School of Physics and Astronomy

Second Year Report

Conor Fitzpatrick
June 2010

Supervisor: Prof. F. Muheim
2nd. Supervisor: Prof. S. Playfer

1 Work To Date

1.1 Introduction

I am currently on Long Term Attachment at CERN, and have been since June of last year. This is an incredible time to be stationed at CERN and I have put my time here to good use by participating in the day-to-day running of the LHCb experiment in a number of roles, described later. I have also used my time at CERN to become active in early analysis work, including laying claim to measuring some of the first particles to be seen by the LHCb experiment. This Section of the report describes my work so far.

1.2 Service Work

Large collaborative experiments such as LHCb require that their member institutes undertake tasks that are not directly involved in physics analysis, but that contribute to the running and performance of the experiment as a whole. This Service Work usually involves taking shifts during the running of the detector and in the early life of the experiment some form of testing, calibration or hardware/software tasks. Due to Edinburgh's strong connection to the RICH detectors of the LHCb experiment, I have chosen to work on the time alignment of the RICHes in addition to taking a combination of RICH and LHCb-wide shifts.

1.2.1 Shift and Piquet work

I am trained and currently active in a number of roles concerning the running of the LHCb experiment. As Data Manager I assess, diagnose and act upon issues affecting the immediate quality of recorded LHC collisions as they leave the detector, working alongside the Shift Leader. As Data Quality Piquet I have been responsible for the production and reconstruction of data prior to analysis, and as RICH Piquet I have been involved in the alignment and stability of the LHCb RICH subdetector and as first point of contact should problems arise. I am the first PhD student to be trained and deployed as Shift Leader, responsible for the daily running of the LHCb experiment as a whole. These roles have given me an insight into the operation of not only our detector, but the LHC machine itself. As a result I have already made significant progress on the detector chapter of my thesis.

1.2.2 RICH Time Alignment

The Ring Imaging CHerenkov (RICH) detectors of the LHCb provide particle identification for the LHCb experiment, offering the $\pi - K$ separation needed to identify the final states of many B and D meson decays[1]. RICH1 covers the momentum range $1 \rightarrow 60$GeV/c and is situated between the Vertex Locator and the first tracking station. RICH2 is downstream of the magnet and covers the range $15 \rightarrow 100$GeV/c upwards. The RICHes consist of gas enclosures containing aerogel, C_4F_{10} (RICH1) and CF_4 radiators. Cherenkov photons from these media are reflected by way of spherical and flat mirrors out of the LHCb acceptance and into rings focussed on enclosures containing Hybrid Photon Detectors (HPDs). The HPDs of the LHCb RICHes are a unique design, tested and calibrated at Edinburgh and Glasgow prior to installation in the experiment. The HPDs are controlled and read-out by on-detector electronics referred to as L0 boards, 2 HPDs per-board. There are 196 HPDs in RICH1 and 288 in RICH2, requiring 242 L0 boards in total. The 40MHz LHCb clock, trigger and calibration signals required to read-out the HPDs are generated once per L0-board, so that both HPDs are read-out synchronously. The L0 boards of the RICH must be timed-in to the global LHC clock in such a way that the HPDs are read-out at their optimal efficiency in time, which is dependent upon the specific characteristics of each HPD, of which the leakage current is the strongest contributor. As such, HPDs are paired according to these characteristics as determined by the automated test procedures devised and implemented at the Photon Detector Test Facility (PDTF) in Edinburgh. This ensures that in general, the smallest unit that must be time-aligned in the LHCb RICH detectors are the L0 boards. My role has been to globally time-align the L0 electronics of the RICH detectors in order to read out data synchronously with the other subdetectors of the LHCb experiment such that the maximum timing efficiency per L0 board is achieved. The L0 boards may be shifted by timing steps of as low as 100ps based on the design of the timing circuitry, but the time profile for HPD readout is approximately a plateau of width O(10)ns. As such we align to 1ns resolution on the center of this plateau in order to allow for drift. I have been responsible for the development of alignment code that reads out the number of accumulated pixel hits per HPD in a range of time steps of varying resolution and calculates the optimal relative delay required per L0-board. During normal functioning of the LHCb RICH detectors these delays are stored and applied to the L0 readout electronics. Prior to colliding protons at the LHC, this alignment procedure was attempted using a pulsed laser, triggered by the LHCb internal clock. The diffuse laser illumination is delivered to the HPD enclosures of RICH1 and RICH2 by way of fibers through a splitter. Due to a non-negligible path length difference between photons reaching the extrema and center of each enclosure the maximum achievable alignment resolution was $\sim$ 5ns. After the LHC start-up this procedure used Cherenkov photons from the products of actual proton-proton collisions, at which point we were able to successfully time-align the L0 electronics to 1ns resolution.

