LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

A model-independent likelihood function for the Belle II $B^+ \to K^+\nu\bar{\nu}$ analysis

Lorenz Gärtner[1,2] *on behalf of Belle II,*

in collaboration with

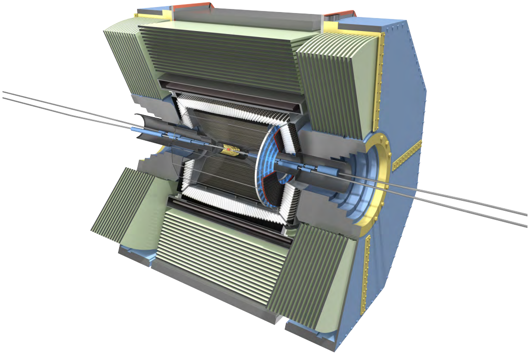
Danny van Dyk[3], Lukas Heinrich[2,4], Méril Reboud[3]

[1]*LMU Munich,* [2]*Excellence Cluster ORIGINS,*
[3]*IPPP Durham,* [4]*TU Munich*

30.08.2023

SuperKEKB & Belle II introduction

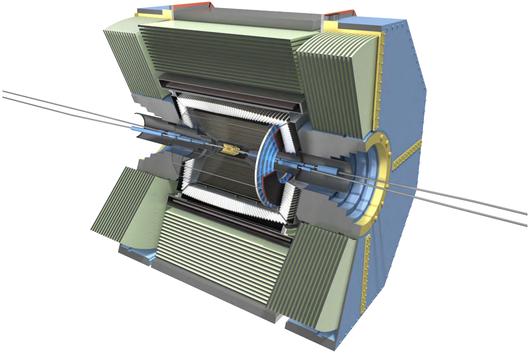
- CP violation / CKM measurements, LFV, BSM FCNC
- **Luminosity** vs. energy frontier



arXiv:1808.10567 [hep-ex]

SuperKEKB & Belle II introduction

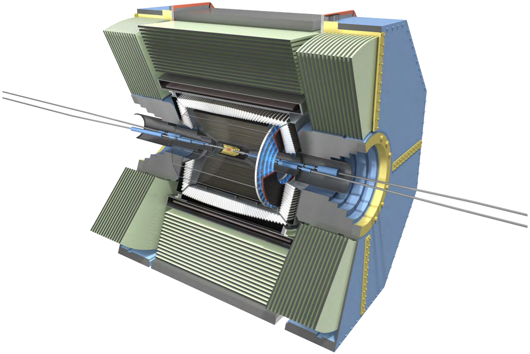
- CP violation / CKM measurements, LFV, BSM FCNC
- **Luminosity** vs. energy frontier
- **SuperKEKB**
  - asymmetric $e^-(7\ \text{GeV}) - e^+(4\ \text{GeV})$
  - $\mathcal{L}_{max} = 8 \times 10^{35}\ \text{cm}^{-2}\text{s}^{-1}$
  - $\sqrt{s} \approx 10.5\ \text{GeV}$
  - b, c, τ factory



arXiv:1808.10567 [hep-ex]

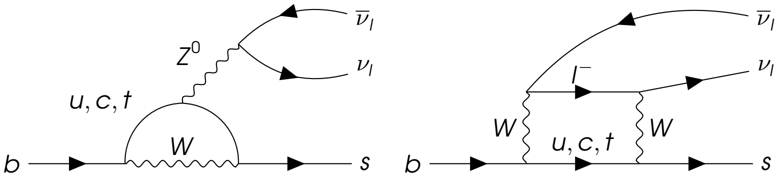
SuperKEKB & Belle II introduction

- CP violation / CKM measurements, LFV, BSM FCNC
- **Luminosity** vs. energy frontier
- **SuperKEKB**
  - asymmetric $e^{-}(7\text{ GeV}) - e^{+}(4\text{ GeV})$
  - $\mathcal{L}_{max} = 8 \times 10^{35}\,\text{cm}^{-2}\text{s}^{-1}$
  - $\sqrt{s} \approx 10.5\text{ GeV}$
  - b, c, τ factory
- **Belle II**
  - Hermetic, azimuthally asymmetric detector
  - ➜ Missing mass analyses possible



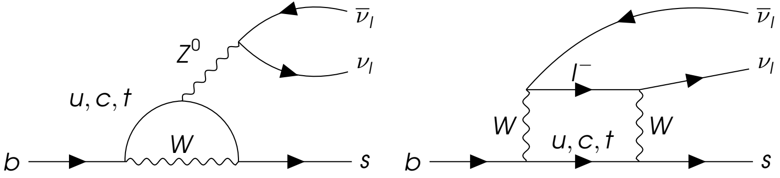
arXiv:1808.10567 [hep-ex]

Why a $B^+ \to K^+\nu\bar{\nu}$ reinterpretation?



