JHEP04 (2020) 007

# Clustering of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ kinematic distributions with ClusterKinG

**Jason Aebischer,**[a] **Thomas Kuhr**[b,c] **and Kilian Lieret**[b,c]

[a]*Excellence Cluster Universe,*
*Garching, Germany*

[b]*Excellence Cluster Origins,*
*Garching, Germany*

[c]*Ludwig Maximilian University,*
*Munich, Germany*

*E-mail:* jason.aebischer@tum.de, thomas.kuhr@lmu.de,
kilian.lieret@lmu.de

ABSTRACT: New Physics can manifest itself in kinematic distributions of particle decays. The parameter space defining the shape of such distributions can be large which is challenging for both theoretical and experimental studies. Using clustering algorithms, the parameter space can however be dissected into subsets (clusters) which correspond to similar kinematic distributions. Clusters can then be represented by benchmark points, which allow for less involved studies and a concise presentation of the results. We demonstrate this concept using the Python package `ClusterKinG`, an easy to use framework for the clustering of distributions that particularly aims to make these techniques more accessible in a High Energy Physics context. As an example we consider $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ distributions and discuss various clustering methods and possible implications for future experimental analyses.

# Contents

# 1 Introduction

New Physics (NP) contributions can influence the kinematic distributions of particle decays. While this opens up possibilities to find and determine the nature of NP, it can also be a nuisance for experimental studies, because most measurements require assumptions on certain kinematic distributions, e.g. to distinguish signal from background or to determine efficiencies.

For example, as was shown in [1], assuming a two-Higgs-doublet model of type II changes the experimental measurement of $R(D^{(*)})$, because discriminating between signal and background requires assumptions on the kinematic shapes of the signal, background and normalisation modes. Such kinematic shapes are in general determined from Monte Carlo simulations.

Thus, many experimental measurements are model-dependent and are often only conducted under the assumption of the Standard Model (SM). Discrepancies between the SM

prediction and the measured values are a good indication for NP. However comparing different NP models based on their predicted results has to be taken with a grain of salt, because the measurements themselves are model-dependent.

A further complication for both theoretical and experimental studies is the high dimensionality of the parameter space of typical NP models. If experimentalists wish to publish model-independent results, the studies have to be repeated for a large sample of parameter points. This can be computationally very expensive. Furthermore numerical results and many visualisations can only be shown for specific (often arbitrary) parameter points, leaving their representative character unknown.

A possibility to reduce the complexity of this problem is to identify regions in the parameter space which lead to similar kinematic distributions. These regions can be found using clustering algorithms. From each cluster, a most representative point (*benchmark point*, BP) can then be chosen.

Experimental studies can focus on these BPs, thereby reducing the multi-dimensional problem to a small number of BPs to be considered. The results can be presented for each BP, allowing for a clear-cut numerical result and simpler visualisations.

Such a strategy has been employed for the first time in the context of Higgs boson pair production in [2–4]. An effective field theory (EFT) approach has been adopted to parametrise the five-dimensional parameter space of anomalous Higgs couplings. It has been shown in [2], that for current and future LHC searches a total of 12 clusters gives a reasonable approximation of the considered parameter space. In [3] several clusters were subjected to experimental limits from the CMS collaboration [5]. Finally, in [4] a method to extend the experimental sensitivity from the BPs to the other cluster members is discussed.

In recent years substantial progress has been made in the EFT description of the SM in form of the SM Effective Theory (SMEFT) [6] and the Weak Effective Theory (WET). The calculation of the complete one-loop SMEFT renormalization group equations (RGEs) [7–10], the complete tree-level and one-loop matching from SMEFT onto WET [11–13] and the complete one-loop QCD and QED RGEs within WET [14, 15] allow for a general NP analysis of low-energy observables.

Various tools are dedicated to the study of Wilson coefficients above and below the EW scale. They include the Match(runn)ers `DsixTools` [10], `MatchingTools` [16] and `wilson` [17], the Wilson coefficient exchange format `WCxf` [18], the fitting tool `smelli` [19], basis codes like `BasisGen` [20], `DEFT` [21] and `abc-eft` [22], the observable calculator `flavio` [23], `FlavorKit` [24], `SuperIso` [25], `SPheno` [26, 27], `FormFlavor` [28], as well as packages related to SMEFT, such as `SMEFT Feynman Rules` [29] and `SMEFTsim` [30].

However public tools to cluster the phenomenology of NP operators in a systematic way are still missing so far.

To fill this gap we have written the Python package `ClusterKinG` (*Cluster*ing of *Kin*ematic *G*raphs), which aims to make clustering techniques more easily accessible in the context of EFTs and High Energy Physics in general. Despite this motivation, `ClusterKinG` is a general tool that can be applied to a diverse set of problems, even outside of physics.

In this article, we use this package to cluster kinematic distributions of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays. These decays are of particular interest in view of the current $B$ anomalies [31–39].

The rest of the paper is organized as follows: in section [2](#) we discuss the clustering method in general terms. In section [3](#) we present for the first time the `ClusterKinG` package and describe its features. In section [4](#) we apply the clustering method to kinematic distributions of the decays $\bar{B} \to D^{(*)} \tau^- (\to \ell \, \bar{\nu}_\ell \, \nu_\tau) \bar{\nu}_\tau$ and perform various consistency tests. Finally we conclude in section [5](#).

## 2   Clustering

In this section we discuss the different steps involved in our clustering approach. As a first step a suitable observable and the corresponding model of interest have to be chosen. After establishing the underlying parameter space the following steps will be performed:

- Sampling the parameter space of the process.

- Computing the kinematic distributions in question.

- Choosing a metric to measure differences between the kinematic distributions.

- Applying a suited clustering algorithm on the parameter space.

- Computing the BPs representing each of the clusters.

The above steps are explained in the following subsections.

### 2.1   Sampling of the parameter space

As discussed in the introduction, typical NP models depend on several parameters. Theoretical considerations (such as symmetry arguments) can often be used to limit the study to a subset of these parameters, thereby reducing the dimensionality of the problem. The considered range of these parameters can be motivated by existing exclusion limits or from theory.

From this subset of the original parameter space, a set of points, *sample points*, are taken and used for the rest of the analysis. While large numbers of sample points will make the following steps more computationally expensive, it is important that the sampling is fine enough to accurately represent the whole parameter space.

`ClusterKinG` allows for an arbitrary selection of sampling points. The examples presented in this paper use a uniform grid in the parameter space for simplicity. In order to limit the number of required points (and thereby computing time), it is also planned to implement adaptive sampling techniques in the future: after an initial run with a coarse grid, regions with large variations in the kinematic distributions are identified and sampled in a finer grid. If needed, this procedure can then be applied several times.

### 2.2   Kinematic distributions

For every sample point, the corresponding kinematic distribution needs to be computed. If analytic expressions of the observable in terms of the parameters are available, this task can be achieved by explicit evaluation of the formulae. Otherwise Monte Carlo (MC) simulations have to be used to generate the distributions. Since the generation of MC samples

is in general resource-intensive, reweighting techniques can be used to generate samples corresponding to different parameter points from existing samples. For semileptonic $B$ decays such methods are already implemented in the `HAMMER` tool [40, 41].

In this article we only consider *binned* observables and our kinematic distributions are thus histograms.

## 2.3 Metric

The objective of the clustering procedure is to partition the parameter space in such a way that parameter points generating similar kinematic distributions are grouped into the same cluster. For this, the "similarity" between kinematic distributions has to be quantified by a metric[1] in the space of kinematic distributions.

