$B \to X_s\gamma$ and $|V_{ub}|$ from SIMBA

Zoltan Ligeti

- Model independent shape function treatment
- Results: $B \to X_s\gamma$

  (F. Bernlochner, H. Lacker, ZL, I. Stewart, F. Tackmann, K. Tackmann,

  PRL 127 (2021) 10, 102001 [2007.04320])
- Ongoing: $B \to X_u\ell\bar{\nu}$
- Conclusions

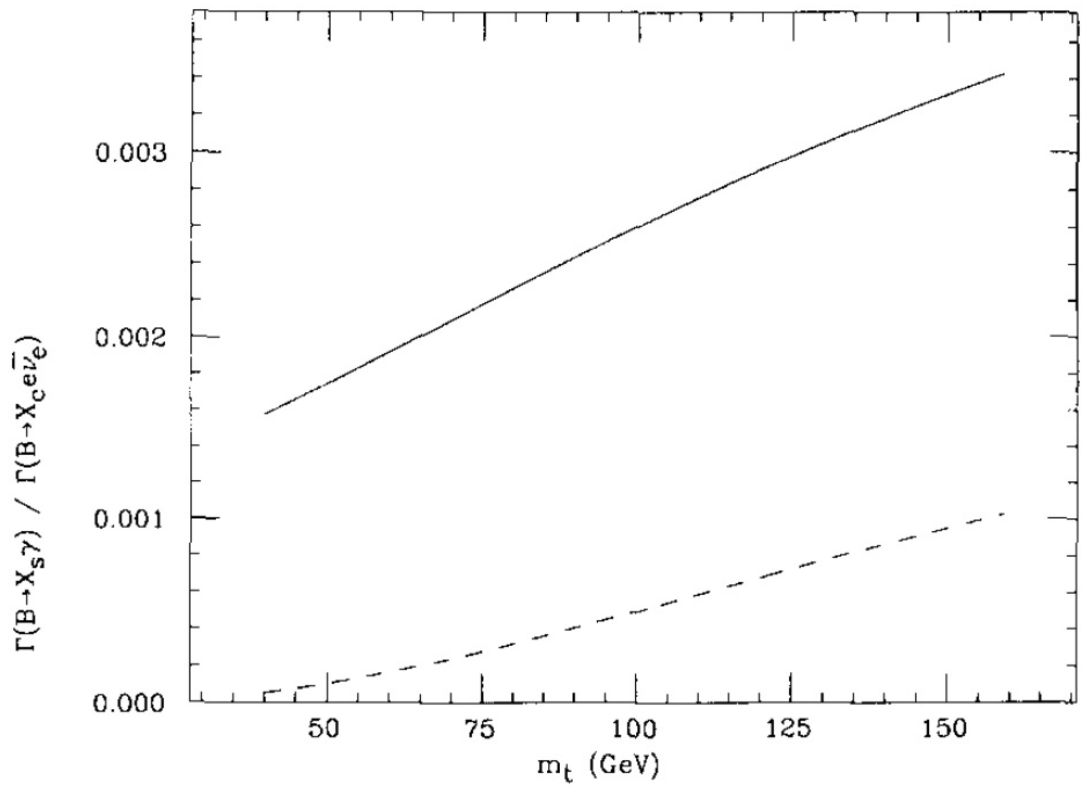
Introduction

- Decades of sophisticated efforts: impacted multi-loop techniques, HQE, BSM

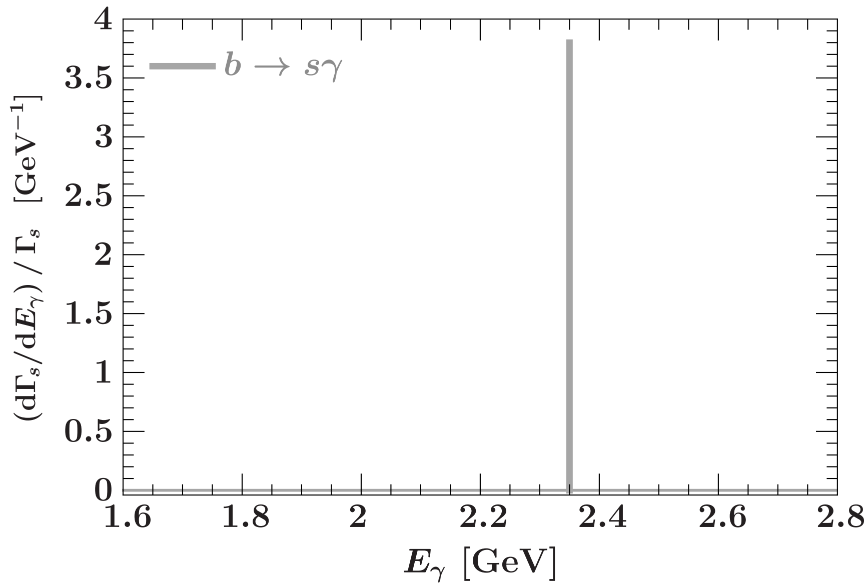
EFFECTIVE HAMILTONIAN FOR WEAK RADIATIVE B-MESON DECAY ☆

Benjamin GRINSTEIN [1], Roxanne SPRINGER and Mark B. WISE [2]
California Institute of Technology, Pasadena, CA 91125, USA

Received 18 November 1987



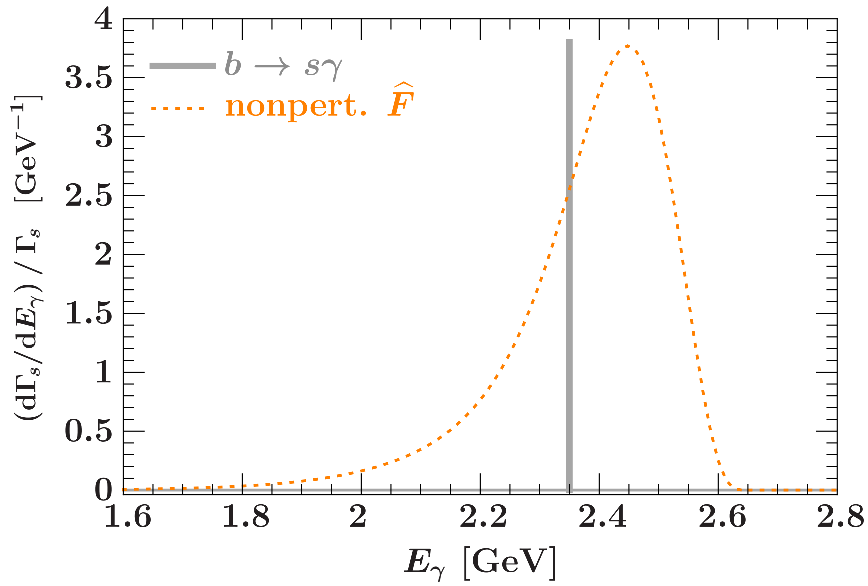
The $B \to X_s\gamma$ challenge at/below the scale m_b



[fig credit: F. Tackmann]

Parton level: $\mathrm{d}\Gamma/\mathrm{d}E_\gamma = |C_7|^2\delta(E_\gamma - m_b/2)$

The $B \to X_s\gamma$ challenge at/below the scale m_b



[fig credit: F. Tackmann]

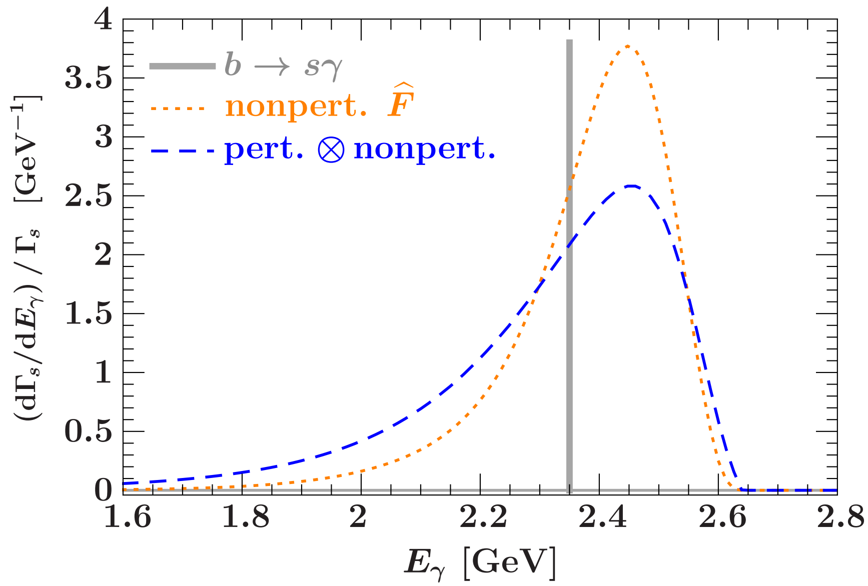
Parton level: $\mathrm{d}\Gamma/\mathrm{d}E_\gamma = |C_7|^2\,\delta(E_\gamma - m_b/2)$

Hadron level: spectrum determined by nonpert. b-quark distribution function

BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

The $B \to X_s\gamma$ challenge at/below the scale m_b



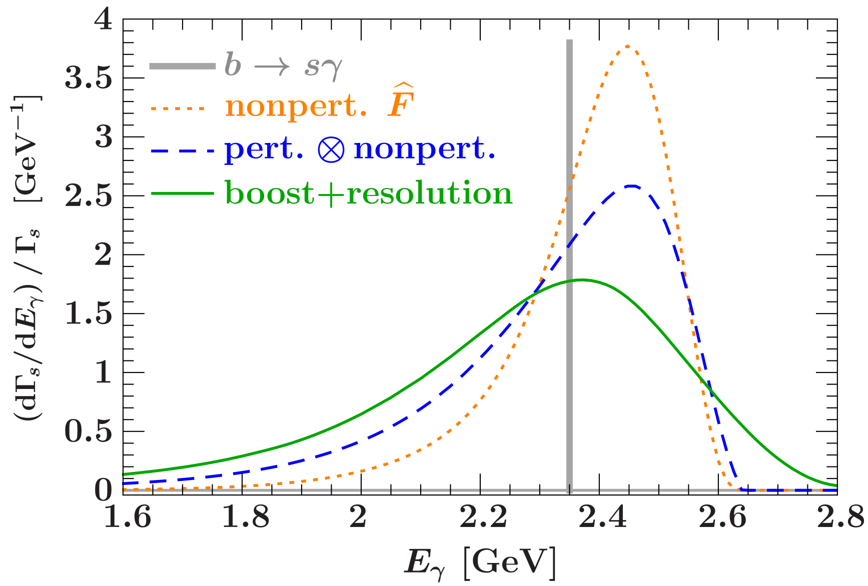
[fig credit: F. Tackmann]

Parton level: $\mathrm{d}\Gamma/\mathrm{d}E_\gamma = |C_7|^2\,\delta(E_\gamma - m_b/2)$

Hadron level: spectrum determined by nonpert. b-quark distribution function

Small E_γ tail (and integral down to it) mostly perturbative

The $B \to X_s\gamma$ challenge at/below the scale m_b



[fig credit: F. Tackmann]

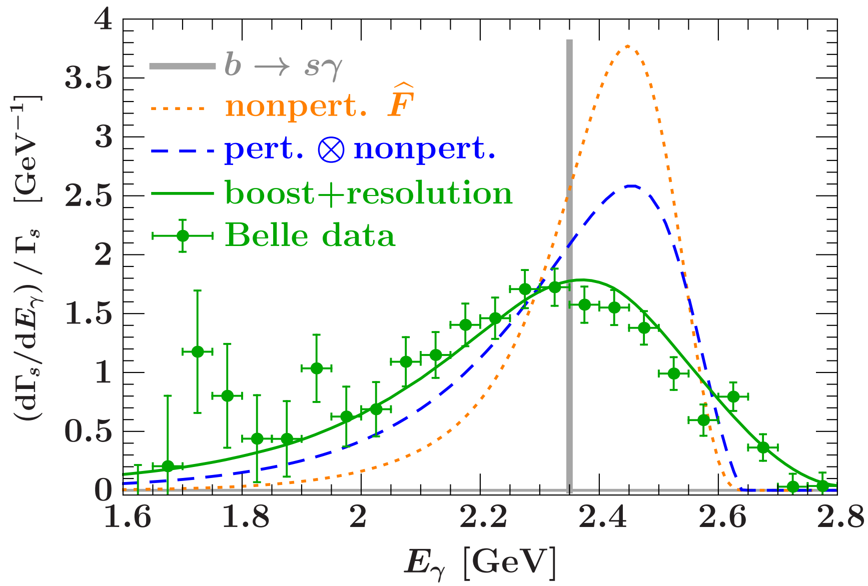
Parton level: $\mathrm{d}\Gamma/\mathrm{d}E_\gamma = |C_7|^2\,\delta(E_\gamma - m_b/2)$

