



Constructing model-agnostic likelihoods, a method for the reinterpretation of particle physics results

Lorenz Gärtner^{1,a} , Nikolai Hartmann^{1,b} , Lukas Heinrich^{2,c} , Malin Horstmann^{2,d} , Thomas Kuhr^{1,e} , MÉRIL Reboud^{3,f} , Slavomira Stefkova^{4,g} , Danny van Dyk^{5,h} 

¹ Ludwig Maximilians University Munich, 85748 Garching b. München, Germany

² Technical University Munich, 85748 Garching b. München, Germany

³ Theoretische Physik 1, Naturwissenschaftliche Fakultät, Universität Siegen, 57068 Siegen, Germany

⁴ Institut für Experimentelle Teilchenphysik, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany

⁵ Institute for Particle Physics Phenomenology and Department of Physics, Durham University, Durham DH1 3LE, UK

Received: 20 February 2024 / Accepted: 22 June 2024
© The Author(s) 2024

Abstract Experimental High Energy Physics has entered an era of precision measurements. However, measurements of many of the accessible processes assume that the final states' underlying kinematic distribution is the same as the Standard Model prediction. This assumption introduces an implicit model-dependency into the measurement, rendering the reinterpretation of the experimental analysis complicated without reanalysing the underlying data. We present a novel reweighting method in order to perform reinterpretation of particle physics measurements. It makes use of reweighting the Standard Model templates according to kinematic signal distributions of alternative theoretical models, prior to performing the statistical analysis. The generality of this method allows us to perform statistical inference in the space of theoretical parameters, assuming different kinematic distributions, according to a beyond Standard Model prediction. We implement our method as an extension to the `pyhf` software and interface it with the `EOS` software, which allows us to perform flavor physics phenomenology studies. Furthermore, we argue that, beyond the `pyhf` or `HistFactory` likelihood specification, only minimal information is necessary to make a likelihood model-agnostic and hence easily rein-

terpretable. We showcase that publishing such likelihoods is crucial for a full exploitation of experimental results.

1 Introduction

The results published using the data produced at high energy physics (HEP) experiments have large scientific potential beyond initial publication. To maximize the scientific impact of the data and corresponding results, facilitating reuse for combination and reinterpretation, should be made standard practice [1].

The importance of this is evident: Most analyses require underlying assumptions. These are, for example, theoretical distributions dictating signatures in the Monte Carlo (MC) data, which acts as a framework for constructing the analysis and provides a basis for comparison with the measured collider data. This also means, that a prior theoretical description has to be chosen, which typically corresponds to a Standard Model (SM) prediction. Therefore, the results obtained in the given analysis will be subject to a model dependency, which does not allow for simple reinterpretation in terms of alternative theories.

The goal of reinterpretation efforts in HEP is to maximize the insight gained from existing collider data, which requires overcoming this model dependency. One can classify these reinterpretation efforts as follows [2]:

- *Kinematic reinterpretation* or *recasting*, which includes testing an alternative physics process with different kinematic distributions. Here, changes of efficiencies and acceptance regions need to be considered.

^a e-mail: lorenz.gaertner@physik.uni-muenchen.de (corresponding author)

^b e-mail: nikolai.hartmann@physik.uni-muenchen.de

^c e-mail: lukas.heinrich@cern.ch

^d e-mail: malin.horstmann@tum.de

^e e-mail: thomas.kuhr@lmu.de

^f e-mail: merilreboud@gmail.com

^g e-mail: slavomira.stefkova@kit.edu

^h e-mail: danny.van.dyk@gmail.com

- *Model updating*, which refines either theoretical predictions or experimental calibrations. This achieves an overall reduction of the uncertainties. Technically, this can be viewed as a subclass of kinematic reinterpretation.
- *Combinations* of datasets and measurements across experiments. This is useful for reducing parameter uncertainties or for deriving global parameter constraints, where different decay channels have possibly different sensitivity to some parameters. For such combinations, it is necessary that the underlying model assumptions are mutually consistent.

Reinterpretation efforts have become a critical component of the research landscape [1, 2]. The main challenge remains the lack of public information on the analyses, that is, details available outside of the respective experimental collaborations. At the same time, the reinterpretation efforts are usually associated with a high computational cost due to the large number of theoretical models. A comprehensive study of all theoretical models, ranging from MC production through analysis to statistical inference, is not feasible.

A review of common reinterpretation methods and tools can be found in Refs. [1–4]. Popular approaches can be classified as [4]:

- *Simulation based reinterpretation* (e.g. CheckMate [5], MadAnalysis5 [6], RECAST [7]), where a full statistical analysis is performed on new MC samples produced according to an alternative theoretical model. This requires access to details of the full analysis strategy, as well as the underlying collider data and potentially also individual MC samples. This information is usually not available outside experimental collaborations. In addition, this approach is very computationally resource-heavy as new MC samples must be produced and analysed for each alternative theory.
- *Simplified model reinterpretation* (e.g. SModelS [8]), where one assumes that acceptances are not significantly affected by kinematic shape differences. Less information and computational resources are required, at the cost of approximations, which potentially lead to biases in the results (see Sect. 4.3).

In this paper, we propose an alternative reinterpretation method based on the reweighting of simulated MC templates, as opposed to a reweighting of individual MC samples. The proposed method strikes a balance between the required information on analysis details and computational cost of bias-free reinterpretation. Our work is an extension of the “brief idea” proposed in reference [9] and provides access to a reinterpretable likelihood function, directly parametrized in terms of any choice of theory parameters.

Reweighting is a standard practice in HEP commonly used for unfolding strategies, see e.g. reference [10]. The HAMMER software provides an application of reweighting for the purpose of reinterpretation of experimental measurements [11, 12]; see also the interface to RooFit [13]. At present, HAMMER allows for the reinterpretation of specifically implemented decays (mostly charged-current semileptonic B -meson decays) in terms of theoretical models of an effective field theory type by performing event-based reweighting. Our proposed reinterpretation method is more generally applicable, and it is not limited to any specific decay type or theoretical model. Furthermore, the proposed method makes use of reweighting on the distribution level, rather than on the event level; it does not require the full set of MC samples to perform the reinterpretation of the measurement; and it is more efficient in terms of computational costs.

One prerequisite for making HEP measurements suitable for reinterpretation and/or combination is the distributability of statistical models. As discussed in Ref. [1], a general recommendation is to make sensibly parametrized likelihood functions publicly available. A standard for likelihood parametrization and preservation has been developed around the pyhf software [14, 15] for statistical inference. The pyhf software is an implementation of the HistFactory model [16], which provides a general functional form for binned likelihoods. This fully parametrized binned likelihood is easily distributable in JSON format. In addition to the reinterpretation method proposed here, we also show that only a minimal amount of additional information allows for distributability of the *reinterpretable* likelihood.

The paper is structured as follows. In Sect. 2 we describe our novel reinterpretation method. We discuss the mathematical description and its applicability to unbinned and binned likelihoods, along with benefits and limitations. In Sect. 3 we describe the implementation of this method within the framework of common analysis tools. Finally, in Sect. 4 we apply the reinterpretation method to two toy examples, which use the model-agnostic framework of the Weak Effective Theory (WET). This effective theory covers all possible beyond Standard Model (BSM) theories that exclusively involve new particles or interactions at or above the scale of electroweak symmetry breaking. We also compare our newly developed method to a more simple approach, commonly used to reinterpret HEP results.

2 Reweighting method

The reinterpretation method described here is based on updating the distributions of the observable variables, given changes in the underlying kinematic distribution.

The probability density function (PDF) of reconstructed events $p(x)$ results from folding the PDF of a theoretical

kinematic prediction $p(z)$ with the conditional distribution $p(x|z)$ and the indicator function $\mathbb{1}_\varepsilon(x)$,

$$p(x) = \frac{1}{\varepsilon} \int dz \mathbb{1}_\varepsilon(x) p(x|z) p(z). \tag{1}$$

Here, the reconstruction variable x represents one or multiple observable variables and the kinematic variable z represents one or multiple kinematic degrees of freedom (d.o.f.). The function $\mathbb{1}_\varepsilon(x)$ models the selection criteria for a reconstructed event and $p(x|z)$ is the conditional probability of measuring a reconstructed configuration x , given an underlying particle configuration z . The overall reconstruction efficiency ε acts as a normalization factor for the PDF $p(x)$. The PDF $p(z)$ corresponds to the normalized kinematic distribution of a theoretical prediction $\sigma(z)$,

$$p(z) = \frac{\sigma(z)}{\sigma}. \tag{2}$$

The number density of expected events, given a total integrated luminosity L is $n(x) = L \sigma \varepsilon p(x)$ and further reads

$$n(x) = L \int dz \varepsilon(x|z) \sigma(z) = \int dz n(x, z), \tag{3}$$

where we combine both reconstruction and selection into $\varepsilon(x|z) = \mathbb{1}_\varepsilon(x) p(x|z)$, and where $n(x, z) = L \varepsilon(x|z) \sigma(z)$ can be thought of as a joint number density, similar to the joint PDF $p(x, z) = p(x|z) p(z)$.