The choice of this metric follows from the interpretation of the distributions as potential measurements. As such, the metric of choice should give more weight to differences in bins that are associated with low expected experimental uncertainties, while being less sensitive to differences in bins of less experimental significance. Having an estimate of experimental uncertainties is also useful when deciding the number of clusters and benchmark points to present: sufficiently many to cover the whole variety of distributions that lead to distinguishable experimental results, but not arbitrarily many. In this way, the number of clusters then also serves as an estimate for the sensitivity of a distribution to NP parameters.

A common choice for a metric measuring the similarity between binned distributions is a two-sample test statistic such as the Kolmogorov-Smirnov test [42], the Anderson-Darling test [43] or the $\chi^2$ test.

In [2] a binned log likelihood test statistic is used to distinguish between two distributions. This likelihood ratio is obtained by taking the ratio of the binomial distribution of the two individual samples and the binomial distribution where both samples are assumed to be equal. The log of this ratio can be shown to be $\chi^2$-distributed up to a minus sign [44]. By basing the test statistic on binomial distributions, the metric incorporates the statistical significance of the different bins.

In this paper, we use a general $\chi^2$ metric operating on distributions with uncertainties. As the distributions are not measured but generated as described in section 2.2, an uncertainty estimate is applied to them, consisting of configurable statistical uncertainties and relative and absolute uncertainties that can be correlated between bins.

Let $n_{1i}$ and $n_{2i}$ be the bin contents of two histograms $H_1$ and $H_2$ with corresponding normalisations $N_1 = \sum_{i=1}^{N} n_{1i}$ and $N_2 = \sum_{i=1}^{N} n_{2i}$ and uncertainties per bin $\sigma_{1i}$ and $\sigma_{2i}$ $(i = 1, \ldots, N)$.

If all uncertainties are uncorrelated, the $\chi^2$ distance between the two histograms is defined by

$$\chi^2(H_1, H_2) = \sum_{i=1}^{N} \frac{(N_1 n_{2i} - N_2 n_{1i})^2}{N_1^2 \sigma_{2i}^2 + N_2^2 \sigma_{1i}^2} \,. \tag{2.1}$$

---

[1]Now and in the following, the term "metric" is used in a rather loose way, emphasizing the intuition of a distance measure while not necessarily fulfilling all requirements to be a metric in the mathematical sense.

In case of correlated errors, the correlation matrix $C$ is the same for both histograms by construction. Thus, one can switch to a simultaneous basis of uncorrelated errors by considering $\tilde{n}_{1i} = \sum_j C_{ij}^{-1} n_{1j}$ and $\tilde{n}_{2i} = \sum_j C_{ij}^{-1} n_{2j}$.

If both histograms are independently generated by the same random variable, then $\chi^2/N$ has an expectation value of 1. In the following we call two distributions distinguishable if their $\chi^2$-distance $\chi^2(H_1, H_2)/N$ is larger than one, corresponding to a $p$-value of 32%. This loose requirement is conservative in the sense that it will lead to more clusters than a stricter criterion.

The metric between the kinematic distributions gives rise to a metric $d$ acting directly on the parameter space. In our case:

$$d(c_1, c_2) \equiv \chi^2(H_1, H_2)/N \,, \tag{2.2}$$

for two sample points $c_{1,2}$ and their respective histograms $H_{1,2}$.

## 2.4 Clustering algorithm

In general a data set can be either clustered *hierarchically* or *partitionally*. Partitional clustering methods [45] such as for example K-means algorithms [46] only perform one single partition of the data. Furthermore the number of resulting clusters has to be chosen beforehand as an input parameter. In the following we focus on hierarchical clustering methods [47].

Hierarchical clustering algorithms group a given set of data in a nested sequence of clusters. Two approaches are common: bottom-up (or agglomerative) algorithms successively merge clusters to form larger clusters, whereas top-down algorithms successively split clusters into smaller ones. In both cases, a stopping criterion is needed to avoid a trivial result. In our analysis we will employ the agglomerative method with the following steps:

1. Associate each sample point to one cluster containing only this element.

2. Merge the nearest two clusters according to a metric $D$.

3. Repeat step 2 until the stopping criterion is fulfilled.

Note that the metric $D$ in step 2 is not between points in the parameter space, but between subsets (existing clusters) of this space. It makes sense to base $D$ on the metric $d$ introduced in eq. (2.2). Two canonical choices for the inter-cluster metric $D$ are

$$D_\infty(C_1, C_2) \equiv \max_{c_1 \in C_1, c_2 \in C_2} d(c_1, c_2) \,, \quad D_1(C_1, C_2) \equiv \frac{1}{|C_1||C_2|} \sum_{c_1 \in C_1, c_2 \in C_2} d(c_1, c_2) \,, \tag{2.3}$$

where $C_1$ and $C_2$ are clusters and $|C_{1,2}|$ denote their number of elements respectively.

While [2] uses the 'average' metric $D_1$, we will employ the $D_\infty$ metric in our analysis. This choice is more conservative in the sense that it usually leads to a larger number of clusters than in the case of $D_1$, simply because $D_\infty \geq D_1$.

The stopping criterion is chosen to be $D_\infty(C_1, C_2) > 1$ for all pairs of clusters. This means that if we merge two clusters the resulting larger cluster does not contain any two distinguishable points. As a consequence all clusters of our final result contain only indistinguishable sample points. This is not the case for the 'average' metric in general.

## 2.5 Benchmark points

After the application of the clustering algorithm BPs have to be determined for all of the resulting clusters. A BP is a cluster element $c$ of a cluster $C$, which is chosen as a representative of that particular cluster. Usually it is taken to be the parameter point that minimizes a certain figure of merit, which is commonly based on the metric $d$ of eq. (2.2). Examples are:

$$f_1(c, C) \equiv \frac{1}{|C|} \sum_{c_i \in C} d(c, c_i), \quad f_2(c, C) \equiv \sqrt{\sum_{c_i \in C} d(c, c_i)^2}, \quad f_\infty(c, C) \equiv \max_{c_i \in C} d(c, c_i), \quad (2.4)$$

which differ in their responsiveness to outliers in the data. The BPs are the key elements of the cluster analysis. They are determined to simplify experimental analyses, which can then lay their focus only on a finite set of BPs instead of the entire parameter space.

## 3 The ClusterKinG package

`ClusterKinG` is publicly developed as open source under MIT license on [github.com/clusterking](github.com/clusterking). The package aims for ease of use, while making it easy for users to manipulate and extend functionality.

The basic building blocks of `ClusterKinG` are worker classes: after initialisation, a set of methods can be called for further configuration, before calling a `run` method, that performs the desired action. By subclassing these worker classes, the functionality can be extended. Furthermore, these worker classes can be passed on to other worker classes to perform e.g. stability checks that require repetitive calling with slight parameter variations.

To demonstrate the ease of use, a fully working example of code to generate clusters, benchmark points and plots similar to the ones presented in section 4 is shown in appendix A.

A typical `ClusterKinG` workflow consists of the following steps:

1. Initialise a `Data` object: this is what the following worker classes read from and write to.

   Internally this class holds a `pandas DataFrame` [48] that contains the bin contents of the kinematic distributions for each sample point. Additional information (*metadata*) is saved as a nested dictionary. Each of the following steps will add information to the metadata, such that any output file contains a description of how it was generated.

   Data objects can be saved to and loaded from an output file in SQL format, allowing to save all data in a single file of comparably small size. Exports of the data to other formats (CSV, XLS, JSON, . . . ) are easily available.

2. Adding uncertainties (optional). After the distributions have been generated, an experimental uncertainty estimate can be added to them. Typically this consists of

   - Statistical uncertainties, modeled as Poisson errors (and thereby dependent on the actual bin content for the distribution for each sample point).

- Systematic uncertainties which can be given in absolute form or relative to the bin content for all sample points at once.

To save memory space and improve performance, the uncertainties on the distributions are only calculated when needed and are not saved to the Data object, but only the procedure of *how* they are calculated. This also means that it is very easy and fast to check the dependency of the clustering results on the uncertainties.

