High Energy Physics
Blackett Laboratory
Imperial College London

First evidence for the violation of lepton flavour universality in rare beauty-quark decays

Răzvan-Daniel Moise

Submitted to Imperial College London
for the degree of Doctor of Philosophy,
July 2021

Abstract

Fundamental particles and their interactions are currently best described by the Standard Model of particle physics. One of the predictions of this theory is the fact that charged leptons (electron, muon, and tau) interact in the same way with other particles, notwithstanding the Higgs interactions that give them different masses. This principle is known as lepton flavour universality, and has been tested through a host of various measurements. One of these is the ratio between the rates at which two beauty-meson decays, $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$, occur. This ratio is known as R_K, and its most precise measurement to date is the subject of this thesis. The result is obtained using 9 fb^{-1} of proton-proton collision data, recorded between the years 2011 and 2018 by the LHCb detector at CERN's Large Hadron Collider. The result is $R_K = 0.846\,^{+0.042}_{-0.039}\,^{+0.013}_{-0.012}$, where the first uncertainty is statistical, and the second systematic. This measurement is in tension with the Standard Model prediction at the level of $3.1\,\sigma$. It therefore constitutes evidence for the violation of lepton flavour universality in $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ decays. Subsequent measurements of R_K and related observables are expected to improve the global picture, potentially leading to the discovery of physics currently beyond the Standard Model.

Statement of originality

This thesis focuses on the measurement presented in Ref. [1]. It is the result of the research I conducted between February 2018 and March 2021, with the support of the High Energy Physics group at Imperial College London, and members of the LHCb collaboration. I performed the analysis work presented in this thesis, with a few exceptions that are highlighted at the beginning of each chapter, and also throughout the main body where relevant. In particular, the procedure heavily relies on the one developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the measurement presented in Ref. [2]. Additional contributions were made by Davide Lancierini and Dr. Konstantinos Petridis. The results and figures that are not a product of my own research are appropriately referenced.

This thesis has not been submitted for any other qualification.

Răzvan-Daniel Moise, July 2021

Copyright declaration

Acknowledgements

I was fortunate to have fantastic company in the four-year journey that culminated in this work. My gratitude is chiefly directed towards my supervisor, Mike McCann. I will forever cherish the depth and breadth of knowledge I've gained by being your student. You've shaped me into a person that is significantly more versed in physics, croquet, and sarcasm (well, at least one of these three may be up for debate). I find myself unable to put into words just how much I appreciate and respect you.

Science is all about standing on the shoulders of giants, and in this context the giants are the proponents of the measurement superseded by this thesis. Mitesh, thank you for being an exceptional leader, especially when it comes to dealing with tricky situations and pushing me to go the extra mile. Paula, I really appreciate that you've shared with me the ins and outs of this analysis, as well as how you were always there with a positive spin and kind words when things got tough. Thibaud, we simply could not have done it without you. I am truly humbled by the dedication and work you've put into this measurement before I joined it.

Thank you Davide for the time spent together working on producing the indescribable amount of intermediary plots, partial results, back-of-the-envelope calculations, things that were just wrong, you name it. My day-to-day activities were made way more enjoyable by having you around to share experiences and results. Huge thanks to Nico, Patrick, Kostas, and Sam for the unwavering support, fruitful discussions, and, of course, the memes. I could not have asked for a better team.

I would also like to thank various members of the the LHCb group at Imperial: Dave, Mark, Tai-Hua, Alex, Malte, Matt B1, Matt B2, Hanae, and Rebecca. I will miss the office and our coffee breaks. Special thanks to Will and Felix for going above and beyond to support me with all sorts of matters, physics-related or otherwise. Thank you Ryan, Ed, Jonno, Sergey, Vilius, Bogdan, and Shameena for being fantastic friends and colleagues. I hope that we will share many more pints together.

Nu în ultimul rând, aș vrea să le mulțumesc părinților mei pentru faptul că m-au susținut și iubit neîncetat, indiferent de cât de dificil a fost drumul. Lucrarea asta v-o dedic vouă.

Contents

I Lepton flavour universality, in theory and in practice 29

III Appendices 227

List of Tables

List of Figures

Chapter 1

Introduction

The fundamental constituents of matter, and their interactions, are currently best described by the Standard Model (SM). This theoretical framework is the result of decades of research, during which it has withstood the test of experimental scrutiny time and time again.

Despite this, the SM is not complete. It cannot explain several effects, such as gravity, and so the SM has to be expanded if it is to describe such phenomena. One way of identifying areas where expansion is needed is by testing the validity of the assumptions made by the SM.

Lepton flavour universality (LFU) is one such assumption. It was tested in the past and found to be valid in a number of processes, however more recent measurements have begun to cast doubt on the scope of LFU. A number of observables that are sensitive to the violation of LFU are showing signs of disagreement with their SM predictions. These quantities are part of a class of measurements that are generally in tension with SM expectation, known as the “flavour anomalies”. Physics beyond the SM does not necessarily obey LFU, so the discrepancies between theory and experiment could be caused by LFU-violating physics beyond the SM.

Individually, the flavour anomalies do not yet have the experimental precision required to definitively rule out the SM. Collectively, however, they can provide complementary information, and hence there are efforts to explain the flavour anomalies together. Such studies involve extending the SM theory to include New Physics (NP) that could explain the anomalies. The results often prefer the NP hypothesis over the SM one. The onus is then the improvement of experimental precision, until new physics can be definitively confirmed, or rejected.

This is where the subject of this thesis comes in: the most precise measurement to date of the LFU-sensitive observable R_K. It provides the first evidence for the violation of LFU in rare decays of beauty quarks. Hence, this measurement of R_K represents a substantial leap towards understanding the flavour anomalies, and any NP that could be behind it.

This thesis is divided into three parts. Part I contains two chapters that cover the physics relevant to the measurement of R_K. Chapter 2 describes the theoretical aspects, and Chapter 3 provides an overview of the experimental setup used to record the data. Emphasis is placed on the subsystems that are key to the success of the measurement. Part II is the main component of this thesis, where the experimental procedure used to measure R_K is described. All aspects of the analysis are covered in detail. Finally, Part III contains appendices. Each chapter begins with a brief overview of the covered topics, and a statement on the originality of the work presented therein.

Natural units, in which $\hbar = c = 1$, are used throughout, with the exception of a few figures where the units are explicitly stated. Charge conjugation is implied whenever a particle or a decay is mentioned, unless otherwise specified.

Part I

Lepton flavour universality, in theory and in practice

Chapter 2

The physics of rare beauty-quark decays

This chapter covers the theoretical aspects most relevant to R_K. Section 2.1 provides an overview of the Standard Model, and then Section 2.2 focuses on the processes upon which R_K is built. They are examples of what are known as $b \to s\ell^+\ell^-$ transitions, and experimental results on such processes are reviewed in Section 2.3. Section 2.4 builds upon this context, by looking beyond the Standard Model. Finally, Section 2.5 introduces the observable R_K. This chapter uses Refs. [3–5] throughout.

2.1 Particles and interactions of the Standard Model

In the SM framework, particles are excitations of quantised fields that interact with each other through the electromagnetic, strong, and weak forces. This means that the SM is a Quantum Field Theory (QFT), whose fields are characterised by several quantum numbers. The electrical charges, Q, and the masses associated to the SM fields are listed in Table 2.1. Depending on the intrinsic angular momentum, *i.e.* the spin, the particles in this table can be classified as follows:

- **Spin-1/2** particles: these fermions are the constituents of matter. They come in three generations, each successive generation being characterised by larger masses than the previous one. Depending on the way they interact, fermions are of two types: quarks and leptons. Quarks are bound together in hadrons, most often in triplets or doublets. In the former case, they form baryons, such as the proton and the neutron. In the latter case, they form mesons, an example of which is the pion. Quarks come in six varieties known as flavours. These are: up (u), down (d), charm (c), strange (s), top (t), and beauty (b); the latter is also known as bottom. Similarly, the leptons also come in six varieties. They are grouped into three pairs of a charged and a neutral particle, and each pair is referred to as a lepton flavour. The charged leptons are the electron (e), the muon (μ), and the tau (τ). Each of them has a corresponding neutrino, ν_ℓ, where $\ell \in \{e, \mu, \tau\}$. For every charged fermion particle f, there exists an antiparticle, $\overline{f}$, that has the same quantum numbers, but the sign of the electric charge is flipped. It is unknown whether neutrinos have antiparticles, or whether they are their own antiparticle. In addition, fermions can be either “right-handed” or “left-handed”, depending on whether their spin is parallel or anti-parallel with their momentum. This is known as helicity, and it is important because only left-handed fermions and right-handed anti-fermions interact through the weak force in the SM.

- **Spin-1** particles: these are the vector bosons, whose exchanges between other particles represent the fundamental forces in the SM. The photon (γ) carries the electromagnetic force between charged particles, the gluon (g) mediates the exchange of the strong force between quarks, and the $W^\pm$ and Z^0 bosons carry the weak force between quarks and leptons.

- **Spin-0** particle: the only fundamental boson in the SM that is a scalar, rather than a vector, is the Higgs (H^0). It plays a central role in the SM, because the Higgs couplings to the other particles give rise to their masses [6–8].

Table 2.1: Constituent particles of the Standard Model. The mass of each particle is taken from Ref. [9], and is indicated below its symbol. Table adapted from Ref. [10].

	Fermions			**Bosons**	
generation	1	2	3	vector	scalar
up-type quarks $(Q = +2/3)$	u 2.2 MeV	c 1.3 GeV	t 173 GeV	γ 0	H^0 125 GeV
down-type quarks $(Q = -1/3)$	d 4.7 MeV	s 93 MeV	b 4.2 GeV	g 0	
charged leptons $(Q = -1)$	e 511 keV	μ 106 MeV	τ 1.8 GeV	$W^\pm$ 80.4 GeV	
neutrinos $(Q = 0)$	ν_e <0.8 eV	ν_μ <0.8 eV	ν_τ <0.8 eV	Z^0 91.2 GeV	

The dynamics of the fields present in the SM are described by a mathematical construct known as the Lagrangian. To highlight some of the properties of the SM Lagrangian, it is worth considering the simplified case where only the electromagnetic force is made manifest. Such is the case of interactions between charged leptons, with photons as mediators. The SM reduces to quantum electrodynamics (QED), whose Lagrangian is:

$$\mathcal{L}^{\text{QED}} = \underbrace{-i\bar{e}\gamma^{\mu}\partial_{\mu}e - m_{e}\bar{e}e - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{\text{kinematics}} \underbrace{-g_{e}\bar{e}\gamma^{\mu}eA_{\mu}}_{\text{interactions}} . \tag{2.1}$$

In the above expression, e represents the electron field, m_e is the mass of the electron, and γ^μ are the Dirac matrices. The electromagnetic field tensor is defined as $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, where A_μ represents the photon field. The first two terms on the right-hand side of Equation (2.1) describe the kinematics of the electrons, as per the Dirac equation. Together with the third term, which encodes the kinematics of the photon, they describe the behaviour of free (non-interacting) electrons and photons. The fourth term is added to the Lagrangian to illustrate electromagnetic interactions. These involve the coupling of electrons to photons, and the strength of this coupling is represented by g_e.

The interaction term in the Lagrangian is used to derive observable quantities, such as the rate at which a given process occurs. In general, an observable is the magnitude squared of an amplitude $\mathcal{A}$, determined as:

$$\mathcal{A} = \langle f \,|\, S \,|\, i \rangle \ . \tag{2.2}$$

In the above expression, $|\,i\rangle$ and $\langle f\,|$ denote the initial and final states of the fields. The operator S is known as the S-matrix, and it is a function of the interaction Lagrangian. In order to fully describe the process $i \to f$, S must include the contributions from all possible intermediary states $i \to m_1 \to m_2 \to ... \to f$. The contribution to the S-matrix of each of these processes is known as the matrix element, typically denoted by $\mathcal{M}$. One way to calculate a matrix element is to represent the corresponding process pictorially, in what are known as Feynman diagrams. Three Feynman diagrams that contribute to the annihilation of an electron and a positron into a photon are shown in Figure 2.1. The main features of a Feynman diagram are lines of particles and interaction vertices. By associating each line and vertex with a particular term, and then multiplying the terms together, it is possible to derive the matrix element associated with the Feynman diagram. The S-matrix is then the sum of the matrix elements of all possible Feynman diagrams. The simplest allowed diagram is known as the leading-order contribution. In the particular case of e^+e^- annihilation, it corresponds to the left-most diagram in Figure 2.1, which leads to the interaction term in Equation (2.1). The other two diagrams in Figure 2.1 contain additional vertices, each of which reduces the contribution to the amplitude by a factor of $g_e \simeq 10^{-2}$. Loops of particles are formed between the additional vertices, and so these diagrams are said to contribute at loop-level. By contrast, the left-most diagram contributes at tree-level.

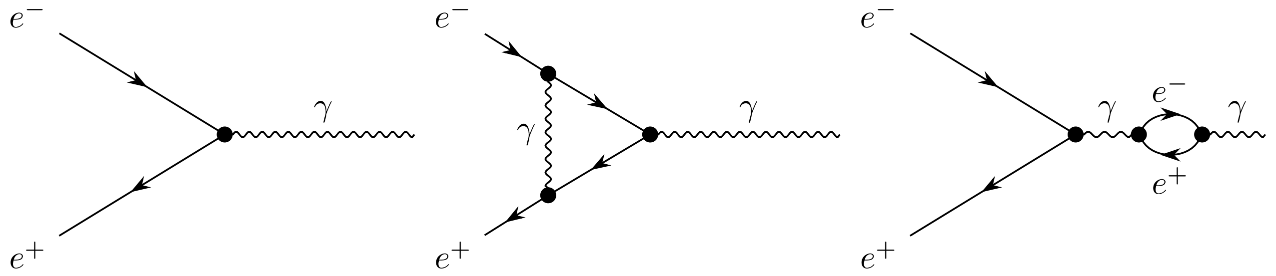

Figure 2.1: Feynman diagrams depicting the production of a photon from electron-positron annihilation, in the presence of a magnetic field (not pictured). The horizontal and vertical axes represent time and space, respectively. Photon (vector) trajectories are shown as wavy lines, whilst electron (fermion) trajectories are represented by straight lines. Antiparticles are interpreted as particles moving backwards in time. The filled dots depict interaction vertices, and are omitted from the other Feynman diagrams presented in this thesis. The leading-order contribution is shown on the left, alongside higher-order diagrams that contain a virtual photon emission (middle) and an e^+e^- loop (right).

2.1.1 Symmetries of the Standard Model

Another approach to making predictions using the SM is through its symmetries. This is by virtue of Noether's theorem, which states that for any differentiable symmetry of a Lagrangian involving conservative forces, there exists a corresponding conserved quantity. One example is the symmetry under translations in space, which leads to conservation of linear momentum. Other examples include transformations that change individual terms of the Lagrangian, without modifying the Lagrangian as a whole. These are known as gauge symmetries, and the QED Lagrangian in particular is invariant under the following pair of simultaneous transformations:

$$e \to e \exp\left[i\theta(x)\right], \qquad A_\mu \to A_\mu - \frac{1}{g_e}\partial_\mu \theta(x)\,. \tag{2.3}$$

The $\theta(x)$ term in these expressions represents a local phase shift of the electron field. Gauge invariance is important because it dictates the form that interaction terms are allowed to take in the Lagrangian. Moreover, physical observables do not depend on the phases of the fields involved, and so gauge invariance ensures predictions are consistent across chosen conventions for the phases.

The gauge transformations that leave a Lagrangian unchanged determine that Lagrangian's symmetry group. In the case of $\mathcal{L}^{\text{QED}}$, the symmetry group is U(1), also known as the circle group. Taking into account other interactions introduces additional symmetries, and therefore changes the symmetry group. For example, QED is unified with the weak interaction at and above energy scales corresponding to the mass of the $W^{\pm}$ boson, $m_W \simeq 80$ GeV. This process is known as electroweak unification [11–13], and it leads to a Lagrangian whose symmetry group is U(1) × SU(2). The former is the circle group characteristic of QED, and the latter is the special unitary group of order 2; it is the symmetry group of the weak interaction. Since the strong interaction is invariant under SU(3) gauge transformations, the symmetry group of the SM is U(1) × SU(2) × SU(3).

Aside from the transformations discussed above, there are symmetries of the SM that are not directly built into the Lagrangian. These are known as accidental symmetries, and they typically emerge through properties of the fields present in the Lagrangian. For example, the fact that neutrinos are massless in the SM leads to a diagonal neutrino mixing matrix. This matrix relates the neutrino mass eigenstates to the flavour eigenstates. A diagonal mixing matrix leads to a lack of terms in the SM Lagrangian that couple between lepton generations, and therefore processes that change the lepton numbers L_e, L_μ, and L_τ are forbidden. Each lepton number is calculated by counting the number of leptons of the corresponding flavour, and subtracting the number of antileptons. This set of conserved quantities forbids the process $\mu^- \to e^+e^-e^-$, because the initial state has $L_e = 0$ and $L_\mu = 1$, whilst the final state has $L_e = 1$ and $L_\mu = 0$. However, the decay $\mu^- \to e^-\overline{\nu}_e\nu_\mu$ is allowed, because $L_e = 0$ and $L_\mu = 1$ in both the initial and final states. Experimentally, the latter has been observed to be the process with the largest rate (nearly 100%), whereas the former's rate is constrained to have an upper limit of 10^{-12} at 90% confidence level [9].

Another accidental symmetry is related to the fact that in the SM Lagrangian, the different lepton flavours couple identically to the vector bosons[1]. This symmetry is known as lepton flavour universality (LFU), and there is no *a priori* motivation behind its presence in the SM. Despite this, it leads to precise predictions on complementary processes, such as $Z^0 \to \mu^+\mu^-$ and $Z^0 \to e^+e^-$, that are identical up to the flavour of the leptons present in the final state. As

---

[1] The couplings to the Higgs are different, since they are proportional to the masses of the different charged leptons. The distinction is therefore made between vector bosons (γ, g, $W^{\pm}$, and Z^0) and gauge bosons (including both the vector bosons and the Higgs).

a result, LFU was tested and found to hold in $W^{\pm}$ and Z^0 decays [14,15], as well as in decays of light mesons [16,17] and of charmonium resonances [18].

2.2 Rare beauty-quark decays

The branch of particle physics where quark flavour plays a key role is known as flavour physics. Among others, it covers the process that is central to the topic of this thesis: the $b \to s\ell^+\ell^-$ transition, whereby a b quark decays into an s quark, with the emission of a charged lepton-antilepton pair. The quarks in the initial and final states have the same electric charge, but different flavours. For this reason, $b \to s\ell^+\ell^-$ transitions are examples of what are known as flavour-changing neutral currents (FCNCs). The fact that electric charge is conserved at every vertex of a Feynman diagram prevents the $W^{\pm}$ bosons from mediating this process at tree level. In addition, tree-level FCNCs cannot proceed via photon or gluon emission, because the three down-type quarks all have the same gauge representations, and therefore have the same gauge interactions. Moreover, the Z^0 boson cannot allow FCNCs at tree level either, by virtue of having universal couplings to up-type and (separately) down-type quarks. This, combined with the fact that the Higgs' Yukawa couplings are aligned with the fermion mass matrices, means that FCNCs cannot proceed at tree-level in the SM.

Given that FCNCs are forbidden in the SM at tree-level, the leading-order SM Feynman diagrams for $b \to s\ell^+\ell^-$ decays involve loops. The examples for $B^+ \to K^+\ell^+\ell^-$, where the beauty and strange quarks are bound in a B^+ and a K^+ meson, respectively, are shown in Figure 2.2. The loops introduce additional electroweak couplings that suppress the rate of $b \to s\ell^+\ell^-$ transitions with respect to related processes that are not FCNCs, such as $b \to c\ell^-\overline{\nu}_\ell$ decays. The FCNC rates are further suppressed by the fact that the up-type quarks in the loop introduce factors proportional to their quark masses in the matrix element. This leads to what is known as the GIM suppression mechanism [19] of FCNCs. In the case of $b \to s\ell^+\ell^-$ transitions, the suppression factors go as:

$$\mathcal{M} \sim \frac{m_i^2}{m_W^2} V_{ib} V_{is}. \tag{2.4}$$

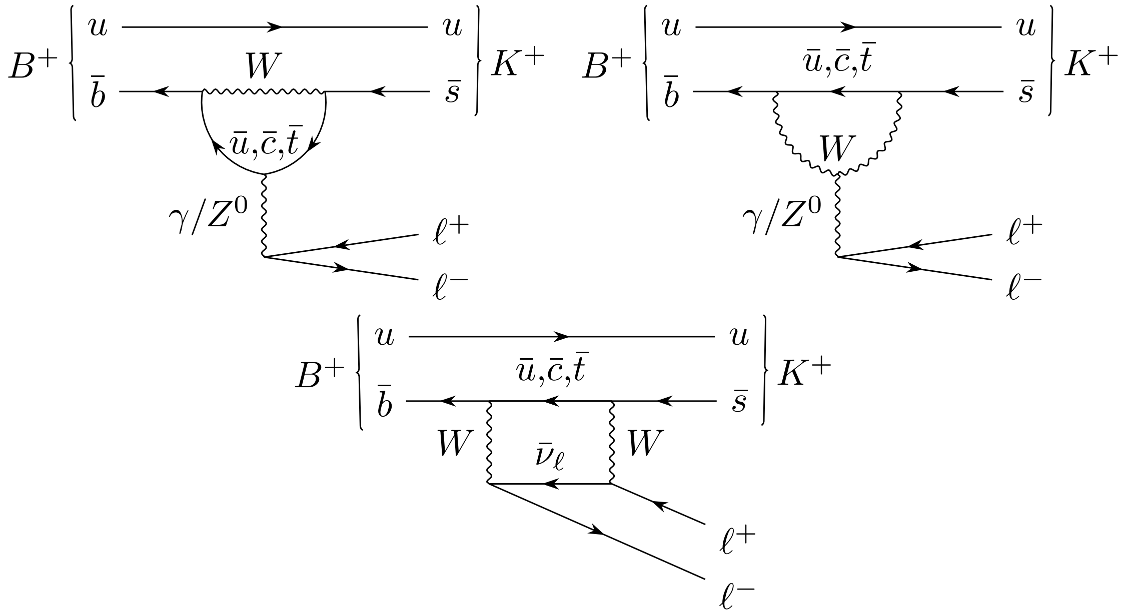

Figure 2.2: Lowest-order Feynman diagrams allowed in the SM for the $B^+ \to K^+\ell^+\ell^-$ decay. Since FCNCs are forbidden at tree-level, these diagrams contain a loop. The diagrams with one W boson in the loop (top) are examples of penguin diagrams, whereas the one with two W bosons (bottom) is an example of a box diagram.

In the above expression, the index i runs over the up-type quarks, and V_{ib} and V_{is} are elements of the quark mixing (CKM) matrix that describes how quarks of different generations mix [20,21]. The top is the only quark that has mass larger than m_W, and the product $V_{ib}V_{is}$ is largest for $i = t$. As a result, the leading contribution to the $b \to s\ell^+\ell^-$ rate comes from loops containing top quarks. This makes $b \to s\ell^+\ell^-$ transitions rare in the SM. In particular for $B^+ \to K^+\ell^+\ell^-$ decays, the expected relative decay rates, also known as branching fractions ($\mathcal{B}$), are $\mathcal{O}(10^{-6})$ [9].

The SM interactions that dictate $B^+ \to K^+\ell^+\ell^-$ decays are of two types. The first corresponds to the electroweak effects that govern the $b \to s\ell^+\ell^-$ transition, which are characterised by energy scales $\sim m_W \simeq 80$ GeV. The second encompasses the strong interactions between the quarks that constitute the B^+ and K^+ hadrons. Such effects have typical energy scales of $\mathcal{O}(10^{-1}$ GeV). Given the separation over three orders of magnitude between short-distance electroweak and long-distance strong effects, the two types of processes can be factorised. Large energy scale effects in $B^+ \to K^+\ell^+\ell^-$ decays can then be described using an effective field theory (EFT) [22]. In such a framework, long-distance (small energy scale) effects are separated from short-distance (large energy scale) ones. This is reflected in the effective

Lagrangian at a given mass scale μ_s [23]:

$$\mathcal{L}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{g_e^2}{4\pi} \sum_i C_i(\mu_s) O_i(\mu_s) \,. \tag{2.5}$$

This expression is analogous to the 4-point interaction of Fermi theory [24]. Fermi's constant is denoted by G_F, g_e is the electromagnetic coupling, and V_{tb} and V_{ts}^* are elements of the quark mixing matrix. Interactions at scales below μ_s are encompasses by the Wilson operators $O_i(\mu_s)$, where the index i runs over a complete basis of operators. Similarly, effects characterised by energy scales above μ_s are encoded by the Wilson coefficients $C_i(\mu_s)$. In the case of $b \to s\ell^+\ell^-$ transitions, the most relevant Wilson operators are:

$$\begin{aligned} O_7 &= \tfrac{m_b}{g_e}\, \bar{s}\sigma^{\mu\nu} b_R \; F_{\mu\nu}\,, & O_7' &= \tfrac{m_b}{g_e}\, \bar{s}\sigma^{\mu\nu} b_L \; F_{\mu\nu}\,, \\ O_9 &= \bar{s}\gamma_\mu b_L \; \bar{\ell}\gamma^\mu \ell\,, & O_9' &= \bar{s}\gamma_\mu b_R \; \bar{\ell}\gamma^\mu \ell\,, \\ O_{10} &= \bar{s}\gamma_\mu b_L \; \bar{\ell}\gamma^\mu\gamma_5 \ell\,, & O_{10}' &= \bar{s}\gamma_\mu b_R \; \bar{\ell}\gamma^\mu\gamma_5 \ell\,. \end{aligned} \tag{2.6}$$

In the above expressions, γ^μ and γ_5 are the Dirac matrices, m_b is the mass of the b quark, and $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The L and R indices denote left and right helicities, respectively. Given that $W^\pm$ bosons only couple to left-handed particles, the coefficients of the primed observables are suppressed by $O(m_s/m_b)$. The other Wilson coefficients have the following SM values at $\mu_s = m_b$ [23]:

$$C_7^{\text{SM}} = -0.3\,, \qquad C_9^{\text{SM}} = +4.2\,, \qquad C_{10}^{\text{SM}} = -4.2\,. \tag{2.7}$$

2.3 Experimental results on beauty-quark decays

Over the past few years, a pattern has been emerging between measurements of b-quark decays [2, 25–64]. They are collectively referred to as the "flavour anomalies", and they manifest themselves as tensions between experimental results and their SM predictions. Particularly for $b \to s\ell^+\ell^-$ transitions, some of these tensions are above $2\,\sigma$. Such levels of departure from the SM are not sufficiently large to rule out statistical fluctuations, and so further studies are needed. The rest of this section presents some of the anomalous observables, grouped according to their type.

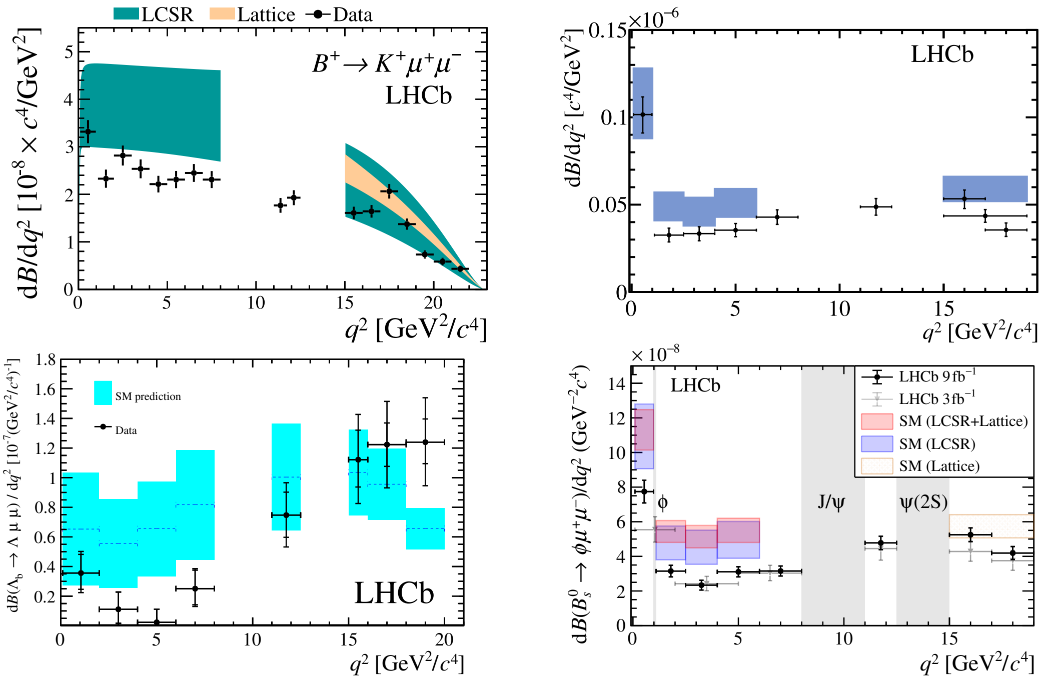

Figure 2.3: Examples of differential branching fraction measurements. Starting from the top left and going clockwise, these are LHCb measurements of: $B^+ \to K^+\mu^+\mu^-$ [26], $B^0 \to K^{*0}\mu^+\mu^-$ [27], $B^0_s \to \phi\mu^+\mu^-$ [28], and $\Lambda^0_b \to \Lambda\mu^+\mu^-$ [29]. The experimental results are depicted in black, whereas the SM predictions are represented by the coloured regions. The gaps in the q^2 spectra are due to selection cuts that veto contributions from the J/ψ and $\psi(2S)$ charmonium resonances.

2.3.1 Differential branching fractions

The SM predicts that decays of beauty hadrons that involve $b \to s\ell^+\ell^-$ transitions are rare, for reasons given in Section 2.2. To test this, the differential branching fractions of several such b-hadron decays have been measured, as a function of the dilepton invariant mass squared (q^2). To date, the experimental precision is driven by the LHCb experiment [65, 66], which is described in Chapter 3. Figure 2.3 shows, as examples, the most precise measurements of the differential branching fractions of four decay modes that involve $b \to s\mu^+\mu^-$ transitions. They are: $B^+ \to K^+\mu^+\mu^-$ [26], $B^0 \to K^{*0}\mu^+\mu^-$ [27], $B^0_s \to \phi\mu^+\mu^-$ [28], and $\Lambda^0_b \to \Lambda\mu^+\mu^-$ [29].

It can be seen that all experimental results at $q^2 < 8$ GeV^2 are consistently below the SM predictions. However, theory and experiment are still compatible to within 2–3 standard deviations, which means that one cannot rule out the SM hypothesis. The experimental results are dominated by systematic uncertainties related to the normalisation channels and theory parameters. Calculating the SM predictions themselves is made complicated by strong-interaction effects that cannot be described using techniques such as perturbation theory, thus leading to what are known as hadronic uncertainties. An example is the internal structure of the b hadron: it is encoded in parameters, known as form factors, that have to be computed non-perturbatively. Another example consists of contributions from Feynman diagrams that are higher-order than the ones depicted in Figure 2.2. Such contributions have to be calculated non-perturbatively if they involve higher-order $b \to c\bar{c}s$ processes, where the $c\bar{c}$ pair forms a loop.

2.3.2 Angular observables

There are $b \to s\ell^+\ell^-$ observables whose SM predictions are more precise than the corresponding differential branching fractions. This includes parameters that describe the distributions of the angles between the particles involved in a $b \to s\ell^+\ell^-$ transition. These are known as angular observables, and they are computed from amplitudes that can be changed by NP contributions in different ways. As a result, angular distributions provide discriminating power between different types of NP. Angular observables have been measured by several experiments in various decay modes [29, 31–42]. One example of an angular observable is P_5', which is designed such that its hadronic uncertainties are small thanks to the form factors cancelling out to first order [67]. The experimental precision on P_5' is driven by LHCb, whose results from $B^0 \to K^{*0}\mu^+\mu^-$ [41] and $B^+ \to K^{*+}\mu^+\mu^-$ [42] data are shown in Figure 2.4. Like with the differential branching fraction results, there are tensions of 2–3 standard deviations between the data and the SM predictions [68–72]. The latter can be systematically biased, because the theory community have not yet reached a consensus on the uncertainties in the predictions caused by contributions from $c\bar{c}$ loops.

2.3.3 Purely leptonic FCNC decays

The observables presented in Sections 2.3.1 and 2.3.2 are difficult to predict in the SM due to the strong interaction. This diminishes the discriminating power between statistical effects and genuine physics beyond the SM. To address this, studies are conducted on FCNC observables for which the theory uncertainties are small. These are known as theoretically-clean observables, two examples of which are the branching fractions of the $B_s^0 \to \mu^+\mu^-$

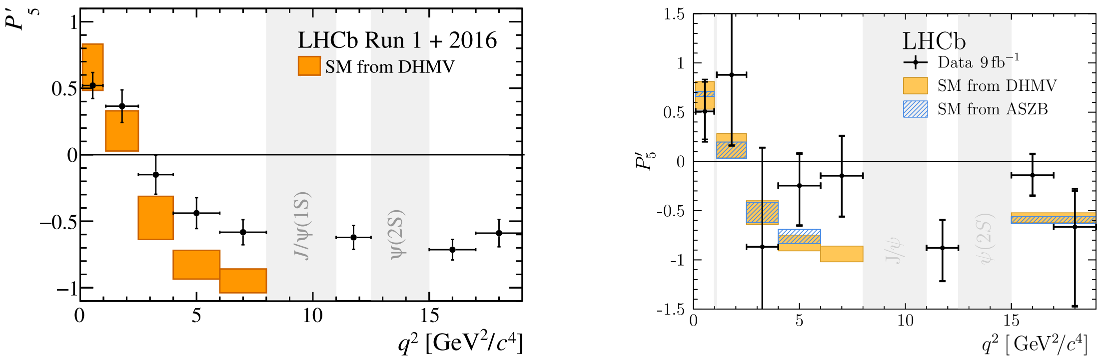

Figure 2.4: Experimental measurements of the angular observable P'_5, as extracted from $B^0 \to K^{*0}\mu^+\mu^-$ [41] (left) and $B^+ \to K^{*+}\mu^+\mu^-$ [42] (right) LHCb data. The SM predictions [68–72] are depicted by the filled boxes. The gaps in the q^2 spectra are due to selection cuts that veto contributions from the J/ψ and $\psi(2S)$ charmonium resonances.

and $B^0 \to \mu^+\mu^-$ decays. The corresponding leading-order SM Feynman diagrams are topologically identical to the ones presented in Figure 2.2, as exemplified on the left-hand side of Figure 2.5. However, the fact that there are only leptons in the final state reduces the theory uncertainty to $\mathcal{O}(1\%)$ [73]. The branching fraction of the B^0_s decay has been measured by the ATLAS, CMS, and LHCb collaborations, who also set upper limits on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ [43–45]. Combining the three sets of measurements leads to an average that is in tension with the SM prediction at the level of 2.1 σ [46], as shown on the right-hand side of Figure 2.5. This is a similar level to the one seen in the differential branching fractions and angular observables. Like with other anomalies, decays of beauty hadrons with muons in the final state are found to have branching fractions below the theoretical predictions. Since the combined result, LHCb has updated their measurement [47, 48], which is in good agreement with both the SM prediction and the result used in the combination with ATLAS and CMS.

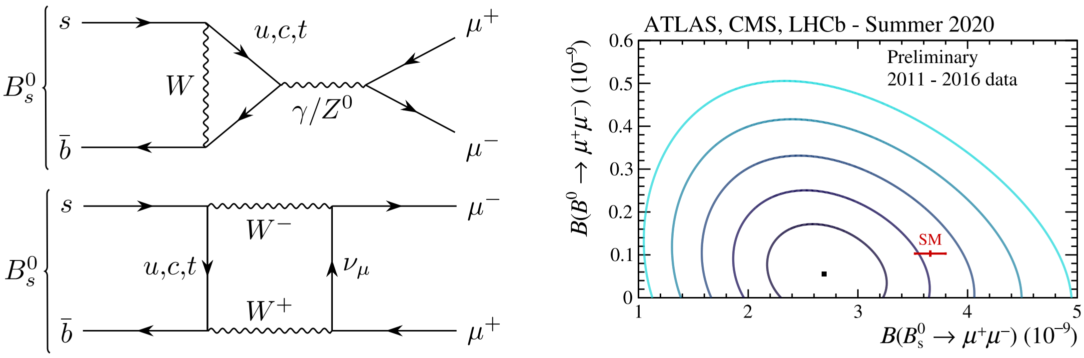

Figure 2.5: (Left) Leading-order Feynman diagrams for $B^0_s \to \mu^+\mu^-$. (Right) Combined average of ATLAS [43], CMS [44], and LHCb [45] results on the branching fractions of $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ [46]. The central value of the average is shown in black, alongside coloured contours depicting 1–5 σ confidence regions. The SM prediction is shown in red.

2.3.4 Ratios of branching fractions

Another way to reduce hadronic uncertainties on FCNCs is to consider the fact that leptons are not affected at leading order by the strong interaction. This means that strong-force effects are expected to be the same in processes that are identical up to the lepton flavour involved. For this reason, theory uncertainties related to form factors and $c\bar{c}$ loops do not impact at leading order the following class of observables:

$$R_H \equiv \frac{\displaystyle\int_{q^2_{\rm min}}^{q^2_{\rm max}} \frac{\mathrm{d}\mathcal{B}(H_b \to H\ell_1^+\ell_1^-)}{\mathrm{d}q^2}\mathrm{d}q^2}{\displaystyle\int_{q^2_{\rm min}}^{q^2_{\rm max}} \frac{\mathrm{d}\mathcal{B}(H_b \to H\ell_2^+\ell_2^-)}{\mathrm{d}q^2}\mathrm{d}q^2} \,. \tag{2.8}$$

In this expression, R_H is a ratio of differential branching fractions, integrated over values of dilepton invariant mass squared $q^2 \in [q^2_{\rm min}, q^2_{\rm max}]$. The particle denoted by H_b can be any hadron with a valence b quark. The H in the final state can be either a particle, such as a K^+, or a system of particles like pK^-. The final states in the numerator and the denominator differ only by the flavour of the leptons $\ell_1^\pm$ and $\ell_2^\pm$. This makes SM predictions accurate, by virtue of the LFU symmetry discussed in Section 2.1.1. Given that electrons and muons have masses that are negligible compared to those of b hadrons, R_H ratios involving electrons and muons are predicted to have values close to 1, with $\mathcal{O}(1\%)$ precision [71,74–82]. Small deviations from unity are expected as a result of effects such as QED corrections, and minute phase space differences arising from the different electron and muon masses. Larger deviations are expected when taus are involved, since their mass is comparable to that of b hadrons.

An example of an R_H ratio is the case where $(H_b, H, \ell_1, \ell_2) = (B^0, K^{*0}, \mu, e)$. This observable, known as $R_{K^{*0}}$, has been measured by LHCb and found to be in tension with the SM [71,74–80] at levels above $2\,\sigma$, as shown on the left-hand side of Figure 2.6 [49]. Ratios where H is a D or a D^*, and ℓ_1 and ℓ_2 are taus and muons, have also been measured [50–58]. They are denoted by R_D and R_{D^*}, and are examples of $b \to c\ell^-\overline{\nu}_\ell$ transitions. Such processes can be mediated at tree-level by $W^\pm$ bosons, which leads to enhanced branching fractions compared to FCNCs. The combination of these results is shown on the right-hand side of Figure 2.6, alongside the SM predictions [83,84]. The average is in tension with the SM at the level of $3.1\,\sigma$ [59].

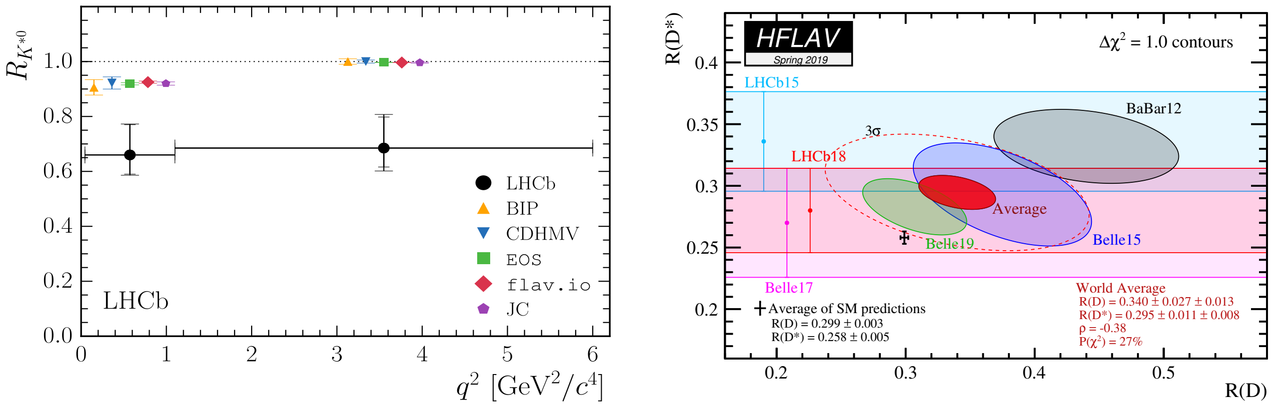

Figure 2.6: Measurements of the LFU-sensitive ratios $R_{K^{*0}}$ [49] (left) and $R_{D^{(*)}}$ [50–58] (right). The $R_{K^{*0}}$ LHCb measurement (black bins) is shown alongside SM predictions [71, 74–80] (coloured bins). The R_D and R_{D^*} results are shown together with their combined average [59] at $1\,\sigma$ and $3\,\sigma$ confidence levels (filled and dotted red, respectively) and the SM predictions [83, 84] (black).

2.4 New Physics in $b \to s\ell^+\ell^-$ transitions

It is shown in the previous section that the SM does not perfectly model experimental results on several b-hadron decays, including observables that are theoretically clean. Even if the discrepancies are caused by statistical fluctuations, there are several other observations that the SM cannot yet describe. It does not incorporate gravity, and it does not provide an explanation for effects attributed to so-called dark matter [85, 86]. Furthermore, the SM does not fully account for the observed imbalance between matter and antimatter [87, 88]. It is therefore clear that the SM is incomplete, and would have to be extended in order to incorporate such phenomena. One way of doing so is to search for new particles and interactions, collectively referred to as New Physics (NP).

Investigating the $b \to s\ell^+\ell^-$ processes described in Section 2.2 represents a potential avenue towards extending the SM, since NP doesn't necessarily exhibit the same suppression mechanisms as the SM. This is illustrated by the Feynman diagrams in Figure 2.7, where the $b \to s\ell^+\ell^-$ process is mediated at tree-level by NP particles. If possible, such transitions would be enhanced with respect to their loop-level SM counterparts. This would lead to deviations of observed quantities away from their SM predictions, such as the ones seen in the case of the flavour anomalies.

---

[2] Baryon number is equal to a third of the difference between the numbers of quarks and antiquarks.

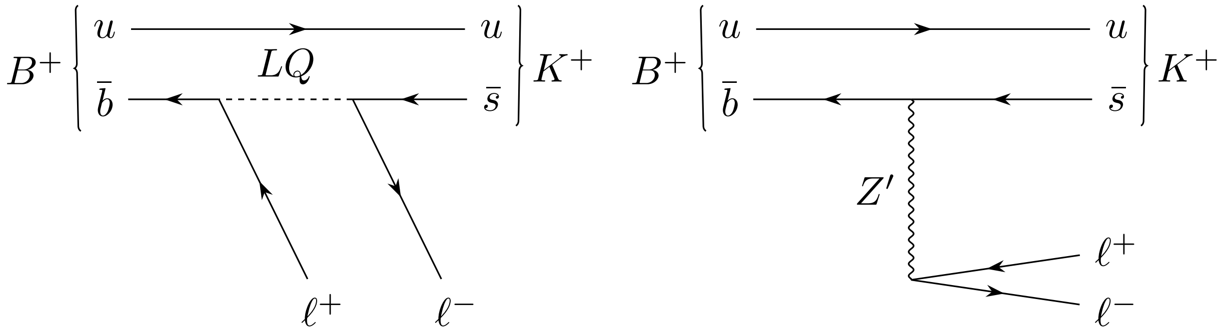

Figure 2.7: NP processes that could allow the $b \to s\ell^+\ell^-$ transition in a $B^+ \to K^+\ell^+\ell^-$ decay to proceed at tree-level. On the left, the decay is mediated by a leptoquark (LQ) that breaks baryon^2 and lepton numbers individually, but conserves their difference. On the right, the mediator is a Z' vector boson with non-universal couplings to the second and third generation of quarks.

In the case of $b \to s\ell^+\ell^-$ and $b \to c\ell^-\overline{\nu}_\ell$ transitions, the mediators are virtual. As such, the anomalies presented in Section 2.3 are examples of indirect searches. Direct searches such as the ones presented in Refs. [89, 90], where NP particles are produced and then decay, have also been conducted. No signal has been observed thus far, indicating that any NP is likely characterised by energy scales above the ones accessible to direct searches; currently, these are $\mathcal{O}(\text{TeV})$. Indirect searches are able to access higher energy scales, since virtual particles can contribute to the loop even if their physical mass is larger than the difference in mass between the final- and initial-state particles. In the EFT context introduced in Section 2.2, this means that at $\mu_s \sim m_b$ NP would manifest itself as a shift of the Wilson coefficients (C_i) away from their SM predictions:

$$C_i = C_i^{\text{SM}} + C_i^{\text{NP}} . \tag{2.9}$$

In the above expression, the Wilson coefficients are written as the sum of their expected SM values — the C_i^{SM} introduced in Equation (2.7) — and NP contributions C_i^{NP}. This allows the definition of two possible scenarios:

- $H_0 \equiv$ "the SM can describe the flavour anomalies": $C_i^{\text{NP}} = 0,\ \forall\, i$; and
- $H_1 \equiv$ "NP is required to explain the flavour anomalies": $\exists\, i$ such that $C_i^{\text{NP}} \neq 0$.

These two hypotheses have been tested based on the observed flavour anomalies, in what are referred to as global fits [72, 91–98]. Generally, this is done using a global likelihood function in the space of Wilson coefficients [99]. After allowing some or all C_i^{NP} of a given complete basis to be nonzero, maximum-likelihood fits to experimental data are performed to find the set of Wilson coefficients that best describes the observations. The significance of H_1 with respect to H_0 is then obtained from the values of the likelihood at the SM and the best-fit points.

Depending on the observables included in the fit, and on the assumptions made about the NP contributions to the Wilson coefficients, global fits generally find that the NP hypothesis, H_1, is favoured over the SM hypothesis, H_0, by more than $5\,\sigma$. Two global fits are given as examples in Figure 2.8 [91]. They use the same data, but make two different assumptions on the nature of possible NP. The fit on the left allows nonzero NP contributions to C_9 and C_{10} in the muon sector. The fit on the right considers cases where the C_9 of all three lepton flavours is universally shifted away from the SM value, and the C_9 and C_{10} of muons are further shifted by contributions that have the same magnitude, but opposite signs. Both NP scenarios are significantly favoured over the SM. It can be seen that different anomalies prefer complementary regions of parameter space, thus leading to preferred NP values that are well constrained.

Global fits performed in an EFT framework, such as the ones shown in Figure 2.8, indicate that the flavour anomalies could be explained coherently by NP that manifests itself through the Wilson coefficients. This motivates the construction of complete theories that can explain the flavour anomalies through new particles and interactions. Some of these new particles could mediate $b \to s\ell^+\ell^-$ transitions at tree-level, as shown in Figure 2.7, and could couple differently to the three quark and lepton generations [100–139]. Different quark-flavour couplings could explain why the anomalies are prevalent in b-quark decays, *i.e.* the heaviest generation. Different lepton-flavour couplings would lead to the violation of LFU, which could explain the anomalies presented in Section 2.3.4. Some NP models, such as those presented in Refs. [140–146], provide explanations for both the flavour anomalies and the recent measurement of the magnetic moment of the muon [147].

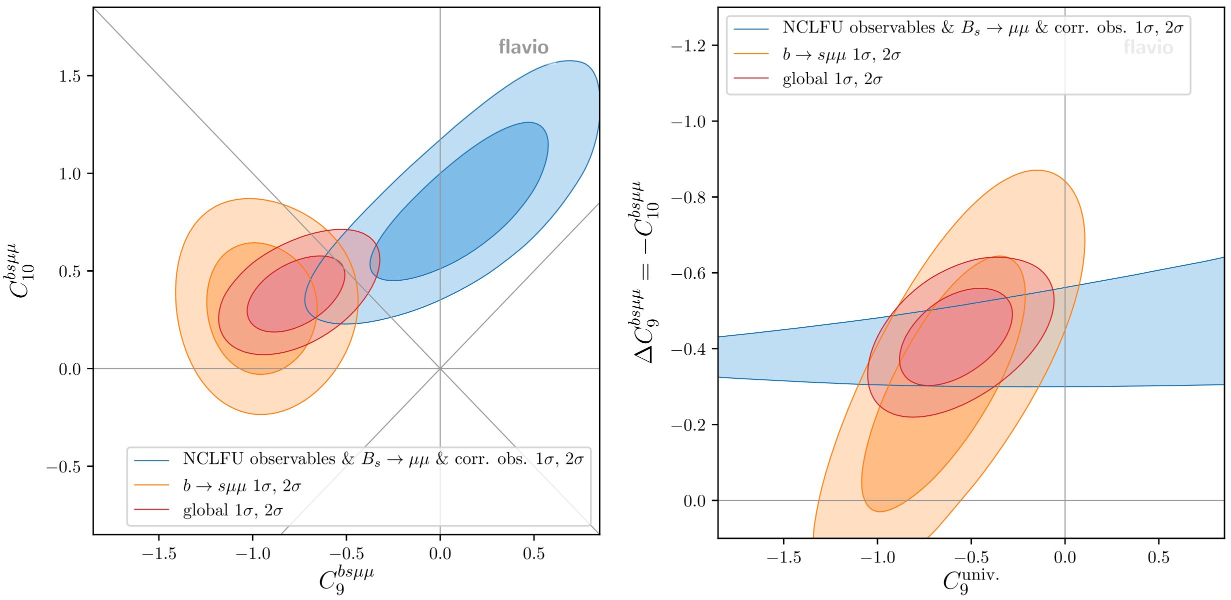

Figure 2.8: Global fits to flavour anomalies [91]. Each plot shows three regions in parameter space that are preferred by the data at $1\,\sigma$ (dark colours) and $2\,\sigma$ (light colours) confidence levels. The orange regions are preferred by results on differential branching fractions [26,27,29,30] and angular observables [31,41,42]. The blue regions are preferred by the measured values of R_K [2] and $R_{K^{*0}}$ [49,148], as well as $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ [46] and correlated observables [36]. The red regions are preferred by all the aforementioned measurements combined. The fits consider NP scenarios that either only affect muons (left), or that affect all lepton flavours universally, on top of a muon-specific contribution (right). The SM lies at the origin of each plot. The fits are performed using the `flavio` software package [79], alongside the global likelihood function provided by `smelli` [99].

2.5 The observable R_K

Given the landscape of possible NP, further studies of the flavour anomalies are required to make a definitive statement on whether the SM can fully describe these processes. Observables of the type defined in Equation (2.8) are particularly important, by virtue of their precise SM predictions. This thesis presents the most precise measurement to date of one such observable:

$$R_K = \frac{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathrm{d}q^2}\mathrm{d}q^2}{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathrm{d}q^2}\mathrm{d}q^2}\,. \tag{2.10}$$

The ratio R_K had been measured previously by the BaBar [60], Belle [61,62], and LHCb [2,63] collaborations. These results are summarised in Figure 2.9, where it can be seen that the

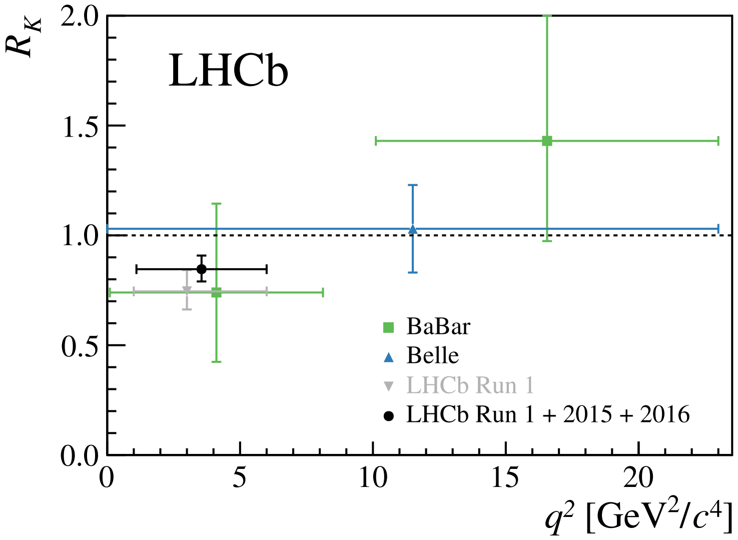

Figure 2.9: Experimental status of R_K prior to the result presented in this thesis. The most precise measurement at the time [2] is shown in black. It was performed by the LHCb collaboration, and it supersedes the previous result [63], which is depicted in grey. Shown in green and blue are the results from the BaBar [60] and Belle [61] collaborations, respectively. The latter has since been updated [62].

experimental precision on R_K is driven by the LHCb measurement[3] that made use of data collected up until the year 2016 [2]. This result is in tension with the SM prediction at the level of $2.5\,\sigma$. The measurement presented in this thesis benefits from an approximate doubling of the available dataset. As such, given that the uncertainty on R_K is dominated by statistics, the expected $\sqrt{2}$ factor gain in precision is crucial for the better understanding of the bigger picture formed by the flavour anomalies.

The following chapter presents the experimental apparatus used for the presently described measurement of R_K. Subsequently, the experimental procedure itself is covered by Part II of this thesis.

---

[3] The q^2 range for this measurement is chosen to be different from the ones employed by other experiments, for reasons discussed in Section 4.3.1.

Chapter 3

The LHCb experiment

This chapter describes the experimental setup used to produce the R_K measurement that represents the topic of this thesis. Section 3.1 describes the Large Hadron Collider, which is used to accelerate and collide protons. These collisions produce B^+ mesons, whose decays are reconstructed by the LHCb detector and used to measure R_K. Section 3.2 describes the processes and techniques employed at LHCb to measure B^+ decays. The flow of information recorded by the LHCb detector is made manageable by means described in Section 3.3, and the techniques used to simulate LHCb data are summarised in Section 3.4.

3.1 The Large Hadron Collider

The currently-largest particle accelerator in the world is the Large Hadron Collider (LHC) [149] at the European Organisation for Nuclear Research (CERN). It is a circular collider 27 km in circumference, situated approximately 100 m below the region surrounding the Franco-Swiss border near Geneva. The majority of LHC's operation time is devoted to accelerating protons to speeds close to the speed of light, and then colliding them in bunches at four points along the circumference of the collider. At each of these interaction points lies a main detector specialised in measuring the particles resulting from the collisions. Two of them, ATLAS [150] and CMS [151], are designed to be general-purpose detectors. The other two, ALICE [152] and LHCb [65,66], specialise in heavy ion collisions and heavy flavour physics, respectively.

The rate at which pairs of beauty hadrons are produced by the proton-proton interactions at the LHC is given by:

$$\frac{\mathrm{d}\mathcal{N}}{\mathrm{d}t} = \mathcal{L} \cdot \sigma_{b\bar{b}}. \tag{3.1}$$

In this expression, $\mathcal{L}$ represents the instantaneous luminosity, which is a measure of how frequently protons collide. When integrated over time, it is referred to as the integrated luminosity, and it corresponds to the amount of data collected in a given period of time. The integrated luminosity recorded by the LHCb detector is presented in Section 3.1.1. The $\sigma_{b\bar{b}}$ term represents the cross-section for the production of pairs of beauty hadrons in the proton-proton collisions. This quantity depends on the centre-of-mass energy at which the protons are collided, $\sqrt{s}$. At scales relevant to the LHC environment, the dependency is found to be approximately linear [153]. Therefore, increasing $\sqrt{s}$ also increases the number of produced beauty hadrons by a similar amount.

Up until the time of writing, there have been two distinct periods devoted to data collection. These are called Run 1 and Run 2. Run 1 started in 2011[1], and involved colliding proton bunches at $\sqrt{s} = 7$ TeV, with a frequency of 20 MHz. The collision energy was raised

---

[1] Some data were recorded in 2010, however they are challenging to calibrate and their statistics are small. They are therefore seldom used in analyses.

to $\sqrt{s} = 8$ TeV during 2012, after which the LHC was turned off for nearly two years. This period, known as Long Shutdown 1 (LS1), saw improvements to the detectors and the accelerator complex, which ultimately allowed proton bunches to be collided at $\sqrt{s} = 13$ TeV, with a frequency of 40 MHz. These conditions persisted throughout Run 2, which began in 2015 and ended in 2018.

3.1.1 LHCb data collection

The LHCb collaboration has collected approximately 9 fb^{-1} of integrated luminosity throughout Run 1 and Run 2, as shown in Figure 3.1. The $b\bar{b}$ cross-section scales approximately linearly with centre-of-mass energy, and so every unit of integrated luminosity from Run 2 contains roughly twice as many b hadrons as the same amount taken during Run 1. To reflect this, four terms are used throughout this thesis to refer to four data-taking periods with similar statistics:

1. "Run 1": data taken during 2011 and 2012;
2. "Run 2.1": data taken during 2015 and 2016;
3. "2017": data taken during 2017;
4. "2018": data taken during 2018.

The latter two years combined are referred to as "Run 2.2". When taken together, Run 1 and Run 2.1 are referred to as "previous data", because they constitute the dataset used in the preceding R_K measurement at LHCb.

3.2 Particle detection at LHCb

The LHCb detector reconstructs an event, such as the decay of a B meson, by extracting as much information as possible from the resulting particles. Different particle properties are best obtained using different technologies, so the LHCb detector consists of several subsystems that complement each other. A schematic of the LHCb detector, showing its

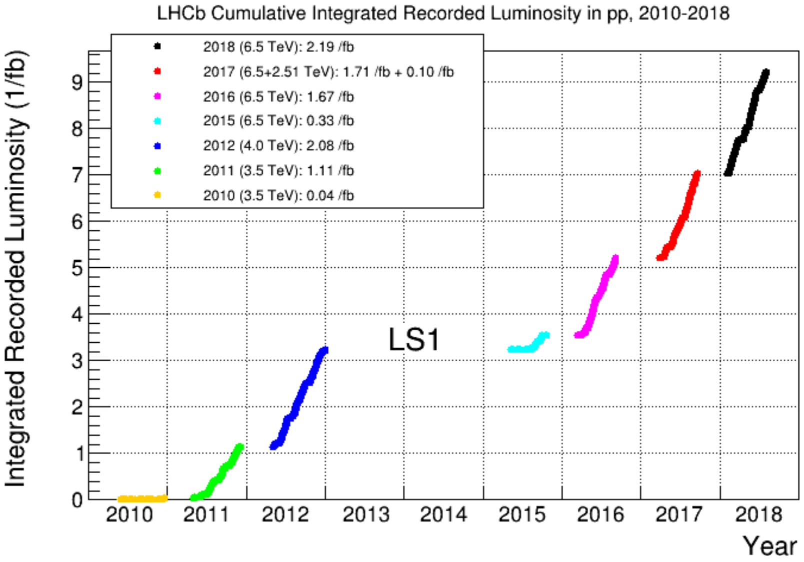

Figure 3.1: Cumulative integrated luminosity recorded by the LHCb experiment from its inception until LS2.

subdetectors, is shown in Figure 3.2. It corresponds to the projection in the (z, y) plane. The z axis follows the direction of the LHC beam line, whereas the y axis points in the vertical direction. The interaction point is located at $z = 0$, and going towards increasing values of z is referred to as the "downstream" direction; the opposite direction is known as "upstream".

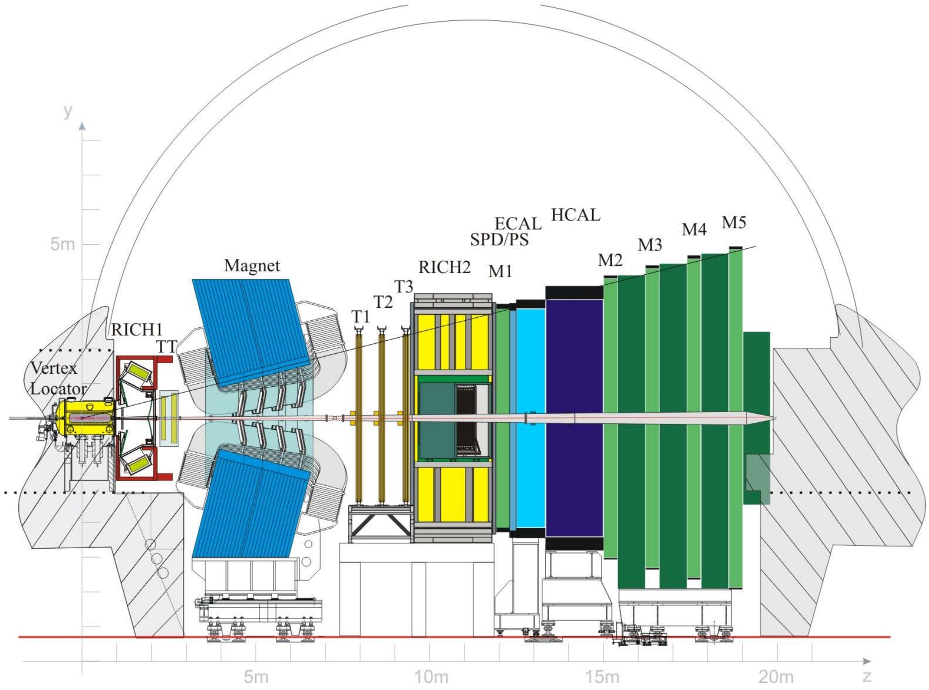

Figure 3.2: Cross-section in the $(z,\ y)$ plane of the LHCb detector. The various labelled subdetectors are described in the main body. Diagram taken from Ref. [65].

In proton-proton collisions, heavy-flavour quarks tend to be produced at low angles with respect to the z axis [154]. Since heavy flavour physics represents the core of the LHCb physics programme, the LHCb detector is designed to cover the very-forward longitudinal angle region $\theta \in (10\text{ mrad}, 250\text{ mrad})$. This corresponds to the pseudorapidity range $\eta \in (2, 5)$, where pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$.

3.2.1 Energy loss mechanisms

There are several ways in which incoming particles interact with the LHCb detector and lose energy. The amount of lost energy depends on properties of both the detector material (*e.g.* the atomic number), and the incoming particle (*e.g.* its speed). The rest of this subsection discusses the energy loss mechanisms that dominate in the LHCb environment.

The quantity under discussion is the mass stopping power, defined as the average energy loss per distance travelled through a medium. The mass stopping power as a function of $\beta\gamma$ is shown in Figure 3.3, alongside the kinematic regions that correspond to typical muons and electrons at LHCb. In the case of charged particles with $\beta\gamma$ approximately between 10^{-1} and 10^3, the dominant process for energy loss is the interaction with the electrons of the medium. Incoming particles pass energy onto these electrons, which then either excite or ionise the atoms. In this kinematic regime, the Bethe formula can be used to calculate the mass stopping power with $\mathcal{O}(\%)$ precision [9]. The resulting spectrum of mass stopping power is characterised by a sharp drop at $\beta\gamma < 1$, followed by a wide minimum, and a slow logarithmic rise. Particles around the minimum are known as minimum ionising particles, and are typically capable of traversing considerable amounts of material before being stopped.

At higher values of $\beta\gamma$, another mechanism leads to average energy losses larger than those caused by ionisation and excitation. The underlying process for this mechanism is the radiation of photons by the traversing particle, in the presence of the electromagnetic potential of the nuclei[2] in the medium. The resulting photons are highly collinear with the passing particle, and are known as bremsstrahlung radiation. The dominant effect that

---

[2] The potential of the electrons may also cause radiative losses, but to a lesser extent.

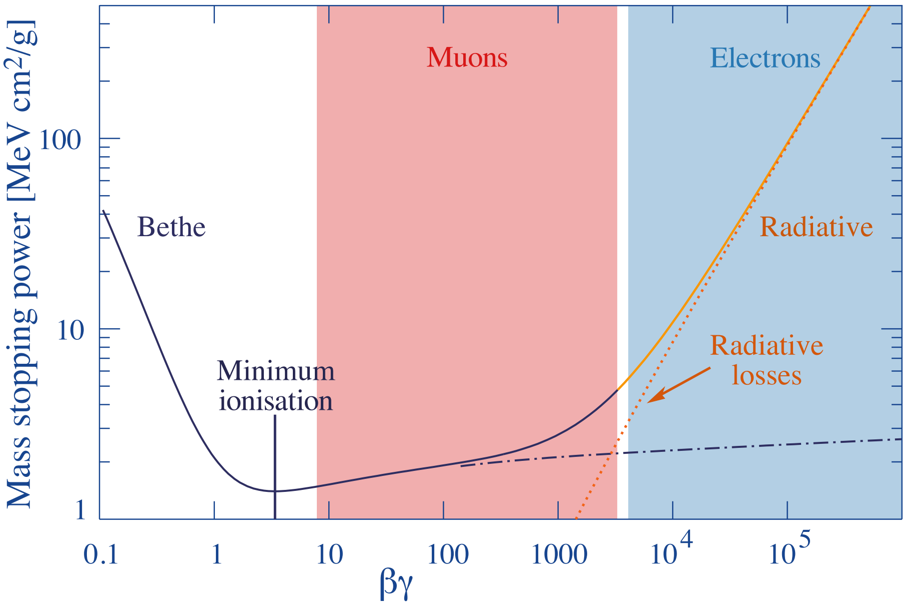

Figure 3.3: Mass stopping power (average energy loss) in copper, as a function of the $\beta\gamma$ of the incident particle. Individual contributions from ionisation and radiative losses are shown by the dashed purple and dotted orange lines, respectively. The solid line shows the total mass stopping power, coloured to reflect the dominant contribution. Regions corresponding to muons and electrons in the LHCb environment are shown in red and blue, respectively. Figure adapted from Ref. [9].

causes these photons to lose energy is the production of e^+e^- pairs, as a result of interactions with the atoms in the traversed material. The resulting electrons and positrons produce further bremsstrahlung radiation, which leads to additional e^+e^- pairs and so on, until the kinetic energy of the initial particle becomes low enough that other energy-loss mechanisms become dominant. This phenomenon is known as an electromagnetic shower.

In particular for the LHCb environment, it can be seen that muons are on the logarithmic rise of the Bethe function, and therefore have small average energy losses. By contrast, electrons lose large amounts of energy through bremsstrahlung radiation. This leads to significant differences between the detection and reconstruction strategies of electrons and muons.

Hadrons such as pions and kaons are more massive than muons, which means they also lose small amounts of energy through ionisation. However, they also interact through the strong force with the detector material. As a result, hadrons produce showers that are qualitatively similar to the electromagnetic showers produced by photons and electrons. They differ through their composition (since they contain some hadrons such as pions) and through the amount of material needed to initiate and contain such showers.

3.2.2 Measuring particle energies

Given the differences between the way photons, electrons, muons, and hadrons interact with matter, the LHCb detector uses different methods to reconstruct the energies of incoming particles. In the case of electrons and photons, the energy is determined using the electromagnetic showers produced in the ECAL [155]. This subdetector, depicted by the light blue rectangle in Figure 3.2, is able to measure shower energies (E) with a resolution of $1\% \oplus 10\%/\sqrt{E}$ (E in GeV) [65]. It consists of layers of active (detection) material, interleaved with passive (absorption) layers. The latter are made of lead, whose density and high atomic number facilitate the development of electromagnetic showers. The former are scintillation plates, meaning that light is emitted when particles pass through the material. The light is sent through wavelength-shifting fibres to photomultiplier tubes that generate an electrical signal proportional to the amount of scintillation light, notwithstanding thresholds and saturation levels. The energy contained in the electromagnetic shower that caused the scintillation is therefore inferred based on the signal from the photomultiplier tubes.

The ECAL is aided by two pads of scintillator that are placed in front of it. These are called the preshower (PS), and the scintillating pad detector (SPD). Their position relative to the ECAL is depicted in Figure 3.4, alongside the locations of the showers produced by different types of particles. Only charged particles are expected to initiate showers in the SPD, and therefore this detector is useful in separating neutral and charged incoming particles. The PS facilitates the formation of electromagnetic showers, thus making electrons and photons easier to stop by the ECAL.

To reflect the different numbers of incoming particles at different angles, the scintillators (referred to as cells) that form the ECAL, PS, and SPD active layers have different dimensions in the (x, y) plane. As shown on the left-hand side of Figure 3.5, there are three different regions. The outer region is farthest from the beam line, where the flux of incoming particles is smallest. As a result, this is where the cells have the largest cross-sectional area (12.12×12.12 cm^2). The segmentation is finer in the middle region, where the cells are 6.06×6.06 cm^2 wide. Closest to the beam line, in the inner region, the cells have an area of 4.04×4.04 cm^2.

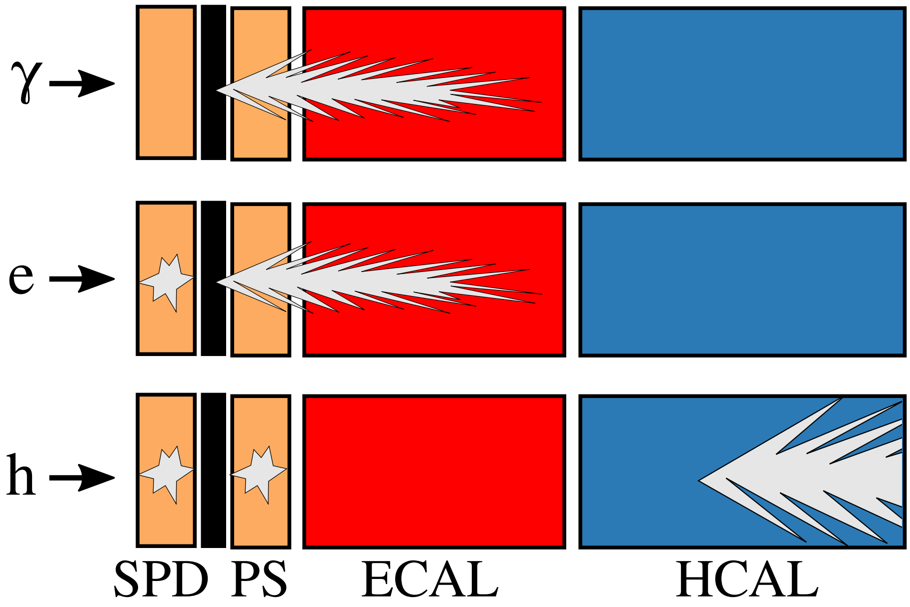

Figure 3.4: Layout of the LHCb calorimeter system in the (z, y) plane. Showers initiated by incoming photons, electrons, and hadrons are shown in white. The coloured rectangles depict the subdetectors that make up the calorimeter system, whilst the black rectangle represents the lead absorption layer placed between the SPD and the PS. Diagram taken from Ref. [10].

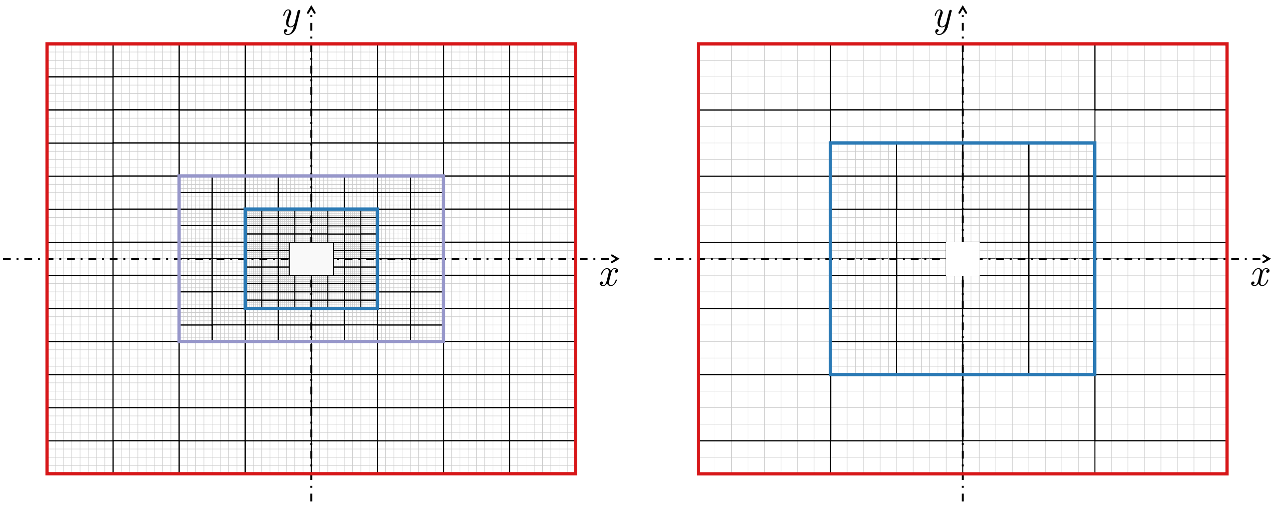

Figure 3.5: Segmentation of the detectors that form the calorimeter system. Shown on the left is the configuration of the ECAL, SPD, and PS, and shown on the right is the layout of the HCAL. The inner, middle, and outer regions are separated by the blue, purple, and red rectangles, respectively. Diagrams adapted from Ref. [156].

Given that hadrons are more difficult to stop than photons and electrons, the majority of them are not expected to initiate showers in the ECAL. Instead, hadronic showers are mostly contained by the HCAL. This subdetector, depicted by the dark blue rectangle in Figure 3.2, has an energy resolution of $(69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$ [65], where E is the shower energy in GeV. It is placed downstream of the ECAL, and employs the same strategy of interleaving layers of detection and absorption material. Like the ECAL, the HCAL also consists of regions with different scintillator sizes, as shown on the right-hand side of Figure 3.5. The difference lies in the number of regions, and the areas of their corresponding cells. The inner region contains cells 13.13×13.13 cm^2 wide, whereas the cells in the outer region have an area of 26.26×26.26 cm^2.

3.2.3 Muon detection

Five detection stations interleaved with iron absorbers are placed at the downstream end of the LHCb detector. They are called the muon stations [157, 158], and are depicted by the green rectangles in Figure 3.2. Most particles that are not muons are expected to stop before they reach the muon stations, thus making this subsystem particularly useful in discriminating muons from other particles. The first station, M1, is placed upstream of the SPD, in order to improve measurements of muon transverse momentum (p_{T}). The other four stations, M2–M5, are situated downstream of the HCAL. Each station consists of 276 chambers that detect incoming charged particles. As with the ECAL and HCAL, the granularity is finer close to the beam line, where most incoming particles are expected to traverse the detector. The 12 inner-most chambers of M1, where radiation damage is expected to be largest, are gas electron multiplier detectors [159]. The other chambers are multi-wire proportional chambers [160]. Both types of chambers are examples of gaseous ionisation detectors, whereby charged particles are detected based on the ionisation they produce as they traverse gas-filled chambers.

3.2.4 Tracking

Aside from estimating the energy of incoming charged particles, ionisation is also used to measure particle momenta. When passing through a magnetic field, a charged particle is bent in a helix whose radius is proportional to the particle's momentum. Therefore, determining the trajectory of a charged particle in a magnetic field leads to a measurement of its momentum. This is called tracking, and the subdetectors that perform this at LHCb employ several techniques to locate the ionisation caused by incoming charged particles. The technologies are chosen based on factors such as required momentum resolution and radiation hardness, however they all rely on ionisation.

There are four tracking stations at LHCb. Three of them, denoted by T1, T2, and T3 in Figure 3.2, are placed downstream of the magnet. They are instrumented differently depending on the proximity to the beam line, to reflect the larger number of tracks at small angles. The large-angles region is known as the Outer Tracker [161, 162], and it consists of straw tube detectors. The region close to the beam line is referred to as the Inner Tracker [163], and it is instrumented with silicon microstrip sensors. The other tracking station is called the Tracker Turicensis (TT in Figure 3.2) [164]. It is placed upstream of the magnet, and like the inner tracker it features layers of silicon microstrip detectors. The overall momentum resolution ranges from 0.5% at low momentum to 1.0% at 200 GeV [66].

The magnet [165] used to bend charged particles is a warm dipole that produces a predominantly vertical magnetic field of up to 1.1 T. The integrated magnetic field along the path of a traversing particle is approximately 4 Tm, leading to an average p_{T} kick of approximately 1.2 GeV. During data taking, the direction of the magnetic field, also known as the polarity, changes periodically between the positive and the negative y-axis direction. The two configurations are referred to as "MagUp" and "MagDown", respectively. Changing the magnet polarity reduces systematic effects induced by positively-charged particles being bent in opposite directions from negatively-charged particles.

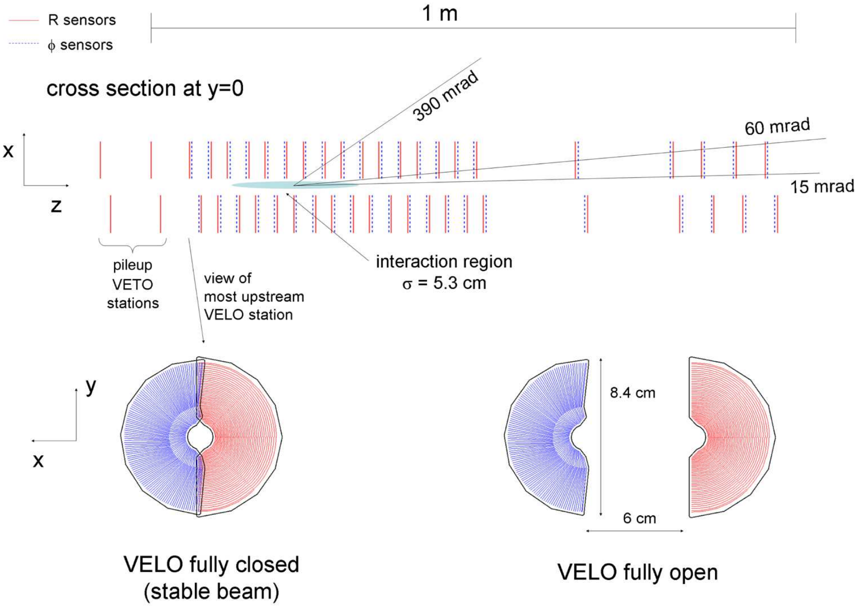

Figure 3.6: Cross-sections of the VELO. Shown at the top is the projection in the (z, x) plane, at $y = 0$, during stable beam conditions. The front of the first module when the two halves are fully closed and fully opened is shown on the bottom left and bottom right, respectively. Diagrams taken from Ref. [65].

3.2.5 Vertexing

Tracking is complemented by a subdetector, known as the Vertex Locator (VELO) [166, 167], designed specifically to determine the points at which particles produced at the interaction point decay. This is important for the LHCb physics programme, because beauty and charm hadrons travel a short distance before decaying, and can therefore be isolated from shorter-lived particles. In particular, the VELO allows LHCb to measure decay times with a resolution of approximately 45 fs for B_s^0 decays [66].

A schematic of the VELO is shown in Figure 3.6. Out of all LHCb subdetectors, it is located closest to the interaction point. It contains 42 modules placed evenly on either side of the beam line, along the z axis. Each module features two semi-circular sensors that have silicon strips arranged along different trajectories. One of the sensors has radial strips, whilst the other has semicircular strips. Therefore, the former sensor measures the polar angle, and the latter determines the radial distance. During data taking, the modules are positioned approximately 8 mm away from the beam line. To prevent damage from potentially unstable beams, this distance is increased to around 3 cm when LHCb is not taking data.

3.2.6 Particle identification

As discussed in Section 3.2.1, the calorimeters are able to distinguish between photons, electrons, and hadrons, based on the locations of the particle showers they initiate. Similarly, muons can be identified from the fact that they tend to be the only particles that are not stopped before the muon stations. To further facilitate the identification of particle species, LHCb also uses two ring-imaging Cherenkov (RICH) detectors [168, 169].

When a charged particle passes through a medium at a speed faster than light would have in said medium, photons are emitted. This is called Cherenkov light, and the angle it forms with respect to the particle's trajectory (θ_c) depends on the speed of the particle (β) and the refractive index (n) of the medium:

$$\theta_c = \arccos(1/n\beta)\,. \tag{3.2}$$

This means that the speed of an incoming particle can be determined by measuring the Cherenkov angle θ_c. When used in conjunction with the momentum measured by the tracking stations, this leads to an estimate of the mass:

$$\beta^2 = \left(\frac{1}{n\cos\theta_c}\right)^2 = \frac{p^2}{p^2+m^2} \tag{3.3}$$

$$\Rightarrow m = p\sqrt{(n\cos\theta_c)^2 - 1}\,. \tag{3.4}$$

In the RICH detectors at LHCb, Cherenkov light is focused by an optical system onto an array of hybrid photon detectors. The optical system contains spherical mirrors that focus the Cherenkov photons emitted by a given particle into a circle, called a Cherenkov ring. The radius of a Cherenkov ring depends on the angle θ_c. The expected pattern under each mass hypothesis is compared to the measured photons, and a likelihood is calculated. Particle type is thus inferred using the likelihood ratio with respect to the pion mass hypothesis.

In order for a particle to produce Cherenkov radiation, it must have a minimum speed: $\beta_{\text{min}} = 1/n$. The larger the refractive index, the lower the momentum threshold for Cherenkov radiation. However, Equation (3.2) shows that larger refractive indices lead to larger

Cherenkov angles, which require bigger detectors to reconstruct. To address this, two RICH stations are used at LHCb, RICH1 and RICH2. They contain different gases, and so they are optimised for complementary momentum ranges. The station closest to the interaction point, RICH1, uses C_4F_{10} gas[3], which has a higher refractive index than the CF_4 gas used by the second station, RICH2. The proximity to the interaction point also allows RICH1 to cover a larger acceptance than RICH2. As a result, RICH1 is optimal for low-momentum particles, whereas RICH2 performs best on high-momentum particles. The characteristics of each RICH station are summarised in Table 3.1, and the particle identification (PID) performance is illustrated in Figure 3.7. It can be seen that there are regions in the $(p,\ \theta_c)$ plane that are populated by particles of certain species.

Table 3.1: Characteristics of the two RICH stations.

