LPHE, Lausanne, June 24 2019

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

Presented by Jason Aebischer

Excellence Cluster
Technische Universität München

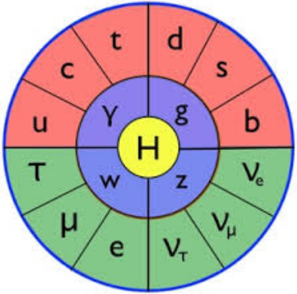
Outline

1. Motivation
2. Clustering
3. ClusterKinG
4. $\bar{B} \to D\tau^-\bar{\nu}_\tau$
5. Summary

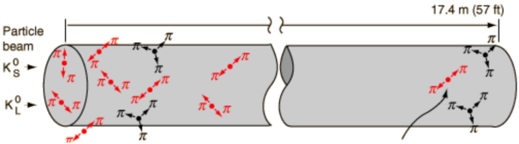
Based on: JA, Thomas Kuhr, Kilian Lieret [1907.xxxx]

Outline

SM and New Physics



CPV in ε'/ε



Measurement

NA48: hep-ex/0208009, KTeV: hep-ex/0208007

$(\varepsilon'/\varepsilon)_{\text{exp}} = (16.6 \pm 2.3) \times 10^{-4}$

SM prediction

Lattice: $(1.4 \pm 6.9) \times 10^{-4}$ RBC-UKQCD: 1502.00263, 1505.07863

DQCD: $\leq (6 \pm 2.4) \times 10^{-4}$ Buras, Gérard 1507.06326

χPT: $(15 \pm 7) \times 10^{-4}$ Gisbert, Pich 1712.06147

$b \to s\,\mu^+\mu^-$ anomaly

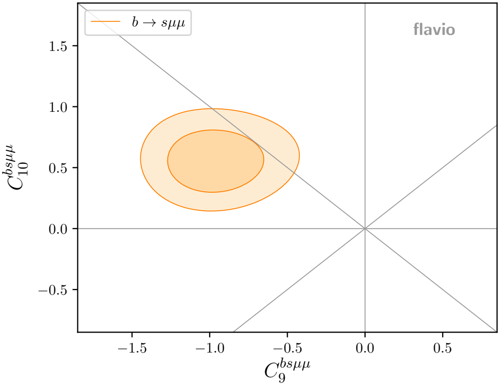
Measurements

Angular observable P_5' in $B \to K^*\mu^+\mu^-$. LHCb, 1512.04442

Branching ratios of $B \to K\mu^+\mu^-$, $B \to K^*\mu^+\mu^-$, and $B_s \to \phi\mu^+\mu^-$.

LHCb, 1403.8044, 1506.08777, 1606.04731

$$O_9^{bs\mu\mu} = (\bar{s}\gamma_\nu P_L b)(\bar{\mu}\gamma^\nu \mu)$$
$$O_{10}^{bs\mu\mu} = (\bar{s}\gamma_\nu P_L b)(\bar{\mu}\gamma^\nu \gamma_5 \mu)$$



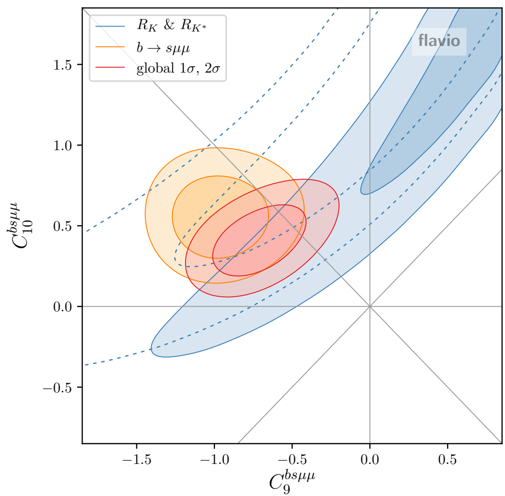
LFU violation in neutral currents

Measurements

$R_K^{[1,6]}$, $R_{K^*}^{[0.045,1.1]}$, $R_{K^*}^{[1.1,6]}$

LHCb: 1406.6482, 1705.05802, 1903.09252, Belle: 1904.0244

$$R_{K^{(*)}} = \frac{BR(B \to K^{(*)}\mu^+\mu^-)}{BR(B \to K^{(*)}e^+e^-)}$$



JA,Altmannshofer,Guadagnoli,Reboud,Stangl,Straub: 1903.10434

LFU violation in charged currents

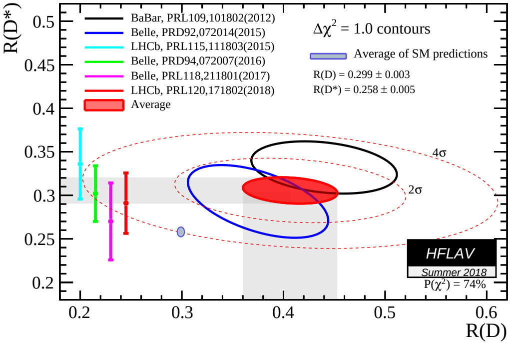
Measurement

R_D and R_{D^*}

BaBar, 1205.5442, 1303.0571, LHCb, 1506.08614, 1708.08856
Belle, 1507.03233, 1607.07923, 1612.00529

