



AFB measurement with $Z \rightarrow q\bar{q}$

Libo Liao

Gang LI, Weimin Song, Jin Zhang, Jiabao Gong, Jiarong Li, Yilun Wang, and Zhaoling Zhang

- Introduction
- Detector, software, and MC samples
- Event selection, background suppression
- Joint flavor and charge identification of a jet
- Statistical analysis
- Conclusion

- The Standard Model is remarkably successful, but NOT a final theory.
 - ◆ Dark matter
 - ◆ Neutrino masses
 - ◆ Matter-antimatter asymmetry
- Two Complementary Search Directions → Strong motivation for Physics Beyond the Standard Model (BSM) :
 - ◆ Direct searches: High energy frontier (e.g. LHC)
 - ◆ Indirect searches: Precision tests of SM predictions
- Future Collider:
 - ◆ CEPC, FCC-ee, ILC...
- CEPC's Advantage:
 - ◆ $> 10^{12}$ Z bosons (orders of magnitude beyond LEP)
 - ◆ Opens new frontier for electroweak precision tests, especially flavor-specific couplings.

Mode	\sqrt{s} (GeV)	L ($10^{34} \text{ cm}^{-2} \text{ s}^{-1}$)	Event yields
H	240	8.3	4.3×10^6
Z	91.5	192 ^(*)	4.1×10^{12}
$W^+ W^-$	155-170	26.7	5.5×10^7
$t\bar{t}$	360	0.80	0.6×10^6

- Key EW Observable: Forward-Backward Asymmetry A_{FB}^f
 - ◆ Highly sensitive to the effective weak mixing angle $\sin^2 \theta_W^{\text{eff}}$

$$A_{FB}^f = \frac{3}{4} \mathcal{A}_e \mathcal{A}_f, \quad \mathcal{A}_f = \frac{2v_f a_f}{v_f^2 + a_f^2}$$

- Core Challenge: Flavor-specific measurements

- ◆ In $Z \rightarrow qq$ decays, cannot tag the quark flavor on an event-by-event basis using conventional methods.
- ◆ Traditional "flavor-blind" analysis dilutes distinct A_{FB}^f values for different flavors, limiting CEPC's physics reach.

$$v_f = I_3^f - 2Q_f \sin^2 \theta_W^{\text{eff}}, \quad a_f = I_3^f$$

Method	Principle	Limitation
Traditional Jet Charge (LEP: OPAL, DELPHI)	Weighted sum of track momenta	Linear, ignores inter-track correlations; poor performance for c, s jets
Maximum Likelihood Fit	Uses full $\cos \theta$ distribution: $d\cos\theta/d\sigma \propto 1 + \cos^2 \theta + \frac{3}{8} A_{FB} \cos\theta$	Superior statistical precision, but still flavor-blind per event

- Method: ParticleNet (Graph Neural Network)
 - ◆ Treats a jet as a particle cloud
 - ◆ Learns optimal representations for joint flavor and charge identification
- Key Advantages:
 - ◆ ParticleNet excels at jet flavor tagging
 - ◆ Flavor-decomposed A_{FB} measurements on an event-by-event basis
 - ◆ Directly overcomes the flavor-blind limitation
- Modern jet charge algorithm + Advanced CEPC detector:
 - ◆ Flavor-decomposed precision measurements: A_{FB}^b , A_{FB}^c , ... with unprecedented statistical power.
 - ◆ Improved $\sin^2\theta_W^{\text{eff}}$ extraction:
 - ◆ Unique BSM probe: Non-universal flavor modifications (e.g., flavor-changing Z' , fourth-generation quarks).
- The power of beam polarization for electroweak precision measurements was convincingly demonstrated by the SLC experiment
 - ◆ Polarization upgrades @ CEPC are under active study

Detector, software, and MC samples

- Detector model: based on the CEPC Reference Detector
- The Monte Carlo (MC) samples:
 - ◆ Generated with whizard 3 @ Z mass
 - ◆ Delphes fast simulation framework with cepec card
 - ◆ Jets are reconstructed using ee-kt algorithm in FastJet
 - ◆ All particles are clustered into exactly two jets.

Event selection, background suppression

Table 1. Relevant branching fractions, cross sections, and expected event yields of the Z boson at the Z-pole energy [12] based on tree-level without initial state radiation.

Decay Mode	Branching Fraction (%)	Cross Section (nb)	Expected Event Yields
$Z \rightarrow \text{hadrons}$	69.911 ± 0.056	39.6	3.96×10^{12}
$Z \rightarrow e^+e^-$	3.363 ± 0.004	1.98	1.98×10^{11}
$Z \rightarrow \mu^+\mu^-$	3.366 ± 0.007	1.98	1.98×10^{11}
$Z \rightarrow \tau^+\tau^-$	3.370 ± 0.008	1.98	1.98×10^{11}

- Other four-fermion backgrounds are negligible
 - ◆ small cross sections and distinct event topologies.