Hadron level: spectrum determined by nonpert. b-quark distribution function

Small E_γ tail (and integral down to it) mostly perturbative

Further smeared by B boost from $\Upsilon(4S)$ decay and resolution

The $B \to X_s\gamma$ challenge at/below the scale m_b



[fig credit: F. Tackmann]

Parton level: $d\Gamma/dE_\gamma = |C_7|^2\,\delta(E_\gamma - m_b/2)$

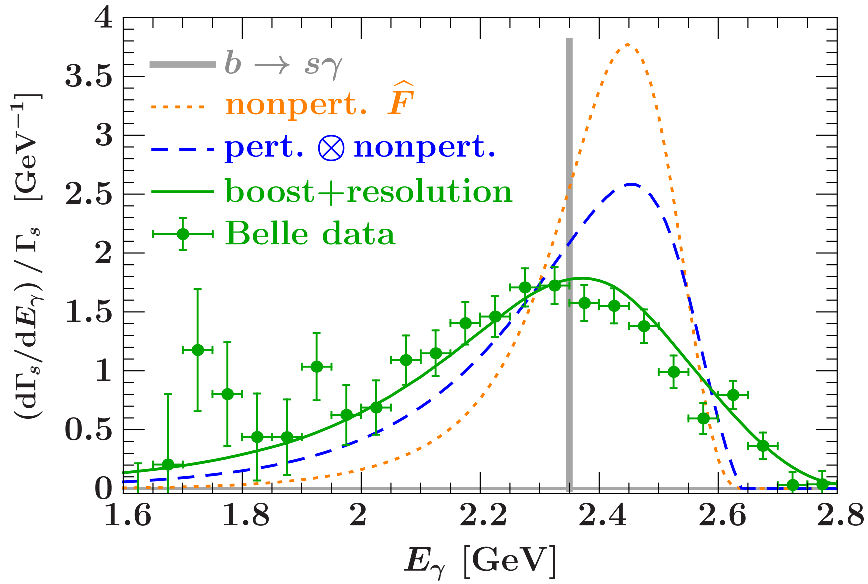
Hadron level: spectrum determined by nonpert. b-quark distribution function

Small E_γ tail (and integral down to it) mostly perturbative

Further smeared by B boost from $\Upsilon(4S)$ decay and resolution

Experimental data most precise in the peak region

Traditional approach to $B \to X_s\gamma$



[fig credit: F. Tackmann]

Past: compare calculations and data for integrated rates, e.g., $\mathcal{B}(E_\gamma > 1.6\,\text{GeV})$

- Data most precise in peak region, this precision is not fully exploited
- Integrating to lower E_γ reduces theory uncertainties but increases experimental ones
- Extrapolation from higher to lower cuts induces hard-to-quantify model dependence

Features & goals of SIMBA

- Optimally combine all measurements (consistently treat uncertainties & correlations)
- Theory:
  - Consistent theory description across E_γ spectrum
  - Model-independent treatment of shape function(s)
- Data:
  - Utilize all $B \to X_s\gamma$ ($B \to X_u\ell\bar{\nu}$) spectra or partial rates
  - Include other constraints on m_b, λ_1, etc.
- Simultaneously determine:
  - Normalization sensitive to short-distance physics: $|C_7^{\rm incl}|$, $|V_{ub}|$
  - Nonperturbative parameters: m_b, shape function(s)
- Same strategy as for inclusive $|V_{cb}|$, just a lot more complicated...

Theory ingredients

Regions of $B \to X_s\gamma$ photon spectrum

- Peak around $E_\gamma \sim 2.3\,\mathrm{GeV}$ $\quad (m_B - 2E_\gamma \sim 0.8\,\mathrm{GeV})$

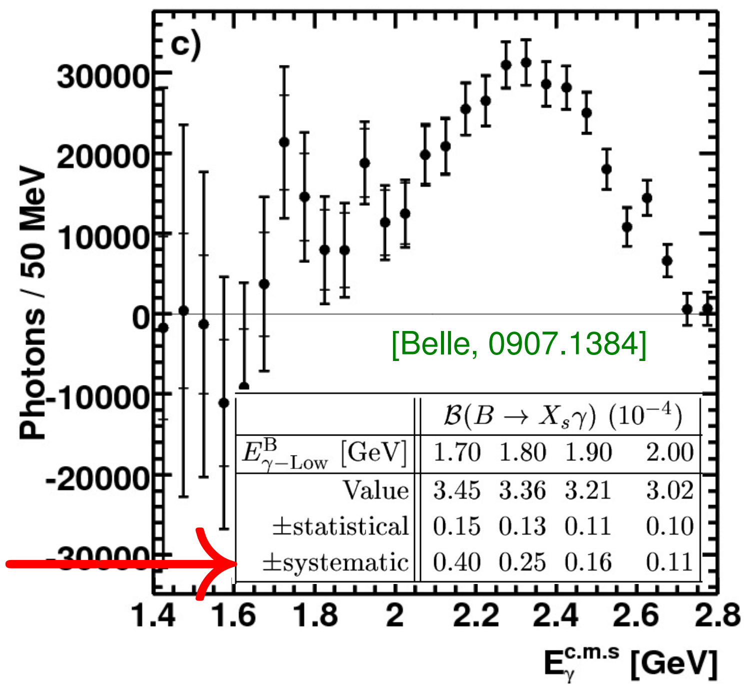
  Three cases: 1) $\Lambda_{\rm QCD} \sim m_B - 2E_\gamma \ll m_B$ ["SCET"]

  2) $\Lambda_{\rm QCD} \ll m_B - 2E_\gamma \ll m_B$ ["MSOPE"]

  3) $\Lambda_{\rm QCD} \ll m_B - 2E_\gamma \sim m_B$

  Expansions and theory uncertainties differ in the 3 regions

  Neither 1) nor 2) is fully appropriate



- Rapid increase of exp. systematic error for smaller $E_\gamma^{\rm cut}$

- At tree level: $\delta(E_\gamma - m_b/2) \;\to\; S(E_\gamma - m_b/2)$

  where $S(\omega) = \langle B|\, \bar b\, \delta(\omega - iD_+)\, b\, |B\rangle$

  Moments of $S(\omega)$ given by local operators' matrix elements $\{1,\, 0,\, -\frac{1}{3}\lambda_1,\, -\frac{1}{3}\rho_1,\, ...\}$

The shape function (b quark PDF in B)

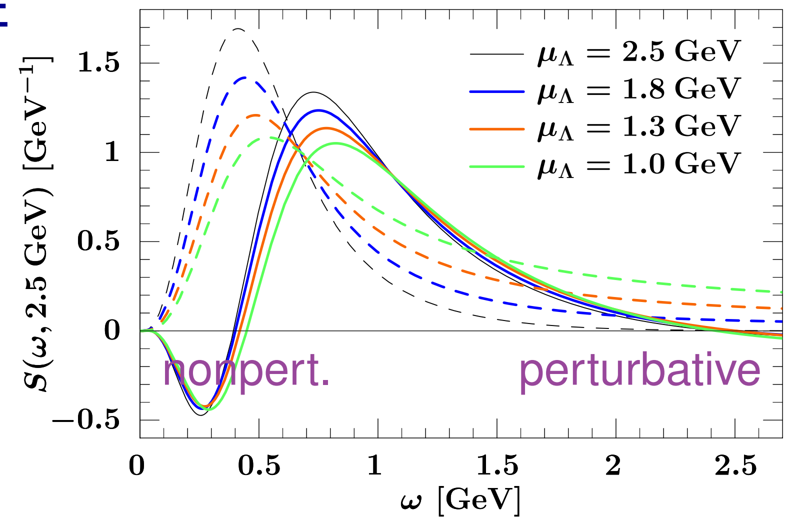
- The shape function $S(\omega,\mu)$ contains nonperturbative physics and obeys an RGE

  Even if $S(\omega,\mu_\Lambda)$ has exponentially small tail, RGE running gives long tail and divergent moments

  $$S(\omega,\mu_i) = \int \mathrm{d}\omega'\, U_S(\omega-\omega',\mu_i,\mu_\Lambda)\, S(\omega',\mu_\Lambda)$$

  [Balzereit, Mannel, Kilian]

  Constraint: moments (OPE) $+\ B \to X_s\gamma$ shape



Model $\begin{cases} S & \text{(dash)} \\ F & \text{(solid)} \end{cases}$ run to $2.5\mathrm{GeV}$

- Derive: $S(\omega,\mu_\Lambda) = \int \mathrm{d}k\, C_0(\omega-k,\mu_\Lambda) F(k)$

  [ZL, Stewart, Tackmann, 0807.1926]

  – Can use any (mass) scheme, work to any order

  – Stable results for varying μ_Λ (SF modeling scale, part of uncertainty, often ignored)

  – Similar to how all matrix elements are defined [e.g., $B_K(\mu) = \widehat{B}_K \times [\alpha_s(\mu)]^{2/9}(1+\ldots)$]

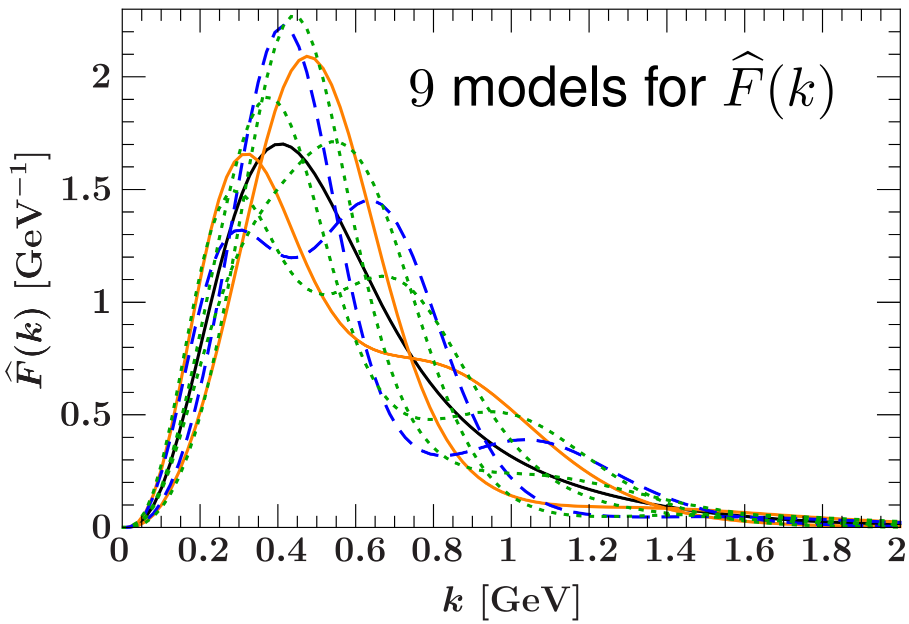
- Consistent to impose moment constraints on $F(k)$, but not on $S(\omega,\mu_\Lambda)$ w/o cutoff

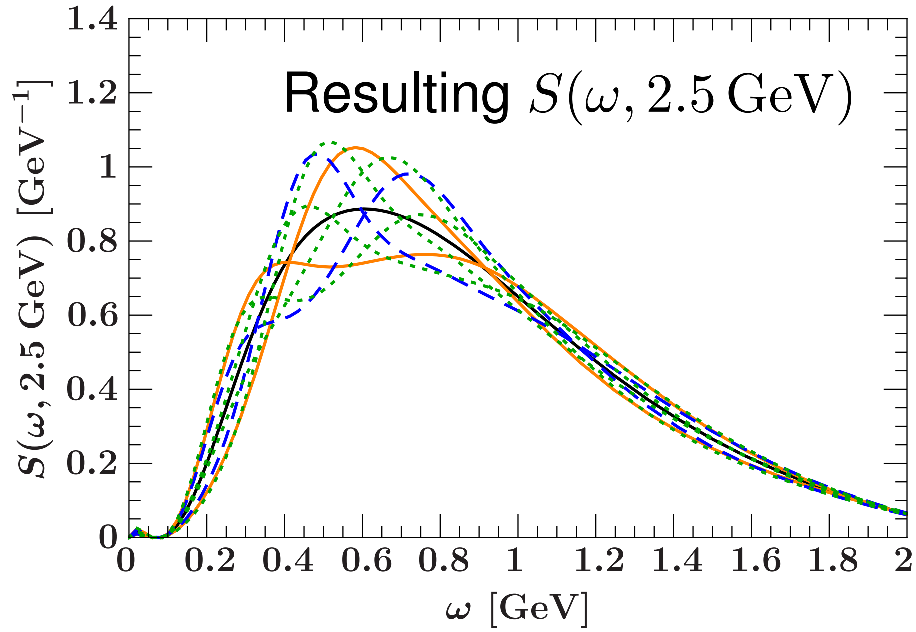
Shape function: the bottom line

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

$\widehat{F}$: nonperturbative
determines peak region
well-defined moments
fit from data

