TODO: update authors and affiliations

Extraction and Validation Framework (EVA)

A. Prokudin,[1] D. Riser,[1] N. Sato,[1] and K. Tezgin[1,*]

[1] *University of Conneticut*

*Electronic address: nsato@jlab.org; Electronic address: kemal.tezgin@uconn.edu; Electronic address: david.riser@uconn.edu

I. SIDIS SETUP

The differential cross section is given by

$$\frac{d\sigma}{dx\ dy\ d\phi_S\ dz\ d\phi_h\ dP_{hT}^2} = \frac{\alpha^2}{xyQ^2}\frac{y^2}{2(1-\varepsilon)}\left(1+\frac{\gamma^2}{2x}\right)\sum_{i=1}^{18} F_i(x,z,Q^2,P_{hT}^2)\beta_i \tag{1}$$

F_i	Standard label	β_i
F_1	$F_{UU,T}$	1
F_2	$F_{UU,L}$	ε
F_3	$F_{UU}^{\cos\phi_h}$	$\sqrt{2\varepsilon(1+\varepsilon)}\ \cos\phi_h$
F_4	$F_{UU}^{\cos 2\phi_h}$	$\varepsilon\ \cos(2\phi_h)$
F_5	$F_{LU}^{\sin\phi_h}$	$\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\sin\phi_h$
F_6	$F_{UL}^{\sin\phi_h}$	$S_{\|\|}\sqrt{2\varepsilon(1+\varepsilon)}\sin\phi_h$
F_7	$F_{UL}^{\sin\ 2\phi_h}$	$S_{\|\|}\varepsilon\sin(2\phi_h)$
F_8	F_{LL}	$S_{\|\|}\lambda_e\sqrt{1-\varepsilon^2}$
F_9	$F_{LL}^{\cos\phi_h}$	$S_{\|\|}\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_h$
F_{10}	$F_{UT,T}^{\sin(\phi_h-\phi_S)}$	$\vert\vec{S}_\perp\vert\sin(\phi_h-\phi_S)$
F_{11}	$F_{UT,L}^{\sin(\phi_h-\phi_S)}$	$\vert\vec{S}_\perp\vert\varepsilon\sin(\phi_h-\phi_S)$
F_{12}	$F_{UT}^{\sin(\phi_h+\phi_S)}$	$\vert\vec{S}_\perp\vert\varepsilon\sin(\phi_h+\phi_S)$
F_{13}	$F_{UT}^{\sin(3\phi_h-\psi_S)}$	$\vert\vec{S}_\perp\vert\varepsilon\sin(3\phi_h-\phi_S)$
F_{14}	$F_{UT}^{\sin\phi_S}$	$\vert\vec{S}_\perp\vert\sqrt{2\varepsilon(1+\varepsilon)}\sin\phi_S$
F_{15}	$F_{UT}^{\sin(2\phi_h-\phi_S)}$	$\vert\vec{S}_\perp\vert\sqrt{2\varepsilon(1+\varepsilon)}\sin(2\phi_h-\phi_S)$
F_{16}	$F_{LT}^{\cos(\phi_h-\phi_S)}$	$\vert\vec{S}_\perp\vert\lambda_e\sqrt{1-\varepsilon^2}\cos(\phi_h-\phi_S)$
F_{17}	$F_{LT}^{\cos\phi_S}$	$\vert\vec{S}_\perp\vert\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\cos\phi_S$
F_{18}	$F_{LT}^{\cos(2\phi_h-\phi_S)}$	$\vert\vec{S}_\perp\vert\lambda_e\sqrt{2\varepsilon(1-\varepsilon)}\cos(2\phi_h-\phi_S)$

- TODO: explicit expression for $|S_\perp|$
- TODO: explicit expression for ε

The 18 structure function in SIDIS at leading-order will be expressed in the context of WW-type approximation in terms of a "minimal" TMD basis using the gaussian ansatz:

$$\mathcal{F}_q(\xi, p_\perp) = \mathcal{K}_q \; \mathcal{C}_q(\xi) \frac{\exp\left(-k_\perp^2/\omega_q\right)}{\pi\omega_q} \tag{2}$$

$$\mathcal{D}_q(\xi, p_\perp) = \mathcal{K}_q \; \mathcal{C}_q(\xi) \frac{\exp\left(-P_\perp^2/\omega_q\right)}{\pi\omega_q}. \tag{3}$$

