Neutral-Current Electroweak Physics and SMEFT Studies at the EIC

Radja Boughezal
Argonne National Laboratory, Lemont, IL, USA

Alex Emmert, Michael Nycz, and Xiaochao Zheng
University of Virginia, Charlottesville, VA, USA

Tyler Kutz
Massachusetts Institute of Technology, Cambridge, MA, USA

Sonny Mantry
University of North Georgia, Dahlonega, GA, USA

Frank Petriello
Argonne National Laboratory, Lemont, IL, USA and
Northwestern University, Evanston, IL, USA

Kağan Şimşek and Daniel Wiegand
Northwestern University, Evanston, IL, USA

(Dated: March 28, 2022)

Abstract

We study the potential for precision electroweak (EW) measurements and beyond-the-Standard Model (BSM) searches using cross section asymmetries in neutral-current deep inelastic scattering at the electron-ion collider (EIC). Our analysis uses a complete and realistic accounting of systematic errors from both theory and experiment, and considers the potential of both proton and deuteron beams for a wide range of energies and luminosities. We also consider what can be learned from a possible future positron beam and a ten-fold luminosity upgrade of EIC. We use the SM effective field theory (SMEFT) framework to parameterize BSM effects, and focus on semi-leptonic four-fermion operators, while for our precision EW study we determine how well the EIC can measure the weak mixing angle. New features of our study include the use of an up-to-date detector design of ECCE (EIC Comprehensive Chromodynamics Experiment) and accurate running conditions of the EIC, the simultaneous fitting of beam polarization uncertainties and Wilson coefficients to improve the sensitivity to SMEFT operators, and the inclusion of the weak mixing angle running in our fit template. We find that the EIC can probe BSM operators at scales competitive with and in many cases exceeding LHC Drell-Yan bounds while simultaneously not suffering from degeneracies between Wilson coefficients.

CONTENTS

I. INTRODUCTION

The Standard Model (SM) of particle physics currently describes all known laboratory phenomena. All particles predicted by the SM have now been found after the discovery of the Higgs boson at the Large Hadron Collider (LHC). No new particles beyond those present in the SM have been discovered, and no appreciable deviation from SM predictions has been conclusively observed. Despite the enormous success of this theory it contains numerous shortcomings. It does not contain an explanation of the dark matter observed in the universe nor the baryon-antibaryon asymmetry, and it does not describe neutrino masses. It additionally suffers from several aesthetic issues such as the hierarchy problem and an extreme hierarchy of fermion Yukawa couplings. Even the sectors of the theory that have been experimentally successful still contain unsatisfying and poorly understood features. For example, the exact composition of the proton spin in terms of the spin and orbital angular momentum of its constituent quarks and gluons is still poorly known.

Numerous experimental programs that attempt to address these residual issues in our understanding of Nature are either running or under design. Our focus in this manuscript will be on the Electron Ion Collider (EIC) to be built at Brookhaven National Lab in Upton, New York. The EIC will be a particle accelerator that collides electrons with protons and nuclei in the intermediate energy range between fixed-target scattering facilities and high energy colliders. It will provide orders of magnitude higher luminosity than HERA, the only electron-proton collider operated to date. It will also be the first lepton-ion collider with the ability to polarize both the electron and the proton (ion) beams, and the first collider with fast spin-flip capacity. These unique design features will allow a direct extraction of parity-violating (PV) asymmetries in the electroweak neutral-current scattering cross section associated with either the electron – $A_{PV}^{(e)}$ – or the proton (ion) spin flip – $A_{PV}^{(p(D))}$. Experimental uncertainties from effects such as luminosity measurement and detector acceptance or efficiency will be substantially reduced due to these capabilities.

Although the EIC was designed primarily to explore outstanding issues in QCD such

as the proton spin issue mentioned above, it additionally has a strong potential to probe several aspects of precision electroweak (EW) and beyond-the-SM (BSM) physics as well. It can measure the value of the weak mixing angle over a wide range of momentum transfers complementary to Z-pole measurements and low-energy determinations. The possibility of polarizing both electron and proton/ion beams gives it unique handles on BSM physics. Our goal in this manuscript is to provide a detailed accounting of the EW and BSM potential of the EIC, with a realistic simulation of anticipated experimental uncertainties. We explore the use of the asymmetries $A_{PV}^{(e)}$ and $A_{PV}^{(p(D))}$. In addition to determining the BSM reach of PV observables, we consider the reach of the lepton-charge asymmetry $A_{LC,p(D)}$ at the EIC for the first time, assuming a positron beam will become available in the future.

Since no new particles beyond the SM have so far been discovered, we adopt the Standard Model Effective Field Theory (SMEFT) for our BSM studies. The SMEFT contains higher-dimensional operators formed from SM fields, assuming all new physics is heavier than both SM states and the accessible collider energy. The leading dimension-6 operator basis of SMEFT for on-shell fields has been completely classified (there is a dimension-5 operator that violates lepton number which we do not consider here). We find that the EIC can probe the full spectrum of SMEFT operators to the few-TeV level or beyond. The wide variety of observables possible at the EIC, which include several asymmetries with either proton or ion beams, ensure that no flat directions remain in the Wilson coefficient parameter space, unlike at the LHC. Our analysis on the determination of the weak mixing angle, assuming a realistic annual luminosity and accounting for all experimental and theoretical uncertainties to the best level that can be reached at pre-EIC running stage, found good precision for this fundamental SM parameter in a kinematic region not explored before. The precision will continue to improve as data are accumulated from the decades-long running of the EIC.

Our paper is organized as follows: In Section II, we first provide a complete description of deep inelastic scattering (DIS) formalism that includes both SM contributions and SMEFT extensions. The DIS cross sections that account for both electron and hadron polarizations are provided in both structure function and parton-model languages. We follow this theoretical framework by presenting a basic strategy to measure the different polarization components of the cross sections and forming the PV asymmetries at the EIC. Measurement of the lepton-charge (LC) asymmetry is also discussed. In Section III, we present data simulation based on the design of the ECCE Detector (now EIC Detector 1) using

a fast-smearing method and event selection criteria, followed by projections of statistical precision for PV and LC asymmetries based on the planned annual luminosity of the EIC. Generation of pseudo data as well as the uncertainty matrix are presented in Section IV, followed by extractions of the EW mixing angle in Section V. In Section VI, we provide extensive description of our SMEFT analysis framework, with representative results on single- and two-Wilson coefficient fits given in Section VII. We conclude in Section VIII. In Appendix A, we present novel analysis methods to simultaneously fit PV asymmetries and the beam polarization, or LC asymmetries and the luminosity difference between e^+ and e^- runs. A complete collection of all SMEFT 1- and 2-Wilson coefficient fit results from this study are given in Appendix B.

II. NEUTRAL-CURRENT DIS MEASUREMENT AT THE EIC

A. Deep Inelastic Scattering and SMEFT Formalism

In this section, we give a brief overview of the formalism for DIS and SMEFT. In particular, we generalize the SM DIS cross section and asymmetry formulae to include contributions from SMEFT operators which encode new physics at an energy level Λ that lies well beyond the electroweak scale. We denote electron scattering off a nucleus as:

$$\ell(k) + H(P) \longrightarrow \ell(k') + X, \tag{1}$$

where ℓ stands for an electron or positron, the hadron H stands for either the proton (p) or the deuteron (D), and X denotes the final state hadronic system. The four momenta of the initial lepton, final lepton, and the initial hadron are denoted as k, k', and P, respectively. Using the momenta of the initial and final state leptons and the initial state hadron, one can define the following Lorentz invariant kinematic variables:

$$s = (P + k)^2, \tag{2}$$

$$Q^2 = -(k - k')^2, \tag{3}$$

$$x = \frac{Q^2}{2P \cdot (k - k')}, \tag{4}$$

$$y = \frac{P \cdot (k - k')}{P \cdot l}, \tag{5}$$

$$W^2 = (P + k - k')^2, \tag{6}$$

where s is the center of mass energy squared, Q^2 is the negative of the lepton four-momentum transfer squared, the Bjorken-x variable is the longitudinal hadron momentum fraction carried by the struck parton, the inelasticity parameter y gives the fractional energy loss of the lepton in the hadron rest frame, and W gives the invariant mass of the final state hadronic system X. The x, y, s, and Q^2 kinematic variables are related to each other as $Q^2 = xy(s - M^2)$ where M is the mass of the nucleon.

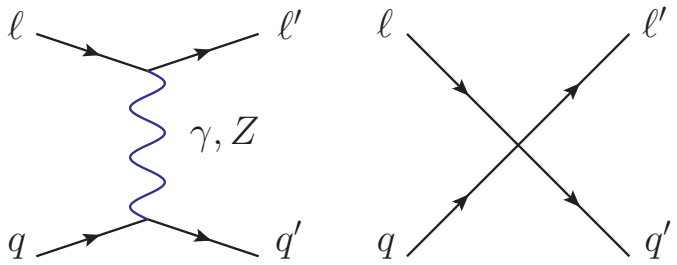

FIG. 1. The Feynman diagrams for $\ell + H \to \ell + X$ at the parton level from one-boson exchange (left) and SMEFT contact interactions (right).

The schematic in Fig. 1 shows the partonic tree-level processes that contribute to Eq. (1). These are the contributions to the total tree-level amplitude from single photon exchange, single Z^0-boson exchange, and the SMEFT contact interactions. The SMEFT Lagrangian that describes these contact interactions has the form

$$\mathcal{L}_{\text{SMEFT}} = \frac{1}{\Lambda^2} \sum_r C_r \mathcal{O}_r + \cdots , \tag{7}$$

where the sum over the index r runs over the set of dimension-6 SMEFT operators and the ellipses denote SMEFT operators of mass dimension greater than 6. We restrict our analysis to include only the effects of dimension-6 SMEFT operators since the higher dimensional operators are formally suppressed by additional powers of E^2/Λ^2, where E is the typical energy scale of the scattering process. $\mathcal{O}_r$ denotes the r-th dimension-6 operator and C_r is the corresponding (dimensionless) Wilson coefficient arising from integrating out the new physics degrees of freedom at the Λ-scale. These Wilson coefficients can be constrained through a comparison of SM predictions with precision measurements of various processes studied in a variety of experiments across a wide range of energy scales.

The subset of dimension-6 operators that we consider in our analysis of DIS are given in Table I. We note that there are additional SMEFT operators but they are known to be far

C_r	$\mathcal{O}_r$	$\tilde{C}_r$	$c^e_{V_r}$	$c^e_{A_r}$	$c^u_{V_r}$	$c^u_{A_r}$	$c^{d,s}_{V_r}$	$c^{d,s}_{A_r}$
$C^{(1)}_{\ell q}$	$\mathcal{O}^{(1)}_{\ell q} = (\bar{L}_L\gamma^\mu L_L)(\bar{Q}_L\gamma_\mu Q_L)$	$C^{(1)}_{\ell q}/4$	1	1	1	1	1	1
$C^{(3)}_{\ell q}$	$\mathcal{O}^{(3)}_{\ell q} = (\bar{L}_L\gamma^\mu\tau^I L_L)(\bar{Q}_L\gamma_\mu\tau^I Q_L)$	$C^{(3)}_{\ell q}/4$	1	1	-1	-1	1	1
C_{eu}	$\mathcal{O}_{eu} = (\bar{e}_R\gamma^\mu e_R)(\bar{u}_R\gamma_\mu u_R)$	$C_{eu}/4$	1	-1	1	-1	0	0
C_{ed}	$\mathcal{O}_{ed} = (\bar{e}_R\gamma^\mu e_R)(\bar{d}_R\gamma_\mu d_R)$	$C_{ed}/4$	1	-1	0	0	1	-1
$C_{\ell u}$	$\mathcal{O}_{\ell u} = (\bar{L}_L\gamma^\mu L_L)(\bar{u}_R\gamma_\mu u_R)$	$C_{\ell u}/4$	1	1	1	-1	0	0
$C_{\ell d}$	$\mathcal{O}_{\ell u} = (\bar{L}_L\gamma^\mu L_L)(\bar{d}_R\gamma_\mu d_R)$	$C_{\ell d}/4$	1	1	0	0	1	-1
C_{qe}	$\mathcal{O}_{qe} = (\bar{Q}_L\gamma^\mu Q_L)(\bar{e}_R\gamma_\mu e_R)$	$C_{qe}/4$	1	-1	1	1	1	1

TABLE I. List of SMEFT operators relevant to DIS in the basis of SM fields before electroweak symmetry breaking and rexpressed in the vector and axial-vector current basis after electroweak symmetry breaking: $C_r\mathcal{O}_r = \tilde{C}_r \sum_f \bar{e}\gamma^\mu(c^e_{V_r} - c^e_{A_r}\gamma_5)e\,\bar{q}_f\gamma^\mu(c^f_{V_r} - c^f_{A_r}\gamma_5)q_f + \cdots$.

better bounded through other data sets such as precision Z-pole observables, and we neglect them here. The above assumptions leave us with the seven Wilson coefficients associated with listed operators which enter the predictions for DIS cross sections and asymmetries.

As seen in Table I, the SMEFT operators $\mathcal{O}_r$ are expressed in terms of the basis of SM fields before electroweak symmetry breaking. For the purposes of DIS phenomenology below the electroweak scale, it is useful to rewrite these SMEFT operators in the vector and axial-vector basis using Dirac fields that describe the massive electrons (e) and quarks (q_f) after electroweak symmetry breaking:

$$\mathcal{L}_{\rm SMEFT} = \frac{1}{\Lambda^2}\sum_r \tilde{C}_r\Big\{\sum_f \bar{e}\gamma^\mu(c^e_{V_r} - c^e_{A_r}\gamma_5)e\,\bar{q}_f\gamma^\mu(c^f_{V_r} - c^f_{A_r}\gamma_5)q_f\Big\} + \cdots\,, \tag{8}$$

where the specific values of the vector and axial-vector couplings $c^{e,q}_{V_k}$ and $c^{e,q}_{A_k}$, respectively, for the r-th SMEFT operator follow from the corresponding chiral and flavor structure of the SMEFT operators. The $\tilde{C}_r$ coefficients are related to the C_r by an overall factor and can be fixed by comparing Eqs. (7) and (8). There is freedom to always redefine the $\tilde{C}_r$ by absorbing an overall factor into the $c^{e,q}_{V_r}, c^{e,q}_{A_r}$ couplings. We specify the exact definitions we use in Table I. These couplings are analogous to the vector and axial-vector couplings, $g^{e,q}_V$ and $g^{e,q}_A$, of the Z^0-boson but are instead generated from integrating out UV physics associated with the scale Λ.

As seen in Fig. 1, the total tree-level amplitude can be decomposed into three contribu-

tions

$$\mathscr{M} = \mathscr{M}_\gamma + \mathscr{M}_Z + \mathscr{M}_\times, \tag{9}$$

where $\mathscr{M}_\gamma, \mathscr{M}_Z$, and $\mathscr{M}_\times$ denote the contributions from single photon exchange, single Z^0-boson exchange, and the SMEFT operators, respectively. In particular, $\mathscr{M}_\times = \sum_r \mathscr{M}_r$, where the sum over the index r runs over the amplitudes arising from the SMEFT operators listed in Table I. Up to leading order in the SMEFT power counting, where only dimension six SMEFT operators that scale as $\sim 1/\Lambda^2$ are kept, the total amplitude squared can be written as:

$$|\mathscr{M}|^2 = \mathscr{M}_{\gamma\gamma} + 2\mathscr{M}_{\gamma Z} + \mathscr{M}_{ZZ} + 2\mathscr{M}_{\gamma\times} + 2\mathscr{M}_{Z\times}, \tag{10}$$

where $\mathscr{M}_{\gamma\gamma} = |\mathscr{M}_\gamma|^2$, $\mathscr{M}_{ZZ} = |\mathscr{M}_Z|^2$, $2\mathscr{M}_{\gamma Z} = \mathscr{M}_\gamma^* \mathscr{M}_Z + \mathscr{M}_\gamma \mathscr{M}_Z^*$, $2\mathscr{M}_{\gamma\times} = \mathscr{M}_\gamma^* \mathscr{M}_\times + \mathscr{M}_\times \mathscr{M}_\gamma^*$, and $2\mathscr{M}_{Z\times} = \mathscr{M}_Z^* \mathscr{M}_\times + \mathscr{M}_\times \mathscr{M}_Z^*$. These denote the single photon exchange amplitude, single Z^0-boson exchange amplitude, interference between the single photon and single Z^0-boson exhange amplitudes, interference between single photon exchange and the SMEFT amplitudes, and interference between the single Z^0-boson exchange and SMEFT amplitudes, respectively. Here we ignore the $|\mathscr{M}_\times|^2$ contribution since it scales as $1/\Lambda^4$, formally the same size as contributions from dimension eight SMEFT operators interfering with the SM. For the hadron-level cross sections and asymmetries, these different contributions will give rise to corresponding structure functions. In particular, in addition to the usual structure functions encountered in SM DIS, new structure functions corresponding to SMEFT contributions will arise. Thus, including SMEFT contributions, the DIS differential cross section takes the general form

$$\frac{d^2\sigma}{dxdy} = \frac{2\pi y\alpha^2}{Q^4}\Big\{ \eta^\gamma L^\gamma_{\mu\nu} W^{\mu\nu}_\gamma + \eta^{\gamma Z} L^{\gamma Z}_{\mu\nu} W^{\mu\nu}_{\gamma Z} + \eta^Z L^Z_{\mu\nu} W^{\mu\nu}_Z + \sum_r \xi^{\gamma r} L^{\gamma r}_{\mu\nu} W^{\mu\nu}_{\gamma r} + \sum_r \xi^{Zr} L^{Zr}_{\mu\nu} W^{\mu\nu}_{Zr} \Big\}, \tag{11}$$

where α is the electromagnetic fine structure constant. The $L^{\gamma,\gamma Z,Z,\gamma r,Zr}_{\mu\nu}$ and $W^{\gamma,\gamma Z,Z,\gamma r,Zr}_{\mu\nu}$ are the leptonic and hadronic tensors, respectively. The first three terms on the RHS correspond to SM contributions from $\mathscr{M}_{\gamma\gamma}$, $2\mathscr{M}_{\gamma Z}$, and $\mathscr{M}_{ZZ}$, respectively, and the last two sets of terms correspond to contributions from SMEFT operators, i.e. $2\mathscr{M}_{\gamma\times}$ and $2\mathscr{M}_{Z\times}$, respectively. For completeness, below we collect some useful results to make the form of the cross section explicit. The dimensionless coefficients $\eta^{\gamma,\gamma Z,Z}, \xi^{\gamma r}$, and ξ^{Zr} are given by

$$\eta^\gamma = 1,$$

$$
\begin{aligned}
\eta^{\gamma Z} &= \frac{G_F M_Z^2}{2\sqrt{2}\pi\alpha}\frac{Q^2}{Q^2+M_Z^2}, \\
\eta^{Z} &= \left(\eta^{\gamma Z}\right)^2, \\
\xi^{\gamma r} &= \frac{\tilde{C}_r}{4\pi\alpha}\frac{Q^2}{\Lambda^2}, \\
\xi^{Z r} &= \eta^{\gamma Z}\frac{\tilde{C}_r}{4\pi\alpha}\frac{Q^2}{\Lambda^2},
\end{aligned}
\tag{12}
$$

where $G_F = 1.1663787(6) \times 10^{-5}$ GeV^{-2} is the Fermi constant and $M_Z = 91.1876 \pm 0.0021$ GeV [1] is the mass of the Z boson. The leptonic tensors in Eq. (11) are

$$
\begin{aligned}
L^{\gamma}_{\mu\nu} &= 2\big[\, k_\mu k'_\nu + k'_\mu k_\nu - k\cdot k' g_{\mu\nu} - i\lambda_e \epsilon_{\mu\nu\alpha\beta} k^\alpha (k')^\beta \,\big], \\
L^{\gamma Z}_{\mu\nu} &= -(g_V^e - \lambda_e g_A^e) L^{\gamma}_{\mu\nu}, \\
L^{Z}_{\mu\nu} &= (g_V^e - \lambda_e g_A^e)^2 L^{\gamma}_{\mu\nu}, \\
L^{\gamma r}_{\mu\nu} &= -(c_{V_r}^e - \lambda_e c_{A_r}^e) L^{\gamma}_{\mu\nu}, \\
L^{Z r}_{\mu\nu} &= (c_{V_r}^e - \lambda_e c_{A_r}^e)(g_V^e - \lambda_e g_A^e) L^{\gamma}_{\mu\nu},
\end{aligned}
\tag{13}
$$

where $\lambda_e = -1, 1$ denotes the lepton helicity. For positrons, one flips the sign of all g_A^e and c_{Ar}^e terms above, and the overall sign of $L^{\gamma Z}$ and $L^{\gamma r}$. Using these identities for the leptonic tensors, Eq. (11) can be written more explicitly as

$$
\begin{aligned}
\frac{d^2\sigma}{dxdy} = \frac{2\pi y\alpha^2}{Q^4} L^{\gamma}_{\mu\nu} \Big\{ & \eta^{\gamma} W_{\gamma}^{\mu\nu} - \eta^{\gamma Z}(g_V^e - \lambda_e g_A^e) W_{\gamma Z}^{\mu\nu} + \eta^{Z}(g_V^e - \lambda_e g_A^e)^2 W_{Z}^{\mu\nu} \\
& - \sum_r \xi^{\gamma r}(c_{V_r}^e - \lambda_e c_{A_r}^e) W_{\gamma r}^{\mu\nu} + \sum_r \xi^{Z r}(c_{V_r}^e - \lambda_e c_{A_r}^e)(g_V^e - \lambda_e g_A^e) W_{Z r}^{\mu\nu} \Big\}.
\end{aligned}
\tag{14}
$$

Based on the general Lorentz tensor structure and the available four momenta and the nucleus spin vector S^μ, the various hadronic tensors are parameterized in terms of structure functions as

$$
\begin{aligned}
W^j_{\mu\nu} = & (-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}) F_1^j + \frac{\hat{P}_\mu \hat{P}_\nu}{(P\cdot q)} F_2^j + \frac{i\epsilon_{\mu\nu\alpha\beta}}{2(P\cdot q)}\Big[P^\alpha q^\beta F_3^j + 2q^\alpha S^\beta g_1^j \Big] \\
& - \frac{S\cdot q}{(P\cdot q)}\Big[\frac{\hat{P}_\mu \hat{P}_\nu}{P\cdot q} g_4^j + \frac{S\cdot q}{P\cdot q} g_{\mu\nu} g_5^j \Big],
\end{aligned}
\tag{15}
$$

where $\hat{P}_\mu \equiv P_\mu - q_\mu (P\cdot q)/q^2$ and $\hat{S}_\mu. \equiv S_\mu - q_\mu (S\cdot q)/q^2$. The index j denotes the possibilities $\{\gamma, \gamma Z, Z, \gamma r, Z r\}$, and $F^j_{1,2,3}$ and $g^j_{1,4,5}$ denote various unpolarized and polarized nuclear structure functions, respectively. We have omitted two additional possible Lorentz structures in the hadronic tensor, typically denoted as the g_2 and g_3 polarized structure functions, since these terms give a contribution to the cross section that is suppressed by

M^2/Q^2 when contracted with the leptonic tensor. Thus, we do not consider the $g_{2,3}$ structure functions in the rest of our analysis. The nucleus spin vector S^μ satisfies the constraints $S^2 = -M^2$ and $S \cdot P = 0$. For longitudinal polarization, it takes the canonical form: $S^\mu = \lambda_H \, (|\mathbf{p}|, \, E\frac{\mathbf{p}}{|\mathbf{p}|})$ where λ_H is the nucleon helicity, $\lambda_H = \pm 1$, and the nucleon four momentum is $P^\mu = (E, \mathbf{p})$.

Based on the structure of the cross section in Eq. (14), in conjunction with the form of the hadronic tensor in Eq. (15), it becomes useful to define the following combinations of structure functions that also include the SMEFT contributions

$$\begin{aligned} F_i &= F_i^{\rm SM,NC} + F_i^{\rm SMEFT}, \\ g_i &= g_i^{\rm SM,NC} + g_i^{\rm SMEFT}, \end{aligned} \tag{16}$$

where the SM contributions are given by the commonly known NC structure functions

$$\begin{aligned} F_i^{\rm SM,NC} &= F_i^\gamma - \eta^{\gamma Z}(g_V^e - \lambda_e g_A^e)F_i^{\gamma Z} + \eta^Z (g_A^e + \lambda_e g_V^e)^2 F_i^Z, \\ g_i^{\rm SM,NC} &= g_i^\gamma - \eta^{\gamma Z}(g_V^e - \lambda_e g_A^e)g_i^{\gamma Z} + \eta^Z (g_A^e + \lambda_e g_V^e)^2 g_i^Z, \end{aligned} \tag{17}$$

and SMEFT contributions are given by

$$\begin{aligned} F_i^{\rm SMEFT} &= -\sum_r \xi^{\gamma r}(c_{V_r}^e - \lambda_e c_{A_r}^e)F_i^{\gamma r} + \sum_r \xi^{Zr}(c_{V_r}^e - \lambda_e c_{A_r}^e)(g_V^e - \lambda_e g_A^e)F_i^{Zr}, \\ g_i^{\rm SMEFT} &= -\sum_r \xi^{\gamma r}(c_{V_r}^e - \lambda_e c_{A_r}^e)g_i^{\gamma r} + \sum_r \xi^{Zr}(c_{V_r}^e - \lambda_e c_{A_r}^e)(g_V^e - \lambda_e g_A^e)g_i^{Zr}. \end{aligned} \tag{18}$$

The parton-model expressions for the SM structure functions are summarized below. We also provide the corresponding expressions for the structure functions arising from the interference of the SM with SMEFT operators:

$$\begin{aligned} \left[F_2^\gamma, F_2^{\gamma Z}, F_2^Z, F_2^{\gamma r}, F_2^{Zr}\right] &= x\sum_f \left[Q_f^2, 2Q_f g_V^f, {g_V^f}^2 + {g_A^f}^2, 2Q_f c_{V_r}^f, 2(g_V^f c_{V_r}^f + g_A^f c_{A_r}^f)\right](q_f + \bar{q}_f), \\ \left[F_3^\gamma, F_3^{\gamma Z}, F_3^Z, F_3^{\gamma r}, F_3^{Zr}\right] &= \sum_f \left[0, 2Q_f g_A^f, 2g_V^f g_A^f, 2Q_f c_{Ar}^f, 2(g_V^f c_{A_r}^f + g_A^f c_{V_r}^f)\right](q_f - \bar{q}_f), \\ \left[g_1^\gamma, g_1^{\gamma Z}, g_1^Z, g_1^{\gamma r}, g_1^{Zr}\right] &= \frac{1}{2}\sum_f \left[Q_f^2, 2Q_f g_V^f, {g_V^f}^2 + {g_A^f}^2, 2Q_f c_{V_r}^f, 2(g_V^f c_{V_r}^f + g_A^f c_{A_r}^f)\right](\Delta q_f + \Delta\bar{q}_f), \\ \left[g_5^\gamma, g_5^{\gamma Z}, g_5^Z, g_5^{\gamma r}, g_5^{Zr}\right] &= \sum_f \left[0, Q_f g_A^f, g_V^f g_A^f, Q_f c_{A_r}^f, g_V^f c_{A_r}^f + g_A^f c_{V_r}^f\right](\Delta q_f - \Delta\bar{q}_f), \end{aligned} \tag{19}$$

where $q_f(x, Q^2)$ and $\Delta q_f(x, Q^2)$ are unpolarized and polarized Parton Distribution Functions (PDFs) of quark flavor f, respectively, and Q_f denotes the electric charge in units of the

proton charge e. In the parton model, at leading-order (LO), one has for the structure functions the Callan-Gross relations $F_2^i = 2xF_1^i$ and $g_4^j = 2xg_5^i$ for $i = \gamma, \gamma Z, Z, \gamma r, Zr$. For an ion beam (or nuclear target), the neutron PDFs can be related to the proton PDFs assuming isospin symmetry for the valence quarks:

$$\begin{aligned} q_{u/n}(x,Q^2) &= q_{d/p}(x,Q^2), \\ q_{d/n}(x,Q^2) &= q_{u/p}(x,Q^2), \\ \Delta q_{u/n}(x,Q^2) &= \Delta q_{d/p}(x,Q^2), \\ \Delta q_{d/n}(x,Q^2) &= \Delta q_{u/p}(x,Q^2), \end{aligned} \tag{20}$$

while the charm and strange sea quark PDFs are assumed to be identical for the proton and the neutron:

$$\begin{aligned} q_{s/n}(x,Q^2) &= q_{s/p}(x,Q^2), \\ q_{c/n}(x,Q^2) &= q_{c/p}(x,Q^2), \\ \Delta q_{s/n}(x,Q^2) &= \Delta q_{s/p}(x,Q^2), \\ \Delta q_{c/n}(x,Q^2) &= \Delta q_{c/p}(x,Q^2), \end{aligned} \tag{21}$$

For the deuteron, an isoscalar bound state of a proton and a neutron, the PDFs can be constructed from the proton and neutron PDFs as

$$\begin{aligned} q_{f/D}(x,Q^2) &= \frac{1}{2}(q_{f/p}(x,Q^2) + q_{f/n}(x,Q^2)), \\ \Delta q_{f/D}(x,Q^2) &= \frac{1}{2}(\Delta q_{f/p}(x,Q^2) + \Delta q_{f/n}(x,Q^2)), \end{aligned} \tag{22}$$

for quark flavor f.

In terms of the generalized structure functions in Eq. (16), which include a dependence on the electron helicity λ_e as seen in Eqs. (17) and (18), one can write the cross section for given electron and nucleon helicities, including SMEFT operator contributions, as

$$\begin{aligned} \frac{d^2\sigma(\lambda_e,\lambda_H)}{dxdy} = \frac{4\pi\alpha^2}{xyQ^2}\Bigg\{ & xy^2F_1 + (1-y)F_2 - \lambda_e\frac{y}{2}(2-y)\,xF_3 + \lambda_e\lambda_H(2-y)\,xy\,g_1 \\ & - \lambda_H(1-y)\,g_4 - \lambda_H\,xy^2\,g_5 \Bigg\}, \end{aligned} \tag{23}$$

where we have ignored the electron mass and all target mass correction terms that are proportional to M^2/Q^2.

To connect to experimentally measured observables, it is convenient to write the scattering cross section of Eq. (23) as the sum of four components that depend on the spin direction of the initial electron and hadron: $d\sigma_0, d\sigma_e, d\sigma_H$, and $d\sigma_{eH}$, where each $d\sigma$ represents the differential cross section such as $d\sigma/(dxdy)$. The quantity $d\sigma_0$ is the unpolarized cross section, $d\sigma_e$ and $d\sigma_H$ denote the cross section differences between initial electron and hadron states of opposite helicity, respectively, and $d\sigma_{eH}$ is the cross sections difference between initial electron-hadron states with the same and opposite helicities defined in the center-of-mass frame. These quantities can be formed from Eq. (23) as:

$$\begin{aligned}
d\sigma_0 &= \frac{1}{4}\Big[\, d\sigma|_{\lambda_e=+1,\lambda_H=+1} + d\sigma|_{\lambda_e=+1,\lambda_H=-1} + d\sigma|_{\lambda_e=-1,\lambda_H=+1} + d\sigma|_{\lambda_e=-1,\lambda_H=-1} \,\Big], \\
d\sigma_e &= \frac{1}{4}\Big[\, d\sigma|_{\lambda_e=+1,\lambda_H=+1} + d\sigma|_{\lambda_e=+1,\lambda_H=-1} - d\sigma|_{\lambda_e=-1,\lambda_H=+1} - d\sigma|_{\lambda_e=-1,\lambda_H=-1} \,\Big], \\
d\sigma_H &= \frac{1}{4}\Big[\, d\sigma|_{\lambda_e=+1,\lambda_H=+1} - d\sigma|_{\lambda_e=+1,\lambda_H=-1} + d\sigma|_{\lambda_e=-1,\lambda_H=+1} - d\sigma|_{\lambda_e=-1,\lambda_H=-1} \,\Big], \\
d\sigma_{eH} &= \frac{1}{4}\Big[\, d\sigma|_{\lambda_e=+1,\lambda_H=+1} - d\sigma|_{\lambda_e=+1,\lambda_H=-1} - d\sigma|_{\lambda_e=-1,\lambda_H=+1} + d\sigma|_{\lambda_e=-1,\lambda_H=-1} \,\Big],
\end{aligned} \tag{24}$$

and can be computed in conjunction with Eqs. (16), (17), and (18).

