INTERNATIONAL DOCTORAL
SCHOOL OF THE USC

Adrián
Casais Vidal

CERN-THESIS-2023-364
27/10/2023

PhD Thesis

Boosting the discovery potential of LHCb by developing a new search for diphoton resonances and an alternative muon reconstruction technique using GPUs

Santiago de Compostela, 2023

Doctoral Programme in Nuclear and Particles Physics

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

DOCTORAL THESIS

Boosting the discovery potential of LHCb by developing a new search for diphoton resonances and an alternative muon reconstruction technique using GPUs

---

Author:
Adrián CASAIS VIDAL

Supervisors:
Prof. Xabier CID VIDAL,
Prof. Diego MARTÍNEZ SANTOS

DECLARACIÓN DO AUTOR/A DA TESE

D./Dna. **Adrián Casais Vidal**

Título da tese: Boosting the discovery potential of LHCb by developing a new search for diphoton resonances and an alternative muon reconstruction technique using GPUs

Presento a miña tese, seguindo o procedemento axeitado ao Regulamento, e declaro que:

1) A tese abarca os resultados da elaboración do meu traballo.
2) De ser o caso, na tese faise referencia ás colaboracións que tivo este traballo.
3) Confirmo que a tese non incorre en ningún tipo de plaxio doutros autores nin de traballos presentados por min para a obtención doutros títulos.
4) A tese é a versión definitiva presentada para a súa defensa e coincide a versión impresa coa presentada en formato electrónico

E comprométome a presentar o Compromiso Documental de Supervisión no caso de que o orixinal non estea na Escola.

En **Santiago de Compostela**, **31 de xullo de 2023**.

Sinatura electrónica

AUTORIZACIÓN DO DIRECTOR/TITOR DA TESE

D./Dna. **Xabier Cid Vidal**

En condición de: **Titor/a e director/a**

Título da tese: Boosting the discovery potential of LHCb by developing a new search for diphoton resonances and an alternative muon reconstruction technique using GPUs

INFORMA:

Que a presente tese, correspóndese co traballo realizado por D/Dna Adrián Casais Vidal, baixo a miña dirección/titorización, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director/titor desta non incorre nas causas de abstención establecidas na Lei 40/2015.

En **Santiago de Compostela**, **31 de Xullo de 2023**

Sinatura electrónica

AUTORIZACIÓN DO DIRECTOR/TITOR DA TESE

D./Dna. **Diego Martínez Santos**

En condición de: **Director/a**

Título da tese: Boosting the discovery potential of LHCb by developing a new search for diphoton resonances and an alternative muon reconstruction technique using GPUs

INFORMA:

Que a presente tese, correspóndese co traballo realizado por D/Dna Adrián Casais Vidal, baixo a miña dirección/titorización, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director/titor desta non incorre nas causas de abstención establecidas na Lei 40/2015.

En **Santiago de Compostela**, **31 de Xullo de 2023**

Sinatura electrónica

"*Quiero entrar en tu garito con zapatillas.*"

El Canto del Loco

Abstract

The main part of this thesis is devoted to describe a search for a diphoton resonance in the $[5,20]$ GeV mass region with pp collision data taken with the LHCb experiment during 2018 at an energy $\sqrt{s} = 13$ TeV. This work describes the first physics analysis using only photons detected by the calorimeter at LHCb. A precise calibration of the different selection steps is performed, particularly for the trigger selections. World best sensitivity for an ALP produced via gluon fusion and decaying to two photons is expected in the $[5,10]$ GeV region. The result is also interpreted as a B_s^0, B^0 or η_b meson decaying to two photons and expected upper limits on the branching fractions are reported.

The thesis is also devoted to the development of a tracking sequence using the VELO and MUON stations of the LHCb detector with a momentum resolution of $\sim 5\%$ for the GPU implementation of the trigger.

Limiar

A meirande parte desta tese está adicada a describir a busca dunha resonancia decaendo a dous fotóns na ventá de masa en $[5,20]$ GeV utilizando os datos provintes de colisións protón-protón, recollidos polo experimento LHCb durante o ano 2018, cunha enerxía $\sqrt{s} = 13$ TeV. Este traballo describe a primeira análise utilizando tan só dous fotóns detectados polo calorímetro en LHCb. Unha calibración precisa das diferentes capas de selección foi realizada, en particular para as seleccións de *trigger*. Como resultado derívase que o experimento LHCb amosa a mellor sensibilidade para atopar un ALP producido por fusión de gluóns e decaendo a dous fotóns na rexión de masa $[5,10]$ GeV. O resultado tamén se interpreta como un mesón B_s^0, B^0 ou η_b decaendo a dous fotóns e os límites superiores na taxa de desintegración son amosados.

Esta tese está tamén adicada ao desenvolvemente dunha secuencia de reconstrución e trazas utilizando o VELO e as estacións MUON do experimento LHCb que conta cunha resolución de $\sim 5\%$, para o implementación en GPUs do trigger.

Resumen

La mayor parte de esta tesis está dedicada a describir la búsqueda de una resonancia decayendo a dos fotones en la ventana de masa $[5,20]$ GeV utilizando los datos provenientes de colisiones protón-protón, recogidos por el experimento LHCb durante el año 2018, con una energía $\sqrt{s} = 13$ TeV. Este trabajo describe el primer anális que utiliza tan solo dos fotones detectados por el calorímetro en LHCb. Una calibración precisa de las diferentes capas de selección fue realizada , en particular para las selecciones de *trigger*. Como resultado, se deriva que el experimento LHCb muestra la mejor sensitividad para encontrar un ALP producido por fusión de gluones y decayendo a dous fotones en la región de masa $[5,10]$ GeV. El resultado también se interpreta como un mesón B_s^0, B^0 o η_b decayendo a dos fotones y los límites superiores esperados en la tasa de desintegración de estos son mostrados.

Esta tesis está también dedicada al desarrollo de una secuencia de reconstrucción de trazas utilizando el VELO y las estaciones MUON del experimento LHCb que cuenta con una resolución de $\sim 5\%$, para la implementación en GPUs del trigger.

Acknowledgements

I would like to start by thanking my thesis advisors, Xabier and Diego. Thank you both for helping me from the beginning up to this point, not only did you teach me a lot about physics but also helped me grow as a person during all these years. I would also like to thank Carlos for encouraging me onto taking this path back when I was deciding which topic to choose for my bachelor thesis, that decision has gone a long way...

I would like to thank my mates at IGFAE and the University of Santiago, especially Alberto, Moncho, Pablo and Asier with whom I've had long talks about work and, probably more importantly, non-work matters. The Allen team holds a special place for me, especially Dani, Roel and Dorothea. Thank you guys for all the effort you have put onto teaching me about the framework, GPUs, computing... you name it. I also want to make a especial mention to Vava. Thanks a lot for all the advice and trust you have given me during these years. I thank Cristina for always noticing right on the spot about new MC requests while on our role as Simulation Liasions of the QEE working group. I will also thank Martino, Lorenzo and Miguel and the rest of the QEE working group for all the advice you have provided. I also want give big thanks to Carla, Max, Maria, Pere and the rest of the people from the Radiative WG: without your help I would have never be able to decode the calorimeter mysteries. Very special thanks go to Renato, thanks for the long discussions, the good ideas and all of the material that gave rise to the VELO-MUON matching.

I want to make a BIG shout out to Titus. You have no idea how much I appreciate all the effort you have put not only to the ALPs analysis but also into teaching me about statistics, reading this very thesis and more stuff that was way over your responsibility. I really owe you big time.

Nun xeito mais persoal gustaríame dirixerme ás persoas máis cercanas e que deixaron pegada en min nestos últimos anos. Empezarei por reiterar a Xabi, quen leva aturándome non só durante esta tese senón dende aquel TFG que puxo as primeiras pedras no camiño. Grazas por toda a paciencia que tiveches conmigo nestes anos. Os compañeiros da carreira, cos que sinto un vínculo vital pois tivemos que ser quen de ir superando obstáculos de maneira simultánea, Alexandre Brea, Gabriel, Julián, Gonzalo e Saúl (quen realmente chegou máis tarde pero peretence a este club por méritos propios). Quenes forman parte dos meus amigos mais cercanos, os que sempre estiveron e os que chegaron para quedarse: Alex Muñiz, Fidel, Berto, Jacobo, Miguel Carrodeguas. Tamén a toda a pandilla de Boiro, Abraham, Abel, Dani, Chevi e Ramiro. Gracias por permanecer ao meu lado. Quero tamén facer unha mención moi especial a Ana Moreiras. Sen aquela viaxe ao CERN no ano 2011 nada desto tería chegado.

Gardo o lugar especial para a miña familia, quenes sempre estiveron aí dende o principio dos tempos para apoiarme e tamén guiarme cando me fixo falta: Mamá, Papá, Marta, Ana.

Finalmente quero agradecer a paciencia, esforzo e resiliencia que a persoa que me acompaña e sempre me ten que aturar. Gracias por apoiarme nos meus proxectos, gracias por poñerme no alto das prioridades e gracias empurrarme a vivir as miñas aventuras de investigador loco. Gracias por todo, Sara.

Preface

This thesis document embodies the summary of my work at the LHCb collaboration from the end of the year 2018 up to the first half the year 2023. The first three chapters serve as an **introduction** to the theory, statistical tools and the experimental setup at the LHCb experiment. Chapters 4 and 5 describe the **objectives**, **methods** and **results** of the two main topics of this thesis. Respectively, Chapter 4 describes the Trigger strategy for the Run 3 data taking period of LHCb and my work on an alternative reconstruction algorithm designed as a part of it. In Chapter 5 I describe the search for an axion-like particle decaying to two photons at the LHCb experiment and **discuss** the results obtained. Finally, Chapter 6 serves as a final wrap-up of the work developed in Chapters 4 and 5 and the **results** are again briefly discussed.

List of Figures

List of Tables

Contents

Chapter 1

Theoretical foundation

1.1 The Standard Model

The Standard Model (SM) [1–4] of Particle Physics provides, to this date, the most precise description of the fundamental pieces of matter and their interactions. The SM is a quantum field theory (QFT), thus the description of nature is done via mathematical objects called fields. These fields can be either fermions, associated to half-integer spin values or bosons, associated to integer spin values. Regular matter is constituted by fermions and the interactions are made possible thanks to the gauge bosons. The Higgs boson does not play a role as an interaction mediator but is a consequence of the remaining SM particles not being massless.

The SM is a quantum field theory [5] governed by the $SU(3)_C \times SU(2)_L \times U(1)_Y$ group. This means that the lagrangian ($\mathscr{L}$) that describes the particle content and their interactions, is invariant under the gauge transformations. Moreover, the fields are classified in terms of the representation of the group that they belong to. On top of the gauge symmetries, the SM is a relativistic field theory that should remain valid for all inertial frames, then Lorentz invariance is also imposed.

1.1.1 Quantum electrodynamics

The $U(1)_{\text{EM}}$ gauge symmetry group describes the interaction of electrically charged fermions. A spin-$\frac{1}{2}$ field, ψ with mass m and charge q has a Lagrangian that can be written as follows, where γ^μ are the Dirac matrices.

$$\mathscr{L} = \bar{\psi}(x)(i\partial_\mu \gamma^\mu - m)\psi(x) \tag{1.1.1}$$

Applying a local phase transformation to a fermion field yields:

$$\psi(x) \to e^{iQ\alpha(x)}\psi(x) \tag{1.1.2}$$

In order to maintain the lagrangian invariant under this transformation one can add a vector field A_μ that replaces the regular ∂_μ derivative to the so-called covariant derivative

$$D_\mu = \partial_\mu + iQA_\mu. \tag{1.1.3}$$

By defining $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, that describes the propagation of the A_μ field one can define the gauge invariant lagragian of Quantum Electrodynamics (QED):

$$\mathscr{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{1.1.4}$$

QED [6] is not a fundamental piece of the SM but a set of rules that are superseded by the electroweak theory, which will be briefly described in the following section.

1.1.2 The electroweak interactions of the Standard Model

The Glashow-Weinberg-Salam (GWS) electroweak theory (EW) [1–3] is based on the $SU(2)_L \times U(1)_Y$ symmetry group. The effect of the $U(1)_Y$ on a fermion field is to change their weak hypercharge Y through a phase shift. On the other hand $SU(2)_L$ implies that fermion fields are composed of the left-handed and right-handed components. Left-handed transform as $SU(2)_L$ doublets, written compactly as a form of vector:

$$L = \left[\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix} \right] \tag{1.1.5}$$

$$Q = \left[\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right] \tag{1.1.6}$$

These are charged under the weak isospin. Left-handed neutrinos have $T_3 = 1/2$ while left-handed leptons have $T_3 = -1/2$. Similarly, left-handed up-type quarks have $T_3 = 1/2$ and left-handed down-type quarks have $T_3 = -1/2$. Right-handed transform as $SU(2)_L$ singlets, which do not carry weak isospin, thus $T_3 = 0$. In the following they are presented as a form of vector, in the same manner as in Eq. 1.1.5 and Eq. 1.1.6:

$$E_R = \begin{pmatrix} e_R \\ \mu_R \\ \tau_R \end{pmatrix}, \; \nu_R = \begin{pmatrix} \nu_{e,R} \\ \nu_{\mu,R} \\ \nu_{\tau,R} \end{pmatrix}, \; U_R = \begin{pmatrix} u_R \\ c_R \\ t_R \end{pmatrix}, \; D_R = \begin{pmatrix} d_R \\ s_R \\ b_R \end{pmatrix} \tag{1.1.7}$$

The electric charge satisfies the following relationship between T_3 and Y:

$$Q = T_3 + \frac{Y}{2} \tag{1.1.8}$$

Similarly to the QED ($U(1)_{\text{EM}}$) case, a covariant derivative is defined to leave the Lagrangian containing $SU(2)_L$ doublets and Y charged fields unchanged:

$$D_\mu = (\partial_\mu + ig\sigma_a W_\mu^a + ig'\frac{Y}{2}B_\mu) \tag{1.1.9}$$

In the former equation the three weak bosons of the $SU(2)_L$ symmetry (W_μ^i) appear in conjunction with the $SU(2)_L$ generators: $T_i = \frac{\sigma_i}{2}$, the Pauli matrices and its coupling strength

g. The hypercharge boson of $U(1)_Y$ symmetry, B_μ, generated by the scalar phase $\frac{Y}{2}$ and its corresponding coupling strength g' are also present.

The kinetic terms of the EW lagrangian, leaving outside the mass terms take the following form, using the compact vector notation introduced in Eqs. [1.1.5 - 1.1.7]:

$$\begin{aligned}\mathscr{L}_{\text{kinematic}} = & i\bar{L}_i\gamma^\mu D_\mu L_i \\ & + i\bar{Q}_i\gamma^\mu D_\mu Q_i \\ & + i\bar{E}^i_R\gamma^\mu D_\mu E^i_R + \bar{\nu}^i_R\gamma^\mu D_\mu \nu^i_R \\ & + i\bar{U}^i_R\gamma^\mu D_\mu U^i_R + \bar{D}^i_R\gamma^\mu D_\mu D^i_R\end{aligned} \tag{1.1.10}$$

1.1.3 Electroweak symmetry breaking

While the structure presented in the previous section preserves the observed different chirality states it does not allow for massive gauge bosons or fermions. In order to enable gauge invariant mass terms for these fields one needs to incorporate an extra $SU(2)_L$ complex doublet field,

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix} \tag{1.1.11}$$

and an extra term to $\mathscr{L}$, considering $V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$

$$\mathscr{L}_{\text{scalar}} = \frac{1}{2}(D_\mu\phi)^\dagger(D^\mu\phi) + V(\phi) \tag{1.1.12}$$

It can be shown that for $-\mu^2 > 0$ and $\lambda > 0$ the minimum of $V(\phi)$, the Vacuum Spectation Value (VEV), occurs for $\|\phi\| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}}$, which allows us to expand the scalar field in terms of v, giving rise to (up to a phase term) the Higgs field:

$$\phi(x) = \begin{pmatrix}0 \\ \frac{1}{\sqrt{2}}(v + h(x))\end{pmatrix} \tag{1.1.13}$$

This process is known as Spontaneous Symmetry Breaking (SSB) and it uses the Higgs-Englert-Brout mechanism [7, 8]. It has the effect of producing three massive gauge bosons (W^+, W^- and Z^0) and a massless one (A, the photon). Invariant formulation of massive fermion fields is also enabled. Residual to this process, an additional massive scalar particle is predicted: the Higgs boson.

Using $\sin\theta_W \equiv \frac{g'^2}{g^2+g'^2}$ to define the Weinberg angle, θ_W, the new fields in terms of the $SU(2)_L \times U(1)_Y$ gauge bosons take the following form:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \tag{1.1.14}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{1.1.15}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{1.1.16}$$

The masses of the $W_\mu^\pm$ and Z_μ bosons can be calculated to be:

$$m_W = \frac{gv}{2}, \;\; m_Z = \frac{\sqrt{g^2+g'^2}v}{2} \tag{1.1.17}$$

whereas the photon field, A_μ remains massless.

Similarly, a gauge invariant Lagrangian involving the scalar field and the left (e_L, u_L, d_L) and right-handed (e_R, u_R, d_R), together with their couplings (λ_e, λ_u, λ_d) can be written as:

$$\mathscr{L}_{\text{Yukawa}} = \lambda_e \tilde{e}_L \phi e_R + \lambda_u \tilde{u}_L \phi u_R + \lambda_d \tilde{d}_L \phi d_R + h.c. \tag{1.1.18}$$

which, after the scalar field takes its minimum value gives rise to gauge invariant mass terms:

$$m_f = \lambda_f \frac{v}{\sqrt{2}} \tag{1.1.19}$$

The $SU(2)_L \times U(1)_Y \to U(1)_{\text{EM}}$ symmetry breaking has taken place. In the process three gauge bosons become massive while the dimension of the gauge group goes from four to one. The three massive gauge bosons acquire longitudinal polarisation degrees of freedom, which are effectively interchanged by three out of the four scalar field of the $SU(2)$ doublet described in Eq. 1.1.11.

1.1.4 Quantum chromodynamics

The $SU(3)_c$ symmetry is associated with the gluons, the corresponding gauge bosons. The corresponding quantum chromodynamics lagrangian (QCD) [4] lagrangian is obtained by replacing the ordinary derivative ∂_μ by the covariant derivative, similar to the QED and EW theories:

$$D_\mu = (\partial_\mu - i\alpha_s \frac{\lambda_a}{2} G_\mu^a(x)) \tag{1.1.20}$$

In this equation, α_s is the coupling strength, λ_a are the $SU(3)_c$ generators and $G_\mu^a(x)$ correspond to the gluon fields. Under these definitions one can define a gauge invariant QCD Lagrangian for n_q fermionic quark fields:

$$\mathscr{L}_{\text{QCD}} = \sum_{n_q} \bar{q}(iD_\mu \gamma^\mu - m)q - \frac{1}{4} G_{\mu\nu}^a(x) G_a^{\mu\nu}(x) \tag{1.1.21}$$

One consequence of the $SU(3)_c$ gauge symmetry group is that both the quark q fields and the gluon $G^a_{\mu\nu}$ gauge bosons are charged under the color quantum number. This fact implies the self interaction of gluons, as opposed to the QED case.

1.1.5 The running of α_s

A very important feature of QFT is that, naturally, the coupling parameters of the theory change depending on the energy scale. This effect is due purely to quantum effects and it is not present in the classical version of the theories. The relative energy scale will be denoted by μ.

This quality is especially interesting for the QCD theory that describes the formation of hadrons, made up by the quarks. The equation that controls the evolution of the coupling constant α_s, at LO, is the following [9]:

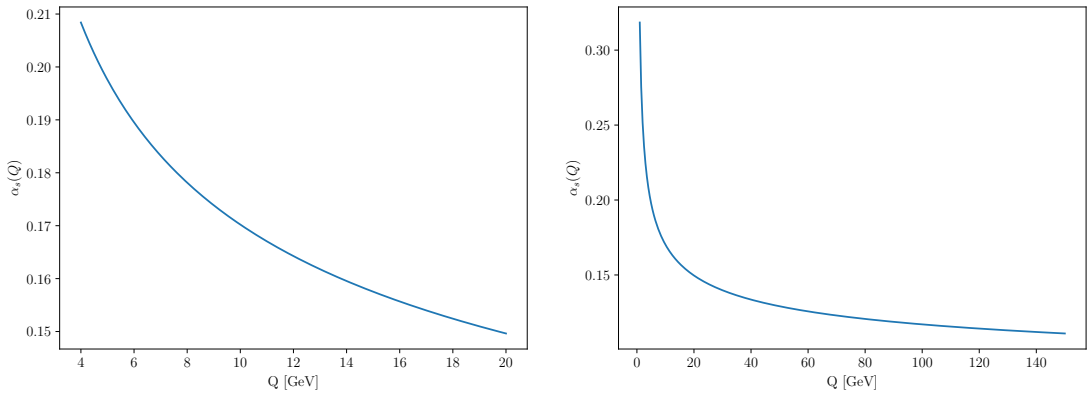

Figure 1.1.1: Value of α_s. Left: α_s values for $Q \in [5, 20]$ GeV. Right α_s values for $Q \in [1, 150]$ GeV calculated with `rundec` [10]

$$\mu^2 \frac{d\alpha_s}{d\mu^2} = \beta(\alpha_s) = -\frac{\alpha_s^2}{4\pi}\left(11 - \frac{2n_f}{3}\right) \tag{1.1.22}$$

The solution to this differential equation in the energy range [1,150] GeV with the initial condition that the value a the Z^0 pole, $\alpha_s(91.8\,\text{GeV}) = 0.1185$ is shown in Fig. 1.1.1.

This particular behaviour of α_s has a number of consequences:

- For interactions as low as $\Lambda_{\text{QCD}} \sim 200\,\text{MeV}$ the interaction strength diverges leading to quarks and gluons not being able to propagate freely and rather be confined within colourless states: hadrons.
- In the high energy regime QCD becomes low-interacting: suited to apply perturbation theory methods. This regime is called *asymptotic freedom* [11, 12] and David Gross,

Frank Wilczek and David Politzer, responsible for its discovery, were awarded with the Nobel Prize in 2004.

1.1.6 Photon production processes

The higher source of photons in pp collisions are π^0 and η meson decays, these type of light mesons are copiously produced in these type of hadronic collisions. Furthermore, there are two main sources of highly energetic photons in pp collisions which are not a byproduct of hadron decays: prompt photons that arise from hard QCD processes, these can be either *direct* photons if they are produced directly from the vertex or *fragmentation* if they are radiated off a final state quark.

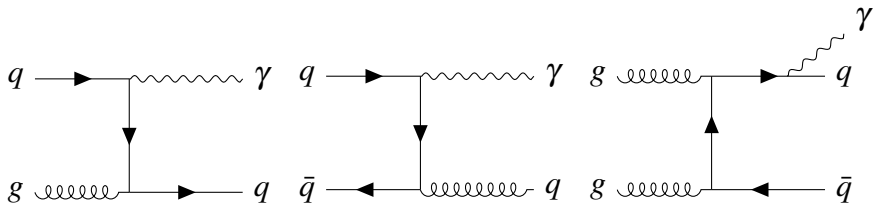

Figure 1.1.2: Diagrams for leading order photon production processes at hadron colliders.
Left: quark-gluon Compton scattering. Middle: quark-antiquark annihilation.
Left: photon radiated off of a final state quark (fragmentation process)

1.1.7 Validity of the SM

The SM is the most complete theory of particles and interactions and yet, it fails to explain all the observed phenomena. The difference between matter and antimatter is measured by the charge-conjugation (C) and parity (P) operators, which together lead to a CP transformation. If the SM was invariant under CP then matter and antimatter would behave in the same manner.

The main source of CP violation (CPV) in the SM arises from $W^{\pm}$ boson-mediated charged-current interactions with quarks. The amount of CPV in the weak sector is not enough to explain the Barion Asymmetry of the Universe (BAU) [13]. In very simple terms, BAU is necessary for matter to exist, otherwise it would annihilate in contact with anti-matter. No CPV is found in the strong sector though non-vanishing terms in the lagrangian suggest it should exist, this idea motivates the concept of axions. This topic will be furtherly developed in Sec. 1.1.8.

The nature of neutrinos is also an open field of research. The SM fails to explain its oscillation [14, 15]. The change in neutrino lepton flavor as they propagate through space

implies they should have a finite mass. In the SM neutrino have a mass which is equal to zero which contradicts the former statement.

Astronomical observations [16] hint the existence of massive particles that would only be allowed to interact very weakly with regular matter. This unobserved phenomena is known as *dark matter* and *dark energy* and its existence can only be hinted through the observation of very massive astronomical objects such as galaxies.

1.1.8 The strong $\mathscr{C}\mathscr{P}$ problem

The QCD lagrangian, in Eq. 1.1.21, can incorporate the operator built on top of the field strength tensor operator $G_{\mu\nu,a}\tilde{G}^{\mu\nu,a}$:

$$\mathscr{L}_{\text{eff}} = \mathscr{L}_{\text{QCD}} + \frac{\theta\alpha_s^2}{32\pi^2} G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} \tag{1.1.23}$$

It can be seen by expressing explicitly the summation terms in θ term, $\theta G_{\mu\nu,a}\tilde{G}^{\mu\nu,a}$ that both parity and time reversal symmetries are violated. This is known as the strong **CP** problem [17, 18]. Even though the term can be expressed as a partial derivative of a current:

$$\partial_\mu K^\mu = G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} \tag{1.1.24}$$

and typically one can argue that the corresponding action associated to this term would vanish, this is not true for QCD. The reason why this term cannot be suppressed through a clever gauge transformation is because the gluon field couples both to left and right-handed fermions. Conversely, that is why no θ term can arise from the EW sector.

The θ term is shifted by the effect of the quark mass terms, encoded as the $\mathscr{M}$ matrix, product of the Higgs mechanism and the Yukawa terms described in Eq. 1.1.18.

$$\bar{\theta} = \theta - \arg\det\mathscr{M} \tag{1.1.25}$$

One of the most relevant consequences of the non-vanishing $\bar{\theta}$ term is that it predicts a non-zero neutron electric dipole moment (EDM-d_N). The current bounds [19] on the d_N value show that:

$$d_N = (0.0 \pm 1.1_{\text{stat}} \pm 0.2_{\text{sys}}) \cdot 10^{-26}\ \text{e}\cdot\text{cm} \implies \bar{\theta} \leq 10^{-9} \tag{1.1.26}$$

Such a tiny value of $\bar{\theta}$ is deemed as requiring an explanation. A plausible solution to this "uglily" small value is to introduce a new symmetry that will give rise to a new observable particle: the QCD axion.

The QCD axion

The idea that such a fine-tuned, close to zero, $\bar{\theta}$ parameter took such value accidentally was not compelling in the community. A plausible path to obtaining such value would be to exploit some symmetry of the theory.

A path towards finding a solution is one that considers $\bar{\theta}$ as a dynamical quantity rather than a fixed parameter that would default to zero under the action of certain forces upon it. A solution of this sort was proposed by Peccei and Quinn [20, 21]. They added an extra global $U(1)_{PQ}$ symmetry which gives rise to the following lagrangian, in which ϕ, a scalar field is introduced:

$$\mathscr{L} = \sum_j \left[\bar{Q}_j i\gamma_\mu D^\mu Q_j - \left(y_j \bar{Q}_{L,j} Q_{R,j}\phi + \text{h.c.}\right)\right] + \frac{\theta\alpha_s^2}{32\pi^2} G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} + \partial_\mu\phi\partial^\mu\phi - V(\phi\phi^\dagger) \tag{1.1.27}$$

It contains both left and right-handed lepton-antilepton interactions, the Yukawa couplings with the ϕ field, the **CP** violating $\bar{\theta}$ term and an even power potential function $V(\phi\phi^\dagger)$, in the same fashion as that of the Higgs mechanism. The global $U(1)_{PQ}$ symmetry protects this Lagrangian from changing under the global the transformation:

$$\phi \xrightarrow{U(1)_{PQ}} e^{i2\alpha}\phi; \quad Q_i \xrightarrow{U(1)_{PQ}} e^{-i\alpha\gamma^5} Q_i \tag{1.1.28}$$

The $U(1)_{PQ}$ is spontaneously broken, and $V(\phi\phi^\dagger)$ has the right terms so that:

$$<\phi(x)> = f_a e^{i<a>/f_a} \tag{1.1.29}$$

Where f_a is the symmetry breaking scale of the axion field, and $<a>$ is the expectation value of the axion field, $a(x)$.

As in the regular Higgs mechanism the fermions of the theory acquire masses following the Yukawa couplings: $m_q = y_q f_a e^{i<a>/f_a}$ such that Eq. 1.1.25 becomes:

$$\bar{\theta} = \theta - \sum_i \arg y_i - N_f <a>/f_a \tag{1.1.30}$$

By this process the $\bar{\theta}$ parameter is no longer a parameter, but a dynamical quantity that depends on the VEV of $a(x)$ which leads to an effective lagrangian containing a $aG_{\mu\nu,a}\tilde{G}^{\mu\nu,a}$ term.

$$\mathscr{L}_{\text{eff}} = \mathscr{L}_{SM} + \frac{\alpha_s^2\bar{\theta}}{32\pi^2} G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} + \frac{\alpha_s^2}{32\pi^2}\frac{\xi}{f_a} a G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} - \frac{1}{2}\partial_\mu a\partial^\mu a + \mathscr{L}_{\text{int}}(\partial_\mu a, \psi) \tag{1.1.31}$$

The ξ and f_a parameters in conjunction with $\mathscr{L}_{\text{int}}$ define the different couplings that the axion field has to any other field in the SM.

It can be shown by exploiting the properties of the $aG_{\mu\nu,a}\tilde{G}^{\mu\nu,a}$, that the VEV under this particular potential is, conveniently:

$$<a> = -\frac{\bar{\theta}}{\xi} f_a \tag{1.1.32}$$

The last step is simply to expand the physical field in terms of this expectation value:

$$a_{\text{phys}} = a(x) - <a> \tag{1.1.33}$$

Finally the lagrangian in Eq. 1.1.31 can be re-written in terms of the physical operator a_{phys} as:

$$\mathscr{L}_{\text{eff}} = \mathscr{L}_{SM} + \frac{\alpha_s^2}{32\pi^2}\frac{\xi}{f_a} a_{\text{phys}} G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} - \frac{1}{2}\partial_\mu a_{\text{phys}}\partial^\mu a_{\text{phys}} + \mathscr{L}_{\text{int}}(\partial_\mu a_{\text{phys}}, \psi) \tag{1.1.34}$$

Where the offending $\bar{\theta}$ term is finally missing and **CP** violation coming from the QCD sector is set to 0.

1.1.9 Axion-like particles

Axion-like particles (ALPs) appear whenever a $U(1)_{PQ}$ symmetry is broken. In the case the current is not conserved and thus the $U(1)$ symmetry is only approximate, ALPs acquire a mass, m_a, and are considered pseudo-Nambu-Goldstone bosons (pNGB). m_a depends on the nature of the $U(1)$ global symmetry and lies below the correspondent NP scale: $M_{NP} \sim 4\pi f_a$. f_a, the scale of spontaneous symmetry breaking is also known as the decay constant and globally controls the ALP coupling to the SM. In general, ALPs interact with the SM and behave like the Axions described in Sec. 1.1.8.

Below the symmetry breaking scale, the effective lagrangian is similar to that described in Eq. C.1.1. The couplings to the gauge sector appear in the form of these terms: $aG_{\mu\nu,a}\tilde{G}^{\mu\nu,a}$, similar to that those of the QCD axion related to the strong CP problem.

$$\mathscr{L}_{\text{int}} = \frac{a}{4\pi f_a}\left[\alpha_s c_3 G_{\mu\nu,a}\tilde{G}^{\mu\nu,a} + g^2 c_2 W_{\mu\nu,a}\tilde{W}^{\mu\nu,a} + g'^2 c_1 B_{\mu\nu}\tilde{B}^{\mu\nu}\right] \tag{1.1.35}$$

$W_{\mu\nu,a}$ and $B_{\mu\nu}$ are the EW field strength tensors, $G_{\mu\nu}$ is the strong force field strength tensor and a is the ALP field. The cross-sections that result from these lagrangian. In general axions would be in reach for the LHC but also tied to a higher energy scale $\Lambda_* \simeq g_* f_a$ in general out of reach to the LHC. This general description is the base for different physics cases in which ALPs would be produced at the LHC and would be within the GeV mass scale.

- **Heavy QCD axion**: In principle the QCD axion emerging as solution of the Strong CP Problem is not necessarily expected to be as massive as a few GeV. Though, by introducing extra complexity to the model the QCD axion would easily be allowed to be in the [2,20] GeV mass range. See, for example [22].
- **ALP mediated Dark Matter (DM)**: in this scenario ALPs would be a key ingredient to the DM freeze out by connecting the SM to the DM sector. In principle ALPs would couple feebly to the gauge sector of the SM allowing for the heavier DM particles to decay to gluons or photons. A recent review can be found in [23].
- **Supersymmetric R-axion**: the $U(1)_{PQ}$ symmetry breaking can be realised as an R-parity symmetry breaking. As in the above example, supersymmetric superpartners would be out of reach to the LHC but the pNGB that arise from their broken symmetries would open a portal at the LHC reach [24].

Considering ALPs that couple both to gluons and photons and considering $K_{gg} \sim 2$, a dimensionless factor that includes Next-to-Next-to-Leading-Order (NNLO) corrections to the gluon field, one can obtain the following partial cross-sections [25]:

$$\Gamma_{\gamma\gamma} = \frac{\alpha_{\rm em}^2 (c_2 + \frac{5}{3}c_1)^2 m_a^3}{64\pi^3 f_a^2}, \quad \Gamma_{gg} = K_{gg} \frac{\alpha_{\rm s}^2 c_3 m_a^3}{8\pi^3 f_a^2} \tag{1.1.36}$$

If one considers the $\mathcal{O}(\,\text{GeV})$ region, thus evaluating the running of $\alpha_{\rm s}$, in Fig. 1.1.1, one obtains:

$$\Gamma_{gg} \simeq 3 \cdot 10^3 \frac{c_3^2}{(c_2 + \frac{5}{3}c_1)^2} \Gamma_{\gamma\gamma} \tag{1.1.37}$$

The latter equation shows that, in general, the cross-section of gluons clearly dominates over that of photons. Only particular models where the coupling to photons is 100 times enhanced over that to gluons would allow for photon dominance.

1.1.10 Current results on ALPs

The parameter space of ALPs has been explored from below the eV level up to, lately, the $\mathcal{O}(\,\text{GeV})$ one. Depending on ALP mass, different finals states are available, the strongest bound is placed by the dielectron threshold, ALPs below this mass are only allowed to decay to photons. The most stringent accelerator based limits below the dielectron threshold come from light shining through a wall (LSW) [26–28] and helioscope [29] experiments. These exploit the Primakoff effect [30], that describes that ALPs passing near an atom nucleus or through a strong magnetic field would regenerate again into photons. For the MeV region and above, fixed target (beamdump) proton and electron experiments are able to place the tougher bounds [31, 32]. These exploit the clean signal that long-lived ALPs interacting with the nucleus of an absorbing target and regenerating into photons leaves. Up to GeV scale LEP

experiments [33] placed the most stringent limits by searching for rare Z decays. Bounds on ALPs coupling only to photons above the GeV range have been set in ultra-peripheral PbPb collisions by the ATLAS [34] and CMS [35] collaborations. A global picture of the many order of magnitude landscape can be checked in Fig. 1.1.3 (left).

So far the landscape for ALPs only coupling to photons has been checked, however, a rich horizon is present when the coupling to gluons is also considered, as it has been mentioned, the coupling to gluons, in general, dominates over that to photons. Over 10 GeV, the strongest limits have been placed recently by the ATLAS experiment in a dedicated search for isolated photons during the Run 2 of the LHC [36]. Between 50 GeV and 70 GeV the strongest bounds are placed by dijet searches by the CMS experiment [37] and from 70 GeV and up to 100 GeV by recasting diphoton searches by the CMS collaboration [38]. Fig. 1.1.3 shows current bounds on ALPs coupling only to photons (left) and ALPs coupling both to photons and gluons (right). The latter plot shows the gap between the best results obtained on flavour experiments for lower masses and what ATLAS can do with the large amount of integrated luminosity they collected from pp collisions. It was shown that LHCb has the best sensitivity to access the parameter space in this gap [25]. Alternative to these results, it was shown in [39] that a decay constant, f_a of 540 GeV can be exluded by doing a non-resonant search of an ALP produced in gluon-fusion that would produce an enhancement in the inclusive $\gamma\gamma$ spectrum. For lower masses the gluon coupling can be better exploited by looking at meson decays with final states that include ALP signatures as it was shown in [40]. Current prospects and limits can be found in Fig. 1.1.3 (bottom). Future experiments aiming for Long Lived signatures such as CODEX-B [41] or Mathusla [42] are especially promising in this particular mass regime.

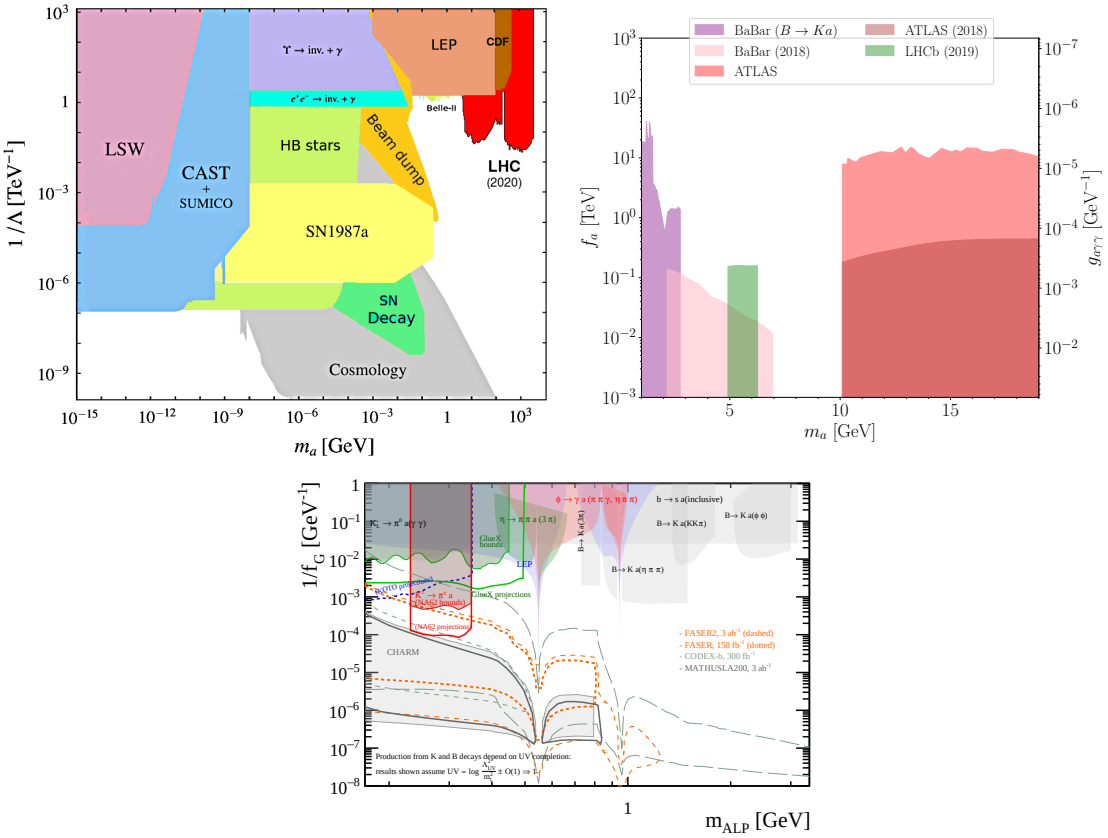

Figure 1.1.3: Left: Limits on the $g_{a\gamma}$ as a function of the ALP mass up to the GeV scale from [43]. Right: Current bounds on the ALP decay constant at the GeV scale. Bottom: Bounds on ALPs coupling to gluons from meson decays, also from [43]. Reproduction of the Figures approved by Springer, the rightful owner.

Chapter 2

Statistical tools for High Energy Physics

Statistical tools are fundamental to bridge theoretical predictions with empirical observation and measurements. In a nutshell, a theory cannot be tested solely based on an observation, a last ingredient has to be added to correctly interpret the limitations of the model predictions and the observable measurements through uncertainties. For a detailed review of the different tools and methods refer to [44, 45].

2.1 Basic concepts

A random process is said to be *random* if its outcome cannot be predicted and only the probabilities of the different possible outcomes can be known. A set of axioms to define the concept of probability was given by Kolmogorov [46] using the tools of set theory. Nowadays two more precise definitions of probability are widely used: the *frequentist* and *Bayesian* approaches.

The frequentist approach assumes that an experiment, E, can be repeated a finite number of times and assigns the probability of a given outcome, A, as the fraction of times that A would happen given that the experiment could be repeated an infinite number of times:

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n} \tag{2.1.1}$$

The Bayesian approach defines the probability as a degree of believe that a certain hypothesis is true. It can be observed that the frequentist approach to some extent can be incorporated in the Bayesian definition, if one interprets the relative frequency of an event as the degree of belief of that outcome. The opposite is in general not true. Not every problem can be evaluated from a frequentist point of view, e.g. the value of a fundamental constant of a theory, which is not associated to a repeatable experiment.

A fundamental piece of the basic theory of probability is the *Bayes theorem*:

$$p(\text{theory}|\text{data}) = \frac{p(\text{data}|\text{theory})p(\text{theory})}{p(\text{data})} \tag{2.1.2}$$

The former equation can be read in the following way, the *posterior probability* of a given theory being true after the outcome of an experiment ($p(\text{theory}|\text{data})$) is equal to the *prior probabilty* that the underlying theory is true ($p(\text{theory})$) multiplied by the *likelihood* that the data follows that theory ($p(\text{data}|\text{theory})$) normalised over the degree of believe assigned to the data outcome ($p(\text{data})$). This definition motivates the Bayesian approach to probabilities as degrees of belief that can be updated when new information is obtained.

2.1.1 Probability density functions

The output of an experiment can be represented by a single continuous variable which will be called x. A natural question to ask is about the probability of observing x within the infinitesimal interval $[x, x+dx]$. The answer to this question is provided by the probability density function (PDF) $f(x)$. Given that x can only take values in the interval $[a,b]$, the probability of finding x within that interval must be 1, which yields the following property:

$$\int_a^b f(x) = 1 \tag{2.1.3}$$

A PDF can in general depend on a larger set of variables, when this happens it is refered the the *joint* PDF: $f(\vec{x})$. Whenever the probability of the variables involved int he joint PDF are independent it is said that they factorize, e.g. $f(x,y) = f_x(x)f_y(y)$.

2.1.2 Maximum likelihood estimators

Consider a random variable, x, that is distributed in a certain range according to the PDF $f(x|\vec{\theta})$. The precise model that controls the behaviour of x is fully described through the set of parameters $\vec{\theta}$. The method of *maximum likelihood* provides a handle to estimating these parameters given a finite sample of the data.

To illustrate the method let us consider that the random variable x has been measured n times: $x_1, ..., x_n$. The joint PDF for all the independent measurements can be expressed as the product of each measurement:

$$\mathscr{L}(\theta) = p(\vec{x}|\vec{\theta}) = \prod_{i=1}^{n} f(x_i|\vec{\theta}) \tag{2.1.4}$$

This is defined as the *likelihood* function, $\mathscr{L}(\theta)$. It is important to point out that the set $x_1, ..., x_n$ is a set of measurements, then those are fixed. The function that emerges is then only dependent on the parameters $\vec{\theta}$. This then sets the definition of Maximum Likelihood (ML) methods: those that maximize the likelihood function. This principle is the basis of two paramount tools in HEP: parameter estimation and hypothesis testing.

2.1.3 Confidence intervals

The confidence interval (CI), in the frequentist approach, is defined as the interval in which the parameter of interest is found with a certain probability $1-\alpha$, usually reported as the confidence level (CL). It is common in HEP results that one of the boundaries is more relevant than the other, for instance when testing the rate or abundancy of a yet to be found process. In this case the relevant information comes from the upper limit that a given experiment can yield on the abundancy or rate of this process.

Given a PDF that models perfectly the random variable x, with a set of known parameters $\vec{\theta}$ the probability of finding it in the range $[a,b]$ is given by:

$$\mathrm{CL} = \int_a^b f(x|\vec{\theta})dx \tag{2.1.5}$$

In general confidence intervals cannot be calculated like this because the parameters $\vec{\theta}$ are not perfectly known, they are estimated. Different techniques can be used to find approximate CI. The certainty that the desired CL is obtained through a given method is called *coverage*.

2.1.4 Hypothesis testing

To obtain confidence intervals, the tool of hypothesis testing is used. Two hypotheses H_0 and H_1 are tested against each other with the outcome of either being compatible or incompatible. For discovery purposes, the null (default) hypothesis is set as background-only processes and is tested against many hypotheses H_1, where signal at varying signal strengths is present. For setting limits however, the situation is reverse: one defines the null hypothesis to be the scenario with background and a given signal strength and tests it against the alternative hypothesis of background-only until one finds a signal strength where the two don't agree anymore.

In order to confront these two hypothesis one should construct a critical region, w, based on the data values $\vec{x}$, such that the probability of finding a measurement $\vec{x}_0$ under the assumption of H_0 is below a given probability, α: the *power* of the test. This means that[1]

$$p(\vec{x} \in w|H_0) = \alpha \tag{2.1.6}$$

It is common place to set a threshold for α, normally 5%, which then narrows down the problem of finding w such that it maximizes the power of the test. The *Neyman-Pearson lemma* [47] states that the critical region should be defined by the likelihood ratio:

$$\lambda(\vec{x}) = \frac{\mathscr{L}(\vec{x}|H_1)}{\mathscr{L}(\vec{x}|H_0)} \tag{2.1.7}$$

[1] A $\leq$ sign is needed in the following equation provided the discrete nature of data does not allow to find a region where the equality holds exactly

such that, for a given α, w is defined such that $c_\alpha > \lambda(\vec{x})$, where c_α is determined by the size of the test. Such a functional of a data measurement $\vec{x}$ like $\lambda(\vec{x})$ is called a *test statistic* and the Neyman-Pearson lemma states that the optimal test-statistic is given by the likelihood-ratio.

Another important tool that helps quantify the level of agreement between a measurement and a given hypothesis (usually the null-hypothesis) is the *p-value* for a given test statistic t, and a measurement, defined as:

$$p-\text{value} = \int_{t_{\text{measured}}}^{+\infty} p(t|H_i)dt \tag{2.1.8}$$

In the context of HEP, the background only hypothesis, H_0 is normally discarded if the p-value is lower than 2.87×10^{-7}, which corresponds to 5σ significance.

In the case where one wants to test the *signal strength* of a given process, for example, a New Physics process one wants to test against the background only hypothesis one needs to find a way to include the rest of the parameters, the *nuisance parameters*. In particular, one actually wants to find a test that yields a result independently of the nuisance parameters. Such a problem can be tackled by *profiling* the parameters, the *profile likelihood ratio*:

$$\lambda(\mu) = \frac{\mathscr{L}(\mu, \hat{\hat{\theta}})}{\mathscr{L}(\hat{\mu}, \hat{\theta})} \tag{2.1.9}$$

In the latter equation μ stands for the scanned parameter, usually some parameter correlated with the amount of signal, the *signal strength*, $\hat{\hat{\theta}}$ are the nuisance parameters best estimates under the assumption of μ; $\hat{\mu}$ and $\hat{\theta}$ are the best estimates from a maximum likelihood fit.

In order to make the process of finding an extreme value easier the following test statistic is defined based on $\lambda(\mu)$:

$$t_\mu = -2\ln\lambda(\mu) \tag{2.1.10}$$

2.1.5 Asymptotic formulae for the likelihood ratio test statistics

The only missing piece of the puzzle is how are the test-statistic values distributed. The most general method to obtain this distribution is to generate pseudo-experiments. The caveat is that usually this method is computationally expensive, such that in the context of a physics analysis is unfeasible. It was shown by Wald [48] and Wilks [49] that, for a large dataset and a single gaussian distributed parameter of interest (in this case μ) the likelihood-ratio based test statistics, $\lambda(\mu)$, asymptotically follows:

$$-2\ln\lambda(\mu) = \frac{(\mu-\hat{\mu})^2}{\sigma^2} + \mathscr{O}(1/\sqrt{N}) \tag{2.1.11}$$

The latter equation when applied to the test statistic defined in Eq. 2.1.10 yields the following relation:

$$f(t_\mu|\mu) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{t_\mu}}e^{-t_\mu/2} \tag{2.1.12}$$

The variance of $\hat{\mu}$, σ can be obtained from the covariance matrix of the parameters $\vec{\theta}$.

2.1.6 The look-elsewhere effect

There are cases where the signal process that is being searched is not, *a priori* located in a particular spot in the parameter space. Normally a scan in a particular region is performed. In case of an excess the fact that signal is, a priori, equally probable anywhere in the range must be taken into account when quoting the significance of the finding. This is known as the look-elsewhere effect [50, 51].

This concept is usually realised as the *trials factor*, which refers to the probability of finding an excess at a given point in the space scanned compared to finding anywhere else in the range.

