Data Preservation:

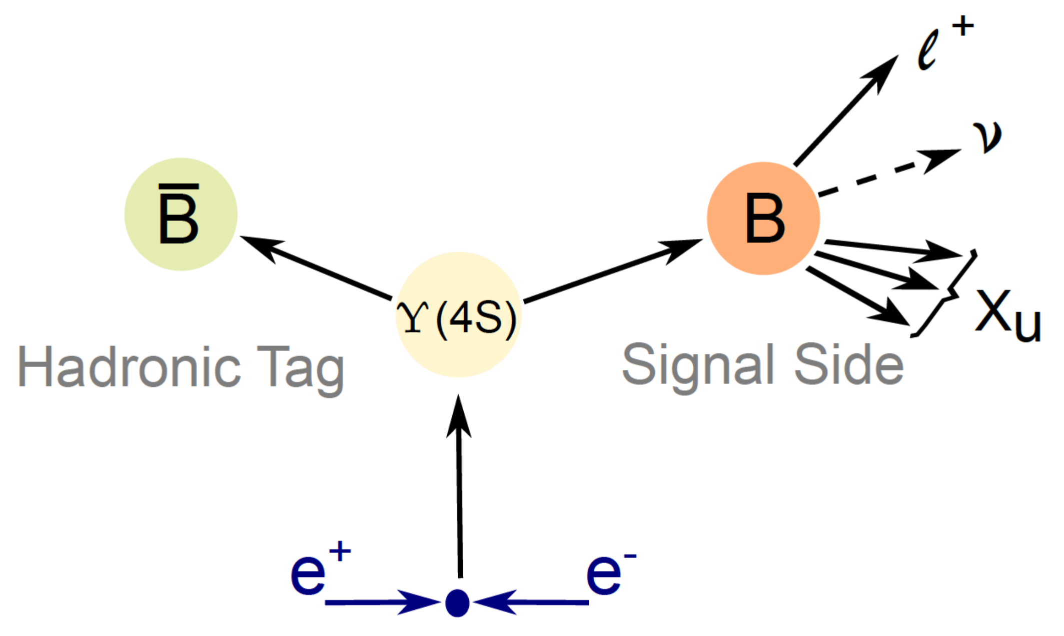
$B \to X_u \ell \nu$ Differential Branching Fractions

Lu Cao

Belle II Data Preservation Workshop, 7 Oct 2022

Measurement of Differential Branching Fractions of Inclusive $B \to X_u \ell \nu$ Decays

Belle: Phys. Rev. Lett. 127 (2021) 26, 261801

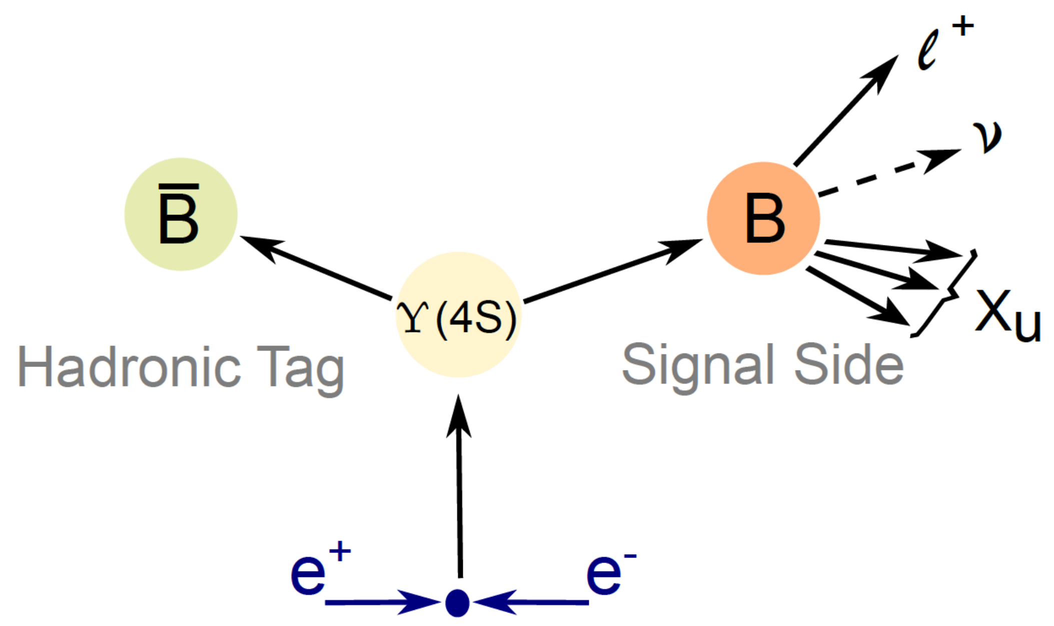


- Using **full Belle** dataset of **711 fb^{-1}**
- **Hadronic tagging** with Neural Networks (~ 0.2-0.3% efficiency)
- Based on **machine learning (BDT)** to suppress backgrounds with 11 training features, e.g. MM^2,#$K^{\pm}$, #K_S.
- Measure differential spectra of 6 kinematic variables in the phase space of $E_l^B > 1$ GeV:

q^2 , E_l^B , M_X , M_X^2 , P_+ , P_- (light-cone momenta: $P_{\pm} = E_X \mp |p_X|$)

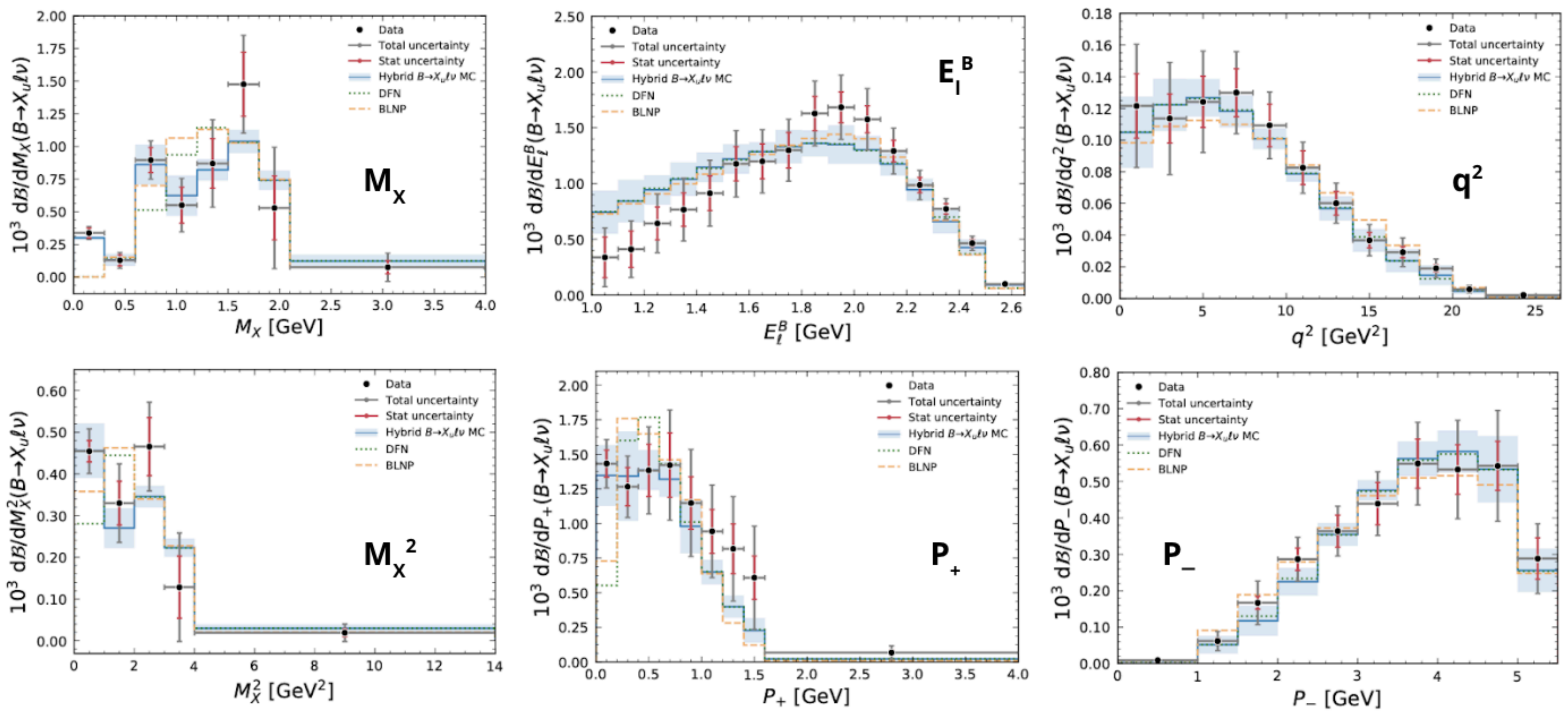
Measurement of Differential Branching Fractions of Inclusive $B \to X_u \ell \nu$ Decays

Belle: Phys. Rev. Lett. 127 (2021) 26, 261801



- Using **full Belle** dataset of **711 fb^{-1}**
- **Hadronic tagging** with Neural Networks (~ 0.2-0.3% efficiency)
- Based on **machine learning (BDT)** to suppress backgrounds with 11 training features, e.g. MM^2,$\#K^\pm$, $\#K_S$.
- Measure differential spectra of 6 kinematic variables in the phase space of $E_l^B > 1$ GeV:

q^2 , E_l^B , M_X , M_X^2 , P_+ , P_- (light-cone momenta: $P_\pm = E_X \mp |p_X|$)

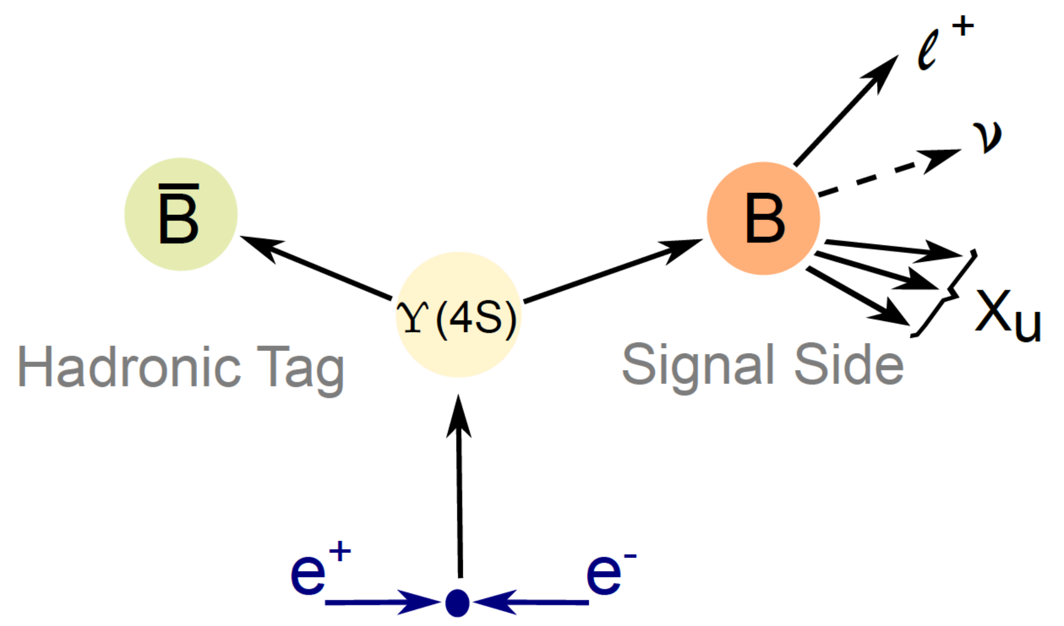


All MC shapes are normalised to 1.59 x 10^{-3} [Belle, PRD 104 , 012008 (2021)]

