# Global Fit Strategy for Inclusive $B \to X L$


Frank Tackmann

Deutsches Elektronen-Synchrotron


CKM 2016 workshop, TIFR Mumbai
November 30, 2016

# Status Overview.

## Inclusive $|V_{cb}|$ [see previous talk]

- Current global moment fits are dominated by theory uncertainties, and in particular theory correlations
- Goal for Belle II will really be to reduce the current uncertainty (50-100%) on the uncertainty ($\sim$ 2%)

## Inclusive $|V_{ub}|$ [see Bob's and Paolo's talks yesterday]

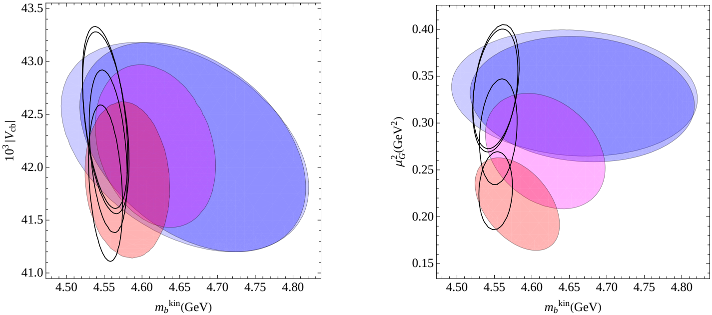
- Current $\sim$ 5-7% uncertainties are probably underestimated (which contributes to the tension with excl. $|V_{ub}|$)
- Current methods do not extrapolate to 3% total uncertainty, need qualitative improvements to get there

Both are (or will be) theory limited, but not in a way that more calculations alone will help

- Overall only little room for improvement in perturbative inputs
- Parametric uncertainties dominate, require coordinated effort between theory and experiment

# Theory Uncertainties in Inclusive $|V_{cb}|$

- For a given moment $M_i(E_{\rm cut})$, results with different lepton energy cut $E_\ell > E_{\rm cut}$ are strongly correlated
  (most events are the same, so they have a high statistical correlation)
  - The independent new information the fit sees is really in the differences $M_i(E_{\rm cut} + 100\,{\rm GeV}) - M_i(E_{\rm cut})$
  - The theory uncertainty on this difference is however never directly evaluated, but only follows indirectly from the assumed correlation for different $E_{\rm cut}$
  - Not surprising that resulting uncertainties (in particular for OPE parameters) strongly depend on theory correlation assumption [Gambino, Schwanda, 1307.4551]



# Theory Uncertainties in Inclusive $|V_{cb}|$.

- Different lepton energy moments at different $E_{\mathrm{cut}}$ are also not independent
  - They all come from the same underlying lepton energy spectrum.
  - For example, higher $E_\ell$ moments are sensitive to in principle the same high-$E_\ell$ information as the rate with high $E_{\mathrm{cut}}$
  - Currently, different moments are assumed to be completely uncorrelated.

Various ways more data can help the theory

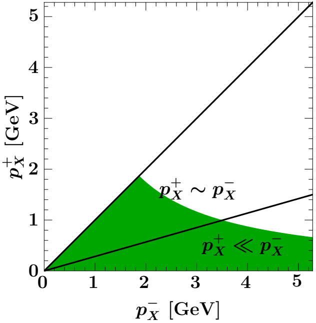
- Measure mass moments in bins of $E_\ell$ instead of different $E_{\mathrm{cut}}$
- Directly measure the lepton energy spectrum as precisely as possible and all the way to the endpoint.
  - Potential avenue to provide important nontrivial constraint on shape function using precise $b \to c$ data (beyond $m_b$ and $\lambda_1$ constraints)
- Also measure the $q^2$ spectrum, $E_\nu$ spectrum, $q^0 = E_\ell + E_\nu$ spectrum
- Could also think about performing helicity decomposition and directly measuring independent hadronic structure functions (e.g. mass moments)

$\Rightarrow$ Belle 2 can improve incl. $|V_{cb}|$ (perhaps contrary to common believe)

# Phase Space Regions for $|V_{ub}|$

Measurements probe different phase-space regions
[see also Bob's talk yesterday]

