Measurement of the angular decay rate of $B^0 \rightarrow K^{*0}ee$ using full available LHCb data samples

Maurice Geijsen

Master thesis defense

Supervisors
Mara Senghi Soares
Alice Biolchini

Nikhef, Amsterdam

13 July 2023

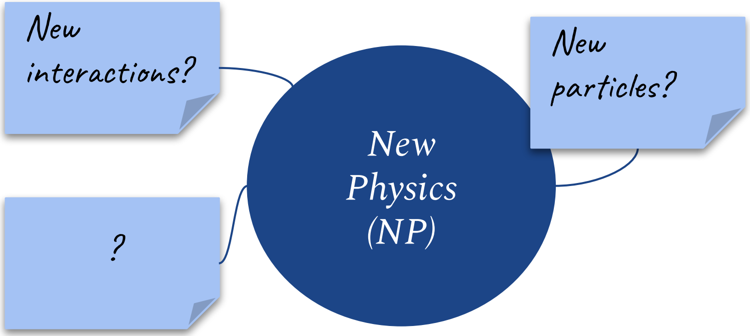
GRavitation AstroParticle Physics Amsterdam

Overview

- Introduction
- Analysis Overview
- Boosted Decision Trees (BDT)
- Semileptonic Background Studies
- Preliminary Mass Fit
- Angular Fit
- Outlook

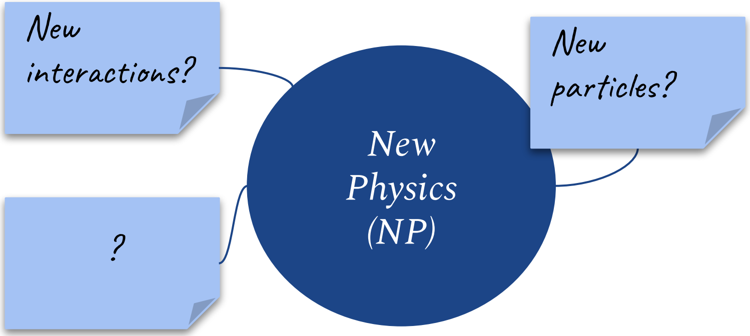
Standard Model (SM)

- SM is an excellent theory for describing particles and their interactions
- Still, it leaves a lot of open questions
- Testing SM predictions to look for New Physics (NP)



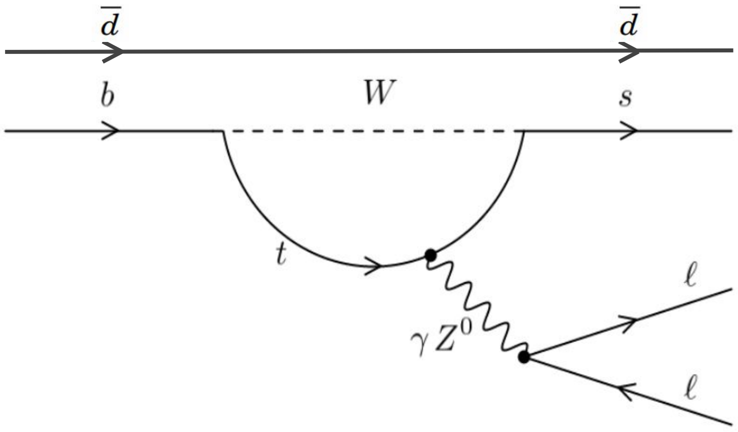
Standard Model (SM)

- SM is an excellent theory for describing particles and their interactions
- Still, it leaves a lot of open questions
- Testing SM predictions to look for New Physics (NP)



Testing the SM using *B* meson decays

- Decays of *B* mesons are great channels to probe the SM for New Physics (NP)
- We look at $B^0 \to K^{*0} ee$
- This is a very rare decay!
  BR of ~ 1.03×10^{-6}!

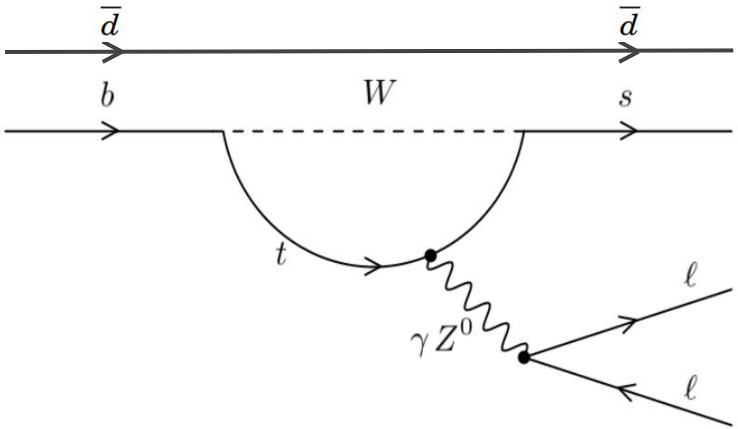


Branching Ratios

Angular coefficients

Testing the SM using *B* meson decays

- Decays of *B* mesons are great channels to probe the SM for New Physics (NP)
- We look at $B^0 \to K^{*0} ee$
- This is a very rare decay!
  BR of ~ 1.03×10^{-6}!



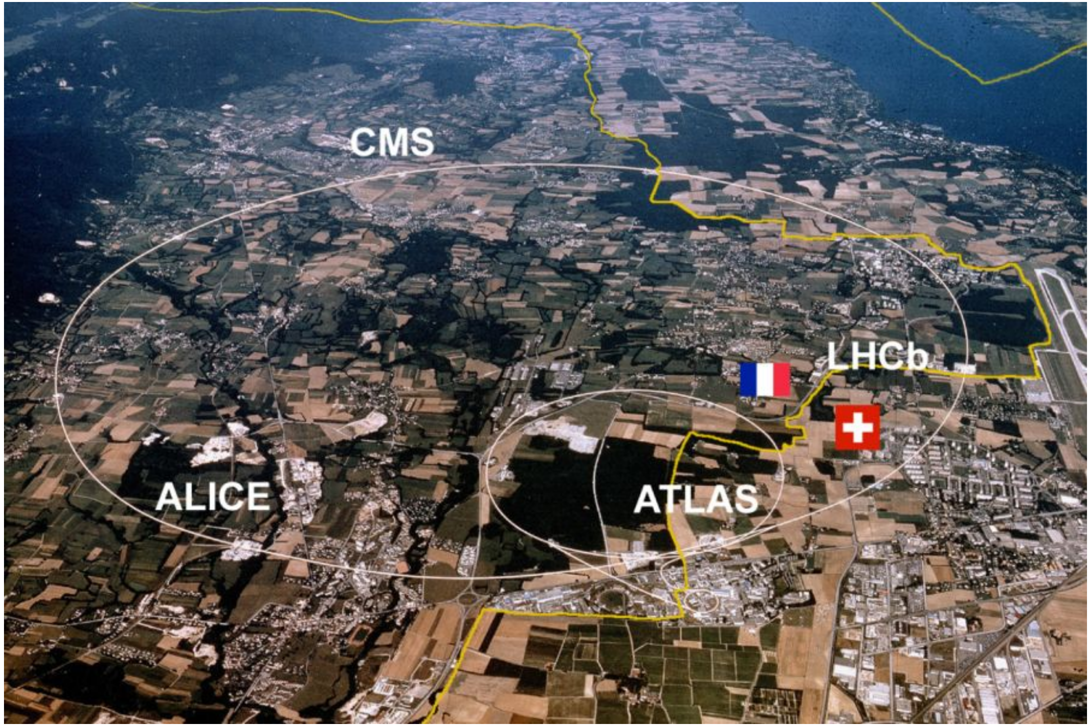
Branching Ratios

Angular coefficients

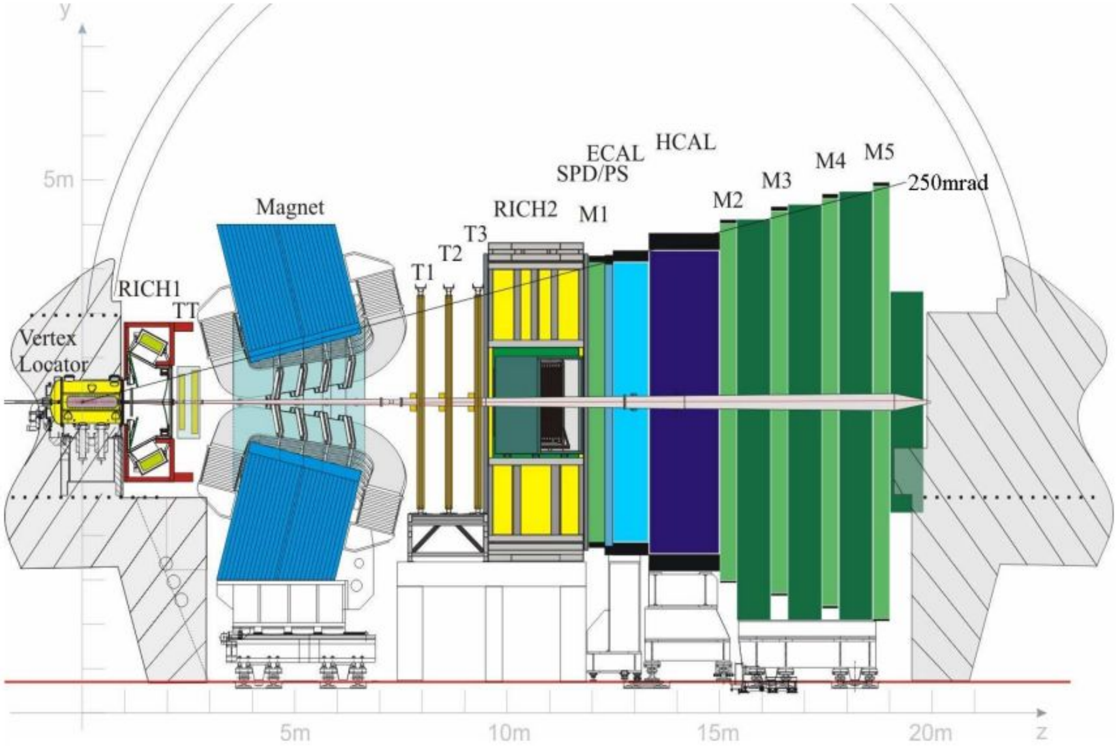
- Compare the SM predictions with the measurements
- Any discrepancies could be a sign of NP

Large Hadron Collider (LHC)

- 27 km circular particle accelerator at CERN
- Proton-proton collisions
- 4 beam collision points
  - ATLAS
  - ALICE
  - CMS
  - LHCb
- 2 data taking periods (Runs 1 & 2)
  - Run 3 ongoing

