Dispersive evaluations of HVP

Martin Hoferichter, Thomas Leplumey

Aug 4, 2026

Ninth Plenary Workshop of the Muon $g-2$ Theory Initiative

Storrs, Connecticut

Dispersive evaluations of HVP: outline

- Main objective: evaluation of $a_\mu^{\text{HVP, LO}}$ from the $e^+e^- \to$ hadrons data base that
  - faithfully propagates the uncertainties from the underlying data
  - optimizes the precision that can be achieved from the combination of data sets

  $\hookrightarrow$ either fully data-driven or including QCD constraints
- Objectives of imposing QCD constraints (analyticity, unitarity, crossing, low-energy theorems)
  - check consistency of data base
  - improve precision
  - establish and explore relations among different channels
- In either case, information on **statistical and systematic uncertainties is crucial**
- Outline for this discussion session:
  - Update on 3π channel MH
  - Update on 2π channel Thomas Leplumey
  - Open discussion

Dispersive evaluations of $e^+e^- \to 3\pi$

- Data base:
  - Data before 2021 SND, CMD-2, ...
  - BaBar 2021, 2110.00520
  - Belle II 2024, 2404.04915
  - SND 2026, 2603.01635
- Dispersive representation based on:
  - Khuri–Treiman equations for $\gamma^* \to 3\pi$
  - Normalization function $a(q^2)$ in terms of $\omega, \phi, \omega', \omega''$ resonances and conformal polynomial
  - Low-energy theorem for $\gamma\pi \to \pi\pi$ (chiral anomaly C_A, fixed in default fit)
  - ρ–ω mixing via coupled-channel formalism
- Errors will be given as:
  1. Experiment (including scale factor $S = \sqrt{\chi^2/\text{dof}}$)
  2. Truncation of conformal expansion
  3. Phase of ρ–ω mixing parameter
  4. Sum in quadrature of 1.–3.

Dispersive evaluations of $e^+e^- \to 3\pi$

	Belle II	BaBar	SND	Global fit (BaBar+SND+earlier) (w/o Belle II)
χ^2/dof	$\frac{218.2}{159} = 1.77$	$\frac{183.7}{129} = 1.42$	$\frac{134.7}{126} = 1.07$	$\frac{672.8}{467} = 1.44$
M_ω [MeV]	782.72(5)(0)(0)[5]	782.542(11)(1)(0)[11]	782.574(10)(4)(2)[11]	782.585(10)(1)(2)[11]
Γ_ω [MeV]	8.43(8)(1)(2)[8]	8.718(26)(2)(3)[26]	8.631(19)(14)(7)[25]	8.672(15)(12)(9)[21]
M_ϕ [MeV]	1019.24(5)(1)(2)[5]	1019.276(6)(0)(0)[6]	1019.258(14)(25)(4)[29]	1019.246(11)(7)(4)[14]
Γ_ϕ [MeV]	3.98(11)(2)(0)[11]	4.289(10)(0)(0)[10]	4.271(26)(16)(4)[31]	4.279(9)(1)(1)[9]
$a_\mu^{3\pi}\big\vert_{\leq 1.8\,\text{GeV}}$	48.7(1.4)(0.1)(0.1)[1.4]	45.99(53)(15)(6)[55]	45.91(35)(15)(2)[38]	45.91(28)(16)(9)[33]
C_A/C_A^{WZW}	1.069(58)(29)(12)[66]	0.945(27)(14)(18)[36]	1.005(19)(51)(16)[57]	1.024(14)(38)(16)[44]

- Belle II data do not pass the test
  - Bad fit quality
  - Resonance parameters off
  - (Strong) pull in C_A, see below

$\hookrightarrow$ enough cause to discard?