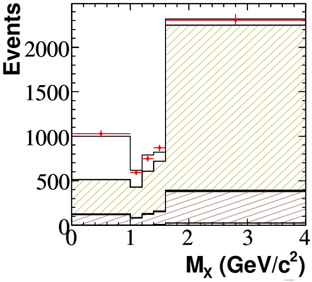
- SF region: $p_X^+ \ll p_X^-$ (large $E_\ell$, $E_\gamma$)
  - Experimentally clean(er) and highest sensitivity
  - Theory more difficult
- Local OPE: $p_X^+ \sim p_X^-$ ($q^2$ spec., small $E_\ell$, $E_\gamma$)
  - Large backgrounds, least sensitivity
  - Theoretically easier
- Something in between: $m_X \sim m_D$, moderately large $E_\ell$ ($E_\gamma$)




There is no single "golden" region

- There are no "optimal" cuts
  - Measuring deep into $b \to c$ hides the issue (e.g. in the MC signal model), dominant sensitivity still comes from region with least $B \to X_c \ell \nu$ background

$\Rightarrow$ Do not choose, measure full spectrum, which gives most information (same for $B \to X_s \gamma$)



# Global Fit Approach for $|V_{ub}|$

Follow same basic strategy as for $|V_{cb}|$ (just more complicated now)

- Simultaneously determine from the data
  - Overall normalization: $|V_{ub}|$, $\mathcal{B}(B \to X_s\gamma)$
  - Input parameters and their uncertainties: $m_b$, shape function(s)
- Combine different decay modes and measurements
  - Different $B \to X_u\ell\nu$ and $B \to X_s\gamma$ spectra
  - Can eventually include/predict also $B \to X_s\ell^+\ell^-$
  - External constraints on $m_b$, $\mu_\pi^2$ ($\lambda_1$), from $B \to X_c\ell\nu$ or elsewhere

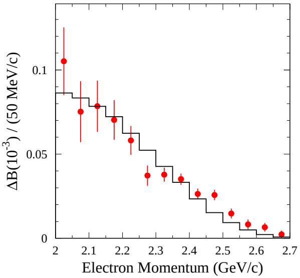
What it achieves

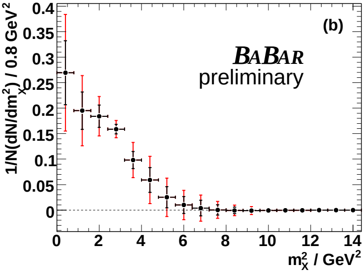
- Minimize uncertainties by making maximal use of all available data
  - Fit automatically "chooses" most sensitive region given exp. and theory unc.
- Qualitatively better inclusive $|V_{ub}|$ with consistent treatment of correlated uncertainties (experimental, theoretical, input parameters)

What it requires

- Experiment: Precise spectra, including full correlations
- Theory: Proper theory description across phase space, model-independent treatment of shape function

# Experimental Side of Global $|V_{ub}|$ Fit




Global fit approach will become very powerful with high statistics

- Measure as many spectra as precisely as possible to maximize the available shape information
  - Detailed shape information is key to constraining subleading corrections
  - Analyses need input on shape in any case, e.g. to improve signal MC
  - $E_\ell$ spectrum (in bins of $m_X$), (high) $q^2$, $m_X$, $p_X^+$ (all in bins of $E_\ell$?)
  - Separate $B^+$ and $B^0$
- Take advantage of large datasets to maximize resolution and to agressively reject backgrounds at the cost of efficiency
  - super-clean full reconstruction sample

# Theory Side of Global $|V_{ub}|$ Fit.

SIMBA [Bernlochner, Lacker, Ligeti, Stewart, FT, K Tackmann, arXiv:1303.0958]

- Global fit combining all available information
- Employs model-independent treatment for SF

  [Ligeti, Stewart, FT, arXiv:0807.1926]

$$\mathrm{d}\Gamma_s = |V_{tb}V_{ts}^*|^2 m_b^2 \left|C_7^{\mathrm{incl}}\right|^2 \int \mathrm{d}k\, \widehat{W}_{77}(E_\gamma; k)\, \widehat{F}(m_B - 2E_\gamma - k) + \cdots$$

$$\mathrm{d}\Gamma_u = |V_{ub}|^2 \int \mathrm{d}k\, \widehat{W}_u(p_X^-, p_X^+, E_\ell; k)\, \widehat{F}(p_X^+ - k) + \cdots$$

- Fit parameters: $|V_{tb}V_{ts}^*|^2 m_b^2$, $|V_{ub}|^2$, $\widehat{F}(\lambda x) = \frac{1}{\lambda}\left[\sum_{n=0}^{\infty} c_n f_n(x)\right]^2$
- Theory input: $\widehat{W}_i(\ldots; k)$ computed to (N)NLL′+(N)NLO in 1S scheme

NNVub [Healey, Mondino, Gambino, arXiv:1604.07598]

