Theoretical Physics Lab-Course 2021S
University of Vienna

Unbiased Fitting

Milutin Popovic
Tim Vogel

Supervisor: Peter Stoffer

April 18, 2021

Contents

1 Abstract

2 Introduction and Motivation

In physics, we often come across the situation, where measured data needs to be approximated by a theoretical model, which requires a set amount of parameters. To determine these parameters a so called "data fit" is required, which can be done via different techniques. One of the most widely used fit-thecnique is "Least-squares fitting", but as soon as we deal with correlated data points, which means, each data point is not a completely independent measurement, a wrong application of the Least-squares fit, can lead to a bias, which, in return, will affect the accuracy of the fit in a negative way. This so called D'Agostini bias, albeit a very situational phenomenon, has to be considered when dealing with correlated data from one or even more experiments. It can be avoided by implementing a iterative fit method and we will consider this in an example from particle physics.

2.1 Motivation

3 Physical description and Findings

3.1 The Vector Form Factor of Pions

In particle physics, one of the best to study reactions of elementary particles, is the collision between an electron (e^-) and it's anti-particle, the positron (e^+). When these two particle collide, they annihilate each other and produce new types of particles. In these experiments very precise measurements can be taken and as such, be a very valuable base of empirical data of the Standard model of physics. A central point of study, of these electron-positron-collisions has been the anomalous magnetic moment g-2 of the muon. The anomaly of this number comes from the fact, that the measured data differs to the theoretical model by quite a large margin. As such it could be the source of exciting discoveries. The theoretical value of the g-2 momentum relies on data from the aforementioned collisions, which is used to reconstruct the so called hadronic vacuum polarization. The hadronic vacuum polarization itself comes from the hadronic final states. About 70% of the contribution to the g-2 momentum comes from the annihilation of an electron and a positron into two pions. The probability of this happening is dependent on the energy of the two particles. The strong interaction between these two pions is given by the so called pion vector form factor (VFF, F_π^V).

3.2 The D'Agostini bias

The D'Agostini bias was first introduced by Giuilo D'Agostini in 1994. It describes a problem with data-fits, when considering data with overall systematic errors, that share a uncertainty on the normalization factor. In such a situation, if the error matrix V of the data points is known, one would normally minimize the χ^2, which can be obtained by $\chi^2 = \vec{\Lambda}^T \cdot V^{-1} \cdot \vec{\Lambda}$. In this formula, Λ denotes the vector between the values of the theoretical model and the measured ones. But, after carrying out such a fit, one often obtains results, which contradict expectations. For example, if we got the results $8.0 \pm 2\%$ and $8.5 \pm 2\%$, from a measurement, which share a 10% normalization error, if we minimized the χ^2 as described-with the matrix V estimated by the data, we would obtain the value 7.87 ± 0.81. This result should immediately take attention, as the result with the highest probability, lies outside the range of the measured values. This error also occurs in a situation, where data is taken from two or more independently conducted experiments, which are afflicted by an additional systematic normalization error, even though the dimensions of the error are not quite as severe as in the situation described before.

3.3   Iterative solution to the D'Agostini bias

3.4   Code Structure

Here the logical structure of the code is shown.

- Construct statistical covariance matrix
- Construct Jacobi matrix of model function in terms of the parameters
- Guess initial parameters $\vec{p}_0$
- Iterate
    - Construct System covariance matrix with $\vec{p}_i$
    - Fill Jacobi matrix with $\vec{p}_i$ which is the Design Matrix
    - Claculate step $\delta \vec{p}_i$
    - update initial parameters

$$\vec{p_{i+1}} = \vec{p}_i + \alpha \cdot \delta \vec{p}_i \quad \alpha = 0.1$$

- Calculate errors
- Calculate χ^2_{min}

The guess used for all fits was determined by standard least-square fit provided by `scipy`.

$$\vec{p}_0 = \big(0.9,\ 0.2,\ 0.81,\ 0.04,\ 0.02,\ -1,\ 0.84,\ 1.55\big) \tag{1}$$

The code can be viewed and/or downloaded from here [1] (including the calculation of the guess parameters).

