

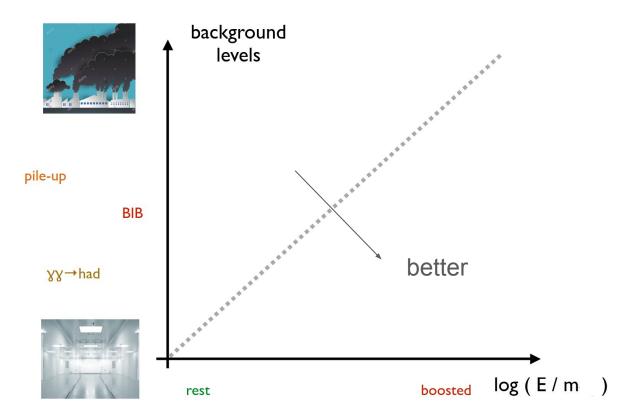
1

Jets at Future Colliders

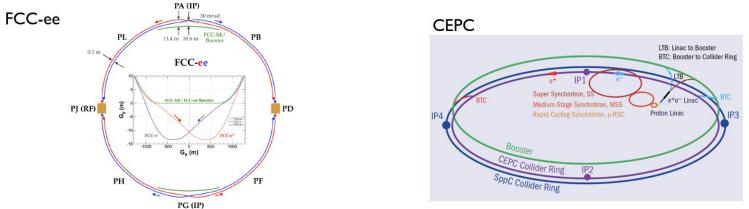
Michele Selvaggi (CERN)

Boost 2024

Collider phase space



FCC-ee/CEPC



91 km storage ring e⁺e⁻ collisions Higgs/EWK/Top factory

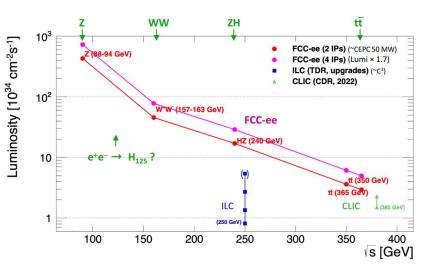
15 (20?) years of operations

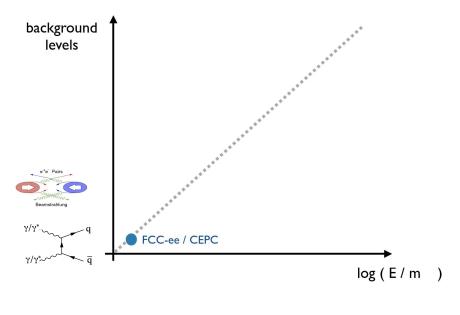
	Z pole	? H pole ?	ww	ZH	ttbar
\sqrt{s} [GeV]	88 - 91 - 94	125	157 - 161	240	350 - 365
Lumi / IP [10 ³⁴ cm ² s ⁻¹]	182	80	19.4	10.8	1.33
Int. lumi / 4IP [ab ⁻¹ / yr]	87	38	9.3	5.2	0.65
N _{years}	4	5	2	3	5
N _{events}	8 Tera	8 K	300 M	2.2 M	2 M

FCC-ee/CEPC

Exquisite luminosity allows for ultimate precision:

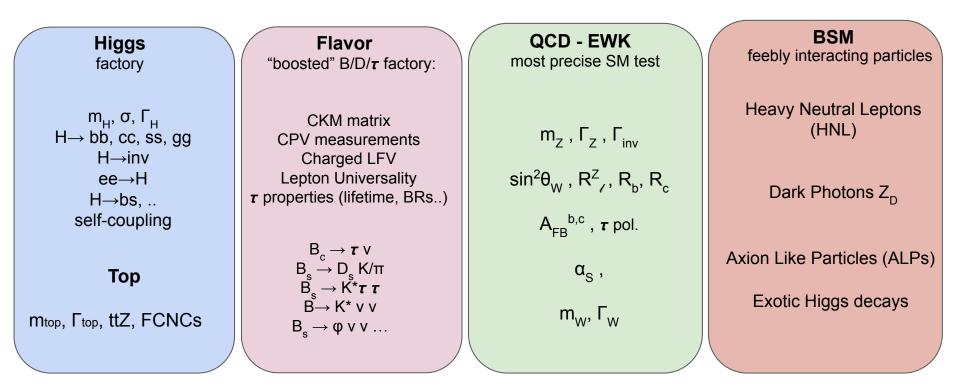
- 100K Z bosons / second
 - LEP dataset in 1 minutes
- 10k W boson / hour
- 2k Higgs bosons / day
- o 3k tops / day





- small backgrounds (mostly forward)
 - Beamstrahlung
 - Incoherent pair production
 - ∎ vv→hadrons
- small boost

Physics landscape at the FCC-ee/CEPC

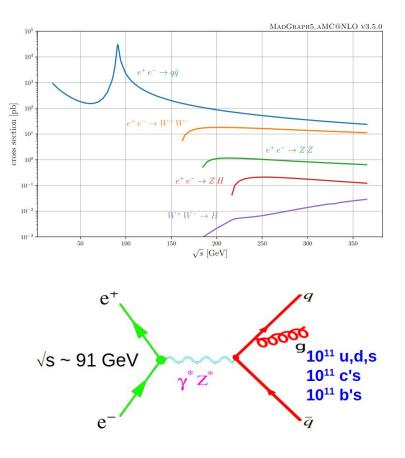


FCC-ee/CEPC

A clean jet factory:

- 10¹² jets at the Z pole
 - \circ 10¹¹ b,c,s,u,d, tau jets
 - 10¹⁰ gluons in 3 jet events
- 10⁸ at the WW threshold
- 10⁶ at the ZH and ttbar thresholds
 - 10^5 H→gg events
- Advantages compared to p-p collisions:
 - QED initial-state with known kinematics
 - QCD radiation only in final-state
 - well-defined heavy-Q, quark, gluon jets
 - no PDFs, no QCD "underlying event",...
- Direct clean parton fragmentation & hadroniz.

Perfect lab to **study QCD** (α_s , fragmentation, jet substructure)



See for a review: [d'enterria FCC week]

Higgs (Factory)

Ο

Ο

0

Ο

•

Most Higgs decays are hadronic

FCC-ee produces ~ 2M (mostly from ZH),

maximal statistical precision!

