Jet or Event? Physics at Future Lepton Colliders

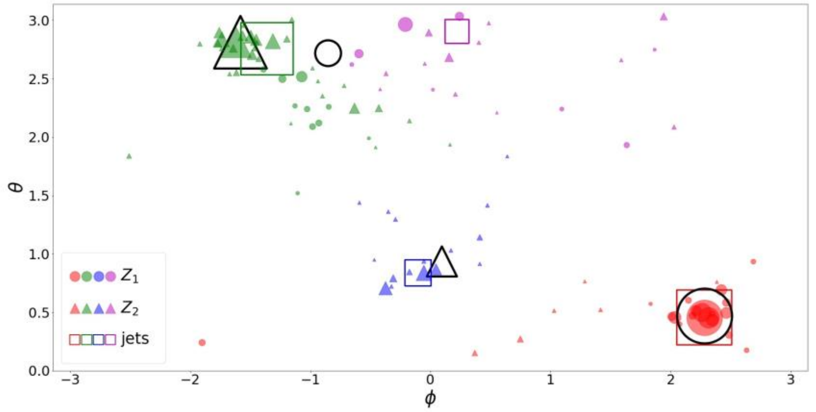
Lingfeng Li (Hong Kong U. of Sci. & Tech.) with Tao Liu, Ying-Ying Li, Si-Jun Xu

LHCP 2020

Introduction

The precision frontier of next decades in Higgs and electroweak (EW) physics is expected to be defined by next-generation ee colliders. However, amongst EW/Higgs processes, the hadronic modes containing (anti-)quarks or/and gluons are dominant. Because of this, the measurement precisions will be limited by our definition of jets:

- Deformation from the truth in due to misclustering of particles in the overlapping region of fragmentation and hadronization inside detector
- Loss of information stem from dimension reduction

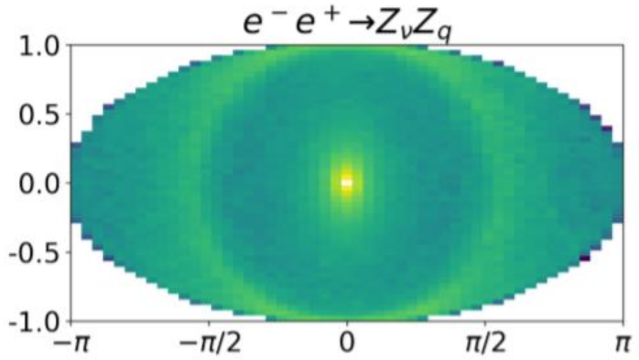


Example event with ee−k_t algorithm. Info distortion makes jet momenta (coloured boxes) deviate from that of quarks (black circles/ triangles).

Two ways to compensate the loss after jet clustering:

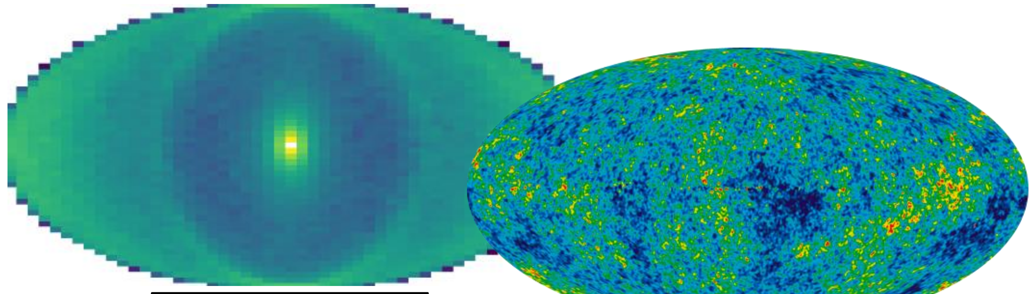
- Pursue jet-level analysis by properly incorporating subjet-scale or event-level observables
  - Simple framework. Physically intuitive
  - Less systematic
- Address the information deformation/loss in jet clustering is to pursue the analysis in a brute-force way, using the event-level data as input.
  - Use the kinematic info at event level to the greatest extent
  - Large complexity, Data structure: need machine learning(ML)

CMB-like Observable Scheme



Cumulative projections of 10^4 events: z-axis being along beam and the most energetic particle rotated towards the center. A"halo" is formed due to the minimal included angle between the two ancestral quarks.