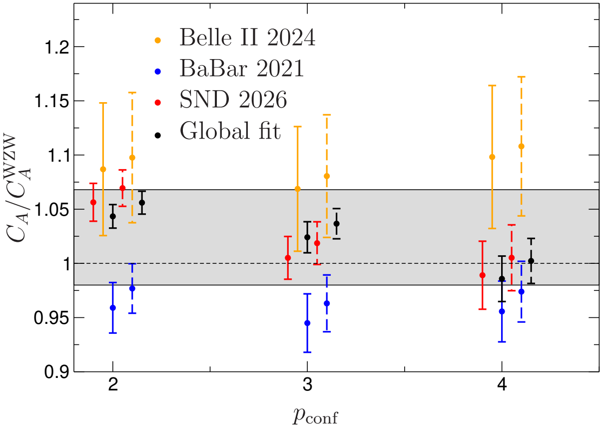
Dispersive evaluations of $e^+e^- \to 3\pi$

	Belle II	BaBar	SND	Global fit (BaBar+SND+earlier) (w/o Belle II)
χ^2/dof	$\frac{218.2}{159} = 1.77$	$\frac{183.7}{129} = 1.42$	$\frac{134.7}{126} = 1.07$	$\frac{672.8}{467} = 1.44$
M_ω [MeV]	782.72(5)(0)(0)[5]	782.542(11)(1)(0)[11]	782.574(10)(4)(2)[11]	782.585(10)(1)(2)[11]
Γ_ω [MeV]	8.43(8)(1)(2)[8]	8.718(26)(2)(3)[26]	8.631(19)(14)(7)[25]	8.672(15)(12)(9)[21]
M_ϕ [MeV]	1019.24(5)(1)(2)[5]	1019.276(6)(0)(0)[6]	1019.258(14)(25)(4)[29]	1019.246(11)(7)(4)[14]
Γ_ϕ [MeV]	3.98(11)(2)(0)[11]	4.289(10)(0)(0)[10]	4.271(26)(16)(4)[31]	4.279(9)(1)(1)[9]
$a_\mu^{3\pi}\big\|_{\leq 1.8\,\text{GeV}}$	48.7(1.4)(0.1)(0.1)[1.4]	45.99(53)(15)(6)[55]	45.91(35)(15)(2)[38]	45.91(28)(16)(9)[33]
C_A/C_A^{WZW}	1.069(58)(29)(12)[66]	0.945(27)(14)(18)[36]	1.005(19)(51)(16)[57]	1.024(14)(38)(16)[44]

- BaBar and SND look amazingly consistent
- Treatment of systematic errors for SND unclear
  - L, $N_{3\pi}$, ϵ, RC each fully correlated or not?
  - If yes, then fit explodes
  - Above results based on best guess for covariance matrix by Aidan Wright from 2603.01635

$\hookrightarrow$ should we really guess the correlations ourselves?