4   Results

In this section the results with consideration of the D'Agostini are shown. Furthermore the findings with fits of two experiments together (6 combinations of two) and also a fit with all experiments are shown. In the end of the section the fitted parameters are compared with the literature values. The plots of the given data and fitted models can be found in Section 4.3.1.

4.1 Single Experiment Fits under consideration of the D'Agostini bias

In this section the data is fitted under consideration of the D'Agostini bias of all experiments separately.

Table 1: Results of all experiment data fitted separately

$\vec{p}$	SND	CMD2	KLOE	BABAR
M_ρ	772.72 ± 0.59	773.93 ± 0.67	773.91 ± 0.25	773.33 ± 0.43
Γ_ρ	149.53 ± 1.15	147.67 ± 1.32	149.72 ± 0.37	149.19 ± 0.81
M_ω	781.94 ± 0.09	782.32 ± 0.07	782.44 ± 0.11	782.18 ± 0.07
Γ_ω	8.55 ± 0.33	8.65 ± 0.44	9.66 ± 0.33	8.17 ± 0.16
ε_ω	2.02 ± 0.09	1.92 ± 0.12	2.07 ± 0.05	1.95 ± 0.03
χ^2_{min}/dof	1.001	1.054	1.443	1.031
p-value	0.530	0.395	0.001	0.377

4.2 Multi Experiment Fits under consideration of the D'Agostini bias

In this section the data is fitted considering the D'Agostini bias, first the data of two experiments together then the data of all experiments.

Table 2: Results of data fits of experimental data fitted in pairs

$\vec{p}$	SND-CMD2	SND-KLOE	SND-BABAR
M_ρ	772.72 ± 0.42	773.92 ± 0.23	773.17 ± 0.36
Γ_ρ	149.53 ± 0.81	149.42 ± 0.35	149.70 ± 0.64
M_ω	781.95 ± 0.07	782.39 ± 0.07	782.07 ± 0.06
Γ_ω	8.56 ± 0.24	9.42 ± 0.20	8.27 ± 0.13
ε_ω	2.02 ± 0.07	2.07 ± 0.05	1.96 ± 0.03
χ^2_{min}/dof	0.904	1.839	0.945
p-value	0.754	0.001	0.763

Table 3: Results of data fits of experimental data fitted in pairs

$\vec{p}$	CMD2-KLOE	CMD2-BABAR	KLOE-BABAR
M_ρ	773.92 ± 0.23	773.17 ± 0.36	773.66 ± 0.20
Γ_ρ	149.42 ± 0.35	149.70 ± 0.64	149.41 ± 0.32
M_ω	782.39 ± 0.07	782.07 ± 0.06	782.49 ± 0.06
Γ_ω	9.42 ± 0.02	8.27 ± 0.13	8.98 ± 0.12
ε_ω	2.07 ± 0.05	1.96 ± 0.03	1.98 ± 0.02
χ^2_{min}/dof	1.838	0.943	1.470
p-value	0.001	0.772	0.001

4.3 Litrature comparison

In the following table 5, we took the results of all experimental data fitted together in table 4 and compared them with the literature values [2].

Table 4: Results of data fit of all experimental data fitted together

M_ρ	773.62 ± 0.18
Γ_ρ	149.42 ± 0.29
M_ω	782.36 ± 0.08
Γ_ω	8.75 ± 0.08
ε_ω	1.96 ± 0.02
χ^2_{min}/dof	1.735
p-value	0.000