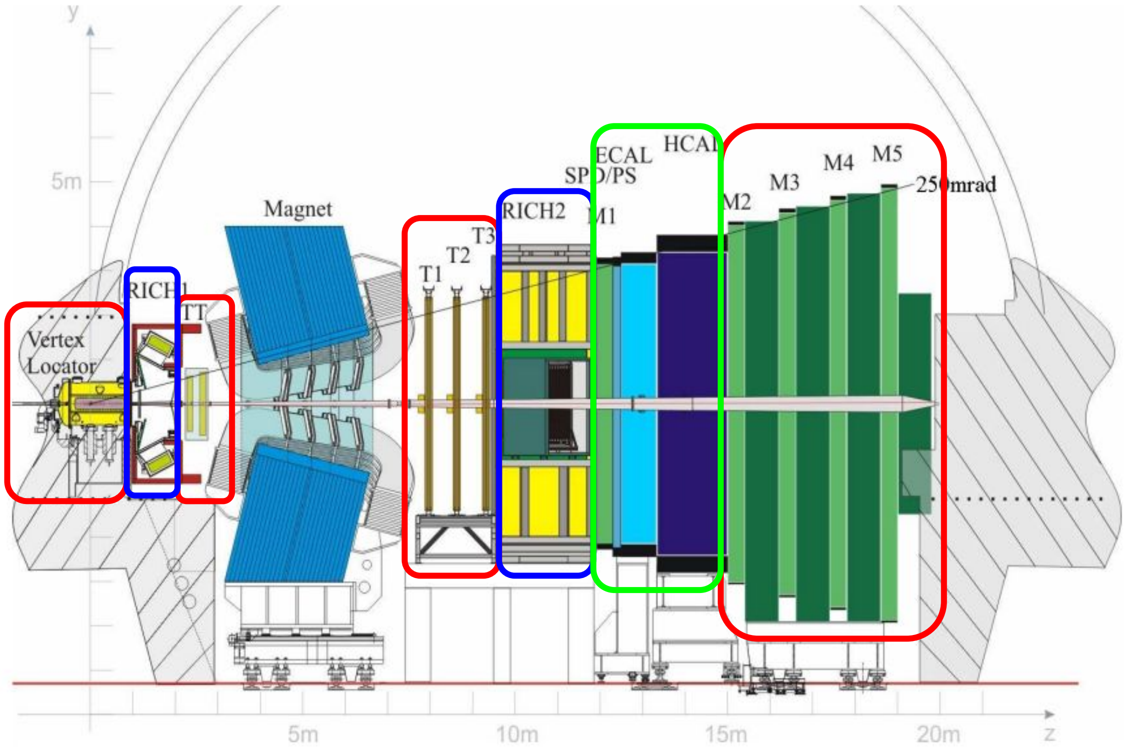


Large Hadron Collider beauty (LHCb)



- Beauty (bottom) quark dedicated experiment
- Composed of several subdetectors

Large Hadron Collider beauty (LHCb)

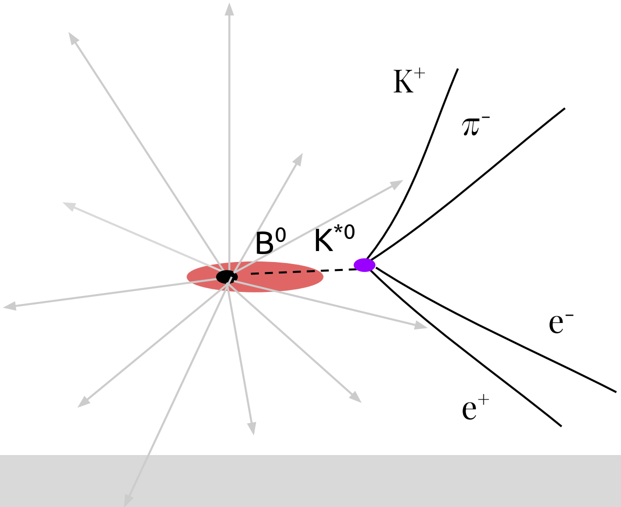


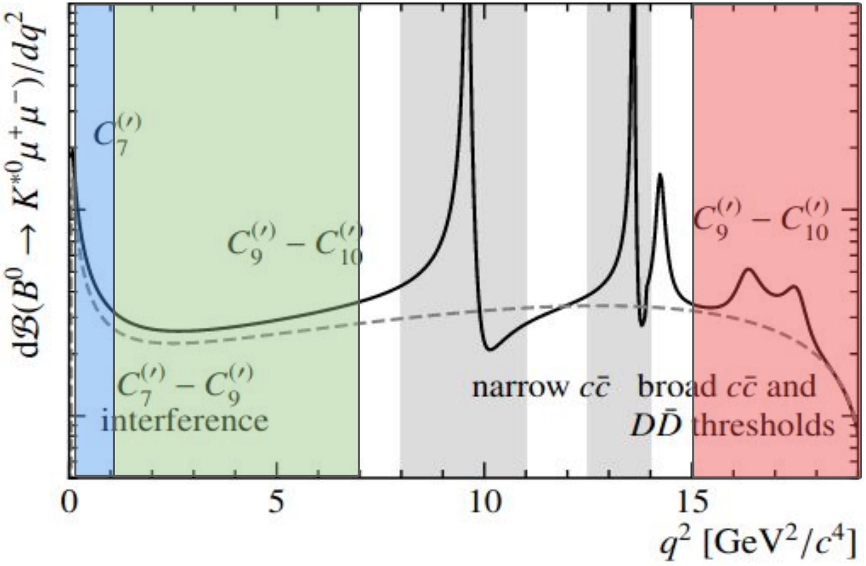
- Beauty (bottom) quark dedicated experiment
- Composed of several subdetectors

- Detect **position** & **momentum** of particle
- Measures **energy** deposition
- Particle **identification** (PID)

$B^0 \rightarrow K^{*0}ee$ topology

- K^{*0} is not stable and decays immediately to K and π
  - We're looking for $B^0 \rightarrow K^+\pi^-e^+e^-$
- q^2 is the di-electron invariant mass
- Data split up into different q^2 regions

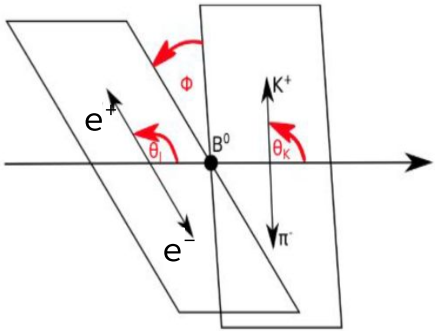




Goal of the analysis

- The decay is described by $\theta_\ell, \theta_K, \phi$ and q^2

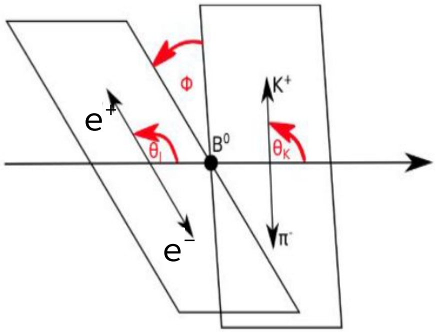
$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\mathrm{d}\vec{\Omega}} = \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-F_L)\sin^2\theta_K + F_L\cos^2\theta_K$$
$$+\tfrac{1}{4}(1-F_L)\sin^2\theta_K\cos 2\theta_\ell$$
$$-F_L\cos^2\theta_K\cos 2\theta_\ell + S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi$$
$$+S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi + S_5\sin 2\theta_K\sin\theta_\ell\cos\phi$$
$$+\tfrac{4}{3}A_{FB}\sin^2\theta_K\cos\theta_\ell + S_7\sin 2\theta_K\sin\theta_\ell\sin\phi$$
$$+S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi + S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]$$



Goal of the analysis

- The decay is described by $\theta_\ell, \theta_K, \phi$ and q^2

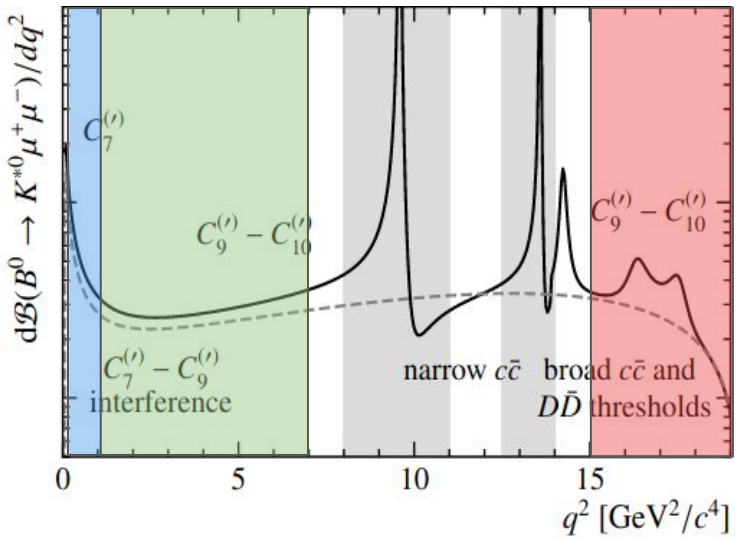
$$\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} = \frac{9}{32\pi}\Big[\tfrac{3}{4}(1-F_L)\sin^2\theta_K + F_L\cos^2\theta_K + \tfrac{1}{4}(1-F_L)\sin^2\theta_K\cos 2\theta_\ell - F_L\cos^2\theta_K\cos 2\theta_\ell + S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi + S_5\sin 2\theta_K\sin\theta_\ell\cos\phi + \tfrac{4}{3}A_{FB}\sin^2\theta_K\cos\theta_\ell + S_7\sin 2\theta_K\sin\theta_\ell\sin\phi + S_8\sin 2\theta_K\sin 2\theta_\ell\sin\phi + S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big]$$



- **Goal of the analysis:** Measure the coefficients describing the angular distribution
  - Compare with SM predictions
  - Perform an angular fit
  - Do this for each q^2 bin

Analysis strategy

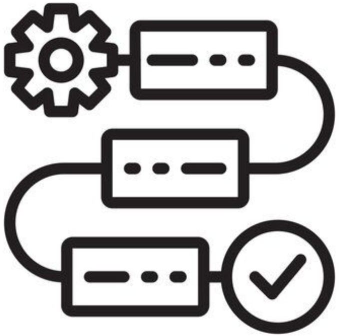
- Run 1 and Run 2 LHCb data
  - Structured in nTuples
- Selection of data very important
  - Separate signal from backgrounds
- Analysis performed in bins of q^2
  - Low q^2: [0.1, 1] GeV^2/c^4
  - Center q^2: [1, 7] GeV^2/c^4
- Make use of Monte Carlo (MC) simulations



Workflow and production of nTuples

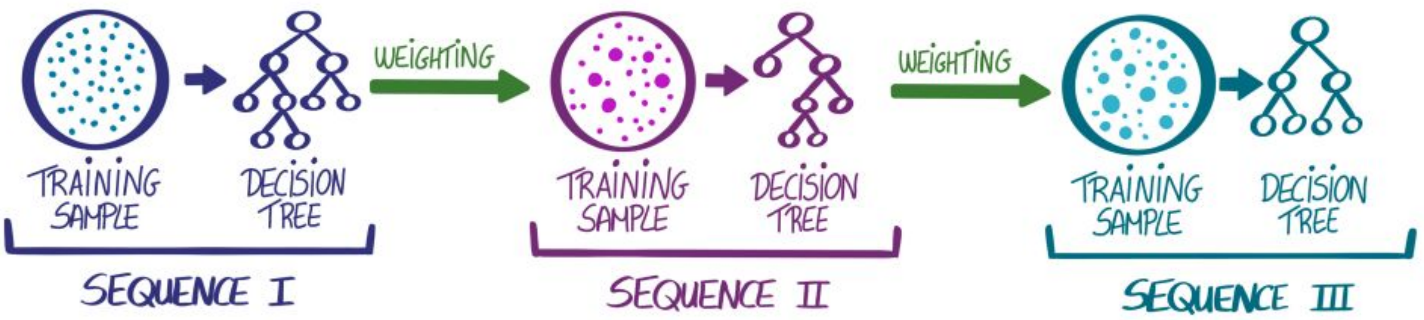
- Workflow
  - Perform corrections to MC
  - Selection of events
- Has some bottlenecks
  - Improve them!
  - Fix issues that come up
- Create nTuples!

snakemake



Nikhef

Boosted Decision Trees (BDT)

