HEPData Example: $B^+ \to K^+ \nu \bar{\nu}$

Slavomira Stefkova

Belle II Data Preservation Workshop
07.10.2022

HEPData

Repository for publication-related High-Energy Physics data

HEPData Format

All details on submission formats: https://hepdata-submission.readthedocs.io/en/latest/

Submission should be done in *.zip, .tar, .tar.gz, .tgz* format and it should contain:

1. Main file *submission.yaml* which tells about your entire submission:
   - content of your submission
   - what data files are in your submission
   - what they contain
   - any related material
   - keywords
2. All data files also in *YAML* or *JSON* format (more details can be found on https://hepdata-submission.readthedocs.io/en/latest/data_yaml.html)
3. All associated figures in png format

You can install hepdata-validator tool to check your submission!

```
name: "Table 1"
description: Describe the data.  The more you say, the easier it'll be
keywords: # used for searching, possibly multiple values for each key
- {name: reactions, values: [P P --> Z0 Z0 X]}
- {name: observables, values: [SIG]}
- {name: cmenergies, values: [7000.0]}
- {name: phrases, values: [Inclusive, Integrated Cross Section, Cross
data_file: data1.yaml
```

YAML data file example

```
independent_variables:
- header: {name: Leading dilepton PT, units: GEV}
  values:
  - {low: 0, high: 60}
  - {low: 60, high: 100}
  - {low: 100, high: 200}
  - {low: 200, high: 600}
dependent_variables:
- header: {name: 10**6 * 1/SIG(fiducial) * D(SIG(fiducial))/DPT, uni
  qualifiers:
  - {name: RE, value: P P --> Z0 < LEPTON+ LEPTON- > Z0 < LEPTON+ LE
  - {name: SQRT(S), units: GEV, value: 7000}
  values:
  - value: 7000
    errors:
    - {symerror: 1100, label: stat}
    - {symerror: 79, label: 'sys,detector'}
    - {symerror: 15, label: 'sys,background'}
  - value: 9800
    errors:
    - {symerror: 1600, label: stat}
    - {symerror: 75, label: 'sys,detector'}
    - {symerror: 15, label: 'sys,background'}
  - value: 1600
    errors:
    - {symerror: 490, label: stat}
    - {symerror: 41, label: 'sys,detector'}
    - {symerror: 2, label: 'sys,background'}
  - value: 80
    errors:
    - {symerror: 60, label: stat}
    - {symerror: 2, label: 'sys,detector'}
    - {symerror: 0, label: 'sys,background'}
```
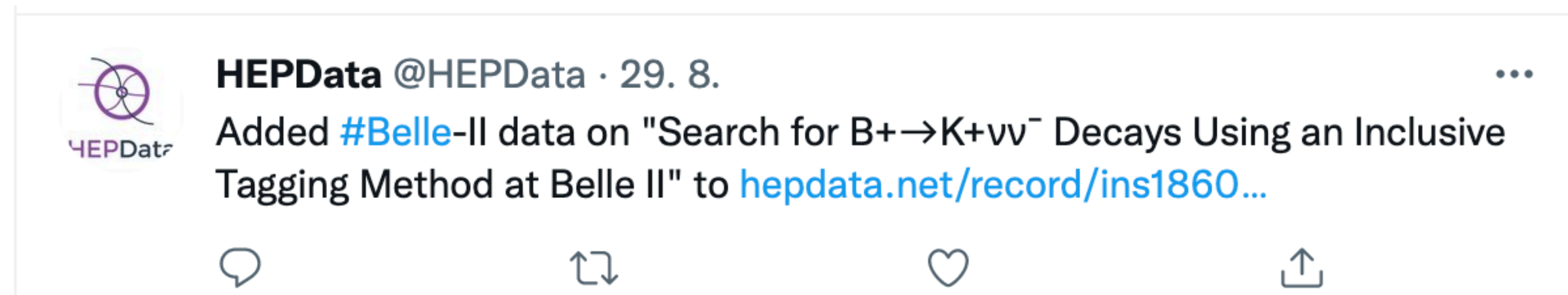
NB: publication information such as the paper title, authors and abstract, or the journal reference and DOI, is pulled from the corresponding INSPIRE record automatically

KIT
Karlsruher Institut für Technologie

HEPData Record for $B^+ \to K^+ \nu \bar{\nu}$

First HEPData Record for Belle II is online since 28.9.2022!

https://www.hepdata.net/record/ins1860766

HEPData @HEPData · 29. 8.

Added #Belle-II data on "Search for B+→K+νν⁻ Decays Using an Inclusive Tagging Method at Belle II" to hepdata.net/record/ins1860...

Search for B+→K+νν⁻ Decays Using an Inclusive Tagging Method at Belle II

The Belle-II collaboration

Abudinén, F. , Adachi, I. , Adamczyk, K. , Ahlburg, P. , Aihara, H. , Akopov, N. , Aloisio, A. , Ky, N. Anh , Asner, D.M. , Atmacan, H.

Phys.Rev.Lett. 127 (2021) 181802, 2021.

https://doi.org/10.17182/hepdata.130199

Journal | INSPIRE | Resources

Abstract (data abstract)
SuperKEKB Belle II. Measurement of the branching fraction of $B^+ \to K^+ \nu \bar{\nu}$ at the Belle II experiment at the SuperKEKB. The analysed data sample corresponds to an integrated luminosity of 63 fb^{-1} collected at the $\Upsilon(4S)$ resonance and a sample of 9 fb^{-1} collected at an energy 60 MeV below the resonance between 2019-2021. Since no significant signal was observed, limit of 4.1×10^{-5} was set using CL$_s$ method.