3. Scanning: (Kinematic) distributions are generated for the chosen sample points. This is done by a `Scanner` object or for convenient clustering in the space of Wilson coefficients its `WilsonScanner` subclass.

Generally this requires three steps:

   (a) Providing a function to generate kinematic distributions: any Python function can be used, in particular any observable from the `flavio` [23] package.

   (b) Defining the sample points, usually in the form of an equidistant grid in the parameter dimensions (though more fine-tunable methods are available). The input of `WilsonScanner` is given in the common `WCxf` format defined in [18]. The Wilson coefficients are specified at a certain scale and in a particular EFT and basis.

   (c) Running the scanning process. `ClusterKinG` supports multiprocessing to speed up calculations.

The resulting distributions and the spread of the bin contents among the sample points are visualised by calling some of the plot methods of the `Data` object.

4. Clustering: different clustering algorithms correspond to different subclasses of the `Cluster` class. Hierarchical clustering is implemented in the `HierarchyCluster` worker which internally uses algorithms of the `scipy` library [49] and can be used with a range of different metrics ($p$-metrics, $\chi^2$ metric, user defined functions).

5. Selection of benchmark points is done by the `Benchmark` class, which can be configured (and subclassed) to allow for different strategies.

Besides visualisation, the data class provides simple methods to find the closest benchmark or sampling points given a point in parameter space, among others.

The technical documentation of this package is available at clusterking.readthedocs.io and multiple usage examples are provided in the form of Jupyter notebooks in the main repository.

## 4 Clustering of $B \to D^{(*)}\tau\bar{\nu}$ distributions

### 4.1 Setup

In this subsection we describe the setup used for our numerical analysis. Motivated by current $B$ anomalies, we perform a clustering analysis on various $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ kinematic

distributions. Such $b \to c$ transitions are described by the following effective Lagrangian:

$$\mathcal{L}_{\text{eff}} = -\frac{4 G_F}{\sqrt{2}} V_{cb} \left[ C_{VL} O_{VL} + C_{VR} O_{VR} + C_{SL} O_{SL} + C_{SR} O_{SR} + C_T O_T \right] + \text{h.c.} \,, \quad (4.1)$$

with the CKM element $V_{cb}$, the Fermi coupling constant $G_F$ and the effective operators

$$\begin{aligned}
O_{VL} &= (\bar{c}\gamma_\mu P_L b)(\bar{\tau}\gamma^\mu P_L \nu_\tau) \,, & O_{VR} &= (\bar{c}\gamma_\mu P_R b)(\bar{\tau}\gamma^\mu P_L \nu_\tau) \,, \\
O_{SL} &= (\bar{c} P_L b)(\bar{\tau} P_L \nu_\tau) \,, & O_{SR} &= (\bar{c} P_R b)(\bar{\tau} P_L \nu_\tau) \,, \\
O_T &= (\bar{c}\sigma_{\mu\nu} P_L b)(\bar{\tau}\sigma^{\mu\nu} P_L \nu_\tau) \,.
\end{aligned} \quad (4.2)$$

We use the notation $P_{R,L} = \frac{1}{2}(\mathbb{1} \pm \gamma_5)$, $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and $b, c, \tau, \nu_\tau$ for the quark and lepton fields. The Wilson coefficients in eq. (4.1) are in general complex quantities. For our analysis we will assume the CP conserving limit, i.e. real Wilson coefficients. This is a common assumption[2] (see f.e. [32, 33]) and is mainly chosen for simplicity. Furthermore, for presentational reasons we study three out of the five Wilson coefficients and choose one for each Dirac structure, namely:

$$C_{VL}, C_{SL}, C_T \,. \quad (4.3)$$

We will assume values for the first two Wilson coefficients in the interval $[-0.5, 0.5]$. The tensor operator is constrained [19] from the longitudinal polarization fraction $F_L$ in $B \to D^{0*}\tau\nu$ [55] and we choose its Wilson coefficient to be in the interval $[-0.1, 0.1]$. For our analysis we have chosen an equidistant grid of 1000 sample points in the three-dimensional parameter space, where each of the Wilson coefficients lies within the aforementioned intervals.

The clustering is performed using the `ClusterKinG` package. As mentioned in section 2 we use the hierarchical clustering algorithm together with the $\chi^2$-metric defined in section 2. The stopping criterion is chosen such that the $\chi^2$ distance between all distributions within the same cluster is $\leq 1$, meaning that they are indistinguishable experimentally. Finally, the BPs are obtained by adopting the figure of merit $f_1$ from eq. (2.4).

The complete code that has been used for the generation of the results and plots below is provided in the `example` directory of the `ClusterKinG` repository together with usage instructions.

## 4.2 Results

### 4.2.1 $\cos(\theta_\tau)$-distribution of $\text{BR}(B \to D^{0*}\tau\nu)$

As a first example we consider nine bins of the $\cos(\theta_\tau)$-distribution of the branching ratio $\text{BR}(B \to D^{0*}\tau\nu)$, where $\theta_\tau$ denotes the angle between the $\tau$ and the $B$ meson in the dilepton mass frame. The kinematic distributions are generated using the `flavio` [23] package. With an assumed systematic uncertainty of 1% and statistical uncertainties corresponding to a yield of 700 events, our clustering procedure leads to a total of six clusters and their corresponding BPs.

---

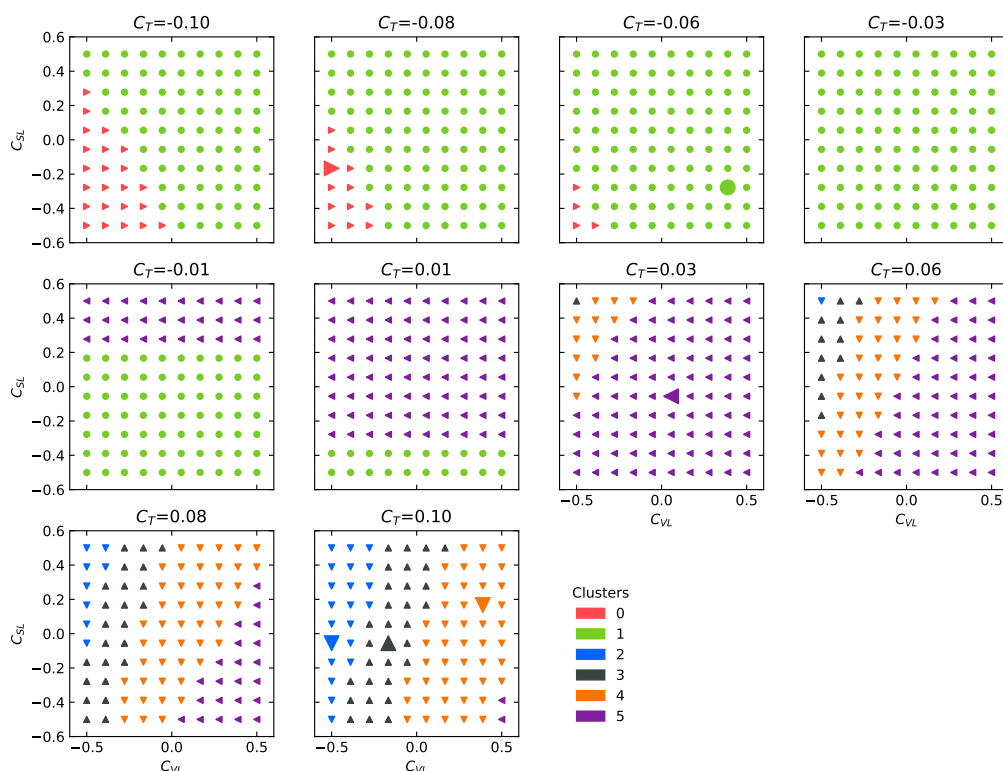[2]Studies with $CP$-violating contributions are analysed in [50–54].