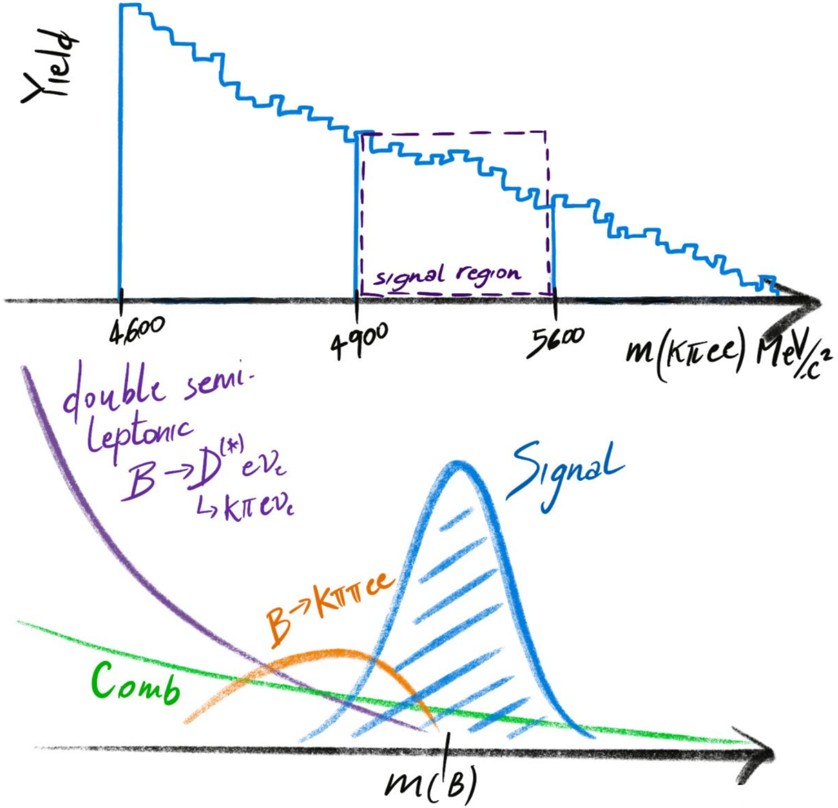


Silvia Ferreres Solé

- Machine Learning tool for separating signal from background
- Identifies patterns in features to predict signal
- Assigns a score to each event based on their ‘signal-likeness’

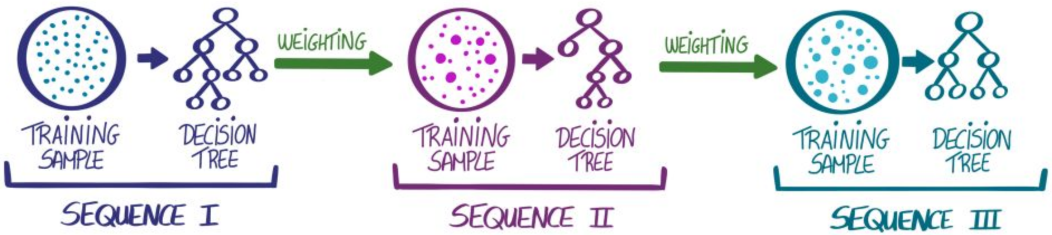
BDT Goals

- Goal is to reduce background while retaining signal
- Train the BDT using training samples
    - MC signal
    - Lower and upper sidebands of data



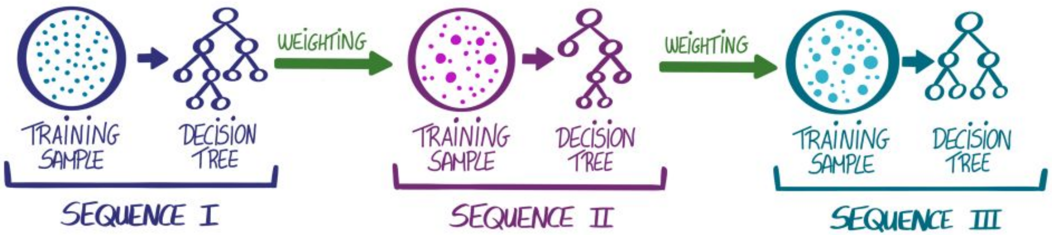
BDT performance and overtraining

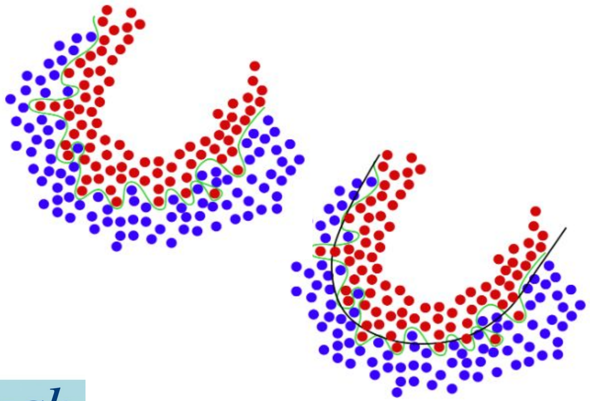
- BDT's performance is based on:
  - the features
  - the hyperparameters
  - the training samples



BDT performance and overtraining

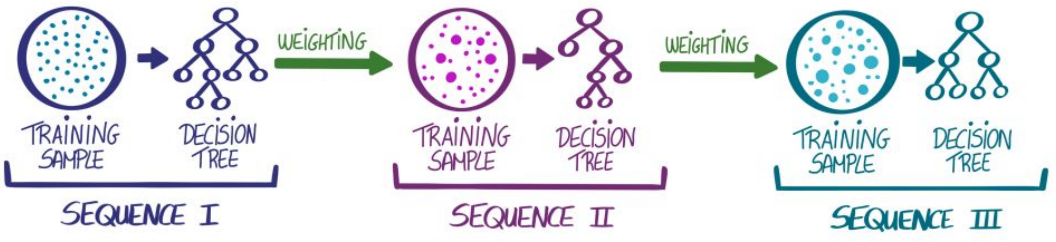
- BDT's performance is based on:
  - the features
  - the hyperparameters
  - the training samples
- Particularly, we want to avoid overtraining

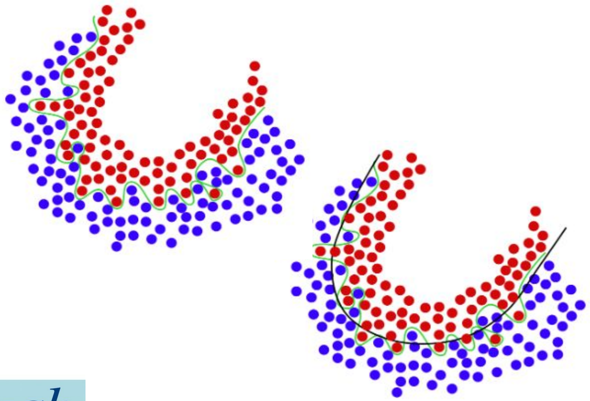




BDT performance and overtraining

- BDT's performance is based on:
  - the features
  - the hyperparameters
  - the training samples
- Particularly, we want to avoid overtraining



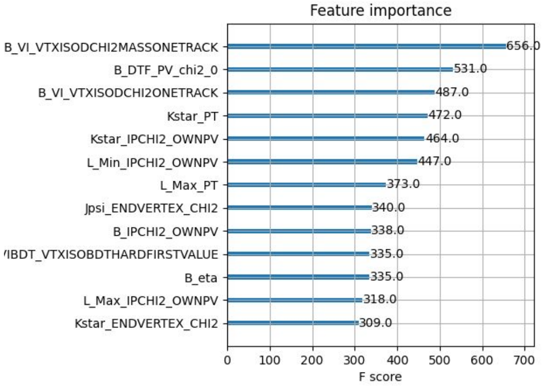


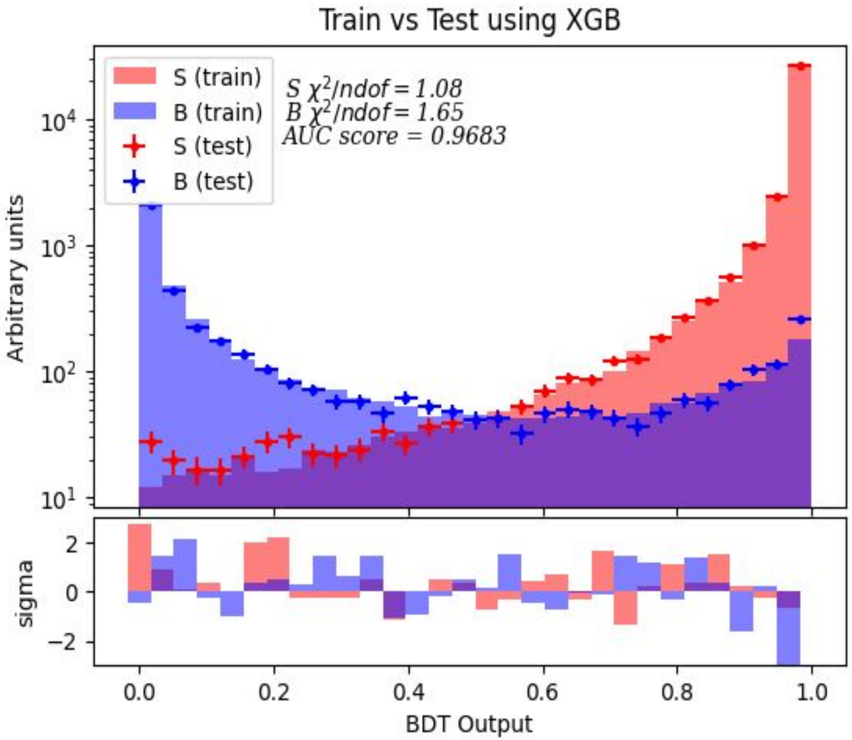
- Tuning the hyperparameters, target the BDT's complexity

Old BDT starting point

- Old BDT as starting point
- XGBoost from scikit library

n_estimators = 800
max_depth = 3
learning_rate = 0.06
subsample = 0.5
min_child_weight = 1

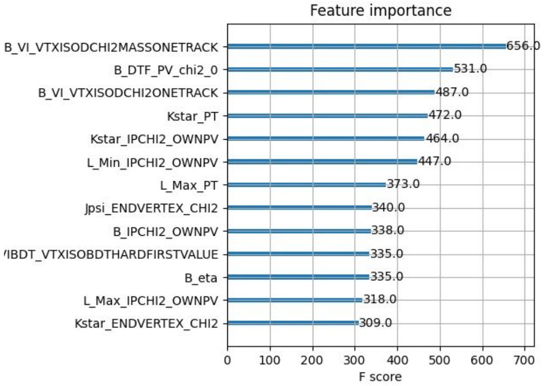


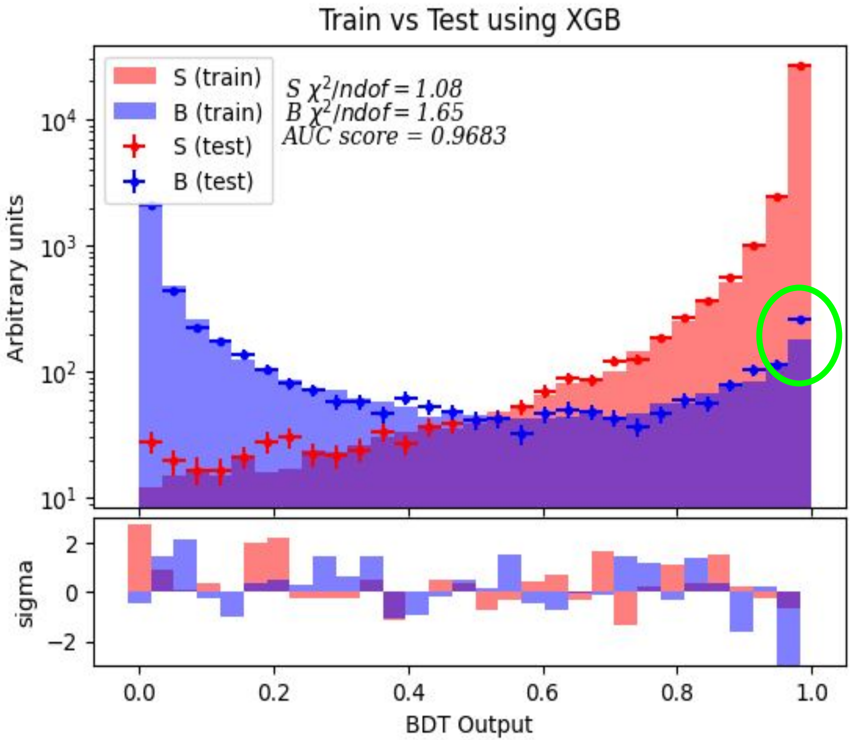


Old BDT starting point

- Old BDT as starting point
- XGBoost from scikit library
- We spot overtraining!