ALEPH

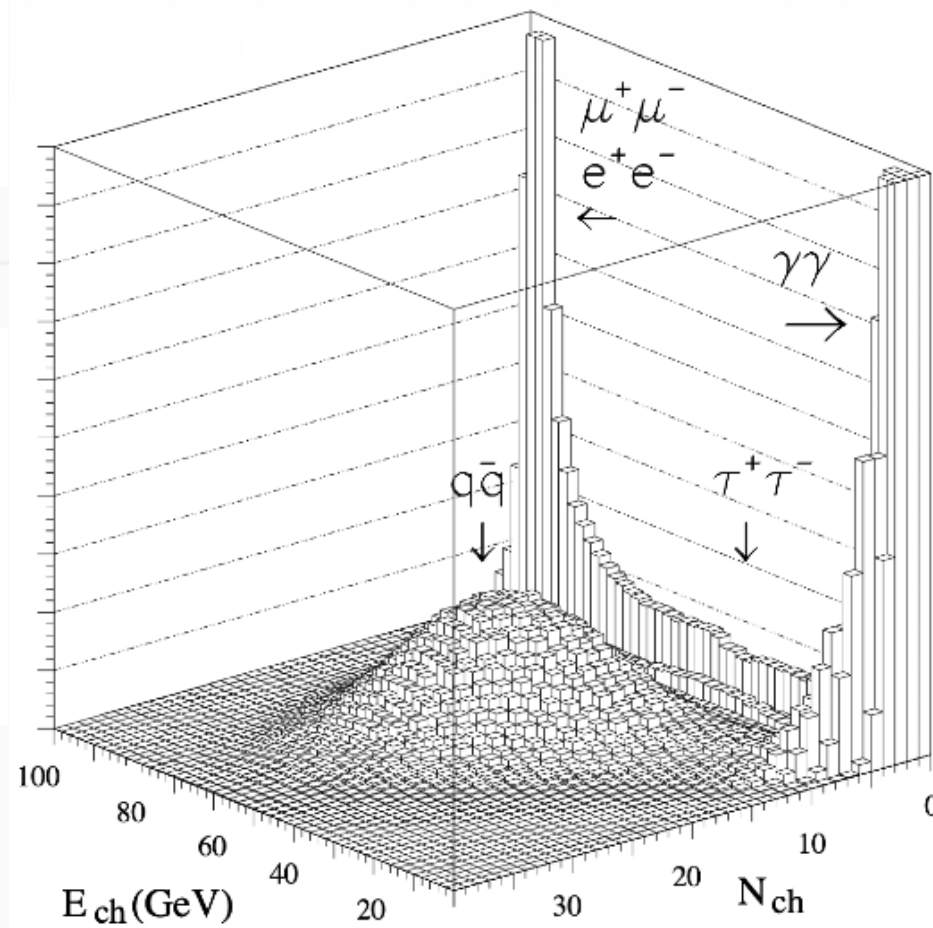


Table 2. Summary of event selection. The columns represent the different quark flavors ($b\bar{b}$, $c\bar{c}$, $s\bar{s}$, $u\bar{u}$, $d\bar{d}$) and $\tau^+\tau^-$ background, and efficiency changes (%).

	$b\bar{b}$	$c\bar{c}$	$s\bar{s}$	$u\bar{u}$	$d\bar{d}$	$\tau^+\tau^-$
Total	100	100	100	100	100	100
$E_{ch} > 20 \text{ GeV}$	99.50	98.95	96.97	98.75	97.94	78.31
$N_{ch} > 8$	98.87	97.15	91.96	94.95	93.99	0.0004

- Effective suppression of leptonic and two-photon backgrounds
- Obtained a high-purity sample of $q\bar{q}$ events
- It motivates the adoption of a deep learning approach for joint flavor and charge identification.

