PROCEEDINGS OF SCIENCE

B-Meson Anomalies: Effective Field Theory Meets Machine Learning

A. Mir,[a,b,*] **J. Alda**[b,c,d] **and S. Peñaranda**[a,b]

[a] *Departamento de Física Teórica, Universidad de Zaragoza, Pedro Cerbuna 12, E-50009 Zaragoza, Spain*

[b] *CAPA, Centro de Astropartículas y Física de Altas Energías, Universidad de Zaragoza, E-50009 Zaragoza, Spain*

[c] *Dipartimento di Fisica ed Astronomia "Galileo Galilei", Università di Padova, via F. Marzolo 8, 35131 Padova, Italy*

[d] *INFN, Sezione di Padova, via F. Marzolo 8, 35131 Padova, Italy*

E-mail: amir@unizar.es, jorge.alda@pd.infn.it, siannah@unizar.es

Discrepancies between experimental measurements and Standard Model predictions in B-meson decays, especially in lepton flavor universality ratios like $R_{D^{(*)}}$, $R_{J/\psi}$ and branching ratios for processes like $B \to K^+ \nu\bar{\nu}$, suggest possible new physics (NP). In this study, we use an effective field theory framework, assuming NP effects only affect a single generation in the interaction basis, leading to non-universal mixing when rotating to the mass basis. We perform a global fit to the current experimental data, exploring three scenarios characterized by different mixing patterns and constraints. Our analysis finds that the best fit involves mixing between the second and third quark generations, with no lepton sector mixing and independent coefficients for singlet and triplet four-fermion operators. To accurately capture the non-Gaussian nature of the resulting parameter distributions, we use a Machine Learning-based Monte Carlo algorithm, enabling the generation of representative samples that reflect the true underlying distributions. This work highlights the valuable role of Machine Learning in accurately modeling complex parameter distributions in particle physics analyses.

The European Physical Society Conference on High Energy Physics (EPS-HEP2025)
7-11 July 2025
Marseille, France

[1] This work is partially supported by Spanish MINECO/FEDER Grants No. PGC2022-126078NBC21 funded by MCIN/AEI/10.13039/ 501100011033 and "ERDF A way of making Europe" and Grant DGA-FSE Grant 2020-E21-17R Aragón Government and the European Union-NextGenerationEU Recovery and Resilience Program on Astrofísica y Física de Altas Energías CEFCA-CAPA13ITAINNOVA.

[*] Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0) All rights for text and data mining, AI training, and similar technologies for commercial purposes, are reserved. ISSN 1824-8039 . Published by SISSA Medialab.

PoS(EPS-HEP2025)349

Semileptonic B meson decays provide sensitive tests of the Standard Model (SM) through lepton flavour universality (LFU) ratios and rare modes with neutrinos in the final state, $b \to c\ell\nu$ and $b \to s\,\ell^+\ell^-$ transitions. Persistent differences between measurements and precise SM predictions could be a sign for physics beyond the SM. Values of the SM predictions and experimental results for $R_D, R_{D^*}, R_{J/\psi}$ and $\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$ are collected in Table 1 and set the scale of the effects discussed here. The effective field theory (EFT) analysis provides a model-independent study of these effects and treats charged and neutral current sectors on equal footing and evaluates their correlations under controlled flavour assumptions. The aim is a minimal description of left-handed semileptonic modifications compatible with present data, together with a clear account of how information from diverse observables constrains the parameter space. The influence of several effective operators on numerous observables is mediated through their connection via the Wilson coefficients. In order to understand the impact of experimental results on flavour physics observables, it is necessary to perform a global fit to the available experimental data. We performed a global fit in previous articles [2–4], where an extensive list of references is included. In [4] the $b \to c\tau\bar{\nu}$ and the $B \to K^{(*)}\nu\bar{\nu}$ decays are analyzed, providing the best fit predictions for R_{D^*} and $R_{J/\psi}$ observables and for the $B^+ \to K^+\nu\bar{\nu}$ and $B^0 \to K^{*0}\nu\bar{\nu}$ branching ratios.