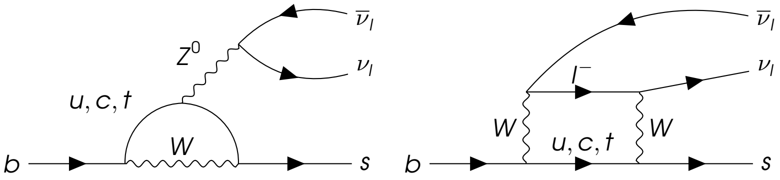
- Due to the suppression of FCNCs in the SM, tree level BSM effects could substantially affect the rate. A $B^+ \to K^+\nu\bar{\nu}$ measurement can help to constrain BSM physics.

Why a $B^+ \to K^+ \nu\bar{\nu}$ reinterpretation?



- Due to the suppression of FCNCs in the SM, tree level BSM effects could substantially affect the rate. A $B^+ \to K^+ \nu\bar{\nu}$ measurement can help to constrain BSM physics.
- **Shortcomings of model-dependence**
  - **Limited interpretability** in terms of any BSM or future SM physics with different kinematic predictions.

Why a $B^+ \to K^+ \nu\bar{\nu}$ reinterpretation?



- Due to the suppression of FCNCs in the SM, tree level BSM effects could substantially affect the rate. A $B^+ \to K^+ \nu\bar{\nu}$ measurement can help to constrain BSM physics.
- **Shortcomings of model-dependence**
  - **Limited interpretability** in terms of any BSM or future SM physics with different kinematic predictions.
- **Benefits of reinterpretation**
  - **Sensitivity** to any current or future (B)SM prediction.
  - **Exclusion** limits in BSM parameter space inferable.
  - **Combinations** with other measurements possible.

Analysis

Where is the model dependence?

The Belle II $B^+ \to K^+ \nu\bar{\nu}$ analysis

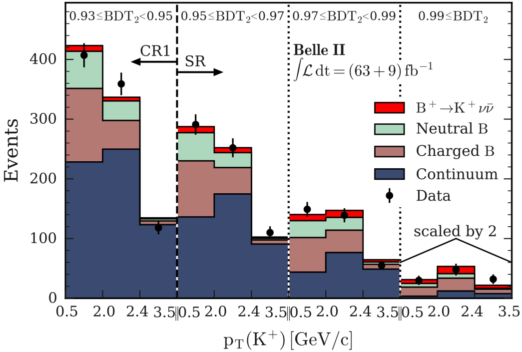
1. Two consecutive BDTs separate signal from background.
2. The signal MC is weighted according to the SM prediction of the kinematic distribution
   $\to$ **model dependence**
3. Max. likelihood fit in bins of $p_T(K^+) \times BDT_2$.

[hepdata.130199]
[Phys.Rev.Lett.127.181802]

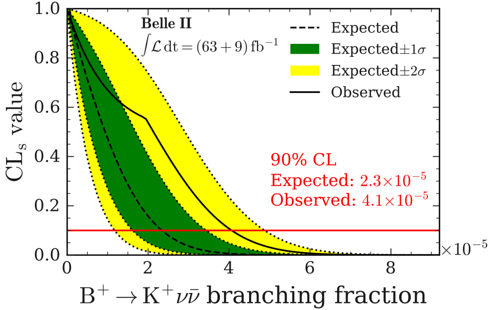
The Belle II $B^+ \to K^+ \nu\bar{\nu}$ analysis

1. Two consecutive BDTs separate signal from background.
2. The signal MC is weighted according to the SM prediction of the kinematic distribution
   $\to$ **model dependence**
3. Max. likelihood fit in bins of $p_T(K^+) \times BDT_2$.

[hepdata.130199]
[Phys.Rev.Lett.127.181802]



CR = control region, SR = signal region



$\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) < 4.1 \times 10^{-5}$@90%CL

Reweighting approach

How do we obtain new signal templates?

Reweighting information

Kinematic d.o.f. of $B^+ \to K^+ \nu\bar{\nu}$

The differential branching ratio is a function of the squared dineutrino invariant mass, $q^2 = (p_\nu + p_{\bar{\nu}})^2$.

Reweighting information

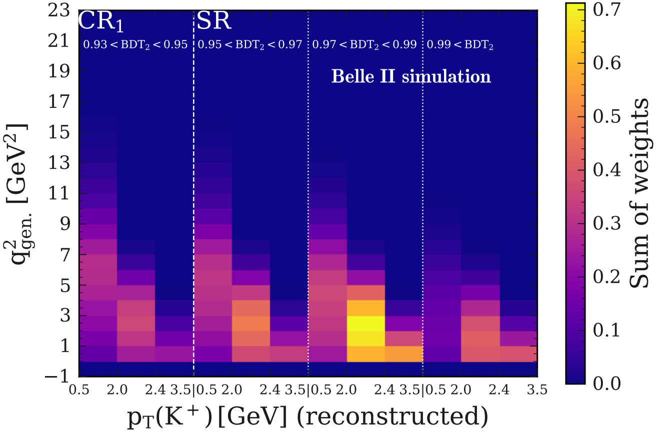
Kinematic d.o.f. of $B^+ \to K^+ \nu\bar{\nu}$

The differential branching ratio is a function of the squared dineutrino invariant mass, $q^2 = (p_\nu + p_{\bar{\nu}})^2$.