- Based on same idea, quite different approach [see Paolo's talk yesterday]

# Factorized Shape Function.

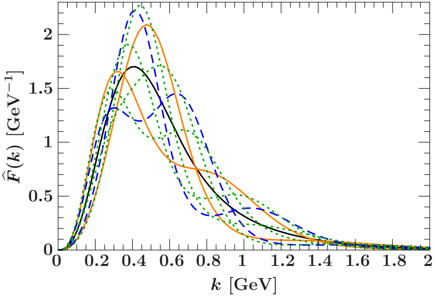
$$S(\omega, \mu_\Lambda) = \int \mathrm{d}k\, \widehat{C}_0(\omega - k, \mu_\Lambda)\, \widehat{F}(k)$$

$\widehat{F}(k)$ nonperturbative part

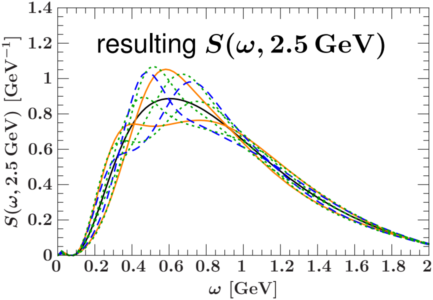
- Determines peak region
- Fit from data

$\widehat{C}_0(\omega, \mu_\Lambda)$ perturbative part

- Generates perturbative tail with correct $\mu_\Lambda$ dependence




$\Rightarrow$



# Factorized Shape Function.

$$S(\omega, \mu_\Lambda) = \int \mathrm{d}k\, \widehat{C}_0(\omega - k, \mu_\Lambda)\, \widehat{F}(k)$$

$\widehat{F}(k)$ nonperturbative part

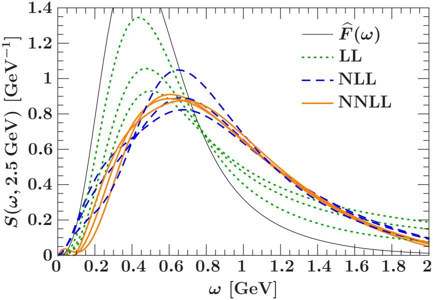
- Determines peak region
- Fit from data

$\widehat{C}_0(\omega, \mu_\Lambda)$ perturbative part

- Generates perturbative tail with correct $\mu_\Lambda$ dependence

- Given $\widehat{F}(k)$ we can *compute* $S(\omega, \mu_\Lambda)$ in perturbation theory
  - Vary $\mu_\Lambda$ to estimate perturbative uncertainty in SF

$\Rightarrow$ Instead of modeling $S(\omega, \mu_\Lambda)$ we can fit for $\widehat{F}(k)$



# Basis Expansion for $\widehat{F}(k)$.

Expand $\widehat{F}(k)$ into suitable orthonormal basis

$$\widehat{F}(\lambda x) = \frac{1}{\lambda}\Big[\sum_{n=0}^{\infty} c_n f_n(x)\Big]^2$$

$$\int \mathrm{d}k\, \widehat{F}(k) = \sum_{n=0}^{\infty} c_n^2 = 1$$

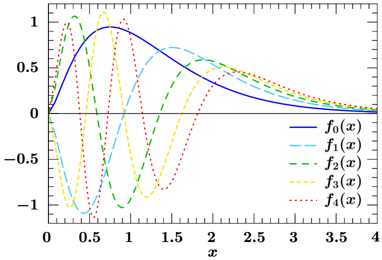
- Provides model-independent description

Fit for $\widehat{F}(k)$ by fitting basis coefficients $c_n$

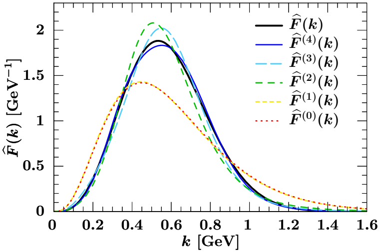
- Experimental uncertainties and correlations can be properly captured in covariance matrix of fitted coefficients $c_n$

⇒ Allows for *data-driven*, reliable estimation of SF uncertainties

Basis functions




Expansion of Gaussian $\widehat{F}(k)$



# Residual Basis Dependence from Series Truncation.

$$\widehat{F}(\lambda x) = \frac{1}{\lambda}\left[\sum_{n=0}^{N} c_n f_n(x)\right]^2$$