Based on publication **[PRL 127, 181802 (2021)]**

git repository for the this record:
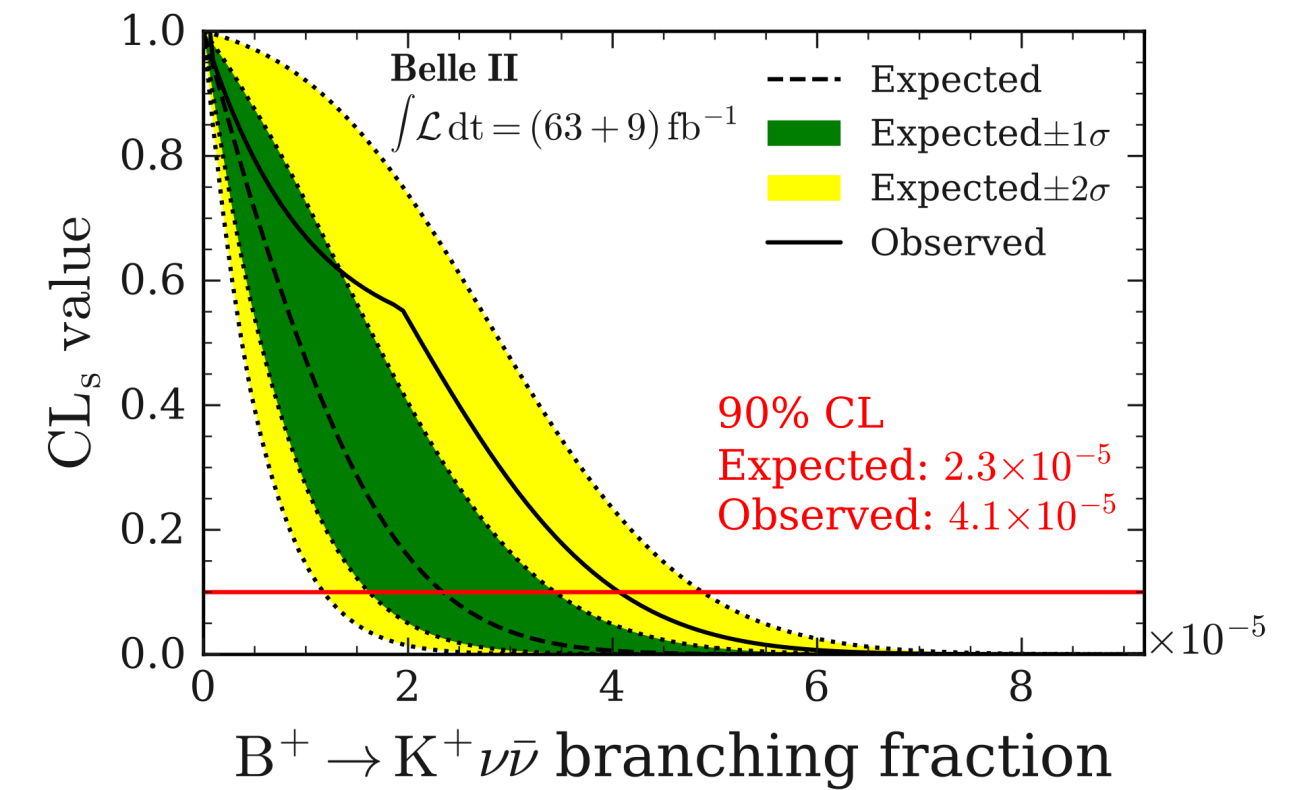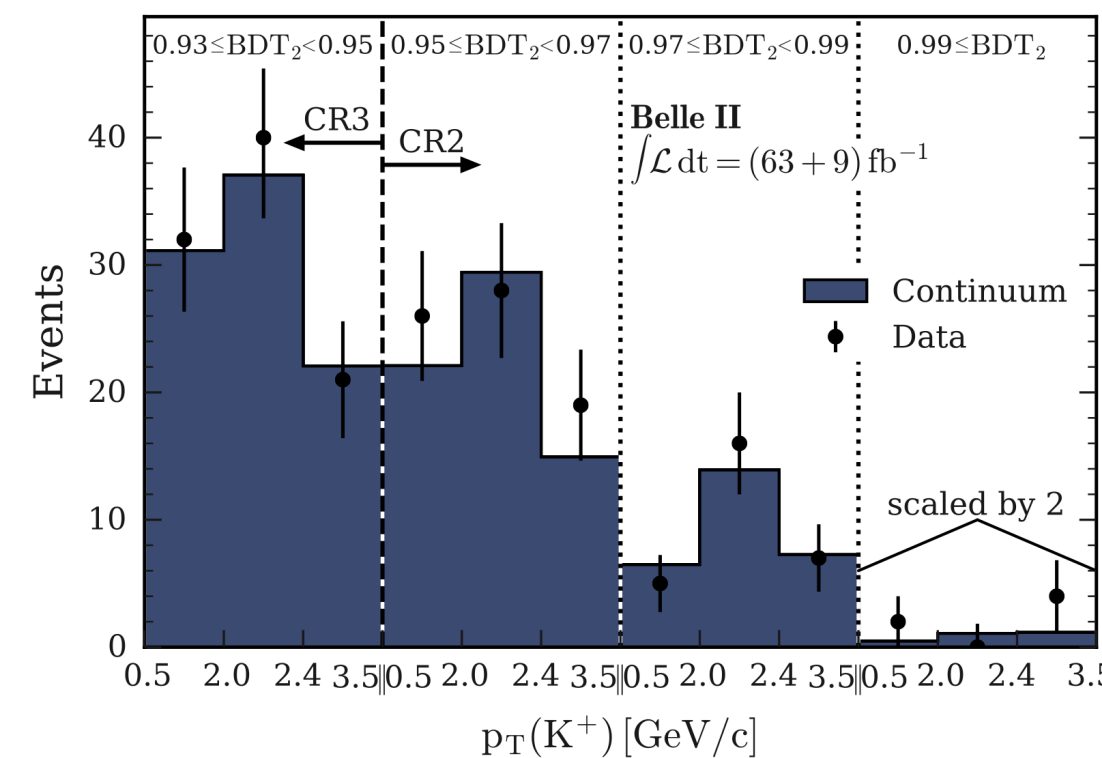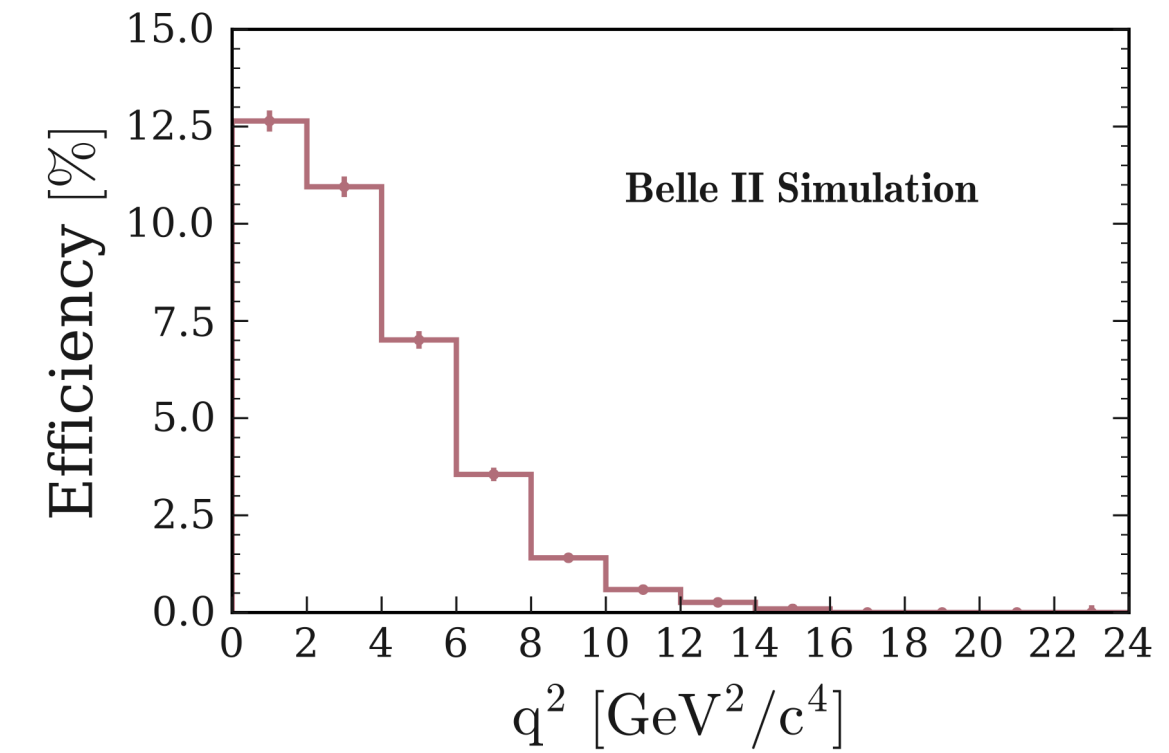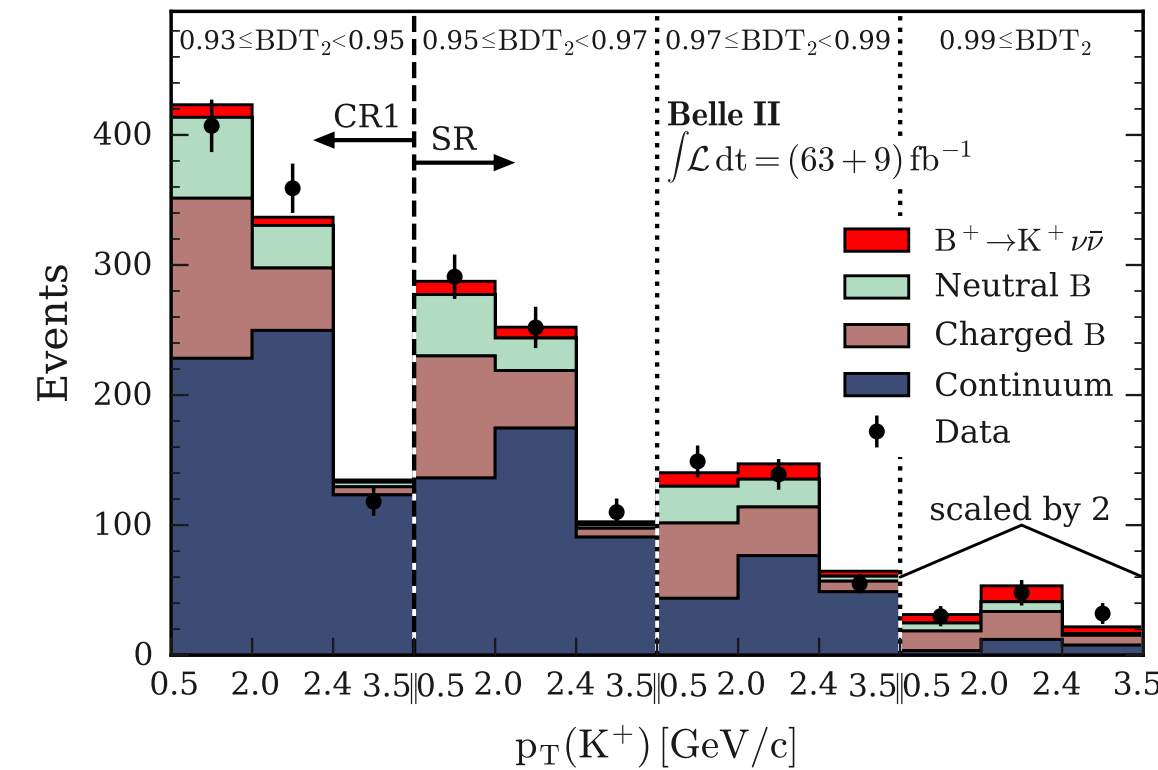**git clone ssh://git@stash.desy.de:7999/~sstefkov/hep_data_b_knunu.git**

