Proposal for B-meson decay analysis

Pere Masjuan

April 21, 2020

1 State-of-the-art

Current analysis of anomalies in flavour physics are based on a linear regression of a χ^2 function. After taking into account correlation between theoretical predictions and experimental observables, a χ^2 function is built and minimized.

Currently, 180 observables are considered from different collaborations [4, 5, 6, 7, 8, 10, 11]. Observable is referred to a measurement of either an angular observable, a branching ratio, or a ratio, in a particular energy bin. For $B \to K\mu\mu$, the available energy of the $\mu\mu$ invariant mass (the so called q^2 variable), runs from 0 to $(m_B - m_K)^2 \sim 22\text{GeV}^2$ while for the $B \to K * \mu\mu$ runs from 0 to $(m_B - m_{K*})^2 \sim 19\text{GeV}^2$. In this exercise, charged and neutral modes of the B meson are considered. This brings, $B^{0,\pm} \to K^{0,\pm}$ which are equivalent up to isospin corrections, thus important to include all of them for enlarging the data set. We also count with two different leptonic final states, $\mu\mu$ and $e + e-$. Rations of the same observable but with different lepton at the final state, the so-called R ratios, are also included and play an important psychological goal since they are a clear indication of the Lepton Flavor Universality Violation - LFUV - (this is, that the Z and W bosons of the SM interact differently with electrons and muons, something which is not expected from the theory) [3]. On the contrary, deviations between theory and experiment for specific lepton, either muons or electrons, cannot distinguish whether the anomaly is due to LFUV phenomenon or new physics come universally for muons and electrons. Only the combination of lepton-flavor dependent and rations of observables with different leptons in the final state may help. But actually, there is no clean way to disentangle such situation and current attempts consider that the most appealing solution to the puzzle is indeed having two different new physics scenarios, two different new-physics particles, one affecting all lepton flavours and a second one affecting exclusively the muonic modes [2].

For a given decay process, for example, the branching ratio $B^0 \to K^{*0}\mu\mu$ we will have five or six energy bins, with each bin having 1 or 2 GeV^2 sizes, thus providing 5 of 6 so-called observables into the fit.

For each measured observable, we have a theory prediction based on the Standard Model of Particle Physics (SM). With 180 observables, the χ^2 value of the SM reaches 225 points [1], which corresponds to a p-value of 1.4%. This indicates the SM to be very far to explain experimental measurements.

The strategy then has been to include on top of the SM, new operators in the effective Hamiltonian to be able to account for such experimental discrepancies.

Standard strategy, the simplest one, considers one-single operator at a time, the so-called 1D fits, starting with $\mathcal{C}_7$, $\mathcal{C}_9$ and $\mathcal{C}_{10}$. For example, in Table 1, some examples can be found from Ref. [1]. The first one, a fit with $\mathcal{C}_{9\mu}^{\rm NP}$ returns $\mathcal{C}_{9\mu}^{\rm NP} = -0.98$. The $\mathcal{C}_{9\mu}^{\rm SM} \sim 4$, so the NP correction to the SM to explain these anomalies are around 25%, pretty large. Pull$_{\rm SM}$ reflects how difficult would be for that model (defined as SM + $\mathcal{C}_{9\mu}^{\rm NP}$) would be able to explain the SM. The p-value is then a criterion to decide whether this fit is rejected or not. Since p-values are very large, the fit is very good compared to the SM p-value of 1.4%.