The reinterpretation task involves determining the number density $n_1(x)$ of an *alternative* theoretical prediction $\sigma_1(z)$. This can be obtained by reweighting the joint number density $n_0(x, z)$ according to the kinematic *null* distribution $\sigma_0(z)$, via

$$\begin{aligned} n_1(x) &= L \int dz \varepsilon(x|z) \sigma_1(z) \\ &= L \int dz \varepsilon(x|z) \sigma_0(z) \frac{\sigma_1(z)}{\sigma_0(z)} \\ &= \int dz n_0(x, z) w(z). \end{aligned} \tag{4}$$

The weight factor $w(z)$ is simply the ratio of the theoretically predicted alternative kinematic distribution to the null distribution.

This reweighting process solely requires the knowledge of the joint null number density $n_0(x, z)$. Together with the weight factor, this is enough to predict the number density according to an alternative theory.

2.1 Discrete reweighting

In practical applications, the continuous joint number density is typically not analytically obtainable and requires estimation through MC simulations. To address this, one can discretize the reweighting method by representing the joint number density as a multidimensional matrix n_{xz} in bins of

$x \times z$. This is done alongside the binning of the theoretically predicted distribution σ_z and weight factor w_z in the kinematic d.o.f. z . Consequently, the discrete joint number density has dimension $\text{dim}(x) \times \text{dim}(z)$.

The discrete joint number density can be obtained by integrating the continuous joint number density over each $x \times z$ bin,

$$n_{xz} = \int_{\text{bin } x} \int_{\text{bin } z} dx' dz' n(x', z'), \tag{5}$$

where the integral boundaries are the bin boundaries of each x and z bin, respectively. The binned weights are given by

$$w_z = \frac{\sigma_{1,z}}{\sigma_{0,z}} = \frac{\int_{\text{bin } z} dz' \sigma_1(z')}{\int_{\text{bin } z} dz' \sigma_0(z')}. \tag{6}$$

The reweighting step of Eq. (4) becomes

$$n_{1,x} = \sum_z n_{0,xz} w_z. \tag{7}$$

We see an advantage in using this discrete approach because of lower computations costs. The price for this simplification is a loss of accuracy due to the binning in the kinematic d.o.f. z . However, this loss is controllable by increasing the number of bins. The binning should be chosen such that the joint number density and weight function factorize approximately in each bin. In principle, an arbitrarily fine binning can be chosen such that the uncertainty due to this loss is negligible compared to other sources of uncertainty (provided enough MC samples are available; see Appendix C.1).

Crucially, only a fixed set of samples from the joint PDF $p_0(x, z)$, based on the null prediction, is required, and no new samples need to be produced for the reinterpretation. Therefore, publishing the (binned) joint null number density and knowledge about the underlying kinematic null distribution is sufficient to perform the reinterpretation of a given measurement.

2.2 Limitation

The proposed reweighting approach is a light-weight and accurate way of obtaining new signal templates, given a joint number density and the kinematic null distribution. Still, it does have one main limitation: if the phase space contains regions with $w(z) \gg 1$, the effective MC sample size decreases. Put differently, we assign large weights to regions that are sparsely populated with MC samples obtained from the null distribution. Further, if the null distribution lacks support in a region of phase space, $\sigma_0(z) \rightarrow 0$, it can happen that $w(z) \rightarrow \infty$ when reweighting to an alternative distribution. In this case, we have no MC samples in the given region. The only solution that we see is reanalysing new samples, produced according to the alternative distribution.

3 Implementation and likelihood construction

Using the reweighting method, we can construct likelihood functions for particle physics analyses, directly parametrized in terms of theory parameters. Even though the reweighting method is independent of any likelihood formalism, we showcase our method in terms of the `HistFactory` formalism [16] as a baseline statistical model.

To build a global likelihood or posterior for a given measurement, including theoretical constraints or priors, we split the likelihood into three parts. The total likelihood is a combination of a data likelihood, L_{data} , the experimental constraint, C_{ex} , and the theory constraint, C_{th} ,

$$L = L_{\text{data}} \cdot C_{\text{ex}} \cdot C_{\text{th}}. \quad (8)$$

The data likelihood is constructed as a product of the Poisson probabilities of experimentally obtaining n events, when ν are expected from MC simulation,

$$L_{\text{data}}(\mathbf{n}|\boldsymbol{\eta}, \boldsymbol{\chi}) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb} | \nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi})). \quad (9)$$

Channels represent disjoint binned distributions, for example signal and control channels. *Bins* correspond to the histogram bins. The expected bin counts ν are a function of unconstrained, $\boldsymbol{\eta}$, and constrained, $\boldsymbol{\chi}$, parameters.

The experimental constraint consists of constraint terms or priors for experimental nuisance parameters $\boldsymbol{\chi}_{\text{ex}} \subset \boldsymbol{\chi}$, including all experimental systematic uncertainties,

$$C_{\text{ex}}(\mathbf{a}|\boldsymbol{\chi}_{\text{ex}}) = \prod_{\chi \in \boldsymbol{\chi}_{\text{ex}}} c_{\chi}(a_{\chi}|\chi). \quad (10)$$

In the frequentist language, constraints, $c_{\chi}(a_{\chi}|\chi)$, are obtained from auxiliary measurements with corresponding auxiliary data \mathbf{a} . This is the frequentist parallel of a *prior distribution*.

The theory constraint consists of constraint terms or priors for theoretical parameters, $\boldsymbol{\chi}_{\text{th}} \subset \boldsymbol{\chi}$,

$$C_{\text{th}}(\mathbf{a}|\boldsymbol{\chi}_{\text{th}}) = \prod_{\chi \in \boldsymbol{\chi}_{\text{th}}} c_{\chi}(a_{\chi}|\chi). \quad (11)$$

3.1 Implementation of the reweighting method

To obtain the data likelihood L_{data} for any theoretical model, one needs to calculate the event rates of the corresponding signal template. This requires an implementation of Eq. (7).

To achieve this, we work with `pyhf` [14, 15], which is an implementation of the `HistFactory` model [16]. Here, a likelihood is constructed by specifying the bin content of all contributing signal and background processes, as usually obtained from MC simulation, and the data measured in an experiment.

Furthermore, to implement uncertainties, one needs to specify the properties of a set of *modifiers*. The modifier settings include the event rate modifications according to each type of uncertainty at the 1σ level and the corresponding constraint type for the modifier parameter. The event rates for each channel and bin are calculated as

$$\nu_{cb}(\boldsymbol{\eta}, \boldsymbol{\chi}) = \sum_{s \in \text{samples}} \prod_{\kappa \in \boldsymbol{\kappa}} \kappa_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) \times \left(\nu_{scb}^0(\boldsymbol{\eta}, \boldsymbol{\chi}) + \sum_{\Delta \in \boldsymbol{\Delta}} \Delta_{scb}(\boldsymbol{\eta}, \boldsymbol{\chi}) \right), \quad (12)$$

where *samples* are used to separate physics processes, ν^0 are the nominal event rates (determined from MC), $\boldsymbol{\kappa}$ and $\boldsymbol{\Delta}$ are the full set of multiplicative and additive modifiers, respectively.

The implementation of the reweighting method (Sect. 2) requires extending the `pyhf` codebase.¹ The method is a prescription on the change of event rates. Therefore, a multiplicative modifier can be used to apply these changes. We extend `pyhf` by a *custom modifier*, which calculates the modifications to the nominal event rates according to the procedure described in Eq. (7). This custom modifier is a function of the underlying theory parameters of the alternative kinematic distribution.

The theory constraint C_{th} contains all constraint terms of these underlying theory parameters, which can be correlated in general. Per definition, modifier parameters in `pyhf` are treated as uncorrelated. To correctly include correlated parameter constraints in our statistical model, we decorrelate the theory parameters using principal component analysis (see Appendix B). We then assign one normally constrained `pyhf` modifier parameter to each of these decorrelated parameters.

4 Example application

A central aim of this work is to motivate the experimental HEP community to make use of the proposed method. This will, in turn, enable subsequent model-agnostic phenomenological analyses of the experimental results. Here, we detail the full reinterpretation of an analysis result of two toy examples from low-energy flavour physics.

In general, one distinguishes between two datasets for a given analysis: the *real data*, measured by a collider experiment and a set of simulated signal and background *MC data*, produced as a means of comparing against the measured collider data.

In this toy example, we do not use any experimental data, but two datasets of simulated samples, where one dataset acts

¹ We plan to contribute our modifications to the `pyhf` codebase.

as *real* data. The *MC data* is produced according to the SM prediction. The *real data* is produced by assuming that BSM physics affects the example process. To simulate detector and other analysis selection effects, observables in both datasets are smeared according to estimated detector resolutions, and event yields are scaled by an efficiency map.

By comparing the produced datasets, using either Bayesian or frequentist methods, we aim to recover the chosen BSM parameters, starting from the SM. This is possible only, because we made the likelihood a function of the theory parameters and a shape change in the kinematic distribution due to BSM physics can be directly taken into account.

As a general result, we want to compute the posterior distribution in the space of theory parameters, given the two simulated datasets and prior parameter constraints. To obtain a posterior from the likelihoods of the form shown in Eq. (8), we use the `bayesian pyhf` [17], an extension to `pyhf`. The posterior is obtained by Markov chain Monte Carlo (MCMC) sampling from the total likelihood, following Bayes' theorem for auxiliary data \mathbf{a} and observations \mathbf{n} ,

$$p(\boldsymbol{\eta}, \boldsymbol{\chi} | \mathbf{n}, \mathbf{a}) \propto L_{\text{data}}(\mathbf{n} | \boldsymbol{\eta}, \boldsymbol{\chi}) p(\boldsymbol{\chi} | \mathbf{a}) p(\boldsymbol{\eta}). \tag{13}$$

The experimental and theoretical constraints are represented by the constraint priors $p(\boldsymbol{\chi} | \mathbf{a})$, and $p(\boldsymbol{\eta})$ contains the priors for the unconstrained parameters $\boldsymbol{\eta}$ as detailed in Ref. [17].

