FACULTY OF SCIENCES

SCHOOL OF PHYSICAL SCIENCES

Template Functional Forms for the Interpretation of Resonance Search Results at the Large Hadron Collider

Edmund TING

Supervisors:

Prof. Martin WHITE

Prof. Anthony WILLIAMS

Submitted towards the degree of:

Master of Philosophy

26 November 2019

Contents

1 Introduction **1**

2 Theoretical Context **3**
2.1 The Standard Model 3
2.2 Beyond the Standard Model 7
2.3 Particle collider experiments 8
2.3.1 Practical application of physics theories 8
2.3.2 The Large Hadron Collider 9
2.3.3 Presentation of LHC results 11

3 A General Functional Form **13**
3.1 Benchmark model 15
3.1.1 The Higgs characterization model 19
3.1.2 Comparison of Monte Carlo with analytic forms 21
3.2 A general parametrization 25
3.2.1 Casting $R(q^2)$ as a parameter 27
3.2.2 Taylor expansion of $R(q^2)$ 29
3.3 An alternative parametrization 34
3.3.1 Analysis of RC degeneracy 38
3.4 Summary 41

4 Global Optimization **43**
4.1 Statistical background 44
4.1.1 The likelihood function 44
4.1.2 Maximum likelihood estimation 44
4.1.3 Bayesian inference 46
4.1.4 MultiNest and the nested sampling algorithm 46
4.2 Fit of RC parameters using MultiNest 48
4.3 The $a_{i,j}$ parametrization 49
4.3.1 MultiNest $a_{i,j}$ toy closure tests 52
4.4 Fit window tests 57
4.4.1 Methodology and results 58
4.4.2 Constructing histograms from fit results 62
4.4.3 Physics model closure fit 64
4.5 Significance of a discovery 64
4.5.1 Toy scan over m and Γ 66
4.6 Summary 67

5 Detector Effects **69**
5.1 DELPHES 3 fast detector simulation . . . 70
5.2 Double-Sided Crystal Ball . . . 72
5.2.1 Narrow-width signal modelling . . . 72
5.2.2 Large-width modelling with PM . . . 73
5.3 Discrepancy in selection efficiencies . . . 77
5.4 General functional form (reconstructed) . . . 80
5.4.1 Fit to reconstructed MC . . . 81
5.5 Summary . . . 84

6 Conclusion **87**

Appendices **89**

A Taylor expansion of $\phi_{\mathcal{SB}}$ **91**

B Toy $a_{i,j}$ fit results, various a **93**

C Fit window tests **103**
C.1 Asimov toys of input distributions . . . 103
C.2 Results for $\Gamma/m = 3\%$ and 10% . . . 104
C.3 PM fit to PM toy results comparison . . . 106

Abstract

The search for unstable resonances produced in the proton–proton collision data of the Large Hadron Collider provides us an avenue through which physics beyond the Standard Model can be probed. A key challenge in current LHC searches is how one can model the interference effects that can severely distort the shape of the invariant mass distribution obtained from the decay products of the resonance, relative to the case where there is no interference. Such effects are strongly dependent on the beyond-Standard Model theory that gives rise to the resonance, which is unknown *a priori*.

This thesis presents a physically-motivated, yet model-independent, template functional form for the purpose of modelling interference effects, both in the characterization of positive discoveries and in the presentation of null results. We select a benchmark Higgs-like scalar resonance decaying to a pair of photons to illustrate the approach of an analysis, utilizing fully simulated Monte Carlo events and toy datasets to exhaustively test the general parametrization. A chapter is also dedicated to the study of detector smearing effects on experimental datasets.

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Edmund Ting
26 November 2019

Acknowledgements

A thank you is owed firstly to my supervisors, Martin White and Anthony Williams, who have shared their expertise with me through my years at university, providing valuable guidance and opportunities for academic growth. I am just as grateful to Lydia Roos, who was essentially my third direct advisor throughout my Masters programme. She has dedicated so much of her time to this project, sharing her expertise of experimental procedures and practices, and even travelled all the way from France specifically to focus some time into making progress on the project—twice! This thesis would also not have been possible without the theorists at CERN: Stefano Frixione, Fabio Maltoni, and Eleni Vryonidou, who (along with Martin and Lydia) devised the topic of this project in the first place. I was simply the fortunate student who happened to commence my candidature at the right time, and the right place, and was given the fantastic opportunity of working in their collaboration. I must also acknowledge CoEPP for providing invaluable computational resources, without which the results and fancy figures in this work would not have been brought to fruition.

Chapter 1

Introduction

The discovery of the Higgs boson at the Large Hadron Collider (LHC) in 2012 provided the final piece of experimental evidence for the particle spectrum of the Standard Model (SM) [1, 2]. Nevertheless, the SM is commonly understood to only be a manifestation of a broader theory, valid at low energies up to the Planck scale. This provides a motivation for seeking physics beyond the Standard Model (BSM) at the TeV scale, where it is hoped that evidence for the overarching theory can be revealed.

A common feature of BSM theories is the existence of new resonances, whose discovery and characterization can be achieved in particle collider experiments through the study of the invariant mass distributions of their visible decay products. However, experimental collaborations currently adopt a top-down strategy in their analysis and presentation of results. Under this approach, the statistically meaningful quantification of a discovery, or indeed a null result, can only be inferred from experimental data after the selection of a benchmark physics model (see, for example, [3–5]). In the case of positive results, a significance of the discovery is reported; for negative results, exclusion limits are presented on the product of the cross section for the resonance production and the branching ratio for decay into the final state of interest. While the top-down strategy prescribes a simple statistical procedure for an analysis, it is not without its drawbacks: the results that one obtains is highly dependent upon the benchmark model selected for the analysis, in that the same data might lead to different conclusions had a different benchmark been selected. Since the choice and treatment of the benchmark is ultimately handled by the experimental collaborations, it can be difficult for a theorist to generalize the reported results to a particular BSM model of interest to them.

Furthermore, it is often the case that the possible interference effects arising between the resonant signal and relevant SM backgrounds are neglected in an analysis. In the absence of interference, it is straightforward to take a lineshape for the signal (which depends on both the mass and width of the resonance), convolve it with the known detector resolution, and perform a fit to the observed data with a simple sum of the signal and background invariant mass distributions. Unfortunately, neglecting interference effects can be a pragmatic compromise rather than a well-motivated assumption. The details of signal–background interference are highly dependent on the new physics that gives rise to the resonance, which is, of course, unknown *a pri-*

ori; different models will generate different patterns and magnitudes of interference, the overall effect of which is to change the production rate of the resonance whilst distorting the invariant mass distribution of the decay products in such a way as to change the apparent mass of the resonance [6, 7]. The potential presence of these effects complicates both the interpretation of a new discovery in a resonance search, and the presentation of null results in the form of cross sections times branching ratio limits, which are not well-defined in the case of interference.

This thesis presents a practical approach for incorporating interference effects in resonance searches in a model-independent way. In contrast to the top-down approach currently employed, we consider the opposite viewpoint, namely that of a bottom-up approach, where the results of a search are presented with minimal assumptions of any particular physics model. The key idea is that, although the precise form of the resonant lineshape depends on the unknown parameters of an unknown physics model, its space of possible variations can be covered by any suitable choice of even and odd functions dictated by general Quantum Field Theory arguments. We develop a physically-motivated functional form that is capable of describing the distortions of the lineshape encountered in the presence of signal–background interference, and illustrate how this functional form can be adopted by the LHC and other collider experiments for the presentation of both positive and null results. In this way, any theoretical model can be quickly checked to be compatible or incompatible with experimental data by means of a simple computation in terms of the same parameters. A paper summarizing our findings is currently being prepared for submission to the Journal of High Energy Physics.

This manuscript is structured as follows. We first set the theoretical context for the work of this thesis in Ch. 2. A general, model-independent functional form will then be derived in Ch. 3. Its description will be tested against fully simulated Monte Carlo events of a benchmark signal model, which serves as our assumed choice of a scenario that exists in Nature.

In Ch. 4, we will explore global optimization techniques for visualizing the parameter space of the general functional form. We demonstrate the viability of the general functional form in describing a large variety of possible lineshapes exhibited under the benchmark signal model, a result that can be generalized to other signal models due to the wide range of behaviour covered.

Ch. 5 investigates a more realistic scenario in which distortions of the data from limited detector resolutions is considered. A description of such effects, in terms of the general functional form, will be presented and verified through the use of Monte Carlo events that include a simulation of generic detector effects.

Finally, we conclude in Ch. 6.

Chapter 2

Theoretical Context

In this chapter, we present a brief overview of the theoretical concepts required to understand the motivation for this work. We begin by introducing the Standard Model—the most complete theory of the microscopic world to date, which experimental efforts have shown to possess remarkable predictive power. Yet, it is not without its shortcomings; we describe some of the motivation for seeking extensions to the Standard Model, and present in more detail a particular example of such a theory. Finally, we discuss how physics beyond the Standard Model can be probed experimentally, focusing particularly on collider experiments and the detectors of the Large Hadron Collider.