	All				LFUV			
1D Hyp.	Best fit	1 σ/2 σ	Pull$_{\rm SM}$	p-value	Best fit	1 σ/ 2 σ	Pull$_{\rm SM}$	p-value
$\mathcal{C}_{9\mu}^{\rm NP}$	-0.98	$[-1.15, -0.81]$ $[-1.31, -0.64]$	5.6	65.4 %	-0.89	$[-1.23, -0.59]$ $[-1.60, -0.32]$	3.3	52.2 %
$\mathcal{C}_{9\mu}^{\rm NP} = -\mathcal{C}_{10\mu}^{\rm NP}$	-0.46	$[-0.56, -0.37]$ $[-0.66, -0.28]$	5.2	55.6 %	-0.40	$[-0.53, -0.29]$ $[-0.63, -0.18]$	4.0	74.0 %
$\mathcal{C}_{9\mu}^{\rm NP} = -\mathcal{C}_{9'\mu}$	-0.99	$[-1.15, -0.82]$ $[-1.31, -0.64]$	5.5	62.9 %	-1.61	$[-2.13, -0.96]$ $[-2.54, -0.41]$	3.0	42.5 %
$\mathcal{C}_{9\mu}^{\rm NP} = -3\mathcal{C}_{9e}^{\rm NP}$	-0.87	$[-1.03, -0.71]$ $[-1.19, -0.55]$	5.5	61.9 %	-0.66	$[-0.90, -0.44]$ $[-1.17, -0.24]$	3.3	52.2 %

Table 1: Most prominent 1D patterns of NP in $b \to s\mu^+\mu^-$. Pull$_{\rm SM}$ is quoted in units of standard deviation. The p-value of the SM hypothesis is 11.0% for the fit "All" and 8.0% for the fit LFUV.

2 Aim of this project

Use a neural network to determine what combination of new operators renders the best possible fit to experimental data.

The intuition can be obtained by looking at the tables of Ref. [1], one of them reproduced here. Forget by now about coefficients with a prime such as $\mathcal{C}_{9'}$ or $\mathcal{C}_{10'}$. The difference between $\mathcal{C}_9^{\rm V}$ and $\mathcal{C}_9^{\rm U}$ is that V refers to muon channel exclusively while U means both muons and electrons. In the long table, different ad-hoc combinations of wilson coefficients are considered. There are, of course, other options. The question is what is the particular combnation that renders the largest agreement with data without using all coeficients free. This is, instead of using, for example 4 parameters $\mathcal{C}_9^{\rm V}$, $\mathcal{C}_9^{\rm U}$ and $\mathcal{C}_{10}^{\rm V}$, $\mathcal{C}_{10}^{\rm U}$, what combination of two parameters (made from the combination of these four, for example $\mathcal{C}_9^{\rm V} = \pm\mathcal{C}_9^{\rm U}$, $\mathcal{C}_9^{\rm V} = \pm\mathcal{C}_{10}^{\rm V}$) yields the best agreement.

The goal is then try to look for these combination via exploring the many options using a neural network. The input shall be a fit with one single parameter at the begining, and explore different χ^2 values using different input coefficients for each observable and then let the neural network decide the best strategy. Combining 2D fits, with two different combinations of coeficients, arbitrary cobination, also provides information for the NN to learn, and it is proved that this is most eficient.

$$\chi^2 = \sum_i \frac{(y_i^{\rm exp} - {\rm Model}_i)^2}{\sigma{\rm Exp}^2 + \sigma{\rm Model}^2}\,, \tag{1}$$

where Model refers to the theoretical calculation. Following the same idea, *pull* is

$$ {\rm Pull}_i = \frac{(y_i^{\rm exp} - {\rm Model}_i)}{\sqrt{\sigma{\rm Exp}^2 + \sigma{\rm Model}^2}} \tag{2}$$

3 Trining exercise

Before starting with full fits with 180 observables, we shall explore the mechanics with a just few of them that can be easily parameterized as a polynomial. The example considers the $\langle P_5' \rangle$ observable, the most famous one, in different bins. I run my codes to generate a polynomial description for this observable. In the table, experimental results and SM predictions for the observable are collected.