$$R_{D^{(*)}} = \frac{BR(B \to D^{(*)}\tau\nu)}{BR(B \to D^{(*)}\ell\nu)}$$

$$\ell \in \{e, \mu\}$$



HFLAV, 1612.07233

NP in kinematic distributions

Shapes influenced by NP parameters

$\longrightarrow$ Wilson coefficients

Theory

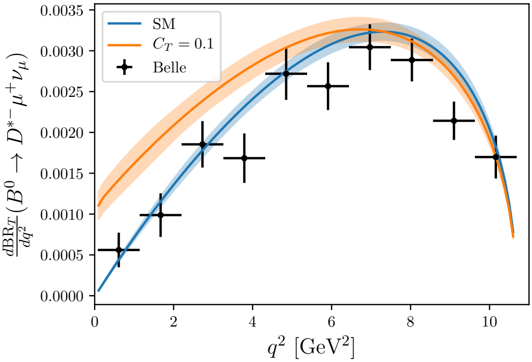
Predict kinematic distribution

Experiment

Exclusions limits (under NP assumptions)

Example: Theory prediction

$$O_T = (\bar{c}\sigma_{\rho\nu} P_L b)(\bar{\mu}\sigma^{\rho\nu} P_L \nu_\mu)$$



Jung, Straub: 1612.07233

Example: Measurement

BaBar

1205.5442

$R(D)$ and $R(D^*)$

SM values

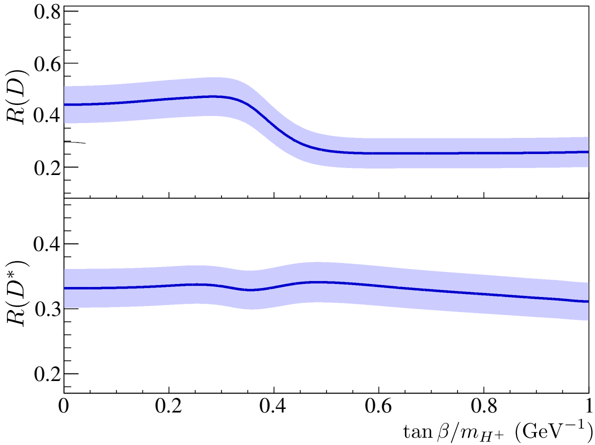
From $m^2_{\text{miss}} - |\vec{p}_\ell|$ distribution

