

Application of profile likelihood method for extraction of $A_L(\eta)$ from multiple datasets

Jan Balewski ^{1*}, Justin Stevens ¹, Jinlong Zhang²

Abstract

The **profile likelihood method** (PLM) allows for simultaneous extraction of pairs of physics observables $A_L(+\eta), A_L(-\eta)$ from quads of yields, N_s , depending on beams polarization directions 's' and the STAR detector η -slice. We used Poisson statistics and multiplied the global likelihoods to combine the data from multiple years. For the most complex case we used 16 yields to extract 2 parameters of interest while marginalizing over 9 nuisance parameters. The RooStats package from ROOT/CERN has been extended to account for the constrain on the support of the likelihood function while computing the confidence interval. The ready to use code examples are included.

This STAR note does NOT contain any physics results for W-boson A_L .

Keywords

spin asymmetry, profile likelihood method, Poisson statistics

¹LNS, Massachusetts Institute of Technology, Cambridge, USA

²Shandong University, Jinan, China

*Corresponding author: balewski@mit.edu

Contents

Introduction	1
1 PLM for 2 spin states and one-observable	2
1.1 Model	2
1.2 Total likelihood function L_Ω	2
1.3 Extracting parameter of interest from L_Ω	3
Profile likelihood • Confidence interval	
1.4 Numerical example	4
2 PLM for 4 spin states and two-observables	4
2.1 Model	4
Model for 8 yields • Model for 4 yields	
2.2 Total likelihood	5
2.3 RooStats implementation	6
RooWorkspace • ProfileLikelihoodCalculator • Confidence Interval	
2.4 Numerical results	6
A^W for 2x8-yields • A^W for 2x4-yields	
2.5 A^W from single symmetry η -bin	7
2.6 A^{LL} extraction	7
2.7 Conclusion	7
References	7

Introduction

This work has been motivated by the desire to combine the 2011 and 2102 STAR data sets to extract the W-boson spin asymmetry $A_L(\eta)$. Due to the low statistics in the 2011 dataset we could not justify the use of Gaussian error propagation. Instead we have used a multi-dimensional likelihood, constructed from many Poisson distributions. For consistency,

the same method will be used for extraction of the double spin asymmetry A_{LL} .

The essence of the problem which we need to address is finding the most probable value (MPV) and the confidence interval (CI) of an unknown parameter A, given a pair of measured yields N_+, N_- obeying a Poisson distribution, and knowing that the true relation between those quantities are

$$N_\pm = N^0 (1 \pm A \cdot P) \quad (1)$$

where $N^0 > 0$ is a free parameter of no interest to us and P is a known constant, $|P| < 1$.

In the limit of large statistics the p.d.f. of measured yields N_\pm are well approximated by the Gaussian distribution. One can then solve Eqn. 1 for A

$$A = \frac{1}{P} \frac{N_+ - N_-}{N_+ + N_-} \quad (2)$$

and propagate the statistical errors of N_\pm into one standard deviation of A, $\sigma(A)$

$$\sigma(A) = \frac{2}{P} \sqrt{\frac{N_+ N_-}{(N_+ + N_-)^3}} \quad (3)$$

In such a case the CI $[A - \sigma, A + \sigma]$ corresponds to CL=0.683.

Let's further assume we repeated the experiment K-times, measuring K pairs of yields $N_{\pm,k}$. However, each time the constant P_k was different. We want to again extract a single, common value of A based on the combined data from these K experiments.

In the limit of large statistics we can compute A_k, σ_k for

each dataset k , then compute weighted average

$$\hat{A} = \sum_{k=1}^K A_k w_k, \quad (4)$$

$$w_k = \sigma_k^{-2} / \sum_{k=1}^K \sigma_k^{-2} \quad (5)$$

The procedure described above fails in the low statistic limit, say for N_{\pm} of few, when the Poisson p.d.f. is not symmetric around the central value any more. Consequently, the $1-\sigma$ CI for A_k is not centered around the MPV. For a given CL we have CI $[A_k^{lo}, A_k^{hi}]$ such that $A_k - A_k^{lo} \neq A_k^{hi} - A_k$. This means Eqn. 5 can't be applied to compute the relative weights needed in Eqn. 4.

The alternative approach to finding the MPV of A from many, low statistics experiments is to apply the likelihood method (LM). Knowing the p.d.f. for each measurement, $N_{\pm,k}$, obeys the Poisson distribution, we construct the likelihood of measuring each pair $N_{\pm,k}$, given A, assuming a physics justified model. Then, multiply the likelihoods from all experiments. Finally, find the global maximum and CI of A by marginalization of nuisance parameters.

