Combining $R(D^{(*)})$ measurements at Belle II

Ilias Tsaklidis on behalf of the Belle II Collaboration

itsaklid@uni-bonn.de

24.09.2024
Challenges in semi-leptonic B decays, Vienna

UNIVERSITÄT BONN
BLUB
Belle group
University of Bonn
Bundesministerium
für Bildung
und Forschung
Belle II

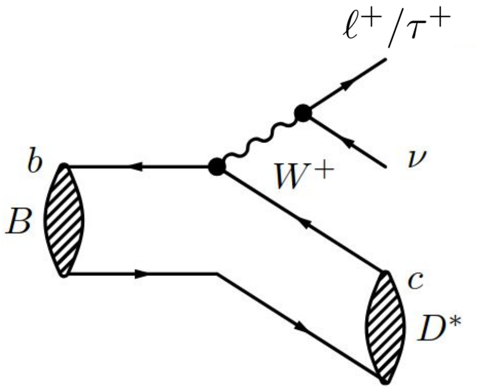
Lepton Flavor Universality

SM: Coupling strengths of e/μ and τ to the electroweak bosons are equal

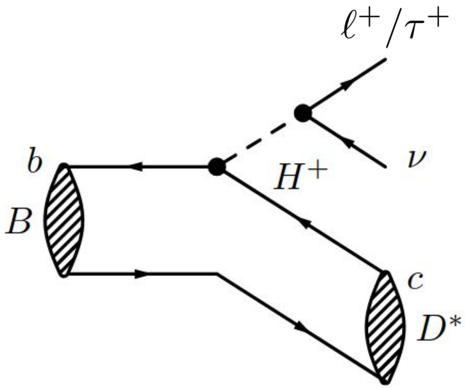
Challenged by experimental measurements
Any deviation would be a **clear sign of BSM** physics processes

Potential tree level contributions:

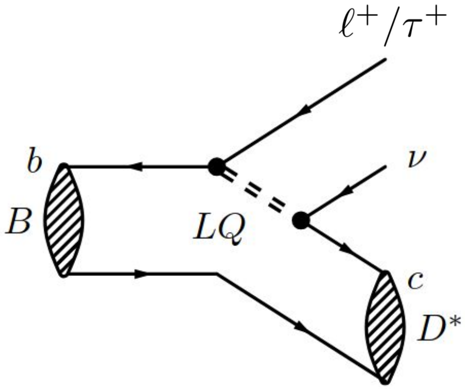
Standard Model



Charged Higgs doublet



LeptoQuarks

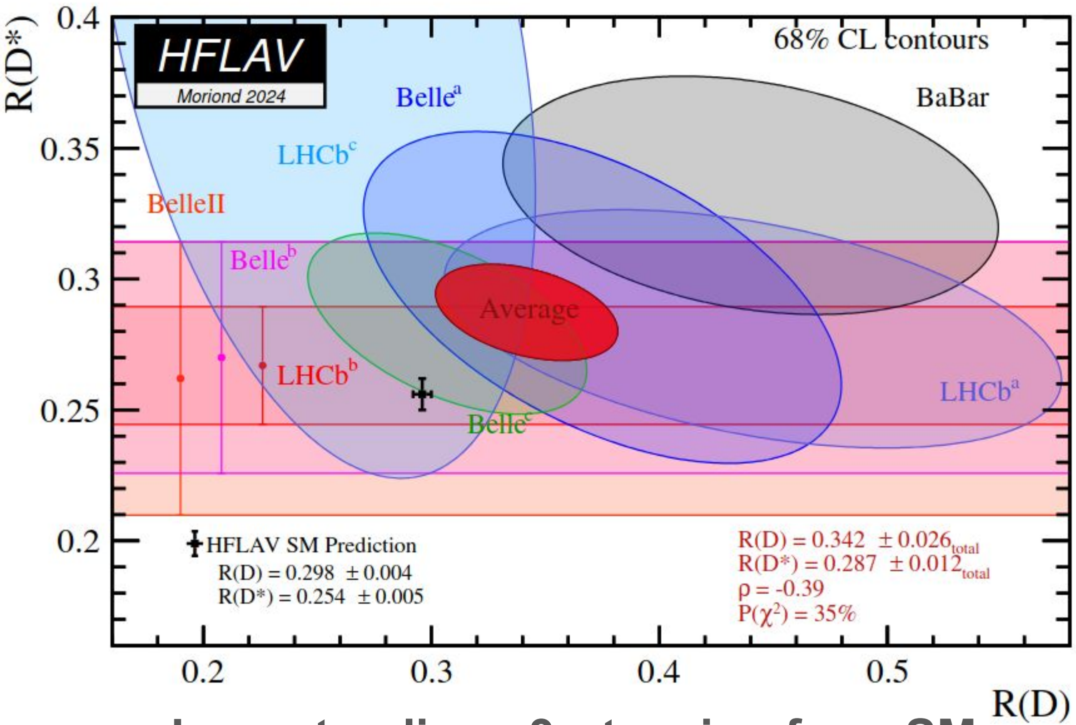


Status of $R(D^{(*)})$

$$R(D^{(*)}) = \frac{\mathcal{B}(\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau)}{\mathcal{B}(\overline{B} \to D^{(*)}\ell^-\overline{\nu}_\ell)}$$

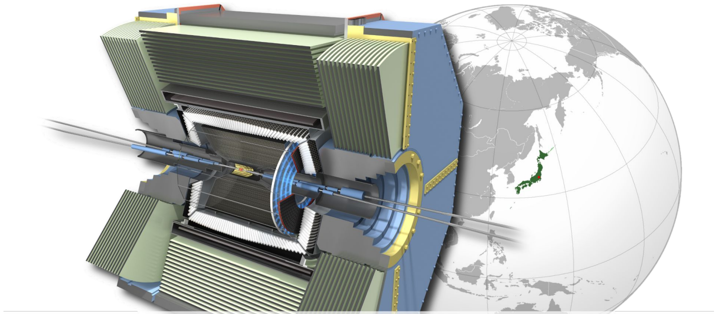
An excellent test of LFU

1. **Direct** test of LFU
2. Precise **theoretical** prediction
3. **Uncertainties** that partially **cancel out** in the ratio:
    a. Hadronic b→c **Form Factor**
    b. **BFs**
    c. Experimental **efficiencies**

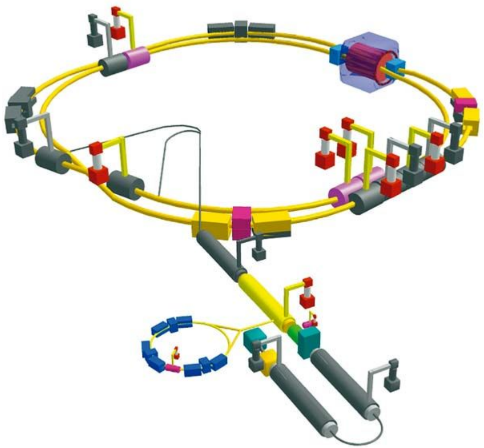


Longstanding ~3σ tension from SM measured by LHCb, BaBar & Belle (II)

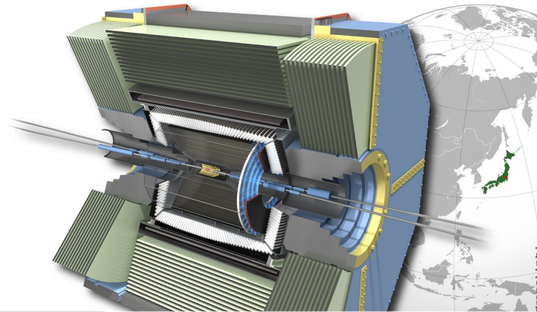
The Belle II detector



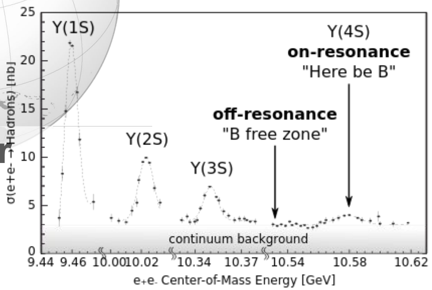
~4π general purpose detector at the interaction point of the SuperKEKB collider



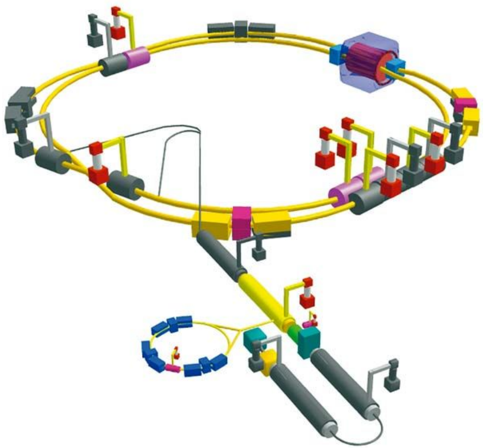
The Belle II detector

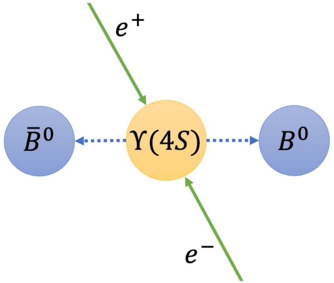


~4π general purpose detector at the interaction point of the SuperKEKB collider

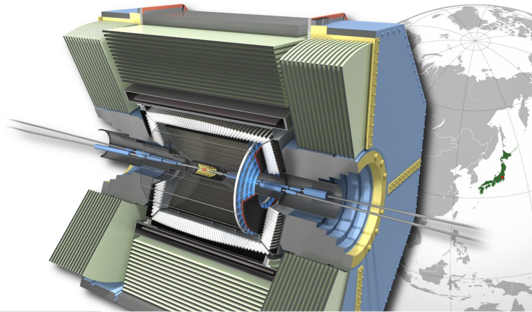


Running at 10.58 GeV

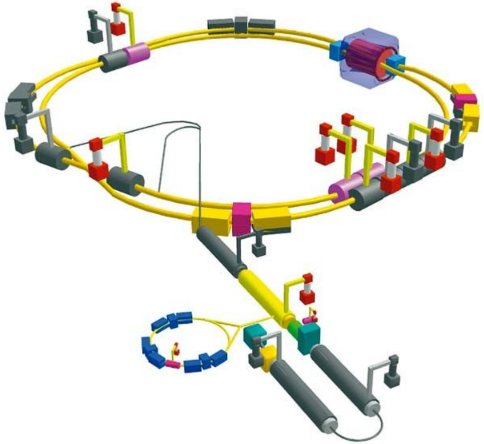


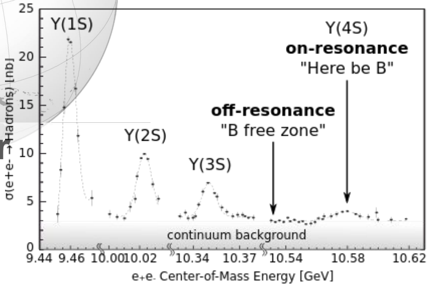


The Belle II detector

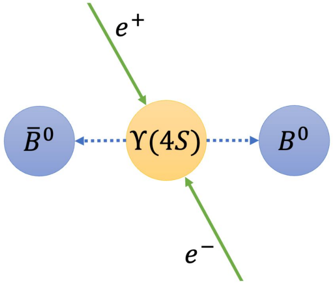


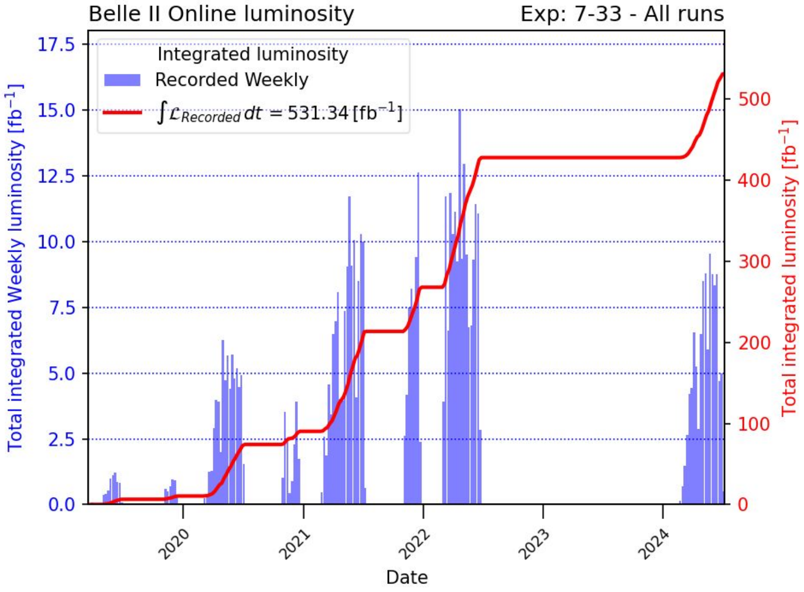
~4π general purpose detector at the interaction point of the SuperKEKB collider





Running at 10.58 GeV

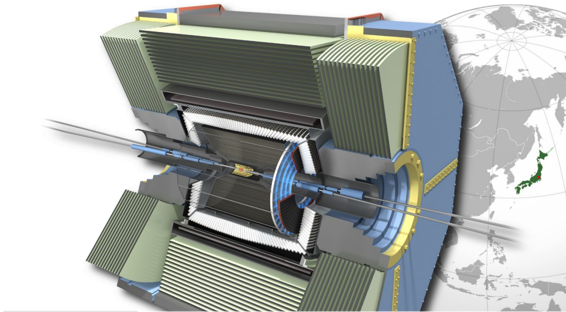




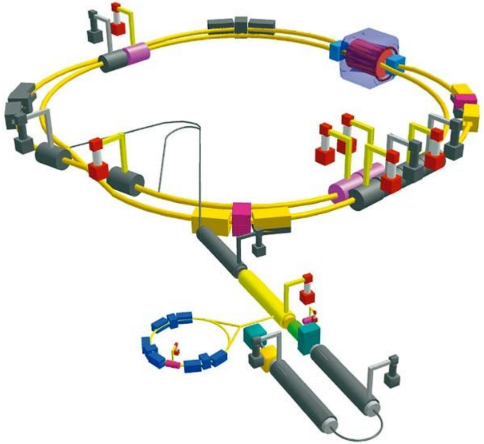
New data collection run starting soon !

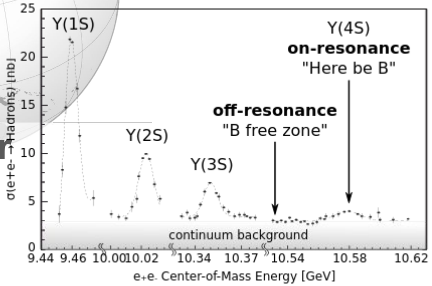
Reaching the total luminosity of Belle 711 fb^{-1} at Y(4S) [1999-2010] in a couple of months

The Belle II detector

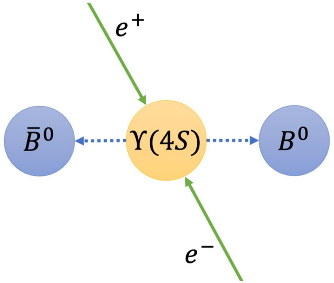


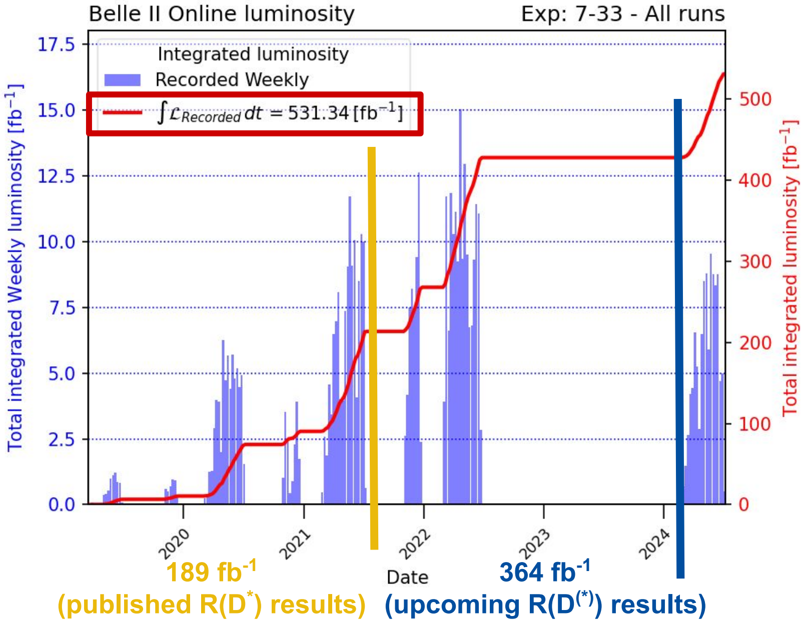
~4π general purpose detector at the interaction point of the SuperKEKB collider





Running at 10.58 GeV





$189\ \mathrm{fb}^{-1}$ (published R(D^*) results)

$364\ \mathrm{fb}^{-1}$ (upcoming R($D^{(*)}$) results)

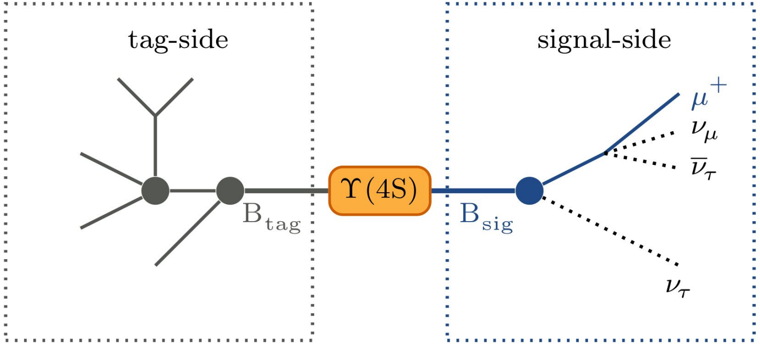
New data collection run starting soon !

Reaching the total luminosity of Belle 711 fb^{-1} at Y(4S) [1999-2010] in a couple of months

R(D$^{(*)}$) opportunities

B-tagging

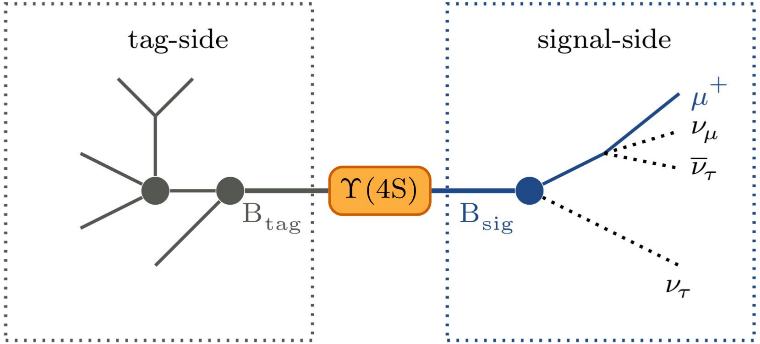
Precise knowledge of the initial state kinematics allows to reconstruct one of the two B mesons and kinematically constrain the second B meson of interest



R(D$^{(*)}$) opportunities

B-tagging

Precise knowledge of the initial state kinematics allows to reconstruct one of the two B mesons and kinematically constrain the second B meson of interest



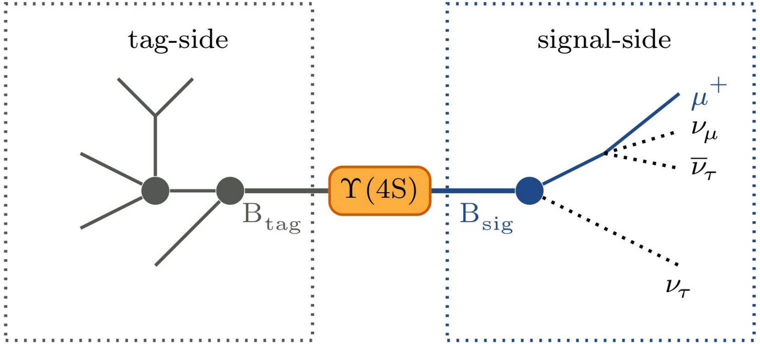
Extremely useful for B-semileptonic decays with missing energy i.e. neutrinos

$$p_{\text{miss}} = (p_{\text{beam}} - p_{B\text{tag}} - p_{D^{(*)}} - p_\ell)$$

R($D^{(*)}$) opportunities

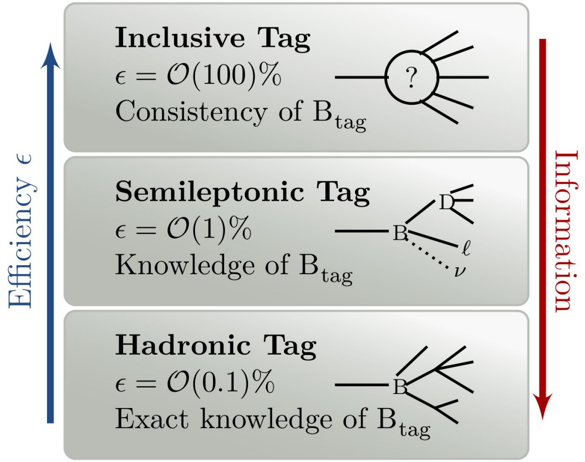
B-tagging

Precise knowledge of the initial state kinematics allows to reconstruct one of the two B mesons and kinematically constrain the second B meson of interest



Extremely useful for B-semileptonic decays with missing energy i.e. neutrinos

$$p_{\text{miss}} = (p_{\text{beam}} - p_{B\text{tag}} - p_{D^{(*)}} - p_\ell)$$



Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of B_{tag}

Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of B_{tag}

Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of B_{tag}

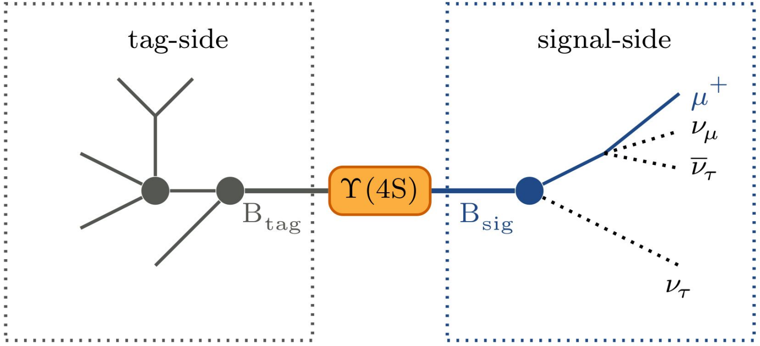
R($D^{(*)}$) opportunities

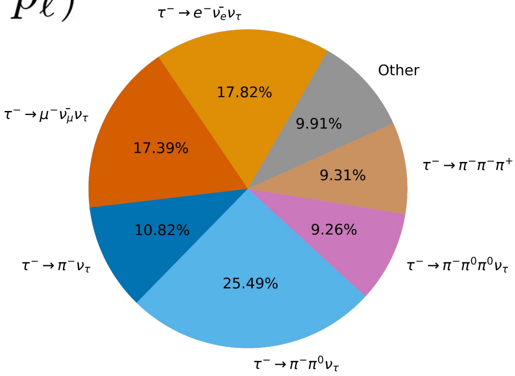
B-tagging

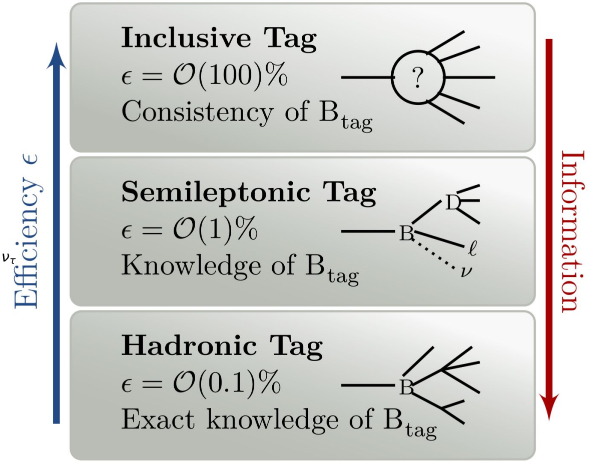
Precise knowledge of the initial state kinematics allows to reconstruct one of the two B mesons and kinematically constrain the second B meson of interest

Extremely useful for B-semileptonic decays with missing energy i.e. neutrinos

$$p_{\text{miss}} = (p_{\text{beam}} - p_{B\text{tag}} - p_{D^{(*)}} - p_\ell)$$







Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of B_{tag}

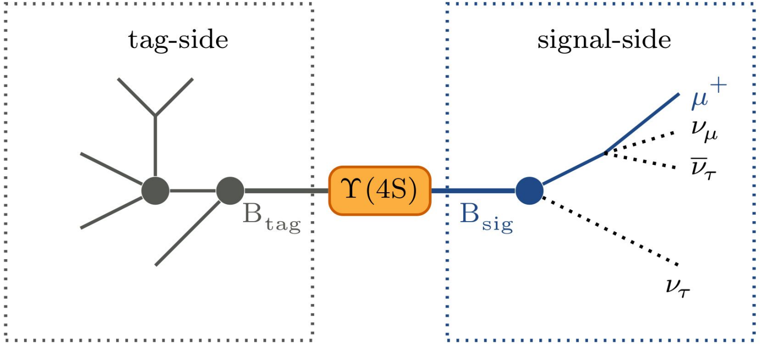
Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of B_{tag}

Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of B_{tag}

R(D$^{(*)}$) opportunities

B-tagging

Precise knowledge of the initial state kinematics allows to reconstruct one of the two B mesons and kinematically constrain the second B meson of interest



Extremely useful for B-semileptonic decays with missing energy i.e. neutrinos

$$p_{\mathrm{miss}} = (p_{\mathrm{beam}} - p_{B\mathrm{tag}} - p_{D^{(*)}} - p_\ell)$$

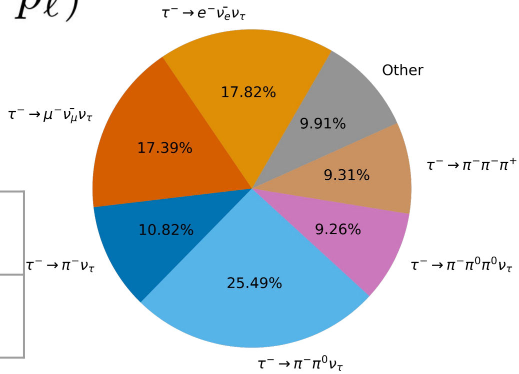
R(D$^{(*)}$) at Belle II

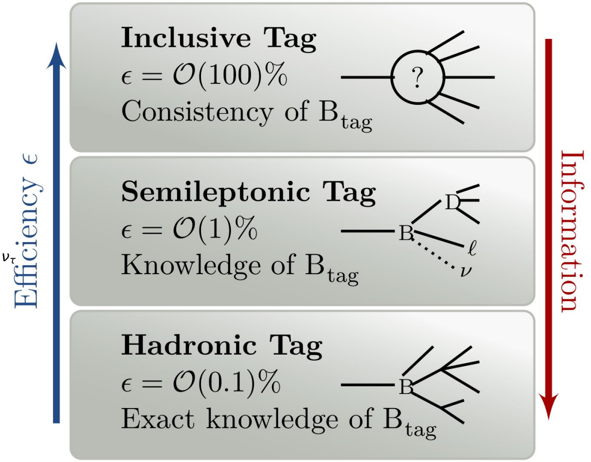
tagging

	hadronic	semileptonic	inclusive
leptonic	✅	✅	✅
hadronic	✅	❌	❌

τ decay

Not impossible but very challenging





Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of $\mathrm{B_{tag}}$

Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of $\mathrm{B_{tag}}$

Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of $\mathrm{B_{tag}}$

Where do we stand on R(D^*) ?

