University of Zurich[UZH]

November 4, 2019

Hunting for $B \to K\tau^+\tau^-$ imprints on the $B \to K\mu^+\mu^-$ dimuon spectrum

Master thesis of

Sascha Liechti

Supervised by

Prof. Nicola Serra

Dr. Patrick Owen

Abstract

The standard model is able to explain many phenomena in physics. However, it looks like it does not explain the full picture as there are still unsolved mysteries. Recently, so called anomalies concerning the sector of lepton flavour universality have been observed at the LHC and since then gained a lot of traction. They became the foundation of many interesting discussions and ideas, including new physics scenarios. One possible consequence of such new physics scenarios would be an enhancement of the ditau contributions in the dimuon spectrum of B decays. As the LHCb is able to measure muons to a very precise level, we wanted to analyse our sensitivity to the $B \to K\tau^+\tau^-$ imprint on the $B \to K\mu^+\mu^-$ spectrum to estimate how sensitive we are to new physics of this kind.

Contents

1 Introduction

1.1 Standard Model of particle physics

The Standard Model (SM) of particle physics (Fig. 1) describes all known elementary particles as well as three of the four fundamental forces. According to the SM there are fermions, which make up all matter in the universe and four gauge bosons enabling particles to interact with each other.

Fermions are further separated into quarks and leptons. Quarks are the only fermions able to interact through the strong force. All fermions can interact through the weak force and as long as they have electrical charge, also through the electromagnetic (em) force.

Fermions are further separated into three generations. The first generation of quarks consists of the up (u) and down (d) quarks while for the leptons it consists of the electron (e) and the corresponding electron neutrino (ν_e) having the same flavour but no electrical charge. Higher generations do have a similar structure but tend to have higher masses than the corresponding particles of lower generations (excluding neutrino masses). Here we also want to pay special attention to the b quark (b), which is the down type quark of the third generation. This particle plays a crucial role in the following analysis and is the heaviest quark that still hadronises.

Gauge bosons function as the carrier of the forces and hence get exchanged in interactions of particles. Additionally, there is the Higgs boson. Being a remnant of the Higgs field, which, through its interaction with particles, grants them the ability to acquire mass. Special attention is paid to the gauge bosons of the weak force, namely the W bosons ($W^{\pm}$) having an electrical charge of ± 1 and the Z boson (Z) being electrically neutral. It is important to note that, according to the SM, the only way to change flavour is through the exchange of W bosons.

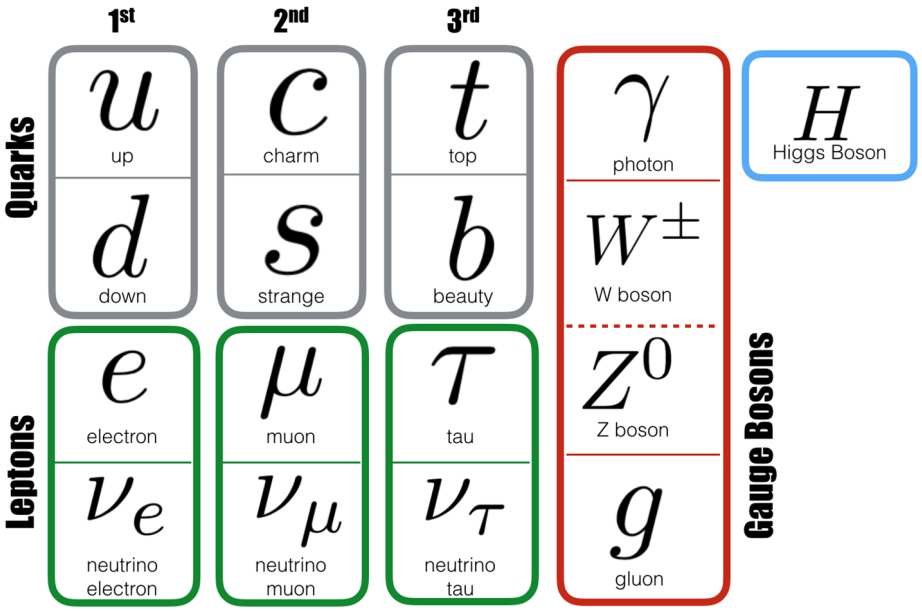


Figure 1: Standardmodel of particle physics

The SM also provides us with a way to describe the likelihood of one quark to change to another. As a consequence of the W boson having either ± 1 electrical charge, no tree level flavour changing neutral currents exist. As a consequence, an up-type quark can only change into a down-type quark and vice versa. We can quantitatively describe the likelihood of such changes in flavour to happen through the CKM matrix (see Fig. 2), where interactions within one generation are preferred over changes in-between generations.

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} \begin{bmatrix} \langle d|D_{init}\rangle \\ \langle s|D_{init}\rangle \\ \langle b|D_{init}\rangle \end{bmatrix} = \begin{bmatrix} \langle u|U_{end}\rangle \\ \langle c|U_{end}\rangle \\ \langle t|U_{end}\rangle \end{bmatrix} \tag{1}$$

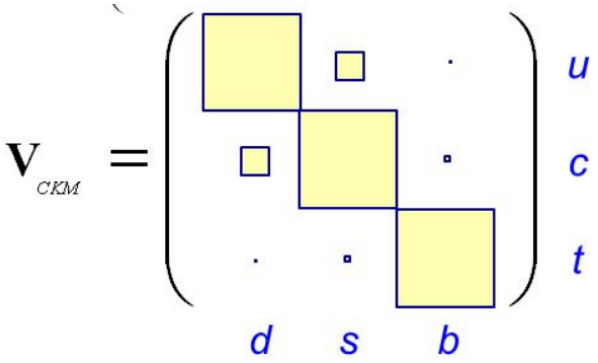


Figure 2: Respective size of entries of CKM matrix

Where Eq. (1) describes for example the transition from an initial state containing exclusively d quarks to a final state U_{end}. As we also demand unitarity for the CKM matrix we can apply additional conditions to its entries. Out of these conditions, physicists have constructed the unitary triangle, which has proven to be an extremely useful way to check several predictions of the SM. Additionally, the SM postulates lepton flavour universality (LFU), which means that all leptons have the exact same quantum numbers and only differentiate in their mass. We will have a deeper look at LFU in Sec. 1.2.

However, while the SM proves to be a robust and good description of the fundamental particles and their interactions, we know that the SM does not describe the full picture. Several phenomena are simply not explained by it. Famous examples are gravity[1], dark matter, dark energy, neutrino masses and the matter-antimatter asymmetry in the universe. Additionally, there are experimental results that cannot be explained by the SM as for example the $R_{K^{(*)}}$ [1–4] and $R_{D^{(*)}}$ [5–9] measurements, whose results point to a violation of LFU. With a significance of only 2-3σ though, they can neither prove nor disprove anything yet. Hence there is a major interest in particle physics to find new ways to test LFU.

1.2 Lepton flavour universality

As already stated, the SM postulates LFU. As a consequence, all interactions with leptons should have the same amplitude independent of the generation of the particle, with the

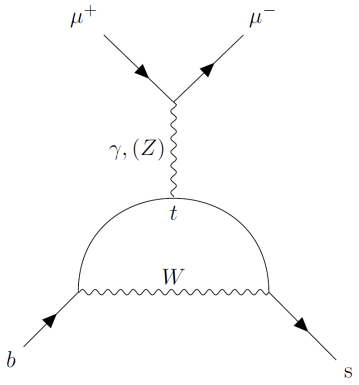
[1]The gravitational force is the one fundamental force not included in the SM and general relativity and the SM cannot be unified

only difference being their mass, which changes their respective phase space accordingly.

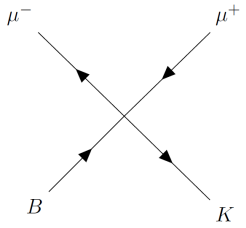
Therefore it comes to no surprise that in recent years the ratios of the semileptonic transitions R_{D^*} ($b \to s\ell\ell$) and R_{K^*} ($b \to c\ell\ell$) have gained a lot of attention. Here physicists have observed discrepancies between the observed and predicted SM values. These so called anomalies seem to point to some interesting new-physics (NP) scenarios, with possible connections to the SM flavor puzzle. Many of these NP models trying to explain these phenomena, propose dominant couplings to third-generation fermions, which should also appear in other semileptonic b-decays.

Hence, most NP models expect an enhancement of $b \to s\tau^+\tau^-$ transitions. Flavour changing neutral current (FCNC) decays with muons as well as electrons have both been observed inclusively and exclusively. However, probing rare decays with a $\tau^+\tau^-$ pair in the final state is currently experimentally very challenging [10, 11]. Here the experimental limits are still far away from corresponding SM predictions, leaving the NP expectations of large enhancements completely untested.

In this thesis we will deeper investigate the decay $B \to K\ell^+\ell^-$, which receives contributions from short-distance FCNC $b \to s\ell^+\ell^-$ contributions (see Fig. 3a) and long-distance contributions from intermediate resonances. As FCNC are forbidden at tree level in the SM, they can only occur via loop-level processes and are suppressed as a result. However, this opens the possibility to detect NP here, as many extensions of the SM have particles that can contribute to the amplitude of these processes and hence change the rate of distribution or the rate of decay of the final particles.



(a) Full $b \to s\ell\ell$ decay



(b) The $b \to s\ell\ell$ decay in its point-like representation used in effective field theory

Figure 3: The full decay in its point-like representation used in effective field theory

Now, let's introduce two additional steps. In the first step, for simplicity, we assume an effective field theory with Wilson coefficients. This only means that we approximate the short-distance interactions in Fig. 3a as a point-like interaction (see Fig. 3b). In the energy range we will be operating in, this has proven itself as a suitable approximation. As a result we encode all the dynamics and interactions in the Wilson coefficients[2]. The diagram in Fig. 3 then describes the non-resonant (NR) part of the decay rate with additional contributions entering through the Wilson coefficients.

The second step is the so called re-scattering. As discussed above, measuring muons in the final state is experimentally very feasible. Measuring taus in the final state though,

[2]Here specifically $\mathcal{C}_7$, $\mathcal{C}_9$ and $\mathcal{C}_{10}$

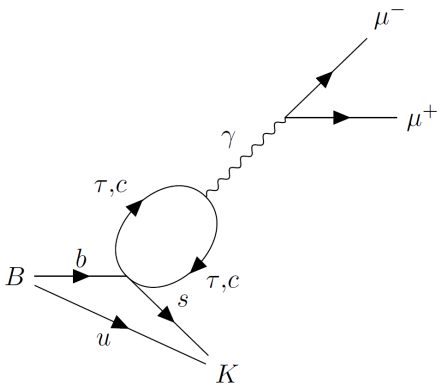


Figure 4: $B \to K\ell\ell$ decay with re-scattering (Fig. 3a represented as a point-like interaction)

it is experimentally extremely challenging. In order to test LFU we need to test both channels to be able to compare them to each other. This is where the second step comes into play. We now assume that in a first step we have taus instead of muons in the final state of Fig. 3a, but this final state is actually just an intermediate state where the taus annihilate again with each other. They can either produce a photon (γ)[3], Z boson (Z) or in the case of some NP models even new intermediate particles. This intermediate state can then decay again into another set of leptons and in the case of a pair of muons being created, we should be able to measure them in the original dimuon spectrum. As this is a long distance contribution, we do not approximate it as a point-like interaction. It is also important to note that the tau leptons do not necessarily need to be on-shell. The change from off-shell to on-shell actually induces some 'cusp'-like shape in the dimuon spectrum. We will discuss this 'cusp' in more depth in Sec. 2.4. We also expect a lot of interference of long-distance $\bar{c}c$-contributions as they could imitate the ditau signal due to their similar shape. We will further address this problem in Sec. 2.3.2.

