Extracting angular observables with Method of Moments

Marcin Chrząszcz[1]
in collaboration with
Frederik Beaujean, Nicola Serra and Danny van Dyk,

based on *arXiv:1503.04100*

[1] University of Zurich

May 8, 2015

1. Motivation.
2. Method of Moments.
3. Systematic uncertainties.
4. MC toy studies.
5. Conclusions.

Motivation

Likelihood(LL) fits even though widely used suffer from couple of draw backs:

1. In case of small number events LL fits suffer from convergence problems. This behaviour is well known and was observed several times in toys when we done $\mathrm{B} \to \mathrm{K}^{*}\mu\mu$.
2. LL can exhibit a bias when underlying physics model is not well known, incomplete or mismodeled.
3. The LL have problems converging when parameters of the p.d.f. are close to their physical boundaries, so-called "boundary problem"
4. Accessing uncertainty in LL in some cases requires application of computationally expensive Feldman-Cousins method.

University of Zurich[UZH]

Method of Moments

MoM solves the above problems:

Advantages of MoM

- Probability distribution function rapidity converges towards the Gaussian distribution.
- MoM gives an unbias result even with small data sample.
- Insensitive to large class of remodelling of physics models.
- Is completely insensitive to boundary problems.

University of Zurich[UZH]

Method of Moments

MoM solves the above problems:

Advantages of MoM

- Each observable can be determined separately from other.
- Uncertainly follows perfectly $1/\sqrt{N}$ scaling.

Method of Moments

MoM solves the above problems:

Advantages of MoM

- Each observable can be determined separately from other.
- Uncertainly follows perfectly $1/\sqrt{N}$ scaling.

Drawback:

Advantages of MoM

- Estimated uncertainty in MoM is larger then the ones from LL.

University of Zurich[UZH]

Introduction to MoM

Let us a define a probability density function p.d.f. of a decay:

$$P(\vec{\nu},\vec{\vartheta}) \equiv \sum_i S_i(\vec{\nu}) \times f_i(\vec{\vartheta}) \tag{1}$$

Let's assume further that there exist a dual basis: $\{f_i(\vec{\vartheta})\}$, $\{\tilde{f}_i(\vec{\vartheta})\}$ that the orthogonality relation is valid:

$$\int_\Omega \mathrm{d}\vec{\vartheta}\, \tilde{f}_i(\vec{\vartheta}) f_j(\vec{\vartheta}) = \delta_{ij} \tag{2}$$

Since we want to use MoM to extract angular observables it's normal to work with Legendre polynomials. In this case we can find self-dual basis:

$$\forall_i \tilde{f}_i = f_i \ , \tag{3}$$

just by applying the ansatz: $\tilde{f}_i = \sum_i a_{ij} f_j$.

University of Zurich[UZH]

Determination of angular observables

Thanks to the orthonormality relation Eq. 2 one can calculate the $S_i(\vec{\nu})$ just by doing the integration:

$$S_i(\vec{\nu}) = \int_\Omega d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \quad (4)$$

We also need to integrate out the $\vec{\nu}$ dependence:

$$\langle S_i \rangle = \int_\Theta d\vec{\nu} \int_\Omega d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \quad (5)$$

MoM is basically performing integration in Eq. 5 using MC method:

$$E[S_i] \to \widehat{E[S_i]} = \frac{1}{N} \sum_{k=1}^{N} \tilde{f}(x_k)$$

University of Zurich[UZH]

Determination of angular observables

Thanks to the orthonormality relation Eq. 2 one can calculate the $S_i(\vec{\nu})$ just by doing the integration:

$$S_i(\vec{\nu}) = \int_{\Omega} d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \tag{4}$$

We also need to integrate out the $\vec{\nu}$ dependence:

$$\langle S_i \rangle = \int_{\Theta} d\vec{\nu} \int_{\Omega} d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \tag{5}$$

MoM is basically performing integration in Eq. 5 using MC method:

$$E[S_i] \to \widehat{E[S_i]} = \frac{1}{N} \sum_{k=1}^{N} \tilde{f}(x_k)$$

University of Zurich[UZH]

Determination of angular observables

Thanks to the orthonormality relation Eq. 2 one can calculate the $S_i(\vec{\nu})$ just by doing the integration:

$$S_i(\vec{\nu}) = \int_\Omega d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \tag{4}$$

We also need to integrate out the $\vec{\nu}$ dependence:

$$\langle S_i \rangle = \int_\Theta d\vec{\nu} \int_\Omega d\vec{\vartheta} P(\vec{\nu}, \vec{\vartheta}) \tilde{f}_i(\vec{\vartheta}) \tag{5}$$

MoM is basically performing integration in Eq. 5 using MC method:

$$E[S_i] \rightarrow \widehat{E[S_i]} = \frac{1}{N} \sum_{k=1}^{N} \tilde{f}(x_k)$$

University of Zurich[UZH]

Uncertainty estimation

MoM provides also a very fast and easy way of estimating the statistical uncertainty:

$$\sigma(S_i) = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\tilde{f}_i(x_k) - \widehat{S}_i)^2} \tag{6}$$

and the covariance:

$$\mathrm{Cov}[S_i, S_j] = \frac{1}{N-1}\sum_{k=1}^{N}[\widehat{S}_i - \tilde{f}_i(x_k)][\widehat{S}_j - \tilde{f}_j(x_k)] \tag{7}$$

Thanks to the CLT both equations are satisfied.