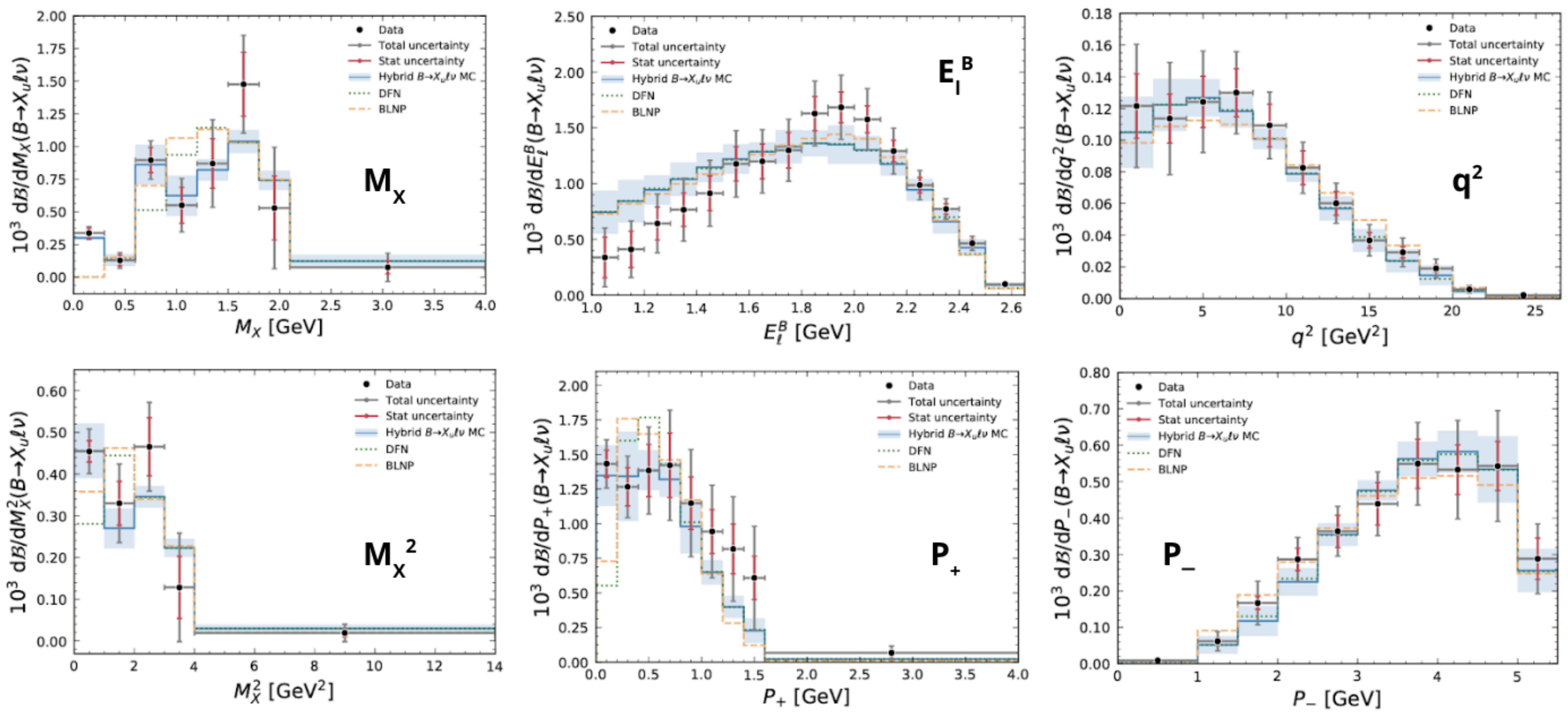
Measurement of Differential Branching Fractions of Inclusive $B \to X_u \ell \nu$ Decays

Belle: Phys. Rev. Lett. 127 (2021) 26, 261801

- Using **full Belle** dataset of **711 fb^{-1}**
- **Hadronic tagging** with Neural Networks (~ 0.2-0.3% efficiency)
- Based on **machine learning (BDT)** to suppress backgrounds with 11 training features, e.g. MM^2,#$K^\pm$, #K_S.
- Measure differential spectra of 6 kinematic variables in the phase space of $E_l^B > 1$ GeV:

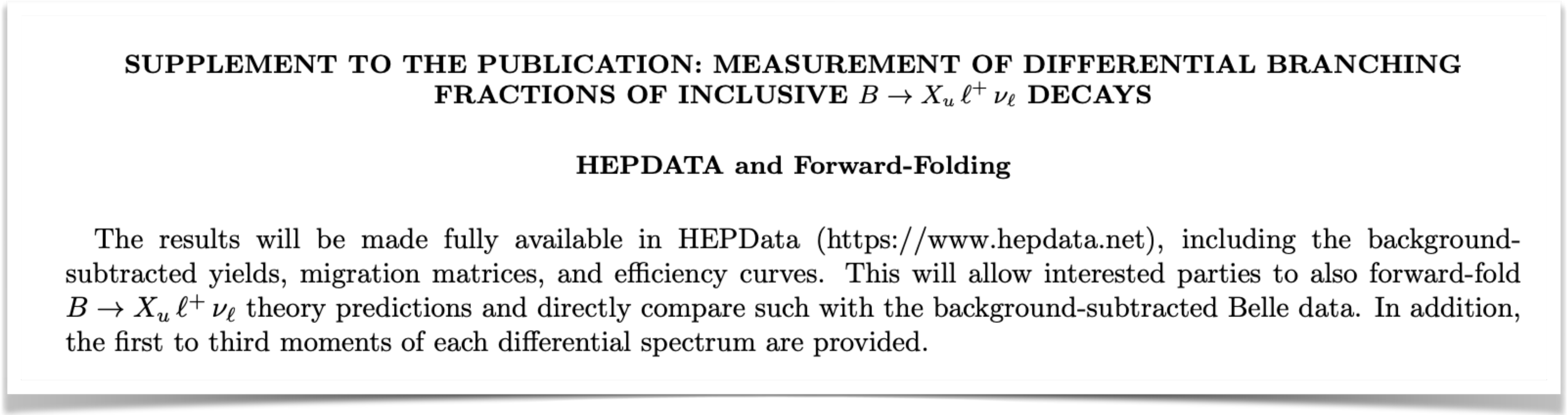


q^2 , E_l^B , M_X , M_X^2 , P_+ , P_- (light-cone momenta: $P_\pm = E_X \mp |p_X|$)



All MC shapes are normalised to 1.59 x 10^{-3} [Belle, PRD 104 , 012008 (2021)]

+ 16 pages supplement and 50 tables on HepData


SUPPLEMENT TO THE PUBLICATION: MEASUREMENT OF DIFFERENTIAL BRANCHING FRACTIONS OF INCLUSIVE $B \to X_u \ell^+ \nu_\ell$ DECAYS

HEPDATA and Forward-Folding

The results will be made fully available in HEPData (https://www.hepdata.net), including the background-subtracted yields, migration matrices, and efficiency curves. This will allow interested parties to also forward-fold $B \to X_u \ell^+ \nu_\ell$ theory predictions and directly compare such with the background-subtracted Belle data. In addition, the first to third moments of each differential spectrum are provided.

Analysis Workflow

Brief Recap

Event Reconstruction

Background subtraction

Unfolding

Eff. & acc. correction

Differential BR & full experimental correlations

Further outcome based on measured spectra

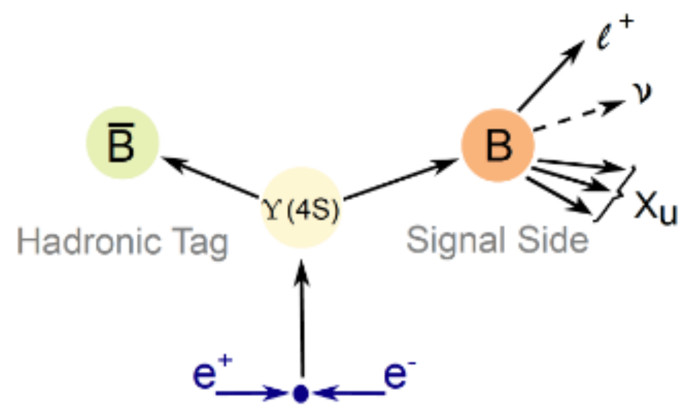
Shortly on analysis details, focus on **which** data is preserved and **why**

Analysis Workflow

Brief Recap

Event Reconstruction → Background subtraction → Unfolding → Eff. & acc. correction → Differential BR & full experimental correlations → Further outcome based on measured spectra

- Reconstruct $B \to X\ell\nu$ decays and key kin. variables
- Apply all necessary selections to enhance signal significancy (e.g. BDT score)
- Strategy inherited from the partial BR measurement [Belle: *Phys. Rev. D* 104 (2021) 1, 012008]



Can fully assign each final state particle to either the tag or signal side

→ Allows to reconstruct X_u

Reconstructed kinematic variables

- Hadronic system X:

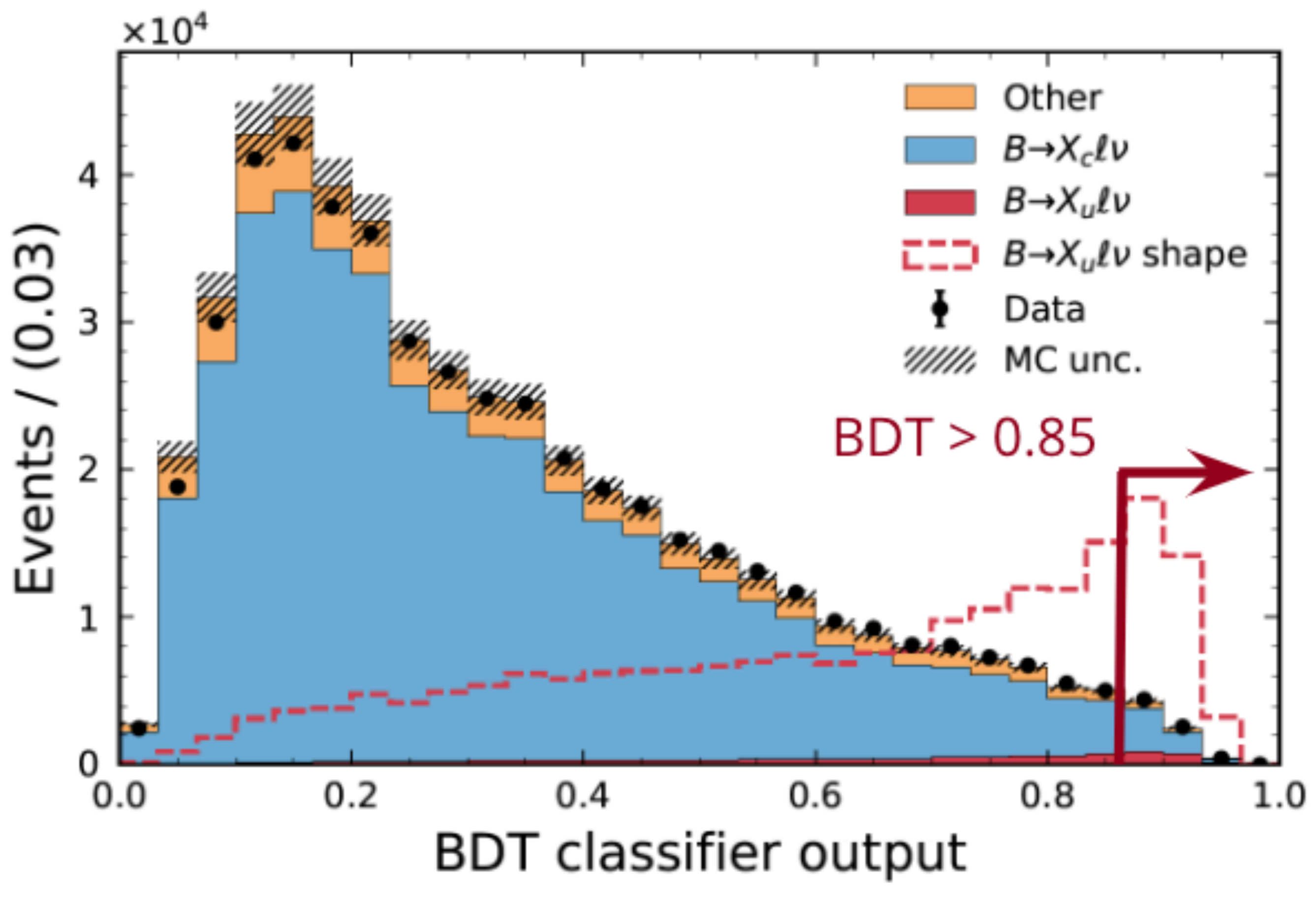
$$p_X = \sum_i \left(\sqrt{m_\pi^2 + |\mathbf{p_i}|^2}, \mathbf{p_i}\right) + \sum_i (E_i, \mathbf{k_i})$$