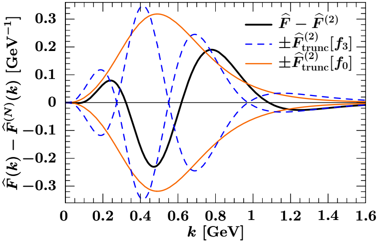
In practice, series must be truncated

- Induces residual basis (model) dependence
- Truncation error scales as $1 - \sum_{n=0}^{N} c_n^2$

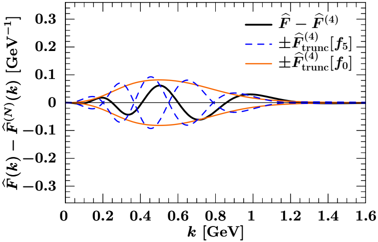
In practice most complications are in choosing good basis ($\lambda$) and $N$

- Want basis so series converges quickly but still unbiased (e.g. iterate)
- Choose $N$ large enough so truncation error is smaller than to exp. uncertainties, but small enough to have stable fit and not waste statistical power
- Add coefficients with more precise data

Truncation error at $N = 2$




Truncation error at $N = 4$



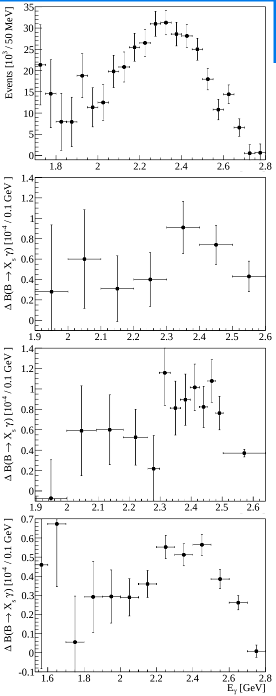
# Global Fit to $B \to X_s\gamma$

## Theory

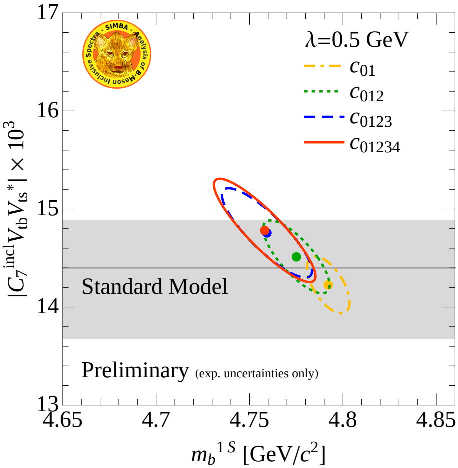
- NNLL′+NNLO
- non-$C_7$ contributions from SM

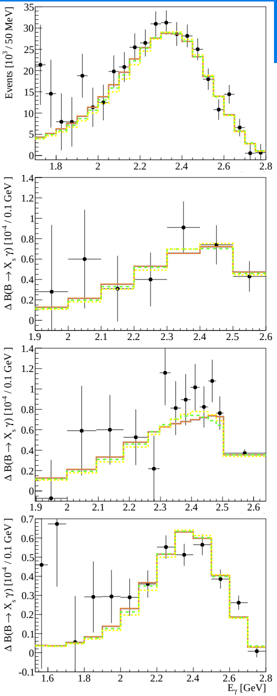
## Experimental Inputs

- Belle inclusive (in $\Upsilon(4S)$ frame)
  [arXiv:0907.1384]
- BaBar hadronic tag (in $B$ frame)
  [arXiv:0711.4889]
- (old) BaBar sum-over-exclusive (in $B$ frame)
  [hep-ex/0508004]
- BaBar inclusive (in $\Upsilon(4S)$ frame)
  [arXiv:1207.5772]



