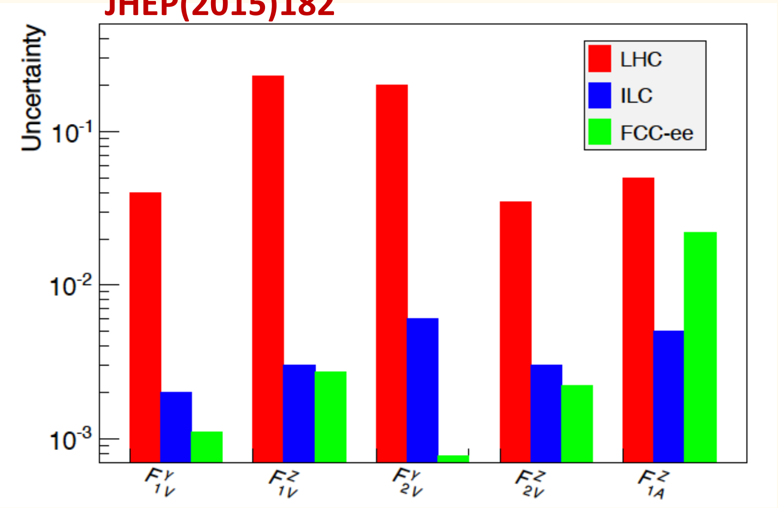
- Direct top mass measurement can be competitive with the threshold scan measurement.

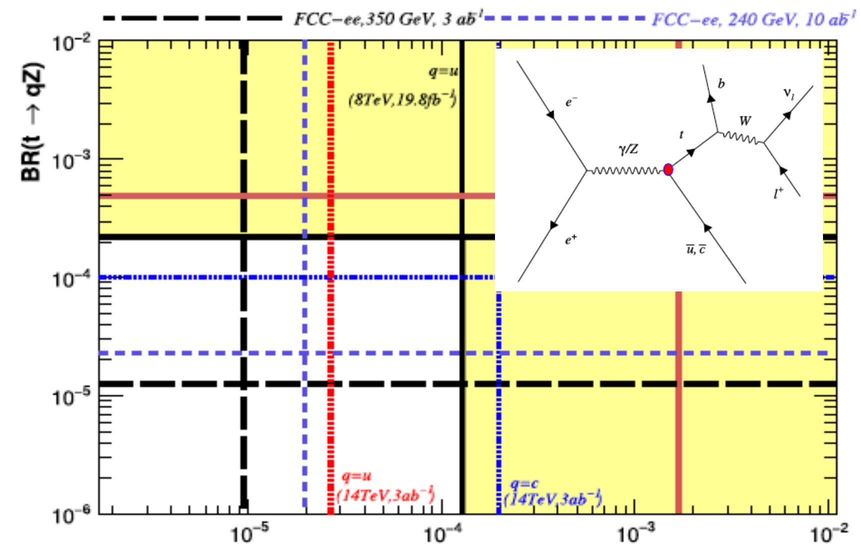
- Other "non" standard measurements can help (dilepton, J/ψ, endpoints, extra jet/γ) => combinations?

Top quark couplings to bosons

- Sensitivity on (anomalous) $t\bar{t}$ EWK couplings.
- Based on lepton energy and polar angle :
  - very low expected experimental uncertainties,
  - dominated by stat. uncertainties (and theory).
- Lower integrated lumi and larger boost at higher energies => better precision at 365 GeV than higher energies.

JHEP(2015)182



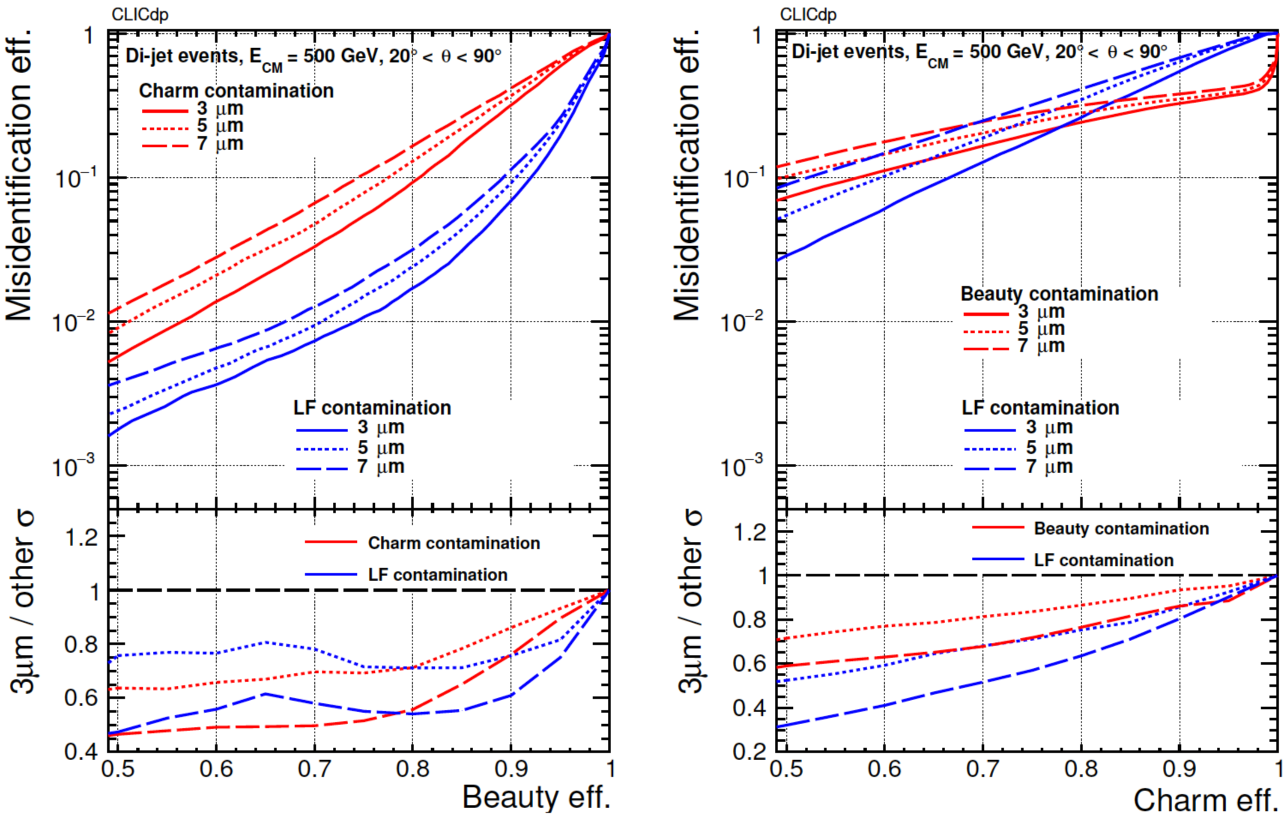


PLB 775(2017) 25-31

- Top-quark FCNC couplings to γ, Z, H usually probed in top quarks decays in $t\bar{t}$ (probably to be updated).
- Interesting channels at lepton colliders : single top production possible for $t\gamma$ and tZ-FCNC.
- Very promising channels : higher cross section, limited by statistics and background contamination (Wjj), $t\bar{t}$ channels still useful to disentangle $t\gamma$ from tZ.
- Large impact of b and c-tagging.

