Muon g-2: data combination, fitting and systematic bias

Alex Keshavarzi

Supervisor: Dr. Thomas Teubner

University of Liverpool

a.i.keshavarzi@liverpool.ac.uk

YTF 8, Durham University

15^{th} January 2016

Overview

1. Muon g-2: the basics
   - Motivation
2. Data combination & fitting: the f_k method
   - Data combination
   - The f_k method
   - Systematic bias
3. Unbiased fitting: the R^I_m method
   - The R^I_m method
4. Results
   - $e^+e^- \to \pi^+\pi^-$
   - Comparison with the f_k method
   - Conclusions

The Standard Model (SM): an incomplete theory

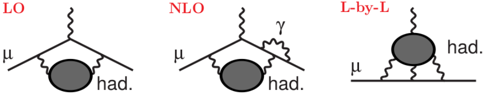
- a_μ is one of the most precisely measured quantities in particle physics, accurate to 0.54ppm.
- BNL experiment revealed $a_\mu^{SM} < a_\mu^{exp}$ by approximately 3.3 standard deviations.
- Any deviation could herald the existence of as-yet-unknown new physics beyond the SM.
- Experiment 4x more accurate after the completion of g-2 experiment at Fermilab!
  $\rightarrow$if mean values stay and with no a_μ^{SM} improvement:
  $\longrightarrow$ 5σ **discrepancy**

The hadronic contribution

$$a_\mu = \frac{(g-2)}{2} = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{had} + a_\mu^{NewPhysics?}$$

Uncertainties from the hadronic sector completely dominate Δa_μ^{SM}!

$$a_\mu^{\rm had} = a_\mu^{\rm had,VP\ LO} + a_\mu^{\rm had,VP\ NLO} + a_\mu^{\rm had,Light-by-Light}$$

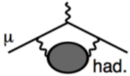


- Most precise prediction by using e^+e^- hadronic cross section data and utilising a dispersion integral.
- Done at LO and NLO (see graphs)
- Now even at NNLO [Steinhauser et. al, PLB734(2014)114]

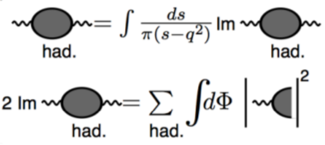
Of these, $a_\mu^{\text{had, VP LO}}$ has the largest uncertainty.

Calculating a_μ^{had}: the basics

Use of data compilation for HVP:



pQCD not useful. Use the dispersion relation and the optical theorem.



$$a_\mu^{\mathrm{had,LO}} = \frac{m_\mu^2}{12\pi^3} \int_{s_{\mathrm{th}}}^{\infty} ds \, \frac{1}{s} \hat{K}(s) \sigma_{\mathrm{had}}(s)$$

- Weight function $\hat{K}(s)/s = \mathcal{O}(1)/s$

$\Longrightarrow$ Lower energies more important

$\Longrightarrow$ $\pi^+\pi^-$ channel: 73% of total $a_\mu^{\mathrm{had,LO}}$

How to get the most precise σ^0_{had}? e^+e^- data:

- Low energies: sum ~ 25 exclusive channels, 2π, 3π, 4π, 5π, 6π, KK, KKπ, KKππ, ηπ, ..., use iso-spin relations for missing channels
- Above ~1.8 GeV: can start to use pQCD (away from flavour thresholds), supplemented by narrow resonances (J/Ψ, Υ)
- Challenge of data combination (locally in √s): from many experiments, in different energy bins, errors from different sources, correlations; must avoid inconsistencies/bias
- σ^0_{had} means `bare' σ, but WITH FSR: RadCorrs [HLMNT: $\delta a_\mu^{\mathrm{had,\,RadCor\ VP+FSR}} = 2 \times 10^{-10}$!]

Fitting global systematic normalisation uncertainties

- Hadronic cross sections σ^0_{had} have to be combined and then fitted. $\rightarrow$ requires statistically valid method!!
- Recent studies (arXiv:0912.2276,1507.02943) have shown that should experimental data include a global normalisation uncertainty, then the choice of fitting method can lead to systematic biases.
- It follows that we must:

1. Review the existing fitting procedure - is there the danger of biased results?
2. Determine a new fit free from bias.
3. Produce and compare results.