Measuring the so called imprint of the ditau signal on the dimuon spectrum circumvents the experimental challenge of reconstructing the tau leptons in the final state. Although the price we pay is having less events overall as we have more vertices to reach the final state.

[3]dominant

2 Theoretical framework

2.1 Effective Hamiltonian

The dimension-six effective Lagrangian describing $b\to s\ell\ell$ transitions, renormalized at low energies [$\mu = O(m_b)$], can be decomposed as:

$$\mathcal{L}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_i \mathcal{C}_i(\mu)\mathcal{O}_i \tag{2}$$

Then, the leading FCNC effective operators can be described as:

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{16\pi^2}\,\mathrm{m}_b(\overline{s}\sigma_{\mu\nu}P_R b)F^{\mu\nu} \\ \mathcal{O}_9^l &= \frac{e}{16\pi^2}(\overline{s}\gamma_\mu P_L b)(\bar{l}\gamma^\mu l) \\ \mathcal{O}_{10}^l &= \frac{e}{16\pi^2}(\overline{s}\gamma_\mu P_L b)(\bar{l}\gamma^\mu\gamma_5 l) \end{aligned} \tag{3}$$

and the most relevant four-quark operators ($q = uc$) as:

$$\begin{aligned} \mathcal{O}_1^q &= (\overline{s}\gamma_\mu P_L q)(\overline{q}\gamma^\mu P_L b) \\ \mathcal{O}_2^q &= (\overline{s}^\alpha\gamma_\mu P_L q^\beta)(\overline{q}^\beta\gamma^\mu P_L b^\alpha) \end{aligned} \tag{4}$$

We should note that for the class of models we are considering, all relevant NP processes should be encoded in the Wilson coefficients $\mathcal{C}_{7,9,10}^l$. Using the normalization from Eq. (2) we see that $\mathcal{C}_{7,9,10}^l$ and $\mathcal{C}_{1,2}^c$ are real and $O(1)$ within the SM, while $\mathcal{C}_{1,2}^u = (V_{ub}V_{us}^*/V_{tb}V_{ts}) \times O(1)$ (see Ref. [12] for the values, derivations and basis of the Wilson coefficients).

We also note that the matrix elements of $\langle K^+\mu^+\mu^+|\mathcal{O}_i|B^+\rangle$ are only non-vanishing at tree level in the case of FCNC operators (with $\ell = \mu$). Considering only these structures for $B^+\to K^+\mu^+\mu^-$, this leads to the following equation for the decay rate:

$$\begin{aligned} \frac{d\Gamma}{dq^2} =& \frac{\alpha_{em}^2 G_F^2 |V_{tb}V_{ts}^*|}{128\pi^5}\kappa(q^2)\beta(q^2)\Bigg\{\frac{2}{3}\kappa^2(q^2)\beta^2(q^2)|\mathcal{C}_{10}^\mu f_+(q^2)|^2 + \frac{4m_\mu^2(m_B^2-m_K^2)^2}{q^2 m_B^2}|\mathcal{C}_{10}^\mu f_0(q^2)|^2 \\ &+ \kappa^2(q^2)\left[1-\frac{1}{3}\beta(q^2)\right]\left|\mathcal{C}_9^\mu f_+(q^2) + 2\mathcal{C}_7\frac{m_b+m_s}{m_B+m_K}f_T(q^2)\right|^2\Bigg\} \end{aligned} \tag{5}$$

Where $\kappa(q^2) = \lambda^{1/2}(m_B^2, m_K^2, q^2)/2m_B$ is the kaon momentum in the B^+ meson rest frame and $\beta(q^2) = 1 - 4m_\mu^2/q^2$. Furthermore $f_i(q^2)$ (with $i = +, 0, T$) are the vector, scalar and tensor $B\to K$ form factors [13]. [4]

[4]For simplification of Eq. (5): $\mathcal{C}_7 \Leftrightarrow$ EM dipole coupling operator strength, $\mathcal{C}_9 \Leftrightarrow$ vector coupling operator strength, $\mathcal{C}_{10} \Leftrightarrow$ axial-vector coupling operator strength

2.2 General considerations

To account for the non-local contributions generated by the non-leptonic operators in $\mathcal{L}_{eff}$, as well as the operator $\mathcal{O}_9^\tau$, we change $\mathcal{C}_9^\mu$ from Eq. (5) to a q^2 dependant function:

$$\mathcal{C}_9^\mu \to \mathcal{C}_9^{\mu,eff}(q^2) = \mathcal{C}_9^\mu + Y_{\bar{c}c}(q^2) + Y_{light}(q^2) + Y_{\bar{\tau}\tau}(q^2) \tag{6}$$

Looking at Eq. (6), $Y_{\mathcal{I}}(q^2)$ encodes the non-local contributions of the intermediate state $\mathcal{I}$, which afterwards annihilates into a $\mu^+\mu^-$ pair via a single photon exchange. Additionally, we need to consider that the functions $Y_{\bar{c}c}(q^2)$ and $Y_{light}(q^2)$ describe non-perturbative hadronic contributions. As a result, they cannot be reliably estimated in perturbation theory for a large fraction of the accessible q^2 spectrum.

Figure 5: Diagrammatic long-distance contributions to $\mathcal{C}_9^{eff}$. On the left is a single vector resonance and on the right a contribution from a two-particle intermediate state

Our main strategy to solve this issue is to write the non-perturbative contributions $Y_{\bar{c}c}(q^2)$ and $Y_{light}(q^2)$ using hadronic dispersion relations. This means, that for one- (1P) and two-particle (2P) intermediate states (see Fig. 5), we use dispersion relations subtracted at $q^2 = 0$. For the subleading $Y_{light}(q^2)$ function we consider only one-particle intermediate states and use unsubtracted dispersion relations.
Using the conditions stated above, we finally arrive at the following form:

$$\mathcal{C}_9^{\mu,eff}(q^2) = \mathcal{C}_9^\mu + Y_{\bar{c}c}^{(0)} + Y_{\bar{c}c}^{1P}(q^2) + Y_{\bar{c}c}^{2P}(q^2) + Y_{light}^{1P}(q^2) + Y_{\bar{\tau}\tau}(q^2) \tag{7}$$

where $\Delta Y_{\bar{c}c}^{1P}(0) + \Delta Y_{\bar{c}c}^{2P}(0) = 0$.
We will further explain the different contributions in Sec. 2.3, 2.4.

2.3 Long distance hadronic contributions

We find that the general structure of the subtracted dispersion relation used to determine $\Delta Y_{\bar{c}c}(q^2)$ is:

$$\Delta Y_{\bar{c}c}(q^2) = 16\pi q^2 \int_{m_{J/\Psi}^2}^{\infty} ds \frac{1}{s(s-q^2)} Im\left[\frac{\mathcal{H}_{\bar{c}c}^{(BK)}(s)}{f_+(s)}\right] \equiv \frac{q^2}{\pi}\int_{m_{J/\Psi}^2}^{\infty} ds \frac{\rho_{\bar{c}c}(s)}{s(s-q^2)} \tag{8}$$

The function $\rho_{\bar{c}c}(s)$ is the spectral-density function describing the hadronic states $\mathcal{I}_{\bar{c}c}$, characterized by $\bar{c}c$ valence quarks and the invariant mass s. These then contribute as real intermediate states to the re-scattering $B \to K\mathcal{I}_{\bar{c}c} \to K\mu^+\mu^-$. As already mentioned before, we decompose $\rho_{\bar{c}c}(s)$ into one- and two-particle intermediate states:

$$\rho_{\bar{c}c}(s) = \rho_{\bar{c}c}^{1P}(s) + \rho_{\bar{c}c}^{2P}(s) \tag{9}$$

with:

$$\rho^{1P}_{\bar{c}c}(s) \propto \sum_j \mathcal{A}(B \to K V^0_j)\delta(s - m_j) \tag{10}$$

$$\rho^{2P}_{\bar{c}c}(s) \propto \sum_j \int dp_j^2 \delta(s - p_j^2) \int \frac{d^3\vec{p}_{j_1} d^3\vec{p}_{j_2}}{16\pi^2 E_{j_1} E_{j_2}} \mathcal{A}(B \to K M^+_{j_1} M^-_{j_2})\delta^4(p_j - p_{j_1} - p_{j_2}) \tag{11}$$

where we are neglecting phase-space suppressed contributions with three or more particles.

2.3.1 Charmonium resonances

For simplicity, the single-particle states in Eq. (10) were treated as infinitely narrow resonances. To implement the effect of finite width we use Breit-Wigner functions:

$$\Delta Y^{1P}_{\bar{c}c}(q^2) = \sum_{j=\Psi(1S),\ldots,\Psi(4415)} \eta_j e^{i\delta_j} \frac{q^2}{m_j^2} A^{res}_j(q^2) \tag{12}$$

with:

$$A^{res}_j(s) = \frac{m_j \Gamma_j}{(m_j^2 - s) - i m_j \Gamma_j}$$

We should note that the sum runs over all the charmonium vector resonances in the accessible kinematical range. Also η_j and δ_j are real parameters, which have to be determined from data. Here we used the values from the analysis performed by LHCb [14]. We can even directly correspond the fitted η_j one-to-one with the product of $B^+ \to K^+ V^0_j$ and $V^0_j \to \mu^+\mu^-$ branching fractions via:

$$\begin{aligned}\mathcal{B}(B^+ \to K^+ V^0_j) \times \mathcal{B}(V^0_j \to \mu^+\mu^-) = \tau_{B^+} \frac{G_F^2 \alpha^2 |V_{tb} V^*_{ts}|^2}{128\pi^5} \int_{4m_\mu^2}^{(m_B - m_K)^2} dq^2 \kappa(q^2)^3 \\ \times \left[\beta(q^2)\ \frac{1}{3}\beta(q^2)^3\right] |f_+(q^2)|^2 |\eta_j|^2 \left| \frac{q^2}{m_j^2} A^{res}_j(q^2) \right|^2\end{aligned} \tag{13}$$

We should note that Eq. (12) differs from the decomposition used in [14] by the q^2/m_j^2 term. It comes from the subtraction procedure in the dispersion relation. This is necessary to ensure the convergence of the integral in the two-particle intermediate states, which will be discussed in more detail in the next subsection (2.3.2). Additionally, choosing the subtraction point at $q^2 = 0$ allows us to decouple the determination of the resonance parameters of the spectrum from the overall normalization rate. This results in an extra term, which is constant and undetermined, $Y^{(0)}_{\bar{c}c} = Y_{\bar{c}c}(0)$ in Eq. 7. It plays no role in the description of the dimuon spectrum. However, it is relevant to determine the value of $\mathcal{C}^\mu_9$.

