Set-subset problem in data analysis for the muon (g–2) experiment at BNL

S. I. Redin [a,b,1] et al.

[a] *Budker Institute of Nuclear Physics, Novosibirsk, Russia*

[b] *Physics Department, Yale University, New Haven, CT 06511, USA*

Abstract

For the muon (g–2) experiment at BNL we use number of systematic checks based on the set-subset comparison. These are, noteably, checks for stability of (g–2) frequency ω versus fit start time (1) and versus energy threshold (2). For the first case it was shown, both explicitly for (g–2) and in general, that expected "random walk" for ω is $\pm\sqrt{\sigma_{\omega 2}^2 - \sigma_{\omega 1}^2}$, where $\sigma_{\omega 1}$ and $\sigma_{\omega 2}$ are statistical errors of ω obtained from chi-squared fit over full set and subset of data, respectively. For the second case some parameters of generic distribution are different for set and subset and thus the theorem is not applicable. Corresponding equation for "random walk" for (g–2) frequency ω for this case was derived explicitly.

Key words: Set-subset problem, Set-subset theorem

1 Introduction

Suppose we have some distribution, say time spectrum, which can be described unambiguously by $L-$parameter function $f(x_1, x_2, ..., x_L; t) \equiv f(\boldsymbol{x}; t)$. Suppose we have some experimental histogram which corresponds to this distribution for some fixed set of parameters $(x_{1\circ}, x_{2\circ}, ..., x_{L\circ}) \equiv \boldsymbol{x}_\circ$. Due to finite statistics, number of events in n^{th} channel of the histogram $\mathcal{N}_n$ differs from "predicted" value $f(\boldsymbol{x}_\circ; t_n)$, which is value of function $f(\boldsymbol{x}; t)$ for fixed parameters $\boldsymbol{x}_\circ$ at the center of n^{th} channel, t_n. For large $\mathcal{N}_n$ mean square of this difference σ_n^2, defined as an average over ensemble of similar histograms (measurements), is equial to

$$\sigma_n^2 \equiv \left\langle \left(f(\boldsymbol{x}_\circ; t_n) - N_n\right)^2 \right\rangle = f(\boldsymbol{x}_\circ; t_n) \tag{1}$$

[1] Corresponding author. E-mail sredin@sun2.bnl.gov

where $\langle ... \rangle$ means ensemble average. Suppose these "true" parameters $\boldsymbol{x}_\circ$ of function $f(\boldsymbol{x};t)$ are not known and we want to find them by the standard least chi squared minimisation fit of the histogram with function $f(\boldsymbol{x};t)$. Due to statistical fluctuations σ_n of number of events in channels of histogram, fitted parameters of $f(\boldsymbol{x};t)$ differs from $\boldsymbol{x}_\circ$ by some amounts $\delta\boldsymbol{x} \equiv (\delta x_1, \delta x_2, ..., \delta x_L)$. We can find these δx_i as functions of σ_n by use of definition of minimun of chi squared : $\partial W/\partial x_i = 0$, where we use notation of [1] $W \equiv \chi^2$.

$$W(\boldsymbol{x}_\circ + \delta\boldsymbol{x}) = \sum_n \frac{[f(\boldsymbol{x}_\circ + \delta\boldsymbol{x}; t_n) - \mathcal{N}_n]^2}{\sigma_n^2} \approx \sum_n \frac{(f_\circ + \sum_i f_i'\,\delta x_i - \mathcal{N}_n)^2}{\sigma_n^2} \tag{2}$$

Here $f_i' = (\partial f/\partial x_i)_{\boldsymbol{x}=\boldsymbol{x}_\circ}$ and $f_\circ = f(\boldsymbol{x}_\circ;t)$. For σ_i^2 we shall use approximation $\sigma_i^2 = f \approx f_\circ$. Then from $\partial W/\partial x_i = 0$ we have system of L linear equations for δx_i :

$$\sum_n f_i' \, \frac{f_\circ + \left(\sum_j f_j' \, \delta x_j\right) - \mathcal{N}_n}{f_\circ} = 0\,, \qquad i = 1, 2, ..., L \tag{3}$$

which can be written in matrix form as

$$\mathcal{A} \times \delta\boldsymbol{x} = \boldsymbol{b} \tag{4}$$

with matix elements a_{ij} of $\mathcal{A}$ and components b_i of vector $\boldsymbol{b}$ being

$$a_{ij} = \sum_n \frac{f_i' \, f_j'}{f_\circ} \qquad \text{and} \qquad b_i = \sum_n \frac{f_i'}{f_\circ}\,(\mathcal{N}_n - f_\circ) \tag{5}$$

respectively. For a histogram with many enough channels and small enough bin width b, sum can be replaced by integral: $\sum_i \to \int dt/b$. Matrix $\mathcal{A}$ is nothing else but the error matrix which we discuss in our previous work [1]. According to the Cramer's rule, solution for eq.5 is

$$\delta x_j = \frac{A_{1j}b_1 + A_{2j}b_2 + ... + A_{Lj}b_L}{det(\mathcal{A})} = \frac{\sum_{i=1}^{L} A_{ij} b_i}{det(\mathcal{A})}, \qquad A_{ij} = (-1)^{i+j} M_{ij} \tag{6}$$

where A_{ij} are cofactors of elements a_{ij} of matrix $\mathcal{A}$ and M_{ij} are its minors.