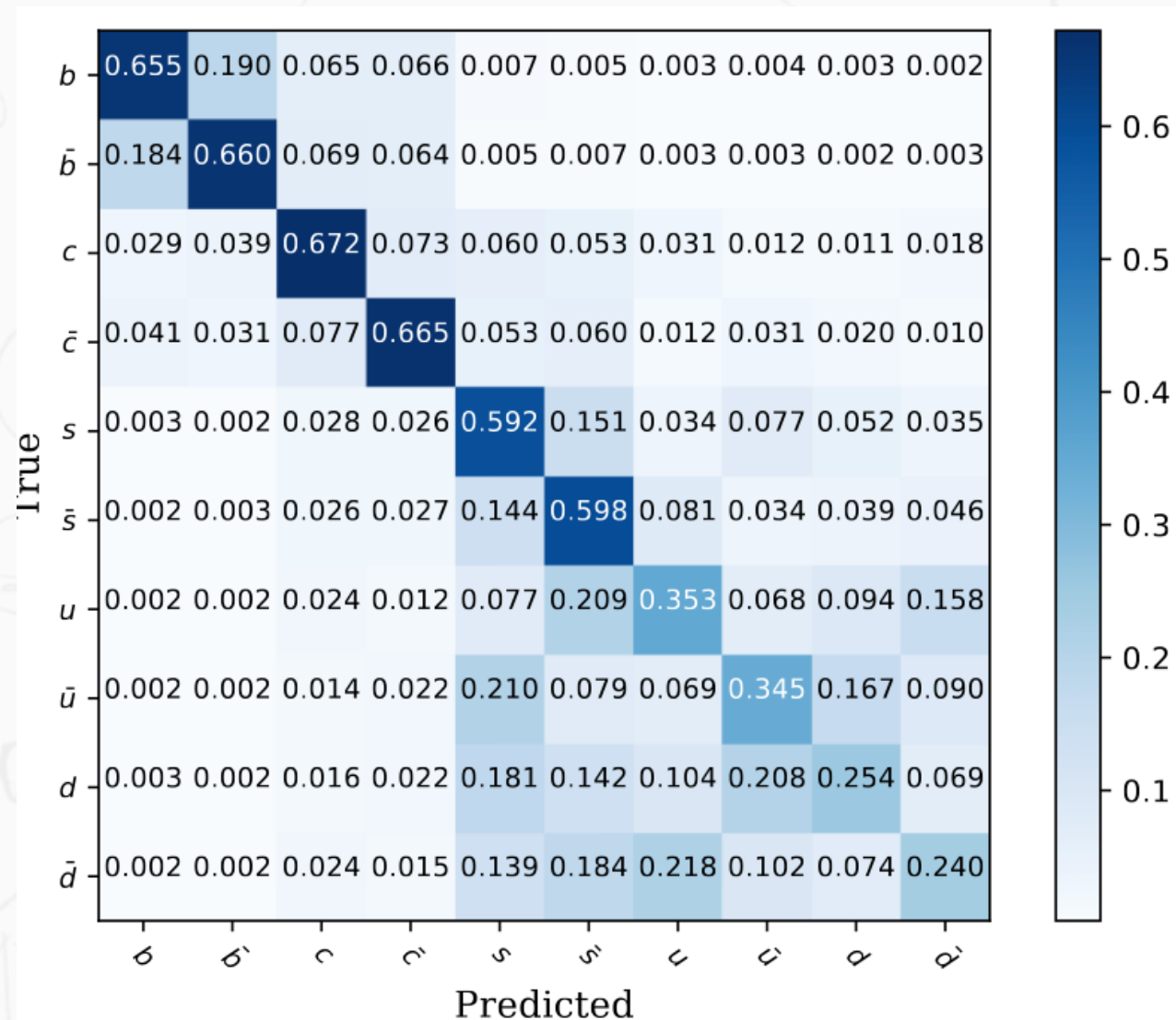
Joint flavor and charge identification

➤ Data sets

- ◆ 10 categories (flavors \times charges)
- ◆ 7,000,000 jets in total
- ◆ train:validation:test = 5:1:1

➤ The model achieves high classification performance

- ◆ The flavor identification accuracy exhibits a clear and expected hierarchy
- ◆ The significant confusion between b-quarks and anti-b-quarks
- ◆ The confusion between c-quark and anti-c-quark is considerably smaller
- ◆ The discrimination between s-quark and anti-s-quark is also challenging