Dispersive evaluations of $e^+e^- \to 3\pi$: chiral anomaly



Erros include scale factor $\sqrt{\chi^2/\mathrm{dof}}$

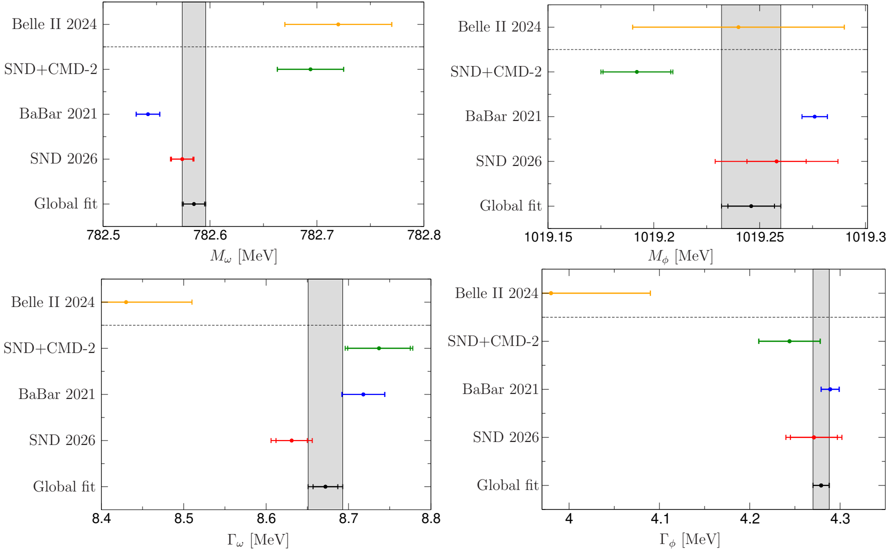
solid: $\delta_\epsilon = 0°$

dashed: $\delta_\epsilon = 3.5°$

band: global fit including truncation error

- Main systematic error from truncation of conformal polynomial p_{conf}
- BaBar, SND, and global fit converge towards WZW expectation
- Belle II systematically above

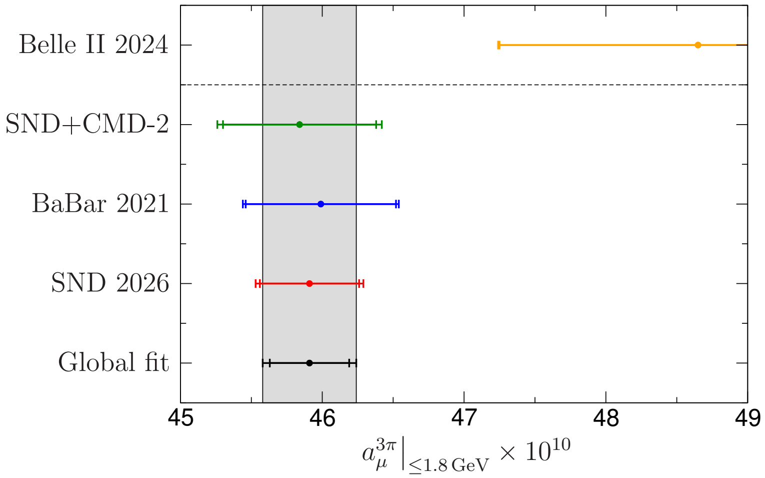
Dispersive evaluations of $e^+e^- \to 3\pi$: resonance parameters



All parameters exclude vacuum polarization

BaBar errors for $M_{\omega,\phi}$ do not account for energy-calibration uncertainty

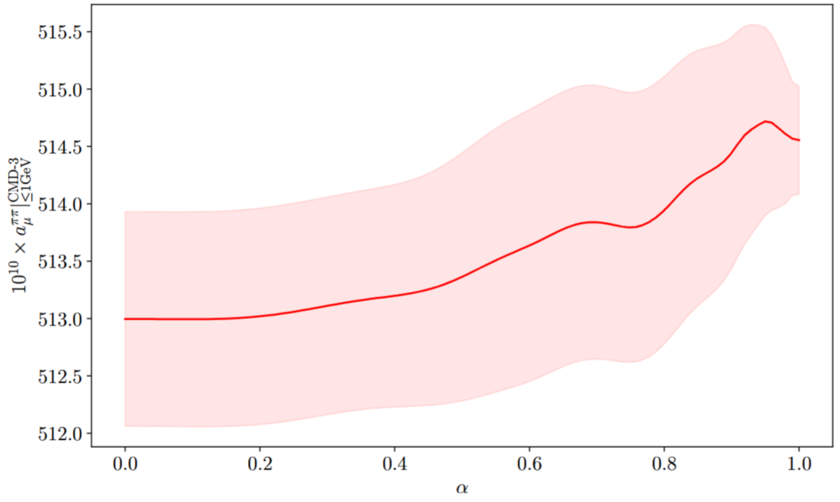
Dispersive evaluations of $e^+e^- \to 3\pi$: contribution to a_μ