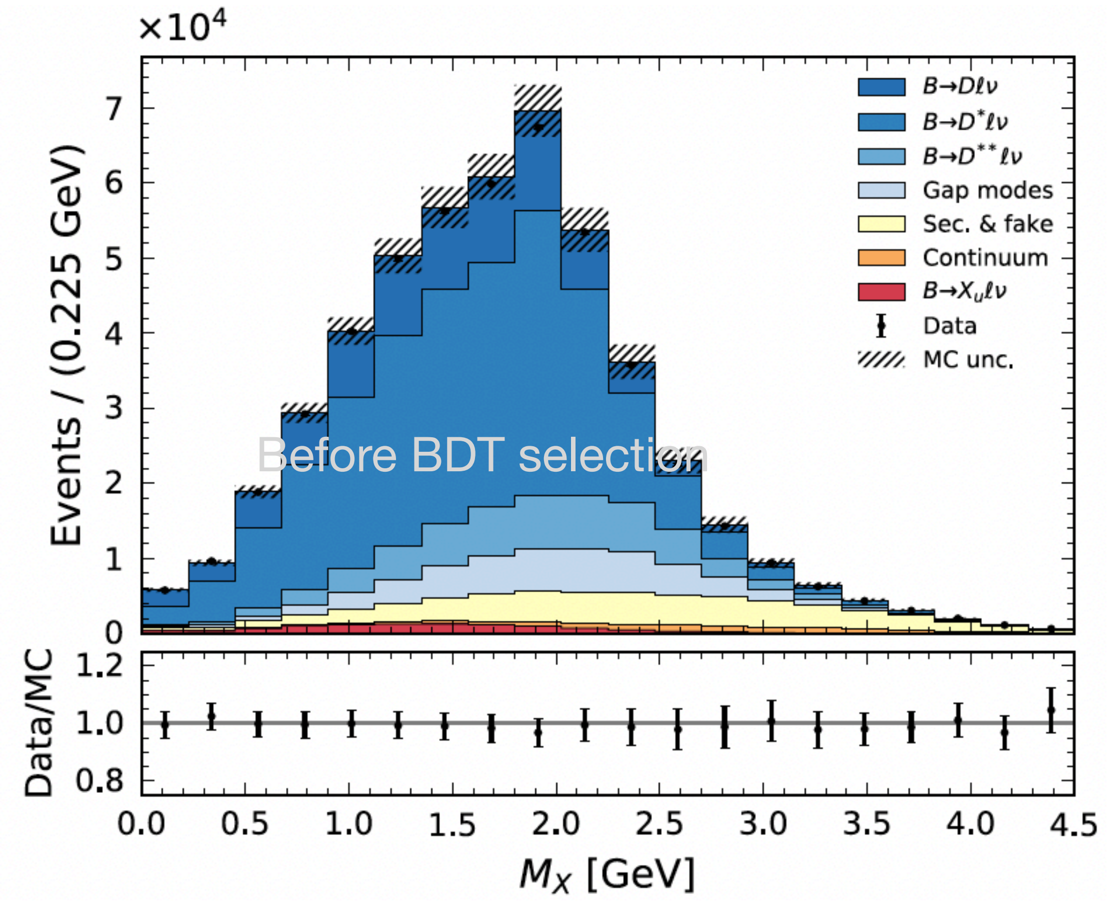
- Missing mass squared:

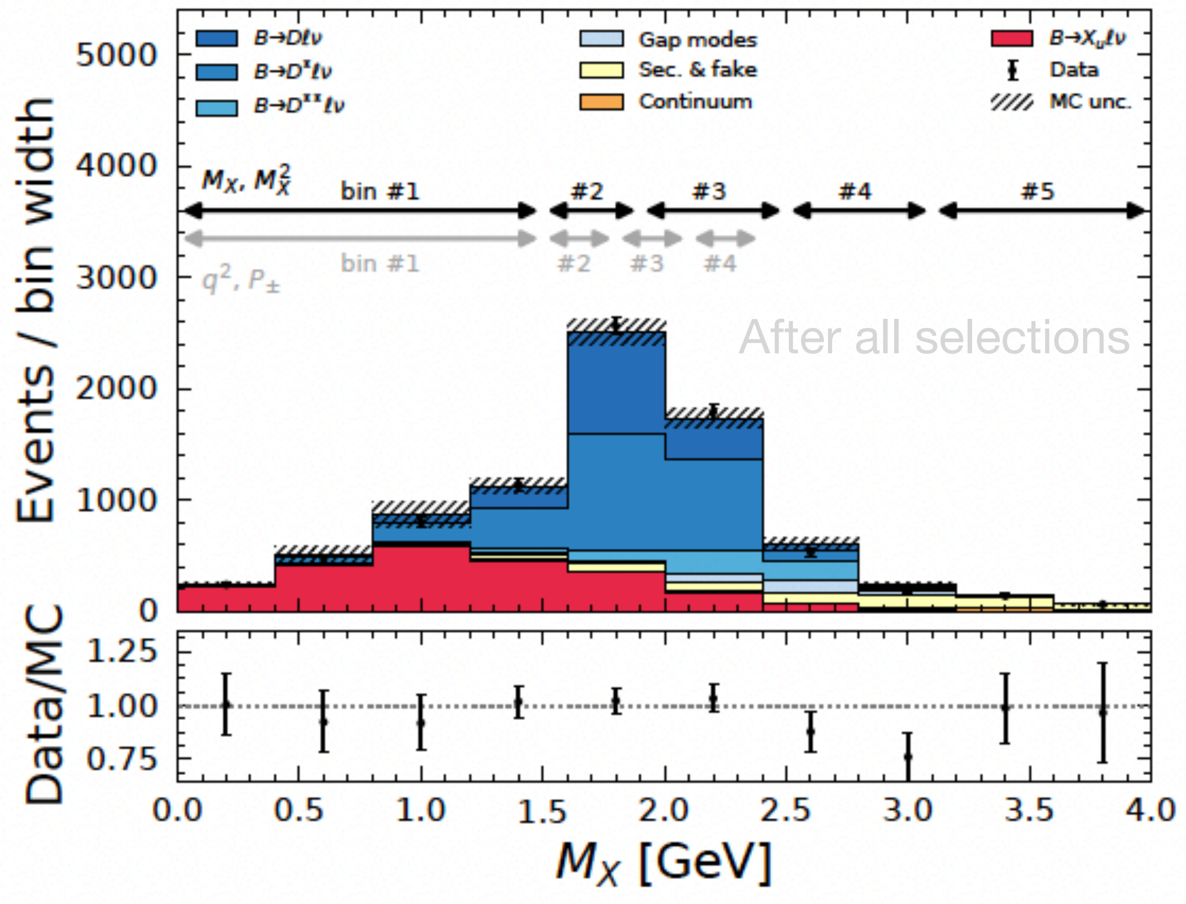
$$MM^2 = \left(P_{\Upsilon(4S)} - P_{\text{tag}} - P_X - P_\ell\right)^2$$

- Leptonic system:

$$q^2 = (P_B - P_X)^2 = (P_\ell + P_\nu)^2$$





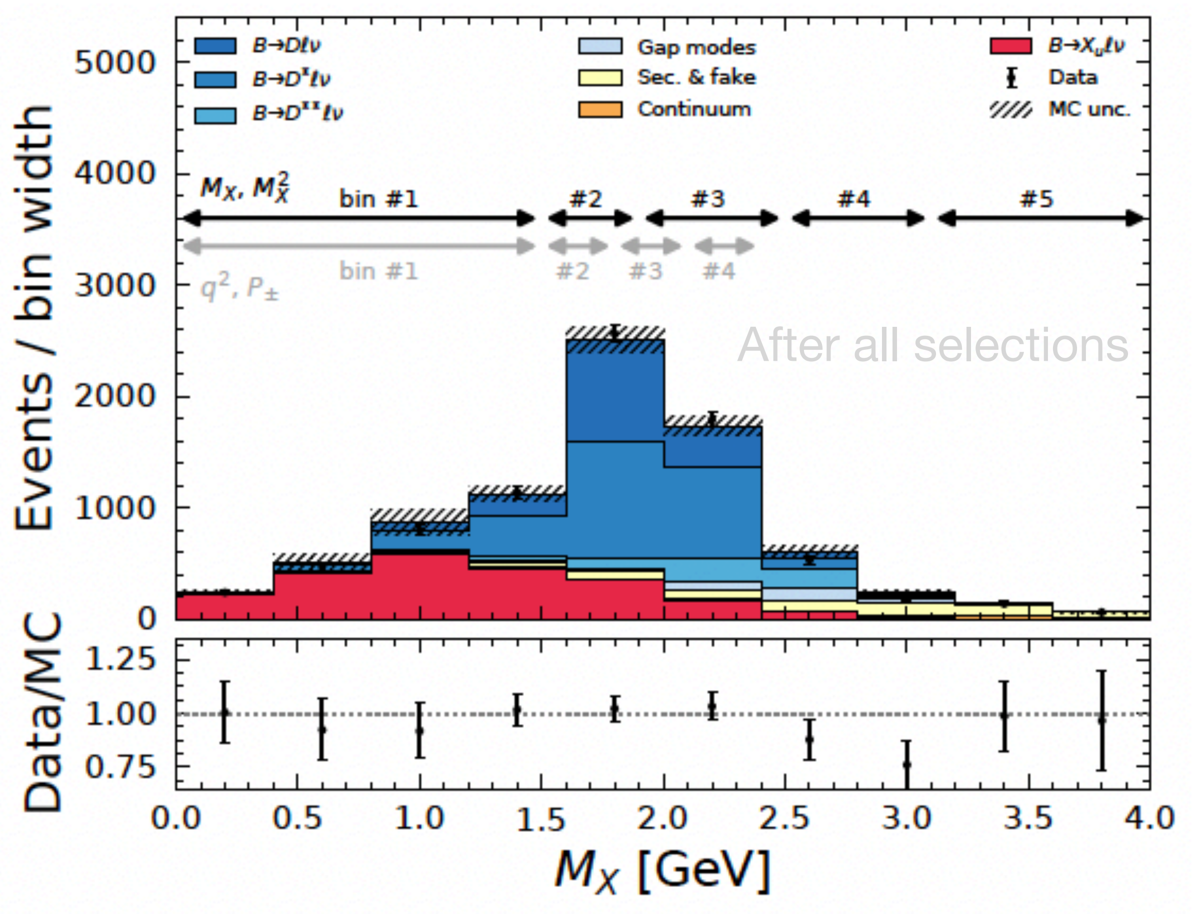


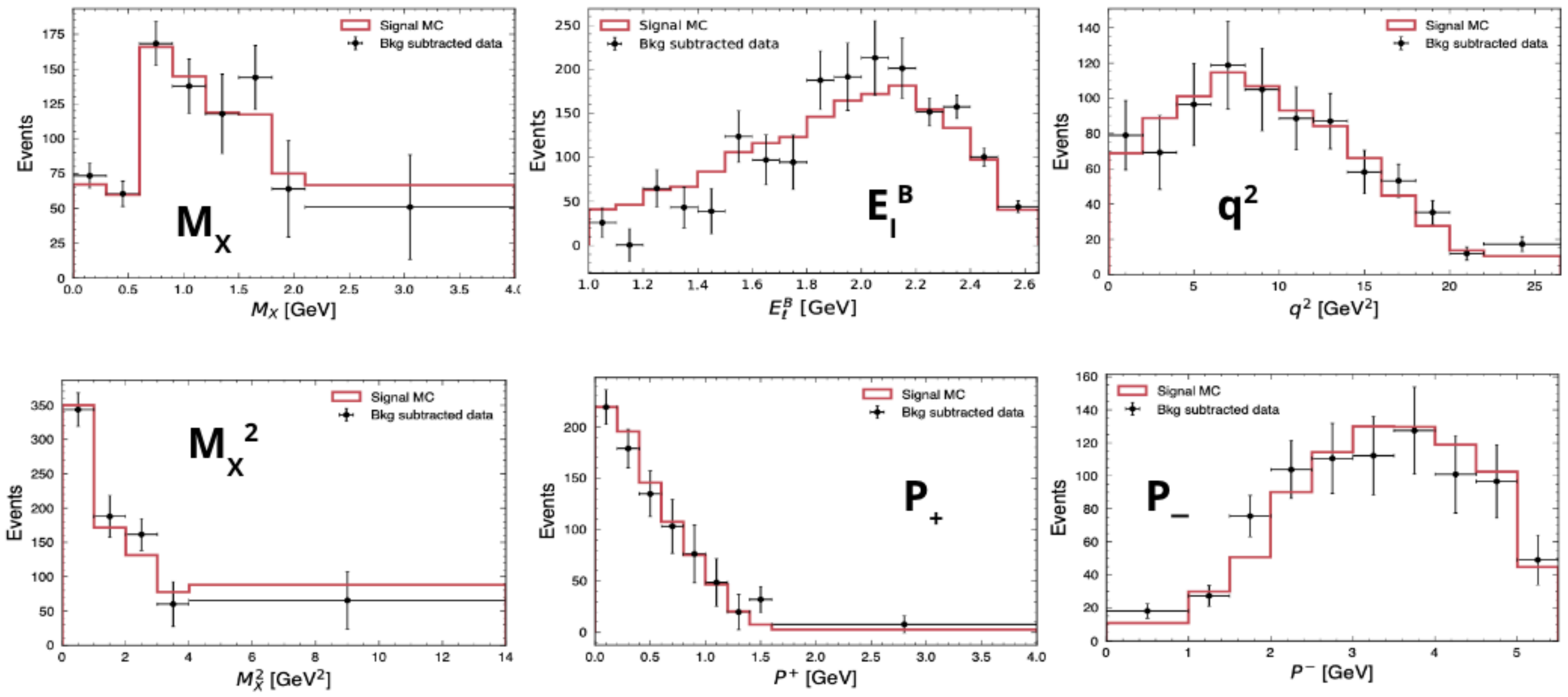
Analysis Workflow

Brief Recap

Event Reconstruction → **Background subtraction** → Unfolding → Eff. & acc. correction → Differential BR & full experimental correlations → Further outcome based on measured spectra

- Fit M_X distribution and subtract background events from data
- Fit uncertainty and correlations fully propagated into bkg-subtracted spectra

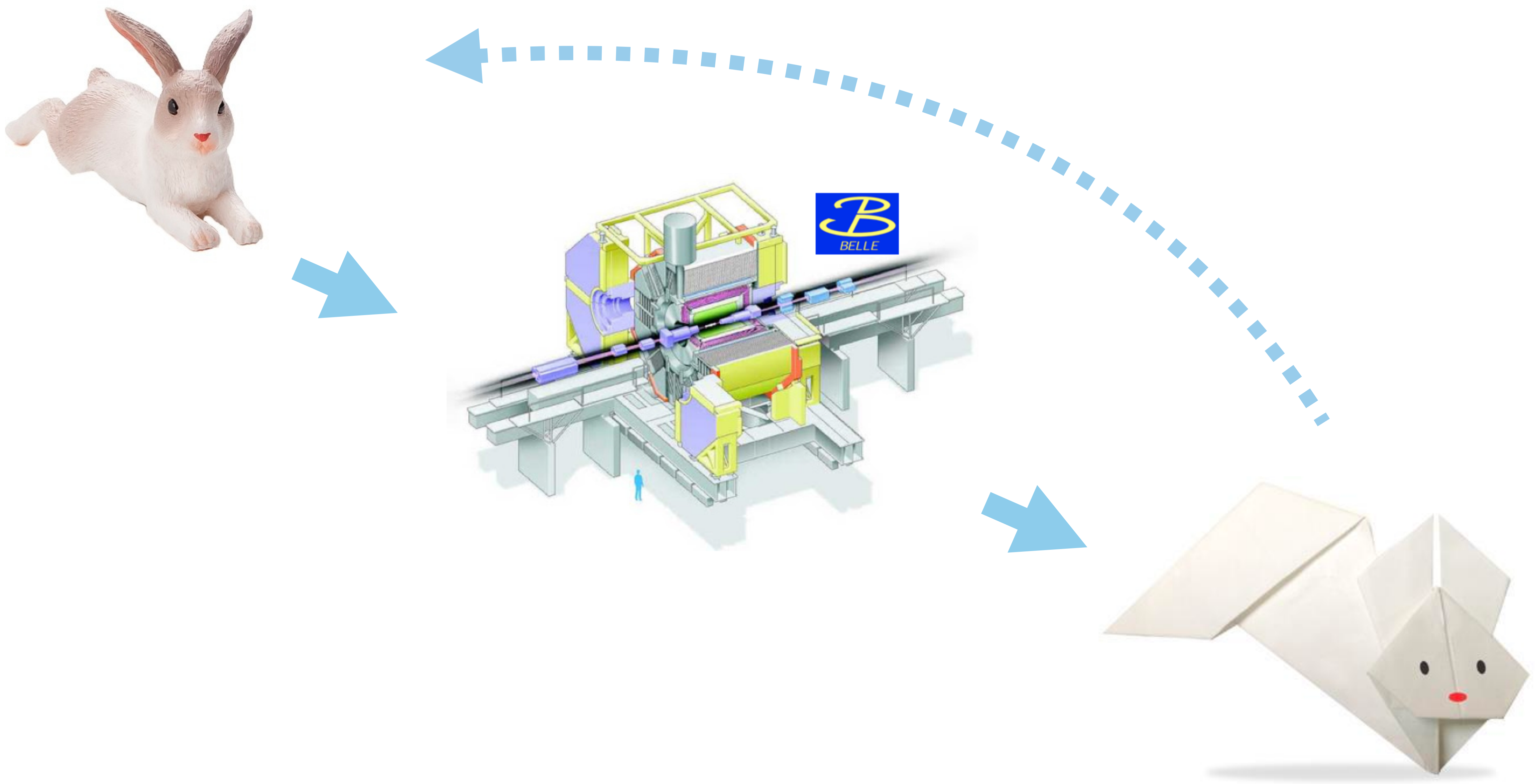




Analysis Workflow

Brief Recap

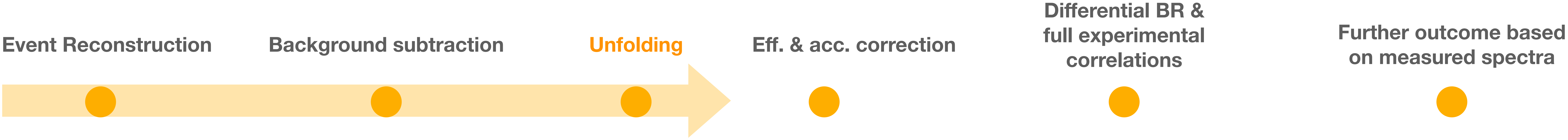
Event Reconstruction → Background subtraction → **Unfolding** → Eff. & acc. correction → Differential BR & full experimental correlations → Further outcome based on measured spectra



- Measured spectra are usually distorted and transformed by various detector effects (resolution, acceptance, etc.) ==> **Need unfolding**
- Singular-Value-Decomposition (SVD) [NIMA 372:469(1996)] is used in this analysis

Analysis Workflow

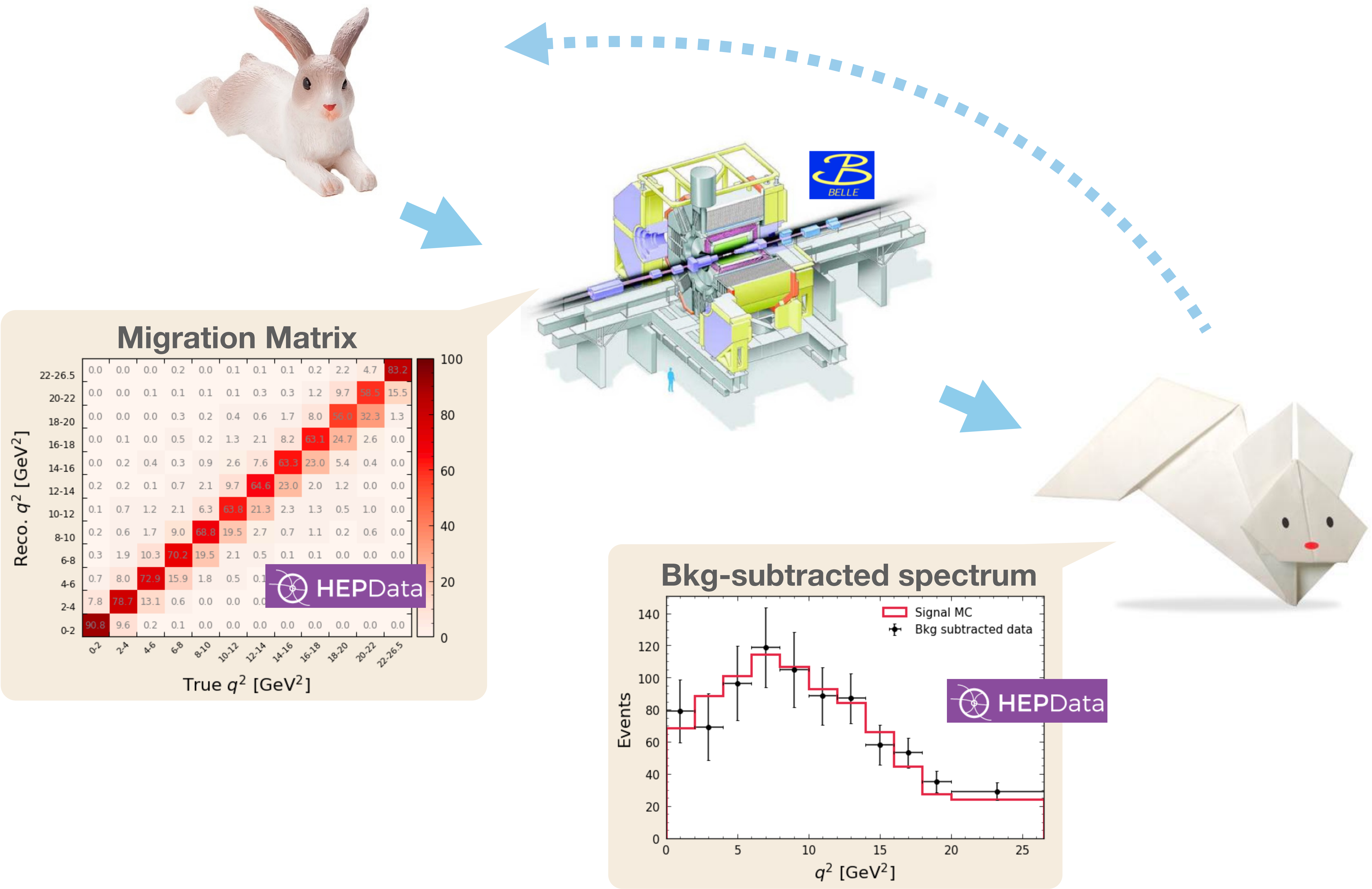
Brief Recap



- Measured spectra are usually distorted and transformed by various detector effects (resolution, acceptance, etc.) ==> **Need unfolding**
- Singular-Value-Decomposition (SVD) [NIMA 372:469(1996)] is used in this analysis