2.1 The Standard Model

The Standard Model (SM) of particle physics is a theory developed to describe our universe at the most fundamental level. It is a study of the most elementary particles in nature—those which, as far as we know, can be treated as point-like and are not composed of other particles—and of their interactions under three of the four fundamental forces: the electromagnetic (EM), weak, and strong forces. Elementary particles can be categorized either as *fermions*, which are particles with half-integer spin, or *bosons*, with integer spin. The matter in our universe is constituted of *quarks* and *leptons*, which are spin-$1/2$ fermions. Their interactions under the three forces prescribed by the SM are achieved through the exchange of mediating force particles. These are the spin-1 gauge bosons: the photon (γ) for the EM force, exchanged between particles possessing *electric* charge; $W^{\pm}$ and Z for the weak force, between particles with weak charge (*weak isospin* and *hypercharge*); and gluon (g) for the strong force, for particles with *colour* charge. The spectrum of SM particles and some of their important properties are shown in Fig. 2.1[1].

Quantum Field Theory (QFT) provides the mathematical foundation for the SM. Under this framework, it is the excitations of quantum fields that manifest as the elementary particles of the model. Because of this, the SM is also a *gauge* theory, due to the fact that QFT dictates all massless vector fields to necessarily be gauge fields [8]. To be more specific, the SM is a gauge theory constructed

[1]Source: MissMJ, Wikimedia Commons (cropped).

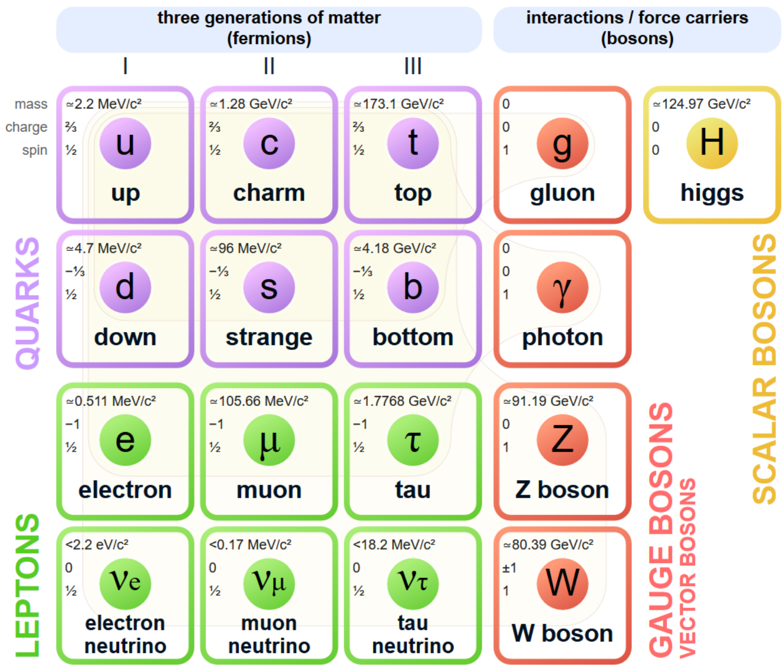

Figure 2.1: The particle spectrum of the Standard Model.

under the $SU(3)_C \times SU(2)_L \times U(1)_{Y_W}$ symmetry group. The $SU(3)_C$ subgroup describes the theory of quarks and gluons and the strong force, referred to as quantum chromodynamics (QCD), with the subscript, C, denoting colour charge. The EM and weak forces are described by a unified electroweak (EW) theory under the $SU(2)_L \times U(1)_{Y_W}$ group, where L indicates a coupling of the weak force to left-handed fields only, and Y_W is the weak hypercharge.

The full SM Lagrangian[2] can thus be written as a sum of QCD and EW contributions:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}}. \tag{2.1}$$

The pure QCD contribution is given by

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G_a^{\mu\nu} G_{\mu\nu}^a + \sum_f \bar{q}_f (i\gamma^\mu D_\mu - m_f) q_f, \tag{2.2}$$

where the Greek indices run over the four spacetime dimensions, while the Latin indices span the dimensionality of the Lie algebra (eight, in the case of $SU(3)_C$). The $G_{\mu\nu}^a$ represent field strength tensors of the gluon fields,

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c, \tag{2.3}$$

[2]To be pedantic, this is a Lagrangian *density*, with the Lagrangian being its integral over all spatial dimensions.

where g_s is the strong coupling constant, f^{abc} the structure constants of the strong symmetry group, and G^a_μ the gluon fields, which are generators of the group. The γ^μ are the Dirac matrices, and q_f (with adjoint $\bar{q}_f \equiv q_f^\dagger \gamma^0$) denotes quark fields of flavour f, with quark mass m_f. The covariant derivative here is given by

$$D_\mu = \partial_\mu + i g_s \frac{\lambda_a}{2} G^a_\mu, \tag{2.4}$$

where λ_a are the eight Gell-Mann matrices.

The EW contribution can itself be separated into distinct parts:

$$\mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}\,. \tag{2.5}$$

The first of these is the kinetic term for the gauge fields of $SU(2)_L \times U(1)_{Y_W}$,

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.6}$$

where $W^a_{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors of the $SU(2)_L$ weak isospin and $U(1)_{Y_W}$ weak hypercharge subgroups respectively:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g_W \varepsilon^{abc} W^b_\mu W^c_\nu, \tag{2.7}$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.8}$$

where g_W is the weak coupling constant, ε^{abc} the Levi-Civita symbol in three dimensions, and $a = 1, 2, 3$, in accordance with the 3-dimensionality of $SU(2)_L$. The gauge fields of Eqs. (2.7) and (2.8) are not directly physical; instead, it is their linear combinations that give rise to the fields of the weak and EM gauge bosons:

$$W^\pm_\mu = \frac{1}{\sqrt{2}} \left(W^1_\mu \mp i W^2_\mu \right), \tag{2.9}$$
$$Z_\mu = W^3_\mu \cos\theta_W - B_\mu \sin\theta_W, \tag{2.10}$$
$$A_\mu = W^3_\mu \sin\theta_W + B_\mu \cos\theta_W, \tag{2.11}$$

where θ_W is the weak mixing angle, which can be written in terms of the coupling constants of the weak isospin (g_W) and weak hypercharge (g_Y) symmetry groups,

$$\sin\theta_W = \frac{g_Y}{\sqrt{g_W^2 + g_Y^2}}\,. \tag{2.12}$$

The second term of Eq. (2.5) is the kinetic term for all SM fermions, including quarks since they also feel the weak and EM forces:

$$\begin{aligned}\mathcal{L}_{\text{fermion}} = {} & \left(\bar{u}_{iL}, \bar{d}_{iL}\right) i\gamma^\mu D_\mu \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix} + \left(\bar{\nu}_{iL}, \bar{l}_{iL}\right) i\gamma^\mu D_\mu \begin{pmatrix} \nu_{iL} \\ l_{iL} \end{pmatrix} \\ & + \bar{u}_{iR} i\gamma^\mu D_\mu u_{iR} + \bar{d}_{iR} i\gamma^\mu D_\mu d_{iR} + \bar{l}_{iR} i\gamma^\mu D_\mu l_{iR} + (\text{h.c.}),\end{aligned} \tag{2.13}$$

where $(u_{iL}, d_{iL})^T$ and $(\nu_{iL}, l_{iL})^T$ are $SU(2)_L$ doublets of left-handed quark and lepton fields respectively, and u_{iR}, d_{iR}, and l_{iR} are $SU(2)_L$ singlets of right-handed fields. The index, i, runs over the three generations of fermions, and for brevity, the Hermitian conjugate of each term has not been written in full. D_μ now denotes the covariant derivative for $SU(2)_L \times U(1)_{Y_W}$,

$$D_\mu = \partial_\mu + ig_W \frac{\sigma_a}{2} W_\mu^a + ig_Y \frac{Y_W}{2} B_\mu, \tag{2.14}$$

where the Pauli matrices, σ_a, are the generators of $SU(2)_L$, and Y_W is the weak hypercharge generating $U(1)_{Y_W}$.

The components of the SM Lagrangian discussed up to this point are relevant for the description of freely propagating and interacting gauge boson and fermion fields. However, note that the theory does not yet contain any mass terms relating to the gauge fields. This is a clear contradiction with Nature: it is known that while the gluon and photon are indeed massless, the $W^\pm$ and Z bosons of the weak force are not. It is precisely to explain this discrepancy with reality that the Higgs and Yukawa terms were introduced to Eq. (2.5); the former is responsible for inducing mass terms for the massive gauge bosons, while the latter introduces a correction to the fermion masses already appearing in the formulation thus far.