Uncertainties on uncertainties

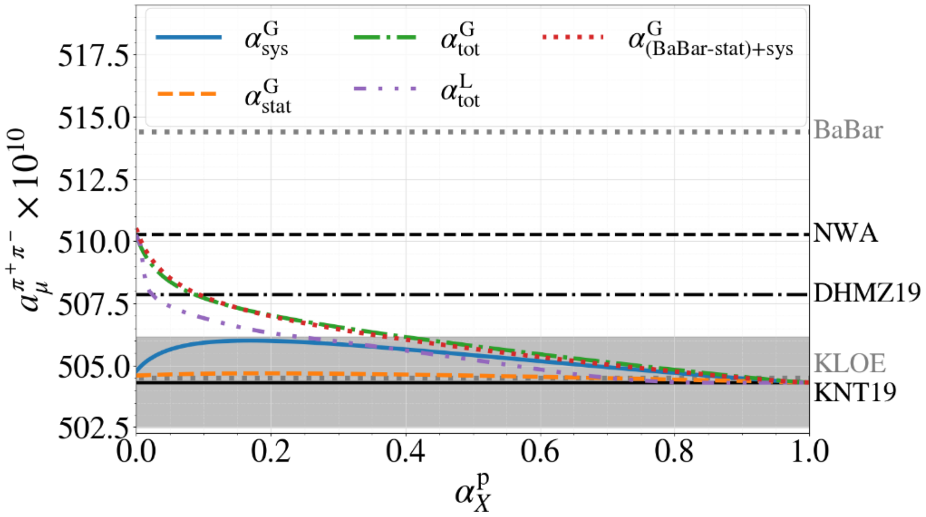
- Nominal CHKLS fit results previously assumed 100% correlated syst. when systematic covariance was not provided by the experiments
- But we are now aware that such an assumption **influences a lot the results**
  - arXiv:2501.09643 (Leplumey, Stoffer): implements a simple local decorrelation scheme
  - arXiv:2604.25004 (KNTW): more detailed studies, interplay local & global decorrelation schemes

