Introduction to the Dispersive HVP Discussion Session

Implications of the existing tensions among the experimental $e^+e^- \to \pi^+\pi^-$ measurements for the dispersive combinations (DHMZ)

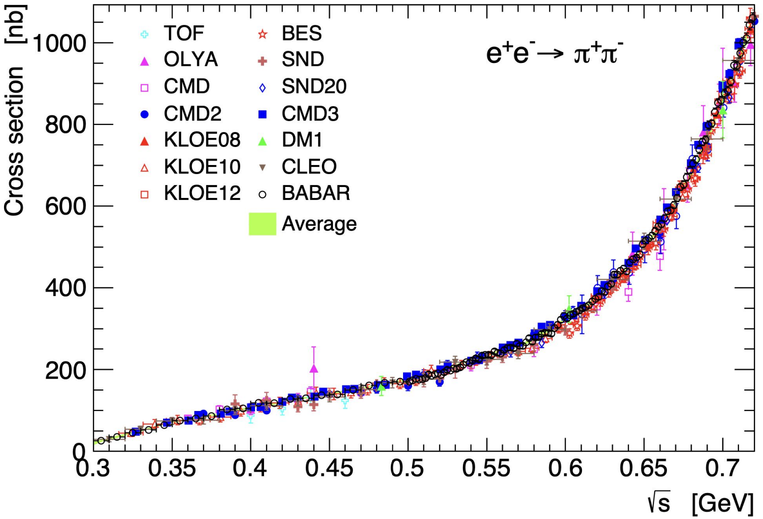
Bogdan MALAESCU

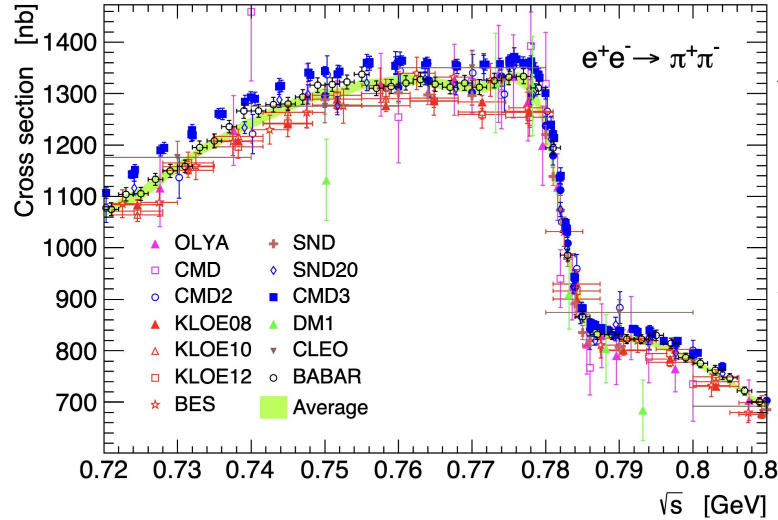
LPNHE, CNRS

6th Plenary Workshop of the Muon g-2 Theory Initiative
University of Bern

05/09/2023

Combining the $e^+e^- \rightarrow \pi^+\pi^-$ data



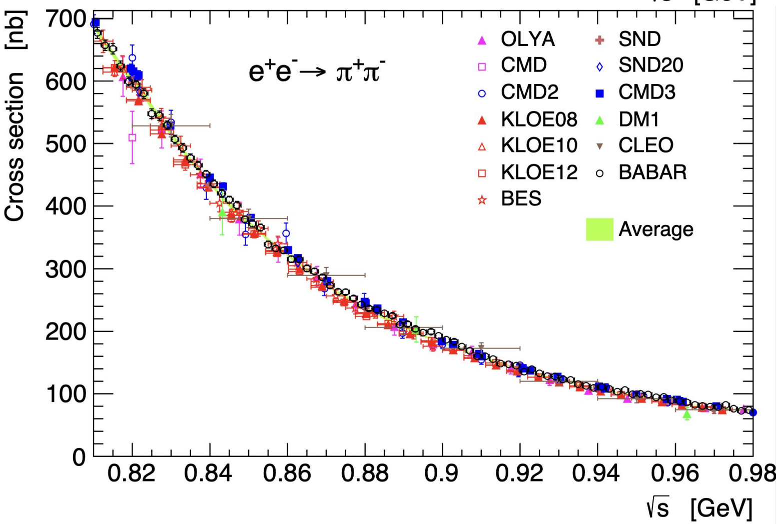


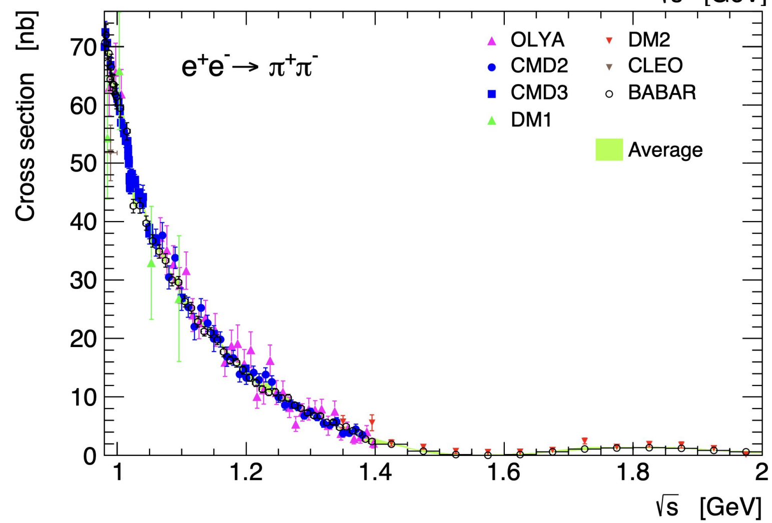
New since WP:

SND20

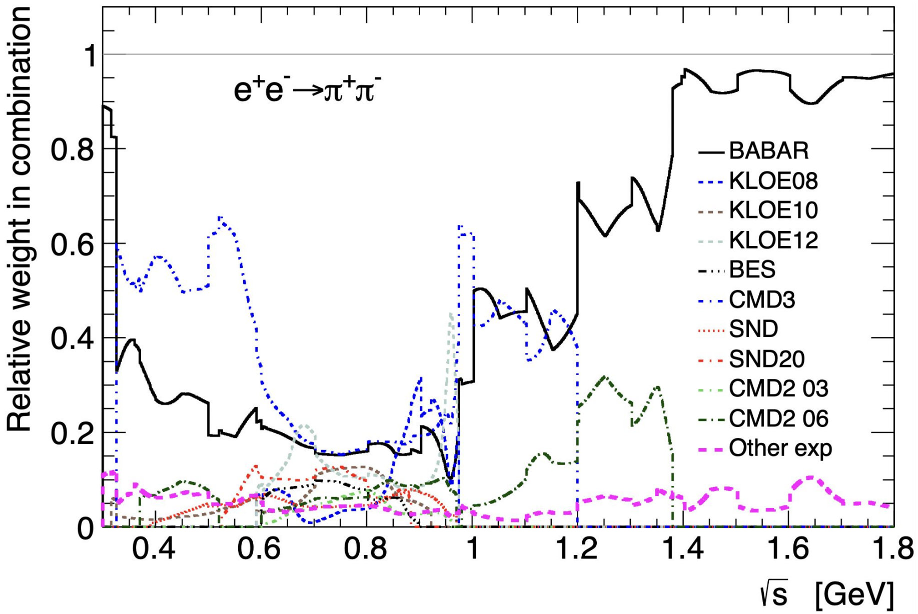
CMD3

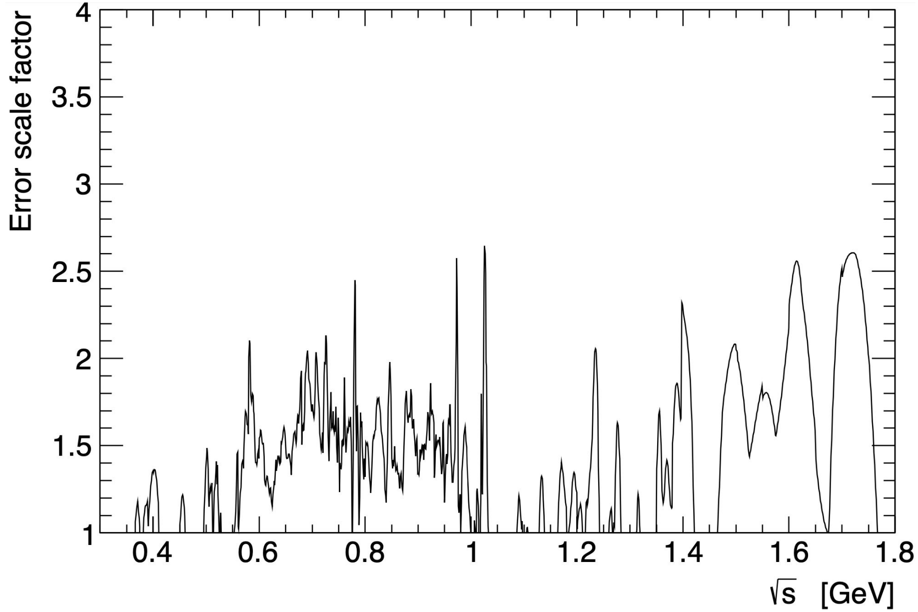
Updated BES cov matrix





Combining the $e^+e^- \to \pi^+\pi^-$ data: weights and tension





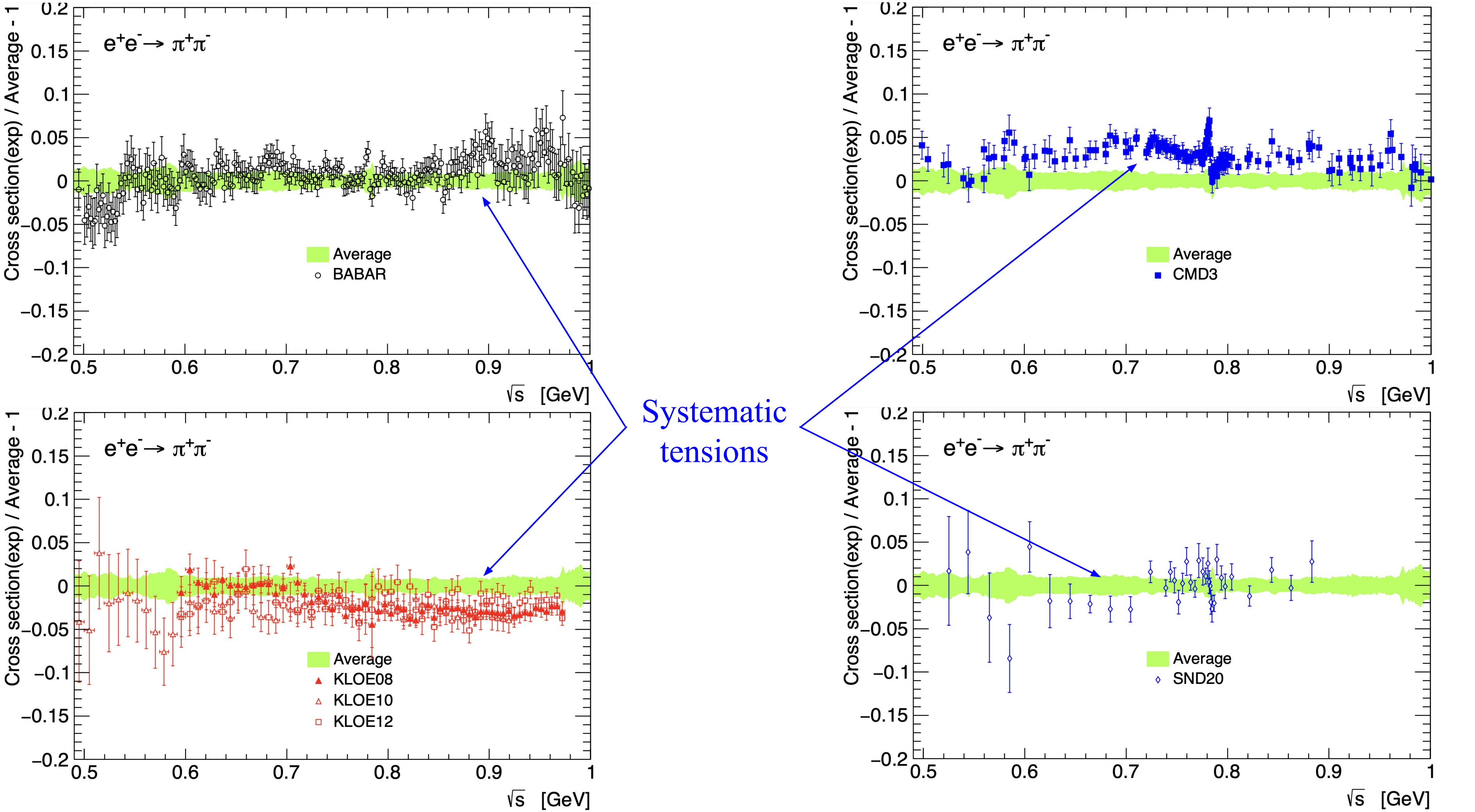
→ The newly added data have important contributions (weights) in the combination