The process by which massive $W^\pm$ and Z bosons are allowed in the SM is known as the *Higgs mechanism*. This is achieved by introducing a new ("Higgs") field to the theory, with the crucial requirement that it attains a nonzero *vacuum expectation value* (vev); as such, this field is necessarily a scalar field, Φ, and transforms according to the particular gauge subgroup for which one requires massive gauge bosons. Thus, the field is chosen to be a $SU(2)_L$ doublet of complex scalar fields:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.15}$$

which yields the EW Lagrangian terms:

$$\mathcal{L}_{\text{Higgs}} = \left(D^\mu \Phi\right)^\dagger \left(D_\mu \Phi\right) - V(\Phi), \tag{2.16}$$

with the covariant derivative according to Eq. (2.14). The most general gauge invariant potential that yields a still-renormalizable theory is of the form

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2, \tag{2.17}$$

where $\mu^2 < 0$ and $\lambda > 0$ are real parameters, with the restrictions on their domains necessary to produce a nonzero vev. This is the famous Mexican hat potential of the Higgs field, with a ring of degenerate minima circularly about the origin in the complex plane, of radius $v = \sqrt{-\mu^2/\lambda}$ corresponding to the acquired vev. The Higgs mechanism gives rise to mass terms for the weak bosons upon spontaneous breaking of this global $U(1)$ symmetry (that of rotational invariance about the origin), which occurs when one chooses to expand the complex Higgs doublet about a particular minimum of the potential.[3]

Finally, the last term of Eq. (2.5) deals with the Yukawa interaction of the SM fermions and the Higgs field:

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} =& - \Gamma^u_{ij} \left(\bar{u}_{iL}, \bar{d}_{iL}\right) \Phi^\dagger u_{jR} + \Gamma^d_{ij} \left(\bar{u}_{iL}, \bar{d}_{iL}\right) \Phi d_{jR} \\ &+ \Gamma^l_{ij} \left(\bar{\nu}_{iL}, \bar{l}_{iL}\right) \Phi l_{jR} + \Gamma^\nu_{ij} \left(\bar{\nu}_{iL}, \bar{l}_{iL}\right) \Phi^\dagger \nu_{jR} + (\text{h.c.}),\end{aligned} \tag{2.18}$$

where Γ^f_{ij} are the 3×3 matrix Yukawa couplings, and the sum over Latin indices again runs over the fermion generations.

2.2 Beyond the Standard Model

While the SM has had incredible success in predicting many of the experimental results in high energy physics, most recently with the discovery of the Higgs boson in 2012 [1, 2], there remain questions to which it does not provide an answer. For example: why are there three generations of quarks and leptons? What leads to the large mass differences between SM particles? Why are the gauge couplings appearing in the SM so different? These questions prescribe the so-called "hierarchy problem" of the SM, and for such reasons, it is often regarded to be a low-energy manifestation of an overlying unified theory [9].

The SM also does not include a description of the fourth fundamental force in the universe: gravity. Instead, its effects are currently explained using general relativity, which is a theory completely separate from the SM. Astronomical observations have also revealed phenomena not explained by the SM, such as the asymmetry in baryon-antibaryon densities, or that most of the universe seems to be composed of invisible, "dark" matter.

These are but some of the reasons that motivate the formulation of beyond-Standard Model (BSM) physics theories. An example of such a theory is supersymmetry (SUSY), which postulates the existence of a fermionic, "supersymmetric" partner for each SM boson, and a bosonic partner for each fermion [10]. SUSY prescribes a means for resolving the hierarchy problem, providing a unified theory for the forces, and also produces particle candidates to explain dark matter—with such attractive features, it is naturally the focus of much research (e.g. [11–13]). However, SUSY requires the existence of two Higgs doublets, and this leads to one of the simplest extensions of the SM: that of including a second Higgs doublet in its description,

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix}, \qquad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix}, \tag{2.19}$$

[3]Often, one might also hear of "spontaneous breaking" of local gauge symmetry, but this is a misnomer owing to the gauge fixing that one usually imposes to remove the redundant degrees of freedom that arise when breaking the global symmetry; indeed, that local gauge symmetries cannot be spontaneously broken is a result known as Elitzur's theorem.

with both transforming under $SU(2)_L$. The introduction of a second complex doublet to the Higgs sector of the SM is itself a common topic of research (e.g. [14, 15]), and is known colloquially as a two Higgs doublet model (2HDM). Since the components of each doublet are complex fields, these two doublets contain eight degrees of freedom in total: three of these result in the longitudinal degrees of freedom of the weak gauge bosons, and the remaining lead to five physical scalar Higgs fields, in contrast to the single field in the case of only one doublet. This gives rise to five physical Higgs bosons: the CP-even neutral scalars, h and H, a CP-odd pseudoscalar, A, and two charged bosons, $H^{\pm}$. In addition to the yet-undiscovered SUSY partner particles, one can thus see that a common feature of BSM theories is the predicted existence of new, non-SM particles.

2.3 Particle collider experiments

The formulation of the SM as a gauge theory constitutes a rigorous and consistent mathematical framework that can be extended theoretically in many directions, as we have seen from the example of SUSY and the simple 2HDM extension. However, at the end of the day, any theory we construct must necessarily provide a description of the physical world; in this section, we shift the discussion from one of somewhat abstract Lagrangians and quantum field theory, to one that is more practically applicable for experimental endeavours. In particular, we discuss the physics of, and relating to, particle collider experiments, and how their results can be used to study particle physics theories (SM or otherwise).

2.3.1 Practical application of physics theories

Let us begin by introducing a few important quantities that can be predicted theoretically from a QFT, and which are also relevant to physical experimental scenarios. Of course, rigorous calculations can be performed under the mathematical framework of QFT, but our goal in this section will only be to present a qualitative sketch of the physics that is a prerequisite to understanding the motivation behind our work in later chapters.

As we have seen in the previous sections, a QFT, such as the SM, can be presented fully in terms of a Lagrangian (for example, Eq. (2.2)). Even without performing any detailed calculations, one can draw insight into the theory simply by inspecting its Lagrangian: using the QCD Lagrangian as an example, one can see terms involving quark-gluon fields, quark-quark fields, etc.

A useful method of visualizing the interactions of a theory are with the use of *Feynman diagrams*. Loosely, these are diagrams that depict the evolution of a particle(s) through spacetime. Particles whose fields are coupled in a theory Lagrangian are allowed to interact under that theory, with some assumed rate given by the coupling strength of the fields. As an example, Fig. 2.2 shows the Feynman diagram of an electron-positron ("Bhabha") scattering process, $e^+e^- \to e^+e^-$. The lines of the

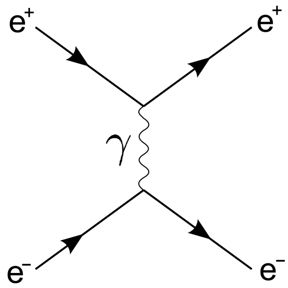

Figure 2.2: Feynman diagram of a Bhabha scattering process.

diagram represent particles traversing spacetime (with different linestyles conventionally used to differentiate types of particles), and typically one assumes a time axis aligned in some direction. The vertices represent their interaction: assuming the arrow of time to point rightwards, the diagram here depicts the EM interaction of an electron-positron pair, which scatter through the exchange of a virtual photon.

In addition to being a useful visualization tool, Feynman diagrams can be used to perform rigorous calculations under the mathematical jurisdiction of *graph theory*. Using the Lagrangian of a theory, one can construct a set of *Feynman rules*, which associates each line ("propagator") and vertex of the Feynman diagram with a corresponding mathematical expression. This can be used to calculate the *amplitude*, $\mathcal{A}$, of a diagram. One can draw an analogy of this quantity with the probability amplitude of quantum mechanics: in itself, it is not a physical quantity. Rather, it is the absolute square of the amplitude that is physically relevant. This is referred to as the *cross section* of the interaction,

$$\sigma \propto |\mathcal{A}|^2, \tag{2.20}$$

which etymologically originates from the classical analogue of scattering: the probability of two macro-objects colliding (e.g. a ball thrown at the side of a barn) is related to the cross-sectional area of those objects—the larger the barn, the more likely it is for the ball to hit. In line with this analogy, quantum cross sections are expressed as areas, typically in units of *barns* (b, with $1\,\mathrm{b} = 10^{-28}\,\mathrm{m}^2$) (which also owes its name to the analogy). While the classical interpretation of scattering does not apply to quantum particles, the cross section of a given Feynman diagram can nevertheless be interpreted as the rate of occurrence of the interaction it represents, with larger values indicating more favourable processes.

