UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
DIPARTIMENTO DI FISICA G. OCCHIALINI
LAUREA MAGISTRALE IN FISICA

Study of the $\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu}$ differential decay width : analysis of the sensitivity of LHCb to the scalar form factor

Relatore:
Prof.ssa: Marta Calvi (Università di Milano-Bicocca)
Correlatori:
Prof: Gino Isidori (University of Zuerich)
Prof: Nicola Serra (University of Zuerich)

Candidato: Davide Lancierini
Matricola: 784555

SESSIONE DI LAUREA: 18 LUGLIO 2017
ANNO ACCADEMICO 2016-2017

Abstract

In this thesis work we will investigate $R_{\mathrm{D}} = \frac{\mathcal{B}(\overline{\mathrm{B}} \to \mathrm{D}\tau\bar{\nu})}{\mathcal{B}(\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu})}$ and the form factors that enter the $\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu}$ decay width, since their precise determination is of crucial importance for both theoretical estimates and experimental measurements of the aforementioned ratio. After reviewing the theoretical tools of Heavy-Quark Effective Theory, we updated the theoretical estimate of the upper bound on the value of the vector form factor f_+ at zero recoil obtaining a result of $f_+(q^2_{\mathrm{max}}) \leq 1.159 \pm 0.024$.

In the second part of this work we elaborate on the Bourrely-Caprini-Lellouch parametrization for the form factors, and we implement it in a fitter whose performance will be first tested on a infinite resolution data sample. In perfect resolution conditions the ratios of the parameters of the form factors can be retrieved with sensitivity of 0.56% and 8.82 % for the vector and the scalar form factor respectively, in a 2×10^6 events sample.

In the last part we will be simulating the LHCb detector resolution, and correct for the effect of bin migration by means of the unfolding techniques. In this case the relative error on the ratio between vector form factor parameters ratio raises to 1.06% while it increases to 24% in a 2×10^6 events sample collected in the worst resolution condition.

Aknowledgements

I would like to cordially thank my relator Marta Calvi for giving me the chance to work on this project, and for her valuable competence (and patience). She constituted a precious advisor in indicating me the path to write my thesis.

Nevertheless, I am also grateful to Gino Isidori and Nicola Serra for their esteemed help and the treasured discussions.

Last but not least I would like to thank the whole Zurich working group which warmly welcomed me for the realisation of this project.

CONTENTS

I. INTRODUCTION

One of the most interesting phenomena reported by particle physics experiments in the last few years are the numerous hints of Lepton Flavour Universality (LFU) violations observed in semileptonic B decays. The very recent LHCb result on the LFU ratios $R^{\mu e}_{K^{(*)}}$ [2] and $R^{\tau\mu}_{D^{(*)}}$ [3] are the last two pieces of a seemingly coherent set of anomalies which involves different observables and experiments. So far, not a single LFU ratio measurement exhibits a deviation with respect to the Standard Model (SM) abode the 3σ level. However, the overall set of observables is very consistent and, once combined, the probability of a mere statistical fluctuation is very low.

The evidences collected so far can naturally be grouped into two categories, according to the underlying quark-level transition:

- deviations from τ/μ (and μ/e) universality in $\mathrm{b} \to \mathrm{c}\ell\bar{\nu}$ charged currents ([3] - [5])
- deviations from μ/e universality in $\mathrm{b} \to \mathrm{s}\ell\bar{\ell}$ neutral currents ([2] - [6])

In both cases the combination of the results lead to an evidence around the 4σ level for LFU violating contribution of non-SM origin, whose size is $\mathcal{O}(10\%)$ compared to the corresponding charged- or neutral- current SM amplitudes. One of the puzzling aspects of the present anomalies is that they have been observed only in semi-leptonic B decays and are quite large compared to the corresponding SM amplitudes [1].

On the contrary, no evidence of deviation from the SM has been seen so far in the precise (per-mil) tests of LFU in semi-leptonic K and π decays, purely leptonic τ decays, and in the electroweak precision observables. The most natural assumption to address this apparent paradox is the hypothesis thet the NP responsible for the breaking of LFU is coupled mainly to the third generations of quarks and leptons, with a small mixing with the light generations [1].

This hypothesis also provides a natural first-order explanation for the different size of the two effects, which compete with a tree-level SM amplitude in charged currents, and with a suppressed loop-induced SM amplitude in neutral currents, respectively.

In this work we will focus on semi-leptonic B decays, and in particular on the ratio

$$R_{\mathrm{D}} = \frac{\mathcal{B}(\overline{\mathrm{B}} \to \mathrm{D}\tau\bar{\nu})}{\mathcal{B}(\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu})} \tag{1}$$

The theoretical estimate of R_{D} within the SM relies dominantly on the hadronic form factors f_+ (the vector form factor) and f_0 (the scalar form factor). For both, precise lattice QCD results have recently been published [24]. In addition, Zero-Recoil Sum Rules (ZRSR) results can be used to complement the lattice QCD results. According to [24] the SM prediction for R_{D} is:

$$R_{\mathrm{D}}^{SM} = 0.300 \pm 0.008 \tag{2}$$

On the experimental side, measurement of R_{D} have been published by both BaBar [3] and, more recently, by Belle [7],

$$R_{\mathrm{D}}^{\mathrm{BaBar}} = 0.440 \pm 0.058 \pm 0.042, \quad R_{\mathrm{D}}^{\mathrm{Belle}} = 0.375 \pm 0.064 \pm 0.026 \tag{3}$$

Combining the two results and normalizing them to the SM, leads to [8]:

$$\Delta R_{\mathrm{D}} = \frac{R_{\mathrm{D}}^{exp}}{R_{\mathrm{D}}^{SM}} - 1 = 0.35 \pm 0.17 \tag{4}$$

A similar effect has been observed also in the R_{D^*} ratios [3]. Combining the two deviations, which are compatible with a enhancement of semileptonic $\mathrm{b} \to \mathrm{c}\tau\nu$ transitions over $\mathrm{b} \to \mathrm{c}\mu\nu$ ones, the discepancy with respect to the SM raises to about 4.1σ [10].

This work is structured as follows: in the first section a review of the calculation of the differential decay width is done, which is followed by analysis of the methods of HQET and their theoretical tools to investigate the form factors. An introduction to the Zero-Recoil Sum Rule calculation is done, and the value at Zero Recoil of the vector form factor is updated with recent values for quark masses and for the parameters of the heavy quark expansion.

Of the existing parametrizations of the form factors the Bourrely - Caprini - Lellouch (BCL) is investigated in detail and the coefficients for the weak unitarity bounds test are calculated. This parametrization is used to test a fit sensitivity to the scalar form factor in the hypothesis of perfect resolution and efficiency.

Subsequently LHCb resolution is simulated through Monte Carlo data, and the techniques of linear inverse ill-posed problems such as unfolding, are applied to give an estimate of the sensitivity to the parameters that describe the form factors in BCL parametrization, given the LHCb detector performances.

II. $\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu}$ DIFFERENTIAL DECAY WITDTH

In a generic NP scenario without right-handed neutrinos that preserves LFU, a coupling from this decay to the scalar, vector and tensor operators can be described via [11]

$$\mathcal{L}^{\mathrm{eff}} = \frac{G_F}{\sqrt{2}} V_{\mathrm{cb}} \left[(\bar{\mathrm{c}}\gamma^\mu \mathrm{b})(\bar{\ell}_{\mathrm{L}}\gamma_\mu \bar{\ell}_{\mathrm{L}}) + g_V(\bar{\mathrm{c}}\gamma^\mu \mathrm{b})(\bar{\ell}_{\mathrm{L}}\gamma_\mu \bar{\ell}_{\mathrm{L}}) + g_S(\mu)(\bar{\mathrm{c}}\ \mathrm{b})(\bar{\ell}_{\mathrm{R}}\bar{\ell}_{\mathrm{L}}) + g_T(\mu)(\bar{\mathrm{c}}\sigma_{\mu\nu}\mathrm{b})(\bar{\ell}_{\mathrm{R}}\sigma^{\mu\nu}\nu_{\mathrm{L}}) \right] + h.c. \tag{5}$$

Where the dimensionless couplings $g_{V,S,T} \propto m_{\mathrm{W}}^2/\mathrm{m}_{NP}^2$ where m_{NP} being the NP scale. As can be verified from eq. (20) of [11] the NP decay width for the process $\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu}$ depends on the scalar, vector and tensor form factors and the NP couplings, thus a model independent direct measurement of the scalar and vector form factors could further constrain the magnitude of those couplings,

A. Dynamics

In order to calculate the amplitude for this decay in the SM framework, it is useful to work in an Effective Field Theory that does not contain the W boson. The Standard Model flavour changing vertex is replaced by a local four-fermion operator $\mathcal{O}_L$ for the transition $\mathrm{b} \to \mathrm{c}\ell\bar{\nu}$, where ℓ denotes a charged lepton $\ell = \mathrm{e}, \mu, \tau$. The Lagrangian is the following:

$$\mathcal{L}^{\mathrm{eff}} = \frac{G_{\mathrm{F}}}{\sqrt{2}} V_{\mathrm{cb}} [\bar{\psi}_{\mathrm{c}}\gamma^\mu(1-\gamma_5)\psi_b][\bar{\psi}_l\gamma_\mu(1-\gamma_5)\psi_\nu] \tag{6}$$

The differential decay amplitude then reads:

$$d\Gamma = \frac{1}{2m_{\mathrm{B}}} |\mathcal{M}|^2 \, d\Phi_3 \tag{7}$$

Where $\mathcal{M}$ is the amplitude for the given process and $d\Phi_3$ is the three-body phase space. To leading order in α_{EW} M factorises as the product of the hadronic and leptonic currents, and choosing the phases that would arise from the time reversal operation on the hadronic states so that the form factors are real, the square of the matrix element yields

$$|\mathcal{M}|^2 \propto H^\mu H^{*\nu} L_\mu L_\nu^* \tag{8}$$

Where

$$H^{\mu} = \langle \mathrm{D}(p^{'}) \left| \bar{\psi_{\mathrm{c}}} \gamma^{\mu} \psi_{\mathrm{b}} \right| \overline{\mathrm{B}}(p) \rangle = f_{+}(q^{2}) \left[(p+p^{'})^{\mu} - \frac{m_{\mathrm{B}}^{2} - m_{\mathrm{D}}^{2}}{q^{2}} q^{\mu} \right] + f_{0}(q^{2}) \frac{m_{\mathrm{B}}^{2} - m_{\mathrm{D}}^{2}}{q^{2}} q^{\mu} \tag{9}$$

with p^{μ} and $p^{'\mu}$ representing the four-momenta respectively of the incoming B and D outgoing mesons, and $q^{\mu} = (p - p^{'})^{\mu}$.

$$\sum_{spins} L_{\mu} L_{\nu}^{*} = 2[(\bar{q}+q)_{\mu}(\bar{q}-q)_{\nu} + (\bar{q}+q)_{\nu}(\bar{q}-q)_{\mu} - g_{\mu\nu}(\bar{q}^{2} - q^{2})]$$

In the above $f_{+}(q^{2})$ and $f_{0}(q^{2})$ are the two form factors, vector and scalar respectively, that can only depend on the unique Lorentz invariant quantity one can build with the four-vectors of the problem, q^{2}. Note that the pseudo-vector current in the hadronic part of the matrix element has been neglected, on account of parity conservation of QCD.

As far as the leptonic current is concerned, $\bar{q}$ represents the sum between the four momenta of the outgoing leptons, and it's totally antisymmetric part has been neglected since it vanishes when contracted with $H^{\mu} H^{*\nu}$, being the latter a totally symmetric tensor in it's two indices.

B. Phase space and differential decay width

The Lorentz invariant phase space element (LIPS) for this process can be written as

$$d\Phi_{3} = -\frac{1}{64\pi^{3}} \frac{\sqrt{\lambda}}{4m_{\mathrm{B}}^{2}} \left(1 - \frac{m_{\mu}^{2}}{q^{2}} \right) d(cos\theta) dq^{2} \tag{10}$$

Where λ is defined as

$$\lambda(q^{2}, m_{\mathrm{B}}^{2}, m_{\mathrm{D}}^{2}) = (q^{2} - m_{\mathrm{B}}^{2} - m_{\mathrm{D}}^{2})^{2} - 4m_{\mathrm{B}}^{2} m_{\mathrm{D}}^{2}$$

And θ is identified as the polar angle between the spatial components of the charged lepton four-momenta and the spatial components of the D (or as well of the B) meson four-momenta in the leptons' centre of mass frame. Note that one is able to integrate over the azimuthal angle φ of the charged lepton in the centre of mass frame, since both the B and the D mesons do not hold any polarization information. Since the isotropy hypothesis holds, one can set $\varphi = 0$.

The differential decay width, as a function of q^{2}, reads [23]:

$$\frac{d\Gamma}{dq^{2}}(\mathrm{B} \to \mathrm{D}\ell\nu_{\ell}) = \frac{\eta_{ew}^{2} G_{F}^{2} \left| V_{cb} \right|^{2} m_{\mathrm{B}} \sqrt{\lambda}}{192\pi^{2}} \left(1 - \frac{m_{\ell}^{2}}{q^{2}} \right)^{2} \left[\frac{\lambda}{m_{\mathrm{B}}^{4}} \left(1 + \frac{m_{\ell}^{2}}{2q^{2}} \right) |f_{+}(q^{2})|^{2} + (1 - r_{\mathrm{D}}^{2})^{2} \frac{3m_{\ell}^{2}}{2q^{2}} |f_{0}(q^{2})|^{2} \right] \tag{11}$$

The factor $\eta_{ew} = 1 + \alpha/\pi \ln(m_{\mathrm{Z}}/m_{\mathrm{B}}) \simeq 1.0066$ [9]. In case of massless leptons the f_{0} contribution becomes irrelevant and it is in the scope of this work to verify at which extent muons can be considered as massless, in the measurement of the decay width, at present LHCb detector capabilities. In order to do so, first we need to review the most updated theoretical previsions for such form factors to which we will dedicate the following sections.

III. THE ZERO-RECOIL SUM RULE

A. Ingredients:

Sum rules in QCD allow to infer information on hadronic parameters such as decay constants, transition form factors, with correlation functions of quark currents. In order to do so, one makes use of three key elements [12]:

- Correlation functions of local quark currents:

  In the simplest case the two-point correlation function is formed by two quark-antiquark current operators of the form:

  $$\langle 0|T\{\bar{q}_i(x)\Gamma^\mu q_j(x)\bar{q}_j(0)\Gamma^\nu q_i(0)\}|0\rangle$$

  This is a function of the 4-momentum transfer between the currents and, in the region of large momentum transfers, it represents a short distance fluctuation of the $q\bar{q}$ fields. In such region the gluon exchanges are suppressed by a small QCD coupling and the propagation of q and $\bar{q}$ at short distances is asymptotically free.

- Operator Product Expansion (OPE):

  OPE of the correlation function of the previous point is worked out, providing an analytical expression for the correlation function at $q^2 > 0$ with a systematic separation of short and long distance effects, the former being described by Feynman diagrams, while the latter are encoded by universal parameters related to non perturbative QCD dynamics. In the case of two-point sum rules, these parameters are the averaged local densities of the QCD vacuum fields, the condensates, whose contributions in the OPE are suppressed by powers of $1/m_q^n$ allowing to truncate the series at some maximal power.

- Hadronic dispersion relations:

  Optical theorem and unitarity condition are used to link the imaginary part of the correlation function in terms of the sum and/or integration over all intermediate hadronic states with quantum numbers of the quark currents. In this way a link between perturbative QCD and hadrons is established and the resulting relation between the OPE expression and the hadronic sum is called "QCD sum rule"

B. The Heavy-Quark Expansion: an introduction

One of the tools that will be needed in order to perform a systematic expansion of the two-point correlation function which stands at the core of the zero-recoil sum rule is an effective theory that allows its OPE. This tool is heavy-quark effective theory.

In analysing the strong interactions of hadrons containing heavy quarks, it's convenient to work in an EFT that naturally separates the scales of the processes through $\Lambda_{\rm QCD}$ [13]. Roughly speaking $\Lambda_{\rm QCD} \sim 0.2$ GeV and quarks fall naturally in two classes: light quarks such as the u, d and s and heavier quarks like c, b and t, the latter being to heavy to form bound states. For heavy quarks the effective coupling constant $\alpha_s(m_Q)$ is small, implying that on the length scales comparable to the Compton wavelength $\lambda_Q \sim 1/{\rm m}_Q$ the strong interactions are perturbative and behave like the electromagnetic interactions. In fact, the quarkonium systems made up by two heavy quarks, whose size is of order $\lambda_Q/\alpha_s(m_Q) << R_{\rm had}$, are very much hydrogen-like [13].

Systems composed of a heavy quark and other light constituents have a size determined by $R_{\rm had} \sim 1/\Lambda_{\rm QCD} \sim 1$fm, and the typical momenta exchanged between the heavy and the light degrees of freedom are of order $\Lambda_{\rm QCD}$. The heavy quark is surrounded by a strong interacting cloud of light quarks, antiquarks and gluons but the fact that $\lambda_Q << R_{\rm had}$ leads to simplifications. Soft gluons are not hard enough to resolve the quantum numbers of the heavy quark, therefore the light degrees of freedom are blind to flavour (mass) and spin orientation of the heavy quark, they experience only its colour field.

It follows that, in the limit $m_Q \to \infty$, hadronic systems which differ only in the flavour or spin quantum numbers of the heavy quark have the same configuration of their light degrees of freedom [13]. This observation provides relations between the properties of heavy mesons such as B, D, B^* and D^*. These relations result from some approximate symmetries of the effective strong interactions of heavy quarks at low energies: for N_h heavy-quark flavours, there is a $SU(2N_h)$ spin-flavour symmetry group, under which the effective strong interactions are invariant. The configuration of light degrees of freedom in a hadron containing a single heavy quark with velocity v does not change if this quark is replaced by another heavy quark with different flavour or spin, but with the same velocity.

Heavy-quark symmetry is an approximate symmetry of an effective theory, which is a good approximation of QCD in a certain kinematic region, the symmetry breaking corrections being powers of $1/m_Q$ and in order to study them systematically it is necessary to cast the QCD Lagrangian for a heavy quark, into a form suitable for taking the limit $m_Q \to \infty$.

C. The Heavy-Quark Expansion

As compared with most effective theories, in which the degrees of freedom of a heavy particle are removed completely from the low energy theory, the Heavy-Quark Effective Theory (HQET) is special in that it's purpose is to describe the properties and decays of hadrons which do contain a heavy quark. Hence it's not possible to remove the heavy quark completely from the effective theory. What is done is to integrate out the "small components" in the full heavy-quark spinor, which describe the fluctuations around the mass shell. In order to do this one singles out the part of the QCD Lagrangian for a heavy quark:

$$\mathcal{L}_{\mathrm{HQ}} = \bar{\Psi}_Q (i\not{D} - m_Q)\Psi_Q + \tag{12}$$

The field operator Ψ_Q defined as:

$$\Psi_Q = \int \frac{d^3 p_Q}{(2\pi)^4} \frac{1}{\sqrt{E_Q}} \left[a(x) u_Q(x) e^{-ip_Q x} + a^\dagger(x) v_Q(x) e^{+ip_Q x} \right]$$

where a, $a^\dagger$ are creation and annihilation operators, and u_Q, v_Q are the heavy quark spinors while p_Q is the momentum of the heavy quark Q. Any momentum is represented by p_Q, although dominant contributions come from those p_Q which come from semiclassical approximation, i.e. the ones which satisfy dispersion relations or are $\hbar$ distant from it.

By rephasing the aforementioned field one gets:

$$\tilde{\Psi}_Q = e^{im_Q v\cdot x}\Psi_Q = \int \frac{d^3 p_Q}{(2\pi)^4 \sqrt{E_Q}} \left[a(x) u_Q(x) e^{-i(p_Q + m_Q v)x} + a^\dagger(x) v_Q(x) e^{+i(p_Q - m_Q v)x} \right] \tag{13}$$

Now let's consider $p_Q^\mu = m_Q v^\mu + k^\mu$ such that $\left|k^\mu\right| = \mathcal{O}(\Lambda_{\mathrm{had}})\ \forall \mu$, note that this choice of is not unique, one option is that it can be chosen as the velocity of the external hadron. Interactions of the heavy quark with light degrees of freedom change the residual momentum by an amount $\Delta k \sim \Lambda_{\mathrm{QCD}}$, but the corresponding changes in the heavy-hadron quark velocity vanish as $\Lambda_{\mathrm{QCD}}/m_Q \to 0$. By defining the two projectors:

$$P_\pm = \frac{1 \pm \not{v}}{2} \tag{14}$$

And by the use of standard spinor relations in momentum space one can easily see that $P_+ u_Q(m_Q v) = u_Q(m_Q v)$ and $P_- u_Q(m_Q v) = 0$. Similar relations hold for v_Q. Such relations receive corrections of $\mathcal{O}(1/m_Q)$ when considering the residual momentum k^μ. It is then straightforward to decompose the heavy quark field operator respectively in the large and small components:

$$h_v(x) = P_+ \tilde{\Psi}_Q(x) \qquad \text{and} \qquad H_v(x) = P_- \tilde{\Psi}_Q(x)$$

Such that $\tilde{\Psi}_Q(x) = h_v(x) + H_v(x)$. Substituting the latter expression in the QCD Lagrangian one finds:

$$\mathcal{L}_{\text{HQ}} = \left(\bar{h}_v + \bar{H}_v\right)(i\not{\mathcal{D}} - 2P_- m_Q)(h_v + H_v) \tag{16}$$

Note that up to this point, full QCD hasn't been left yet and no approximations were applied to the Lagrangian. By noting that any four-vector can be decomposed as: $\mathcal{D}^\mu = (v\cdot\mathcal{D}v^\mu + \mathcal{D}_\perp^\mu)$, one can see that terms appearing in the QCD Lagrangian of the form $\bar{h}_v(i\not{\mathcal{D}})h_v$ can be decomposed along the components of the four-velocity v^μ obtaining terms of the form:

$$\bar{h}_v\left(i\not{\mathcal{D}}_\perp + i(v\cdot\mathcal{D})\not{v}\right)h_v$$

And by noting that terms such as $i\not{\mathcal{D}}_\perp h_v$ are equivalent to $P_- i\not{\mathcal{D}}_\perp h_v$ one can make use of the orthogonality relations associated to $P_\pm$ to rewrite $\mathcal{L}_{\text{QCD}}$ as:

$$\mathcal{L}_{\text{HQ}} = \bar{h}_v i(v\cdot\mathcal{D})h_v + \bar{H}_v i(\not{\mathcal{D}}_\perp)h_v + \bar{h}_v i(\not{\mathcal{D}}_\perp)H_v - \bar{H}_v\left[i(v\cdot\mathcal{D}) + 2m_Q\right]H_v \tag{17}$$

The first term being the kinetic term of a massless mode of the heavy quark field while the last representing the one of a massive $2m_Q$ quark. These will be the heavy degrees of freedom that will be eliminated in the construction of the effective field theory. The fields are mixed by the second and third term which describe pair creation or annihilation of heavy quarks and antiquarks. On a classical level, the heavy degrees of freedom H_v can be eliminated by using the Euler-Lagrange equations:

$$\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial_\mu\bar{H}_v} - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{H}_v)}\right] = 0 \tag{18}$$

Thus we obtain

$$(iv\cdot\mathcal{D} + 2m_Q)H_v = i\not{\mathcal{D}}_\perp h_v$$

which can be formally solved to give:

$$H_v = \frac{1}{2m_Q + iv\cdot\mathcal{D}}i\not{\mathcal{D}}_\perp h_v$$

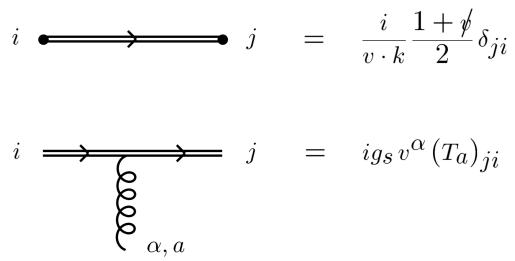
By substituitng the latter expression in equation (17) one obtains the Lagrangian:

$$\mathcal{L}_{\text{HQ}} = \bar{h}_v\left(iv\cdot\mathcal{D} + i\not{\mathcal{D}}_\perp\frac{1}{2m_Q + iv\cdot\mathcal{D}}i\not{\mathcal{D}}_\perp\right)h_v \tag{19}$$

Which contains the inverse derivative operator, thus introducing an integration at all possible locations which makes the $\mathcal{L}_{\text{QCD}}$ non-local. In momentum space, derivatives acting on h_v yeld powers of the residual momentum k, which is much smaller than m_Q. Hence, the non-local effective Lagrangian 17 allows for a derivative expansion:

$$\mathcal{L}_{\text{HQ}} = \mathcal{L}^{\text{HQET}}_{m_Q\to\infty} + \sum_{n=1}^{\infty}\mathcal{L}_n^{\text{HQET}} \tag{20}$$

FIG. 1. In the vertex $V^{\mu,a}_{m_Q\to\infty}$, γ^μ is replaced by the velocity v^μ, while the propagator for a heavy quark has a pole for $v\perp k$ [13]



being

$$\mathcal{L}^{\rm HQET}_{(0),m_Q\to\infty} = \bar{h}_v (iv\cdot\mathcal{D}) h_v \qquad \text{and} \qquad \mathcal{L}^{\rm HQET}_{(n)} = \frac{1}{2m_Q}\bar{h}_v i\not{\mathcal{D}}_\perp \left(-\frac{iv\cdot\mathcal{D}}{2m_Q}\right)^{n-1} i\not{\mathcal{D}}_\perp h_v$$

By use of the identity:

$$\not{\mathcal{D}}_\perp \not{\mathcal{D}}_\perp = \mathcal{D}^2_\perp + \sigma_{\mu\nu}\left(\frac{g}{2}G^{\mu\nu}\right) \tag{22}$$

with $\sigma_{\mu\nu} = \frac{i}{2}\left[\gamma^\mu,\gamma^\nu\right]$ and $G^{\mu\nu} = ig_s\left[\mathcal{D}^\mu,\mathcal{D}^\nu\right]$ one finds that [13]:

$$\mathcal{L}^{\rm HQET}_{(1)} = \bar{h}_v (iv\cdot\mathcal{D}) h_v - \frac{1}{2m_Q}\bar{h}_v (i\mathcal{D}_\perp)^2 h_v - \frac{g_s}{4m_Q}\bar{h}_v \sigma_{\mu\nu} G^{\mu\nu} h_v + \mathcal{O}(1/m_Q^2) \tag{23}$$

In the limit $m_Q\to\infty$ only the first term remains

$$\mathcal{L}^{\rm HQET}_{(0),m_Q\to\infty} = \bar{h}_v (iv\cdot\mathcal{D}) h_v \tag{24}$$

this is the effective Lagrangian of HQET which, when compared with the full QCD one gives rise to the Feynman rules of fig.II:

D. Heavy Quark Spin-Flavour Symmetry

When studying the symmetries of this Lagrangian (24) one can see that there appear no Dirac matrices, thus interactions of the heavy quark with gluons leave its spin unchanged. In fact if we introduce a set of spin operators:

$$S^{(i)} = \frac{1}{2}\gamma^5 \not{v}\not{e}^{(i)} \tag{25}$$

Where $e^{(i)}\cdot e^{(j)} = -\delta_{ij}$ and $e^{(i)} = 0\ \forall i$, is a basis of spacelike components for the velocity v, one can verify by Dirac anticommutation that they obey the SU(2) algebra: $\left[S^{(i)},S^{(j)}\right] = i\epsilon^{ijk}S^{(k)}$. An infinitesimal SU(2) transformation $h_v \to (1+d\vec{\theta}\cdot\vec{S})h_v$ leaves (24) unchanged. If one considers the $1/m_Q$ terms appearing in in the effective Lagrangian (23), the first term

$$\mathcal{O}_{\rm kin} = \frac{1}{2m_Q}\bar{h}_v (i\mathcal{D}_\perp)^2 h_v$$

is the gauge-covariant extension of the kinetic energy arising from the residual motion of the heavy quark, in other words the term $\langle \mathrm{B}|\bar{h}_v(i\mathcal{D}_\perp)^2 h_v|\mathrm{B}\rangle$ tells us how much of the B meson energy is given by the Brownian motion of the b quark in B. This operator does not violate Heavy Quark Spin Symmetry [14]. The second operator is the non-Abelian analogue of the Pauli interaction [13], which describes the chromo-magnetic coupling of the heavy-quark spin to the gluon field

$$\mathcal{O}_{\mathrm{chromo}} = \frac{g_s}{4m_Q}\bar{h}_v(\sigma_{\mu\nu}G^{\mu\nu})h_v \to -\frac{g_s}{m_Q}\bar{h}_v(\vec{S}\cdot\vec{B}_c)h_v$$

where the i-th component of the chromomagnetic field $\vec{B}_c$ is given by $B_c^i = -\frac{1}{2}\epsilon^{ijk}G^{jk}$ and $\vec{S}$ is the spin operator defined in (25) [13]. The expectation value $\langle \mathrm{B}|\mathcal{O}_{chromo}|\mathrm{B}\rangle \neq 0$ since the light degrees of freedom have a net light quark flavour quantum number of 1 and thus $\mathrm{I} = \pm\frac{1}{2}$. This term is the first breaking the Heavy Quark Spin-Flavour Symmetry.