n_estimators = 800
max_depth = 3
learning_rate = 0.06
subsample = 0.5
min_child_weight = 1

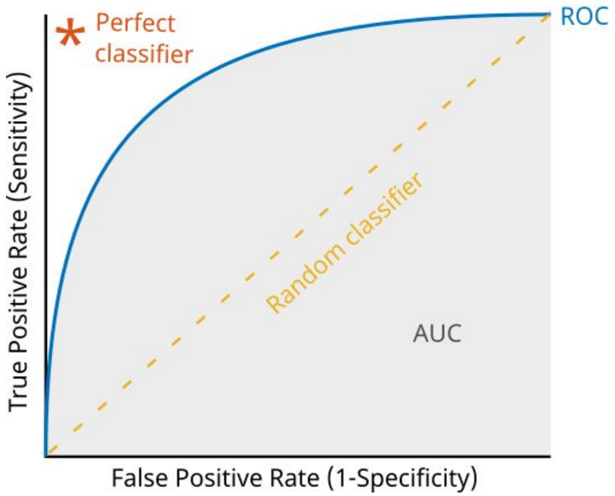




Optimizing BDT

- We tune hyperparameters
- Choose those that target the BDT's complexity
- Each combination of values gets assigned an AUC score

n_estimators
max_depth
learning_rate
subsample
min_child_weight

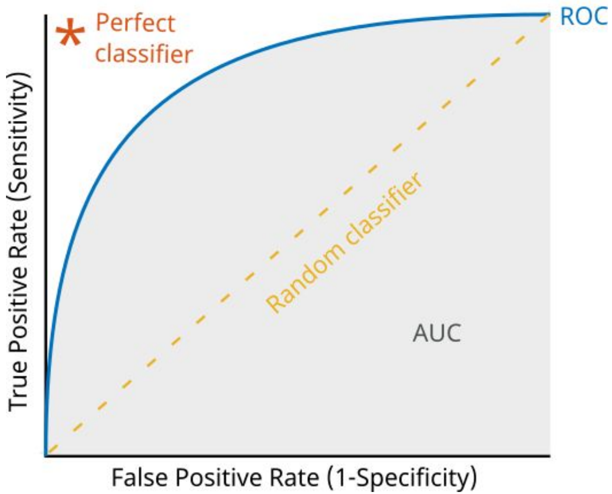


Optimizing BDT

- We tune hyperparameters
- Choose those that target the BDT's complexity
- Each combination of values gets assigned an AUC score

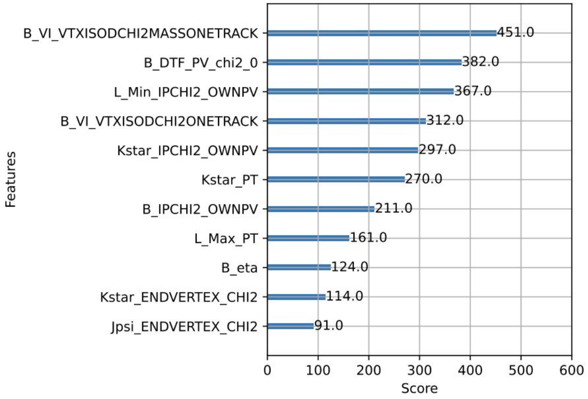
- Use GridSearchCV function to find the best combination given a range
  - Find the highest scoring combination!

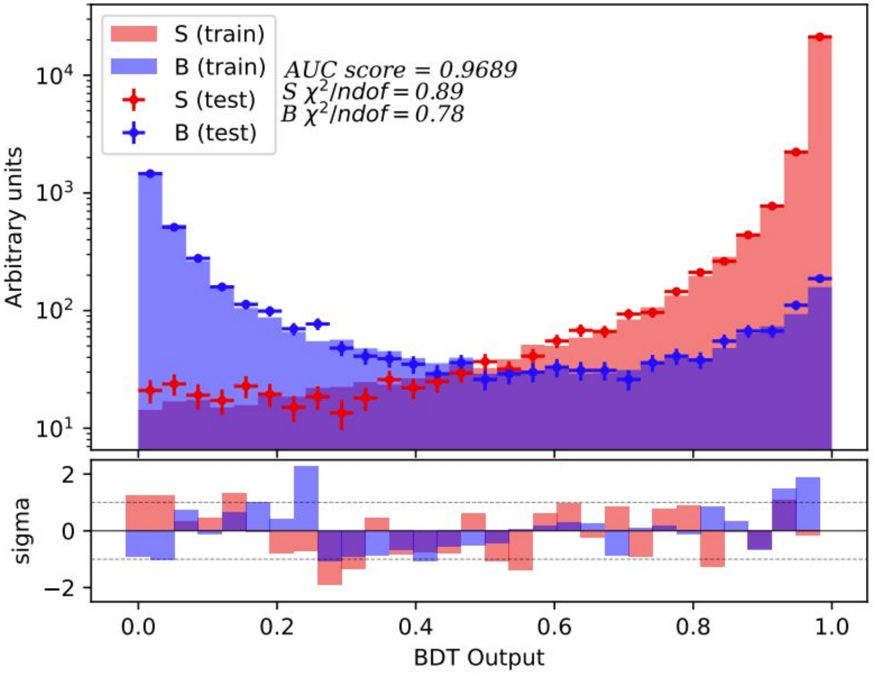
n_estimators	=	800	200
max_depth	=	3	4
learning_rate	=	0.06	0.05
subsample	=	0.5	0.5
min_child_weight	=	1	5



Optimized BDT in central q^2

$n\_estimators = 200$
$max\_depth = 4$
$learning\_rate = 0.05$
$subsample = 0.5$
$min\_child\_weight = 5$





BDT for low q^2

- Now we want the same, but for the low q^2 bin
- But, before starting from scratch we check how the central q^2 BDT performs here

BDT for low q^2

- Now we want the same, but for the low q^2 bin
- But, before starting from scratch we check how the central q^2 BDT performs here

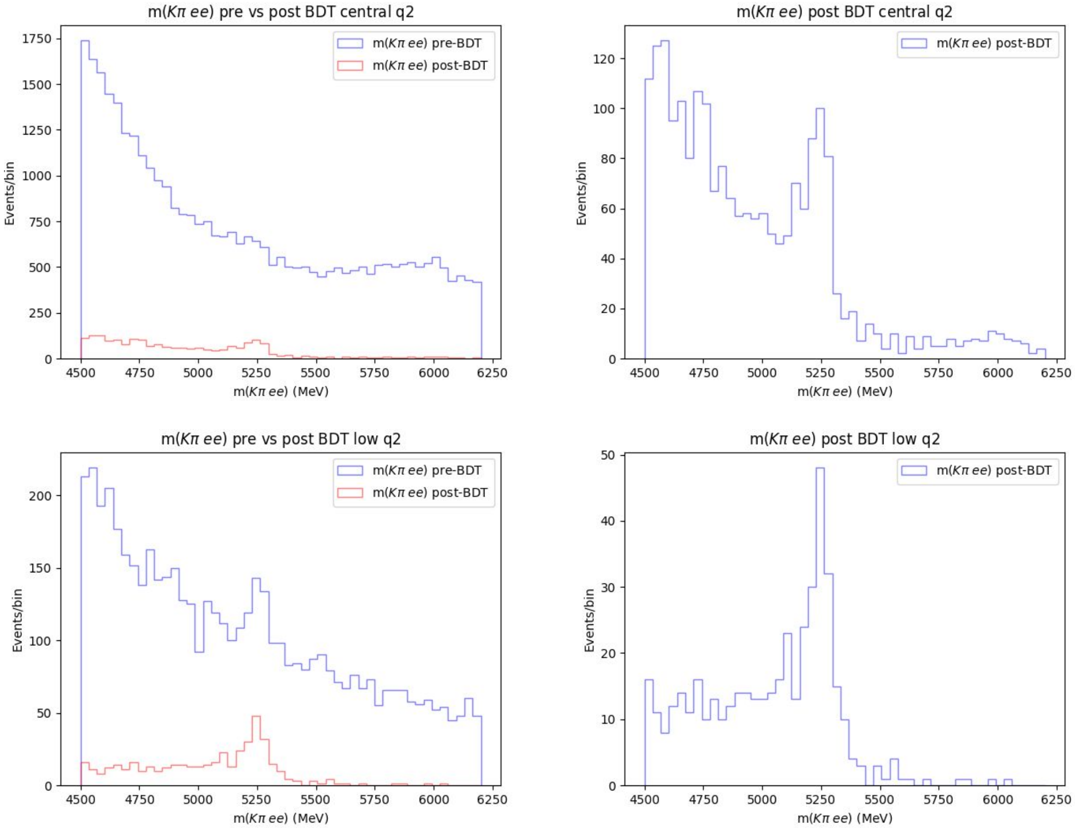
Retention	central q^2	low q^2
Signal	0.883±0.002	0.880±0.002
Combinatorial	0.0212±0.002	0.0212±0.002
Partially Reconstructed	0.702±0.012	0.712± 0.012

We can work with this!