Section 1 describes the application of the profile likelihood method (PLM) for a simple 2-yield experiment. Section 2 will expand the PLM for a more realistic case of a series 8-yield measurements and simultaneous extraction of multiple parameters of interest.

1. PLM for 2 spin states and one-observable

Let's start with a very simple case of the **profile likelihood method** applied to extract the single-spin asymmetry (SSA) from an experiment using a polarized beam hitting an unpolarized target and involving one detector.

1.1 Model

Let N_{\pm} be the measured yields of W-boson events in our experiment for two opposite polarizations of the beam (\pm). Let μ_{\pm} be the expected values of the yields from our model of the experiment, discussed below.

The W-boson reconstruction algorithm accepts a small fraction of non-W events (*i.e.* background) - this impacts the value of the measured SSA and needs to be corrected for. We have identified 3 dominant background sources indexed by the subscript $i=Z, E, Q$, and W:

'Z' labels $Z \rightarrow e^+e^-$ events accepted if one of leptons misses the BEMC or EEMC,

'E' labels QCD \rightarrow jet-jet events for which one jet heads toward the non-existent East EMC endcap,

'Q' labels other subset of QCD events which sometimes hadronize in such a way that they pass the W reconstruction algorithm. For completeness, 'W' labels W-boson events of interest.

Let f_i denote the fraction of reconstructed event yield, n_i , of a given type :

$$f_i = \frac{n_i}{\sum n_i} ; \quad \sum_i f_i = 1 \quad (6)$$

In general, the SSA for each background process, A_i , may be different and non-zero which leads to the following **model of spin dependent yields**, μ_{\pm} , for all events accepted by the W algorithm:

$$N_{\pm} \rightarrow \mu_{\pm} = \mu_{W\pm} + \mu_{Z\pm} + \mu_{E\pm} + \mu_{Q\pm} \quad (7)$$

$$N_{W\pm} \rightarrow \mu_{W\pm} = l_{\pm} N^0 f_W (1 \pm A^W P) \quad (8)$$

$$N_{Z\pm} \rightarrow \mu_{Z\pm} = l_{\pm} N^0 f_Z (1 \pm A^Z P)$$

$$N_{E\pm} \rightarrow \mu_{E\pm} = l_{\pm} N^0 f_E (1 \pm A^E P)$$

$$N_{Q\pm} \rightarrow \mu_{Q\pm} = l_{\pm} N^0 f_Q (1 \pm A^Q P)$$

Eqn. 7 implies the full model for the spin dependent yields, μ_{\pm} , is the sum of all possible processes

$$\mu_{\pm} = l_{\pm} N^0 [f_W + f_Z + f_E + f_Q \pm P(f_W A^W + f_Z A^Z + f_E A^E + f_Q A^Q)] \quad (9)$$

$$= l_{\pm} N^0 [1 \pm P(\beta A^W + \alpha)] \quad (10)$$

where

$$\beta = \frac{f_W}{f_W + f_Z + f_E + f_Q} \quad (11)$$

$$\alpha = \frac{f_Z A^Z + f_E A^E + f_Q A^Q}{f_W + f_Z + f_E + f_Q} \quad (12)$$

In practice the α -term is much smaller than the statistical uncertainty of the experiment so we will ignore it in this paper. The final model of spin dependent yields is

$$\mu_{\pm} = l_{\pm} N^0 (1 \pm P \beta A^W) \quad (13)$$

The beam polarization, P , is a **constant**. The relative luminosities l_{\pm} are also assumed to be **constants**. l_{\pm} depend on the (very large) number of events recorded by the luminosity monitor for both spin states $N_{LUM\pm}$.

$$l_{\pm} = \frac{2N_{LUM\pm}}{N_{LUM+} + N_{LUM-}}; \quad \frac{1}{2}(l_+ + l_-) = 1 \quad (14)$$

$$\sigma_{l_{\pm}} = 1/\sqrt{N_{LUM+} + N_{LUM-}} \text{ are small} \quad (15)$$

1.2 Total likelihood function L_{Ω}

The total likelihood $L_{\Omega}(A^W, N^0, \beta)$ is constructed as the joint probability using all the information we gather from various sources:

$$L_{\Omega}(A^W, N^0, \beta) \equiv L_{PHY}(A^W) \cdot L_{SPIN}(A^W, N^0, \beta) \cdot L_{BCK}(\beta) \quad (16)$$

where:

- A^W is the SSA we want to extract from the experiment (*i.e.* the variable of interest),

- $L_{PHY}(A^W) = H(1 - |A^W|)$: restricts the range of the physically allowed values of SSA, where $H(x)$ is the step function,
- N_0 and β are nuisance parameters describing the unpolarized expected yield and unpolarized background, respectively,
- $L_{SPIN}(A^W, N_0, \beta) = \prod_i^2 f(N_i | \mu_i(A^W, N_0, \beta))$ is product of Poisson distributions $f(N | \mu)$, describing the probability of measuring N events given the expected value was μ from Eqn. 13,
- $L_{BCK}(\beta) = g(\beta - \hat{\beta}, \sigma_\beta)$ is the probability distribution function for the unpolarized background magnitude, here parametrized as a gaussian with a mean $\hat{\beta}$ and standard deviation σ_β .