Weight updating

- Need information on the q^2 distribution per analysis bin.
- 3d binning:

$$\underbrace{p_T(K^+) \times BDT_2}_{\text{analysis binning (reconstructed)}} \times \underbrace{q^2_{gen.}}_{\text{kinematic d.o.f. (generated)}}$$



Reweighting approach

Recipe

1. Get distributions of kinematic d.o.f (q^2)

$$N_{k\,l\,m} = \underbrace{p_T \times BDT_2}_{\text{analysis binning}} \times \underbrace{q^2_{gen.}}_{\text{kinematic d.o.f}}$$

2. Apply weights in bins of kinematic d.o.f.

$$N_{k\,l} = \sum_{m \in q^2} N^{\text{PHSP}}_{k\,l\,m} w_m = \sum_{m \in q^2} N^{\text{PHSP}}_{k\,l\,m} \int_m dq^2 \frac{d\Gamma^{(B)SM}}{dq^2} \left(\frac{d\Gamma^{\text{PHSP}}}{dq^2} \right)^{-1}$$

Reweighting approach

Recipe

1. Get distributions of kinematic d.o.f (q^2)

$$N_{klm} = \underbrace{p_T \times BDT_2}_{\text{analysis binning}} \times \underbrace{q^2_{gen.}}_{\text{kinematic d.o.f}}$$

2. Apply weights in bins of kinematic d.o.f.

$$N_{kl} = \sum_{m \in q^2} N^{\text{PHSP}}_{klm} w_m = \sum_{m \in q^2} N^{\text{PHSP}}_{klm} \int_m dq^2 \frac{d\Gamma^{(B)SM}}{dq^2} \left(\frac{d\Gamma^{\text{PHSP}}}{dq^2} \right)^{-1}$$

Benefits

+ Good accuracy with sufficient number of bins
+ Very versatile
+ Easily publishable

Theory

How can we parametrize our model dependence?

Weak Effective Theory for $B \to K\nu\bar{\nu}$

Contribution operators

The effective Lagrangian is

$$\mathcal{L}^{WET} = \sum_{X=L,R} C_{\mathrm{VX}}\, \mathcal{O}_{\mathrm{VX}} + \sum_{X=L,R} C_{\mathrm{SX}}\, \mathcal{O}_{\mathrm{SX}} + C_{\mathrm{TL}}\, \mathcal{O}_{\mathrm{TL}} + \text{h.c.}$$

The $d = 6$ contributing operators in and beyond the SM are given by

$$\mathcal{O}_{\mathrm{VL}} = \left(\overline{\nu_L}\gamma_\mu \nu_L\right)\left(\overline{s_L}\gamma^\mu b_L\right) \qquad \mathcal{O}_{\mathrm{VR}} = \left(\overline{\nu_L}\gamma_\mu \nu_L\right)\left(\overline{s_R}\gamma^\mu b_R\right)$$

$$\mathcal{O}_{\mathrm{SL}} = \left(\overline{\nu_L^C}\nu_L\right)\left(\overline{s_R} b_L\right) \qquad \mathcal{O}_{\mathrm{SR}} = \left(\overline{\nu_L^C}\nu_L\right)\left(\overline{s_L} b_R\right)$$

$$\mathcal{O}_{\mathrm{TL}} = \left(\overline{\nu_L^C}\sigma_{\mu\nu}\nu_L\right)\left(\overline{s_R}\sigma^{\mu\nu} b_L\right)$$

[arXiv:2111.04327 [hep-ph]]

Weak Effective Theory for $B \to K\nu\bar{\nu}$

Decay width

Decay width dependence on the Wilson coefficients is given by

$$\frac{d\Gamma\left(B \to K\nu\bar{\nu}\right)}{dq^2} = \frac{\sqrt{\lambda_{BK}}q^2}{(4\pi)^3 m_B^3}\left[\frac{\lambda_{BK}}{24q^2}\left|f_+(q^2)\right|^2 \left|C_{\mathrm{VL}} + C_{\mathrm{VR}}\right|^2 + \frac{\left(m_B^2 - m_K^2\right)^2}{8\left(m_b - m_s\right)^2}\left|f_0(q^2)\right|^2 \left|C_{\mathrm{SL}} + C_{\mathrm{SR}}\right|^2 + \frac{2\lambda_{BK}}{3\left(m_B + m_K\right)^2}\left|f_T(q^2)\right|^2 \left|C_{\mathrm{TL}}\right|^2\right]$$

valid for $J^P = 0^-$ kaon states.