University of Zurich[UZH]

MoM provides also a very fast and easy way of estimating the statistical uncertainty:

$$\sigma(S_i) = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\tilde{f}_i(x_k) - \widehat{S}_i)^2} \tag{6}$$

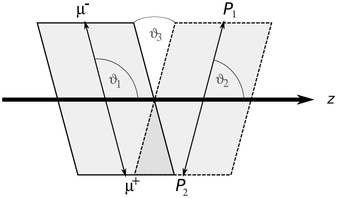
and the covariance:

$$\mathrm{Cov}[S_i, S_j] = \frac{1}{N-1}\sum_{k=1}^{N}[\widehat{S}_i - \tilde{f}_i(x_k)][\widehat{S}_j - \tilde{f}_j(x_k)] \tag{7}$$

Thanks to the CLT both equations are satisfied.

Partial Waves mismodeling

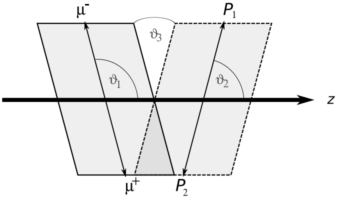
- Let us consider a decay of $B \to P_1 P_2 \mu^- \mu^+$.
- In terms of angular p.d.f. is expressed in terms of partial-wave expansion.
- For $B \to K\pi\mu^-\mu^+$ system, S,P,D waves have been studied.
  - The muon system of this kind of decays has a fixed angular dependence in terms of ϑ_1 and ϑ_3.
  - The hadron system can have arbitrary large angular momentum.



Partial Waves mismodeling

- One can write the p.d.f. separating the hadronic system:

$$P(\cos\vartheta_1, \cos\vartheta_2, \vartheta_3) = \sum_i S_i(\vec{\nu}, \cos\vartheta_2) f_i(\cos\vartheta_1, \vartheta_3) \quad (8)$$



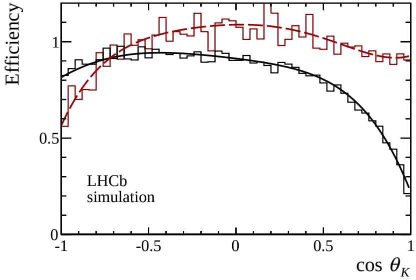
  - $S_i(\vec{\nu}, \cos\vartheta_2)$ can be further expend in terms of Legendre polynomials $p_l^{|m|}(\cos\vartheta_2)$:

$$S_i(\vec{\nu}, \cos\vartheta_2) = \sum_{l=0}^{\text{inf}} S_{k,l}(\vec{\nu}) p_l^{|m|}(\cos\vartheta_2) \quad (9)$$

  - Experimentally the $S_{k,l}$ are easily accessible, but there is a theoretical difficulty as one would need to sum over infinite number of partial waves.

Detector effects

- Since our detectors are not a perfect devices the angular distribution observed by them are not the distributions that the physics model creates.



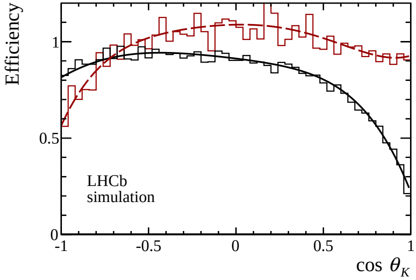
- To take into account the acceptance effects one needs to simulate the a large sample of MC events.
  Try to figure out the efficiency function.
- Try to figure out the efficiency function.
- Number of possibilities.
- Then you can just weight events:

$$\widehat{E[S_i]} = \frac{1}{\sum_{k=1}^{N} w_k} \sum_{k=1}^{N} w_k \tilde{f}(x_k), \;\; w_k = \frac{1}{\epsilon(x_k)}$$

University of Zurich[UZH]

Detector effects

- Since our detectors are not a perfect devices the angular distribution observed by them are not the distributions that the physics model creates.