We denote the transverse momentum of the quark inside a fast moving proton by $\mathbf{k}_\perp$. We use the notation $\mathbf{P}_\perp$ for the transverse momentum of the quark relative to the original parton motion. The structure functions are expressed as

$$F = \sum_q e_q^2 \; \mathcal{K}_q \; \mathcal{F}_q(x) \; \mathcal{D}_q(z) \frac{\exp\left(-P_{hT}^2/\Omega_q\right)}{\pi\Omega_q} \tag{4}$$

$$\Omega_q = z^2 \langle k_\perp^2 \rangle_q + \langle P_\perp^2 \rangle_q \tag{5}$$

type	Name	$\mathcal{K}_q$	$\mathcal{C}_q$
$\mathcal{F}_q$	upol. PDF	1	f_1^q
$\mathcal{F}_q$	pol. PDF	1	g_1^q
$\mathcal{F}_q$	Transversity	1	h_1^q
$\mathcal{F}_q$	Sivers	$\frac{2M^2}{\omega_q}$	$f_{1T}^{\perp(1)q}$
$\mathcal{F}_q$	Boer-Mulders	$\frac{2M^2}{\omega_q}$	$h_1^{\perp(1)q}$
$\mathcal{F}_q$	Pretzelosity	$\frac{2M^4}{\omega_q^2}$	$h_{1T}^{\perp(2)q}$
$\mathcal{F}_q$	Worm Gear	1	$g_{1T}^{\perp q}$
$\mathcal{F}_q$	Worm Gear	1	$h_{1L}^{\perp q}$
$\mathcal{C}_q$	FF	1	D_1^q
$\mathcal{C}_q$	Collins	$\frac{2z^2m_h^2}{\omega_q}$	$H_1^{\perp(1)q}$

TABLE I: The minimal basis

		$\mathcal{K}_q$	$\mathcal{F}_q(x)$	$\mathcal{D}_q(z)$
F_1	$F_{UU,T}$	x	f_1^q	D_1^q
F_2	$F_{UU,L}$	0		
F_3	F_{LL}	x	g_1^q	D_1^q
F_4	$F_{UT}^{\sin(\phi_h+\phi_S)}$	$\frac{2xzP_{hT}m_h}{w_q}$	h_1^q	$H_1^{\perp(1)q}$
F_5	$F_{UT,T}^{\sin(\phi_h-\phi_S)}$	$-\frac{2xzMP_{hT}}{w_q}$	$f_{1T}^{\perp(1)q}$	D_1^q
F_6	$F_{UT,L}^{\sin(\phi_h-\phi_S)}$	0		
F_7	$F_{UU}^{\cos(2\phi_h)}$	$\frac{4xz^2MP_{hT}^2m_h}{w_q^2}$	$h_1^{\perp(1)q}$	$H_1^{\perp(1)q}$
F_8	$F_{UT}^{\sin(3\phi_h-\phi_S)}$	$\frac{2xz^3P_{hT}^3m_hM^2}{w_q^3}$	$h_{1T}^{\perp(2)q}$	$H_1^{\perp(1)q}$
F_9	$F_{LT}^{\cos(\phi_h-\phi_S)}$	$\frac{2xzMP_{hT}}{w_q}$	$g_{1T}^{\perp q}$	D_1^q
F_{10}	$F_{UL}^{\sin(2\phi_h)}$	$\frac{4xz^2MP_{hT}^2m_h}{w_q^2}$	$h_{1L}^{\perp q}$	$H_1^{\perp(1)q}$
F_{11}	$F_{LT}^{\cos\phi_S}$	$-\frac{2M}{Q}x\frac{z^2\langle k_\perp^2\rangle_q\left[P_{hT}^2+\langle P_\perp^2\rangle_q\right]+\langle P_\perp^2\rangle^2}{w_q^2}$	$g_{1T}^{\perp q}$	D_1^q
F_{12}	$F_{LL}^{\cos\phi_h}$	$-\frac{2xzP_{hT}}{Q}\frac{\langle k_\perp^2\rangle_q}{w_q}$	g_1^q	D_1^q
F_{13}	$F_{LT}^{\cos(2\phi_h-\phi_S)}$	$-\frac{2xz^2MP_{hT}^2}{Q}\frac{\langle k_\perp^2\rangle_q}{w_q^2}$	$g_{1T}^{\perp q}$	D_1^q
F_{14}	$F_{UL}^{\sin\phi_h}$	$-\frac{8M^3}{Q}x\frac{z^2\langle k_\perp^2\rangle_q(P_{hT}^2-z^2\langle k_\perp^2\rangle_q)+\langle P_\perp^2\rangle_q^2}{w_q^3}$	$h_{1L}^{\perp q}$	$H_1^{\perp(1)q}$
F_{15}	$F_{LU}^{\sin\phi_h}$	0		
F_{16}	$F_{UU}^{\cos\phi_h}(i)$	$-\frac{8M}{Q}xzP_{hT}m_h\frac{\left[\langle P_\perp^2\rangle_q^2+z^2\langle k_\perp^2\rangle_q(P_{hT}^2-z^2\langle k_\perp^2\rangle_q)\right]}{w_q^3}$	$h_1^{\perp(1)q}$	$H_1^{\perp(1)q}$
F_{16}	$F_{UU}^{\cos\phi_h}(ii)$	$-\frac{2M}{Q}\frac{xzP_{hT}}{M}\frac{\langle k_\perp^2\rangle_q}{w_q}$	f_1^q	D_1^q
F_{17}	$F_{UT}^{\sin\phi_S}(i)$	$-\frac{2M}{Q}x\frac{z^2\langle k_\perp^2\rangle_q\left(P_{hT}^2+\langle P_\perp^2\rangle_q\right)+\langle P_\perp^2\rangle_q^2}{w_q^2}$	$f_{1T}^{\perp(1)q}$	D_1^q
F_{17}	$F_{UT}^{\sin\phi_S}(ii)$	$\frac{4xz^2m_h}{Q}\frac{\langle k_\perp^2\rangle_q\left(-P_{hT}^2+w_q\right)}{w_q^2}$	h_1^q	$H_1^{\perp(1)q}$
F_{18}	$F_{UT}^{\sin(2\phi_h-\phi_S)}(i)$	$-\frac{2M^2}{Q}x\frac{\langle k_\perp^2\rangle_qM}{w_q^2}$	$f_{1T}^{\perp(1)q}$	D_1^q
F_{18}	$F_{UT}^{\sin(2\phi_h-\phi_S)}(ii)$	$-\frac{2M^2}{Q}x\frac{4z^2P_{hT}^2m_h}{w_q^2}$	$h_{1T}^{\perp(2)q}$	$H_1^{\perp(1)q}$
F_{19}	$F_{CAHN}^{\cos(2\phi_h)}$	$\frac{1}{Q^2}\frac{2xz^2P_{hT}^2\langle k_\perp^2\rangle_q^2}{w_q^2}$	f_1^q	D_1^q

