Learning Physics at Future e-e+ Colliders with Machine

Sijun Xu

The Hong Kong University of Science and Technology

28 Oct, 2020

Based on JHEP:10(2020)018 with Lingfeng Li, Ying-Ying Li and Tao Liu

Outline

- Limitations of Jets @Precision Frontier
- Jet + X vs Event-level ML
- Benchmark: 2 jets
- Application: Measure Γ(h) @240 GeV, 5ab-1

Precision Frontier of Next Decades

The precision frontier of next decades in Higgs and electroweak physics is expected to be defined by future e^-e^+ colliders.

Measurements (%)	$\text{CEPC}_{240(250)}$ [17, 19]	FCC_{240} [18]	FCC_{365} [18]	CILC_{350} [20]	ILC_{250} [21–23, 78]
$\sigma(Zh)$	0.5 (0.5)	0.5	0.9	1.6	2.8
$\sigma(Zh_b)$	0.27 (0.26)	0.3	0.5	0.86	1.2
$\sigma(Zh_c)$	3.3 (3.1)	2.2	3.5	14	8.3
$\sigma(Zh_g)$	1.3 (1.2)	1.9	6.5	6.1	7.0
$\sigma(Zh_W)$	1.0 (0.9)	1.2	2.6	5.1	6.4
$\sigma(Zh_Z)$	5.1 (4.9)	4.4	12	-	19
$\sigma(\nu\nu h_b)$	3.2 (2.9)	3.1	0.9	1.9	10.5
$\sigma(\nu\nu h_c)$	-	-	10	26	-
$\sigma(\nu\nu h_W)$	-	-	3.0	-	-

[F. An et al., 1810.09037; A. Abada et al., (2019); H. Abramowicz et al., 1608.07538]

Question: whether they fully represent the capability of the machine?

Precision Frontier of Next Decades

These precisions are typically set up by the measurement of hadronic events and analysis at jet level.

Primary Higgs and Electroweak processes

Jet Number	0	2	4	6
$e^-e^+ \to WW$	11%	44%	45%	0%
$e^-e^+ \to ZZ$	9%	42%	49%	0%
$e^-e^+ \to ZH$	3%	32%	55%	11%
$e^-e^+ \to H\nu\nu$	20%	69%	11%	0%
$e^-e^+ \to t\bar{t}$	0%	11%	44%	45%

Hadronic mode dominant

How would jet clustering affect these precisions?

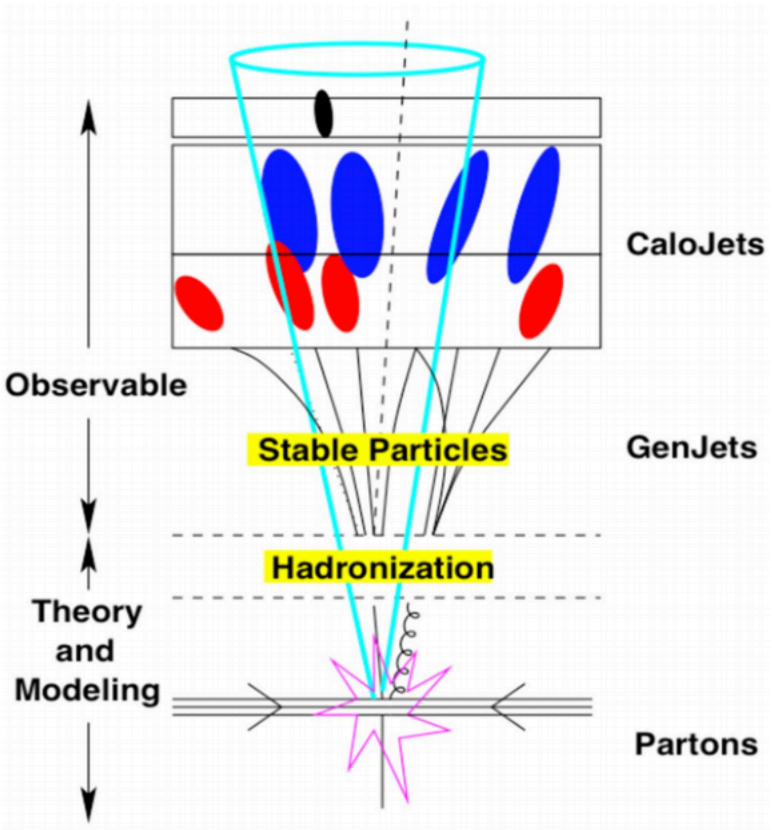
Jet Clustering ($\boldsymbol{ee\text{-}k_T}$)

Jet: A collimated spray of stable particles from the fragmentation and hadronization of a parton.