4.1 $B \rightarrow K \nu \bar{\nu}$

The recent measurements of the total rate of $B \rightarrow K \nu \bar{\nu}$ decays by the Belle II collaboration [18, 19] hint at an excess of signal events compared to the SM expectation. This has triggered a substantial interest in the HEP phenomenology community to interpret this excess as a sign of BSM physics and to extract the corresponding model parameters [20–22]. In this subsection, we study the performance of our proposed approach at the hand of simulated $B \rightarrow K \nu \bar{\nu}$ data.

4.1.1 Weak effective theory parametrization

While we cannot achieve a general model-independent theoretical description of the $B \rightarrow K \nu \bar{\nu}$ decay, it is nevertheless possible to capture the effects of a large number of BSM theories under mild assumptions, as mentioned previously. Here, we assume that potential new BSM particles and force carriers have masses at or above the scale of electroweak symmetry breaking. In this scenario, it is useful to work within an effective quantum field theory that describes both the SM and the potential BSM effects using a common set of parameters; this effective field theory is commonly known as the Weak Effective Theory (WET) [23–25].

For the description of $b \rightarrow s \nu \bar{\nu}$ transitions, it suffices to discuss the *sbvν sector* of the WET. It is spanned by a subset of local operators of mass-dimension six, which is

closed under the renormalization group [26]. Since the mass of the initial on-shell B meson limits the maximum momentum transfer in this process, the matrix elements of operators with mass dimension eight or above are suppressed by at least a factor of $M_B^2/M_W^2 \simeq 0.004$, which are hence commonly neglected in these types of analyses. The corresponding Lagrangian density for the *sbvν sector* reads [27]

$$\mathcal{L}^{\text{WET}} = -\frac{4G_F}{\sqrt{2}} \frac{\alpha}{2\pi} V_{ts}^* V_{tb} \sum_i C_i(\mu_b) \mathcal{O}_i + \text{h.c.}, \tag{14}$$

with G_F the Fermi constant, α the fine structure constant, and V the Cabibbo–Kobayashi–Maskawa quark mixing matrix, respectively. The separation scale is chosen to be $\mu_b = 4.2$ GeV. Matrix elements of the operators \mathcal{O}_i describe the dynamics of the process at energies below μ_b , while the dynamics at energies above μ_b are encoded in the (generally complex-valued) Wilson coefficients $C_i(\mu_b)$ in the modified minimal subtraction ($\overline{\text{MS}}$) scheme. This enables a simultaneous description of SM-like and BSM-like dynamics in $b \rightarrow s \nu \bar{\nu}$ processes, as long as all BSM effects occur at scales larger than μ_b ; the different dynamics result simply in different values of the Wilson coefficients.

Assuming massless neutrinos, the full set of dimension-six operators is given by [27],

$$\begin{aligned} \mathcal{O}_{\text{VL}} &= (\bar{\nu}_L \gamma_\mu \nu_L) (\bar{s}_L \gamma^\mu b_L) \\ \mathcal{O}_{\text{VR}} &= (\bar{\nu}_L \gamma_\mu \nu_L) (\bar{s}_R \gamma^\mu b_R) \\ \mathcal{O}_{\text{SL}} &= (\bar{\nu}_L^c \nu_L) (\bar{s}_R b_L) \\ \mathcal{O}_{\text{SR}} &= (\bar{\nu}_L^c \nu_L) (\bar{s}_L b_R) \\ \mathcal{O}_{\text{TL}} &= (\bar{\nu}_L^c \sigma_{\mu\nu} \nu_L) (\bar{s}_R \sigma^{\mu\nu} b_L), \end{aligned} \tag{15}$$

with

$$\bar{\nu}_L \equiv \nu_L^\dagger \gamma^0, \nu_L^c \equiv C \bar{\nu}_L^T, \tag{16}$$

and where $C = i\gamma^2\gamma^0$ is the charge conjugation operator. In the above, the subscripts V, S, T represent vectorial, scalar, and tensorial operators, respectively; $\nu_{L/R}$ represent left- or right-handed neutrino fields; and $q_{L/R}$ represent left- or right-handed quark fields. The spin structure of the operators is expressed in terms of the Dirac matrices γ^μ and their commutator $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The operators are defined as sums over the neutrino flavors, since this is a property that we cannot determine experimentally. If one assumes the existence of only left-handed massless neutrinos, all operators except VL and VR vanish. The SM point in the parameter space of the WET Wilson coefficients reads

$$C_{\text{VL}} \simeq 6.6, \quad C_i = 0 \quad \forall i \neq \text{VL}. \tag{17}$$

Presently, the only measured observable is the differential decay rate for $B \rightarrow K \nu \bar{\nu}$, which we simulate in this example. Since the B meson is a pseudoscalar, the decay is

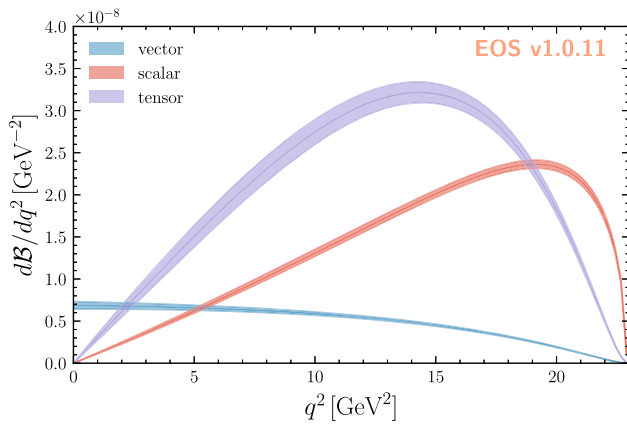


Fig. 1 Illustration of the variety of shapes of the $B \rightarrow K\nu\bar{\nu}$ decay rate due to purely vectorial, scalar, or tensorial interactions. Each curve corresponds to setting a single non-zero Wilson Coefficient in Eq. (18) to unity while keeping all other coefficients at zero

isotropic in the rest frame of the B meson and hence the only kinematically free variable is the squared dineutrino invariant mass, $q^2 = (p_B - p_K)^2 = (p_\nu + p_{\bar{\nu}})^2$. The differential decay rate is given by [27,28]

$$\begin{aligned} \frac{d\Gamma}{dq^2} = & 3 \left(\frac{4G_F \alpha}{\sqrt{2} 2\pi} \right)^2 |V_{ts}^* V_{tb}|^2 \frac{\sqrt{\lambda_{BK}} q^2}{(4\pi)^3 M_B^3} \\ & \times \left[\frac{\lambda_{BK}}{24q^2} |f_+(q^2)|^2 |C_{VL} + C_{VR}|^2 \right. \\ & + \frac{(M_B^2 - M_K^2)^2}{8(m_b - m_s)^2} |f_0(q^2)|^2 |C_{SL} + C_{SR}|^2 \\ & \left. + \frac{2\lambda_{BK}}{3(M_B + M_K)^2} |f_T(q^2)|^2 |C_{TL}|^2 \right], \end{aligned} \quad (18)$$

where M_B, M_K are the masses of the B meson and the kaon, respectively, m_b, m_s are the masses of the b and s quarks in the \overline{MS} scheme, respectively, and $\lambda_{BK} \equiv \lambda(M_B^2, M_K^2, q^2)$ is the Källén function.

In order to highlight the individual contributions of vectorial, scalar and tensorial terms from Eq. (18) to the differential decay rate, an illustration where individual left-handed Wilson coefficients are set to unity, is shown in Fig. 1.

As can be inferred from Eq. (18), the decay is only sensitive to the magnitude of three linear combinations of Wilson coefficients

$$|C_{VL} + C_{VR}|, |C_{SL} + C_{SR}|, |C_{TL}|. \quad (19)$$

The hadronic matrix elements of their operators are described by three independent hadronic form factors commonly known as $f_+(q^2)$, $f_0(q^2)$ and $f_T(q^2)$, which are functions of q^2 . In this work, the form factors are parametrized following the BCL parametrization [29], which is truncated at order $K = 2$. The values for the corresponding 8 hadronic parameters are extracted from a joint theoretical prior PDF

comprised of the 2021 lattice world average based on results by the Fermilab/MILC and HPQCD collaborations [30], and recent results by the HPQCD collaboration [31]. Correlations between the hadronic parameters are taken into account through their respective covariance matrices and implemented as discussed in Sect. 3.1.

The Belle II experiment, which found the first evidence for this decay, observes more events than expected in the SM. The ratio of observed to expected events is 4.6 ± 1.3 [19]. For latter use, we define a benchmark point in the space of Wilson coefficients that roughly reproduce the measured branching fraction, after correcting for the efficiency. Assuming all Wilson coefficients to be real, it reads

$$|C_{VL}|=10, |C_{VR}|=4, |C_{SL}|=3, |C_{SR}|=1, |C_{TL}|=1. \quad (20)$$

As shown in Eq. (18), the decay rate of $B \rightarrow K\nu\bar{\nu}$ is only sensitive to three linear combinations of Wilson coefficients shown in Eq. (19). The projection of our benchmark point onto this subspace reads

$$|C_{VL} + C_{VR}| = 14, |C_{SL} + C_{SR}| = 4, |C_{TL}| = 1. \quad (21)$$

4.1.2 Datasets

To make this example as realistic as possible, we design our setup similar to what has been done in the Belle II analysis.