The SM contribution to the DIS cross sections are (omitting target-mass terms):

$$\begin{aligned}
\frac{d^2\sigma_0}{dxdy} &= \frac{4\pi\alpha^2}{xyQ^2}\Big\{(1-y)\left[F_2^\gamma - g_V^e \eta_{\gamma Z} F_2^{\gamma Z} + ({g_V^e}^2 + {g_A^e}^2)\eta_Z F_2^Z\right] \\
&\quad + xy^2\left[F_1^\gamma - g_V^e \eta_{\gamma Z} F_1^{\gamma Z} + ({g_V^e}^2 + {g_A^e}^2)\eta_Z F_1^Z\right] \\
&\quad - \frac{xy}{2}(2-y)\left[g_A^e \eta_{\gamma Z} F_3^{\gamma Z} - 2 g_V^e g_A^e \eta_Z F_3^Z\right]\Big\}\ , \\
\frac{d^2\sigma_e}{dxdy} &= \frac{4\pi\alpha^2}{xyQ^2}\Big\{(1-y)\left[g_A^e \eta_{\gamma Z} F_2^{\gamma Z} - 2 g_V^e g_A^e \eta_Z F_2^Z\right] + xy^2\left[g_A^e \eta_{\gamma Z} F_1^{\gamma Z} - 2 g_V^e g_A^e \eta_Z F_1^Z\right] \\
&\quad + \frac{xy}{2}(2-y)\left[g_V^e \eta_{\gamma Z} F_3^{\gamma Z} - ({g_V^e}^2 + {g_A^e}^2)\eta_Z F_3^Z\right]\Big\}\ , \\
\frac{d^2\sigma_H}{dxdy} &= \frac{4\pi\alpha^2}{xyQ^2}\Big\{(2-y)\,xy\left[g_A^e \eta_{\gamma Z} g_1^{\gamma Z} - 2 g_V^e g_A^e \eta_Z g_1^Z\right] \\
&\quad - (1-y)\left[-g_V^e \eta_{\gamma Z} g_4^{\gamma Z} + ({g_V^e}^2 + {g_A^e}^2)\eta_Z g_4^Z\right] \\
&\quad - xy^2\left[-g_V^e \eta_{\gamma Z} g_5^{\gamma Z} + ({g_V^e}^2 + {g_A^e}^2)\eta_Z g_5^Z\right]\Big\}\ , \\
\frac{d^2\sigma_{eH}}{dxdy} &= \frac{4\pi\alpha^2}{xyQ^2}\Big\{(2-y)\,xy\left[g_1^\gamma - g_V^e \eta_{\gamma Z} g_1^{\gamma Z} + ({g_V^e}^2 + {g_A^e}^2)\eta_Z g_1^Z\right] \\
&\quad - (1-y)\left[g_A^e \eta_{\gamma Z} g_4^{\gamma Z} - 2 g_V^e g_A^e \eta_Z g_4^Z\right] + xy^2\left[g_A^e \eta_{\gamma Z} g_5^{\gamma Z} - 2 g_V^e g_A^e \eta_Z g_5^Z\right]\Big\}\ .
\end{aligned} \tag{25}$$

And if also including SMEFT contributions:

$$\frac{d^2\sigma_0^{\rm SMEFT}}{dxdy} = \frac{4\pi\alpha^2}{xyQ^2}\left[(1-y)(c_V^e \xi_{\gamma r} F_2^{\gamma r} + (c_V^e g_V^e + c_A^e g_A^e)\xi_{Zr} F_2^{Zr}\right)$$

$$
\begin{aligned}
&+xy^2(c_V^e\xi_{\gamma r}F_1^{\gamma r}+(c_V^e g_V^e+c_A^e g_A^e)\xi_{Zr}F_1^{Zr})\\
&+\frac{xy}{2}(2-y)(c_A^e\xi_{\gamma r}F_3^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}F_3^{Zr})\Big]\\
\frac{d^2\sigma_e^{\rm SMEFT}}{dxdy}&=-\frac{4\pi\alpha^2}{xyQ^2}\left[(1-y)(c_A^e\xi_{\gamma r}F_2^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}F_2^{Zr})\right.\\
&+xy^2(c_a^e\xi_{\gamma r}F_1^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}F_1^{Zr})\\
&\left.+\frac{xy}{2}(2-y)(c_V^e\xi_{\gamma r}F_3^{\gamma r}+(c_A^e g_A^e+c_V^e g_V^e)\xi_{Zr}F_3^{Zr})\right]\\
\frac{d^2\sigma_H^{\rm SMEFT}}{dxdy}&=-\frac{4\pi\alpha^2}{xyQ^2}\sum_r\left[\frac{xy}{2}(2-y)(c_A^e\xi_{\gamma r}g_1^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}g_1^{Zr})\right.\\
&+(1-y)(c_V^e\xi_{\gamma r}g_4^{\gamma r}+(c_A^e g_A^e+c_V^e g_V^e)\xi_{Zr}g_4^{Zr})\\
&\left.+xy^2(c_V^e\xi_{\gamma r}g_5^{\gamma r}+(c_A^e g_A^e+c_V^e g_V^e)\xi_{Zr}g_5^{Zr})\right]\\
\frac{d^2\sigma_{eH}^{\rm SMEFT}}{dxdy}&=\frac{4\pi\alpha^2}{xyQ^2}\left[\frac{xy}{2}(2-y)(c_V^e\xi_{\gamma r}g_1^{\gamma r}+(c_A^e g_A^e+c_V^e g_V^e)\xi_{Zr}g_1^{Zr})\right.\\
&+(1-y)(c_A^e\xi_{\gamma r}g_4^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}g_4^{Zr})\\
&\left.+xy^2(c_A^e\xi_{\gamma r}g_5^{\gamma r}+(c_V^e g_A^e+c_A^e g_V^e)\xi_{Zr}g_5^{Zr})\right]\ . \qquad (26)
\end{aligned}
$$

If a positron beam becomes available at the EIC, one can measure cross sections of both e^+H and e^-H cross sections and study the differences. Again neglecting target mass terms and writing SM and SMEFT contributions all together, we have:

$$
\begin{aligned}
\frac{d^2\sigma_0^{e^+}}{dxdy}-\frac{d^2\sigma_0^{e^-}}{dxdy}&=\frac{4\pi\alpha^2}{xyQ^2}g_A^e\eta_{\gamma Z}\left[xy(2-y)\left(F_3^{\gamma Z}-2g_V^e\eta_{\gamma Z}F_3^Z\right)\right]\\
&-\frac{8\pi\alpha^2}{xyQ^2}\left[(1-y)c_A^e g_A^e\xi_{Zr}F_2^{Zr}+xy^2c_A^e g_A^e\xi_{Zr}F_1^Z\right.\\
&\left.+\frac{xy}{2}(2-y)(c_A^e(\xi_{\gamma r}F_3^{\gamma r}+g_V^e F_3^{Zr}\xi_{Zr}))\right]\\
\frac{d^2\sigma_e^{e^+}}{dxdy}-\frac{d^2\sigma_e^{e^-}}{dxdy}&=\frac{4\pi\alpha^2}{xyQ^2}g_A^e\eta_{\gamma Z}\left[2(1-y)\left(-F_2^{\gamma Z}+2g_V^e\eta_{\gamma Z}F_2^Z\right)-2xy^2\left(F_1^{\gamma Z}-2g_V^e\eta_{\gamma Z}F_1^Z\right)\right]\\
&+\frac{8\pi\alpha^2}{xyQ^2}\left[(1-y)c_A^e(\xi_{\gamma r}F_2^{\gamma r}+g_V^e\xi_{Zr}F_2^{Zr})+xy^2c_A^e(\xi_{\gamma r}F_1^{\gamma r}+g_V^e\xi_{Zr}F_1^{Zr})\right.\\
&\left.+\frac{xy}{2}(2-y)(c_A^e g_A^e\xi_{Zr}F_3^{Zr})\right]\\
\frac{d^2\sigma_H^{e^+}}{dxdy}-\frac{d^2\sigma_H^{e^-}}{dxdy}&=\frac{4\pi\alpha^2}{xyQ^2}g_A^e\eta_{\gamma Z}\left[2y\,(2-y)\left(-g_1^{\gamma Z}+2g_V^e\eta_{\gamma Z}g_1^Z\right)\right]\\
&+\frac{8\pi\alpha^2}{xyQ^2}\left[\frac{xy}{2}(2-y)c_A^e(g_1^{\gamma r}\xi_{\gamma r}+g_V^e g_1^{Zr}\xi_{Zr})\right.\\
&\left.+(1-y)c_A^e g_A^e\xi_{Zr}g_4^{Zr}+xy^2c_A^e g_A^e\xi_{Zr}g_5^{Zr}\right]\\
\frac{d^2\sigma_{eH}^{e^+}}{dxdy}-\frac{d^2\sigma_{eH}^{e^-}}{dxdy}&=\frac{4\pi\alpha^2}{xyQ^2}g_A^e\eta_{\gamma Z}\left[2(1-y)\left(-g_4^{\gamma Z}+2g_V^e\eta_{\gamma Z}g_4^Z\right)-2xy^2\left(\eta_{\gamma Z}g_5^{\gamma Z}-2g_V^e\eta_{\gamma Z}g_5^Z\right)\right]
\end{aligned}
$$

$$+\frac{8\pi\alpha^2}{xyQ^2}\Big[\frac{xy}{2}(2-y)c_A^e g_A^e \xi_{Zr} g_1^{Zr} + (1-y)c_A^e(\xi_{\gamma r} g_4^{\gamma r} + g_V^e \xi_{Zr} g_4^{Zr}) + xy^2 c_A^e g_A^e (\xi_{\gamma r} g_5^{\gamma r} + g_V^e \xi_{Zr} g_5^{Zr})\Big] \quad (27)$$

In this study, we focus on measurements of both parity-violating and lepton-charge asymmetries. The parity-violating asymmetry can be formed either by comparing right-handed and left-handed electron scattering from unpolarized hadrons, referred to as "unpolarized PV asymmetry":

$$A_{PV}^{(e)} \equiv \frac{d\sigma_e}{d\sigma_0}, \quad (28)$$

or by comparing unpolarized electron scattering off right-handed and left-handed hadrons, referred to as "polarized PV asymmetry":

$$A_{PV}^{(H)} \equiv \frac{d\sigma_H}{d\sigma_0} \ . \quad (29)$$

If a positron beam becomes available in the future, the lepton-charge asymmetry, defined as the unpolarized DIS cross section asymmetry between electron and positron beams:

$$A_{\mathrm{LC,H}} = \frac{d\sigma_0^{e^+} - d\sigma_0^{e^-}}{d\sigma_0^{e^+} + d\sigma_0^{e^-}} \ , \quad (30)$$

will provide additional constraints on SMEFT interactions. On the other hand, the double-spin asymmetry, $A^{(eH)} \equiv \frac{d\sigma_{eH}}{d\sigma_0}$, is the primary observable to study the nucleon spin structure but is not within the scope of this work. Similarly, a complete list of lepton-charge asymmetries that includes lepton polarization dependence can be found in [2], but they provide similar constraints to SM and SMEFT studies as the unpolarized asymmetry defined in Eq. 30 and are not discussed in this work.

B. Measurement of Parity-Violating Asymmetry at the EIC

In DIS experiments utilizing an electron beam of polarization P_e and a hadron beam of polarization P_H, the measured differential cross section is

$$d\sigma = d\sigma_0 + P_e\, d\sigma_e + P_H\, d\sigma_H + P_e P_H\, d\sigma_{eH}, \quad (31)$$

where P_e and P_H have the same sign as the respective beam helicity, λ_e and λ_H, and can take the values $-1 \leqslant P_e, P_H \leqslant 1$. The various cross section components in Eq. (31) are given in Eqs. (25).

The PVDIS asymmetry can be formed by flipping the spin direction of either the electron beam or the ion beam. For the EIC, beams of opposite polarizations will be injected into the storage rings alternately, and thus each of the signs of both electron and ion polarizations is flipped periodically on a short time scale. This is in contrast to HERA, where data were taken with positive, then negative electron polarization, with such long time intervals in between that runs with opposite electron polarizations are essentially two independent experiments.

We express the measured DIS event counts during a certain beam helicity state as

$$N^{++} = a_{det} L^{++} \left(d\sigma_0 + |P_e^{++}| d\sigma_e + |P_H^{++}| d\sigma_H + |P_e^{++}||P_H^{++}| d\sigma_{eH} \right) \tag{32}$$

$$N^{+-} = a_{det} L^{+-} \left(d\sigma_0 + |P_e^{+-}| d\sigma_e - |P_H^{+-}| d\sigma_H - |P_e^{+-}||P_H^{+-}| d\sigma_{eH} \right) \tag{33}$$

$$N^{-+} = a_{det} L^{-+} \left(d\sigma_0 - |P_e^{-+}| d\sigma_e + |P_H^{-+}| d\sigma_H - |P_e^{-+}||P_H^{-+}| d\sigma_{eH} \right) \tag{34}$$

$$N^{--} = a_{det} L^{--} \left(d\sigma_0 - |P_e^{--}| d\sigma_e - |P_H^{--}| d\sigma_H + |P_e^{--}||P_H^{--}| d\sigma_{eH} \right) \ , \tag{35}$$

where $ij = ++, +-, -+, --$ represents the electron and the proton helicity states with their time sequence depending on the helicity pattern of the beam injection, L^{ij} stands for the integrated luminosity, and P_e^{ij} and P_H^{ij} are the electron and the proton (or ion) beam polarizations during the corresponding helicity bunch ij. The a_{det} factor represents the detector phase space, acceptance and efficiency. In the simplest case, if we assume both beam polarizations, the luminosity, and detector efficiency and acceptance do not vary with time, then

$$d\sigma_0 = \frac{1}{4} \left(d\sigma^{++} + d\sigma^{+-} + d\sigma^{-+} + d\sigma^{--} \right) \ , \tag{36}$$

$$d\sigma_e = \frac{1}{4|P_e|} \left(d\sigma^{++} + d\sigma^{+-} - d\sigma^{-+} - d\sigma^{--} \right) \ , \tag{37}$$

$$d\sigma_H = \frac{1}{4|P_H|} \left(d\sigma^{++} - d\sigma^{+-} + d\sigma^{-+} - d\sigma^{--} \right) \ , \tag{38}$$

$$d\sigma_{eH} = \frac{1}{4|P_e||P_H|} \left(d\sigma^{++} - d\sigma^{+-} - d\sigma^{-+} + d\sigma^{--} \right) \ , \tag{39}$$

where we have defined the experimentally measured cross section $d\sigma^{ij} \equiv N^{ij}/L^{ij}/a_{det}$. The PVDIS asymmetry due to electron spin flip can be extracted from data by taking the ratio of the cross sections. Because spin flips of both electron and hadron beams will be carried out at very short time scale, the a_{det} term can be assumed as a constant and cancels out when forming the asymmetry, and we can extract the asymmetry from experimentally measured

yields, defined as $Y^{ij} \equiv N^{ij}/L^{ij}$:

$$A_{PV}^{(e)} \equiv \frac{d\sigma_e}{d\sigma_0} = \frac{1}{|P_e|}\frac{Y^{++} + Y^{+-} - Y^{-+} - Y^{--}}{Y^{++} + Y^{+-} + Y^{-+} + Y^{--}}, \tag{40}$$

and that due to proton (ion) spin flip can be similarly extracted as

$$A_{PV}^{(H)} \equiv \frac{d\sigma_H}{d\sigma_0} = \frac{1}{|P_H|}\frac{Y^{++} - Y^{+-} + Y^{-+} - Y^{--}}{Y^{++} + Y^{+-} + Y^{-+} + Y^{--}} \ . \tag{41}$$

The design of the EIC requires that the point-to-point luminosity uncertainty to be at 10^{-4} level. Therefore, the dominant experimental uncertainty would come from electron and proton (ion) polarimetry, for $A_{PV}^{(e)}$ and $A_{PV}^{(H)}$, respectively.

C. Measurement of Lepton-Charge Asymmetry at the EIC

Unlike PV asymmetries which can be formed by comparing scattering yields of right-handed vs left-handed electron or hadron scattering on a short time scale, measurement of the LC asymmetry requires comparison between electron runs and positron runs, and thus relies on two independent cross section measurements. To reduce the uncertainty in the measurement of $A_{LC,H}$, we can reverse the polarity of the magnet to minimize the systematic uncertainty due to differences in e^- and e^+ detection. In this case, the main experimental systematic uncertainty will come from the luminosity difference between e^- and e^+ running, which is assumed to be 2% (relative in luminosity, absolute in $A_{LC,H}$) in this analysis.

III. PROJECTION OF PARITY-VIOLATION AND LEPTON-CHARGE ASYMMETRY DATA

A. EIC Simulation with ECCE Detector Configuration

We used `Djangoh` event generator [3] (version 4.6.16 [4]) that includes full electromagnetic and electroweak radiative effects to generate 20-million (20 M) Monte-Carlo (MC) events for each of the four beam energy and two beam type combinations: $18\times275(137)$, $10\times275(137)$, 10×100, and 5×100 GeV for ep (eD) collisions, respectively. For the deuterium ion beam, the energy specified is per nucleon. In lieu of a full GEANT-based simulation, a fast smearing method was applied to inclusive electron events in the `Djangoh` output, and the

physics cross section and parity-violating asymmetries were calculated event-by-event using a modified user routine of `Djangoh`. The number of the scattered DIS electrons was then calculated using the cross section information and the expected integrated luminosity after correcting for bin migration.

The detector fast smearing was obtained from a single-electron gun simulation. Resolution spectra were determined for 57 evenly-spaced bins for the pseudo-rapidity range $\eta = (-3.5625, 3.5625)$ and 1 GeV-wide bins in the transverse momentum p_T. For each `Djangoh`-simulated event, smearing in the electron momentum p and polar and azimuthal angles θ and ϕ were randomly picked from the corresponding spectrum and applied to the event, which were used to determine the detected kinematics of the event. While the smearing spectra were not exactly Gaussian-shaped, they were fitted with a Gaussian function and the fitted RMS values extracted for illustration purposes, see Fig. 2.

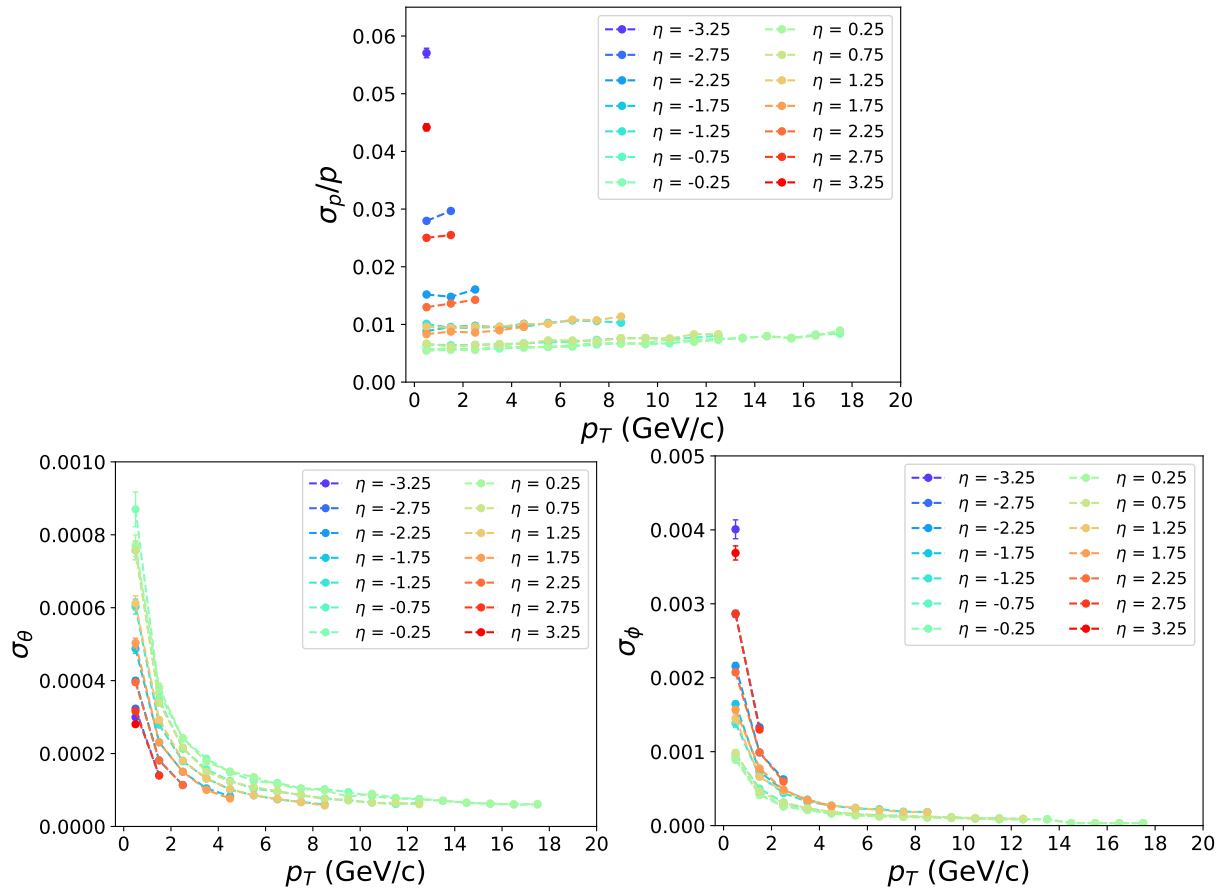

FIG. 2. RMS values for fast-smearing spectra obtained from single electron-gun simulation of July 2021 concept of ECCE. The unit for σ_θ and σ_ϕ are radians.

Using the fast-smearing method, we generated 10 M MC events to study the kinematic coverage over the full phase space. An additional 10 M events were generated with $Q^2_{\rm min} = 50$ GeV^2 for where DIS events have the most impact on the extraction of the weak mixing angle. The drawback of the fast smearing method is that no selection of the hadronic state was implemented. Methods utilizing hadronic final states such as the double-angle method may provide better DIS event identification for certain kinematic range and thus improve analysis precision.

Bin migration of inclusive scattering electrons due to internal and external radiative effects were studied with fast-smearing simulation and treated using the "R matrix" unfolding method [5]. Background reactions were studied using the hadronic final state generated by `Djangoh` (with $Q^2_{\rm min} = 1.0$ GeV^2), and another Monte-Carlo simulation of photoproduction events generated by `Pythia` (version 6.428 with $Q^2_{\rm min} = 0$). All events were passed through the ECCE full simulation. We found the highest background events to occur at high y values, and these events were rejected at the event selection stage, see next section.

B. Event Selection

For the 20 M fast-smearing events, event selection criteria were applied to choose DIS events ($Q^2_{\rm det} > 1.0$ GeV^2), to avoid regions with severe bin migration and unfolding uncertainty ($y_{\rm det} > 0.1$), to avoid regions with high photoproduction background ($y_{\rm det} < 0.90$), to restrict events in the main acceptance of the ECCE detector where the fast-smearing method is applicable ($\eta_{\rm det} > -3.5$ and $\eta_{\rm det} < 3.5625$), and to ensure high purity of electron samples ($E' > 2.0$ GeV). Here the subscript "det" implies the variables were calculated using the detected information of the electron. The projected values and statistical uncertainty for $A^{(e)}_{PV}$ and $A^{(H)}_{PV}$ after unfolding are shown respectively in Fig. 3 and 4 for 18×275 GeV ep for an integrated uncertainty of 100 fb^{-1}.

C. Integrated Luminosity

To account for realistic running conditions, the annual luminosity – ten times the "high divergence configuration" value as shown in Table 10.1 of the Yellow Report (YR) [6] – were used. More specifically, the integrated luminosity values are assumed to be 15.4, 100, 44.8,

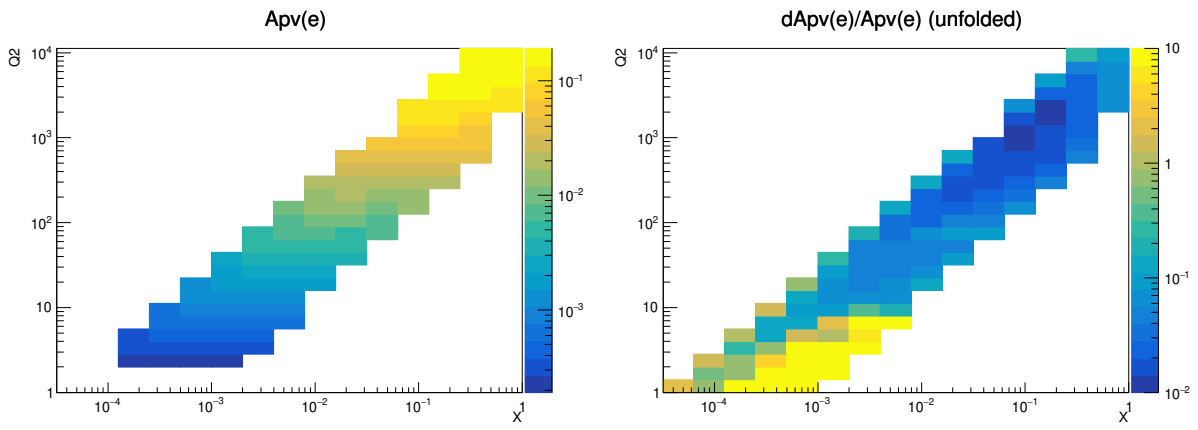

FIG. 3. Projection for $A_{PV}^{(e)}$ (left), and $dA_{PV,\text{stat}}^{(e)}/A_{PV}^{(e)}$ after unfolding (right) for 18×275 GeV ep collision, with event selections criteria applied. An integrated luminosity of 100 fb^{-1} and an 80% electron polarization were assumed.

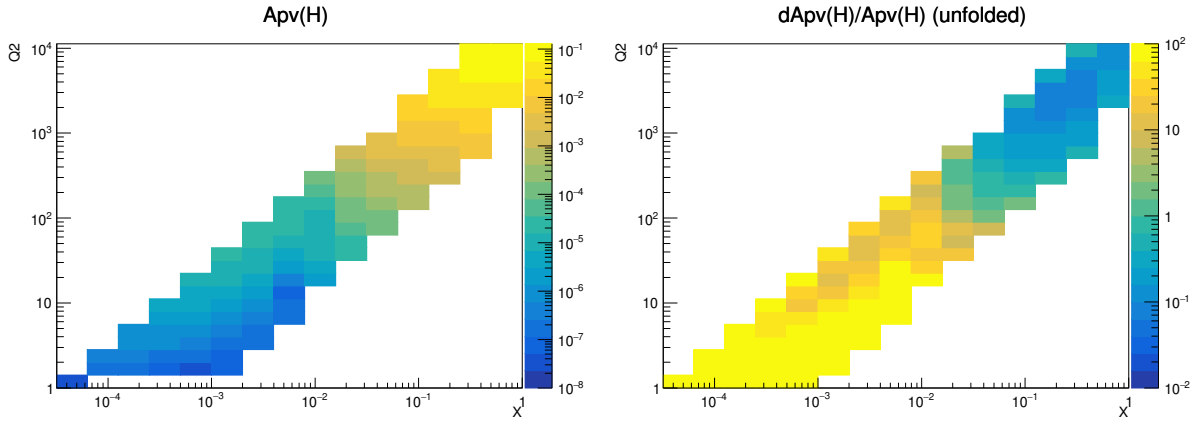

FIG. 4. Projection for $A_{PV}^{(p)}$ (left), and $dA_{PV,\text{stat}}^{(p)}/A_{PV}^{(p)}$ after unfolding (right) for 18×275 GeV ep collision, with event selections criteria applied. An integrated luminosity of 100 fb^{-1} and an 70% proton polarization were assumed.

36.8, and 4.4 fb^{-1} for $18 \times 275(137)$, $10 \times 275(137)$, 10×100, 5×100, and 5×41 GeV ep (eD) collisions, respectively. As a comparison with the weak mixing angle extraction presented in the YR, we also carried out projections for 100 fb^{-1} 18×275 GeV ep and 10 fb^{-1} 18×137 GeV eD collision as the "YR reference point". We abbreviate the ep pseudo data sets as P1, P2, P3, P4, P5 and the eD pseudo-data sets as D1, D2, D3, D4, D5, see Table II. The YR reference point is denoted P6. Simulated pseudo-data sets with polarized hadrons are indicated as ΔD1–5 and ΔP1–6, while positron data sets are referred

385 to as LD1–5 and LP1–6 (with "L" for Lepton charge).

D1	5 GeV × 41 GeV eD, 4.4 fb^{-1}	P1	5 GeV × 41 GeV ep, 4.4 fb^{-1}
D2	5 GeV × 100 GeV eD, 36.8 fb^{-1}	P2	5 GeV × 100 GeV ep, 36.8 fb^{-1}
D3	10 GeV × 100 GeV eD, 44.8 fb^{-1}	P3	10 GeV × 100 GeV ep, 44.8 fb^{-1}
D4	10 GeV × 137 GeV eD, 100 fb^{-1}	P4	10 GeV × 275 GeV ep, 100 fb^{-1}
D5	18 GeV × 137 GeV eD, 15.4 fb^{-1}	P5	18 GeV × 275 GeV ep, 15.4 fb^{-1}
		P6	18 GeV × 275 GeV ep, 100 fb^{-1}

TABLE II. Energy and luminosity configurations assumed for the EIC in our analysis. P6 is the YR reference setting.

386 D. Statistical Uncertainty Projection for PV Asymmetries

387 For a given value of integrated luminosity, the statistical uncertainty of an asymmetry
388 measurement is

$$dA_{\text{stat,measured}} = \frac{1}{\sqrt{N}} \tag{42}$$

389 where N is the total number of events detected, assumed to be approximately equally divided
390 between the two scattering types – either between left- and right-handed electron beam, or
391 between left- and right-handed proton (ion) beam, or between positron and electron runs.
392 The unfolding process increases the statistical precision only slightly for the region where
393 the relative statistical uncertainty on the asymmetry is the most precise.

394 If the asymmetry originates from polarization (as for the case of PV asymmetries), we
395 must correct for the beam polarization:

$$dA^{(e)}_{\text{stat,PV}} = \frac{1}{|P_e|}\frac{1}{\sqrt{N}} \ , \quad \text{and} \quad dA^{(H)}_{\text{stat,PV}} = \frac{1}{|P_H|}\frac{1}{\sqrt{N}} \ . \tag{43}$$

396 For $A^{(e)}_{PV}$ projection, an electron beam polarization of $P_e = 80\%$ with relative 1% systematic
397 uncertainty from the electron polarimetry were assumed. Similarly, for $A^{(H)}_{PV}$ projection,
398 a proton (ion) beam polarization of $P_H = 70\%$ with relative 2% systematic uncertainty
399 from the proton (ion) polarimetry were used. An illustration of the relative precision of PV
400 asymmetries is provided in Figs. 3 and 4. The statistical uncertainty of $A^{(H)}_{PV}$ is rather large
401 because of the much smaller size of $A^{(H)}_{PV}$ than $A^{(e)}_{PV}$.

E. Statistical and QED Uncertainty Projection for Lepton-Charge Asymmetries

As described in Section II C, to measure lepton-charge asymmetry $A_{LC,H}$, one can reverse the polarity of the magnet to minimize the systematic uncertainty due to differences in e^- and e^+ detection. In this case, the main experimental systematic uncertainty will come from the luminosity difference between e^- and e^+ running, which is assumed to be 2% (relative in luminosity, absolute in A_{LC}) in this analysis. If the detector magnet polarity is reversed, then the detection of DIS positrons would be very similar to that of DIS electrons, and all data simulation, event selection, unfolding, etc, described in Section III A applies. The statistical uncertainty in A_{LC} is thus determined by the luminosity of e^+ running, which we assume to be one-tenth that of the electron beam. Note that beam polarization and thus polarimetry uncertainties do not affect A_{LC} measurements.

The EW physics reach of A_{LC} is further clouded by the difference in e^- vs. e^+ DIS cross sections due to QED higher order effects. We calculated the value of A_{LC} using `Djangoh` version 4.6.19 in both the Born LO (that includes one-boson exchange only) and NLO radiated mode (that includes higher order EW and QED effects), see Fig. 5. The difference of NLO −(minus) Born is taken as an estimate of QED NLO effects and the uncertainty is assumed to be a relative 5%.

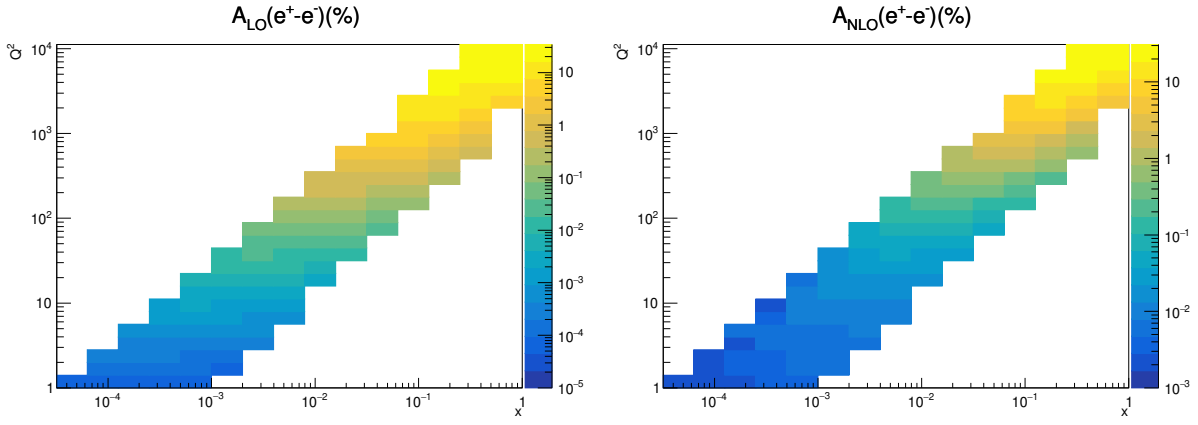

FIG. 5. Calculation for A_{LC} at the Born (LO) (left) and NLO (right) level for e^+p vs. e^-p collision at 18×275 GeV. The LO calculation includes only γZ interference term which is of main interest of this study. The NLO calculation includes box-diagrams which introduces a large QED effect to the asymmetry and is effectively a background to the EW and SMEFT study presented here.

Because of the moderate Q^2 reach of the EIC, the 2% absolution uncertainty from lu-

minosity measurement is a dominating systematic effect for the uncertainty of $A_{LC,H}$. In Appendix A 1, we present a method to simultaneously fit the luminosity term with SMEFT coefficients; however, we found this method yields 15 to 20% weaker SMEFT constraints.

F. Projection for High-Luminosity EIC

In addition to the nominal luminosity expected for the EIC, we also carried out projections considering the possibility of an additional factor ten increase in the luminosity, the so-called high-luminosity EIC (HL-EIC). Assuming all experimental systematic effects remain the same, we scale the projected statistical uncertainty of asymmetry observables described in the previous section by the factor $1/\sqrt{10}$. For beam energies with lower luminosity (hence larger statistical uncertainty) or asymmetries of smaller sizes such as $A_{PV}^{(H)}$, the factor 10 increase in luminosity will push the physics reach one step further. On the other hand, for beam energies with already-high luminosity and observables where systematic effects dominate over statistical ones, such as $A_{PV}^{(e)}$ for 10×275 GeV ep and 10×137 GeV eD collisions and $A_{LC,H}$, the impact from the luminosity increase of HL-EIC on the physics reach is marginal.

IV. PSEUDO DATA GENERATION AND UNCERTAINTY MATRIX

A. Pseudo Data for Parity-Violating Asymmetries

We discuss first the two PV asymmetry cases: the polarized electron asymmetries with unpolarized hadrons $A_{PV}^{(e)}$ and polarized hadron asymmetries with unpolarized electron $A_{PV}^{(H)}$. The uncertainties are from three sources: statistical $\sigma_{\rm stat}$, experimental systematic $\sigma_{\rm syst}$ (mainly due to particle background and also includes other imperfection of the measurement) that is assumed to be fully uncorrelated, and beam polarimetry $\sigma_{\rm pol}$ that is assumed to be fully correlated within data of the same $\sqrt{s}$ and beam type. For the $b^{\rm th}$ bin, with given $\sqrt{s}$, x, and Q^2 values and using the nominal PDF set under consideration, we first compute the theoretical SM prediction, $(A_{\rm PV})^{\rm theo}_{{\rm SM},0,b}$. Combining the given uncertainties in quadrature separately for uncorrelated and correlated ones, we obtain a pseudo-experimental asymmetry

value by

$$(A_{\rm PV})_b^{\rm pseudo} = (A_{\rm PV})_{{\rm SM},0,b}^{\rm theo} + r_b \sqrt{\sigma_{{\rm stat},b}^2 + \left[(A_{\rm PV})_{{\rm SM},0,b}^{\rm theo} \left(\frac{\sigma_{\rm sys}}{A}\right)_b\right]^2} + r' \sqrt{\left[(A_{\rm PV})_{{\rm SM},0,b}^{\rm theo} \left(\frac{\sigma_{\rm pol}}{A}\right)_b\right]^2} \tag{44}$$

where r_b and r' are random numbers chosen from a normal distribution of mean 0 and standard deviation 1. Note that the correlated errors are incorporated using a single random number (r') across all the bins. The systematic uncertainties are $\sigma_{\rm syst}/A = 1\%$, $\sigma_{\rm pol}/A = 1\%$ for $A_{PV}^{(e)}$, and $\sigma_{\rm pol}/A = 2\%$ for $A_{PV}^{(e)}$.

B. Pseudo Data for Lepton-Charge Asymmetries

We next consider unpolarized electron-positron asymmetries with unpolarized hadrons, the lepton-charge (LC) asymmetries. The uncertainties are from three sources: statistical $\sigma_{\rm stat}$, experimental systematic $\sigma_{\rm syst}$ mailing due to background that is assumed to be fully uncorrelated, luminosity difference between e^+ and e^- runs $\sigma_{\rm lumi}$ that is fully correlated within data of the same $\sqrt{s}$ and ion beam type, and QED higher-order effect $\sigma_{\rm QED\ NLO}$, taken as 5% of the A_{LC} difference between calculated NLO and Born (LO) values. In analogy with Eq. (44), for the LC asymmetries, we write

$$(A_{\rm LC})_b^{\rm pseudo} = (A_{\rm LC})_{{\rm SM},0,b}^{\rm theo} + r_b \sqrt{\sigma_{\rm stat}^2 + \left[(A_{\rm LC})_{{\rm SM},0,b} \left(\frac{\sigma_{\rm sys}}{A}\right)_b\right]^2 + \sigma_{{\rm QED\ NLO},b}^2} + r' \sigma_{{\rm lumi},b} \tag{45}$$

C. Uncertainty Matrix

The uncertainty matrix, Σ^2, for a given data set with $N_{\rm bin}$ bins is an $N_{\rm bin} \times N_{\rm bin}$ symmetric matrix. It consists of two parts, which we call Σ_0^2 and $\Sigma_{\rm pdf}^2$:

$$\left(\Sigma^2\right)_{bb'} = \left(\Sigma_0^2\right)_{bb'} + \left(\Sigma_{\rm pdf}^2\right)_{bb'} . \tag{46}$$

The first part of the matrix, Σ_0^2, is constructed using all the uncertainty components (statistical, systematic, polarimetry or luminosity, QED) other than the PDF uncertainties. All uncertainties that enter Σ_0^2 must be absolute; relative uncertainties are converted to absolute ones by multiplying the theoretical SM prediction, $A_{{\rm SM},0,b}^{\rm theo}$, computed using the

central member of the PDF set taken into account. The first part of the matrix then takes the form

$$\Sigma_0^2 = \begin{pmatrix} \sigma_1^2 & \rho_{12}\tilde{\sigma}_1\tilde{\sigma}_2 & \cdots & \rho_{1N_{\rm bin}}\tilde{\sigma}_1\tilde{\sigma}_{N_{\rm bin}} \\ & \sigma_2^2 & \cdots & \rho_{2N_{\rm bin}}\tilde{\sigma}_2\tilde{\sigma}_{N_{\rm bin}} \\ & & \ddots & \vdots \\ & & & \sigma_{N_{\rm bin}}^2 \end{pmatrix}_{\rm sym} \tag{47}$$

where, for the PV asymmetries, we have for the diagonal elements

$$\sigma_b^2 = \sigma_{{\rm stat},b}^2 + \left[(A_{\rm PV})_{{\rm SM},0,b}^{\rm theo}\left(\frac{\sigma_{\rm sys}}{A}\right)_b\right]^2 + \left[(A_{\rm PV})_{{\rm SM},0,b}^{\rm theo}\left(\frac{\sigma_{\rm pol}}{A}\right)_b\right]^2 \tag{48}$$

and for the off-diagonal elements

$$\tilde{\sigma}_b = (A_{\rm PV})_{{\rm SM},0,b}^{\rm theo}\left(\frac{\sigma_{\rm pol}}{A}\right)_b . \tag{49}$$

For the LC asymmetries, we have for the diagonal elements

$$\sigma_b^2 = \sigma_{{\rm stat},b}^2 + \left[(A_{\rm LC})_{{\rm SM},0,b}^{\rm theo}\left(\frac{\sigma_{\rm sys}}{A}\right)_b\right]^2 + \sigma_{{\rm lumi},b}^2 + \sigma_{{\rm QEDNLO},b}^2 \tag{50}$$

and for the off-diagonal elements

$$\tilde{\sigma}_b = \sigma_{{\rm lumi},b}. \tag{51}$$

Here, b and b' are bin numbers and we assume full correlation for uncertainties originating from beam polarimetry or luminosity: $\rho_{bb'} = 1$ for all b and b'.