Another symmetry of the HQET Lagrangian (24) arises since the mass of the heavy-quark does not explicitly appear in the Lagrangian. For N_h heavy quarks moving at the same velocity v, eq. (24) can be extended by writing

$$\mathcal{L}_\infty^{(0)} = \sum_{i=1}^{N_h}\bar{h}_v^i(iv\cdot\mathcal{D})h_v^i$$

This Lagrangian is invariant under rotations in flavour space. When combined with the spin symmetry, the symmetry group is promoted to $\mathrm{SU}(2\mathrm{N}_h)$ which is called the heavy-quark spin-flavour symmetry group [13].

E. Loop corrections and reparametrization invariance

Equation (23) has been derived at tree level. Including loop corrections changes the Lagrangian to [14]:

$$\mathcal{L}_{\mathrm{eff}}^{(1)} = \bar{h}_v(iv\cdot\mathcal{D})h_v + \frac{1}{2m_Q}\bar{h}_v(i\mathcal{D}_\perp)^2 h_v + a(\mu)\frac{g_s}{4m_Q}\bar{h}_v\sigma_{\mu\nu}G^{\mu\nu}h_v + \mathcal{O}(1/m_Q^2) \tag{26}$$

The tree level matching with equation (23) implies that:

$$a(m_Q) = 1 + \mathcal{O}[\alpha_s(m_Q)]$$

The μ dependence of $\mathcal{O}_{\mathrm{chromo}}$ is canceled by the μ dependence of $a(\mu)$. In the leading logarithmic approximation [14]:

$$a(\mu) = \left[\frac{\alpha_s(m_Q)}{\alpha_s(\mu)}\right]^{9/(33-2N_q)}$$

where N_q is the number of light quark flavours. Loop effects do not change the coefficient of the heavy quark kinetic energy term due to reparametrization invariance. As previously stated , the choice of the residual momentum v^μ, the heavy-quark four-velocity is not unique, being $p_Q^\mu = m_Q v^\mu + k^\mu$ and $|k| \sim \Lambda_{\mathrm{QCD}} << m_Q$. A small change in v^μ of the order of $\Lambda_{\mathrm{QCD}}/m_Q$ will be compensated by a change in the residual momentum:

$$v^\mu \to v^\mu + \frac{1}{m_Q}\epsilon^\mu \qquad \text{produces a change in } k^\mu \text{ of} \qquad k^\mu \to k^\mu - \epsilon^\mu$$

One can see that being $v^2 = 1$, $v\cdot\epsilon = 0$ up to orders $\mathcal{O}(1/m_Q^2)$. In addition, since the heavy-quark spinor h_v must preserve the constraint $\not{v}h_v = h_v$, the variation of the heavy-quark spinor δh_v under a change in the heavy-quark four-velocity by an amount ϵ^μ/m_Q must satisfy:

$$(1-\not{v})\delta h_{v} = \frac{\not{\ell}}{m_{Q}} h_{v} \tag{28}$$

A suitable ansatz for a change in h_{v} is $\delta h_{v} = \frac{\not{\ell}}{2m_{Q}} h_{v}$ since $\not{v}\delta h_{v} = -\delta h_{v}$ and (28) is satisfied. Note however that this solution is not unique, we have chosen the one that preserves the norm of the $i(v\cdot\mathcal{D})$ term of the effective Lagrangian and other choices are equivalent to the above by a redefinition of the fields [14].

It can be shown that the $\mathcal{L}_{\rm HQET} = \mathcal{L}^{(0)} + \mathcal{L}^{(1)} + ...$ is invariant under the combined changes:

$$v^{\mu} \to v^{\mu} + \frac{1}{m_{Q}}\epsilon^{\mu} \qquad \text{and} \qquad h_{v} \to e^{i\epsilon\cdot x}\left(1 + \frac{\not{\ell}}{2m_{Q}}\right) h_{v}$$

the prefactor $e^{i\epsilon\cdot x}$ causing a shift in the residual momentum of $k^{\mu} \to k^{\mu} - \epsilon^{\mu}$. Thus the $\mathcal{L}_{\rm HQET} = \mathcal{L}^{(0)} + \mathcal{L}^{(1)} + ...$ is reparametrization invariant [14]. This would not be the case if the coefficient of the $\mathcal{O}_{\rm kin}$ deviated from 1, there can be no radiative corrections to $\mathcal{O}_{\rm kin}$ as long as the theory is regularized in a way that preserves reparametrization invariance and dimensional regularization is such a regulator, since the arguments made hold in n dimensions.

F. HQET Applications: heavy to heavy transitions

The elastic scattering of a B meson, $\mathrm{B}(v) \to \mathrm{B}(v')$ induced by a vector current coupled to the b quark can be described in the HQET framework through the following hadronic matrix element:

$$\frac{1}{m_{\rm B}}\langle \mathrm{B}(v')|\overline{\mathrm{b}}_{v'}\gamma^{\mu}\mathrm{b}_{v}|\mathrm{B}(v)\rangle = \xi(v\cdot v')(v+v')^{\mu} \tag{30}$$

Where b_{v} and $\overline{\mathrm{b}}_{v'}$ are the velocity dependent heavy-quark fields of HQET. In the limit of $m_{\rm b} \to \infty$, the form factor can only depend on $v\cdot v'$, reflecting the fact that, as the velocities become more and more different, the probability for an elastic transition decreases. Thus in this limit, the dimensionless probability function $\xi(v\cdot v')$, named Isgur-Wise function [14], describes the transition. No term proportional to $(v-v')^{\mu}$ appears in (30) since contracting the matrix element for the transition with such term must give zero given that $\not{v}\mathrm{b}_{v} = \mathrm{b}_{v}$ and $\overline{\mathrm{b}}_{v'}\not{v}' = \overline{\mathrm{b}}_{v'}$ and the conventional normalization of HQET meson states [14]

$$\langle \mathrm{B}(v)|\mathrm{B}(v')\rangle = 2v^{0}(2\pi)^{3}\delta^{3}(\vec{v}-\vec{v}') \tag{31}$$

When writing the above matrix element in terms of an elastic form factor $F_{el}(q^{2})$ depending only on the momentum transfer $q^{\mu} = (p-p')^{\mu}$ one finds

$$\langle \mathrm{B}(m_{\rm B}v')|\overline{\mathrm{b}}_{v'}\gamma^{\mu}\mathrm{b}_{v}|\mathrm{B}(m_{\rm B}v)\rangle = F_{\rm el}(q^{2})(p+p')^{\mu} \tag{32}$$

where $p^{(i)} = m_{\rm B}v^{(i)}$. When compared with (30) one finds that:

$$F_{el}(q^{2}) = \xi(v\cdot v') \qquad , \qquad q^{2} = 2m_{\rm B}^{2}(1-v\cdot v')$$

Because of current conservation the elastic form factor is normalized to unity at $q^{2} = 0$, which implies the normalization of the Isgur-Wise function at $v\cdot v' = 1$. This reflects the intuitive argument that the probability for an elastic

transition is unity if there is no velocity change between the two states, and since $v = v'$ implies that the final-state meson is at rest in the intial-meson rest frame, this point is referred to as zero-recoil limit.

Heavy-quark symmetry can be used to replace the b quark with a c in the final-state meson, therefore turning the B meson into a D meson, then the scattering process become a weak decay process. In the infinite mass limit, being the replacement $\mathrm{b}_{v'} \to \mathrm{c}_{v'}$ a symmetry transformation of the effective Lagrangian, the matrix element is still defined by the same function $\xi(v \cdot v')$

$$\langle \mathrm{D}(v')|\overline{\mathrm{c}}_{v'}\gamma^{\mu}\mathrm{b}_{v}|\overline{\mathrm{B}}(v)\rangle = \sqrt{m_{\mathrm{B}} m_{\mathrm{D}}}\xi(v \cdot v')(v + v')^{\mu}$$

But in general the matrix element of a flavour-changing current between two pseudoscalar mesons is described by two form factors which are conventionally referred to respectively as vector and scalar form factor

$$\langle \mathrm{D}(p')|\bar{\mathrm{c}}\gamma^{\mu}\mathrm{b}|\mathrm{B}(p)\rangle = f_{+}(q^2)(p + p')^{\mu} + f_0(q^2)\frac{m_{\mathrm{B}}^2 - m_{\mathrm{D}}^2}{q^2}q^{\mu} \tag{35}$$

Thus in the limit of $m_{\mathrm{b}}, m_{\mathrm{c}} \to \infty$ or equivalently at zeroth order in $1/m_{\mathrm{b}}, 1/m_{\mathrm{c}}$, HQET establishes a link between the Isgur-Wise function and the form factors, by comparing eq. (**??**) and eq. (35), one can see that

$$f_{+}(q^2) = \frac{m_{\mathrm{B}} + m_{\mathrm{D}}}{2\sqrt{m_{\mathrm{B}} m_{\mathrm{D}}}}\xi(w) \qquad \text{and} \qquad f_0(q^2) = \frac{(m_{\mathrm{B}} + m_{\mathrm{D}})^2 - q^2}{2(m_{\mathrm{B}} + m_{\mathrm{D}})\sqrt{m_{\mathrm{B}} m_{\mathrm{D}}}}\xi(w)$$

Where $w = v \cdot v'$ and $q^{\mu} = (p - p')^{\mu}$. From the equations above one can see that the two form factors are equal at $q^2 = 0$. This equality is valid at all orders of $1/m_{\mathrm{b}}, 1/m_{\mathrm{c}}$. Moreover, the normalization of the Isgur-Wise function at $w = 1$ (i.e. at zero-recoil) implies the normalization of the vector form factor at $q^2_{\mathrm{max}} = (m_{\mathrm{B}} - m_{\mathrm{B}})^2$ [13]

$$f_{+}(q^2_{\mathrm{max}}) = \frac{m_{\mathrm{B}} + m_{\mathrm{D}}}{2\sqrt{m_{\mathrm{B}} m_{\mathrm{D}}}} \tag{37}$$

This upper bound for the vector form factor receives non-perturbative corrections through a systematic expansion in $1/m_{\mathrm{B}}$ which is provided within the framework of the zero-recoil sum rule. In order to do so let's define the time ordered product: [15]

$$t^{\mu\nu} = -i\int d^4x e^{-iq\cdot x}\mathcal{T}\{\overline{\mathrm{b}}(x)\Gamma^{\nu\dagger}\mathrm{c}(x); \overline{\mathrm{c}}(0)\Gamma^{\mu}\mathrm{b}(0)\} \tag{38}$$

Where Γ^{μ} are the correct Dirac structures for a purely left-handed weak vertex. $\Gamma^{\mu} = \gamma^{\mu}(1 - \gamma^5)$. The $T^{\mu\nu}$ is defined as in [15]

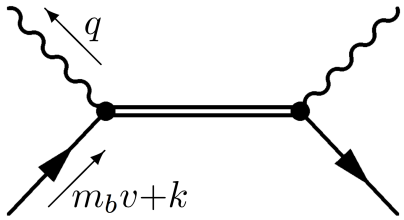
$$T^{\mu\nu}(v \cdot q) = \frac{1}{2m_{\mathrm{B}}}\langle \mathrm{B}|t^{\mu\nu}|B\rangle \tag{39}$$

It's imaginary part is related to the hadronic tensor appearing in eq. (8) by the optical theorem which takes the form [15]

$$H^{\mu}H^{*\nu} = 2\,\mathrm{Im}\,T_{\mu\nu} \tag{40}$$

The tree-level nonperturbative expansion in $1/m_{\mathrm{b}}$ is set up by looking at the Feynman Diagram in fig. (III F). The double line denotes the Green function of the charm quark propagating in the background field of soft gluons in the B meson. By rephasing the fields $\mathrm{b}(x) \to e^{-im_{\mathrm{b}}(v\cdot x)}\mathrm{b}_{v}(x)$, the momentum operator of the b-quark becomes

FIG. 2. Tree level diagram for the hadronic tensor in inclusive semileptonic decays [13]



$$p_{\rm b} = m_{\rm b} v + i\mathcal{D} \tag{41}$$

With the term $i\mathcal{D}$ playing the role of the 'residual' momentum. Thus we write

$$t^{\mu\nu} = -i \int d^4x e^{-iq\cdot x} e^{im_{\rm b}(v\cdot x)} \mathcal{T}\{\overline{\rm b}_v(x)\Gamma^{\nu\dagger}{\rm c}(x); \overline{\rm c}(0)\Gamma^{\mu}{\rm b}_v(0)\} \tag{42}$$

The phase factor from the rephased b-quark field combines with the c-quark propagator to yeld the background-field charm propagator [15]

$$t^{\mu\nu} = -i \int d^4x e^{-iq\cdot x} e^{-i(m_{\rm b} v - q + i\mathcal{D})\cdot x} \mathcal{T}\{\overline{\rm b}_v(0)\Gamma^{\nu\dagger}{\rm c}(0); \overline{\rm c}(0)\Gamma^{\mu}{\rm b}_v(0)\} \tag{43}$$

$$S_{\rm BGF} = \frac{1}{\not p + i\not{\mathcal{D}} - m_{\rm c}} \qquad \text{with} \qquad p_\mu = m_{\rm b} v_\mu - q_\mu$$

For semileptonic processes at tree level, one only needs to multiply the $S_{\rm BGF}$ by the appropriate Dirac matrices for the left handed current. A calculation of the OPE series to order $1/m_{\rm b}^n$ requires an expansion of $S_{\rm BGF}$ to the $n^{\rm th}$ order in the covariant derivative $i\mathcal{D}$ according to [15]

$$S_{\rm BGF} = \frac{1}{\not p - m_{\rm c}} + \frac{1}{\not p - m_{\rm c}}(-i\not{\mathcal{D}})\frac{1}{\not p - m_{\rm c}} + \frac{1}{\not p - m_{\rm c}}(-i\not{\mathcal{D}})\frac{1}{\not p - m_{\rm c}}(-i\not{\mathcal{D}})\frac{1}{\not p - m_{\rm c}} + \dots \tag{45}$$

The covariant derivatives do not commute in general, and the above expansion takes care of their ordering. The first term of this expansion corresponds to operators of dimension 3 in the OPE and takes care of corrections of zeroth order $1/m_{\rm b}$ to $T^{\mu\nu}$. They are calculated by evaluating the structure

$$\overline{\rm b}\gamma^\mu(1-\gamma^5)\frac{1}{\not p - m_{\rm c}}\gamma^\nu(1-\gamma^5){\rm b} \tag{46}$$

Which by the use of the hadronic matrix elements of the operators, [14]

$$\begin{aligned} \langle {\rm B}|\overline{\rm b}\gamma^\mu{\rm b}|B\rangle &= 2p_{\rm B}^\mu = 2m_{\rm B} v^\mu \\ \langle {\rm B}|\overline{\rm b}\gamma^\mu\gamma^5{\rm b}|B\rangle &= 0 \end{aligned} \tag{47}$$

The second one being zero due to parity conservation of strong interactions, give the zeroth order in the expansion of $T^{\mu\nu}$

$$T^{\mu\nu}_{(0)} = \frac{2}{\Delta}\left[2m_{\rm b}v^{\mu}v^{\nu} - g^{\mu\nu}(m_{\rm b} - v\cdot q) - (q^{\mu}q^{\nu} + q^{\nu}q^{\mu})\right] \tag{48}$$

Where $\Delta = m_{\rm b}^2 - m_{\rm c}^2 + q^2 - 2m_{\rm b}(v\cdot q)$

It is important to note that equations (47) are exact, there hasn't yet been the transition to the HQET fields. The second term of the expansion (45) will yeld terms of order $1/m_{\rm b}$ to $T^{\mu\nu}$. By evaluating the structure

$$T^{\mu\nu}_{(1)} = -\frac{1}{\Delta^2}\overline{\rm b}\gamma^{\mu}(1-\gamma^5)\not{p}(i\not{\mathcal{D}})\not{p}\gamma^{\nu}(1-\gamma^5){\rm b} \tag{49}$$

For the vectorial current, one will encounter terms of the form

$$\langle {\rm B}|\overline{\rm b}\gamma^{\mu}(i\mathcal{D}^{\nu}){\rm b}|B\rangle \tag{50}$$

When including $1/m_{\rm b}$ corrections to the relationship between QCD and HQET operators, dimension 4 and dimension 5 operators appear in the expansion, to order α_s^0 [14]

$${\rm b}(x) = e^{-im_{\rm b}v\cdot x}\left(1 + \frac{i\not{\mathcal{D}}}{2m_{\rm b}}\right){\rm b}_v(x) \tag{51}$$

Recalling eq. (26) one can write the operator matching condition that links QCD and HQET to $1/m_{\rm b}$

$$\overline{\rm b}\gamma^{\mu}(i\mathcal{D}^{\nu}){\rm b} = \overline{\rm b}_v\gamma^{\mu}i\mathcal{D}^{\nu}{\rm b}_v + \overline{\rm b}_v\left(-\frac{i\overleftarrow{\not{\mathcal{D}}}}{2m_{\rm b}}\right)\gamma^{\mu}(i\mathcal{D}^{\nu}){\rm b}_v + \overline{\rm b}_v\gamma^{\mu}(i\mathcal{D}^{\nu})\left(\frac{i\not{\mathcal{D}}}{2m_{\rm b}}\right){\rm b}_v + i\int d^4x\mathcal{T}\{\overline{\rm b}_v\gamma^{\mu}i\mathcal{D}^{\mu}{\rm b}_v(0); \mathcal{L}_1(x)\} \tag{52}$$

To this order the heavy-quark expansion contains $\mathcal{O}_4$ and $\mathcal{O}_5$ HQET operators. It can be shown that $\mathcal{O}_4$ operators do not contribute to this expansion, being those of the form

$$\langle {\rm B}|\overline{\rm b}_v(i\mathcal{D}^{\mu}){\rm b}_v|B\rangle = Av^{\mu} \tag{53}$$

By contracting this structure with v_{μ} and by the use of the equations of motion of HQET $i(\mathcal{D}\cdot v){\rm b}_v = 0$ one finds that A = 0

The time-ordered product appearing in (52) is evaluated by noting that

$$i\int d^4x\mathcal{T}\{\overline{\rm b}_v v^{\mu}i\mathcal{D}^{\nu}{\rm b}_v(0); \mathcal{L}_1(x)\} = Xv^{\mu}v^{\nu} \tag{54}$$

Which by repeated contraction with v_{μ} yields

$$i\int d^4x\mathcal{T}\{\overline{\rm b}_v(i\mathcal{D}\cdot v){\rm b}_v(0); \mathcal{L}_1(x)\} = X \tag{55}$$

At order α_s^0 eq. (55) is evaluated by the use of

$$(\mathcal{D}\cdot v)S_{\rm HQET}(x-y) = \delta^4(x-y) \tag{56}$$

Yielding

$$X = \langle \mathrm{B}(v)|\mathcal{L}(0)|\mathrm{B}(v)\rangle = \langle \mathrm{B}(v)|\overline{\mathrm{b}}_v \frac{(\mathcal{D}_\perp)^2}{2m_\mathrm{b}} \mathrm{b}_v|\mathrm{B}(v)\rangle + g_s\langle \mathrm{B}(v)|\overline{\mathrm{b}}_v \frac{\sigma^{\alpha\beta}G_{\alpha\beta}}{4m_\mathrm{b}} \mathrm{b}_v|\mathrm{B}(v)\rangle \tag{57}$$

Adopting the same notation as in [16]

$$\mu_G^2 = \frac{1}{2m_\mathrm{B}}\langle \mathrm{B}|\mathrm{b}\frac{i}{2}\sigma_{\mu\nu}G^{\mu\nu}\mathrm{b}|\mathrm{B}\rangle \quad , \qquad \mu_\pi^2 = \frac{-1}{2m_\mathrm{B}}\langle \mathrm{B}|\mathrm{b}(i\mathcal{D}_\perp)^2\mathrm{b}|\mathrm{B}\rangle \tag{58}$$

The X coefficient of (57) can be evaluated to give

$$\frac{1}{2m_\mathrm{B}}X = \frac{\mu_G^2 - \mu_\pi^2}{2m_\mathrm{b}}$$

The second and third terms of (52) are evaluated by noting that:

$$\overline{\mathrm{b}}_v \left(\frac{i\not{\mathcal{D}}_\perp}{2m_\mathrm{b}}\right)\gamma^\mu(i\mathcal{D}_\perp^\nu)\mathrm{b}_v + \overline{\mathrm{b}}_v\gamma^\mu(i\mathcal{D}_\perp^\nu)\left(\frac{i\not{\mathcal{D}}_\perp}{2m_\mathrm{b}}\right)\mathrm{b}_v = \overline{\mathrm{b}}_v\left[\frac{i\mathcal{D}_\perp^{\{\mu}i\mathcal{D}_\perp^{\nu\}}}{2m_\mathrm{b}} + \frac{ig_s}{4m_\mathrm{b}}\sigma_\alpha{}^\mu G^{\alpha\nu}\right]\mathrm{b}_v \tag{59}$$

The first of the terms on the r.h.s of the above equation can be decomposed in Lorentz covariant structures as follows

$$\langle \mathrm{B}|\overline{\mathrm{b}}_v i\mathcal{D}^{\{\mu}i\mathcal{D}^{\nu\}}\mathrm{b}_v|B\rangle = Y(g^{\mu\nu} - v^\mu v^\nu) \tag{60}$$

Since it must vanish when contracted with either v^μ or v^μ due to equations of motion of HQET. By contraction with the metric tensor eq. (60) yields

$$\langle \mathrm{B}|\overline{\mathrm{b}}_v(i\mathcal{D}_\perp)^2\mathrm{b}_v|B\rangle = 3Y \tag{61}$$

From which one can write

$$\frac{1}{2m_\mathrm{B}}Y = -\frac{\mu_\pi^2}{3} \tag{62}$$

Analogously

$$\langle \mathrm{B}|\overline{\mathrm{b}}_v\sigma_\alpha{}^\mu G^{\alpha\nu}\mathrm{b}_v|B\rangle = Z(g^{\mu\nu} - v^\mu v^\nu) \tag{63}$$

due to the identity $\overline{\mathrm{b}}_v\sigma_\alpha{}^\mu v_\mu \mathrm{b}_v = 0$ [14].
From which we derive

$$\frac{1}{2m_\mathrm{B}}Z = \frac{2}{3}\mu_G^2 \tag{64}$$

So, at order $1/m_\mathrm{b}$ one can write:

$$T_{(1)}^{\mu\nu} = T_{(0)}^{\mu\nu} + T_{(1)}^{\mu\nu} \tag{65}$$

Where

$$T^{\mu\nu} = \frac{-2p^2}{\Delta^2}\left[\frac{2}{3}\left(\frac{\mu_\pi^2 - \mu_G^2}{m_{\rm b}}\right)v^\mu v^\nu \ + \frac{1}{3}\left(\frac{\mu_\pi^2 - \mu_G^2}{m_{\rm b}}\right)g^{\mu\nu}\right] \tag{66}$$

Eq. (39) can be alternatively expanded by inserting a complete set of hadronic intermediate states in the time-ordered product. By defining the operator $J^\mu_{(i)}(x) = \bar{\rm c}\Gamma^\mu_{(i)}{\rm b}(x)$ where $i = V, A$ corresponds repectively to the vector and axial-vector current, the time ordered product is expanded as follows:

$$\begin{aligned} T_i^{\mu\nu} = \sum_X \frac{1}{2m_{\rm B}}\int d^4x \Big[& e^{i(p_{\rm B}-p_X)\cdot x}\langle {\rm B}|{J^\mu_{(i)}}^\dagger(0)|X\rangle\langle X|J^\nu_{(i)}(0)|{\rm B}\rangle\theta(x_0) \\ & - e^{-i(p_{\rm B}-p_X)\cdot x}\langle {\rm B}|J^\nu_{(i)}(0)|X\rangle\langle X|{J^\mu_{(i)}}^\dagger(0)|{\rm B}\rangle\theta(-x_0) \end{aligned} \tag{67}$$

Where the matrix elements have been shifted to the same point by the use of

$$\langle {\rm B}|J^\mu(x)|X\rangle = \langle {\rm B}|J^\mu(0)|X\rangle e^{i(p_{\rm B}-p_X)\cdot x} \tag{68}$$

And its conjugate. The sum over the intermediate states is now independent of x, so the integral can be evaluated in the B rest frame by making use of the integral representation of the Heaviside theta function as in [14]

$$\theta(x_0) = \frac{1}{2\pi i}\lim_{\delta\to 0}\int_{-\infty}^{+\infty}\frac{e^{-iwx_0}}{w+i\delta} \tag{69}$$

yielding

$$\begin{aligned} T_i^{\mu\nu} = & \sum_{X_{\rm c}}\frac{1}{2m_{\rm B}}\frac{\langle {\rm B}|J_i^{\mu\dagger}(0)|X_{\rm c}\rangle\langle X_{\rm c}|J_i^\nu(0)|{\rm B}\rangle}{m_{\rm B} - E_{X_{\rm c}} - (v\cdot q)}(2\pi)^3\delta^3(\vec{q}+\vec{p}_{X_{\rm c}}) \\ & + \sum_{X_{\bar{\rm c}{\rm b}{\rm b}}}\frac{\langle {\rm B}|J_i^\nu(0)|X_{\bar{\rm c}{\rm b}{\rm b}}\rangle\langle X_{\bar{\rm c}{\rm b}{\rm b}}|J_i^{\mu\dagger}(0)|{\rm B}\rangle}{E_{X_{\bar{\rm c}{\rm b}{\rm b}}} - m_{\rm B} - (v\cdot q)}(2\pi)^3\delta^3(\vec{q}-\vec{p}_{X_{\bar{\rm c}{\rm b}{\rm b}}}) \end{aligned} \tag{70}$$

Where the first summatory includes resonant states, while the second contains inelastic contributions. Introducing the variable $\varepsilon = m_{\rm b} - m_{\rm c}w + (v\cdot q)$, where w is the scalar product of the four-velocities v and v' of the B and D mesons respectively.
Passing to the zero recoil limit and contracting with $g_{\mu\nu}$, eq. (70) yields

$$T_{i\ \ \mu}^{\ \mu} = \frac{1}{2m_{\rm B}(\varepsilon - (E - m_{\rm c}))}(2\pi)^3\delta^{(3)}(\vec{q}+\vec{p}_{\rm D})|\langle {\rm B}|J_i^\mu(0)|{\rm D}\rangle|^2_{\rm ZR} + \sum_{\text{resonant } {\rm X_c}\neq {\rm D}} + \sum_{\rm inel} \tag{71}$$

It is found that in the case of a D pseudoscalar meson as a final state only the vector current $J^\mu_{\rm V}$ contributes to the ZRSR, due to parity invariance of QCD interactions.