[arXiv:2111.04327 [hep-ph]]

(B)SM theory predictions

We can capture BSM physics that lives exclusively above the scale of electroweak symmetry breaking within 3 linear combinations of Wilson coefficients*

$C_{VL} + C_{VR}$ $C_{SL} + C_{SR}$ C_{TL}

eos.github.io

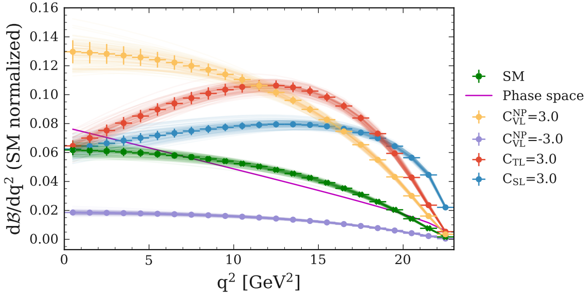
*$C_{VL} = C_{VL}^{SM} + C_{VL}^{NP}$

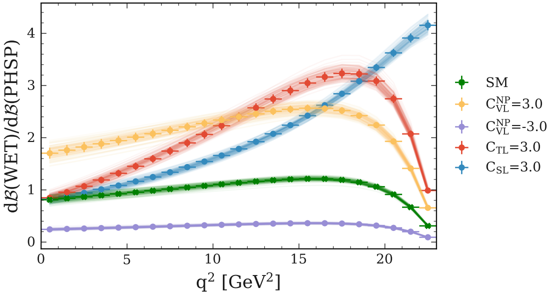
(B)SM theory predictions

We can capture BSM physics that lives exclusively above the scale of electroweak symmetry breaking within 3 linear combinations of Wilson coefficients*

$C_{VL} + C_{VR}$ $C_{SL} + C_{SR}$ C_{TL}

eos.github.io





Uncertainties originate from form factors and are q^2 dependent.

$^*C_{VL} = C_{VL}^{SM} + C_{VL}^{NP}$

Reinterpretation

How do we bring it all together?

Implementation

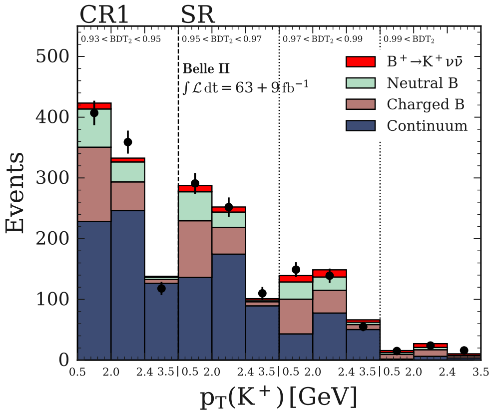
eos.github.io

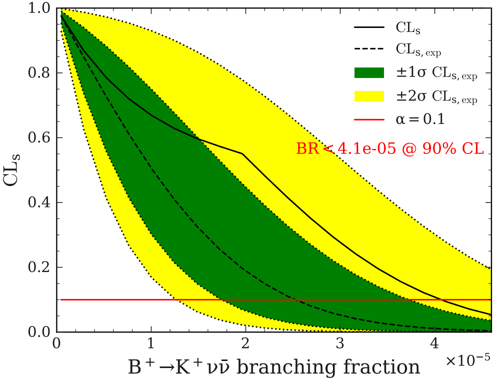
- Calculate theoretical predictions
- Theory parameters: Wilson coefficients & hadronic parameters

pyhf.readthedocs.io

- Built a "custom modifier" that generates new signal template from theory parameters.
- Theory parameters become fitting parameters.

Cross check: Reproducing upper limit





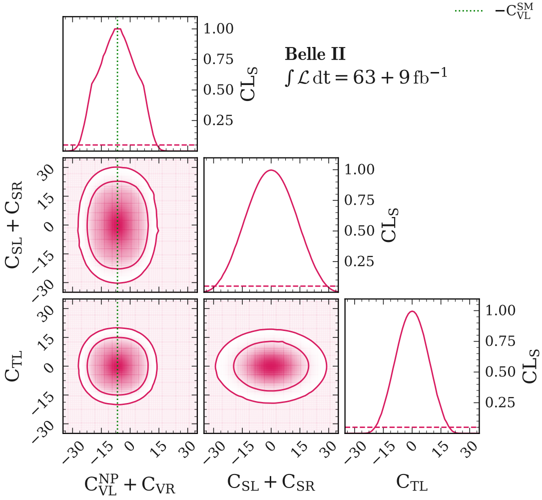
$\mathcal{B}(B^+ \to K^+ \nu\bar{\nu}) < 4.1 \times 10^{-5}$@90%CL

Wilson coefficient exclusion limits

- We computed CL_s values for Wilson coefficients on a 1d grid, profiling over the other two Wilson coefficients (diagonals).
- The horizontal line corresponds to $CL_s = 0.05$.
- Exclusion limits @95%CL

