Searching for beyond the Standard Model physics using direct and indirect methods at LHCb

Samuel C. P. Hall

Imperial College London
Department of High Energy Physics

A thesis to be submitted to Imperial College London
for the degree of Doctor of Philosophy

Abstract

It is known that the Standard Model of particle physics is incomplete in its description of nature at a fundamental level. For example, the Standard Model can neither incorporate dark matter nor explain the matter-dominated nature of the Universe. This thesis presents three analyses undertaken using data collected by the LHCb detector. Each analysis searches for indications of physics beyond the Standard Model in different decays of B mesons, using different techniques. Notably, two analyses look for indications of new physics using indirect methods, and one uses a direct approach.

The first analysis shows evidence for the rare decay $B^+ \to D_s^+ \phi$ with greater than $3\,\sigma$ significance; this also constitutes the first evidence for a fully-hadronic annihilation-type decay of a B^+ meson. A measurement of the branching fraction of the decay $B^+ \to D_s^+ \phi$ is seen to be higher than, but still compatible with, Standard Model predictions. The CP-asymmetry of the decay is also measured, and its value is precisely in line with the Standard Model expectations.

The second analysis claims the first observations of the decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$ which are both flavour changing neutral currents, forbidden at leading order in the Standard Model. Branching fractions of both these decays are measured, and for the high statistics channel $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ the differential branching fraction, as a function of the invariant mass squared of the dimuon system, is also presented.

These first two analyses both constitute indirect searches for physics beyond the scope of the Standard Model, where the observables are sensitive to contributions from new physics entering at loop-level. In contrast, the third analysis presented in this thesis involves the direct search for a new dark boson, χ, which is a messenger particle between a dark sector and the Standard Model particles. Using a frequentist technique, the dimuon component of candidates of the decay $B^0 \to K^*(892)^0\mu^+\mu^-$ for an excess consistent with $\chi \to \mu^+\mu^-$.

Chapter 2

The Standard Model and beyond

"Before beginning a Hunt, it is wise to ask someone what you are looking for before you begin looking for it."

Winnie the Pooh, A.A. Milne

This thesis contains the work undertaken in three analyses; each of which concerns a different area of interest in high energy physics. The following chapter aims to motivate each analysis in turn after introducing the Standard Model of particle physics.

Firstly, the formulation of the SM will be outlined, with particular detail paid to the flavour sector. Various successes of the SM will then be discussed before going on to identify its shortcomings using arguments from both experiment and theory. These shortcomings will then be used to motivate the three analyses: a search for the decay $B^+ \to D_s^+ \phi$ (Chap. 5); a search for the two related decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$ (Chap. 6); a search for dark sector particles in $B^0 \to K^*(892)^0$ (Chap. 7). Theory specific to each of these analyses will be detailed in the relevant chapter.

2.1 The Standard Model

The behaviour of fundamental particles and forces are described by the SM of particle physics, which was concocted in the 1970s, when the Higgs mechanism was incorporated into Glashow's electroweak theory by Salam and Weinberg. The theory prescribes a treatment as to how fundamental particles interact via three of the four fundamental forces, namely: the strong, weak and electromagnetic forces.

Mathematically, the SM is a locally gauge invariant quantum field theory. It inhabits a space-time with a global Poincaré symmetry that obeys a local $SU_C(3) \otimes SU_L(2) \otimes U_Y(1)$

Perhaps say where Heathen you are following.

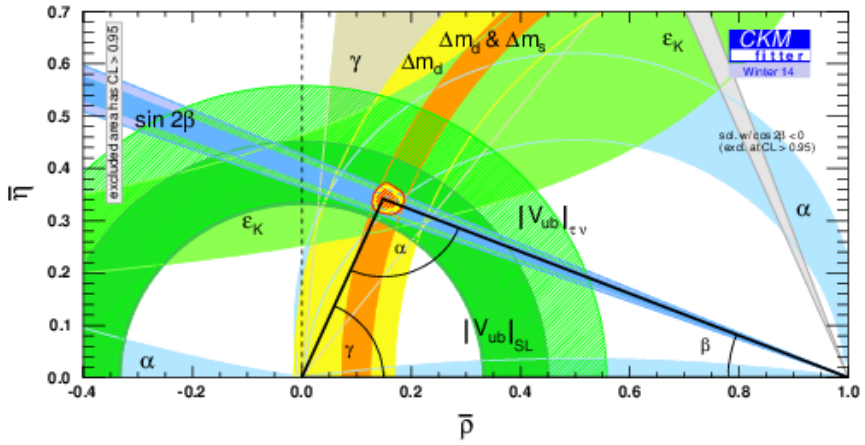

Figure 2.3: Diagram of the UT with coloured bands indicating various constraints on side lengths, angles and position of the apex, which is taken from the CKMfitter group in Ref. [25]. The constraints on V_{ub} from the combination of inclusive and exclusive modes ($|V_{ub}|_{\mathrm{SL}}$) is given separately to a value obtained using $\mathcal{B}\left(B^+ \to \tau^+\nu_\tau\right)$, ($|V_{ub}|_{\tau\nu}$).

measurements. A measurement from the LHCb experiment uses the baryonic decay $\Lambda_b^0 \to p\mu^-\overline{\nu}_\mu$ calculated a value of $|V_{ub}|$ to be $(3.27 \pm 0.23) \times 10^{-3}$ [22]. This is an exclusive measurement, and is in agreement with other exclusive measurements.

Another method to access the CKM matrix parameter $|V_{ub}|$ is via the annihilation-type decay $B^+ \to \tau^+\nu_\tau$. Measurements from both the BaBar and Belle experiments of $\mathcal{B}\left(B^+ \to \tau^+\nu_\tau\right)$ [23, 24] suffer from small statistics, but are found to be in better agreement with values of $|V_{ub}|$ determined using inclusive measurements than exclusive. Searching for the decay $B^+ \to \tau^+\nu_\tau$ is not viable at LHCb; instead, decays of the same topologies can be searched for. The decay $B^+ \to D_s^+\phi$ is also an annihilation-type decay in which V_{ub} appears in the amplitude; an analysis of this decay is described in Chap. 5.

Current measurements of angles and side lengths of the UT, from Ref. [25], are shown in Fig. 2.3. This figure also shows global V_{ub} measurements from the semi-leptonic and $B^+ \to \tau^+\nu_\tau$ modes alongside one another.