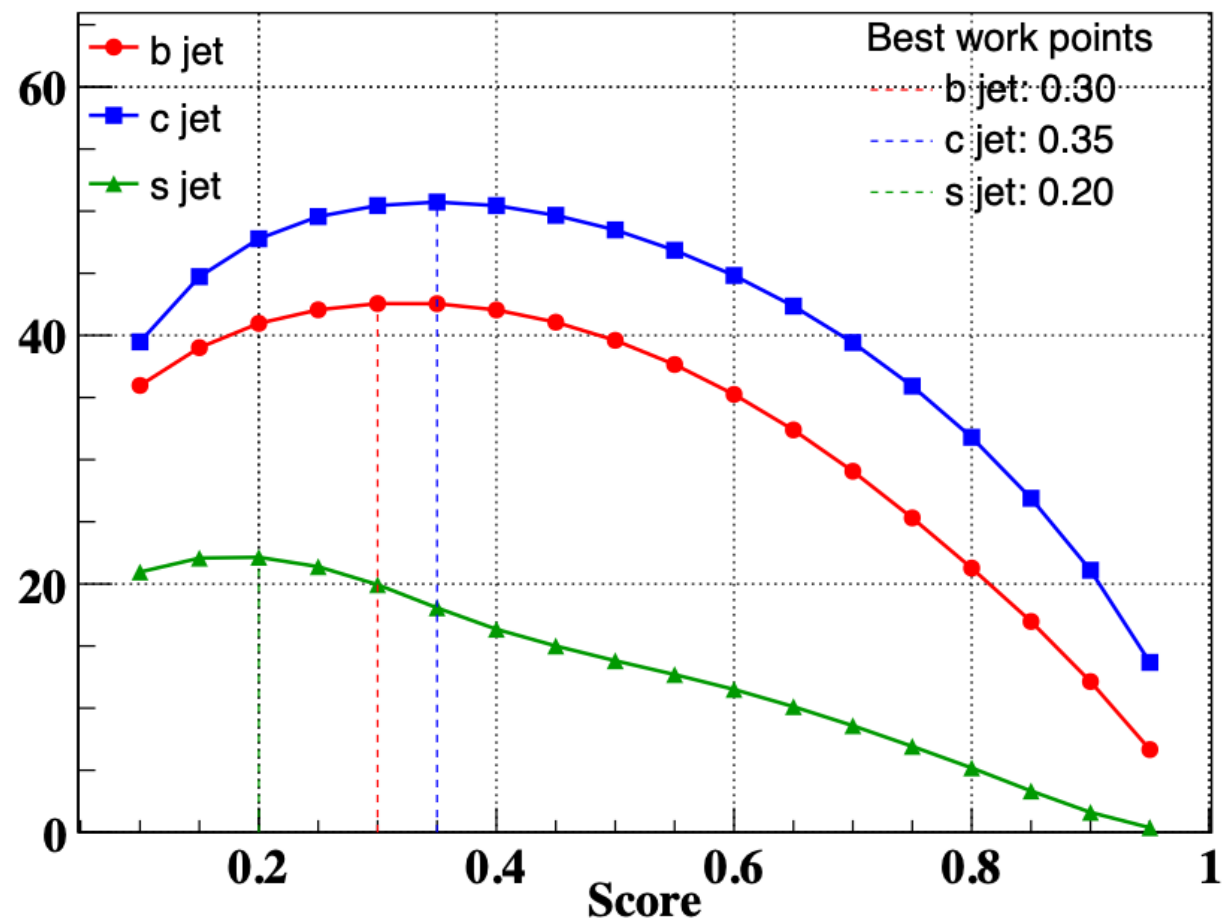


Determination of working points

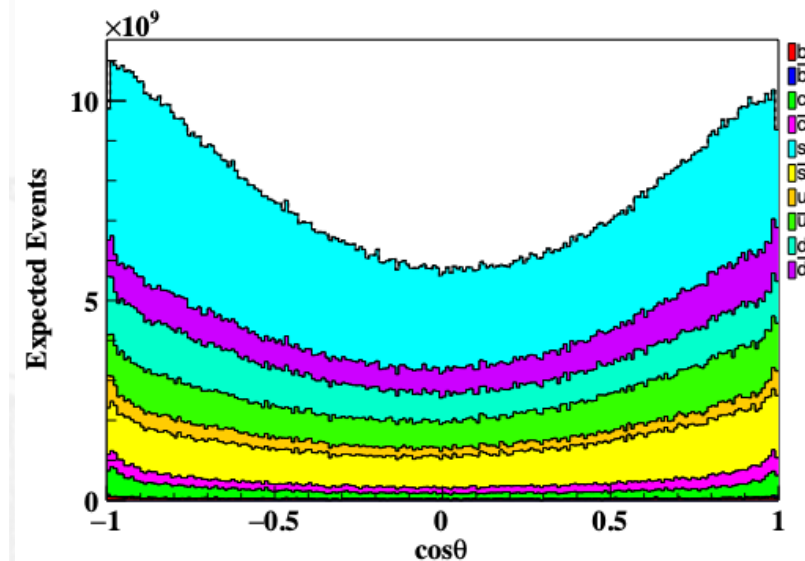
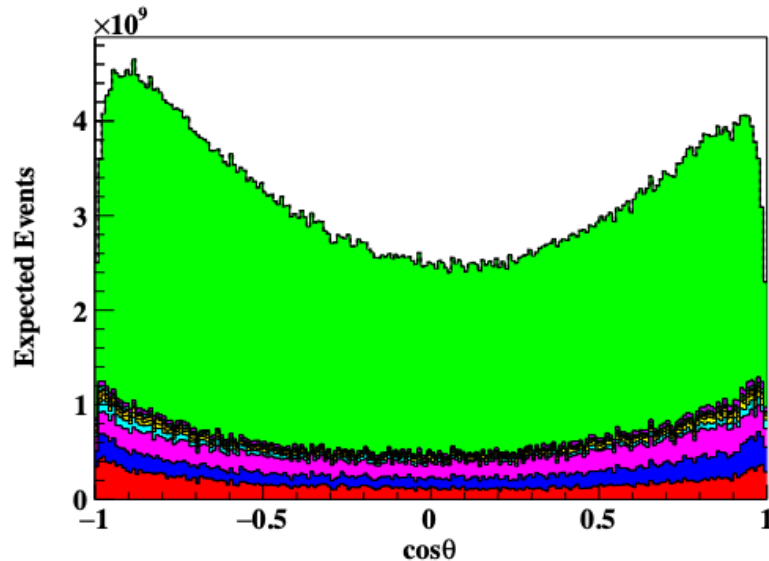
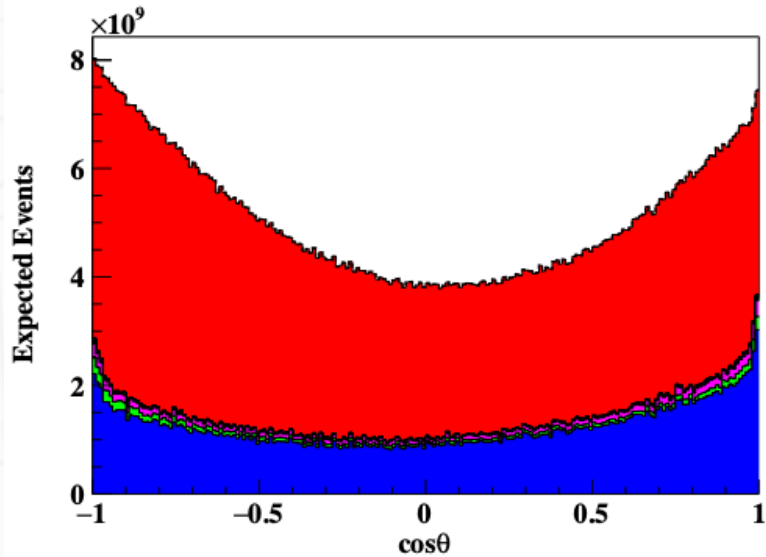
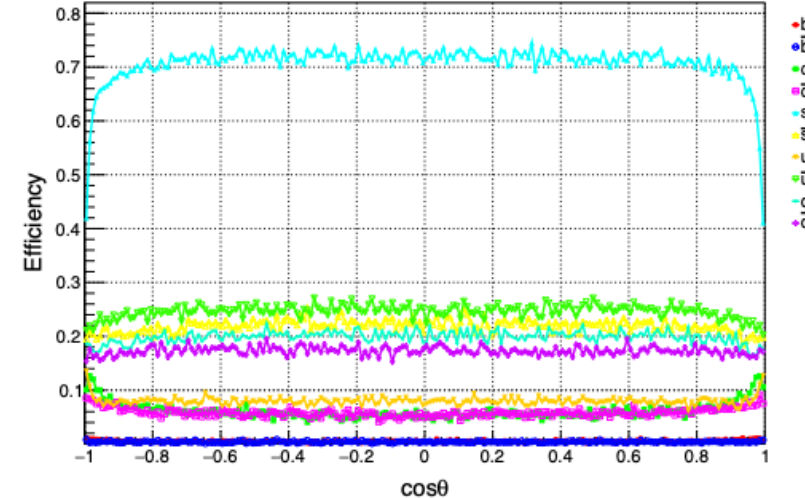