Eqs.4-6 give statistical fluctuations of fit parameters as explicit functions of statistical fluctuations of number of events in all individual channels of given histogram. These equations are particular useful for ensemble averaging, where we can use fundamental relations :

$$\langle\,(f(\boldsymbol{x};t_n)-\mathcal{N}_n)\,\rangle \equiv 0 \tag{7}$$

$$\begin{aligned}
&\langle\,(f(\boldsymbol{x};t_n)-\mathcal{N}_n)\,(f(\boldsymbol{x};t_k)-\mathcal{N}_k)\,\rangle = 0 && \text{for}\quad n\neq k \qquad \text{and}\\
&\langle\,(f(\boldsymbol{x};t_n)-\mathcal{N}_n)\,(f(\boldsymbol{x};t_k)-\mathcal{N}_k)\,\rangle \equiv \sigma_n^2 = f(\boldsymbol{x};t_n) && \text{for}\quad n = k
\end{aligned} \tag{8}$$

As an example of use of eqs.7,8 we can find statistical error of parameter x_j as ensemble r.m.s. value of $(\delta x_j)^2$:

$$\begin{aligned}
\sigma_j^2 \equiv \left\langle(\delta x_j)^2\right\rangle &= \frac{1}{det^2(\mathcal{A})}\ \left\langle\sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij}\,b_i\,A_{kj}\,b_k\ \right\rangle = \\
&= \frac{1}{det^2(\mathcal{A})}\ \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij}\,A_{kj}\ \left\langle\ \sum_{n}\frac{f'_i}{f_\circ}(\mathcal{N}_n-f_\circ)\sum_{m}\frac{f'_k}{f_\circ}(\mathcal{N}_m-f_\circ)\ \right\rangle = \\
&= \frac{1}{det^2(\mathcal{A})}\ \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij}\,A_{kj}\ \sum_{n}\frac{f'_i\,f'_k}{f_\circ^2}\left\langle\,(\mathcal{N}_n-f_\circ)^2\right\rangle = \\
&= \frac{1}{det^2(\mathcal{A})}\ \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij}\,A_{kj}\ \sum_{n}\frac{f'_i\,f'_k}{f_\circ} = \frac{1}{det^2(\mathcal{A})}\ \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij}\,A_{kj}\,a_{ki}
\end{aligned} \tag{9}$$

Now recall that $det(\mathcal{A}) = \sum_{k=1}^{L} A_{kj}\,a_{kj}$ and hence expression $\sum_{k=1}^{L} A_{kj}\,a_{ki}$ is determinant of a matrix constructed from maxtrix $\mathcal{A}$ by replacing j^{th} row of $\mathcal{A}$ by i^{th} row of the very same matrix $\mathcal{A}$. Obviously, such a determinant vanishes unless $i=j$. Thus

$$\sigma_j^2 = \frac{1}{det^2(\mathcal{A})}\ A_{jj}\ det(\mathcal{A}) = \frac{M_{jj}}{det(\mathcal{A})} \tag{10}$$

which is same as in our previous work [1], obtained by somewhat different approach.

Another useful value to calculate is coefficient of statistical correlations between some pare of fit parameters (x_p,x_q), ρ_{pq}, defined as

$$\rho_{pq} \equiv \frac{\langle\delta x_p\,\delta x_q\rangle}{\sigma_p\,\sigma_q} \tag{11}$$

Calculation of $\langle\delta x_p\,\delta x_q\rangle$ is similar to that of $\langle\,(\delta x_j)^2\rangle$ in eq.9 :

$$\begin{aligned}
\langle\delta x_p\,\delta x_q\rangle &= \frac{1}{det^2(\mathcal{A})}\ \left\langle\sum_{i=1}^{L}\sum_{k=1}^{L} A_{ip}\,b_i\,A_{kq}\,b_k\ \right\rangle = \\
&= \frac{1}{det^2(\mathcal{A})}\ \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ip}\,A_{kq}\ \left\langle\ \sum_{n}\frac{f'_i}{f_\circ}(\mathcal{N}_n-f_\circ)\sum_{m}\frac{f'_k}{f_\circ}(\mathcal{N}_m-f_\circ)\ \right\rangle =
\end{aligned}$$

$$= \frac{1}{det^2(\mathcal{A})} \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ip}\, A_{kq} \sum_n \frac{f_i' \, f_k'}{f_\circ^2} \left\langle (\mathcal{N}_n - f_\circ)^2 \right\rangle =$$
$$= \frac{1}{det^2(\mathcal{A})} \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ip}\, A_{kq} \sum_n \frac{f_i' \, f_k'}{f_\circ} = \frac{1}{det^2(\mathcal{A})} \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ip}\, A_{kq}\, a_{ki} =$$
$$= \frac{1}{det^2(\mathcal{A})}\, A_{qp} \times det(\mathcal{A}) = \frac{A_{pq}}{det(\mathcal{A})} \tag{12}$$

and hence

$$\rho_{pq} = \frac{A_{pq}}{det(\mathcal{A})} \sqrt{\frac{det(\mathcal{A})}{M_{pp}}} \sqrt{\frac{det(\mathcal{A})}{M_{qq}}} = \frac{A_{pq}}{\sqrt{M_{pp}\, M_{qq}}} = (-1)^{p+q} \frac{M_{pq}}{\sqrt{M_{pp}\, M_{qq}}} \tag{13}$$

In the next section we consider some applications of derived equations for the muon (g–2) experiment.