Ingredients provided on HepData allow:

- **EXP: utilise another unfolding method**
- **TH: forward-fold to smear theoretical predictions**
  (can be used to compare with observed spectra w/o unfolding)

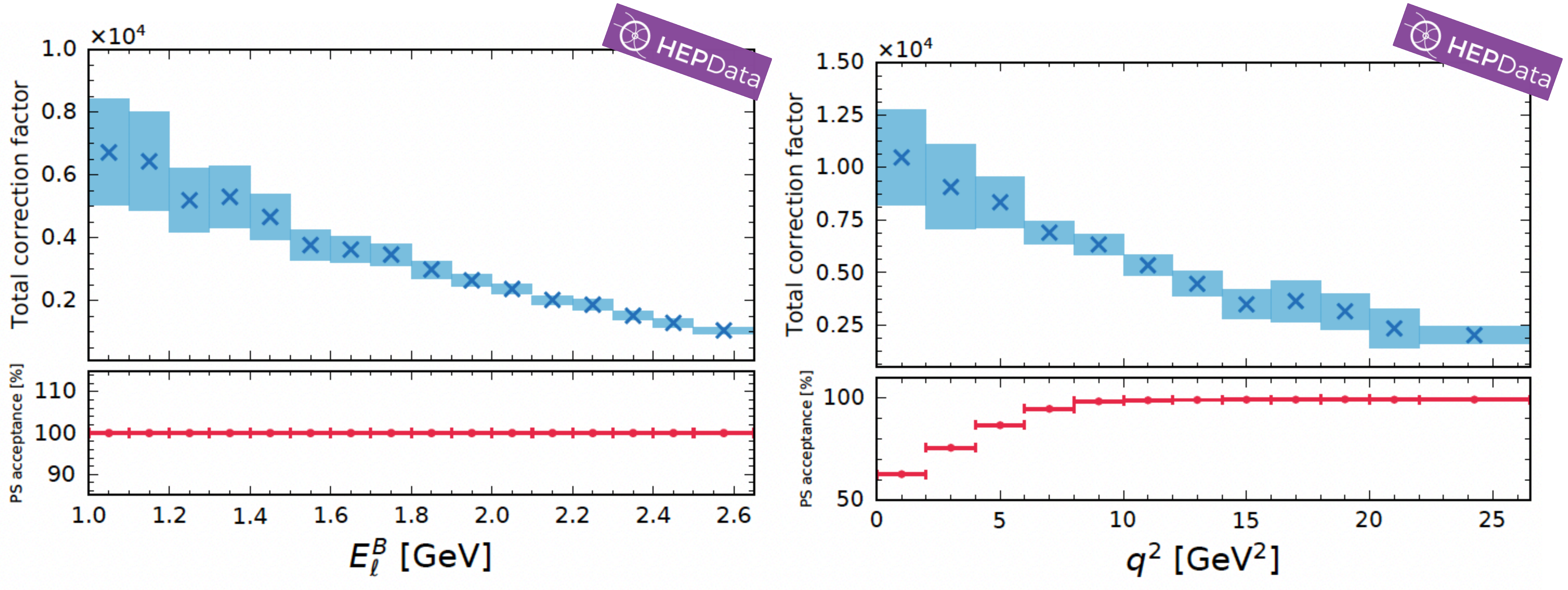


Analysis Workflow

Brief Recap

$$\Delta^i \mathcal{B}\left(B \to X_u \ell^+ \nu_\ell;\ \text{PS}\right) = \frac{\hat{\eta}^i_{\text{unfolded}} \cdot \left(\epsilon^i_{\text{tag}} \cdot \epsilon^i_{\text{sel}}\right)^{-1} \cdot \epsilon^i_{\Delta\mathcal{B}(\text{PS})}}{4 \cdot N_{B\bar{B}}}$$

- Measured phase space (PS): $E_\ell^B > 1$ GeV
- Correction factors for selection efficiency and PS acceptance are provided for all kin. variables



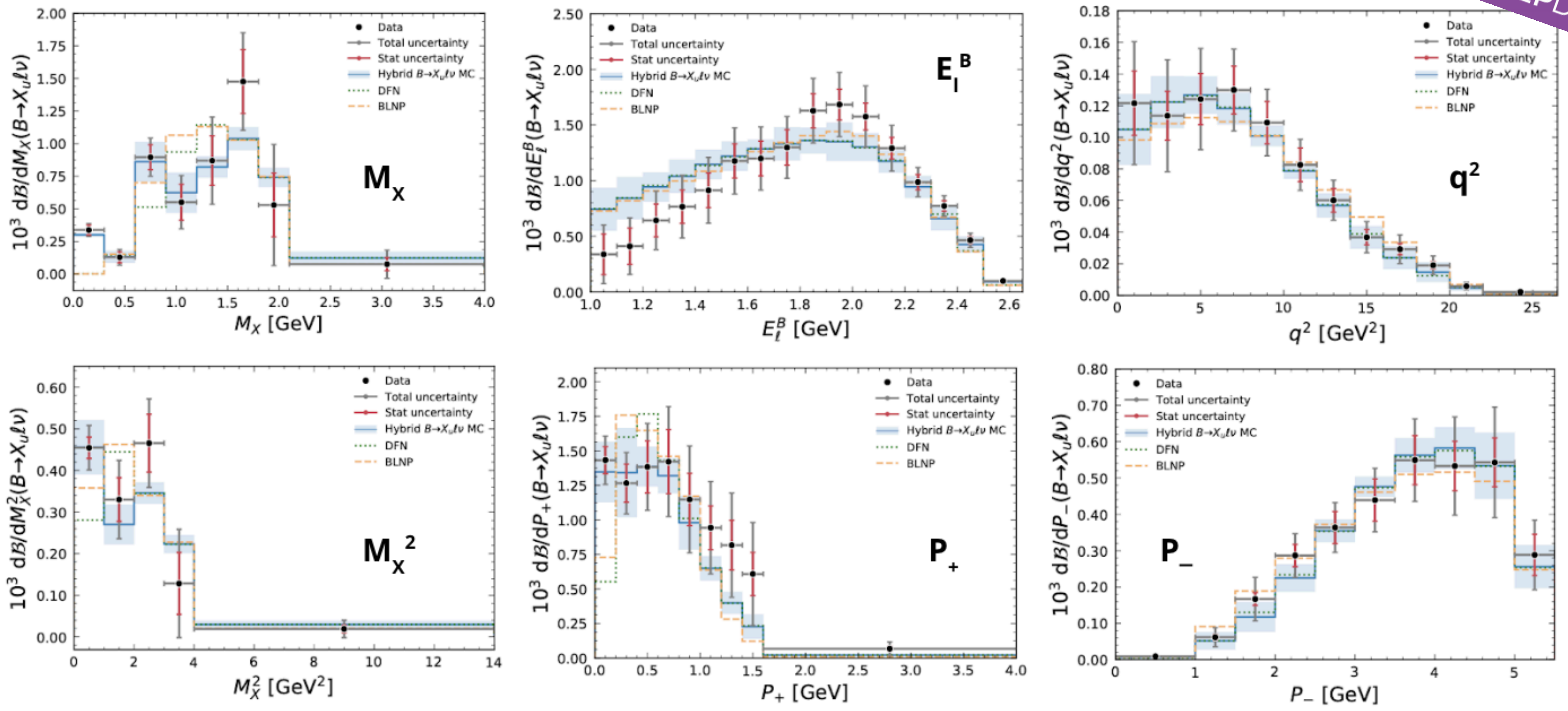
Analysis Workflow

Brief Recap

Event Reconstruction → Background subtraction → Unfolding → Eff. & acc. correction → **Differential BR & full experimental correlations** → Further outcome based on measured spectra

$$\Delta^i \mathscr{B}\left(B \to X_u \ell^+ \nu_\ell;\ \text{PS}\right) = \frac{\hat{\eta}^i_{\text{unfolded}} \cdot \left(\epsilon^i_{\text{tag}} \cdot \epsilon^i_{\text{sel}}\right)^{-1} \cdot \epsilon^i_{\Delta\mathscr{B}(\text{ PS })}}{4 \cdot N_{B\bar{B}}}$$

HEPData

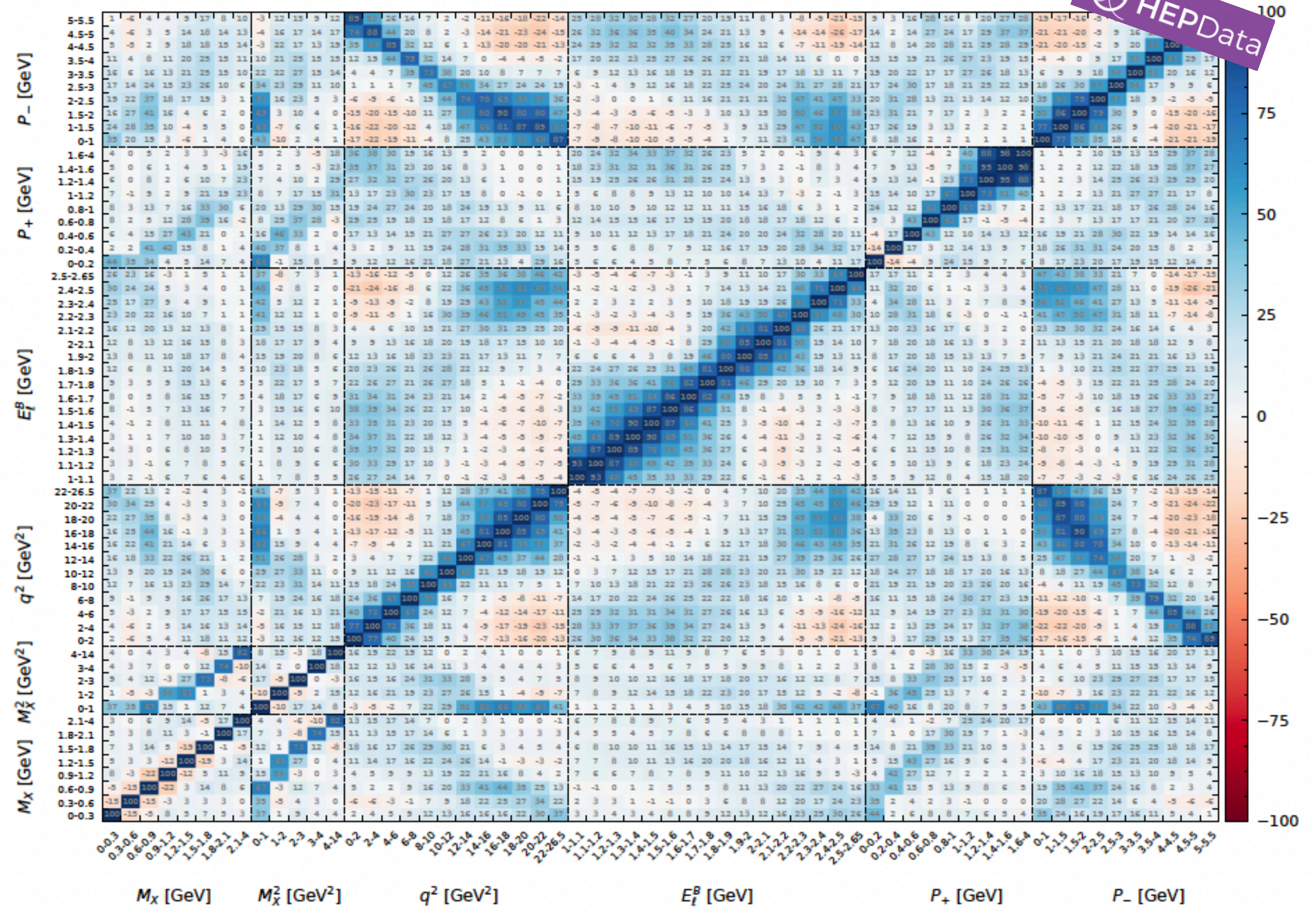


All MC shapes are normalised to 1.59 x 10^{-3} [Belle, PRD 104 , 012008 (2021)]

