Parity Violation in Deep Inelastic Scattering (electron-deuterium) with the SoLID detector at JLab

Felipe Villazon, Cristhian Calderon

Abstract

Solenoidal Large Intensity Detector (SoLID) is a large acceptance high luminosity device in the Jefferson Lab. One of the five experiments in the Jefferson Lab is the measurement of Parity-Violating Asymmetry (APV) in a Parity-Violating Deep Inelastic Scattering (PVDIS) on a deuterium target. The SoLID offers anticipated statistical uncertainties of APV for a given x and Q^2. In this project, we have used this anticipated data to simulate a deviation from the Standard Model APV predictions. The deviation from the Standard Model (SM) could be associated with new physics effects, in this case, Charge Symmetry Violation, and Higher-twist effects. These two contributions can be parameterized by the parameters β_{CSV} and β_{HT}. The idea of this project is to compare the sensitivity of SoLID to the new physics effects. We find, how the SM parameters involved in APV processes, such that, a_1 and a_3 are correlated to β_{CSV} and β_{HT} depending on the uncertainty in the polarization beam.

Contents

1 Introduction

Symmetries play an important role in physics. There are different types of symmetries used in physics. For instance, Lorentz symmetry tells us that the law of physics, i.e. the equation of motion, is independent of the reference frame. Then, we also have the gauge symmetry which describe the interactions between the particles in the Standard Model (SM). There are also some operators associated with discrete symmetries: Charge Conjugation (C), Parity (P), and Time reversal (T) operators. Charge conjugation operator takes particles and transforms them into their antiparticles or vice-versa, we say that if the processes described by the Lagrangian are invariant under this transformation, the theory is C-invariant or that converses C-symmetry. The parity operator, on the other hand, takes particles or antiparticles of a given helicity and returns the same particle or antiparticle having a different helicity, it could be left-handed to right-handed and vice versa. Again, if the Lagrangian remains invariant, we say that the theory is Parity conserving. Time reversal means if we switch the flow of time, the Lagrangian remains invariant. C, P, and T were naturally assumed to be a symmetry of nature until T.D. Lee and C.N. Yang suggested that P-symmetry had to not be conserved in order to explain the identical mass and lifetimes of two different mesons [1]. This hypothesis was tested by C.S Wu and their collaborators, using a nuclear β decay experiment, which confirmed parity violation [2]. After the discovery of parity violation, there were several theories trying to explain this phenomenon. Among this plethora of theories, there was the Glashow-Weinberg-Salam (GWS) model [3–5] which unifies the electromagnetic and weak interaction using the gauge symmetry group $SU(2)_L \times U(1)_Y$ and also predicts the chiral behavior of weak interactions. It means that only left-handed particles and right-handed antiparticles interact via weak forces, producing a difference between the cross-section of left-handed and right-handed particles.

Spinors carry two types of couplings: vector and axial coupling. The vector coupling (g_V) describes the average of the strength of neutral-weak interaction for the left- and right-handed fields, while the axial coupling (g_A) describes the difference between the two. In Deep Inelastic Scattering (DIS), where an electron scatters with a nucleus (proton or deuterium), the virtual boson exchange should be electrically neutral like a photon and a Z-boson. For a virtual γ-boson exchange, there is no difference between left- and right-handed couplings, so there is only vector coupling. For the W-boson, it interacts only with the left-handed particles. For the virtual Z boson exchange, it interacts with both left- and right-handed particles, therefore, for a Z boson exchange, there are both axial and vector coupling [6].

One of the first Parity-Violating Deep-Inelastic Scattering (PVDIS) was performed at the Stanford Linear Accelerator Center (SLAC) on a deuterium target [7]. PVDIS has been proposed as an important new tool for testing the Standard Model and beyond of the Standard Mode. In addition, PVDIS promises to play a relevant role in possible new discoveries regarding the partonic structure of the nucleon and charge symmetry violation.

The goal of the Solenoidal Large Intensity Detector (SoLID) PVDIS program [9] is to measure the cross-section asymmetry, A_{PV}, between right- and left-handed beam electrons with high precision. This asymmetry comes from chiral behavior in weak interactions. At JLab

energies, it can be determined from the interference between γ and Z_0 exchange processes in DIS. SoLID will provide data on A_{PV} with sub-percent relative precision over a wide (x, Q^2) range. Measured on a deuteron target, the A_{PV} data can be used to determine parameters of the electroweak Standard Model and to set limits on new physics, such as Higher-Twisst (HT) and Charge Symmetry Violation (CSV), up to an energy scale that is comparable to the reach of the LHC. The goal of this project is to study the contribution of Higher-twist and Charge Symmetry Violation to the Parity-Violating Asymmetry in DIS with the anticipated statistical uncertainty offered by [8].

This report is structured as follows: in Section 2 we review the Parity Violating Asymmetry in the Standard Model and the two new effects: Higher-twist and Charge Symmetry Violation. In Section 3 we review the concepts and tools we need for the analysis and the fitting of the new physics parameters and we also show and discuss the results. In Section 5, we summarize our discussion and state our conclusions.

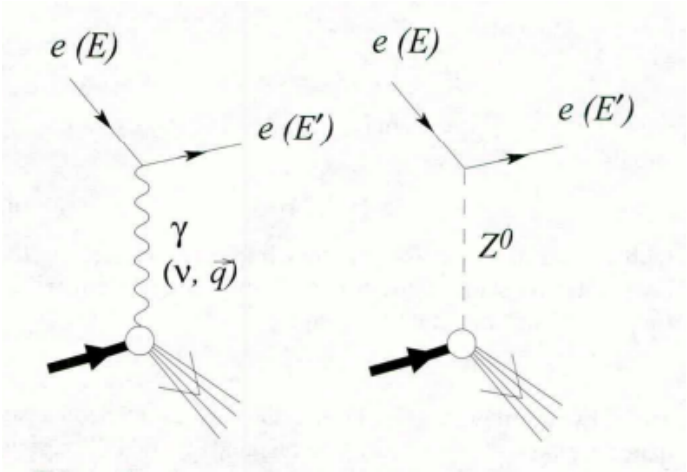

Figure 1: *The electron exchanges either a virtual photon (left) or a virtual Z° (right) with the target. The interference between these two processes leads to a parity-violating asymmetry between left- and right-handed electrons. [10]*

2 Review of the Theory

PVDIS at JLab with the SoLID detector provides a great opportunity to explore Parity Violating Asymmetry. A polarized 11 GeV electron beam scatters with an unpolarized deuterium target in the kinematic region of large Bjorken $x = Q^2/2M\nu$. Parity violation implies that left-handed particles don't interact in the same way that right-handed particles do. Therefore, APV measures the difference between the cross-section of left-handed electrons that scatter the nucleus and the cross-section of right-handed electrons.

$$A_{PV} = \frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}, \tag{1}$$

where $\sigma_R(\sigma_L)$ is the cross section for incident right-(left-) handed electrons. If right-handed electrons scatter like left-handed electrons, the APV would be null.