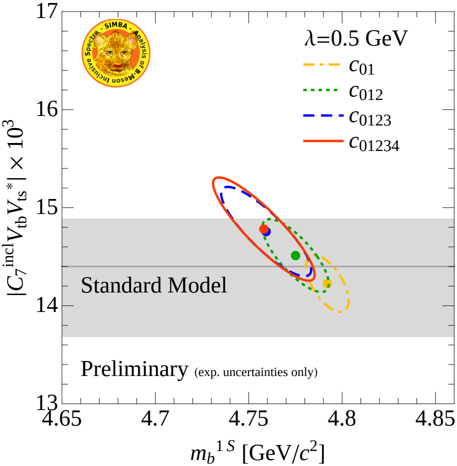
# Global Fit to $B \to X_s\gamma$

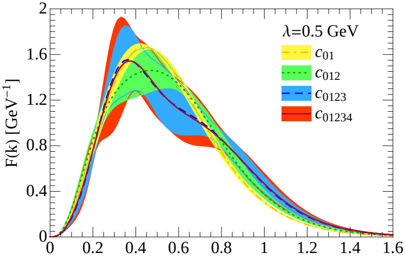


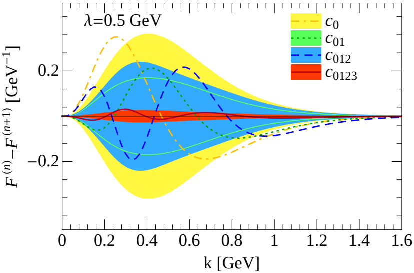




- Too few coefficients lead to clear bias and underestimated uncertainties
- Extracted $|C_7^{\rm incl} V_{tb} V_{ts}^*|$ consistent with SM

# Global Fit to $B \to X_s\gamma$



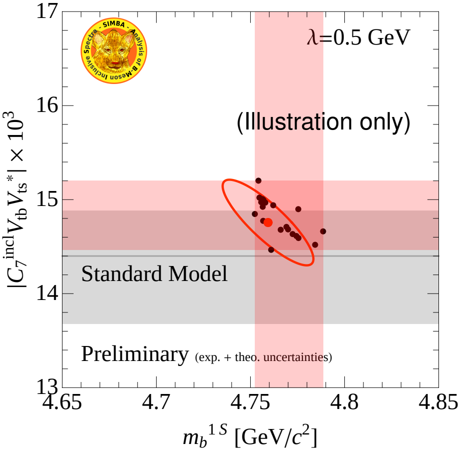




- Too few coefficients lead to clear bias and underestimated uncertainties
- Extracted $|C_7^{\rm incl} V_{tb} V_{ts}^*|$ consistent with SM

# Work in Progress.

## Perturbative uncertainties

- Dominant source of theory uncertainties
- Important to take into account correlations across spectrum
  - Integrating resummed NNLL′+NNLO spectrum should reproduce smaller uncertainties in total NNLO rate (quite nontrivial)
- Evaluated via large set of profile scale variations
- Expect theory uncertainties of comparable size to fit uncertainties

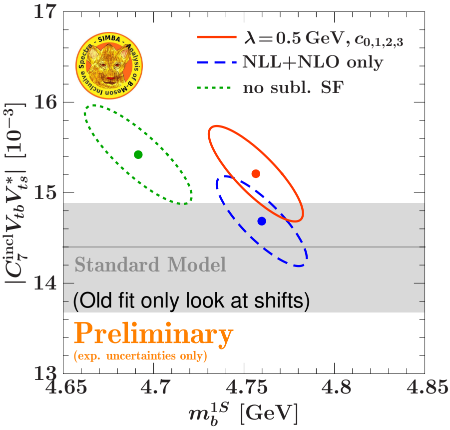


# Work in Progress.

## Subleading shape functions

- $|C_7|^2$-like ones can be absorbed into leading SF
  - Irrelevant for fit
  - Important for interpretation: Cause substantial shift in $m_b$ given by their total 1st moment

$$\frac{-\lambda_1 + 3\lambda_2}{2m_b} \sim 70\,\mathrm{MeV}$$




- Four-quark shape functions:
  - Formally $\alpha_s/m_b$ suppressed
  - Do not find large ($\sim 5\%$) effects as in [Benzke, Lee, Neubert, Paz, arXiv:1003.5012]
  - Dominant effect from $O_{1,2}O_7$ interference ($c\bar{c}$ loops) can be included via single subleading shape function → only minor effect on fit

# Work in Progress.

## Consistent treatment of charm contributions

- Integrate out charm loops ($n_f = 3$) vs. keeping charm dynamic ($n_f = 4$)
- Include known massive results
- In the end small effect (most $m_c$ dependence is absorbed into $C_7^{\mathrm{incl}}$)

## (Indirect) dependence on $\lambda_2$ and $\rho_2$

- Only small effect on uncertainties

## Fit strategy

- Study systematically sensitivity to basis choice and truncation
- Study dependence on experimental inputs

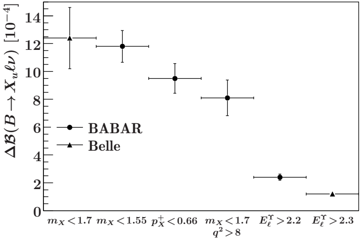
# Global $|V_{ub}|$ Fit.