2 Muon (g–2) distribution

In our previous works [1] and [2] we discuss statistical errors and correlations of parameters and systematic errors associated with small unidentified background for time spectrum (distribution) of decay electrons in the muon (g–2) experiment [3]. This distribution can be approximated by 5-parameter function

$$G(t) = N_\circ \, e^{-t/\tau} \, \left[1 + A \, \cos(\omega t + \phi)\right] \tag{14}$$

In [1] it was shown that error matrix $\mathcal{A}$ for distribution (14) can be made diagonal if time origin $t = 0$ is chosen to be at center of gravity of this distribution, i.e. at one life time τ after the histogram start time. In this case eq.4 can be written explicitely as

$$\begin{pmatrix} \frac{N}{N_\circ^2} & 0 & 0 & 0 & 0 \\ 0 & \frac{N}{\tau^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{N}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{N\, A^2 \tau^2}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{N\, A^2}{2} \end{pmatrix} \times \begin{pmatrix} \delta N_\circ \\ \delta\tau \\ \delta A \\ \delta\omega \\ \delta\phi \end{pmatrix} = \begin{pmatrix} \sum_n \frac{\partial G}{\partial N_\circ} \frac{\mathcal{N}_n - G}{G} \\ \sum_n \frac{\partial G}{\partial \tau} \frac{\mathcal{N}_n - G}{G} \\ \sum_n \frac{\partial G}{\partial A} \frac{\mathcal{N}_n - G}{G} \\ \sum_n \frac{\partial G}{\partial \omega} \frac{\mathcal{N}_n - G}{G} \\ \sum_n \frac{\partial G}{\partial \phi} \frac{\mathcal{N}_n - G}{G} \end{pmatrix} \tag{15}$$

Solution for statistical fluctuation of (g–2) frequency ω, which is of main concern in the muon (g–2) experiment, is

$$\delta\omega = \frac{2}{N\,A^2\,\tau^2}\sum_n \frac{\partial G}{\partial\omega}\,\frac{\mathcal{N}_n - G}{G} \tag{16}$$

As an useful for further applications exercise, we find ensemble average for $(\delta\omega)^2$, which is by definition equal to σ_ω^2 :

$$\begin{aligned}
\langle(\delta\omega)^2\rangle &= \left(\frac{2}{N\,A^2\tau^2}\right)^2 \left\langle \sum_n \left(\frac{\partial G}{\partial\omega}\,\frac{\mathcal{N}_n - G}{G}\right) \sum_k \left(\frac{\partial G}{\partial\omega}\,\frac{\mathcal{N}_k - G}{G}\right) \right\rangle = \\
&= \left(\frac{2}{N\,A^2\tau^2}\right)^2 \sum_n \sum_k \left(\frac{\partial G}{\partial\omega}\,\frac{1}{G}\right)_n \left(\frac{\partial G}{\partial\omega}\,\frac{1}{G}\right)_k \langle\, (\mathcal{N}_n - G_n)(\mathcal{N}_k - G_k)\,\rangle = \\
&= \left(\frac{2}{N\,A^2\tau^2}\right)^2 \sum_n \left(\frac{\partial G}{\partial\omega}\,\frac{1}{G}\right)^2 G = \left(\frac{2}{N\,A^2\tau^2}\right)^2 \sum_n \left(\frac{\partial G}{\partial\omega}\right)^2 \frac{1}{G}
\end{aligned} \tag{17}$$

For small enough bin width b, sum over histogram channels $\sum_n$ can be replaced by integral $\int dt/b$:

$$\sum_n \left(\frac{\partial G}{\partial\omega}\right)^2 \frac{1}{G} \approx \frac{1}{b}\int_{-\tau}^{\infty} \left(\frac{\partial G}{\partial\omega}\right)^2 \frac{1}{G}\,dt = \frac{1}{b}\,\frac{e\,N_\circ A^2\tau^3}{2} = \frac{N\,A^2\tau^2}{2} \tag{18}$$

In the last step of eq.18 we replace parameter b by total number of events N through the normalization relation:

$$N = \sum_n \mathcal{N}_n \approx \sum_n G(t_n) \approx \frac{1}{b}\int_{-\tau}^{\infty} G(t)\,dt\ ,\quad \text{hence}\ \frac{1}{b} = \frac{N}{\int G(t)\ dt} = \frac{N}{e\,N_\circ\tau} \tag{19}$$

Evaluation of integrals in eqs.18,19 can be found in Appendix II of [1].

Finally, from eqs.17,18 we have

$$\sigma_\omega \equiv \sqrt{\langle(\delta\omega)^2\rangle} = \frac{\sqrt{2}}{A\,\tau\,\sqrt{N}} \tag{20}$$

which is same as in our previous work [1], obtained by somewhat different approach. For further applications we note :

$$\sum_n \left(\frac{\partial G}{\partial\omega}\right)^2 \frac{1}{G} \approx \frac{N\,A^2\tau^2}{2} = \sigma_\omega^{-2} \tag{21}$$

which follows from eqs.18,20, and also

$$\delta\omega = \sigma_\omega^2 \sum_n \frac{\partial G}{\partial\omega}\,\frac{\mathcal{N}_n - G}{G} \tag{22}$$

which follows from eqs.16,20.

