University of Ljubljana
Faculty of Mathematics and Physics

IJS

Measurement of the $B^+ \to K^+K^-\ell^+\nu_\ell$ Decay with the Belle Detector

Matic Lubej

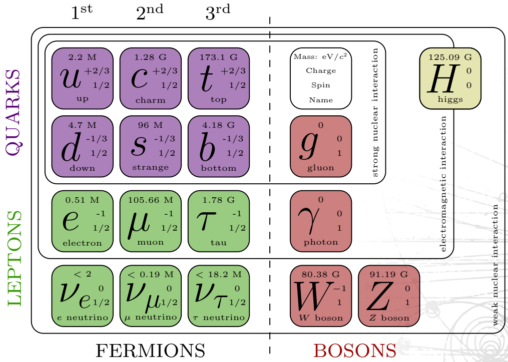
PhD Thesis Defense

Ljubljana, November 22$^{\text{th}}$, 2018

Overview

Introduction I – Standard Model

- The Standard Model (SM) is a basic theory which describes elementary particles and interactions between them
- Quarks: can be grouped into baryons ($q_1q_2q_3$) or mesons ($\bar{q}_1q_2$)
- Focus of this analysis are B mesons: $B^+(\bar{b}u)$

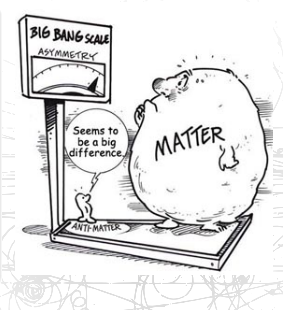


Introduction II – *CP* Violation

- *C* (charge conjugation), *P* (parity) and *T* (time reversal) symmetries were believed to be conserved individually (based on EM interaction, 1954)
- Observation of *P* symmetry violation confused physicists (1956)
- Observation of *CP* symmetry violation confused physicists even more (1964)

CP symmetry? $\rightarrow$ A sort of a "physics mirror", stating that the laws of physics should be the same if observed through said mirror.

- *CP* violation one of the necessary conditions for matter and antimatter asymmetry in the universe
- This analysis is one of the many steps toward a better understanding of our universe

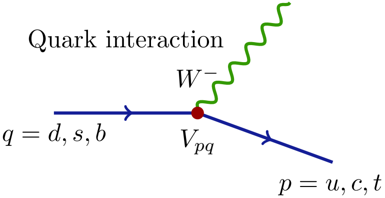


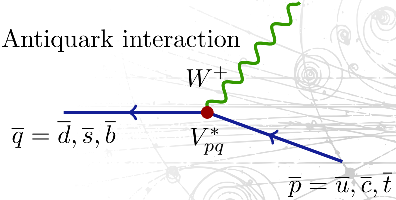
http://pprc.qmul.ac.uk/~still/homepage/Matter_and_Anti-Matter.html

Introduction III – CKM Matrix

- Quarks can transform to other quarks (weak nuclear interaction)
- Transition probabilities are contained in the complex Cabibbo-Kobayashi-Maskawa (CKM) matrix
- CKM matrix is strongly connected to *CP* violation

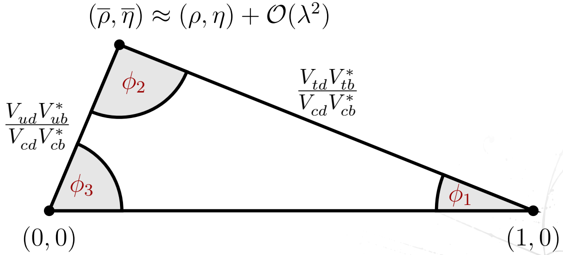
$$V_{CKM} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix}, \quad |V_{ij}|^2 \propto q_i \leftrightarrow q_j \ \text{transition probability}$$





Introduction IV – Unitarity Triangle

- The most relevant unitarity condition of the CKM matrix for this analysis is
$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$

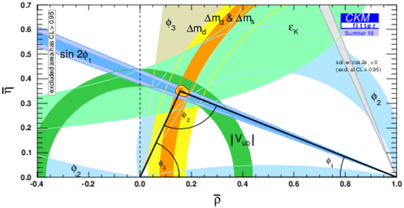


- By measuring the sides and angles of the triangle, we can overconstrain it and check if all the sides meet
- CKM matrix elements are not determined by theory or experiment alone, but by their joint effort

Introduction V – Unitarity Triangle Determination

- The most relevant unitarity condition of the CKM matrix for this analysis is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$

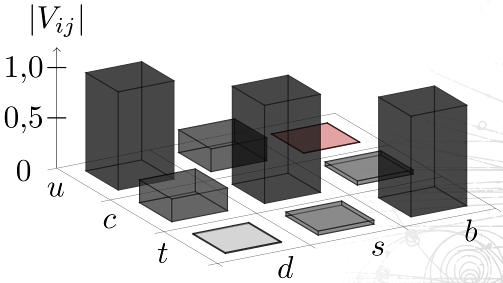


- By measuring the sides and angles of the triangle, we can overconstrain it and check if all the sides meet
- CKM matrix elements are not determined by theory or experiment alone, but by their joint effort
- In this analysis we focus on decays involving V_{ub}

Overview

Motivation I – Why V_{ub}?

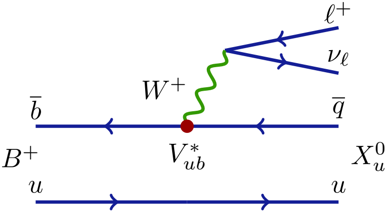
- The magnitude of *CP* violation that we know of is not large enough to account for the matter and antimatter asymmetry
- We are searching for new physics (NP) processes, which are not described by our current model
- $|V_{ub}|$ has the smallest value and the largest uncertainty of all the CKM matrix elements, precision measurements require better accuracy



Motivation II – Why B Mesons?