100% correlated



Uncorrelated

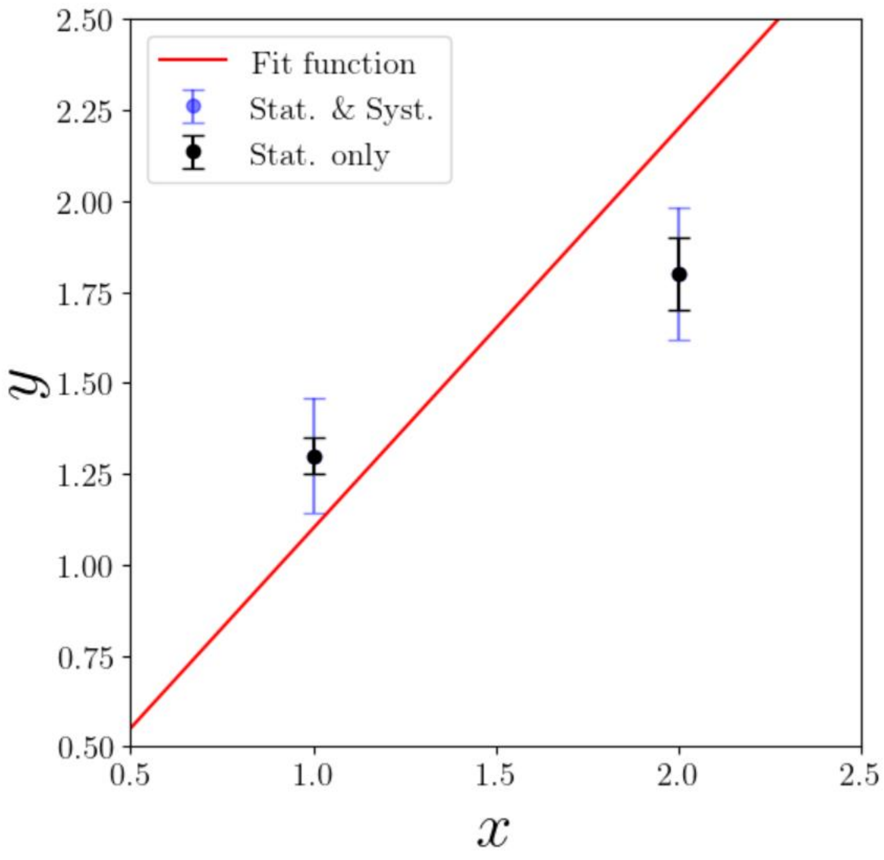
Uncorrelated



100% correlated

1. We would like to find a procedure that **faithfully covers** the post-fit constraints given by **all possible correlations** at realistic computation costs
2. To what extent do we need to consider « all possible correlations »? Can we define the coverage with respect to a **simpler / more relevant ensemble of correlations**?

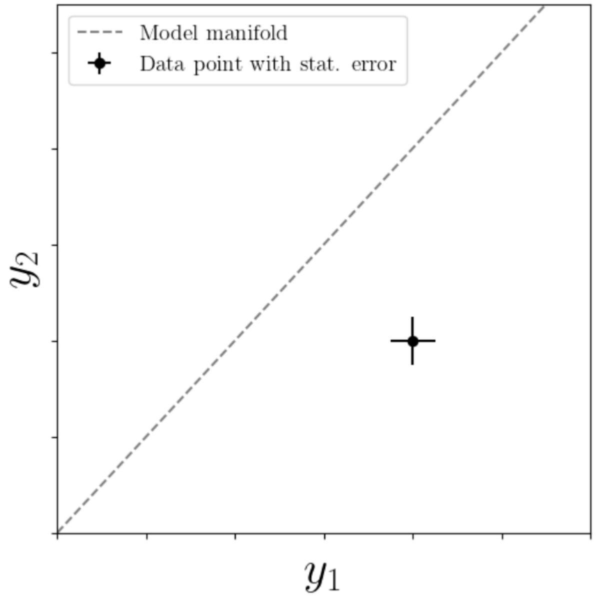
- Consider a toy model:
  - 2 data points, diagonal statistical covariance
  - Fit with a linear function $y = ax$
  - Given systematic uncertainty for each data point, but unknown correlation



2D Observable space: (y_1, y_2)

1D Model manifold: $\{(ax_1, ax_2), a \in \mathbb{R}\}$

We can represent the observed data point and the model manifold in this 2D observable space