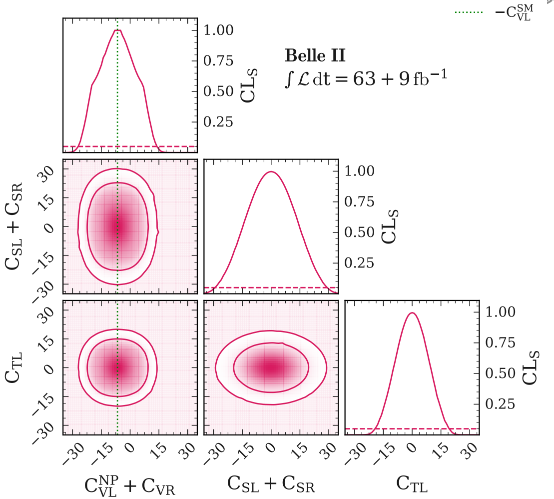
$$|C_{VL} + C_{VR}| < 20.6$$
$$|C_{SL} + C_{SR}| < 29.3$$
$$|C_{TL}| < 19.4$$

$$C_{VL} = C_{VL}^{SM} + C_{VL}^{NP}$$



Wilson coefficient exclusion limits

- We computed CL_s values for Wilson coefficients on a 1d grid, profiling over the other two Wilson coefficients (diagonals).
- The horizontal line corresponds to $CL_s = 0.05$.
- We computed CL_s values for Wilson coefficients on a 2d grid, profiling over the third Wilson coefficient (off-diagonals). contours correspond to $CL_s = 0.32$ (inner) and $CL_s = 0.05$ (outer) limits.
- The region outside $CL_s = 0.05$ is excluded at 95% confidence level.
- Grid range is taken loosely from limits in arXiv:2111.04327 [hep-ph] (backup)



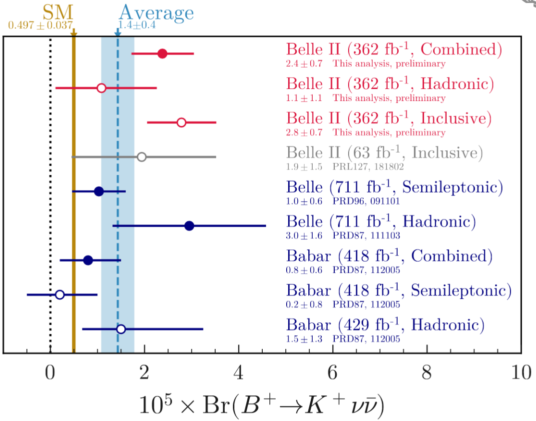
Analysis update: $B^+ \to K^+ \nu\bar{\nu}$ @ 362 fb^{-1}

Inclusive	$\mathcal{B} = \left(2.8^{+0.5}_{-0.5}{}^{+0.5}_{-0.5}\right) \times 10^{-5}$
Hadronic	$\mathcal{B} = \left(1.1^{+0.9}_{-0.8}{}^{+0.8}_{-0.5}\right) \times 10^{-5}$
Combined	$\mathcal{B} = \left(2.4^{+0.5}_{-0.5}{}^{+0.5}_{-0.4}\right) \times 10^{-5}$

Significance of the **combined** result:

- 3.6σ wrt. null hypothesis
- 2.8σ wrt. SM

First evidence of $B^+ \to K^+ \nu\bar{\nu}$



Model-independent likelihood method will be applied and published once paper is accepted.

Presented at EPS 2023

Summary

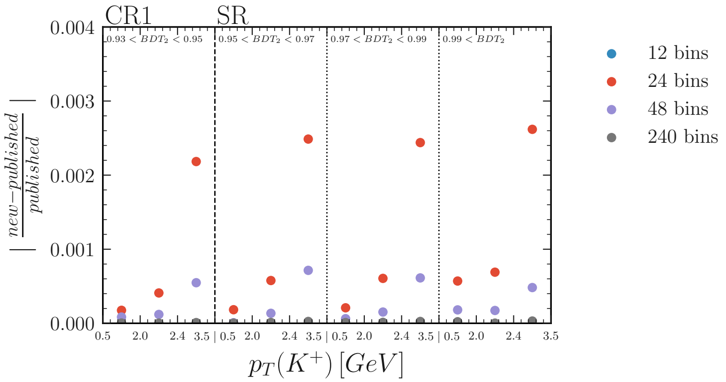
- **Challenge**: Neutrino-induced experimental complexities in $B^+ \to K^+ \nu\bar{\nu}$ lead to model-dependent results due to kinematic assumptions and hadronic matrix element description.
- **Solution**: A model-independent likelihood function enables maximum likelihood fits for any given (B)SM signal prediction, using the supplied information about the q^2 distribution.
- **Tool integration**:
  - Extend `pyhf` and interface it with `EOS` for run-time template updating.
  - Method fully applicable to other decay channels and results.
- **Benefits**:
  - **Exploration of exclusions in BSM parameter space.**
  - Individual model studies with provided decay rate predictions.
  - ...
- **Significance**: Publishing such likelihoods is crucial for a full exploitation of experimental results.

---

lorenz.gaertner@physik.uni-muenchen.de

Binning choice

We compared the relative accuracy of the binned weighting (new) with the event-by-event weighting (published).



