Combination of highly correlated measurements of the muon precession frequency in magnetic field for the FNAL measurement of the muon magnetic anomaly

Alberto Lusiani
for the FNAL Muon $g-2$ collaboration
https://orcid.org/0000-0002-6876-3288

FNAL-E989 measurement of muon magnetic anomaly

muon magnetic anomaly

$$a_\mu = \left[\frac{\omega_a}{\tilde{\omega}'_p(T)}\right] \cdot \left[\frac{\mu'_p(T)}{\mu_e(\mathrm{H})}\right] \left[\frac{\mu_e(H)}{\mu_e}\right] \left[\frac{m_\mu}{m_e}\right] \left[\frac{g_e}{2}\right]$$

- ω_a: muon precession frequency
- $\tilde{\omega}'_p(T)$: proton spin precession in magnetic field (NMR)
- The other factors: precise external inputs

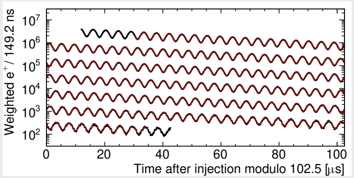
$$\frac{\omega_a}{\tilde{\omega}'_p(T)} = \frac{\omega_a^m}{\tilde{\omega}'_p(T)^m} \frac{1 + C_e + C_p + C_{\mathrm{ml}} + C_{\mathrm{pa}} + C_{\mathrm{dd}}}{1 + B_k + B_q}$$

- ω_a^m, $\tilde{\omega}'_p(T)^m$: measured frequencies
- Numerator terms: ω_a corrections
- Denominator terms: $\tilde{\omega}'_p(T)$ corrections

C_e	electric field correction
C_p	pitch correction
C_{ml}	muon loss correction
C_{pa}	phase acceptance correction
C_{dd}	differential decay correction
B_k	kickers' eddy currents correction
B_q	quadrupoles' vibration correction

- statistical uncertainty on a_μ primarily determined by ω_a^m, 0.43 ppm in Run 1, 0.20 ppm in Run 2+3
- systematic uncertainties from ω_a^m, $\tilde{\omega}'_p(T)^m$, C_i, B_i, external inputs

ω_a FNAL-E989 Run 2+3 measurements



- ω_a measured with wiggle plot fit by 7 groups with 19 analyses

datasets

Run	d	dataset
Run2	Run2	Run 2
Run3	Run3a	Run 3a
Run3	Run3b	Run 3b

measurement methods

m	ω_a measurement method
T	Threshold
A	Asymmetry weighted
RT	Ratio Threshold
RA	Ratio Asymmetry weighted
Q	Charge
RQ	Ratio Charge

analysis groups

g	group
B	Boston
C	Cornell
E	Europa
I	IRMA
K	Kentuky
S	Shanghai
W	Washington

reconstructions

r	reconstruction method
RW	ReconWest
RW2	ReconWest v2
RE	ReconEast
RI1	ReconITA v1
RQ	Q method recostruction

analyses

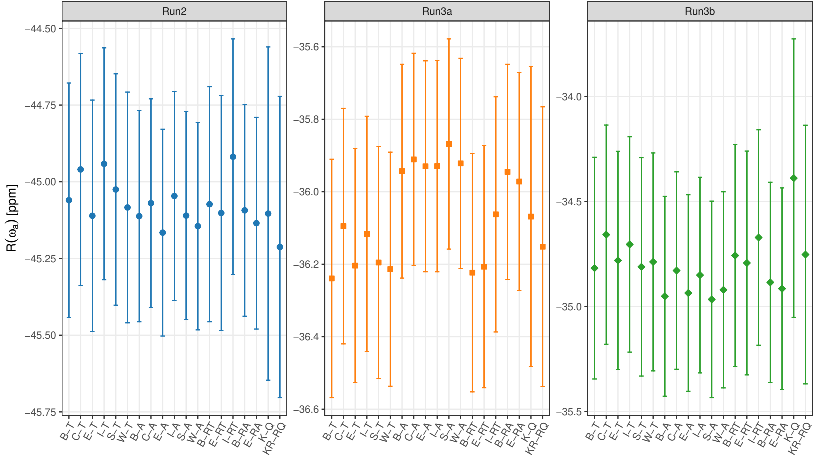
g	r	m
B	RW2	T, A, RT, RA
C	RE	T, A
E	RI1	T, A, RT, RA
I	RE	T, A, RT
K	RQ	Q, RQ
S	RW	T, A
W	RW2	T, A

ratio randomization methods

rr	ratio randomization methods
R	random
K	kernel

ω_a measurements consistency checks

Run 2,3 datasets measurements, $R(\omega_a)$ = ppm deviation from reference, blind



variance / uncertainty of difference between two analyses (residual)