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \to D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \to D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$\boxed{R(D^{*}) = 0.262^{+0.041}_{-0.039}(\text{stat})^{+0.035}_{-0.032}(\text{syst})}$$

Consistent with SM !

Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$R(D^*) = 0.262^{+0.041}_{-0.039}(\text{stat})^{+0.035}_{-0.032}(\text{syst})$$

Consistent with SM !

Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

First ever R(X) measurement at a B factory !

Using **hadronic tag** $R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$
reconstruct a **single lepton** and combine
the rest into an X system inclusively

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \to D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both charged and neutral B mesons

$$R(D^{*}) = 0.262^{+0.041}_{-0.039}(\mathrm{stat})^{+0.035}_{-0.032}(\mathrm{syst})$$

Consistent with SM !

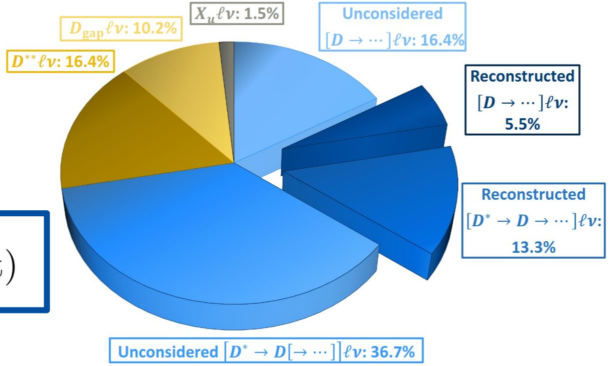
Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

First ever R(X) measurement at a B factory !

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$

Using **hadronic tag**
reconstruct a **single lepton** and combine the rest into an X system inclusively



Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \rightarrow D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$R(D^*) = 0.262^{+0.041}_{-0.039}(\text{stat})^{+0.035}_{-0.032}(\text{syst})$$

Consistent with SM !

Similar precision to Belle
with **25%** of the data

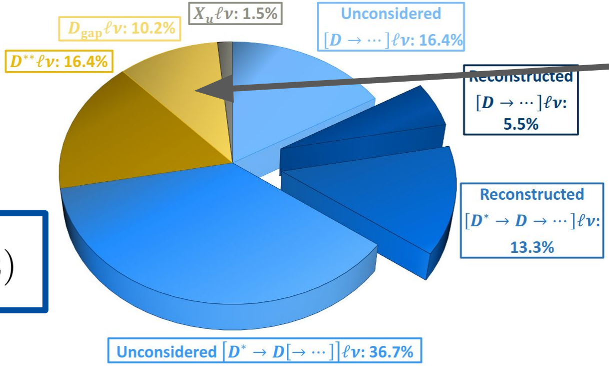
arXiv: 2401.02840

First ever R(X) measurement at a B factory !

Using **hadronic tag**

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$

reconstruct a **single lepton** and combine the rest into an X system inclusively



Gap modes:
The difference between the
sum of **exclusive BFs** to the **inclusive BF**.
Filled in MC with an educated guess

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \rightarrow D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$R(D^{*}) = 0.262^{+0.041}_{-0.039}(\mathrm{stat})^{+0.035}_{-0.032}(\mathrm{syst})$$

Consistent with SM !

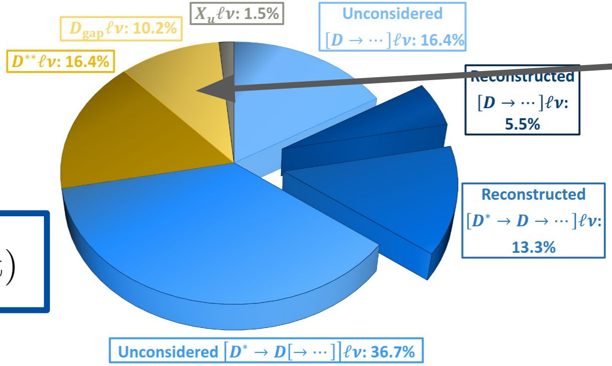
Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

First ever R(X) measurement at a B factory !

Using **hadronic tag**
reconstruct a **single lepton** and combine
the rest into an X system inclusively

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$



Gap modes:
The difference between the
sum of **exclusive BFs** to the **inclusive BF**.
Filled in MC with an educated guess

Consistent with SM !

$$R(X_{\tau/\ell}) = 0.228 \pm 0.016(\mathrm{stat}) \pm 0.036(\mathrm{syst})$$

PhysRevLett.132.211804

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \rightarrow D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$R(D^{*}) = 0.262^{+0.041}_{-0.039}(\text{stat})^{+0.035}_{-0.032}(\text{syst})$$

Consistent with SM !

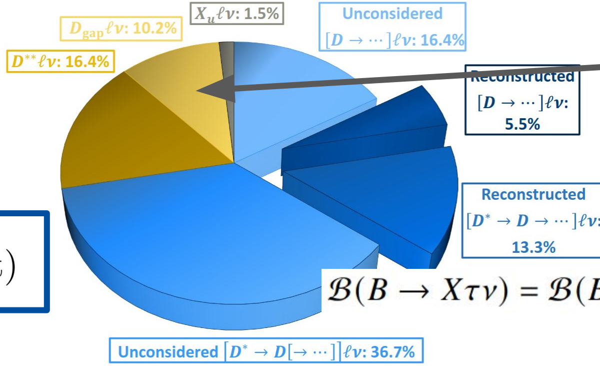
Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

First ever R(X) measurement at a B factory !

Using **hadronic tag**
reconstruct a **single lepton** and combine
the rest into an X system inclusively

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$



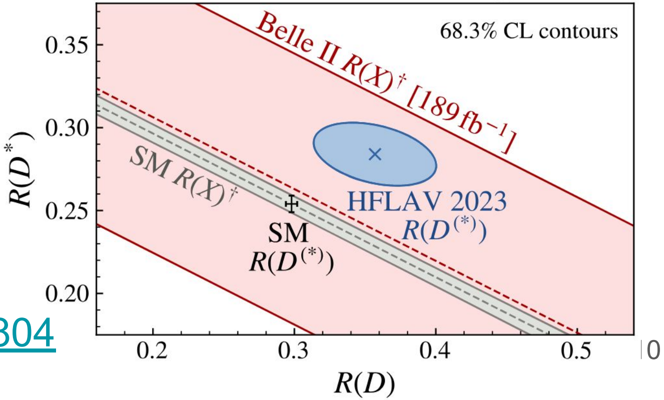
Gap modes:
The difference between the
sum of **exclusive BFs** to the **inclusive BF**.
Filled in MC with an educated guess

Consistent with SM !

$$R(X_{\tau/\ell}) = 0.228 \pm 0.016(\text{stat}) \pm 0.036(\text{syst})$$

$$\mathcal{B}(B \rightarrow X\tau\nu) = \mathcal{B}(B \rightarrow D\tau\nu) + \mathcal{B}(B \rightarrow D^{*}\tau\nu) + \mathcal{B}(B \rightarrow D^{**}_{(\text{gap})}, X_u\tau\nu)$$

† = with expected SM contributions of $D^{**}_{(\text{gap})}, X_u$ removed



PhysRevLett.132.211804

Published results of R(D*) at Belle II using 189 fb^{-1}

First R(D*) measurement at Belle II !

Using **hadronic tag**
Reconstruct $\overline{B} \rightarrow D^{(*)}\tau^{-}\overline{\nu}_{\tau}$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

$$R(D^{*}) = 0.262^{+0.041}_{-0.039}(\mathrm{stat})^{+0.035}_{-0.032}(\mathrm{syst})$$

Consistent with SM !

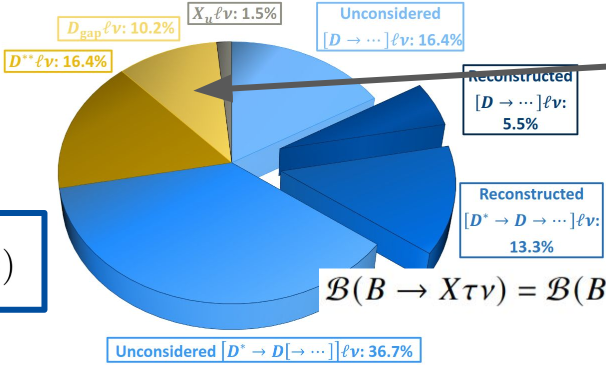
Similar precision to Belle
with **25%** of the data

arXiv: 2401.02840

First ever R(X) measurement at a B factory !

Using **hadronic tag**
reconstruct a **single lepton** and combine
the rest into an X system inclusively

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$



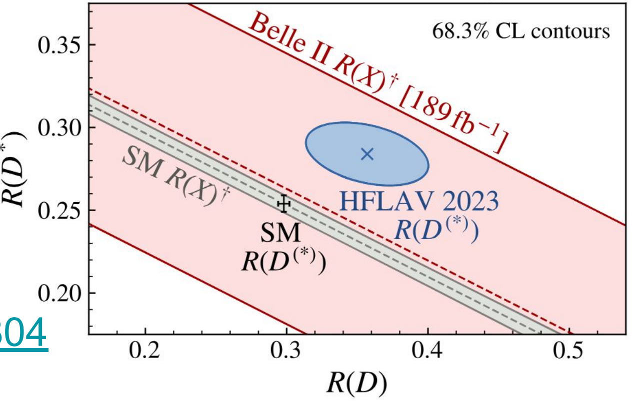
Gap modes:
The difference between the
sum of **exclusive BFs** to the **inclusive BF**.
Filled in MC with an educated guess

Consistent with SM !

$$R(X_{\tau/\ell}) = 0.228 \pm 0.016(\mathrm{stat}) \pm 0.036(\mathrm{syst})$$

$$\mathcal{B}(B \rightarrow X\tau\nu) = \mathcal{B}(B \rightarrow D\tau\nu) + \mathcal{B}(B \rightarrow D^{*}\tau\nu) + \mathcal{B}(B \rightarrow D^{**}_{(\mathrm{gap})}, X_u\tau\nu)$$

† = with expected SM contributions of $D^{**}_{(\mathrm{gap})}, X_u$ removed



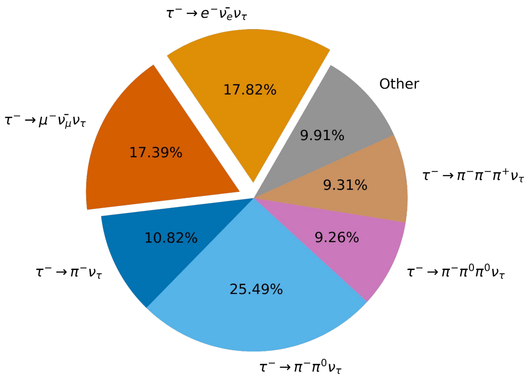
Statistical correlation with R(D*) ~0.02

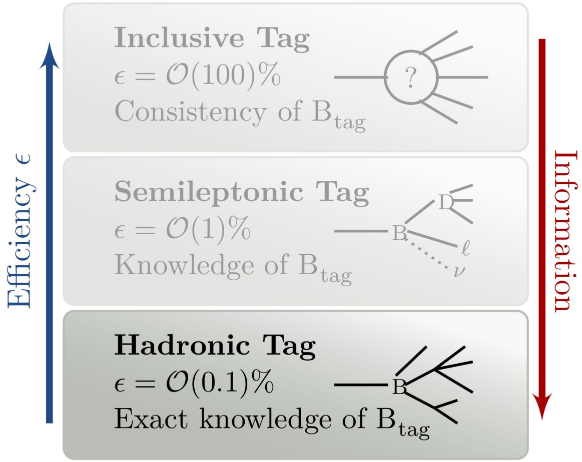
Systematic correlation (mainly D** BFs)
non trivial

PhysRevLett.132.211804

What are we working on?

- **Hadronic tag, leptonic τ**
  - Update R(D*) with full 364 fb^{-1}
  - Measure R(D) simultaneously
  - Further optimize selection
  - Revisit signal extraction strategy





Efficiency ϵ

Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of $\mathrm{B}_{\mathrm{tag}}$

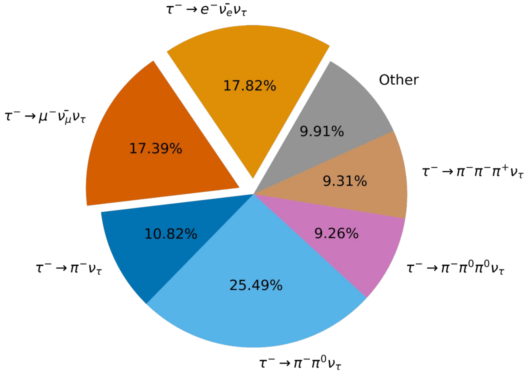
Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of $\mathrm{B}_{\mathrm{tag}}$

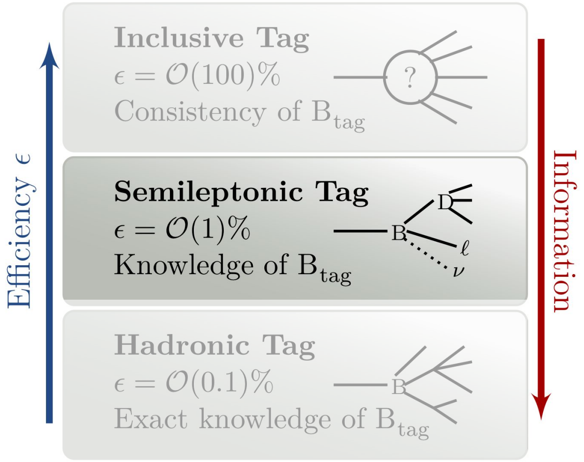
Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of $\mathrm{B}_{\mathrm{tag}}$

Information

What are we working on?

- **Hadronic tag, leptonic τ**
    - Update R(D*) with full 364 fb^{-1}
    - Measure R(D) simultaneously
    - Further optimize selection
    - Revisit signal extraction strategy
- **Semileptonic tag, leptonic τ**
    - Simultaneous measurement of R(D*) and R(D)
    - Completely orthogonal measurement





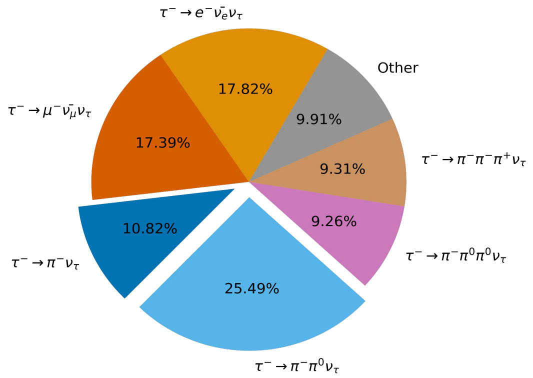
Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of B_{tag}

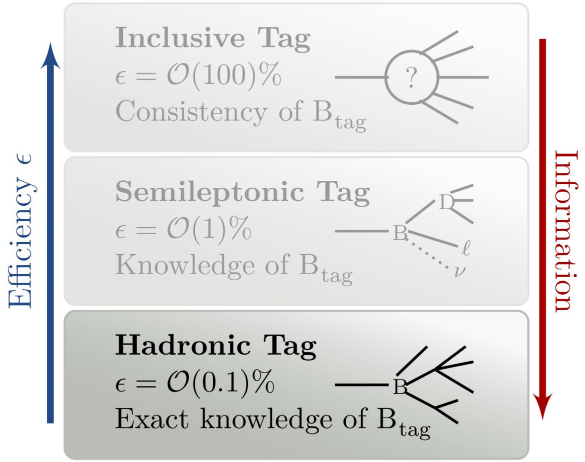
Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of B_{tag}

Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of B_{tag}

What are we working on?

- **Hadronic tag, leptonic τ**
  - Update R(D*) with full 364 fb^{-1}
  - Measure R(D) simultaneously
  - Further optimize selection
  - Revisit signal extraction strategy
- **Semileptonic tag, leptonic τ**
  - Simultaneous measurement of R(D*) and R(D)
  - Completely orthogonal measurement
- **Hadronic tag, hadronic 1-prong τ**
  - Measure R(D*). R(D) challenging due to backgrounds
  - Simultaneous measurement of τ polarization





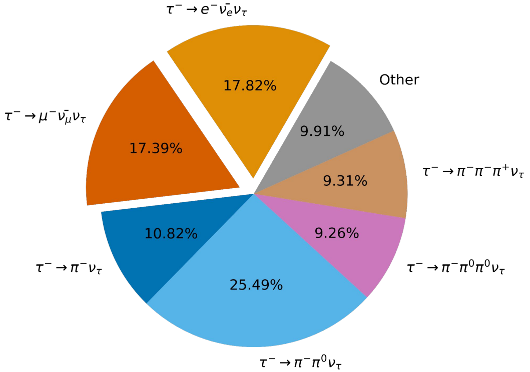
Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of $\mathrm{B}_{\mathrm{tag}}$

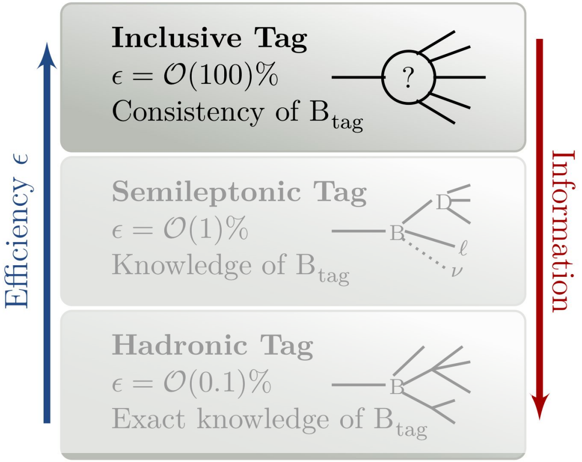
Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of $\mathrm{B}_{\mathrm{tag}}$

Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of $\mathrm{B}_{\mathrm{tag}}$

What are we working on?

- **Hadronic tag, leptonic τ**
  - Update R(D*) with full 364 fb^{-1}
  - Measure R(D) simultaneously
  - Further optimize selection
  - Revisit signal extraction strategy
- **Semileptonic tag, leptonic τ**
  - Simultaneous measurement of R(D*) and R(D)
  - Completely orthogonal measurement
- **Hadronic tag, hadronic 1-prong τ**
  - Measure R(D*). R(D) challenging due to backgrounds
  - Simultaneous measurement of τ polarization
- **Inclusive tag, leptonic τ**
  - Simultaneous measurement of R(D*) and R(D)
  - High reconstruction efficiency but low purity

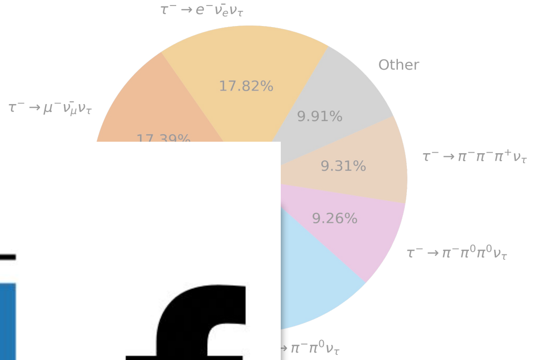


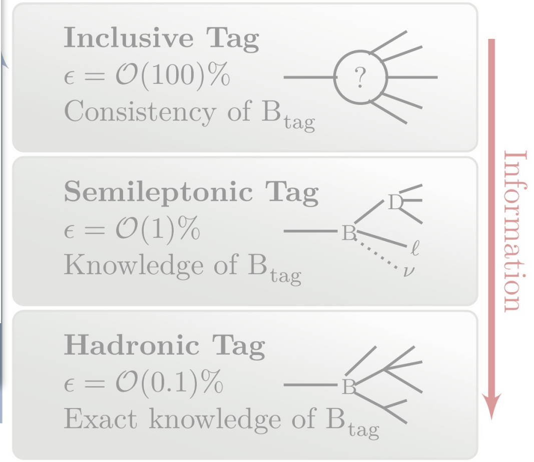


What are we working on?

- **Hadronic tag, leptonic τ**
  - Update R(D*) with full 364 fb^{-1}
  - Measure R(D) si
  - Further optimize
  - Revisit signal ex
- **Semileptonic tag, le**
  - Simultaneous m
  - Completely orth
- **Hadronic tag, hadro**
  - Measure R(D*).
  - Simultaneous m
- **Inclusive tag, leptonic τ**
  - Simultaneous measurement of R(D*) and R(D)
  - High reconstruction efficiency but low purity

pyhf ∂ifferentiable ℒikelihoods





Inclusive Tag
$\epsilon = \mathcal{O}(100)\%$
Consistency of $\mathrm{B_{tag}}$

Semileptonic Tag
$\epsilon = \mathcal{O}(1)\%$
Knowledge of $\mathrm{B_{tag}}$

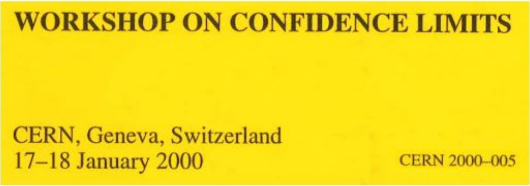
Hadronic Tag
$\epsilon = \mathcal{O}(0.1)\%$
Exact knowledge of $\mathrm{B_{tag}}$

R(D*) combinations by likelihood

pyhf is python based statistical inference library that implements the HistFactory method.
Well established in the LHC experiments with large user community

R(D*) combinations by likelihood

pyhf is python based statistical inference library that implements the HistFactory method. Well established in the LHC experiments with large user community



WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000

CERN 2000–005

Massimo Corradi

Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

Any disagreement ? Carried unanimously. That's actually quite an achievement for this Workshop.