The MC data is produced according to the SM prediction (*null* hypothesis) of the differential branching ratio $d\mathcal{B}(B \rightarrow K\nu\bar{\nu})/dq^2$, where the Wilson coefficient are set to the values in Eq. (17). The number of samples produced is equivalent to the number of signal events seen or expected for a given integrated luminosity. We produce samples for 362 fb^{-1} integrated luminosity, which corresponds to the total collider data used in Ref. [19], and for 50 ab^{-1} integrated luminosity, corresponding to the total target luminosity of the Belle II experiment [32], respectively. We multiply the estimated number of $B\bar{B}$ events with the SM branching fraction, $BR(B \rightarrow K\nu\bar{\nu}) \approx 4.81 \times 10^{-6}$ [33,34] to get a rough estimate for the number of MC samples to produce, prior to the efficiency correction (see below).

The real data is produced according to a BSM prediction (*alternative* hypothesis). Following the observations of more events than predicted in the SM [19], we use the previously defined benchmark point in Eq. (20). We multiply the estimated number of $B\bar{B}$ events with the BSM branching fraction, $BR(B \rightarrow K\nu\bar{\nu}) \approx 2.71 \times 10^{-5}$ [33,34] to get a rough estimate for the number of data samples to produce, prior to the efficiency correction (see below).

We list the number of produced samples in Table 1. Unless stated otherwise, all numerical values, figures, and tables provided in the following are obtained from studies that assume

Table 1 The number of $B \rightarrow K\nu\bar{\nu}$ samples produced for this study, corresponding to an equivalent of 362 fb^{-1} and 50 ab^{-1} integrated luminosity at the SuperKEKB collider. *Generated* and *reconstructed* samples correspond to the numbers prior and post efficiency correction

Luminosity	$B\bar{B}$ events	MC generated \ reconstructed	Data generated \ reconstructed
362 fb^{-1}	$\sim 3.87 \times 10^8$	1.86×10^3 241	1.05×10^4 1.14×10^3
50 ab^{-1}	$\sim 5.35 \times 10^9$	2.57×10^5 3.21×10^4	1.45×10^6 1.58×10^5

the 50 ab^{-1} sample size. We produce samples of the decay’s probability distribution for both the null and the alternative hypothesis using the EOS software in version 1.0.11 [34].

To simulate the detector resolution, we shift the q^2 value of each sample by a value drawn from a normal distribution of width 1 GeV^2 . This roughly corresponds to the Belle II detector resolution.

Furthermore, we apply an efficiency map to the samples according to the function

$$\varepsilon(q^2) = 0.4 \exp\left(-5 q^2/M_B^2\right), \tag{22}$$

which mimics the efficiency obtained in reference [19].

The reconstruction variable is chosen to be the reconstructed momentum transfer q_{rec}^2 , obtained from the kinematic variable q^2 , by applying detector and efficiency correction. The binnings of our kinematic and our reconstruction variables differ:

- For the reconstruction variable, we need to strike a balance between the number of events in each bin (to ensure sufficient statistical power) and the number of bins to ensure sensitivity toward differences in the shape of the q^2 distribution. We choose 8 equally spaced bins for the reconstruction variable.
- For the kinematic variable, we determine the number of bins as follows. We study the convergence of the expected yields from the reweighted model, as we increase the number of kinematic bins. For this study, we remove the detector resolution smearing. This is done for 100 randomly chosen theoretical models (see Appendix C.1 for further details). These models correspond to normally-distributed variations $\sim \mathcal{N}(0, 10)$ of the WET parameters with respect to the SM parameter point. We aim to ensure convergence at the level of 1% accuracy. We find that using 24 equally spaced bins for the kinematic variable ensures this aim. Figure 2 illustrates this procedure for the benchmark point in Eq. (20).

Both datasets, according to the null (SM) and alternative (BSM) hypothesis, and their corresponding changes after detector resolution smearing and efficiency correction are shown in Fig. 3.

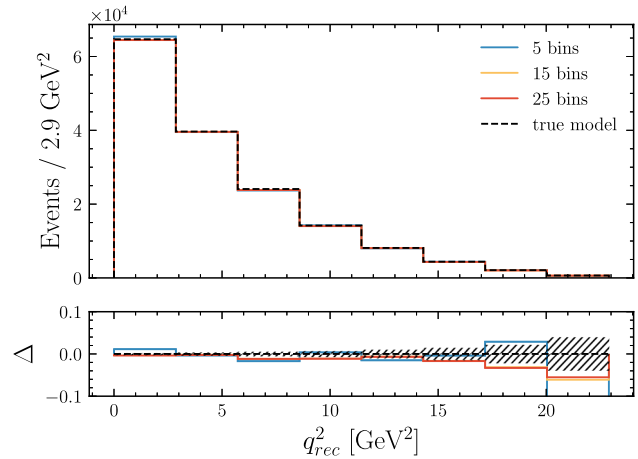


Fig. 2 The null histogram yields, reweighted to the benchmark point in Eq. (20), as a function of kinematic bins (red). The histogram yields of the true model, sampled from the probability density function at the benchmark point (black). The bottom plot shows the relative difference Δ of the reweighted models to the true model. The statistical uncertainty of the true model is shown as the hatched region

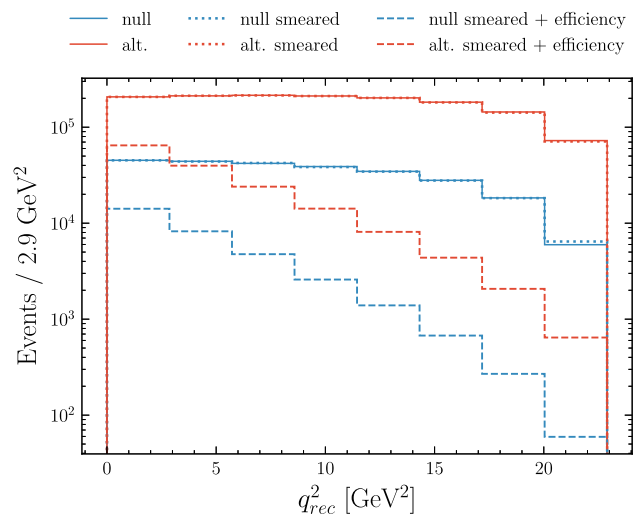


Fig. 3 Both the null/SM (blue) and alternative/BSM (red) $B \rightarrow K\nu\bar{\nu}$ datasets, according to the pure theoretical prediction, after detector resolution smearing and efficiency correction

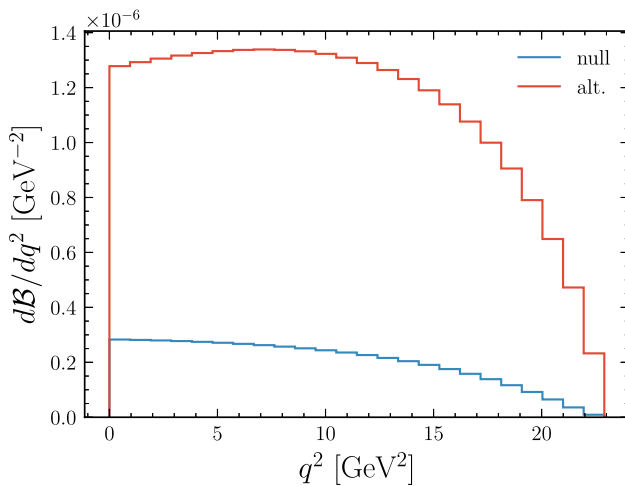


Fig. 4 The bin-integrated null/SM (blue) and alternative/BSM (red) predictions for the differential branching ratio $d\mathcal{B}(B \rightarrow K\nu\bar{\nu})/dq^2$

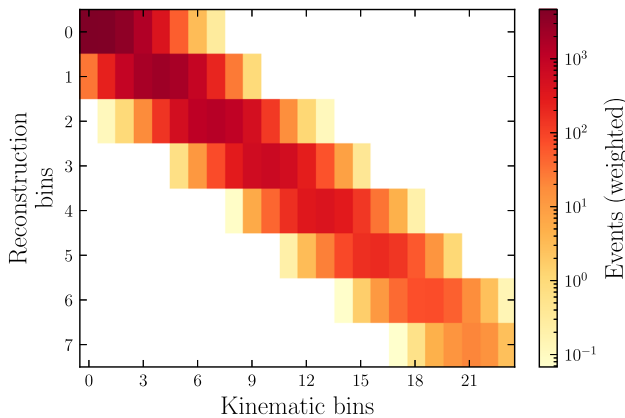


Fig. 5 The null joint number density, where we see the 8 bins of the reconstruction variable on the vertical axis and the 24 bins of the kinematic variable on the horizontal axis

The null (SM) and alternative (BSM) predictions have also been calculated using the EOS software [33,34] and are shown in Fig. 4.

The null joint number density is obtained by binning the MC data in a 2-dimensional histogram of the reconstruction variable q_{rec}^2 against the kinematic variable q^2 . This is shown in Fig. 5.

4.1.3 Full statistical model

To build the posterior according to the statistical model described in Sect. 3, we collect all parameters of our likelihood and their corresponding priors.