- $\sigma^2\left(R_{\text{analysis } j} - R_{\text{analysis } i}\right) = \sigma^2_{j-i} = \sigma^2_i + \sigma^2_j - 2\rho\sigma_i\sigma_j$ with $\rho \lesssim 1$, high correlation

Estimating uncertainty on residual σ_{j-i}

residual uncertainty between different analysis methods A, T, Q

- different event weighting: A = asymmetry, T = threshold, Q = charge
- A method known to be optimal, T & Q performed for checks and possible different systematics
- T, Q measure same data as A with additional uncorrelated uncertainty: $\sigma_T^2 = \sigma_A^2 + \sigma^2_{\text{extra T vs. A}}$
  $\Rightarrow \quad \sigma^2_{T-A} = \sigma^2_T - \sigma^2_A, \quad \sigma^2_{Q-A} = \sigma^2_Q - \sigma^2_A$
- for Run 2,3 data checks we found acceptable to also estimate $\sigma^2_{Q-T} = \sigma^2_Q - \sigma^2_T$
- for Run 2,3 data, σ_T is ∼10% larger than σ_A, σ_Q is ∼40% larger than σ_A,

residual uncertainty for analyses that differ by one feature

- e.g., analysis i uses RW reconstruction, j uses RE reconstruction
- we found: $\sigma_{\text{analysis using RW}} \approx \sigma_{\text{analysis using RE}}$ when all other features are the same
- model:
  - $\sigma^2_{\text{analysis using RW}} = \sigma^2_0 + \sigma^2_{RW}$
  - $\sigma^2_{\text{analysis using RE}} = \sigma^2_0 + \sigma^2_{RE}$

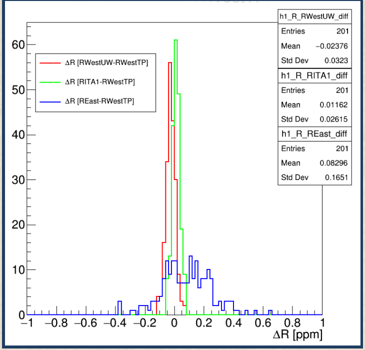
  $\sigma_{RW} = \sigma_{RE} = \sigma_{RERW}$, with σ_{RW}, σ_{RE} uncorrelated
- with this model, for two analysis differing only for reconstruction, $\sigma^2_{RE-RW} = \sigma^2_{RW} + \sigma^2_{RE} = 2\sigma^2_{RERW}$

Estimating residual uncertainties on data with bootstrap method

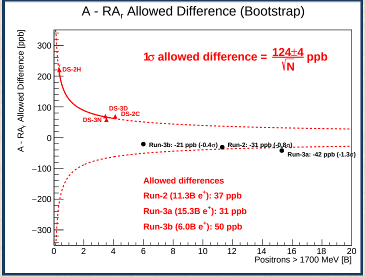
bootstrap method

- divide ω_a data sample in $\sim$200 equal-size subsamples
- assemble $\sim$200 full-size "bootstrap" data samples by random extraction, with repetition
- compute $\sigma^2_{j-i} = \left\langle \left(\omega^j_a - \omega^i_a\right)^2 \right\rangle$ on bootstrap samples