$\widehat{C}_0$: perturbatively calculable
generates tail consistent with RGE
divergent moments
solves question of model scale





BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

Master formula for decay rate

- Write decay rate as:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E_\gamma} = \frac{G_F^2 \widehat{m}_b^2}{8\pi^3}\,\frac{\alpha_{\rm em}}{4\pi} |V_{tb}V_{ts}|^2 E_\gamma^3 \int \mathrm{d}k \Big[\widehat{P}(k)\mathcal{F}(m_B - 2E_\gamma - k) + (\text{subleading})\Big]$$

  Perturbatively calculable:

$$\widehat{P}(k) = |C_7^{\rm incl}|^2 \Big[W_{77}^{\rm s}(k) + W_{77}^{\rm ns}(k)\Big] + 2\,{\rm Re}C_7^{\rm incl} \sum_{i\neq 7} \mathcal{C}_i\, W_{7i}^{\rm ns}(k) + \sum_{i,j\neq 7} \mathcal{C}_i \mathcal{C}_j\, W_{ij}^{\rm ns}(k)$$

- $W_{77}^{\rm s}(k) \sim HJ \otimes \widehat{C}_0$ contains factorized singular contributions
  - Resummed using SCET to NNLL' $+$ NNLO
  - Use profile scales to turn off resummation away from endpoint

- Non-singular terms are suppressed in the peak region

Split matching and $C_7^{\rm incl}$

- Decouple perturbative series above and below $\mu = m_b$ [Lee, ZL, Stewart, Tackmann, '05–'06]

$$C_7^{\rm incl} = \overline{C}_7(\mu) + \sum_{i\neq 7} \overline{C}_i(\mu) \Big[s_i(\mu, \widehat{m}_b) + r_i(\mu, \widehat{m}_b, \widehat{m}_c) \Big]$$

The s_i terms cancel the μ-dependence of $\overline{C}_7(\mu)$ and satisfy $s_i(\widehat{m}_b, \widehat{m}_b) = 0$

The $C_i\, r_i$ contain all virtual corrections $\propto C_{i\neq 7}$ that give singular contributions

- Integrate out b and c quarks at the hard scale $\mu = m_b$

Consistent with measurements removing charm states as backgrounds; simplifies SCET setup

- Better to compare theory and data for $|C_7^{\rm incl}|$ than for $\mathcal{B}(E_\gamma > 1.6\, {\rm GeV})$

- From Misiak *et al.* (SM, NNLO): $|C_7^{\rm incl}|_{\rm SM} = 0.3624 \pm 0.0128_{c\bar c} \pm 0.0080_{\rm scale}$

Nonsingular contributions

- $C_7^{\rm incl}$ captures all singular and some nonsingular terms $\propto C_{i\neq 7}$
  - Remainder gives $W_{i7}^{\rm ns}$ and $W_{ij}^{\rm ns}$, numerically subdominant
  - Include $ij = 78$ to $\mathcal{O}(\alpha_s^2)$; all others for $i = 1, 2, 8$ to $\mathcal{O}(\alpha_s^2 \beta_0)$
- Relevant for tails, especially $W_{77}^{\rm ns}(k)$, included to $\mathcal{O}(\alpha_s^2)$

  Precise form depends on whether overall E_γ^3 is kept exact

  If expanded, E_γ^3 dependence arises as cancellations between sing. and nonsing.

  Better to keep as prefactor, effectively resums some kinematic power corrections
- Not trivial to extract $W_{i7}^{\rm ns}$ and $W_{ij}^{\rm ns}$, e.g.,

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E_\gamma}\bigg|_{7i} = \frac{\alpha_s}{\pi}\Big[r_i^{(1)}\delta(k) + w_{7i}^{\rm ns(1)}(k)\Big] + \frac{\alpha_s^2}{\pi^2}\Big\{r_i^{(2)}\delta(k) + r_i^{(1)}\Big[w_{77}^{sing(1)}(k) + w_{77}^{\rm ns(1)}(k)\Big] + w_{7i}^{\rm ns(2)}(k)\Big\}$$

Basis expansion: designer orthonormal functions

- Devise suitable orthonormal basis functions

$$\mathcal{F}(k) = \frac{1}{\lambda}\Big[\sum c_n f_n(\tfrac{k}{\lambda})\Big]^2, \quad n\text{ th moment} \sim \Lambda_{\rm QCD}^n$$

$$f_n(x) \sim P_n[y(x)] \leftarrow \text{Legendre polynomials}$$

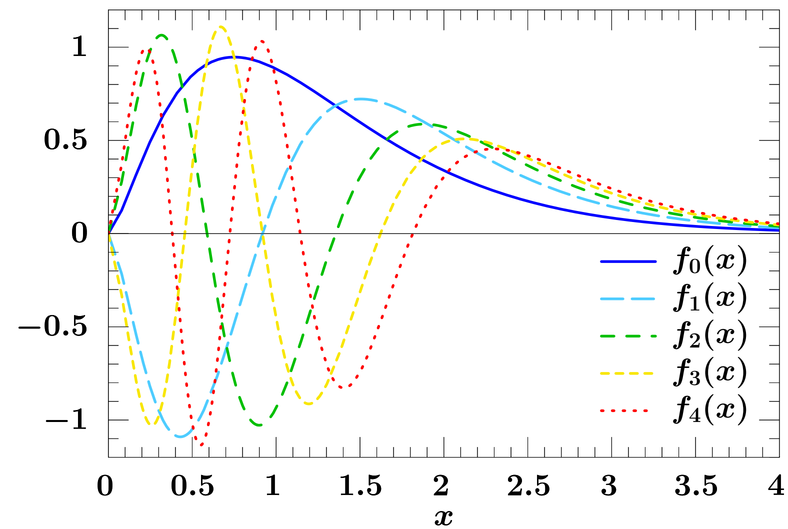
- Can construct an orthonormal basis, where f_0 is any model shape function

Better to add a new term in an orthonormal basis than a new parameter to a model:

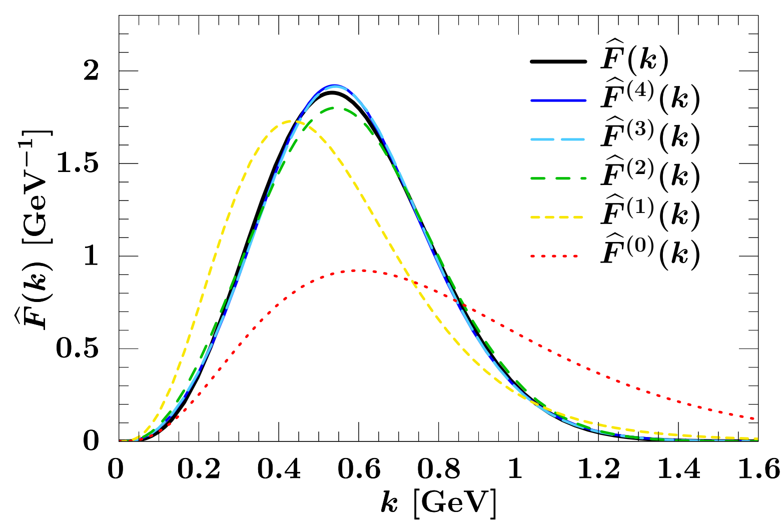
– less parameter correlations

– uncertainties easier to quantify

"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk." (John von Neumann)



[ZL, Stewart, Tackmann, 0807.1926]



BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

Global fit

Theory inputs to fit

- $\mathcal{F}(k) = \sum c_m c_n F_{mn}(k)$ enters the spectra linearly $\qquad F_{mn}(k) = \frac{1}{\lambda} f_m(\frac{k}{\lambda})\, f_n(\frac{k}{\lambda})$

  $\Rightarrow$ can calculate independently the contribution of $f_m\, f_n$ in the expansion of $\mathcal{F}(k)$:

$$\mathrm{d}\Gamma = \sum_{m,n=0}^{N} \underbrace{c_m\, c_n}_{\text{fit}} \underbrace{\mathrm{d}\Gamma_{mn}}_{\text{compute}}$$

  Precompute $\mathrm{d}\Gamma_{ij}$ for each measured bin, using fixed theory inputs

  (In subleading terms use SM values for $C_{i\neq 7}$ and SM sign for $C_7^{\rm incl}$)

- Fit $|C_7^{\rm incl}|$ and c_i coefficients from all data (similar to inclusive $|V_{cb}|$ fit)

  Redo fit with different theory inputs to estimate theory uncertainties

- Perturbative uncertainties dominate on the theory side
  - Independently vary hard, jet, soft, nonsingular scales, profile parameter
  - Consider all possible $3^5 = 243$ variations

BERKELEY CENTER FOR THEORETICAL PHYSICS

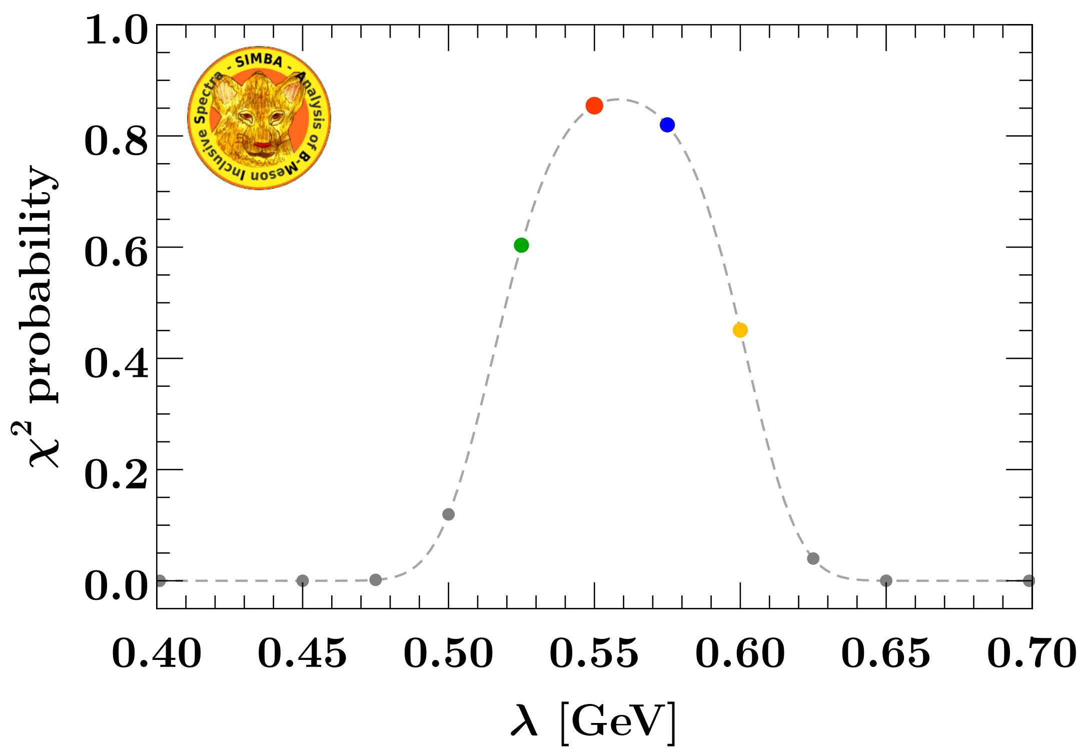
BERKELEY LAB

Choosing a basis

- Want $\mathcal{F}(k)$ well approximated by $F_{00}(k)$ $\Rightarrow$ quick convergence with small N