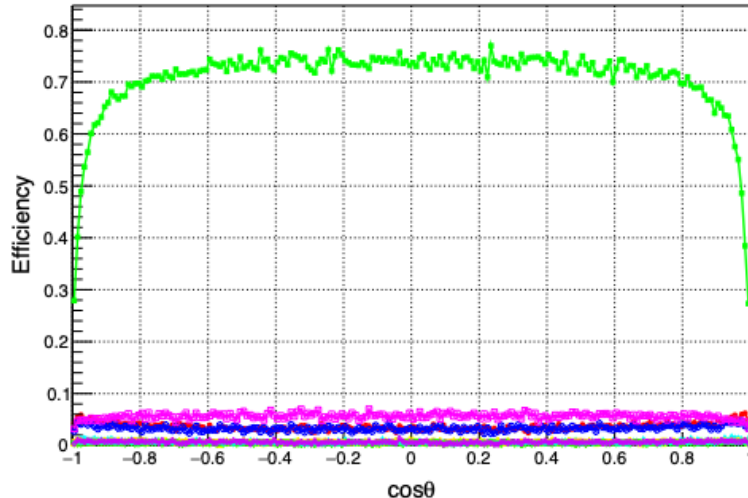
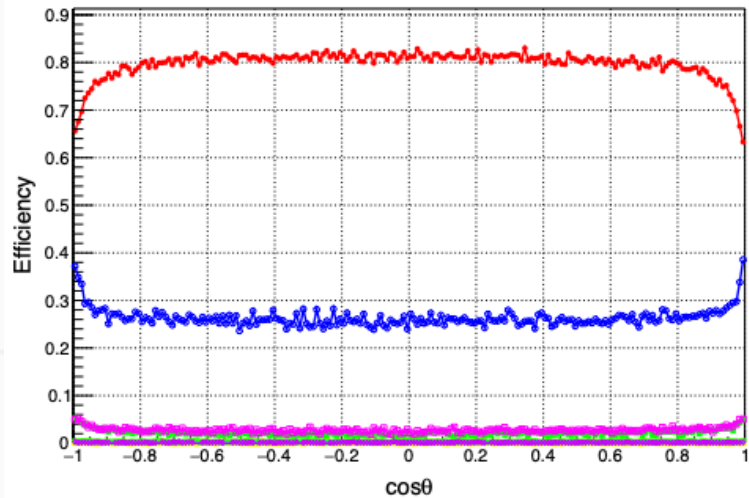
- The working point is chosen to maximize the overall jet charge identification performance, weighted by the production cross sections