Jet Clustering: Algorithms used to combine the calorimetry and tracking information to define jets.

$\boldsymbol{ee\text{-}k_T}$:
[S. Catani et. al., 1991]

- Originally designed for e^-e^+ colliders
- Priority for soft components
- Fixed jet number



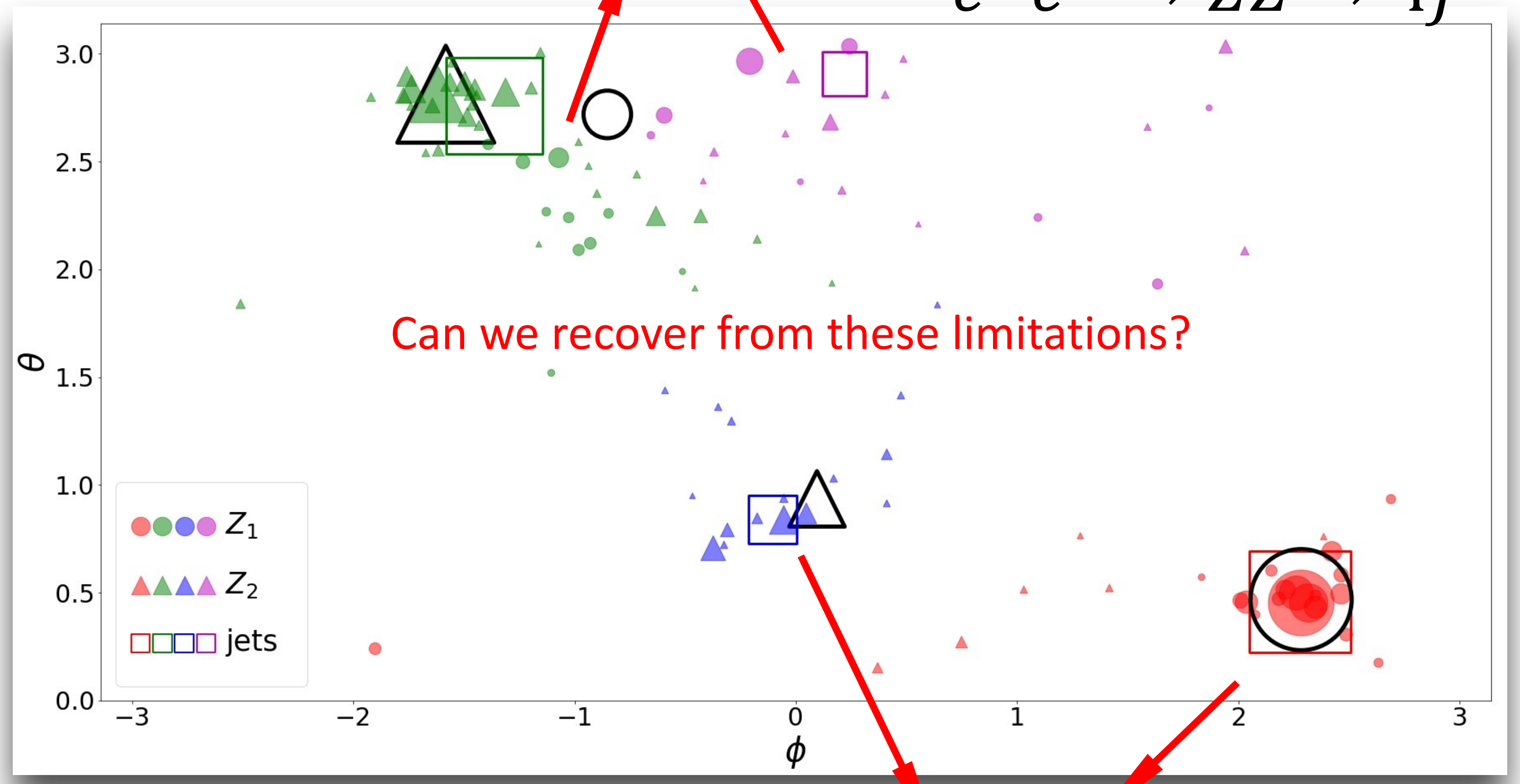
[Ryan Atkin 2015 J. Phys.: Conf. Ser. 645 012008]

Consider an event of $e^-e^+ \to ZZ \to 4j$ with $\boldsymbol{ee\text{-}k_T}$ algorithm

Limitation of Jet Clustering

Hadrons from different Z clustered in a same jet (info distortion)

$e^- e^+ \rightarrow ZZ \rightarrow 4j$



Can we recover from these limitations?