Scenario		Best-fit point	1 σ	2 σ	Pull_{SM}	p-value
Scenario 5	$\mathcal{C}_{9\mu}^{\rm V}$	-0.54	$[-1.06, -0.06]$	$[-1.68, +0.39]$	6.0	39.4 %
	$\mathcal{C}_{10\mu}^{\rm V}$	$+0.58$	$[+0.13, +0.97]$	$[-0.48, +1.33]$		
	$\mathcal{C}_9^{\rm U} = \mathcal{C}_{10}^{\rm U}$	-0.43	$[-0.85, +0.05]$	$[-1.23, +0.67]$		
Scenario 6	$\mathcal{C}_{9\mu}^{\rm V} = -\mathcal{C}_{10\mu}^{\rm V}$	-0.56	$[-0.65, -0.47]$	$[-0.75, -0.38]$	6.2	41.4 %
	$\mathcal{C}_9^{\rm U} = \mathcal{C}_{10}^{\rm U}$	-0.41	$[-0.53, -0.29]$	$[-0.64, -0.16]$		
Scenario 7	$\mathcal{C}_{9\mu}^{\rm V}$	-0.84	$[-1.15, -0.54]$	$[-1.48, -0.26]$	6.0	36.5 %
	$\mathcal{C}_9^{\rm U}$	-0.25	$[-0.59, +0.10]$	$[-0.92, +0.47]$		
Scenario 8	$\mathcal{C}_{9\mu}^{\rm V} = -\mathcal{C}_{10\mu}^{\rm V}$	-0.34	$[-0.44, -0.25]$	$[-0.54, -0.16]$	6.5	48.4 %
	$\mathcal{C}_9^{\rm U}$	-0.80	$[-0.98, -0.60]$	$[-1.16, -0.39]$		
Scenario 9	$\mathcal{C}_{9\mu}^{\rm V} = -\mathcal{C}_{10\mu}^{\rm V}$	-0.66	$[-0.79, -0.52]$	$[-0.93, -0.40]$	5.7	28.4 %
	$\mathcal{C}_{10}^{\rm U}$	-0.40	$[-0.63, -0.17]$	$[-0.86, +0.07]$		
Scenario 10	$\mathcal{C}_{9\mu}^{\rm V}$	-1.03	$[-1.18, -0.87]$	$[-1.33, -0.71]$	6.2	41.5 %
	$\mathcal{C}_{10}^{\rm U}$	$+0.28$	$[+0.12, +0.45]$	$[-0.04, +0.62]$		
Scenario 11	$\mathcal{C}_{9\mu}^{\rm V}$	-1.11	$[-1.26, -0.95]$	$[-1.40, -0.78]$	6.3	43.9 %
	$\mathcal{C}_{10'}^{\rm U}$	-0.29	$[-0.44, -0.15]$	$[-0.58, -0.01]$		
Scenario 12	$\mathcal{C}_{9'\mu}^{\rm V}$	-0.06	$[-0.21, +0.10]$	$[-0.37, +0.26]$	2.1	2.2 %
	$\mathcal{C}_{10}^{\rm U}$	$+0.44$	$[+0.26, +0.62]$	$[+0.09, +0.81]$		
Scenario 13	$\mathcal{C}_{9\mu}^{\rm V}$	-1.16	$[-1.31, -1.00]$	$[-1.46, -0.83]$	6.2	49.2 %
	$\mathcal{C}_{9'\mu}^{\rm V}$	$+0.56$	$[+0.27, +0.83]$	$[-0.02, +1.10]$		
	$\mathcal{C}_{10}^{\rm U}$	$+0.28$	$[+0.08, +0.49]$	$[-0.11, +0.70]$		
	$\mathcal{C}_{10'}^{\rm U}$	$+0.01$	$[-0.19, +0.22]$	$[-0.40, +0.42]$		

Table 2: Most prominent patterns for LFU and LFUV NP contributions from Fit "All". Scenarios 5 to 8 were introduced in Ref. [2]. Scenarios 9 (motivated by 2HDMs [13]) and 10 to 13 (motivated by Z' models with vector-like quarks [14]) are newly introduced in the main text.