Solution based on pseudo-experiments

A straight forward method to account for the trials factor is to perform N_{toys} pseudo-experiment based MC simulations of the background only model and apply the search procedure on each one of them, for the whole mass range. From every pseudo-experiment and for each mass hypothesis, the test-statistic of the result is drawn in order to learn its distribution. The global p-value is then obtained by integrating from the measured test-statistic value up to infinity. The main shortcoming of this approach is that the number of pseudo-experiments needed increases for higher local significance values.

Asymptotic formulation

In order to lower the computation load for accounting the trials factor there is an asymptotic formulation of the problem which is described in [52] Let us take the profile likelihood, λ, as defined in Eq. 2.1.9, which it has been shown in Eq. 2.1.12 to follow a χ^2 distribution with one degree of freedom. The interest here lies not in how it depends on the signal strength μ but rather how it changes over the mass range to yield a global p-value. This is, the interest lies in the maximum of $\lambda(m)$ denoted as $\lambda(\hat{m})$. Then, let c be a constant that defines the p-value of the test and the following can be written:

$$P(\lambda(\hat{m}) > c) \leq P(\chi^2_1 > c) + \langle N(c)\rangle \tag{2.1.13}$$

The former equation is a formal way of saying that the excess significance at the mass that maximizes the test statistic is given by the asymptotic behaviour plus mean number of upcrossings over the mass range, $\langle N(c)\rangle$. From [49], and using the asymptotic behaviour of the functions involved, backed up by the fact that $c >> 1$,

$$
\begin{aligned}
N_{\text{trials}} &= \frac{P(q(\hat{\theta})) > c}{P(q(\theta)) > c} \\
&\simeq 1 + \mathscr{N} \frac{\chi^2_2 > c}{\chi^2_1 > c} \\
&\simeq 1 + \mathscr{N} \frac{\sqrt{\pi c}}{2}
\end{aligned}
\tag{2.1.14}
$$

$\mathscr{N}$ represents the effective value of upcrossings of the test statistic λ over a certain threshold. Using the asymptotic behaviour of c one can derive the following relation:

$$
\mathscr{N} = e^{c_0/2} \langle N(c_0) \rangle \tag{2.1.15}
$$

where $\langle N(c_0) \rangle$ is the mean number of times that λ goes over a certain threshold c_0 which is supposed to be chosen as close to 0 as possible while still being able to correctly estimate λ, a common choice is $c_0 = 0.5$. In order to calculate $\langle N(c_0) \rangle$ one can use pseudo-experiments and calculate the number of upcrossings and average over the number of pseudo-experiments used.

The trials factor gives then a handle to "normalise" the significance of parameter estimation given the global shape of the search in a certain mass range.

Chapter 3

The LHCb experiment at the LHC

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [53] is the world's largest and most powerful particle accelerator. It is located at the European Organization for Nuclear Research (CERN), one of the most important research centers in Europe and the world. The LHC is situated in a 26.7 km-long tunnel that previously housed the Large Electron-Positron Collider (LEP) until its discontinuation in 2000. The LHC is a two-ring superconducting proton-proton collider designed to operate at the TeV scale in terms of center-of-mass energy. Protons are accelerated to a maximum energy of 6.8 TeV per beam and collide at a rate of 40 MHz at various interaction points (IPs). Different experiments are positioned at these points to collect data from the particles resulting from these collisions. During special runs, one or both of the LHC beams can be filled with specific ion nuclei, such as lead, instead of protons.

The most significant milestone achieved by the LHC was the discovery of the Higgs boson in 2012 [54, 55]. This accomplishment was made possible by the high instantaneous luminosity of approximately $10^{34}, \mathrm{cm}^{-2}\mathrm{s}^{-1}$ in proton-proton collisions.

The most important experiments placed at the LHC are:

- **ATLAS** (*A Toroidal LHC ApparatuS* [56]): the largest general purpose detector at the LHC. It is designed to directly search for heavy new particles. It is responsible for the Higgs discovery in 2012 [54].
- **CMS** (*Compact Muon Solenoid* [57]): a large general purpose detector tasked to fullfil similar goals as the ATLAS detector. It is also responsible for the discovery of the Higgs boson in 2012 [55].
- **ALICE** (*A Large Ion Collider Experiment* [58]): designed to study Quark Gluon Plasma (QGP) evidence in heavy ion collisions.
- **LHCb** (*Large Hadron Collider beauty experiment* [59]): a forward region detector. Initially designed to study the decays of mesons containing b quarks. A detailed description of the experiment will be provided in the following text.

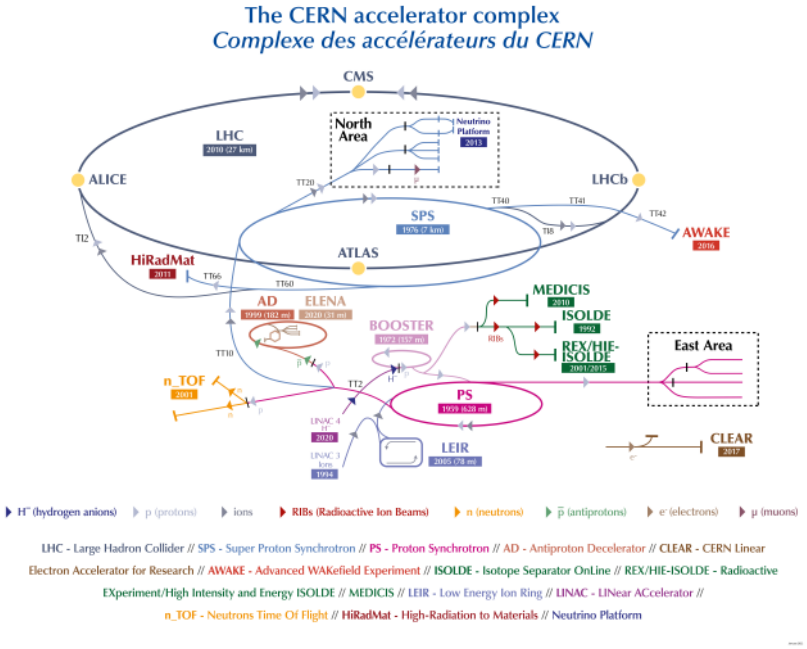

Figure 3.1.1: The LHC complex layout. Appart from the main accelerator, the four main interactions points together with the complementary acceleration facilities are highlighted.

3.2 The LHCb experiment

The LHCb experiment is a forward spectrometer instrumented in the forward region. Designed to detect mesons containing b quarks, it has shown the ability to produce world leading results in many other physics scenarios such as the study of decays of charmed baryons [60], CP violation in the charm sector [61, 62], world leading results in the strange sector [63, 64], spectroscopy of exotic states, like the predicted pentaquarks [65, 66], electroweak precision measurements - such as the W boson mass measurement [67] or shown unique sensitivity to compelling dark sectors: dark photons [68, 69], heavy neutral leptons [70], long lived particles [71] or low mass dimuon resonances in general [72, 73].

LHCb is instrumented in the forward region, covering the angular acceptance ranging from 10 mrad to 300 (250) in the bending (non-bending) plane. This design is chosen to cover region where the most of the hadrons containing b or $\bar{b}$ are produced in high energy pp collisions.

In the following, the detector layout during Run 1 and Run 2 will be described. The relevant detectors that were improved or replaced during the Upgrade and that is running during Run 3 will also be described.

3.2.1 Magnet

The dipole magnet [74] of the LHCb experiment provides a bending power ($\int \vec{B} \cdot d\vec{l}$) of 4 T·m. The aim of the magnet is to bend the trajectory of relativistic charged particles in order to measure the values of their charge and momentum.

The dipole magnet is composed of an iron (Fe) yoke surrounded by two identical aluminum (Al-99.7) coils.

The magnetic field has been measured with a relative precision of $\mathcal{O}(10^{-4})$ in two different campaigns in 2012 and 2014.

3.2.2 Tracking detectors

The tracking detectors play a paramount role in the LHCb operation. They provide the majority of the information related to charged particles passing through the detector: the momentum and the charge. Precise measurements of this information is needed to then correctly identify those in the different charged species, task that is achieved by combining this information with that of other subdetectors.

Vertex Locator

The Vertex Locator (VELO) [75] is a silicon microstrip detector consisting of 21 stations positioned along and perpendicular to the beampipe within a region of approximately 1 meter in length. Each module is composed of two submodules mounted back-to-back. The first submodule contains azimuthal strips, which measure the radial distance (r) from the beam pipe, while the second submodule is composed of radial strips that are sensitive to the angular coordinate (ϕ). The disc-shaped modules have a radius of 42 mm. The strip pitch varies, with a finer pitch at the inner end ranging from 38 μm and linearly increasing to 101.6 μm in the outer region. The VELO material is located as close as 7 mm to the LHC beam, which is smaller than the typical beam width during injection. To prevent radiation damage to sensitive areas of the device, the VELO detector is mounted on a movable structure, allowing it to remain open until the beam reaches stable conditions.

The VELO sensors are separated from the LHC vacuum system by a 0.3 mm thick aluminum shield known as the RF foil. This additional layer introduces additional material that particles from the proton-proton collisions must traverse before being detected, resulting in a degradation of tracking and vertexing observables.

The primary task of the VELO is to reconstruct tracks of charged particles, enabling the construction of primary vertices (PV) and secondary vertices (SV), with a particular emphasis on effectively distinguishing between the two. This capability is essential since the core of the LHCb physics program relies on accurately reconstructing and identifying secondary vertices originating from the decays of b- and c-hadrons.

The upgraded LHCb detector counts on a renovated VELO subdetector [76]. It is designed similarly to the original, with the major change of using silicon pixels instead of microstrips.

The detector consists of 26 L-shaped layers with two modules mounted back to back. Each of these modules is then subdivided in 12 square sensors containing a grid of 256×256 square pixels with a side size of 55 μm. This new design allows the beamline to be at a minnimum distance of 5.1 mm as opposed to the 7 mm distance that was obtained during Run 1 and Run 2. This new design will allow to cope with the higher radiation intensity while also allowing for better secondary vertex discrimination thanks to the closer position to the beamline and, also, fastest readout speed thanks to the pixel technology.

The tracking system

The first stations located upstream the magnet are called Tracker Turicensis or Trigger Tracker (TT), these are followed, downstream the magnet, by the T-stations. This last is composed by the cross-shaped Inner Tracker (IT) [77] that is surrounded by the Outer Tracker (OT) [78]. The TT and IT share the same microstrip silicon fiber technology and that is why together they are referred as the Silicon Tracker (ST). The ST has a strip pitch of 200 μm that implies a single hit resolution of 50 μm. This particular arrangement provides enough information to reconstruct a charged particle hit in the three dimensions. The TT is 150 cm wide and 130 cm high covering the full acceptance of the experiment. The IT covers a 120 cm wide and 40 cm high cross-shaped region in the center of the T-stations. The OT is a drift-time detector using as counting gas a mixture of Argon (70%) and CO2 (30%), with which a drift time below 50 ns and a drift-coordinate resolution of 200 μm is achieved. The OT in each station covers an active area of 6×5 m^2. Each T-station is segmented in four layers: the first and last layers are instrumented with vertical strips while the strips corresponding to the two sitting in the middle are tilted 5° and -5° alternatively, in a $x-u-v-x$ fashion.

For the Run 3, starting in 2022, the tracking system was removed and a new one installed. The Upstream Tracker (UT) [79] is the new silicon strip detector situated upstream the magnet and designed to replace the TT. The four planes are arranged in a $x-u-v-x$ way, like the previous T-stations. The four planes of silicon strips are made of up 16 or 18 vertical staves which are further subdivided in 14 98.98×98.98 mm^2 sensors containing 512 strips with a 190 μm pitch. The Scintillating Fibre Tracker (SciFi), commonly known as the forward tracker is located downstream the beamline. The three tracking stations are subdivided in four layers arranged in the $x-u-v-x$ fashion. The detector layers are formed by 2.4 m long and with a 250 μm diameter scintillating fibers.

3.2.3 Calorimeters

The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is constructed using the “shashlik” technology [80]. Each individual module consists of lead absorber layers followed by scintillator tiles, with readout accomplished through fibers connected to photomultipliers.

The ECAL is positioned 12.5 meters away from the interaction point (IP). The detector itself forms a wall with dimensions of 7.8×6.3 square meters, featuring a hole in the middle with a $\theta < 30$ milliradian aperture to accommodate the beam pipe. The subdetector is composed of 3312 square-section submodules. The ECAL is divided into three regions with equally sized modules but varying granularities. Each module in the Inner/Middle/Outer regions is further subdivided into 9/4/1 readout cells, with sizes of 4.04/6.06/12.12 square centimeters each. In total, there are 6016 individual cells. The multiplicity can vary significantly depending on the azimuthal angle, which influenced the choice of granularity. Consequently, the gain of the calorimeter cells is set to be uniform in $E_{\mathrm{T}} = E \sin\theta$ [81]. Each module is composed of 66 layers, consisting of 2 millimeters thick lead, 120 micrometers reflecting paper, and 4 millimeters scintillating fibers, resulting in a total radiation length of 25 X_0. The Molière radius (defined in Sec. 3.3.1) of the ECAL is measured to be 3.5 centimeters.

The nominal resolution of the ECAL, determined through electron reconstruction during the CERN test beam in 2007, is reported as [82]:

$$\frac{\sigma(E)}{E} = \frac{(9.0 \pm 0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E_{\mathrm{T}}} \tag{3.2.1}$$

Due to the limited number of readout channels, a compromise must be made between energy resolution and the maximum measurable value. During Run 2, the configuration of the subdetector was adjusted to measure E_{T} in the range of $[0, 12]$ GeV.

Upstream of the ECAL, two additional calorimeter systems are situated: the Preshower Detector (PS) and the Scintillator Pad Detector (SPD). The primary role of the PS is to distinguish between charged and neutral particles. This information is rapidly read out and utilized to differentiate electrons from photons without relying on tracking information. Subsequently, a 15 millimeter lead sheet initiates an electromagnetic shower for electrons and photons, whereas charged hadrons are less affected due to their longer interaction lengths. The SPD then detects electromagnetic showers using a similar principle as the ECAL. This subdetector is primarily used to measure event multiplicities.

The hadronic calorimeter

The hadronic calorimeter (HCAL) is a sampling device using scintillating tiles as active material and steel as absorber. The most particular feature of this subdetector is that the scintillating tiles run parallel to the beam axis. This choice is made to enhance light collecting efficiency. The readout cells are defined by grouping together different sets of fibers onto one photomultiplier tube. Depending on the size of the square cells two regions are defined - the inner region has cells of size 131.3 mm whereas in the outer region they are of 232.6 mm. The inner (outer) section is bounded by $x \in [-2101, +2101]([-4202, 4202])$mm and $y \in [-1838, +1838]([-3414, 3414])$mm.

The bulk of the HCAL is positioned at a distance $z = 13.33$ m from the interaction point and has overall dimensions of 8.4 m in height, 6.8m in width and 1.65 m in depth.

The HCAL baseline resolution was measured to be, using test beams with pions [83]:

$$\frac{\sigma(E)}{E} = \frac{(67 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \tag{3.2.2}$$

The limited space left dictates that the HCAL equivalent depth is of $5.6\lambda_{\text{int}}$. This fact is also a limiting factor in how good resolution it can be achieved though it was showed this not affect significantly the trigger efficiencies, its main purpose.

3.2.4 PID detectors

RICH system

The RICH detectors [84] in LHCb exploit the phenomenon of Cherenkov radiation. This occurs when a charged particle traverses a dielectric medium with a refractive index (n) different from that of free space, at a velocity (v_p) higher than the speed of light in that medium (c). The emitted photons form a cone with an angle θ determined by the values of n, c, and v_p:

$$\cos\theta = \frac{c}{v_p n} \tag{3.2.3}$$

This information, combined with the momentum measurement of a track, can be used to make hypotheses about the mass of the detected particle.

LHCb employs two RICH detectors. The upstream detector, RICH1, utilizes C_4F_{10} gas as the radiator medium and covers the momentum range of [1, 60] GeV. The downstream detector, RICH2, employs CF_4 as the radiator medium and is designed to cover a higher momentum range of [15, 100] GeV.

The radiated light is focused using an array of mirrors, which directs the photons toward the photon detectors. These detectors are shielded from the dipole magnet, which can interfere with the relativistic charged particles.

MUON system

The MUON system [85] is the last detector downstream the beampipe. It is composed of 5 stations (M1-M5) that use multi-wire proportional chambers (MWPC) to detect charged particles. The first station (M1) is located upstream the calorimeter system while the remaining four (M2-M5) are located downstream of it. These 4 subdetectors are interleaved with 80 cm iron absorbers. This setup maximizes the probability that the flux of particles that reach this far will be mostly composed by muons, due to its minimum ionising properties. The angular acceptance is designed projectively to match that of the tracking system in the 20 (16) mrad to 306 (258) mrad in the bending (non-bending) plane. The thick absorbers put a constraint on the minimum momentum that muons need to reach M5: ~ 10 GeV.

For the upgrade meant for Run 3 minimal changes are made to the MUON system. The most relevant being the removal of M1, situated before the calorimeter system, since it was intended to measure momentum of muons at the L0, system which run without a dedicated tracking reconstruction. This change also removes unnecessary material upstream the calorimeters.

3.2.5 The trigger system during Run 2

The event rate that LHCb processes in real time during the Run 2 data taking period is of 40 MHz but most of the events lack any interesting signal. The progamme extends to strange physics or the study of hadrons containing *s* quarks, EW measurements or even searches for signals outside the Standard Model; although, the bulk of the LHCb physics programme is dedicated to heavy flavour physics: the study of hadrons containing *b* or *c* quarks. This is why the trigger system that filters the data and lowers the input rate down to 3 kHz before saving it to disk is optimised for this type of signatures [86], though enabling enough flexibility to enlarge the scope of the experiment. The trigger system is divided in three different stages or levels. In each increasing level the reconstruction of the event is more detailed, this means signal specific selections are more suited for later stages of trigger selection. An outline of the Run 2 trigger can be found in Fig. 4.0.1 (left).

The first level of the trigger, L0, is a hardware system that partially reconstructs the event based on information that mainly comes from the calorimeters and the MUON system. The information is barely processed in order to yield decisions at the 40 MHz input rate. This stage of the trigger mainly applies selection per-particle or on the general event multiplicity. This last point is mainly available by measuring the occupancy of the SPD (nSPDHits). The PS also helps distinguish between photons and electrons at this stage. The fully readout events that successfully pass the L0 requirements are shipped to a farm of CPUs that run the High Level Trigger (HLT) which is divided in two parts: HLT1 and HLT2. The HLT1 applies track reconstruction to those charged particles that go through the full detector, i.e. long tracks. These tracks are then used to reconstruct the primary vertex (PV)[1] and the secondary vertices left by particle decays. The output rate after the HLT1 is shrank down to $\mathcal{O}(100)$ kHz. The HLT2 is the last trigger step. The application that steers it picks up events that were previously buffered in disks. It not only serves the purpose of applying the last filter to the input rate but also calibrates and aligns the detector for the next run. At this stage the reconstruction of the event matches that obtained in offline conditions, PID information from the RICH detectors is also available.

The information of which part of the detector fired a certain decision is kept, this allows to select events based, basically, on whether our signal of interest fired the trigger or, rather, it was something else in the event that did it. A event is said to be Triggered On Signal (TOS) if

[1] The mean number of PVs during the Run 2 was set to 1.6. This is why normally one only considers one PV per event.

the signal of interest satisfies certain trigger requirements. Whenever an event is kept because a different signal triggered the event it is said to be Triggered Independent of Signal (TIS).

The trigger configuration while it is meant to be stable during the data taking period it is subject to change. Each run is associated with a hexadecimal value: the Trigger Configuration Key (TCK). Each TCK identifies the trigger: what lines were running and with which configuration settings.

After the trigger, data is deferred for offline scrutiny. This last step is usually performed centrally in different campaigns, known in the LHCb collaboration as *Stripping* campaigns. For the Run 2 a new technique was introduced to implement exclusive selections already at the trigger level: *Turbo* [87]. This is especially indicated for those physics analyses with very high output. For calibration purposes, a small fraction of the trigger rate is dedicated to the *TurCal* stream, which is processed in the same way as Turbo but is dedicated for calibration purposes.

3.2.6 The LHCb simulation

Simulation samples play a crucial role in High Energy Physics. They are paramount in order to calibrate selections of yet-to-discover signals and even help fine-tune different reconstruction algorithms. LHCb uses a framework that ties together different pieces of software that produce the simulated data samples. At the generation stage Pythia 8 [88] is chosen to simulate pp collisions in order to obtain the different Standard Model hadrons with the correct kinematics and within the correct underlying event context. For this, work, Madgraph5@NLO [89] was also used within the framework to generate ALP candidates. The events are nonetheless processed with Pythia in order to transform the basic event information into the different hadrons. EvtGen [90] is also extensively used to force the decay of certain particles to a given final state. This is particularly interesting for heavily suppressed decays. The material interactions with the different layers of the detector is simulated with Geant4 [91]. The detector response is further digitised and passed on to a software that simulates the L0 response and then to the same software that applies the trigger decisions on data. By following this procedure real data and simulated samples are kept as similar as possible structure wise.

During Run 2, the widely used simulation framework version is called Sim09. This was replaced in 2022 by Sim10. The major difference is the migration from Geant4 v9 to Geant4 v10.

3.3 Photon reconstruction and identification at LHCb

Photons are reconstructed using the ECAL subdetector, described in Sec. 3.2.3. Since the ECAL detects electromagnetic showers that are produced just before, tracks are needed to distinguish neutral from charged clusters. The basic idea is that the hypothesis to match a neutral cluster to a track is disfavored while the opposite happens with a charged cluster.

Using information from the shape of the cluster in the ECAL, in combination with the tracking system (Sec. 3.2.2) and the information provided by the PRS and SPD calorimeter systems one can distinguish between different types of particle hypothesis. Tracks compatible with circularly shaped clusters hint towards an electron passing through the detector while the opposite is a hint of either a γ or a collimated $\pi^0 \to \gamma\gamma$ decay hitting the ECAL.

3.3.1 Electromagnetic showers

The measurement of the energy of photons (and also electrons) is based on how these particles interact with matter. The main interaction process for electrons for collider energy regimes is *bremsstrahlung*, which produces secondary photons. The main interaction process for photons is, conversely, pair production which on turn produces more electrons that will continue the process when any of these enter a dense material, like the LHCb calorimeters. This collective process is known as electromagnetic shower. It is useful at this point to define the interaction length X_0 which is the longitudinal distance that a shower has when its energy goes from E_0 to E_0/e. The shower is expected to die off when the system reaches the critical energy, E_c. At this point the main interaction regime for electrons is ionisation. It is convenient to define length and energy variables relative to these, that define the material scale:

$$t = \frac{x}{X_0}; \quad y = \frac{E}{E_c} \tag{3.3.1}$$

The evolution of the shower can be parametrised with two material dependent parameters: a and b:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} \tag{3.3.2}$$

This equation defines a maximum penetration depth which is maximized approximately for $t_{\max} \simeq \ln \frac{E}{E_c} \pm 0.5$.

On the transverse plane the shower can be parameterised with so called *Molière Radius* [92]:

$$R_M \simeq X_0 \frac{21\,\text{MeV}}{E_c}. \tag{3.3.3}$$

It is shown experimentally that almost 99% of the energy is contained within 3.5 R_M [93].

3.3.2 Photon reconstruction during Run 2

Energy deposits on the ECAL are gathered into clusters around each local maximum. The clusterisation is obtained thanks to a Cellular Automation algorithm [94]. To avoid double counting of energy, if a cell is shared by many clusters its energy is redistributed in proportionality to the total cluster energy.

The first rough parameterisation of this cluster is performed by calculating the total energy and the position of the barycenter, defined as follows:

$$\varepsilon = \sum_i \varepsilon_i, \quad x_b = \frac{1}{\varepsilon}\sum_i x_i\varepsilon_i, \quad y_b = \frac{1}{\varepsilon}\sum_i y_i\varepsilon_i \tag{3.3.4}$$

The transverse dispersion matrix parameterises the 2D shape of the cluster:

$$\mathscr{S}_{xx} = \frac{1}{\varepsilon}\sum_i \varepsilon_i(x_i - x_b)^2, \quad \mathscr{S}_{yy} = \frac{1}{\varepsilon}\sum_i \varepsilon_i(y_i - y_b)^2, \quad \mathscr{S}_{xy} = \mathscr{S}_{yx} = \frac{1}{\varepsilon}\sum_i \varepsilon_i(x_i - x_b)(y_i - y_b) \tag{3.3.5}$$

Clusters are then classified as neutral by checking that these do not satisfy a χ^2 matching criteria with the reconstructed tracks of the event:

$$\chi^2(\vec{p}) = (\vec{r}_{\text{track}} - \vec{r}_{\text{cluster}})^T (C_{\text{track}} + S_{\text{cluster}})^{-1} (\vec{r}_{\text{track}} - \vec{r}_{\text{cluster}}) \tag{3.3.6}$$

3.3.3 Energy and cluster position corrections

The first hypothesis to the photon energy is calculated by adding all the individual depositions in the 3x3 cluster. This first approximation has to be corrected for leakages in the ECAL and also leakages in the PRS, the corrected energy is parametrised as follows:

$$E_c = \alpha E_{3x3} + \beta E_{\text{PRS}} \tag{3.3.7}$$

The parameter α encodes information about leakages because of mismodelling of the energy-weighted barycenter which results in the shower not being well contained in the 3x3 cluster. It also encodes the energy that is lost in the absorber and longitudinal leakages. The β parameter encodes the so called *passiv-to-active* energy factor and it accounts for the energy that is used to start that shower and that might not be completely captured by the calorimeter active material.

As a function of the corrected energy, E_c, a correction on the depth of the barycenter is also used:

$$z_c = z_{\text{ECAL}} + \alpha \ln E_c + \beta E_{\text{PRS}} \tag{3.3.8}$$

The transversal profile of the shower is found to be gaussian-shaped. This modelling motivates a different parameterisation of the barycenter position in the XY plane. This new definition takes into account the energy weighted barycenter vector position (x_b, y_b), a tuned constant, b, that encodes the gaussian shape of the cluster and the half-size of the cluster, Δ.

$$(x_c, y_c) = \mathscr{S}_0[(x_b, y_b), b] = b \,\text{asinh}\left[\frac{(x_b, y_b)}{\Delta}\cosh\frac{\Delta}{b}\right] \tag{3.3.9}$$

3.4 Photon identification

In Sec. 3.3.2 it was shown that neutral clusters are classified as such by vetoing those that satisfy a χ^2 matching criteria. This classification is not free from mis-classification and thus further information from the detector is needed to correctly classify these clusters as real photon candidates.

The main background categories are:

- $e^{\pm}$: cases where the electron tracks are poorly reconstructed can lead into classifying the EM cluster as a neutral when it was originated from these charged leptons.
- Hadronic energy depositions. Light hadrons are very common species in pp collisions. Some of them can interact in the ECAL and behave as if they were misidentified γ.
- Merged $\pi^0 \to \gamma\gamma$ decays. The neutral pion decays predominantly to photons. Boosted hadrons can decay in a pair of photons that propagate very collimated to each another producing clusters in the ECAL that greatly overlap.

In order to mitigate these sources of background a set of multivariate classifiers were developed [95, 96]. The first two categories of background are discriminated thanks to the following variables:

- The ratio of the energy of the seed cell (i.e. the one that sits in the middle of the 3x3 cluster) with respect to the total energy of the ECAL cluster: $(E_{\text{seed}}/E_{3x3})_{\text{ECAL}}$.
- The ratio of the energy of the PS cluster in from of the seed ECAL cell with respect to the total amount of energy deposited in front of the 3x3 ECAL cluster: $(E_1/E_{3x3})_{\text{PS}}$.
- The relative energy leakage to the ECAL in the projective area in the HCAL: $E_{\text{ECAL}}/E_{\text{HCAL}}$. This variable will also be refered as HCAL2ECAL.
- The total energy of the 2x2 array of PS cells that sits in from of the 3x3 ECAL cluster: PSE4Max
- The second order momentum of the cluster (as defined in Eq. 3.3.5): $\mathscr{S}_{xx}$, $\mathscr{S}_{yy}$ and $\mathscr{S}_{xy}$.
- The χ^2 matching criteria to a track: χ^2TrackMatch

This set of variables is chosen because they serve as measurement of how likely an ECAL cluster is to have been originated by a shower produced by a photon or an electron. These show different phenomenology with respect to those generated by charged hadrons. This results in different energy deposits in the PS and different cluster transversal size. In order to veto charged clusters the χ^2TrackMatch profiled is paramount.

In order to achieve a γ/π^0 separation variables that encode the shape of the cluster in the ECAL and the PS are the most important ones. The amount of energy deposited in both the ECAL and the PS together with the PS multiplicities is also used.

These methods are implemented through a multivariate classifier that is trained with signal and background species containing radiative B meson decay control channels. A detailed review of the tools that implement this selections can be found in [95].

The tools that implement the signal/background separation are:

- Electron-photon separation: `IsNotE`
- Discrimination against hadronic energy deposition: `IsNotH`
- γ/π^0 separation: `IsPhoton`

Chapter 4

An alternative muon reconstruction algorithm for the GPU HLT1

The Upgraded LHCb detector takes data without a hardware level trigger. Instead, a *Graphics Processing Unit* (GPU) implementation of the HLT1, capable of processing the 30 MHz full event rate of the LHCb experiment, was chosen. This is the first time that a trigger application is fully developed using GPUs for a general purpose HEP experiment. Fig. 4.0.1 shows the main differences between the trigger that was used during the Run 2 data taking period of LHCb and the trigger used with the upgraded detector that began taking data at the end of 2021. This *hardware-triggerless* system allows for a more flexible selection at the readout level at the same time that it allows for a reduction of the differences between HLT1 and HLT2.

In this chapter, the Allen [97] project is presented and a track reconstruction sequence that uses the VELO and MUON detectors is described.

4.1 Basic GPU concepts

GPUs are devices designed to display images on a screen, this task can be narrowed down to calculating the color of multiple pixels in parallel. GPU devices are connected to work together with a Central Processing Unit (CPU) with which they exchange data. Historically, the first graphic processors did not count on parallelisation capabilities but simply consisted a pipeline of operations that would be combined with a buffer in order to show graphics on display. The first modern GPUs that incorporated programmable cores, using a technology comparable to the state of the art devices, were those of the NVIDIA GeForce 3 series, in 2001. These devices consisted of 16 programmable cores. Although the base technology has evolved and the comparison only holds approximately as modern GPUs use thousands of cores to process pixels on a screen.

The programmable cores opened a gate to expand the usability of these devices. The key aspect that the programmable cores incorporated is that the arithmetic operations performed on the input information (e.g. formerly color) was completely controlled by the programmer. This meant that what was interpreted as color could actually be *any* data. This data could then

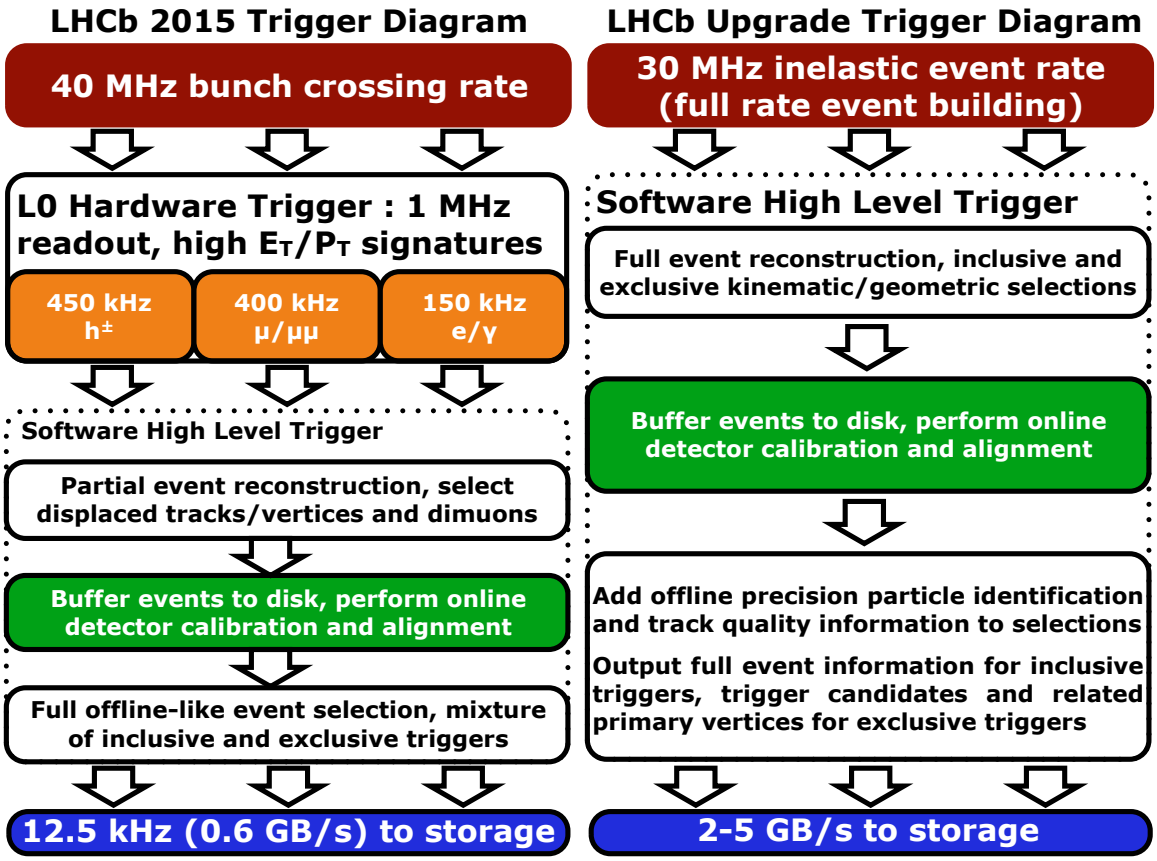

Figure 4.0.1: Comparison between the Run 2 trigger configuration (left) and the configuration of the Upgraded LHCb detector for Run 3.

be handed back to the CPU and the calculations interpreted. Graphic handling *Application Programming Interfaces* (APIs) such as DirectX or OpenGL were used in the early days of this technology, which came with the limitations of using tools that were not designed for the desired purpose. It would be with the development of CUDA and OpenCL APIs, in 2006 and 2009 respectively, that the computing power of these devices could be more readily exploited.

The huge number of cores that modern GPUs provided are structured in a scalable way. Each of the individual processing units is called a Streaming Multiprocessor (SM). Each SM has many Streaming Processors (SPs) together with buffers of shared memory within the SM. The GPU has a global memory buffer accesible from all SMs. Different processors within the SM can launch multiple threads that execute independently and run concurrently. This last aspect is a key factor: concurrent threads can execute in parallel and synchronize with a single SM instruction. The processor architechture is described as *single-instruction, multiple-thread* (SIMT). The multithreaded instructions scheduled by the SMs are executed in groups of 32 threads called *warps*. As stated before, threads within a warp can run concurrently but are

free to branch if the code they execute diverges at a certain point. If branching happens the execution of diverging threads is effectively serialised. The peak efficiency is found when the divergences are minimised and concurrency is thus maximised. The SIMT architechture can be regarded as similar to the *single-instruction, multiple-data* (SIMD) architecture that is widely used in multicore CPUs. The main difference is that SIMD focuses on applying the same instruction to different data locations, while SIMT applies a single instruction to different independent threads. While SIMD focuses on applying an operation to a vectorised data structure, SIMT launches independent threads and handles their branching behaviour effectively.

Memory plays a significant role in the efficiency of GPUs. Similarly to other processor architectures, a cache hierarchy optimizes memory accesses according to spatial and temporal locality principles. In the case of NVIDIA GPUs, the first cache layer is local to each SM and part of it is configurable as indirectionable memory referred to as *shared memory*. While all threads in a GPU, from different SMs, can access the bigger sized *global memory*, it is faster for threads within the same SM make use of the shared memory buffer. Memory efficiency is a primary concern when developing efficient GPU algorithms. The idiosyncratic memory hierarchy combined with the low memory available (a total of $\sim$ 20 GB in modern GPUs with $\sim$ 5000 cores) calls for algorithms that make a clever use of the device resources and a specific data-structure to map onto its characteristics. Memory structures are a relevant consideration to achieve good efficiency in GPU algorithms, where access patterns should influence the data layout to maximise coalesced accesses. The Structure of Arrays (SoA) data structure is typically preferred for data-parallel workloads such as the ones in GPUs, however it is best to custom-tailor the data structure according to the use case through testing and profiling.

4.2 The Allen project

The effort to develop a full implementation of the HLT1 trigger on GPUs, resulted in the *Allen project* [97, 98], named after Frances E. Allen [99]. The framework implements all the reconstruction algorithms and can be compiled using the CUDA [100] toolchain to target NVIDIA GPUs or the HIP [101] toolchain to target AMD devices. It is also possible to compile and execute all of the algorithms on CPUs, only requiring a `C++17` installation. The software can run as a standalone application and also as part of the LHCb software stack.

The Allen design pursues the goal of being a self-contained application that takes in the raw data and outputs the decision whether to store or discard events while maximising the algorithm performance. LHCb events are small enough ($\sim$ 100 kB) so that typically $\sim$ 10 bunches of 1000 events each are being processed simultaneously on a given GPU. Each of those bunches is then handled by a single *CPU thread* that is in turn associated to a *GPU stream*. Each of these bunches is handled asychronously in order to hide the latency incurred by copying data to the GPU. The memory needed to run the sequence per stream is allocated before running it and a dedicated memory manager is used to reserve and release from the

GPU specific buffers in between kernel execution. Algorithm dependencies are handled with a scheduler that produces a static schedule upon launching the application. A static schedule allows to identify when must buffers be reserved and freed, which is automatically handled by the framework without user interaction. This precludes requiring any other dynamic memory allocations, which are otherwise avoided as a best practice throughout the framework. Users must specify the sizes of the buffers, following a philosophy that could be described as "count first, write later". In short, before the launching of an algorithm the memory buffers size (or at least an informed guess of it) is required upfront. This usually translates on interleaving algorithms that are tasked with counting the relevant objects. These counters are then processed by a *prefix sum* that allows to reserve sufficient memory for the SoA structure that hosts the data in a contiguous way. The prefix sum algorithm makes an opportunistic use of the CPU by taking these counters back to it, executing the prefix sum, and then also shipping back the SoA indices to the GPU device. The combination of this simple strategy allows to use as much as possible the resources of the GPU by maximising the concurrent reading and writing of data.

4.3 The HLT1 trigger requirements

The goal of the HLT1 is to reduce the the rate of events to a level where data can be buffered in disk for real-time alignment, calibration and further filtering in HLT2. This process should be performed in an efficient way to comply with the LHCb physics programme, which has a rough estimated output rate of a few kHz.. This means reducing the event rate by a factor 120, starting from the input rate of 30 MHz. The detailed requirements of the HLT1 can be found in [98]. Those can be narrowed down to:

- Reconstruction of tracks:

  Upstream and Long tracks should be recontructed without severe p_{T} and p cuts. While for beauty and charm programs a $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ is adequate, for strange physics it is very important not to apply any p_{T} cut but to only apply a $p > 3$ GeV cut.

  Muon leptons have to be identified.

- Reconstruction of vertices:

  All of the tracks in the VELO detector must be reconstructed in order to reconstruct all of the primary vertices and secondary vertices of long lived particles.

The baseline reconstruction sequence is summarised in Fig. 4.3.1. Downstream tracking, reconstruction of neutral particles and electrons with information from the ECAL and particle identification through the RICH are not required though some particular areas of the LHCb physics programme could greatly benefit from them.

The Allen reconstruction data model is flexible enough to host alternative reconstruction techniques that might fall outside of the baseline HLT1 requirements. In the following section an alternative muon reconstruction algorithm within the Allen framework will be described.

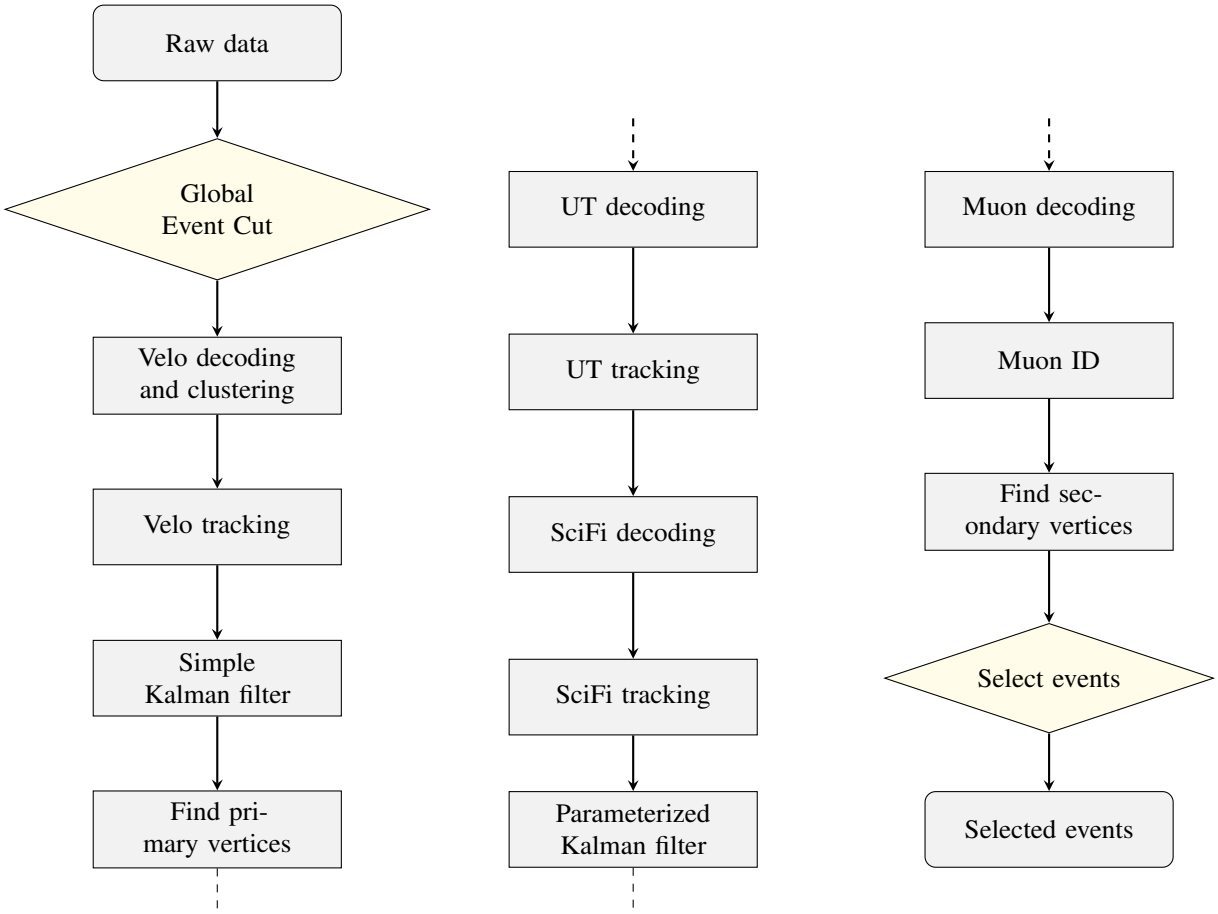

Figure 4.3.1: Full HLT1 sequence implemented in CUDA to run on GPUs.
Figure adapted from [97].

4.4 VELO-MUON tracking

In this work a VELO-MUON reconstruction sequence was designed. The first step consists on reconstructing standalone MUON tracks, consisting only from hits left at the MUON stations (Sec. 3.2.4). Then these MUON tracks are associated with VELO tracks to build VELO-MUON track candidates. These tracks describe the path of charged particles that go through the whole detector, even across all of the MUON chambers, this is why they are given a strong hypothesis of corresponding to $\mu^{\pm}$ leptons. These tracks can be then associated to form vertices. This is showcased with MC samples by showing the reconstruction of displaced signals, like $B_s^0 \to \mu^+\mu^-$ and prompt ones, showcased with the example of $J/\psi \to \mu^+\mu^-$

4.4.1 Standalone MUON tracks

The first part of the reconstruction chain is to obtain standalone muon stubs that are reconstructed without seeding from any other subdetector.

The algorithm looks for hits in the MUON stations in the opposite direction to that of charged particles in the LHCb detector, meaning that hits are searched for in the outermost stations first. This strategy is chosen in order to minimise the timing requirements of the algorithm. Because of the thick iron absorber between the different stations of the MUON chambers, outer stations will show lower multiplicities. The maximum number of candidates then falls back to the number of hits in the outermost station considered.

The algorithm does the task of searching for the hits that build up the track, fitting the two spatial projections of the track and then rejecting low quality tracks and those that share hits, clone tracks.

First, tracks are built by searching for hits in all four stations: M2, M3, M4 and M5. In the first iteration a hit from the outer station is picked and a straight line connecting that hit with the origin of the LHCb coordinate system is drawn[1]. A first search window is open searching for hits within it, the closest is picked, the third and fourth hits are searched by extrapolating linearly from the last two hits. Search window sizes can be found in Tab. 4.4.1. Then the two projections of the three-dimensional lines are fitted to straight lines. The result of this fit serves two purposes: finding the t_x and t_y slopes and measuring the quality of the track. Tracks that share hits are compared by means of the fit χ^2 result and the lower quality one is discarded. An example of a MUON track candidate is demonstrated in Fig. 4.4.1.

Region	$x_{\max}$ [mm]	$y_{\max}$ [mm]
R1	400	60
R2	400	120
R3	400	240
R4	400	480

Region	$x_{\max}$ [mm]	$y_{\max}$ [mm]
R1	100	60
R2	200	120
R3	300	240
R4	400	480

Region	$x_{\max}$ [mm]	$y_{\max}$ [mm]
R1	100	60
R2	200	120
R3	300	240
R4	400	480

Table 4.4.1: Left (right) [right] M4 (M3) [M2] search window dimensions in the x, y plane per MUON station region.

[1]The origin of the coordinate system is centered in the region where pp collisions are expected to happen

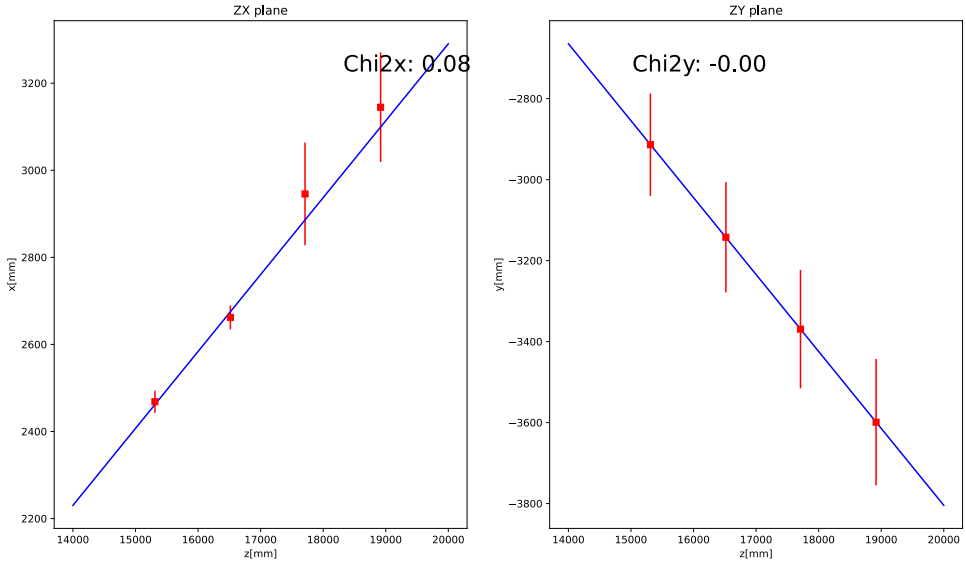

Figure 4.4.1: Muon track hits and linear fits to both XZ and YZ projections using MC simulations.

4.4.2 VELO-MUON tracks

Muons that are produced near the interaction point within the LHCb acceptance with a total momentum $p \gtrsim 10\,\text{GeV}$ will leave enough information in the detector to be reconstructed by the VELO and by the four MUON stations. Putting together these two segments would constitute the reconstruction of a VELO-MUON track that would have the power to reconstruct vertices very precisely thanks to the VELO with a decent momentum resolution thanks to the long lever arm.

VELO-MUON tracks open an alternative to SciFi tracks that would enable alternative muon reconstruction at the HLT1 trigger level. These are particularly useful to measure tracking efficiencies already at the trigger level through the *tag-and-probe* method while also being particularly useful for high-multiplicity events, where the reconstruction of UT and SciFi tracks becomes unfeasible while the MUON stations should be much cleaner due to the material bulk that sits in the pathway of the tracks. The VELO reconstruction should also be accessible thanks to the nature of the VELO as a pixel detector and the fact that tracks are well described by straight lines inside it, as well as the dedicated hardware clustering [102]. This alternative reconstruction also shows the flexibility that the Allen framework satisfies by allowing, at readout level, such a special topology to be implemented.