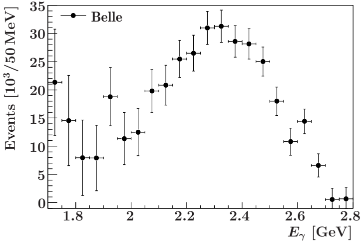
## Theory

- NLL′+NLO
- ignoring subleading SFs

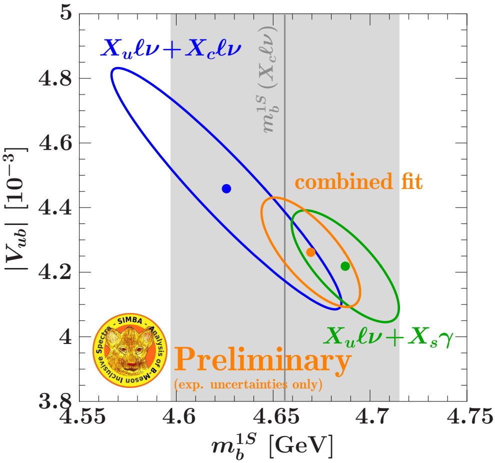
## Experimental Inputs

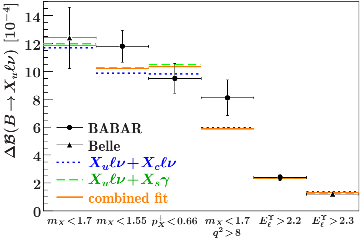
- $B \to X_u \ell \nu$ partial branching fractions
  - picked measurements for which we are sure enough that they have negligible (SF) model dependence
  - BaBar and Belle hadronic tag
  - BaBar and Belle lepton endpoint
- $B \to X_s \gamma$
  - Belle inclusive (shown)
  - (old) BaBar sum-over-exclusive (not shown)
  - BaBar hadronic tag (not shown)
- $B \to X_c \ell \nu$
  - $m_b^{1S}$, $\lambda_1$ from moment fits

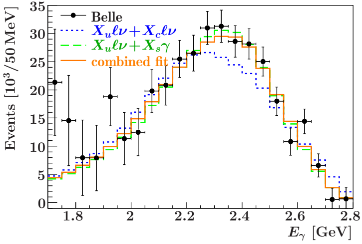


# Global $|V_{ub}|$ Fit.




- Parametric unc. (SF, $m_b$) are part of fit, no pert. unc. included yet
- Without full $B \to X_s\gamma$: $\sim$ 10% uncertainties on $|V_{ub}|$
- Including $B \to X_s\gamma$: halves uncertainties but also shifts $|V_{ub}|$

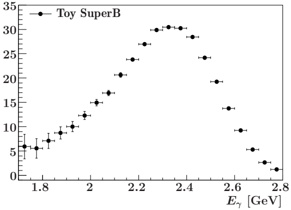
# Projections for Belle 2.

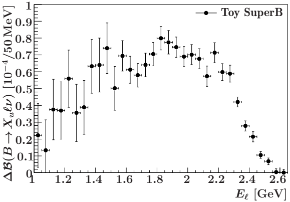
## Theory

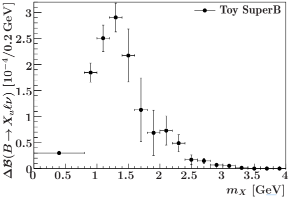
- NLL′+NLO
- ignoring subleading SFs

## Toy study

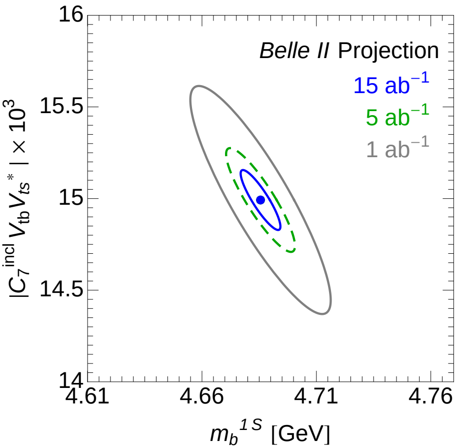
- Generated $m_X$, $E_\ell$, and $E_\gamma$ spectra from theory
- Smeared with uncertainties and correlations inspired by BaBar hadronic tag analysis, Belle 2 hadronic tagging efficiency is much better by now
- Originally for 75/ab, scaled down to 1/ab, 5/ab, 15/ab
- Caveats:
  - No resolution effects considered
  - Should be done more thoroughly by Belle 2



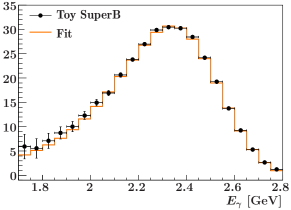



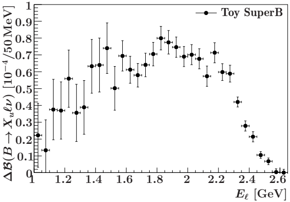



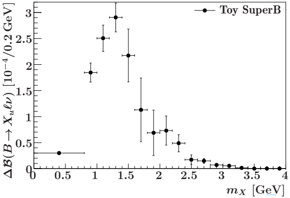
# Projections for Belle 2.