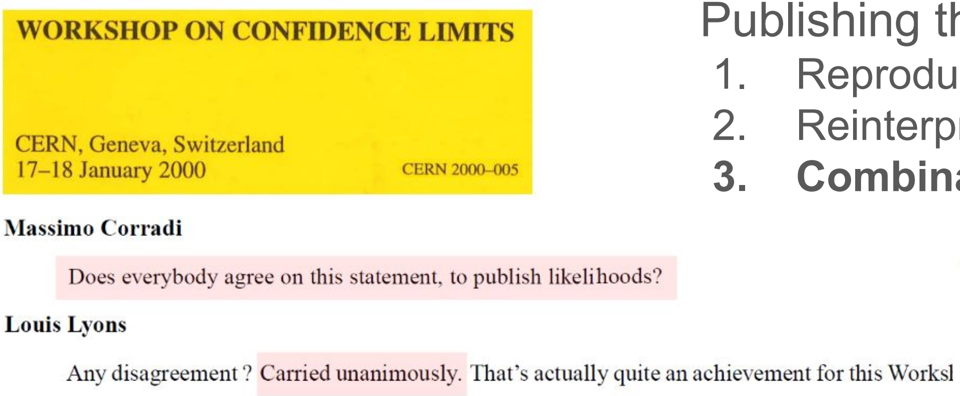
K. Cranmer at PHYSTAT seminar

Publishing the likelihood enables:

1. Reproducibility of research
2. Reinterpretation in a model independent way
3. **Combination of results**

R(D*) combinations by likelihood

pyhf is python based statistical inference library that implements the HistFactory method.
Well established in the LHC experiments with large user community



WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000
CERN 2000–005

Massimo Corradi

Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

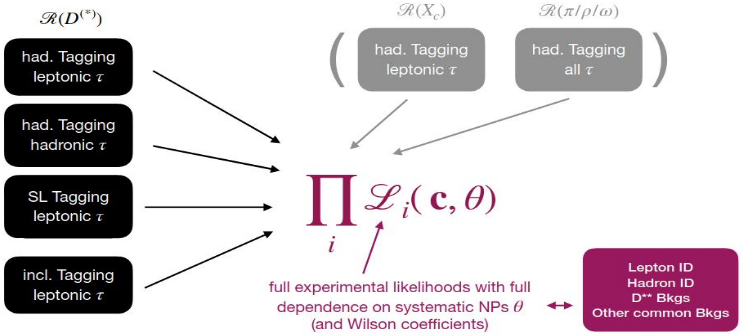
Any disagreement ? Carried unanimously. That's actually quite an achievement for this Worksl

K. Cranmer at PHYSTAT seminar

Publishing the likelihood enables:

1. Reproducibility of research
2. Reinterpretation in a model independent way
3. **Combination of results**

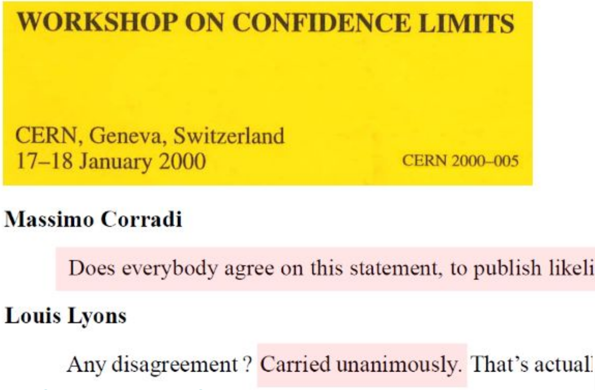
e.g. Belle II



F. Bernlochner at Challenges in SL B decays, 2022

R(D*) combinations by likelihood

pyhf is python based statistical inference library that implements the HistFactory method.
Well established in the LHC experiments with large user community



WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000
CERN 2000–005

Massimo Corradi

Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

Any disagreement ? Carried unanimously. That's actual

K. Cranmer at PHYSTAT seminar

Publishing the likelihood enables:

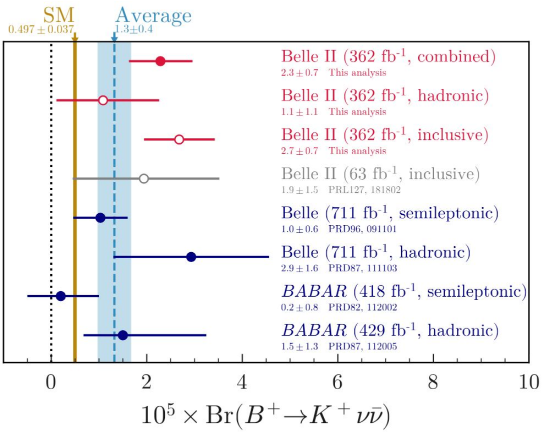
1. Reproducibility of research
2. Reinterpretation in a model independent way
3. **Combination of results**

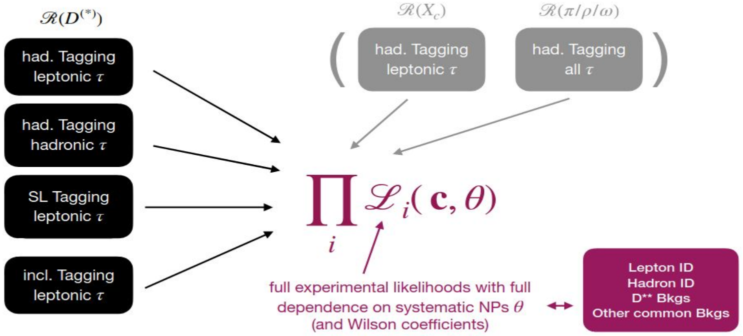
Methodology already tested at a recent Belle II publication

Evidence for $B^+ \to K^+\nu\bar{\nu}$ decays

arXiv:2311.14647

e.g. Belle II



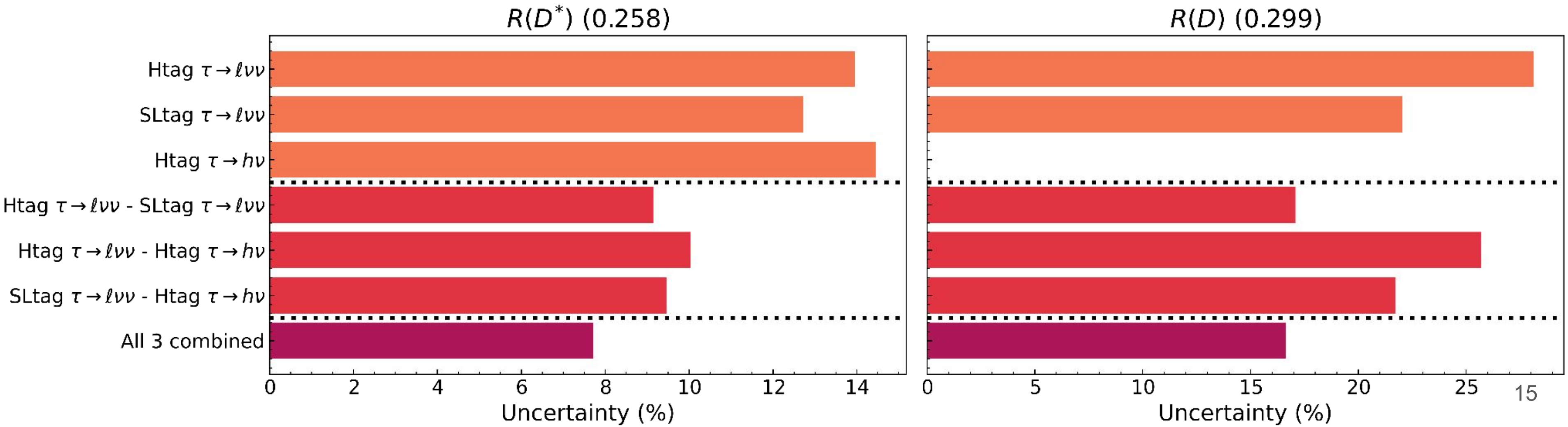


F. Bernlochner at Challenges in SL B decays, 2022

Simple likelihood combination

Combining the likelihoods of 3 preliminary R(D$^{(*)}$) analyses at Belle II

No systematic uncertainties assumed



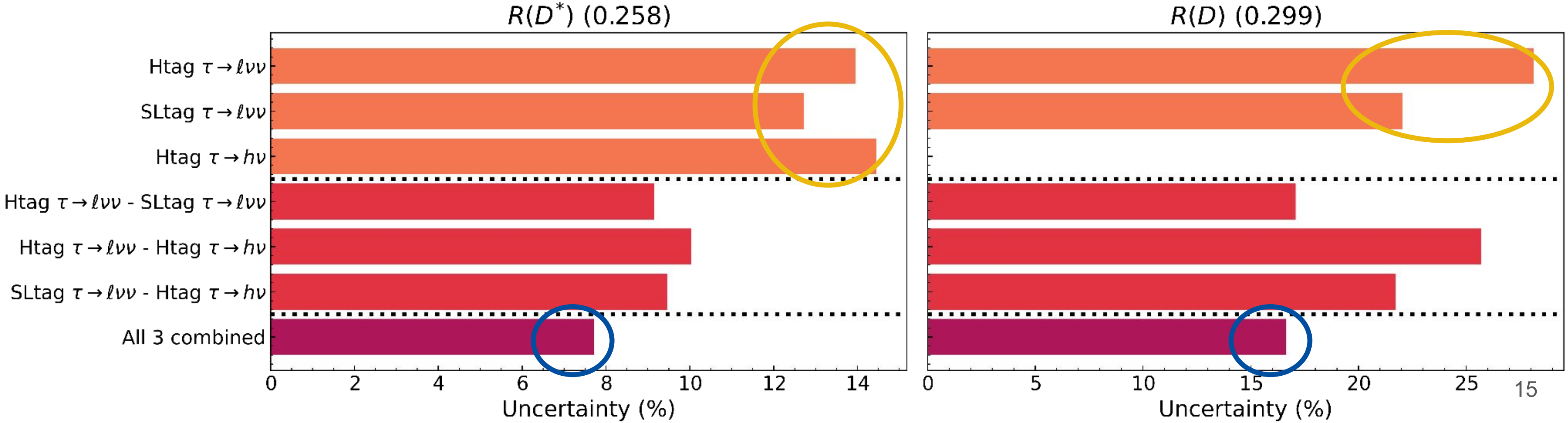
Simple likelihood combination

Combining the likelihoods of 3 preliminary R(D$^{(*)}$) analyses at Belle II

No systematic uncertainties assumed

Combined statistical uncertainty drops from

- **~ 14%** to **< 8%** for R(D*)
- **> 25%** to (**<17%**) for R(D)



What about systematic uncertainties ?

R(D*) combinations by HFLAV

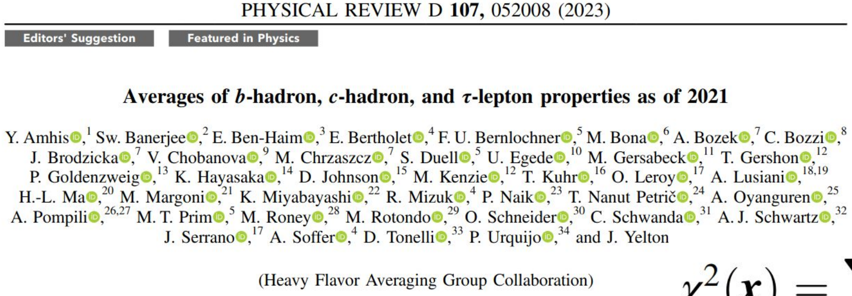
HFLAV

Our **optimal** method so far for:
Combining independent, but potentially **correlated** measurements

R(D*) combinations by HFLAV

Our **optimal** method so far for:
Combining independent, but potentially **correlated** measurements

The averaging method
Is outlined in detail in:



PHYSICAL REVIEW D **107,** 052008 (2023)

Editors' Suggestion | Featured in Physics

Averages of *b*-hadron, *c*-hadron, and *τ*-lepton properties as of 2021

Y. Amhis,[1] Sw. Banerjee,[2] E. Ben-Haim,[3] E. Bertholet,[4] F. U. Bernlochner,[5] M. Bona,[6] A. Bozek,[7] C. Bozzi,[8] J. Brodzicka,[7] V. Chobanova,[9] M. Chrzaszcz,[7] S. Duell,[5] U. Egede,[10] M. Gersabeck,[11] T. Gershon,[12] P. Goldenzweig,[13] K. Hayasaka,[14] D. Johnson,[15] M. Kenzie,[12] T. Kuhr,[16] O. Leroy,[17] A. Lusiani,[18,19] H.-L. Ma,[20] M. Margoni,[21] K. Miyabayashi,[22] R. Mizuk,[4] P. Naik,[23] T. Nanut Petrič,[24] A. Oyanguren,[25] A. Pompili,[26,27] M. T. Prim,[5] M. Roney,[28] M. Rotondo,[29] O. Schneider,[30] C. Schwanda,[31] A. J. Schwartz,[32] J. Serrano,[17] A. Soffer,[4] D. Tonelli,[33] P. Urquijo,[34] and J. Yelton

(Heavy Flavor Averaging Group Collaboration)

global χ^2 statistic
$\mathbf{x}_\mathbf{i}$ set of N independent measurements
$\mathbf{V}_\mathbf{i}$: their covariance matrix

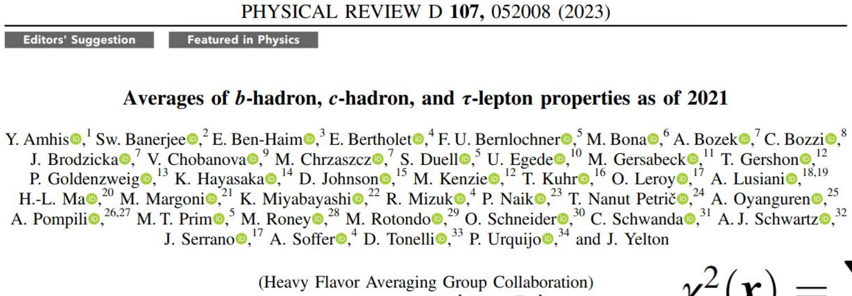
$$\chi^2(\boldsymbol{x}) = \sum_i^N (\boldsymbol{x}_i - \boldsymbol{x})^\mathrm{T} \boldsymbol{V}_i^{-1} (\boldsymbol{x}_i - \boldsymbol{x})$$

R(D*) combinations by HFLAV

HFLAV

Our **optimal** method so far for:
Combining independent, but potentially **correlated** measurements

The averaging method
Is outlined in detail in:



PHYSICAL REVIEW D **107,** 052008 (2023)

Editors' Suggestion Featured in Physics

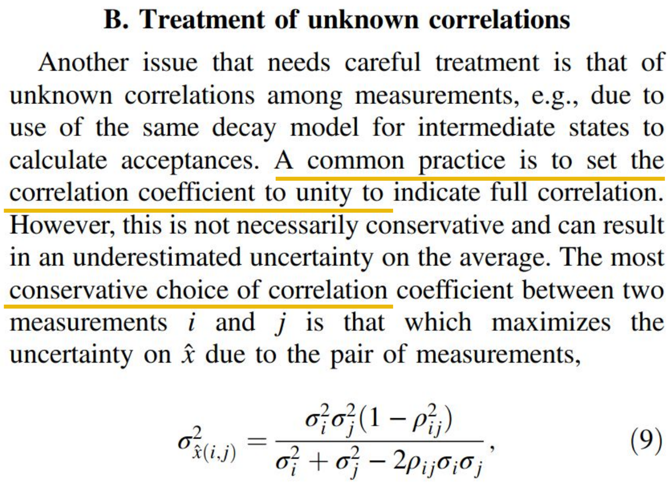
Averages of *b*-hadron, *c*-hadron, and *τ*-lepton properties as of 2021

Y. Amhis,[1] Sw. Banerjee,[2] E. Ben-Haim,[3] E. Bertholet,[4] F. U. Bernlochner,[5] M. Bona,[6] A. Bozek,[7] C. Bozzi,[8] J. Brodzicka,[7] V. Chobanova,[9] M. Chrzaszcz,[7] S. Duell,[5] U. Egede,[10] M. Gersabeck,[11] T. Gershon,[12] P. Goldenzweig,[13] K. Hayasaka,[14] D. Johnson,[15] M. Kenzie,[12] T. Kuhr,[16] O. Leroy,[17] A. Lusiani,[18,19] H.-L. Ma,[20] M. Margoni,[21] K. Miyabayashi,[22] R. Mizuk,[4] P. Naik,[23] T. Nanut Petrič,[24] A. Oyanguren,[25] A. Pompili,[26,27] M. T. Prim,[5] M. Roney,[28] M. Rotondo,[29] O. Schneider,[30] C. Schwanda,[31] A. J. Schwartz,[32] J. Serrano,[17] A. Soffer,[4] D. Tonelli,[33] P. Urquijo,[34] and J. Yelton

(Heavy Flavor Averaging Group Collaboration)

global χ^2 statistic
x_i set of N independent measurements
V_i : their covariance matrix

$$\chi^2(\boldsymbol{x}) = \sum_i^N (\boldsymbol{x}_i - \boldsymbol{x})^{\mathrm{T}} \boldsymbol{V}_i^{-1} (\boldsymbol{x}_i - \boldsymbol{x})$$

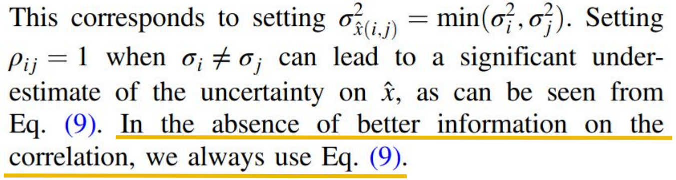


B. Treatment of unknown correlations

Another issue that needs careful treatment is that of unknown correlations among measurements, e.g., due to use of the same decay model for intermediate states to calculate acceptances. A common practice is to set the correlation coefficient to unity to indicate full correlation. However, this is not necessarily conservative and can result in an underestimated uncertainty on the average. The most conservative choice of correlation coefficient between two measurements i and j is that which maximizes the uncertainty on $\hat{x}$ due to the pair of measurements,

$$\sigma^2_{\hat{x}(i,j)} = \frac{\sigma_i^2\sigma_j^2(1-\rho_{ij}^2)}{\sigma_i^2+\sigma_j^2-2\rho_{ij}\sigma_i\sigma_j}, \qquad (9)$$

$$\rho_{ij} = \min\left(\frac{\sigma_i}{\sigma_j}, \frac{\sigma_j}{\sigma_i}\right).$$



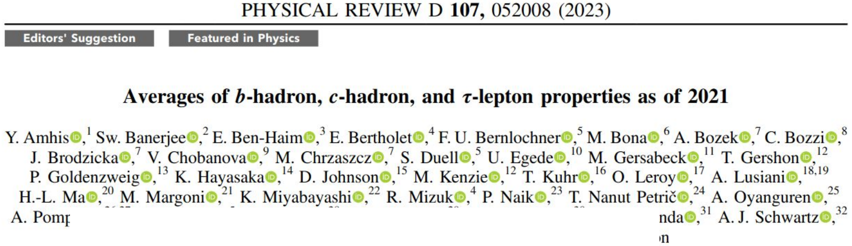
This corresponds to setting $\sigma^2_{\hat{x}(i,j)} = \min(\sigma_i^2, \sigma_j^2)$. Setting $\rho_{ij} = 1$ when $\sigma_i \neq \sigma_j$ can lead to a significant underestimate of the uncertainty on $\hat{x}$, as can be seen from Eq. (9). In the absence of better information on the correlation, we always use Eq. (9).

Requires **assumptions** and/or **approximations** regarding the **correlation between two systematic uncertainties**

R(D*) combinations by HFLAV

Our **optimal** method so far for:
Combining independent, but potentially **correlated** measurements

The averaging method
Is outlined in detail in:



PHYSICAL REVIEW D **107,** 052008 (2023)

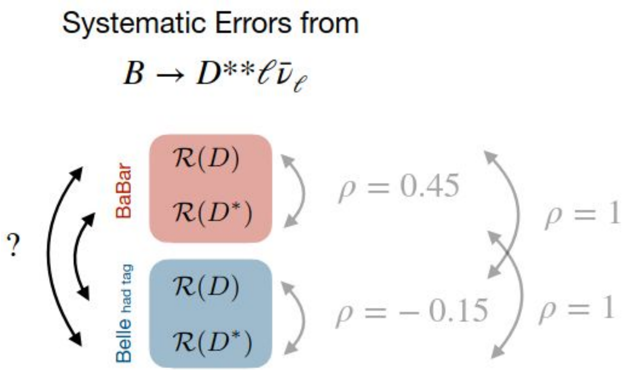
Editors' Suggestion
Featured in Physics

Averages of *b*-hadron, *c*-hadron, and *τ*-lepton properties as of 2021

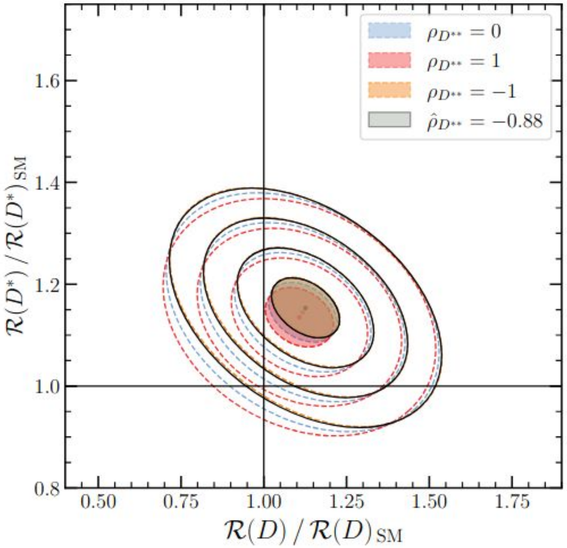
Y. Amhis,[1] Sw. Banerjee,[2] E. Ben-Haim,[3] E. Bertholet,[4] F. U. Bernlochner,[5] M. Bona,[6] A. Bozek,[7] C. Bozzi,[8]
J. Brodzicka,[7] V. Chobanova,[9] M. Chrzaszcz,[7] S. Duell,[5] U. Egede,[10] M. Gersabeck,[11] T. Gershon,[12]
P. Goldenzweig,[13] K. Hayasaka,[14] D. Johnson,[15] M. Kenzie,[12] T. Kuhr,[16] O. Leroy,[17] A. Lusiani,[18,19]
H.-L. Ma,[20] M. Margoni,[21] K. Miyabayashi,[22] R. Mizuk,[4] P. Naik,[23] T. Nanut Petrič,[24] A. Oyanguren,[25]
A. Pomp
nda,[31] A. J. Schwartz,[32]

global χ^2 statistic
x_i set of N independent measurements
V_i : their covariance matrix

$$\chi^2(\boldsymbol{x}) = \sum_i^N (\boldsymbol{x}_i - \boldsymbol{x})^{\mathrm{T}} \boldsymbol{V}_i^{-1} (\boldsymbol{x}_i - \boldsymbol{x})$$



F. Bernlochner at Challenges in SL B decays, 2022

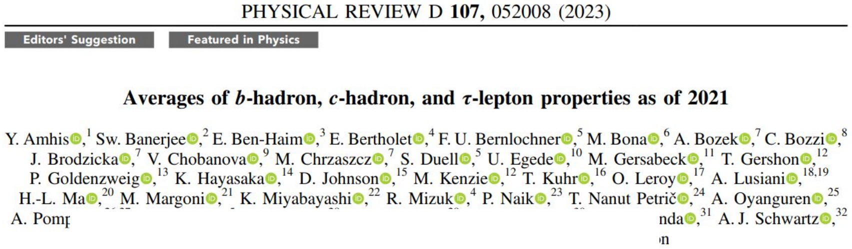


Requires **assumptions** and/or **approximations** regarding the **correlation between two systematic uncertainties**

R(D*) combinations by HFLAV

Our **optimal** method so far for:
Combining independent, but potentially **correlated** measurements

The averaging method
Is outlined in detail in:



PHYSICAL REVIEW D **107,** 052008 (2023)

Editors' Suggestion   Featured in Physics

Averages of *b*-hadron, *c*-hadron, and *τ*-lepton properties as of 2021

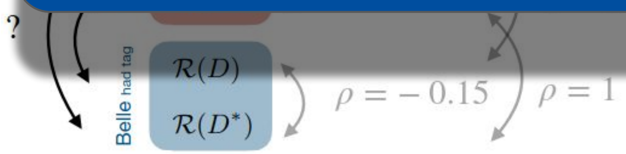
Y. Amhis,[1] Sw. Banerjee,[2] E. Ben-Haim,[3] E. Bertholet,[4] F. U. Bernlochner,[5] M. Bona,[6] A. Bozek,[7] C. Bozzi,[8]
J. Brodzicka,[7] V. Chobanova,[9] M. Chrzaszcz,[7] S. Duell,[5] U. Egede,[10] M. Gersabeck,[11] T. Gershon,[12]
P. Goldenzweig,[13] K. Hayasaka,[14] D. Johnson,[15] M. Kenzie,[12] T. Kuhr,[16] O. Leroy,[17] A. Lusiani,[18,19]
H.-L. Ma,[20] M. Margoni,[21] K. Miyabayashi,[22] R. Mizuk,[4] P. Naik,[23] T. Nanut Petrič,[24] A. Oyanguren,[25]
A. Pomp nda,[31] A. J. Schwartz,[32]
n

global χ^2 statistic
x_i set of N independent measurements
V_i : their covariance matrix

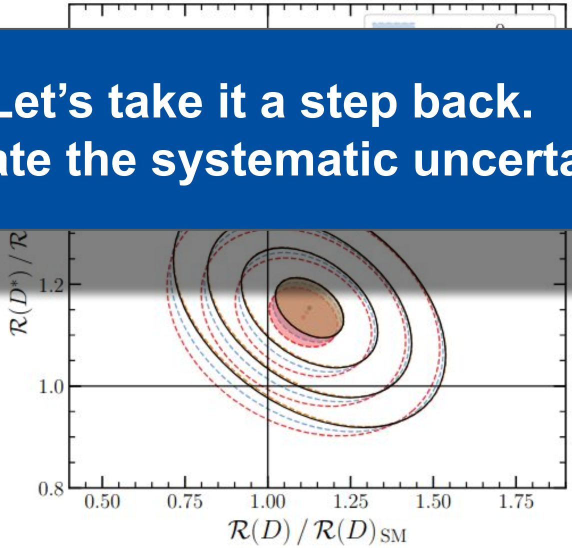
$$\chi^2(x) = \sum^{N} (x_i - x)^T V_i^{-1} (x_i - x)$$

Systematic Errors from

Let's take it a step back.
How do we even calculate the systematic uncertainty in one analysis?