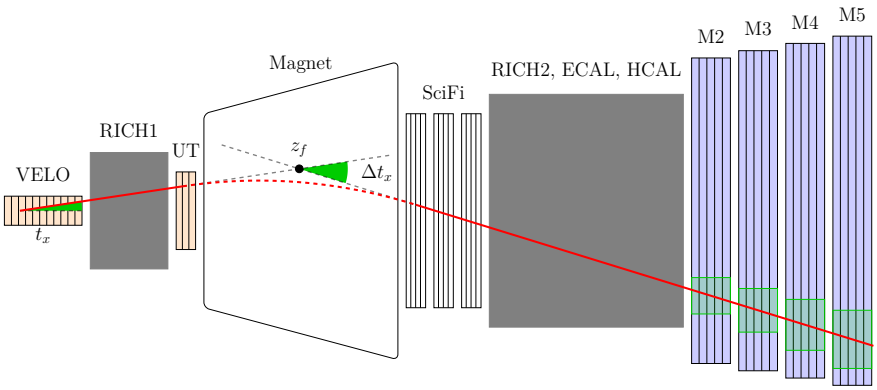

Figure 4.4.2: Trajectory of a charged particle traversing the LHCb detector. The change in slope, Δt_x that is then used to calculate the momentum of the charged particle is highlighted. Original figure from [103].

Matching VELO and MUON segments

The next step towards reconstructing VELO-MUON tracks is putting together both segments. The basics of this matching lies on two hypotheseses: the magnet deflection of the charged particle only happens in the ZX plane, this means that t_y should remain constant along the path of the track: $t_{y,\text{VELO}} \simeq t_{y,\text{MUON}}$; the deflection on the bending plane can be accurately parameterised by a second degree polynomial on $t_{x,\text{VELO}}$. In this work the parameters obtained in [103] are used.

$$z_{\text{deflection}} = a - b\, t^2_{x,\text{VELO}} \tag{4.4.1}$$

It is useful to define the following distance quantities,Δx, Δy by extrapolating the VELO and MUON segments to $z_{\text{deflection}}$:

$$\begin{cases} x_{\text{VELO,extrapol}} = x_{\text{VELO}} + (z_{\text{deflection}} - z_{\text{VELO}}) \cdot t_{x,\text{VELO}} \\ x_{\text{MUON,extrapol}} = x_{\text{MUON}} + (z_{\text{deflection}} - z_{\text{MUON}}) \cdot t_{x,\text{MUON}} \end{cases} \tag{4.4.2}$$

The matching point for the ZY plane is chosen arbitrarily near to the MUON stations, this is because $t_{y,\text{VELO}}$ has a better resolution than $t_{y,\text{MUON}}$ and the extrapolation from the VELO is then more accurate. The similar extrapolation is performed as in Eq. 4.4.2 but instead of extrapolating to $z_{\text{deflection}}$ it is instead performed to $z = 17500$ mm. This point is chosen by scanning as it shows the best matching performance.

$$\begin{cases} \Delta x = x_{\text{VELO,extrapol}} - x_{\text{MUON,extrapol}} \\ \Delta y = y_{\text{VELO,extrapol}} - y_{\text{MUON,extrapol}} \end{cases} \tag{4.4.3}$$

The following quantity is used to match both segments:

$$q^2 = \frac{(\Delta x)^2}{(10\text{ mm})^2} + \frac{(\Delta y)^2}{(10\text{ mm})^2} + (\Delta t_y)^2 \tag{4.4.4}$$

The reconstruction algorithm puts together two segments and measures the agreement by evaluating q^2 as it is described in Eq. C.4.1, it exits early if the evaluation of any of the four variables entering the definition of q^2 are outside physical bounds. The best match is then saved, and the muon slope is refitted by adding to the ZX projection the point $(x_{\text{VELO,extrapol}}, z_{\text{deflection}})$. This last point helps to improve the determination of $t_{x,\text{MUON}}$.

The momentum is then determined by mapping the change in slope onto momentum by following the next relation:

$$\frac{q}{p} = \frac{1}{\int |d\vec{l} \times \vec{B}|_x} \left(\frac{t_{x,\text{MUON}}}{\sqrt{1 + t_{x,\text{MUON}}^2 + t_{y,\text{MUON}}^2}} - \frac{t_{x,\text{VELO}}}{\sqrt{1 + t_{x,\text{VELO}}^2 + t_{y,\text{VELO}}^2}} \right) \tag{4.4.5}$$

The matching algorithm provides all the slope parameters at pattern recognition, the only missing ingredient is the determination of the line integral of the B field: $\int |d\vec{l} \times \vec{B}|_x$. This integral is evaluated over the path of the track, so it depends not only on the final slopes but on the actual trajectory of the track over the B field.

The method used to obtain this integral is by simulating toy tracks on the span of all physical slopes and momentum values within the LHCb volume and propagate them through the magnetic field from the VELO up to the first MUON station. This method allows to isolate the integral in terms of the $t_{x,\text{VELO}}$, $t_{y,\text{VELO}}$ and also of $\Delta t = \frac{t_{x,\text{MUON}}}{\sqrt{1+t_{x,\text{MUON}}^2+t_{y,\text{MUON}}^2}} - \frac{t_{x,\text{VELO}}}{\sqrt{1+t_{x,\text{VELO}}^2+t_{y,\text{VELO}}^2}}$. Then, this parameterisation takes the following shape: $\int |d\vec{l} \times \vec{B}|_x = f(t_{x,\text{VELO}}, t_{y,\text{VELO}}, \Delta t)$, this means that the difference in slope maps onto a momentum value differently for each VELO track. In order to capture the different nuances a double fitting procedure is applied:

First, for each VELO track ($t_{x,\text{VELO}}$,$t_{y,\text{VELO}}$) the effect of the B field is fitted with a fourth degree polynomial. Then, the following expansion is used:

$$f(t_{x,\text{VELO}}, t_{y,\text{VELO}}, \Delta t) = \sum_{i=0}^{4} = c_i \, (\Delta t)^i \tag{4.4.6}$$

which is fitted to a new 2D polynomial in $(t_{x,\text{VELO}}, t_{y,\text{VELO}})$ for each coefficient expansion term, c_i, that respects the observed symmetries, as can be seen in Fig. 4.4.3.

A closure test is applied to the toy-track based parameterisation, showing that the momentum resolution, ideally is bounded by $\sim 1\%$ as can be seen in Fig. 4.4.4 (left). In the actual implementation the resolution is significantly smeared due to the multiple scattering

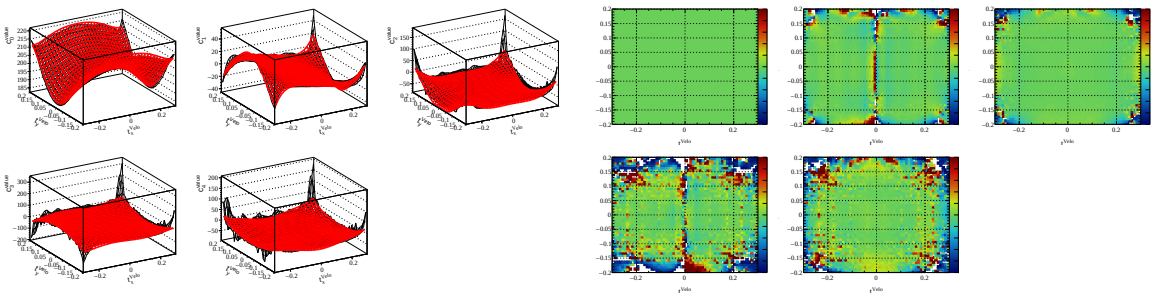

Figure 4.4.3: Left:2D polynomial expansion of the coefficients defined in Eq. 4.4.6. Right: residuals of the fits.

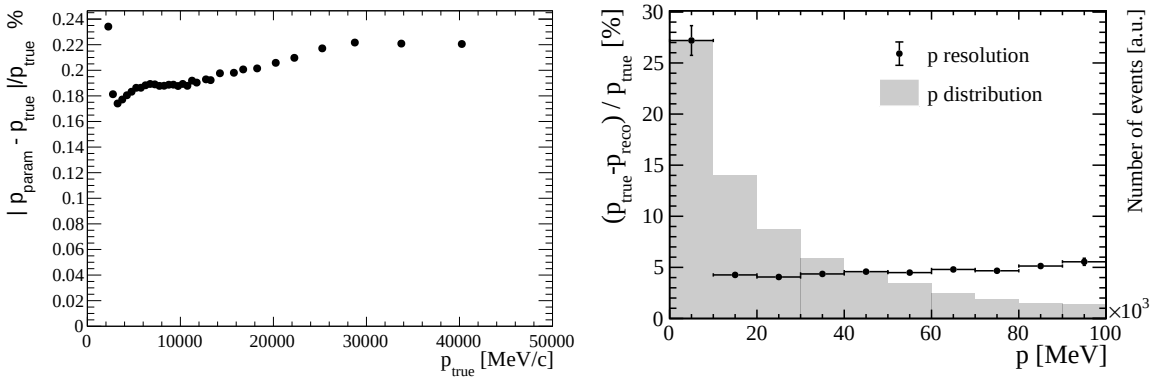

Figure 4.4.4: Left: Optimal resolution obtained with the toy track parameterisation through a closure test with the toy tracks. Right: resolution on $B_s^0 \to \mu^+\mu^-$ simulated events.

and the limited spatial resolution of the MUON stations, the resolution evaluated in simulated $B_s^0 \to \mu^+\mu^-$ decays can be seen in Fig. 4.4.4 (right). A very visible effect is the drop in quality that happens for $p < 10\,\text{GeV}$, this is because the implicit momentum cut applied by requiring the four MUON stations to be populated, so most of the candidates in this momentum bin are actually fake candidates. VELO-MUON tracks are then propagated to the same vertex fitter that produces secondary vertices for regular long tracks. This is possible because the VELO segment is the relevant piece of information.

A configurable dimuon trigger selection was set up, in this selection one can calibrate both kinematic and displacement cuts. To showcase the full reconstruction chain with simulated data, both a displaced and a prompt topology were chosen: $B_s^0 \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$, respectively the output of this selection can be found in Fig. 4.4.5 with a Double Crystal fit overlaid.

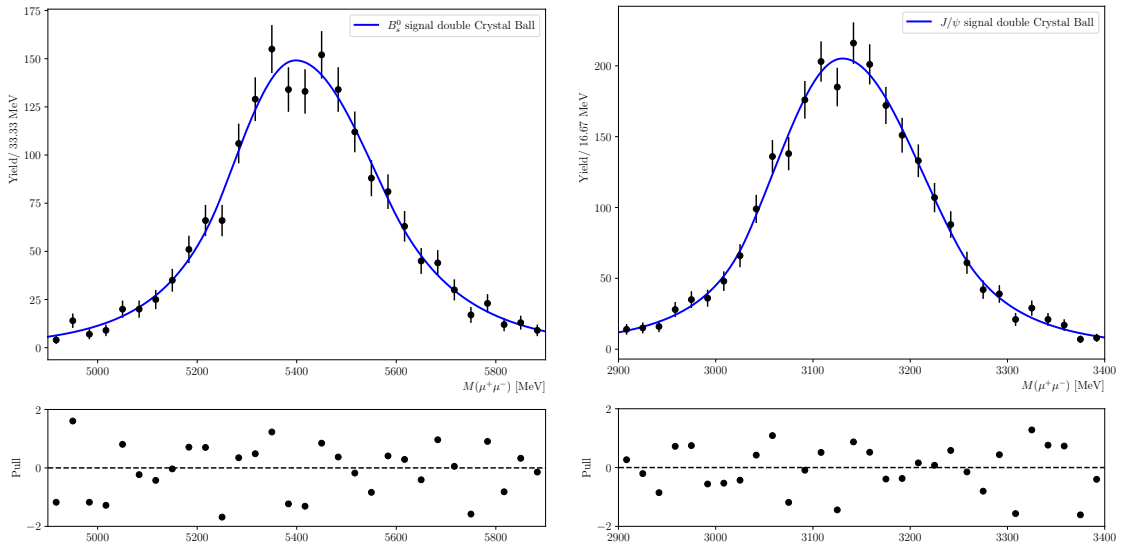

Figure 4.4.5: Mass spectrum of J/ψ $\rightarrow \mu^+\mu^-$ and $B_s^0 \rightarrow \mu^+\mu^-$ signals which showcase the ability to reconstruct both prompt and displaced muon signals.

4.4.3 Conclusions

A full VELO-MUON reconstruction with the standalone reconstruction of MUON stubs has been developed. It uses a custom B-field parameterisation for the momentum calcuation based on the technique implemented for the SciFi tracker. It can be especially useful for exceptionally busy events, where the MUON would still be much cleaner than the trackers. VELO-MUON tracks could also serve to calibrate the tracking efficiencies at the trigger level. It has been demonstrated that dimuon candidates can be resolved with an acceptable momentum resolution of 5%.

The sequence is designed to work as part of the Allen framework and optimised to run on a GPU. The port to a CPU platform would be straightforward. At this stage there are some obvious improvements that could be added on top, both at the standalone muon reconstruction stage and at the matching stage.

The standalone MUON reconstruction at the time requires the four stations to fire, this requirement implies a momentum cut of $p > 10$ GeV which might be inefficient for most physics cases. Lower momentum muons will only leave of hits in the three first stations. The further matching of VELO and MUON segments is controlled by the variable defined in Eq. C.4.1. This parameters could be further fine-tuned in order to avoid killing real muons. Furthermore, these parameters were initially chosen to show-case the reconstruction chain, so a clean signal was favored over an efficient tool. In terms of putting this on a trigger line it might be beneficial to endure more fake tracks for the sake of saving more muons.

A possible follow-up work would be to adapt this algorithm for VELO-UT-MUON tracks. This idea was already explored in [103], but without standalone MUON reconstruction. The

key advantage of VELO-UT tracks over VELO tracks is that these count on an estimate of the momentum before the matching to a MUON segment. This extra piece of information would allow knowing beforehand what region of the MUON stations is crucial, also on the bending plane.

Chapter 5

Search for ALP $\to \gamma\gamma$ at LHCb

In this chapter the search for a diphoton resonance in the LHCb experiment, interpreted as an ALP decaying to two photons is described. The prospects are also interpreted as a B_s^0, B^0 and η_b meson decaying to two photons and the sensitivity is reported. The author of the text is the main contributor to the analysis.

5.1 Data samples

In this section the recorded data (Sec. 5.1.1) as well as Monte-Carlo (MC) (Sec. 5.1.2) samples are presented. The photon kinematics are processed with the post-calibration tool implemented through the `CaloPostCalib` algorithm, which recalibrates the calorimeter clusters offline and refits the decay tree afterwards

5.1.1 Recorded data samples

This analysis uses the full 2018 pp data sample, collected at $\sqrt{s} = 13\,\text{TeV}$. This amounts to an integrated luminosity of $2.02 \pm 0.04\,\text{fb}^{-1}$ plus about $0.05\,\text{fb}^{-1}$ where the integrated luminosity could not be precisely calculated and therefore a generous uncertainty of 50 % is assigned. This leads to a total luminosity of $2.07 \pm 0.05\,\text{fb}^{-1}$, which is used for the analysis. The signal and control channels are taken from different stripping/TurCal streams due to their availability:

- the signal ALP $\to \gamma\gamma$ selection was reconstructed with `Stripping34`. A small fraction of this data set ($4\,\% \approx 80\,\text{pb}^{-1}$) is unblinded and used for the training of the final selection classifier. The size of the data corresponds to data that was made public together with the publication of the of the first diphoton trigger [104].
- the $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi\gamma$ samples are selected through `Stripping34r0p1`
- the $\eta \to \mu^+\mu^-\gamma$ sample was reconstructed through the TurCal `Turbo05` selection.

5.1.2 MC samples

Simulated samples are used to determine efficiencies, investigate backgrounds and calibrate efficiencies.

- Two sets of minimum bias simulation samples (Event types 30000051 and 30000052), simulated with Sim09h, filtered with the cuts of the stripping line without applying the trigger and a TCK 0x617d18a4. Events of type 30000051 contain at least one photon or π^0 in acceptance with $p_T > 4$ GeV. After the stripping filter about 1M events are available. Events of type 30000052 contain at least two photon or two π^0 in acceptance with each $p_T > 3$ GeV. After the stripping filter 663000 events are available for analysis.

- ALP $\to \gamma\gamma$ samples with the masses $[5,6,7,8,9,10,11,13,15,17,19,20]$ GeV. The generation process of ALPs is modelled by the authors of [25] and implemented through an UFO model using MADGRAPH [89], while the underlying event is generated from Pythia 8 [88]. The UFO model provides the production mechanism of ALPs through gluon fusion and allows them to decay to a $\gamma\gamma$ final state. Generation cuts are applied: $p_T(\gamma) > 500$ MeV, $p(\gamma) > 5500$ MeV and inside LHCb acceptance. Currently relying on privately simulated samples, this followed the Sim10a model, designed for the Sim10 validation campaign.

- $B_s^0 \to \gamma\gamma$ (Event Type 13100212) samples, Stripping filtered, TCK: 0x617d18a4. 8M events Sim09k; 2M events Sim10b. No simulated samples for $B^0 \to \gamma\gamma$ were generated, since the kinematics is extremely similar to $B_s^0 \to \gamma\gamma$.

- $\eta \to \mu^+\mu^-\gamma$ (Event Type 39112231) with all decay products and the η in acceptance, as well as μ $p_T > 500$ MeV, η $p_T > 1$ GeV and γ $p_T > 2.5$ GeV. Offline a more strict p_T cut is applied to the signal samples because the signal has harder p_T distributions, such that this p_T cut cannot bias the calibration. 4M events, Sim09k and Sim10b each, TCK: 0x617d18a4. Since the ALP signal samples have to be Sim10, the Sim10b sample is used as baseline.

- $B^0 \to K^{*0}\gamma$ (Event type 11102204) with all final state particles in the LHCb acceptance and the photon inside the ECAL acceptance and with $p_T > 1.8$ GeV; produced with Sim09j and Sim10b, stripping filtered, TCK: 0x617d18a4, after the Stripping filter there are 750k events (4M for Sim10b) available. Since the ALP signal samples have to be Sim10, the Sim10b sample is used as baseline.

- $B_s^0 \to \phi\gamma$ (Event type 13102263) with the same requirements as on $B^0 \to K^{*0}\gamma$. The TCK is 0x617d18a4 and the simulation version Sim09j and Sim10b were used. After the stripping filter about 720k events (4.5M for Sim10b) are available. Since the ALP signal samples have to be Sim10, the Sim10b sample is used as baseline.

All simulated samples are truth-matched via the TRUEID [93] of the full decay chain (the background category tool breaks in some corner cases in the diphoton final state). The ALP

samples are required to have a TRUEID of 22 for the photons and 54 of the photon ancestor and the ALP candidate (and ± 531 for the $\overleftrightarrow{B}{}^0_s$). For the control modes mostly the same matching strategy is applied: all reconstructed particles in the decay chain $B^0 \to K^{*0}\gamma$ and $B^0_s \to \phi\gamma$ is truth matched and the photon ancestor is additionally matched to the B^0 or B^0_s. For the $\eta \to \mu^+\mu^-\gamma$ a slightly different truth matching is applied: the background events are removed through the *sPlot* method. The tail parameters of the signal mass distribution are fixed from a fit to the distribution with the BKGCAT 0 [105] and then the non-matched distribution with a signal plus background model to extract the background component. This procedure was found to have a more stable behaviour in the tails of the mass distribution. Details can be found in Sec. 5.2.3.

5.2 Selection

5.2.1 Signal selection

The ALP $\to \gamma\gamma$ selection consists on the trigger and stripping selections and offline a multivariate classifier based on isolation variables, as well as fiducial requirements. A summary of the selections for the ALP $\to \gamma\gamma$ mode is given in Tab. 5.2.1.

The L0 selection is given in Tab. 5.2.2 and consists of an effective p_T cut on the hardest photon of the candidate, where a photon candidate is made out of a cluster of 2×2 cells. At the same time a standard global event cut is applied.

The HLT1 selection consists of a unique selection in the Run 2 configuration of LHCb. It makes a unique usage of the `L0` ECAL reconstruction (i.e. the 2×2 clusters from the L0 stage are reused) to build diphoton candidates (denoted by $\gamma\gamma$). The selection requirements of the HLT1 selection can be seen in Tab. 5.2.3.

The selection implemented at the HLT2 level is separated in two different parts: the first one implements some kinematic cuts as detailed in Tab. 5.2.4, mainly on the E_T of the single photons and that of the $\gamma\gamma$ candidate; the second one makes use of a neural network classifier based on kinematic and PID variables. In contrast to the `L0` and HLT1 level, photon candidates at HLT2 level and offline are composed of clusters made by 3×3 calorimeter cells.

The online selection of diphotons is discussed in detail in [106]. A multilayer perceptron with a simple hidden layer topology was chosen: 3 neurons in the first and 2 the second. The corresponding activation function for this hidden layers is the logistic function. For the training as a signal proxy simulated $B^0_s \to \gamma\gamma$ and ALP $\to \gamma\gamma$ samples were used, while the background consisted of MinBias pp collision data collected in 2017. The list of variables that are fed to the MLP classifier are the following:

- $\gamma\gamma(p_T)$
- γ `IsNotH` (aka `gamma_CL`), a multivariate classifier to separate photons from hadrons
- γ `ShowerShape`, the second momenta of the deposited energy in the ECAL

Table 5.2.1: Summary of the ALP $\rightarrow \gamma\gamma$ selection. Details on the trigger and higher level selections are given in the text. For $B_s^0 \rightarrow \gamma\gamma$ a dedicated BDT is trained with a separate working point.

Cut	Units	Selection
Trigger		
L0		`L0Photon_TOS` ‖ `L0Electron_TOS`
HLT1		`Hlt1Bs2GammaGamma_TOS` ‖ `Hlt1Bs2GammaGammaHighMass_TOS`
HLT2		`Hlt2RadiativeB2GammaGamma_TOS`
Stripping		
$\gamma\gamma_M$	MeV	[4800,20000]
$\gamma\gamma_{p_T}$	MeV	2000
γ_{p_T}	MeV	> 1100
γ_p	MeV	> 6000
γ_{CL}	-	> 0.3
$\gamma_{HCAL2ECAL}$	-	< 0.1
Offline		
γ_{p_T}	MeV	> 3000
γ Saturation		no cluster cell saturated
γ IsPhoton		> 0.85
BDT ($BDT_{B_s^0}$)		> 0.9 (> 0.85)

Table 5.2.2: L0 cuts applied on the `0x617d18a4` TCK, standard for simulation productions under 2018 conditions. The transformation between ADCs and E_T is such that each ADC corresponds to 24 MeV.

Line	nSPDHits	ADC counts	E_T [MeV]	$\sum E_T$ previous event [MeV]
L0Photon	≤ 450	> 123	> 2952	< 24000
L0Electron	≤ 450	> 99	> 2376	< 24000

Table 5.2.3: Selection of the HLT1 trigger. The `Hlt1Bs2GammaGammaHighMass` line was only available for 2018 data taking, while the other line ran also in 2017.

Line	$\gamma(p_T)$ [MeV]	$\sum\gamma(p_T)$ [MeV]	$\gamma\gamma(p_T)$ [MeV]	$\gamma\gamma$(M) [MeV]
`Hlt1Bs2GammaGamma`	≥ 3500	≥ 8000	≥ 2000	[3500,6000]
`Hlt1Bs2GammaGammaHighMass`	≥ 5000	≥ 11000	≥ 5000	[6000,11000]

Table 5.2.4: Selection of the ALP $\to \gamma\gamma$ HLT2 trigger. It was improved for 2018 data taking. A different version ran also in 2017 data taking.

Cut	Units	`Hlt2RadiativeB2GammaGamma`
$\gamma\gamma_M$	MeV	$\in [3500, 20000]$
$\gamma\gamma_{p_T}$	MeV	> 2000
γ_{p_T}	MeV	> 2000
$\sum\gamma_{p_T}$	MeV	> 6000
MVA		> 0.8

The stripping selection applies further kinematic requirements on the individual photons and the diphoton candidates, as well as particle identification requirements: the photons are required to have a good identification as a photon ($\gamma_{CL} = $ `IsNotH`) and the ratio of deposited energy in the HCAL over the ECAL cells which correspond to the area of the ECAL cluster is supposed to be small. On top of this, the minimum p_T on the photons is tightened to > 3 GeV (to allow a safe calibration with $\eta \to \mu^+\mu^-\gamma$) and signal candidates are vetoed where at least one photon saturates at least one calorimeter cell (this can happen for $p_T > 12$ GeV). Such saturation causes the signal peak to be almost impossible to separate from background due to a sizeable tail on the left side, as can be seen on Fig. 5.2.1. However, for high masses

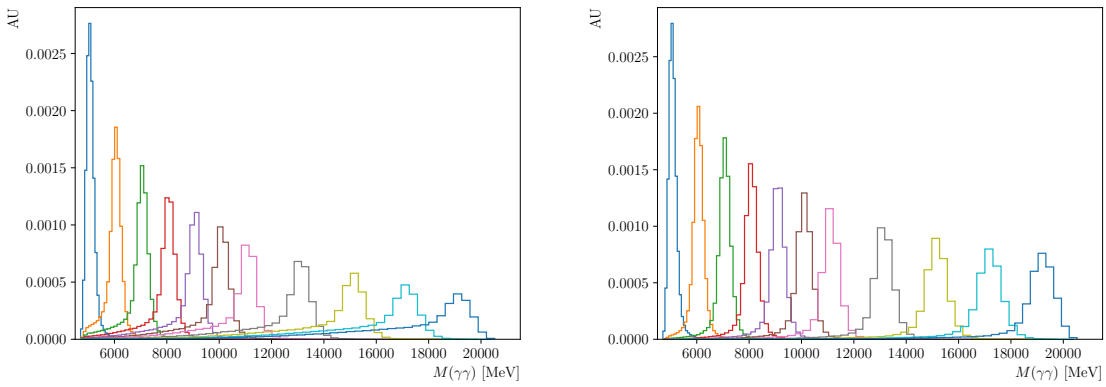

Figure 5.2.1: Reconstructed ALPs peaks without (left) and with (right) vetoing saturation. Distributions are normalized.

this requirement also creates a sizeable efficiency penalty. The saturation efficiencies on simulation after stripping, truth matching and offline selection (no trigger applied) are listed in Tab. 5.2.5.

Table 5.2.5: Efficiency of the saturation veto on the several simulated signal samples. The efficiency is evaluated after Stripping, truth matching and the photon p_T requirement (no trigger selection applied). The efficiencies of the 5 GeV and the 20 GeV ALP are not reliable as part of the signal distribution reaches out of the mass window.

m_{ALP} [GeV]	ε [%]
5	90.10(12)
6	78.90(14)
7	75.20(14)
8	74.16(13)
9	73.72(12)
10	73.16(12)
11	71.77(11)
13	67.50(12)
15	61.38(13)
17	54.79(13)
19	46.65(14)
20	31.06(18)

5.2.2 Multivariate classifier and further particle identification

In order to obtain a higher signal to background discrimination a model was developed using an XGBoost classifier [107]. The classifier optimizes a logistic function for binary classification and as metric evaluates the area under the *receiver operating characteristic* (ROC) curve. It was chosen because of its simplicity, robustness and because the comparative between different classifiers did not show any big discrimination power differences.

It is trained using the full ensemble of ALP$\to \gamma\gamma$ simulation samples as signal proxy and a small (4 %, as large as the open data set) data sample as background. To account for the difference size of the samples that are in an 18:1 relation, the imbalance factor of XGBoost is used that weights the under-represented sample to effectively match the sample sizes.

Cone isolation variables are chosen as features to discriminate between signal and background, which is efficient in suppressing combinatorial background (which vastly dominates all other background contributions). They inspect the charged tracks in cones in η and φ of the sizes $[1.0, 1.35, 1.7]$ around the diphoton candidate. The features explored are the numbers of charged tracks, the sum of the absolute momenta and transverse momenta, as well as the asymmetry of the sum of the momenta and transverse momenta with respect to the diphoton momentum and transverse momentum:

- Cone track mulitplicity
- Cone p

- $\frac{p_{\text{Cone}}-p_\gamma}{p_{\text{Cone}}+p_\gamma}$
- $\frac{p_{\text{T}}(\text{Cone})-p_{\text{T}}(\gamma)}{p_{\text{T}}(\text{Cone})+p_{\text{T}}(\gamma)}$

The signal vs. background distributions of these features are displayed in Fig. 5.2.2. The features can be seen in Fig. 5.2.2 and their correlations are displayed in Fig. 5.2.3.

To check the overtraining of the classifier a k-fold strategy [108] with 5 folds is followed. Each fold is validated separately both by scoring the are under the ROC curve and by showing the training/test sample response of the operator. The scores can be found in Tab. 5.2.6 and an example of the performance of one of the folds in Fig. 5.2.4 (Top Left).

Table 5.2.6: Cross validation scores of the ALP classifier.

ALP $\rightarrow \gamma\gamma$ BDT	0.920	0.920	0.921	0.922	0.919

The performance can be checked in the ROC curve presented in Fig. 5.2.4. As proof that the classifier is not producing fake peaks on the background, Fig. 5.2.5 shows that the classifier, although it changes a bit the background profile, it does not create any peaking structure.

The best working point was selected by maximizing the figure of merit [109]

$$\text{FOM} = \frac{\varepsilon}{3/2+\sqrt{N_{\text{bkg}}}} \tag{5.2.1}$$

for several ALP masses. In order to calculate ε and N_{bkg} the search window is restricted to a 3σ region around the signal peak. As demonstrated in Fig. 5.2.4 (bottom), the optimum varies slightly across the investigated mass range, which is due to the strongly varying amount of background depending on the mass region.

To define the final selection requirement an optimisation study is performed in the same way together with a selection on the particle identification variables. Large backgrounds come from π^0 combinations. Therefore the variable `IsPhoton` has significant extra separation power as illustrated in Fig. 5.2.6. The optimal cut point has been investigated simultaneously together with the BDT with the same figure of merit as above. The results are displayed in Fig. 5.2.7. The optimisation shows on the one hand little correlation between the BDT and the `IsPhoton` selection. On the other hand it again shows some variation across the scanned mass region due to the varying backgrounds. Following this optimisation strategy the final selection is set to BDT > 0.9 and `IsPhoton` > 0.85 (for both photons). For the $\text{B}^0_\text{s} \rightarrow \gamma\gamma$ search the BDT is retrained and reoptimised, taking into account also the $\min(\gamma p_{\text{T}})$ and $\max(\gamma p_{\text{T}})$, using the same `IsPhoton` > 0.85 cut, finding a cut point at $\text{BDT}_{\text{B}^0_\text{s}} > 0.85$.

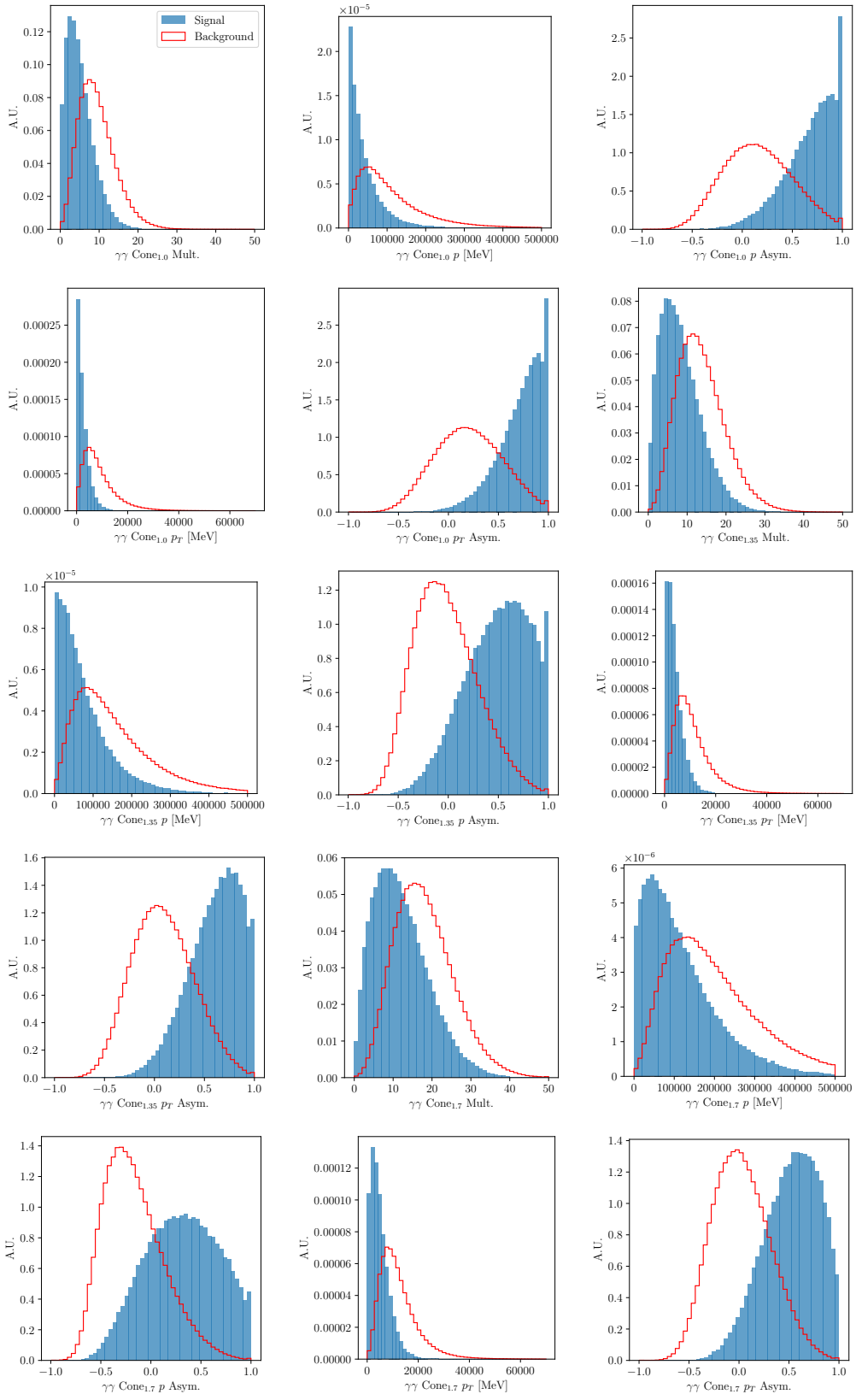

Figure 5.2.2: Signal versus background distributions of the features used in the signal classifier.

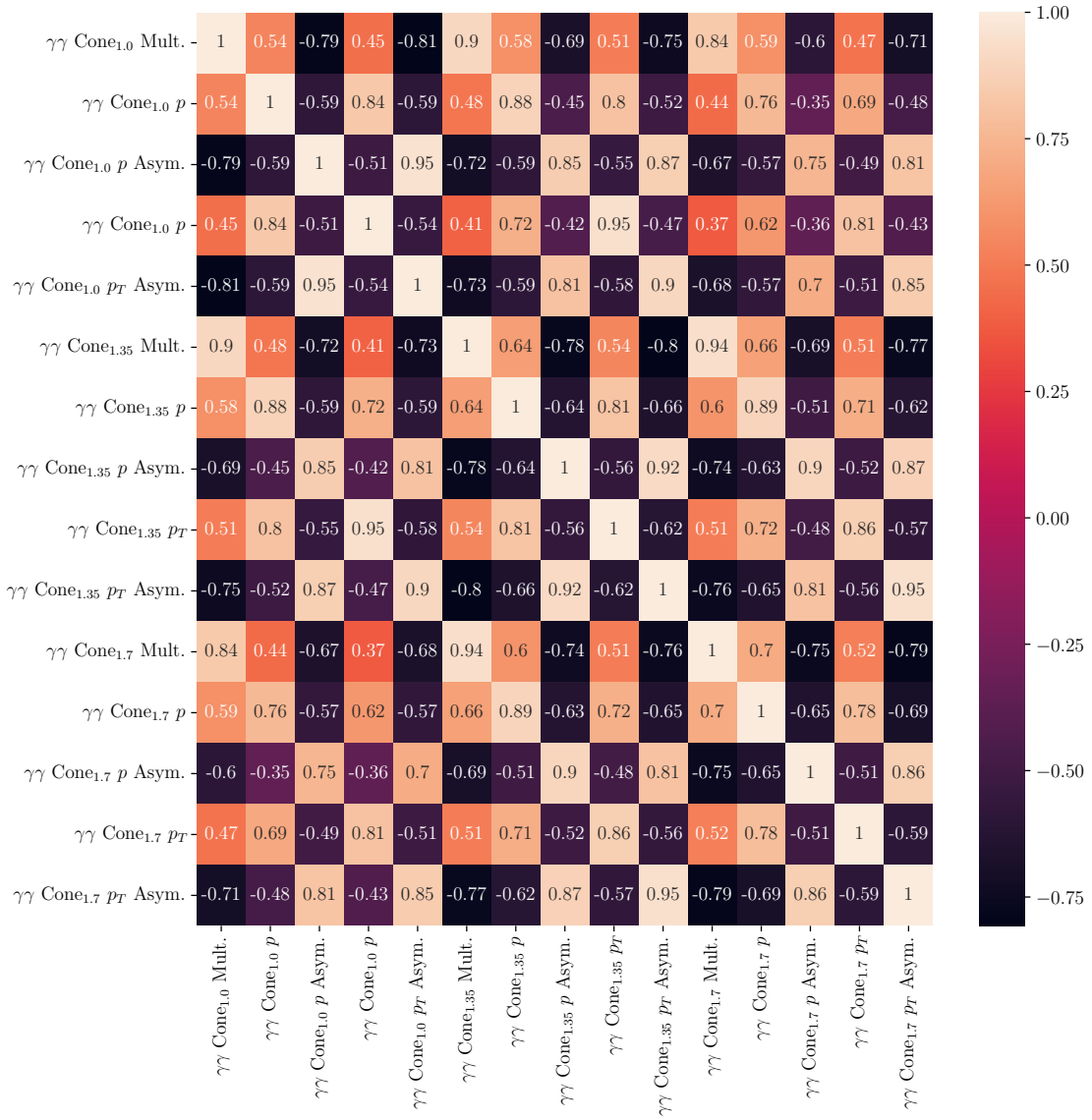

Figure 5.2.3: Correlation matrix of the variables involved in the classifier for the signal dataset.

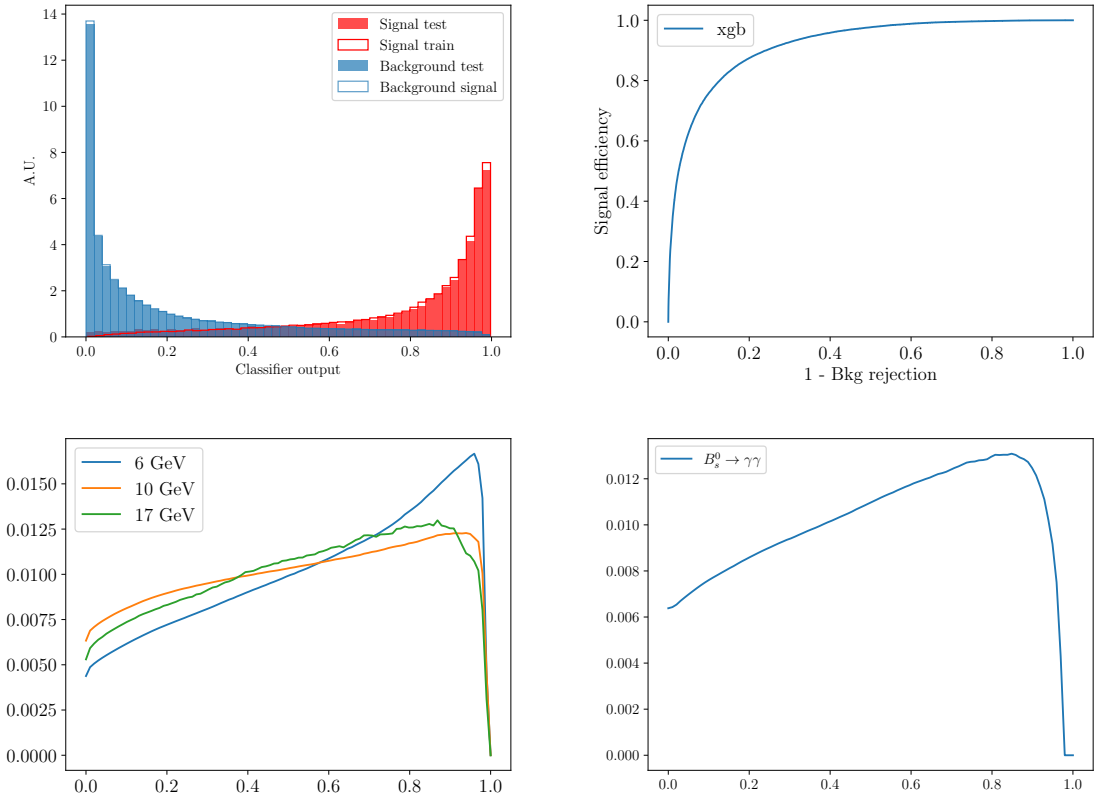

Figure 5.2.4: Upper left: performance of the ALP $\to \gamma\gamma$ classifier. Upper right: peformance ROC curve. Bottom: Normalised Punzi figure histograms for difference mass hypotheses based on a single requirement on the BDT for different ALP hypotheses (left) and $B_s^0 \to \gamma\gamma$ (right).

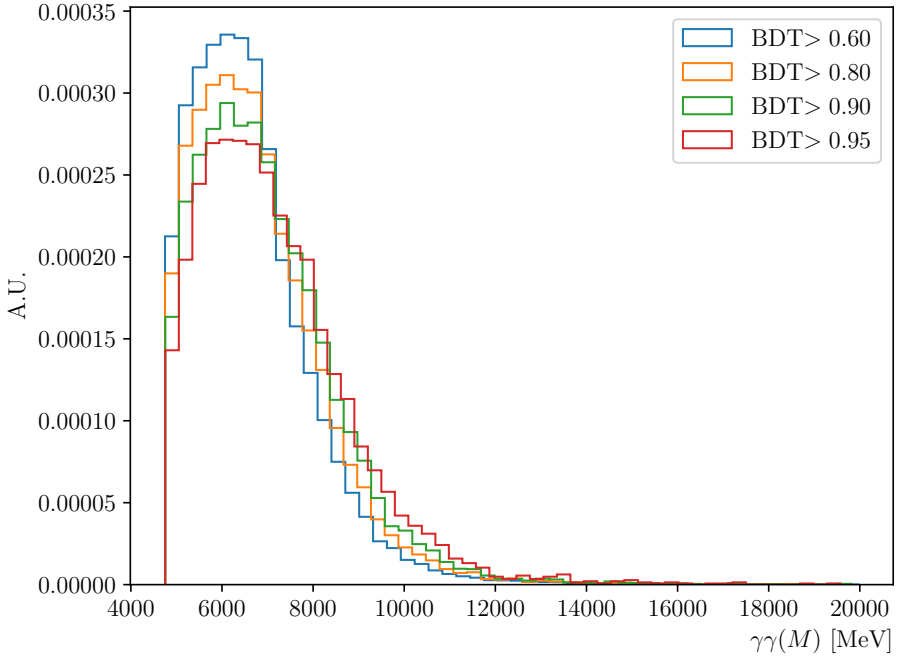

Figure 5.2.5: Effect of different BDT cuts on the mass shape of the background data sample. While tight cuts lead to a wider background, it doesn't create any peaking structure.

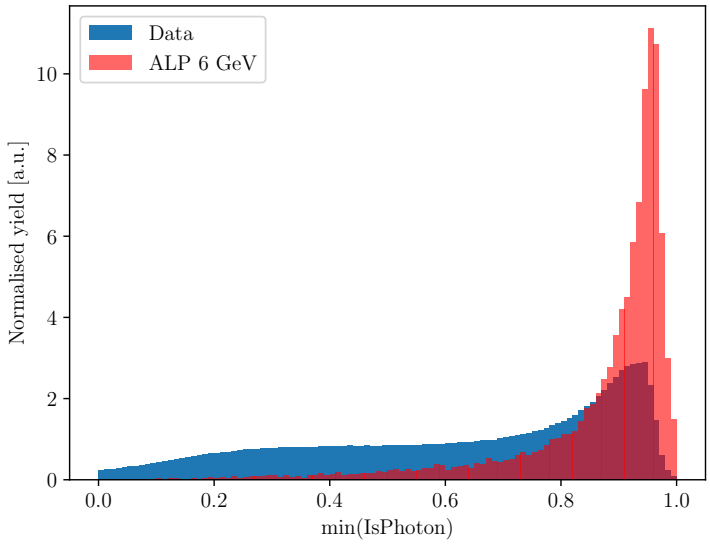

Figure 5.2.6: Distributions of the minimum of the `IsPhoton` variable for the photon candidates, comparing background and ALP signal.

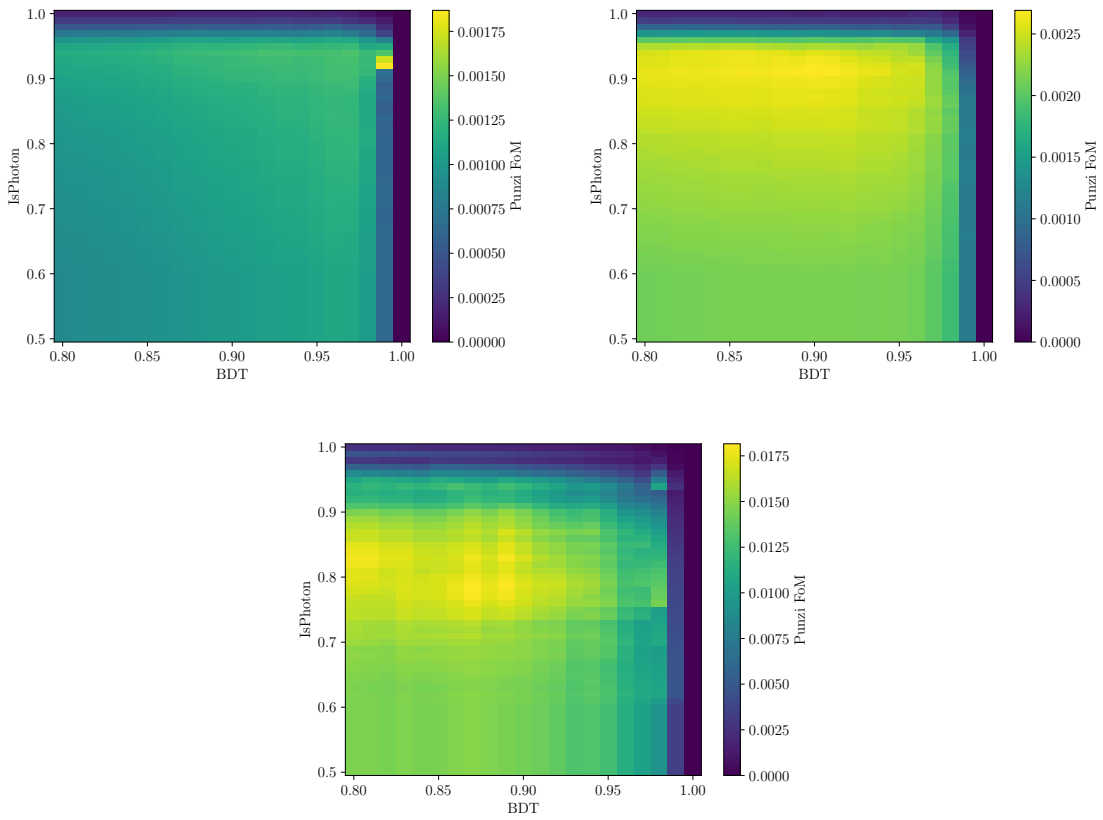

Figure 5.2.7: Two-dimensional optimisation of the Punzi figure of merit for the 6 GeV ALP (left), the 10 GeV ALP (middle) and the 17 GeV ALP (right).

5.2.3 Control channels

The efficiencies and photon distributions are calibrated with three control modes:

- $B_s^0 \to \phi\gamma$
- $B^0 \to K^{*0}\gamma$
- $\eta \to \mu^+\mu^-\gamma$

Their selection criteria are summarised in Tab. 5.2.7. The $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decays are selected using the `StrippingBeauty2XGammaExclTDCPVBs2PhiGammaLine` `StrippingBeauty2XGammaExclTDCPVBd2KstGammaLine` and from the `Stripping34r0p1` re-stripping campaign of 2018 data, while $\eta \to \mu^+\mu^-\gamma$ is selected as TurCal sample through the `Hlt2CaloPIDEta2MuMuGammaTurboCalibDecision` line. On top of that the selection summarised in Tab. 5.2.8 is applied. After this selection the $\eta \to \mu^+\mu^-\gamma$ samples are already very pure. Only to simplify the data-MC corrections an additional cut at the dimuon mass < 430 MeV is added.

For the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ channels however, in order to obtain an effective background subtraction, an additional selection was applied. For this selection an xgboost [107] classifier was trained for both channels. Due to the similarities of the decay topology, similar strategies were followed. The variables that showed higher separation power: separation of the hadronic secondary vertex and photon isolation. Since the photon is reconstructed using only the energy deposited in the ECAL its reconstruction is not involved in the secondary vertex determination. This way, the secondary vertex of the prompt hadron is effectively the same as that of the neutral B meson.

The specific variables that were used as features of the classifier are the following, cone isolation variables were finally not included, since the $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi\gamma$ candidates were used to evaluate MC/Data differences regarding the multiplicity.

