JET OR EVENT?
- PHYSICS AT FUTURE LEPTON COLLIDERS

Sijun Xu

The Hong Kong University of Science and Technology

ML4Jets2020@NYU

Based on arXiv:2002.xxxxx
in collaboration with Lingfeng Li, Ying-Ying Li and Tao Liu

Precision Frontier of Next Decades

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

Measurements	CEPC$_{250}$ [61]	FCC$_{240}$ [62]	FCC$_{365}$ [62]	CILC$_{350}$ [63]	ILC$_{250}$ [60, 64, 65]
$\sigma(Zh)$	0.5%	0.5%	0.9%	1.6%	2.6%
$\sigma(Zh)\mathrm{BR}(h \to bb)$	0.3%	0.3%	0.5%	0.86%	1.2%
$\sigma(Zh)\mathrm{BR}(h \to cc)$	3.1%	2.2%	3.5%	14%	8.3%
$\sigma(Zh)\mathrm{BR}(h \to gg)$	1.2%	1.9%	6.5%	6.1%	7.0%
$\sigma(Zh)\mathrm{BR}(h \to WW^*)$	0.9%	1.2%	2.6%	5.1%	6.4%
$\sigma(Zh)\mathrm{BR}(h \to ZZ^*)$	4.9%	4.4%	12%	-	19%
$\sigma(h\nu\nu)\mathrm{BR}(h \to bb)$	2.9%	3.1%	0.9%	1.9%	10.5%
$\sigma(h\nu\nu)\mathrm{BR}(h \to cc)$	-	-	10%	26%	-
$\sigma(h\nu\nu)\mathrm{BR}(h \to WW^*)$	-	-	3.0%	-	-
$\sigma(h\nu\nu)\mathrm{BR}(h \to ZZ^*)$	-	-	10%	-	-

[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 measurements of hadronic events and analysed 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

The precision is generically limited by jet-level analysis if jet is not perfect.

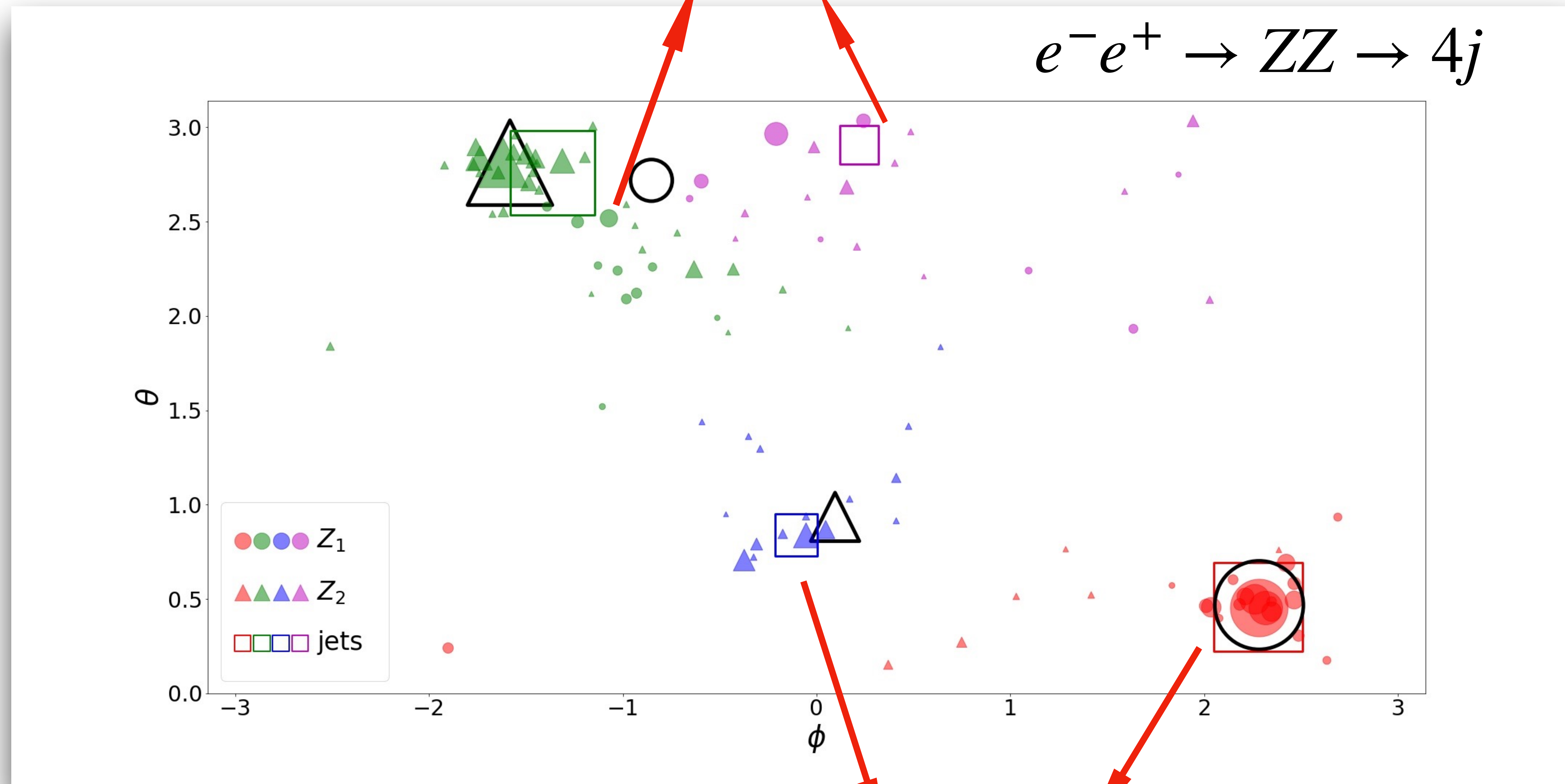
We should look at performance of jet algorithms.