Detailed structures are gone after clustering (info loss)

First Way: Jet + Event-Level Obs.

Find a systematic way to organize event-level information.

- Angular distribution of energy

$$E(\Omega) = \sum_{i \in \text{event}} E_i \delta(\Omega - \Omega_i)$$

Spherical harmonics expansion

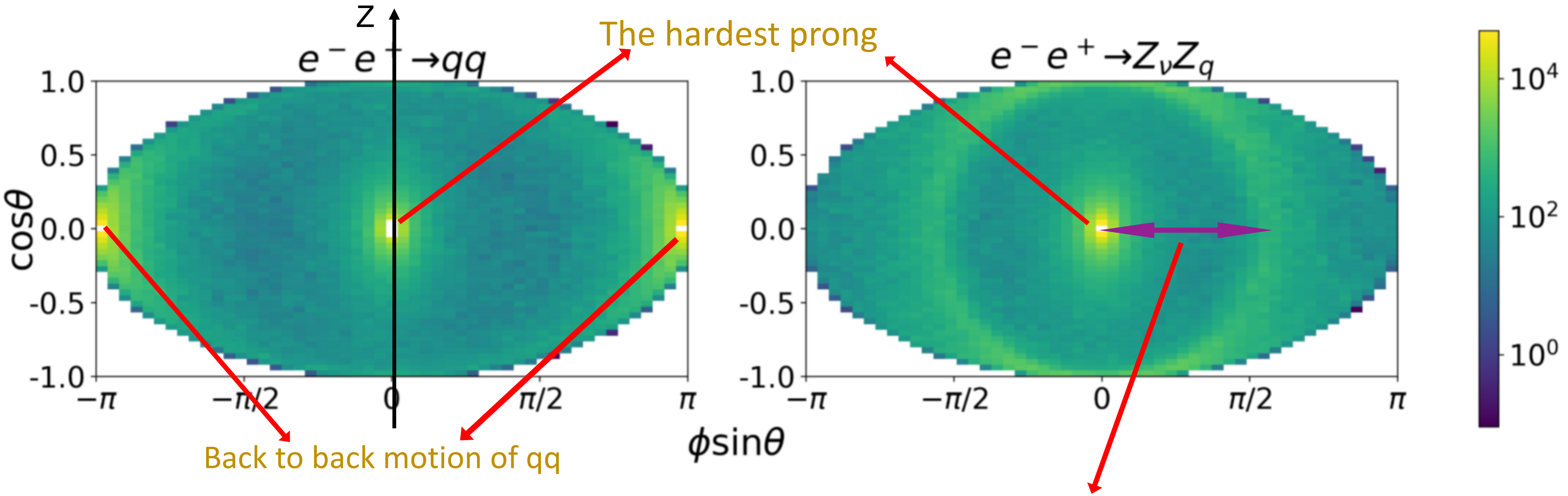
$$E(\Omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_l^m Y_l^m(\Omega) \qquad a_l^m = \sum_{i \in \text{event}} E_i Y_l^{m*}(\Omega_i)$$

- Fox-Wolfram moments [G. C. Fox and S. Wolfram, 1978] and extensions:

$$H_{EE;l} = \frac{1}{s} \frac{4\pi}{2l+1} \sum_{m=-l}^{l} a_l^{m*} a_l^m = \sum_{i,j \in \text{event}} \frac{E_i E_j}{s} P_l(\cos(\Omega_{ij}))$$

the event-level information is encoded as the FW moments at leading order and multi-spectra at higher orders.