We will use the following parameterization for the observables.

$$
\begin{aligned}
\langle P_5' \rangle_{[0.1,0.98]} &= 0.674 - 0.060 C_{10} - 0.012 C_{10}^2 - 0.104 C_9 - 0.010 C_9 C_{10} + 0.005 C_9^2 \\
\langle P_5' \rangle_{[1.1,2.5]} &= 0.196 - 0.001 C_{10} - 0.004 C_{10}^2 - 0.210 C_9 + 0.001 C_9 C_{10} + 0.002 C_9^2 \\
\langle P_5' \rangle_{[2.5,4]} &= -0.491 - 0.042 C_{10} + 0.001 C_{10}^2 - 0.285 C_9 - 0.020 C_9 C_{10} + 0.033 C_9^2 \\
\langle P_5' \rangle_{[4,6]} &= -0.826 - 0.066 C_{10} + 0.004 C_{10}^2 - 0.207 C_9 + 0.011 C_9 C_{10} + 0.058 C_9^2 \\
\langle P_5' \rangle_{[6,8]} &= -0.937 - 0.060 C_{10} + 0.009 C_{10}^2 - 0.131 C_9 + 0.036 C_9 C_{10} + 0.055 C_9^2 \\
\langle P_5' \rangle_{[15,19]} &= -0.572 - 0.021 C_{10} + 0.010 C_{10}^2 - 0.040 C_9 + 0.029 C_9 C_{10} + 0.025 C_9^2
\end{aligned}
$$

Notice that for $\mathcal{C}_9 = \mathcal{C}_{10} = 0$, we recover the SM prediction.

Table 3: $\langle P_5' \rangle$ bins, SM predictions, experimental results from LHCb collaboration, and Pulls between SM and experiment.

bin	SM	experiment	Pull
$(0.1, 0.98)$	0.67 $\pm$0.14	0.52 $\pm$0.10	+0.9
$(1.1, 2.5)$	0.20 $\pm$0.12	0.37 $\pm$0.12	-1.0
$(2.5, 4)$	-0.49 $\pm$0.12	-0.15 $\pm$0.15	-1.8
$(4, 6)$	-0.83 $\pm$0.08	-0.44 $\pm$0.12	-2.7
$(6, 8)$	-0.94 $\pm$0.08	-0.58 $\pm$0.09	-2.9
$(15, 19)$	-0.57 $\pm$0.05	-0.67 $\pm$0.06	+1.2

Imagine, you solve $\langle P_5' \rangle_{[0.1,0.98]} = 0.52 \pm 0.10$ but using only $\mathcal{C}_9$, you get $\mathcal{C}_9 = 1.58$. However, if you take $\langle P_5' \rangle_{[4,6]} = -0.83 \pm 0.08$, you get $\mathcal{C}_9 = -1.32$ which has nothing to do. A combined fit to all these observables returns $\mathcal{C}_9 = -1.06$ with $\chi^2_{\text{min}} = 9.2$, which is not really a good fit. Moreover, if we neglect the last bin, the fit returns $\mathcal{C}_9 = -1.24$ with $\chi^2_{\text{min}} = 4.0$, much better. The question is then, what observable weights more in the fit and how can be put into agreement the different tensions.

Figure 1 represents the final philosophy. Each observable (yellow circle) has an experimental value (x_i) depending on the bin you are looking at. For $\langle P_5' \rangle$ we have 6 different bins, so we will have 6 different x_i. Each of it can include $\mathcal{C}_9$, $\mathcal{C}_{10}$, with V, U, and all these options are represented with y_i. The combination of observable in a particular bin, with a fit with a $\mathcal{C}$ coefficient, returns a χ^2 value.

The NN should learn how to reach the best χ^2 by exploring all the possible combinations of $\mathcal{C}$ wilson coefficients with the experimental values and learn what is the optimal weights of the observables.

Two different strategies to be compared:

- look for the minimum χ^2
- look for the solution with closer $\chi^2/dof = 1$ possible.

References

[1] M. Algueró, B. Capdevila, A. Crivellin, S. Descotes-Genon, P. Masjuan, J. Matias, M. Novoa and J. Virto, Eur. Phys. J. C **79** (2019) no.8, 714 [arXiv:1903.09578 [hep-ph]].