As we can see in Fig.3, SoLID proposes to obtain data over $x > 0.2$, $2 < Q^2 < 10$ [GeV]. In a deuterium target and over this range, the asymmetry is approximately directly proportional to electron-quark electroweak couplings [8]. However, it is possible that the following physics could be observed in our data:

1. Charge Symmetry violation at the quark level.

2. Higher-twist effects in the Deep Inelastic Scattering.

Before discussing the new physics, we should discuss the Parity Violating Asymmetry in the Standard Model. In the DIS region, the Parity-Violating Asymmetry can be written as [9]

$$A_{PV} = -\frac{G_F Q^2}{4\sqrt{2}\pi\alpha}[a_1 Y_1 + a_3 Y_3], \tag{2}$$

where $G_F = 1.166 \cdot 10^{-5} GeV^{-2}$ is the Fermi constant, α is the fine structure constant, and

$$a_1(x) = 2g_A^e \frac{F_1}{F_1^\gamma}, \tag{3}$$

$$a_3(x) = g_V^e \frac{F_3^{\gamma Z}}{F_1^\gamma}. \tag{4}$$

with,

$$F_1^\gamma(x, Q^2) = \frac{1}{2}\sum e_i^2 [q_i(x, Q^2) + \bar{q}_i(x, Q^2)], \tag{5}$$

$$F_1^{\gamma Z}(x, Q^2) = \sum e_i g_V^i [q_i(x, Q^2) + \bar{q}_i(x, Q^2)], \tag{6}$$

$$F_3^{\gamma Z}(x, Q^2) = 2\sum e_i g_A^i [q_i(x, Q^2) - \bar{q}_i(x, Q^2)]. \tag{7}$$

$F_{1,3}^{\gamma,\gamma Z}$ are structure functions and e_i is the electromagnetic charge of the i^{th} quark. These structure functions can be written in terms of the Parton distribution functions (PDFs), $q_i(x, Q^2)$ and $\bar{q}_i(x, Q^2)$ of the target. The $g_{V,A}^{e,i}$ are the vector and axial couplings of the electron or quark of flavor i in the Standard Model and are related to the weak mixing angle, the electric and weak hypercharge of the particle. The Y_i are functions of the kinematic variable $y = \nu/E$ and the ratios of structure functions $R_j(x, Q_2)$,

$$Y_1(x, y, Q^2) = \frac{1 + (1-y)^2 - y^2(1 + r^2 - \frac{2r^2}{(1+R^{\gamma Z})})}{1 + (1-y)^2 - y^2(1 + r^2 - \frac{r^2}{(1+R^\gamma)})} \left(\frac{1 + R^{\gamma Z}}{1 + R^\gamma}\right), \tag{8}$$

$$Y_3(x, y, Q^2) = \frac{1 - (1-y)^2}{1 + (1-y)^2 - y^2(1 + r^2 - \frac{2r^2}{1+R^\gamma})} \left(\frac{r^2}{1 + R^\gamma}\right). \tag{9}$$

where $r = 1 + \frac{Q^2}{\nu^2}$ and $R^{\gamma(\gamma Z)}(x, Q^2)$ is the ratio of the longitudinal to transverse virtual photon electromagnetic absorption cross sections. The above expression is general; however, in the DIS region, we can approximate the kinetic terms to a more straightforward expression.

$$Y_1 \longrightarrow 1, \tag{10}$$

$$Y_3 \longrightarrow \frac{1 - (1-y)^2}{1 + (1-y)^2}. \tag{11}$$

Therefore, the expression of the Parity Violation Asymmetry in the DIS region simplifies to

$$A_{PV} = -\frac{G_F Q^2}{4\sqrt{2}\pi\alpha}\left[a_1 + a_3\left(\frac{1-(1-y)^2}{1+(1-y)^2}\right)\right], \tag{12}$$

where,

$$y = \frac{Q^2}{x}\frac{1}{s-m_p^2}. \tag{13}$$

2.1 New Physics

The Parity-Violating Asymmetry calculated in the Standard Model does not consider new effects that could have contribution in the measurement of Parity-Violating Asymmetry. We will discuss in detailed two contributions of new physics.

2.1.1 Higher-Twist effect

The higher-twist effects refer to the interaction between quarks inside the nucleon at low Q^2, where QCD perturbation theory breaks down, but not too low where the effective QCD coupling diverges. In order to explain the definition of twist we need to discuss about the factorization and Operator Product Expansion (OPE) in QCD.

For practical purposes, the parton model is enough to perform perturbative QCD calculations for high-energy scattering involving hadrons. This phenomenological approach is based on *factorization*: that Parton Distrbution Functions (PDFs) are universal objects, and any scattering process involving protons can be calculated using the same PDFs with a different perturbative calculation. We will discuss classical factorization for Deep Inelastic Scattering using the OPE. When this expansion is performed, several operators appear with a new quantity called "*twist*". We will see more in detail how this expansion is performed.

To apply the OPE to DIS, we first need to express the hadronic tensor $W^{\mu\nu}$ in terms of matrix elements of the electromagnetic current constructed from quarks.

$$W_{\mu\nu} = \sum_X \int d\Pi_X \int d^4x e^{i(q+P-p_X)x}\langle P|J_\mu(0)|X\rangle\langle X|J_\nu(0)|P\rangle. \tag{14}$$

The current-current matrix element in a quark state is the same as the forward scattering matrix for Compton Scattering. At leading order in perturbation theory, the result is then

$$i\int d^4x e^{iqx}\langle p|T\{J_\mu(0)J_\nu(0)\}|P\rangle = -\bar{u}(p)\frac{\gamma^\mu(\not{p}+\not{q})\gamma^\nu}{(p+q)^2+i\epsilon}u(p) - \bar{u}(p)\frac{\gamma^\nu(\not{p}-\not{q})\gamma^\nu}{(p-q)^2+i\epsilon}u(p). \tag{15}$$

We perform an expansion in the denominator of the propagator and we will be able to read off the Wilson coefficients and operators in the OPE.

$$\frac{1}{(p+q)^2} = \frac{1}{-Q^2+2q\cdot p+p^2} = -\frac{1}{Q^2}\sum_{n=0}^{\infty}\left(\frac{2p\cdot q+p^2}{Q^2}\right). \tag{16}$$

For example, a term $\left(\frac{p^2}{Q^2}\right)^n$ should come from an operator $\mathcal{O}_n = \bar{\Psi}\Box^n\Psi$. Plugging this expansion 16 in the Eq. 15, we get the matrix elements of the hadronic tensor as a sum of operators

that carry a twist number. The twist is defined as the difference between the dimension of the operator and the spin, where spin is the rank of the operator. In order to illustrate the quantity t, we have the following operator that comes from the OPE.

$$\mathcal{O}^{\mu\nu}_{2,0} = \bar{\Psi}_q \left(i\gamma^\mu \partial^\nu + i\gamma^\nu \partial^\mu - \frac{1}{2} i g^{\mu\nu} \not{\partial} \right) \Psi_q. \tag{17}$$

The dimension of this operator is four and it has rank two, therefore the twist of this operator is two. The expansion is suppressed by $1/Q^2$ so for higher energies the leading operators contribution is bigger than the higher-twist operators. The leading order operator has twist equal two and it turns out the following operator with higher-twist has twist equal four.

$$\mathcal{O}^{\mu\nu}_{ud} = \frac{1}{2} \left[\bar{u}(x)\gamma^\mu u(x) d(0)\gamma^\nu d(0) + (u \Longleftrightarrow d \right]. \tag{18}$$

We will study how this higher-twist operator that comes from the OPE of the hadronic tensor contributes to the Parity-Violating Asymmetry.

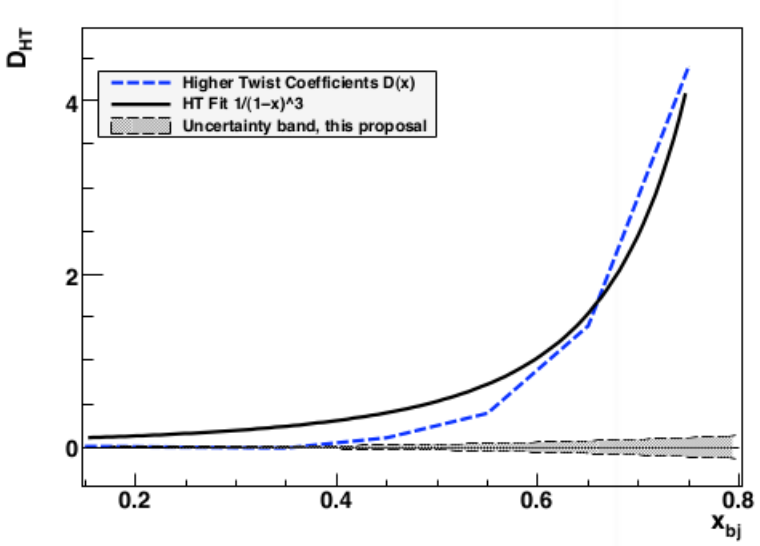

Figure 2: *Possible Contribution to A_{PV} due to higher twist contributions. Plot taken from the SoLID report [8].*

The Higher-Twist effects modify the PVDIS asymmetry by affecting the absorption cross-section ratio R^γ in Eq. 8. It could also affect the structure functions ratios a_1 and a_3 in Eq. 3. According to the study in [10], the effect on R^γ and a_1 was found to be very small and there is no theoretical estimation of HT effects on the a_3 structure constant ratio. A representative plot on the higher twist effect in the unpolarized electromagnetic cross sections is shown in Fig. 2. We see in Fig. 2 that HT coefficient has a polynomial dependency on x and it also goes as $1/Q^2$ according to [11]

$$A_{PV} \longrightarrow A_{PV} \left(1 + \beta_{HT} \frac{1}{(1-x)^3 Q^2} + ... \right), \tag{19}$$

where the correction is due to Higher-Twist effects and the ellipsis refers to other effects such as Charge Symmetry Violation.