The theoretical parameters include the WET parameters and the hadronic parameters. The WET parameters correspond to the three independent linear combinations of Wilson coefficients that enter the theoretical description of $B \rightarrow K\nu\bar{\nu}$ decays; see Eq. (18). These are $|C_{VL} + C_{VR}|$,

$|C_{SL} + C_{SR}|$, and $|C_{TL}|$. While the Wilson coefficients are – in general – complex-valued parameters, we note that the overall phase of the WET Lagrangian Equation (14) is not observable. Moreover, an inspection of the differential decay rate in Eq. (18) shows only sensitivity to the absolute values of the three linear combinations of Wilson coefficients. Hence, we represent each linear combination as a positive real-valued number. Their prior is chosen as the uncorrelated product of uniform distributions with support

$$\begin{aligned} 5 &< |C_{VL} + C_{VR}| < 20, \\ 0 &< |C_{SL} + C_{SR}| < 15, \\ 0 &< |C_{TL}| < 15. \end{aligned} \quad (23)$$

The correlated hadronic parameters describe the $B \rightarrow K$ form factors as discussed in Sect. 4.1.1. Their prior is a multivariate normal distribution, which is implemented as a sequence of independent univariate normal distributions, as discussed at the end of Sect. 3.1.

The experimental constraint includes one parameter per bin of the reconstruction variable, representing the statistical uncertainty of the MC yields. The prior for these parameters are normal distributions $\mathcal{N}(1, 1/\sqrt{N_b})$, where N_b is the total yield in reconstruction bin b . For the purpose of this proof-of-concept study, we do not account for further (systematic) sources of uncertainty.

4.1.4 Reinterpretation results

Having built a model-agnostic likelihood function from our toy data, we investigate the potential of our approach to constrain the Wilson coefficients. Using MCMC sampling, we obtain the 3-dimensional marginal posterior distribution of the Wilson coefficients. The values at the mode of the full posterior agree with those of the benchmark point outlined in Eq. (20). We show the full set of 2-dimensional marginalizations of this posterior and the resulting intervals at 68% and 95% probability in Fig. 6.

We find that the marginal posterior peaks at the expected point, Eq. (21). The anti-correlation of the scalar and tensorial Wilson coefficients can be seen in their marginalized 2-dimensional distribution. This behaviour is not surprising, as the tensorial and scalar terms in Eq. (18) peak at larger values of q^2 , where the efficiency (Eq. (22)) is low. Moreover, the observed behaviour weakens as the statistical power of the data increases.

Overall, we see a good agreement with the expected Wilson coefficients, which acts as a closure test for our method.

4.2 Combination of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$

A limitation of studying solely the $B \rightarrow K\nu\bar{\nu}$ process is that its sensitivity to the WET Wilson coefficients is lim-

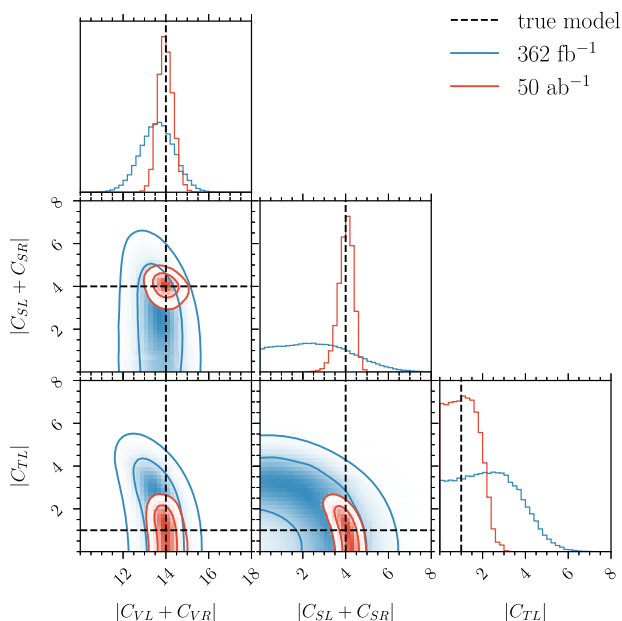


Fig. 6 The marginalized posterior distributions, obtained by MCMC sampling from the $B \rightarrow K \nu \bar{\nu}$ likelihood. On the diagonal, we see the 1-dimensional marginal distributions of the Wilson coefficients in Eq. (19). The contours on the 2-dimensional plots correspond to 68% (inner) and 95% (outer) probability. The dashed lines indicate the true underlying model (Eq. (21))

ited to the three linear combinations shown in Eq. (19). In this example, we showcase the power of combining data on $B \rightarrow K \nu \bar{\nu}$ and $B \rightarrow K^* \nu \bar{\nu}$ decays. These decays exhibit *complementary sensitivity* to the Wilson coefficients, due to their different hadronic spin and orbital angular momentum configurations. For the sake of simplicity of this example, we neglect effects of additional kinematic variables in the decay chain $B \rightarrow K^*(\rightarrow K\pi)\nu\bar{\nu}$, such as the helicity angle θ_K of the kaon and the $K\pi$ invariant mass. For the application of our proposed method to a real-world example, all kinematic variables should be included in the joint number density for full reinterpretability.

Moreover, this example shows from a technical perspective that our method and its implementation work also for combined `pyhf` models, providing full access to the complementarity in the sensitivity.

4.2.1 $B \rightarrow K^* \nu \bar{\nu}$ WET parametrization

The decay $B \rightarrow K^* \nu \bar{\nu}$ is governed by the same WET Lagrangian as described by Eqs. (14) and (15). Its differential decay rate reads [27, 28]

$$\frac{d\Gamma}{dq^2} = 3 \left(\frac{4G_F \alpha}{\sqrt{2} 2\pi} \right)^2 |V_{ts}^* V_{tb}|^2 \frac{\sqrt{\lambda_{BK^*} q^2}}{(4\pi)^3 M_B^3} \times \left[|\mathcal{A}_V|^2 |C_{VL} + C_{VR}|^2 + |\mathcal{A}_A|^2 |C_{VL} - C_{VR}|^2 \right]$$

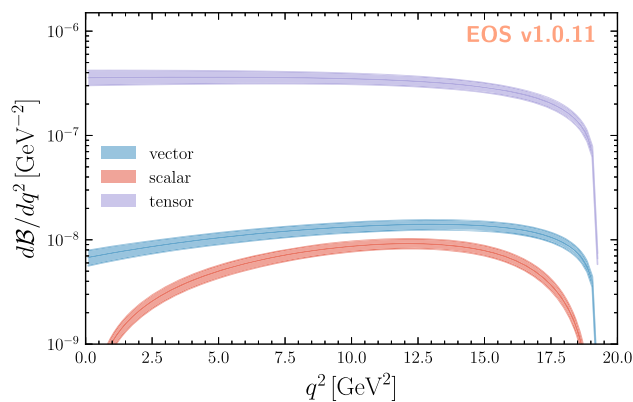


Fig. 7 Illustration of the variety of shapes of the $B \rightarrow K^* \nu \bar{\nu}$ decay rate due to purely vectorial, scalar, or tensorial interactions. Each curve corresponds to setting a single (left-handed) non-zero Wilson coefficient in Eq. (24) to unity while keeping all other coefficients at zero

$$+ |\mathcal{A}_P|^2 |C_{SR} - C_{SL}|^2 + |\mathcal{A}_T|^2 |C_{TL}|^2 \Big], \quad (24)$$

where the reduced amplitudes multiplying the Wilson coefficients read

$$\begin{aligned} |\mathcal{A}_V|^2 &= \frac{\lambda_{BK^*} |V(q^2)|^2}{12 (M_B + M_{K^*})^2}, \\ |\mathcal{A}_A|^2 &= \frac{8M_B^2 M_{K^*}^2}{3q^2} |A_{12}(q^2)|^2 \\ &\quad + \frac{(M_B + M_{K^*})^2 |A_1(q^2)|^2}{12}, \\ |\mathcal{A}_P|^2 &= \frac{\lambda_{BK^*}}{8 (m_b + m_s)^2} |A_0(q^2)|^2, \\ |\mathcal{A}_T|^2 &= \frac{32M_B^2 M_{K^*}^2 |T_{23}(q^2)|^2}{3 (M_B + M_{K^*})^2} + \frac{4\lambda_{BK^*} |T_1(q^2)|^2}{3q^2} \\ &\quad + \frac{4 (M_B^2 - M_{K^*}^2)^2 |T_2(q^2)|^2}{3q^2}. \end{aligned} \quad (25)$$

The description below Eq. (18) applies here as well.

In order to understand the individual contributions to the differential decay rate in Eq. (18) by vectorial, scalar, and tensorial operators, we provide an illustration of their relative sizes and their shapes in Fig. 7. This is achieved by setting their respective Wilson coefficients to unity.