TABLE II: SIDIS structure functions up to twist 3

II. SIA SETUP

TODO: add table of SIA structure functions

III. TUTORIAL

TODO: Add more details

Here is a general workflow

- Identify which parameters are needed to be fitted for a given set of TMDs
- Identify available observables to extract the TMDs
- Create xlsx files for each data sets (see the database repo)
- Create and input file (see i.e. fitlab/inputs/upol.py). Here you will specify which data sets are going to be used
- Test in a single fit (fitter.py) that a reasonable χ^2 is obtained
- Once the test fit is ready, proceed to run a multinest which is an nested sampling algorithm to map out the likelihood function. In the input file you must specify the output. Look for the line "output dir " in the input file. It is recommended to store the results in the repo analysis.
- Once the multinest finishes it produces a nestout file at the path that was specified. Use one of the jupyter templates to proceed to analyze the output. You need to place the notebook at the root of analysis. Otherwise the paths wont match.

If you want to run the codes but you want to modify things here and there, create a new workspace via pacman/gen-fitpack, and cd to fitpack. The latter has an exact copy of all the repos, and you can do with that whatever you want. Once you see that some of your modifications should be placed in the actual repositories, proceed to do it.

IV. UNPOLARIZED TMDS

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

V. SIVERS

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

VI. COLLINS FUNCTION FROM e^+e^-

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

VII. TRANSVERSITY

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

VIII. BOER-MULDERS

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

IX. PRETZELOSITY

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

X. WORM-GEAR

- TODO: table of data sets with the following columns: exp, observable, num points, chi2
- TODO: what TMDS are going to be fitted? what is the parametrization
- TODO: distribution of the parameters
- TODO: chi2 distributions for each data sets
- TODO: plot the x or z dependece
- TODO: plot the $k_\perp$ or $P_\perp$ dependece

XI. MONTE CARLO EVENT GENERATOR (MCEG)

In this section we present details of how to construct an MCEG using the analytic expression for the SIDIS cross sections given in Eq. (1). The goal is to generate events for the reaction

$$l + P \to l' + P_h + P_X$$

in the laboratory frame where l and l' are the momenta of the incoming and outgoing leptons respectively, P is the target momentum, P_h is the momentum of the observed hadron and P_X is unobserved system. In this frame, the initial momenta are given by

$$\begin{aligned} l &= (E, 0, 0, E) \\ P &= (M, 0, 0, 0) \ , \end{aligned} \tag{6}$$

where E is the incoming electron beam energy and M is the target nucleon mass.