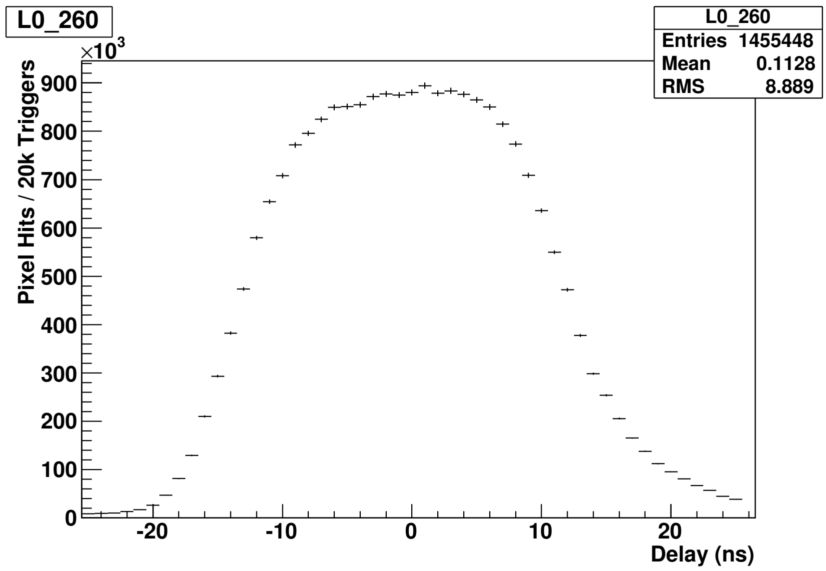

Figure 1: Time alignment profile of a typical L0 board. The alignment code applies an initial delay of -25ns, integrates the pixel hits in both HPDs on the board over the course of 20,000 triggers and then steps the delay by 1ns, repeating this process over a 51ns range, centered on the currently stored delay value. The midpoint to which the board should then be delayed is defined as half the width of the plateau, whose turn-on and turn-off points are at 90% of the background-subtracted maximum.

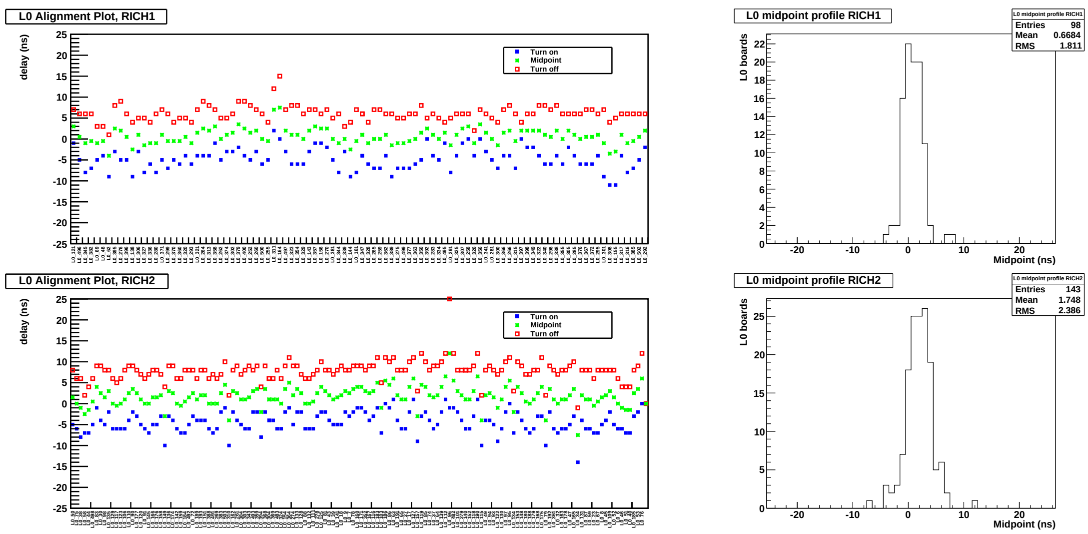

Figure 2: Time alignment plots (left) and midpoint profiles (right) for (top) RICH1 and (bottom) RICH2 after initial alignment using limited statistics with proton-proton collisions.