**Figure 1**. Two-dimensional cuts of the clustered parameter space resulting from $\mathrm{dBR}(B \to D^{0*}\tau\nu)/\mathrm{d}(\cos(\theta_\tau))$ distributions are shown. The parameter space is spanned by the three Wilson coefficients $C_{VL}$, $C_{SL}$ and $C_T$, varied within their respective ranges. Six different clusters are found in our approach, which are indicated with different markers and colours. BPs are given in boldface.

The clustered parameter space is shown as two dimensional cuts in the $C_T$-direction in figure 1 and numeric values for the BPs are reported in table 1. As can be seen in figure 1, the parameter space exhibits a strong cluster variation in the direction of $C_T$. This fact is not surprising considering the explicit dependence of the kinematic distribution and agrees with the findings in [56], where this "flat term" observable has been proposed in the context of charged $b \to c$ transitions involving light leptons.

The distributions corresponding to the sample points are visualised as a box plot in figure 2. As expected, different clusters correspond to significantly different distributions. Furthermore, the distributions of the benchmark points are similar to the distributions given by the means of the bin contents of all distributions of the corresponding cluster.

### 4.2.2 $\cos(\theta_V)$-distribution of $\mathrm{BR}(B \to D^{0*}\tau\nu)$

As a second example we consider the $\cos(\theta_V)$-distribution of the process $B \to D^{0*}\tau\nu$. Here $\theta_V$ denotes the angle between $D^{0*}$ and the $B$ meson. The kinematic distributions for this process are again generated using `flavio` [23]. Assuming a signal yield of 700 events and a relative systematic uncertainty of 1%, the clustering procedure leads to a total of

| BP | $C_{VL}$ | $C_{SL}$ | $C_T$ |
|----|----------|----------|-------|
| 0  | $-0.50$  | $-0.17$  | $-0.08$ |
| 1  | $0.39$   | $-0.28$  | $-0.06$ |
| 2  | $-0.50$  | $-0.06$  | $0.10$  |
| 3  | $-0.17$  | $-0.06$  | $0.10$  |
| 4  | $0.39$   | $0.17$   | $0.10$  |
| 5  | $0.06$   | $-0.06$  | $0.03$  |

**Table 1**. List of benchmark points for the distribution $\mathrm{dBR}(B \to D^{0*}\tau\nu)/\mathrm{d}(\cos(\theta_\tau))$ obtained from the clustering procedure given in terms of the left-handed vector, scalar and tensor Wilson coefficients $C_{VL}, C_{SL}$ and $C_T$.
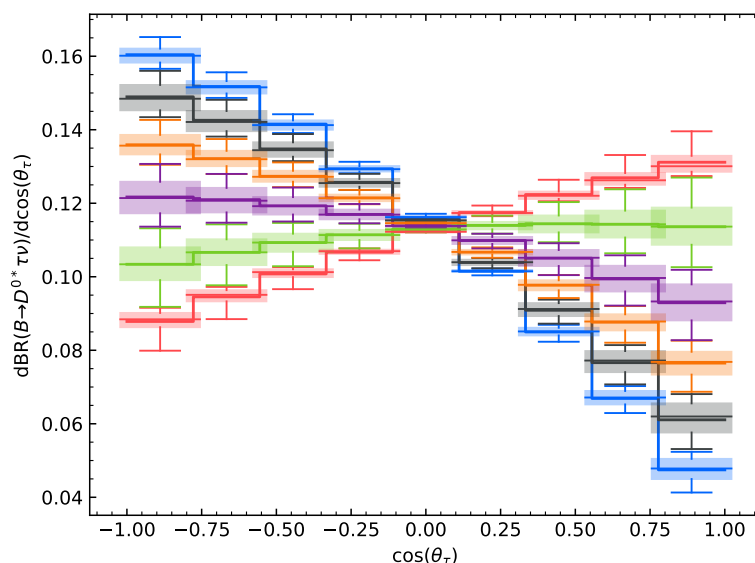


**Figure 2**. The distributions of the observable $\mathrm{dBR}(B \to D^{0*}\tau\nu)/\mathrm{d}(\cos(\theta_\tau))$ for the six different clusters (with colours matching these of figure 1). The histograms corresponding to the the BPs are shown as solid lines. The boxes extend from the upper to the lower quartile of the distribution of the bin contents within a cluster and a horizontal line indicates the median. Whiskers are used to further indicate the span of the data, covering six times the interquartile range.

three clusters. The clustered three-dimensional sample space is shown in figure 3 and $C_T$ can again be identified to be the most influential Wilson coefficient. A large subset of the parameter space belongs to cluster 2 (blue), whereas only a few sample points are contained in the first cluster (red), which is found at the edges of the sample space. The three BPs are reported in table 2.

Finally we show several example distributions for each cluster together with the BP distributions in figure 4. While the distinction between the red cluster and the other two clusters is very clear, the kinematic distributions of the blue and green clusters are more similar.
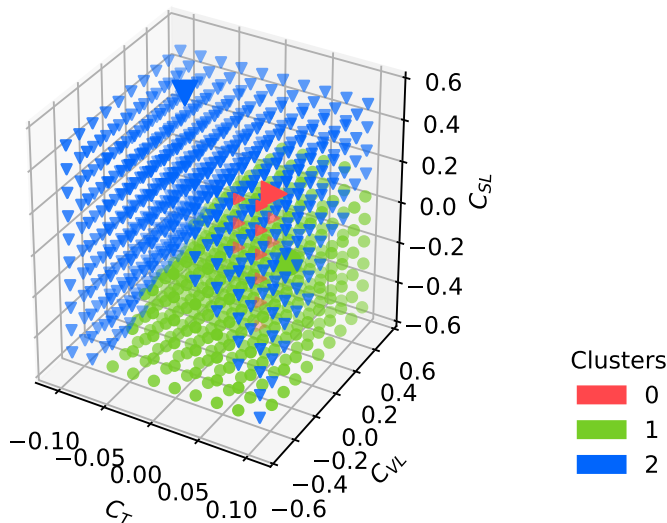
**Figure 3**. The clustered three-dimensional parameter space resulting from dBR($B \to D^{0*}\tau\nu$)/d(cos($\theta_V$)) distributions is shown. The parameter space is spanned by the three Wilson coefficients $C_{VL}$, $C_{SL}$ and $C_T$, varied within their respective ranges. Three different clusters are found in our approach, which are indicated with different markers and colours. BPs are given in boldface.

| BP | $C_{VL}$ | $C_{SL}$ | $C_T$ |
|----|----------|----------|-------|
| 0 | $-0.39$ | $0.50$ | $0.10$ |
| 1 | $0.17$ | $-0.17$ | $0.06$ |
| 2 | $0.17$ | $0.50$ | $-0.06$ |

**Table 2**. List of benchmark points for the distribution dBR($B \to D^{0*}\tau\nu$)/d(cos($\theta_V$)) obtained from the clustering procedure given in terms of the left-handed vector, scalar and tensor Wilson coefficients $C_{VL}, C_{SL}$ and $C_T$.

Compared to $\theta_\tau$, less clusters are found, but the shapes of the clusters are different from the previous ones. The two observables can therefore be considered complementary in their respective sensitivity to NP models.