Necessary input for a global fit

Full experimental correlation matrix

HEPData

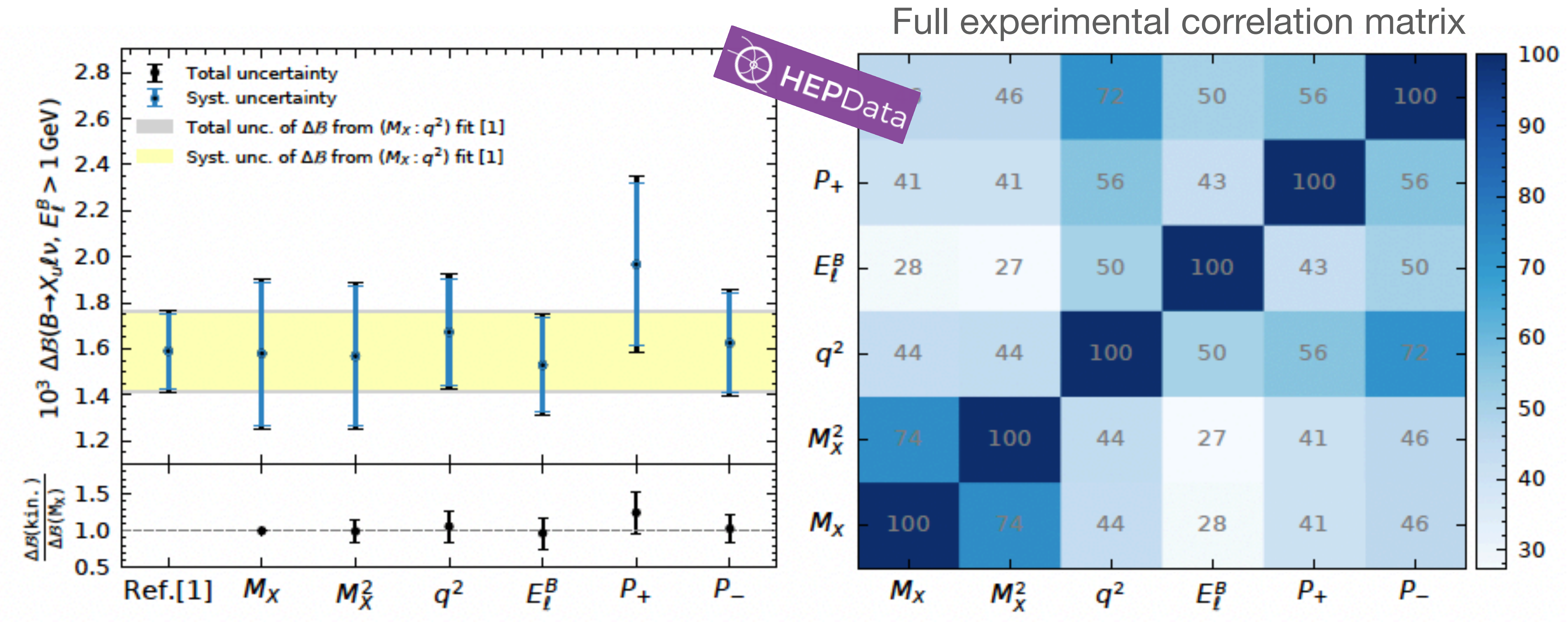


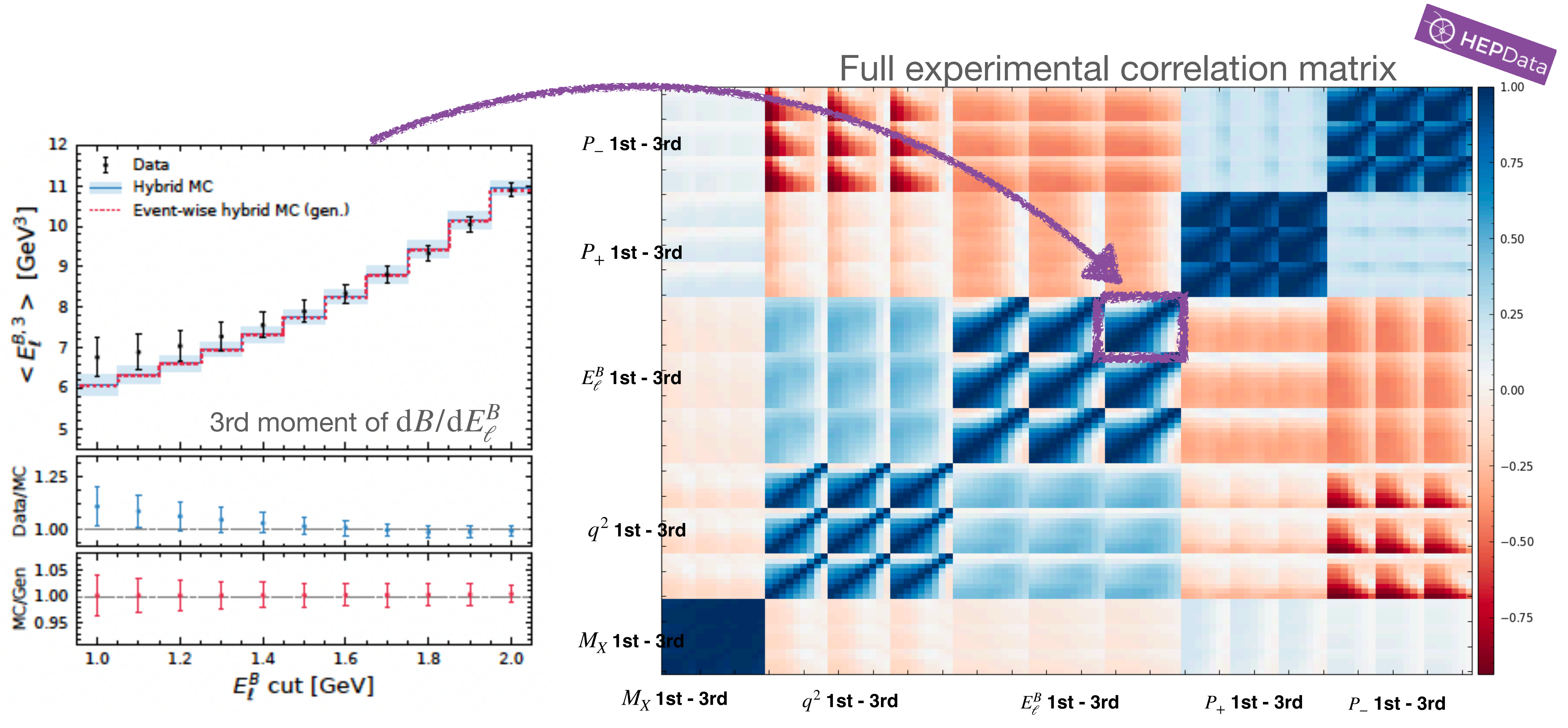
Analysis Workflow

Brief Recap

Event Reconstruction → Background subtraction → Unfolding → Eff. & acc. correction → Differential BR & full experimental correlations → Further outcome based on measured spectra

- Provide **direct** input for theory community:
  - Summed differential BRs & full correlations
  - 1st-3rd moments of each measured differential spectrum & full correlations