The following additional parameters: l_\pm, P , needed to compute the numerical values of μ (see Eqn. 13), are assumed to be constant.

Note: for the practical reasons the ranges of all nuisance parameters are bracketed to $\pm 10\sigma$ around the respective central values.

1.3 Extracting parameter of interest from L_Ω

The following inputs are required to extract the asymmetry of interest A^W

- N_\pm : spin sorted yields from STAR experiment,
- l_\pm : relative luminosities from STAR experiment,
- $\hat{\beta}, \sigma_\beta$: description of background p.d.f., based on embedding, simulations, and theory,
- P : beam polarization from RHIC.

With these inputs we can build the 3-dimensional total likelihood function $L_\Omega(A^W, N^0, \beta)$. To find the central value of the A^W we need to remove the nuisance parameters (N^0 and β) from the problem. One method to accomplish this is to marginalize L_Ω (or integrate over) the nuisance parameters to produce the 1-dimensional likelihood vs. the variable of interest, A^W :

$$L_{\text{marg}}(A^W) = \int dN^0 \int d\beta L_\Omega(A^W, N^0, \beta). \quad (17)$$

Another method to treat the nuisance parameters, described in the PDG statistics review [1] (specifically Sec. 36.3.2.3) and a longer review from Cowan [2], is the profile likelihood method which we will use and should yield the same central value and confidence intervals as the marginalization method.

The profile likelihood method consists of two steps: (i) construction of the profile likelihood $L_{\text{prof}}(A^W)$ and (ii) extraction of the central value and confidence interval for A^W .

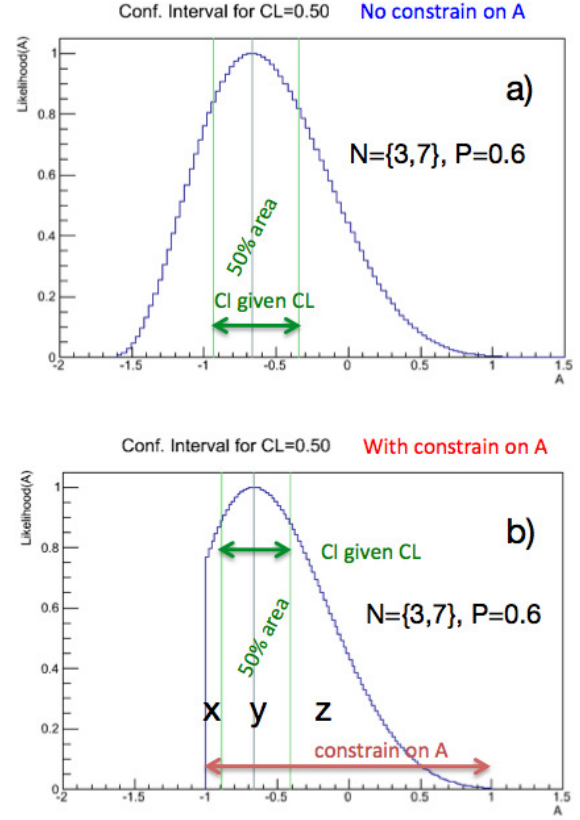


Figure 1. Definition of confidence interval for the case w/o constrain on support (a) and with additional constrain (b).

1.3.1 Profile likelihood

Let's group all nuisance parameters as a vector $\mathbf{v} \equiv (N^0, \beta)$. For each value of A^W , there exist $\hat{\mathbf{v}}(A^W)$ which maximizes $L_\Omega(A^W, \hat{\mathbf{v}})$ defined by Eqn. 16. The **profile likelihood** $L_{\text{prof}}(A^W)$ is defined as

$$L_{\text{prof}}(A^W) = L_\Omega(A^W, \hat{\mathbf{v}}) \quad (18)$$

It is a 1-dimensional likelihood, depending only on A^W . Often one conveniently normalizes the profile likelihood by constructing the **profile likelihood ratio** defined as

$$\lambda_{\text{prof}}(A^W) \equiv \frac{L_{\text{prof}}(A^W)}{L_{\Omega 0}}, \quad (19)$$

where $L_{\Omega 0}$ is the global maximum of the 3D likelihood in the $[A^W, N^0, \beta]$ parameter space.