Table 5: Result comparison with literature

M_ρ	775.26 ± 0.25
Γ_ρ	147.8 ± 0.9
M_ω	782.65 ± 0.12
Γ_ω	8.49 ± 0.08

4.3.1 Plots
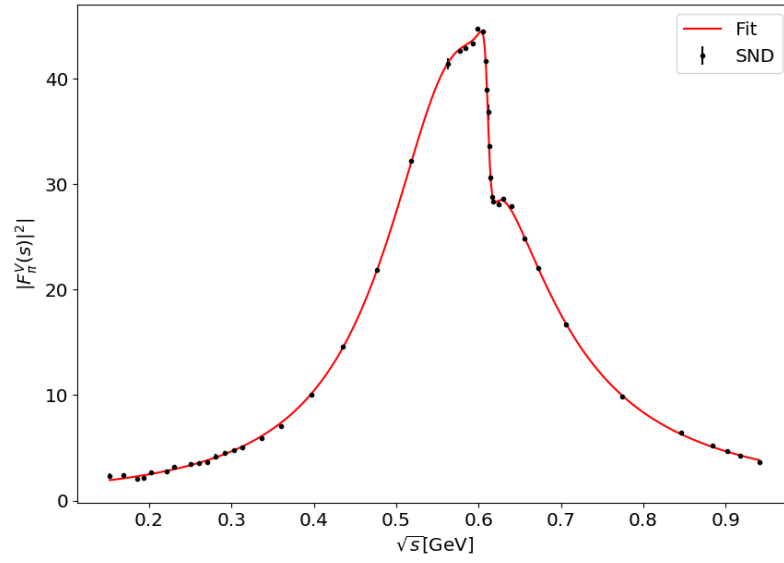
Figure 1: SND data fit
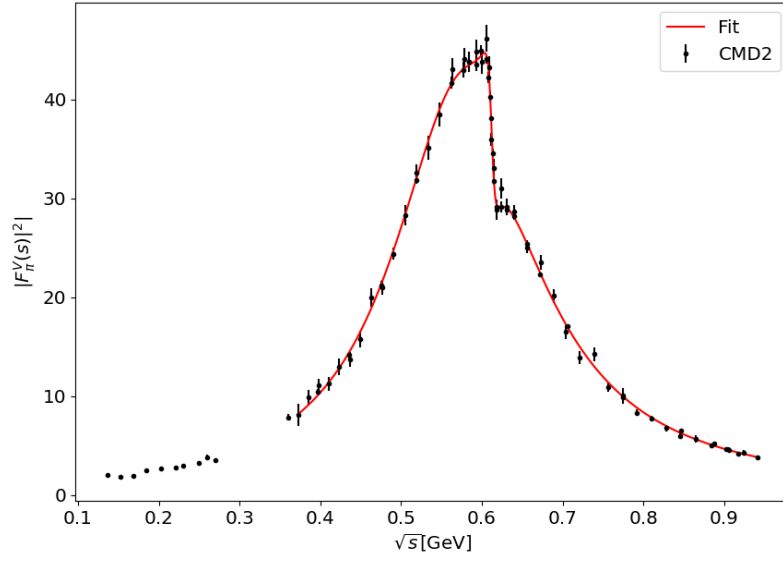
Figure 2: CMD2 data fit
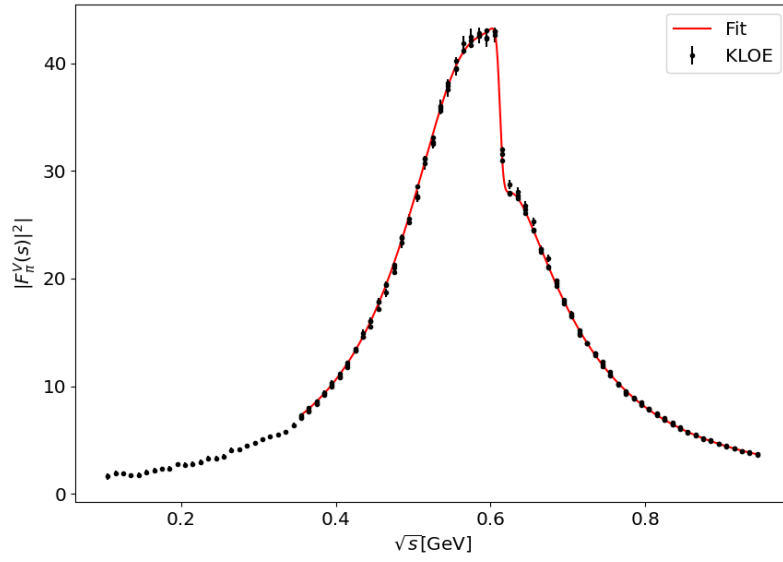
Figure 3: KLOE data fit