Observable	SM	Experiment
R_D	0.299 ± 0.004	0.342 ± 0.026
R_{D^*}	0.257 ± 0.005	0.287 ± 0.012
$R_{J/\psi}$	0.258 ± 0.004	0.52 ± 0.20
$\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$	$(0.429 \pm 0.021) \times 10^{-5}$	$(2.3 \pm 0.7) \times 10^{-5}$

Table 1: Representative inputs (SM vs. experiment) used in the analysis [1].

New-physics (NP) effects are encoded through two left-handed semileptonic structures organised as weak-isospin singlet and triplet, with flavour projectors selecting a dominant component in lepton and quark spaces:

$$\mathcal{L}_{\mathrm{NP}} = \frac{\lambda_\ell^{ij}\lambda_q^{kl}}{\Lambda^2}\left[C_1(\bar{\ell}_i\gamma_\mu\ell_j)(\bar{q}_k\gamma^\mu q_l) + C_3(\bar{\ell}_i\gamma_\mu\tau^I\ell_j)(\bar{q}_k\gamma^\mu\tau^I q_l)\right], \tag{1}$$

where λ^ℓ and λ^q are the matrices that determine the flavour structure of the NP interactions. The projectors are hermitian and idempotent with $\mathrm{tr}\,\lambda = 1$, parametrised by two complex numbers, $\alpha^{\ell,q}$ and $\beta^{\ell,q}$, as

$$\lambda_{\ell,q} = \frac{1}{1 + |\alpha_{\ell,q}|^2 + |\beta_{\ell,q}|^2}\begin{pmatrix} |\alpha_{\ell,q}|^2 & \alpha_{\ell,q}\overline{\beta}_{\ell,q} & \alpha_{\ell,q} \\ \overline{\alpha}_{\ell,q}\beta_{\ell,q} & |\beta_{\ell,q}|^2 & \beta_{\ell,q} \\ \overline{\alpha}_{\ell,q} & \overline{\beta}_{\ell,q} & 1 \end{pmatrix}. \tag{2}$$

We examine three benchmark hypotheses: Scenario I with $C_1 = C_3$ and quark mixing restricted to the second and third generations, no lepton mixing; Scenario II with $C_1 = C_3$ and generic mixing in both sectors; Scenario III with $C_1 \neq C_3$, 2–3 quark mixing only, no lepton mixing ($\alpha^\ell = \beta^\ell = 0$). In this setup, LFU ratios in $b \to c\ell\nu$ and the neutrino modes $B \to K^{(*)}\nu\bar{\nu}$ probe complementary combinations of (C_1, C_3) along components selected by $\lambda_{\ell,q}$.

The fit includes all observables implemented by `smelli` version 2.3.3 [5] plus the observable $R_{J/\psi}$ that we implemented using the form factors in [6]. This choice ensures that charged and

PoS(EPS-HEP2025)349

neutral current observables are fitted coherently since the coefficients C_1 and C_3 affect both types of currents, constrains the flavour mixing introduced by the quark projector λ_q after rotating to the mass basis, and prevents improvements in one sector from reappearing as tensions elsewhere. While LFU ratios in $b \to c\ell\nu$ and the neutrino modes $B \to K^{(*)}\nu\bar{\nu}$ provide the main pull on (C_1, C_3) under the hypotheses above, null tests, electroweak precision tests, CKM constraints and light-lepton channels contribute to breaking approximate degeneracies, profiling nuisances and validating the stability of the preferred region.

The analysis combines two complementary procedures. First, we determine the optimal parameters (C_1, C_3, β_q) by minimizing the χ^2 of the difference between experimental and theoretical values of the observables. This global fitting process provides a reference set of confidence intervals. Second, we construct a regression tree algorithm of the exact log–likelihood to accelerate high resolution mapping and uncertainty propagation.

We generate 10,500 parameter points and evaluate the exact log–likelihood at each point. The sample is split evenly into training (84%) and validation (16%) sets. We use a gradient boosted ensemble of shallow decision trees (XGBoost) with learning rate 0.03 and early stopping after 5 rounds; training converges after approximately 776 boosting iterations. On the validation set the emulator reproduces the targets with high linear correlation (Pearson coefficient $r \simeq 0.96$), and its level sets coincide with those of the exact likelihood within the plotting resolution used in the figures.

