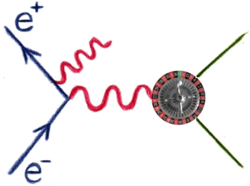
The Why's and How's of covariance matrices in the KLOE ISR analyses

S. E. Müller



Radio MonteCarLOW Meeting
ECT Trento, 11-12 April 2013*

The covariance matrix

A covariance matrix describes the correlation between the fluctuations (uncertainties) of *n* observables (data points)

$$V = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \dots & \sigma_{2n} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \dots & \sigma_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \dots & \sigma_{nn} \end{pmatrix}$$

The diagonal elements contain the squared uncertainties (variances). The off-diagonal elements contain the correlation coefficient ρ:

$$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$$

$$-1 \leq \rho \leq +1$$

If the observables are uncorrelated, the ρ's are zero and the covariance matrix is diagonal.

The covariance matrix

The uncertainty of a function $f = f(x_1, x_2, \ldots, x_n)$ which depends on the (correlated) observables is constructed via

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \ldots & \sigma_{2n} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \ldots & \sigma_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \ldots & \sigma_{nn} \end{pmatrix} \begin{pmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{pmatrix}$$

If the covariance matrix is diagonal, one recovers the familiar formula

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 \sigma_1^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 \sigma_2^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2 \sigma_n^2$$

If the observables are fully correlated ($\rho = +1$), we get

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x_1}\sigma_1 + \frac{\partial f}{\partial x_2}\sigma_2 + \cdots + \frac{\partial f}{\partial x_n}\sigma_n\right)^2$$

Dispersion integral for $a_\mu^{\rm had}$

Evaluating $\mathrm{a}_\mu^{\rm had}$ between $s_{\rm min}$ and $s_{\rm max}$ with a binned data set of n points with binwidth Δs:

$$\mathrm{a}_\mu^{\rm had}[s_{\rm min}, s_{\rm max}] = \frac{1}{4\pi^3}\int_{s_{\rm min}}^{s_{\rm max}} \sigma^{\rm had}(s) K(s)\, ds \simeq \frac{1}{4\pi^3}\sum_{i=1}^{n} \sigma_i^{\rm had} K_i \Delta s$$

With $c_i \equiv \partial \mathrm{a}_\mu^{\rm had}/\partial\sigma_i = (1/4\pi^3)K_i\Delta s$ we obtain for the uncertainty on $\mathrm{a}_\mu^{\rm had}$

$$\begin{aligned}\left(\sigma_{\mathrm{a}_\mu^{\rm had}}\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j V_{ij} \\ &= \underbrace{\sum_{i=1}^{n} c_i^2\sigma_i^2}_{\substack{i=j,\\ \text{diagonal term}}} + \underbrace{\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} c_i c_j \sigma_{ij}}_{\substack{i\neq j,\\ \text{off}-\text{diagonal term}}}\end{aligned}$$

The KLOE ISR analysis

From the observed spectrum to the ISR cross section:

$$\frac{d\sigma^{\pi\pi+\gamma}}{ds_\pi} = \frac{\Delta N_{Obs} - \Delta N_{Bkg}}{\Delta s_\pi} \cdot \frac{1}{\int \mathcal{L} dt} \cdot \frac{1}{\epsilon_{Sel}}$$

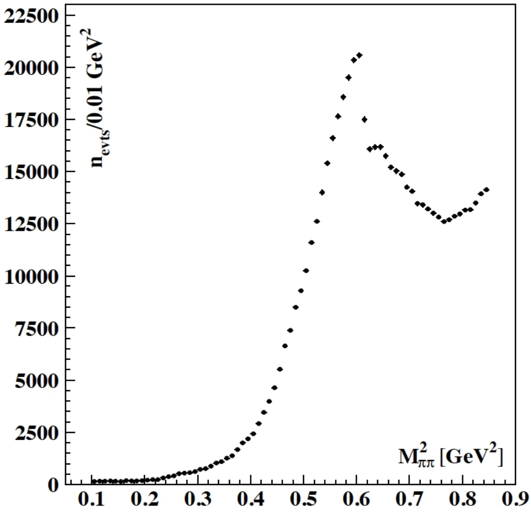
Applying the radiator function $H(s_\pi, s)$:

$$\sigma^{\pi\pi}(s_\pi) = \frac{d\sigma^{\pi\pi+\gamma}}{ds_\pi} \cdot s \cdot \frac{1}{H(s_\pi, s)}$$