HEPData Record for $B^+ \to K^+ \nu \bar{\nu}$

Content of the current $B^+ \to K^+ \nu \bar{\nu}$ submission

1. Main file:
   - *submission.yaml*
2. All data files:
   - *postfit_yields_Y4S.yaml*
   - *postfit_yields_offres.yaml*
   - *limit_B_Knunu.yaml*
   - *efficiency_vs_q2.yaml*
3. All associated figures in png format:
   - *Y4S_postfit.png*
   - *offres_postfit.png*
   - *limit_BF_BKnunu.png*
   - *q2_eff.png*

Y(4S): (Figure 3 in *https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.127.181802*)

Off-resonance: (Figure 1 in *https://journals.aps.org/prl/supplemental/10.1103/PhysRevLett.127.181802/suppl_mat.pdf*)

Expected limit and observed limit: (Figure 2 in *https://journals.aps.org/prl/supplemental/10.1103/PhysRevLett.127.181802/suppl_mat.pdf*)

Selection efficiency: (Figure 3 in *https://journals.aps.org/prl/supplemental/10.1103/PhysRevLett.127.181802/suppl_mat.pdf*)

Main file: *submission.yaml*

HEPData Record for $B^+ \rightarrow K^+\nu\bar{\nu}$

‹ Hide Publication Information

Search for B+→K+νν̄ Decays Using an Inclusive Tagging Method at Belle II

The Belle-II collaboration

Abudinén, F. , Adachi, I. , Adamczyk, K. , Ahlburg, P. , Aihara, H. , Akopov, N. , Aloisio, A. , Ky, N. Anh , Asner, D.M. , Atmacan, H.

Phys.Rev.Lett. 127 (2021) 181802, 2021.

https://doi.org/10.17182/hepdata.130199

Journal | INSPIRE | Resources

4

Abstract (data abstract)
SuperKEKB Belle II. Measurement of the branching fraction of $B^+ \to K^+\nu\bar{\nu}$ at the Belle II experiment at the SuperKEKB. The analysed data sample corresponds to an integrated luminosity of 63 fb^{-1} collected at the $\Upsilon(4S)$ resonance and a sample of 9 fb^{-1} collected at an energy 60 MeV below the resonance between 2019-2021. Since no significant signal was observed, limit of 4.1×10^{-5} was set using CL$_s$ method.