2j and 4j final states modes

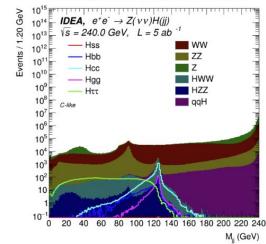
For rare channels ($H \rightarrow cc, H \rightarrow ss$) need:

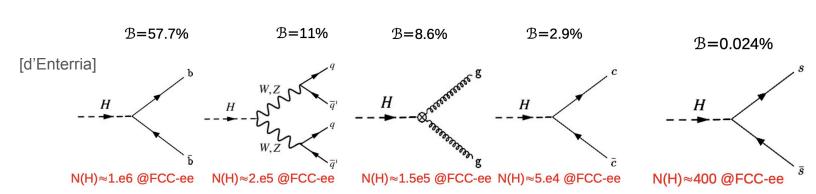
high tagging efficiency

crucial to exploit all of them for

excellent visible mass resolution

sensitivity ~ $\sqrt{\sigma(m_{vis})}$





FCCAnalyses: FCC-ee Simulation (Delphes)

Jet Clustering at lepton colliders

- **spherical** symmetry
- beam direction "z" not special
 - (as opposed to LHC where long. boost invariance along beam axis)
- distance measure should use
 - \circ **E**_i, **θ**_{ij}
- relative **absence** of machine **backgrounds**
 - every hadron should be clustered in a jet

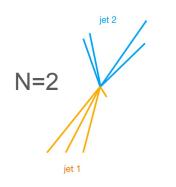
baseline: Durham k_{T} algorithm:

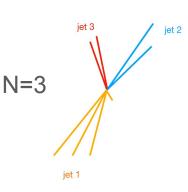
$$d_{ij} = 2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$$

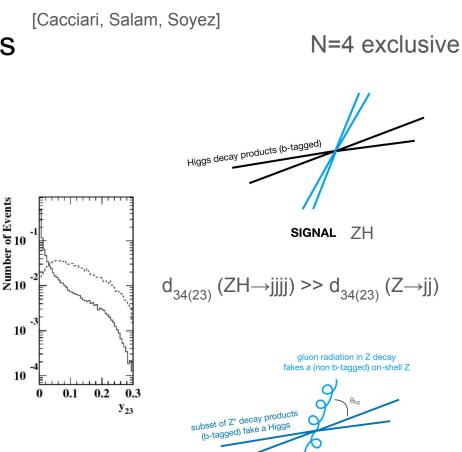
Jet Clustering at lepton colliders

virtually no ISR (only QED):

- **number of jets** in the final states well known both for signal and backgrounds
- baseline: exclusive clustering N mode:
 - stop when number of required jets reached
 - store merging distances d_{ij} for further background discrimination







Particle Flow and detector requirements

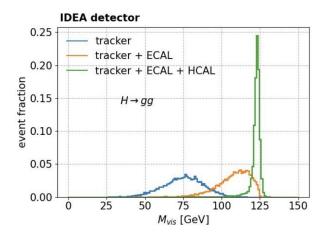
- To maximise visible energy/mass resolution:
 - every final state particle should be reconstructed

$$\sigma_{\rm PF}^2(E) = \sum_{i \in \rm tracks} \sigma_{\rm track}^2(E_i) + \sum_{i \in \gamma} \sigma_{\rm E}^2(E_i) + \sum_{i \in \rm had} \sigma_{\rm H}^2(E_i)$$

$$\frac{60\%}{30\%} \frac{30\%}{10\%}$$

Requirements:

- hadron energy resolution
- **100% tracking eff.**, photon, neutral hadrons reco. (n, KL) efficiency
 - low mat budget in front of calo, low noise
 - excellent granularity for optimal charged component and neutral hadron identification



Ideal PF

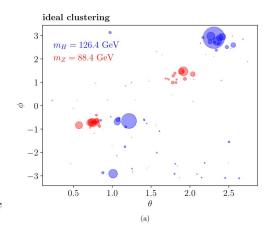
Resolution [GeV]	Crystal Cu/Brass (CMS)	LAr TileCal (ATLAS)	Dual Readout	Dual Readout +Crystal
S_{ECAL}	5%	10%	10%	5%
S_{HCAL}	100%	50%	30%	30%
σ_{ECAL}	0.3 GeV	0.6 GeV	0.6 GeV	0.3 GeV
σ_{HCAL}	3.7 GeV	1.8 GeV	1.1 GeV	1.1 GeV
σ	3.7 GeV	1.9 GeV	1.2 GeV	1.1 GeV

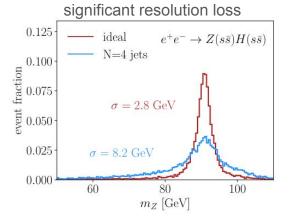
Color Singlet Clustering (FCC-ee)

- **boosted** resonances are "easy":
 - jet substructure techniques
- resolved multi-resonance events are harder (e.g. $ZH \rightarrow jjjj$):
 - jet clustering mixes decay products from the two resonances
 - ad-hoc criteria to producing pairings from N=4 jets to N=2 singlets
 - standard heuristics pairs of jets that match singlets mass
 - risk of sculpting backgrounds

$$\chi^2 = \left(\frac{m_{i_1i_2} - m_{S_1}}{\sigma_{S_1}}\right)^2 + \left(\frac{m_{i_3i_4} - m_{S_2}}{\sigma_{S_2}}\right)^2$$

[Garcia, Gergaud, MS]