Station	Refractive index	Optimal momentum range [GeV]	Horizontal acceptance [mrad]	Vertical acceptance [mrad]
RICH1	$1 + 1.4 \times 10^{-3}$	[10, 60]	[±25, ±300]	[±25, ±250]
RICH2	$1 + 4.8 \times 10^{-4}$	[15, 100]	[±15, ±120]	[±15, ±100]

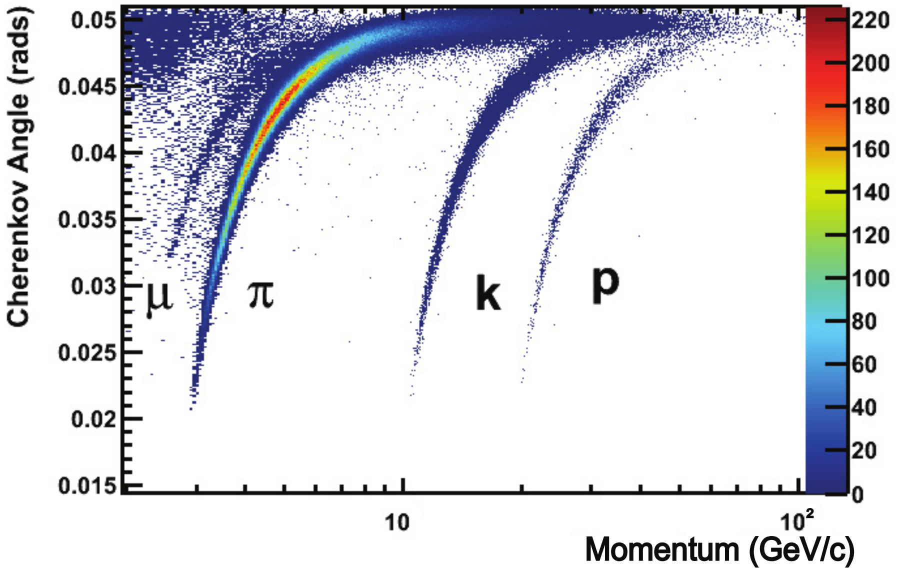

Figure 3.7: Angles at which Cherenkov light is produced in RICH1 by particles with various momenta. The annotations indicate regions in the plane that are mostly populated by particles of the corresponding species. Plot taken from Ref. [169].

---

[3] Aerogel was also used in the RICH1 during Run 1. It was removed to facilitate ring reconstruction.

3.3 The flow of data at LHCb

Given that a full LHCb event has a size of around 100 kB, storing the information recorded by the detector is made challenging by the 40 MHz rate at which particle bunches are collided by the LHC. To address this, data are required to pass several filtering stages before they are stored and made available for physics analyses.

The first filtering layer is the trigger [170], which consists of two stages. The first one is hardware-based, and is called the L0 trigger. Its role is to reduce the flow of data to a rate that allows the entire detector to be read out: from 40 MHz to 1 MHz. This imposes a upper limit of 25 ns on the time window of each subdetector used by the L0 trigger. Hits in the muon chambers are found to have a time resolution between 2.5 ns and 4.0 ns [171]. In the calorimeters, the detected signal pulses are generally longer than the nominal read-out window of. To take this into account, the signal is clipped to fit within 25 ns.

Events that pass the L0 trigger are reconstructed and further analysed by the HLT, which is the next step in the trigger selection. It is software-based, and reduces the rate to 5 kHz in Run 1, and 12.5 kHz in Run 2. This allows the events to be fully reconstructed offline and stored. To ease the burden on computing resources, an additional filtering stage is executed before the data is made available to analysts. It is called the stripping, and it consists of loose cuts that improve the quality of selected candidates. By the end of 2020, LHCb data and simulation amounted to 74.6 PB of tape storage, and 35.4 PB of disk storage [172]. The former covers raw data, whilst the other contains simulated samples and processed data.

An additional data stream was introduced in Run 2 to allow more events to be stored. It is called Turbo, and it involves directly saving the candidates reconstructed by the HLT to disk. This means that Turbo events do not undergo full reconstruction, however the HLT reconstruction during Run 2 is identical to the offline reconstruction. Since the offline reconstruction stage is skipped, data can be collected at a higher rate, thus enabling the Turbo stream to collect more events. However, only candidates reconstructed by the trigger can be recorded, and therefore Turbo data cannot be used in cases where information about the underlying event is necessary to the analysis.

3.4 LHCb simulation

Simulated events undergo the same filtering stages described in Section 3.3 for the data. Proton-proton collisions are generated using PYTHIA [173, 174] configured specifically for the LHCb environment [175]. The resulting hadronic particles, as well as their decays, are simulated by EVTGEN [176]. Furthermore, PHOTOS [177] is used to account for final-state radiation. Finally, the interactions between the simulated particles and the LHCb detector are modelled by the GEANT4 [178] toolkit.

For analysis purposes, the distinction is made between two types of simulation samples. When the final-state particles are not propagated through the detector and reconstructed, the samples are referred to as generation-level simulation. This allows the study of detector effects such as geometrical acceptance and reconstruction. If, instead, the final-state particles are treated in the same way as data, the result is reconstruction-level simulation.

Part II

Measuring R_K

Chapter 4

Measurement strategy

Following the theoretical overview of R_K and the description of the LHCb detector in Part I, the experimental procedure for measuring R_K at LHCb is presented in Part II. It opens with this chapter, which summarises the practical aspects that drive the measurement strategy. Section 4.1 describes the final states relevant to the analysis. Then, Section 4.2 explains how the definition of R_K from Equation (2.10) is adapted to address certain experimental challenges. Finally, Section 4.3 provides an overview of the selection employed to collect the data used in the R_K measurement.

The measurement strategy closely follows the one designed and implemented by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous LHCb R_K measurement [2]. Results that constitute original work are highlighted where appropriate.

4.1 $K^+\ell^+\ell^-$ final states

The general principles employed by the LHCb experiment to process the data it records are presented in Section 3.3. This section starts from this foundation and concentrates on the specifics of the R_K analysis.

Consider Equation (2.10), which expresses R_K as the ratio between the branching fractions of two B^+ decays. The daughters from both decay types can be fully reconstructed by the LHCb detector. This has two major implications: first, the R_K analysis strategy revolves around selecting data where the final state contains a charged kaon and either a pair of oppositely-charged muons, or a pair of oppositely-charged electrons. Second, for a given signal candidate, the invariant mass of the $K^+\ell^+\ell^-$ system, denoted by $m(K^+\ell^+\ell^-)$, is expected to be approximately equal to the mass of the B^+ meson, m_B. This invariant mass is given by:

$$m(K^+\ell^+\ell^-)^2 = \left\|p_K + p_{\ell^+} + p_{\ell^-}\right\|^2 = (E_K + E_{\ell^+} + E_{\ell^-})^2 - \left\|\vec{p}_K + \vec{p}_{\ell^+} + \vec{p}_{\ell^-}\right\|^2 , \tag{4.1}$$

where p_X is the 4-momentum of particle X, E_X is its energy, and $\vec{p}_X$ is its 3-momentum. Signal candidates have $m(K^+\ell^+\ell^-)$ values close to m_B, so the spectrum of $m(K^+\ell^+\ell^-)$ in data contains an accumulation around $m(K^+\ell^+\ell^-) = m_B$ that corresponds to signal events. This accumulation, known as a peak, makes $m(K^+\ell^+\ell^-)$ crucial in separating signal from background.

The extent to which $m(K^+\ell^+\ell^-)$ provides separation power between signal and background can be better appreciated by also taking into account the square of the invariant mass of the two leptons, q^2. This is portrayed in Figure 4.1, where the $m(K^+\ell^+\ell^-)$ and q^2 distributions of partially-selected data candidates reveal certain features:

- a vertical band centred around $m(K^+\ell^+\ell^-) = m_B \simeq 5.28$ GeV that stretches across q^2. This is the signal mode, $B^+ \to K^+\ell^+\ell^-$, which has a branching fraction of $\mathcal{O}(10^{-6})$ [9];
- an accumulation at $m(K^+\ell^+\ell^-) \simeq 5.28$ GeV and $q^2 = m^2_{J/\psi} \simeq 9.6$ GeV^2. This corresponds to the tree-level decay $B^+ \to K^+J/\psi$, followed by the decay of the J/ψ into two oppositely-charged leptons. The total branching fraction for this process is $\mathcal{O}(10^{-4})$ [9],

making $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ events more abundant than thei $B^+ \to K^+\ell^+\ell^-$ counterparts by a factor of $\mathcal{O}(100)$;

- another accumulation at $m(K^+\ell^+\ell^-) \simeq 5.28$ GeV, but at $q^2 = m^2_{\psi(2S)} \simeq 13.6$ GeV^2 instead. This is another tree-level decay, $B^+ \to K^+\psi(2S)$, with the subsequent decay of the $\psi(2S)$ into two leptons;

- horizontal bands that have the same q^2 as the two resonances described above, but with different $m(K^+\ell^+\ell^-)$. The regions where $m(K^+\ell^+\ell^-) > 5.28$ GeV are dominated by combinations between a random J/ψ or $\psi(2S)$ resonance, and a random kaon in the event. These are called combinatorial events. By contrast, the regions where $m(K^+\ell^+\ell^-) < 5.28$ GeV are mostly populated by events where a beauty hadron has decayed, but not all resulting particles were reconstructed. These are known as partially-reconstructed events, and examples include $B^0 \to K^{*0} J/\psi(\ell^+\ell^-)$ processes where the K^{*0} decays into a kaon and a pion, and the latter escapes reconstruction; and

- diagonal bands that extend down to lower $m(K^+\ell^+\ell^-)$ and q^2 from the two resonances. Here, the B^+ and/or one of its daughters emits one or several photons through bremsstrahlung. These photons take some energy away from the final-state particles, hence the lower $m(K^+\ell^+\ell^-)$ and q^2 values. Events where this happens are called radiative events, and the diagonals themselves are referred to as radiative tails.

The features enumerated above are all noticeably blurred in the electron channel, compared to their muon counterparts. The loss in resolution is related to the fundamental differences between the ways muons and electrons interact with the LHCb detector. As explained in Section 3.2.1, muons lose little to no energy — $\mathcal{O}(1\text{ MeVcm}^2/\text{g})$ — as they traverse the detector. By contrast, electrons are expected to lose considerable fractions of their energy due to bremsstrahlung radiation. The differences between electron and muon detection are what drive most of the decisions taken by the R_K measurement strategy.

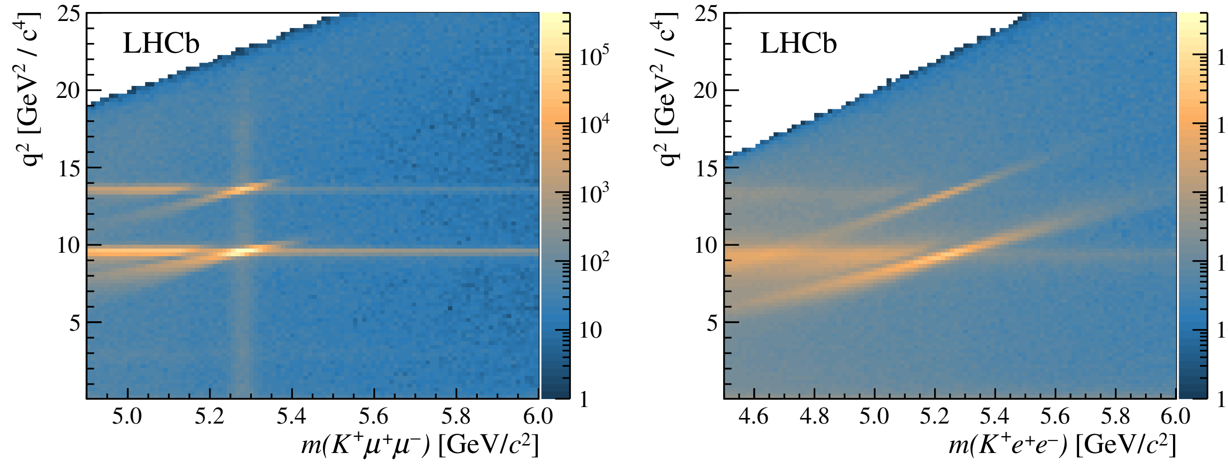

Figure 4.1: Distributions in $m(K^+\ell^+\ell^-)$ and q^2 of events used in the R_K analysis. These are muon (left) and electron (right) signal candidates from all data-taking periods, to which a partial selection has been applied in order to highlight the features discussed in Section 4.1. This partial selection is formed by the requirements listed in Table 4.2. The vertical band corresponding to the $B^+ \to K^+e^+e^-$ signal is not visible, for reasons discussed in the main body.

Electron energy loss is mitigated by the bremsstrahlung recovery process [179], whereby photon clusters in the ECAL have their energies added to electrons whose trajectories before the magnet match the locations of the clusters. Tracks are most affected by bremsstrahlung radiation emitted before the magnet, *i.e.* before the curvature is measured by the tracking stations. Bremsstrahlung radiation is approximately collinear to the electron track, making energy losses negligible in the case of radiation once the track no longer changes direction. For this reason, the algorithm focuses on the recovery of bremsstrahlung photons emitted upstream of the magnet. To do so, the tangent to a given electron track is extrapolated to the ECAL (x, y) plane, based on the origin vertex of the track and its intersection with the TT. A 2σ confidence area is then calculated, based on the precision of the extrapolation, and the uncertainty on the position of a given photon cluster. Clusters with centres inside this area that have $p_{\mathrm{T}} > 75$ MeV and satisfy loose photon identification requirements are considered to have come from bremsstrahlung radiation emitted by the associated electron track. The 4-momentum of the photon is calculated based on the cluster energy (assuming the photon originates from the primary vertex and that its direction points to the barycentre of the shower) and added to the 4-momentum of the electron.

Even after the bremsstrahlung recovery process, the resolution of electrons is not on par with that of muons. The following two sections describe how the R_K analysis strategy is designed to mitigate the impact of differences between electrons and muons. Section 4.2 presents a method that substantially diminishes systematic uncertainties induced by differences between electrons and muons. Then, Section 4.3 explains the event selection process, highlighting steps taken to optimise the procedure for electrons.

4.2 R_K as an experimental observable

In light of the differences between detecting muons and electrons at LHCb, it becomes clear that the expression for R_K given in Equation (2.10),

$$R_K = \frac{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathrm{d}q^2}\mathrm{d}q^2}{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathrm{d}q^2}\mathrm{d}q^2},$$

is susceptible to large systematic uncertainties. Measuring a muon process and an electron process, and then comparing the two by taking the ratio, would lead to uncertainties related to the differences between how muons and electrons are measured at LHCb. This would not be the case if the definition of R_K were adjusted to depend on ratios of two muon processes and of two electron processes, instead of one of each. For this reason, one more muon process and one more electron process are chosen to act as control channels, and R_K is measured relative to these two processes.

The chosen control channels are $B^+ \to K^+J/\psi(\ell^+\ell^-)$. Their branching fractions are larger than those of the rare $B^+ \to K^+\ell^+\ell^-$ modes, owing to the $b \to c$ transition that produces one of the valence quarks of the J/ψ resonance. In addition, the flavour-changing $W^\pm$ boson preferentially decays into the $c\bar{s}$ pair needed to create $J/\psi K^+$. Another advantage of this channel is that the resonant J/ψ structure leads to the overwhelming majority of $B^+ \to K^+J/\psi(\ell^+\ell^-)$ events to be characterised by a q^2 around the square of the mass of the J/ψ. This, combined with the fact that final-state particles from the two channels have similar kinematics, allows the selection strategies for the $B^+ \to K^+\ell^+\ell^-$ and $B^+ \to K^+J/\psi(\ell^+\ell^-)$ channels to be identical up to the cuts on q^2 and on $m(K^+\ell^+\ell^-)$. By introducing information from the $B^+ \to K^+J/\psi(\ell^+\ell^-)$ modes, the definition of the experimental observable R_K becomes:

$$R_K = \frac{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathrm{d}q^2}\mathrm{d}q^2}{\displaystyle\int_{q^2_{\text{min}}}^{q^2_{\text{max}}} \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathrm{d}q^2}\mathrm{d}q^2} \cdot \frac{\displaystyle\int_{q^2_{ee}} \frac{\mathrm{d}\mathcal{B}\,(B^+ \to K^+J/\psi(e^+e^-))}{\mathrm{d}q^2}\mathrm{d}q^2}{\displaystyle\int_{q^2_{\mu\mu}} \frac{\mathrm{d}\mathcal{B}\,(B^+ \to K^+J/\psi(\mu^+\mu^-))}{\mathrm{d}q^2}\mathrm{d}q^2}\,, \tag{4.2}$$

where q^2_{ee} and $q^2_{\mu\mu}$ are the q^2 selection regions for $B^+ \to K^+J/\psi(e^+e^-)$ and $B^+ \to K^+J/\psi(\mu^+\mu^-)$, respectively; these ranges, alongside q^2_{min} and q^2_{max}, are listed in Table 4.1. The final step in expressing R_K as a function of experimental quantities is to write the branching fractions in terms of yields and efficiencies. The yield of the final state X from a B^+ decay, $N(X)$, can be expressed as a function of the branching fraction of the process, $\mathcal{B}(X)$:

$$N(X) = \varepsilon(X) \cdot \mathcal{N}(X) = \varepsilon(X) \cdot \mathcal{B}(X) \cdot N(B)\,. \tag{4.3}$$

Here, $\varepsilon(X)$ is the efficiency to select a candidate for process X, and $\mathcal{N}(X)$ is the total number of times B^+ mesons decayed into X during data taking. The total number of produced B^+ mesons, $N(B)$, is a property that does not depend on the subsequent B^+ decay. This means that $N(B)$ cancels out in the ratios on the right-hand side of Equation (4.2), and therefore the branching fractions can be expressed in terms of their recorded yields and estimated efficiencies. This leads to the double-ratio expression for R_K:

$$R_K = \frac{N(K^+\mu^+\mu^-)}{\varepsilon(K^+\mu^+\mu^-)} \cdot \frac{\varepsilon(K^+e^+e^-)}{N(K^+e^+e^-)} \cdot \underbrace{\frac{\varepsilon(K^+J/\psi(\mu^+\mu^-))}{N(K^+J/\psi(\mu^+\mu^-))} \cdot \frac{N(K^+J/\psi(e^+e^-))}{\varepsilon(K^+J/\psi(e^+e^-))}}_{r_{J/\psi}}\,. \tag{4.4}$$

The yields are obtained through fits to the invariant mass of the final-state particles. These are described in Chapter 5 and Chapter 9 for the control and signal modes, respectively. The efficiencies are obtained based on simulated events, as described in Chapter 6. The terms in Equation (4.2) that are obtained from the control channels are grouped into the single ratio $r_{J/\psi}$. As detailed in Section 8.1, this observable is a stringent test of the experimental procedure. A related check, which uses the $\psi(2S)$ modes, is described in Section 8.2.

4.3 Data selection

In total, six processes are central to the R_K analysis: $B^+ \to K^+e^+e^-$, $B^+ \to K^+J/\psi(e^+e^-)$, and $B^+ \to K^+\psi(2S)(e^+e^-)$, together with their muon counterparts. Throughout, they are referred to as the "signal", "control", and "$\psi(2S)$" modes, respectively. The signal data is kept blind up until the full validation of the experimental procedure. Cuts are applied to select candidate events that match the requirements imposed by each channel, and remove as much background as possible. The selection strategy for this analysis is essentially identical to the one developed by Dr. Thibaud Humair and Dr. Paula Álvarez Cartelle for the previous measurement of R_K. A few modifications were made when necessary, and are highlighted where relevant.

The selection contains a number of criteria that target specific characteristics desirable of signal candidates. These are listed in Tables 4.1, 4.2, 4.3, and 4.5, where they are grouped by their purpose. For example, requirements that ensure particles are contained by the LHCb detector acceptance are listed under "Fiducial cuts". The following subsections provide details on the selection requirements, according their roles.

4.3.1 Invariant-mass cuts

The signal, control, and $\psi(2S)$ selections are identical, up to the cuts on q^2 and the invariant mass of all three final-state particles. These cuts are summarised in Table 4.1. The value of q^2 in control and $\psi(2S)$ events peaks at $m^2_{J/\psi}$ and $m^2_{\psi(2S)}$, respectively. For this reason, the q^2 windows for these channels are chosen to be around the values of the two masses quoted in the PDG [9]. To account for bremsstrahlung radiation, the boundaries of a given window are chosen such that their average is lower than the mass of the resonance. In addition, the $K^+\ell^+\ell^-$ invariant-mass windows have different widths between electron and muon modes, to take into account the different resolutions. It can be seen from projections of the fit to $B^+ \to K^+\ell^+\ell^-$ data, such as the ones shown in Figures 11.3 to 11.5, that the resolutions of $m(K^+\mu^+\mu^-)$ and $m(K^+e^+e^-)$ are approximately 20 MeV and 70 MeV, respectively.

The q^2 selection for the rare modes is designed to reduce background contamination. On

Table 4.1: The reconstructed q^2 and mass ranges used to separate the six channels used in this analysis. They are the only requirements that differ between the signal, control, and $\psi(2S)$ modes.

	q^2 selections	
	Electron channel	Muon channel
	Bounds [GeV^2]	Bounds [GeV^2]
signal	(1.10, 6.00)	(1.10, 6.00)
control	(6.00, 12.96)	(8.68, 10.09)
$\psi(2S)$	(9.92, 16.40)	(12.50, 14.20)

		mass selections	
		Electron channel	Muon channel
	Quantity	Bounds [GeV]	Bounds [GeV]
signal	$m(K^+\ell^+\ell^-)$	(4.88, 6.20)	(5.18, 5.60)
control	$m_{J/\psi}(K^+\ell^+\ell^-)$	(5.08, 5.68)	(5.18, 5.60)
$\psi(2S)$	$m_{\psi(2S)}(K^+\ell^+\ell^-)$	(5.08, 5.68)	(5.18, 5.60)

one hand, the lower bound of 1.1 GeV^2 rejects contributions from low-mass resonances, such as the $\phi(1020)$. On the other hand, the upper bound of 6.0 GeV^2 is chosen to minimise the background formed by $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ events in the low-mass tail of the resonance. It is estimated, by means described in Chapter 6, that this q^2 window contains approximately 25% of all $B^+ \to K^+\ell^+\ell^-$ events.

Following the same line of reasoning, the lower bounds of the invariant-mass windows are chosen to reduce contributions from physics backgrounds situated at low $m(K^+\ell^+\ell^-)$, whilst still being efficient at selecting the signal. The upper bound is chosen to enable a good description of the contribution from combinatorial events, which are seen in Figure 4.1 to dominate the high-$m(K^+\ell^+\ell^-)$ region. In the resonant J/ψ and $\psi(2S)$ modes, the $K^+\ell^+\ell^-$ invariant-mass estimate is improved during reconstruction by constraining the dilepton system to have mass equal to the PDG central values. This results in estimates, denoted by $m_{J/\psi}$ and $m_{\psi(2S)}$ respectively, that have better resolutions than the unconstrained mass [180].

4.3.2 Ensuring the quality of the decay

Loose requirements are applied at the beginning of the selection chain to reject combinatorial background events. These requirements ensure that the candidate has a topology compatible with a $B^+ \to K^+\ell^+\ell^-$ process, which is depicted in Figure 4.2. The B^+ is produced

Table 4.2: Offline selection cuts applied to the electron (left) and muon (right) samples.

Event quality		
$\chi^2_{\mathrm{FD}}(B^+)$	$>$	100
$\delta(B^+)$	$<$	arccos(0.995)
$\chi^2_{\mathrm{IP}}(B^+)$	$<$	25
$\chi^2_{\mathrm{DV}}/\mathrm{ndof}(\mathrm{B}^+)$	$<$	9
$\chi^2_{\mathrm{FD}}(e^+e^-)$	$>$	16
$\chi^2_{\mathrm{DV}}/\mathrm{ndof}(\mathrm{e}^+\mathrm{e}^-)$	$<$	9
$\chi^2_{\mathrm{IP}}(e^\pm)$	$>$	9
$\chi^2_{\mathrm{IP}}(K^+)$	$>$	9
$p_{\mathrm{T}}(K^+)$	$>$	400 MeV
`nSPDHits`	$<$	600 (Run 1)
	$<$	450 (Run 2)
$\texttt{prob}_{\mathrm{ghost}}(K^+, e^\pm)$	$<$	0.3
Cascade & mis-ID vetoes		
$m(K^+e^-)$	$>$	1885 MeV
$m(K^+e^-_{[\to\pi^-]})$	$\notin$	$m(D^0) \pm 40$ MeV
Fiducial cuts		
$\texttt{hasRich}(K^+, e^\pm)$	$=$	`true`
$\texttt{hasCalo}(e^\pm)$	$=$	`true`
$p_{\mathrm{T}}(e^\pm)$	$>$	0.5 GeV
$p(e^\pm)$	$>$	3 GeV
$\lvert x_{\mathrm{ECAL}}(e^\pm)\rvert$	$>$	363.6 mm
or $\lvert y_{\mathrm{ECAL}}(e^\pm)\rvert$	$>$	282.6 mm
PID cuts		
$\texttt{probNN}_K(K^+)$	$>$	0.2
$\texttt{DLL}_e(K^+)$	$<$	0
$\texttt{DLL}_e(e^\pm)$	$>$	3

Event quality		
$\chi^2_{\mathrm{FD}}(B^+)$	$>$	121
$\delta(B^+)$	$<$	arccos(0.9999)
$\chi^2_{\mathrm{IP}}(B^+)$	$<$	16
$\chi^2_{\mathrm{DV}}/\mathrm{ndof}(\mathrm{B}^+)$	$<$	8
$\chi^2_{\mathrm{FD}}(\mu^+\mu^-)$	$>$	9
$\chi^2_{\mathrm{DV}}/\mathrm{ndof}(\mu^+\mu^-)$	$<$	12
$\chi^2_{\mathrm{IP}}(\mu^\pm)$	$>$	9
$\chi^2_{\mathrm{IP}}(K^+)$	$>$	6
`nSPDHits`	$<$	600 (Run 1)
	$<$	450 (Run 2)
$\texttt{prob}_{\mathrm{ghost}}(\mu^\pm)$	$<$	0.3
Cascade & mis-ID vetoes		
$m(K^+\mu^-)$	$>$	1885 MeV
$m(K^+\mu^-_{[\to\pi^-]})$	$>$	1885 MeV
$m(K^+_{[\to\mu^+]}\mu^-)$	$\notin$	$m(J/\psi) \pm 60$ MeV
$m(K^+_{[\to\mu^+]}\mu^-)$	$\notin$	$m(\psi(2S)) \pm 60$ MeV
Fiducial cuts		
$\texttt{hasRich}(K^+, \mu^\pm)$	$=$	`true`
$\texttt{inMuonAcc}(K^+, \mu^\pm)$	$=$	`true`
$p_{\mathrm{T}}(\mu^\pm)$	$>$	0.8 GeV
PID cuts		
$\texttt{probNN}_K(K^+)$	$>$	0.2
$\texttt{isMuon}(K^+)$	$=$	`false`
$\texttt{DLL}_\mu(\mu^\pm)$	$>$	-3
$\texttt{isMuon}(\mu^\pm)$	$=$	`true`

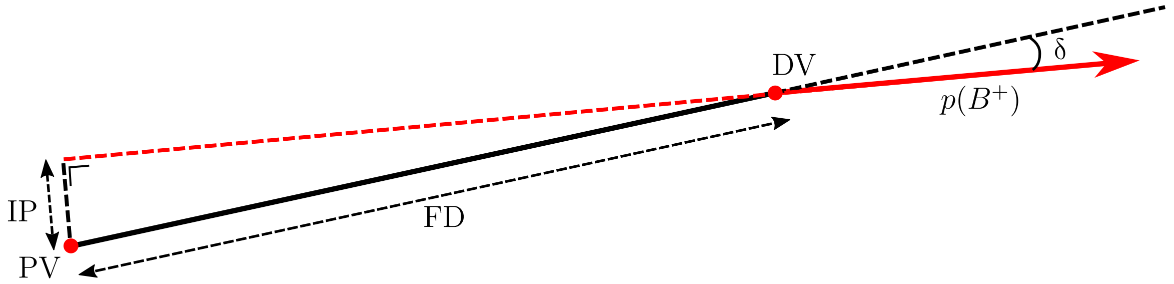

Figure 4.2: Schematic of a B^+ decay. The annotations represent quantities relevant to the R_K selection, as explained in the main body. Diagram adapted from Ref. [10].

at the primary vertex (PV) and decays at the decay vertex (DV). The distance between the PV and the DV represents the flight distance (FD), and the angle between it and the reconstructed B^+ momentum is denoted by δ. Following the momentum direction and drawing the perpendicular line that contains the PV is what defines the impact parameter (IP).

The B^+ decay vertex is required to be of good quality, and to be well separated from the primary vertex. The former requirement is ensured by imposing an upper limit on the χ^2 per degrees of freedom obtained from the fit for the DV. The latter requirement is enforced by selecting only events where the B^+ travels a significant distance away from the PV before decaying. In addition, since the true momentum of the B^+ is collinear with the FD, requiring the angle δ to be small ensures that the decay is well aligned. Finally, all final-state particles are required to be inconsistent with being produced in a proton-proton collision, and so tracks whose impact parameter with respect to any PV is not significant are rejected. Since that is not the case for the B^+, it is required to be compatible with having been produced from a PV. The significance of the IP with respect to the PV is quantified by χ^2_{IP}.

The requirements thus far are applied during the stripping stage, which was mentioned in Section 3.3. Two more cuts are applied after the stripping to further improve the purity of the data samples. The first one rejects overly crowded events by placing an upper limit on `nSPDHits`, which represents the number of hits in the SPD. The second one removes candidates containing tracks that could be fictitious. These are called ghosts, and they arise when the reconstruction algorithm uses hits produced either by noise, or by other particles. The variable used to reject ghosts is denoted by $\texttt{prob}_{\text{ghost}}$ in Table 4.2. The `nSPDHits` and $\texttt{prob}_{\text{ghost}}$ cuts, together with the q^2 selection and the cuts discussed in Sections 4.3.5 and 4.3.6, are collectively known as the preselection.

4.3.3 Trigger strategy

Data collected by the LHCb detector during Run 1 and Run 2 is required to pass a two-stage trigger selection: the `L0` and HLT that are introduced in Section 3.3. The trigger can fire on particles in the candidate (the kaon and the two leptons considered to have originated from a

B^+ decay), on particles that are not part of the candidate, and on both. These three scenarios are referred to as "TOS", "TIS", and "TOB" (trigger on signal, trigger independently of the signal, and trigger on both, respectively). The rest of this subsection covers the specific trigger strategies employed by the R_K analysis.

Muon data is collected using an `L0` trigger line that requires at least one track, with p_{T} above a threshold, whose trajectory is compatible with energy deposits in the muon stations. This is known as the `L0Muon` line. A similar strategy is used to select electron data, through the `L0Electron` line. It requires at least one track whose trajectory is compatible with energy deposits, above a certain threshold, in the ECAL. This line is less efficient than `L0Muon`, by a factor of approximately 2 – 3, as exemplified by the trigger efficiencies listed in Appendix D. For this reason, two more `L0` strategies are used to increase the electron-mode yields. The first strategy requires at least one track whose trajectory is compatible with energy deposits in the HCAL, and is thus called `L0Hadron`. The second strategy requires at least one of several `L0` lines to be TIS. Hence, this strategy is known as `L0TIS`.

Each of the four `L0` strategies used in this analysis is illustrated in Figure 4.3. These trigger lines use information from the calorimeters and muon stations to apply a fast reconstruction algorithm. This results in rough estimates of transverse momenta and energy deposits[1], denoted by $p_{\mathrm{T}}^{\mathrm{L0}}$ and $E_{\mathrm{T}}^{\mathrm{L0}}$, respectively. Although their resolution is not as good compared to their fully-reconstructed counterparts, p_{T} and E_{T}, they are calculated more quickly and therefore allow the `L0` to make fast decisions. More specifically, `L0Muon` places a lower threshold on the highest $p_{\mathrm{T}}^{\mathrm{L0}}$ in the event, whilst `L0Electron` and `L0Hadron` use thresholds on the highest $E_{\mathrm{T}}^{\mathrm{L0}}$ in a given event. These thresholds are listed in the top three rows of Table 4.3.

During data taking, the thresholds fluctuate by a few percent to maintain consistent efficiencies in spite of effects such as changes to the collision environment and detector ageing. This can cause disagreement with simulation, because the simulated samples corresponding to one particular year and one particular magnet polarity are generated using only one trigger configuration, due to computing constraints. The only exception is 2018, when data-taking

---

[1] The transverse component is calculated using the polar angle of the line joining the primary vertex and the centre of the cells that make up a cluster [156].

conditions were kept constant. To improve the agreement in the electron samples, events at the threshold are rejected after the preselection, by applying fiducial cuts on the E_{T} of particles that fire `L0Electron` or `L0Hadron`. These fiducial cuts are listed in the bottom two rows of Table 4.3. The values chosen for 2017 and 2018 data are based on studies that constitute original work. The agreement in the muon samples is improved by means that differ between data-taking periods. In 2017, the simulation is generated with the loosest conditions used to take data, and so the agreement is improved by selecting the simulation such that the L0 conditions are similar. In 2015 and 2016, some data is taken with conditions looser than the ones used in the simulation. Therefore, only events collected with the `L0Muon` configuration used to generate the simulated samples are kept. In Run 1, the fluctuations in $p_{\mathrm{T}}^{\mathrm{L0}}$ thresholds are small enough to not require alignment with the simulation.

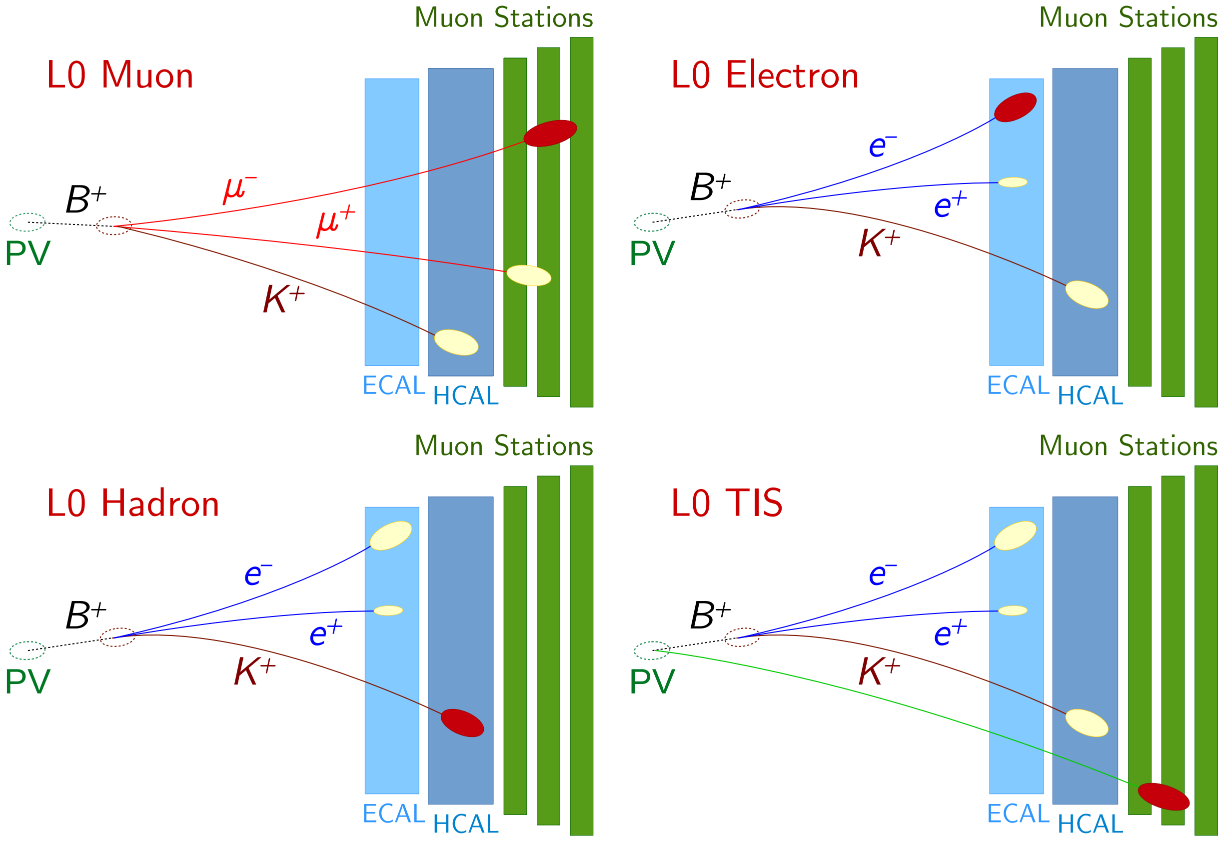

Figure 4.3: Diagrams depicting the L0 trigger strategies employed in the R_K analysis. The red-filled ellipse in each diagram indicates the particle that fires the corresponding trigger. The example given for `L0TIS` shows the line firing due to energy deposits in the muon station, however it can also be triggered by the ECAL and HCAL. Diagrams created by Dr. Paula Álvarez Cartelle.

Table 4.3: Requirements placed by the `L0` selection, in the majority of data collected in each year. The thresholds are imposed at the trigger-level, whilst the fiducial cuts are applied offline to improve the agreement between data and simulation. All values are given in GeV.

	2011	2012	2015	2016	2017	2018
`L0Muon` $p_{\mathrm{T}}^{\mathrm{L0}}$ threshold	1.5	1.8	2.9	1.9	1.5	1.8
`L0Electron` $E_{\mathrm{T}}^{\mathrm{L0}}$ threshold	2.5	2.7	2.7	2.4	2.1	2.4
`L0Hadron` $E_{\mathrm{T}}^{\mathrm{L0}}$ threshold	3.5	3.6	3.6	3.7	3.5	3.8
`L0Electron` E_{T} fiducial	3.0	3.0	2.7	2.7	2.9	3.2
`L0Hadron` E_{T} fiducial	3.0	3.0	3.5	3.5	3.5	3.5

After this process of improving the agreement between data and simulation, muon-channel events fall into what is referred to as the μTOS trigger category. This contains candidates where `L0Muon` fires on at least one of the signal muons. The events in the electron samples fall into one of three exclusive trigger selections. The dominant one, which accounts for roughly two thirds of all $B^+ \to K^+e^+e^-$ data, is called eTOS. It contains events where `L0Electron` fires on at least one of the signal electrons, and that electron has E_{T} above the appropriate fiducial cut. Candidates that are not eTOS fall into the hTOS! category if they contain a kaon that triggers `L0Hadron` and has E_{T} above the appropriate fiducial cut. This category contains 15–20% of all $B^+ \to K^+e^+e^-$ data. The rest is found in the TIS! category, consisting of events that are neither eTOS nor hTOS!, but where the `L0` fires independently of the signal. The exclamation marks at the end of hTOS! and TIS! indicate that the trigger strategies are exclusive: an event can be assigned to only one of the three trigger categories, prioritising eTOS and then hTOS!.

The next step in the trigger selection is the HLT, which is done in two stages: HLT1 and HLT2. The former partially reconstructs tracks in the event, in order to make a fast decision on whether the event is likely to contain interesting physics. Particularly for the R_K analysis, at least one of the three tracks in Run 1 events must have large enough p_{T} and χ^2_{IP} estimates to pass the HLT1 selection. In Run 2, the decision is made by a multivariate classifier, based on information such as p_{T} and χ^2_{IP}. Events that pass the HLT1 are sent to the HLT2, which performs a full reconstruction of all tracks in the event. The lines used in the R_K measurement search for two- or three-track topologies compatible with originating from the decay of a heavy object, such as a B^+ meson [181].

4.3.4 Particle identification requirements

The primary reconstruction object is a charged track, which could come from any charged particle species. To reject background events where one or several tracks in the candidate are misidentified as other particles, several requirements are placed on variables that distinguish between different particle hypotheses. The selection in the R_K measurement makes use of three types of PID-discriminating variables. The first one uses information from the RICH, calorimeters, and muon stations to construct the likelihood that a given track is of a given species, relative to the pion hypothesis. This PID-discriminating variable is denoted in Table 4.2 by $\texttt{DLL}_X$, where X is the particle hypothesis to be compared to the pion one. Following the same notation, the second variable is called $\texttt{probNN}_X$. It is the subunitary output of a neural network trained on information from several subdetectors to distinguish between different particle species; this information includes the $\texttt{DLL}_X$ variables. The third and final variable is the boolean `isMuon` decision, which uses information from the muon stations to disentangle muons from other particle species.

As a result of the PID requirements, background events containing misidentified pions are reduced to negligible levels. This is particularly important in LHCb analyses, given that proton-proton collisions produce a considerable amount of pions. Contributions from the Cabibbo-suppressed $B^+ \to \pi^+\ell^+\ell^-$ mode, where the pion is misidentified as a kaon, are expected to amount to around 0.4% of the signal yield. This estimate is obtained based on the branching fraction averages listed in the PDG [9], and the PID efficiencies obtained from simulation. Given that the total $B^+ \to K^+\ell^+\ell^-$ signal yield is expected to be $\mathcal{O}(10^3)$, this contribution is deemed negligible. Candidates originating from $B^+ \to K^+\pi^+\pi^-$ events, where the pions are misidentified as leptons, are also expected to be negligible. Their levels in muon and electron data are estimated to be 0.6% and 0.8%, respectively. These estimates were obtained by Dr. Thibaud Humair during the previous R_K measurement [10]. As part of the current measurement, a cross-check of the expected $B^+ \to K^+\pi^+\pi^-$ contributions is conducted and presented in Section 8.5.

4.3.5 Fiducial cuts

In order to apply the PID and trigger requirements described in the previous two subsections, all final-state particles in the $K^+\ell^+\ell^-$ candidate are required to be within the geometric acceptance of the RICH detectors and the calorimeters. In the muon samples, the particles are also required to be within the acceptance of the muon stations. The relevant variables are denoted in Table 4.2 by `hasRich`, `hasCalo`, and `inMuonAcc`, respectively. In the electron data, a portion of the inner ECAL is vetoed, because it contains cells that are not read out. This is done by requiring the intersection of electron candidate tracks with the ECAL plane to have (x, y) coordinates $|x_{\mathrm{ECAL}}| > 363.6\,\mathrm{mm}$ or $|y_{\mathrm{ECAL}}| > 282.6\,\mathrm{mm}$. Furthermore, requirements are applied on the momenta of electrons and muons, in order to align the selection with the one employed to obtain the samples used to calibrate PID efficiencies, as described in Section 6.2.

4.3.6 Vetoes against specific backgrounds

One potential source of background events is the semileptonic decay of a D meson originating from a B^+ that had also decayed semileptonically. These are referred to as cascade backgrounds, and examples include $B^+ \to \overline{D}^0(K^+e^-\overline{\nu}_e)e^+\nu_e$ processes, where the neutrinos are not detected. Cascade backgrounds are expected to accumulate at invariant masses below the mass of the B^+, as a result of the undetected energy carried away by the neutrinos. However, the tails that could overlap with the signal are enhanced by the tree-level nature of $B^+ \to DX$ decays. For this reason, cuts that specifically target cascade backgrounds are added to the selection. Their discrimination power comes from the fact that the invariant mass of the kaon and the opposite-sign lepton in a candidate, $m(K^+\ell^-)$, cannot be larger than the mass of the D^0 in the case of cascade background events, notwithstanding resolution effects. This mass is known with good precision to be $\sim 1865\,\mathrm{MeV}$ [9], therefore candidate electron and muon events are required to have $m(K^+\ell^-) > 1885\,\mathrm{MeV}$.

It can be seen on the left-hand side of Figure 4.4 that this cut reduces contamination from $B^+ \to \overline{D}^0(K^+e^-\overline{\nu}_e)\,e^+\nu_e$ events to a negligible level. The same is true for events where the B^+ decays into a $\overline{D}^0$ and a π^+, the latter being misidentified as a signal electron; these are labelled as $B^+ \to \bar{D}^0(K^+e^-\bar{\nu}_e)\,\pi^+{}_{[\to e^+]}$. However, the application of the $m(K^+\ell^-) > 1885\,\text{MeV}$ cut still leaves behind a significant tail from events where the $\bar{D}^0$ decays hadronically into a K^+ and a π^-. These are labelled as $B^+ \to \bar{D}^0(K^+\pi^-{}_{[\to e^-]})\,e^+\nu_e$, and are removed by computing $m(K^+e^-)$ under the assumption that the electron candidate is a pion. This leads to $m(K^+e^-{}_{[\to\pi^-]})$, whose distribution in simulated events is shown on the right-hand side of Figure 4.4. It can be seen that $B^+ \to \bar{D}^0(K^+\pi^-{}_{[\to e^-]})\,e^+\nu_e$ events peak in this mass around m_{D^0}. Therefore, cutting ±40 MeV around m_{D^0} is removes this background source, whilst retaining the vast majority of signal events.

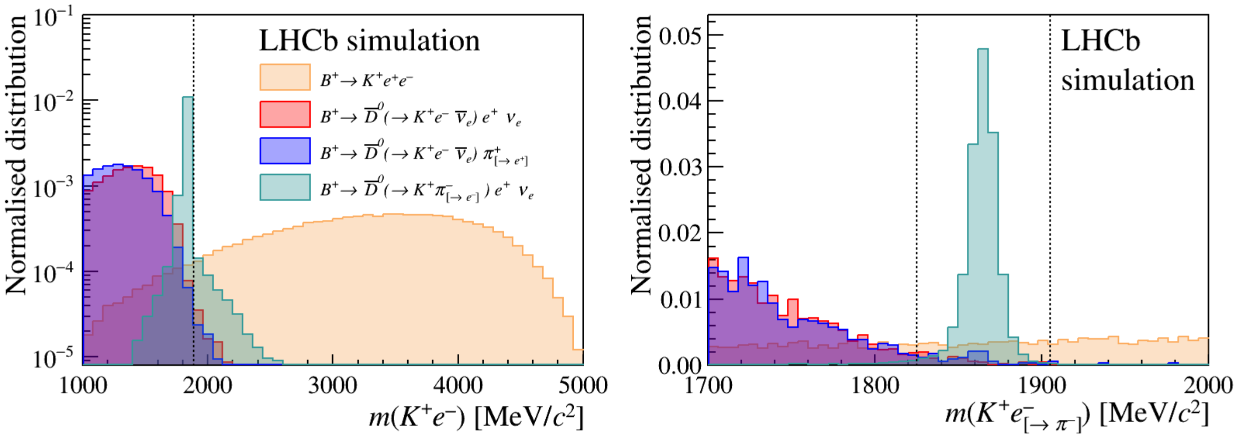

Figure 4.4: Simulated signal (orange) and background (red, green, blue) distributions of the invariant mass of the kaon and the opposite-sign electron. The latter is reconstructed assuming an electron (left) and a pion (right) mass hypothesis. The dotted lines show the locations of the mass vetoes described in the main body.

In the case of muons, the vetoes are different for two reasons. First, cascade backgrounds featuring pion decays in flight are more prevalent, and can lead to $m(K^+\mu^-)$ values below m_{D^0}. For this reason, all values of $m(K^+\mu^-{}_{[\to\pi^-]})$ below 1885 MeV are rejected, instead of removing just the D^0 peak as is done for electrons. Second, muons can be more kaon-like than electrons, and so two additional cuts are applied to the muon samples. They reject J/ψ and $\psi(2S)$ decays into pairs of muons, where one of the muons is misidentified as a kaon. This is done by reconstructing the invariant mass of the candidate kaon and opposite-sign muon, under the assumption that the kaon is a muon. This mass is denoted by $m(K^+{}_{[\to\mu^+]}\mu^-)$ in Table 4.2, and is expected to peak at the mass of either the J/ψ or the $\psi(2S)$ resonance if it corresponds to a background process of this kind. For this reason, candidate muon events are required to not have $m(K^+{}_{[\to\mu^+]}\mu^-)$ values within 60 MeV of either resonance.

4.3.7 Multivariate selection

Some combinatorial events are expected to pass the requirements presented so far. Therefore, a dedicated selection is implemented to further suppress this background. Boosted decision trees (BDTs) [182] are trained to distinguish between signal and combinatorial background. This multivariate selection was designed and implemented by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair during the previous R_K measurement. For Run 2.2 data, the procedure is repeated on the new samples by Dr. Konstantinos Petridis.

The BDTs are trained on the variables listed in Table 4.4. As discussed in Section 4.3.2, they contain information about the decay topology. This makes them useful in rejecting events containing tracks that are not formed by the decay products of a B^+. Kinematic information is kept to a minimum, and is provided only in the form of the p_T of the particles. This ensures that the BDT does not learn how to reconstruct invariant masses, such as q^2 and $m(K^+\ell^+\ell^-)$; that would lead to sculpting of these variables.

Table 4.4: List of variables used by the BDT classifiers.

B^+	$p_\mathrm{T}, \log\chi^2_\mathrm{IP}, \chi^2_\mathrm{DV}, \delta, \chi^2_\mathrm{FD}$
$\ell^+\ell^-$	$p_\mathrm{T}, \log\chi^2_\mathrm{IP}$
K^+	$p_\mathrm{T}, \log\chi^2_\mathrm{IP}$
$\ell^\pm$	$\min,\max(p_\mathrm{T}), \min,\max(\log\chi^2_\mathrm{IP})$

The BDTs are trained separately for each data-taking period, and for electron and muon samples. Particularly for the electrons, BDTs are trained independently on samples from each of the three trigger selection, as well as on the combined samples. The background training sample is taken from data with $m(K^+\ell^+\ell^-) > 5.4$ GeV. This is known as the upper sideband, and it consists of purely combinatorial events; this is depicted by the horizontal bands in Figure 4.1. The signal training samples are fully selected and calibrated $B^+ \to K^+\ell^+\ell^-$ simulated events, because the BDTs are designed to be optimal on the rare samples. To prevent statistical biases, the k-folding method [183] is employed during training and validation, with $k = 10$ folds. Figure 4.5 showcases the BDT performance. The equivalent curves obtained from other data-taking periods are similar, and are thus omitted.

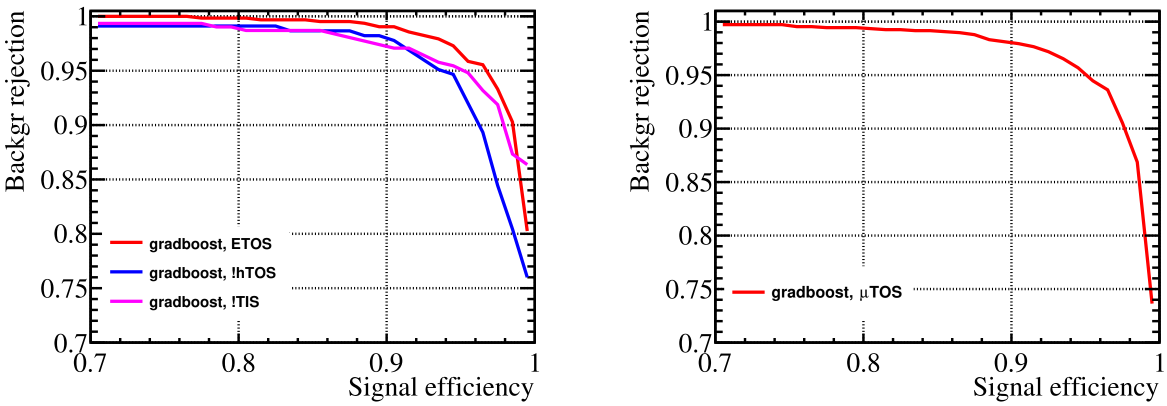

Figure 4.5: BDT background rejection rate, as a function of signal efficiency, as obtained from the training and testing of a fold from the 2018 electron (left) and muon (right) samples. The optimised working point has a signal efficiency of around 90%. Plots created by Dr. Konstantinos Petridis.

Each BDT is optimised separately, based on the expected significance:

$$\mathcal{S} = \frac{S}{\sqrt{S+B}}\,. \tag{4.5}$$

In the above expression, S and B are the expected signal and background yields, respectively. The former is obtained from the estimated efficiencies of the BDT and selection requirements on the signal, normalised to the control mode:

$$S = \varepsilon^{\mathrm{BDT}}_{\mathrm{rare}} \cdot \frac{\varepsilon^{\mathrm{sel}}_{\mathrm{rare}}}{\varepsilon^{\mathrm{sel}}_{J/\psi}} \cdot \frac{\mathcal{B}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{B}(B^+ \to K^+J/\psi)\mathcal{B}(J/\psi \to \ell^+\ell^-)} \cdot N^{\mathrm{sel}}_{J/\psi}\,. \tag{4.6}$$

In this context, the selection refers to the application of the trigger and preselection requirements. The branching fractions in the above equation are taken from the PDG. The yields of the control modes, $N^{\mathrm{sel}}_{J/\psi}$, are obtained from invariant-mass fits to partially-selected $B^+ \to K^+J/\psi(\ell^+\ell^-)$ data. These fits are presented in Sections 5.2 and 5.5.

The background estimate is calculated differently for muon and electron BDTs. The higher statistics of the former channel allow the parametrisation of the upper sideband as an exponential function. As a result, B is estimated by extrapolating the best-fit exponential curve down to the signal region and integrating the area underneath it. The statistics in the electron sideband are too low to lead to an accurate parametrisation, so a proxy has to be used instead. The proxy is a sample with higher statistics, and with shape that is expected to be similar to the one found in the signal sample. The chosen proxy consists of $B^+ \to K^+e^+\mu^-$ events. These are mostly combinatorial, owing to the PID and mass vetoes successfully removing events from other background sources. The estimated background yield in the signal region is then:

$$B = N^{\mathrm{SB}}_{K^+e^+e^-} \cdot \frac{N^{\mathrm{SR}}_{K^+e^+\mu^-}}{N^{\mathrm{SB}}_{K^+e^+\mu^-}}, \tag{4.7}$$

where $N^{\mathrm{SB}}_{K^+e^+e^-}$ is the yield of the $K^+e^+e^-$ sideband. The ratio between $K^+e^+\mu^-$ yields in the signal region (SR) and the sideband (SB) is obtained by integrating the exponential function which best fits the $K^+e^+\mu^-$ invariant-mass spectrum. The fit results are shown in Figure 4.6.

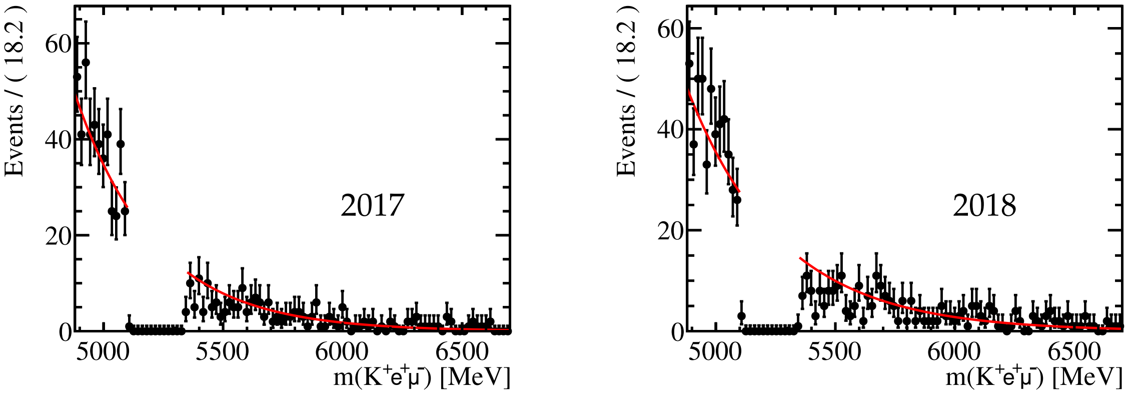

Figure 4.6: Exponential fit (red) to the $K^+e^+\mu^-$ samples (black) used to extrapolate the number of events from the sideband (m > 5400 MeV) to the signal region ($m \in$ (5000 MeV, 5380 MeV)). To avoid potential contributions from lepton flavour violating events, the region $m \in$ (5100 MeV, 5350 MeV) is blinded and excluded from the fit. Plots created by Dr. Konstantinos Petridis.

The expected significance, $\mathcal{S}$, is computed for a series of cuts on the output of each BDT. The working point of each classifier is chosen to correspond to the maximum significance. It is found that eTOS reaches the optimal performance when using a BDT specifically trained on eTOS samples. However, because hTOS! and TIS! have fewer events, BDTs trained on just these samples tend to overfit. For this reason, the BDT trained on all electron data is used in these trigger categories. The expected significance as a function of the working points of the four BDTs used on 2018 data are presented in Figure 4.7. The results from the other data-taking periods are similar, and thus omitted. The working points used in all data-taking periods and trigger selections are listed in Table 4.5, alongside the estimated signal and combinatorial yields in the signal window. It can be seen that the electron combinatorial estimates based on $B^+ \to K^+e^+\mu^-$ data are more precise than, and compatible with the results coming from only using the $K^+e^+e^-$ sideband.

The BDTs have one more purpose, other than the rejection of combinatorial events. As shown in Table 4.6, a sub-percent fraction of selected events contain multiple candidates. It is expected that only one candidate in an event is a genuine signal process. Therefore, in any given event with multiple candidates, only the one with the highest BDT output is retained.

The BDT represents the final stage of the selection chain. Once the data are fully selected, the next step in the analysis consists of understanding the control modes. This is presented in the next chapter.

Table 4.5: Expected signal and combinatorial yields in the signal window. The optimal BDT cuts are listed under "WP", and found by means described in the main body. The estimates under "Comb. 2" are calculated by extrapolating the number of events from the upper sideband to the signal region, using an exponential fit to the distribution of $m(K^+\ell^+\ell^-)$. For electron samples, a more accurate estimate, listed under "Comb. 1", is derived using $B^+ \to K^+e^+\mu^-$ data, as discussed in the main body.

	WP	Signal	Comb. 1	Comb. 2
		Run 1		
eTOS	0.89	181 ± 2	56 ± 11	100 ± 50
hTOS!	0.79	58 ± 1	83 ± 17	76 ± 31
TIS!	0.87	64 ± 1	24 ± 9	30 ± 28
μTOS	0.81	989 ± 6		78 ± 17
		Run 2.1		
eTOS	0.86	294 ± 1	95 ± 14	80 ± 25
hTOS!	0.86	86 ± 2	40 ± 11	60 ± 40
TIS!	0.85	85 ± 2	40 ± 9	47 ± 20
μTOS	0.79	873 ± 5		59 ± 19
		2017		
eTOS	0.81	270 ± 2	85 ± 8	77 ± 23
hTOS!	0.80	74 ± 2	39 ± 8	22 ± 10
TIS!	0.80	78 ± 2	31 ± 6	59 ± 24
μTOS	0.65	942 ± 6		55 ± 17
		2018		
eTOS	0.78	377 ± 3	115 ± 10	110 ± 26
hTOS!	0.75	114 ± 1	80 ± 13	45 ± 14
TIS!	0.80	115 ± 1	45 ± 9	112 ± 44
μTOS	0.69	1150 ± 5		60 ± 15

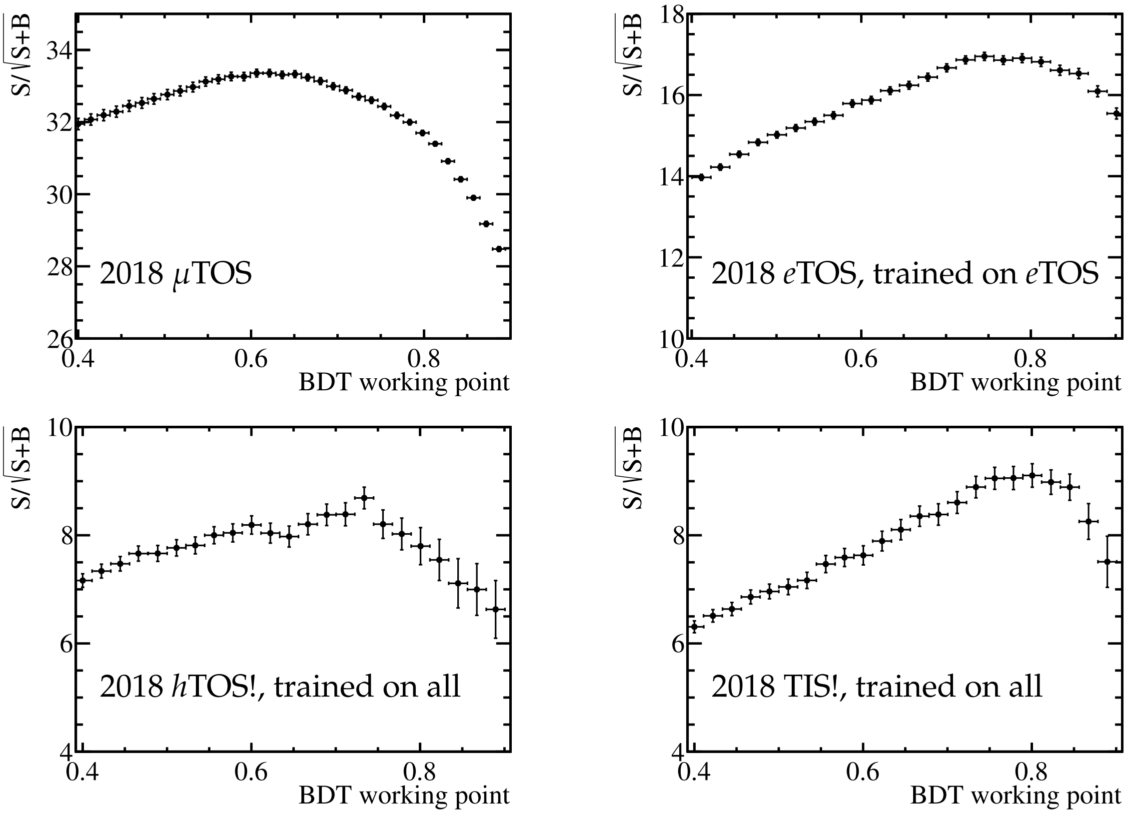

Figure 4.7: Expected significance in 2018 muon and electron data, as a function of the BDT working point.

Table 4.6: Fraction of events for which more than one μTOS or eTOS candidate is reconstructed, when the full selection chain is applied. The equivalents for hTOS! and TIS! are similar, and therefore omitted.

	simulation				data	
	rare ee	rare $\mu\mu$	control ee	control $\mu\mu$	control ee	control $\mu\mu$
Run 1	0.30%	0.01%	0.20%	0.04%	0.10%	0.01%
Run 2.1	0.47%	0.03%	0.32%	0.03%	0.17%	0.03%
2017	0.51%	0.06%	0.35%	0.05%	0.34%	0.04%
2018	0.52%	0.04%	0.37%	0.04%	0.37%	0.03%

Chapter 5

Fits to the control modes

The $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ data is used in the definition of R_K in Equation (4.4), and to calibrate the simulation before the determination of efficiencies. This chapter describes the extraction of information from the control channels. This is done using fits performed separately on samples from each data-taking period and trigger selection.

The electron and muon fit strategies are described in Sections 5.1 and 5.4, respectively. The validation of the fits is presented in Section 5.7. The fits are first conducted before the application of the multivariate selection, in order to obtain a background-subtracted sample from which the corrections in Chapter 6 are derived. The results from muon and electron data are presented in Sections 5.2 and 5.5, respectively. Sections 5.3 and 5.6 present subsequent fits performed on the fully-selected muon and electron samples to obtain the yields that serve as inputs to R_K.

The fit procedure was developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous R_K measurement. The Run 2.2 results presented throughout the chapter constitute original work. They are obtained using the same fit procedure, with a few minor adjustments that are highlighted where relevant. The results in Section 5.7 also constitute original work.

Throughout the R_K analysis, fits are performed using the extended maximum likelihood method [184], as implemented by the `RooFit` package [185]. The procedure involves the optimisation of a likelihood, for which the `Minuit` algorithm [186] is used. The likelihood is calculated by modelling the distribution of events in data as the sum of a signal and several background components:

$$\mathcal{P}^{rt}(m) = N^{rt}_{\text{sig}}\mathcal{S}^{rt}(m) + \sum_i N^{rt}_{\text{bkg},i}\mathcal{B}^{rt}_i(m)\,. \tag{5.1}$$

Here, $\mathcal{P}^{rt}(m)$ is the model for the distribution of mass m in data taken during period r using trigger t. The signal is modelled by $\mathcal{S}^{rt}(m)$, and the background models (indexed by i) are labelled $\mathcal{B}^{rt}_i(m)$. The yields of the signal and background components are N^{rt}_{sig} and $\{N^{rt}_{\text{bkg},i}\}$, respectively. The distributions $\mathcal{S}^{rt}(m)$ and $\mathcal{B}^{rt}_i(m)$ are not identical between fits to muon data and electron data, hence they are described separately in the following two sections.

5.1 Fit strategy for $B^+ \to K^+J/\psi(\mu^+\mu^-)$ data

The fits to $B^+ \to K^+J/\psi(\mu^+\mu^-)$ data are performed separately for Run 1, Run 2.1, 2017, and 2018 data from the μTOS trigger. The independent variable is the invariant mass of the $K^+\ell^+\ell^-$ system, as obtained from a kinematic fit where the mass of the dilepton system has been constrained to the literature value of the J/ψ mass [180]. This mass is denoted by $m_{J/\psi}(K^+\mu^+\mu^-)$, and has better resolution than its unconstrained counterpart, $m(K^+\mu^+\mu^-)$.

The fit model consists of one signal component and two background components. The signal is modelled using a Hypatia distribution [187], which has non-Gaussian tails that are able to model the radiative tails of the signal. In fits to Run 2.2 data, a Gaussian distribution is added to the signal model to improve the quality of the fit. It is shown in Appendix A.1 that the signal model accurately describes simulated $B^+ \to K^+J/\psi(\mu^+\mu^-)$ samples. Since the effects that lead to the tails are well simulated, the Hypatia tail parameters are fixed to values obtained from these fits to simulation. However, imperfections in the simulated momentum calibration and detector resolution prevent the extraction of the mean and widths of the signal components from simulation. Instead, these quantities in data are reparametrised in terms of the shift ($\Delta\mu$) and, respectively, scale (s_σ) with respect to their simulated counterparts:

$$\mu^{\text{data}} = \mu^{\text{sim}} + \Delta\mu, \qquad \sigma^{\text{data}} = s_\sigma \cdot \sigma^{\text{sim}}. \tag{5.2}$$

Thus, the mean $\Delta\mu$ and scale s_σ are variables in the fit to data, and are shared by the Hypatia and Gaussian components of the signal model.

The first background component accounts for $B^+ \to \pi^+ J/\psi(\mu^+\mu^-)$ events where the pion is incorrectly identified as a kaon. This is expected to peak approximately 50 MeV above the signal, and to have non-Gaussian tails due to the incorrect mass hypothesis. For these reasons, the $B^+ \to \pi^+ J/\psi(\mu^+\mu^-)$ component is modelled using the sum of two Crystal Ball (CB) distributions [188] that have the same mean and width, but exponential tails on opposite sides. These tails are constrained using simulated $B^+ \to \pi^+ J/\psi(\mu^+\mu^-)$ samples, following the same ethos as the signal shape. The mean and width are parametrised in terms of the same shift $\Delta\mu$ and scale s_σ that are used by the signal model. The $B^+ \to \pi^+ J/\psi(\mu^+\mu^-)$ yield, $N_{\text{mis-ID}}$, is constrained relative to the yield of the signal, N_{sig}, based on the branching fractions of the two B^+ decays, and estimated efficiencies. The constraint takes the form of a Gaussian with mean given by:

$$\frac{N_{\text{mis-ID}}}{N_{\text{sig}}} = \frac{\mathcal{B}(B^+ \to \pi^+ J/\psi(\mu^+\mu^-))}{\mathcal{B}(B^+ \to K^+ J/\psi(\mu^+\mu^-))} \cdot \frac{\varepsilon_{J/\psi K}}{\varepsilon_{J/\psi \pi}}, \tag{5.3}$$

and width equal to 5% of the mean; this is driven by the uncertainty on the branching fractions, which dominates the error on the efficiencies.

The final background component consists of random combinations of kaon and muon (combinatorial events). The shape of this background is modelled by an exponential function:

$$\mathcal{B}^{rt}_{\text{combi}}(m) \propto e^{-\lambda m}. \tag{5.4}$$

5.2 Calibration fits to $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data

This section presents the results of the first fits to $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data, as outlined in the introduction to this chapter. These fits are conducted before the multivariate and multiple-candidate selections, and before any efficiency-correcting weights are applied. This is because said weights require the output of these fits: clean samples of control-mode data. These samples are obtained by separating the signal component from the backgrounds using the *sWeight* technique [189]. The fit results are presented in Figure 5.1. The pulls between the data and the model distribution, $\mathcal{P}^{rt}(m)$, are shown below each plot.

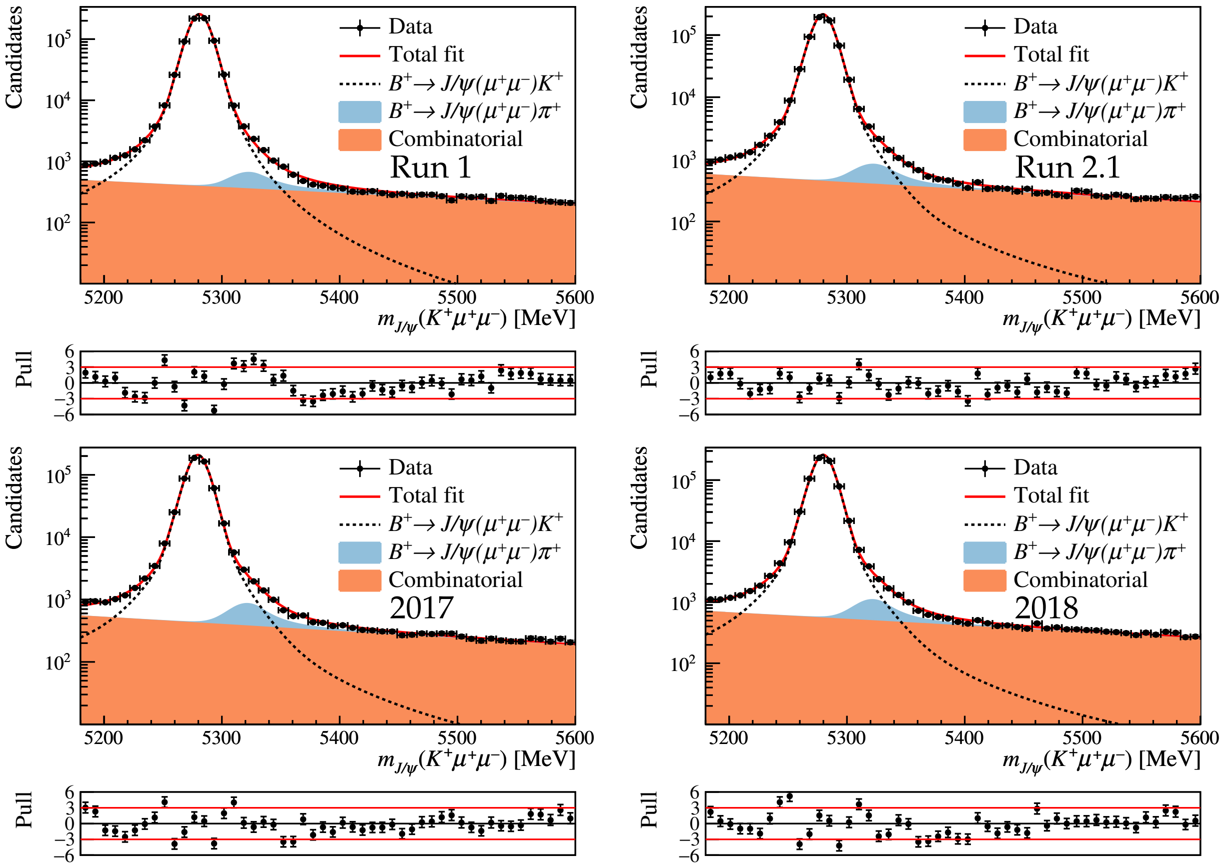

Figure 5.1: Fits to the $m_{J/\psi}(K^+\mu^+\mu^-)$ distribution of $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ partially-selected candidates in the samples corresponding to each data-taking period. The red solid line shows the fit model, the dotted black line represents the signal, the blue-filled area shows misidentified $B^+ \to \pi^+ J/\psi(\mu^+\mu^-)$ events, and the orange-filled area depicts the combinatorial background. Shown below each plot are the pulls between the data and the total fit model.

5.3 Fits to fully-selected $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data

After applying the rest of the selection to the $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data samples, the fits are performed once more to determine the yields $N(K^+ J/\psi(\mu^+\mu^-))$. The procedure is left unchanged, however the result is different in two ways: first, the signal and misidentified shape constraints are recomputed, because the corrections to the simulation samples can now be used to improve the agreement with data. Second, the additional selection changes the yields of the components. In particular, the multivariate selection considerably reduces the combinatorial background, whilst leaving the other components almost untouched. As explained in Section 4.3.7, this is the intended behaviour of the multivariate classifier.

The fits are shown in Figure 5.2, alongside the pulls between the fit model and the data. These pulls are generally small, indicating that the fit suitably models the data. As expected, the combinatorial background is greatly suppressed with respect to the fits to partially-selected data. The signal yields are listed in Table 5.1, and the values of the fit parameters that are found to maximise the likelihood are provided in Appendix A.3. It can be seen that Run 2.2 contains nearly the same amount of events as the sample used in the previous R_K analysis, with 2018 and Run 2.1 containing around 600 000 $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ events each, and 2017 and Run 1 having over 500 000 events each.

5.4 Fit strategy for $B^+ \to K^+ J/\psi(e^+e^-)$ data

The procedure employed to fit $B^+ \to K^+ J/\psi(e^+e^-)$ data is similar to the one used for muons, with a few notable adaptations imposed by the different resolution of electrons. Both signal and background distributions are wider, to the point where a non-negligible amount of background events is expected to overlap with the signal. Such background events are decay chains, started by a hadron, that produce a kaon, two oppositely-charged electrons, and other particles. If the decay is reconstructed using the kaon and the electrons, but not the other particles, the event can pass the selection even though it is not signal. These are called partially-reconstructed events. Given that the decay products that escape reconstruction

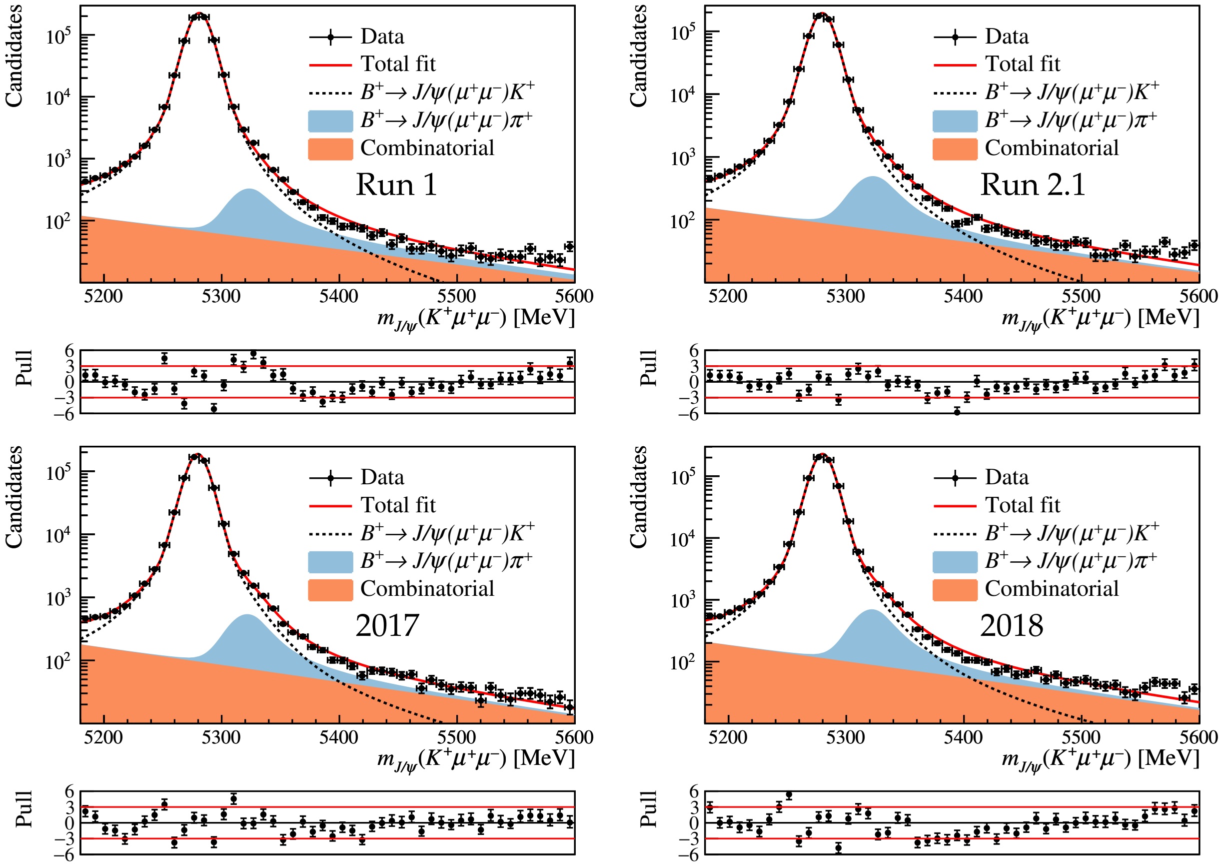

Figure 5.2: Fits to the $m_{J/\psi}(K^+\mu^+\mu^-)$ distribution of $B^+ \to K^+J/\psi(\mu^+\mu^-)$ fully-selected candidates in the samples corresponding to each data-taking period. The red solid line shows the fit model, the dotted black line represents the signal, the blue-filled area shows misidentified $B^+ \to \pi^+J/\psi(\mu^+\mu^-)$ events and the orange-filled area depicts the combinatorial background. Shown below each plot are the pulls between the data and the total fit model.

take away some of the initial-state energy, partially-reconstructed events are characterised by $K^+\ell^+\ell^-$ invariant masses below the mass of the B^+. The separation is sufficiently large to disentangle these background events from the signal in the muon modes, however the poorer resolution of the electron channels leads to overlap with the signal.

In the fits to $B^+ \to K^+J/\psi(e^+e^-)$ data, around 97% of all partially-reconstructed events consist of $H_b \to Y(K^+X)J/\psi$ processes, where a beauty hadron (H_b) decays into a pair of electrons and a strange resonance (Y), such as the K^{*0} or the ϕ meson. These resonances then decay into a charged kaon, and other particles (X) that escape detection. Such events are referred to as strange partially-reconstructed backgrounds. They are joined by decay chains containing charmed hadrons (such as charmonium resonances) that also produce a pair of electrons, a charged kaon, and other particles that escape detection. Such processes, which include

$B^+ \to \psi(2S)(J/\psi\pi^0)K^+$ events where the π^0 is not reconstructed, are referred to as charmed partially-reconstructed backgrounds. Their mass distributions are different from those of their strange counterparts, so the two backgrounds are included in the fit as separate components. The shapes of these components are obtained from simulated events, using an adaptive kernel density estimation method [190]. The same simulation samples are used to apply a Gaussian constraint to the relative abundances of charm and strange partially-reconstructed events. The central value of the constraint is the ratio between the numbers of simulated events of each type that survive the selection. The width of the constraint is equal to 15% of the central value.

Another consequence of the poor electron resolution is the necessity to use the bremsstrahlung recovery process described in Section 4.1 to reclaim some of the lost energy. Since electrons with associated bremsstrahlung radiation have different energies and resolutions compared to electrons without recovered bremsstrahlung radiation, the invariant mass $m(K^+e^+e^-)$ — and, by extension, $m_{J/\psi}(K^+e^+e^-)$ — is expected to have different distributions in events where bremsstrahlung photons are found for either none, one, or both electrons. This is captured in the fit through the splitting of the signal component into three distributions:

$$\mathcal{S}^{rt}(m) = f^{rt}_{0\gamma}\,\mathcal{S}^{rt}_{0\gamma}(m) + f^{rt}_{1\gamma}\,\mathcal{S}^{rt}_{1\gamma}(m) + (1 - f^{rt}_{0\gamma} - f^{rt}_{1\gamma})\,\mathcal{S}^{rt}_{2\gamma}(m)\,. \tag{5.5}$$

Here, $\mathcal{S}^{rt}_{0\gamma}(m)$ is the distribution that models signal events (from data-taking period r and trigger t) where none of the electrons have added bremsstrahlung clusters; these events fall into the so-called 0γ photon category. The other photon categories, 1γ and 2γ, represent events where one and, respectively, both electrons have added bremsstrahlung clusters. Their distributions are modelled by $\mathcal{S}^{rt}_{1\gamma}(m)$ and $\mathcal{S}^{rt}_{2\gamma}(m)$, respectively. Each of the three distributions that make up $\mathcal{S}^{rt}(m)$ is the sum of two CB distributions of the same mean, with exponential tails on separate sides of the central region. The low-mass tail models the radiative energy losses below the signal peak, which are non-Gaussian in nature. The high-mass tail takes into account the fact that the bremsstrahlung recovery algorithm sometimes overestimates the amount of radiated energy, thus leading to a non-Gaussian contribution above the signal peak. Like with the muons, the simulation provides a suitable description of the processes which lead to the tails, as exemplified in Appendix A.1. For this reason, the tail parameters in the fits to data are fixed using simulated events. The simulation also provides estimates of the fractions of 0γ and 1γ events, relative to the total. These fractions are denoted in Equation (5.5) by $f_{0\gamma}^{rt}$ and $f_{1\gamma}^{rt}$, respectively. In the fit, each of these parameters is constrained to the simulated values, by means of a Gaussian distribution whose mean is the estimated value from simulation, and whose width is equal to 1% of the mean.