- DIRA(B): the cosine of the angle between the momentum of the B meson and the direction of flight from the best PV to the decay vertex.
- Vertex$_{\chi^2}(hh)$
- IP$_{\chi^2}$(Tracks)

As signal, $B^0 \to K^{*0}\gamma$ ($B_s^0 \to \phi\gamma$) MC samples were used. For the background proxy the right side band in data, outside the fit range - $m(h\gamma) > 5700(5800)$ MeV - is used, where no signal contributions are expected. The distributions of the input variables are displayed in Figs. 5.2.8 and 5.2.9.

To train the classifier a k-fold technique with 5 folds was used. Then the model was cross validated and scores of that validation were compatible up to very small fluctuation effects, these can be checked on Tab. 5.2.9. The overall performance of each classifier can be observed in Fig. 5.2.10. The decays are selected requiring BDT > 0.6.

Table 5.2.7: $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ stripping selections.

Cut	Units	`Stripping34r0p1`	Offline
Track χ^2/ndf		≤ 4	-
Track Ghost prob		≤ 0.6	-
Track χ^2_{IP}		≥ 9	-
Tracks min p	MeV	≥ 3000	-
Tracks min $(p_T(h^+), p_T(h^-))$	MeV	≥ 250	-
Tracks $\sum p_{\text{T}}$	MeV	≥ 500	-
Tracks and γ $\sum p_{\text{T}}$	MeV	≥ 3000	-
$K^{\pm}$ ProbNNk		-	-
$K^{\pm}$ ProbNNpi		-	-
$\pi^{\pm}$ ProbNNk		-	-
V χ^2_{vtx}	MeV	≤ 25	-
ϕ ΔM	MeV	≤ 15	-
K^* ΔM	MeV	≤ 100	-
γ E_{T}	MeV	≥ 2500	-
γ CL		-	≥ 0.3
B p_{T}	MeV	$\geq$2500	-
B χ^2_{IP}		≤ 16	-
B χ^2_{vtx}		≤ 16	-
B Mass	MeV	[4000,7000]	-
Trigger			
L0		`L0Photon_TOS` $\|\|$ `L0Electron_TOS`	
HLT1		`Hlt1TrackMVA_TOS`	
HLT2		`Hlt2RadiativeBd2KstGammaDecision_TOS` (`Hlt2RadiativeBs2PhiGammaUnbiasedDecision_TOS`)	
BDT			> 0.6

Table 5.2.8: Selection on the $\eta \to \mu^+\mu^-\gamma$ sample on top of the TurCal selection. The selection in the lower part is applied to facilitate the data-MC corrections.

Cut	Units	Value
γ p_{T}	MeV	≥ 500
μ p_{T}	MeV	≥ 500
μ ProbNNmu		≥ 0.8
μ χ^2_{IP}		≤ 6
$m(\eta)$	MeV	$\in [405, 695]$
$m(\mu\mu)$	MeV	< 430

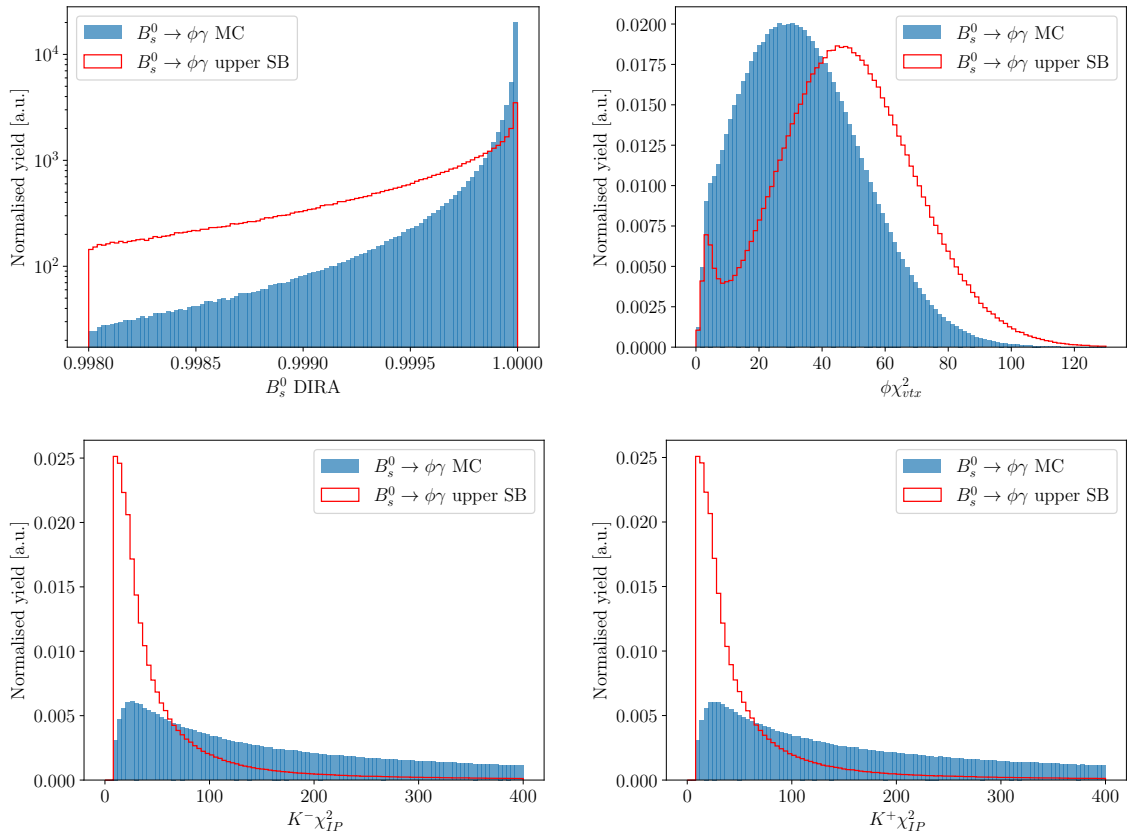

Figure 5.2.8: Distributions of the background (red) and signal (blue) distributions of the input variables for the BDT used to extract a more pure $B_s^0 \to \phi\gamma$ signal. Trigger and truth matching are applied.

Table 5.2.9: Cross validation scores of the control channel BDTs.

$B^0 \to K^{*0}\gamma$	0.895	0.893	0.894	0.8941	0.894
$B_s^0 \to \phi\gamma$	0.920	0.919	0.919	0.919	0.920

From the thus selected samples, pure signal distributions are extracted with the *sPlot* technique [110], using the mass distributions as discriminating variable. In the fit to the mass distribution in the boundaries m_1 and m_2 the signal shape is described as a double sided Crystal Ball function (Eq. 5.2.2, with the conditions on Eq. 5.2.3 and the transformation on

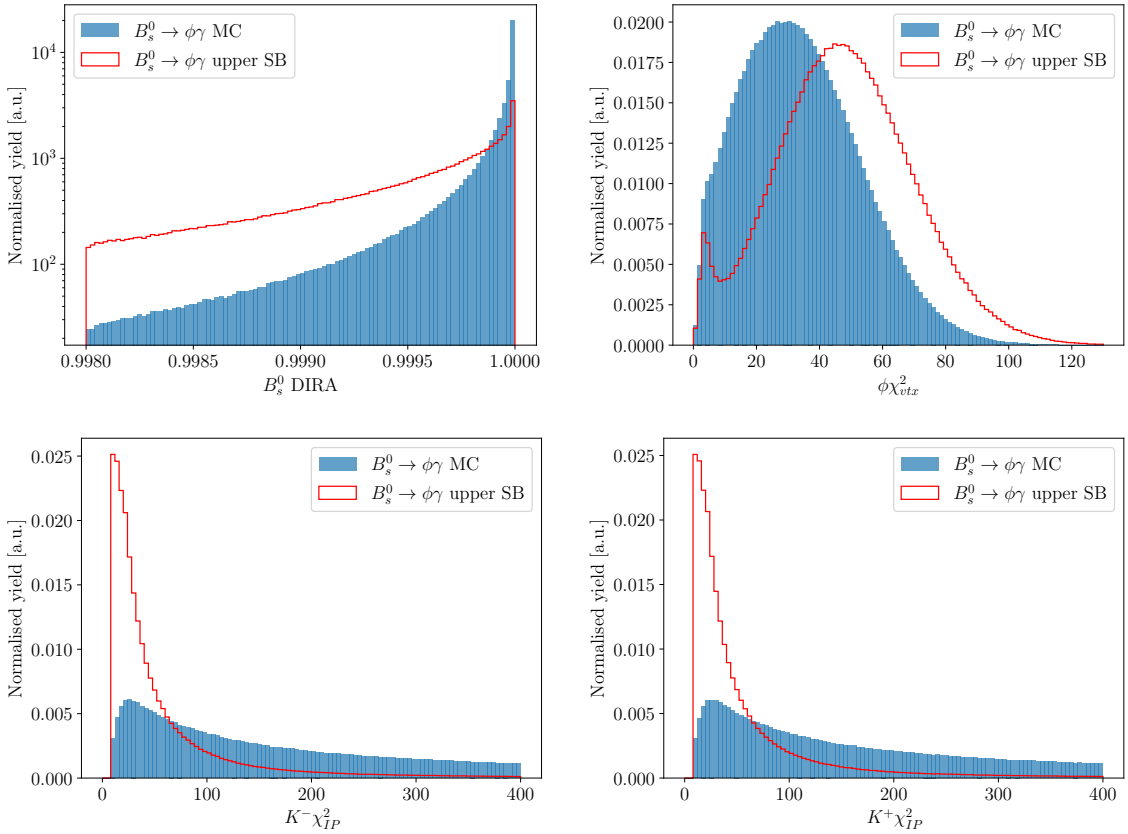

Figure 5.2.9: Distributions of the background (red) and signal (blue) distributions of the input variables for the BDT used to extract a more pure $B^0 \to K^{*0}\gamma$ signal. Trigger and truth matching are applied.

Eq. 5.2.4) [111].

$$f(m;m_0,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} A_L \cdot (B_L - \frac{m-m_0}{\sigma_L})^{-n_L}, & \text{for } \frac{m-m_0}{\sigma_L} < -\alpha_L \\ \exp\left(-\frac{1}{2}\cdot\left[\frac{m-m_0}{\sigma_L}\right]^2\right), & \text{for } \frac{m-m_0}{\sigma_L} \leq 0 \\ \exp\left(-\frac{1}{2}\cdot\left[\frac{m-m_0}{\sigma_R}\right]^2\right), & \text{for } \frac{m-m_0}{\sigma_R} \leq \alpha_R \\ A_R \cdot (B_R + \frac{m-m_0}{\sigma_R})^{-n_R}, & \text{otherwise,} \end{cases} \tag{5.2.2}$$

$$\begin{aligned} A_i &= \left(\frac{n_i}{|\alpha_i|}\right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2}\right) \\ B_i &= \frac{n_i}{|\alpha_i|} - |\alpha_i| \end{aligned} \tag{5.2.3}$$

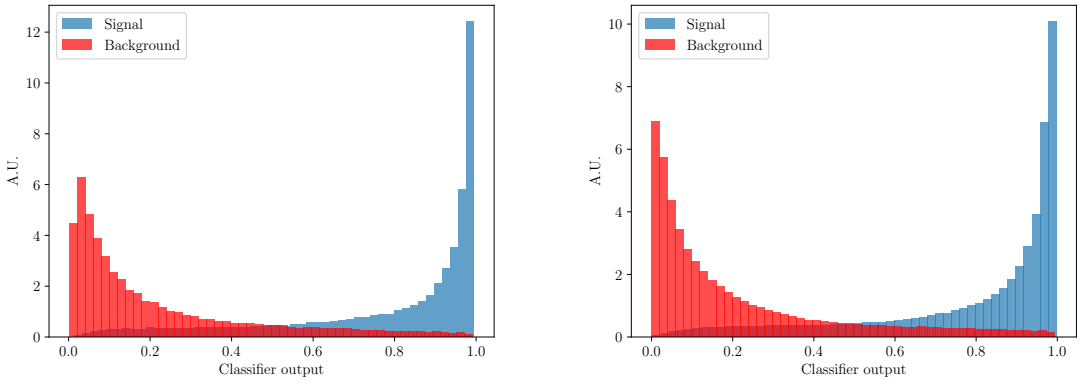

Figure 5.2.10: Left: $B_s^0 \to \phi\gamma$ performance. Right: $B^0 \to K^{*0}\gamma$ classifier performance.

$$\begin{aligned} \sigma_R &= \sigma + \Delta(\sigma) \\ \sigma_L &= \sigma - \Delta(\sigma) \end{aligned} \tag{5.2.4}$$

As background model, a linear function like Eq. 5.2.5 is used, where the slope parameter is denoted with a:

$$f(m) = a \cdot \left(2\frac{m - m_1}{m_2 - m_1} - 1 \right) + C \tag{5.2.5}$$

For fitting the $\eta \to \mu^+\mu^-\gamma$ samples the background model used is an exponential with the slope parameter λ. Due to the high statistics the signal model is extended by two additional Gaussian cores to improve the description for this mode.

The fit proceeds by first determining the shape parameters from simulation with a signal-only model and then fitting the signal plus background model to data with fixed tail parameters (for $\eta \to \mu^+\mu^-\gamma$ the tail parameters are readily determined with the available data and don't need to be fixed). The resulting distributions are shown in Figs. 5.2.11, 5.2.12 and 5.2.13 with the parameters reported in Tabs. 5.2.10, 5.2.11 and 5.2.12. With this fit model the *sPlot* [110] technique is applied to the data and the sWeights obtained that project out the pure signal distributions.

The main control mode $\eta \to \mu^+\mu^-\gamma$ exploits a similar strategy in simulation to obtain the optimal truth matching: the signal model only is fitted to the fraction of data with BKGCAT 0 first. Then the tail parameters of the Double Crystal Ball are fixed from that fit and the full non-matched simulation is fit by the signal model and a single exponential for the background candidates. Following this fit the *sPlot* method is applied to extract the pure signal distribution.

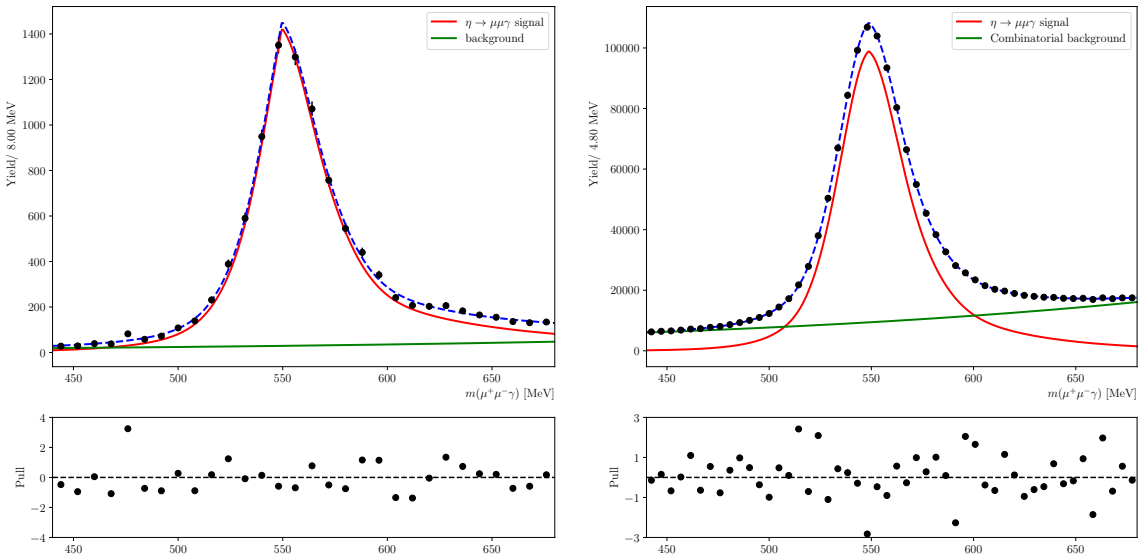

Figure 5.2.11: $\eta \to \mu^+\mu^-\gamma$ MC (left) and data (right) fit plot to determine sweights. The resulting fit parameters are summarised in Tab. 5.2.10.

Table 5.2.10: Results of the fit to $\eta \to \mu^+\mu^-\gamma$ distributions, corresponding to Fig. 5.2.11.

Parameter	Units	Value MC	Value Data
m/m_{PDG}		1.0029 ± 0.0026	$1.001\,54 \pm 0.000\,26$
σ	MeV	3.0 ± 1.9	11.797 ± 0.028
$\Delta\sigma$	MeV	2.1 ± 1.5	-9.891 ± 0.028
α_l		0.052 ± 0.012	1.394 ± 0.013
n_l		6.9 ± 4.3	18.5 ± 3.1
α_r		0.17 ± 0.06	0.0500 ± 0.0007
n_r		1.8 ± 0.6	28 ± 4
fraction CB		0.61 ± 0.15	0.5297 ± 0.0035
fraction Gauss 1		0.08 ± 0.09	0.2435 ± 0.0022
m/m_{PDG} Gauss 1		1.007 ± 0.006	$1.001\,25 \pm 0.000\,13$
σ Gauss 1.	MeV	9.9 ± 2.9	11.95 ± 0.05
m/m_{PDG} Gauss 2		1.0124 ± 0.0029	$1.017\,99 \pm 0.000\,31$
σ Gauss 2	MeV	18.9 ± 1.9	18.15 ± 0.10
λ		0.0038 ± 0.0018	$0.004\,10 \pm 0.000\,04$
n_{Bkg}		939 ± 440	$512\,054 \pm 1700$
n_{Sig}		9378 ± 450	$1\,006\,320 \pm 1700$

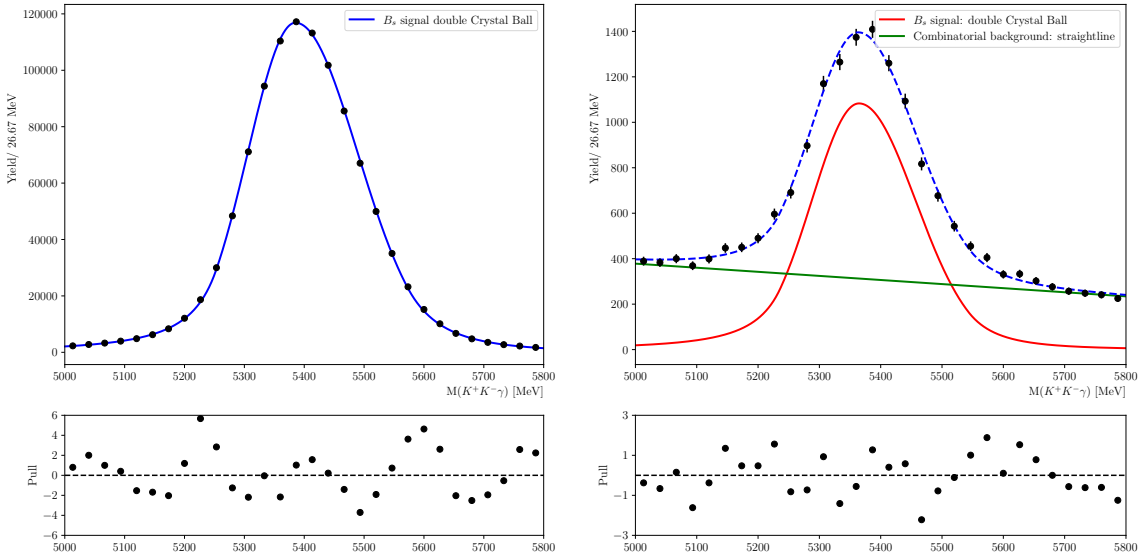

Figure 5.2.12: $B_s^0 \to \phi\gamma$ MC (left) and data (right) fit plot to determine sweights. The resulting fit parameters are summarised in Tab. 5.2.11.

Table 5.2.11: Results of the fit to $B_s^0 \to \phi\gamma$ distributions, corresponding to Fig. 5.2.12.

Parameter	Units	Value MC	Value Data
m/m_{PDG}		$1.003\,64 \pm 0.000\,07$	0.9997 ± 0.0007
σ	MeV	92.16 ± 0.12	84.0 ± 1.7
$\Delta(\sigma)$	MeV	11.90 ± 0.31	7.4 ± 3.2
α_L		1.670 ± 0.012	-
α_R		1.764 ± 0.017	-
n_L		1.99 ± 0.05	-
n_R		3.82 ± 0.21	-
n_{Signal}		-	9570 ± 190
$n_{\text{Background}}$		-	$11\,840 \pm 200$
a		-	-0.238 ± 0.020

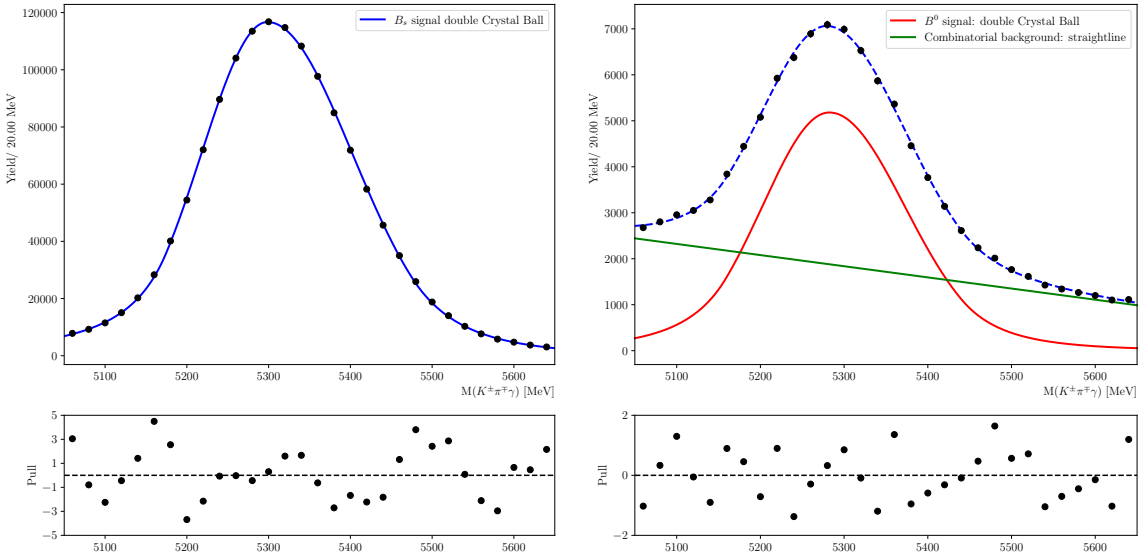

Figure 5.2.13: $B^0 \to K^{*0}\gamma$ MC (left) and data (right) fit plot to determine sweights. The resulting fit parameters are summarised in Tab. 5.2.12.

Table 5.2.12: Results of the fit to $B^0 \to K^{*0}\gamma$ distributions, corresponding to Fig. 5.2.13.

Parameter	Units	Value MC	Value Data
m/m_{PDG}		$1.004\,04 \pm 0.000\,10$	$1.000\,97 \pm 0.000\,32$
σ	MeV	92.63 ± 0.17	85.9 ± 0.9
$\Delta(\sigma)$	MeV	10.6 ± 0.4	3.3 ± 1.5
α_L		1.591 ± 0.026	-
α_R		1.640 ± 0.026	-
n_L		1.91 ± 0.18	-
n_R		5.9 ± 0.7	-
n_{Signal}		-	$60\,100 \pm 700$
$n_{\text{Background}}$		-	$59\,100 \pm 700$
a		-	-0.409 ± 0.011

5.3 Correlation between the photons

In this analysis the aim is to search for particles that have two photons in the final state. These photons are reconstructed mainly by the ECAL, though some global event and particle identification cuts involve the usage of the PRS and SPD detectors. In order to keep under control the different efficiencies it was decided to use high yield radiative (single-photon) decay channels, as there are no suitable diphoton control modes. In order to propagate accordingly the efficiency calibrations one has to understand whether there is a strong correlation factor between the photons or if, rather, their reconstruction efficiency behaves for each photon independently.

A first study consists in addressing what is the chance for overlap between the 3x3 cluster portion of the ECAL that is used to reconstruct each photon. For this one should retrieve the center of the clusters and construct a 3x3 array of ECAL squared cells that, depending on what ECAL region the photon reached, will have a different sizes:

- Inner: 40 mm
- Middle: 60 mm
- Outer: 120 mm

Within the size of our signal (ALP, B_s^0) simulation samples, no candidate is found in which the two photons share any cell. This first fact already shows that the measurement of energy and that the shape of the electromagnetic shower produced by each photon should not affect the other - meaning that the energy resolution or the particle identification quantities should not be affected by having a pair of highly energetic photons in each event.

A complementary way of seeing this is by exploiting the the definition of independent events. Let A and B, be independent events then it means that the following equation holds:

$$p(B|A) = p(B) \tag{5.3.1}$$

It follows then that the probability of the intersection of two events factors out as the product of their probabilities:

$$p(A \cap B) = p(A)p(B|A) = p(A)p(B) \tag{5.3.2}$$

This fact shows that, in general single photon efficiencies can be propagated and distributions as if each photon was independent from the other without adding a relevant bias.

This effectively tells us that the union probabilty follows:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = p(A) + p(B) - p(A)p(B) \tag{5.3.3}$$

Eqs. 5.3.2 and 5.3.1 are tested on the 6 ALP mass hypothesis. The stripping and saturation selection are applied as well as the global event cut, as described in Sec. 5.2.1 and then check

$\varepsilon_{\texttt{L0}}(\mathrm{B}_\mathrm{s}^0)$	$\varepsilon_{\texttt{L0}}(\gamma\gamma)$	$\frac{\varepsilon_{\texttt{L0}}(\mathrm{B}_\mathrm{s}^0)-\varepsilon_{\texttt{L0}}(\gamma\gamma)}{\varepsilon_{\texttt{L0}}(\mathrm{B}_\mathrm{s}^0)}$
89.59 ± 0.09	89.78 ± 0.06	-0.0021 ± 0.0013

Table 5.3.1: Comparison of the union efficiency approximation using the L0 trigger selection. The number $\varepsilon_{\texttt{L0}}(\mathrm{B}_\mathrm{s}^0)$ is obtained from requiring the `L0` selection on either of the two photons, while $\varepsilon_{\texttt{L0}}(\gamma\gamma)$ uses Eq. 5.3.4. An excellent agreement is obtained.

$\varepsilon_{p_\mathrm{T}}(\mathrm{ALP}\to\gamma\gamma)$	$\varepsilon_{p_\mathrm{T}}(\gamma\gamma)$	$\frac{\varepsilon_{p_\mathrm{T}}(\mathrm{ALP}\to\gamma\gamma)-(\varepsilon_{p_\mathrm{T}}(\gamma\gamma)}{\varepsilon_{p_\mathrm{T}}(\gamma\gamma)}$
68.33 ± 0.17	68.06 ± 0.14	0.0039 ± 0.0013

Table 5.3.2: Comparison of the intersection efficiency approximation using the photon kinematic selection. The number $\varepsilon_{p_\mathrm{T}}(\mathrm{ALP}\to\gamma\gamma)$ is obtained from requiring both photons to fulfill the cut and $\varepsilon_{p_\mathrm{T}}(\gamma\gamma)$ uses Eq. 5.3.5. An excellent agreement is obtained.

that the relations between the two photons γ_1 and γ_2 hold:

$$\varepsilon(\gamma\gamma) = \varepsilon(\gamma_1) + \varepsilon(\gamma_2) - \varepsilon(\gamma_1)\varepsilon(\gamma_2) \tag{5.3.4}$$

for the `L0` trigger efficiency and

$$\varepsilon(\gamma\gamma) = \varepsilon(\gamma_1) \times \varepsilon(\gamma_2) \tag{5.3.5}$$

for the photon $p_\mathrm{T} > 3$ GeV cut.

(5.3.6)

An excellent agreement is found. It is immediate to also show that the efficiency of the PID cuts on the stripping selection follows Eq. 5.3.2.

5.4 Data-simulation corrections on the control modes

To be able to understand the data-simulation differences in the photon properties, the control modes are used. However, these suffer also from mismodelled production kinematics in simulation. So before performing data-simulation difference studies for the signal, the $\eta \to \mu^+\mu^-\gamma$ kinematics is reweighted in simulation in the $(\eta, p_\mathrm{T}, M(\mu^+\mu^-))$ space to match the data. For this a GBreweighter implementation from `hepml` has been used. The sweights coming from the fit described in Sec. 5.2.3 are properly propagated to the reweighter.

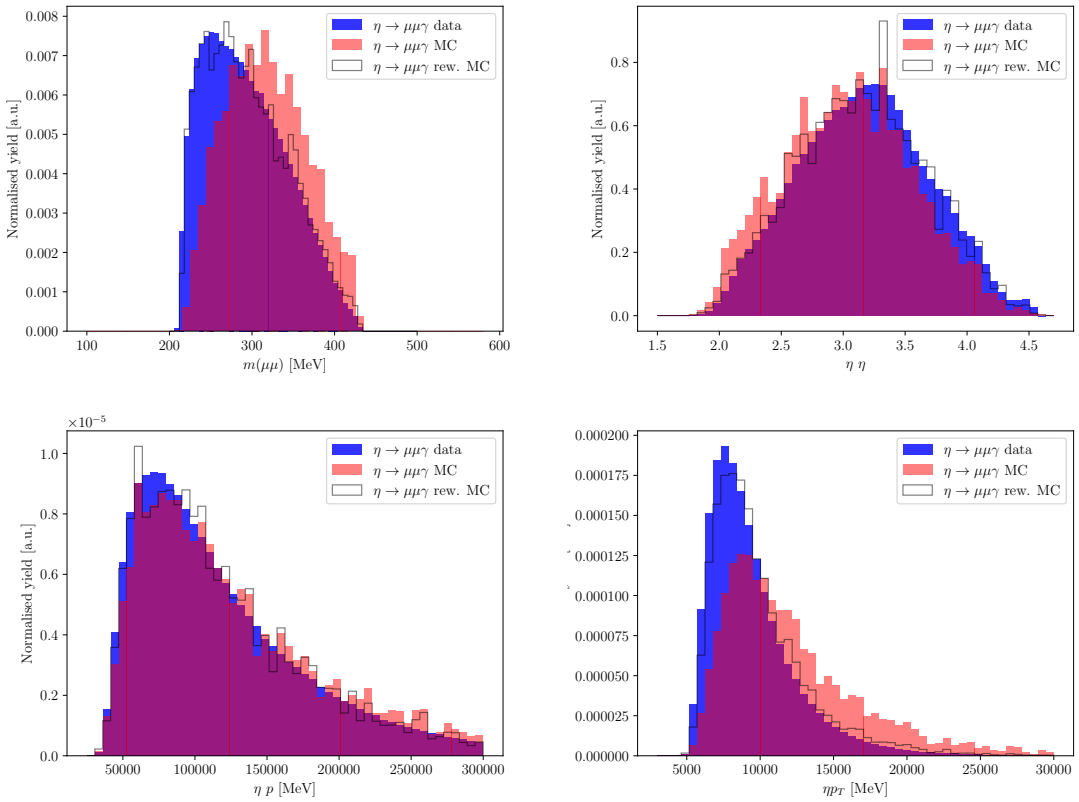

Figure 5.4.1: Comparison of $\eta \to \mu^+\mu^-\gamma$ kinematics between simulation and sWeighted Data before and after reweighting.

A comparison of the kinematic variables before and after the reweighting can be found in Fig. 5.4.1

5.5 Efficiencies and normalisation factor

The selection efficiencies are taken partially from simulation with cross checks performed with the control modes (geometric efficiency at generation level, kinematic stripping selection, selection BDT) and partially calibrated from data with $\eta \to \mu^+\mu^-\gamma$ (PID selection, trigger selection, saturation efficiencies). The mode $\eta \to \mu^+\mu^-\gamma$ is chosen as it has the highest number of high p_T photons amomg the photon modes. The following subsections go through all selection steps and discuss the efficiencies and corresponding systematics in the order they are applied in simulation. Every efficiency is calculated conditional to the previous, i.e. the stripping efficiency is calculated given the generator level cuts etc. The efficiencies of the generator level cuts are determined at the generation time with the Gauss framework

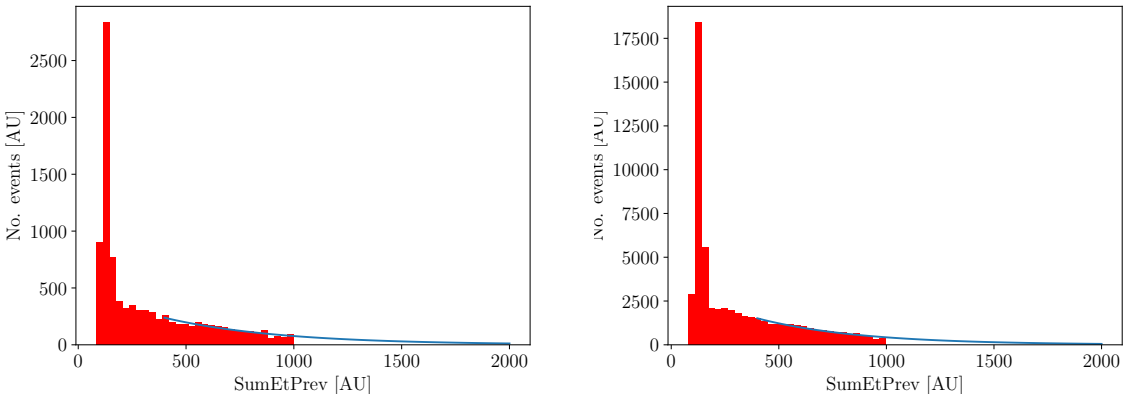

Figure 5.5.1: Completion of `SumEtPrev` distribution. Left: $B_s^0 \to \phi\gamma$. Right: $B^0 \to K^{*0}\gamma$.

directly and no systematic uncertainties are assigned to that procedure. Also no systematic uncertainties are assigned to the photon reconstruction itself as studies have shown perfect data-simulation agreement in the kinematic regime of this analysis ($p_T > 0.7$ GeV) [112].

5.5.1 Global Event Cut

The hardware L0 trigger implements the selection `L0Photon_TOS` || `L0Electron_TOS`. It is useful to discuss the Global Event Cut that this selection implements (`nSPDHits` < 450 & `SumEtPrev` < 1000) independently before the other selections. The reason behind this is to more easily handle the control channels to calibrate the remaining efficiencies given the global event cut already applied.

In order to calculate the `SumEtPrev` cut efficiency, the $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi\gamma$ channels were used. For both, an exponential pdf is fit to the tail of the `SumEtPrev` distribution and extrapolated into the removed region. This pdf is integrated over the removed region to obtain the efficiency of the cut. This procedure yields an efficiency of 94.1 % and 94.8 % respectively. As nominal value the mean is taken and the small difference between the two values (0.7 %) is taken as systematic uncertainty.

To obtain the efficiency of the requirement on `nSPDHits`, the distribution in simulation has to be scaled to match the data. While in data the distributions look similar across all control modes, in simulation they vary significantly. It is however not to be expected that the nSPDHits distribution stays constant across the whole mass range. Indeed, in the ALP simulations a trend is found towards higher occupancies with higher ALP masses. The `nSPDHits` efficiency is therefore calculated cutting on the signal distribution after applying a scale factor obtained from the control channels $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi\gamma$. These vary from 1.24 to 1.36, respectively. The best scaling factor was obtained by minimising the χ^2 between the original and the target distribution following the method described in Ref. [113]. Not

knowing the degree to which the ALP production is modelled, the true scale factor for the ALPs is assumed to be within that range, while for $B_s^0 \to \gamma\gamma$ the factor from $B_s^0 \to \phi\gamma$ (1.36) is assumed. A systematic uncertainty is assigned by calculating the efficiencies with both values and applying as systematic the difference between the efficiency from the mean scaling factor and the bound that yields are higher difference in efficiency, obtaining a systematic of around 5%, detailed values per mass can be checked at Tab. 5.5.1 (the efficiencies range from 78 % at the lowest mass to 67 % at the highest mass). Fig. 5.5.2 shows the nSPDHits distribution in simulation and after applying the optimal scaling factor compared to the distributions in data.

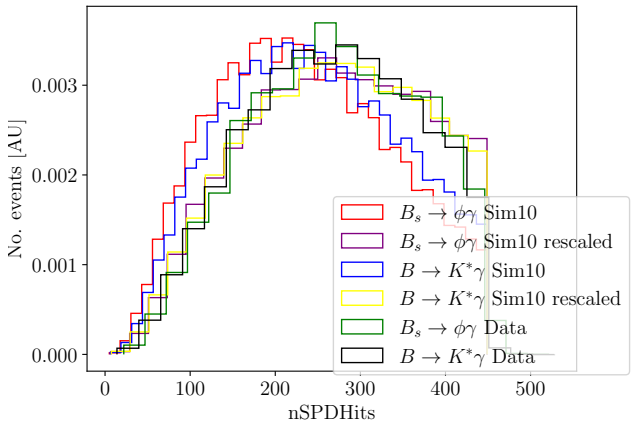

Figure 5.5.2: Obtention of the nSPDHits scale factor using $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ control channels.

Mass [GeV]	nSPDHits
$B_s^0 \to \gamma\gamma$	68.0 ± 0.6
5	78.1 ± 2.8
6	75.2 ± 3.0
7	72.8 ± 3.2
8	71.2 ± 3.3
9	70.1 ± 3.4
10	69.4 ± 3.5
11	68.5 ± 3.5
13	68 ± 4
15	67 ± 4
17	67 ± 4
19	67 ± 4
20	66 ± 4

Table 5.5.1: nSPDHits efficiency calculated for each mass hypothesis

5.5.2 Stripping and preliminary offline selection

The efficiencies of the stripping and offline selection are calculated from signal simulation. Here, together with the stripping selection also the extra p_T requirement on the photons as well as the `IsPhoton` requirement is considered. The stripping efficiency systematics are determined from the $\eta \to \mu^+\mu^-\gamma$ sample. The largest effect come from the γ_{CL}, `HCAL2ECAL` and `IsPhoton` cuts (which are PID cuts), as they are not perfectly modelled in simulation. A plot showing the `HCAL2ECAL` and `IsNotH` (aka γ_{CL}) and `IsPhoton` variables from the $\eta \to \mu^+\mu^-\gamma$ photon is shown in Fig. 5.5.3.

The effect from possible momentum resolution effects has been investigated by smearing the photon momenta with the 9 % different resolution in data before applying the kinematic cuts. The difference in efficiency is $< 1\%$. This is a negligible effect compared to the PID selection systematics, no correction and neither a systematic uncertainty for this is applied. In order to understand how the efficiencies of the PID cuts behave between data and

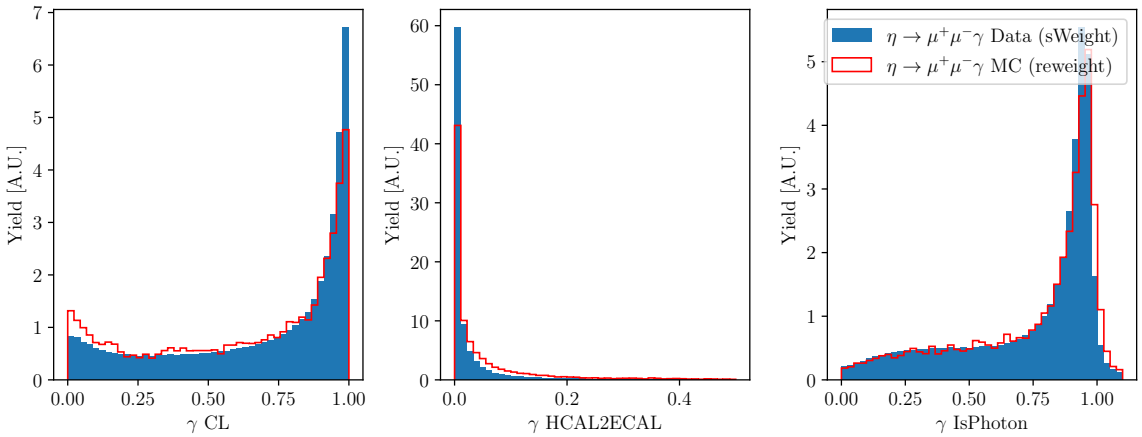

Figure 5.5.3: MC-data differences on photon PID variables using the $\eta \to \mu^+\mu^-\gamma$ decay channel.

simulation single photon efficiencies are calculated with $\eta \to \mu^+\mu^-\gamma$ in bins of the cluster-track matching χ^2 (shortened to χ^2_{match}) in the bins $[0, 50, 100, 150, 200, 400, 600, 800, \infty]$. Although this binning might seem not typical when compared to the charged particle binning schemes of PIDCalib (they use the particle kinematics and event occupancy), this is the result of many attempts. The χ^2_{match} is a direct input to the classifiers for the `IsNotH` and `IsPhoton` variables

The binned efficiencies are then convoluted with the single photon distributions of the ALPs. This exercise yields Tab. 5.5.2. The efficiencies based on ALP simulation are denoted with $\varepsilon^{\text{plain}}$ ($\varepsilon^{\text{plain}}(\gamma\gamma)$ for the combined), while the efficiencies calibrated from $\eta \to \mu^+\mu^-\gamma$ (simulation or data) are denoted with $E^{\text{calib}}_{\text{MC}}$ and $E^{\text{calib}}_{\text{Data}}$. As expected from Sec. 5.3 the product of the efficiencies is compatible with the combined efficiency, which shows fully uncorrelated efficiencies. The efficiencies calibrated from $\eta \to \mu^+\mu^-\gamma$ simulation E_{MC} reproduce the single photon efficiencies from the ALP simulation, albeit with large uncertainties. These uncertainties arise from the comparatively low number of available $\eta \to \mu^+\mu^-\gamma$ simulated events and the large discrepancy in the χ^2_{match} between $\eta \to \mu^+\mu^-\gamma$ and ALP$\to \gamma\gamma$. The nominal value for the efficiencies is chosen as the one calibrated from data and then the difference between the combined and uncorrelated efficiencies ($\varepsilon^{\text{plain}}(\gamma\gamma)$-$(\varepsilon(\gamma_1) \times \varepsilon(\gamma_2))^{\text{plain}}$) is added in quadrature to the relative uncertainty of the efficiencies calibrated from $\eta \to \mu^+\mu^-\gamma$ simulation as systematic uncertainty. This approach is chosen as on the one hand it takes into account the precision of the closure test that proves the method to be working and on the other hand the (negligible) correlations. The full efficiencies are given in the last column of the table together with the kinematic requirements and have systematic uncertainties of up to 13 % (smaller towards higher masses).

In the summary of efficiencies, in Tab. C.5.1, these selections are factorised in **Reconstruction** and **PID** where the first refers to the efficiency of reconstructing the generator level

Table 5.5.2: Per photon efficiencies of the PID cuts using MC and data proxies. The efficiencies calculated from signal simulation $\varepsilon^{\text{plain}}(\gamma)$ are compared to the efficiencies calibrated from $\eta \to \mu^+\mu^-\gamma$ Data ($E^{\text{calib}}_{\text{Data}}(\gamma)$) and $\eta \to \mu^+\mu^-\gamma$ simulation ($E^{\text{calib}}_{\text{Data}}(\gamma)$). The first columns only contain the particle identification requirements, but the last column summarises the final efficiency and also contains the kinematic cuts.

Mass [GeV]	$\varepsilon^{\text{plain}}(\gamma)$	$E^{\text{calib}}_{\text{MC}}(\gamma)$	$E^{\text{calib}}_{\text{Data}}(\gamma)$	$\varepsilon^{\text{plain}}(\gamma\gamma)$	$(\varepsilon(\gamma_1)\times\varepsilon(\gamma_2))^{\text{plain}}$	$E^{\text{calib}}_{\text{MC}}(\gamma\gamma)$	$E^{\text{calib}}_{\text{Data}}(\gamma\gamma)$
$B^0_s \to \gamma\gamma$	75.24 ± 0.17	71 ± 5	76.2 ± 0.4	56.51 ± 0.19	56.49 ± 0.18	51 ± 5	58.1 ± 0.4
5	75.47 ± 0.16	72 ± 6	77.2 ± 0.5	56.98 ± 0.18	56.82 ± 0.17	52 ± 7	59.6 ± 0.5
6	74.68 ± 0.13	72 ± 6	76.9 ± 0.4	55.88 ± 0.15	55.80 ± 0.14	52 ± 6	59.2 ± 0.5
7	74.19 ± 0.12	72 ± 6	76.7 ± 0.4	54.98 ± 0.14	54.89 ± 0.13	52 ± 6	58.8 ± 0.4
8	74.16 ± 0.11	72 ± 6	76.5 ± 0.4	55.08 ± 0.13	54.95 ± 0.12	51 ± 6	58.6 ± 0.4
9	73.79 ± 0.10	72 ± 6	76.5 ± 0.4	54.65 ± 0.12	54.55 ± 0.11	51 ± 6	58.5 ± 0.4
10	73.64 ± 0.10	72 ± 6	76.4 ± 0.4	54.33 ± 0.11	54.24 ± 0.10	51 ± 6	58.4 ± 0.4
11	73.43 ± 0.10	72 ± 5	76.4 ± 0.4	54.11 ± 0.11	54.00 ± 0.10	51 ± 6	58.3 ± 0.4
13	72.85 ± 0.09	71 ± 5	76.19 ± 0.35	53.42 ± 0.10	53.28 ± 0.10	51 ± 5	58.0 ± 0.4
15	72.73 ± 0.09	71 ± 5	76.12 ± 0.35	53.01 ± 0.10	52.90 ± 0.10	51 ± 5	57.9 ± 0.4
17	72.27 ± 0.09	71 ± 5	76.15 ± 0.35	52.42 ± 0.10	52.21 ± 0.09	51 ± 5	58.0 ± 0.4
19	72.03 ± 0.09	71 ± 5	76.23 ± 0.35	52.13 ± 0.10	52.02 ± 0.09	51 ± 5	58.1 ± 0.4
20	72.16 ± 0.11	72 ± 5	76.3 ± 0.4	52.38 ± 0.12	52.02 ± 0.11	51 ± 6	58.3 ± 0.4

decays and applying the kinematic cuts and the former relates to only the PID cuts, as the name suggests.

5.5.3 Saturation veto

To determine a systematic uncertainty for the accuracy of the MC description of the saturation veto, the $\eta \to \mu^+\mu^-\gamma$ control channel is used, as it has the largest amount of "higher p_{T}" photons in data. However, it is severely statistically limited, especially in the simulation size. Due to its very different kinematics compared to the ALP photons it also has very few photons that saturate the ECAL. The efficiency of vetoing photons saturating an ECAL cluster is in simulation $(99.49 \pm 0.17)\,\%$ and in data $(99.716, 0.009)\,\%$, which is in good agreement. To cross check the agreement at the ALP systematic, the $\eta \to \mu^+\mu^-\gamma$ simulation sample is reweighted to the ALP simulation and the saturation veto efficiency is reevaluated. The reweighting is performed with a Gradient Boosting reweighter in the photon p_{T}, p, as well as the ECAL region of the cluster (inner, central or outer). The same classifier is applied to the $\eta \to \mu^+\mu^-\gamma$ data sample. The saturation efficiency of the ALPs is then in good agreement $(1-2\sigma)$ with the reweighted $\eta \to \mu^+\mu^-\gamma$ simulation. Therefore the saturation efficiencies from the reweighted $\eta \to \mu^+\mu^-\gamma$ data sample are used as nominal value for the efficiencies and the uncertainties comprise the data-uncertainties as well as the relative difference between the efficiency from the saturation cut on both photons and the product of the efficiency of

Table 5.5.3: Numbers for the determination of the saturation efficiencies and systematic uncertainty. The "true $\varepsilon(\gamma)$" denotes the per-photon efficiency as determined on the signal simulation, $\varepsilon_{\text{rew}}(\gamma)$ the per-photon efficiency as determined from reweighted $\eta \to \mu^+\mu^-\gamma$ MC and data samples, while "Total" describes the diphoton efficiency and the resulting combined statistical and systematic uncertainty. In the 20 GeV ALP a significant effect from the correlation appears due to the proximity to the upper mass cut from the stripping.

Mass [GeV]	true $\varepsilon(\gamma)$	$\varepsilon_{\text{rew}}(\gamma)$ MC	$\varepsilon_{\text{rew}}(\gamma)$ Data	Total
$B_s^0 \to \gamma\gamma$	0.97213 ± 0.00034	0.984 ± 0.008	0.9884 ± 0.0004	0.9769 ± 0.0006
5	0.9590 ± 0.0009	0.966 ± 0.022	0.9674 ± 0.0013	0.9359 ± 0.0019
6	0.8909 ± 0.0012	0.95 ± 0.05	0.9417 ± 0.0026	0.8868 ± 0.0035
7	0.8691 ± 0.0012	0.95 ± 0.05	0.9272 ± 0.0035	0.860 ± 0.010
8	0.8626 ± 0.0011	0.94 ± 0.04	0.9365 ± 0.0028	0.877 ± 0.013
9	0.8584 ± 0.0011	0.95 ± 0.05	0.9338 ± 0.0029	0.872 ± 0.018
10	0.8554 ± 0.0010	0.94 ± 0.04	0.9267 ± 0.0033	0.859 ± 0.020
12	0.8483 ± 0.0010	0.93 ± 0.04	0.9279 ± 0.0029	0.861 ± 0.021
14	0.8217 ± 0.0010	0.92 ± 0.05	0.908 ± 0.004	0.824 ± 0.019
15	0.7867 ± 0.0011	0.91 ± 0.06	0.888 ± 0.004	0.789 ± 0.011
17	0.7451 ± 0.0012	0.87 ± 0.07	0.864 ± 0.012	0.746 ± 0.015
19	0.6986 ± 0.0012	0.68 ± 0.20	0.846 ± 0.010	0.715 ± 0.030
20	0.6054 ± 0.0014	0.86 ± 0.11	0.819 ± 0.007	0.67 ± 0.11

both individual photons. The resulting values are listed in Tab. 5.5.3. The relative uncertainty of the $\eta \to \mu^+\mu^-\gamma$ simulated efficiencies are in this case not added as extra uncertainty as it is already taken into account in the PID efficiency, where a similar procedure is used.