Impact of different correlation assumptions on toy model

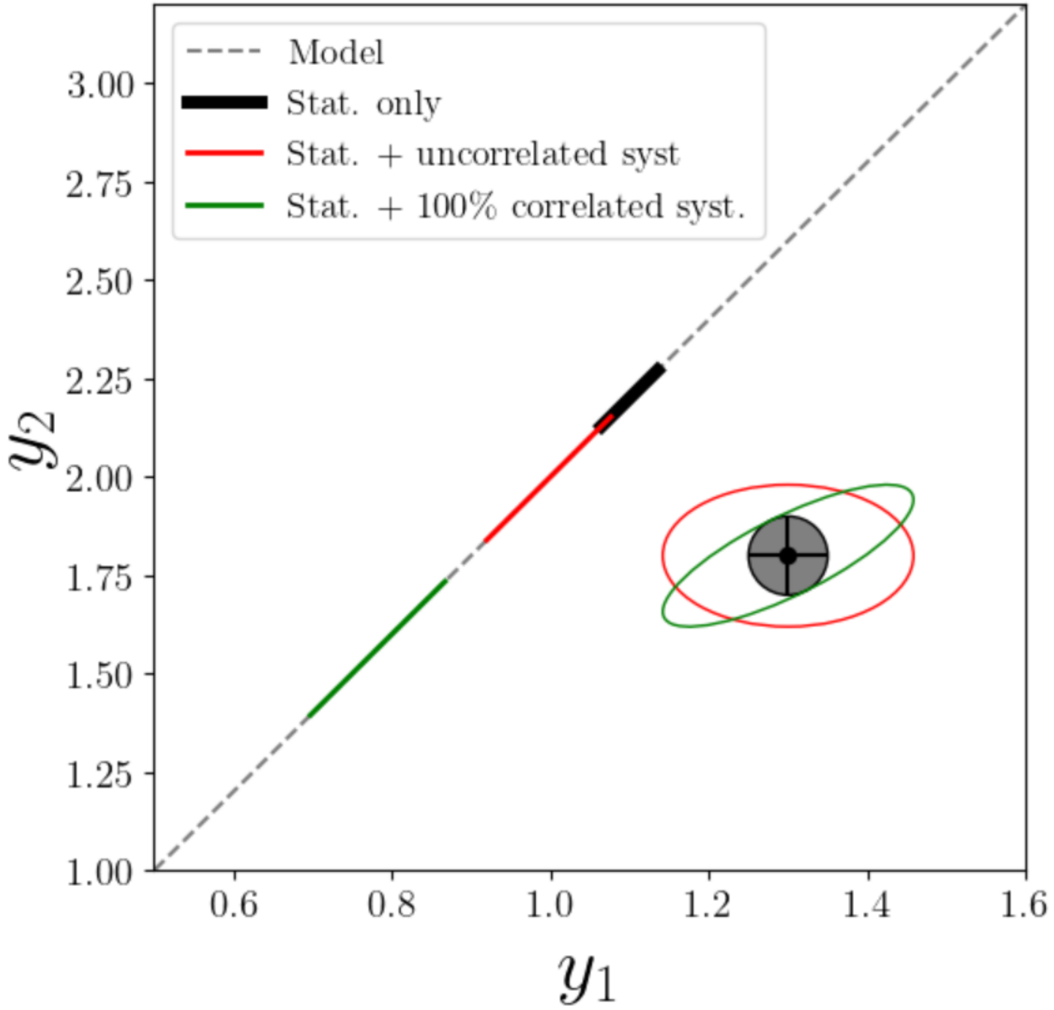
- This toy model illustrates the effect of **changing the syst. correlation factor**
  - **Size of the post-fit uncertainty**
  - **Shift in the best-fit parameters**
  - **Goodness of fit**

- Fit ≈ projection of data on model manifold wrt given covariance matrix

- Distance to model manifold = goodness of fit

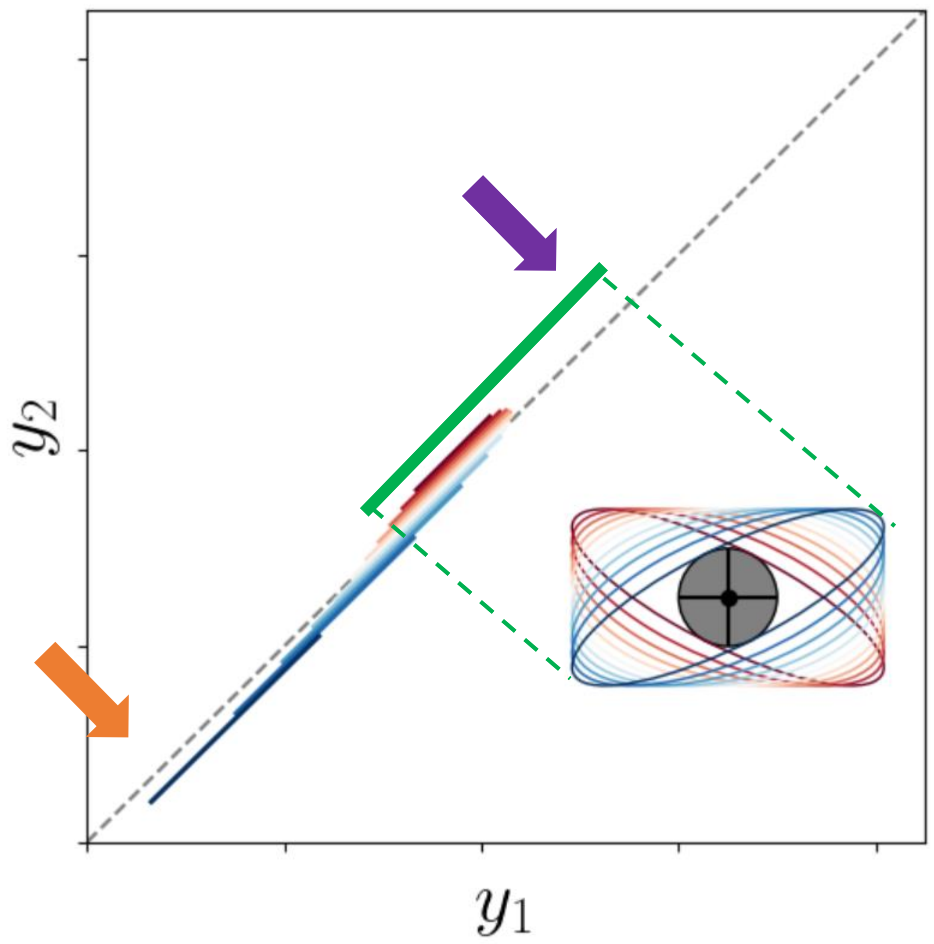
  - **Stat only: very bad GoF**
  - **Uncorrelated syst: good GoF**
  - 100% correlated: poorer GoF