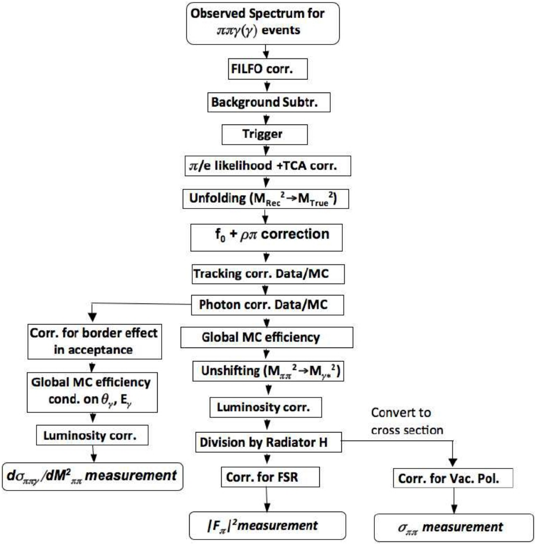
Dispersion integral:

$$a_\mu^{\pi\pi}[s_{\min}, s_{\max}] = \frac{1}{4\pi^3} \int_{s_{\min}}^{s_{\max}} \sigma^{\pi\pi}(s) K(s)\, ds$$

The observed spectrum (KLOE10)



Data analysis flow (KLOE10)



Uncertainties

Uncorrelated uncertainties:
Uncertainties independent between different data points (bins) - enter only in diagonal elements of covariance matrix.

- e.g. stat. errors of observed spectrum

"Sum of squares" in the uncertainty estimation of $a_\mu^{\pi\pi}$:

$$\left(\sigma_{\mathrm{a}_\mu^{\pi\pi}}\right)^2 = \sum_{i=1}^{n} \left(\frac{1}{4\pi^3} (\sigma_{\sigma_{\pi\pi}})_i K_i \Delta s \right)^2$$

$(\sigma_{\sigma_{\pi\pi}})_i$ can combine several uncorrelated uncertainties

$$(\sigma_{\sigma_{\pi\pi}})_i = \sqrt{\left(\sigma_i^{\mathrm{spectrum}}\right)^2 + \left(\sigma_i^{\mathrm{eff}}\right)^2 + \ldots}$$

Uncertainties

Fully correlated uncertainties:

Uncertainties affect all data points (bins) in a similar way. Covariance matrix elements are constructed as $V_{ij} = +1 \cdot \sigma_i \sigma_j$.

- e.g. error on int. luminosity measurement $\int \mathcal{L} dt$

"Square of sum" in the uncertainty estimation of $a_\mu^{\pi\pi}$:

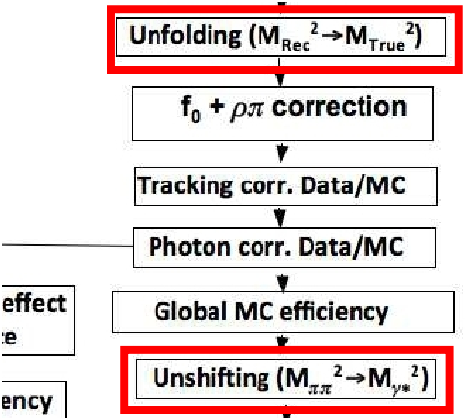
$$\left(\sigma_{\mathrm{a}_\mu^{\pi\pi}}\right)^2 = \left(\sum_{i=1}^{n} \frac{1}{4\pi^3} (\sigma_{\sigma_{\pi\pi}})_i K_i \Delta s\right)^2$$