Unnatural NP models with parameters that differ wildly in magnitude tend to lead to parameters or processes that must cancel to absurdly high precision in order to agree with experimental observations. These precise cancellations are known as *fine tuning*. In the SM, quantum loop corrections to the Higgs mass are of the order 10^{19} for $m_H \simeq 125\,\mathrm{GeV}$ [8, 9]. This means that the cancellations required to result in a Higgs mass comparable to the masses of the weak vector bosons must be exact to 17 orders of magnitude. This instance of fine tuning is known as the *hierarchy problem*. A solution for the hierarchy problem is to introduce NP particles, whose contributions to loop level processes reduce the magnitude of fine tuning required to a level deemed acceptable. SUSY immediately solves the hierarchy problem because for every SM particle that contributes to the Higgs mass, a SUSY particle

Chapter 3

The LHCb experiment

The following chapter first introduces the LHC and the LHCb detector before briefly describing data collection and processing. Other important features, such as particle identification and triggering, of the LHCb experiment will also be outlined.

3.1 The LHC

The LHC is a superconducting synchrotron which can simultaneously accelerate beams of proton bunches in opposite directions. Physically, the LHC is located at CERN, near Geneva, Switzerland; it is 27 km in circumference and spans the Franco-Swiss border at a depth of about 100 m. Protons are supplied to the LHC from the Super Proton Synchrotron (SPS) with an energy of 450 GeV, they can then be accelerated and collided with a centre-of-mass energy of up to 14 TeV. In the years 2011 and 2012 collisions operated with a centre-of-mass energy was 7 and 8 TeV respectively. Once the desired energy is reached the beams are collided at four interaction points. The LHCb detector is situated at one of them [32]. Collisions of proton bunches occur every 50 ns reaching luminosities of up to $7 \times 10^{32}\,\mathrm{cm^2 s^{-1}}$, however the beams entering LHCb must be luminosity levelled, to $3(4) \times 10^{32}\,\mathrm{cm^2 s^{-1}}$ in 2011(2012), in order to reduce detector occupancy. These high energy collisions produce a vast number of $b\bar{b}$ pairs which subsequently hadronize into b hadrons. It is the prospect of studying these bound states of b (and other heavy flavour) quarks that has motivated the design of the LHCb detector.

3.2 The LHCb detector

Before introducing the LHCb detector, it is helpful to first define the Cartesian coordinate system around which the LHCb detector is built. The z direction is defined by the LHC

collected in 2012. In total, the data collected in 2011 and 2012 is known as Run-1 data.

Even the, much reduced, HLT2 output rate of 5 kHz is a vast amount of data for an analyst to sift through in a timely manner. To improve the speed to access data, additional selections are applied to the dataset biannually which further categorize each event. This is known as stripping. Stripped datasets are the only ones accessible to analysts, which makes the process of retrieving data of interest fast. Stripping selections in this thesis vary, and will be described when appropriate.

Reliable analysis of real data would not be possible without selections of simulated events. These allows collaborator's access to pure samples of specific, requested decays to aid their research. This can be for the evaluation of efficiencies, understanding effects in data, or making analysis decisions without compromising blinded data. These events are generated in two independent phases: generation and simulation. Proton-proton collisions are generated using Pythia [36] with a specific LHCb configuration [37], and subsequent hadronic decays are handled by the EvtGen [38] package. The simulation phase is designed to mimic the LHCb detector's response to particles, this is done with Geant4 [39] as described in Ref. [40]. Simulated events after the hadronization stage and before detector modelling are known as *generator level* simulation.

Chapter 4

Multivariate selection techniques

The analyses detailed in Chap. 5, Chap. 6 and Chap. 7 make prodigious use of multivariate techniques to reduce combinatorial backgrounds. Combinatorial backgrounds are formed from random combinations of tracks which appear to form a vertex, pass selection criteria, and satisfy relevant PID assignments. To remove these backgrounds Multivariate Analysis (MVA) techniques can be employed. A multivariate discriminator exploits correlations between weakly discriminating variables to produce a single, more separating, classifier.

MVA techniques used in HEP tend to be supervised learning algorithms, whereby a selection of events are given input, and an algorithm produces a response based on how best to separate them. Input into the algorithms to separate a background are: a sample of the signal and background candidates that must be separated, and a set of variables to be used to do so. Samples of events are split in two, some are used for training the MVA, and the remaining are used for testing it. The input, or training, variables define an n-dimensional space populated by the input samples. The algorithm then classifies regions in this n-dimensional space as signal- or background-like; such that an arbitrary event placed somewhere in the space would also be classified based on the point it inhabits. The Boosted Decision Tree (BDT) algorithm is used throughout this thesis because it can handle a weighted training sample, including negative weights, and can exploit non-linear correlations between variables [41, 42].

A BDT is composed of a combination of numerous Decision Trees (DTs), each of which is a classifier in its own right — albeit a weak one — being able to distinguish between high density regions of signal and background populations.

Training a DT begins with a single parent node populated by the whole training sample, which inhabits the parameter space defined by the variables x_i, whose true distribution is $f(x_i)$. The sample on the parent node is split by selecting a cut based on maximizing some figure-of-merit. Child nodes are then split, and the process is repeated until there is no possible improvement in the separation of signal and background. The definition of

improvement is usually related to the signal and background purity of a node:

$$p_{\text{sig}} = \big(1 - \varepsilon_{\text{bkg}}\big)$$
$$p_{\text{bkg}} = \big(1 - \varepsilon_{\text{sig}}\big). \tag{4.1}$$

The final child nodes, or leaves, are each associated with signal or background depending on the purity of the sample which populates it. Each leaf therefore maps out an area in n-dimensions, and is classified as a signal or background leaf depending on the purity of the training sample enclosed by that area. The hypothesised category, as output by the DT, $h(x_i)$, will ideally be equal to $f(x_i)$, but in reality there will be events which are misclassified. A figure-of-merit which is often used to determine the cut used at each node is the G_{ini} index, which is defined as

$$\mathrm{G}_{ini} = 2p_{\text{bkg}}p_{\text{sig}} = 2\big(1 - \varepsilon_{\text{sig}}\big)\big(1 - \varepsilon_{\text{bkg}}\big) = \frac{2sb}{(s+b)^2}, \tag{4.2}$$

where s and b are the weighted sum of signal and background candidates, respectively, after a given cut.

Decision Trees have the advantages over other machine learning algorithms — such as neural nets — of being able to deal with weighted training samples, and being insensitive to variables with very little separation power because the G_{ini} index never identifies a cut on them as being profitable. However, DTs are sensitive to statistical fluctuations in the training sample. To negate this problem DTs can be *Boosted* using any one of a number of algorithms. The procedure of boosting removes the power that statistical fluctuations has over the final BDT.