2.3.2 The Large Hadron Collider

CERN's Large Hadron Collider (LHC) [16], located near Geneva at the French–Swiss border, is the largest and most ambitious scientific experiment to date. As its name may suggest, its operational design involves the acceleration of hadrons (specifically protons) and facilitating their high energy collisions for the purpose of testing the theories of particle physics. This is done by accelerating two anti-parallel proton beams on a circular path to near light-speed, and allowing these beams to intersect at four different points on their rings. Its accelerator apparatus, with a circumference of approximately 27 km, lies in the underground tunnel previously occupied by the Large Electron-Positron (LEP) collider. The LHC began operations in 2008, and since then has produced data corresponding to collision energies of $\sqrt{s} = 7$, 8, and 13 TeV over its two runs, with plans to continue towards its design goal of $\sqrt{s} = 14$ TeV beam collisions. A notable outcome of its endeavours thus far is, of course, the discovery of the SM Higgs boson in 2012 [1, 2].

The accelerator itself, however, does not an experiment make; one requires methods for (indirectly) observing the collisions, by using detectors that are sensitive to the final state of their scattering. This is achieved at the LHC primarily by its four main detectors, each located at an intersection point of the proton beams: the ATLAS and CMS general purpose detectors, the ALICE heavy ion experiment, and LHCb for flavour physics studies [17–20].

Note that protons are not "clean" particles, but instead a QCD "soup" of quarks and gluons. Thus, an *event* at the LHC is defined as an interaction of the constituent quarks and gluons during a bunch crossing. The rate at which events occur depends on both the total cross section of all possible interactions and the *instantaneous luminosity* of the beams:

$$\frac{dN}{dt} = \sigma \mathcal{L}(t), \tag{2.21}$$

where, since N is dimensionless and σ has units of area, it follows that $\mathcal{L}(t)$ is expressed as $(\text{area} \times \text{time})^{-1}$. Instantaneous luminosity can be interpreted loosely as the cross-sectional density of protons in the beams; it is then intuitive that a larger $\mathcal{L}(t)$ results in a greater probability of an event occurring. The LHC is designed to achieve instantaneous luminosities of 10^{34} cm^{-2}s^{-1}. Another important and related quantity is the *integrated luminosity*, which is given by the integral of instantaneous luminosity over time,

$$\mathcal{L}_{\text{int}} = \int_{t_{\text{start}}}^{t_{\text{end}}} dt\, \mathcal{L}(t), \tag{2.22}$$

which is often expressed in units of inverse femtobarns (fb^{-1}), and is used to characterize the total output of a collider over some operational duration.

Since the detectors of the LHC do not observe the collisions of protons (or more accurately, of its constituents) directly, but instead the final state produced in the scattering, one can also write an event rate separately for each possible final state:

$$\frac{dN_{(f)}}{dt} = \sum_i \sigma_i^{(f)} \mathcal{L}(t), \tag{2.23}$$

where f denotes some final state, and the cross section of all possible interactions resulting in f contributes to the sum. Introducing a sum over f then yields Eq. (2.21).

Eq. (2.23) is what enables the study of BSM physics at collider experiments. As we have mentioned in the previous section, a common trait of BSM theories is the existence of new particles. The decay of these unstable resonances into SM particles can be predicted from the particular theory. Thus, if an excess of events not predicted by the SM is observed in some final state, one can infer a possible BSM contribution to $\sigma_i^{(f)}$ in Eq. (2.23), corresponding to

$$\sigma_{\mathrm{BSM}} \times \mathrm{BR}_{(f)}, \tag{2.24}$$

where σ_{BSM} is the production cross section of the BSM resonance, which is predictable from its theory, and $\mathrm{BR}_{(f)}$ denotes the branching ratio of the resonance decay into the particular final state, f.

To infer meaningful results from the collision data collected, it has to be converted to a form on which a statistical analysis can be performed. For this, it is useful to present the total number of collected events as a distribution against some observable, $\mathcal{O}$, of the final state particles:

$$\frac{dN_{(f)}}{d\mathcal{O}_{(f)}} = \frac{d\sigma_{(f)}}{d\mathcal{O}_{(f)}} \mathcal{L}_{\mathrm{int}}, \tag{2.25}$$

where $d\sigma_{(f)}/d\mathcal{O}_{(f)}$ is the *differential cross section* with respect to the observable, which of course satisfies:

$$\int_{-\infty}^{\infty} d\mathcal{O}_{(f)} \frac{d\sigma_{(f)}}{d\mathcal{O}_{(f)}} = \sigma_{(f)}. \tag{2.26}$$

In the context of BSM resonance searches, a particularly useful observable is the *invariant mass* of the final state particles, $m_{(f)}$. Simplistically, an excess of events at a particular invariant mass value, if assumed to arise due to a BSM resonant contribution, corresponds to the mass of the BSM particle. This is because the invariant mass is derived from the inner product of 4-momentum vectors:

$$m_{(f)} = \frac{P^{\mu} P_{\mu}}{c^2}, \tag{2.27}$$

where the P^{μ} are evaluated in the rest frame of the collision. Owing to the conservation of 4-momentum, one can then associate the invariant mass of the final state with that of the intermediary BSM resonance.

2.3.3 Presentation of LHC results

Realistically, the procedure of an analysis is not as simple as presented in the previous section: due to QFT complications, one does not see an excess of events akin to a δ-distribution at the mass of the assumed resonance, but instead a smeared excess across the invariant mass distribution near the expected resonance mass.[4] Ultimately, this distortion of the invariant mass lineshape can be predicted from theory, but herein lies the problem: to draw conclusions on the resonance from the distribution observed, one has to make a choice of a particular BSM theory, and assume the

[4]There is further smearing due to the physical limitations of detector equipment, but we will leave this discussion to a later chapter.

resonance to be one of the new particles it predicts. Only then can the lineshape of the data be characterized for the extraction of physical quantities—mass, width, etc. However, these quantities can vary greatly depending on the theory assumed, yielding a final result that can be heavily model-dependent and always non-generalizable to the plethora of existing BSM theories.

Another issue is associated with Eq. (2.23), which carelessly assumes that one can simply obtain a total cross section by summing the individual cross sections of all contributing processes. However, the cross section is obtained by evaluating the absolute square of an amplitude (Eq. (2.20)). As one might recall from quantum mechanics, a probability is calculated as the absolute square of the sum of wavefunctions—not the sum of the absolute square of each wavefunction (though one can sometimes show that the cross terms cancel anyway, yielding an equivalence between the two approaches). Similarly in this case, one has to consider the sum of amplitudes before squaring the result to obtain the cross section. Thus, a more technically correct form of Eq (2.23) can be written as:

$$\frac{dN_{(f)}}{dt} = \mathcal{L}(t) \sum_i \sigma^{(i \to f)} \propto \mathcal{L}(t) \sum_i \left(\left| \sum_j \mathcal{A}_j^{(i \to f)} \right|^2 \right), \tag{2.28}$$

where the sum over cross sections now applies only to interactions that begin with different initial states, i. For processes with the same initial and final state, the sum of each of their scattering amplitudes is obtained before performing the absolute square for the cross section. Importantly, the mixing of the $\mathcal{A}_j^{(i \to f)}$ in Eq. (2.28) gives rise to *quantum interference* between interactions of like initial and final states. Interference can modify the apparent cross section of an interaction, either constructively, destructively, or to distort the lineshape of the final state invariant mass distribution.

Currently, LHC resonance searches are performed using "simplified" models [21], while also neglecting possible interference effects that can arise between the assumed resonance and relevant SM "background" processes (i.e. SM interactions with the same initial and final state as the BSM signal). In the scenario that a positive signal is detected, the parameters of the chosen model are reported; for null results, an exclusion limit is presented on the production cross section times branching ratio ($\sigma \times \mathrm{BR}$) of the signal. This renders the reported results model-dependent, and thus possibly difficult for a given theorist in the field to generalize to their particular BSM model of interest, especially if interference is predicted to play an important role.

Chapter 3

A General Functional Form

In this chapter, we will derive a general functional form usable for the characterization of invariant mass distributions of any (directly visible) final state. To begin, we seek a baseline template describing an invariant mass distribution. The general interaction involving the production of an unstable resonance can be written as

$$i_1 \dots i_m \to X \to f_1 \dots f_n, \tag{3.1}$$

where $i_1 \dots i_m$ represent the m particles in the initial state, and $f_1 \dots f_n$ the n decay products of a particular decay mode of a new particle, X. We will refer to such an interaction as the "resonant" or "signal" process. In the scenario that no intermediary particle is produced, the interaction can be written as

$$i_1 \dots i_m \to f_1 \dots f_n, \tag{3.2}$$

which will be referred to as the "background", typically associated with physics that can be assumed to be well understood (for example, the SM). For the purpose of our work, any background contribution to the $f_1 \dots f_n$ final state that does not interfere with the signal process will be ignored.