- To take into account the acceptance effects one needs to simulate the a large sample of MC events.
  Try to figure out the efficiency function.
- Try to figure out the efficiency function.
- Number of possibilities.
- Then you can just weight events:

$$\widehat{E[S_i]} = \frac{1}{\sum_{k=1}^{N} w_k} \sum_{k=1}^{N} w_k \tilde{f}(x_k), \;\; w_k = \frac{1}{\epsilon(x_k)}$$

University of Zurich[UZH]

Unfolding matrix

In general one can write the distribution of events after the detector effects:

$$P^{\rm Det}(x_d) = N \int \int dx_t \ dx_d \ P^{\rm Phys}(x_t) E(x_d|x_t), \qquad (10)$$

where $N^{-1} = \int \int dx_t \ dx_d \ P^{\rm Phys}(x_t) E(x_d|x_t)$ and $(x_d|x_t)$ denotes the efficiency $\epsilon(x_t)$ and resolution of the detector $R(x_d|x_t)$:

$$E(x_d|x_t) = \epsilon(x_t) R(x_d|x_t) \qquad (11)$$

One can define the raw moments:

$$Q_i^{(m)} = \int \int dx_t \ dx_d \ \tilde{f}_i(x_d) P^{(m)}(x_t) E(x_d|x_t) \qquad (12)$$

The m index corresponds to simulation sample that has S_0 and S_m observables set to $\frac{1}{2}$ and rest to zero.

University of Zurich[UZH]

Unfolding matrix

Once again we can use MC estimator:

$$Q_i^{(m)} \to \widehat{Q}_i^{(m)} = \frac{1}{N_t} \sum_i^{N_d} \tilde{f}_i(x_d^{i,m}) \tag{10}$$

Linearity of the integral ensures that there has to exists a linear transformation:

$$\vec{Q} = M\vec{S}, \tag{11}$$

where M is so-called unfolding matrix, given by the formula:

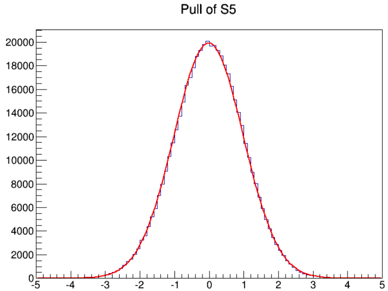
$$M_{ij} = \begin{cases} 2Q_i^{(0)} & j = 0\,, \\ 2\left(Q_i^{(j)} - Q_i^{(0)}\right) & j \neq 0\,, \end{cases} \tag{12}$$

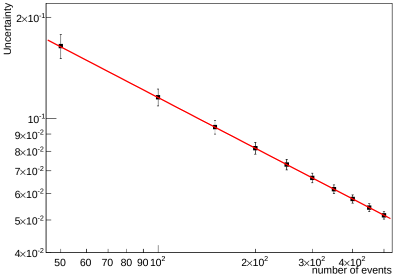
Once we measured the moments Q in data we can invert Eq. 11 and get the $\vec{S}$:

$$\widehat{\vec{S}} = M^{-1}\widehat{\vec{Q}}, \tag{13}$$

Toy Validation

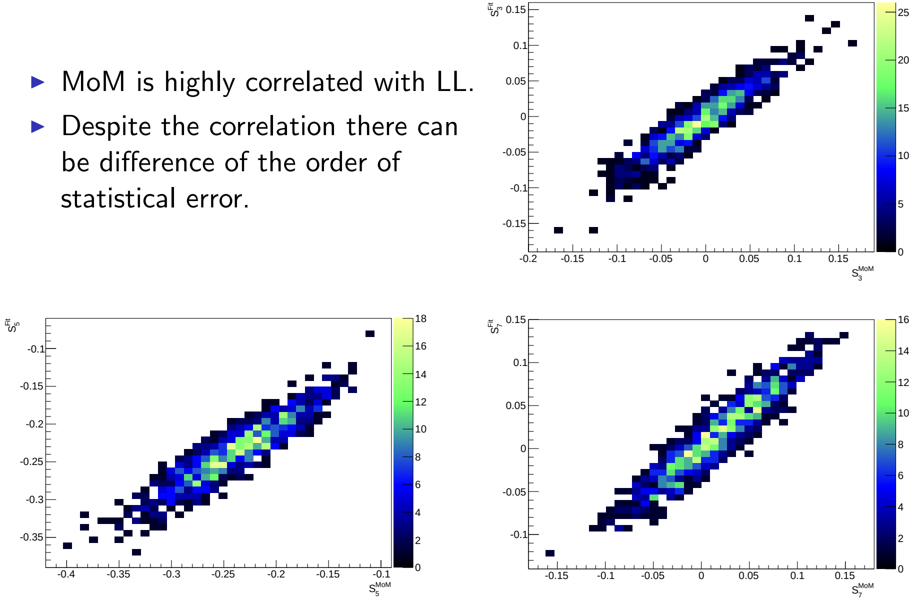
- All the statistics properties of MoM have been tested in numbers of TOY MC.
- As long as you have ~ 30 events your pulls are perfectly gaussian.
- Uncertainty scales with $\frac{\alpha}{\sqrt{n}}$, $\alpha = \mathcal{O}(1)$.
- Never observed any boundary problems.





Correlation of MoM with Likelihood

- MoM is highly correlated with LL.
- Despite the correlation there can be difference of the order of statistical error.



of

Conclusions

1. MoM posses several big advantages with one drawback which is larger statistical uncertainty.
2. Allows us to go smaller q^2 bins (get ready for 1 GeV^2 soon!).
3. Alternative method of extracting the detector effects.
4. Can be applied to various rare decays.