A different boosting method is used to train the BDTs in each analyses in this thesis. The algorithms used are outlined in the remainder of this chapter.

4.1 The bagging algorithm

Bootstrap aggregating, or bagging[1], is a method of boosting whereby the effects of statistical fluctuations are negated by making many independent DTs and using them to make a decision based on the average response. Training n DTs, it is possible to define three errors, namely: square of the error of a single estimator

$$\epsilon_t(x_i) = \big(f(x_i) - h_t(x_i)\big)^2 \tag{4.3}$$

[1] This section is based on Ref. [43].

Chapter 5

Search for the decay $B^+ \to D_s^+ \phi$

5.1 Introduction

In the SM the decay $B^+ \to D_s^+ \phi$ proceeds via the annihilation of the constituent b and u quarks of a B^+ meson forming a virtual W^+ boson from the CC interaction, the processes is suppressed by the CKM matrix element V_{ub}[1]. To achieve the final state, the W^+ decays into a $c\bar{s}$ pair and an additional $s\bar{s}$ pair must be created from the QCD field. This is the only diagram that can perpetuate such a decay at tree-level, because the initial state quarks are all different to those in the final state. A Feynman diagram of the decay $B^+ \to D_s^+ \phi$ is shown in Fig. 5.1, where the final state mesons can be formed in the way indicated, or the $s\bar{s}$ pair from the QCD field can form the ϕ, although this is colour-suppressed. Also, the gluon that forms the $s\bar{s}$ pair can originate from any of the initial or final state quarks. This analysis was published in Ref. [1].

Annihilation decays of B^+ mesons are rare in the SM due to the magnitude of $|V_{ub}| \sim 4 \times 10^{-3}$. In fact, no fully hadronic decays proceeding via annihilation-type diagrams have yet been observed.

Predictions for the branching fraction $\mathcal{B}\big(B^+ \to D_s^+ \phi\big)$ are calculated using the OPE defined by the effective Hamiltonian [47–50]:

$$\mathcal{H}_{\text{eff}} = -4\frac{G_F}{\sqrt{2}} V_{ub} V_{cs}^* \big[C_1(\Lambda)\mathcal{O}_1 + C_2(\Lambda)\mathcal{O}_2\big] \tag{5.1}$$

where

$$\begin{aligned} \mathcal{O}_1 &= \big(\bar{b}\gamma_\mu P_L u\big)\big(\bar{c}\gamma_\mu P_L s\big) \\ \mathcal{O}_2 &= \big(\bar{b}\gamma_\mu P_L s\big)\big(\bar{c}\gamma_\mu P_L u\big). \end{aligned} \tag{5.2}$$

[1] All mentions of the ϕ meson refers to the $\phi(1020)$.

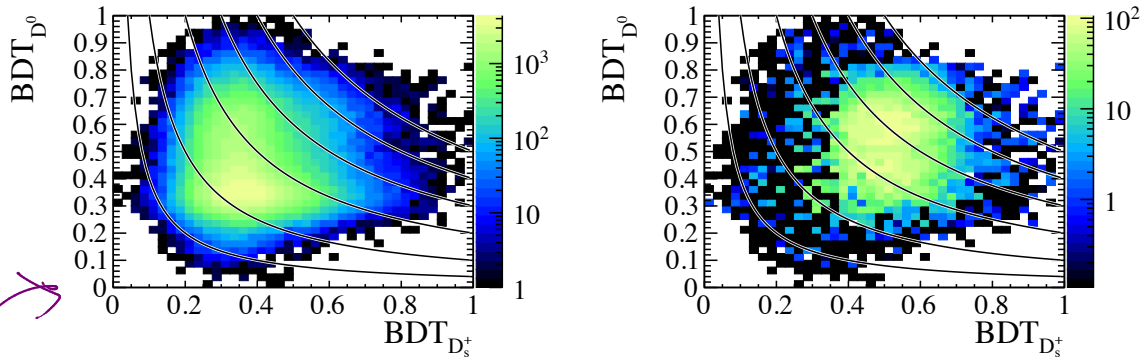

Figure 5.2: An example distribution of the response for a BDT$_{D^0}$ against a BDT$_{D_s^+}$ for a sample of (left) $B^+ \to D_s^+ \overline{D}^0$ candidates where the background dominates and occupies the lower left of the plots, and (right) the same candidates after sWeighting has been applied where the signal peak is observed further towards the upper right. The lines overlayed on the plot are show the boundaries for the cuts of BDT$_{D_s^+}$ × BDT$_{D^0} > 0.01$, 0.04, 0.10, 0.20, 0.30, 0.40, and 0.50. It is seen that using the product of the BDT discriminants is more effective for rare decays where a tighter cut will be needed, in this region candidates are selected if their response is particularly positive, at the expense of the other meson.

of using a BDT to identify each meson is also used in Ref. [53], which measures branching fraction ratios of various $B \to DD$ decays. The bagging boosting technique used gives a response between zero and one. Therefore, it is natural to cut on the product of the two BDT responses, BDT$_{D_s^+}$ × BDT$_\phi > X$, as opposed to BDT$_{D_s^+} > X_1$ and BDT$_\phi > X_2$. Cutting on the product of the BDTs improves the performance of the selection, because a very strong D_s^+ candidate (for example) will be selected at the expense of a slightly weaker ϕ selection, this is particularly true for tighter cuts. Figure 5.2 shows the effect of cutting on BDT$_{D_s^+}$ × BDT$_{D^0}$ in the normalization channel $B^+ \to D_s^+ \overline{D}^0$.

The $D_s^+ \to K^+K^-\pi^+$BDT was trained using $D_s^+ \to K^+K^-\pi^+$ decays from data taken from the high statistics channel $\overline{B}_s^0 \to D_s^+\pi^-$. The signal sample of D_s^+ decays came from the $D_s^+ \to K^+K^-\pi^+$ candidates that fell within 3σ of the known D_s^+ mass, the sample was background subtracted using the sWeighting technique [54] on the B^+ mass spectrum. Background data was taken from candidates falling in the upper mass sideband of the B^+, and either sideband of the D_s^+. Similarly, for the ϕ BDT, the signal sample was sWeighted and the background comes from the ϕ mass sidebands; but the sample is taken from the high statistics $B_s^0 \to J/\psi\,\phi$ mode.

In total, there are five kinematic and geometric training variables for the parent meson. For the daughter tracks there are a total of 23 variables, including kinematic, geometric and PID variables. Since the BDT was trained using data, it is possible to use PID variables that are poorly described in simulation. A summary of all training variables is given in Table 5.2.