$$\begin{aligned} \text{FoM} &= \sum_i \sigma_i \times \epsilon_i \times \rho_i \times \left(\frac{1}{\sigma_i} \frac{\partial \sigma_i}{\partial A_{\text{FB}}} \right)^2 \\ &= \sum_i \epsilon_i \times \rho_i \times \frac{1}{\sigma_i} \left(\frac{8}{3} \cos \theta_i \right)^2, \end{aligned}$$

- The results are consistent with expectation.
 - ◆ The lower threshold for s-jets reflects the greater challenge in distinguishing strange quarks due to the absence of a heavy quark decay vertex.



Efficiency curves



Measurement of A_{FB}

- The expected number of events for reconstructed flavor j (5 flavor) in bin i (200) is:

$$N_{i,j}^{\text{exp}} = \mathcal{L} \sum_k \sigma_{i,j} \epsilon_{i,j,k},$$

$$\sigma_{i,j} = (1 + \cos^2 \theta_i + \frac{8}{3} A_{FB,j} \cos \theta_i) \Delta,$$

- A_{FB} could be extracted via a binned maximum likelihood fit

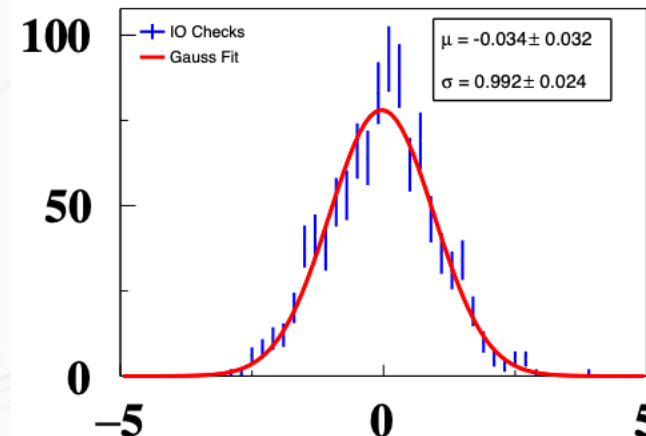
$$\mathcal{L}(A_{FB,j}, N_j^{\text{prod}}) = \prod_i \frac{\mu_{i,j}^{n_{i,j}} e^{-\mu_{i,j}}}{n_{i,j}!},$$

- To estimate the statistical precision, and joint flavor and charge identification

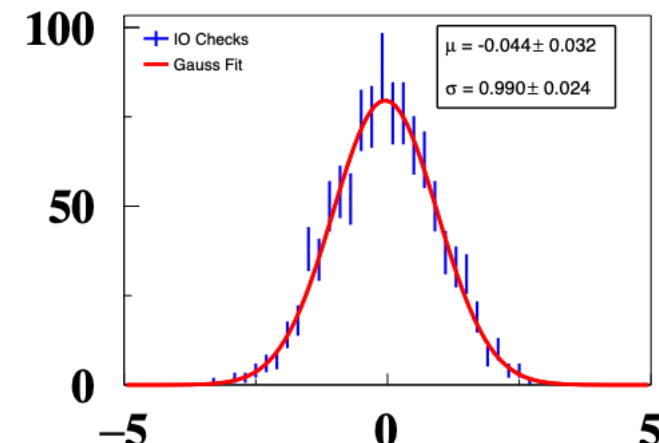
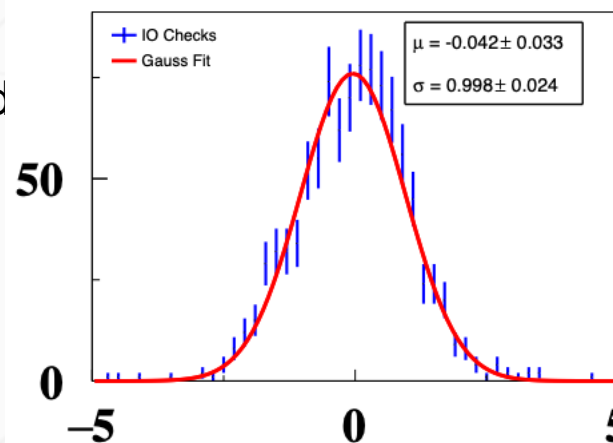
- ◆ Generate 1000 toy Monte Carlo samples.