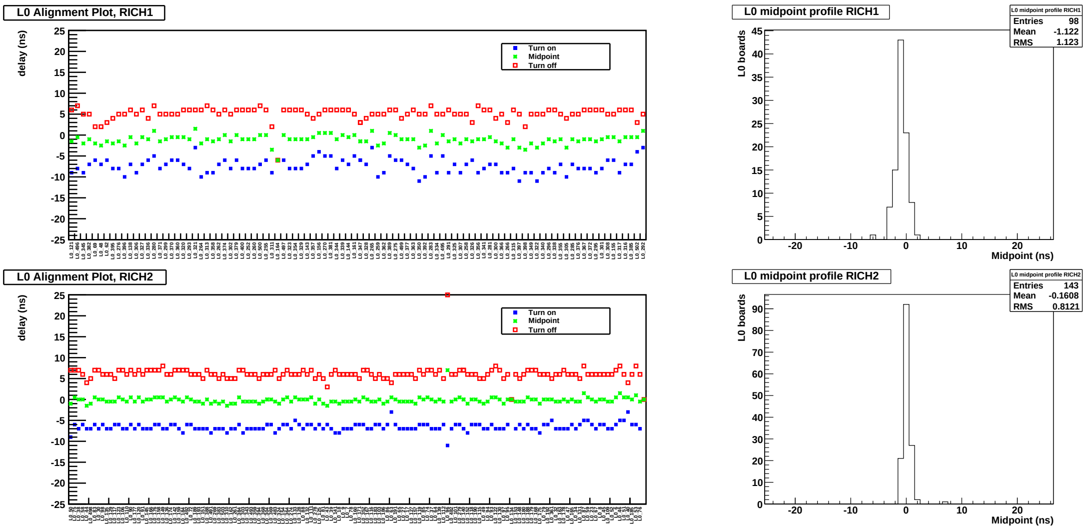

Figure 3: Time alignment plots (left) and midpoint profiles (right) for (top) RICH1 and (bottom) RICH2 after final alignment using high statistics with proton-proton collisions. The final resolution is O(1ns) for both RICHes

1.3 (Re)Discovery of charmed strange mesons at the LHCb experiment

With the first data from the LHCb I have been the first within the collaboration to present a preliminary measurement of $D_s^{\pm}$ mesons in combination with D^+ mesons decaying to $\phi\{K^+, K^-\}\pi^{\pm}$. This was shared with the collaboration less than 48 hours after the first physics datasets were processed, for which I was also responsible as Data Quality Piquet. The rediscovery of previously measured particles is of great importance for any new experiment as it confirms calibration and alignment in addition to tuning the analysis framework. The LHCb experiment can probe pseudorapidity and momentum ranges inaccessible to the other experiments at the LHC (From 0GeV/c upwards in P_t and $2.0 < y < 4.5$), which allows us to measure the production cross-section of particles decaying within the LHCb acceptance at the energy frontier.

1.3.1 A Novel approach to selection optimisation

In order to obtain a clean sample of particles with which to measure the production cross section, I have implemented a new method of selection optimisation that is purely data-driven. In a selection, cuts are applied which discriminate signal from background. Determining which variables offer discrimination power and which values of cut to use traditionally relies upon monte-carlo (MC) simulation: Input to any cut optimisation tool requires some form of flag that denotes samples as signal and background so that the optimiser can tell how efficient and pure the selection is at each working point. It is a common optimisation strategy to maximise the signal statistical significance, defined as $S/\sqrt{S+B}$ where S denotes signal candidates and B background. Previously I have devised and written a C++ binary application called CROP, the Cut Recursive OPtimiser[2]. This determines the optimal signal significance achievable by an ensemble of discriminating variables when implemented as rectangular cuts in a hypervolume using Monte-Carlo. While this technique is not new, CROP is optimised for speed and simplicity, directly interfacing the data storage format and powerful processing techniques