- B mesons exhibit a rich spectrum of decay modes, out of which many allow the study of the underlying physics processes
- Decays are deeply connected to the CKM matrix

- We focus on the charmless semileptonic B meson decays of the form $B^+ \to X_u^0 \ell^+ \nu_\ell$
- Such decays are used to determine the $|V_{ub}|$ CKM matrix element



Reliable experimental measurements along with precise theoretical calculations enable the determination of the $|V_{ub}|$

$$\mathrm{d}\Gamma \propto G_F^2 |V_{ub}|^2 |L^\mu \langle X_u | \bar{u} \gamma_u \frac{1}{2}(1-\gamma_5) b | B \rangle |^2,$$

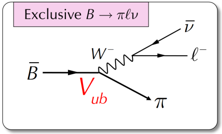
Γ is the decay width, L^μ is the leptonic current and $\langle \dots \rangle$ is the hadronic current.

Motivation III – Why This Analysis?

There are two common methods of $|V_{ub}|$ determination

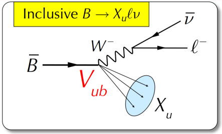
Exclusive

- B decays to a specific hadronic final state X_u (such as π or ρ)



Inclusive

- B meson decays to any hadronic final state X_u



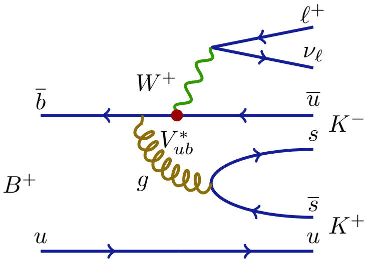
Both methods require different experimental and theoretical approaches, therefore yield largely independent results

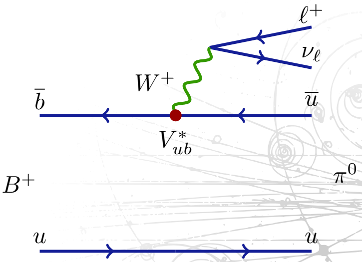
$$|V_{ub}|_{\text{excl.}} = (3.65 \pm 0.09 \pm 0.11) \times 10^{-3},$$
$$|V_{ub}|^{\text{GGOU}}_{\text{incl.}} = \left(4.52 \pm 0.15\,^{+0.11}_{-0.14}\right) \times 10^{-3},$$

However, they agree only at a 3σ level $\rightarrow$ The V_{ub} puzzle

Motivation IV – Why This Decay?

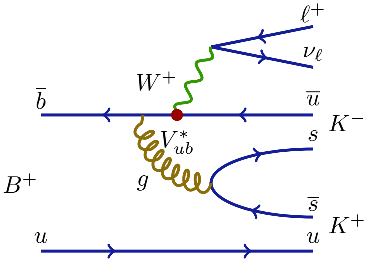
- The decay $B \to KK\ell\nu$ is similar to the $B \to \pi\ell\nu$ decay, so a similar analysis process can be applied
- The decay has not been observed yet
- Kaons ($K^+(u\bar{s})$) are usually present in $b \to c(\to s)$ decays, so a K-veto is used to remove such background cases in inclusive V_{ub} studies
  $\to$ but this is a charmless process ($b \to u$) with kaons in the final state!

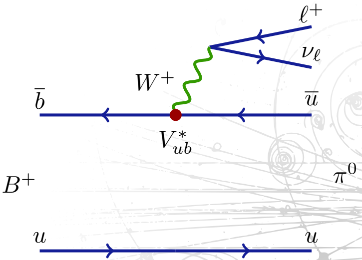




Motivation IV – Why This Decay?

- The decay $B \to KK\ell\nu$ is similar to the $B \to \pi\ell\nu$ decay, so a similar analysis process can be applied
- The decay has not been observed yet
- Kaons ($K^+(u\bar{s})$) are usually present in $b \to c(\to s)$ decays, so a K-veto is used to remove such background cases in inclusive V_{ub} studies
  $\to$ but this is a charmless process ($b \to u$) with kaons in the final state!

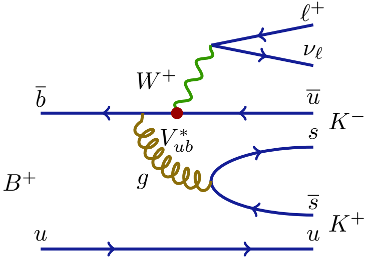


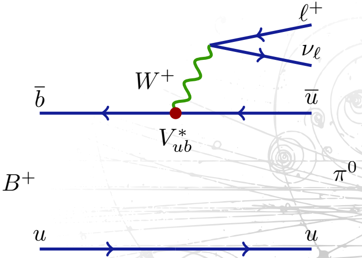


Can change the value of $|V_{ub}|$ with inclusive method!

Motivation IV – Why This Decay?

- The decay $B \to KK\ell\nu$ is similar to the $B \to \pi\ell\nu$ decay, so a similar analysis process can be applied
- The decay has not been observed yet
- Kaons ($K^+(u\bar{s})$) are usually present in $b \to c(\to s)$ decays, so a K-veto is used to remove such background cases in inclusive V_{ub} studies
  $\to$ but this is a charmless process ($b \to u$) with kaons in the final state!



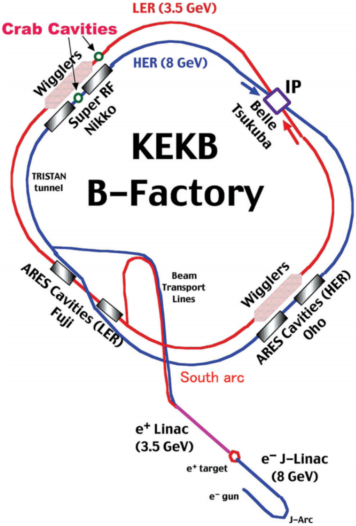


Impact of not taking these decays into account?