- The results are listed below:

	This work (10^{-3})	PDG values(10^{-3})
ΔA_{FB}^b	0.013	16
ΔA_{FB}^c	0.021	50
ΔA_{FB}^s	0.018	148



$$\text{Pull} = (O - I) / \text{error}$$



Measurement of $\sin^2 \theta_W^{eff}$

- The published sensitivities of $\sin^2 \theta_{eff}^{lept}$ are at the 10^{-3} level
 - ◆ The sensitivity obtained from $A_{FB}^{0,b}$ is better than that from $A_{FB}^{0,l}$, because the cross section for quark final states is larger than that for leptonic decays.

	$\sin^2 \theta_{eff}^{lept}$	$\Delta \sin^2 \theta_{eff}^{lept} (10^{-3})$
LEP, $A_{FB}^{0,b}$	0.23221 ± 0.00029	1.3
LEP, $A_{FB}^{0,l}$	0.23099 ± 0.00053	2.3
SLC, A_{LR}	0.23098 ± 0.00026	1.2
LEP+SLC	0.23153 ± 0.00016	0.70

- To extract $\sin^2 \theta_W^{eff}$ from the simulated CEPC data, a similar binned extended maximum likelihood is performed
- 1000 toy Monte Carlo samples, pull distributions are also standard
- The results (10^{-6}) are listed
 - ◆ Improvements in both statistical precision and analysis methods

	b	c	s
$\Delta \sin^2 \theta_{eff}^{lept}$	1.0	1.5	1.4

Extension to polarization scenario

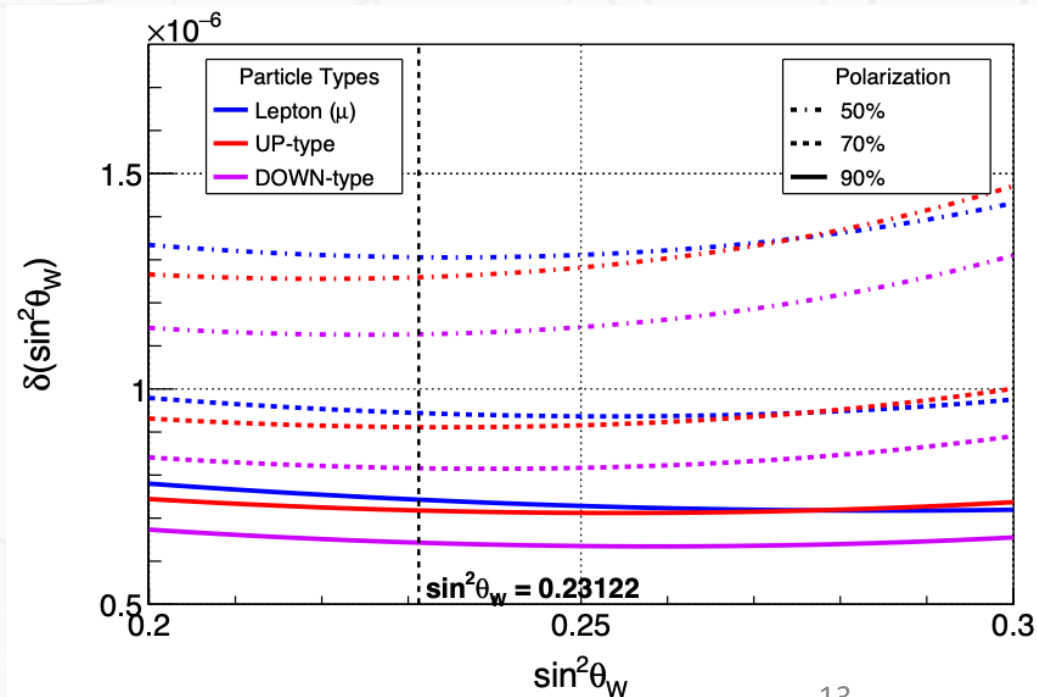
- The power of beam polarization for electroweak precision measurements was convincingly demonstrated by the SLC experiment
 - ◆ With less statistics using a polarized electron beam
 - ◆ A precision comparable to — and in some respects exceeding