3 Set-subset problem for muon (g–2): ω versus fit start time

In the muon (g–2) data analysis, one of the most common checks for systematic errors is a test for stability of (g–2) frequency ω versus histogram fit start time. For that test we compare our result ω_1, obtained from "standard" largest possible fitting time interval, with another result, ω_2, obtained from fitting over smaller time interval. The "standard" time interval starts from some time T_1 after muon beam injection and lasts about 10 lifetimes. For "smaller" time interval we shift fit start time forward to some time T_2. The upper time limit is not very crucial for these studies due to exponential decay of number of events and therefore was fixed and same for both cases. In practice both T_1 and T_2 are lower limits of channels P_1 and P_2, respectively, of the entire histogram.

Due to limited statistics we do not expect ω_2 to be equial to ω_1 exactly, but rather fluctuate around ω_1 with r.m.s.$=\sqrt{\langle\,(\omega_1-\omega_2)^2\rangle}$. Since "true" value $\omega_\circ$ for both cases is the same, $\omega_1-\omega_2$ can be replaced by $\delta\omega_1-\delta\omega_2 \equiv (\omega_1-\omega_\circ)-(\omega_2-\omega_\circ)=\omega_1-\omega_2$ and then eq.6 or eq.16 can be used. The latter is simpler, of course, but requires usage of center of gravity time frame. For that we rewrite function (14) as $G_1(t)$ and $G_2(t)$ in such a way that they would effectively have time of origin at one lifetime τ after T_1 and T_2 respectively:

$$
\begin{aligned}
G_1(t) &= N_\circ\, e^{(-T_1-\tau)/\tau}\, e^{-(t-T_1-\tau)/\tau}\,[1+A\,\cos\{\,\omega(t-T_1-\tau)+\omega(T_1+\tau)+\phi\}\,]= \\
&= N_{\circ 1}\, e^{-(t-T_1-\tau)/\tau}\,[1+A\,\cos\{\,\omega(t-T_1-\tau)+\phi_1\,\}\,] \qquad (23)\\
G_2(t) &= N_{\circ 2}\, e^{-(t-T_2-\tau)/\tau}\,[1+A\,\cos\{\,\omega(t-T_2-\tau)+\phi_2\,\}\,] \qquad (24)
\end{aligned}
$$

As it's evident from eqs.(23) and (24), we just have redefined normalization constant $N_\circ$ and phase ϕ, but in a special way: we have hidden some amount of parameters τ and ω into new $N_\circ$ and ϕ, respectively.

Now construct $\delta\omega_1-\delta\omega_2$:

$$
\begin{aligned}
\delta\omega_1-\delta\omega_2 &= \sigma^2_{\omega 1}\sum_{n=P_1}^{n_{max}}\frac{\partial G_1}{\partial\omega}\frac{\mathcal{N}_n-G_1}{G_1} \; - \; \sigma^2_{\omega 2}\sum_{n=P_2}^{n_{max}}\frac{\partial G_2}{\partial\omega}\frac{\mathcal{N}_n-G_2}{G_2} = \\
&= \sigma^2_{\omega 1}\sum_{n=P_1}^{P_2}\left(\frac{\partial G_1}{\partial\omega}\frac{\mathcal{N}_n-G_1}{G_1}\right)+ \\
&\quad + \sum_{n=P_2}^{n_{max}}\left(\sigma^2_{\omega 1}\frac{\partial G_1}{\partial\omega}\frac{\mathcal{N}_n-G_1}{G_1}-\sigma^2_{\omega 2}\frac{\partial G_2}{\partial\omega}\frac{\mathcal{N}_n-G_2}{G_2}\right)= \qquad (25)
\end{aligned}
$$

$$= \sigma^2_{\omega 1} \sum_{i=P_1}^{P_2} \left(\frac{\partial G_1}{\partial \omega} \frac{\mathcal{N}_n - G_1}{G_1} \right) + \sum_{n=P_2}^{n_{max}} \frac{\mathcal{N}_n - G}{G} \left(\sigma^2_{\omega 1} \frac{\partial G_1}{\partial \omega} - \sigma^2_{\omega 2} \frac{\partial G_2}{\partial \omega} \right)$$

and find $\langle\, (\delta\omega_1 - \delta\omega_2)^2 \,\rangle$:

$$\begin{aligned} \langle\, (\delta\omega_1 - \delta\omega_2)^2 \,\rangle = {} & \sigma^4_{\omega 1} \sum_{n=P_1}^{P_2} \left(\frac{\partial G_1}{\partial \omega} \right)^2 \frac{\langle\, (\mathcal{N}_n - G_1)^2 \rangle}{G_1^2} + \\ & + \sum_{n=P_2}^{n_{max}} \frac{\langle\, (\mathcal{N}_n - G)^2 \rangle}{G^2} \left(\sigma^2_{\omega 1} \frac{\partial G_1}{\partial \omega} - \sigma^2_{\omega 2} \frac{\partial G_2}{\partial \omega} \right)^2 = \\ & = \sigma^4_{\omega 1} \sum_{n=P_1}^{P_2} \frac{1}{G_1} \left(\frac{\partial G_1}{\partial \omega} \right)^2 + \sum_{n=P_2}^{n_{max}} \frac{1}{G} \left(\sigma^2_{\omega 1} \frac{\partial G_1}{\partial \omega} - \sigma^2_{\omega 2} \frac{\partial G_2}{\partial \omega} \right)^2 \end{aligned} \tag{26}$$