B mass peak pre and post BDT

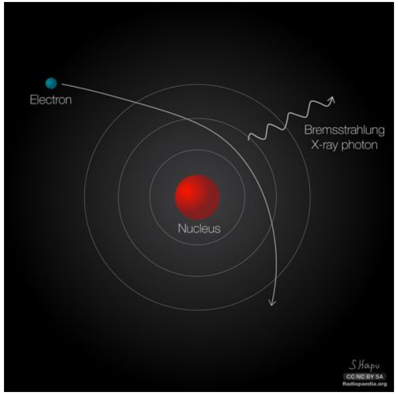
Central

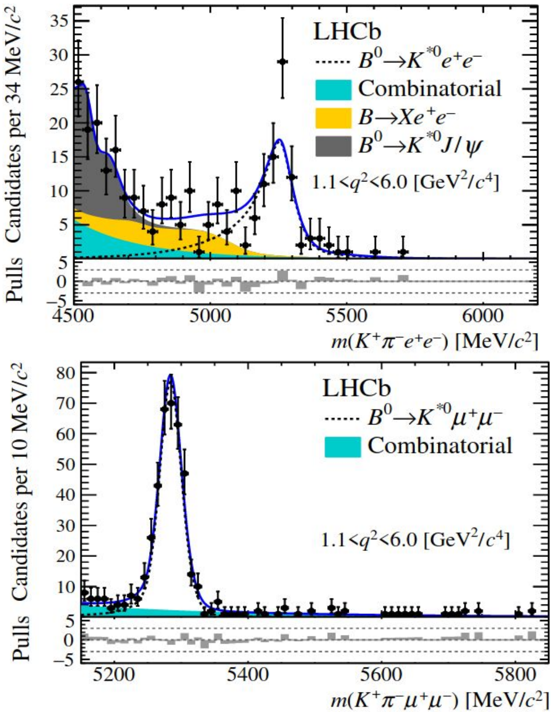
Low



Troubles with reconstruction of electrons

- Electron Bremsstrahlung 10^8 higher than muons!
- Worse mass resolution
- Worse reconstruction efficiency

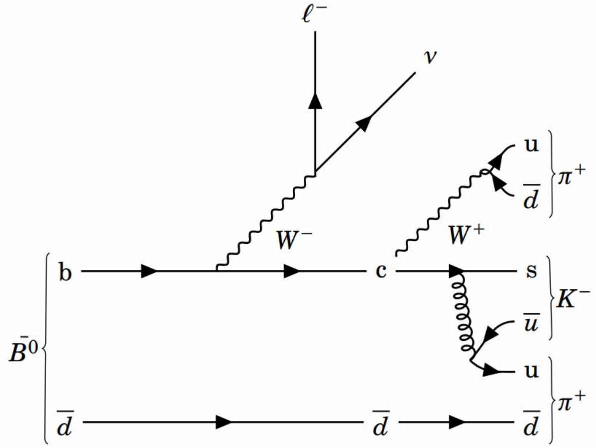


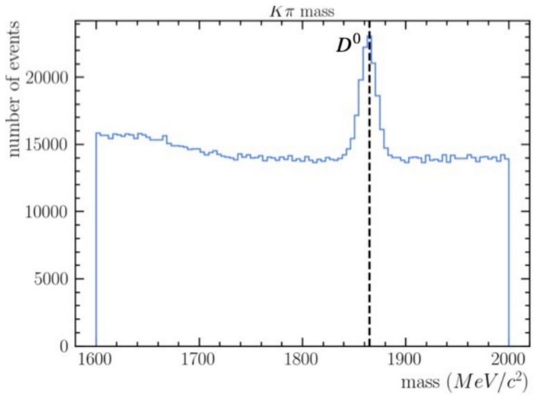


JHEP 08 (2017) 055

Semileptonic background studies

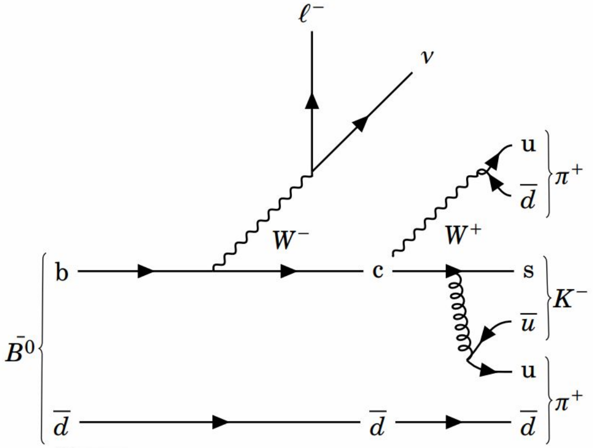
- Semileptonic decays is one of the most common b quark decays
- It is a large background in our data
- We used to have a cut in place to deal with this, however …

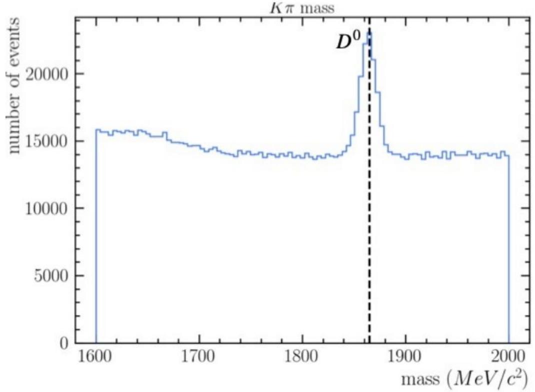




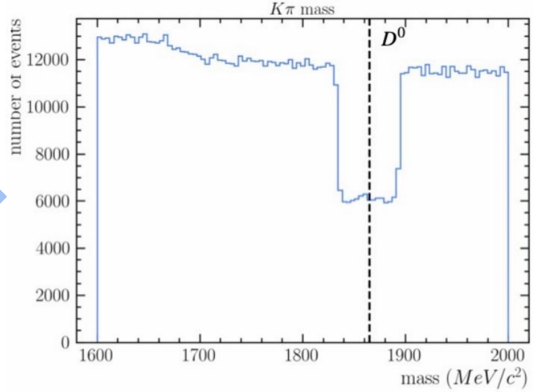
Semileptonic background studies

- Semileptonic decays is one of the most common *b* quark decays
- It is a large background in our data
- We used to have a cut in place to deal with this, however ...





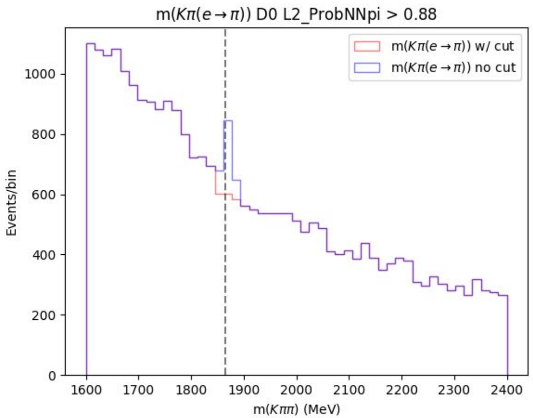
Old cut

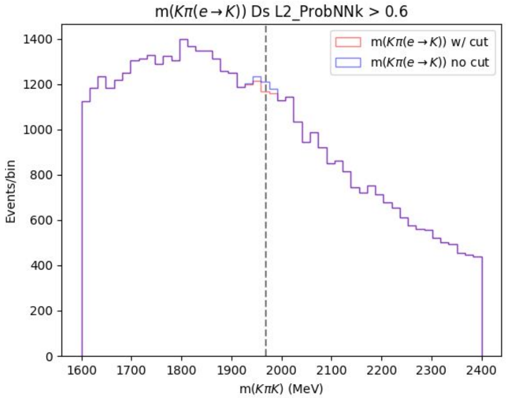


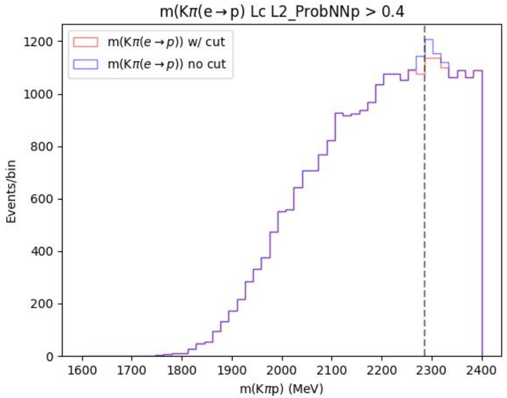
[Nina Oskam 2023]

New cut for (some) semileptonic decays in central q^2

- New cut using brem 0 events and probability of electron being …
  - Pion
  - Kaon
  - Proton

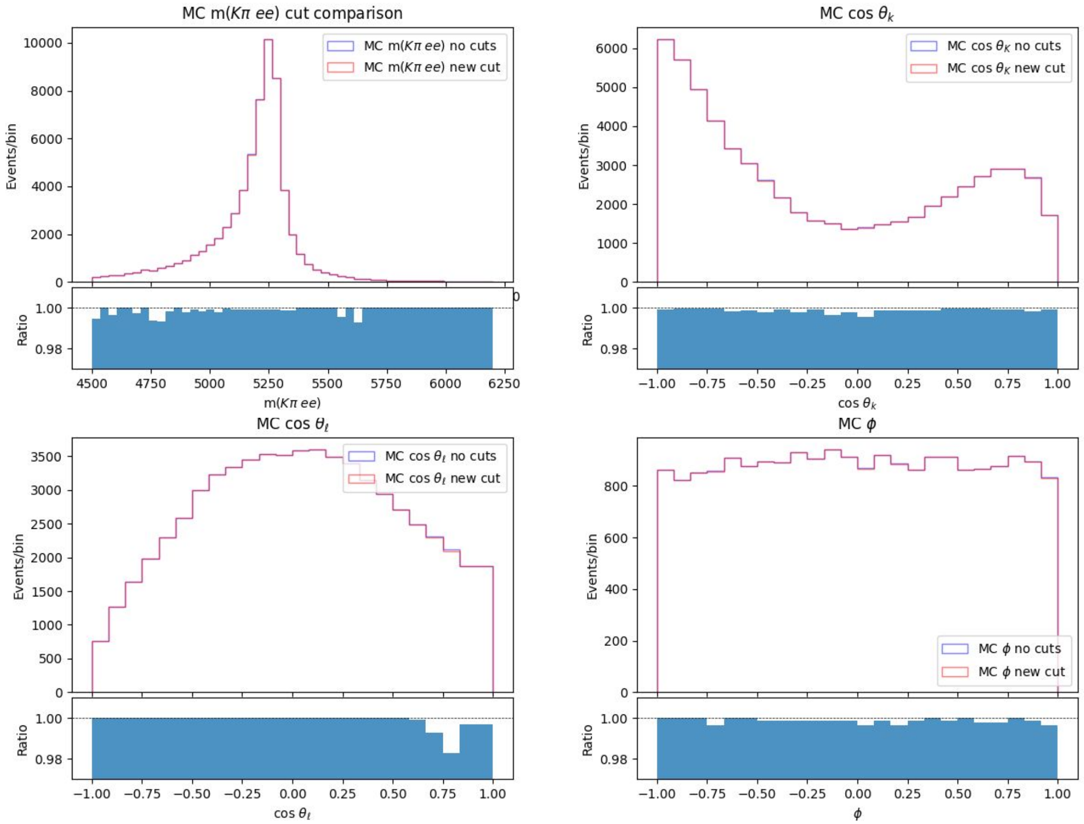






- More on this later…

Effects of new cut



Efficiency	MC
New cut	99.90%
Old cut	96.66%

Preliminary Mass Fit

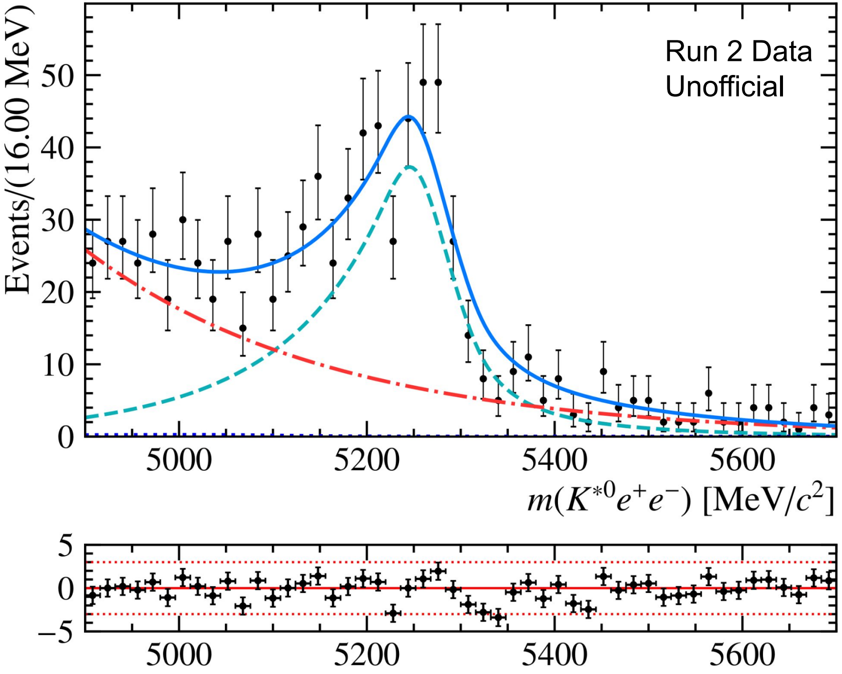
- Making use of Likelihood function
- Using Minuit to minimize this to find best pdf