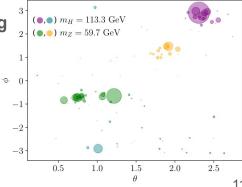




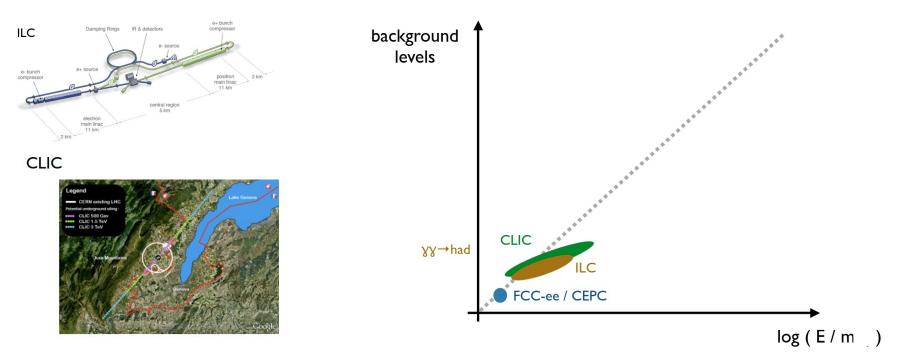
\rightarrow supervised (resolved) color singlet clustering

Method	Resolution Higgs in HZ [GeV]	Z1 in ZZ $[GeV]$	S/B	Significanc
Transformer	10.6	10.5	7.713e-04	1.804e-01
Gatr	9.2	10.4	1.294e-03	2.070e-01
Z	19.7	21.3	1.167e-03	1.826e-01
ZH	6.8	17.5	5.609e-04	1.532e-01
MCJets	6.8	21.3	2.150e-03	2.997e-01

N=4 jet clustering



High energy linear colliders



- Linear collider can reach 1-3 TeV
- Gives access to ttH, HH

High energy lepton colliders

Jet clustering "Valencia Linear Collider" (VLC)

• γγ→hadrons background (isolated energetic, forward)

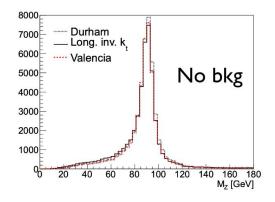
• beta exponent additional parameter which allows for tuning algorithm

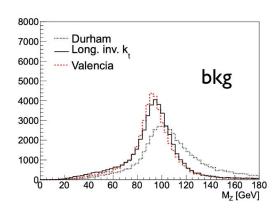
· governs likelihood of clustering background

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

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

[Boronat, Garcia, Vos]





Kinematic Fit at lepton colliders

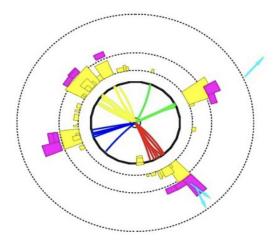
precise **knowledge of center of mass energy**: \rightarrow kinematic fit can be used to improve resolution

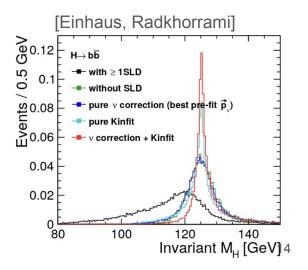
e.g for 4 jet final states (ZH, WW, ..), kinematics overconstrained

for i = 1, ..., N observed objects, $\sum_{i} \vec{p}_{i} = 0$ and $\sum_{i} E_{i} = 2E_{beam} = \sqrt{s}$

Derive covariance matrix C for $(\textbf{p}_1, \ldots, \textbf{p}_4)$ and minimize with Lagrange multipliers

$$\chi^{2}(\hat{p}_{1},\hat{p}_{2},\hat{p}_{3},\hat{p}_{4}) = \sum_{k=1}^{4} \left[\sum_{\alpha\beta} \left(p_{k}^{\alpha} - \hat{p}_{k}^{\alpha} \right) (\mathbf{C}_{k}^{-1})_{\alpha\beta} \left(p_{k}^{\beta} - \hat{p}_{k}^{\beta} \right) \right] \\ + \lambda_{0} \left[\sum_{k} \hat{E}_{k} - \sqrt{s} \right] + \sum_{a=1}^{3} \lambda_{a} \left[\sum_{k} \hat{p}_{k}^{a} \right] + \mu \left[(\hat{E}_{a} + \hat{E}_{b}) - (\hat{E}_{c} + \hat{E}_{d}) \right]$$





Jet tagging at FCC

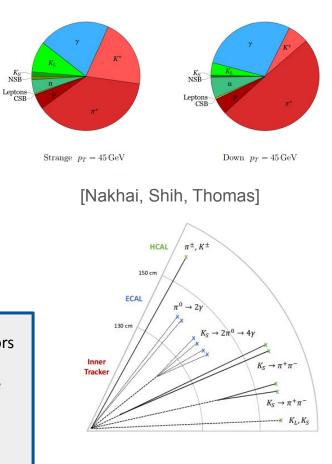
- b/c-tagging:
 - Large lifetime, (2-3) mm for ~50 GeV boost Ο
 - Displaced vertices/tracks (Large impact parameters) 0
 - Large track multiplicity Ο
 - Presence of non-isolated e/μ (20 (10)% in B (C) decays) 0
- s-tagging:
 - Large Kaon content:
 - as tracks (K/pi separation ToF, dEdx, dNdx)
 - Neutral Kaons:
 - K_s 2 tracks
 - K, ToF vs n

Detector constraints:

Need power pixel/tracking detectors

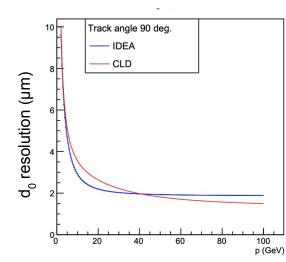
K_S-NSB-

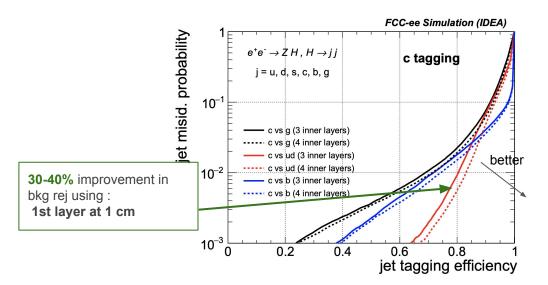
- Good spatial resolution
- As little material as possible _
- Precise track alignment
- Timing detectors -
- Charged energy loss (gas/silicon)