The cut for the BDT was optimized using the metric $S/\sqrt{S+B}$, In this case, the number

Table 5.2: List of training variables used in the D_s^+ and ϕ BDTs. Each BDT uses five variables associated with the parent particle and 23 variables from each daughter track.

Particle		Variable
D_s^+, ϕ	Kinematic variables	p, p_{T}
	Geometric variables	χ^2_{vtx}, χ^2_{IP}, χ^2_{FD}
Tracks	Kinematic variables	p, p_{T}
	Geometric variables	$\min(\chi^2_{\mathrm{IP}})$
	Track variables	4 variables characterizing the track quality
	PID variables	16 variables containing PID information, such as `isMuon` and DLL variables from the RICH detectors

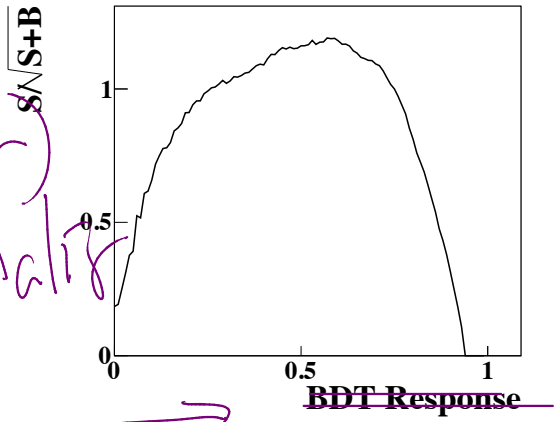

Figure 5.3: Value of the figure of merit $S/\sqrt{S+B}$ is shown as a function of the BDT response, $\mathrm{BDT}_{D_s^+} \times \mathrm{BDT}_{\phi}$. The maximum value of the figure of merit is 0.57, which is chosen as the final BDT cut.

of signal events, S, was estimated from the yield from the decay $B_s^0 \to D_s^- \pi^+$, according to:

$$S = \frac{\mathcal{B}\left(B^+ \to D_s^+ \phi\right)}{\mathcal{B}\left(\overline{B}_s^0 \to D_s^+ \pi^-\right)} \frac{\varepsilon_{\mathrm{gen}}\left(B^+ \to D_s^+ \phi\right)}{\varepsilon_{\mathrm{gen}}\left(\overline{B}_s^0 \to D_s^+ \pi^-\right)} \frac{f_d}{f_s} N\left(\overline{B}_s^0 \to D_s^+ \pi^-\right), \tag{5.5}$$

f_s/f_d quantifies the fraction of B_s^0 mesons produced relative to B^0 mesons. The generator-level efficiency, $\varepsilon_{\mathrm{gen}}$, is the efficiency introduced by the acceptance region of the LHCb detector, and the necessity that all daughter particles must travel through the detector. Background yield for a given cut is estimated as:

$$B = c \cdot N_{\mathrm{c}}\left(B_s^0 \to D_s^- \pi^+\right) \cdot N_{\mathrm{c}}\left(B_s^0 \to J/\psi\, \phi\right), \tag{5.6}$$

where N_{c} indicates the yield of the combinatoric background for the indicated decay, and c is a constant scaled such that $N_{\mathrm{c}}\left(B_s^0 \to D_s^- \pi^+\right) \cdot N_{\mathrm{c}}\left(B_s^0 \to J/\psi\, \phi\right) = N_{\mathrm{c}}\left(B^+ \to D_s^+ \phi\right)$ with no BDT cut. The optimization procedure results in the optimal cut as $\mathrm{BDT}_{D_s^+} \times \mathrm{BDT}_{\phi} > 0.57$, as is shown in Fig. 5.3.

Both the signal and normalization channels have a $D_s^+ \to K^+K^-\pi^+$, and therefore efficiencies from the D^+ and Λ_c^+ vetoes cancel to a large extent. Given that these cuts are very efficient, the difference is negligible, and the efficiency ratio is assumed to be unity for the calculation. The only differences between the two modes are due the decay kinematics, since the ϕ is lighter than the D^0. There is a difference of 1.5 % between $\varepsilon_{\text{veto}}$ for the decays $\overline{B}_s^0 \to D_s^+\pi^-$ and $B^+ \to D_s^+\overline{D}^0$, and since the mass of the ϕ is nearer the D^0 mass than the pion mass, a systematic uncertainty of 1 % is assigned.

There are other systematic uncertainties that affected the selection. The BDT cut is assigned an uncertainty of 3 %, which is due to the sizes of the validation samples used to calculate the efficiencies. Mass windows around the D_s^+ and ϕ lead to a 3 % systematic uncertainty. Also, the low statistics of the simulation samples used to deduce efficiencies led to a 3 % uncertainty.

The mass fits introduced systematic uncertainties from each component. If the parameters describing the signal shape are allowed to float the fit results in a yield which is 5% higher than the nominal fit. This is assigned as a systematic uncertainty.

The total uncertainty from the background shape is 5 %; and is estimated by making changes to the background model. By removing either the $B^+ \to D_s^{*+}\phi$ or $\overline{B}_s^0 \to D_s^{*+}K^{*0}K^-$ components, the yield changes by only 1 %. Changing the constraints on **A**/**B** and **C**/**D** for $\overline{B}_s^0 \to D_s^+K^{*0}K^-$ by a factor of 2 results in a 1 % change in signal yield. The combinatorial background is estimated by allowing the slope to float free, this led to an approximate 3.5 % systematic uncertainty.

Contributions from all sources of systematic uncertainties are summarized in Table 5.6. The dominant systematic uncertainty — discounting the uncertainties on the branching fraction of the normalization channel — is from the mass fits, which is unsurprising considering the complexity of the fit, and the treatment of the backgrounds. Regardless of the constraints that are, or are not, included in the fit: the lowest significance obtained is still greater than $3\,\sigma$. Therefore the significance is quoted as greater than $3\,\sigma$.

Describe iterative procedure, Is constrain effect of BDT cut value

5.5 Direct *CP* asymmetry

The *CP* asymmetry is defined in Eq. 5.4, but this must be modified to account for background in the sample:

$$\mathcal{A}_{CP}\left(B^+ \to D_s^+\phi\right) = \frac{N(B^- \to D_s^-\phi) - N(B^+ \to D_s^+\phi)}{N(B^- \to D_s^-\phi) + N(B^+ \to D_s^+\phi) - N_{\text{bkg}}}. \tag{5.10}$$

Here, the yields, N, refer to the decay with the indicated charge, and N_{bkg} denotes the amount of background contaminating the signal.