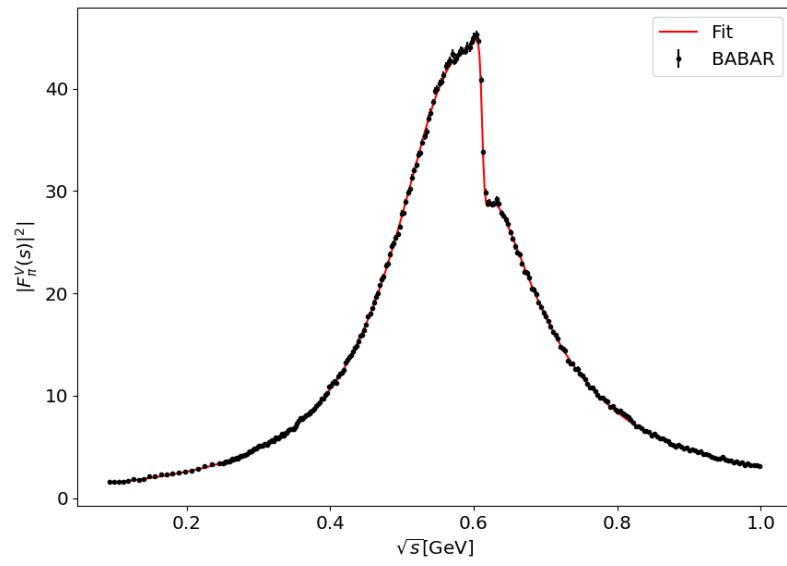
Figure 4: BABAR data fit
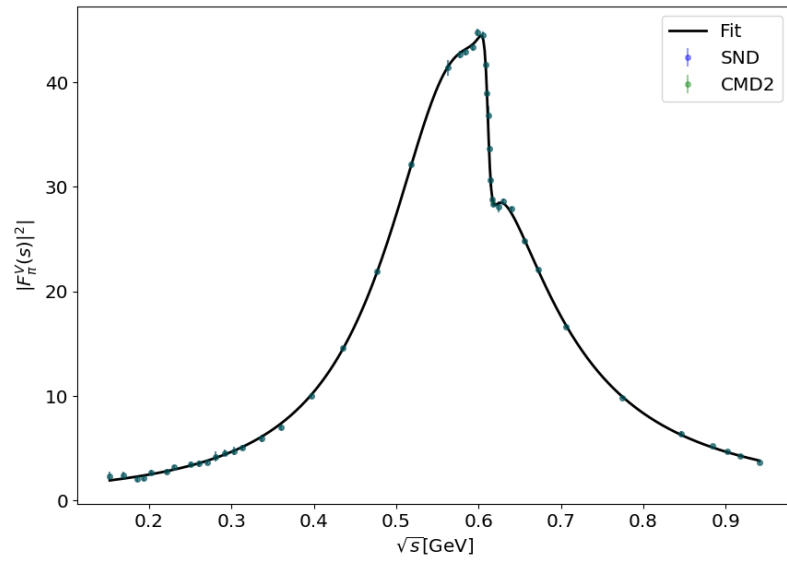
Figure 5: SND and CMD2 fitted togther
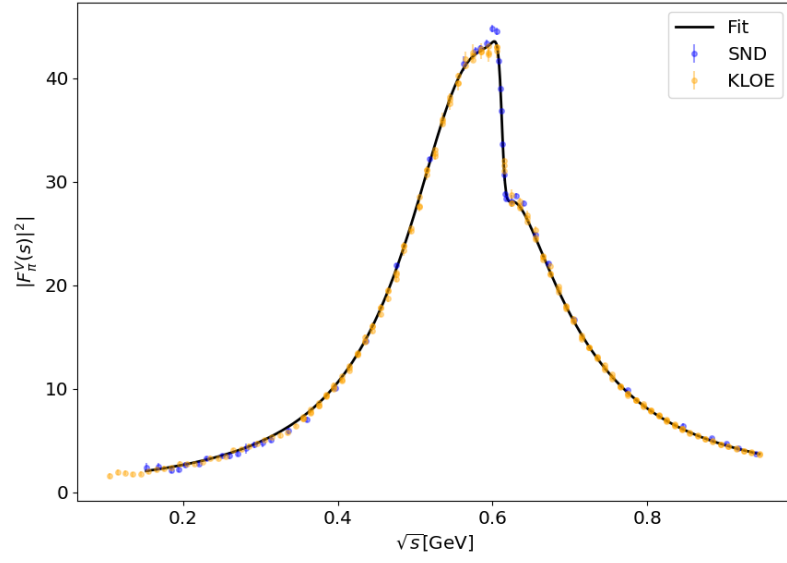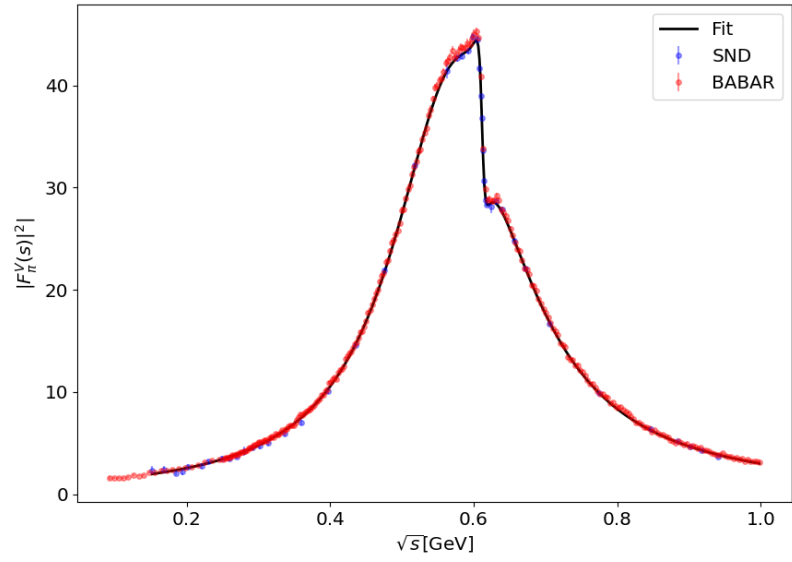
Figure 6: SND and KLOE fitted togther
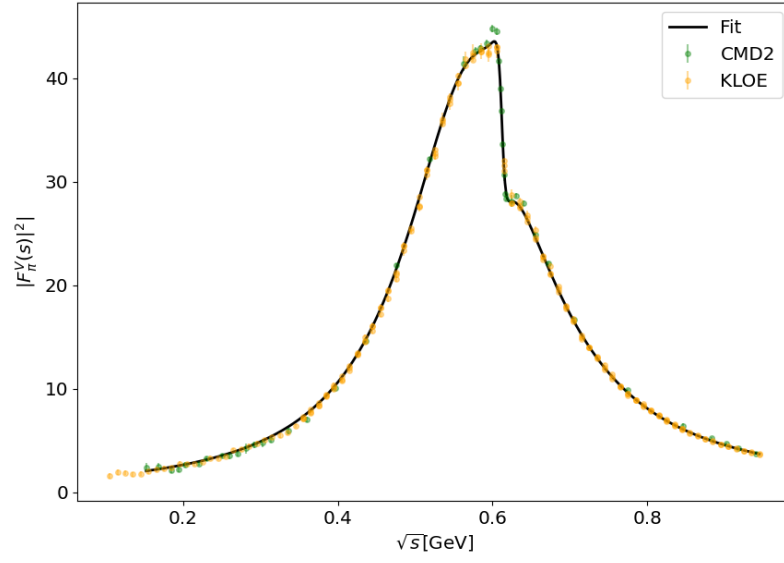