The second part of the uncertainty matrix, $\Sigma_{\rm pdf}^2$, is built using the same procedure for both PV and LC asymmetries by taking into account differences between the theoretical SM asymmetry prediction computed at the nominal PDF member, $A_{{\rm SM},0}^{\rm theo}$, and theoretical SM asymmetry predictions evaluated at all other members of the PDF set under consideration, $A_{{\rm SM},m}^{\rm theo}$, where $m = 1, 2 \ldots N_{\rm PDF}$ with $N_{\rm PDF}$ the total number of PDF sets or replicas available. For Hessian-based PDF sets, the diagonal and off-diagonal elements can be collectively written as

$$\left(\Sigma_{\rm PDF}^2\right)_{bb',{\rm Hessian}} = \frac{1}{4}\sum_{m=1}^{N_{\rm PDF}/2}(A_{{\rm SM},2m,b} - A_{{\rm SM},2m-1,b})(A_{{\rm SM},2m,b'} - A_{{\rm SM},2m-1,b'}) \ . \tag{52}$$

And for replica-based PDF sets:

$$\left(\Sigma_{\rm PDF}^2\right)_{bb',{\rm replica}} = \frac{1}{N_{\rm PDF}}\sum_{m=1}^{N_{\rm PDF}}(A_{{\rm SM},m,b}^{\rm theo} - A_{{\rm SM},0,b}^{\rm theo})(A_{{\rm SM},m,b'}^{\rm theo} - A_{{\rm SM},0,b'}^{\rm theo}). \tag{53}$$

D. Comparison of uncertainty components

We present in this section the various uncertainty components that enter the SMEFT analysis. We also investigate the total uncertainties combined in quadrature that contribute to the diagonal entries of the uncertainty matrix.

1. Individual uncertainty components

We begin by considering the individual components of the uncertainties. We investigate the effects of sea quarks in the analysis by defining a valence-only approximation for the PDFs. The tag `ud` in the plot labels is the valence-only approximation where only up and down quark contributions are considered in the hadronic cross section, whereas `uds` indicates that up, down, strange, and their antiquarks are taken into account. Note that for the data sets involving unpolarized deuteron with the `ud` tag, there will be no uncertainty from PDF as deuteron PDFs, defined in terms of proton and neutron PDFs using isospin symmetry, cancel when analytically forming asymmetries in the valence-only approximation. Note that for experimental systematic uncertainties other than those from beam polarimetry, both 1% and 2% are shown in all figures of this section, although the 1% value is used in the results presented.

Fig. 6 shows the comparison of the uncertainty components for the data set D4 in the `ud` and `uds` scenarios. For PDFs, we use NNPDF3.1 NLO [7] in the unpolarized case and NNPDFPOL1.1 [8] in the polarized case throughout. Only (x, Q^2) region relevant for SMEFT analysis is shown, though the full region is used for the extraction of the weak mixing angle. The x-axis of these plots is ordered by bin number; these are ordered first from low to high Q^2, then from small to large x within each Q^2 bin, leading to the observed oscillatory behavior. When we turn on the sea quark contributions, the unpolarized deuteron data sets receive nonzero but highly suppressed PDF uncertainties, indicating that the assumption of deuteron PDFs completely cancelling is a reasonably good approximation. The right panel shows that even after including sea quarks, the PDFs are still the smallest uncertainty component. This indicates that potentially poorly determined sea quark and strange quark distributions have little effect on this analysis. The largest single uncertainty component is the statistical uncertainty (shown as a dark red line). This is larger than both the 1%

beam polarization uncertainty (light blue line), and either of the 1% or 2% uncorrelated systematic uncertainty assumptions (solid and dotted blue lines, respectively). When we switch to the high-luminosity (HL-EIC) scenario (dotted red line), the statistical uncertainty becomes comparable to the systematic ones. All uncertainties are significantly smaller than the predicted value of the asymmetry, shown as the solid black line in the plots.

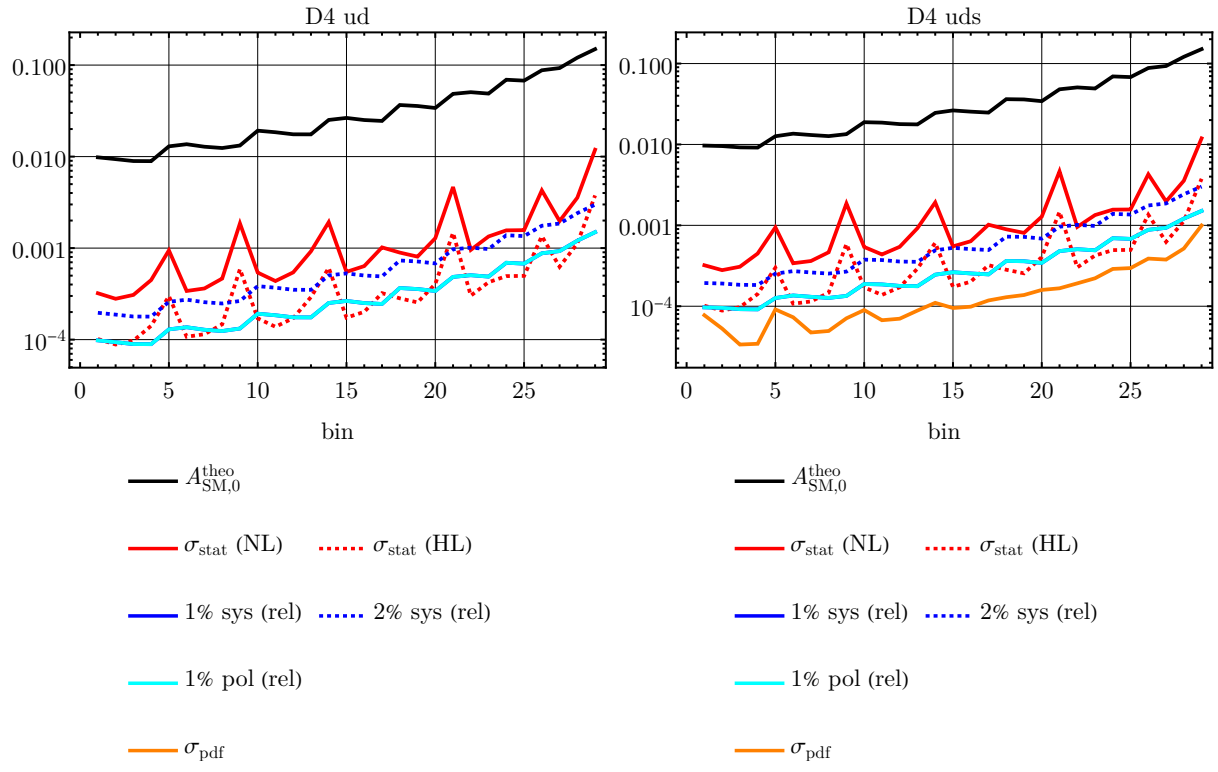

FIG. 6. Comparison of the uncertainty components for the data set D4 in the valence-only scenario (`ud`) and with the contributions from the sea quarks (`uds`). Here "NL" refers to the currently planned annual luminosity of the EIC, while "HL" refers to a possible ten-fold luminosity upgrade.

In Fig. 7, we display the different contributions to the diagonal entries of the uncertainty matrix of the data sets P5 and ΔP5. The pattern of uncertainties for P5 is very similar to that observed for D4. The statistical ones are the largest single uncertainty source, while the PDFs are the smallest. Assuming high luminosity, the statistical uncertainties become comparable to the anticipated systematic ones. The pattern is different for ΔP5: the statistical uncertainties are largest for all bins, even assuming high luminosity. The PDF uncertainties are also non-negligible, consistent with the expectation that spin-dependent PDFs are not known as precisely as spin-independent ones. The anticipated experimental

systematic uncertainties are negligible for all bins.

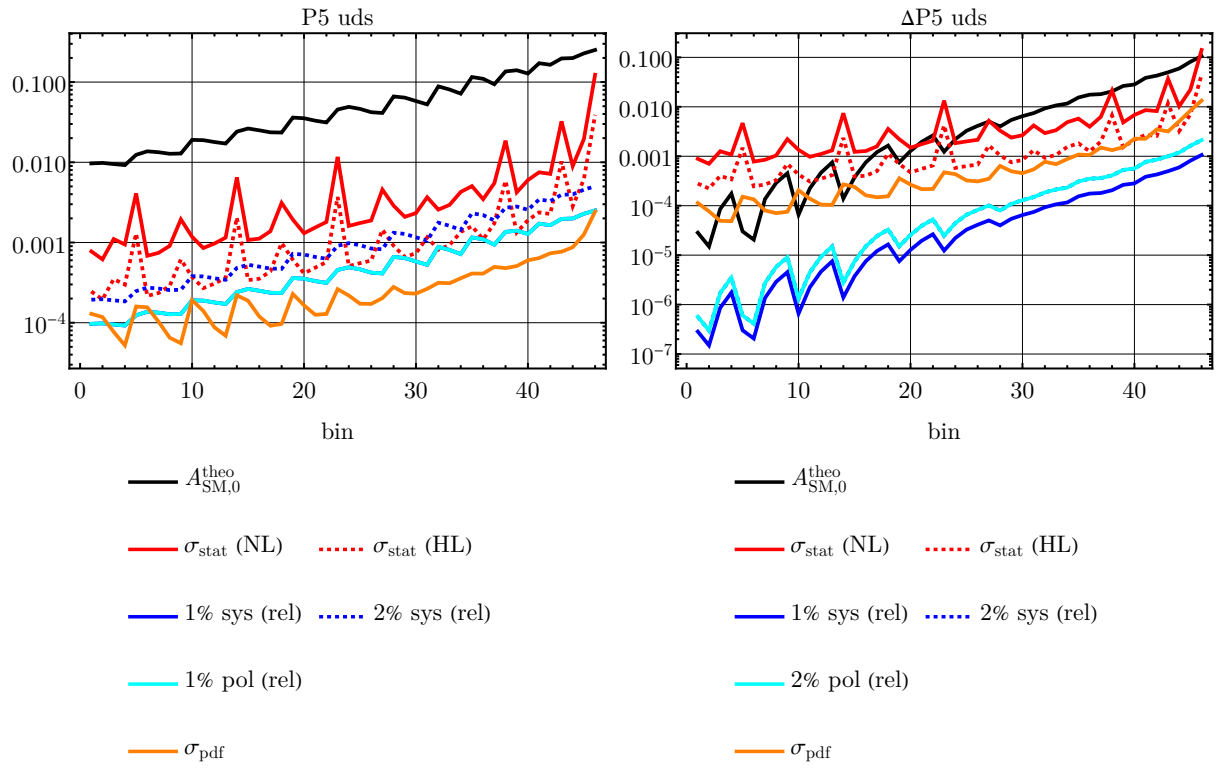

FIG. 7. Uncertainty components for the P5 and ΔP5 data sets.

Finally, we show in Fig. 8 the individual uncertainties for the electron-positron asymmetry data set LP5. The error budget is different for this scenario compared to PV asymmetries. Since both beams are unpolarized, there is no uncertainty related to beam polarization. However, since electron and positron runs occur with different beams there is the possibility of a significant overall luminosity difference between the two runs that can lead to an apparent asymmetry. We assume an absolute 2% uncertainty, two times the luminosity uncertainty requirement of [9]. Finally, we consider the possible errors arising from higher-order QED corrections that may differentiate between electron and positron scattering. We estimate this uncertainty by taking 5% of the difference between the Born-level and NLO QED results, obtained using `Djangoh`. The two largest sources of uncertainty throughout the entire kinematic range are the luminosity and the statistical uncertainties. PDFs, higher-order QCD, and the anticipated systematic uncertainties are all significantly smaller.

Summarizing all figures presented in this section, we can make the following main points:

- The expected statistical uncertainties are the dominant ones for the nominal EIC

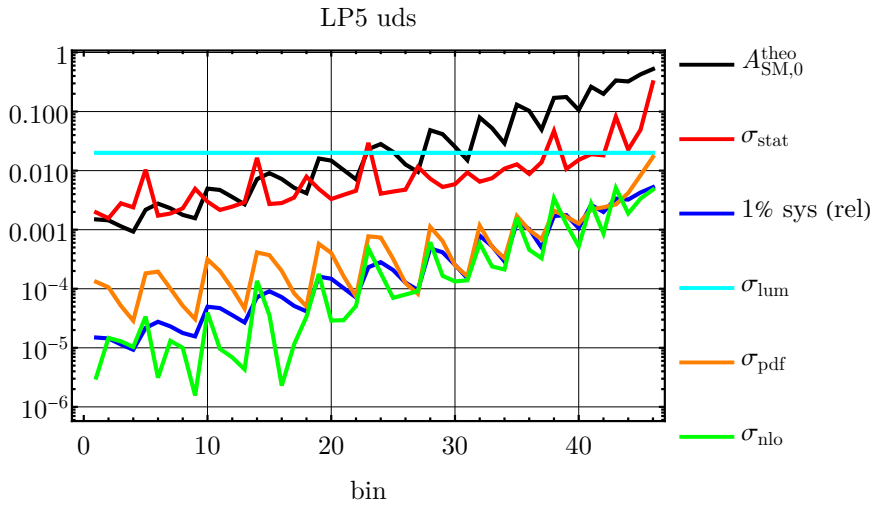

FIG. 8. The same as in Fig. 7 but for LP5.

luminosity. If a high luminosity (HL-EIC) upgrade becomes possible, they become comparable to experimental systematic uncertainties for PV asymmetries of the unpolarized hadron, $A_{PV}^{(e)}$.

- PDF uncertainties are nearly irrelevant for the asymmetries of unpolarized hadrons $A_{PV}^{(e)}$. They become significant, second to statistical uncertainties, for PV asymmtery of polarized hadrons $A_{PV}^{(H)}$.

- The luminosity effect dominates the statistical uncertainty for the majority of the phase space in the case of electron-positron asymmetries $A_{LC,H}$, particularly at low x and low Q^2. On the other hand, uncertainty from higher-order QED corrections are expected to be small.

2. Total uncertainties for nominal luminosity vs. high luminosity

We now investigate the total uncertainties for the data sets D4, ΔD4, P5, ΔP5. We consider four different scenarios: nominal luminosity planned for the EIC or a ten-fold high luminosity upgrade, combined with 1% or 2% relative experimental systematic uncertainties due to particle background. We show the results in Figs. 9 and 10. We first observe that the dominant uncertainty component in all cases is statistical. The four uncertainty scenarios, namely 1% or 2% systematic uncertainty with nominal or high luminosity, can be in fact

reduced to just the luminosity comparison, i.e. nominal vs. high. Next, for both D4 and P5, the asymmetries $A_{PV}^{(e)}$ are measured to percent-level throughout the considered phase space. This is not the case for the polarized sets ΔD4 and ΔP5. Particularly in the ΔP5 scenario, at lower Q^2, the anticipated errors are larger than the asymmetry for all choices of systematic error and luminosity. Only in the very high Q^2 bins does a measurement of the asymmetry $A_{PV}^{(H)}$ become meaningful.

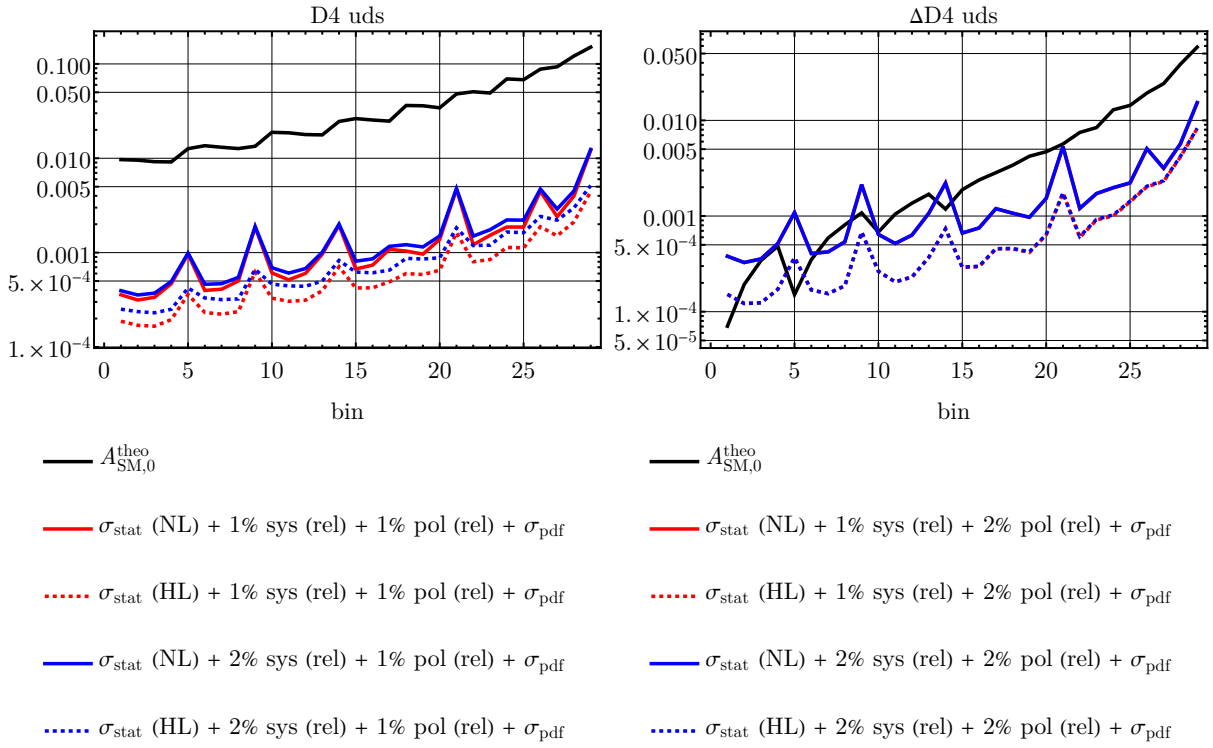

FIG. 9. Total uncertainties combined in quadrature for the data sets D4 and ΔD4 in the `uds` scenario.

Our evaluation of the uncertainties indicate that using 1% or 2% relative systematic uncertainties makes practically no difference, as the total errors are mostly dominated by the statistical uncertainties for the PV asymmetries or the luminosity difference for the LC asymmetries. We also show that one can take into account the contribution of only the valence quarks to the asymmetries or include the sea quarks up to strange flavor and its antiquark, both of which lead to the same size and rank of PDF errors for the data sets under consideration. In our best-fit analyses we thus focus on the data sets with 1% relative systematic uncertainty and nominal luminosity in the `uds` scenario. Comparisons will be

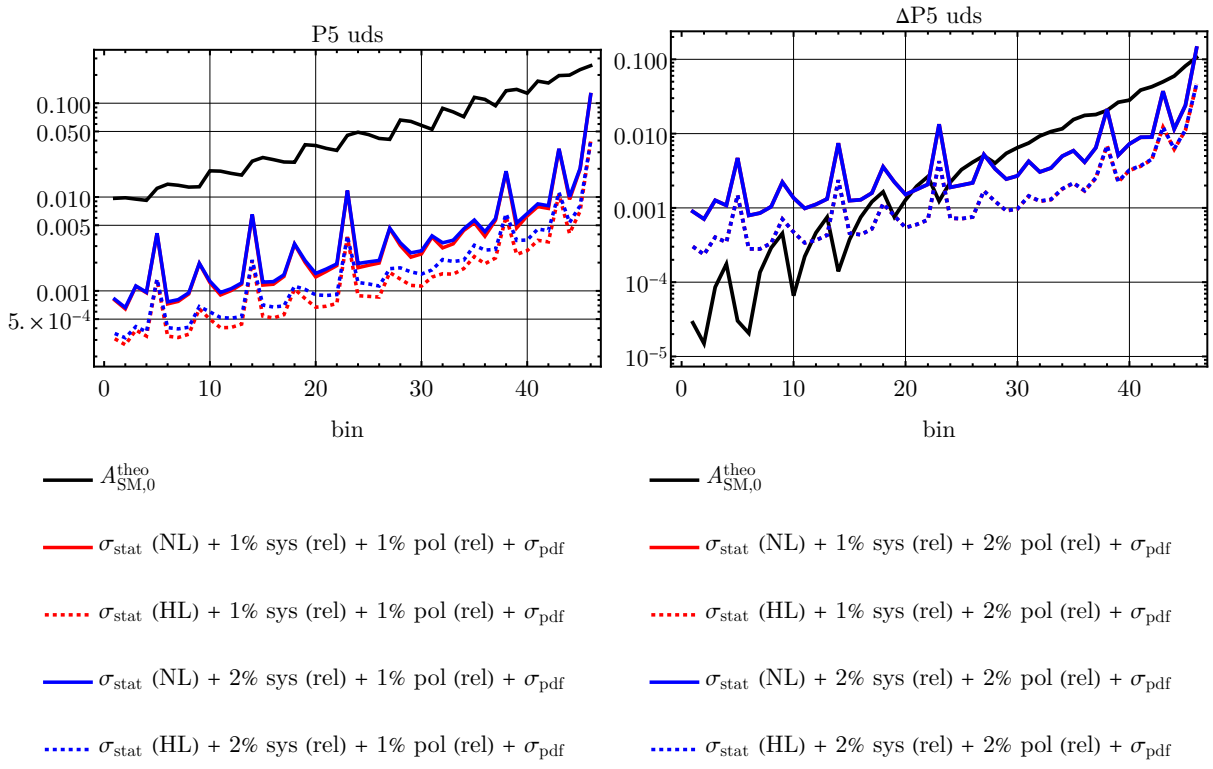

FIG. 10. The same as in Fig. 9 but for P5 and ΔP5.

performed with the ones having high luminosity, keeping the rest of the configuration the same.

V. EXTRACTION OF THE SM WEAK MIXING ANGLE

The weak mixing angle, often represented as $\sin^2\theta_W$, is a fundamental parameter of the SM and has been measured in experiments ranging from atomic parity violation at eV energy levels, to high energy colliders at the Z-pole. The EIC will provide constraints on $\sin^2\theta_W$ in the intermediate energy range that resides between the reach of fixed-target and collider facilities. For the extraction of the weak mixing angle, we focus on $A_{PV}^{(e)}$ where $\sin^2\theta_W$ enters through the electron coupling $g_{V,A}^e$ and quark couplings in the structure functions. If including all target-mass correction terms, we can write

$$A_{PV} = \frac{|P_e|\eta_{\gamma Z}\left[g_A^e 2yF_1^{\gamma Z} + g_A^e\left(\frac{2}{xy}-\frac{2}{x}-\frac{2M^2xy}{Q^2}\right)F_2^{\gamma Z} + g_V^e(2-y)F_3^{\gamma Z}\right]}{2yF_1^{\gamma} + \left(\frac{2}{xy}-\frac{2}{x}-\frac{2M^2xy}{Q^2}\right)F_2^{\gamma} - \eta_{\gamma Z}\left[g_V^e 2yF_1^{\gamma Z} + g_V^e\left(\frac{2}{xy}-\frac{2}{x}-\frac{2M^2xy}{Q^2}\right)F_2^{\gamma Z} + g_A^e(2-y)F_3^{\gamma Z}\right]} \tag{54}$$

where M is the nucleon mass. Note that given the moderate Q^2 values of the EIC, the pure-Z contribution to the structure functions has been omitted for the precision relevant to our analysis.

The running value of $\sin^2\theta_W = 0.231$ was used in the pseudo-data generation and only one pseudo data set was produced. We then fit the value of $\sin^2\theta_W$ by minimizing the χ^2 defined as

$$\chi^2 = [\mathcal{A}^{\text{pseudo-data}} - \mathcal{A}^{\text{theory}}][(\Sigma^2)^{-1}][\mathcal{A}^{\text{pseudo-data}} - \mathcal{A}^{\text{theory}}]^T \tag{55}$$

where $\mathcal{A}$ is a dimension-N_{bin} vector with N_{bin} the total number of (x, Q^2) bins, Σ^2 is the uncertainty matrix of dimension $N_{\text{bin}} \times N_{\text{bin}}$, described in Section IV C, and the $\sin^2\theta_W$ to be fitted enters A^{theory}. The PDF portion of the uncertainty matrix was evaluated using the CT18NLO [10] (LHAPDF [11] ID 14400–14458), MMHT2014nlo_68cl [12] (ID 25100–25150) and NNPDF31_nlo_as_0118 [7] (ID 303400–303500) PDF sets.

Our results for $\sin^2\theta_W$ are shown in Tables III and IV for five energy and nominal annual luminosity combinations for ep and eD collisions, respectively. These results are illustrated in Fig. 11. The inner error bars show the combined uncertainty from statistical and 1% uncorrelated experimental systematics (due to particle background); the median error bars show the experimental uncertainty that includes statistical, 1% uncorrelated experimental systematics, and 1% electron polarimetry. The outer-most error bars (which almost coincide with the median error bars) include all the above and the PDF uncertainty evaluated using the CT18NLO sets. Results evaluated with the MMHT2014 and NNPDF31NLO sets are similar. Along with our projection with EIC annual nominal luminosity, we show the "YR reference point" (blue diamond), obtained from combining 100 fb^{-1} ep 18×275 GeV and 10 fb^{-1} eD 18×137 GeV pseudo-data. Also shown are the expected precision from near-future P2 [13], MOLLER [14] and SoLID [15] PVDIS [16, 17] experiments, respectively, that will dominate the landscape of low to medium energy scales.

We note that our results have larger uncertainties than in the YR [6] that fitted PDFs and $\sin^2\theta_W$ simultaneously using the JAM framework [18], possibly due to using realistic detector simulation and accurate running conditions. On the other hand, we found that PDF uncertainties are likely not a dominant uncertainty for EIC projections, but the electron polarization is, for settings where the integrated luminosity approaches 100 fb^{-1}. Consequently, upgrading the luminosity of EIC does not bring significant improvement on the

uncertainty of $\sin^2\theta_W$ and therefore we will not show our fitting results for the ten-fold luminosity upgrade.

Beam type and energy	ep 5 × 100	ep 10 × 100	ep 10 × 275	ep 18 × 275	ep 18 × 275
Label	P2	P3	P4	P5	P6
Luminosity (fb^{-1})	36.8	44.8	100	15.4	(100 YR ref)
$\langle Q^2\rangle$ (GeV^2)	154.4	308.1	687.3	1055.1	1055.1
$\langle A_{PV}\rangle$ ($P_e = 0.8$)	−0.00854	−0.01617	−0.03254	−0.04594	−0.04594
$(\mathrm{d}A/A)_{\mathrm{stat}}$	1.54%	0.98%	0.40%	0.80%	(0.31%)
$(\mathrm{d}A/A)_{\mathrm{stat+syst(bg)}}$	1.55%	1.00%	0.43%	0.81%	(0.35%)
$(\mathrm{d}A/A)_{1\%\mathrm{pol}}$	1.0%	1.0%	1.0%	1.0%	(1.0%)
$(\mathrm{d}A/A)_{\mathrm{tot}}$	1.84%	1.42%	1.09%	1.29%	(1.06%)
Experimental					
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{stat+syst(bg)}}$	0.002032	0.001299	0.000597	0.001176	0.000516
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{stat+syst+pol}}$	0.002342	0.001759	0.001297	0.001769	0.001244
with PDF					
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,CT18NLO}}$	0.002388	0.001807	0.001363	0.001823	0.001320
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,MMHT2014}}$	0.002353	0.001771	0.001319	0.001781	0.001270
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,NNPDF31}}$	0.002351	0.001789	0.001313	0.001801	0.001308

TABLE III. Projected PVDIS asymmetry and fitted results for $\sin^2\theta_W$ using ep collision data and the nominal annual luminosity. Here $\langle Q^2\rangle$ denotes the value averaged over all (x, Q^2) bins, weighted by $(dA/A)^{-2}_{\mathrm{stat}}$ for each bin. The electron beam polarization is assumed to be 80% with a relative 1% uncertainty. The total ("tot") uncertainty is from combining all of statistical, 1% systematic (background), 1% beam polarization, and the PDF uncertainty evaluated using three different PDF sets. The right-most column is for comparison with the YR.

Our results show that the EIC will provide determination of $\sin^2\theta_W$ at an energy scale that bridges higher energy colliders with low to medium-energy SM tests. Additionally, data points of different $\sqrt{s}$ values of EIC can be combined, or the Q^2-dependence of the EW parameter can be explored, depending on the runplan of the EIC. Furthermore, one could study the exploratory potential of EIC beyond the scope of a single SM parameter, and we

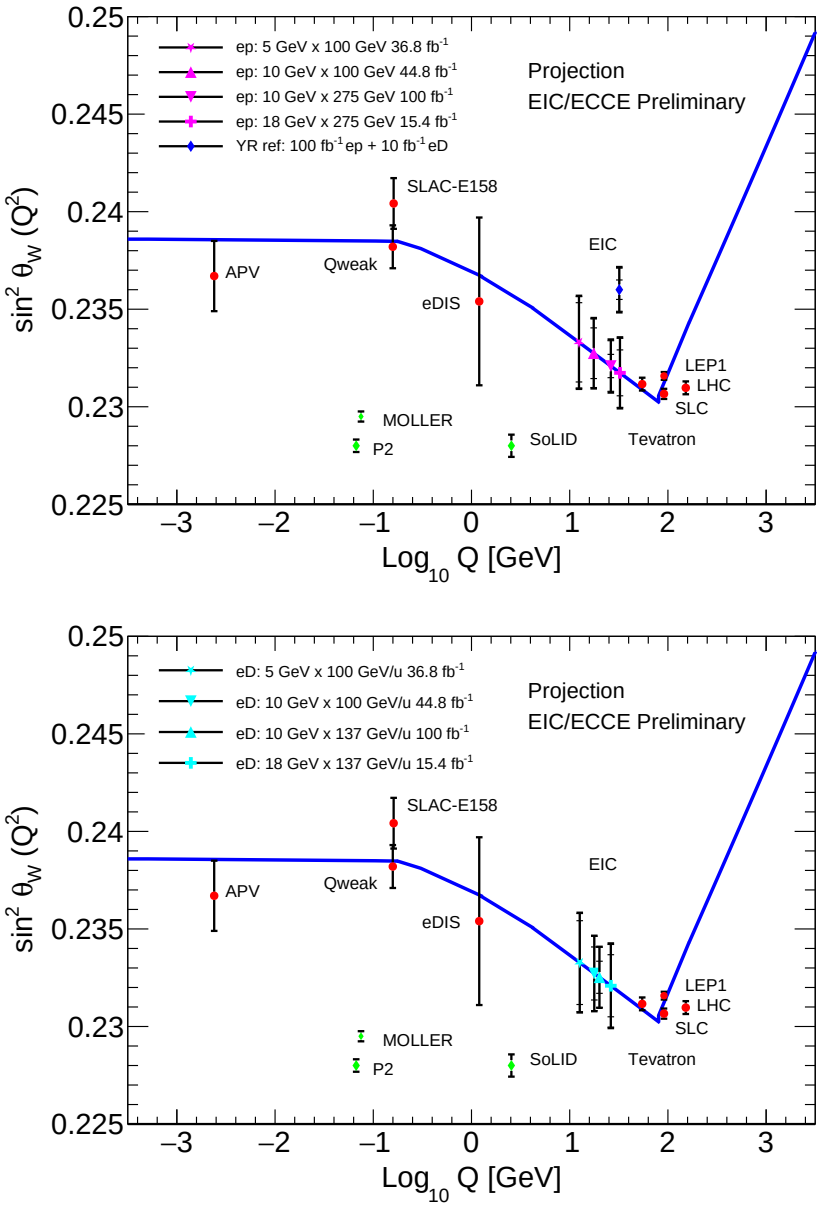

FIG. 11. Projected results for $\sin^2\theta_W$ using ep (top, solid magenta markers) and eD (bottom, solid cyan markers) collision data and the nominal annual luminosity given in Table 10.1 of the Yellow Report [6], along with existing world data (red solid circles) and hear-future projections (green diamonds), see text for details. Data points for Tevatron and LHC are shifted horizontally for clarity. The script used to produce this plot was inherited from [19], and the scale dependence of the weak mixing angle expected in the SM (blue curve) is defined in the modified minimal subtraction scheme ($\overline{\text{MS}}$ scheme) [20].

Beam type and energy	eD 5 × 100	eD 10 × 100	eD 10 × 137	eD 18 × 137	eD 18 × 137
Label	D2	D3	D4	D5	N/A
Luminosity (fb^{-1})	36.8	44.8	100	15.4	(10 YR ref)
$\langle Q^2 \rangle$ (GeV^2)	160.0	316.9	403.5	687.2	687.2
$\langle A_{PV} \rangle$ ($P_e = 0.8$)	−0.01028	−0.01923	−0.02366	−0.03719	−0.03719
$(\mathrm{d}A/A)_{\mathrm{stat}}$	1.46%	0.93%	0.54%	1.05%	(1.31%)
$(\mathrm{d}A/A)_{\mathrm{stat+bg}}$	1.47%	0.95%	0.56%	1.07%	(1.32%)
$(\mathrm{d}A/A)_{\mathrm{syst,1\%pol}}$	1.0%	1.0%	1.0%	1.0%	(1.0%)
$(\mathrm{d}A/A)_{\mathrm{tot}}$	1.78%	1.38%	1.15%	1.46%	(1.66%)
Experimental					
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{stat+bg}}$	0.002148	0.001359	0.000823	0.001591	0.001963
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{stat+bg+pol}}$	0.002515	0.001904	0.001544	0.002116	0.002414
with PDF					
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,CT18}}$	0.002558	0.001936	0.001566	0.002173	0.00247
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,MMHT2014}}$	0.002527	0.001917	0.001562	0.002128	0.002424
$\mathrm{d}(\sin^2\theta_W)_{\mathrm{tot,NNPDF31}}$	0.002526	0.001915	0.001560	0.002127	0.002423

TABLE IV. Projected PVDIS asymmetry and fitted results for $\sin^2\theta_W$ using eD collision data and the nominal annual luminosity. The uncertainty evaluation is the same as Table III.

provide results using the SMEFT framework in the next section.