Combining & fitting data

- Combine the data in a given channel before integrating.
- Re-bin the data points into *energy clusters* (piecewise constant R).
- Use adaptive clustering algorithm to produce *target clusters*.
  - $\rightarrow$ too small a cluster = precise data overwhelmed
  - $\rightarrow$ too large a cluster = data missed about resonance peaks
- Weighted average for the cross section value R_m is given by

$$R_m = \left[\sum_k \sum_{i=1}^{N^{(k,m)}} \frac{R_i^{(k,m)}}{\left(\mathrm{d}\tilde{R}_i^{(k,m)}\right)^2} \right] \left[\sum_k \sum_{i=1}^{N^{(k,m)}} \frac{1}{\left(\mathrm{d}\tilde{R}_i^{(k,m)}\right)^2} \right]^{-1}$$

$\longrightarrow$ taken as initial values for fit parameters.

where

$$\mathrm{d}\tilde{R}_i^{(k,m)} = \sqrt{\left(\mathrm{d}R_i^{(k,m)}\right)^2 + \left(\mathrm{d}f_k R_i^{(k,m)}\right)^2}$$

$\longrightarrow$ $\mathrm{d}f_k$ is the global normalisation uncertainty of experiment k.

Fitting: the f_k method

- Fit data and minimise a non-linear χ^2-function [HLMNT, 2012]
  $\longrightarrow$ Two fitting parameters: initial cluster values R_m, normalisation factor f_k of each experiment k

$$\chi^2(R_m, f_k) = \sum_{k=1}^{N_{exp}} \left\{ \left(\frac{1-f_k}{df_k} \right)^2 + \left[\sum_{m=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \left(\frac{R_i^{(k,m)} - f_k R_m}{dR_i^{(k,m)}} \right)^2 \right]_{\text{w/o cov. mat}} \right.$$
$$\left. + \left[\sum_{m=1}^{N_{clu}} \sum_{n=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \sum_{j=1}^{N^{(k,n)}} (R_i^{(k,m)} - f_k R_m) C^{-1}(m_i, n_j) (R_j^{(k,n)} - f_k R_n) \right] \right\}$$

The f_k's are multiplicative re-normalisation factors for the data which vary as the χ^2-function is minimised.

$\longrightarrow$ Penalty Trick Method!!

The origin of bias

What is a biased result?

→ *D'Agostini Bias* = Fit favours more precise measurement.

(DOI: 10.1016/0168-9002(94)90719-6)

How do we include a global normalisation uncertainty whilst avoiding a D'Agostini bias?

⟶ Penalty Trick Method!!

PENALTY TERM

NORMALISATION FACTOR

$$\chi^2(R_m, f_k) = \sum_{k=1}^{N_{exp}} \left\{ \left(\frac{1 - f_k}{df_k} \right)^2 + \left[\sum_{m=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \left(\frac{R_i^{(k,m)} - f_k R_m}{dR_i^{(k,m)}} \right)^2 \right]_{\text{w/o cov. mat}} \right.$$

FIT PARAMETERS

Is the f_k method truly free from bias?

Model data: a biased calculation of $a_\mu^{\text{had, LOVP}}$

Consider two measurements: $R_i^{(k,m)}$ and $R_j^{(l,m)}$ with equal (uncorrelated) errors...

$$\mathrm{d}f_k = \mathrm{d}f_l \equiv \mathrm{d}f \quad ; \quad \mathrm{d}R_i^{(k,m)} = \mathrm{d}R_j^{(l,m)} \equiv \mathrm{d}R$$

$$\text{Unbiased solution} \longrightarrow R_m = \bar{R}_m = \frac{1}{2}(R_i^{(k,m)} + R_j^{(l,m)})$$