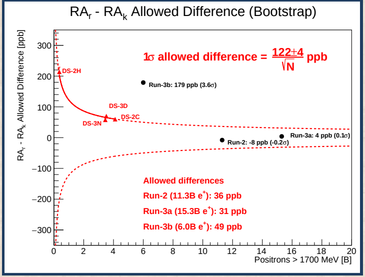
different reconstructions



ratio vs. non-ratio



ratio-kernel vs. ratio-random



Modeling residual uncertainties

findings from bootstrap studies

- confirmed $\sigma^2_{T-A} = \sigma^2_T - \sigma^2_A$ for T vs. A method measurements
- for measurements i, j with same σ_{ij}, $\sigma_{T-A} = \text{constant} \cdot \sigma_{ij}$ (scales with σ_{ij})

  $\Rightarrow$ model residual uncertainties with $\sigma_{r,\, j-i} = \sigma_{j-i}/\sigma_{ij}$, relative uncorrelated uncertainty

further modeling assumptions

- when two analyses differ for >1 features of different categories, sum related variances

  $$\sigma^2_{r,\,\text{reco}=RW,\text{ratio}-\text{reco}=RE,\text{non-ratio}} = \sigma^2_{r,\,\text{reco}=RW-\text{reco}=RE} + \sigma^2_{r,\,\text{ratio}-\text{non-ratio}}$$

- when two analyses total uncertainties are different, multiply relative variance σ^2_r by mean variance

  $$\sigma^2_{A,\text{reco}=RW-T,\text{reco}=RE} = \left(\sigma^2_{T-A}\right) + \left(\sigma^2_{r,\,\text{reco}=RW-\text{reco}=RE} \cdot \frac{\sigma^2_A + \sigma^2_T}{2}\right)$$

  consistent with model that each measurement has one variance contribution per feature
  and this variance is product of relative variance σ^2_r and total variance σ^2

Bootstrap studies results expressed as relative uncorrelated uncertainties

A, RA methods

	RW	RW2	RE
RW2	0.03788		
RE	0.08334	0.07450	
RI1	0.03157	0.02020	0.07955

T, RT methods

	RW	RW2	RE
RW2	0.03632		
RE	0.18728	0.18047	
RI1	0.02951	0.02270	0.18160

A method

	non-ratio	ratio-random
ratio-random	0.07730	
ratio-kernel	0.09289	0.07606

T method

	non-ratio	ratio-random
ratio-random	0.06421	
ratio-kernel	0.08393	0.07266

Modeled covariance matrix between two ω_a measurents

$$V_{\omega_a^A,\omega_a^B} = \begin{pmatrix} \sigma_A^2 & \frac{\sigma_A^2+\sigma_B^2}{2} - \frac{\left|\sigma_B^2-\sigma_A^2\right|}{2} - \left(\sum_i \sigma_{r,i}^2\right)\cdot\frac{\sigma_A^2+\sigma_B^2}{2} \\ \frac{\sigma_A^2+\sigma_B^2}{2} - \frac{\left|\sigma_B^2-\sigma_A^2\right|}{2} - \left(\sum_i \sigma_{r,i}^2\right)\cdot\frac{\sigma_A^2+\sigma_B^2}{2} & \sigma_B^2 \end{pmatrix}$$

- using above covariance, uncertainty on residual between measurements
  - T, ratio, reconstruction RW
  - A, non-ratio, reconstruction RE

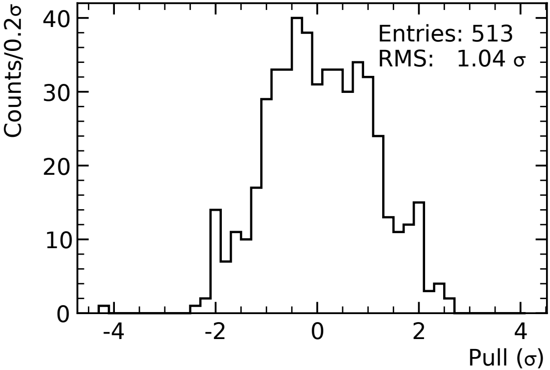
would be: $\sigma_{j-i} = \sigma_T^2 - \sigma_A^2 + \sigma^2_{r,\text{ RW vs. RE}}\dfrac{\sigma_A^2+\sigma_T^2}{2} + \sigma^2_{r,\text{ ratio vs. non-ratio}}\dfrac{\sigma_A^2+\sigma_T^2}{2}$

critical correlation

- by construction, for above covariance, correlation $\rho_{A,B} \leq \sigma_{\text{min}}/\sigma_{\text{max}}$ (critical correlation)
  - discussed later