VI. FRAMEWORK FOR THE SMEFT ANALYSIS

A. Data Generation and Selection

We use the procedure described in Section III to determine the uncertainty of our data projection and the uncertainty matrix. We consider both ep and eD collisions and we focus on the two highest-energy settings listed in Table II. Because collisions with higher center-of-mass energy are more sensitive to SMEFT operators, we choose four data families with the two highest $\sqrt{s}$ to focus on:

$$18 \text{ GeV} \times 100 \text{ GeV } eD \text{ } 100 \text{ fb}^{-1}\text{: D4, } \Delta\text{D4, LD4}$$

$$18 \text{ GeV} \times 137 \text{ GeV } eD \text{ } 15.4 \text{ fb}^{-1}\text{: D5, } \Delta\text{D5, LD5}$$

$$18 \text{ GeV} \times 100 \text{ GeV } ep \text{ } 100 \text{ fb}^{-1}\text{: P4, } \Delta\text{P4, LP4}$$

$$18 \text{ GeV} \times 275 \text{ GeV } ep \text{ } 15.4 \text{ fb}^{-1}\text{: P5, } \Delta\text{P5, LP5.}$$

For the highest $\sqrt{s}$ but lower-luminosity families D5, LD5, P5 and LP5 we consider two data sets: the nominal luminosity as indicated above and in Table II, and the high luminosity option denoted with an "HL" label with ten-fold higher statistics.

We use Eq. (44) to generate $N_{\text{exp}} = 1000$ pseudodata sets for each of the data families. We then impose the following selection criteria on x and Q^2, and the inelasticity y:

$$x < 0.5, \quad Q^2 > 100 \text{ GeV}^2, \quad 0.1 < y < 0.9. \tag{56}$$

These restrictions are designed to remove large uncertainties from non-perturbative QCD and nuclear dynamics that occur at low Q^2 and high x, where sensivity to SMEFT effects is anyways expected to be reduced. We note that the condition on y was already applied in the data generation and unfolding stage as described in Section III B.

B. Structure of the SMEFT asymmetry corrections

In the computation of SMEFT asymmetry values, A_{SMEFT}, we use the central member of the PDF set under consideration. We use PDF sets `NNPDF31_nlo_as_0118` [21] and `NNPDFpol11_100` [8] for the computation of unpolarized $A_{PV}^{(e)}$ and polarized $A_{PV}^{(H)}$ PV asymmetries, respectively. We factor out the ultraviolet (UV) cut-off scale from all the seven Wilson coefficients, $C_k \to C_k/\Lambda_{\text{UV}}^2$, and set $\Lambda_{\text{UV}} = 1$ TeV. *Kaan: ~~We focus on one or two~~*

~~*Wilson coefficients at a time for simplicity of presentation*~~ We turn on only one or two Wilson coefficients at a time and set the remaining ones to zero and linearize SMEFT expressions with respect to the Wilson coefficient(s) of interest. SMEFT asymmetry expressions then generically take the form

$$A_{\rm SMEFT}(x,Q^2) = A^{\rm theo}_{\rm SM,0}(x,Q^2) + C\delta(x,Q^2) \tag{57}$$

or

$$A_{\rm SMEFT}(x,Q^2) = A^{\rm theo}_{\rm SM,0}(x,Q^2) + C_1\delta_1(x,Q^2) + C_2\delta_2(x,Q^2). \tag{58}$$

Comparing Eq. (57) to (44) or (45), we see that at the end of a multi-run analysis, the distribution of the best-fit values for any single Wilson coefficient should be a Gaussian centered at the origin.

C. Best-fit analysis of Wilson coefficients

Generating pseudo-data values, $A^{\rm pseudo}_{\rm SM}$ and obtaining the SMEFT asymmetry expressions, $A_{\rm SMEFT}$, we define a χ^2 test statistic as

$$\chi^2 = \sum_{b=1}^{N_{\rm bin}}\sum_{b'=1}^{N_{\rm bin}} \left[A_{{\rm SMEFT},b} - A^{\rm pseudo}_{{\rm SM},b}\right] [(\Sigma^2)^{-1}]_{bb'} \left[A_{{\rm SMEFT},b'} - A'^{\rm pseudo}_{{\rm SM},b}\right] \tag{59}$$

where $N_{\rm bin}$ is the number of bins in a given data set. Generically, it will look like

$$\chi^2(C) = k_0 + k_1 C + k_2 C^2 \tag{60}$$

for a single-parameter fit of Wilson coefficient C, or

$$\chi^2(C_1,C_2) = k_{00} + k_{10}C_1 + k_{01}C_2 + k_{11}C_1C_2 + k_{20}C_1^2 + k_{02}C_2^2 \tag{61}$$

for a two-parameter fit of Wilson coefficients C_1 and C_2. The χ^2 function is minimized with respect to C or to C_1 and C_2. This gives us the best-fit values, $\bar{C}$ or $\bar{C}_1$ and $\bar{C}_2$. We obtain the inverse square of the error of the single-parameter best-fit value via

$$\frac{1}{\sigma_C^2} = \frac{1}{2}\frac{{\rm d}^2\chi^2}{{\rm d}C^2} \tag{62}$$

evaluated at $\bar{C}$. The inverse covariance matrix, V^{-1}, of the two-parameter fit is constructed in such a way that its $ij^{\rm th}$ component is given by

$$(V^{-1})_{ij} = \frac{1}{2}\frac{\partial\chi^2}{\partial C_i \partial C_j} \tag{63}$$

for $i, j = 1, 2$, evaluated at the best-fit values of C_1 and C_2. Inverting V^{-1}, we obtain the individual errors and the correlation of the fit:

$$V = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ & \sigma_2^2 \end{pmatrix}_{\rm sym} \tag{64}$$

1. *Averaging over multiple pseudodata sets*

When we repeat $N_{\rm exp}$ times the single-parameter best-fit analysis described in Sec. VI C, we obtain $N_{\rm exp}$ best-fit values, $\bar{C}_e$, with corresponding uncertainties, $\sigma_{C,e}$, for each pseudo-experiment e. The mean of the best-fit values is obtained by averaging individual best-fit values weighted by the inverse square of the uncertainties:

$$\bar{C} = \left(\sum_{e=1}^{N_{\rm exp}} \frac{1}{\sigma_{C,e}^2} \right)^{-1} \left(\sum_{e=1}^{N_{\rm exp}} \frac{1}{\sigma_{C,e}^2} \bar{C}_e \right) \ , \tag{65}$$

and the average uncertainty of this mean value is obtained via

$$\frac{1}{\sigma_C^2} = \frac{1}{N_{\rm exp}} \sum_{e=1}^{N_{\rm exp}} \frac{1}{\sigma_{C,e}^2} \ . \tag{66}$$

When we repeat $N_{\rm exp}$ times the two-parameter best-fit analysis on Wilson coefficients described in Sec. VI C, we obtain $N_{\rm exp}$ pairs of best-fit values, $\bar{C}_{1,e}$ and $\bar{C}_{2,e}$, and inverse covariance matrices, $(V^{-1})_e$, for each pseudo-experiment, e. The best-fit values are averaged similarly to the one-dimensional case but with the inverse square of uncertainties replaced by inverse covariance matrices:

$$\begin{pmatrix} \bar{C}_1 \\ \bar{C}_2 \end{pmatrix} = \left[\sum_{e=1}^{N_{\rm exp}} (V^{-1})_e \right]^{-1} \left[\sum_{e=1}^{N_{\rm exp}} (V^{-1})_e \begin{pmatrix} \bar{C}_{1,e} \\ \bar{C}_{2,e} \end{pmatrix} \right] \ . \tag{67}$$

The average inverse covariance matrix of the resultant best fit is calculated using

$$V^{-1} = \frac{1}{N_{\rm exp}} \sum_{e=1}^{N_{\rm exp}} (V^{-1})_e \ . \tag{68}$$

We note the presence of the factor $1/N_{\rm exp}$ in Eqs. (66) and (68). Without it, we would be effectively increasing the luminosity of the corresponding central data set by factor $N_{\rm exp}$. We avoid this by including this factor in computing the average uncertainty or inverse covariance matrix.

2. Definition of confidence intervals

The result of a single-parameter multi-run fit can be expressed as

$$\frac{(C-\bar{C})^2}{\sigma_C^2} = \Delta\chi^2 \tag{69}$$

so we can express the fitted result and the uncertainty of coefficient C as

$$C = C_{\text{best}} \pm \sqrt{\Delta\chi^2}\sigma_C \ . \tag{70}$$

For a two-parameter multi-run fit, the ellipse equation reads

$$\begin{pmatrix} C_1-\bar{C}_1 \\ C_2-\bar{C}_2 \end{pmatrix}^{\text{T}} V^{-1} \begin{pmatrix} C_1-\bar{C}_1 \\ C_2-\bar{C}_2 \end{pmatrix} = \Delta\chi^2 \tag{71}$$

in the (C_1, C_2) plane. The $\Delta\chi^2$ values that determines the size of the best-fit interval for an arbitrary confidence level are well-known. For 95% CL, we have $\Delta\chi^2 = 3.841$, 5.991, and 7.815 for one-, two-, and three-parameter fits, respectively.

3. Combination of best-fits from distinct data sets

Suppose we have two data sets, say T1 and T2, from which we obtain the single-parameter best-fit values of Wilson coefficient, C, to be $\bar{C}_{\text{T1}}$ and $\bar{C}_{\text{T2}}$, together with the errors $\sigma_{C,\text{T1}}$ and $\sigma_{C,\text{T2}}$. Assuming the said data sets can be treated uncorrelated to a good approximation, we obtained the combined best-fit value and the corresponding uncertainty by using Eqs. (65) and (66) with slight modifications. Firstly, the summation index e now runs from 1 to 2, representing number of data sets. Secondly, the $1/N_{\text{exp}}$ factor should be removed from Eq. (66) because we now have indeed two independent (uncorrelated) measurements. This method can be generalized to the combination of the best-fit values from more than two data sets such as different beam energies, and to the case of multi-parameter fits in a straightforward manner.

4. Simultaneous fit of Wilson coefficients and beam polarization or luminosity difference

We observed in Section IV D that experimental uncertainties such as the beam polarization and luminosity difference between e^+ and e^- runs can be limiting factors for some of

the data sets. When the data statistical uncertainty is very precise, there is the possibility that one could use data themselves to constrain these systematic effects. We present in Appendix A 2 a method to simultaneously fit the SMEFT coefficient(s) and the beam polarization for PV asymmetries, and in Appendix A 1 a method to simultaneously fit the SMEFT coefficient(s) and the luminosity difference for the LC asymmetries.

VII. SMEFT FIT RESULTS

A. Fits of single Wilson coefficients

In this section, we discuss the 95% CL intervals for the Wilson coefficients in single-parameter fits averaged over 1000 pseudo-experiments. The bounds on the Wilson coefficient C_{eu} across numerous data sets are representative and exhibit the common features of single-parameter fits. We therefore show only the bounds on C_{eu} to illustrate the main observations and include the remaining Wilson coefficients in Appendix B 1. Figure 12 displays the 95% CL intervals of C_{eu} for the four data families in which we are primarily interested in this paper. The intervals are grouped by asymmetries, namely electron PV asymmetries $A_{PV}^{(e)}$ of unpolarized hadrons ("unpolarized $A_{\rm PV}$"), hadron PV asymmetries $A_{PV}^{(H)}$ with unpolarized electrons ("polarized $A_{\rm PV}$"), and unpolarized electron-positron asymmetries $A_{LC,H}$ of unpolarized hadrons ("lepton-charge A"). PV asymmetries are then grouped into two, showing the fits in the nominal- and high-luminosity scenarios. In each block of intervals, there are four double lines in the case of PV asymmetries and four single lines in LC asymmetries. These four lines correspond to the data families D4 (black and its shades), D5 (red), P4 (blue), and P5 (orange). The darker of the two lines indicate the bounds from single-parameter fits with the Wilson coefficient C_{eu}, whereas the lighter ones show the bounds on the said Wilson coefficient from simultaneous $(1+1)$-parameter fits with C_{eu} and the beam polarization. We describe the details of the fits involving the beam polarization as an additional free variable in Appendix A 2.

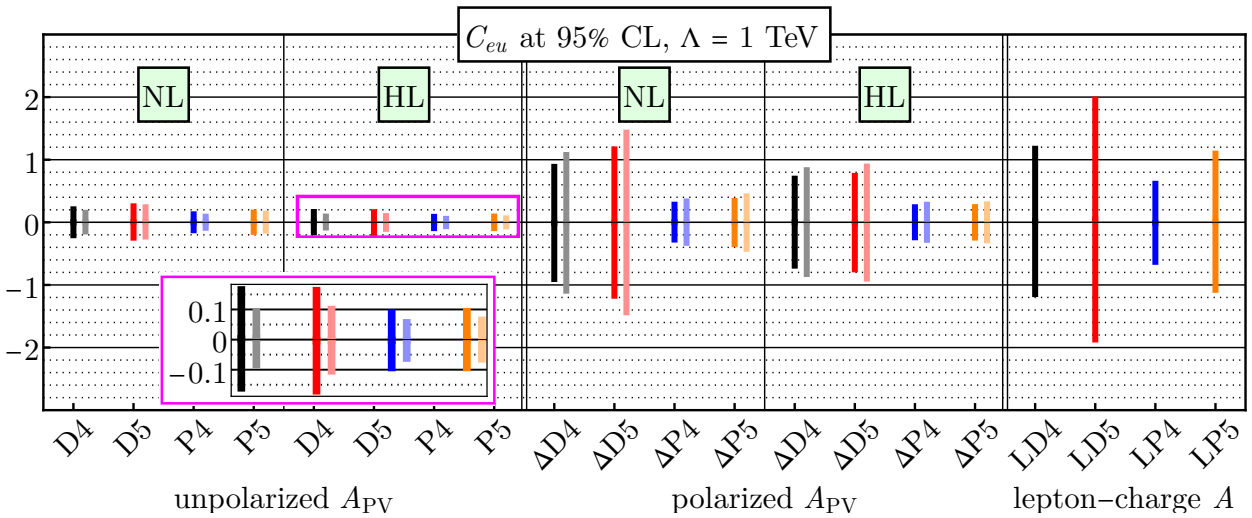

FIG. 12. 95% CL bounds of C_{eu} from single-parameters fits (darker) and from the (1+1)-parameter fits with beam polarization as an additional fitting parameter (lighter) using the families of data sets D4, D5, P4, and P5 at $\Lambda = 1$ TeV.

From Fig. 12, we can extract the following main points.

- Proton asymmetries of all the three types, namely unpolarized and polarized PV asymmetries and LC asymmetries, impose considerably stronger bounds than deuteron.
- High-energy low-luminosity data sets D5 and P5 lead to slightly weaker bounds than the less energetic but higher-luminosity ones, D4 and P4, respectively.
- Unpolarized PV asymmetries $A_{PV}^{(e)}$ offer much stricter bounds than the polarized ones $A_{PV}^{(H)}$; however, it should be noted that for some Wilson coefficients, unpolarized proton asymmetries yield nearly the same bounds as the corresponding polarized ones.
- Data sets in the high-luminosity scenario make a noticeable difference in the size of bounds. The improvement due to increased luminosity is slightly more significant for polarized deuteron asymmetries.
- Bounds from electron-positron asymmetries $A_{LC,H}$ are comparable to or looser than the ones from polarized hadron asymmetries. Never do they offer stricter bounds than high-luminosity hadron PV asymmetries.
- If introducing the beam polarization as a new free fitting parameter, unpolarized hadron asymmetries give considerably stronger bounds. The improvement is more

significant in the high-luminosity scenario. However, the same fitting method yields weaker bounds with polarized hadron asymmetries.

Assuming weak correlations, one can also combine the bounds within a given family of data sets, e.g. D4, ΔD4, and LD4. We find that the resultant bound is never stronger than the strongest one obtained from the individual family members, which is the electron PV asymmetry data.

In Fig. 13, we present the effective UV cut-off scales, $\Lambda/\sqrt{C_{eu}}$, with $\Lambda = 1$ TeV, corresponding to the bounds shown in Fig. 12. The organization of this plot in terms of asymmetries and data sets is the same as Fig. 12. Improved bounds on C_{eu} with the addition of the beam polarization to the fits are equivalent to higher energy scales in the unpolarized PV asymmetries, which are indicated by the lighter columns in the background; on the other hand, weaker bounds from the fits with beam polarization are depicted by the lighter columns in the foreground for the polarized PV asymmetries.

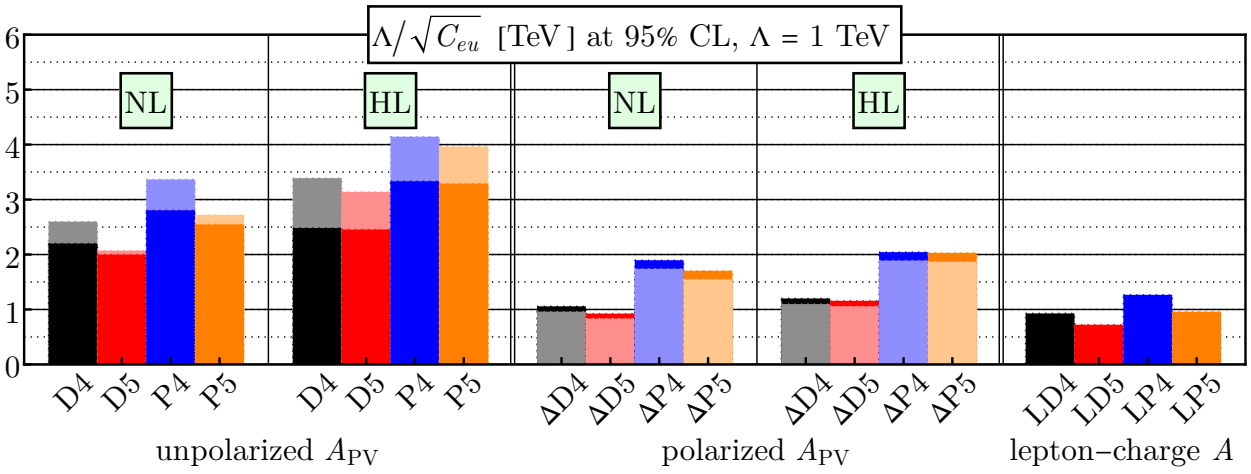

FIG. 13. Effective UV cut-off scales, $\Lambda/\sqrt{C_k}$, defined in terms of the 95% CL bounds on the Wilson coefficients and with $\Lambda = 1$ TeV.

One can observe that scales reaching 3 TeV can be probed with nominal luminosity, while scales exceeding 4 TeV can be probed for others. We remark that care must be taken in comparing these mass limits with others found in the literature, which sometimes assume a strong coupling limit equivalent to setting $C_i = 4\pi$ and assume also maximally constructive interference between different quark contributions. For example, converting our results to the notation of [22] would yield a bound on $\Lambda/\sqrt{C_{eu}}$ of 19 TeV, instead of 3 TeV quoted here,

which is only very approximate and is calculated by multiplying 3 TeV by $\sqrt{4\pi}$ and $\sqrt{\sqrt{5}}$, where the latter is to account for the constructive interference between quark contributions, and by another factor to convert 90% CL to 95% CL.

B. Fits of two Wilson coefficients

In this section, we discuss fits on pairs of Wilson coefficients in order to determine how well the EIC can break degeneracies between parameters that occur in the LHC Drell-Yan data [23, 24]. We emphasize that just as we show only representative examples, these examples are from the simultaneous fits with beam polarization in light of the results of the one-parameter fits in the previous section. The description of the beam-polarization fits is presented in Appendix A 2. The complete set of plots of confidence ellipses are given in Appendix B 2.

In Fig. 14, we compare the 95% CL ellipses for the pair (C_{eu}, C_{qe}) between the D4 and P4 data families. Each asymmetry type gives a distinct correlation pattern, complementary to one another. Electron-positron asymmetries give rise to wide and elongated, band-like ellipses compared to PV asymmetries. As in the case of one-parameter fits, electron PV asymmetries of unpolarized hadrons offer the strongest bounds on the pairs of Wilson coefficients. Comparing deuteron to proton, one can see that proton data are significantly more constraining.

Figure 15 shows the comparison of the simultaneous fit on the Wilson coefficients $(C_{eu}, C_{\ell u})$ projected for the EIC to the corresponding fit with the LHC data adapted from [24]. The LHC fits exhibit a flat direction, i.e. a particular linear combination of the two coefficients cannot be determined. A similar comparison is given in Fig. 16 for the pair $(C_{eu}, C_{\ell q}^{(1)})$, using the nominal- and high-luminosity P4 set of the EIC. We observe that in both figures, projected EIC fits have different correlation patterns from the LHC. More importantly, the EIC projected data show the capability of resolving flat directions and significantly constraining the aforementioned pairs of Wilson coefficients.

Finally, in Fig. 17, we present fits from the P4 data set and the LHC adapted from [23], for the pair $(C_{\ell q}^{(1)}, C_{\ell q}^{(3)})$. This figure shows that when the LHC data imposes tight bounds on a pair of Wilson coefficients, the EIC preliminary data can introduce far stronger bounds on the same pair of Wilson coefficients. Moreover, fits from EIC and LHC have distinct

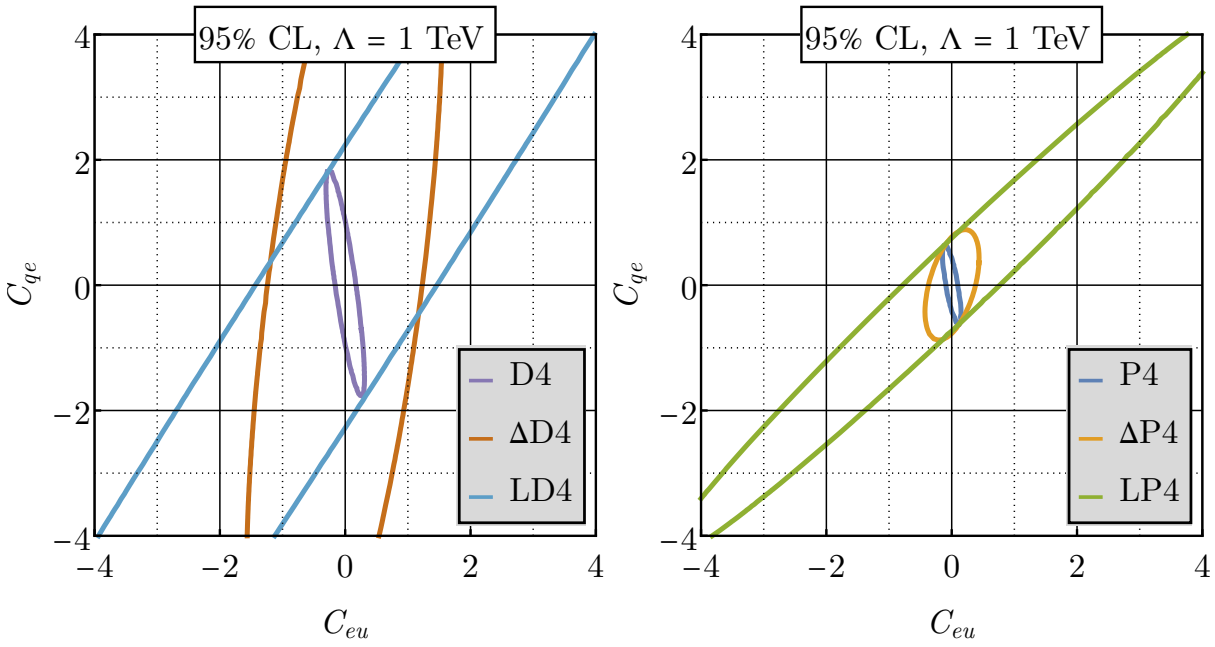

FIG. 14. 95% CL ellipses for the Wilson coefficients C_{eu} and C_{qe} using the families of data sets D4 and P4 in the simultaneous $(2+1)$ fits that includes the beam polarization as an additional free fitting parameter.

correlations, which indicates the complementarity of the EIC to the LHC as a future collider. Treating the projected EIC and the LHC data to be uncorrelated, we also plot the combined fit of the two, which turns out to even more strongly constrain the chosen pair of Wilson coefficients. We remark that the effective UV scales probed with the combined data set exceed 2 TeV.

It should be noted that there appear flat directions in the fits of certain pairs of Wilson coefficients with the projected EIC data that utilize the deuteron beam. Examples include (C_{eu}, C_{ed}) and $(C_{\ell u}, C_{\ell d})$. We can explain these observations analytically. We find that these pairs always appear in a specific way in asymmetry expressions; to wit, $2C_{eu} - C_{ed}$ for electron PV asymmetries with unpolarized deuteron. In all such cases only one of the data families exhibits this behavior, with the degeneracy broken by another data family.

Our results on the bounds from Wilson coefficients in simultaneous $(2+1)$-parameter fits with the beam polarization as an additional parameter can be summarized as follows:

- Proton asymmetries impose much stricter bounds than deuteron.
- Unpolarized hadron asymmetries lead to stronger correlations than polarized ones.

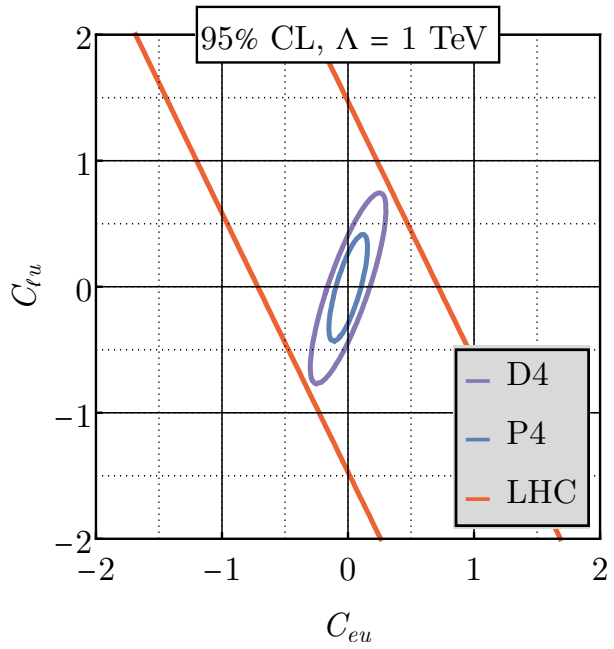

FIG. 15. 95% CL ellipses for the Wilson coefficients C_{eu} and C_{qe} using the data sets D4 and P4 in the $(2+1)$ fit that includes the beam polarization as an additional fitting parameter, compared with the corresponding two-parameter fit from the LHC data [24].

- The three types of asymmetries of deuteron and proton considered in this work, together with the LHC data, are complementary to each other in the sense that they offer distinct correlation patterns.

- The projected EIC data are capable of resolving all flat directions that appear in the LHC Drell-Yan data.

- The bounds from the projected EIC data can be much stronger than the LHC data, advertising the EIC as an excellent future collider.

VIII. CONCLUSIONS

In this manuscript we have analyzed the potential of testing the electroweak SM and exploring BSM physics of the future EIC. We have focused on the precision determination of the weak mixing angle over a wide range of momentum transfers, and probes of heavy new physics. We have provided all formulae for NC DIS and simulation details that will be needed for future studies of these areas. Our BSM analysis utilizes the model-independent SMEFT

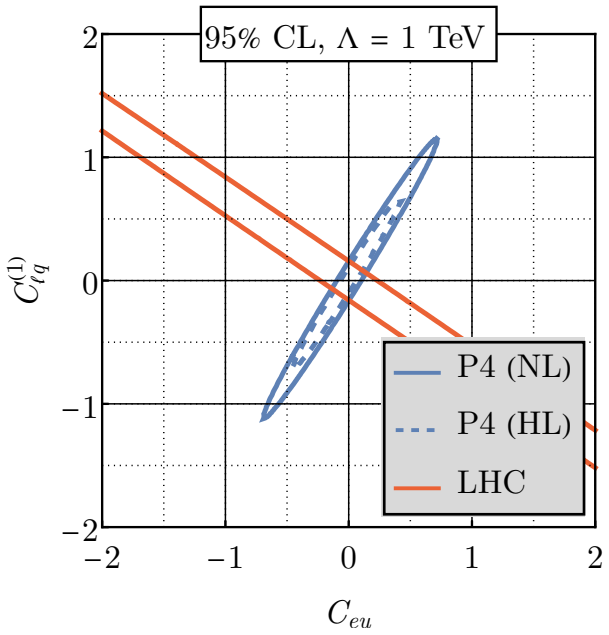

FIG. 16. 95% CL ellipses for the Wilson coefficients C_{eu} and $C_{\ell q}^{(1)}$ using the nominal- and high-luminosity data set P4 in the $(2+1)$ fit that includes the beam polarization as an additional fitting parameter, compared with the corresponding two-parameter fit from the LHC data [23].

framework, and focuses on the semi-leptonic four-fermion operator sector of the theory. We translate our formalism into the DIS language in terms of parity couplings and structure functions to facilitate communication between the high energy and medium energy physics communities. We provide a detailed accounting of uncertainties from statistics, experimental systematic effects, beam polarimetry for PV asymmetries, QED higher-order corrections for LC asymmetries, and finally from PDFs. Additionally, we explore simultaneously fitting the beam polarization with the anticipated high-precision PV asymmetry data as a possible analysis technique to improve upon the experimental limitation from beam polarimetry.

Our BSM analysis finds that UV scales in excess of 3 TeV can be probed with the currently planned (nominal) annual luminosity of the EIC, with scales above 4 TeV possible with a ten-fold high luminosity upgrade. The most stringent bounds come from polarized electron scattering off of unpolarized protons. Constraints from polarized hadrons, deuterons, and from a possible future positron beam provide important complementary probes. Our complete study of correlations between Wilson coefficients finds that no degeneracies remain upon combining all EIC data sets. This is not the case with LHC Drell-Yan measurements, in which numerous degeneracies exist, and will continue to occur even after LHC's high lu-

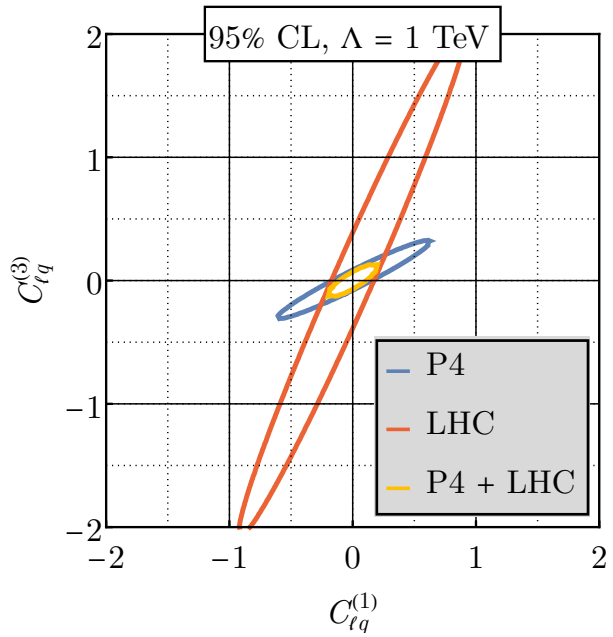

FIG. 17. 95% CL ellipses for the Wilson coefficients $C_{\ell q}^{(1)}$ and $C_{\ell q}^{(3)}$ using the nominal-luminosity data set P4 in the $(2+1)$ fit that includes the beam polarization as an additional fitting parameter, compared with the corresponding fit from the LHC data [23], and the combined fit of the two.

minosity running. This demonstrates that although the EIC is primarily thought of as a QCD machine, it is in fact a powerful probe of potential BSM effects with a broad coverage of heavy new physics parameter space, and is in many ways competitive with the higher energy LHC. We hope that our work motivates future studies of the unexpected power of the EIC for new physics searches.

Appendix A: Additional fits

1. Luminosity difference fits

Since electron and positron data would be taken at different times with different beam configurations, there is the possibility of a significant offset between the absolute luminosities of the two data sets. In the main text, we include this uncertainty in the error matrix as the luminosity error, $\sigma_{\rm lum} = 0.02$, which is assumed absolute. We study here the possibility of simultaneously fitting this luminosity difference together with the Wilson coefficients.

We fit the pseudodata for the LC asymmetries with an overall shift, $A_{\rm lum}$, added to the pseudo-data. Then, we define the χ^2 test statistics as

$$\chi^2 = \sum_{b=1}^{N_{\rm bin}} \sum_{b'=1}^{N_{\rm bin}} [A_{{\rm SMEFT},b} - A_{{\rm SM},b}^{\rm pseudo}][(\tilde{\Sigma}^2)^{-1}]_{bb'}[A_{{\rm SMEFT},b'} - A_{{\rm SM},b'}^{\rm pseudo}] \tag{A1}$$

where we omit the uncertainty in the luminosity difference between e^+ and e^- runs from the uncertainty matrix:

$$\tilde{\Sigma}^2 = \Sigma^2\big|_{\sigma_{\rm lum}\to 0}. \tag{A2}$$

However, we keep the luminosity uncertainty in the pseudo-data generation. By introducing the luminosity difference, $A_{\rm lum}$, as a new variable, we extend our one-parameter and two-parameter Wilson-coefficient fits to $(1+1)$-parameter and $(2+1)$-parameter fits.

We find that there are mild correlations, $|\rho_k| \lesssim 0.4$, between $A_{\rm lum}$ and any C_k in the $(1+1)$ and $(2+1)$ fits. In addition, the fitted results for Wilson coefficients have slightly larger uncertainty when the luminosity difference is treated as a fitting parameter. In Fig. 18, we show the 95% CL intervals with and without $A_{\rm lum}$ for the Wilson coefficient C_{eu} in all the four LC asymmetry data sets. In Fig. 19, we compare the 95% CL ellipses of the Wilson coefficients (C_{eu}, C_{qe}) with the data sets LD4 and LP5 with and without the luminosity difference as a fitted parameter. From these figures, we see that the 95% CL bounds on C_{eu} become 15 to 20% weaker. The difference is less noticeable in the confidence ellipses.