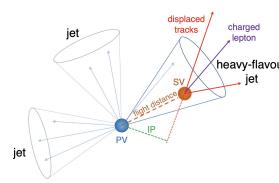


Track impact parameter resolution and vertexing

- Impact parameter resolution major driver of jet charm and bottom jet identification
- precise IP determination driven by:
 - single point resolution
 - **radial distance of first tracking layer** from the interaction point (at large momentum)
 - need small radius beam-pipe
 - material budget X/X₀ (at low p)







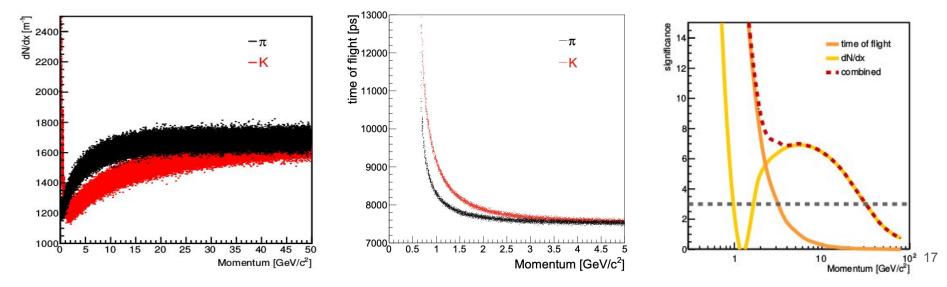
16

Strange tagging FCC

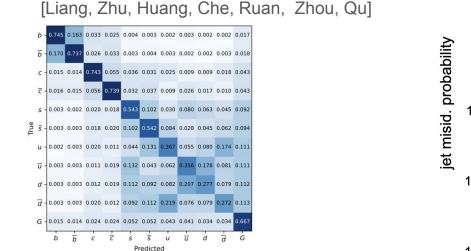
- Kinematic
- Displacement (important for b-tagging)
- Particle Identification:
 - Number of ionization clusters (dN/dx)
 - \circ ToF results in good K/ π separation at low-momenta

input

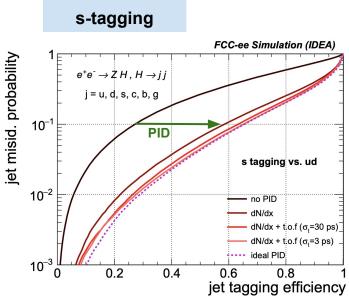
[Bedeschi, Gouskos, MS]



Jet tagging FCC



Final state	Comb. [%]
$H \rightarrow bb$	0.22
$H \rightarrow cc$	1.70
$H \rightarrow gg$	0.9
$H \rightarrow ss$	120



strange vs light rejection crucial vs Z(ss)Z(jj) and Z(ss)H(gg) backgrounds

eff = 65% , mistag = 10%

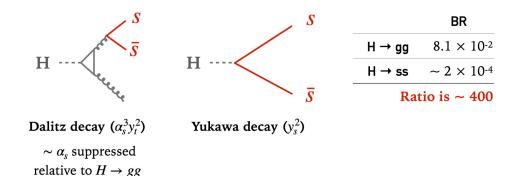
Can approach SM strange Yukawa sensitivity at FCC-ee l_{18}

Open questions in Hss and strange fragmentation

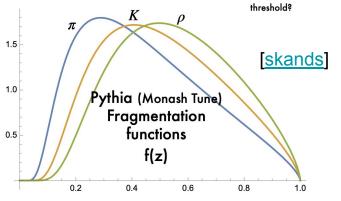




- contamination seems under control since lives in a different phase space
 - $H \rightarrow gg$ at N³LO required
 - with NLL showers

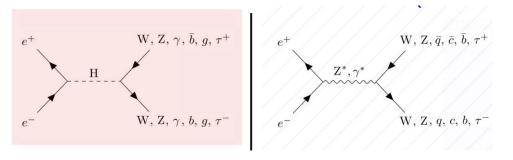


- strange fragmentation functions?
 - used in our taggers
 - average FF well constrained
 - width is not
 - High precision hadron data required
 - $\circ \quad 10^{11} Z \rightarrow qq \text{ events } !$

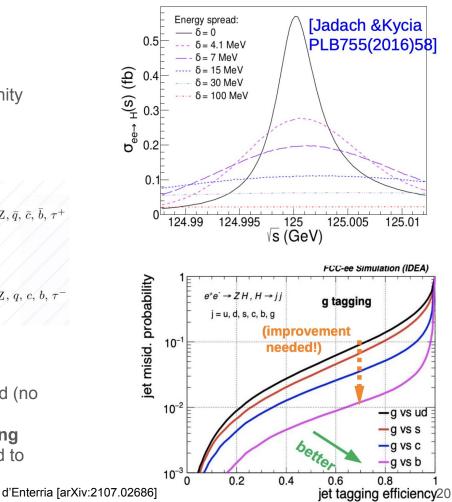


$H \rightarrow gg$ and gluon tagging

- s-channel ee→H production provides unique opportunity to probe electron Yukawa
- very challenging, even with mono-chromatisation
 - ISR and BES smear the exp. width

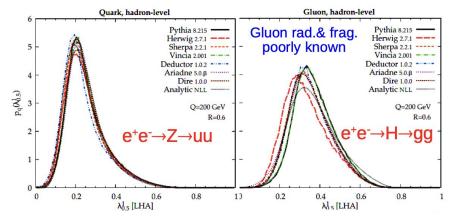


- exploit H→gg final state against Z→qq background (no prompt gg)
- requires 1% light quark mistag rate in gluon tagging
 - 10x light rejection vs state-of-the-art required to approach electron yukawa
 - SM sensitivity in ~ 2 years

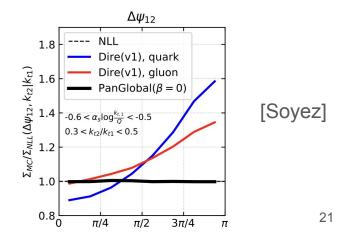


Gluon tagging, open questions

- (g vs q) taggers make use of LL parton showers, that differ substantially from one MC to another, in particular for gluons
- quark showers well constrained from LEP, gluons less-so