Overview

Experimental Setup I – KEKB Accelerator

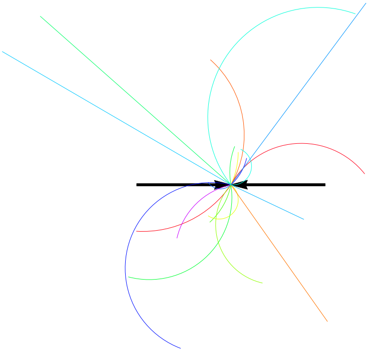
- Two rings (e^+ and e^-) with a diameter ≈ 1 km
- Particles are produced in events when e^+ and e^- collide
- Energy in the center-of-mass frame is 10.58 GeV
  - corresponds to $\Upsilon(4S)$ meson mass
- B factories are known for mass production of B mesons



Experimental Setup II – Particle Detection

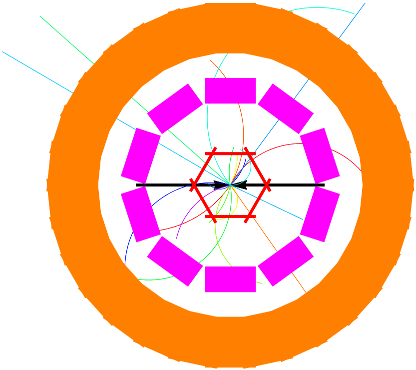
- Event – collision of e^+ and e^-

Experimental Setup II – Particle Detection



- Event – collision of e^+ and e^-
- Production of new particles

Experimental Setup II – Particle Detection



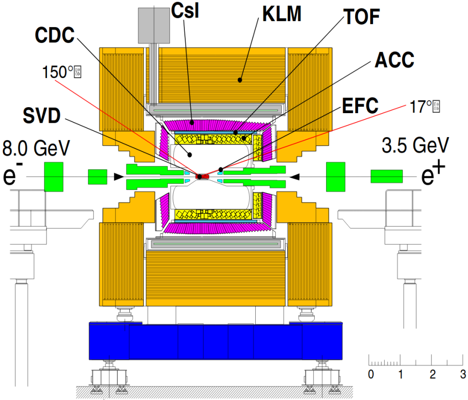
- Event – collision of e^+ and e^-
- Production of new particles
- Detection of final, stable particles

Experimental Setup III – Belle Detector

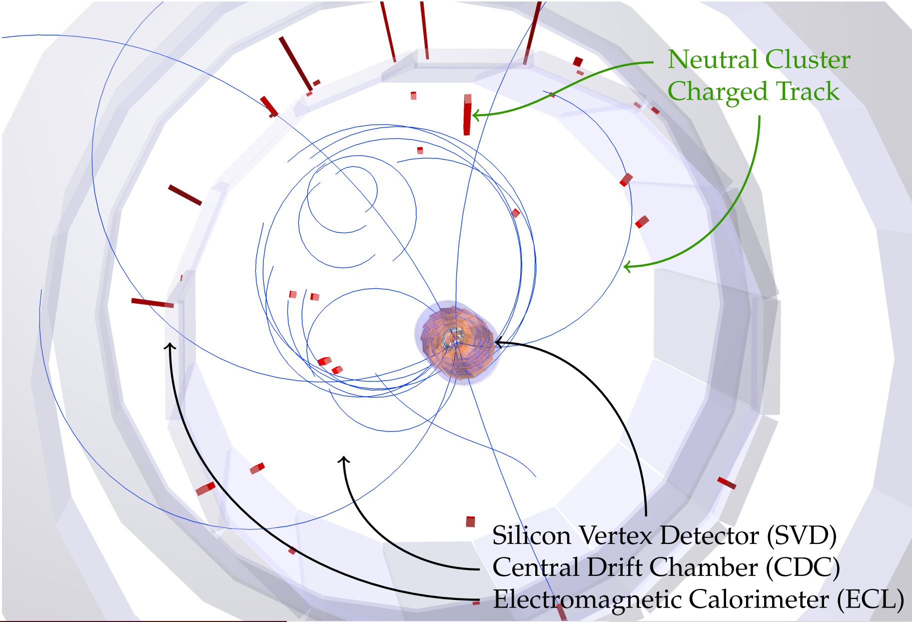
- A cylindrically symmetric magnetic spectrometer
- Wide solid angle coverage ($\sim 92\%$)
- Specialized for e^+e^- collisions
- Constructed from several subdetectors, each with its own purpose

Belle II detector subsystems:

- Decay vertex determination
- Tracking
- Particle identification
- Calorimetry



Experimental Setup III – Belle Detector Subsystems



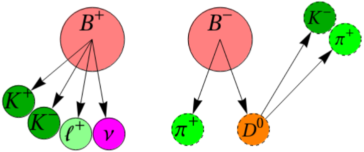
Overview

Analysis I – Method overview

- Initial state well known: $e^+e^- \to \Upsilon(4S)$ @ $E_{CMS} \approx M_{\Upsilon(4S)}$
- $\Upsilon(4S)$ at rest $\to B\overline{B}$
- Kaons K and the leptons e and μ produce tracks $\to$ easily detectable
- Neutrinos ν interact weakly and escape the detector $\to$ missing energy and momentum

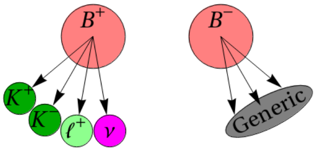
Reconstruction methods

Tagged measurement



Example scheme
of a tagged mode

Untagged measurement



Companion B
not reconstructed

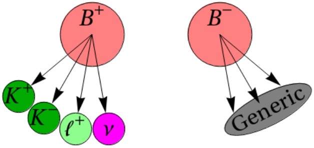
Analysis I – Method overview

- Initial state well known: $e^+e^- \to \Upsilon(4S)$ @ $E_{CMS} \approx M_{\Upsilon(4S)}$
- $\Upsilon(4S)$ at rest $\to B\overline{B}$
- Kaons K and the leptons e and μ produce tracks $\to$ easily detectable
- Neutrinos ν interact weakly and escape the detector $\to$ missing energy and momentum

Reconstruction methods

Untagged measurement

We opt for this method $\to$ Neutrino 4-momentum is inferred from missing momentum in event (assuming only 1 neutrino missing)



Companion B not reconstructed

Analysis II – Particle Reconstruction

Part I: Final State Particles (FSP)

- Charged particles like K, e and μ are reconstructed from tracks, a quality selection is performed
- Neutrinos are weakly interacting and escape detection