The zero-recoil sum rule is obtained by equating two representations of the contour integral

$$I_n = \frac{-1}{2\pi i}\oint_{\varepsilon=|\varepsilon_M|} d\varepsilon\ \varepsilon^n T^\mu_{\ \mu} \tag{72}$$

Where $|\varepsilon_{M}|$ is some arbitrary parameter chosen in such a way that only the components of the sum (71) that we are interested in are taken into account, leaving out the resonances situated at higher energy. In our case $0 < \varepsilon_{M} < m_{\mathrm{D}} - m_{\mathrm{D}^*}$.

Moreover it is interesting to note that, using the decomposition of (35) the ZRSR directly constrains only the scalar form factor being

$$|\langle \mathrm{B}|J^{\mu}_{\mathrm{V}}(0)|\mathrm{D}\rangle|^2_{ZR} = (m_{\mathrm{B}} + m_{\mathrm{D}})^2 |f_0(q^2 = q^2_{\mathrm{max}})| \tag{73}$$

although the literature [17] relates the ZRSR to the value of $f_{+}(q^2 = q^2_{\mathrm{max}})$. The two representations coincide when the f_0 is expressed in terms of f_{+} through the Isgur-Wise function (36). This choice has a historic motivation: since the ZRSR for the B to D transition was at first motivated by the measurement of $|V_{\mathrm{cb}}|$, the interest on f_0 grew when the $R_{\mathrm{D}^{(*)}}$ discrepancies where first measured and an better estimate of f_0 to evaluate the size of the τ contribution to the ratio was needed.

G. Results for the Sum Rule at Zero Recoil

The literature [17] lists non-perturbative corrections up to $\mathcal{O}(1/m^3_{\mathrm{b}})$ and perturbative corrections up to $\mathcal{O}(\alpha_s)$.

$$\mathcal{F}_{+} = \frac{2(m_{\mathrm{B}} m_{\mathrm{D}})^{1/2}}{m_{\mathrm{B}} + m_{\mathrm{D}}} f_{+}(q^2_{\mathrm{max}}) \leq \xi^{\mathrm{pert}}_{V}(\mu) - \frac{\mu^2_{\pi} - \mu^2_{G}}{4}\left(\frac{1}{m_{\mathrm{c}}} - \frac{1}{m_{\mathrm{b}}}\right)^2 - \frac{\rho^3_{D} - \rho^3_{LS}}{4}\left(\frac{1}{m_{\mathrm{c}}} + \frac{1}{m_{\mathrm{b}}}\right)\left(\frac{1}{m_{\mathrm{c}}} - \frac{1}{m_{\mathrm{b}}}\right)^2 \tag{74}$$

where $\xi^{\mathrm{pert}}_{\mu}$ of eq. (27) of ([17]) includes the order $\mathcal{O}(\alpha_s)$ to the vector current.

A fit to the various kinematic moments of the inclusive $\mathrm{B} \to X_{\mathrm{c}} \ell \nu$ yields [18]

$$\begin{aligned} m^{\mathrm{kin}}_{b}(1\mathrm{GeV}) &= (4.561 \pm 0.021)\,\mathrm{GeV}\,, \qquad m^{\mathrm{kin}}_{c}(1\mathrm{GeV}) = (1.092 \pm 0.020)\,\mathrm{GeV}\,, \\ \mu^2_{\pi}(1\mathrm{GeV}) &= +(0.464 \pm 0.067)\,\mathrm{GeV}\,, \qquad \mu^2_{G}(1\mathrm{GeV}) = +(0.333 \pm 0.061)\,\mathrm{GeV}\,, \\ \rho^3_{D}(1\mathrm{GeV}) &= +(0.175 \pm 0.040)\,\mathrm{GeV}\,, \qquad \rho^3_{LS}(1\mathrm{GeV}) = -(0.146 \pm 0.096)\,\mathrm{GeV}\,, \end{aligned} \tag{75}$$

whose central values have been used as an input to evaluate $\xi^{pert}_{V}(\mu = 1GeV)$. The errors on ξ^{pert}_{V} are assessed by varying the scale $0 \leq \mu \leq 2GeV$

$$\xi^{pert}_{V}(\mu = 1GeV) = 1.038^{0.021}_{-0.011} \tag{76}$$

We obtained the upper limit on $f_{+}(q^2_{\mathrm{max}})$ from a gaussian fit run on a 10^6 events sample, obtained by normally variating the input parameters listed in (75) and the ξ^{pert}_{V} found in (76). The fit is displayed in figure (3).

The obtained value of upper bound to f_{+} at the zero recoil kinematic point is

$$f_{+}(q^2 = q^2_{\mathrm{max}}) \leq 1.159 \pm 0.024 \tag{77}$$

with the meson's masses used as input listed in table (I).

If we compare this result with the most recent Latice QCD results of [24] , $f_{+}(q^2_{\mathrm{max}}) = 1.1775(45)$, which is calculated using the central values for the Lattice QCD parameters and a normal variation of the errors listed in the paper, we find that the latter saturate the zero-recoil sum rule upper bounds, leaving almost no room for inelastic contributions.

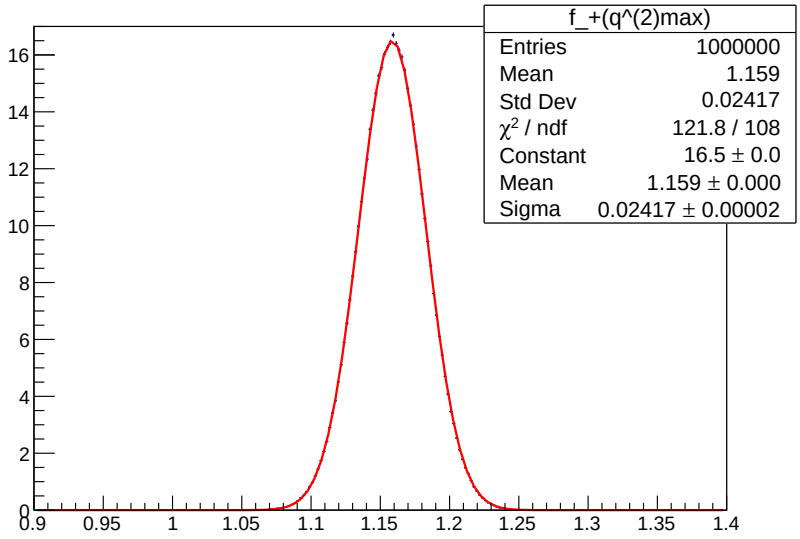


FIG. 3. Fitted value of $f_+(q^2 = q^2_{\rm max})$ from a 10^6 events sample generated from normal variation of parameters in (75, 76).

TABLE I. Values for the mesons masses used to calculate (77)

Type	Mass (GeV)	Ref
m_B	5.279	[19]
m_D	1.870	[19]

IV. THE BCL PARAMETRIZATION

A. The z-expansion

The parametrization of the form factors $f^{B\to D}_{+,0}(q^2)$ has been the subject of intense investigation, motivated in particular by the need to extrapolate the information obtained in a restricted q^2 to the whole q^2 range since lattice QCD calculations and the zero-recoil sum rule are limited to the highest q^2 values. The knowledge of f_+ and f_0 in the whole kinematic range $m^2_\mu \leq q^2 \leq (m_B - m_D)^2$ allows for the calculation of $R(D)$.

The BCL parametrization was originally introduced in [22] in order to model-independently describe the form factors involved in the transition $B \to \pi\ell\nu$. It consists of a systematic expansion which incorporates the constraints of analitycity, unitarity and crossing symmetry. By an analogue approach, one can obtain a parametrization for the $f^{B\to D}_{+,0}(q^2)$ by exploiting the pole structure and positivity of the correlator (39).

In the case of semileptonic B decays q^2 ranges from m^2_ℓ to $(m_B - m_D)^2$ but the form factors can be continued analytically in the q^2 complex plane. They have a cut at $q^2 = (m_B + m_D)^2$ and various lower lying poles corresponding to B_c resonances with the appropriate quantum numbers [23]. By adopting the same notation as in [22] the following variables are defined:

$$s = q^2 = (p - p')^2, \qquad s_+ = (m_B + m_D)^2, \qquad s_- = (m_B - m_D)^2$$

$$z(s, s_0) = \frac{\sqrt{s_+ - s} - \sqrt{s_+ - s_0}}{\sqrt{s_+ - s} + \sqrt{s_+ - s_0}} \tag{78}$$

Where $z(s, s_0)$ conformally maps the q^2 plane cut for $q^2 \geq s_+$ onto a unit disk $|z(s, s_0)| \leq 1$ in the z complex variable, such that $z(s_+, s_0) = -1$ and $z(\infty, s_0) = 1$. The arbitrary parameter $s_0 < s_+$ determines the point of q^2 mapped onto the origin in the z plane, i.e. $z(s_0, s_0) = 0$.

With the choice $s_0 = s_- = (m_B - m_D)^2$ as in [23] the range of variation of z is $0 \leq z \leq 0.0644$ and the point of zero recoil is at $z = 0$.

Alternatively s_0 can be chosen in such a way that $|z_{min}| = |z_{max}|$, which in the limit of massless leptons reads $s_0 = s_m = s_+ - N(s_+ - s_-)$ where $N = (1 + r_D)/(2\sqrt{r_D})$. With this prescription $|z_{max}|$ is minimized to $|z_{max}| \sim 0.032$. The form factors can thus be expanded as in [23]:

$$f_+(q^2) = \frac{1}{1 - q^2/m_+^2} \sum_{k=0}^{k=N_+} a_k z(s, s_-)^k, \qquad f_0(q^2) = \frac{1}{1 - q^2/m_0^2} \sum_{k=0}^{k=N_0} b_k z(s, s_-)^k \tag{79}$$

Where $m_{+,0}$ are the masses of the lowest lying B_c resonances in the vector and scalar channels, their values and the corresponding references are listed in table (II). This terms embed the distant poles at $q^2 = m_{+,0}^2$. No further resonances below threshold are considered relevant to this expansion [23].

TABLE II.

Type	Mass(GeV)	Ref
1^-	6.329(3)	[19] [20]
0^+	6.716	[21]

Analyticity implies that near threshold

$$\mathrm{Re}\, f_+(q^2) \sim c_+^1 + c_+^2 (s - s_+)$$

where c_+^1 and c_+^2 are constants. Moreover, angular momentum conservation imposes the behaviour

$$\mathrm{Im}\, f_+(q^2) \sim (q^2 - s_+)^{3/2}$$

near threshold. Since the threshold s_+ is mapped onto the point $z = -1$ and $(z + 1) \sim \text{const.} \times (q^2 - s_+)^{1/2}$, f_+ near $z = -1$ can be written as

$$f_+(z)|_{z=-1} = c_+^1 + c_+^2 (z + 1)^2 + i\ \text{const} \times (z + 1)^3$$

And must satisfy [22]:

$$\left[\frac{df_+}{dz}\right]_{z=-1} = 0 \tag{80}$$

By plugging eq. (80) in eq. (79) one can express the N_+^{th} order parameter of the vector form factor expansion as a function of the remaining $N_+ - 1$ parameters obtaining the relation [22]:

$$a_{N_+} = -\frac{(-1)^{N_+}}{N_+} \sum_{k=0}^{N_+-1} (-1)^k \, k a_k \tag{81}$$

Another constrain on the parameters of this expansion derives from the equality $f_+(0) = f_0(0)$, (see eq. (36)), which, when plugged in eq. (79) constrains the N_0^{th} parameter of the scalar form factor as:

$$b_{N_0} = \sum_{k=0}^{N_+-1} a_k \left[z_0^{k-N_0} - (-1)^{k-N_+} \frac{k}{N_+} z_0^{N_+-N_0} \right] - \sum_{k=0}^{N_0-1} b_k z_0^{k-N_0} \tag{82}$$

Where $z_0 = z(0, s_0)$. From the above relations it follows that:

$$\begin{aligned} f_+(q^2) &= \frac{1}{1 - q^2/m_+^2} \sum_{k=0}^{N_+-1} a_k \left[z^k - (-1)^{k-N_+} \frac{k}{N_+} z^{N_+} \right] \\ f_0(q^2) &= \frac{1}{1 - q^2/m_0^2} \left[\sum_{k=0}^{N_0-1} b_k + z^{N_0} \left(\sum_{k=0}^{N_+-1} a_k \left[z_0^{k-N_0} - (-1)^{k-N_+} \frac{k}{N_+} z_0^{N_+-N_0} \right] - \sum_{k=0}^{N_0-1} b_k z_0^{k-N_0} \right) \right] \end{aligned} \tag{83}$$

In principle any quantity depending on the two form factors, when those are expanded respectively to the N_+^{th} and N_0^{th} order, would contain $N_0 + N_+ + 2$ free parameters, which are reduced to at most $N_0 + N_+$ by the implementation of the above constraints. Moreover, the way those constraints are implemented is not unique, as long as the constraints are fulfilled at all orders of the expansion.

B. Unitarity Bounds

The main advantage of the BGL class parametrizations, such as the BCL parametrization, is that its coefficients can be explicitly constrained by unitarity bounds. Since the correlator (39) is a sum of positive contributions, the assumption that it is saturated by BD intermediate states, unitarity and crossing symmetry guarantee that the coefficients of the BGL expansion:

$$\begin{aligned} f_+(q^2) &= \frac{1}{B_+(s)\phi_+(s, s_0)} \sum_{k=0}^{\infty} \tilde{a}_k z(s, s_0)^k \\ f_0(q^2) &= \frac{1}{B_0(s)\phi_0(s, s_0)} \sum_{k=0}^{\infty} \tilde{b}_k z(s, s_0)^k \end{aligned} \tag{84}$$

Are upper bounded by:

$$\sum_{k=0}^{\infty} \tilde{a}_k^2 \leq 1, \qquad \sum_{k=0}^{\infty} \tilde{b}_k^2 \leq 1 \tag{85}$$

Where the *Blaschke factors* $B_{+,0}(s) = z(s, m_{+,0}^2)$ embed the poles that lie above the semileptonic domain of the B_c vector and scalar resonances below threshold; and the *outer functions* $\phi_{+,0}(s, s_0)$ reflect the way in which the form factors enter the dispersive integral (eq. (2.7) of [23]). Their normalization is set to depend on the q^2 derivatives

of the longitudinal and transverse component of (39), evaluated at $q^2 = 0$ in order to give the unitarity bound the form (85).

The unitarity condition (85) can also be expressed in terms of the coefficients a_k and b_k of eq. (79). By comparing representations (79) and (84), we have for the vector form factor [22]:

$$\sum_{k=0}^{\infty} \tilde{a}_k z^k = \Psi^+(z) \sum_{k=0}^{N_+} a_k z^k \tag{86}$$

Where [23]

$$\Psi^+(z) = \frac{B_+(z)\phi_+(z, s_0)}{1 - s(z)/m_+^2}$$

depends both on z and s_0 and is analytic in $|z| < 1$. Thus it can be expanded around $z = 0$ as:

$$\Psi^+(z) = \sum_{k=0}^{\infty} \eta_k^+ z^k \tag{87}$$

By inserting this expansion in (86) the following equation is obtained [22]:

$$\tilde{a}_n = \sum_{k=0}^{\min[N_+, n]} \eta_{n-k}^+ a_k, \qquad n \geq 0 \tag{88}$$

Which, when inserted in (85), gives an upper bound for the coefficients of the BCL parametrization that reads:

$$\sum_{j,k=0}^{N_+} A_{j,k} a_j a_k \leq 1, \qquad A_{j,k} = \sum_{n=0}^{\infty} \eta_n^+ \, \eta_{n+|j-k|}^+ \tag{89}$$

It follows from the definition that $A_{i,j} = A_{j,i}$, $A_{i,(i+j)} = A_{0,j}$ and $A_{i,i} = A_{i+1,i+1}$ for $0 < i < N_+ - 1$.

By virtue of these properties it will suffice to determine $A_{0,i}$ for $i < N_+ - 1$. Similar relations hold for the scalar form factor and one can define $B_{i,j}$ in a similar fashion as for the vectorial case. Including the masses of the first vectorial and scalar resonances at $\mathrm{B_c}(1^-)$ and $\mathrm{B_c}(0^+)$ of table and by the use of the outer functions as expressed in [23] in this work we calculate the values for $A_{0,i}$ and $B_{0,j}$ and they are listed in Table (III).

TABLE III. Values of the first five coefficients $A_{i,j}$ and $B_{i,j}$, the summation over n is performed up to $n = 20$

$s_0(\mathrm{GeV}^2)$	$A_{0,0}$	$A_{0,1}$	$A_{0,2}$	$A_{0,3}$	$A_{0,4}$	$A_{0,5}$
s_-	0.0199	-0.0037	-0.0122	0.0051	0.0018	-0.0011
s_m	0.0168	-0.0034	-0.0101	0.0046	0.0012	0.0009

$s_0(\mathrm{GeV}^2)$	$B_{0,0}$	$B_{0,1}$	$B_{0,2}$	$B_{0,3}$	$B_{0,4}$	$B_{0,5}$
s_-	0.2920	-0.1698	-0.0376	0.1189	-0.0873	0.0438
s_m	0.2594	-0.1552	-0.0250	0.1021	-0.0811	0.0447

While in principle the choice of $s_0 = s_m$, by reducing the range in z forces a faster convergence of the z-expansion, for the rest of the work we have chosen $s_0 = s_-$ in order to be consistent with the parametrization of [24], since no appreciable change in sensitivity of the fit has been noticed.

C. Upper bound on error from series truncation

Unitarity bounds allow one to derive an upper limit on the remainder of expansion (83). For the vectorial form factor, for example, by writing the remainder of the expansion as:

$$\delta f_+(q^2) = \frac{1}{1 - q^2/m_+^2} \sum_{k=N_++1}^{\infty} a_k z^k \tag{90}$$

Eq. (86) can be used to express each coefficient a_k as:

$$a_k = \sum_{j=0}^{k} \tilde{\eta}_{k-j}^{+} \tilde{a}_j, \qquad k \geq 0 \tag{91}$$

where the $\tilde{\eta}_j$ are the coefficients of the expansion:

$$1/\Psi^+(z) = \sum_{j=0}^{\infty} \tilde{\eta}_j^+ z^j \tag{92}$$

By the use of the Cauchy inequality we can write an upper bound for the a_k^{th} term of the expansion in terms of the coefficients $\tilde{a}_k$ and $\tilde{\eta}_k$ as:

$$|a_k| \leq \left(\sum_{j=0}^{k} \tilde{\eta}_j^{+2} \sum_{j=0}^{k} \tilde{a}_j^2 \right)^{1/2}, \qquad k \geq 0 \tag{93}$$

That can be simplified by using the first of (85)

$$|a_k| \leq \left(\sum_{j=0}^{k} \tilde{\eta}_j^{+2} \right)^{1/2} \tag{94}$$

Which establishes an upper bound for the remainder (90) in terms of calculable quantities:

$$|\delta f_+(q^2)| \leq \frac{1}{|1 - q^2/m_+^2|} \sum_{k=N_++1}^{\infty} \left(\sum_{j=0}^{k} \tilde{\eta}_j^{+2} \right)^{1/2} |z|^k \tag{95}$$

As observed in [22] this upper bound can be made sufficiently small for a certain N_+ on $|z| < 1$. This follows from the analyticity of $1/\Psi^+(z)$ inside the disk $|z| < 1$. Therefore, although the Taylor coefficients $\tilde{\eta}_j^{+2}$ increase with j, the increase is such that the sum $\{\sum_{j=0}^{k} \tilde{\eta}_j^{+2}\}^{1/2} |z|^k$ can be made arbitrarily small for a certain N_+ and $|z| < 1$.

V. THE FIT

A. Method

Now that a parametrization of the form factors has been introduced, our goal will be to establish if, with the present LHCb detector capabilites, the scalar form factor can be measured directly from data, and if yes with how much sensitivity. As a first study, perfect resolution is assumed, while a realistic reconstruction efficiency and finite momentum resolution are introduced in chapter (VI).
The current values for the coefficients of the BCL expansion for the scalar form factor of the $\overline{\mathrm{B}} \to \mathrm{D}\ell\bar{\nu}$ transition $f_0(q^2)$ are obtained exclusively from Lattice QCD and the latest results are listed in [24]. The fit strategy consists of the following steps:

1. Generate several data samples with known parameters
2. Fit to the differential decay width probability density function (which, from now on, will be referred to as PDF) to retrieve the parameters the sample was generated with.
3. Perform statistical analysis of the retrieved parameters

B. Monte Carlo generation

Though the m_ℓ^2 factor multiplying f_0 would favour the $\overline{\mathrm{B}} \to \mathrm{D}\ell\bar{\nu}$ channel as far as the sensitivity to its parameters is concerned, the low statistic and the deteriorated resolution due to the missing neutrinos' energy makes this an unviable strategy at hadron facilities. On the other hand, the muon channel is characterised by a strong suppression of the scalar form factor, but the number of collected events (presently $\mathcal{O}(10^6)$), which will grow during RUN II phase, could compensate for such suppression. In this chapter we'll first try to determine how much statistic is needed to determine the form factors parameters from the $\overline{\mathrm{B}} \to \mathrm{D}\ell\bar{\nu}$ channel in the hypothesis of perfect resolution.

The Monte Carlo data is generated with an accept/reject procedure: a random q^2 between the lower and higher kinematic endpoints is given by the TRandom3() pseudo-casual ROOT number generator and it is accepted if the value of the PDF (96) with known parameters of table (IV) at such random q^2 is lower than some threshold determined from the PDF itself. The dataset is then stored and distributed in histograms composed by different number of samples and in several binning schemes as in table (V).

At this stage of the analysis the samples are still collected in the hypothesis of perfect detector efficiency. An example of a histogram generated through this method is represented in figure (4).

$$\frac{1}{\Gamma(\mathrm{B} \to \mathrm{D}\mu\nu)} \frac{d\Gamma(\mathrm{B} \to \mathrm{D}\mu\nu)}{dq^2} \tag{96}$$

The fit to the parameters of f_+ andf_0 in the BCL parametrization will be performed in several binning schemes and number of events for each sample, being our aim to estimate the relative influence of these fit conditions on the error to the fitted parameters. The different fit conditions are listed in table (V).

TABLE IV. List of the parameters used for the generation of MC samples, as presented in [24]

f_+	a_0	a_1	a_2
	0.836(29)	-2.66(52)	-0.07(2.96)
f_0	b_0	b_1	b_2
	0.647(29)	0.27(30)	-0.09(2.94)

TABLE V. Different fit conditions: for each number of samples, all the bin configurations have been fitted

# of bins	14	20	40	60	80
# of events per sample ($\cdot 10^6$)	0.5	1	2	3	4

C. Configuration of the fit

The script used to fit is TMinuit [25] and was implemented through a python script which has been iterated on several samples. The fit method used in this analysis is the chi-square minimization. In order to check for the fit stability and the uniqueness of the minimun, for each sample two fits have been performed: one in which the fitted

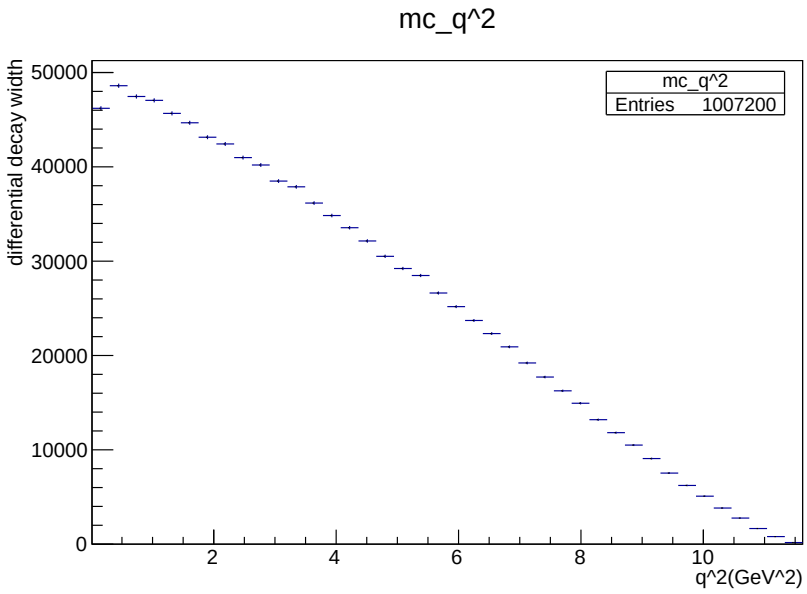


FIG. 4. MC generated sample of 10^6 events for the differential decay width B $\to$ D$\mu\nu$ with form factors parameters listed in [24]

parameters were initialised at the same arbitrary value (0.1) and are set unbounded, while the following in which parameters have been initialised to the results of the previous fit and bound between ± 4 times the absolute value of the parameters from the previous fit results. After running this test, the fit resulted stable and the minimum unique.

The function used to fit is the normalized differential decay rate (96) which, by inserting the form factors parametrization as in (83), and integrating on the kinematic z-variable will yeld a PDF whose shape will depend on the form factors' parameters. Note that the PDF shape is invariant if the all form factors' parameters are rescaled by a constant. As a consequence of that, the fit is only sensitive to ratios of BCL parameters, if no further input on the form factors' overall normalization is provided. Thus the fitting function will be of the form:

$$\frac{1}{\Gamma(1,\tilde{a}_1,\ldots,\tilde{a}_{N_+};\tilde{b}_0,\tilde{b}_1,\ldots,\tilde{b}_{N_0})}\frac{d\Gamma(1,\tilde{a}_1,\ldots,\tilde{a}_{N_+};\tilde{b}_0,\tilde{b}_1,\ldots,\tilde{b}_{N_0})}{dq^2} \tag{97}$$

Where $\tilde{a}_k = a_k/a_0$ and $\tilde{b}_k = b_k/a_0$

It is instructive to note how the shape of the width changes when including either both or only the vector form factor in eq. (96) also by computing the ratio of the PDFs shapes.

$$\frac{\frac{d\Gamma}{dq^2}_{\text{without } f_0}}{\frac{d\Gamma}{dq^2}_{\text{with } f_0}} \tag{98}$$

it is found that the differences are of order percent, as shown in fig (5).