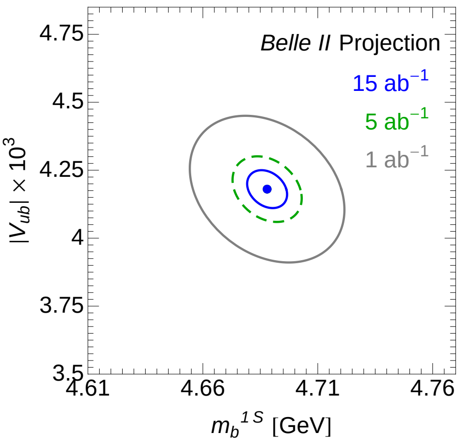


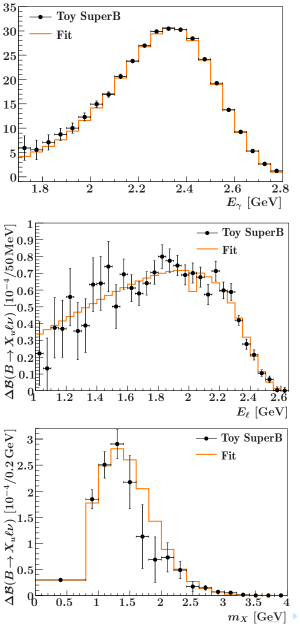

- No perturbative uncertainties included
  (but they clearly won't scale with statistics)



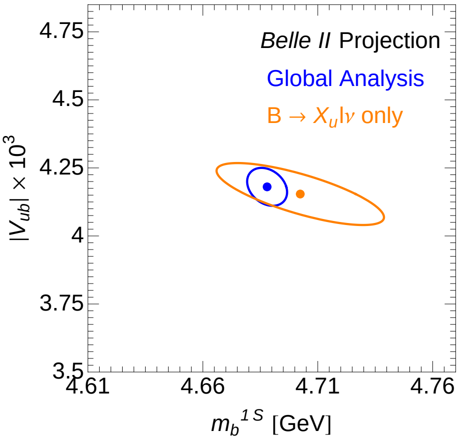



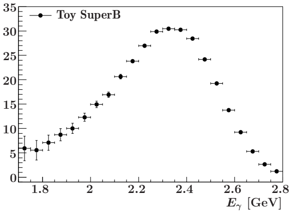


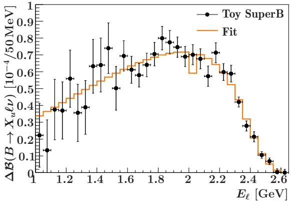



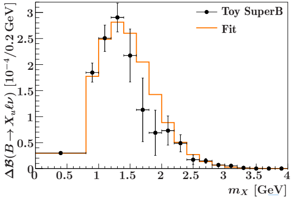

- No perturbative uncertainties included
  (but they clearly won't scale with statistics)

# Projections for Belle 2.



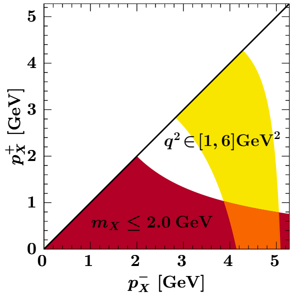




- No perturbative uncertainties included (but they clearly won't scale with statistics)
- At Belle 2 can use $B \to X_u \ell\nu$ alone to determine SF, $m_b$, and $|V_{ub}|$

# Combined analysis of $B \to X_s \ell\ell$ and $B \to X \ell\nu$.

## Experimental kinematic cuts for $B \to X_s \ell\ell$

- $1 < q^2 < 6\,\mathrm{GeV}^2$, $m_X < m_X^{\mathrm{cut}} \sim 2\,\mathrm{GeV}$
- Unavoidable to suppress huge $b \to c\ell^-\bar{\nu} \to s\ell^+\ell^-\nu\bar{\nu}$ background
- Shape function effects must be taken into account to retain NP sensitivity