Residuals' pulls, all comparisons between analyses in same dataset

Run 2,3 datasets



- added irreducible systematic uncertainty of 28 ppb to all residuals
  (otherwise analyses by different groups but same features have expected residual uncertainty = 0)
- 1 outlier, one Q vs. RQ comparison – no Q-RQ bootstrap study, used findings for T-RT
  - Q measurements not used for a_μ measurements, used just for checks

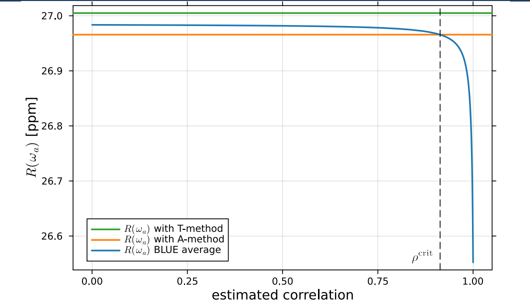
Critical correlation

- as mentioned earlier, estimated covariancs for checks, by construction, has correlation $\rho_{A,B} \leq \sigma_{\text{min}}/\sigma_{\text{max}}$ (critical correlation)
- this condition alone does not insure that whole correlation matrix may be nevertheless "super-critical" i.e., when used to average ω_a measurements with BLUE / minimum χ^2, results in negative weights; literature on the topic:
  - for averages of two measurements, statistics books e.g. G. Cowan, Statistical data analysis, 1998
  - for averages of many measurements A.Valassi & R.Chierici, EPJC 74 (2014) 2717

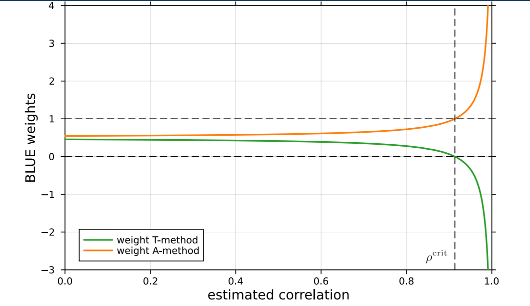
BLUE average vs. assumed correlation

- average of 1 A and 1 T method measurements
- T uncertainty 10% larger than A
- true correlation $\rho = \sigma_A/\sigma_T$, $\Rightarrow$ w_A=1, w_T=0
- unstable outcomes for assumed $\rho > \sigma_A/\sigma_T$

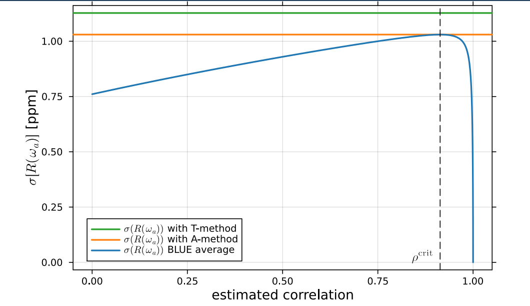
BLUE average value vs. estimated correlation