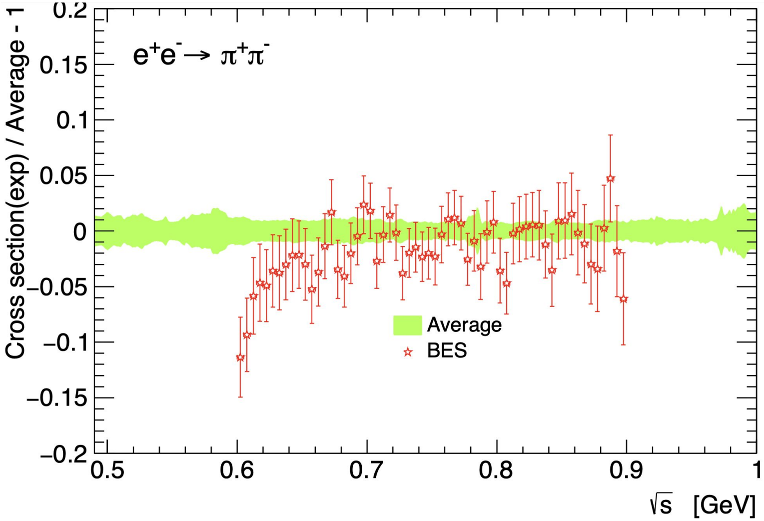
→ Enhanced tensions, especially between KLOE & CMD3, which provide the smallest / largest cross-sections in the ρ region: *yet another indication of underestimated uncertainties*

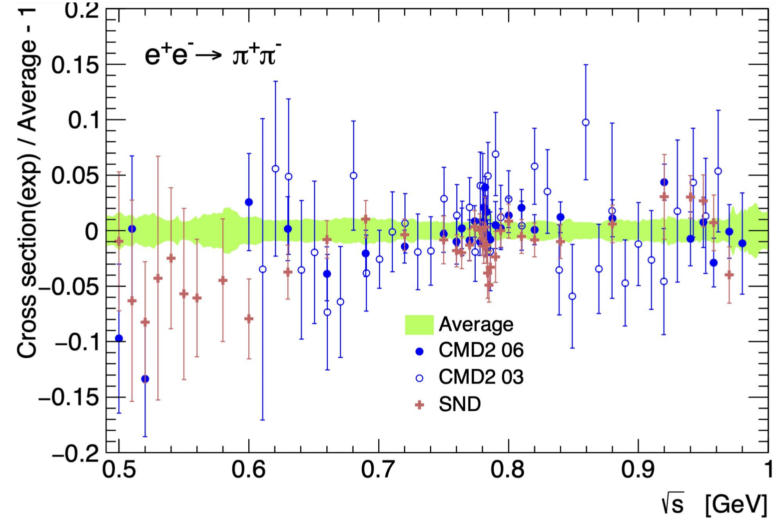
→ *Calls for conservative uncertainty treatment* in combination fit (fits / evaluation of weights)

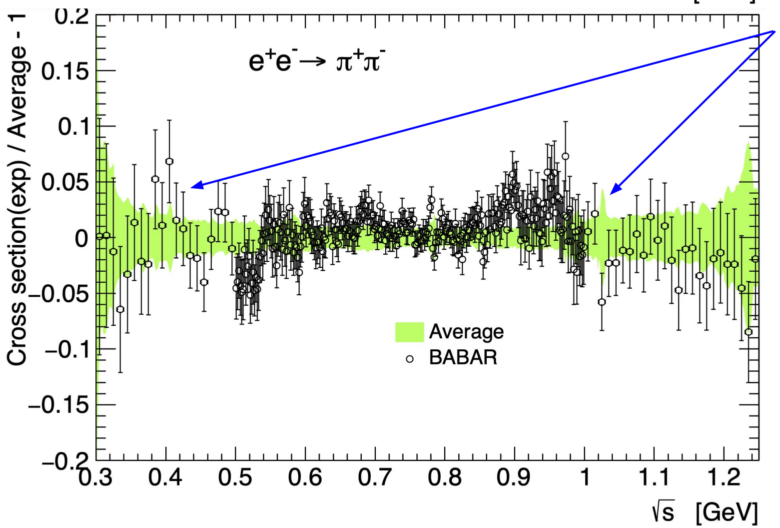
→ Systematic effects beyond the local χ^2 /ndof rescaling: had already motivated the inclusion of the dominant BABAR-KLOE systematic by DHMZ since 2019, but *tensions are larger now*