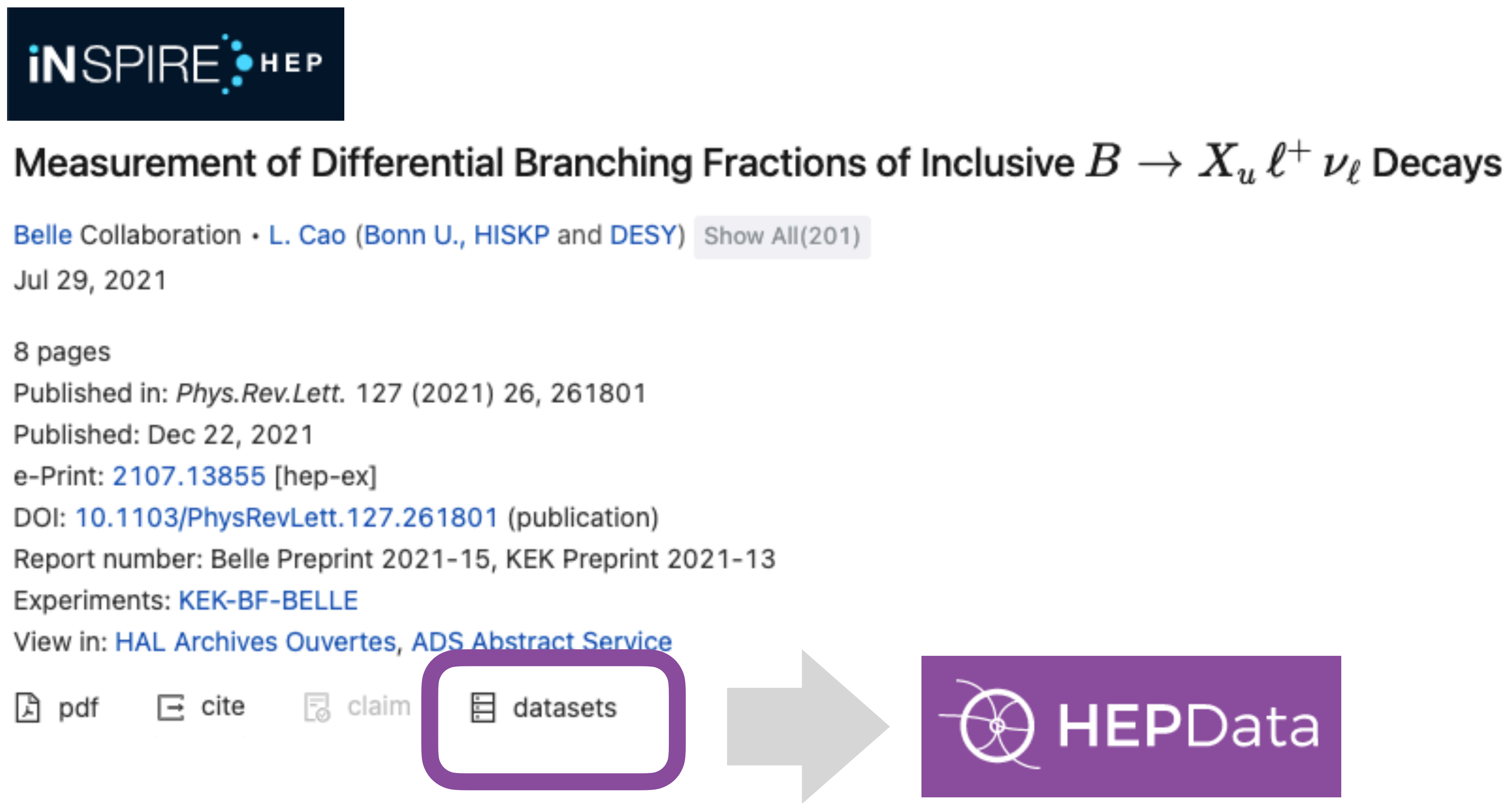


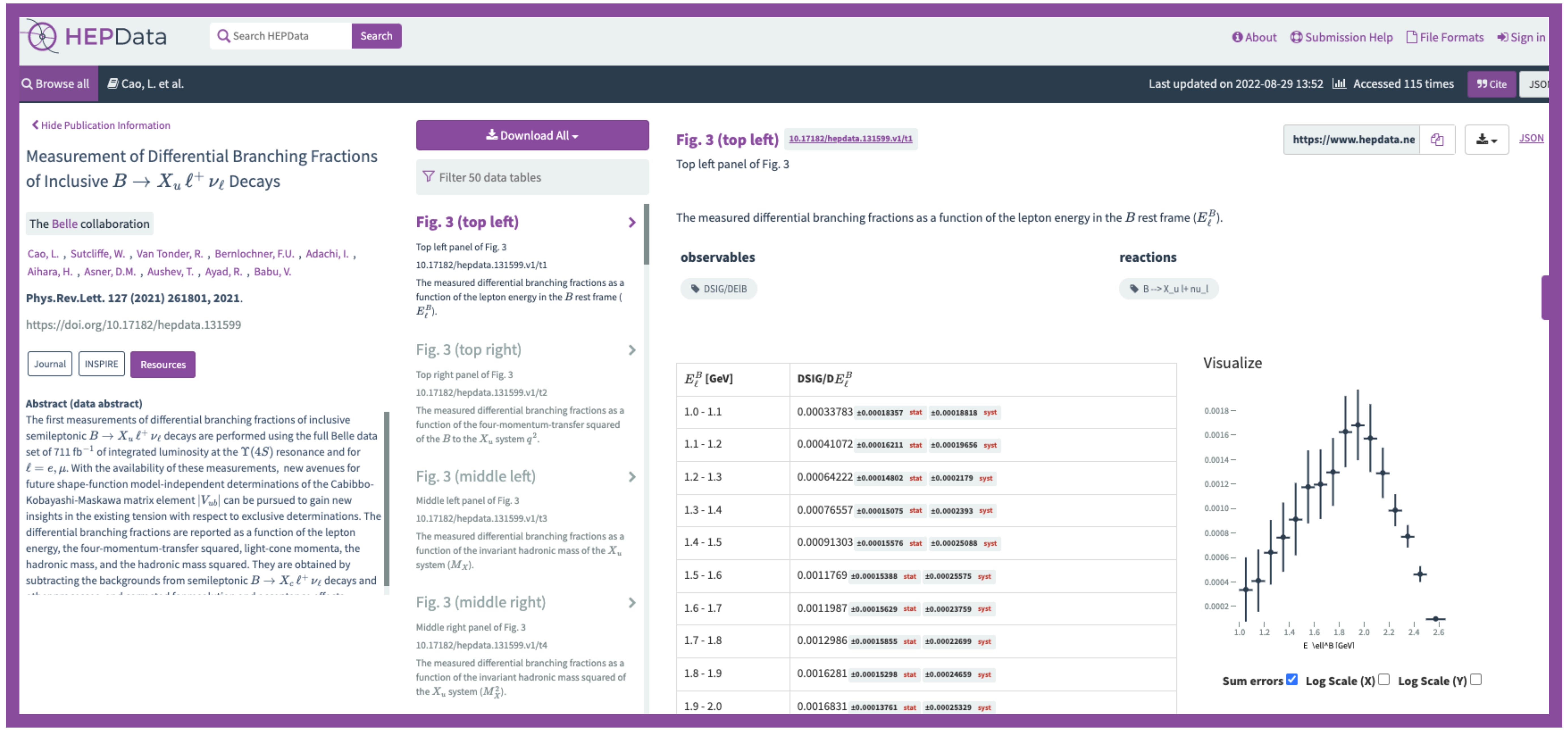


Crucial input for theoretical extraction of non-perturbative shape function & incl. $|V_{ub}|$

HepData Resources

HepData Resources





Hands-on Example

1D, 2D, binned, unbinned

- **1D**: yields, efficiency, differential branching fractions, moments, etc.
- **2D**: all matrices, i.e. covariance, correlation, migration
- **Binned**: differential branching fractions, efficiency, migration, etc.
- **Unbinned**: moments (as function of threshold), matrix given with elements' index, etc.

ALL filled & provided as flat tables

HEPData

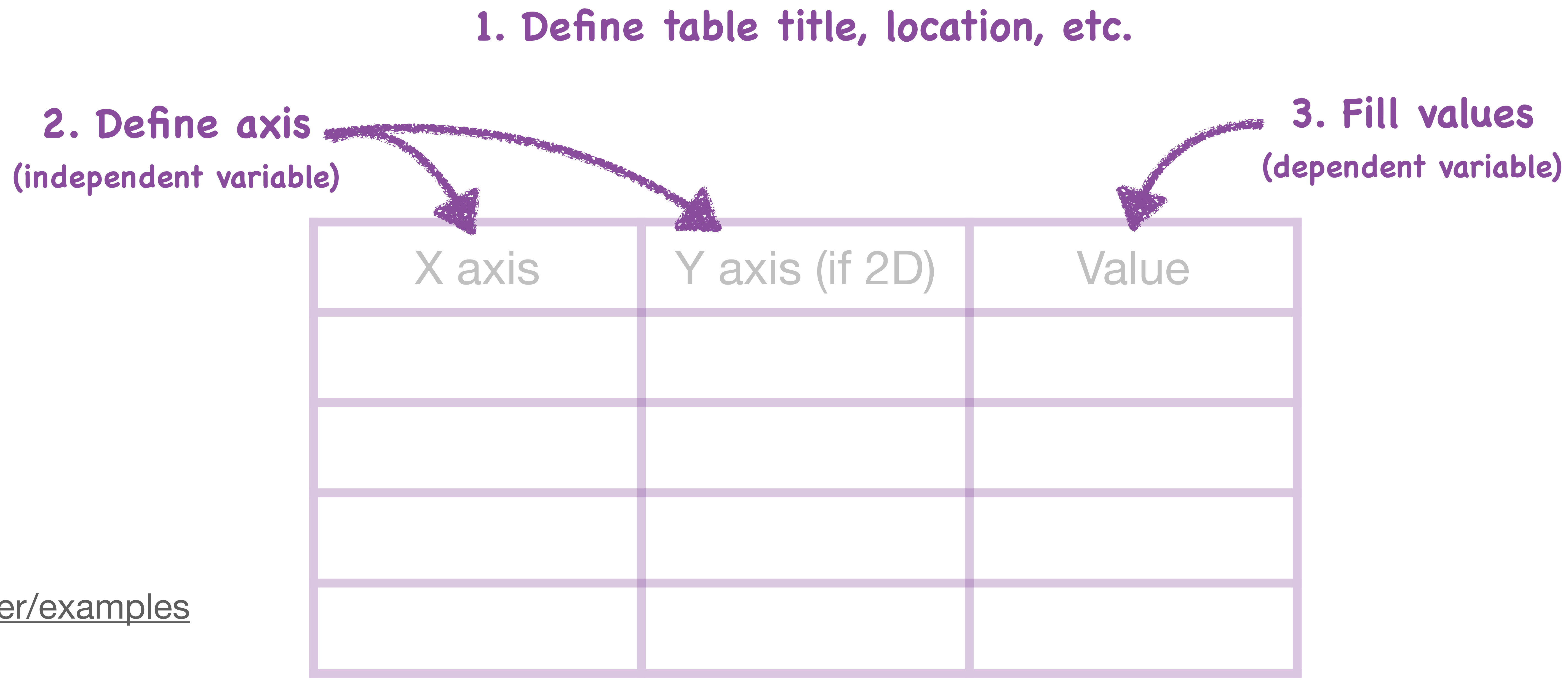
Hands-on Example

1D, 2D, binned, unbinned

- **1D**: yields, efficiency, differential branching fractions, moments, etc.
- **2D**: all matrices, i.e. covariance, correlation, migration
- **Binned**: differential branching fractions, efficiency, migration, etc.
- **Unbinned**: moments (as function of threshold), matrix given with elements' index, etc.

ALL filled & provided as flat tables

HEPData



HepData tutorial: https://github.com/HEPData/hepdata_lib/tree/master/examples

Hands-on Example

1D binned: $\mathrm{d}B/\mathrm{d}q^2$

```python
import hepdata_lib
from hepdata_lib import Submission, Table, Variable, Uncertainty

submission = Submission()

submission.read_abstract("./B1_BtoXulnu_PRL/abstract_hepdata.txt")
submission.add_link("Publication website", "https://journals.aps.org/prl
submission.add_link("arXiv", "https://arxiv.org/abs/2107.13855")
submission.add_record_id(1895149, "inspire")
```

```python
## define table title, location, etc.
table_fig3b = Table("Fig. 3 (top right)")
table_fig3b.description = "The measured differential branching fractions as a function of t
table_fig3b.location = "Top right panel of Fig. 3"
table_fig3b.keywords["observables"] = ["DSIG/DQ**2"]
table_fig3b.keywords["reactions"] = ["B --> X_u l+ nu_l"]

# define 1D axis
x_fig3b = Variable("$q^{2}$", is_independent=True, is_binned=True, units="GeV$^2$")
x_fig3b.values = list(zip(q2_unfold_bin[:-1], q2_unfold_bin[1:]))
table_fig3b.add_variable(x_fig3b)

# fill central result
data_fig3b = Variable("DSIG/D$q^{2}$", is_independent=False, is_binned=False, units="")
data_fig3b.values = dBF_q2['dBF_cen']/get_bin_widths(q2_unfold_bin)

# fill stat. uncertainty
unc_data_fig3b_stat = Uncertainty("stat", is_symmetric=True)
unc_data_fig3b_stat.values = dBF_q2['dBF_err_stat']/get_bin_widths(q2_unfold_bin)
data_fig3b.add_uncertainty(unc_data_fig3b_stat)

# fill syst. uncertainty
unc_data_fig3b_syst = Uncertainty("syst", is_symmetric=True)
unc_data_fig3b_syst.values = dBF_q2['dBF_err_sys']/get_bin_widths(q2_unfold_bin)
data_fig3b.add_uncertainty(unc_data_fig3b_syst)
table_fig3b.add_variable(data_fig3b)
```

Fig. 3 (top right) 10.17182/hepdata.131599.v1/t2 https://www.hepdata.ne JS

Top right panel of Fig. 3

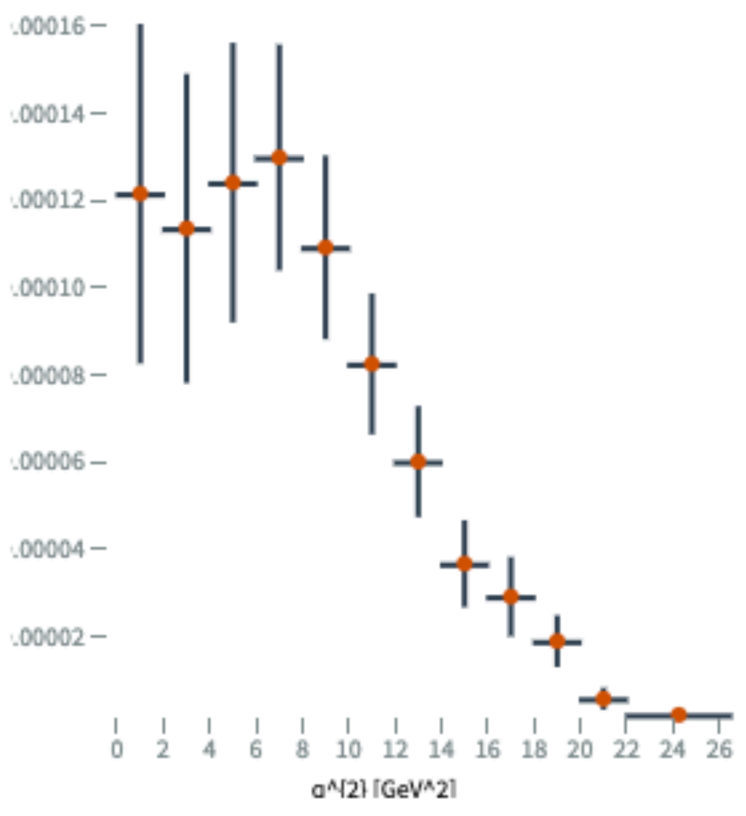
The measured differential branching fractions as a function of the four-momentum-transfer squared of the B to the X_u system q^2.

observables DSIG/DQ**2

reactions B-->X_u l+ nu_l

q^2 [GeV^2]	DSIG/Dq^2
0.0 - 2.0	0.00012158 ±2.0314e-05 stat ±3.3209e-05 syst
2.0 - 4.0	0.00011361 ±1.5535e-05 stat ±3.1849e-05 syst
4.0 - 6.0	0.00012413 ±1.6196e-05 stat ±2.7661e-05 syst
6.0 - 8.0	0.00012988 ±1.5083e-05 stat ±2.1042e-05 syst
8.0 - 10.0	0.00010926 ±1.3397e-05 stat ±1.6234e-05 syst
10.0 - 12.0	8.2558e-05 ±1.0624e-05 stat ±1.2165e-05 syst
12.0 - 14.0	6.0146e-05 ±7.4884e-06 stat ±1.0255e-05 syst
14.0 - 16.0	3.6742e-05 ±4.8875e-06 stat ±8.6703e-06 syst
16.0 - 18.0	2.9185e-05 ±3.4902e-06 stat ±8.378e-06 syst
18.0 - 20.0	1.8975e-05 ±2.2995e-06 stat ±5.4628e-06 syst
20.0 - 22.0	5.7594e-06 ±9.0366e-07 stat ±2.3605e-06 syst
22.0 - 26.5	2.0731e-06 ±4.1392e-07 stat ±5.2062e-07 syst

Visualize



Sum errors ☑ Log Scale (X) ☐ Log Scale (Y) ☐

Deselect variables or hide different error bars by clicking on them.