$(\sigma_{\sigma_{\pi\pi}})_i$ can not be a combined uncertainty (unless it's constant) - need to evaluate $(\sigma_{\mathrm{a}_\mu^{\pi\pi}})^2$ for each individual uncertainty source (background, efficiencies, etc.) - then combine quadratically:

$$\left(\sigma_{\mathrm{a}_\mu^{\pi\pi}}\right) = \sqrt{\left(\sigma_{\mathrm{a}_\mu^{\mathrm{background}}}\right)^2 + \left(\sigma_{\mathrm{a}_\mu^{\mathrm{eff}_1}}\right)^2 + \left(\sigma_{\mathrm{a}_\mu^{\mathrm{eff}_2}}\right)^2 + \ldots}$$

Uncertainty

Partially correlated uncertainties:



Unfolding

Correction for detector resolution

- Passage from reconstructed quantitity $s_\pi^{\mathrm{rec}} \longrightarrow s_\pi^{\mathrm{true}}$ requires re-distribution of events:

$$N_i^{\mathrm{true}} = \sum_{j=1}^{n} P(N_i^{\mathrm{true}} | N_j^{\mathrm{rec}}) \cdot N_j^{\mathrm{rec}}$$

- $P(N_i^{\mathrm{true}} | N_j^{\mathrm{rec}})$ from unfolding method using Bayesian approach (G. D'Agostino, NIM A362 (1995) 487)

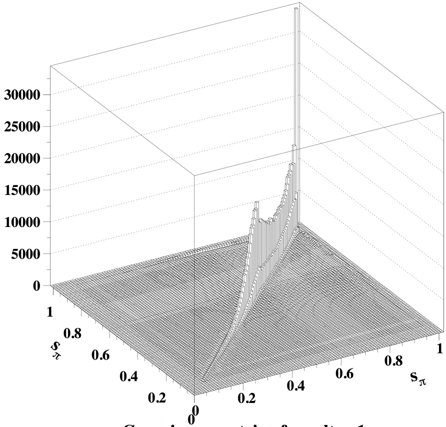
$$P(N_i^{\mathrm{true}} | N_j^{\mathrm{rec}}) = \underbrace{\frac{P(N_j^{\mathrm{rec}} | N_i^{\mathrm{true}}) \cdot P_0(N_i^{\mathrm{true}})}{\sum_{k=1}^{n_{\mathrm{true}}} P(N_j^{\mathrm{rec}} | N_k^{\mathrm{true}}) \cdot P_0(N_k^{\mathrm{true}})}}_{\text{from MC}}$$

- Need to find contributions to covariance matrix of the N_i^{true} due to $P(N_i^{\mathrm{true}} | N_j^{\mathrm{rec}})$ and N_j^{rec}.

Unfolding

Contribution to covariance matrix

- D'Agostini's code provides the covariance matrix after the unfolding procedure:

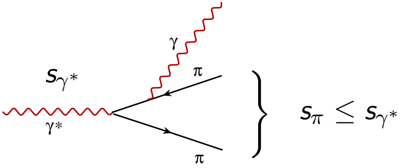


Covariance matrix of results - 1 -

Unshifting

Correction for shift in s_π due to FSR events

The presence of FSR shifts the observed value of s_π (evaluated from the 2 pion tracks' momenta) away from the invariant mass squared of the virtual photon s_{γ^*}:



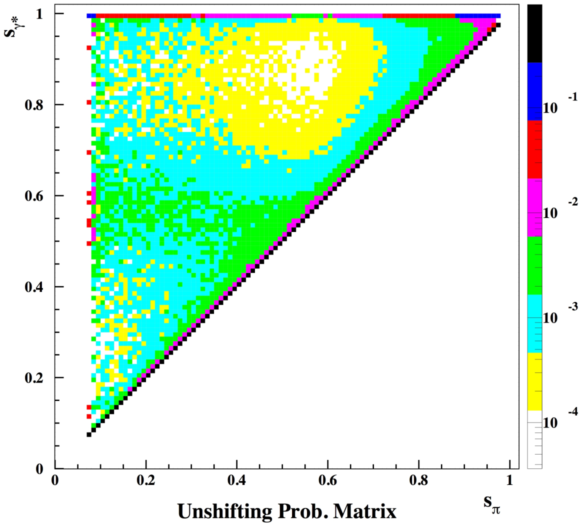
Redistribute events to obtain "unshifted" distribution:

$$N_i^{s_{\gamma^*}} = \sum_{j=1}^{n} P(N_i^{s_{\gamma^*}} | N_j^{s_\pi}) \cdot N_j^{s_\pi}$$