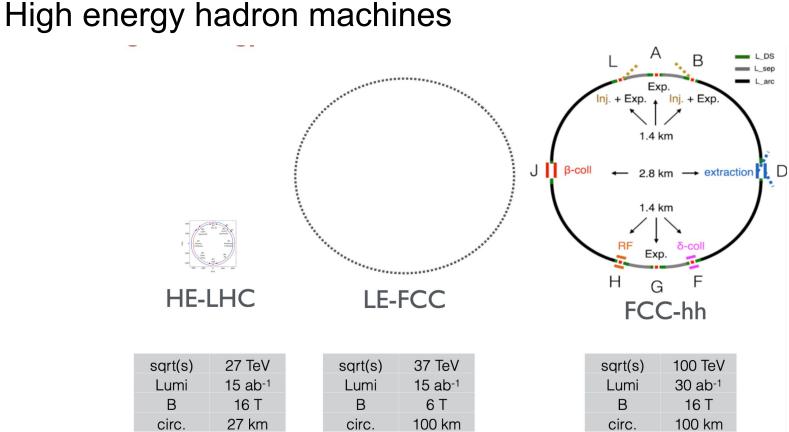


un-physical differences in q vs g in LL shower



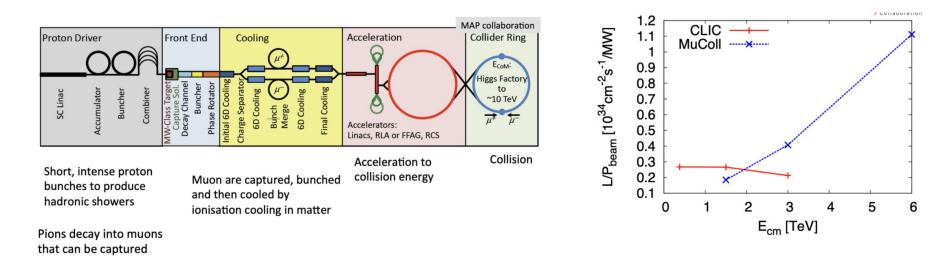
Need:

- (N)NLL accurate showers (e.g. Panscale)
- clean data for tuning (in varying kin. regimes exploiting ISR):
 - \circ 10⁵ H \rightarrow gg events
 - $\circ \quad 10^{11} \text{ Z} \rightarrow \text{bbg}$



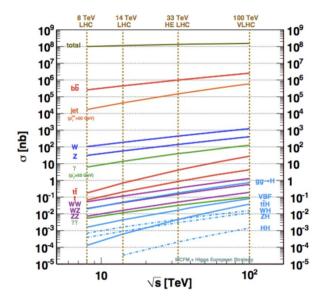
Main challenge: high field superconducting > 14 T magnets , high PU

High energy muon collider



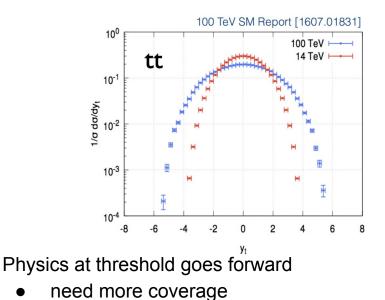
- only lepton machine capable of reaching multi TeV energies in a circular ring
 - electrons radiate too strongly
- many challenges (cooling, neutrino hazard, beam induced background ..)

Processes at FCC-hh



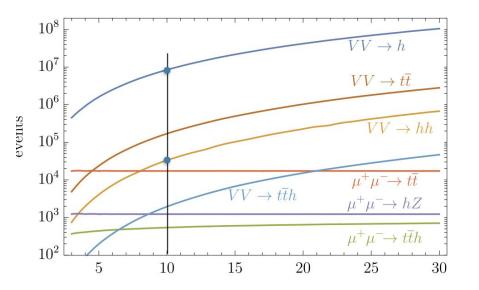
- Among all SM "backgrounds", ttbar production gains the most in rate @100 TeV
 - tt > HH > VV > H > V

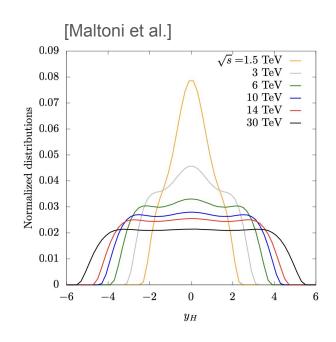
Massive boosted samples (30 ab^{-1}): 10¹² tops 10⁹ top with $p_T > 1$ TeV 100k top with $p_T > 5$ TeV 1000 top with $p_T > 10$ TeV



FCC-hh physics benchmarks with jets : HH \rightarrow bbXX, ttH, ttZ, ...

Processes at Muon collider





Physics at threshold goes forward

- Production modes:
 - direct mass BSM via s-channel (plus recoil eventually)
 - \circ Weak boson fusion

Muon physics benchmarks with jets : H, HH, HHH, EWK multiplets

Hadron Machines specs and detector requirements

lumi & pile-up

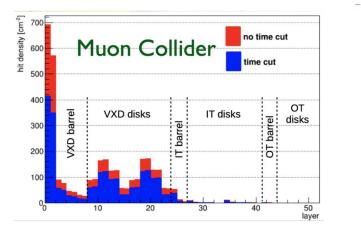
	parameter		unit	LHC	HL-LHC	HE-LHC	FCC-hh
	E_{cm}		TeV	14	14	27	100
	circumference		km	26.7	26.7	26.7	97.8
	peak $\mathcal{L} \times 10^{34}$		$\mathrm{cm}^{-2}\mathrm{s}^{-1}$	1	5	25	30
	bunch spacing		ns	25	25	25	25
	number of bunches			2808	2808	2808	10600
	goal $\int \mathcal{L}$		ab^{-1}	0.3	3	10	30
	σ_{inel}		mbarn	85	85	91	108
	σ_{tot}		mbarn	111	111	126	153
	BC rate		MHz	31.6	31.6	31.6	32.5
	peak pp collision rate		GHz	0.85	4.25	22.8	32.4
	peak av. PU events/BC			27	135	721	997
	rms luminous region σ_z line PU density time PU density $dN_{ch}/d\eta _{\eta=0}$		mm	45	57	57	49
			mm^{-1}	0.2	0.9	5	8.1
			ps ⁻¹	0.1	0.28	1.51	2.43
				7	7	8	9.6
	charged tracks per collision N_{ch}	- 1		95	95	108	130
	Rate of charged tracks $< p_T >$		GHz	76	380	2500	4160
			GeV/c	0.6	0.6	0.7	0.76
Number	lumber of pp collisions		10^{16}	2.6	26	91	324
Charged	Charged part. flux at 2.5 cm est.(FLUKA)		$Hz cm^{-2}$	0.1	0.7	2.7	8.4 (12)
1 MeV-n	MeV-neq fluence at 2.5 cm est.(FLUKA)		$1^{16}{ m cm}^{-2}$	0.4	3.9	16.8	84.3 (60)
Total ior	Fotal ionising dose at 2.5 cm est.(FLUKA)		MGy	1.3	13	54	270 (400)
$dE/d\eta _{\eta}$	$dE/d\eta _{\eta=5}$		GeV	316	316	427	765
$dP/d\eta _{\eta}$	$dP/d\eta _{\eta=5}$		kW	0.04	0.2	1.0	4.0