5.5.4 `L0` efficiency

The L0 selection effectively applies a p_{T} cut on the γ as well as putting an upper level on the number of SPD hits (nSPDHits). The cuts for the `0x617d18a4` can be seen in Tab. 5.2.2

To estimate the efficiencies, the $\eta \to \mu^+\mu^-\gamma$ channel is studied using the TURCAL stream.

The method to calculate and propagate the `L0` efficiencies from the signal photon $\eta \to \mu^+\mu^-\gamma$ decay to the ALP $\to \gamma\gamma$ diphoton topology is separated in two main parts which are independent of each other: a kinematic cut on $\gamma_{E_{\text{T}}}$ and a global event cut: `nSPDHits` < 450 & `SumEtPrev` < 1000, discussed in Sec. 5.5.1.

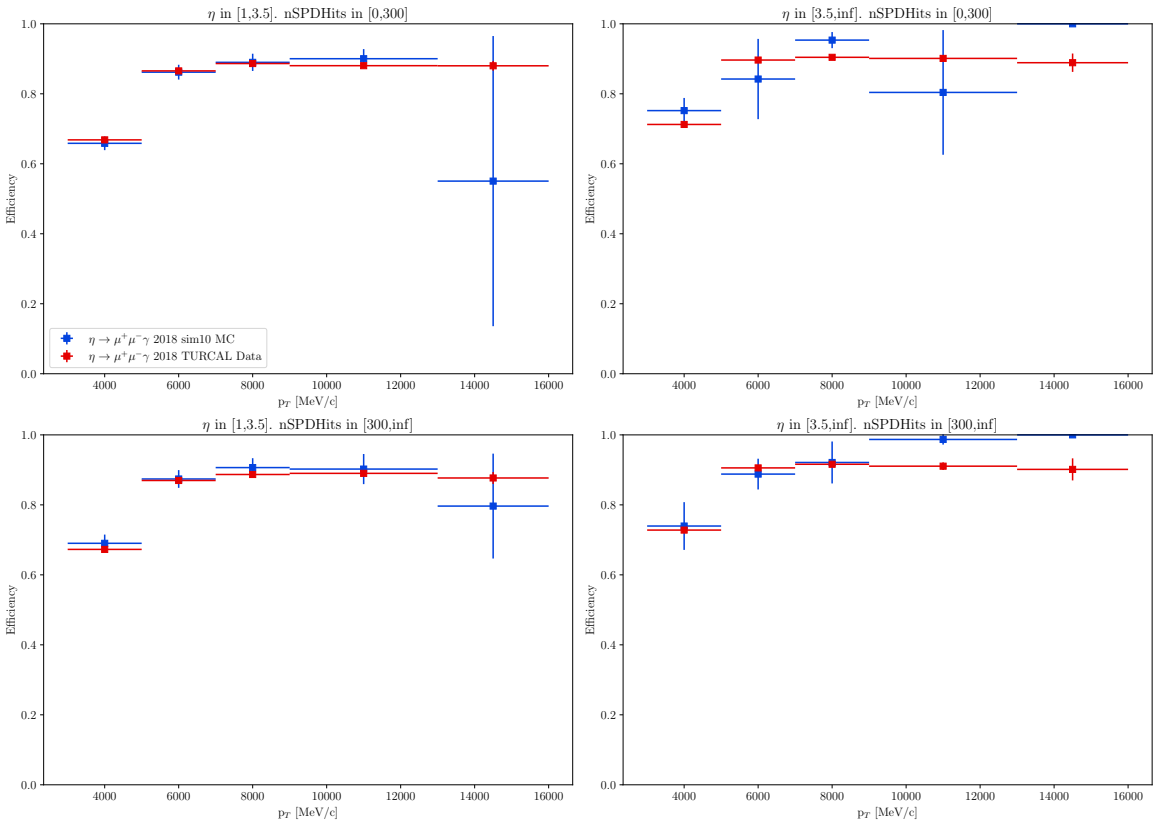

Figure 5.5.4: Single photon MC vs Data L0 efficiency comparison.

In order to calculate the L0 efficiency (modulo nSPDHits and SumEtPrev) in the following quantity is calculated in bins of photon p_{T}, pseudorapidity and nSPDHits:

$$\varepsilon_{\mathtt{L0}} = \frac{N((\mathtt{L0Photon\_TOS} \parallel \mathtt{L0Electron\_TOS}) \;\&\; \mathtt{L0Muon\_Dec} \;\&\; \text{kinematic bin})}{N(\mathtt{L0Muon\_Dec} \;\&\; \text{kinematic bin})} \tag{5.5.1}$$

The inclusion of the L0Muon_Dec cut is a way to have an unbiased sub-sample with respect to the photon, though this line also applies the same global event cuts. In order to calculate the data and MC efficiencies the kinematics of the signal η were reweighted using the sWeighted data, using the fit model described in Sec. 5.2.3. By looking at Fig. 5.5.4 one can see that the bigger impact appears at the lower p_{T} bins. The way to proceed is such that the binned L0 efficiencies are used to recalculate the efficiency for the ALP$\rightarrow \gamma\gamma$ candidates. This is, the efficiencies will be integrated in the kinematic $\times$ global event space weighting by the

ALP$\to\gamma\gamma$ distributions. Thus, the total L0 efficiency is recalculated as:

$$\begin{aligned} \varepsilon_{\texttt{L0}} &= \varepsilon_{\texttt{SumEtPrev}} \times \varepsilon_{\texttt{nSPDHits}} \times \sum_{i\in\text{kin}} \varepsilon_i f_i \\ \text{kin} &= \gamma_{p_\text{T}} \times \gamma_\eta \times \texttt{nSPDHits} \end{aligned} \tag{5.5.2}$$

where f_i and ε_i are respectively the signal fraction and L0 efficency in the i-labelled kinematic bin. The bin boundaries are

- $\gamma_{p_\text{T}} \in [3,5,7,10,12,\infty]$ GeV
- $\gamma_\eta \in [1,2.5,3.5,\infty]$
- nSPDHits $\in [0,200,300,\infty]$

The L0 efficiency study for different ALPs masses can be checked from calibrating with $\eta\to\mu^+\mu^-\gamma$ MC and Data and also cutting simply on the signal MC. The full candidate efficiency is calculated from the per-photon efficiencies using the relation given in Sec. 5.3 in Eq. 5.3.3.

Then, the probability of the intersection is well known that can be calculated with Eq. 5.3.2. The calculated efficiencies can be written as follows, provided ε are the cut-based efficiencies, E are the calibrated efficiencies. Since there is no way to differentiate γ and γ^0, they are not oredered in anyway, it follows that: $\varepsilon(\gamma^0)=\varepsilon(\gamma)$, equivalently, $E(\gamma)=E(\gamma^0)$:

$$\begin{cases} \varepsilon(\gamma\gamma) = \varepsilon(\gamma)+\varepsilon(\gamma^0)-\varepsilon(\gamma)\varepsilon(\gamma^0) = 2\varepsilon(\gamma)-\varepsilon^2(\gamma) \\ E(\gamma\gamma) = E(\gamma)+E(\gamma^0)-E(\gamma)E(\gamma^0) = 2E(\gamma)-E^2(\gamma) \end{cases} \tag{5.5.3}$$

Using this nomenclature, the L0 efficiencies result as given in Tab. 5.5.4.

Using the study summed up in Tab. 5.5.4, the L0 efficiency is calculated for all mass hypotheses, using the data-corrected efficiencies ($E_\text{Data}(\text{L0})$) as nominal efficiency. As a systematic uncertainty for that approach, the relative difference between the cut-based effciciencies $\varepsilon(\gamma\gamma)$ and the efficiencies calibrated on $\eta\to\mu^+\mu^-\gamma$ simulation ($E_\text{MC}(\gamma\gamma)$) are assigned. This difference mainly comes from the very different p_T distribution the $\eta\to\mu^+\mu^-\gamma$ and ALPs photons have and the limited statistics on the $\eta\to\mu^+\mu^-\gamma$ simulation does not allow for fine enough binning. This is not critical since uncertainties are way smaller than the one arising from the nSPDHits correction.

5.5.5 Data-driven HLT1 efficiency

The cuts imposed for this selection can be seen at Tab. 5.2.3. When looking at them one should remember that the reconstruction used to obtain this quantities is not a complete one, so usually these quantities underestimate the true kinematic value they refer to. In order to illustrate this idea, a plot of the difference between the p_T^{L0} of the simulated signal photons as

Table 5.5.4: Detailed efficiencies for the L0 study.

Mass [GeV]	$\varepsilon(\gamma)$	$E_{\text{Data}}(\gamma)$	$E_{\text{MC}}(\gamma)$	$\varepsilon(\gamma\gamma)$	$E_{\text{Data}}(\gamma\gamma)$	$E_{\text{MC}}(\gamma\gamma)$
$B_s^0 \to \gamma\gamma$	82.63 ± 0.20	76.77 ± 0.27	76.8 ± 0.8	97.09 ± 0.09	94.6 ± 2.5	94.6 ± 2.5
5	82.80 ± 0.19	79.42 ± 0.28	79.1 ± 0.9	97.56 ± 0.08	95.8 ± 1.9	95.6 ± 2.0
6	83.32 ± 0.17	79.02 ± 0.25	78.8 ± 0.8	97.52 ± 0.07	95.6 ± 2.0	95.5 ± 2.0
7	84.28 ± 0.16	78.81 ± 0.23	78.6 ± 0.8	97.70 ± 0.06	95.5 ± 2.3	95.4 ± 2.3
8	84.89 ± 0.14	79.20 ± 0.22	79.0 ± 0.8	97.92 ± 0.06	95.7 ± 2.3	95.6 ± 2.4
9	85.41 ± 0.13	80.18 ± 0.21	80.0 ± 0.8	98.04 ± 0.05	96.1 ± 2.1	96.0 ± 2.1
10	86.20 ± 0.12	81.14 ± 0.20	81.0 ± 0.8	98.26 ± 0.05	96.4 ± 1.9	96.4 ± 1.9
11	87.20 ± 0.12	82.06 ± 0.20	81.9 ± 0.8	98.46 ± 0.04	96.8 ± 1.7	96.7 ± 1.8
13	88.19 ± 0.11	84.18 ± 0.21	83.9 ± 0.9	98.64 ± 0.04	97.5 ± 1.2	97.4 ± 1.3
15	88.91 ± 0.11	86.20 ± 0.23	85.7 ± 1.1	98.72 ± 0.04	98.1 ± 0.8	98.0 ± 0.8
17	88.73 ± 0.12	87.50 ± 0.24	86.9 ± 1.4	98.75 ± 0.04	98.4 ± 0.5	98.3 ± 0.6
19	88.79 ± 0.13	88.13 ± 0.28	87.2 ± 1.7	98.78 ± 0.05	98.6 ± 0.4	98.4 ± 0.6
20	89.39 ± 0.21	88.1 ± 0.4	87.6 ± 1.6	98.77 ± 0.08	98.59 ± 0.32	98.5 ± 0.5

seen by the 2×2 L0 clusters (used in the HLT1 selection) and the 3×3 offline ones ($p_{\text{T}}^{\text{Offline}}$), using the $\eta \to \mu^+\mu^-\gamma$ decay channel, can be seen in Fig. 5.5.6 in bins of the offline p_{T}. Figure 5.5.5 shows a comparison of the p_{T}^{L0} as seen by the L0 in data vs. simulation of the $\eta \to \mu^+\mu^-\gamma$ decay.

From the mentioned plots one can draw several conclusions: the calibration differs from data and simulation, as the simulation distributions are more shifted to the right; also, it was checked that for Simulated events for $p_{\text{T}} > 7600\,\text{MeV}$, the 99% of the p_{T}^{L0} values are saturated to the maximum value available: 6120 MeV. In Data it was checked that it is after 7400 MeV that 99 % of p_{T}^{L0} values are saturated. It was checked that this difference does not bias the efficiencies within a 0.5 % effect. These facts indicate the need to study how the determination of the p_{T}^{L0} differs for data and simulation up to $p_{\text{T}}^{\text{Offline}} < 7600\,\text{MeV}$, since above this threshold both will behave the same: set to the maximum value available. This fact also explains why starting from this threshold the plots on Fig. 5.5.6 no longer have a peak centered at a value near 0 MeV, rather they look like an offset regular p_{T} distribution.

The HLT1 efficiency is built from two components. A kinematic selection and also both offline photons must be matched to an L0 cluster:

$$\varepsilon_{\text{HLT1}} = \varepsilon_{\text{Matching L0 cluster}} \times \varepsilon_{\text{Kinematic cuts}} \tag{5.5.4}$$

The L0 cluster matching efficiency is taken from data and only statistical uncertainty is taken into account given the $\eta \to \mu^+\mu^-\gamma$ sample limited size. The matching efficiency is calculated in three p_{T} bins. Of course when a photon fires the L0 selection there is always an L0 cluster associated, the relevant efficiency is then associated when one of the photons does not fire it. With the definitions $\frac{N(\texttt{L0 TOS})_{p_{\text{T}}}}{N_{p_{\text{T}}}} = \varepsilon_{p_{\text{T}}}^{\texttt{L0}}$ as the L0 efficiency in a given p_{T} bin,

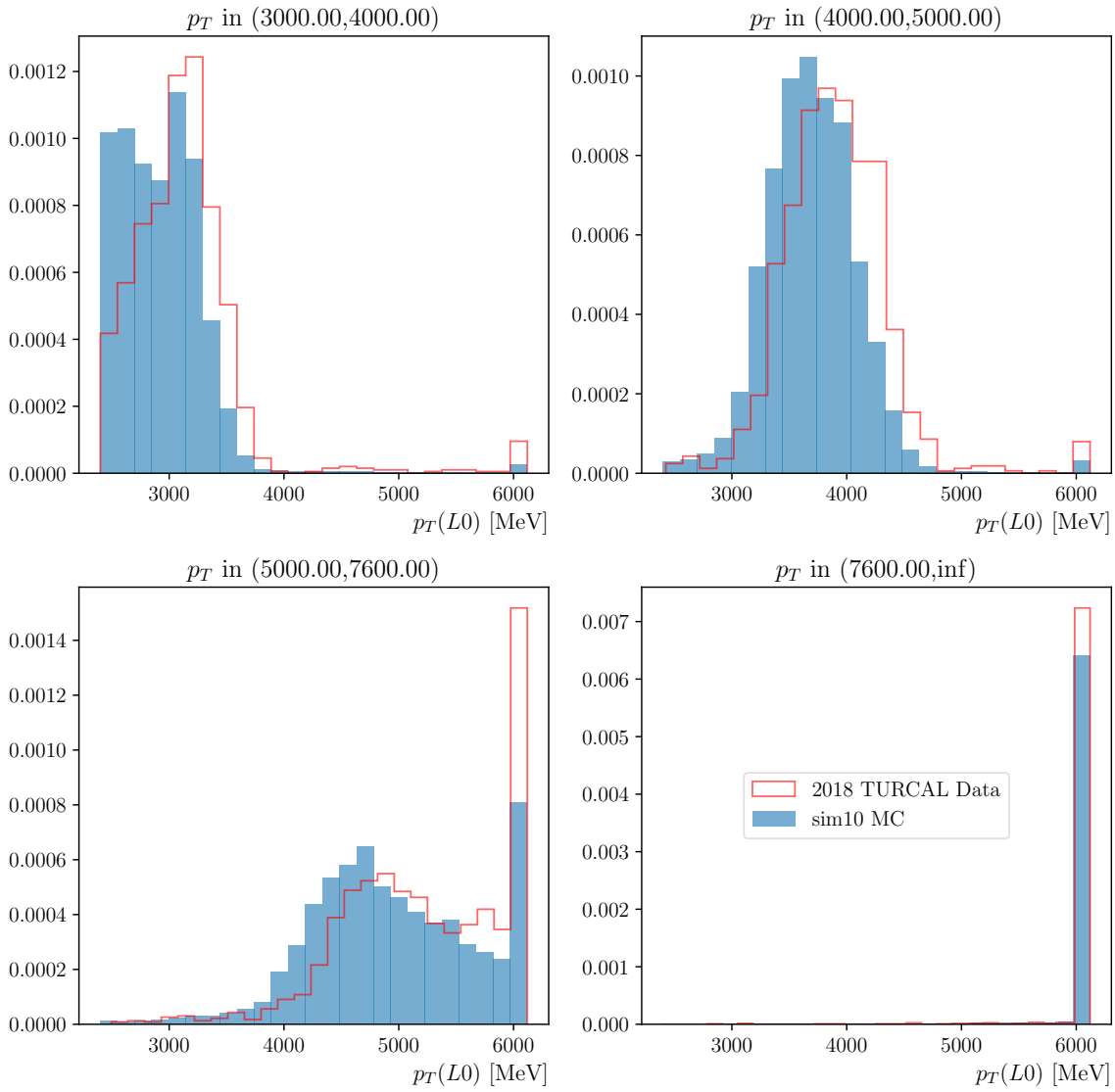

Figure 5.5.5: $p_{\mathrm{T}}^{\mathrm{L0}}$ (the p_{T} as seen by the HLT1) in bins of the offline p_{T} for sim10 simulation and TURCAL data from the $\eta \to \mu^+\mu^-\gamma$ decay channel.

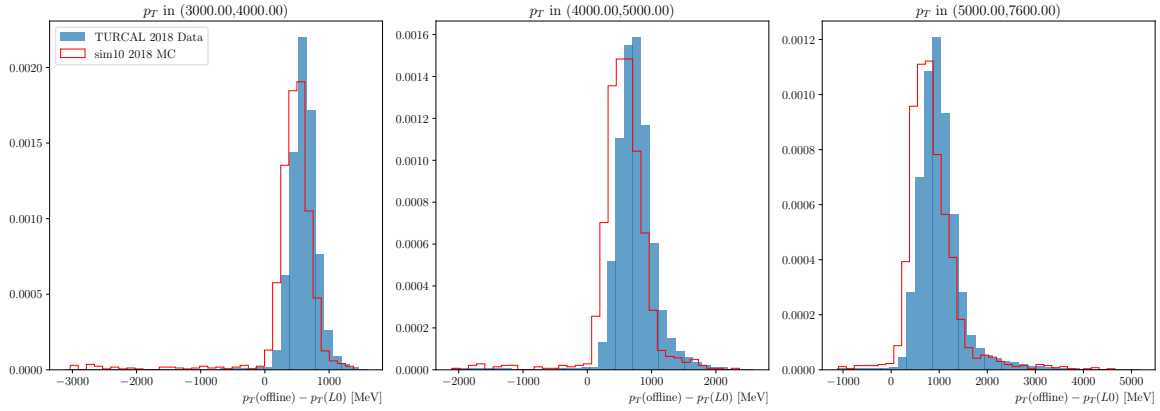

Figure 5.5.6: $p_{\mathrm{T}}(\mathtt{L0}) - p_{\mathrm{T}}(\mathrm{offline})$ distributions using $\eta \to \mu^+\mu^-\gamma$ decay.

$\varepsilon_{\mathrm{Matching!L0}}$ as the probability of finding an associated L0 cluster when L0 does not fire and $f_{p_{\mathrm{T}}} = \frac{N_{p_{\mathrm{T}}}}{N_{\mathrm{Total}}}$ as the fraction of signal in a given p_{T} bin, the global probability that both photons are associated to an L0 cluster is found as:

$$\varepsilon_{\mathrm{Matching\ L0\ cluster}} = \sum_{p_{\mathrm{T}}} \left(\varepsilon^{\mathtt{L0}}_{p_{\mathrm{T}}} + (1 - \varepsilon^{\mathtt{L0}}_{p_{\mathrm{T}}})\varepsilon_{\mathrm{Matching!L0}}\right) \times f_{p_{\mathrm{T}}} \tag{5.5.5}$$

The $\eta \to \mu^+\mu^-\gamma$ data sample is again used to calculate these efficiencies. A reweighting procedure is applied to the χ^2_{match} variable. The efficiency is then calculated with the thus reweighted $\eta \to \mu^+\mu^-\gamma$ MC and compared to that of ALPs MC. These values can be checked in Tab. 5.5.5. Within the limited statistics of the $\eta \to \mu^+\mu^-\gamma$ MC the method indeed yields compatible results between the two MC samples. Compared to the (reweighted) $\eta \to \mu^+\mu^-\gamma$ data sample, this calibration yields numbers that after applying Eq. 5.5.5 (i.e. integrating over p_{T}) only differ within the statistical uncertainties, for all masses $< 0.5\%$. It is concluded that the matching efficiency is well described for the purpose of this analysis.

The HLT1 kinematic efficiency is obtained from data by sampling the difference between the $p_{\mathrm{T}}^{\mathtt{L0}}$ and $p_{\mathrm{T}}^{\mathrm{Offline}}$ of the photon on $\eta \to \mu^+\mu^-\gamma$ in bins of $p_{\mathrm{T}}^{\mathrm{Offline}}$

$$\Delta p_{\mathrm{T}} = p_{\mathrm{T}}^{\mathtt{L0}} - p_{\mathrm{T}}^{\mathrm{Offline}}. \tag{5.5.6}$$

These are then transformed back into a proxy for $p_{\mathrm{T}}^{\mathtt{L0}}$ by applying

$$\Pi_{\mathrm{T}} = p_{\mathrm{T}}^{\mathrm{Offline}} - \Delta p_{\mathrm{T}}, \tag{5.5.7}$$

where $p_{\mathrm{T}}^{\mathrm{Offline}}$ is taken from the photons of the ALP $\to \gamma\gamma$ sample.

The same procedure can be applied to data and simulation from $\eta \to \mu^+\mu^-\gamma$. To prove the validity of this "smearing procedure", Fig. 5.5.7 shows both the distributions of $p_{\mathrm{T}}^{\mathtt{L0}}$ and

Table 5.5.5: $\varepsilon_{\text{Matching !L0}}$ efficiency for the photon candidates the relevant p_{T} bins. The $\eta \to \mu^+\mu^-\gamma$ data value (reweighted by χ^2_{match})is used to calculate the efficiency while the $\eta \to \mu^+\mu^-\gamma$ MC (reweighted by χ^2_{match}) and ALPs MC (exemplary only the 6 GeV ALP is displayed) are shown as means to cross check the method.

p_{T} bin [MeV]	$\varepsilon^{\text{rew}}_{\text{MC}}$	$\varepsilon_{\text{MC,ALPs}}$	$\varepsilon^{\text{rew}}(\text{Data})$
3000,3500	70 ± 9	73.9 ± 1.5	74.0 ± 0.5
3500,4000	15 ± 7	19.2 ± 2.3	23.1 ± 0.9
4000,5000	12 ± 4	8.0 ± 1.3	8.2 ± 0.4
5000,7000	14 ± 5	3.5 ± 0.8	4.10 ± 0.32
7000,9000	6 ± 4	1.6 ± 0.7	1.05 ± 0.22
9000,∞	4.7 ± 3.3	0.5 ± 0.4	0.20 ± 0.12

$\Pi^{\text{MC}}_{\text{T}}$. Some differences can be spotted in the lower p_{T} bin, where the cut imposed by the L0 selection is not perfectly described by the newly constructed $\Pi^{\text{MC}}_{\text{T}}$ variable. I can also be seen that the entirety of candidates have a saturated p^{L0}_{T} value of 6120 MeV whereas $\Pi^{\text{MC}}_{\text{T}}$ is not able to capture it. For the latter the solution is simple, since this fact applies also for data all the $\Pi^{\text{MC}}_{\text{T}}$ and $\Pi^{\text{Data}}_{\text{T}}$ values can be set with $p^{\text{Offline}}_{\text{T}} > 6120$ MeV to the 6120 MeV threshold without biasing the MC/data difference.

After demonstrating the feasibility of the method, the variables that make up the HLT1 selection can be constructed. Extra information is needed to transform our scalar transverse momentum values into 3-vectors, that is the (x, y, z) position of the ECAL clusters. The next step is to set $z_{ECAL} = 12500\,\text{mm}$, as it is used in the L0. These angles will be needed in order to re-calculate the HLT1 mass which follows the following approximation:

$$\begin{aligned} M^2(\gamma_0\gamma_1) &= 2E(\gamma_0)E(\gamma_1)(1-\cos\theta) \simeq E(\gamma_0)E(\gamma_1)\theta^2 \\ &= \frac{E_{\text{T}}(\gamma_0)D_0}{\sqrt{(x_0^2+y_0^2)}}\frac{E_{\text{T}}(\gamma_1)D_1}{\sqrt{(x_1^2+y_1^2)}}\theta^2 \\ &\simeq E_{\text{T}}(\gamma_0)E_{\text{T}}(\gamma_1)\frac{(x_0-x_1)^2+(y_0-y_1)^2}{\sqrt{(x_0^2+y_0^2)(x_1^2+y_1^2)}} \end{aligned} \tag{5.5.8}$$

where (x_i, y_i) is the position of the cluster on the calorimeter plane and D_i is the distance from the di-photon vertex to the ECAL. Since such vertexing is not possible, then, $D = D_0 = D_1$:

$$\theta^2 = \frac{(x_0-x_1)^2+(y_0-y_1)^2}{D^2} \tag{5.5.9}$$

which rounds up Eq. 5.5.8.

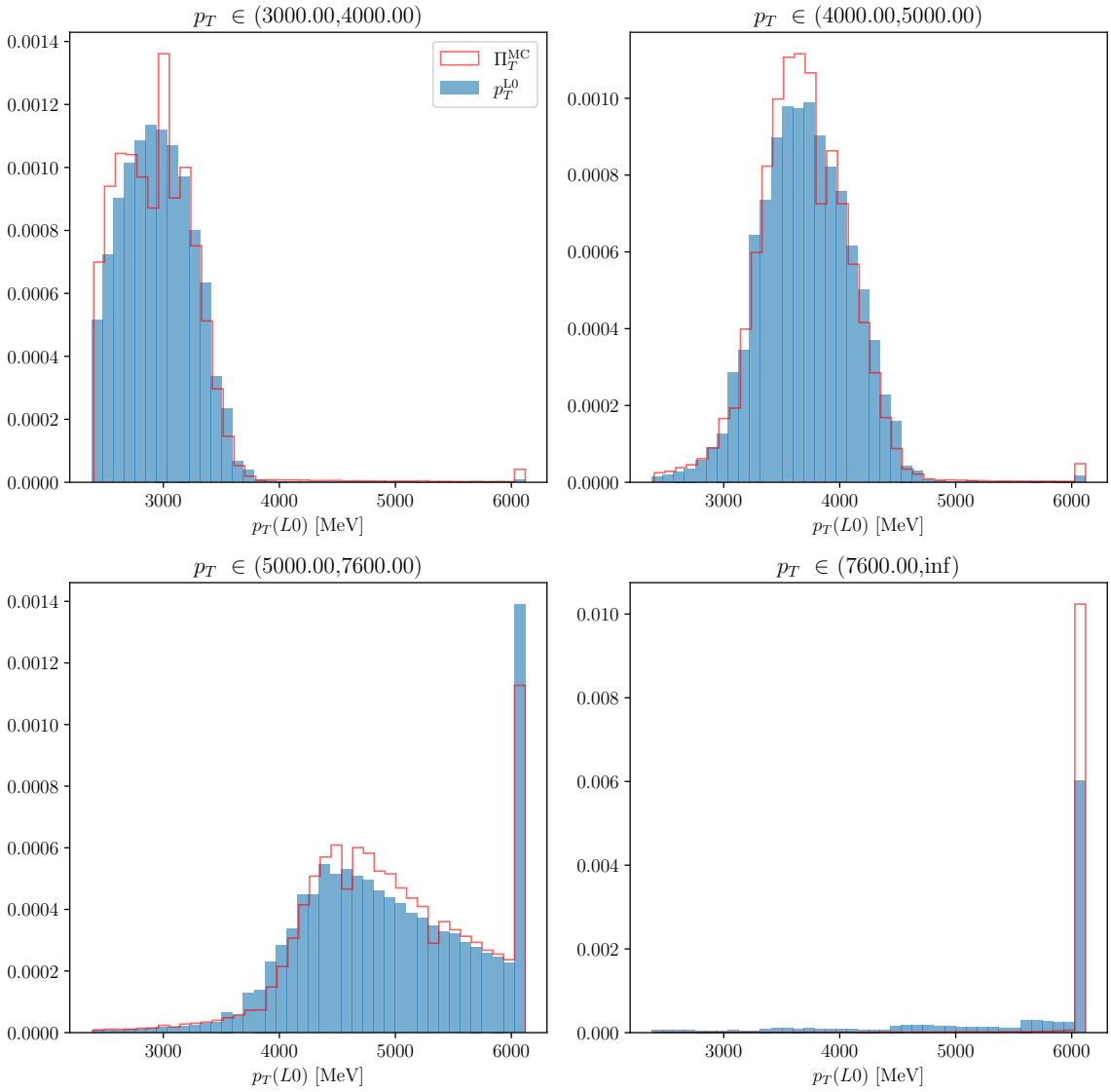

Figure 5.5.7: Comparison between p_T^{L0} and Π_T^{MC} as closure test using the $\eta \rightarrow \mu^+\mu^-\gamma$ decay channel in simulation.

It is only left to calculate the diphoton E_{T}: $\gamma\gamma(E_{\mathrm{T}})$ and for that it will be needed to do the vector sum $\vec{p} = \vec{p_0} + \vec{p_1}$ that will be calculated through the slopes t_x and t_y:

$$\begin{cases} t_x &= x_{\mathrm{ECAL}}/z_{\mathrm{ECAL}} \\ t_y &= y_{\mathrm{ECAL}}/z_{\mathrm{ECAL}} \end{cases} \tag{5.5.10}$$

that will allow to calculate p_x and p_y:

$$\begin{cases} p_x^{\mathrm{L0}} &= p_{\mathrm{T}}^{\mathrm{L0}} \sqrt{\frac{t_x^2}{t_x^2+t_y^2}} \\ p_y^{\mathrm{L0}} &= p_{\mathrm{T}}^{\mathrm{L0}} \sqrt{\frac{t_y^2}{t_x^2+t_y^2}} \end{cases} \tag{5.5.11}$$

Provided the (x,y) position of the L0 clusters is used, all ingredients are in place to prepare the HLT1 selection using three different p_{T} proxies: $p_{\mathrm{T}}^{\mathrm{L0}}$, $\Pi_{\mathrm{T}}^{\mathrm{MC}}$ and $\Pi_{\mathrm{T}}^{\mathrm{Data}}$. Table 5.5.6 compares the selection efficiencies from ALP $\to \gamma\gamma$ simulation once determined from cutting on the HLT1 trigger line and once by rebuilding the selection based on $p_{\mathrm{T}}^{\mathrm{L0}}$ with the above discussed considerations to build the variables. Perfect agreement within the uncertainties is achieved, demonstrating that the above assumptions are valid. The small discrepancy is mostly driven by a small chance of getting the wrong cluster when matching them to the offline one. It was checked that the effect of the reconstruction on the angles can be considered negligible with respect to the effect on the energy resolution by comparing the gen-level directions with the L0 and offline reconstruction, obtaining very similar efficiencies.

Table 5.5.6: Comparison between the HLT1 cut-based efficiency and the efficiency based on the reconstruction by hand on ALP simulation.

Mass	ε(TOS)	$\varepsilon(p_{\mathrm{T}}^{\mathrm{L0}})$
$\mathrm{B_s^0} \to \gamma\gamma$	32.77 ± 0.25	32.69 ± 0.25
5	35.62 ± 0.24	35.45 ± 0.24
6	41.15 ± 0.22	41.06 ± 0.22
7	30.31 ± 0.20	29.92 ± 0.20
8	22.28 ± 0.17	22.10 ± 0.17
9	18.31 ± 0.15	18.22 ± 0.15
10	18.22 ± 0.14	18.17 ± 0.14
11	19.97 ± 0.14	19.93 ± 0.14
13	20.94 ± 0.14	20.84 ± 0.14
15	17.70 ± 0.14	17.58 ± 0.14
17	12.31 ± 0.13	12.18 ± 0.13
19	8.16 ± 0.12	8.06 ± 0.12
20	5.75 ± 0.16	5.65 ± 0.16

Table 5.5.7: The first columns serve as crosscheck that the smearing method yields compatible efficiencies, when comparing with the baseline cut based HLT1 efficiency. The nominal value is taken from data under the label $\varepsilon(\Pi_T^{\text{Data}})$. The last column shows the relative systematic uncertainty of the method.

Mass [GeV]	$\varepsilon_{\text{Matching L0 cluster}}$	$\varepsilon(\text{TOS})$	$\varepsilon(\Pi_T^{\text{MC}})$	$\varepsilon(\Pi_T^{\text{Data}})$	$\frac{\varepsilon(\Pi_T^{\text{MC}})-\varepsilon(\text{TOS})}{\varepsilon(\text{TOS})}$
$B_s^0 \to \gamma\gamma$	85.82 ± 0.06	32.77 ± 0.25	32.02 ± 0.24	27.11 ± 0.23	-2.3 ± 1.0
5	84.29 ± 0.09	34.90 ± 0.22	33.99 ± 0.22	28.99 ± 0.21	-2.6 ± 0.9
6	85.26 ± 0.08	40.58 ± 0.20	40.02 ± 0.20	35.89 ± 0.20	-1.4 ± 0.7
7	86.02 ± 0.08	30.34 ± 0.18	31.86 ± 0.18	31.52 ± 0.18	5.0 ± 0.9
8	86.51 ± 0.07	22.44 ± 0.15	22.69 ± 0.15	21.34 ± 0.15	1.1 ± 1.0
9	86.67 ± 0.08	18.61 ± 0.13	18.63 ± 0.13	17.14 ± 0.13	0.1 ± 1.0
10	86.74 ± 0.09	18.29 ± 0.13	17.87 ± 0.13	16.00 ± 0.12	-2.3 ± 1.0
11	86.79 ± 0.10	20.07 ± 0.13	19.21 ± 0.13	17.32 ± 0.12	-4.3 ± 0.9
13	86.63 ± 0.11	21.17 ± 0.13	20.73 ± 0.13	19.80 ± 0.13	-2.1 ± 0.8
15	86.09 ± 0.13	18.17 ± 0.13	18.19 ± 0.13	17.85 ± 0.13	0.1 ± 1.0
17	85.09 ± 0.15	13.08 ± 0.12	13.21 ± 0.12	13.19 ± 0.12	0.9 ± 1.2
19	84.58 ± 0.16	8.79 ± 0.11	8.91 ± 0.11	8.93 ± 0.11	1.3 ± 1.7
20	85.46 ± 0.15	6.33 ± 0.15	6.38 ± 0.16	6.33 ± 0.16	0.9 ± 3.5

The systematic uncertainty for the efficiency corresponding to the kinematic part of the HLT1 selection is determined from the relative difference between the efficiency based on the Π_T variables (obtained from the $\eta \to \mu^+\mu^-\gamma$ simulation) and the HLT1 baseline efficiency, that obtained from the emulation of the trigger directly. The matching efficiency, the efficiency from emulating the trigger directly, the efficiencies calibrated from $\eta \to \mu^+\mu^-\gamma$ simulation and data and the systematic uncertainty are listed in Tab. 5.5.7.

5.5.6 HLT2 efficiency

The HLT2 efficiency is obtained through a data driven method using 2018 data from the $\eta \to \mu^+\mu^-\gamma$ control channel. The input variables involved in the classifier are γ IsNotH and γ ShowerShape together with $\gamma\gamma$ p_T. These are chosen to distinguish against hadronic background and misidentified π^0 meson decays. In order to obtain the efficiencies from data a 2D histogram involving γ IsNotH and γ ShowerShape is used. This map keeps the correlations betweeen the PID variables. The sweighted $\eta \to \mu^+\mu^-\gamma$ data candidates are then reweighted with respect to ALP photon χ^2_{match} variable, which shows to be more relevant variable when comparing final states which charged particles and purely neutral modes. By following this procedure the effect of having extra tracks on the $\eta \to \mu^+\mu^-\gamma$ decay is mitigated. From the sweighted and reweighted $\eta \to \mu^+\mu^-\gamma$ sample new IsNotH and ShowerShape values are

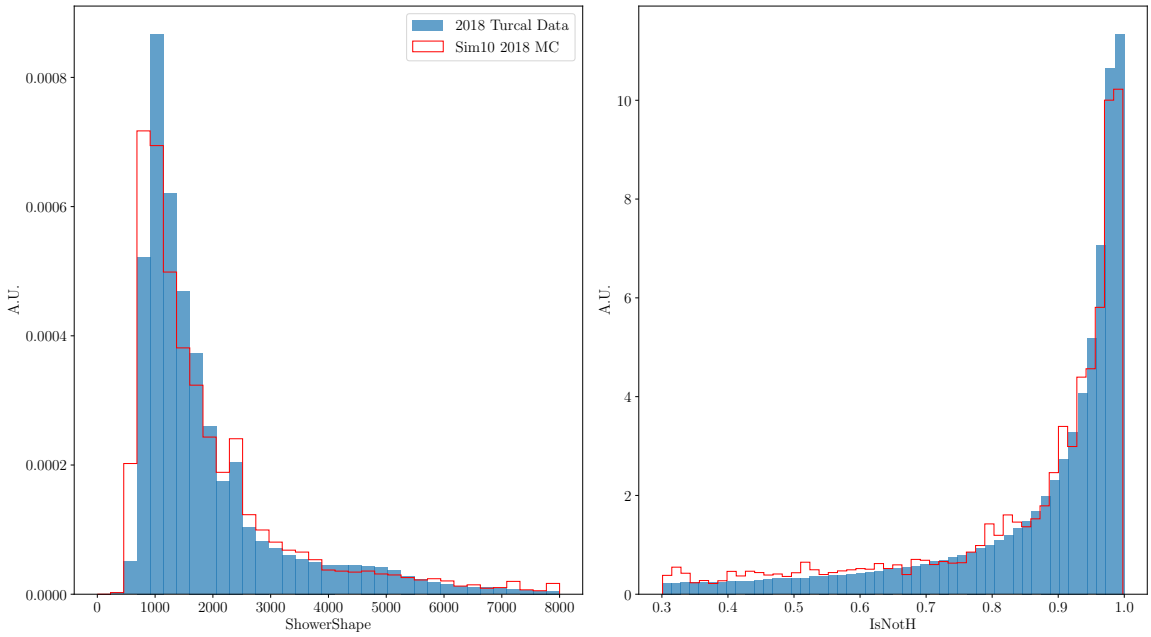

Figure 5.5.8: MC/data comparison of relevant variables for the HLT2 selection using the $\eta \to \mu^+\mu^-\gamma$ control channel.

drawn, following the 2D histogram. These are then used to feed the classifier. This way the HLT2 efficiency can be obtained from data.

A systematic uncertainty is calculated from investigating the efficiency obtained directly from simulation with that obtained from applying the method to reweighted $\eta \to \mu^+\mu^-\gamma$ MC candidates. The distributions of the input features of $\eta \to \mu^+\mu^-\gamma$ data and simulation are shown in Fig. 5.5.8. The diphoton p_T is not modified since no sizeable differences are expected in the profile of the transverse momentum.

Comparing these numbers with the plain MC efficiency in Tab. 5.5.8, agreement is found within a few percent. Since this variation is an instrinsic effect of the reweighting and sampling (note that the uncertainties from the reweighting itself were not determined, which would require a costly bootstrapping procedure), it was decided to assess as systematic a combination of the efficiency difference and the uncertainty of these efficiencies (Eq. 5.5.12), in order to describe the accuracy of the method, which should be stable through the whole mass range. In Figs. [A.2.2,A.2.11] the `IsNotH` and `ShowerShape` distributions as seen in the ALPs MC is overlayed on top of the counterpart as probed from the $\eta \to \mu^+\mu^-\gamma$ MC and data candidates.

$$s_{\text{HLT2}} = \left|\varepsilon_{\text{MC}} - \varepsilon_{\text{MC-resampled}}\right| \tag{5.5.12}$$

Table 5.5.8: Comparing plain MC HLT2 efficiency with that after rescaling variables with data information.

Mass	ε_{MC}	$\varepsilon(\text{MC}^{\text{calib}})$	$\varepsilon(\text{Data}^{\text{calib}})$
$B_s^0 \to \gamma\gamma$	76.9 ± 0.4	76.1 ± 0.4	81.5 ± 0.4
5	76.8 ± 0.4	74.0 ± 0.4	79.6 ± 0.4
6	78.90 ± 0.30	75.46 ± 0.32	82.10 ± 0.28
7	82.04 ± 0.31	79.34 ± 0.33	84.84 ± 0.29
8	83.09 ± 0.33	80.73 ± 0.34	86.60 ± 0.30
9	82.62 ± 0.34	80.7 ± 0.4	85.83 ± 0.31
10	82.12 ± 0.33	79.53 ± 0.35	84.76 ± 0.31
11	80.49 ± 0.32	76.62 ± 0.34	82.95 ± 0.30
13	80.58 ± 0.31	77.02 ± 0.33	82.46 ± 0.30
15	81.07 ± 0.35	77.9 ± 0.4	84.22 ± 0.32
17	81.2 ± 0.4	79.5 ± 0.4	85.0 ± 0.4
19	81.0 ± 0.6	82.0 ± 0.6	86.5 ± 0.5
20	80.7 ± 1.2	89.6 ± 0.9	84.4 ± 1.1

5.5.7 Isolation variables classifier

The efficiency of the final selection BDT is also taken directly from cutting and counting on signal simulation. In order to address the classifier systematic uncertainty a classifier for the $B^0 \to K^{*0}\gamma$ channel was used. In this exercise an xgboost classifier is trained using simulation and apply it on MC and in sWeighted data. For this training the same cone isolation variables that were used for our ALP $\to \gamma\gamma$ exercise are chosen, but applied to the photon only. Then the efficiencies are computed, which can be checked in Tab. 5.5.9 on both samples and draw a systematic for the difference in the numeric values where the efficiency matches the efficiency of the BDT cut on the ALP samples. It is used as systematic the deviation on these efficiencies. This yields a very small uncertainty of $< 1\%$ ($1\,\%$ is assigned as a conservative but totally negligible systematic uncertainty), as the cone isolation variables are well described in simulation. To sustain this statement Fig. 5.5.9 shows that indeed the cone variables agree very well between data and simulation.

5.5.8 Signal model systematics

Different production mechanisms can produce very different diphoton kinematics and isolations. The bumphunt is performed with the MadGraph model described in the introduction. To facilitate recasts the model will be provided in the appendix of the paper. Since the search is thus rather model dependent, no extra model systematic will be associated.

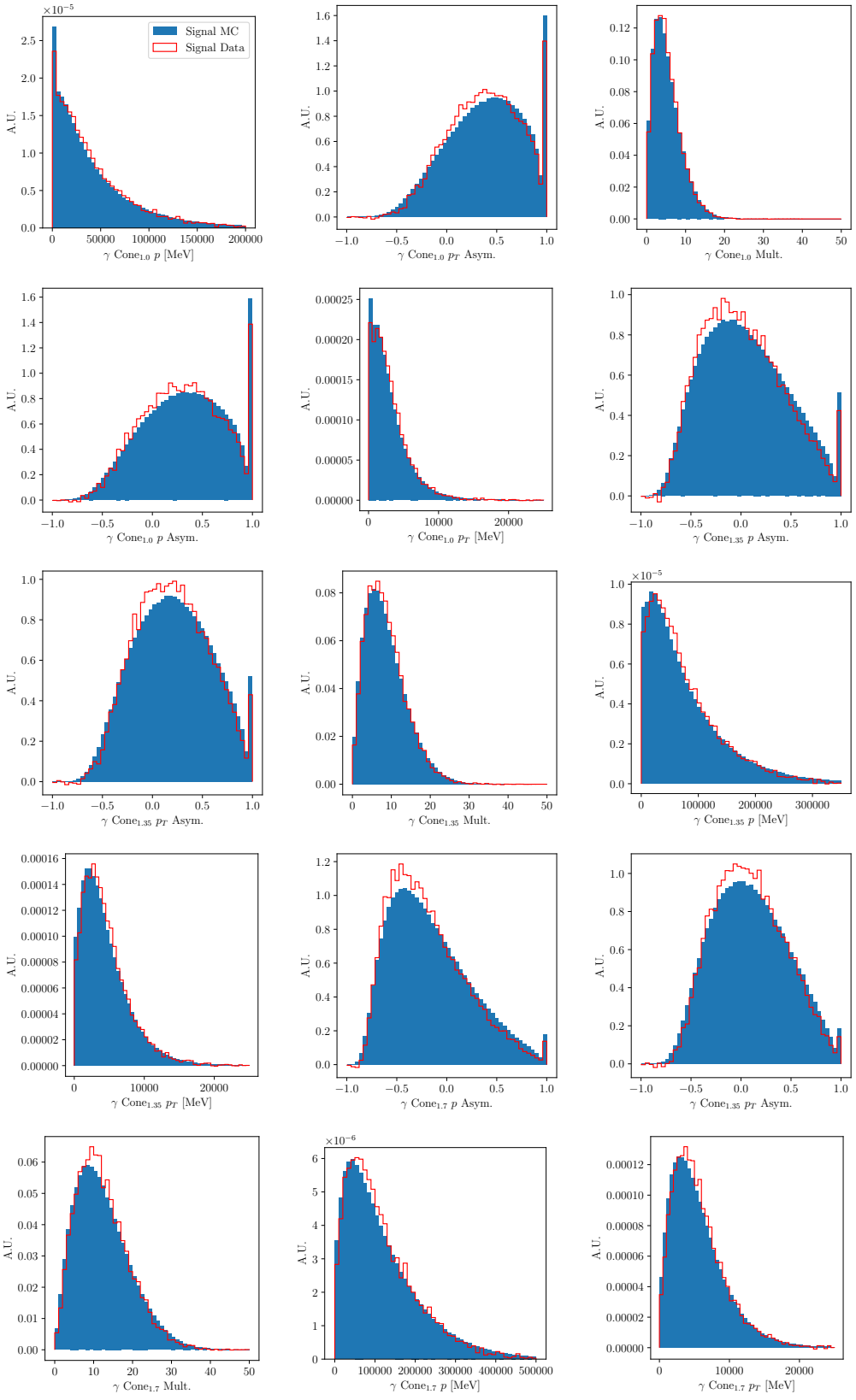

Figure 5.5.9: Comparison between MC and data isolation variables of the $B^0 \to K^{*0}\gamma$ decay.

Table 5.5.9: Comparison between MC and sWeighted $B^0 \to K^{*0}\gamma$ data xgboost classifier efficiencies.

Cut	ε(MC)	ε(Data)
0.10	98.71 ± 0.01	100.25 ± 0.25
0.20	95.74 ± 0.01	97.97 ± 0.32
0.30	91.68 ± 0.01	94.19 ± 0.36
0.40	86.84 ± 0.01	89.43 ± 0.39
0.50	81.35 ± 0.01	83.66 ± 0.40
0.60	75.05 ± 0.01	76.82 ± 0.40
0.70	67.39 ± 0.02	68.74 ± 0.40
0.80	57.62 ± 0.04	58.62 ± 0.39
0.90	42.95 ± 0.07	42.87 ± 0.35

5.5.9 Efficiencies for $B_s^0 \to \phi\gamma$

The efficiencies for the normalisation mode for $B_s^0 \to \gamma\gamma$, $B_s^0 \to \phi\gamma$, are directly calculated from simulation. For the L0 trigger efficiency a correction factor for the nSPDHits and SumEtPrev cut is determined in the same as for the $B_s^0 \to \gamma\gamma$ efficiencies as discussed in Sec. 5.5.4. The other efficiencies are expected to be well reproduced in simulation and thus the systematic uncertainties on the $B_s^0 \to \gamma\gamma$ efficiencies will fully dominate the uncertainty on the normalisation factor of $B^0_{(s)} \to \gamma\gamma$. The full breakdown of the efficiencies is listed in Tab. 5.5.10.

Table 5.5.10: Efficiencies $\varepsilon(\mathrm{B}_\mathrm{s}^0 \to \phi\gamma)$ and yield of the normalisation decay mode $\mathrm{B}_\mathrm{s}^0 \to \phi\gamma$, used to determine the normalisation factor for $B^0_{(s)} \to \gamma\gamma$. The stripping filter is performed with very high stats and thus the uncertainty is negligible.