Two-Higgs-Doublet model

Charged Higgs mass: m_{H^+}
Mixing angle: β

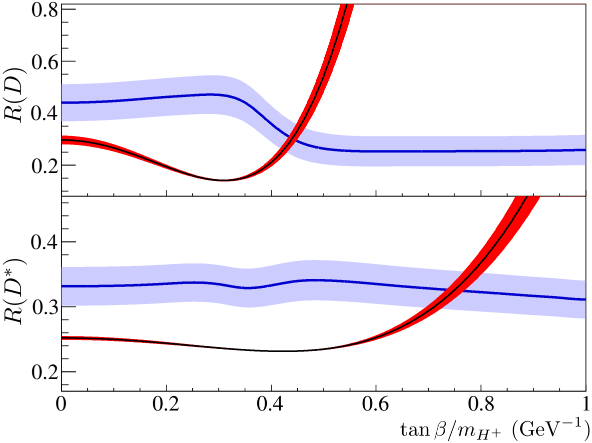
Example: Measurement

BaBar: 1205.5442



Example: Measurement

BaBar: 1205.5442



Caveats

Many NP parameters

SM effective theory: 2499

Weak effective theory: 5963

Model-dependent extraction

Relies on NP assumptions

Mutlidimensional problem

Hard to visualise

Solution: Clustering

Parameter space

Devided into regions of similar kinematic final states

Clustering

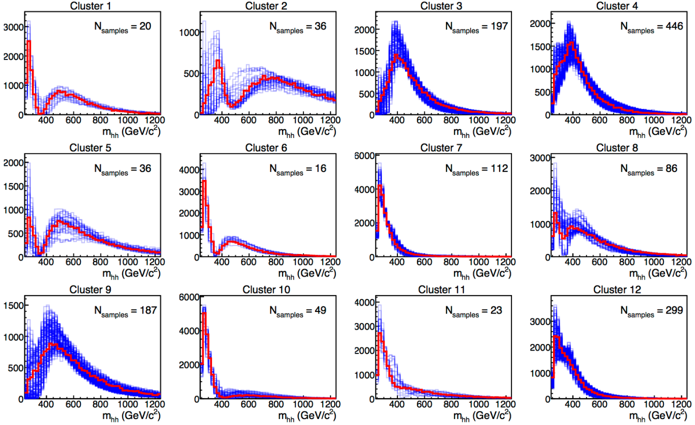
Group all similar points into a cluster

Benchmark point

Most representative point in cluster $\rightarrow$ used for analysis

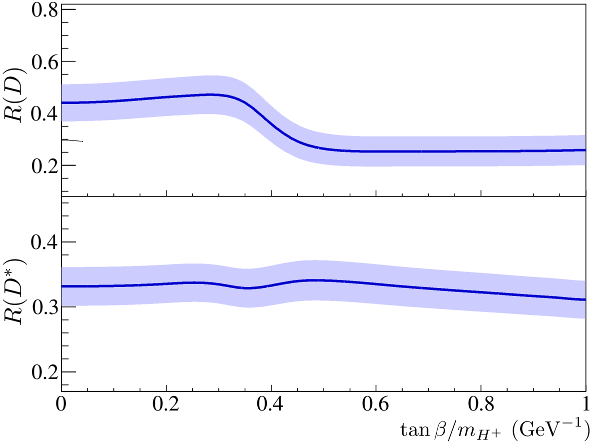
Example: Higgs pair production

1507.02245, 1608.06578, 1710.08261

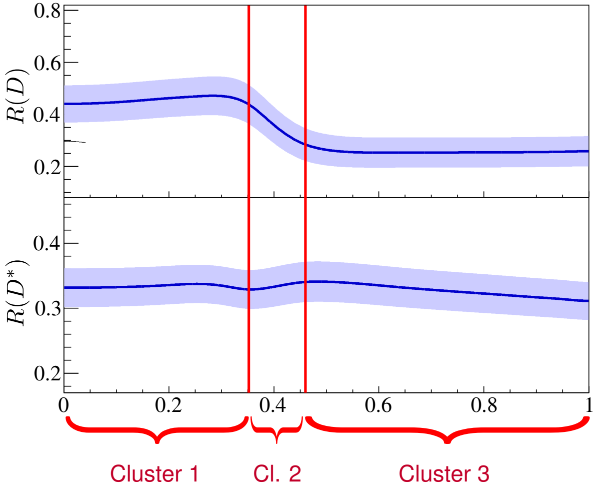


Carvalho, Dall'Osso, Dorigo, Goertz, Gottardo, Tosi: 1507.02245

Example: $R(D^{(*)})$ from BaBar



Example: $R(D^{(*)})$ from BaBar



Outline

Clustering: Prerequisites

Similarity between distributions

Metric: χ^2 test, Kolmogorov test, etc.

Benchmark points

Criterion: closest to all other points

Clustering algorithm

Example: Hierarchical, K-Means

Clustering algorithm: Hiearchical

Step 1

Each point in own cluster

Step 2

Merge nearest points

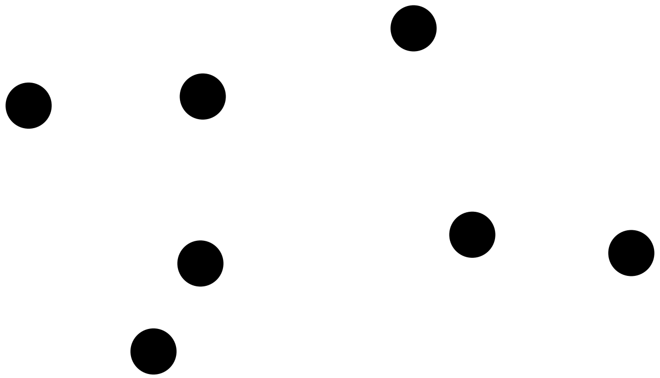
Step 3

Compute benchmark point of each cluster

Step 4 and following

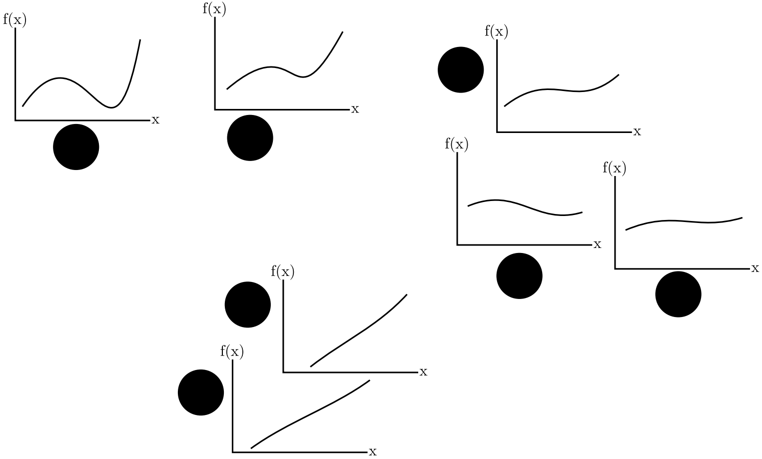
Repeat 2 and 3 until desired N_{clusters} reached