1

⬇ Download All ▾

▽ Filter 5 data tables

Table of contents ›

10.17182/hepdata.130199.v1/t1

- - - - - - - Overview of HEPData Record - - - - - -...

Postfit yields Y(4S) ›

Figure 3 in https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.127.1 10.17182/hepdata.130199.v1/t2
Yields in on-resonance data and as predicted by the simultaneous fit to the on- and off-resonance data, corresponding to an...

Postfit yields off-resonance ›

Figure 1 in https://journals.aps.org/prl/supplemental/10.1103/PhysRe 10.17182/hepdata.130199.v1/t3
Yields in off-resonance data and as predicted by the simultaneous fit to the on- and off-resonance data, corresponding to an...

Expected and observed limit ›

Figure 2 in https://journals.aps.org/prl/supplemental/10.1103/PhysRe 10.17182/hepdata.130199.v1/t4
CL$_s$ value as a function of the branching fraction of $B^+ \to K^+\nu\bar{\nu}$ for expected and observed signal...

Selection efficiency ›

Figure 3 in https://journals.aps.org/prl/supplemental/10.1103/PhysRe 10.17182/hepdata.130199.v1/t5

2

Table of contents 10.17182/hepdata.130199.v1/t1

- - - - - - - Overview of HEPData Record - - - - - - -

Post-fit yields:

- Y(4S)
- Off-resonance

Exclusion limit:

- Expected limit and observed limit

Efficiency:

- Selection efficiency as a function of q^2

Overview

3

HEPData $B^+ \rightarrow K^+ \nu \bar{\nu}$: *submission.yaml*

1: Abstract

‹ Hide Publication Information

Search for B+→K+νν̄ Decays Using an Inclusive Tagging Method at Belle II

The Belle-II collaboration

Abudinén, F. , Adachi, I. , Adamczyk, K. , Ahlburg, P. , Aihara, H. , Akopov, N. , Aloisio, A. , Ky, N. Anh , Asner, D.M. , Atmacan, H.

Phys.Rev.Lett. 127 (2021) 181802, 2021.

https://doi.org/10.17182/hepdata.130199

| Journal | INSPIRE | Resources |

Abstract (data abstract)
SuperKEKB Belle II. Measurement of the branching fraction of $B^+ \rightarrow K^+ \nu \bar{\nu}$ at the Belle II experiment at the SuperKEKB. The analysed data sample corresponds to an integrated luminosity of 63 fb^{-1} collected at the $\Upsilon(4S)$ resonance and a sample of 9 fb^{-1} collected at an energy 60 MeV below the resonance between 2019-2021. Since no significant signal was observed, limit of 4.1×10^{-5} was set using CL$_s$ method.

How?
In *submission.yaml* via comment field

```
comment: | # Information that applies to all data tables.
  SuperKEKB Belle II.  Measurement of the branching fraction of $B^{+} \rightarrow K^{+} \nu \bar{\nu}$ at the B
elle II experiment at the SuperKEKB. The analysed data sample corresponds to an integrated luminosity of 63 fb$^
{-1}$ collected at the $\Upsilon(4S)$ resonance and a sample of 9 fb$^{-1}$ collected at an energy 60 MeV below
the resonance between 2019-2021. Since no significant signal was observed, limit of $4.1\times 10^{-5}$ was set
using CL$_{s}$ method.
---
```

KIT Karlsruher Institut für Technologie

HEPData $B^+ \rightarrow K^+ \nu\bar{\nu}$: $submission.yaml$

2: Clickable links for data tables:

How?

In *submission.yaml* **description** field with HTML links


```
description: '- - - - - - - - Overview of HEPData Record - - - - - - - - <br/><br/>
    </ul>
    <b>Post-fit yields:</b>
    <ul>
    <li><a href="130199?version=1&table=Postfit%20yields%20Y(4S)">Y(4S)</a>
    <li><a href="130199?version=1&table=Postfit%20yields%20off-resonance">Off-resonance</a>
    </ul>
    <b>Exclusion limit:</b>
    <ul>
    <li><a href="130199?version=1&table=Expected%20and%20observed%20Limit">Expected limit and observed limit</a>
    </ul>
    <b>Efficiency:</b>
    <ul>
    <li><a href="130199?version=1&table=Selection%20efficiency">Selection efficiency as a function of $q^{2}$</a>
    </ul>'
```

Table of contents   10.17182/hepdata.130199.v1/t1

-------- Overview of HEPData Record --------

Post-fit yields:

- Y(4S)
- Off-resonance

Exclusion limit:

- Expected limit and observed limit

Efficiency:

- Selection efficiency as a function of q^2

KIT — Karlsruher Institut für Technologie

HEPData $B^+ \rightarrow K^+ \nu\bar{\nu}$: *submission.yaml*

3: Table of contents

How?

In *submission.yaml* via name, additional resources, data_file, location, description, keywords

```
keywords: []
name: Table of contents
---
# Start of table entries.
# This is Table 1.
name: "Postfit yields Y(4S)"
additional_resources:
- {description: Image file, location: Y4S_postfit.png}
- {description: Thumbnail image file, location: thumb_Y4S_postfit.png}
data_file: postfit_yields_Y4S.yaml
location: Figure 3 in https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.127.181802
description: Yields in on-resonance data and as predicted by the simultaneous fit to the on- and off-resonance d
ata, corresponding to an integrated luminosity of 63 and 9 fb$^{-1}$, respectively. The predicted yields are sho
wn individually for charged and neutral B-meson decays and  the five continuum background categories. The leftmo
st three bins belong to the first control region (CR1) with BDT$_{2} \in [0.93; 0.95]$ and the other nine bins c
orrespond to the signal region (SR), three for each range of BDT$_{2} \in [0.95; 0.97; 0.99; 1.0]$. Each set of
three bins is defined by $p_{T}(K^{+}) \in [0.5; 2.0; 2.4; 3.5] \rm{GeV}/c^{2}$.
keywords: # used for searching, possibly multiple values for each keyword
- {name: reactions, values: [$B^+ \rightarrow K^+\nu\bar\nu$]}
- {name: observables, values: [signal strength $\mu$]}
- {name: phrases, values: [FCNC, b --> s l l transition, electroweak penguin decay, missing energy]}
```