The general helicity amplitude that describes such signal and background processes can be written as

$$\mathcal{A}(q^2) = \frac{\mathcal{S}(q^2)}{q^2 - m^2 + im\Gamma} + \mathcal{B}(q^2), \tag{3.3}$$

where q^2 denotes the four-momentum transfer of the scattering process; $\mathcal{S}(q^2)$ denotes the amplitude of the signal process, which produces an unstable intermediary particle with mass m and decay width Γ, up to a factor of the propagator which has been written explicitly; and $\mathcal{B}(q^2)$ represents the amplitude of all interfering SM background processes. Note that we assume a single resonance within the invariant mass range considered.

The signal and background amplitudes appearing in Eq. (3.3) can be separated in terms of a real component and a complex phase,

$$\mathcal{S}(q^2) = |\mathcal{S}(q^2)| \exp\left(i\phi_{\mathcal{S}}(q^2)\right), \quad \mathcal{B}(q^2) = |\mathcal{B}(q^2)| \exp\left(i\phi_{\mathcal{B}}(q^2)\right). \tag{3.4}$$

Substituting these expressions into the total amplitude, we then calculate its absolute square:

$$
\begin{aligned}
|\mathcal{A}(q^2)|^2 = &\frac{|\mathcal{S}(q^2)|^2}{(q^2-m^2)^2+m^2\Gamma^2} + |\mathcal{B}(q^2)|^2 \\
&+ \frac{2|\mathcal{S}(q^2)||\mathcal{B}(q^2)|}{(q^2-m^2)^2+m^2\Gamma^2}\left\{ \left(q^2-m^2\right)\cos\phi_{\mathcal{SB}} + m\Gamma \sin\phi_{\mathcal{SB}} \right\}. \qquad (3.5)
\end{aligned}
$$

The first term in this equation is the contribution from the signal process, and we see the familiar Breit-Wigner distribution arising from the squaring of the propagator, up to a constant factor:

$$
f_{\mathrm{BW}}(q^2) \equiv \frac{1}{(q^2-m^2)^2+m^2\Gamma^2}\,. \qquad (3.6)
$$

The second term is the contribution of the background; finally, the third term is a description of the interference, in terms of a relative phase between the signal and background amplitudes which has been defined as

$$
\phi_{\mathcal{SB}} = \phi_{\mathcal{S}} - \phi_{\mathcal{B}}. \qquad (3.7)
$$

While Eq. (3.5) has been derived assuming a single helicity configuration for the signal and background, one can extend its description to accommodate multiple helicities by making a distinction between the two phase terms: $\cos\phi_{\mathcal{SB}} \to \cos\phi_c$ and $\sin\phi_{\mathcal{SB}} \to \sin\phi_s$, where

$$
\cos\phi_c = \frac{\sum_i^N |\mathcal{S}_i||\mathcal{B}_i|\cos\phi^i_{\mathcal{SB}}}{\sqrt{\sum_i^N |\mathcal{S}_i|^2 \sum_i^N |\mathcal{B}_i|^2}}\,, \qquad (3.8)
$$

$$
\sin\phi_s = \frac{\sum_i^N |\mathcal{S}_i||\mathcal{B}_i|\sin\phi^i_{\mathcal{SB}}}{\sqrt{\sum_i^N |\mathcal{S}_i|^2 \sum_i^N |\mathcal{B}_i|^2}}\,, \qquad (3.9)
$$

with the summation running over the N helicity states. The amplitudes appearing in Eq. (3.5) are then implicitly understood to also denote a similar sum,

$$
|\mathcal{S}(q^2)|^2 = \sum_i^N |\mathcal{S}_i(q^2)|^2\,, \qquad |\mathcal{B}(q^2)|^2 = \sum_i^N |\mathcal{B}_i(q^2)|^2\,. \qquad (3.10)
$$

With this small correction to Eq. (3.5), we can now find a description for the final state invariant mass distribution. In the case of two particles in the initial and final states, the differential cross section describing a single event at energy q can be written as [22]

$$
\frac{d\sigma}{d\Omega_{\mathrm{CM}}} = \frac{1}{64\pi^2 q^2}\frac{|\boldsymbol{k}'|}{|\boldsymbol{k}|}|\mathcal{A}(q^2)|^2, \qquad (3.11)
$$

where $d\Omega_{\rm CM}$ is the differential solid angle in the centre of mass frame, and $\boldsymbol{k}$ and $\boldsymbol{k}'$ are the 4-momenta of incoming and outgoing particles respectively:

$$|\boldsymbol{k}'| = \frac{1}{2q}\sqrt{q^4 - 2q^2(m_{1'}^2 + m_{2'}) + (m_{1'}^2 - m_{2'})^2}\,, \tag{3.12}$$

$$|\boldsymbol{k}| = \frac{1}{2q}\sqrt{q^4 - 2q^2(m_1^2 + m_2^2) + (m_1^2 - m_2)^2}\,, \tag{3.13}$$

with m_i, $i = 1, 2$, denoting particle masses, and primed subscripts indicating outgoing particles. In the context of hadronic collisions such as those carried out by the LHC, Eq. (3.11) is modified by a weighting with a parton luminosity function,

$$\frac{d\sigma}{d\Omega_{\rm CM}} = \int dq^2\, \mathcal{L}(q^2)\frac{1}{64\pi^2 q^2}\frac{|\boldsymbol{k}'|}{|\boldsymbol{k}|}|\mathcal{A}(q^2)|^2, \tag{3.14}$$

necessary due to the composite nature of hadrons. This differential cross section can be re-parametrized to one in terms of the invariant mass:

$$\frac{d\sigma}{dq} = \int d\Omega_{\rm CM}\, \mathcal{L}(q^2)\frac{1}{32\pi^2 q}\frac{|\boldsymbol{k}'|}{|\boldsymbol{k}|}|\mathcal{A}(q^2)|^2. \tag{3.15}$$

Eq. (3.15), in combination with Eq. (3.5), provides a description for the invariant mass distribution of the decay products in a hadronic 2-to-2 process. Note that at this stage, the model-dependent terms in the amplitude are simply denoted collectively as $|\mathcal{S}(q^2)|$; a particular model for the signal has not yet been assumed.

In the following section, we introduce a benchmark signal model that will serve as our assumed choice of a scenario which exists in nature. The template of Eq. (3.15) will be used to obtain an analytical description of the benchmark signal, by finding an appropriate form for the model-dependent $|\mathcal{S}(q^2)|$ term in the amplitude. We then introduce a convenient code implementation of the benchmark physics model, which will be used to simulate pseudo-realistic datasets using Monte Carlo methods, and verify that the analytical forms obtained indeed parametrize the Monte Carlo samples generated.

We will then proceed to derive a general, model-independent functional form using the baseline description of Eq. (3.15), and perform various tests of its parametrization against the benchmark physics model in Secs. 3.2 and 3.3.

A summary of our results is presented in Sec. 3.4.

3.1 Benchmark model

We choose to perform our study in the diphoton final state. A Higgs-like signal model is chosen as the benchmark, prescribing a CP-even scalar resonance produced via 2-gluon fusion and decaying to two photons (Fig. 3.1, left). To find a description of the diphoton invariant mass distribution for this process, we first refer to Eq. (3.15). The factor $|\boldsymbol{k}'|/|\boldsymbol{k}|$ reduces to 1, as both the incoming (g) and outgoing (γ) particles

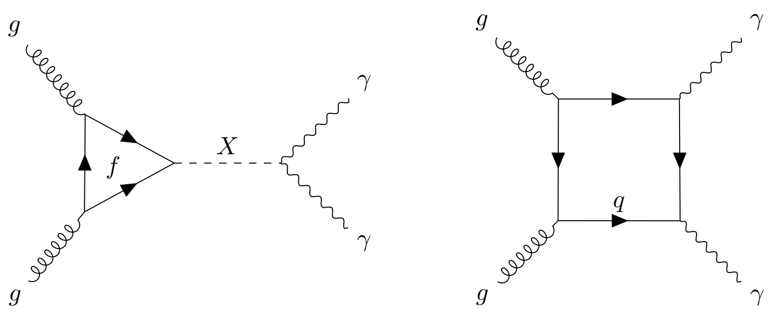

Figure 3.1: Left: the $gg \to X \to \gamma\gamma$ signal process. The resonance is a CP-even, spin-0 particle; the f denotes virtual fermions in the loop. Right: the leading order SM $gg \to \gamma\gamma$ background process with circulating quarks, q.

are massless (see Eqs. (3.12) and (3.13)). We will also neglect terms that do not contain explicit q-dependence; this yields a simple expression for the differential cross section,

$$\frac{d\sigma}{dq} \propto \frac{\mathcal{L}_{gg}(q^2)}{q} |\mathcal{A}(q^2)|^2, \tag{3.16}$$

where $\mathcal{L}_{gg}(q^2)$ is now specifically the gluon-gluon luminosity function. Eq. (3.16) holds generally for a 2-to-2 process with incoming gluons and outgoing massless particles.