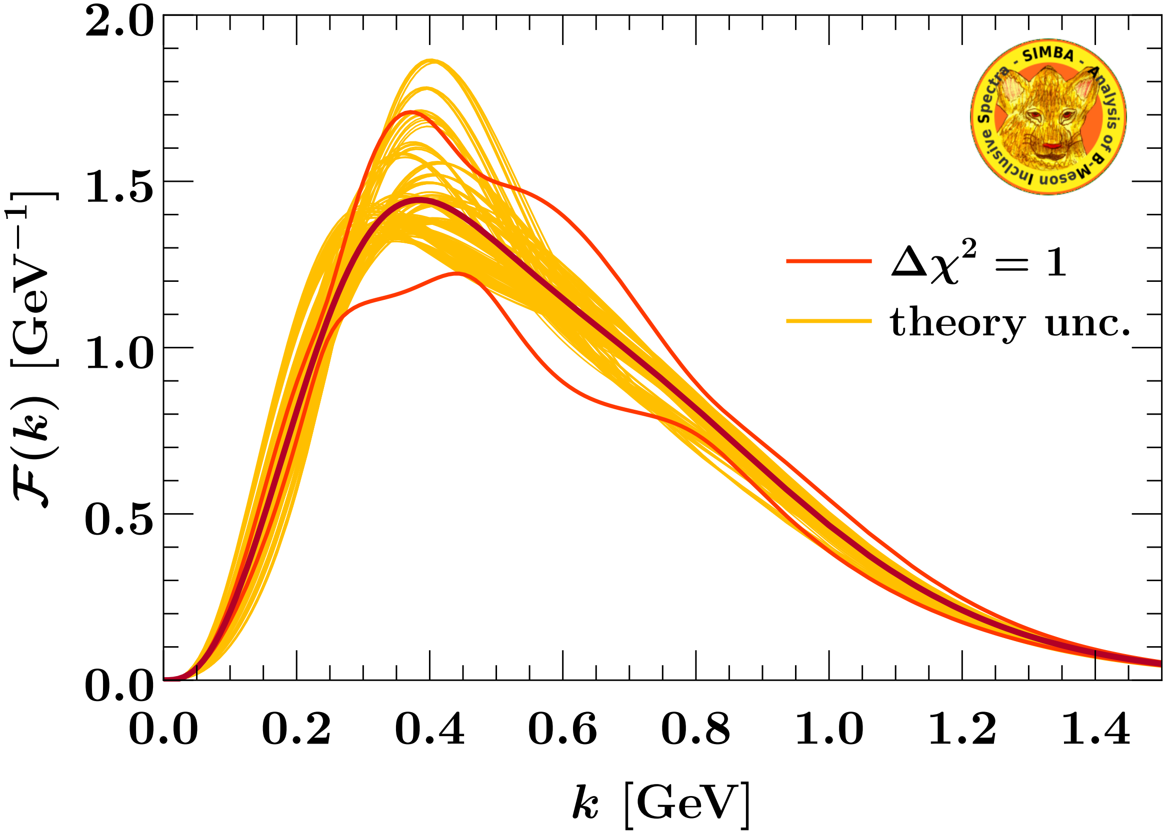
Take $F_{00}(k) \sim \dfrac{k^3}{\lambda^4} e^{-4k/\lambda}$

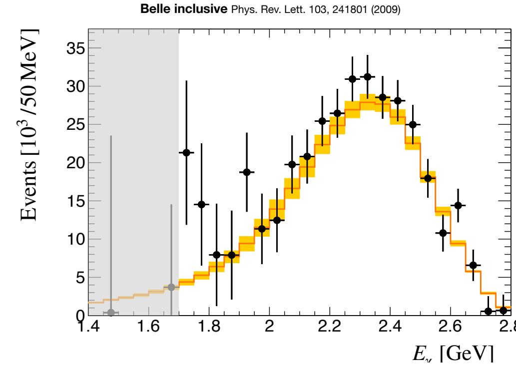
and do pre-fits to data

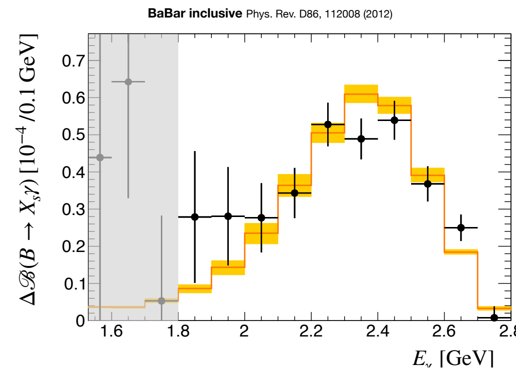


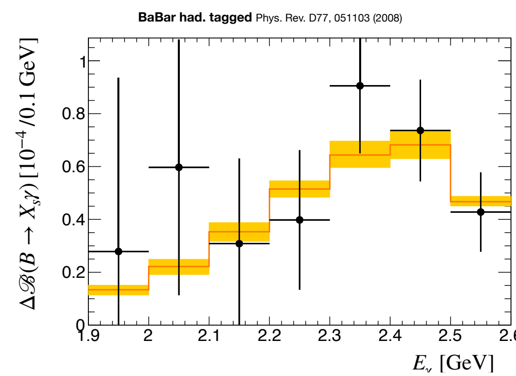
- Use as default $\lambda = 0.55\,\mathrm{GeV}$, and $\lambda = 0.525,\, 0.575,\, 0.6\,\mathrm{GeV}$ to test the basis independence

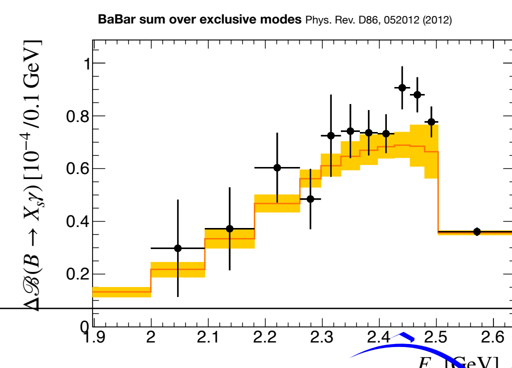
Fits to $B \to X_s\gamma$





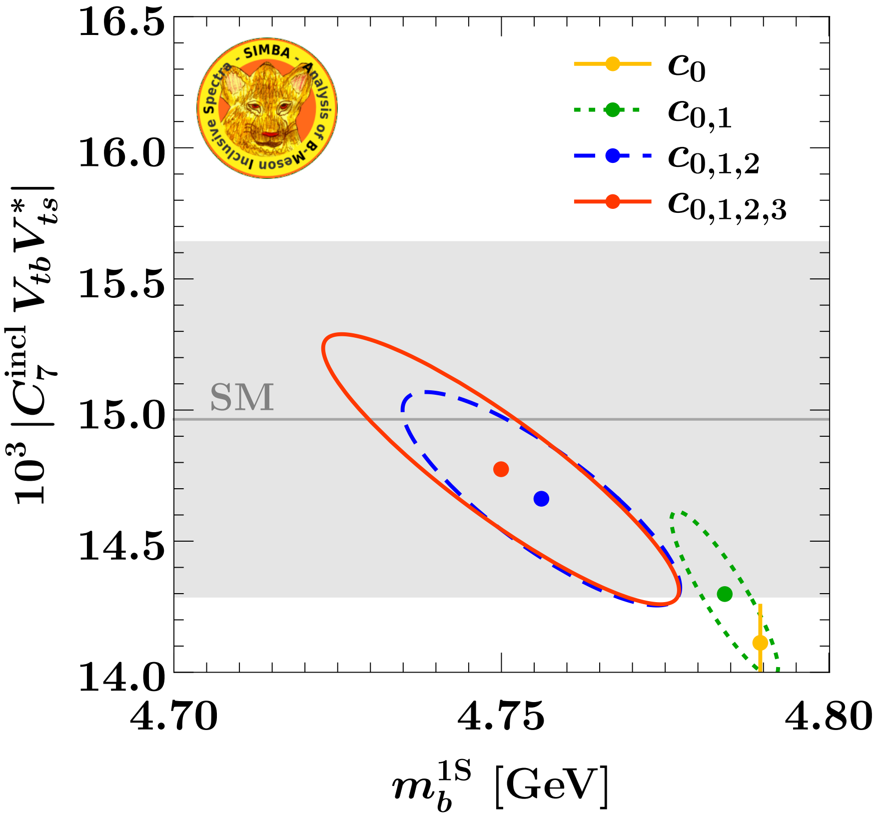


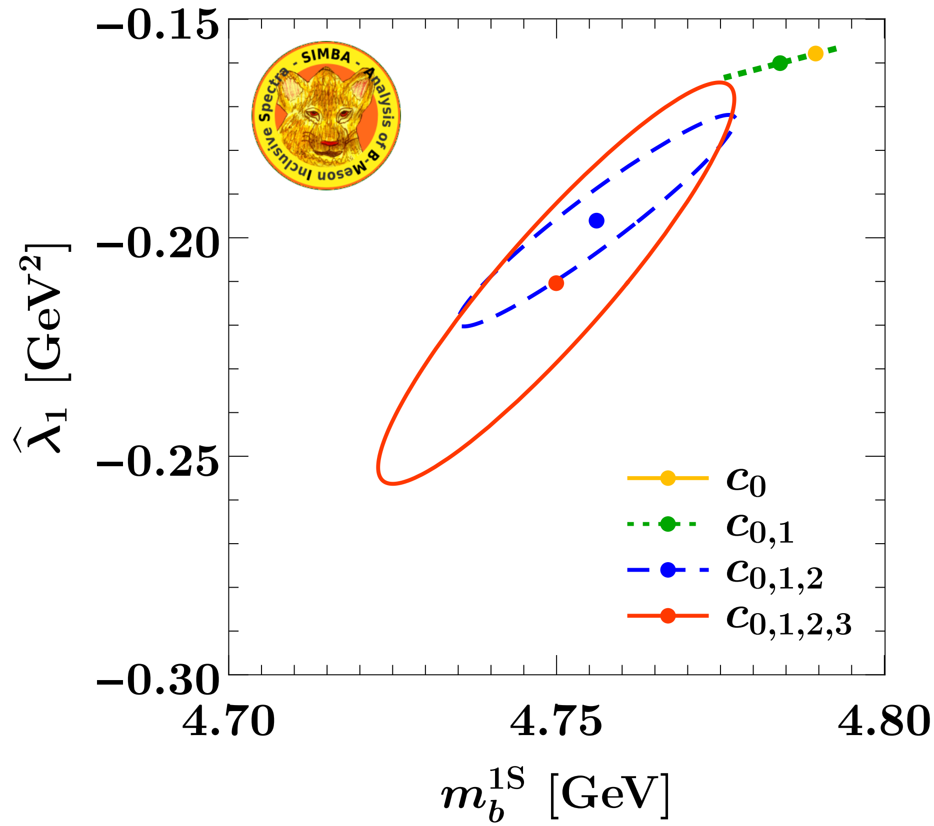




Using moment relations for $\mathcal{F}(k)$, can obtain fit results for m_b, λ_1, etc.