Chapter 6

Search for the decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$

6.1 Introduction

The decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$ both are $b \to s\mu^+\mu^-$ FCNC transitions, which are forbidden at tree-level in the SM[1]. Therefore, these processes are sensitive to virtual NP particles contributing to the decay amplitude in loops. The analysis described in the following chapter makes a measurement of the branching fractions of both decays and the differential branching fraction of $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ in bins of q^2, where q^2 is the invariant mass of the dimuon system squared. This analysis was published in Ref. [2].

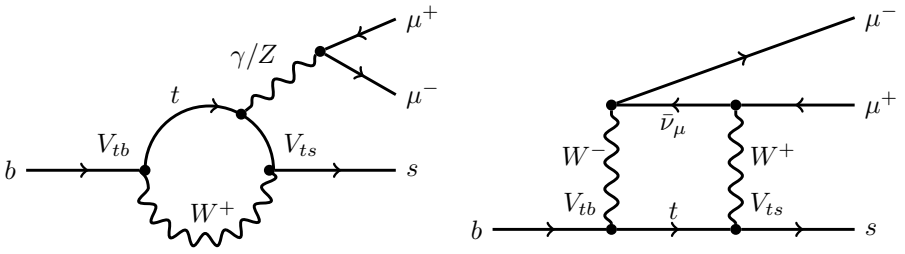

Figure 6.1: Schematic Feynman diagrams for the operators $\mathcal{O}_7$, $\mathcal{O}_9$, and $\mathcal{O}_{10}$ which are most sensitive to the $b \to s\mu^+\mu^-$ FCNC. The propagators are the (left) photonic and Z penguin diagram, and the (right) W^+-mediated box diagram. Operator $\mathcal{O}_7$ describes the photonic pengiuin diagram; while $\mathcal{O}_9$ and $\mathcal{O}_{10}$ are the vector and axial-vector parts of both the Z and W^+ diagrams.

[1] All mentions of the ϕ refer implicitly to the $\phi(1020)$ meson.

Chapter 7

Search for the decay of a dark sector particle $\chi \to \mu^+\mu^-$ in $B^0 \to K^{*0}\mu^+\mu^-$

7.1 Introduction

This chapter describes a procedure for searching for a dark boson, χ, of unknown mass and lifetime[1]. A frequentist method is applied to the dimuon distribution of $B^0 \to K^{*0}\mu^+\mu^-$ candidates, to search for an excess of events above the SM background, consistent with a χ decaying into a pair of muons. Lifetime information is added by splitting candidates into two bins of decay time: those which are prompt, and the χ vertex is the same as the K^{*0} vertex; and those which are displaced.

Chapter 2 explains that the SM cannot explain the numerous experimental observations of DM. Little is known about dark matter, except that it interacts gravitationally, and does not interact with electromagnetic radiation to any significant extent. A possible extension to the SM is to introduce a dark sector, which can contain a rich variety of distinct particles operating through forces that are hitherto unknown. Dark sector particles would be gauge singlet states with respect to the SM, and only be able to communicate with known particles via weakly interacting messenger particles through one of four *portals*: the vector, axion, Higgs, and neutrino portals [73]. Interaction terms for messengers in each of these portals are given in Table 7.1.

Theories involving dark sectors are tremendously attractive because it is relatively easy to construct a complex theory that explains various unexplained phenomena. Yet, these can have little impact on the SM observables since the interaction between the two sectors can be extremely weak.

The Higgs portal has a scalar messenger particle which can mix with the SM Higgs. There

[1] Throughout this chapter, the symbol χ shall denote a general dark boson and all references to a K^{*0} will be implicitly referencing the $K^*(892)^0$; unless explicitly stated otherwise.

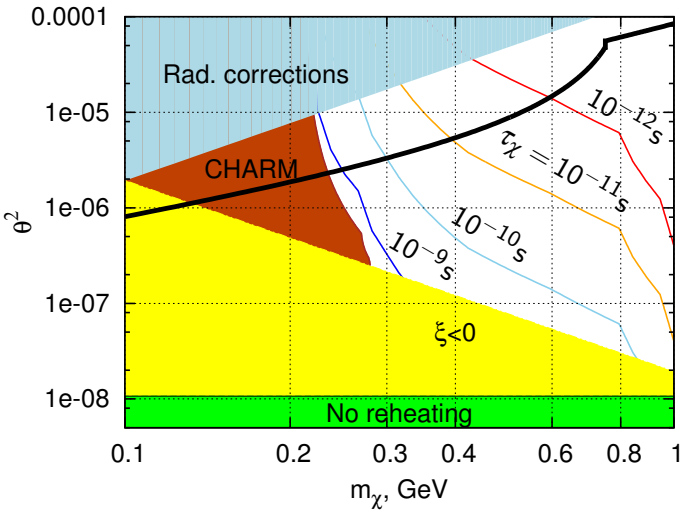

Figure 7.1: Allowed and excluded regions of the Higgs-inflaton mixing parameter squared as a function of inflaton mass, taken from Ref. [77].

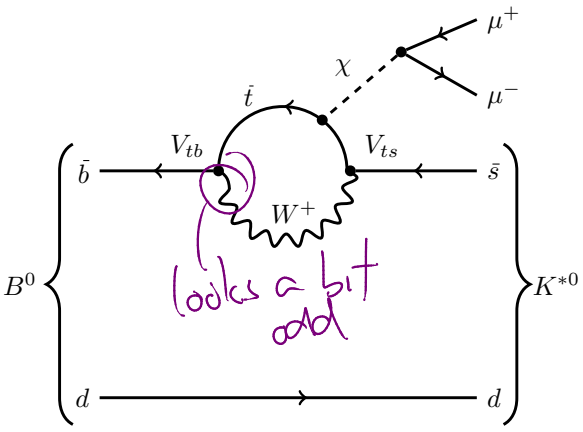

Figure 7.2: Feynman diagram showing the decay $B^0 \to K^{*0}\chi$, and $\chi \to \mu^+\mu^-$. Depending on the portal through which the χ acts, it couple directly to the muons, or need to mix with a SM Higgs, Z, or photon.

Table 7.2: Selected particles properties for mesons which decay into a dimuon pair final state, and the branching fractions for the relevant decays [3]. Central values of mass and width are given in MeV.