The total amplitude of the signal process will receive a contribution each from the production loop, the effective decay vertex, and the propagator of the resonance. By comparison with Eq. (3.3), we thus have that

$$\mathcal{S}(q^2) = \mathcal{A}_{ggX} \mathcal{A}_{X\gamma\gamma}, \tag{3.17}$$

where $\mathcal{A}_{ggX}$ and $\mathcal{A}_{X\gamma\gamma}$ are amplitudes of the production loop and decay vertex respectively. The effective decay contributes a simple (q-dependent) factor,

$$\mathcal{A}_{X\gamma\gamma} \propto q^2. \tag{3.18}$$

The production term is more complicated; due to the fermion loop, a form factor is introduced:

$$\mathcal{A}_{ggX} \propto q^2 \sum_f A(\tau_f), \tag{3.19}$$

where f denotes the fermions in the loop, and is $A(\tau)$ the form factor, which for a CP-even resonance is [6]

$$A(\tau_f) = 2\tau_f(1 + (1 - \tau_f) g(\tau_f)), \tag{3.20}$$

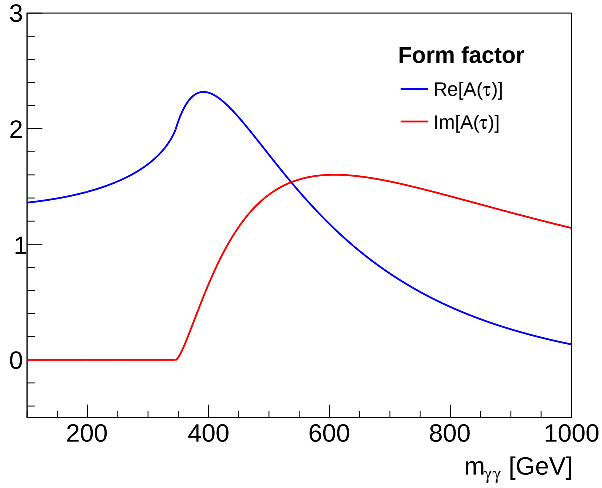

Figure 3.2: Real (blue) and imaginary (red) components of the production loop form factor containing a fermion with a 173 GeV mass.

where

$$\tau_f = \left(\frac{2m_f}{q}\right)^2, \tag{3.21}$$

and

$$g(\tau_f) = \begin{cases} \arcsin^2\left(\frac{1}{\sqrt{\tau_f}}\right) & \text{for } \tau_f \geq 1, \\ -\frac{1}{4}\left[\log\left(\frac{1+\sqrt{1-\tau_f}}{1-\sqrt{1-\tau_f}}\right) - i\pi\right]^2 & \text{for } \tau_f < 1. \end{cases} \tag{3.22}$$

The real and imaginary contributions of Eq. (3.20) are plotted against the diphoton invariant mass in Fig. 3.2, for a fermion of mass $m_f = 173$ GeV corresponding to the SM top quark.

From Eq. (3.16) and the first term of Eq. (3.5), we can thus posit the diphoton invariant mass distribution of the signal to be described by[1]

$$\frac{d\sigma_{\mathcal{S}}}{dq} \propto \mathcal{L}_{gg}(q^2) \times q^7 \times f_{\mathrm{BW}} \times |A(\tau)|^2, \tag{3.23}$$

with the Breit-Wigner given by Eq. (3.6). The gluon luminosity function can be approximated by a parametrization found using APFEL [23, 24] and the CT10 set of parton distribution functions (PDFs) [25], for a centre of mass energy of 13 TeV in accordance with LHC run-2 specifications [22]:

$$\mathcal{L}_{gg}(q^2) \propto \left(1 - \left(\frac{q}{13\,000}\right)^{1/3}\right)^{11.4407} \left(\frac{q}{13\,000}\right)^{-2.5785}. \tag{3.24}$$

[1]The sum of $A(\tau)$ over all fermions is left implicit.

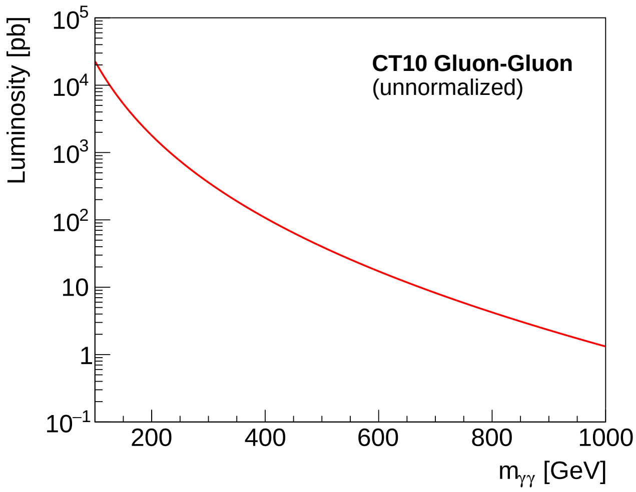

Figure 3.3: The shape of the gluon-gluon luminosity obtained with the CT10 PDF set, as a function of the invariant mass.

The shape[2] of this function is shown in Fig. 3.3.

An analytic form for the interference between the signal model and an interfering SM background process, such as the leading order $gg \to \gamma\gamma$ "box" diagram (Fig. 3.1, right), can be found using Eqs. (3.16)–(3.19) and the final term of Eq. (3.5). Direct substitution of the relevant terms yields:

$$\frac{d\sigma_{\mathcal{I}}}{dq} \propto \frac{\mathcal{L}_{gg}(q^2)}{q}\, 2 f_{\mathrm{BW}}\, q^4\, |A(\tau)| \left|\mathcal{B}(q^2)\right| \left\{ \left(q^2 - m^2\right) \cos\phi_{\mathcal{SB}} + m\Gamma \sin\phi_{\mathcal{SB}} \right\}. \quad (3.25)$$

While the amplitude of SM background diagrams can be calculated explicitly, these processes are not the main interest of a resonance search. Instead, an *ad hoc* parametrization of the background invariant mass distribution will typically be found; let us denote such a parametrization as

$$F_{\mathcal{B}}(q^2) \equiv \frac{d\sigma_{\mathcal{B}}}{dq} \propto \frac{\mathcal{L}_{gg}(q^2)}{q} \left|\mathcal{B}(q^2)\right|^2 . \quad (3.26)$$

The background amplitude appearing in Eq. (3.25) can then be re-cast in terms of this function:

$$\frac{d\sigma_{\mathcal{I}}}{dq} \propto \sqrt{\frac{\mathcal{L}_{gg}(q^2)}{q}}\, \sqrt{F_{\mathcal{B}}(q^2)}\, 2 f_{\mathrm{BW}}\, q^4\, |A(\tau)| \left\{ \left(q^2 - m^2\right) \cos\phi_{\mathcal{SB}} + m\Gamma \sin\phi_{\mathcal{SB}} \right\}. \quad (3.27)$$

Note, however, that the constant of proportionality in this equation is not arbitrary; if the normalizations of the signal (Eq. (3.23)) and background (Eq. (3.26)) differential cross sections are respectively denoted f_s and f_b, then we necessarily require

[2]Its normalization is not accounted for.

the interference normalization to be

$$f_i \equiv \sqrt{f_s f_b}\,. \tag{3.28}$$

A caveat to Eq. (3.28) is that it holds only when considering a single helicity configuration for the signal and background. As previously discussed, to accommodate multiple helicity states, it is sufficient to simply introduce two independent phases, whose interpretations are given by Eqs. (3.8) and (3.9).

3.1.1 The Higgs characterization model

A convenient computational implementation of the benchmark model (Fig. 3.1, left) is provided by the Higgs characterization (HC) model [26], a framework originally developed for studying the properties of the Higgs boson. In this section, we briefly summarize the relevant parts of the HC model, before describing, in more detail, our procedure for obtaining fully simulated Monte Carlo (MC) samples of the benchmark signal and the SM background.

For the loop-induced production of a new CP-even scalar particle, the Lagrangian,

$$\mathcal{L}_0^f = -\sum_{f=t,b,\tau} \bar{\psi}_f \left(c_\alpha \kappa_{Xff} g_{Xff}\right) \psi_f X, \tag{3.29}$$

describes the interaction of the scalar field, X, with fermion fields, ψ_f, within the framework. The parameter $c_\alpha \equiv \cos\alpha$ is the cosine of the mixing angle between CP-even and CP-odd states, with $c_\alpha = 1$ (0) for the CP-even (odd) case; κ_{Xff} are dimensionless coupling parameters; and g_{Xff} is the strength of the coupling between the resonance X and fermion f. The effective decay to $\gamma\gamma$ is described by the dimension-5 operator,

$$\mathcal{L}_0^V = -\frac{1}{4}\left[c_\alpha \kappa_{X\gamma\gamma} g_{X\gamma\gamma} A_{\mu\nu} A^{\mu\nu}\right] X, \tag{3.30}$$

where $A_{\mu\nu}$ is the electromagnetic field strength tensor, and $\kappa_{X\gamma\gamma}$ and $g_{X\gamma\gamma}$ are as previously described, the only difference being that the coupling is to two photons instead of fermions.