This tree based approach is advantageous here for three reasons aligned with our use case: (i) The confidence intervals are non Gaussian and can develop curved ridges in (C_1, C_3, β_q); boosted trees approximate such structure without assuming global smoothness. (ii) Mixed parameter scales and weakly constrained combinations are handled robustly by tree partitions without preconditioning. (iii) Once trained, evaluations are fast and enable dense sampling to draw 1σ-2σ contours and to propagate uncertainties to the observable planes. For diagnostics and interpretability, we also compute SHAP (Shapley Additive Explanations) values on the trained tree ensemble, which quantify the contribution of each parameter (C_1, C_3, β_q) to the emulator output both globally and point by point.

Other Machine Learning algorithms, such as neural networks, can emulate the likelihood as well, but in this low dimensional, non Gaussian setting they typically require larger training samples and stronger regularisation to avoid over-smoothing narrow ridges. Regression trees achieve the required accuracy with modest tuning and provide stable performance under reseeding.

Table 2 collects the best–fit results for the three benchmarks; Scenario III exhibits the largest pull and is therefore the preferred case, which is the one studied in the remainder of this proceedings.

Figure 1 presents predictions in the preferred region of Scenario III for the key observables used to confront the EFT hypotheses. For completeness, results for Scenarios I and II are also included. The left panel compiles LFU ratios from $b \to c\ell\nu$, while the right panel displays $\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$ and $\mathrm{BR}(B^0 \to K^{*0}\nu\bar{\nu})$. Enforcing $C_1 = C_3$ over-correlates charged and neutral current sectors and hinders a simultaneous description. Relaxing this relation decouples the combinations of (C_1, C_3) probed by the two classes of observables, yielding a coherent fit without inducing tension in channels that are SM-like at the order considered.

The comparison between likelihood structure of the fit and the Machine Learning algorithm prediction is shown in Figure 2 in the planes (C_1, C_3), (C_3, β_q) and (C_1, β_q) for Scenario III,

PoS(EPS-HEP2025)349

	Scenario I	Scenario II	Scenario III
C_1	$-0.11^{+0.03}_{-0.04}$	-0.12 ± 0.03	-0.205 ± 0.015
C_3	$-0.11^{+0.03}_{-0.04}$	-0.12 ± 0.03	$-0.12^{+0.02}_{-0.01}$
α^ℓ	—	0.0 ± 0.07	—
β^ℓ	0.00 ± 0.02	0.000 ± 0.014	—
α^q	—	-0.076	—
β^q	$0.78^{+1.22}_{-0.36}$	$0.85^{+1.05}_{-0.60}$	$0.64^{+1.36}_{-0.24}$
Pull	$5.71\,\sigma$	$5.51\,\sigma$	$6.25\,\sigma$
$\Delta\chi^2_{\rm SM}$	39.8	43.12	46.66
p-value	1.2×10^{-8}	3.5×10^{-8}	4.1×10^{-10}

Table 2: Summary of best–fit parameters and goodness–of–fit for the three benchmark scenarios. Uncertainties are one–standard–deviation intervals.

