

DataBase for STRONG-2020



Alberto Lusiani
Scuola Normale Superiore and INFN, sezione di Pisa



PrecisionSM meeting, 3 June 2020

STRONG-2020 goals (my view)

- ▶ low energy precision data (cross-sections for muon $g-2$ HVP, HLBL, possibly tau spectral functions)
 - ▶ identify relevant useful experimental data
 - ▶ upload into a public DB: [HEPData.net](https://hepdata.net)
 - ▶ verify that data are correct and complete
- ▶ provide example code to download and elaborate the data
- ▶ elaborate code for dedicated studies, e.g. muon $g-2$ HVP contribution
- ▶ make code public on github.com
- ▶ publish documentation on a dedicated web site (github.io)

HEPData.net

- ▶ HEP data public storage web site, mainly used by LHC experiments
- ▶ well defined submission data format, functionally adequate for our data of interest
- ▶ uses YAML, can export to Root format and other formats
- ▶ links inspirehep.net publication, updates inspirehep.net with reference to HEPData.net
- ▶ it is possible to download data and metadata programmatically
- ▶ note: inherits an old realization where some data were harvested on inspirehep.net
 - ▶ must verify that past submissions are correct and complete for our relevant data

data submission

- ▶ coordinator has privileges for submission (A.L. is STRONG-2020 coordinator)
- ▶ coordinator appoints **uploader** and **reviewer**, giving appropriate permissions
- ▶ Stefan Mueller has prepared a working submission of data provided by a KLOE paper
 - ▶ tested OK for me
 - ▶ noticed some too heavy CPU load when loading the web pages, origin not clear to me
 - ▶ will soon proceed to complete submission

Code repository, web site

github.com

- ▶ Microsoft appears to be strongly committed to opensource
- ▶ github free functionality is very good and Microsoft extended it (unlimited private repositories)
- ▶ heavily used by HEP
- ▶ free web sites on github.io, can be versioned as well
- ▶ high quality collaborative functionality
- ▶ possible alternative could be gitlab.cern.ch (with CERN agreement)

web site on github.io

- ▶ collaborative web site with versioning with static web site generators
 - ▶ e.g. Hugo, Jekyll, **Nikola**
- ▶ edit and test web pages locally on your PC, then commit to github
- ▶ collaboration proceeds similarly as with code
- ▶ updating and publishing a web site is straightforward also for non-experts after proper setup

Example of published notebook with code example

<https://alusiani.github.io/webnotes/strong-2020-jun20.html>

BaBar $\sigma(e^+e^- \rightarrow \pi^+\pi^-(\gamma))$

The latest BaBar measurements are published in two papers, a PRL and a later PRD containing more detailed information. Both papers report the cross-section information in the supplemental material, in ASCII files that are identical.

- B. Aubert et al. [BaBar Collaboration], [Phys. Rev. Lett. 103 \(2009\) 231801](#), [inspirehep](#), "Precise measurement of the $e^+e^- \rightarrow \pi^+\pi^-(\gamma)$ cross section with the Initial State Radiation method at BABAR"
- J. P. Lees et al. [BaBar Collaboration], [Phys. Rev. D 86 \(2012\) 032013](#), [inspirehep](#), "Precise Measurement of the $e^+e^- \rightarrow \pi^+\pi^-(\gamma)$ Cross Section with the Initial-State Radiation Method at BABAR"
 - [data](#)

The data report the **"bare cross section including FSR"** in nb, and in detail:

- the cross-section and its total uncertainty in variable-width bins of energy
- the per-mil relative systematic uncertainty (per energy bin, 100% correlated on all bins)
- the statistical correlation between any two bins of cross-section

In the following the data are used to show a few plots.

```
In [1]: using DataFrames
using TextParse
using CSV

using FileIO

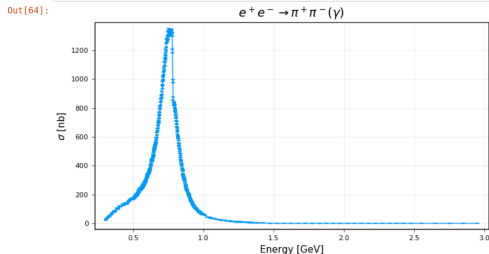
## using Plots
using StatsPlots

using LaTeXStrings
```

Example of published notebook with code example

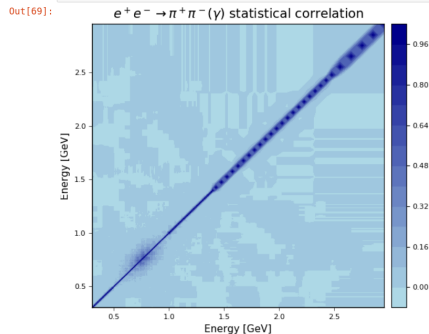
cross-section

```
In [64]: pl1 = @df df plot(
    :E,
    :sigma_val,
    yerror=:sigma_unc,
    title=L"Se+e- \rightarrow \pi^+\pi^-(\gamma)*",
    xlabel="Energy [GeV]",
    ylabel=L"\sigma [nb]",
    markerstrokecolor = :auto,
    legend = false
)
## savefig(pl1, "plot.pdf")
## display(pl1)
```



correlation

```
In [69]: ##
## plot statistical correlation
##
pl3 = @df df contourf(
    :E, :E, stat_corr,
    ## xlims = cov_xlims,
    color = :blues,
    title=L"Se+e- \rightarrow \pi^+\pi^-(\gamma) statistical correlation",
    xlabel="Energy [GeV]",
    ylabel="Energy [GeV]",
    size=(575, 500)
)
```



Next steps

- ▶ prepare a list of relevant measurements to be uploaded
- ▶ solicit collaborations to upload their data, collaborate in the uploading
 - ▶ beware: submitting to HEPData.net is a somewhat non-trivial and demanding work
 - ▶ should build momentum to promote as HEP standard the data sharing on HEPData.net
 - ▶ note: only few relevant collaborations have HEPData coordinators (BaBar, BESIII)
- ▶ setup collaborative web site for STRONG-2020 on github.io and github repository for code
- ▶ work on code for examples and studies