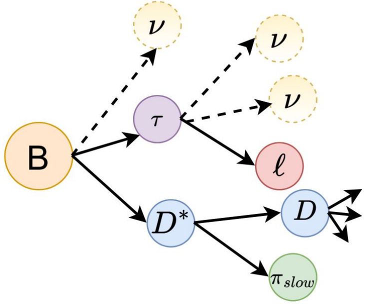


F. Bernlochner at Challenges in SL B decays, 2022



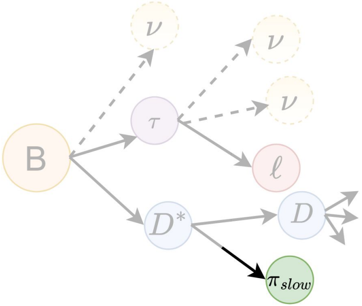
and/or **approximations** regarding the **correlation between two systematic uncertainties**

An example of π_{slow} efficiency systematics

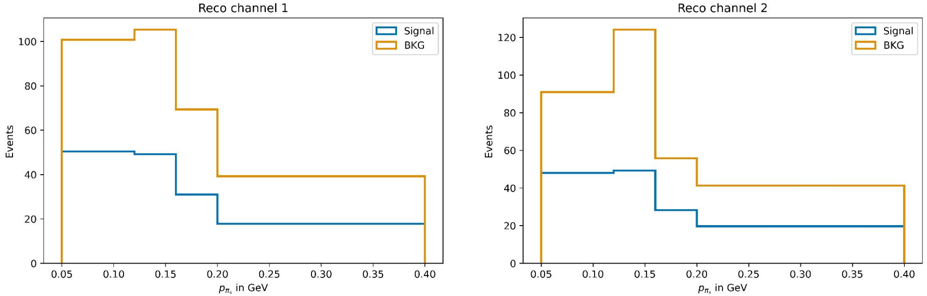


The track finding efficiency of π_{slow} typically differs on Data and MC.
We correct those in bins of the momentum of π_{slow}

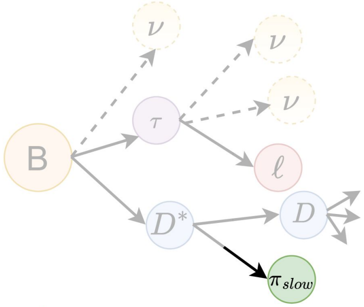
An example of π_{slow} efficiency systematics

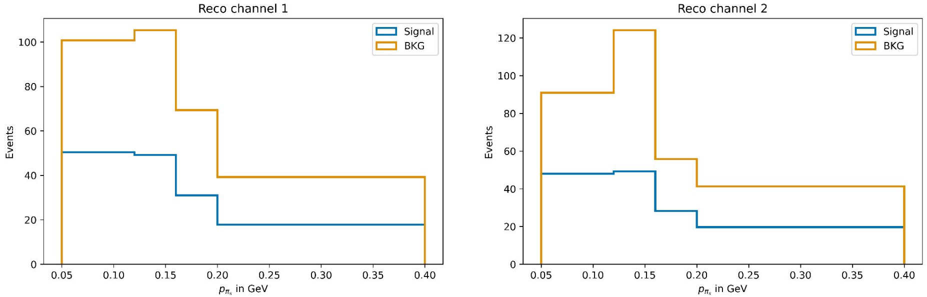


Apply $\frac{\epsilon^i_{\text{Data}}}{\epsilon^i_{\text{MC}}}$ factors as correction weights on MC with i the π_{slow} momentum bin

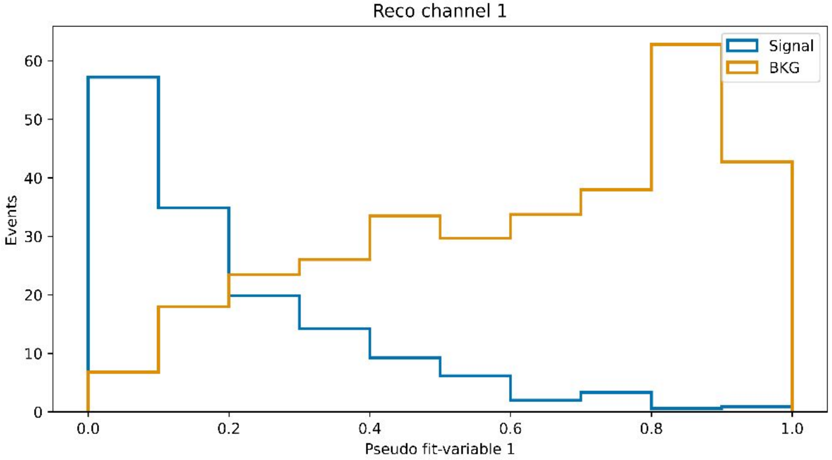


An example of π_{slow} efficiency systematics



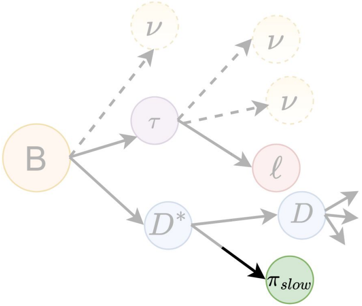


Apply $\frac{\epsilon^i_{\text{Data}}}{\epsilon^i_{\text{MC}}}$ factors as correction weights on MC with i the π_{slow} momentum bin

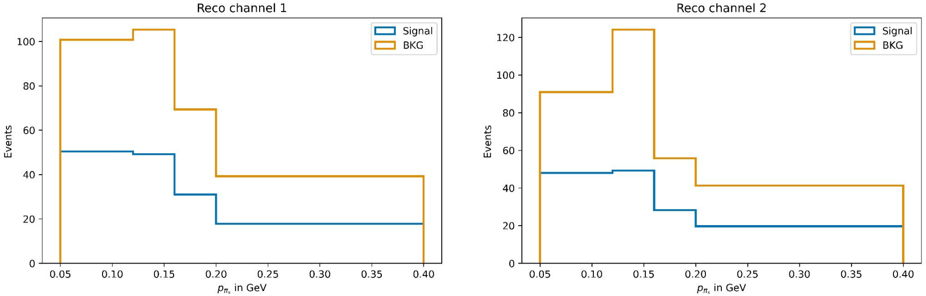


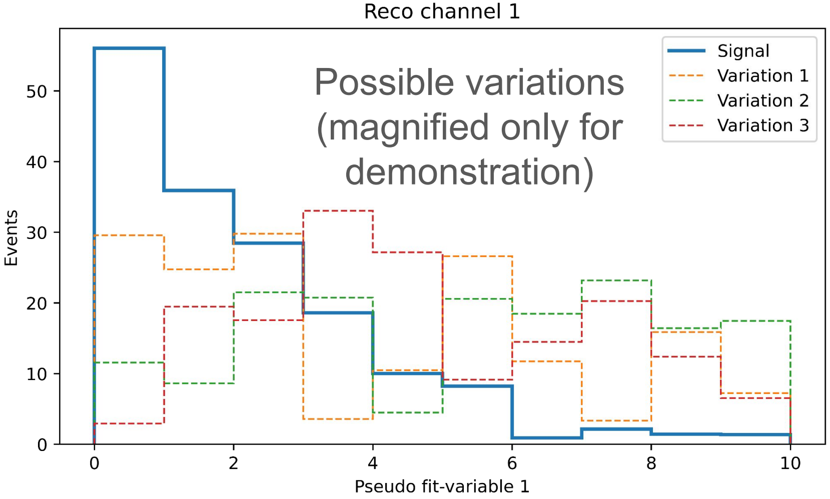
How do these correction weights affect the signal extraction variable ?

An example of π_{slow} efficiency systematics



Apply $\frac{\epsilon^i_{\text{Data}}}{\epsilon^i_{\text{MC}}}$ factors as correction weights on MC with i the π_{slow} momentum bin

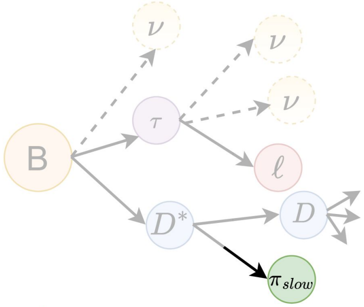


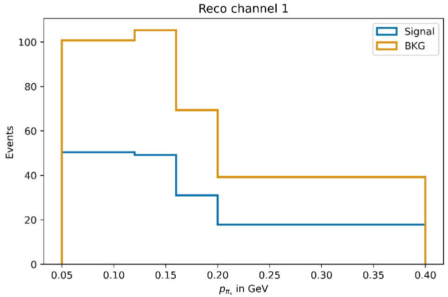


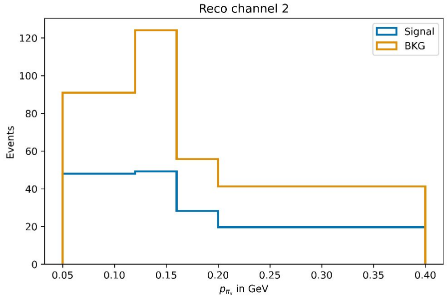
Such corrections can have two effects on the signal extraction variable

1. Overall normalization (trivial to calculate)
2. Shape effect
   (non trivial especially in simultaneous fits)

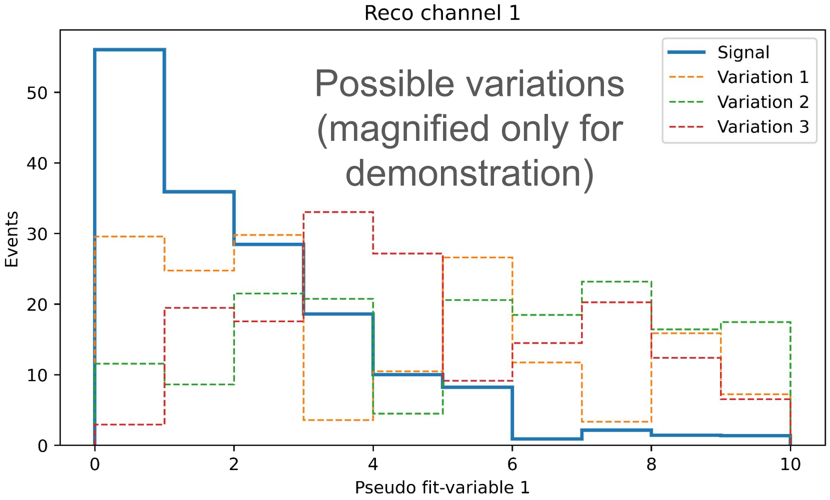
An example of π_{slow} efficiency systematics







Apply $\frac{\epsilon^i_{\text{Data}}}{\epsilon^i_{\text{MC}}}$ factors as correction weights on MC with i the π_{slow} momentum bin



Such corrections can have two effects on the signal extraction variable

1. Overall normalization (trivial to calculate)
2. Shape effect
   (non trivial especially in simultaneous fits)

In case no arbitrary correlation matrices can be used in the fitter to **properly correlate the different bins** (like in pyhf) one has to

1. Build a **covariance matrix** between all the bins
   Dimensions: (channels x templates x bins)
2. **Diagonalize** it
3. Calculate the **eigenvariations** of those.

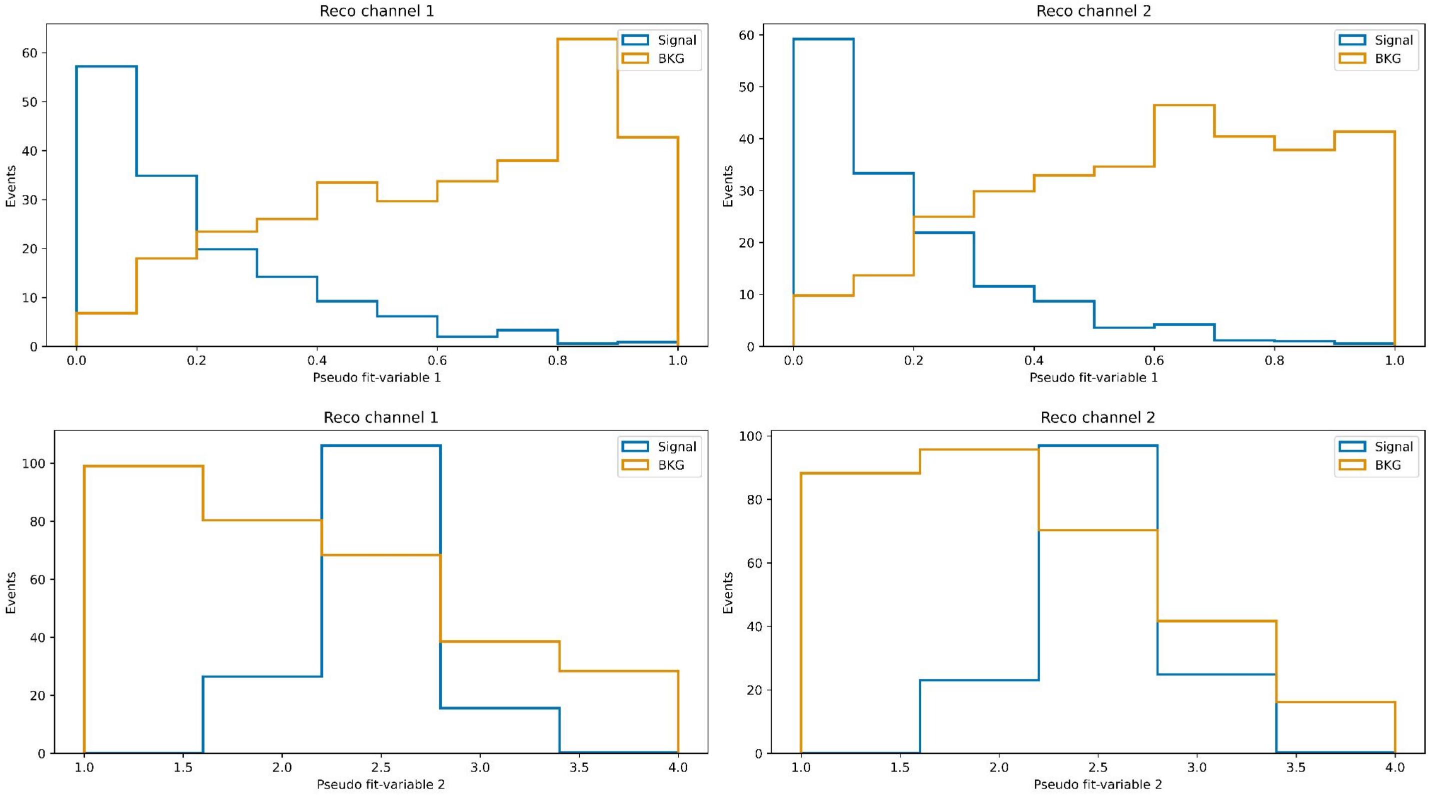
Every eigendirection is then implemented as **a fully correlated nuisance parameter across all bins.**

SysVar

A New Tool for Enhancing Consistency in the Treatment of Systematic Uncertainties

User interface example

For demonstration let’s consider a 2D simultaneous fit in two reconstructions channels

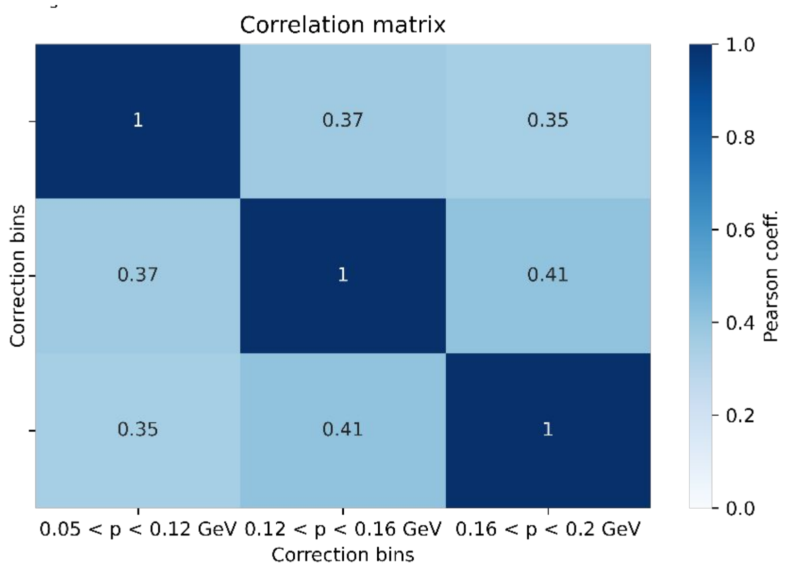


Slow π efficiency systematics need to be considered

SysVar: A tool enhancing consistency in the treatment of systematics

SysVar is a python based tool that allows to:

1. Apply Data/MC corrections to a DataFrame



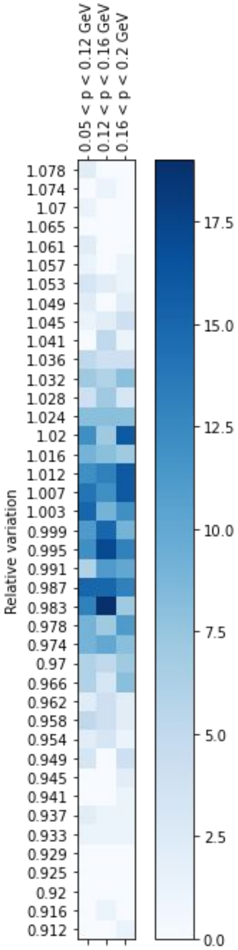
SysVar: A tool enhancing consistency in the treatment of systematics

SysVar is a python based tool that allows to:

1. Apply Data/MC corrections to a DataFrame
2. Generate Variations of Data/MC Corrections

200 variations

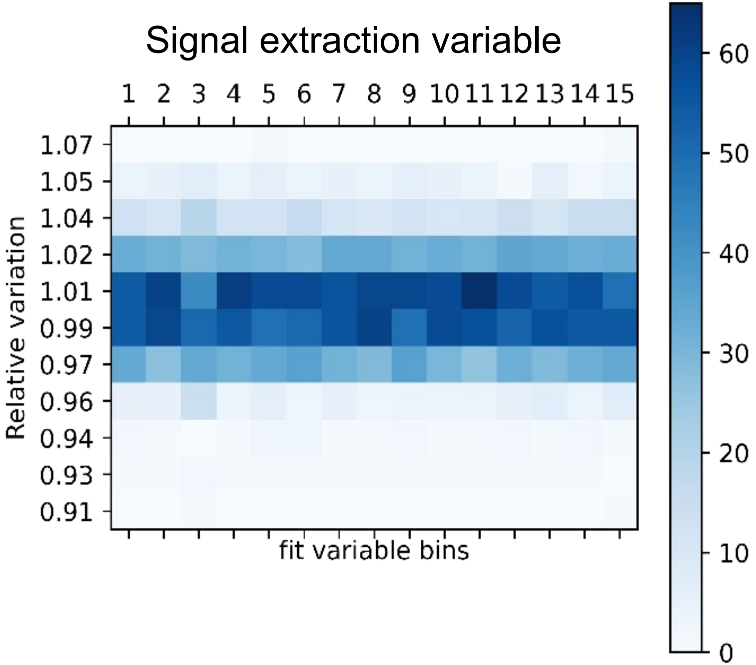
Sampled from a multidimensional gaussian



SysVar: A tool enhancing consistency in the treatment of systematics

SysVar is a python based tool that allows to:

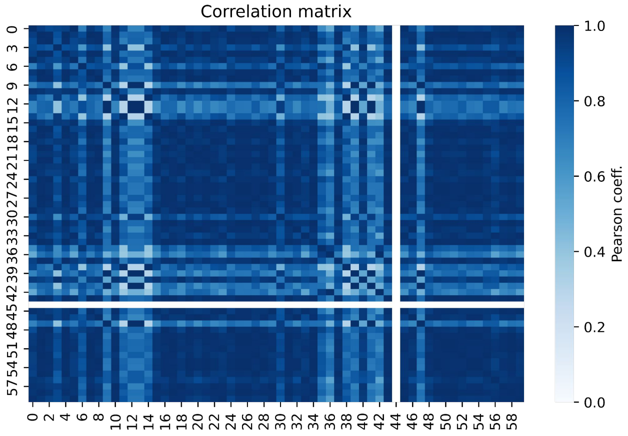
1. Apply Data/MC corrections to a DataFrame
2. Generate Variations of Data/MC Corrections
3. Histogram MC and Data and build templates and template variations for a non-parametric fit.



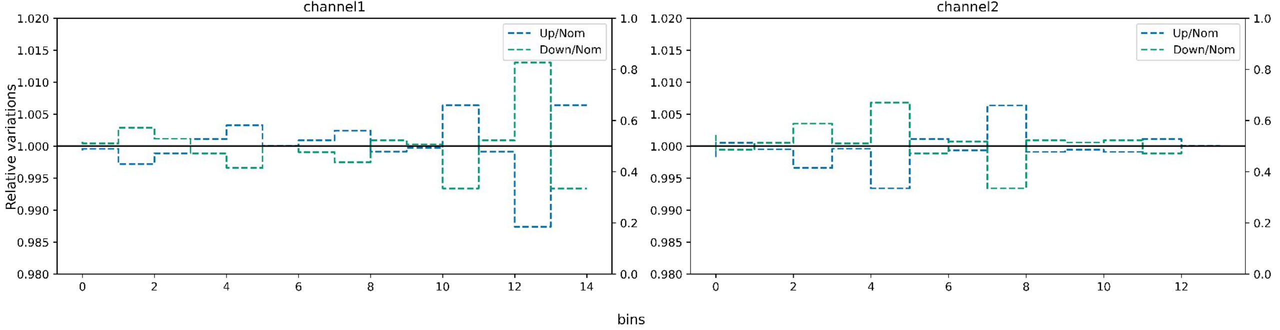
SysVar: A tool enhancing consistency in the treatment of systematics

SysVar is a python based tool that allows to:

1. Apply Data/MC corrections to a DataFrame
2. Generate Variations of Data/MC Corrections
3. Histogram MC and Data and build templates and template variations for a non-parametric fit.
4. Diagonalize a cov matrix (channels x templates x bins) to produce Eigenvariations and save them to implement Correlated Shape Variations



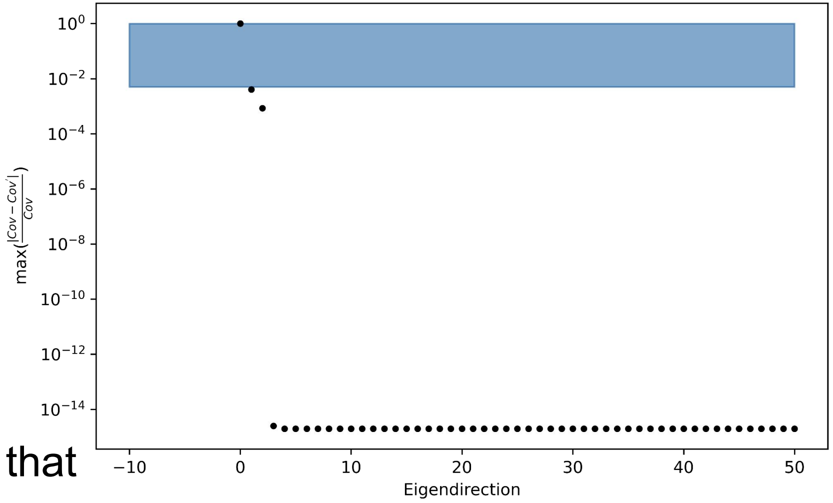
Charged slow pion efficiency var1 for: signal



SysVar: A tool enhancing consistency in the treatment of systematics

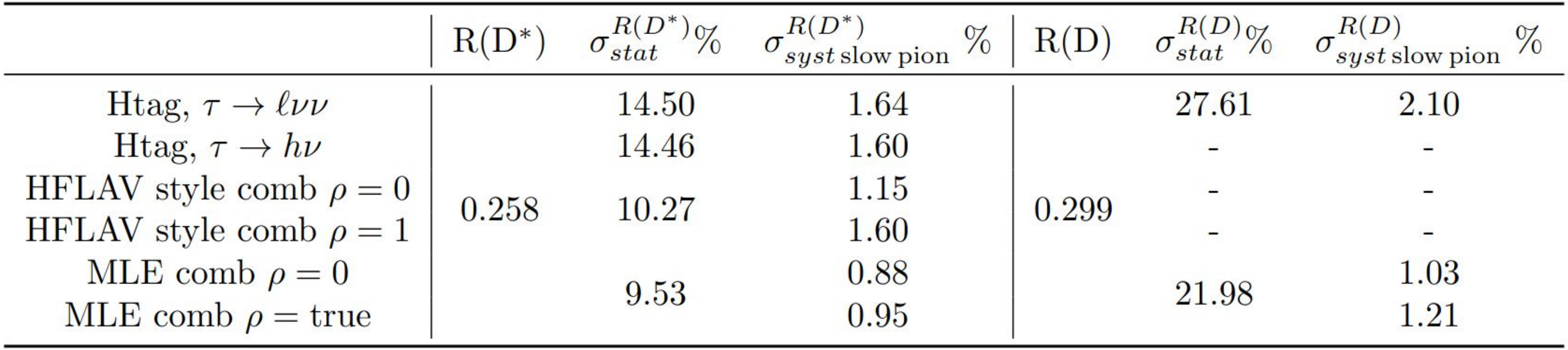
SysVar is a python based tool that allows to:

1. Apply Data/MC corrections to a DataFrame
2. Generate Variations of Data/MC Corrections
3. Histogram MC and Data and build templates and template variations for a non-parametric fit.
4. Diagonalize a cov matrix (channels x templates x bins) to produce Eigenvariations and save them to implement Correlated Shape Variations
5. Identify the number of necessary Eigendirections that the user should to consider for an accurate analysis (multiple criteria are available)