Using the estimate of Ref. [12], based on a Λ^2/m_c^2, while also taking into account next-to-leading $\mathcal{O}(\alpha_s)$ corrections on the pure partonic result, gives us:

$$Y^{(0)}_{\bar{c}c} \approx -0.10 \pm 0.05 \tag{14}$$

which is around $-(2 \pm 1)\%$ of $\mathcal{C}^{\mu,SM}_9 \approx 4.23$

2.3.2 Two-particle intermediate states

Similarly as in Sec. 2.3.1 we get for the two-particle contributions:

$$\Delta Y_{\bar{c}c}^{2P}(q^2) = \sum_j \eta_j e^{i\delta_j} A_j^{2P}(q^2) \tag{15}$$

with

$$A_j^{2P}(q^2) = \frac{q^2}{\pi} \int_{s_0^j}^{\infty} \frac{ds}{s} \frac{\hat{\rho}(s)}{(s-q^2)} \tag{16}$$

$\hat{\rho}(s)$ stands for the normalised spectral densities of the two-body intermediate states, which are characterised by the threshold $s_0^j = (m_{j_1} m_{j_2})^2$.

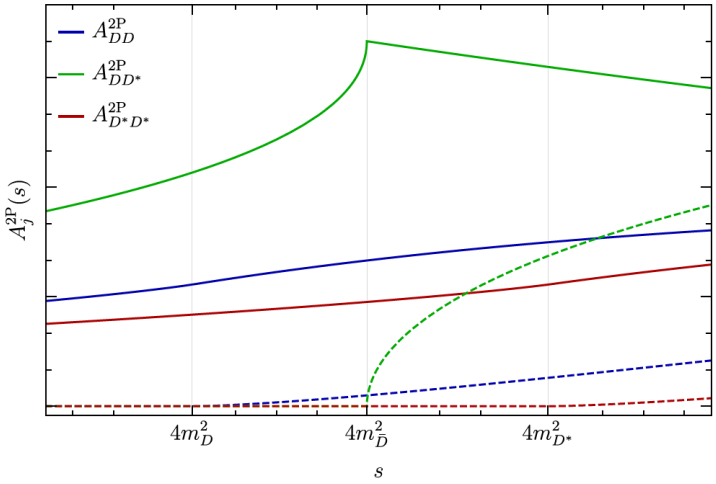


Figure 6: Real (solid) and imaginary (dashed) parts of the normalised hadronic two-particle contributions to $Y_{\bar{c}c}(q^2)$, as defined in Eq. (15, 16)

We do not have a precise estimate of these spectral densities at generic kinematical points, however, we can get a precise description of their behaviour around the respective thresholds by approximating them with powers of the Källén function[5] (see Fig. 6). We then obtain the following estimates for the normalised spectral densities of the two-particle intermediate states of lowest mass:

$$\begin{aligned}
\hat{\rho}_{DD}(s) &= \left(1 - \frac{4m_D^2}{s}\right)^{\frac{3}{2}} \\
\hat{\rho}_{D^*D^*}(s) &= \left(1 - \frac{4m_{D^*}^2}{s}\right)^{\frac{3}{2}} \\
\hat{\rho}_{DD^*}(s) &= \left(1 - \frac{4m_{\bar{D}}^2}{s}\right)^{\frac{1}{2}}
\end{aligned} \tag{17}$$

[5]The exponent is determined by the lowest partial wave allowed in the $B \to M_1 M_2 K \to K\mu^+\mu^-$ re-scattering. This is because higher-order partial waves, characterized by higher orders of the Källén function, are phase space suppressed and give rise to a less singular behavior around the threshold.

For the DD^* intermediate state, we replaced the complete expression depending on both masses with a simplified one depending only on $m_{\overline{D}} = (m_D + m_{D^*})/2$. Using the estimates above we find:

$$\Delta Y^{2P}_{\bar{c}c}(q^2) = \eta_{\overline{D}} e^{i\delta_{\overline{D}}} h_S(m_{\overline{D}}, q^2) + \sum_{j=D,D^*} \eta_j e^{i\delta_j} h_P(m_j, q^2) \tag{18}$$

with:

$$\begin{aligned} h_P(m, q^2) &= \frac{2}{3} + \left(1 - \frac{4m^2}{q^2}\right) h_S(m, q^2) \\ h_S(m, q^2) &= 2 - G\left(1 - \frac{4m^2}{q^2}\right) \end{aligned} \tag{19}$$

and

$$G(y) = \sqrt{|y|} \left\{ \Theta(y) \left[\ln\left(\frac{1+\sqrt{y}}{1-\sqrt{y}}\right) - i\pi \right] + 2\Theta(-y) \arctan\left(\frac{1}{\sqrt{-y}}\right) \right\} \tag{20}$$

We should note that the lowest threshold is at $q^2 = 4m_D^2$. However, the contribution from the DD^* intermediate state is the only one which can occur in an S-Wave state, giving rise to a singular behaviour[6] at the threshold (see Fig. 6).

2.3.3 Light resonances

The Wilson coefficients of the effective operators appearing in $\mathcal{H}^{(BK)}_{light}(q^2)$ are either loop- or CKM-suppressed. As a consequence, we can limit ourselves to only one-particle hadronic intermediate states. As we fit the hadronic coefficients η_j from the data, additional parameters, like the ones describing transitions to valence charm quarks, get absorbed in $\Delta Y_{\bar{c}c}(q^2)$. What is left are only vector resonances containing light valence quarks.

Additionally, we can restrict ourselves even further to only consider the ρ, ω and ϕ states, as the leptonic decay rates of heavier states are very small.

Contrary to our treatment of the charm-quark contributions, we describe the intermediate states employing unsubtracted dispersion relations for two reasons. Firstly, since we are considering only single-particle states the convergence of the dispersive integrals does not pose a problem. Secondly, a subtraction at $q^2 = 0$ simply does not prove to be particularly beneficial, as the light-quark contributions are truly no-perturbative in this regime. This leads us to:

$$Y^{1P}_{light}(q^2) = \sum_{j=\rho,\omega,\phi} \eta_j e^{i\delta_j} A^{res}_j(q^2) \tag{21}$$

In analogy to EQ. (15, 16).

[6] Square root like behavior

2.3.4 Theoretical constraints on the hadronic parameters

The hadronic decompositions in Eq. (12, 15 and 21) contain 12 free complex parameters: 6 in $\Delta Y^{1P}_{\bar{c}c}(q^2)$, 3 in $\Delta Y^{2P}_{\bar{c}c}(q^2)$ and 3 in $\Delta Y^{1P}_{light}(q^2)$. In theory, all of these parameters should be fitted from data. However, as they all correspond to different functional forms, an unconstrained fit would in practice leave significant degeneracies in the parameter space. In the following we will introduce three conservative constraints, which can be imposed using perturbative arguments.

I. *Constraint on the slope of* $\Delta Y_{\bar{c}c}(q^2)$ *at* $q^2 = 0$

The lowest-order perturbative estimate of $\Delta Y_{\bar{c}c}(q^2)$ is obtained by factorising the matrix element $\langle K(p)|\bar{s}\gamma^\mu b|B(p+q)\rangle$ and computing the charm-loop at $\mathcal{O}(\alpha_s^0)$:

$$\begin{aligned}\Delta Y^{pert}_{\bar{c}c}(q^2) &= 2\left(\mathcal{C}_2 + \frac{1}{3}\mathcal{C}_1\right) \times Q_c \times q^2 \int_{4m_c^2}^{\infty} ds \frac{\sqrt{1-\frac{4m_c^2}{s}}\left(1+\frac{2m_c^2}{s}\right)}{s(s-q^2)} \\ &= 2\left(\mathcal{C}_2 + \frac{1}{3}\mathcal{C}_1\right)\left[h_S(m_c,q^2) - \frac{1}{3}h_P(m_c,q^2)\right]\end{aligned} \tag{22}$$

This is expected to be a reasonable approximation at $q^2 \approx 0$ up to $\mathcal{O}(\Lambda_{QCD}/m_c^2)$ corrections, but is not a good approximation close to the resonance region. However, we can use this to set bounds on the slope of $\Delta Y_{\bar{c}c}(q^2)$ in the vicinity of $q^2 = 0$. The perturbative result is then:

$$\left.\frac{d}{dq^2}\Delta Y^{pert}_{\bar{c}c}(q^2)\right|_{q^2=0} = \frac{4}{15}\left(\mathcal{C}_2 + \frac{1}{3}\mathcal{C}_1\right)\frac{1}{m_c^2} \approx (1.7 \pm 1.7) \times 10^{-2}\,\text{GeV}^{-2} \tag{23}$$

where the numerical value has been obtained setting $m_b/2 < \mu < 2m_b$ and $m_c = 1.3\,\text{GeV}$. According to the analysis of Ref. [12] the inclusion of $\mathcal{O}(\Lambda_{QCD}/m_c^2, \alpha_s)$ corrections[7] modifies the above prediction to $-(0.5 \pm 0.2) \times 10^{-2}\,\text{GeV}^{-2}$. This leads us to the following constraints:

$$\begin{aligned} &Re\left[\sum_{j=\Psi(1S),\ldots} \eta_j e^{i\delta_j}\frac{\Gamma_j}{m_j^3} + \eta_{\overline{D}} e^{i\delta_j}\frac{1}{6m_{\overline{D}}^2}\sum_{j=D,D^*}\eta_j e^{i\delta_j}\frac{1}{10m_j^2}\right] = (1.7 \pm 2.2) \times 10^{-2}\,\text{GeV}^{-2} \\ &\left|\sum_{j=\Psi(1S),\ldots} \eta_j e^{i\delta_j}\frac{\Gamma_j}{m_j^3} + \eta_{\overline{D}} e^{i\delta_j}\frac{1}{6m_{\overline{D}}^2}\sum_{j=D,D^*}\eta_j e^{i\delta_j}\frac{1}{10m_j^2}\right| \leq 5 \times 10^{-2}\,\text{GeV}^{-2}\end{aligned} \tag{24}$$

where we slightly enlarged the error from Eq. (23), such that the 1σ range covers the difference between the central value in Eq. (23) and the one including $\mathcal{O}(\Delta_{QCD}/m_c^2, \alpha_s)$, corrections estimated in [12].

II. *Upper bound on the* $|\eta_j|$ *in* $\Delta Y^{2P}_{\bar{c}c}(q^2)$

Comparing the perturbative result with $\Delta Y^{2P}_{\bar{c}c}(q^2)$ also allows us to define the natural range for the $\eta_{\overline{D},D,D^*}$ parameters, which are poorly constrained by data. If we focus on

[7] Which involve new hadronic matrix elements

the leading S-wave contribution and set $\eta_{\overline{D}} = 2(\mathcal{C}_2 + \mathcal{C}_1/3) \approx (0.2 \pm 0.2)$ in the limit $m_c \to m_{\overline{D}}$, it turns out that the perturbative quark loop can be saturated by the leading DD^* meson loop. On general grounds, each of the exclusive meson contributions should be significantly smaller than the inclusive quark contribution. As a result, we set an upper limit of:

$$|\eta_{\overline{D},D,D^*}| \leq 0.2 \tag{25}$$

III. *Upper bound on* $|Y^{1P}_{light}(q^2 = 0)|$

As we used an unsubtracted dispersion relation in the case of $Y^{1P}_{light}(q^2)$, $Y^{1P}_{light}(q^2 = 0)$ is non-vanishing. At the same time, we know that the Wilson coefficients entering Y^{1P}_{light} are very suppressed. The suppression either comes from loop factors or from subleading CKM factors. By simple dimensional arguments we expect $|Y^{1P}_{light}(0)| \ll |Y^{(0)}_{\bar{c}c}|$. As a result, we set the conservative bound:

$$|Y^{1P}_{light}(0)| \approx \left| \sum_{j=\rho,\omega,\phi} \eta_j \frac{\Gamma_j}{m_j} \right| \leq 0.02 \tag{26}$$

This corresponds to around 20% of the perturbative charm contribution at $q^2 = 0$.