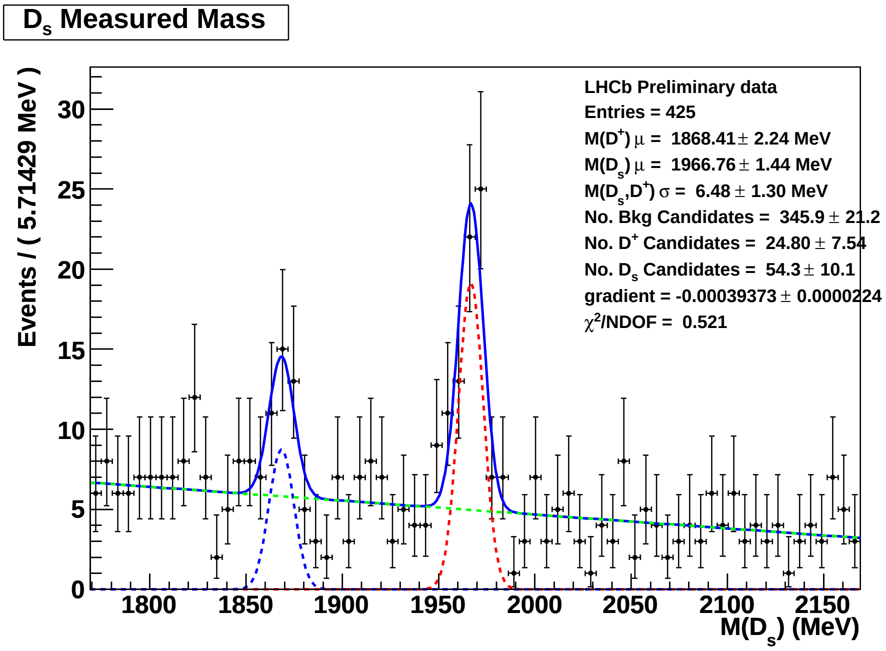

Figure 4: A fit to the first $D_s^{\pm}$ and $D^{\pm}$ candidates measured by the LHCb experiment. The leftmost Gaussian (Blue, dashed) is a fit to the $D^{\pm}$ mass (PDG value 1869.62±0.20MeV). The rightmost Gaussian (Red, dashed) is a fit to the $D_s^{\pm}$ (PDG value 1968.49±0.34MeV). The widths of both Gaussians are a single parameter in the fit. The background (Green, dashed) is modelled as a simple first-order polynomial with the specified gradient.

offered by the ROOT HEP framework. CROP along with a number of additional tools formed the basis of my first LHCb collaboration internal publication, and is in use by a number of institutes in LHCb.

At the early stages of data taking, Monte Carlo is not fully mature and so the data can differ significantly from MC. Cuts optimised on MC can therefore be less efficient or pure when applied to data than expected. In addition Monte Carlo is resource-intensive, requiring significant quantities of time and storage space to generate. As such, limited MC samples are available to train optimisation strategies on. The largest source of minimum bias data we have generated is actual collision data- the available MC is several orders of magnitude smaller. It is therefore interesting to investigate an approach to selection optimisation that does not rely on Monte Carlo. In order to do this, we must be able to obtain or somehow unfold the distributions of signal and background in a data sample; something that in MC is straightforward by looking at the "truth". In data no such truth flag is available, much to the lament of many PhD students! A powerful technique to unfold the distribution of multiple species (eg: signal, background) is that of the sPlot[3]. An sPlot for a given variable is a binned histogram derived from the sum of per-event sWeights. sWeights are obtained from a discriminating variable whose distributions in each species are well enough defined to be fit to by an unbinned log-likelihood fit. Each candidate is assigned an sWeight per species such that the sum of the sWeights of a given species in bins of any other variable reflects the yield of the given species in these bins.

While the sPlot technique is not new and nor is rectangular cut optimisation, combining the two to form an entirely MC-free optimisation technique is certainly a new development within the LHCb collaboration. Rather than maximising the signal statistical significance $S/\sqrt{S+B}$, CROP is now modified to optimise $\sum_{i=1}^{N} sw_i/\sqrt{sw_i + bw_i}$ Where N is the total number of candidates passing the ensemble of cuts, sw_i is the signal sWeight assigned to candidate i and bw_i is the background sWeight assigned to the same candidate. The only requirement is that some form of preselection is required in order to achieve discrimination in one variable.