When performing the fit of (97) with the form factors (83) expanded to $N_+ = 3$ and $N_0 = 2$, TMinuit encounters difficulties in fitting the higher order free parameters $\tilde{a}_2$ and $\tilde{b}_1$. Since the fitter is not sensitive to their variation, its attempt to minimise chi-square with respect to their variation pushes them to the boundaries or yields results compatible with zero, this has a deteriorating effect on the determination of the error on the other parameters the fit is sensitive to. So we have decided to expand the form factors to order $N_+ = 2$ and $N_0 = 1$, which fixes the parameters $\tilde{a}_2$ and $\tilde{b}_1$ to be a function of the remaining free parameters $\tilde{a}_1$ and $\tilde{b}_0$ of the fit, by means of relations (81) and (82).

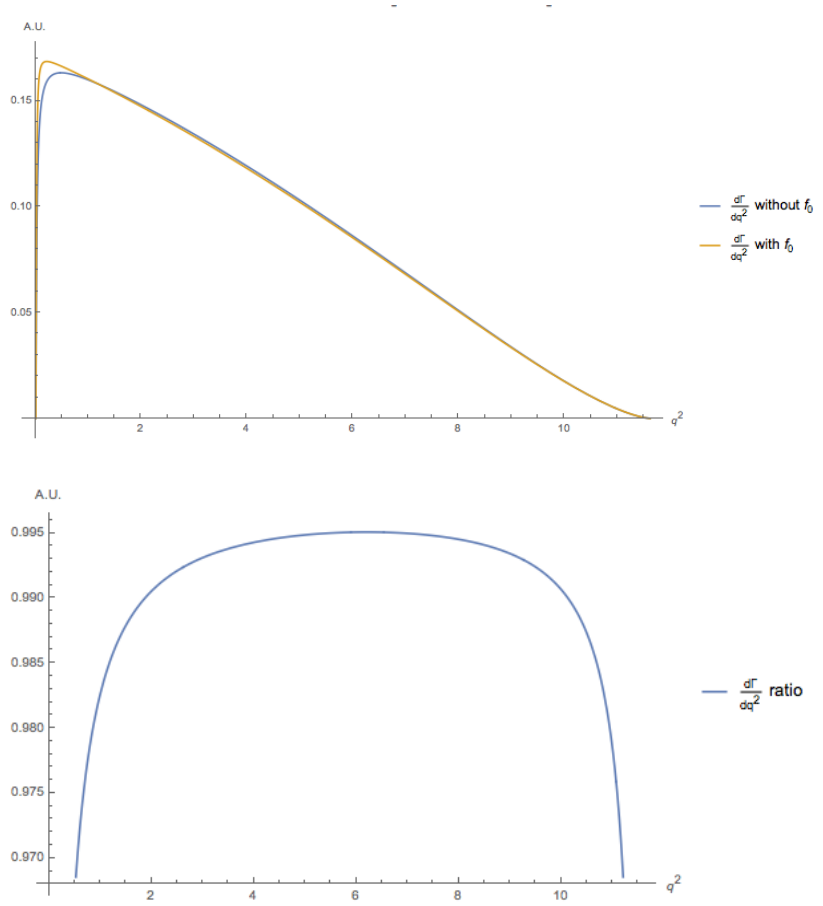


FIG. 5. Comparisons of the PDFs for the B $\to$ D$\mu\nu$ decay width when including either only the vector form factors or both, 5(b) displays their ratio

The error due to this order truncation can be estimated by the use of (95). We calculated that for the two form factors amounts to the value shown in fig. (6).

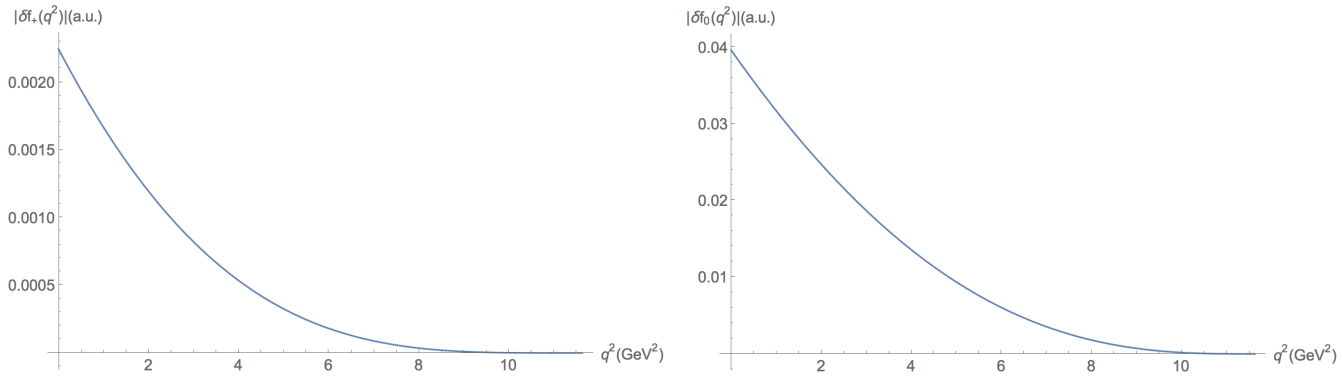
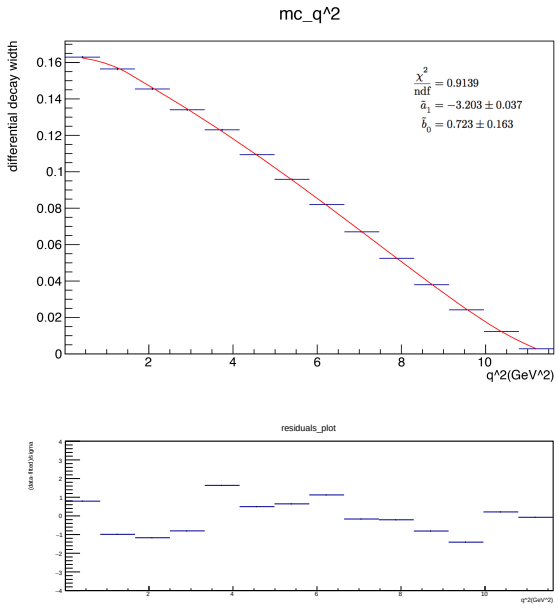


FIG. 6. Error estimate due to series truncation of the form factors as in (83)

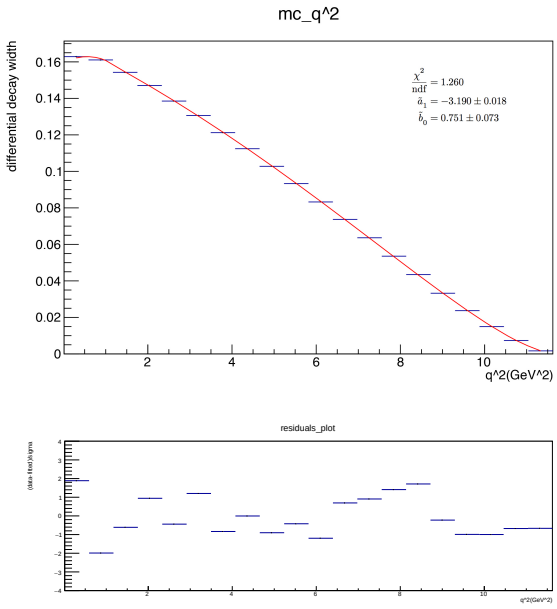
This plots shows that this systematic error affects the determination of both form factors parameters, especially at low q^2, but while on the vector form factor the absolute error due to the series truncation is of order per mille, the one on the scalar form factor is of order percent.

D. Fit results with perfect resolution hypothesis

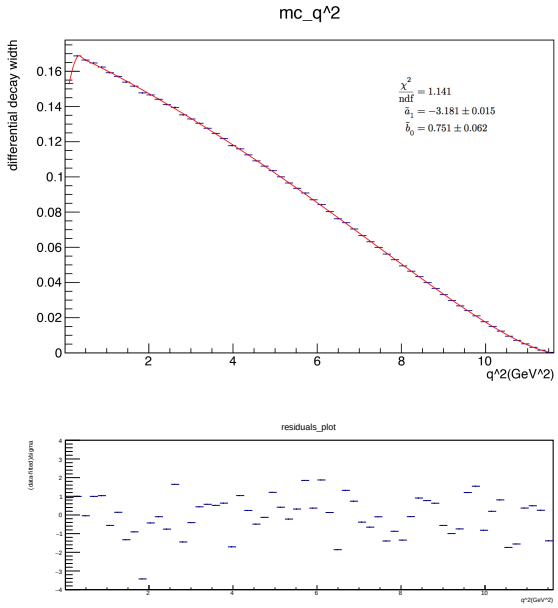
An exemplifying group of fit results is presented in this subsection. The group of figures (7) shows the fit and the corresponding residuals in four different fit conditions.



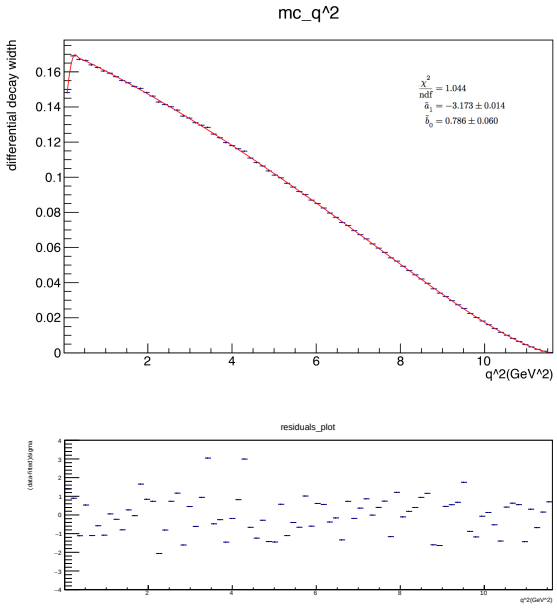
(c) Fit results of a 5×10^5 events sample plotted in 14 bins



(d) Fit results of a 2×10^6 events sample plotted in 20 bins

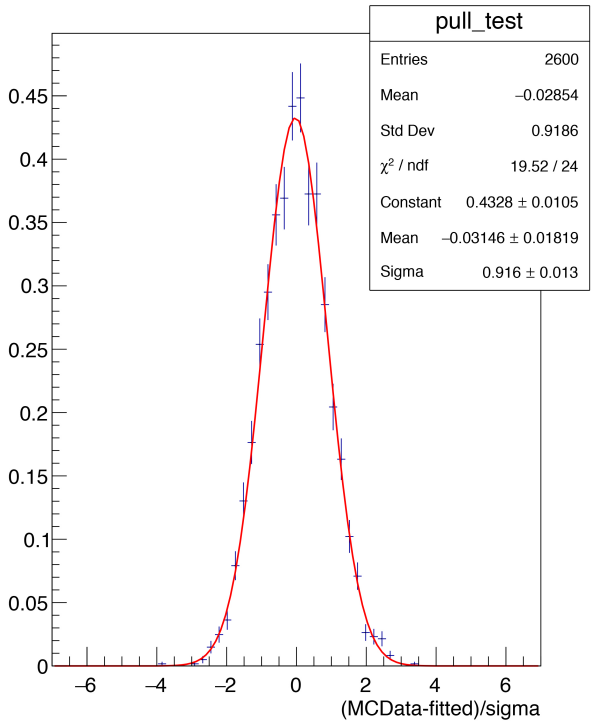


(g) Fit results of a 3×10^6 events sample plotted in 60 bins

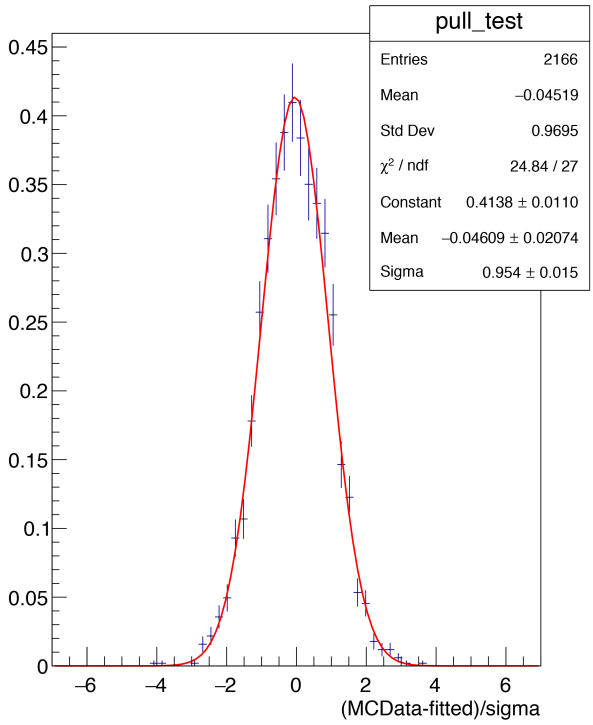


(h) Fit results of a 4×10^6 events sample plotted in 80 bins

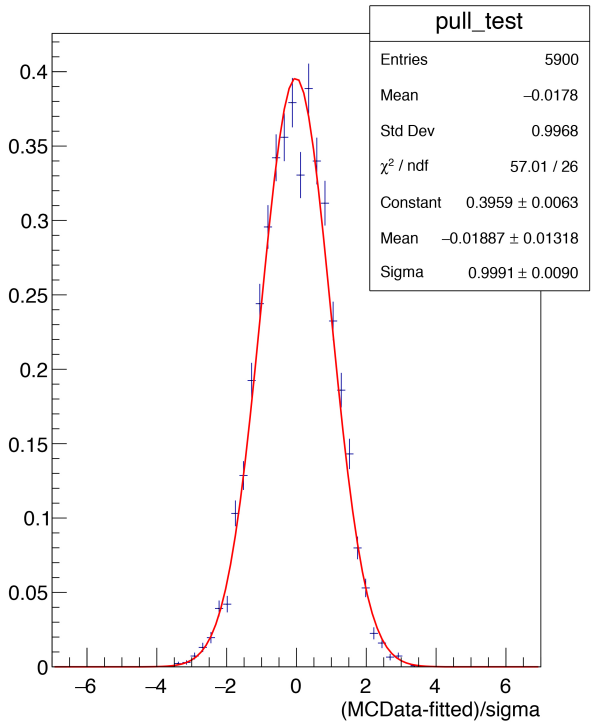
FIG. 7. Fit results in several fit conditions, the samples where generated through MC generation with an accept/reject method.



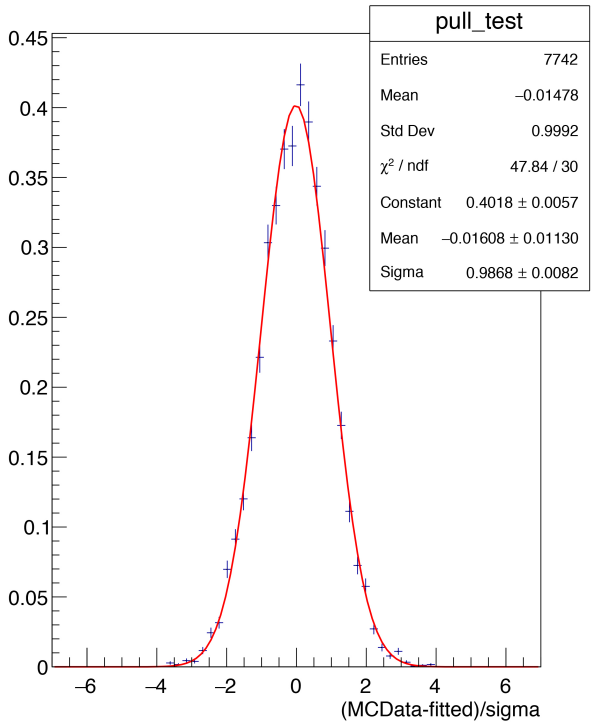
(a) Results of the pull test on 100 samples of 5×10^5 events in 14 bins



(b) Results of the pull test on 100 samples of 2×10^6 events in 20 bins



(c) Results of the pull test on 100 samples of 3×10^6 events in 60 bins



(d) Results of the pull test on 100 samples of 4×10^6 events in 80 bins

FIG. 8. Different pull tests run on some of the fit conditions listed in fig. (7), each of which is listed below the pull plot.

By executing a pull test on 100 fit iterations on statistically independent samples in each binning scheme and number of events, we can check the performance of the fitter script. The pull tests for the fit configurations of the samples shown in fig. (7) are shown in fig. (8).

For each bin, the theoretical function to be compared with the histogram value, was averaged several times in order to compare a better estimate of the function with the MC data in the chi-square minimizing function, simulating a finer binning even when the number of bins in which the histogram is plotted is low. While this choice is not different from averaging the value of the function at the bin edges where the slope of the function with true parameters as input is constant, it provides better sensitivity to the deviations from constant slope where those are important to retrieve the true parameters, i.e. at low and high q^2 as can be seen from fig. (5). We will refer to the number of times per bin the function was averaged with the term "finer" binning.

The inputted value in the events generation for $\tilde{a}_1$ is given by the ratio of a_1/a_0 as given in table (IV) which amounts to $\tilde{a}_1 = -3.182$, while for $\tilde{b}_1$ it is given by the ratio of b_0/a_0 of the same table, which amounts to $\tilde{a}_1 = 0.774$. Lattice QCD results ([24]) estimate of $\tilde{a}_1$ has a relative error of $\sim 20\%$ while it is reduced to $\sim 5\%$ as far as $\tilde{b}_0$ is concerned. Those values will have to be compared with the fit results of tables (VI) and (VII) respectively.

The scaling of the error on the fitted parameters with the number of events is shown in the plots of fig. (10), while varying the bins number but fixing the events number per samples influences the errors on the fitted parameters as shown in fig. (11)

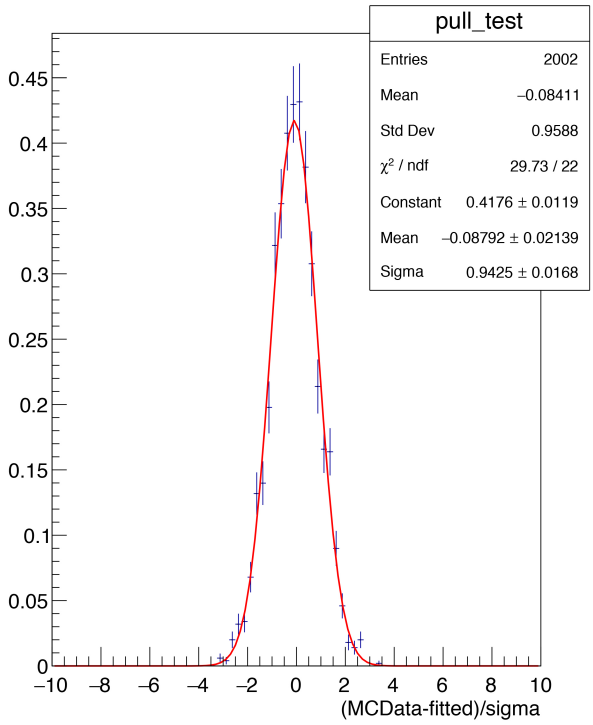
As far as the σ of the gaussian fitted from the pulls of the fit iterations is concerned, it shows an overestimation of the error which amounts up to 10% in low binning schemes. Moreover, we noted that the fit exhibits a small bias on the average value of the pulls which becomes more important when fitting at a low binning scheme. This in turn translates in a slight under/overestimation of the fitted values with respect to the real ones the data was generated with. This effect is found to be (at least) influenced by the finer sampling of the fitting function we adopt. In order to quantify this influence and verify it's dependence with the number of times per bin the fit function was averaged, we executed four pull tests at a low binning scheme (14 bins) with a fixed number of events in the sample (2×10^6) but varying the number of further bin partitioning. The variation of the bias is shown in fig. (9).

The bias tends to decrease as the averaging of the fit function per bin becomes finer. The number of times the function gets averaged in one bin is upper bound by machine precision and cannot be set too large otherwise the fit would loose sensitivity to the parameters, thus an optimal per bin finer partitioning of 300 is chosen as a compromise of computing time and low biased results. A numerical study on the machine precision needed to reach zero bias would be one way to further investigate this result. In any case this difficulty could be overcame for example by the use of a variable partitioning of the fit function evaluation, made finer only where the PDF changes it's first derivative.

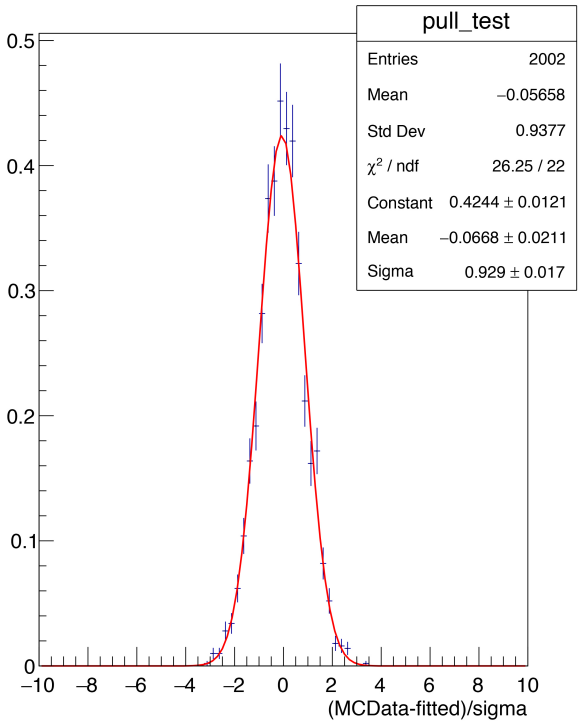
Tables (VI) and (VII) list the average values of $\tilde{a}_1^{fit} - \tilde{a}_1^{MC}$ and $\tilde{b}_0^{fit} - \tilde{b}_0^{MC}$ together with their root means square $\tilde{a}_{1,rms}^{fit}$ and $\tilde{b}_{0,rms}^{fit}$, respectively.

TABLE VI. Average value of $\tilde{a}_1^{fit} - \tilde{a}_1^{MC}$ and root mean square of $\tilde{a}_1^{fit}$ in the different fit conditions

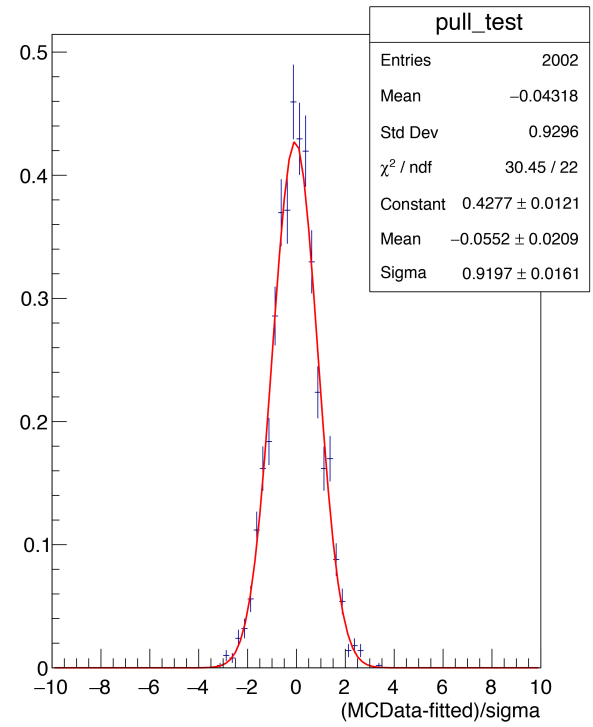
$\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$, $\tilde{a}_{1,rms}^{fit}$	# of ev.ts $\times 10^6$				
	0.5	1	2	3	4
# of bins					
5	0.050, 0.061	0.044, 0.044	0.046, 0.039	0.036, 0.021	0.035, 0.020
14	0.025, 0.040	0.018, 0.030	0.014, 0.020	0.015, 0.016	0.014, 0.015
20	0.020, 0.039	0.013, 0.028	0.010, 0.020	0.010, 0.015	0.009, 0.014
40	0.015, 0.038	0.007, 0.027	0.006, 0.020	0.005, 0.015	0.005, 0.013
60	0.013, 0.037	0.006, 0.027	0.005, 0.020	0.004, 0.015	0.003, 0.014
80	0.013, 0.037	0.006, 0.027	0.004, 0.019	0.003, 0.014	0.003, 0.014



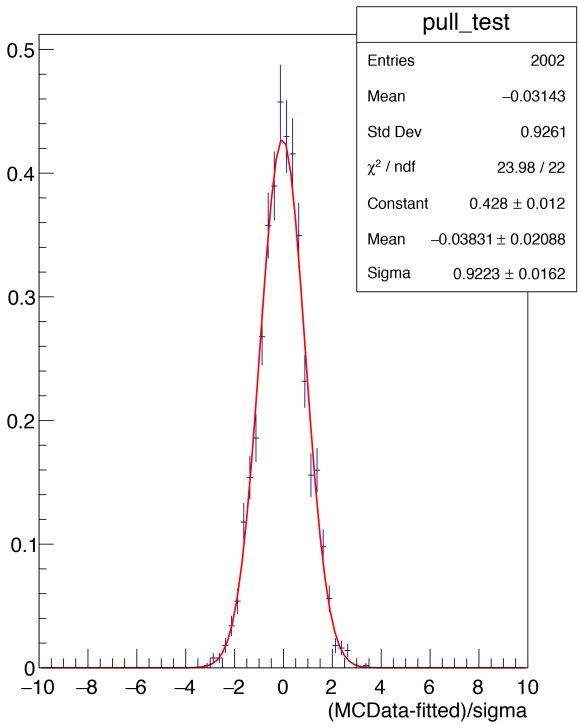
(a) Finer partitioning = 100



(b) Finer partitioning = 200



(c) Finer partitioning = 300



(d) Finer partitioning = 500

FIG. 9. Results of the pull test on the same 100 MC samples composed by 2×10^6 events in 14 bins, each of the pulls is executed with a finer averaging of the fit function, while keeping the bin width constant. The finer partitioning of the fit function of each pull is shown in the caption. The μ of the gaussian is progressively shifting towards zero indicating a decreasing bias.