$$\mathcal{L}(\vec{x}_i, \vec{\theta}) = \prod_i f(\vec{\theta}|\vec{x}_i)$$

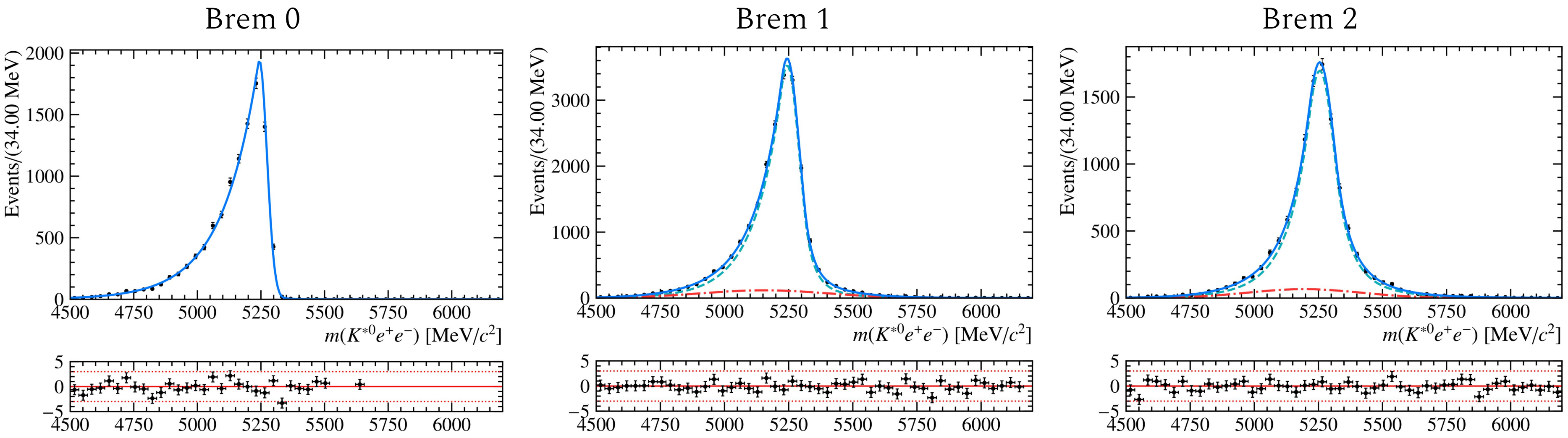
$$-\ln \mathcal{L}(\vec{x}_i, \vec{\theta}) = -\sum_i \ln f(\vec{\theta}|\vec{x}_i)$$

Preliminary Mass Fit

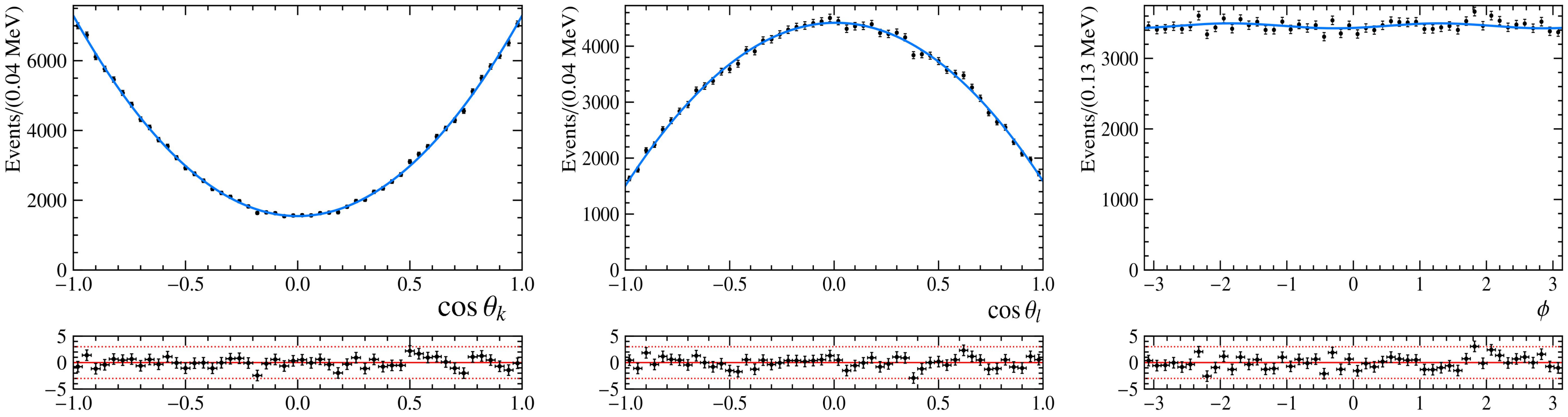
- Data signal
- Combinatorial
- Partially Reconstructed



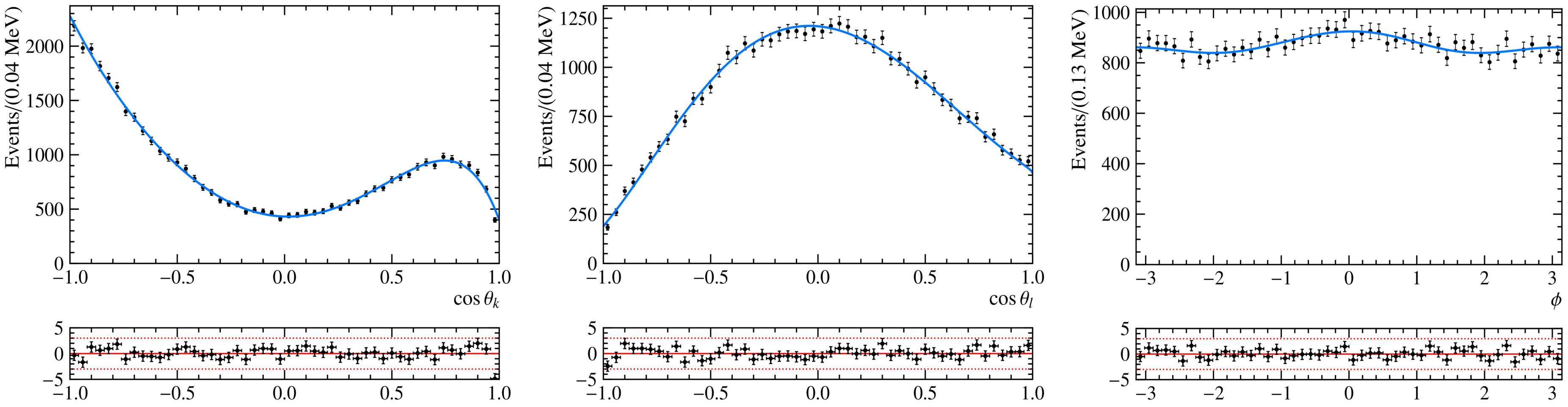
MC Signal Fits



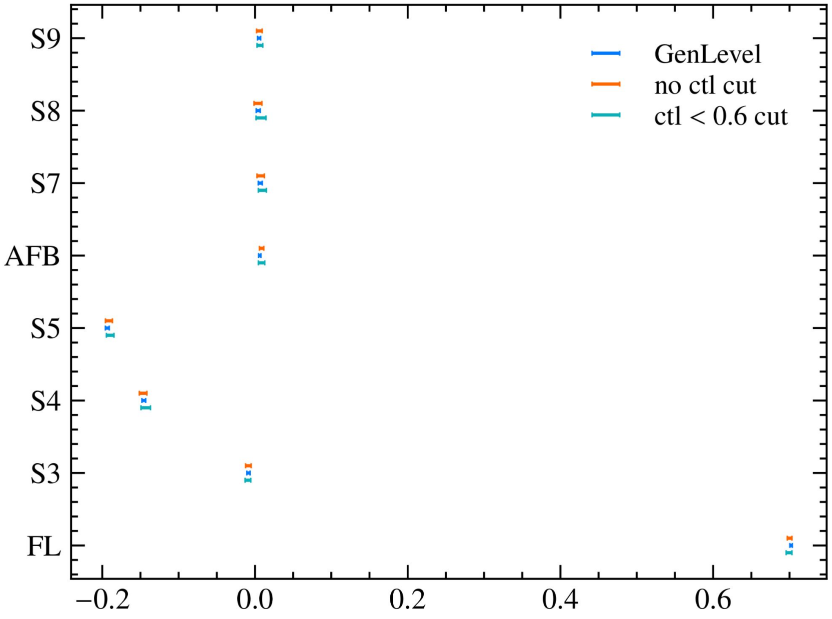
MC Generator Level Decay Angles

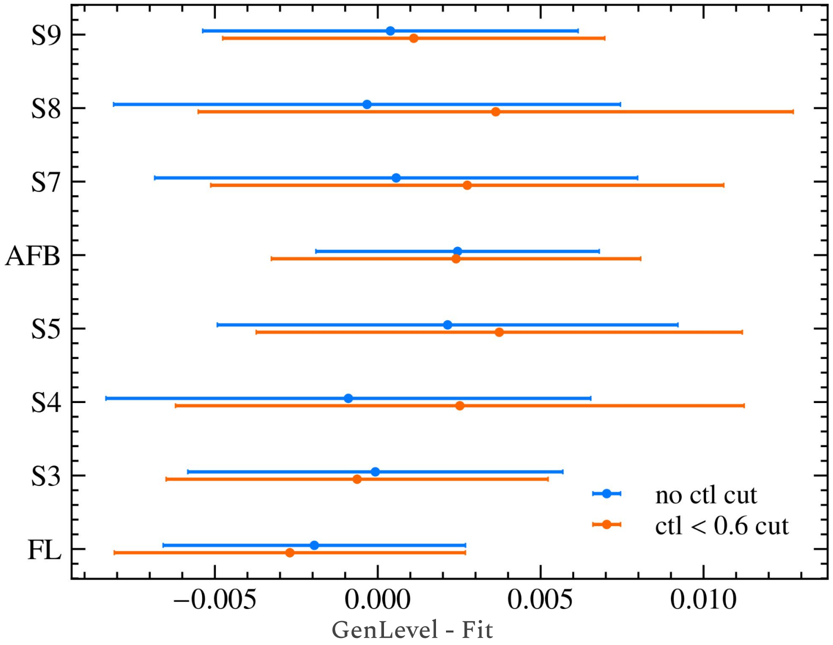


MC Angular Fit w/ Acceptance



Comparison effects of cut in ctl on angular coefficients

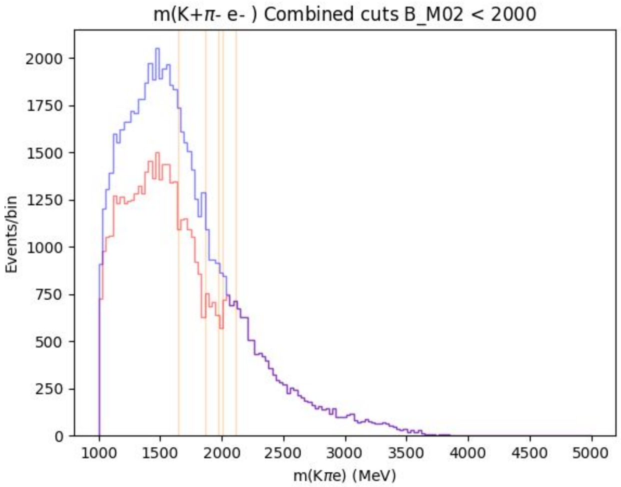


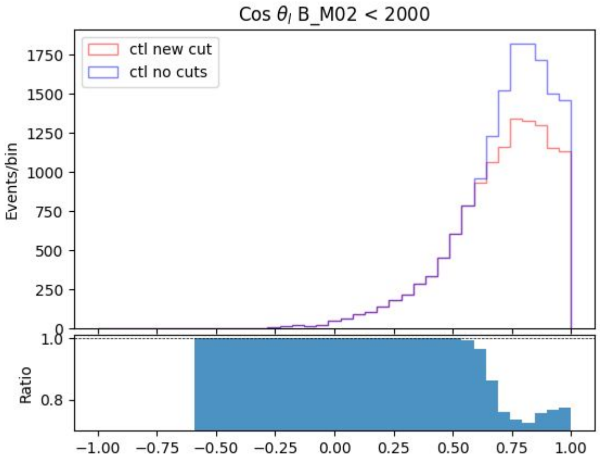


Outlook for semileptonic backgrounds

- Target SL decays below m(K^+e^-) < 2000 MeV^2/c^4
- Peak in high cos $\theta\ell$
- Use the χ^2 on the electron vertex

$$B^0 \to (D^- \to (K^* \to K^+\pi^-)\pi^-)e^+\nu_e$$
$$B^0 \to (D^* \to (\bar{D} \to K^+\pi^-)\pi^-)e^+\nu_e$$
$$B^0 \to (D^- \to (K^* \to K^+\pi^-)e^-\bar{\nu}_e)e^+\nu_e$$