BLUE average weights

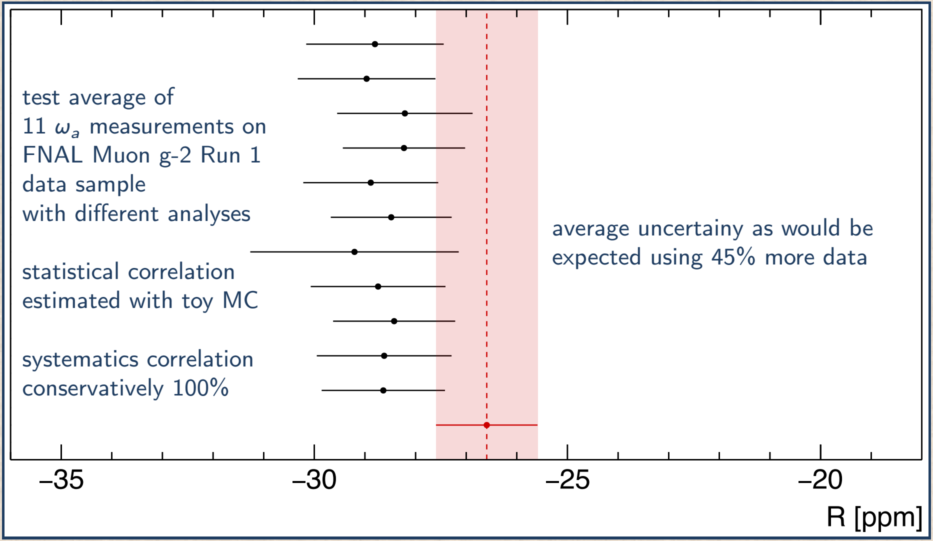


BLUE average uncertainty vs. est...d correlation



Averaging highly correlated measurements with imperfect correlation

- minimum χ^2 average of correlated measurement using imprecise correlation can be "'un-physical"
  - especially when conservatively estimating fully correlated systematic uncertainties



FNAL Muon $g-2$ "safe" averaging of highly correlated ω_a measurements

BLUE / minimum χ^2 averaging weights and uncertainty

- $w_i = (A^t V^{-1} A)^{-1} A^t V^{-1}$, $\sigma^2_{\text{av}} = w_i^t V w_i$, with V covariance, A model matrix (column of 1's)

safe averaging weights and uncertainty (used for FNAL Muon $g-2$ Run 1, and Run 2+3)

- only use most precise ω_a analyses (A, RA), with in principle exactly the same uncertainty
- $w_i =$ series of equal weigths $\dfrac{1}{N_{\text{measurements}}}$, $\sigma^2_{\text{av}} = w_i^t V w_i$
- conservative covariance $V_{ij} = \sigma_i \sigma_j$
  - (100% correlation, supercritical if $\sigma_i \neq \sigma_j$, critical if $\sigma_i = \sigma_j$), leading to $\sigma_{\text{av}} = \langle \sigma_i \rangle$
- if $V_{ij} = \sigma_i \sigma_j$ (100% correlation, supercritical if $\sigma_i \neq \sigma_j$, critical if $\sigma_i = \sigma_j$), then $\sigma_{\text{av}} = \langle \sigma_i \rangle$

sub-optimality of safe averaging used in FNAL Muon $g-2$

covariance	weights	uncertainty scale
conservative	equal weights	100.0%
modeled for consistency checks	equal weights	∼99.5%
modeled for consistency checks	optimal (BLUE) weights	∼98.5%

- small degradation of uncertainty for using sub-optimal but safe weights
- at least one negative optimal weight on all 3 FNAL Muon $g-2$ Run 2,3 datasets

End

Backup Slides

Calculation of the muon magnetic anomaly

$$a_\mu = \left[\frac{\omega_a}{\tilde{\omega}'_p(T)}\right] \cdot \left[\frac{\mu'_p(T)}{\mu_e(\mathrm{H})}\right] \left[\frac{\mu_e(H)}{\mu_e}\right] \left[\frac{m_\mu}{m_e}\right] \left[\frac{g_e}{2}\right]$$

measurements by the Muon $g-2$ collaboration

- ω_a — precession of muon spin relative to momentum rotation in magnetic field
- $\tilde{\omega}'_p(T)$ — precession frequency of shielded proton spin in spherical water sample at $T = 34.7\,°\mathrm{C}$ in muon-beam-weighted magnetic field, $\tilde{\omega}'_p(T) = \langle \omega'_p(T)(x, y, \varphi) \times M(x, y, \varphi) \rangle$