Limitation of Jet-Level Analysis

- Jet confusion: deformation from the truth in kinematics due to mis-clustering of particles in the overlapping region of fragmentation and hadronization inside detector
- Loss of information on jet sub-structure and superstructure
  - dimension reduction of jet clustering
  - particle missing: unclustered

Event Display (ee-kt)

Jet confusion



Information loss due to dimension reduction

Can we recover from these limitations?

First Way: Jet + Extra Info

- Jet substructure observables: extensively applied in boost kinematics
- Event shape: relatively intuition-based, e.g. thrust [E. Farhi, 1977]
- Fox-Wolfram moments [G. C. Fox and S. Wolfram, 1978]and their extensions: more systematic, but relatively less intuitive.

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

- Pros: Simple framework. Physically intuitive.
- Cons: Less systematic (except F-W moments) .

Another Way: Event-Level ML

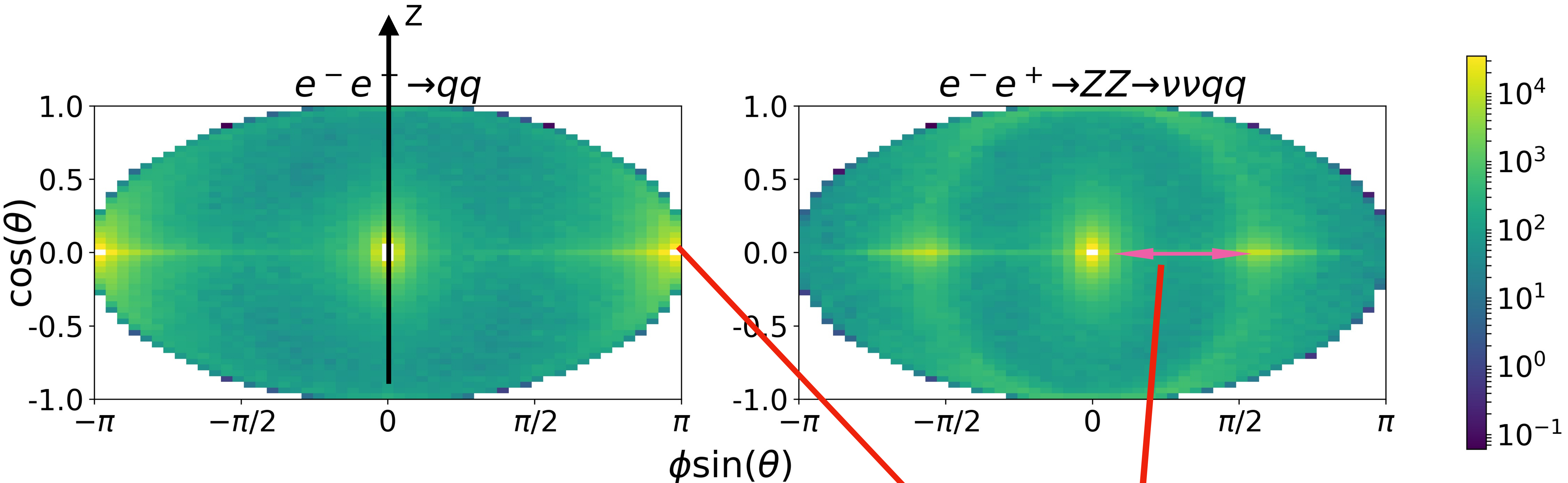
Pursue analysis directly at event level

- Pro: Most information.
    - Lepton Collider: negligible pileups, colorless beam and approximate spherical symmetry
- Con: Large complexity, Data structure.

DL is a solution. Below we will pursue a comparative study between these two methods, using the tool of DL (jet-level: FCNN; event-level: ResNet-50 based CNN)

- Goal 1: provide a more complete evaluation on the capability of e-e+ collider
- Goal 2: explore the possibility of reconstructing the lost information in jet-level analyses