Let’s combine some $R(D^{(*)})$ measurements

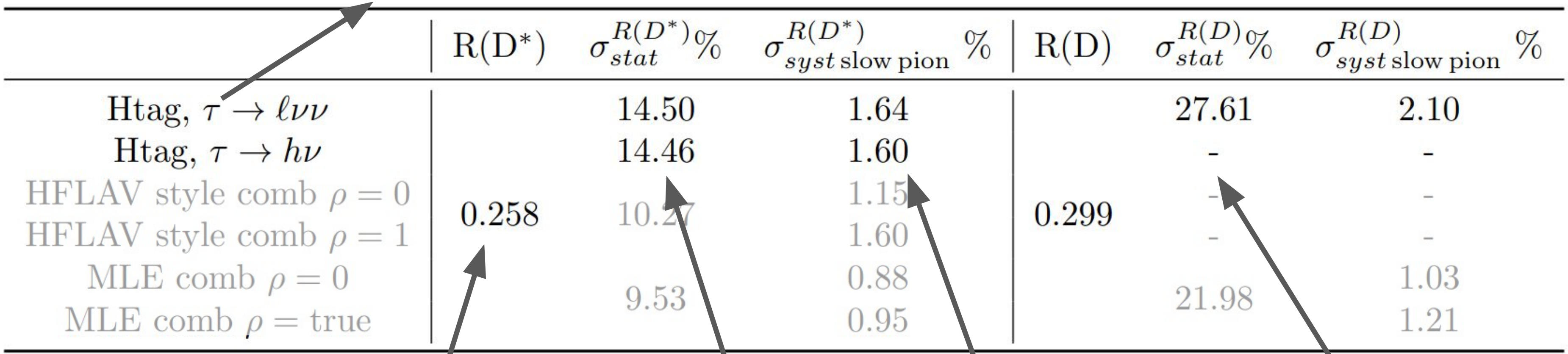
Comparison of combination methods



	R(D*)	$\sigma^{R(D^*)}_{stat}$ %	$\sigma^{R(D^*)}_{syst\,\text{slow pion}}$ %	R(D)	$\sigma^{R(D)}_{stat}$ %	$\sigma^{R(D)}_{syst\,\text{slow pion}}$ %
Htag, $\tau \to \ell\nu\nu$	0.258	14.50	1.64	0.299	27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho = \text{true}$			0.95			1.21

Comparison of combination methods

This is an independent and preliminary cross-check effort
to the published result with completely different reconstruction and signal extraction strategies



	R(D*)	$\sigma^{R(D^*)}_{stat}$ %	$\sigma^{R(D^*)}_{syst\,\text{slow pion}}$ %	R(D)	$\sigma^{R(D)}_{stat}$ %	$\sigma^{R(D)}_{syst\,\text{slow pion}}$ %
Htag, $\tau \to \ell\nu\nu$	0.258	14.50	1.64	0.299	27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho =$ true			0.95			1.21

Two independent measurements

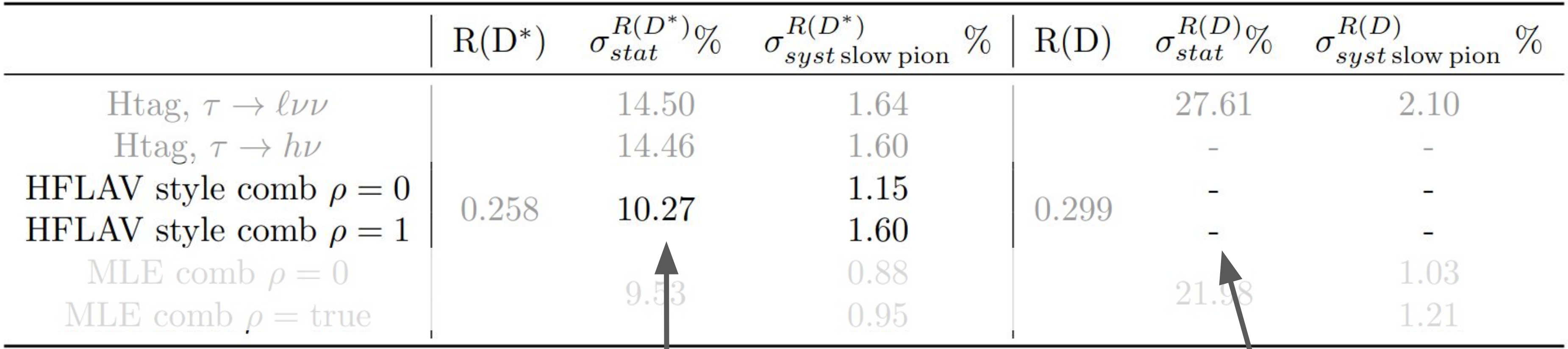
Preliminary Asimov fits

Similar statistical precision (assuming 364 fb^{-1})

Similar effect of slow π systematics

had τ does not measure R(D)

Comparison of combination methods



	R(D*)	$\sigma_{stat}^{R(D^*)}$ %	$\sigma_{syst\,\text{slow pion}}^{R(D^*)}$ %	R(D)	$\sigma_{stat}^{R(D)}$ %	$\sigma_{syst\,\text{slow pion}}^{R(D)}$ %
Htag, $\tau \to \ell\nu\nu$	0.258	14.50	1.64	0.299	27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho =$ true			0.95			1.21

Calculate the uncertainty HFLAV style

$$\sigma^2_{\hat{x}(i,j)} = \frac{\sigma_i^2 \sigma_j^2 (1 - \rho_{ij}^2)}{\sigma_i^2 + \sigma_j^2 - 2\rho_{ij}\sigma_i\sigma_j}$$

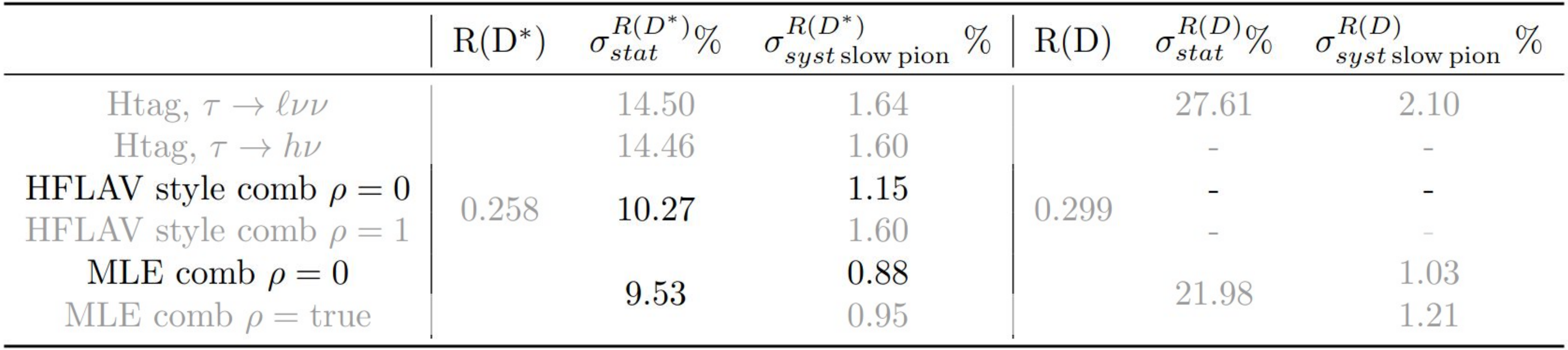
$$\rho_{ij} = \min\left(\frac{\sigma_i}{\sigma_j}, \frac{\sigma_j}{\sigma_i}\right)$$

Increased statistical precision

Not a full HFLAV style global χ^2 fit here

No average of R(D)

Comparison of combination methods

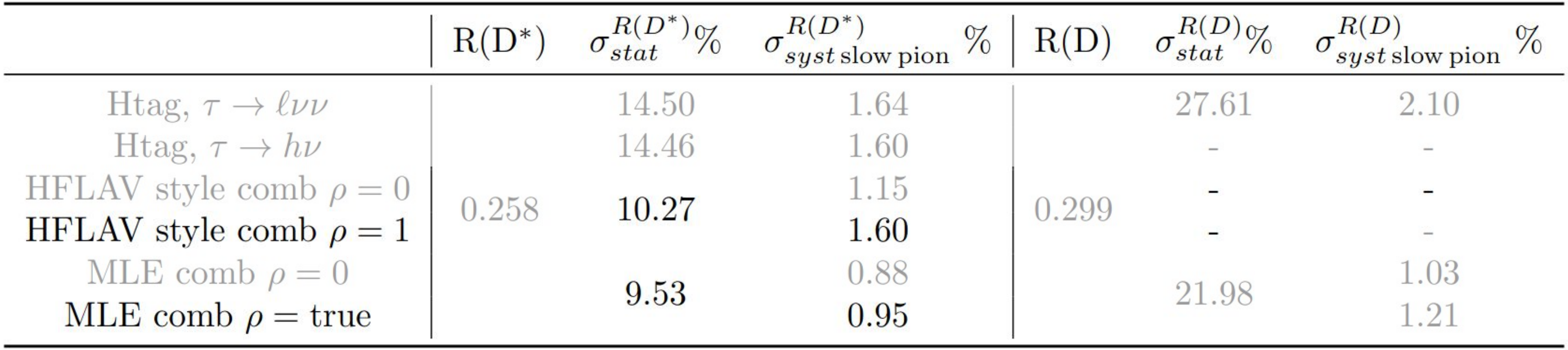


	R(D*)	$\sigma_{stat}^{R(D^*)}$%	$\sigma_{syst\,\text{slow pion}}^{R(D^*)}$ %	R(D)	$\sigma_{stat}^{R(D)}$%	$\sigma_{syst\,\text{slow pion}}^{R(D)}$ %
Htag, $\tau \to \ell\nu\nu$	0.258	14.50	1.64	0.299	27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho =$ true			0.95			1.21

The simultaneous MLE reduces the combined systematic uncertainty

For a simultaneous MLE we have a smoother likelihood, avoiding extreme fluctuations in the NP

Comparison of combination methods



	R(D*)	$\sigma_{stat}^{R(D^*)}$ %	$\sigma_{syst\,\text{slow pion}}^{R(D^*)}$ %	R(D)	$\sigma_{stat}^{R(D)}$ %	$\sigma_{syst\,\text{slow pion}}^{R(D)}$ %
Htag, $\tau \to \ell\nu\nu$		14.50	1.64		27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$	0.258	10.27	1.15	0.299	-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho =$ true			0.95			1.21

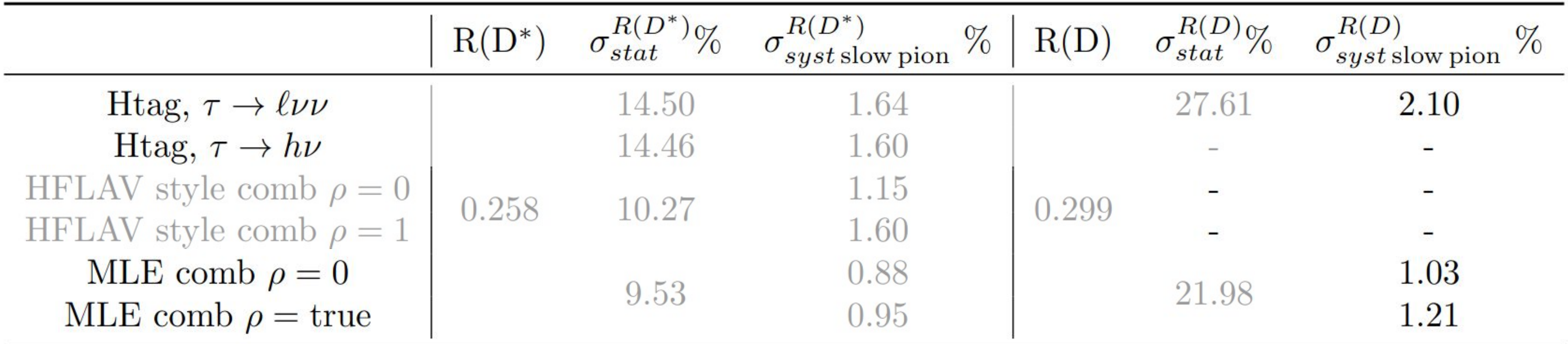
Full correlation in HFLAV style is a conservative approach

$$\rho_{ij} = \min\left(\frac{\sigma_i}{\sigma_j}, \frac{\sigma_j}{\sigma_i}\right)$$

NP don't float independently leading to a smaller reduction of the systematic uncertainty

The combined uncertainty from MLE takes into account both normalization and **shapes**

Comparison of combination methods

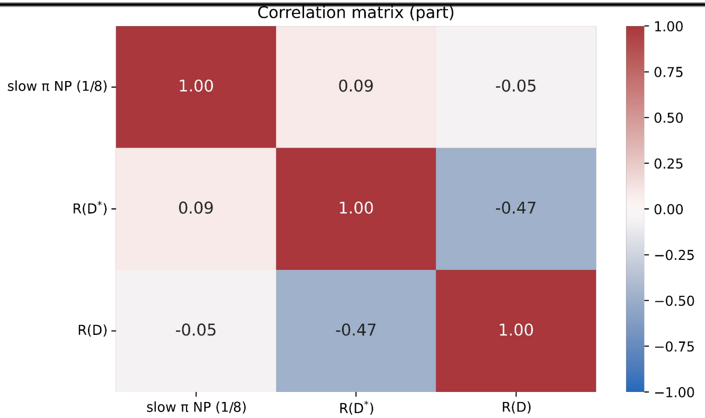


	R(D*)	$\sigma_{stat}^{R(D^*)}$ %	$\sigma_{syst\,\mathrm{slow\,pion}}^{R(D^*)}$ %	R(D)	$\sigma_{stat}^{R(D)}$ %	$\sigma_{syst\,\mathrm{slow\,pion}}^{R(D)}$ %
Htag, $\tau \to \ell\nu\nu$	0.258	14.50	1.64	0.299	27.61	2.10
Htag, $\tau \to h\nu$		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho = \mathrm{true}$			0.95			1.21

No R(D) is determined from had τ

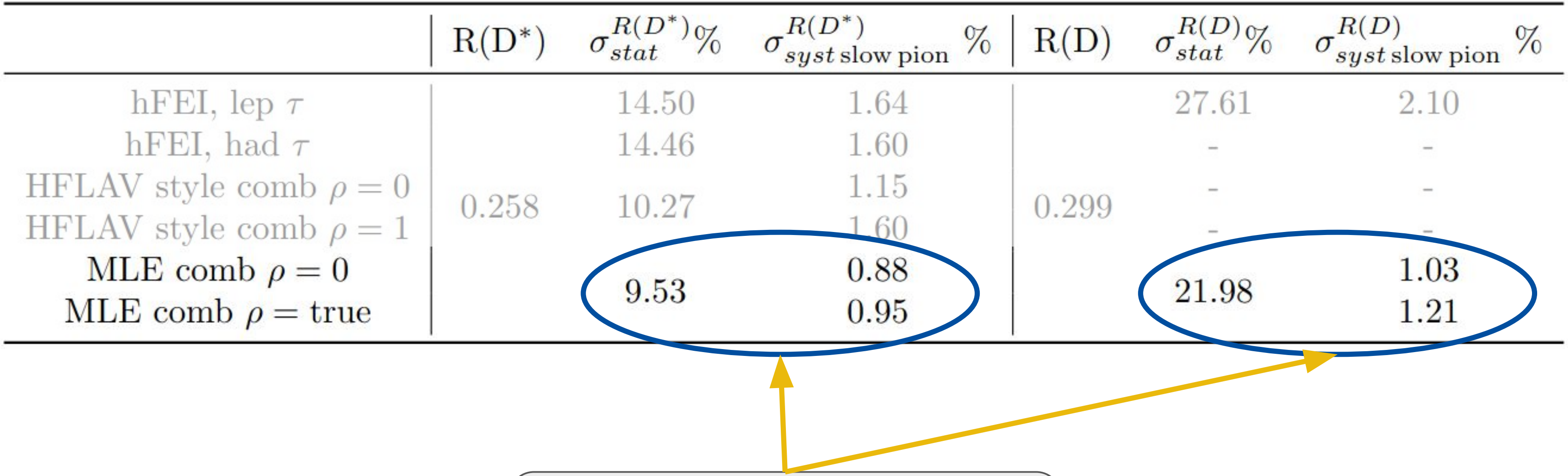
Shared NP constrain the uncertainties more tightly

Correlation matrix (part)



	slow π NP (1/8)	R(D*)	R(D)
slow π NP (1/8)	1.00	0.09	-0.05
R(D*)	0.09	1.00	-0.47
R(D)	-0.05	-0.47	1.00

Comparison of combination methods

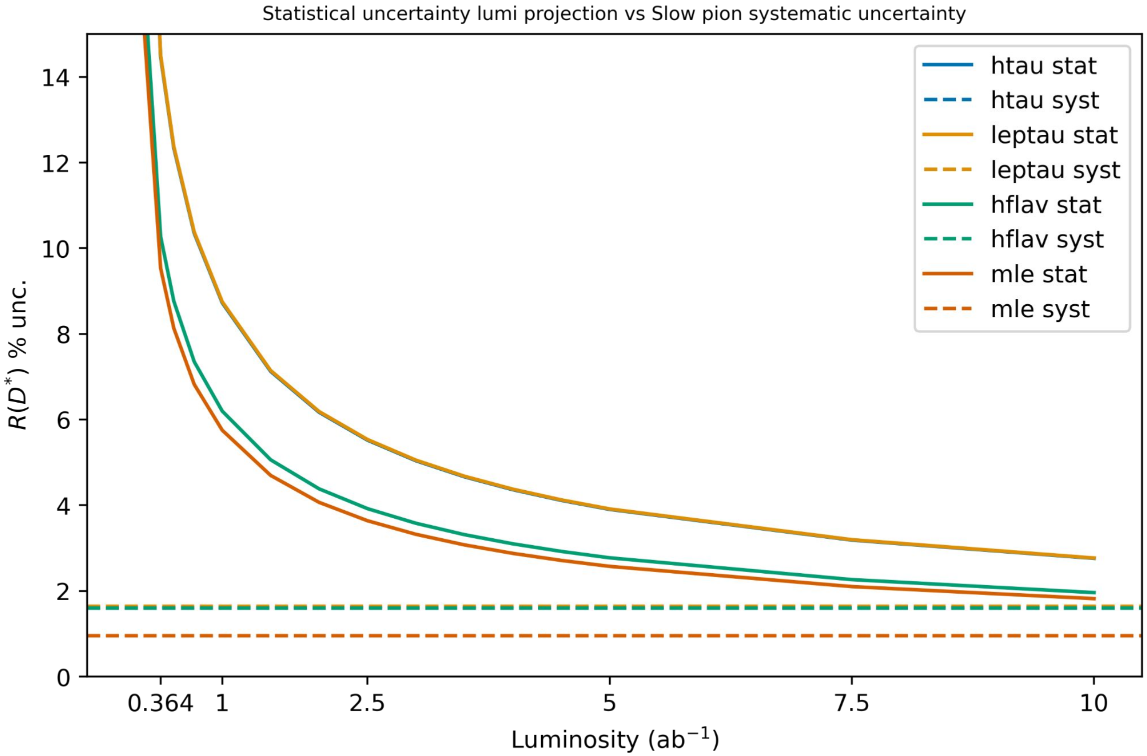


	R(D*)	$\sigma_{stat}^{R(D^*)}$%	$\sigma_{syst\,\text{slow pion}}^{R(D^*)}$ %	R(D)	$\sigma_{stat}^{R(D)}$%	$\sigma_{syst\,\text{slow pion}}^{R(D)}$ %
hFEI, lep τ	0.258	14.50	1.64	0.299	27.61	2.10
hFEI, had τ		14.46	1.60		-	-
HFLAV style comb $\rho = 0$		10.27	1.15		-	-
HFLAV style comb $\rho = 1$			1.60		-	-
MLE comb $\rho = 0$		9.53	0.88		21.98	1.03
MLE comb $\rho =$ true			0.95			1.21

stat error >> syst error

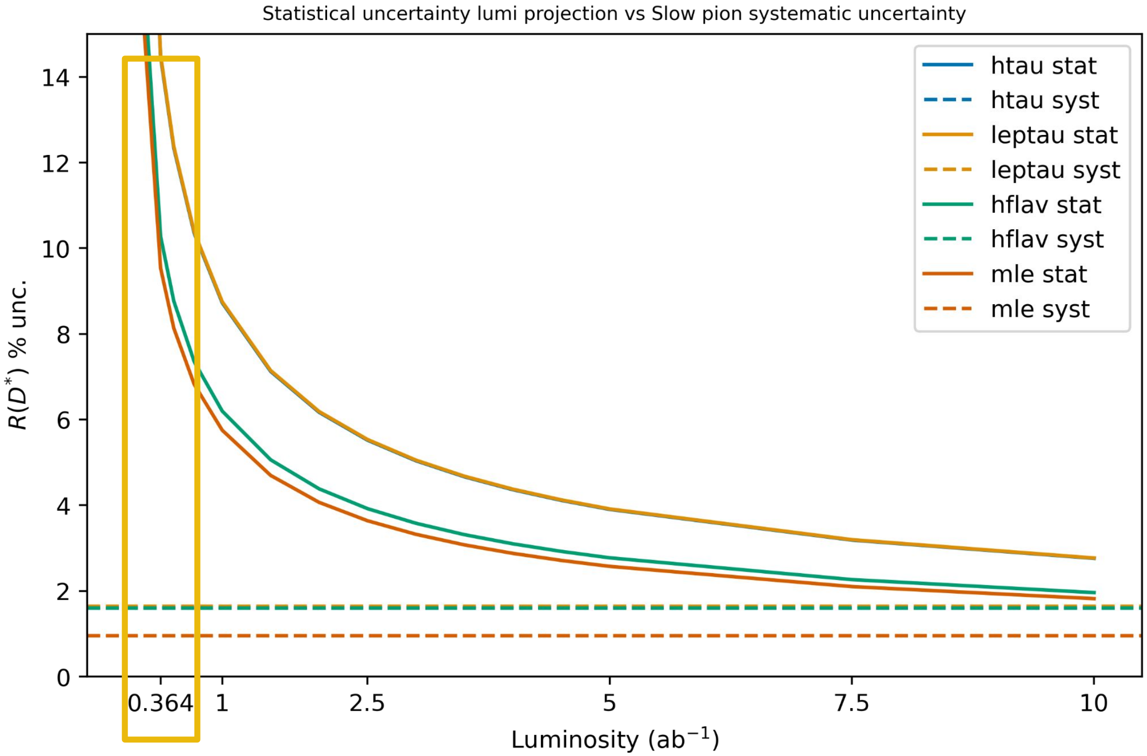
Why even bother ?