### 4.2.3  $q^2$-distribution of BR($B \to D^0\tau\nu$)

The $q^2$-distribution of BR($B \to D^0\tau\nu$) has already been studied extensively in the literature. For our purpose we consider this observable to study the influence of the systematic and statistical uncertainties on the resulting number of clusters. This is relevant for future experiments such as Belle II, reaching a higher luminosity than previously attained. The $q^2$-distributions were computed using `flavio` [23]. In figure 5 we show the number of clusters as a function of the signal yield for various systematic uncertainties. As expected, the number of resulting clusters is dominated by the signal yield for small yields, while systematic uncertainties become the limiting factor for larger yields.
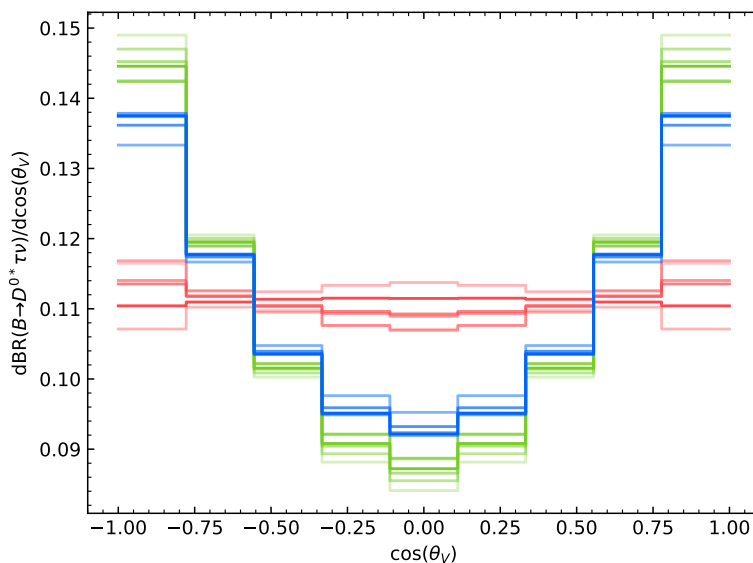
**Figure 4**. $\cos(\theta_V)$-distributions of the decay $\bar{B} \to D^{0*}\tau^-\bar{\nu}_\tau$ corresponding to three sample points and the BPs. Distributions corresponding to the same cluster are shown in the same colour (matching the colour scheme of figure 3), with the sample distributions faded out slightly.
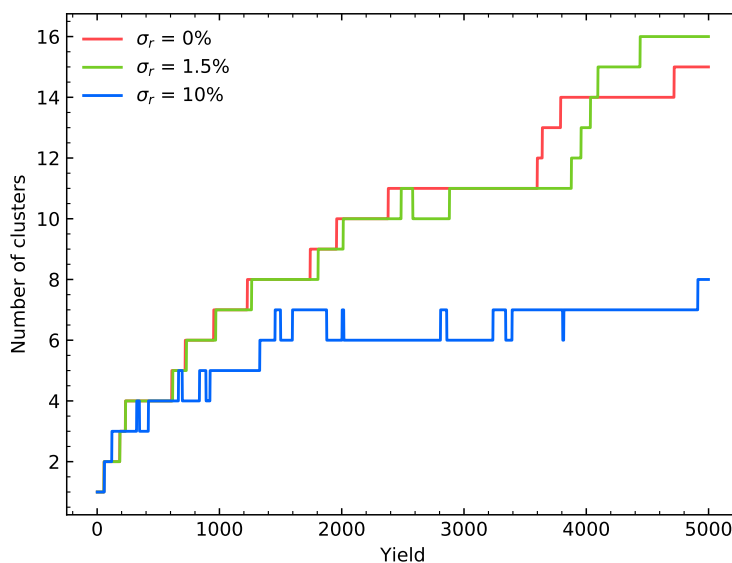


**Figure 5**. The number of clusters as a function of the signal yield for the observable dBR($B \to D^0\tau\nu$)/d$q^2$ for three different relative systematic uncertainties $\sigma_r$.

### 4.2.4 $E_\ell$-distribution of $\Gamma(\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)$

Finally, we consider $E_\ell$, the energy of the light lepton $\ell$ from the 5-body decay $\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau$. To generate the kinematic distributions in terms of the chosen set of Wilson coefficients, we use the explicit expressions given in [57] as further outlined in appendix B.
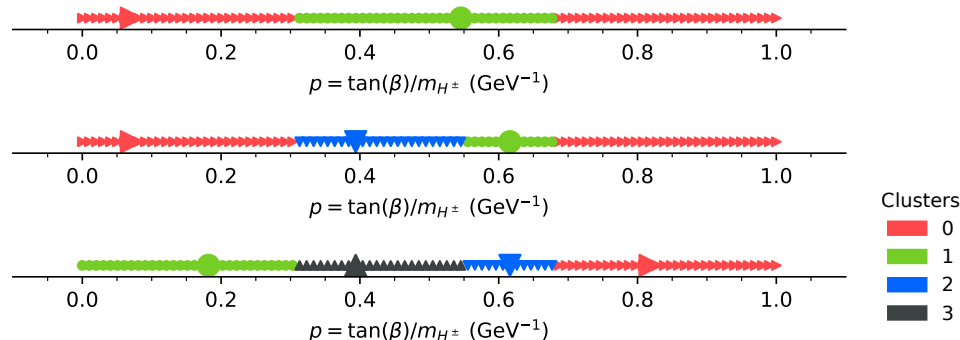
**Figure 6**. Clustering of the one-dimensional parameter space $p$ of the observable $\mathrm{d}\Gamma(\bar{B} \to D\tau^- (\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)/\mathrm{d}(E_\ell)$, with $E_\ell$ denoting the lepton energy. The NP parameter $p$ stems from a two-Higgs-doublet model. The clustering is performed assuming signal yields of 1000, 1800 and 2000 events as well as a relative uncertainty of 1% and leads to a total of two, three and four clusters respectively.

The study of this observable is motivated by the BaBar analysis in [1], where the experimental value of $R(D)$ was extracted under the assumption of a two-Higgs-doublet model. Signal and background yields were extracted with a fit to the two-dimensional $m_{\mathrm{miss}}^2, |\vec{p}_\ell^{\,*}|$ distribution. Here, $m_{\mathrm{miss}}$ and $\vec{p}_\ell^{\,*}$ denote the mass of the undetected neutrinos and the three-momentum of the light lepton in the rest frame of the $B$ meson. The shape of these distributions for signal and background were taken from MC simulations. Such MC simulations are usually assumed to be SM like. In [1] however, the (SM) MC simulations were reweighted assuming non-zero values for the parameter $p \equiv \tan\beta/m_{H^\pm}^2$ of the two-Higgs-doublet model of type II.[3] The (model-dependent) experimental values of $R(D)$ were then extracted for 20 different values of $p$. The resulting distribution of $R(D)$ measurements shows a sharp transition at $p = 0.35\,\mathrm{GeV}^{-1}$, but is otherwise relatively independent of $p$.

This result motivates to perform a clustering analysis to investigate the model dependency of the input kinematics.

In this paper, we consider the $E_\ell$-distribution, which can be taken as an approximation of $|\vec{p}_\ell^{\,*}|$ as considered in [1]. The results of our clustering analysis with respect to the parameter $p$ are shown in figure 6. The one-dimensional parameter space is clustered three times, assuming signal yields of 1000, 1800 and 2000 events as well as a relative systematic uncertainty of 1%. The first sub-figure shows the two resulting clusters for a yield of 1000 events, which coincide with the findings of [1], where two different values of $R(D)$ are obtained. However, in [1] the first value for $R(D)$ is obtained for $p \leq 0.3$ and the second one for $p \geq 0.45$, whereas the first sub-figure suggests to have the same $R(D)$ for $0.3 \leq p \leq 0.7$ and another value for the rest of the parameter space.