Part II: FSP Combinations

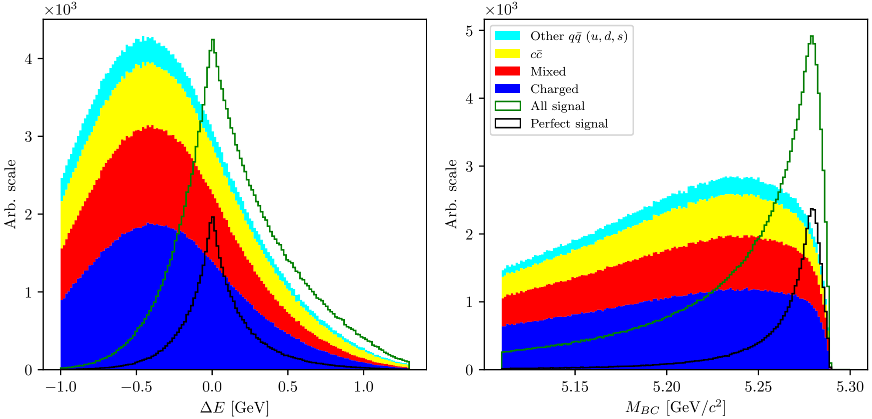
- Combinations of FSP particles KKe and $KK\mu$ represent first B meson candidates with a missing neutrino

Part III: Loose Neutrino Reconstruction

- Rest of Event (ROE) are all the tracks (charged particles) and clusters (neutral particles) which were not used in the reconstruction
- ROE is needed to reconstruct the missing neutrino momentum, where all tracks and clusters from ROE are summed up together

Analysis III – B meson specific variables

$$\Delta E = E_B - E_{CMS}/2, \quad M_{BC} = \sqrt{(E_{CMS}/2)^2 - |\vec{p}_B|^2}$$

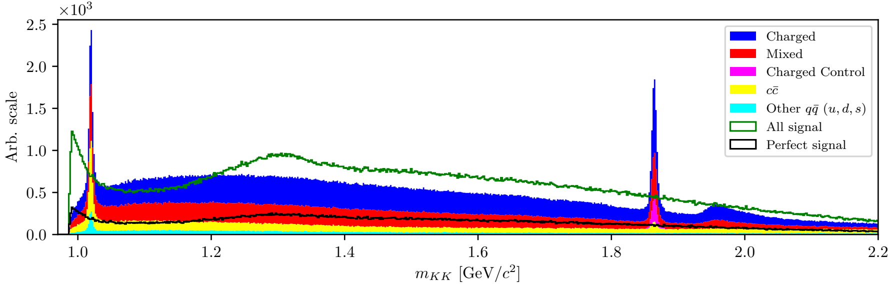


Signal is scaled up. Correctly reconstructed candidates have $\Delta E \approx 0$ and $M_{BC} \approx m_B$

Analysis IV – Control Decay

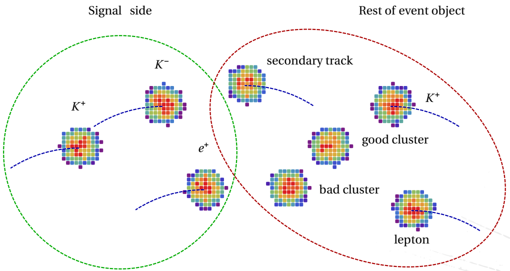
We define a control decay mode, similar to our signal decay mode.

- The decay mode is $B^+ \to \bar{D}^0 \ell^+ \nu$, $D^0 \to K^+ K^-$
- The properties of the control and signal decay are very similar
- Its purpose is to continuously check the consistency between simulation (Monte Carlo or MC) and measured data



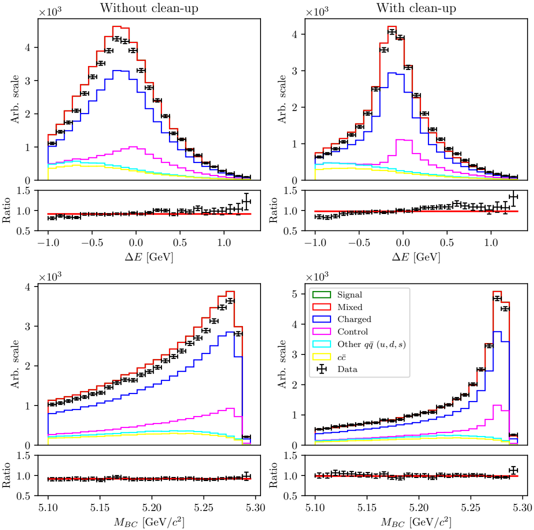
Analysis V – ROE Clean-up

- Rest of Event (ROE) are all the tracks (charged particles) and clusters (neutral particles) which were not used in the reconstruction
- ROE is needed to reconstruct the missing neutrino momentum



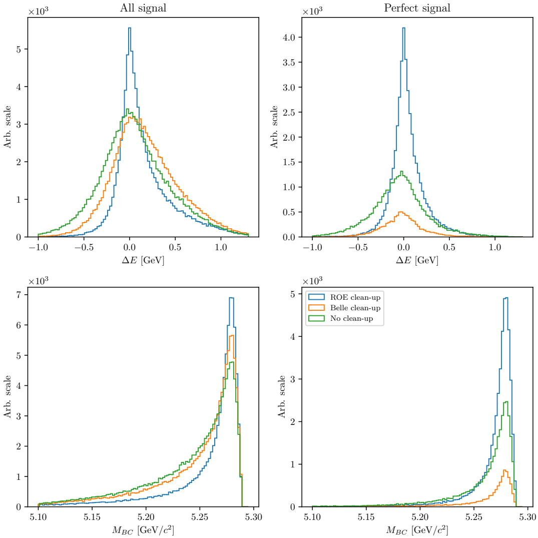
- Why? Extra tracks and clusters should not be taken into account when calculating the neutrino 4-momentum
- How? We apply machine learning in several steps of the ROE clean-up in order to efficiently clean it

Analysis VI – ROE Clean-up Results



- ROE clean-up performs well, distributions are more prominent
- The clean-up procedure does not introduce differences between MC and Data
- The control (and therefore signal) distribution is much more significant after the cleanup
- The charged background more "signal" like because phase-space similar to the one for the control decay

Analysis VI – ROE Clean-up Results



- The procedure **improves the resolution** of signal distribution
- It performs **better** than the standard Belle procedure
- **The efficiency** of perfectly reconstructed signal **increases**

	$\mathcal{E}$	FWHM
Belle	28.5 %	75.0 %
ROE	140.1 %	35.0 %

Perfectly reconstructed signal candidates are signal candidates from events, where all tracks and clusters are accounted for.