Selection	$\varepsilon(\mathrm{B}_\mathrm{s}^0 \to \phi\gamma)\,[\%]$
Generation	13.200 ± 0.018
Stripping filter	7.945
Truth matching	91.889 ± 0.015
PID	93.931 ± 0.014
L0	71.408 ± 0.025
HLT1	85.950 ± 0.022
HLT2	89.279 ± 0.021
BDT	79.351 ± 0.030
Total	0.3936 ± 0.0006
Yield (not %)	$(5.80 \pm 0.07) \times 10^4$

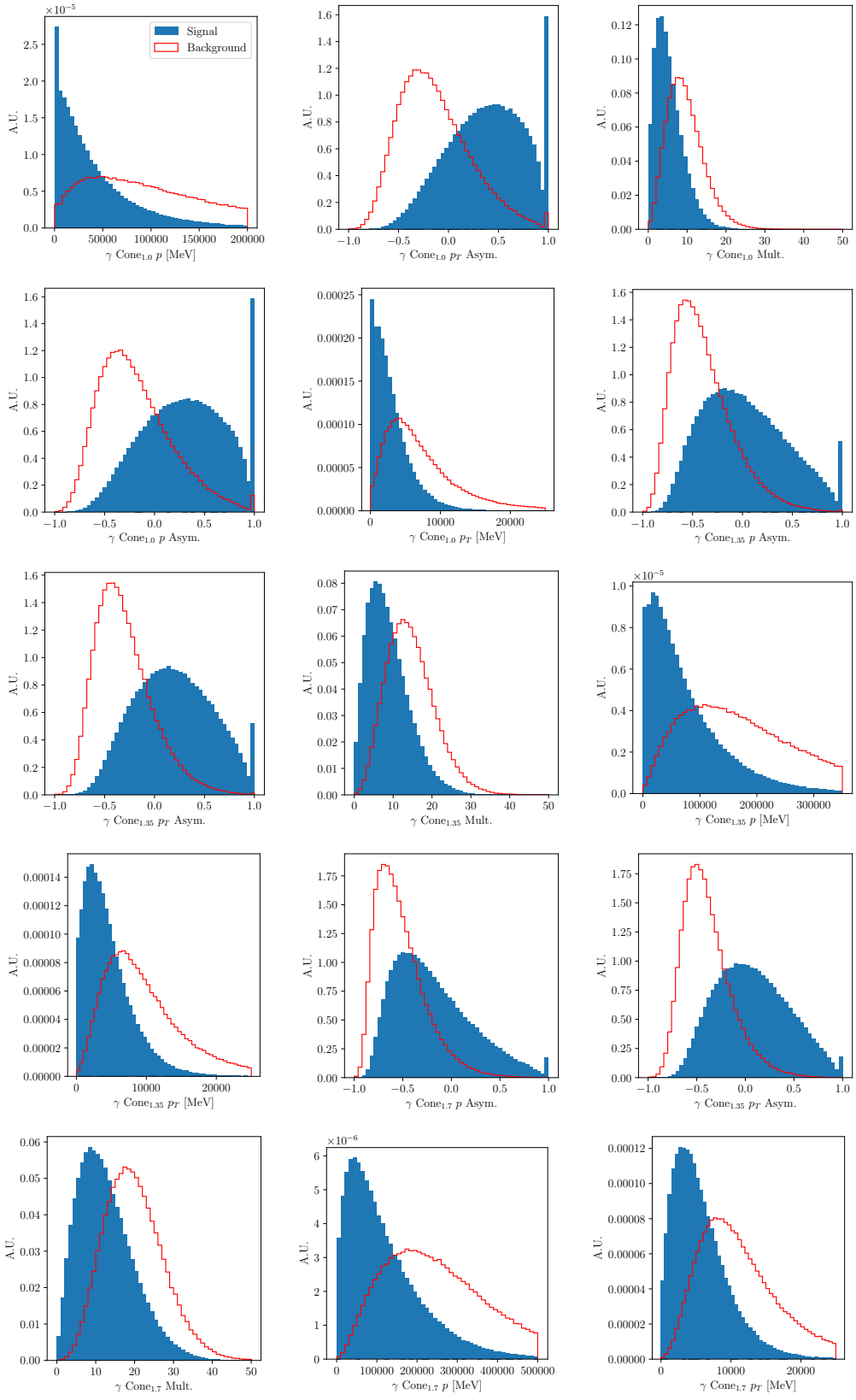

Figure 5.5.10: Input variables of $B^0 \to K^{*0}\gamma$ (signal simulation versus data side bands) that are used in classifier to calculate systematic.

5.5.10 Efficiency summary

The full summary of the selection efficiencies is found in Tab. C.5.1, already including the systematic uncertainties. The trends are illustrated in Fig. 5.5.11.

The efficiencies at 5 GeV and at 20 GeV are significantly lower than the one examined at other energies due to the mass range of $[4.8, 20]$ GeV imposed in the stripping selection. Apart from that, several trends in the efficiencies can be observed that roughly cancel each other: The generator efficiency (detector acceptance) shrinks slightly with growing mass due to the more central production of heavier masses. The reconstruction efficiency (in this summary including the p_T cut at 3 GeV) grows with mass (due to the harder p_T spectra of higher mass particles). This however increases also the event multiplicity and thus the global event cut has slightly lower efficiencies with increasing mass. The PID efficiency shrinks with mass, but is relatively stable. The saturation veto leads to decreasing efficiencies especially at masses beyond 11 GeV. Note that also here the fact that the two most extreme simulation samples are not fully inside the stripping range, increases the effect of the veto, because saturated cells lead to strong tails towards lower diphoton masses. The L0 kinematic efficiencies increase slightly with higher ALP masses (again the p_T spectra become harder). The HLT1 efficiency shows two peaks due to the overlap of the two trigger lines and then a decrease where the diphoton mass cut hits (significantly after 11 GeV because the mass in the HLT1 is made of the L0 clusters). The HLT2 efficiency mainly relies on p_T spectra and photon ID and thus shows an increasing trend as discussed before. Finally, the ultimate selection BDT has a non-monotonous trend, possibly following the unevenly distributed events in the background sample (see Fig. 5.7.1), which makes the BDT concentrate more on these events.

In the implementation of the bump hunt, the efficiencies and uncertainties for the calculation of the cross section at a given mass point are estimated through a linear interpolation between the two closest simulated points as shown in Fig. 5.5.11. The samples at 5 GeV and

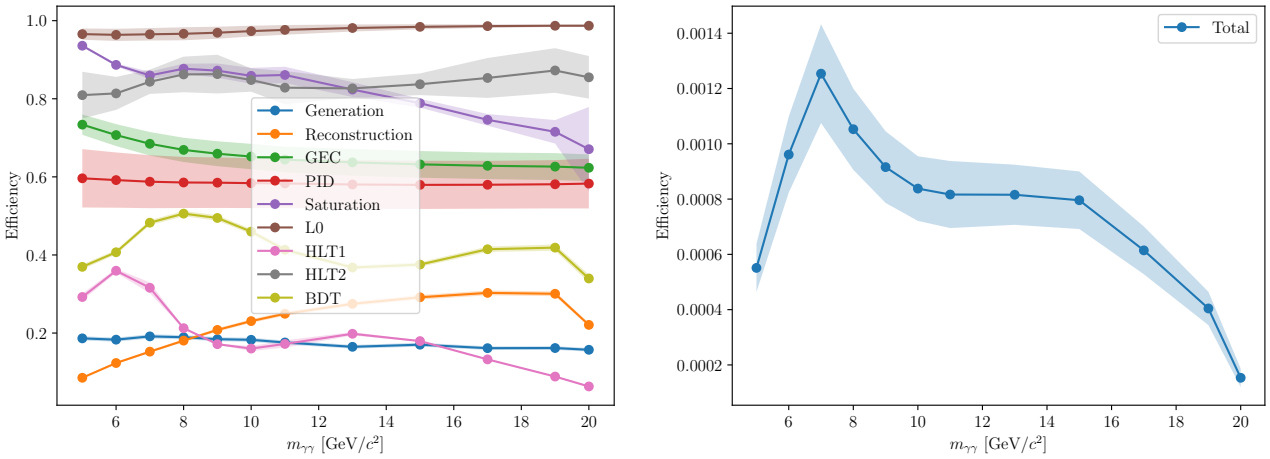

Figure 5.5.11: Trend of the ALPs efficiencies over the [5,20] GeV mass range.

the one of 20 GeV are ignored for the reasons mentioned above and instead the interpolation from the two next reliable points is used to estimate the efficiencies below 6 GeV and above 19 GeV. The efficiency for the $B^0_{(s)} \to \gamma\gamma$ signal is directly estimated from the simulated $\mathrm{B}^0_\mathrm{s} \to \gamma\gamma$ sample with the same procedures as the ALP signal. Its efficiencies are significantly lower as the isolation properties behave differently (see App. **??**). From that the normalisation constants transforming the signal yields into cross sections and/or branching fractions is determined through the formulae (assuming $f_u = f_d$)

$$\begin{aligned} N_{\mathrm{ALP}} &= \sigma(pp \to \mathrm{ALP}) \times \mathscr{B}(\mathrm{ALP} \to \gamma\gamma) \times \mathscr{L} \times \varepsilon_{\mathrm{ALP}\to\gamma\gamma}, \\ N_{\mathrm{ALP}} &\equiv \sigma(pp \to \mathrm{ALP}) \times \mathscr{B}(\mathrm{ALP} \to \gamma\gamma)/\alpha_{\mathrm{ALP}}, \end{aligned} \tag{5.5.13}$$

$$\begin{aligned} N_{B^0_{(s)}\to\gamma\gamma} &= \mathscr{B}(B^0_{(s)} \to \gamma\gamma) \times \mathscr{L} \times \sigma(pp \to \mathrm{b\bar{b}}X) \times 2 \times \frac{f_{d,s}}{f_d} \times f_u \times \varepsilon_{\mathrm{B}^0_\mathrm{s}\to\gamma\gamma}, \\ N_{B^0_{(s)}\to\gamma\gamma} &\equiv \mathscr{B}(B^0_{(s)} \to \gamma\gamma)/\alpha_{d,s} \end{aligned} \tag{5.5.14}$$

The luminosity ($\mathscr{L}$) is well known, as discussed in Sec. 5.1.1. The production cross section of b-quarks has been measured at 13 TeV to be

$$\sigma(pp \to \mathrm{b\bar{b}}X) = (495 \pm 2 \pm 52)\,\mu\mathrm{b}\ [114]$$

and the b-hadronisation fractions are measured to be

$$\frac{f_s}{f_d} = 0.2539 \pm 0.0079\ [115] \text{ and}$$
$$f_u = 0.405 \pm 0.006\ [116].$$

Thus the normalisation factors (single event sensitivities) result to

$$\alpha_s = (1.5 \pm 0.4) \times 10^{-8}, \text{ and}$$
$$\alpha_d = (3.9 \pm 0.9) \times 10^{-9},$$

combining statistical and systematic uncertainties.

Note that f_u has been measured with Z decays at LEP and the assumption was made that the same number would hold for pp collisions at high energy. However, it has been shown that there is a relevant dependence on the b-hadron kinematics for these fractions and thus an extra uncertainty has to be considered to this assumption. A generous ad-hoc uncertainty of 10 % is assigned. A more accurate normalisation is performed using $\mathrm{B}^0_\mathrm{s} \to \phi\gamma$ with the formula

$$N_{\mathrm{B}^0_\mathrm{s}\to\gamma\gamma} = \mathscr{B}(\mathrm{B}^0_\mathrm{s} \to \gamma\gamma) \times \frac{N_{\mathrm{B}^0_\mathrm{s}\to\phi\gamma}}{\mathscr{B}(\mathrm{B}^0_\mathrm{s} \to \phi\gamma)} \frac{\varepsilon_{\mathrm{B}^0_\mathrm{s}\to\gamma\gamma}}{\varepsilon_{\mathrm{B}^0_\mathrm{s}\to\phi\gamma}} \equiv \mathscr{B}(\mathrm{B}^0_\mathrm{s} \to \gamma\gamma)/\alpha, \tag{5.5.15}$$

Table 5.5.11: ALPs efficiencies (in %) calculated sequentially . The systematic uncertainties are already included. It is important to note that the Reconstruction column covers the number of events that are reconstructed and that pass the kinematic cuts with respect to the generator level.

Mass [GeV]	Generation	Reconstruction	GEC	PID	Saturation	L0	HLT1	HLT2	BDT	Total
$B_s^0 \to \gamma\gamma$	17.0 ± 2.5	7.34 ± 0.15	63.9 ± 0.6	58 ± 6	97.69 ± 0.06	95.7 ± 1.8	26.8 ± 0.4	82 ± 5	33.0 ± 0.6	0.031 ± 0.006
5.000	18.6 ± 0.5	8.55 ± 0.17	73.4 ± 2.6	60 ± 7	93.59 ± 0.19	96.6 ± 1.4	29.0 ± 0.4	82 ± 5	37.3 ± 0.6	0.056 ± 0.008
6.000	18.3 ± 0.5	12.33 ± 0.25	70.7 ± 2.8	59 ± 7	88.68 ± 0.35	96.4 ± 1.6	35.84 ± 0.30	81.6 ± 2.7	41.7 ± 0.6	0.098 ± 0.013
7.000	19.2 ± 0.5	15.24 ± 0.31	68.5 ± 3.0	59 ± 7	86.0 ± 1.0	96.5 ± 1.6	31.6 ± 1.5	81.6 ± 0.6	49.9 ± 0.7	0.125 ± 0.017
8.000	18.9 ± 0.5	18.1 ± 0.4	66.9 ± 3.1	59 ± 7	87.7 ± 1.3	96.6 ± 1.7	21.38 ± 0.33	81.6 ± 1.5	52.7 ± 0.7	0.104 ± 0.014
9.000	18.4 ± 0.5	20.8 ± 0.4	65.9 ± 3.2	59 ± 6	87.2 ± 1.8	96.9 ± 1.5	17.21 ± 0.26	81.6 ± 1.1	52.0 ± 0.7	0.091 ± 0.012
10.000	18.3 ± 0.5	23.1 ± 0.5	65.2 ± 3.3	58 ± 6	85.9 ± 2.0	97.3 ± 1.4	16.0 ± 0.4	81.6 ± 0.6	48.3 ± 0.7	0.085 ± 0.011
11.000	17.6 ± 0.5	24.9 ± 0.5	64.4 ± 3.3	58 ± 6	86.1 ± 2.1	97.6 ± 1.2	17.3 ± 0.7	81.6 ± 1.2	43.6 ± 0.6	0.085 ± 0.012
13.000	16.5 ± 0.5	27.5 ± 0.6	63.7 ± 3.4	58 ± 6	82.4 ± 1.9	98.1 ± 0.9	19.8 ± 0.4	81.6 ± 1.1	39.7 ± 0.6	0.087 ± 0.011
15.000	17.0 ± 0.5	29.2 ± 0.6	63.2 ± 3.4	58 ± 6	78.9 ± 1.1	98.4 ± 0.7	17.91 ± 0.13	81.6 ± 0.6	41.2 ± 0.6	0.085 ± 0.011
17.000	16.1 ± 0.5	30.3 ± 0.6	62.8 ± 3.4	58 ± 6	74.6 ± 1.5	98.6 ± 0.5	13.23 ± 0.17	81.6 ± 0.5	45.4 ± 0.7	0.064 ± 0.008
19.000	16.2 ± 0.5	30.1 ± 0.6	62.6 ± 3.5	58 ± 6	71.5 ± 3.0	98.70 ± 0.31	8.91 ± 0.14	81.6 ± 0.7	46.3 ± 0.8	0.042 ± 0.006
20.000	15.7 ± 0.5	22.1 ± 0.4	62.3 ± 3.4	58 ± 6	67 ± 11	98.70 ± 0.31	6.36 ± 0.16	81.6 ± 0.9	37.5 ± 1.0	0.0163 ± 0.0034

knowing the $B_s^0 \to \phi\gamma$ branching fraction as [93]

$$\mathscr{B}(B_s^0 \to \phi\gamma, \phi \to K^+K^-) = (3.4 \pm 0.4) \times 10^{-5} \times (49.1 \pm 0.5)\,\%.$$

Together with the numbers listed in Tab. 5.5.10 from Sec. 5.5.9, this results in the normalisation factors

$$\alpha_s = (2.3 \pm 0.5) \times 10^{-8}, \text{ and}$$
$$\alpha_d = (5.8 \pm 1.3) \times 10^{-9},$$

roughly compatible, but more reliable than the numbers obtained with using the cross sections as input. Thus the latter numbers are used for normalising the $B^0_{(s)} \to \gamma\gamma$ decays.

5.5.11 Offline momentum and mass resolution

A complete review of the performance of the ECAL during the Run 2 can be checked in [117].

Nevertheless, this study was not updated to 2018 conditions where the ECAL parametrisation was improved for data, resulting in a mismatch between the two calibrations.

Furthermore, on top of the regular offline calibration a common-use post calibration tool was applied: `CaloPostCalib`. This tool implements some corrections to the energy calibration using as handle the $B^0 \to K^{*0}\gamma$ decay and adjusting it to minimize the energy resolution and to adjust the central value of the peak to be as close as possible to the PDG mass value of the B^0. In order to have a quantitative measure of how different is the energy resolution between 2018 Data and Simulation the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ decay channels are used. To both the `CaloPostCalib` tool is applied.

In order to establish this comparison the ratio of the the widths of the mass peak will be compared: $R_\sigma = \frac{\sigma_{\text{Data}}}{\sigma_{\text{MC}}}$. The results to both channels can be checked in Tab. 5.5.12

Decay	R_σ	$m_{\text{PDG}} - \mu_{\text{Data}}$ [MeV]	$m_{\text{PDG}} - \mu_{\text{MC}}$ [MeV]
$B^0 \to K^{*0}\gamma$	92.73 %	5.20	21.46
$B_s^0 \to \phi\gamma$	91.11 %	1.58	15.84

Table 5.5.12: Comparison of σ and μ parameters in the $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi\gamma$ channels. R_σ is the ratio between the σ parameter of the data fit and the fit to simulation.

Both channels show a discrepancy ($\mathcal{O}(9\%)$) in the σ and μ resolution, which is assigned as a correction factor to the ALP masses and resolutions in the bump hunt. In absolute value the measured center is closer to the PDG value in data than it is on MC. The p_T cut systematic associated to this discrepancy is found to be negligible when comparing it with other more sizeable differences.

To show how this affects to the ALPs, a plot of different simulated masses with the corresponding Double Crystal ball is shown in Fig. A.3.1. From the results in Tab. 5.5.13 a number of conclusions can be drawn. Firstly that the mean of the Gaussian in the reconstructed distribution, μ is such that $\mu = 1.01 \times m_{\text{True}}$, so slightly shifted with respect to the generated one.

A linear trend is visible between the mass resolution, $\sigma(M)$, and the generated mass, M. Knowing that the resolution follows this expression:

$$\sigma(E) = \sqrt{a^2 E + b^2 E^2 + c^2} \tag{5.5.16}$$

then mass resolution will follow the following expression, following uncorrelated propagation:

$$\begin{aligned} \sigma^2(M) = a^2 & \left[E(\gamma_0) + E(\gamma_1)\right](1 - \cos\theta) \\ & + b^2 \, E(\gamma_0)E(\gamma_1)(1 - \cos\theta) \\ & + c^2 \, \frac{E^2(\gamma_0) + E^2(\gamma_1)}{E(\gamma_0)E(\gamma_1)}(1 - \cos\theta) \end{aligned} \tag{5.5.17}$$

which, in the limit where $(E(\gamma_0), E(\gamma_1)) \to (\infty, \infty)$

$$\sigma(M) \simeq b \, \sqrt{E(\gamma_0)E(\gamma_1)(1 - \cos\theta)} = \frac{b}{\sqrt{2}} M \tag{5.5.18}$$

this means that the same resolution factor that affects the energy resolution affects in the same way the mass resolution and that the only relevant term of the enrgy resolution is the background term. A regression, which can be checked in Fig. 5.5.12 was performed in order to obtain this parameter, which is found to be $b = \sqrt{2} \times 2.39\% = 3.16\%$ which is very similar to the reported number in [117]. Combining this fact with the knowledge that data and MC

resolution differ $\sim 9\%$ from the $B_s^0 \to \phi\gamma$ and $B^0 \to K^{*0}\gamma$ analysis it can be infered how the resolution will translate in data.

A cross-check is performed to the fact that the resolution is in fact reproduced only by the parameter just calculated. In Fig. 5.5.13 the fully reconstructed peaks are overlayed to those that were reproduced by smearing the energy using a Gaussian p.d.f. In the former the following parameters are used: $\sigma = 3.35\%E$ and $\mu = 1.01 \times m_{\text{ALP}}$. One can see that this model almost perfectly reproduces the behaviour. Knowing that the difference between MC and Data is around 9% which shows that the actual relevant resolution term is around 3% , rather than 5% as it is noted in [117].

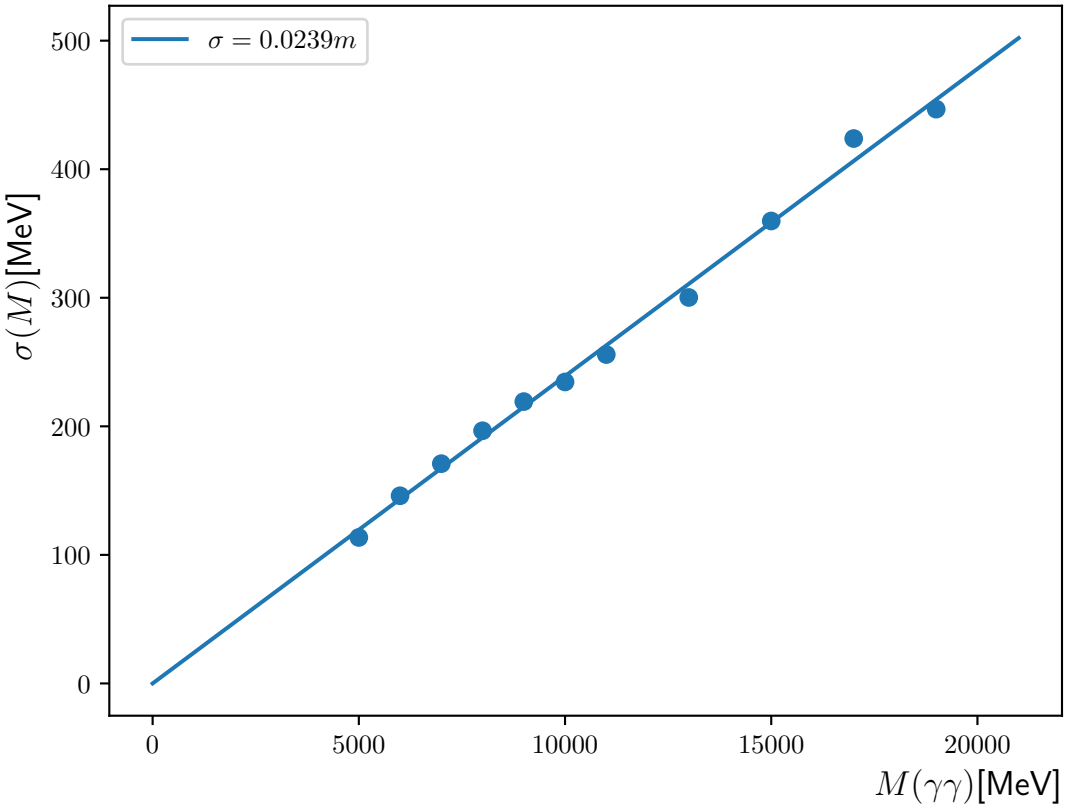

Figure 5.5.12: Variance of the Gaussian cores of the fits to the ALP samples in dependence of the ALP mass. A linear regression is overlaid to determine the accuracy of the resolution dependency. A clear linear behaviour is observed.

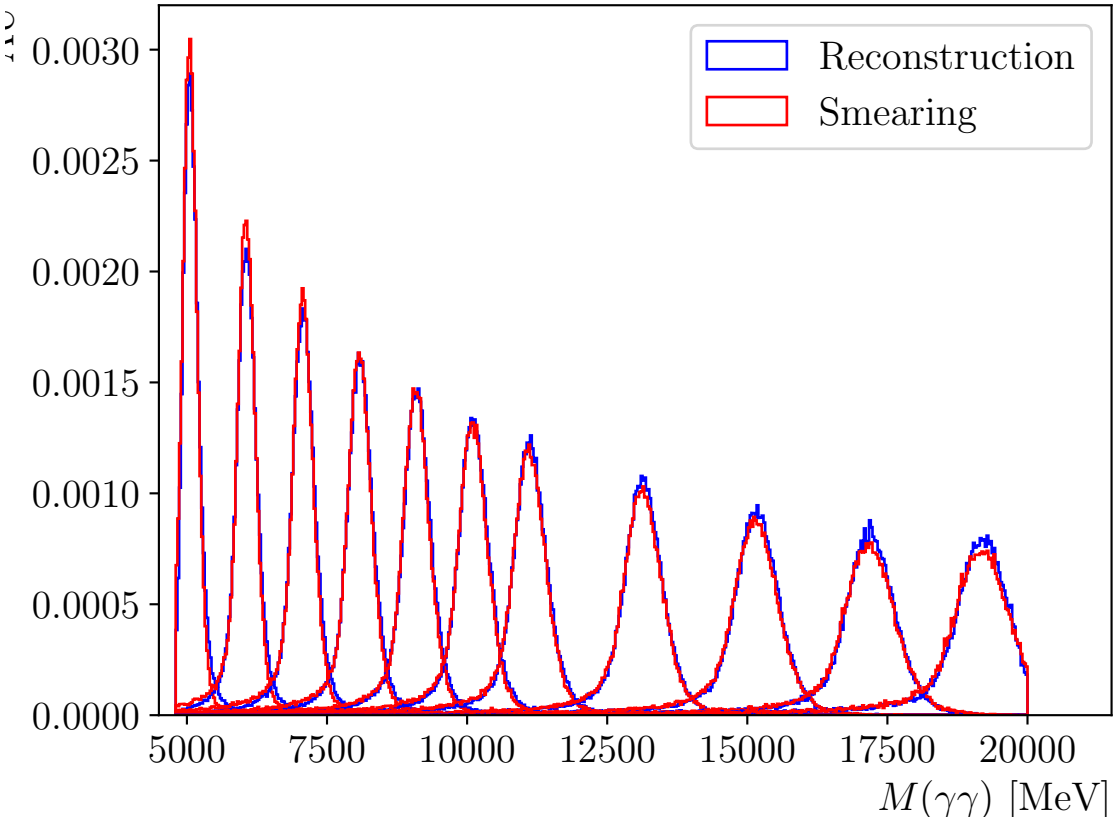

Figure 5.5.13: Reconstructed ALP peaks overlayed to smeared versions using a double Crystal Ball p.d.f.

Table 5.5.13: (Double) Crystal Ball parameters from the fits to the ALP simulation samples.

Mass [GeV]	m/m_{gen}	σ [MeV]	$\Delta\sigma$ [MeV]	α_L	α_R	n_L	n_R
5	$1.011\,24 \pm 0.000\,29$	109.0 ± 1.4	-	-	1.18 ± 0.04	-	150 ± 60
6	1.0076 ± 0.0009	147.0 ± 1.7	15 ± 4	1.55 ± 0.13	1.70 ± 0.11	3.7 ± 1.5	130 ± 100
7	1.0082 ± 0.0009	172.6 ± 2.4	6 ± 6	1.41 ± 0.14	1.63 ± 0.10	4.4 ± 2.4	150 ± 290
8	1.0107 ± 0.0009	197.5 ± 2.2	3 ± 5	1.66 ± 0.10	1.68 ± 0.11	2.0 ± 0.7	150 ± 260
9	1.0097 ± 0.0008	220.4 ± 2.2	11 ± 5	1.69 ± 0.09	2.09 ± 0.20	1.9 ± 0.6	2.9 ± 2.6
10	1.0092 ± 0.0009	236.1 ± 2.9	20 ± 7	1.48 ± 0.12	1.76 ± 0.11	3.2 ± 1.1	150 ± 220
11	1.0091 ± 0.0007	257.0 ± 2.8	23 ± 6	1.54 ± 0.10	1.78 ± 0.09	2.9 ± 0.9	160 ± 180
13	1.0083 ± 0.0008	305 ± 4	27 ± 8	1.41 ± 0.10	1.87 ± 0.19	3.2 ± 1.0	100 ± 700
15	1.0095 ± 0.0008	357 ± 5	13 ± 10	1.68 ± 0.17	1.88 ± 0.13	1.1 ± 0.6	$158.703\,44 \pm 0.000\,32$
17	1.0117 ± 0.0010	420 ± 5	5 ± 13	1.63 ± 0.04	1.91 ± 0.14	1.0000 ± 0.0004	158.0 ± 1.3
19	1.0151 ± 0.0006	406 ± 8	-	1.07 ± 0.13	-	4.3 ± 2.4	-
20	0.9997 ± 0.0007	241 ± 10	-	$0.500\,00 \pm 0.000\,04$	-	5.6 ± 2.3	-

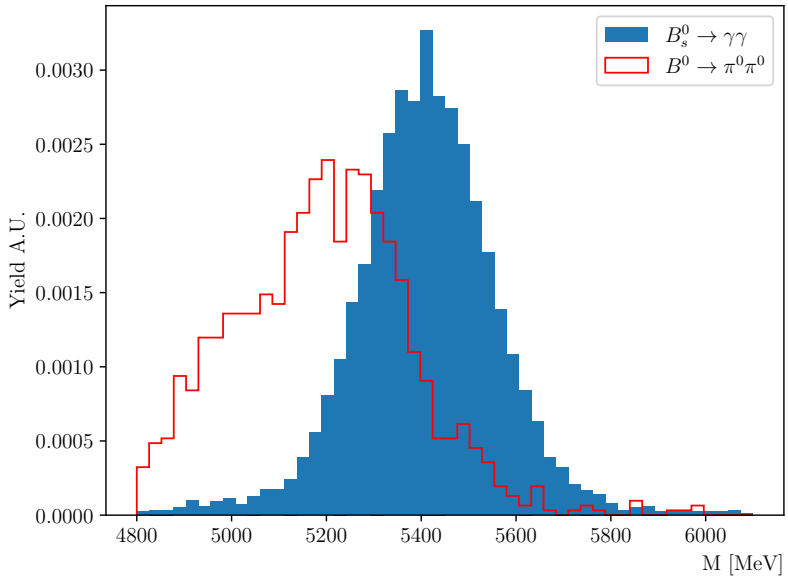

Figure 5.6.1: Invariant mass distribution of $B_s^0 \to \gamma\gamma$ candidates and misreconstructed $B^0 \to \pi^0\pi^0$ candidates.

5.6 Potential background contributions

In principle there are very few background sources in the studied mass region that could potentially peak. Most dangerous backgrounds would be direct diphoton resonances or dipion resonances. Of that type only the decay $B^0 \to \pi^0\pi^0$ exists in the mass range with a branching fraction of $\mathscr{B}(B^0 \to \pi^0\pi^0) = (1.59 \pm 0.26) \times 10^{-6}$ [93]. The $B^0 \to \pi^0\pi^0$ selection efficiency is estimated to be of $\mathscr{O}(10^{-5})$ using uncalibrated simulation, which lays one order of magnitude below the $B_s^0 \to \gamma\gamma$ efficiency value estimated in the same manner ($\mathscr{O}(10^{-4})$). This degraded efficiency can be understood to come from the PID selection tailored to discriminate against π^0 mesons and the fact that only a fraction of π^0 appear as merged π^0 in the calorimeter (which is what fakes the photon candidate), while others are resolved and make up two photon clusters. To illustrate this, Fig. 5.6.1 shows the invariant mass distribution of $B_s^0 \to \gamma\gamma$ and the invariant mass distribution of $B^0 \to \pi^0\pi^0$ decays. Though the resulting $B^0 \to \pi^0\pi^0$ shape shows a peaking structure it is considerably wider and shifted because of the partial reconstruction. Following the efficiency calculations about $\mathscr{O}(10)$ $\pi^0\pi^0$ decays are expected in the final data set which will be fully overwhelmed by the combinatorial background.

This also gives confidence that partially reconstructed backgrounds that could arise from heavy flavour resonances into $2\gamma[\pi^0] + X$ do not create any peaking structure that could be confused for the signal. The only known resonances in the investigated range are $B^0, B_s^0, \Lambda_b^0, B_c^+, \eta_b, \Upsilon(1S), \Upsilon(2S), \Upsilon(3S), \Upsilon(4S), \Upsilon(5S)$. Apart from the decay discussed

above none of these particles are measured to decay into two neutrals. Λ_b^0 decays should be negligible as there is "no" misidentification possible between proton and photon and thus at least a proton will be not reconstructed in the decay chain. B_c^+ decays have to be partially reconstructed due to electric charge conservation. The production fraction $\frac{f_c}{f_d}$ is $\sim 7.6 \times 10^{-3}$ [118]. No direct radiative decays have been observed (nor decays with π^0). The highest observed decay branching fraction with consecutive neutral decays is $B_c^+ \to J/\psi D^{*+}$. Assuming 100 % branching fraction for $D^{*+} \to \gamma/\pi^0 + X$ and a conservative 10 % for $J/\psi \to \gamma/\pi^0 + X$ [93, 118], this results in an effective branching fraction of $\frac{f_c}{f_s} x \mathscr{B}(B_c^+ \to (J/\psi \to \gamma/\pi^0 X)(D^{*+} \to \gamma/\pi^0 X')) \sim 3 \times 10^{-5}$, which is negligible given the background and the partially reconstructed nature of the decays (which will also have a negative impact on the isolation features). Finally, to discuss about the quarkonia: correcting for the dimuon branching fraction the $\Upsilon(1S)$ cross section is $\sim 189\,\text{nb}$ [119], while the B_s^0 cross section is $\sim 14.4\,\text{b}$ [120]. Combining this with the normalisation factor of $B_s^0 \to \gamma\gamma$ one reaches a normalisation factor for $\Upsilon(1S)$ decays of 1.8×10^{-6}. This factor assumes a BF of 100 % for $\Upsilon(1S)$ decays into $2\pi^0/\gamma$ (and does not take into account further efficiency losses due to lower kinematics of partially reconstructed decays and π^0 suppression). The best candidate for these decays are $\Upsilon(1S) \to \gamma\pi^0\pi^0$ decays ($\mathscr{B}(\Upsilon(1S) \to \gamma\pi^0\pi^0) = 1.7 \times 10^{-5}$ [93]), which could have a yield of $\mathscr{O}(10)$ events. Note that due to the unfortunately quite high combinatorial background level a visible contribution of at least 1 sigma would have to contain at least $100 - 200$ decays in that mass regime and thus also these decays would not be visible.

In a more general approach, the composition of the background was studied with simulated minimum bias samples, applying the Stripping and Trigger selection. By requiring high p_T neutrals at generation time high statistics was ensured. After applying the trigger selection 2714 candidates are available in the higher statistics sample (the conclusions with the other sample are compatible).

For each of the photon candidates its true ID, the ID of its mother particle and its grandmother was investigated. This study shows that the main source of background comes from π^0 and η meson decays, in $\sim$ 98% of the cases at least one of those is present and in $\sim 90\%$ of them a π^0 is present while in $\sim 28\%$ of the total cases and η decay is there. Finally, $\sim 70\%$ of the background is composed by either of these two meson decays, the 30 % remaining includes more light meson decays together with prompt photon production.

The π^0 particles themselves have been checked to either come directly from the pp collision (40 %), or other non-heavy flavour decays (56 %). Only about 4 % of the π^0 are from heavy flavour decays.

The above study however has the caveat that the truth matching is only very limited in precision. Because a neutral object has only a 3×3 cluster in the ECAL as detector information, it can easily happen that the truth matching algorithm combines a merged π^0 cluster with one of the true MC photons originating from that particle (the other way round with identifying a photon cluster with a random merged π^0 is much less likely, so there is

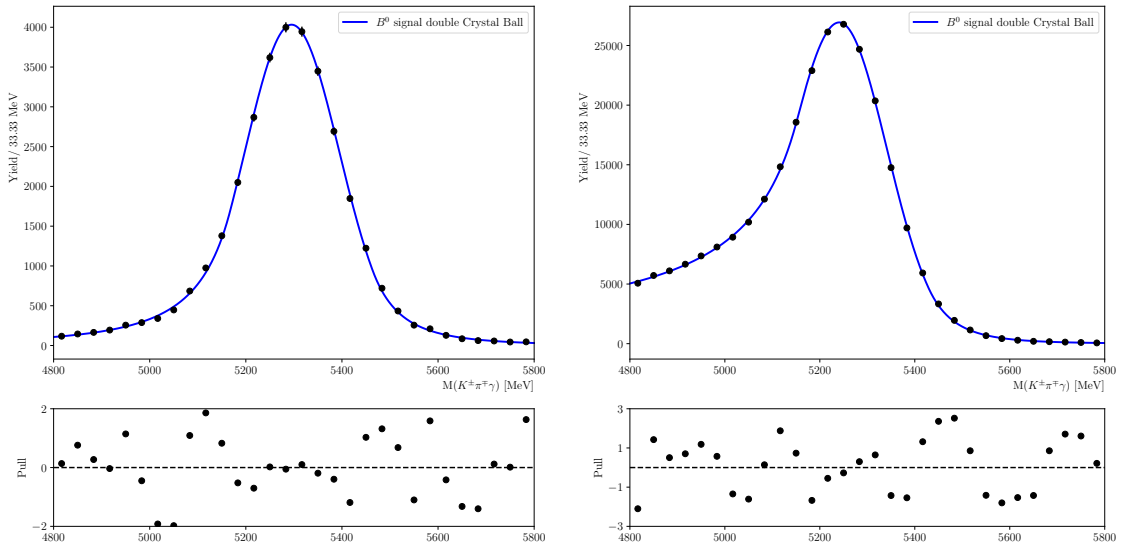

Figure 5.6.2: $B^0 \to K^*\pi^0$ simulation reconstructed as $B^0 \to K^*\gamma$. Left: enriched merged π^0 selection. Right: enriched resolved π^0 selection, which shows a strong tail indicating the missing photon from $\pi^0 \to \gamma\gamma$ decays.

no issue with our signal truth matching). That is why the background study showed a large fraction of background being photons from π^0 decays.

A simulated sample of $B^0 \to K^*\pi^0$ decays, reconstructed via $B^0 \to K^*\gamma$, was studied to investigate that effect. By applying γIsPhoton < 0.1 it can be ensured that mostly merged π^0 mesons pass the selection. This statement is backed up in Fig 5.6.2 (left). The peak shows no relevant tails and the center is precisely centered at the B^0 mass. On the contrary if one applies γIsPhoton > 0.9 as seen in Fig 5.6.2 (right) a prominent left-tail appears, accounting for a more predominant resolved π^0 presence. In the first case, though it has been shown that a negligible fraction of resolved π^0 decays are present, only $\sim 30\%$ are identified as merged π^0 using the truth matching machinery. Following this study one can deduce that most of the background candidates are actually not photons from π^0, but merged π^0 particles themselves and therefore a tight cut on the IsPhoton variable is adopted in the nominal selection of the diphoton candidates as discussed above. This helps in further reducing the background.

Taking all the above studies together it is fair to assume a purely combinatorial nature of the background, without any peaking structure. For this reason, the background is empirically modelled with Chebychev polynomials.

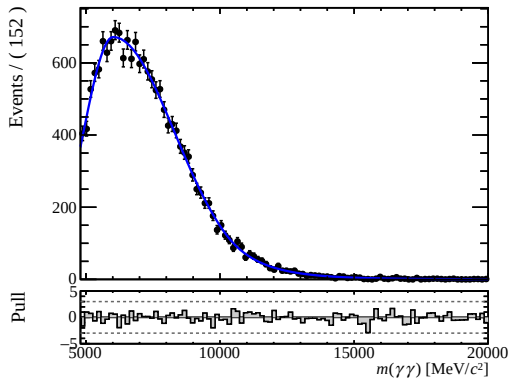

Figure 5.7.1: Fit to the small data subset to obtain a data distribution proxy. The data are very well described by a double Crystal Ball function.

5.7 Expected sensitivity

A scan across the fully selected mass range is performed to obtain a limit on the (production $\times$ decay) cross section of the process depending on the diphoton mass. A mass window of $\pm 5\sigma$ around the inspected mass is selected and projected to a range of $[-1, 1]$. The width at the inspected mass is determined from the linear fit performed in Fig. 5.5.12.

The background distribution is described with a sum of odd Chebychev polynomials up to the ninth order and even polynomials up to the second order. Chebychev polynomials are chosen for their great computational performance and orthogonality between even and odd modes.

The signal distribution is described with a Double Crystal Ball, centered at 0 of the determined range by definition, where the width is taken from the linear fit in Fig. 5.5.12 and the other parameters are averaged from Tab. 5.5.13, as no clear dependency is observed.

Each data set is then fitted with a an unbinned maximum likelihood fit to obtain the cross section and an upper limit is determined with the CLs method using the one-sided test statistic as implemented in the `Prob` method (the teststatistic distribution is calculated using asymptotic formulae [121]) as implemented in the `GammaCombo` or `RooStats` frameworks [122, 123] (both have been implemented and cross checked against each other). The efficiencies and normalisation of the signal yield to obtain the cross sections are included as nuisance parameters in the fit which are Gaussian constrained to the numbers determined above. The width of this Gaussian constraint corresponds to the systematic uncertainty on the efficiencies.

The procedure is validated using a small subset (4 %) of the data as proxy. The fully selected data is fitted with a wide double crystal ball model (empirically chosen to not select any signal-like peak) as shown in Fig. 5.7.1. From this fitted model, a toy data set corresponding to the full data size is generated and used to determine the expected upper limits,

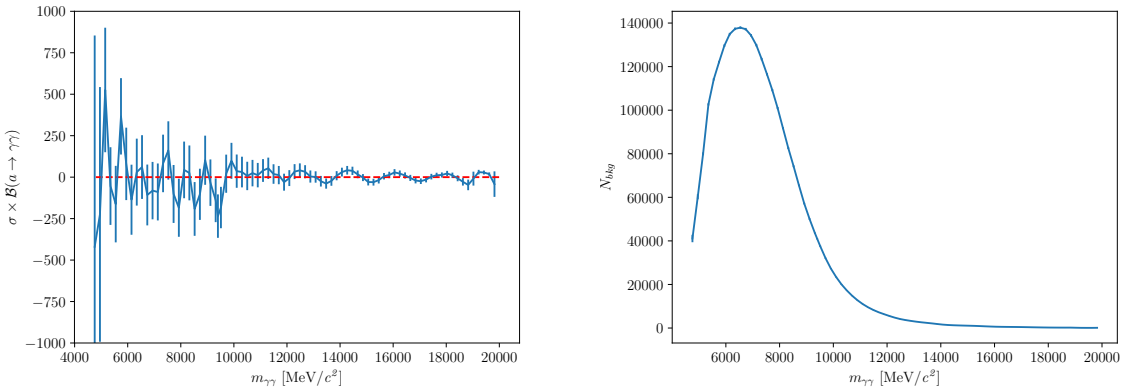

Figure 5.7.2: Signal (left) and background (right) yield distributions for the proxy bumphunt. As expected, no significant signal deviation from zero is observed. All signal yield estimates are unbiased around zero.

such that the full procedure with reliable expectations can be tested without disclosing the data.

Figure 5.7.2 shows the distribution of the signal yields and background yields, demonstrating that despite a non-monotonous slope the fitting is unbiased. Two example plots of the fits are shown in Fig. 5.7.3, with the full list of plots in App. B. Given the data can be well described with the chosen model and do not strongly depend on the amount of included odd polynomials (see also Ref. [124]), no extra systematic on the background is considered. The resulting expected limit distributions for this test are given in Fig. 5.7.4. At the edges of the investigated mass range the limits soften significantly because there the signal distribution is not central anymore but closer to the boundary (because not the full 10 sigma mass window is available anymore). Thus the orthogonality between the odd polynomials and the mostly even signal distribution breaks down and background and signal become hardly distinguishable, leading to soft bounds. Therefore in this region only a lower complexity model is chosen: up to the 5th odd order to mitigate the effect.

Additional cross checks for the fit stability are performed. The benchmark masses at $[5.2, 6.0, 6.2, 12, 15, 17, 19]$ GeV are investigated and for each point 1000 pseudo-experiments are thrown, injecting signal with a significance of $0, 0.5, 1.0$ and 2.0 σ (where σ is defined as the square root of the background yield in full the search window). These data from these pseudo-experiments is then refitted and pull distributions are produced, which are then themselves fitted with a Gaussian. The results of the fits to the pull distributions of all tests are given in Tab. 5.7.1. No significant fit bias is found.

The results are interpreted in the context of the model discussed in [25] as limits on the ALP decay constant as shown in Fig. 5.7.4, where the translation between f_a and $\sigma \times \mathscr{B}$ is taken directly from the ones calculated in [25]. In [39] strong bounds on the decay constant are also placed. This work follows a different approach by exploiting the non-resonant nature

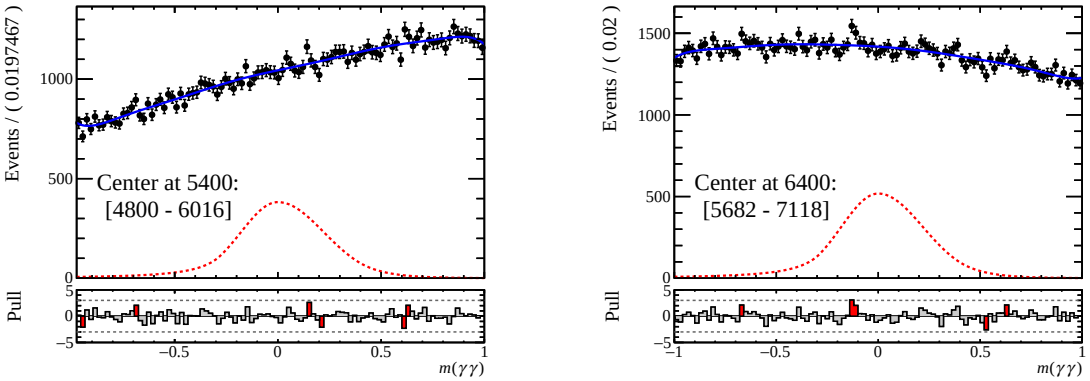

Figure 5.7.3: Example fits to two mass ranges of the proxy bump hunt. The data are well described by the chosen model.

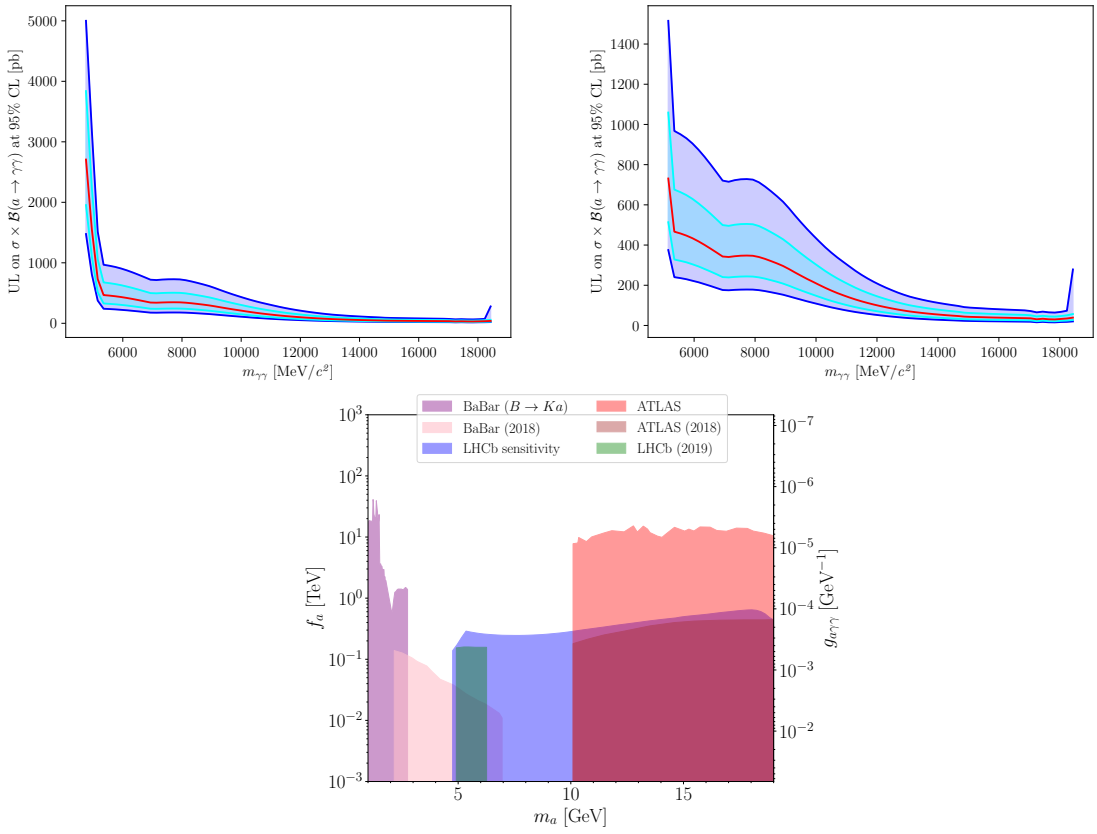

Figure 5.7.4: Left: distribution of the expected limits on the signal yield of the proxy bump hunt. Right: reduced region, less affected by boundary effects. Bottom: projection of these limits in the model discussed in [25].

Table 5.7.1: Results of fits to the pull distribution of the fit bias toys. For each configuration the investigated mass, the number of injected signal candidates (N_{inj}), as well as the mean (for a perfect pull supposed to be 0) and the width (for a perfect pull supposed to be 1) of the Gaussian fitted to the pull distributions.