Like in the muon case, the yield of the misidentified $B^+ \to \pi^+ J/\psi(e^+e^-)$ component is constrained with respect to the signal, based on the known branching fractions and estimated efficiencies. Again, the shape is modelled by two CB distributions with a shared mean and exponential tails on opposite sides. The parameters of the tails are modelled using $B^+ \to K^+ J/\psi(e^+e^-)$ events, by virtue of a method that estimates the effect of the change between kaon and pion mass, m_K and m_π:

$$m_{\text{mis-ID}}^{J/\psi} = \sqrt{(m_{J/\psi}(K^+e^+e^-))^2 + \frac{E_{B^+}}{E_{K^+}} \cdot (m_{K^+}^2 - m_{\pi^+}^2)}\,. \tag{5.6}$$

In the above expression, E_{B^+} and E_{K^+} are the energies of the B^+ and K^+ candidates. This departure from the muon fits is due to simulated $B^+ \to \pi^+ J/\psi(e^+e^-)$ decays not being available for all data-taking periods. Nevertheless, $B^+ \to \pi^+ J/\psi(e^+e^-)$ and $B^+ \to K^+ J/\psi(e^+e^-)$ have very similar kinematics, making the distribution of $m_{\text{mis-ID}}^{J/\psi}$ in $B^+ \to K^+ J/\psi(e^+e^-)$ events a good approximation (up to small corrections induced by kaon-pion mass differences) to that of $m_{J/\psi}(K^+e^+e^-)$ in misidentified $B^+ \to \pi^+ J/\psi(e^+e^-)$ events. This technique was introduced and validated by Dr. Thibaud Humair for the previous R_K measurement.

Following the same procedure as in the muon fits, the means and widths of the signal and misidentified data distributions are reparametrised in terms of the shift $\Delta\mu$ and scale s_σ with respect to their values from simulation. Also following the example of the muons, the shape formed by combinatorial events is modelled by an exponential function, with freely-floating normalisation and exponent. In cases where the multivariate selection is highly efficient at removing combinatorial events, the fit is performed without including this component.

Checking this against a fit strategy that allowed the combinatorial yield to take negative values found that the signal yield does not change significantly. This is the only departure from the fitting strategy developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous R_K measurement. Note that in the fits to fully-selected Run 1 and Run 2.1 data, the typical number of combinatorial background events is compatible with 0.

5.5 Calibration fits to $B^+ \to K^+ J/\psi(e^+e^-)$ data

Like with the muon channels, and for the same reasons, fits to $B^+ \to K^+ J/\psi(e^+e^-)$ data are first performed before the application of the multivariate and multiple-candidate selections. The *sWeight* technique is employed to extract clean $B^+ \to K^+ J/\psi(e^+e^-)$ samples used to correct efficiencies and to optimise the multivariate selection. The result of the fits to eTOS data from all data-taking periods is shown in Figure 5.3. The fit components are qualitatively similar between the four samples, and the pulls depicted below each plot demonstrate that the fits are of adequate quality for the use of the *sWeight* method.

5.6 Fits to fully-selected $B^+ \to K^+ J/\psi(e^+e^-)$ data

Following the application of the full selection chain, the $B^+ \to K^+ J/\psi(e^+e^-)$ fit is performed once again to extract the yields. The fits to data taken with the most efficient electron trigger strategy, eTOS, are shown in Figure 5.4. The results from the other triggers are presented in Appendix A.2. The pulls between the data and the fit model are shown below each plot, and indicate good fit quality. The signal yields from each data-taking period and trigger selection are listed in Table 5.1, and the values of the fit parameters that maximise the likelihood are presented in Appendix A.3. The Run 2.2 sample contains slightly more events than the sample used in the previous R_K analysis: around 275 000 events compared to approximately 250 000. Combined with the fact that the muon yields are roughly equally split between the Run 2.2 and the previous samples, this supports the expectation that adding Run 2.2 effectively doubles the dataset.

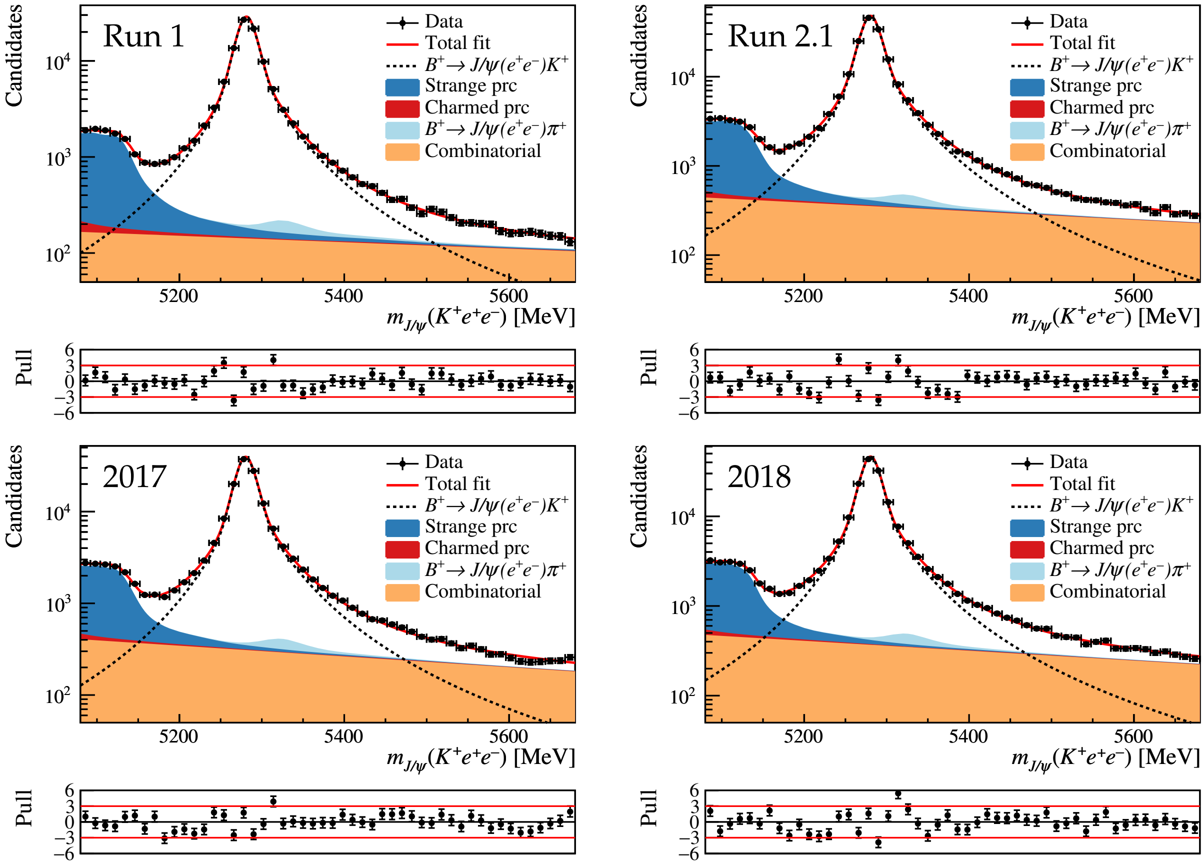

Figure 5.3: Fits to the $m_{J/\psi}(K^+e^+e^-)$ distribution of $B^+ \to K^+J/\psi(e^+e^-)$ partially-selected eTOS candidates in the samples corresponding to each data-taking period. The red solid line represents the fit model, the dotted black line is the signal component, the light-blue filled area represents misidentified $B^+ \to \pi^+J/\psi(e^+e^-)$ events, and the orange-filled area shows the combinatorial background. The strange and charm partially-reconstructed backgrounds (referred to as "prc" in the legend) are represented by the filled areas filled with dark blue and red, respectively. Shown below each plot are the pulls between the data and the total fit model.

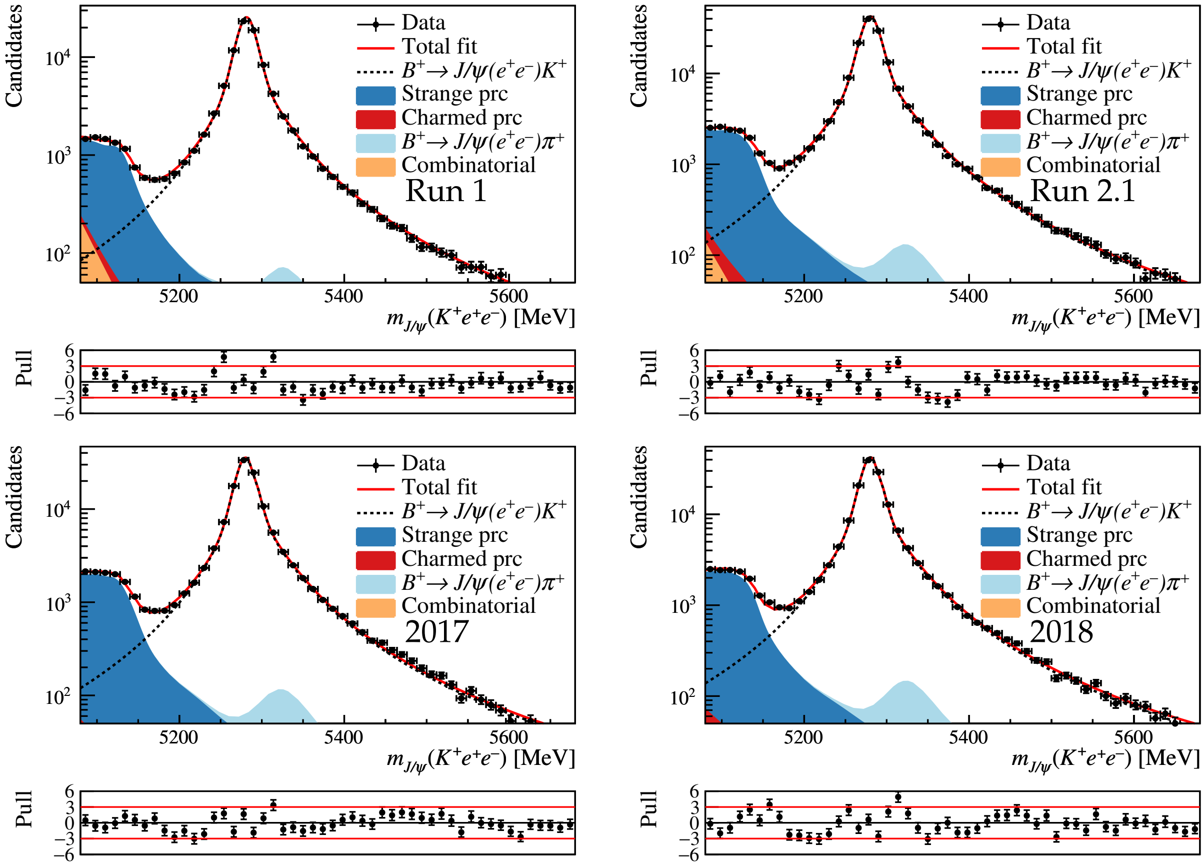

Figure 5.4: Fits to the $m_{J/\psi}(K^+e^+e^-)$ distribution of $B^+ \to K^+ J/\psi(e^+e^-)$ fully-selected eTOS candidates in the samples corresponding to each data-taking period. The red solid line represents the fit model, the dotted black line is the signal component, the light-blue filled area represents misidentified $B^+ \to \pi^+ J/\psi(e^+e^-)$ events, and the orange-filled area shows the combinatorial background. The strange and charm partially-reconstructed backgrounds (referred to as "prc" in the legend) are represented by the filled areas filled with dark blue and red, respectively. Shown below each plot are the pulls between the data and the total fit model.

Table 5.1: Control-mode yields in data from each data-taking period and trigger selection. The first column corresponds to muon data, whilst the other three represent the electron samples taken using the three trigger categories defined in Section 4.3.3. The uncertainties on the total yields represent the addition in quadrature of the individual yields.

	μTOS	eTOS	hTOS!	TIS!
Run 1	618332 ± 796	90215 ± 309	15394 ± 127	30361 ± 178
Run 2.1	543457 ± 749	153363 ± 408	16961 ± 131	37810 ± 200
2017	507094 ± 722	126293 ± 367	17470 ± 136	37888 ± 199
2018	619616 ± 800	148885 ± 401	22767 ± 156	45855 ± 230
Total	2288499 ± 1535	518756 ± 747	72592 ± 276	151914 ± 405

5.7 Validation of the fit

The estimates of signal yields (N_{sig}) obtained from the control-mode fits could be biased with respect to the true ones. If these biases are not negligible, they might affect the correct determination of R_K. To assess the size of any potential biases, the fit results are used to generate pseudoexperiments, also known as toys, that mimic the data samples. The shapes of the fit components are kept identical, and their normalisation factors are allowed to float according to their Poisson statistics, in order to ensure correct coverage of the fit parameters. The fit procedure is repeated on each pseudoexperiment, and the extracted signal yield, $N_{\text{sig}}^{\text{fit}}$, is compared to the one used to generate the toy, $N_{\text{sig}}^{\text{gen}}$. The difference between the two is referred to as the residual, and its distribution across pseudoexperiments is expected to tend towards a Gaussian of mean 0 if the fit is unbiased. The distribution of the residuals obtained from pseudoexperiments generated to represent data collected during the year 2018 are shown in Figure 5.5. The other data-taking periods are also checked and found to be qualitatively similar, and are therefore omitted. For each distribution, χ^2 minimisation is employed to determine the best-fit Gaussian, whose mean is interpreted as the bias of the fit. All biases are found to be around 10–110 events, which represents a negligible amount compared to the expected 10^4–10^5 yields.

5.8 Summary of control-mode fits

Understanding the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ samples is an important step towards measuring R_K. The first round of fits, presented in Sections 5.2 and 5.5, are vital to the optimisation of the multivariate selection, and to the calculating of efficiencies in Chapter 6. The subsequent fits, which are covered by Sections 5.3 and 5.6, are needed to test the validity of the efficiencies, by means described in Chapter 8. As shown in Section 5.7, the fit procedure is found to lead to no significant biases in the signal yields. Once calculated, the efficiencies and yields are incorporated in the likelihood used by the fit to the rare modes, from which R_K is derived. As expanded upon in Chapter 9, the model for the rare-mode data has several similarities with the fits described in this chapter. This means that the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ fits represent an important exercise towards understanding the $B^+ \to K^+\ell^+\ell^-$ data.

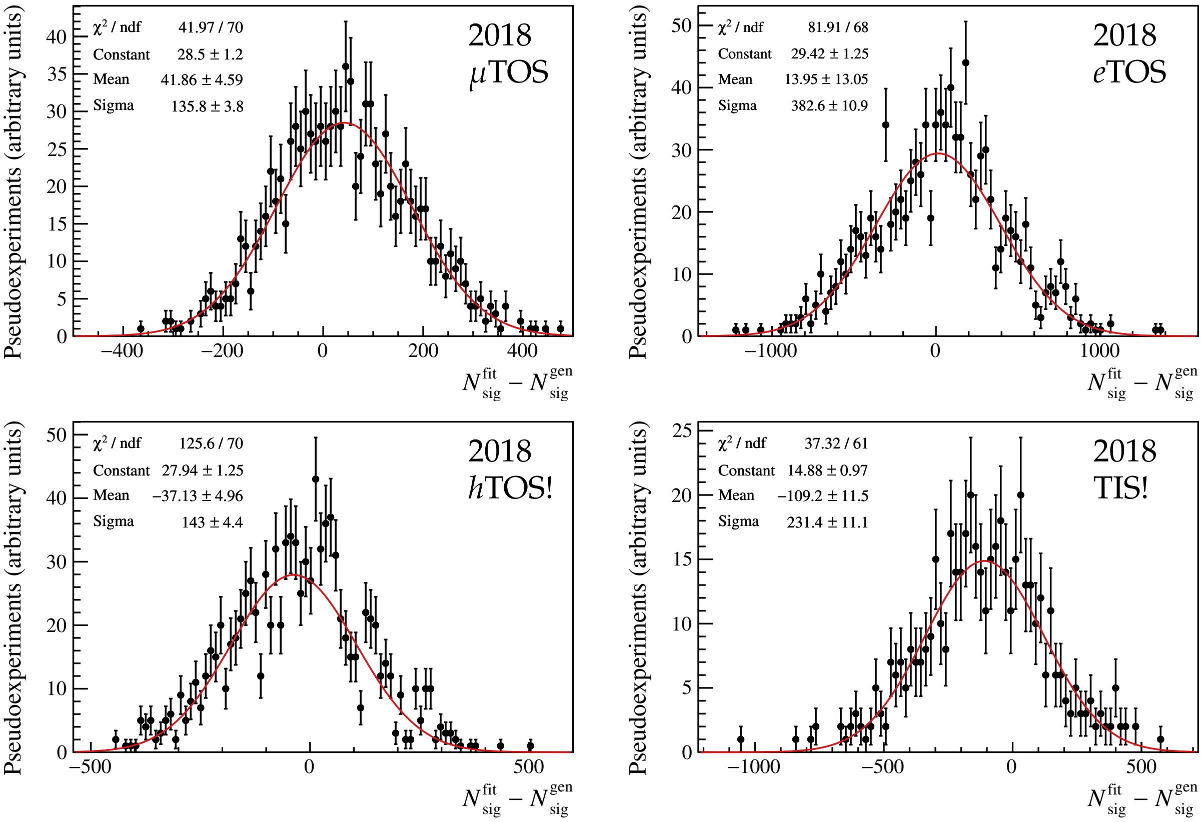

Figure 5.5: Distributions of the residuals from 2018 $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ pseudoexperiments (black), alongside the best-fit Gaussian distribution (red). The minimum χ^2 from the Gaussian fit is presented on the top left of each plot, alongside the best-fit parameters of the Gaussian.

Chapter 6

Calculation of efficiencies

Half of the factors in equation Equation (4.4), where the experimental observable R_K is defined, represent the efficiencies with which data are selected. Therefore, correctly determining efficiencies is vital to the R_K measurement. This is done based on simulated samples that are corrected to account for known imperfections of the simulation. This chapter presents the methods used to calibrate simulated events, as well as the resulting efficiency estimates.

The procedure was developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous R_K measurement, with a few changes that are highlighted where relevant. As such, Run 2.2 results are presented preferentially over their Run 1 and Run 2.1 counterparts.

Data used to conduct the R_K measurement is required to have passed multiple selection criteria, each targeting specific properties that candidates must have. This means that the total efficiency, ε_{tot}, depends on the performance of each part of the selection described in Section 4.3. Since the simulation models these effects with different levels of accuracy and precision, they need to be grouped to reflect the calibration methods employed to compute the efficiency. To this end, the total efficiency is factorised as:

$$\varepsilon_{\text{tot}} = \varepsilon_{\text{geom}} \cdot \varepsilon_{\text{rec,strip}} \cdot \varepsilon_{\text{presel}} \cdot \varepsilon_{\text{PID}} \cdot \varepsilon_{\text{trig}} \cdot \varepsilon_{\text{BDT}} \,. \tag{6.1}$$

In the expression above,

- $\varepsilon_{\text{geom}}$ is the geometric acceptance of the LHCb detector: all tracks are required to have a polar angle between 10 mrad and 400 mrad;
- $\varepsilon_{\text{rec,strip}}$ is the efficiency of the reconstruction and of the stripping selection (omitting the PID cuts therein);
- $\varepsilon_{\text{presel}}$ is the efficiency of the preselection (defined in Section 4.3.2) and the q^2 cut;
- ε_{PID} is the efficiency of all PID cuts;
- $\varepsilon_{\text{trig}}$ is the trigger efficiency; and
- ε_{BDT} is the efficiency of the BDT selection and the invariant-mass fit window cut.

Each term on the right-hand side of the above expression is computed on events that fulfil the requirements imposed by the preceding terms. This means that, for example, $\varepsilon_{\text{presel}}$ should strictly speaking be represented as $\varepsilon(\text{presel} \,|\, \text{geom, rec, strip})$. For brevity, the short-hand notation in Equation (6.1) is used for the rest of this thesis.

In the case of the resonant modes, all events are produced at q^2 equal to the square of the mass of either the J/ψ or the $\psi(2S)$. However, the true dilepton invariant mass squared, q^2_{true}, of $B^+ \to K^+\ell^+\ell^-$ events can take any values from $4m_\ell^2$ to $(m_{B^+} - m_{K^+})^2$. An additional term is therefore calculated to take into account the fact that R_K is not measured across the entire

q^2 spectrum. This factor, denoted by f^{q^2}, appears in the following expression for the number of selected $B^+ \to K^+\ell^+\ell^-$ events:

$$\begin{aligned} N_{\text{sel}}(B^+ \to K^+\ell^+\ell^-) &= \varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-) \cdot \mathcal{N}_{\text{all}}(B^+ \to K^+\ell^+\ell^-) && (6.2)\\ &= \varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-) \cdot \mathcal{N}_{\text{in}}(B^+ \to K^+\ell^+\ell^-) \cdot \frac{1}{f^{q^2}} \, . && (6.3) \end{aligned}$$

Here, $\mathcal{N}_{\text{all}}$ is the total number of events across q^2_{true}, and $\mathcal{N}_{\text{in}}$ is the number of events with $q^2_{\text{true}} \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$. This leads to the following expression for the ratio between the differential rare-mode branching fraction, averaged over $q^2_{\text{true}} \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$, and the control-mode branching fraction:

$$\begin{aligned} \frac{\mathcal{B}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{B}(B^+ \to K^+J/\psi(\ell^+\ell^-))} &= \frac{\mathcal{N}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{N}(B^+ \to K^+J/\psi(\ell^+\ell^-))} && (6.4)\\ &= \frac{\mathcal{N}_{\text{all}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{N}(B^+ \to K^+J/\psi(\ell^+\ell^-))} \cdot f^{q^2} && (6.5)\\ &= \frac{N_{\text{sel}}(B^+ \to K^+\ell^+\ell^-)}{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)} \cdot \frac{\varepsilon_{\text{tot}}(B^+ \to K^+J/\psi(\ell^+\ell^-))}{N_{\text{sel}}(B^+ \to K^+J/\psi(\ell^+\ell^-))} \cdot f^{q^2} \, . && (6.6) \end{aligned}$$

The efficiencies are computed using simulated events that are weighted to correct known imperfections in the simulation. In total, there are four sets of weights:

- w_{PID}: these represent the product of the PID efficiencies of the final-state particles: $w_{\text{PID}} = \varepsilon_{\text{PID}}(K^+) \cdot \varepsilon_{\text{PID}}(\ell^-) \cdot \varepsilon_{\text{PID}}(\ell^+)$, where $\varepsilon_{\text{PID}}(K^+)$ and $\varepsilon_{\text{PID}}(\ell^\pm)$ are the PID efficiencies of the kaon and leptons, respectively, in each candidate;
- w_{trig}: weights that calibrate the simulated trigger response;
- $w^{\text{rec}}_{\text{kin}}$: weights that correct the reconstructed kinematics of the B^+; and
- $w^{\text{gen}}_{\text{kin}}$: weights that adjust the generated kinematics of the B^+.

Section 6.5 explains why two sets of kinematic weights are needed. The presence of weights turns each efficiency in Equation (6.1) into a ratio between two sums of weights. The sum in the numerator runs over simulated events that pass the relevant selection criterion, whereas the sum in the denominator runs over the events that pass all requirements in the selection chain, up until the relevant criterion. The first two efficiencies, $\varepsilon_{\text{geom}}$ and $\varepsilon_{\text{rec,strip}}$, are treated separately from the rest, because they require samples to which the reconstruction is not applied. As outlined in Section 3.4, these are referred to as generation-level simulation samples. There are only four sums that don't cancel out in Equation (6.1), meaning that the final expression for the total efficiency is:

$$\varepsilon_{\text{tot}} = \underbrace{\frac{\sum\limits_{\text{rec}} w^{\text{gen}}_{\text{kin}}}{\sum\limits_{\text{gen}} w^{\text{gen}}_{\text{kin}}}}_{\varepsilon_{\text{geom}}\cdot\varepsilon_{\text{rec,strip}}} \cdot \underbrace{\frac{\sum\limits_{\text{sel}} w^{\text{rec}}_{\text{kin}} \cdot w_{\text{PID}} \cdot w_{\text{trig}}}{\sum\limits_{\text{rec}} w^{\text{rec}}_{\text{kin}}}}_{\varepsilon_{\text{presel}}\cdot\varepsilon_{\text{PID}}\cdot\varepsilon_{\text{trig}}\cdot\varepsilon_{\text{BDT}}}, \tag{6.7}$$

where $\sum_{\text{rec}}$, $\sum_{\text{gen}}$, and $\sum_{\text{sel}}$ run over all reconstructed, generated, and fully-selected events, respectively. The fraction f^{q^2} is calculated by taking the ratio between the sums of $w^{\text{gen}}_{\text{kin}}$ over generation-level simulated events:

$$f^{q^2} = \frac{\sum\limits_{\text{sel}\, q^2_{\text{true}}} w^{\text{gen}}_{\text{kin}}}{\sum\limits_{\text{all}\, q^2_{\text{true}}} w^{\text{gen}}_{\text{kin}}}. \tag{6.8}$$

The sum in the numerator runs over events with $q^2_{\text{true}} \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$, and the sum in the denominator covers all generated events. Note that the q^2 value used here is prior to any resolution effect, and also prior to radiation emitted by particles in the final state, in the presence of the magnetic fields surrounding the decay. This is known as final-state radiation, and is predominantly emitted by electrons due to their small mass. The effect on kaons and muons is approximately 1‰ [74], and so q^2_{true} is taken to be equal to $\left\| p^{\text{true}}_{B^+} - p^{\text{true}}_{K^+} \right\|^2$. As stated in Section 3.4, final-state radiation is described by PHOTOS, which has been tested and validated in Ref. [74].

The following section introduces what is known as the truth-matching algorithm, on which the calculation of efficiencies relies. Subsequent sections describe each of the individual corrections to the simulation. A summary of the corrected estimates of the efficiencies is then provided in Section 6.7.

6.1 Truth-matching and ghosts

The LHCb simulation software propagates all generated particles through the detector, and runs the entire reconstruction chain to obtain simulated events that match the data as closely as possible. An important feature provided by the simulation is the access to the "true" parameters with which a reconstructed particle has been generated. This is called the truth-matching algorithm, and it enables the study of effects such as detector resolution.

In some cases, the truth-matching algorithm does not successfully retrieve the generation-level information for all particles in the candidate. This causes some signal events to be mis-classified as ghosts. The rate at which this happens in simulated signal and control samples is at the percent level, as shown in Table 6.1. Therefore, ghosts are taken into account in the calculation of efficiencies, in order to avoid underestimating the efficiencies by as much as a few percent. This opens up the possibility of a bias being introduced by events classified as ghosts that are not signal. Such background events are suppressed in three ways. First, only candidates with one track classified as a ghost are retained. Second, the application of the $\texttt{prob}_{\text{ghost}}$ requirements listed in Table 4.2 suppresses most ghost events. Third, the multivariate selection described in Section 4.3.7 removes combinatorial events almost entirely. Since ε_{BDT} is the last term on the right-hand side of Equation (6.1), simulated events classified as ghosts are used only when calculating the total efficiency. They are excluded from the calculation of correction weights, because at that stage the samples still contain non-negligible levels of combinatorial background events, which would lead to potential biases.

Table 6.1: Relative fractions of simulated events being classified as ghosts in fully selected $B^+ \to K^+\ell^+\ell^-$ and $B^+ \to K^+J/\psi(\ell^+\ell^-)$ events, not counting multiple candidates.

	$K^+e^+e^-$	$K^+J/\psi(e^+e^-)$	$K^+\mu^+\mu^-$	$K^+J/\psi(\mu^+\mu^-)$
Run 1	4.2%	3.7%	1.5%	1.6%
Run 2.1	4.9%	4.2%	2.0%	1.9%
2017	4.7%	4.0%	1.9%	1.8%
2018	4.9%	4.0%	2.0%	1.8%

6.2 PID efficiencies

The PID techniques employed at LHCb make use of information from most subdetectors. This makes the PID-discriminating variables particularly susceptible to imperfections in the simulation of the LHCb detector and of the proton-proton collision environment. To prevent such effects from influencing the evaluation of the PID performance, the efficiencies are calibrated using data — rather than simulation — samples, where the species of final-state particles are known unambiguously. The efficiency of a given PID selection can then be obtained by applying it to the data, and using invariant-mass fits to obtain the yields before and after the selection. Since the PID performance depends on the kinematics of the particle, the fits are performed in several regions of phase space.

Depending on the particle species, one of two methods is used to calculate efficiencies. The method outlined in Section 6.2.1 is employed to estimate the kaon and muon identification efficiencies, as well as pion misidentification rates. The PID efficiency of electrons is estimated using the method described in Section 6.2.2.

6.2.1 PID efficiency of kaons, muons, and pions

As part of the LHCb data-taking programme, dedicated PID calibration samples are recorded and made available for all analyses [191]. All steps, from data selection to generation of PID calibration tables, are executed centrally. The procedure uses $J/\psi \to \mu^+\mu^-$ data to extract muon PID efficiencies, and $D^{*+} \to D^0(K^-\pi^+)\pi^+_{\text{slow}}$ data to evaluate the performance of kaons and pions. The former benefits from the excellent resolution of muons, which leads to a clean peak at $m(J/\psi)$ in the dimuon invariant-mass spectrum. In the latter, the small difference between the mass of the D^{*+} and that of the D^0 leads the pion labelled as "slow" to have low momentum. This allows its charge to uniquely determine the charge of the kaon and pion from the D^0 decay. Since the kaon's charge is of opposite sign with respect to the charge of the pions, samples of kaons and pions are obtained unambiguously. To further suppress backgrounds, the two final-state particles in the J/ψ and D^0 decays are treated as a "tag" and a "probe". PID efficiencies are evaluated on the probe, whilst a tight PID requirement is applied on the tag to ensure sample purity. To avoid biasing the efficiencies, the same fiducial cuts applied to $B^+ \to K^+\ell^+\ell^-$ data, listed in Section 4.3.5, are applied to the PID calibration data.

The dedicated samples are used to evaluate the efficiency of the kaon and muon PID cuts listed in Table 4.2. The `isMuon` efficiency is better modelled by the simulation, and is therefore treated separately from the `DLL` and `probNN` cuts. The ratio between `isMuon` efficiencies in calibration and simulation samples is used to obtain weights that strengthen the agreement between data and simulation. The PID efficiencies are also evaluated over pions, in order to constrain the $B^+ \to \pi^+ J/\psi(\ell^+\ell^-)$ components of the control-mode fits, and to conduct the cross-checks covered by Section 8.5. Given their relatively limited usage, the pion misidentification efficiencies are beyond the scope of this thesis.

The *sWeight* method [189] is employed to select signal events in the calibration data. The PID efficiency of a given cut is then obtained by summing the *sWeight*s of events before and after the application of said cut. Several factors, such as data-taking conditions and kinematics, are known to impact the PID efficiencies. For this reason, the data is separated in bins of: data-taking period, magnet polarity, track momentum, and track pseudorapidity. The efficiency is also expected to depend on the amount of particles in the event, known as the occupancy. However, this quantity is known to be imperfectly modelled in the simulation, as discussed in Section 6.4. For this reason, PID efficiencies are integrated over the occupancy. The systematic uncertainty induced by the modelling of the occupancy is evaluated in Section 7.4. The efficiency of a given PID cut, in a bin b, is given by:

$$\varepsilon(\text{cut}) = \frac{\sum\limits_{b,\,\text{pass cut}} sWeight}{\sum\limits_{b} sWeight}\,. \tag{6.9}$$

The binning in momentum and pseudorapidity is optimised using a procedure adapted from Ref. [192]. First, the $(p,\ \eta)$ space is divided into 100×10 bins of approximately equal populations, and efficiencies are computed. Then, the efficiencies are projected onto the momentum and pseudorapidity separately. In each of the two projections, two adjacent bins are merged if their corresponding efficiency values are compatible within a certain threshold. For the R_K analysis, the threshold is chosen to be at $2.5\,\sigma$. The systematic uncertainty induced by this choice is evaluated in Section 7.8, and found to be small.

The projections on momentum and pseudorapidity of the kaon and muon efficiencies, for 2018 MagUp data-taking conditions, are shown in Figure 6.1. Their equivalents for other conditions are similar, and are therefore omitted.

6.2.2 PID efficiency of electrons

The performance of the $\texttt{DLL}_e > 3$ requirement applied to electrons is evaluated using $B^+ \to K^+ J/\psi(e^+e^-)$ data. This calibration sample has additional selection cuts that improve the signal purity, and so it is not identical to the one used throughout the rest of the analysis. In particular, the tag electron is required to have high p_{T} (above 1.5 GeV) and to be very electron-like ($\texttt{DLL}_e > 5$). In particular for Run 2.2 samples, the cut on electron χ^2_{IP} is slightly tighter in the PID calibration samples than in the data used throughout the analysis. This leads to a small systematic effect, evaluated in Section 7.8.

The electron PID efficiency tables are obtained separately for each data-taking period. They are binned in three dimensions: momentum (p), pseudorapidity (η), and whether the probe has a bremsstrahlung photon associated with it or not (`hasBrem`). The two magnet polarities are not treated separately because their corresponding electron PID efficiencies are compatible within statistical uncertainty.

The electron PID calibration samples have higher levels of combinatorial background compared to their kaon and muon counterparts. In addition, the poor resolution of electrons causes the shape of the signal to be correlated with the variables in which the PID efficiency is binned. The *sWeight* method is only valid when considering quantities that are not correlated with the variable used to obtain the weights. Because of this, unbiased electron PID efficiencies cannot be obtained through the *sWeight* method. Therefore, the *sWeight*s are only used to determine the optimal binning in p and η, based on the algorithm outlined in Section 6.2.1. The efficiencies are instead determined through fits to the distribution of the J/ψ-constrained

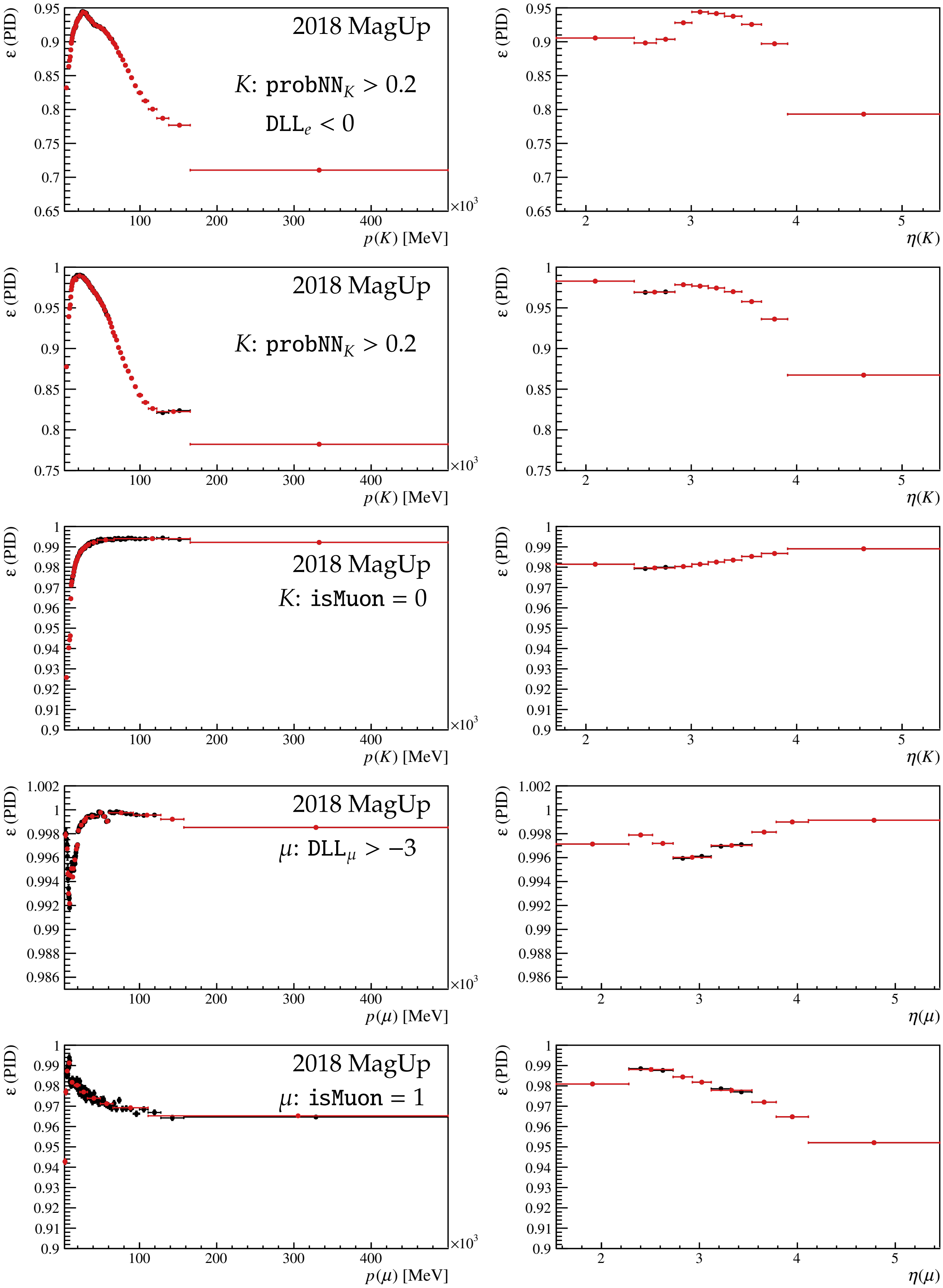

Figure 6.1: Efficiencies of the PID cuts applied to kaons and muons, projected onto particle momentum (left) and pseudorapidity (right). The red bins show the effect of the binning optimisation algorithm on the black bins. This algorithm is used to obtain the efficiencies used throughout the measurement, and is described in the main body.

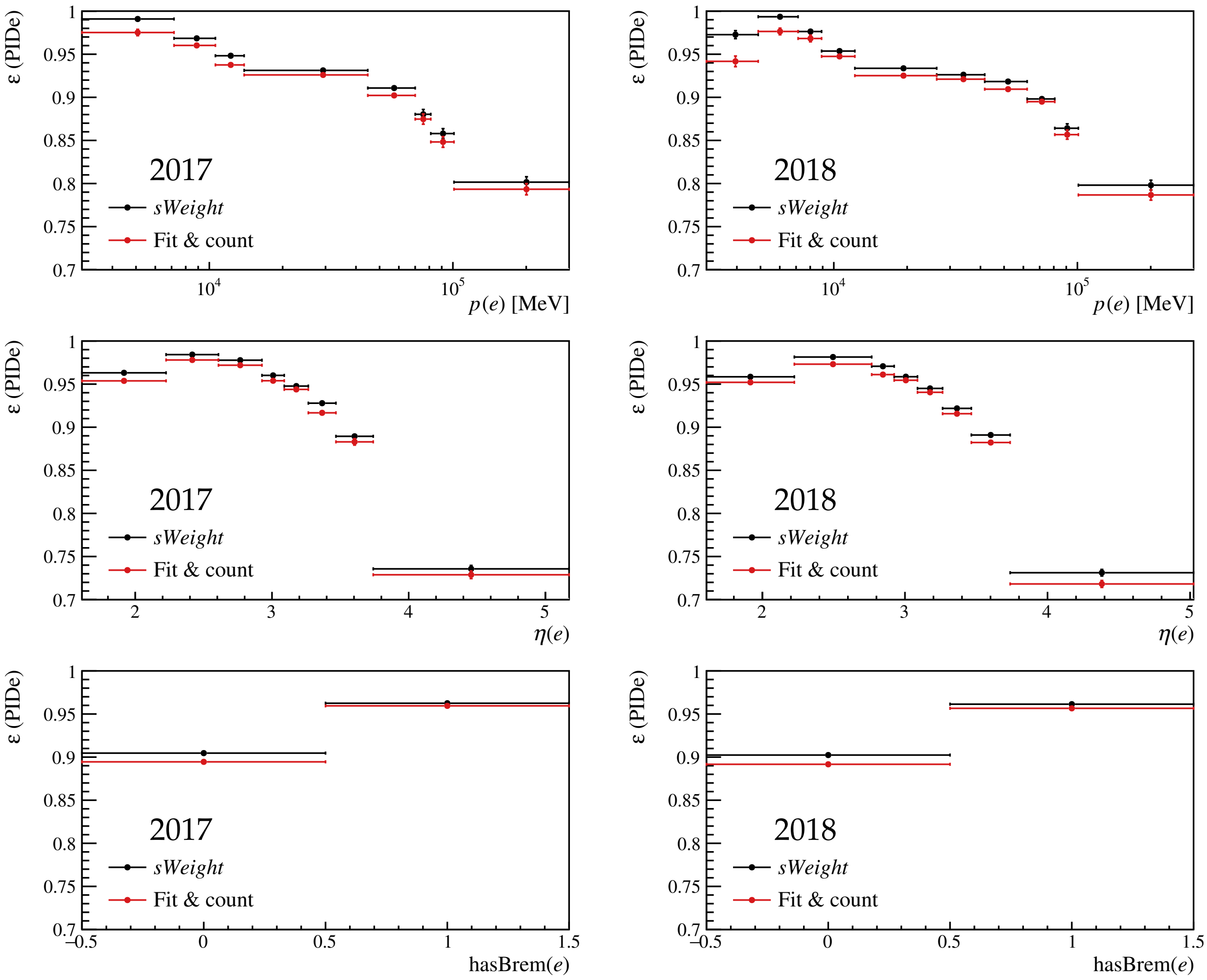

Figure 6.2: Electron PID performance in 2017 (left) and 2018 (right) data-taking conditions, projected onto particle momentum, pseudorapidity, and `hasBrem`. The values used are the ones obtained from the fit & count technique (red), with the *sWeight* results (black) shown for comparison.

$K^+e^+e^-$ invariant mass in each kinematic bin, before and after the electron PID cut. This is referred to as the "fit & count" method, and its result is presented in Figure 6.2. It can be seen that using the *sWeight* technique to extract electron efficiencies would lead to overestimation by a few percent.

6.2.3 Combination of PID efficiencies

The PID weight used to compute efficiencies is the product of the PID efficiencies of the three final-state particles:

$$w_{\mathrm{PID}} = \varepsilon_{\mathrm{PID}}(K^+) \cdot \varepsilon_{\mathrm{PID}}(\ell^+) \cdot \varepsilon_{\mathrm{PID}}(\ell^-)\,. \tag{6.10}$$

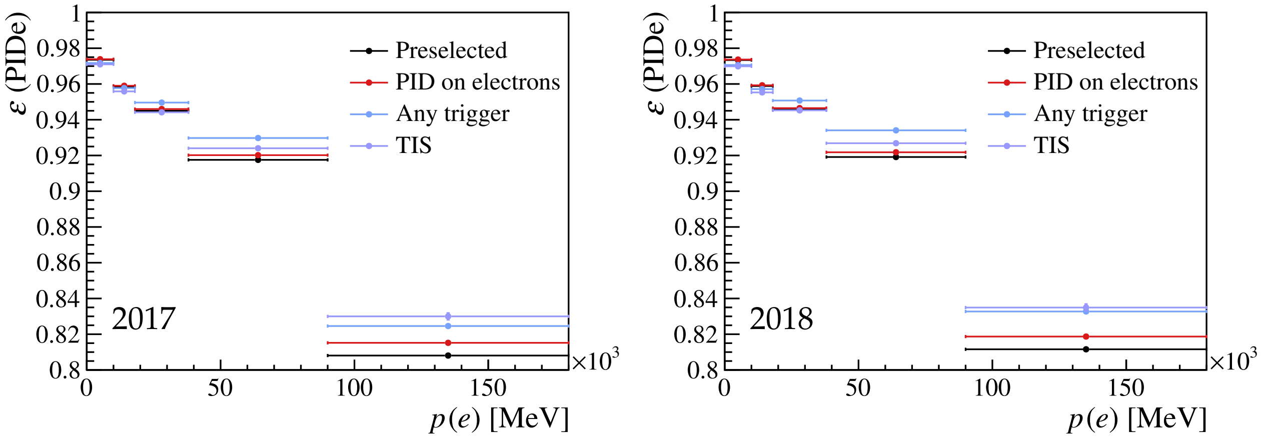

Figure 6.3: Efficiencies of the PID cut applied to electrons in 2017 (left) and 2018 (right) simulated control samples, projected onto particle momentum and evaluated on all preselected events (black), events where both electrons satisfy the PID requirements (red), events that fire any trigger (blue), and events where the trigger fires independently of the signal; this is labelled as TIS (purple).

Equation (6.10) assumes that the three efficiencies are independent of each other. This is checked by choosing one particle at a time, and recomputing its PID efficiency based only on events where the other two particles pass their PID requirements. If there are no correlations, this leads to results that are compatible with the ones obtained by not placing any cuts on the other two particles. It is also checked whether there are any effects induced by the trigger selection. The kaon and muon efficiencies are found to be unbiased. However, a trigger bias is found in the electron PID, as shown in Figure 6.3. The fact that the efficiencies with and without cuts on the other electron agree with each other shows that the PID efficiencies of the two electrons are independent. However, they disagree with the results obtained by applying trigger requirements. This indicates that a trigger bias is present, and so a systematic uncertainty is assigned by means described in Section 7.8.

6.3 Trigger calibration

The simulated efficiency of the L0 trigger is known to be at odds with the real performance. The trigger has to accept or reject events at a high rate, and so it uses information that's read quickly by the detector. This information is simulated imperfectly, hence the performance of the trigger is expected to disagree between data and simulation. Imperfections are also caused by variations in running conditions: in any given year during which LHCb took data, the configuration of the trigger changes several times. These configurations apply different cuts to the variables used by the `L0`, so some setups are by construction more efficient than others. However, the simulation of any particular year and magnet polarity only uses one configuration, and that leads to imperfect modelling of events close to trigger thresholds. It is only during the year 2018 that the trigger conditions were kept constant.

The simulated trigger performance is corrected by assigning weights to simulated events. These weights are determined based on how the efficiency changes in data and simulation as a function of the variables relevant to the detector response, such as the muon p_{T} for the `L0Muon` trigger. This information is readily accessible in simulation, where events are stored regardless of the output of the trigger. However, data that do not pass the trigger are not stored and reconstructed, and so the efficiency of any trigger selection has to be evaluated relative to another set of trigger requirements. This is done via what is referred to as the "tag & probe" method, where the efficiency of the "probe" trigger selection is measured after applying the "tag" requirements. The tags are validated by comparing their results with the true simulated performance, and then used as proxies of the probe efficiency. Weights are then computed and assigned to simulated events, based on how the efficiency of the probe differs between data and simulation. The following subsections cover the specifics of how each of the four triggers used in the R_K analysis is calibrated.

6.3.1 Calibration of the `L0Muon` trigger

The probability of each muon track to fire the `L0Muon` trigger is estimated in simulated and data $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ samples through the use of three tags. Events from all tags are required to have passed the preselection and HLT requirements. The three tags, and the additional requirements that distinguish them, are:

- μTIS tag: events where the `L0Muon` fires independently of the candidate;
- μ tag: events where `L0Muon` is triggered by the other muon in the candidate; and
- K tag: events where the `L0Hadron` fires on the kaon.

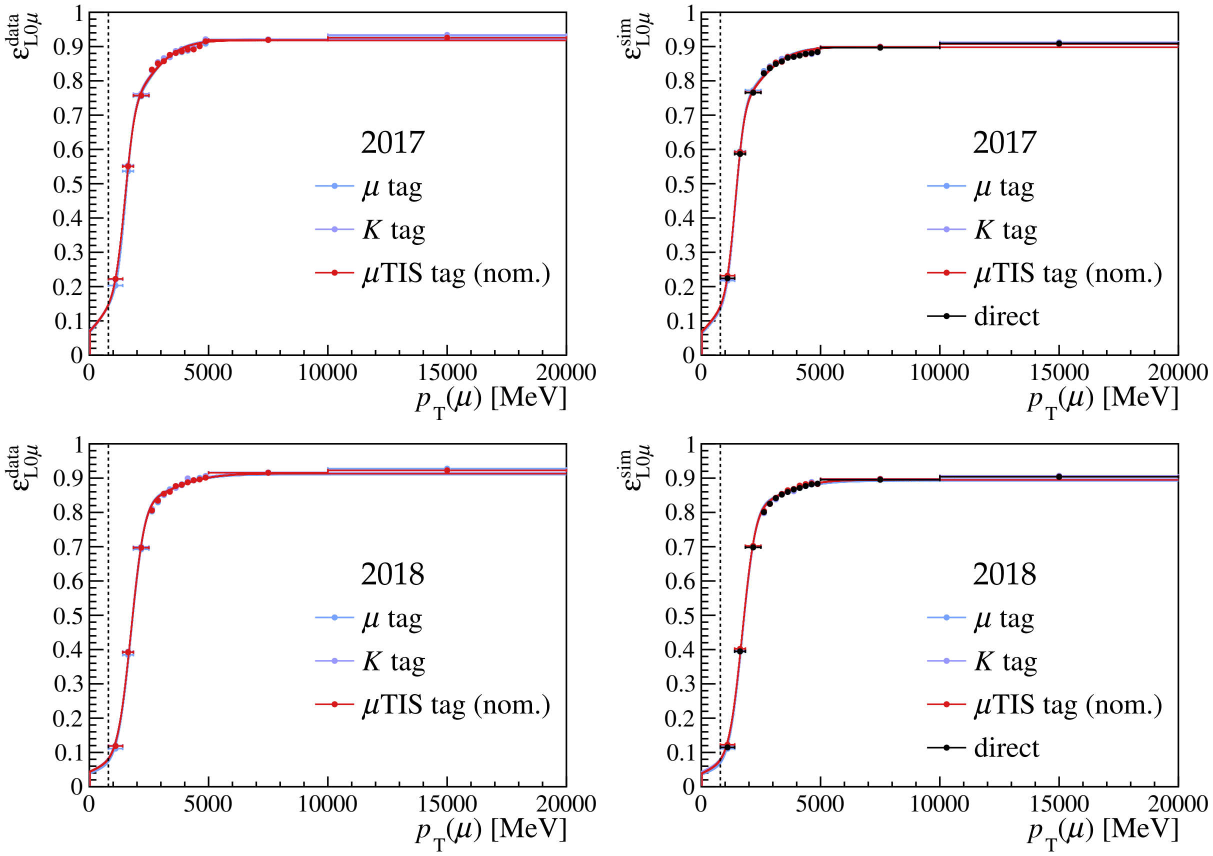

Figure 6.4: Efficiency with which a muon from data (left) and simulation (right) fires the `L0Muon` trigger, as a function of its reconstructed transverse momentum. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.11). The tag labelled as "nom." is the one used to estimate the nominal efficiencies, the other tags being used to evaluate systematic uncertainties. The dashed line corresponds to the $p_{\text{T}} > 800$ MeV fiducial cut applied to all muons. The fit does not extend below this threshold.

The μTIS tag is the one that has the highest statistics and hence the best precision. For this reason, the nominal efficiency estimates come from the μTIS tag, whilst the other two tags are used to evaluated the systematic effect induced by any biases present in the tags.

The `L0Muon` line makes a requirement on the highest transverse momentum among the muon tracks in the event. Therefore, the `L0Muon` efficiency ($\varepsilon_{\text{L0}\mu}$) is a step function in this variable. However, this estimate of p_{T} is only used at the trigger level, since the full reconstruction makes an improved measurement of the event kinematics. In addition, the trigger-level tracks are not matched with the reconstructed candidates, and there is no guarantee that a given particle has the highest p_{T} in the event at both trigger-level and reconstruction-level. As a result, the `L0Muon` efficiency is a smeared function of fully-reconstructed transverse momentum. This is shown by the bins in Figure 6.4, for both data and simulation samples.

The latter also depicts the efficiency obtained without the use of any tag, demonstrating that all three tags are unbiased and in good agreement.

To better interpolate between the bins used to correct the `L0Muon` efficiency, a function is used to model the performance of each tag in data and simulation. The chosen model is the sum of two sigmoids:

$$\varepsilon_{\mathrm{L0}\mu}(p_{\mathrm{T}}) = N\left[\left(1+\operatorname{erf}\left(\frac{p_{\mathrm{T}}-t}{\sigma_1\sqrt{2}}\right)\right)+f\left(1+\operatorname{erf}\left(\frac{p_{\mathrm{T}}-t}{\sigma_2\sqrt{2}}\right)\right)\right]. \tag{6.11}$$

In the above expression, t is the turn-on point of each sigmoid, and $\sigma_{1,2}$ are their widths. The normalisation is dictated by N and f. The muon p_{T} resolution is better close to the beam pipe, where the gas electron multiplier detectors are used, hence the use of two sigmoids of different widths. The parameters that lead to the best description of each tag are found through χ^2 minimisation. The resulting functions are depicted in Figure 6.4 by solid lines of the same colour as the bins they model.

The functions obtained from the fit lead to the determination of the `L0Muon` efficiency, $\varepsilon_{\mathrm{L0}\mu}(\mu^{\pm})$, of a muon with a particular p_{T}. Since each event contains two muons, the efficiency on an event-by-event basis is equal to $\varepsilon_{\mathrm{L0}\mu}(\mu^+) + \varepsilon_{\mathrm{L0}\mu}(\mu^-) - \varepsilon_{\mathrm{L0}\mu}(\mu^+)\varepsilon_{\mathrm{L0}\mu}(\mu^-)$. This leads to the following expression for the weights assigned to each simulated muon-channel event to correct the performance of the `L0Muon` trigger:

$$w_{\mu\mathrm{TOS}} = \frac{\varepsilon^{\mathrm{data}}_{\mathrm{L0}\mu}(\mu^+) + \varepsilon^{\mathrm{data}}_{\mathrm{L0}\mu}(\mu^-) - \varepsilon^{\mathrm{data}}_{\mathrm{L0}\mu}(\mu^+)\varepsilon^{\mathrm{data}}_{\mathrm{L0}\mu}(\mu^-)}{\varepsilon^{\mathrm{sim}}_{\mathrm{L0}\mu}(\mu^+) + \varepsilon^{\mathrm{sim}}_{\mathrm{L0}\mu}(\mu^-) - \varepsilon^{\mathrm{sim}}_{\mathrm{L0}\mu}(\mu^+)\varepsilon^{\mathrm{sim}}_{\mathrm{L0}\mu}(\mu^-)}. \tag{6.12}$$

The above equation assumes that the performances of the two muons in a given candidate event are not correlated. This assumption allows the two muon efficiencies to be factorised. Its validity is verified by studying the `L0Muon` performance of muons in simulated $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ events, as a function of whether the other muon also fires `L0Muon` or not. The result of this study is presented in Figure 6.5, where it can be seen that the efficiency of a given muon in the event does not depend on the other's performance.

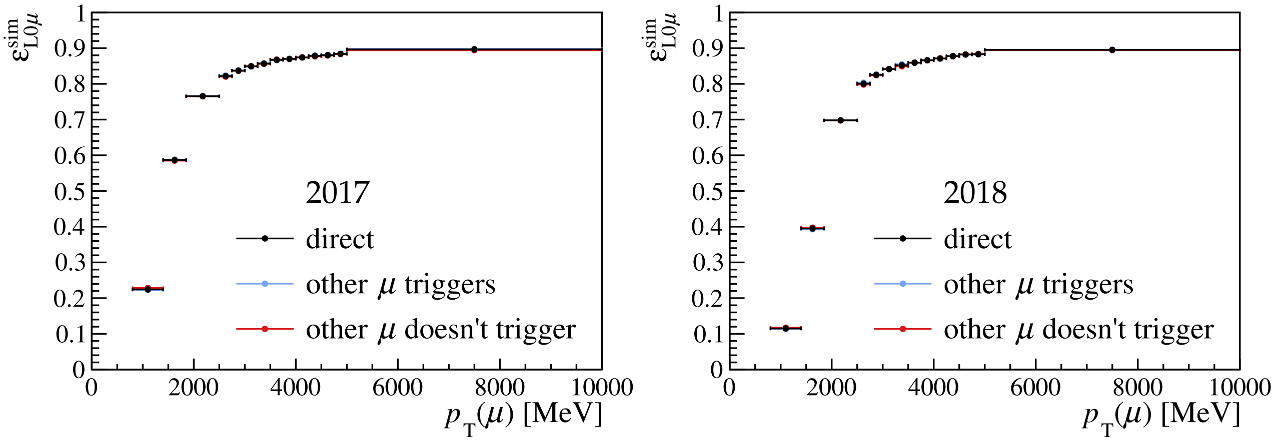

Figure 6.5: Efficiency with which a simulated muon fires the `L0Muon` trigger, calculated separately for candidates where: the other muon in the event also fires the `L0Muon` line (blue), the other muon does not fire `L0Muon` (red), and where no requirement is placed on the other muon (black).

6.3.2 Calibration of the `L0Electron` trigger

The `L0Electron` efficiency of electrons from $B^+ \to K^+ J/\psi(e^+e^-)$ data and simulation is estimated through the use of two tags. Events from both tags are required to have passed the preselection and HLT requirements. The two tags are:

- eTIS tag: events where the `L0Electron` line fires independently of the candidate; and
- K tag: events where the `L0Hadron` line fires on the kaon.

The nominal values come from the eTIS tag, because it has higher statistics. The K tag is used to evaluate the systematic uncertainty due to any tag bias, as described in Section 7.6.

Similarly to how the `L0Muon` line bases its decision on the transverse momentum of the particle, `L0Electron` makes use of the transverse energy deposited in the ECAL, $E_{\rm T}^{\rm L0}$. As such, the `L0Electron` efficiency of an electron ($\varepsilon_{{\rm L0}e}$) is a smeared step function of the electron's reconstructed $E_{\rm T}$. Given that cells in the three ECAL regions have different dimensions and hence different $E_{\rm T}$ resolutions, the amount of smearing depends on the ECAL region in which the cluster is located. To account for this, the `L0Electron` efficiencies are estimated separately for the inner, middle, and outer regions. The efficiency in the inner-most ECAL region is presented in Figure 6.6, whilst the others are shown in Appendix B.1.

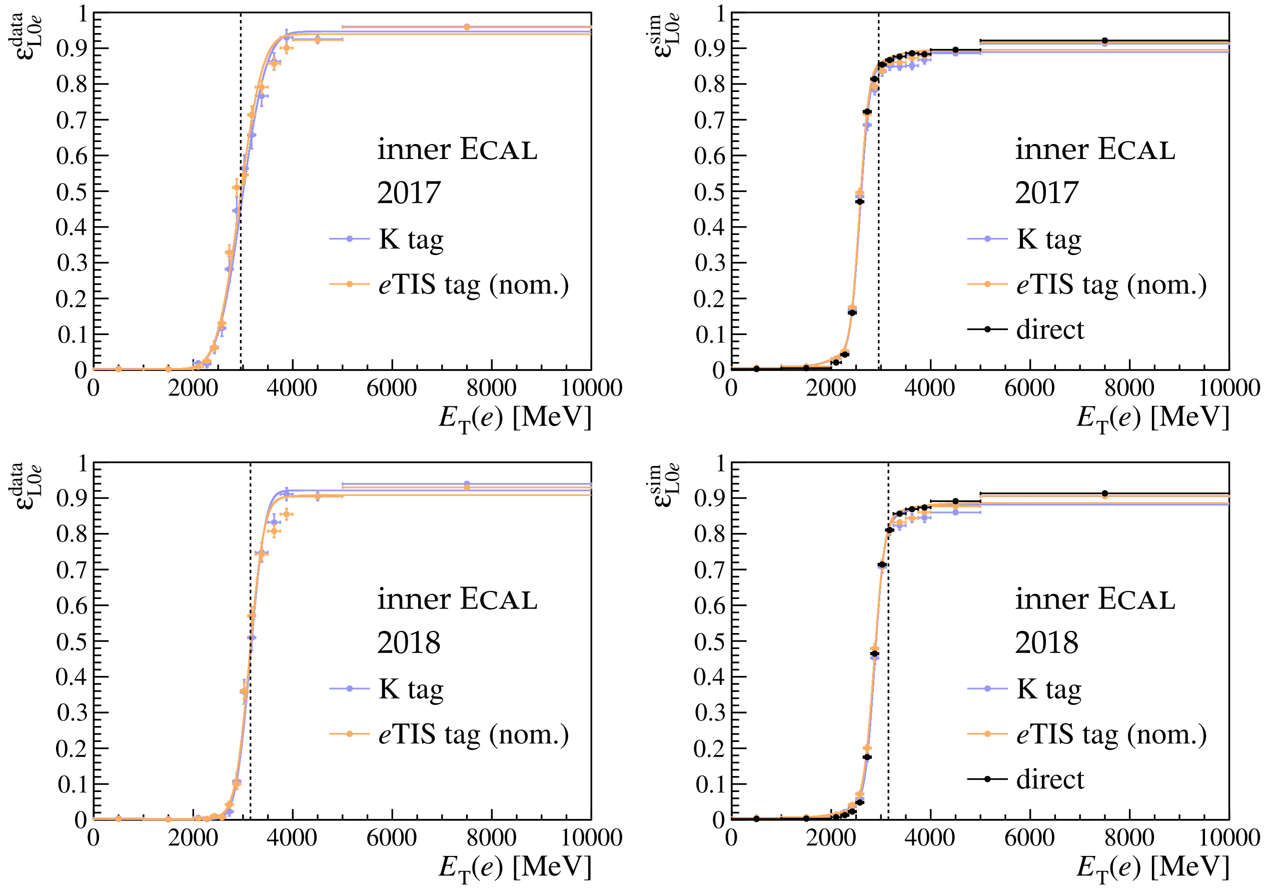

Figure 6.6: Efficiency with which an electron traversing the inner ECAL region in data (left) and simulation (right) fires the `L0Electron` trigger, as a function of the reconstructed transverse energy deposited in the ECAL. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.13). The tag labelled as “nom.” is the one used to estimate the nominal efficiencies, the other tag being used to evaluate systematic uncertainties. The dashed line corresponds to the fiducial cut applied to all electrons that fire `L0Electron`, in order to improve the level of agreement between selected data and simulated events.

Like for the `L0Muon` line, the `L0Electron` efficiencies are modelled by the sum of two sigmoids:

$$\varepsilon_{\mathrm{L0}e}(E_{\mathrm{T}}) = N\left[\left(1+\mathrm{erf}\left(\frac{E_{\mathrm{T}}-t}{\sigma_1\sqrt{2}}\right)\right) + f\left(1+\mathrm{erf}\left(\frac{E_{\mathrm{T}}-t}{\sigma_2\sqrt{2}}\right)\right) + a\right]. \tag{6.13}$$

Equation (6.13) contains an additional parameter (a) that accounts for noise in the ECAL. The best-fit functions are again obtained through χ^2 minimisation, and are depicted by solid lines in Figure 6.6. Following the same line of reasoning employed when calibrating the `L0Muon` trigger, correction weights are computed on an event-by-event basis, using the best-fit functions. These weights are given by:

$$w_{e\mathrm{TOS}} = \frac{\varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^+) + \varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^-) - \varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^+)\varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^-)}{\varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^+) + \varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^-) - \varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^+)\varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^-)}\,. \tag{6.14}$$

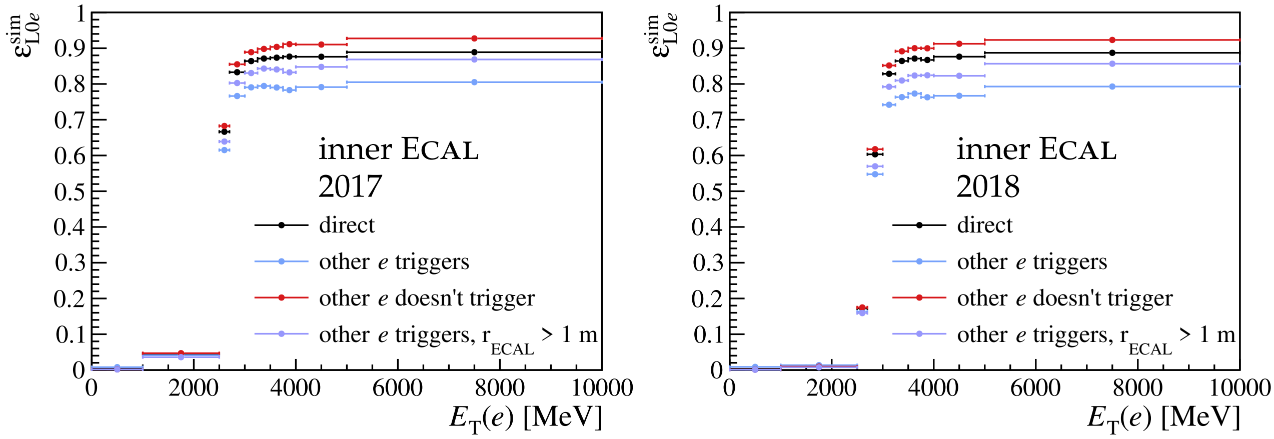

Figure 6.7: Efficiency with which a simulated electron fires `L0Electron`, calculated separately for candidates where: the other electron in the event also fires the `L0Electron` line (blue), the other electron fires `L0Electron` and is well separated from the probed electron (purple), the other electron does not fire `L0Electron` (red), and where no requirement is placed on the other electron (black).

This expression is functionally identical to the one defined in Equation (6.12). As a consequence, it too assumes the efficiencies of the two leptons can be factorised. Like with the muons, this assumption is tested by studying the `L0Electron` performance of simulated electrons from $B^+ \to K^+ J/\psi(e^+e^-)$ events, as a function of whether the other electron in the reconstructed candidate also fires `L0Electron`. Although the efficiencies of the muons were found to be factorisable, Figure 6.7 shows that the trigger performance of electrons depends on whether the other electron in the event also fires `L0Electron` or not. This bias is found to be smaller when the two electron clusters in the ECAL are separated by more than 1 m. If that is not the case, *i.e.* if there is significant overlap between the two clusters, the ECAL will only save the one with higher $E_{\mathrm{T}}^{\mathrm{L0}}$ and discard the other. This explains why candidates where both electrons fire `L0Electron` are not as efficient as expected. A systematic uncertainty is assigned to account for this dependence of the `L0Electron` efficiency on the distance between the two electron ECAL clusters. This systematic effect is estimated in Section 7.6.

6.3.3 Calibration of the `L0Hadron` trigger

The performance of the `L0Hadron` line can be estimated from both $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ and $B^+ \to K^+ J/\psi(e^+e^-)$ samples. As a result, several tags are used on events that have passed the preselection and HLT requirements. The tags are:

- μTIS tag: $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ events where `L0Muon` fires independently of the candidate;
- eTIS tag: $B^+ \to K^+ J/\psi(e^+e^-)$ events where `L0Electron` fires independently of the candidate;
- μ tag: $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ events where the other muon in the event fires `L0Muon`; and
- e tag: $B^+ \to K^+ J/\psi(e^+e^-)$ events where the other electron in the event fires `L0Electron`.

Out of the four, the μ tag is found to have the best statistical precision, and is therefore used to obtain the nominal efficiency estimates. The other tags are used to evaluate the systematic effect induced by tag biases.

The `L0Hadron` trigger decision is based on the transverse energy deposited in the HCAL. Hence, the performance of this trigger line is parametrised in terms of the reconstructed E_T deposited in the HCAL. Like with `L0Electron`, this dependence has the shape of a smeared step function, and the amount of smearing depends on the HCAL region in which the cluster is located. This is accounted for by computing the `L0Hadron` efficiencies separately for each of the two HCAL regions. The result for the inner region is presented in Figure 6.8, with its outer-region counterpart shown in Appendix B.2.

To improve the accuracy and precision of the `L0Hadron` calibration weights, the efficiency is modelled by a sigmoid with normalisation N, turn-on point t, and width σ:

$$\varepsilon_{\mathrm{L0}h}(E_\mathrm{T}) = N\left[\left(1 + \mathrm{erf}\left(\frac{E_\mathrm{T} - t}{\sigma\sqrt{2}}\right)\right) + a\right]. \tag{6.15}$$

As with `L0Electron`, the constant factor a is introduced to account for noise in the calorimeter. Similarly to the other trigger lines, χ^2 minimisation is employed to obtain the best-fit parameters for each tag. The resulting shapes are depicted as solid lines in Figure 6.8. These curves are used to determine the `L0Hadron` efficiencies, in data and simulation, as a function of hadron E_T. Taking the ratio leads to weights that correct the `L0Hadron` performance:

$$w_{h\mathrm{TOS}} = \frac{\varepsilon^{\mathrm{data}}_{\mathrm{L0}h}(K^+)}{\varepsilon^{\mathrm{sim}}_{\mathrm{L0}h}(K^+)}. \tag{6.16}$$

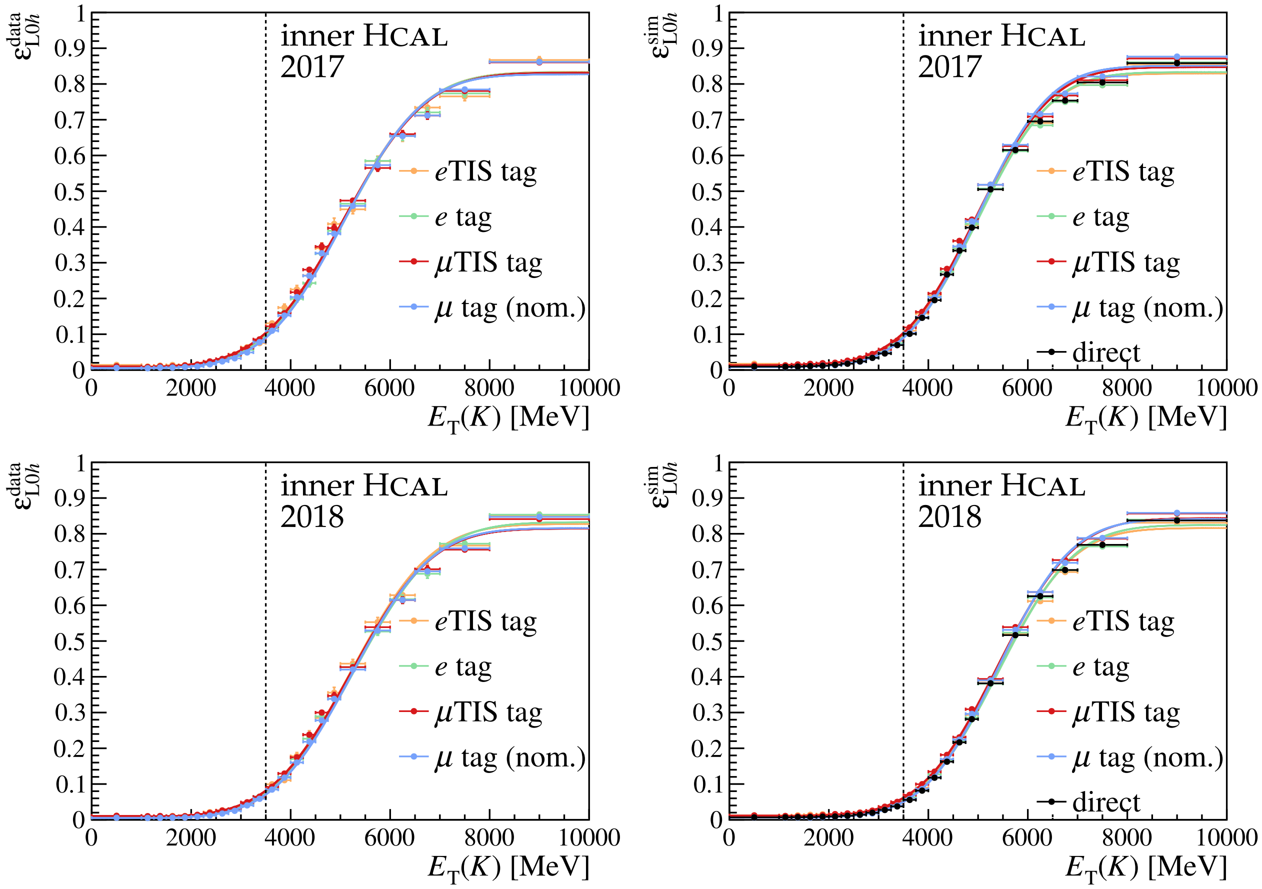

Figure 6.8: Efficiency with which a kaon traversing the inner HCAL region in data (left) and simulation (right) fires the `L0Hadron` trigger, as a function of the reconstructed transverse energy deposited in the HCAL. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.15). The tag labelled as "nom." is the one used to estimate the nominal efficiencies, the other tag being used to evaluate systematic uncertainties. The dashed line corresponds to the $E_{\mathrm{T}} > 3.5$ GeV fiducial cut applied to all kaons that fire `L0Hadron`, in order to improve the level of agreement between selected data and simulated events.

Since hTOS! is defined as an exclusive trigger category, its performance depends on that of eTOS. However, the inclusive hTOS category is used to calibrate the trigger, owing to its larger statistics. This means that an additional factor has to be introduced to account for the different eTOS performance in data and simulation. The final expression for the weights that correct the trigger efficiency of simulated hTOS! events is then:

$$w_{h\mathrm{TOS!}} = \frac{\varepsilon_{\mathrm{L0}h}^{\mathrm{data}}(K^+)}{\varepsilon_{\mathrm{L0}h}^{\mathrm{sim}}(K^+)} \cdot \frac{\left(1-\varepsilon_{\mathrm{L0}e}^{\mathrm{data}}(e^+)\right)\cdot\left(1-\varepsilon_{\mathrm{L0}e}^{\mathrm{data}}(e^-)\right)}{\left(1-\varepsilon_{\mathrm{L0}e}^{\mathrm{sim}}(e^+)\right)\cdot\left(1-\varepsilon_{\mathrm{L0}e}^{\mathrm{sim}}(e^-)\right)}\,. \tag{6.17}$$

6.3.4 Calibration of the `L0TIS` trigger

In general, when the `L0` triggers independently of the signal, it fires on the other beauty hadron in the event. Since the $b\bar{b}$ system is predominantly produced at low p_{T}, the transverse momenta of the two B mesons are correlated. As a result, the performance of `L0TIS` is highly dependent on the p_{T} of the signal B. For this reason, the previous R_K measurement corrected the efficiency of `L0TIS` based on its dependency on $p_{\mathrm{T}}(B^+)$. For Run 2.2, the procedure is improved by taking into account the fact that some events are classified as TIS, but in reality it is the signal that causes the trigger to fire. This happens, for example, when the `L0` fires on photons emitted by signal electrons as bremsstrahlung radiation. These are referred to as "spurious TIS" events, and their trigger performance is qualitatively different to that of events where the trigger genuinely fires independently of the signal. In practice, `L0TIS` requires at least one of four `L0` lines to fire independently of the signal. These are: `L0Electron`, `L0Photon`, `L0Hadron`, and `L0Muon`. The former two are expected to be impacted by spurious TIS events, and so are calibrated separately from the latter two. The two lines with high contributions from spurious TIS are collectively referred to as $e\gamma$TIS, and the other two as $h\mu$TIS.

The efficiencies of $e\gamma$TIS and $h\mu$TIS in $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ data and simulation are shown in Figure 6.9 and in Figure 6.10, respectively. The $e\gamma$TIS efficiency is computed as a function of the maximum p_{T} of the two leptons in the event, because the number of spurious TIS events depends on the kinematics of the signal leptons. The tag used to obtain the $e\gamma$TIS efficiency, labelled "μTIS tag", requires the `L0Muon` line to have fired independently of the signal. This tag has a small bias, as evidenced by the direct efficiency extracted from $B^+ \to K^+ J/\psi(e^+e^-)$ simulation. The $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ counterpart is significantly lower, since muons don't emit as much bremsstrahlung radiation as electrons. Given that spurious TIS is expected to have a negligible impact on $h\mu$TIS, its efficiency is parametrised in the p_{T} of the B^+. The tag chosen to obtain the $h\mu$TIS efficiency, e tag, requires at least one of the electrons in $B^+ \to K^+ J/\psi(e^+e^-)$ preselected samples to fire `L0Electron`. The alternative μ tag requires muons in $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ preselected samples to fire `L0Muon`, and is used to estimate the systematic effect of tag biases. The expression for the weights changes between new and

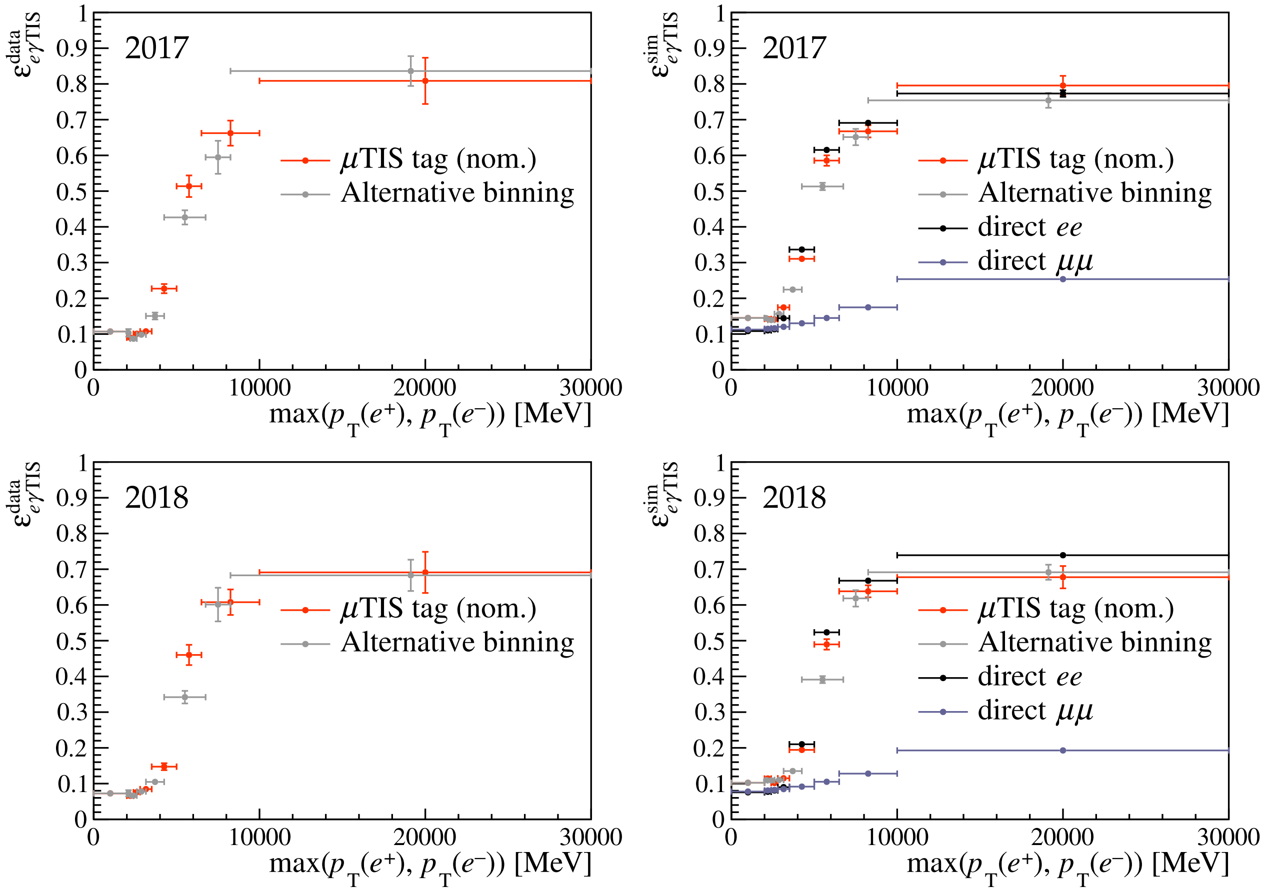

Figure 6.9: Efficiency with which the `L0Electron` and `L0Photon` lines fire independently of the signal, in data (left) and simulation (right) samples. These two lines are collectively referred to as $e\gamma$TIS, and are expected to be impacted by spurious TIS events. The alternative binning scheme is used to evaluate systematic uncertainties, as described in Section 7.6.

previous samples, as a result of the differences between the calibration histograms. In Run 1 and Run 2.1, the `L0TIS` efficiency is only parametrised in the p_{T} of the B^{+}, so the weight that corrects the `L0TIS` performance is:

$$w^{\text{prev}}_{\texttt{L0TIS}} = \frac{\varepsilon^{\text{data}}_{\texttt{L0TIS}}(B^{+})}{\varepsilon^{\text{sim}}_{\texttt{L0TIS}}(B^{+})} \, . \tag{6.18}$$

For Run 2.2 samples, Equation (6.18) is adjusted to take into account the different corrections schemes for $e\gamma$TIS and $h\mu$TIS:

$$w^{\text{Run 2.2}}_{\texttt{L0TIS}} = \frac{\varepsilon^{\text{L0TIS}}_{\text{data}}(e\gamma\text{TIS}) + \varepsilon^{\text{L0TIS}}_{\text{data}}(\text{h}\mu\text{TIS}) - \varepsilon^{\text{L0TIS}}_{\text{data}}(\text{e}\gamma\text{TIS}) \cdot \varepsilon^{\text{L0TIS}}_{\text{data}}(\text{h}\mu\text{TIS})}{\varepsilon^{\text{L0TIS}}_{\text{sim}}(e\gamma\text{TIS}) + \varepsilon^{\text{L0TIS}}_{\text{sim}}(\text{h}\mu\text{TIS}) - \varepsilon^{\text{L0TIS}}_{\text{sim}}(\text{e}\gamma\text{TIS}) \cdot \varepsilon^{\text{L0TIS}}_{\text{sim}}(\text{h}\mu\text{TIS})} \, . \tag{6.19}$$

Like hTOS!, TIS! is an exclusive category whose performance is corrected using the inclusive variant owing to the better statistics. For this reason, two additional factors are introduced

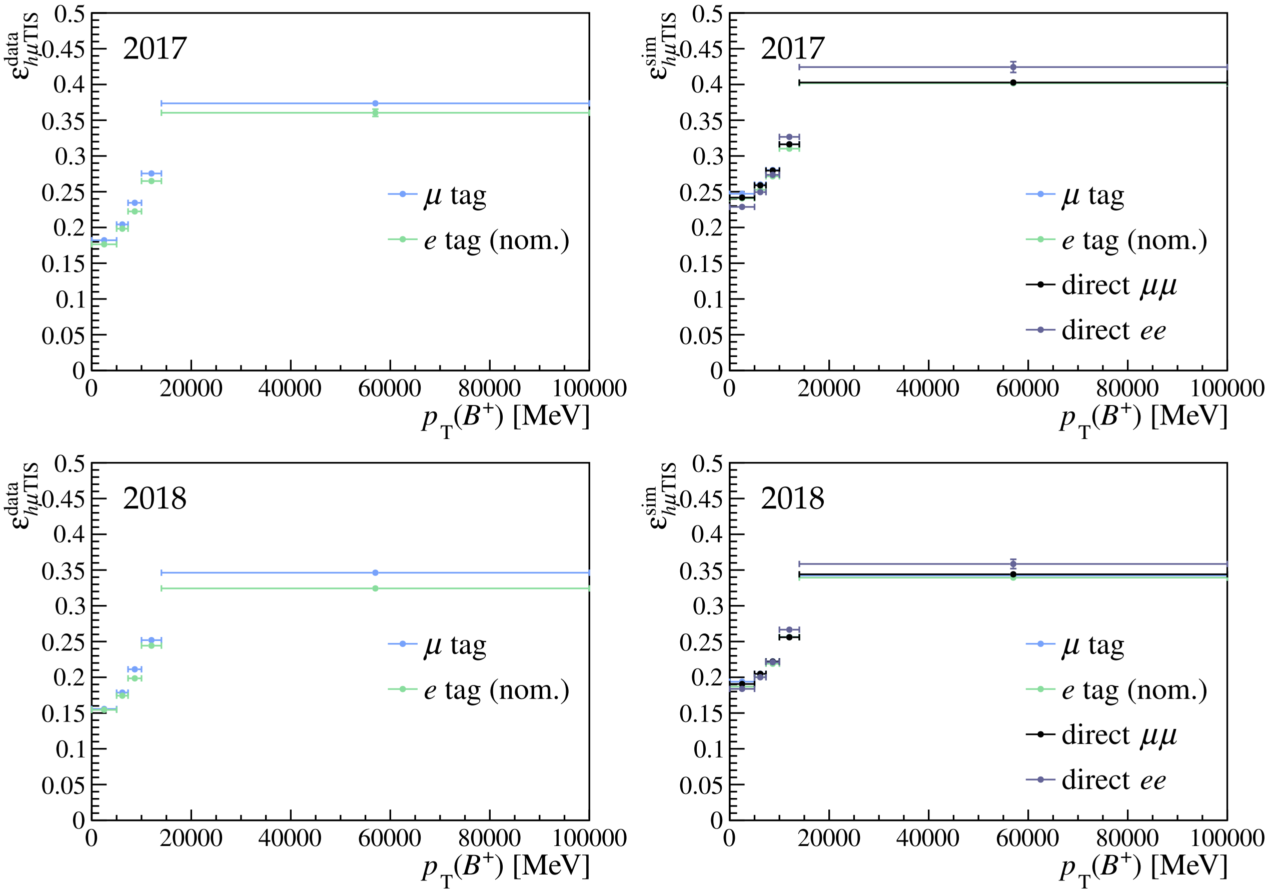

Figure 6.10: Efficiency with which the `L0Hadron` and `L0Muon` lines fire independently of the signal, in data (left) and simulation (right) samples. These two lines are collectively referred to as $h\mu$TIS, and are not expected to be impacted by fake TIS events.

to account for the different performances of the eTOS and hTOS! trigger strategies in data compared to simulation. This leads to the following expression for the weights applied to simulated TIS! events to correct their `L0` trigger performance:

$$w_{\mathrm{TIS!}} = w_{\mathtt{L0TIS}} \cdot \frac{\left(1-\varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^{+})\right)\cdot\left(1-\varepsilon^{\mathrm{data}}_{\mathrm{L0}e}(e^{-})\right)}{\left(1-\varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^{+})\right)\cdot\left(1-\varepsilon^{\mathrm{sim}}_{\mathrm{L0}e}(e^{-})\right)} \cdot \frac{1-\varepsilon^{\mathrm{data}}_{\mathrm{L0}h}(K^{+})}{1-\varepsilon^{\mathrm{sim}}_{\mathrm{L0}h}(K^{+})}. \tag{6.20}$$

6.3.5 Calibration of the HLT

The simulated description of the high-level trigger in Run 2 conditions is found to be in good agreement with the data. However, the Run 1 counterpart requires additional calibration via weights. These were obtained by Dr. Thibaud Humair, and are outside the scope of this thesis, since such weights are not required for Run 2.2 data-taking conditions.