This technique has proven to be powerful not just in the case of two species, but in several. Independent fits to multiple signal species result in sWeights that can be summed such that $sw_i = sw_i^1 + sw_i^2$ permitting simultaneous optimisation for control channels, etc. sWeights are also multiplicative, so it is possible to discriminate several species of a single signal from background and their sister species using a second discriminating variable to further split the signal sample such that $sw_i^{ab} = sw_i^a \times sw_i^b$. I addition to this, parameter extraction can be performed from an sWeighted dataset using a weighted likelihood fit[4].

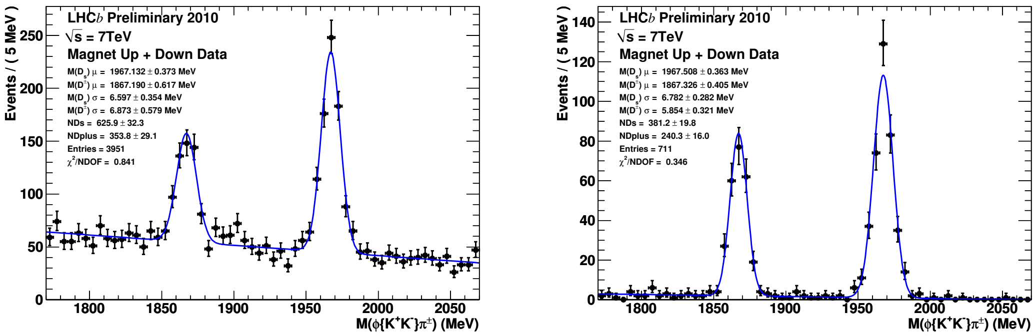

Figure 5: $D_s^{\pm}$ and $D^{\pm}$ candidates before (left) and after (right) selection optimisation using sWeights on data. The initial B/S in this mass window is 3.0. After optimisation this drops to 0.14.

Using sWeights in CROP has allowed me to obtain a very pure and efficient sample of D_s, D^+ mesons without using MC, and to separate them into prompt and secondary candidates. Other analyses are also using this tool for their optimisations with promising results being shown for $B_s \to J/\psi\phi$.

1.3.2 Towards a production Cross-section Measurement of D_s, D^+ Mesons

Having selected an optimal sample of D_s and D^+ mesons, my priorities have turned to making a production cross section measurement. This is progressing at a somewhat hurried pace in preparation for the summer conferences. Presently I have made a preliminary differential cross-section measurement in bins of y, Pt independently due to limited statistics, and without analysis of systematic errors. Work is ongoing to develop the result into one which will be suitable for external publication.

1.4 Thesis Layout

I have already devoted some spare time to creating a thesis template, and made progress in writing a portion of the chapters that do not rely on analysis results. The thesis outline so far consists of the following chapters and sections:

- Chapter 1: CP violation in the standard model and beyond

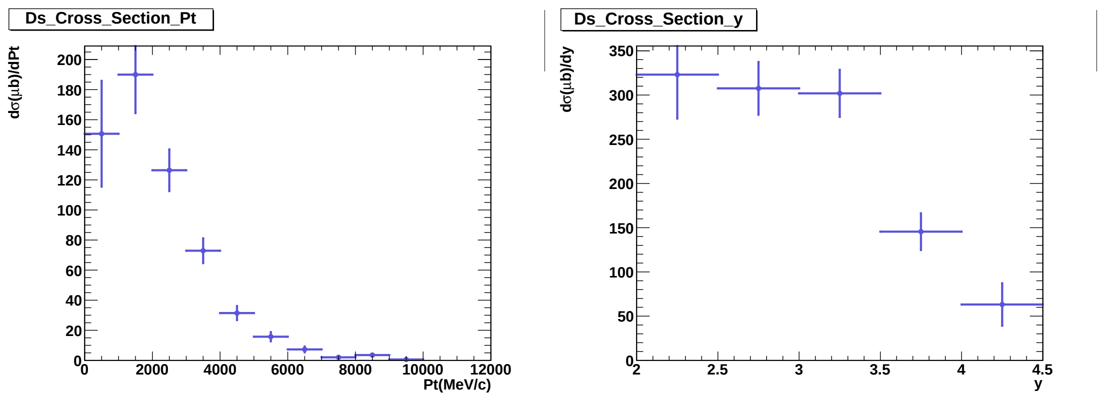

Figure 6: D_s Production cross sections as measured in the mode $D_s \to \phi\{K^+K^-\}\pi^+$ in (left) Pt on the range $0 \to 12\mathrm{GeV/c}$ and (right) in y on the range $2.0 \to 4.5$. Due to limited statistics it is only possible at present to show each variable integrated over the range of the other. These results are preliminary and do not include systematic uncertainties.