Meson (X)	Mass	Width	$\mathcal{B}(B^0 \to K^{*0}X)$	$\mathcal{B}(X \to \mu^+\mu^-)$	$\mathcal{B}_{\text{tot}}(B^0 \to K^{*0}\mu^+\mu^-)$
η	547.9	0.001	$(1.59 \pm 0.10) \times 10^{-5}$	$(5.8 \pm 0.8) \times 10^{-6}$	$(9.2 \pm 1.4) \times 10^{-11}$
ρ	775.3	147.8	$(3.9 \pm 1.3) \times 10^{-6}$	$(4.55 \pm 0.28) \times 10^{-5}$	$(1.8 \pm 0.6) \times 10^{-10}$
ω	782.7	8.5	$(2.0 \pm 0.5) \times 10^{-6}$	$(9.0 \pm 3.1) \times 10^{-5}$	$(1.8 \pm 0.8) \times 10^{-10}$
ϕ	1019.5	4.3	$(1.00 \pm 0.05) \times 10^{-5}$	$(2.87 \pm 0.19) \times 10^{-4}$	$(2.9 \pm 0.2) \times 10^{-9}$
D^0	1864.8	10.1	$(4.2 \pm 0.6) \times 10^{-5}$	$< 6.2 \times 10^{-9}$	$< 2.6 \times 10^{-14}$
J/ψ	3096.9	0.093	$(5.96 \pm 0.03) \times 10^{-2}$	$(1.32 \pm 0.06) \times 10^{-3}$	$(7.9 \pm 0.4) \times 10^{-5}$
$\psi(2S)$	3686.1	0.299	$(7.9 \pm 0.9) \times 10^{-3}$	$(6.0 \pm 0.4) \times 10^{-4}$	$(4.7 \pm 0.6) \times 10^{-6}$

low and high q^2 and reduce experimental uncertainties by removing the region around the ϕ meson which is centred at $q^2 \simeq 1.04\,\text{GeV}^2$.

7.2.5 Resonant backgrounds

In the absence of resonances, the dimuon background from the SM decay $B^0 \to K^{*0}\mu^+\mu^-$ is expected to be locally-linear to within 1 % over the entire mass range. However, the inclusion of resonances can lead to significant departures from linearity, dependent upon the resonance's width (Γ), magnitude, and the value of x. Table 7.2 lists a number of mesons that decay to a dimuon final state and could contribute as background to the decay $B^0 \to K^{*0}\mu^+\mu^-$. Wide resonances, $\Gamma \gtrsim 20\sigma_m$, such as the ρ, are sufficiently wide not to be problematic, even if they dominate the local background distribution. Conversely, narrow resonances where $\Gamma \lesssim 5\sigma_m$ lead to significant deviations from a locally-linear background and must be vetoed. Dimuon decays of the ϕ, J/ψ, and $\psi(2S)$ mesons have the highest branching fractions, and are also among the narrowest resonances; they are therefore vetoed. Intermediate resonances, $5 \lesssim \Gamma \lesssim 20\sigma_m$, are considered on a case-by-case basis since they are only troublesome if they account for a large fraction of the local background. For this analysis, these ranges roughly translate to requiring that resonances with $\Gamma < 25\,\text{MeV}$ will be vetoed, and those with $\Gamma > 125\,\text{MeV}$ shall be ignored. Other resonances in Table 7.2 are broad, and contribute to the dimuon structures at low mass. It is shown in Ref. [83] that the local-linearity approximation is accurate to ∼5 % in regions where there may be contributions from wide resonances. Below the mass of the J/ψ the values of $x = 5$ and $\sigma_y = 0.05$ are chosen.

Resonances that have a natural width in the intermediate region, $5 \lesssim \Gamma \lesssim 20\sigma_m$, are various $c\bar{c}$ states with masses above the mass of the $\psi(2S)$. For example, there is contribution from the $\psi(4160)$ which was observed by LHCb in the decay $B^+ \to K^+\mu^+\mu^-$ [86]. All known charmonium resonances in this region are wide, $\Gamma \sim 70\,\text{MeV}$, and are dealt with by reducing the size of the sidebands and increasing the uncertainty on the background

Table 7.5: Selection criteria applied to signal candidates in the stripping. Criteria definitions are defined in text. While the B^0 mass is constrained in the fit, the selection makes a cut on the unconstrained mass.

Candidate	Selection criterion			
B^+	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	25	
	χ^2_{IP}	$<$	50	
	τ	$>$	0.2	ps
	m	$\in$	$[4800, 5800]$	MeV
	p_{T}	$>$	1000	MeV
	$\cos\theta_{\text{dir}}$	$>$	0	
χ	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	10	
	χ^2_{FD}	$<$	25	
	p_{T}	$>$	250	MeV
	DOCA	$<$	0.2	mm
	DOCA χ^2	$<$	25	
Tracks	$\chi^2_{\text{trk}}/\text{ndf}$	$<$	3	
	$\min(\chi^2_{\text{IP}})$	$>$	9	
	$\mathcal{P}_{\text{gh}}$	$<$	0.3	
K^+, π^+	p_{T}	$>$	250	MeV
	p	$>$	2000	MeV
	χ^2_{IP}	$>$	9	
K^+	$\mathcal{P}_K$	$>$	0.1	
π^+	$\mathcal{P}_\pi$	$>$	0.2	
μ^+	p_{T}	$>$	100	MeV
	`PIDmu`	$>$	-5	
	`isMuon`		True	

$\text{DLL}_{K\pi}(\pi^+) > 10$.

The decay $K^0_{\text{S}} \to \pi^+\pi^-$ has a branching fraction of $(69.20 \pm 0.05) \times 10^{-2}$ [3], and is removed in the preselection by requiring that $|m_{\pi^+_\mu\pi^-_\mu} - m^{\text{PDG}}_{K^0_{\text{S}}}| < 25\,\text{MeV}$. This roughly translates to a cut in the dimuon invariant mass spectrum of $436 < m_{\mu^+\mu^-} < 490\,\text{MeV}$.

Since the χ can be displaced from the B^0 decay vertex, a potential background for the decay $\chi \to \mu^+\mu^-$ is from a $\mu^+\mu^-$ pair directly from a PV. This is suppressed by requiring that the transverse flight distance (FD_T) of the χ vertex, with respect to the PV, is greater than 0.1 mm.

The high branching fraction of $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ means that there is also contamination from decays where a hadron is misidentified as a muon, and vice versa. This background can be suppressed by requiring that neither of the K^{*0} daughters satisfy the `isMuon` criteria. Figure 7.6, shows the invariant mass of the hadron-muon pair before and after the requirement that the hadrons fail the `isMuon` criterion. A summary of all preselection cuts is shown in Table 7.6.