Statistical vs Systematic uncertainty



$$\sigma_{\mathrm{R(D^*)}} = \sigma^{\mathrm{LS1}}_{\mathrm{R(D^*)}} \cdot \sqrt{\frac{\mathcal{L}_{\mathrm{LS1}}}{\mathcal{L}}}$$

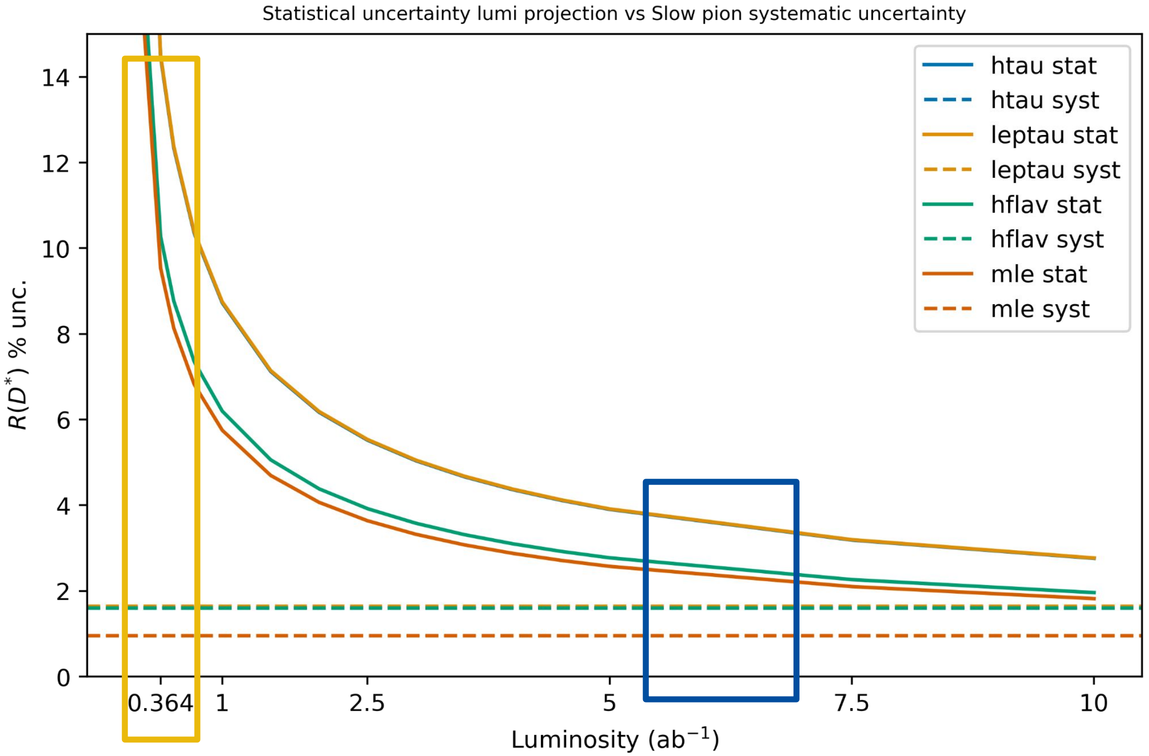
Statistical vs Systematic uncertainty



Now
we're here

$$\sigma_{\mathrm{R(D^*)}} = \sigma^{\mathrm{LS1}}_{\mathrm{R(D^*)}} \cdot \sqrt{\frac{\mathcal{L}_{\mathrm{LS1}}}{\mathcal{L}}}$$

Statistical vs Systematic uncertainty

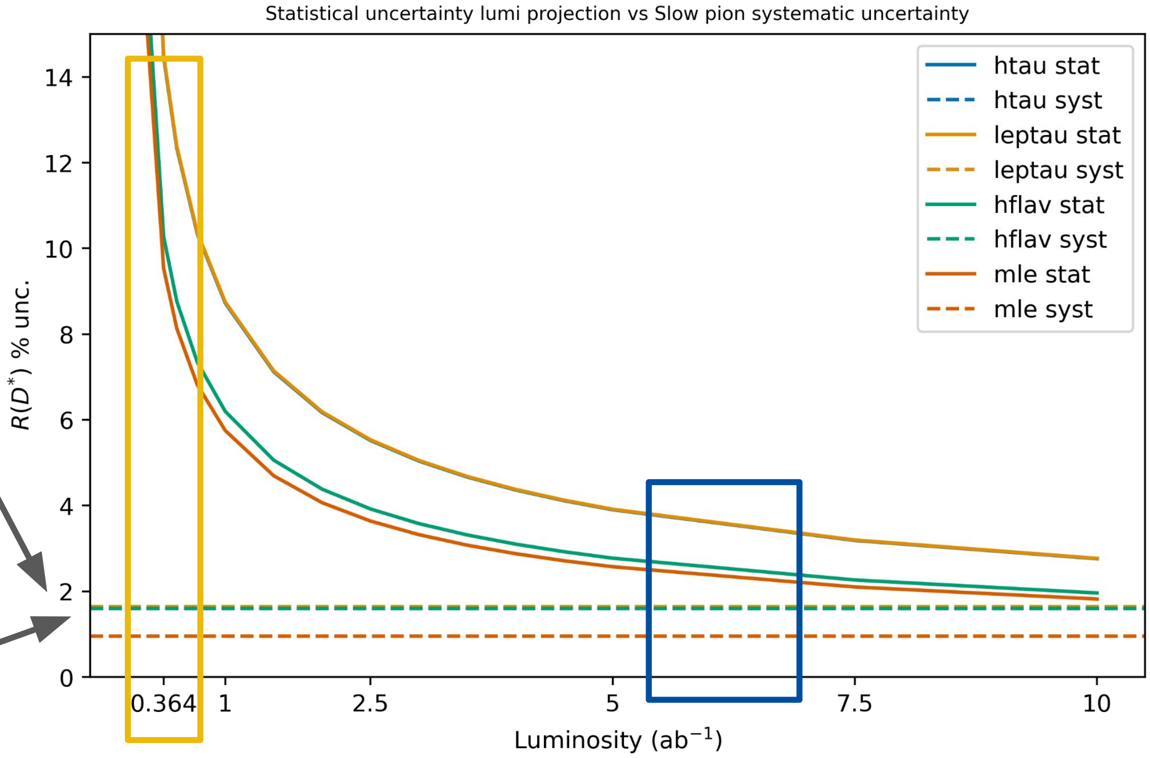


$$\sigma_{\mathrm{R(D^*)}} = \sigma^{\mathrm{LS1}}_{\mathrm{R(D^*)}} \cdot \sqrt{\frac{\mathcal{L}_{\mathrm{LS1}}}{\mathcal{L}}}$$

Now we're here

In a couple of years we'll be there

Statistical vs Systematic uncertainty



$$\sigma_{\mathrm{R(D^*)}} = \sigma^{\mathrm{LS1}}_{\mathrm{R(D^*)}} \cdot \sqrt{\frac{\mathcal{L}_{\mathrm{LS1}}}{\mathcal{L}}}$$

not the total systematic error

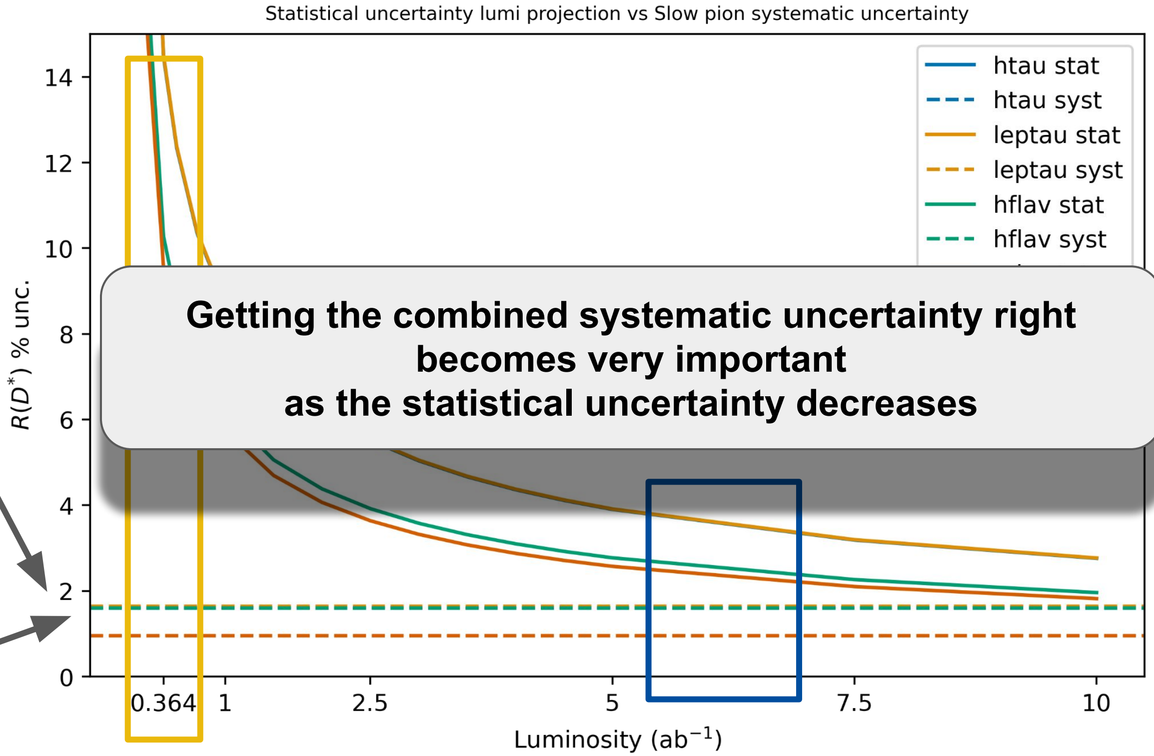
not the leading systematic uncertainty

Now we're here

In a couple of years we'll be there

Need to add a few of those in quadrature

Statistical vs Systematic uncertainty



Getting the combined systematic uncertainty right becomes very important as the statistical uncertainty decreases

not the total systematic error

not the leading systematic uncertainty

Now we're here

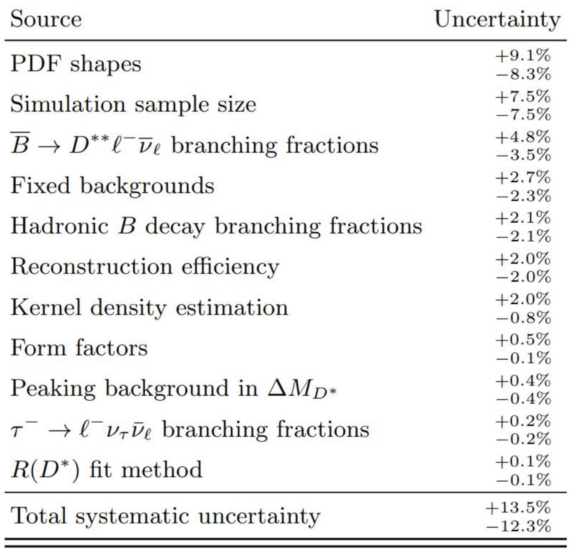
In a couple of years we'll be there

$$\sigma_{\mathrm{R(D^*)}} = \sigma^{\mathrm{LS1}}_{\mathrm{R(D^*)}} \cdot \sqrt{\frac{\mathcal{L}_{\mathrm{LS1}}}{\mathcal{L}}}$$

Need to add a few of those in quadrature

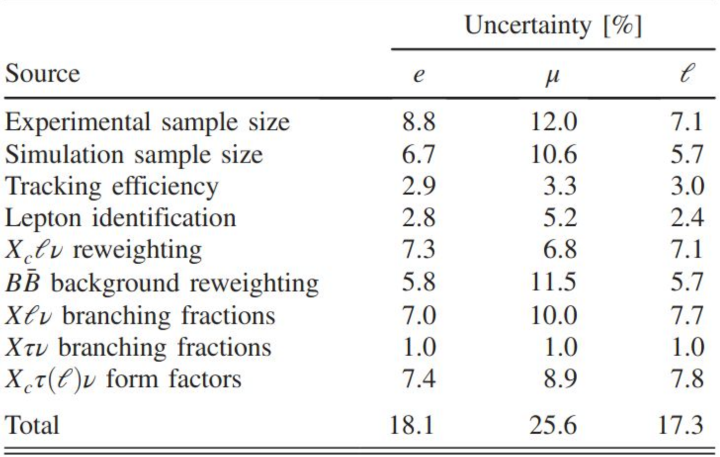
Full systematic budgets of R(D*) measurements

Belle II R(D*) with hFEI 189 fb^{-1}



Source	Uncertainty
PDF shapes	$^{+9.1\%}_{-8.3\%}$
Simulation sample size	$^{+7.5\%}_{-7.5\%}$
$\overline{B} \to D^{**}\ell^- \overline{\nu}_\ell$ branching fractions	$^{+4.8\%}_{-3.5\%}$
Fixed backgrounds	$^{+2.7\%}_{-2.3\%}$
Hadronic B decay branching fractions	$^{+2.1\%}_{-2.1\%}$
Reconstruction efficiency	$^{+2.0\%}_{-2.0\%}$
Kernel density estimation	$^{+2.0\%}_{-0.8\%}$
Form factors	$^{+0.5\%}_{-0.1\%}$
Peaking background in ΔM_{D^*}	$^{+0.4\%}_{-0.4\%}$
$\tau^- \to \ell^- \nu_\tau \overline{\nu}_\ell$ branching fractions	$^{+0.2\%}_{-0.2\%}$
$R(D^*)$ fit method	$^{+0.1\%}_{-0.1\%}$
Total systematic uncertainty	$^{+13.5\%}_{-12.3\%}$

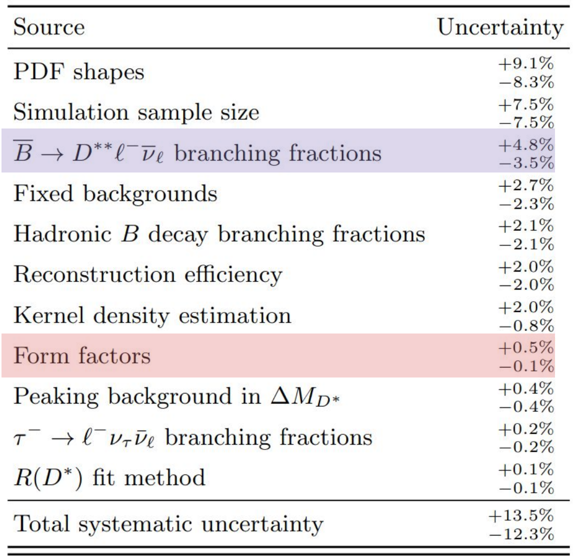
Belle II R(X) with hFEI 189 fb^{-1}



	Uncertainty [%]		
Source	e	μ	ℓ
Experimental sample size	8.8	12.0	7.1
Simulation sample size	6.7	10.6	5.7
Tracking efficiency	2.9	3.3	3.0
Lepton identification	2.8	5.2	2.4
$X_c\ell\nu$ reweighting	7.3	6.8	7.1
$B\bar{B}$ background reweighting	5.8	11.5	5.7
$X\ell\nu$ branching fractions	7.0	10.0	7.7
$X\tau\nu$ branching fractions	1.0	1.0	1.0
$X_c\tau(\ell)\nu$ form factors	7.4	8.9	7.8
Total	18.1	25.6	17.3

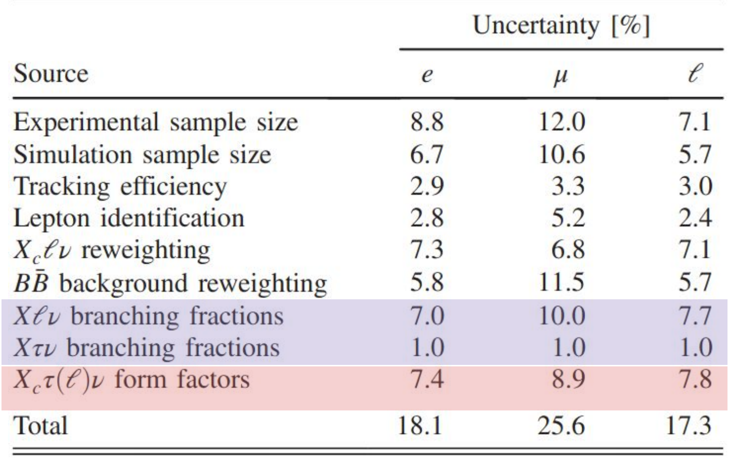
Fully correlated systematic uncertainties

Belle II R(D*) with hFEI 189 fb^{-1}

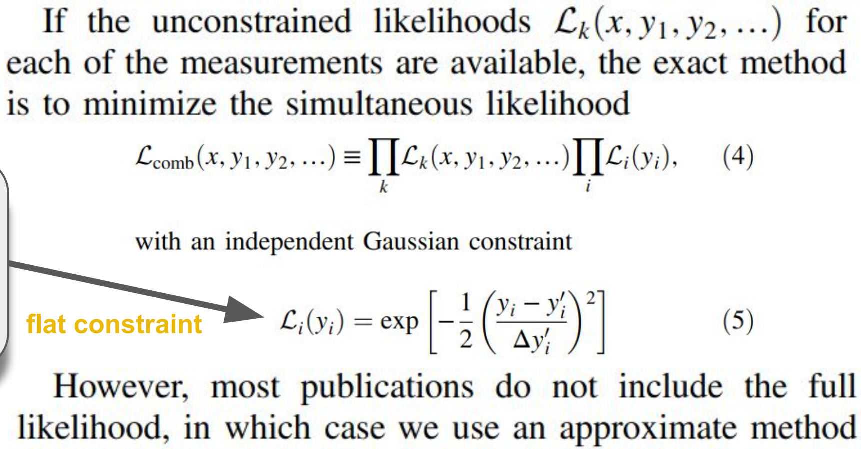


Source	Uncertainty
PDF shapes	$^{+9.1\%}_{-8.3\%}$
Simulation sample size	$^{+7.5\%}_{-7.5\%}$
$\overline{B} \to D^{**}\ell^-\overline{\nu}_\ell$ branching fractions	$^{+4.8\%}_{-3.5\%}$
Fixed backgrounds	$^{+2.7\%}_{-2.3\%}$
Hadronic B decay branching fractions	$^{+2.1\%}_{-2.1\%}$
Reconstruction efficiency	$^{+2.0\%}_{-2.0\%}$
Kernel density estimation	$^{+2.0\%}_{-0.8\%}$
Form factors	$^{+0.5\%}_{-0.1\%}$
Peaking background in ΔM_{D^*}	$^{+0.4\%}_{-0.4\%}$
$\tau^- \to \ell^-\nu_\tau\overline{\nu}_\ell$ branching fractions	$^{+0.2\%}_{-0.2\%}$
$R(D^*)$ fit method	$^{+0.1\%}_{-0.1\%}$
Total systematic uncertainty	$^{+13.5\%}_{-12.3\%}$

Belle II R(X) with hFEI 189 fb^{-1}



	Uncertainty [%]		
Source	e	μ	ℓ
Experimental sample size	8.8	12.0	7.1
Simulation sample size	6.7	10.6	5.7
Tracking efficiency	2.9	3.3	3.0
Lepton identification	2.8	5.2	2.4
$X_c\ell\nu$ reweighting	7.3	6.8	7.1
$B\bar{B}$ background reweighting	5.8	11.5	5.7
$X\ell\nu$ branching fractions	7.0	10.0	7.7
$X\tau\nu$ branching fractions	1.0	1.0	1.0
$X_c\tau(\ell)\nu$ form factors	7.4	8.9	7.8
Total	18.1	25.6	17.3



If the unconstrained likelihoods $\mathcal{L}_k(x, y_1, y_2, \ldots)$ for each of the measurements are available, the exact method is to minimize the simultaneous likelihood

$$\mathcal{L}_{\text{comb}}(x, y_1, y_2, \ldots) \equiv \prod_k \mathcal{L}_k(x, y_1, y_2, \ldots) \prod_i \mathcal{L}_i(y_i), \quad (4)$$

with an independent Gaussian constraint

$$\mathcal{L}_i(y_i) = \exp\left[-\frac{1}{2}\left(\frac{y_i - y_i'}{\Delta y_i'}\right)^2\right] \quad (5)$$

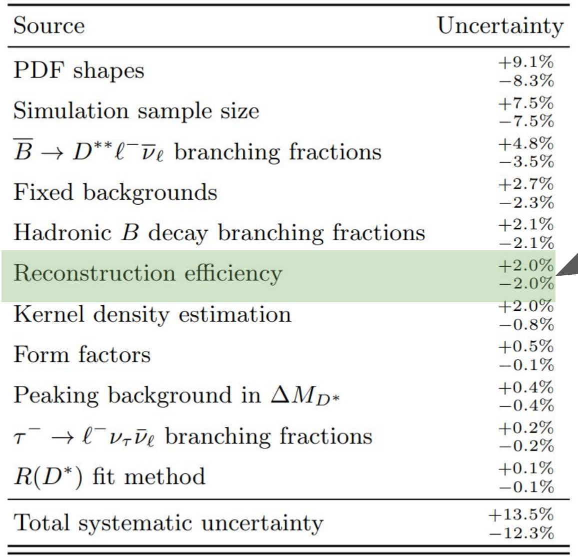
However, most publications do not include the full likelihood, in which case we use an approximate method

flat constraint

The **HFLAV** approximate method is valid in most cases but **fails to profit** from any information encoded in the **shape effects** of the **correlated systematics**

Arbitrarily correlated systematic uncertainties

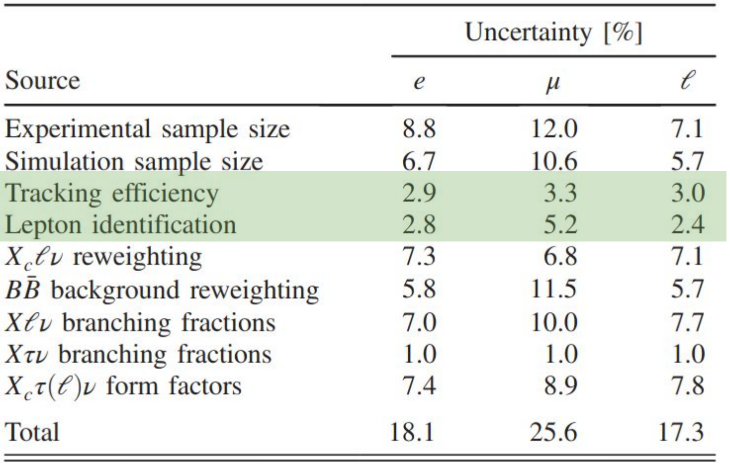
Belle II R(D*) with hFEI 189 fb^{-1}



Source	Uncertainty
PDF shapes	$^{+9.1\%}_{-8.3\%}$
Simulation sample size	$^{+7.5\%}_{-7.5\%}$
$\overline{B} \to D^{**}\ell^- \overline{\nu}_\ell$ branching fractions	$^{+4.8\%}_{-3.5\%}$
Fixed backgrounds	$^{+2.7\%}_{-2.3\%}$
Hadronic B decay branching fractions	$^{+2.1\%}_{-2.1\%}$
Reconstruction efficiency	$^{+2.0\%}_{-2.0\%}$
Kernel density estimation	$^{+2.0\%}_{-0.8\%}$
Form factors	$^{+0.5\%}_{-0.1\%}$
Peaking background in ΔM_{D^*}	$^{+0.4\%}_{-0.4\%}$
$\tau^- \to \ell^- \nu_\tau \overline{\nu}_\ell$ branching fractions	$^{+0.2\%}_{-0.2\%}$
$R(D^*)$ fit method	$^{+0.1\%}_{-0.1\%}$
Total systematic uncertainty	$^{+13.5\%}_{-12.3\%}$

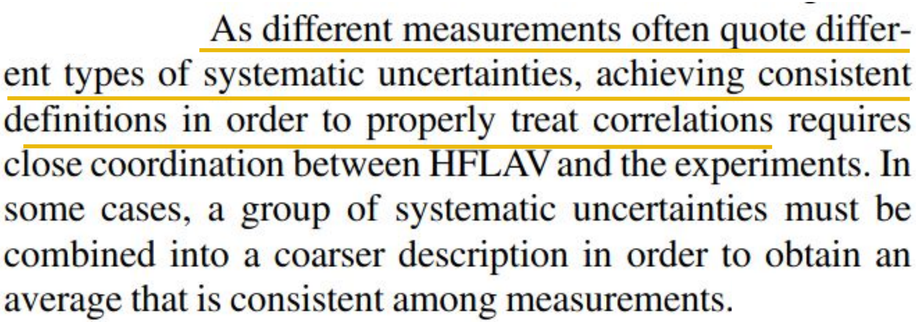
tagging
tracking
LID
HID
π_{slow}
K_{short}
π^0

Belle II R(X) with hFEI 189 fb^{-1}



	Uncertainty [%]		
Source	e	μ	ℓ
Experimental sample size	8.8	12.0	7.1
Simulation sample size	6.7	10.6	5.7
Tracking efficiency	2.9	3.3	3.0
Lepton identification	2.8	5.2	2.4
$X_c\ell\nu$ reweighting	7.3	6.8	7.1
$B\bar{B}$ background reweighting	5.8	11.5	5.7
$X\ell\nu$ branching fractions	7.0	10.0	7.7
$X\tau\nu$ branching fractions	1.0	1.0	1.0
$X_c\tau(\ell)\nu$ form factors	7.4	8.9	7.8
Total	18.1	25.6	17.3

HFLAV's conservative approach: $\rho_{ij} = \min\left(\frac{\sigma_i}{\sigma_j}, \frac{\sigma_j}{\sigma_i}\right)$



As different measurements often quote different types of systematic uncertainties, achieving consistent definitions in order to properly treat correlations requires close coordination between HFLAV and the experiments. In some cases, a group of systematic uncertainties must be combined into a coarser description in order to obtain an average that is consistent among measurements.

With collaboration internal knowledge
and the **proper tools** (i.e. SysVar)
the **exact correlation** can be obtained to
treat common systematics consistently and
benefit from shape effects

And in practice ?