  - 1.1: Introduction to the Standard Model
  - 1.2: Manifestation of CP in the Standard Model
  - 1.3: $-2\beta_s$ in $B_s \to J/\psi\phi$

- Chapter 2: The LHCb Experiment
  - 2.1: b-physics at colliders
  - 2.2: The LHC
  - 2.3: b-production at the LHC
  - 2.4: The LHCb Detector

Additional chapters shall consist of:

- Chapter 3: A measurement of the production cross-section of charmed, strange mesons at LHCb
- Chapter 4: Time Alignment of the LHCb RICH detectors
- Chapter 5: Preliminary measurements of ϕ_s in $B_s \to J/\psi\phi$

While subject to change in terms of layout, these chapter headings serve as placeholders for work that I have either already completed, or that I will complete in the course of my final year.

1.5 Conclusions

This year has been a busy one for both myself and the LHCb experiment as a whole. I have capitalised on my proximity to the experiment at CERN to apply myself to work that will form a substantial component of my thesis. I have thoroughly enjoyed being immersed in the productive atmosphere and environment of CERN, and look forward to contributing further to the body of work produced by the LHCb.

2 Future Plans and Final Year

2.1 Introduction

This section will discuss preparation of my thesis, planning and topics of analysis for my final year. The LHC and the LHCb have both had a remarkable year with the first recorded collisions being towards the end of 2009 and a new energy record set in early 2010. The LHCb expects to have recorded $\sim 100\text{pb}^{-1}$ of collisions by the end to 2010, and the LHC plans to deliver another 1fb^{-1} throughout 2011. With the current data on-disk there are many opportunities for research at the energy frontier, but with the 2011 data the LHCb will be able to make some of its flagship measurements.

2.2 ϕ_s in $B_s \to J/\psi\phi$

The main body of my analysis topic is an extraction of $\phi_s = -2\beta_s$[5], where ϕ_s is the measurable phase in the decay of B_s mesons to $J/\psi\phi$. In the Standard Model CP violation in the interference between mixing and decay of B_s mesons is highly suppressed. This is quantified by the CP violating weak phase $-2\beta_s = (3.68 \pm 0.17) \times 10^2$. Should physics processes exist that are not a part of the Standard Model, the measurable phase ϕ_s may deviate from the SM value of $-2\beta_s$. Current measurements of ϕ_s from the Tevatron are compatible with $-2\beta_s$ but have a large uncertainty, which leaves room for the discovery of new physics.

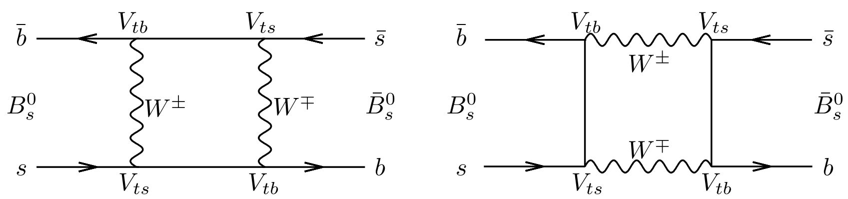

Figure 7: B_s mixing s and t channel diagrams.

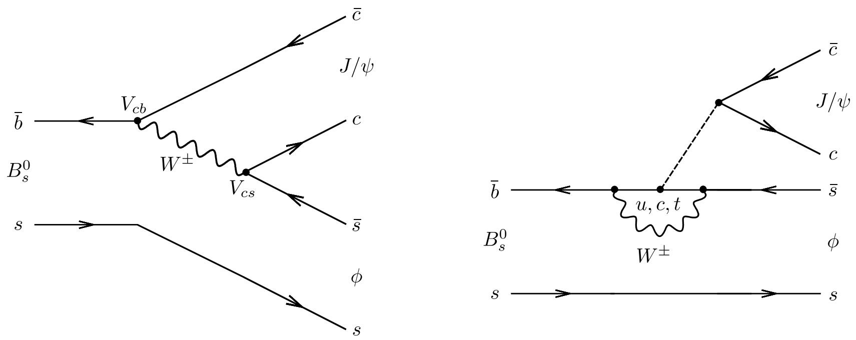

Figure 8: $B_s \to J/\psi\phi$ feynman diagrams at (left) tree and (right) penguin level. Tree level is CKM favoured and dominates by a factor of λ^2 with respect to the penguin.