2.1.2 Charge Symmetry Violation

Charge symmetry (CS) is a form of Isospin symmetry (IS). At the level of quarks and gluons of QCD, Charge Symmetry interchanges up and down quarks $P_{cs}|d\rangle = |u\rangle$, $P_{cs}|u\rangle = -|d\rangle$. In QCD, the only sources of Charge Symmetry Violation (CSV) are: the mass difference between u and d quarks and the electromagnetic interaction.

Naively, one would expect that CSV would be of the same order of $(m_d - m_u)/\langle M\rangle$, where $\langle M\rangle = [0.5 - 1]GeV$, giving a CSV effect of 1%. CSV is related to our understanding of the flavor dependence of the quark masses (one of the key unsolved problems in Physics – why is $m_d \sim m_u \neq m_s \neq m_c \neq m_b \neq m_t$). *One very important point is that Charge Symmetry has been universally assumed in Parton distribution functions.* For example,

$$\begin{aligned} u^p(x,Q^2) =& d^n(x,Q^2), \\ d^p(x,Q^2) =& u^n(x,Q^2). \\ s^p(x,Q^2) =& s^n(x,Q^2). \end{aligned}$$

The charge symmetry violation affects the Parity-violating Asymmetry and it also depends in the variables x and Q^2. The parametrization of how charge symmetry violation deviates the Parity Violating asymmetry is offered by Solid [8].

$$A_{PV} \longrightarrow A_{PV}(1 + \beta_{CSV}x^2 + ...), \tag{20}$$

where the ellipsis represent other effects of new physics that could be consider and modifies the Parity-Violating Asymmetry.

3 Data Sample and Analysis

When an experiment or a simulation is performed we have to deal with and analyse the data we obtain. This is crucial, how we process and understand it might modify the way we interpret it or the conclusions we come up with.
We have made the analysis in two ways. The first one consists in that we have written a code in python from the scratch that can perform a minimization of the χ^2 and compute the covariance matrix in order to find the errors and correlations. The second method was used to verify our first method, there, we used a python module called *iminuit* that has already implemented the minimization of the χ^2.

The anticipated statistical uncertainty for A_{PV} in percent is given in Fig. 3 and it is summarized in Table 1. This anticipated statistical uncertainty will be useful to simulate new physics in the A_{PV} values. Before simulating the data that contains new physics, we should calculated the APV in the Standard Model. We know that in the Standard Model the amount of parity violation asymmetry is given by Eq. 12. Thus, using the values for x and Q^2 from Table 1 we can evaluate Equation 12 and then obtain the values of APV in the Standard Model, A_{PV}^{EW} as you can see in Table 1. Once we have the Standard Model precision of our experiment, we need to find a way to generate new possible data that could contain new physics. For this, we have simulated it for different values of x and Q^2. The procedure is explained below.

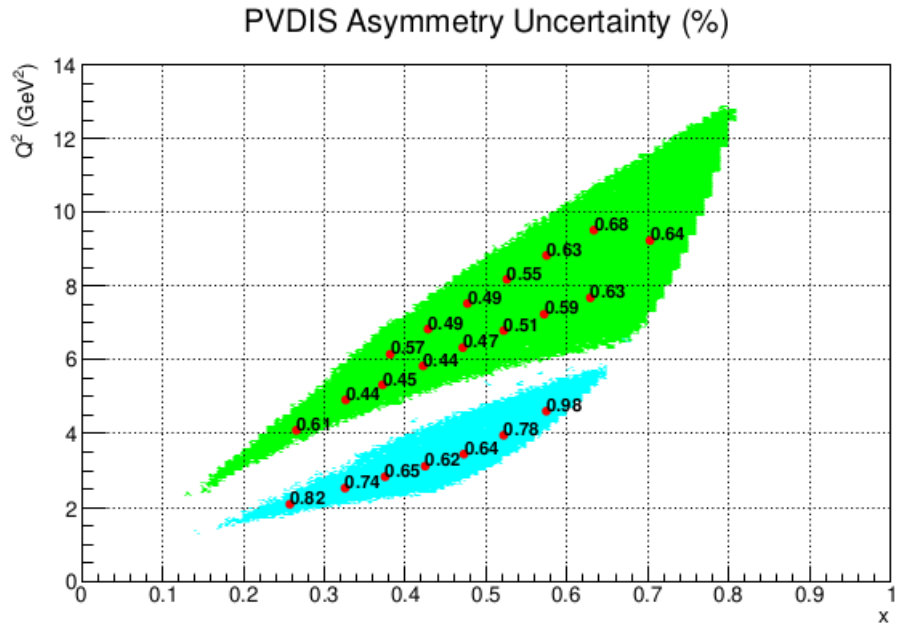

Figure 3: *Anticipated statistical uncertainty for A_{PV} in percent versus Q^2 and x, 11 GeV data (green region) are based on a simulation including the electron calorimeter trigger. For the 6.6 GeV data, the effect of the trigger is estimated. Plot taken from [8]*

We now generate data containing possibly new physics by simulating them using random Gaussian distributed numbers, which means they will follow the distribution,

$$A_{PV} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\left(\frac{x-\mu}{\sigma}\right)^2\right], \tag{21}$$

where $\mu = A_{PV}^{EW}$ and $\sigma = \frac{\delta A_{PV}}{A_{PV}^{EW}} \cdot 100$, both set of values given by Table 1. Eq. 21 tells that the new Parity-Violating Asymmetry is generated by the Gaussian random distribution with the mean value in the Standard Model APV and a standard deviation offered by Solid [8]. These deviations of the APV might contain new physics.

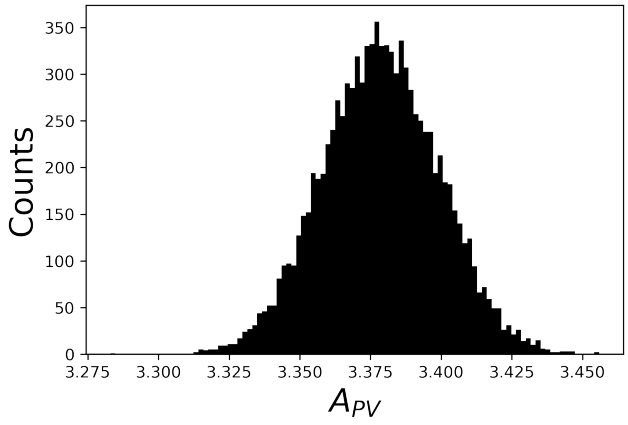

Figure 4: *Gaussian distribution of the simulated data for A_{PV}.*

To illustrate the simulation, we do it with one δAPV data point at fixed x and Q^2. For example Fig. 4 shows us the case how data is distributed using $A_{PV}^{EW} = 3.3788 \cdot 10^{-4}$ and $100 \cdot \frac{\delta A_{PV}}{A_{PV}} = 0.61\%$. Table 1 shows an example of the simulated A_{PV} data. It is important to stress that from Fig. 4 we are just getting one new APV data point for a given x and Q^2. It means for one A_{PV}^{EW} value at fixed x and Q^2 we have only one simulated A_{PV} at the same fixed x and Q^2.