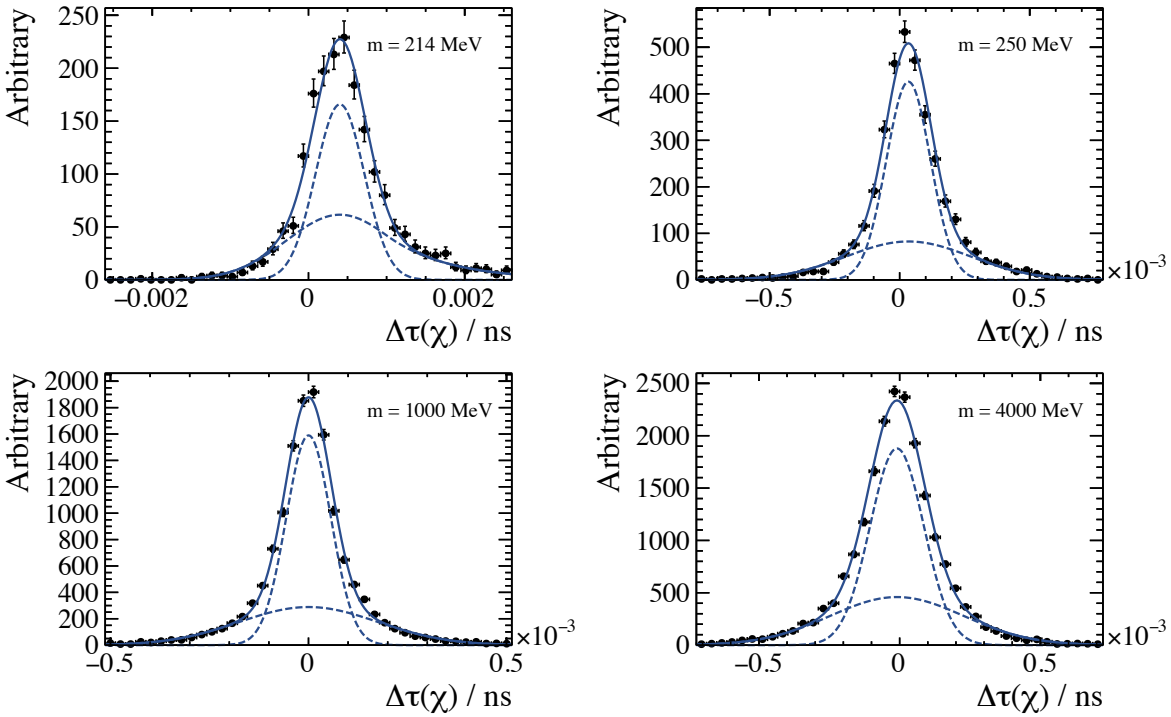

Figure 7.12: Fits to the lifetime resolution parameter, $\Delta\tau$, for individual mass samples. Each fit for $m \geqslant 250\,\text{MeV}$ is made using a double Gaussian function, and for $m < 250\,\text{MeV}$ the wider Gaussian has an exponential tail on the right-hand side. The solid line shows the total fit, and the dashed lines indicate the two components.

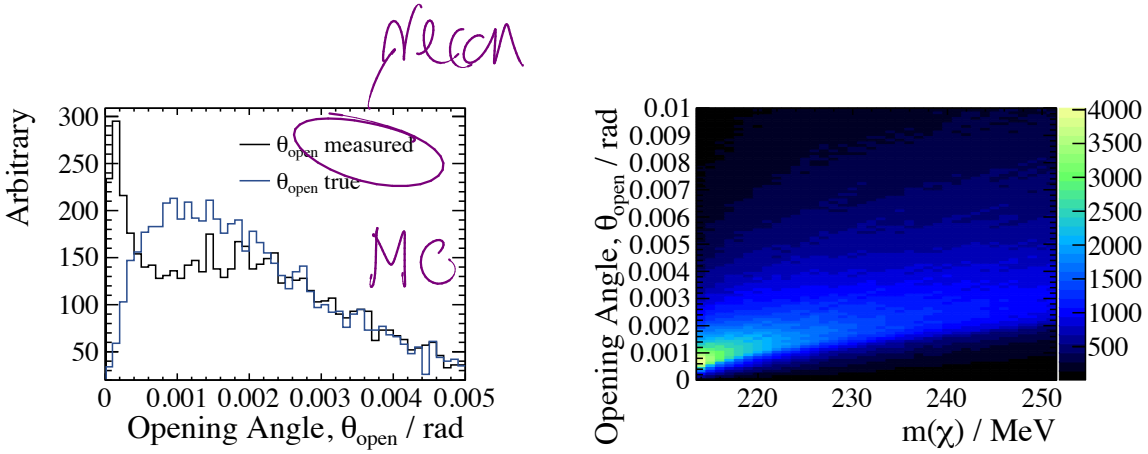

Figure 7.13: Opening angle of the two muons decaying from the χ is highly dependent on m_χ; for low dimuon masses the opening angle measurement is biased. For $m_\chi = 214\,\text{MeV}$ the difference between generated and reconstructed opening angles are shown, (left) the greatest difference being below $\theta_{\text{open}} = 0.002$. Generator level distributions of opening angles for a range of masses, (right) show that most χ decays have $\theta_{\text{open}} > 0.002$ for $m_\chi \gtrsim 235\,\text{MeV}$.

Chapter 8

Conclusions

This thesis presents three analyses: two have been published, claiming first observations [1, 2], and the third is currently under review; all were undertaken using data collected by the LHCb experiment [32]. Each has the objective of finding evidence of physics BSM in the decays of B mesons. The analysis techniques employed are all slightly different, most obviously the first two analyses are indirect searches for NP, while the third is a direct search for a dark boson of unknown mass and lifetime.

An analysis of the decay $B^+ \to D_s^+ \phi$ is presented in Chap. 5. First evidence for the decay was seen with a statistical significance of greater than $3\,\sigma$; this also constitutes the first evidence for a fully hadronic decay via an annihilation-type diagram. The branching fraction measurement made in this analysis is sensitive to NP effects contributing to the decay $B^+ \to D_s^+ \phi$ can only propagate via one diagram at tree-level, and this is suppressed by a factor $|V_{ub}|^2$. The element V_{ub} has the largest uncertainties in the CKM matrix and there are historic tensions between values of V_{ub} made using inclusive an exclusive modes.