Fit results for $B \to X_s\gamma$





Increase N until no significant improvement in fit quality, use nested hypothesis test

$$|C_7^{\rm incl} V_{tb} V_{ts}| = (14.77 \pm 0.51_{\rm fit} \pm 0.59_{\rm theory} \pm 0.08_{\rm param}) \times 10^{-3}$$

Larger uncert. than HFLAV

$$m_b^{1S} = (4.750 \pm 0.027_{\rm fit} \pm 0.033_{\rm theory} \pm 0.003_{\rm param})\ {\rm GeV}$$

$$\widehat{\lambda}_1 = (-0.210 \pm 0.046_{\rm fit} \pm 0.040_{\rm theory} \pm 0.056_{\rm param})\ {\rm GeV}^2$$

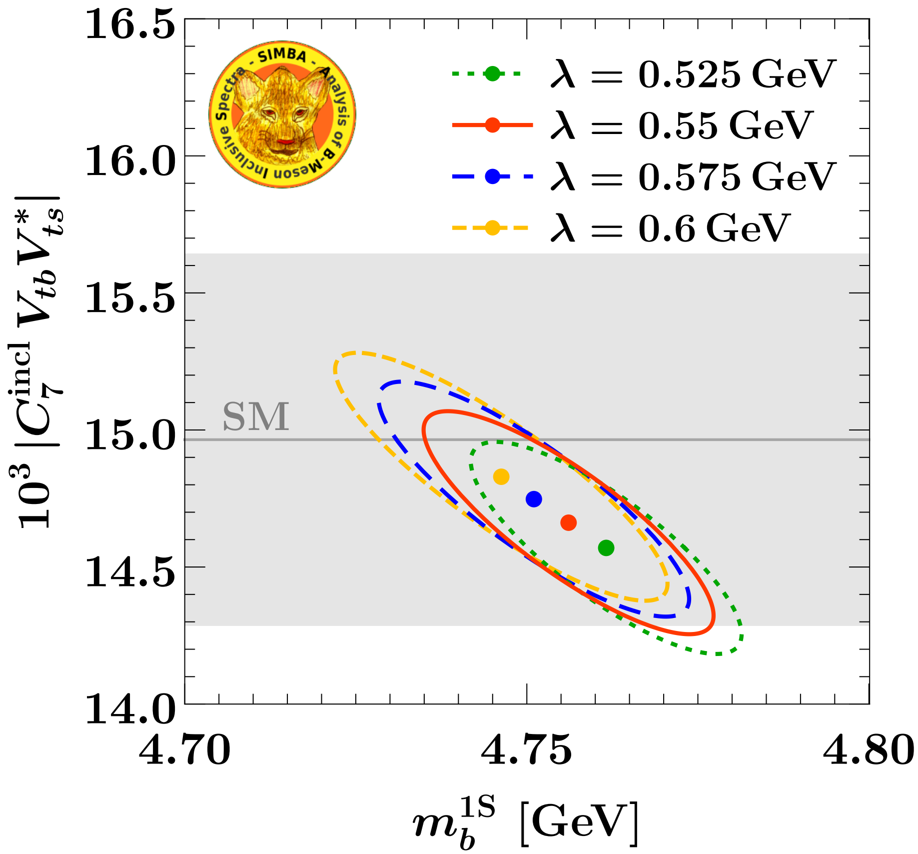
BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

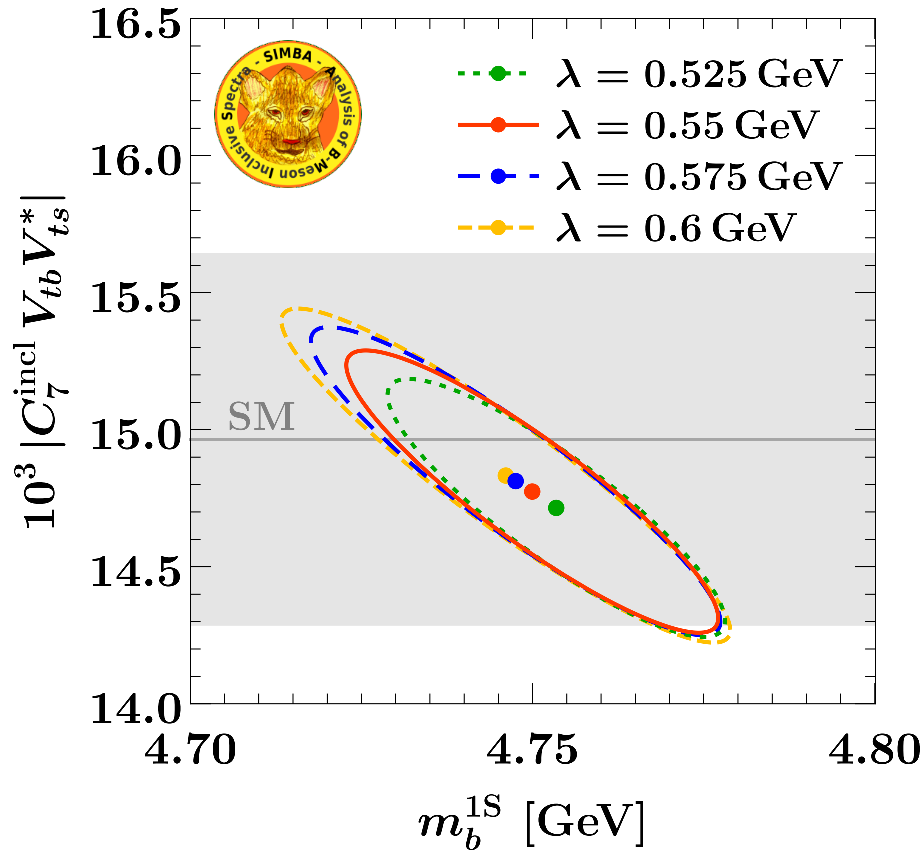
Verify basis independence

- Fits with 4 different values of λ:

Fits with c_0, c_1, c_2,



Fits with c_0, c_1, c_2, c_3



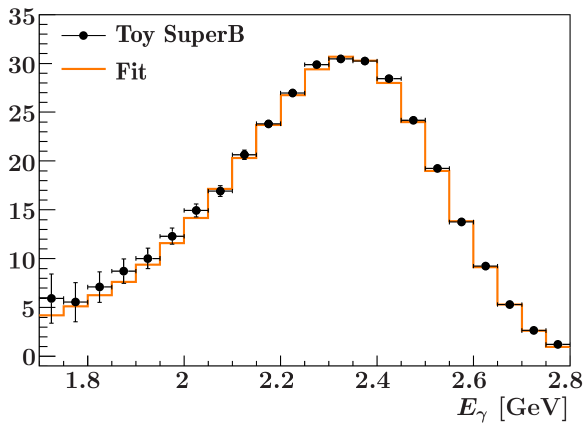
Future of $B \to X_s\gamma$

- Toy fits few years ago for 75/ab:

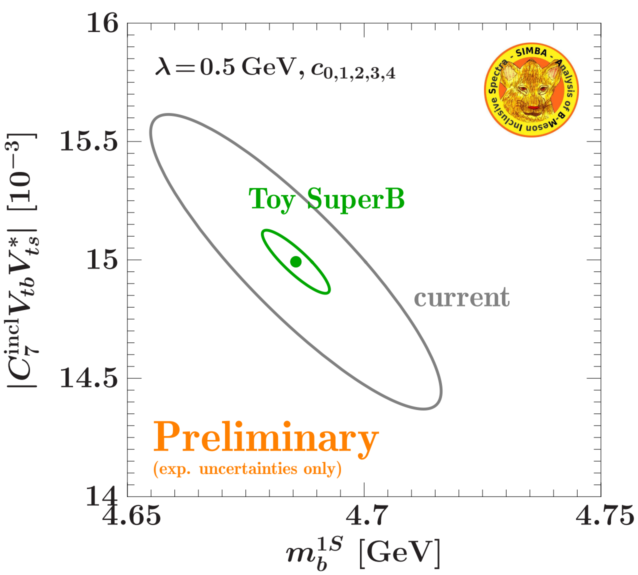
5 coefficients

$\lambda = 0.5\,\text{GeV}$

Theory uncert. will dominate



[BELLE2-NOTE-0021]	Statistical	Systematic (reducible, irreducible)	Total Exp
$\mathcal{B}(B \to X_s\gamma)$ inclusive (untagged)			
605 fb^{-1}	4.2	(10.3, 5.3)	12.3
5 ab^{-1}	1.5	(3.6, 5.3)	6.6
50 ab^{-1}	0.5	(1.1, 5.3)	5.4
$\mathcal{B}(B \to X_s\gamma)$ inclusive (hadron tagged)			
210 fb$^{-1\dagger}$	23.2	(15.7, 4.8)	28.4
5 ab^{-1}	4.8	(3.2, 4.8)	7.5
50 ab^{-1}	1.5	(1.0, 4.8)	5.1



We assumed factor of 3 reduction in systematic uncertainty, slightly (but not vastly) optimistic

High precision data can be used to fit with more coefficients and constrain subleading effects

$$B \to X_u \ell \bar{\nu}$$

Everything below are preliminary & based on (old) toys

V_{ub} — the beginning

CLEO, PRL **64** (1990) 16, Received 8 Nov 1989, $(212+101)/\mathrm{pb}$

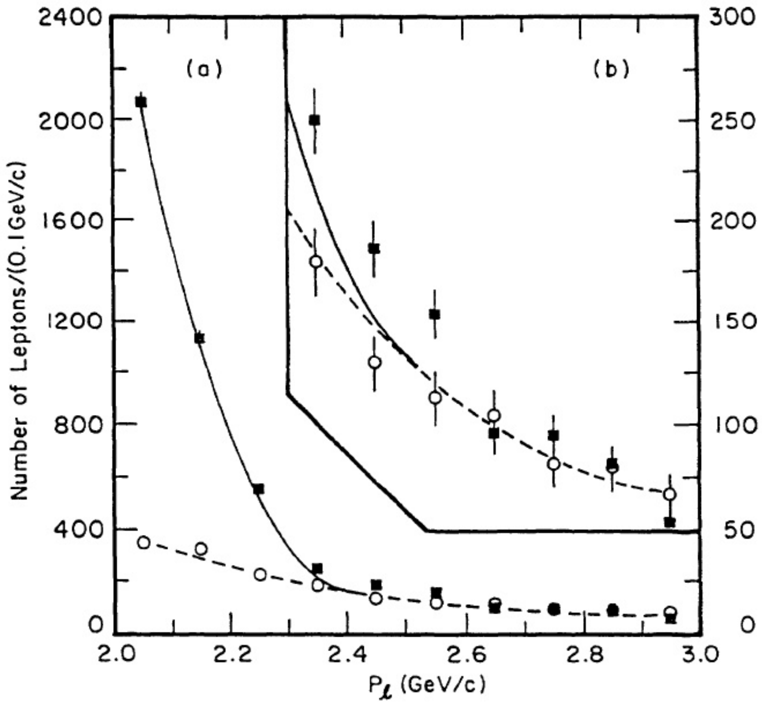


FIG. 1. Sum of the e and μ momentum spectra for ON data (filled squares), scaled OFF data (open circles), the fit to the OFF data (dashed line), and the fit to the OFF data plus the $b \to c l \nu$ yield (solid line). Note the different vertical scales in (a) and (b).

"$|V_{ub}/V_{cb}|$ ... is approximately 0.1; it is sensitive to the theoretical model."

ARGUS, PLB **234** (1990) 409, Received 28 Nov 1989, $(201+69)/\mathrm{pb}$

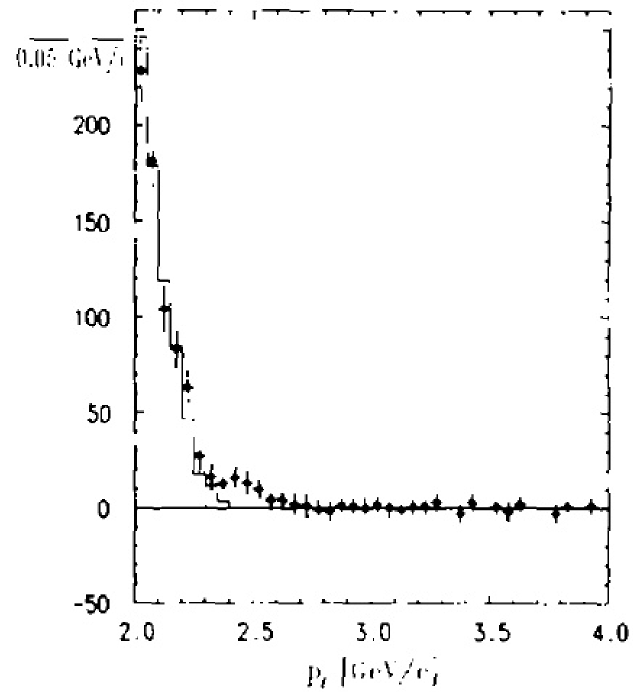


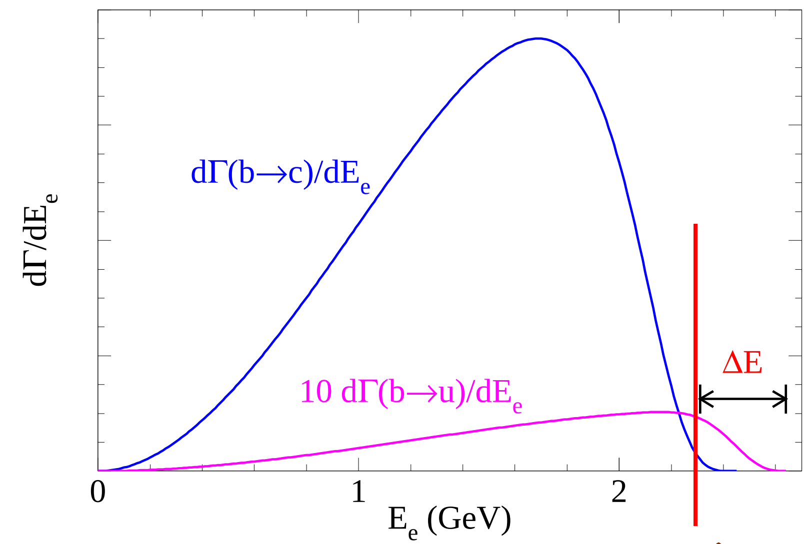
Fig. 5. Combined lepton momentum spectrum for direct $\Upsilon(4S)$ decays: the histogram is a $b \to c$ contribution normalized in the region 2.0–2.3 GeV/c.