Possible combination scheme

Belle II analysts

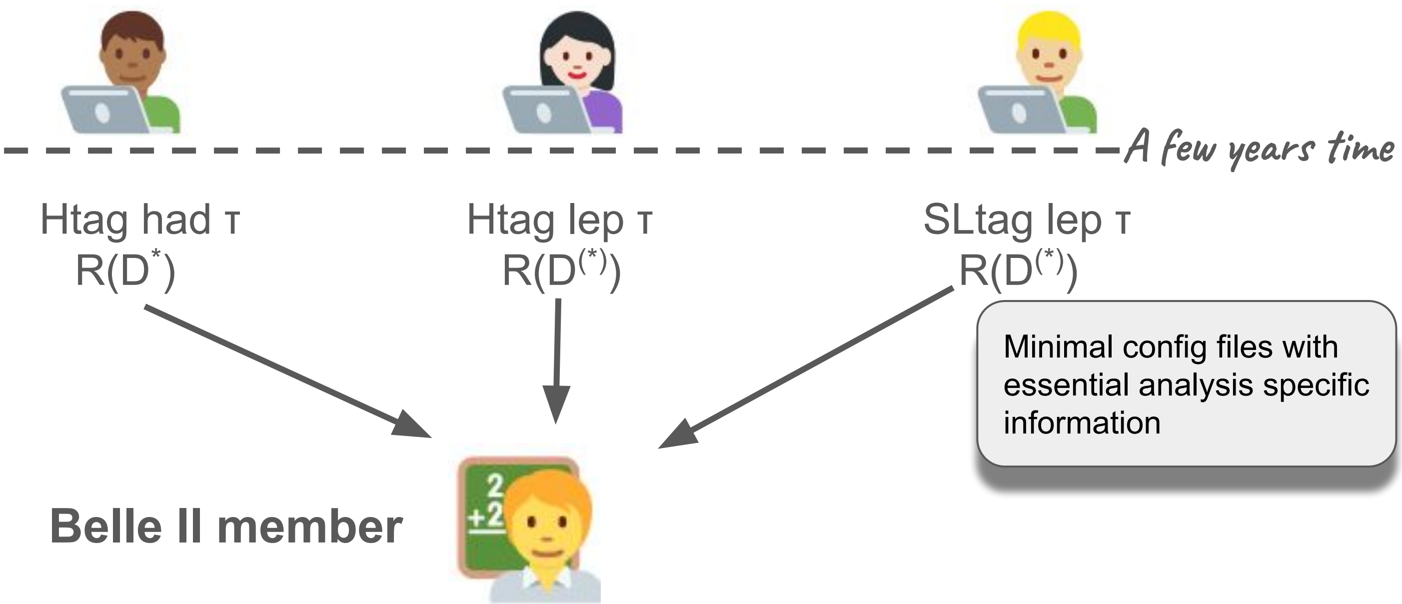
Htag had τ
$R(D^*)$

Htag lep τ
$R(D^{(*)})$

SLtag lep τ
$R(D^{(*)})$

Possible combination scheme

Belle II analysts



A few years time

Htag had τ
$R(D^*)$

Htag lep τ
$R(D^{(*)})$

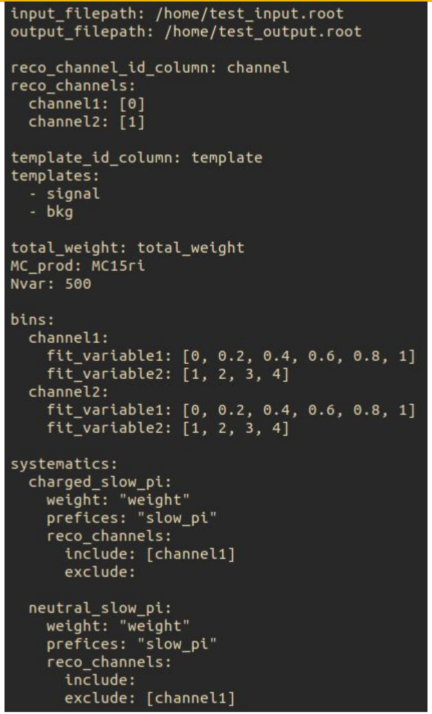
SLtag lep τ
$R(D^{(*)})$

Minimal config files with essential analysis specific information

Belle II member

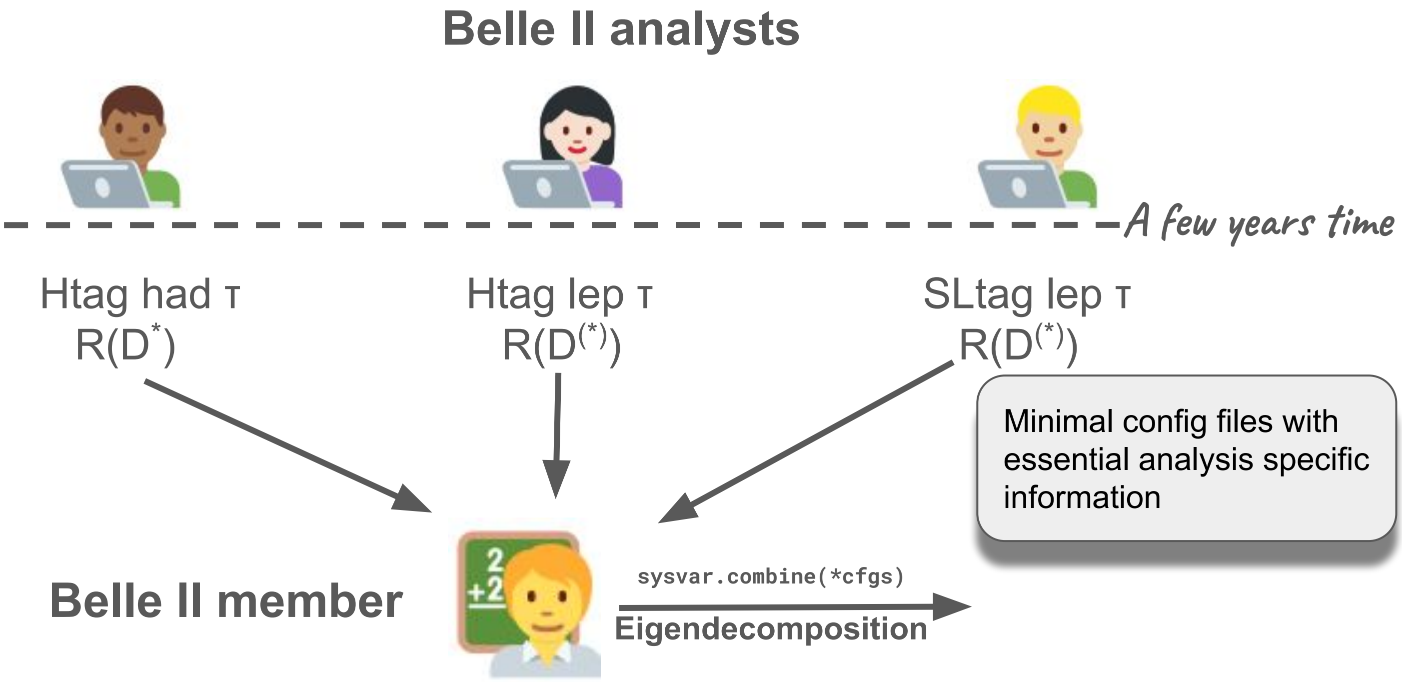
Current requirements:

1. The tuples are stored on the **same machine**
2. The analysts prepare a **SysVar config file**

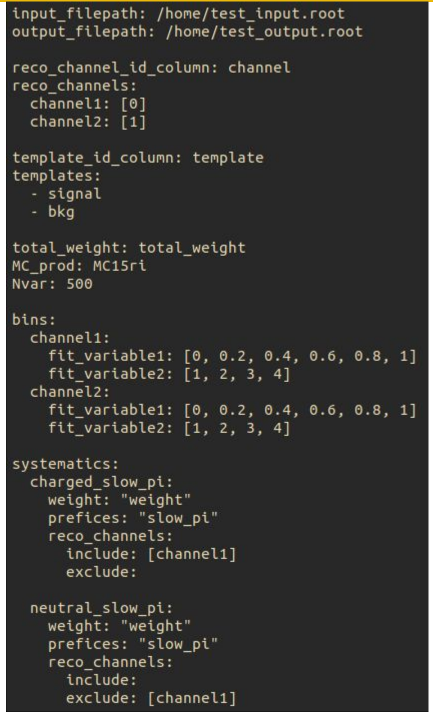


```
input_filepath: /home/test_input.root
output_filepath: /home/test_output.root

reco_channel_id_column: channel
reco_channels:
  channel1: [0]
  channel2: [1]

template_id_column: template
templates:
  - signal
  - bkg

total_weight: total_weight
MC_prod: MC15ri
Nvar: 500

bins:
  channel1:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]
  channel2:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]

systematics:
  charged_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include: [channel1]
      exclude:

  neutral_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include:
      exclude: [channel1]
```

Possible combination scheme



Minimal config files with essential analysis specific information



```
input_filepath: /home/test_input.root
output_filepath: /home/test_output.root

reco_channel_id_column: channel
reco_channels:
  channel1: [0]
  channel2: [1]

template_id_column: template
templates:
  - signal
  - bkg

total_weight: total_weight
MC_prod: MC15ri
Nvar: 500

bins:
  channel1:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]
  channel2:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]

systematics:
  charged_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include: [channel1]
      exclude:

  neutral_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include:
      exclude: [channel1]
```

Current requirements:
1. The tuples are stored on the **same machine**
2. The analysts prepare a **SysVar config file**

Possible combination scheme

Belle II analysts

A few years time

Htag had τ
R(D*)

Htag lep τ
R(D$^{(*)}$)

SLtag lep τ
R(D$^{(*)}$)

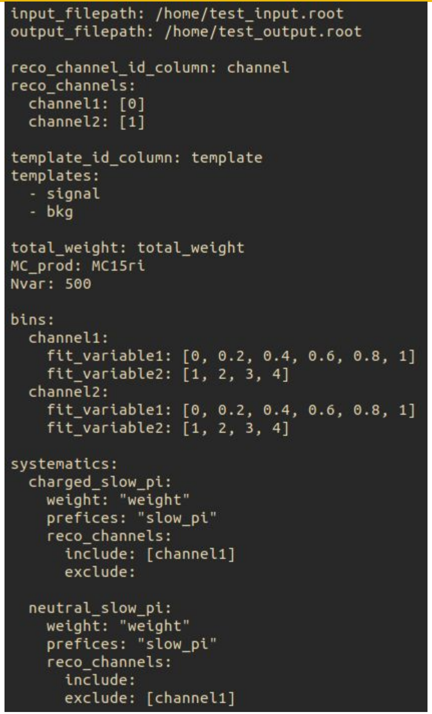
Belle II member

`sysvar.combine(*cfgs)`

Eigendecomposition

Minimal config files with essential analysis specific information

pyhf ∂ifferentiable ℒikelihoods



```
input_filepath: /home/test_input.root
output_filepath: /home/test_output.root

reco_channel_id_column: channel
reco_channels:
  channel1: [0]
  channel2: [1]

template_id_column: template
templates:
  - signal
  - bkg

total_weight: total_weight
MC_prod: MC15ri
Nvar: 500

bins:
  channel1:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]
  channel2:
    fit_variable1: [0, 0.2, 0.4, 0.6, 0.8, 1]
    fit_variable2: [1, 2, 3, 4]

systematics:
  charged_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include: [channel1]
      exclude:

  neutral_slow_pi:
    weight: "weight"
    prefices: "slow_pi"
    reco_channels:
      include:
      exclude: [channel1]
```

Current requirements:
1. The tuples are stored on the **same machine**
2. The analysts prepare a **SysVar config file**

Full likelihood

$$\prod_i \mathscr{L}_i(\mathbf{c}, \theta)$$

statistically validate fit!

A few months time

Belle II combined R(D*)

SysVar’s combination API

Can we handle all the data ?

Eigendecomposition on the likelihood level should not be preferred

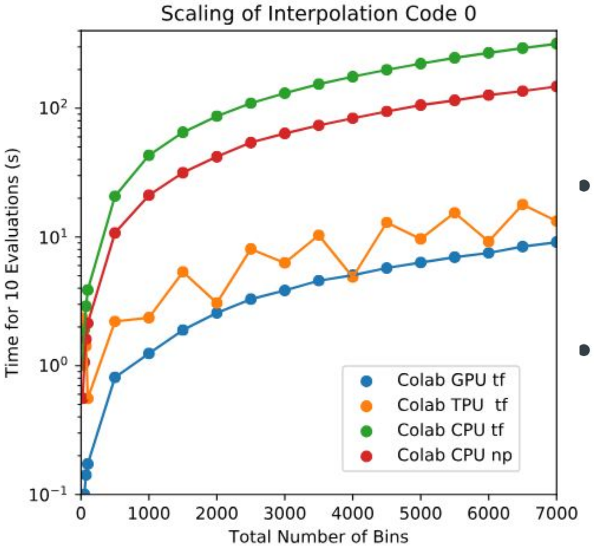
SysVar's Combination API: A Powerful Tool for Streamlined Combined EigenDecomposition

1. **Selective column collection**
    - Automatically retrieve only essential columns, as defined in the cfg file, minimizing memory consumption. Currently supporting a growing list of file formats.
2. **Input merging and unification**
    - Combine and standardize input data structures to ensure consistency across analyses.
3. **Automatic cfg generation**
    - Build a new cfg file for all analyses, simplifying configuration management.
4. **Flexible multi-channel processing**
    - Handle multiple reconstruction channels (even from different analyses) as a single unified workflow.

Impact of large number of NP

Can we fit an increasing number of nuisance parameters ?

pyhf is using **deep learning frameworks** as **computational backends** which allows for exploitation of auto differentiation (**autograd**) and **GPU acceleration**



- Show hardware acceleration giving **order of magnitude speedup** for some models!
- Improvements over traditional
  - 10 hrs to 30 min; 20 min to 10 sec

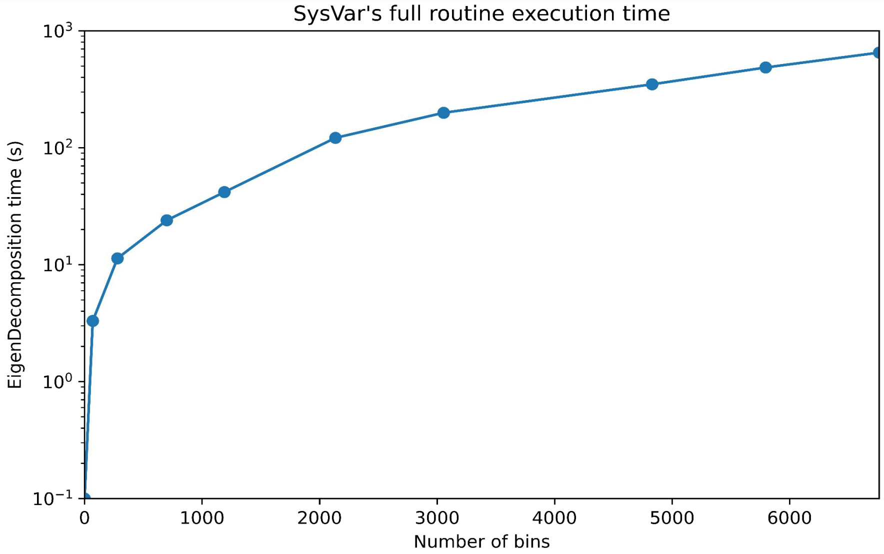
[L. Heinrich at ACAT 2019](#)
[M. Feickert at Scipy 2020](#)

SysVar timing benchmarks

How computationally intensive is the EigenDecomposition ?

Execution time includes:

1. Building templates
2. Calculating variations
3. Performing Eigendecomposition
4. Determining important eigendirections

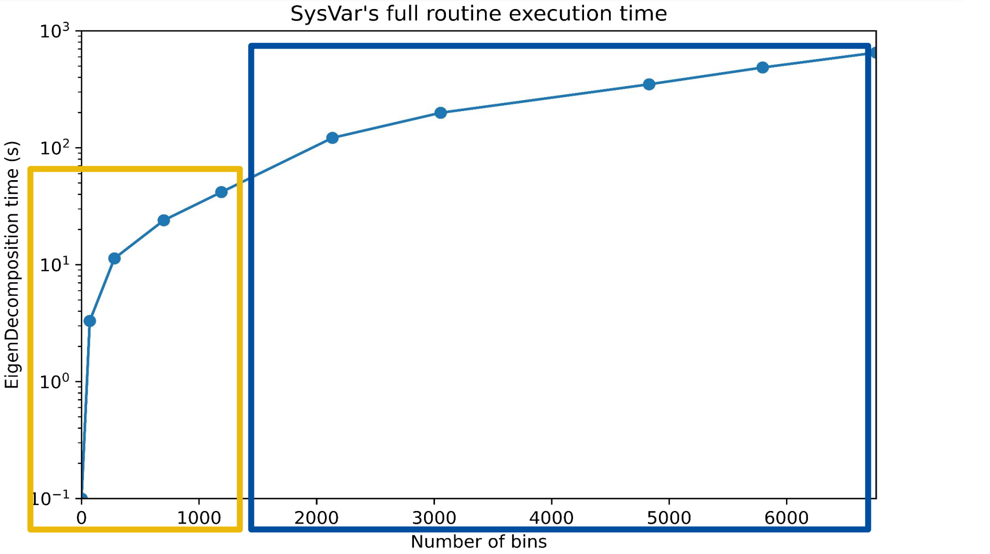


SysVar timing benchmarks

How computationally intensive is the EigenDecomposition ?

Execution time includes:

1. Building templates
2. Calculating variations
3. Performing Eigendecomposition
4. Determining important eigendirections

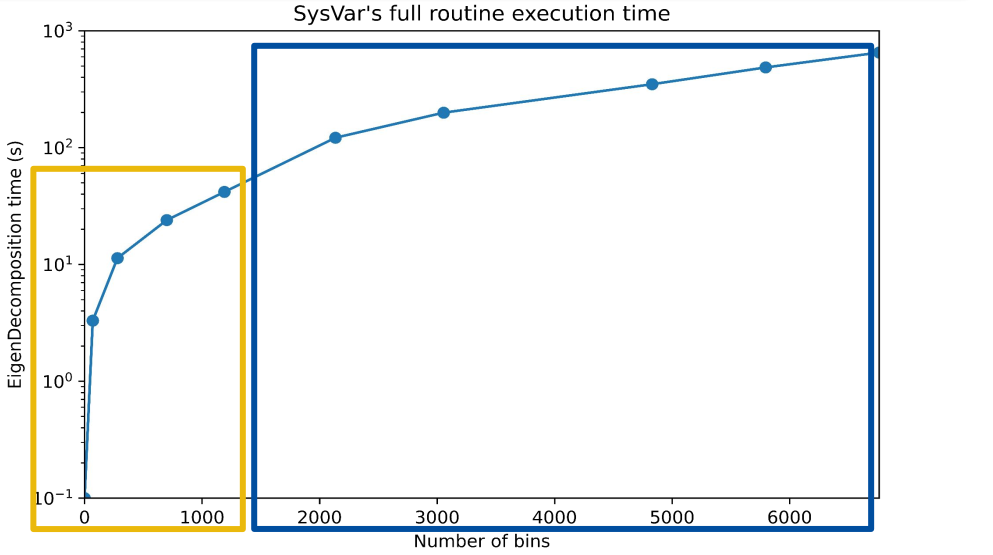


SysVar timing benchmarks

How computationally intensive is the EigenDecomposition ?

Execution time includes:

1. Building templates
2. Calculating variations
3. Performing Eigendecomposition
4. Determining important eigendirections



Depends on the criterion used.
There's room for improvement in the implementation
but still the execution time doesn't make the task impractical

Summary

1. Recent results on R(D$^{(*)}$) by Belle II
2. More orthogonal and complementary measurements are expected soon.
3. A likelihood-based approach has been demonstrated to combine systematics when averaging R(D$^{(*)}$) measurements.
   This method offers key advantages compared to the **HFLAV** approach:
   a. Precise correlations
   b. Tighter constraints of nuisance parameters
   c. Leveraging shape effects in the data for greater precision
4. The approach has been shown to be practical and feasible for real-world application

Summary

1. Recent results on $R(D^{(*)})$ by Belle II
2. More orthogonal and complementary measurements are expected soon.
3. A likelihood-based approach has been demonstrated to combine systematics when averaging $R(D^{(*)})$ measurements.
   This method offers key advantages compared to the **HFLAV** approach:
   1. Precise correlations
   2. Tighter constraints of nuisance parameters
   3. Leveraging shape effects in the data for greater precision
4. The approach has been shown to be practical and feasible for real-world application

Stay tuned for more exciting results on $R(D^{(*)})$ by Belle II !

Considering the true correlations when combining our measurements can improve our precision

Opens road for LHCb/Belle II combinations

Back-up

Hadronically tagged R(D*) at Belle II

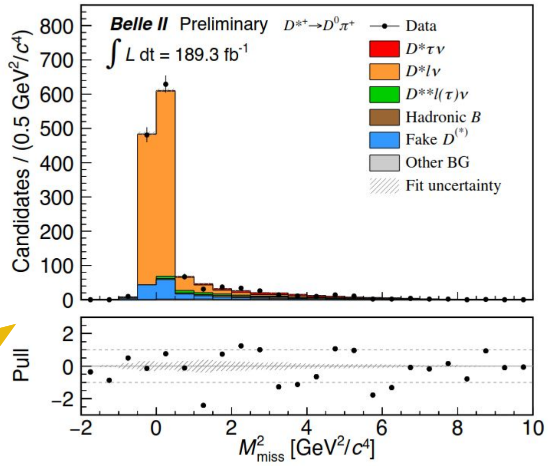
First R(D*) measurement at Belle II !

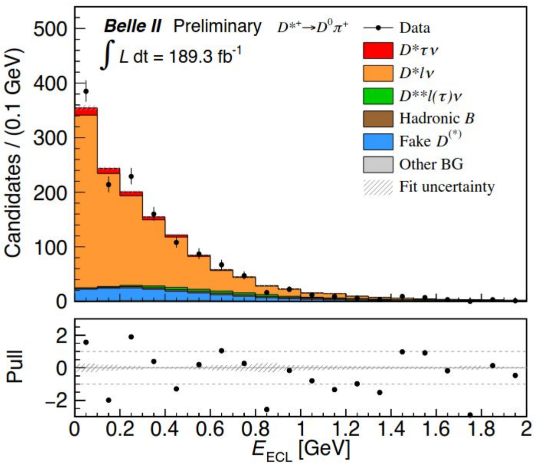
Using **hadronic tag**
Reconstruct $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$
with remaining tracks

leptonic τ decays in both
charged and neutral B mesons

missing mass squared
and
unassigned energy in the calorimeter
to extract signal

Use control regions to constrain main backgrounds
(fake D^*, D^{**} etc)





$$\boxed{R(D^*) = 0.262^{+0.041}_{-0.039}(\text{stat})^{+0.035}_{-0.032}(\text{syst})}$$

Consistent with SM !

Source	Uncertainty
PDF shapes	$^{+9.1\%}_{-8.3\%}$
Simulation sample size	$^{+7.5\%}_{-7.5\%}$
$\overline{B} \to D^{**}\ell^-\overline{\nu}_\ell$ branching fractions	$^{+4.8\%}_{-3.5\%}$
Fixed backgrounds	$^{+2.7\%}_{-2.3\%}$
Hadronic B decay branching fractions	$^{+2.1\%}_{-2.1\%}$
Reconstruction efficiency	$^{+2.0\%}_{-2.0\%}$
Kernel density estimation	$^{+2.0\%}_{-0.8\%}$
Form factors	$^{+0.5\%}_{-0.1\%}$
Peaking background in ΔM_{D^*}	$^{+0.4\%}_{-0.4\%}$
$\tau^- \to \ell^-\nu_\tau\overline{\nu}_\ell$ branching fractions	$^{+0.2\%}_{-0.2\%}$
$R(D^*)$ fit method	$^{+0.1\%}_{-0.1\%}$
Total systematic uncertainty	$^{+13.5\%}_{-12.3\%}$

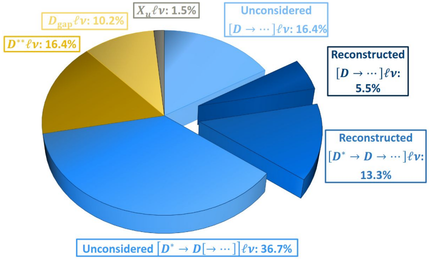
Similar precision to Belle
with **25%** of the data

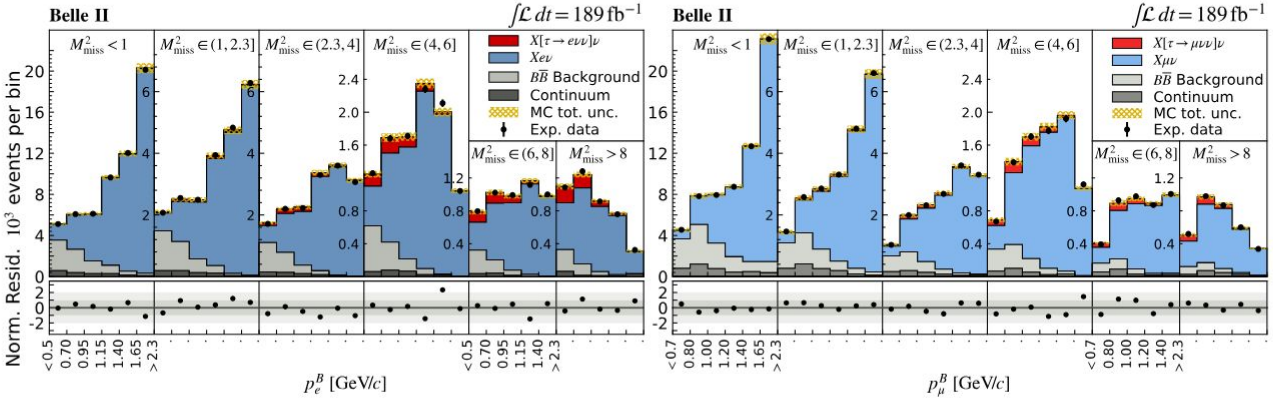
arXiv: 2401.02840

R(D*) from R(X) at Belle II

Using **hadronic tag** reconstruct a **single lepton** and combine the rest into an X system inclusively

$$R(X_{\tau/\ell}) = \frac{\mathcal{B}(X\tau\nu)}{\mathcal{B}(X\ell\nu)}$$





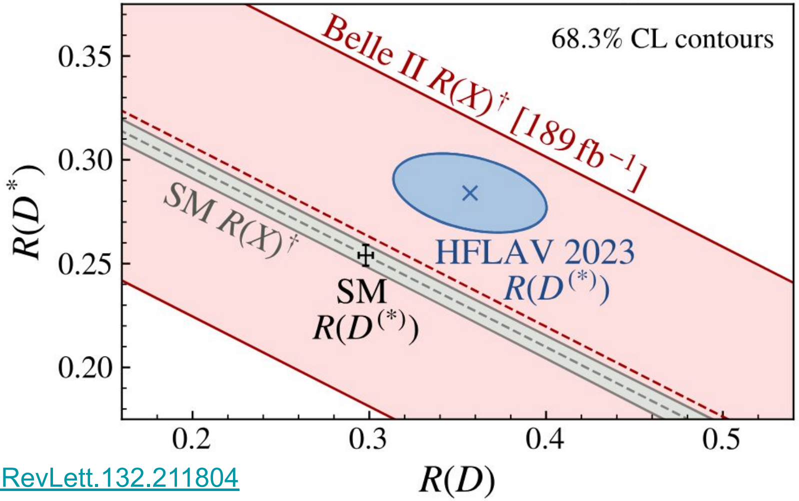
missing mass squared

and

lepton momentum

for signal extraction

$\dagger$ = with expected SM contributions of $D^{**}_{(\text{gap})}, X_u$ removed



$$R(X_{\tau/\ell}) = 0.228 \pm 0.016(\text{stat}) \pm 0.036(\text{syst})$$

Consistent with SM !

$$\mathcal{B}(B \to X\tau\nu) = \mathcal{B}(B \to D\tau\nu) + \mathcal{B}(B \to D^*\tau\nu) + \mathcal{B}(B \to D^{**}_{(\text{gap})}, X_u\tau\nu)$$

Assuming SM like D** BFs.