2.4 Tau lepton contribution

We can directly compute the contribution from intermediate τ leptons in perturbation theory:

$$Y_{\tau\bar{\tau}}(q^2) = -\frac{\alpha_{em}}{2\pi}\mathcal{C}_9^\tau \left[h_S(m_\tau, q^2) - \frac{1}{3} h_P(m_\tau, q^2) \right], \alpha_{em} \approx \frac{1}{128} \tag{27}$$

with the functions h_L from Eq. (19). The functional form is identical to the one of the perturbative charm contributions as well as the DD^* meson contribution shown in Fig. 6. Nonetheless, the cusp is well separated from the various hadronic contributions, as it is located at $q^2 = 4m_\tau^2$. Theoretically, the short-distance amplitude does not need to be controlled by the CKM matrix. However, in most realistic NP frameworks the weak phases of all $b \to s\ell^+\ell^-$ amplitudes are aligned to the SM one. This implies $Im(\mathcal{C}_9^\tau) = Im(\mathcal{C}_9^\mu) = 0$. In the following we use this simplifying assumption.

We can derive a model-independent estimate of the maximal allowed size of $|\mathcal{C}_9^\tau|$ can be derived from the experimental upper bound on $\mathcal{B}(B^+ \to K^+\tau^+\tau^-) < 2.25 \times 10^{-3}$ by BaBar [15]. Note that this is more than four orders of magnitude bigger than the SM prediction of $\mathcal{B}(B^+ \to K^+\tau^+\tau^-)_{SM} \approx 1.2 \times 10^{-7}$. Neglecting the contributions from operators other than $\mathcal{O}_9^\tau$ and $\mathcal{O}_{10}^\tau$ and assuming $\mathcal{C}_9^\tau = -\mathcal{C}_{10}^\tau$, we find:

$$|\mathcal{C}_9^\tau| \approx \left| \frac{\mathcal{B}(B^+ \to K^+\tau^+\tau^-)}{6.7 \times 10^{-9}} \right|^{1/2} \lesssim 580 \tag{28}$$

This is then to be compared to the SM prediction of $\mathcal{C}_9^{\tau,SM} \approx 4.2$. However, calculating a BR for the ditau contribution as it is done in Eq. (28) should be done with caution as by translating $\mathcal{C}_9^\tau$ into a BR we introduce some model dependence.

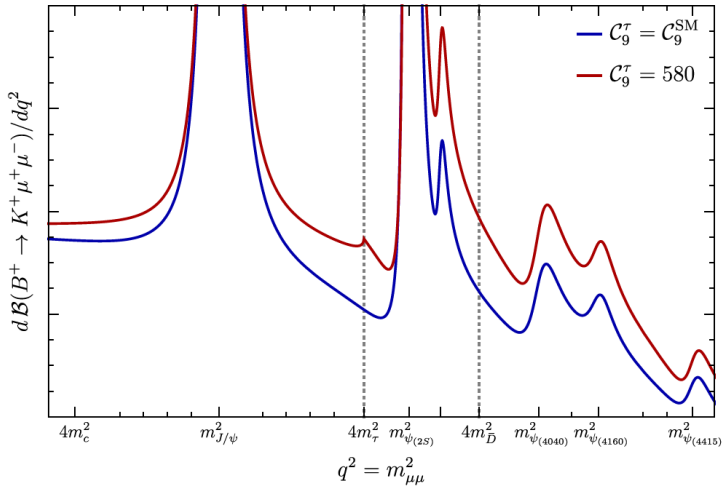


Figure 7: Illustration of the maximal impact of the $B^+ \to K^+\tau^+\tau^- \to K^+\mu^+\mu^-$ re-scattering on the $B^+ \to K^+\mu^+\mu^-$ dimuon spectrum, taking into account the present experimental upper limit on $\mathcal{B}(B^+ \to K^+\tau^+\tau^-)$.

In Fig. 7 we show the potential impact on the $B^+ \to K^+\tau^+\tau^- \to K^+\mu^+\mu^-$ re-scattering on the $B^+ \to K^+\mu^+\mu^-$ dimuon spectrum using the maximal value of $\mathcal{C}_9^\tau$ in Eq. (28), the resonance parameters determined in Ref. [16] and neglecting two-particle contributions. We see that for the extreme choice of the maximum values of the parameters, the effect is very well visible. This clearly illustrates the possibility of extracting of a significant upper bound $(B^+ \to K^+\tau^+\tau^-)$ from a detailed analysis of the dimuon spectrum.

3 Analysis of the expected sensitivity at LHCb

3.1 Toys

As our goal was to generate single points generated from Eq. (5), we implemented the function as tensorflow function [17]. This enabled us to use the zfit library [18]. Additionally, having the tensorflow framework in the back-end gives us extreme boosts in speed, that made the fitting of possible future scenarios with more data even feasible within a reasonable time-frame, which will be further discussed in Sec. 3.2.7.

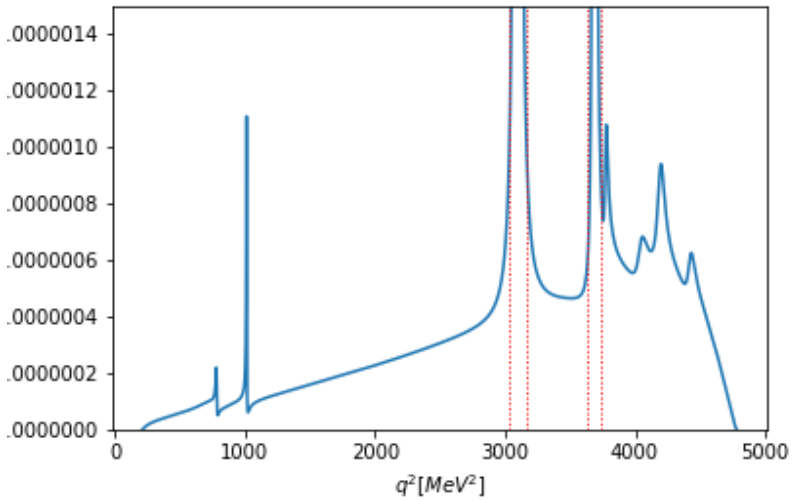


Figure 8: Cuts on the q^2-region showing the five regions

As a next step we split the q^2 spectrum into five parts, which can be seen in Fig. 8. We will further elaborate our choice of regions in Sec. 3.2.3. For now we should know that we solely generated particles in the low-q^2 region, in the region in between the J/Ψ and $\Psi(2S)$ resonances and in the high-q^2 region, so outside of the the second (J/Ψ resonance) and the fourth region($\Psi(2S)$ resonance).

According to LHCb, we expect around 5.4 million events in the whole q^2 region. However, now that we exclude the J/Ψ region and the $\Psi(2S)$ region, we are only left with around 37′000 events. This is due to the fact that the J/Ψ and $\Psi(2S)$ resonances are by around 4 and 3 order of magnitudes bigger than the other resonances (see branching ratios (BR) in table below and Fig. 9).

Resonance	BR
ρ	1.7×10^{-10}
ω	4.9×10^{-10}
ϕ	2.5×10^{-9}
J/Ψ	6.02×10^{-5}
$\Psi(2S)$	4.97×10^{-6}
$\Psi(3770)$	1.38×10^{-9}
$\Psi(4040)$	4.2×10^{-10}
$\Psi(4160)$	2.6×10^{-9}
$\Psi(4415)$	6.1×10^{-10}
NR	4.37×10^{-7}

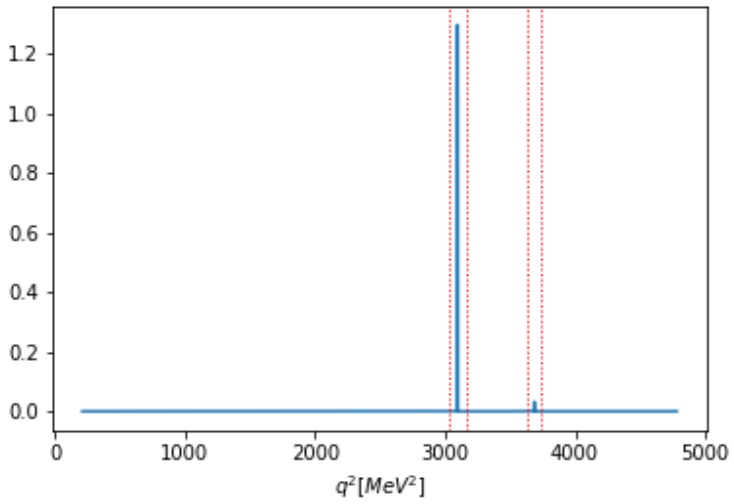


Figure 9: Difference in order of magnitude of the J/Ψ (left) and $\Psi(2S)$ (right) to the rest of the decay rate

Note that, even though the NR part has a relatively big BR, it has still a relatively small amplitude. This is explained by the fact that the NR spans over the whole q^2 region, while the resonances are much narrower, which then leads to a lower amplitude overall.

3.2 Fitting procedure

3.2.1 Fitting framework

For the fitting we chose to use the zfit library [18]. Using the tensorflow framework in the back-end, we can make full use of the speed gain that GPU's provide for calculations like fitting, which can easily reach a few orders of magnitude. Additionally, having the whole mathematical structure available in the form of graphs proves to be useful for the fitting procedure as well. We also opted to fit with a negative log-likelihood (NLL). For the minimization, zfit uses the iminuit and MIGRAD libraries.

Before every fit we randomised the variables with 5% around the original value that we generated the data from. With the purpose to introduce some randomness into the fit,

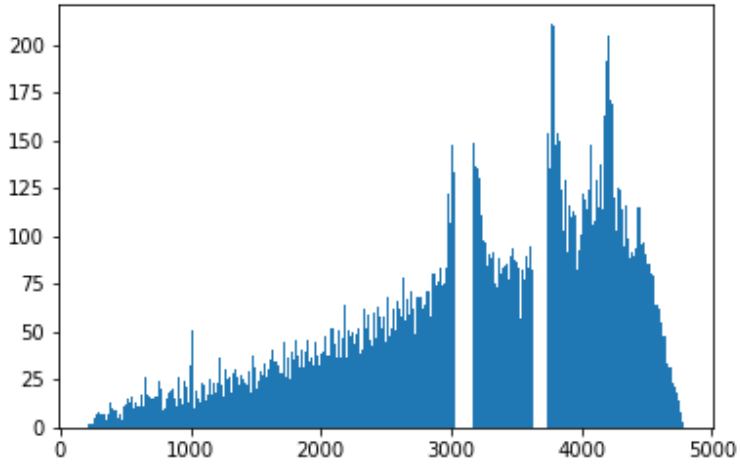


Figure 10: Example toy consisting of around 37′000 events

as for the fit with real data we cannot expect to start on the true values as well, which should give us a more realistic fitting over all.