TABLE VII. Average value of $\tilde{b}_0^{fit} - \tilde{b}_0^{MC}$ and root mean square of $\tilde{b}_0^{fit}$ in the different fit conditions

$\langle \tilde{b}_0^{fit} - \tilde{b}_0^{MC} \rangle$, $\tilde{b}_{0,rms}^{fit}$	# of ev.ts $\times 10^6$				
	0.5	1	2	3	4
# of bins					
5	0.029, 0.215	0.018, 0.183	0.035, 0.169	-0.025, 0.115	-0.015, 0.107
14	-0.026, 0.140	-0.006, 0.109	-0.007, 0.076	-0.003, 0.066	-0.003, 0.055
20	-0.021, 0.124	-0.009, 0.094	-0.006, 0.068	-0.002, 0.059	-0.003, 0.047
40	-0.020, 0.115	-0.017, 0.084	-0.005, 0.064	-0.007, 0.057	-0.005, 0.047
60	-0.033, 0.110	-0.023, 0.086	-0.007, 0.067	-0.007, 0.058	-0.005, 0.047
80	-0.035, 0.108	-0.023, 0.085	-0.010, 0.066	-0.006, 0.059	-0.007, 0.050

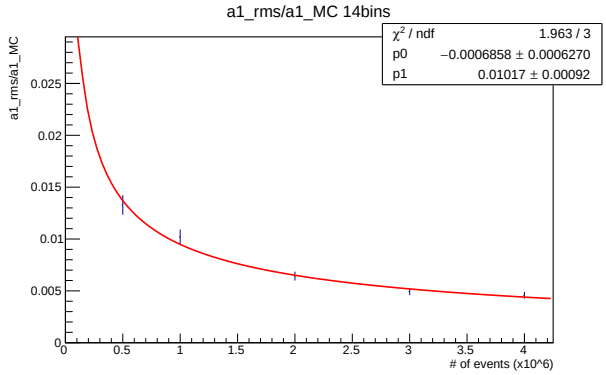
Both errors on parameters scale correctly as $1/\sqrt{N}$ when increasing the number of events in the fit sample, and tend to decrease while increasing the number of bins the sample is distributed into, as seen in fig (10). In those plots the superposed red line is a fit to the function

$$f(x) = p_0 + p_1 x^{1/2}$$

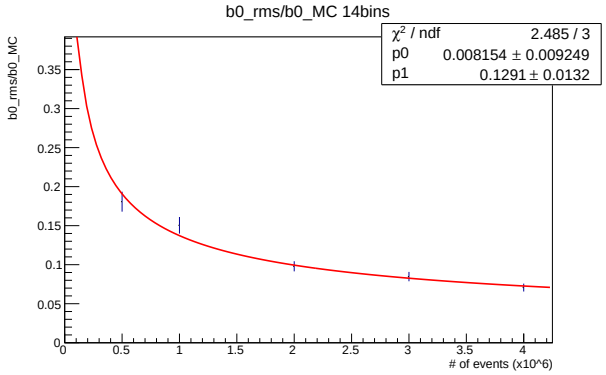
In order to retrieve the value under which the error on the parameter extracted from the fit can not decrease, even at infinite number of events.

The effect of error decrease with the increase of bins number, keeping fixed the number of events, is not as pronounced as expected, this is most likely due to the fact that in this case we kept the finer binning constant throughout the bin variation. This trend can be verified by looking at fig. (11). The extra line of 5 bins is indicative of at which lower binning threshold the rms on the parameters starts to become too large, or the fitter yields a result incompatible with the real one.

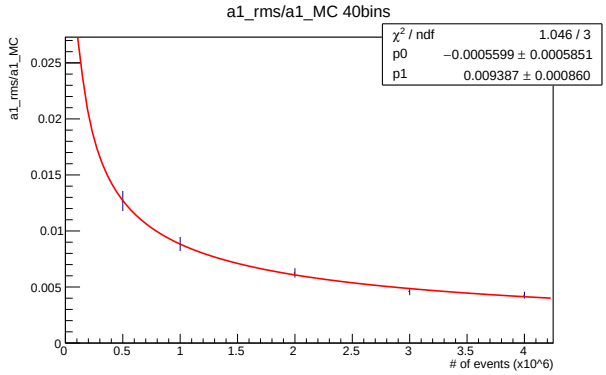
It is interesting to observe how for $\tilde{a}_0$ a lower binning scheme causes $\tilde{a}_{1,rms} \sim \langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$ and thus several times the fitter yields a value for $\tilde{a}_1^{fit}$ at at least 1σ from the real value, while increasing the bins number has a consistent influence on lowering the value of $\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$, especially at a high number of events. On average the extracted values of $\tilde{b}_0$ behave in a somewhat different way: the rms is halved when passing from 5 to 14 bins, especially at a high number of events, but remains always compatible with the value the samples where generated with.



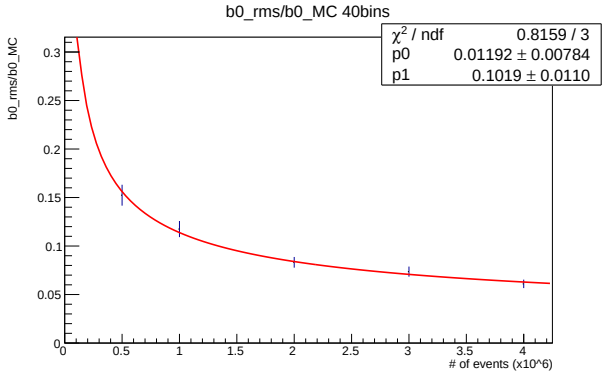
(a) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of events per sample, 14 bins



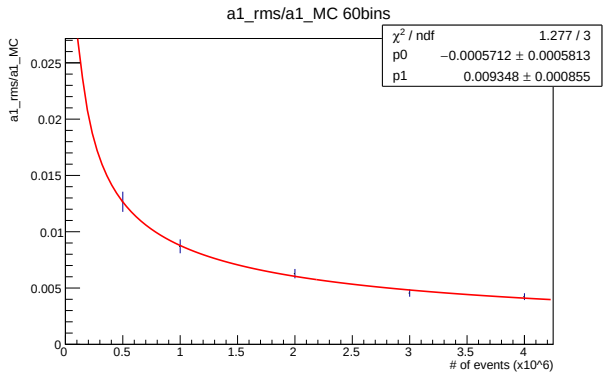
(b) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of events per sample, 14 bins



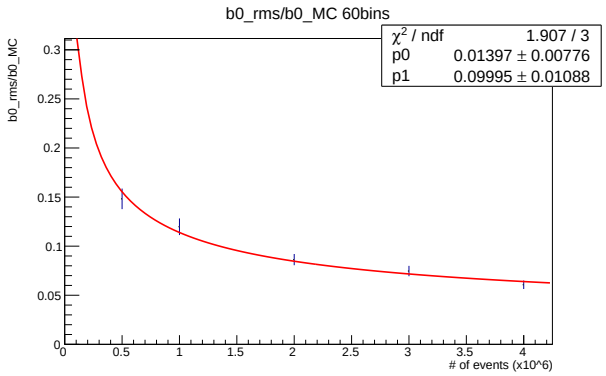
(c) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of events per sample, 40 bins



(d) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of events per sample, 40 bins

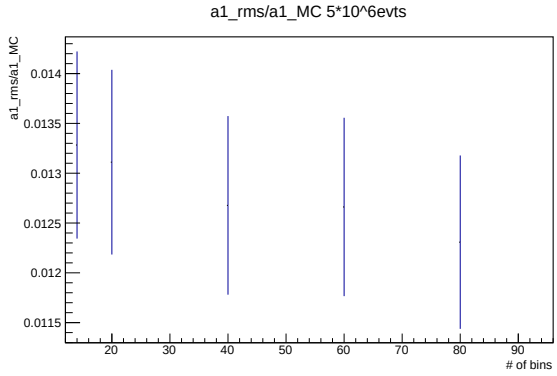


(e) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of events per sample, 60 bins

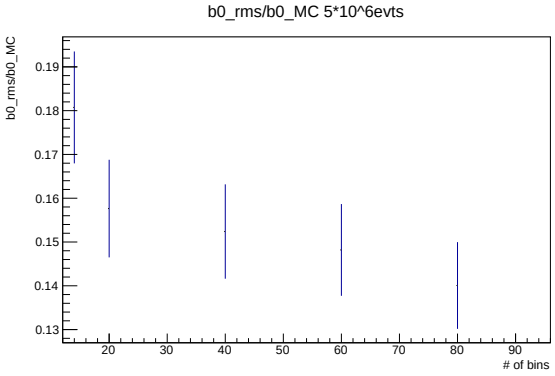


(f) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of events per sample, 60 bins

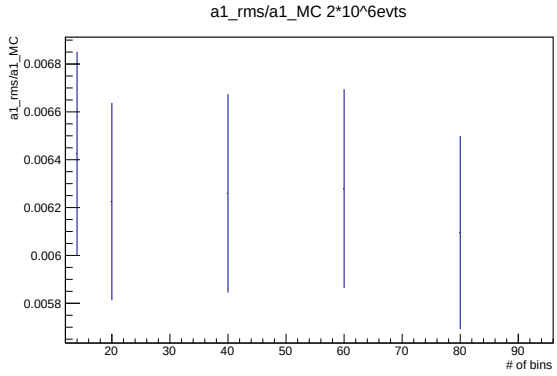
FIG. 10. Scaling of the relative error on the parameters given a fit with fixed bins number but variable number of events per MC sample. The parameter and the bins number are listed in each caption, the superposed fit function is $f(x) = p_0 + p_1 x^{1/2}$.



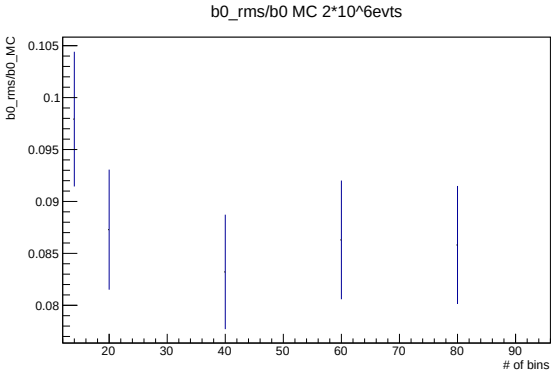
(a) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of bins per sample, 5×10^5 events



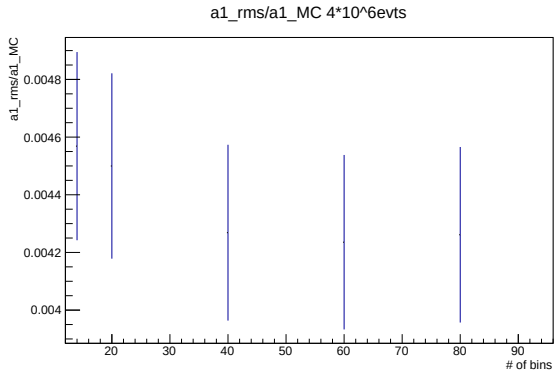
(b) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of bins per sample, 5×10^5 events



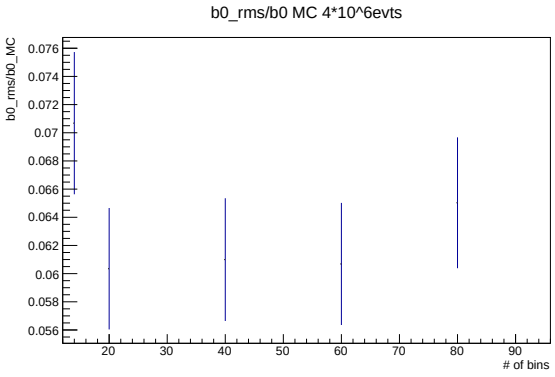
(c) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of bins per sample, 2×10^6 events



(d) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of bins per sample, 2×10^6 events



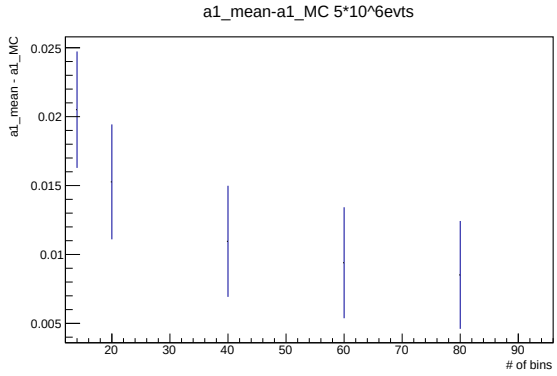
(e) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of bins per sample, 4×10^6 events



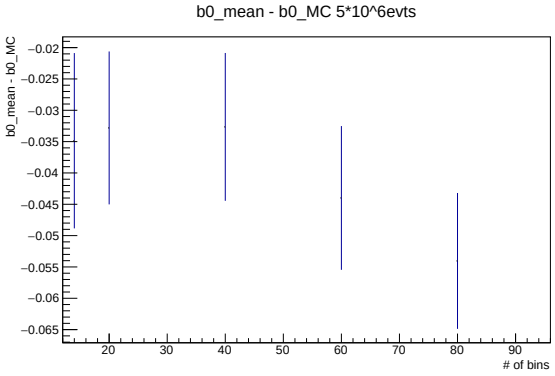
(f) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of bins per sample, 4×10^6 events

FIG. 11. Scaling of the relative error on the parameters given a fit with fixed number of events per MC sample, but variable number of bins. The parameter and the number of events are listed in each caption.

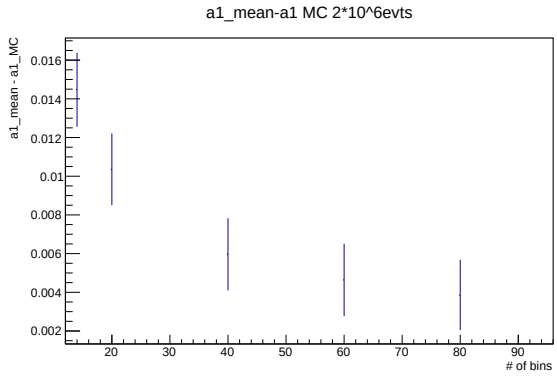
While for $\tilde{a}_1$ the mean over the fit iterations of the difference between fitted parameter value and the input one decreases as the number of bin increases, it is not the case for $\tilde{b}_0$. Nevertheless in most cases the mean value of $\tilde{b}_0^{fit} - \tilde{b}_0^{MC}$ is compatible with zero or at less than two sigmas from it, except for the 5×10^5 and 1×10^6 events samples which exhibit a growing bias when increasing the bins number. This could be the effect of the bias noted before, but in no case the mean value of $\tilde{b}_0^{fit} - \tilde{b}_0^{MC}$ is greater than the error on the fitted parameter. The plots showing the scaling of the mean values of the fitted parameters with the bins number variation are shown in fig. (12)



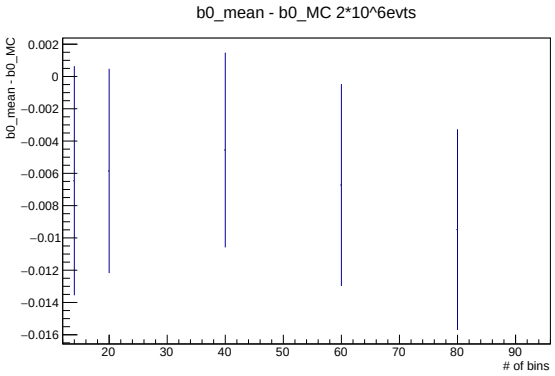
(a) Variation of $\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$ with respect to the number of bins, given a sample of 5×10^5 events



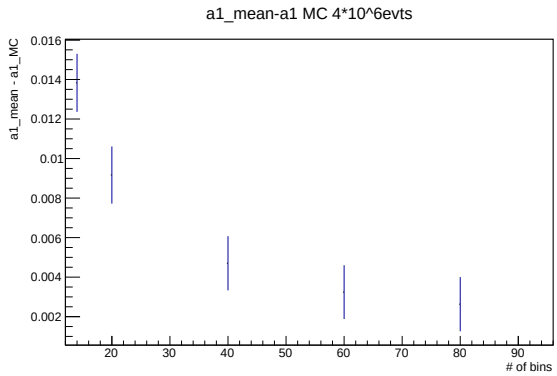
(b) Variation of $\langle \tilde{b}_0^{fit} - \tilde{b}_0^{MC} \rangle$ with respect to the number of bins, given a sample of 5×10^5 events



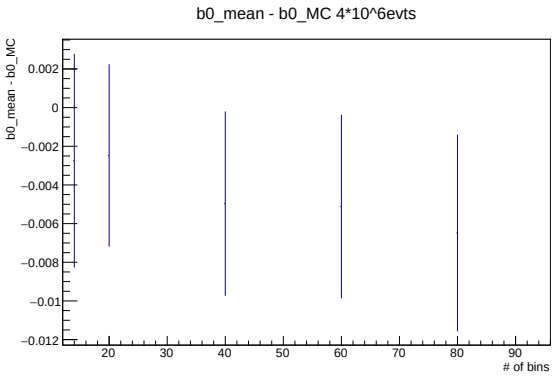
(c) Variation of $\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$ with respect to the number of bins, given a sample of 2×10^6 events



(d) Variation of $\langle \tilde{b}_0^{fit} - \tilde{b}_0^{MC} \rangle$ with respect to the number of bins, given a sample of 2×10^6 events



(e) Variation of $\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$ with respect to the number of bins, given a sample of 4×10^6 events



(f) Variation of $\langle \tilde{b}_0^{fit} - \tilde{b}_0^{MC} \rangle$ with respect to the number of bins, given a sample of 4×10^6 events

FIG. 12. Scaling of $\tilde{b}_0^{fit} - \tilde{b}_0^{MC}$ averaged over 100 fit interations, with respect to the variation of the number of bins.

The study of the fit in different conditions allows us to conclude that in perfect resolution and ideal efficiency conditions, both form factors can be resolved from the B $\to$ D$\mu\nu$ decay, even though with different sensitivity. They can both be accessed in a model independent way through a fit to an arbitrarily normalized PDF which allows to determine ratios of parameters that describe the form factors in the BCL parametrization.

As far as $f_+(q^2)$ is concerned, the ratio $\tilde{a}_1 = a_1/a_0$ could be measured, given a sample of 2×10^6 events distributed in 14 bins, with a sensitivity of around 0.56 %. The sensitivity to the ratio $\tilde{b}_0 = b_0/a_0$ would amount to 8.82 % given

the same fit conditions. Both sensitivities have been corrected by subtracting a 10% to their values given the error overestimation of the pulls as shown in fig. (8).

From those values, and comparing with the Lattice QCD results [24], we can conclude that, in perfect resolution condition the value of the ration $\tilde{a}_1$ can be retrieved with 20 times better sensitivity even with as low as 5×10^5 events while it is not the case for the scalar form factor parameters ratio which, in 14 bins can, with the same number of events, be retrieved with roughly 3 times worse sensitivity than the Lattice QCD simulation, when including the error from the series truncation as expressed in (95).

VI. DEALING WITH THE RESOLUTION

A. The unfolding method: introduction

LHCb [26] is one of the four main experiments at the Large Hadron Collider (LHC); its main goal is the study of CP-violation and rare decays of beauty and charm hadrons. In proton-proton collisions $b\bar{b}$ and $c\bar{c}$ pairs are produced mainly through gluon fusion and quark-antiquark annihilation with the two incident partons having very different momenta in the laboratory frame, for this reason the $b\bar{b}$ and the $c\bar{c}$ pairs are preferentially produced around the beam axis in the same forward or backward direction, as shown in fig. (13). The LHCb detector is thus built as a single arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, where $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$. A schematic of the LHCb detector can be seen in fig. (14)

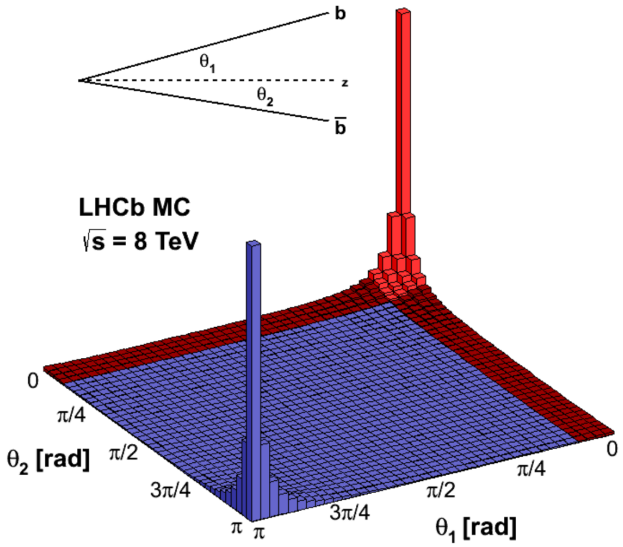


FIG. 13. Simulated $b\bar{b}$ production angles in pp collisions at $\sqrt{s} = 7$ TeV. The region highlighted in red is the LHCb acceptance and θ_i are the angles from the beam directions [26]

In this chapter we will deal with the finite detector resolution in order to give an estimate of the sensitivity to the form factors' parameters in more realistic data simulation. Due to the presence of a missing neutrino, the measurement of the visible products of the decay energies and momenta relate to q^2 through a second order equation, giving a twofold ambiguity on which is the correct solution to accept. This, and the finite detector resolution, combine in a net effect of event migration between bins, which causes the q^2 of an event to be attributed to a different value from the true one. This effect is also referred to as data "smearing" in literature.
Moreover we will take into account the limited acceptance of the detector, which consists of the probability to observe an event. The acceptance is typically lower than 1 and depends on the kinematic variable q^2.

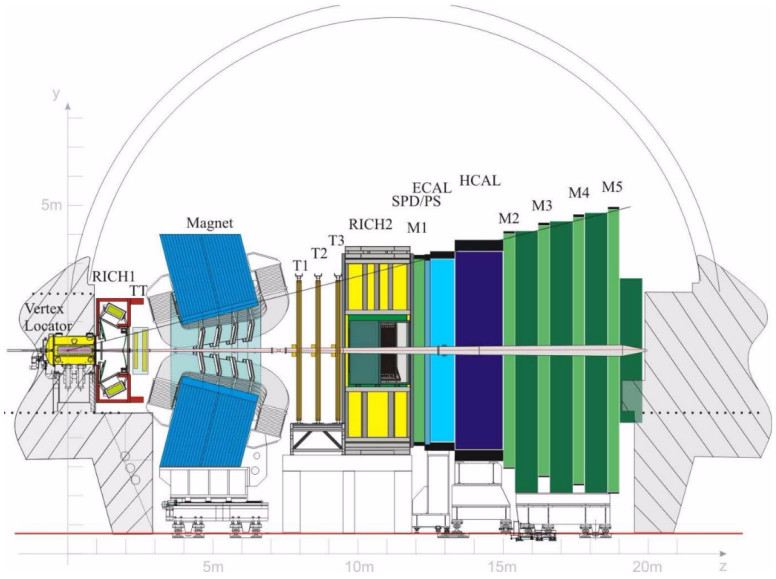


FIG. 14. Schematic of the LHCb detector [26].

In order to correct for those effects, the unfolding procedure is developed. Mathematically the relation between the distribution $f(x)$ of the true variable x, to be determined in an experiment, and the measured distribution $g(y)$ of the measured quantity y is given by the integral equation [27]:

$$g(y) = \int dx\ A(y,x) f(x) \tag{99}$$

Which is a Fredholm integral equation of the first kind. The resolution function $A(y,x)$ represents the effect of the detector. The problem to determine the distribution $f(x)$ from the measured distributions $g(y)$ is called unfolding [27] and it consists of an inverse problem. It requires the knowledge of the resolution function $A(y,x)$, i.e. the effects of limited acceptance and finite resolution. If we represent the distributions by histograms, the resolution function becomes a matrix and equation (99) can then be represented by the matrix equation

$$\mathbf{y}_{\text{meas}} = A\ \mathbf{x}_{\text{true}} \tag{100}$$

which has to be solved for the vector $\mathbf{x}_{\text{true}}$, given the measured vector $\mathbf{y}_{\text{meas}}$. The vector $\mathbf{y}_{\text{meas}}$ with n entries represents a histogram of the measured quantity y_{meas}, and the distribution $f(x)$ is represented by a histogram of the vector $\mathbf{x}_{\text{true}}$ with m entries. The transition from $\mathbf{x}_{\text{true}}$ to $\mathbf{y}_{\text{meas}}$ is described by the $n-$ by $-m$ matrix A whose elements a_{ij} are the probabilities of observing an entry in bin i of histogram $\mathbf{y}_{\text{meas}}$, while its true value was generated in bin j of histogram $\mathbf{x}_{\text{true}}$. Note that an additional row of the matrix has to be considered, namely a_{0j}. It consists of the number of events that were generated in bin j but failed detection, divided by the total number of events generated in bin j. Thus the row $a_{0,j}$ encodes information on the reconstruction efficiency of the detector.

Since the resolution function $A(x,y)$, or the matrix A is not known analytically, it is recovered from the statistical analysis of a Monte Carlo simulation of the response of the detector $g_{\text{MC}}(y)$, based on some assumed input distribution $f_{\text{MC}}(x)$

$$g_{\text{MC}}(y) = \int dx\ A(y,x) f_{\text{MC}}(x) \qquad \text{(Monte Carlo Simulation)} \tag{101}$$

and thus is also statistically limited. Standard methods for the solution of integral equations or linear equations can not be used in this case [27].

B. Simulation of the resolution: building the unfolding matrix

In this subsection, the process of building the unfolding matrix A through MC simulation is presented. The measurement process can be schematized in 3 phases:

- An event with $q^2_{\rm MC}$ is generated with perfect resolution.
- Some of the MC generated events are rejected according to LHCb acceptance criteria.
- If the event passes the selection, it is reconstructed with two values of q^2. The measured value will coincide with the correct solution in $q^2_{\rm measured} = q^2_{\rm correct}$ in the hypothesis of non ambiguity between the two solutions, otherwise $q^2_{\rm measured} = q^2_{\rm rec}$ and one of the two reconstructed solutions will be chosen according to some arbitrary selection criteria.

The tuples used for extracting the resolution information are two and independently generated. They are generated by EvtGen and the LHCb simulation software. The first contains the MC simulated values of q^2 with perfect resolution over the whole kinematic range and solid angle, while the second contains the values of q^2 with perfect resolution that pass LHCb acceptance selections and the corresponding pairs of values of reconstructed q^2 smeared by finite detector resolution. Although the data used to extrapolate resolution information is generated with differently parametrized form factors, this won't affect our resolution study since we're only considering ratios of widths and reconstruction related probabilities.

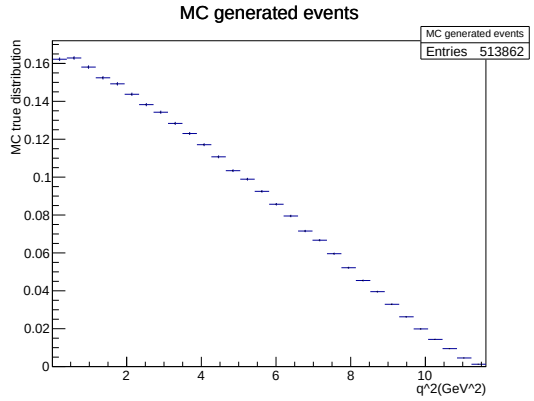
Four distributions of $d\Gamma/dq^2$ can be plotted, and can be seen in fig. (15). The decision criteria used to choose between the two q^2 solutions in fig. (15(d)) is related to the decay time of the D meson. It is found in fact that 62% of the times, the lower of the two reconstructed q^2 value is the correct solution when the decay-time of the D meson in the laboratory frame is greater than some threshold. A longer decay time in the laboratory frame corresponds to a greater Lorentz boost given to the D meson and consequently to less q^2 available to the lepton pair in the decay.

Ratios of the differential decay widths of fig. (15) give an estimate of the per bin efficiencies, and are plotted in fig. (16). Fig. (16(a)) represents the efficiency due to LHCb acceptance cuts while fig. (16(b)) takes into account the reconstruction efficiency of an accepted event. Since not for all events a real solution of the second order equation in q^2 exists due to detector resolution, some of the accepted events are discarded. Fig. (16(c)) is the overall efficiency and represents the product of the two efficiencies described above. Given an overall efficiency of 56 % a sample with N entries will tipycally contain half of them.