6.4 The underlying event occupancy

The proton-proton collision environment at the LHC is difficult to model, due to the non-perturbative nature of the strong-interaction effects involved. As a consequence, there are limitations to how well the underlying events at LHCb can be described by simulation. One of these limitations is the improper modelling of the total number of particles in an event, known as the occupancy.

Different quantities correlated with the occupancy are used to assess how busy an event is. These are called occupancy proxies, and are known to be modelled improperly by simulation. A study was performed by Dr. Thibaud Humair to evaluate the impact of correcting occupancy proxies to match the data. It was found that making one proxy match the data increases the discrepancies found in other proxies. For this reason, no corrections to occupancy proxies are applied when evaluating efficiencies. This leads to a systematic uncertainty that is evaluated in Section 7.4.

6.5 Corrections to B^+ kinematics

Another known imperfection of the simulation lies in the kinematic variables with which the B^+ mesons are generated. The spectra of kinematic variables, such as transverse momentum and pseudorapidity, are not reproduced perfectly by the simulation. This happens at the stage where the B^+ are generated, so the p_T and η are different with respect to data both before and after the reconstruction takes place.

Additional discrepancies arise after reconstruction between the simulated and actual distributions of the χ^2_IP and χ^2_DV of the B^+. These variables are related to the quality of the reconstructed primary and decay vertices, respectively, and are thus sensitive to any tracking-related effect that is not perfectly modelled by the simulation. An example of this is the imperfect modelling of the underlying event occupancy, as mentioned in Section 6.4. The occupancy is correlated with the total number of tracks in the event, and the more particles there are, the more affected the tracking performance is. Another example is the imperfect modelling of the VELO, given its central role in tracking and vertexing.

The distributions of kinematic variables in data are obtained from the $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ μTOS samples obtained by means described in Section 5.2. The statistics are high, and the background is subtracted through the *sWeight* method. The resolution of the muons is also better than that of the electrons, and, as expanded upon in Section 7.6 and Section 7.8, the trigger and PID corrections to muon samples are under better control.

To correct the simulated B^+ kinematics, two sets of weights are computed after the calibration of the PID and trigger performances. The first set of weights address the imperfections in the generated two-dimensional distribution of $(p_{\mathrm{T}}(B^+),\ \eta(B^+))$. These are the $w_{\mathrm{kin}}^{\mathrm{gen}}$ weights introduced in Equation (6.7). The second set of weights correct the reconstructed $(p_{\mathrm{T}}(B^+),\ \eta(B^+))$ distribution, together with the spectra of $\chi^2_{\mathrm{IP}}(B^+)$ and $\chi^2_{\mathrm{DV}}(B^+)$. In Equation (6.7), these are denoted by $w_{\mathrm{kin}}^{\mathrm{rec}}$.

Both $w_{\mathrm{kin}}^{\mathrm{gen}}$ and $w_{\mathrm{kin}}^{\mathrm{rec}}$ are calculated through the following procedure:

1. data obtained via the *sWeight* method are used to populate histograms of kinematic variables; the histogram for the $(p_{\mathrm{T}}(B^+),\ \eta(B^+))$ distribution has 20 bins along the p_{T} axis and 10 bins along the η axis, whilst the χ^2_{IP} and χ^2_{DV} histograms used for $w_{\mathrm{kin}}^{\mathrm{rec}}$ have 150 bins each;

2. the previous step is repeated on simulated $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ μTOS generation- and reconstruction-level samples; the comparison between generation-level simulation and reconstructed data is possible thanks to the good muon resolution;

3. the histograms obtained in the first step are divided by the ones obtained in step 2., thus obtaining histograms of weights;

4. to account for correlations between weighted reconstruction-level variables, steps 1 to 3 are repeated three times to obtain the $w_{\mathrm{kin}}^{\mathrm{rec}}$ weights.

The effect of the kinematic correction weights on muon samples is exemplified in Figure 6.11. Each plot shows the simulated distribution, before and after the calibration of the kinematics, together with the spectrum in background-subtracted data. This includes both variables that are corrected directly, *e.g.* $p_{\mathrm{T}}(B^+)$ and $\eta(B^+)$, and other kinematic variables, *e.g.* the opening angle between the two signal leptons ($\alpha_{\ell^+\ell^-}$). It can be seen that the kinematic weights lead to near-perfect agreement between data and simulation. These plots correspond to 2018 data-taking conditions. Additional kinematic distributions, as well as equivalents from 2017 samples, are shown in Appendix C. The calibration is of the same quality in the two years.

Equivalent distributions in electron samples are presented in Figure 6.12. The calibration improves the agreement between the data and simulation, however there are a few residual discrepancies in some of the variables. This is because the weights are obtained from muon samples and applied to the electron channels, where the kinematics are not identical. The systematic effect induced by these residual differences is evaluated in Sections 7.2 and 7.7. They are found to be sub-dominant on R_K, thanks to cancellation in the double ratio. Additional distributions in 2018 samples, together with their 2017 counterparts, are presented in Appendix C. The calibration is of similar quality in the two years.

6.6 Momentum and q^2 resolution calibration

The resolution of the dielectron invariant mass, $m(e^+e^-)$, is found to be better in simulation than in data. If left uncorrected, the efficiency of any selection that cuts in q^2, or in any other derived variable such as $m(K^+e^+e^-)$, would be overestimated. To prevent this, the distribution of $m(e^+e^-)$ in simulated control-mode events is modified to match its equivalent from data. This process is called smearing. The calculation of efficiencies takes into account this correction by using the smeared variables to evaluate the performance of the q^2 and invariant-mass cuts.

The mismatch in resolution was already hinted at in Chapter 5, where templates for peaking components of the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ fits could not be used immediately to determine the resolution. Instead, the mean of the data peak has to be reparametrised in terms of a shift from its simulated value, and the width is a scaled version of the one in simulation:

$$\mu^{\mathrm{data}} = \mu^{\mathrm{sim}} + \Delta\mu\,, \qquad \sigma^{\mathrm{data}} = s_\sigma \cdot \sigma^{\mathrm{sim}}\,. \tag{6.21}$$

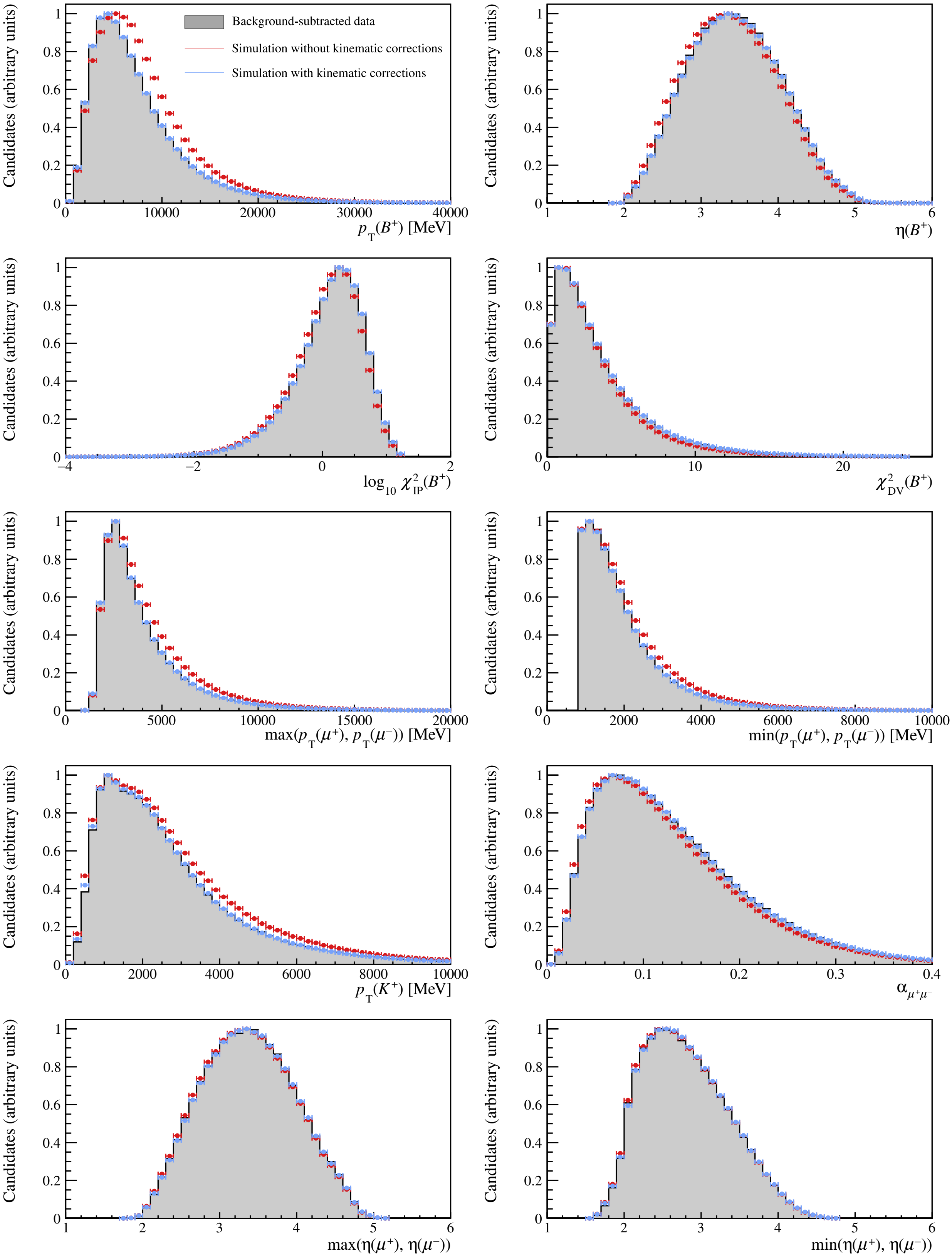

Figure 6.11: Distributions of kinematic variables in $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ 2018 μTOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively. The histograms are normalised based on the bin that takes the highest value.

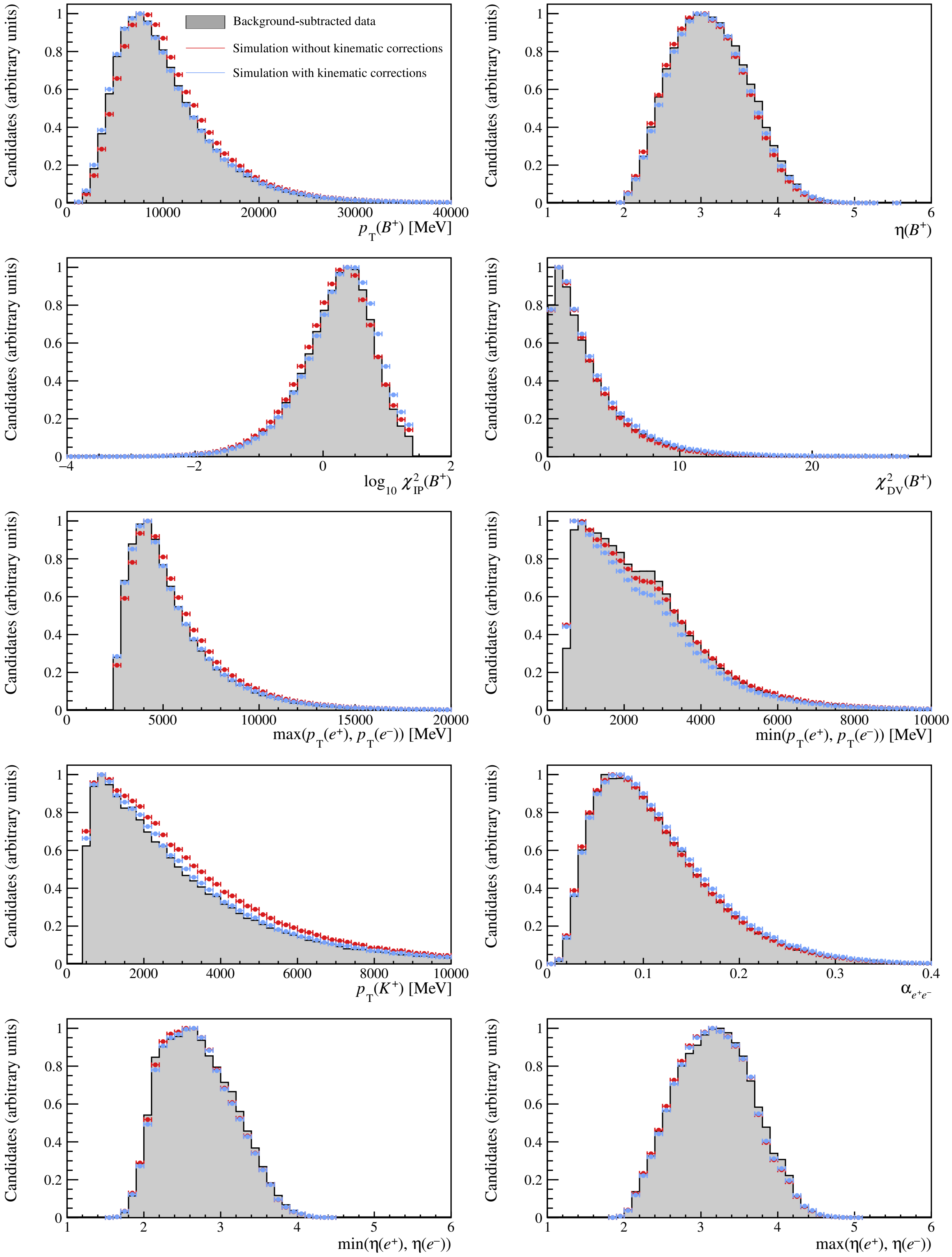

Figure 6.12: Distributions of kinematic variables in $B^+ \to K^+ J/\psi(e^+e^-)$ 2018 eTOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively. The histograms are normalised based on the bin that takes the highest value.

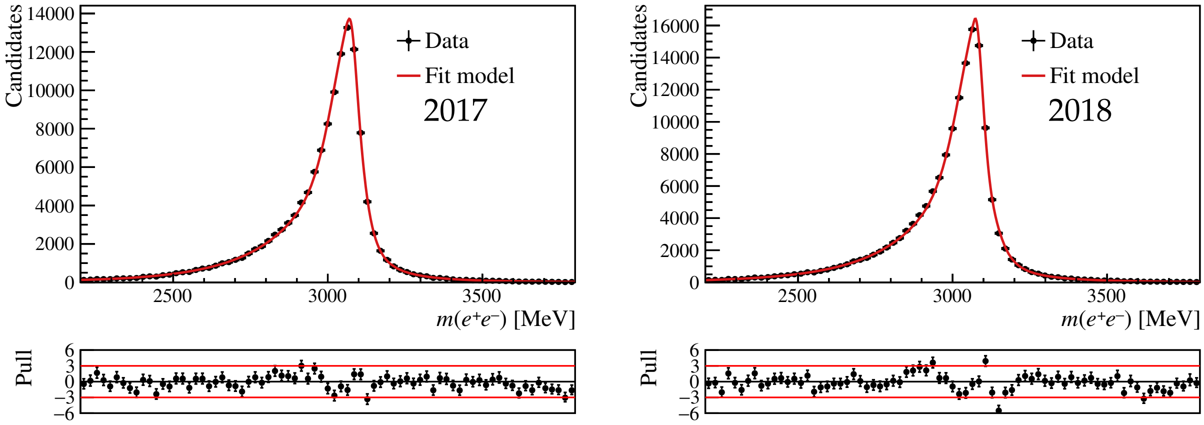

Figure 6.13: Fits to the distribution of $m(e^+e^-)$ in 2017 (left) and 2018 (right) $B^+ \to K^+J/\psi(e^+e^-)$ eTOS data. The photon categories are fitted separately, and then combined to obtain these figures.

The shift $\Delta\mu$ and the scale s_σ are used to perform the smearing. They are obtained from fits to the distribution of $m(e^+e^-)$ in $B^+ \to K^+J/\psi(e^+e^-)$ data and simulation. To minimise the impact of any potential q^2 dependency on the smearing factors, the $B^+ \to K^+J/\psi(e^+e^-)$ events are selected with $q^2 \in (3.0\ \text{GeV}^2,\ 15.0\ \text{GeV}^2)$. In addition, the impact of partially-reconstructed background events is reduced by tightening the selection on the J/ψ-constrained invariant mass to $m_{J/\psi}(K^+e^+e^-) \in (5.20\ \text{GeV},\ 5.68\ \text{GeV})$. These fits are performed separately in each data-taking period, electron trigger, and photon category. The latter division is motivated by the fact that bremsstrahlung radiation affects both the mean and the resolution of the mass peak. The simulated $m(e^+e^-)$ distributions are modelled by two CB distributions with opposite tails. The low-mass tail provides an accurate description of the radiative effects that impact this region of $m(e^+e^-)$. The high-mass tail is able to model events where the bremsstrahlung recovery algorithm overestimates the energy that has to be added to the electron. In general, both effects are suitably simulated, so the tail parameters are fixed to the values extracted from simulation. The exceptions are the upper tails in the 1γ and 2γ categories, which are allowed to float in the fit to data to improve the quality of the fit.

The results of the fits to eTOS data, in all three photon categories combined, are presented in Figure 6.13. The pulls between the data and the fit model are shown below each plot. Most of them are small, and there are no indications of any systematic deviations. This suggests that the fits are of good quality. The scales and shifts obtained from all fits to data are listed in Table 6.2. The values of s_σ, $\Delta\mu$, and μ^{sim} are averaged across the three electron triggers,

Table 6.2: Width scale factor and mean shift parameters extracted by fitting $m(e^+e^-)$ in $B^+ \to K^+ J/\psi(e^+e^-)$ data. The values are obtained separately for each trigger and photon category. The average among the three trigger categories is also listed. The error cited for the average corresponds to the standard deviation between the three trigger categories.

	0γ		**1γ**		**2γ**	
Trigger	s_σ	**$\Delta\mu$ [MeV]**	s_σ	**$\Delta\mu$ [MeV]**	s_σ	**$\Delta\mu$ [MeV]**
			Run 1			
*e*TOS	1.055 ± 0.008	1.1 ± 0.5	1.092 ± 0.007	4.8 ± 0.4	1.098 ± 0.012	9.6 ± 0.7
*h*TOS!	1.106 ± 0.030	2.6 ± 1.5	1.136 ± 0.021	3.7 ± 1.5	1.140 ± 0.032	9.1 ± 2.1
TIS!	1.108 ± 0.017	2.8 ± 0.9	1.089 ± 0.012	2.3 ± 1.0	1.103 ± 0.021	10.4 ± 1.5
Average	1.068 ± 0.023	1.5 ± 0.7	1.094 ± 0.012	4.3 ± 0.9	1.103 ± 0.012	9.7 ± 0.4
			Run 2.1			
*e*TOS	1.112 ± 0.007	-1.7 ± 0.3	1.135 ± 0.007	-8.1 ± 0.2	1.215 ± 0.004	-8.8 ± 0.3
*h*TOS!	1.178 ± 0.031	-1.3 ± 1.6	1.110 ± 0.021	-15.1 ± 1.3	1.290 ± 0.050	-18.3 ± 2.4
TIS!	1.143 ± 0.020	-0.8 ± 0.9	1.178 ± 0.015	-13.8 ± 1.0	1.183 ± 0.024	-15.7 ± 1.4
Average	1.118 ± 0.016	-1.6 ± 0.3	1.140 ± 0.018	-8.5 ± 1.7	1.215 ± 0.008	-9.4 ± 2.0
			2017			
*e*TOS	1.082 ± 0.008	-3.0 ± 0.4	1.125 ± 0.007	-5.3 ± 0.4	1.121 ± 0.014	-6.4 ± 0.6
*h*TOS!	1.077 ± 0.032	-1.7 ± 1.4	1.152 ± 0.022	-10.0 ± 1.4	1.161 ± 0.034	-14.0 ± 2.0
TIS!	1.093 ± 0.016	-1.6 ± 0.8	1.160 ± 0.016	-10.6 ± 1.0	1.117 ± 0.023	-13.4 ± 1.4
Average	1.084 ± 0.003	-2.7 ± 0.5	1.132 ± 0.013	-6.2 ± 2.0	1.125 ± 0.014	-8.0 ± 3.0
			2018			
*e*TOS	1.098 ± 0.007	-2.1 ± 0.3	1.113 ± 0.007	-5.0 ± 0.3	1.153 ± 0.013	-2.7 ± 0.6
*h*TOS!	1.134 ± 0.031	-2.1 ± 1.3	1.172 ± 0.017	-7.1 ± 1.2	1.127 ± 0.029	-11.6 ± 1.5
TIS!	1.104 ± 0.018	-1.9 ± 0.9	1.169 ± 0.014	-8.0 ± 0.9	1.192 ± 0.022	-10.0 ± 1.2
Average	1.101 ± 0.007	-2.1 ± 0.1	1.129 ± 0.026	-5.5 ± 1.0	1.158 ± 0.020	-5.0 ± 4.0

for each photon category individually. These averages, denoted $\overline{s_\sigma}$, $\overline{\Delta\mu}$, and $\overline{\mu^{\text{sim}}}$, are used to obtain the smeared simulated dielectron mass:

$$m^{\text{smeared}} = m^{\text{true}} + \overline{s_\sigma} \cdot (m - m^{\text{true}}) + \overline{\Delta\mu} + (1 - \overline{s_\sigma}) \cdot (\overline{\mu^{\text{sim}}} - m_{J/\psi}) \, . \tag{6.22}$$

The true and reconstructed dilepton invariant masses, m^{true} and m respectively, are computed for each simulated event, whereas $\overline{s_\sigma}$, $\overline{\Delta\mu}$, and $\overline{\mu^{\text{sim}}}$ come from the fit to data. The mass of the J/ψ resonance, $m_{J/\psi}$, comes from the PDG [9]. Equation (6.22) ensures that if m follows a Gaussian distribution of width σ and mean μ, m^{smeared} follows a Gaussian distribution of width $s_\sigma \cdot \sigma$ and mean $\mu + \Delta\mu$. The impact of the smearing on the control samples is presented in Figure 6.14. The procedure is found to successfully match the simulated distributions of $m(e^+e^-)$ and $m(K^+e^+e^-)$ to the ones in data. There is a small region in the upper-mass tail where the smearing is not perfect, indicating the presence of one or more effects to which the procedure is not sensitive. The systematic uncertainties induced by these effects are estimated in Section 7.9, and found to have a negligible impact on R_K.

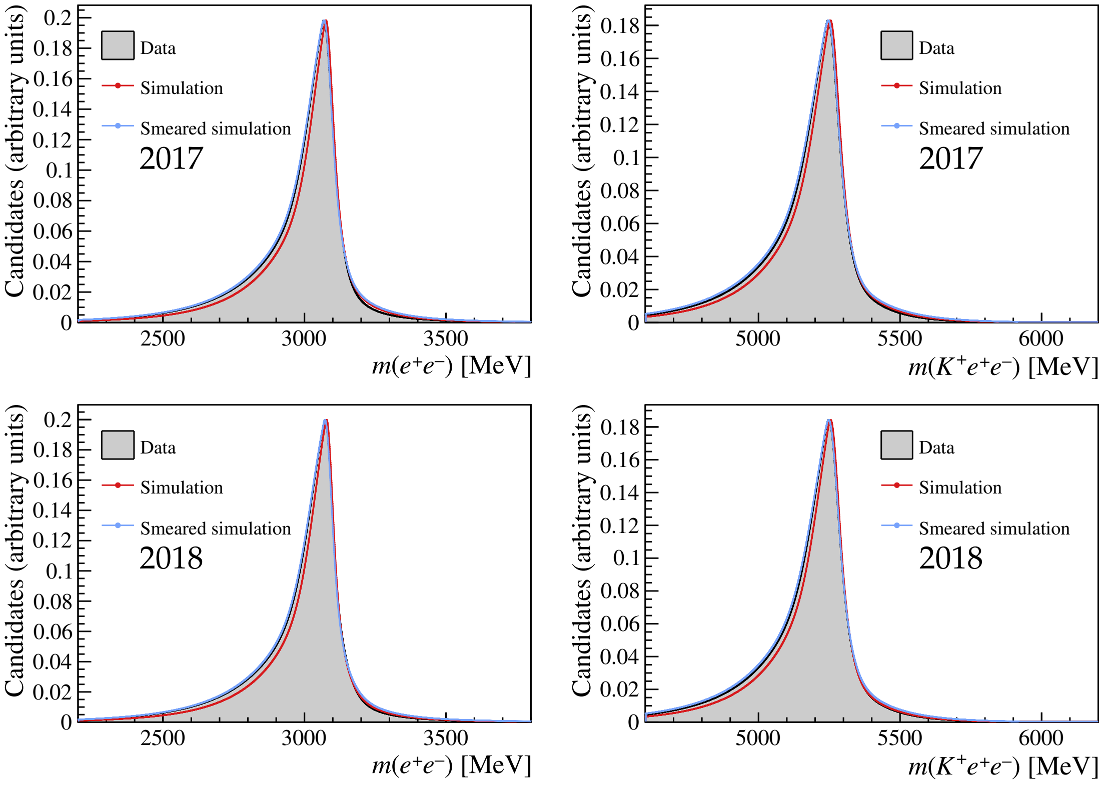

Figure 6.14: Effect of the q^2 smearing procedure on the shapes of $m(e^+e^-)$ (left) and $m(K^+e^+e^-)$ (right) in 2017 (top) and 2018 (bottom) control samples. The black histograms show the distributions in data, whilst the red and blue shapes depict the simulated distributions, before and after the resolution calibration, respectively.

Table 6.3: Estimated number of $B^+ \to K^+e^+e^-$ events migrating in and out of the range $q^2 \in (1.1\ \text{GeV}^2, 6.0\ \text{GeV}^2)$, expressed as a percentage of the total number of events in the true range $q^2_{\text{true}} \in (1.1\ \text{GeV}^2, 6.0\ \text{GeV}^2)$. The columns corresponding to events whose q^2 is in the allowed interval, but whose q^2_{true} is either above the maximum or below the minimum, are denoted by "up→in" and "down→in", respectively. Similarly, the columns corresponding to events whose q^2_{true} is in the allowed interval, but whose q^2 is either above the maximum or below the minimum, are denoted by "in→up" and "in→down", respectively. The column denoted by "in→in" lists the percentages of events where both q^2 and q^2_{true} fall into the allowed interval.

	up→in [%]	down→in [%]	in→in [%]	in→up [%]	in→down [%]
Run 1					
No smearing	8.00 ± 0.23	0.34 ± 0.05	96.94 ± 0.14	1.31 ± 0.10	1.75 ± 0.11
smearing	7.90 ± 0.23	0.43 ± 0.06	96.83 ± 0.15	1.53 ± 0.10	1.65 ± 0.11
Run 2.1					
No smearing	7.77 ± 0.18	0.48 ± 0.05	96.86 ± 0.12	1.40 ± 0.08	1.74 ± 0.09
smearing	8.78 ± 0.20	0.36 ± 0.04	96.50 ± 0.13	1.39 ± 0.08	2.11 ± 0.10
2017					
No smearing	8.62 ± 0.24	0.33 ± 0.05	96.60 ± 0.16	1.60 ± 0.11	1.81 ± 0.11
smearing	9.26 ± 0.25	0.32 ± 0.05	96.20 ± 0.17	1.66 ± 0.11	2.14 ± 0.12
2018					
No smearing	7.60 ± 0.23	0.35 ± 0.05	96.57 ± 0.16	1.64 ± 0.11	1.79 ± 0.12
smearing	8.50 ± 0.25	0.31 ± 0.05	96.22 ± 0.17	1.73 ± 0.12	2.05 ± 0.13

Table 6.4: Total rare, control, and $\psi(2S)$ efficiencies, for each data-taking period and trigger category. The right-most column lists the ratios of efficiencies between the rare and control channels. The rows labelled as "μTOS" correspond to muon samples, with the other three labels representing the three electron trigger strategies. The uncertainties are statistical only.

	$\varepsilon_{\text{rare}}$ [%]	$\varepsilon_{\text{control}}$ [%]	$\varepsilon_{\psi(2S)}$ [%]	Ratio [%]
Run 1				
μTOS	0.3165 ± 0.0016	1.3964 ± 0.0027	1.5066 ± 0.0035	22.7 ± 0.1
eTOS	0.0544 ± 0.0005	0.2167 ± 0.0009	0.2413 ± 0.0019	25.1 ± 0.2
hTOS!	0.0178 ± 0.0002	0.0350 ± 0.0003	0.0120 ± 0.0004	50.7 ± 0.8
TIS!	0.0193 ± 0.0003	0.0696 ± 0.0004	0.0614 ± 0.0009	27.7 ± 0.4
Run 2.1				
μTOS	0.3236 ± 0.0015	1.4322 ± 0.0028	1.5038 ± 0.0034	22.6 ± 0.1
eTOS	0.1041 ± 0.0007	0.4188 ± 0.0013	0.4268 ± 0.0028	24.9 ± 0.2
hTOS!	0.0264 ± 0.0004	0.0466 ± 0.0003	0.0148 ± 0.0005	56.8 ± 0.9
TIS!	0.0316 ± 0.0004	0.1095 ± 0.0006	0.0933 ± 0.0013	28.9 ± 0.4
2017				
μTOS	0.3452 ± 0.0017	1.6174 ± 0.0027	1.7400 ± 0.0040	21.4 ± 0.1
eTOS	0.0813 ± 0.0007	0.3957 ± 0.0009	0.4221 ± 0.0018	20.6 ± 0.2
hTOS!	0.0232 ± 0.0004	0.0528 ± 0.0003	0.0172 ± 0.0003	43.9 ± 0.7
TIS!	0.0288 ± 0.0004	0.1210 ± 0.0004	0.1039 ± 0.0008	23.8 ± 0.4
2018				
μTOS	0.3300 ± 0.0017	1.5609 ± 0.0027	1.6768 ± 0.0029	21.1 ± 0.1
eTOS	0.0751 ± 0.0007	0.3719 ± 0.0009	0.4015 ± 0.0009	20.2 ± 0.2
hTOS!	0.0239 ± 0.0004	0.0542 ± 0.0003	0.0184 ± 0.0003	44.0 ± 0.7
TIS!	0.0255 ± 0.0004	0.1136 ± 0.0004	0.1016 ± 0.0006	22.5 ± 0.4

A direct consequence of the smearing is the migration of $B^+ \to K^+e^+e^-$ events into and out of the $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$ range. This migration is caused by the different resolution of the reconstructed q^2 compared to its true counterpart (q^2_{true}). The fractions of events going in and out of the rare-mode q^2 window are listed in Table 6.3, before and after the application of the smearing. It is found that the smearing has a small effect, of only around 1%.

6.7 Summary of efficiencies

The estimated efficiencies of the resonant and signal modes, based on the entire selection and correction chain, are summarised in Table 6.4. The intermediary efficiencies in terms of which the ε_{tot} are factorised are presented in Appendix D. The fractions f^{q^2} of events that fall into the signal window $q^2_{\text{true}} \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$ are listed in Table 6.5. Comparing them with the ratios between signal and control efficiencies shows that most of the difference in

Table 6.5: The fraction f^{q^2} of generation-level signal events that have $q^2_{\text{true}} \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$. The values corresponding to Run 2.2 data-taking conditions are slightly different as a result of an update to the $B^+ \to K^+\ell^+\ell^-$ model used to generate the simulated samples.

	$f^{q^2}_{\text{muons}}$ [%]	$f^{q^2}_{\text{electrons}}$ [%]
Run 1	26.75 ± 0.04	29.72 ± 0.05
Run 2.1	26.85 ± 0.04	29.67 ± 0.04
2017	25.19 ± 0.04	25.15 ± 0.04
2018	25.20 ± 0.04	25.18 ± 0.04

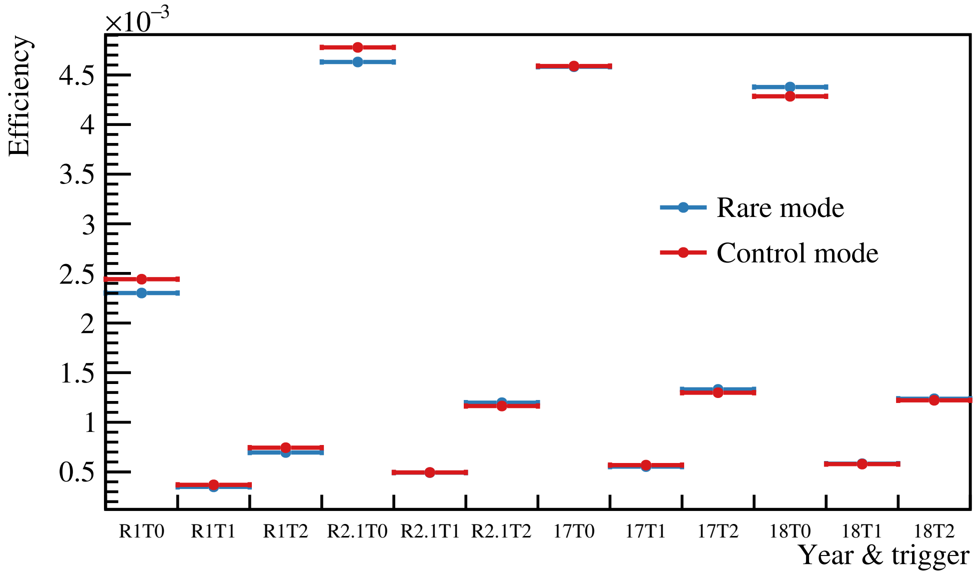

Figure 6.15: Efficiencies estimated from signal and control events with $q^2 \in (8.0\ \text{GeV},\ 10.0\ \text{GeV})$. Each group of three bins represents a particular data-taking period, which are from left to right: Run 1, Run 2.1, 2017, and 2018. In each group, the three bins correspond to the three electron triggers (eTOS, hTOS!, and TIS!).

efficiencies between the control mode and the signal mode stems from the q^2 selection. This can be examined further by evaluating the efficiency of simulated signal events with q^2 close to $m^2_{J/\psi}$. The same efficiency calculation method is applied to simulated signal and control electron events with $q^2 \in (8.0\ \text{GeV},\ 10.0\ \text{GeV})$. Here, the two channels are expected to give compatible results, however due to the fact that the q^2 distributions in this window are not identical, the agreement is not expected to be perfect. Nevertheless, the result of this test, as shown in Figure 6.15, indicates good agreement overall. This suggests that the efficiency corrections, as obtained from the J/ψ modes, can be ported onto other channels.

6.7.1 Estimate of the gain in precision

This is a study conducted before unblinding to estimate the expected gain in precision with respect to the previous R_K measurement, based on the control-mode yields and the rare and resonant efficiencies. Starting from Equation (6.6), the number of selected rare-mode events can be obtained as follows:

$$\frac{\mathcal{B}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{B}(B^+ \to K^+ J/\psi(\ell^+\ell^-))} = \frac{N_{\text{sel}}(B^+ \to K^+\ell^+\ell^-)}{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)} \cdot \frac{\varepsilon_{\text{tot}}(B^+ \to K^+ J/\psi(\ell^+\ell^-))}{N_{\text{sel}}(B^+ \to K^+ J/\psi(\ell^+\ell^-))} \cdot f^{q^2}$$

$$\begin{aligned}\Rightarrow N_{\text{sel}}(B^+ \to K^+\ell^+\ell^-) &= \frac{\mathcal{B}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{B}(B^+ \to K^+ J/\psi(\ell^+\ell^-))} \cdot \frac{N_{\text{sel}}(B^+ \to K^+ J/\psi(\ell^+\ell^-))}{\varepsilon_{\text{tot}}(B^+ \to K^+ J/\psi(\ell^+\ell^-))} \cdot \frac{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)}{f^{q^2}} \\ &= \frac{\mathcal{B}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)}{\mathcal{B}(B^+ \to K^+ J/\psi(\ell^+\ell^-))} \cdot \mathcal{N}(B^+ \to K^+ J/\psi(\ell^+\ell^-)) \cdot \frac{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)}{f^{q^2}} \,.\end{aligned}$$

It can be seen that the rare-mode yield is the product between a component that does not change with data-taking period — the fraction $\mathcal{B}_{\text{in}}(B^+ \to K^+\ell^+\ell^-)/\mathcal{B}(B^+ \to K^+ J/\psi(\ell^+\ell^-))$ — and a component that is expected to vary with run conditions:

$$\kappa \equiv \mathcal{N}(B^+ \to K^+ J/\psi(\ell^+\ell^-)) \cdot \frac{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)}{f^{q^2}} \,. \tag{6.23}$$

Therefore, computing κ for the various data-taking periods leads to an approximation for how much the rare-mode yields are expected to increase with the addition of Run 2.2 data. This is done in Table 6.6, where the estimated gain is the sum of the values of κ over all runs and all trigger categories, divided by the sum over Run 1 and Run 2.1 only. The electron dataset is expected to be approximately 2.1 times larger, whilst the statistics of the muon dataset are expected to improve by a factor of nearly 2. This is in agreement with what one would expect based on the increase in total integrated luminosity, and the improved electron-mode efficiencies in Run 2 compared to Run 1.

Table 6.6: Quantities used to estimate the expected gain in precision with respect to the previous analysis, as explained in the main body, for every data-taking period and trigger selection. The expected increase in electron and muon statistics is given in bold.

electron modes

		$\kappa \times 10^{-5}$	$\mathcal{N}(B^+ \to K^+ J/\psi(\ell^+\ell^-))$	$\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)$	f^{q^2}	$\frac{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)}{f^{q^2}}$
Run 1	eTOS	0.762	4.166×10^{7}	0.544×10^{-3}	0.2972	1.830×10^{-3}
	hTOS!	0.262	4.390×10^{7}	0.178×10^{-3}	0.2972	0.597×10^{-3}
	TIS!	0.282	4.365×10^{7}	0.193×10^{-3}	0.2972	0.647×10^{-3}
Run 2.1	eTOS	1.265	3.606×10^{7}	1.041×10^{-3}	0.2967	3.508×10^{-3}
	hTOS!	0.321	3.603×10^{7}	0.264×10^{-3}	0.2967	0.891×10^{-3}
	TIS!	0.363	3.412×10^{7}	0.316×10^{-3}	0.2967	1.065×10^{-3}
Previous		3.258				
2017	eTOS	1.033	3.195×10^{7}	0.813×10^{-3}	0.2515	3.233×10^{-3}
	hTOS!	0.306	3.315×10^{7}	0.232×10^{-3}	0.2515	0.922×10^{-3}
	TIS!	0.359	3.136×10^{7}	0.288×10^{-3}	0.2515	1.145×10^{-3}
2018	eTOS	1.194	4.003×10^{7}	0.751×10^{-3}	0.2518	2.983×10^{-3}
	hTOS!	0.399	4.200×10^{7}	0.239×10^{-3}	0.2518	0.949×10^{-3}
	TIS!	0.409	4.037×10^{7}	0.255×10^{-3}	0.2518	1.013×10^{-3}
Run 2.2		3.699				
Est. gain		**2.135**				

muon modes

		$\kappa \times 10^{-5}$	$\mathcal{N}(B^+ \to K^+ J/\psi(\ell^+\ell^-))$	$\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)$	f^{q^2}	$\frac{\varepsilon_{\text{tot}}(B^+ \to K^+\ell^+\ell^-)}{f^{q^2}}$
Run 1	μTOS	5.272	4.456×10^{7}	3.165×10^{-3}	0.2675	1.183×10^{-3}
Run 2.1	μTOS	4.571	3.793×10^{7}	3.236×10^{-3}	0.2685	1.205×10^{-3}
Previous		9.844				
2017	μTOS	4.296	3.135×10^{7}	3.452×10^{-3}	0.2519	1.370×10^{-3}
2018	μTOS	5.198	3.969×10^{7}	3.300×10^{-3}	0.2520	1.311×10^{-3}
Run 2.2		9.494				
Est. gain		**1.964**				

Chapter 7

Systematic uncertainties induced by the calculation of efficiencies

Efficiencies can be calculated by means other than those presented in Chapter 6. For this reason, systematic uncertainties have to be assigned whenever a particular method is chosen over others. This chapter covers the systematic uncertainties related to the procedure employed to calculate efficiencies. Section 7.1 provides an overview of the considered systematic effects, as well as their net contribution to the determination of R_K. Subsequent sections describe each systematic effect in detail, including the methods by which their impact is estimated.

The procedure was developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous R_K analysis. The current measurement takes into account the same sources of systematic uncertainty, with the addition of a new effect described in Section 7.8; this is specific to Run 2.2. In order to take into account the correlations between data-taking periods and trigger selections, the entire procedure is rerun with the addition of Run 2.2 data. Therefore, all results presented here constitute original work. The Run 2.1 results are the same as in the previous measurement. Their Run 1 counterparts change by a few permille, due to an update in the simulation of 2011 samples.

Let $\overline{\varepsilon}$ be an efficiency estimate obtained by means described in Chapter 6. This represents the "nominal" estimation of a true efficiency, ε. To assign a systematic uncertainty on $\overline{\varepsilon}$, several methods are used to derive n "alternative" estimates of ε, denoted by $\{\varepsilon^i\}_{i=\overline{1,n}}$. These values are used to evaluate the variance on $\overline{\varepsilon}$, and the square root of this variance is interpreted as the systematic uncertainty:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\varepsilon^i - \overline{\varepsilon})^2}\,. \tag{7.1}$$

This exercise is performed on the rare and resonant modes. Two efficiencies that correspond to the same channel, but not necessarily the same trigger and/or data-taking period, may be correlated. To take this into account, efficiencies are calculated individually for each data-taking period and trigger selection. The covariance between two such efficiencies, ε_1 and ε_2, is then given by:

$$V_{1,2} = V_{2,1} = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_1^i - \overline{\varepsilon_1})\cdot(\varepsilon_2^i - \overline{\varepsilon_2})\,. \tag{7.2}$$

Similarly, the correlation factor between ε_1 and ε_2 is:

$$\mathrm{corr}_{1,2} = \mathrm{corr}_{2,1} = \frac{V_{1,2}}{\sigma_1\sigma_2}\,. \tag{7.3}$$

When considered together, the variances and the correlations encode information on how particular choices of efficiency calculation methods affect the overall result. This information is used in two places. The first is the combination of the $r_{J/\psi}$ and $R_{\psi(2S)}$ ratios, as described in Chapter 8. The second is the fit to the rare modes, as presented in Chapter 9.

In this thesis, systematic uncertainties on the ratios $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$ are presented through fractional error matrices S, defined as:

$$S = \begin{pmatrix} \sigma_1/\overline{\varepsilon}_1 & \mathrm{corr}_{1,2} & \dots & \mathrm{corr}_{1,n} \\ \mathrm{corr}_{2,1} & \sigma_2/\overline{\varepsilon}_2 & \dots & \mathrm{corr}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{corr}_{n,1} & \mathrm{corr}_{n,1} & \dots & \sigma_n/\overline{\varepsilon}_n \end{pmatrix}. \tag{7.4}$$

Given that the R_K measurement considers 4 distinct data-taking periods and 3 electron triggers, S is a 12×12 matrix. Its diagonal contains fractional uncertainties, whilst the off-diagonal terms represent correlations between different selections.

The following section presents the total systematic uncertainties on $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$, together with the contributions from every considered effect. The rest of the chapter describes how each individual systematic effect is assessed. The effects are presented in descending order of their impact on the total uncertainty. They are: the kinematic calibration method; the finite size of the calibration samples; the detector occupancy; the signal decay model; the trigger calibration; the modelling of the material budget; the PID efficiency correction; and the resolution of q^2 and $m(K^+e^+e^-)$.

7.1 Summary of efficiency systematics

Covariance matrices are obtained individually for all considered sources of systematic uncertainty on the efficiencies, and then added together into the total covariance matrices for $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$. The total covariance matrix for R_K is then used in the fit to $B^+ \to K^+\ell^+\ell^-$ data, whereas the other two total covariance matrices lead to the combined $r_{J/\psi}$ and $R_{\psi(2S)}$ estimates presented in Chapter 8. The fractional error matrices that result from applying Equations (7.3) and (7.4) to the total covariance matrices are presented in Table 7.1. Note that the effect of the fit model is treated separately, in Section 10.2. As expected, $r_{J/\psi}$ has the largest systematic uncertainty, given that it does not benefit from the double-ratio cancellation inherent to R_K and $R_{\psi(2S)}$.

Table 7.2 lists all considered systematic effects on the ratios $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$. The total systematic uncertainty on R_K is found to be around 1.5%, which is predominantly due to the $B^+ \to K^+\ell^+\ell^-$ fit model. Thanks to the double ratio, all systematic effects associated with the calculation of efficiencies are reduced below the percent-level. This does not contradict the values listed in Table 7.1, where the overall impact on each data-taking period and trigger category is presented.

Table 7.1: Total fractional error matrices for $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$, obtained by propagating all systematic uncertainties related to the calculation of efficiencies. These matrices are symmetric, so the elements below the main diagonal are omitted. All entries are listed as percentages.

$r_{J/\psi}$

	Run 1			Run 2.1			2017			2018	
eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!
6.35	−9.87	−7.99	63.24	49.47	45.42	57.17	27.73	5.65	49.75	34.19	26.42
	6.78	36.04	38.80	28.16	5.94	36.73	66.97	27.65	40.26	57.36	10.98
		7.43	18.49	−7.55	49.45	26.69	18.09	38.39	27.57	2.41	24.24
			6.08	25.98	14.56	72.80	45.65	26.50	74.08	37.63	23.20
				8.85	59.16	21.81	55.50	22.47	13.56	64.73	42.59
					8.33	21.55	25.49	35.33	11.18	24.34	51.38
						4.67	68.53	11.52	96.63	60.08	13.69
							6.72	10.29	68.10	90.18	12.57
								3.17	6.88	−1.71	60.87
									4.84	59.23	5.17
										6.77	8.35
											3.18

R_K

	Run 1			Run 2.1			2017			2018	
eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!
2.05	46.35	5.99	3.51	1.10	4.74	−6.26	3.97	14.24	3.72	15.21	4.23
	3.35	9.24	−11.92	6.47	4.71	−23.12	6.65	5.70	−12.85	8.24	−2.00
		1.92	0.02	−0.09	2.27	−3.69	−1.31	2.24	9.03	9.75	12.05
			1.51	17.52	46.23	49.66	28.62	26.27	42.75	21.16	14.89
				7.90	44.64	−0.59	3.20	−0.55	1.25	9.53	3.51
					2.20	22.75	26.25	7.77	26.35	1.90	3.48
						1.68	42.26	22.18	49.15	12.55	10.43
							3.00	18.90	29.97	29.73	−2.21
								2.41	28.23	16.61	27.18
									1.51	34.83	29.70
										2.72	31.11
											2.23

$R_{\psi(2S)}$

	Run 1			Run 2.1			2017			2018	
eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!
1.49	20.28	48.99	36.42	1.99	42.25	44.64	16.05	27.31	32.46	23.84	24.30
	4.21	29.40	2.64	31.98	37.82	24.78	39.63	32.76	3.75	30.53	26.42
		3.10	19.34	−2.73	44.69	43.44	1.14	9.57	18.80	11.80	13.75
			1.21	8.03	18.51	41.92	20.63	30.56	56.15	21.42	28.82
				4.87	0.60	−2.43	22.31	19.20	9.88	32.34	19.44
					2.36	41.30	24.53	25.02	28.64	13.73	17.25
						1.14	11.48	35.17	57.98	3.67	21.55
							3.67	57.54	16.31	49.03	37.25
								2.26	37.60	38.50	49.20
									0.96	19.27	38.23
										2.83	39.45
											1.86

Table 7.2: Fractional systematic uncertainties on $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$. The contributions are sorted by their effect on R_K, in descending order. Indented entries are related to each other and given separately, in addition to the combined effect above them. The total is not identical to the sum in quadrature of the individual entries, since correlations are taken into account.

Source	$r_{J/\psi}$ [%]	R_K [%]	$R_{\psi(2S)}$ [%]
1. Fit model	—	1.00	—
1a. Signal	—	0.70	—
1b. Background	—	0.71	—
2. Kinematic corrections	1.57	0.59	0.52
3. Finite size	1.41	0.47	0.33
4. Occupancy	1.17	0.39	0.09
5. Decay model	—	0.39	—
6. Trigger calibration	0.78	0.37	0.15
6a. *e*TOS	0.63	0.36	0.14
6b. *h*TOS!	0.39	0.07	0.03
6c. μTOS	0.28	0.04	0.06
6d. TIS!	0.40	0.03	0.01
7. Material and tracking	—	0.29	0.08
8. PID calibration	0.80	0.25	0.07
8a. Electron PID	0.80	0.25	0.07
8b. Muon and kaon PID	0.03	0.01	0.01
9. q^2 and mass resolution	0.53	0.19	0.57
9a. Parametrisation	0.43	0.14	0.44
9b. Trigger bias	0.30	0.10	0.14
9c. Upper mass tail	0.44	0.08	0.34
Total	**2.18**	**1.51**	**0.87**

7.2 Kinematic corrections

As described in Section 6.5, the simulated distributions of kinematic variables are calibrated through correction weights to match the data. These weights are derived using $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ μTOS samples, and residual imperfections in the `L0Muon` calibration may affect the kinematic weights. In addition, corrections are extracted from muon modes and applied to electrons as well. This means that the calibration may not necessarily account for effects such as electron bremsstrahlung emission in the VELO. As a result, the calibration of the χ^2_{DV} and χ^2_{IP} of the B^+ may not be perfect in the electron samples.

To evaluate the size of such systematic effects, the procedure described in Section 6.5 is repeated to produce alternative sets of kinematic weights. These are then used to obtain alternative efficiency estimates, which are compared to their nominal values to determine the size of the systematic effect. In the muon modes, two sets of weights are compared:

1. μTOS weights: these are the nominal weights; they come from $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ samples obtained using the μTOS trigger strategy; and
2. μTIS weights: these are computed using $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ samples where the L0 trigger fires independently of the signal.

In the electron modes, there are five sets of weights:

1. μTOS weights: like for the muon samples, these are the nominal weights;
2. μTIS weights: these are obtained from the strategy used to derive the alternative weights for the muon channels;
3. eTOS weights: these are computed using $B^+ \to K^+ J/\psi(e^+e^-)$ eTOS samples;
4. eTIS weights: these also use $B^+ \to K^+ J/\psi(e^+e^-)$ events, but from the inclusive TIS trigger instead of the eTOS trigger; and
5. mixed weights: $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ μTOS samples are used to correct the p_{T}, η and χ^2_{IP} of the B^+, with χ^2_{DV} corrections coming from $B^+ \to K^+ J/\psi(e^+e^-)$ eTOS samples.

To study the effect of the different calibration strategies, control-mode yields are computed in bins of phase space. Then, they are divided by the efficiencies of each bin, as obtained using each set of kinematic weights. This leads to efficiency-corrected yields, which are expected to be uniform across phase space, provided the kinematics are calibrated correctly. Figure 7.1 shows efficiency-corrected 2018 yields, as a function of p_{T} and χ^2_{DV} of the B^+. It can be seen that the nominal μTOS corrections lead to flat efficiency-corrected muon yields, however χ^2_{DV} shows a trend in the electron sample. If mixed weights are used instead, the efficiency-corrected electron yield becomes uniform in χ^2_{DV}. Therefore, the alternative calibration strategies are expected to provide good coverage of the systematic effects related to the kinematic corrections. The same conclusions are valid for the other data-taking periods, because they yield qualitatively similar results; for this reason, they are omitted.

The systematic effect of the kinematic corrections is listed in row 2 of Table 7.2. The impact on $r_{J/\psi}$ is around 1.6%, which gets suppressed to approximately 0.5% in the double ratios.

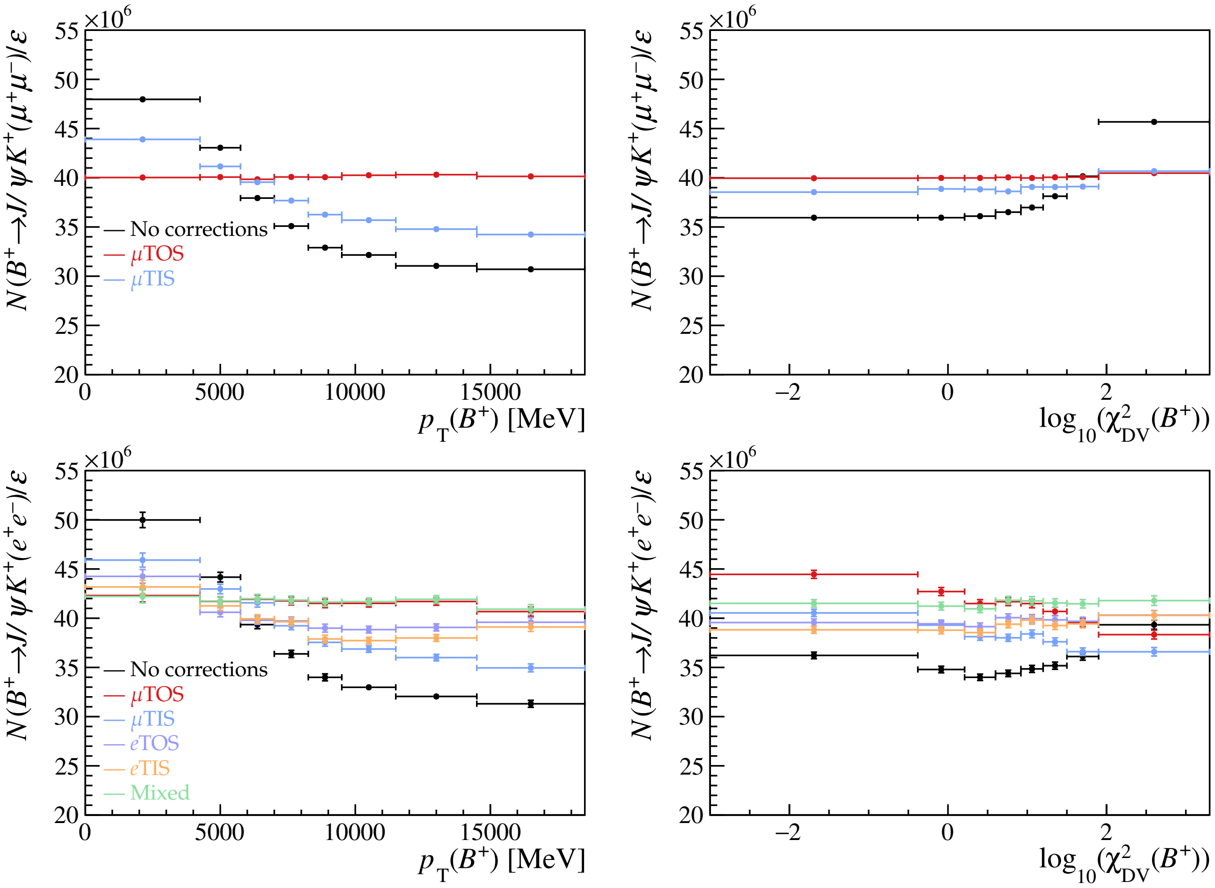

Figure 7.1: Efficiency-corrected yields of 2018 μTOS (top) and eTOS (bottom) data. The efficiencies used in the denominator are computed based on the different kinematic calibration strategies described in the main body.

7.3 Finite size of simulation and calibration data

The $B^+ \to K^+J/\psi(\ell^+\ell^-)$ modes are used extensively to compute calibration weights. This means that the efficiency corrections are correlated with the statistics of the control modes. For this reason, the errors on the $B^+ \to K^+J/\psi(\ell^+\ell^-)$ yields are treated as systematic uncertainties, and their correlations with the efficiencies are taken into account by employing a bootstrap method [193]. This involves assigning each event in simulation and data a weight drawn from a Poisson distribution of mean 1, and then repeating all selection and calibration steps based on these Poisson weights. This is done 100 times, thus leading to 100 versions of the efficiencies and control-mode yields. These are subsequently used in Equations (7.1) and (7.2) to estimate the associated systematic uncertainty.

The net impact of the size of the simulation and calibration data is listed in row 3 of table Table 7.2. It is smaller than the effect induced by the calibration of the B^+ kinematics. The simulated Run 2.2 samples benefit from increased statistics, leading to a smaller systematic effect than in the preceding R_K measurement.

7.4 Occupancy proxies

For reasons outlined in Section 6.4, the event occupancy is simulated imperfectly, and cannot be accessed directly. This effect cannot be mitigated through correction weights, because strategies that make one occupancy proxy agree between data and simulation lead to disagreements in other proxies. For this reason, no corrections are applied to the occupancy of simulated events. Instead, a systematic uncertainty is assigned to account for imperfections in the modelling of the occupancy. The procedure described in Section 6.5 is repeated after the addition of three occupancy proxies, one at a time, to the kinematic calibration scheme. This leads to three sets of alternative efficiencies, from which systematic uncertainties are calculated using Equations (7.1) and (7.2). The three considered occupancy proxies are:

1. `nSPDHits`: the number of hits in the scintillating pad detector;
2. `nTracks`: the number of tracks in the event; and
3. `nPVs`: the number of primary vertices reconstructed in the VELO.

The procedure is hence similar to the one presented in Section 7.2, with an additional effect that is taken into account. The performance of the TIS! trigger strategy depends on the occupancy, so the calibration histograms presented in Section 6.3.4 are recomputed after each occupancy proxy has been corrected. The efficiencies are then calculated using these histograms, rather than the nominal kinematic corrections histograms.

The impact of the occupancy proxies is presented in row 4 of Table 7.2. The single ratio $r_{J/\psi}$ is affected by approximately 1%, whilst the double ratios are only impacted at the permille-level, thanks to the in-built cancellation.

7.5 Signal decay model

The q^2 distribution in simulated $B^+ \to K^+\ell^+\ell^-$ events depends on the theoretical model used to describe these decays. This in turn has an effect on the efficiencies, so the systematic effect of the model parameters has to be evaluated. Dr. Paula Álvarez Cartelle used the `flavio` software package [79] to fluctuate these parameters according to their uncertainties 100 times, thus obtaining 100 alternative q^2 distributions. The rest of the procedure, which constitutes original work, involves correcting the simulated q^2 distribution to match each of the fluctuated distributions in turn. This leads to 100 alternative sets of efficiency estimations, which are combined according to Equations (7.1) and (7.2) to evaluate the bias induced by the uncertainties in the decay model. The alternative distributions are presented in Figure 7.2. For $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$, where R_K is measured, the four data-taking periods lead to similar results. However, there are differences at higher values of q^2, as a result of the different models used to simulate $B^+ \to K^+\ell^+\ell^-$ decays. This is reflected in the different estimates of f^{q^2} listed in Table 6.5. The overall impact on R_K is presented in row 5 of Table 7.2. The effect is found to be of a few permille.

7.6 Trigger calibration

The trigger calibration method has several aspects that could induce systematic uncertainties. One of them is the choice of tag employed to compute efficiencies via the “tag & probe” method described in Section 6.3. As shown in Figures 6.4, 6.6, 6.8, and 6.10, using the nominal tags leads to simulated efficiencies that are in good agreement with those extracted without the use of a tag (hence being labelled as “direct”). The other tagging strategies shown in these figures are used to derive alternative trigger corrections, from which the systematic effect is evaluated using Equations (7.1) and (7.2). In the case of $e\gamma$TIS, which is shown in Figure 6.9, no alternative unbiased tag is found. For this reason, the systematic effect it induces is evaluated using an alternative binning scheme for the efficiency as a function of the maximum electron p_{T}.

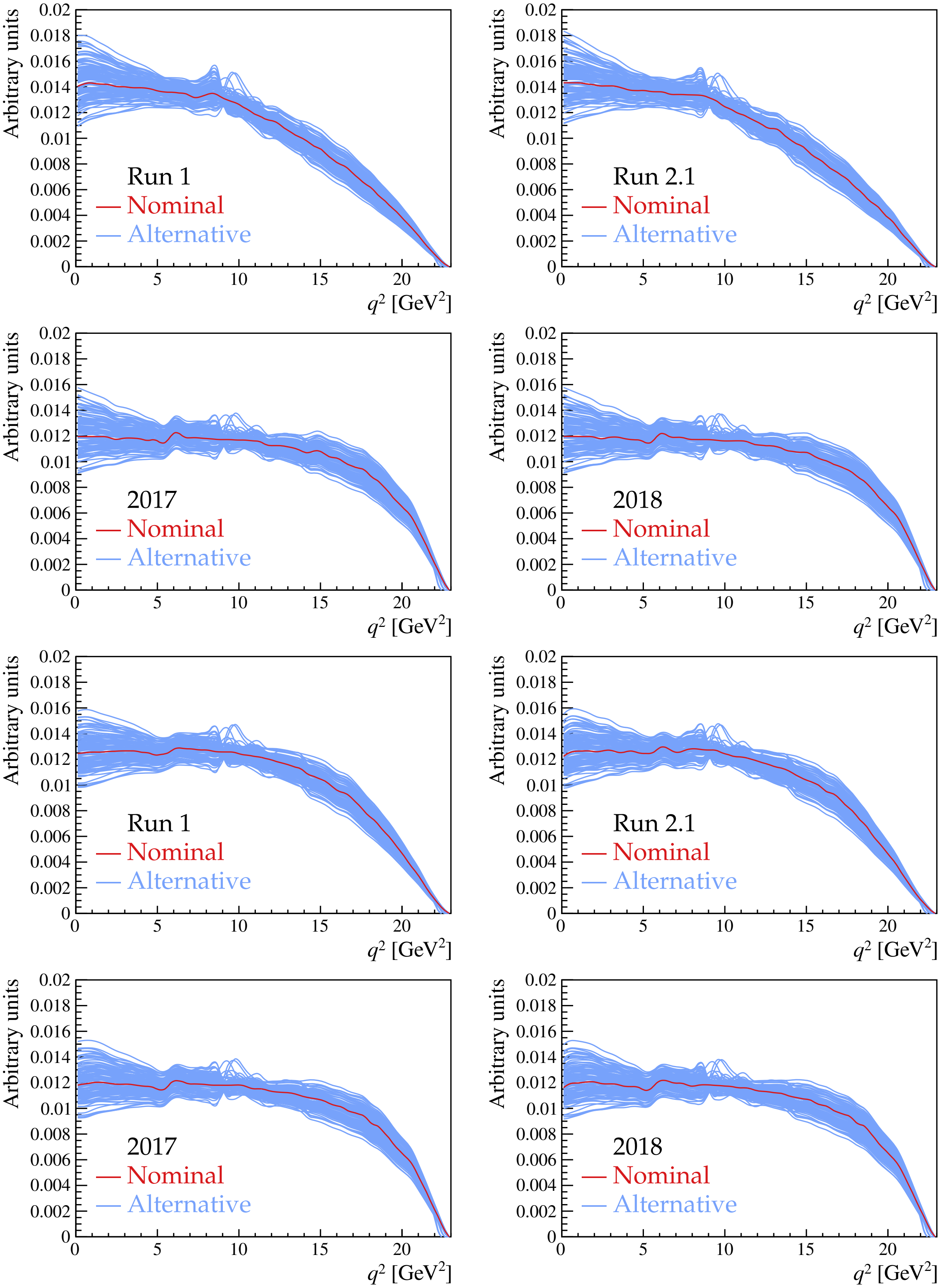

Figure 7.2: Simulated q^2 distributions of generated $B^+ \to K^+e^+e^-$ (top two rows) and $B^+ \to K^+\mu^+\mu^-$ (bottom two rows) events, before and after the application of weights derived using `flavio`. The original distributions are shown in red, whilst the changes induced by applying the weights derived using `flavio` are shown in blue. The predictions in the region of narrow charmonium resonances are not meaningful.

Another potential source of systematic uncertainty is the assumption that the trigger efficiencies of the two leptons in the candidate are independent. The `L0Muon` and `L0Electron` correction weights are computed based on this factorisation, which is shown in Figure 6.5 to hold. However, Figure 6.7 indicates that the electron trigger efficiency depends on how well-separated the two candidate electrons are in the ECAL. To assess the impact of this effect, `L0Electron` efficiencies are recalculated using an alternative binning scheme. Instead of the nominal E_T of each electron, this scheme uses the distance between the two ECAL clusters, r_calo, and the maximum E_T of the two candidate electrons. The efficiencies resulting from this method are shown in Figure 7.3 for 2018 data and simulation. Their equivalents for the other data-taking periods are similar, and therefore omitted.

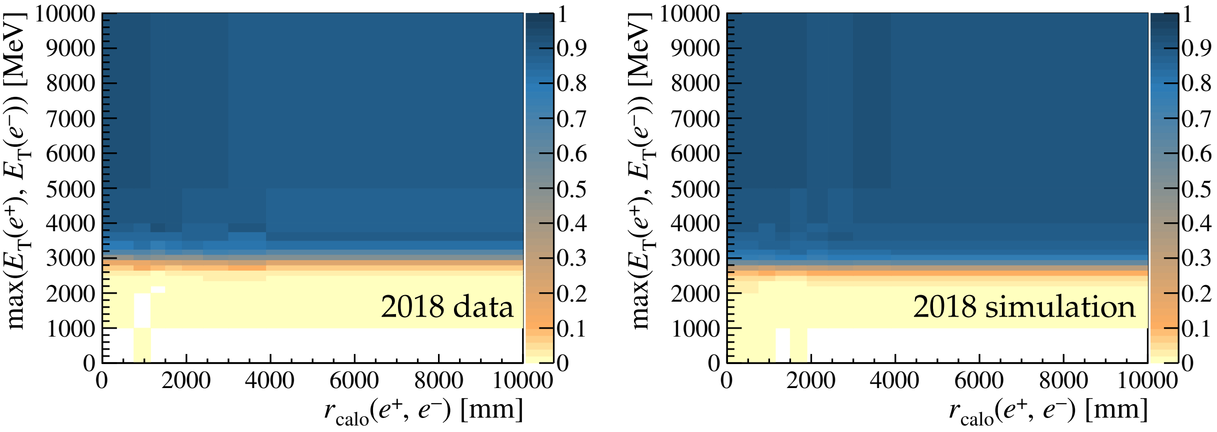

Figure 7.3: Efficiency with which the two electrons in $B^+ \to K^+ J/\psi(e^+e^-)$ data (left) and simulation (right) trigger the `L0Electron` line, as a function of their maximum E_T and the distance between their ECAL clusters. All efficiencies are computed using the nominal eTIS tag, described in Section 6.3.2.

The total systematic uncertainty induced by the trigger calibration is listed in row 6 of Table 7.2. The individual contributions of each trigger selection are presented as indented subpoints. The double ratio leads to good cancellation of these uncertainties, which are mostly at the sub-permille level. The dominant effect is the eTOS factorisation, which has a permille effect on R_K.

7.7 Material and tracking

It can be seen in Figure 6.12 that the kinematic calibration weights do not fully account for discrepancies between the distributions of electron pseudorapidity (η) in data and simulation.

This is attributed to imperfections in the simulated material budget of the detector. If the simulation underestimates the material budget as a function of η, the electron interaction probability in the detector is also underestimated. This leads to an overestimated tracking efficiency in those particular regions of η, because not as much bremsstrahlung radiation is emitted. In other words, mismodelling of the electron pseudorapidity and tracking efficiency is caused by imperfections in the simulated material budget of the detector.

A systematic uncertainty attributed to this mismatch is assigned using weights that calibrate the distribution of simulated electron pseudorapidity to match the data. Alternative efficiency estimates are obtained based on these weights, and then compared to the nominal efficiencies that do not use the pseudorapidity calibration weights. The overall impact on R_K is then calculated using Equations (7.1) and (7.2).

The pseudorapidity calibration weights are calculated as data-over-simulation ratios of two-dimensional $(\eta(e^+),\, \eta(e^-))$ distributions in $B^+ \to K^+ J/\psi(e^+e^-)$ eTOS samples. The ratios are computed separately for electrons that have or have not had a photon added by the bremsstrahlung recovery procedure. Alternative binning schemes, which also take into account the polarity of the magnet, are found to yield equivalent results. The calibration weights for 2018 samples are illustrated in Figure 7.4. The results corresponding to the other data-taking periods are qualitatively similar, and therefore omitted. The effect of the weights is illustrated in Figure 7.5, where the simulated pseudorapidity distribution in control samples, after the application of corrections, is found to be similar to the one in data. The distribution in the signal samples is not expected to be identical, due to the different kinematics. However, the corrections are not expected to depend strongly on kinematics, and so the pseudorapidity correction weights can be applied to the simulated signal samples.

Both the rapidity corrections, $w(\eta(e^+),\, \eta(e^-))$, and the electron tracking efficiencies, $\varepsilon_{\text{tracking}}$, depend on the material budget. As a result, they can be related through an unknown normalisation factor k:

$$w(\eta(e^+),\, \eta(e^-)) = k \cdot \frac{\varepsilon^{\text{data}}_{\text{tracking}}(\eta(e^+)) \cdot \varepsilon^{\text{data}}_{\text{tracking}}(\eta(e^-))}{\varepsilon^{\text{sim}}_{\text{tracking}}(\eta(e^+)) \cdot \varepsilon^{\text{sim}}_{\text{tracking}}(\eta(e^-))}\,. \tag{7.5}$$

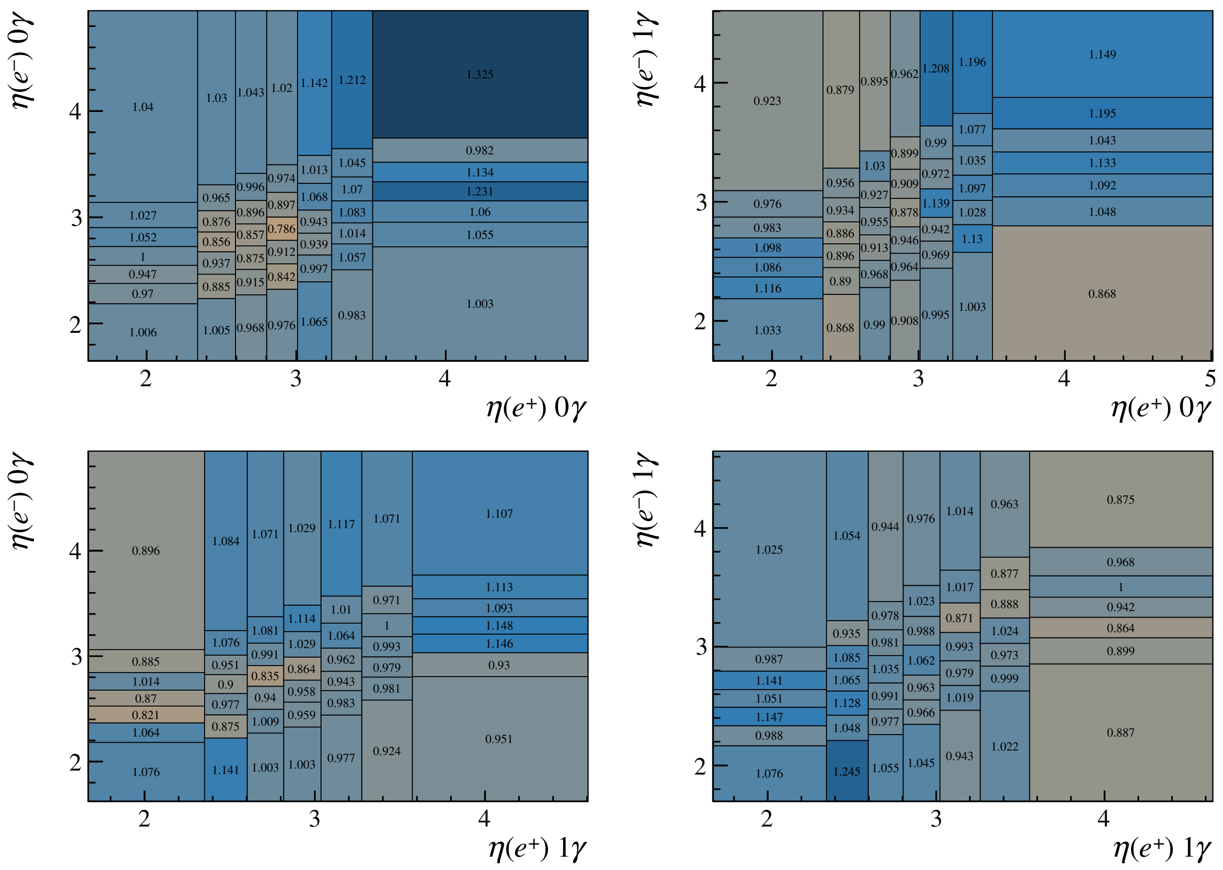

Figure 7.4: Weights used to calibrate the simulated pseudorapidity distribution in simulated 2018 electron samples. The axis labels indicate whether the weights are applied to electrons that have undergone bremsstrahlung recovery (1γ), or to electrons that have not (0γ).

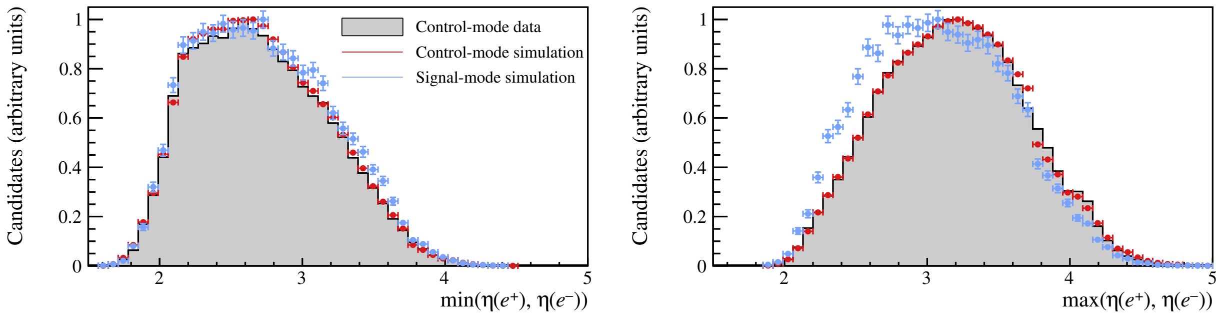

Figure 7.5: Minimum and maximum pseudorapidity of the two electrons in 2018 eTOS candidates. The grey histograms show the distribution in background-subtracted $B^+ \to K^+ J/\psi(e^+e^-)$ data. The red and blue lines depict the distributions in fully-selected $B^+ \to K^+ J/\psi(e^+e^-)$ and, respectively, $B^+ \to K^+e^+e^-$ simulated events, to which the pseudorapidity calibration procedure is applied.

The factor k depends on the statistics of the data and simulation samples in each data-taking period, as well as on the global ratio of efficiencies. This has two implications: first, k is expected to vary across data-taking periods; second, the pseudorapidity weights cannot be used to compute the impact of electron tracking discrepancies on the total efficiency of $B^+ \to K^+ J/\psi(e^+e^-)$, and by extension on $r_{J/\psi}$. However, k cancels out in the ratio of electron efficiencies which is built into R_K and $R_{\psi(2S)}$. For this reason, the pseudorapidity correction weights are used to assign a systematic uncertainty on the two double ratios, but not on $r_{J/\psi}$.

The result is presented in row 7 of Table 7.2. It is found to be at the permille level, thanks to the cancellation of this effect in the double ratio. The size of the effect on the double ratios is validated through an alternative method of assessing this systematic uncertainty, which is presented in Section 8.4.

7.8 PID corrections

There are three effects that lead to systematic uncertainties in the calculation of the PID efficiencies. The first one is induced by the chosen binning of the PID calibration tables. On one hand, increasing the number of bins would in principle lead to more accurate efficiency estimates. On the other hand, using fewer bins reduces the impact of random statistical fluctuations. To assess how the efficiency histograms for kaon and muon PID are affected by the binning, alternative calibration tables are produced by changing the final step in the generation procedure. As outlined in Section 6.2.1, this step consists of merging adjacent bins if their corresponding efficiencies are compatible within a certain threshold. This threshold is nominally set to $2.5\,\sigma$, so alternative tables are generated by changing this threshold to 1.5, 2.0, 3.0, and $3.5\,\sigma$. For the electrons, the calibration samples are not statistically powerful enough to allow the use of this method. As a consequence, only one alternative binning scheme is used. The boundaries of the bins in this scheme correspond to the central values of the nominal bins.

Another systematic effect is illustrated in Figure 6.3. It can be seen that changing the trigger selection of the electron calibration samples leads to variations in the resulting efficiencies.

To quantify this effect, alternative electron PID efficiencies are obtained based on samples where the L0 fires independently of the signal.

The final systematic uncertainty is specific to the Run 2.2 electron calibration samples. As mentioned in Section 6.2.2, their selection cuts harsher on the χ^2_{IP} of the electrons than the selection used by the R_K measurement. Nominally, the effect of this cut is accounted for using correction factors, $w_i^{\chi^2_{\text{IP}}}$, for the efficiencies obtained from data. These factors are calculated on a bin-by-bin basis, using the simulated effect of the tighter χ^2_{IP} cut on the efficiencies. The PID efficiency in each bin then becomes:

$$\varepsilon^i_{\text{PID}} \mapsto \varepsilon^i_{\text{PID}} \cdot w_i^{\chi^2_{\text{IP}}} = \varepsilon^i_{\text{PID}} \cdot \frac{\varepsilon^i_{\text{PID}}(\text{sim} \mid \text{loose cut})}{\varepsilon^i_{\text{PID}}(\text{sim} \mid \text{tight cut})}\,, \tag{7.6}$$

where the efficiencies in the ratio on the right-hand side are obtained from simulated $B^+ \to K^+ J/\psi(e^+e^-)$ events. The loose cut corresponds to the one in the R_K selection, whilst the latter is obtained by tightening the cut to the value used by the selection of the calibration data. The expression above does not rely on the absolute efficiency values in the simulation, which are known to be imperfect. Instead, the underlying assumption is that dependencies of the PID efficiency on the χ^2_{IP} of electrons are modelled correctly. This assumption is tested by generating alternative calibration histograms, using both data and simulated control samples to which a tighter cut in χ^2_{IP} is applied. The correction factors $w_i^{\chi^2_{\text{IP}}}$ are adjusted to take into account data and simulation differences in the modelling of the PID efficiency as a function of χ^2_{IP}. The adjustment is done in two ways:

$$w_i^{\chi^2_{\text{IP}}} \mapsto w_i^{\chi^2_{\text{IP}}} \cdot \frac{\varepsilon^i_{\text{PID}}(\text{data} \mid \text{tight cut}) \,/\, \varepsilon^i_{\text{PID}}(\text{data} \mid \text{tighter cut})}{\varepsilon^i_{\text{PID}}(\text{sim} \mid \text{tight cut}) \,/\, \varepsilon^i_{\text{PID}}(\text{sim} \mid \text{tighter cut})}\,, \text{ and} \tag{7.7}$$

$$w_i^{\chi^2_{\text{IP}}} \mapsto w_i^{\chi^2_{\text{IP}}} \,/\, \frac{\varepsilon^i_{\text{PID}}(\text{data} \mid \text{tight cut}) \,/\, \varepsilon^i_{\text{PID}}(\text{data} \mid \text{tighter cut})}{\varepsilon^i_{\text{PID}}(\text{sim} \mid \text{tight cut}) \,/\, \varepsilon^i_{\text{PID}}(\text{sim} \mid \text{tighter cut})}\,. \tag{7.8}$$

Using both correction factors accounts for the fact that it is unknown in which direction the simulated efficiency in the cut-out region disagrees with data. This is therefore a conservative way of assessing this systematic effect.