The HC model was employed for its FeynRules [27, 28] implementation of the Lagrangian terms of Eqs. (3.29) and (3.30), which provided an automated computation of the Feynman rules of the model. This was imported into MadGraph5_aMC@NLO [29–31] for a calculation of matrix elements and event generation, then interfaced with Pythia8 [32] for a parton shower simulation. Finally, Delphes 3.4.1 [33] was used to perform a generic fast detector simulation.

In the interest of increasing the statistical accuracy of our simulations while maintaining good computational efficiency, we chose to separately generate samples of signal, interference, and background events. Signal and interference events were generated for a $m = 400$ GeV resonance, with width $\Gamma/m = 10\%$. To obtain a well-defined relative phase between the signal and background processes, we also

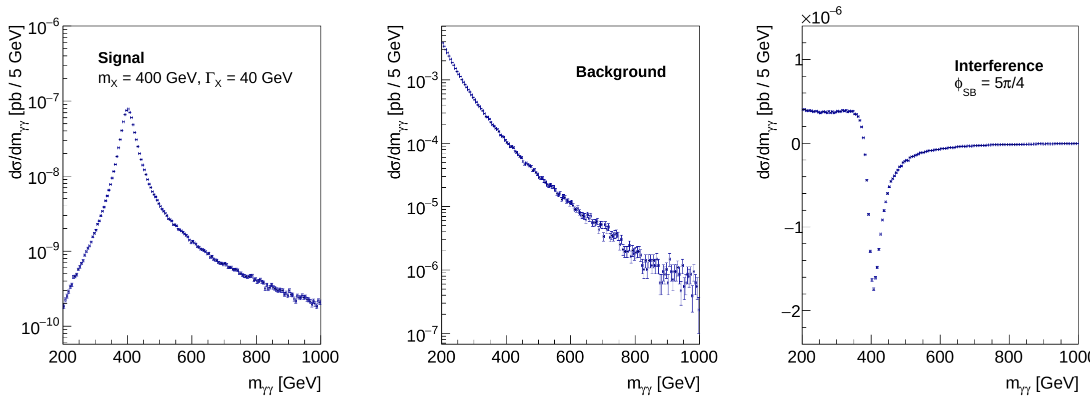

Figure 3.4: Cross-section-per-bin histograms constructed from MC samples. The signal corresponds to a scalar resonance with a 400 GeV mass and 40 GeV width. Interference with the SM continuum occurs with a relative phase of $\phi_{\mathcal{SB}} = 5\pi/4$.

introduced a modification to the HC model to define two new parameters, `KHttim` and `KHttre`, whose values respectively specify the sine and cosine of the nominal relative phase. However, the validity of this modification required a fermion in the resonant production loop to have a mass much greater than that of the resonance, to shift the non-trivial contributions of the form factor (in particular, of its imaginary component (see Fig. 3.2)) to higher invariant masses. Within the SM and in the unmodified HC model, circulating fermions are dominantly the t and b quarks and the τ lepton; we set the mass of the t quark to $m_t = 2$ TeV, such that contributions to the production loop form factor are almost entirely from its effect. The value `KHttim` = `KHttre` = -0.7071 was chosen, corresponding to an artificial relative phase $\phi_s = \phi_c = 5\pi/4 \equiv \phi_{\mathcal{SB}}$.

One million diphoton events each were generated for the signal and interference, and 500 000 events for the background sample. The relevant output of our simulations is a file produced by DELPHES 3.4.1, containing information on both the "generation" or "truth"-level events imported directly from the PYTHIA8 simulation, as well as the "reconstructed" events that result from the DELPHES 3.4.1 detector simulation. For now, we will work with the truth-level events, deferring the study of detector effects to Ch. 5.

CERN's ROOT data analysis framework [34] was used to extract and calculate the invariant mass of diphoton events in the DELPHES 3.4.1 particle record. We also extracted the weight of each event, the sum over which yields the cross section of the generated process,[3]

$$\sum_{\text{events}} w = \sigma. \tag{3.31}$$

This information was necessary to understand the contribution of interference photons to the overall invariant mass distribution; in particular, it is possible for the

[3]Depending on how the MC generator normalizes its events, it can also be necessary to divide by the total number of events generated.

events in this record to be associated with a negative weight, as interference can not only enhance but also suppress a signal. The weighted events of each sample were then sorted into binned histograms of their invariant mass (Fig. 3.4).

Typically, one needs to be conscious of relative normalisations when combining separately generated samples, since the contribution of the interference is physically dependent on the signal and background. It would be unphysical to directly combine the samples of 1 million signal and interference and 500 000 background events, for example. However, by taking into account the weight of each event when constructing the histograms, the number of generated events (which is arbitrary) has been converted into a cross section (which is only process-dependent, and does not depend on the number of events generated). Thus, histograms at the level of cross sections are already normalized according to the physics given as input to the event generator; the correct relative normalization between the three components is ensured, and a physically consistent total diphoton invariant mass distribution can be obtained by a simple sum of their histograms.

A comparison of our signal, interference, and background cross-section-per-bin histograms is shown in the top-leftmost plot of Fig. 3.5. While having a correct relative normalization, the signal and interference distributions are several orders of magnitude lower than that of the background. However, note that we can apply a scaling factor to enhance the signal and interference contributions; to preserve the correct relative normalization, a scaling of the interference histogram by a factor $\sqrt{\mu}$ necessarily accompanies a signal scaling by μ. This post-generation enhancement of our sample cross sections is equivalent to modifying the Yukawa coupling parameter, `KHtt`, within the HC model before generating any events. The remaining plots in Fig. 3.5 show a comparison of the three components for scaling factors μ= 75, 150, and 250.

3.1.2 Comparison of Monte Carlo with analytic forms

In this section, we will perform fits to the cross-section-per-bin histograms constructed from our MC samples using the analytical forms we have derived, to verify that they are indeed representative of the physics we wish to study. We will also summarize the results of [35], which tests several *ad hoc* template functional forms to find one that provides a suitable parametrization of the MadGraph5_aMC@NLO background.

ROOT's 1-dimensional histogram fitting function, `TH1::Fit`, was used to perform the fits. This implements the Minuit minimization program for the optimization of an objective function, `FCN`, using the Migrad algorithm [36]. By default, `FCN` corresponds to the chi-square statistic; assuming no correlation between bins, this is given by

$$\chi^2 = \sum_i \frac{(y(x_i;\boldsymbol{\theta}) - O_i)^2}{E_i^2}\,, \tag{3.32}$$