With the data available at the end of 2010 it should be possible to measure ϕ_s to a precision close to that of the Tevatron experiments at the LHCb. My work in this area will start with selection optimisation as soon as sufficient quantities of data become available, again using CROP. We plan to modify CROP in such a way as to minimise the error on ϕ_s, rather than maximising $S/\sqrt{S+B}$ as is currently the case. Thereafter work on a measurement of ϕ_s will begin in earnest, with the majority of my work being in fitting to extract the weak phase. It is expected that this can be done in such a way as to be prepared for the data taken in 2011 so that my analysis chapter will contain the most precise measurement of ϕ_s to date.

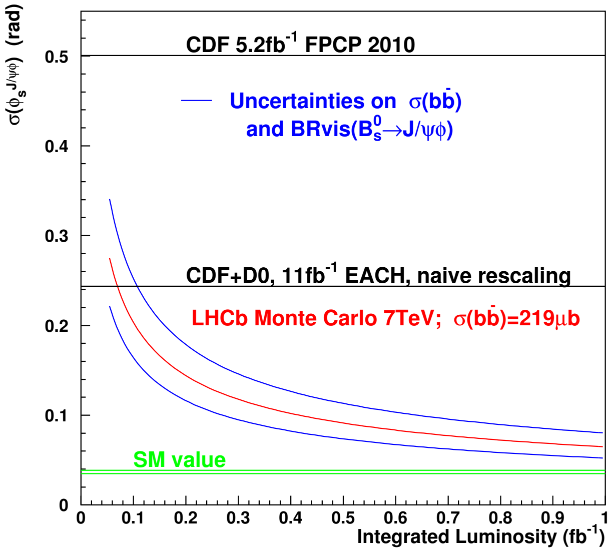

Figure 9: Estimated sensitivity to ϕ_s as a function of integrated luminosity. The red line denotes the nominal sensitivity to $2\beta_s$ for the LHCb experiment. In blue, the error on this estimate due to uncertainties in the $b\bar{b}$ cross-section and branching ratio $B_s \to J/\psi\{\mu^+\mu^-\}\phi\{K^+K^-\}$. The upper black line denotes the sensitivity achieved by CDF as made public at FPCP2010, and the lower denotes the sensitivity achievable by the CDF and D0 experiments at the Tevatron with 11fb^{-1} each assuming no further improvements in analysis save for increased dataset size. The green bars denote the uncertainty on the Standard Model value of $\phi_s = -2\beta_s$.

2.3 Thesis writeup

My theory and detector chapters are already partially written and shall be ready for a first proof by around Christmas. The analysis and results of the time alignment chapter are finished, and will shortly be written into my thesis in parallel with an internal LHCb note. The production cross-section measurement is ongoing and will be written up after it is finalised. Currently an LHCb note is being written on this in preparation for the summer conferences, and a significant portion of this note will serve as input to the thesis. Ideally, for my 3rd year report I would like to have several mature chapters written and submitted for proofreading.

References

[1] The LHCb Collaboration. The lhcb detector at the lhc. *Journal of Instrumentation*, 3(08):S08005, 2008.

[2] C Fitzpatrick. simpletools: Handy command line tools for ntuple manipulation and

analysis. Technical Report LHCb-INT-2009-029. CERN-LHCb-INT-2009-029, CERN, Geneva, Nov 2009.

[3] M. Pivk and F. R. Le Diberder. sPlots: A statistical tool to unfold data distributions. *Nuclear Instruments and Methods in Physics Research A*, 555:356–369, December 2005.

[4] Y. Xie. sFit: a method for background subtraction in maximum likelihood fit. *ArXiv e-prints*, May 2009.

[5] LHCb β_s and mixing workgroup. Road map for the measurement of mixing induced CP violation in $B_s^0 \to J/\psi\phi$ at LHCb. Feb 2009. LHCb-ROADMAP3-001.