The central value of A^W is the one which maximizes the profile likelihood $L_{\text{prof}}(A^W)$, or equivalently minimizes the negative log-likelihood, $-\ln L_{\text{prof}}$.

1.3.2 Confidence interval

For a given confidence level (CL) the confidence interval of A^W is computed as the pair $[A_{lo}^W, A_{hi}^W]$ satisfying the integral

$$\int_{A_{lo}^W}^{A_{hi}^W} L_{\text{prof}}(A) dA = CL \cdot \int_{\text{support}} L_{\text{prof}}(A) dA \quad (20)$$

In the absence of constrain on A and for a non-symmetric p.d.f., as show in fig. 1a), we need to impose additionally

$$L_{\text{prof}}(A_{lo}^W) = L_{\text{prof}}(A_{hi}^W) \quad (21)$$

for unambiguous definition of CI.

The presence of the constrain on the support $L_{PHY}(A^W)$, see Fig. 1b), complicates this picture slightly. For clarity, let's assume that the lower bound is closer to the maximum (as in the figure). In general two CI values divide the whole area on 3 parts, labeled x,y,z.

$$\int_x + \int_y + \int_z = \int_{-1}^{+1} L_{\text{prof}}(A) dA \quad (22)$$

In particular, \int_x may be zero if chosen CL is too large. In such case we set $A_{lo}^W = -1$ (i.e. the lower boundary of the constrain) and A_{hi}^W is defined by the modified relation

$$\int_{-1}^{A_{hi}^W} L_{\text{prof}}(A) dA = CL \cdot \int_{-1}^{+1} L_{\text{prof}}(A) dA \quad (23)$$

1.4 Numerical example

For numerical computation we will use the RooStats [3] (extension of CERN root). For educational purposes we have prepared several ready to use macros placed in the MIT disc at BNL [4].

The code used to produce Fig. 1 is named `spin2Asy_constrain.C`. Since the original RooStat did not properly handle the constraint on the asymmetry to be within the physically allowed range, we developed our own after-burner macro

`getIntervGivenConstrSimple.C`, used in all sections of this paper.

2. PLM for 4 spin states and two-observables

In section 1 we have applied the profile likelihood method to a simplified case of extracting one parameter of interest (A^W) out of a pair of measurements (N_+, N_-), using a model (Eqn. 13) with two additional nuisance parameters (N^0, β) and 3 fixed parameters (l_{\pm}, P). The complexity of the real world problem discussed in this section, extraction of $A_L(\eta)$ for W-boson measured at STAR, is much larger.

2.1 Model

In the following we will retain the naming convection of physical quantities, however we will add additional indexing. The indexes s, η, k denote the following:

s spin state of colliding beams, $s = [++, +-, -+, --]$,
 η pseudorapidity of 2 detector regions, $\eta = [\eta_1, \eta_2]$,
 k labels the datasets, $k = [1, 2, \dots, K]$.

Let's review the previously defined quantities with extended indexing:

$N_{s\eta k}$ are yields **measured** by STAR for spin state (s), detector (η), dataset (k),

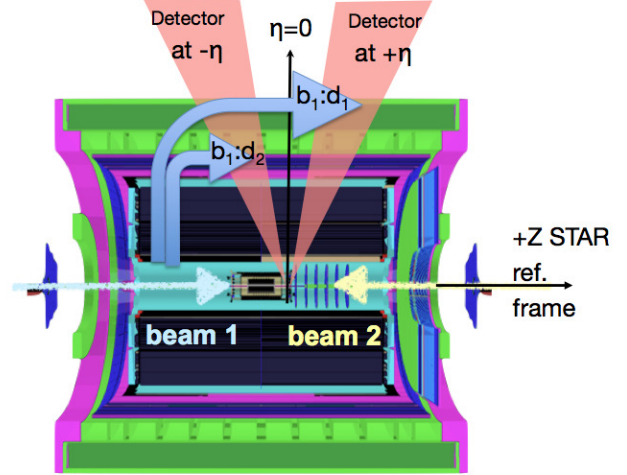


Figure 2. Definition of beam direction and signs of the angles in the detector with respect to polarized beam needed to define dependence of polarized yields on SSA & DSA in eqs. 25.

$\mu_{s\eta k}$ are yields **predicted** by the model defined below.