The overall effect of the methods used to assess the PID performance is presented in row 8 of Table 7.2. As shown by the indented sub-entries in the table, the electron PID is significantly more impacted by systematic effects than the kaon and muon PID. Even so, these systematics cancel almost entirely in the double ratio, so the net effect on R_K is of a few permille.

7.9 Resolution of q^2 and $m(K^+e^+e^-)$

The smearing method described in Section 6.6 introduces three sources of systematic uncertainty. The first one is attributed to the assumption that the disagreement between q^2 in data and simulation is uniform across phase space. Since candidate electrons have momenta much larger than their mass, the J/ψ invariant mass can be written as:

$$m(e^+e^-) = \sqrt{2\,p_{e^+}\,p_{e^-}\,(1 - \cos\alpha_{\ell^+\ell^-})}\,, \tag{7.9}$$

where p_{e^+} and p_{e^-} are the momenta of the electrons, and $\alpha_{\ell^+\ell^-}$ is the opening angle between them. Depending on its magnitude, the momentum has a resolution of around 10%. Since $\alpha_{\ell^+\ell^-}$ is found to have a resolution two orders of magnitude better than that of the momentum, any q^2-dependent smearing effects would be driven by momentum dependencies. To assess the systematic effect of any such trends, the smearing factors are recomputed using $K^+J/\psi(e^+e^-)$ data and simulation binned in minimum and maximum candidate electron momentum. The resulting factors are then used to obtain alternative efficiency estimates, which are used in Equations (7.1) and (7.2) to evaluate the systematic effect. Projections of the ratios between resolutions in data and in simulation are shown in Figure 7.6.

Another consideration is the fact that the smearing is done using the scales and shifts averaged across trigger selections[1]. A systematic effect may be induced by the small fluctuations present across samples taken using different trigger selections. To assess the size of the effect, the efficiencies are recomputed using smearing factors fluctuated one standard deviation above and below their uncertainties. These uncertainties are the RMS between the factors computed in the three trigger categories.

---

[1] These factors are presented in Table 6.2.

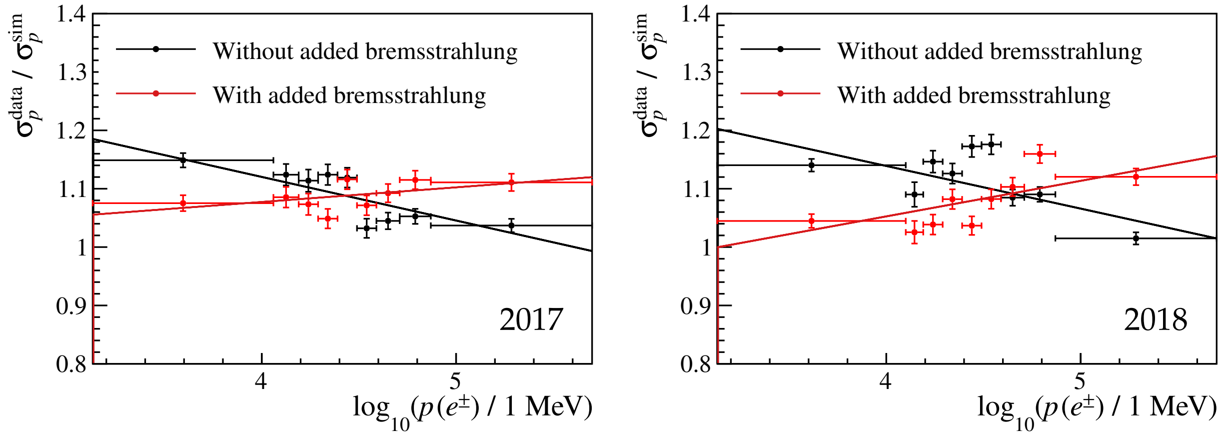

Figure 7.6: Data-over-simulation ratio of the momentum resolution in $B^+ \to K^+ J/\psi(e^+e^-)$ 2017 (left) and 2018 (right) samples. Linear minimum χ^2 fits are superimposed to highlight the momentum dependency.

Another systematic effect is induced by the imperfect modelling of the upper mass tail of the $m(e^+e^-)$ distribution. This tail is dominated by events where the bremsstrahlung recovery algorithm incorrectly adds too much energy back to the signal electrons. This can happen when the added energy comes from an ECAL cluster corresponding to a stray photon in the event. Since the number of stray photons changes with occupancy, the size of the upper mass tail is also expected to depend on the occupancy. For this reason, the $m(e^+e^-)$ fit used to derive the smearing parameters is repeated by using three different signal templates. These are obtained from simulated samples with small, medium, and large values of `nSPDHits`. It can be seen in Figure 7.7 that the tails of the signal templates change with `nSPDHits`. Taking this into account leads to better modelling of the upper mass tail in data, as shown in Figure 7.8.

The total systematic uncertainties induced by the smearing procedure are reported in row 9 of Table 7.2. The subsequent indented entries show the contributions from the three considered effects, which are: the momentum dependence, trigger bias, and upper mass tail. Each of them is found to impact R_K only by approximately 0.1%.

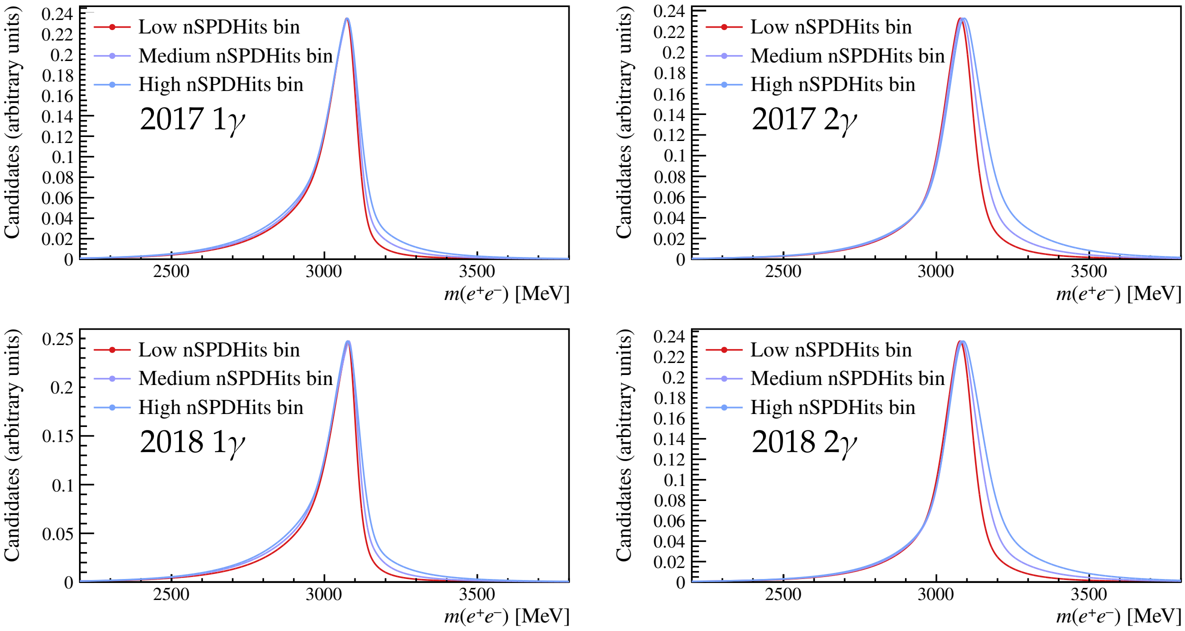

Figure 7.7: Shape templates for $m(e^+e^-)$ in simulated Run 2.2 eTOS events from photon categories 1γ (left) and 2γ (right), shown separately for small, medium, and large values of `nSPDHits`. The normalisation is such that the peaks have the same height.

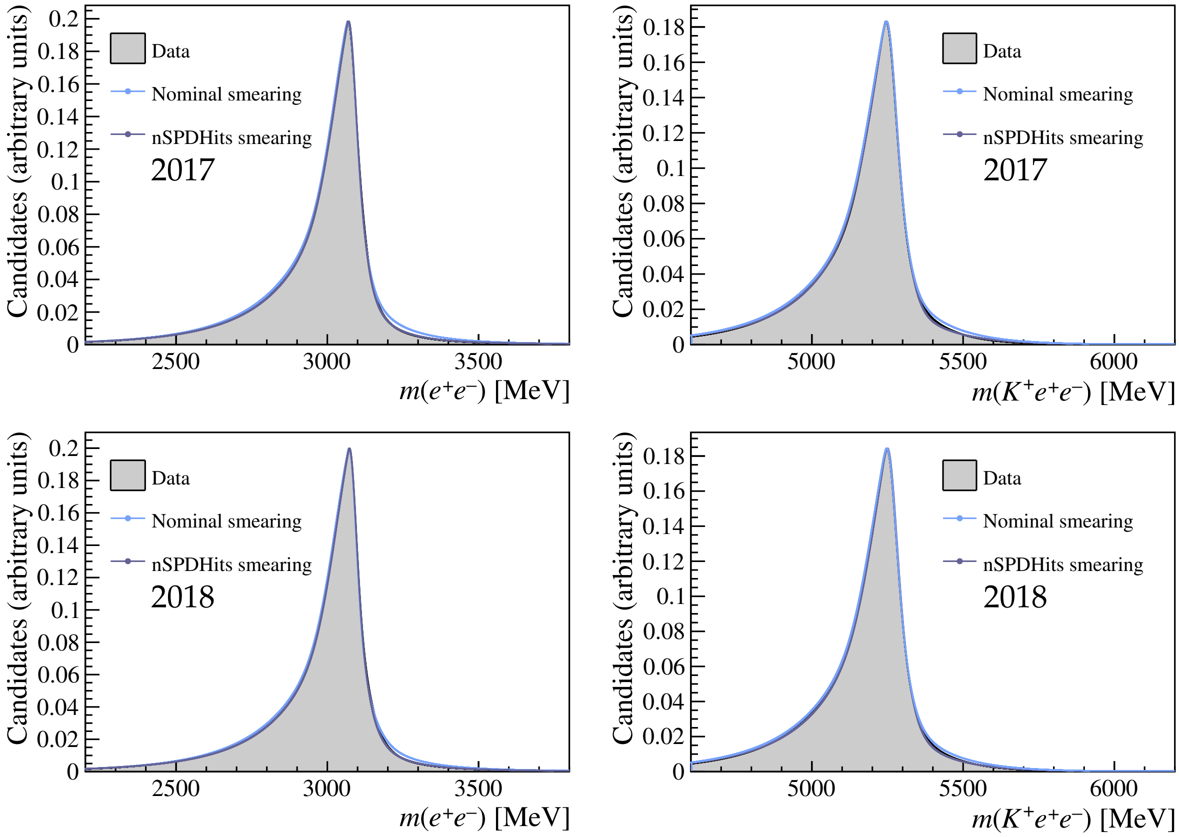

Figure 7.8: Distribution of $m(e^+e^-)$ (left) and $m(K^+e^+e^-)$ (right) in Run 2.2 data and simulation. The smearing is applied to the simulation through through two methods: the nominal one (blue), as described in Section 6.6, and one that takes into account the dependency on occupancy of the signal shape (purple). The distribution in calibration data is shown in black, for comparison. The equivalent results from the other data-taking periods are similar, and therefore omitted.

Chapter 8

Cross-checks

Understanding the efficiencies across phase space is vital to the measurement of R_K. To this end, several cross-checks are performed. Each of them tests the experimental prodecure in different, complementary ways. This chapter covers the results of the most important tests, starting in Section 8.1 with the single ratio $r_{J/\psi}$. This quantity is a stringent test of the modelling of the efficiencies, which is further verified by the double ratio $R_{\psi(2S)}$ calculated in Section 8.2. In Section 8.3, a test of the robustness of the kinematic corrections is presented. This is followed by Section 8.4, which covers the validation of the systematic uncertainty assigned due to imperfections in the simulated material budget. Finally, Section 8.5 demonstrates that the selection suppresses particular backgrounds as expected.

The procedure employed to obtain the results from the first two sections was developed by Dr. Paula Álvarez Cartelle and Dr. Thibaud Humair for the previous R_K measurement. They used it to obtain all Run 1 and Run 2.1 results presented in these sections. In addition, the calculation of $R_{\psi(2S)}$ was conducted jointly with Davide Lancierini. Everything else in this chapter represents original work.

8.1 The single ratio $r_{J/\psi}$

By grouping the terms in Equation (4.4) based on whether they relate to the rare or the control mode, R_K can be expressed as:

$$R_K = \left(\frac{N(K^+\mu^+\mu^-)}{\varepsilon(K^+\mu^+\mu^-)} \cdot \frac{\varepsilon(K^+e^+e^-)}{N(K^+e^+e^-)} \right) \Big/ \, r_{J/\psi} \,, \tag{8.1}$$

where the parameter $r_{J/\psi}$ is defined as:

$$\begin{aligned} r_{J/\psi} &= \frac{N(B^+ \to K^+J/\psi(\mu^+\mu^-))}{\varepsilon(B^+ \to K^+J/\psi(\mu^+\mu^-))} \cdot \frac{\varepsilon(B^+ \to K^+J/\psi(e^+e^-))}{N(B^+ \to K^+J/\psi(e^+e^-))} && (8.2) \\ &= \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(J/\psi \to e^+e^-)} \,. && (8.3) \end{aligned}$$

Since this quantity can be obtained purely from the control channels, it is useful in verifying efficiencies before unblinding R_K. In addition, $r_{J/\psi}$ is known to respect LFU down to a few permille, as evidenced by the estimate obtained from the branching fractions listed in the PDG [9]:

$$r_{J/\psi}^{\text{PDG}} = 0.998 \pm 0.008. \tag{8.4}$$

Note that $r_{J/\psi}$ is a single ratio, and hence it does not benefit from the same cancellation of muon-electron detection differences that double ratios benefit from. This makes $r_{J/\psi}$ a stringent test of efficiencies: this quantity is measured accurately only if the electron-channel efficiencies are under control with respect to their muon counterparts.

Estimates of $r_{J/\psi}$ are computed separately for each data-taking period and electron trigger selection in the analysis. To check the correlations between the different selections, these individual $r_{J/\psi}$ results are averaged in several ways, as described in Section 8.1.1. Then, the modelling of the efficiency corrections across phase space is validated by recomputing $r_{J/\psi}$ in bins of variables relevant to the kinematics and detector response; this is covered by Sections 8.1.2 and 8.1.3.

8.1.1 Integrated value of $r_{J/\psi}$

Estimates of $r_{J/\psi}$ are obtained using Equation (8.2) to combine $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ yields and efficiencies. This is done separately for the 4 major data-taking periods, and for the 3 electron triggers, thus leading to 12 estimates in total. All individual results, alongside any combinations that take into account correlations and uncertainties, are expected to agree with $r_{J/\psi}^{\mathrm{PDG}}$. The combinations are weighted means of N individual values, where the weights $\{w^i\}_{i=\overline{1,N}}$ are assigned based on the systematic uncertainties and correlations of the $r_{J/\psi}$ estimates. The weights are calculated using the covariance matrix V (derived by means covered by Chapter 7) that quantifies the systematic effects induced by efficiencies:

$$w^i = \frac{\sum\limits_{j=1}^{N} V^{-1}_{i,j}}{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} V^{-1}_{i,j}}\,. \tag{8.5}$$

This means that a combination of N estimates of $r_{J/\psi}$ has the central value:

$$\overline{r_{J/\psi}} = \sum_{i=1}^{N} w^i r^i_{J/\psi} = \frac{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} V^{-1}_{i,j} r^i_{J/\psi}}{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} V^{-1}_{i,j}}\,, \tag{8.6}$$

and the uncertainty:

$$\sigma(\overline{r_{J/\psi}}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w^i w^j V_{i,j}\,. \tag{8.7}$$

The individual measurements of $r_{J/\psi}$, for each data-taking period and electron trigger, are presented in the first column of Table 8.1. Alongside them, various combinations across data-taking periods are also listed. In particular, the average value of all twelve individual estimates is found to be:

$$r_{J/\psi} = 0.981 \pm 0.020\,. \tag{8.8}$$

This value is in good agreement with $r_{J/\psi}^{\mathrm{PDG}}$. As shown in Table 8.1, the individual estimates, as well as the combinations, are also compatible with the PDG value. This is good indication that the efficiencies are well understood.

Table 8.1: Integrated values of $r_{J/\psi}$ and $R_{\psi(2S)}$, as obtained from the four data-taking periods and the three electron strategies. Also listed are the results of several combinations. The error on $r_{J/\psi}$ is systematic, whereas for $R_{\psi(2S)}$ the first uncertainty is statistical, and the second is due to systematic effects that also impact R_K. The two $R_{\psi(2S)}$ columns correspond to the two q^2 selections presented in the main body.

	$\boldsymbol{r_{J/\psi}}$	$\boldsymbol{R_{\psi(2S)}}$ narrow	$\boldsymbol{R_{\psi(2S)}}$ wide
Run 1			
*e*TOS	1.063 ± 0.064	0.999 ± 0.017 ± 0.016	1.011 ± 0.018 ± 0.016
*h*TOS!	1.008 ± 0.071	0.970 ± 0.066 ± 0.042	1.002 ± 0.070 ± 0.044
TIS!	1.015 ± 0.070	1.078 ± 0.035 ± 0.033	1.062 ± 0.035 ± 0.033
Run 1 combination			
	1.034 ± 0.038	1.004 ± 0.015 ± 0.016	1.016 ± 0.016 ± 0.016
Run 2.1			
*e*TOS	1.052 ± 0.063	0.995 ± 0.013 ± 0.012	0.979 ± 0.014 ± 0.012
*h*TOS!	1.053 ± 0.090	0.932 ± 0.060 ± 0.045	0.974 ± 0.065 ± 0.048
TIS!	1.112 ± 0.092	0.992 ± 0.027 ± 0.023	0.982 ± 0.028 ± 0.023
Run 2.1 combination			
	1.066 ± 0.056	0.991 ± 0.012 ± 0.011	0.979 ± 0.012 ± 0.011
Run 1 + Run 2.1 combination			
	1.014 ± 0.036	0.995 ± 0.009 ± 0.011	0.989 ± 0.009 ± 0.011
2017			
*e*TOS	0.981 ± 0.046	0.984 ± 0.013 ± 0.011	0.991 ± 0.014 ± 0.011
*h*TOS!	0.946 ± 0.064	0.958 ± 0.055 ± 0.035	1.037 ± 0.064 ± 0.038
TIS!	1.000 ± 0.032	1.020 ± 0.027 ± 0.023	1.031 ± 0.028 ± 0.023
2017 combination			
	0.994 ± 0.017	0.987 ± 0.012 ± 0.011	0.998 ± 0.012 ± 0.011
2018			
*e*TOS	0.992 ± 0.048	0.991 ± 0.012 ± 0.010	1.003 ± 0.013 ± 0.010
*h*TOS!	0.946 ± 0.064	0.991 ± 0.050 ± 0.028	1.039 ± 0.054 ± 0.029
TIS!	0.983 ± 0.031	1.031 ± 0.024 ± 0.019	1.030 ± 0.024 ± 0.019
2018 combination			
	0.984 ± 0.027	0.997 ± 0.011 ± 0.009	1.008 ± 0.011 ± 0.009
Run 2.2 combination			
	0.994 ± 0.024	0.992 ± 0.008 ± 0.009	1.003 ± 0.008 ± 0.009
Total combination			
	0.981 ± 0.020	0.992 ± 0.006 ± 0.008	0.997 ± 0.006 ± 0.009

8.1.2 Single-variable dependencies in $r_{J/\psi}$

Although the $r_{J/\psi}$ estimates in Table 8.1 indicate that efficiencies are well understood, they are not sensitive to effects that cancel when integrated over phase space. Mismodelling can occur, for example, if efficiencies of low-p_{T} electron events are overestimated, whilst high-p_{T} electron efficiencies are underestimated in such a way that the effect cancels out when integrated across the p_{T} spectrum. Since some variables do not have the same distribution in signal and control samples, the cancellation that would happen at resonant q^2 may not occur in the signal region. This could then compromise the application of efficiency corrections from the control channel to the signal modes.

Thanks to the high statistics of the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ samples, the $r_{J/\psi}$ test can be adapted to check for such effects. This is done by splitting the range of a given variable into several bins, and computing $r_{J/\psi}$ in each bin individually[1]. In a given bin of index i, this is done based on Equation (8.2) as follows:

1. the estimated $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ yield is obtained from a fit to the data in the bin;
2. the efficiency of control-mode events falling into that bin is evaluated using fully-calibrated simulation;
3. steps 1 and 2 are repeated for each bin, using both electron and muon control samples, thus obtaining the efficiency-corrected yields $\mathcal{N}_i^e$ and $\mathcal{N}_i^\mu$ in each bin;
4. the resulting $r_{J/\psi}$ estimate is then $r^i_{J/\psi} = \mathcal{N}_i^\mu / \mathcal{N}_i^e$.

This is referred to as the "1D $r_{J/\psi}$" check. Ideal efficiency corrections would lead to identical $r^i_{J/\psi}$ in each bin of the studied variable. In other words, $r_{J/\psi}$ would be flat in that variable. Since the corrections are not expected to be perfect, some deviations from flatness are expected. The problematic cases are those where the double ratio does not guarantee perfect cancellation between signal and control channels. This happens when the distribution of the binned variable differs between signal and control samples, and deviations from flatness are due to genuine mismodelling of the efficiency as a function of the studied variable. As a result, the estimated signal efficiencies in the different bins may not necessarily reflect the actual yields. Since ideally the ratio between the expected signal and control yields would be equal to the ratio between their respective efficiencies, any deviations from unity of the quantity

$$d_f^\ell = \frac{\sum_i N_i^{\text{rare}}}{\sum_i N_i} \Bigg/ \frac{\sum_i \varepsilon_i^{\text{rare}}}{\sum_i \varepsilon_i} \tag{8.9}$$

---

[1] The binning scheme is optimised such that equivalent populations are expected in each bin.

would point towards an effect that does not cancel in the double ratio. Here, N_i and ε_i are the binned $B^+ \to K^+J/\psi(\ell^+\ell^-)$ yields and efficiencies obtained in steps 1 and 2, whilst N_i^{rare} and $\varepsilon_i^{\text{rare}}$ are their $B^+ \to K^+\ell^+\ell^-$ counterparts. Note that here the signal yield is an estimate based on the control-mode, *i.e.* $N_i^{\text{rare}} = \varepsilon_i^{\text{rare}}\mathcal{N}_i$. The quantity d_f^ℓ can be computed for both electrons and muons, leading to the definition of the following observable:

$$d_f = \left(\frac{\sum\limits_i \varepsilon_{i,\mu}^{\text{rare}} \cdot \mathcal{N}_i^{\mu}}{\sum\limits_i \varepsilon_{i,\mu} \cdot \mathcal{N}_i^{\mu}} \cdot \frac{\sum\limits_i \varepsilon_{i,\mu}}{\sum\limits_i \varepsilon_{i,\mu}^{\text{rare}}} \right) \Bigg/ \left(\frac{\sum\limits_i \varepsilon_{i,e}^{\text{rare}} \cdot \mathcal{N}_i^{e}}{\sum\limits_i \varepsilon_{i,e} \cdot \mathcal{N}_i^{e}} \cdot \frac{\sum\limits_i \varepsilon_{i,e}}{\sum\limits_i \varepsilon_{i,e}^{\text{rare}}} \right) - 1\,. \tag{8.10}$$

This is called the flatness parameter. It encodes the impact on R_K of $r_{J/\psi}$ fluctuations across bins, assuming these fluctuations are due to genuine mismodelling of the efficiencies in terms of the binned variable. For every $r_{J/\psi}$ test, this flatness parameter is required to be comparable, if not smaller than, the expected systematic uncertainty on R_K. If that is not the case, it has to be taken into account as a separate systematic effect.

A selection of results from the 1D $r_{J/\psi}$ check is presented in Figure 8.1. The binned variables describe the decay kinematics, such as the transverse momenta of the leptons. This makes the selected variables particularly important for the correct calibration of efficiencies. The plots in Figure 8.1 use 2018 eTOS and μTOS events. The equivalent results with electron samples from hTOS! and TIS! are presented in Appendix E. Also shown in this appendix are the results from 2017 data, which are qualitatively similar to the 2018 ones. Note that this test targets the flatness of $r_{J/\psi}$ across bins, and hence the figure of merit is d_f. The value to which the bins average is not important, and so the bins in each check are normalised to the average $r_{J/\psi}$ value across phase space, denoted by $\overline{r_{J/\psi}}$.

For the shown variables, the ratio $r_{J/\psi}$ is found to exhibit little to no variation between bins. This is a clear indication that the efficiency dependencies on these variables are modelled correctly. However, trends are observed in two other places. The first place is in the χ^2_{DV} of the B^+, as shown in Figure 8.2. This is one of the variables whose simulated distributions are calibrated using the kinematic corrections described in Section 6.5. The systematic effect induced by the calibration method is evaluated in Section 7.2, using alternative correction

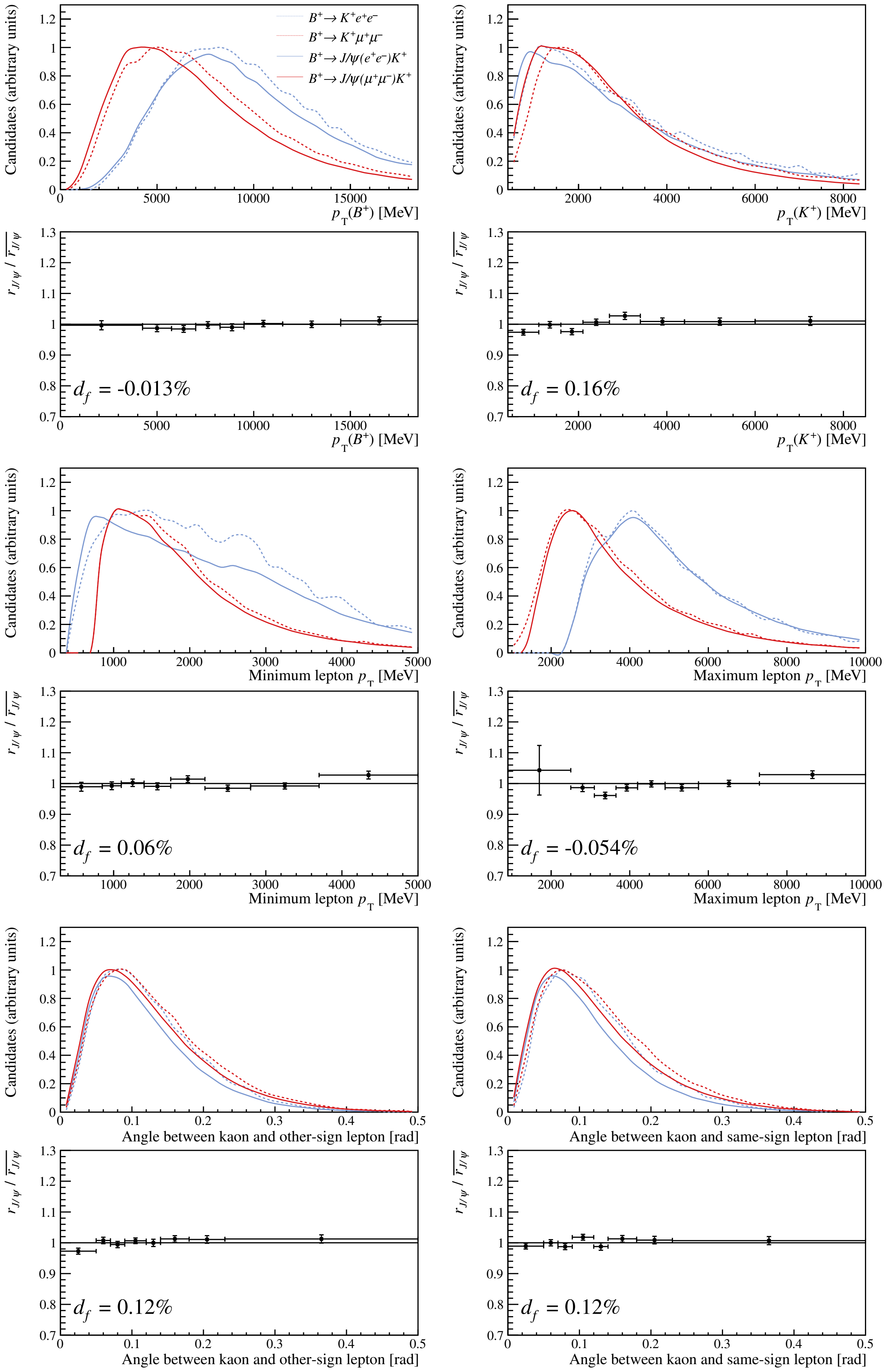

Figure 8.1: Plots of $r_{J/\psi}$ obtained from 2018 eTOS and μTOS data, as a function of variables relevant to the detector response. The corresponding flatness parameters are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control samples. The distributions in electron and muon control samples are depicted in blue and red solid lines, respectively. Their signal-mode counterparts are shown in blue and red dotted lines, respectively. The uncertainties shown are purely statistical.

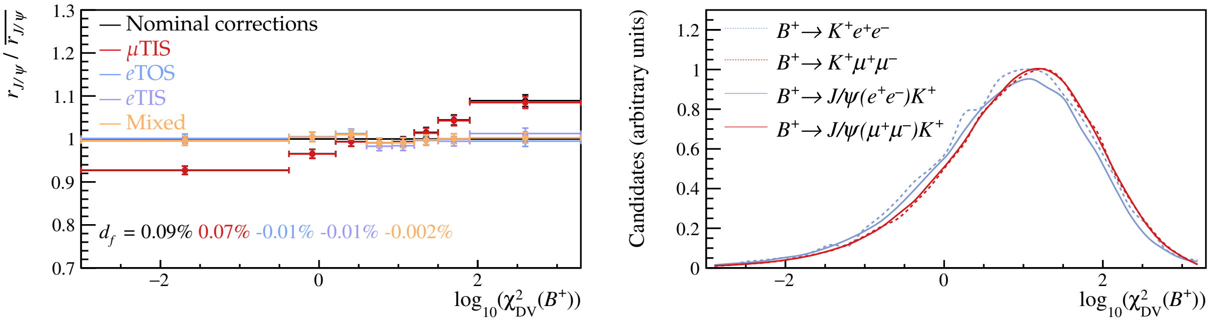

Figure 8.2: Plot of $r_{J/\psi}$ as a function of the $\chi^2_{\rm DV}$ of the B^+ (left), obtained from 2018 eTOS and μTOS data. The black bins show the result based on the nominal efficiency corrections. The coloured bins represent the alternative schemes described in Section 7.2, which are used to assess the systematic uncertainty due to calibration of the kinematics. The corresponding flatness parameters are displayed at the bottom, in matching colours. The kinematic distributions of the binned variable, in simulated samples, are shown on the right. The distributions in electron and muon control samples are depicted in blue and red solid lines, respectively. Their signal-mode counterparts are shown in blue and red dotted lines, respectively.

strategies. To check whether the trend in $r_{J/\psi}$ is encompassed by the systematic uncertainty, the binned exercise is repeated using the alternative kinematic correction strategies. The results are represented as coloured bins in Figure 8.2. All flatness parameters are found to be well below the permille-level, the smallest of them resulting from the strategy that explicitly takes into account differences in $\chi^2_{\rm DV}$ between muon samples and electron samples. This means that the trend in $r_{J/\psi}$ is covered by the systematic uncertainty assigned due to kinematic calibration strategy.

The other place where trends are observed is in variables related to the pseudorapidity of the leptons, as shown in Figure 8.3. As explained in Section 7.7, the material budget of the detector is not expected to be modelled perfectly by the simulation. To check the effect on R_K of these imperfections, weights are derived to correct the simulated pseudorapidity distributions of electrons, to match the data. The impact on R_K is found to be small, thanks to the cancellation inherent to the double ratio. Hence, the pseudorapidity corrections are not used nominally when computing efficiencies. However, because the imperfect modelling of the material budget affects electrons more than muons, the single ratio $r_{J/\psi}$ is more susceptible to the effect of these corrections. For this reason, the 1D $r_{J/\psi}$ tests shown in Figure 8.3 are conducted both with and without the η corrections. It is found that applying them improves the values taken by d_f. In particular, the flatness parameter associated to the angle between the two leptons, $\alpha_{\ell^+\ell^-}$, drops from nearly 2% to a few permille after the application of the η corrections. This reduction suggests that the effect is not a genuine mismodelling of the efficiencies, but a reflection of the imperfections in the simulated pseudorapidity distributions. Since the definition of d_f assumes any deviation from flatness is due to genuine mismodelling in the variable under study, the flatness parameter cannot be used in this particular case to assess the effect on R_K. The associated systematic uncertainty is evaluated based on the η corrections, and validated via the check described in Section 8.4.

In summary, the modelling of efficiencies across phase space is tested by computing $r_{J/\psi}$ as a function of variables relevant to the detector response. The distribution of $r_{J/\psi}$ in each of these variables is generally found to be flat. The small departures from flatness are most likely statistical in nature. If, instead, these deviations are assumed to stem from genuine mismodelling of the efficiencies, the systematic effect on R_K is evaluated through the flatness parameter d_f. All results for d_f are found to be at the permille-level at most. The largest effects are discussed in detail and found to be within the associated systematic uncertainties assigned in Chapter 7. This gives confidence that the efficiencies are suitably modelled.

8.1.3 Double-variable dependencies in $r_{J/\psi}$

The 1D $r_{J/\psi}$ test covered by the previous subsection is extended to two dimensions, in order to validate the modelling of efficiency corrections in terms of variables — defined in the detector reference frame — that describe a $B^+ \to K^+\ell^+\ell^-$ event. These variables are illustrated in Figure 8.4. The momentum of the B^+ is encoded by three parameters: its magnitude (p_{B^+}), the polar angle (α_{B^+}), and the azimuthal angle (ϕ_{B^+}). The decay of the B^+ into a kaon and a dilepton system can be fully described by three variables: the dilepton invariant mass (q^2) and polar angle ($\phi_{\ell^+\ell^-}$), along with the angle between the kaon and the dilepton (α_{K^+}). Parametrising the individual leptons requires two further kinematic variables: the angle between the two ($\alpha_{\ell^+\ell^-}$), and the polar coordinate of one of the leptons (ϕ_ℓ). In summary, a $B^+ \to K^+\ell^+\ell^-$ event can be described using a set of 8 variables:

$$A_0 = \left\{p_{B^+},\ \alpha_{B^+},\ \phi_{B^+},\ \alpha_{K^+},\ \phi_{\ell^+\ell^-},\ q^2,\ \alpha_{\ell^+\ell^-},\ \phi_\ell\right\} . \tag{8.11}$$

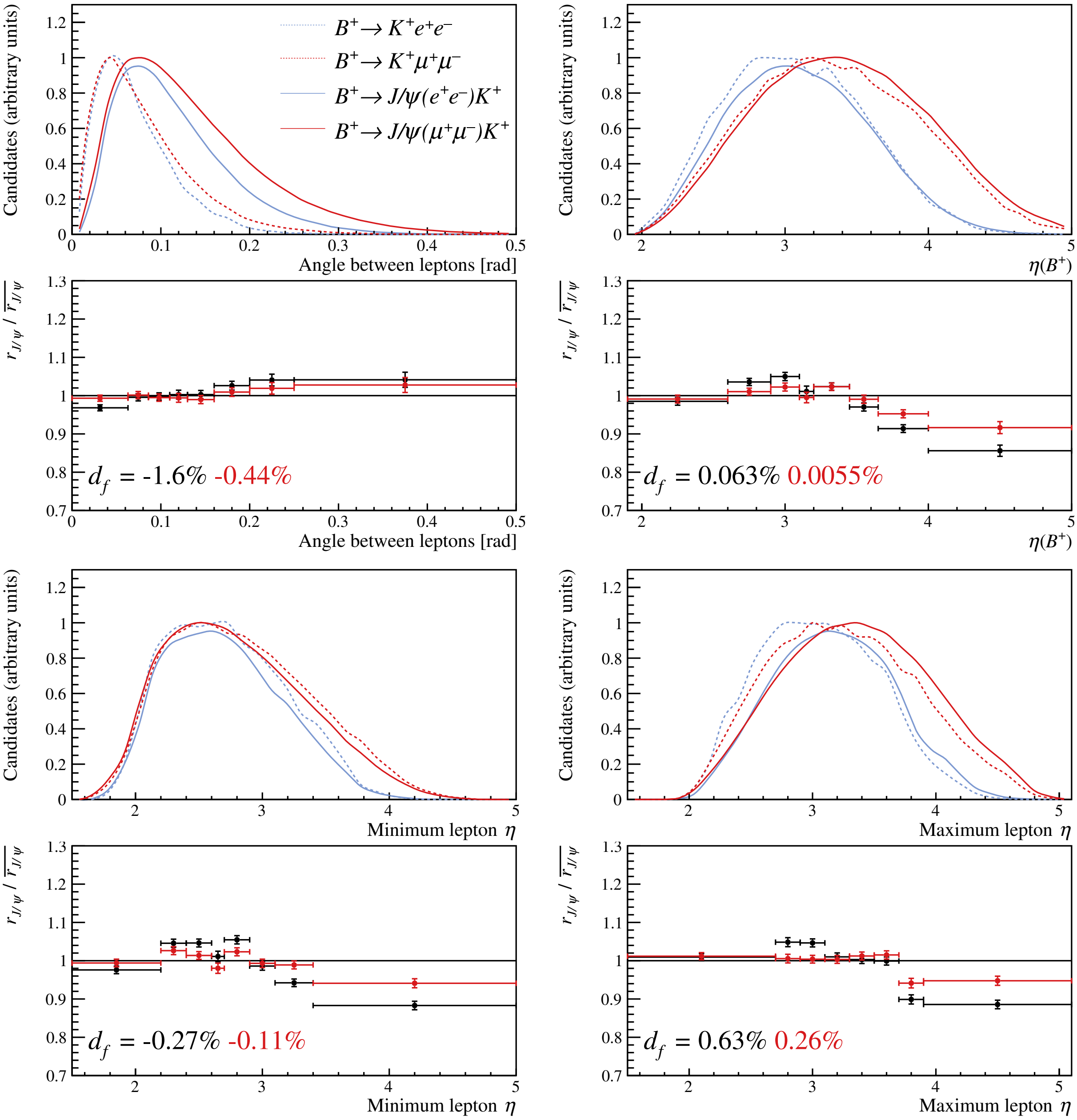

Figure 8.3: Plot of $r_{J/\psi}$ as a function of variables related to lepton pseudorapidity, obtained from 2018 eTOS and μTOS data. The black bins show the result based on the nominal efficiency corrections. The red bins show how the result changes as a result of the weights derived in Section 7.7 to correct the simulated distributions of electron pseudorapidity. The corresponding flatness parameters are displayed at the bottom of each plot, in matching colours. At the top, the kinematic distributions of the binned variable are shown. The distributions in electron and muon control samples are depicted in blue and red solid lines, respectively. Their signal-mode counterparts are shown in blue and red dotted lines, respectively.

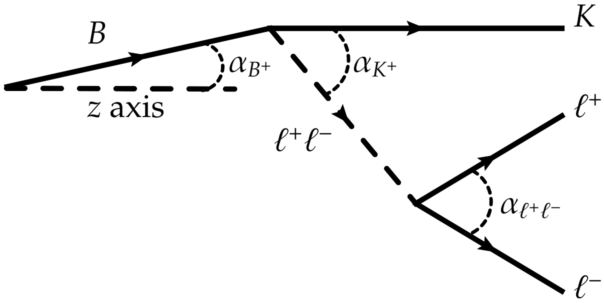

Figure 8.4: Illustration of variables used to parametrise a $B^+ \to K^+\ell^+\ell^-$ event.

The dilepton invariant mass is related to the momenta of the leptons and the angle between their tracks: $q^2 = 2\,p_{\ell^+}\,p_{\ell^-}\,(1 - \cos\alpha_{\ell^+\ell^-})$. This means that q^2 and p_{B^+} can be replaced in the above list of variables with the maximum and minimum momenta of the two leptons, $\max \mathrm{p}_\ell$ and $\min \mathrm{p}_\ell$ respectively. Furthermore, the distribution of α_{B^+} is not related to the internal structure of the decay, and the LHCb detector can be assumed to be polar symmetric in the decay frame. This brings the list of variables describing a $B^+ \to K^+\ell^+\ell^-$ event down to four. They form the set:

$$A = \{\max \mathrm{p}_\ell,\ \min \mathrm{p}_\ell,\ \alpha_{\ell^+\ell^-},\ \alpha_{\mathrm{K}^+}\}\ . \tag{8.12}$$

In the particular case of resonant decays, q^2 is constrained to the mass of the resonance squared. This reduces the number of independent kinematic variables to three. Note that $q^2 \notin A$: although the rare and resonant modes are well separated in this variable, they exhibit good overlap in the four kinematic variables contained by the set A. This is why grouping these variables into pairs to form a two-dimensional phase space, and computing $r_{J/\psi}$ in regions of this space, provides a powerful check of the validity of the efficiency corrections. This is called the "2D $r_{J/\psi}$" check, and it follows the same procedure as the one-dimensional variant described in the previous subsection.

The four kinematic variables $\max \mathrm{p}_\ell$, $\min \mathrm{p}_\ell$, $\alpha_{\ell^+\ell^-}$, and α_{K^+} are grouped into six pairs, and the $r_{J/\psi}$ check is conducted in the phase space spanned by each combination. A 4×4 binning scheme in each pair of variables is chosen such that the expected number of $B^+ \to K^+J/\psi(e^+e^-)$ candidates in each bin is roughly similar. The estimated populations are determined using simulated $B^+ \to K^+J/\psi(e^+e^-)$ events. Like for the one-dimensional check, $B^+ \to K^+J/\psi(\ell^+\ell^-)$ yields and efficiencies are determined individually for each bin, and the flatness parameter d_f is computed to assess the systematic effect of deviations from flatness.

The results of the 2D $r_{J/\psi}$ check, conducted on 2017 eTOS and μTOS data, is shown in Figure 8.5. The equivalents from 2018 data are similar, and are hence omitted. The samples from the other electron triggers are not sufficiently high in statistics to lead to conclusive results. Since this check is interested in the flatness of $r_{J/\psi}$, rather than the average, the normalisation is not important. For convenience, the $r_{J/\psi}$ values are normalised to the average across phase space. All flatness parameters are generally found to be at most at the permille-level, with the exception of those where the dilepton angle, $\alpha_{\ell^+\ell^-}$, is one of the variables. As discussed in the previous subsection, this is not an effect of mismodelling in $\alpha_{\ell^+\ell^-}$. Instead, it is due to imperfections in the simulated electron pseudorapidity. If these are corrected using calibration weights, the corresponding flatness parameters decrease below the permille-level, thus indicating that the systematic uncertainty calculated in Section 7.7 covers the trend in $\alpha_{\ell^+\ell^-}$.

To summarise, the modelling of efficiencies is further tested by computing $r_{J/\psi}$ in two dimensions. Like in the 1D variant of this check, all flatness parameters are found to be at the permille-level at most, notwithstanding the already-discussed effect on $\alpha_{\ell^+\ell^-}$ of imperfections in the simulation of η. This strengthens the confidence in the modelling of the efficiencies.

8.2 The double ratio $R_{\psi(2S)}$

The tests conducted on $r_{J/\psi}$ prove that both muon and electron efficiencies are well understood across phase space. Although q^2 separates the signal and resonant channels, the detector performance itself depends on other kinematic variables, such as those that make up the set A defined in Equation (8.12). In other words, q^2 is not a variable on which efficiencies depend directly, and since the different modes overlap well in variables relevant to the detector response, the efficiencies derived from the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ modes are expected to be modelled accurately across phase space, including across q^2.

The portability of efficiency corrections across q^2 cannot be tested using selected $B^+ \to K^+\ell^+\ell^-$ candidates before unblinding R_K. The rare-mode efficiencies could in principle be cross-checked using control-mode events in the corresponding q^2 ranges, as is done in Figure 6.15. However, a more powerful test of the modelling of efficiencies across q^2 is possible at the

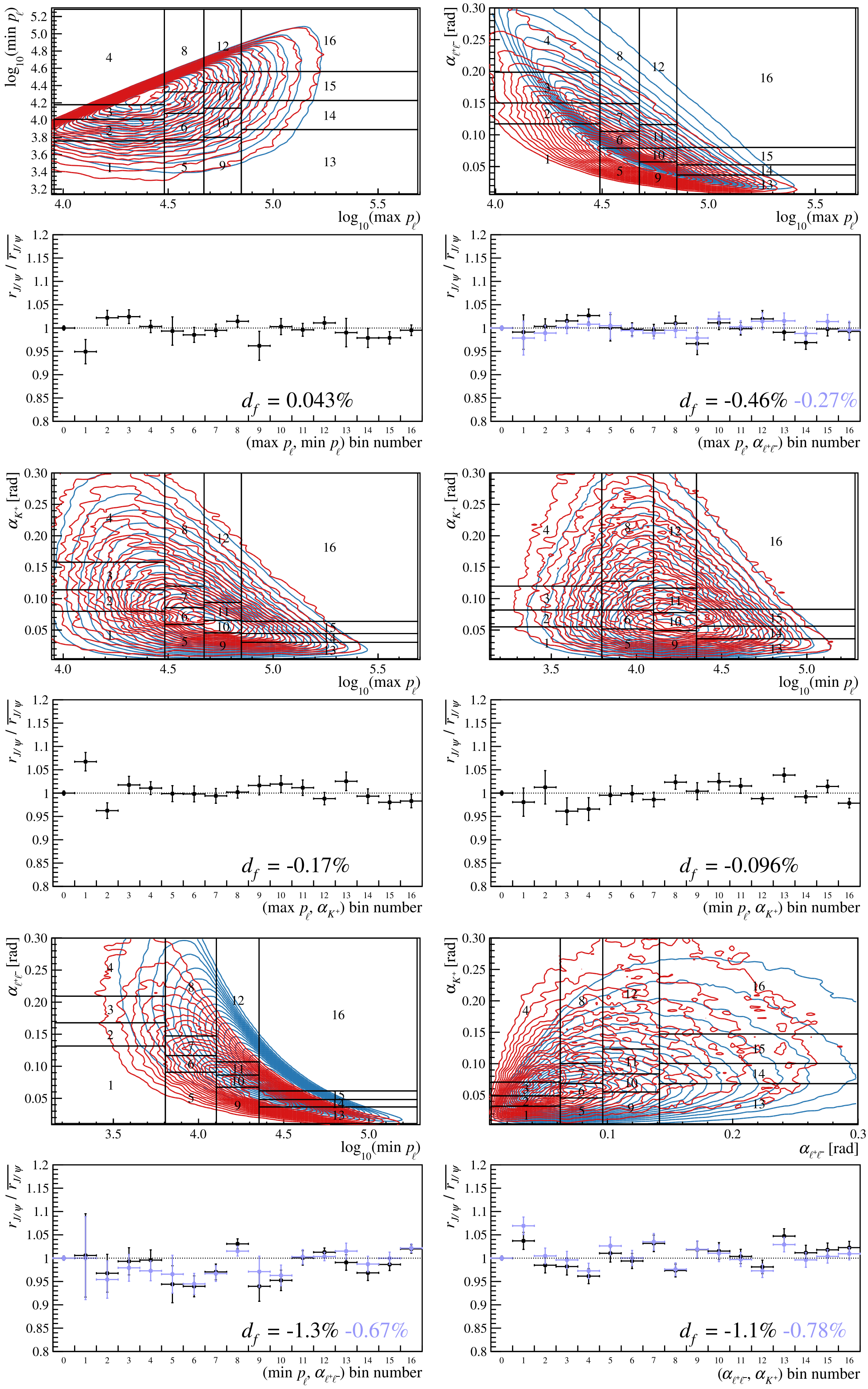

Figure 8.5: The ratio $r_{J/\psi}$ computed in two-dimensional bins of kinematic variables, normalised to the average value across phase space. The uncertainties are purely statistical. The 0th bin corresponds to the normalisation and is hence 1 by definition. The binning scheme is shown at the top of each plot, alongside distributions of simulated $B^+ \to K^+e^+e^-$ (red) and $B^+ \to K^+J/\psi(e^+e^-)$ (blue) events. The flatness parameters from each test are shown at the bottom of the plots. Where one of the variables is the dilepton angle, $\alpha_{\ell^+\ell^-}$, the exercise is repeated with the addition of the η corrections discussed in the main body; this is shown in purple.

$\psi(2S)$ resonance. For this reason, an observable similar to R_K is constructed using the same procedure, with the exception that the q^2 and $m(K^+\ell^+\ell^-)$ selection ranges are different. This amounts to replacing the $B^+ \to K^+\ell^+\ell^-$ numerator in the definition of R_K with its $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ equivalent. The resulting observable is:

$$\begin{aligned}
R_{\psi(2S)} &= \frac{\mathcal{B}(B^+ \to K^+\psi(2S)(\mu^+\mu^-))}{\mathcal{B}(B^+ \to K^+\psi(2S)(e^+e^-))} \Bigg/ \frac{\mathcal{B}(B^+ \to K^+J/\psi(\mu^+\mu^-))}{\mathcal{B}(B^+ \to K^+J/\psi(e^+e^-))} && (8.13)\\
&= \frac{N(B^+ \to K^+\psi(2S)(\mu^+\mu^-))}{\varepsilon(B^+ \to K^+\psi(2S)(\mu^+\mu^-))} \cdot \frac{\varepsilon(B^+ \to K^+\psi(2S)(e^+e^-))}{N(B^+ \to K^+\psi(2S)(e^+e^-))}\\
&\cdot \frac{\varepsilon(B^+ \to K^+J/\psi(\mu^+\mu^-))}{N(B^+ \to K^+J/\psi(\mu^+\mu^-))} \cdot \frac{N(B^+ \to K^+J/\psi(e^+e^-))}{\varepsilon(B^+ \to K^+J/\psi(e^+e^-))}\,. && (8.14)
\end{aligned}$$

This is designed to check the portability of the efficiency corrections into q^2 ranges other than those corresponding to the control-mode windows. For this reason, only systematic effects that are relevant to R_K efficiencies are assessed for $R_{\psi(2S)}$. Their cancellation is checked through the fact that $R_{\psi(2S)}$ is a double ratio.

The $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ yields are obtained from fits to data, for each data-taking period and trigger selection separately. These fits were performed by Davide Lancierini, who also improved the background model with respect to the previous R_K analysis. The results are shown in Appendix F. The efficiencies are calculated using $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ simulated events, after applying the full calibration chain. Estimates of $R_{\psi(2S)}$ are obtained based on $B^+ \to K^+\psi(2S)(e^+e^-)$ candidates selected in the "wide" range $q^2 \in (9.92\ \text{GeV}^2, 16.40\ \text{GeV}^2)$. The q^2 cut is expected to affect some fit components more than others, and so the background model is validated by recomputing $R_{\psi(2S)}$ after tightening the selection to the "narrow" range $q^2 \in (11.22\ \text{GeV}^2, 16.40\ \text{GeV}^2)$. The results are presented in Table 8.1. Like with $r_{J/\psi}$, estimates of $R_{\psi(2S)}$ are obtained separately for each data-taking period and electron trigger, and then combined in several ways. All Run 2.2 results represent original work, the only external contributions being the $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ yields obtained by Davide Lancierini. Conversely, all combinations from Table 8.1 that include estimates for Run 2.2 also constitute original work. The combinations are based on the $R_{\psi(2S)}$ covariance matrix $V^{R_{\psi(2S)}}$, obtained by means described in Chapter 7, together with the statistical uncertainties arising from the fits. More specifically, the matrix $C^{R_{\psi(2S)}}$ used to assign weights to the values in a given combination,

$$w^i = \frac{\sum_{j=1}^{N} \left(C_{i,j}^{R_{\psi(2S)}}\right)^{-1}}{\sum_{i=1}^{N}\sum_{j=1}^{N} \left(C_{i,j}^{R_{\psi(2S)}}\right)^{-1}}\,, \tag{8.15}$$

is given by:

$$C^{R_{\psi(2S)}} = V^{R_{\psi(2S)}} + \mathrm{diag}(\delta_{\mathrm{i}}^2)\,. \tag{8.16}$$

Here, δ_i is the uncertainty induced by the fits to data on the i^{th} $R_{\psi(2S)}$ estimate used in the combination. In particular, when averaging over the entire data set, the resulting $R_{\psi(2S)}$ values from the "wide" and "narrow" q^2 ranges are:

$$R_{\psi(2S)}^{\text{wide}} = 0.997 \pm 0.006 \pm 0.009\,, \tag{8.17}$$

$$R_{\psi(2S)}^{\text{narrow}} = 0.992 \pm 0.006 \pm 0.008\,. \tag{8.18}$$

The first uncertainty is statistical, and the second is due to systematic effects relevant to the measurement of R_K. These averages, together with the individual results and the other combinations, are compatible with the PDG average, $R_{\psi(2S)}^{\text{PDG}} = 0.991 \pm 0.078$ [9]. Given this excellent agreement, the efficiency corrections extracted from the control modes are found to be valid in q^2 regions other than around the J/ψ.

8.3 Calibrating B^+ kinematics via scaling

The disagreement between data and simulation concerning the distributions of the χ^2_{IP} and χ^2_{DV} of the reconstructed B^+ may be caused by imperfectly simulated resolutions. If that is the case, an alternative calibration method, different from the one described in Section 6.5, might be better motivated. This method involves scaling simulated variables, instead of assigning weights on an event-by-event basis, and is expected to lead to corrections equivalent to the nominal ones. Hence, it is used as an additional check of the validity of the kinematic weights employed to correct the efficiencies corresponding to Run 2.2 data.

The scaling method relies on the principle of inverse transform sampling, in which the cumulative density function (CDF) of a variable $k = g(x_1)$ is identical to the CDF of another variable, x_2. In the present situation, x_1 is the observable in simulation, x_2 is its counterpart in data, and $g(x_1)$ is the transformation function that maps x_1 onto a scaled variable, x_{scaled}, whose distribution matches the one of x_2. This can be represented as:

$$\underbrace{\int_{-\infty}^{x} \mathrm{d}x\, \mathcal{P}_{\text{data}}(x)}_{F_{\text{data}}(x)} = \underbrace{\int_{-\infty}^{x} \mathrm{d}x\, \mathcal{P}_{\text{sim}}(g(x))}_{F_{\text{sim}}(g(x))}\,, \tag{8.19}$$

where $x \in \{\chi^2_{\text{IP}}(B^+),\, \chi^2_{\text{DV}}(B^+)\}$, $\mathcal{P}_{\text{data}}(x)$ represents the distribution of x in data, and $\mathcal{P}_{\text{sim}}(g(x))$ (equivalent to $\mathcal{P}_{\text{sim}}(x_{\text{scaled}})$) is the distribution of the scaled variable in simulation. Their cumulative density functions are denoted by $F_{\text{data}}(x)$ and $F_{\text{sim}}(g(x))$, respectively. The scaling function is therefore given by:

$$g(x) = F^{-1}_{\text{sim}}(F_{\text{data}}(x))\,. \tag{8.20}$$

Note that the stripping selection involves cuts on both $\chi^2_{\text{IP}}(B^+)$ and $\chi^2_{\text{DV}}(B^+)$, which are listed in Table 4.2. Because of this, $\mathcal{P}_{\text{data}}$ and $\mathcal{P}_{\text{sim}}$ are parametrised using two analytic functions that are extrapolated into the region removed by the stripping. These functions are chosen based on how well they are able to capture the features of the distributions they model. The chosen parametrisations of $\chi^2_{\text{IP}}(B^+)$ and $\chi^2_{\text{DV}}(B^+)$ are:

$$\mathcal{P}^{\chi^2_{\text{IP}}(B^+)} = C\left(e^{-k_1 x} + q e^{-k_2 x}\right)\,, \text{ and, respectively} \tag{8.21}$$

$$\mathcal{P}^{\chi^2_{\text{DV}}(B^+)} = C x^n \left[e^{-k_1 x} + q\left(e^{-k_2 x} + r e^{-k_3 x}\right)\right]\,. \tag{8.22}$$

The parameters $C,\ n,\ q,\ r,\ k_1,\ k_2$, and k_3 are obtained from χ^2-minimisation fits conducted individually on Run 2.2 eTOS and μTOS data and simulation. The fits, together with the cumulative distributions, are shown in Figure 8.6. The distributions of the scaled variables, compared to the data and to the unscaled variables, are illustrated in Figure 8.7.

To quantify the effect of the scaling, the efficiencies of the rare and resonant modes are computed in two different ways. The first method uses kinematic weights to correct $p_{\text{T}}(B^+)$

and $\eta(B^+)$, as well as $\chi^2_{\mathrm{IP}}(B^+)$ and $\chi^2_{\mathrm{DV}}(B^+)$. The second method calibrates only $p_{\mathrm{T}}(B^+)$ and $\eta(B^+)$ using kinematic weights, whilst $\chi^2_{\mathrm{IP}}(B^+)$ and $\chi^2_{\mathrm{DV}}(B^+)$ are scaled as described above. The fractional differences between the efficiency estimates derived using the two methods are listed in Table 8.2. The scaling-based estimates are found to be compatible at the permille level with the results obtained from the nominal reweighting method. Moreover, the variation between the two sets of efficiencies is covered by the systematic uncertainty assigned to the kinematic calibration method. Note that this is a cross-check and not an evaluation of systematic uncertainty, since the effect induced by the kinematic calibration method is already estimated via the method described in Section 7.2. For this reason, the results of this cross-check are not included in the evaluation of the total systematic uncertainties.

Table 8.2: Fractional differences, expressed as percentages, between computing $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$ with and without scaling the simulated distributions of $\chi^2_{\mathrm{IP}}(B^+)$ and $\chi^2_{\mathrm{DV}}(B^+)$ to match the data.

	2017			2018		
	*e*TOS	*h*TOS!	TIS!	*e*TOS	*h*TOS!	TIS!
$r_{J/\psi}$	1.30	0.86	0.71	0.58	0.49	0.45
R_K	0.11	0.13	0.48	0.23	0.49	0.31
$R_{\psi(2S)}$	0.14	0.31	0.01	0.12	0.14	0.11

8.4 Validation of the material and tracking corrections

As mentioned in Section 7.7, the tracking performance of electrons is modelled imperfectly by the simulation. However, electron tracks are expected to be similar between the signal and control channels, leading to cancellation of tracking effects in the double ratio. In the R_K measurement, this cancellation is tested by calibrating the pseudorapidity distributions of simulated electrons to match the control data. This ensures that the amount of material electrons interact with, as a function of pseudorapidity, is accurately described by the simulation. Since the electron tracking performance depends on the amount of radiated bremsstrahlung energy, which in turn depends on the material budget, the pseudorapidity is a suitable proxy for assessing the impact of tracking-related discrepancies between data and simulation. As presented in Section 7.7, effects due to material and tracking are found to cancel well in the double ratio. This subsection presents an additional test of this cancellation.

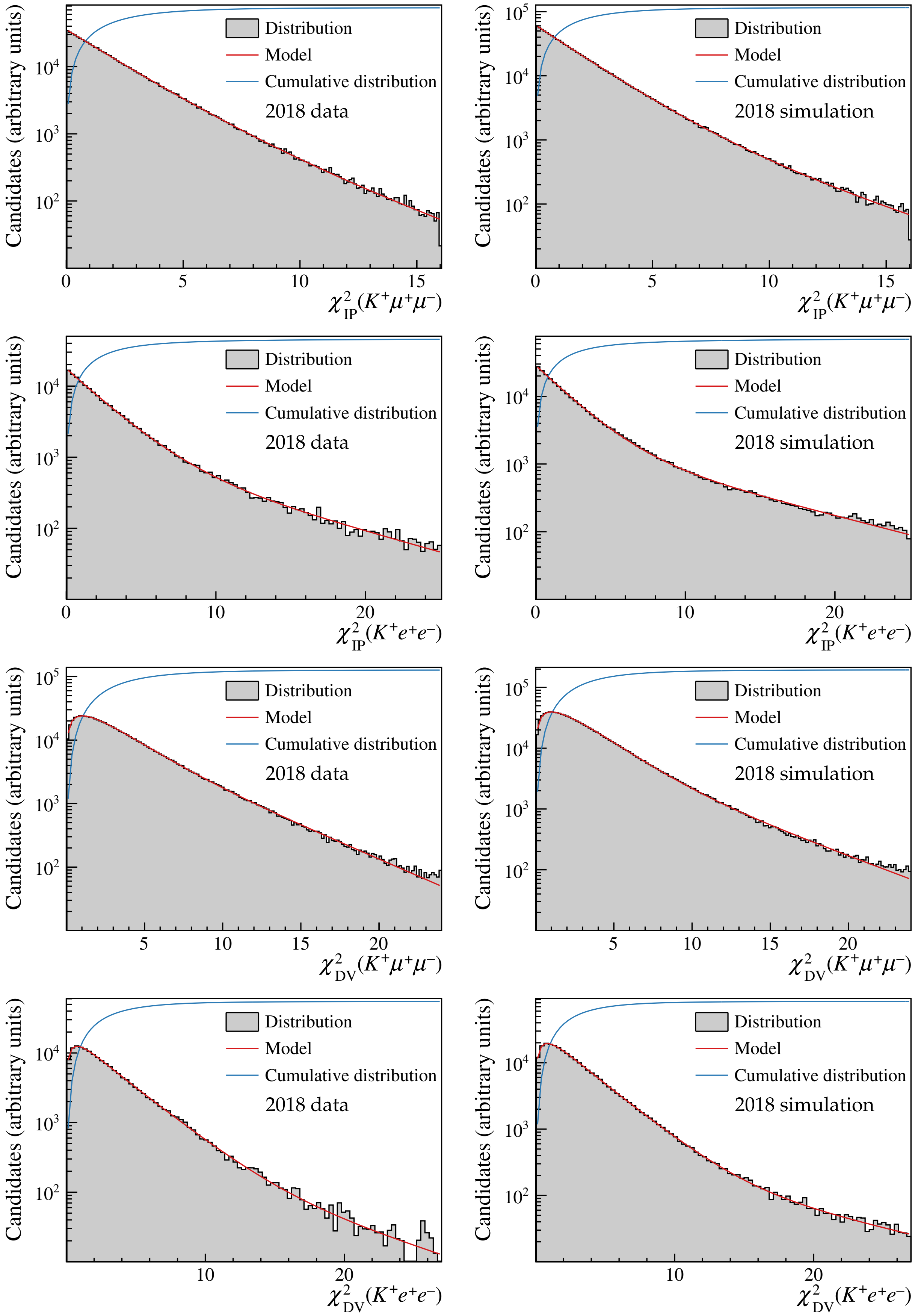

Figure 8.6: Distributions of $\chi^2_{\mathrm{IP}}(B^+)$ (top two rows) and $\chi^2_{\mathrm{DV}}(B^+)$ (bottom two rows) in 2018 μTOS and eTOS data (left) and simulation (right). The red curves represent the $\mathcal{P}^{\chi^2_{\mathrm{IP}}(B^+)}$, as obtained from fits to the distributions in black. The blue lines show the cumulative distributions of $\mathcal{P}^{\chi^2_{\mathrm{IP}}(B^+)}$.

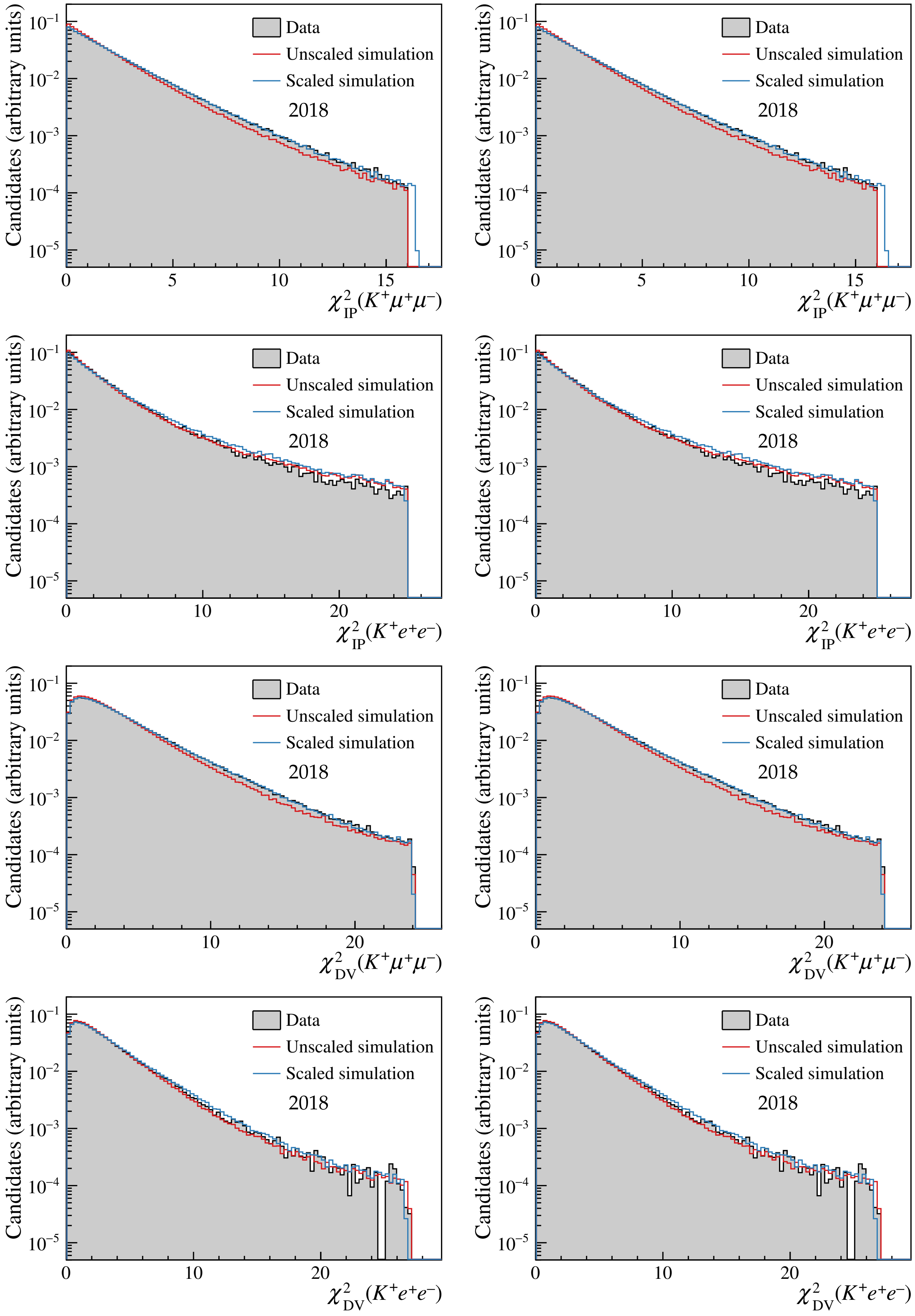

Figure 8.7: Comparison between 2018 $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ data (black) and simulation before (red) and after (blue) scaling the $\chi^2_{\rm IP}$ and $\chi^2_{\rm DV}$ of the reconstructed B^+.

Table 8.3: Fractional differences, expressed as percentages, between computing $r_{J/\psi}$, R_K, and $R_{\psi(2S)}$ with and without tracking corrections.

	Run 1			Run 2.1			2017			2018		
	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!	eTOS	hTOS!	TIS!
$r_{J/\psi}$	1.98	2.67	2.73	2.08	2.10	2.21	1.36	1.68	1.82	2.22	2.19	2.23
R_K	0.12	0.47	0.14	0.01	0.65	0.33	0.16	0.21	0.19	0.01	0.64	0.22
$R_{\psi(2S)}$	0.11	0.78	0.48	0.01	0.14	0.22	0.21	0.21	0.15	0.03	0.29	0.21

This study employs a data-driven method [194] to measure the tracking and reconstruction efficiency of electrons, in both data and simulation. This information leads to calibration of simulated samples through correction weights assigned on an event-by-event basis. The effect of these weights on efficiencies, and by extension on the ratios $r_{J/\psi}$, $R_{\psi(2S)}$, and R_K, is studied following the procedure described in Chapter 7. The ratios are calculated using tracking corrections, and are compared with the nominal results (for which the tracking is not calibrated). The fractional differences are presented in Table 8.3, where cancellation down to the permille-level is observed in the double ratios. Therefore, the impact on R_K of tracking effects is well covered by the systematic uncertainties calculated in Section 7.7.

The details of this data-driven method were yet to be finalised when the R_K measurement was ready to unblind. In particular, it was not possible to align the selection of the calibration samples to that used by the R_K analysis. For this reason, the correction weights derived here are not part of the nominal efficiency calibration, being used only to check the coverage of the systematic effect calculated using the corrections to the pseudorapidity distributions.

8.5 Contamination due to double misidentification

The cuts presented in Sections 4.3.1, 4.3.4, and 4.3.6 reduce misidentified backgrounds to negligible levels. In the case of double misidentification, where both signal leptons are in fact pions, the estimated efficiencies and the branching fractions listed in the PDG [9] indicate that the contamination is expected to be at the level of only a few permille. The branching fraction used for the $B^+ \to K^+\pi^+\pi^-$ estimate is $\mathcal{B}(B^+ \to K^+\pi^+\pi^-) = \left(1.63\,^{+0.21}_{-0.15}\right) \times 10^{-5}$, which corresponds to the non-resonant component. This is motivated by the fact that the mass vetoes and the q^2 selection are expected to reject resonant $B^+ \to K^+\pi^+\pi^-$ events. Since contamination levels are similar in the muon and the electron channels, the overall impact of $B^+ \to K^+\pi^+\pi^-$ events on R_K is expected to be negligible. This is further tested through several checks, presented in this section. The first one was conducted in collaboration with Dr. Konstantinos Petridis, and the second represents original work.

The first test validates the branching fraction value used to calculate the estimated contamination. The differential $B^+ \to K^+\pi^+\pi^-$ branching fraction, integrated over the phase space region relevant to the R_K measurement, is obtained from the model employed by the $B^+ \to K^+\pi^+\pi^-$ simulation. This is referred to as the effective branching fraction, and the result is $\mathcal{B}^{\text{eff}}(B^+ \to K^+\pi^+\pi^-) = (3.83 \pm 0.42) \times 10^{-6}$. This is compared to the value obtained using the amplitude measurement from the BaBar collaboration [195], as implemented in `Laura++` [196]. This data-driven method finds the effective branching fraction to be $\mathcal{B}^{\text{eff}}(B^+ \to K^+\pi^+\pi^-) = (3.26 \pm 0.15) \times 10^{-6}$. The values from the simulated and the data-driven models are found to be compatible, the former being larger by a few percent. Since this is the one used for the contamination estimate, it is not expected to lead to an underestimation of the $B^+ \to K^+\pi^+\pi^-$ background levels.

The second test targets the size of the effect in the electron sample, where the contamination is expected to be larger than in the muons. The fully-selected $B^+ \to K^+e^+e^-$ candidates are reconstructed by changing the mass hypothesis of each electron to that of the pion; this is only possible after unblinding. The new mass hypothesis leads to the invariant mass $m(K^+e^+_{[\to\pi^+]}e^-_{[\to\pi^-]})$, whose shape in data is compared to the expectation from simulation. To do this, simulated $B^+ \to K^+e^+e^-$, $B^+ \to K^+\pi^+\pi^-$, and $B^0 \to K^{*0}e^+e^-$ events are used to obtain the expected shapes, which are scaled to their estimated contributions as determined from their yields and efficiencies. It is checked whether the simulated model is able to suitably describe the data, with or without the estimated contribution from $B^+ \to K^+\pi^+\pi^-$ events. To facillitate the comparison, combinatorial events are also taken into account. Their level is estimated based on a fit to the upper $m(K^+e^+_{[\to\pi^+]}e^-_{[\to\pi^-]})$ sideband, using an exponential model.

The left-hand plot in Figure 8.8 shows the distribution in data, alongside the simulated model with and without the estimated $B^+ \to K^+\pi^+\pi^-$ contamination. Both models are compatible

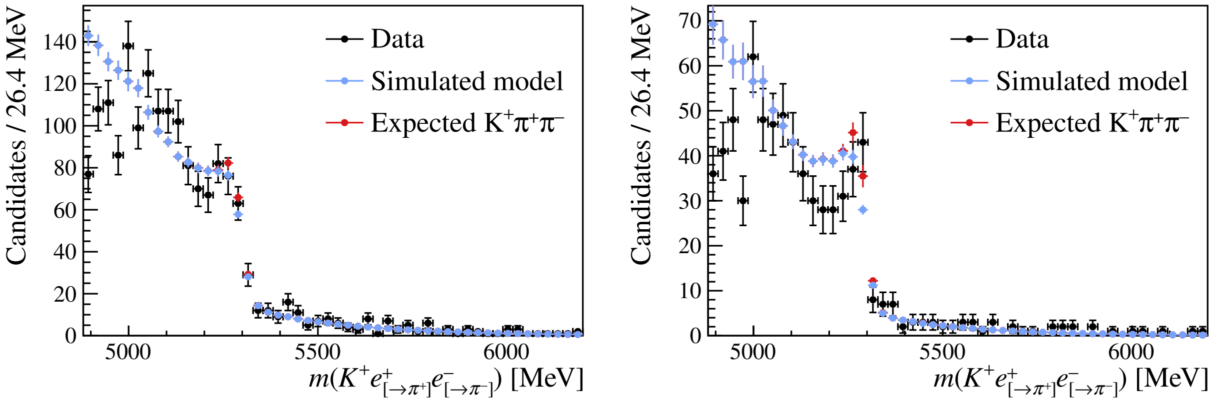

Figure 8.8: Distributions of the $m(K^+e^+_{[\to\pi^+]}e^-_{[\to\pi^-]})$ invariant mass. The black corresponds to data, whereas the blue contains the expected contributions from $B^+ \to K^+e^+e^-$, $B^0 \to K^{*0}e^+e^-$, and combinatorial events, as predicted based on simulation and data. The estimated contribution from $B^+ \to K^+\pi^+\pi^-$ events is shown in red, stacked on top of the blue. The exercise is performed after the nominal selection (left), and then repeated after additioally requiring both electrons to have `ProbNNpi` > 0.2 (right).

with the data in the region of interest, *i.e.* around the mass of the B^+. The exercise is repeated after applying a cut at `ProbNNpi` > 0.2 on the electrons, in order to enhance any potential $B^+ \to K^+\pi^+\pi^-$ contamination. The resulting distribution is presented on the right-hand side of Figure 8.8. No significant excess is observed.

Note that both tests presented in this section result in upper limits on the size of the potential impact of misidentified $B^+ \to K^+\pi^+\pi^-$ events. This, combined with the fact that the second test is only possible after unblinding, means that the double misidentification studies should not be used to assign a systematic uncertainty.

In summary, several tests are performed to verify that $B^+ \to K^+\pi^+\pi^-$ events where both pions are misidentified as leptons are suppressed as expected by the PID requirements, the q^2 window, and the mass vetoes. As shown in this section, the test results are compatible with the expectation that the levels of $B^+ \to K^+\pi^+\pi^-$ events are sufficiently low to not warrant explicit modelling in the fit to $B^+ \to K^+\ell^+\ell^-$ data, which is the subject of the next chapter.

Chapter 9

The fit to the rare modes

Once the efficiencies and their uncertainties are well understood, the next step towards measuring R_K is the fit to $B^+ \to K^+\ell^+\ell^-$ data. As outlined in Chapter 5, the fit is based on the extended maximum likelihood method [184] implemented in the RooFit package [185]. The expression for the likelihood is presented in Section 9.1, and then the distributions that model the signal and background in each $B^+ \to K^+\ell^+\ell^-$ data sample are described in Section 9.2. Next, Section 9.3 describes the constraints that enter the likelihood, in order to improve the result on R_K. A summary is provided in Section 9.4.

The $B^+ \to K^+\ell^+\ell^-$ fit procedure is designed and implemented by Dr. Paula Álvarez Cartelle. The results that constitute original work are: the constraints presented in Tables 9.1 to 9.4, and the shape templates that enter the likelihood as described in Section 9.1. This includes the high-$m(K\pi)$ weights illustrated in Figures 9.2 and 9.3, which represent an improvement with respect to the previous R_K analysis.

9.1 Description of the likelihood surface

The selected $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ data are fitted simultaneously, by maximising a likelihood built upon models used to describe the data. This is equivalent to minimising the negative of the logarithm of the likelihood, referred to as the NLL and used throughout this thesis for convenience. This section shows how the NLL is constructed step by step. For the muon and electron samples, respectively, the NLLs are:

$$-\log\mathcal{L}^r_\mu = -\sum_i \log\mathcal{P}^r_\mu(m^i_r|N^r_{K\mu\mu})\,,\text{ and} \tag{9.1}$$

$$-\log\mathcal{L}^{rt}_e = -\sum_i \log\mathcal{P}^{rt}_e(m^i_{rt}|N^{rt}_{Kee})\,. \tag{9.2}$$

In the above expressions, the index i runs over all selected events, r runs over the four data-taking periods, and t runs over the three electron trigger selections. The independent variable is the $K^+\ell^+\ell^-$ invariant mass, m, whose distribution in each mode, period, and trigger category is modelled independently by $\mathcal{P}^{r,t}_{e,\mu}$. The yields $N^{rt}_{K\mu\mu}$ and N^{rt}_{Kee} depend on R_K by virtue of Equation (4.4). At this stage, these yields are the only unknown quantities in the expression for R_K, meaning that it can be written as:

$$R^{rt}_K = \frac{N^r_{K\mu\mu}}{N^{rt}_{Kee}}\cdot\frac{N^{rt}_{J/\psi ee}}{N^r_{J/\psi\mu\mu}}\cdot\frac{\varepsilon^{rt}_{Kee}}{\varepsilon^r_{K\mu\mu}}\cdot\frac{\varepsilon^r_{J/\psi\mu\mu}}{\varepsilon^{rt}_{J/\psi ee}} = \frac{N^r_{K\mu\mu}}{N^{rt}_{Kee}}\cdot c^{rt}_K\,. \tag{9.3}$$

In the above expression, the efficiencies, together with the control-mode yields, are grouped into the c^{rt}_K factors. They allow the likelihood to be reparametrised such that R_K is obtained as one of the fit's parameters of interest. The sum of muon and electron NLLs is then:

$$-\log\mathcal{L} = -\sum_r\sum_i \log\mathcal{P}^r_\mu(m^{rt}_i|N^r_{K\mu\mu}) - \sum_{rt}\sum_i \log\mathcal{P}^{rt}_e\left(m^{rt}_i\middle|N^r_{K\mu\mu}\cdot c^{rt}_K/R_K\right)\,. \tag{9.4}$$

The models for $\mathcal{P}^{rt}_e$ and $\mathcal{P}^{rt}_\mu$ are described in Section 9.2. To improve the fit procedure, constraints on various parameters of the fit are added to the NLL. These constraints take the form of multi-dimensional Gaussian distributions $\left\{\mathcal{G}_j(\vec{x}_j|\vec{\mu}_j,\Sigma_j)\right\}$, where the index j runs over the different constraints. The means and widths of the constraints are denoted by $\vec{\mu}_j$ and Σ_j, respectively. The NLL to be minimised by the fit can then be written as:

$$-\log \mathcal{L} = -\sum_{r}\sum_{i} \log \mathcal{P}^{r}_{\mu}(m_i^{rt}|N^{r}_{K\mu\mu}) - \sum_{rt}\sum_{i} \log \mathcal{P}^{rt}_{e}\left(m_i^{rt}\middle|N^{r}_{K\mu\mu}\cdot c^{rt}_{K}/R_K\right) - \sum_{j} \log \mathcal{G}_j(\vec{x}_j|\vec{\mu}_j,\Sigma_j)\,. \tag{9.5}$$

Based on the type of parameter they target, the $\mathcal{G}_j$ in the formula above can be classified into constraints on shapes, and constraints on yields. The former are described in Section 9.2, together with the $\mathcal{P}^{r,t}_{e,\mu}$ used in the model. The latter are covered by Section 9.3.

9.2 Models for the fit components

The model used to describe $B^+ \to K^+\ell^+\ell^-$ data has several parallels with the one presented in Chapter 5, which is used for $B^+ \to K^+J/\psi(\ell^+\ell^-)$ data. The signal is expected to have a peaking structure, with tails that are non-Gaussian due to effects such as bremsstrahlung energy loss. Combinatorial events, consisting of random combinations of tracks in the candidate, follow a distribution that can be accurately modelled by an exponential function. The excellent resolution of the muon channels allows the lower limit of the invariant mass to be chosen such that the signal and the combinatorial are the only components in the muon data. In the $B^+ \to K^+e^+e^-$ samples, the poorer mass resolution caused by radiative losses leads to contributions from two additional backgrounds. The first one consists of partially-reconstructed events, such as $B^+ \to K^{*0}(K^+\pi^-)e^+e^-$ decays, where at least one particle (in this case the π^-) escapes detection. The second one represents $B^+ \to K^+J/\psi(e^+e^-)$ decays whose reconstructed q^2 falls into the $q^2 \in (1.1\ \text{GeV}^2, 6.0\ \text{GeV}^2)$ interval, due to failures in the bremsstrahlung recovery process. In summary, the model for $B^+ \to K^+\mu^+\mu^-$ data has two components: the peaking signal and the exponential combinatorial background. The $B^+ \to K^+e^+e^-$ data also has signal and combinatorial components, but there are two additional background sources: partially-reconstructed events, and leakage from the J/ψ.

In $B^+ \to K^+\mu^+\mu^-$ data, the signal is modelled by the sum of two Crystal Ball (CB) distributions [188]. They share the same mean, and their power-law tails are on opposite sides of the peak, in order to account for radiative effects. Like in the control mode, the tail parameters are obtained from fits to simulated samples, however the mean and the width of the peak are expected to be different in data. For this reason, they are reparametrised into the shift $\Delta\mu$ and the scale s_σ introduced in Equation (5.2). Their values are obtained from fits to the unconstrained $m(K^+\mu^+\mu^-)$ invariant-mass distribution in $B^+ \to K^+J/\psi(\mu^+\mu^-)$ data. This is motivated by the larger statistics of the control modes, and by the fact that $\Delta\mu$ and s_σ are not expected to vary across q^2. These fits are performed by Dr. Paula Álvarez Cartelle, and are outside the scope of this thesis.