Combining the $e^+e^- \to \pi^+\pi^-$ data: relative differences

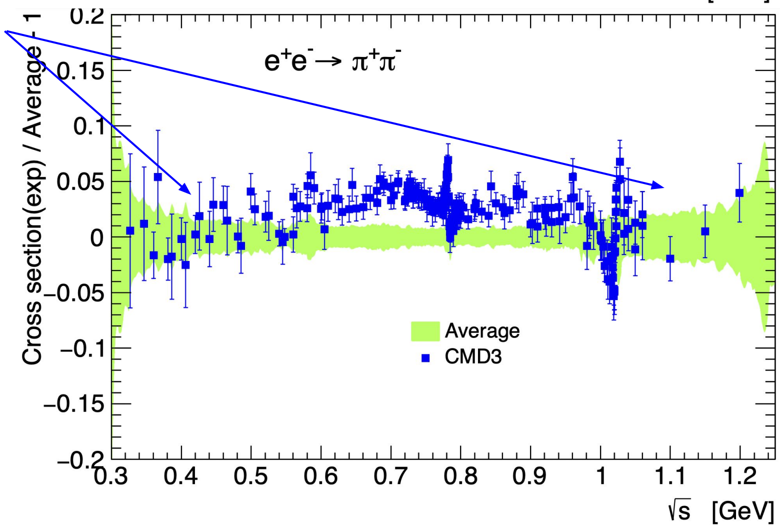








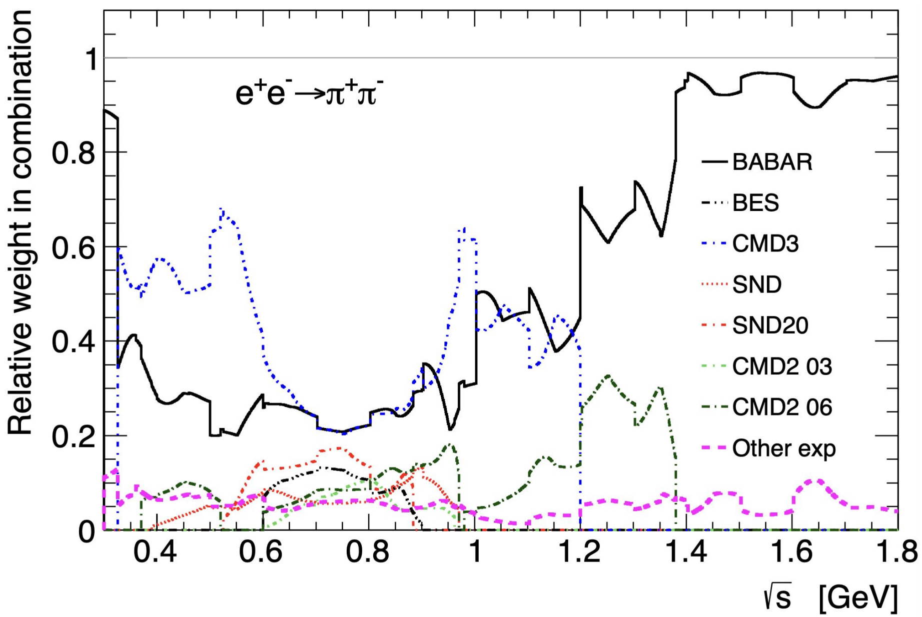
Reasonable BABAR/CMD3 agreement at low & high E

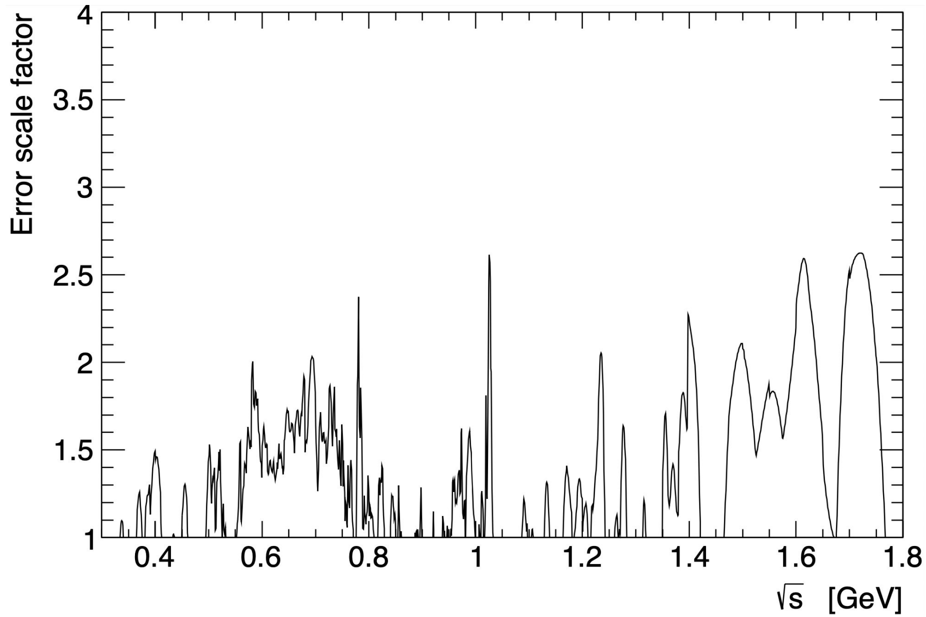


Combining the $e^+e^- \to \pi^+\pi^-$ data, *exercise* without KLOE

→ Motivated by recent NNLO measurement by BaBar and the findings about the description of extra radiation in Phokhara (see talks by M. Davier and Z. Zhang)