3.2.2 Parameters

For the resonances we used the following values taken from Ref. [14]. It is also important to note, that in [14] there are four different possible variations of phases for the J/Ψ and $\Psi(2S)$. Both can either have a positive or a negative phase, giving us four possible combinations, each with its corresponding set of values for the remaining variables of the resonances. For our analysis we chose the combination where the phase of the J/Ψ resonance takes a negative value and the phase of the $\Psi(2S)$ takes a positive value. Note, that none of the combinations is preferred as all of them are equally good, which will further elaborated in Sec. 3.2.5. The table below shows the values of the resonances used for the fit and hence to generate toys.

Resonance	Mass m [MeV]	Amplitude η [MeV]	Width Γ [MeV]	Phase δ [rad]
$\rho(770)$	743.2	1.05	149.0	-0.3
$\omega(782)$	782.7	8.5	6.8	0.3
$\phi(1020)$	1013.5	19.2	4.25	-1.5
J/Ψ	3096.1	9897.0	0.09	-1.5
$\Psi(2S)$	3686.0	1396.0	0.3	2.08
$\Psi(3770)$	3773.0	2.5	27.2	-2.89
$\Psi(4040)$	4039.0	1.01	80.0	-2.69
$\Psi(4160)$	4191.0	2.2	70.0	-2.13
$\Psi(4416)$	4421.0	1.24	62	-2.43

Out of these variables we were floating all the phases δ and the amplitudes η, as these were the ones to have the biggest impact on our sensitivity. The mass (center) of the resonances as well as the width have already been determined to a high precision by

numerous experiments (like [14]), so we defined them to be fixed. To accurately describe the interplay of the resonances with the D contributions as well as the ditau contribution, we were floating the phases and amplitudes as they directly interact with each other. Also we wanted to provide a complete fit of the full spectrum to have a full picture of the spectrum.

However, as we were excluding the J/Ψ and $\Psi(2S)$ regions we have no means to accurately determine their amplitude. As a consequence, we fixed the amplitude of these two resonances and, as it will be more difficult to find the phases of the two resonances as well, we will implement some additional constraints on them, which will be discussed in more detail in Sec. 3.2.4. As the cusp is located in between the J/Ψ and $\Psi(2S)$ resonances and as their amplitudes are extremely large in comparison, we expect their phases to have a lot of impact on the cusp region. With the constraints on the phases in place we can float the phases, while still expecting to have a reasonable result.

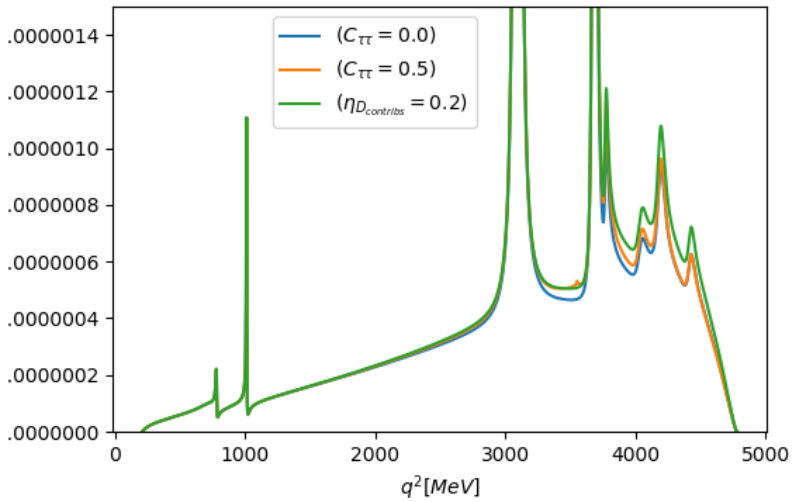


Figure 11: Comparison of the decay rate with no signal to a signal of the ditau contribution and the D contributions

If we now move on to the D contributions described in Eq. (29), we see that they also each have a phase and an amplitude associated with them. As the mass of these contributions is quite well known, we assume these as fixed, while simultaneously floating their amplitudes and phases. This is crucial, as the D contributions provide a signal being able to imitate signal of ditau contribution fairly well (see Fig. 11). Looking at Eq. (29), we made one simplification for the fitting. Namely we merged the D and the D^* contributions to one, as their shapes were extremely similar, which made it very difficult to be differentiated by the fit. This resulted in a change of $\Delta Y_{\bar{c}c}^{2P}$ to:

$$\Delta Y_{\bar{c}c}^{2P}(q^2) = \eta_{\overline{D}} e^{i\delta_{\overline{D}}} h_S(m_{\overline{D}}, q^2) + \eta_{D,D^*} e^{i\delta_{D,D^*}} h_P(m_{D,D^*}, q^2) \tag{29}$$

with:

$$m_{D,D^*} = \frac{m_D + m_{D^*}}{2} \tag{30}$$

When generating toys, we always assumed the phases and the amplitudes of these contributions to be 0. However, the system still had the freedom to find the optimal minimum of the system. This also gave rise to some peculiar solutions for the minimum with nonzero amplitudes for the D contributions, which will be further discussed in Sec. 3.2.8 and will have to be addressed for the final fit of the LHCb data.

For the form factors (FF) we used the parametrisation presented in [13]. We can then describe the FF like:

$$\begin{aligned} f_+(q^2) &= \frac{1}{P_+(q^2)} \sum_{m=0}^{K-1} b_m^+ \left[z^m - (-1)^{m-K} \frac{m}{K} z^K \right] \\ f_0(q^2) &= \frac{1}{P_0(q^2)} b_m^0 z^m \\ f_T(q^2) &= \frac{1}{P_T(q^2)} \sum_{m=0}^{K-1} b_m^T \left[z^m - (-1)^{m-K} \frac{m}{K} z^K \right] \end{aligned} \tag{31}$$

with:

$$\begin{aligned} z(q^2, t_0) &= \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}} \\ t_+ &= (M_B + M_K)^2 \\ t_0 &= (M_B + M_K)(\sqrt{M_B} - \sqrt{M_K})^2 \\ P_{+,0,T}(q^2) &= 1 - \frac{q^2}{M^2} \text{with:} \begin{cases} 0, & M = M_{B_0^*} = 5711.0\,\text{MeV} \\ +, T, & M = M_{B^*} = 5415.4\,\text{MeV} \end{cases} \end{aligned} \tag{32}$$

Looking at Eq. (31), we set $K = 3$ and then were floating the parameters b_m^+, because they were the one connected to $\mathcal{C}_9$ (see Eq. (5)). As our starting values we used the values stated in the same paper:

b_0^+	b_1^+	b_2^+	b_0^0	b_1^0	b_2^0	b_0^T	b_1^T	b_2^T
0.466	-0.885	-0.213	0.292	0.281	0.150	0.460	-1.089	-1.114

3.2.3 Regions

As already mentioned above, we split the q^2 region into 5 parts (see Fig. 8).

The first part is the low-q^2 region. We define it as the q^2 region below the J/Ψ resonance. In this region we find the light-resonances ρ, ω and ϕ that have already been discussed in Sec. 2.3.3. Even though the cusp-like shape of the ditau enhancement does not lie in this region, it still provides a lot of usefulness in giving us bounds on the imaginary tails of the cusp as well as generally providing more control on the shape of the decay rate.

The second region consists of the J/Ψ-resonance itself. We chose to not include this region in the final fit. There were two main factors that lead to this decision. The first one was to compensate for resolution, meaning that, as we assume infinite precision for our toys, we compensate this by excluding the J/Ψ and $\Psi(2S)$ resonances, which are the

biggest contributions by far. The second factor was more a consideration of the time constraints to finish the whole fit. With the number of events discussed in Sec. 3.1, we can comfortably finish a full analysis of the current scenario within a few days. However, increasing the number of events from around 37′000 to 5.4 million per toy also drastically increases the computational time needed.

The third region is the region in between the two big resonances. It is of great importance, as this is the region where we expect the cusp of the ditau enhancement to appear.

We also chose to exclude the fourth ($\Psi(2S)$) region following similar arguments that lead to our exclusion of the J/Ψ region.

Similar to the first region, the fifth (high-q^2) region provides us with more control over the shape of the decay rate. It mainly includes the NR part as well as higher excited states of the Ψ meson.

3.2.4 Constraints

For completeness we will briefly list the constraints listed in Sec. 3.2.4, before we will move on to the additional constraints used in the fit:

I. *Constraint on the slope of* $\Delta Y_{\bar{c}c}(q^2)$ *at* $q^2 = 0$

$$Re\left[\sum_{j=\Psi(1S),\dots} \eta_j e^{i\delta_j}\frac{\Gamma_j}{m_j^3} + \eta_{\overline{D}} e^{i\delta_j}\frac{1}{6m_{\overline{D}}^2}\sum_{j=D,D^*}\eta_j e^{i\delta_j}\frac{1}{10m_j^2}\right] = (1.7 \pm 2.2)\times 10^{-2}\,\mathrm{GeV}^{-2}$$

$$\left|\sum_{j=\Psi(1S),\dots} \eta_j e^{i\delta_j}\frac{\Gamma_j}{m_j^3} + \eta_{\overline{D}} e^{i\delta_j}\frac{1}{6m_{\overline{D}}^2}\sum_{j=D,D^*}\eta_j e^{i\delta_j}\frac{1}{10m_j^2}\right| \leq 5\times 10^{-2}\,\mathrm{GeV}^{-2} \tag{33}$$

II. *Upper bound on the* $|\eta_j|$ *in* $\Delta Y_{\bar{c}c}^{2P}(q^2)$

$$|\eta_{\overline{D},D,D^*}| \leq 0.2 \tag{34}$$

III. *Upper bound on* $|Y_{light}^{1P}(q^2 = 0)|$

$$|Y_{light}^{1P}(0)| \approx \left|\sum_{j=\rho,\omega,\phi}\eta_j\frac{\Gamma_j}{m_j}\right| \leq 0.02 \tag{35}$$

However, we found that this constraint is impossible to match with current values of the light resonances, as the sum was always exceeding the limit set in the constraint. We will further address this issue in Sec. 3.2.9, for now we should note that due to this peculiar result, we did not include this constraint in the final fit and listed it here only for completeness.

IV. *Covariance matrix of the form factors*

We also used the uncertainties on the FF and covariance matrix given in [12] as a constraint. The FF are what gives rise to the NR part of the decay rate in Eq. (5). As

we expect the cusp-like shape to be fairly small and laying on top of the NR part the sensitivity is rather dependant on the FF.