Minimising w.r.t R_m and f_k and substituting, we find

$$R_m = \bar{R}_m(1+\beta_m) \ ,$$

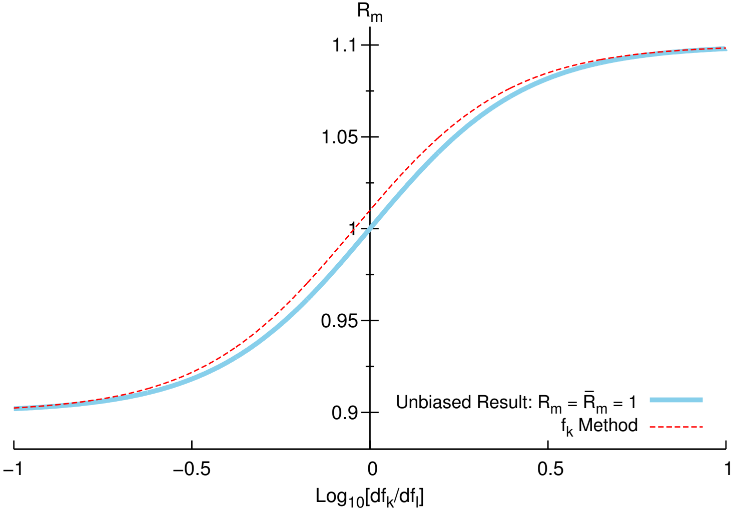
$$\beta_m = \frac{1}{2\bar{R}_m^2(\mathrm{d}f)^2}\Bigg(-(\mathrm{d}R)^2 - R_i^{(k,m)}R_j^{(l,m)}(\mathrm{d}f)^2 + \sqrt{4\bar{R}_m^4(\mathrm{d}f)^4 - 4\bar{R}_m^2 R_i^{(k,m)}R_j^{(l,m)}(\mathrm{d}f)^4 + \Big(R_i^{(k,m)}R_j^{(l,m)}(\mathrm{d}f)^2 + (\mathrm{d}R)^2\Big)^2}\Bigg)$$

$\Rightarrow$ β_m is the *bias contribution* to the cluster centre R_m.

(due to non-linear nature of χ^2 function)

Model data: a biased calculation of $a_\mu^{\text{had, LOVP}}$

$$R_i^{(k,m)} = 0.9 \text{ and } R_j^{(l,m)} = 1.1 \quad ; \quad \mathrm{d}R_i^{(k,m)} = \mathrm{d}R_j^{(l,m)} \equiv \mathrm{d}R = 0$$



Fixing the covariance matrix

Covariance matrices defined incorrectly!

Covariance matrix dependence on normalisation uncertainties comes from measurements themselves.

→ allows error propagation in the fit to skew the theory value!

1. Remove normalisation factors f_k and penalty term from penalty trick definition of χ^2-function - Linear error function!
2. Remove any previous treatment of normalisation uncertainties from all covariance matrices.
3. Fix covariance matrices with normalisation uncertainties throughout fit - choose to fix with guess value for cluster R_m^0:

$$C_k(m_i, n_j) = \mathbf{c}_k(m_i, n_j) + (\mathsf{d}f_k)^2 R_m^0 R_n^0$$

New definition of our now linear χ^2 function:

$$\chi^2(R_m) = \sum_{m=1}^{N_{clu}} \sum_{n=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \sum_{j=1}^{N^{(k,n)}} \left(R_i^{(k,m)} - R_m\right) C_k^{-1}(m_i, n_j) \left(R_j^{(k,n)} - R_n\right)$$

Model data: a bias free example

Consider two measurements: $R_i^{(k,m)}$ and $R_j^{(l,m)}$ with equal (uncorrelated) errors...

$$\chi^2(R_m) = \sum_{i=1}^{N^{(k,m)}} \frac{(R_i^{(k,m)} - R_m)^2}{(\mathrm{d}R_i^{(k,m)})^2 + (\mathrm{d}f_k)^2 (R_m^0)^2}$$

Minimising w.r.t R_m, we find

$$R_m = \left[\sum_{i=1}^{N^{(k,m)}} \frac{R_i^{(k,m)}}{(\mathrm{d}R_i^{(k,m)})^2 + (\mathrm{d}f_k)^2 (R_m^0)^2} \right] \left[\sum_{i=1}^{N^{(k,m)}} \frac{1}{(\mathrm{d}R_i^{(k,m)})^2 + (\mathrm{d}f_k)^2 (R_m^0)^2} \right]^{-1}$$

Reintroduce $\rightarrow \mathrm{d}f_k = \mathrm{d}f_l \equiv \mathrm{d}f \;\; ; \;\; \mathrm{d}R_i^{(k,m)} = \mathrm{d}R_j^{(l,m)} \equiv \mathrm{d}R$