x	$Q^2[GeV^2]$	$\delta A_{PV}/A_{PV}[\%]$	$A_{PV}^{EW}[10^{-4}]$	$A_{PV}[10^{-4}]$
0.2654	4.0887	0.61	3.3788	3.3872
0.3262	4.9033	0.44	4.0721	4.0794
0.3728	5.3263	0.45	4.4654	4.4735
0.3813	6.1566	0.57	5.0373	5.0490
0.4228	5.8276	0.44	4.9150	4.9237
0.4286	6.8459	0.49	5.6145	5.6256
0.4708	6.3289	0.47	5.3580	5.3683
0.4772	7.5352	0.49	6.1950	6.2073
0.5223	6.7989	0.51	5.7813	5.7932
0.5259	8.1932	0.55	6.7551	6.7702
0.5717	7.2375	0.59	6.1749	6.1896
0.5753	8.8355	0.63	7.3062	7.3249
0.6289	7.6762	0.63	6.5737	6.5905
0.6332	9.5091	0.68	7.8986	7.9204
0.7026	9.2428	0.64	7.8486	7.8689

Table 1: *Relative δA_{PV} is the anticipated statistical uncertainty of A_{PV} in percent. [8]*

The task now is to see how our physical observables are affected when new physics is considered, that is, how they become more or less correlated if new physics is considered. In our case, we have taken into account that our new physics should follow the following parametrization [8],

$$A_{PV}(\beta_{HT}, \beta_{CSV}, a_1, a_3) = A_{PV}^{EW} \cdot \left(1 + \beta_{HT} \frac{1}{(1-x)^3 Q^2} + \beta_{CSV} x^2\right) \tag{22}$$

One should expect that if there is no new physics, the true values of our parameters to be fitted should be the ones given by the Standard Model, that is,

$$a_1 = 0.864 \qquad a_3 = 0.096 \qquad \beta_{HT} = 0 \qquad \beta_{CSV} = 0, \tag{23}$$

what are exactly the mean value of the parameters we want to fit when we perform the minimization of the χ^2.

3.1 Systematic and Statistical errors

When an experiment is performed, the way we have to know how reliable our results are is by assigning or calculating a given error to the measurements. These uncertainties can be divided into two kinds of errors, **Systematic errors** and **Statistical errors**.

The statistical error is calculated from the size of the sample and various measures of how it might (statistically) vary. There's no single formula for that, but there are a few well-recognized methods: Poisson statistics for small numbers of events, normal (Gaussian) statistics for large numbers of samples from a single distribution, etc. The statistical error basically tracks the possibility of random fluctuations during the measurement: the more repetitions of the measurement, the smaller the statistical uncertainties are.

Systematic errors are not due to random processes in the measurement itself and they can't be reduced when the experiment is repeated several times. Systematic errors capture the errors that can appear in realistic experiments. Systematic uncertainties arise typically when instruments are poorly calibrated, or are influenced, in a way not known exactly, by the environment (temperature, pressure. . .).

For us, it is important to identify the possible systematic errors that could affect the data, *for this, one of the main goals of this work is to see how systematic errors could change, increase or decrease the correlation between our observables.*

By assuming that our data have Gaussian errors, the simplest covariance matrix V is when systematic errors are not considered and it should be diagonal

$$V = \begin{pmatrix} \sigma_1^2 & 0 & 0 & . & . \\ 0 & \sigma_2^2 & 0 & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix} \tag{24}$$

As defined in Equation 24, the covariance matrix is just a diagonal matrix with the statistical Gaussian errors of the measurements in the ii components. However, experiments cannot be perfect, there will be always some factors like temperature, calibration, etc, problems that could affect the measured data. It is important to take them into account since we could get more realistic conclusions. In the SoLID experiment, it could be that we are not having a perfectly polarized beam of electrons, which would produce the wrong conclusion about the asymmetry parameter A_{PV}. We will start our data analysis using that simpler case, with no systematic uncertainties and then we will consider systematic errors related to the polarimeter.

In order to have systematic errors in the second case, we are going to consider that we have a systematic error in the polarimeter that would let us explore the correlation between the experiments. We have a set of experiments $\{x_1, x_2, ..., x_n\}$, and these experiments are correlated through the polarimeter systematic error. We would like to know how the correlations between the measurements depend on the polarimeter parameter and its uncertainty. We introduce the parameter η which describes how many of the electrons are really polarized.

$$A_{PV}^{\text{real}} = \eta \cdot A_{PV}, \tag{25}$$

if $\eta = 1$ means that we have a perfect polarized beam. We will assume that the polarization parameter follows a Gaussian Distribution. Let x be an experiment from our set $\{x_1, x_2, ..., x_n\}$.

The mean values are now given by,

$$\langle x\rangle = \frac{1}{2\pi\sigma_x\sigma_\eta}\int x\cdot\exp\left[\frac{-(x-\eta x_o)^2}{\sigma_x^2}\right]dx\exp\left[\frac{-(\eta-\eta_o)^2}{\sigma_\eta^2}\right]d\eta \tag{26}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_\eta}\int \eta\cdot x_o\exp\left[\frac{-(\eta-\eta_o)^2}{\sigma_\eta^2}\right]d\eta = x_o\eta_o, \tag{27}$$

where x_o is the mean value in the case of perfect polarization. We can realize that if the mean value of the measurements do not change, the mean value of the polarimeter has to be 1. In our case, it means that the A_{PV} does not change when the mean value of the polarimeter is 1. The main reason to introduce the polarimeter in our analysis is to see how the experiments are correlated. Let x and y be two different experiments from our set $\{x_1, x_2, ..., x_n\}$, hence the correlation between these two experiments is given by,

$$\text{corr(x,y)} = \frac{\text{cov(x,y)}}{\sigma_x\cdot\sigma_y} = \frac{1}{\sigma_x\cdot\sigma_y}\int(x-\eta_o x_o)(y-\eta_o y_o)P(x,y)dxdy. \tag{28}$$

The polarization parameter is implicitly in the distribution function and it is given by

$$P(x,y) = \frac{1}{((2\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_\eta)}\int\exp\left[\frac{-(x-\eta x_o)^2}{\sigma_x^2}\right]\exp\left[\frac{-(y-\eta y_o)^2}{\sigma_y^2}\right]\exp\left[\frac{-(\eta-\eta_o)^2}{\sigma_x^2}\right] \tag{29}$$

Let us first calculate the covariance between the experiments x and y.

$$\text{cov(x,y)} = \int(xy-\eta_o y_o x-\eta_o x_o y+\eta_o^2 x_o y_o)P(x,y)dxdy \tag{30}$$

$$= \langle xy\rangle - \eta_o y_o\langle x\rangle - \eta_o x_o\langle y\rangle + \eta_o^2 x_o y_o = \langle xy\rangle - \eta_o^2 y_o x_o - \eta_o^2 x_o y_o + \eta_o^2 x_o y_o \tag{31}$$

$$= \langle xy\rangle - \eta_o^2 x_o y_o. \tag{32}$$

We need to compute the mean value of the product between the experiment x and y.

$$\langle xy\rangle = \int x\cdot y\cdot P(x,y)dxdy \tag{33}$$

$$= \frac{1}{((2\pi)^{\frac{3}{2}}\sigma_x\sigma_y\sigma_\eta)}\int x\cdot\exp\left[\frac{-(x-\eta x_o)^2}{\sigma_x^2}\right]y\cdot\exp\left[\frac{-(y-\eta y_o)^2}{\sigma_y^2}\right]\exp\left[\frac{-(\eta-\eta_o)^2}{\sigma_x^2}\right] \tag{34}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_\eta}\int\eta x_o\cdot\eta y_o\exp\left[\frac{-(\eta-\eta_o)^2}{\sigma_x^2}\right] = x_o\cdot y_o\cdot\left\langle\eta^2\right\rangle = x_o\cdot y_o(\eta_o^2+\sigma_\eta^2). \tag{35}$$

Plugging this last result in the covariance expression between x and y, we find.

$$\text{cov(x,y)} = \langle xy\rangle - \eta_o^2 x_o y_o = \sigma_\eta^2 x_o y_o. \tag{36}$$

Finally, in order to get the correlation, we have to normalize by the uncertainties of x and y. We can see that the correlation between two experiments x and y, for instance two points from the Fig. 3, are in function of the uncertainty of the polarimeter and it is independent of the mean value of the polarimeter.

$$\text{corr(x,y)} = \frac{\sigma_\eta^2\langle x\rangle\langle y\rangle}{\sigma_x\cdot\sigma_y} \tag{37}$$

3.2 The Least squares method χ^2

We already mentioned that we need to fit our data to the function 22 which includes new physics effects. For this, we are going to use the χ^2-method.