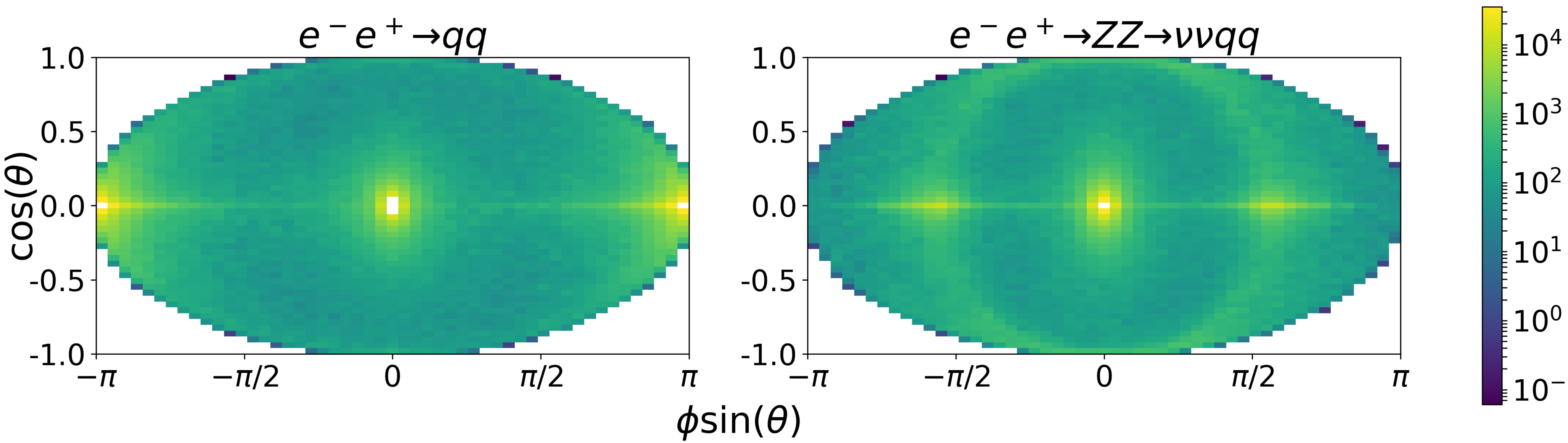
Cumulative Mollweide Projection

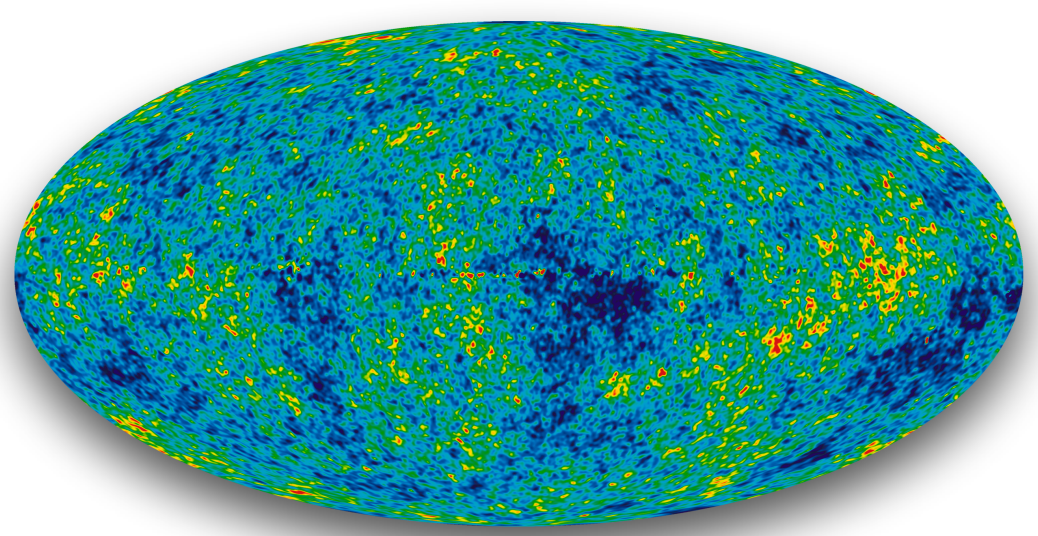


- Define a Cartesian coordinate system: z-axis being along beam line and x − y plane (equatorial plane) overlapping with its transverse plane
- Rotate the motion direction of the most energetic particle to be along x-axis, and the plane expanded together with that of the second-most energetic particle (if they are not in the same straight line) to overlap with the x − y plane.
- Project the particles to ``detector sphere''

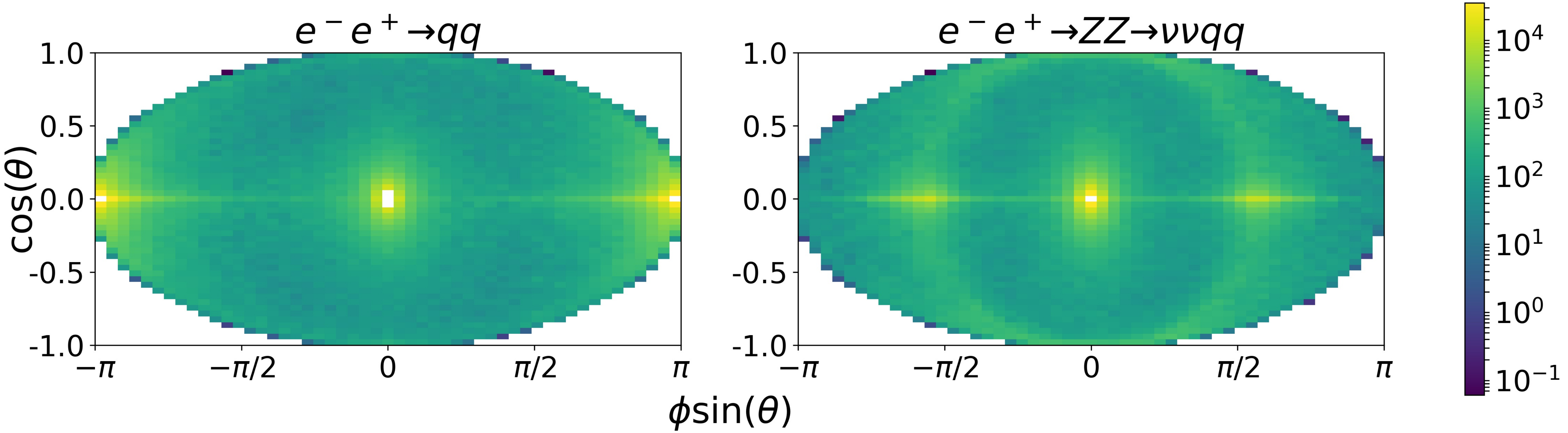
- Back-to-back motion of qq
- Minimal included angle between qq
- Halo structure (carries information missed at jet level)