Statistical correlation with R(D*) ~0.02

Systematic correlation (mainly D** BFs) **non trivial**

PhysRevLett.132.211804

Extracting R(D*)

Status quo: Fit the yields and combine them into an R(D^*) ratio.

$$\mathcal{R}(\mathrm{D}^{(*)}) = 2 \cdot \frac{\nu_{\mathrm{D}^{(*)0}\tau^+} \cdot \epsilon^{-1}_{\mathrm{D}^{(*)0}\tau^+} + \nu_{\mathrm{D}^{(*)-}\tau^+} \cdot \epsilon^{-1}_{\mathrm{D}^{(*)-}\tau^+}}{\nu_{\mathrm{D}^{(*)0}\ell^+} \cdot \epsilon^{-1}_{\mathrm{D}^{(*)0}\ell^+} + \nu_{\mathrm{D}^{(*)-}\ell^+} \cdot \epsilon^{-1}_{\mathrm{D}^{(*)-}\ell^+}}$$

2 light leptons

yield

reconstruction efficiency

Certain uncertainties associated to reconstruction efficiencies may not fully cancel as they don't fully factorize between B^0 and B^+

New approach: Parameterize the yields by assuming isospin symmetry and fit R(D*) directly

This leads to a safer treatment of the efficiencies that now appear directly as ratios

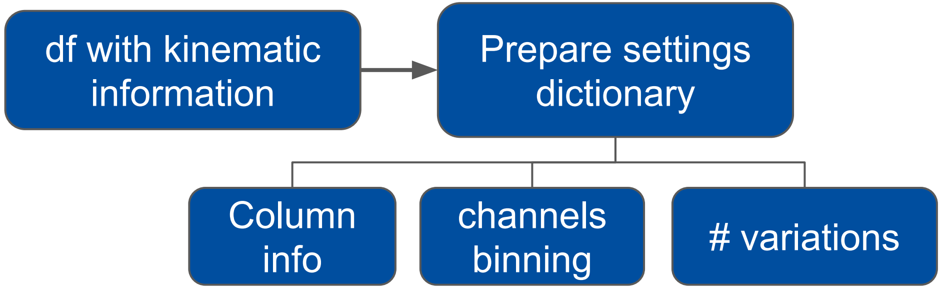
$$\nu_{\mathrm{D}^{*-}\tau^+} = \frac{1}{2}\mathcal{R}(\mathrm{D}^*) \cdot \nu_{\mathrm{D}^{*0}\ell^+} \cdot \frac{\epsilon_{\mathrm{D}^{*-}\tau^+}}{\epsilon_{\mathrm{D}^{*0}\ell^+}} \cdot \tau_{0+}$$

$$\nu_{\mathrm{D}^{*0}\tau^+} = \frac{1}{2}\mathcal{R}(\mathrm{D}^*) \cdot \nu_{\mathrm{D}^{*0}\ell^+} \cdot \frac{\epsilon_{\mathrm{D}^{*0}\tau^+}}{\epsilon_{\mathrm{D}^{*0}\ell^+}}.$$

SysVar: A tool enhancing consistency in the treatment of systematics

df with kinematic information

Quick flow chart of producing EigenVariations



df with kinematic information

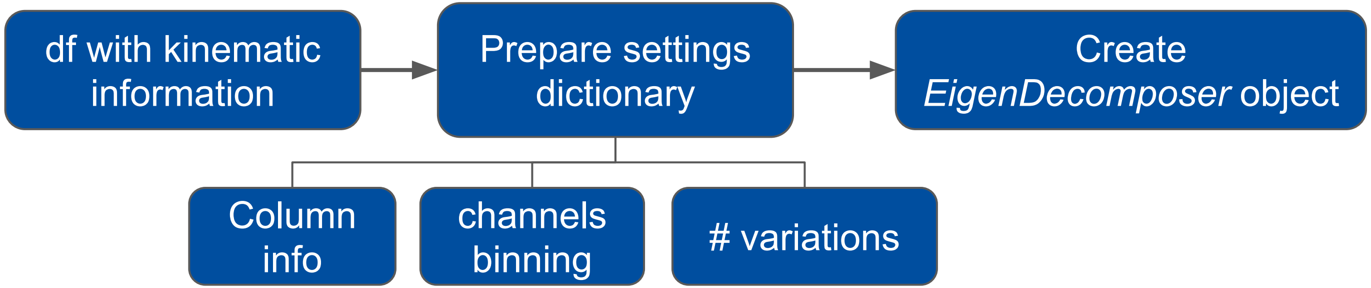
Prepare settings dictionary

Column info

channels binning

variations

Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

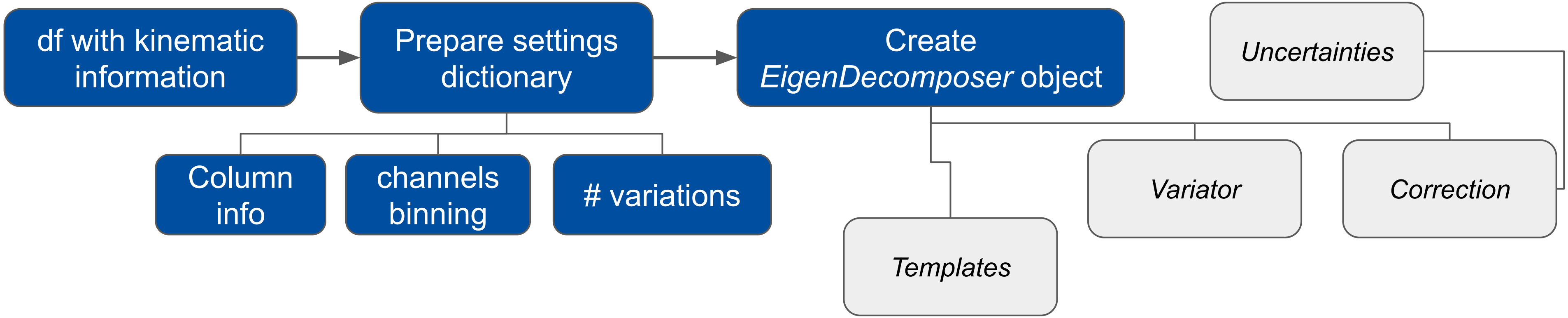
Create *EigenDecomposer* object

Column info

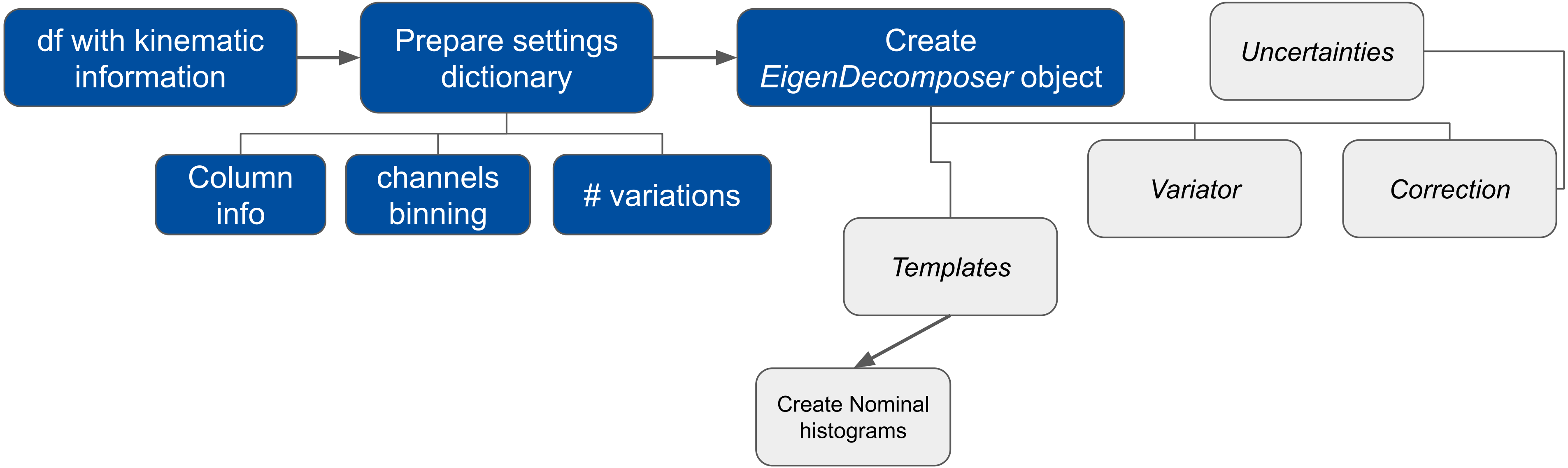
channels binning

variations

Quick flow chart of producing EigenVariations



Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

Create *EigenDecomposer* object

Uncertainties

Column info

channels binning

variations

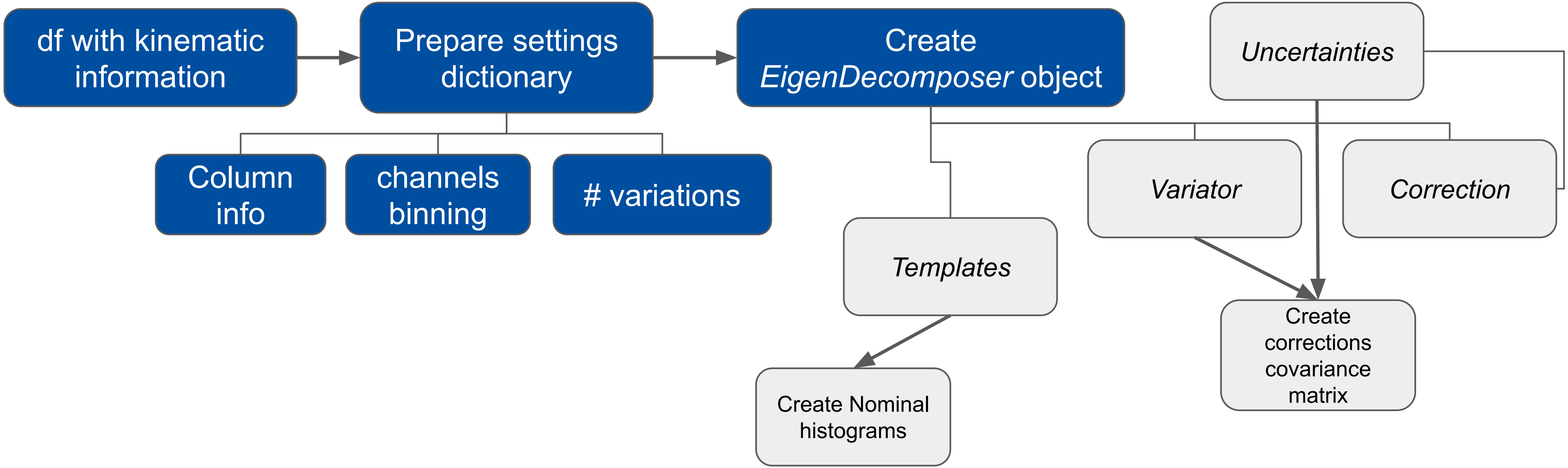
Variator

Correction

Templates

Create Nominal histograms

Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

Column info

channels binning

variations

Create *EigenDecomposer* object

Uncertainties

Variator

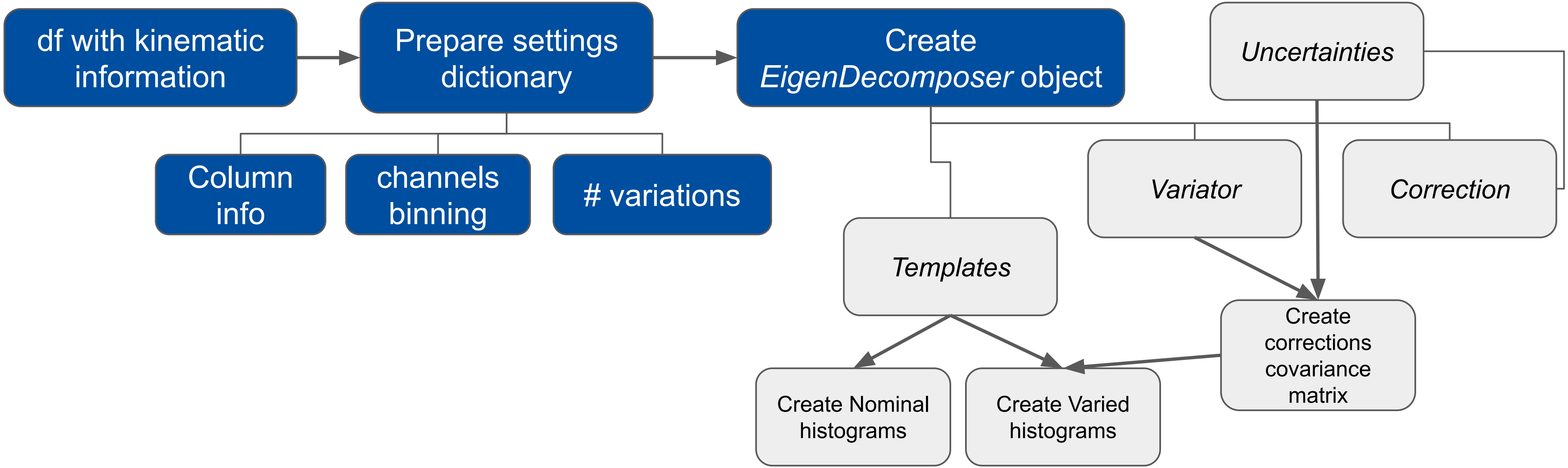
Correction

Templates

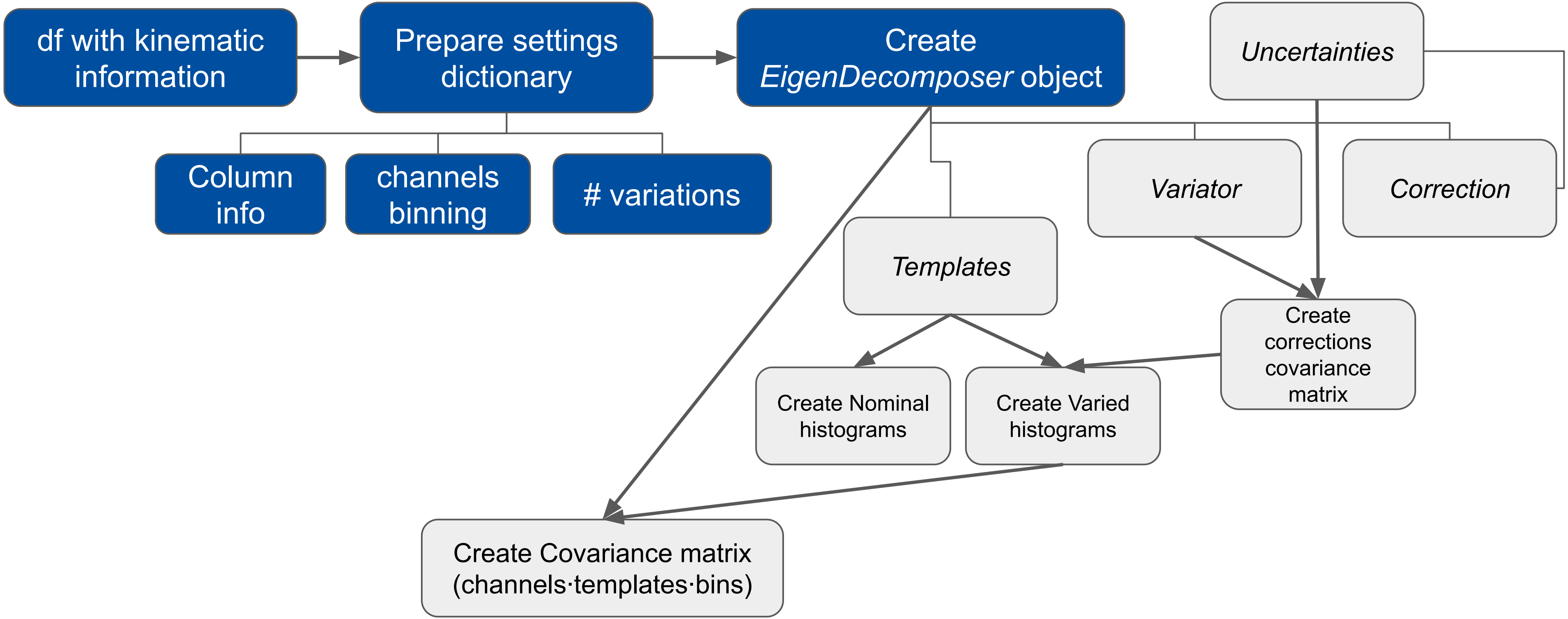
Create Nominal histograms

Create corrections covariance matrix

Quick flow chart of producing EigenVariations



Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

Create *EigenDecomposer* object

Uncertainties

Column info

channels binning

variations

Variator

Correction

Templates

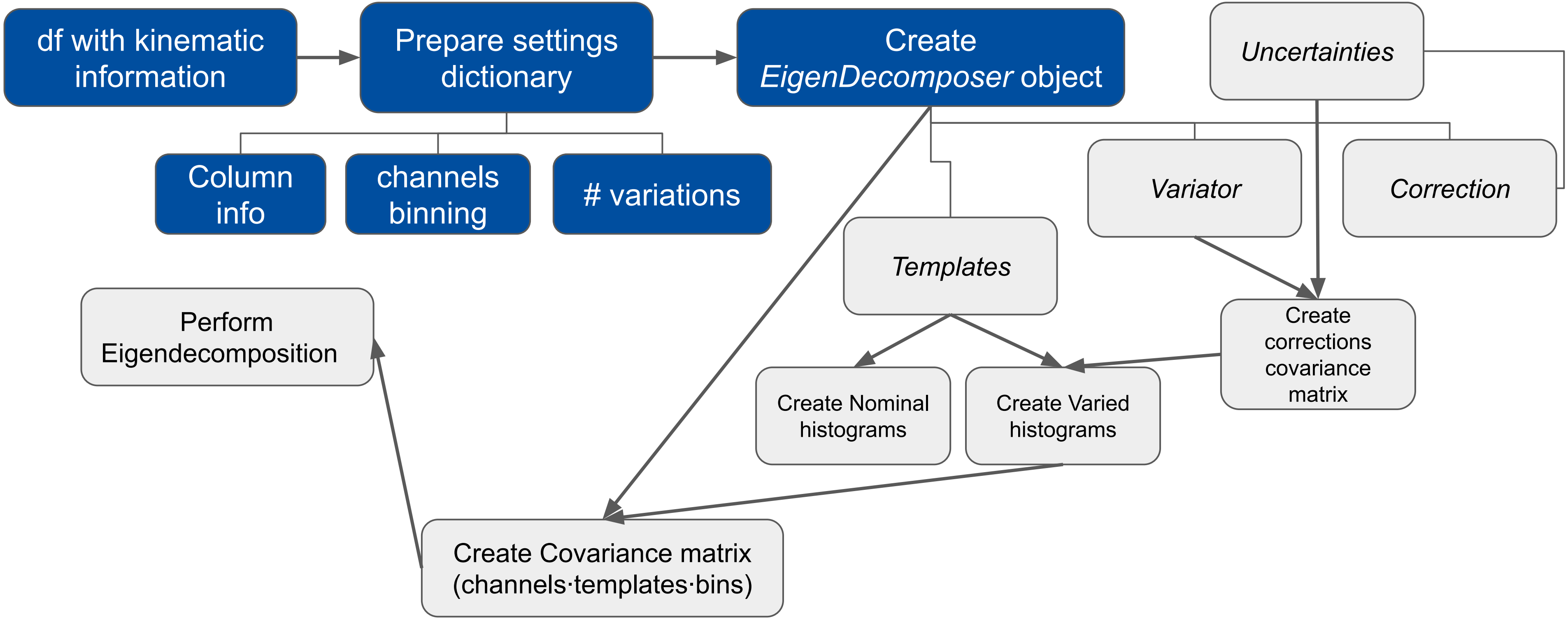
Create corrections covariance matrix

Create Nominal histograms

Create Varied histograms

Create Covariance matrix (channels·templates·bins)

Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

Create *EigenDecomposer* object

Uncertainties

Column info

channels binning

variations

Variator

Correction

Templates

Perform Eigendecomposition

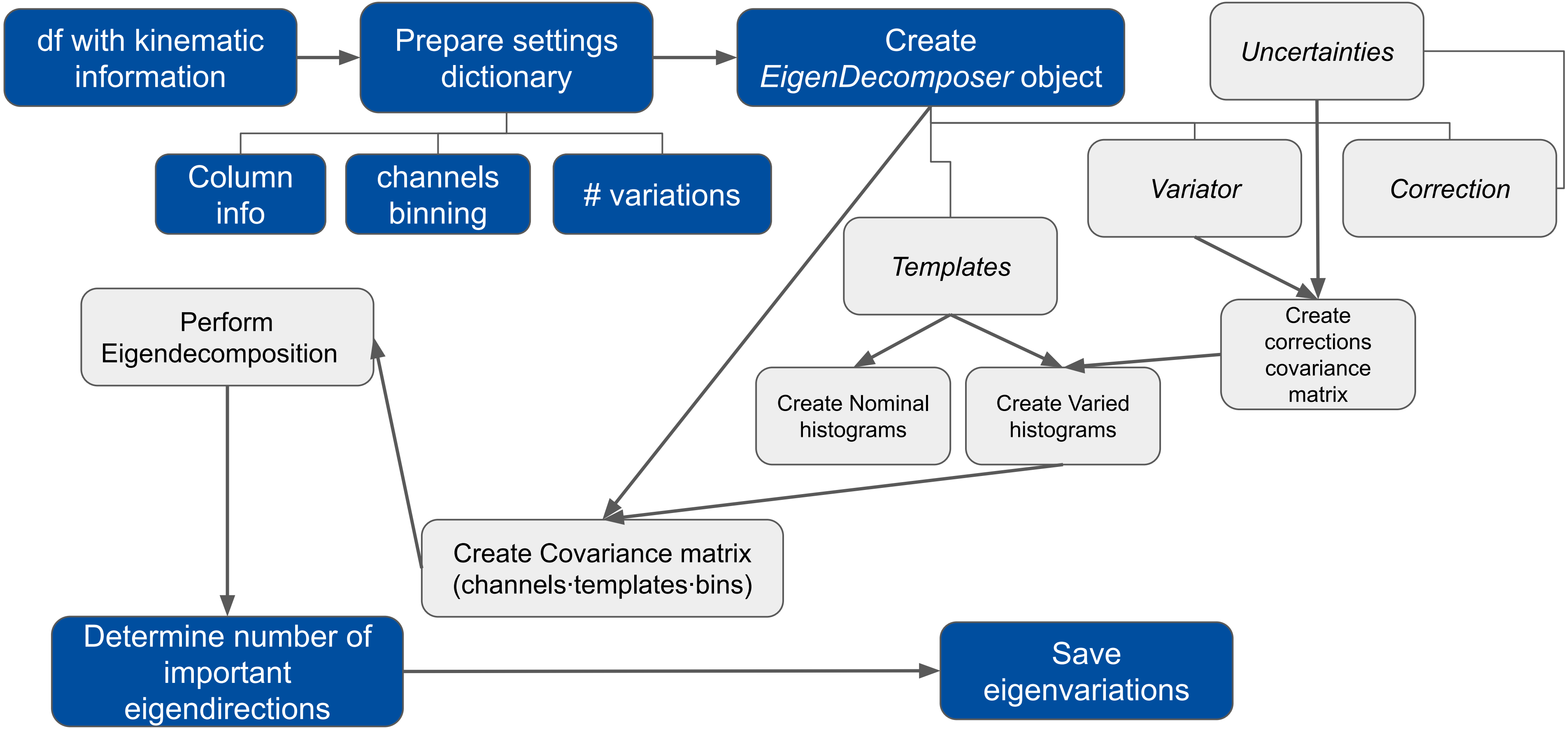
Create corrections covariance matrix

Create Nominal histograms

Create Varied histograms

Create Covariance matrix (channels·templates·bins)

Quick flow chart of producing EigenVariations



df with kinematic information

Prepare settings dictionary

Create *EigenDecomposer* object

Uncertainties

Column info

channels binning

variations

Variator

Correction

Templates

Perform Eigendecomposition

Create corrections covariance matrix

Create Nominal histograms

Create Varied histograms

Create Covariance matrix (channels·templates·bins)

Determine number of important eigendirections

Save eigenvariations