Increasing the yield (and thereby reducing the uncertainties on the distributions) results in more clusters in the middle region (see figure 6), again indicating that the shape of the kinematic distribution significantly changes between $0.3 < p < 0.7$. However, as can

---

[3]$\tan\beta$ denotes the ratio of the vacuum expectation values of the two Higgs fields and $m_{H^\pm}$ is the mass of the charged Higgs boson.
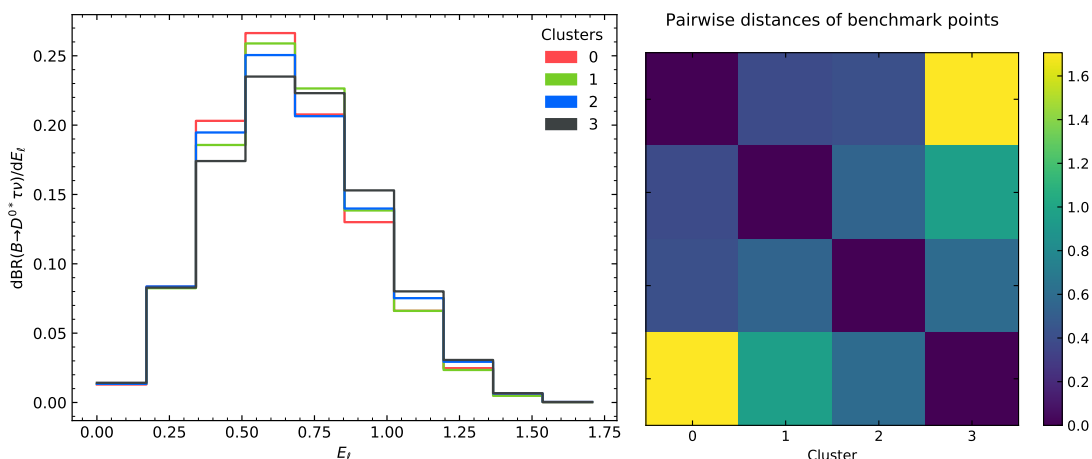
**Figure 7**. Benchmark $E_\ell$-distributions of the decay width $\Gamma(\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau)$ for a yield of 2000 events and a relative uncertainty of 1%. The matrix plot shows the pairwise distances between the kinematic distributions.

be seen from figure 7, the kinematics for low and high $p$ are still very similar, incompatible with the result of [1].

However, since [1] used the $m_{\mathrm{miss}}^2$ distribution together with the $|\vec{p}_\ell^{\,*}|$ distribution, it is not too surprising to arrive at a rather different result, as the shape of the distributions in $m_{\mathrm{miss}}^2$ could behave very differently than $|\vec{p}_\ell^{\,*}|$. Applying clustering techniques to the 2D $m_{\mathrm{miss}}^2, |\vec{p}_\ell^{\,*}|$ distribution will allow for a more thorough comparison and is left for future work.

## 5 Conclusions

In this article we discussed cluster analyses of kinematic distributions. These analyses divide the parameter space governing the distribution into subsets (*clusters*), in which all points correspond to similar kinematic distributions. Each cluster is then reduced to its most representative point (*benchmark point, BP*). Analyses relying on these kinematic distributions can then be carried out for the BPs only, rather than using the full parameter space. This can drastically reduce the required computing power and makes it easier to present numerical results as well as visualisations.

The results of the cluster analyses depend on the sampling of the parameter space, the clustering algorithm and the metric measuring differences between kinematic distributions.

This paper introduced the Python package `ClusterKinG` which implements the above steps and allows to perform clustering analyses without technical overhead. While it particularly aims to make clustering techniques more accessible for the High Energy Physics community, the software can also be applied to more general problems outside of particle physics. `ClusterKinG` is available at [github.com/clusterking](github.com/clusterking) together with usage examples and technical documentation.

We used the `ClusterKinG` package to study several kinematic distributions of the decays $\bar{B} \to D^{0(*)}\tau^-\bar{\nu}_\tau$. The $\theta_\tau$ and $\theta_V$ distribution of $\bar{B} \to D^{0*}\tau^-\bar{\nu}_\tau$ were studied, showing the clustered parameter space, the BPs as well as the corresponding distributions. A strong dependence of the $\theta_\tau$-distribution on the tensor Wilson coefficient $C_T$ has been shown, which agrees with previous findings in the literature [56].

The influence of statistical and systematic uncertainties on the clustering result is shown on the example of the $q^2$-distribution of $\bar{B} \to D^0\tau^-\bar{\nu}_\tau$.

Finally we analysed the $E_\ell$-distribution of the 5-body decay $\bar{B} \to D\tau^-(\to \ell\,\bar{\nu}_\ell\,\nu_\tau)\bar{\nu}_\tau$. The shape of this variable is an important input for some experimental measurements of $R(D)$. The resulting model dependency that was observed in [1] on the example of type II two-Higgs-doublet models should also be seen from clustering the input kinematics. While not entirely consistent with the results of [1], our simplified approach correctly hints at the significant change of $R(D)$ at $\tan\beta/m^2_{H^\pm} \approx 0.3\,\text{GeV}^{-1}$. However, a full analysis including also the shape of the $m^2_{\text{miss}}$ distributions remains to be done in order to fully compare the results with those of [1].

## Acknowledgments

## A    ClusterKinG example

The following code example shows a full `ClusterKinG` workflow: kinematic distributions for $\text{dBR}(B^+ \to D^0\tau^-\bar{\nu}_\tau)/\text{d}q^2$ are generated for a selection of sample points in Wilson space. Uncertainties are added to the data and the distributions are clustered. Finally BPs are selected and several plots are generated. Similar (and more systematic) examples can be found in the `examples` folder of the `ClusterKinG` repository.

```python
import flavio
import numpy as np
import clusterking as ck
from clusterking.maths.metric import chi2_metric


# Define kinematic function using the flavio package
def dBrdq2(w, q):
    return flavio.np_prediction("dBR/dq2(B+->Dtaunu)", w, q)


# Set up and configure Scanner
s = ck.scan.WilsonScanner(scale=5, eft="WET", basis="flavio")
# Set kinematic function
s.set_dfunction(
    dBrdq2,
```

```
17        binning=np.linspace(3.2, 11.6, 10),
18        normalize=True
19    )
20    # Set sampling points in Wilson space
21    s.set_spoints_equidist({
22        "CVL_bctaunutau": (-0.5, 0.5, 10),
23        "CSL_bctaunutau": (-0.5, 0.5, 10),
24        "CT_bctaunutau": (-0.1, 0.1, 10)
25    })
26
27    # Run scanner and add errors
28    d = ck.DataWithErrors()     # Create data object to write results to
29    r = s.run(d)                # Run scanner
30    r.write()                   # Write results back to data object
31    d.add_err_poisson(1000)     # statistical uncertainties
32    d.add_rel_err_uncorr(0.1)   # 0.1% relative system uncertainties,
          uncorrelated
33
34    # Clustering
35    c = ck.cluster.HierarchyCluster()  # Initialize worker class
36    c.set_metric(chi2_metric)
37    c.set_max_d(1)     # "Cut off" value for hierarchy
38    r = c.run(d)       # Run clustering on d
39    r.write()          # Write results back to data object
40
41    # Benchmarking
42    b = ck.Benchmark()  # Initialize worker class
43    b.set_metric(chi2_metric)
44    r = b.run(d)        # Run benchmarking
45    r.write()           # Write results back to data object
46
47    # Optional: Save data (kinematic distributions, clusters, BPs, ...)
48    d.write("btaunu_q2.sql")
49
50    # Find closest benchmark point to new parameter point
51    # Similar function for spoints: find_closest_spoints
52    d.find_closest_bpoints(
53    {
54        "CVL_bctaunutau": 0.1,
55        "CSL_bctaunutau": 0.1,
56        "CT_bctaunutau": 0.1,
57    },
58    n=1
59    )
60
61    # Generate plots similar to figures 1, 2, 3, 4, 7 respectively
62    d.plot_clusters_scatter(["CVL_bctaunutau", "CSL_bctaunutau"])
63    d.plot_dist_box()
64    d.plot_clusters_scatter(["CT_bctaunutau", "CVL_bctaunutau", "
          CSL_bctaunutau"])
65    d.plot_dist()
66    d.plot_bpoint_distance_matrix()
```