HEPData $B^+ \rightarrow K^+ \nu \bar{\nu}$: *submission.yaml*

4: Resources

- How?
  - In *submission.yaml* via additional_resources fields
- What?
  - List of all additional resources:
    - Link to supplemental material page : PRL
    - png files (figures)


```
additional_resources: # additional references (e.g. experiment TWiki page for analysis)
- {location: "https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.127.181802#supplemental",
  description: "PRL web page with sumplementary material"}
```

```
additional_resources:
- {description: Image file, location: limit_BF_BKnunu.png}
- {description: Thumbnail image file, location: thumb_limit_BF_BKnunu.png}
```

Data Files:
postfit_yields_Y4S.yaml
limit_B_Knunu.yaml
efficiency_vs_q2.yaml

Post-fit yields in Y(4S)

[PRL 127, 181802 (2021)]

Postfit yields Y(4S) 10.17182/hepdata.130199.v1/t2

Figure 3 in https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.127.181802
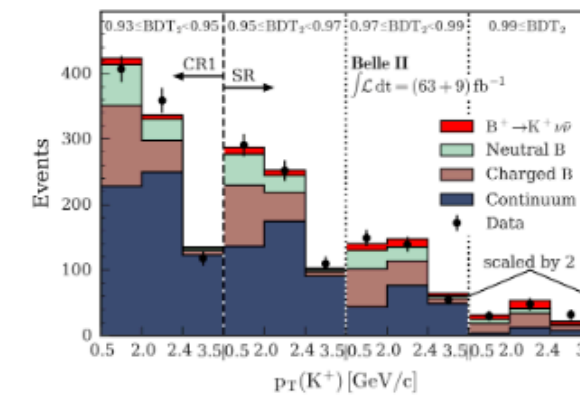
Resources   https://www.hepdata.net   JS

Yields in on-resonance data and as predicted by the simultaneous fit to the on- and off-resonance data, corresponding to an integrated luminosity of 63 and 9 fb^{-1}, respectively. The predicted yields are shown individually for charged and neutral B-meson decays and the five continuum background categories. The leftmost three bins belong to the first control region (CR1) with BDT $_2 \in [0.93; 0.95]$ and the other nine bins correspond to the signal region (SR), three for each range of BDT$_2 \in [0.95; 0.97; 0.99; 1.0]$. Each set of three bins is defined by $p_T(K^+) \in [0.5; 2.0; 2.4; 3.5]\text{GeV/c}^2$.

observables

🏷 signal strenghth μ

phrases

🏷 FCNC   🏷 b --> s l l transition

🏷 electroweak penguin decay   🏷 missing energy

reactions

🏷 $B^+ \rightarrow K^+ \nu\bar{\nu}$

See next page for details of implementing data tables!

Visualize

	yield			
$p_T \times \mathrm{BDT}_2$ **bins**	**Observed data**	**Number of signal events** $B^+ \rightarrow K^+ \nu\bar{\nu}$	**Number of events from charged** B **backgrounds**	**Number of events from neutral** B **backgrounds**	**Number of events from** $c\bar{c}$ **backgrounds**
$[0.5; 2.0] \times [0.93; 0.95]$	407.0 ±20.2	9.6767794	123.073966	62.1993461	144.517922
$[2.0; 2.4] \times [0.93; 0.95]$	359.0 ±18.9	6.32168858	47.9504222	32.7334206	150.685436
$[2.4; 3.5] \times [0.93; 0.95]$	118.0 ±10.9	1.6782059	6.20851002	3.56989512	63.613553
$[0.5; 2.0] \times [0.95; 0.97]$	291.0 ±17.1	10.2102786	93.8582929	47.179443	82.5892068

Post-fit yields in Y(4S):

- **Histogram in with several components**
- **In HEPData encoded with 1 independent variable and 1 dependent variable**

postfit_yields_Y4S.yaml
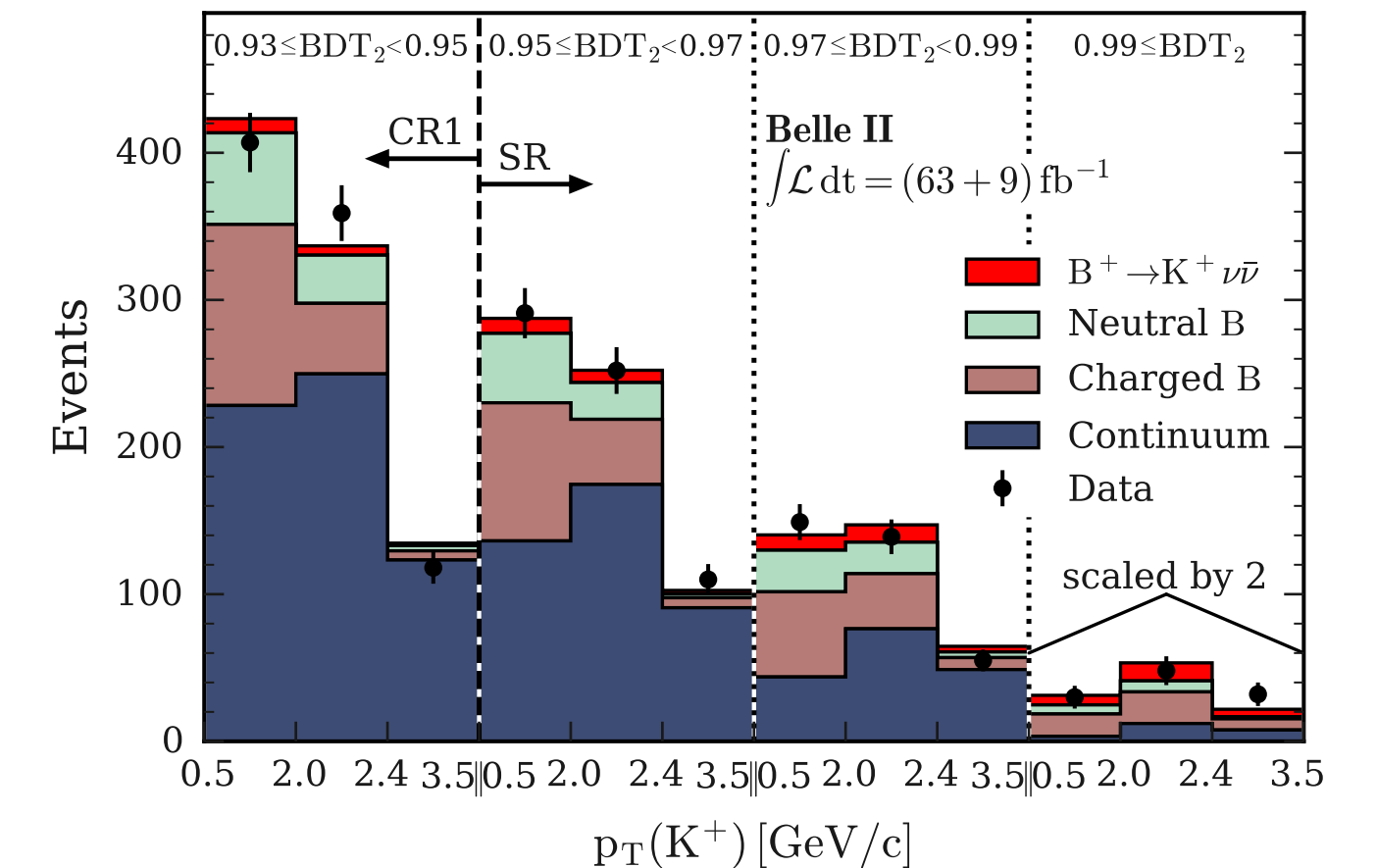
- **independent_variables** (x-axis)
  - header: {name: …} (x-axis name)
  - values:
    - value: … (bin name)**
- **dependent_variables** (y-axis)
  - header: {name: …} (y-axis component 1 name)
  - qualifiers: {name: …} (x-tra metadata, yield)
  - values:
    - value: … (value 1)
    - errors:
      - symerror: … (error on the value 1) ***
    - value: …. (value 2)
  - header: {name: …} (y-axis component 2 name)