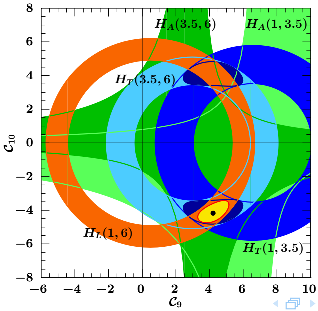



## Helicity decomposition for inclusive rate

[Lee, Ligeti, Stewart, FT (2008)]

$$\frac{\mathrm{d}^3\Gamma}{\mathrm{d}p_X^+\mathrm{d}p_X^-\mathrm{d}z} = \frac{3}{8}\Big[(1+z^2)H_T(p_X^\pm) + 2zH_A(p_X^\pm) + 2(1-z^2)H_L(p_X^\pm)\Big]$$

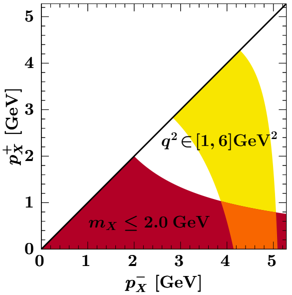
- $z = \cos\theta = 2\,\frac{E_\ell - E_{\bar{\ell}}}{p_X^- - p_X^+}$ is angle between lepton and $B$ meson in $W$ rest frame



# Combined analysis of $B \to X_s \ell\ell$ and $B \to X \ell\nu$.

## Experimental kinematic cuts for $B \to X_s \ell\ell$

- $1 < q^2 < 6\,\mathrm{GeV}^2$, $m_X < m_X^{\mathrm{cut}} \sim 2\,\mathrm{GeV}$
- Unavoidable to suppress huge $b \to c\ell^- \bar\nu \to s\ell^+\ell^-\nu\bar\nu$ background
- Shape function effects must be taken into account to retain NP sensitivity




## Same basic structure

$$\mathrm{d}H_{T,A,L} = \sum_{ij} C_i^{\mathrm{incl}} C_j^{\mathrm{incl}} \int \mathrm{d}k\, \widehat{W}_{ij}^{A,T,L}(p_X^+, E_\ell, E_{\bar\ell}; k)\, \widehat{F}(p_X^+ - k) + \cdots$$

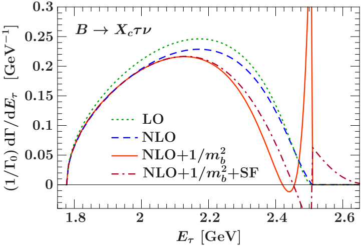
## Combined fit of $B \to X_s \ell\ell$ and $B \to X \ell\nu$

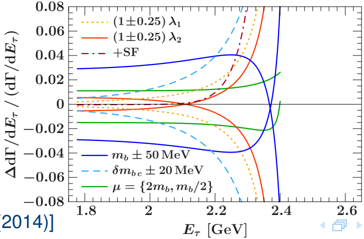
- Best (perhaps only) way to get clean extraction of $C_9^{\mathrm{incl}}$, $C_{10}^{\mathrm{incl}}$
- Using the same inclusive helicity decomposition for $B \to X \ell\nu$ would allow to fully disentangle dominant subleading SF effects

# Combination with $B \to X_c \ell \nu$.

## Combined analysis of $B \to X_u \ell \nu$ and $B \to X_c \ell \nu$

- Measure very precise lepton energy spectrum
- Allows for fully consistent and correlated treatment of both channels
  - Moves separation between the modes from analysis level to interpretation step
  - Can constrain leading SF from $b \to c$
- Combined fit to directly extract $|V_{ub}/V_{cb}|$




## $B \to X \tau \nu$ and $R(X)$

- Belle 2 should obviously measure $R(X)$
  - If $R(D^{(*)})$ is due to new physics it must show up in $R(X)$
- Theory for inclusive decay is as clean
- Combined analysis of $B \to X \ell \nu$ and $B \to X \tau \nu$ to measure $R(X)(q^2)$

[Ligeti, FT (2014)]

# Summary.

- Inclusive $|V_{cb}|$ and $|V_{ub}|$ with current approaches are theory limited, but not in a way that more calculations alone will help
- Strategy for Belle 2 should be to exploit increased data sets to help theory by providing maximal amount of information in the form of differential and as model-independent as possible measurements
- Global fit to inclusive rare and semileptonic data with model-independent treatment of shape function will be key to reach ultimate precision for inclusive $|V_{ub}|$
- Global analysis will also be essential to fully exploit new-physics sensitivity of inclusive $B \to X_s \gamma$ and $B \to X_s \ell\ell$