	$\sin^2 \theta_{\text{eff}}^{\text{lept}}$	$\Delta \sin^2 \theta_{\text{eff}}^{\text{lept}} (10^{-3})$
LEP, $A_{FB}^{0,b}$	0.23221 ± 0.00029	1.3
LEP, $A_{FB}^{0,l}$	0.23099 ± 0.00053	2.3
SLC, A_{LR}	0.23098 ± 0.00026	1.2
LEP+SLC	0.23153 ± 0.00016	0.70

- The statistical advantage of the polarized measurement could be estimated using Fisher Information
 - The factor $\sqrt{2}$ accounts for the fact that the total integrated luminosity is split between the left-handed and right-handed polarization states.

$$\frac{\Delta(\sin^2 \theta_W^{\text{eff}})_{\text{pol}}}{\Delta(\sin^2 \theta_W^{\text{eff}})_{\text{unpol}}} = \frac{\sqrt{2}}{P_e} \cdot \frac{|\partial A_{FB} / \partial \sin^2 \theta_W^{\text{eff}}|}{|\partial A_{LR} / \partial \sin^2 \theta_W^{\text{eff}}|},$$

- The expected precision of $\sin^2 \theta_W^{\text{eff}}$ measurements using the Fisher information method



Extension to polarization scenario

- The expected bin-by-bin number of left(right)-handed events

$$N_{L(R),i,j}^{\text{exp}} = \mathcal{L} \Delta \sum_k \sigma_{L(R),i,j} \epsilon_{i,j,k},$$

- $\sin^2 \theta_W^{\text{eff}}$ could be extracted via a binned maximum likelihood fit

$$\mathcal{L}(\sin^2 \theta_W^{\text{eff}}, N_{L,j}^{\text{prod}}, N_{R,j}^{\text{prod}}) = \prod_i \frac{\mu_{L,i,j}^{n_{L,i,j}} e^{-\mu_{L,i,j}}}{n_{L,i,j}!} \times \frac{\mu_{R,i,j}^{n_{R,i,j}} e^{-\mu_{R,i,j}}}{n_{R,i,j}!},$$

- 1000 toy Monte Carlo samples, pull distributions are also standard
- The results (10^{-6}) are listed

	$P_e = 0.5$	$P_e = 0.6$	$P_e = 0.7$	$P_e = 0.8$
b	0.55	0.54	0.52	0.51
c	0.71	0.69	0.68	0.65
s	0.83	0.82	0.81	0.79

Discussion of systematic uncertainties

- Detector modeling
 - ◆ The abundant Bhabha and dimuon events at the Z-pole enable a data-driven validation of the detector simulation
 - ◆ The corresponding systematic uncertainties to be controlled at the same level as the statistical.
- Flavor-tagging calibration
 - ◆ A straightforward choice is taking a half the data as control sample
 - ◆ The efficiency and mistag rates could be controlled to the same level of the statistical uncertainty
- Beam energy and polarization
 - ◆ The systematic uncertainty from beam energy is estimated to be 1.7×10^{-6} for A_{FB} and 1.0×10^{-6} for $\sin^2 \theta_W^{\text{eff}}$
 - ◆ Assuming a 0.1% relative uncertainty on the polarization, the corresponding systematic uncertainty is 1.7×10^{-5} .
 - ◆ Affects of luminosity (10^{-4}) are still studying, but it would be negligible reasonably.

Discussion of systematic uncertainties

- QCD and hadronization modeling
 - ◆ The systematic uncertainty from QCD corrections is estimated to be at the 10^{-4} level
 - ◆ These effects must be modeled into next generation event generators, which must be validated and tuned with high statistics data.
- Background contamination
 - ◆ The uncertainty in the normalization and shape of these backgrounds must be controlled to a level well below the statistical precision
- Beamstrahlung effects
 - ◆ The associated systematic uncertainty — mainly from mis-modeling of photon emission — does not affect jet energy, direction, or identification
 - ◆ Be safely ignored as a background in precision jet analyses

- the potential for measuring the forward-backward asymmetry A_{FB} and the effective weak mixing angle $\sin^2 \theta_W^{\text{eff}}$ @ CEPC, based on a sample of 4×10^{12} Z bosons.
- The statistical precision achievable is unprecedented.

	This work (10^{-3})	PDG values(10^{-3})		b	c	s		$P_e = 0.5$	$P_e = 0.6$	$P_e = 0.7$	$P_e = 0.8$
ΔA_{FB}^b	0.013	16	$\Delta \sin^2 \theta_{\text{eff}}^{\text{lept}}$	1.0	1.5	1.4	b	0.55	0.54	0.52	0.51
ΔA_{FB}^c	0.021	50		c	0.71	0.69	0.68	0.65			
ΔA_{FB}^s	0.018	148		s	0.83	0.82	0.81	0.79			

- The control of systematic uncertainties will become the dominant challenge.
 - ◆ Several important sources has been addressed.