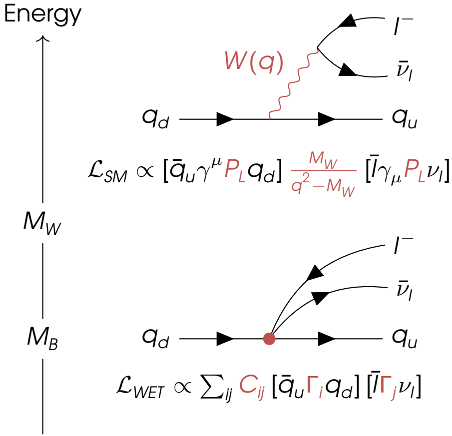
Effective field theory

- **High energy collisions**
  Enough energy to radiate off an on-shell (massive) W boson.
- **Lower energy quark decays**
  The W boson is always off-shell.
- **Weak effective theory**
  The W is integrated out, and its effects are encoded in new couplings

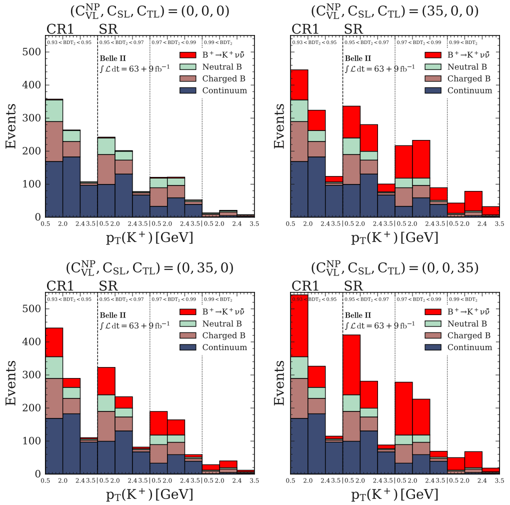
$$\mathcal{L}_{SM} \to \mathcal{L}_{WET} = \sum C_i \mathcal{O}_i$$

- $\mathcal{O}_i$ are completely model independent.

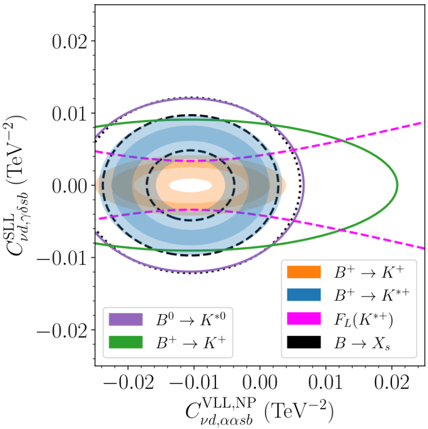
➜ Model independent **parametrization**, constrained only by Wilson coefficients C_i.

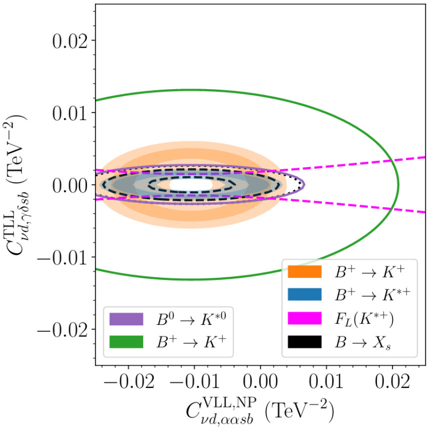


Expected yields



Bounds from `arXiv:2111.04327 [hep-ph]`





Caution: Wilson coefficients have a slightly different interpretation.

Parameter space selection

The definition of Wilson coefficients in `arXiv:2111.04327 [hep-ph]` compared to the values used in EOS are

$$C_{paper} = -\frac{4G_F}{\sqrt{2}}\frac{\alpha}{2\pi}V_{ts}^* V_{tb}\left(\frac{X}{\sin^2\theta_W}\right) = -\frac{4G_F}{\sqrt{2}}\frac{\alpha}{2\pi}V_{ts}^* V_{tb} C_{EOS} \approx \frac{1}{615 TeV^2} C_{EOS}.$$

We get a rough estimate of the parameter space from `arXiv:2111.04327 [hep-ph]`:

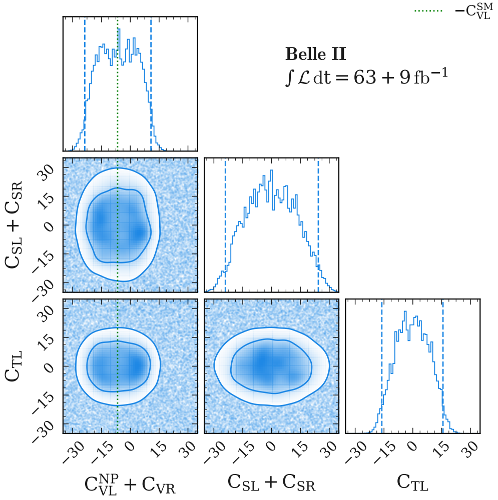
Operator	Value (paper) $[\mathrm{TeV}^{-2}]$	Value (EOS)	NP scale $[\mathrm{TeV}]$	Observable
$\mathcal{O}_{\nu d,\alpha\alpha sb}^{\mathrm{VL,NP}}$	0.028	17.2	6	$B \to K^* \nu\nu$
$\mathcal{O}_{\nu d,\alpha\alpha sb}^{\mathrm{VR}}$	0.021	12.9	7	$B \to K \nu\nu$
$\mathcal{O}_{\nu d,\gamma\delta sb}^{\mathrm{VL}}$	0.014	8.61	9	$B \to K^* \nu\nu$
$\mathcal{O}_{\nu d,\gamma\gamma sb}^{\mathrm{SL}}$	0.012	7.38	10	$B \to K^{(*)} \nu\nu$
$\mathcal{O}_{\nu d,\gamma\delta sb}^{\mathrm{SL}}$	0.009	5.54	10	$B \to K^{(*)} \nu\nu$
$\mathcal{O}_{\nu d,\gamma\delta sb}^{\mathrm{TL}}$	0.002	1.23	25	$B \to K^* \nu\nu$