$P(N_i^{s_{\gamma^*}} | N_j^{s_\pi})$ obtained from MC (Phokhara_omega)

Unshifting

Probability matrix:



Unshifting

Contribution to covariance matrix:

Rewrite the unshifting procedure as

$$\vec{y}_{s_{\gamma^*}} = \mathbf{M} \cdot \vec{x}_{s_\pi} \Leftrightarrow y_k = \sum_{i=1}^{n} M_{ki} \cdot x_i$$

Then error propagation gives

$$V^y_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ki} \underbrace{\mathrm{cov}(x_i, x_j)}_{V^x_{ij}} M_{lj} + \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \, \mathrm{cov}(M_{ki}, M_{lj}) \, x_j$$

$\mathrm{cov}(x_i, x_j)$ is the element V^x_{ij} of the covariance matrix of $\vec{x}_{s_\pi}$.
Need to evaluate $\mathrm{cov}(M_{ki}, M_{lj})$

Unshifting

Evaluation of $\mathrm{cov}(M_{ki}, M_{lj})$:

The transfer matrix **M** is constructed from a 2 dim population histogram in s_{γ^*} and s_π which is normalized to the sum $N_i^{\mathrm{tot}} = \sum_{k=1}^{n} N_{ik}$ for a given bin i in the shifted variable s_π.

- Each column M_{ki} for given i follows a multinomial distribution:

$$M_{ki} = N_{ki}/N_i^{\mathrm{tot}} \ , \ \sum_{k=1}^{n} M_{ki} = 1$$

- $\mathrm{cov}(M_{ki}, M_{lj}) = 0$ for $i \neq j$

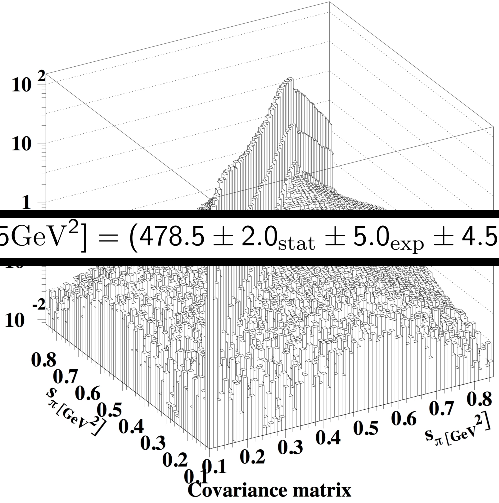
Therefore we get

$$\mathrm{cov}(M_{ki}, M_{lj}) = \begin{cases} M_{ki}(1 - M_{ki})/N_i^{\mathrm{tot}} & \text{if } i = j,\ k = l \\ -(M_{ki}M_{li})/N_i^{\mathrm{tot}} & \text{if } i = j,\ k \neq l \\ 0 & \text{else} \end{cases}$$

And finally...

The final covariance matrix:

Including also all uncorrelated uncertainties in the diagonal elements



$$\Delta a_\mu^{\pi\pi}[0.1 - 0.85\mathrm{GeV}^2] = (478.5 \pm 2.0_{\mathrm{stat}} \pm 5.0_{\mathrm{exp}} \pm 4.5_{\mathrm{theo}}) \times 10^{-10}$$

How to publish?

Long list of numbers

- KLOE08/12: 3600 entries
- KLOE10: 5625 entries
- BaBar: 113569 entries

- Don't publish
  - Ask people to contact you via email
- Publish on institute's webpage
  - `http://www.lnf.infn.it/kloe/ppg/` for KLOE data
- EPAPS depository
  - BaBar result:
    `ftp://ftp.aip.org/epaps/phys_rev_lett/E-PRLTAO-103-045950/`
- ...