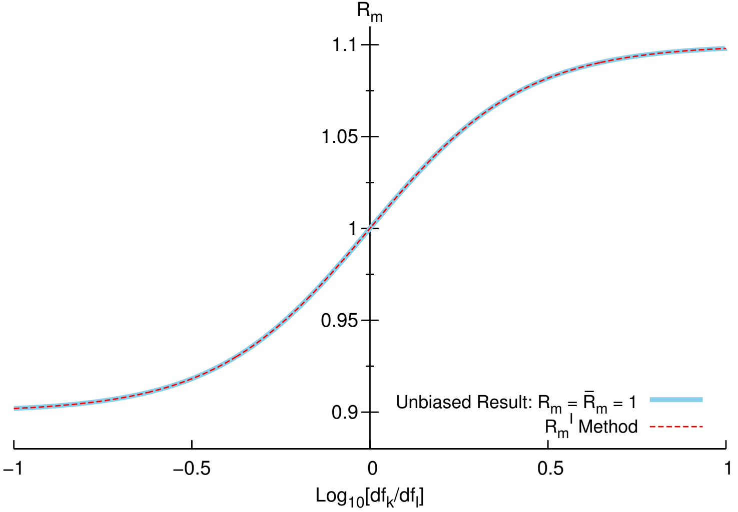
$$\left(\text{Unbiased solution} \longrightarrow R_m = \bar{R}_m = \frac{1}{N^{(k,m)}} \sum_{i=1}^{N^{(k,m)}} R_i^{(k,m)} \right)$$

$$R_m = \left[\frac{R_i^{(k,m)} + R_j^{(l,m)}}{(\mathrm{d}R)^2 + (\mathrm{d}f)^2 (R_m^0)^2} \right] \left[\frac{2}{(\mathrm{d}R)^2 + (\mathrm{d}f)^2 (R_m^0)^2} \right]^{-1} = \frac{1}{2}(R_i^{(k,m)} + R_j^{(l,m)})$$

$\Rightarrow$ Fixing covariance matrix ensures unbiased solution!

Model data: a bias free example

$$R_i^{(k,m)} = 0.9 \text{ and } R_j^{(l,m)} = 1.1 \quad ; \quad \mathrm{d}R_i^{(k,m)} = \mathrm{d}R_j^{(l,m)} \equiv \mathrm{d}R = 0$$



An iterated fit

Iterating the fit ensures an unbiased solution!

$\rightarrow$ Forces the fit to converge to an unbiased result.

$$\chi_1^2(R_m) = \sum_{m=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \sum_{j=1}^{N^{(k,n)}} \left(R_i^{(k,m)} - R_m\right) C_{k,0}^{-1}(m_i, n_j) \left(R_j^{(k,n)} - R_n\right)$$

$$C_{k,0}(m_i, n_j) = \mathbf{c}_k(m_i, n_j) + (\mathrm{d}f_k)^2 R_m^0 R_n^0$$

$\Rightarrow$ Feed the fitted R_m values into the next iteration...

$$\chi_2^2(R_m) = \sum_{m=1}^{N_{clu}} \sum_{i=1}^{N^{(k,m)}} \sum_{j=1}^{N^{(k,n)}} \left(R_i^{(k,m)} - R_m\right) C_{k,1}^{-1}(m_i, n_j) \left(R_j^{(k,n)} - R_n\right)$$

$$C_{k,1}(m_i, n_j) = \mathbf{c}_k(m_i, n_j) + (\mathrm{d}f_k)^2 R_m^1 R_n^1$$

Repeat until fit converges and returns final fitted values for clusters, $R_m = R_m^I$.

The $\pi^+\pi^-$ channel

$\pi^+\pi^-$ channel contributes to over 70% of $a_\mu^{\text{had, LOVP}}$

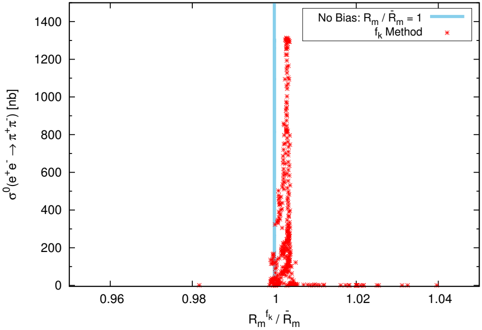
- Consider real cluster with one measurement:
  $E_{cm}(\text{GeV}) = 0.4600$, $\sigma^0(\text{nb}) = 123.6418 \pm 22.4855 \quad \Rightarrow \quad R_m^0 = 123.6418$
- Experimental data includes three more measurements in different clusters.
  $\rightarrow$ Normalisation uncertainty provides weighting to cluster value through correlations.