2. Beam polarization fits

In the same spirit as the previous section, we now consider fitting the beam polarization simultaneously with the Wilson coefficients in an attempt to reduce the uncertainty asso-

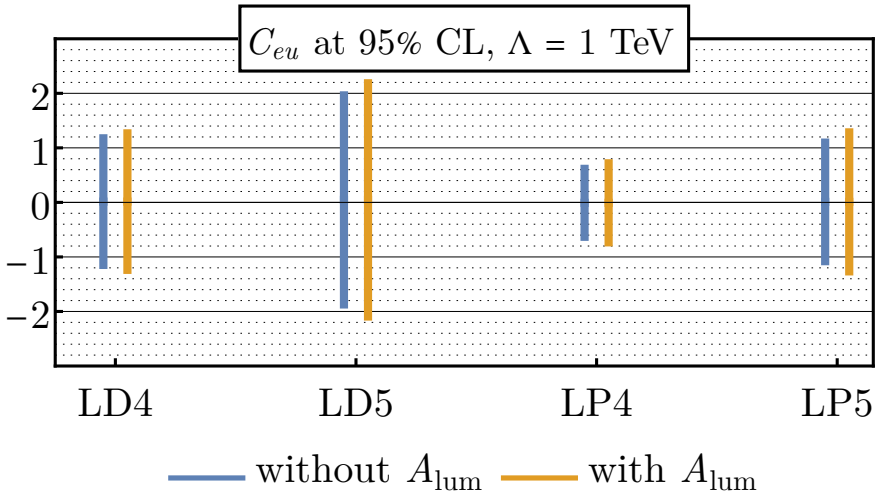

FIG. 18. Comparison of the bounds in single-parameter fits of Wilson coefficient C_{eu} with all the LC asymmetry data sets in the absence and presence of the luminosity difference as a new free fitting parameter.

ciated with the experimental limitation from beam polarimetry. We fit the pseudodata for the PV asymmetries by including a factor of P in the SMEFT asymmetries. We then define a χ^2 test statistics as

$$\chi^2 = \sum_{b=1}^{N_{\rm bin}} \sum_{b'=1}^{N_{\rm bin}} [P A_{{\rm SMEFT},b} - A_{{\rm SM},b}^{\rm pseudo}][(\tilde{\Sigma}^2)^{-1}]_{bb'}[P A_{{\rm SMEFT},b'} - A_{{\rm SM},b'}^{\rm pseudo}] + \frac{(P-\bar{P})^2}{\delta P^2} \ . \quad \text{(A3)}$$

In this approach, we omit the beam polarization uncertainty, $\sigma_{\rm pol}$, from the uncertainty matrix because it is now treated as a fitting parameter:

$$\tilde{\Sigma}^2 = \Sigma^2\big|_{\sigma_{\rm pol}\to 0} \ , \quad \text{(A4)}$$

but not during pseudo-data generation. The second term on the RHS of Eq. (A3) is added by hand, where $\bar{P}$ and δP are the beam polarization value and its uncertainty provided by the polarimetry, presumably uncorrelated to the asymmetry measurements. The logic behind this addition is that experimentally, the polarimetry does provide knowledge on the beam polarization, but we hope to obtain a better determination of the polarizations within the uncertainty provided by the polarimetry, by fitting data with high statistical precision. As for the beam polarization itself, we use a normalized value of $\bar{P} = 1$ in this study for simplicity. Treating the new term to be the contribution of a new observable, we increase the degrees of freedom of the χ^2 distribution by 1. As in the case of luminosity difference fits,

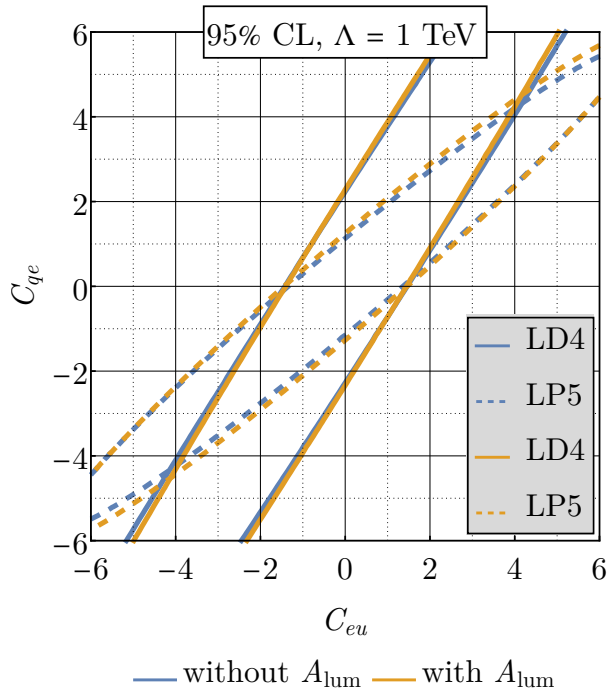

FIG. 19. Comparison of the 95% CL ellipses for the Wilson coefficients (C_{eu}, C_{qe}) with the data sets LD4 and LP5 in the absence and presence of the luminosity difference as an additional free fitting parameter.

we extend our 1- and 2-parameter fits of Wilson coefficients to (1+1)- and (2+1)-parameter simultaneous fits by including the beam polarization as a new variable.

From $(1+1)$ fits, we find that P and any C_k are rather weakly correlated, $|\rho_k| \lesssim 0.1$, in the polarized hadron data sets, whereas there are strong correlations, $|\rho_k| \gtrsim 0.7$, in the unpolarized hadron asymmetries. We observe similar correlations in the $(2+1)$ fits.

In Fig. 20, we present the allowed intervals of the Wilson coefficient C_{eu} for the nominal- and high-luminosity data sets P4 and ΔP4, while Fig. 21 displays the 95% CL ellipse of the Wilson coefficients (C_{eu}, C_{qe}) for the same data sets in the nominal-luminosity scenario. We find that bounds from unpolarized hadron data sets become stronger by 30 to 50%, yet the ones from polarized hadron asymmetries become 15 to 20% weaker. The improvement is sharper in the high-luminosity unpolarized hadron sets, whereas the worsening is significant for the nominal-luminosity polarized hadron sets.

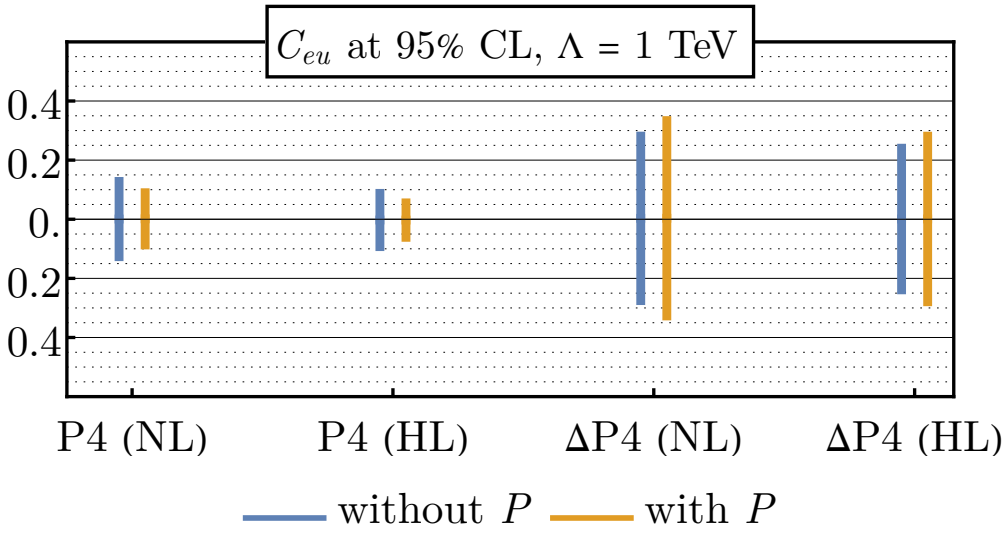

FIG. 20. 95% CL bounds on the Wilson coefficient C_{eu} with the nominal- and high-luminosity data sets P4 and ΔP4 in the absence and presence of the beam polarization, P, as an additional free variable in the fits.

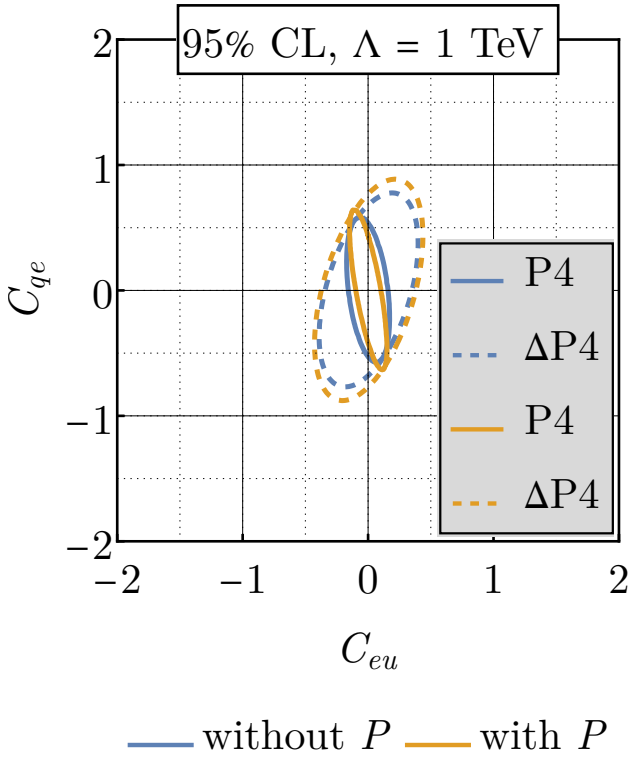

FIG. 21. 95% CL ellipse of the Wilson coefficients C_{eu} and C_{qe} for the data sets P4 and ΔP4 in the absence and presence of the beam polarization, P, as a new free variable in the fits.

876 One can explain why the bounds become weaker in the polarized hadron sets by referring

to the correlations. Since in these data sets, the beam polarization and the Wilson coefficients are found to be weakly correlated, one would naively expect the bounds obtained from single-parameter fits of Wilson coefficients to roughly remain the same on the grounds that P and C_k can be thought of almost fully independent so that they will not affect each other in the fits. Thus, any increase in the allowed limits of Wilson coefficient can be attributed to the increase in the number of parameters fitted, which is reflected as the normalization of the uncertainties of the fit.

Appendix B: Complete set of fitted results on Wilson coefficients

1. Fits of single Wilson coefficients

In this section, we present the 95% CL intervals and the corresponding effective UV cut-off scales for all the seven Wilson coefficients in single-parameter fits averaged over 1000 pseudo-experiments. For the self-containment of this section, we remind the following abbreviations for the EIC preliminary data sets:

- electron PV asymmetries $A_{PV}^{(e)}$ of unpolarized deuteron:
  - D4: eD 10 GeV $\times$ 137 GeV, 100 fb^{-1}
  - D5: eD 18 GeV $\times$ 137 GeV, 15.4 fb^{-1}
- electron PV asymmetries $A_{PV}^{(e)}$ of unpolarized proton:
  - P4: ep 10 GeV $\times$ 275 GeV, 100 fb^{-1}
  - P5: ep 18 GeV $\times$ 275 GeV, 15.4 fb^{-1}
- hadron PV asymmetries $A_{PV}^{(H)}$ with unpolarized electron: ΔD4, ΔD5, ΔP4, and ΔP5 with the same energy and luminosity configuration as the corresponding D- and P-sets
- unpolarized electron-positron asymmetries of unpolarized hadrons $A_{LC,H}$: LD4, LD5, LP4, and LP5 with the same energy configuration as the corresponding D- and P-sets, but the luminosity of the positron beam is assumed to be 10 times smaller than that of the electron beam.

Figures 22–28 display the 95% CL bounds of each Wilson coefficient for the four data families in which we are primarily interested in this work. As in the main part of the manuscript, the intervals are grouped by asymmetries, namely electron PV asymmetries of unpolarized hadrons $A_{PV}^{(e)}$ ("unpolarized $A_{\rm PV}$"), hadron PV asymmetries with unpolarized electrons ("polarized $A_{\rm PV}$"), and unpolarized electron-positron asymmetries of unpolarized hadrons $A_{LC,H}$ ("lepton-charge A"). PV asymmetries are then grouped into two, showing the fits in the nominal- and high-luminosity scenarios. The nominal luminosity ("NL") refers to the annual integrated luminosity of Table 10.1 of YR [6]. The high luminosity ("HL") is assumed to be 10 times higher than the nominal one and requires a luminosity upgrade of the EIC. In each block of intervals, there are four double lines in the case of PV asymmetries and four single lines in LC asymmetries. These four lines correspond to the data families D4 (black and its shades), D5 (red), P4 (blue), and P5 (orange). The darker of the two lines indicate the bounds from single-parameter fits with the Wilson coefficient C_k, whereas the lighter ones show the bounds on the said Wilson coefficient from simultaneous two-parameter fits with C_k and the beam polarization. The details of the fits involving the beam polarization as an additional free variable are described in Appendix A 2.

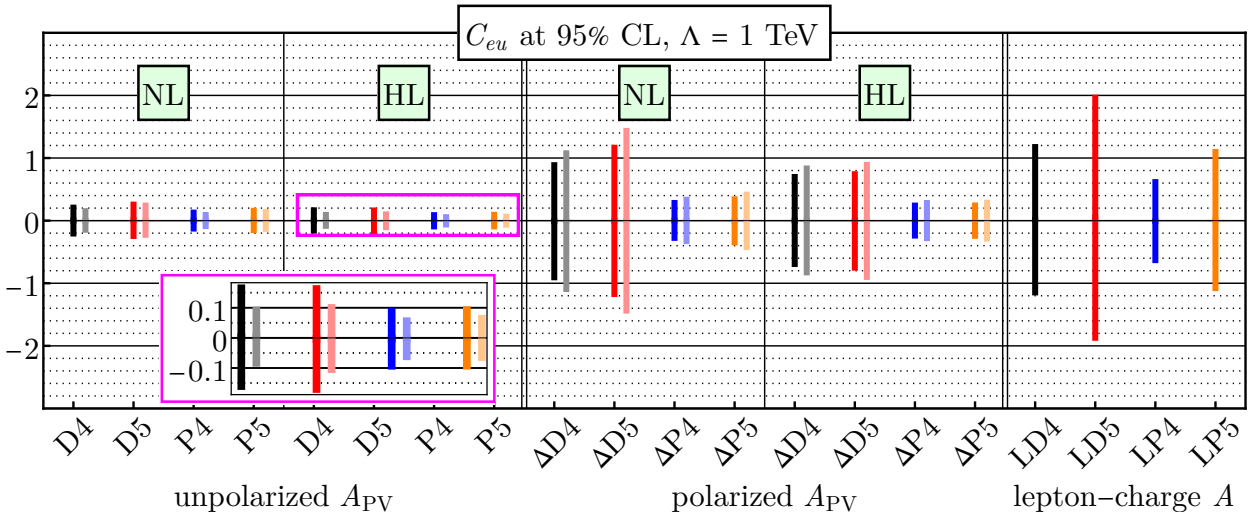

FIG. 22. 95% CL bounds of C_{eu} from 1-parameter fits (darker) and from simultaneous $(1+1)$-parameter fits with beam polarization (lighter) using the families of data sets D4, D5, P4, and P5.

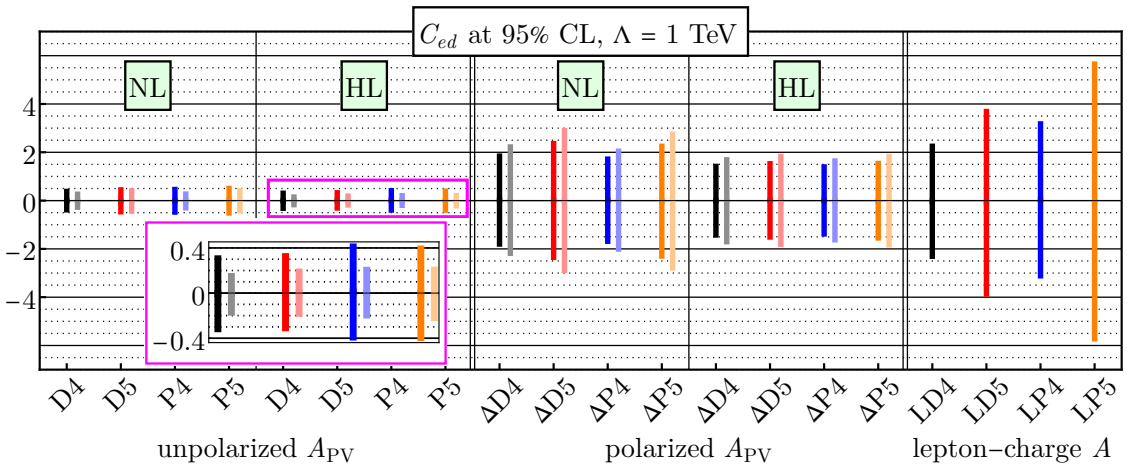

FIG. 23. The same as in Fig. 22 but for C_{ed}.

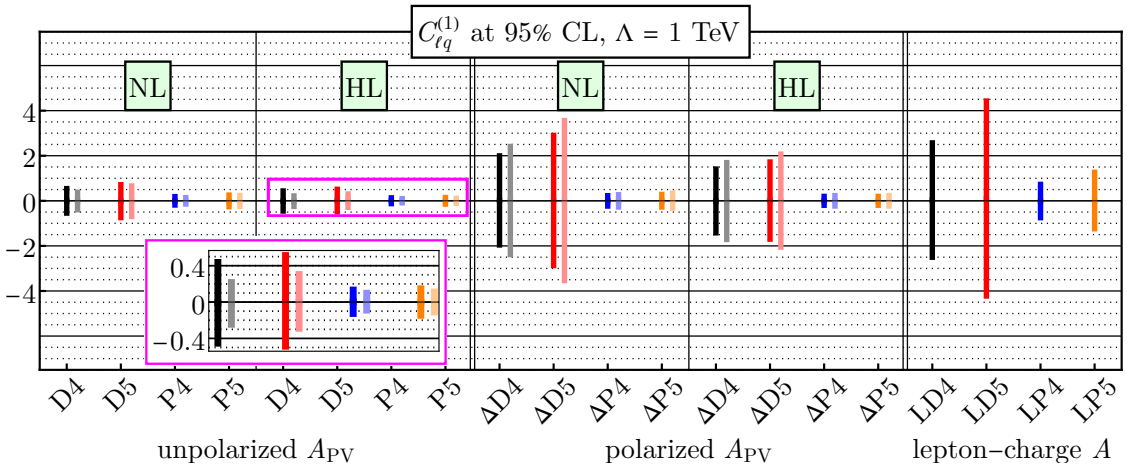

FIG. 24. The same as in Fig. 22 but for $C_{\ell q}^{(1)}$.

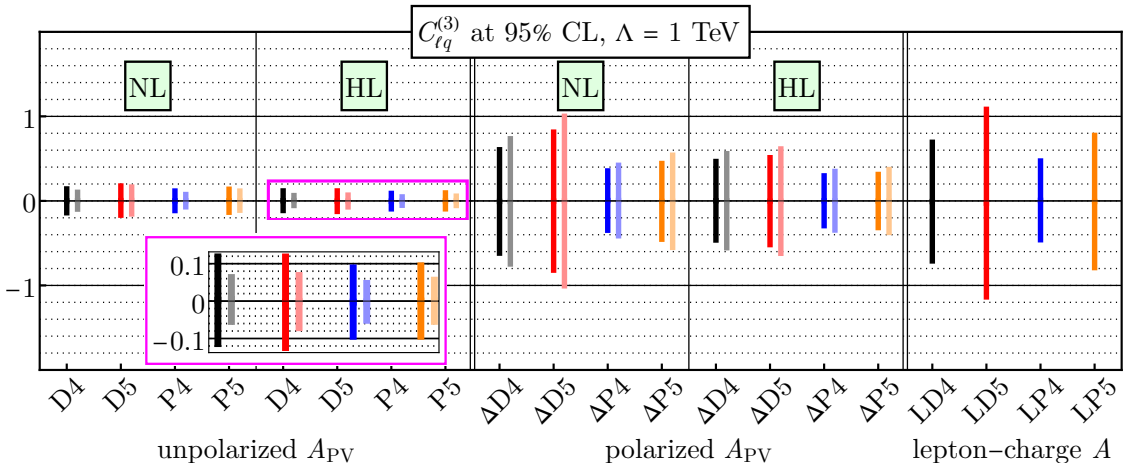

FIG. 25. The same as in Fig. 22 but for $C_{\ell q}^{(3)}$.

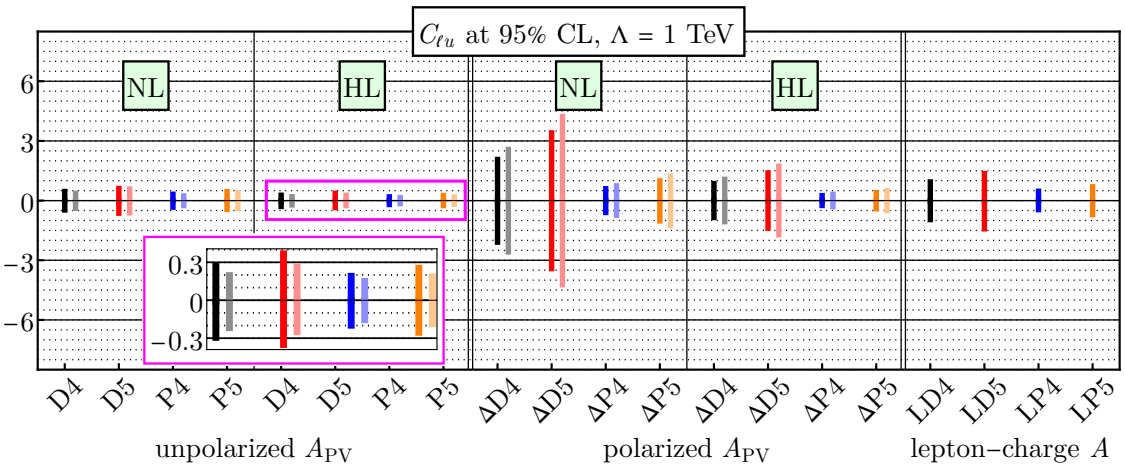

FIG. 26. The same as in Fig. 22 but for $C_{\ell u}$.

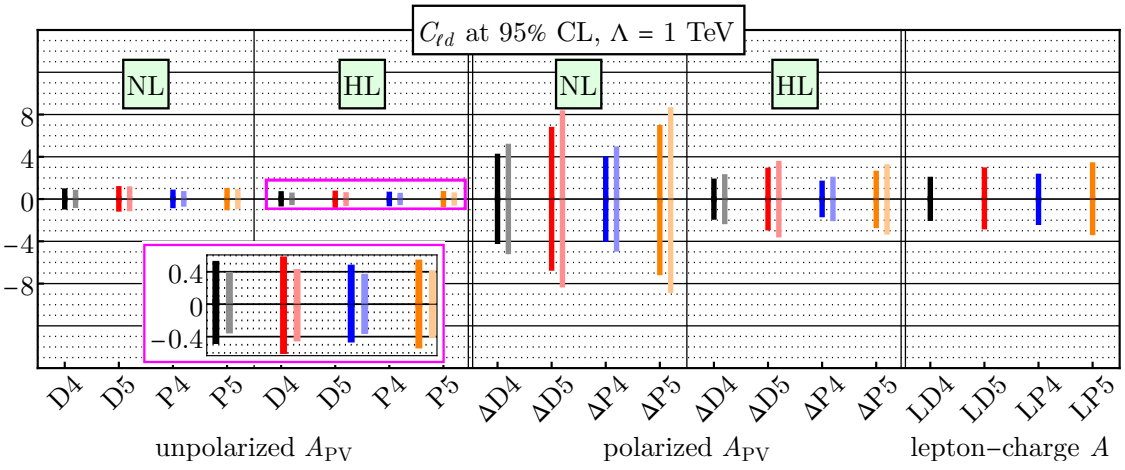

FIG. 27. The same as in Fig. 22 but for $C_{\ell d}$.

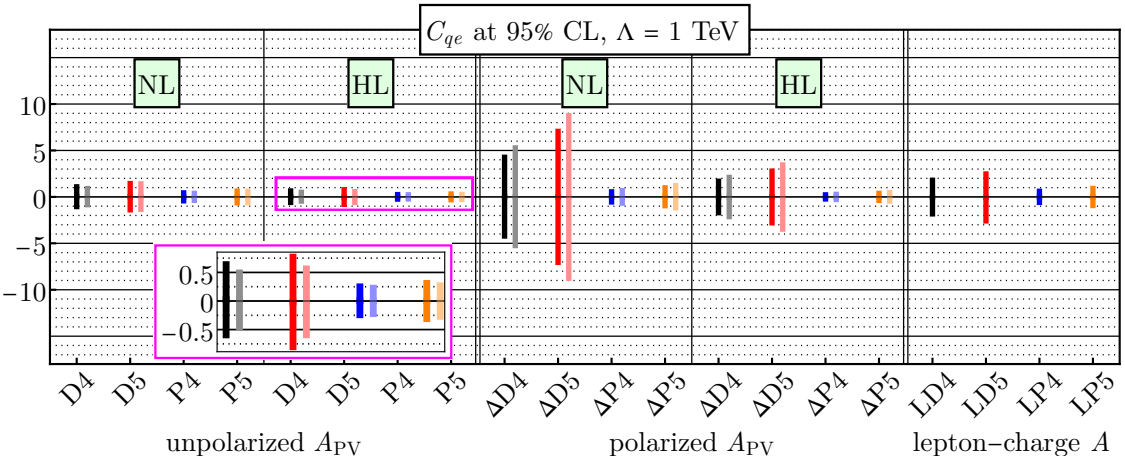

FIG. 28. The same as in Fig. 22 but for C_{qe}.

In Figs. 29–35, we present the effective UV cut-off scales, $\Lambda/\sqrt{C_k}$, with $\Lambda = 1$ TeV, corresponding to the bounds shown in Figs. 22–28. The organization of these plots in terms of asymmetries and data sets is the same as the one in Figs. 22–28. Improved bounds on C_k with the addition of the beam polarization to the fits are equivalent to higher energy scales in the unpolarized PV asymmetries, which are indicated by the lighter columns in the background; on the other hand, weaker bounds from the fits with beam polarization are depicted by the lighter columns in the foreground for the polarized PV asymmetries.

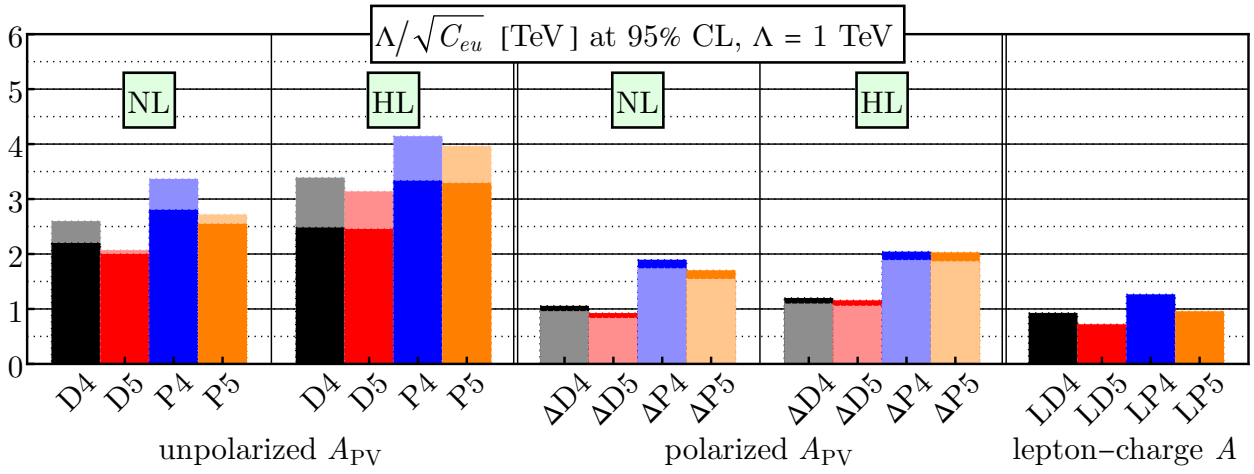

FIG. 29. Effective UV cut-off scales, $\Lambda/\sqrt{C_{eu}}$, defined in terms of the 95% CL bounds on the Wilson coefficient C_{eu} with $\Lambda = 1$ TeV. The darker columns in the foreground of unpolarized PV asymmetries and in the background of polarized PV asymmetries indicate the results of single-parameter fits on the Wilson coefficient, C_{eu}, where the lighter columns in the background of unpolarized PV asymmetries and in the foreground of polarized PV asymmetries denote the results of simultaneous $(1 + 1)$-parameter fits of C_{eu} with the beam polarization, P.

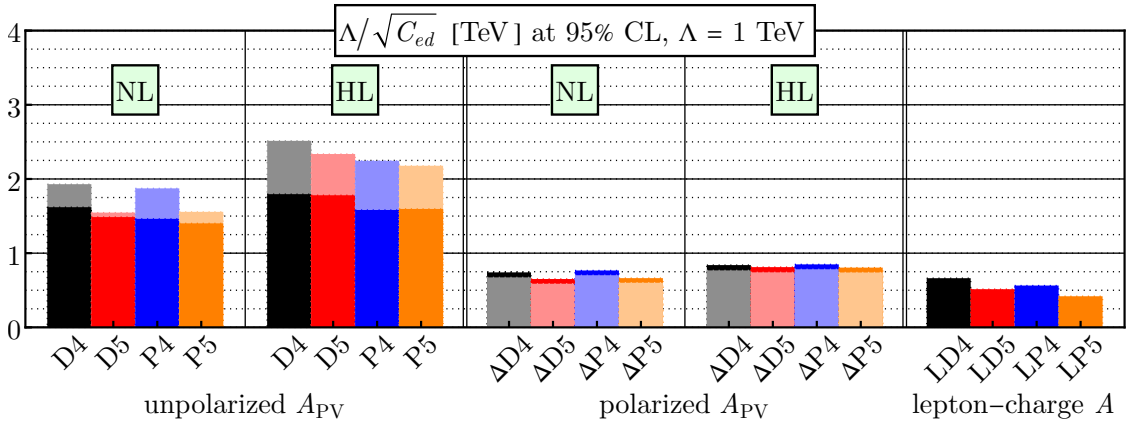

FIG. 30. The same as in Fig. 29 but for C_{ed}.

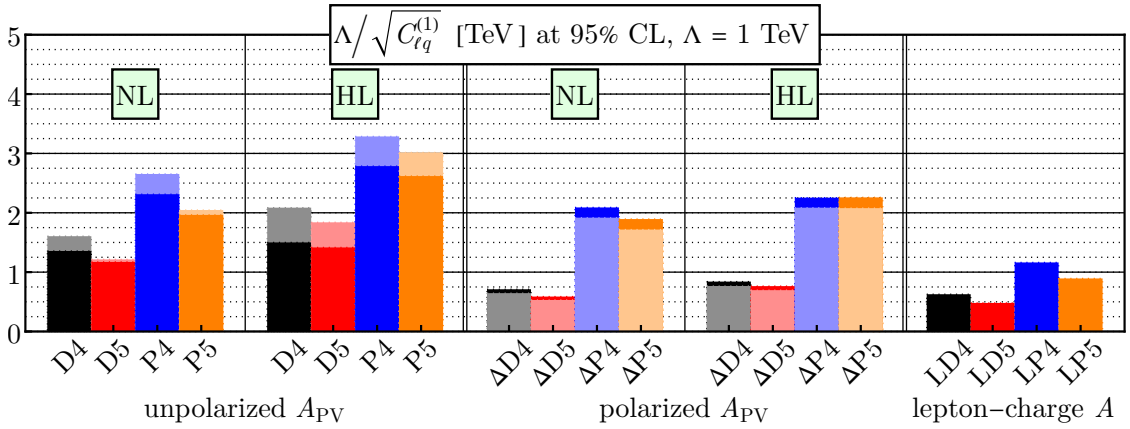

FIG. 31. The same as in Fig. 29 but for $C_{\ell q}^{(1)}$.

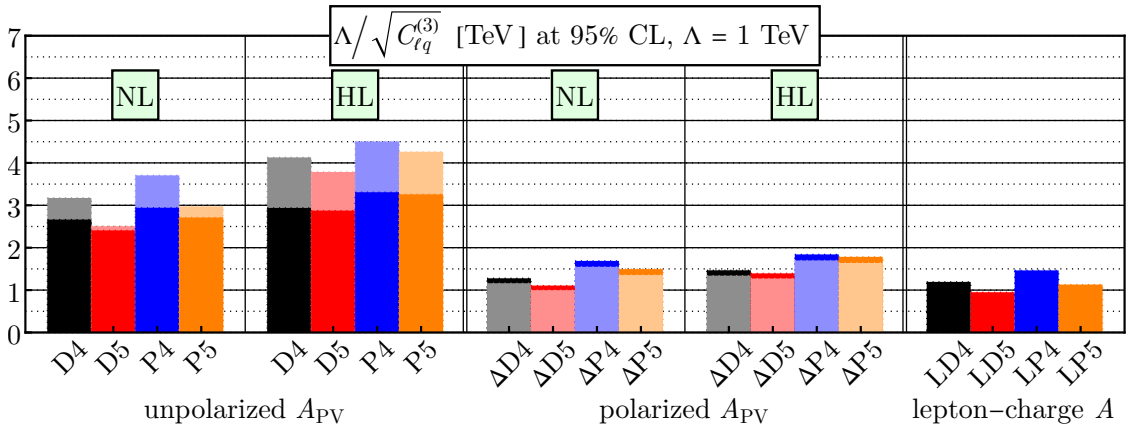

FIG. 32. The same as in Fig. 29 but for $C_{\ell q}^{(3)}$.

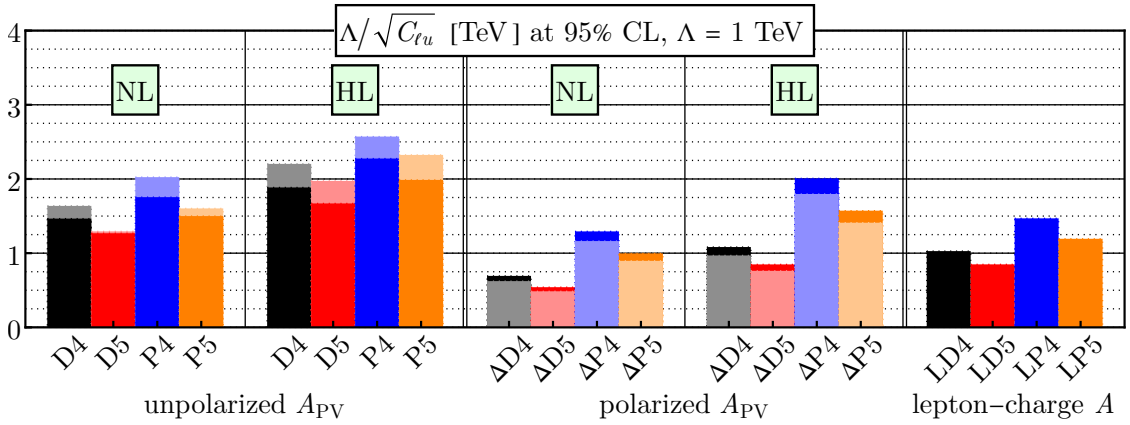

FIG. 33. The same as in Fig. 29 but for $C_{\ell u}$.

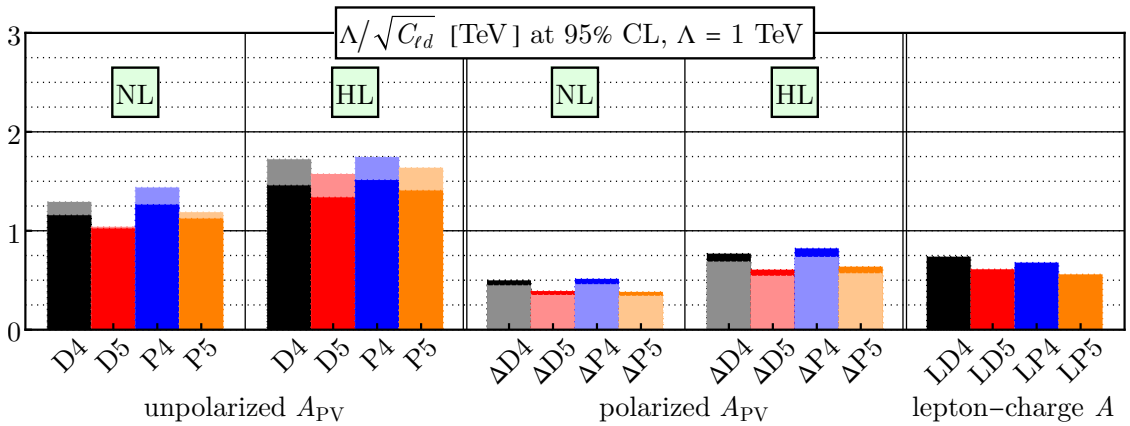

FIG. 34. The same as in Fig. 29 but for $C_{\ell d}$.

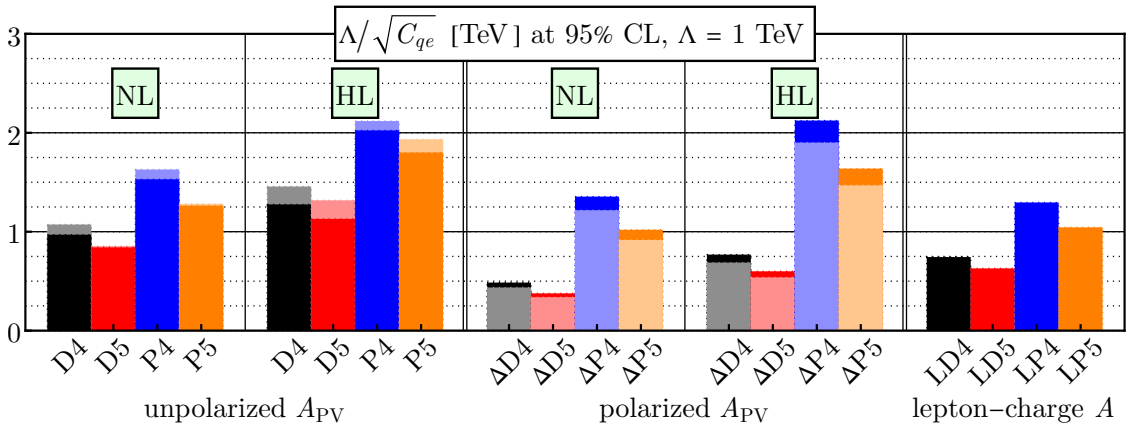

FIG. 35. The same as in Fig. 29 but for C_{qe}.