Conclusion

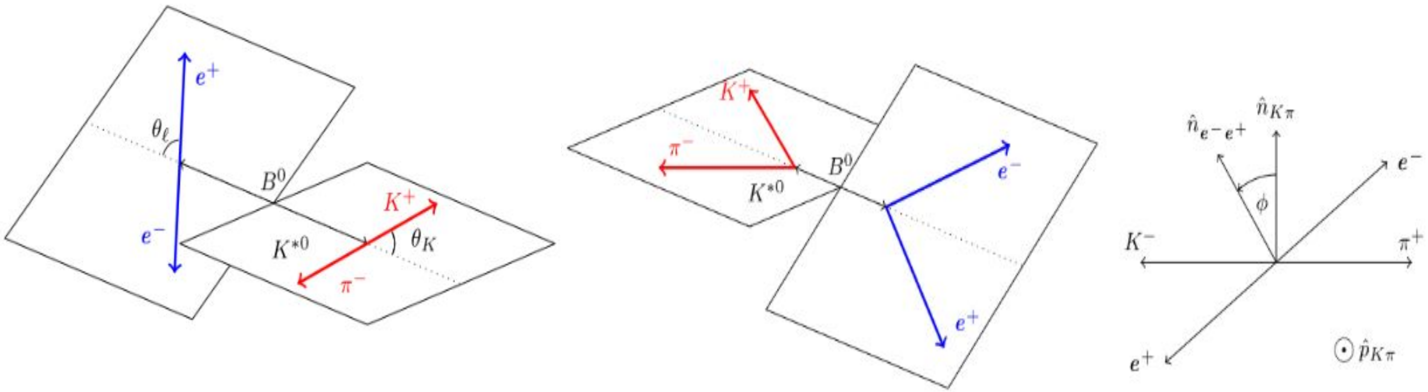
- Provided contributions to the collaboration
    - Optimizing steps in workflow
    - Creation of nTuples
- Improved the BDT by optimizing hyperparameters
    - Both central and low q2
- Replaced a cut on semileptonic decays
    - To be used as a starting point
- Provided a preliminary mass and angular fit
    - Mass fit on data
    - Angular fit on MC

Thank you for your attention!

Thank you for your attention!

BACK UP SLIDES

Decay angles definitions

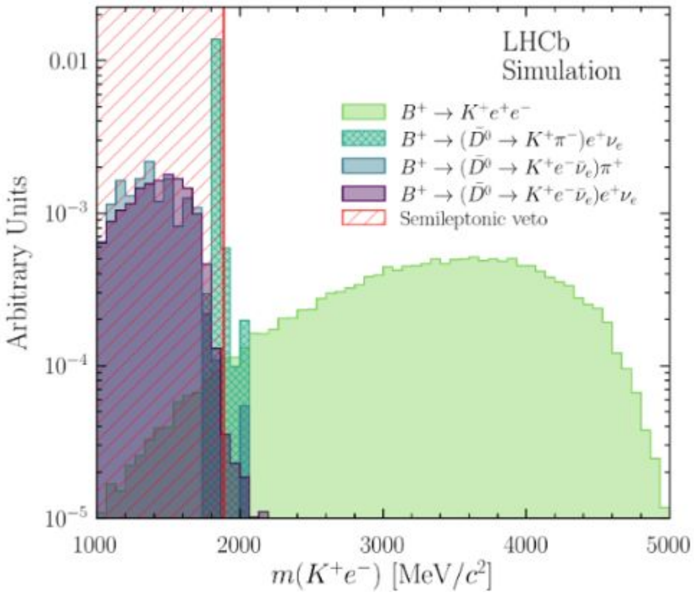


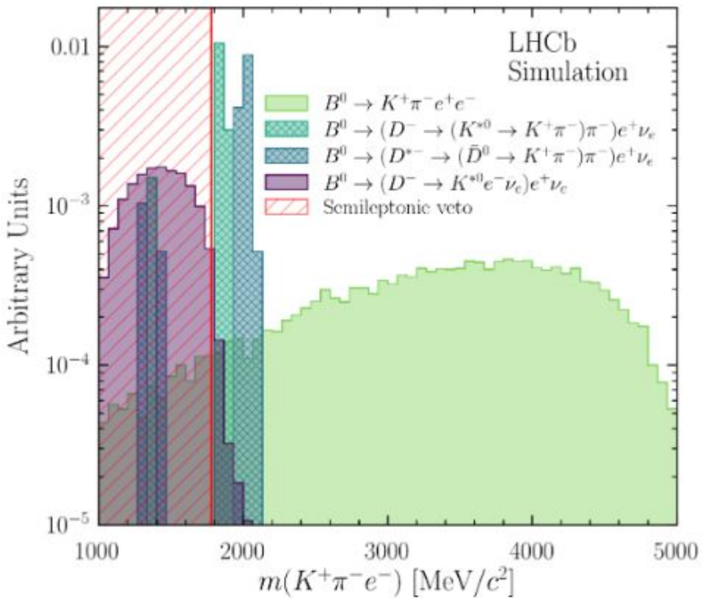
θ_l → between the direction of the e^+ and the direction opposite to that of the B^0 in the rest frame of the dimuon system
θ_K → between the direction of the K^+ and the direction of the B^0 in the rest frame of the K^{*0}
φ → between the plane defined by the electrons pair and the plane defined by the kaon and pion in the B^0 rest frame

Wilson coefficients

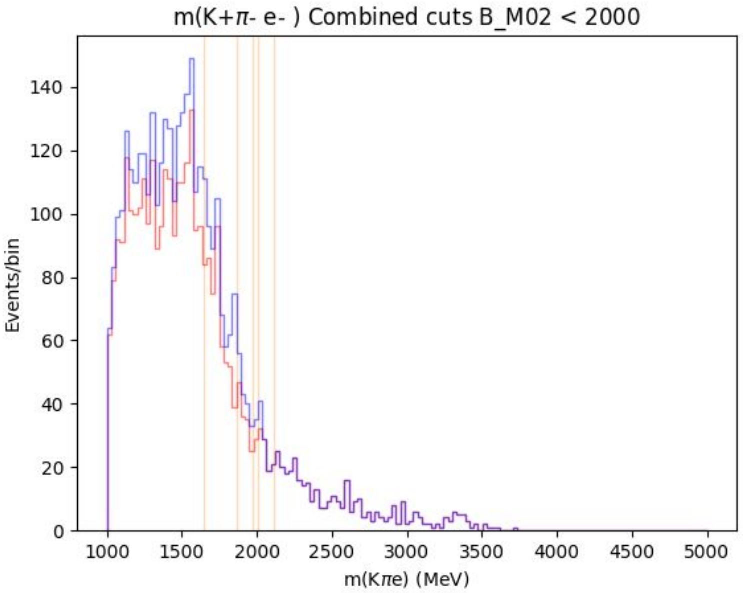
- The operator C7 describes the radiative decay $b \to s\gamma$
- The operators C9 and C10 both describe the $b \to q\ell\ell$ transition. C9 corresponds to the vector current, and C10 to the axial current
- Finally, C10 describes the $B \to \ell^+\ell^-$ decays (in the SM)