“Dictionary” between Event Projection and CMB



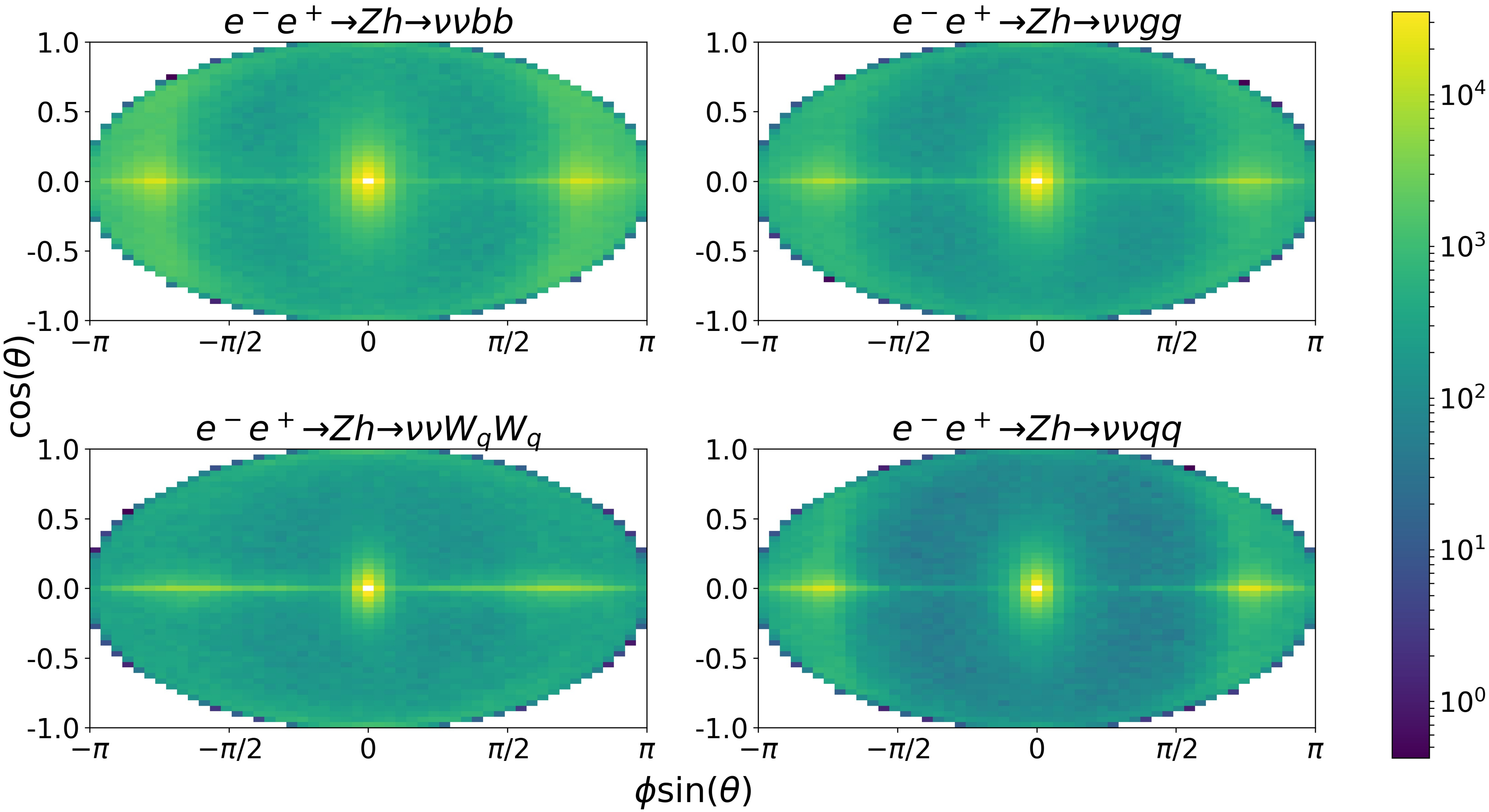


“Dictionary” between Event Projection and CMB



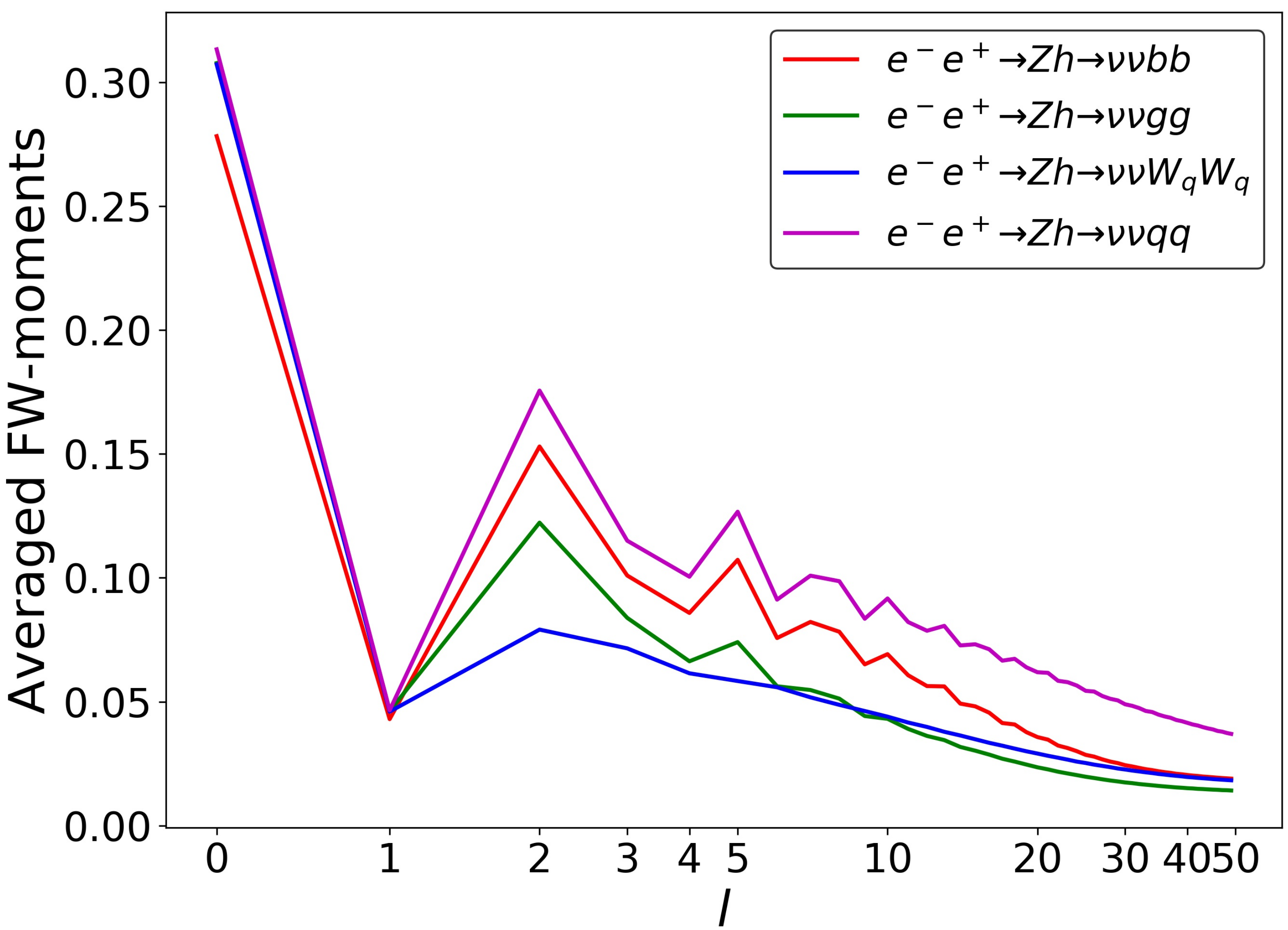
Event distribution at e^-e^+ collider	All-sky CMB map
Detector projection sphere	Celestial sphere
Equatorial plane	Galactic plane
Energy (p_T, charge, etc.) distributions	Temperature (polarization) fields
Event-level kinematics	Anisotropic features
Fox-Wolfram moments	CMB power spectrum (TT, TB, BB, etc.)
Multiple-point correlators	Non-Gaussianity
... ...	... ...