Mass [GeV]	N_{inj}	Pull mean	Pull sigma
5.2	0.0	-0.036 ± 0.031	0.987 ± 0.022
	143.4	-0.014 ± 0.031	0.980 ± 0.022
	286.8	0.008 ± 0.030	0.947 ± 0.021
	573.5	-0.020 ± 0.031	0.981 ± 0.022
6.0	0.0	0.032 ± 0.032	0.998 ± 0.022
	180.4	0.029 ± 0.032	1.010 ± 0.023
	360.7	0.026 ± 0.032	0.998 ± 0.022
	721.5	-0.086 ± 0.031	0.976 ± 0.022
6.2	0.0	-0.032 ± 0.030	0.940 ± 0.021
	182.8	-0.041 ± 0.032	1.008 ± 0.023
	365.6	-0.069 ± 0.030	0.941 ± 0.021
	731.3	-0.055 ± 0.030	0.950 ± 0.021
12.0	0.0	0.005 ± 0.030	0.961 ± 0.022
	36.9	0.021 ± 0.031	0.992 ± 0.022
	73.7	-0.031 ± 0.030	0.935 ± 0.022
	147.4	-0.052 ± 0.030	0.963 ± 0.022
15.0	0.0	-0.071 ± 0.030	0.946 ± 0.021
	15.7	-0.043 ± 0.030	0.960 ± 0.022
	31.4	-0.017 ± 0.030	0.960 ± 0.022
	62.9	-0.066 ± 0.031	0.984 ± 0.022
17.0	0.0	-0.028 ± 0.032	0.996 ± 0.022
	10.2	-0.031 ± 0.030	0.958 ± 0.021
	20.5	-0.031 ± 0.031	0.982 ± 0.022
	41.0	-0.035 ± 0.031	0.993 ± 0.022
19.0	0.0	-0.036 ± 0.034	1.065 ± 0.024
	6.1	-0.038 ± 0.033	1.032 ± 0.023
	12.3	-0.025 ± 0.034	1.061 ± 0.024
	24.6	-0.072 ± 0.033	1.027 ± 0.023

of off-shell ALPs, which would increase the yield of highly energetic photons. In this model the expected results exceed by far the existing limits in the range 4.8 GeV to 10 GeV from a direct search. Above 10 GeV the recent search performed by the ATLAS experiment [36] is superior. The comparison between the existing limits, found in [25] and this projection cannot be made fairly because the reconstruction efficiency assumption was not accurate, this is why the projection estimation could not be met. In turn, the background suppression was highly improved which partially compensates for it.

For the $B_s^0 \to \gamma\gamma$ and $B^0 \to \gamma\gamma$ decay branching fractions, dedicated limits are determined by fitting for a peak at the known B^0 and B_s^0 masses with a similar model as the bump hunt. With the normalisation factors from Sec. 5.5.10 (using the efficiencies of the $\mathrm{B_s^0} \to \gamma\gamma$ sample), expected limits of

$$\mathscr{B}(\mathrm{B_s^0} \to \gamma\gamma) < 2.5 \times 10^{-5}$$
$$\mathscr{B}(\mathrm{B^0} \to \gamma\gamma) < 0.72 \times 10^{-5}$$

at 95 % CL are calculated, which are far from the current best limits [125, 126] on these decays due to the poor $\mathrm{B_s^0} \to \gamma\gamma$ efficiency.
The expected upper limit on the $\eta_b \to \gamma\gamma$ signal using the ALP simulation as proxy is estimated as

$$\sigma(pp \to \eta_b X) \times \mathscr{B}(\eta_b \to \gamma\gamma) < 263\,\mathrm{pb}$$

which is about $2\times$ higher than the best exclusion limit [127] in the partial width of 466 eV. Together with the total η_b width of 10 MeV [128, 129] and a production cross section of 3000 nb [130], yields an effective best limit in the literature of $\sigma(pp \to \eta_b X) \times \mathscr{B}(\eta_b \to \gamma\gamma) <$ 140 pb.

Chapter 6

Conclusions

In this thesis the first analysis with only unconverted photons in the final state using LHCb data has been described. Also, trigger developments using GPUs towards reconstructing muons using only the VELO and MUON stations have been detailed.

A detailed study of the efficiency calibration of the different selection layers of a search for a diphoton resonance has been shown. The diphoton final state constitutes a unique use of the calorimeter as the only handle to reconstruct a final state. This particular topology is accompanied by a very special trigger selection with unique use of hardware level reconstruction clusters and also novel use of machine learning techniques. The asset which allowed to get the greater separation from background is the isolation of the final state photons. Finally, the sensitivity of LHCb in the $[5,20]$ GeV mass region has been shown, with world best prospects in the $[5,10]$ GeV region.

The Allen project embodies the research effort of putting together a trigger system based on GPUs for the LHCb experiment during the Run 3 data-taking period and beyond. It is the first time that such a system is designed to run in a fully standalone way on GPUs. A dedicated VELO-MUON tracking algorithm was designed to run on this framework and in order to prove the power of it, $\mu^+\mu^-$ trigger selections that have been developed. The tracking sequence consists on a standalone reconstruction of tracks in the MUON stations for posterior matching to already reconstructed VELO tracks. The momentum resolution is at the level of 5%, thanks to a parameterisation of the magnetic field designed using simulated tracks. This reconstruction technique offers the possibility of tracking muons for high multiplicity events, where the tracker detectors might be too busy to even be decoded. It also serves as alternative tracking algorithm which runs independently of the other tracking sequence. It can also serve as handle to measure efficiencies through the tag-and-probe method.

These two topics while showcasing two very different aspects of LHCb, they portray the flexibility of it by using alternative approaches to study physics phenomena.

Appendices

Appendix A

Efficiency determination

This appendix shows plots that are relevant for the efficiency determination of the HLT1 and HLT2 trigger selections.

A.1 HLT1

This appendix is intended to show the HLT1 distributions that are discussed in Sec. 5.5.5.

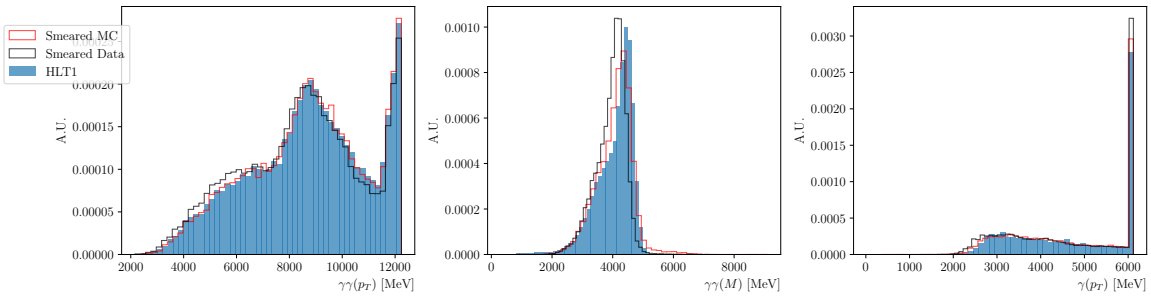

Figure A.1.1: HLT1 kinematic distributions for a 5 GeV ALP.

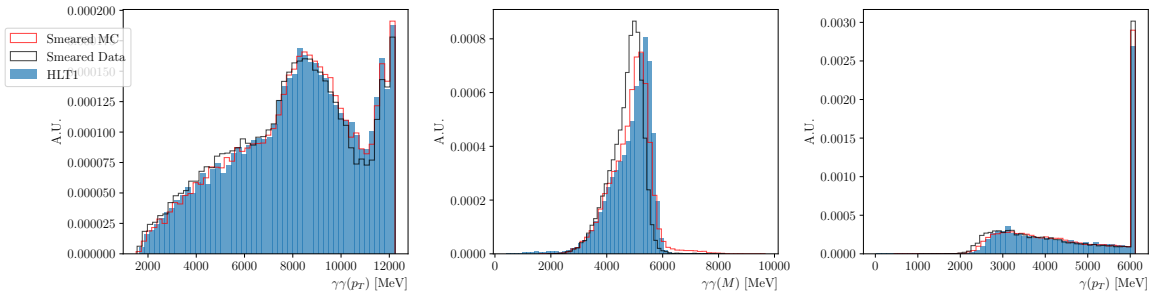

Figure A.1.2: HLT1 kinematic distributions for a 6 GeV ALP.

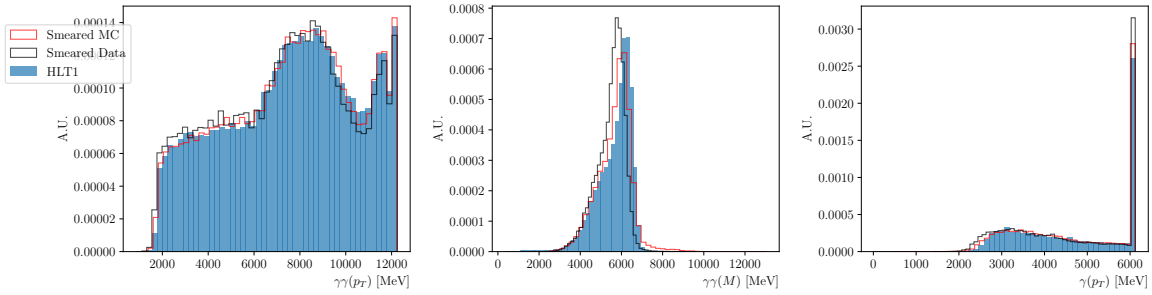

Figure A.1.3: HLT1 kinematic distributions for a 7 GeV ALP.

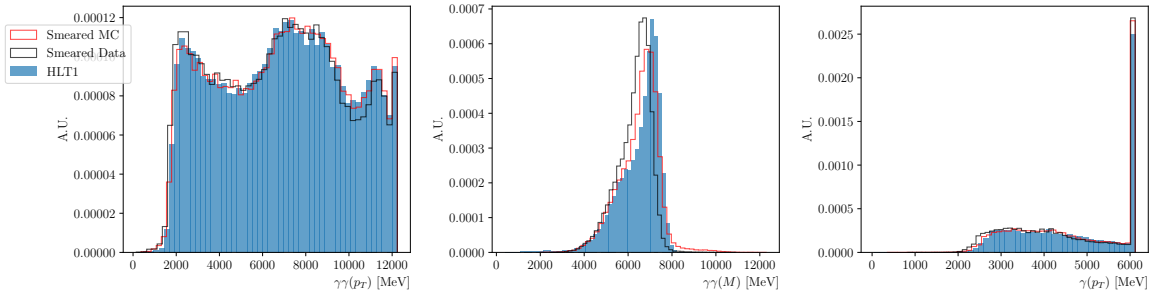

Figure A.1.4: HLT1 kinematic distributions for a 8 GeV ALP.

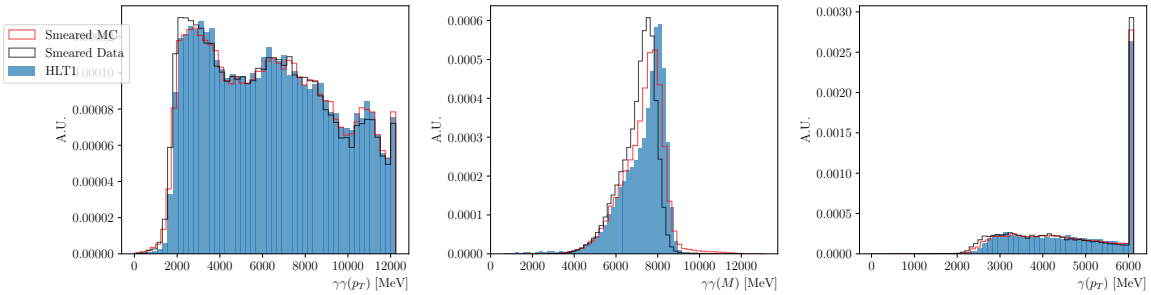

Figure A.1.5: HLT1 kinematic distributions for a 9 GeV ALP.

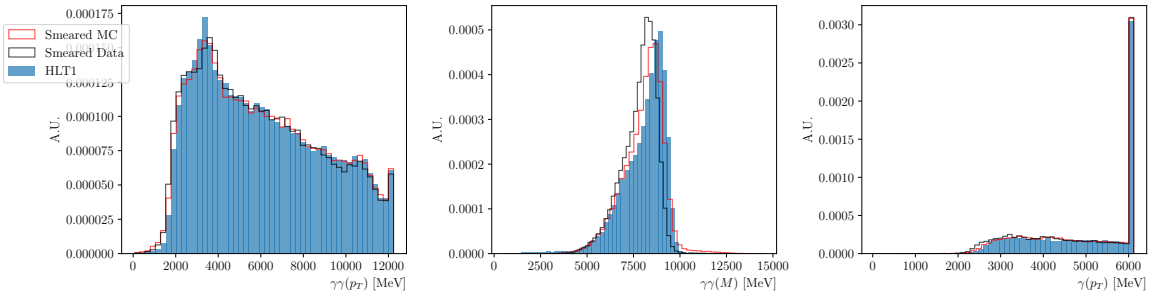

Figure A.1.6: HLT1 kinematic distributions for a 10 GeV ALP.

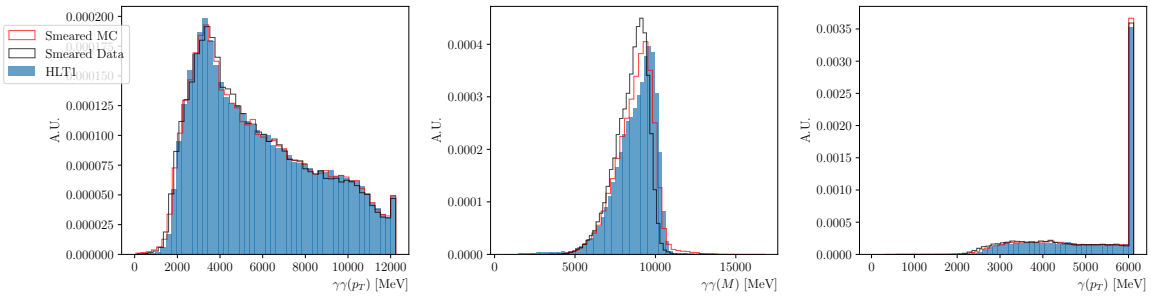

Figure A.1.7: HLT1 kinematic distributions for a 11 GeV ALP

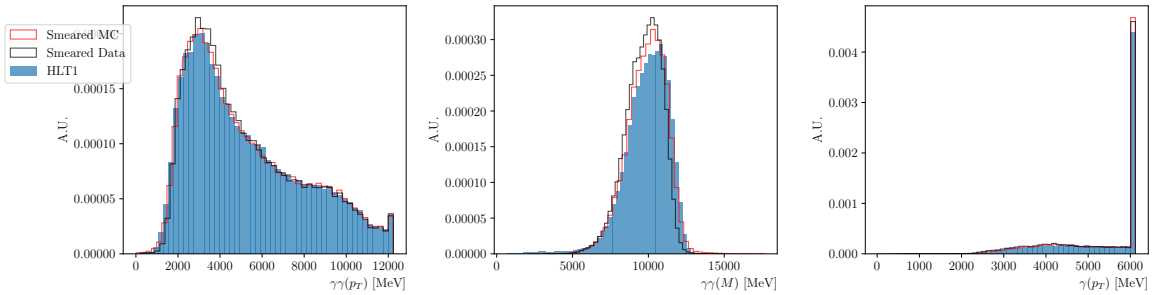

Figure A.1.8: HLT1 kinematic distributions for a 13 GeV ALP.

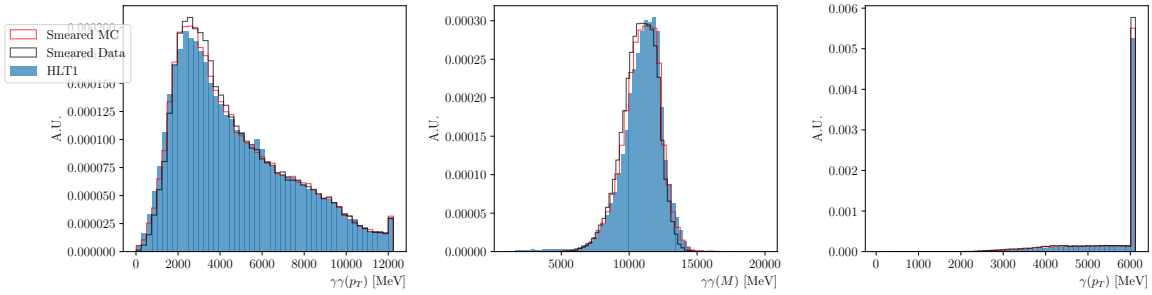

Figure A.1.9: HLT1 kinematic distributions for a 15 GeV ALP.

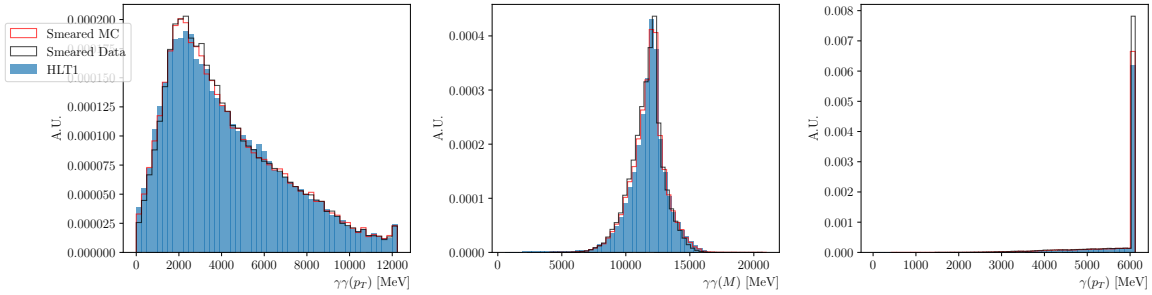

Figure A.1.10: HLT1 kinematic distributions for a 17 GeV ALP.

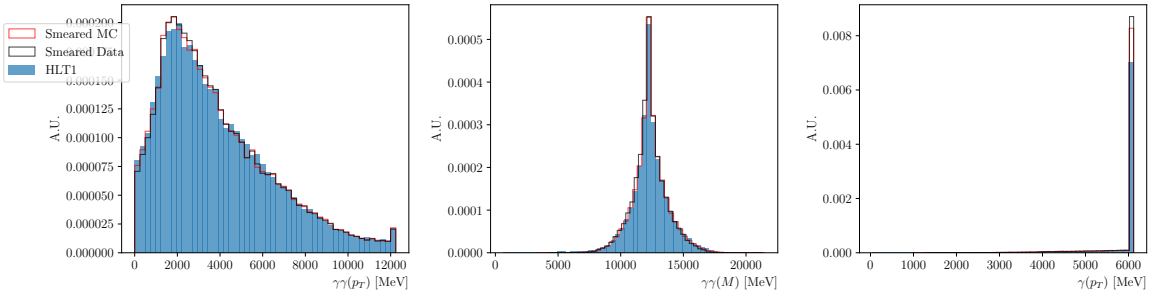

Figure A.1.11: HLT1 kinematic distributions for a 19 GeV ALP.

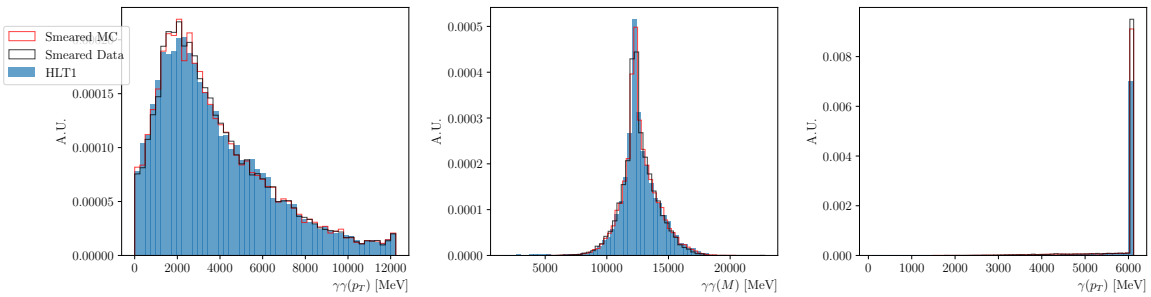

Figure A.1.12: HLT1 kinematic distributions for a 20 GeV ALP.

A.2 HLT2

This appendix is intended to show the HLT2 distributions that are discussed in Sec. 5.5.6.

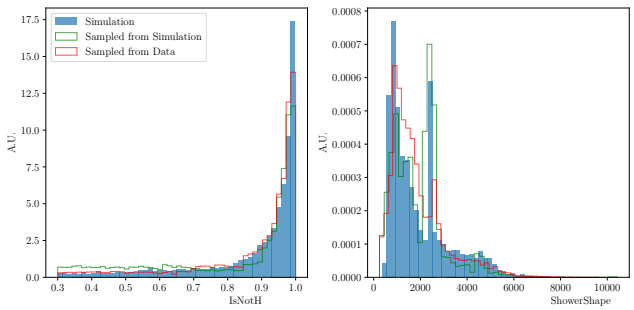

Figure A.2.1: HLT2 `IsNotH` 5 GeV ALP.

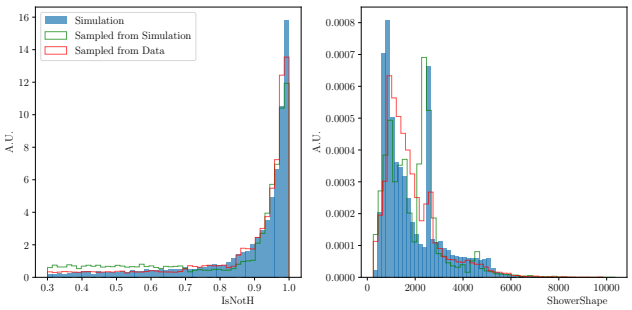

Figure A.2.2: HLT2 `IsNotH` 6 GeV ALP.

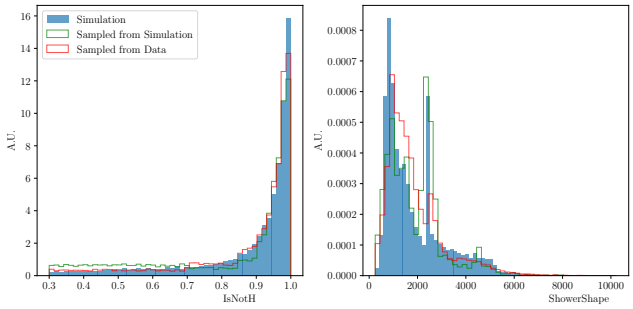

Figure A.2.3: HLT2 `IsNotH` 7 GeV ALP.

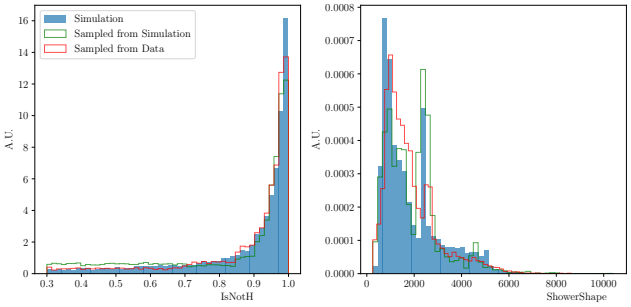

Figure A.2.4: HLT2 `IsNotH` 8 GeV ALP.

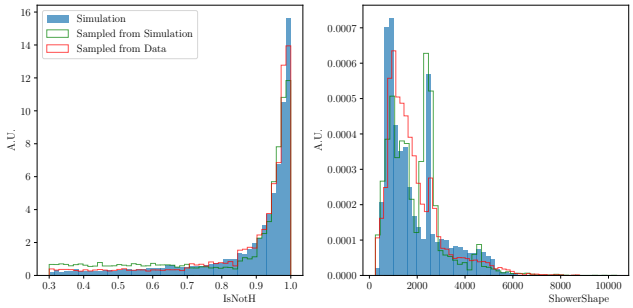

Figure A.2.5: HLT2 `IsNotH` 9 GeV ALP.

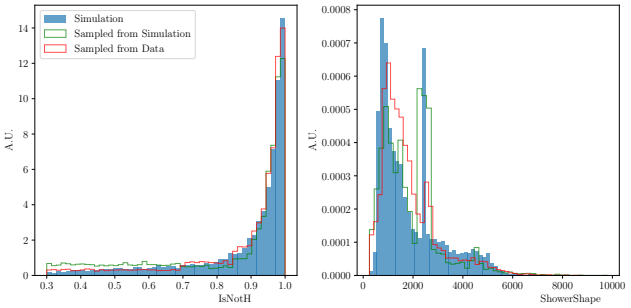

Figure A.2.6: HLT2 `IsNotH` 10 GeV ALP.

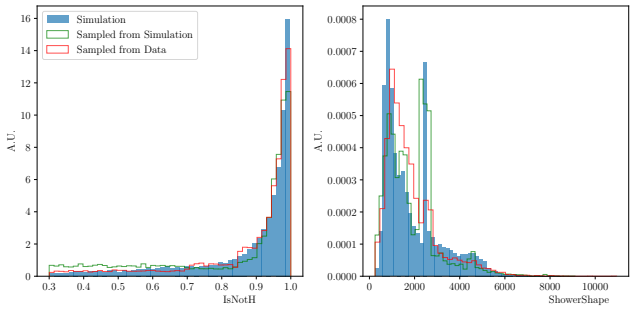

Figure A.2.7: HLT2 `IsNotH` 11 GeV ALP.

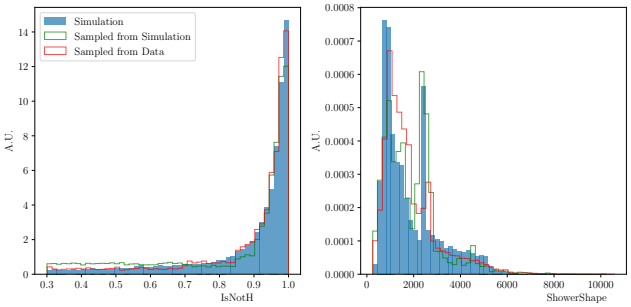

Figure A.2.8: HLT2 `IsNotH` 13 GeV ALP.

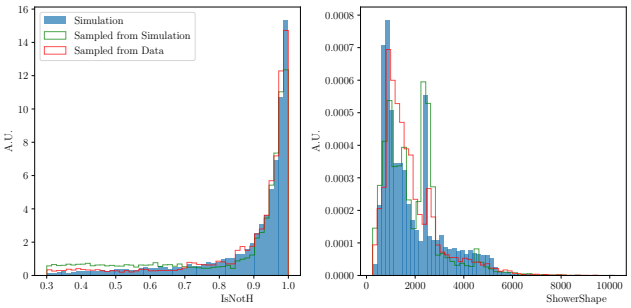

Figure A.2.9: HLT2 `IsNotH` 15 GeV ALP.

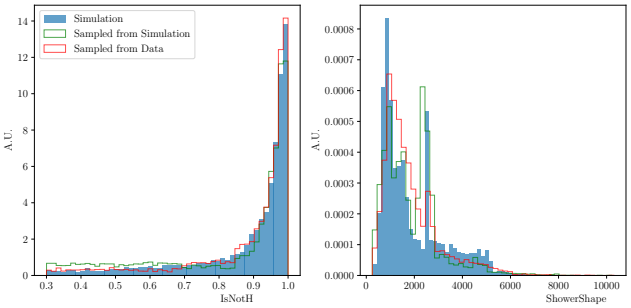

Figure A.2.10: HLT2 `IsNotH` 17 GeV ALP.

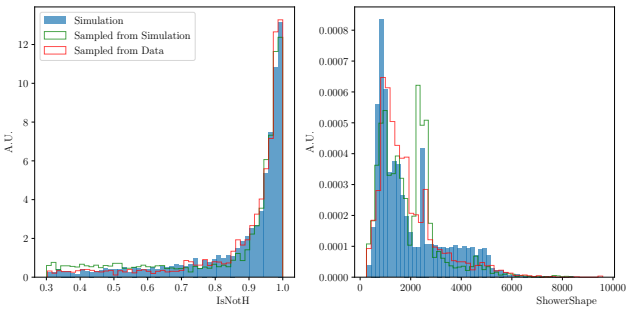

Figure A.2.11: HLT2 `IsNotH` 19 GeV ALP.

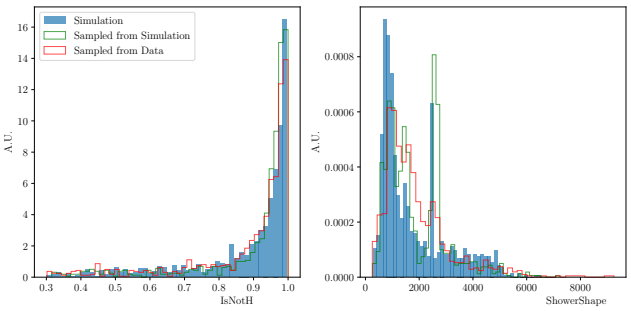

Figure A.2.12: HLT2 `IsNotH` 20 GeV ALP.

A.3 Resolution

Plots of the ALP peak for different mass hypotheses are shown in Fig. A.3.1.

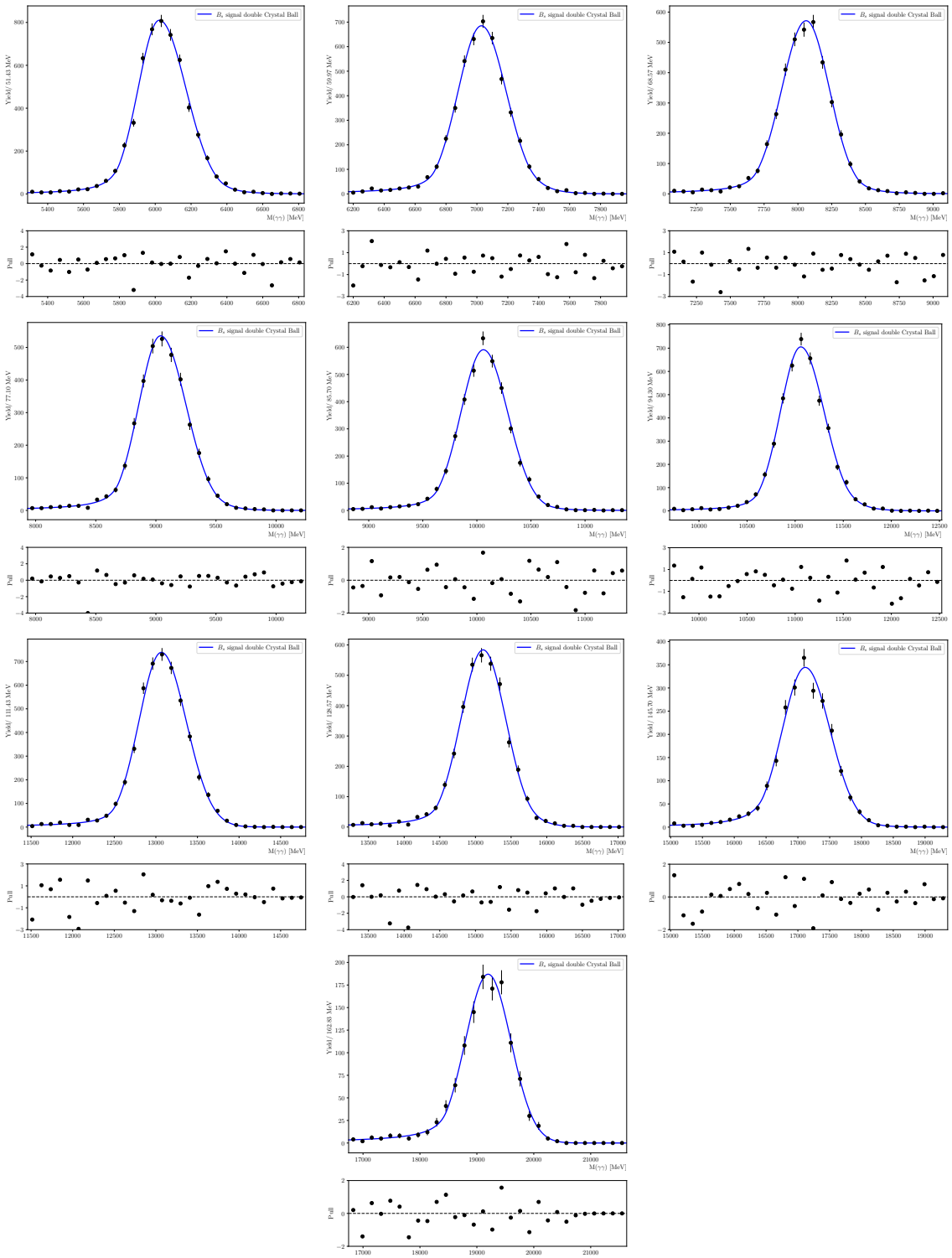

Figure A.3.1: Figure showing the ALP peaks fitted to their correspondent signal pdf.

Appendix B

Bump hunt fits on the data proxy

These plots show all fits performed to determine the expected upper limit distribution in Sec. 5.7.

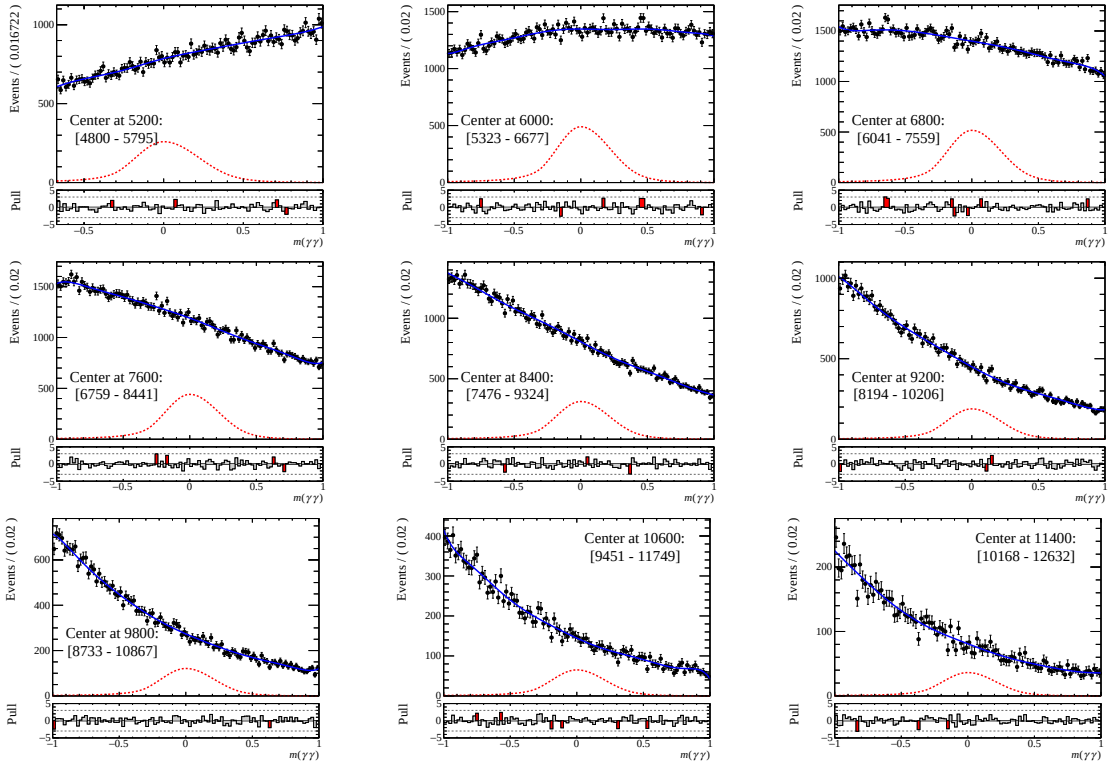

Figure B.0.1: These plots show all fits performed to determine the expected upper limit distribution in Sec. 5.7.

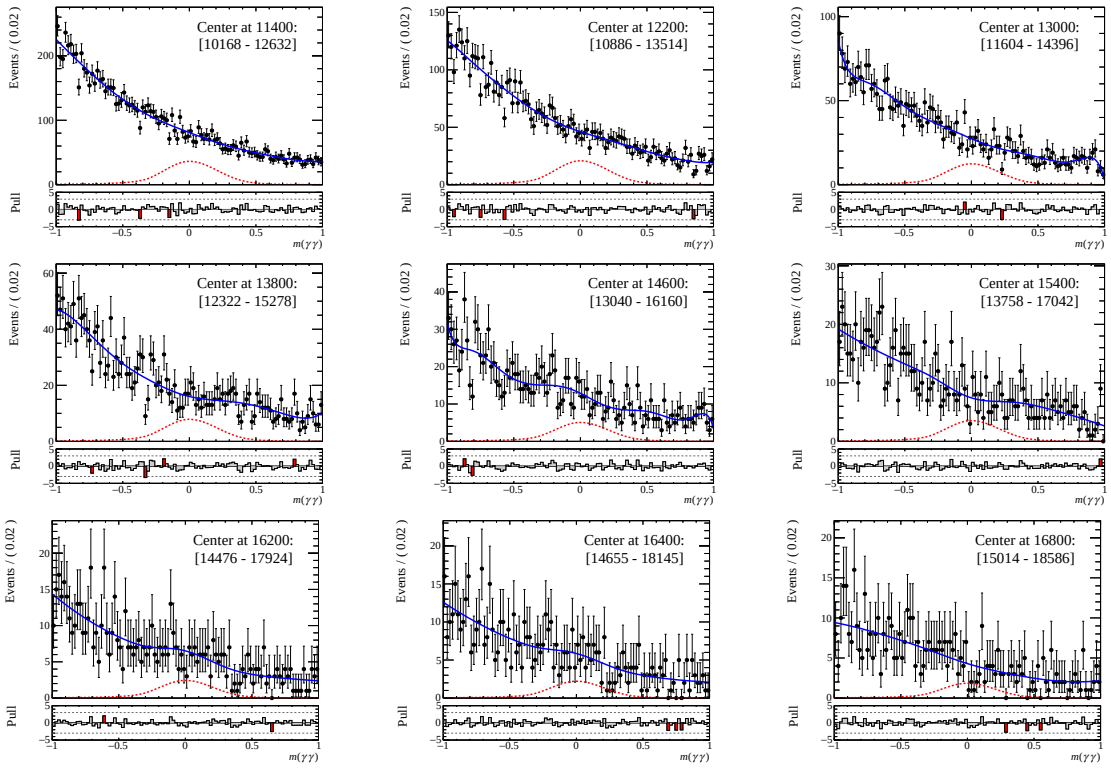

Figure B.0.2: These plots show all fits performed to determine the expected upper limit distribution in Sec. 5.7 (ctd.).

Appendix C

Resumo da tese en galego

O Modelo Estándar (SM) da física de partículas probou, ata o día de hoxe, ser a teoría mais exacta para explicar as interaccións entre as partículas elementais que constituen o noso universo. A pesares disto, sabemos que o SM é unha teoría como mínimo incompleta, xa que ten algunhas fraquezas, e existen diferentes propostas de modelos que buscan constituintes do universo que non están recollidos no SM.

C.1 Fundamentos teóricos

Se o SM fose invariante baixo conxugación de carga (C) e paridade (P) a materia e a antimateria serían totalmente indistinguibles, pero non é así. A violación CP presente no sector electrofeble non é suficiente para explicar toda a que observamos. O feito de non atopar violación CP no sector forte indica que os termos que a poden producir deben desaparecer do lagranxiano dalgún xeito.

Ao lagranxiano de QCD podemos engadirlle un termo proporcional ao tensor enerxía-momento da teoría.

$$\mathscr{L}_{\text{eff}} = \mathscr{L}_{\text{QCD}} + \frac{\theta \alpha_s^2}{32\pi^2} G_{\mu\nu,a} \tilde{G}^{\mu\nu,a} \tag{C.1.1}$$

O termo θ debería ser extremadamente pequeno pero non exactamente 0, o cal fai pensar a posibilidade dunha nova simetría que dea lugar aos axións, a cal estaría espontánemanete rota a través dun mecanismo similar ao de Higgs que caracterizaría este tipo de partículas de acordo co seguinte lagranxiano [20, 21]:

$$\mathscr{L}_{\text{eff}} = \mathscr{L}_{\text{SM}} + \frac{\alpha_s^2}{32\pi^2} \frac{\xi}{f_a} a_{\text{phys}} G_{\mu\nu,a} \tilde{G}^{\mu\nu,a} - \frac{1}{2} \partial_\mu a_{\text{phys}} \partial^\mu a_{\text{phys}} + \mathscr{L}_{\text{int}}(\partial_\mu a_{\text{phys}}, \psi) \tag{C.1.2}$$

Deste xeito, o parámetro θ non estaría presente e e a violación CP do sector de QCD sería 0.

Outros modelos dan lugar a partículas con características similares, coñecidas como $e^{\mp}$Axion-like Particles (ALPs), algúns modelos que as explicarían:

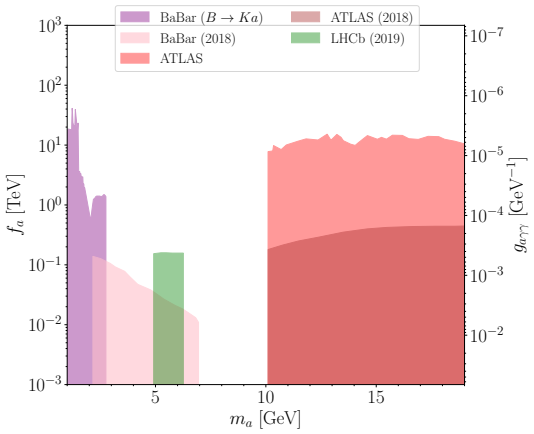

Figura C.1.1: Límites actuais na búsqueda dos ALPs en relación á súa masa e á súa constante de acoplo.

Este tipo de partículas poden aparecer en diferentes extensións do SM. Normalmente coñécense na literatura como *Axion-like Particles*. Normalmente os ALPs poden alcanzar masas que serían visibles no LHC, mentras que outras partículas que aparecerían en conxunto con estes ALPs poderían estar fóra do alcance. Algúns modelos con ALPs son os seguintes:

- Axión masivo de QCD: este sería un candidato a axión, que dados certos modelos pode chegar ao rango dos GeV [22].
- ALPs como conectores ca Materia Escura: neste caso os axión xogarían o papel de conectar o sector visible co sesctor escuro [23].
- R-axión supersimétrico: os ALPs destes modelos aparecen a partir da ruptura de simetría dunha supersimetría [24].

Diferentes experimentos están a buscar este tipo de partículas establecendo límites en todo o espazo de parámetros (Fig. C.1.1).

C.2 Ferramentas estatísticas en Física de Altas Enerxías

As ferramentas estatísticas xogan un rol fundamental á hora de conectar as observacións experimentais coas teorías. Para determinar se unha observación está dacordo ou non cunha teoría ou proposta precisamos facer uso da estatística.

Son especialmente importantes os tests de hipótese que nos permiten determinar se una sinal, por exemplo, alén do SM está presente nos nosos datos ou non. Tamén, baixo o suposto de que ese sinal non está presente, cal é a capacidade do detector para excluir a súa abundancia.

C.3 O experimento LHCb no LHC

O LHC [53] é o acelerador de partículas máis grande e enerxético construido ata a o de agora, onde protóns son acelerados e feitos colisionar en diferentes experimentos, sendo os catro máis importantes: ATLAS [56], CMS [57], ALICE [58] e LHCb [59]; sendo este último o marco deste traballo.

LHCb é un espectrómetro cun só brazo instrumentado na rexión posterior, especializado en medir toda a física que provén dos quarks *b*, aínda que probou que o seu alcance é moito máis amplo. Cubre a rexión angular de, aproximadamente, 10 mrad a 300 (250) mrad no plano perpendicular (paralelo) ao campo magnético. Está composto polos seguintes elementos e subdetectores:

- Imán: dipolo magnético cunha $\int \vec{\mathbf{B}} \cdot d\vec{l} = 4$ T·m, serve para caracterizar a carga das partículas.
- VELO: subdetector central do experimento, primeiro elemento orientado a localizar os vértices onde decaen as partículas.
- Sistema de trazas: é un conxunto de detectores de silicio enfocados a reconstruir o camiño que seguen as partículas no experimento.
- Calorímetro electromagnético (ECAL): utilizado para medir a deposición enerxética dalgunhas partículas.
- Calorímetro hadrónico (HCAL): usado tamén para realizar medicións da enerxía das partículas.
- RICH: caracteriza o tipo de partícula segundo a súa radiación Cerenkov.
- Sistema MUON: utilizado para identificar muóns e reconstruir a súa traxectoria.

C.3.1 Sistema de trigger durante o Run 2

Durante o Run 2 procésanse eventos a unha taxa de 40 MHz, pero a maioria deles non teñen ningunha clase de sinal relevante. Este sistema ten por obxectivo reducir esa taxa para só gardar eventos potencialmente interesante reducindo a taxa a 3 MHz [86]. Divídese en tres niveis:

- L0: Procesa eventos a 40 MHz e discrimina segundo a información dos calorímetros e do sistema de muóns principalmente. Todo a información do evento que pasa o L0 avanza ao seguinte nivel.
- HLT1: Reconstrúee a traxectoria das partículas cargadas que atravesan todo o detector, reducindo no proceso a taxa a uns $\sim$ 100 kHz.

- HLT2: Conta coa reconstrucción completa do detector, incluído o alineamento que se fai en tempo real. Permite aplicar seleccións con detectores encargados de identificar as partículas, como o RICH.

Finalmente, hai un proceso de selección offline, é dicir, unha vez os eventos de interese están gardados en discos, coñecido como Stripping.

C.3.2 A simulación en LHCb

Para poder modelizar e comprender os efectos que aparecen nos datos, recúrrese á simulación, que está baseada nos comportamentos que se observan nos datos reais, pero sabendo que sucede exactamente en cada instante. Nesta tese é especialmente relevante a inclusión por primeira vez do xerador Madgraph [89]. Este é necesario para xerar correctamente os procesos que hipotéticamente dan lugar aos ALPs.

C.3.3 Reconstrución e selección de fotóns no LHCb

Utilízase para isto o ECAL. A información deste detector é crucial para identificar os fotóns, elemento central desta tese. A enerxía dos mesmos mídese pola súa interacción co material do detector, sendo a xeración de pares o proceso de interacción que domina. Os electróns subsecuentes interaccionan maiormente xerando fotóns creándose así un ciclo que da lugar a estruturas coñecidos como *fervenzas electromagnéticas*. A enerxía deste proceso aparece entón depositada en aglomeracións [93, 94], o que permite o seu estudo e posterior caracterización.

C.4 Un algoritmo alternativo para reconstruir muóns para a implementación do HLT1 en GPU

A finais do Run 2 o experimento LHCb foi mellorado na camapaña coñecida como *Upgrade I*, nesta eliminouse o trigger a nivel de hardware e decideuse empregar GPUs para implementar o HLT1, capaz de procesar eventos a unha taxa de 30 MHz. Para isto existe o proxecto Allen [97].

C.4.1 Traxectorias VELO-MUON

No marco de Allen, deseñouse un algoritmo de reconstrución de trazas tan só utilizando información do VELO e das estacións MUON. Estas trazas serían deixadas por muóns enerxéticos que atravesan o detector.

O algoritmo consiste en asociar ás traxectorias reconstruidas no sistema MUON, unha traxectoria reconstruida no VELO, obtendo así unha reconstrución da traxectoria completa destas partículas.

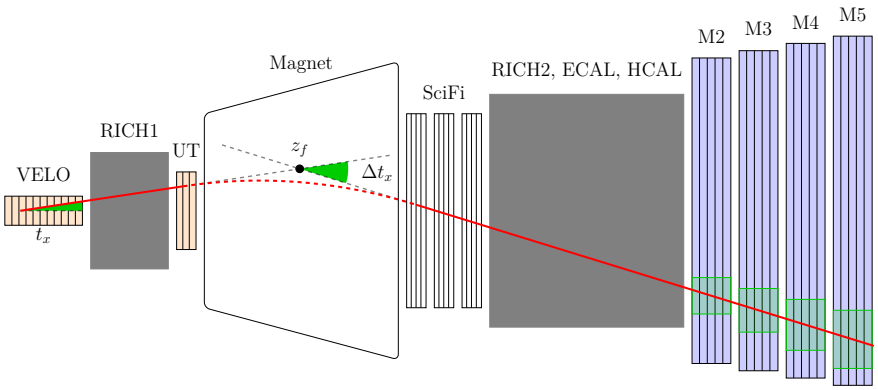

Figura C.4.1: Traxectoria dun muón nunha traza tipo VELO-Muón.

En primeiro lugar tómanse muóns reconstruidos no sistema MUON, os cales teñen sinal en todas as estacións do sistema (M2, M3, M4 e M5). Na primeira iteración considérase o punto máis externo no que hai sinal e trázase unha liña recta ca orixe de coordenadas. Búscase entón o seguinte sinal máis cercano e o terceiro e cuarto extrapólanse (Fig. C.4.1). Con isto obténse a traxectoria e a calidade da mesma a través dun test χ^2.

Para asignar a traza do muón no VELO ca trazas do muón no sistema de muóns temos en conta que a deflexión unicamente sucede no plano ZX. Avaliando a seguinte cantidade pódese realizar a asociación entre a traxectoria no VELO e a traxectoria no sistema MUON:

$$q^2 = \frac{(\Delta x)^2}{(10\ \text{mm})^2} + \frac{(\Delta y)^2}{(10\ \text{mm})^2} + (\Delta t_y)^2 \tag{C.4.1}$$

Este algoritmo proba ser válido para a reconstrución tanto de muóns que se xeran de xeito desplazados (por exemplo estos xerados polo decaemento raro $B_s^0 \to \mu^+\mu^-$) ou en muóns xerados no punto de interacción (como aqueles provenientes do decamento $J/\psi \to \mu^+\mu^-$) ambos decaementos son probados con sinais simuladas (Fig. C.4.2) e obtense unha medida do momento cunha parametrización cunha resolución de ata o 5% (Fig. C.4.3).