Evaluation of sums in the right hand side of eq.26 is given in Appendix, the result is

$$\sum_{n=P_1}^{P_2} \frac{1}{G_1} \left(\frac{\partial G_1}{\partial \omega} \right)^2 \approx \sigma^{-2}_{\omega 1} - \sigma^{-2}_{\omega 2} - \sigma^{-2}_{\omega 2} \left(\frac{T_2 - T_1}{\tau} \right)^2 \qquad \text{and} \tag{27}$$

$$\sum_{n=P_2}^{n_{max}} \frac{1}{G} \left(\sigma^2_{\omega 1} \frac{\partial G_1}{\partial \omega} - \sigma^2_{\omega 2} \frac{\partial G_2}{\partial \omega} \right)^2 \approx \sigma^4_{\omega 1} \left(\frac{T_2 - T_1}{\tau} \right)^2 \sigma^{-2}_{\omega 2} + (\sigma^2_{\omega 1} - \sigma^2_{\omega 2})^2 \, \sigma^{-2}_{\omega 2} \tag{28}$$

That gives:

$$\begin{aligned} \langle\, (\delta\omega_1 - \delta\omega_2)^2 \,\rangle = {} & \sigma^2_{\omega 1} - \sigma^4_{\omega 1} \sigma^{-2}_{\omega 2} - \sigma^4_{\omega 1} \sigma^{-2}_{\omega 2} \left(\frac{T_2 - T_1}{\tau} \right)^2 + \\ & + (\sigma^2_{\omega 1} - \sigma^2_{\omega 2})^2 \sigma^{-2}_{\omega 2} + \sigma^4_{\omega 1} \sigma^{-2}_{\omega 2} \left(\frac{T_2 - T_1}{\tau} \right)^2 = \sigma^2_{\omega 2} - \sigma^2_{\omega 1} \end{aligned} \tag{29}$$

and finally :

$$\langle\, (\omega_1 - \omega_2)^2 \,\rangle = \langle\, (\delta\omega_1 - \delta\omega_2)^2 \,\rangle = \sigma^2_{\omega 2} - \sigma^2_{\omega 1} \tag{30}$$

Fig.1 from [3] shows (g–2) frequency ω as a function of fit start time. Statistical errors of ω are shown representively by vertical bars. Smooth solid line is centred at initial value of frequency ω_1 and indicates one sigma band allowed for "random walk" of ω as calculated within eq.30.

Eq.30 was derived for (g–2) frequency ω essentially for the simplest case of no correlations between ω and other parameters. It can be proved, however, that eq.30 is far more general and valid for the fit start time check no matter of how many fitted paramers function $f(\boldsymbol{x}; t)$ has and whether or not they correlate with each other.

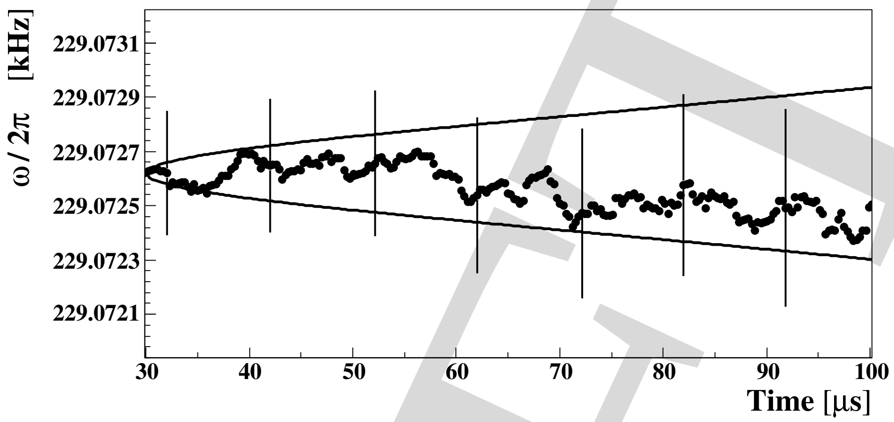

Fig. 1. (g–2) frequency ω as a function of fit start time

Denote x_{j1} and x_{j2} to be solutions for the full set and for some subset, respectively. Since

$$\left\langle (x_{j1} - x_{j2})^2 \right\rangle = \left\langle (\delta x_{j1} - \delta x_{j2})^2 \right\rangle = \left\langle (\delta x_{j1})^2 \right\rangle - 2\,\langle \delta x_{j1}\,\delta x_{j2} \rangle + \left\langle (\delta x_{j2})^2 \right\rangle =$$
$$= \sigma_1^2 - 2\,\langle \delta x_{j1}\,\delta x_{j2} \rangle + \sigma_2^2 \qquad (31)$$

find $\langle \delta x_{j1}\,\delta x_{j2} \rangle$:

$$\begin{aligned}
\langle \delta x_{j1}\,\delta x_{j2} \rangle &= \Bigg\langle \frac{1}{det(\mathcal{A}_1)} \sum_{i=1}^{L} A_{ij\,1} \sum_{n=P_1}^{n_{max}} \frac{f_i'}{f_\circ}\,(\mathcal{N}_n - f_\circ) \times \\
&\times \frac{1}{det(\mathcal{A}_2)} \sum_{k=1}^{L} A_{kj\,2} \sum_{m=P_2}^{n_{max}} \frac{f_k'}{f_\circ}\,(\mathcal{N}_m - f_\circ) \Bigg\rangle = \\
&= \frac{1}{det(\mathcal{A}_1)}\,\frac{1}{det(\mathcal{A}_2)} \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij\,1}\;A_{kj\,2} \sum_{n=P_2}^{n_{max}} \frac{f_i'}{f_\circ}\frac{f_k'}{f_\circ} \left\langle (\mathcal{N}_n - f_\circ)^2 \right\rangle = \\
&= \frac{1}{det(\mathcal{A}_1)}\,\frac{1}{det(\mathcal{A}_2)} \sum_{i=1}^{L}\sum_{k=1}^{L} A_{ij\,1}\;A_{kj\,2} \sum_{n=P_2}^{n_{max}} \frac{f_i' f_k'}{f_\circ} = \\
&= \frac{1}{det(\mathcal{A}_1)}\,\frac{1}{det(\mathcal{A}_2)} \sum_{i=1}^{L} A_{ij\,1} \times \sum_{k=1}^{L} A_{kj\,2}\;a_{kj\,2} = \\
&= \frac{1}{det(\mathcal{A}_1)}\,\frac{1}{det(\mathcal{A}_2)}\;A_{jj\,1} \times det(\mathcal{A}_2) = \frac{M_{jj\,1}}{det(\mathcal{A}_1)} = \sigma_1^2 \qquad (32)
\end{aligned}$$

Here we use $\sum_{i=1}^{L} A_{ij\,1} \times \sum_{k=1}^{L} A_{kj\,2}\;a_{ki\,2} = A_{jj\,1} \times det(\mathcal{A}_2)$, see derivation of eqs.9,10. In eq.32 we denote $\mathcal{A}_1$ and $\mathcal{A}_2$ to be the error matrix $\mathcal{A}$ from eqs.4,5 calculated for the full set and for the subset respectively. Same for cofactors A_{ij}, minors M_{ij} and matrix elements a_{ij}.

Finally, from eqs.31 and 32 we have:

$$\left\langle (x_{j1} - x_{j2})^2 \right\rangle = \sigma_1^2 - 2\sigma_1^2 + \sigma_2^2 = \sigma_2^2 - \sigma_1^2 \tag{33}$$

which is, in a sense, generalization of eq.30.

4 Set-subset problem for muon (g–2): ω versus energy freshold

We now consider another kind of the set-subset problem arising in muon (g–2) experiment, namely stability of (g–2) frequency ω versus energy treshold E_{thr} of decay electrons. Suppose we have 5-parameter distribution (14) for some "full" set and for a subset for which we subtracted some particular energy bin (rise E_{thr}) from the full set. Chi-squared optimization gives us $\delta\omega_1$ for the first case (full set) and $\delta\omega_2$ for the second case (subset). 5-parameter functions $G_1(t)$ and $G_2(t)$, describing set and subset, have different normalization constants ($N_{\circ 1}$, $N_{\circ 2}$), asymmetries (A_1, A_2), phases (ϕ_1, ϕ_2), total number of events (N_1, N_2) and statistical errors ($\sigma_{\omega 1}$, $\sigma_{\omega 2}$, etc.) but have same parameters ω and τ and same fit start time. We also introduce function $G_3(t) = G_1(t) - G_2(t)$ which describes this energy bin alone.

As before, we choose time origin at one τ after beginning of histogram, which allows us to use simple eq.22. Then

$$\delta\omega_1 = \sigma_{\omega 1}^2 \sum_n \frac{\partial G_1}{\partial \omega} \frac{\mathcal{N}_{n1} - G_1}{G_1} = \sigma_{\omega 1}^2 \sum_n \frac{\partial G_1}{\partial \omega} \frac{\mathcal{N}_{n2} - G_2 + \mathcal{N}_{n3} - G_3}{G_1} \tag{34}$$

$$\delta\omega_2 = \sigma_{\omega 2}^2 \sum_n \frac{\partial G_2}{\partial \omega} \frac{\mathcal{N}_{n2} - G_2}{G_2} \tag{35}$$

Construct $\delta\omega_1 - \delta\omega_2$:

$$\begin{aligned}\delta\omega_1 - \delta\omega_2 = & \sum_n \left[\sigma_{\omega 1}^2 \frac{\partial G_1}{\partial \omega} \frac{1}{G_1} - \sigma_{\omega 2}^2 \frac{\partial G_2}{\partial \omega} \frac{1}{G_2}\right] (\mathcal{N}_{n2} - G_2) + \\ & + \ \sigma_{\omega 1}^2 \sum_n \frac{\partial G_1}{\partial \omega} \frac{\mathcal{N}_{n3} - G_3}{G_1}\end{aligned} \tag{36}$$