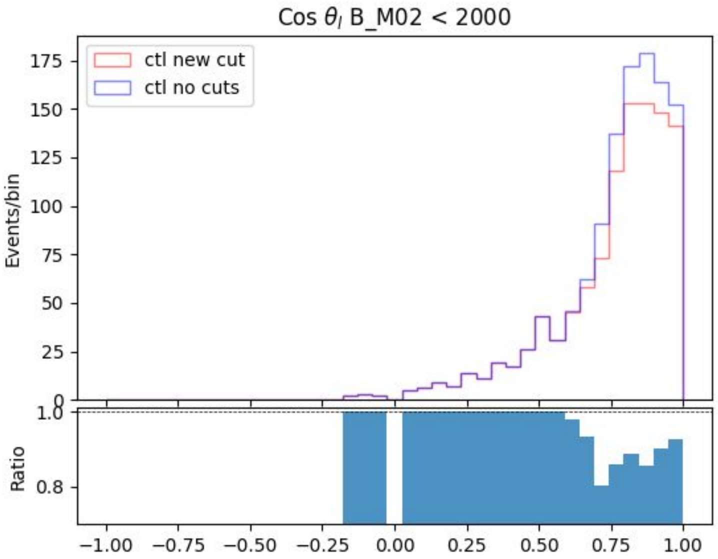
Semileptonic backgrounds





Outlook cuts after BDT



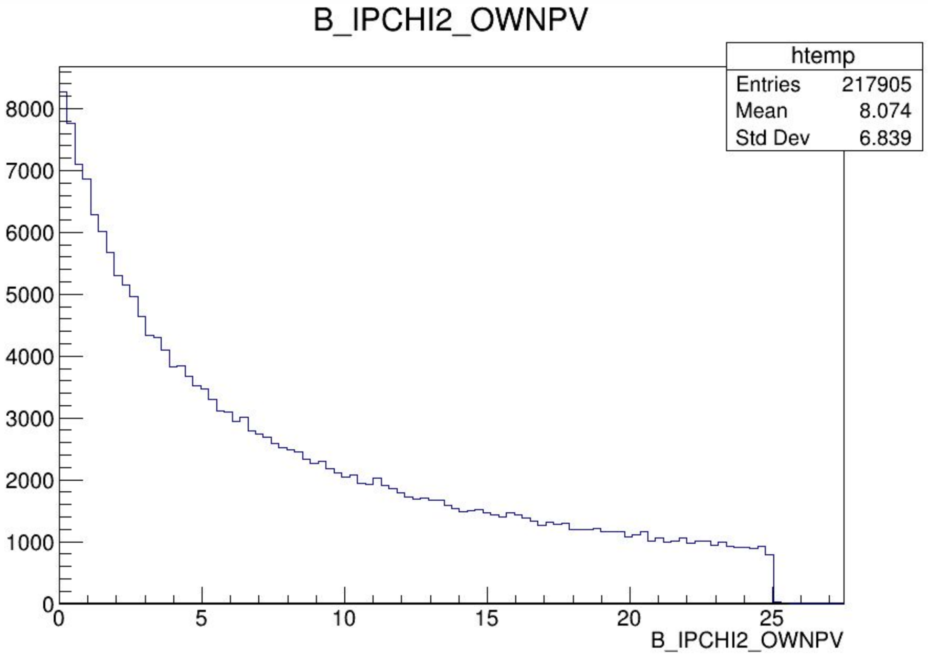


Hyperparameter definitions

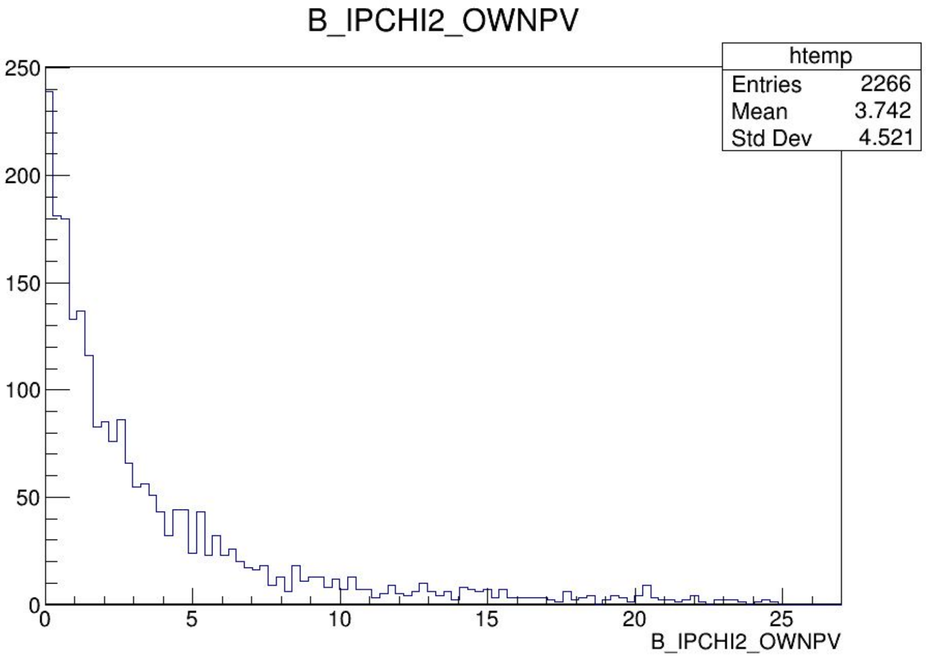
- n_estimator: amount of trees or rounds in the model
- max_depth: maximum depth of a tree
- learning_rate: step size shrinkage
- subsample: ratio of training instance, selection of the training data
- min_child_weight: minimum sum of instance weight needed in a child

Feature differences Data vs MC

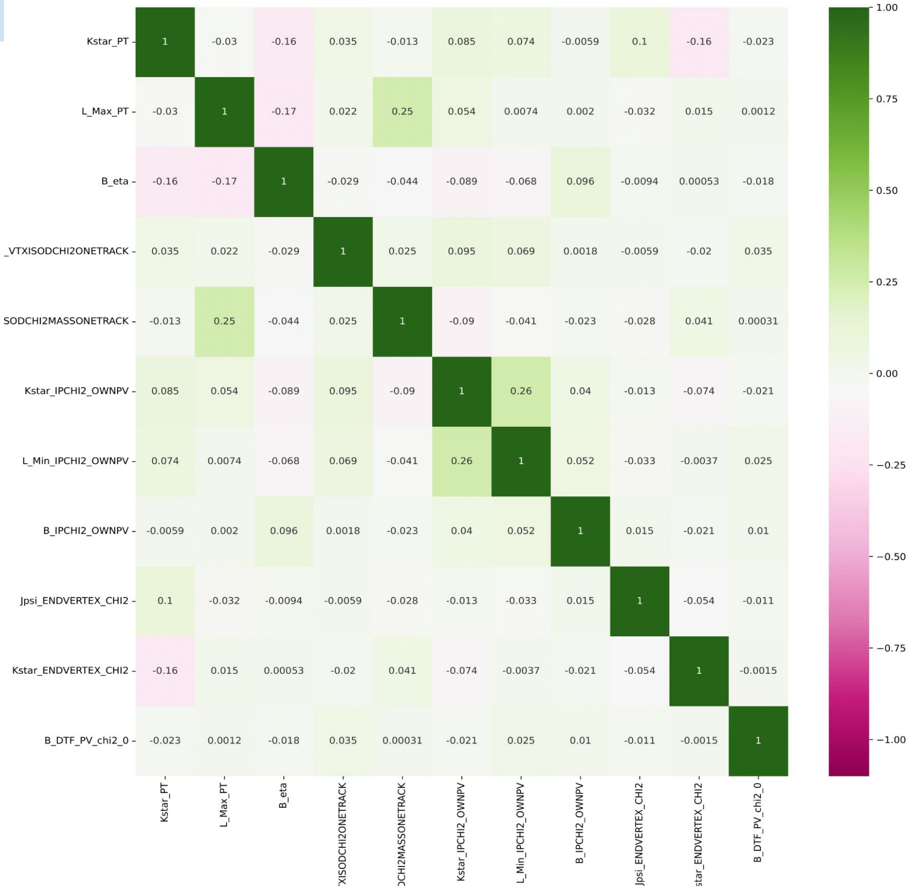
Data



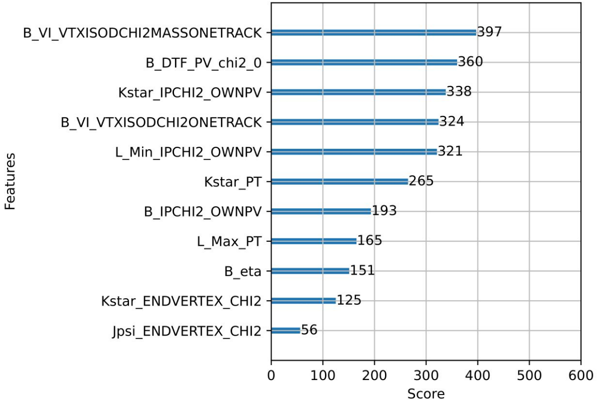
MC

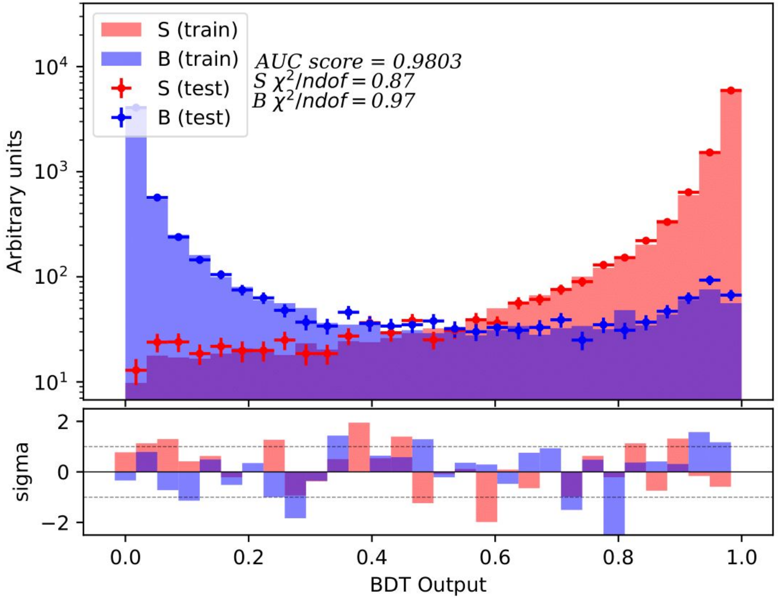


Correlation Matrix

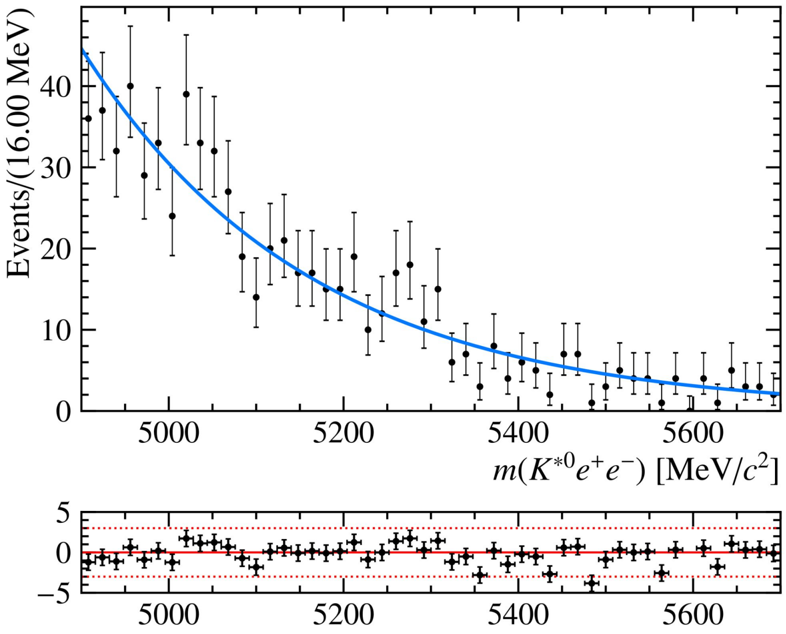


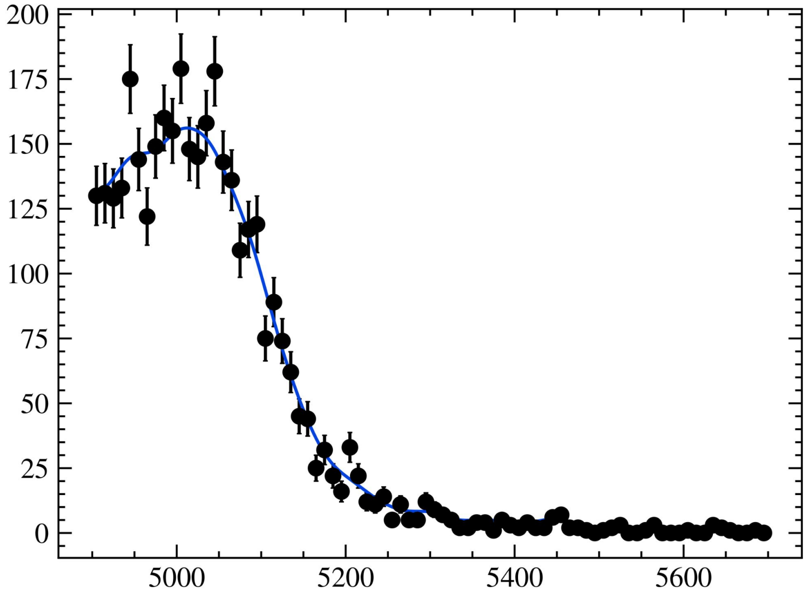
BDT train vs test low q2





Combinatorial and Partially Reconstructed Fits





BDT training samples

Cuts:

Signal: 4900 < B_M < 5600
q2 [1, 7] or [0.1, 7] for central or low

Backgr: 4600 < B_M < 4900 and B_M > 5600

Sample	Size [1, 7]	[0.1, 7]
Signal (MC)	56618	78171
Background (Data)	22722	25686

GenericPresel & GenericPresel_Additional & TighterKst0Presel & VetoesPresel & TriggerPresel & CloseVeto & MeerkatPresel_Tight & PIDPresel

B0_BKGCAT = 10, 50 or 60

BDT efficiency sample sizes

Sample	Size [1, 7]	Size [0.1, 1]
Signal (MC) total	19171	38742
Signal (MC) Brem 0	5357	10699
Signal (MC) Brem 1	9582	19480
Signal (MC) Brem 2	4232	85663
Combinatorial (Data)	5459	5178
Part Reco (MC) total	1505	1424
Part Reco (MC) Brem 0	314	275
Part Reco (MC) Brem 1	817	793
Part Reco (MC) Brem 2	374	356