** In general: `low` and `high` bin limits, together with a central `value` e.g

```
values:
- {low: 0, high: 60}
```

*** In general could also have asymmetric errors e.g. `asymerror: {plus: 0.4, minus: -0.3}`

HEPData $B^+ \rightarrow K^+ \nu\bar{\nu} : limit\_B\_Knunu.yaml$

Limits:

- Curves with observed central values, expected central values, 1 and 2 sigma expected values
- In HEPData encoded with 1 independent variable, 2 dependent variables
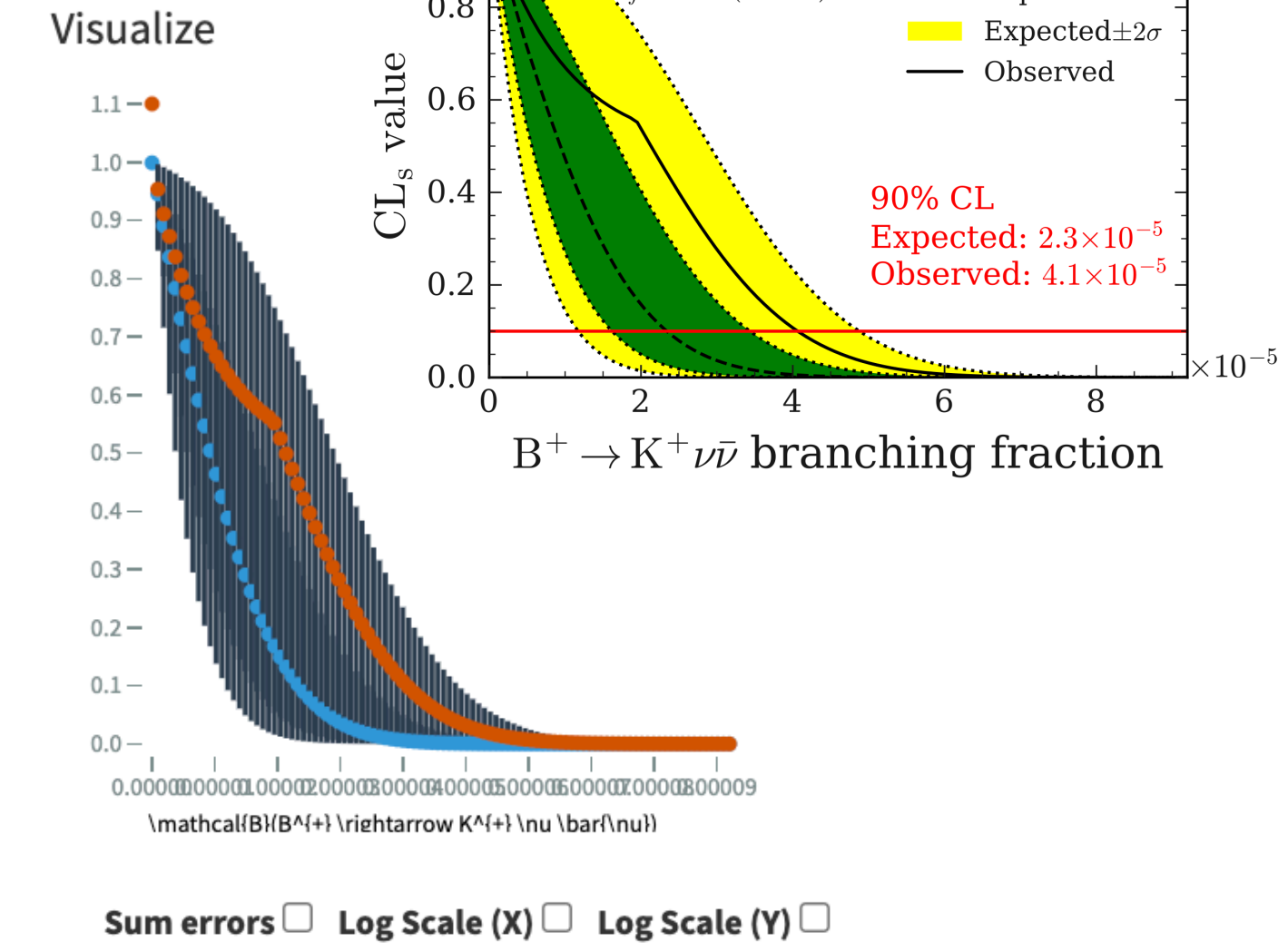
[PRL 127, 181802 (2021)]

```
independent_variables:
- header: {name: '$\mathcal{B}(B^{+} \rightarrow K^{+} \nu \bar{\nu})$'}
  values:
    - {value: 0.0}
    - {value: 9.292929292929293e-07}
    - {value: 1.8585858585858586e-06}
```

```
dependent_variables:
- header: {name: 'CL$_{s}$ value'}
  qualifier:
  {name: 'Limit' , value: 'Observed'}
  {name: 'Luminosity' , units: 'fb$^{-1}$', value: 63+9}
  values:
    - {value: 1.1}
    - {value: 0.953245962}
    - {value: 0.910805129}
    - {value: 0.872237598}
```

```
- header: {name: 'CL$_{s}$ value'}
  qualifier:
  {name: 'Limit' , value: 'Expected'}
  {name: 'Luminosity' , units: 'fb$^{-1}$', value: 63+9}
  values:
  - value: 0.99590567
    errors:
    {assymerror: plus: 0.999502538 , minus: 0.999904355}, label: '1 sigma'}
    {assymerror: plus: 0.997366525 , minus: 0.998621089}, label: '2 sigma'}
  - value: 0.847729031
    errors:
    {assymerror: plus: 0.979556031 , minus: 0.995933079}, label: '1 sigma'}
    {assymerror: plus: 0.898783907 , minus: 0.945208474}, label: '2 sigma'}
  - value: 0.715606801
```