Detector impact on flavour tagging

CLICdp-Note-2018-005



- Flavour (b/c)-tagging is a key element for top quark physics.
  - $\varepsilon_{t\bar{t}} \propto \varepsilon_b^2$,
  - Top-FCNC , $t \to cH(b\bar{b})$, $\varepsilon_{tHc} \propto \varepsilon_b^2 \varepsilon_c$.
- B-tagging and c-tagging performances for various single point resolutions.
- From 7μ to 3μ:
  - ε_b: ~8%(abs.) improvement at $\varepsilon_l \approx 1\%$,
  - ε_c: ~18%(abs) improvement at $\varepsilon_l \approx 10\%$.
- ➔ increase of ~10% abs (**20% rel**) of $\varepsilon_{t\bar{t}}$ (for Medium P…) and ~15% abs (**75% rel**) of ε_{tHc}.
- Flavour tagging systematics ⬄ data driven estimations of efficiencies.

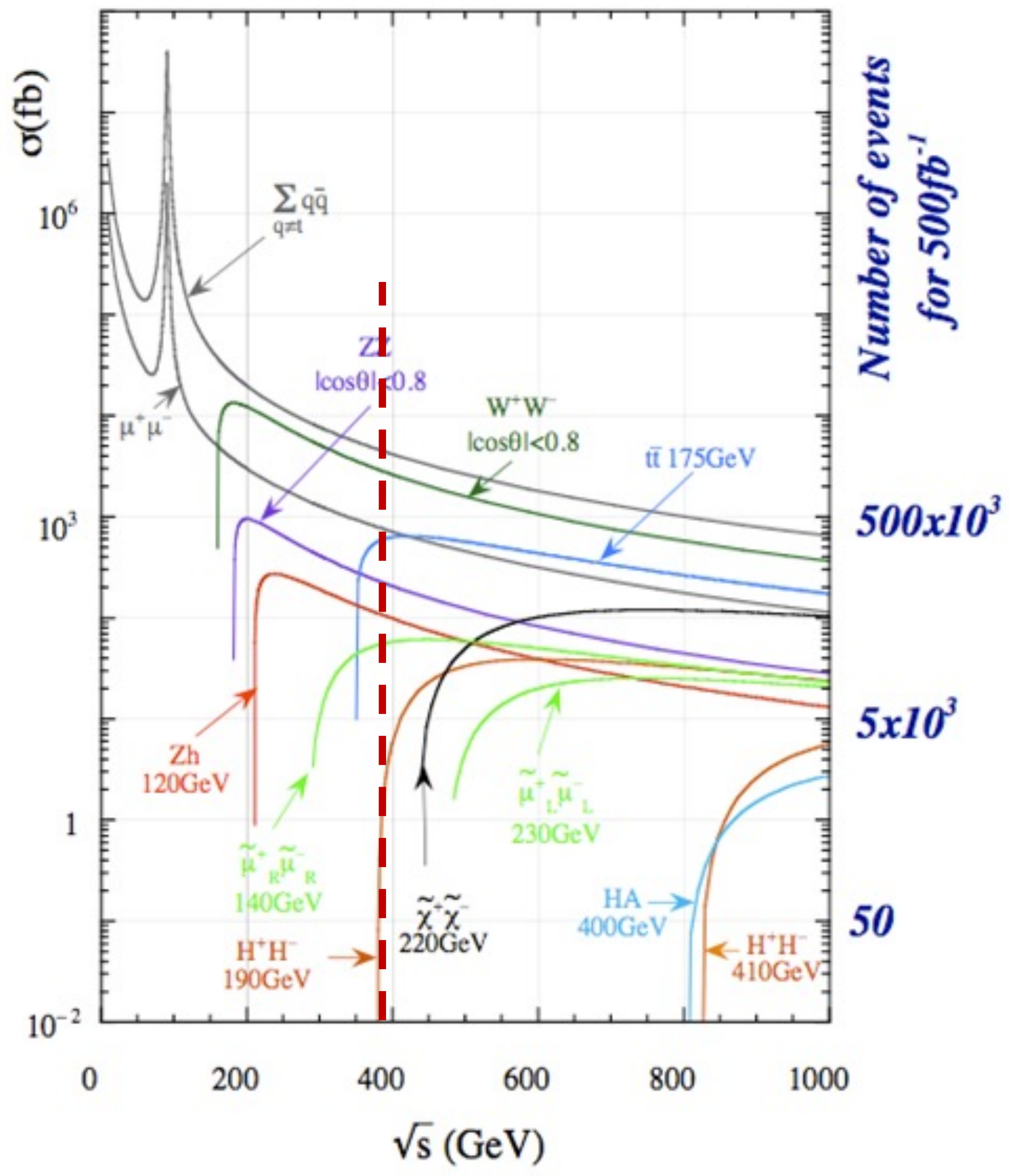
Backup

Discussions on backgrounds

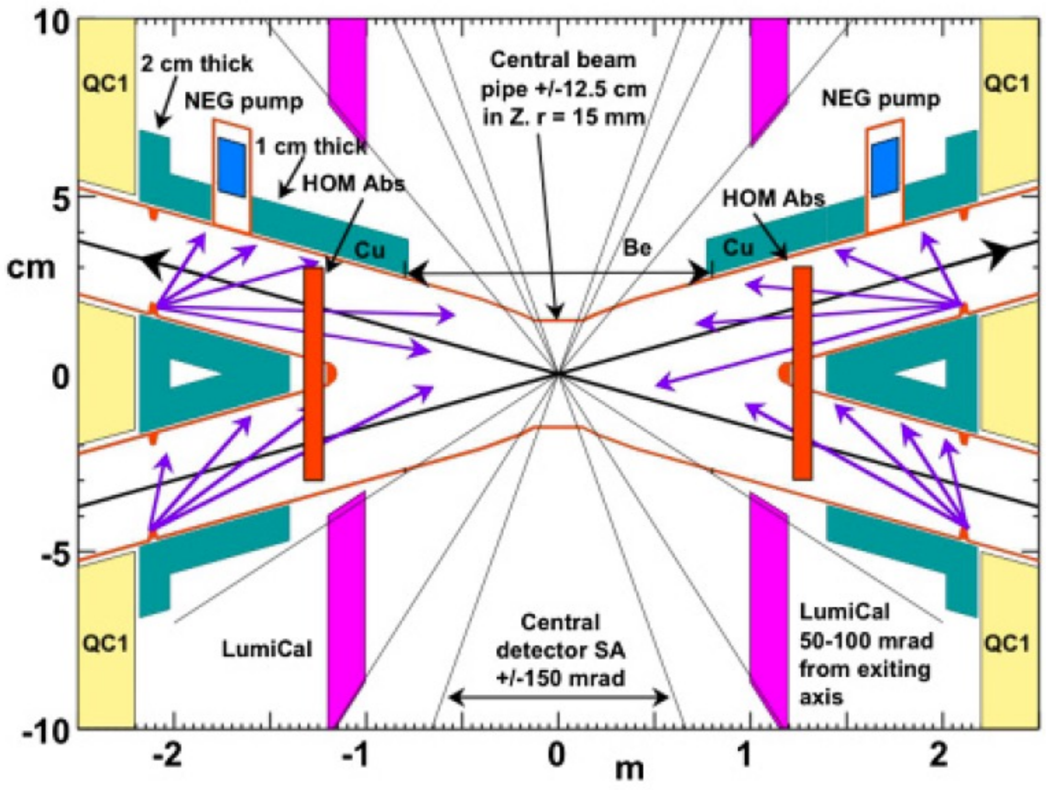
- List of the main background and cross sections.

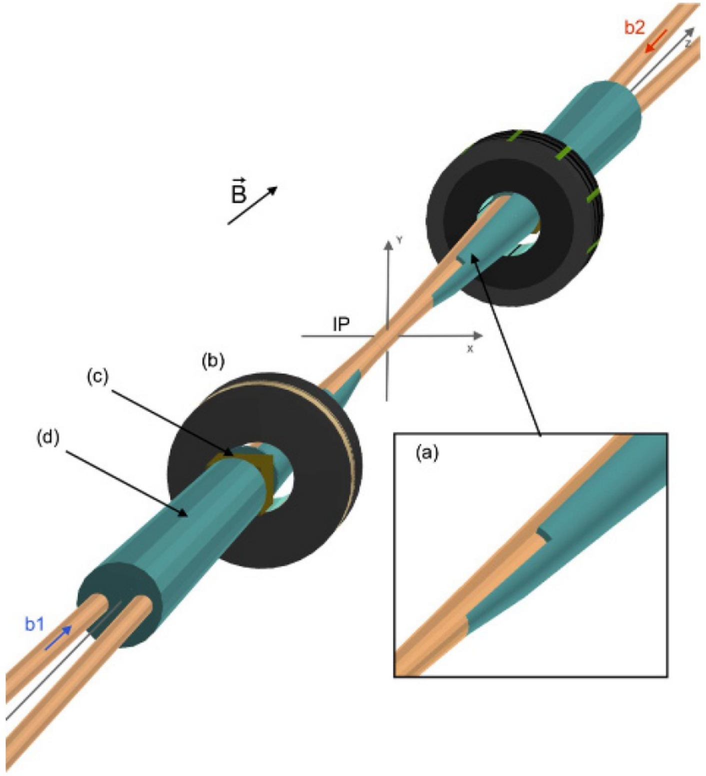
Beam backgrounds (large angle) M.Dam link

Energy	Process	Cross Section	Large angle $e^+e^- \to \gamma\gamma$	Large angle $e^+e^- \to e^+e^-$
90 GeV	$e^+e^- \to Z$	40 nb	0.039 nb	2.9 nb
160 GeV	$e^+e^- \to W^+W^-$	4 pb	15 pb	301 pb
240 GeV	$e^+e^- \to ZH$	0.2 pb	5.6 pb	134 pb
350 GeV	$e^+e^- \to tt$	0.5 pb	2.6 pb	60 pb