925 For completeness, we summarize in Table V the fitting results of all the seven Wilson
926 coefficients with the four families of data sets of interest. The values indicated in this table

are the 95% CL bounds around zero.

TABLE V. 95% CL bounds of all the seven Wilson coefficients around zero at $\Lambda = 1$ TeV with the four families of data sets, D4, D5, P4, and P5 in various configurations.

eD				C_{eu}	C_{ed}	$C_{\ell q}^{(1)}$	$C_{\ell q}^{(3)}$	$C_{\ell u}$	$C_{\ell d}$	C_{qe}
10 GeV × 137 GeV, 100 fb^{-1}	unpol. $A_{\rm PV}$	NL	without P	0.21	0.38	0.54	0.14	0.47	0.75	1.06
			with P	0.15	0.27	0.39	0.10	0.37	0.60	0.87
		HL	without P	0.16	0.31	0.44	0.12	0.28	0.47	0.61
			with P	0.09	0.16	0.23	0.06	0.21	0.34	0.47
	pol. $A_{\rm PV}$	NL	without P	0.90	1.82	1.98	0.61	2.08	4.02	4.24
			with P	1.08	2.20	2.40	0.74	2.57	4.97	5.26
		HL	without P	0.69	1.42	1.41	0.46	0.85	1.69	1.69
			with P	0.83	1.70	1.70	0.56	1.06	2.11	2.11
	lepton−charge A			1.16	2.28	2.54	0.70	0.95	1.84	1.81
18 GeV × 137 GeV, 15.4 fb^{-1}	unpol. $A_{\rm PV}$	NL	without P	0.25	0.45	0.73	0.17	0.62	0.96	1.41
			with P	0.23	0.42	0.68	0.16	0.60	0.92	1.37
		HL	without P	0.17	0.32	0.50	0.12	0.36	0.56	0.78
			with P	0.10	0.18	0.30	0.07	0.26	0.40	0.58
	pol. $A_{\rm PV}$	NL	without P	1.17	2.36	2.89	0.82	3.41	6.55	7.05
			with P	1.43	2.89	3.55	1.00	4.23	8.12	8.75
		HL	without P	0.75	1.52	1.71	0.51	1.39	2.72	2.78
			with P	0.89	1.82	2.06	0.62	1.72	3.37	3.46
	lepton−charge A			1.92	3.78	4.32	1.11	1.39	2.68	2.53

ep				C_{eu}	C_{ed}	$C_{\ell q}^{(1)}$	$C_{\ell q}^{(3)}$	$C_{\ell u}$	$C_{\ell d}$	C_{qe}
10 GeV × 275 GeV, 100 fb^{-1}	unpol. $A_{\rm PV}$	NL	without P	0.13	0.47	0.19	0.12	0.32	0.63	0.43
			with P	0.09	0.28	0.14	0.07	0.24	0.48	0.38
		HL	without P	0.09	0.40	0.13	0.09	0.19	0.44	0.24
			with P	0.06	0.20	0.09	0.05	0.15	0.33	0.22
	pol. $A_{\rm PV}$	NL	without P	0.28	1.70	0.23	0.35	0.60	3.79	0.54
			with P	0.33	2.03	0.27	0.42	0.74	4.72	0.68
		HL	without P	0.24	1.39	0.20	0.29	0.25	1.48	0.22
			with P	0.28	1.63	0.23	0.35	0.31	1.85	0.28
	lepton−charge A			0.62	3.15	0.74	0.47	0.46	2.17	0.60
18 GeV × 275 GeV, 15.4 fb^{-1}	unpol. $A_{\rm PV}$	NL	without P	0.15	0.51	0.26	0.14	0.44	0.79	0.63
			with P	0.14	0.41	0.24	0.11	0.39	0.71	0.61
		HL	without P	0.09	0.39	0.15	0.09	0.25	0.50	0.31
			with P	0.06	0.21	0.11	0.06	0.19	0.37	0.27
	pol. $A_{\rm PV}$	NL	without P	0.34	2.27	0.28	0.45	1.01	6.85	0.96
			with P	0.42	2.77	0.34	0.55	1.24	8.53	1.19
		HL	without P	0.24	1.54	0.20	0.31	0.40	2.46	0.37
			with P	0.29	1.83	0.23	0.37	0.50	3.07	0.46
	lepton−charge A			1.09	5.69	1.25	0.78	0.70	3.19	0.92

2. Fits of two Wilson coefficients

In this section, we present the complete set of confidence ellipses for all possible pairs of Wilson coefficients that we take into account in this work. The ellipses are plotted at 95% CL and $\Lambda = 1$ TeV. For the self-containment of this section, we remind the following abbreviations for the EIC preliminary data sets:

- electron PV asymmetries $A_{PV}^{(e)}$ of unpolarized deuteron:
  - D4: eD 10 GeV × 137 GeV, 100 fb^{-1}
  - D5: eD 18 GeV × 137 GeV, 15.4 fb^{-1}
- electron PV asymmetries $A_{PV}^{(e)}$ of unpolarized proton:
  - P4: ep 10 GeV × 275 GeV, 100 fb^{-1}

938 – P5: *ep* 18 GeV × 275 GeV, 15.4 fb^{-1}

939 • hadron PV asymmetries $A_{PV}^{(H)}$ with unpolarized electron: ΔD4, ΔD5, ΔP4, and ΔP5
940 with the same energy and luminosity configuration as the corresponding D- and P-sets

941 • unpolarized electron-positron asymmetries of unpolarized hadrons $A_{LC,H}$: LD4, LD5,
942 LP4, and LP5 with the same energy configuration as the corresponding D- and P-sets,
943 but the luminosity of the positron beam is assumed to be 10 times smaller than that
944 of the electron beam.

945 We refer to electron PV asymmetries collectively as unpolarized $A_{\rm PV}$, hadron PV asym-
946 metries as polarized $A_{\rm PV}$, and electron-positron asymmetries as lepton-charge A. Data sets
947 with the label NL or HL indicate the luminosity: The nominal luminosity ("NL") refers to
948 the annual integrated luminosity of Table 10.1 of YR [6]. The high luminosity ("HL") is
949 assumed to be 10 times higher than the nominal one and requires a luminosity upgrade of
950 the EIC.

951 Each figure consists of four panels, containing one of each of the four families of data sets,
952 namely D4, D5, P4, and P5. We plot the fits from polarized and unpolarized PV asymmetry
953 data sets in both nominal- and high-luminosity scenarios for comparison. We remark that
954 the ellipses for the polarized and unpolarized PV asymmetry data sets indicate the results of
955 simultaneous fits on Wilson coefficients with the beam-polarization parameter, P, in light of
956 significant improvements in the results with unpolarized PV asymmetries. Moreover, we in-
957 clude, when available, fitted results for the LHC Drell-Yan data, adapting from [23] and [24].
958

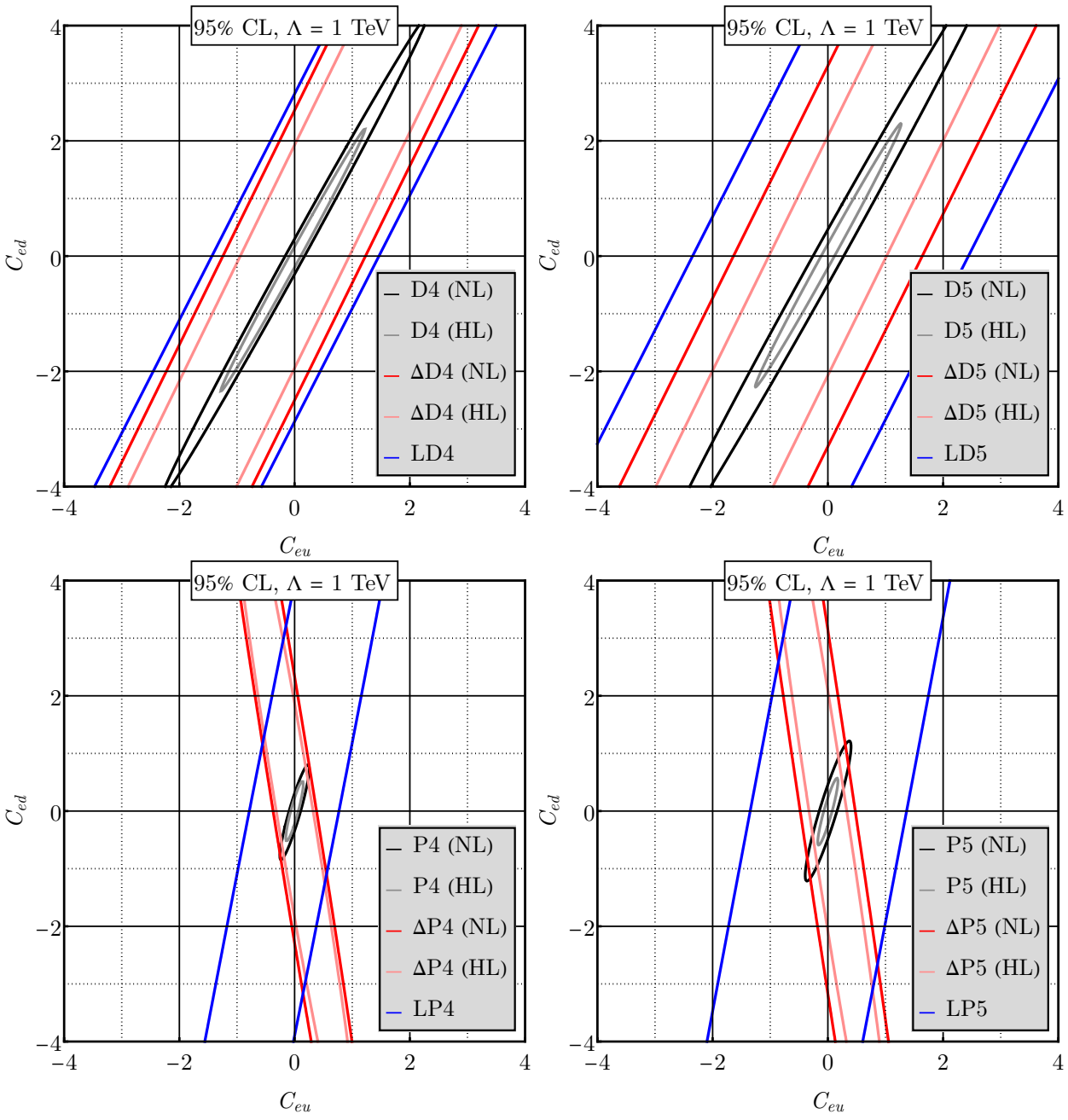

FIG. 36. 95% CL ellipses for the Wilson coefficients C_{eu} and C_{ed} using the families of data sets D4, D5, P4, and P5 at $\Lambda = 1$ TeV.

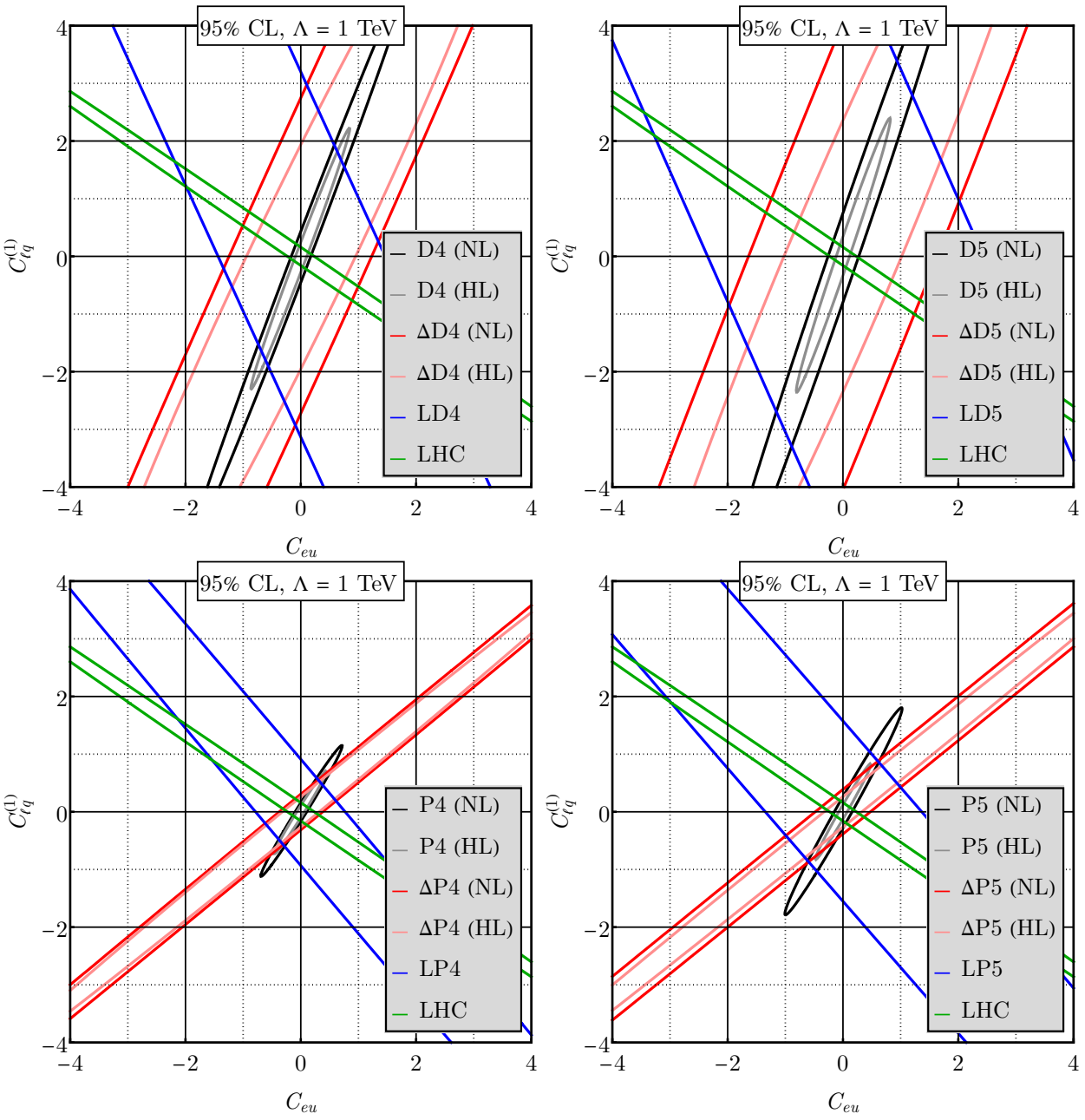

FIG. 37. The same as in Fig. 36 but for C_{eu} and $C_{\ell q}^{(1)}$.

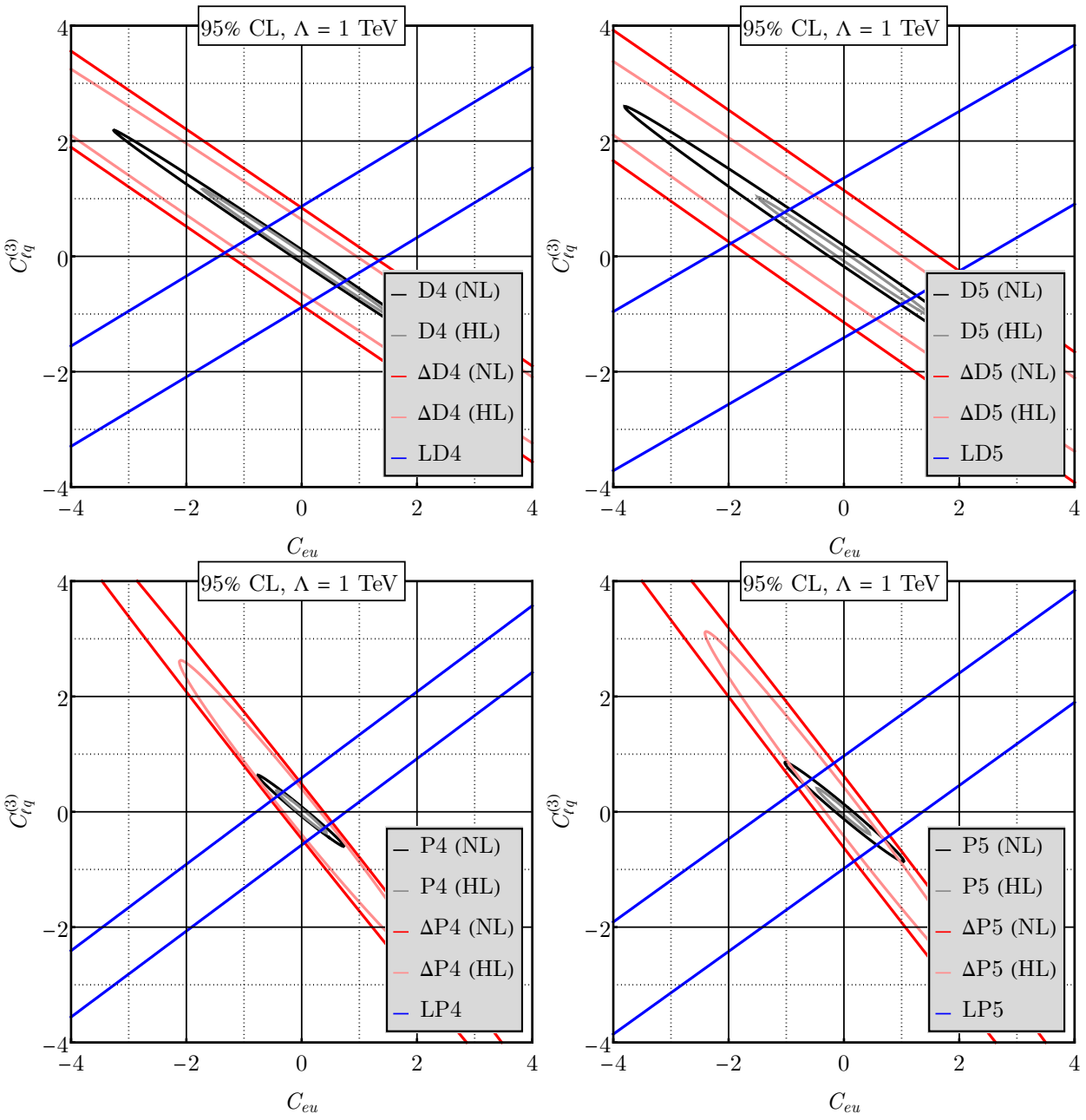

FIG. 38. The same as in Fig. 36 but for C_{eu} and $C_{\ell q}^{(3)}$.

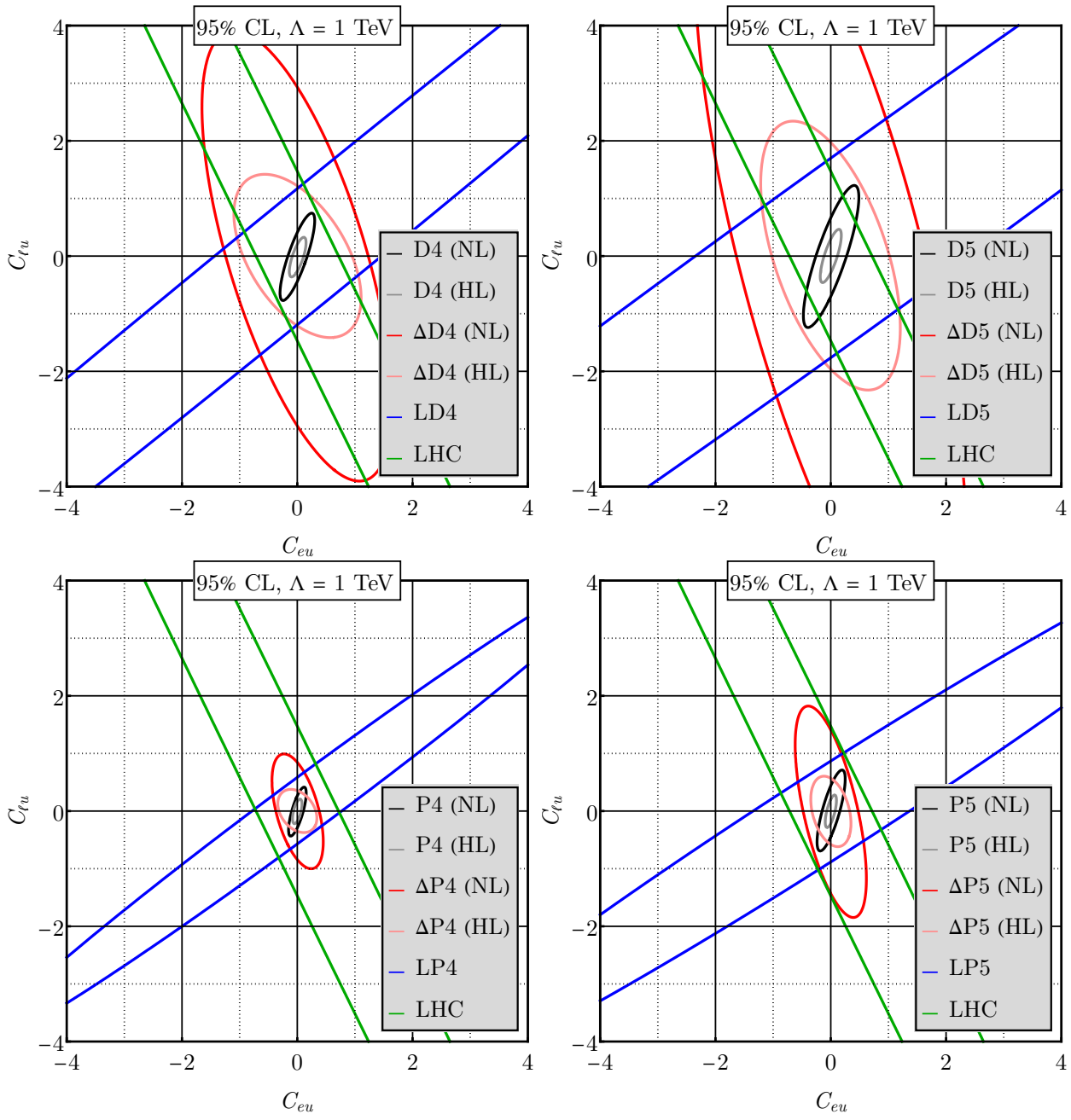

FIG. 39. The same as in Fig. 36 but for C_{eu} and $C_{\ell u}$.

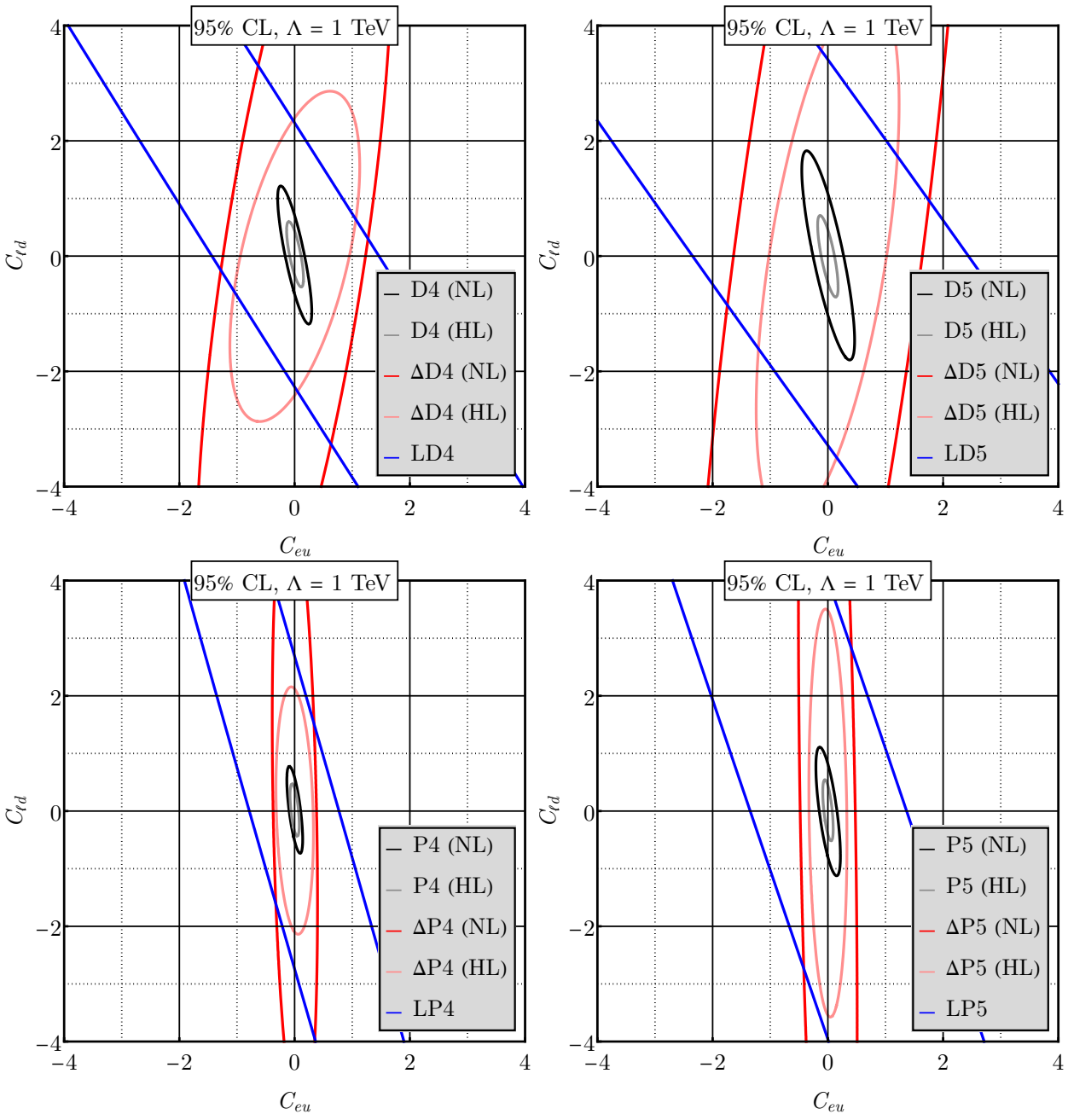

FIG. 40. The same as in Fig. 36 but for C_{eu} and $C_{\ell d}$.

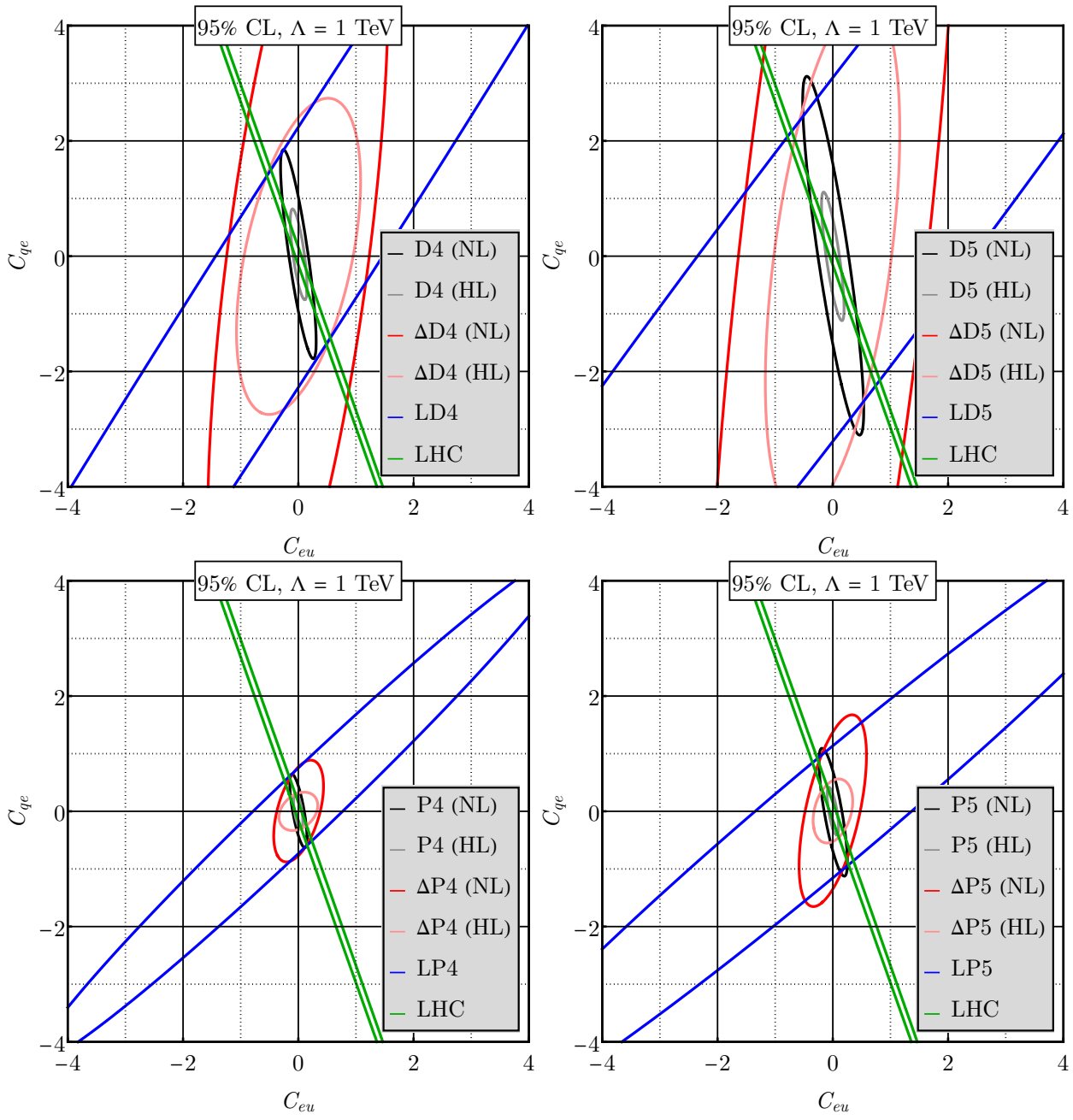

FIG. 41. The same as in Fig. 36 but for C_{eu} and C_{qe}.

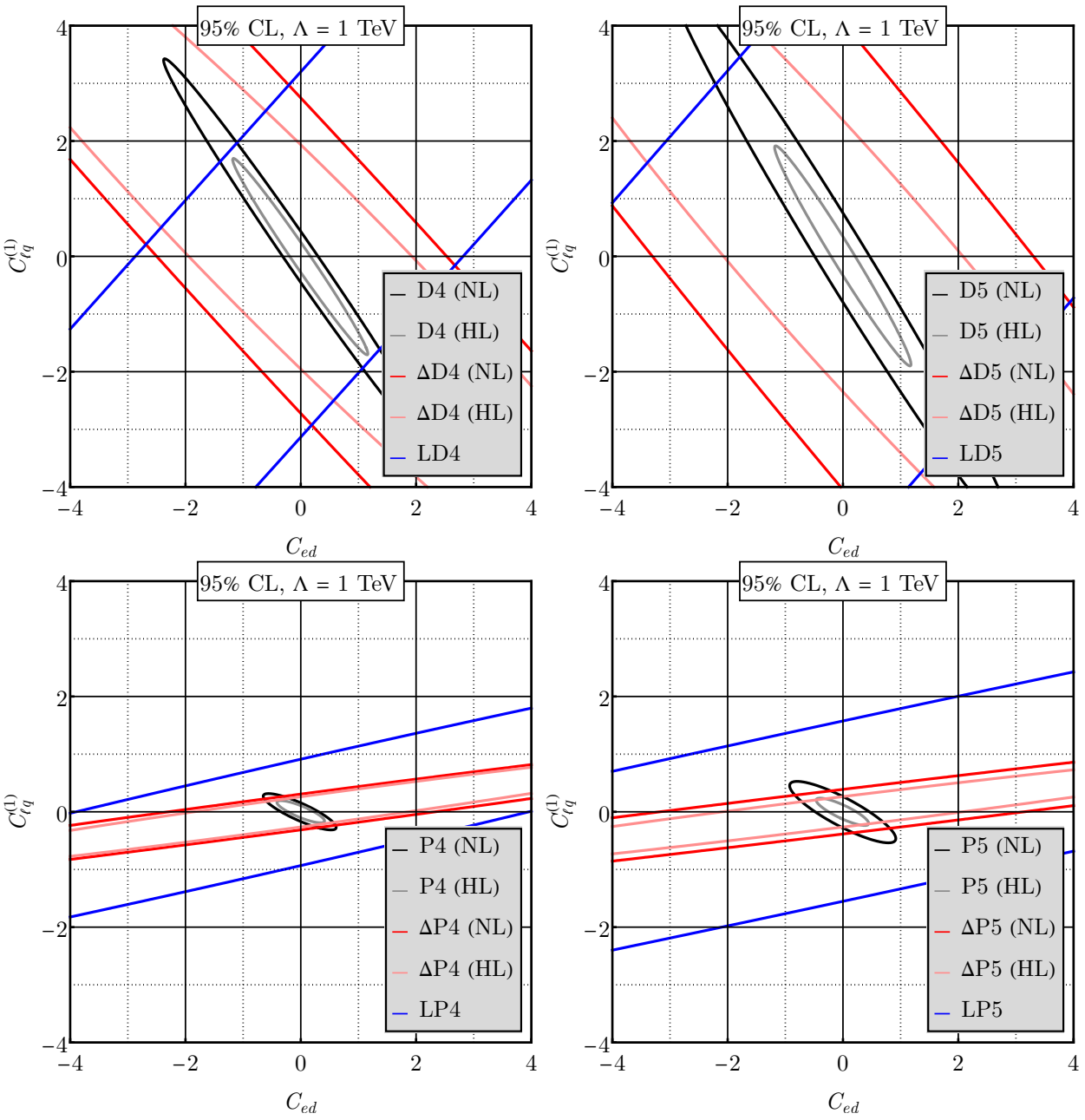

FIG. 42. The same as in Fig. 36 but for C_{ed} and $C_{\ell q}^{(1)}$.