# B   Differential decay rates

The differential decay rates for the decays $\bar{B} \to D^{0(*)} \tau^- \bar{\nu}_\tau$ are given by:

$$\frac{\mathrm{d}^2 \Gamma(B \to D^{0(*)} \tau \nu)}{\mathrm{d}q^2 \, \mathrm{d}(\cos \theta_\ell)} = \frac{3}{8}(I_1^c + 2I_1^s) + \cos(\theta_\ell)\frac{3}{8}(I_6^c + 2I_6^s) + \cos(2\theta_\ell)\frac{3}{8}(I_2^c + 2I_2^s), \quad \text{(B.1)}$$

$$\frac{\mathrm{d}^2 \Gamma(B \to D^{0*} \tau \nu)}{\mathrm{d}q^2 \, \mathrm{d}(\cos \theta_V)} = -\cos^2(\theta_V)\frac{3}{8}(-3I_1^c + I_2^c) - \sin^2(\theta_V)\frac{3}{8}(-3I_1^s + I_2^s), \quad \text{(B.2)}$$

$$\frac{\mathrm{d}\Gamma(B \to D^{0*} \tau \nu)}{\mathrm{d}q^2} = \frac{3}{4}(2I_1^s + I_1^c) - \frac{1}{4}(2I_2^s + I_2^c), \quad \text{(B.3)}$$

where $q^2$ is the invariant mass of the lepton pair and $I_i^j$ denote the angular coefficients, which depend on $q^2$ and the Wilson coefficients. The angle $\theta_\ell$ is defined as the angle between the direction of the tau in the dilepton rest frame and the direction of the dilepton in the $B$ rest frame, whereas $\theta_V$ is defined as the angle between the direction of the $D^{0*}$ meson in the dilepton rest frame and the direction of the $D^{0*}$ in the $B$ rest frame. A general discussion of the kinematics of semileptonic meson decays (in the context of $B \to K^* \bar{\ell}\ell$) together with the angular coefficients can be found in [58]. For $\bar{B} \to D^{0(*)} \tau^- \bar{\nu}_\tau$ decays the corresponding distributions are discussed for example in [57, 59]. Our implementation of the above decay rates can be found at `github.com/clusterking/clusterking_physics`.

## References

[1] BABAR collaboration, *Evidence for an excess of $\bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau$ decays*, *Phys. Rev. Lett.* **109** (2012) 101802 [`arXiv:1205.5442`] [INSPIRE].

[2] A. Carvalho, M. Dall'Osso, T. Dorigo, F. Goertz, C.A. Gottardo and M. Tosi, *Higgs Pair Production: Choosing Benchmarks With Cluster Analysis*, *JHEP* **04** (2016) 126 [`arXiv:1507.02245`] [INSPIRE].

[3] A. Carvalho et al., *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*, `arXiv:1608.06578` [INSPIRE].

[4] A. Carvalho, F. Goertz, K. Mimasu, M. Gouzevitch and A. Aggarwal, *On the reinterpretation of non-resonant searches for Higgs boson pairs*, `arXiv:1710.08261` [INSPIRE].

[5] CMS collaboration, *Search for two Higgs bosons in final states containing two photons and two bottom quarks in proton-proton collisions at 8 TeV*, *Phys. Rev.* **D 94** (2016) 052012 [`arXiv:1603.06896`] [INSPIRE].

[6] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [`arXiv:1008.4884`] [INSPIRE].

[7] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators I: Formalism and lambda Dependence*, *JHEP* **10** (2013) 087 [`arXiv:1308.2627`] [INSPIRE].

[8] E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators II: Yukawa Dependence*, *JHEP* **01** (2014) 035 [arXiv:1310.4838] [INSPIRE].

[9] R. Alonso, E.E. Jenkins, A.V. Manohar and M. Trott, *Renormalization Group Evolution of the Standard Model Dimension Six Operators III: Gauge Coupling Dependence and Phenomenology*, *JHEP* **04** (2014) 159 [arXiv:1312.2014] [INSPIRE].

[10] A. Celis, J. Fuentes-Martin, A. Vicente and J. Virto, *DsixTools: The Standard Model Effective Field Theory Toolkit*, *Eur. Phys. J.* **C 77** (2017) 405 [arXiv:1704.04504] [INSPIRE].

[11] J. Aebischer, A. Crivellin, M. Fael and C. Greub, *Matching of gauge invariant dimension-six operators for $b \to s$ and $b \to c$ transitions*, *JHEP* **05** (2016) 037 [arXiv:1512.02830] [INSPIRE].

[12] E.E. Jenkins, A.V. Manohar and P. Stoffer, *Low-Energy Effective Field Theory below the Electroweak Scale: Operators and Matching*, *JHEP* **03** (2018) 016 [arXiv:1709.04486] [INSPIRE].

[13] W. Dekens and P. Stoffer, *Low-energy effective field theory below the electroweak scale: matching at one loop*, *JHEP* **10** (2019) 197 [arXiv:1908.05295] [INSPIRE].

[14] J. Aebischer, M. Fael, C. Greub and J. Virto, *B physics Beyond the Standard Model at One Loop: Complete Renormalization Group Evolution below the Electroweak Scale*, *JHEP* **09** (2017) 158 [arXiv:1704.06639] [INSPIRE].

[15] E.E. Jenkins, A.V. Manohar and P. Stoffer, *Low-Energy Effective Field Theory below the Electroweak Scale: Anomalous Dimensions*, *JHEP* **01** (2018) 084 [arXiv:1711.05270] [INSPIRE].

[16] J.C. Criado, *MatchingTools: a Python library for symbolic effective field theory calculations*, *Comput. Phys. Commun.* **227** (2018) 42 [arXiv:1710.06445] [INSPIRE].

[17] J. Aebischer, J. Kumar and D.M. Straub, *Wilson: a Python package for the running and matching of Wilson coefficients above and below the electroweak scale*, *Eur. Phys. J.* **C 78** (2018) 1026 [arXiv:1804.05033] [INSPIRE].

[18] J. Aebischer et al., *WCxf: an exchange format for Wilson coefficients beyond the Standard Model*, *Comput. Phys. Commun.* **232** (2018) 71 [arXiv:1712.05298] [INSPIRE].

[19] J. Aebischer, J. Kumar, P. Stangl and D.M. Straub, *A Global Likelihood for Precision Constraints and Flavour Anomalies*, *Eur. Phys. J.* **C 79** (2019) 509 [arXiv:1810.07698] [INSPIRE].

[20] J.C. Criado, *BasisGen: automatic generation of operator bases*, *Eur. Phys. J.* **C 79** (2019) 256 [arXiv:1901.03501] [INSPIRE].

[21] B. Gripaios and D. Sutherland, *DEFT: A program for operators in EFT*, *JHEP* **01** (2019) 128 [arXiv:1807.07546] [INSPIRE].

[22] I. Brivio et al., *Computing Tools for the SMEFT*, in *Computing Tools for the SMEFT*, J. Aebischer et al. eds., arXiv:1910.11003 [INSPIRE].

[23] D.M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, arXiv:1810.08132 [INSPIRE].

[24] W. Porod, F. Staub and A. Vicente, *A Flavor Kit for BSM models*, *Eur. Phys. J.* **C 74** (2014) 2992 [arXiv:1405.1434] [INSPIRE].

[25] F. Mahmoudi, *SuperIso: A Program for calculating the isospin asymmetry of $B \to K^*\gamma$ in the MSSM*, *Comput. Phys. Commun.* **178** (2008) 745 [arXiv:0710.2067] [INSPIRE].