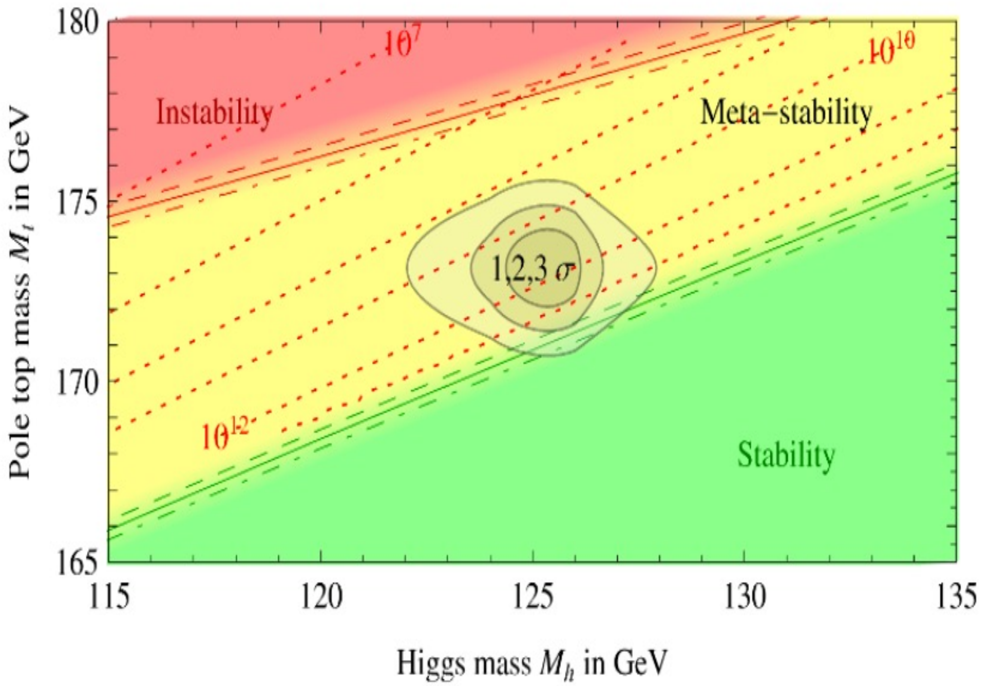
Interacting points





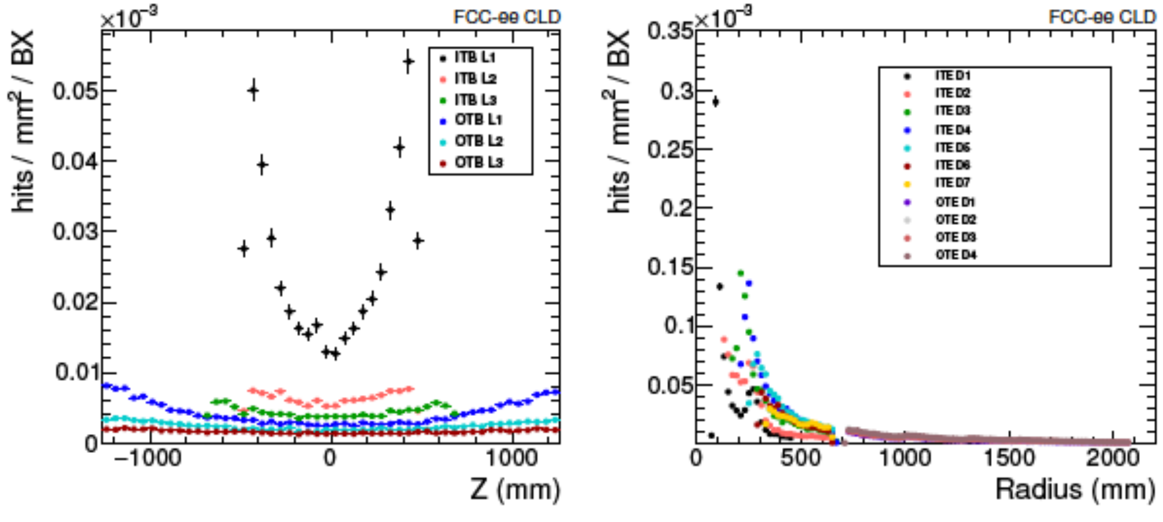
Top mass : target

JHEP08 98 (2012)



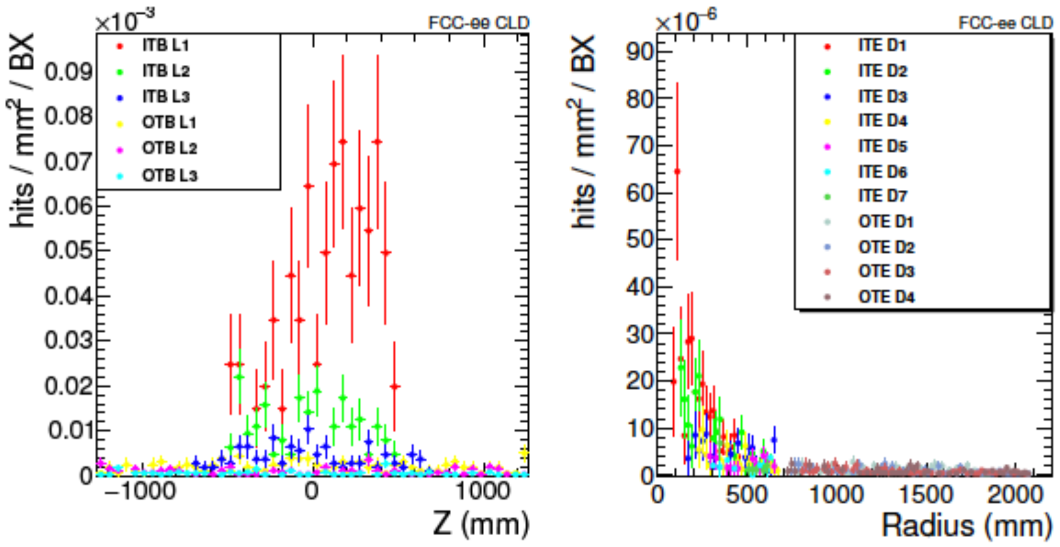
- Objectives of top mass measurement :
  - Test of the SM, yukawa couplings and top mass,
  - Confront pole mass to the “MC” mass (differences of a coupe f hundreds MeV),
  - Study of the stability of the vacuum, differentiations between stable and meta-stable universe.

Beam background



(a) (b)

Figure 17: Hit densities in the CLD tracking detector barrel layers (a) and discs (b) for particles originating from incoherent pairs, for operation at 365 GeV. Vertical error bars show the statistical uncertainty, horizontal bars indicate the bin size. Safety factors for the simulation uncertainties are not included.



(a) (b)

Figure 18: As Figure 17 but for hits related to synchrotron radiation photons.

Breit-Wheeler

$$\gamma + \gamma \rightarrow e^- + e^+$$

Bethe-Heitler

$$\gamma + e^\pm \rightarrow e^\pm + e^- + e^+$$

Landau-Lifshitz

$$e + e \rightarrow e + e + e^- + e^+$$

Bremsstrahlung

$$e + e \rightarrow e + e + \gamma$$

$\sqrt{s}$ (GeV)	91.2	365
Total particles	800	6200
Total E (GeV)	500	9250
Particles with $p_T \geq 5\,\mathrm{MeV}$ and $\theta \geq 8^\circ$	6	290