As pointed out in Section 5.4, the resolution of electrons depends on whether they contain added energy from the bremsstrahlung recovery process. This motivates the splitting of electron data into photon categories, which are modelled independently as shown in Equation (5.5). In the case of no bremsstrahlung recovery, *i.e.* in 0γ data, the lower tail is comparatively larger, due to the missing energy carried away by bremsstrahlung photons. This is showcased in Figure 9.1, where simulated signal shapes in the three photon categories are plotted together. The success of the bremsstrahlung recovery algorithm leads to reduced low-mass tails and more accurate peak locations. However, the high-mass tails become more populated, due to candidates where the bremsstrahlung recovery algorithm has overestimated the energy loss of the signal electrons. Like in the $B^+ \to K^+\mu^+\mu^-$ data, the 0γ signal is modelled by the sum of two CBs that share the same mean, and whose power-law tails lie on opposite sides of the peak. The same process is repeated on the 1γ and 2γ samples, noting that the addition of energy from bremsstrahlung radiation broadens the peak and reduces the size of the low-mass tail. To better capture this behaviour, a Gaussian distribution is added to the pair of CBs in each photon category. The Gaussian shares the same mean shift $\Delta\mu$ and width scale s_σ as the CBs.

The shape of the J/ψ leakage in the electron window is taken from simulated $B^+ \to K^+J/\psi(e^+e^-)$ events that pass the signal selection. It is modelled using an adaptive kernel density estimation method [190]. The same technique is used to describe the partially-reconstructed component, where simulated $B^0 \to K^{*0}e^+e^-$ events are used. These samples only contain the P-wave contribution, which corresponds to the peak at the K^{*0} resonance in the distribution of the invariant mass of the kaon and pion, $m(K\pi)$. This peak is shown on the

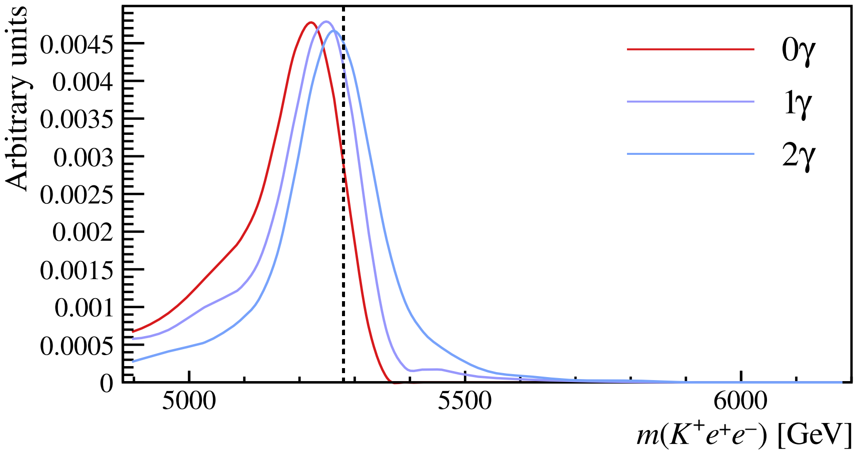

Figure 9.1: Shape templates for the signal in the 2018 eTOS sample, shown separately for the three bremsstrahlung categories. The vertical line shows the $m(K^+e^+e^-)$ values corresponding to m_B. The range of the horizontal axis coincides with the fit window.

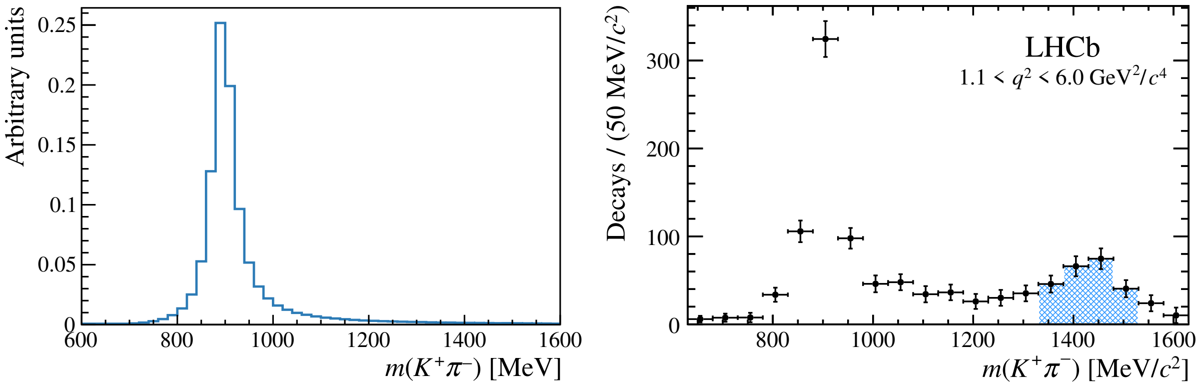

Figure 9.2: Distribution of $m(K\pi)$ in simulated $B^0 \to K^{*0}e^+e^-$ events (left), compared to the measured spectrum [39] (right).

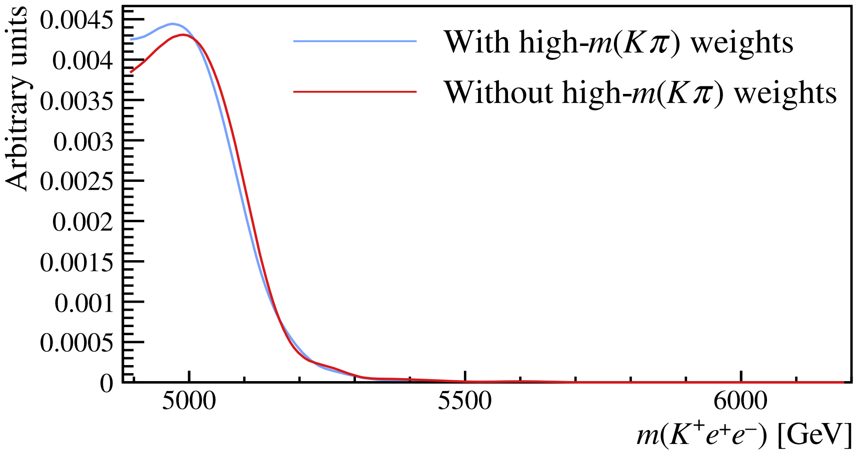

Figure 9.3: Shape templates for partially-reconstructed events in the 2018 eTOS sample, with and without the weights that account for high-$m(K\pi)$ states. The range of the horizontal axis coincides with the fit window.

left-hand side of Figure 9.2. However, events at higher $m(K\pi)$ are expected to contribute to the shape in data, although with kinematic suppression induced by the q^2 selection. To account for this, weights are assigned to the $B^0 \to K^{*0}e^+e^-$ simulation such that the $m(K\pi)$ spectrum matches the one measured in data [39]. This is shown on the right-hand side plot of Figure 9.2. Figure 9.3 shows the effect of these weights on the expected invariant-mass distribution of partially-reconstructed events. It can be seen that the high-$m(K\pi)$ weights lead to larger populations at very low $m(K^+e^+e^-)$. This is expected, since the missing pion makes it likely for events with large $m(K\pi)$ to have low $m(K^+e^+e^-)$. In addition, the template that includes the weights drops off at lower $m(K^+e^+e^-)$, indicating that not accounting for high-$m(K\pi)$ states could have lead to overestimation of the overlap between the models for the signal and the partially-reconstructed background.

9.3 Constraints on component yields

The accuracy of R_K is improved by increasing the amount of information available to the fit. For example, the c_K^{rt} terms introduced in Equation (9.3) are predicted using the measured yields of the control data, along with the estimated efficiencies. The c_K^{rt} factors are anti-correlated with R_K, so constraining the fit to prefer values close to their expectations encourages it to change R_K, until it finds the value that minimises the NLL. This is how adding constraints as opposed to freely floating parameters improves the accuracy of the result.

In total, there are four multidimensional Gaussian constraints $\mathcal{G}_j$ that enter Equation (9.5) to constrain the relative yields of the fit components. The first constraint is on the c_K^{rt} parameters. The central values, $\vec{\mu}$, are the estimates obtained from the nominal efficiencies and control-mode yields; they are listed in Table 9.1. The width of the constraint, Σ, is dictated by the total covariance matrix V obtained in Chapter 7.

The second constraint is on the normalisation of the partially-reconstructed background relative to the signal. This component is dominated by $B^0 \to K^{*0}e^+e^-$ events, which are similar to $B^+ \to K^+e^+e^-$ decays. For this reason, the partially-reconstructed yields with respect to the signal can be estimated from simulation, through the following ratios:

$$r^{rt}_{\text{prc}} = \frac{N^{rt}_{\text{prc}}/N^{r0}_{\text{prc}}}{N^{rt}_{\text{sig}}/N^{r0}_{\text{sig}}} = \frac{\varepsilon^{rt}_{\text{trig,mass,BDT}}(B^0 \to K^{*0}e^+e^-)/\varepsilon^{r0}_{\text{trig,mass,BDT}}(B^0 \to K^{*0}e^+e^-)}{\varepsilon^{rt}_{\text{trig,mass,BDT}}(B^+ \to K^+e^+e^-)/\varepsilon^{r0}_{\text{trig,mass,BDT}}(B^+ \to K^+e^+e^-)}\,. \tag{9.6}$$

These terms express the ratio between partially-reconstructed and signal yields in each data sample, N^{rt}_{prc} and N^{rt}_{sig} respectively, as a function of the eTOS yields from each data-taking period (N^{r0}_{prc} and N^{r0}_{sig}). Thanks to the similar kinematics of $B^+ \to K^+e^+e^-$ and $B^0 \to K^{*0}e^+e^-$ decays, most efficiencies are expected to cancel in the ratio. The exception is the efficiency of the trigger, mass, and multivariate selection, denoted by $\varepsilon_{\text{trig,mass,BDT}}$ in the above expression.

In the fit to $B^+ \to K^+\ell^+\ell^-$ data, the r^{rt}_{prc} factors are constrained via a multi-dimensional Gaussian to the efficiency ratios estimated from simulation. The central values of these constraints, which represent the mean of the Gaussian, are listed in Table 9.2. They are obtained through Equation (9.6), using efficiency estimates derived from fully-calibrated $B^+ \to K^+e^+e^-$ and $B^0 \to K^{*0}e^+e^-$ samples. To determine the width of the constraint, estimates of r^{rt}_{prc} are recalculated based on $B^+ \to K^{*+}e^+e^-$ simulation instead of $B^0 \to K^{*0}e^+e^-$. The contributions from the two decay modes are expected to have similar shapes, however their normalisation relative to the signal could be different due to systematic effects induced by the neutral pion coming from the K^{*+} decay (these would not cancel in the efficiency ratio). In addition, the $B^+ \to K^{*+}e^+e^-$ simulation samples have lower statistics than their $B^0 \to K^{*0}e^+e^-$, leading to relative uncertainties of up to 20% on the efficiency estimates. For this reason, the uncertainties on the efficiencies calculated using $B^+ \to K^{*+}e^+e^-$ samples are explicitly taken into account when calculating the diagonal elements of the r^{rt}_{prc} covariance matrix. This represents a change with respect to the previous R_K measurement, where the width of the r^{rt}_{prc} constraint did not take into account the statistics of the $B^+ \to K^{*+}e^+e^-$ simulation. In addition, the only sample available at the time corresponded to Run 1 data-taking conditions, and so the constraint on Run 2.1 values was taken to be fully correlated to its Run 1 counterpart.

The i^{th} element on the diagonal of the r^{rt}_{prc} covariance matrix is given by:

$$V^{r^{rt}_{\text{prc}}}_{i,i} = \left(\xi(B^+ \to K^{*+}e^+e^-) - \xi(B^0 \to K^{*0}e^+e^-)\right)^2 + (\delta\xi)^2\,, \tag{9.7}$$

where $\xi(X) \equiv \varepsilon^{rt}_{\text{trig,mass,BDT}}(X)/\varepsilon^{r0}_{\text{trig,mass,BDT}}(X)$, and $\delta\xi$ is the uncertainty on $\xi(B^+ \to K^{*+}e^+e^-)$. Equations (7.3) and (7.4) are then used to convert the covariance matrix into the fractional error matrix, which is presented in Table 9.3.

Table 9.1: Central values of the constraint on the c^{rt}_K terms that enter the fit for R_K.

	c^{rt}_K		
	eTOS	hTOS!	TIS!
Run 1	0.1436	0.0509	0.0548
Run 2.1	0.2809	0.0710	0.0806
2017	0.2403	0.0709	0.0835
2018	0.2296	0.7652	0.0788

Table 9.2: Central values of the constraint on the r^{rt}_{prc} terms that enter the fit for R_K.

	r^{rt}_{prc}	
	hTOS!	TIS!
Run 1	0.936916	1.06515
Run 2.1	0.840481	0.907575
2017	0.852263	0.938888
2018	0.884121	0.995914

Table 9.3: Fractional error matrix that determines how tightly the r^{rt}_{prc} terms are constrained to their central values in the fit for R_K.

r^{rt}_{prc}							
Run 1		Run 2.1		2017		2018	
$\frac{h\text{TOS!}}{e\text{TOS}}$	$\frac{\text{TIS!}}{e\text{TOS}}$	$\frac{h\text{TOS!}}{e\text{TOS}}$	$\frac{\text{TIS!}}{e\text{TOS}}$	$\frac{h\text{TOS!}}{e\text{TOS}}$	$\frac{\text{TIS!}}{e\text{TOS}}$	$\frac{h\text{TOS!}}{e\text{TOS}}$	$\frac{\text{TIS!}}{e\text{TOS}}$
27.48	−0.01	0.74	−34.58	−13.95	−0.91	26.93	−32.27
	24.16	−0.00	0.02	0.01	0.00	−0.01	0.02
		−0.20	9.55	3.85	0.25	−7.44	8.91
		16.75	−1.50	−0.60	−0.04	1.17	−1.40
			22.21	28.28	1.84	−54.61	65.44
				−2.10	−0.14	4.05	−4.85
				18.97	0.74	−22.03	26.40
					15.88	−1.44	1.72
						−47.42	56.83
						34.40	−50.97
							17.22

The third constraint is on the normalisation of the background component corresponding to leakage from the J/ψ, denoted by N_{leak}. The efficiencies of the rare and resonant q^2 selections are determined from simulated $B^+ \to K^+ J/\psi(e^+e^-)$ events. The ratio of the two efficiencies, f_{leak}, is expected to be equal to the ratio between the leakage yield in the $B^+ \to K^+e^+e^-$ fit, and the yield of the control mode as obtained in Section 5.6. Therefore, the normalisation of the J/ψ leakage background, in each data-taking period and trigger selection, is constrained to:

$$N^{rt}_{\text{leak}} = f_{\text{leak}} \cdot N^{rt}_{J/\psi ee} \,. \tag{9.8}$$

The central values of these constraints are listed in Table 9.4, alongside their uncertainties. They are driven by the statistics of the simulated $B^+ \to K^+ J/\psi(e^+e^-)$ samples in the q^2 region corresponding to the rare mode, and are used as the widths of the constraints. The individual constraints are treated as uncorrelated, following an investigation conducted by Dr. Thibaud Humair during the previous R_K analysis.

Table 9.4: Central values of the constraint on the J/ψ leakage normalisation. The width of the constraint is dictated by the uncertainties, which are dominated by the statistics of the simulated $B^+ \to K^+ J/\psi(e^+e^-)$ sample. All data-taking periods and trigger selections are treated as uncorrelated.

		$\boldsymbol{N^{rt}_{\text{leak}}}$	
	eTOS	hTOS!	TIS!
Run 1	4.5 ± 2.1	4.0 ± 1.9	2.8 ± 2.0
Run 2.1	12.1 ± 3.1	8.0 ± 2.7	6.4 ± 2.4
2017	12.1 ± 2.2	2.4 ± 1.0	3.8 ± 1.1
2018	12.0 ± 2.4	3.3 ± 1.4	5.7 ± 1.8

The fourth and final constraint is on the fractions of signal candidates in each photon category, denoted by $f_{0\gamma}$, $f_{1\gamma}$, and $f_{2\gamma}$. It is checked with control-mode simulation and data that these fractions are simulated accurately. As a result, the central values of the constraint, which are presented in Table 9.4, come from $B^+ \to K^+e^+e^-$ simulation. Each dimension of the constraint is treated as uncorrelated with the others, and has an assigned width of 1% of its central value. This is a conservative uncertainty that covers the statistical precision offered by the simulation samples.

Table 9.5: The central value of the constraints on the fractions of events falling in each photon category. The uncertainties shown are statistical. Each individual constraint is assumed to be uncorrelated with the rest.

		$f_{0\gamma}$	
	eTOS	hTOS!	TIS!
Run 1	0.2383 ± 0.0015	0.1558 ± 0.0030	0.2052 ± 0.0024
Run 2.1	0.2509 ± 0.0005	0.1356 ± 0.0011	0.1724 ± 0.0008
2017	0.2498 ± 0.0009	0.1434 ± 0.0019	0.1692 ± 0.0014
2018	0.2520 ± 0.0010	0.1460 ± 0.0020	0.1692 ± 0.0015
		$f_{1\gamma}$	
	eTOS	hTOS!	TIS!
Run 1	0.4962 ± 0.0018	0.4982 ± 0.0043	0.4957 ± 0.0030
Run 2.1	0.5008 ± 0.0005	0.5018 ± 0.0016	0.5000 ± 0.0010
2017	0.5020 ± 0.0010	0.4990 ± 0.0027	0.5013 ± 0.0019
2018	0.5028 ± 0.0011	0.5007 ± 0.0029	0.4957 ± 0.0020
		$f_{2\gamma}$	
	eTOS	hTOS!	TIS!
Run 1	0.2655 ± 0.0016	0.3460 ± 0.0042	0.2991 ± 0.0027
Run 2.1	0.2483 ± 0.0005	0.3626 ± 0.0016	0.3276 ± 0.0010
2017	0.2482 ± 0.0009	0.3576 ± 0.0026	0.3295 ± 0.0018
2018	0.2452 ± 0.0010	0.3532 ± 0.0028	0.3350 ± 0.0019

9.4 Outcome of the fit

The result of the fit is the value of R_K found to minimise the NLL, R_K^{fit}. The uncertainty on this estimate, $\sigma(R_K)$, is determined by minimising the NLL for a range of R_K values around the minimum; this is referred to as profiling. Based on Wilks' theorem [197], the 68% confidence interval is assigned to be the range of R_K values for which $-\log\left(\mathcal{L}(R_K)/\mathcal{L}(R_K^{\text{fit}})\right) < 0.5$. The resulting confidence interval covers the statistical uncertainty of the $B^+ \to K^+\ell^+\ell^-$ data, as well as the systematic uncertainty induced by the c_K^{rt} constraints. To finalise the R_K measurement, two further effects are incorporated into the profiled NLL. These effects are: the bias induced on R_K^{fit} by the fit procedure, and the systematic uncertainty induced by the fit model. They are covered by the following chapter.

Chapter 10

Validation of the fit to signal data

The previous chapter describes the fit employed to describe the signal data, and thus infer R_K. This chapter presents the tests conducted to validate the fit procedure. First, in Section 10.1, the bias on R_K as extracted from the fit is evaluated. This leads to a correction of the central value of the R_K measurement. Second, Section 10.2 evaluates the systematic effect due to the chosen model for the data. Third, Section 10.3 presents a host of checks designed to verify that the fit is able to model variations between subsamples of the data.

The tests presented here are performed in collaboration with Dr. Paula Álvarez Cartelle, who designed the procedure during the previous R_K measurement. In this chapter, results that constitute original work are found in Section 10.2.2 and Section 10.3.

10.1 Bias on the R_K estimate from the fit

It is possible that the fit is biased, meaning that it prefers values of R_K that are not identical to the true value. To assess the size of this bias, 10 000 simulated samples are generated based on the expected fit parameters (including R_K) and shapes. These pseudoexperiments, also referred to as toys, use samples that are statistically representative of the expected data to probe the behaviour of the fit. The toys are generated by fluctuating the expected yields according to their Poisson errors. In addition, the parameters that are constrained in the fit are resampled according to the widths of their constraints. The shape templates are kept constant, with the exception of the total $B^+ \to K^+e^+e^-$ signal shape that changes according to the sampled constraints on $f_{0\gamma}$, $f_{1\gamma}$, and $f_{2\gamma}$. The fit is then performed on each generated pseudoexperiment, and its behaviour is inferred by comparing the outcome to the model used to generate the toys. The bias and coverage are then obtained from the distribution across toys of the pull between the value of R_K obtained from the fits, R_K^{fit}, and the value used to generate the toys, R_K^{gen}. The pull is defined as:

$$\text{pull} = \frac{R_K^{\text{fit}} - R_K^{\text{gen}}}{\sigma(R_K)}, \tag{10.1}$$

where $\sigma(R_K)$ is the uncertainty on the R_K estimate coming from the fit. If the fit is unbiased, the average value across the toys of R_K^{fit} is equal to R_K^{gen} by definition. This means that, in the limit of infinitely-many generated toys, the distribution of the pull has a mean of 0. In addition, the standard deviation of the pull distribution is equal to 1 if the statistical fluctuations of R_K^{fit} with respect to R_K^{gen} are successfully captured by the precision $\sigma(R_K)$. In other words, the pulls of R_K from fits to toy datasets are expected to be distributed according to a standard normal distribution. The mean of the pull distribution represents the bias on R_K^{fit}, relative to the uncertainty $\sigma(R_K)$.

The validation procedure is performed twice. The first iteration is conducted before unblinding, to ensure the fit bias on R_K is kept within a few percent relative to the uncertainty coming from the fit. The yields and shapes corresponding to the fit components in the previous data are generated based on the fit result from the previous R_K measurement. Their Run 2.2

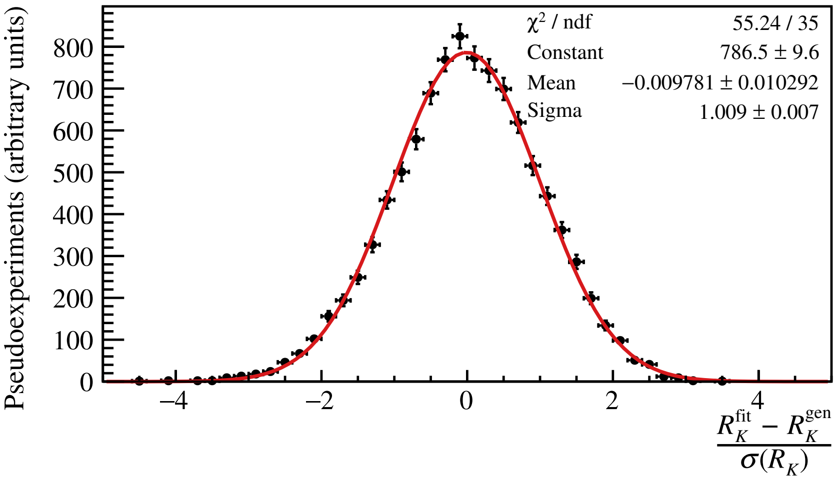

Figure 10.1: Distribution across pseudoexperiments of the pull on R_K (black), alongside the best-fit Gaussian distribution (red). The minimum χ^2 from the fit for the Gaussian is presented on the top right, alongside the best-fit parameters.

equivalents are blinded at this stage of the procedure, so they are estimated by scaling the Run 2.1 yields from the previous measurement by the expected gain in luminosity. This iteration found the fit to be biased by around 5.7% of $\sigma(R_K)$. In addition, the width of the pull distribution is found to be compatible with 1, indicating good coverage.

The second iteration of the validation procedure takes place after unblinding, thus allowing the toys to be generated using the central values extracted directly from the fit. The distribution of the resulting pulls is shown in Figure 10.1. The estimated bias is obtained from the Gaussian distribution that is found — through χ^2 minimisation — to best model the pull distribution. The mean of this Gaussian is $(-0.98 \pm 1.03)\% \cdot \sigma(R_K)$, showing that the fit bias is small. In addition, the width of the pull distribution is compatible with 1, thus demonstrating good coverage of the precision on R_K. As agreed with the internal review committee, the central value of R_K is corrected according to the estimated bias, and the sum in quadrature of half of the size of the bias and the bias′ uncertainty is treated as an additional systematic uncertainty on R_K.

10.2 Systematic effects related to the fit procedure

There are models equally valid to those described in Section 9.2 that can be used to describe the signal and background contributions to the invariant-mass distributions in data. The choice of which model to use can induce a systematic uncertainty on the extracted value of R_K. Such effects are negligible, except when they are due to:

1. the modelling of the $B^+ \to K^+e^+e^-$ signal shape;
2. partially-reconstructed $K\pi$ states with high $m(K\pi)$ invariant masses; and
3. partially-reconstructed events originating from higher-order K^* resonances.

The following subsections go through each of the above three effects in turn. The results are summarised in Table 10.1. The total systematic uncertainty is the sum in quadrature over the three effects. It is found to be 1%, making it the dominant systematic effect on R_K.

Table 10.1: Systematic uncertainties on R_K, in percent, arising from variations of the fit model.

Source	Uncertainty (%)
Signal model	0.70
High-$m(K\pi)$ states	0.67
Additional resonances	0.23
Total	1.00

10.2.1 Validation of the signal model

It is checked whether the $B^+ \to K^+e^+e^-$ data in the three photon categories can be accurately described by the chosen models. As detailed in Section 9.2, the sum of two CBs is used to model the signal in the 0γ data, whereas the 1γ and 2γ signals are described by the combination of two CBs and a Gaussian distribution. Pseudoexperiments are generated based on these parametrisations and the expected yields. Then, two fits are performed on each toy. The first fit uses the nominal signal model described so far. The second fit uses alternative models for the $B^+ \to K^+e^+e^-$ signal. These consist of: a single CB in the 0γ data, and the sum of three Gaussian distributions in each of the other photon categories. The two fits lead to two estimates, R_K^{nom} and R_K^{alt}. The mean across pseudoexperiments of $\left(R_K^{\text{alt}} - R_K^{\text{nom}}\right)\big/ R_K^{\text{gen}}$ represents the assigned systematic uncertainty.

10.2.2 Impact of high-$m(K\pi)$ states

The $m(K\pi)$ spectrum in simulated $B^0 \to K^{*0}e^+e^-$ events is calibrated, using weights, to match the one measured in data [39]. The $m(K\pi)$ spectra are shown in Figure 9.2, and the effect of the calibration procedure is exemplified by Figure 9.3. The statistical uncertainty on the weights induces a systematic effect on R_K. This is evaluated using toys with bootstrapped partially-reconstructed templates. The fractional shift in R_K is taken to be the systematic uncertainty.

Because this calibration was not conducted for the previous R_K measurement, it is checked whether the result from Ref. [2] changes as expected after the introduction of the $m(K\pi)$ correction weights. Pseudoexperiments are generated based on the known Run 1 and Run 2.1 yields and shapes, taking into account high-$m(K\pi)$ states. Two fits are performed on each pseudoexperiment, and on the Run 1 and Run 2.1 data. One of the fits is in the configuration used by the previous R_K measurement, and the other is adjusted to take into account the effect of the $m(K\pi)$ calibration. The difference between the values of R_K obtained from the two fit configurations, R_K^{prev} and R_K^{new}, respectively, is used as the test statistic. The distribution of $R_K^{\text{new}} - R_K^{\text{prev}}$ in pseudoexperiments is shown in Figure 10.2, alongside vertical lines depicting several confidence intervals. It is agreed with the internal review committee that if the value of $R_K^{\text{new}} - R_K^{\text{prev}}$ obtained by running the check over data falls in the $2.5\,\sigma$ confidence interval of the distribution from toys, the test is considered a success. The results of the two fits to data are:

$$R_K^{\text{new}} = 0.843\,^{+\,0.062}_{-\,0.055}\,, \text{ and} \tag{10.2}$$

$$R_K^{\text{prev}} = 0.848\,^{+\,0.061}_{-\,0.055}\,. \tag{10.3}$$

This means that the calibration of the $m(K\pi)$ distribution in partially-reconstructed background events induces a shift of -0.005 in the previous R_K result. This is compatible with the expectation from pseudoexperiments. Note that the two fit configurations are assumed to have the same bias, leading to its cancellation in the difference. For this reason, the bias of the fit is not evaluated for the results above. This is why the result in Equation (10.3) is not identical to the previous R_K result.

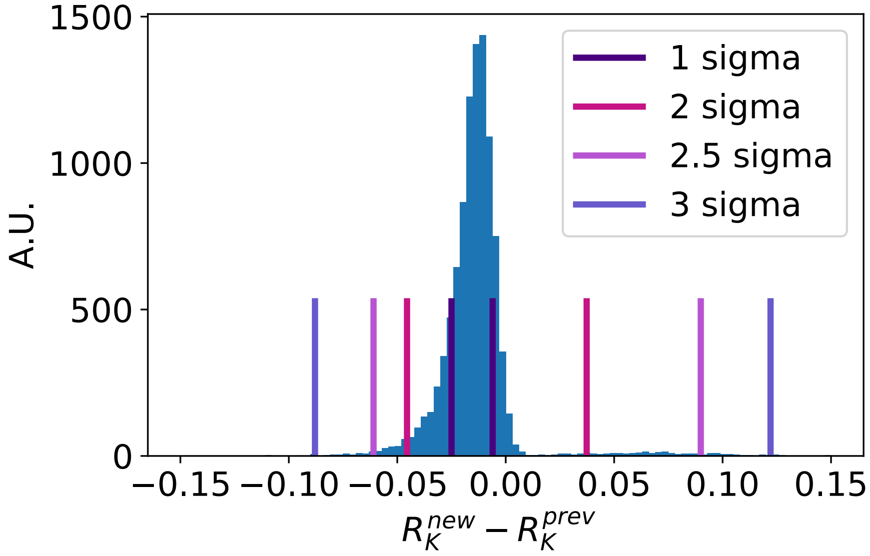

Figure 10.2: Expected shift in the value of R_K, based on Run 1 and Run 2.1 data, induced by the calibration of the distribution of $m(K\pi)$ in partially-reconstructed background events. The distribution of the shift is shown in blue, whilst several confidence intervals are delimited by vertical lines in shades of purple.

10.2.3 Effect induced by additional resonances

To further test the robustness of the model used to describe partially-reconstructed contributions to the $m(K^+e^+e^-)$ data, the mass-shape templates are recomputed to include contributions from additional K^* resonances that decay into a K^+ and two pions, the latter of which escape detection. An example is the process $B^+ \to K_1^+(K^+\pi^+\pi^-)e^+e^-$. The contributions are normalised assuming the hadronic system in the rare channel is the same in the J/ψ channel. For the example above, this means that $N(K_1^+e^+e^-) = N(K^+e^+e^-) \cdot N(K_1^+J/\psi)/N(K^+J/\psi)$, where $N(X)$ is the yield of mode X. The new templates are used as an alternative model for the partially-reconstructed background component of the data. Pseudoexperiments are generated based on this model, and then two fits are performed on each toy. One of the fits is in the nominal configuration, whereas the other uses the alternative partially-reconstructed background templates. Like with the effect of the signal modelling, the mean across pseudoexperiments of the R_K estimates from the two fits represents the assigned systematic uncertainty.

10.3 Compatibility checks

The cross-checks presented in Chapter 8 reinforce the robustness of the efficiencies in a number of different ways. The results on $r_{J/\psi}$ show that the electron and muon channels are in agreement, and that the efficiencies are parametrised correctly. The $R_{\psi(2S)}$ check demonstrates the portability of efficiency corrections across q^2. The particularities of some corrections are tested in further detail through the scaling and tracking cross-checks. The tests presented in this section go a step further, by targeting the fit procedure in addition to the efficiencies. This is done by verifying that variations between data samples, induced by factors such as trigger strategies and data-taking conditions, are suitably taken into account by the fit.

For example, it is shown in Chapter 6 that the *e*TOS electron trigger selection is more efficient than the TIS! selection. However, the measured value of R_K must not depend on whether one uses *e*TOS or TIS! data. The same argument applies to individual data-taking periods: R_K is not expected to vary with time, and the efficiencies are required to capture changes in data-taking conditions. This is what motivates the investigation of the compatibility between subsets of the data used to measure R_K.

This study comprises several checks of whether distinct subsets of the data prefer significantly different values of R_K. Given the invariant nature of R_K, each subset should lead to approximately the same value of R_K, regardless of the criteria used to split the total sample into subsets. If that is the case, the fit does not gain much from the liberty of choosing separate estimates of R_K for each subset. For this reason, the compatibility checks are performed by changing the nominal configuration of the fit to accommodate distinct values of R_K in different partitions of the data. If the addition of these R_K values does not substantially improve the fit, then the subsets under scrutiny are compatible. By contrast, if the fit benefits from the freedom of multiple R_K values, then the subsets under question may not be compatible with each other.

The checks are conducted before unblinding R_K, therefore a test statistic is chosen to quantify the compatibility between the fits. The outcome of each check is obtained from the *p*-value of the test statistic from data. The following subsection describes how this p-value is determined, alongside the general aspects of the procedure. Then, the main results are presented in Section 10.3.2, followed by complementary checks given in Section 10.3.3.

10.3.1 General procedure

The final objective of each compatibility check is to determine if distinct partitions of the data prefer different values of R_K, without inspecting the resulting values. To this end, a test statistic is chosen for each check, and its distribution is obtained by running the check on toy datasets. Once this is done, the check is run on data to obtain the corresponding value of the test statistic. This is used to calculate a p-value, based on the distribution obtained from the fits to toy data. If this p-value corresponds to a significance below $2.5\,\sigma$, the data partitions are considered compatible, thus passing the check successfully. This $2.5\,\sigma$ threshold, together with the partitioning scheme for the checks, was chosen *a priori* through discussion with the internal review committee.

An inherent risk of the assessment based on the test statistic is the possibility that a small p-value from data arises from statistical fluctuations, rather than from a flaw of the fit. Given the $2.5\,\sigma$ threshold, this is unlikely to happen when considering one individual check. However, the number of data partitions for which compatibility is assessed increases, the more likely false negatives become. To keep the probability of this scenario below 10%, a partitioning scheme consisting of three checks is agreed upon with the internal review committee. The data is required to pass these three checks before unblinding, whilst tests involving other partitions are only considered in case of unusual results, such as a potential false negative. The three checks are:

1. “polarity”: compatibility between data taken with the two magnet configurations;
2. “window”: compatibility between the nominal fit, and a fit that excludes electron candidates with $m(K^+e^+e^-) < 5$ GeV; and
3. “selection”: compatibility between data-taking periods and trigger categories.

The polarity test is motivated by the fact that the configuration of the LHCb magnet is not expected to have a noticeable impact on R_K. The window check studies the effect of reducing the contamination from partially-reconstructed events, at the cost of signal precision. Finally, the selection test probes whether the electron trigger strategies, alongside any effects that are expected to change between data-taking periods, are modelled correctly. The probability of at least one false negative in any of these checks is estimated by running the checks over the same toy datasets (in order to account for correlations) and found to be around 6%. Adding the checks presented in Section 10.3.3 would have increased this to approximately 11%, which was deemed too high. As such, these checks are only performed to provide additional information in case of failure.

To summarise, the compatibility checks are conducted as follows: 10 000 pseudoexperiments are generated using the blinded result of the fit to data. Each compatibility check is performed on the toys by fitting the data twice with different configurations. Every check's test statistic is obtained for each toy dataset, by comparing the outcomes of the two fits involved. Finally, the checks are run on the data, and whether the data passes or fails the test is determined based on the distribution of the test statistic in toys. Dr. Paula Álvarez Cartelle generated the toys and performed the fit to data, whilst the fits to toys and the determination of the test statistics' distributions represent original work.

10.3.2 Results of the compatibility checks

This subsection presents the outcome of the polarity, window, and selection compatibility checks. In the first two checks, the chosen test statistic is the absolute difference between the R_K values from the two fits, weighted by their uncertainties:

$$t_{\text{stat}} = \frac{||R_{K1} - R_{K2}||}{\sqrt{\sigma_1^2 + \sigma_2^2}} \,. \tag{10.4}$$

Here, $R_{K1,2}$ are the values of R_K obtained from the two fits, and $\sigma_{1,2}$ are their respective uncertainties. This test statistic accounts for the fact that the two fits in each check are performed on different sets of candidates. The distributions of t_{stat}, as obtained by running the

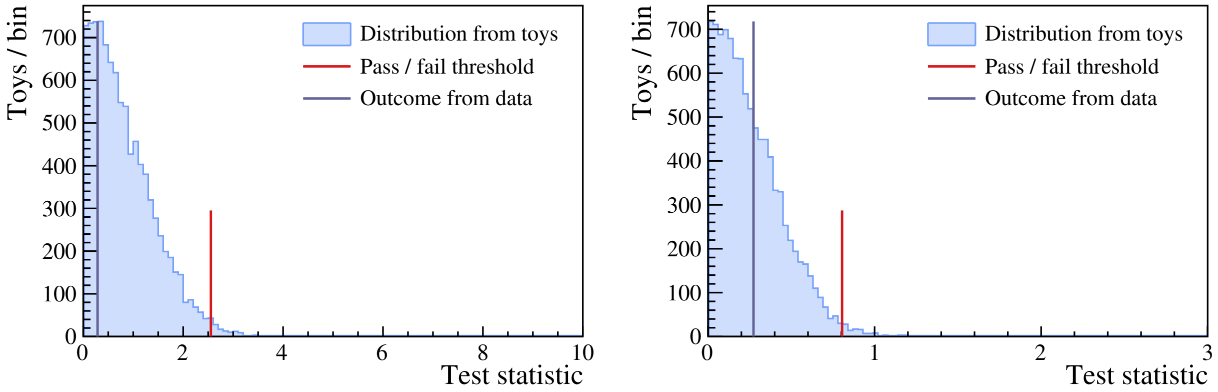

Figure 10.3: Outcome of the polarity (left) and window (right) compatibility checks. The corresponding test statistic is defined in Equation (10.4). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). The result obtained from the data (purple) is within the region for success in both checks.

polarity and window checks over toy datasets, are shown in Figure 10.3. The corresponding *p*-values for data are listed in Table 10.2. In both checks, the result obtained from running over the data lies in the region corresponding to good compatibility. This implies that data taken with the two magnet configurations are compatible, and that no mismodelling of the partially-reconstructed background is found.

The selection check considers three data-division schemes, with the fit being allowed to prefer an independent value of R_K for each subset. This is then compared to the nominal fit, which only allows one value of R_K. Scheme (a) splits the samples by data-taking period, leading to 4 separate R_K values for: Run 1, Run 2.1, 2017, and 2018. Scheme (b) distinguishes between the 3 electron trigger selection (*e*TOS, *h*TOS!, and TIS!). Finally, scheme (c) divides the data both by data-taking period and electron trigger, resulting in a fit with 12 independent R_K values.

In a given scheme, the only difference between the alternative and nominal fits is the number of degrees of freedom, since the two fits differ only in the number of independent R_K values. Due to the different numbers of dimensions between nominal and alternative fits, the test statistic used for the polarity and window checks is not appropriate. This leads to the choice of the following test statistic:

$$t_{\text{stat}} = \min(\log \mathcal{L}_{\text{alt}}) - \min(\log \mathcal{L}_{\text{nom}})\,, \tag{10.5}$$

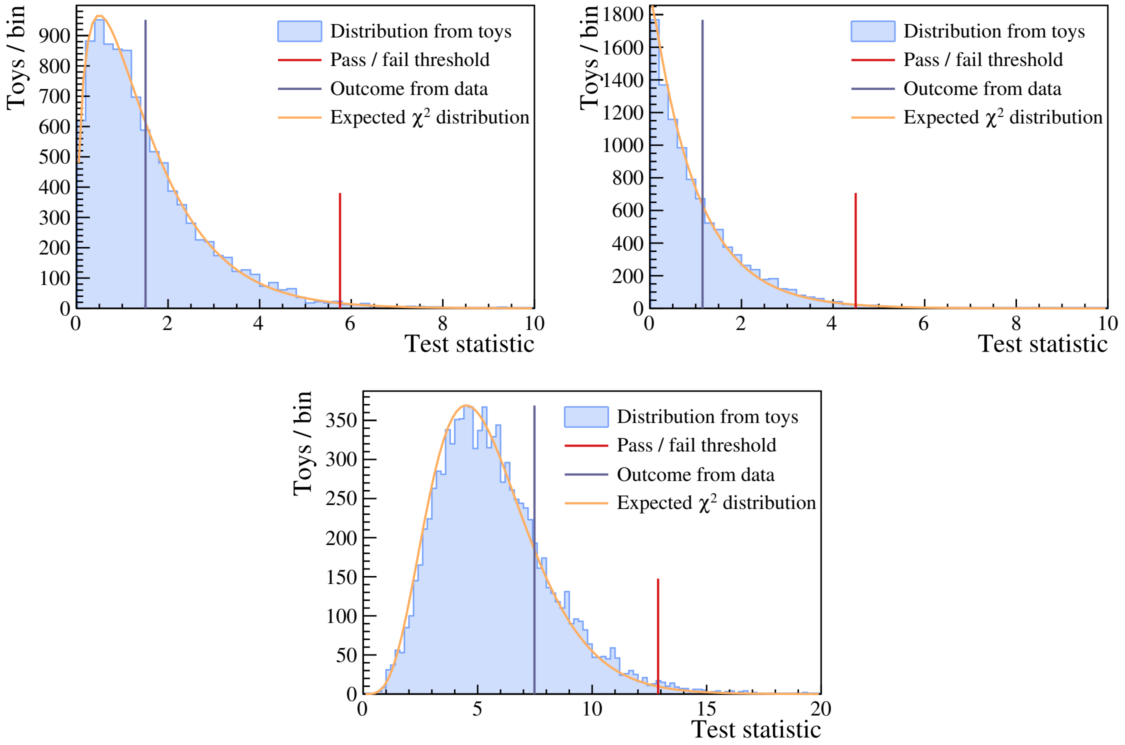

Figure 10.4: Outcome of the selection compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The corresponding test statistic is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.

where $\min(\log\mathcal{L}_{\text{nom}})$ is the minimum value of the log-likelihood, as found by the nominal fit (with one R_K estimate), and $\min(\log\mathcal{L}_{\text{alt}})$ is its counterpart from the alternative fit (with multiple values of R_K). This test statistic is the difference between two log-likelihoods constructed from the same data. This means that, in the limit of infinitely-many toys, $2t_{\text{stat}}$ follows a χ^2 distribution with number of degrees of freedom equal to the difference between the number of variables each likelihood has. This is equal to the difference between the number of R_K values each fit is allowed to find, so for example the test statistic for the scheme that splits by electron trigger follows a χ^2 distribution of $3-1=2$ degrees of freedom.

The result of the selection check is presented in Figure 10.4. Running each of the three considered splitting schemes over the data leads to significances that lie comfortably in the region for success. This demonstrates that the different data-taking periods and electron trigger strategies produce datasets that are compatible with each other.

Table 10.2: Results obtained from running the compatibility checks over the data. The table lists the values of the corresponding test statistic, its p-value, and the significance in number of standard deviations. The final three entries correspond to the three splitting schemes that make up the selection check, *i.e.* by data-taking period, by trigger, and by both.

	t_{stat}	p-value	$Z\ [\sigma]$
Polarity test,	0.2863	0.8481	0.19
Window test,	0.2767	0.4057	0.83
Selection test, scheme (a)	1.5086	0.3949	0.79
Selection test, scheme (b)	1.1501	0.3416	0.95
Selection test, scheme (c)	7.4811	0.2048	1.26

A summary of the compatibility checks is provided in Table 10.2. It shows that running the checks over the data results in t_{stat} values that are compatible with expectation from toys. This demonstrates that effects induced by changing data-taking conditions and selection strategies are suitably taken into account by the fit.

10.3.3 Additional compatibility checks

The conclusions drawn from conducting the checks presented in the previous subsection are reinforced by considering additional ways of splitting the data into subsets. These checks are functionally identical to the selection test: t_{stat} is defined in the same way, and the same three splitting schemes are used for each check. The difference lies in how the data-taking periods are defined. The five additional compatibility checks are:

1. “Run 1 vs. Run 2.1”: compatibility between data taken during Run 1 and the subset of Run 2 data that was used in the previous R_K measurement;
2. “2017 vs. 2018”: compatibility between data taken during the years 2017 and 2018;
3. “Run 2.1 vs. Run 2.2”: compatibility between the Run 2 data that were and were not used in the previous R_K measurement;
4. “Run 1 vs. Run 2”: compatibility between data taken during Run 1 and Run 2; and
5. “old vs. new”: compatibility between the data that was and was not used in the previous R_K measurement.

Table 10.3: Outcome of running the additional compatibility checks over the data. The table lists the values of the corresponding test statistic, its p-value, and the significance in number of sigmas. The schemes correspond to splitting the data by (a) data-taking period, (b) trigger, and (c) both.

Check	Scheme	t_{stat}	p-value	$Z\ [\sigma]$
Run 1	(a)	1.5005	0.0863	1.72
vs	(b)	0.2769	0.7711	0.29
Run 2.1	(c)	2.8204	0.3666	0.90
2017	(a)	0.0001	0.9871	0.02
vs	(b)	3.6903	0.0306	2.16
2018	(c)	4.6669	0.1048	1.62
Run 2.1	(a)	0.2341	0.4940	0.68
vs	(b)	2.9128	0.0608	1.87
Run 2.2	(c)	4.1763	0.1548	1.42
Run 1	(a)	1.2749	0.1165	1.57
vs	(b)	1.1500	0.3429	0.95
Run 2	(c)	5.2308	0.0780	1.76
old	(a)	0.0017	0.9529	0.06
vs	(b)	1.1501	0.3416	0.95
new	(c)	3.9591	0.1817	1.33

As a consequence of the different ways in which data-taking periods are defined, scheme (a) uses two separate R_K values, as opposed to the four used by the selection test. In addition, that the number of free R_K parameters in scheme (c) is six instead of twelve. Checks 1, 2, and 3 differ from selection also in the fact that they are not run over the entire data.

The results of these checks are summarised in Table 10.3, which lists the values of t_{stat} obtained by running the tests over the data, alongside the corresponding p-values and significances. The distributions of t_{stat} obtained from toys are presented in Appendix G. All tests show good compatibility between the studied partitions of the data. This further demonstrates that the $B^+ \to K^+\ell^+\ell^-$ fit successfully takes into account effects induced by changing data-taking conditions and selection strategies.

Chapter 11

Results

The compatibility checks presented in Section 10.3 conclude the validation of the R_K measurement strategy. The next and final step is the calculation of the final R_K estimate, as presented in this chapter. The initial three sections cover the unblinding process. First, Section 11.1 describes how the differential branching fraction of $B^+ \to K^+\mu^+\mu^-$ is estimated from the muon data. Second, Section 11.2 covers the unblinding of the Run 2.2 data and the resulting R_K estimate. Third, in Section 11.3, all data are unblinded, and R_K is obtained from the fit. As explained in Section 10.1, the resulting value has a small fit bias, and does not yet include all systematic effects. Section 11.4 contains the final result, obtained after the fit bias is corrected, and after adding the systematic uncertainties attributed to the fit model. Then, Section 11.5 presents a measurement of the differential branching fraction of $B^+ \to K^+e^+e^-$, for which the R_K result is combined with the $B^+ \to K^+\mu^+\mu^-$ result from Ref. [26]. The rest of the chapter is devoted to a few additional results: the effect of turning off all efficiency corrections is presented in Section 11.6, and the R_K estimates obtained during the blinded compatibility checks are listed in Section 11.7. This is followed by Section 11.8, which puts the updated R_K result into the context of the global EFT fits introduced in Section 2.4. The results in this chapter that constitute original work are found in Sections 11.1, 11.6, and 11.8, the rest being obtained by Dr. Paula Álvarez Cartelle.

11.1 Differential branching fraction of $B^+ \to K^+\mu^+\mu^-$

A cautious step undertaken at the beginning of the unblinding procedure is looking at just the $K^+\mu^+\mu^-$ data. The muon samples are expected to be better modelled and cleaner than their electron counterparts, so potential issues with the modelling of the $K^+e^+e^-$ data could be pre-emptively identified by checking $B^+ \to K^+\mu^+\mu^-$ candidates. For this test, the differential branching fraction of $B^+ \to K^+\mu^+\mu^-$ is estimated using each data-taking period separately. The procedure was developed and used by Dr. Thibaud Humair to obtain the Run 1 and Run 2.1 results, which are presented together with the original Run 2.2 results for comparison. Starting from the efficiency-corrected yield of $K^+J/\psi(\mu^+\mu^-)$ control-mode events, one can estimate the total number of produced B^+ mesons:

$$\mathcal{N}_{B^+} = \frac{N(K^+J/\psi(\mu^+\mu^-))}{\varepsilon(K^+J/\psi(\mu^+\mu^-))} \Bigg/ \left[\mathcal{B}(B^+ \to K^+J/\psi) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-)\right] . \tag{11.1}$$

Multiplying this expression by the differential branching fraction of $B^+ \to K^+\mu^+\mu^-$, averaged over a given q^2 interval of length Δq^2, leads to an estimate of the total number of $B^+ \to K^+\mu^+\mu^-$ decays in that given q^2 interval:

$$\mathcal{N}_{K^+\mu^+\mu^-} = \frac{N(K^+J/\psi(\mu^+\mu^-))}{\varepsilon(K^+J/\psi(\mu^+\mu^-))} \cdot \frac{\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathrm{d}q^2 \cdot \Delta q^2}{\mathcal{B}(B^+ \to K^+J/\psi) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-)} . \tag{11.2}$$

This same estimate can also be obtained by efficiency-correcting the measured $B^+ \to K^+\mu^+\mu^-$ yield, taking into account the fraction f^{q^2} of events in the chosen q^2 interval:

$$\mathcal{N}_{K^+\mu^+\mu^-} = \frac{N(K^+\mu^+\mu^-)}{\varepsilon(K^+\mu^+\mu^-)} \cdot f^{q^2} . \tag{11.3}$$

This leads to the following expression for the differential branching fraction of $B^+ \to K^+\mu^+\mu^-$, averaged over a given q^2 interval:

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}q^2} = \frac{N(K^+\mu^+\mu^-)}{\varepsilon(K^+\mu^+\mu^-)} \cdot \frac{\varepsilon(K^+J/\psi(\mu^+\mu^-))}{N(K^+J/\psi(\mu^+\mu^-))} \cdot \frac{f^{q^2}}{\Delta q^2} \cdot \mathcal{B}(B^+ \to K^+J/\psi) \cdot \mathcal{B}(J/\psi \to \mu^+\mu^-) . \tag{11.4}$$

Therefore, the quantities needed to compute the branching fraction are:

- the control-mode yields and efficiencies, as listed in Tables 5.1 and 6.4, respectively;
- the rare-mode efficiencies, also presented in Table 6.4;
- the branching fractions $\mathcal{B}(B^+ \to K^+J/\psi)$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-)$, as listed in the PDG [9];
- the fraction of events in the chosen q^2 range, f^{q^2}, which is computed based on the true $q^2 = \left\|p(\mu^+) + p(\mu^-)\right\|^2$ of simulated events. This allows the results of this check to be compared to the ones from LHCb's Run 1 measurement [26], where the q^2 was also defined in terms of the muons' 4-momenta. However, the values listed in Table 6.5 are calculated by defining q^2 as $q^2 = \left\|p(B^+) - p(K^-)\right\|^2$, so the f^{q^2} estimates used to calculate $\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$ are expected to be slightly different.

The first part of the check uses the nominal q^2 interval, $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$. The fits to signal candidates in this window are shown in Figure 11.1. The resulting yields, together with the corresponding estimates for $\mathrm{d}\mathcal{B}/\mathrm{d}q^2$, are listed in Table 11.1. The results are encouraging: not only are the branching fraction estimates in good agreement with each other, and with Ref. [26], but the yields demonstrate that, as expected, the muon dataset doubles in statistics with respect to the previous R_K measurement.

Table 11.1: Results of the check on $B^+ \to K^+\mu^+\mu^-$ data with $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$, for each data-taking period. The uncertainties are statistical only. The entry at the top of the table corresponds to the result of Ref. [26], with the first uncertainty being statistical, and the second systematic.

	$\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathrm{d}q^2$	$N(B^+ \to K^+\mu^+\mu^-)$
Ref. [26]	$(24.2 \pm 0.7 \pm 1.2) \times 10^{-9}$	
Run 1	$(24.5 \pm 0.9) \times 10^{-9}$	1024 ± 35
Run 2.1	$(24.9 \pm 0.9) \times 10^{-9}$	914 ± 34
2017	$(24.7 \pm 1.0) \times 10^{-9}$	850 ± 33
2018	$(25.3 \pm 0.9) \times 10^{-9}$	1050 ± 36

The second part of the check repeats the procedure, after splitting the nominal q^2 range into five bins. The fits are conducted separately in each bin, and Equation (11.4) is used to obtain the differential branching fraction of $B^+ \to K^+\mu^+\mu^-$, averaged over each bin. As presented in Figure 11.2, the results are found to be in good agreement with each other, and with Ref. [26].

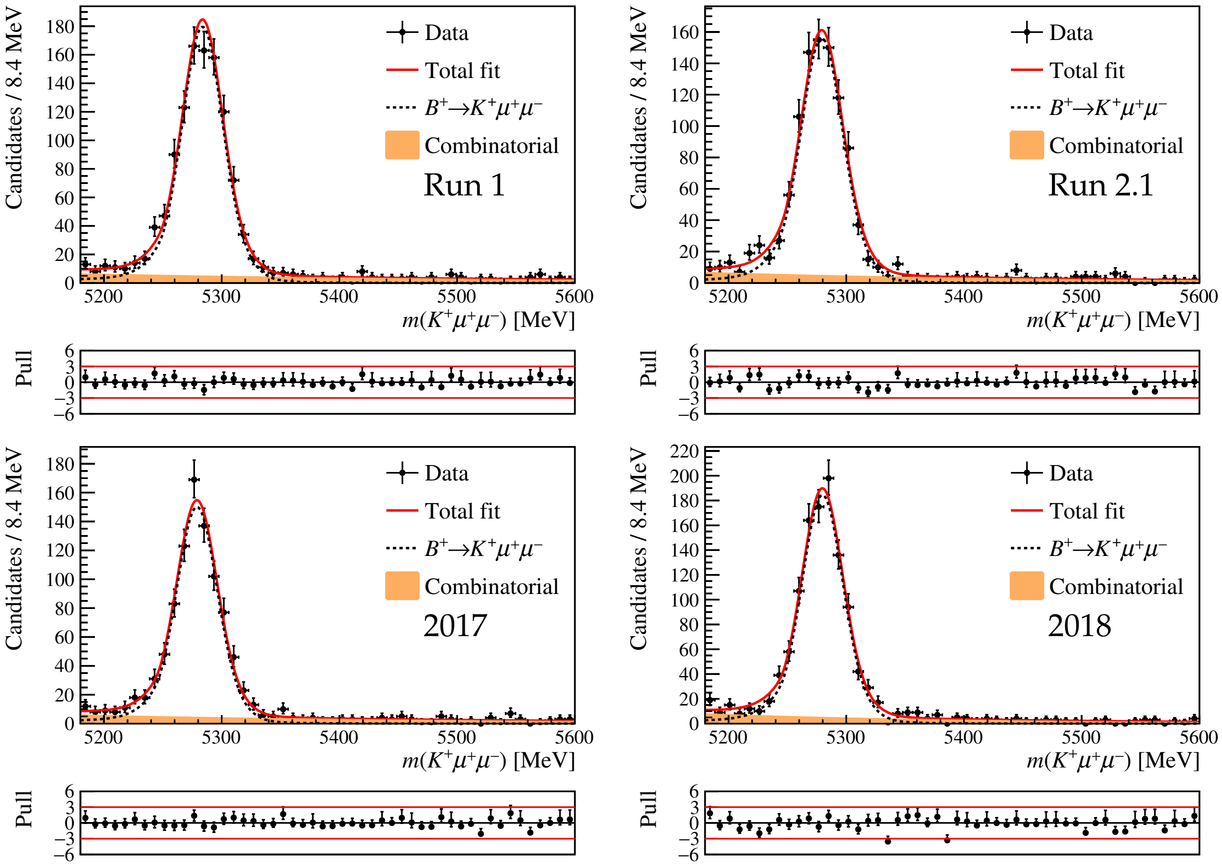

Figure 11.1: Fit to the $m(K^+\mu^+\mu^-)$ distribution of $B^+ \to K^+\mu^+\mu^-$ selected candidates in the four data samples. The red solid line shows the fit model, the black dotted line is the signal component, and the orange band depicts the combinatorial background.

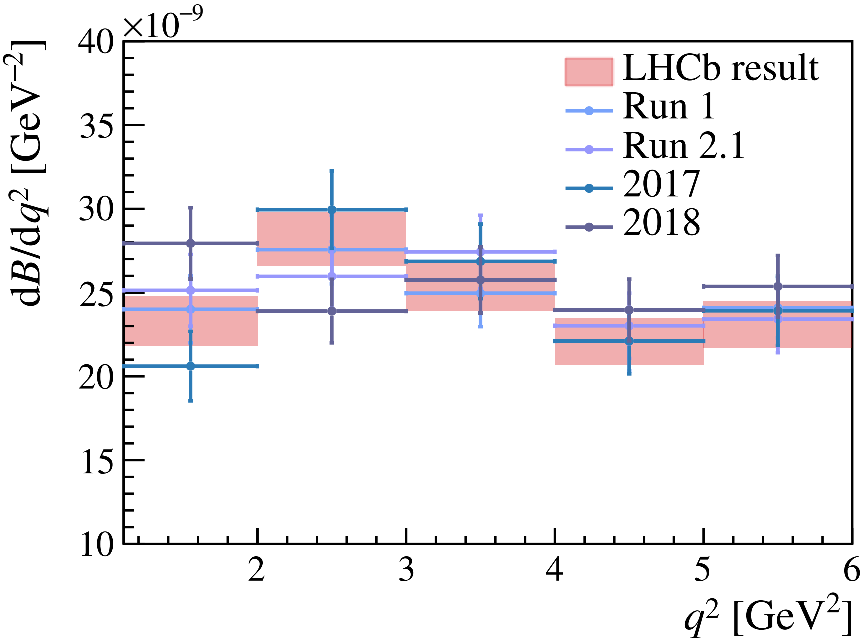

Figure 11.2: Differential branching fraction of $B^+ \to K^+\mu^+\mu^-$, averaged over several q^2 intervals. The coloured bins depict the results from the four data-taking periods. Only the statistical uncertainties are shown. The entry labelled "LHCb result" corresponds to Ref. [26].

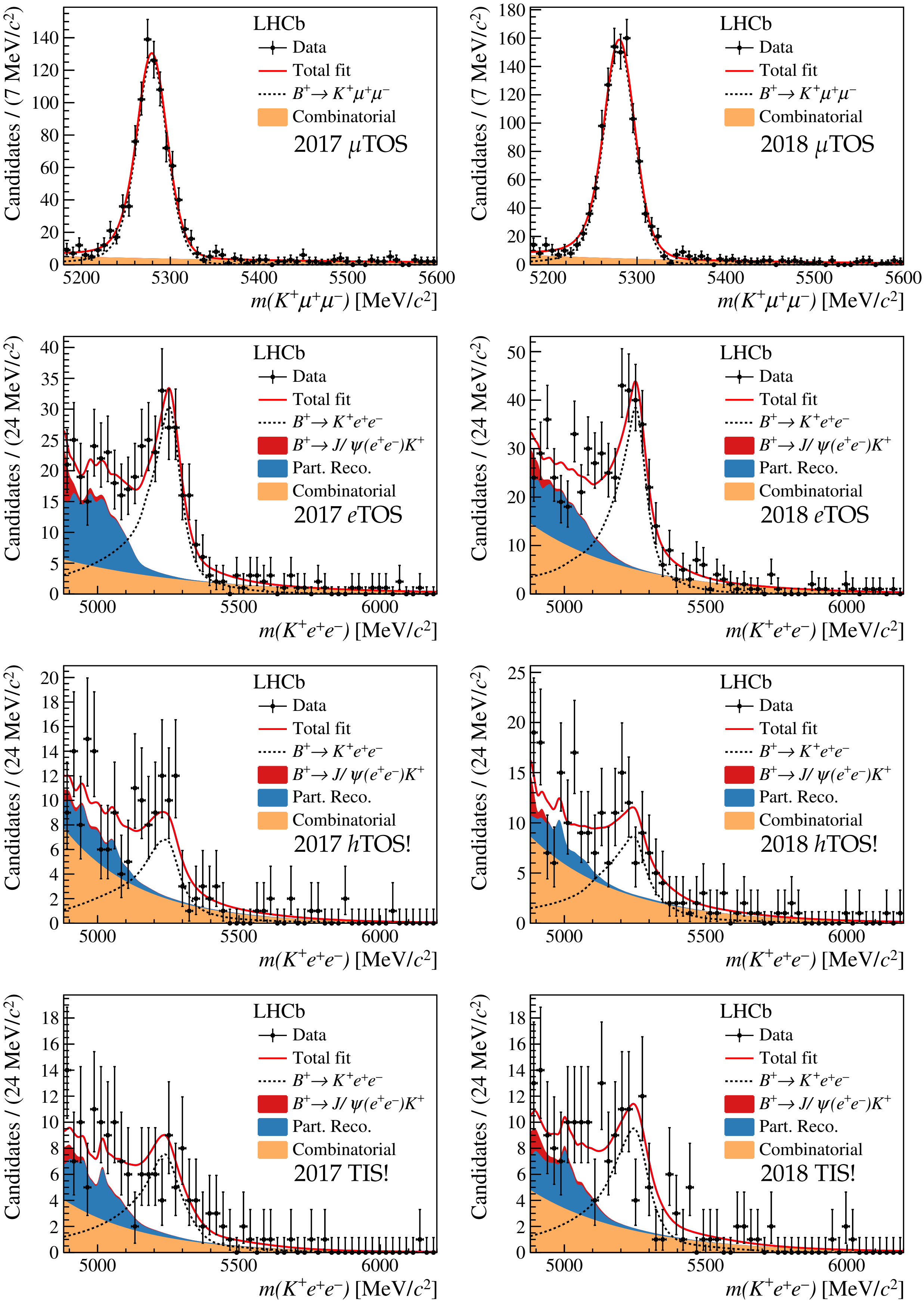

Figure 11.3: Projections of the fit to the $m(K^+\ell^+\ell^-)$ invariant-mass distribution of selected rare-mode events in 2017 (left) and 2018 (right) data. The distributions in the muon data are shown on the top row, and the subsequent rows present the electron data from each trigger selection.

11.2 Result of the fit to Run 2.2 data

The second step in the unblinding procedure is performing the fit on Run 2.2 alone. This is done to ensure the new data samples are free of any issues that may have gone undetected by the cross-checks. The projections of the fit are depicted in Figure 11.3. All pulls between the data and the fit model are found to be small, and are therefore omitted. This indicates that the fit is of good quality. The value of R_K preferred by the fit is:

$$R_K^{\text{Run 2.2}} = 0.849\,^{+\,0.062}_{-\,0.059}\,, \tag{11.5}$$

where the uncertainty includes statistical and systematic effects, with the exception of the uncertainty due to the fit model. The central value is not yet corrected to account for the fit bias, however this effect is expected to be of a few permille of the uncertainty from the fit. The precision on this estimate of R_K is comparable to the one from the previous measurement, as expected from the estimated gain in statistics.

11.3 Result of the fit to all data

With the validation of the Run 2.2 samples complete, R_K is determined using the entire set of available data. The fit is performed simultaneously on the candidates from each data-taking period and trigger selection, and the projections are merged together for illustration purposes. The projections for new and previous data from each trigger selection are shown in Figure 11.4, whilst the projections of the fit across all $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ candidates are presented in Figure 11.5.

Based on the estimated gains listed in Table 6.6 and the yields from the previous analysis, 1 631 and 3 816 signal events are expected in the electron and muon samples, respectively. From the fit, they are found to be 1 640 ± 73 and 3 845 ± 69, respectively. Note that these quantities are derived from parameters of interest, meaning that their central values and uncertainties are not extracted directly from the fit. Nevertheless, they are in very good agreement with the expected gain estimated in Section 6.7.1. Note that the total muon yield is not expected to be identical to the sum of the individual yields in Table 11.1. This is because

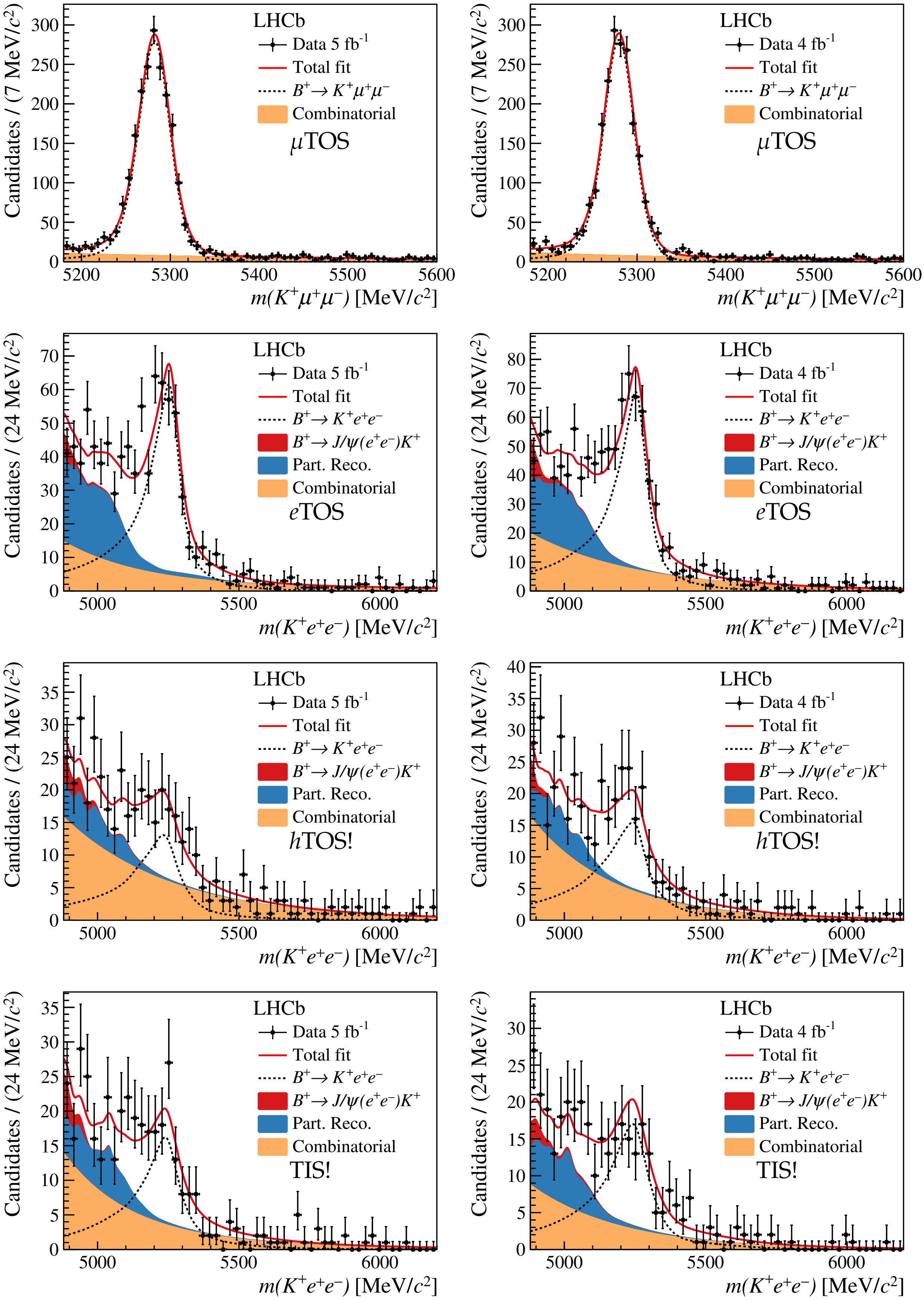

Figure 11.4: Projections of the fit to the $m(K^+\ell^+\ell^-)$ invariant-mass distribution of selected rare-mode events in previous (left) and new (right) data. The distributions in the muon data are shown on the top row, and the subsequent rows present the electron data for each trigger selection.

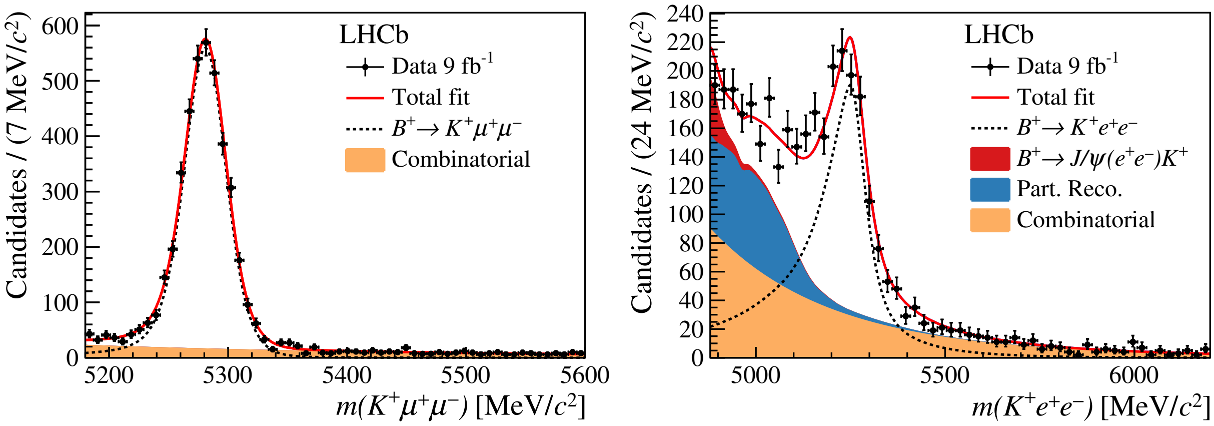

Figure 11.5: Projections of the fit to the $m(K^+\ell^+\ell^-)$ invariant-mass distribution of selected $B^+ \to K^+\mu^+\mu^-$ (left) and $B^+ \to K^+e^+e^-$ (right) data.

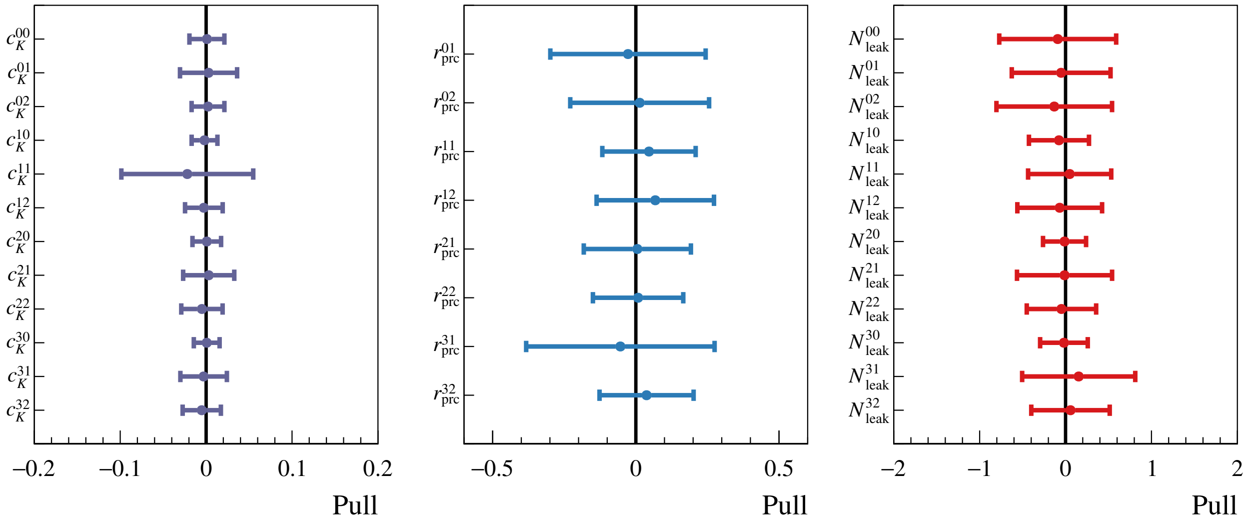

Figure 11.6: Pulls between the values of c_K^{rt} (left), r_{prc}^{rt} (middle), and N_{leak}^{rt} (right) at the point where the likelihood reaches its maximum, and the central values of the constraints on said parameters. The index r runs from 0 to 3 to represent, in order, Run 1, Run 2.1, 2017, and 2018 data. Similarly, the index t is 0 for eTOS, 1 for hTOS!, and 2 for TIS!.

in the latter case the NLL to be minimised only contains the model for each individual muon sample, rather than the full expression given in Equation (9.5).

It is also checked *a posteriori* how much, if at all, the parameters presented in Section 9.3 deviate from the central values of their constraints. It can be seen in Figure 11.6 that all c_K^{rt}, r_{prc}^{rt}, and N_{leak}^{rt} parameters show close to no shift from the central values of their constraints to the values that maximise the likelihood. This further demonstrates that the simulation accurately describes the data, and that, as a consequence, the fit prefers to adjust R_K as it converges onto the point in parameter space where the likelihood is maximal.

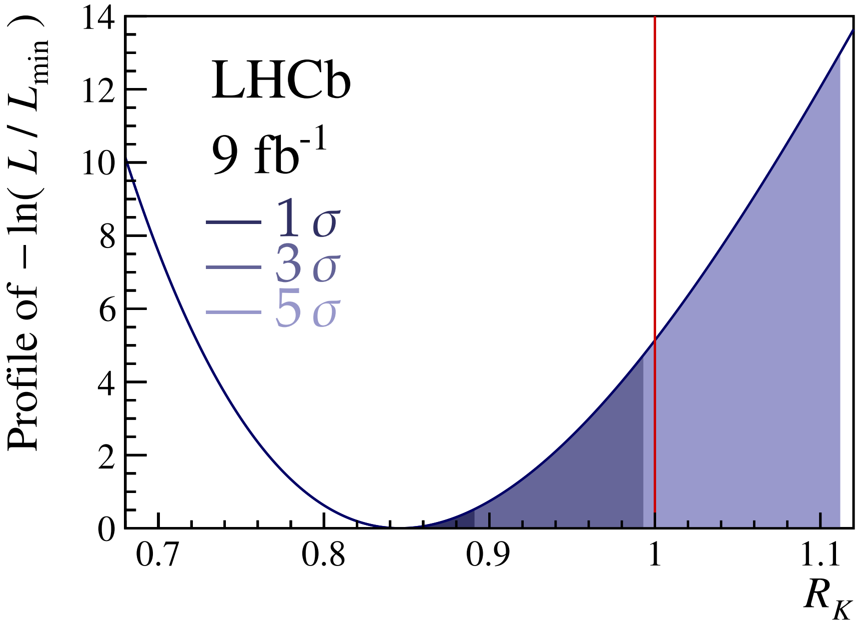

Figure 11.7: Log-likelihood function profiled as a function of R_K. The dark, medium and light purple regions show the R_K values allowed at 1σ, 3σ, and 5σ, respectively. The red line indicates the prediction from the SM.

11.4 Measured value of R_K

The fit induces a small bias on the extracted value of R_K. This bias is evaluated using the procedure described in Section 10.1, and found to be $(-0.98 \pm 1.03)\%$ of the uncertainty from the fit. This means that R_K is slightly underestimated by the fit, and the central value is corrected accordingly.

The final step towards the R_K result is the inclusion of the systematic effects induced by the fit model, presented in Section 10.2. These uncertainties are combined with those from the fit by convolving the profiled likelihood with a Gaussian distribution, whose width is equal to the size of the systematic effects. This smeares the profiled likelihood to reflect the impact of the systematic uncertainty induced by the fit. Figure 11.7 shows the final NLL profiled in R_K, together with the SM expectation. The departure from a symmetric distribution arises from the definition of R_K. The denominator is affected by larger statistical uncertainties than the numerator, owing to the different efficiencies of the muon and electron channels. However, the intervals of the likelihood distribution are found to be identical when estimated with $1/R_K$ as the fit parameter.

The result is obtained from the point that maximises the likelihood, and is found to be:

$$R_K = 0.846 \,^{+\,0.044}_{-\,0.041}\,, \tag{11.6}$$

where the uncertainties combine statistical and systematic effects. To disentangle the statistical component of the total uncertainty, the NLL minimisation is repeated after fixing the c_K^{rt} parameters to their values at the minimum. All systematic effects are hence turned off, resulting in an R_K estimate whose uncertainty is purely statistical. This leads to the final result on R_K, with separated statistical and systematic uncertainties:

$$R_K = 0.846 \,^{+\,0.042}_{-\,0.039}\,^{+\,0.013}_{-\,0.012}\,. \tag{11.7}$$

The statistical precision on this result is approximately 5%. This represents an improvement by a factor of 1.4 with respect to the previous measurement, as expected from the doubling of statistics. The measurement is dominated by the statistical uncertainty, given that the systematic uncertainty of 1.5% is much smaller.

The level of compatibility with the SM is evaluated using the profiled NLL. The SM prediction is expected to vary by 1% as a result of possible imperfections in the QED corrections generated by PHOTOS [71,74–82]. This is taken into account by reparametrising the likelihood in terms of the SM prediction, R_K^{SM}, and the distance ΔR_K between the experimental result and R_K^{SM}. The likelihood is multiplied by a Gaussian constraint on R_K^{SM}, of mean 1 and standard deviation 1%. The profile of this likelihood in ΔR_K is interpreted as the posterior PDF of ΔR_K, for which a uniform prior is assumed. The p-value of the SM hypothesis is computed by integrating this posterior PDF above $\Delta R_K = 0$, and found to be equal to 10^{-3}. The corresponding significance in terms of standard deviations is calculated using the inverse Gaussian cumulative distribution function for a one-sided conversion. The R_K measurement is found to be $3.1\,\sigma$ away from the SM prediction, giving evidence for the violation of lepton flavour universality in $B^+ \to K^+\ell^+\ell^-$ decays.

11.5 Differential branching fraction of $B^+ \to K^+e^+e^-$

Combining the value of R_K with the measurement in $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$ from Ref. [26],

$$\frac{\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathrm{d}q^2} = (24.2 \pm 0.7\,(\text{stat.}) \pm 1.2\,(\text{syst.})) \times 10^{-9}\ \text{GeV}^{-2}\,, \tag{11.8}$$

the differential branching fraction of $B^+ \to K^+e^+e^-$ is measured over $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$. The uncertainties incorporated by this combination are summarised in Table 11.2. Assuming that the Run 1 $B^+ \to K^+\mu^+\mu^-$ sample in the R_K analysis overlaps entirely with the sample used in Ref. [26], the correlation between the statistical uncertainty of $\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$ and R_K is estimated to be 21%. The dominant systematic effect in $\mathcal{B}$ $(B^+ \to K^+\mu^+\mu^-)$ is the uncertainty on the branching fraction of the $B^+ \to K^+J/\psi(\mu^+\mu^-)$ normalisation channel. It is not correlated with R_K, and found to be approximately 4.2%. The remaining systematic effects, associated with imperfections in the simulation and the physics model, are assumed to be fully correlated with the equivalent systematics in R_K.

Table 11.2: Uncertainties on R_K and $\mathcal{B}(B^+ \to K^+\mu^+\mu^-)$ that are relevant to the calculation of $\mathcal{B}(B^+ \to K^+e^+e^-)$. All values are percentages.

	$\mathrm{d}\mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathrm{d}q^2$	R_K
Statistical	2.9	$^{+5.0}_{-4.6}$
Efficiencies	2.6	$^{+1.2}_{-1.0}$
R_K fit model & bias	-	1.0
$\mathcal{B}$ $(B^+ \to K^+J/\psi(\mu^+\mu^-))$	4.2	-
Total	5.7	$^{+5.2}_{-4.8}$

The result for $\mathrm{d}\mathcal{B}(B^+ \to K^+e^+e^-)/\mathrm{d}q^2$, averaged over the q^2 range used in the R_K analysis, is:

$$\frac{\mathrm{d}\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathrm{d}q^2} = \left(28.6\,^{+1.5}_{-1.4} \pm 1.3\right) \times 10^{-9}\ \text{GeV}^{-2}\,, \tag{11.9}$$

where the first uncertainty is statistical, and the second systematic. This measurement is depicted in Figure 11.8, alongside results from the muon modes [26] and the SM predictions [198, 199]. The electron data is found to be in better agreement with the SM than the muon data.

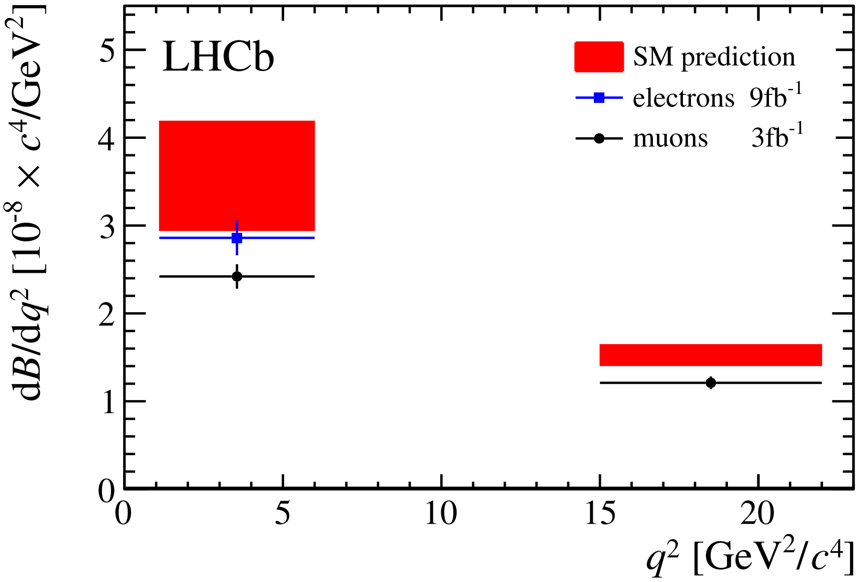

Figure 11.8: Differential branching fractions of $B^+ \to K^+e^+e^-$ (blue) and $B^+ \to K^+\mu^+\mu^-$ (black), where the latter come from Ref. [26]. The SM predictions (red) from Refs. [198, 199] are also shown for comparison. Plot created by Dr. Patrick Owen.

11.6 Effect of turning off the efficiency corrections

A significant portion of the R_K analysis is devoted to understanding and correctly measuring the efficiencies. As expanded upon in Chapter 6, some aspects of the data are not reproduced by the simulation perfectly, hence the need for efficiency corrections. This section describes an exercise that quantifies the effect these corrections have on R_K.