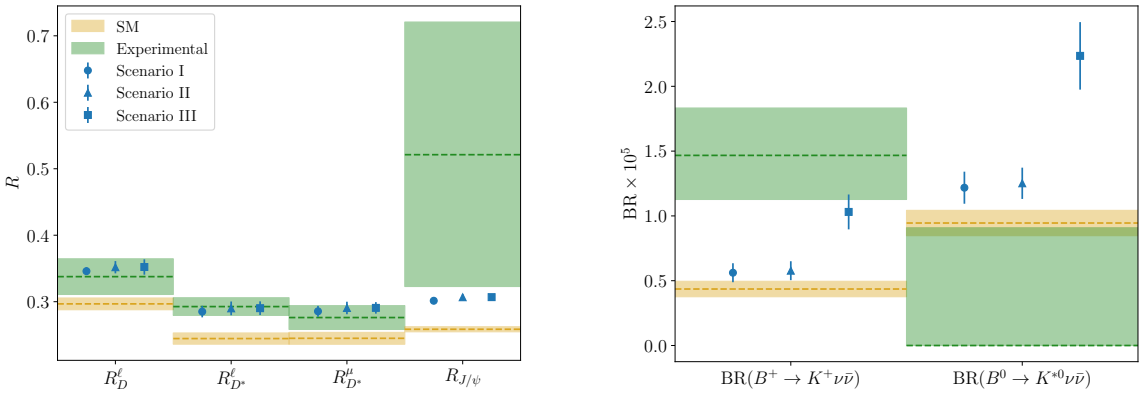

Figure 1: Scenario III: selected LFU ratios in $b \to c\ell\nu$ (left) and the neutrino-mode branching fractions (right) in the preferred region and across scenarios.

showing a high degree of agreement. Therefore, we can conclude that ML reproduces correctly the general features of the fit and it is a suitable strategy for this kind of analysis.

A compact global view is provided in Figure 3. The left panel compares the pull of each observable with respect to its experimental value in the SM (orange line) and in Scenario III (blue line). The numbers attached to some entries are indices of observables used in the full paper's appendix table (not reproduced here); they are shown only to identify which observables differ by more than $1\,\sigma$. The highlighted observables are all related to the B meson anomalies: $R^\ell_{D^*}$ (index 4) and R^ℓ_D (index 92) improve by $3.3\,\sigma$ and $1\,\sigma$, respectively; $\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$ (index 9) improves by $1.7\,\sigma$ and $\mathrm{BR}(B^+ \to K^{*+}\nu\bar{\nu})$ (index 161) by $1\,\sigma$, while $\mathrm{BR}(B^0 \to K^{*0}\nu\bar{\nu})$ (index 144) becomes worse by $1\,\sigma$. For $b \to s\mu^+\mu^-$ at high q^2, $\mathrm{BR}(B^\pm \to K^\pm\mu^+\mu^-)$ in the $[16, 17]$ GeV^2 (index 51) and $[15, 22]$ GeV^2 (index 87) bins, as well as $\mathrm{BR}(B_s^\pm \to \phi^\pm\mu^+\mu^-)$ in the $[15, 19]$ GeV^2 bin (index 111), improve by about $1\,\sigma$, whereas $\mathrm{BR}(B^\pm \to K^\pm\mu^+\mu^-)$ in the $[17, 18]$ GeV^2 bin (index 211) worsens by about $1\,\sigma$. Within a setup with only left-handed effective operators it is not possible to accommodate simultaneously $B^+ \to K^+\nu\bar{\nu}$ and $B^0 \to K^{*0}\nu\bar{\nu}$ as shown in Figure 1, which explains the deterioration

PoS(EPS-HEP2025)349

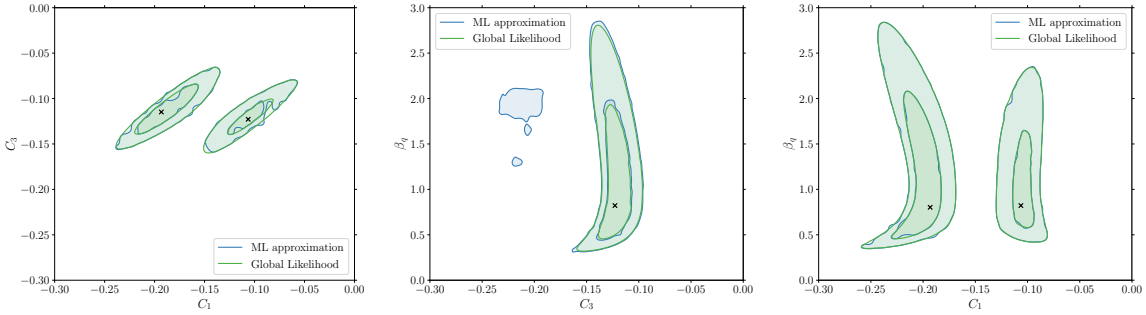

Figure 2: Scenario III: two–parameter likelihood slices in (C_1, C_3), (C_3, β_q) and (C_1, β_q). Regression–tree emulator and exact contours agree across the three planes.