Cumulative Mollweide Projection

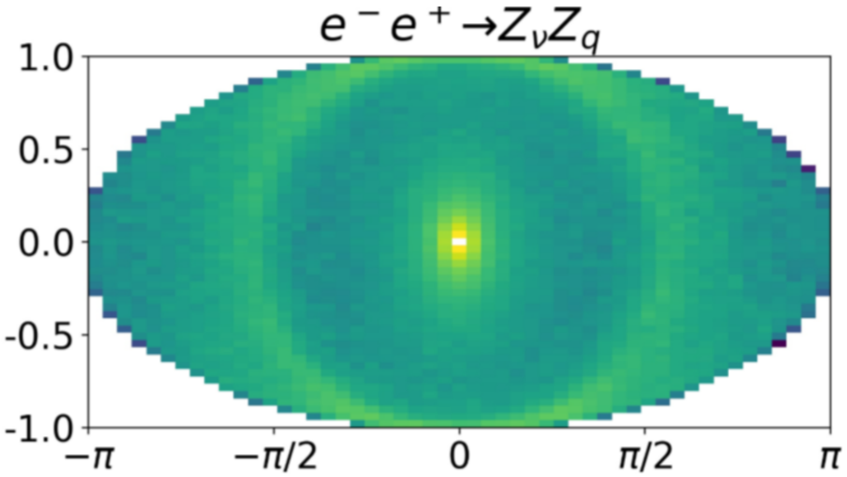


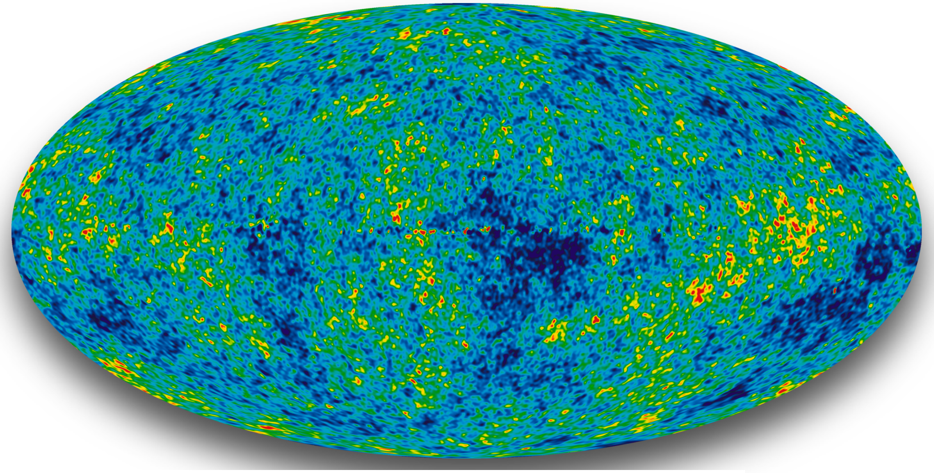
Preprocessing

1. Define a Cartesian coordinate system: z-axis being along beam line and x-y plane (equatorial plane) overlapping with its transverse plane
2. Rotate the motion direction of the most energetic particle to be along the x-axis
3. Project the particles to “detector sphere”

Halo size and structure: minimal included angle of quarks, info missing at jet level.

"Dictionary" between Event Projection and CMB





Mollweide projection at e^-e^+ colliders	All-sky CMB map
Projection sphere	Celestial sphere
Equatorial plane	Galactic plane
Energy (p_T, timing, charge, d_0, etc.) projection	Temperature (polarization) map
Event-level kinematics	Anisotropy
Fox-Wolfram moments	Power spectrum (TT, TB, BB, etc.)
Multi-spectra	Bispectrum, trispectrum, etc.
… …	… …

Another Way: Event-Level ML

Pursue analysis directly at event level

- Lepton collider: negligible pileups, colorless beam and fixed energy.
- Project event to images, using CNN to extract features.