Hands-on Example

2D unbinned: full experimental correlations of differential spectra

- Giant correlations among all kin. variables
- In this case, matrix element is given with bin index instead of bin range (due to numerical overlap for various variables)

```
# define table title, location, etc.
table_fig8 = Table("Fig. 8")
table_fig8.description = "The full experimental (statistical and systematical) correlations of the
table_fig8.keywords["reactions"] = ["B --> X_u l+ nu_l"]

# define 2D X-axis
x = Variable("Element index (X)", is_independent=True, is_binned=False)
x.values = get_2Dbinedges(np.linspace(0,59,60),"X", binindex=True)

# define 2D Y-axis
y = Variable("Element index (Y)", is_independent=True, is_binned=False)
y.values = get_2Dbinedges(np.linspace(0,59,60),"Y", binindex=True)

# fill matrix elements
correlation_full = Variable("Correlation coefficient", is_independent=False, is_binned=False)
correlation_full.values = giant_cor_tot.transpose().reshape(len(giant_cor_tot)**2)

table_fig8.add_variable(x)
table_fig8.add_variable(y)
table_fig8.add_variable(correlation_full)

submission.add_table(table_fig8)
```

Fig. 8 10.17182/hepdata.131599.v1/t32

https://www.hepdata.ne

Example location

The full experimental (statistical and systematical) correlations of the differential branching fractions are shown. The matrix elements are ordered for $M_X, M_X^2, q^2, E_\ell^B, P_+$ and P_-.

reactions

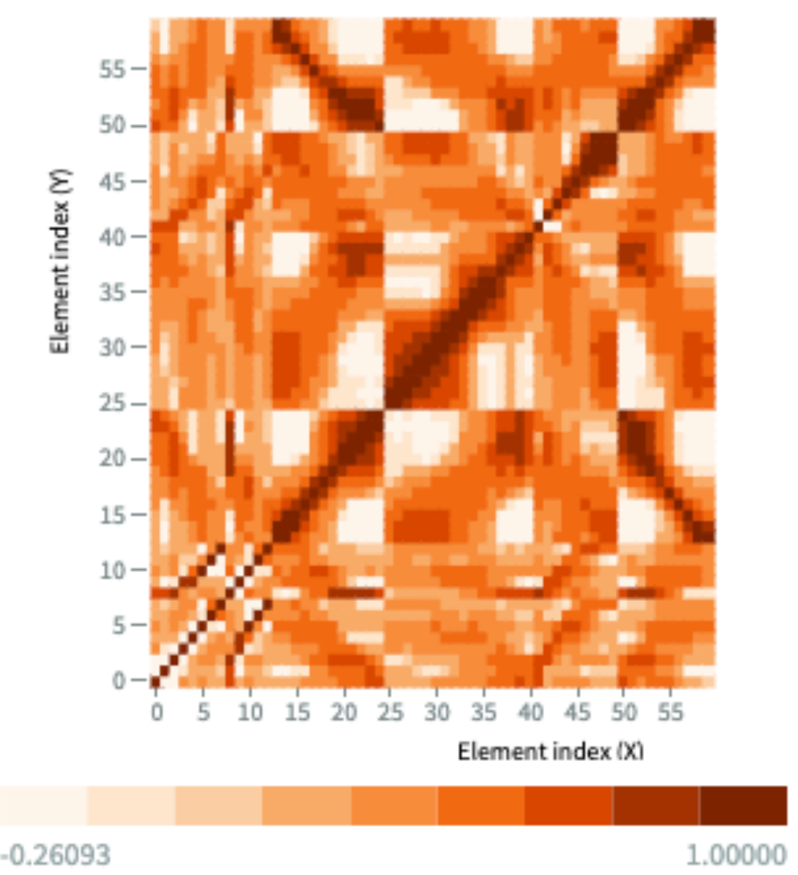
B --> X_u l+ nu_l

Showing 50 of 3600 values

Show All 3600 values

Element index (X)	Element index (Y)	Correlation coefficient
0.0	0.0	1.0
0.0	1.0	-0.15023
0.0	2.0	-0.046279
0.0	3.0	0.078509
0.0	4.0	0.050481
0.0	5.0	0.067942
0.0	6.0	0.053793
0.0	7.0	0.029232
0.0	8.0	0.36972
0.0	9.0	0.013695
0.0	10.0	0.090197
0.0	11.0	0.036792
0.0	12.0	0.03691

Visualize



Brushing Enabled? ☐

X Axis Element index (X)

Y Axis Element index (Y)

Hands-on Example

2D binned: migration matrix

```python
# define table title, location, etc.
table_mig_q2 = Table("Fig. 5 (middle right)")
table_mig_q2.description = "Migration matrix of $q^2$ [in %]."
table_mig_q2.location = "Middle right panel of Fig. 5"
table_mig_q2.keywords["reactions"] = ["B --> X_u l+ nu_l"]

# define 2D X-axis
x_mig_q2 = Variable("True $q^2$ bin", is_independent=True, is_binned=True, units="GeV$^2$")
x_mig_q2.values = get_2Dbinedges(q2_unfold_bin,"X")

# define 2D Y-axis
y_mig_q2 = Variable("Reco. $q^2$ bin", is_independent=True, is_binned=True, units="GeV$^2$")
y_mig_q2.values = get_2Dbinedges(q2_unfold_bin,"Y")

# fill matrix elements
data_mig_q2 = Variable("Migration matrix", is_independent=False, is_binned=False, units="%")
data_mig_q2.values = 100*mig_q2['nominal_migration_matrix'].transpose().reshape(len(mig_q2['nominal_migration_matrix'])

table_mig_q2.add_variable(x_mig_q2)
table_mig_q2.add_variable(y_mig_q2)
table_mig_q2.add_variable(data_mig_q2)
```

Tutorial notebook: https://stash.desy.de/users/lcaocn/repos/hepdata_example/browse

Example: convert results into yaml for HepData submission

```python
In [1]: import numpy as np
        import pickle

        import hepdata_lib
        from hepdata_lib import Submission, Table, Variable, Uncertainty
```

Welcome to JupyROOT 6.24/06

1. Initialise submission with details on related publication

```python
In [2]: submission = Submission()

        submission.read_abstract("./abstract.txt")
        submission.add_link("Publication website", "https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.127.261801")
        submission.add_link("arXiv", "https://arxiv.org/abs/2107.13855")
        submission.add_record_id(1895149, "inspire")
```

Some useful functions

```python
In [3]: def get_bin_widths(bins):
            widths = [b-a for a, b in zip(bins[0:-1], bins[1:])]
            return widths

        def get_2Dbinedges(bins, axis="X", binindex=False):
        # prepare 2D binning array for hepdata matrix

            if binindex==False:
                a=list(zip(bins[:-1], bins[1:]))
                b=[]
                if axis=="X":
```

Fig. 5 (middle right) 10.17182/hepdata.131599.v1/t20

https://www.hepdata.ne

Middle right panel of Fig. 5

Migration matrix of q^2 [in %].

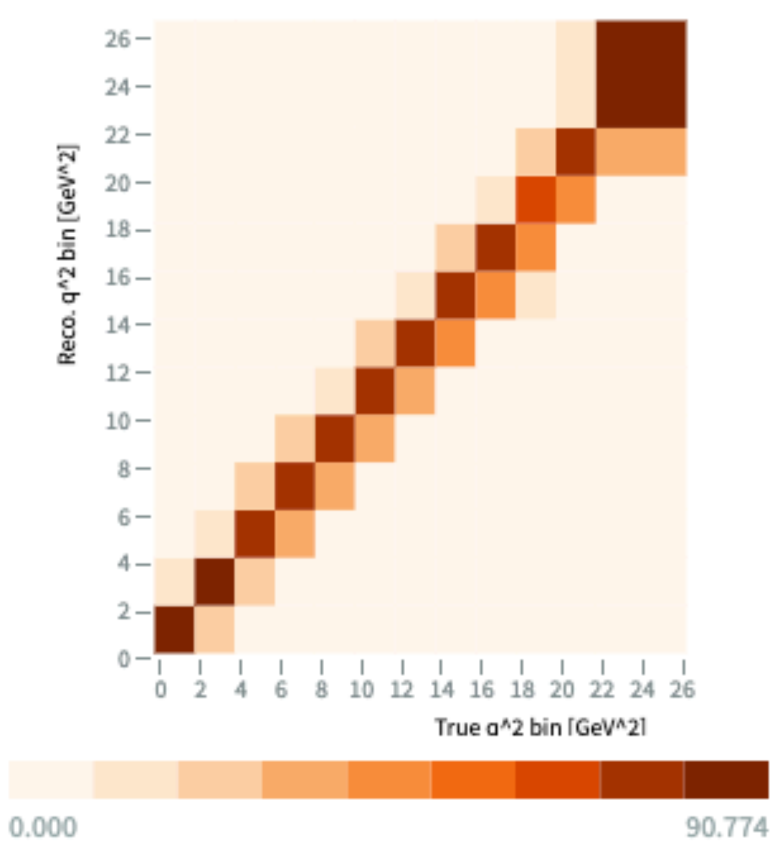
reactions

B --> X_u l+ nu_l

Showing 50 of 144 values — Show All 144 values

True q^2 bin [GeV^2]	Reco. q^2 bin [GeV^2]	Migration matrix [%]
0.0 - 2.0	0.0 - 2.0	90.774
0.0 - 2.0	2.0 - 4.0	7.7856
0.0 - 2.0	4.0 - 6.0	0.7107
0.0 - 2.0	6.0 - 8.0	0.27675
0.0 - 2.0	8.0 - 10.0	0.15665
0.0 - 2.0	10.0 - 12.0	0.070682
0.0 - 2.0	12.0 - 14.0	0.22555
0.0 - 2.0	14.0 - 16.0	0.0
0.0 - 2.0	16.0 - 18.0	0.0
0.0 - 2.0	18.0 - 20.0	0.0
0.0 - 2.0	20.0 - 22.0	0.0
0.0 - 2.0	22.0 - 26.5	0.0
2.0 - 4.0	0.0 - 2.0	9.5878

Visualize



Brushing Enabled? ☐

X Axis: True q^2 bin [GeV^2]

Y Axis: Reco. q^2 bin [GeV^2]