→ Probe the hypothesis of a possible impact of "N(N)LO" photons for KLOE, through the use of Phokhara (pending extra studies by KLOE): investigate a possible combination without KLOE

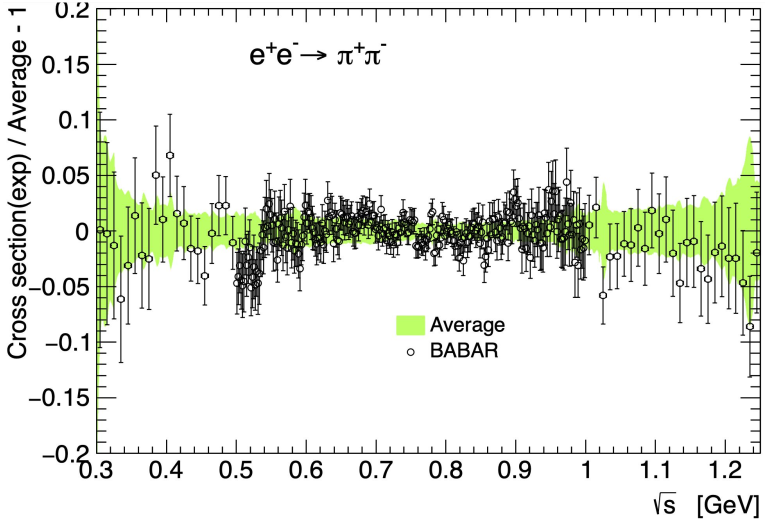


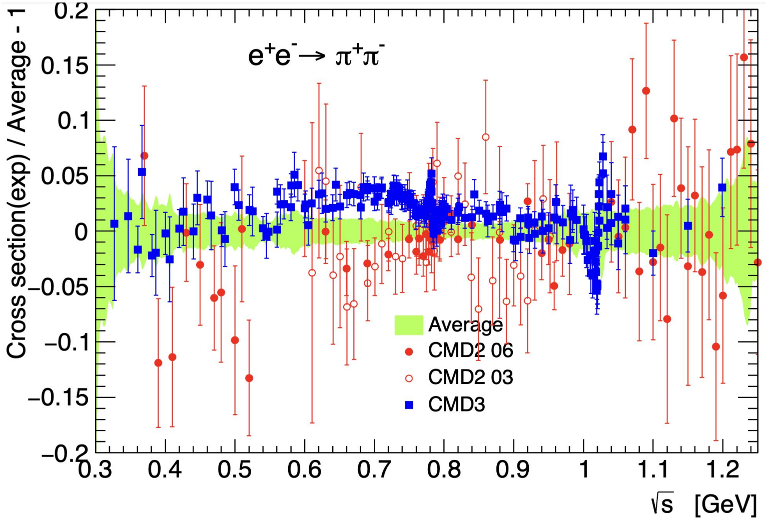


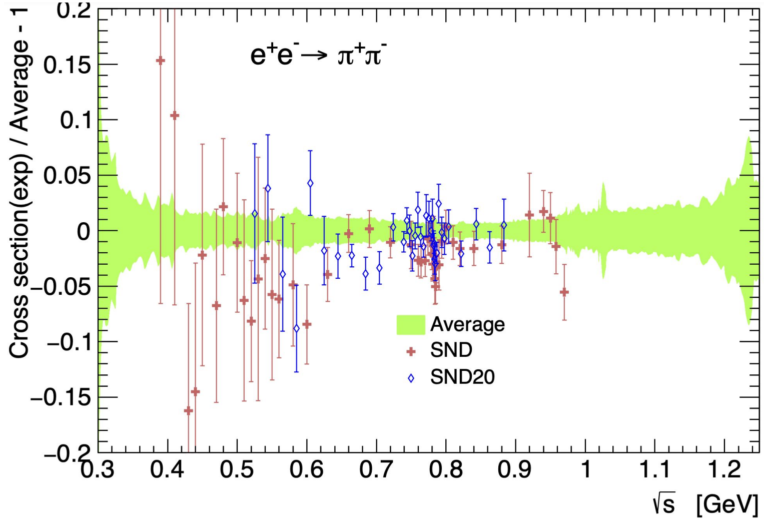
→ Preserved hierarchy of weights for remaining experiments

→ Significantly reduced tension in the region ~0.80 - 0.97 GeV

Combining the $e^+e^- \to \pi^+\pi^-$ data, *exercise* without KLOE







→ BABAR overlaps well with the average, while the larger values of CMD3 in the ρ region are balanced by lower values of CMD2, SND20, SND etc.

→ Systematic tension between CMD2 and CMD3 especially important to clarify

Quantitative comparisons for a_μ^{HVP}

$\times 10^{-10}$

All experiments:
a_μ [0.3 ; 1.8 GeV] = 510.9 ±2.5 (±0.8 (stat) ±2.3 (syst))

Without applying the χ^2/ndof rescaling of uncertainties:
a_μ [0.3 ; 1.8 GeV] = 510.9 ±1.7 (±0.6 (stat) ±1.6 (syst))

Exercise without KLOE:
a_μ [0.3 ; 1.8 GeV] = 513.9 ±2.6 (±1.0 (stat) ±2.5 (syst))

Without applying the χ^2/ndof rescaling of uncertainties:
a_μ [0.3 ; 1.8 GeV] = 513.9 ±2.0 (±0.7 (stat) ±1.8 (syst))

All experiments:
a_μ [0.5251 ; 0.8832 GeV] = 410.9 ±2.0 (±0.7 (stat) ±1.8 (syst))

Without applying the χ^2/ndof rescaling of uncertainties:
a_μ [0.5251 ; 0.8832 GeV] = 410.9 ±1.3 (±0.5 (stat) ±1.2 (syst))