In Equation (9.3), it is shown that information from the efficiencies is encoded into the fit (and hence into R_K) through the c_K^{rt} parameters. For this reason, the effect of the corrections is evaluated by fitting the $B^+ \to K^+\ell^+\ell^-$ data using c_K^{rt} values obtained without applying any corrections to the simulation. This is then compared to the outcome of a fit that's given the nominal c_K^{rt} values. In both cases, the c_K^{rt} are fixed to their estimates from simulation, in order to isolate the effect of the efficiency calibration. The quantity of interest is the shift in R_K induced by not applying the corrections to simulation.

The c_K^{rt} parameters are recalculated using Equation (9.3) by turning off the efficiency corrections, and by rerunning the fits to the control data using signal and background models that are not adjusted by the calibration weights. These templates are found to be similar to their

Table 11.3: The c_K^{rt} factors used to evaluate the impact of efficiency corrections on R_K. The values in the left-most column take into account the corrections, so are identical to the central values of the constraints used in the nominal fit. The middle column shows how these parameters change by turning off the efficiency corrections. The changes listed in the right-most column are defined as the differences between the values without and with corrections, relative to the corrected ones.

Data-taking period		c_K^{rt} With corrections	c_K^{rt} Without corrections	Change [%]
Run 1	eTOS	0.1436	0.1399	−2.58
	hTOS!	0.0509	0.0478	−6.09
	TIS!	0.0548	0.0530	−3.28
Run 2.1	eTOS	0.2809	0.2773	−1.28
	hTOS!	0.0710	0.0697	−1.83
	TIS!	0.0806	0.0772	−4.22
2017	eTOS	0.2403	0.2335	−2.83
	hTOS!	0.0709	0.0675	−4.80
	TIS!	0.0835	0.0785	−5.99
2018	eTOS	0.2296	0.2200	−4.18
	hTOS!	0.0765	0.0771	0.78
	TIS!	0.0788	0.0748	−5.08

nominal, corrected counterparts, so the impact on the $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ yields is negligible. However, the efficiencies change by a few percent, as reflected in the resulting c_K^{rt} estimates. They are presented in Table 11.3 together with the nominal values. The dominant trigger category, eTOS, shows downwards changes of around 3%, so it is expected that R_K would shift by a similar amount. Indeed, the R_K values obtained from the two fits,

$$R_K^{\text{corr.}} = 0.848^{+0.042}_{-0.039}, \text{ and} \tag{11.10}$$

$$R_K^{\text{uncorr.}} = 0.823^{+0.041}_{-0.038}, \tag{11.11}$$

differ by 3.1%. This is comparable to the overall uncertainty on R_K, and the estimate that ignores corrections, $R_K^{\text{uncorr.}}$, has a smaller value than the one which does take into account corrections, $R_K^{\text{corr.}}$. This means that the efficiency corrections are necessary to avoid a result that's biased away from the SM expectation by an amount comparable to the overall precision. The effect on $r_{J/\psi}$ is found to be larger: 15% – 40%, depending on the data-taking period and trigger selection. This further demonstrates the power of the double ratio to suppress systematic biases that affect resonant and control data similarly.

11.7 Unblinded compatibility checks

With the unblinding of the final result on R_K, it is possible to inspect the values resulting from the compatibility checks presented in Section 10.3. The agreement between these values was assessed while they were still blinded, using suitable test statistics whose distributions were obtained from statistically-representative pseudoexperiments. No significant tension was found in any of the tests performed, and so it was decided to move forward to the unblinding procedure. The values of R_K preferred by the fit configurations described in Section 10.3.1 are summarised in Figure 11.9. Additional compatibility checks were conducted and presented in Section 10.3.3.

Two noteworthy results are the large p-values for compatibility between 2017 and 2018 data, and between these two samples combined (Run 2.2) and the previous dataset. Before unblinding, this suggested that the addition of Run 2.2 data would not change the central value by a substantial amount. However, it could not completely rule out unusual features of the likelihood, such as multiple minima or correlations in the fit that would be able to accommodate several R_K estimates. Even so, Figure 11.9 shows that the 2017 and 2018 data prefers nearly the same value of R_K, which is almost identical to the one obtained from the fit to the entire dataset. This explains why the result presented in Equation (11.7) is very close to the one given in Equation (10.3), thus providing a remarkable demonstration of the robustness of the R_K measurement.

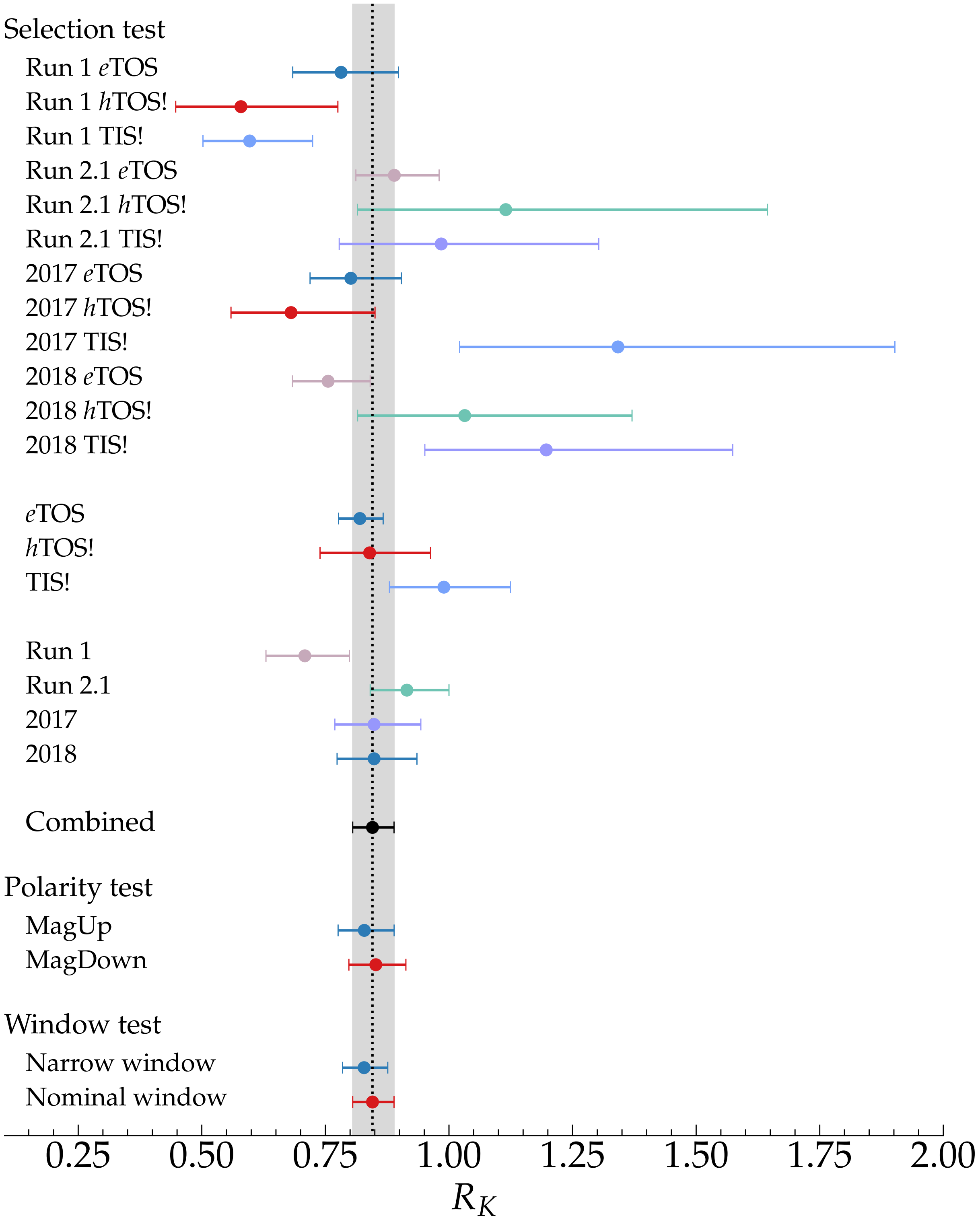

Figure 11.9: Unblinded values of R_K obtained from the three compatibility checks presented in Section 10.3. The uncertainties do not contain systematic effects induced by the fit procedure, and the central values are not corrected to account for fit biases. The dashed vertical line represents the combined results, and its uncertainty is depicted by the grey band.

11.8 Impact on the landscape of anomalies

As a result of the approximate doubling of statistics with respect to the previous analysis, the tension between the measured value of R_K and the SM prediction increases from $2.5\,\sigma$ to $3.1\,\sigma$. This subsection covers an additional study of how impactful the improvement in precision is. The study represents original work, and is performed in the context of the EFT framework introduced in Section 2.4. Likelihood surfaces that depend on Wilson coefficients are calculated using the R_K measurements before and after the addition of Run 2.2 data. The contours of the likelihood surfaces are then compared, in order to put the increased statistical power into the context of global fits to Wilson coefficients.

Like the fits presented in Figure 2.8, the study in this section is performed using the `flavio` software package [79]. The SM nuisance parameters are set in their default configuration, and NP is assumed to only impact certain Wilson coefficients. At leading order, the observable R_K is only sensitive to the C_9 and C_{10} coefficients in the electron and muon sectors. In the context of the results presented in Figure 11.8, $B^+ \to K^+e^+e^-$ is assumed to be SM-like. Therefore, this study only considers NP contributions to the muonic coefficients, denoted by $C_9^{bs\mu\mu}$ and $C_{10}^{bs\mu\mu}$.

Figure 11.10 shows the likelihood surface as a function of the two NP contributions to the Wilson coefficients, before and after the addition of Run 2.2 data. It can be seen that the increased statistics reduce the $1\,\sigma$ and $3\,\sigma$ confidence regions, leading to the SM being disfavoured by more than $3\,\sigma$. This is compatible with the $3.1\,\sigma$ tension obtained from the likelihood shown in Figure 11.7. The best-fit points from both R_K results are compatible, and are situated close to the $C_9^{bs\mu\mu} = -C_{10}^{bs\mu\mu}$ line. This is consistent with a $V-A$ NP interaction[1]. To test this hypothesis against the SM, the likelihood is recalculated after fixing NP contributions to C_{10} to have equal magnitudes and opposite signs to C_9 contributions. The result is shown in Figure 11.11. The minimum from the current 9 fb^{-1} R_K result is found to be at $C_9^{bs\mu\mu} = -0.32 \pm 0.10$, with the 5 fb^{-1} equivalent at $C_9^{bs\mu\mu} = -0.30 \pm 0.13$. The uncertainty scales by a factor of approximately $\sqrt{2}$, notwithstanding the effect of the theory uncertainties. The updated and previous R_K results prefers the $C_9^{bs\mu\mu} = -C_{10}^{bs\mu\mu}$ NP hypothesis over the SM

---

[1] In this context, $V - A$ stands for "vector minus axial-vector". An example of a $V - A$ interaction in the SM is the charged-current weak interaction.

hypothesis by approximately 3.5 and 2.3 standard deviations, respectively. These tensions are higher than in the case where both Wilson coefficients are left freely floating, due to the reduction by one degree of freedom. This is another example of the power provided by the doubling of statistics compared to the previous measurement, whilst also showing that the SM is even more disfavoured by the data.

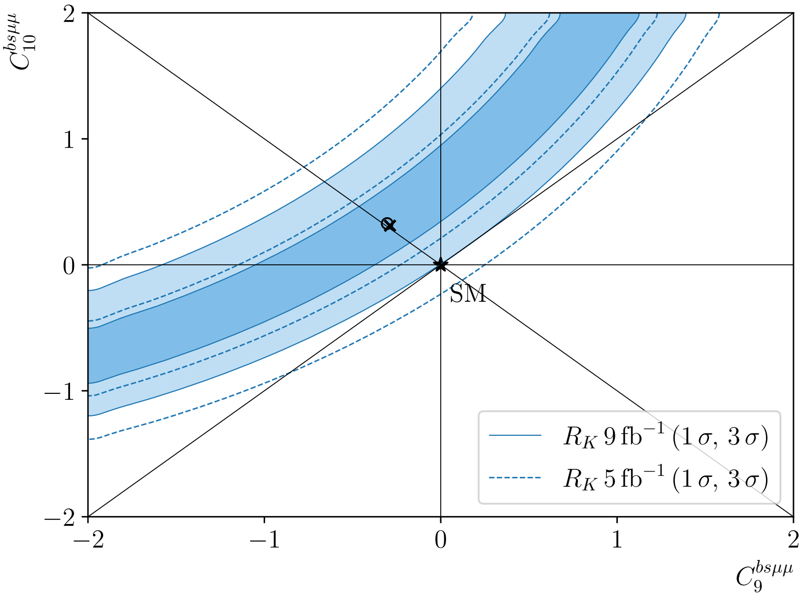

Figure 11.10: The $1\,\sigma$ and $3\,\sigma$ contours obtained from the current and previous R_K measurements by allowing NP contributions to the muonic C_9 and C_{10} Wilson coefficients. The circle and the cross depict the best-fit points to the current and previous LHCb R_K measurements, respectively.

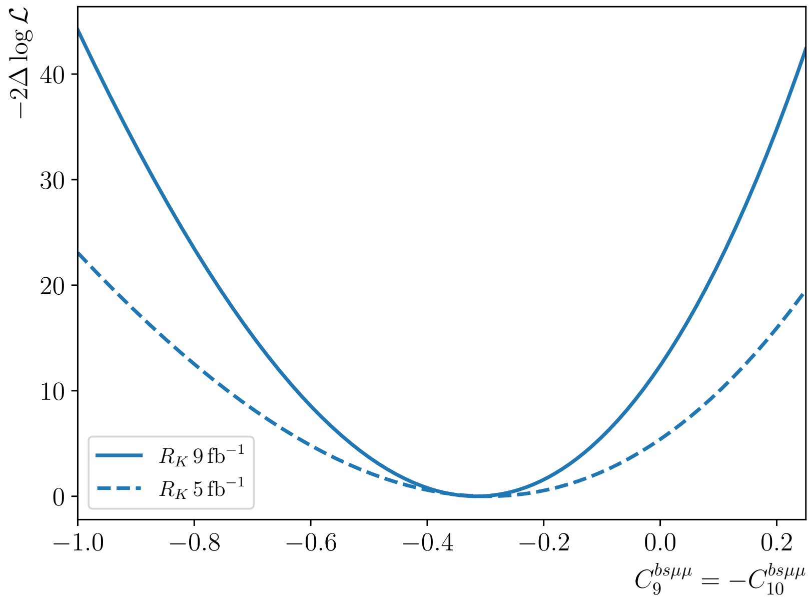

Figure 11.11: Likelihood scan of equal but opposite-sign NP contributions to C_9 and C_{10}, using the current and previous LHCb R_K results.

Chapter 12

Conclusion

This thesis gives the first evidence for the violation of lepton flavour universality, in beauty-quark decays [1]. The ratio of branching fractions R_K is measured using 9 fb^{-1} of proton-proton collision data recorded by the LHCb experiment at CERN. The result is:

$$R_K = 0.846\,^{+\,0.042}_{-\,0.039}\,^{+\,0.013}_{-\,0.012}\,, \tag{12.1}$$

which is in tension with the Standard Model prediction at the level of 3.1 σ. Using EFT at scales close to m_b to test the SM null hypothesis against certain NP scenarios, it is found that the R_K measurement prefers NP by over 3 σ. Such EFT fits provide useful information when building complete NP models that could explain R_K and other flavour anomalies, such as leptoquarks or Z' vector bosons.

Verification from other experiments of the LHCb R_K measurement is expected in the near future. The Belle II collaboration [200, 201] begun taking data shortly after LHCb's Run 2 ended. As the successor of the Belle collaboration that provided some of the results discussed in Section 2.3, Belle II is expected to have a significant impact on the landscape of flavour anomalies, and beyond. Furthermore, the CMS collaboration has collected samples of b-hadron decays [202], which could be used to measure LFU ratios such as R_K and R_{D^*}.

By the time other experiments will have measured the anomalous flavour observables, LHCb is expected to have conducted additional studies of the anomalies. Measuring R_K in other q^2 regions, such as $q^2 > m^2_{\psi(2S)}$, would provide complementary information to the measurement at $q^2 \in (1.1\ \mathrm{GeV}^2,\ 6.0\ \mathrm{GeV}^2)$. Such high-$q^2$ regions are expected to feature background sources that are different to the ones present in the measurement described by this thesis. Therefore, measuring R_K at high-q^2 would provide important complementary information in the study of LFU. In addition, the current LHCb measurement of $R_{K^{*0}}$ only uses Run 1 data, therefore a result approximately twice as precise is expected in the near future. Since $R_{K^{*0}}$ is sensitive to violation of LFU, an updated result is expected to heavily influence the landscape of the flavour anomalies. Another ratio to consider is R_{pK}, which has been measured by LHCb [64] using Run 1 and 2016 data. Given that it involves baryons in the final state, R_{pK} provides complementary information to the other ratios, at the cost of reduced precision.

Data collection by the LHCb collaboration is expected to restart at the beginning of 2022, when Run 3 is scheduled to begin. As a result of the upgraded trigger [203], LHCb is expected to collect data at higher rates, leading to as much as 23 fb^{-1} of total integrated luminosity. This is estimated to increase to 50 fb^{-1} during Run 4, when the LHC is scheduled to be upgraded to run at higher instantaneous luminosities. After Run 4, it is projected that the LHCb collaboration could be able to collect as much as 300 fb^{-1} of proton-proton collision data [204]. Since results on FCNC R_H ratios are dominated by statistical uncertainties, such increases to the amount of available data are important to the understanding of the flavour anomalies. This is reflected in the reduction of uncertainties on the already-measured ratios, as well as in the enabling of R_H tests based on other hadronic systems H. In particular, $b \to d\ell^+\ell^-$ observables, such as R_π, where the b decays into the first quark generation rather than the second, would be able to provide complementary constraints on the current NP models that attempt to explain the anomalies. Such a measurement is currently challenging at LHCb, because $B^+ \to \pi^+\ell^+\ell^-$ modes are suppressed with respect to their kaon counterparts by a factor of approximately 25. However, it can be seen in Table 12.1 that several R_H ratios, including R_π, are expected to have statistical precision of a few percent with 300 fb^{-1} of data. In particular, R_K and $R_{K^{*0}}$ are expected to only become systematically-dominated after Run 4. Note that since the R_K result presented in this thesis has a statistical uncertainty of 5%, it

Table 12.1: Extrapolation, based on Run 1 results, of statistical uncertainties on R_H ratios and corresponding expected electron-mode yields. The $b\overline{b}$ production cross-section is assumed to scale linearly with centre-of-mass energy, and the detector performance is assumed to be unchanged with respect to Run 1. The extrapolated $B^+ \to K^+e^+e^-$ yield at 9 fb^{-1} differs from the yield given in Section 11.3. This is attributed to the different central values of R_K, alongside the better detector performance in Run 2 compared to Run 1. Table adapted from Ref. [204].

R_H precision	Run 1 result	9 fb^{-1}	23 fb^{-1}	50 fb^{-1}	300 fb^{-1}
R_K	$0.745 \pm 0.090 \pm 0.036$	0.043	0.025	0.017	0.007
$R_{K^{*0}}$	$0.69 \pm 0.11 \pm 0.05$	0.052	0.031	0.020	0.008
R_ϕ	–	0.130	0.076	0.050	0.020
R_{pK}	–	0.105	0.061	0.041	0.016
R_π	–	0.302	0.176	0.117	0.047
Yield	**Run 1 result**	**9 fb^{-1}**	**23 fb^{-1}**	**50 fb^{-1}**	**300 fb^{-1}**
$B^+ \to K^+e^+e^-$	254 ± 29	1 120	3 300	7 500	46 000
$B^0 \to K^{*0}e^+e^-$	111 ± 14	490	1 400	3 300	20 000
$B^0_s \to \phi e^+e^-$	–	80	230	530	3 300
$\Lambda^0_b \to pKe^+e^-$	–	120	360	820	5 000
$B^+ \to \pi^+e^+e^-$	–	20	70	150	900

surpasses the extrapolated precision of 6% presented in Table 12.1.

Another strategy not yet fully exploited by LHCb is to measure ratios of FCNC branching fractions involving taus. If NP couples preferentially to both the heaviest quark and the heaviest lepton generations, it is expected that the largest deviations from the SM are to be found in flavour anomalies containing taus in the final state, such as the $R_D^{(*)}$ anomalies discussed in Section 2.3.4. However, the tau decays into at least one neutrino, which typically escapes detection. This results in experimental challenges that are qualitatively different from the ones faced by, for example, the electron modes. Nevertheless, once experimental techniques are developed to address these difficulties, tauonic FCNCs are expected to lead to an increased understanding of potential NP in beauty-quark decays.

Flavour physics is at a notable point in its history. Through the measurement presented in this thesis, evidence has been provided for the first time for the violation of lepton flavour universality, in decays of beauty quarks. This result is related to other observations, and the global picture points towards possible particles and interactions that are not currently accounted for in the Standard Model of particle physics. The situation is not yet conclusive, however studies in the near future are expected to make considerable contributions to our understanding of the flavour anomalies. These are certainly exciting times.

Bibliography

[1] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, `arXiv:2103.11769`, to appear in Nature Physics.

[2] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **122** (2019) 191801, `arXiv:1903.09252`.

[3] M. Thomson, *Modern particle physics*, Cambridge University Press, New York, 2013.

[4] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[5] Y. Grossman and P. Tanedo, *Just a Taste: Lectures on Flavor Physics,* in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics (TASI 2016): Boulder, CO, USA, June 6-July 1, 2016*, pp. 109–295, 2018. `arXiv:1711.03624`. doi: 10.11429789813233348_0004.

[6] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964), no. 9 321.

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964), no. 16 508.

[8] T. W. B. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967), no. 5 1554.

[9] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020), no. 8 083C01.

[10] T. Humair, *Testing lepton universality in penguin decays of beauty mesons using the LHCb detector,* CERN-THESIS-2019-044.

[11] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[12] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[13] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.

[14] ALEPH, DELPHI, L3, OPAL, SLD, LEP collaborations Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, `arXiv:hep-ex/0509008`.

[15] ALEPH, DELPHI, L3, OPAL, LEP collaborations, S. Schael *et al.*, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119, `arXiv:1302.3415`.

[16] PiENu collaboration, A. Aguilar-Arevalo *et al.*, *Improved Measurement of the* $\pi \to e\nu$ *Branching Ratio*, Phys. Rev. Lett. **115** (2015) 071801, `arXiv:1506.05845`.

[17] NA62 collaboration, C. Lazzeroni *et al.*, *Precision Measurement of the Ratio of the Charged Kaon Leptonic Decay Rates*, Phys. Lett. **B719** (2013) 326, `arXiv:1212.4012`.

[18] KEDR collaboration, V. M. Aulchenko *et al.*, *Measurement of the ratio of the leptonic widths* $\Gamma_{ee}/\Gamma_{\mu\mu}$ *for the* J/ψ *meson*, Phys. Lett. **B731** (2014) 227, `arXiv:1311.5005`.

[19] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[20] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[21] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[22] H. Georgi, *Effective field theory*, Ann. Rev. Nucl. Part. Sci. **43** (1993) 209.

[23] T. Blake, T. Gershon, and G. Hiller, *Rare b hadron decays at the LHC*, Ann. Rev. Nucl. Part. Sci. **65** (2015) 113, `arXiv:1501.03309`.

[24] F. L. Wilson, *Fermi's theory of beta decay*, American Journal of Physics **36** (1968) 1150.

[25] LHCb collaboration, R. Aaij *et al.*, *First measurement of the differential branching fraction and CP asymmetry of the* $B^+ \to \pi^+\mu^+\mu^-$ *decay*, JHEP **10** (2015) 034, `arXiv:1509.00414`.

[26] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of* $B \to K^{(*)}\mu^+\mu^-$ *decays*, JHEP **06** (2014) 133, `arXiv:1403.8044`.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in* $B^0 \to K^+\pi^-\mu^+\mu^-$ *decays and the* $B^0 \to K^*(892)^0\mu^+\mu^-$ *differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, `arXiv:1606.04731`.

[28] LHCb collaboration, R. Aaij *et al.*, *Branching fraction measurements of the rare* $B^0_s \to \phi\mu^+\mu^-$ *and* $B^0_s \to f_2'(1525)\mu^+\mu^-$ *decays*, `arXiv:2105.14007`, to appear in PRL.

[29] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of* $\Lambda^0_b \to \Lambda\mu^+\mu^-$ *decays*, JHEP **06** (2015) 115, Erratum ibid. **09** (2018) 145, `arXiv:1503.07138`.

[30] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay* $B^0_s \to \phi\mu^+\mu^-$, JHEP **09** (2015) 179, `arXiv:1506.08777`.

[31] LHCb collaboration, R. Aaij *et al.*, *Angular moments of the decay* $\Lambda^0_b \to \Lambda\mu^+\mu^-$ *at low hadronic recoil*, JHEP **09** (2018) 146, `arXiv:1808.00264`.

[32] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay using* $3\,fb^{-1}$ *of integrated luminosity*, JHEP **02** (2016) 104, `arXiv:1512.04442`.

[33] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2018) 047, `arXiv:1805.04000`.

[34] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, `arXiv:1710.02846`.

[35] CMS collaboration, A. M. Sirunyan *et al.*, *Angular analysis of the decay $B^+ \to K^*(892)^+\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **04** (2021) 124, `arXiv:2010.13968`.

[36] Belle collaboration, S. Wehle *et al.*, *Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, `arXiv:1612.05014`.

[37] BaBar collaboration, B. Aubert *et al.*, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$*, Phys. Rev. **D73** (2006) 092001, `arXiv:hep-ex/0604007`.

[38] CDF collaboration, T. Aaltonen *et al.*, *Measurements of the angular distributions in the decays $B \to K^{(*)}\mu^+\mu^-$ at CDF*, Phys. Rev. Lett. **108** (2012) 081807, `arXiv:1108.0695`.

[39] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^+\pi^-\mu^+\mu^-$ in the $K^*_{0,2}(1430)^0$ region*, JHEP **12** (2016) 065, `arXiv:1609.04736`.

[40] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^*e^+e^-$ decay in the low-q^2 region*, JHEP **04** (2015) 064, `arXiv:1501.03038`.

[41] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **125** (2020) 011802, `arXiv:2003.04831`.

[42] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **126** (2021) 161802, `arXiv:2012.13241`.

[43] ATLAS collaboration, M. Aaboud *et al.*, *Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector*, JHEP **04** (2019) 098, `arXiv:1812.03017`.

[44] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of properties of $B_s^0 \to \mu^+\mu^-$ decays and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment*, JHEP **04** (2020) 188, `arXiv:1910.12127`.

[45] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017) 191801, `arXiv:1703.05747`.

[46] LHCb collaboration, *Combination of the ATLAS, CMS and LHCb results on the $B^0_{(s)} \to \mu^+\mu^-$ decays*, LHCb-CONF-2020-002. ATLAS-CONF-2020-049, CMS PAS BPH-20-003, LHCb-CONF-2020-002.

[47] LHCb collaboration, R. Aaij *et al.*, *Analysis of neutral B-meson decays into two muons*, `arXiv:2108.09284`, submitted to PRL.

[48] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ decay properties and search for the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-\gamma$ decays*, `arXiv:2108.09283`, submitted to PRD.

[49] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, `arXiv:1705.05802`.

[50] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, Publisher's Note ibid. **115** (2015) 159901, `arXiv:1506.08614`.

[51] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ-lepton decays*, Phys. Rev. Lett. **120** (2018) 171802, `arXiv:1708.08856`.

[52] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavor universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, Phys. Rev. **D97** (2018) 072013, `arXiv:1711.02505`.

[53] Belle collaboration, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017), no. 21 211801, `arXiv:1612.00529`.

[54] Belle collaboration, Y. Sato *et al.*, *Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method*, Phys. Rev. **D94** (2016), no. 7 072007, `arXiv:1607.07923`.

[55] Belle collaboration, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. **D92** (2015), no. 7 072014, `arXiv:1507.03233`.

[56] BaBar collaboration, J. P. Lees *et al.*, *Measurement of an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays and implications for charged Higgs bosons*, Phys. Rev. **D88** (2013), no. 7 072012, `arXiv:1303.0571`.

[57] BaBar collaboration, J. P. Lees *et al.*, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802, `arXiv:1205.5442`.

[58] Belle collaboration, A. Abdesselam *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, `arXiv:1904.08794`.

[59] HFLAV collaboration, *Average of $R(D)$ and $R(D^*)$ for Spring 2019*, `https://hflav-eos.web.cern.ch/hflav-eos/semi/spring19/html/RDsDsstar/RDRDs.html`. Accessed: May 2021.

[60] BaBar Collaboration, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$* , Phys. Rev. **D86** (2012) 032012.

[61] Belle collaboration, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \to K^{(*)}l^+l^-$*, Phys. Rev. **D103** (2009) 12.

[62] Belle collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays*, JHEP **03** (2021) 105, `arXiv:1908.01848`.

[63] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, `arXiv:1406.6482`.

[64] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays*, JHEP **05** (2020) 040, `arXiv:1912.08139`.

[65] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[66] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[67] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of* $B \to K^*\mu^+\mu^-$ *at large recoil*, JHEP **01** (2013) 048, `arXiv:1207.2753`.

[68] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of* $B \to K^*\mu^+\mu^-$ *observables*, JHEP **12** (2014) 125, `arXiv:1407.8526`.

[69] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in* $B \to K^{(*)}\ell^+\ell^-$ *and* $B \to K^*\gamma$, JHEP **09** (2010) 089, `arXiv:1006.4945`.

[70] W. Altmannshofer and D. M. Straub, *New physics in* $b \to s$ *transitions after LHC run 1*, Eur. Phys. J. C **75** (2015) 382.

[71] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of* $b \to s\ell\ell$ *anomalies*, JHEP **06** (2016) 092, `arXiv:1510.04239`.

[72] M. Algueró *et al.*, *Emerging patterns of new physics with and without lepton flavour universal contributions*, Eur. Phys. J. **C79** (2019), no. 8 714, `arXiv:1903.09578`, [Addendum: Eur.Phys.J.C 80, 511 (2020)].

[73] M. Beneke, C. Bobeth, and R. Szafron, *Power-enhanced leading-logarithmic QED corrections to* $B_q \to \mu^+\mu^-$, JHEP **10** (2019) 232, `arXiv:1908.07011`.

[74] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for* R_K *and* R_{K^*}, Eur. Phys. J. **C76** (2016), no. 8 440, `arXiv:1605.07633`.

[75] B. Capdevila, S. Descotes-Genon, J. Matias, and J. Virto, *Assessing lepton-flavour non-universality from* $B \to K^*\ell\ell$ *angular analyses*, JHEP **10** (2016) 075, `arXiv:1605.03156`.

[76] D. van Dyk *et al.*, *eos/eos: Eos version 0.3.3*, Mar., 2021. doi: 10.5281/zenodo.4586379.

[77] A. Bharucha, D. M. Straub, and R. Zwicky, $B \to V\ell^+\ell^-$ *in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, `arXiv:1503.05534`.

[78] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *Status of the* $B \to K^*\mu^+\mu^-$ *anomaly after Moriond 2017*, Eur. Phys. J. **C77** (2017), no. 6 377, `arXiv:1703.09189`.

[79] D. M. Straub, *flavio: a python package for flavour and precision phenomenology in the Standard Model and beyond*, `arXiv:1810.08132arXiv:1810.08132`.

[80] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the* $B \to K^*\ell^+\ell^-$ *decays in the large-recoil region: SM challenges and BSM opportunities*, Phys. Rev. **D93** (2016), no. 1 014028, `arXiv:1412.3183`.

[81] G. Isidori, S. Nabeebaccus, and R. Zwicky, *QED corrections in* $\overline{B} \to \overline{K}\ell^+\ell^-$ *at the double-differential level*, JHEP **12** (2020) 104, `arXiv:2009.00929`.

[82] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of* $\bar{B} \to K\bar{\ell}\ell$ *decays*, JHEP **07** (2007), no. 12 040, `arXiv:0709.4174`.

[83] S. Aoki *et al.*, *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J. C **77** (2017), no. 2 112, `arXiv:1607.00299`.

[84] S. Fajfer, J. F. Kamenik, and I. Nisandzic, *On the* $B \to D^* \tau \bar{\nu}_\tau$ *Sensitivity to New Physics*, Phys. Rev. D **85** (2012) 094025, `arXiv:1203.2654`.

[85] M. Schumann, *Dark Matter 2014*, EPJ Web Conf. **96** (2015) 01027, `arXiv:1501.01200`.

[86] Planck collaboration, P. A. R. Ade *et al.*, *Planck 2013 results. XVI. Cosmological parameters*, Astron. Astrophys. **571** (2014) A16, `arXiv:1303.5076`.

[87] J. M. Cline, *Baryogenesis*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime*, 9, 2006. `arXiv:hep-ph/0609145`.

[88] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. D **51** (1995) 379, `arXiv:hep-ph/9404302`.

[89] ATLAS collaboration, G. Aad *et al.*, *Search for pair production of third-generation scalar leptoquarks decaying into a top quark and a* τ*-lepton in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, `arXiv:2101.11582`.

[90] CMS collaboration, A. M. Sirunyan *et al.*, *Search for leptoquarks coupled to third-generation quarks in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **121** (2018), no. 24 241802, `arXiv:1809.05558`.

[91] W. Altmannshofer and P. Stangl, *New Physics in Rare B Decays after Moriond 2021*, `arXiv:2103.13370`.

[92] M. Algueró *et al.*, $b \to s\ell\ell$ *global fits after Moriond 2021 results*, in *55th Rencontres de Moriond on QCD and High Energy Interactions*. `arXiv:2104.08921`.

[93] M. Ciuchini *et al.*, *Lessons from the* $B^{0,+} \to K^{*0,+}\mu^+\mu^-$ *angular analyses*, Phys. Rev. **D103** (2021), no. 1 015030, `arXiv:2011.01212`.

[94] J. Aebischer *et al.*, *B-decay discrepancies after Moriond 2019*, Eur. Phys. J. **C80** (2020), no. 3 252, `arXiv:1903.10434`.

[95] A. Arbey *et al.*, *Update on the* $b \to s$ *anomalies*, Phys. Rev. **D100** (2019), no. 1 015045, `arXiv:1904.08399`.

[96] M. Ciuchini *et al.*, *New Physics in* $b \to s\ell^+\ell^-$ *confronts new data on lepton universality*, Eur. Phys. J. **C79** (2019), no. 8 719, `arXiv:1903.09632`.

[97] K. Kowalska, D. Kumar, and E. M. Sessolo, *Implications for new physics in* $b \to s\mu\mu$ *transitions after recent measurements by Belle and LHCb*, Eur. Phys. J. **C79** (2019), no. 10 840, `arXiv:1903.10932`.

[98] C. Hati, J. Kriewald, J. Orloff, and A. M. Teixeira, *The fate of vector leptoquarks: the impact of future flavour data*, `arXiv:2012.05883`.

[99] J. Aebischer, J. Kumar, P. Stangl, and D. M. Straub, *A Global Likelihood for Precision Constraints and Flavour Anomalies*, Eur. Phys. J. C **79** (2019), no. 6 509, `arXiv:1810.07698`.

[100] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. **D89** (2014) 095033, `arXiv:1403.1269`.

[101] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, Phys. Rev. Lett. **114** (2015) 151801, `arXiv:1501.00993`.

[102] A. Celis, J. Fuentes-Martín, M. Jung, and H. Serôdio, *Family nonuniversal Z' models with protected flavor-changing interactions*, Phys. Rev. **D92** (2015), no. 1 015007, `arXiv:1505.03079`.

[103] A. Falkowski, M. Nardecchia, and R. Ziegler, *Lepton flavor non-universality in B-meson decays from a U(2) flavor model*, JHEP **11** (2015) 173, `arXiv:1509.01249`.

[104] B. C. Allanach, J. M. Butterworth, and T. Corbett, *Collider constraints on Z' models for neutral current B-anomalies*, JHEP **08** (2019) 106, `arXiv:1904.10954`.

[105] B. C. Allanach and J. Davighi, *Naturalising the third family hypercharge model for neutral current B-anomalies*, Eur. Phys. J. **C79** (2019), no. 11 908, `arXiv:1905.10327`.

[106] J. Kawamura, S. Raby, and A. Trautner, *Complete vectorlike fourth family and new U(1)' for muon anomalies*, Phys. Rev. **D100** (2019), no. 5 055030, `arXiv:1906.11297`.

[107] S. Dwivedi, D. Kumar Ghosh, A. Falkowski, and N. Ghosh, *Associated Z' production in the flavorful $U(1)$ scenario for $R_{K^{(*)}}$*, Eur. Phys. J. **C80** (2020), no. 3 263, `arXiv:1908.03031`.

[108] Z.-L. Han, R. Ding, S.-J. Lin, and B. Zhu, *Gauged $U(1)_{L_\mu - L_\tau}$ scotogenic model in light of $R_{K^{(*)}}$ anomaly and AMS-02 positron excess*, Eur. Phys. J. **C79** (2019), no. 12 1007, `arXiv:1908.07192`.

[109] B. Capdevila, A. Crivellin, C. A. Manzari, and M. Montull, *Explaining $b \to s\ell^+\ell^-$ and the Cabibbo angle anomaly with a vector triplet*, Phys. Rev. **D103** (2021), no. 1 015032, `arXiv:2005.13542`.

[110] W. Altmannshofer, J. Davighi, and M. Nardecchia, *Gauging the accidental symmetries of the Standard Model, and implications for the flavor anomalies*, Phys. Rev. **D101** (2020), no. 1 015004, `arXiv:1909.02021`.

[111] S.-L. Chen *et al.*, *Signatures of a flavor changing Z' boson in $B_q \to \gamma Z'$*, Nucl. Phys. **B962** (2021) 115237, `arXiv:2006.03383`.

[112] A. Carvunis, D. Guadagnoli, M. Reboud, and P. Stangl, *Composite dark matter and a horizontal symmetry*, JHEP **02** (2021) 056, `arXiv:2007.11931`.

[113] A. Karozas, G. K. Leontaris, I. Tavellaris, and N. D. Vlachos, *On the LHC signatures of $SU(5) \times U(1)'$ F-theory motivated models*, Eur. Phys. J. **C81** (2021), no. 1 35, `arXiv:2007.05936`.

[114] D. Borah, L. Mukherjee, and S. Nandi, *Low scale $U(1)_X$ gauge symmetry as an origin of dark matter, neutrino mass and flavour anomalies*, JHEP **12** (2020) 052, `arXiv:2007.13778`.

[115] B. C. Allanach, *$U(1)_{B_3 - L_2}$ explanation of the neutral current B−anomalies*, Eur. Phys. J. **C81** (2021), no. 1 56, `arXiv:2009.02197`.

[116] J.-H. Sheng, *The analysis of $b \to s\ell^+\ell^-$ in the family non-universal Z' Model*, Int. J. Theor. Phys. **60** (2021), no. 1 26.

[117] G. Hiller and M. Schmaltz, *R_K and future $b \to s\ell\ell$ physics beyond the Standard Model opportunities*, Phys. Rev. **D90** (2014) 054014, `arXiv:1408.1627`.

[118] B. Gripaios, M. Nardecchia, and S. A. Renner, *Composite leptoquarks and anomalies in B-meson decays*, JHEP **05** (2015) 006, `arXiv:1412.1791`.

[119] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, `arXiv:1503.01084`.

[120] R. Barbieri, C. W. Murphy, and F. Senia, *B-decay anomalies in a composite leptoquark model*, Eur. Phys. J. **C77** (2017), no. 1 8, `arXiv:1611.04930`.

[121] M. Bordone, C. Cornella, J. Fuentes-Martín, and G. Isidori, *Low-energy signatures of the PS3 model: from B-physics anomalies to LFV*, JHEP **10** (2018) 148, `arXiv:1805.09328`.

[122] B. Fornal, S. A. Gadam, and B. Grinstein, *Left-right SU(4) vector leptoquark model for flavor anomalies*, Phys. Rev. **D99** (2019), no. 5 055025, `arXiv:1812.01603`.

[123] S. Balaji and M. A. Schmidt, *Unified SU(4) theory for the $R_{D^{(*)}}$ and $R_{K^{(*)}}$ anomalies*, Phys. Rev. **D101** (2020), no. 1 015026, `arXiv:1911.08873`.

[124] C. Cornella, J. Fuentes-Martin, and G. Isidori, *Revisiting the vector leptoquark explanation of the B-physics anomalies*, JHEP **07** (2019) 168, `arXiv:1903.11517`.

[125] A. Datta, D. Sachdeva, and J. Waite, *Unified explanation of $b \to s\mu^+\mu^-$ anomalies, neutrino masses, and $B \to \pi K$ puzzle*, Phys. Rev. **D100** (2019), no. 5 055015, `arXiv:1905.04046`.

[126] O. Popov, M. A. Schmidt, and G. White, *R_2 as a single leptoquark solution to $R_{D^{(*)}}$ and $R_{K^{(*)}}$*, Phys. Rev. **D100** (2019), no. 3 035028, `arXiv:1905.06339`.

[127] I. Bigaran, J. Gargalionis, and R. R. Volkas, *A near-minimal leptoquark model for reconciling flavour anomalies and generating radiative neutrino masses*, JHEP **10** (2019) 106, `arXiv:1906.01870`.

[128] J. Bernigaud, I. de Medeiros Varzielas, and J. Talbert, *Finite family groups for fermionic and leptoquark mixing patterns*, JHEP **01** (2020) 194, `arXiv:1906.11270`.

[129] L. Da Rold and F. Lamagna, *A vector leptoquark for the B-physics anomalies from a composite GUT*, JHEP **12** (2019) 112, `arXiv:1906.11666`.

[130] J. Fuentes-Martín, M. Reig, and A. Vicente, *Strong CP problem with low-energy emergent QCD: The 4321 case*, Phys. Rev. **D100** (2019), no. 11 115028, `arXiv:1907.02550`.

[131] C. Hati, J. Kriewald, J. Orloff, and A. M. Teixeira, *A non-unitary interpretation for a single vector leptoquark combined explanation to the B-decay anomalies*, JHEP **12** (2019) 006, `arXiv:1907.05511`.

[132] A. Datta, J. L. Feng, S. Kamali, and J. Kumar, *Resolving the $(g-2)_\mu$ and B anomalies with leptoquarks and a dark Higgs boson*, Phys. Rev. **D101** (2020), no. 3 035010, `arXiv:1908.08625`.

[133] A. Crivellin, D. Müller, and F. Saturnino, *Flavor phenomenology of the leptoquark singlet-triplet model*, JHEP **06** (2020) 020, `arXiv:1912.04224`.

[134] C. Borschensky, B. Fuks, A. Kulesza, and D. Schwartländer, *Scalar leptoquark pair production at hadron colliders*, Phys. Rev. **D101** (2020), no. 11 115017, `arXiv:2002.08971`.

[135] S. Saad, *Combined explanations of* $(g-2)_\mu$, $R_{D^{(*)}}$, $R_{K^{(*)}}$ *anomalies in a two-loop radiative neutrino mass model*, Phys. Rev. **D102** (2020), no. 1 015019, `arXiv:2005.04352`.

[136] J. Fuentes-Martín and P. Stangl, *Third-family quark-lepton unification with a fundamental composite Higgs*, Phys. Lett. **B811** (2020) 135953, `arXiv:2004.11376`.

[137] P. S. Bhupal Dev, R. Mohanta, S. Patra, and S. Sahoo, *Unified explanation of flavor anomalies, radiative neutrino masses, and ANITA anomalous events in a vector leptoquark model*, Phys. Rev. **D102** (2020), no. 9 095012, `arXiv:2004.09464`.

[138] B. Fornal, *Gravitational wave signatures of lepton universality violation*, Phys. Rev. **D103** (2021), no. 1 015018, `arXiv:2006.08802`.

[139] J. Davighi, M. Kirk, and M. Nardecchia, *Anomalies and accidental symmetries: charging the scalar leptoquark under* $L_\mu - L_\tau$, JHEP **12** (2020) 111, `arXiv:2007.15016`.

[140] B. Dutta, S. Ghosh, P. Huang, and J. Kumar, *Explaining* $g_\mu - 2$ *and* $R_{K^{(*)}}$ *using the light mediators of* $U(1)_{T3R}$, `arXiv:2105.07655`.

[141] J. Davighi, *Anomalous* Z' *bosons for anomalous B decays*, `arXiv:2105.06918`.

[142] J. S. Alvarado, S. F. Mantilla, R. Martinez, and F. Ochoa, *A non-universal* $U(1)_X$ *extension to the Standard Model to study the B meson anomaly and muon* $g-2$, `arXiv:2105.04715`.

[143] W. Altmannshofer, S. A. Gadam, S. Gori, and N. Hamer, *Explaining* $(g-2)_\mu$ *with Multi-TeV Sleptons*, `arXiv:2104.08293`.

[144] D. Marzocca and S. Trifinopoulos, *A Minimal Explanation of Flavour Anomalies: B-Meson Decays, Muon Magnetic Moment, and the Cabbibo Angle*, `arXiv:2104.05730`.

[145] G. Arcadi, L. Calibbi, M. Fedele, and F. Mescia, *Muon* $g-2$ *and B-anomalies from Dark Matter*, `arXiv:2104.03228`.

[146] A. Greljo, P. Stangl, and A. E. Thomsen, *A Model of Muon Anomalies*, `arXiv:2103.13991`.

[147] Muon g-2 collaboration, B. Abi *et al.*, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021), no. 14 141801, `arXiv:2104.03281`.

[148] Belle collaboration, A. Abdesselam *et al.*, *Test of lepton flavor universality in* $B \to K^*\ell^+\ell^-$ *decays at Belle*, `arXiv:1904.02440`.

[149] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004. doi: 10.5170/CERN-2004-003-V-1.

[150] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[151] CMS collaboration, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[152] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[153] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, `arXiv:1612.05140`.

[154] LHCb collaboration, C. Elsässer, *$b\bar{b}$ production angle plots*, `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html`. Accessed: 2021-05-28.

[155] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[156] A. Martin Sanchez, P. Robbe, and M.-H. Schune, *Performances of the LHCb L0 Calorimeter Trigger*, Tech. Rep. LHCb-PUB-2011-026. CERN-LHCb-PUB-2011-026, CERN, Geneva, Jun, 2012. `http://cds.cern.ch/record/1407893`.

[157] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010.

[158] A. A. Alves, Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[159] G. Bencivenni *et al.*, *A triple-GEM detector with pad readout for the inner region of the first LHCb muon station*, Tech. Rep. LHCb-2001-051, CERN, Geneva, Jul, 2001. `https://cds.cern.ch/record/691604`.

[160] A. Kashchuk *et al.*, *Asymmetric and double-cathode-pad wire chambers for the LHCb muon system*, Nucl. Instrum. Meth. **A555** (2005) 48.

[161] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024.

[162] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, `arXiv:1708.00819`.

[163] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[164] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[165] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007.

[166] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[167] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[168] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037.

[169] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[170] T. Head, *The LHCb trigger system*, JINST **9** (2014) C09015.

[171] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[172] C. Bozzi, *LHCb Computing Resource usage in 2020*, tech. rep., CERN, Geneva, Feb, 2021.

[173] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[174] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[175] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[176] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[177] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, `arXiv:hep-ph/0506026`.

[178] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[179] M. van Veghel, *Pursuing forbidden beauty: Search for the lepton-flavour violating decays* $B^0 \to e^{\pm}\mu^{\mp}$ *and* $B^0_s \to e^{\pm}\mu^{\mp}$ *and study of electron-reconstruction performance at LHCb*, PhD thesis, University of Groningen, 2020, doi: 10.33612/diss.128123609.

[180] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, `arXiv:physics/0503191`.

[181] M. Williams *et al.*, *The HLT2 Topological Lines*, tech. rep., CERN, Geneva, Jan, 2011.

[182] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[183] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: Bounds for k-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), p. 203–208, Association for Computing Machinery, 1999. doi: 10.1145/307400.307439.

[184] R. Barlow, *Extended maximum likelihood*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **297** (1990), no. 3 496 .

[185] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, `arXiv:physics/0306116`.

[186] F. James, *MINUIT: Function Minimization and Error Analysis Reference Manual*, 1998. CERN Program Library Long Writeups, `https://cds.cern.ch/record/2296388`.

[187] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **764** (2014) 150 .

[188] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, PhD thesis, Cracow, INP, 1986.

[189] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[190] K. Cranmer, *Kernel estimation in high-energy physics*, Computer Physics Communications **136** (2001) 198–207.

[191] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, EPJ Tech. Instrum. **6** (2019), no. 1 1, `arXiv:1803.00824`.

[192] C. Burr *et al.*, *Prompt charm production in pp collisions at* $\sqrt{s} = 13\,\text{TeV}$ *in 2015 data*, tech. rep., Apr, 2015. `https://cds.cern.ch/record/2012800`.

[193] B. Efron, *Bootstrap methods: Another look at the jackknife*, The Annals of Statistics **7** (1979), no. 1 1.

[194] LHCb collaboration, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, JINST **14** (2019) P11023, `arXiv:1909.02957`.

[195] BaBar collaboration, B. Aubert *et al.*, *Dalitz-plot analysis of the decays* $B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$, Phys. Rev. D **72** (2005) 072003, `arXiv:hep-ex/0507004`, [Erratum: Phys.Rev.D 74, 099903 (2006)].

[196] J. Back *et al.*, *Laura ++: A dalitz plot fitter*, Computer Physics Communications **231** (2018) 198.

[197] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Stat. **9** (1938) 60.

[198] J. A. Bailey *et al.*, $B \to Kl^{+}l^{-}$ *decay form factors from three-flavor lattice QCD*, Phys. Rev. **D93** (2016) 025026, `arXiv:1509.06235`.

[199] D. Du *et al.*, *Phenomenology of semileptonic B-meson decays with form factors from lattice QCD*, Phys. Rev. **D93** (2016), no. 3 034005, `arXiv:1510.02349`.

[200] Belle-II collaboration, W. Altmannshofer *et al.*, *The Belle II physics book*, PTEP **2019** (2019), no. 12 123C01, Erratum ibid. **2020** (2020) 029201, `arXiv:1808.10567`.

[201] Belle-II collaboration, T. Abe *et al.*, *Belle II Technical Design Report*, `arXiv:1011.0352`.

[202] CMS collaboration, S. Mukherjee, *Data Scouting and Data Parking with the CMS High level Trigger*, PoS **EPS-HEP2019** (2020) 139.

[203] LHCb collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016.

[204] LHCb collaboration, *Physics case for an LHCb Upgrade II — Opportunities in flavour physics, and beyond, in the HL-LHC era*, `arXiv:1808.08865`.

Part III

Appendices

Appendix A

Additional information on the fits to resonant data

This appendix is divided into three sections. Appendix A.1 shows fits to simulated $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ samples, which are used to constrain signal and background shapes in data. Appendix A.2 contains fits to $B^+ \to K^+ J/\psi(e^+e^-)$ hTOS! and TIS! data. Appendix A.3 lists the fit parameters that maximise the likelihood in each fit to fully-selected data.

A.1 Fits to simulated control-mode samples

Fits to simulated $B^+ \to K^+ J/\psi(\ell^+\ell^-)$ 2018 samples are shown in Figures A.1 and A.2. The results from other data-taking conditions are similar, and are thus omitted. The pulls indicate that the models suitably describe the signal shapes.

A.2 Fits to hTOS! and TIS! data

The fits to hTOS! and TIS! data are presented in Figure A.3. Shown below each plot are the pulls between the data and the total fit model. The pulls are generally small, indicating good agreement between the data and the model.

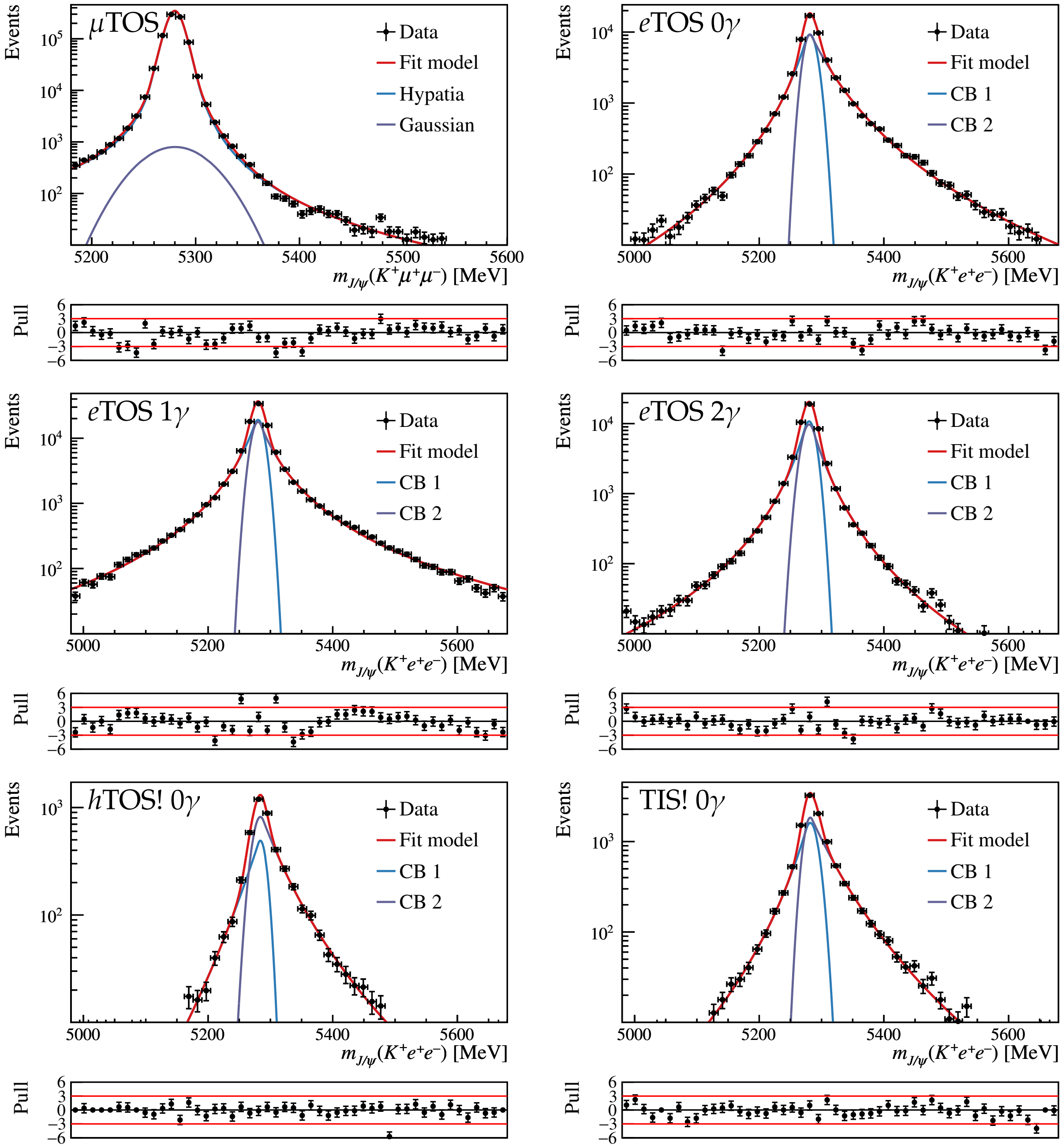

Figure A.1: Fits to the $m_{J/\psi}(K^+\ell^+\ell^-)$ distribution of simulated $B^+ \to K^+J/\psi(\ell^+\ell^-)$ events that pass the entire selection chain. The pulls between the simulation samples (black) and the models (red) are shown at the bottom of each plot. The components of the signal shape, *viz.* a Hypatia and a Gaussian for muons and two Crystal Ball distributions for electrons, are shown in blue and purple. Shown below each plot are the pulls between the data and the total fit model.

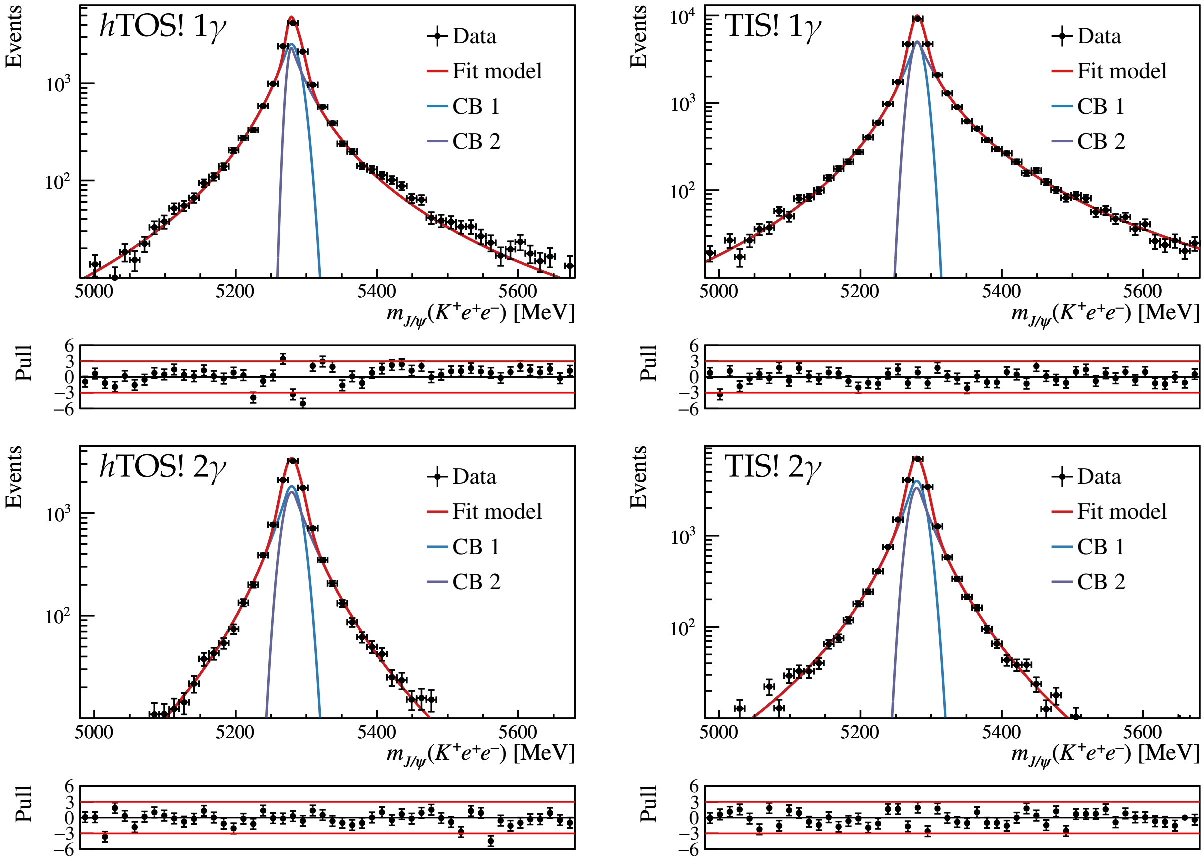

Figure A.2: Continued from the previous page.

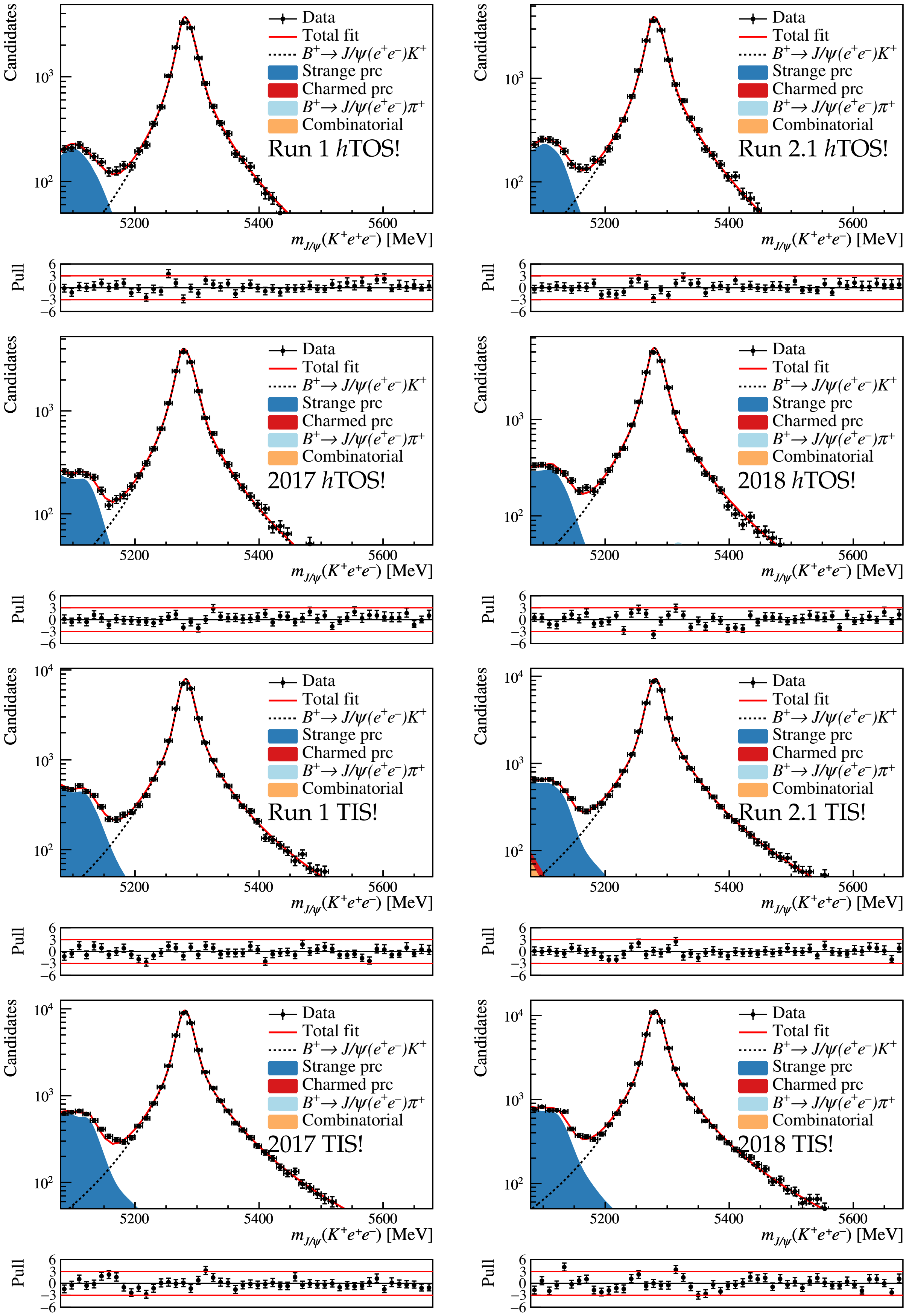

Figure A.3: Fits to the $m_{J/\psi}(K^+e^+e^-)$ distribution of $B^+ \to K^+J/\psi(e^+e^-)$ fully-selected hTOS! and TIS! candidates in the samples corresponding to each data-taking period. The red solid line represents the fit model, the dotted black line is the signal component, the light-blue filled area represents misidentified $B^+ \to \pi^+J/\psi(e^+e^-)$ events, and the orange-filled area shows the combinatorial background. The strange and charm partially-reconstructed backgrounds (referred to as "prc" in the legend) are represented by the filled areas filled with dark blue and red, respectively. Shown below each plot are the pulls between the data and the total fit model.

A.3 Best-fit parameters

Table A.1 lists the parameters that maximise the likelihood in the fits to fully-selected $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data from each data-taking period. Their equivalents from the electron data, from each of the three trigger categories, are given in Tables A.2 to A.4.

Table A.1: Results of the final fits to $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ data taken during the four data-taking periods. The notation is explained in Section 5.1.

	Run 1		**Run 2.1**
$N_{\text{sig}} =$	618300 ± 800	$N_{\text{sig}} =$	543500 ± 700
$N_{\text{combi}} =$	2280 ± 140	$N_{\text{combi}} =$	2970 ± 140
$N_{\text{mis-ID}}/N_{\text{sig}} =$	0.00288 ± 0.00011	$N_{\text{mis-ID}}/N_{\text{sig}} =$	0.00473 ± 0.00016
$\lambda =$	0.0057 ± 0.0004	$\lambda =$	0.00565 ± 0.00030
$\Delta\mu =$	1.219 ± 0.013	$\Delta\mu =$	-0.215 ± 0.014
$s_\sigma =$	1.1600 ± 0.0014	$s_\sigma =$	1.1531 ± 0.0016
	2017		**2018**
$N_{\text{sig}} =$	507100 ± 700	$N_{\text{sig}} =$	619600 ± 800
$N_{\text{combi}} =$	3170 ± 130	$N_{\text{combi}} =$	3780 ± 150
$N_{\text{mis-ID}}/N_{\text{sig}} =$	0.00523 ± 0.00017	$N_{\text{mis-ID}}/N_{\text{sig}} =$	0.00562 ± 0.00018
$\lambda =$	0.00623 ± 0.00027	$\lambda =$	0.00603 ± 0.00025
$\Delta\mu =$	-0.275 ± 0.014	$\Delta\mu =$	-0.078 ± 0.013
$s_\sigma =$	1.1239 ± 0.0016	$s_\sigma =$	1.1231 ± 0.0014

Table A.2: Results of the final fits to $B^+ \to K^+ J/\psi(e^+e^-)$ data taken during the four data-taking periods, using the *e*TOS trigger strategy. The notation is explained in Sections 5.1 and 5.4.

	Run 1		**Run 2.1**
$N_{\text{sig}} =$	90220 ± 310	$N_{\text{sig}} =$	153400 ± 400
$N_{\text{prc}} =$	7680 ± 280	$N_{\text{prc}} =$	13500 ± 400
$N_{\text{charm}}/N_{\text{strange}} =$	0.027 ± 0.004	$N_{\text{charm}}/N_{\text{strange}} =$	0.028 ± 0.004
$N_{\text{mis-ID}}/N_{\text{K}} =$	0.0047 ± 0.0005	$N_{\text{mis-ID}}/N_{\text{K}} =$	0.0052 ± 0.0005
$\Delta_\mu =$	1.32 ± 0.06	$\Delta_\mu =$	-0.20 ± 0.05
$s_\sigma =$	1.112 ± 0.006	$s_\sigma =$	1.124 ± 0.005
$f_{0\gamma} =$	0.2397 ± 0.0024	$f_{0\gamma} =$	0.2507 ± 0.0025
$f_{1\gamma} =$	0.492 ± 0.005	$f_{1\gamma} =$	0.497 ± 0.005
$N_{\text{combi}} =$	470 ± 250	$N_{\text{combi}} =$	300 ± 400
$\lambda =$	0.038 ± 0.011	$\lambda =$	0.036 ± 0.022
	2017		**2018**
$N_{\text{sig}} =$	126300 ± 400	$N_{\text{sig}} =$	148900 ± 400
$N_{\text{prc}} =$	11640 ± 120	$N_{\text{prc}} =$	13670 ± 140
$N_{\text{charm}}/N_{\text{strange}} =$	0.029 ± 0.004	$N_{\text{charm}}/N_{\text{strange}} =$	0.037 ± 0.006
$N_{\text{mis-ID}}/N_{\text{K}} =$	0.0059 ± 0.0006	$N_{\text{mis-ID}}/N_{\text{K}} =$	0.0062 ± 0.0006
$\Delta_\mu =$	-0.11 ± 0.05	$\Delta_\mu =$	0.02 ± 0.05
$s_\sigma =$	1.105 ± 0.005	$s_\sigma =$	1.108 ± 0.005
$f_{0\gamma} =$	0.2511 ± 0.0025	$f_{0\gamma} =$	0.2541 ± 0.0025
$f_{1\gamma} =$	0.502 ± 0.005	$f_{1\gamma} =$	0.500 ± 0.005

Table A.3: Results of the final fits to $B^+ \to K^+ J/\psi(e^+e^-)$ data taken during the four data-taking periods, using the hTOS! trigger strategy. The notation is explained in Sections 5.1 and 5.4.

Run 1

$N_{\mathrm{sig}} = 15390 \pm 130$
$N_{\mathrm{prc}} = 1260 \pm 50$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.035 \pm 0.005$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.0119 \pm 0.0012$
$\Delta_{\mu} = 1.03 \pm 0.19$
$s_{\sigma} = 1.081 \pm 0.014$
$f_{0\gamma} = 0.1569 \pm 0.0012$
$f_{1\gamma} = 0.500 \pm 0.005$
$N_{\mathrm{combi}} = -13 \pm 20$
$\lambda = 0.078 \pm 0.099$

Run 2.1

$N_{\mathrm{sig}} = 16960 \pm 130$
$N_{\mathrm{prc}} = 1360 \pm 40$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.032 \pm 0.003$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.0133 \pm 0.0009$
$\Delta_{\mu} = -0.30 \pm 0.13$
$s_{\sigma} = 1.140 \pm 0.014$
$f_{0\gamma} = 0.1341 \pm 0.0013$
$f_{1\gamma} = 0.501 \pm 0.005$
$N_{\mathrm{combi}} = -10 \pm 260$
$\lambda = 0.00 \pm 0.08$

2017

$N_{\mathrm{sig}} = 17470 \pm 140$
$N_{\mathrm{prc}} = 1380 \pm 60$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.032 \pm 0.005$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.0138 \pm 0.0013$
$\Delta_{\mu} = -0.67 \pm 0.19$
$s_{\sigma} = 1.152 \pm 0.014$
$f_{0\gamma} = 0.1424 \pm 0.0014$
$f_{1\gamma} = 0.496 \pm 0.005$

2018

$N_{\mathrm{sig}} = 22770 \pm 160$
$N_{\mathrm{prc}} = 1820 \pm 50$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.033 \pm 0.005$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.0140 \pm 0.0014$
$\Delta_{\mu} = -0.28 \pm 0.16$
$s_{\sigma} = 1.118 \pm 0.012$
$f_{0\gamma} = 0.1490 \pm 0.0014$
$f_{1\gamma} = 0.500 \pm 0.005$

Table A.4: Results of the final fits to $B^+ \to K^+ J/\psi(e^+e^-)$ data taken during the four data-taking periods, using the TIS! trigger strategy. The notation is explained in Sections 5.1 and 5.4.

Run 1

$N_{\mathrm{sig}} = 30360 \pm 180$
$N_{\mathrm{prc}} = 2820 \pm 60$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.041 \pm 0.006$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.00312 \pm 0.00031$
$\Delta_{\mu} = 1.36 \pm 0.12$
$s_{\sigma} = 1.0728 \pm 0.0098$
$f_{0\gamma} = 0.1920 \pm 0.0019$
$f_{1\gamma} = 0.498 \pm 0.005$
$N_{\mathrm{combi}} = -17 \pm 18$
$\lambda = 0.007 \pm 0.009$

Run 2.1

$N_{\mathrm{sig}} = 37810 \pm 200$
$N_{\mathrm{prc}} = 3490 \pm 260$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.042 \pm 0.006$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.00299 \pm 0.00030$
$\Delta_{\mu} = -0.03 \pm 0.11$
$s_{\sigma} = 1.127 \pm 0.009$
$f_{0\gamma} = 0.1697 \pm 0.0017$
$f_{1\gamma} = 0.499 \pm 0.005$
$N_{\mathrm{combi}} = 210 \pm 250$
$\lambda = 0.030 \pm 0.017$

2017

$N_{\mathrm{sig}} = 37890 \pm 200$
$N_{\mathrm{prc}} = 3790 \pm 80$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.044 \pm 0.007$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.00305 \pm 0.00031$
$\Delta_{\mu} = -0.19 \pm 0.11$
$s_{\sigma} = 1.101 \pm 0.009$
$f_{0\gamma} = 0.1722 \pm 0.0017$
$f_{1\gamma} = 0.502 \pm 0.005$
$N_{\mathrm{combi}} = -14 \pm 22$
$\lambda = 0.08 \pm 0.17$

2018

$N_{\mathrm{sig}} = 45860 \pm 230$
$N_{\mathrm{prc}} = 4590 \pm 80$
$N_{\mathrm{charm}}/N_{\mathrm{strange}} = 0.039 \pm 0.006$
$N_{\mathrm{mis-ID}}/N_{\mathrm{K}} = 0.00329 \pm 0.00033$
$\Delta_{\mu} = 0.03 \pm 0.10$
$s_{\sigma} = 1.115 \pm 0.009$
$f_{0\gamma} = 0.1737 \pm 0.0017$
$f_{1\gamma} = 0.496 \pm 0.005$
$N_{\mathrm{combi}} = 30 \pm 80$
$\lambda = 0.004 \pm 0.015$

Appendix B

Additional trigger calibration tables

This appendix contains additional material used to calibrate the performance of the `L0` trigger lines employed in the R_K analysis. The calibration procedure is described in Section 6.3.

B.1 Calibration of `L0Electron` in the other ECAL regions

The main body of the thesis presents in Figure 6.6 the `L0Electron` performance of electrons passing through the inner-most region of the ECAL. The corresponding tables for electrons which traverse the outer and middle regions are listed in Figure B.1 and Figure B.2, respectively.

B.2 Calibration of `L0Hadron` in the other HCAL regions

The main body of the thesis presents in Figure 6.8 the `L0Hadron` performance of electrons passing through the inner region of the HCAL. The corresponding tables for kaons which traverse the outer region are listed in Figure B.3.

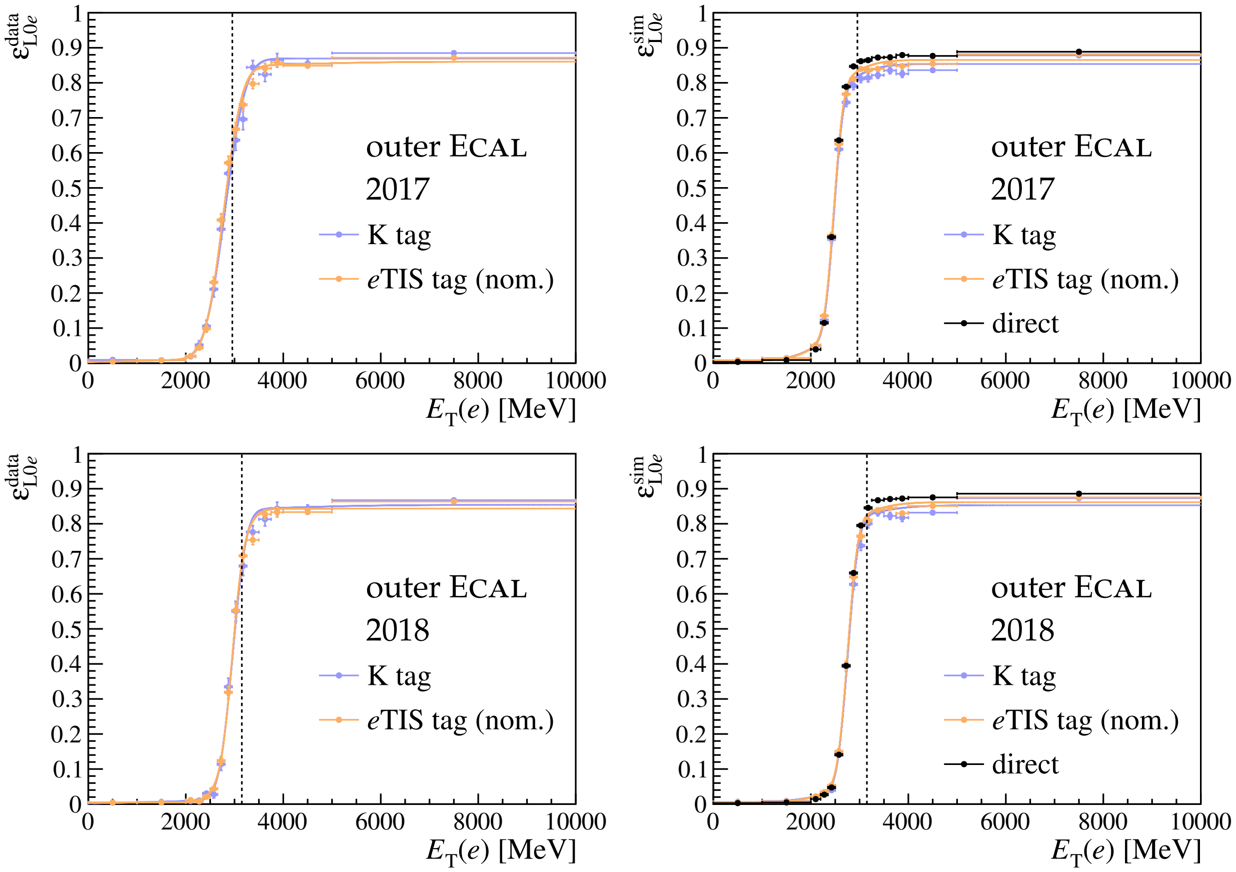

Figure B.1: Efficiency with which an electron traversing the outer ECAL region in data (left) and simulation (right) fires the `L0Electron` trigger, as a function of its reconstructed transverse momentum. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.13). The tag labelled as "nom." is the one used to estimate the nominal efficiencies, the other tag being used to evaluate systematic uncertainties. The dashed line corresponds to the fiducial cut applied to all electrons that fire `L0Electron`, in order to improve the level of agreement between selected data and simulated events.

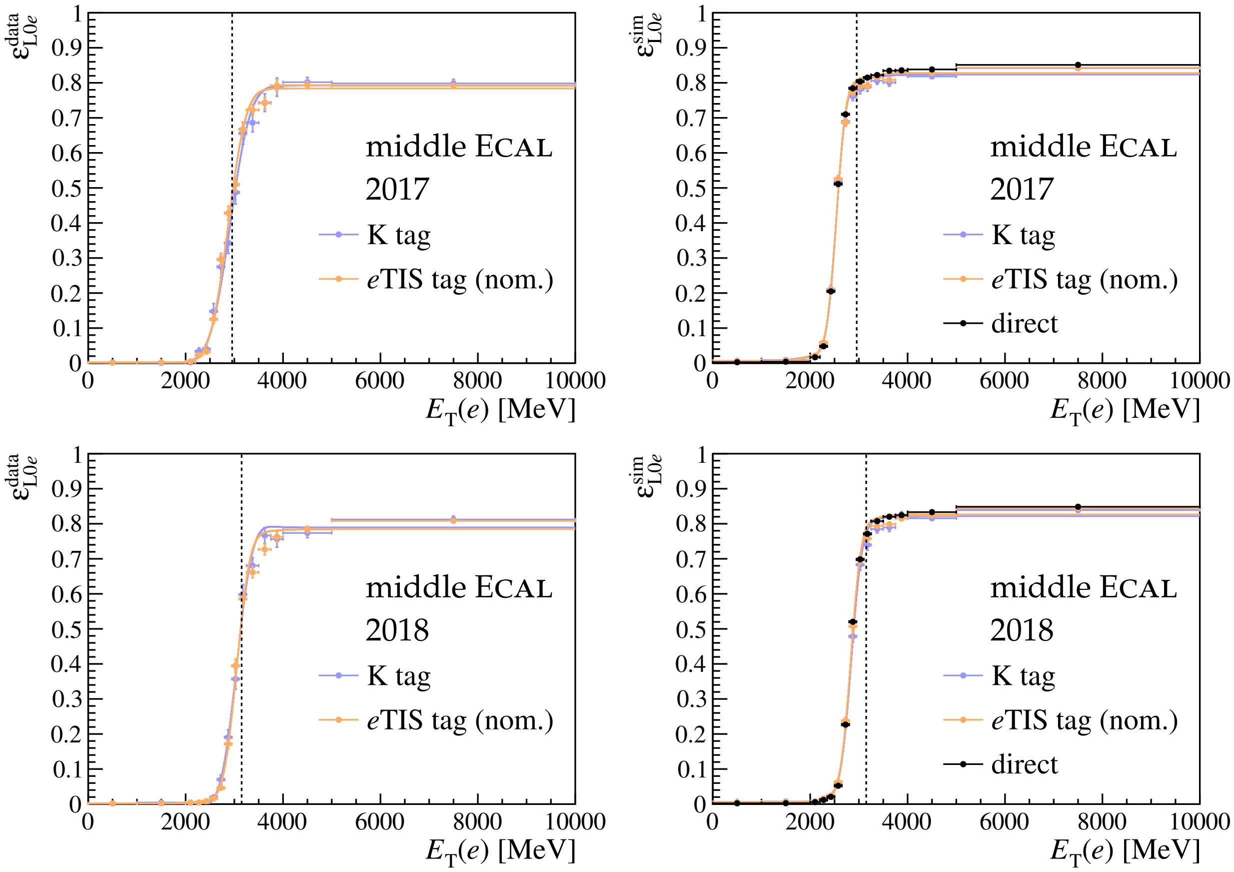

Figure B.2: Efficiency with which an electron traversing the middle ECAL region in data (left) and simulation (right) fires the `L0Electron` trigger, as a function of its reconstructed transverse momentum. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.13). The tag labelled as "nom." is the one used to estimate the nominal efficiencies, the other tag being used to evaluate systematic uncertainties. The dashed line corresponds to the fiducial cut applied to all electrons that fire `L0Electron`, in order to improve the level of agreement between selected data and simulated events.