Rates of electron pair backgruonds

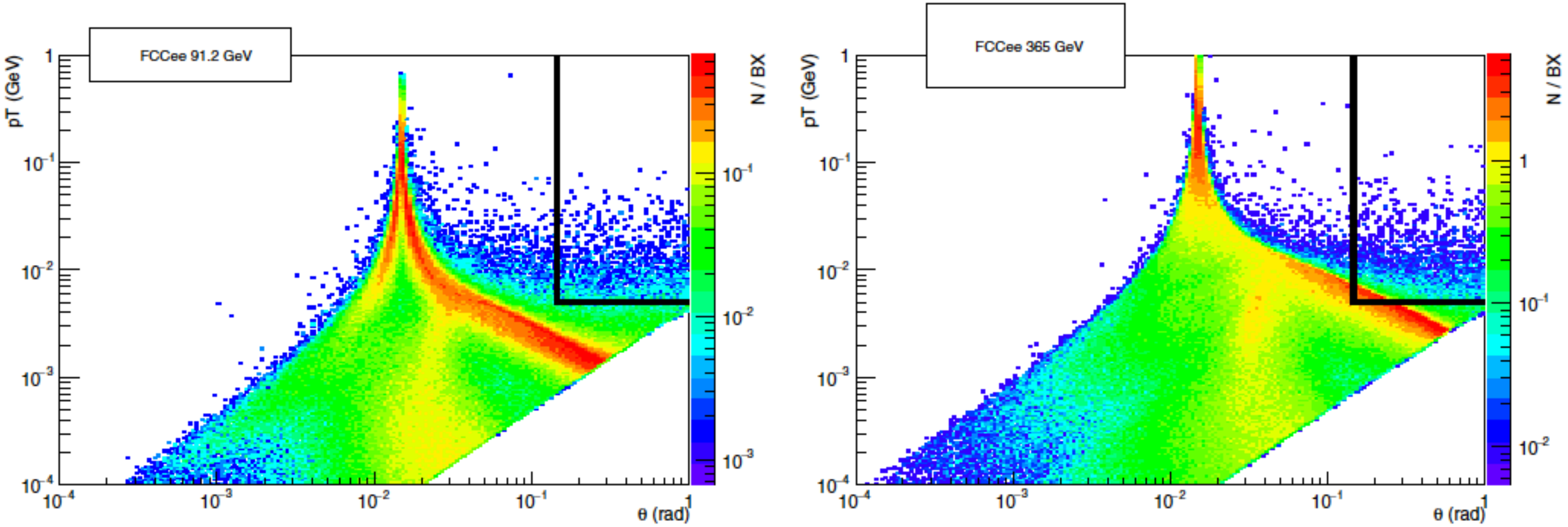


Fig. 7.2. Rates of $e^{\pm}$ from IPC in the (p_T, θ) plane, in the detector frame, for $\sqrt{s} = 91.2$ GeV (left) and 365 GeV (right). The black line in the upper-right corner delineates the CLD vertex detector acceptance within a field of 2 T.

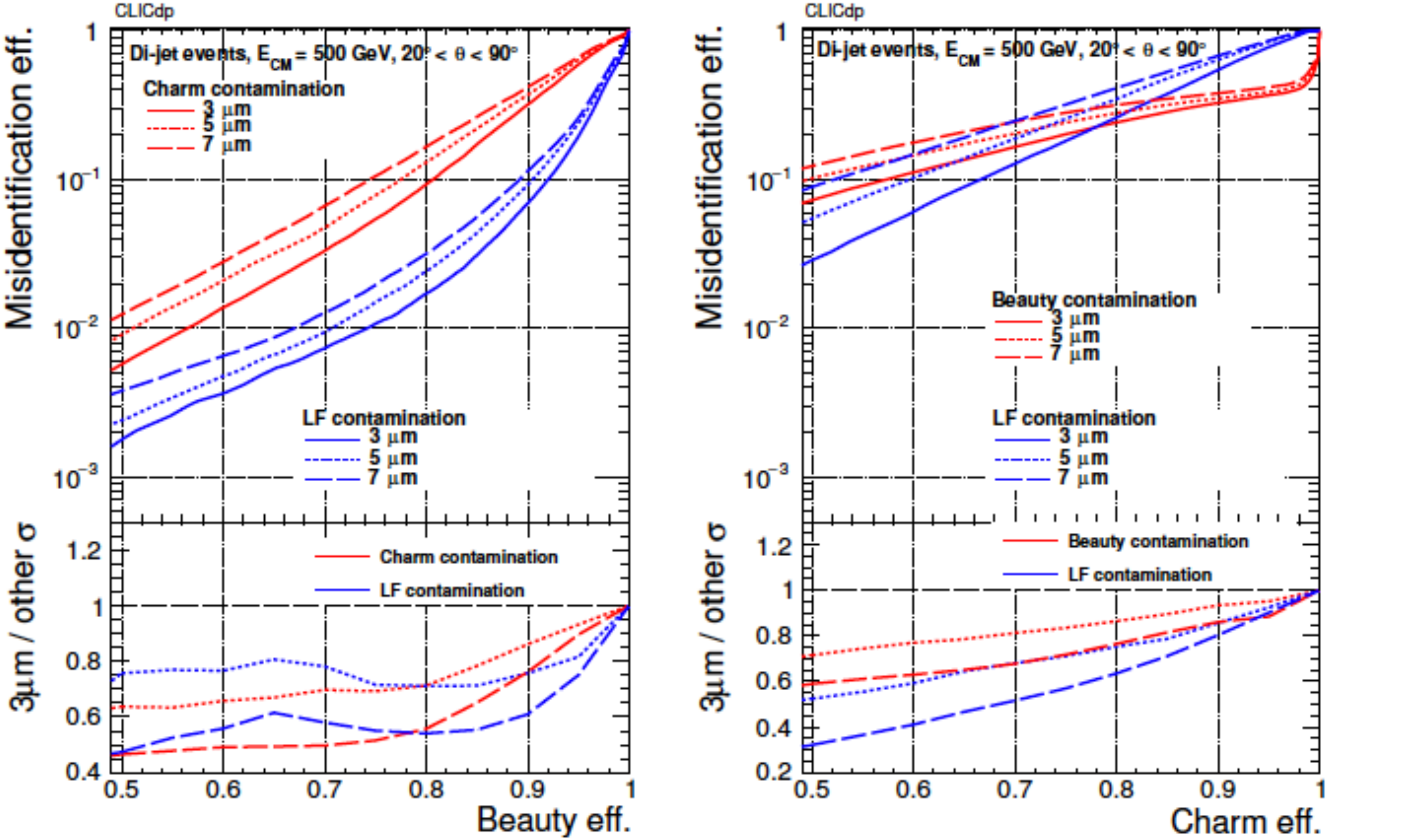
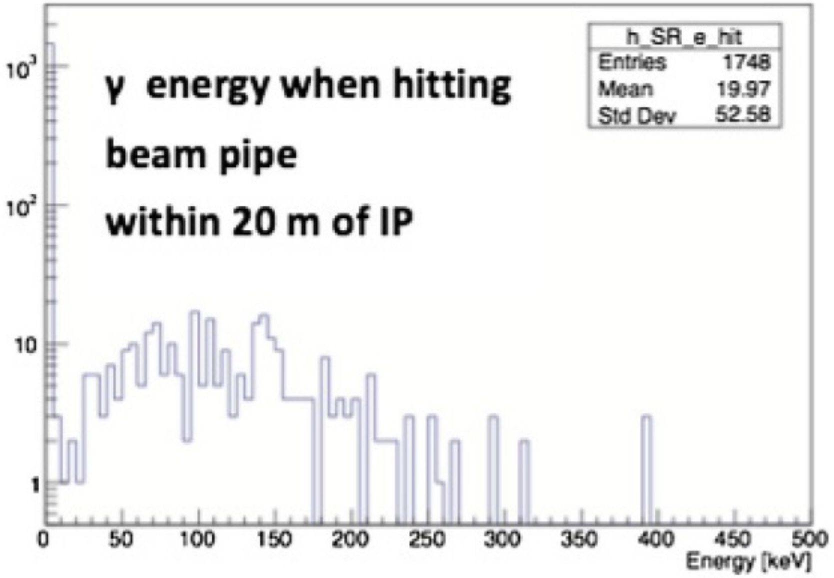


Figure 62: Global performance of beauty tagging (left) and charm tagging (right) for jets in di-jet events at $\sqrt{s} = 500\,\text{GeV}$ with a mixture of polar angles between 20° and 90°. A comparison of performance obtained with different single point resolutions in the vertex detector is presented. On the y-axis, the misidentification probability and the ratio of misidentification probabilities with respect to the nominal (3 μm) single point resolution are given.