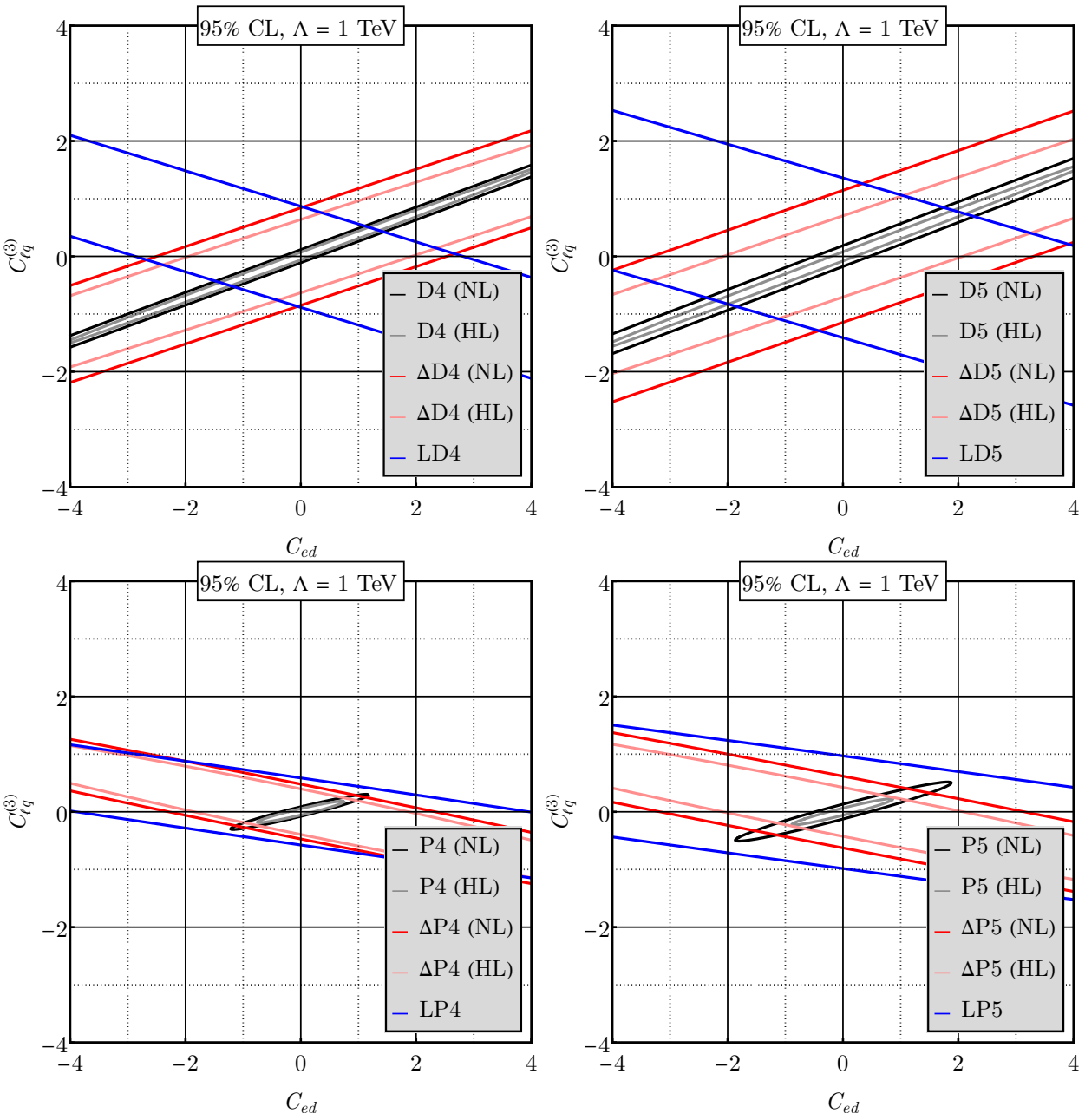

FIG. 43. The same as in Fig. 36 but for C_{ed} and $C_{\ell q}^{(3)}$.

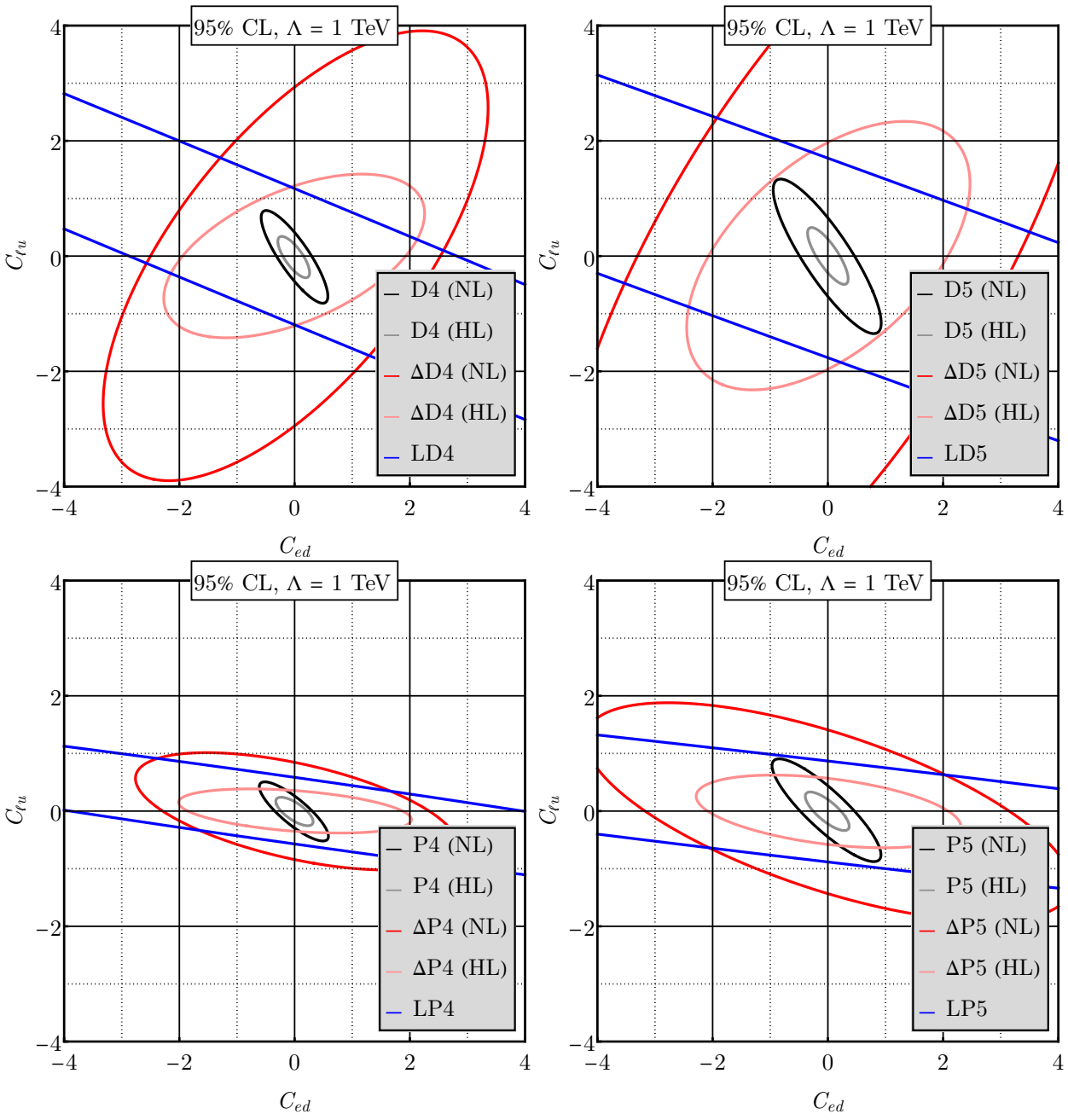

FIG. 44. The same as in Fig. 36 but for C_{ed} and $C_{\ell u}$.

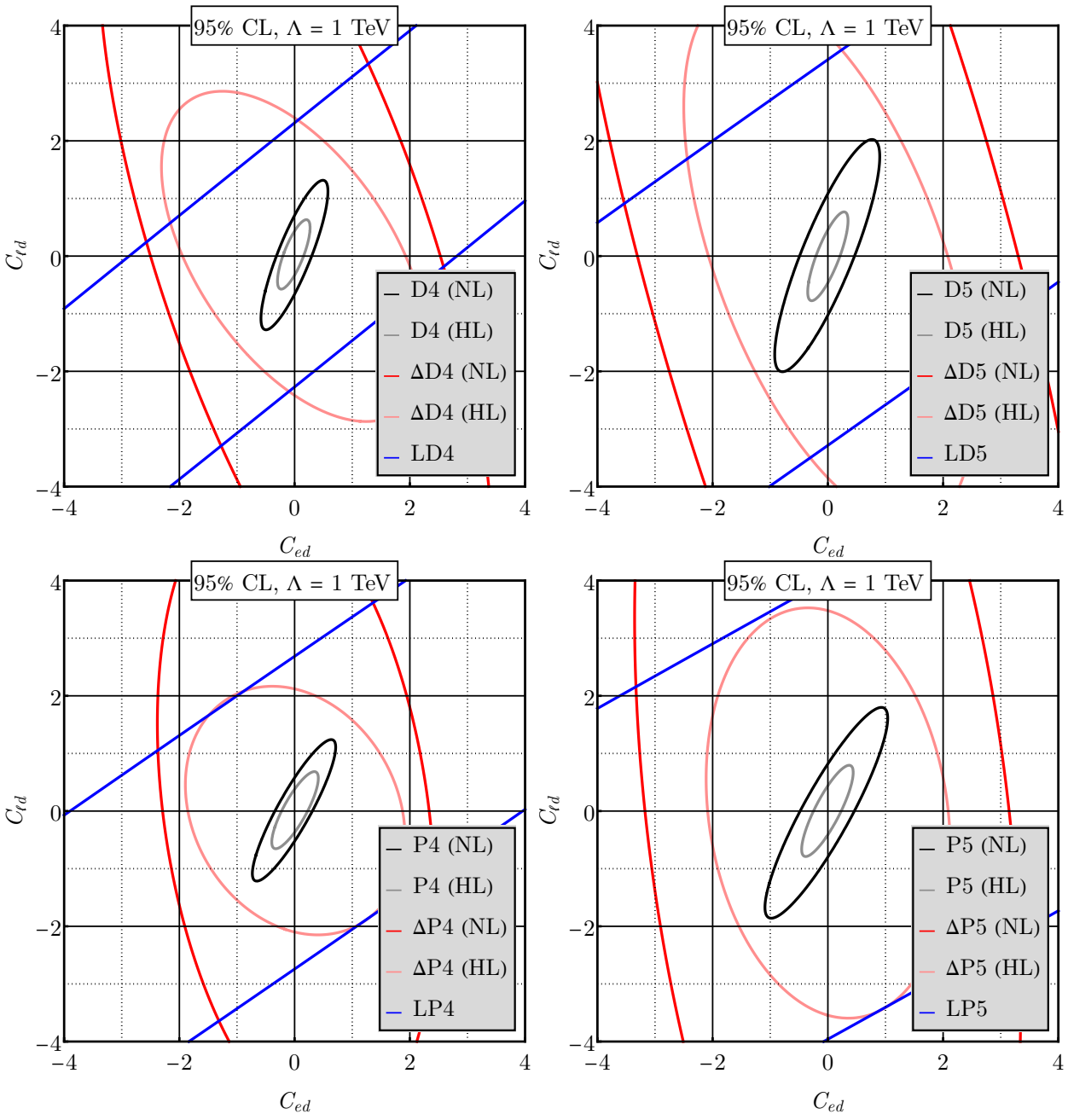

FIG. 45. The same as in Fig. 36 but for C_{ed} and $C_{\ell d}$.

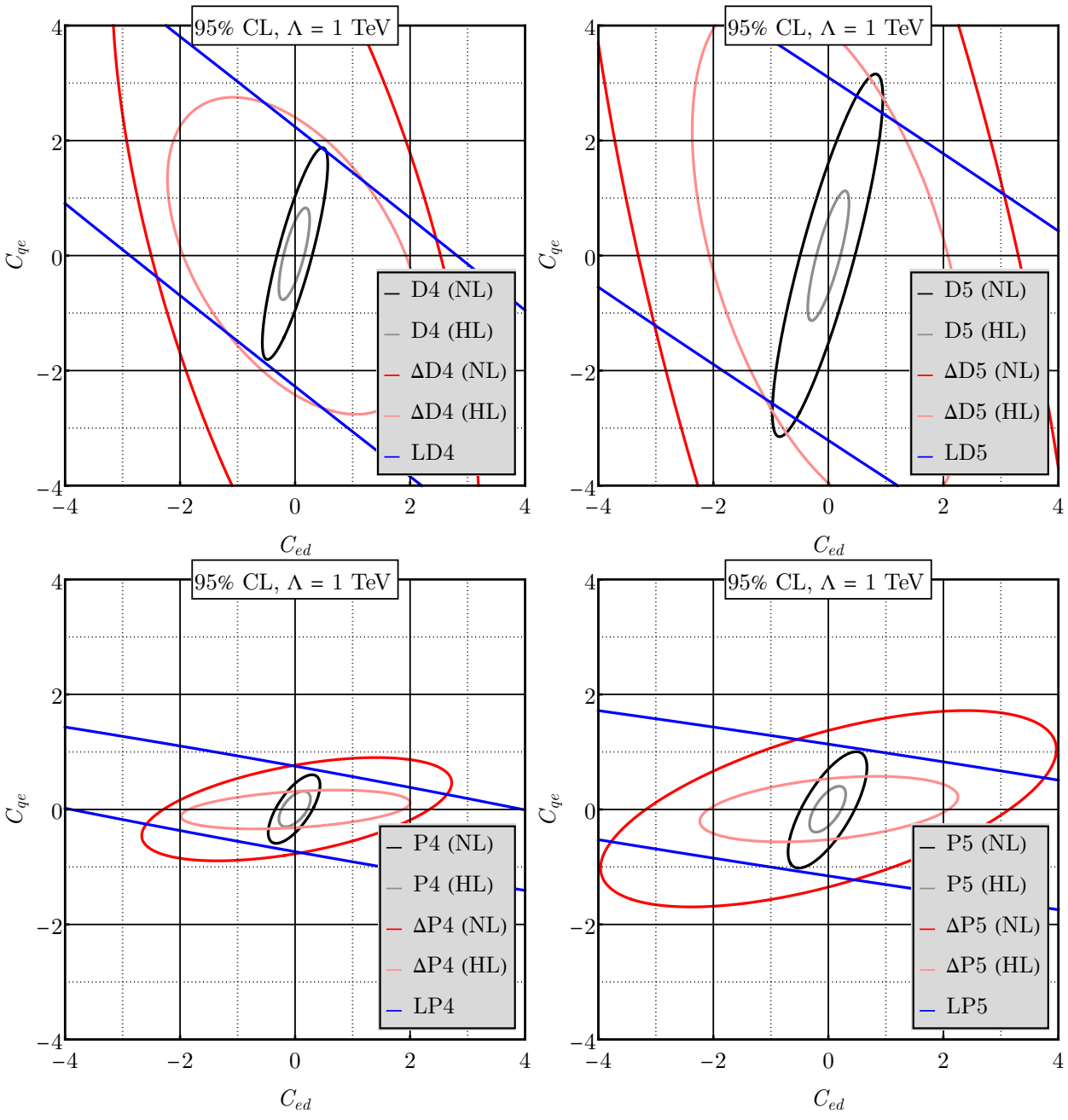

FIG. 46. The same as in Fig. 36 but for C_{ed} and C_{qe}.

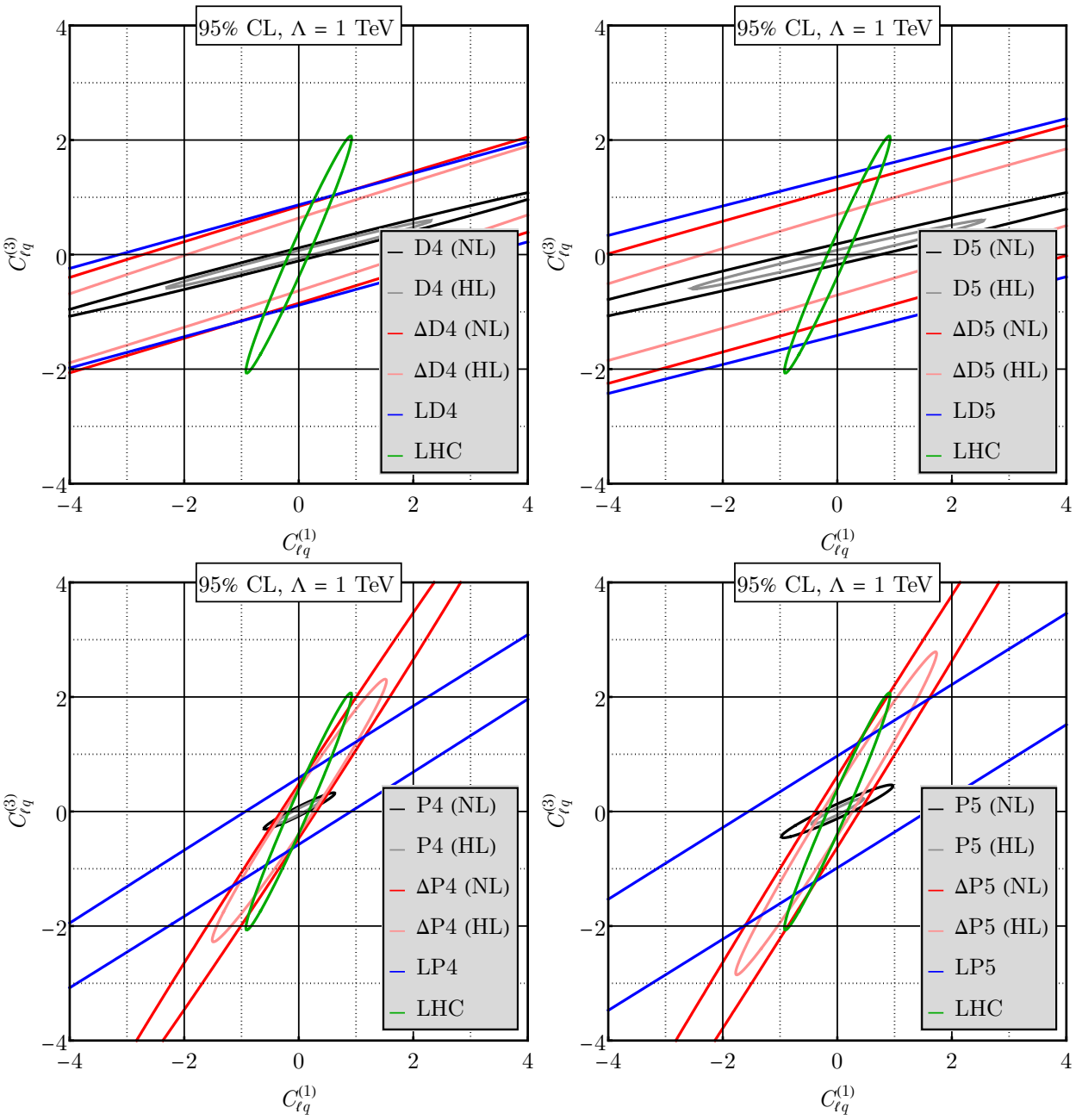

FIG. 47. The same as in Fig. 36 but for $C_{\ell q}^{(1)}$ and $C_{\ell q}^{(3)}$.

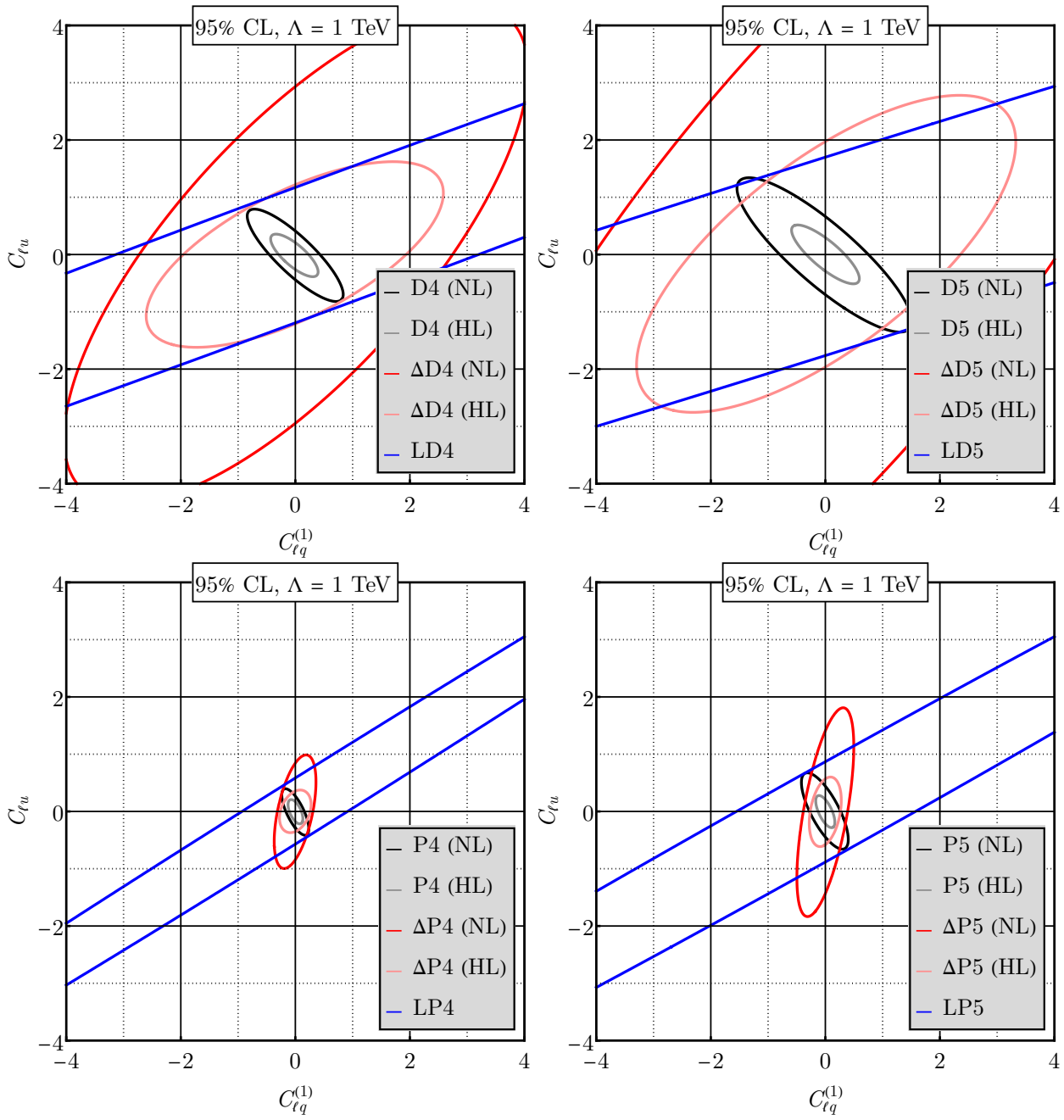

FIG. 48. The same as in Fig. 36 but for $C_{\ell q}^{(1)}$ and $C_{\ell u}$.

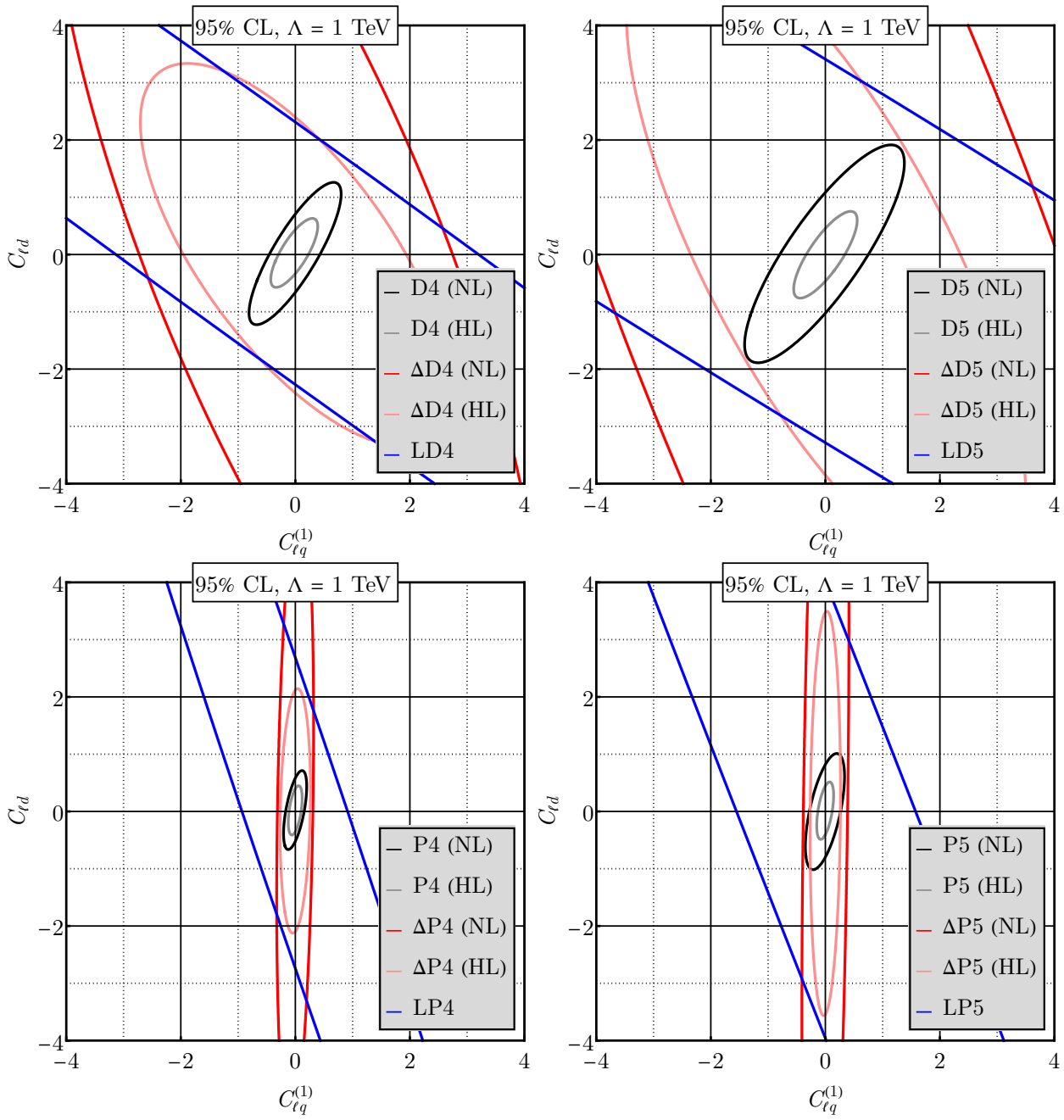

FIG. 49. The same as in Fig. 36 but for $C_{\ell q}^{(1)}$ and $C_{\ell d}$.

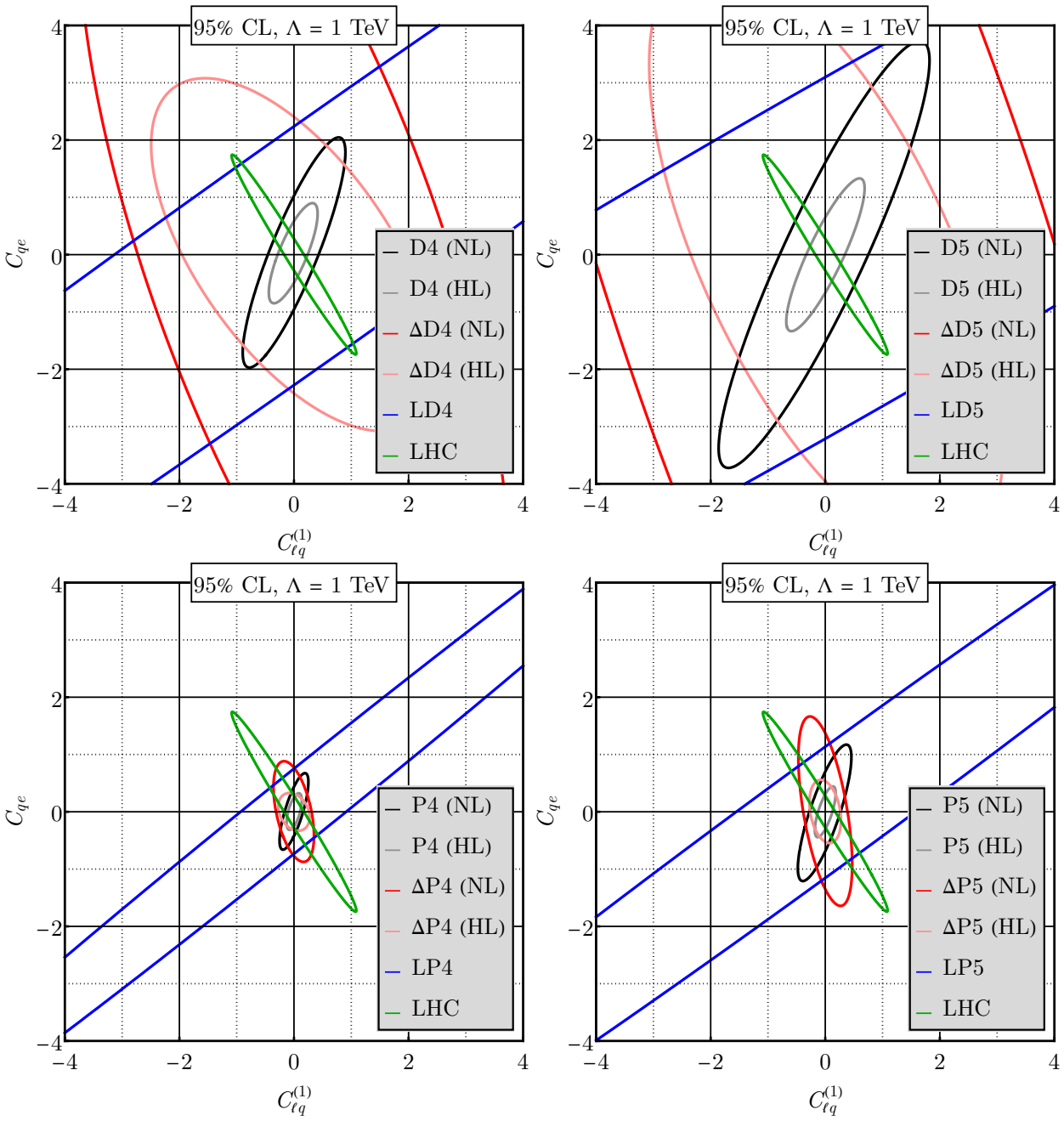

FIG. 50. The same as in Fig. 36 but for $C_{\ell q}^{(1)}$ and C_{qe}.

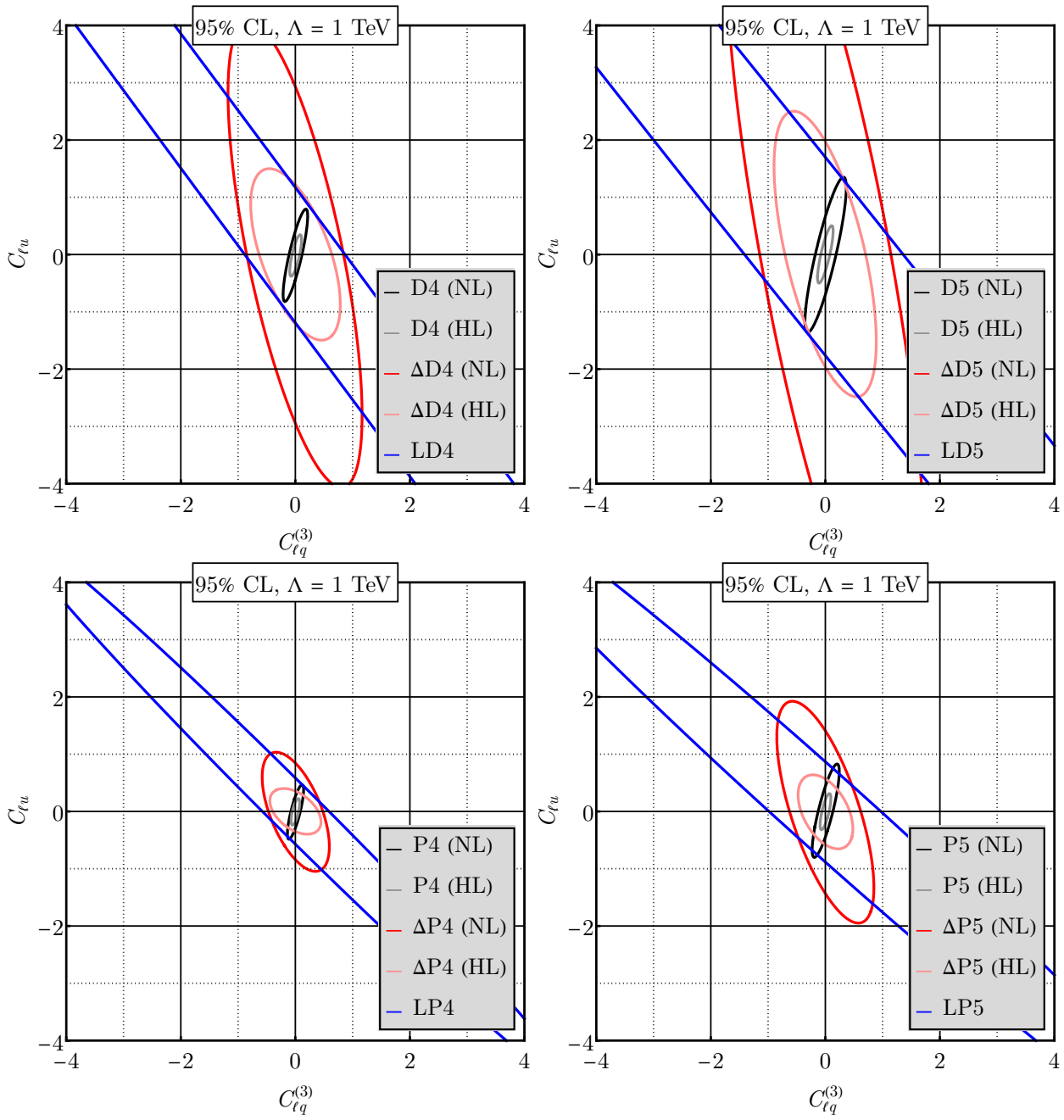

FIG. 51. The same as in Fig. 36 but for $C_{\ell q}^{(3)}$ and $C_{\ell u}$.

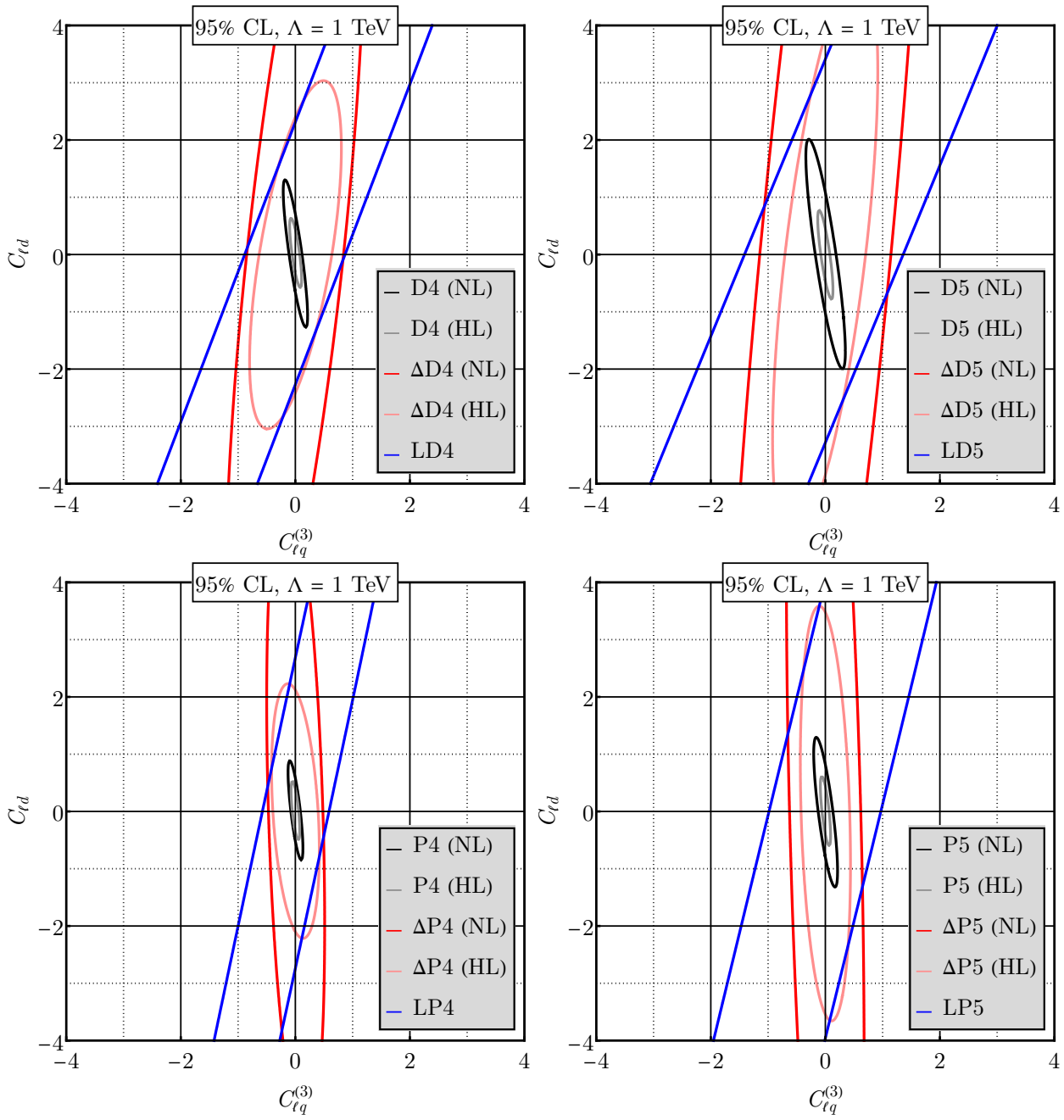

FIG. 52. The same as in Fig. 36 but for $C_{\ell q}^{(3)}$ and $C_{\ell d}$.