Hence, choosing an upper bound of $C_{EOS} \leq 35$ completely covers branching ratio values of up to $Br \leq 1.1 \times 10^{-3}$.

Cross check: Bayesian sampling

As an alternative approach, we found exclusion limits in the space of Wilson coefficients by sampling random points in theory space.

- We performed a ML fit for each sample, with fixed signal strength.
- The likelihood is used as a weight for each sample.
- The dashed lines include the 95% central region for each distribution.
- The contours correspond to 68% (inner) and 95% (outer) intervals.

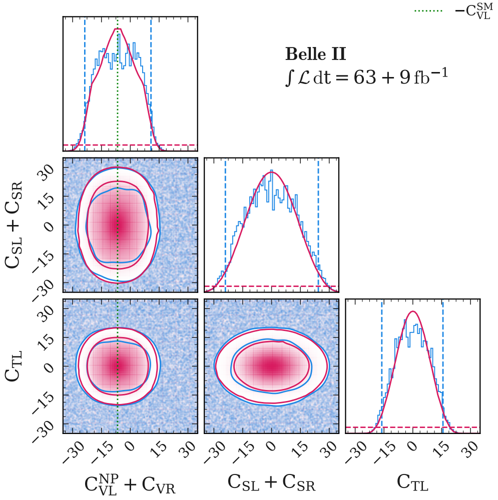


Contour overlay

We compare the two different methods by overlaying their contours.

- The 2d contours overlay very well.
- The 1d contours have a slightly different peak structure.
- The frequentist approach has slightly more conservative limits (1d distributions).

Since the exclusion criterium is different for the two cases, we should only compare the exclusion contours.



Hadronic parameters

Form factors are parameterized using the BCL parametrization

$$f_0\left(q^2\right) = \frac{\mathcal{L}}{1-\frac{q^2}{M^2_{B^*_{s0}}}} \sum_{n=0}^{N-1} a_n^0 z^n$$

$$f_+\left(q^2\right) = \frac{\mathcal{L}}{1-\frac{q^2}{M^2_{B^*_{s}}}} \sum_{n=0}^{N-1} a_n^+ \left(z^n - \frac{n}{N}(-1)^{n-N} z^N\right)$$

$$f_T\left(q^2\right) = \frac{\mathcal{L}}{1-\frac{q^2}{M^2_{B^*_{s}}}} \sum_{n=0}^{N-1} a_n^T \left(z^n - \frac{n}{N}(-1)^{n-N} z^N\right),$$

The correlation matrix between the hadronic parameters.

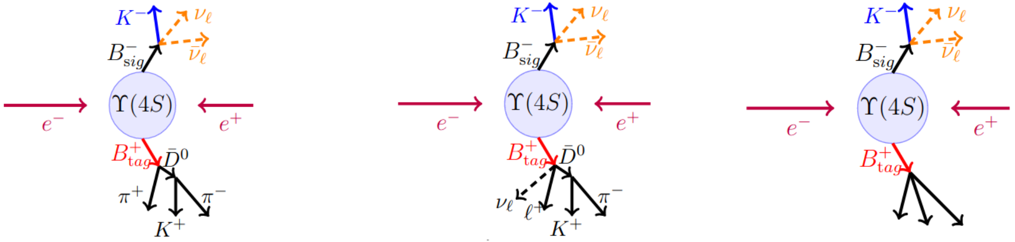
	a_0^+	a_1^+	a_2^+	a_1^0	a_2^0	a_0^T	a_1^T
a_0^+	1.00	0.67	0.33	0.94	0.83	0.43	0.34
a_1^+	0.67	1.00	0.86	0.73	0.70	0.22	0.39
a_2^+	0.33	0.86	1.00	0.39	0.41	0.04	0.26
a_1^0	0.94	0.73	0.39	1.00	0.96	0.40	0.37
a_2^0	0.83	0.70	0.41	0.96	1.00	0.34	0.33
a_0^T	0.43	0.22	0.04	0.40	0.34	1.00	0.89
a_1^T	0.34	0.39	0.26	0.37	0.33	0.89	1.00
a_2^T	0.21	0.42	0.44	0.23	0.23	0.71	0.91