Analysis VII – Background Suppression

In order to further suppress various sources of background, we use *machine learning algorithms*:

These algorithms take multiple properties of candidates as input and produce an output variable, similar to a signal probability

Boosted Decision Trees (BDT) algorithms are commonly used in the field

Continuum Suppression

- Events of the form $e^+e^- \to q\bar{q}$, where $q \in [u, d, s, c]$
- Energy and momentum distribution of such events differ from $e^+e^- \to B\bar{B}$ events

$B\bar{B}$ Suppression

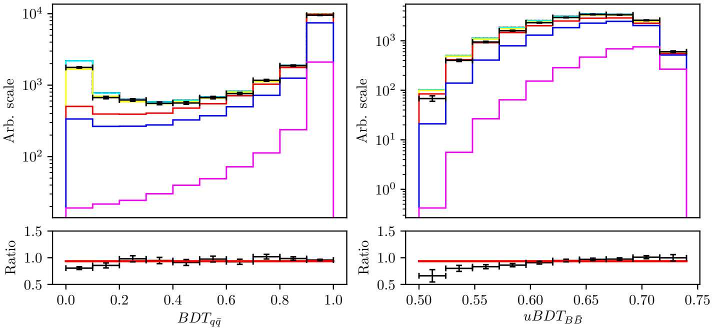
- Deeper look into the properties of the signal decay enable separation of signal from other $e^+e^- \to B\bar{B}$ events

Analysis VII – Background Suppression

In order to further suppress various sources of background, we use *machine learning algorithms*:

These algorithms take multiple properties of candidates as input and produce an output variable, similar to a signal probability

Boosted Decision Trees (BDT) algorithms are commonly used in the field

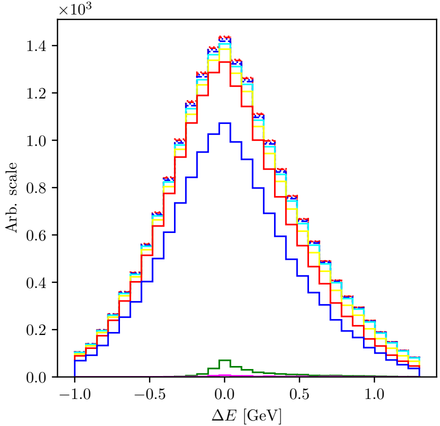


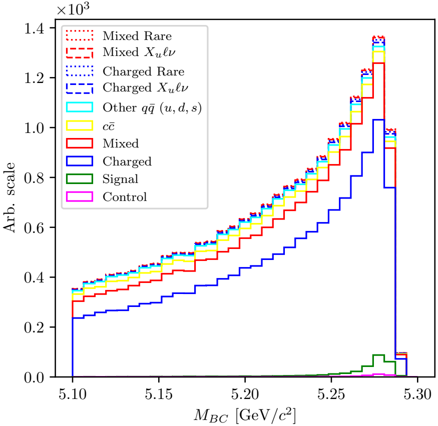
Continuum (left) and $B\bar{B}$ (right) suppression BDT variable. Signal is more likely distributed on the right side of the variables.

Analysis VIII – Final Selection

Sample composition after background suppression:

$$N_{\text{sig}} = 264, \quad \frac{N_{\text{sig}}}{N_{\text{bkg}}} = 1.33$$





Overview

Parameter Extraction I – Fit Method

- Define "templates" which describe distribution shapes (signal, background, . . .)
- For each histogram bin define Poisson probability for it's content w.r.t. the measurement
- Obtain the combination of template yields which maximise the probability

Templates in the signal fit:

- signal
- continuum
- well-defined sources →
- other $B\bar{B}$ background

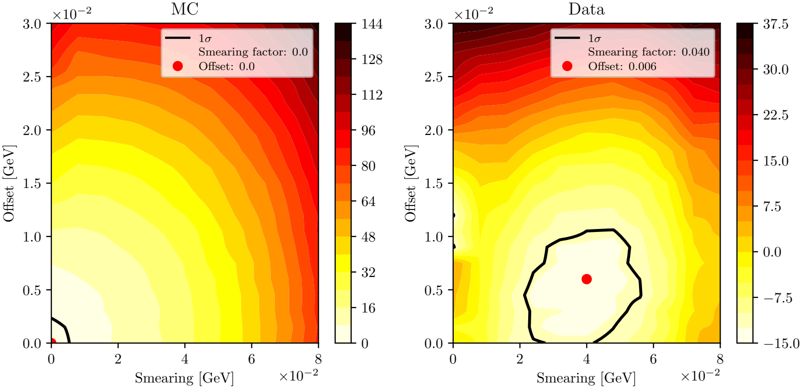
- control decay,
- $B \to \bar{D}^{*} \ell^{+} \nu,\ D^{0} \to K^{-} K^{+},$
- $B \to \bar{D}^{(*)} \ell^{+} \nu,\ D^{0} \to K^{-} \pi^{+},$
- $B \to \bar{D}^{(*)} \ell^{+} \nu,\ D^{0} \to K^{-} K^{+} \pi^{0},\ K^{-} \pi^{+} \pi^{0},$
- $B \to \bar{D}^{(*)} \ell^{+} \nu,\ D^{0} \to K^{-} \ell^{+} \nu,$
- $B^{0} \to D^{(*)-} \ell^{+} \nu,\ D^{+} \to K^{-} K^{+} \pi^{+},\ K^{-} \pi^{+} \pi^{+},$
- other $B \to \bar{D}^{(*)} \ell^{+} \nu$ decays,

Yields of all templates are floated, except in the well-defined cases, where yields are constrained by world measurements.