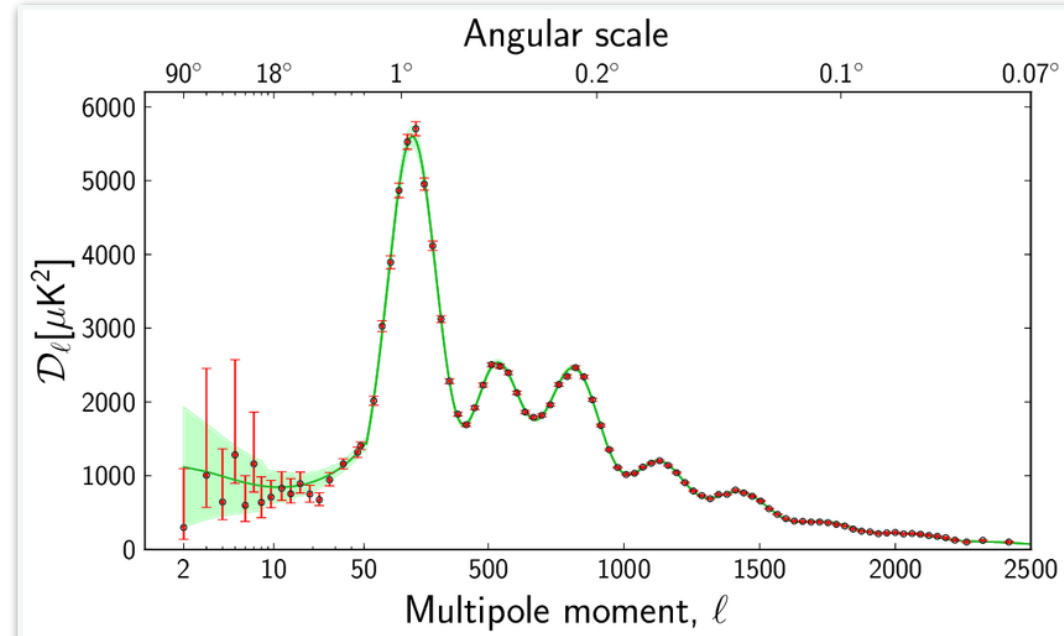
Benchmark



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

F-W Moments of Energy Distribution



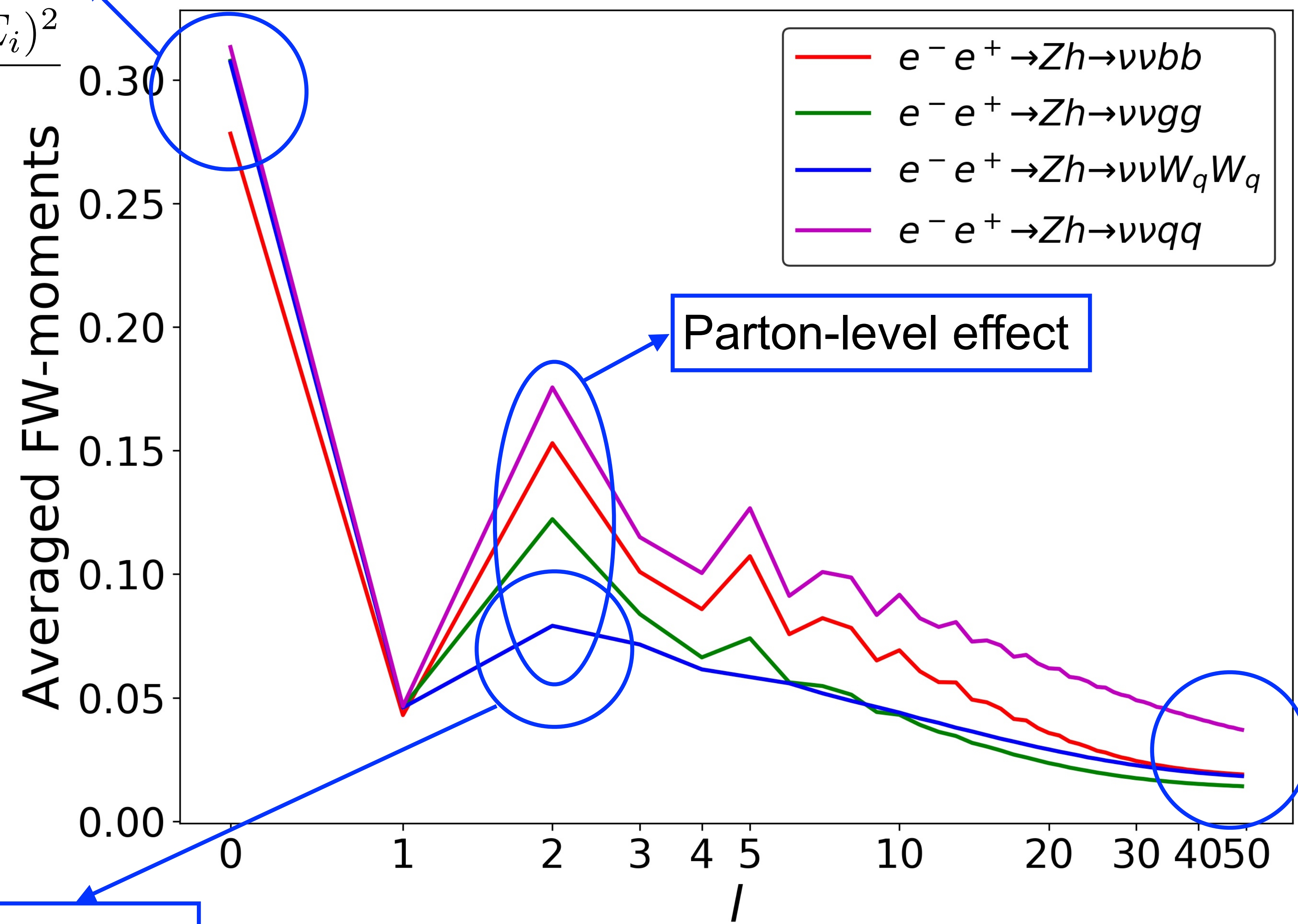


- Analogue to CMB power spectrum
- Difference: suppressed sample (``cosmic'') variance, due to large size of data sample
- Similarity: physics at characteristic sales may result in ``acoustic peaks''