Example: Hierarchical Clustering



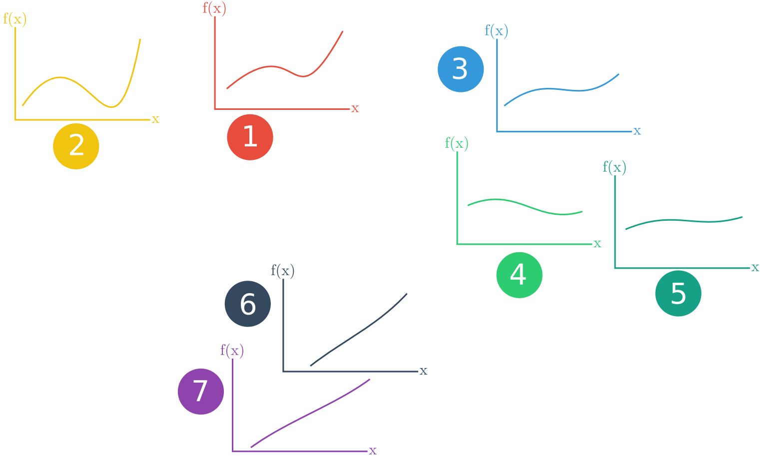
points in the parameter space

Example: Hierarchical Clustering



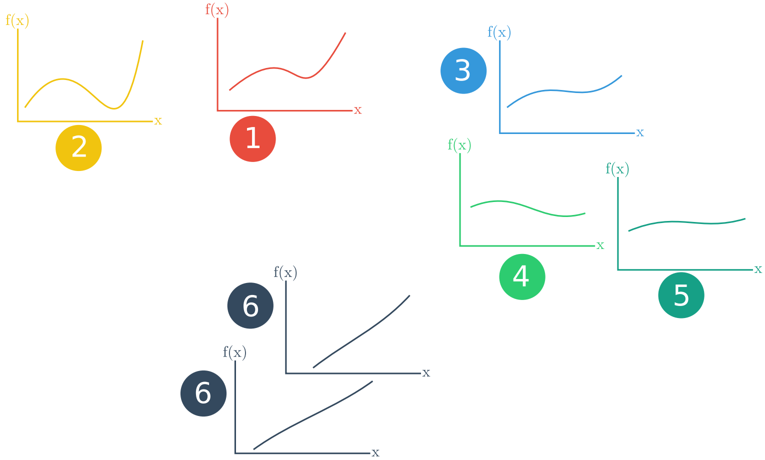
compute corresponding distributions

Example: Hierarchical Clustering



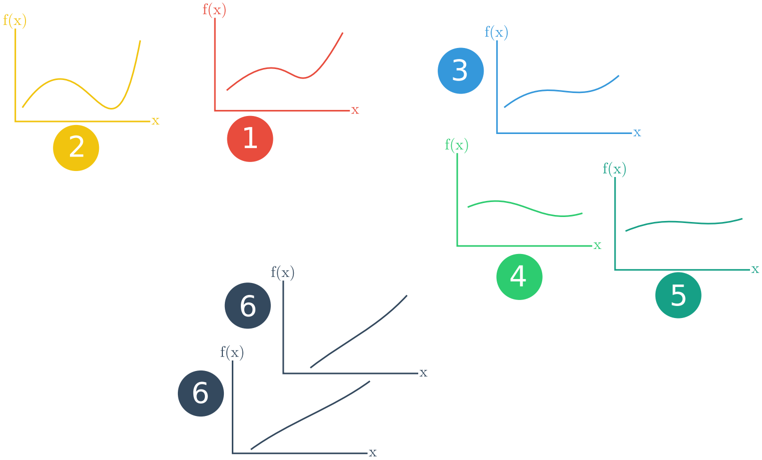
Step 1: Every point in own cluster

Example: Hierarchical Clustering



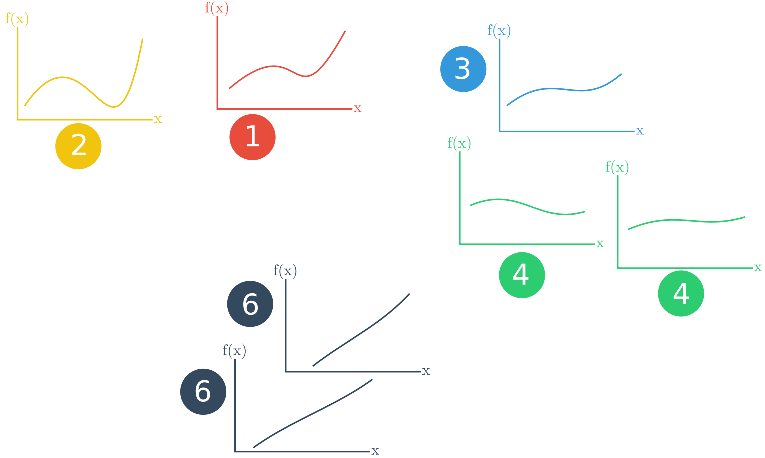
Step 2: Distributions from 6 and 7 merged

Example: Hierarchical Clustering



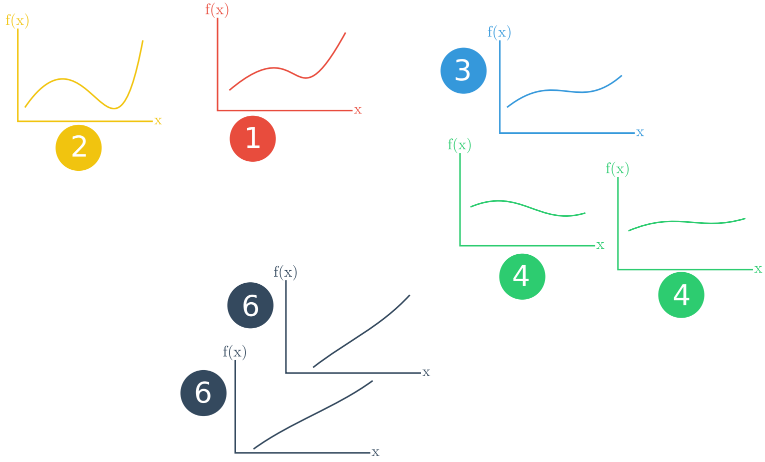
Step 3: Compute Benchmark points

Example: Hierarchical Clustering



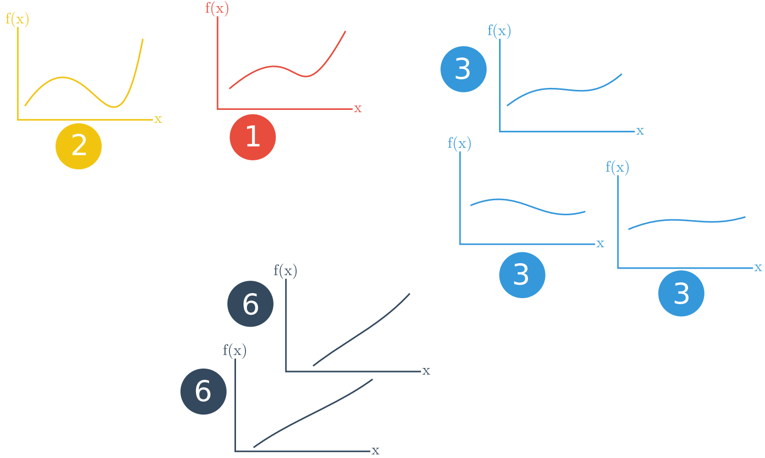
Step 4: Distributions from 4 and 5 merged