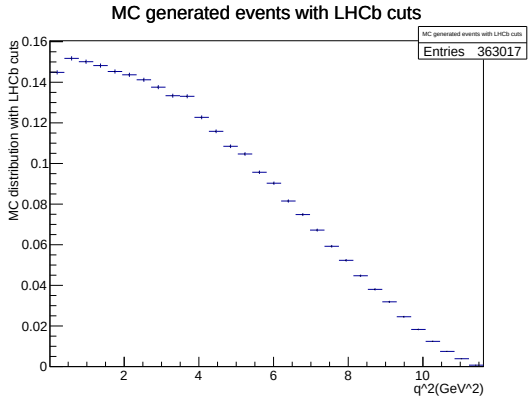
The resolution information is then extracted from the second tuple where the detector response for an event passing the LHCb selection cuts is simulated.The number of rows and columns of the unfolding matrix are resolution dependent. In our case we chose a 30×14 matrix. If the correct q^2 solution was known, the typical σ is of order 0.8 GeV^2 (VI B), and when lowering the number of bins from 30 to 14 the value of σ/binwidth decreases from ~ 2.2 to 1. The same matrix configuration is used for the case in which the ambiguity holds. In this case the typical σ is roughly doubled with respect to the previous case and thus we expect lower unfolding performances.

In order to build the unfolding matrix each $q^2_{\rm MC}$ and its corresponding reconstructed $q^2_{\rm reco}$ are stored in a 2-D histogram; its i,j-th bin content is divided by the total number of $q^2_{\rm MC}$ of the j-th bin and the result is assigned to the $a_{i,j}$-th element of A. In addition to this process, 30 histograms are built by storing the value $\Delta q^2 = q^2_{\rm reco} - q^2_{\rm MC}$ for each bin to which $q^2_{\rm MC}$ belongs. The plots resulting from this first analysis are shown in fig. (17), while in fig. (VI B) the histograms of Δq^2 for 4 sample bins are displayed. The root mean square of this difference is collected and plotted in the histogram of fig. (18).

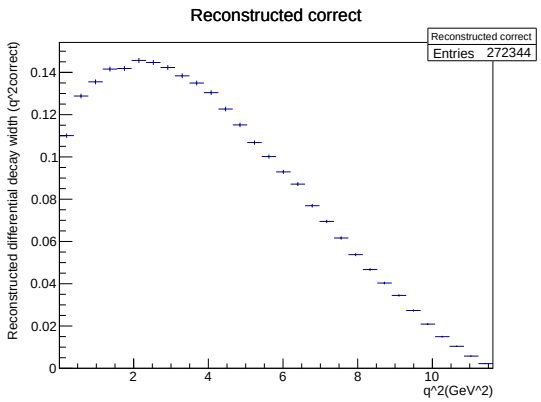
As can be noted from fig. (VI B) the Δq^2 mean value differs from zero, shifting gradually from an overestimation of the reconstructed value of q^2 with respect to it's true value, to its underestimation. Moreover the shape of such



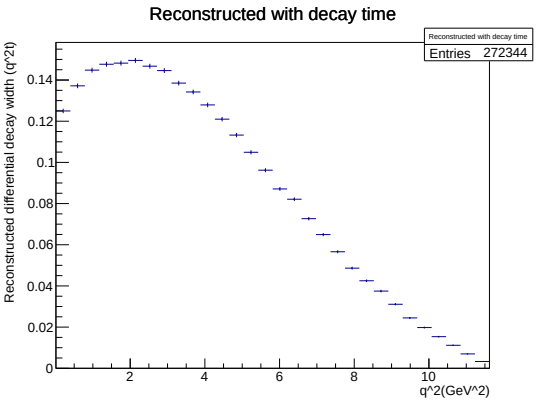
(a) Monte Carlo truth for the decay width, no cuts applied



(b) Monte Carlo truth for the decay width, LHCb cuts applied



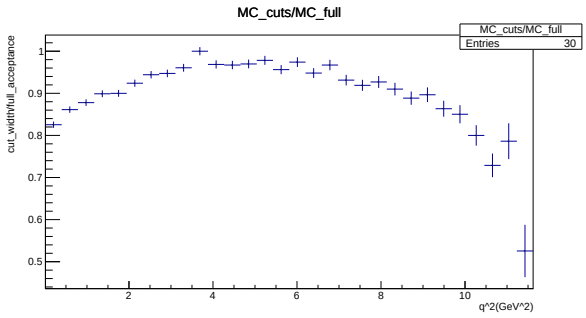
(c) Simulation of q^2 reconstruction, in the hypothesis of no solution ambiguity



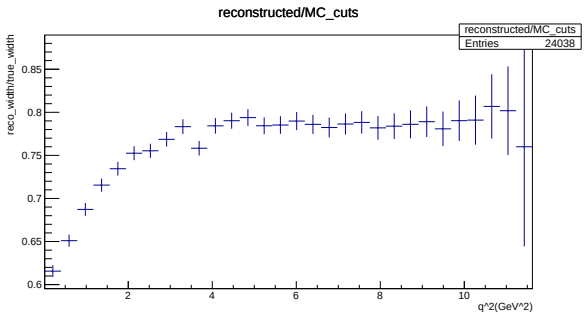
(d) Simulation of q^2 reconstruction, in the hypothesis of solution ambiguity. The selection criteria is based on the τ_D decay time

FIG. 15. Four different stages of the shape of the decay width, from generation to measurement, plotted in 30 bins

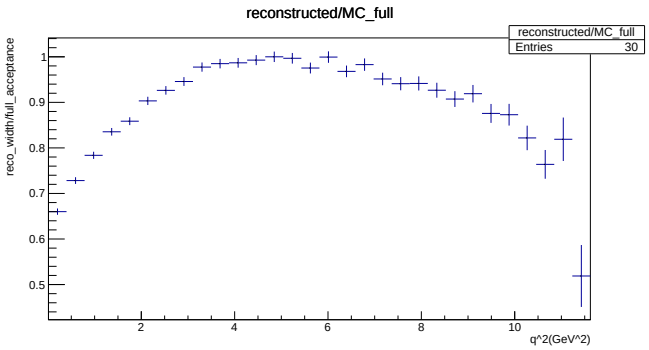
histograms is asymmetric, especially for the q^2 with the solution ambiguity at high q^2, which suggests a deteriorated resolution close to the kinematic endpoint. This can be verified by looking at fig. (18): while for the no ambiguity in q^2 case (blue line) the resolution is roughly independent on q^2 and stable on average at 0.8 GeV^2, it grows almost linearly in q^2 from 1.5 to 2.2 GeV^2 for the case of ambiguity between the solutions (red line). Due to poor statistics in the last bins, the last 5 histograms of Δq^2 where summed and their rms has been set to the same value.



(a) Ratio between differential widths of fig. (15(b)) and (15(a)) plotted as a function of q^2

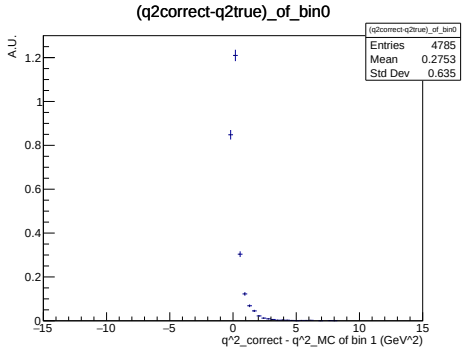


(b) Ratio between differential widths of fig. (15(c)) and (15(b)) plotted as a function of q^2_{true}

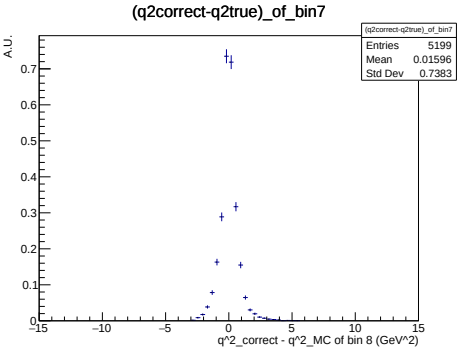


(c) Ratio between differential widths of fig. (15(c)) and (15(a)) plotted as a function of q^2

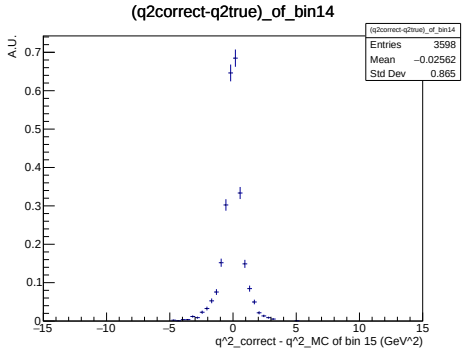
FIG. 16. Ratios of differential widths after selection and after reconstruction are displayed, in fig. (16(c)) and (16(b)) existence of a real q^2 solution is required.



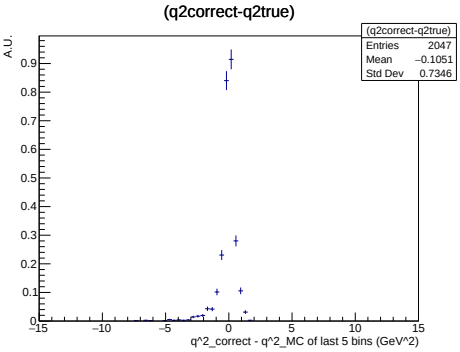
(a) Δq^2 histogram of bin 1 in case of non ambiguity of the correct q^2 solution



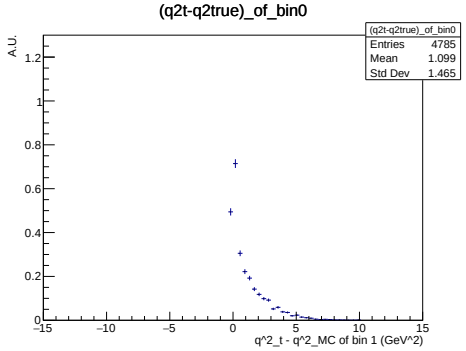
(b) Δq^2 histogram of bin 8 in case of non ambiguity of the correct q^2 solution



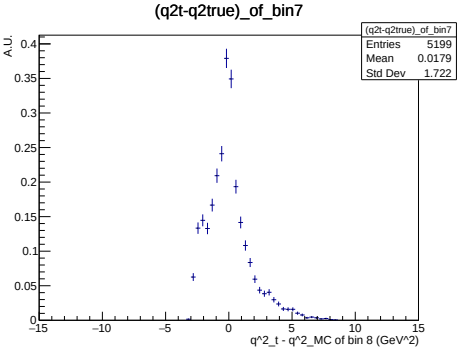
(c) Δq^2 histogram of bin 15 in case of non ambiguity of the correct q^2 solution



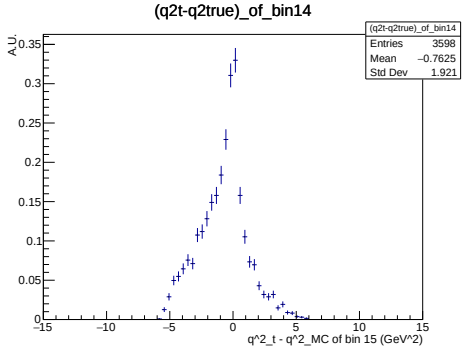
(d) Δq^2 histogram of the last 5 bins in case of non ambiguity of the correct q^2 solution



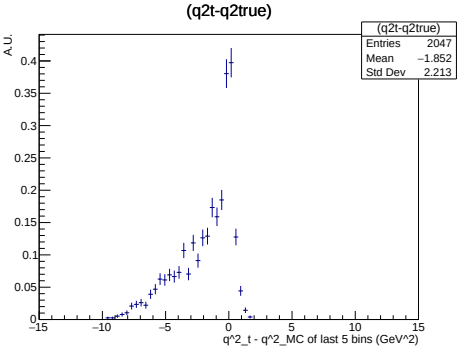
(e) Δq^2 histogram of bin 1 in case of ambiguity of the correct q^2 solution



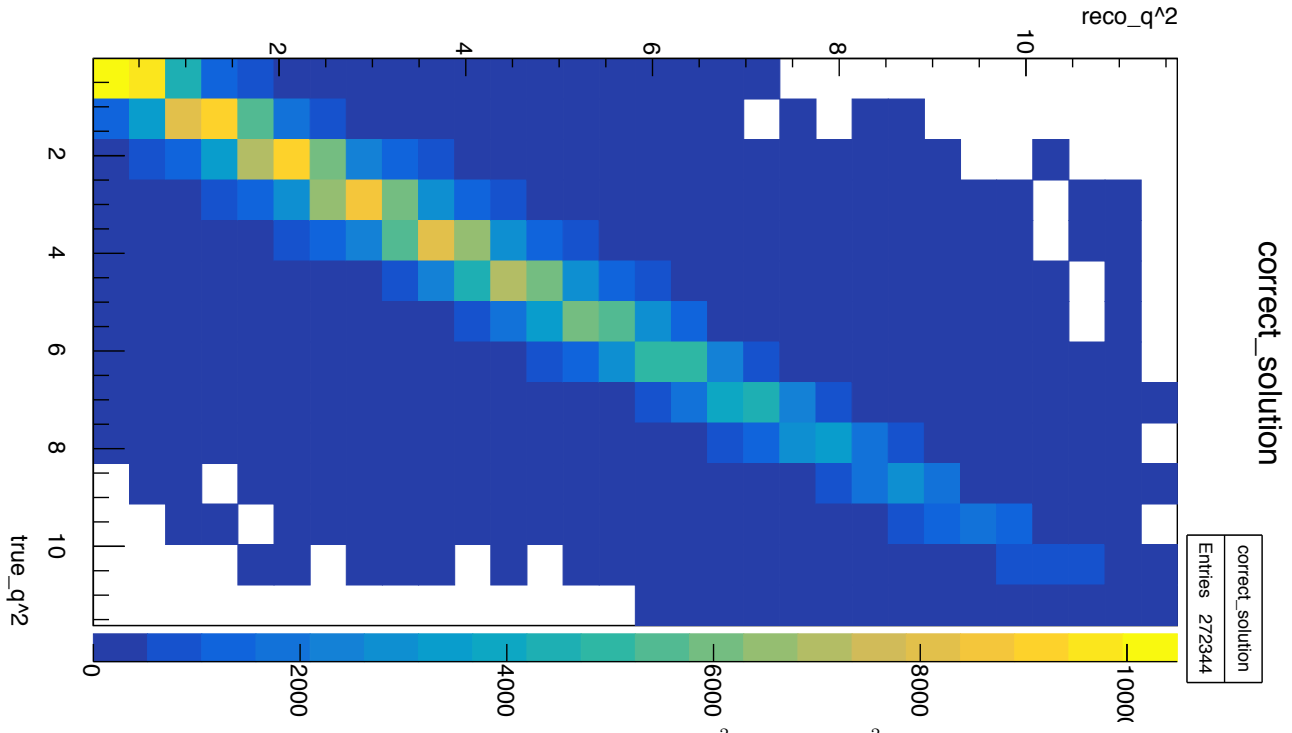
(f) Δq^2 histogram of bin 8 in case of ambiguity of the correct q^2 solution



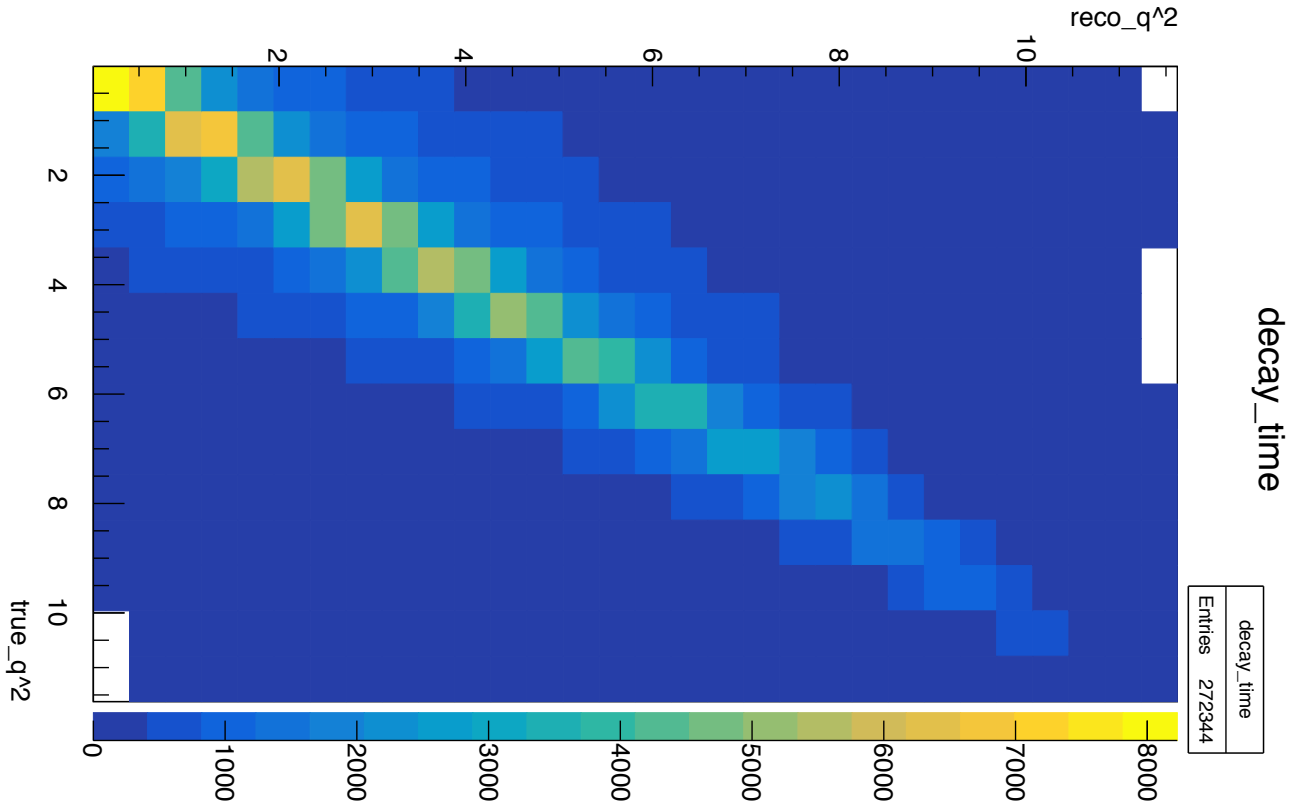
(g) Δq^2 histogram of bin 15 in case of ambiguity of the correct q^2 solution



(h) Δq^2 histogram of last 5 bins in case of ambiguity of the correct q^2 solution



(i) LHCb simulated scatter plot of $q^2_{correct}$ versus q^2_{MC}



(j) LHCb simulated scatter plot of q^2_t versus q^2_{MC}

FIG. 17. Scatter plots for the tuples from LHCb simulation software, plot in fig. (17(i)) represents the case in which the correct solution is known while the one in fig. (17(j)) displays the case in which the decay time of the D meson is used as a choice criterium

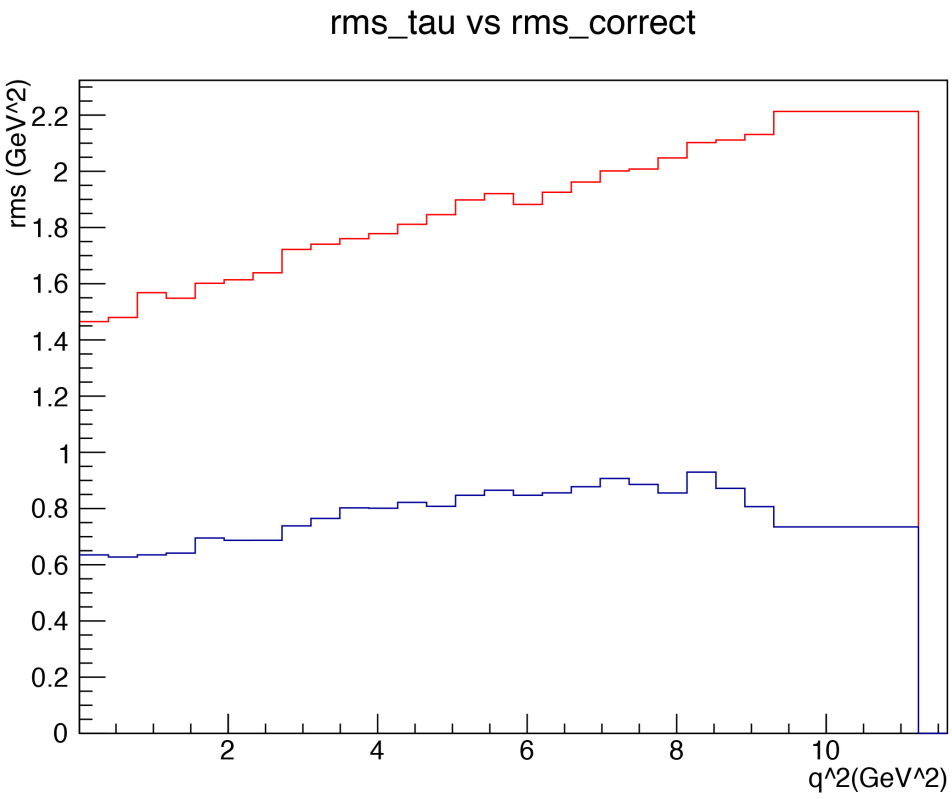


FIG. 18. Dependence on q^2 of the root mean squares of the histograms of fig. (VI B) which model detector resolution. The red solid curve represents the case of q^2 solution chosen with the τ_D criterium, while the blue curve represents the case of non ambiguity between the two solutions

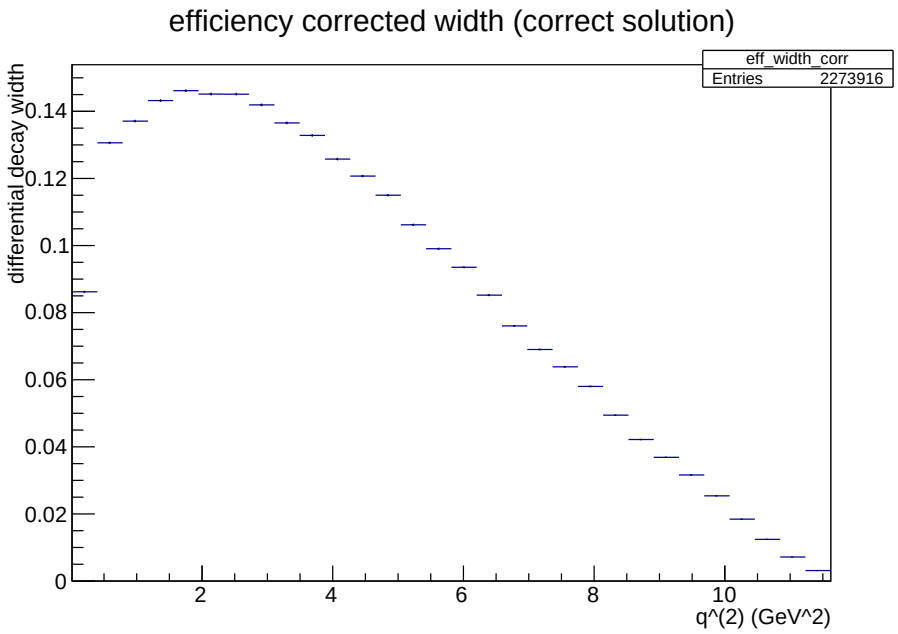
Now that the detector response has been modelled, we can produce several smeared histograms that simulate the reconstructed distributions both in case of known or guessed q^2 correct solution, and an unfolding matrix with arbitrary statistics. The followed procedure to build the unfolding matrix and the smeared histograms is analogous to the one described above:

- A q^2 value is taken from the dataset of the previous infinite resolution analysis.

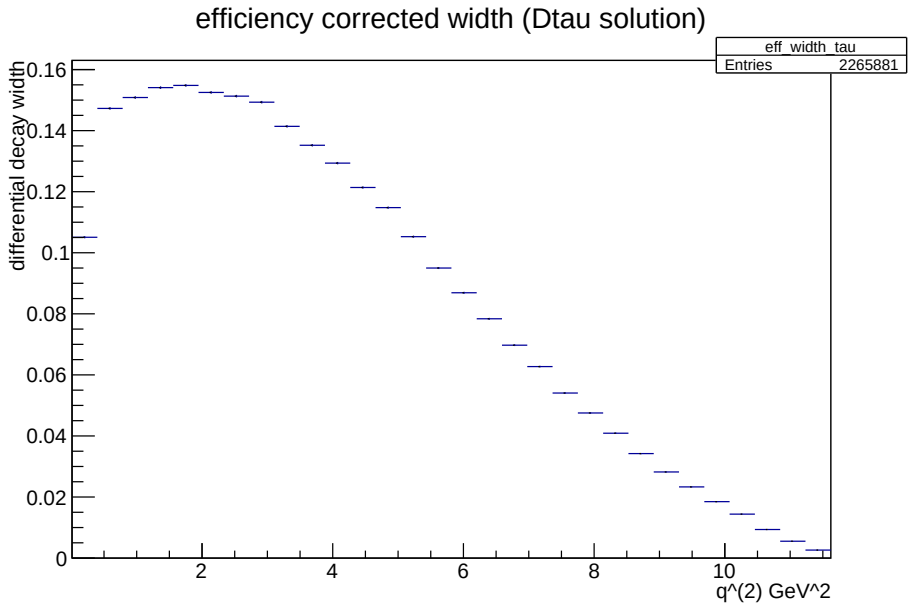
- A random generated number between 0 and 1 is compared with the binned acceptance efficiency of (16(a)), and the q^2 value is accepted if such number is lower than the efficiency value of the corresponding bin.

- A random number is generated between 0 and 1 and determines the accept/reject criterium when compared to the binned reconstruction efficiency of (16(b)). If this number is lower than the efficiency of the corresponding bin, the q^2 value is smeared by adding to its value a random value taken from the corresponding bin resolution PDF of (VI B). If the smeared value is still inside the kinematic boundaries, it is stored. In case we are building the unfolding matrix the original MC q^2 is also stored.

The described process is iterated to obtain several samples like the ones displayed in fig. (19) and the matrices of (20). The produced samples are data-like, they will be used to simulate the measurement process with arbitrary statistics and as an input for the unfolding procedure. The matrices calculated through the process described above now have enough statistics to unfold samples containing a high number of events.
Moreover, the information on the elements $a_{0,j}$ is stored separately, since it consists of the probability of an event to be accepted by LHCb cuts but to be reconstructed with invalid or imaginary q^2 value, it will be used in the unfolding procedure.