Comparative studies to compare the two approaches using ML tools

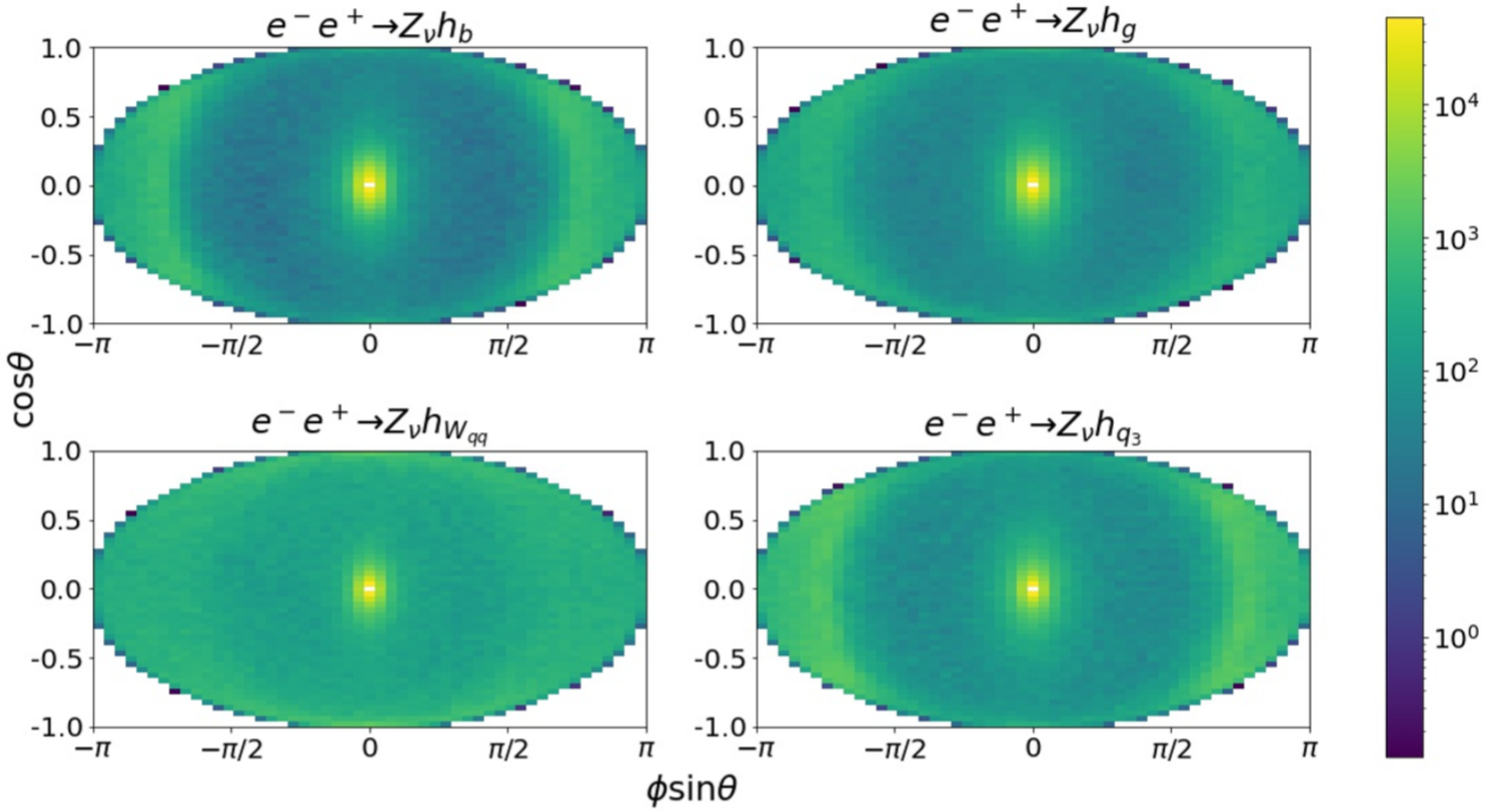
- Jet Level: Fully Connected Network (FCN)

  Input: jet 4-momenta (and FW moments $l<50$ / track info)
- Event Level: Convolutional Neural Network (CNN), Based on ResNet-50

  Input: 50×50 pixelized event-level image (and track info)