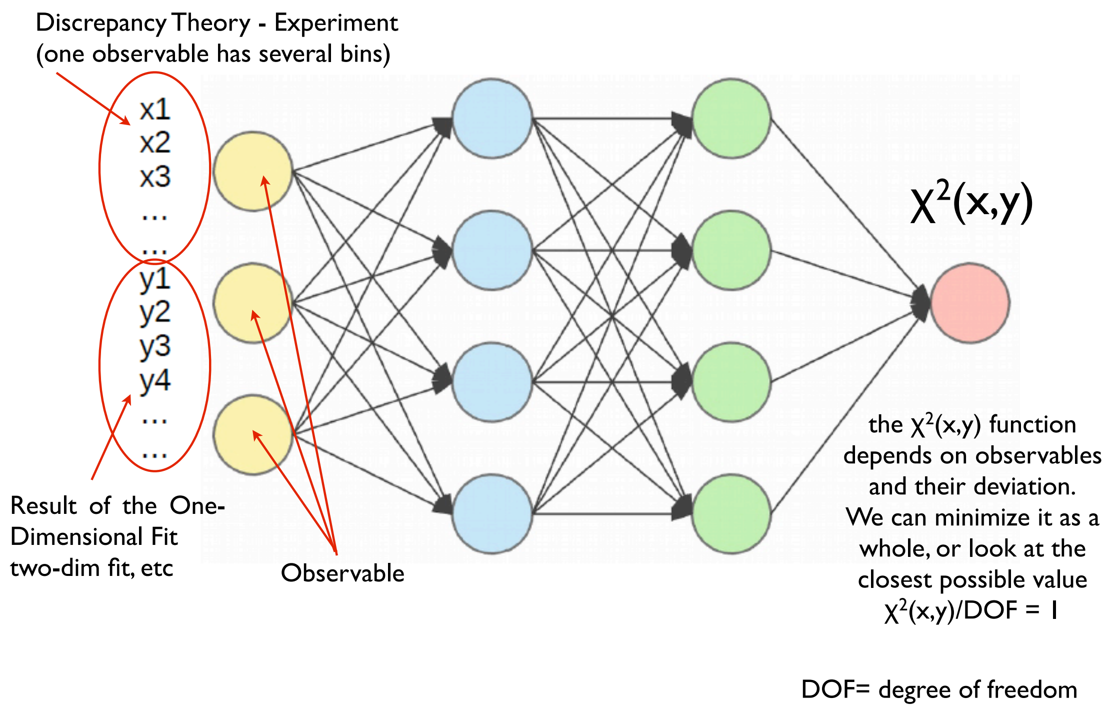

Figure 1: Simplistic view of a neural network.

[2] M. Algueró, B. Capdevila, S. Descotes-Genon, P. Masjuan and J. Matias, Phys. Rev. D **99** (2019) no.7, 075017 [arXiv:1809.08447 [hep-ph]].

[3] A. Vicente, [arXiv:2001.04788 [hep-ph]].

[4] R. Aaij *et al.* [LHCb Collaboration], arXiv:1903.09252 [hep-ex].

[5] S. Choudhury for the Belle collaboration, "Measurement of Lepton Flavour Universality in B decays at Belle", Talk at 'EPS-HEP Conference 2019'.

[6] R. Aaij *et al.* [LHCb Collaboration], JHEP **1406** (2014) 133 [arXiv:1403.8044 [hep-ex]].

[7] A. Abdesselam *et al.* [Belle Collaboration], arXiv:1904.02440 [hep-ex].

[8] M. Aaboud *et al.* [ATLAS Collaboration], [arXiv:1812.03017 [hep-ex]].

[9] S. Aoki *et al.* [Flavour Lattice Averaging Group], arXiv:1902.08191 [hep-lat].

[10] S. Wehle *et al.* [Belle Collaboration], Phys. Rev. Lett. **118** (2017) no.11, 111801 [arXiv:1612.05014 [hep-ex]].

[11] A. Abdesselam *et al.* [Belle Collaboration], arXiv:1604.04042 [hep-ex].

[12] L. Calibbi, A. Crivellin and T. Li, Phys. Rev. D **98** (2018) no.11, 115002 [arXiv:1709.00692 [hep-ph]].

[13] A. Crivellin, D. Muller and C. Wiegand, arXiv:1903.10440 [hep-ph].

[14] C. Bobeth, A. J. Buras, A. Celis and M. Jung, JHEP **1704** (2017) 079 [arXiv:1609.04783 [hep-ph]].