The least squares method is a form of mathematical regression analysis used to determine the line of best fit for a set of data, providing a visual demonstration of the relationship between the data points. Each point of data represents the relationship between a known independent variable and an unknown dependent variable. Indeed, all is resume in minimizing the χ^2 function, defined as,

$$\chi^2 = \sum_{i,j} (y_i - f_i) V_{ij}^{-1} (y_j - f_j) \tag{38}$$

where f_i represents the fitting function, y_i is the data and V_{ij} is the covariance matrix. The parameters we would like to fit are inside f_i, as seen in Equation 22. In the case of no correlation and no systematic errors, the χ^2 function is reduced to,

$$\chi^2 = \sum_i \left(\frac{y_i - f_i}{\sigma_i} \right)^2 \tag{39}$$

where σ_i represents the statistical errors in the data.
Then, the χ^2 function is,

$$\chi^2(a_1, a_2, \beta_{HT}, \beta_{CSV}) = \sum_i \left(\frac{[A_{PV} - A_{PV}(x, Q^2, a_1, a_2, \beta_{HT}, \beta_{CSV})]_i}{\sigma_i} \right)^2 \tag{40}$$

where $\sigma_i = \left(\frac{\delta A_{PV}}{A_{PV}^{EW}} \cdot 100 \right)_i$. We can see that from Table 1, we have all the data necessary to evaluate in the χ^2, leaving it as a function of only four parameters. These parameters are physical observables and we will be fitting and finding correlations between them.

A more realistic analysis is done by adding systematic errors and correlations. A source of systematic errors in PVDIS is the polarization of the electrons. In a realistic experiment, the polarimeter does not polarize 100% of the electrons. This systematic error must be considered in our χ^2 analysis. The polarization factor was introduced in Eq. 25 and the χ^2 is modified to

$$\chi^2(a_1, a_2, \beta_{HT}, \beta_{CSV}, \eta) = \sum_i \left(\frac{[A_{PV} - \eta \cdot A_{PV}(x, Q^2, a_1, a_2, \beta_{HT}, \beta_{CSV})]_i}{\sigma_i} \right)^2 + \frac{(\eta - 1)^2}{\sigma_\eta^2}, \tag{41}$$

where we have assumed that the error of the polarimeter follows a Gaussian distribution and the mean value of the polarimeter is 1. It is relevant to stress that σ_i contains statistical uncertainties only since the systematic errors and correlations are implicitly contained in the polarimeter parameter. Now we have to fit five instead of four parameters since the polarization parameter becomes a variable in the χ^2 function. We are not interested in fitting the polarization parameter; however this unwanted parameter influences the estimate of the former parameters, this is the definition of *nuisance-parameter*.

3.3 Error Estimation

Once we have obtained the fitted values of the parameters, the next task is to find the uncertainties, covariance, and correlations between these parameters. To remember, covariance measures the direction of a relationship between two variables, while correlation measures the strength of that relationship. Both correlation and covariance are positive when the variables move in the same direction, and negative when they move in opposite directions. However, a correlation coefficient must always be between -1 and +1, with extreme values indicating a strong relationship.

To do so, we will apply two different methods in the case of no systematic errors case: the $\chi^2_{min}+1$ method and the Monte Carlo method [1] . While for the case of the systematic error we will use the most used one, the $\chi^2_{min}+1$ method. These two methods are discussed in detail below.

3.3.1 $\chi^2_{min}+1$ Method

This will be used for both cases, with and without systematic errors since it is the most used and reliable one. Here, we basically will take the Hessian of the χ^2 function and evaluate it in the fit-best values. For example in the case of $\chi^2(a,b)$ function depending on two variables we would have,

$$Hessian\left[\chi^2(a,b)\right] = \begin{pmatrix} \frac{\partial^2\chi^2}{\partial a\partial a} & \frac{\partial^2\chi^2}{\partial a\partial b} \\ \frac{\partial^2\chi^2}{\partial b\partial b} & \frac{\partial^2\chi^2}{\partial b\partial a} \end{pmatrix}. \tag{42}$$

and finally, the covariance matrix, which contains all the information we are searching for is,

$$\left(V^{-1}\right) = \frac{1}{2}\begin{pmatrix} \frac{\partial^2\chi^2}{\partial a\partial a} & \frac{\partial^2\chi^2}{\partial a\partial b} \\ \frac{\partial^2\chi^2}{\partial b\partial b} & \frac{\partial^2\chi^2}{\partial b\partial a} \end{pmatrix}_{a=\hat{a},b=\hat{b}} \tag{43}$$

where $\hat{a}$ and $\hat{b}$ stand for the estimated or the best-fit-values. Perform second order differentiation sometimes is hard and a second way to calculate the errors could help to verify that our first method is correct and reliable.

3.3.2 Monte Carlo Error Estimation

Unlike the method above, Monte Carlo has the advantage that does not have to deal with differentiation in the error estimation but has to repeat the χ^2 minimization several times. This means that in order to obtain better results, we need to repeat the process as much as we can. We accumulate statistics and use unbiased estimators to calculate the errors. We use the following estimators,

[1] We repeat the fitting several times and with unbiased estimator, the error is calculated.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{for mean value of x ,} \tag{44}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2 \quad \text{for statistical error of x} \tag{45}$$

$$V_{xy} = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) \quad \text{for covariance between x and y variables} \tag{46}$$

For example, let's say we did the minimization and we found the best-fit-values to be $\hat{a}_1$ and $\hat{b}_1$, but as we can see since the error go as $1/\sqrt{n}$ we need more statistics, therefore we repeat the minimization again and again until we obtain a data set like,

$$x = (\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{a}_5, ...) \tag{47}$$

$$y = (\hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4, \hat{b}_5, ...) \tag{48}$$

It is important to stress that every component of x or y variables are best-fit-values obtained from the χ^2 method.

As we already mentioned, both methods will be used in the case of no systematic errors in order to check the code for the χ^2. Then, we will focus on the more realistic case considering systematic errors and correlations between the measurements using only the method explained in Section 3.3.1.

4 Application and Results

4.1 Case without systematic error

As a first step, we start the analysis not considering systematic errors, therefore we will be using the χ^2 defined in Eq. 40 to do the analysis. For the case of using the method $\chi^2_m in + 1$ explained in Section 3.3.1, the correlation between our physical observables is given by Figure 5 and Table 2.

We check our results using the method Monte Carlo from Section 3.3.2. From Table 3a to Table 3d we can see how the correlation between the parameters is affected by having more or fewer statistics. As expected, we can see by looking at Table 3d, which is the Monte Carlo estimation with 10000 iterations, and at Table 2, we realize that both methods tend to give the same results. In addition, what is common even in the case with few MC iterations case is that β_{HT} and β_{CSV} tend to be strongly negatively correlated as well as a_1 and a_3. We can also see that a_3 and β_{CSV} are weakly correlated.

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.88	0.43	-0.28
β_{CSV}	-0.88	1	-0.37	0.04
a_1	0.43	-0.37	1	-0.88
a_3	-0.28	0.04	-0.88	1

Table 2: *Correlation matrix between the fitted parameters: β_{HT}, β_{CSV}, a_1 and a_3. Values close to +1 means that they are positive strongly correlated, values close to -1 means they are negative strongly correlated and values close to 0 means that they are weakly correlated*

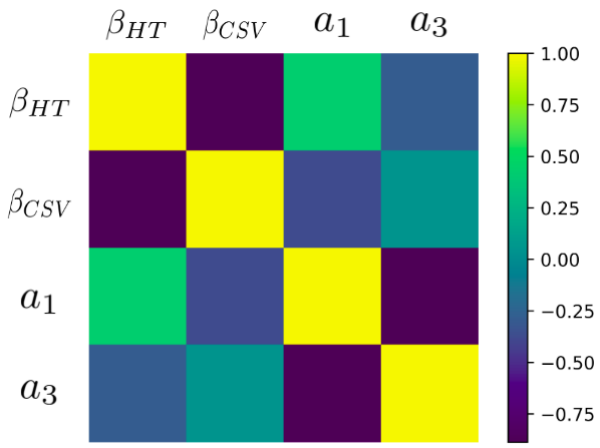

Figure 5: *Correlation matrix between the fitted parameters represented in colors.*

The case we just showed is not quite realistic since we are not considering possible systematic contributions to the errors in the measurements, therefore in Section 4.2, we will analyze what happens when the polarization of the beam of electrons is not perfectly known.