For clarity let's ignore, for the moment, the dataset index k . The generic formula for the model $\mu_{s\eta}$ depends on similar parameters as Eqn. 13:

$$\mu_{s\eta} = l_s N_{\eta}^0 \left[1 \oplus_s P_1 \beta_{\eta} A_{\eta'}^W \oplus_s P_2 \beta_{\eta} A_{\eta''}^W \oplus_s P_1 P_2 \beta_{\eta} A^{LL} \right] \quad (24)$$

where \oplus_s means the sign switch depending on the spin state 's'. The index of A^W depends on the angle between polarized beam and detector η -bin. The definitions of the parameters used in Eqn. 24 are below:

l_s are the relative luminosity corrections,
 normalization: $\sum_s l_s = 4$, do not depend on the detector,

N_{η}^0 are the predicted spin-average yields, nuisance params,
 change with the detector

P_1, P_2 are the beam polarization magnitudes,

$A_{\eta'}, A_{\eta''}^W$ are the SSAs for a pair of symmetric η -bins with respect to the polarized beam; which are parameters of interest,

β_{η} are the unpolarized corrections to A_{η}^W , depend on detector angle, nuisance parameter,

A^{LL} is the DSA nuisance param, which has no $\pm\eta$ detector dependence.

2.1.1 Model for 8 yields

In total there are 8 different possibilities for the index $s\eta$ defined in the full model (Eqn. 24). Below we will write them

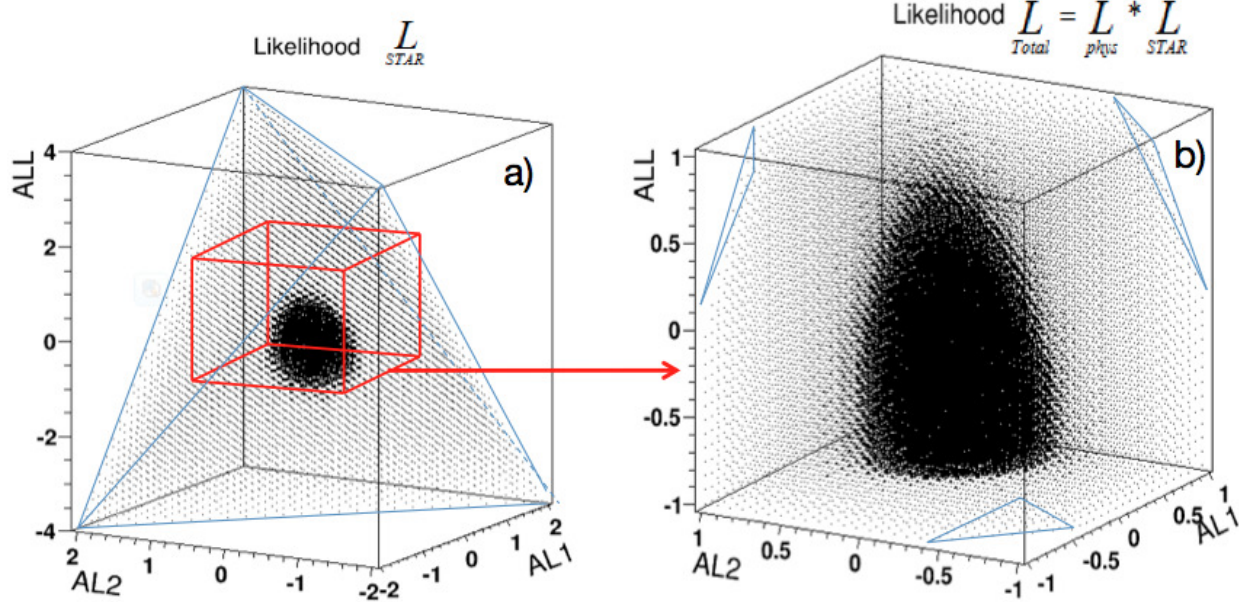


Figure 3. Illustration of applying the physics constrain, L_{PHY} , on the support of the total likelihood for Eqn. 16 . a) no constrains, b) after constrains are applied the support has been reduced.