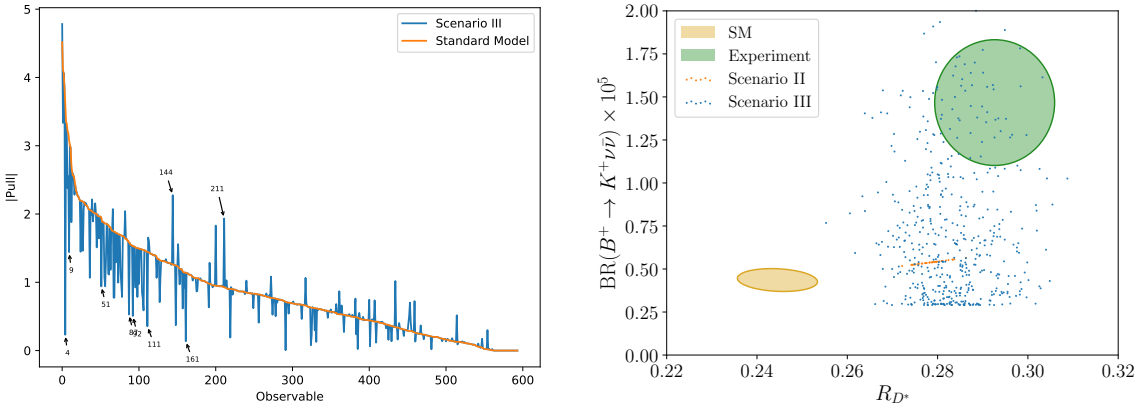

Figure 3: Scenario III: left, pulls for the SM and the preferred configuration; right, samples in $(R_{D^*}, \mathrm{BR}(B^+ \to K^+\nu\bar{\nu}))$ showing that the trend present for $C_1 = C_3$ vanishes for independent (C_1, C_3).

observed for index 144. The right panel displays samples in $(R_{D^*}, \mathrm{BR}(B^+ \to K^+\nu\bar{\nu}))$: the tight positive trend that appears when $C_1 = C_3$ disappears once the singlet and triplet coefficients are allowed to vary independently.

Summarising, an EFT with two left-handed semileptonic operators and flavour projectors provides a compact description of the data. The full set of relevant observables is required to impose gauge consistent relations across sectors, constrain the mixing pattern encoded in λ_q and prevent improvements in one channel from reappearing as tensions elsewhere. Within this framework, a focused subset of measurements provides the dominant sensitivity to (C_1, C_3, β_q). A regression tree algorithm reproduces the non Gaussian shapes seen in the fit and enables efficient uncertainty propagation and correlation studies. Allowing the singlet and triplet coefficients to vary independently, with mixing confined to the second and third quark generations and no lepton mixing, yields regions consistent with current measurements and removes the spurious relation between R_{D^*} and $\mathrm{BR}(B^+ \to K^+\nu\bar{\nu})$ from previous studies.

PoS(EPS-HEP2025)349

References

[1] Y. S. Amhis *et al.* [HFLAV], Phys. Rev. D **107** (2023) no.5, 052008, arXiv:2206.07501.

[2] J. Alda, J. Guasch, S. Peñaranda, JHEP 07 (2022) 115, arXiv:2109.07405.

[3] J. Alda, J. Guasch and S. Penaranda, Eur. Phys. J. Plus **137** (2022) no.2, 217, arXiv:2012.14799.

[4] J. Alda, A. Mir, S. Peñaranda, *A comparative analysis of results on B-anomalies using a machine learning algorithm*, arXiv:2412.15830.

[5] J. Aebischer, J. Kumar, P. Stangl and D. M. Straub, Eur. Phys. J. C **79** (2019) no.6, 509, 1810.07698.

[6] J. Harrison *et al.* [HPQCD], Phys. Rev. D **102** (2020) no.9, 094518, 2007.06957.

PoS(EPS-HEP2025)349