Example: Hierarchical Clustering



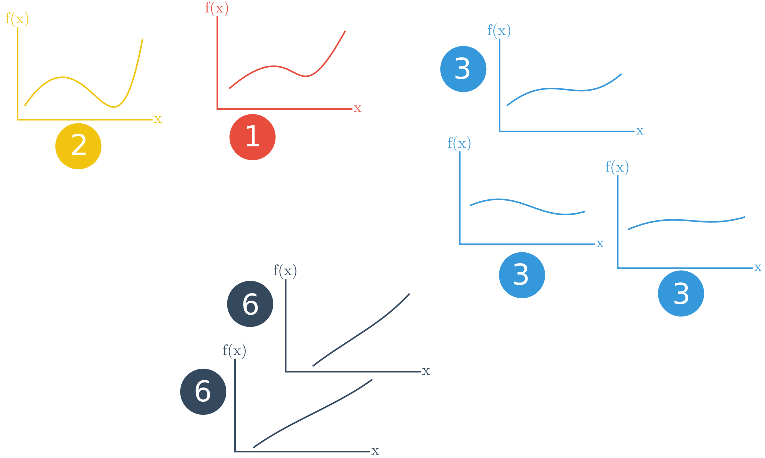
Step 5: Compute Benchmark points

Example: Hierarchical Clustering



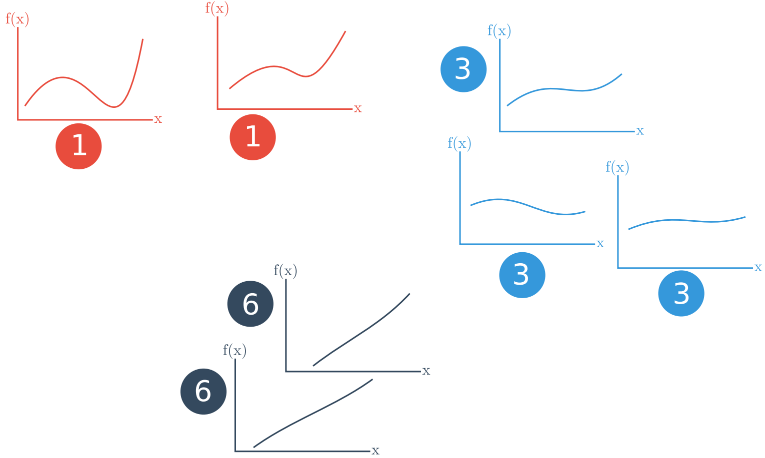
Step 6: Distributions from 4 and 3 merged

Example: Hierarchical Clustering



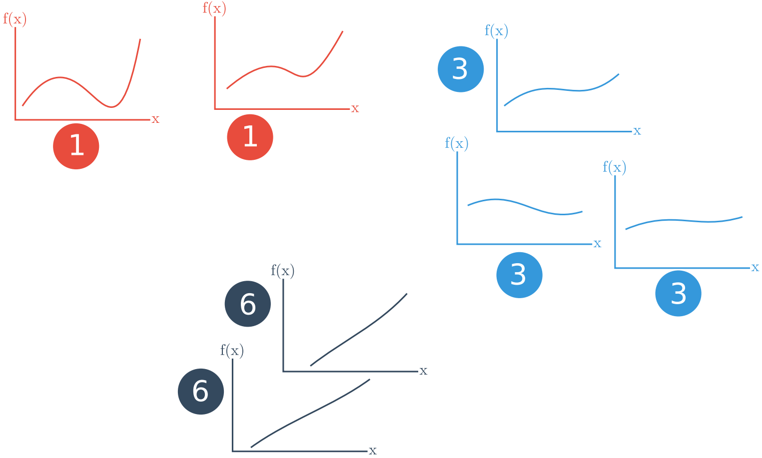
Step 7: Compute Benchmark points

Example: Hierarchical Clustering



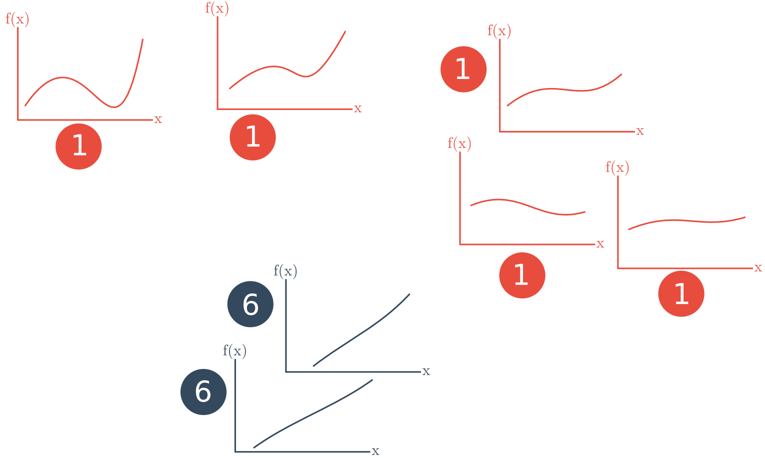
Step 8: Distributions from 1 and 2 merged