One readily finds that the dependence of the observables on the Wilson coefficients is very different in Eqs. (18) and (24), respectively. Compared to $B \rightarrow K \nu \bar{\nu}$ decays, the differential $B \rightarrow K^* \nu \bar{\nu}$ decay rate exhibits additional sensitivity to the quantities

$$|C_{VL} - C_{VR}|, |C_{SL} - C_{SR}|. \quad (26)$$

As a consequence, a simultaneous analysis of both decays allows constraining a total of five real-valued (out of ten total real-valued) parameters in the $sb\nu\nu$ sector. Assuming all WET Wilson coefficients to be real-valued, this corresponds

Table 2 The number of $B \rightarrow K^* \nu \bar{\nu}$ samples produced for this study, corresponding to an equivalent of 50 ab^{-1} integrated luminosity at the SuperKEKB collider. *Generated and reconstructed* samples correspond to the numbers prior and post efficiency correction

Luminosity	$B\bar{B}$ events	MC generated \ reconstructed	Data generated \ reconstructed
362 fb^{-1}	$\sim 3.87 \times 10^8$	3.61×10^3 846	1.05×10^4 2.49×10^3
50 ab^{-1}	$\sim 5.35 \times 10^9$	4.99×10^5 1.16×10^5	1.45×10^6 3.43×10^5

to constraining the magnitudes of all Wilson coefficients. For an illustrative example, we apply this assumption here.

The hadronic matrix elements of the WET operators in this decay are expressed in terms of seven independent form factors $V(q^2)$, $A_0(q^2)$, $A_1(q^2)$, $A_{12}(q^2)$, $T_1(q^2)$, $T_2(q^2)$ and $T_{23}(q^2)$, which are functions of the momentum transfer q^2 . In this work, these form factors are parametrized following the BSZ parametrization [35], which is truncated at order $K = 2$. The values for the corresponding 19 hadronic parameters arise from the Gaussian likelihood provided in Ref. [36]. Correlations between the hadronic parameters are taken into account through their covariance matrix and implemented as discussed in Sect. 3.1.

4.2.2 Datasets

To produce the $B \rightarrow K^* \nu \bar{\nu}$ datasets, we adapt the same procedure as in Sect. 4.1.2.

The MC data is produced according to the SM prediction (*null* hypothesis). The number of samples is calculated by multiplying the estimated number of $B\bar{B}$ events in a 50 ab^{-1} Belle II dataset with the predicted SM branching fraction, $BR(B \rightarrow K \nu \bar{\nu}) \approx 9.34 \times 10^{-6}$ [33,34].

The real data is produced according to the BSM prediction of the benchmark point in Eq. (20) (*alternative* hypothesis). The number of data samples is calculated by multiplying the estimated number of $B\bar{B}$ events in a 50 ab^{-1} Belle II dataset with the predicted BSM branching fraction, $BR(B \rightarrow K^* \nu \bar{\nu}) \approx 2.72 \times 10^{-5}$ [33,34].

We list the number of samples in Table 2. We produce MC samples of the decay’s probability distribution for both the null and the alternative hypothesis using the EOS software in version 1.0.11 [34].

The efficiency map in this case is chosen to be

$$\varepsilon(q^2) = 0.3 \left(1 - 0.08 \exp \left(-2.5 q^2 / M_B^2 \right) \right), \quad (27)$$

which is an approximate expectation for an inclusive $B \rightarrow K^* \nu \bar{\nu}$ analysis. For the sake of simplicity, we assume that the efficiency is independent of the helicity angle θ_K and the $K\pi$ invariant mass.

We choose 10 bins in the reconstruction variable (q_{rec}^2) and find that 25 bins in the kinematic variable (q^2) pro-

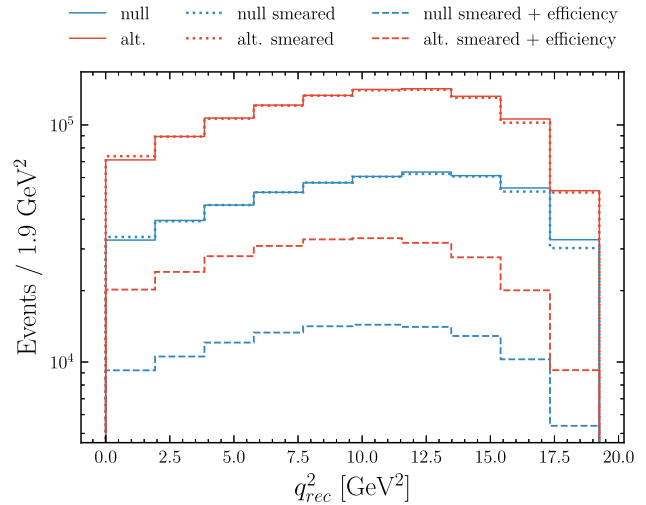


Fig. 8 Both the null/SM (blue lines) and alternative/BSM (red lines) $B \rightarrow K^* \nu \bar{\nu}$ datasets, according to the pure theoretical prediction, after detector resolution smearing and efficiency correction

vide a sufficient accuracy, using the procedure described in Appendix C.1 and Sect. 4.1.2.

Both datasets, according to the null (SM) and alternative (BSM) hypothesis, and their corresponding changes after detector resolution smearing and efficiency correction are shown in Fig. 8.

4.2.3 Full statistical model

In order to derive the statistical model encompassing $B \rightarrow K \nu \bar{\nu}$ and $B \rightarrow K^* \nu \bar{\nu}$, we construct individual posteriors for each channel following the methodology outlined in Sect. 4.1.3. The combined posterior arises from their product. The WET theory parameters are the only parameters shared by the individual posteriors.

These parameters correspond to the five magnitudes of Wilson coefficients that enter the theoretical description of $B \rightarrow K \nu \bar{\nu}$ and $B \rightarrow K^* \nu \bar{\nu}$ decays; see Eq. (18), (24). They are C_{VL} , C_{VR} , C_{SL} , C_{SR} , and C_{TL} . Their prior is chosen as the uncorrelated product of uniform priors with support

$$5 < |C_{VL}| < 15,$$

$$0 < |C_{VR}| < 10,$$

$$0 < |C_{SL}| < 10, 0 < |C_{SR}| < 10, 0 < |C_{TL}| < 10. \quad (28)$$

The hadronic parameters, describing the $B \rightarrow K$ and $B \rightarrow K^*$ form factors are discussed in Sects. 4.1.1 and 4.2.1. Their prior is a multivariate normal distribution, which is implemented as a sequence of independent univariate normal distributions, as discussed at the end of Sect. 3.1.

In the context of HistFactory models, the combined likelihood of $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ is a combination on the channel level, as discussed in Sect. 3. One custom modifier is added to each channel, which are functions of the same WET parameters, but different hadronic parameters.

4.2.4 Reinterpretation results

From the constructed model-agnostic likelihood function, we investigate the power of constraining the full set of Wilson coefficients appearing in Eqs. (18) and (24), under the assumption that they are real-valued. The decay rates in Eqs. (18) and (24) exhibit two discrete symmetries; one under the exchange $C_{VL} \leftrightarrow C_{VR}$ and another under the exchange $C_{SL} \leftrightarrow C_{SR}$. The combination of both symmetries leads to a four-fold ambiguity for the extraction of the Wilson coefficients from data and therefore a multimodal posterior density. To avoid computational issues in sampling from the posterior, we select one of the four fully equivalent modes for sampling. We do so by imposing the additional constraints $C_{VL} > C_{VR}$ and $C_{SL} > C_{SR}$. We use MCMC sampling and initialize the chains with the mode of the full posterior. The values at the mode of the full posterior align with those of the benchmark point outlined in Eq. (20). To obtain the full multimodal posterior, we restore the original symmetry manually. From the symmetrized samples, we obtain the 5-dimensional marginal posterior distribution of the Wilson coefficients. We illustrate it in Fig. 9 by showing the full set of 1- and 2-dimensional marginalizations and the resulting regions at 68% and 95% probability.

Most significantly, we see that it is now possible to probe the magnitudes of all 5 Wilson coefficients by combining results for $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$. In terms of accuracy, the benefits of this combination are especially visible for $|C_{TL}|$, compared to the $B \rightarrow K\nu\bar{\nu}$ result in Fig. 6. The improvement in precision of the 50 ab^{-1} over the 362 fb^{-1} datasets is also clearly visible. This is especially prominent in the scalar sector. For the smaller dataset, the peaks overlap such that the modes are not clearly separated. In addition, the fact that we are sampling the magnitudes of the Wilson coefficients causes an asymmetry in the scalar distributions. A tail of $|C_{TL}|$ towards lower values is present, as in the previous example in Fig. 6.

As in the previous example, this study serves as a further successful closure-test for our reinterpretation method.

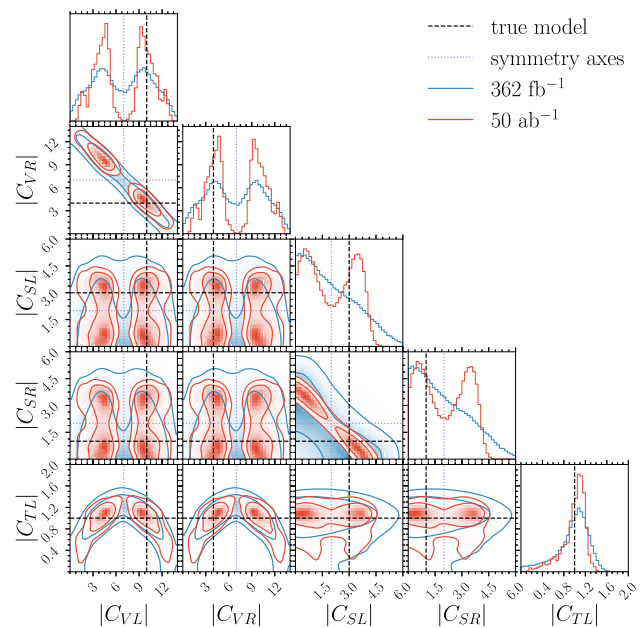


Fig. 9 The marginal posterior distributions, obtained by MCMC sampling from the combined $B \rightarrow K\nu\bar{\nu}$ and $B \rightarrow K^*\nu\bar{\nu}$ likelihood. On the diagonal, we see the 1-dimensional marginal distributions of the Wilson coefficients appearing in Eqs. (18) and (24). The contours on the 2-dimensional plots correspond to 68% (inner) and 95% (outer) probability. The dashed lines indicate the true underlying model. The dotted lines indicate the symmetry axes of the global likelihood

Furthermore, it shows how efficiently combinations of measurements can be performed with this method.