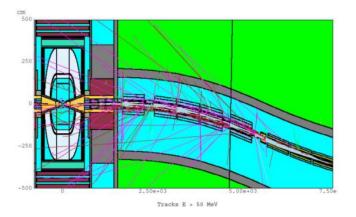
→ x6 HL-LHC

LHC: 30 PU events/bc HL-LHC: 140 PU events/bc FCC-hh: 1000 PU events/bc

but also $\times 10$ integrated luminosity w.r.t to HL-LHC

Muon collider machine specs

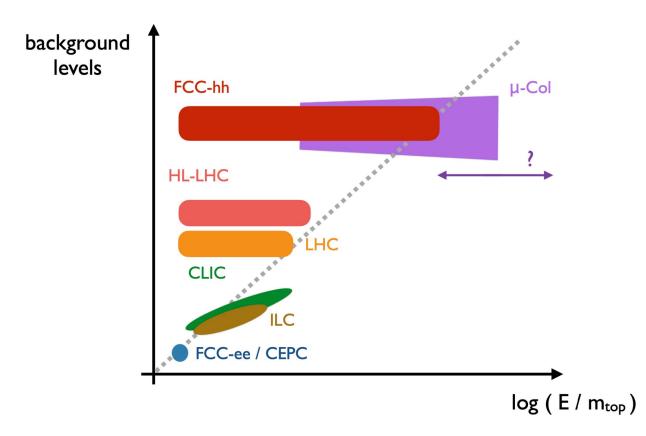




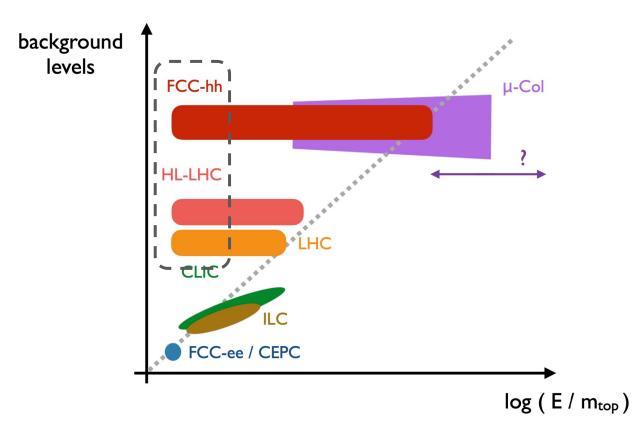
charged fluence: 400-700 (cm⁻² / BX)

- high occupancy (HL-LHC < muCol < FCC-hh), but low collision rate (~ 50 kHz vs 40 MHz)
- at threshold (or low energy) jet reconstruction will suffer from similar limitations as the FCC-hh (large PU → large Beam induced background)
 - despite some conceptual differences (directionality, energy $\dots \rightarrow$ timing cuts)
 - jet definition: probably Valencia like algorithm in exclusive mode since BIB (but no ISR, no UE)

Future high energy facilities

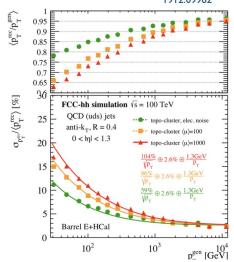


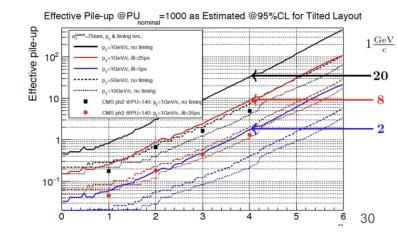
Muon colliders



Experimental challenges for jets (at threshold)

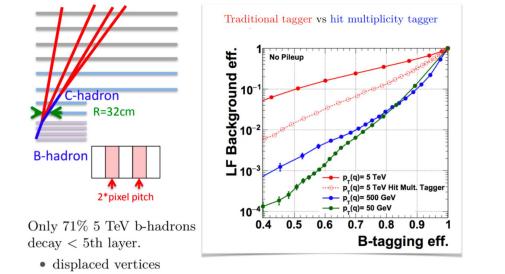
- relative impact of PU is large on:
 - jet energy resolution and scale
 - HF-tagging (b/c-tagging)
- PU subtraction techniques
 - charged hadron subtraction
 - timing information (5-10 ps resolution)
 - forward!
 - Residual:
 - area-subtraction
 - PUPPI reconstruction
 - advanced graph based-ML





1912.09962

High pT flavor tagging



Perez Codina, Roloff [CERN-ACC-2018-0023]

To be verified in high pile-up environment.

arXiv:1701:06832

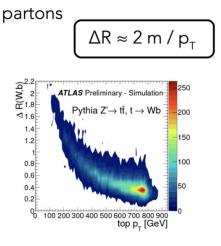
Change in paradigm: heavy flavour tagging

• multi-TeV b-Hadrons decay outside the pixel volume ($p_T(b) = 2 \text{ TeV} \rightarrow \gamma c \tau = 50 \text{ cm}$)

Need to adapt identification algorithms for identifying multi-TeV tops

Boosted topologies at multi-TeV energies

min. distance to resolve two



ex for top:

 $\begin{array}{rcl} p_{T} = & 200 \; \text{GeV} & \rightarrow & \text{R} \sim 2 \\ p_{T} = & 1 \; \text{TeV} & \rightarrow & \text{R} \sim 0.4 \\ \textbf{p}_{T} = & \textbf{10} \; \text{TeV} & \rightarrow & \text{R} \sim 0.05 \end{array}$

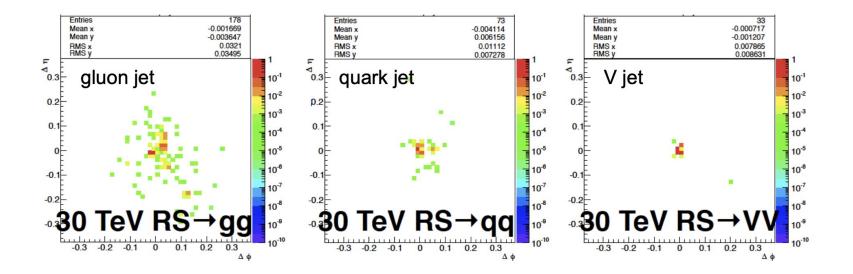
- At 10 TeV whole jet core within 1 calo cell
 - neutrals possibly un-resolvable
 - B field "helps" with charged
 - PF reconstruction will be severely affected
 - Total jet energy OK, calo does good job
 - reed to be studied and rethought for
- Naive approach:
 - use calo for energy measurement
 - tracking for substructure identification

in CMS:

 $\begin{array}{rcl} \mbox{Tracking} & \rightarrow & \Delta R \sim 0.002 \\ \mbox{ECAL} & \rightarrow & \Delta R \sim 0.02 \\ \mbox{HCAL} & \rightarrow & \Delta R \sim 0.1 \end{array}$

Color Singlets (W/Z/H)

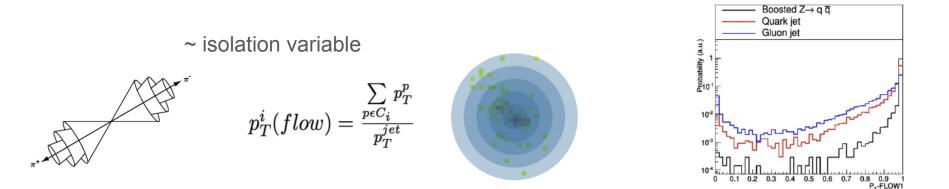
[Pierini]



- **Gluon/quark** jet looks the same at 50 GeV and 5 TeV (**QCD** is ~ scale invariant)
- Color Singlets look like taus (do not radiate, a part from occasional QED/EWK shower)
 - high mass, highly isolated, highly collimated tracks

Boosted Color Singlet ID

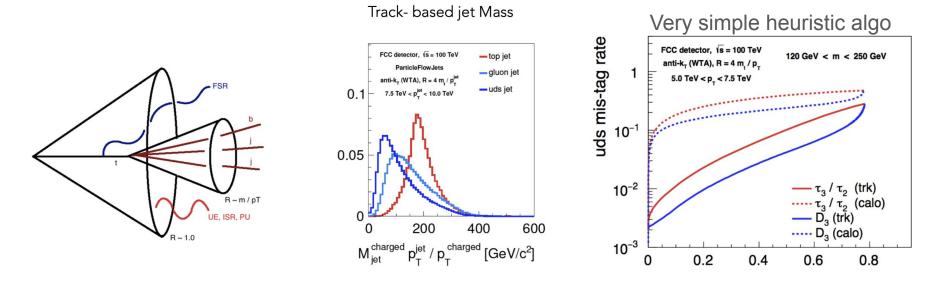
[Pierini]



Loss in performance, but no show stoppers

Very simple heuristic based , can probably do much better with today's techniques

Boosted Colored Resonances



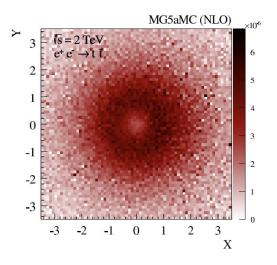
- Multi TeV top radiates FSR at a typical scale angular scale ~ m / pT (deadcone)
- Large cone FSR can spoil mass by adding $\Delta m \sim m_{top}$ even for 1 GeV emission
 - $\circ \rightarrow$ use shrinking cone algo by reclustering with R ~ 4m/pT
 - use tracking for substructure

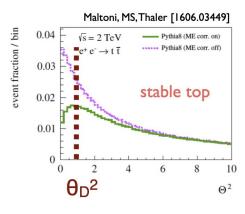
The deadcone effect for massive colored res.

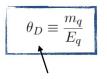
FSR in soft and collinear limit :

$$\frac{1}{\sigma} \frac{\mathrm{d}^2 \sigma}{\mathrm{d}z \,\mathrm{d}\theta^2} \simeq \frac{\alpha_S}{\pi} C_F \frac{1}{z} \frac{\theta^2}{(\theta^2 + \theta_D^2)^2}$$

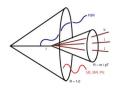
- effect can be observed at HL-LHC
- rather than treated as a nuisance can be exploited for top tagging at multi TeV energies







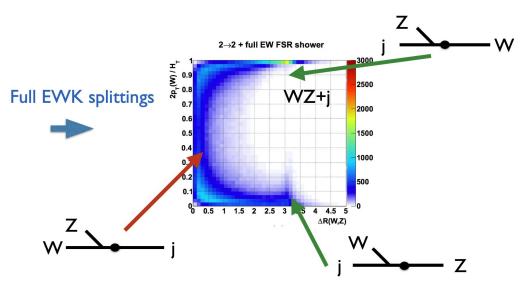
for the top can be pretty large angle



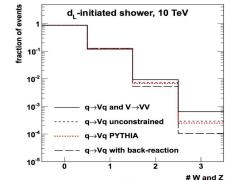
Electroweak showers

3000 2p_1(W) / H_1 8.0 8.0 8.0 2500 WZ+j 0.7 2000 0.6 0.5 1500 0.4 1000 0.3 0.2 500 0.1 0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 ∆R(W,Z)

2→2 + PYTHIA weak FSR shower



Chen, Han and Tweedie [1611.00788]



- EWK shower become sizeable log-enhanced at multi-TeV energies
 - $\circ \quad j \to j W \text{ can fake a top jet}$
- can and have to be included and studied in multi-TeV jet tagging
- Neutrino showers?