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.90	0.24	-0.17
β_{CSV}	-0.90	1	-0.42	0.19
a_1	0.24	-0.42	1	-0.89
a_3	-0.17	0.19	-0.89	1

(a) *MC with 10 iterations.*

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.87	0.33	-0.19
β_{CSV}	-0.87	1	-0.25	-0.05
a_1	0.33	-0.25	1	-0.89
a_3	-0.19	-0.05	-0.89	1

(b) *MC with 100 iterations.*

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.88	0.42	-0.29
β_{CSV}	-0.88	1	-0.36	0.05
a_1	0.42	-0.36	1	-0.88
a_3	-0.29	0.05	-0.88	1

(c) *MC with 1000 iterations.*

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.88	0.44	-0.29
β_{CSV}	-0.88	1	-0.38	0.05
a_1	0.44	-0.38	1	-0.88
a_3	-0.29	0.05	-0.88	1

(d) *MC with 10000 iterations.*

Table 3: *Correlation between the fitted parameters using Monte Carlo error estimation 3.3.2.*

4.2 Case with systematic error

From [8], we find that the uncertainty in the polarization of the beam reaches 0.4%. We have seen that we can get access to the correlations through the error of the polarization beam. In Table 4 we are comparing how the correlation between the fitted parameters has changed by adding the systematic error on the polarization beam. At first sight, we can note that by looking at Table 4a and Table 4b, the relation between the parameters of the possible new physical effects β_{HT} and β_{CSV} is not affected by the systematic beam uncertainty. On the other hand, what is appreciable are the changes in the correlation between all the other parameters.

We just have seen what happens with the correlations when we use the systematic errors given by the experiment. Sometimes, it is possible that the systematic errors are not detected or determined correctly, as could be the case for the polarization beam. Therefore, we also consider

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.88	0.34	-0.28
β_{CSV}	-0.88	1	-0.29	0.04
a_1	0.34	-0.29	1	-0.67
a_3	-0.28	0.04	-0.67	1

(a) *Beam with 0.4% error in the polarization.*

	β_{HT}	β_{CSV}	a_1	a_3
β_{HT}	1	-0.88	0.43	-0.28
β_{CSV}	-0.88	1	-0.37	0.04
a_1	0.43	-0.37	1	-0.88
a_3	-0.28	0.04	-0.88	1

(b) *Beam with perfect symmetric polarization.*

Table 4: *Covariance matrix comparison between the ideal case with symmetric beam polarization (left) and the realistic case with asymmetric polarization (right).*

the possibility of having larger and smaller uncertainty in the polarization beam. It could be the case that it was overestimated or underestimated. Fig. 6 shows how the correlation in the parameters tends to be modified when different errors for the polarization beam are considered.

The error of the polarization beam is not only related to the correlation of the parameters as we can see in Fig. 6, but it is also related to the correlation between the experiments. In section 3.1 we have explored how the experiments are correlated through the error of the polarization beams. This dependency is given by Eq. 36 and using the statistical uncertainties given by Table 1 we find,

$$\text{corr (x,y)} = \frac{\sigma_\eta^2 \langle x\rangle \langle y\rangle}{\sigma_x \sigma_y} = \frac{\sigma_\eta^2 \langle x\rangle \langle y\rangle}{\delta A_{PV}(x) \langle x\rangle \cdot \delta A_{PV}(y) \langle y\rangle} = \frac{\sigma_\eta^2}{\delta A_{PV}(x) \delta A_{PV}(y)}, \tag{49}$$

where $\delta A_{PV}(x)$ is the relative anticipated uncertainty of an experiment x given in the Table 1. We can get access to the correlation between the experiments and correlations between the fitted parameters through the error of the polarimeter as we can see in the following illustration.

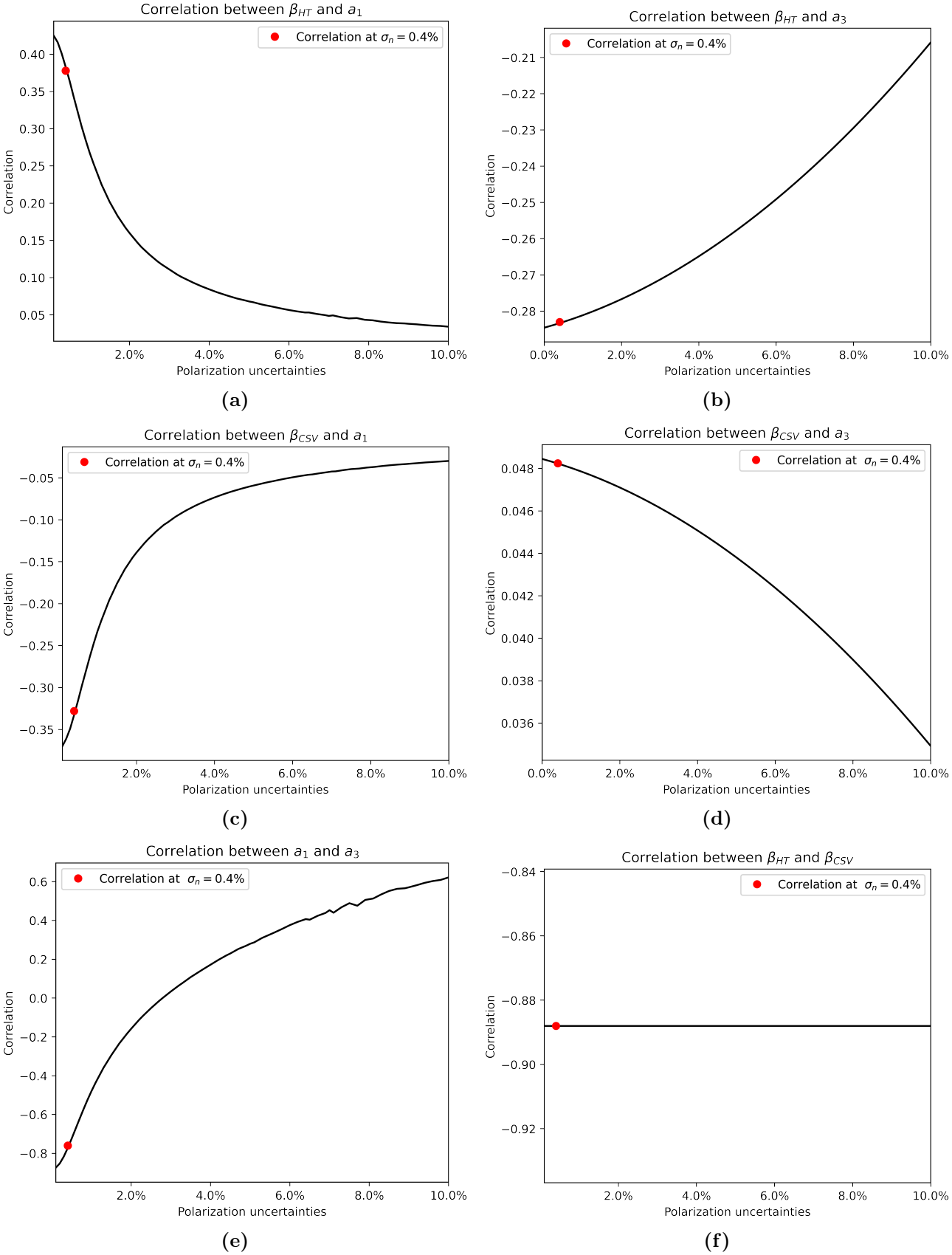

Figure 6: *We simulate an error in the polarization beam between 0.1% to 10%. The red dot represents the correlation at 0.4% as given by [8].*

5 Conclusions

In this project, we have studied the Parity-Violating Asymmetry in Deep Inelastic Scattering between electrons and a nucleus of deuterium at the JLab using the SoLID device. We have explored and considered two new physics effects: Higher-twist and Charge Symmetry Violation. These two effects could affect the Parity-Violating Asymmetry through either the structure functions ratio or the kinetic terms. We have simulated A_{PV} data from the Standard Model values and the anticipated statistical uncertainty offered by the Preliminary SoLID report (2019). With the simulation in our hands, we have fitted the contribution of the Higher-Twist and Charge Symmetry Violation effects as well as the structure constants ratios a_1 and a_3. We have found that the fitted values of a_1 and a_3 are consistent with the values of the Standard Model and that the parameters of the new effects are around zero which is a consistent result since these two new effects should be some kind of corrections of the Standard Model.