Exercise without KLOE:
a_μ [0.5251 ; 0.8832 GeV] = 413.6 ±2.1 (±0.9 (stat) ±1.9 (syst))

Without applying the χ^2/ndof rescaling of uncertainties:
a_μ [0.5251 ; 0.8832 GeV] = 413.6 ±1.5 (±0.6 (stat) ± 1.4(syst))

→ Including/removing KLOE induces shift of ~3 units for the total integral

→ Precision gain in presence of extra data largely compensated by the local χ^2/ndof rescaling;
In addition, an extra (dominant) uncertainty accounting for systematic deviations between measurements has to be added

Quantitative comparisons for a_μ^{HVP}

→ Comparison of integrals computed on restricted energy ranges, for individual experiments:
significance of the difference between different experiments, taking into account correlations

Δa_μ[0.32698 ; 1.19917 GeV] :
BABAR - CMD3: -11.9 ± 5.6 ; Significance: 2.1 σ

Δa_μ[0.5251 ; 0.8832 GeV] :
SND - KLOE10: 4.3 ± 6.7 Significance: 0.6 σ
SND20 - KLOE10: 7.8 ± 5.4 Significance: 1.4 σ
SND20 - SND: 3.4 ± 6.3 Significance: 0.5 σ
CMD3 - KLOE10: 22.7 ± 5.2 Significance: 4.4 σ
CMD3 - SND: 18.4 ± 6.3 Significance: 2.9 σ
CMD3 - SND20: 15.0 ± 4.9 Significance: 3.1 σ
BABAR - KLOE10: 12.1 ± 5.0 Significance: 2.4 σ
BABAR - SND: 7.9 ± 6.2 Significance: 1.3 σ
BABAR - SND20: 4.4 ± 4.7 Significance: 0.9 σ
BABAR - CMD3: -10.6 ± 4.5 Significance: 2.4 σ

Δa_μ[0.7 ; 0.8 GeV] :
SND20 - CMD2-2004: -0.7 ± 2.5 Significance: 0.3 σ
SND20 - CMD2-2006: 1.2 ± 2.4 Significance: 0.5 σ
SND20 - KLOE08: 4.0 ± 2.4 Significance: 1.7 σ
SND20 - KLOE10: 4.9 ± 2.4 Significance: 2.1 σ
SND20 - KLOE12: 4.8 ± 2.4 Significance: 2.0 σ
SND20 - SND: 1.5 ± 3.2 Significance: 0.5 σ

CMD3 - CMD2-2004: 5.7 ± 2.6 Significance: 2.2 σ
CMD3 - CMD2-2006: 7.6 ± 2.6 Significance: 3.0 σ

CMD3 - KLOE08: 10.4 ± 2.3 Significance: 4.4 σ
CMD3 - KLOE10: 11.2 ± 2.3 Significance: 4.9 σ
CMD3 - KLOE12: 11.1 ± 2.3 Significance: 4.8 σ
CMD3 - BES2pi: 8.6 ± 2.7 Significance: 3.2 σ
CMD3 - SND: 7.8 ± 3.2 Significance: 2.5 σ
CMD3 - SND20: 6.3 ± 2.4 Significance: 2.7 σ

BABAR - CMD3: -5.6 ± 2.3 Significance: 2.5 σ
BABAR - SND20: 0.7 ± 2.3 Significance: 0.3 σ
BABAR - KLOE08: 4.7 ± 2.3 Significance: 2.1 σ
BABAR - KLOE10: 5.6 ± 2.3 Significance: 2.5 σ
BABAR - KLOE12: 5.5 ± 2.3 Significance: 2.4 σ

→ Largest tensions between CMD3 and KLOE

→ Important to clarify tension between CMD3 and CMD2

Comparison of / consequences for combination methods

Analysis aspect	DHMZ	KNT
Blinding	Not necessary (No ad-hoc choices to make)	Included for upcoming update
Binning	Fine ($\leq$ 1 MeV) final binning for average and integrals. Large (O(100 MeV) or less) common binning @ intermediate step: compare statistics of experiments coherently for deriving weights in fine bins.	Re-bin data into "clusters". Scans over cluster configurations for optimisation.
Closure test	Using model for spectrum: negligible bias. (since 2010)	Not performed
Additional constraints	Analyticity constraints for 2π channel.	None
Fitting	χ^2 minimisation with correlated uncertainties incorporated locally (in fine & large bins), for deriving weights. Full propagation of uncertainties & correlations.	χ^2 minimisation with correlated uncertainties incorporated globally.
Integration / interpolation	Av. of quadratic splines (3rd order polynomial), integral preservation in bins of measurements. Analyticity-based function for 2π ($<$ 0.6 GeV).	Trapezoidal for continuum, quintic for resonances.
Uncertainty inflation	Local χ^2 uncertainty inflation. (since 2009) Extra BABAR-KLOE systematic. (since 2019)	Local χ^2 uncertainty inflation. (adopted since 2017)
Inter-channel correlations	Taken into account. (since 2010)	Not included.
Missing channels	Estimated based on isospin symmetry. (since 1997 - ADH)	Adopted in subsequent updates

→ Large DHMZ/KNT differences for the resulting uncertainties, as well as for the shapes of the combined spectra (backup)

WP TI	DHMZ19	KNT19
$a_\mu^{\text{HVP, LO}} \times 10^{10}$	694.0(4.0)	692.8(2.4)

→ CHS approach for 2π and 3π: Analyticity and global χ^2 fit (See talk by Peter Stoffer)

Points for discussion

→ Questions on the combination results/exercise above ?