These observations are reminiscent of the all-sky CMB map where the message on the early Universe is encoded as its power spectrum and multi-spectra. Quite generally, we can build up a dictionary between the Mollweide projection of each ee collision event and the all-sky CMB map:

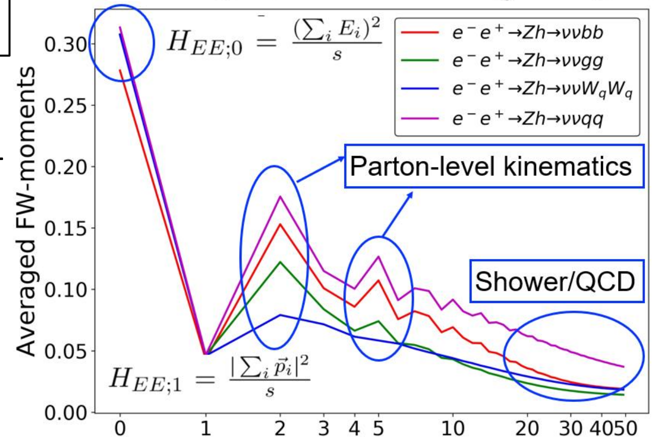


ee→WW->4q

Jet- vs. Event-Level Information

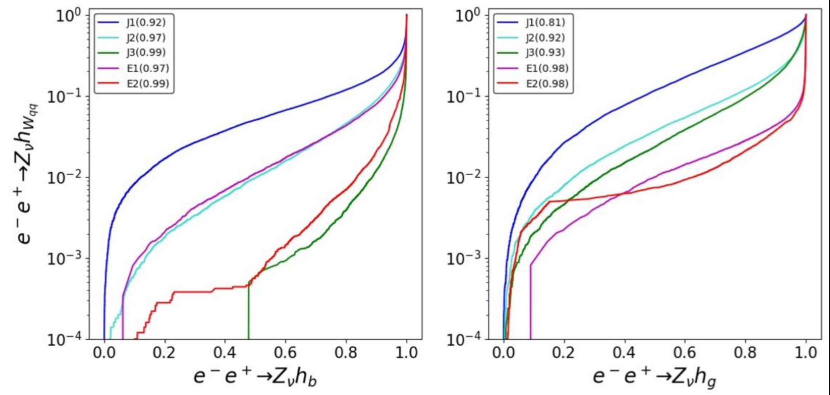
Now the Fox-Wolfram moments[2] as the leading order event-level input, analogous to the power spectra in CMB studies. There definition is

$$H_{AB;l} = \sum_{m=-l}^{l} H_{AB;l,m} = \frac{4\pi}{2l+1} \sum_{i,j} \frac{A_i B_j}{s} \sum_{m=-l}^{l} \left(Y_l^m(\Omega_i)^* Y_l^m(\Omega_j)\right) = \sum_{i,j} \frac{A_i B_j}{s} P_l(\cos\Omega_{ij}) \,.$$



Similar to the latter, they are characterized by a series of "acoustic peaks and valleys" containing rich physical information. Unlike the CMB power spectrum, the FW moments are free from big sample variance, because of the large size of collider data.

Using ML to better understand/evaluate jet- vs. event-level input:



Example ROC and AUC for our benchmark study, comparing different ee->Zh final states. In general, the AUC of J1 < J2 < J3 and E1 < E2 holds.

For jet-level ML, we use Fully Connected Network (FCN) and for event level we use Convolutional Neural Network (CNN) to take the whole event as image:

- **J1** :Jet momenta only.
- J2: Jet momenta, with FW moments with l ≤ 50
- J3: Jet momenta, with FW moments and track info
- E1: Event Image(50×50), without track information.
- E2: Event Image(50×50), with track information

It is noticeable that the AUC gap between **J1** and **E1** classifiers is not fully addressed by FW moments in most cases->multi-spectra?

Precisions of Higgs Width Measurement

We apply ML classifiers to measuring the SM Γ_h, one of the most important tasks at future ee colliders, with the data of 5ab^{-1}@240GeV using:

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

Using all 5 ML models, expected precisions of measuring Γ_h:

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)
$\sigma(\nu\nu h_b)$	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)

The best outcome of 1.9% improves the baseline precisions, i.e., ~3.5% at both CEPC_{240}[3] and FCC_{240}[4], by a factor about 1.8.

Conclusion

The event-level classifiers perform better compared to the jet-level ones; but, incorporating the FW moments into the jet-level classifiers can significantly reduce the performance gap between them. As an application of such classifiers, we analyzed the precision of measuring the SM Γ_h at ee colliders with the data of 5ab^{-1}@240GeV. The precisions obtained are significantly better than the baseline ones.

References

[1] Li, L., Li, Y.-Y., Liu, T., and Xu, S.-J., 2004.15013

[2] G. C. Fox and S. Wolfram, Phys. Rev. Lett., vol. 41, p. 1581, 1978

[3] F. An et al., Chin. Phys., vol. C43, no. 4, p. 043002, 2019, 1810.09037

[4] A. Abada et al., Eur. Phys. J., vol. C79, no. 6, p. 474, 2019