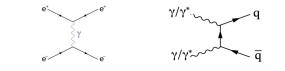
37

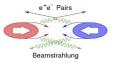
Summary

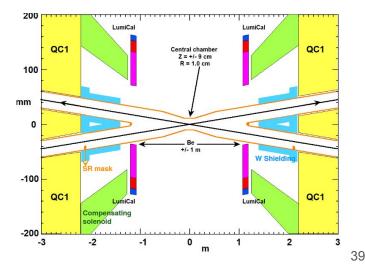
- Circular ee (FCC-ee/CEPC)
 - small boost, small background, well known initial state
 - Huge statistics 10¹² jets of any flavor (including tau's)
 - study jets (Q vs G), HF jets and calibrate taggers in data
- Linear ee machines (ILC/CLIC)
 - Low to moderate boost/backgrounds
- High energy lepton (*µ*-Col) and hadron collider (FCC-hh)
 - at threshold:
 - SM Physics is forward, challenging machine backgrounds (PU, BIB)
 - precise tracking/timing
 - Hyper boosted regime ($p_T > 10 \text{ TeV}$)
 - calorimeters cannot resolve substructure
 - tracking is key
 - new handles:
 - Isolation for color singlets
 - deadcone radiation

FCC-ee Detector requirements - general considerations

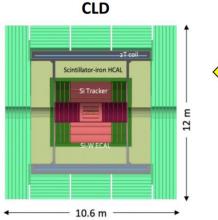
- Requirements for Higgs and above have been studied to some extent by LC:
 - we want a detector that is able to withstand a large dynamic range:
 - in energy (√s = 90 365 GeV)
 - in luminosity (L = $10^{34} 10^{36} \text{ cm}^2/\text{s}$)
- most of the machine induced limitations are imposed by the Z pole run:
 - large collision rates ~ 33 MHz and continuous beams
 - no power pulsing possible
 - large event rates ~ 100 kHz
 - fast detector response / triggerless design challenging (but rewarding)
 - high occupancy in the inner layers/forward region (Bhabha scattering/γγ hadrons)
 - o beamstrahlung
- complex MDI: last focusing quadrupole is ~ 2.2m from the IP
 - magnetic field limited to B = 2T at the Z peak (to avoid disrupting vertical emittance/inst. Lumi via SR)
 - limits the achievable track momentum resolution
 - o "anti"-solenoid
 - limits the acceptance to ~ 100 mrad



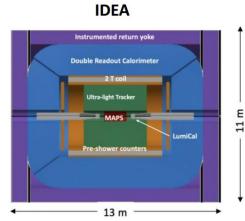




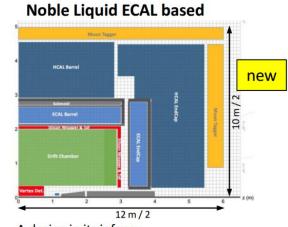
FCC-ee Detector Benchmarks



- Well established design
 - ILC -> CLIC detector -> CLD
- Full Si vtx + tracker;
- CALICE-like calorimetry;
- Large coil, muon system
- Engineering still needed for operation with continuous beam (no power pulsing)
 - Cooling of Si-sensors & calorimeters
- Possible detector optimizations
 - σ_p/p, σ_E/E
 - PID (**O**(10 ps) timing and/or RICH)?

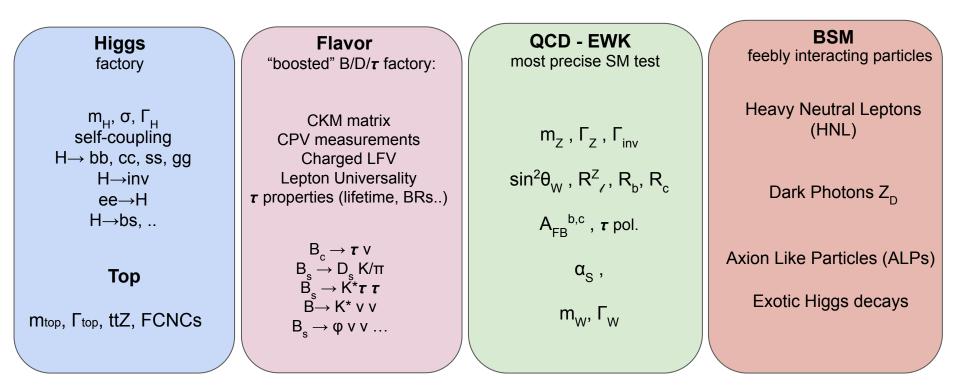


- A bit less established design
 - But still ~15y history
- Si vtx detector; ultra light drift chamber w powerful PID; compact, light coil;
- Monolithic dual readout calorimeter;
 - Possibly augmented by crystal ECAL
- Muon system
- Very active community
 - Prototype designs, test beam campaigns, ...



- A design in its infancy
- Si vtx det., ultra light drift chamber (or Si)
- High granularity Noble Liquid ECAL as core
 - Pb/W+LAr (or denser W+LKr)
- CALICE-like or TileCal-like HCAL;
- Coil inside same cryostat as LAr, outside ECAL
- Muon system.
- Very active Noble Liquid R&D team
 - Readout electrodes, feed-throughs, electronics, light cryostat, ...
 - Software & performance studies

Physics landscape at the FCC-ee/CEPC



Detector requirements at the FCC-ee/CEPC

Higgs factory

track momentum resolution (low X_0)

IP/vertex resolution for flavor tagging

PID capabilities for flavor tagging

jet energy/angular resolution (stochastic and noise) and PF **Flavor** "boosted" B/D/**τ** factory:

track momentum resolution (low X_0)

IP/vertex resolution

PID capabilities

Photon resolution, pi0 reconstruction QCD - EWK most precise SM test

acceptance/alignment knowledge to 10 µm

luminosity

Momentum resolution

BSM feebly interacting particles

Large decay volume

High radial segmentation - tracker - calorimetry - muon

> impact parameter resolution for large displacement

> > timing

triggerless