  } *Big shift*



Effect of the method proposed at last workshop

- By varying the correlation factor between -1 and 1, we can determine the range spanned within the model manifold



- In reality: $\mathcal{O}(10^2)$ data points i.e. $\mathcal{O}(10^4)$ correlation factors
  $\rightarrow$ we can't just scan all of them
- Proposal at last Workshop: Shift the data points one by one within their syst errors and add the shifts linearly
  $\rightarrow$ **result shown on toy model**
- Not representative of the fit (disregards model constraints)
  - **Over-coverage:** covered regions with very bad GoF
  - **Under-coverage:** uncovered regions favored by the model

How would we want to treat this problem?

- Mathematical reason why this procedure does not do what we want:

$$\Sigma_{\text{params}} = \left[P^{\mathrm{T}}\left(\Sigma_{\text{data}}^{\text{stat}} + \Sigma_{\text{data}}^{\text{syst}}\right)^{-1}P\right]^{-1}$$

Non-linear!

(could work if Σ_{syst} was small compared to Σ_{stat}, but it's not the case)

- In general, finding a robust bound that is not over-conservative and without making any assumption about the correlations is non-trivial
  - Attempts using e.g. Woodbury identity, but not converged yet

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U\left(C^{-1}+VA^{-1}U\right)^{-1}VA^{-1},$$

- A more realistic approach is to scan only a sub-class of correlations matrices
  - KNTW approach in arXiv:2604.25004: parameterized "global" and "local" decorrelations
  - We can extend this to other or larger classes, if we have good motivations

- But when/how do we consider that we have a representative coverage?

Another practical attempt

- Early attempts we made tried to find the correlations that maximize the post-fit uncertainty
  - Assume some correlations as a starting point
  - Using first order approx, find the corr matrix that maximize $V\left[a_\mu^{\pi\pi}\right]$
  - Re-fit using this new correlation matrix, and iterate until it's stable

$$\mathrm{Corr}(\sigma_i, \sigma_j) = \mathrm{sign}\left(\frac{\partial a_\mu^{\pi\pi}}{\partial \sigma_i} \times \frac{\partial a_\mu^{\pi\pi}}{\partial \sigma_j}\right)$$

- But we observed that the result changes when we change the starting point
  - This stresses out that 1st order approximation fails for this use case (non-linear manifold)
  - Furthermore, we now understand that this is not a good approach due to best-fit shift

- Follow-up, to be checked: find correlations that **min(max)imize the value of $a_\mu^{\pi\pi}$ and take full interval as syst. uncertainty** (incorporates shifting effects)
  - But no guarantee yet that we wouldn't have the same issue...