Showing 50 of 100 values   Show All 100 values

Limit	Observed	Expected		
Luminosity	63+9 fb^{-1}			
$\mathcal{B}(B^+ \rightarrow K^+\nu\bar{\nu})$	**CL$_s$ value**			
0.0	1.1	0.998621089 +0.0008814489999999786 / -0.0012545640000000136	1 sigma	
		+0.0012832659999999496 / -0.002715419000000052	2 sigma	
9.292929292929293e-07	0.953245962	0.945208474 +0.03434755700000003 / -0.04642456699999997	1 sigma	+0.05072460500000009 / -0.097479443 2 sigma
1.8585858585858586e-06	0.910805129	0.890664314 +0.06709263499999996 / -0.085949371	1 sigma	+0.1006238590000004 / -0.17505751300000005 2 sigma
2.787878787878788e-06	0.872237598	0.836616884 +0.09800288000000001 / -0.1189236729999995	1 sigma	+0.14939054699999998 / -0.235116985999999997 2 sigma
3.7171717171717173e-06	0.837177085	0.783342012 +0.12684882100000006 / -0.14572960699999993	1 sigma	+0.19669320499999998 / -0.279873672 2 sigma
4.646464646464647e-06	0.805284872	0.730882399 +0.153539105 / -0.16687991299999994	1 sigma	+0.2424039460000004 / -0.31153037399999994 2 sigma
5.575757575757576e-06	0.776264559	0.683497347 +0.17605125100000008 / -0.181713468	1 sigma	+0.28282856 / -0.3306736629999995 2 sigma
6.505050505050505e-06	0.749850745	0.636447247 +0.19676304700000002 / -0.192593318	1 sigma	+0.322018241 / -0.341707891 2 sigma
7.434343434343434e-06	0.725805403	0.590765945 +0.21515509499999996 / -0.19961746200000002	1 sigma	+0.35900842099999997 / -0.34554721 2 sigma
8.363636363636363e-06	0.703914061	0.54663455 +0.231151279 / -0.20319673400000005	1 sigma	+0.39357704299999996 / -0.34342043200000005 2 sigma
9.292929292929294e-06	0.683984266	0.504201266 +0.244713371 / -0.2037486450000006	1 sigma	+0.42553774199999994 / -0.336447649 2 sigma

Visualize

\mathcal{B}(B^{+} \rightarrow K^{+} \nu \bar{\nu})

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

Deselect variables or hide different error bars by clicking o

Variables

CL$_s$ value
Limit:Observed

CL$_s$ value
Limit:Expected

1 sigma error
2 sigma error

Belle II $\int \mathcal{L} dt = (63+9) fb^{-1}$

- - - Expected
▓ Expected±1σ
▓ Expected±2σ
— Observed

90% CL
Expected: 2.3×10^{-5}
Observed: 4.1×10^{-5}

$B^+ \rightarrow K^+\nu\bar{\nu}$ branching fraction

Slavomira Stefkova, Belle II Data Preservation Workshop , 07.10.2022   14

# HEPData $B^+ \to K^+ \nu\bar\nu$: *efficiency_vs_q2.yaml*

Efficiency as a function of $q^2(\nu\bar\nu)$:

- **Values with error bars**
- **In HEPData encoded with 1 independent variable and 1 dependent variable**

```
independent_variables:
- header: {name: '$q^{2}$',units: 'GeV$^2/c^4$'}
  values:
  - {low: 0.0, high: 2.0}
  - {low: 2.0, high: 4.0}
  - {low: 4.0, high: 6.0}
  - {low: 6.0, high: 8.0}
  - {low: 8.0, high: 10.0}
  - {low: 10.0, high: 12.0}
  - {low: 12.0, high: 14.0}
  - {low: 14.0, high: 16.0}
  - {low: 16.0, high: 18.0}
  - {low: 18.0, high: 20.0}
  - {low: 20.0, high: 22.0}
  - {low: 22.0, high: 24.0}
dependent_variables:
- header: {name: 'Efficiency'}
  qualifiers:
  - {name: Luminosity , units: 'fb$^{-1}$', value: 63+9}
  values:
  - value: 12.66745696
    errors:
    - symerror: 0.27207295
  - value: 10.82571692
    errors:
    - symerror: 0.26463688
```


Belle II Simulation
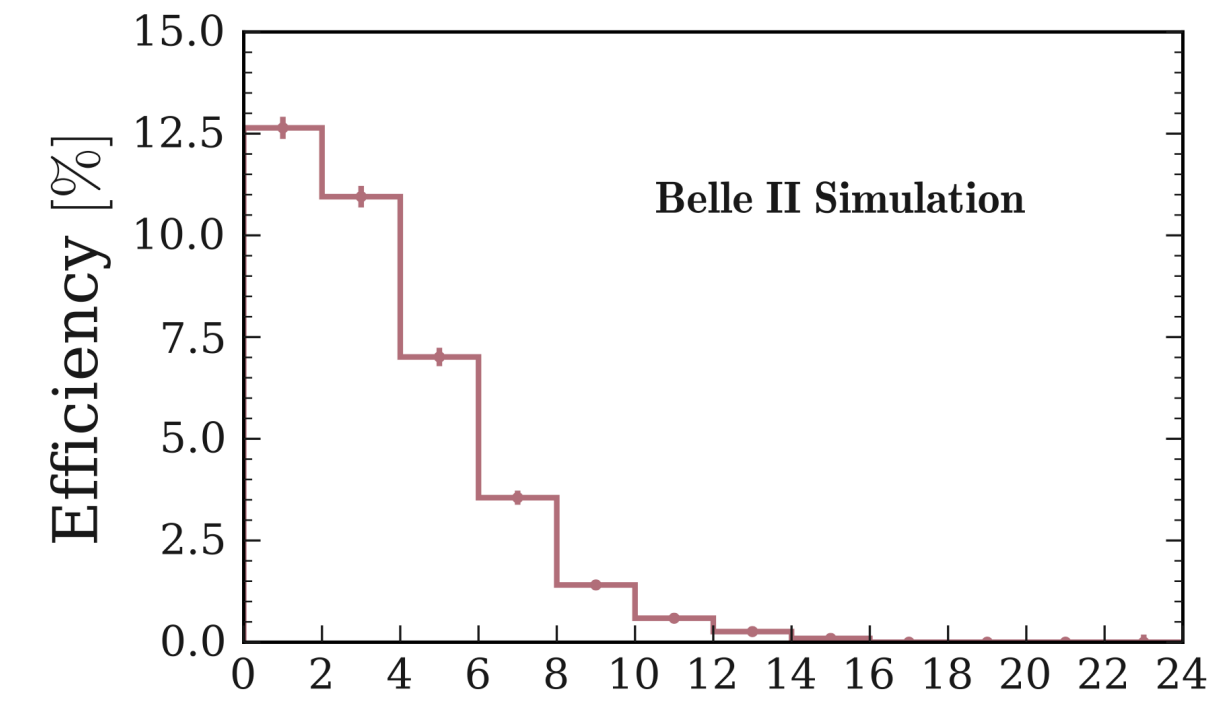
Luminosity	$63{+}9\ \mathrm{fb}^{-1}$
q^2 [GeV$^2/c^4$]	**Efficiency**
0.0 - 2.0	12.66745696 ±0.27207295
2.0 - 4.0	10.82571692 ±0.26463688
4.0 - 6.0	7.04488885 ±0.2278063
6.0 - 8.0	3.51769225 ±0.1711566
8.0 - 10.0	1.46683133 ±0.11813559
10.0 - 12.0	0.68175914 ±0.08670158
12.0 - 14.0	0.27954132 ±0.06074916
14.0 - 16.0	0.12747088 ±0.04571342