Figure 7: SND and BABAR fitted togther
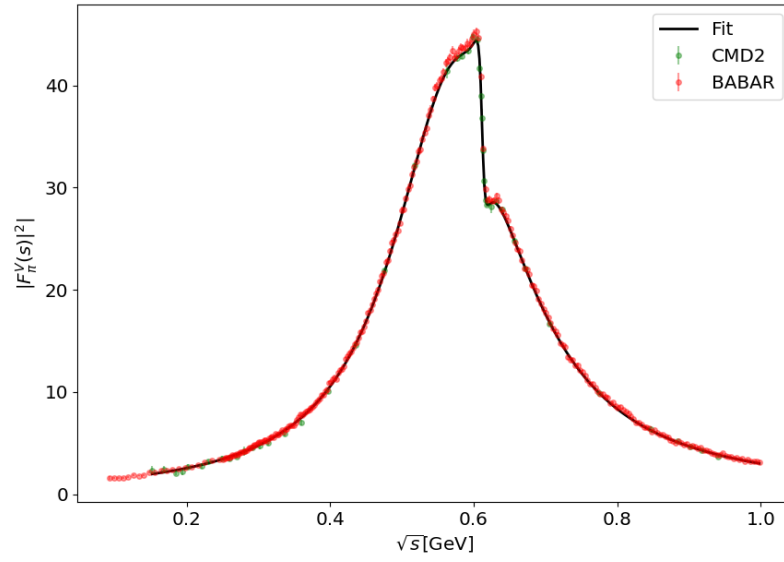
Figure 8: CMD2 and KLOE fitted togther
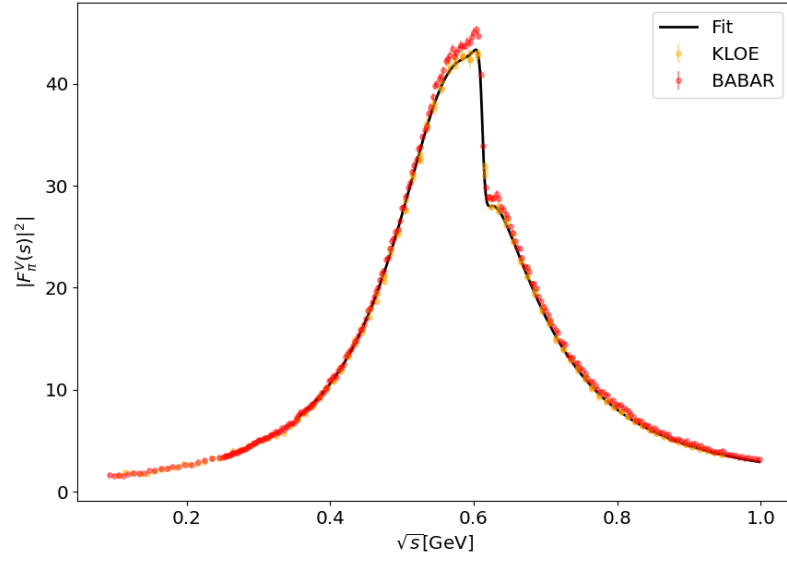