$\sigma_{b_0^+}$	$\sigma_{b_1^+}$	$\sigma_{b_2^+}$	$\sigma_{b_0^0}$	$\sigma_{b_1^0}$	$\sigma_{b_2^0}$	$\sigma_{b_0^T}$	$\sigma_{b_1^T}$	$\sigma_{b_2^T}$
0.014	0.128	0.548	0.010	0.125	0.441	0.019	0.236	0.971

$$\begin{aligned}\mathrm{Cov}\left(\begin{bmatrix} b_0^+ \\ b_1^+ \\ b_2^+ \\ b_0^0 \\ b_1^0 \\ b_2^0 \\ b_0^T \\ b_1^T \\ b_2^T \end{bmatrix}\right) &= \begin{bmatrix} 1. & 0.45 & 0.19 & 0.857 & 0.598 & 0.531 & 0.752 & 0.229 & 0.117 \\ 0.45 & 1. & 0.677 & 0.708 & 0.958 & 0.927 & 0.227 & 0.443 & 0.287 \\ 0.19 & 0.677 & 1. & 0.595 & 0.770 & 0.819 & -0.023 & 0.07 & 0.196 \\ 0.857 & 0.708 & 0.595 & 1. & 0.83 & 0.766 & 0.582 & 0.237 & 0.192 \\ 0.598 & 0.958 & 0.770 & 0.83 & 1. & 0.973 & 0.324 & 0.372 & 0.272 \\ 0.531 & 0.927 & 0.819 & 0.766 & 0.973 & 1. & 0.268 & 0.332 & 0.269 \\ 0.752 & 0.227 & -0.023 & 0.582 & 0.324 & 0.268 & 1. & 0.59 & 0.515 \\ 0.229 & 0.443 & 0.07 & 0.237 & 0.372 & 0.332 & 0.59 & 1. & 0.897 \\ 0.117 & 0.287 & 0.196 & 0.192 & 0.272 & 0.269 & 0.515 & 0.897 & 1. \end{bmatrix} \\ &= B_{ij}\end{aligned} \tag{36}$$

However, as we were only floating the b_m^+, we were only using a part of the covariance matrix. Having three parameters, we then constructed a three dimensional gaussian constraint:

$$\begin{aligned} L_{FF} =& -2\frac{t}{b} \\ t =& x_0^2(B_{12}^2-1) + x_1^2(B_{02}^2-1) + x_2^2(B_{01}^2-1) + 2x_0x_1(B_{01}-B_{02}B_{12}) \\ &+ 2x_0x_2(B_{02}-B_{01}B_{12}) + 2x_1x_2(B_{12}-B_{01}B_{02}) \\ b =& 2(B_{01}^2+B_{02}^2+B_{12}^2-2B_{01}B_{02}B_{12}-1) \end{aligned} \tag{37}$$

with:

$$x_m = \frac{b_{m,current}^+ - b_{m,true}^+}{\sigma_{b_m^+}} \tag{38}$$

where L_{FF} describes the loss added to the NLL during the fit, $b_{m,current}^+$ corresponds to the current value of the parameter during the fitting process and $b_{m,true}^+$ corresponds to the true value listed in Sec. 3.2.2.

V. *Constraint on the phases of the J/Ψ and $\Psi(2S)$ resonances*

As we are cutting out the J/Ψ and $\Psi(2S)$ regions, it becomes increasingly difficult for the fitting algorithm to find the phases of the J/Ψ and $\Psi(2S)$ resonances. To account for cutting out these regions, we added a gaussian constraint for each of the phases. As uncertainty we used the uncertainties mentioned in [14] and as the true values the ones listed in Sec. 3.2.2.

$$L_j = \frac{1}{2}\left(\frac{\delta_{j,current}-\delta_{j,true}}{\sigma_{\delta_j}}\right)^2, \text{with } j = J/\Psi, \Psi(2S) \tag{39}$$

3.2.5 Assumptions

As already mentioned in Sec. 3.1, 3.2.3, we assume infinite precision. To account for this assumption, we exclude the J/Ψ and $\Psi(2S)$ regions. Additionally, we artificially "smear" the curve afterwards to have a more realistic expectation. To artificially introduce a finite precision after fitting, we integrate the decay rate after a fit in small steps (100 keV wide bins) and then spread them out into a gaussian with the same area but a 7 MeV width. Doing this gives us a curve of the decay rate like Fig. 12, which looks more like what we expect to see in the real case.

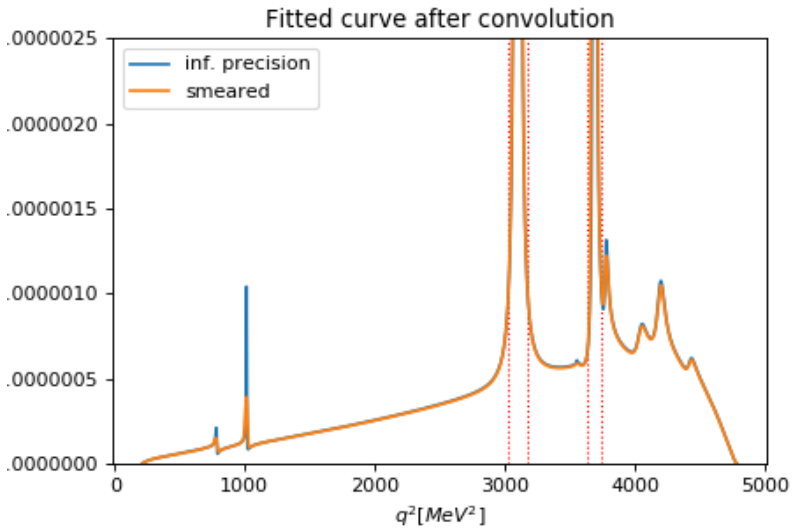


Figure 12: Example of how the shape of the curve changes after the smearing process (injected signal of $\mathcal{C}_9^\tau = 595$)

As outlined earlier in Sec. 3.1, we expect around 37′000 events. Expecting around 5.4 million events on the total q^2 range, we simply got this number by subtracting the events we expect to see from the J/Ψ and $\Psi(2S)$ resonances.

Also, as stated before, if you have a close look at [14], you will notice that there are four different phase combinations. We assumed all of them to deliver equal results. However, to test if our assumption is not too naive, we did some small testing. Therefore, we ran small analysis runs with around 200 toys each. There, all our four combinations were pointing to a similar sensitivity, while also having similar numbers and uncertainties on the parameters we were fitting.

3.2.6 CLs method

To calculate our sensitivity, we used the CL_S method presented in [19]. It is a statistical tool, allowing us to differentiate between actually seeing a signal or only a random fluctuation in the signal. In this method we are comparing fits with injected signal to one without injected signal or to express it more precisely, our confidence level (CL) to reject the null hypothesis (here: 1−CL = sensitivity).

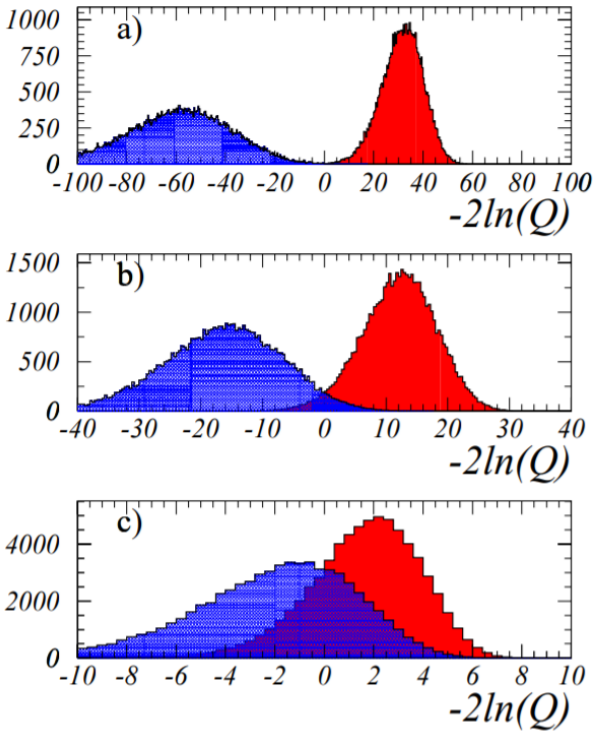


Figure 13: Example image showing the increasing separation of the difference of the NLLs (taken from [19])

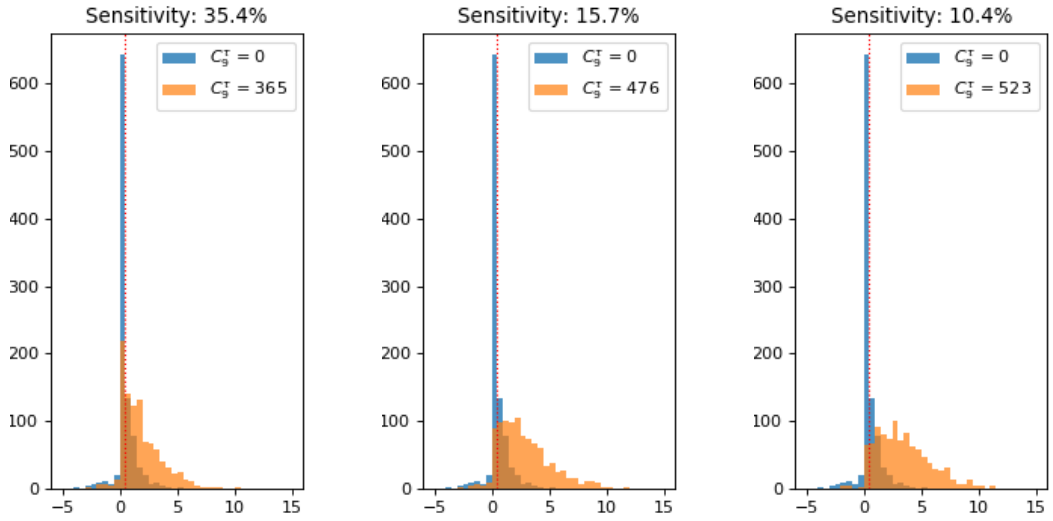
First we determine a range in which we want to probe the sensitivity and then choose specific points in this range. In this analysis this meant choosing and then probing different values of $\mathcal{C}_9^\tau$. So for example for the current scenario, we opted to probe the range of $\mathcal{C}_9^\tau$ in $[197, 680]$ with 20 steps, as this was the range we expected to lie in based on the hessian uncertainty of $\mathcal{C}_9^\tau$ seen in previous fits. Additionally we also need a step with $\mathcal{C}_9^\tau = 0$ as a point of reference with no injected signal, which will be used as a reference. We then generate a toy for each of the steps (including $\mathcal{C}_9^\tau = 0$) and fit the toy once with the value of $\mathcal{C}_9^\tau$ floating and once with the value of $\mathcal{C}_9^\tau$ fixed to zero. For the case, where we actually generated the toy with no injected signal, we then expect the NLLs of the fixed and floating fits to be very close to each other. The difference between these two is the number of comparison needed for our analysis of the sensitivity. As a next step, we then continue this procedure but now use toys with injected signal of the corresponding $\mathcal{C}_9^\tau$ size of the step we are analysing. This means we do one fit with the value of $\mathcal{C}_9^\tau$ fixed to zero and one with the value of $\mathcal{C}_9^\tau$ floating. However, now as the injected signal increases, we expect the fit with $\mathcal{C}_9^\tau$ set to zero to start representing the data worse than the fully

floating fit (as the curve has less freedom to adapt to the shape of the data). This means that with increasing $\mathcal{C}_9^\tau$, we expect the NLLs of the fits to separate more and more and hence the difference between the two NLLs to increase as well. If we repeat this procedure enough times to gain a sufficient amount of statistics, we will start to see the distributions of the NLLs separating (see Fig. 13). To reach enough statistics in our analysis we opted for 1000 toys.