Notes. The second column refers to the number of photons incident at 500 μm from mask tip and with an energy >1 keV, the third and fourth columns give the incident number of photons in the central beam pipe per beam crossing and per second, respectively. Solenoid fields and collimators were not taken into account. Note that this table was calculated for an older version of the beam optics with a maximum beam energy of 175 GeV. For the more recent optics of Section 2.4 even at a beam energy of 182.5 GeV the critical photon energy is below 100 keV.

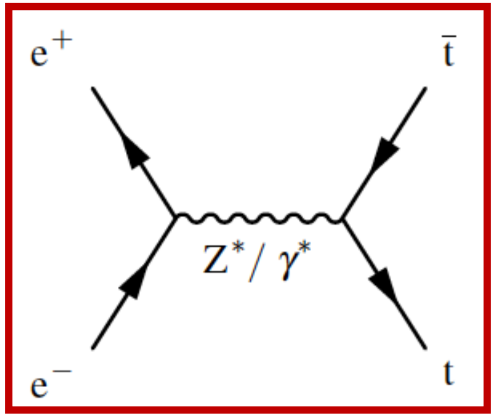
Table 2.7. Summary of the SR coming from the last soft bend upstream of the IP.

E_{beam} GeV	$E_{critical}$ keV	Incident γ/crossing (500 μm from tip)	Incoming on central pipe/crossing	γ rate on central pipe (Hz)
182.5	113.4.	3.32×10^9	1195	1.18×10^8
175	100	3.06×10^9	1040	1.25×10^8
125	36.4	1.05×10^9	10.3	1.01×10^7
80	9.56	6.11×10^8	0.18	7.02×10^5
45.6	1.77	9.62×10^7	1.92×10^{-4}	9.58×10^3

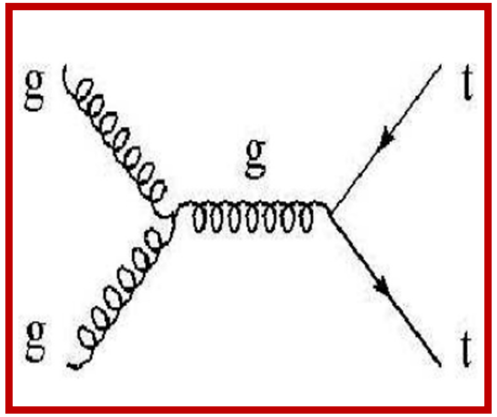
$t\bar{t}$ events properties : reminder

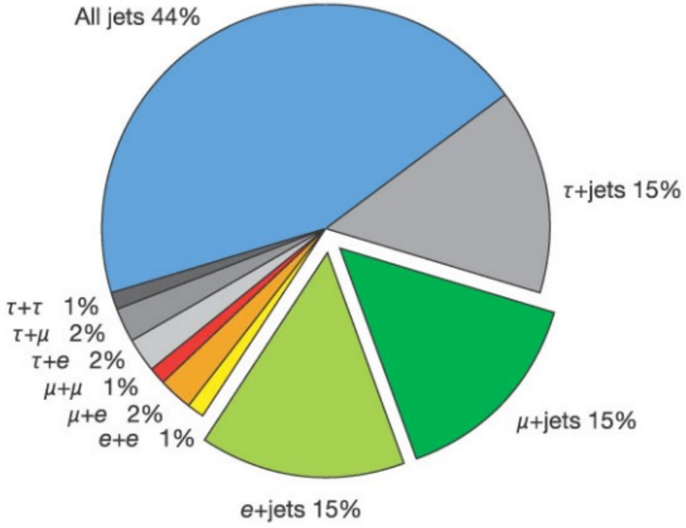
- Heaviest particle known so far.
- Decays before hadronises => top quark can be reconstructed precisely from decay products.
- Decays almost entirely into a W boson and a b-quark (although it is interesting to measure ratio to $|V_{tb}| + |V_{td}| + |V_{ts}|$).