Parameter Extraction II – Smearing and Offset

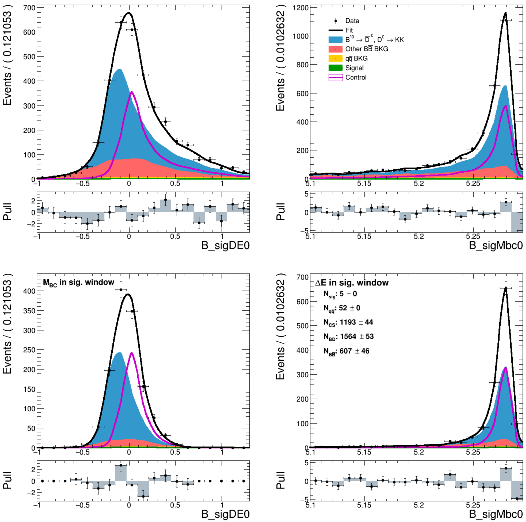
We introduced 2 additional parameters to the ΔE variable:

$$f_{\text{offset}} : x \mapsto x + a, \quad f_{\text{smearing}} : x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},$$



Consistency check on MC, data fit better with introduced parameters:
Smearing: 40^{+15}_{-17} MeV, Offset: $6^{+4.6}_{-6}$ MeV.

Results of the control decay fit to DATA

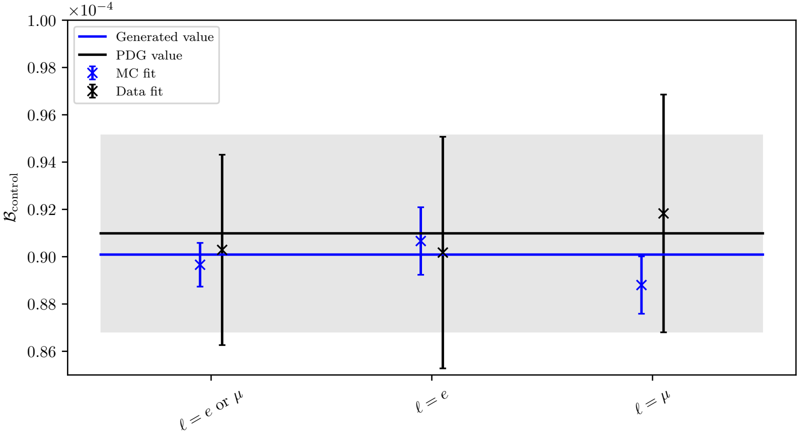


- Templates are appropriately summed to best fit the measurement
- The fit behaves well, no strange artefacts in pulls

pulls ∝ differences between the function and the points

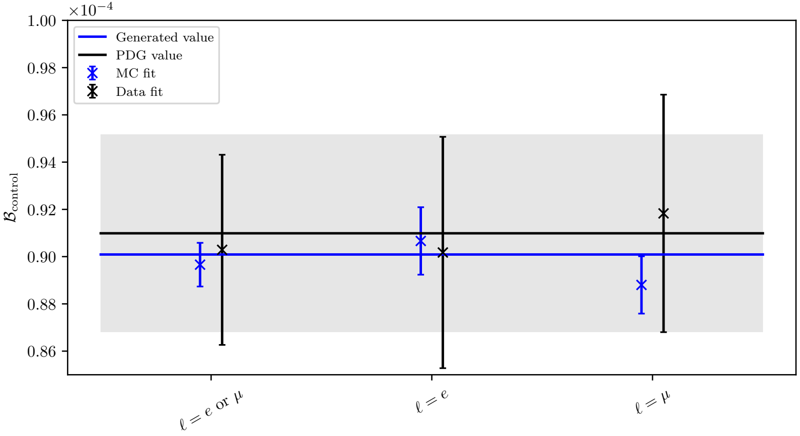
Parameter Extraction III – Control Fit Results

Simulated and measured branching ratio of the control decay for joint and split lepton modes.



Parameter Extraction III – Control Fit Results

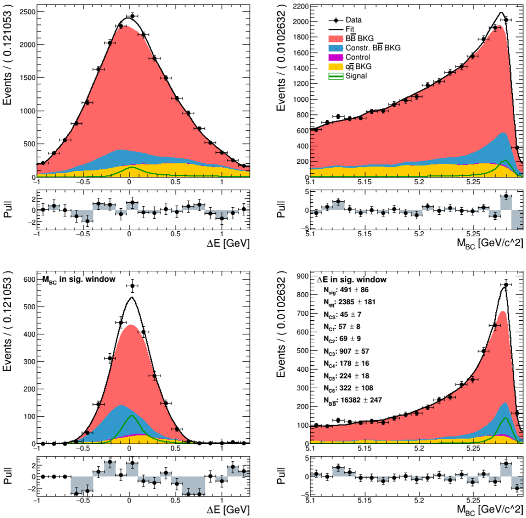
Simulated and measured branching ratio of the control decay for joint and split lepton modes.



The simulated and measured values of the control decay branching ratio seems to agree well with the generated value and the world average from PDG (`http://pdglive.lbl.gov`).