"If interpreted as a signal of $b \to u$ coupling ... $|V_{ub}/V_{cb}|$ of about 10%."

The challenge of inclusive $|V_{ub}|$ measurements

- Total rate calculable with $\sim 4\%$ uncertainty, similar to $\mathcal{B}(B \to X_c \ell \bar\nu)$
- To remove the huge charm background $(|V_{cb}/V_{ub}|^2 \sim 100)$, need phase space cuts

  Phase space cuts can enhance perturbative and nonperturbative corrections drastically



- Hadronic parameters are functions (like PDFs)

  Leading order: universal & related to $B \to X_s\gamma$;
  $\mathcal{O}(\Lambda_{\rm QCD}/m_b)$: several new unknown functions

- Nonperturbative effects shift endpoint $\frac{1}{2}m_b \to \frac{1}{2}m_B$ & determine its shape

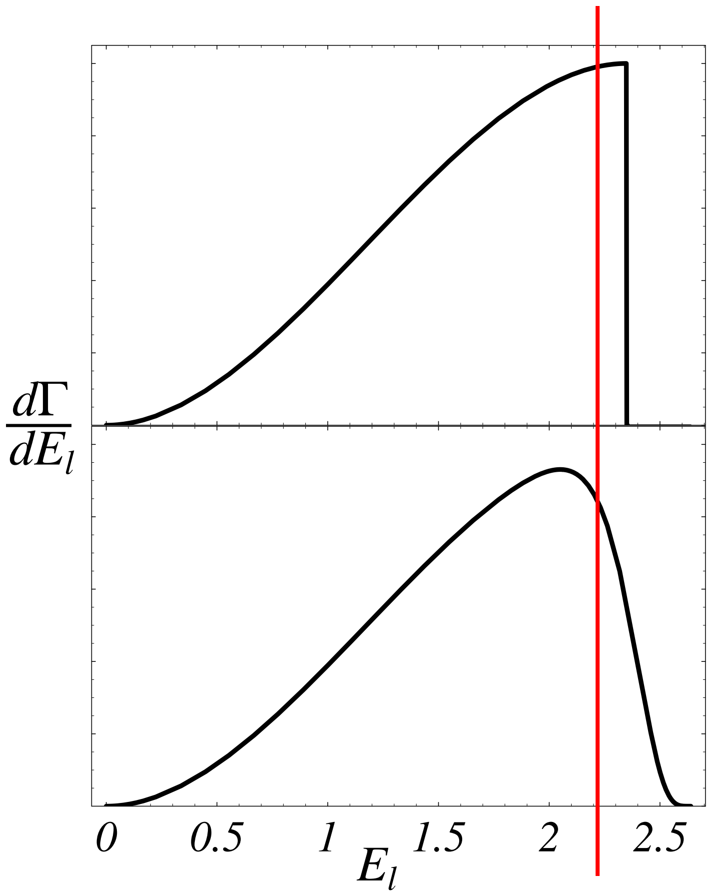
- Shape in the endpoint region is determined by b quark PDF in B

  Related to $B \to X_s\gamma$ photon spectrum at lowest order [Bigi, Shifman, Uraltsev, Vainshtein; Neubert]

Shape function: lepton endpoint vs. $B \to X_s\gamma$

b quark decay spectrum

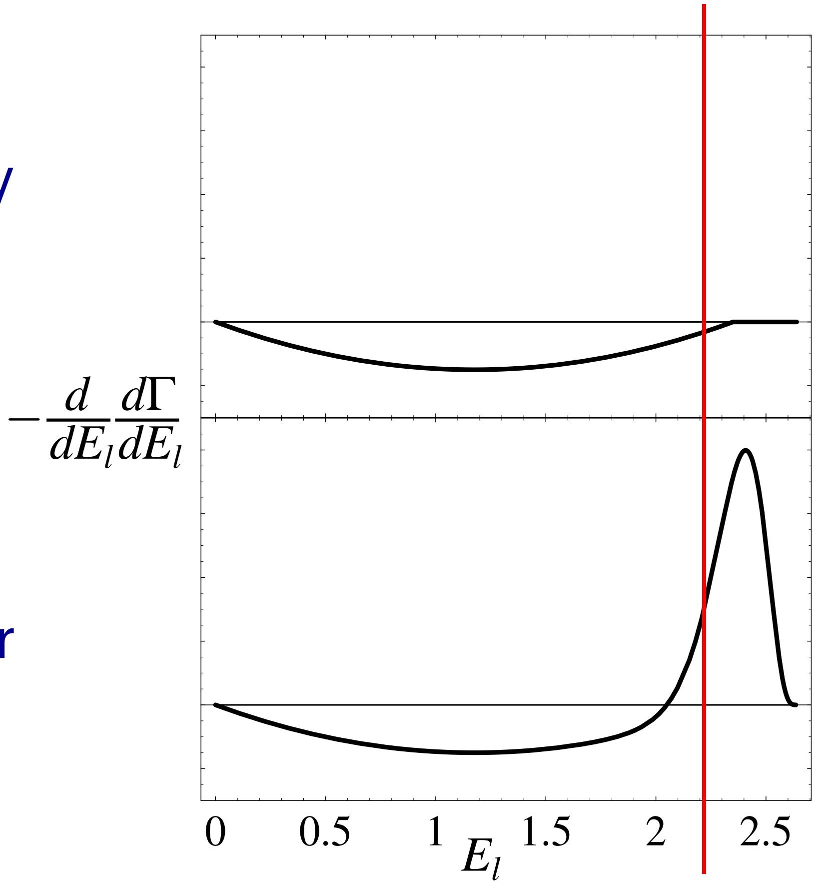
with a model for b quark PDF



Shape function: lepton endpoint vs. $B \to X_s\gamma$

b quark decay spectrum

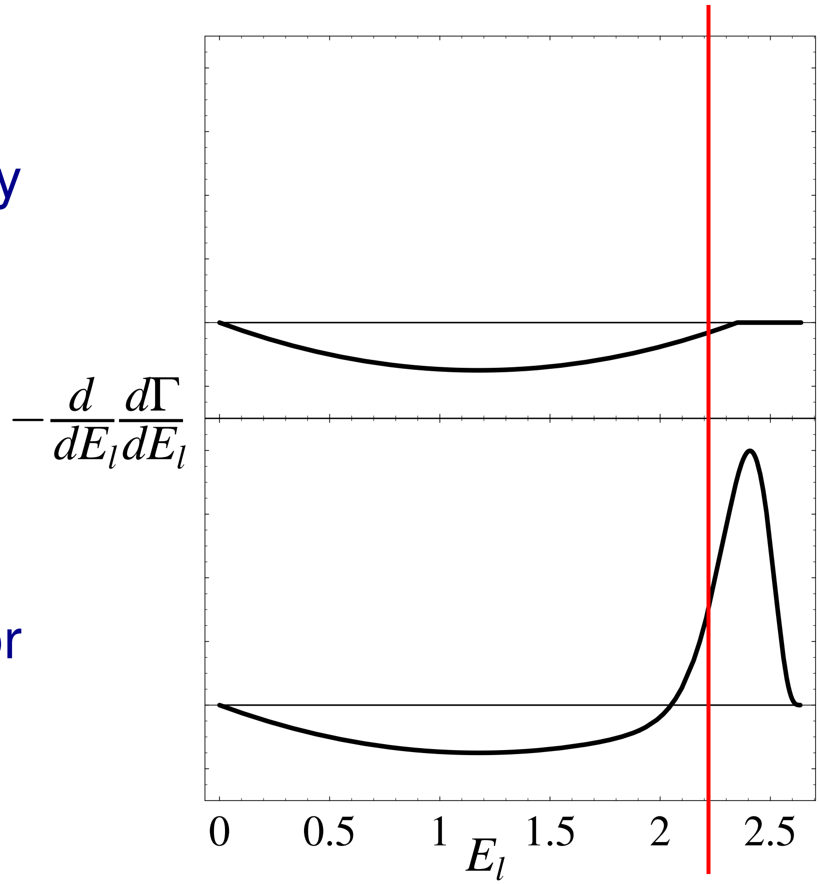
with a model for b quark PDF



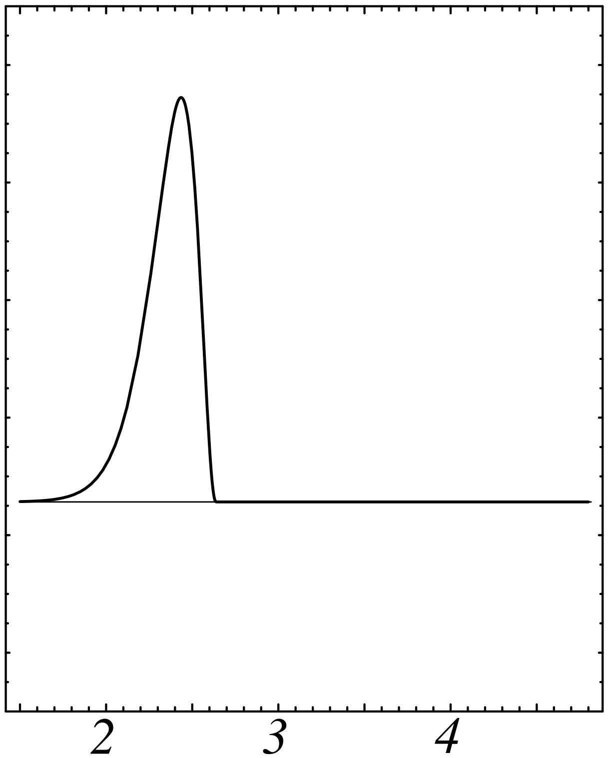
Shape function: lepton endpoint vs. $B \to X_s\gamma$

b quark decay spectrum

with a model for b quark PDF



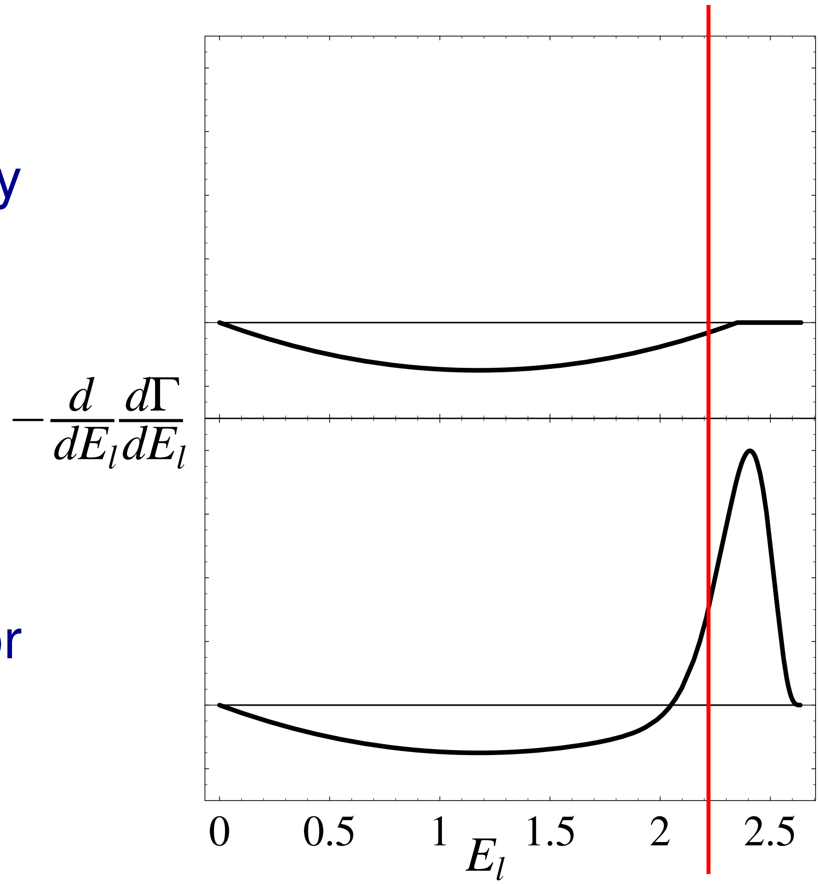
difference:



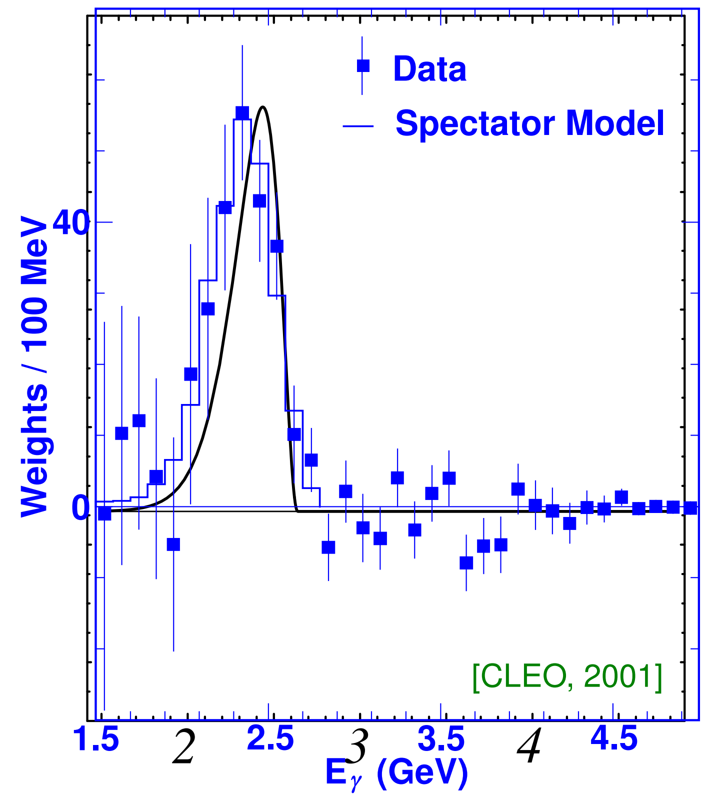
Shape function: lepton endpoint vs. $B \to X_s\gamma$

b quark decay spectrum

with a model for b quark PDF



difference:



- Both spectra determined at lowest order by the b quark PDF in B meson

$B \to X_u \ell \bar{\nu}$: more complicated kinematics

- "Natural" kinematic variables: $p_X^{\pm} = E_X \mp |\vec{p}_X|$ (ratio is "jettiness" of hadrons)

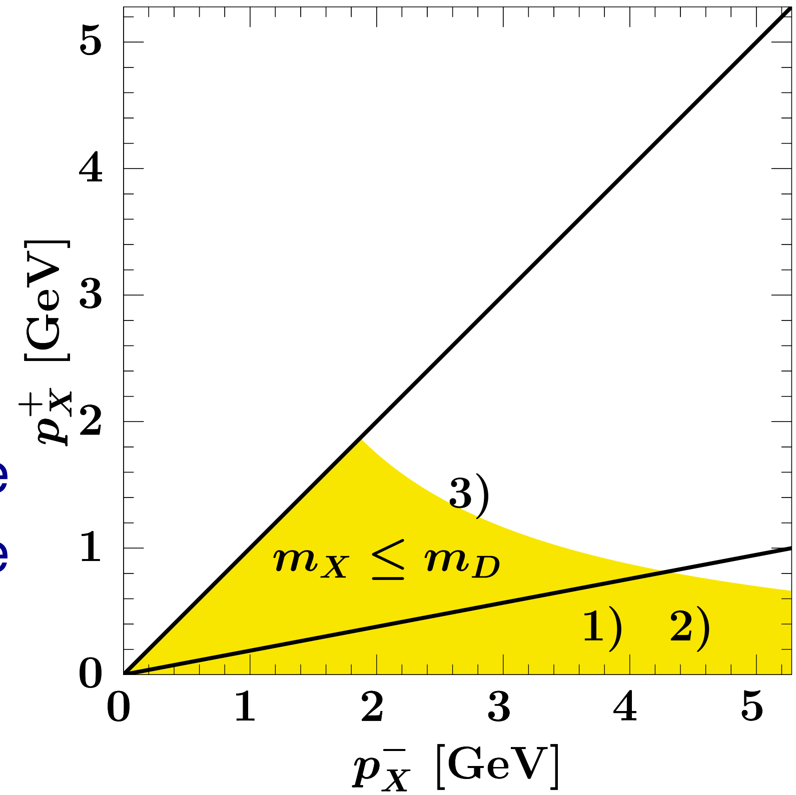
  $B \to X_s \gamma$: $p_X^+ = m_B - 2E_\gamma$ & $p_X^- \equiv m_B$ — independent variables in $B \to X_u \ell \bar{\nu}$

- Three cases: 1) $\Lambda \sim p_X^+ \ll p_X^-$
  2) $\Lambda \ll p_X^+ \ll p_X^-$ } Shape Fn region
  3) $\Lambda \ll p_X^+ \sim p_X^-$ local OPE region

  Want to make no assumptions how p_X^- compares to m_B

- $B \to X_u \ell \bar{\nu}$: 3-body final state, appreciable rate in region 3), where hadronic final state not jet-like

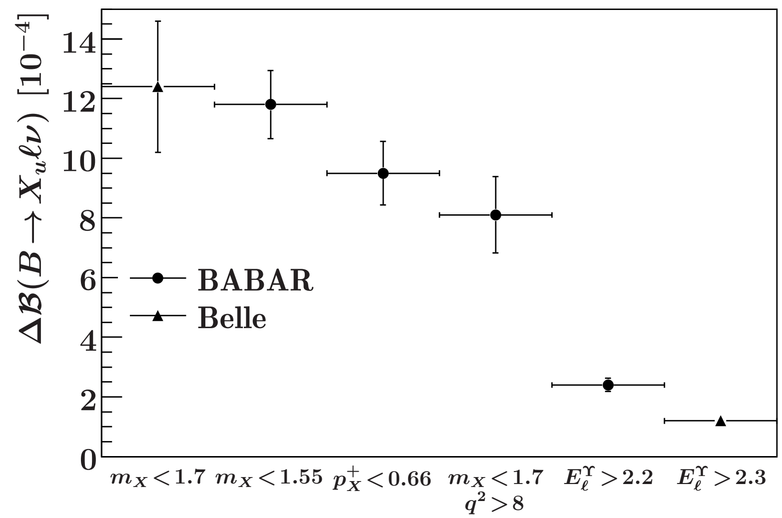
  E.g., $m_X^2 < m_D^2$ does not imply $p_X^+ \ll p_X^-$

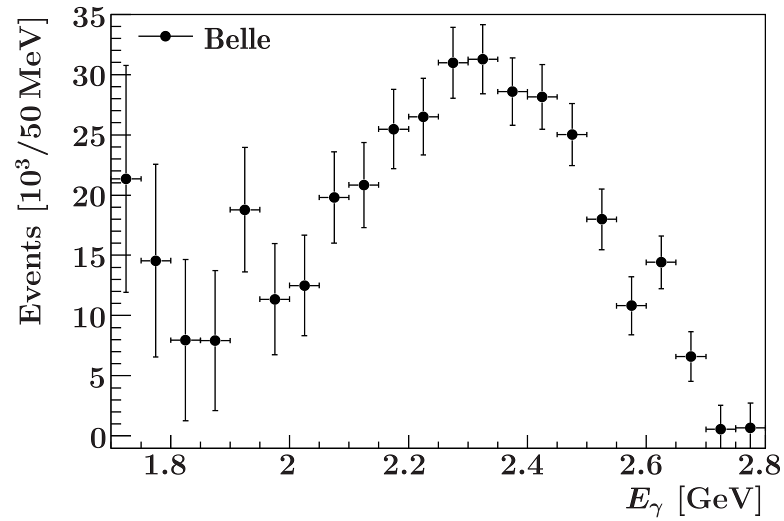


- Existing results based on theory in one region, extrapolated / modeled to rest

Exploratory: $|V_{ub}|$ w/ NLO + NLL' only

- $B \to X_u \ell \bar{\nu}$ hadronic tag
  - BaBar $m_X,\ m_X - q^2,\ p_X^+$
  - Belle m_X
- $B \to X_u \ell \bar{\nu}$ lepton endpoint
  - BaBar $E_\ell^\Upsilon > 2.2\,\mathrm{GeV}$
  - Belle $E_\ell^\Upsilon > 2.3\,\mathrm{GeV}$
- $B \to X_s \gamma$ spectra
  - Belle latest result (shown)
  - BaBar sum over exclusive + hadronic tag
- m_b^{1S}, λ_1 from $B \to X_c \ell \bar{\nu}$ fit
  - $m_b^{1S} = (4.66 \pm 0.05)\,\mathrm{GeV}$
  - $\lambda_1 = (-0.34 \pm 0.05)\,\mathrm{GeV}^2$