[26] W. Porod, *SPheno, a program for calculating supersymmetric spectra, SUSY particle decays and SUSY particle production at $e^+e^-$ colliders*, *Comput. Phys. Commun.* **153** (2003) 275 [hep-ph/0301101] [INSPIRE].

[27] W. Porod and F. Staub, *SPheno 3.1: Extensions including flavour, CP-phases and models beyond the MSSM*, *Comput. Phys. Commun.* **183** (2012) 2458 [arXiv:1104.1573] [INSPIRE].

[28] J.A. Evans and D. Shih, *FormFlavor Manual*, arXiv:1606.00003 [INSPIRE].

[29] A. Dedes, M. Paraskevas, J. Rosiek, K. Suxho and L. Trifyllis, *SmeftFR — Feynman rules generator for the Standard Model Effective Field Theory*, *Comput. Phys. Commun.* **247** (2020) 106931 [arXiv:1904.03204] [INSPIRE].

[30] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [arXiv:1709.06492] [INSPIRE].

[31] W. Altmannshofer, P.S. Bhupal Dev and A. Soni, *$R_{D^{(*)}}$ anomaly: A possible hint for natural supersymmetry with R-parity violation*, *Phys. Rev.* **D 96** (2017) 095010 [arXiv:1704.06659] [INSPIRE].

[32] C. Murgui, A. Peñuelas, M. Jung and A. Pich, *Global fit to $b \to c\tau\nu$ transitions*, *JHEP* **09** (2019) 103 [arXiv:1904.09311] [INSPIRE].

[33] R.-X. Shi, L.-S. Geng, B. Grinstein, S. Jäger and J. Martin Camalich, *Revisiting the new-physics interpretation of the $b \to c\tau\nu$ data*, *JHEP* **12** (2019) 065 [arXiv:1905.08498] [INSPIRE].

[34] D. Bečirević, M. Fedele, I. Nišandžić and A. Tayduganov, *Lepton Flavor Universality tests through angular observables of $\overline{B} \to D^{(*)}\ell\overline{\nu}$ decay modes*, arXiv:1907.02257 [INSPIRE].

[35] J.D. Gómez, N. Quintero and E. Rojas, *Charged current $b \to c\tau\bar{\nu}_\tau$ anomalies in a general $W'$ boson scenario*, *Phys. Rev.* **D 100** (2019) 093003 [arXiv:1907.08357] [INSPIRE].

[36] M. Algueró et al., *Emerging patterns of New Physics with and without Lepton Flavour Universal contributions*, *Eur. Phys. J.* **C 79** (2019) 714 [arXiv:1903.09578] [INSPIRE].

[37] M. Ciuchini et al., *New Physics in $b \to s\ell^+\ell^-$ confronts new data on Lepton Universality*, *Eur. Phys. J.* **C 79** (2019) 719 [arXiv:1903.09632] [INSPIRE].

[38] A. Datta, J. Kumar and D. London, *The B anomalies and new physics in $b \to se^+e^-$*, *Phys. Lett.* **B 797** (2019) 134858 [arXiv:1903.10086] [INSPIRE].

[39] J. Aebischer, W. Altmannshofer, D. Guadagnoli, M. Reboud, P. Stangl and D.M. Straub, *B-decay discrepancies after Moriond 2019*, arXiv:1903.10434 [INSPIRE].

[40] F.U. Bernlochner, S. Duell, Z. Ligeti, M. Papucci and D.J. Robinson, *Das ist der HAMMER: Consistent new physics interpretations of semileptonic decays*, arXiv:2002.00020 [INSPIRE].

[41] S. Duell, F. Bernlochner, Z. Ligeti, M. Papucci and D. Robinson, *HAMMER: Reweighting tool for simulated data samples*, PoS(ICHEP2016)1074.

[42] F.J. Massey, *The Kolmogorov-Smirnov test for goodness of fit*, *J. Am. Statist. Assoc.* **46** (1951) 68.

[43] A.N. Pettitt, *A two-sample Anderson-Darling rank statistic*, *Biometrika* **63** (1976) 161.

[44] S.S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, *Ann. Math. Statist.* **9** (1938) 60.

[45] J.A. Hartigan, *Clustering algorithms*, *Wiley series in probability and mathematical statistics*, Wiley, New York U.S.A. (1975).

[46] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume 1: Statistics*, Berkeley U.S.A. (1967), pg. 281, University of California Press, Berkeley U.S.A. (1967).

[47] L. Kaufman and P. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, John Wiley & Sons, New York U.S.A. (1990).

[48] W. McKinney, *Data structures for statistical computing in python*, in *Proceedings of the 9th Python in Science Conference*, S. van der Walt and J. Millman eds., Austin U.S.A. (2010), pg. 51.

[49] E. Jones et al., *SciPy: Open source scientific tools for Python*, (2001).

[50] A. Celis, M. Jung, X.-Q. Li and A. Pich, *Scalar contributions to $b \to c(u)\tau\nu$ transitions*, *Phys. Lett.* **B 771** (2017) 168 [arXiv:1612.07757] [INSPIRE].

[51] M. Blanke, A. Crivellin, S. de Boer, T. Kitahara, M. Moscati, U. Nierste et al., *Impact of polarization observables and $B_c \to \tau\nu$ on new physics explanations of the $b \to c\tau\nu$ anomaly*, *Phys. Rev.* **D 99** (2019) 075006 [arXiv:1811.09603] [INSPIRE].

[52] S. Bhattacharya, S. Nandi and S. Kumar Patra, *$b \to c\tau\nu_\tau$ Decays: a catalogue to compare, constrain and correlate new physics effects*, *Eur. Phys. J.* **C 79** (2019) 268 [arXiv:1805.08222] [INSPIRE].

[53] D. Bečirević, I. Doršner, S. Fajfer, N. Košnik, D.A. Faroughy and O. Sumensari, *Scalar leptoquarks from grand unified theories to accommodate the B-physics anomalies*, *Phys. Rev.* **D 98** (2018) 055003 [arXiv:1806.05689] [INSPIRE].

[54] B. Bhattacharya, A. Datta, S. Kamali and D. London, *CP Violation in $\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu$*, *JHEP* **05** (2019) 191 [arXiv:1903.02567] [INSPIRE].

[55] BELLE, BELLE-II collaboration, *Semitauonic B decays at Belle/Belle II*, in *10th International Workshop on the CKM Unitarity Triangle (CKM 2018) Heidelberg, Germany, September 17-21, 2018*, 2019, arXiv:1901.06380 [INSPIRE].

[56] M. Jung and D.M. Straub, *Constraining new physics in $b \to c\ell\nu$ transitions*, *JHEP* **01** (2019) 009 [arXiv:1801.01112] [INSPIRE].

[57] R. Alonso, A. Kobach and J. Martin Camalich, *New physics in the kinematic distributions of $\bar{B} \to D^{(*)}\tau^-(\to \ell^-\bar{\nu}_\ell\nu_\tau)\bar{\nu}_\tau$*, *Phys. Rev.* **D 94** (2016) 094021 [arXiv:1602.07671] [INSPIRE].

[58] J. Gratrex, M. Hopfer and R. Zwicky, *Generalised helicity formalism, higher moments and the $B \to K_{J_K}(\to K\pi)\bar{\ell}_1\ell_2$ angular distributions*, *Phys. Rev.* **D 93** (2016) 054008 [arXiv:1506.03970] [INSPIRE].

[59] D. Becirevic, S. Fajfer, I. Nisandzic and A. Tayduganov, *Angular distributions of $\bar{B} \to D^{(*)}\ell\bar{\nu}_\ell$ decays and search of New Physics*, *Nucl. Phys.* **B 946** (2019) 114707 [arXiv:1602.03030] [INSPIRE].