The branching fraction measurement

$$\mathcal{B}\big(B^+ \to D_s^+ \phi\big) = \big(1.87\,^{+1.25}_{-0.73}(\text{stat}) \pm 0.19(\text{syst}) \pm 0.32(\text{norm})\big) \times 10^{-6}$$

is somewhat higher than SM predictions, which are of order 10^{-7} [47–50], but not incompatible considering large theoretical uncertainties. This value of $\mathcal{B}(B^+ \to D_s^+ \phi)$ sheds no light on the true value of V_{ub}. Reference [90] asserts that these discrepancies cannot be explained by physics BSM, and is rather due to underestimated uncertainties in either theory or experiment.

Since the decay $B^+ \to D_s^+ \phi$ is mediated by a W^+ boson, in NP scenarios another charged boson, such as a H^+ from a 2HDM, could contribute to the decay amplitude. These additional processes could alter the branching fraction significantly, which is not observed, or introduce extra phases into the decay, causing the CP-asymmetry to deviate from $\mathcal{A}_{CP} = 0$, as expected in the SM. The value measured, after correcting for production and

detection asymmetries, is

$$\mathcal{A}_{CP}\big(B^+ \to D_s^+\phi\big) = -\big(0.01 \pm 0.41(\text{stat}) \pm 0.03(\text{syst})\big),$$

which is perfectly consistent with SM expectations.

Chapter 6 presents an analysis leading to the first observations of the two decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$, their branching fractions were measured, as was the differential branching fraction of $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ in bins of q^2. The integrated branching fractions of these decays are

$$\mathcal{B}\big(B^+ \to K^+\pi^+\pi^-\mu^+\mu^-\big) = \big(4.36\,^{+0.29}_{-0.27}(\text{stat}) \pm 0.21(\text{syst}) \pm 0.18(\text{norm})\big) \times 10^{-7},$$
$$\mathcal{B}\big(B^+ \to \phi K^+\mu^+\mu^-\big) = \big(0.82\,^{+0.19}_{-0.17}(\text{stat})\,^{+0.10}_{-0.04}(\text{syst}) \pm 0.27(\text{norm})\big) \times 10^{-7},$$

and both have statistical significances of greater than $5\,\sigma$. The decay $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ has a large branching fraction and could be used for future analyses similar to those of interest in $B^0 \to K^{*0}\mu^+\mu^-$, since both are sensitive to the same operators. An angular analysis would help constrain the scalar, pseudoscalar, and tensor amplitudes of the decay, all of which are vanishingly small in the SM. This is made difficult by the number of contributing states to the $K^+\pi^+\pi^-$ system, but with more statistics it will be possible to gain a better understanding of strange states that decay into kaons and pions.

Given larger statistics, the additional channels $B^+ \to K^+K^-\pi^+\mu^+\mu^-$ and $B^+ \to \pi^+\pi^+\pi^-\mu^+\mu^-$ may be observable, and give access to the ratio of CKM matrix elements V_{td}/V_{ts}. These would be complimentary to the current measurements from B-meson oscillations.

The large uncertainties in the measurement of $\mathcal{B}(B^+ \to \phi K^+\mu^+\mu^-)$ are primarily due to uncertainties propagated from the branching fraction of the normalization channel $B^+ \to J/\psi\,\phi K^+$. The paper in Ref. [2] quotes the ratio of branching fractions in order for $\mathcal{B}(B^+ \to \phi K^+\mu^+\mu^-)$ to be calculated given an improved measurement.

Finally, a direct search for a NP particle, χ, belonging to some dark sector is presented. Using $B^0 \to K^{*0}\mu^+\mu^-$ candidates, the dimuon invariant mass sectrum is searched for a signal indicative of a dark boson decaying via $\chi \to \mu^+\mu^-$. The selection is designed specifically not to bias any corner of the mass-lifetime space that the χ might inhabit, this is done primarily with the aid of the uBoost algorithm. Efficiencies and resolutions are parameterized using discrete simulated samples of $B^0 \to K^{*0}\chi$ and spline interpolation is used to understand selection and resolution effects for any value of m_χ. A novel frequentist strategy was employed which scans in mass for and, at each point, calculates the local p-value that the observed signal is consistent with the null hypothesis of zero signal. One the look-elsewhere effect was accounted for, the significance of the minimum local p-value was equvilent to a significance of $0.48\,\sigma$ at $0.48\,\sigma$ at $m_{\mu^+\mu^-} = 4285.0\,\text{MeV}$. This

is consistent with with no new particle. Further studies will push tested masses to the boundaries of vetoed regions. The projected limits for an inflaton model indicate that much of the allowed parameter space will be excluded for $m_\chi < 1000\,\text{MeV}$. A similar approach could be used in any arbitrary mass spectrum to search for a multitude of particles appearing above a smoothly varying background.

In conclusion, the SM continues in its resilience, seeming to be in agreement to the limit of accuracy that experimental high energy particle physics can reach. There is a complementarity that exists between indirect and direct measurements. Historically, it has often been the case that indications of future discoveries were first anticipated by observations made by indirect experiments. This has been the case because indirect measreuments are sensitive to loop-level processes and therefore virtual particles. This pattern looks set to endure, as precision measurements in the flavour sector continue to play an important role in cornering the nature of NP. As of Run 1 of the LHC, there has been no clear indication of where new physics may lie; but, there are a number of discrepancies between SM predictions and measurements of observables in the flavour sector with significances greater than $3\,\sigma$. Considering the current landscape, it is perhaps increasingly difficult to interpret these discrepancies, and how they impact various NP models.

Unfortunately, interpretation of precise measurements of B physics observables that can be made experimentally are often made difficult by the form-factor parameterization which must be adopted. It is these QCD effects, particularly the form-factors, that are the dominant sources of theoretical uncertainty. Difficulties in dealing with QCD has been demonstrated by a lack of consistent predictions for $\mathcal{B}(B^+ \to D_s^+ \phi)$, and the absnece of any predictions for either $\mathcal{B}(B^+ \to K^+\pi^+\pi^-\mu^+\mu^-)$ and $\mathcal{B}(B^+ \to \phi K^+\mu^+\mu^-)$.

The idea that nature is natural, is an attractive one. As such, it is not unreasonable to expect — or at least hope — that NP lives just around the corner. Run-2 of the LHC will collide protons with a centre-of-mass energy of about 14 TeV with the aim to see signals indicative of NP. That being said, there are still areas of parameter space of various theoretical scenarios that are accessible at lower masses to probe. In the absence of direct evidence from the high energy frontier, and the plethora of dark sector models which contain weakly interacting messenger particles, the intensity frontier may be a good place to search.

The interplay between direct and indirect searches in the arena of high energy physics has long been important, and will continue to be so.