Using R_m^I method,

$$R_m^I = 130.3513 \ ,$$

whereas, the f_k method returns

$$R_m^{f_k} = 130.8127 \ .$$



Did the f_k method incur a bias?

Are previous results still reliable?

Compare f_k method and R_m^I method with only multiplicative normalisation uncertainties.

$\rightarrow$ If we see **differences** in mean value, then bias previously influenced the fit.
$\longrightarrow$ Previous results unreliable!

$\rightarrow$ If we see **no differences** in mean value, then bias did not influence fit (any change come from the inclusion of extra systematics).
$\longrightarrow$ Previous results reliable!

Example - $\pi^+\pi^-$
Set 1 - CMD-2(06) (0.7% Systematic Uncertainty), Set 2 - CMD-2(06) (0.8% Systematic Uncertainty), Set 3 - SND(04) (1.3% Systematic Uncertainty)

From $0.37 \rightarrow 0.97$ GeV

Fit Method:	f_k Method		R_m^I Method		
Channel	a_μ	χ^2_{min}/d.o.f.	a_μ	χ^2_{min}/d.o.f.	Difference
$\pi^+\pi^-$	481.42 ± 4.26	1.10	481.43 ± 4.04	1.02	$+0.01$

Results for dominant hadronic channels

Comparative results for the f_k method and the R^I_m method:

Fit Method:	f_k Method		R^I_m Method		
Channel	a_μ	χ^2_{min}/d.o.f.	a_μ	χ^2_{min}/d.o.f.	Difference
$\pi^+\pi^-$	505.77 ± 3.09	1.39	504.42 ± 2.24	1.35	-1.35
$\pi^+\pi^-\pi^0$	47.51 ± 0.98	3.04	47.47 ± 0.91	3.33	-0.04
$\pi^+\pi^-\pi^0\pi^0$	20.73 ± 1.28	1.29	20.40 ± 1.16	1.16	-0.33
$\pi^+\pi^-\pi^+\pi^-$	14.73 ± 0.48	1.81	14.49 ± 0.48	1.72	-0.24
K^+K^-	22.12 ± 0.41	1.95	22.08 ± 0.42	1.71	-0.04
$K^0_s K^0_l$	13.46 ± 0.17	1.10	13.46 ± 0.17	0.93	0.00
$\pi^+\pi^-\pi^+\pi^-\pi^0$	1.42 ± 0.09	1.21	1.38 ± 0.08	1.09	-0.04
$\pi^+\pi^-\pi^+\pi^-\pi^+\pi^-$	0.30 ± 0.01	1.67	0.30 ± 0.01	1.52	0.00
Total:	626.04 ± 3.55		624.00 ± 2.76		-2.04

- Changes in almost each channel due to inclusion of extra systematics.
- Improved error estimate and goodness-of-fit.
- Much more statistically reliable and trustworthy method.
- Reduction in overall mean value would mean increased $g-2$ discrepancy, Δa_μ.

Conclusions

- Hadronic sector provides biggest uncertainty to $\Delta a_\mu = a_\mu^{\mathrm{EXP}} - a_\mu^{\mathrm{SM}} \rightarrow$ need statistically trustworthy method.
- Potential bias can occur through the fitting of experimental data due to global systematic normalisation uncertainties.
- Specifically, allowing normalisation uncertainties to vary as parameters in a fit can incur bias.
- Redefine our fit function so covariance matrices are fixed. Normalisation uncertainties are then not free to vary in the fit and thus cause a bias.
- Iterative fit procedure ensures convergence to correct solution.
- R_m^I method provides us with a trustworthy and much improved fitting method **BUT** previous results still reliable.
- Results show reduced mean value with improved uncertainty.

If Δa_μ is larger than originally thought, where is new physics...?

Thank You

a.i.keshavarzi@liverpool.ac.uk
(alexkesh@liv.ac.uk)

https://www.linkedin.com/in/alexkesh