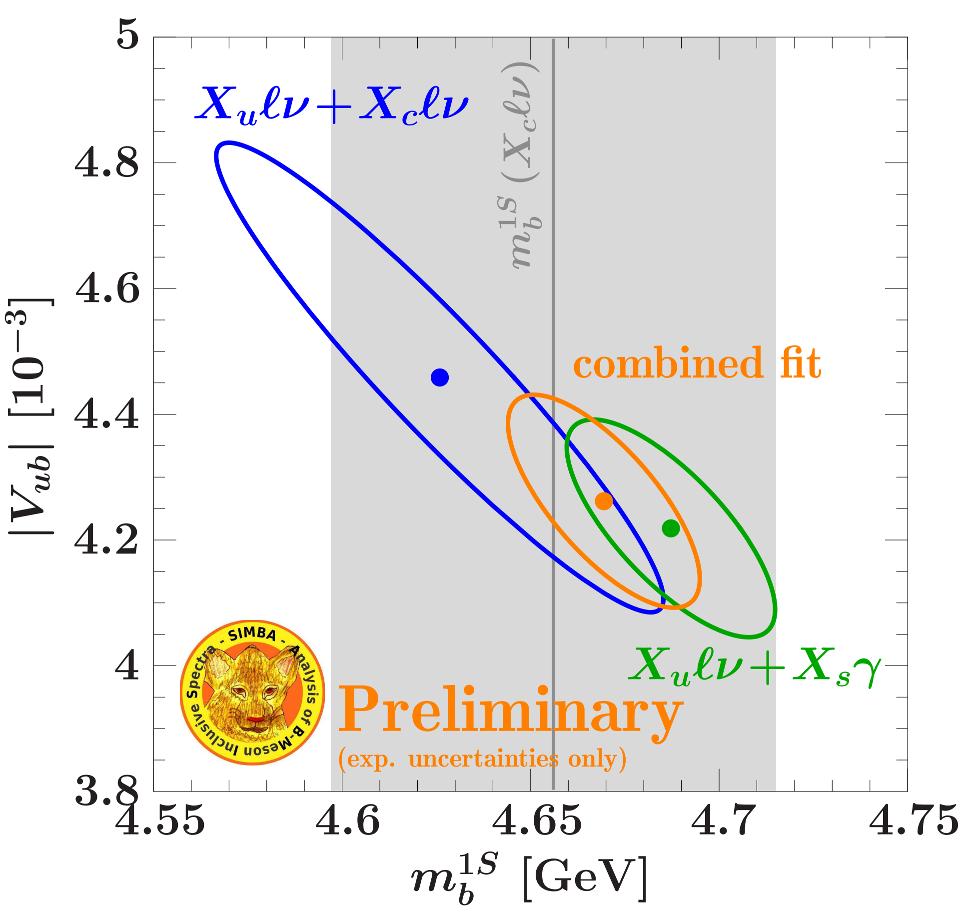


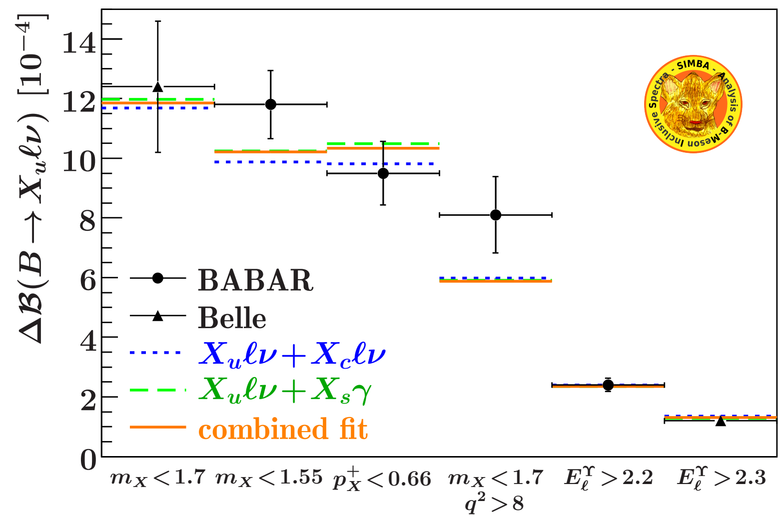


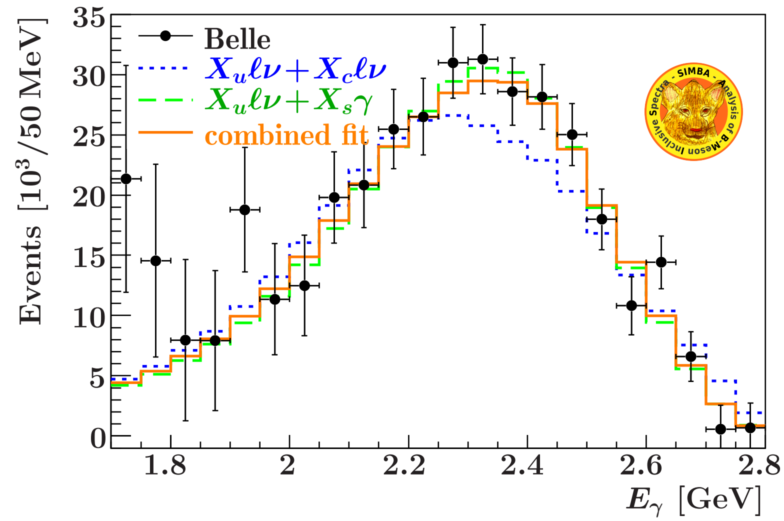
BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

Exploratory: $|V_{ub}|$ w/ NLO + NLL' only







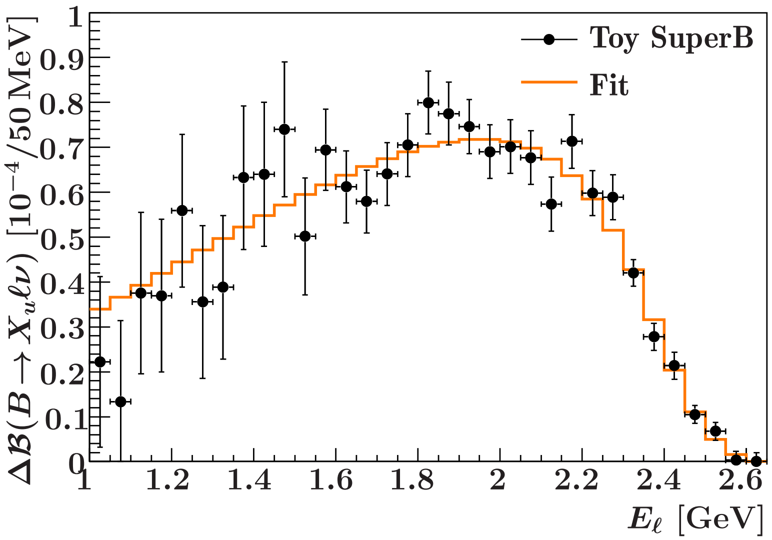
- E_γ spectrum is off without $B \to X_s\gamma$ in the fit
- Including it, favors lower values of $|V_{ub}|$

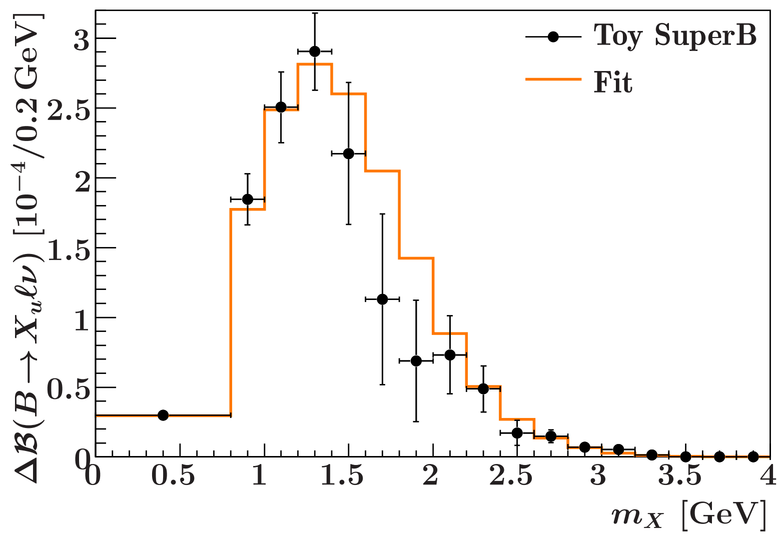
BERKELEY CENTER FOR THEORETICAL PHYSICS

BERKELEY LAB

Future of $B \to X_u \ell \bar{\nu}$

- Spectra generated with $\lambda = 0.6\,\mathrm{GeV}$ and $c_0 = 1$ (Assumed uncertainties & correlations similar to BaBar full reco analysis, 1112.0702 — by now Belle hadronic tagging efficiency is better)

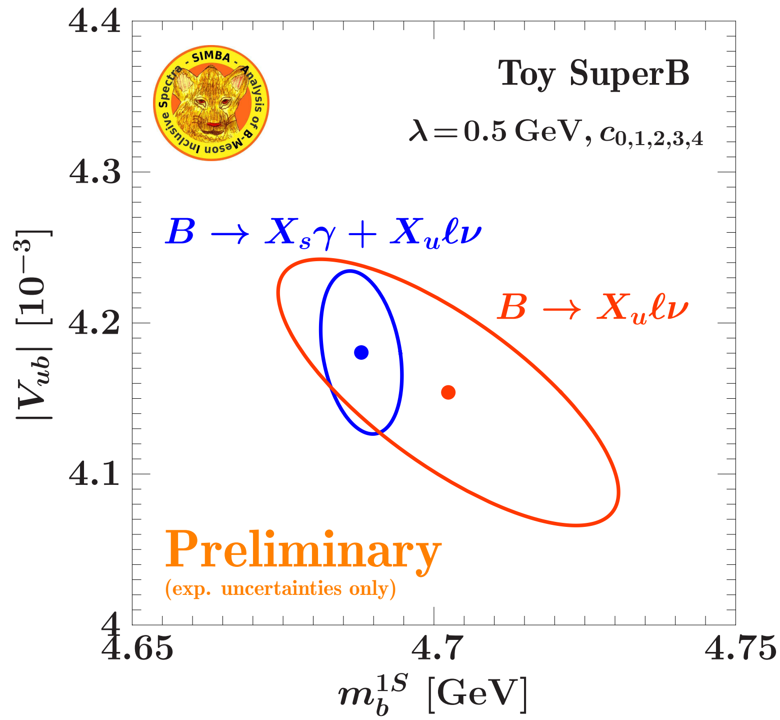




- Measure spectra — the rate with low E_ℓ or high m_X cut cannot give optimal $|V_{ub}|$
  - Uncertainties grow, as for $\mathrm{d}\Gamma(B \to X_s \gamma)/\mathrm{d}E_\gamma$
  - Experimental analysis needs input on shape in any case
- Large data sets will push analysis to the limits, constrain subleading SF effects

Future of $B \to X_u \ell \bar{\nu}$ (2)

- Toy fit with 5 coefficients for $75/\text{ab}$:



- With Belle II data sets:
  - Combination with $B \to X_s\gamma$ will be essential for ultimate sensitivity
  - Combination with $B \to X_c\ell\bar{\nu}$ (moments, shapes?) also possible

Final comments

Conclusions

- First gloal fit of inclusive $B \to X_s \gamma$
  - Model independent and data-driven treatment of shape function
  - More reliable than using $\mathcal{B}(E_\gamma > 1.6\,\text{GeV})$
  - Precise extraction of $|C_7^{\text{incl}}|$
- Larger uncertainty than HFLAV analysis, more room for BSM at present
  Belle II can yield significant improvements
- Current status of $|V_{ub}|$ unsettled — improvement crucial to better constrain NP
  Hope to see measurements w/ different uncertainties agree (incl., excl., leptonic)

  Qualitatively better inclusive $|V_{ub}|$ analysis possible than those implemented so far

Backup slides

Derivation of the magic formula (1)

- The shape function is the matrix element of a nonlocal operator:

$$S(\omega,\mu) = \langle B|\,\underbrace{\bar b_v\,\delta(iD_+ - \delta + \omega)\,b_v}_{O_0(\omega,\mu)}\,|B\rangle, \qquad \delta = m_B - m_b$$

Integrated over a large enough region, $0 \le \omega \le \Lambda$, one can expand O_0 as

$$O_0(\omega,\mu) = \sum C_n(\omega,\mu)\,\underbrace{\bar b_v\,(iD_+ - \delta)^n\,b_v}_{Q_n} + \ldots = \sum C_n(\omega-\delta,\mu)\,\underbrace{\bar b_v\,(iD_+)^n\,b_v}_{\widetilde Q_n} + \ldots$$

The C_n are the same for Q_n and $\widetilde Q_n$ (since O_0 only depends on $\omega - \delta$)

- Matching: $\langle b_v|O_0(\omega{+}\delta,\mu)|b_v\rangle = \sum C_n(\omega,\mu)\,\langle b_v|\widetilde Q_n|b_v\rangle = C_0(\omega,\mu), \quad \langle b_v|\widetilde Q_n|b_v\rangle = \delta_{0n}$

$$\langle b_v(k_+)|O_0(\omega+\delta,\mu)|b_v(k_+)\rangle = C_0(\omega+k_+,\mu) = \sum \frac{k_+^n}{n!}\,\frac{\mathrm{d}^n C_0(\omega,\mu)}{\mathrm{d}\omega^n}$$

$$\langle b_v(k_+)|O_0(\omega+\delta,\mu)|b_v(k_+)\rangle = \sum C_n(\omega,\mu)\langle b_v|\widetilde Q_n|b_v\rangle = \sum C_n(\omega,\mu)\,k_+^n$$

- Comparing last two lines: $C_n(\omega,\mu) = \dfrac{1}{n!}\,\dfrac{\mathrm{d}^n C_0(\omega,\mu)}{\mathrm{d}\omega^n}$ [Bauer & Manohar]

Derivation of the magic formula (2)

- Define the nonperturbative function $F(k)$ by: [ZL, Stewart, Tackmann; Lee, ZL, Stewart, Tackmann]

$$S(\omega, \mu_\Lambda) = \int \mathrm{d}k\, C_0(\omega - k, \mu_\Lambda)\, F(k), \qquad C_0(\omega, \mu) = \langle b_v | O_0(\omega + \delta, \mu) | b_v \rangle$$

uniquely defines $F(k)$: $\widetilde{F}(y) = \widetilde{S}(y, \mu) / \widetilde{C}_0(y, \mu)$

- Expand in k: $S(\omega, \mu) = \sum_n \frac{1}{n!} \frac{\mathrm{d}^n C_0(\omega, \mu)}{\mathrm{d}\omega^n} \int \mathrm{d}k\, (-k)^n F(k)$

Compare with previous page $\Rightarrow$ $\int \mathrm{d}k\, k^n F(k) = (-1)^n \langle B | Q_n | B \rangle$

$$\langle B | Q_0 | B \rangle = 1\,, \quad \langle B | Q_1 | B \rangle = -\delta\,, \quad \langle B | Q_2 | B \rangle = -\frac{\lambda_1}{3} + \delta^2$$

More complicated situation for higher moments, so stop here

- This treatment is fully consistent with the OPE