LO e+e-



LO pp



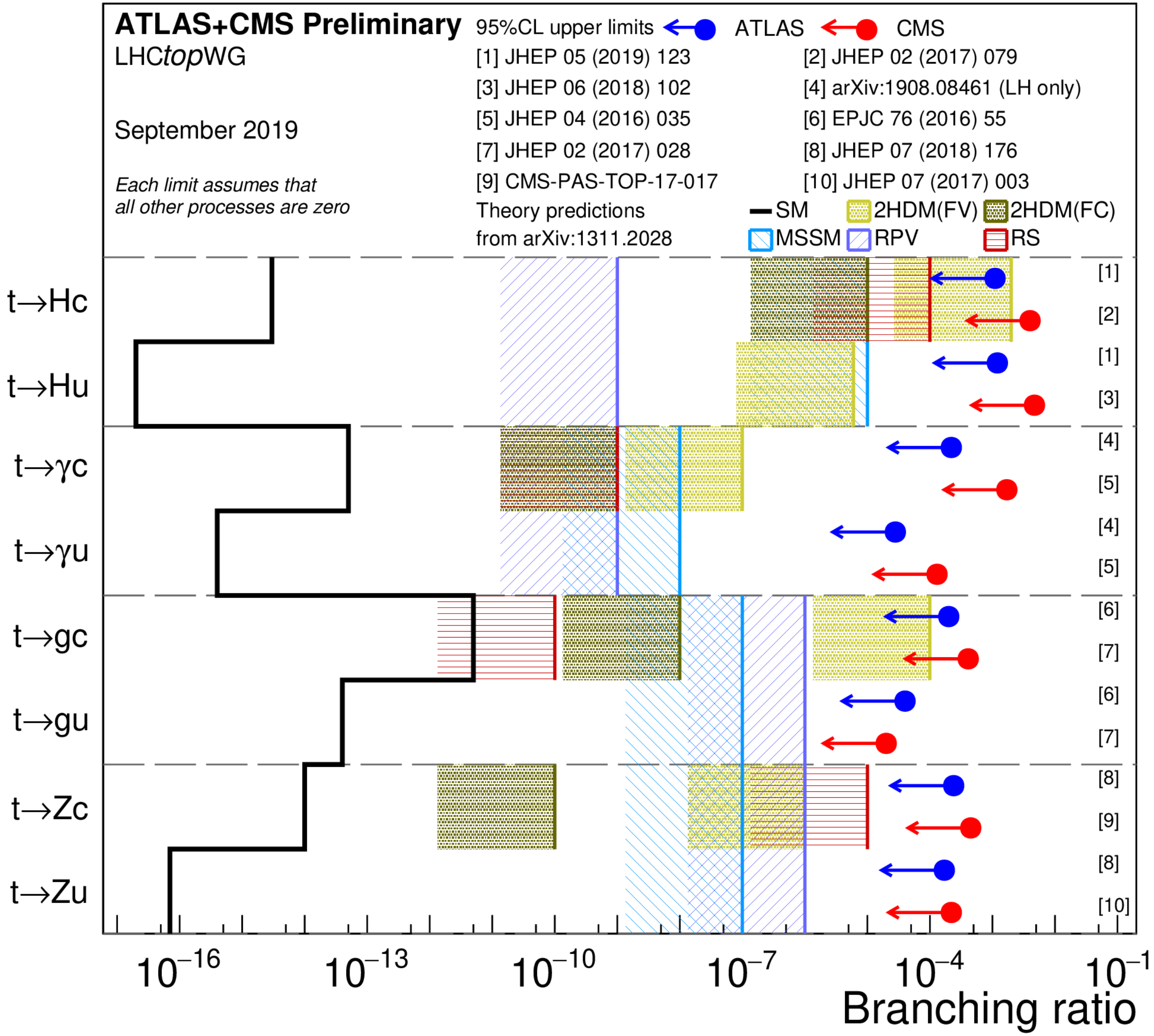


- At the LHC :
  - dileptonic channels are very precise => lower background and large lumi compensate the lower Br, top full reconstruction more challenging
  - Full hadronic very challenging because large QCD- multijet background,
  - Semi-leptonic channel shows a good compromise.

- At FCCee, the situation is totally different ! Moderate backgrounds for all channels, mainly from WW,
  - ➔ Easier to exploit all channels.

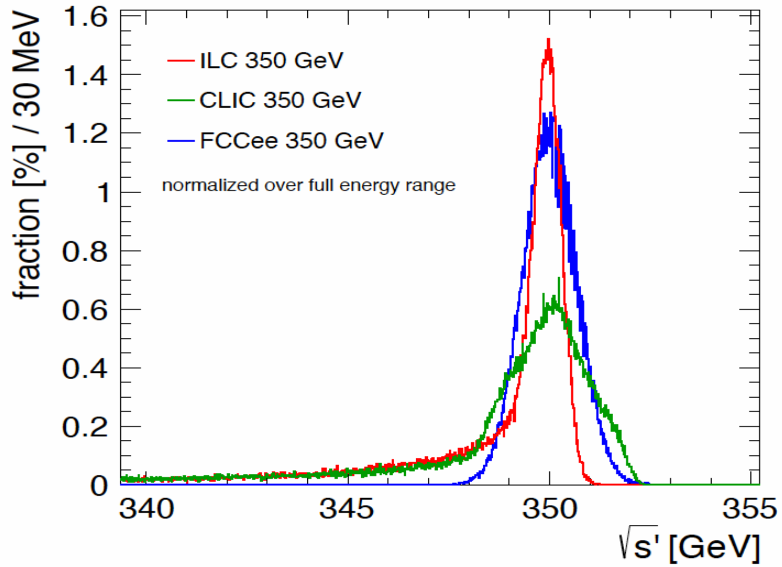
$t\bar{t}$ decays contained all objects ! Very relevant for determining detector requirements.