Our main goal of this project was to study the correlations of the fitted parameters: $\beta_{HT}, \beta_{CSV}, a_1$, and a_3 in different scenarios. These scenarios differ by the value of the polarization uncertainties. We can see in Fig. 6a that the correlation between β_{HT} and a_1 is indirectly proportional to the polarization uncertainties. It means that well-calibrated polarimetry will give around a 50% correlation and this correlation drops down while the polarization uncertainties increases. A similar behavior we can see in Fig. 6d in the correlation between β_{CSV} and a_1. There is a direct positive correlation between β_{HT} and a_3 in Fig. 6b. It means that well-calibrated polarimetry will give a negative correlation around -0.28 and it increases when the polarization uncertainties increase. A similar behavior with the next couples parameters: (β_{CSV}, a_1) and (a_1, a_3). A very interesting result in Fig. 6f we found in the couple $(\beta_{HT}, \beta_{CSV})$, we can see that the correlation is constant when we vary the polarization uncertainties. It means that the correlation of these two new effects is independent of polarization.

We have also studied how the error of the polarization beam induces the systematic correlations between the measurements $\{x_1, x_2, ..., x_n\}$ and is given by Eq. 36. The correlation is independent of the mean value of the polarization beam and it is direct proportional to the square of the uncertainty of the polarimeter, σ_η.

References

[1] T. D. Lee and C. N. Yang, Phys. Rev. 104, 254 (Oct 1956)

[2] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, Phys. Rev. 105, 1413 (Feb 1957)

[3] S. L. Glashow, Nucl. Phys. 22, 579 (1961)

[4] A. Salam, 8th Nobel Symposium Lerum, Sweden, May 19-25, 1968, Conf. Proc. C680519, 367 (1968)

[5] S. Weinberg, Phys. Rev. Lett. 19, 1264 (1967)

[6] Y. X. Zhao 2017 (*Preprint* arXiv:1701.02780v1)

[7] C. Y. Prescott, *et al.* Phys. Lett. B 77, 347-352 (1978)[doi:10.1016/0370-2693(78)90722-0]

[8] SoLID (Solenoidal Large Intensity Device) Updated Preliminary Conceptual Design Report, The SoLID Collaboration, November 2019.

[9] J. Arrington it et al. The Solenoid Large Intensity Device (SoLID) for JLab 12 GeV, 2022 (*Preprint* arXiv:2209.13357)

[10] D. Wang *et al.* Measurement of parity-violating asymmetry in electron-deuteron inelastic scattering, The PVDIS collaboration (2015)

[11] A. D. Martin, , R. G. Roberts, W. J. Stirlinga, and R. S. Thorned, Scheme dependence, leading order and higher twist studies of MRST partons (*Preprint* arXiv:hep-ph/9808371)