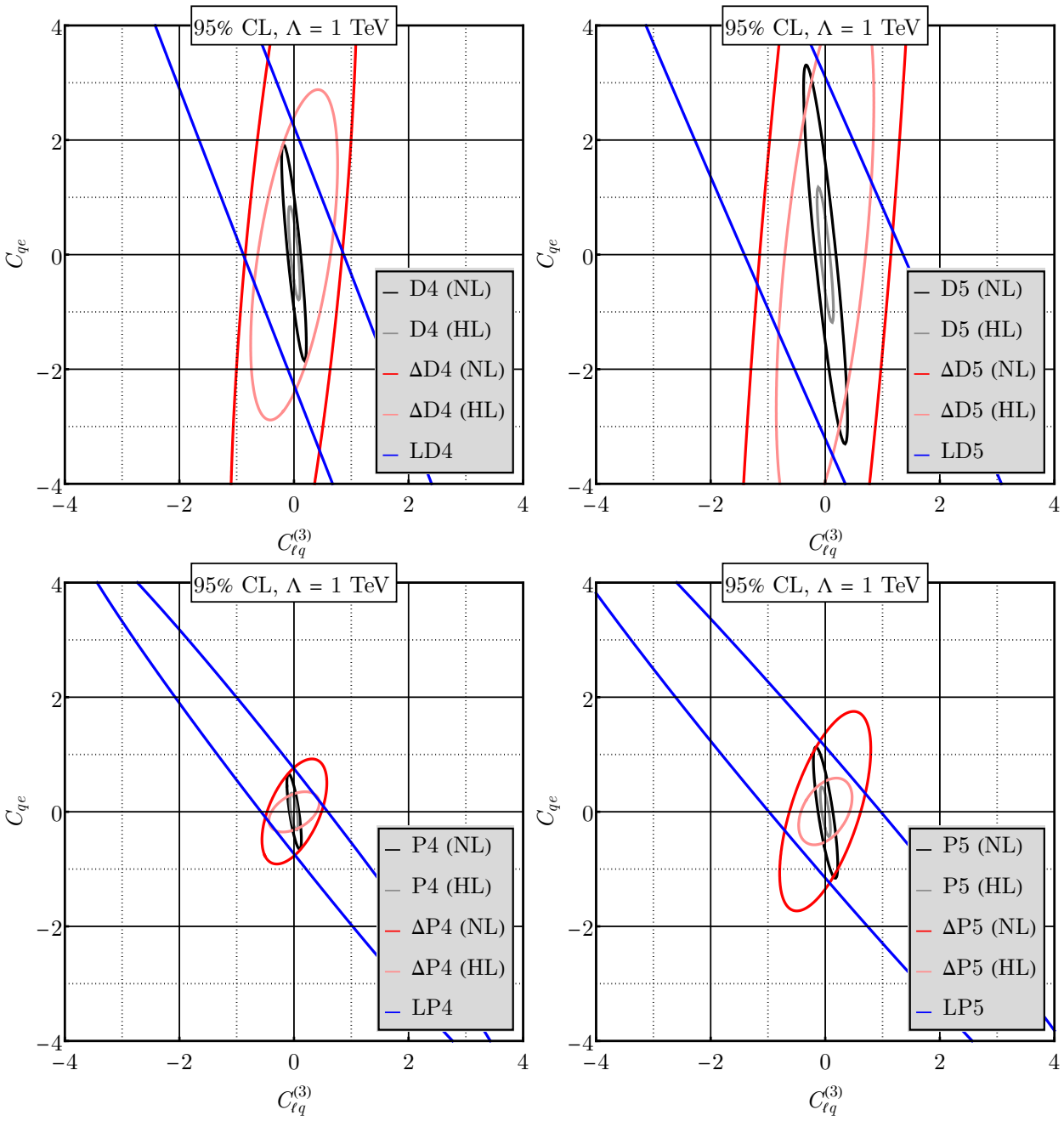

FIG. 53. The same as in Fig. 36 but for $C_{\ell q}^{(3)}$ and C_{qe}.

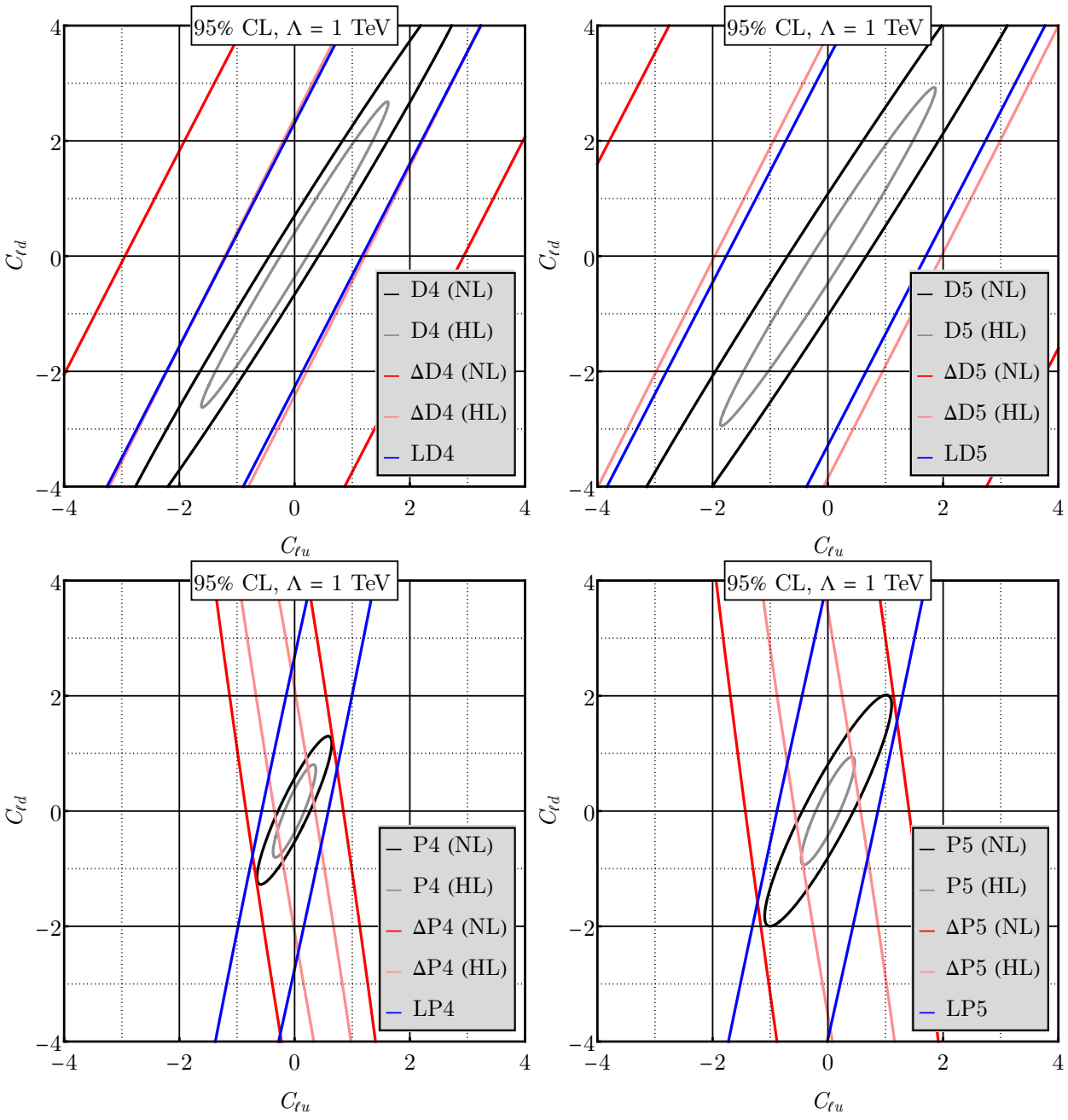

FIG. 54. The same as in Fig. 36 but for $C_{\ell u}$ and $C_{\ell d}$.

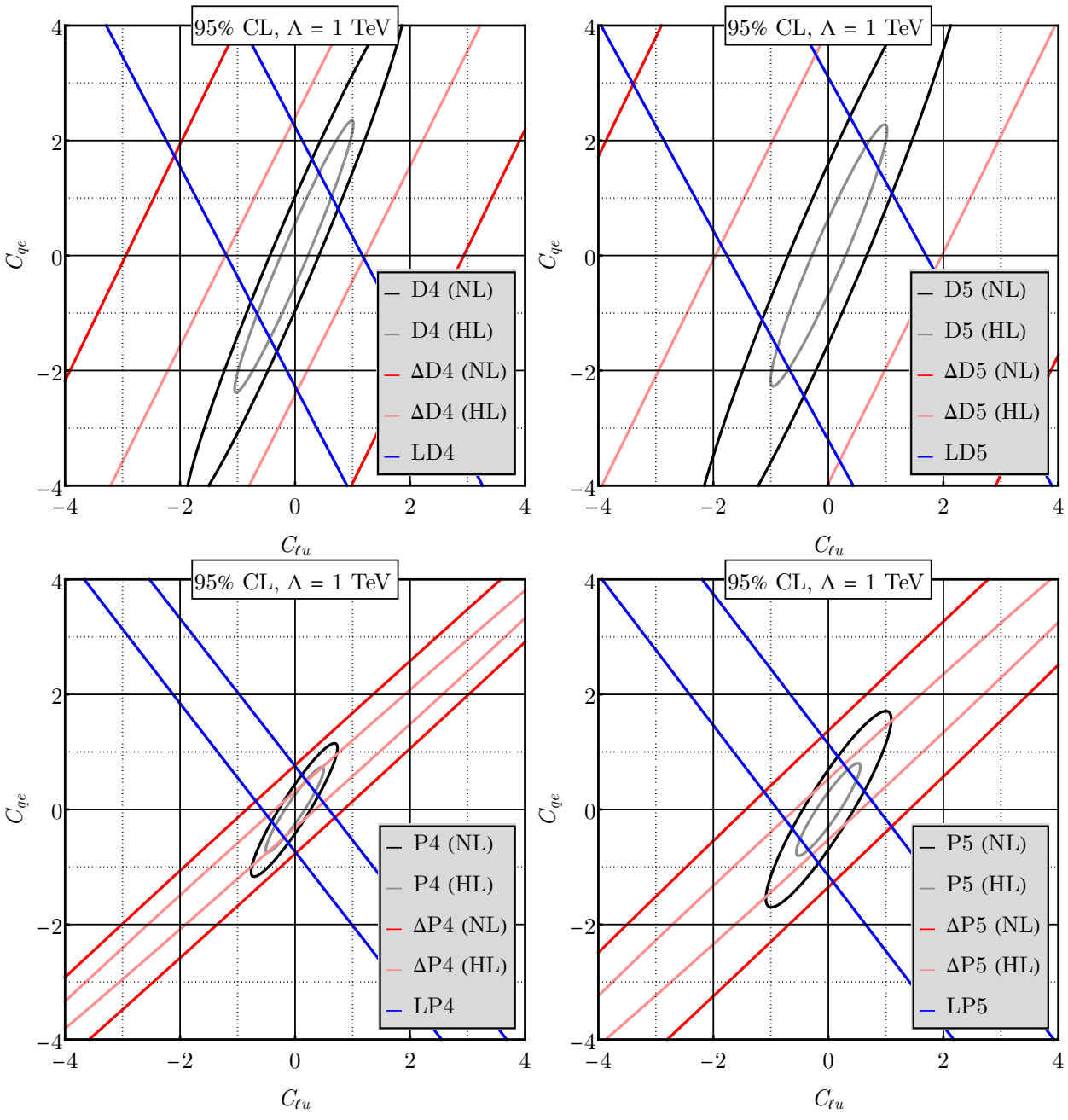

FIG. 55. The same as in Fig. 36 but for $C_{\ell u}$ and C_{qe}.

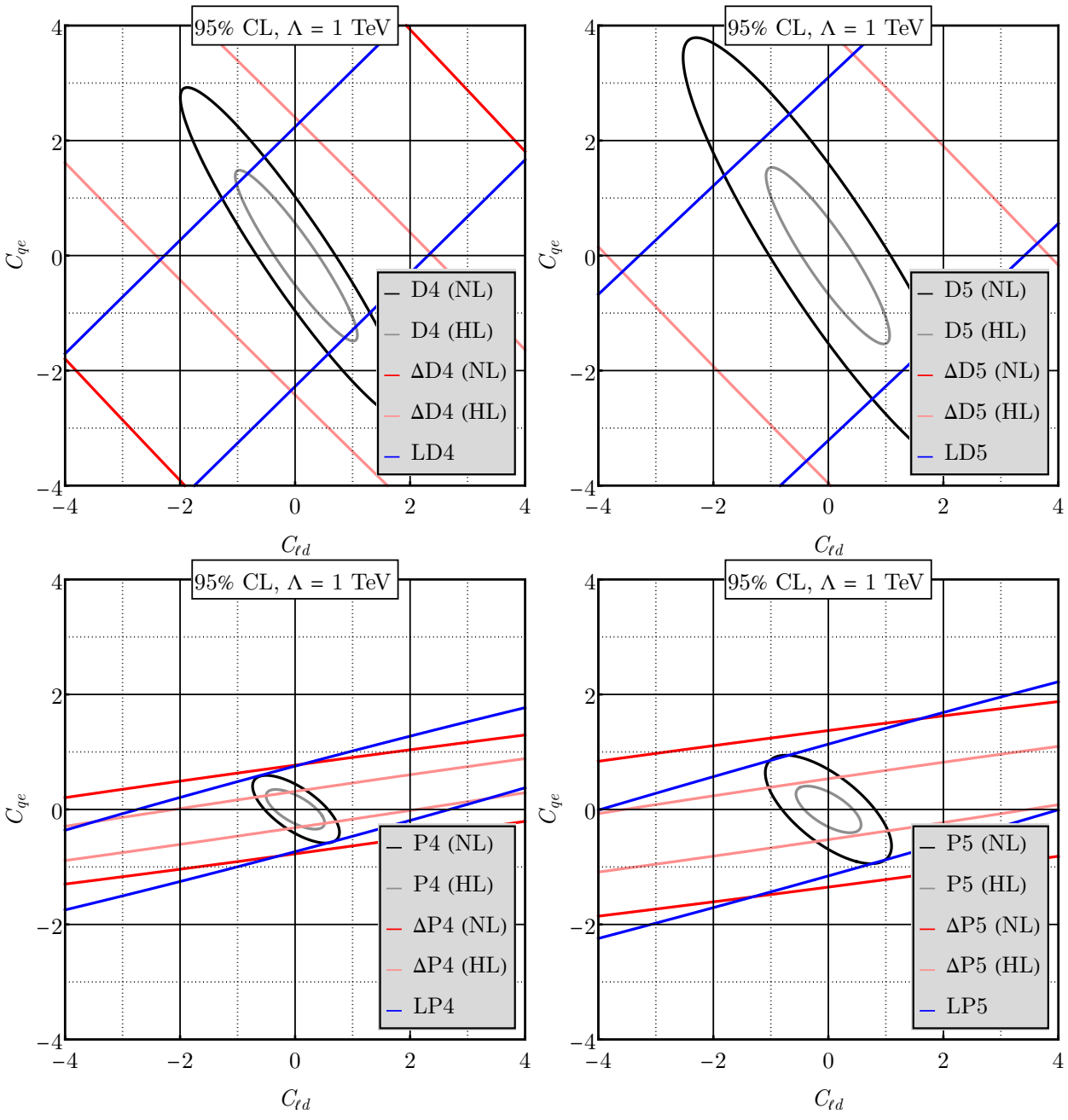

FIG. 56. The same as in Fig. 36 but for $C_{\ell d}$ and C_{qe}.

For completeness, we show the correlations of Wilson coefficients in Figs. 57–60.

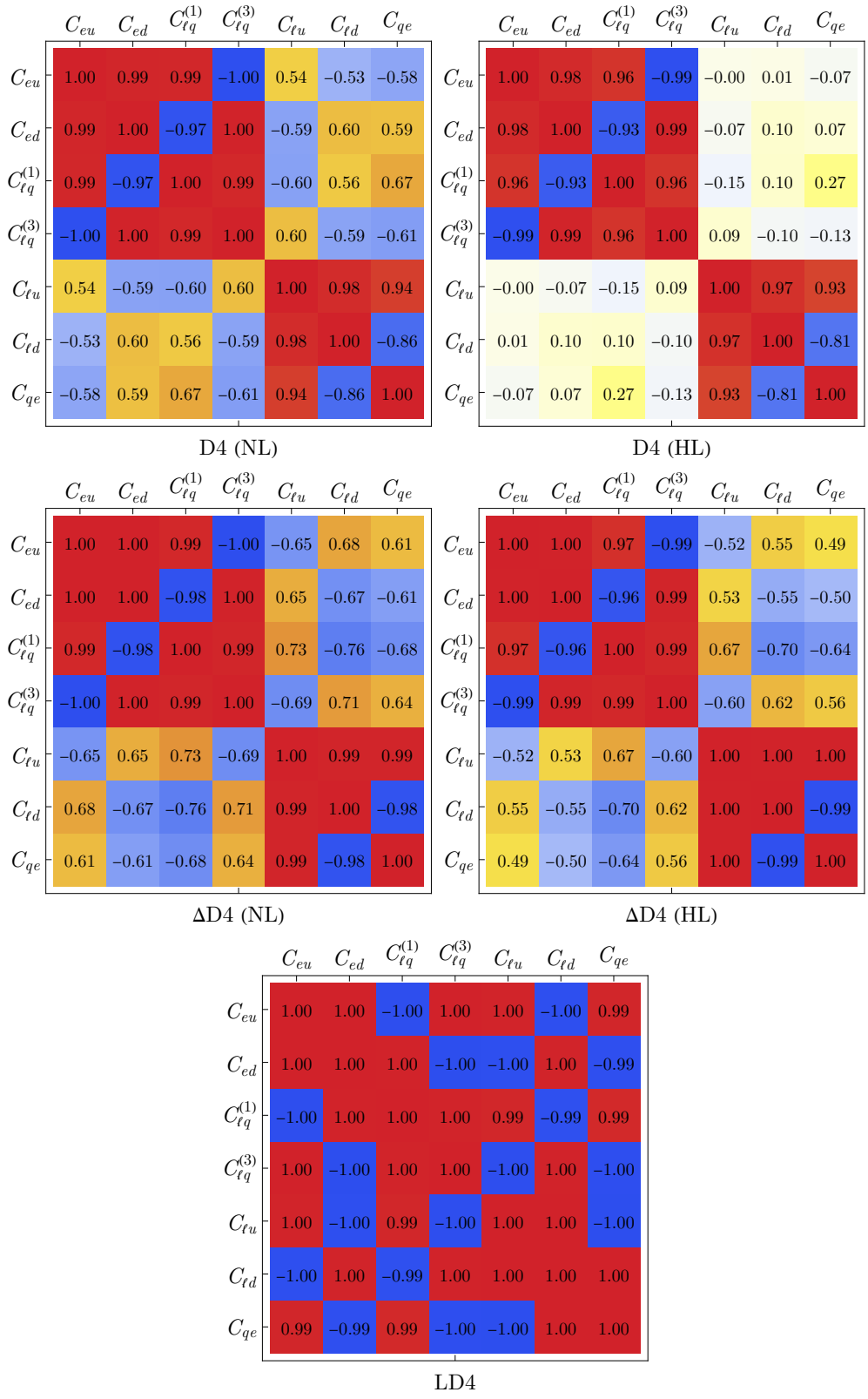

FIG. 57. The correlation table of Wilson coefficients in the D4 data family. The off-diagonal entries are from the results of simultaneous fits of the $(2+1)$ fit of two Wilson coefficients plus the beam-polarization parameter P.

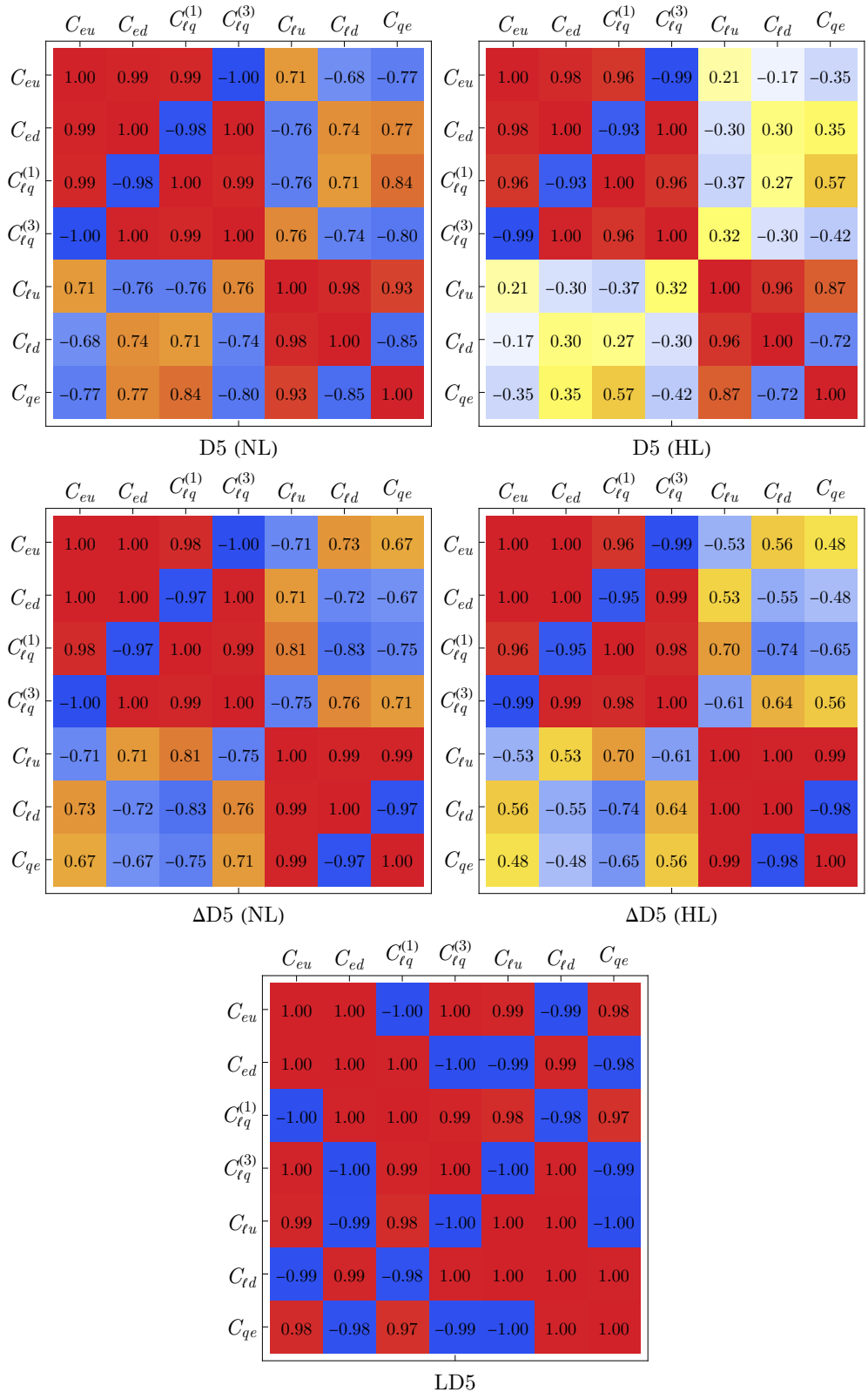

FIG. 58. The same as in Fig. 57 but for D5.

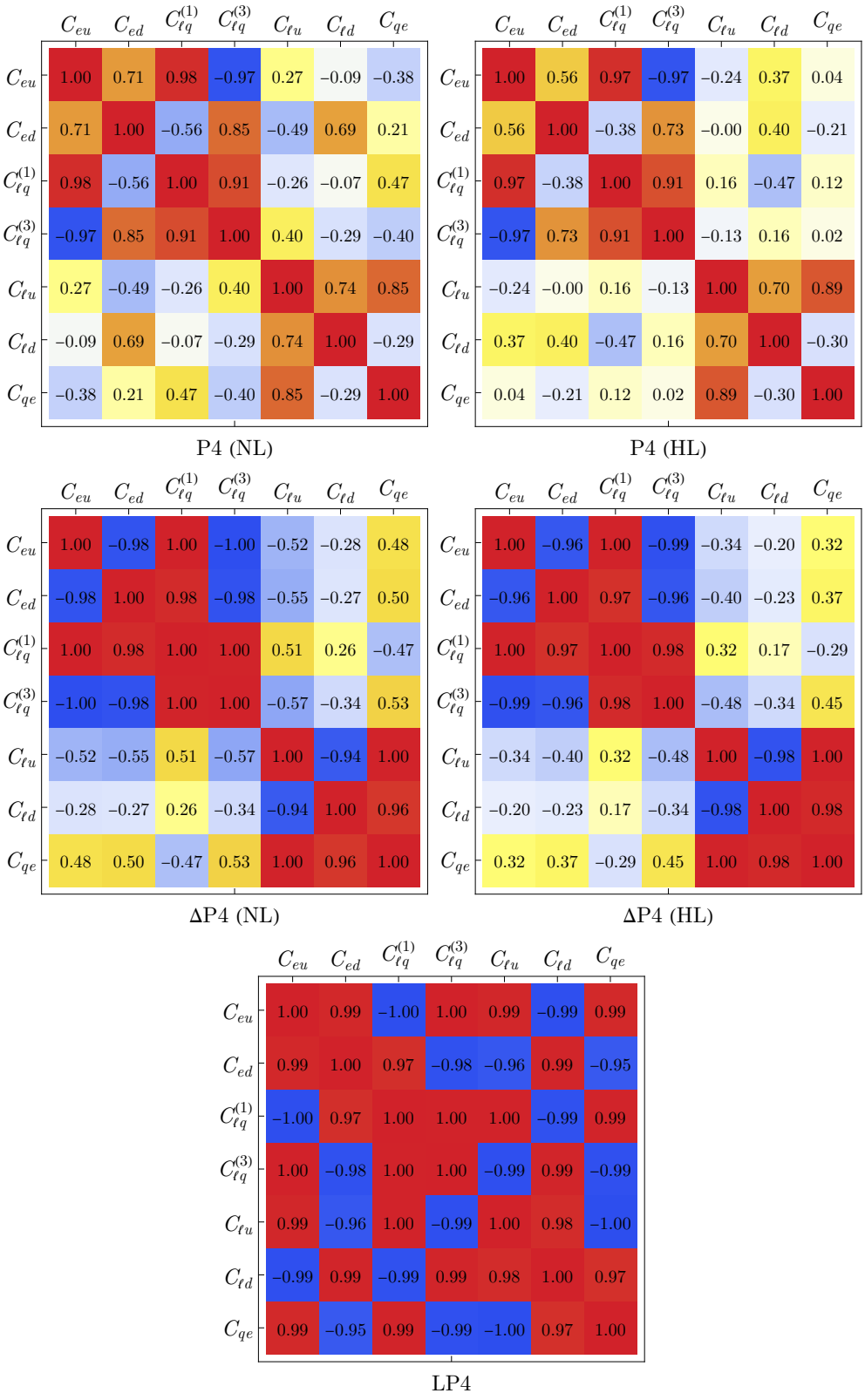

FIG. 59. The same as in Fig. 57 but for P4.

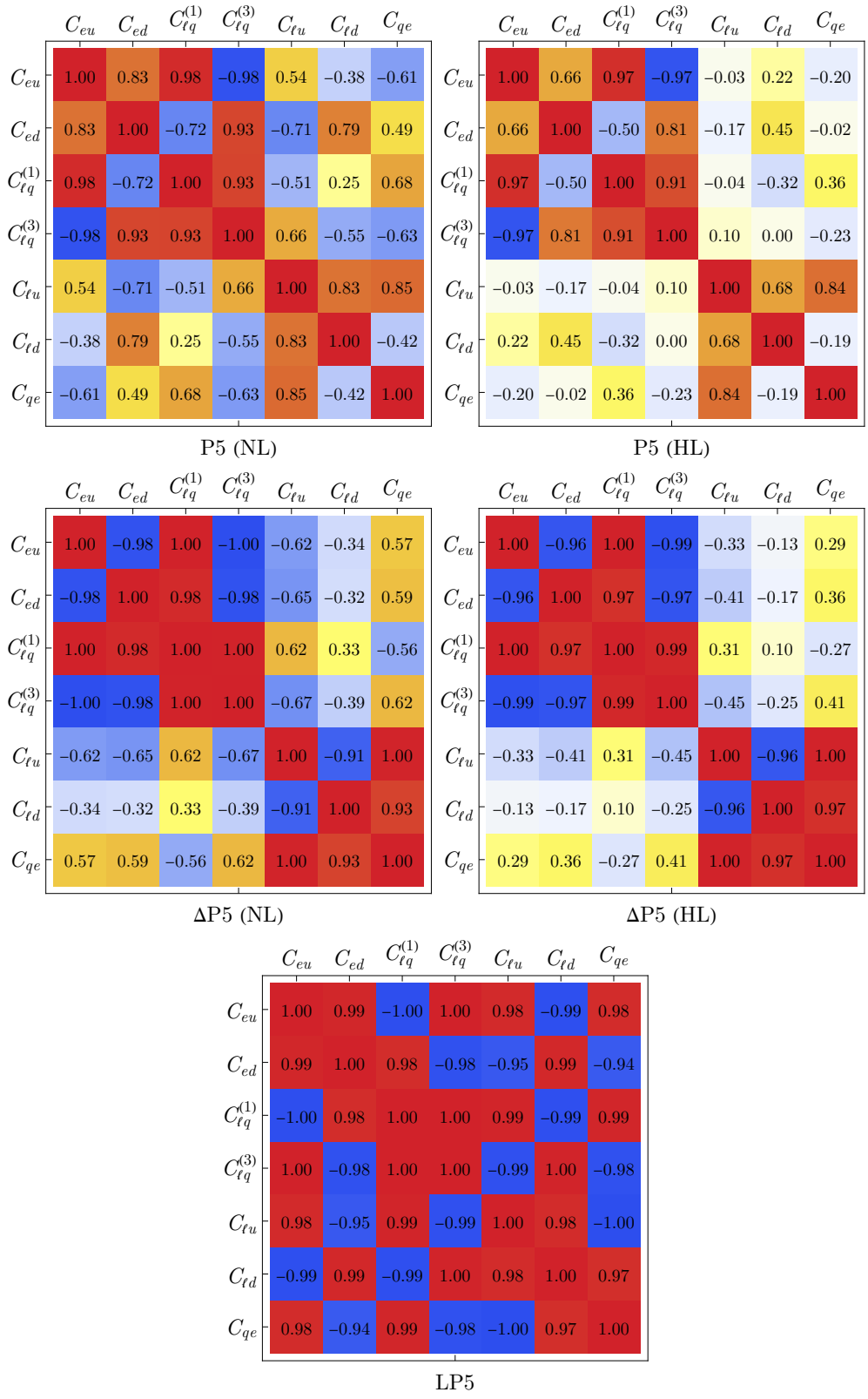

FIG. 60. The same as in Fig. 57 but for P5.

Acknowledgement: This work is supported by the U.S. Department of Energy (DOE) under Award number DE–SC0014434. [add more here]. The authors would like to thank the ECCE Consortium for performing a realistic simulation of their detector design and for providing up-to-date information on EIC run conditions. X. Z. would like to thank H. Spiesberger for suggestions on the use of the Djangoh generator and useful discussions related to the analysis.

---

[1] P. A. Zyla *et al.* (Particle Data Group), Review of Particle Physics, PTEP **2020**, 083C01 (2020).

[2] X. Zheng, J. Erler, Q. Liu, and H. Spiesberger, Accessing weak neutral-current coupling g_{AA}^{eq} using positron and electron beams at Jefferson Lab, Eur. Phys. J. A **57**, 173 (2021), arXiv:2103.12555 [nucl-ex].

[3] K. Charchula, G. A. Schuler, and H. Spiesberger, Combined QED and QCD radiative effects in deep inelastic lepton - proton scattering: The Monte Carlo generator DJANGO6, Comput. Phys. Commun. **81**, 381 (1994).

[4] https://github.com/spiesber/DJANGOH.

[5] S. Schmitt, Data Unfolding Methods in High Energy Physics, EPJ Web Conf. **137**, 11008 (2017), arXiv:1611.01927 [physics.data-an].

[6] R. Abdul Khalek *et al.*, Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report, (2021), arXiv:2103.05419 [physics.ins-det].

[7] R. D. Ball *et al.* (NNPDF), Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017), arXiv:1706.00428 [hep-ph].

[8] E. R. Nocera *et al.* (NNPDF Collaboration), Nucl. Phys. **B887**, 276 (2014).

[9] Electron Ion Collider, Conceptual Design Report, (2021).

[10] T.-J. Hou *et al.*, Progress in the CTEQ-TEA NNLO global QCD analysis, (2019), arXiv:1908.11394 [hep-ph].

[11] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, LHAPDF6: parton density access in the LHC precision era, Eur. Phys. J. C **75**, 132 (2015), arXiv:1412.7420 [hep-ph].

[12] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Parton distributions in

the LHC era: MMHT 2014 PDFs, Eur. Phys. J. C **75**, 204 (2015), arXiv:1412.3989 [hep-ph].

[13] D. Becker *et al.*, The P2 experiment, Eur. Phys. J. A **54**, 208 (2018), arXiv:1802.04759 [nucl-ex].

[14] The MOLLER Project, Measurement Of a Lepton Lepton Electroweak Reaction An Ultra-precise Measurement of the Weak Mixing Angle using Møller Scattering, Conceptual Design Report (2020).

[15] J. P. Chen, H. Gao, T. K. Hemmick, Z. E. Meziani, and P. A. Souder (SoLID), A White Paper on SoLID (Solenoidal Large Intensity Device), (2014), arXiv:1409.7741 [nucl-ex].

[16] P. Souder *et al.* (JLab SoLID-PVDIS, experiment E12-10-007), Precision Measurement of Parity-violation in Deep Inelastic Scattering Over a Broad Kinematic Range (2010).

[17] J. Erler, C. J. Horowitz, S. Mantry, and P. A. Souder, Weak Polarized Electron Scattering, Ann. Rev. Nucl. Part. Sci. **64**, 269 (2014), arXiv:1401.6199 [hep-ph].

[18] Https://www.jlab.org/theory/jam.

[19] Y. X. Zhao, A. Deshpande, J. Huang, K. S. Kumar, and S. Riordan, Neutral-Current Weak Interactions at an EIC, Eur. Phys. J. A **53**, 55 (2017), arXiv:1612.06927 [nucl-ex].

[20] J. Erler and M. J. Ramsey-Musolf, The Weak mixing angle at low energies, Phys. Rev. D **72**, 073003 (2005), arXiv:hep-ph/0409169.

[21] R. P. Ball *et al.* (NNPDF Collaboration), Eur. Phys. J. C **77**, 663 (2017).

[22] J. Arrington *et al.*, Physics with CEBAF at 12 GeV and Future Opportunities, (2021), arXiv:2112.00060 [nucl-ex].

[23] R. Boughezal, F. Petriello, and D. Wiegand, Phys. Rev. D **101**, 116002 (2020).

[24] R. Boughezal, F. Petriello, and D. Wiegand, Phys. Rev. D **104**, 116005 (2021).