4.3 The necessity for reinterpretation

The availability of open-datasets for most particle physics results is currently very limited, although improving, thanks to the popularity of novel statistical approaches such as HistFactory and tools such as pyhf. These current limitations regularly hinder theorists to fully interpret existing experimental results in their BSM analyses. In particular, BSM changes to the distribution of the reconstruction variable are routinely neglected. In fact, the most common approach in a BSM analysis is to constrain the ratio of BSM prediction over SM prediction from branching ratio measurements or upper limits. This approach is only valid if the BSM changes to the shape of the distribution of the kinematic variable can be accounted for by a systematic experimental uncertainty in the reconstruction space. As we show in the following, this does not hold for measurements of the branching ratio of $B \rightarrow K\nu\bar{\nu}$.

To illustrate the issue, we compare our results in Sect. 4.1 based on simulated data with those obtained from a naive rescaling of the branching fraction. In the language of the presented reinterpretation method, the latter corresponds to using only a single bin in the kinematic d.o.f., covering the

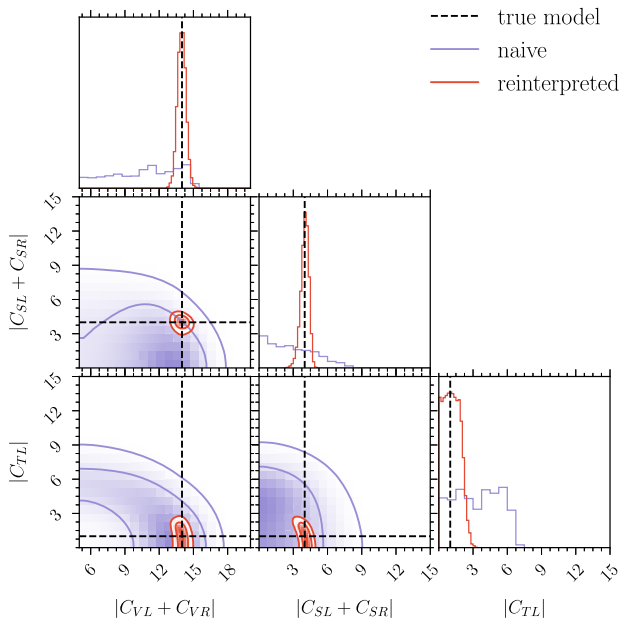


Fig. 10 The comparison of the posterior distribution resulting from a model with only one bin in the kinematic d.o.f. to the proposed reinterpretation method, respecting shape changes in the kinematic distribution (see Fig. 6)

full kinematic range. This further translates to a single weight applied to all bins of the reconstruction variable, corresponding to the ratio of the alternative to the null prediction integrated over the full kinematic range. We therefore construct a further “naive” $B \rightarrow K \nu \bar{\nu}$ posterior, which deviates from the setup in Sect. 4.1 only by using a single bin in the kinematic range.

After sampling from this “naive” posterior, we compare the marginal distributions for the Wilson coefficients to those presented in Fig. 6. This comparison is shown in Fig. 10.

We find a striking difference in the overall shape of the distributions and the central intervals at 68% (95%) probability. Clearly, the “naive” procedure fails to validate, yielding large deviations from the benchmark point in Eq. (21) in all three sectors. Our results illustrate that our approach is essential for a faithful reinterpretation of the experimental results of $B \rightarrow K \nu \bar{\nu}$.

We want to emphasize that our method provides a means to ensure an *accurate* interpretation of the existing likelihood beyond the assumptions of the underlying signal model. This does not imply, however, that our interpretation is more *precise* than a naive BSM interpretation. Put differently, our approach eliminates a bias introduced by using an incorrect template for the decay’s kinematic distribution, however, at the expense of potentially larger uncertainties on the theory parameters.

5 Discussion and significance of the method

We present a novel reinterpretation method for particle physics results, which is simple in its application and requires only minimal information in addition to published likelihoods.

Our proposed method avoids biases that are introduced in the naive reinterpretation of the data at a negligible increase of compute time. As such, it provides most of the benefits of reinterpretation using full analysis preservation. Therefore, this method provides good trade-off between accuracy and speed, which also has the potential to be used for improving the accuracy of global effective field theory fits to many analysis results.

To showcase the method, we apply it to a simulated dataset of the $B \rightarrow K \nu \bar{\nu}$ decay, inspired by the recent Belle II analysis [19] but without resorting to using any public or private Belle II data.

Using the two examples discussed in Sect. 4, we validate our method by successfully recovering the benchmark theory point from the underlying synthetic data. This outcome underscores the accuracy and self-consistency of our approach. We further investigate the bias introduced by naive rescaling of the $B \rightarrow K \nu \bar{\nu}$ branching ratio. For our benchmark point, we find a sizable bias when determining the WET Wilson coefficient without the application of our method.

In conclusion, this paper illustrates the ease of applicability of and the urgent necessity for shape-respecting reinterpretation over the traditional approach. We hope that this work will motivate experimental collaborations and analysts to consider future reinterpretability of their results and to publish the necessary material (for further details, see Appendix C). This will, in turn, enable the whole community to use the analysis results with accuracy.

Funding L.G. is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 460248186 (PUNCH4NFEDI). M.H. and L.H. are supported by the Excellence Cluster ORIGINS, which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094-39078331. D.v.D. acknowledges support by the UK Science and Technology Facilities Council (grant numbers ST/V003941/1 and ST/X003167/1).

Data Availability Statement This manuscript has no associated data. [Author’s comment: Data sharing not applicable to this article as no datasets were generated or analysed during the current study.]

Code Availability Statement This manuscript has associated code/software in a data repository. [Author’s comment: The software produced or modified as part of this work is available in the following public repositories: [redist, pyhf, EOS].

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, pro-

vide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Funded by SCOAP³.

Appendix A: Code repository and examples

The code is available at this [repository](#) [37]. In the `examples` folder, one can find the examples described in this work.

Statistical inference is performed using `pyhf` [14,15]. Theoretical predictions are obtained from EOS [33,34].

Appendix B: Singular value decomposition

Singular value decomposition is a useful method for decorrelating a set of parameters by a unitary transformation.

In our case, we start with a covariance matrix C , which is symmetric. Hence, we can always decompose it as

$$C = USU^H$$

where $UU^H = \mathbb{1}$. The columns of the transformation matrix U are the eigenvectors of C . The eigenvalues, $s_i = S_{ii}$, are the variances in the rotated space. The standard deviations are $\sigma_i = \sqrt{s_i}$.

If we want to incorporate the variances, we can define a new transformation matrix $Z = U\sqrt{S}$. Z is column-wise composed of the eigenvectors of C , each of which is now scaled by the corresponding standard deviation.

The `pyhf` modifier parameters, \mathbf{p} , describe the contribution of each of these scaled eigenvectors. Hence, to rotate from these parameters to the standard deviation vector for the correlated parameters, $\boldsymbol{\alpha}$, we can use

$$\boldsymbol{\sigma}_\alpha = Z\mathbf{p}.$$

That way, the modifier parameters are all interpretable in the same way as the usual `pyhf` modifier parameters, i.e. that a $p_i = \pm 1$ corresponds to a shift of $\pm\sigma_i$ along the i^{th} eigenvector direction.

Appendix C: A recipe for application of this reinterpretation method

To assist with an easy application of this reinterpretation method to any analysis, we provide a simple 4-step guide on what needs to be done to reinterpret a result from high

energy physics. We focus here on the discrete approach of Sect. 2.1.

1. *Samples*. Gather your post-reconstruction samples and ensure that they contain information of all kinematic d.o.f. as well as the reconstruction variable.
2. *Null joint number density*. From these samples, build the null joint number density by simply binning samples in bins of the reconstruction variable times the kinematic d.o.f. (see Appendix C.1 on how to optimize the kinematic binning).
3. *Weights*. Identify your null prediction used for producing the original MC samples. Chose your alternative theoretical prediction(s) and ensure that the support of the null distribution covers the full range of the alternative distribution (this can also be done by setting an upper bound on the weights). Compute the weights as the ratio of the bin-integrated alternative to the bin-integrated null distribution (as in Eq. (6)).
4. *Inference*. Either making use of the code in [37] or by implementing Eq. (7), compute the expected yields, given the alternative prediction, making use of the joint number density and the computed weights. Using either `pyhf` [14,15] or alternative tools, statistical inference can be used to compute results for the alternative theory.

C.1. Kinematic binning

To obtain suggestions on the number of bins to use for the kinematic variable(s), one can follow a similar procedure, as already mentioned in Sect. 4.1.2.