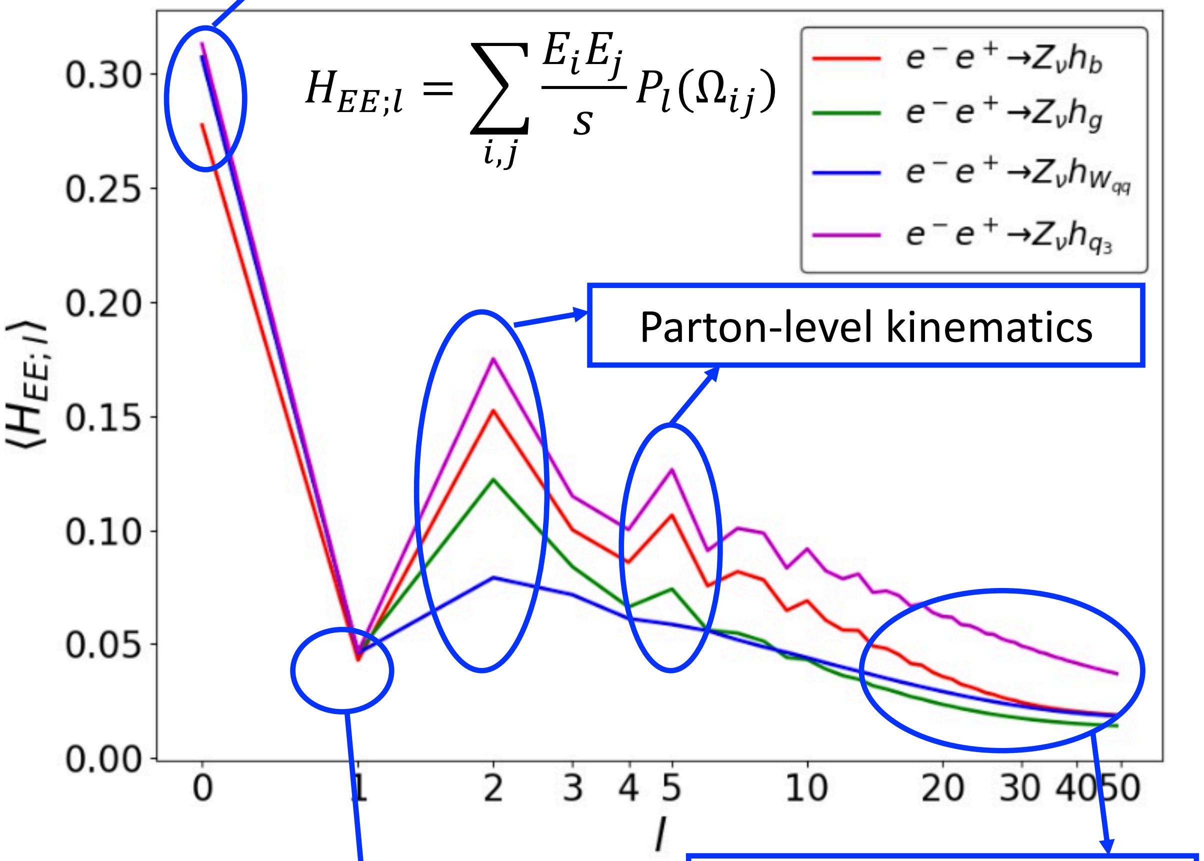
Benchmark

$$e^-e^+ \rightarrow Zh \rightarrow \nu\nu + (bb, jj, gg, W_q W_q^*)$$

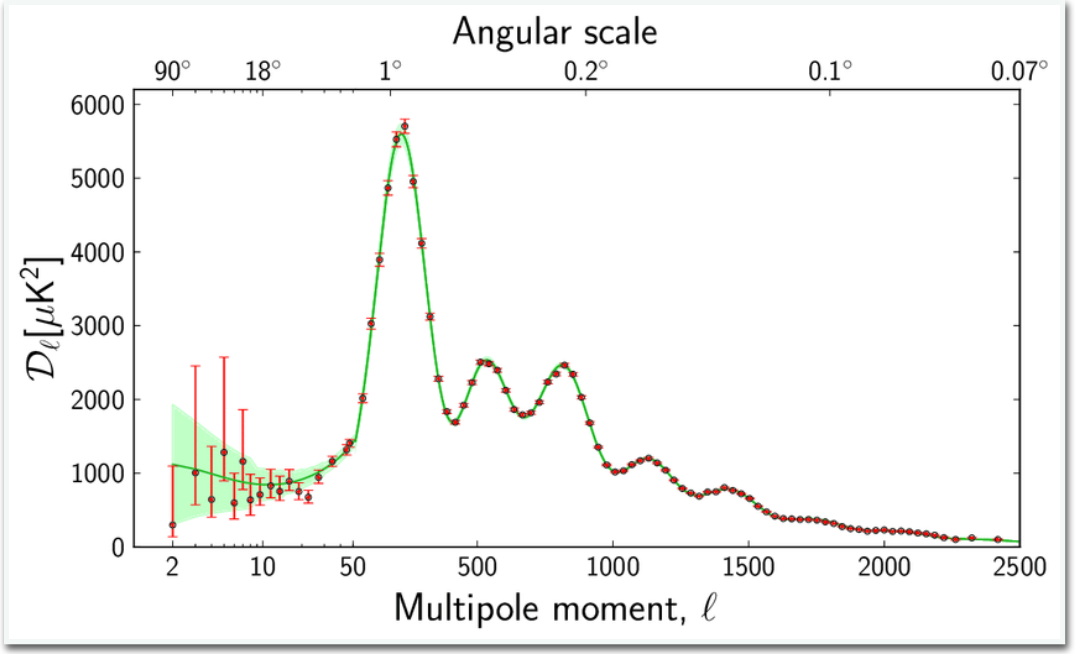


FW Moments of Energy Distribution

$$H_{EE;0} = \frac{(\sum_i E_i)^2}{s}$$



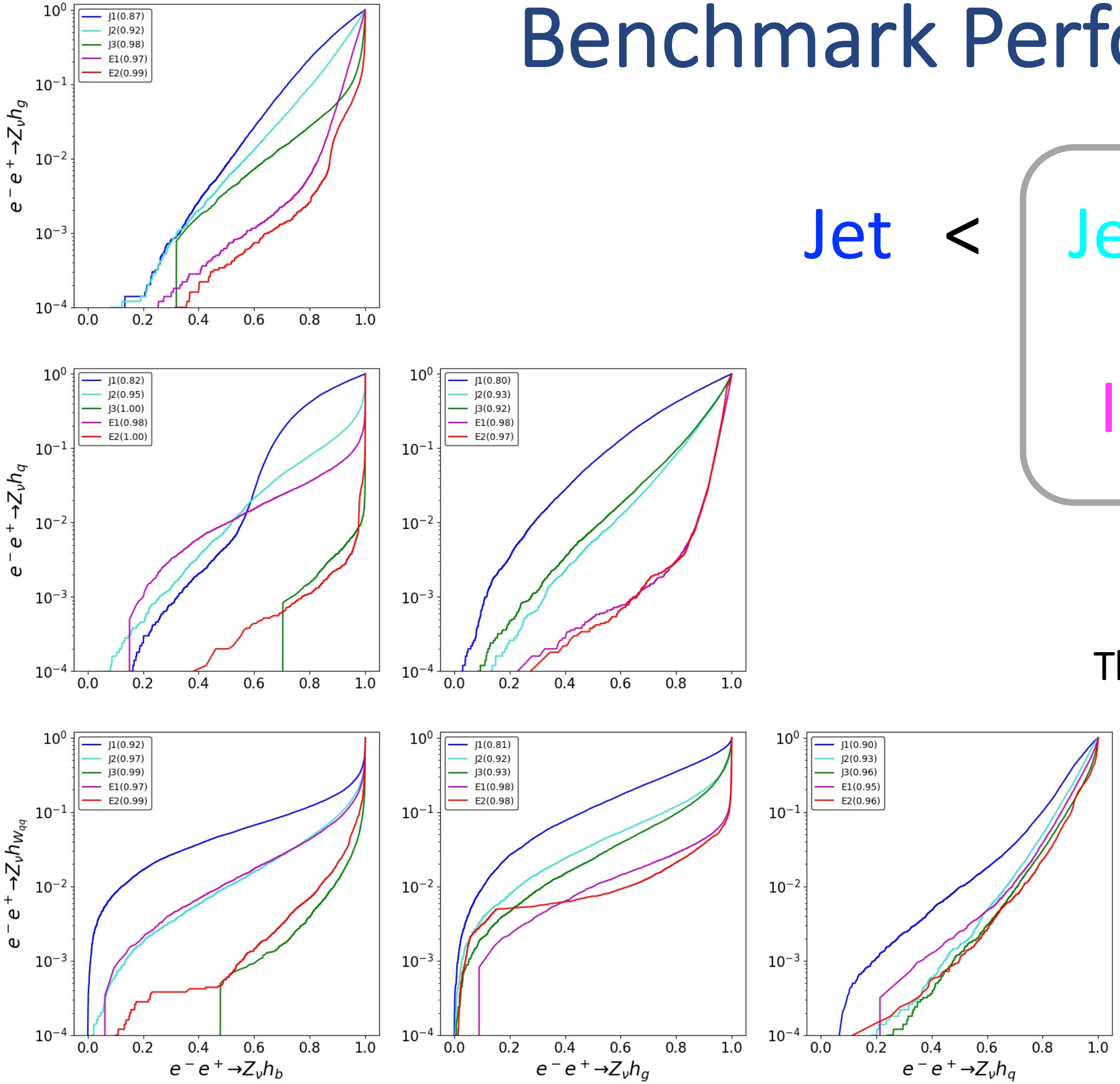
$$H_{EE;1} = \frac{|\sum_i \vec{p}_i|^2}{s}$$



Analogue to CMB power spectrum

- Difference: suppressed sample ("cosmic") variance, due to large size of data sample
- Similarity: physics at characteristic scales may result in "acoustic peaks"

Benchmark Performance



Jet < Jet + FW ~ Image < Jet + FW + track ~ Image + track

The performance gap between Jet+FW and Image may be explained by higher order correlation terms.

Application: Measurement of Γ(h)

- Mainstream Method

Bottleneck

$$\Gamma_h = \frac{\Gamma(h \to WW^*)}{\mathrm{BR}(h \to WW^*)} \propto \frac{\sigma(\nu\nu h)}{\mathrm{BR}(h \to WW^*)} = \frac{[\sigma(\nu\nu h_b)][\sigma(Zh)]^2}{[\sigma(Zh_b)][\sigma(Zh_W)]}$$

- Pros: relatively big signal rates for ee -> Zh -> ZWW* at low energy runs
- Cons: relatively small signal rate for vvh, h->bb for low energy beams
- Probably the most important method @ 240 GeV and 365 GeV

Mainstream Method

Γ_h (%)	$\mathrm{CEPC}_{240(250)}$ [17, 68]	FCC_{240} [18]	$\mathrm{FCC}_{240+365}$ [18]	CLIC_{350} [72]	ILC_{250} [71, 73, 74]
Method A	5.1 (5.0)	4.5*	4.2*	-	20*
Method B	3.5 (3.2)	3.5*	1.7*	6.7	13
Method C	-	-	3.4*	-	-
Combined	2.8 (2.7)	2.7	1.3	6.7	11