explicitly

$$\begin{aligned}
 \mu_{++1} &= l_{++1} N_1^0 [1 + P_1 \beta_1 A_1^W + P_2 \beta_1 A_2^W + P_1 P_2 \beta_1 A^{LL}] \\
 \mu_{+-1} &= l_{+-1} N_1^0 [1 + P_1 \beta_1 A_1^W - P_2 \beta_1 A_2^W - P_1 P_2 \beta_1 A^{LL}] \\
 \mu_{-+1} &= l_{-+1} N_1^0 [1 - P_1 \beta_1 A_1^W + P_2 \beta_1 A_2^W - P_1 P_2 \beta_1 A^{LL}] \\
 \mu_{--1} &= l_{--1} N_1^0 [1 - P_1 \beta_1 A_1^W - P_2 \beta_1 A_2^W + P_1 P_2 \beta_1 A^{LL}] \\
 \mu_{++2} &= l_{++2} N_2^0 [1 + P_1 \beta_2 A_1^W + P_2 \beta_2 A_2^W + P_1 P_2 \beta_2 A^{LL}] \\
 \mu_{+-2} &= l_{+-2} N_2^0 [1 + P_1 \beta_2 A_1^W - P_2 \beta_2 A_2^W - P_1 P_2 \beta_2 A^{LL}] \\
 \mu_{-+2} &= l_{-+2} N_2^0 [1 - P_1 \beta_2 A_1^W + P_2 \beta_2 A_2^W - P_1 P_2 \beta_2 A^{LL}] \\
 \mu_{--2} &= l_{--2} N_2^0 [1 - P_1 \beta_2 A_1^W - P_2 \beta_2 A_2^W + P_1 P_2 \beta_2 A^{LL}]
 \end{aligned} \quad (25)$$

From the mathematical perspective, the model of $\mu_{s\eta}$ defined by Eqns. 25 does not need any justification. However, if you are a curious physicist, Fig. 2 defines directions of both beams and signs of the angles of the detector with respect to polarized beam.

2.1.2 Model for 4 yields

In certain cases we only have a single detector η . Then, we need only 4 equations for the model:

$$\mu_s = l_s N^0 [1 \oplus_s P_1 \beta A_{\eta'}^W \oplus_s P_2 \beta A_{\eta''}^W \oplus_s P_1 P_2 \beta A^{LL}] \quad (26)$$

and,

$$\begin{aligned}
 \mu_{++} &= l_{++} N^0 [1 + P_1 \beta A_1^W + P_2 \beta A_2^W + P_1 P_2 \beta A^{LL}] \\
 \mu_{+-} &= l_{+-} N^0 [1 + P_1 \beta A_1^W - P_2 \beta A_2^W - P_1 P_2 \beta A^{LL}] \\
 \mu_{-+} &= l_{-+} N^0 [1 - P_1 \beta A_1^W + P_2 \beta A_2^W - P_1 P_2 \beta A^{LL}] \\
 \mu_{--} &= l_{--} N^0 [1 - P_1 \beta A_1^W - P_2 \beta A_2^W + P_1 P_2 \beta A^{LL}]
 \end{aligned} \quad (27)$$

2.2 Total likelihood

The total likelihood for one dataset consisting of 8 measured yields $N_{s\eta}$ is a product of all p.d.f.s, in analogy to Eqn. 16,

$$L_{\Omega 8}(A_1^W, A_2^W, v_8) = \prod_{s,\eta}^8 f(N_{s\eta} | \mu_{s\eta}) \prod_{\eta}^2 g(\beta_{\eta}) \prod_{\eta'}^2 H(1 - |A_{\eta'}^W|) H(1 - |A_{LL}|) \quad (28)$$

where v_8 represent 5 nuisance parameters $v_8 = [N_{\eta}^0, \beta_{\eta}, A^{LL}]$. The functions $f(\dots), g(\dots), H(\dots)$ were previously defined in the Sec. 1.2. Fig. 3 illustrates the impact of the constrains, $H(x)$, on the allowed parameter space of the total likelihood function.

Finally, lets allow for multiple datasets and restore the index 'k'. For 2 datasets, $k=1,2$, we measure a total of 16 yields $N_{s\eta k}$ and need to almost double the number of nuisance parameters for the total likelihood. This is the final formula:

$$L_{\Omega 16}(A_1^W, A_2^W, v_{16}) = \prod_{s,\eta,k}^{16} f(N_{s\eta k} | \mu_{s\eta k}) \prod_{\eta,k}^4 g(\beta_{\eta k}) \prod_{\eta'}^2 H(1 - |A_{\eta'}^W|) H(1 - |A_{LL}|) \quad (29)$$

where v_{16} represent 9 nuisance parameters $[N_{\eta k}^0, \beta_{\eta k}, A^{LL}]$.