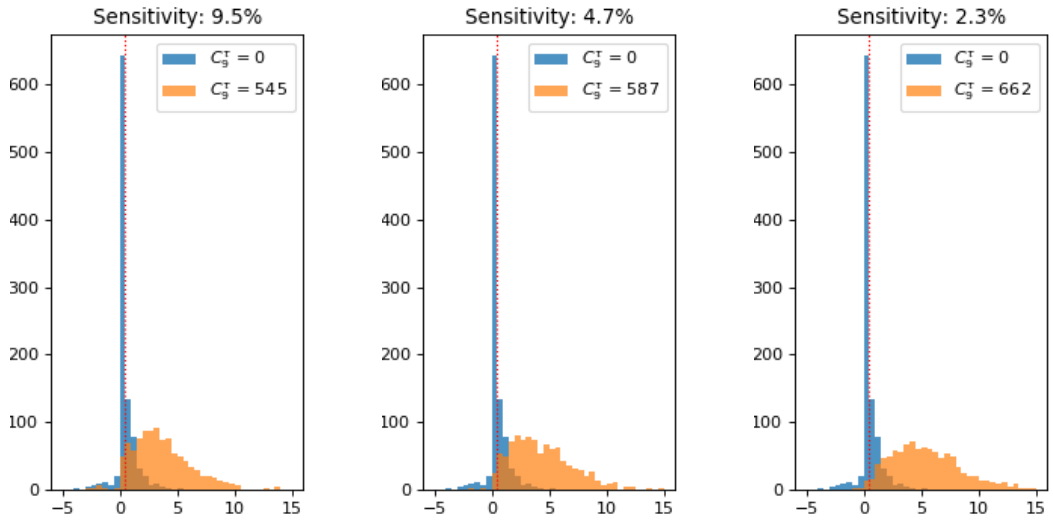
To get some measure of the separation of the two distributions, we then check how much of the distribution with the injected signal passes the mean of the distribution with no signal. We then get for our sensitivity S:

$$S = \frac{\#\text{events}_{\text{inj. Sig., left of mean}}}{\#\text{events}_{\text{no. Sig., right of mean}}} \tag{40}$$

Our goal was to see, at which values of $\mathcal{C}_9^\tau$ lies our 5% and 10% sensitivity. Scanning over different values of $\mathcal{C}_9^\tau$ gave us:



(a) Sensitivity at $\mathcal{C}_9^\tau = 365$ (b) Sensitivity at $\mathcal{C}_9^\tau = 476$ (c) Sensitivity at $\mathcal{C}_9^\tau = 523$



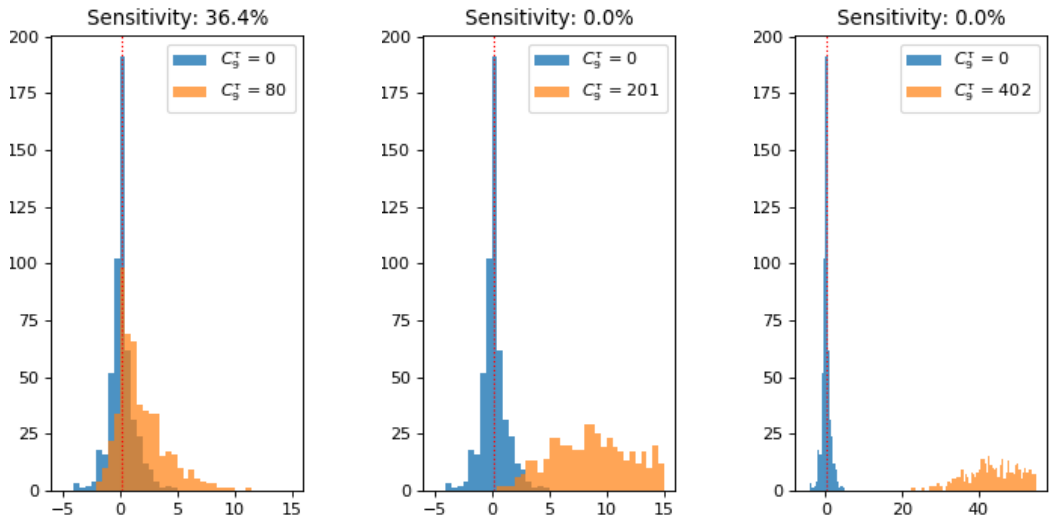
(d) Sensitivity at $\mathcal{C}_9^\tau = 545$ (e) Sensitivity at $\mathcal{C}_9^\tau = 587$ (f) Sensitivity at $\mathcal{C}_9^\tau = 662$

Figure 14: The CLs distributions over several values of $\mathcal{C}_9^\tau$ with 1000 toys each and the red-dotted line marking the mean the non-signal distribution

Fig. 15 shows very nicely the increasing separation of the two distributions as expected. Special attention here goes to Fig. 14c showing our sensitivity at 10% and to Fig. 14e showing our sensitivity at the 5% range. The 5% sensitivity proves to be especially useful, as it marks our CL to reject the null-hypothesis at 2σ. By interpolating in between results we then find $\mathcal{C}_9^\tau = 585$ as our 5% limit and $\mathcal{C}_9^\tau = 532$ as our 10% limit.

3.2.7 Upgrade scenario

It is always interesting to see how our sensitivity could improve in the near future. There are two main factors that could yield a significant improvement in sensitivity in the near future. The first one is simply having more data and hence more statistics. As the LHC and with it also the LHCb are in the middle of the upgrade period for run III, we expect to collect significantly more data during run III than we do now. According to [20], we can expect to collect up to 40× more data than during run I and II combined. Including the data we already have, this would increase the number of events to around 1.5 million events, which should definitely yield some improvement in sensitivity. Additionally, we can also expect progress on the theoretical front by giving us more constraints and smaller uncertainties. Improving the uncertainty on the form factors for example, could give us much better control on the NR part, which proves to be a large background for the ditau enhancement. As a possible future outlook and improvement we expect the uncertainty of the form factors to decrease by a factor of 1/3. These are the two possible improvements we want to analyse that create a plausible near future upgrade scenario.



(a) Sensitivity at $\mathcal{C}_9^\tau = 365$ (b) Sensitivity at $\mathcal{C}_9^\tau = 201$ (c) Sensitivity at $\mathcal{C}_9^\tau = 402$

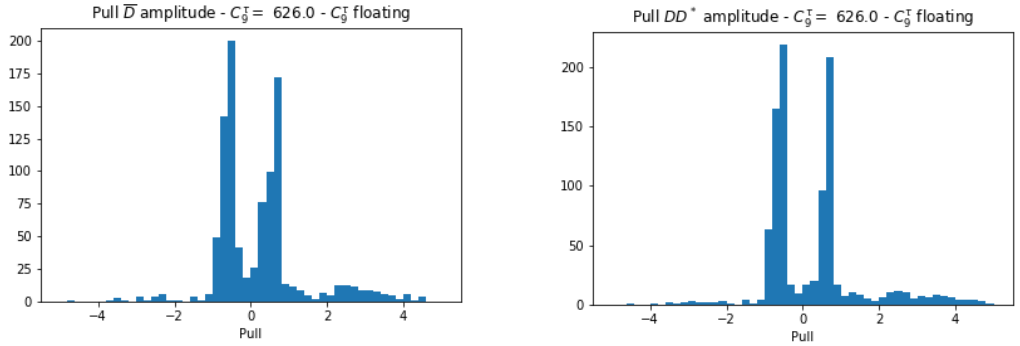
Figure 15: The CLs distributions of the upgraded scenario over several values of $\mathcal{C}_9^\tau$ with around 600 toys each and the red-dotted line marking the mean of the non-signal distribution

Due to time constraints in the analysis and the analysis being very time intensive, we only probed 3 points for the upgraded scenario. To actually find our 5% and 10% sensitivity, we interpolated in between the 3 data points. For simplification, we assumed a more or less linear behaviour for the interpolation, which will probably not describe the real behaviour. However, as the purpose of this outlook into an upgraded scenario is solely to determine the future potential of this way to probe LFU and not getting a precise result, this should suffice.

Interpolating linearly then yields a 5% sensitivity at $\mathcal{C}_9^\tau = 185$ and a 10% sensitivity at $\mathcal{C}_9^\tau = 168$.

3.2.8 Problem with D contributions

As already shown in Fig. 11, the D contributions can imitate the ditau signal fairly well. This originates from the DD^* and the $\overline{D}$ contributions having a similar shape as the ditau signal. As we generated with zero signal for the D contributions, the amplitude of both D contributions often took nonzero values during the fit. More precisely, the fit often found a positive value for the amplitude of one contribution and an equally big but negative value for the amplitude of the other contribution in the minimum (see Fig. 16 showing the pulls of the amplitudes). This resulted in them mostly cancelling each other out because of their similar shape. This bias seems intuitively correct, as we generated the toys with no injected signal of the D contributions, so them cancelling each other out would lead to a similar shape of the curve. The same thing also happens with significantly more statistics in the upgraded scenario. We expect them to be better controlled during the fit of the real data, as there we will have a non-zero signal for the D contributions. However, further analyses will definitely need to pay special attention to this behaviour.



(a) Pull of the amplitude of the $\overline{D}$ contribution

(b) Pull of the amplitude of the DD^* contribution

Figure 16: Pulls of the amplitudes of the D contributions, showing the peculiar behaviour of sometimes preferring nonzero values (Here an extreme case was chosen to demonstrate the problem.)

3.2.9 Problem with the constraint on the light contributions

While investigating the constraint of Eq. (26), which was derived by theoretical considerations, we found another peculiar result. When using the values presented in [14], we found ourselves to be always over the limit of the presented constraint. Inserting the values yields:

$$\begin{aligned} \textbf{Theory: } |Y^{1P}_{light}(0)| &\approx \left| \sum_{j=\rho,\omega,\phi} \eta_j \frac{\Gamma_j}{m_j} \right| \leq 0.02 \\ \textbf{Exp. data: } |Y^{1P}_{light}(0)| &\approx \left| \sum_{j=\rho,\omega,\phi} \eta_j \frac{\Gamma_j}{m_j} \right| \approx 0.36 \end{aligned} \tag{41}$$

Looking at Eq. (41) and comparing the theoretical prediction and the experimental results we see a clear discrepancy between the two, which cannot be explained by experimental uncertainties as well. We did not find a solution for this problem, which is also the reason why we did not include this constraint in the final fit. Nonetheless, we felt it was important to note this discrepancy, which should definitely be further investigated.

4 Results

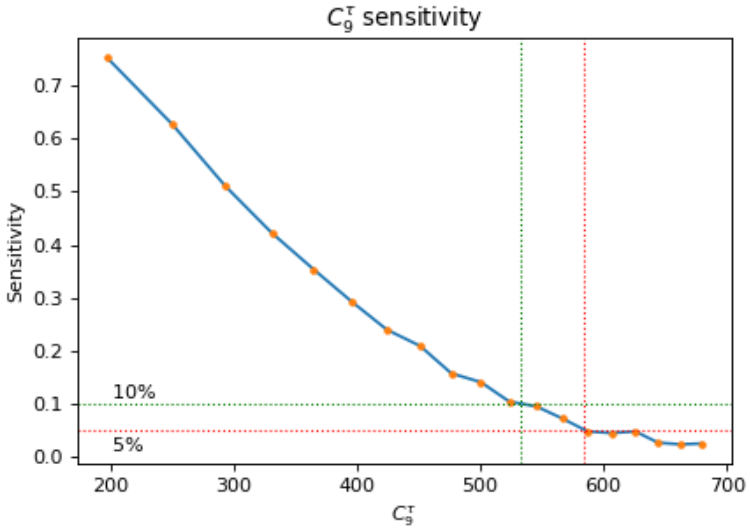


Figure 17: Sensitivity of the current scenario with green marking our 10% sensitivity and red marking our 5% sensitivity

Using the CLs method, we determined the 5% sensitivity to determine a ditau enhancement at $\mathcal{C}_9^\tau = 585$ and the 10% sensitivity at $\mathcal{C}_9^\tau = 532$ (see Fig. 17). Comparing this to the SM prediction of $\mathcal{C}_9^\tau \approx 4.2$ we would be able to detect an enhancement of a factor of 139 with a sensitivity of 2σ at LHCb.