Application: Measurement of Γ(h)

- Using inclusive Higgs decay in VBF
  - Include h->bb/cc/gg/ττ decays
  - Pros: fully make use the Higgs hadronic decay modes at event level

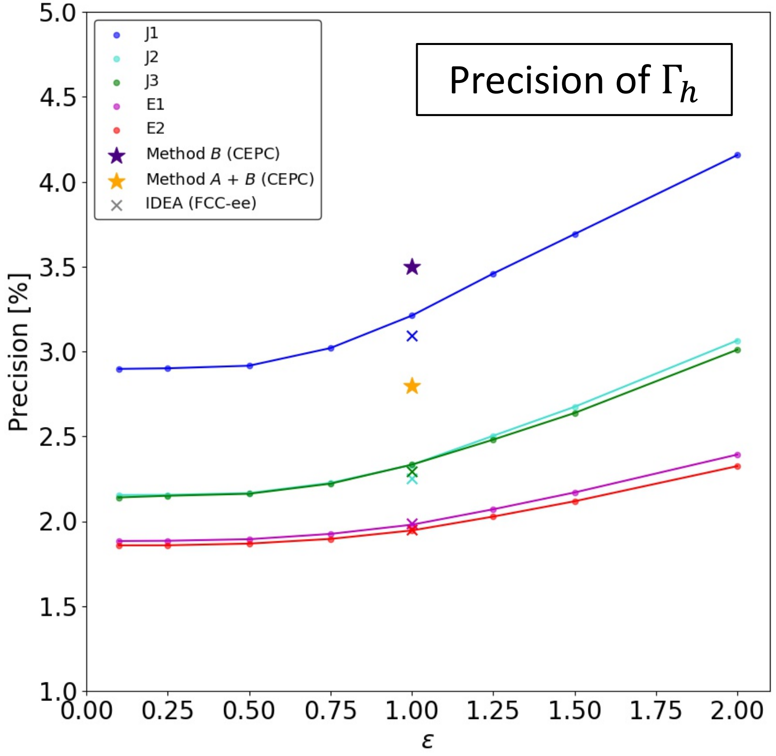
$$\Gamma_h = \frac{\Gamma(h \to WW^*)}{\mathrm{BR}(h \to WW^*)} \propto \frac{\sigma(\nu\nu h)}{\mathrm{BR}(h \to WW^*)} = \frac{[\sigma(\nu\nu h_h)][\sigma(Zh)]^2}{[\sigma(Zh_h)][\sigma(Zh_W)]}$$

Can be well-measured with sub percent precision; let's focus on the other two

h_h denotes the inclusive two-body Higgs decays $h \to bb$, cc, gg and $\tau\tau$.

Application: Results @240GeV, 5ab-1

	Jet	Jet+FW	Jet+FW+track	Image	Image+track
Precision (%)	J1	J2	J3	E1	E2
$\sigma(Z_\nu h_{W_{lq}})$	1.7 (1.6)	1.4 (1.6)	1.5 (1.6)	1.5 (1.4)	1.5 (1.4)
$\sigma(Z_\nu h_{W_{qq}})$	1.6 (1.6)	1.2 (1.2)	1.1 (1.1)	1.1 (1.1)	1.1 (1.1)
$\sigma(\nu\nu h_h)$	2.8 (2.7)	1.8 (1.7)	1.9 (1.8)	1.4 (1.4)	1.3 (1.3)
Γ_h	$3.2^{+0.9}_{-0.3}$ (3.1)	$2.3^{+0.7}_{-0.2}$ (2.2)	$2.3^{+0.7}_{-0.2}$ (2.3)	$1.9^{+0.5}_{-0.1}$ (1.9)	$1.9^{+0.4}_{-0.1}$ (1.9)



- W decay channels
  FW-moments capture event-level info
- $\nu\nu h_h$ channel
  FW-moments compensate partially, ~1% precision off from jet-level
- Γ_h
  achieve 2.3% with J3 classifier, 1.9% with E2 classifier

The precision achieved is robust against the rescaling of the detector energy/momentum resolutions and different detector templetes.

Outlook I

Can the Higgs decay width be measured at sub percent level @ 240+365 GeV or even @ 240 GeV, given the currently proposed detector baseline?

- Apply event-level ML to multiple channels
- Extra information: charge, PID, displacement, etc
- Advantaged ML techniques
- … …

We expect event-level analysis with ML to be broadly applied to other hadronic-event measurements at future e-e+ colliders. To what extent one can benefit from it?

- Higgs coupling to quarks/gluons
- CP property of Higgs boson
- Flavor physics
- … …

Thank You!