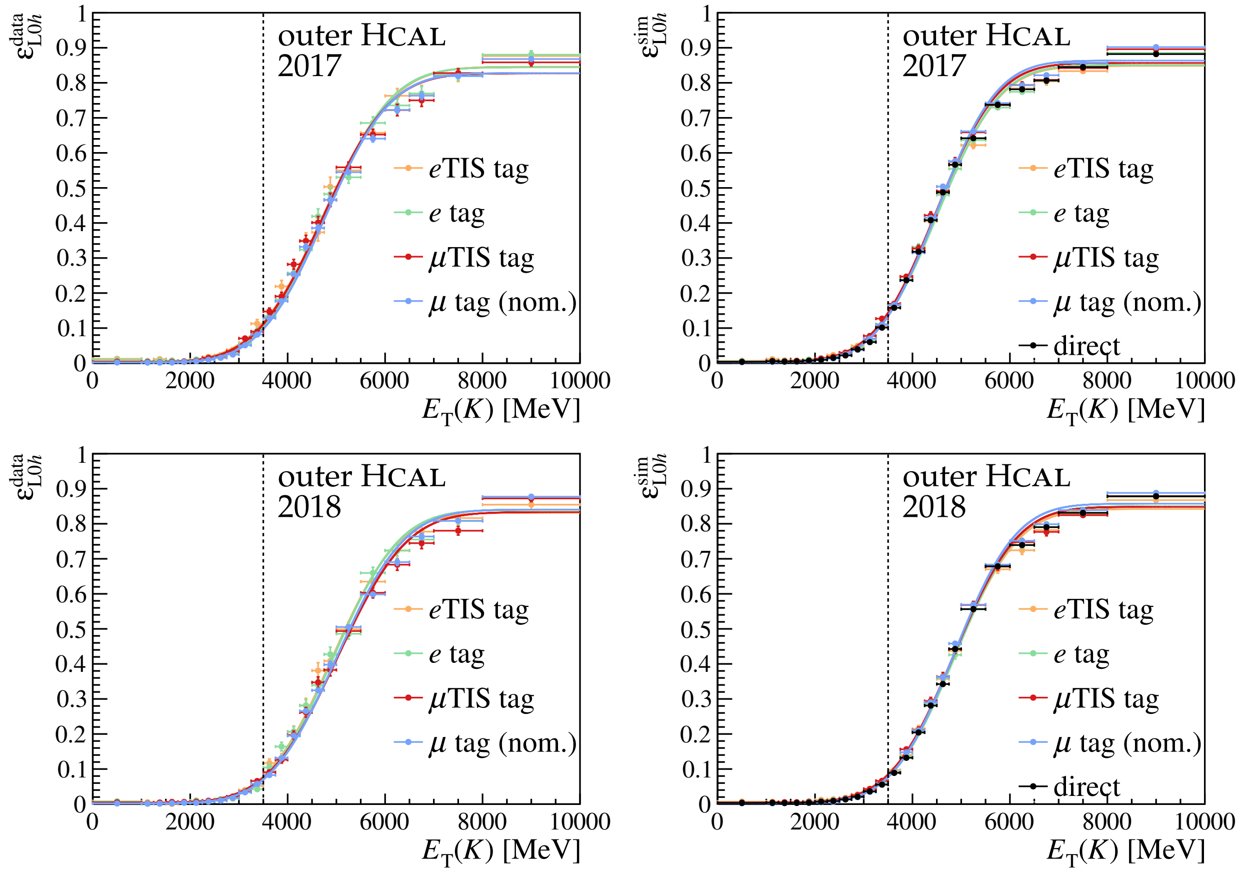

Figure B.3: Efficiency with which a kaon traversing the outer HCAL region in data (left) and simulation (right) fires the `L0Hadron` trigger, as a function of the reconstructed transverse energy deposited in the HCAL. The bins represent the tags described in the main body, whilst the lines are the corresponding fits to the function defined in Equation (6.15). The tag labelled as “nom.” is the one used to estimate the nominal efficiencies, the other tag being used to evaluate systematic uncertainties. The dashed line corresponds to the $E_{\text{T}} > 3.5$ GeV fiducial cut applied to all kaons that fire `L0Hadron`, in order to improve the level of agreement between selected data and simulated events.

Appendix C

Supplementary kinematic distributions

The main body of the thesis discusses in Section 6.5 the calibration of the simulated B^+ kinematics. Its effect on the distributions of several variables in muon and electron samples in 2018 data-taking conditions is exemplified in Figures 6.11 and 6.12, respectively. This appendix presents in Figures C.1 and C.2 the distributions of additional variables in the 2018 muon and electron samples, respectively. The equivalent spectra obtained from 2017 samples are shown in Figures C.3 to C.6.

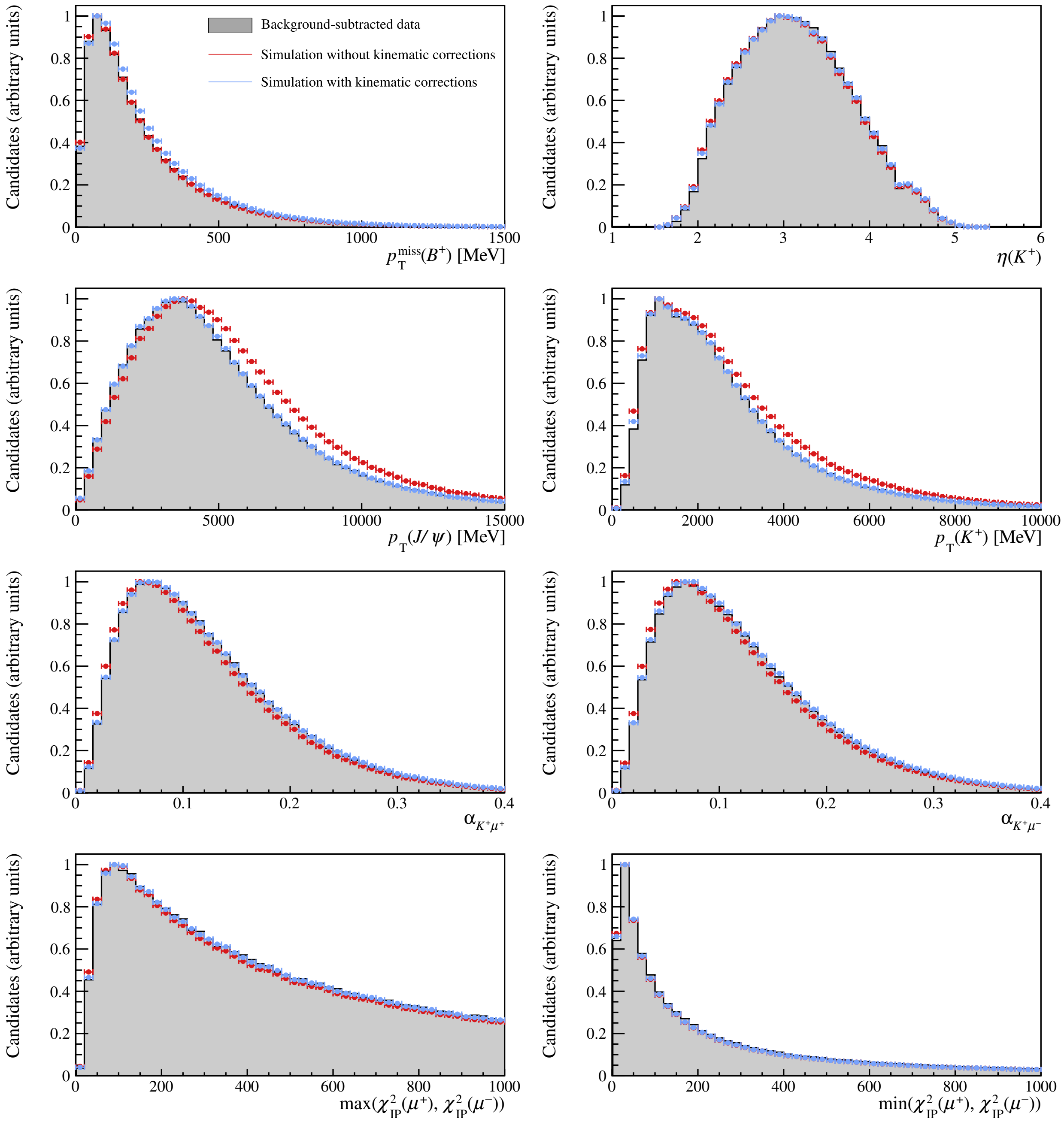

Figure C.1: Supplementary distributions of kinematic variables in $B^{+} \to K^{+}J/\psi(\mu^{+}\mu^{-})$ 2018 μTOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

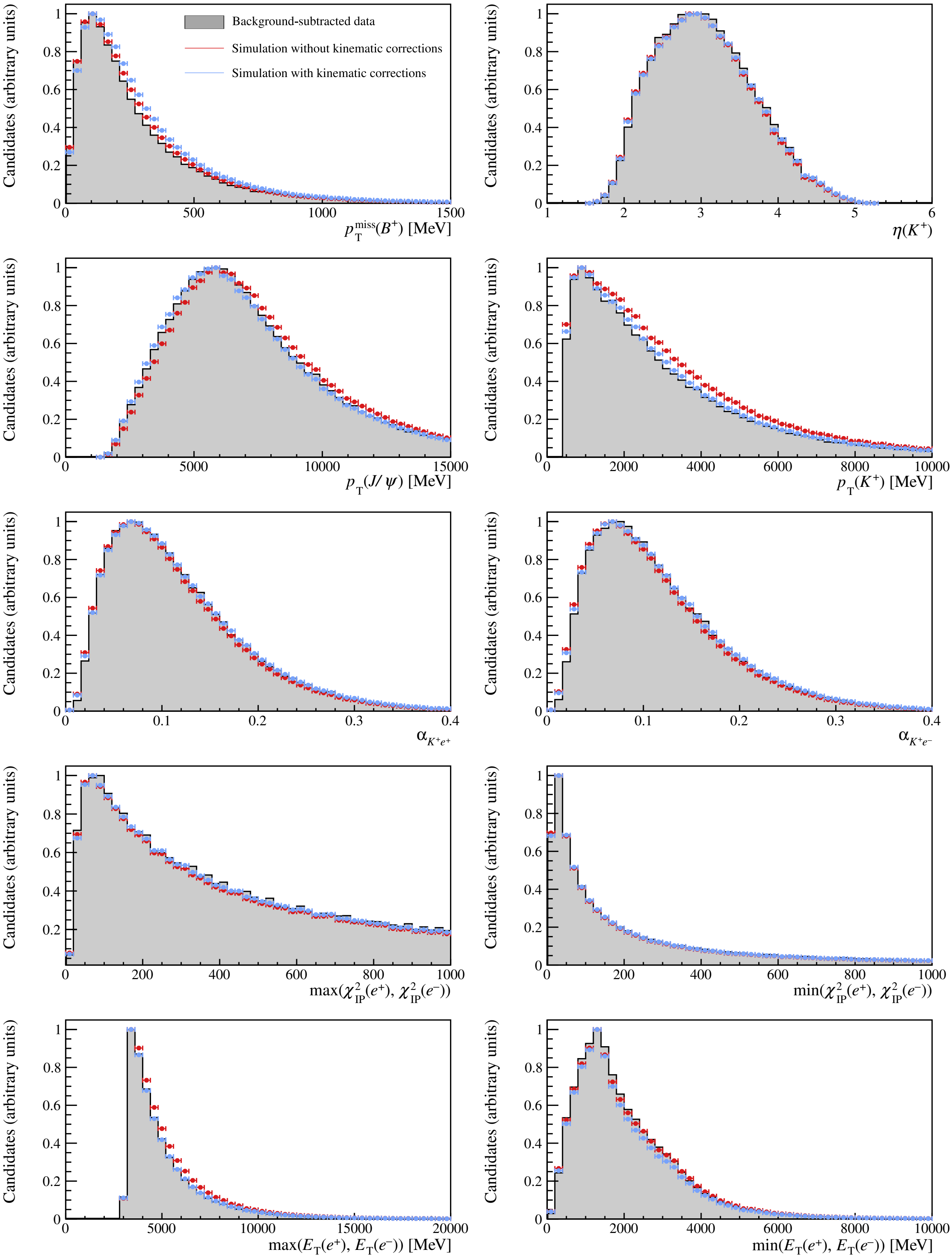

Figure C.2: Supplementary distributions of kinematic variables in $B^+ \to K^+ J/\psi(e^+e^-)$ 2018 *e*TOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

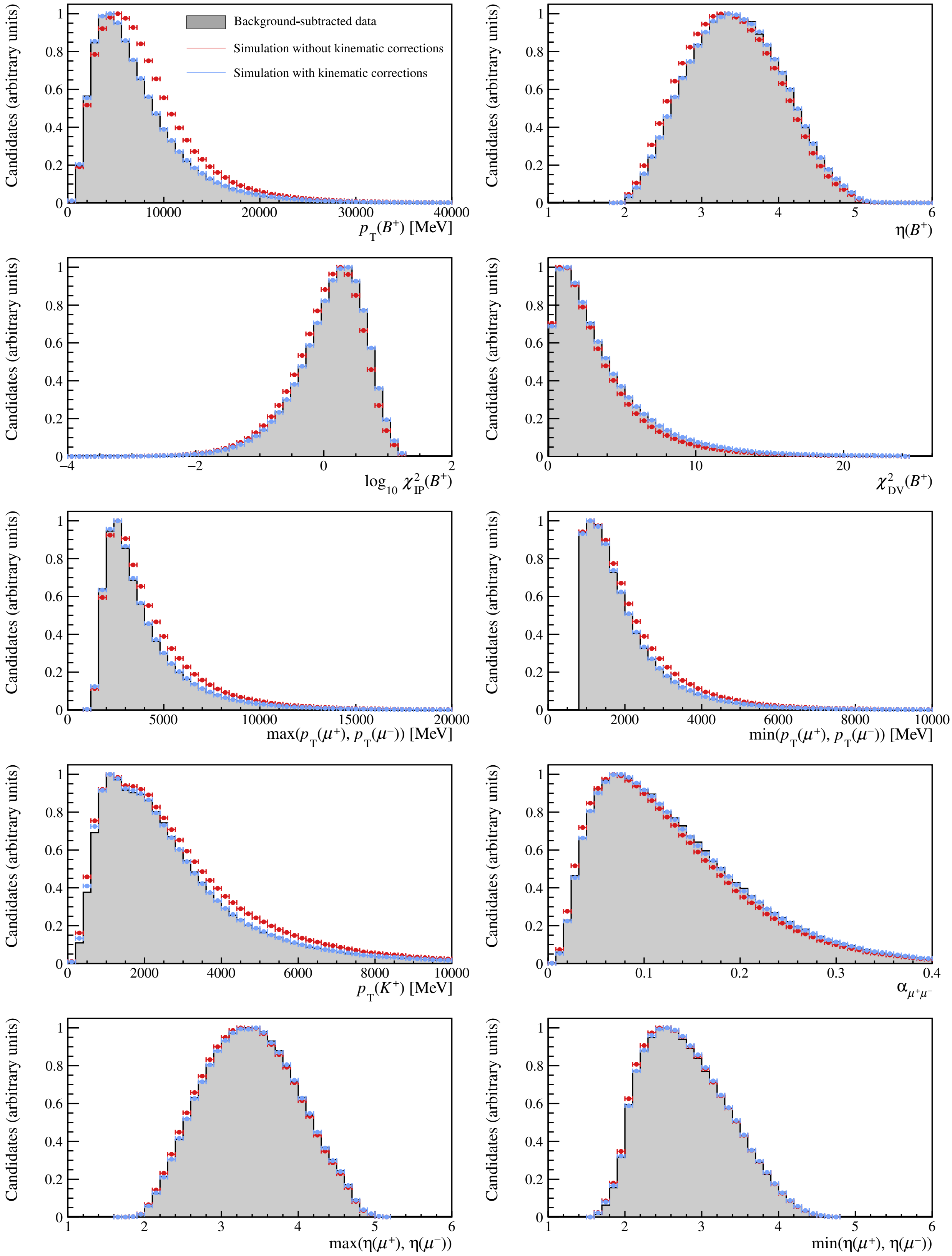

Figure C.3: Distributions of kinematic variables in $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ 2017 μTOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

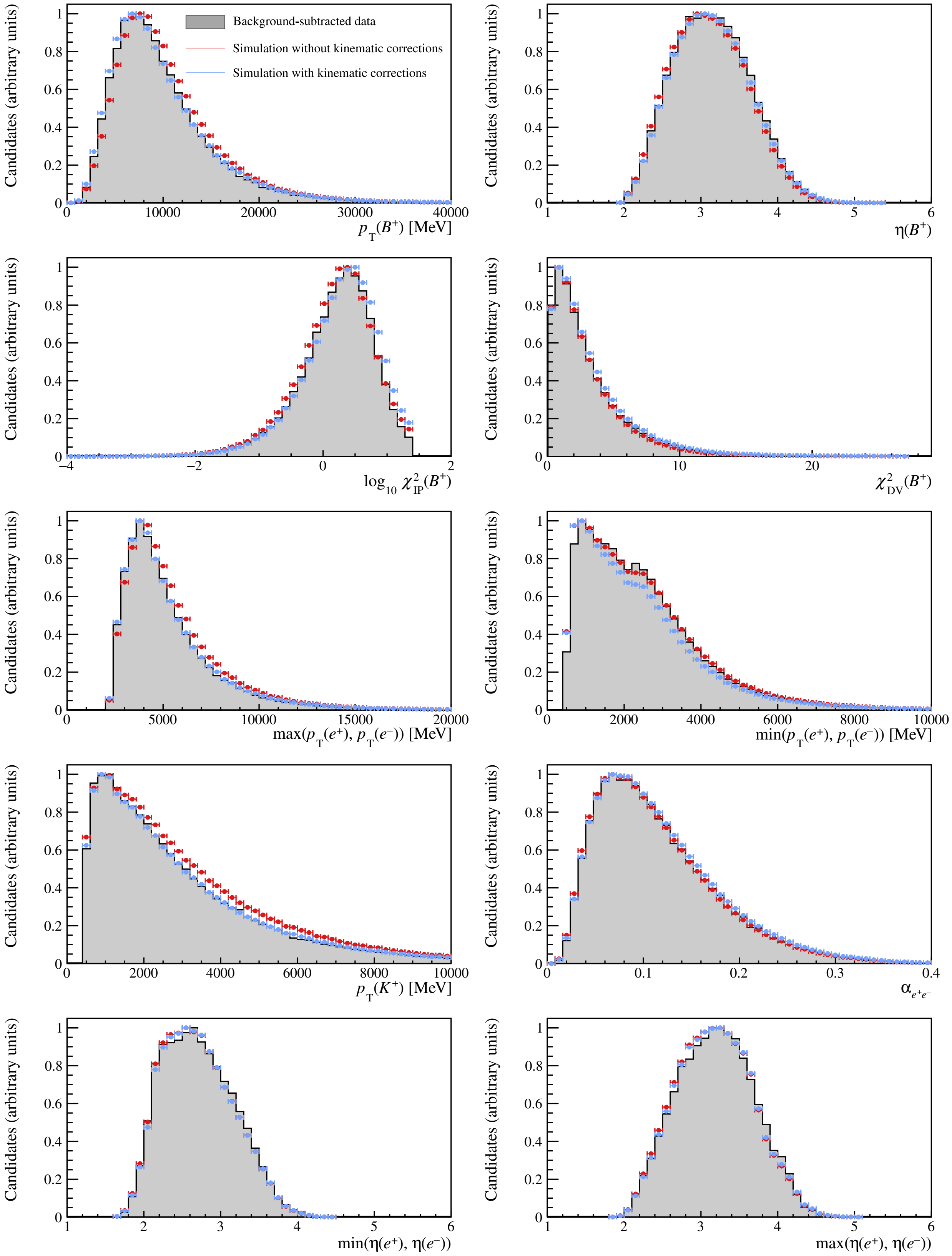

Figure C.4: Distributions of kinematic variables in $B^+ \to K^+J/\psi(e^+e^-)$ 2017 *e*TOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

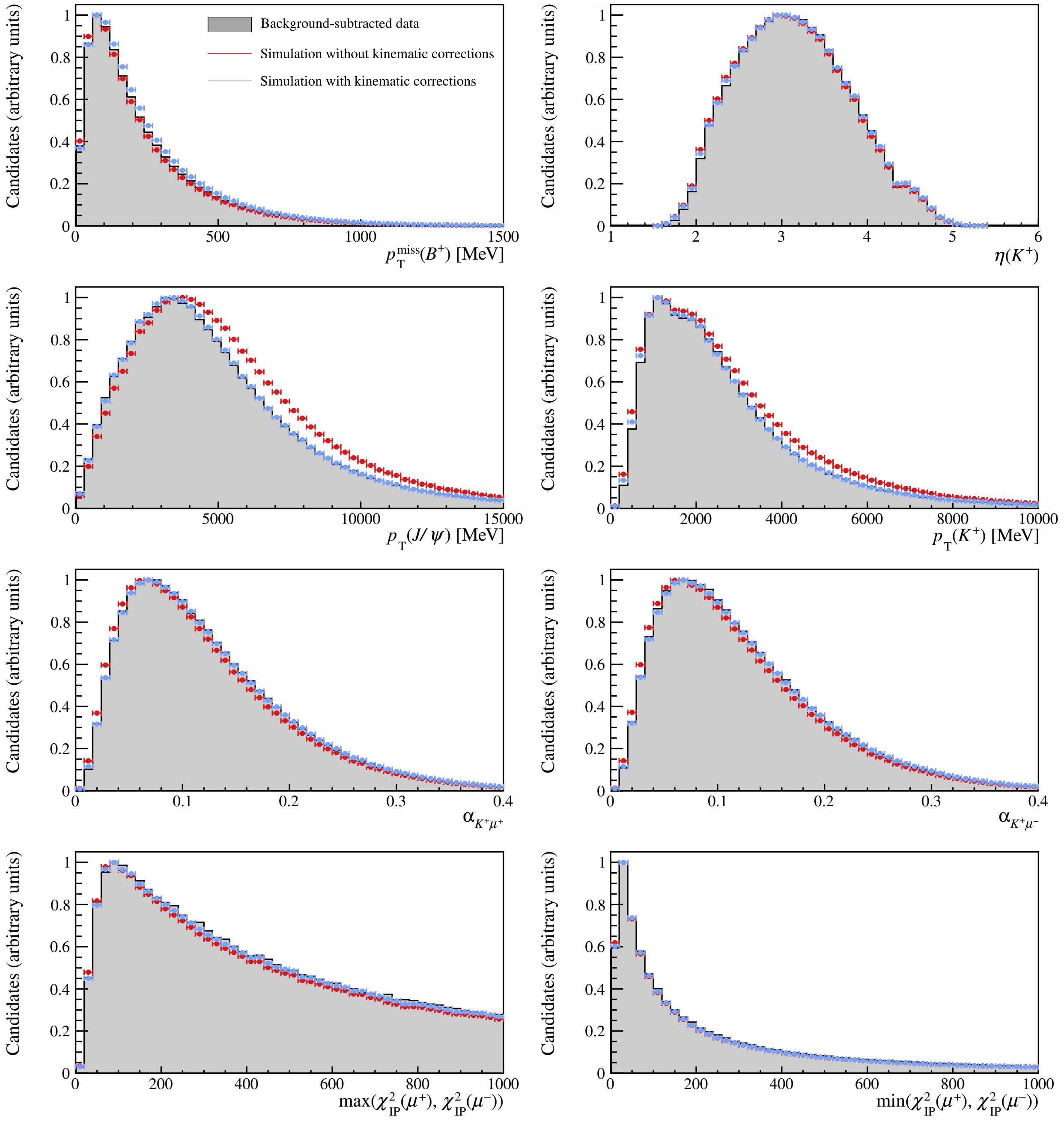

Figure C.5: Supplementary distributions of kinematic variables in $B^+ \to K^+ J/\psi(\mu^+\mu^-)$ 2017 μTOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

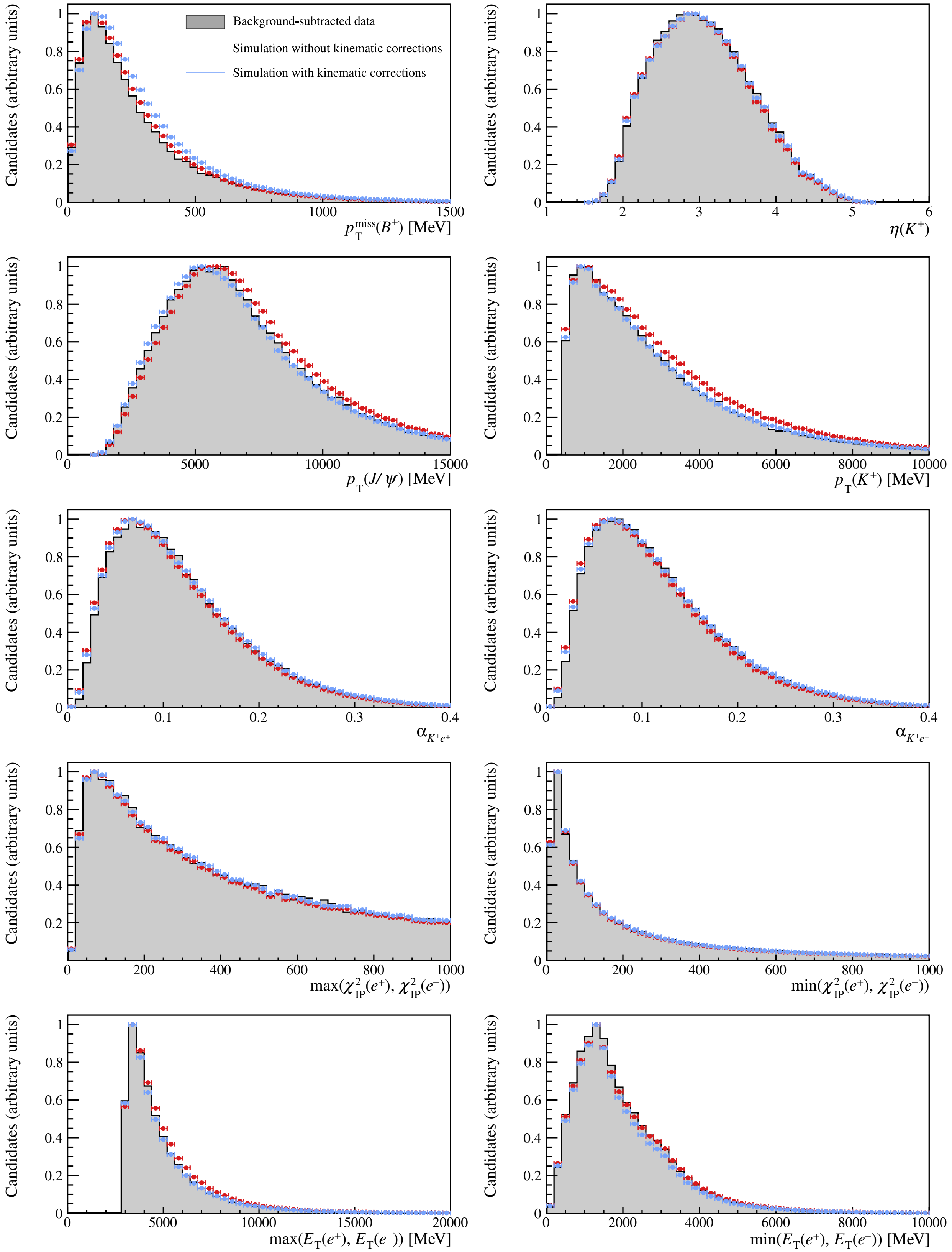

Figure C.6: Supplementary distributions of kinematic variables in $B^+ \to K^+ J/\psi(e^+e^-)$ 2017 *e*TOS samples. The black histograms show the distributions in data, whilst the red and blue bins depict the simulated distributions before and after the kinematic calibration, respectively.

Appendix D

Tables of intermediary efficiencies

The total efficiencies listed in Table 6.4 are factorised in several terms, as per Equation (6.1). These intermediary efficiencies are listed in Tables D.1 and D.2. The largest discrepancy in rare and control efficiencies lies in $\varepsilon_{\text{presel}}$. This is the one that includes the q^2 cut efficiency, and so it reflects the fact that several $B^+ \to K^+\ell^+\ell^-$ events lie outside the $q^2 \in (1.1\ \text{GeV}^2,\ 6.0\ \text{GeV}^2)$ window. Other smaller differences can be found in $\varepsilon_{\ell\text{TOS}}$, due to the slightly different kinematics, and in ε_{BDT}, as a result of the mass cuts.

Table D.1: Individual efficiencies of the resonant and signal μTOS and eTOS events. The total efficiency is listed as a percentage.

	$\varepsilon_{\text{rare}}$	$\varepsilon_{\text{control}}$	$\varepsilon_{\psi(2S)}$	$\varepsilon_{\text{rare}}$	$\varepsilon_{\text{control}}$	$\varepsilon_{\psi(2S)}$
	Run 1 μTOS			**Run 1 eTOS**		
$\varepsilon_{\text{geom}}$	0.1646 ± 0.0003	0.1617 ± 0.0003	0.1668 ± 0.0003	0.1635 ± 0.0003	0.1609 ± 0.0003	0.1668 ± 0.0003
$\varepsilon_{\text{rec,strip}}$	0.2346 ± 0.0003	0.2386 ± 0.0001	0.2398 ± 0.0002	0.1268 ± 0.0002	0.1312 ± 0.0002	0.1346 ± 0.0004
$\varepsilon_{\text{presel}}$	0.1452 ± 0.0005	0.6001 ± 0.0003	0.6225 ± 0.0004	0.2618 ± 0.0007	0.7546 ± 0.0006	0.6688 ± 0.0014
ε_{PID}	0.9061 ± 0.0014	0.9132 ± 0.0003	0.9127 ± 0.0004	0.7661 ± 0.0017	0.7649 ± 0.0008	0.7719 ± 0.0019
$\varepsilon_{\text{trig}}$	0.7159 ± 0.0018	0.7543 ± 0.0004	0.7925 ± 0.0004	0.1663 ± 0.0011	0.1991 ± 0.0006	0.2386 ± 0.0015
ε_{BDT}	0.8702 ± 0.0016	0.8758 ± 0.0004	0.8367 ± 0.0004	0.7869 ± 0.0031	0.8933 ± 0.0011	0.8722 ± 0.0024
ε_{tot} [%]	0.3165 ± 0.0016	1.3964 ± 0.0027	1.5066 ± 0.0035	0.0544 ± 0.0005	0.2167 ± 0.0009	0.2413 ± 0.0019
	Run 2.1 μTOS			**Run 2.1 eTOS**		
$\varepsilon_{\text{geom}}$	0.1794 ± 0.0003	0.1766 ± 0.0003	0.1756 ± 0.0003	0.1787 ± 0.0003	0.1758 ± 0.0003	0.1756 ± 0.0003
$\varepsilon_{\text{rec,strip}}$	0.2270 ± 0.0003	0.2332 ± 0.0001	0.2345 ± 0.0002	0.1317 ± 0.0002	0.1366 ± 0.0002	0.1440 ± 0.0004
$\varepsilon_{\text{presel}}$	0.1437 ± 0.0005	0.5939 ± 0.0003	0.6154 ± 0.0004	0.2646 ± 0.0008	0.7299 ± 0.0005	0.6234 ± 0.0014
ε_{PID}	0.9281 ± 0.0011	0.9318 ± 0.0003	0.9316 ± 0.0004	0.7746 ± 0.0017	0.7739 ± 0.0007	0.7779 ± 0.0019
$\varepsilon_{\text{trig}}$	0.6552 ± 0.0016	0.6900 ± 0.0004	0.7325 ± 0.0005	0.2783 ± 0.0015	0.3430 ± 0.0006	0.3994 ± 0.0018
ε_{BDT}	0.9094 ± 0.0013	0.9112 ± 0.0003	0.8696 ± 0.0005	0.7755 ± 0.0026	0.9001 ± 0.0007	0.8717 ± 0.0020
ε_{tot} [%]	0.3236 ± 0.0015	1.4322 ± 0.0028	1.5038 ± 0.0034	0.1041 ± 0.0007	0.4188 ± 0.0013	0.4268 ± 0.0028
	2017 μTOS			**2017 eTOS**		
$\varepsilon_{\text{geom}}$	0.1789 ± 0.0002	0.1761 ± 0.0002	0.1792 ± 0.0002	0.1779 ± 0.0003	0.1754 ± 0.0002	0.1793 ± 0.0004
$\varepsilon_{\text{rec,strip}}$	0.2626 ± 0.0003	0.2705 ± 0.0001	0.2707 ± 0.0002	0.1332 ± 0.0002	0.1382 ± 0.0001	0.1438 ± 0.0002
$\varepsilon_{\text{presel}}$	0.1196 ± 0.0004	0.5184 ± 0.0003	0.5421 ± 0.0005	0.2264 ± 0.0008	0.7301 ± 0.0004	0.6266 ± 0.0009
ε_{PID}	0.9427 ± 0.0012	0.9480 ± 0.0003	0.9491 ± 0.0004	0.7466 ± 0.0021	0.7440 ± 0.0005	0.7477 ± 0.0012
$\varepsilon_{\text{trig}}$	0.7300 ± 0.0018	0.7627 ± 0.0004	0.7913 ± 0.0006	0.2682 ± 0.0018	0.3286 ± 0.0005	0.3841 ± 0.0011
ε_{BDT}	0.8927 ± 0.0015	0.9058 ± 0.0003	0.8815 ± 0.0005	0.7569 ± 0.0033	0.9147 ± 0.0005	0.9100 ± 0.0010
ε_{tot} [%]	0.3452 ± 0.0017	1.6174 ± 0.0027	1.7400 ± 0.0040	0.0813 ± 0.0007	0.3957 ± 0.0009	0.4221 ± 0.0018
	2018 μTOS			**2018 eTOS**		
$\varepsilon_{\text{geom}}$	0.1791 ± 0.0003	0.1764 ± 0.0002	0.1794 ± 0.0002	0.1782 ± 0.0003	0.1756 ± 0.0002	0.1798 ± 0.0004
$\varepsilon_{\text{rec,strip}}$	0.2622 ± 0.0003	0.2699 ± 0.0001	0.2691 ± 0.0002	0.1328 ± 0.0002	0.1384 ± 0.0001	0.1441 ± 0.0002
$\varepsilon_{\text{presel}}$	0.1191 ± 0.0004	0.5197 ± 0.0003	0.5437 ± 0.0003	0.2268 ± 0.0008	0.7325 ± 0.0004	0.6284 ± 0.0007
ε_{PID}	0.9460 ± 0.0012	0.9519 ± 0.0002	0.9529 ± 0.0003	0.7435 ± 0.0021	0.7413 ± 0.0005	0.7444 ± 0.0009
$\varepsilon_{\text{trig}}$	0.7087 ± 0.0019	0.7485 ± 0.0004	0.7850 ± 0.0004	0.2452 ± 0.0017	0.3026 ± 0.0004	0.3574 ± 0.0009
ε_{BDT}	0.8799 ± 0.0016	0.8855 ± 0.0003	0.8541 ± 0.0004	0.7672 ± 0.0034	0.9312 ± 0.0005	0.9274 ± 0.0008
ε_{tot} [%]	0.3300 ± 0.0017	1.5609 ± 0.0027	1.6768 ± 0.0029	0.0751 ± 0.0007	0.3719 ± 0.0009	0.4015 ± 0.0009

Table D.2: Individual efficiencies of the resonant and signal *h*TOS! and TIS! events. The total efficiency is listed as a percentage.

	$\varepsilon_{\text{rare}}$	$\varepsilon_{\text{control}}$	$\varepsilon_{\psi(2S)}$	$\varepsilon_{\text{rare}}$	$\varepsilon_{\text{control}}$	$\varepsilon_{\psi(2S)}$
		Run 1 *h*TOS!			**Run 1 TIS!**	
$\varepsilon_{\text{geom}}$	0.1635 ± 0.0003	0.1609 ± 0.0003	0.1668 ± 0.0003	0.1635 ± 0.0003	0.1609 ± 0.0003	0.1668 ± 0.0003
$\varepsilon_{\text{rec,strip}}$	0.1268 ± 0.0002	0.1312 ± 0.0002	0.1346 ± 0.0004	0.1268 ± 0.0002	0.1312 ± 0.0002	0.1346 ± 0.0004
$\varepsilon_{\text{presel}}$	0.2618 ± 0.0007	0.7546 ± 0.0006	0.6688 ± 0.0014	0.2618 ± 0.0007	0.7546 ± 0.0006	0.6688 ± 0.0014
ε_{PID}	0.7499 ± 0.0017	0.7671 ± 0.0008	0.7858 ± 0.0019	0.7652 ± 0.0016	0.7761 ± 0.0008	0.7787 ± 0.0019
$\varepsilon_{\text{trig}}$	0.0520 ± 0.0006	0.0288 ± 0.0002	0.0099 ± 0.0003	0.0639 ± 0.0007	0.0658 ± 0.0004	0.0611 ± 0.0008
ε_{BDT}	0.8380 ± 0.0050	0.9962 ± 0.0004	1.0220 ± 0.0050	0.7260 ± 0.0050	0.8555 ± 0.0019	0.8590 ± 0.0050
ε_{tot} [%]	0.0178 ± 0.0002	0.0350 ± 0.0003	0.0120 ± 0.0004	0.0193 ± 0.0003	0.0696 ± 0.0004	0.0614 ± 0.0009
		Run 2.1 *h*TOS!			**Run 2.1 TIS!**	
$\varepsilon_{\text{geom}}$	0.1787 ± 0.0003	0.1758 ± 0.0003	0.1756 ± 0.0003	0.1787 ± 0.0003	0.1758 ± 0.0003	0.1756 ± 0.0003
$\varepsilon_{\text{rec,strip}}$	0.1317 ± 0.0002	0.1366 ± 0.0002	0.1440 ± 0.0004	0.1317 ± 0.0002	0.1366 ± 0.0002	0.1440 ± 0.0004
$\varepsilon_{\text{presel}}$	0.2646 ± 0.0008	0.7299 ± 0.0005	0.6234 ± 0.0014	0.2646 ± 0.0008	0.7299 ± 0.0005	0.6234 ± 0.0014
ε_{PID}	0.7696 ± 0.0017	0.7807 ± 0.0007	0.7879 ± 0.0018	0.7765 ± 0.0017	0.7853 ± 0.0007	0.7876 ± 0.0018
$\varepsilon_{\text{trig}}$	0.0771 ± 0.0008	0.0365 ± 0.0002	0.0127 ± 0.0004	0.0900 ± 0.0010	0.0875 ± 0.0004	0.0824 ± 0.0010
ε_{BDT}	0.7160 ± 0.0050	0.9319 ± 0.0018	0.941 ± 0.00805	0.7270 ± 0.0050	0.9091 ± 0.0014	0.9120 ± 0.0040
ε_{tot} [%]	0.0264 ± 0.0004	0.0466 ± 0.0003	0.0148 ± 0.0005	0.0316 ± 0.0004	0.1095 ± 0.0006	0.0933 ± 0.0013
		2017 *h*TOS!			**2017 TIS!**	
$\varepsilon_{\text{geom}}$	0.1779 ± 0.0003	0.1754 ± 0.0002	0.1793 ± 0.0004	0.1779 ± 0.0003	0.1754 ± 0.0002	0.1793 ± 0.0004
$\varepsilon_{\text{rec,strip}}$	0.1332 ± 0.0002	0.1382 ± 0.0001	0.1438 ± 0.0002	0.1332 ± 0.0002	0.1382 ± 0.0001	0.1438 ± 0.0002
$\varepsilon_{\text{presel}}$	0.2264 ± 0.0008	0.7301 ± 0.0004	0.6266 ± 0.0009	0.2264 ± 0.0008	0.7301 ± 0.0004	0.6266 ± 0.0009
ε_{PID}	0.7459 ± 0.0021	0.7528 ± 0.0005	0.7637 ± 0.0011	0.7536 ± 0.0021	0.7585 ± 0.0005	0.7596 ± 0.0012
$\varepsilon_{\text{trig}}$	0.0845 ± 0.0011	0.0431 ± 0.0002	0.0146 ± 0.0002	0.1019 ± 0.0012	0.1006 ± 0.0002	0.0932 ± 0.0006
ε_{BDT}	0.6850 ± 0.0060	0.9199 ± 0.0013	0.9550 ± 0.0040	0.6990 ± 0.0060	0.8962 ± 0.0010	0.9089 ± 0.0020
ε_{tot} [%]	0.0232 ± 0.0004	0.0528 ± 0.0003	0.0172 ± 0.0003	0.0288 ± 0.0004	0.1210 ± 0.0004	0.1039 ± 0.0008
		2018 *h*TOS!			**2018 TIS!**	
$\varepsilon_{\text{geom}}$	0.1782 ± 0.0003	0.1756 ± 0.0002	0.1798 ± 0.0004	0.1782 ± 0.0003	0.1756 ± 0.0002	0.1798 ± 0.0004
$\varepsilon_{\text{rec,strip}}$	0.1328 ± 0.0002	0.1384 ± 0.0001	0.1441 ± 0.0002	0.1328 ± 0.0002	0.1384 ± 0.0001	0.1441 ± 0.0002
$\varepsilon_{\text{presel}}$	0.2268 ± 0.0008	0.7325 ± 0.0004	0.6284 ± 0.0007	0.2268 ± 0.0008	0.7325 ± 0.0004	0.6284 ± 0.0007
ε_{PID}	0.7370 ± 0.0021	0.7463 ± 0.0005	0.7577 ± 0.0009	0.7458 ± 0.0021	0.7516 ± 0.0005	0.7533 ± 0.0009
$\varepsilon_{\text{trig}}$	0.0827 ± 0.0011	0.0428 ± 0.0002	0.0152 ± 0.0002	0.0939 ± 0.0011	0.0958 ± 0.0003	0.0912 ± 0.0005
ε_{BDT}	0.7300 ± 0.0060	0.9540 ± 0.0011	0.9819 ± 0.0024	0.6800 ± 0.0060	0.8861 ± 0.0010	0.9095 ± 0.0018
ε_{tot} [%]	0.0239 ± 0.0004	0.0542 ± 0.0003	0.0184 ± 0.0003	0.0255 ± 0.0004	0.1136 ± 0.0004	0.1016 ± 0.0006

Appendix E

Additional 1D $r_{J/\psi}$ tests

This appendix contains results from the 1D $r_{J/\psi}$ check described in Section 8.1.2. The main body shows the outcome from 2018 *e*TOS data, whereas their equivalents from *h*TOS! and TIS! data are presented in Figures E.1 and E.2, respectively. The results from 2017 *e*TOS, *h*TOS!, and TIS! data are shown in Figures E.3, E.4, and E.5, respectively.

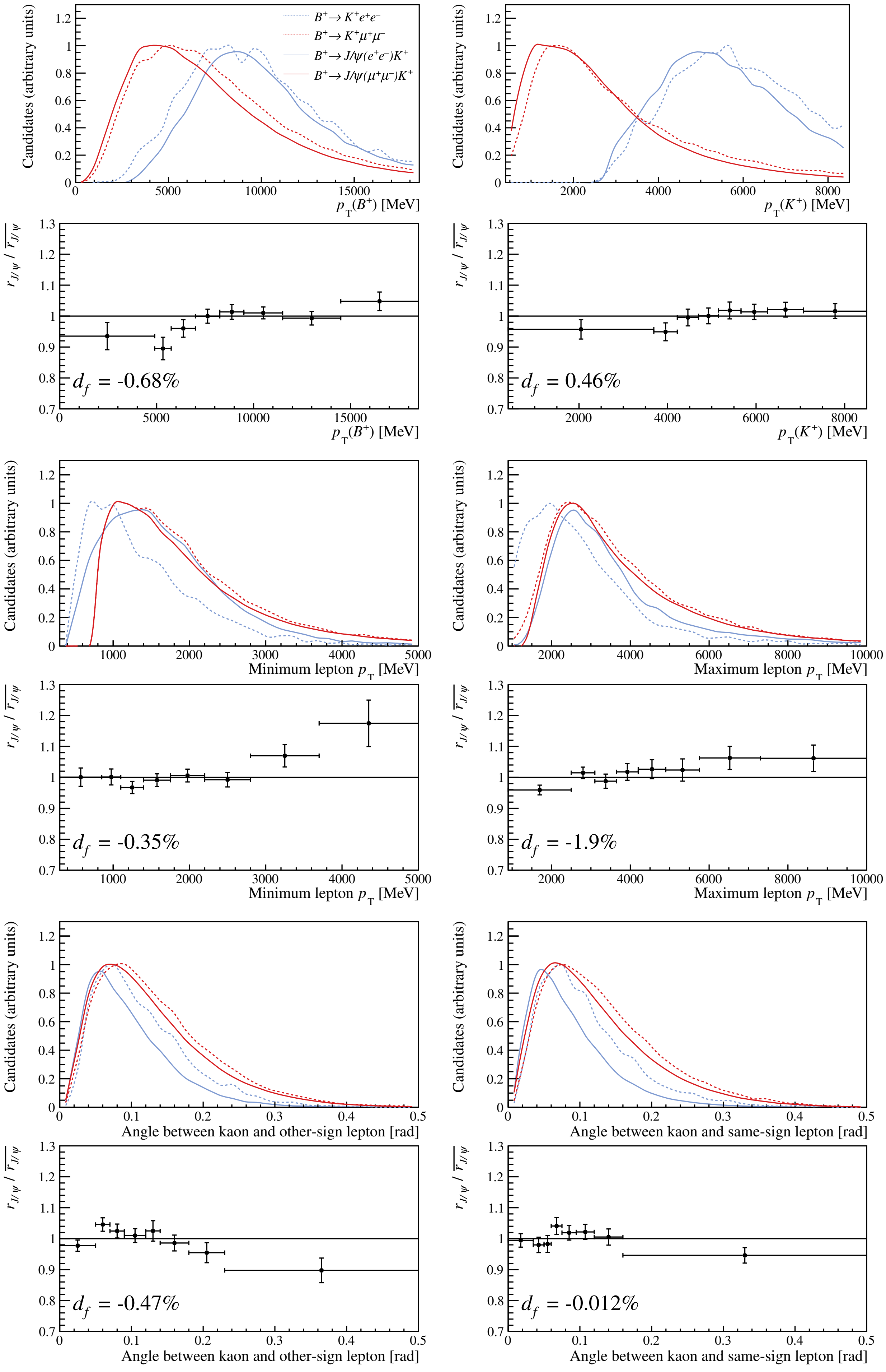

Figure E.1: Plots of $r_{J/\psi}$ obtained from 2018 hTOS! and μTOS data, as a function of variables relevant to the detector response. The flatness parameters d_f are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control-mode samples. The uncertainties shown are purely statistical.

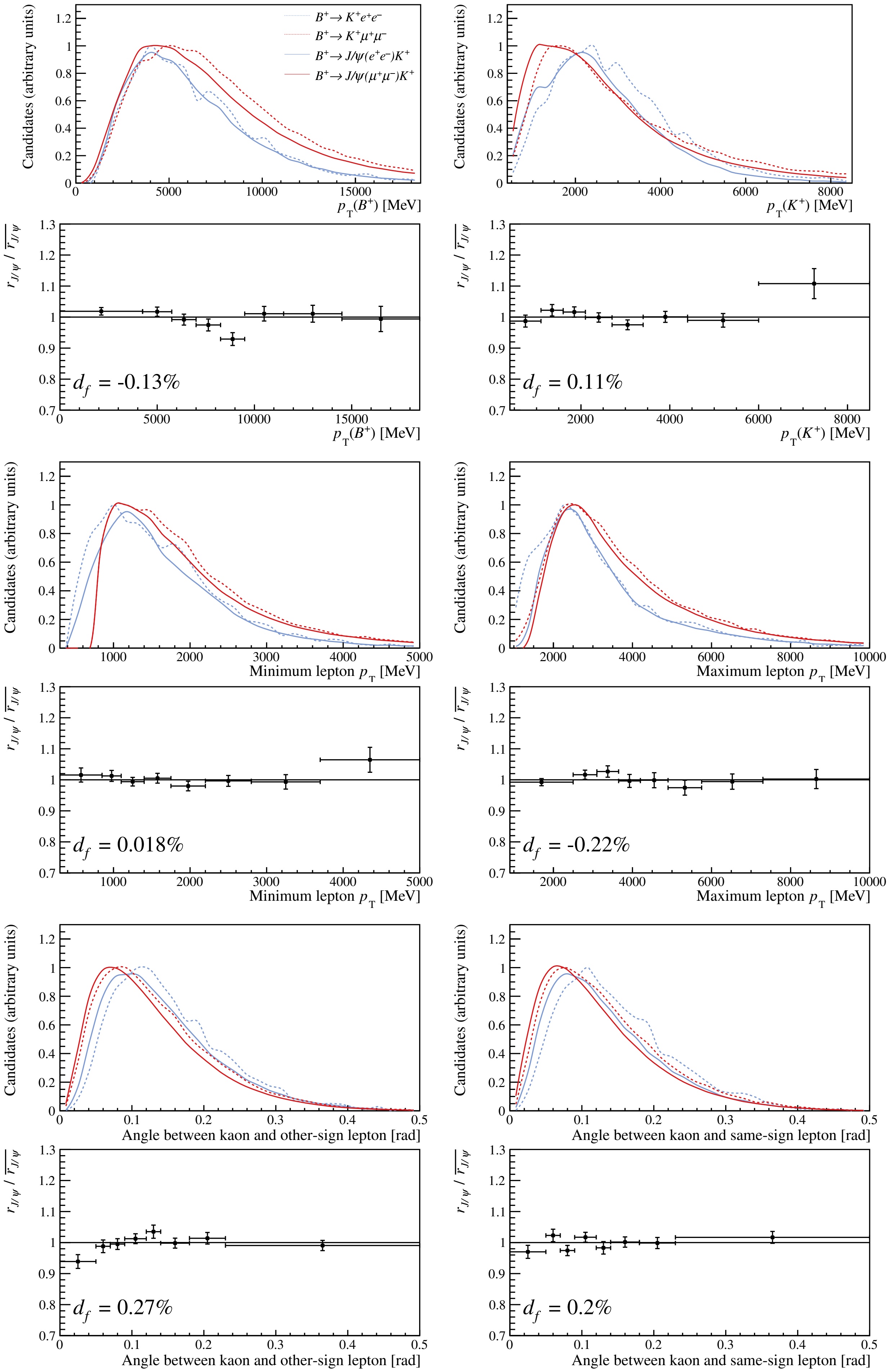

Figure E.2: Plots of $r_{J/\psi}$ obtained from 2018 TIS! and μTOS data, as a function of variables relevant to the detector response. The flatness parameters d_f are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control-mode samples. The uncertainties shown are purely statistical.

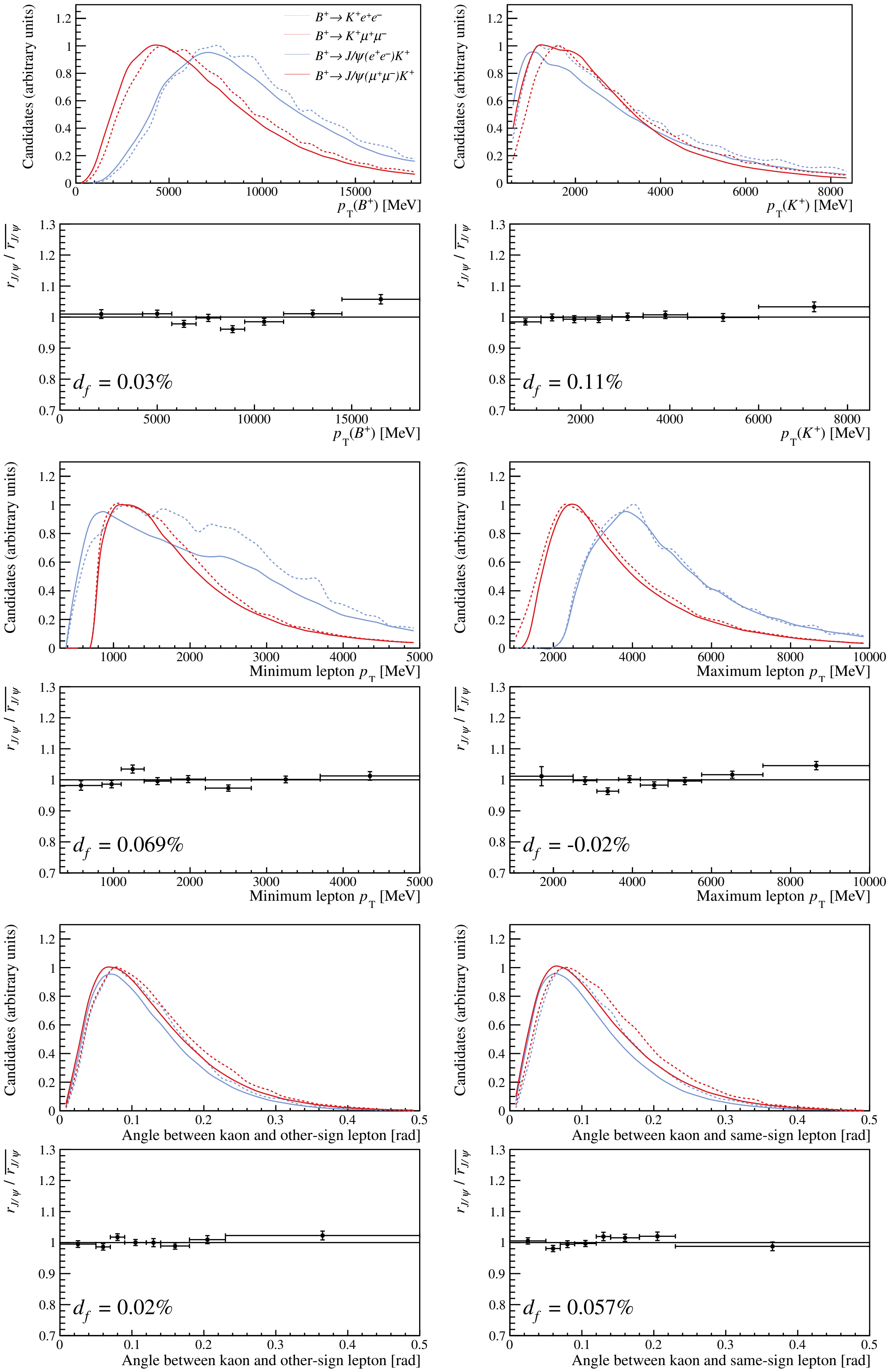

Figure E.3: Plots of $r_{J/\psi}$ obtained from 2017 eTOS and μTOS data, as a function of variables relevant to the detector response. The flatness parameters d_f are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control-mode samples. The uncertainties shown are purely statistical.

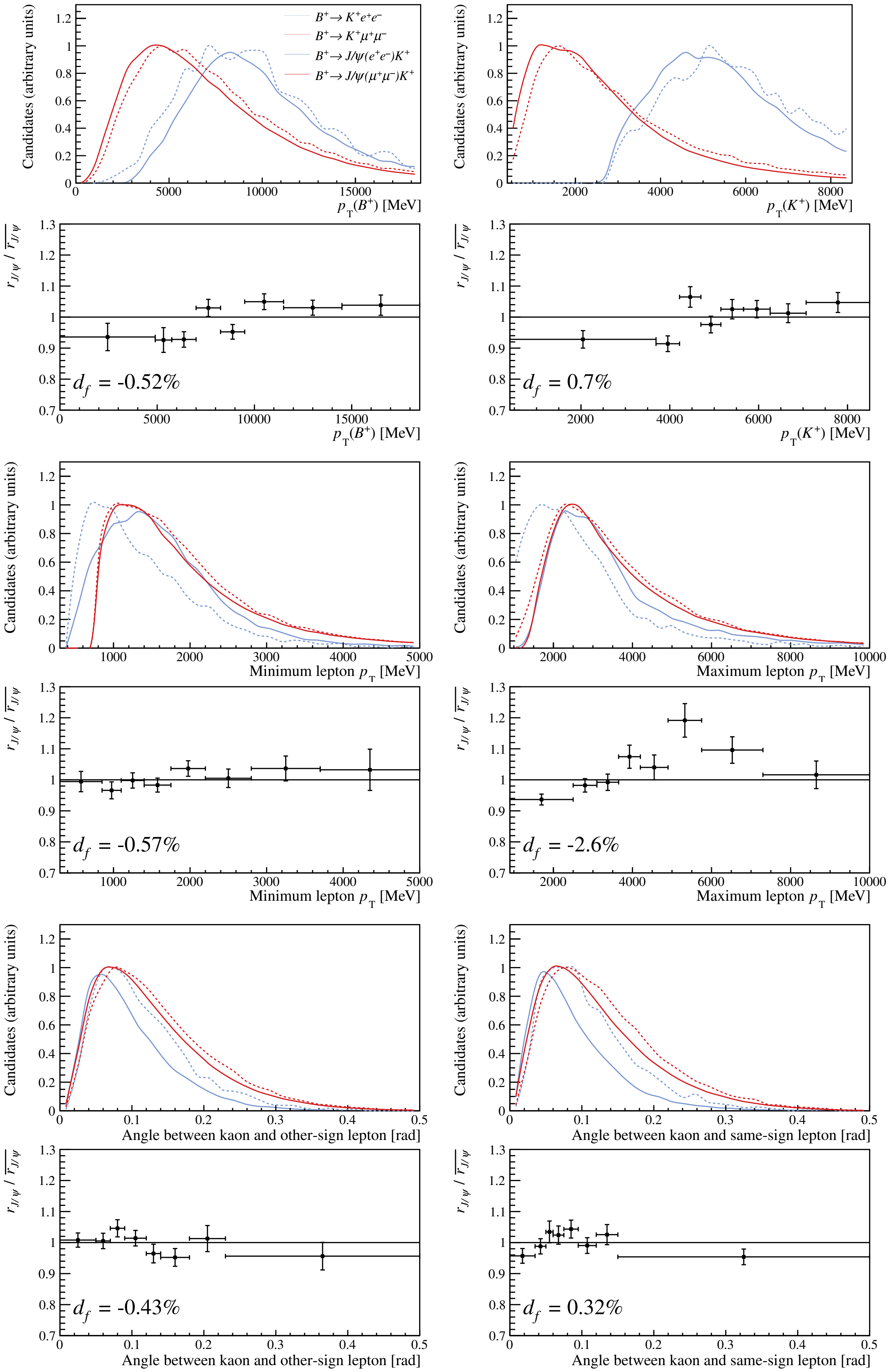

Figure E.4: Plots of $r_{J/\psi}$ obtained from 2017 hTOS! and μTOS data, as a function of variables relevant to the detector response. The flatness parameters d_f are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control-mode samples. The uncertainties shown are purely statistical.

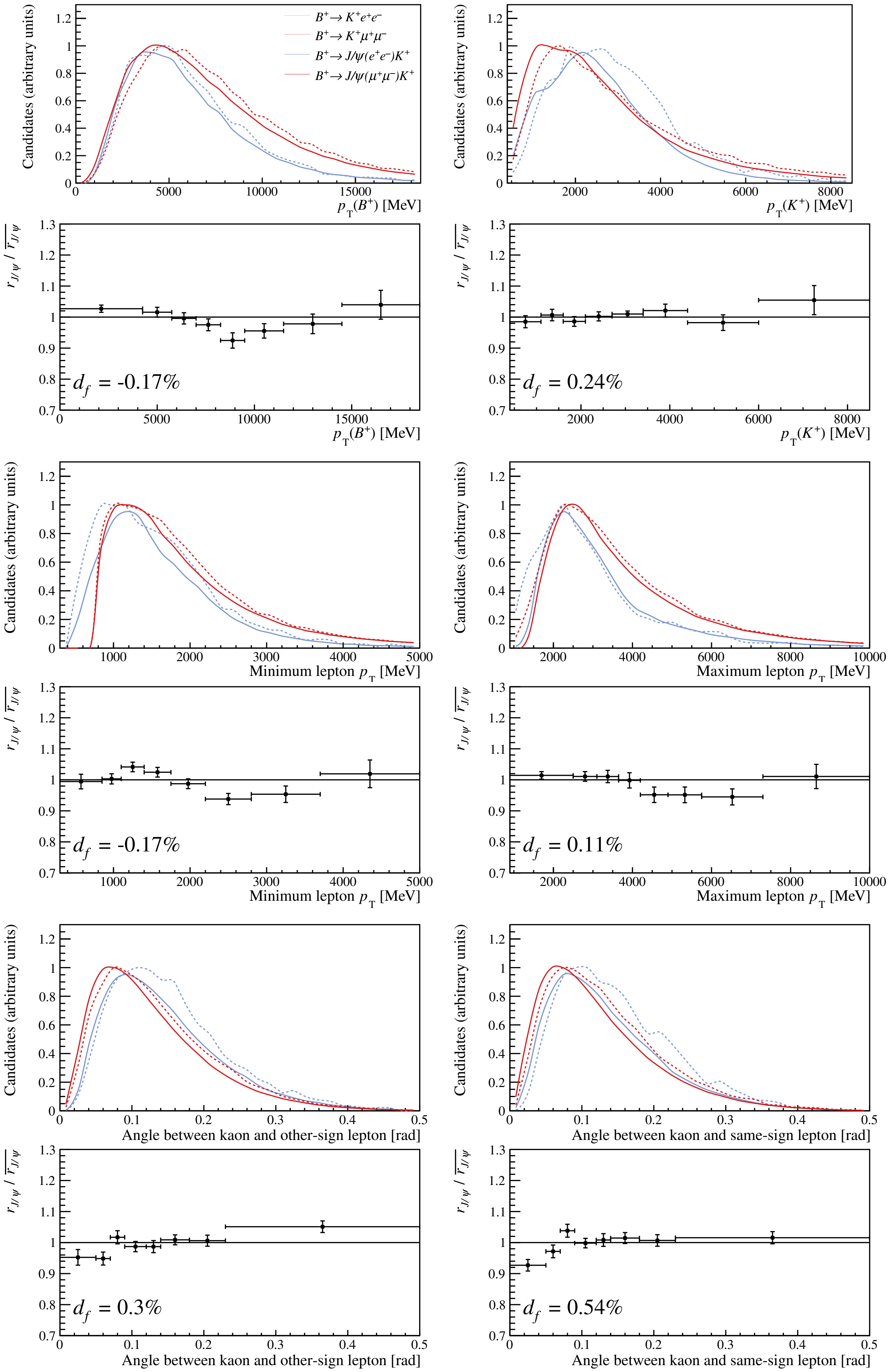

Figure E.5: Plots of $r_{J/\psi}$ obtained from 2017 TIS! and μTOS data, as a function of variables relevant to the detector response. The flatness parameters d_f are displayed at the bottom of each plot. The top of each $r_{J/\psi}$ plot shows the kinematic distributions of the binned variable, in simulated signal and control-mode samples. The uncertainties shown are purely statistical.

Appendix F

Fits to $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ data

This appendix presents in Figure F.1 the fits to $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ data, which were performed by Davide Lancierini. Similarly to the J/ψ modes, the independent variable is calculated by constraining the dilepton system to have an invariant mass compatible with the mass of the $\psi(2S)$ resonance. The fits are used to extract the yields of the $B^+ \to K^+\psi(2S)(\ell^+\ell^-)$ modes, which are used to calculate $R_{\psi(2S)}$ by virtue of Equation (8.14).

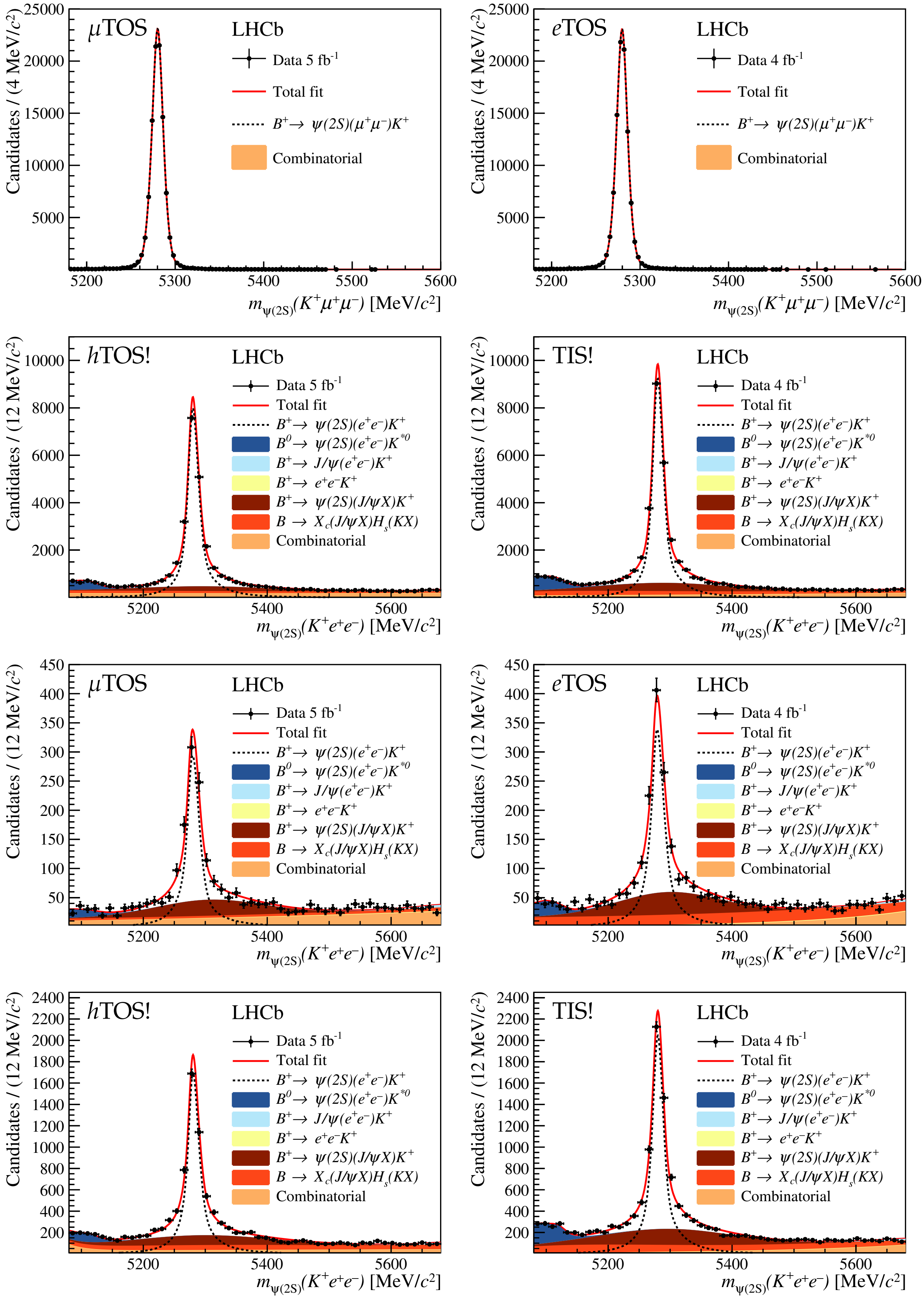

Figure F.1: Invariant-mass distributions of fully-selected $B^+\to K^+\psi(2S)(\ell^+\ell^-)$ candidates, shown separately for previous (left) and Run 2.2 (right) data, and for each trigger category. The red solid line shows the fit model, the dotted black line represents the signal component, and the orange-filled distribution depicts the combinatorial background model. Additional backgrounds are present in the electron samples, and are represented by components filled in various colours, as listed in the legend.

Appendix G

Additional compatibility checks

This appendix presents in Figures G.1 to G.5 the results of the compatibility checks presented in Section 10.3.3. They are presented in the following order: “Run 1 vs. Run 2.1”, “17 vs. 18”, “Run 2.1 vs. Run 2.2”, “Run 1 vs. Run 2”, and “old vs. new”.

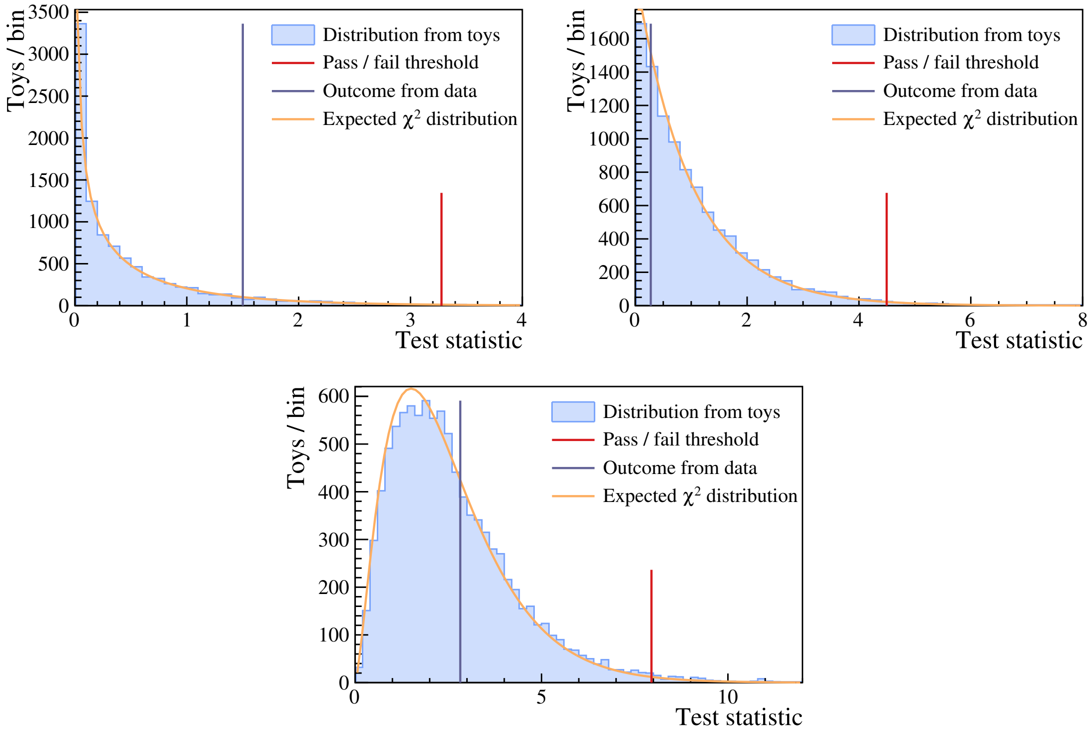

Figure G.1: Outcome of the “Run 1 vs. Run 2.1” compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The test statistic for this check is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.

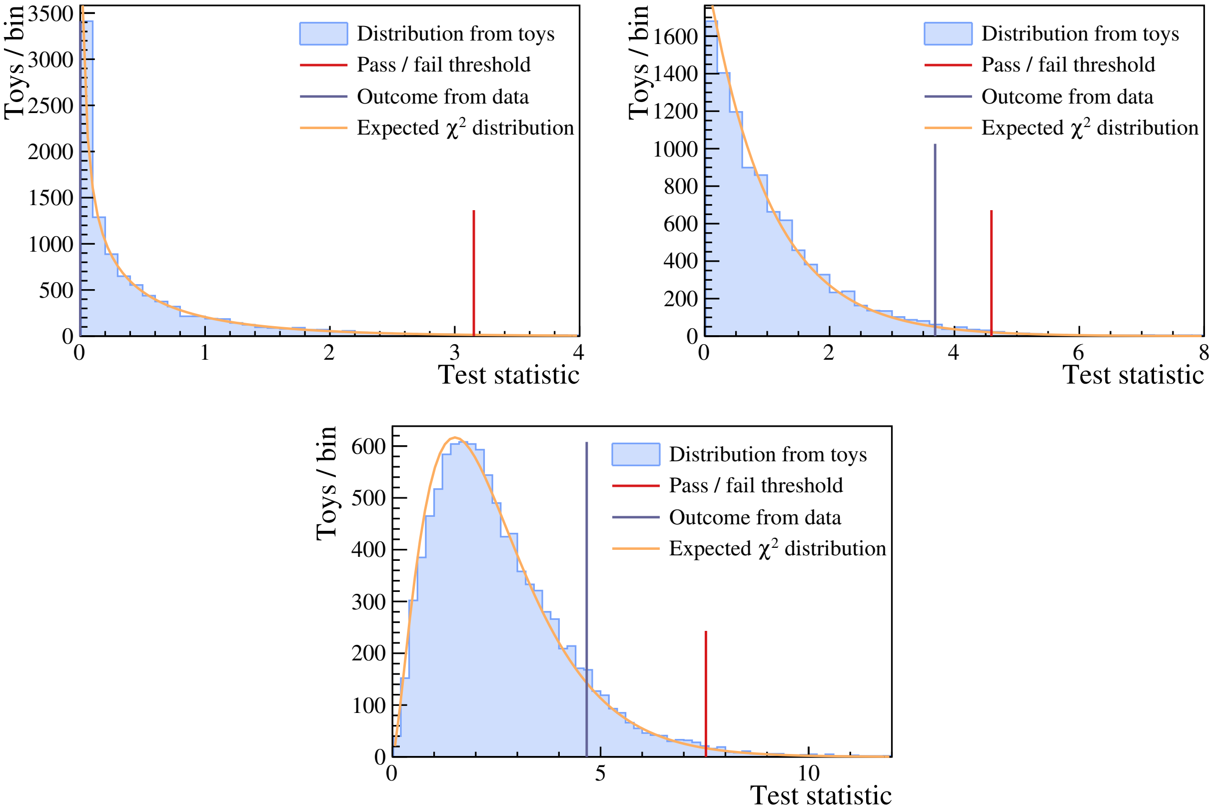

Figure G.2: Outcome of the “17 vs. 18” compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The test statistic for this check is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.

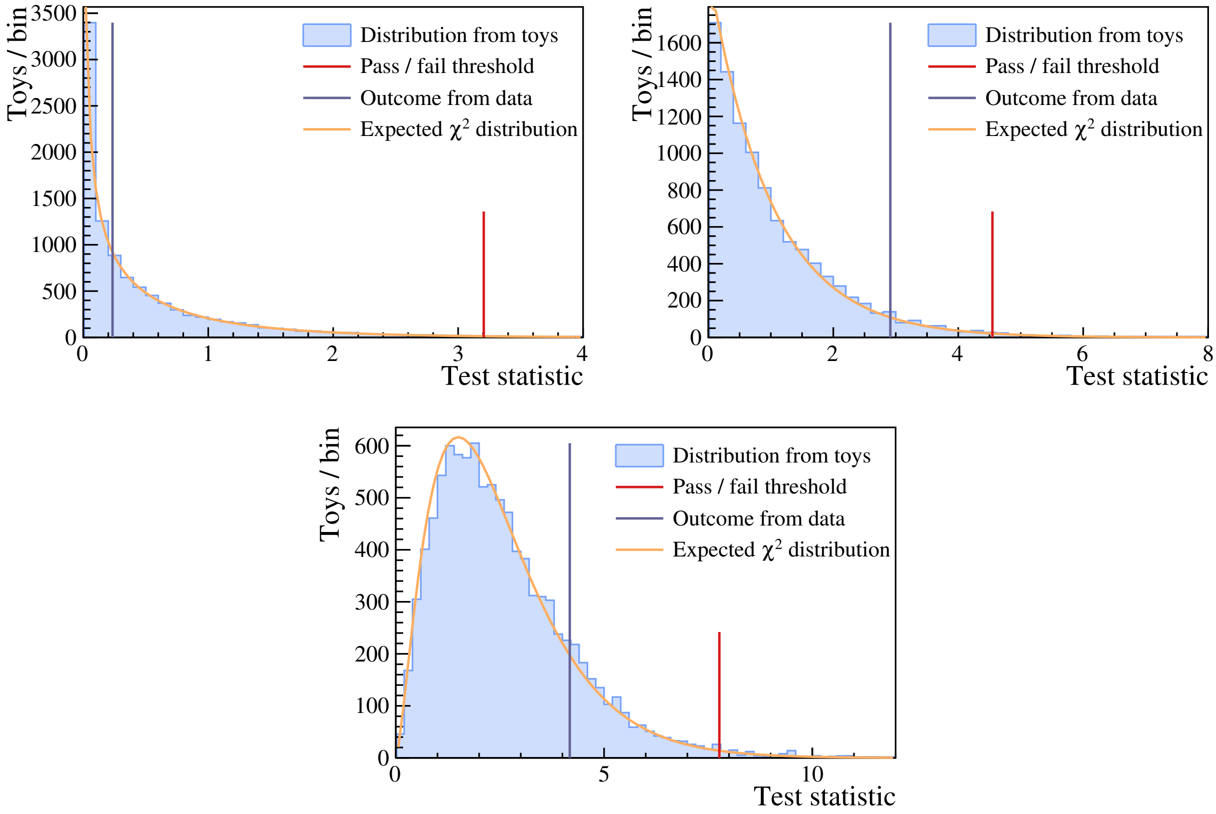

Figure G.3: Outcome of the “Run 2.1 vs. Run 2.2” compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The test statistic for this check is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.

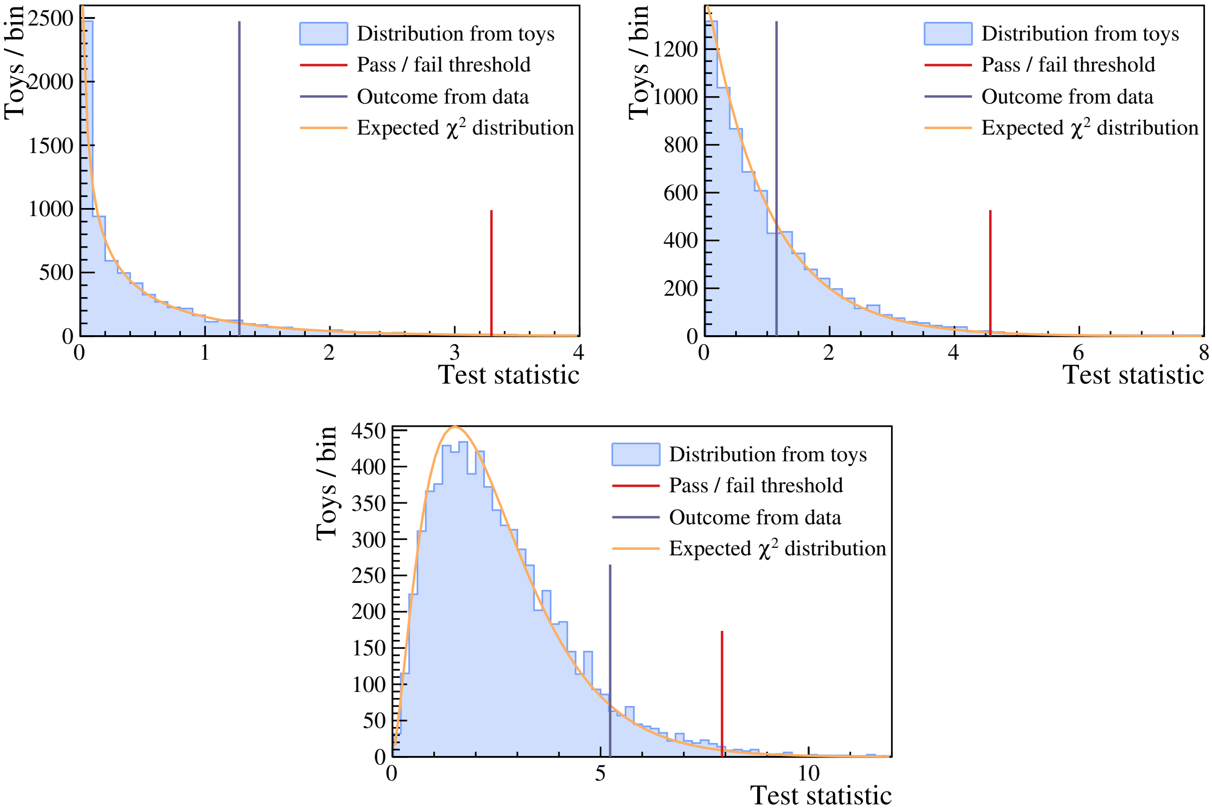

Figure G.4: Outcome of the "Run 1 vs. Run 2" compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The test statistic for this check is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.

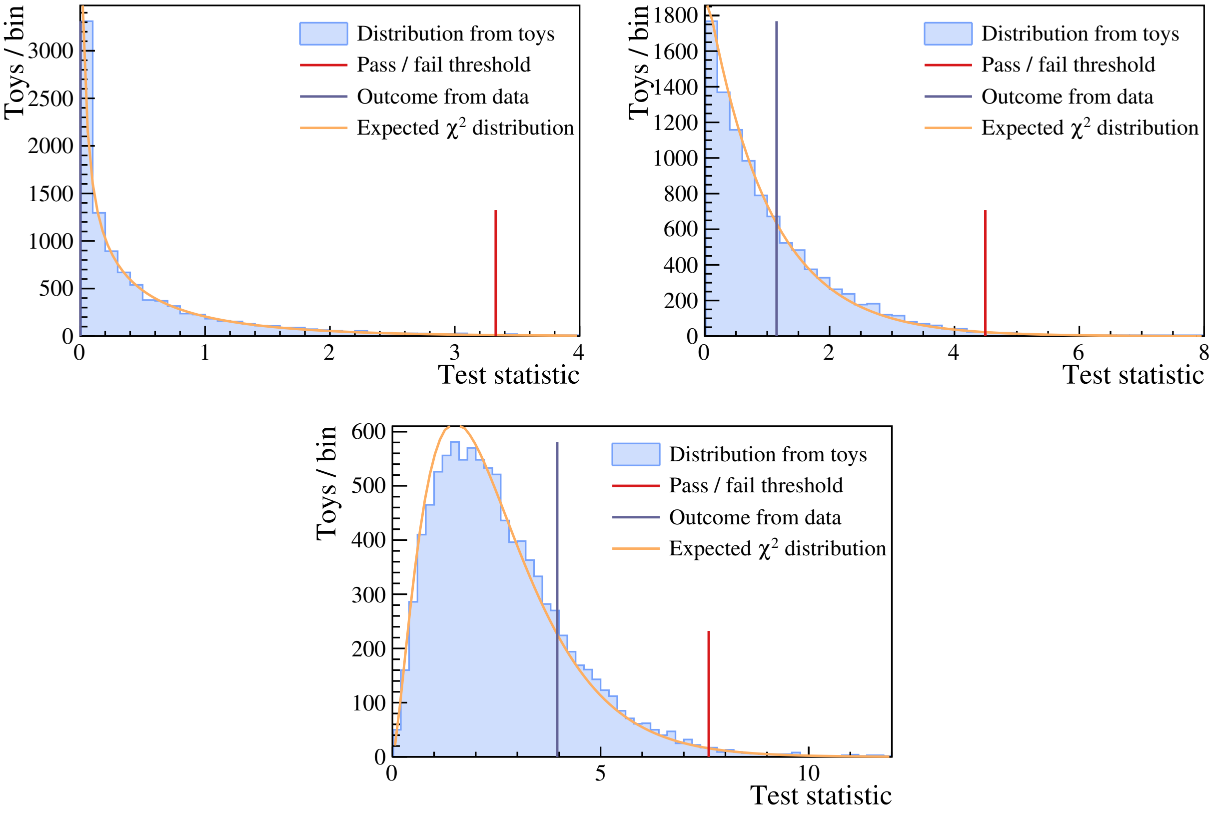

Figure G.5: Outcome of the "old vs. new" compatibility check, which involves splitting the data by data-taking periods (top left), electron trigger strategies (top right), and by both period and trigger (bottom). The test statistic for this check is defined in Equation (10.5). Its distribution from fits to toy datasets (light blue) is used to set the maximum allowed value for a success (red). In the limit of infinitely-many toys, the test statistic is expected to follow a χ^2 distribution (orange) whose number of degrees of freedom depends on the number of R_K values in the fits. The result obtained from the data (purple) is within the region for success in all three splittings.