(→ Discussions / checks / open points for CMD3 discussed yesterday)

→ Extra ISR/FSR photons in data & simulation: impact in the context of the observed tensions

→ Enhanced evidence for uncertainties on uncertainties

→ …

Backup

Combine cross section data: goal and requirements

→ Goal: combine experimental spectra with arbitrary point spacing / binning

→ Requirements:

- Properly propagate uncertainties and correlations

  \- *Between measurements (data points/bins) of a given experiment*

  (covariance matrices and/or detailed split of uncertainties in sub-components)

  \- *Between experiments* (common systematic uncertainties, e.g. VP)

  based on detailed information provided in publications

  \- *Between different channels* – motivated by understanding of the meaning of systematic uncertainties and identifying the common ones

  BABAR luminosity (ISR or BhaBha), efficiencies (photon, Ks, Kl, modeling);
  BABAR radiative corrections; $4\pi 2\pi^0 - \eta\omega$
  CMD2 $\eta\gamma - \pi^0\gamma$; CMD2/3 luminosity; SND luminosity;
  FSR; hadronic VP (old experiments)

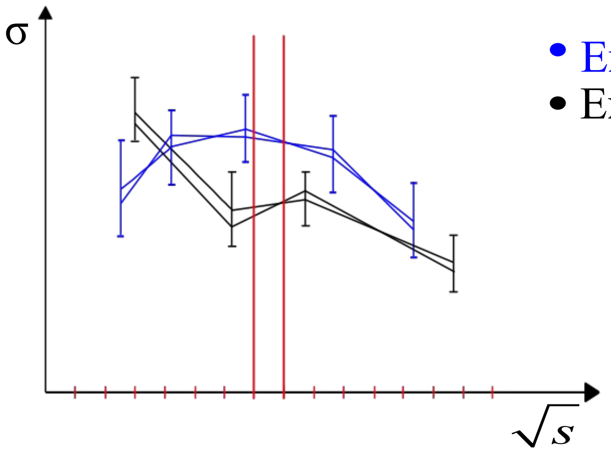
  (1^{st} motivation for using DHMZ uncertainties as "baseline" in the g-2 TI White Paper)

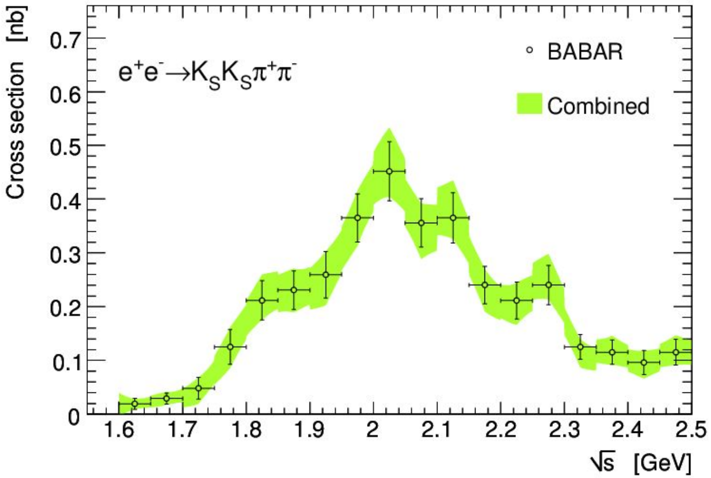
- Minimize biases

- Optimize g-2 integral uncertainty

  (*without overestimating the precision with which the uncertainties of the measurements are known*)

Combination procedure implemented in HVPTools software





→ Define a (fine) final binning (to be filled and used for integrals etc.)

→ Linear/quadratic splines to interpolate between the points/bins of each experiment
- for binned measurements: preserve integral inside each bin
- closure test: replace nominal values of data points by Gounaris-Sakurai model and re-do the combination
  → *(non-)negligible bias for (linear)quadratic interpolation*

→ Fluctuate data points taking into account correlations & re-do the splines for each (pseudo-)experiment
- each uncertainty fluctuated coherently for all the points/bins that it impacts
- eigenvector decomposition for (statistical) covariance matrices

Combination procedure implemented in HVPTools software

For each final bin:

→ Compute an average value for each measurement and its uncertainty

→ Compute correlation matrix between experiments

→ Minimize χ^2 and get average coefficients (weights)

→ Compute average between experiments and its uncertainty

Evaluation of integrals and propagation of uncertainties:

→ Integral(s) evaluated for nominal result and for each set of toy pseudo-experiments;
uncertainty of integrals from RMS of results for all toys

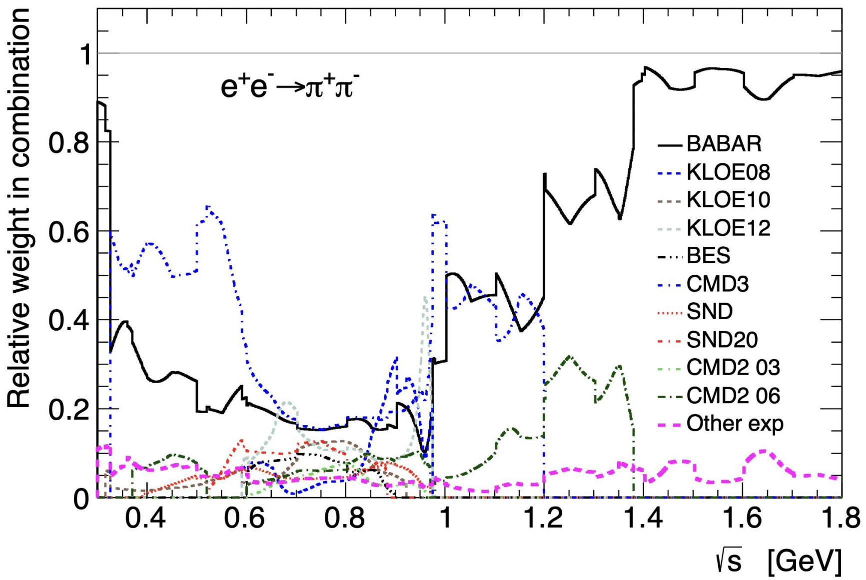
→ The pseudo-experiments also used to derive (statistical & systematic) covariance matrices of combined cross sections → Integral evaluation