Reconstruction techniques

Efficiency

$\epsilon \sim 0.1 - 1\,\%$ | $\epsilon \sim 1 - 3\,\%$ | $\epsilon \sim 1 - 100\,\%$

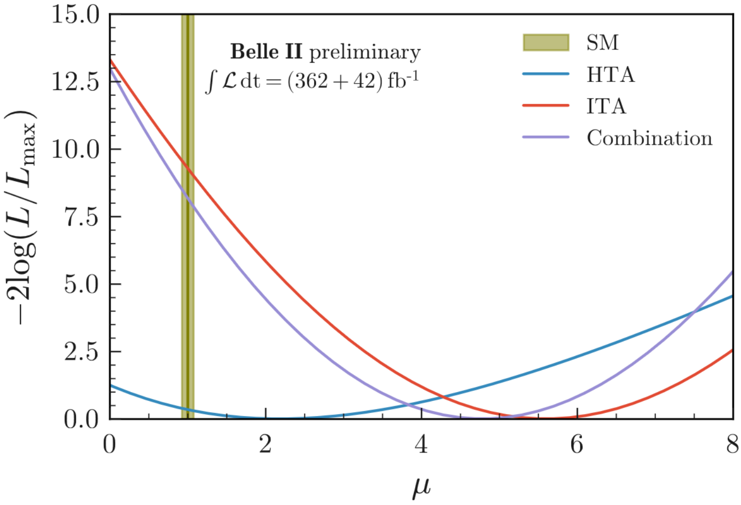
Exclusive hadronic | Exclusive semileptonic | Inclusive



Purity, Resolution

Different reconstruction techniques lead to nearly orthogonal data samples

Analysis update: $B^+ \to K^+\nu\bar{\nu}$ @ 362 fb^{-1}



pyhf

For observed event counts $\boldsymbol{n}$ the likelihood function is composed of

$$L(\boldsymbol{n},\boldsymbol{a}\mid\boldsymbol{\eta},\boldsymbol{\chi}) = \underbrace{\prod_{c\in\text{ channels}}\prod_{b\in\text{ bins}} \text{Pois}\left(n_{cb}\mid\nu_{cb}(\boldsymbol{\eta},\boldsymbol{\chi})\right)}_{\substack{\text{Simultaneous measurement}\\ \text{of multiple channels}}} \quad \underbrace{\prod_{\chi\in\boldsymbol{\chi}} c_{\chi}\left(a_{\chi}\mid\chi\right)}_{\substack{\text{constraint terms}\\ \text{for "auxiliary measurements"}}}$$

with free and constrained parameters $\boldsymbol{\eta},\boldsymbol{\chi}$, respectively,

$$L(\boldsymbol{x}\mid\phi) = L(\boldsymbol{x}\mid\overbrace{\boldsymbol{\eta}}^{\text{free}},\underbrace{\boldsymbol{\chi}}_{\text{constrained}}) = f(\boldsymbol{x}\mid\overbrace{\boldsymbol{\psi}}^{\text{parameters of interest}},\underbrace{\boldsymbol{\theta}}_{\text{nuisance parameters}})$$

The auxiliary measurements $\boldsymbol{a}$ are a frequentist approach to count modification.
The expected number of events for each channel and in each bin is

$$\nu_{cb}(\phi) = \sum_{s\in\text{ samples}}\nu_{scb}(\boldsymbol{\eta},\boldsymbol{\chi}) = \sum_{s\in\text{ samples}}\underbrace{\prod_{\kappa\in\boldsymbol{\kappa}}\kappa_{scb}(\boldsymbol{\eta},\boldsymbol{\chi})}_{\text{multiplicative modifiers}}\left(\nu^{0}_{scb}(\boldsymbol{\eta},\boldsymbol{\chi}) + \underbrace{\sum_{\Delta\in\boldsymbol{\Delta}}\Delta_{scb}(\boldsymbol{\eta},\boldsymbol{\chi})}_{\text{additive modifiers}}\right).$$

Modifiers and constraints

Description	Modification	Constraint Term c_χ	Input
Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Pois}\left(r_b = \sigma_b^{-2} \middle| \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
Correlated Shape	$\Delta_{scb}(\alpha) = f_p\left(\alpha \middle| \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\Delta_{scb,\alpha=\pm 1}$
Normalisation Unc.	$\kappa_{scb}(\alpha) = g_p\left(\alpha \middle| \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\mathrm{Gaus}\left(a = 0 \middle| \alpha, \sigma = 1\right)$	$\kappa_{scb,\alpha=\pm 1}$
MC Stat. Uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \mathrm{Gaus}\left(a_{\gamma_b} = 1 \middle| \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\mathrm{Gaus}\left(l = \lambda_0 \middle| \lambda, \sigma_\lambda\right)$	$\lambda_0, \sigma_\lambda$
Normalisation	$\kappa_{scb}(\mu_b) = \mu_b$		
Data-driven Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$		