Goodness of fit issues

- On top of giving **different model constraints**, the different assumptions on correlations give **extremely different goodness of fit** (sometimes absurd)
  - e.g. $\frac{\chi^2}{\text{dof}} \sim 0.5$ in CMD-3 uncorrelated case, while $\frac{\chi^2}{\text{dof}} \sim 1.1$ for 100% correlated case…

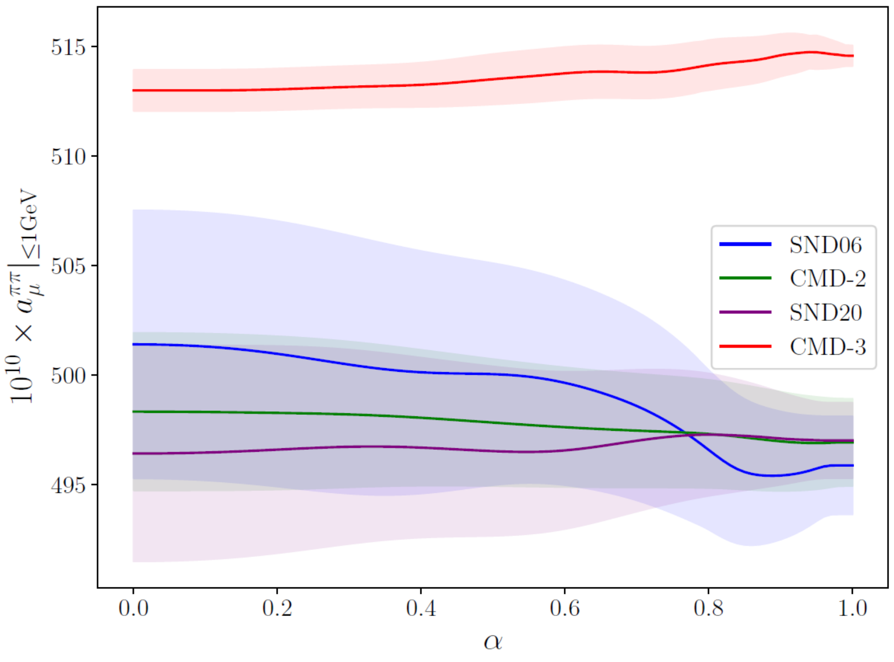
- Thus, defining 1σ region as the enveloppe of all possible cov. matrices would not work (only the edge with best GoF would matter)

  → Difficult to find a prescription inherently **compatible with frequentist approach**…

- Also an issue to determine the **compatibility of different datasets** with unitarity, analiticity constraints, etc.
  - Would it be enough if one set of correlations gives a good GoF to say it's compatible?

Our conclusions so far

- Empirically, uncorrelated and 100% correlated syst. seem to provide nearly maximal coverage, though we cannot provide a solid proof of this
- For CMD-3, decorrelation pulls $a_\mu^{\pi\pi}$ upward, while for most other experiments it pulls $a_\mu^{\pi\pi}$ downward
  $\rightarrow$ Claiming discrepancies with 100% correlations is somewhat conservative



Our wish list

- Nevertheless, guessing what the correlations should be cannot be fully a theorist task
- Experiments should provide all the information they have as much as possible
- For instance, releasing several cov. matrices for each systematic source, and provide correlations for those which are available (e.g. detector effects)

Discussion points

- How can we improve the situation regarding systematic uncertainties?
  - Models for decorrelation
  - Casting information into the form of covariance matrices
- Interplay between systematic and statistical correlations
- Under which conditions can we discard experiments in global compilations?
  SND 20 for 2π, Belle II for 3π
- Can the evidence derived from inelastic channels be made more significant?
- ...