For a large set of models, covering your parameter space as thoroughly as possible, compute the expected yields for a finer and finer binning in the kinematic d.o.f. (a new joint number density and new weights need to be computed every time). At each step, compute the difference to the results of the previous step and stop when reaching a pre-defined convergence condition. The maximum number of bins over all the looped models should give a good estimate on the number of kinematic bins to use.

References

1. K. Cranmer et al., SciPost Phys. **12**, 037 (2022). <https://doi.org/10.21468/SciPostPhys.12.1.037>. arXiv:2109.04981 [hep-ph]
2. S. Bailey et al., Data and analysis preservation, recasting, and reinterpretation (2022). Contribution to the Snowmass process. arXiv:2203.10057 [hep-ph]. <https://inspirehep.net/literature/2054747>
3. G. Stark, C.A. Ots, M. Hance (2023). arXiv:2306.11055 [hep-ex] <https://inspirehep.net/literature/2669860>
4. W. Abdallah et al., LHC reinterpretation forum. SciPost Phys. **9**, 022 (2020). <https://doi.org/10.21468/SciPostPhys.9.2.022>. arXiv:2003.07868 [hep-ph]

5. D. Dercks, N. Desai, J.S. Kim, K. Rolbiecki, J. Tattersall, T. Weber, *Comput. Phys. Commun.* **221**, 383 (2017). <https://doi.org/10.1016/j.cpc.2017.08.021>. arXiv:1611.09856 [hep-ph]
6. E. Conte, B. Fuks, G. Serret, *Comput. Phys. Commun.* **184**, 222 (2013). <https://doi.org/10.1016/j.cpc.2012.09.009>. arXiv:1206.1599 [hep-ph]
7. K. Cranmer, I. Yavin, *JHEP* **04**, 038 (2011). [https://doi.org/10.1007/JHEP04\(2011\)038](https://doi.org/10.1007/JHEP04(2011)038). arXiv:1010.2506 [hep-ex]
8. M. Mahdi Altakach, S. Kraml, A. Lessa, S. Narasimha, T. Pascal, W. Waltenberger, *SciPost Phys.* **15**, 185 (2023). <https://doi.org/10.21468/SciPostPhys.15.5.185>. arXiv:2306.17676 [hep-ph]
9. K. Cranmer, L. Heinrich, *Recasting through reweighting* (2017). <https://doi.org/10.5281/zenodo.1013926>
10. M. Pivk, F.R. Le Diberder, *Nucl. Instrum. Methods A* **555**, 356 (2005). <https://doi.org/10.1016/j.nima.2005.08.106>. arXiv:physics/0402083
11. S. Duell, F. Bernlochner, Z. Ligeti, M. Papucci, D. Robinson, *PoS ICHEP2016*, 1074 (2017). <https://doi.org/10.22323/1.282.1074>
12. F.U. Bernlochner, S. Duell, Z. Ligeti, M. Papucci, D.J. Robinson, *Eur. Phys. J. C* **80**, 883 (2020). <https://doi.org/10.1140/epjc/s10052-020-8304-0>. arXiv:2002.00020 [hep-ph]
13. J. García Pardiñas, S. Meloni, L. Grillo, P. Owen, M. Calvi, N. Serra, *JINST* **17**, T04006 (2022). <https://doi.org/10.1088/1748-0221/17/04/T04006>. arXiv:2007.12605 [hep-ph]
14. L. Heinrich, M. Feickert, G. Stark, pyhf: v0.7.6 (2024). <https://doi.org/10.5281/zenodo.1169739>. See also the GitHub webpage. <https://github.com/scikit-hep/pyhf/releases/tag/v0.7.6>
15. L. Heinrich, M. Feickert, G. Stark, K. Cranmer, *J. Open Source Softw.* **6**, 2823 (2021). <https://doi.org/10.21105/joss.02823>
16. K. Cranmer, G. Lewis, L. Moneta, A. Shibata, W. Verkerke (ROOT), HistFactory: a tool for creating statistical models for use with RooFit and RooStats. Technical Report (New York U., New York, 2012). <https://cds.cern.ch/record/1456844>
17. M. Feickert, L. Heinrich, M. Horstmann, in *26th International Conference on Computing in High Energy & Nuclear Physics* (2023). arXiv:2309.17005 [stat.CO]. <https://inspirehep.net/literature/2704887>
18. F. Abudinén et al. (Belle-II), *Phys. Rev. Lett.* **127**, 181802 (2021). <https://doi.org/10.1103/PhysRevLett.127.181802>. arXiv:2104.12624 [hep-ex]
19. I. Adachi et al. (Belle-II) (2023). arXiv:2311.14647 [hep-ex] <https://inspirehep.net/literature/2725943>
20. K. Fridell, M. Ghosh, T. Okui, K. Tobioka (2023). arXiv:2312.12507 [hep-ph]. <https://inspirehep.net/literature/2739211>
21. W. Altmannshofer, A. Crivellin, H. Haigh, G. Inguglia, J. Martin Camalich (2023). arXiv:2311.14629 [hep-ph] <https://inspirehep.net/literature/2725980>
22. E. Gabrielli, L. Marzola, K. Mürsepp, M. Raidal (2024). arXiv:2402.05901 [hep-ph] <https://inspirehep.net/literature/2756697>
23. J. Aebischer, M. Fael, C. Greub, J. Virto, *JHEP* **09**, 158 (2017). [https://doi.org/10.1007/JHEP09\(2017\)158](https://doi.org/10.1007/JHEP09(2017)158). arXiv:1704.06639 [hep-ph]
24. E.E. Jenkins, A.V. Manohar, P. Stoffer, *JHEP* **03**, 016 (2018). [https://doi.org/10.1007/JHEP03\(2018\)016](https://doi.org/10.1007/JHEP03(2018)016). arXiv:1709.04486 [hep-ph]
25. E.E. Jenkins, A.V. Manohar, P. Stoffer, *JHEP* **01**, 084 (2018). [https://doi.org/10.1007/JHEP01\(2018\)084](https://doi.org/10.1007/JHEP01(2018)084). arXiv:1711.05270 [hep-ph]
26. J. Aebischer et al., *Comput. Phys. Commun.* **232**, 71 (2018). <https://doi.org/10.1016/j.cpc.2018.05.022>. arXiv:1712.05298 [hep-ph]
27. T. Felkl, S.L. Li, M.A. Schmidt, *JHEP* **12**, 118 (2021). [https://doi.org/10.1007/JHEP12\(2021\)118](https://doi.org/10.1007/JHEP12(2021)118). arXiv:2111.04327 [hep-ph]
28. J. Gratex, M. Hopfer, R. Zwicky, *Phys. Rev. D* **93**, 054008 (2016). <https://doi.org/10.1103/PhysRevD.93.054008>. arXiv:1506.03970 [hep-ph]
29. C. Bourrely, I. Caprini, L. Lellouch, *Phys. Rev. D* **79**, 013008 (2009). <https://doi.org/10.1103/PhysRevD.82.099902>. arXiv:0807.2722 [hep-ph] [Erratum: *Phys. Rev. D* **82**, 099902 (2010)]
30. Y. Aoki et al., Flavour lattice averaging group (FLAG). *Eur. Phys. J. C* **82**, 869 (2022). <https://doi.org/10.1140/epjc/s10052-022-10536-1>. arXiv:2111.09849 [hep-lat]
31. W.G. Parrott, C. Bouchard, C.T.H. Davies (HPQCD), *Phys. Rev. D* **107**, 014510 (2023). <https://doi.org/10.1103/PhysRevD.107.014510>. arXiv:2207.12468 [hep-lat]
32. E. Kou et al., *Prog. Theor. Exp. Phys.* **2019**, 123C01 (2019). <https://doi.org/10.1093/ptep/ptz106>. <https://academic.oup.com/ptep/article-pdf/2019/12/123C01/32693980/ptz106.pdf>
33. D. van Dyk et al. (EOS Authors), *Eur. Phys. J. C* **82**, 569 (2022). <https://doi.org/10.1140/epjc/s10052-022-10177-4>. arXiv:2111.15428 [hep-ph]
34. D. van Dyk, M. Reboud, N. Gubernari, P. Lüghausen, D. Lejnak, S. Kürten, A. Kokulu, M. Kirk, L. Gärtner, F. Novak, V. Kuschke, C. Bobeth, E. Graverini, C. Bolognani, M. Ritter, T. Blake, M. Bordone, S. Meiser, E. Eberhard, E. Romero, I. Toijala, K.K. Vos, F. Beaujean, J. Eschle, M. Rahimi, R. O'Connor, EOS Version 1.0.11 (2024). <https://doi.org/10.5281/zenodo.10600399>. See also the GitHub webpage. <https://github.com/eos/eos/releases/tag/v1.0.11>
35. A. Bharucha, D.M. Straub, R. Zwicky, *JHEP* **08**, 098 (2016). [https://doi.org/10.1007/JHEP08\(2016\)098](https://doi.org/10.1007/JHEP08(2016)098). arXiv:1503.05534 [hep-ph]
36. N. Gubernari, M. Reboud, D. van Dyk, J. Virto, *JHEP* **12**, 153 (2023). [https://doi.org/10.1007/JHEP12\(2023\)153](https://doi.org/10.1007/JHEP12(2023)153). arXiv:2305.06301 [hep-ph]
37. L. Gärtner, *redist* (2024). <https://doi.org/10.5281/zenodo.10630165>. See also the GitHub webpage <https://github.com/lorenzennio/redist/tree/v1.0>