Example: Hierarchical Clustering



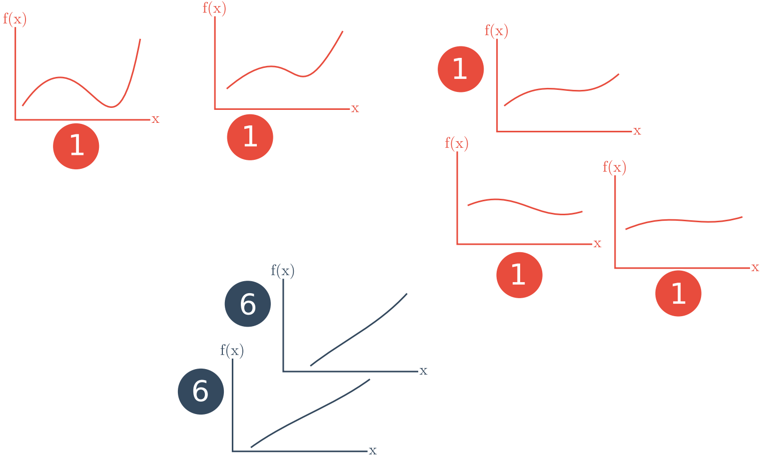
Step 9: Compute Benchmark points

Example: Hierarchical Clustering



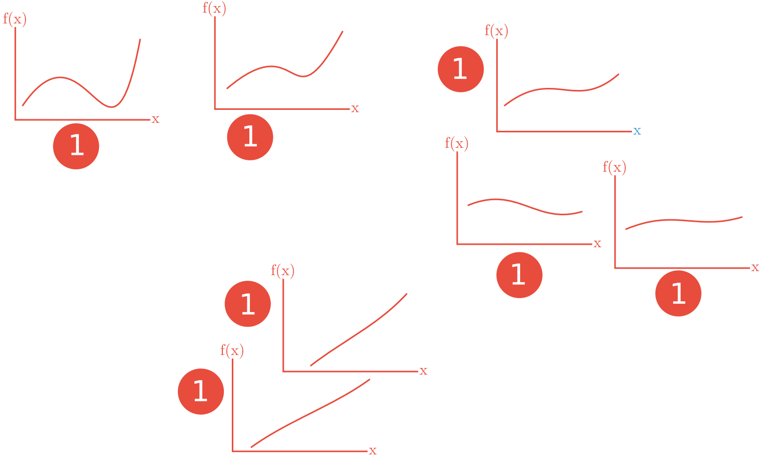
Step 10: Distributions from 1 and 3 merged

Example: Hierarchical Clustering



Step 11: Compute Benchmark points

Example: Hierarchical Clustering



Step 12: 1 cluster remaining

Problems

Dimensionality

Large number of points, many parameters

Repetition

Redo clustering for new analysis

Solution

Software: ClusterKinG

Outline

CLUSTER KING

Clustering of Kinematic Graphs

Python module

Automatic clustering of parameter space

Open source

Actively developed on GitHub: `https://github.com/clusterking`

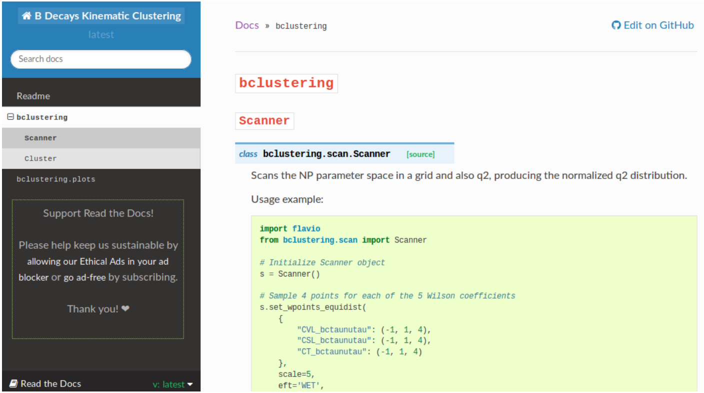
Turorials

Interactive using `jupyter notebooks`

In browser using `binder`

Documentation

Read the Docs:



- manual: coming soon

Features

Scanner

Scan automatically over parameter space

Clustering

Algorithms: Hierarchical, K-Means

Costumizable

Various metrics, clustering criterions

Visualisation

Clusters

2D or 3D cuts through parameter space

Benchmark points

Boxplots, histograms, bundle plots

Errors

Systematic, statistical

Outline

$\bar{B} \to D\tau^-(\to \ell^-\bar{\nu}_\ell\nu_\tau)\bar{\nu}_\tau$

Charged current

Only 5 Wilson coefficients

Distributions known

Alonso/Camalich/Kobach: 1602.07671

$q^2, cos(\theta), E_\ell$

Parameter ranges

can be constrained

Parameter space

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

V_{cb} = CKM element G_F = Fermi coupling constant

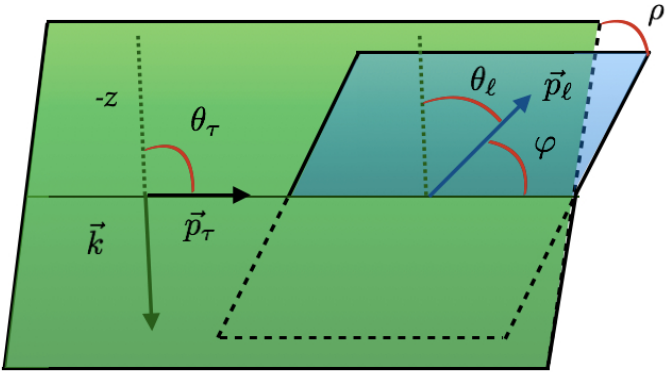
$$O_{VL} = (\bar{s}\gamma_\mu P_L b)(\bar{\tau}\gamma^\mu P_L \nu_\tau)\,, \qquad O_{VR} = (\bar{s}\gamma_\mu P_R b)(\bar{\tau}\gamma^\mu P_L \nu_\tau)\,,$$