Parameter Extraction IV – Signal Fit Results

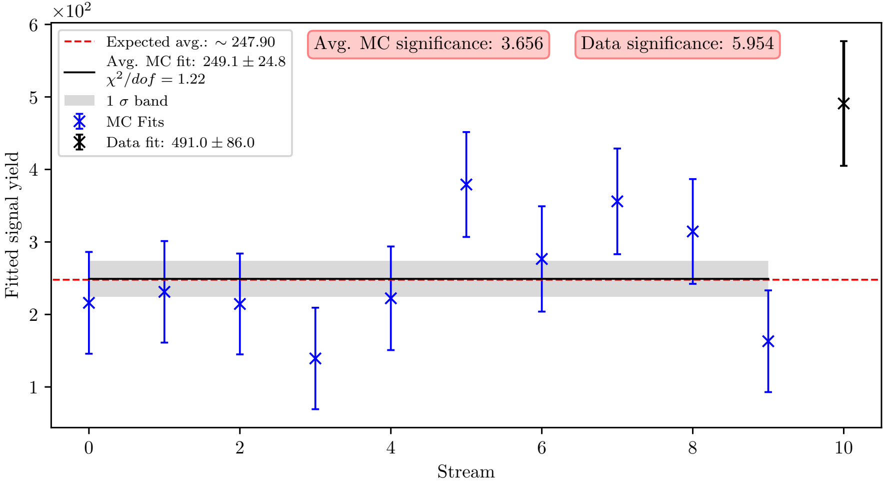
Results of the signal decay fit to DATA



Category	Fit Yield
Signal	491 ± 86
$q\bar{q}$ bkg	2385 ± 181
C_0	45 ± 7
C_1	57 ± 8
C_2	69 ± 9
C_3	907 ± 57
C_4	178 ± 16
C_5	224 ± 18
C_6	322 ± 108
Other $B\bar{B}$ bkg	16382 ± 247

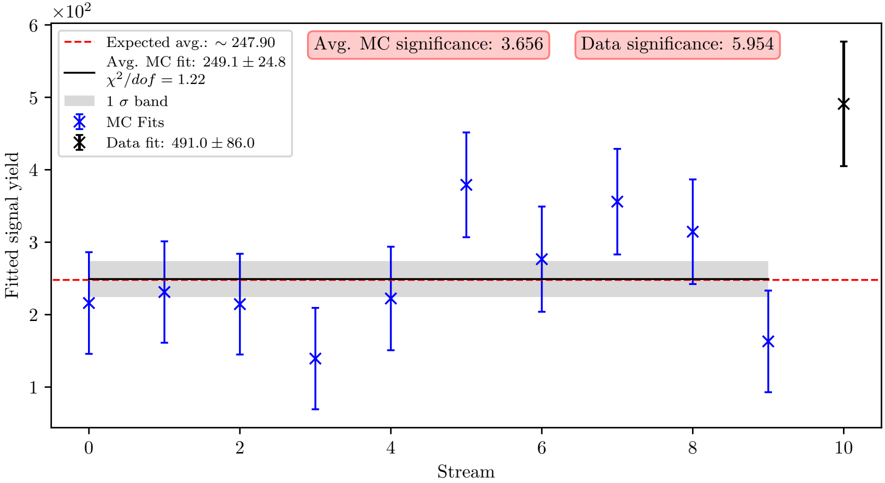
Parameter Extraction IV – Signal Fit Results

Signal fit to measured data and 10 equal samples of simulated data (streams)



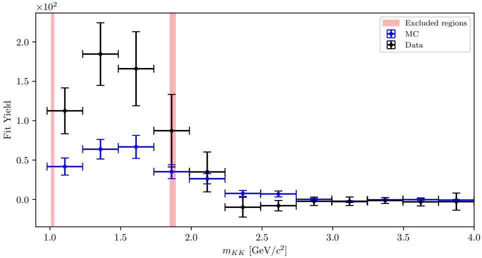
Parameter Extraction IV – Signal Fit Results

Signal fit to measured data and 10 equal samples of simulated data (streams)

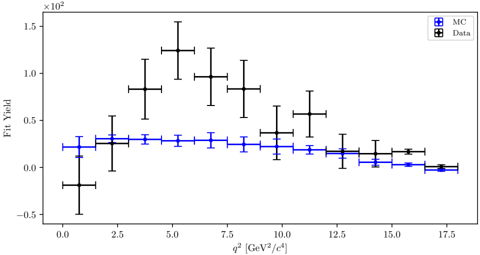


Much more signal in measured data than expected from simulation!

Parameter Extraction V – Signal Distributions



- Invariant mass of two kaons
- Distributions are similar
- Signal much more abundant in DATA



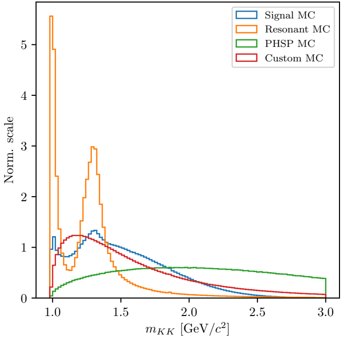
- q is the 4-momentum transferred to the lepton pair
- Distributions are quite different
- Signal much more abundant in DATA

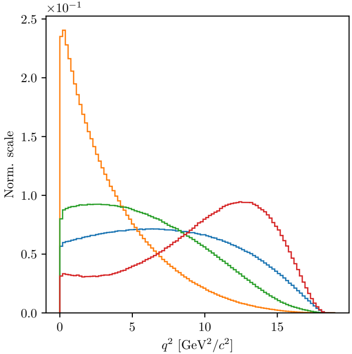
Such measurements help improve future models and minimize difference between data and simulations.

Overview

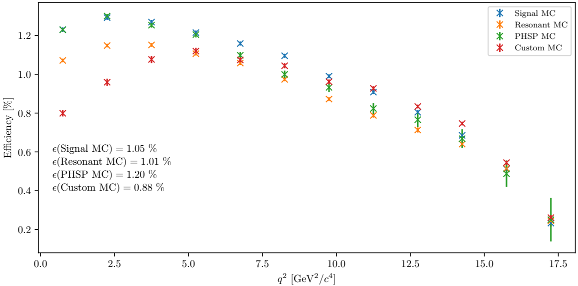
Systematic Uncertainties I – Model Dependency

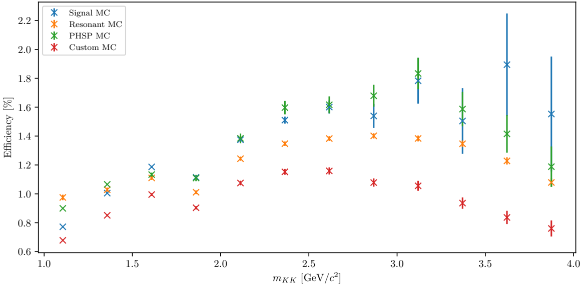
- Largest source of uncertainty in this analysis
- `ISGW2` model for signal decay is not the most reliable, try to not depend on it
- Three additional signal models are used to estimate dependency, used as substitute templates
- Different models have different signal efficiencies