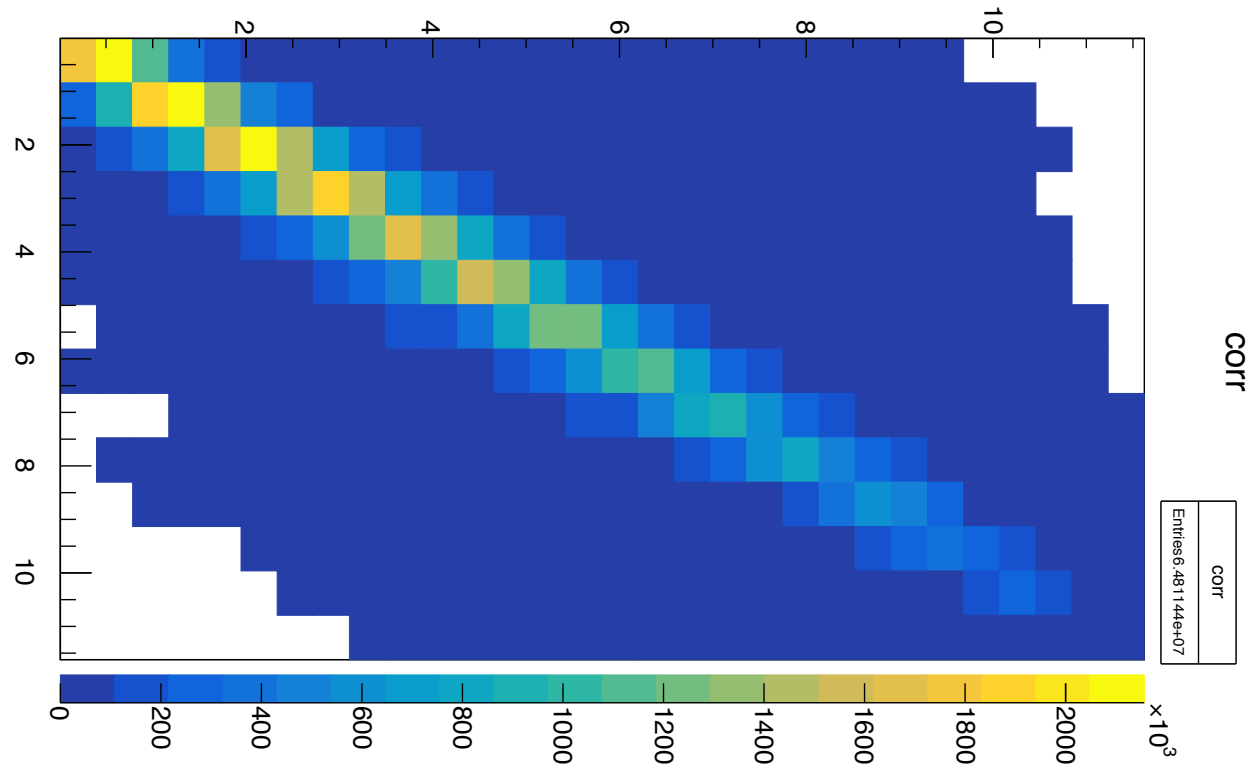


(a) Simulation of the measured differential decay width, in case of correct solution known

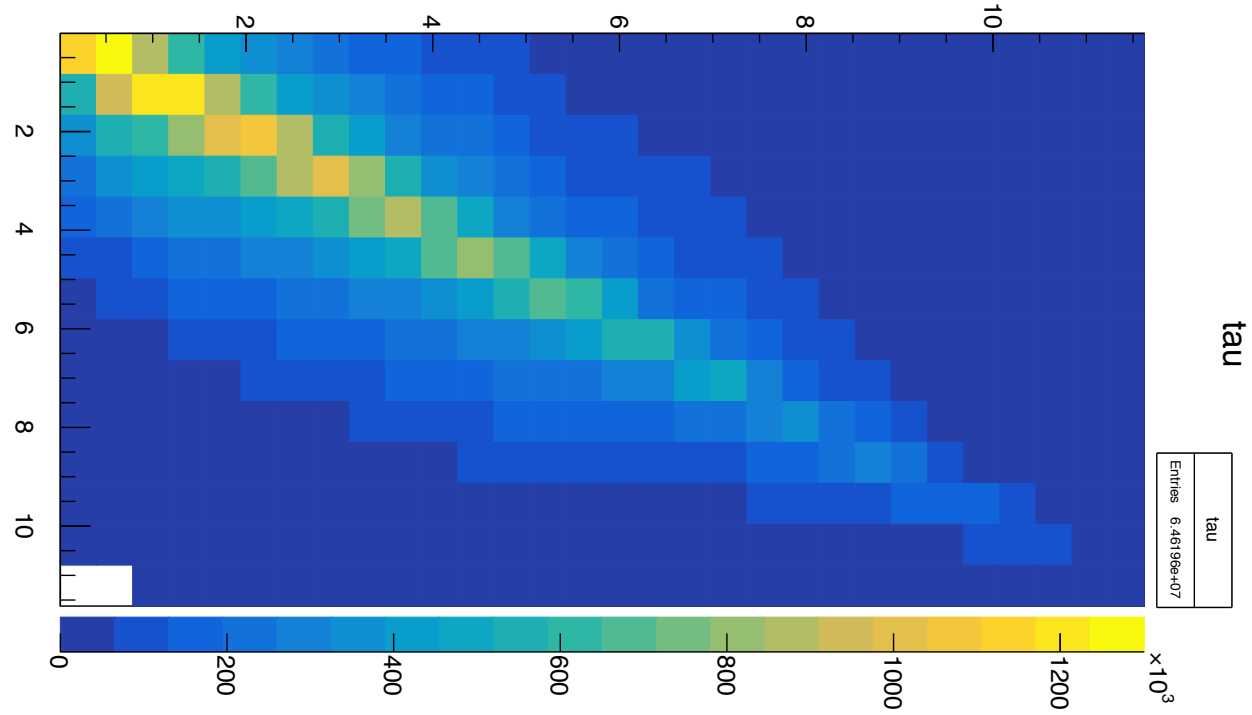


(b) Simulation of the measured differential decay width, in case of ambiguity between the correct solution

FIG. 19. Examples of histograms of datalike collected q^2 samples. The effect of bin migration changes the shape of the PDF deteriorating the estimate of the form factors parameters in a fit



(a) This work's simulated scatter plot of $q^2_{correct}$ versus q^2_{MC}



(b) This work's simulated scatter plot of q^2_t versus q^2_{MC}

FIG. 20. Scatter plots generated in this work, plot in fig. (20(a)) represents the case in which the correct solution is known while the one in fig. (20(b)) displays the case in which the decay time of the D meson is used as a choice criterium

C. Unfolding as an ill-posed problem

The problems inherent to unfolding can be discussed in a simple special case, assuming a resolution matrix A with some smearing of data into nearest neighour bins as in [28].

$$A = \frac{1}{2}\begin{pmatrix} 1+\varepsilon & 1-\varepsilon \\ 1-\varepsilon & 1+\varepsilon \end{pmatrix} \tag{102}$$

with $0 \leq \varepsilon \leq 1$ determining the "quality" of the detector: $\varepsilon = 1$ means an ideal detector with response matrix equal to the identity matrix, while small ε corresponds to a poor detector, almost unable to distinguish the two bins. The measurement process is now simulated by multiplying matrix (102) over the true distribution $\mathbf{x}_{\text{true}}$, resulting in the measured histogram $\mathbf{y}_{\text{meas}}$ as in (100).

In this example efficiency is ideal, so that no event escapes the detector (the sum of elements in each column equals 1), thus the A matrix should represent a rotation matrix which when applied to the true distribution won't change its euclidean norm.

However it is interesting to note that $\det A = \varepsilon$ and the deviation of ε from one represents the deviation of the unfolding matrix A from orthogonality. This can be seen in the dependence of the norm of the vector $\mathbf{y}_{meas}$ on ε which can be written as:

$$||\mathbf{y}_{meas}|| = \frac{1}{2}\sqrt{2\left(1+\varepsilon^2\right)\left(x^2_{\text{true},1}+x^2_{\text{true},2}\right)+4\left(1-\varepsilon^2\right)x^2_{\text{true},1}x^2_{\text{true},2}} \tag{103}$$

Which shows norm conservation of the vectors after matrix multiplication only in case $\varepsilon = 1$.

If we suppose now that the apparatus described by matrix (102) has been used to measure numbers of events in a two-bin histogram, with corresponding covariance matrix $V_{\mathbf{y}}$ given by purely statistical errors in independent entries y_1 and y_2:

$$\mathbf{y}_{\text{meas}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad V_{\mathbf{y}} = \begin{pmatrix} y_1 & 0 \\ 0 & y_2 \end{pmatrix} \tag{104}$$

Since the inverse matrix A^{-1} exists for any $\varepsilon \neq 0$, the unfolded vector $\mathbf{x}_{\text{true}}$ can be obtained by matrix inversion [28]:

$$\mathbf{x}_{\text{true}} = \frac{y_1 - y_2}{2\varepsilon}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{y_1 + y_2}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{105}$$

If all components of the r.h.s. of equation (105) are statistically significant and if ε is too small, the system (100) can be solved without any problem. But if ε is small the problem becomes ill determined, and when in addition the r.h.s. is affected by measurement errors, the exact solution does not make any sense. In this case conventional methods of solving linear systems do not work [28].

D. Least Squares Method

The solution of system (100) represents the numerical solution of the following least square problem [28]

$$\sum_{i=1}^{n}\left(\frac{\sum_{j=1}^{m} A_{ij}x_j - y_i}{\sigma_i}\right)^2 = \min \tag{106}$$

Where we omitted the subscripts *true* and *meas* in favour of readability, and σ_i represents the error with which the $i-$th component of $\mathbf{y}$ is known. Eq. (106) can be cast in a more general matrix form, valid also in case the covariance matrix V_y is not diagonal, it reads:

$$(Ax-y)^T V_y^{-1}(Ax-y) = min \tag{107}$$

This equation is minimized by those y that satisfy [27]:

$$\begin{aligned}(A^T V_y^{-1} A)x &= (A^T V_y^{-1})y \\ Cx &= b\end{aligned} \tag{108}$$

Where $C = A^T V_y^{-1} A$ is a symmetric matrix and consists of the unfolding A matrix reweighted by the measurement errors, while $b = A^T V_y^{-1} y$. The product of the matrices $A^{\#} = (A^T V_y^{-1} A)^{-1} A^T V_y^{-1}$ is called pseudo-inverse. The solution x is a linear transformation of the measurement vector y, which allows for standard error propagation that reads [27]:

$$V_x = A^{\#} V_y A^{\#T} = C^{-1} \tag{109}$$

Although the estimator $\hat{x} = A^{\#} y$ with $E[y] = Ax$ is unbiased being

$$E[\hat{x}] = A^{\#} E[y] = (A^{\#} A)x = x, \qquad \text{with} \quad A^{\#} A = I \tag{110}$$

The solution x presents large unwanted flutctuations [27]. In order to better understand the problem it is instructive to perform singular value decomposition (SVD) on the reweighted unfolding matrix C. Being C a symmetric matrix, it's singular value decomposition will be of the form:

$$C = U\ \Lambda\ U^T$$

Where U is an orthogonal matrix whose columns are the eigenvectors u_j and Λ is a diagonal matrix of eigenvalues λ_i in non increasing order, i.e. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. If all $\lambda_i > 0$ the inverse C^{-1} exists and it can be decomposed as:

$$x = C^{-1} b = U\Lambda^{-1/2}\left(\Lambda^{-1/2} U^T\right) b = U\Lambda^{-1/2} c, \qquad \text{where} \qquad c = \left(\Lambda^{-1/2} U^T\right) b$$

The measured vector y has been transformed to c whose covariance, V_c, is by construction the identity matrix. The vector c is referred to as vector of Fourier coefficients [27] since its components c_j represent the projection of the rotated y vector over the eigeinvectors u_j of matrix U. The solution x can be decomposed as a sum of eigenvectors of U as in [27]:

$$x = \sum_j^m \frac{1}{\sqrt{\lambda_j}} c_j \vec{u}_j \qquad \text{with} \quad c_j = \frac{1}{\sqrt{\lambda_j}}(u_j^T b) \qquad \text{and} \quad V_x = U\Lambda U^T \tag{111}$$

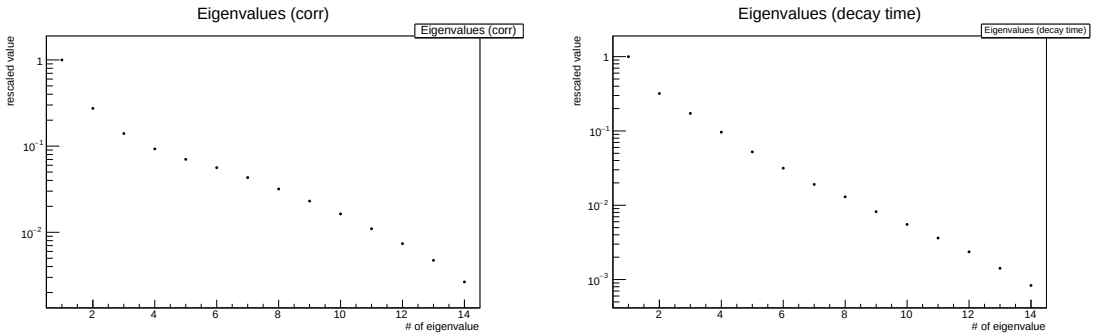
Fourier coefficients which are uncompatible with zero within one sigma ($\sigma_{c_j} = 1 \quad \forall j$) are referred to as significant, while the presence of non-significant ones can make a dominating contribution to the solution x if the eigenvalues λ_i's are small.

In the case of the unfolding matrices of fig. (20), which from now on will be referred to as A_{corr} (20(a)) and A_{t_D} (20(b)) respectively, the eigenvalues of the SVD are plotted in fig. (21). In both cases they decrease up to 3 orders of magnitude, although the Fourier coefficients, which are shown in fig. (22), have a different number of significant components. In both cases their number is lower than the dimension of the unfolded vector, thus the summatory of eq. (111) should be truncated to $k < m$ which, being lower the rank of V_x makes it a singular matrix and thus impossible to use to fit correlated data with the chi-square mnimization method. The number k to which the series in (111) is sample dependent and here is chosen as the number of the first Fourier coefficient lower than 2.

The corrected covariance matrices of the two unfolded solutions (each for each unfolding matrix simulating a different performance in resolution) for a sample of 2×10^6 events are shown in fig 23. They show a high bin to bin correlation, and the average over the bins of the bin correlation coefficient defined as in [27]

$$\rho_j = \sqrt{1 - \left[(V_x)_{jj} \cdot (V_x^{-1})_{jj} \right]^{-1}} \tag{112}$$

Yelds a value of $\rho_{avg}^{corr} = 0.795$ for the A_{corr} case, and a higher $\rho_{avg}^{t_D} = 0.922$ for the A_{t_D} case. The unfolded widths for two 2×10^6 events samples, each one in a different resolution condition, are presented in (unfoldedwidths):



(a) Eigenvalues of the reweighted A_{corr}

(b) Eigenvalues of the reweighted A_{t_D}

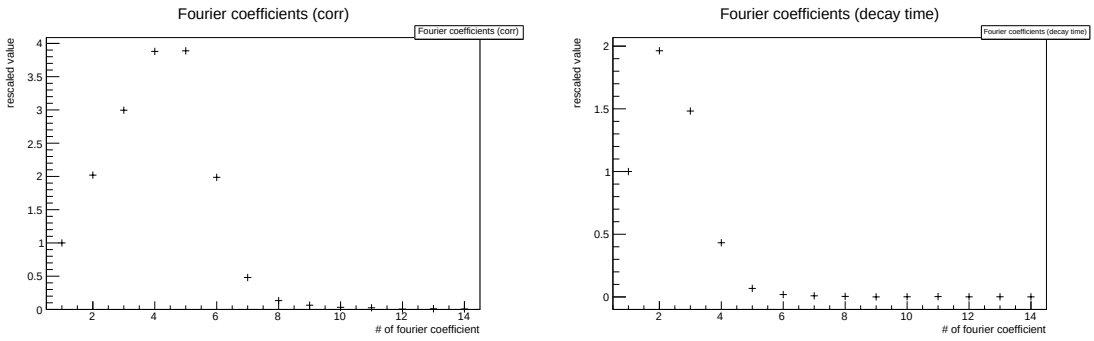
FIG. 21. List of the singular values of the reweighted unfolding matrices for the two resolution cases

E. Regularization techniques: second derivative

The complete solution of eq. (111) leads to an oscillating solution. This spurious oscillatory component is suppressed using some *a priori* knowledge about the solution. Technically this can be achieved by adding the *regularization* term to the expression to be minimized ([28],[27]). In this work we will be incorporating assumptions about the size and the smoothness of the solution. In order to do so, we will add to eq. (107) a curvature limiting requirement [28]:

$$(Ax - y)^T V_y^{-1} (Ax - y) + \tau L^T L x = min \tag{113}$$

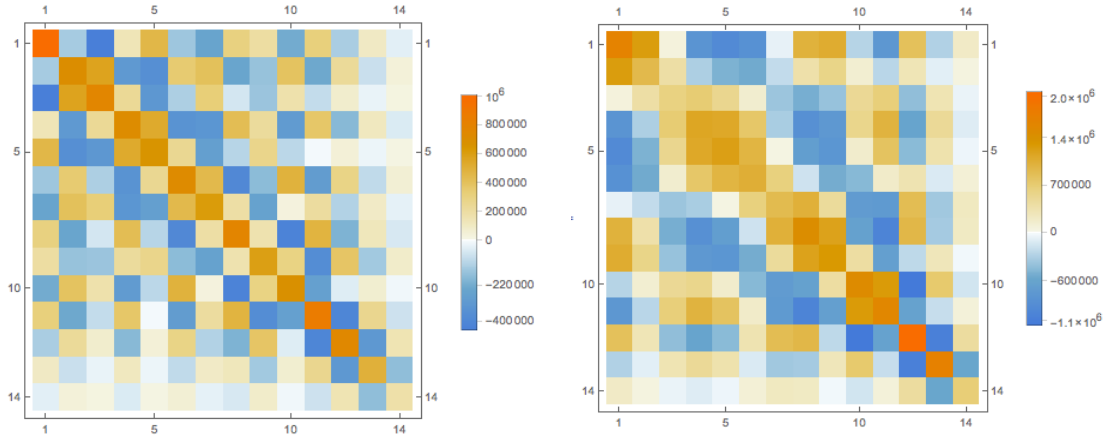
Where L is a matrix which implements the sum of the squares of the second derivatives of the unfolded distribution x [28], and τ is an arbitrary scalar regularization parameter to be determined. In order to implement the square of the discretised second derivative:



(a) Case of correct solution known

(b) Case of guessed solution using t_{D}

FIG. 22. List of the Fourier coefficient of the projection of the measurement vector on the eigenvectors of U



(a) V_x associated to the case of correct solution known

(b) V_x associated to the case of guessed solution using t_{D}

FIG. 23. Correlation matrices of the unfolded vectors calculated as in (111) with a truncated summation

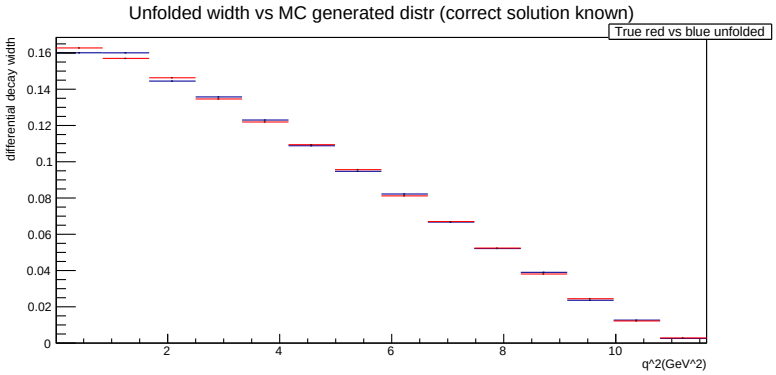
$$\sum_i \Big[(x_{i+1} - x_i) - (x_i - x_{i-1})\Big]^2 \tag{114}$$

the choice of L is:

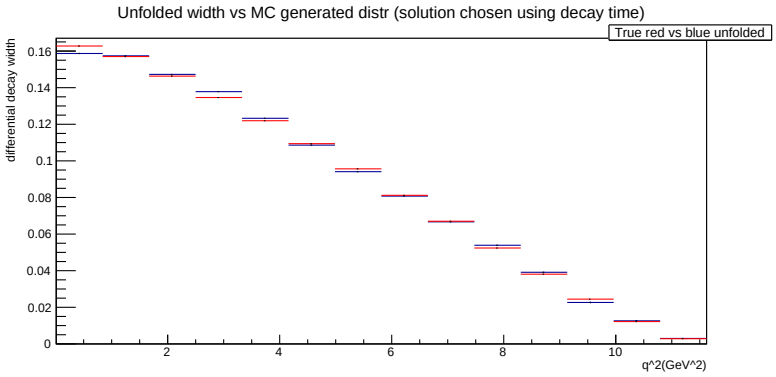
$$L = \begin{pmatrix} -1 & 1 & 0 & 0 & \dots & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & 1 & -2 & 1 \\ 0 & \dots & & & 1 & -1 \end{pmatrix} \tag{115}$$

Adding the extra term $(\tau L^T Lx)$ to eq. (108) will suppress solutions x having large curvatures. Recovering the previous notation, eq. (113) becomes[27]:

$$(C + \tau L^T L)x = b \tag{116}$$



(a) Case of correct solution known



(b) Case of guessed solution using t_{D}

FIG. 24. Unregularized unfolding method, two samples of 2×10^6 events are shown in the two resolution conditions

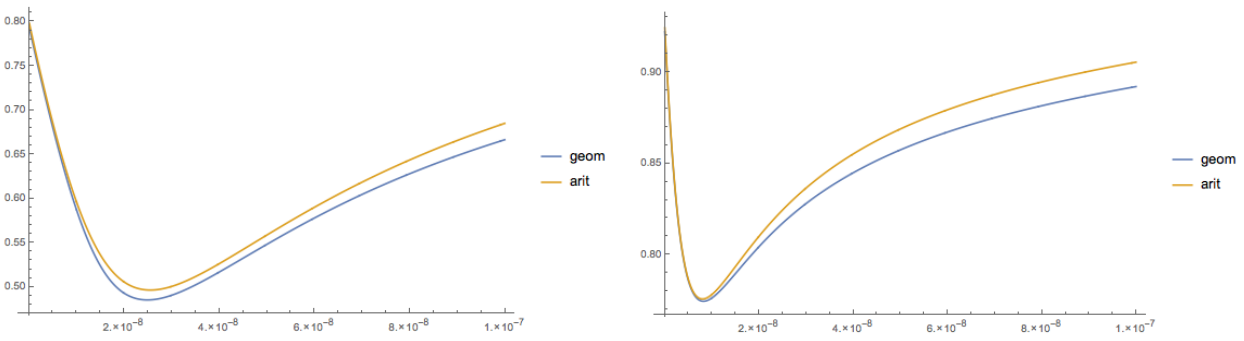
Two simultaneous SVDs cast equation (116) in a form analogous to (111) [27]:

$$x = (R^T)^{-1}(I + \tau S)^{-1}(R^{-1})b \tag{117}$$

Where R is an orthogonal matrix, and S is a diagonal matrix, both built from the simultaneous decompositions of $L^T L$ and C. If it wasn't for the regularization parameter, the solution (117) would be formally the same as (111). The factor $(I + \tau S)$ is referred to as $filter\ factor$ [27] while $c = (R^{-1})b$ are the new fourier coefficients. There is no generally accepted and unique method to determine the regularization parameter τ, but since it explicitly appears in the covariance matrix of the unfolded vector x, it can be chosen to minimise the average correlation between bins of eq. (112)

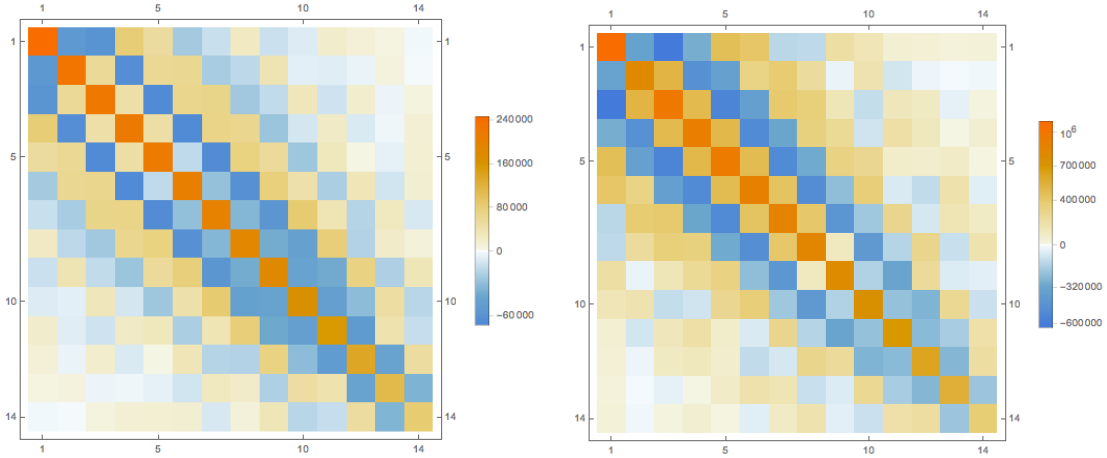
As expected, this choice of tuning of the τ-parameter makes it dependent on the sample to unfold. Given a sample to unfold, the arithmetic and the geometric means over the measured bins of eq. (112) are calculated for a wide range of τ values and the ones that minimises both is chosen. In the case of a 2×10^6 sample the plots of the curves for the average correlation coefficient as a function of the regularization parameter are shown in fig. (25). The inversion matrix (116) is calculated and the covariance matrix of the unfolded solution is derived through standard error propagation, and represented in fig. (26).

An example of two unfolded differential decay widths, in the two resolution conditions, after second derivative regularization is presented in fig. (27). The plotted histograms contain 2×10^6 events and have been corrected for the reconstruction efficiency of fig (16(b)) which stands for the binned efficiency for an event to pass LHCb cuts but to fail reconstruction due to invalid q^2 solution (either immaginary or out of the kinematic boundaries). The two unfolded distributions are characterized by a less significant oscillatory component as compared to the unregularized



(a) Correct solution in q^2 known (b) Ambiguity on the correct solution in q^2

FIG. 25. Variation of the global correlation coefficient (112) with the magnitude of the τ regularization, parameter. Each of the resolution conditions is listed below the plot.



(a) Correct solution in q^2 known (b) Ambiguity on the correct solution in q^2

FIG. 26. Covariance matrix of the regularized unfolded vector x. Each of the resolution conditions is listed below the plot

case, and the errors associated to their entries is reduced although it still differs of a factor 2.5 in the two resolution cases.