$$O_{SL} = (\bar{s} P_L b)(\bar{\tau} P_L \nu_\tau)\,, \qquad O_{SR} = (\bar{s} P_R b)(\bar{\tau} P_L \nu_\tau)\,,$$

$$O_T = (\bar{s}\sigma_{\mu\nu} P_L b)(\bar{\tau}\sigma^{\mu\nu} P_L \nu_\tau)\,.$$

Distributions

Alonso/Camalich/Kobach: 1602.07671



$$\frac{d^3\Gamma_5}{dq^2\,dE_\ell\,d(\cos\theta_\ell)} = B[\tau_\ell]\,\frac{G_F^2|V_{cb}|^2}{32\pi^3}\,\frac{|\vec{k}|}{m_B^2}\left(1-\frac{m_\tau^2}{q^2}\right)^2\frac{E_\ell^2}{m_\tau^3}\left[I_0(q^2,E_\ell)+I_1(q^2,E_\ell)\cos\theta_\ell+I_2(q^2,E_\ell)\cos\theta_\ell\right]$$

Parameter regions

CP conserving limit

$C_i \in \mathbb{R}$

Linear combinations

only $C_{VL} + C_{VR}$ and $C_{SL} + C_{SR}$

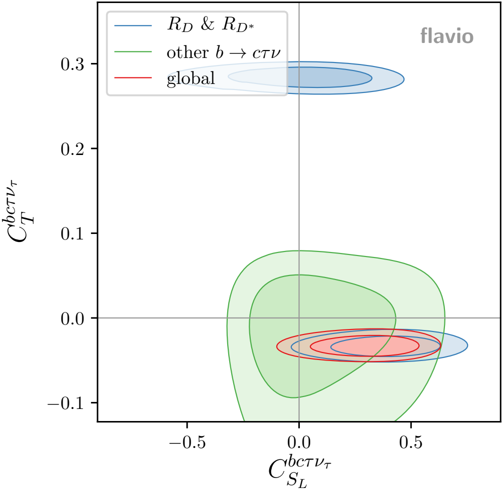
Perturbativiy

$C_i \sim \mathcal{O}(1)$

Parameter regions

C_T, C_{SL} **Belle: 1901.06380**

Constrained from F_L: $C_T \in [-0.1, 0.1]$, $C_{SL} \in [-0.5, 0.5]$



JA/Kumar/Stangl/Straub: 1810.07698

Setting up scan

```python
import clusterking as ck

d = ck.Data()
s = ck.WilsonScanner(scale=5, eft='WET', basis='flavio')
```

Setting up scan

```
import clusterking as ck

d = ck.Data()
s = ck.WilsonScanner(scale=5, eft='WET', basis='flavio')

s.set_dfunction(
    dGq2,
    binning=linspace(q2min, q2max, 10),
)
```

Setting up scan

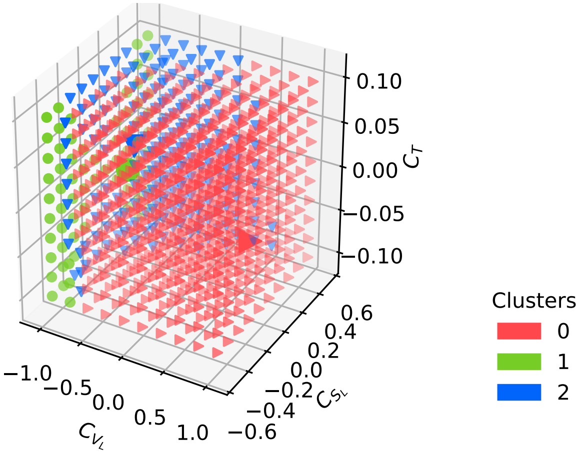
```
import clusterking as ck

d = ck.Data()
s = ck.WilsonScanner(scale=5, eft='WET', basis='flavio')

s.set_dfunction(
    dGq2,
    binning=linspace(q2min, q2max, 10),
)

s.set_spoints_equidist({
    "CVL_bctaunutau": (-1, 1, 10),
    "CSL_bctaunutau": (-0.5, 0.5, 10),
    "CT_bctaunutau": (-0.1, 0.1, 10)
})

s.run(d)
```

Clustering

```
c = ck.cluster.HierarchyCluster(d)
c.build_hierarchy()     # Build up clustering hierarchy
c.write()               # Write results to d
d.plot_clusters_scatter(["CVL_bctaunutau","CSL_bctaunutau","
    CT_bctaunutau"])
```

Clustering

```
c = ck.cluster.HierarchyCluster(d)
c.build_hierarchy()      # Build up clustering hierarchy
c.write()                # Write results to d
d.plot_clusters_scatter(["CVL_bctaunutau","CSL_bctaunutau","
    CT_bctaunutau"])
```