Evaluate $\langle (\delta\omega_1 - \delta\omega_2)^2 \rangle$:

$$\left\langle (\delta\omega_1 - \delta\omega_2)^2 \right\rangle \quad =$$

$$= \sum_n \left[\sigma_{\omega 1}^2 \frac{\partial G_1}{\partial \omega} \frac{1}{G_1} - \sigma_{\omega 2}^2 \frac{\partial G_2}{\partial \omega} \frac{1}{G_2}\right]^2 \langle (\mathcal{N}_{n2} - G_2)^2 \rangle + \sigma_{\omega 1}^4 \sum_n \left(\frac{\partial G_1}{\partial \omega}\right)^2 \frac{\langle (\mathcal{N}_{n3} - G_3)^2 \rangle}{G_1^2} =$$

$$
\begin{aligned}
&= \sum_n \left[\sigma^2_{\omega 1} \frac{\partial G_1}{\partial \omega} \frac{1}{G_1} - \sigma^2_{\omega 2} \frac{\partial G_2}{\partial \omega} \frac{1}{G_2} \right]^2 G_2 \quad + \quad \sigma^4_{\omega 1} \sum_n \left(\frac{\partial G_1}{\partial \omega} \right)^2 \frac{G_3}{G_1^2} = \\
&= \sum_n \left[\sigma^4_{\omega 1} \left(\frac{\partial G_1}{\partial \omega} \right)^2 \frac{G_2}{G_1^2} - 2\, \sigma^2_{\omega 1}\, \sigma^2_{\omega 2} \left(\frac{\partial G_1}{\partial \omega} \right) \left(\frac{\partial G_2}{\partial \omega} \right) \frac{1}{G_1} + \right. \\
&\qquad \left. + \, \sigma^4_{\omega 2} \left(\frac{\partial G_2}{\partial \omega} \right)^2 \frac{1}{G_2} + \sigma^4_{\omega 1} \left(\frac{\partial G_1}{\partial \omega} \right)^2 \frac{G_3}{G_1^2} \right] = \\
&= \sigma^4_{\omega 1} \sum_n \left(\frac{\partial G_1}{\partial \omega} \right)^2 \frac{1}{G_1} - 2\, \sigma^2_{\omega 1}\, \sigma^2_{\omega 2} \sum_n \left(\frac{\partial G_1}{\partial \omega} \right) \left(\frac{\partial G_2}{\partial \omega} \right) \frac{1}{G_1} + \sigma^4_{\omega 2} \sum_n \left(\frac{\partial G_2}{\partial \omega} \right)^2 \frac{1}{G_2} = \\
&= \sigma^2_{\omega 1} - 2\, \sigma^2_{\omega 1}\, \sigma^2_{\omega 2} \sum_n \left(\frac{\partial G_1}{\partial \omega} \right) \left(\frac{\partial G_2}{\partial \omega} \right) \frac{1}{G_1} + \sigma^2_{\omega 2}
\end{aligned}
\tag{37}
$$

Evaluate remaining sum explicitly:

$$
\begin{aligned}
\sum_n \left(\frac{\partial G_1}{\partial \omega} \right) \left(\frac{\partial G_2}{\partial \omega} \right) \frac{1}{G_1} &= \int_{-\tau}^{\infty} \frac{N_{\circ 2}\, A_1\, A_2\, t^2\, e^{-t/\tau}\, \sin(\omega t + \phi_1)\, \sin(\omega t + \phi_2)}{1 + A_1\, \cos(\omega t + \phi_1)} \frac{dt}{b} \approx \\
&\approx \frac{N_{\circ 2}\, A_1\, A_2\, \cos(\phi_1 - \phi_2)}{2\, b} \int_{-\tau}^{\infty} t^2\, e^{-t/\tau}\, dt = \\
&= \frac{N_{\circ 2}\, A_1\, A_2\, \cos(\phi_1 - \phi_2)}{2\, b} \tau^3\, e = \frac{A_1}{A_2} \cos(\phi_1 - \phi_2) \frac{N_2\, A_2^2\, \tau^2}{2} = \\
&= \frac{A_1}{A_2} \cos(\phi_1 - \phi_2) \quad \sigma^{-2}_{\omega 2}
\end{aligned}
\tag{38}
$$

And finally:

$$
\left\langle (\delta\omega_1 - \delta\omega_2)^2 \right\rangle = \sigma^2_{\omega 2} - \sigma^2_{\omega 1} \left(2 \frac{A_1}{A_2} \cos(\phi_1 - \phi_2) - 1 \right) \tag{39}
$$

We use equation $\delta\omega_1 - \delta\omega_2 = \omega_1 - \omega_2$ and rewrite eq.39 in terms of ω_1, ω_2:

$$
\left\langle (\omega_1 - \omega_2)^2 \right\rangle = \sigma^2_{\omega 2} - \sigma^2_{\omega 1} \left(2 \frac{A_1}{A_2} \cos(\phi_1 - \phi_2) - 1 \right) \tag{40}
$$

which reduces to eq.30 when $A_1 = A_2$ and $\phi_1 = \phi_2$.