The SIDIS cross section in Eq. (1) is defined in terms of $x, y, z, \phi_h, \phi_S, P_h^T$. In order to generate the final state momenta, one needs to generate samples of $x, y, z, \phi_h, \phi_S, P_h^T$ that are distributed according to Eq. (1) and covert each sample into the final state momenta l', P_h. Notice that P_X is fixed by momentum conservation.

ϕ_h, ϕ_S and P_h^T in Eq. (1) are defined in the Breit frame. To transform the vectors into such frame we need to perform a rotation R followed by a Lorentz boost B. The rotation can be done by using the Euler–Rodrigues formula by rotating $\vec{q}$ around the vector $\vec{u} = \vec{q} \times (-\hat{e}_z)$ by an angle $\alpha = \arccos(\vec{q} \cdot \hat{e}_z |\vec{q}|))$ with $\hat{e}_z = (0, 0, 1)$. The rotation of a any vector $\vec{x}$ is given by

$$R(\vec{x}) = \vec{x} + 2a(\vec{w} \times \vec{x}) + 2(\vec{w} \times (\vec{w} \times \vec{x}))$$

where $a = \cos(\alpha/2)$ and $\vec{w} = \sin(\alpha/2)\vec{u}/|\vec{u}|$. The boost B is along the z-axis such that $B(R(q)) = (0, 0, 0, -Q)$ and it given by

$$B = \frac{1}{Q} \begin{pmatrix} -q_R^3 & q_R^0 \\ q_R^0 & -q_R^3 \end{pmatrix} \tag{7}$$

where $q_R = R(q)$.

We shall proceed to define a "flat" generator of the phase space needed to be used in a given generic sampling algorithm. The generation of a phase space point proceeds with following steps:

1. Generate $x \in [1/(2ME), 1]$, $Q2 \in [1, 2xME]$ and $\psi \in [0, 2\pi]$ using uniform distributions and compute outgoing lepton momentum

$$l' = E'\left(1, \sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)\right)\ . \tag{8}$$

where $E' = E - Q^2/(2xM)$ and $\cos(\theta) = 1 - Q^2/(2EE')$. In addition compute $q = l - l'$.

2. Generate $z \in [M_h/q^0, 1]$ using uniform distribution and compute the energy of the outgoing hadron $P_h^0 = zq^0$ and the norm of outgoing hadron's 3-vector $|\vec{P}_h| = \sqrt{P_h^0 - M_h^2}$.

3. Generate $P_h^T \in [0, |\vec{P}_h|]$ and $\phi \in [0, 2\pi]$ using uniform distributions and a random sign $\in \{1, -1\}$ in order to generate outgoing hadron's momentum

$$P_h = \left(P_h^0, P_h^T\cos(\phi), P_h^T\sin(\phi), \text{sign}\ \sqrt{|\vec{P}_h|^2 - (P_h^T)^2}\right) \tag{9}$$

4. Rotate and boost all the 4-vectors l, P, l', P_h, S as explained above and proceed to compute the phase space parameters that enters in Eq. (1)

$$x = \frac{Q^2}{2P\cdot q} \tag{10}$$

$$y = \frac{P\cdot q}{P\cdot l} \tag{11}$$

$$z = \frac{P\cdot P_h}{P\cdot q} \tag{12}$$

$$\phi_h = \pi - \arccos((\vec{l}\times\vec{l'})\cdot(\vec{P}_h\times\vec{q})) \tag{13}$$

$$\phi_S = \frac{\pi}{2} - \arccos((\vec{l}\times\vec{l'})\cdot S) \tag{14}$$

$$P_h^T = \sqrt{(P_h^x)^2 + (P_h^z)^2} \tag{15}$$

The flat sampler outlined above can generate events for arbitrary small x, Q^2, z values. However, the analytic expressions for the SIDIS cross section in Eq. (1) are only valid for $Q^2 > 1$ GeV^2 to stay in the deep inelastic regime, $z > 0.2$ to generate events in the current fragmentation region. The range of x probed by the generator will depend on the beam energy E and for fixed target experiments such as the Jefferson Lab 12 GeV the smallest value at $Q^2 = 1$ GeV^2 is still in the domain of applicability of Eq. (1). The sampler can be embedded within a simple while loop that generate events and exit whenever the phase space parameters are within the desired range.

Issues:

- The flat sampler does not produce flat ϕ_h and ϕ_S
- Maybe we can to do uniform sampling on ϕ_h and ϕ_S and constrain azymuthal angles for l' and p_h