6 Code Implementation

Below we can find the code implementation for Monte Carlo method from Section 4.1.

```
import numpy as np
import matplotlib.pyplot as plt
import math as m
import csv
from random import gauss
from scipy.optimize import fmin,fmin_cg,curve_fit
from scipy import misc
from sympy import diff
from sympy import symbols
from sympy import lambdify
from sympy import evalf
import scipy as sp
from sympy import symbols, Matrix, Function, simplify, exp, hessian, solve,
                                            init_printing
init_printing()
from sympy import lambdify

file = open('/home/felipe/labpro/solid.csv')

csvreader = csv.reader(file)
x = []
Q2 = []
APV_err = []
for row in csvreader:
    x.append(float(row[0]))
    Q2.append(float(row[1]))
    APV_err.append(float(row[2])/100)

x = np.array(x)
Q2 = np.array(Q2)
APV_err = np.array(APV_err)

def Apvew(Q2,x,a1,a3): #Apv from the Standard Model, Ekectroweak sector

    mp = 0.93828 #GeV
    s = 11 #GeV
    GF = 1.1663787*(10**-5) #GeV^-2
    sc = 0.007297351 #structure constant
    coeff = ( GF*Q2 )/( 4*np.sqrt(2)*np.pi*sc ) # coefficient
    y = ( Q2/x )*( 1/( s-mp**2 ) ) # from deep inelastic scattering
                                                relations
    z = coeff*(a1 + a3*( (1 - (1-y)**2 )/( 1 + (1-y)**2) ) )

    return z

a10 = 0.864
a30 = 0.096

i = 0
MVpvew = []
for i in range(15):
```

```
    MVpvew.append( Apvew( Q2[i] , x[i] , a10 , a30 ) )

MVpvew = np.array(MVpvew)

APV_err = np.array(MVpvew*APV_err)
plt.hist(Adpv[0],bins=1000)
plt.show()

squared_APV_err = [Err_Adpv ** 2 for Err_Adpv in Err_Adpv]
stat = np.diag( squared_APV_err )

def Adpvf( Q2 , x , a1 , a3 , bht, bcsv ):

    x1 = 1/((1-x)**3)

    x2 = 1 + bht*x1*(1/Q2) + bcsv*(x**2)

    x3 =  Apvew(Q2,x,a1,a3)*x2

    return x3


def incov(s):

    def cov(s):

    z = stat + err(s)

    return z

    y = np.linalg.inv( cov(s) )

    return y


l = 5
for n in range(l):

    i = 0
    Adpv = []
    for i in range(15):
        Adpv.append( gauss(MVpvew[i], APV_err[i]))

    Adpv = np.array(Adpv)



    def chi2(bht , bcsv , a1 , a3, s ):

        y = Adpv-Adpvf(Q2,x,a1,a3,bht,bcsv)
        p = 0

        for i in range(15):
            for j in range(15):
                p += y[i]*incov(s)[i,j]*y[j]
```

```
    return p


s=0

def best(cbht , cbcsv , ca1 , ca3 , s):

    bestbht,bestbcsv,besta1,besta3 =fmin(lambda x,x1: chi2(x[0],x[1],x[2
                                        ],x[3],x1),
                                    [cbht,cbcsv,ca1,ca3],args=(s,))

    p = bestbht,bestbcsv,besta1,besta3

    chi2_min = chi2(bestbht,bestbcsv,besta1,besta3,s)


    return p

bvalue = best(0.,0.,a10,a30,s)

x1, x2, x3, x4 = symbols('x1 x2 x3 x4')
f, g, h = symbols('f g h', cls=Function)

X = Matrix([x1,x2,x3,x4])

def chi2(x1,x2,x3,x4):

    y = Adpv-Adpvf(Q2,x,x4,x3,x1,x2)
    p = 0

    for i in range(15):
        for j in range(15):
            p += y[i]*incov(0)[i,j]*y[j]

    return Matrix([p])

h = simplify(hessian(chi2(x1,x2,x3,x4), X))
s = (x1, x2, x3, x4)
g_func = lambdify(s, h, modules='numpy')

U = np.linalg.inv(0.5*g_func(bvalue[0],bvalue[1],bvalue[2],bvalue[3]))

err_ = np.sqrt(np.diag(U))


if (1 > 0):

    #params = (bht,bcsv,a1,a3,ebht,ebcsv,ea1,ea3,U)
    params = ( bvalue[0], bvalue[1], bvalue[2], bvalue[3], err_[0], err_
                                        [1], err_[2], err_[3] )
        #rint(("{} "*len(params)).format(*params))
    file = open('/home/felipe/labpro/pd10000.txt',"a")
    print(("{} "*len(params)).format(*params))
    file.write(("{},"*len(params)).format(*params)+"\n")
```

```
        file.close()

#%%

file1 = open("/home/felipe/labpro/pd10.txt")
bht1 = np.genfromtxt("/home/felipe/labpro/pd10.txt",dtype=float,usecols=0,
                                        delimiter=",")
bcsv1 = np.genfromtxt('/home/felipe/labpro/pd10.txt',dtype=float,usecols=1,
                                        delimiter=",")
a11 = np.genfromtxt('/home/felipe/labpro/pd10.txt',dtype=float,usecols=2,
                                        delimiter=",")
a31 = np.genfromtxt('/home/felipe/labpro/pd10.txt',dtype=float,usecols=3,
                                        delimiter=",")

file2 = open("/home/felipe/labpro/pd100.txt")
bht2 = np.genfromtxt("/home/felipe/labpro/pd100.txt",dtype=float,usecols=0,
                                        delimiter=",")
bcsv2 = np.genfromtxt('/home/felipe/labpro/pd100.txt',dtype=float,usecols=1,
                                         delimiter=",")
a12 = np.genfromtxt('/home/felipe/labpro/pd100.txt',dtype=float,usecols=2,
                                        delimiter=",")
a32 = np.genfromtxt('/home/felipe/labpro/pd100.txt',dtype=float,usecols=3,
                                        delimiter=",")

file3 = open("/home/felipe/labpro/pd1000.txt")
bht3 = np.genfromtxt("/home/felipe/labpro/pd1000.txt",dtype=float,usecols=0,
                                         delimiter=",")
bcsv3 = np.genfromtxt('/home/felipe/labpro/pd1000.txt',dtype=float,usecols=1
                                        , delimiter=",")
a13 = np.genfromtxt('/home/felipe/labpro/pd1000.txt',dtype=float,usecols=2,
                                        delimiter=",")
a33 = np.genfromtxt('/home/felipe/labpro/pd1000.txt',dtype=float,usecols=3,
                                        delimiter=",")

file4 = open("/home/felipe/labpro/pd10000.txt")
bht4 = np.genfromtxt("/home/felipe/labpro/pd10000.txt",dtype=float,usecols=0
                                        , delimiter=",")
bcsv4 = np.genfromtxt('/home/felipe/labpro/pd10000.txt',dtype=float,usecols=
                                        1, delimiter=",")
a14 = np.genfromtxt('/home/felipe/labpro/pd10000.txt',dtype=float,usecols=2,
                                         delimiter=",")
a34 = np.genfromtxt('/home/felipe/labpro/pd10000.txt',dtype=float,usecols=3,
                                         delimiter=",")

#%%

def mean(list):
    y = 1/len(list)
    return y*sum(list)

def s2(list):
    y = 1/(len(list)-1)
    return y*sum((list-mean(list))**2)

def covm(x1,x2,x3,x4):
```

```
    def cov(list1,list2):
        y = 1/(len(list1)-1)
        z = sum((list1 - mean(list1))*(list2 - mean(list2)))
        return y*z

    U = [[cov(x1,x1),cov(x1,x2),cov(x1,x3),cov(x1,x4)],[cov(x2,x1),cov(x2,x2
                                                ),cov(x2,x3),cov(x2,x4)]
         ,[cov(x3,x1),cov(x3,x2),cov(x3,x3),cov(x3,x4)],[cov(x4,x1),cov(x4,
                                                    x2),cov(x4,x3),cov(x4,x4)]]
    return U

U1 = covm(bht1,bcsv1,a11,a31)
U2 = covm(bht2,bcsv2,a12,a32)
U3 = covm(bht3,bcsv3,a13,a33)
U4 = covm(bht4,bcsv4,a14,a34)

#%%

def corrm(x1,x2,x3,x4):

    def cov(list1,list2):
        y = 1/(len(list1)-1)
        z = sum((list1 - mean(list1))*(list2 - mean(list2)))
        return y*z

    U = [[cov(x1,x1),cov(x1,x2),cov(x1,x3),cov(x1,x4)],[cov(x2,x1),cov(x2,x2
                                                ),cov(x2,x3),cov(x2,x4)]
         ,[cov(x3,x1),cov(x3,x2),cov(x3,x3),cov(x3,x4)],[cov(x4,x1),cov(x4,
                                                    x2),cov(x4,x3),cov(x4,x4)]]

    U = np.array(U)

    rows, cols = np.shape(U)

    corr_mat = np.zeros((cols, cols))

    for i in range(cols):

        for j in range(cols):

            # not here that we are just normalizing the covariance matrix
            corr_mat[i,j] = U[i,j] / (np.sqrt(U[i,i]) * np.sqrt(U[j,j]) )

    return corr_mat

#%%

plt.imshow(corrm(bht1,bcsv1,a11,a31))
plt.colorbar()
plt.axis('off')
plt.plot()
plt.savefig('d1.png', bbox_inches='tight',pad_inches = 0,dpi=700)
plt.show()
```

Below the final code from which we have analyzed the realistic case as shown in Section 4.2

```
from matplotlib import pyplot as plt
import numpy as np
from random import seed
from random import gauss
from iminuit import Minuit
from iminuit.cost import LeastSquares
import csv

#a1 = 0.864 in the Standard Model
#a3 = 0.096 in the Standard Model

def APVew(x,Q2,a1,a3): #Apv from the Standard Model, Ekectroweak sector
    mp = 0.93828 #GeV
    s = 11 #GeV
    GF = 1.1663787*(10**-5) #GeV^-2
    sc = 0.007297351 #constant
    coeff = ( GF*Q2 )/( 4*np.sqrt(2)*np.pi*sc ) # coefficient
    y = ( Q2/x )*( 1/( s-mp**2 ) ) # from deep inelastic scattering
                                   relations
    z = coeff*(a1 + a3*( (1 - (1-y)**2 )/( 1 + (1-y)**2) ) )
    return z

file = open('/home/felipe/labpro/solid.csv')

csvreader = csv.reader(file)
x = []
Q2 = []
APV_err = []

for row in csvreader:
    x.append(float(row[0]))
    Q2.append(float(row[1]))
    APV_err.append(float(row[2])/100)
#remember the APV_err is in percent, that is why I divide by 100
x = np.array(x)
Q2 = np.array(Q2)
APV_err = np.array(APV_err)
APVSM = APVew(x,Q2,0.864,0.096)
APV_err = APVSM*APV_err #APV_err was in percent, now it has the right scale
APV = gauss(APVSM,APV_err) #generating simulated data: APV = APVSM + APV_err

def AD_PVchi2( x,Q2, b1, b2,a1,a3 ):
    x1 = 1/((1-x)**3)

    x2 = 1 + b1*x1*(1/Q2) + b2*(x**2)

    x3 =  APVew(x,Q2,a1,a3)*x2

    return x3

stat = np.diag( APV_err*APV_err )

def chi2(b1 , b2 , a1 , a3, p):
```

```
    y = (APV- p*AD_PVchi2(x,Q2,b1,b2,a1,a3))
    k = 0

    sp=-0.25
    for i in range(15):
      k += y[i]*y[i]/stat[i,i]

    z = ((p-1)**2)/(sp**2)


    return z + k


m = Minuit(chi2,b1=0,b2=0,a1=0.864 ,a3=0.48, p = 1)
m.migrad()
m.hesse()
print(m.covariance.correlation())
print("b_ht =", m.values[0], "+-", m.errors[0],"\n")
print("b_csv =", m.values[1], "+-", m.errors[1],"\n")
print("a_1 =", m.values[2], "+-", m.errors[2],"\n")
print("a_3 =", m.values[3], "+-", m.errors[3],"\n")
print("p =", m.values[4], "+-", m.errors[4],"\n")
print('case with  30')
```