Similarly, for the two dataset consisting of only 4 measured yields N_s for one pseudorapidity bin of the detector the total likelihood is:

$$L_{\Omega 8}(A_1^W, A_2^W, v_8) = \prod_{s,k}^8 f(N_{sk} | \mu_{sk}) \prod_k^2 g(\beta_k) \prod_k^2 H(1 - |A_{\eta'}^W|) H(1 - |A_{LL}|) \quad (30)$$

2.3 RooStats implementation

In section 1.4 RooStats [3] was mentioned. It is a powerful tool to deal with the profile likelihood of multiple parameters, among many other things. If we define all the models and likelihood PDFs, RooStats will do all the rest of the work for us. To call the RooStats package, 5.28.00 or higher version of ROOT is required. An example of the solution to problem defined by Eqns. 29 and 30 is available at [4].

2.3.1 RooWorkspace

The RooWorkspace is a persistent container for RooFit projects. A workspace can contain and own variables, p.d.f.s, functions and datasets. All objects that live in the workspace are owned by the workspace. The import() method enforces consistency of objects upon insertion into the workspace (e.g. no duplicate object with the same name are allowed) and makes sure all objects in the workspace are connected to each other. The code `creatPDF.C` is an example for how to define a RooWorkspace which contains all the variables, p.d.f.s and datasets that will be used in our analysis.

```
> root -l creatPDF.C
```

Executing this code will create the RooWorkspace and print it out.

2.3.2 ProfileLikelihoodCalculator

In the calculation, all the relative variables should be initialized with the experimental parameters. The function `RooStats::ProfileLikelihoodCalculator` will do the main computation of the profile likelihood.

```
ProfileLikelihoodCalculator plC(*dataY,
*modelConfig);
```

where the `dataY` is `RooDataSet` which contains all the observables (namely, the spin sorted yields for the $W A_L$ analysis), and `modelConfig` is a `RooStats::ModelConfig` which contains the full likelihood function (e.g. Eqn. 29 and Eqn. 30) and the definition of parameters of interest (A^W) for the model (μ).

2.3.3 Confidence Interval

We can't use the output `ProfileLikelihoodCalculator` directly since it does not account correctly for the reduced support due to constraints. We instead call our "after-burner" code, discussed in section 1.3.2. Taking out the profile likelihood ratio from the `ProfileLikelihoodCalculator`, we can get the central value and the confidence interval with given confidence level. The code `getIntervalGivenConstrain.C` is an example to get the result from a profile likelihood ratio that comes from `ProfileLikelihoodCalculator`.

2.4 Numerical results

To allow a cross check by the reader we will report a few results for synthetic data.

2.4.1 A^W for 2x8-yields

Lets assume the following 16 yields, $N_{\eta d}$, for the pair of 2 detectors, for 2 years:

$N_{++1,1}=18, N_{+-1,1}=24, N_{-+1,1}=27, N_{--1,1}=21$
 $N_{++2,1}=25, N_{+-2,1}=14, N_{-+2,1}=33, N_{--2,1}=43$
 $N_{++1,2}=87, N_{+-1,2}=184, N_{-+1,2}=161, N_{--1,2}=226$
 $N_{++2,2}=104, N_{+-2,2}=186, N_{-+2,2}=182, N_{--2,2}=269$

Other parameters are:

$L_{s,1} = \{ 1.0180, 0.9891, 0.9926, 1.0002 \}$
 $L_{s,2} = \{ 0.9950, 1.0077, 0.9933, 1.0040 \}$
 $P_{1,1}=0.49, P_{2,1}=0.49; P_{1,2}=0.55, P_{2,2}=0.57$
 $\beta_{\eta,1} = \{ 0.976, 0.971 \}; \beta_{\eta,2} = \{ 0.967, 0.962 \}$

A^W	Profile likelihood, CL=68.3%				Gaussian method*	
	mpv A^W	A_{lo}^W	A_{hi}^W	δA^W	A^W	σA^W
A_1^W	-0.345	-0.391	-0.299	0.046	-0.345	0.046
A_2^W	-0.424	-0.470	-0.378	0.046	-0.425	0.045

Table 1. A_1^W, A_2^W from RooStats and a comparison with the Gaussian method for 16-yields, input discussed in Sec. 2.4.1. *Gaussian method: calculate A^W for each detector eta bin of each year dataset and then average them with error weight defined by Gaussian error propagation.

Implementation of the RooStats based code for the likelihood function of 16 yields (Eqn. 29) was applied on this numerical example. The most probable values and confidence intervals for a 68% confidence level of A_1^W and A_2^W were obtained simultaneously.

$A_1^W = -0.345$, with confidence interval $[-0.391, -0.299]$
 $A_2^W = -0.424$, with confidence interval $[-0.469, -0.378]$

Table 1 lists results of A_1^W, A_2^W and a comparison with the Gaussian method. Results from both methods are consistent. To reproduce execute:

```
> root -l rdAprofF.C' (0, 2, "AL")'
```

2.4.2 A^W for 2x4-yields

To test the case of one detector we assumed:

$N_{++1}=3, N_{+-1}=2, N_{-+1}=1, N_{--1}=3$
 $N_{++2}=16, N_{+-2}=15, N_{-+2}=16, N_{--2}=10$

Other parameters as in section 2.4.1, except
 $\beta_1 = 0.991, \beta_2 = 0.962$.