5 Appendix

Evaluate sums in eq.26:

$$\sum_{n=P_1}^{P_2} \frac{1}{G_1} \left(\frac{\partial G_1}{\partial\omega}\right)^2 \approx \quad \frac{N_{\circ 1}\,A^2}{2\,b} \int\limits_{T_1}^{T_2} e^{-(t-T_1-\tau)/\tau}\,(t-T_1-\tau)^2\,dt =$$

$$= \frac{N_{\circ 1}\,A^2}{2\,b} \int\limits_{-\tau}^{-\tau-T_1+T_2} e^{-x/\tau}\,x^2\,dx = \frac{N_{\circ 1}\,A^2}{2\,b} \quad e^{-x/\tau}\,(-x^2\tau - 2x\tau - 2\tau^3)\,\Big|_{-\tau}^{-\tau-T_1+T_2} =$$

$$= \frac{N_{\circ 1}\,A^2}{2\,b} \left(e^{(\tau+T_1-T_2)/\tau}\,\left[-\tau(-\tau-T_1+T_2)^2 - 2\tau^2(-\tau-T_1+T_2) - 2\tau^3\right] - \right.$$

$$\left. - \,e\,\left[-\tau^3 + 2\tau^3 - 2\tau^3\right]\right) = \frac{N_{\circ 1}\,A^2}{2\,b} \left(e\,\tau^3 - e^{(\tau+T_1-T_2)/\tau}\,\tau\,\left[\tau^2 + (T_2-T_1)^2\right]\right) =$$

$$= \frac{N_{\circ 1}\,A^2\,\tau^3\,e}{2\,b} \left(1 - e^{-(T_2-T_1)/\tau}\,\left[1 + \left(\frac{T_2-T_1}{\tau}\right)^2\right]\right) =$$

$$= \frac{N_1\,A^2\,\tau^2}{2} \left(1 - e^{-(T_2-T_1)/\tau}\,\left[1 + \left(\frac{T_2-T_1}{\tau}\right)^2\right]\right) =$$

$$= \frac{A^2\,\tau^2}{2} \left(N_1 - N_2 - N_2\,\left(\frac{T_2-T_1}{\tau}\right)^2\right) =$$

$$= \sigma_{\omega 1}^{-2} - \sigma_{\omega 2}^{-2} - \sigma_{\omega 2}^{-2}\,\left(\frac{T_2-T_1}{\tau}\right)^2 \tag{41}$$

$$\sum_{n=P_2}^{n_{max}} \frac{1}{G} \left(\sigma_{\omega 1}^2\,\frac{\partial G_1}{\partial\omega} - \sigma_{\omega 2}^2\,\frac{\partial G_2}{\partial\omega}\right)^2 \approx \quad \frac{1}{b} \int\limits_{T_2}^{\infty} \frac{dt}{N_\circ\,e^{-t/\tau}\,[1 + A\,\cos(\omega t + \phi_\circ)\,]} \times$$

$$\times \left[\sigma_{\omega 1}^2\,N_{\circ 1}\,e^{-(t-T_1-\tau)/\tau}\,A\,(t-T_1-\tau)\,\sin(\omega(t-T_1-\tau) + \phi_1) - \right.$$

$$\left. - \,\sigma_{\omega 2}^2\,N_{\circ 2}\,e^{-(t-T_2-\tau)/\tau}\,A\,(t-T_2-\tau)\,\sin(\omega(t-T_2-\tau) + \phi_2)\,\right]^2 =$$

$$= \frac{N_\circ\,A^2}{b} \int\limits_{T_2}^{\infty} dt\,e^{-t/\tau}\,\sin^2(\omega t + \phi_\circ)\,\left[\sigma_{\omega 1}^2\,(t-T_1-\tau) - \sigma_{\omega 2}^2\,(t-T_2-\tau)\,\right]^2 \approx$$

$$\approx \frac{N_\circ\,A^2}{2\,b} \int\limits_{T_2}^{\infty} dt\,e^{-t/\tau}\,\left[\sigma_{\omega 1}^2(T_2-T_1) + (\sigma_{\omega 1}^2 - \sigma_{\omega 2}^2)\,(t-T_2-\tau)\,\right]^2 =$$

$$= \frac{N_\circ\,A^2}{2\,b} \int\limits_{T_2}^{\infty} dt\,e^{-t/\tau}\,\left[\sigma_{\omega 1}^4\,(T_2-T_1)^2 + (\sigma_{\omega 1}^2 - \sigma_{\omega 2}^2)^2\,(t-T_2-\tau)^2\,\right] =$$

$$= \frac{A^2}{2\,b}\,\left[\sigma_{\omega 1}^4\,(T_2-T_1)^2\,N_{\circ 2}\,\tau\,e + (\sigma_{\omega 1}^2 - \sigma_{\omega 2}^2)^2\,N_{\circ 2}\,\tau^3\,e\,\right] =$$

$$= \sigma_{\omega 1}^4\,\left(\frac{T_2-T_1}{\tau}\right)^2\,\sigma_{\omega 2}^{-2} + (\sigma_{\omega 1}^2 - \sigma_{\omega 2}^2)^2\,\sigma_{\omega 2}^{-2} \tag{42}$$

References

[1] S. I. Redin et al., Nucl. Instr. and Meth. **AXXX**, XXXX (200X)

[2] S. I. Redin et al., Nucl. Instr. and Meth. **AXXX**, XXXX (200X)
[3] H. N. Brown et al., Phys. Rev. Lett. **86**, 2227 (2001)