→ Uncertainties also propagated through $\pm 1\sigma$ shifts of each uncertainty:
- allows to account for correlations between different channels (for integrals and spectra)

→ *Checked consistency between the different approaches*

For each final bin:
→ Minimize χ^2 and get average coefficients

Note: average weights must account for bin sizes / point spacing of measurements
(do not over-estimate the weight of experiments with large bins)
→ weights in fine bins evaluated using a common (large) binning for measurements + interpolation
→ compare the precisions on the same footing



→ Bins used by KLOE larger than the ones by BABAR in ρ-ω interference region (factor ~3)

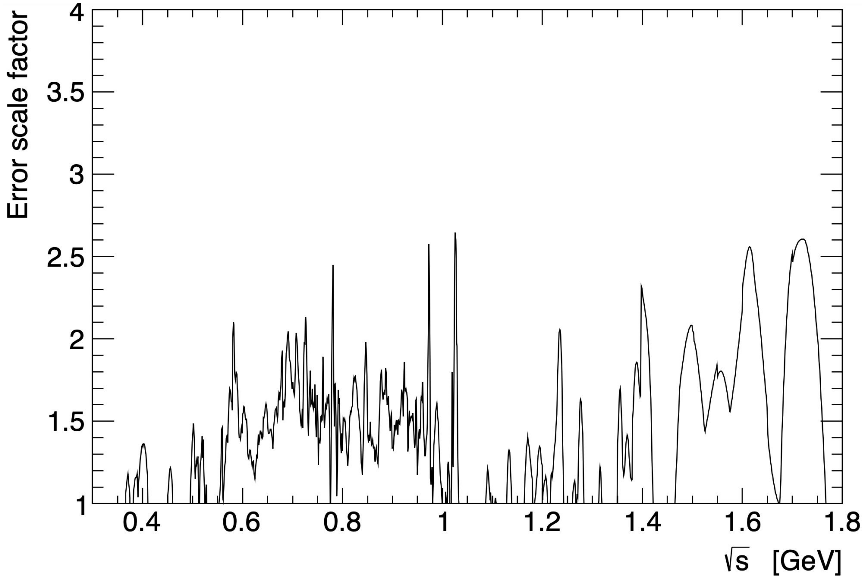
→ Average dominated by BaBar, CMD3 KLOE, SND20
BaBar covering full range

Combination procedure: compatibility between measurements

For each final bin:

→ χ^2 /ndof: test locally the level of agreement between input measurements, *taking into account the correlations*

→ Scale uncertainties in bins with χ^2 /ndof > 1 (PDG): *locally* conservative; Adopted by KNT since '17



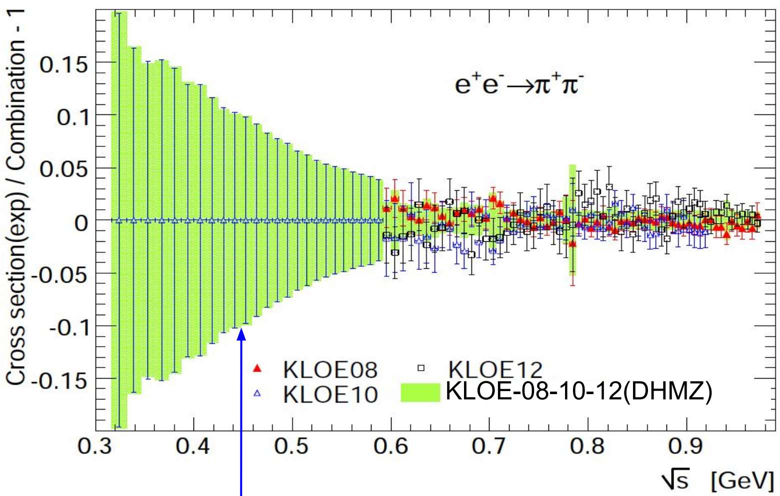
→ Tension between measurements: *indication of underestimated uncertainties* Motivates conservative uncertainty treatment in combination fit (evaluation of weights)

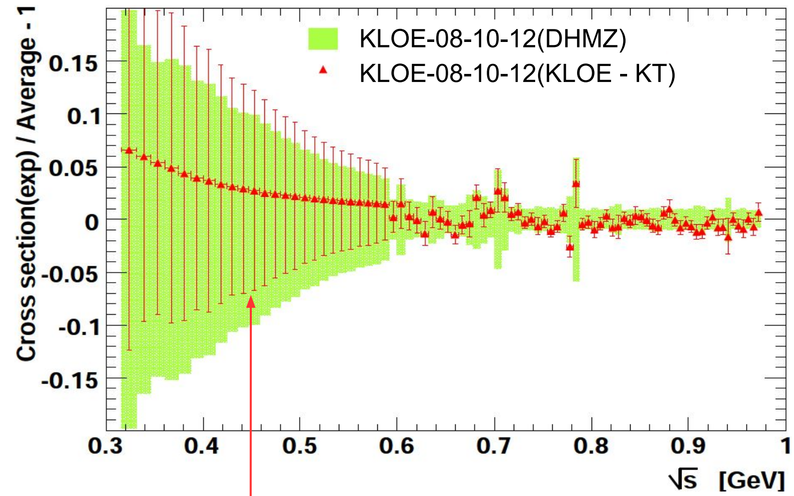
→ Observed (systematic) tension between measurements

→ (*Since 2019*) Included extra (dominant) uncertainty: 1/2 difference between integrals w/o either BABAR or KLOE *(2nd motivation for using DHMZ uncertainties as "baseline" in the TI WP)*

Extra uncertainty starts to be adopted in other studies (2205.12963)

Combining the 3 KLOE measurements





Local combination (DHMZ)

Information propagated between mass regions, through shifts of systematics - relying on correlations, amplitudes and shapes of systematics (KLOE-KT)