Figure 9: CMD2 and BABAR fitted togther
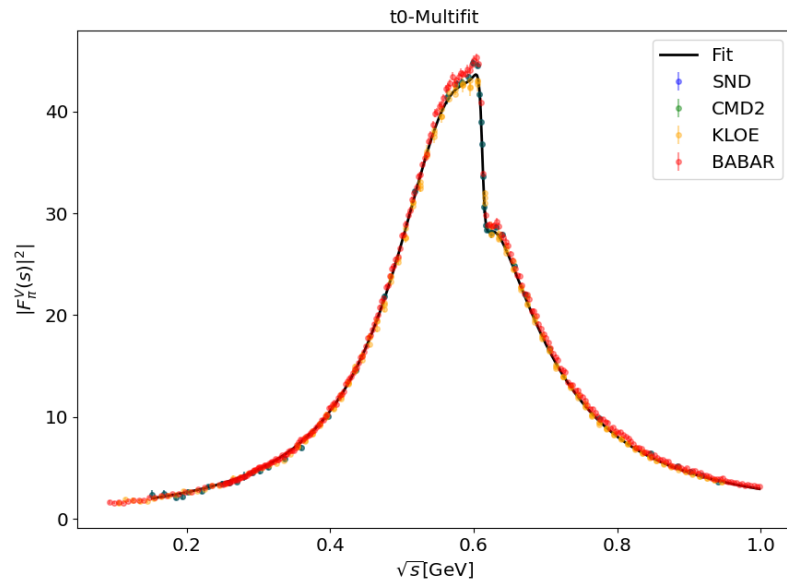
Figure 10: KLOE and BABAR fitted togther

Figure 11: SND, CMD, KLOE and BABAR fitted togther

References

[1] *Git Instance, Implementation of the T0-Method*. 2021. URL: git://popovic.xyz/tprak.git (visited on 04/17/2021).

[2] Particle Data Group. *Particle Data Group*. 2020. URL: https://pdg.lbl.gov/ (visited on 04/17/2021).