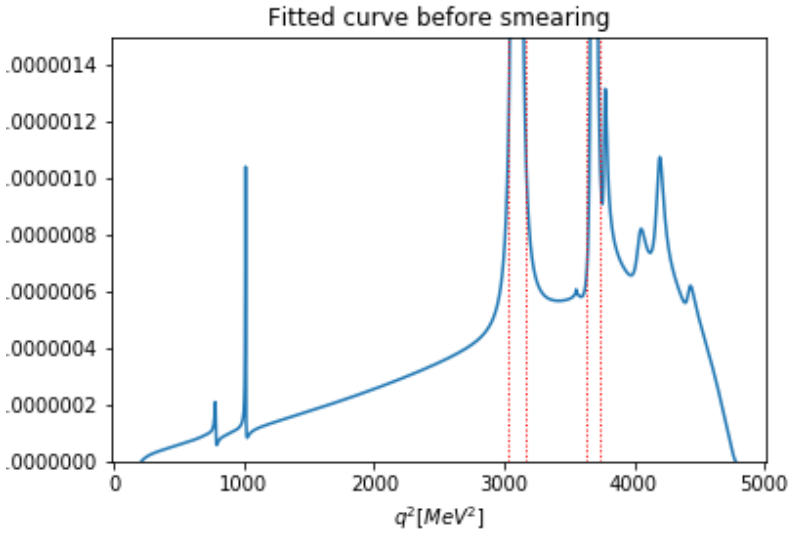


Figure 18: Example of a fitted curve before smearing with an injected signal of $\mathcal{C}_9^\tau = 585$

Fig. 18 shows the decay rate of the dimuon spectrum with an injected signal[8] of

[8]Corresponding to our 5% sensitivity

$\mathcal{C}_9^\tau = 585$ with around 37′000 events. Note that this fit still assumes infinite precision. To showcase a more realistic scenario, we convoluted the curve as described in Sec. 3.2.5 with a finite precision of 7 MeV after fitting. Applying this procedure returns a curve like Fig. 19.

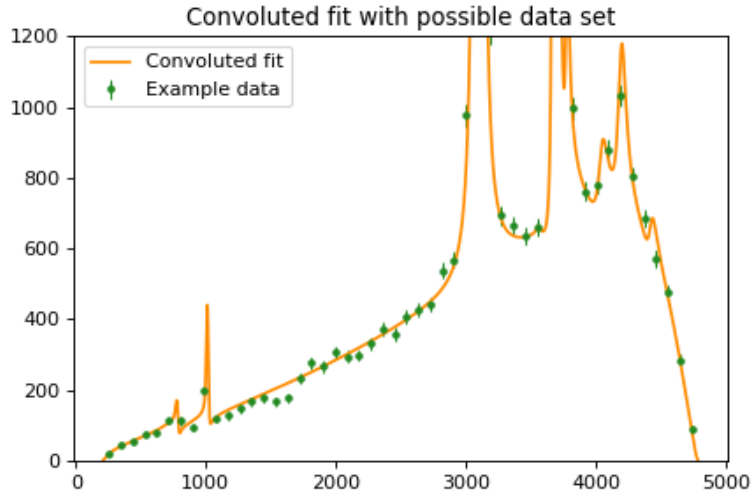


Figure 19: Example of a convoluted fit with data (injected signal of $\mathcal{C}_9^\tau = 585$)

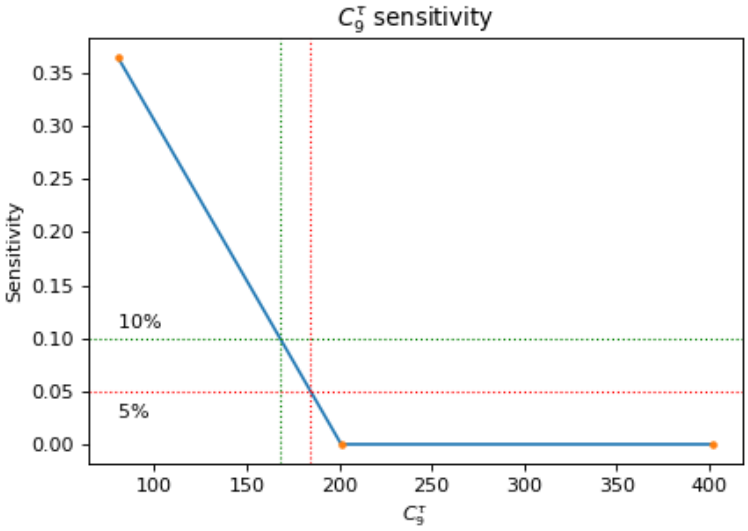


Figure 20: Sensitivity of the upgraded scenario with green marking our 10% sensitivity and red marking our 5% sensitivity (each data point consists of around 600 toys)

To estimate the sensitivities of the upgraded scenario[9], we used interpolation and only a few data points (see Fig. 20), so they should be used with caution. Nonetheless, this

[9] 41 times amount of data, uncertainty on FF decreased by 1/3

should give us an idea of the increase in sensitivity we can expect in the near future. We estimate the 5% sensitivity to be at $\mathcal{C}_9^\tau = 185$ and the 10% sensitivity to be at $\mathcal{C}_9^\tau = 168$. This implies that we can improve our current sensitivity by a factor of 3 and be able to detect an enhancement of a factor of 44 with a sensitivity of 2σ at LHCb.

5 Conclusion

We have demonstrated that probing LFU by searching for imprints of the ditau rescattering in the dimuon spectrum is already a viable method at LHCb. LHCb is even able to take a leading position[10] ($\mathcal{B}(B \to K\tau\tau) \approx 1.78 \times 10^{-3}$ at 90% CL) together with Belle II [21] ($\mathcal{B}(B \to K\tau\tau) = 2.25 \times 10^{-3}$ at 90% CL) in this sort of search. Currently, both experiments are only able to detect enhancements of $\mathcal{C}_9^\tau$ at three order of magnitudes with a CL of 90%. However, using our estimate of the upgraded scenario including run III of the LHC and improvements on the FF uncertainties, we expect to be able to increase our sensitivity significantly. We should even be able to probe enhancements of $\mathcal{O}(2)$ in the near future.

[10]The conversion from $\mathcal{C}_9^\tau$ to a BR introduces model dependence, here we used the conversion presented in Sec. 2.4.

6 Acknowledgment

I would like to first express my gratitude to Professor Nicola Serra for letting me do this thesis in his group and being very supportive and trusting overall.

A huge thank you goes to Dr. Patrick Owen, who was a great supervisor providing amazing support and just generally providing a productive atmosphere. All the discussions, quick responses to emails and inputs helped immensely to advance and develop this thesis.

I also want to thank Professor Gino Isidori as well as Dr. Matthias König and Dr. Claudia Cornella for their work providing all the theoretical framework and all the discussions. Without their work, there would be no foundation to build this thesis upon.

Also special thanks goes to Jonas Eschle, for always being open and taking time for questions concerning zfit and physics in general. Without his inputs and smart ideas, many parts of the script would be less elegant and have taken more time to figure out on how to solve them.

A great thanks goes to the University of Zürich and CERN for offering the opportunities to do this kind of research by providing the adequate resources.

Last but not least I would like to thank three people not only professionally for their support but also privately for various things, too many as they could even remotely be summarised here: Janina Justus, Peter Liechti and Barbara Liechti.

References

[1] R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+ l^+ l^-$ *decays*, Physical review letters **113** (2014) 151601.

[2] R. Aaij *et al.*, *Test of lepton universality with* $B_0 \to K_0^* l^+ l^-$ *decays*, Journal of High Energy Physics **2017** (2017) 55.

[3] R. Aaij *et al.*, *Search for lepton-universality violation in* $B^+ \to K^+ l^+ l^-$ *decays*, Physical review letters **122** (2019) 191801.

[4] A. Abdesselam *et al.*, *Test of lepton flavor universality in* $B \to K_0^* l^+ l^-$ *decays at belle*, arXiv preprint arXiv:1904. 02440 (2019).

[5] R. Aaij *et al.*, *Measurement of the ratio of branching fractions* $\mathrm{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathrm{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\tau)$, Physical review letters **115** (2015) 111803.

[6] J. Lees *et al.*, *Measurement of an excess of* $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ *decays and implications for charged higgs bosons*, Physical Review D **88** (2013) 072012.

[7] S. Hirose *et al.*, *Measurement of the* τ *lepton polarization and* $\mathrm{R}(\mathrm{D}^*)$ *in the decay* $\overline{B} \to D^*\tau^-\overline{\nu}_\tau$, Physical review letters **118** (2017) 211801.

[8] 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*, Physical Review D **97** (2018) 072013.

[9] A. Abdesselam *et al.*, *Measurement of R(D) and R(D*) with a semileptonic tagging method*, arXiv preprint arXiv:1904. 08794 (2019).

[10] R. Aaij *et al.*, *Search for the decays* $B_s^0 \to \tau^+\tau^-$ *and* $B \to \tau^+\tau^-$, Physical Review Letters **118** (2017) .

[11] J. ees *et al.*, *Search for* $B_s^0 \to K\tau^+\tau^-$ *at the* BaBar *experiment*, Physical Review Letters **118** (2017) .

[12] A. Khodjamirian, T. Mannel, and Y.-M. Wang, $B^+ \to K^+ l^+ l^-$ *decay at large hadronic recoil*, Journal of High Energy Physics **2013** (2013) .

[13] J. A. Bailey *et al.*, $B^+ \to K^+ l^+ l^-$ *decay form factors from three-flavor lattice qcd*, Physical Review D **93** (2016) .

[14] R. Aaij *et al.*, *Measurement of the phase difference between short- and long-distance amplitudes in the* $B^+ \to K^+\mu^+\mu^-$ *- decay*, The European Physical Journal C **77** (2017) .

[15] J. ees *et al.*, *Search for* $B^+ \to K^+\tau^+\tau^-$ *babar experiment*, Physical Review Letters **118** (2017) .

[16] R. Aaij *et al.*, *Measurement of the phase difference between short- and long-distance amplitudes in the* $B^+ \to K^+\mu^+\mu^-$ *decay*, The European Physical Journal C **77** (2017) .

[17] M. Abadi *et al.*, *Tensorflow: A system for large-scale machine learning*, doi, pp. 265–283, 2016.

[18] J. Eschle *et al.*, *zfit: scalable pythonic fitting*, 2019. doi: 10.5281/zenodo.3356645.

[19] A. L. Read, *Modified frequentist analysis of search results (the CL_s method)*, tech. rep., Cern, 2000.

[20] LHCb Collaboration, R. Aaij *et al.*, *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*, Tech. Rep. CERN-LHCC-2017-003, CERN, Geneva, 2017.

[21] S. Sandilya, *Radiative and EWP B decays at B-factories*, PoS **BEAUTY2018** (2018) 046.