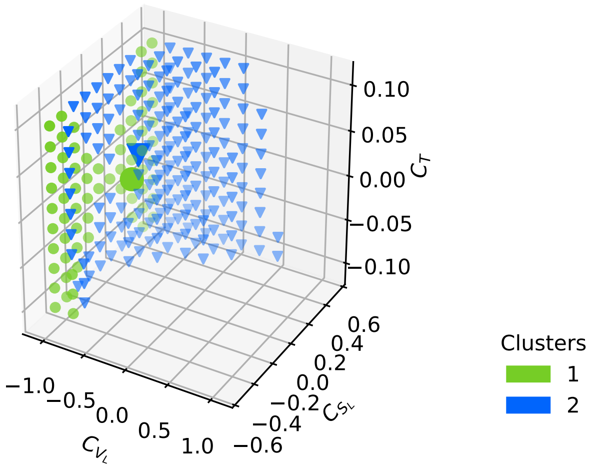


3D plot

```
d.plot_clusters_scatter(["CVL_bctaunutau","CSL_bctaunutau","
    CT_bctaunutau"], clusters=[1,2])
```

3D plot

```
d.plot_clusters_scatter(["CVL_bctaunutau","CSL_bctaunutau","
    CT_bctaunutau"], clusters=[1,2])
```

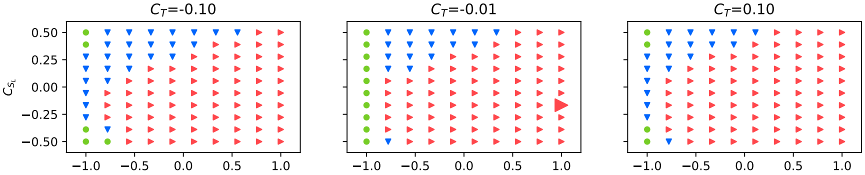


2D plots

```
d.plot_clusters_scatter(['CVL_bctaunutau','CSL_bctaunutau'])
```

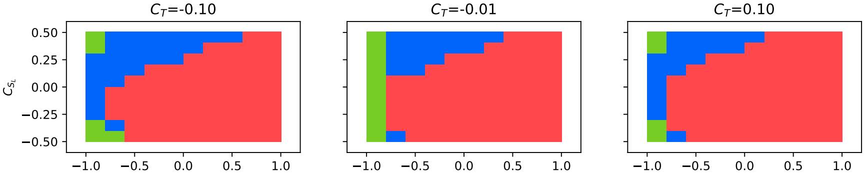
2D plots

```
d.plot_clusters_scatter(['CVL_bctaunutau','CSL_bctaunutau'])
```



2D plots filled

```
d.plot_clusters_fill(['CVL_bctaunutau', 'CSL_bctaunutau'])
```

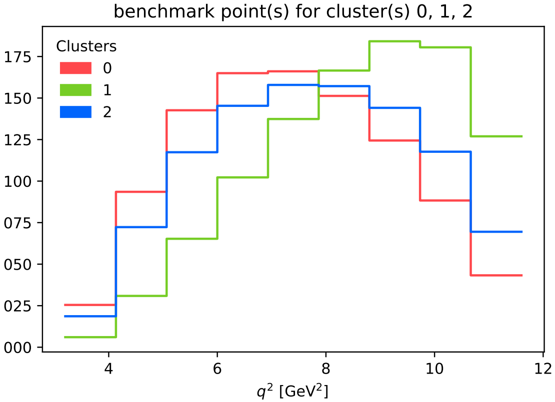


Distribution plots

```
d.plot_dist()
```

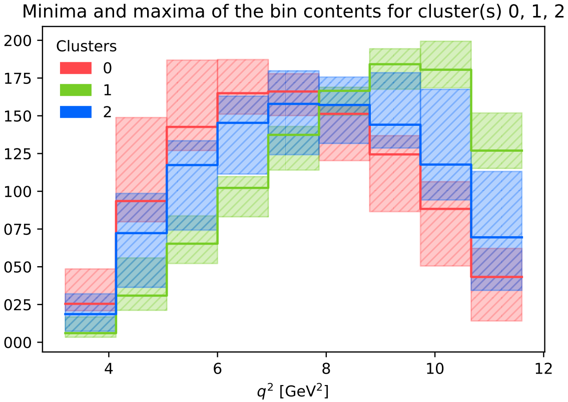
Distribution plots

```
d.plot_dist()
```



Distribution plots

```
d.plot_dist_minmax()
```



Outline

Summary

NP in kinematical distributions

Few benchmark points

CLUSTER KING

Clustering and visualisation

Example

$\bar{B} \to D\tau^- \bar{\nu}_\tau$ distribution

Project outline

Identify observable(s)

Find models, which ones are doable

Implementation

flavio, wcxf, wilson, tree-models

Clustering

Metric, Benchmark points, $N_{Clusters}$

Do analysis

For given Benchmark points

Tools for the numerical analysis

- Computing hundreds of relevant flavour observables properly accounting for theory uncertainties
  - **flavio** `https://flav-io.github.io` Straub, 1810.08132
- Representing and exchanging thousands of Wilson coefficient values, different EFTs, possibly different bases
  - **Wilson coefficient exchange format (WCxf)** `https://wcxf.github.io/` JA et al., 1712.05298
- RG evolution above and below the EW scale, matching from SMEFT to the weak effective theory (WET)
  - **wilson** `https://wilson-eft.github.io` JA, Kumar, Straub, 1804.05033