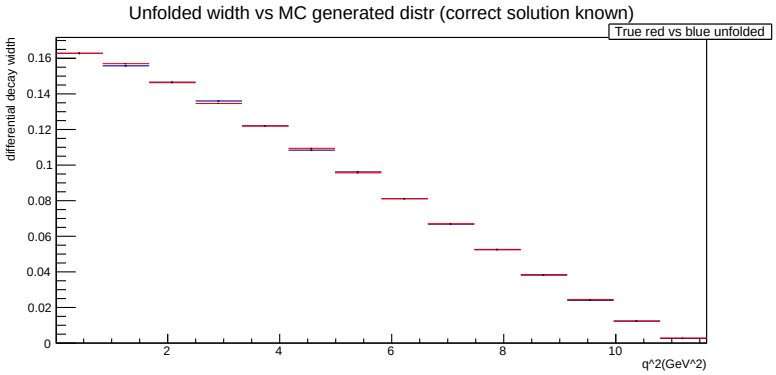
F. Fit results with finite resolution hypothesis

The unfolding procedure is iterated on groups of 100 samples, each group containing samples of an different number of events, all plotted in histograms of 14 bins. The obtained histograms are successively corrected for LHCb geometric and kinematic acceptance of fig. (16(a)). The results are finally fitted with a python script based on TMinuit(). The chis-square minimization function has been changed to take into account bin-to-bin correlations, i.e. in the case of non-diagonal covariance matrix, according to [29]

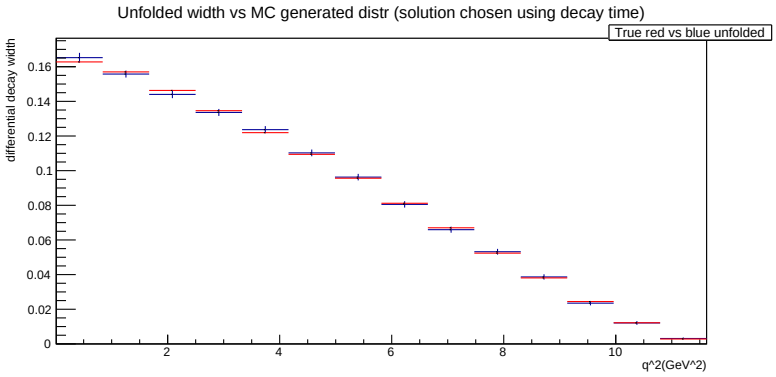
$$\chi^2 = \Delta^T V_x^{-1} \Delta \tag{118}$$

The pulls for the correct q^2 resolution case are presented in fig. (28), while the ones for the ambiguity in the q^2 solution case are displayed in fig. (29).

The pull tests present the same μ deviation at more than one sigma from zero indicating the some bias has been introduced in the unfolding process; This effect is seen as the number of events grows, together with the appearance of



(a) Case of correct solution known



(b) Case of guessed solution using t_{D}

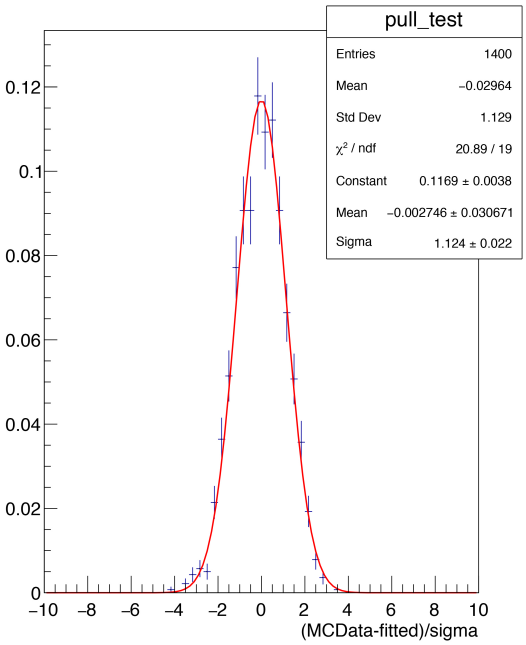
FIG. 27. Regularized unfolding method, two samples of 2×10^6 events are shown in the two resolution conditions

a tail indicating that the fitted value excesses the data, introducing the bias. This effect is seen especially prominent for the case of solution ambiguity, 2×10^6 events. In all cases the errors attributed from the fitter to the parameters is underestimated, the amount of this effect growing with the number of events, up to reaching a 40% value in the ambiguity case when fitting a 2×10^6 events sample.

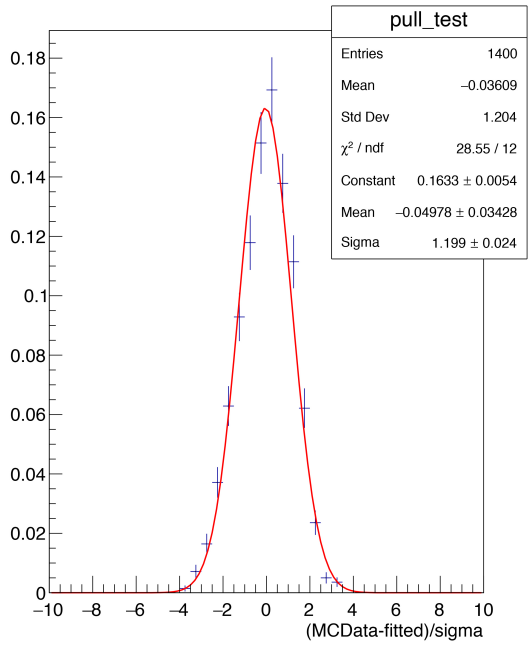
Two exemplificative plots of the fit results, with their relative pulls are shown in fig (30), while the results of the iteration of the fit, with analogous notation as the one introduced in table (VI) and table (VII) are listed in (VIII) and (IX). In order to compare the results with the perfect resolution ones, the parameters have been listed under the effective number of events each sample is composed by, and scaled by the standard deviation of the pull data in order to take into account the error underestimation.

TABLE VIII. Average value of $\tilde{a}_1^{fit} - \tilde{a}_1^{MC}$ and root mean square of $\tilde{a}_1^{fit}$ with different number of events per sample. The first row corresponds to non ambiguity of the correct q^2 solution, while the second row indicates the choice of correct solution based on the D meson decay time

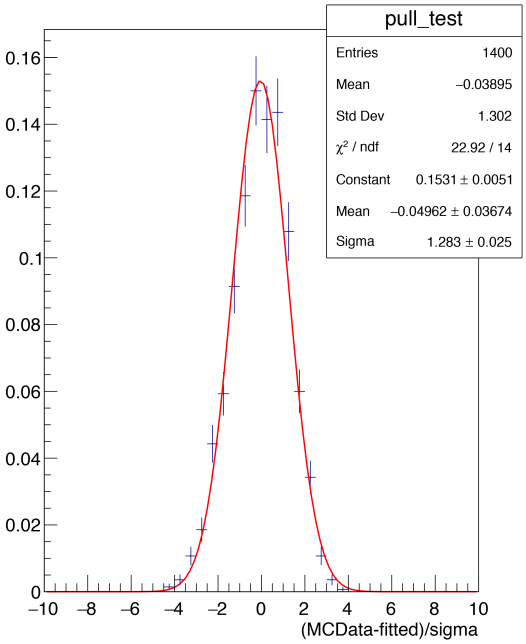
$\langle \tilde{a}_1^{fit} - \tilde{a}_1^{MC} \rangle$, $\tilde{a}_{1,rms}^{fit}$	# of ev.ts $\times 10^6$				
	0.25	0.5	1	1.5	2
# of bins					
14	0.042, 0.067	0.031, 0.046	0.022, 0.036	0.015, 0.035	0.023, 0.033
14	0.052, 0.085	0.038, 0.069	0.027, 0.049	0.021, 0.045	0.028, 0.040



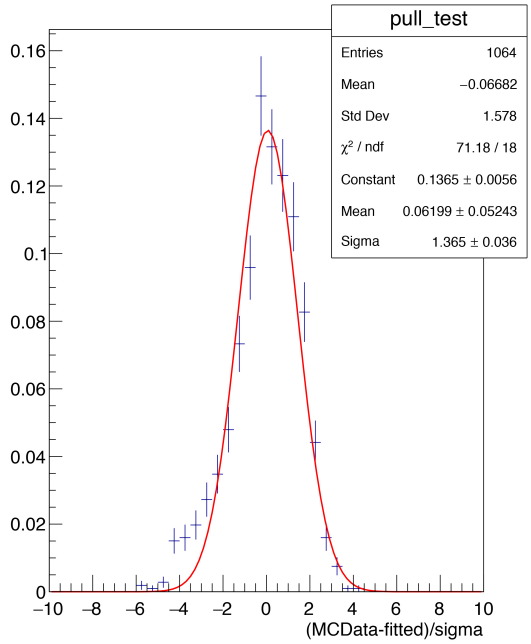
(a) Pull test on 100 samples of 0.25×10^6 effective events in 14 bins, correct solution known



(b) Pull test on 100 samples of 0.5×10^6 effective events in 14 bins, correct solution known

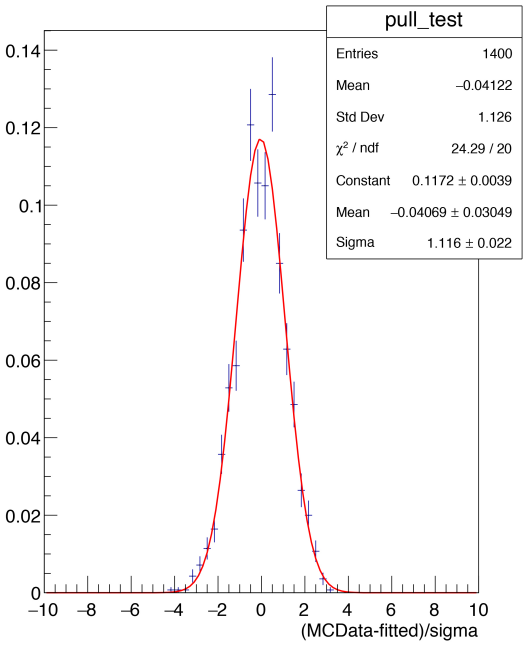


(c) Pull test on 100 samples of 1×10^6 effective events in 14 bins, correct solution known

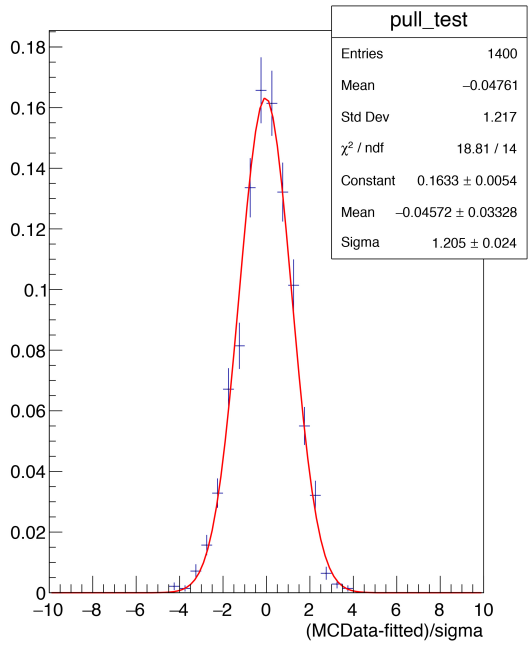


(d) Pull test on 100 samples of 2×10^6 effective events in 14 bins, correct solution known

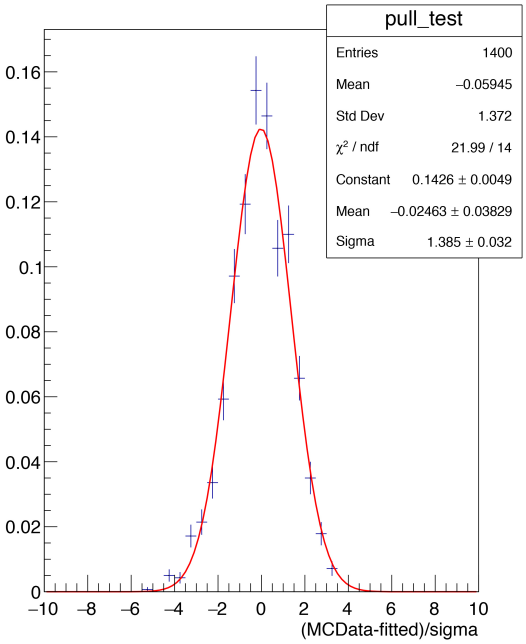
FIG. 28. Different pull tests run on the fit, with fit conditions listed in figure with correct solution known



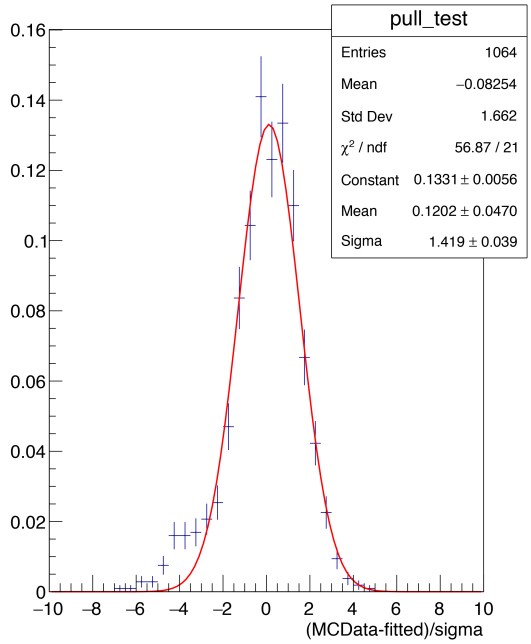
(a) Pull test on 100 samples of 0.25×10^6 effective in 14 bins, solution ambiguity



(b) Pull test on 100 samples of 0.5×10^6 effective events in 14 bins, solution ambiguity



(c) Pull test on 100 samples of 1×10^6 effective events in 14 bins, solution ambiguity

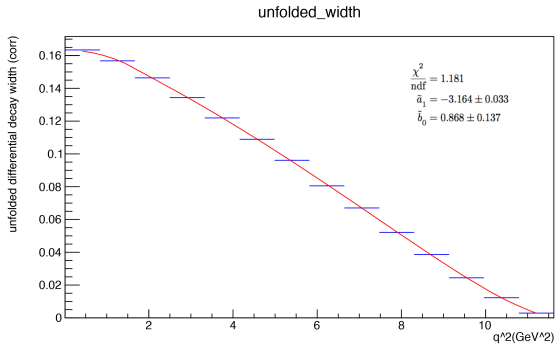


(d) Pull test on 100 samples of 2×10^6 effective events in 14 bins, solution ambiguity

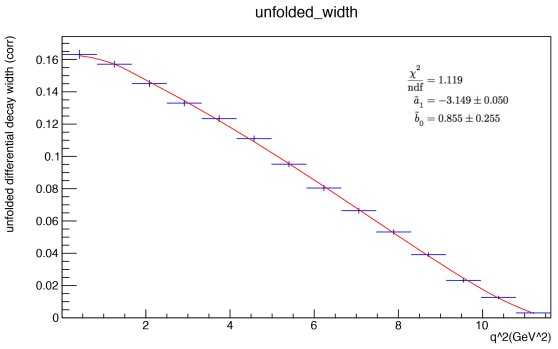
FIG. 29. Different pull tests run on the fit, with fit conditions listed in figure with solution chosen using the D decay-time

TABLE IX. Average value of $\tilde{a}_1^{fit} - \tilde{a}_1^{MC}$ and root mean square of $\tilde{a}_1^{fit}$ with different number of events per sample. The first row corresponds to non ambiguity of the correct q^2 solution, while the second row indicates the choice of correct solution based on the D meson decay time

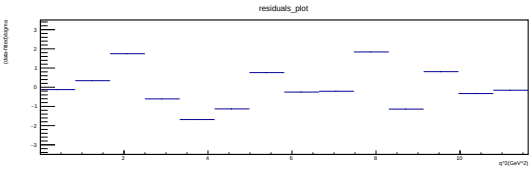
$\langle \tilde{b}_0^{fit} - \tilde{b}_0^{MC} \rangle$, $\tilde{b}_{0,rms}^{fit}$	# of ev.ts $\times 10^6$				
	0.25	0.5	1	1.5	2
# of bins					
14	-0.041, 0.127	-0.050, 0.139	-0.077, 0.116	-0.074, 0.096	-0.077, 0.0817
14	-0.040, 0.150	-0.040, 0.136	-0.052, 0.136	-0.037, 0.095	-0.040, 0.105



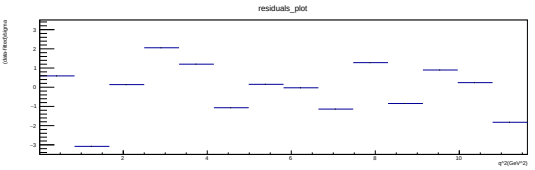
(a) Fit of sample of 2×10^6 events in 14 bins, correct solution known



(b) Fit of sample of 2×10^6 events in 14 bins, correct solution guessed using t_D



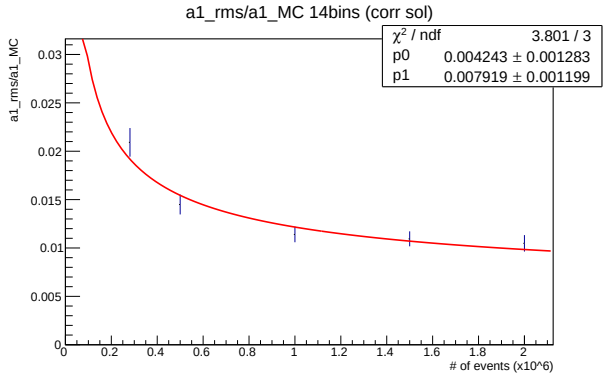
(c) Pull test on the above fit



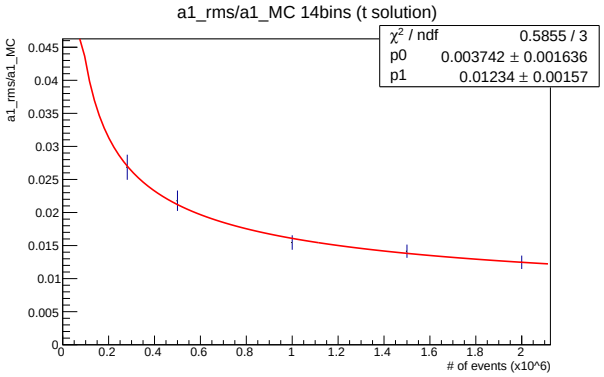
(d) Pull test on the above fit

FIG. 30. Fits run in different resolution conditions

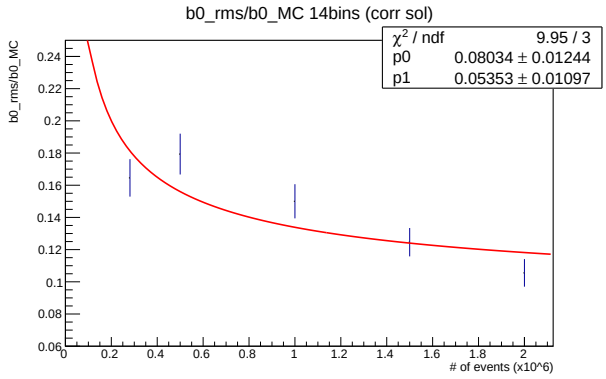
The corresponding scaling of the fitted parameters is shown in fig (31) with a fit executed in an analogous fashion as in the plots of perfect resolution (10).



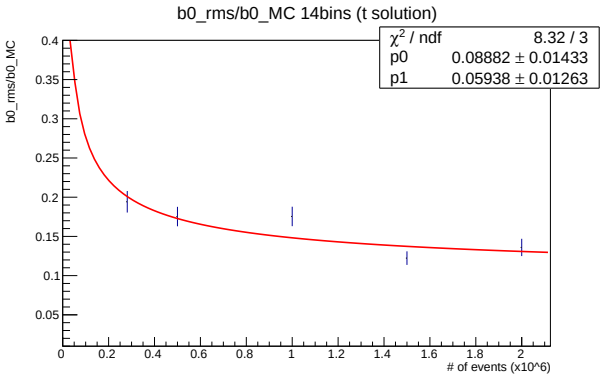
(a) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of events per sample, 14 bins, correct solution in q^2 known



(b) Scaling of $\tilde{a}_{1,rms}^{fit}/\tilde{a}_1^{MC}$ with respect to the # of events per sample, 14 bins, solution in q^2 guessed using t_{D}



(c) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of events per sample, 14 bins, correct solution in q^2 known



(d) Scaling of $\tilde{b}_{0,rms}^{fit}/\tilde{b}_0^{MC}$ with respect to the # of events per sample, 14 bins, solution in q^2 guessed using t_{D}

FIG. 31. Scaling of the relative error on the parameters given a fit with fixed bin number but variable number of events per MC sample.

VII. CONCLUSIONS

As far as the ratio $\tilde{a}_1$ is concerned, the present detector resolution shows good sensitivity, and its extraction from a fit of the differential decay width could improve its value yielding up to 12 times better estimate of its value comparing to the Lattice QCD value [24], if a sample of 1×10^6 events is collected, in case of q^2 solution guessed from the D decay time. The sensitivity would improve to 1% if the correct solution was known.

Moreover, with this simulation we can conclude that a model independent measurement of the ratio between the vector and the scalar form factor parameters, $\tilde{b}_0$, at the present LHCb detector performances would need more than 2×10^6 events, even in case of knowing which is the correct solution of q^2. This would in turn make possible a model independent measurement of R_{D} at LHCb since the scalar form factor contribution is decisive for the $\overline{\mathrm{B}} \to \mathrm{D}\tau\bar{\nu}$ decay width.

Although the error on the fitted $\tilde{a}_0$ shows an irreducible component due to the non compatibility of the p_0 parameter with zero at less than 1σ in both cases, they approximatively show a good behaviour with the increase of the number of events, especially in case of correct solution in q^2 is known. This is not the case for the $\tilde{b}_0$ parameter whose scaling does not decrease with the number of events, in both resolution cases.

This could be due to a bias introduced by the unfolding procedure. In order to investigate if this effect is of systematic nature, one could test the closure of the folding/unfolding procedure used in this work by executing it on

different resolution models and fit functions. Alternatively, one could run the same procedure on a higher number of samples, each one with higher number of events, in order to reduce the statistical fluctuations of $\tilde{b}_{0,rms}$ and verify if at some number of events per sample the error on $\tilde{b}_0$ vanishes.

The parameters of Lattice QCD value [24], or the results from the ZRSR, could be used as an input for the fit to further constrain the overall normalization of the width, but at the price of rendering the measurement model-dependent, and, in the case of the ZRSR, it could only be done after a more precise estimate of the inelastic contributions to the value of $f_+(0)$.

[1] D. Buttazzo, A. Greljo, G. Isidori, D. Marzocca B -physics anomalies: a guide to combined explanations arXiv:1706.07808v1 [hep-ph] Submitted 23 Jun 2017

[2] LHCb Collaboration, R. Aaij et al. LHCb-PAPER-2017-013, CERN-EP-2017-100, arXiv:1705.05802

[3] BaBar Collaboration, J. P. Lees et al. Phys. Rev. D88 (2013), no. 7 072012, BABAR-PUB-13/001, SLAC-PUB-15381, arXiv:1303.0571

[4] Belle Collaboration, S. Hirose et al. Phys. Rev. Lett. 118, 211801 (2017), arXiv:1612.00529

[5] LHCb Collaboration, R. Aaij et al. Phys. Rev. Lett. 115 (2015), no. 11 111803

[6] LHCb Collaboration, R. Aaij et al. Phys. Rev. Lett. 113 (2014) 151601, arXiv:1406.6482

[7] Belle Collaboration, M. Huschle et al., Phys. Rev. D92, 072014 (2015), arXiv:1507.03233 [hep-ex].

[8] M. Bordone,G. Isidori, D. Van Dyk Impact of leptonic τ decays on the distribution B $\to P\mu\bar{\nu}$ decays Eur. Phys. J. C (2016) 76: 360 arXiv:1602.06143v2 [hep-ph] 11 Jul 2016

[9] A. Sirlin, Nucl. Phys. B 196 (1982) 83.

[10] Y. Amhis et al., Averages of b-hadron, c-hadron, and tau-lepton properties as of summer 2016 arXiv:1612.07233

[11] D. Becirevic, N. Kosnik, A. Tayduganov $\overline{\mathrm{B}} \to \mathrm{D}\tau\bar{\nu}$ vs.$\overline{\mathrm{B}} \to \mathrm{D}\mu\bar{\nu}$, 10.1016/j.physletb.2012.08.016, arXiv:1206.4977v4 [hep-ph]

[12] A. Khodjamirian, *Applications of QCD Sum Rules to Heavy Quark Physics*, Lectures at the Helmholtz International Summer School "Physics of Heavy Quarks and Hadrons", July 2013, Dubna, Russia. SI-HEP-2013-15 arXiv:1312.6480v1 *23 Dec 2013*.

[13] M. Neubert, *B Decays and the Heavy-Quark Expansion*, Adv.Ser.Direct.High Energy Phys.15:239-293,1998 arXiv:hep-ph/9702375v1.

[14] A. V. Manohar, M. B. Wise, Heavy Quark Physics Cambridge University Press, 2007

[15] T. Mannel, S. Turczyk and N. Uraltsev Higher Order Power Corrections in Inclusive B Decays, JHEP 1011:109,2010 arXiv:1009.4622v1 [hep-ph] *23 Sept 2010*

[16] I. Bigi, M. Shifman, N.G. Uraltsev, A. Vainshtein. Sum Rules for Heavy Flavor Transitions in the SV Limit Phys.Rev.D52:196-235,1995 12 Dec 1994

[17] N. Uraltsev A 'BPS expansion' for B and D mesons, Phys.Lett.B585:253-262,2004 arXiv:hep-ph/0312001v1 30 Nov 2003

[18] A. Alberti, P. Gambino, K. J. Healey, S. Nandi Precision determination of the CKM element V_{cb} Phys. Rev. Lett. 114, 061802 (2015), arXiv:1411.6560 [hep-ph] 31 Jan 2015

[19] C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016).

[20] R. J. Dowdall, C. T. H. Davies, T. C. Hammant and R. R. Horgan, 10.1103/PhysRevD.86.094510, arXiv:1207.5149 [hep-lat] Phys. Rev. D 86 (2012) 094510

[21] A. K. Rai and N. Devlani, texitProceedings of the XV International Conference on Hadron spectroscopy-Hadron 2013, Nara, Japan, PoS Hadron 2013 (2013) 045; N. Devlani, V. Kher and A. K. Rai, Eur. Phys. J. A 50 (2014) no.10, 154

[22] C. Bourrely, I. Caprini and L. Lellouch *Model-independent description of* B $\to \pi\ell\nu$ *decays and determination of* $|V_{\mathrm{ub}}|$, Phys.Rev.D79:013008,2009; Erratum-ibid.D82:099902,2010 arXiv:0807.2722v3 [hep-ph] 4 Oct 2010

[23] D. Bigi, P. Gambino Revisiting B $\to \mathrm{D}\ell\nu$, Phys. Rev. D 94, 094008 (2016), arXiv:1606.08030v2 [hep-ph] 29 Aug 2016

[24] H. Na, C. Bouchard, G. Lepage, C. Monahan, J. Shigemitsu B $\to \mathrm{D}\ell\nu$ form factors at nonzero recoil and extraction of $|V_{\mathrm{bc}}|$, Phys. Rev. D 92, 054510 (2015), arXiv:1505.03925v3 [hep-lat] 13 Jun 2016

[25] F. James Computing and Network division Minuit: function minimization and error analysis CERN Program Library Long Writeup D506, 1994

[26] LHCb collaboration, A. A. Alves Jr. et al., The LHCb detector at the LHC JINST 3 2008 S08005

[27] Volker Blobel An Unfolding Method for High Energy Physics Experiments, DESY 02-078 (June 2002) arXiv:hep-ex/0208022v1

[28] A. Hoecker, V. Kartvelishvili SVD Approach to Data Unfolding, Nucl.Instrum.Meth.A372:469-481,1996, arXiv:hep-ph/9509307

[29] G. D'Agostini, On the use of the covariance matrix to fit correlated data Nuclear Instruments and Methods in Physics Research A 346 (1994) 306-311