F-W Moments of Energy Distribution

Squared share of visible energy

$$C_0 = \frac{(\sum_i E_i)^2}{s}$$



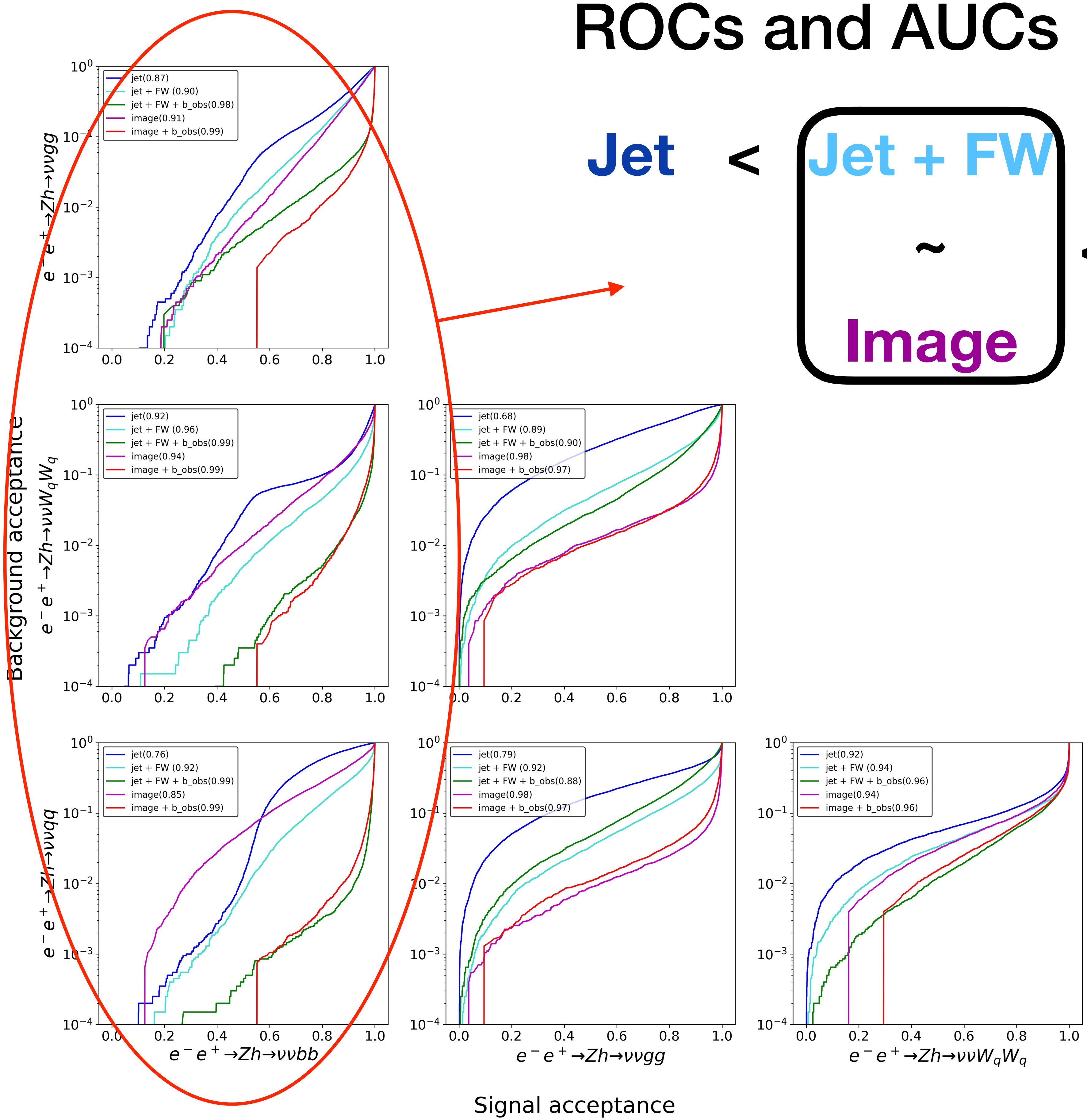
Parton-level effect

Relative broad energy distribution

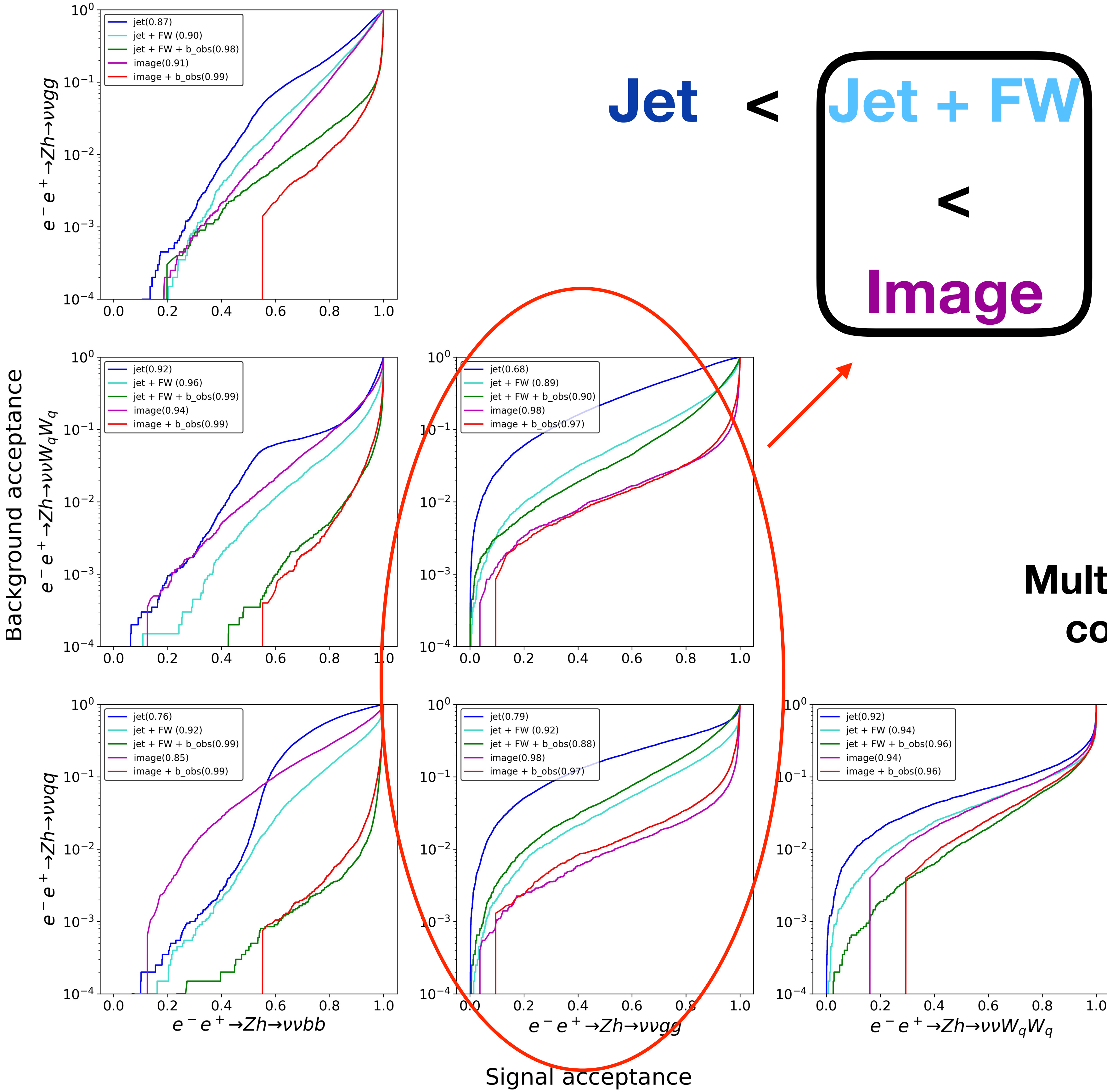
Collinear and IR safe

ROCs and AUCs

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



ROCs and AUCs



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

Multiple-point Correlators may compensate the loss here.