For CL of 68% we got:
 $A_1^W = 0.167$, with confidence interval $[-0.065, 0.399]$
 $A_2^W = 0.181$, with confidence interval $[-0.046, 0.403]$

Table 2 lists results of A_1^W, A_2^W and a comparison with the Gaussian method. As expected, the result also compare well for this case. Due to the low statistics, the difference between two methods is more significant. To reproduce execute:

```
> root -l rdAprofF.C' (1, 7, "AL")'
```

A^W	Profile likelihood, CL=68.3%				gauss method	
	mpv A^W	A_{lo}^W	A_{hi}^W	δA^W	A^W	σA^W
A_1^W	0.167	-0.065	0.399	0.232	0.169	0.234
A_2^W	0.182	-0.046	0.403	0.224	0.179	0.224

Table 2. A_1^W, A_2^W from RooStats and comparison with the Gaussian method for 8-yields, input discussed in Sec. 2.4.2.

2.5 A^W from single symmetry η -bin

For the case of single symmetry η -bin, it is similar with the model for 4 yields discribed in Sec. 2.1.2. The only difference is $A_{\eta'}^W = A_{\eta''}^W$, namely, the index η is skipped. So, the parameters of interest are reduced to only one, A^W . To test this case we assumed:

$$N_{++,1}=64, N_{+-,1}=68, N_{-+,1}=89, N_{--,1}=105 \\ N_{++,2}=339, N_{+-,2}=586, N_{-+,2}=559, N_{--,2}=818$$

Other parameters as in section 2.4.1, except $\beta_1 = 0.950, \beta_2 = 0.990$.

For CL of 68% we got:

$$A^W = -0.363, \text{ with confidence interval } [-0.398, 0.348]$$

We can reproduce the result by executing below line,

```
> root -l rdAprofF.C' (0, 8, "AL")'
```

For the $Z A^L$ calculation, we can also use this method by changing the input yields.

2.6 A^{LL} extraction

Previous subsections are focused on A^W , the single spin asymmetry. For A^{LL} , the double spin asymmetry, the case is similar. By setting A^W s as a nuisance parameters and setting A^{LL} as the parameter of interest, we can extract the most probable value and confidence interval of A^{LL} .

We used the sample yields from Sec. 2.4.1 and 2.4.2 to obtain the numerical results of A^{LL} . The most probable value and confidence interval with 68% confidence level for 2x8-yields and 2x4-yields respectively are shown below.

```
> root -l rdAprofF.C' (0, 2, "ALL")'
```

$$A^{LL} = -0.050, \text{ with confidence interval } [-0.134, 0.035].$$

```
> root -l rdAprofF.C' (1, 7, "ALL")'
```

$$A^{LL} = -0.155, \text{ with confidence interval } [-0.556, 0.251]$$

2.7 Conclusion

Based on the knowledge from previous section, this method can be applied to extract $W A_L$ from STAR data from multiple dataset. The Barrel part of STAR detector can be divided into pairs of symmetric pseudorapidity bins, where Eqn 29 applies. Since there is only one endcap at STAR the Eqn. 30 should be applied for this case. Applying the model and likelihood described in Sec. 2.1 and 2.2, we can obtain the

$W A_L$ pseudorapidity dependence in the full η range of the STAR detector. The $W A_{LL}$ and $Z A_L$ also can be extracted via this method.

Acknowledgments

We would like to thank Christine Spinka for recommending use of the profile likelihood method as a reliable tool to extract information from statistically starving event sets.

References

- [1] G.Cowan, "PDG Statistics Review", <http://pdg.lbl.gov/2013/reviews/rpp2012-rev-statistics.pdf>
- [2] G.Cowan, "Statistics for Searches at the LHC", <http://arxiv.org/abs/1307.2487v1>
- [3] Statistical tool developed for LHC <https://twiki.cern.ch/twiki/bin/view/RooStats>
- [4] Directory at MIT disc at BNL containing all example macros is [/star/institutions/mit/balewski/freezer/2013-SN-590](http://star.institutions/mit/balewski/freezer/2013-SN-590). Note, we used `starver SL13d` or a later library at `rcas6nnn` to execute the examples.
- [5] Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0), <http://creativecommons.org/licenses/by-nc-sa/3.0/>