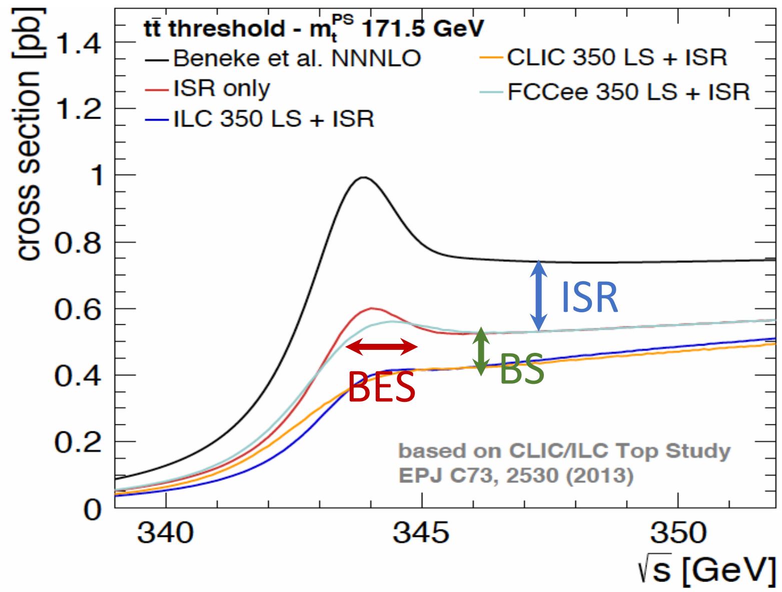
Status top FCNC at LHC



Beam characteristics and impact on $t\bar{t}$ at FCCee

F.Simon, PoS (ICHEP 2016) 872



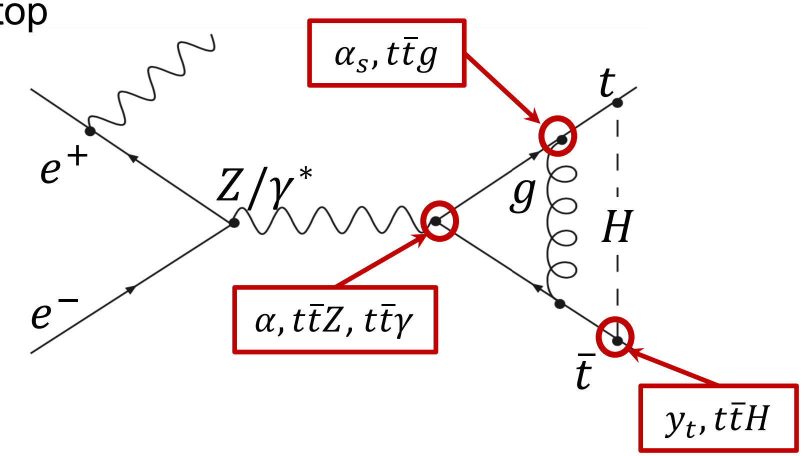


- Due to the energies involved and beam optimisation, there is some beam energy spread at $t\bar{t}$ thresholds.
- Beam energy spread, systematics ?
  - beam energy (spread) measured with dimuon-events at top energies to very precise values,
  - high muon resolution required, known better than ~10%.
- Can be adjusted with machine optimisation.
- At FCCee : narrow and no tails toward lower energies
  - Beamstrahlung effects on beam profile small, energy loss recovered by RF.
- Impacts of energy spread :
  - lower effective cross sections (fraction of the lumi below the threshold),
  - broader ”turn-on” ,
  - interplay between statistics and energy shape,
  - FCCee in a favourable position.

Has to be accounted in the simulation/interpretation !

$t\bar{t}$ precision measurements

- Inclusive and differential => probe of $t\bar{t}Z$ and $t\bar{t}\gamma$ couplings (EFT related)., measure of α and top yukawa.

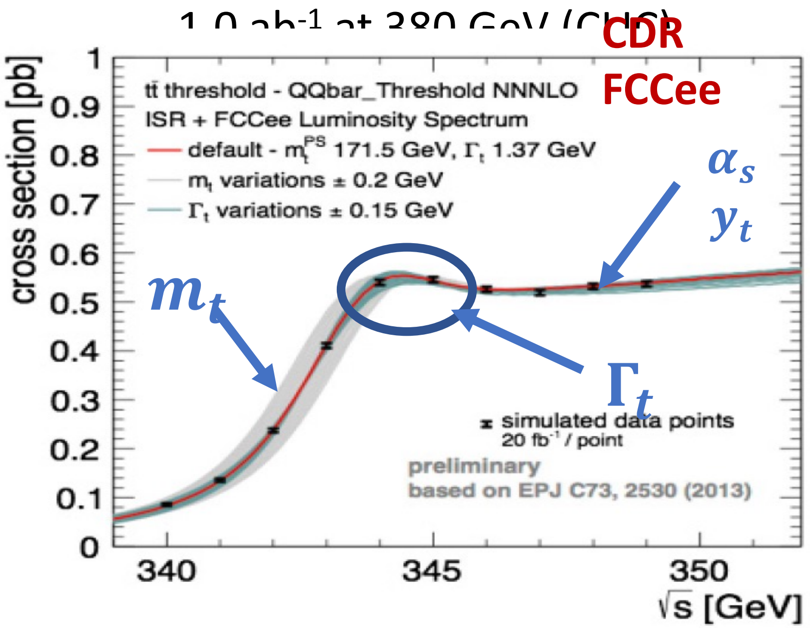


- Dominant backgrounds (lepton+jets):
  - WW(dominant)/ZZ
  - WWZ, ZH => more difficult to reject, but much lower cross section (/20).

- "typical" events selection in the lepton+jet channel :
  - one (relatively loose) isolated lepton with E>10 GeV, 80-90% efficiency,
  - 4 jets reconstructed using an exclusive algorithm (VLC),
  - b-tagging requirements,
  - jets and lepton association to top-quark, with a kin-fit (W and top mass, initial state!).

- Overall efficiency >70% can be achieved (JHEP 11 (2019) 003), very high purity (>90%).

- Target **systematics ~few %** (even below ?)
  - physics backgrounds very small,
  - High selection efficiency : related to detector performance (lepton/jets selection, flavour tagging) =: impact on acceptance and modelling uncertainties.
  - Excellent control of selection efficiencies (from data).



Top mass measurement

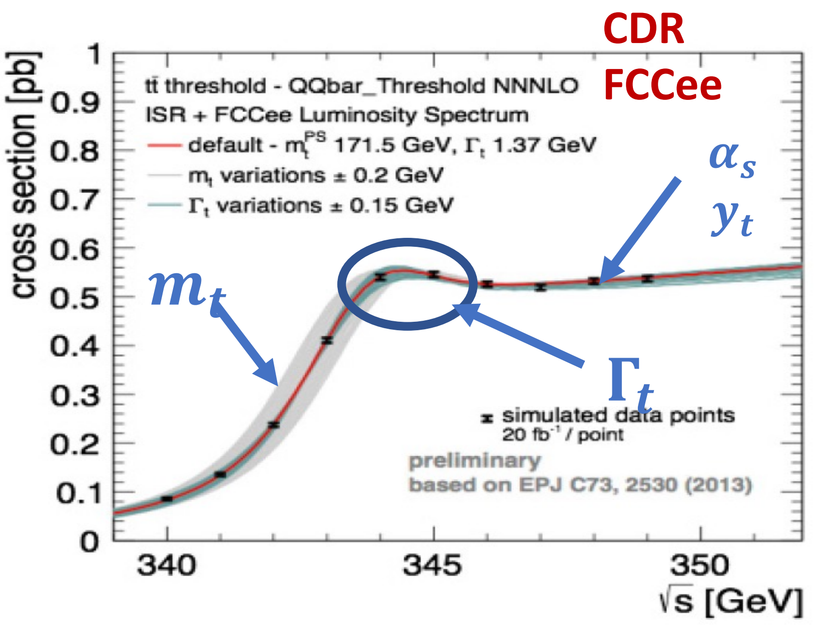
- Top mass measurement from cross sections => resolving top mass "ambiguities" : MC mass vs mass in various renorm. scheme. Also important to keep the universe safe against vacuum instability !

- Typical mass difference in the various renorm. schemes ~200 MeV.

- Mass extracted from various cross section measurements while scanning $\sqrt{s}$, and then compared to theoretical predictions.

- Cross section measurement precision : 1-2% to reach <200 MeV.

- Expected precisions (CLIC analysis revisited for FCCee):
  - Stat uncertainty at ~15 MeV,
  - Beam energy, reconstruction efficiency and background contamination ~50 MeV ,
  - And luminosity ... ~10 MeV,
  - **Total uncertainty below 100 MeV**, previous measurements of α_s => **reduction to < 50 MeV could be achievable**!



- Experimental uncertainties (close to be) dominated by statistics is possible at the FCCee !

- Still significant impact of theory uncertainties : requires a significant effort from theory community.