Measurement of Higgs Decay Width

- Mainstream Method
  - Pros: relatively big signal rates for ee -> Zh -> ZWW* for low beam energy
  - Probably the most important method @ 240 GeV and 365 GeV

$$\Gamma_h = \frac{\Gamma(hWW^*)}{\mathrm{BR}(hWW^*)} \propto \frac{\sigma(\nu\bar{\nu}h)}{\mathrm{BR}(hWW^*)} = \frac{[\sigma(\nu\bar{\nu}h)\mathrm{BR}(hbb)]}{\mathrm{BR}(hbb)\mathrm{BR}(hWW^*)}$$
$$= \frac{[\sigma(\nu\bar{\nu}h)\mathrm{BR}(hbb)][\sigma(Zh)]^2}{[\sigma(Zh)\mathrm{BR}(hbb)][\sigma(Zh)\mathrm{BR}(hWW^*)]} \, .$$

Mainstream Method

$\Delta\Gamma_h/\Gamma_h$	CEPC_{240} [1]	FCC_{240} [70]	$\mathrm{FCC}_{240+365}$ [70]	CILC_{350} [71]	ILC_{250} [68, 72, 73]
Method A	5.1%	4.5*%	4.2*%	-	9.0*%
Method B	3.5%	3.4*%	1.6*%	6.7%	6.0*%
Method C	-	-	3.3*%	-	-
Combined	2.8%	2.7%	1.3%	6.7%	5.4%

Our Method

- Using inclusive Higgs decay in VBF.
  - Pros: fully make use of the Higgs hadronic decay modes, at event level

$$\Gamma_h = \frac{\Gamma(hWW^*)}{\mathrm{BR}(hWW^*)} \propto \frac{\sigma(\nu\bar{\nu}h)}{\mathrm{BR}(hWW^*)} = \frac{[\sigma(\nu\bar{\nu}h)\mathrm{BR}(h \to \mathrm{hadrons})]}{\mathrm{BR}(h \to \mathrm{hadrons})\mathrm{BR}(hWW^*)}$$

$$= \frac{[\sigma(\nu\bar{\nu}h)\mathrm{BR}(h \to \mathrm{hadrons})][\sigma(Zh)]^2}{[\sigma(Zh)\mathrm{BR}(h \to \mathrm{hadrons})][\sigma(Zh)\mathrm{BR}(hWW^*)]} \, .$$

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

Analysis Results

$\Delta\Gamma_h/\Gamma_h$	$CEPC_{240}$ [1]	FCC_{240} [70]	$FCC_{240+365}$ [70]	$CILC_{350}$ [71]	ILC_{250} [68, 72, 73]
Method A	5.1%	4.5*%	4.2*%	-	9.0*%
Method B	3.5%	3.4*%	1.6*%	6.7%	6.0*%
Method C	-	-	3.3*%	-	-
Combined	2.8%	2.7%	1.3%	6.7%	5.4%

Mainstream Method

Measurements	Jet ML	Jet ML + Vertex	Jet ML + FW + Vertex	image ML	image ML + Vertex
$\sigma(Zh)\mathrm{BR}(h \to W_q W_q^*)$	1.53%	1.36%	1.25%	1.30%	1.26%
$\sigma(Zh)\mathrm{BR}(h \to W_l W_q^* / W_l^* W_q)$	2.71%	2.59%	1.56%	1.97%	1.96%
$\sigma(h\nu\nu)\mathrm{BR}(h \to \text{hadronic})$	3.67%	2.33%	1.93%	1.21%	1.10%
Γ_h	3.98(3.86)%	2.73(2.60)%	2.28(2.02)%	1.80(1.65)%	1.71(1.56)%

- Jet-level: Jet ML < Mainstream < Jet ML + Vertex < Jet ML + FW + Vertex
- Image-level: image ML ~ image ML + Vertex

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?

Measurements	Jet ML	Jet ML + Vertex	image ML	image ML + Vertex
$\sigma(Zh)\mathrm{BR}(h \to W_q W_q^*)$	1.53%	1.36%	1.30%	1.26%
$\sigma(Zh)\mathrm{BR}(h \to W_l W_q^* / W_l^* W_q)$	2.71%	2.59%	1.97%	1.96%
$\sigma(h\nu\nu)\mathrm{BR}(h \to \text{hadronic})$	3.67%	2.33%	1.21%	1.10%
Γ_h	3.98(3.86)%	2.73(2.60)%	1.80(1.65)%	1.71(1.56)%

ZH final state		Precision
$Z \to e^+e^-$	$H \to WW^* \to \ell\nu\ell'\nu, \ell\nu q\bar{q}$	2.6%
$Z \to \mu^+\mu^-$	$H \to WW^* \to \ell\nu\ell'\nu, \ell\nu q\bar{q}$	2.4%
$Z \to \nu\bar{\nu}$	$H \to WW^* \to \ell\nu q\bar{q}, q\bar{q}q\bar{q}$	1.5%
$Z \to q\bar{q}$	$H \to WW^* \to q\bar{q}q\bar{q}$	1.7%
Combination		0.9%

- The measurements above need to be further improved
- Might be achievable by combining with missed channels

Outlook II

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 couplings to quarks/gluons
- CP properties of Higgs boson
- Flavor physics
- Precision measurements of QCD parameters
- … …

Thank you!