Systematic Uncertainties I – Model Dependency





Systematic Uncertainties II – All Sources

Source	σ	δ [%]
PID	10	2.0
Fit Bias	$^{+7}_{-10}$	$^{+1.5}_{-2.0}$
Gaussian Constraints	26	5.4
Template Smearing	$^{+41}_{-33}$	$^{+8.3}_{-6.7}$
Template Offset	$^{+41}_{-31}$	$^{+8.4}_{-6.3}$
Finite MC Effects	26	5.3
MVA Selection	5	1.0
Model Shape	$^{+45}_{-39}$	$^{+9.3}_{-8.0}$
Model Efficiency	$^{+70}_{-79}$	$^{+14.3}_{-16.2}$
Total	$^{+109}_{-107}$	$^{+22.2}_{-21.9}$

Overview

Results I – Branching Ratio of Signal Decay

Calculated branching ratio from simulated data is:

$$\mathcal{B}^{\mathrm{MC}}(B^+ \to K^+K^-\ell^+\nu) = (1.55 \pm 0.15) \times 10^{-5},$$

And on data, after taking all uncertainties into account:

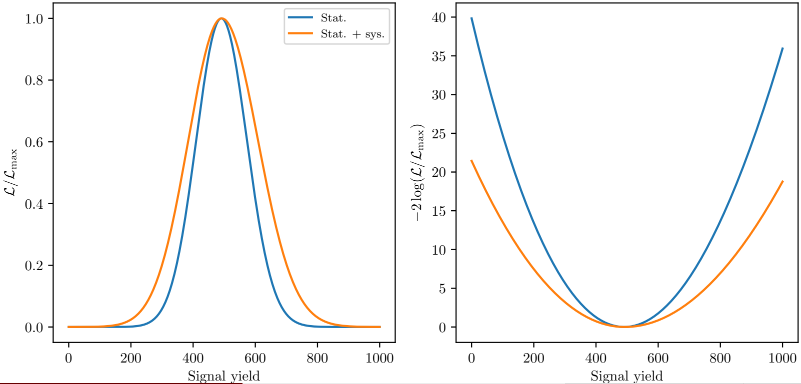
$$\mathcal{B}(B^+ \to K^+K^-\ell^+\nu) = (3.04 \pm 0.51 \pm {}^{+0.67}_{-0.66}) \times 10^{-5},$$

where the first and the last error are statistical and systematic, respectively.

Almost a factor of 2 difference!

Results II – Signal Significance

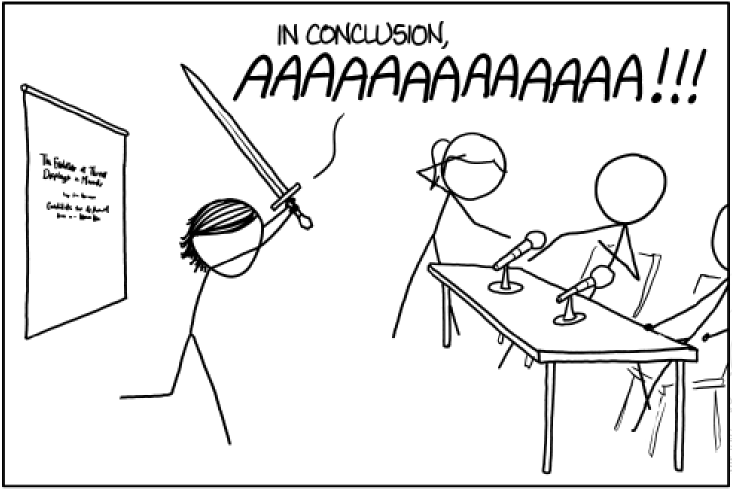
- Profile likelihood gives the value of likelihood at different expected yields of signal, used for significance estimation
- Statistical significance corresponds to 6.3σ
- Systematic uncertainties are incorporated via a convolution
- Overall signal significance $4.6\sigma \rightarrow$ evidence!



Summary

- We present an improvement in the untagged method with the ROE clean-up
- Machine learning algorithms aid substantially in such complex analyses
- The evidence for the decay mode $B^+ \to K^+K^-\ell^+\nu_\ell$ was found
- Further insight into signal distributions helps to improve future simulation models
- **Discarding decays of the form $B \to X_u\ell\nu$ with kaons in the final state might affect the inclusive measurements more than previously believed.**

Thank you!



THE BEST THESIS DEFENSE IS A GOOD THESIS OFFENSE.

https://xkcd.com/1403/

BACKUP

Analysis III – Selection Criteria

Two modes are reconstructed, neutrinos escape detection:
$$B^+ \to K^+K^-e^+ \quad \text{and} \quad B^+ \to K^+K^-\mu^+$$

Selection:

- FSP particles:
  - electrons: $|d_0| < 0.1$ cm, $|z_0| < 1.5$ cm, $p > 0.6$ GeV$/c$, $p_{CMS} \in [0.4,\ 2.6]$ GeV$/c$, $eID > 0.9$,
  - muons:
    $|d_0| < 0.1$ cm, $|z_0| < 1.5$ cm, $p_{CMS} \in [0.6,\ 2.6]$ GeV$/c$, $\mu ID > 0.97$,
  - kaons: $|d_0| < 0.15$ cm, $|z_0| < 1.5$ cm, $p_{CMS} < 2.5$ GeV$/c$, $K/\pi\ ID > 0.6$, $K/p\ ID > 0.1$,
- B meson candidates:
  - standard selection:
    $P(\chi^2,\ DOF) > 6 \times 10^{-3}$, $|\cos\theta_{BY}| < 1.05$, $|m^2_{miss}| < 0.975$ GeV$/c^2$,
  - fit region selection:
    $\Delta E \in [-1.0, 1.3]$ GeV, $M_{BC} \in [5.1, 5.295]$ GeV$/c^2$,
  - signal region selection: $|\Delta E| < 0.126$ GeV, $M_{BC} > 5.271$ GeV$/c^2$,
  - charge categorization: $q_{B^\pm} q_{B^\mp} = -1$.