Este algoritmo pode ser unha alternativa ás traxectorias reconstruidas cos subdetectores VELO-SciFi-UT o cal pode ser especialmente útil en eventos con moitas trazas.

C.5 Búsqueda de ALP $\to \gamma\gamma$ no LHCb

Realízase a búsqueda de dous fotóns resonantes os cales poidan ser interpretados coma un ALP nos datos de 2018 cunha enerxía de $\sqrt{s} = 13$ TeV e unha luminosidade integrada de 2 fb^{-1}. Os resultados tamén se reinterpretan como un mesón B_s^0, B^0 ou η_b decaendo a dous fotóns. Estimando a senstividade na medida do coeficiente de ramificación ou na sección eficaz do proceso radiativo.

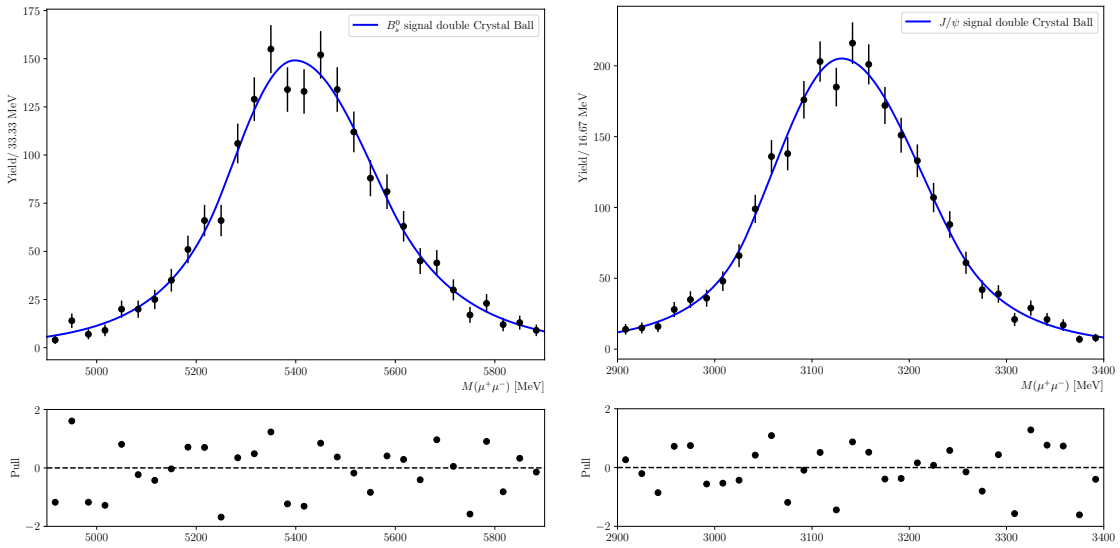

Figura C.4.2: Espectro de masa de $J/\psi \to \mu^+\mu^-$ e $B_s \to \mu^+\mu^-$ que amosa a capacidade para reconstruir sinais desprazadas ou non con respecto ao punto de interacción.

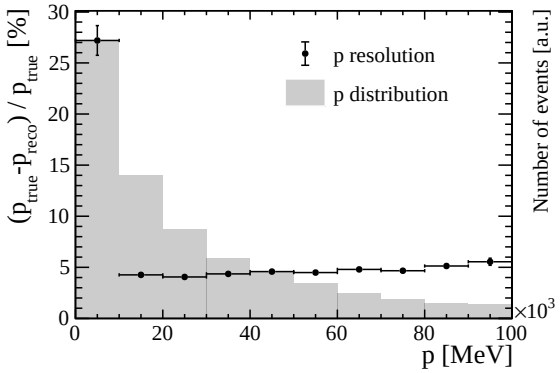

Figura C.4.3: Resolución no momentro dos muóns en datos simulados do decaemento $B_s^0 \to \mu^+\mu^-$.

Para realizar a análise utilízanse como canles de control $\eta \to \mu^+\mu^-\gamma$, $B^0 \to K^{*0}\gamma$ e $B_s^0 \to \phi\gamma$ por ser canles con polo menos un fotón no seu estado final.

C.5.1 Selección

Deséñase unha selección adicada a maximizar a significancia de sinal fronte a fondo, baseada nuns cortes no Stripping sobre variables cinemáticas e de PID de cada un dos fotóns así como de ambos á vez e nunha selección de trigger optimizada para a detección de fotóns [104, 106].

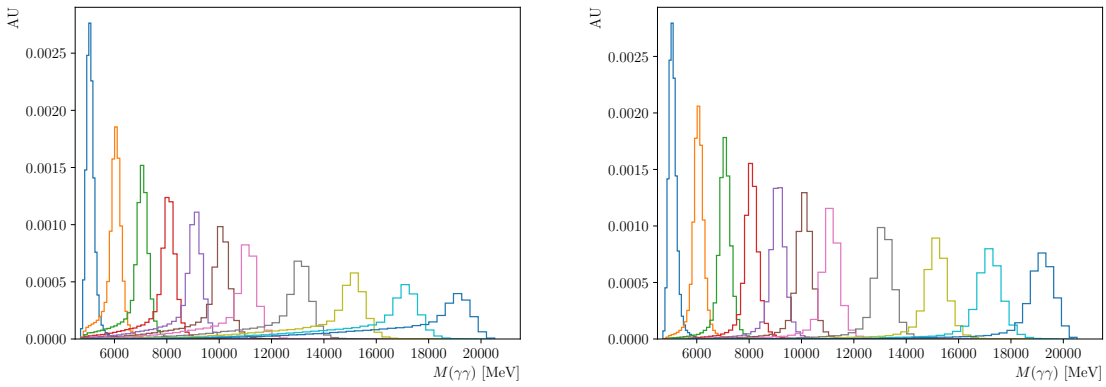

Figura C.5.1: Reconstrución para diferentes hipóteses de masa dos ALPs sen (esquerda) e con (dereita) veto por saturación.

Establécese un veto sobre os eventos nos cales un dos fotóns satura polo menos unha célula do calorímetro. Isto provoca que o pico de sinal sexa máis ancho pola esquerda (Fig. C.5.1) facendo que sexa case imposible discriminalo do fondo, isto é especialmente acusado para hipóteses de masas altas.

C.5.2 BDT

Desenvólvese un clasificador XGBoost [107] para aumentar a capacidade de discriminación. Entrénase cunha simulación de ALP $\to \gamma\gamma$ como sinal e con datos como fondo. Utilízanse variables de illamento tipo cono, xa que son especialmente boas para reducir o fondo combinatorio.

Obténse o mellor rendemento para o clasificador maximizando a seguinte figura de mérito [109] para diferentes hipóteses de masa:

$$\text{FOM} = \frac{\varepsilon}{3/2 + \sqrt{N_{\text{bkg}}}} \tag{C.5.1}$$

obtendo un corte de $BDT > 0.9$ para a procura de ALPs e $BDT_{B_s^0} > 0.85$ para a procura de $B_s^0 \to \gamma\gamma$.

C.5.3 Canles de control

Entrénase tamén un clasificador adicional para dúas das tres canles de control ($B_s^0 \to \phi\gamma$ e $B^0 \to K^{*0}\gamma$) baseado en variables topolóxicas sobre o vértice secundario desprazado e illamento sobre o fotón, obtendo un corte de BDT > 0.6.

Con isto realízanse os axustes pertinentes (Fig. C.5.2, Fig. C.5.3 e Fig. C.5.4) e aplícase a técnica do *sPlot* [110] aos datos e os sWeights obtidos que amosan a distribución de sinal.

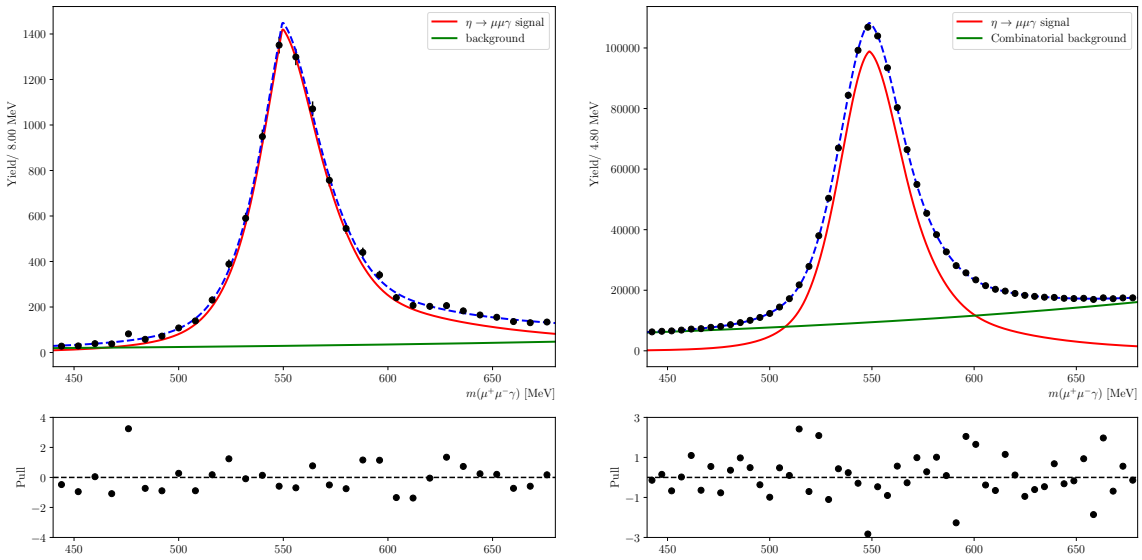

Figura C.5.2: $\eta \to \mu^+\mu^-\gamma$. Axustes de simulación (esquerda) e datos (dereita) para obter os sweights.

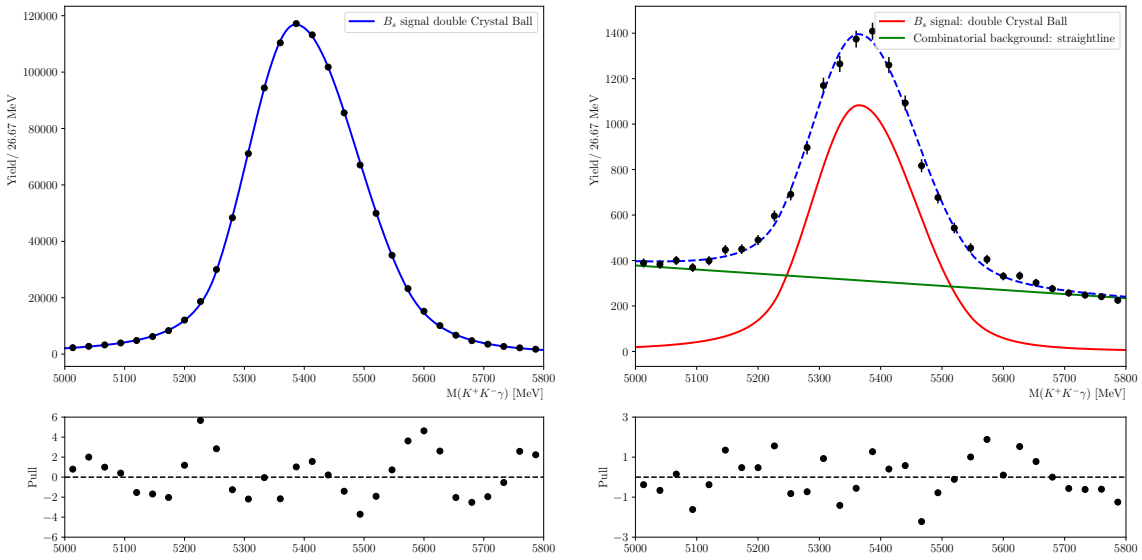

Figura C.5.3: $B_s^0 \to \phi\gamma$. Axustes de simulación (esquerda) e datos (dereita) para obter os sweights.

Estas canles de control son utilizados tanto para calibrar as eficiencias de sinal como para establecer a comparativa entre a resolución de enerxía do calorímetro en datos e simulación.

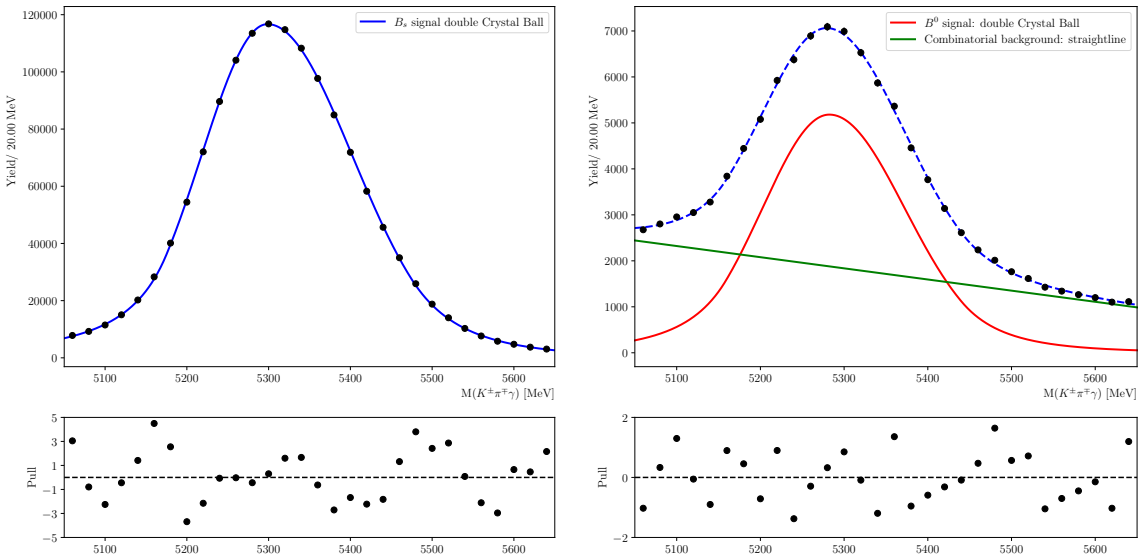

Figura C.5.4: $B_s^0 \to K^{*0}\gamma$. Axustes de simulación (esquerda) e datos (dereita) para obter os sweights.

C.5.4 Correlación entre fotóns

Estúdanse as características dos fotóns para saber se estes se comportan de xeito independente ou estan altamente correlacionados. Isto adquire especial relevancia xa que as canles de control só teñen un fotón presente no seu estado final.

Non se atopan candidatos nas simulacións de sinal nas cales ningunha parella de fotóns compartan célula do detector, co cal non se espera que a medida da enerxía ou do tipo de partícula sobre un deles afecte ao outro. Tamén se fan estudos para comprobar se as eficiencias tanto de indentificación como de medida de enerxía sufren correlacións cando se aplican aos dous fotóns ou son, contrariamente, independentes.

C.5.5 Correccións sobre as canles de control

Debido a pequenos defectos na simulación, algunhas distribucións cinemáticas non se comportan comon nos datos, aplícase un proceso de repesado para alinear os datos simulados co esperado nos datos (a comparativa entre as distribucións de masa atópase en Fig. C.5.2, Fig. C.5.3 e Fig. C.5.4), obtendo un mellor resultado tras o proceso (Fig. C.5.5).

C.5.6 Eficiencias e factor de normalización

As eficiencias calcúlanse a través de tests realizados sobre as canles de control (eficiencia xeométrica, eficiencia do stripping, eficiencia do corte en BDT) e calibrando segundo os datos

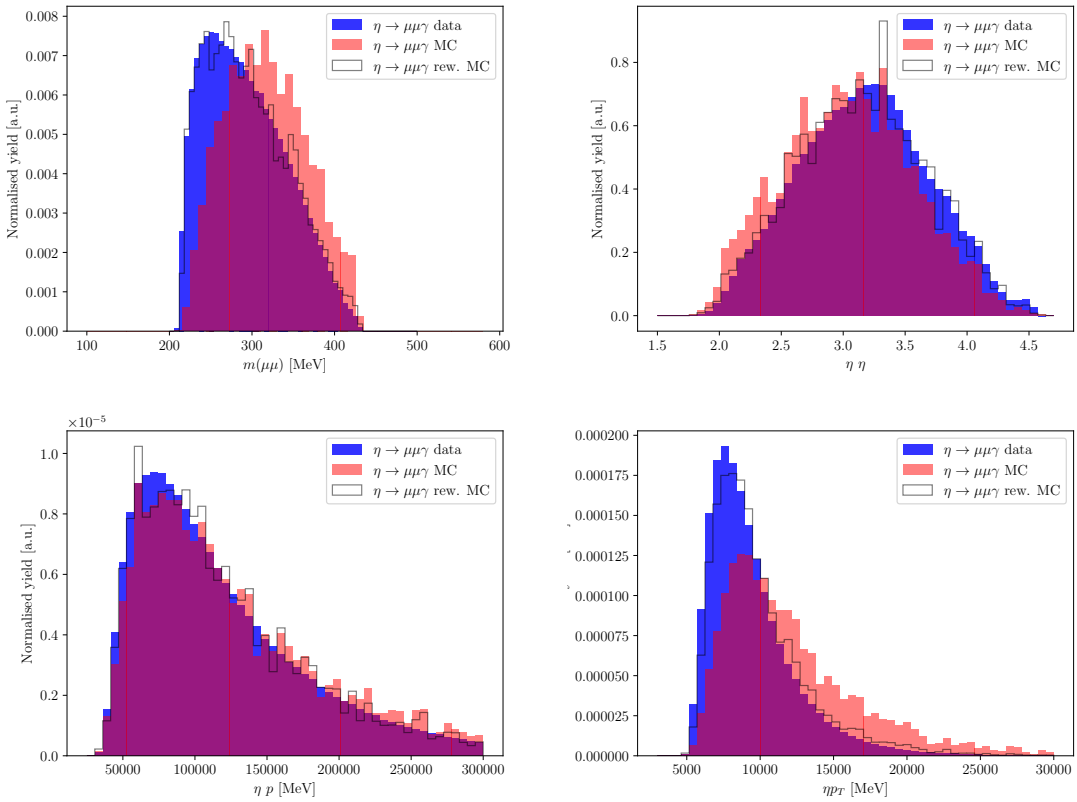

Figura C.5.5: Comparación das variables cinemáticas entre simulación e datos antes de despois de reponderar os datos.

de $\eta \to \mu^+\mu^-\gamma$ (cortes no tipo de partícula, trigger e efectos de saturación), obtendo así a táboa C.5.1.

Coñecendo o número de veces de que decae

$$\mathscr{B}(B_s^0 \to \phi\gamma, \phi \to K^+K^-) = (3.4 \pm 0.4) \times 10^{-5} \times (49.1 \pm 0.5)\% \tag{C.5.2}$$

podemos obter os factores de normalización para $B^0_{(s)} \to \gamma\gamma$:

$$\alpha_s = (2.3 \pm 0.5) \times 10^{-8}, \text{ and}$$
$$\alpha_d = (5.8 \pm 1.3) \times 10^{-9}.$$

C.5.7 Potenciais contribucións ao fondo

A maioría de fondos exclusivos que poderíamos esperar, como por exemplo $B^0 \to \pi^0\pi^0$, case non aparece no final, e estarían moi suprimidos polas contribucións de fondo combinatorio.

Táboa C.5.1: Eficiencias para ALPs (en %) calculadas de xeito secuencial . Os erros sistemáticos están incluidos. A columna reconstrución inclue tanto a eficiencia de reconstrución así como os cortes cinemático a nivel xerador.

Mass [GeV]	Generation	Reconstruction	GEC	PID	Saturation	L0	HLT1	HLT2	BDT	Total
$B_s^0 \to \gamma\gamma$	17.0 ± 2.5	7.34 ± 0.15	63.9 ± 0.6	58 ± 6	97.69 ± 0.06	95.7 ± 1.8	26.8 ± 0.4	82 ± 5	33.0 ± 0.6	0.031 ± 0.006
5.000	18.6 ± 0.5	8.55 ± 0.17	73.4 ± 2.6	60 ± 7	93.59 ± 0.19	96.6 ± 1.4	29.0 ± 0.4	82 ± 5	37.3 ± 0.6	0.056 ± 0.008
6.000	18.3 ± 0.5	12.33 ± 0.25	70.7 ± 2.8	59 ± 7	88.68 ± 0.35	96.4 ± 1.6	35.84 ± 0.30	81.6 ± 2.7	41.7 ± 0.6	0.098 ± 0.013
7.000	19.2 ± 0.5	15.24 ± 0.31	68.5 ± 3.0	59 ± 7	86.0 ± 1.0	96.5 ± 1.6	31.6 ± 1.5	81.6 ± 0.6	49.9 ± 0.7	0.125 ± 0.017
8.000	18.9 ± 0.5	18.1 ± 0.4	66.9 ± 3.1	59 ± 7	87.7 ± 1.3	96.6 ± 1.7	21.38 ± 0.33	81.6 ± 1.5	52.7 ± 0.7	0.104 ± 0.014
9.000	18.4 ± 0.5	20.8 ± 0.4	65.9 ± 3.2	59 ± 6	87.2 ± 1.8	96.9 ± 1.5	17.21 ± 0.26	81.6 ± 1.1	52.0 ± 0.7	0.091 ± 0.012
10.000	18.3 ± 0.5	23.1 ± 0.5	65.2 ± 3.3	58 ± 6	85.9 ± 2.0	97.3 ± 1.4	16.0 ± 0.4	81.6 ± 0.6	48.3 ± 0.7	0.085 ± 0.011
11.000	17.6 ± 0.5	24.9 ± 0.5	64.4 ± 3.3	58 ± 6	86.1 ± 2.1	97.6 ± 1.2	17.3 ± 0.7	81.6 ± 1.2	43.6 ± 0.6	0.085 ± 0.012
13.000	16.5 ± 0.5	27.5 ± 0.6	63.7 ± 3.4	58 ± 6	82.4 ± 1.9	98.1 ± 0.9	19.8 ± 0.4	81.6 ± 1.1	39.7 ± 0.6	0.087 ± 0.011
15.000	17.0 ± 0.5	29.2 ± 0.6	63.2 ± 3.4	58 ± 6	78.9 ± 1.1	98.4 ± 0.7	17.91 ± 0.13	81.6 ± 0.6	41.2 ± 0.6	0.085 ± 0.011
17.000	16.1 ± 0.5	30.3 ± 0.6	62.8 ± 3.4	58 ± 6	74.6 ± 1.5	98.6 ± 0.5	13.23 ± 0.17	81.6 ± 0.5	45.4 ± 0.7	0.064 ± 0.008
19.000	16.2 ± 0.5	30.1 ± 0.6	62.6 ± 3.5	58 ± 6	71.5 ± 3.0	98.70 ± 0.31	8.91 ± 0.14	81.6 ± 0.7	46.3 ± 0.8	0.042 ± 0.006
20.000	15.7 ± 0.5	22.1 ± 0.4	62.3 ± 3.4	58 ± 6	67 ± 11	98.70 ± 0.31	6.36 ± 0.16	81.6 ± 0.9	37.5 ± 1.0	0.0163 ± 0.0034

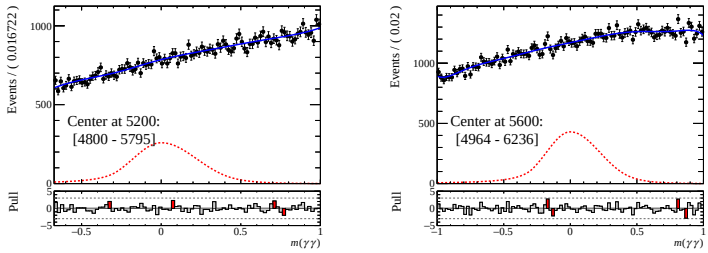

Figura C.5.6: Dous dos axustes realizados a unha potencial sinal e a un fondo combinatorio.

Sucede algo similar con fondos parcialmente reconstruidos que involucran pións ou fotóns, os cales non xeran ningunha estrutura similar a unha resonancia.

Utilizando un 4% dos datos modelízase a forma dos eventos de fondo e xérase unha mostra segue a forma atopada nos datos. Utilizarase este modelo para establecer a sensitividade de LHCb cos datos recollidos durante o ano 2018.

C.5.8 Resultados

Tendo en conta todo o exposto ata o de agora, búscase axustar os datos a unha función *double crystal ball* para o sinal e a un fondo combinatorio modelizado con polinomios de Chebychev. Na Fig. C.5.6 amósanse os resultados dos axustes utilizando os datos que recrean a hipótese de só fondo anteriormente modelizada.

As bandas de exlución esperadas no rango de masa $[5, 20]$ GeV atópanse na Fig. C.5.7. Espérase a mellor exclusión para o rango de masa $[5, 10]$ GeV.

Reinterpretando os datos, establecemos tamén a exclusión agardada para as canles de desintegración $Bs \to \gamma\gamma$ e $\mathrm{B}^0 \to \gamma\gamma$:

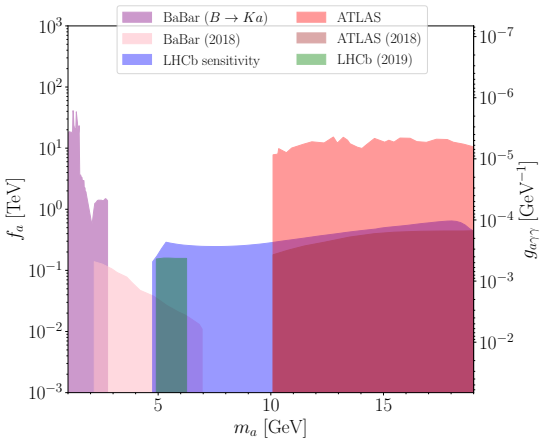

Figura C.5.7: Proxección dos novos límites tras esta tese.

$$\mathscr{B}(\mathrm{B}_\mathrm{s}^0 \to \gamma\gamma) < 2.5 \times 10^{-5}$$
$$\mathscr{B}(\mathrm{B}^0 \to \gamma\gamma) < 0.72 \times 10^{-5}.$$

C.6 Conclusións

Esta tese describe a primeira procura dunha resonancia decaendo a dous fotóns e reconstruida usando soamente información proveniente dos calorímetros do experimento LHCb. Os datos son recollidos durante o ano 2018 e corresponden a unha luminosidade integrada aproximada de $2.0\,\mathrm{fb}^{-1}$. Modelízase a sensibilidade e utilizando as eficiencias calibradas das diferentes hipóteses de masa obtense a sensitividade de LHCb para atopar estas resonancias na ventá de masa $[5, 20]$ GeV. Os mellores resultados nunha procura directa son esperados na ventá $[5, 10]$ GeV. Estes resultados tamén se reinterpretan os resultados como un mesón B_s^0 ou B^0 decaendo a dous fotóns e a sensibilidade agardada ponse en función do límite superior que se podería establecer no coeficiente de ramificación do decaemento a dous fotóns dos citados mesóns.

Tamén se describe unha secuencia de reconstrución de trazas con información do VELO e das estacións MUON. Estes algoritmos deséñanse para funcionar en GPUs cun uso óptimos dos recursos.

Bibliography

[1] "Partial symmetries of weak interactions," *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961.

[2] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[3] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc. C*, vol. 680519, pp. 367–377, 1968.

[4] M. Gell-Mann, "A Schematic Model of Baryons and Mesons," *Phys. Lett.*, vol. 8, pp. 214–215, 1964.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p.

[6] P. A. M. Dirac, "Quantum theory of emission and absorption of radiation," *Proc. Roy. Soc. Lond. A*, vol. 114, p. 243, 1927.

[7] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[8] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[9] A. Buckley, C. White, and M. White, *Practical Collider Physics*. IOP, 12 2021.

[10] K. G. Chetyrkin, J. H. Kuhn, and M. Steinhauser, "RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses," *Comput. Phys. Commun.*, vol. 133, pp. 43–65, 2000.

[11] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973.

[12] H. D. Politzer, "Reliable perturbative results for strong interactions?," *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, Jun 1973.

[13] A. D. Sakharov, "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe," *Pisma Zh. Eksp. Teor. Fiz.*, vol. 5, pp. 32–35, 1967.

[14] Q. R. Ahmad *et al.*, "Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory," *Phys. Rev. Lett.*, vol. 87, p. 071301, 2001.

[15] Q. R. Ahmad *et al.*, "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory," *Phys. Rev. Lett.*, vol. 89, p. 011301, 2002.

[16] V. C. Rubin, "A Century of Galaxy Spectroscopy," *The Astrophysical Journal*, vol. 451, p. 419, Oct. 1995.

[17] I. I. Bigi and A. I. Sanda, *CP violation*, vol. 9. Cambridge University Press, 9 2009.

[18] G. 't Hooft, "Symmetry Breaking Through Bell-Jackiw Anomalies," *Phys. Rev. Lett.*, vol. 37, pp. 8–11, 1976.

[19] C. Abel *et al.*, "Measurement of the Permanent Electric Dipole Moment of the Neutron," *Phys. Rev. Lett.*, vol. 124, no. 8, p. 081803, 2020.

[20] R. D. Peccei and H. R. Quinn, "CP conservation in the presence of pseudoparticles," *Phys. Rev. Lett.*, vol. 38, pp. 1440–1443, Jun 1977.

[21] R. D. Peccei and H. R. Quinn, "Constraints imposed by CP conservation in the presence of pseudoparticles," *Phys. Rev. D*, vol. 16, pp. 1791–1797, Sep 1977.

[22] V. A. Rubakov, "Grand unification and heavy axion," *JETP Lett.*, vol. 65, pp. 621–624, 1997.

[23] B. Bellazzini, M. Cliche, and P. Tanedo, "Effective theory of self-interacting dark matter," *Phys. Rev. D*, vol. 88, no. 8, p. 083506, 2013.

[24] B. Bellazzini, A. Mariotti, D. Redigolo, F. Sala, and J. Serra, "*R*-axion at colliders," *Phys. Rev. Lett.*, vol. 119, no. 14, p. 141804, 2017.

[25] X. Cid Vidal, A. Mariotti, D. Redigolo, F. Sala, and K. Tobioka, "New Axion Searches at Flavor Factories," *JHEP*, vol. 01, p. 113, 2019. [Erratum: JHEP 06, 141 (2020)].

[26] J. Redondo and A. Ringwald, "Light shining through walls," *Contemp. Phys.*, vol. 52, pp. 211–236, 2011.

[27] K. Ehret *et al.*, "New ALPS Results on Hidden-Sector Lightweights," *Phys. Lett. B*, vol. 689, pp. 149–155, 2010.

[28] R. Ballou *et al.*, "New exclusion limits on scalar and pseudoscalar axionlike particles from light shining through a wall," *Phys. Rev. D*, vol. 92, no. 9, p. 092002, 2015.

[29] P. Sikivie, "Experimental tests of the "invisible" axion," *Phys. Rev. Lett.*, vol. 51, pp. 1415–1417, Oct 1983.

[30] H. Primakoff, "Photo-production of neutral mesons in nuclear electric fields and the mean life of the neutral meson," *Phys. Rev.*, vol. 81, pp. 899–899, Mar 1951.

[31] E. M. Riordan, M. W. Krasny, K. Lang, P. de Barbaro, A. Bodek, S. Dasu, N. Varelas, X. Wang, R. Arnold, D. Benton, P. Bosted, L. Clogher, A. Lung, S. Rock, Z. Szalata, B. W. Filippone, R. C. Walker, J. D. Bjorken, M. Crisler, A. Para, J. Lambert, J. Button-Shafer, B. Debebe, M. Frodyma, R. S. Hicks, G. A. Peterson, and R. Gearhart, "Search for short-lived axions in an electron-beam-dump experiment," *Phys. Rev. Lett.*, vol. 59, pp. 755–758, Aug 1987.

[32] J. D. Bjorken, S. Ecklund, W. R. Nelson, A. Abashian, C. Church, B. Lu, L. W. Mo, T. A. Nunamaker, and P. Rassmann, "Search for neutral metastable penetrating particles produced in the slac beam dump," *Phys. Rev. D*, vol. 38, pp. 3375–3386, Dec 1988.

[33] M. A. e al., "Search for anomalous z $\rightarrow$ gamma gamma gamma events at lep," *Physics Letters B*, vol. 345, no. 4, pp. 609–616, 1995.

[34] G. Aad *et al.*, "Measurement of light-by-light scattering and search for axion-like particles with 2.2 nb^{-1} of Pb+Pb data with the ATLAS detector," *JHEP*, vol. 03, p. 243, 2021. [Erratum: JHEP 11, 050 (2021)].

[35] A. M. Sirunyan *et al.*, "Evidence for light-by-light scattering and searches for axion-like particles in ultraperipheral PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV," *Phys. Lett. B*, vol. 797, p. 134826, 2019.

[36] "Search for boosted diphoton resonances in the 10 to 70 GeV mass range using 138 fb^{-1} of 13 TeV pp collisions with the ATLAS detector," 11 2022.

[37] A. M. Sirunyan *et al.*, "Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV," *JHEP*, vol. 01, p. 097, 2018.

[38] A. Mariotti, D. Redigolo, F. Sala, and K. Tobioka, "New LHC bound on low-mass diphoton resonances," *Phys. Lett. B*, vol. 783, pp. 13–18, 2018.

[39] M. B. Gavela, J. M. No, V. Sanz, and J. F. de Trocóniz, "Nonresonant Searches for Axionlike Particles at the LHC," *Phys. Rev. Lett.*, vol. 124, no. 5, p. 051802, 2020.

[40] D. Aloni, Y. Soreq, and M. Williams, "Coupling QCD-Scale Axionlike Particles to Gluons," *Phys. Rev. Lett.*, vol. 123, no. 3, p. 031803, 2019.

[41] G. Aielli *et al.*, "Expression of interest for the CODEX-b detector," *Eur. Phys. J. C*, vol. 80, no. 12, p. 1177, 2020.

[42] C. Alpigiani *et al.*, "A Letter of Intent for MATHUSLA: A Dedicated Displaced Vertex Detector above ATLAS or CMS.," 7 2018.

[43] P. Agrawal *et al.*, "Feebly-interacting particles: FIPs 2020 workshop report," *Eur. Phys. J. C*, vol. 81, no. 11, p. 1015, 2021.

[44] G. Cowan, *Statistical data analysis*. 1998.

[45] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J. C*, vol. 71, p. 1554, 2011. [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[46] A. N. Kolmogorov, *Foundations of the Theory of Probability*. Chelsea Pub Co, 2 ed., June 1960.

[47] J. Neyman and E. S. Pearson, "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Phil. Trans. Roy. Soc. Lond. A*, vol. 231, no. 694-706, pp. 289–337, 1933.

[48] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Transactions of the American Mathematical Society*, vol. 54, no. 3, pp. 426–482, 1943.

[49] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. 60 – 62, 1938.

[50] L. Lyons, "Open statistical issues in Particle Physics," *The Annals of Applied Statistics*, vol. 2, no. 3, pp. 887 – 915, 2008.

[51] CERN, *PHYSTAT-LHC Workshop on Statistical Issues for LHC Physics: CERN, Geneva, Switzerland 27 - 29 Jun 2007. PHYSTAT-LHC Workshop on Statistical Issues for LHC Physics*, (Geneva), CERN, 2008.

[52] E. Gross and O. Vitells, "Trial factors for the look elsewhere effect in high energy physics," *Eur. Phys. J. C*, vol. 70, pp. 525–530, 2010.

[53] "LHC Machine," *JINST*, vol. 3, p. S08001, 2008.

[54] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, pp. 1–29, 2012.

[55] S. Chatrchyan *et al.*, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett. B*, vol. 716, pp. 30–61, 2012.

[56] W. W. Armstrong *et al.*, "ATLAS: Technical proposal for a general-purpose p p experiment at the Large Hadron Collider at CERN," 12 1994.

[57] "CMS, the Compact Muon Solenoid: Technical proposal," 12 1994.

[58] "ALICE: Technical proposal for a large ion collider experiment at the CERN LHC," 12 1995.

[59] S. Amato *et al.*, "LHCb technical proposal: A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays," 2 1998.

[60] R. Aaij *et al.*, "Observation of the doubly charmed baryon Ξ_{cc}^{++}," *Phys. Rev. Lett.*, vol. 119, no. 11, p. 112001, 2017.

[61] R. Aaij *et al.*, "Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates," *Phys. Rev. Lett.*, vol. 108, p. 111602, 2012.

[62] R. Aaij *et al.*, "Observation of CP Violation in Charm Decays," *Phys. Rev. Lett.*, vol. 122, no. 21, p. 211803, 2019.

[63] R. Aaij *et al.*, "Constraints on the $K_S^0 \to \mu^+\mu^-$ Branching Fraction," *Phys. Rev. Lett.*, vol. 125, no. 23, p. 231801, 2020.

[64] "Search for $K^0_{\mathrm{S(L)}} \to \mu^+\mu^-\mu^+\mu^-$ decays at LHCb," 12 2022.

[65] R. Aaij *et al.*, "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays," *Phys. Rev. Lett.*, vol. 115, p. 072001, 2015.

[66] R. Aaij *et al.*, "Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$," *Phys. Rev. Lett.*, vol. 122, no. 22, p. 222001, 2019.

[67] R. Aaij *et al.*, "Measurement of the W boson mass," *JHEP*, vol. 01, p. 036, 2022.

[68] R. Aaij *et al.*, "Search for Dark Photons Produced in 13 TeV pp Collisions," *Phys. Rev. Lett.*, vol. 120, no. 6, p. 061801, 2018.

[69] R. Aaij *et al.*, "Search for $A' \to \mu^+\mu^-$ Decays," *Phys. Rev. Lett.*, vol. 124, no. 4, p. 041801, 2020.

[70] R. Aaij *et al.*, "Search for heavy neutral leptons in $W^+ \to \mu^+\mu^\pm$jet decays," *Eur. Phys. J. C*, vol. 81, no. 3, p. 248, 2021.

[71] R. Aaij *et al.*, "Search for long-lived particles decaying to $e^\pm\mu^\mp\nu$," *Eur. Phys. J. C*, vol. 81, no. 3, p. 261, 2021.

[72] R. Aaij *et al.*, "Search for a dimuon resonance in the Υ mass region," *JHEP*, vol. 09, p. 147, 2018.

[73] R. Aaij *et al.*, "Searches for low-mass dimuon resonances," *JHEP*, vol. 10, p. 156, 2020.

[74] S. Amato *et al.*, *LHCb magnet: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2000.

[75] P. R. Barbosa-Marinho *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2001.

[76] L. Collaboration, "LHCb VELO Upgrade Technical Design Report," tech. rep., 2013.

[77] P. R. Barbosa-Marinho *et al.*, *LHCb inner tracker: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2002. revised version number 1 submitted on 2002-11-13 14:14:34.

[78] P. R. Barbosa-Marinho *et al.*, *LHCb outer tracker: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2001.

[79] L. Collaboration, "LHCb Tracker Upgrade Technical Design Report," tech. rep., 2014.

[80] A. C. Benvenuti *et al.*, "An Electromagnetic Shashlik calorimeter with longitudinal segmentation," *Nucl. Instrum. Meth. A*, vol. 432, pp. 232–239, 1999.

[81] C. Abellán Beteta *et al.*, "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC," 8 2020.

[82] A. Arefev, S. Barsuk, I. Belyaev, B. Bobchencko, L. Camilleri, V. Egorychev, Y. Gilitsky, A. Golutvin, O. Gouchtchine, I. Korolko, T. Kvaratskheliia, I. Machikhilian, M. Martemyanov, E. Melnikov, A. Morozov, M. Prokudin, D. Roussinov, V. Rusinov, A. Schopper, S. Schuvalov, A. Soldatov, E. Tarkovski, and K. Voronchev, "Beam Test Results of the LHCb Electromagnetic Calorimeter.," tech. rep., CERN, Geneva, 2008. revised version submitted on 2008-05-15 09:09:53.

[83] C. Coca, T. Preda, A. Rosca, I. Ajinenko, A. E. Dorokhov, R. I. Dzhelyadin, A. K. Konoplyannikov, V. Matveev, V. Novikov, O. P. Yushchenko, and Y. Ranyuk, "The hadron calorimeter prototype beam-test results," tech. rep., CERN, Geneva, 2000.

[84] S. Amato *et al.*, *LHCb RICH: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2000.

[85] P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2001.

[86] R. Antunes-Nobrega *et al.*, *LHCb trigger system: Technical Design Report*. Technical design report. LHCb, Geneva: CERN, 2003. revised version number 1 submitted on 2003-09-24 12:12:22.

[87] R. Aaij *et al.*, "Tesla : an application for real-time data analysis in High Energy Physics," *Comput. Phys. Commun.*, vol. 208, pp. 35–42, 2016.

[88] C. Bierlich *et al.*, "A comprehensive guide to the physics and usage of PYTHIA 8.3," 3 2022.

[89] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[90] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth. A*, vol. 462, pp. 152–155, 2001.

[91] J. Allison, K. Amako, *et al.*, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science*, vol. 53, no. 1, pp. 270–278, 2006.

[92] W. R. Nelson, T. M. Jenkins, R. C. McCall, and J. K. Cobb, "Electron-induced cascade showers in copper and lead at 1 gev," *Phys. Rev.*, vol. 149, pp. 201–208, Sep 1966.

[93] R. L. Workman *et al.*, "Review of particle physics," *Prog. Theor. Exp. Phys.*, vol. 2022, no. 8, p. 083C01, 2022.

[94] V. Breton, N. Brun, and P. Perret, "A clustering algorithm for the LHCb electromagnetic calorimeter using a cellular automaton," tech. rep., CERN, Geneva, 2001.

[95] M. Hoballah, "Measurement of the photon polarization using $B_s^0 \to \phi\gamma$ at LHCb.," 2015. Presented 03 Mar 2015.

[96] M. Calvo Gomez, E. Cogneras, O. Deschamps, M. Hoballah, R. Lefevre, S. Monteil, A. Puig Navarro, and V. J. Rives Molina, "A tool for γ/π^0 separation at high energies," tech. rep., CERN, Geneva, 2015.

[97] R. Aaij, J. Albrecht, M. Belous, P. Billoir, T. Boettcher, A. B. Rodríguez, D. vom Bruch, D. H. C. Pérez, A. C. Vidal, D. C. Craik, P. F. Declara, L. Funke, V. V. Gligorov, B. Jashal, N. Kazeev, D. M. Santos, F. Pisani, D. Pliushchenko, S. Popov, R. Quagliani, M. Rangel, F. Reiss, C. S. Mayordomo, R. Schwemmer, M. Sokoloff, H. Stevens, A. Ustyuzhanin, X. V. Cardona, and M. Williams, "Allen: A high-level trigger on GPUs for LHCb," *Computing and Software for Big Science*, vol. 4, apr 2020.

[98] C. LHCb Collaboration, "LHCb Upgrade GPU High Level Trigger Technical Design Report," tech. rep., CERN, Geneva, 2020.

[99] Wikipedia, "Frances Elizabeth Allen — Wikipedia, the free encyclopedia." `http://es.wikipedia.org/w/index.php?title=Frances%20Elizabeth%20Allen&oldid=149714313`, 2023. [Online; accessed 15-July-2023].

[100] NVIDIA, P. Vingelmann, and F. H. Fitzek, "Cuda, release: 10.2.89," 2020.

[101] A. M. Devices, "Rocm: Hip." `https://github.com/ROCm-Developer-Tools/HIP`, 2016.

[102] G. Bassi, L. Giambastiani, K. Hennessy, F. Lazzari, M. J. Morello, T. Pajero, A. Fernandez Prieto, and G. Punzi, "A FPGA-Based Architecture for Real-Time Cluster Finding in the LHCb Silicon Pixel Detector," *IEEE Trans. Nucl. Sci.*, vol. 70, no. 6, pp. 1189–1201, 2023.

[103] M. Ramos Pernas, "Search for $K^0_S \to \mu^+\mu^-$ and trigger developments at LHCb," 2020. Presented 24 Jul 2020.

[104] S. Benson and A. Puig Navarro, "Triggering $B^0_s \to \gamma\gamma$ at LHCb," tech. rep., CERN, Geneva, 2018.

[105] V. V. Gligorov, "Reconstruction of the Channel $B^0_d \to D^+\pi^-$ and Background Classification at LHCb (revised)," tech. rep., CERN, Geneva, 2007. revised version submitted on 2008-01-24 12:46:44.

[106] S. Benson, A. Casais Vidal, X. Cid Vidal, and A. Puig Navarro, "Real-time discrimination of photon pairs using machine learning at the LHC," *SciPost Phys.*, vol. 7, no. 5, p. 062, 2019.

[107] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, (New York, NY, USA), pp. 785–794, ACM, 2016.

[108] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. Springer Series in Statistics, New York, NY, USA: Springer New York Inc., 2001.

[109] G. Punzi, "Sensitivity of searches for new signals and its optimization," *eConf*, vol. C030908, p. MODT002, 2003.

[110] M. Pivk and F. R. Le Diberder, "sPlot: A statistical tool to unfold data distributions," *Nucl. Instrum. Meth.*, vol. A555, pp. 356–369, 2005.

[111] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02.

[112] V. Belyaev, V. Egorychev, and D. Golubkov, "Study of π^0/γ reconstruction efficiency with 2011 data," tech. rep., CERN, Geneva, 2012.

[113] L. Sestini and D. Lucchesi, "Measurement of the $Z \to b\bar{b}$ cross section in the forward region and determination of the b-jet energy scale, using b-dijet events with a balancing jet in pp collisions at 8 TeV.," 2016.

[114] R. Aaij *et al.*, "Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s}$ =13 TeV," *JHEP*, vol. 10, p. 172, 2015.

[115] R. Aaij *et al.*, "Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics," *JHEP*, vol. 04, p. 001, 2013.

[116] Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016," *Eur. Phys. J.*, vol. C77, p. 895, 2017.

[117] C. Abellán Beteta *et al.*, "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC," tech. rep., Aug 2020. All figures and tables, along with any supplementary material and additional information, are available at http://lhcbproject.web.cern.ch/lhcbproject/Publications/LHCbProjectPublic/LHCb-DP-2020-001.html (LHCb public pages).

[118] R. Aaij *et al.*, "Measurement of the B_c^- production fraction and asymmetry in 7 and 13 TeV pp collisions," *Phys. Rev.*, vol. D100, p. 112006, 2019.

[119] R. Aaij *et al.*, "Measurement of Υ production cross-section in pp collisions at $\sqrt{s}$ =13 TeV," *JHEP*, vol. 07, p. 134, 2018.

[120] R. Aaij *et al.*, "Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions," *Phys. Rev. Lett.*, vol. 118, p. 052002, 2017.

[121] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J. C*, vol. 71, p. 1554, 2011. [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[122] M. Kenzie, M. Karbach, T. Mombächer, M. Schlupp, and K. Schubert, "GammaCombo: A statistical analysis framework for combining measurements, fitting datasets and producing confidence intervals."

[123] R. Aaij *et al.*, "Measurement of the CKM angle γ from a combination of LHCb results," *JHEP*, vol. 12, p. 087, 2016.

[124] M. Williams, "A novel approach to the bias-variance problem in bump hunting," *JINST*, vol. 12, no. 09, p. P09034, 2017.

[125] D. Dutta *et al.*, "Search for $B_s^0 \to \gamma\gamma$ and a measurement of the branching fraction for $B_s^0 \to \phi\gamma$," *Phys. Rev. D*, vol. 91, no. 1, p. 011101, 2015.

[126] P. del Amo Sanchez *et al.*, "Search for the Decay $B^0 \to \gamma\gamma$," *Phys. Rev. D*, vol. 83, p. 032006, 2011.

[127] J. Abdallah, "Search for eta(b) in two-photon collisions at LEP II with the DELPHI detector," *Phys. Lett. B*, vol. 634, pp. 340–346, 2006.

[128] R. Mizuk *et al.*, "Evidence for the $\eta_b(2S)$ and observation of $h_b(1P) \to \eta_b(1S)\gamma$ and $h_b(2P) \to \eta_b(1S)\gamma$," *Phys. Rev. Lett.*, vol. 109, p. 232002, 2012.

[129] U. Tamponi *et al.*, "First observation of the hadronic transition $\Upsilon(4S) \to \eta h_b(1P)$ and new measurement of the $h_b(1P)$ and $\eta_b(1S)$ parameters," *Phys. Rev. Lett.*, vol. 115, no. 14, p. 142001, 2015.

[130] J.-P. Lansberg and M. A. Ozcelik, "Curing the unphysical behaviour of NLO quarkonium production at the LHC and its relevance to constrain the gluon PDF at low scales," *Eur. Phys. J. C*, vol. 81, no. 6, p. 497, 2021.

The main part of this thesis is devoted to describe a search for a diphoton resonance in the [5,20] mass region with pp collision data taken with the LHCb experiment during 2018 at a center of mass energy of 13 TeV. This work describes the first physics analysis using only photons detected by the calorimeter at LHCb. A precise calibration of the different selection steps is performed, particularly for the trigger selections. World best sensitivity for an ALP produced via gluon fusion and decaying to two photons is expected in the [5,10] GeV region. The result is also interpreted as a Bs0, B0 or EtaB meson decaying to two photons and expected upper limits on the branching fractions are reported. The thesis is also devoted to the development of a tracking sequence using the VELO and MUON stations for the GPU trigger.