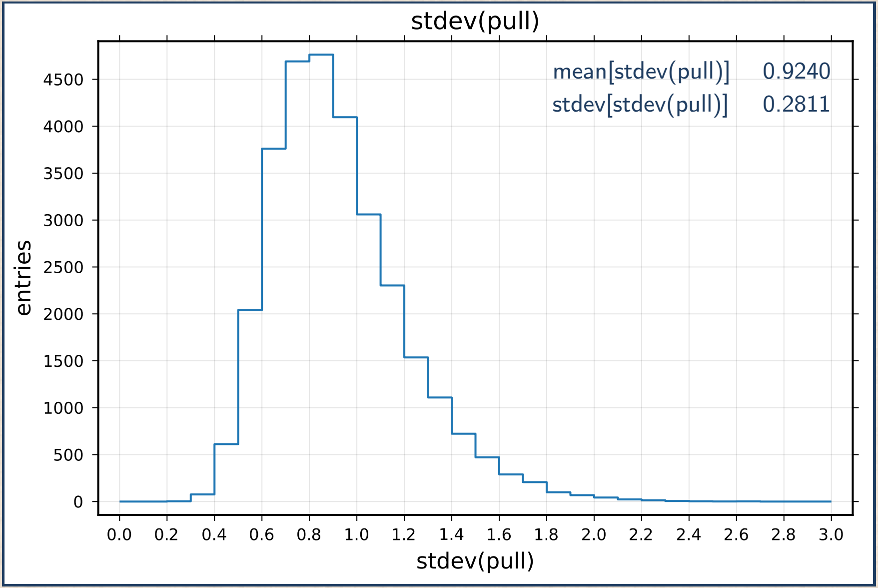
notation

- $\mu'_p(T)$ — magnetic moment of proton in spherical water sample at 34.7 °C
- $\mu_e(\mathrm{H})$ — magnetic moment of electron in hydrogen atom

external measurements

- $\mu'_p(T)/\mu_e(H)$ — 10.5 ppb precision, Metrologia 13, 179 (1977)
- $\mu_e(H)/\mu_e$ — 100 ppt (for n. digits) theory QED calculation, Rev. Mod. Phys. 88 035009 (2016)
- m_μ/m_e — 22 ppb precision — CODATA 2018 fit, primarily driven by LAMPF 1999 measurements of muonium hyperfine splitting, Phys. Rev. Lett. 82, 711 (1999)
- $g_e/2$ — 0.12 ppt, PDG 2023

Toy MC expected distribution of pulls standard deviation



Modeled correlation between different ω_a measurements

	C_T	E_T	I_T	S_T	W_T	B_A	C_A	E_A	I_A	S_A	W_A	B_RT	E_RT	I_RT	B_RA	E_RA	K_Q	K_RQ
B_T	0.967	0.999	0.967	0.999	1.000	0.900	0.871	0.884	0.867	0.884	0.884	0.993	0.995	0.963	0.895	0.904	0.765	0.824
C_T		0.967	1.000	0.965	0.967	0.891	0.900	0.875	0.896	0.874	0.875	0.961	0.963	0.996	0.887	0.895	0.756	0.815
E_T			0.967	0.999	0.999	0.913	0.885	0.898	0.880	0.898	0.897	0.993	0.996	0.963	0.909	0.918	0.753	0.811
I_T				0.965	0.967	0.897	0.906	0.881	0.902	0.880	0.880	0.961	0.963	0.996	0.892	0.901	0.751	0.809
S_T					0.999	0.915	0.886	0.900	0.882	0.902	0.899	0.992	0.995	0.961	0.911	0.920	0.751	0.809
W_T						0.915	0.887	0.899	0.882	0.899	0.899	0.993	0.995	0.963	0.911	0.919	0.752	0.810
B_A							0.994	1.000	0.994	0.999	1.000	0.890	0.886	0.887	0.991	0.994	0.688	0.740
C_A								0.994	1.000	0.993	0.994	0.862	0.857	0.896	0.986	0.988	0.681	0.732
E_A									0.994	0.999	1.000	0.875	0.871	0.871	0.991	0.994	0.676	0.727
I_A										0.993	0.994	0.858	0.853	0.892	0.986	0.988	0.678	0.729
S_A											0.999	0.875	0.871	0.870	0.990	0.993	0.677	0.728
W_A												0.875	0.870	0.871	0.991	0.994	0.676	0.727
B_RT													0.994	0.962	0.902	0.907	0.758	0.825
E_RT														0.967	0.895	0.901	0.767	0.837
I_RT															0.895	0.901	0.750	0.819
B_RA																0.994	0.682	0.743
E_RA																	0.689	0.754
K_Q																		0.994