Visualize

Sum errors ☐ Log Scale (X) ☐ Log Scale (Y) ☐
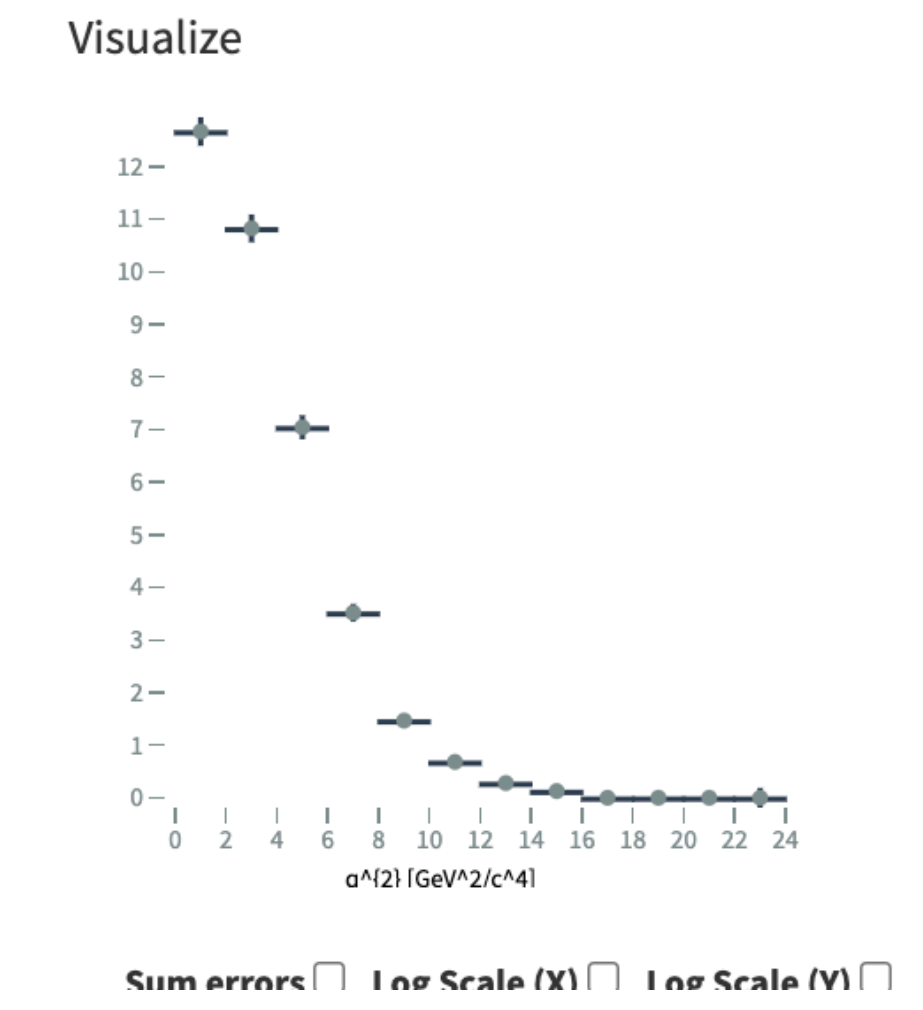
Future plans for $B^+ \to K^+ \nu \bar{\nu}$ submission

HEPData $B^+ \to K^+ \nu \bar{\nu}$: Likelihood

At Belle II we want to provide likelihoods for the measurements:

1. Discuss and agree within collaboration about what format we want to provide likelihoods
2. Submit a revision to HEPData entry with likelihood and/or other information

How to do it in HEPData?

- No forced format: can be uploaded under additional_resources
- $B^+ \to K^+ \nu \bar{\nu}$ analysis used *pyhf (binned likelihood)* → can provide the pyhf json file
- This has already been done for other experiments: e.g examples from ATLAS https://www.hepdata.net/record/ins2072870

Repository for publication-related High-Energy Physics data

How to make submission in Belle II

Conclusion

We want more examples from Belle II!

- Make HEPData submissions
- Please keep data of your plots ideally in json or yaml format
- Please keep your likelihood in a format (also ideally json)

- In Belle II we will prepare confluence page: https://confluence.desy.de/display/BI/HEPData
- Lu Cao will show you an example from Belle, encoding 2D measurements

Purposes of HEPData
1. Facilitate data preservation
2. Give everyone access to HEP results in data format →combinations
3. Serve as a platform for physics reinterpretation

Backup

HEPData Submission in Belle II

If you would like to make HEPData record:

1. Email mgt-hepdata-coordinators@belle2.org to request a submission with following information:
   - **Inspire ID** of your record: e.g https://inspirehep.net/literature/**1860766**
   - Your name and email address
2. Coordinators will create a dashboard for you and you will be able to submit your record after receiving an automatised email from submissions@hepdata.net
3. Upload and submit your record
4. Coordinators will review your record
5. After successful review, coordinators will publish the HEPData record