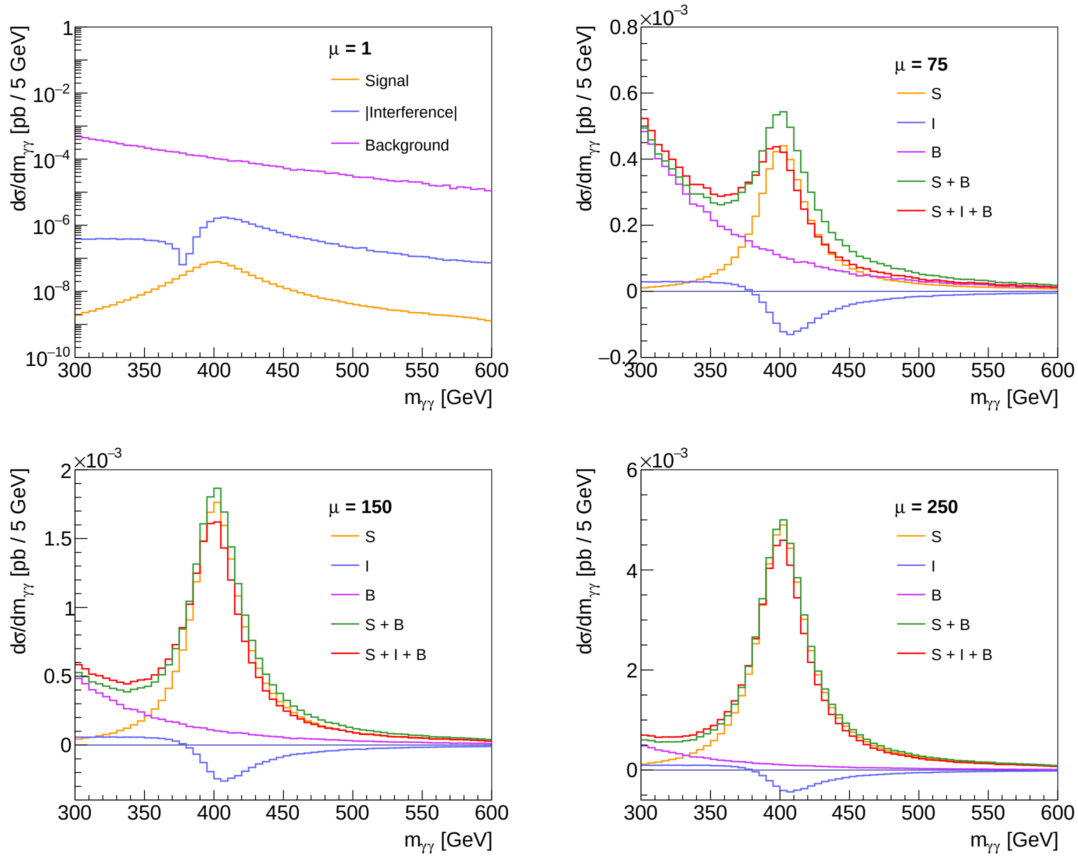

Figure 3.5: Top left: a comparison between the signal, (absolute value of the) interference, and background histograms at the cross section level, with no post-generation scaling factors applied. The remaining plots show the result of applying scaling factors, μ and $\sqrt{\mu}$, respectively to the signal and interference components.

where O_i denotes the value stored in the i^{th} bin, with uncertainty E_i. The function being fitted to the histogram, $y(x;\boldsymbol{\theta})$, with free parameters $\boldsymbol{\theta}$, is evaluated discretely at the central value of each bin. The minimization of the χ^2 yields best-fit values for each parameter, along with an estimate of their uncertainties [37].

The signal histogram was fitted using Eq. (3.23), with the form factor receiving only a single contribution from a fermion of mass 2 TeV. This corresponds to the choice we made for the t quark in the modification of the HC model. The free parameters of the fit are m, Γ, and f_s; these are respectively the mass and width of the resonance, appearing in the Breit-Wigner, and an overall normalization factor.

The result of the fit is presented in Fig. 3.6. Values extracted for the m and Γ parameters are in good agreement with their true values of 400 GeV and 40 GeV respectively, despite not being strictly within the one standard deviation (1σ) uncertainty ranges reported by the fitter. The metric of fit quality is the chi-square per degree of freedom (χ^2/ndf), where the number of degrees of freedom is equal to the number of datapoints (histogram bins) less the number of fitted parameters. Overall,

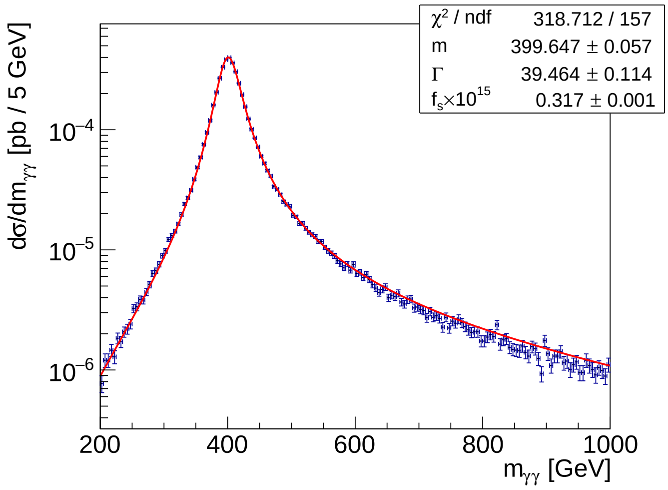

Figure 3.6: A fit of the physics model functional form (red) to the MC sample of signal events. Best-fit parameter values and uncertainties are shown in the box.

we find the analytical description of Eq. (3.23) to provide a good parametrization of the data distribution, with $\chi^2/\mathrm{ndf} \approx 2$.

For the background, instead of performing a possibly complicated calculation for the amplitude of the background processes, it is typical to seek an *ad hoc* parametrization,

$$F_{\mathcal{B}}(q^2) \equiv \frac{d\sigma_{\mathcal{B}}}{dq} \propto \frac{\mathcal{L}_{gg}(q^2)}{q} |\mathcal{B}(q^2)|^2. \qquad \text{(3.26 revisited)}$$

Here, we present a summary of the results from [35], which tested three different functional forms with the goal of identifying a suitable description of the MadGraph5_aMC@NLO diphoton background distribution. The first of these was based on a family of functional forms used in an analysis by the ATLAS collaboration for a centre of mass energy of 13 TeV [4],

$$F_{\mathcal{B}}^{\mathrm{ATLAS}}(q^2) \propto \frac{1}{q}\left(1-\left(\frac{q}{13\,000}\right)^{1/3}\right)^{A}\left(\frac{q}{13\,000}\right)^{B}, \tag{3.33}$$

with free parameters A and B. In particular, $A = 6.3$ and $B = -3.5$ are the values used in the ATLAS analysis. The second functional form was one used in an analysis by the CMS collaboration [5],

$$F_{\mathcal{B}}^{\mathrm{CMS}}(q^2) \propto q^{A+B\log q}, \tag{3.34}$$

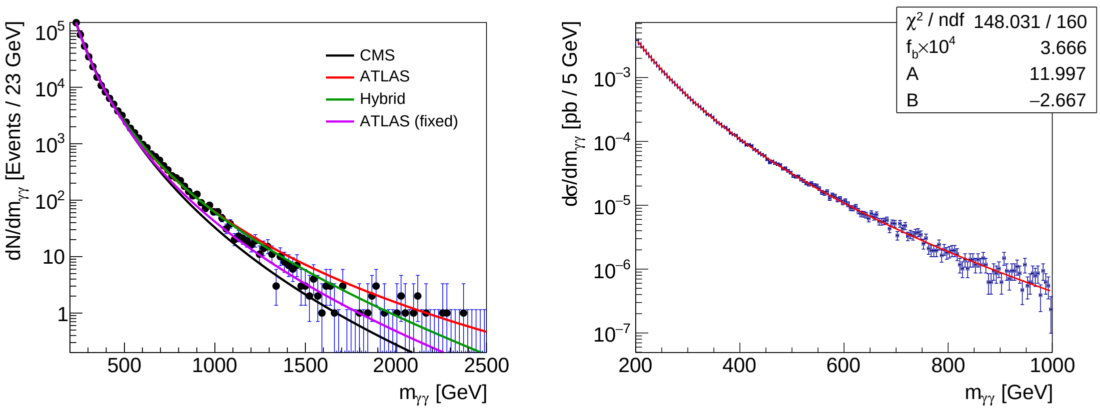

Figure 3.7: Left: a test of various template functional forms using a sample of MadGraph5_aMC@NLO background events. This result is extracted from [35]. Right: a fit of the background sample we generated, using the "ATLAS" functional form.

with free parameters again labelled as A and B. Finally, a combination of the ATLAS and CMS forms was defined:

$$F_{\mathcal{B}}^{\text{combined}}(q^2) \propto \frac{1}{q}\left(1-\left(\frac{q}{13\,000}\right)^{1/3}\right)^{A}\left(\frac{q}{13\,000}\right)^{B+C\log q}, \tag{3.35}$$

with three free parameters A, B, and C.

Eqs. (3.33)–(3.35) were tested against a large sample of approximately 10 million background-only events generated with MadGraph5_aMC@NLO. ROOT's RooFit fitting toolkit [38] was employed to perform fits over the 200–2500 GeV invariant mass range. Results are shown in Fig. 3.7 (left). The CMS parametrization (black) was found to underestimate the distribution generated by the MC, but both the ATLAS (red) and hybrid (green) forms provided good descriptions, performing equally well in the invariant mass regions below approximately 1200 GeV. At higher invariant masses, the hybrid parametrization was found to provide a slightly more accurate description of the data. The exact parametrization used by ATLAS is also plotted as a comparison (purple), but its description was found to be incompatible with the MadGraph5_aMC@NLO background.

Let us now make use of these results; since we have chosen to generate signal events at the $m = 400$ GeV mass point, the invariant mass region we will study lies in the range for which the hybrid and ATLAS parametrizations perform almost equivalently. We will thus use the simpler form of the latter, with an additional normalization parameter, f_b, to parametrize our histogram of background events. A good description is found (Fig. 3.7, right), verifying the suitability of this template for our study.

Finally, the analytical description of the interference (Eq. (3.27), with $F_{\mathcal{B}}(q^2)$ fixed to the background description found above) was verified against the histogram of interference events. The result of the fit is shown in Fig. 3.8. The characteristic

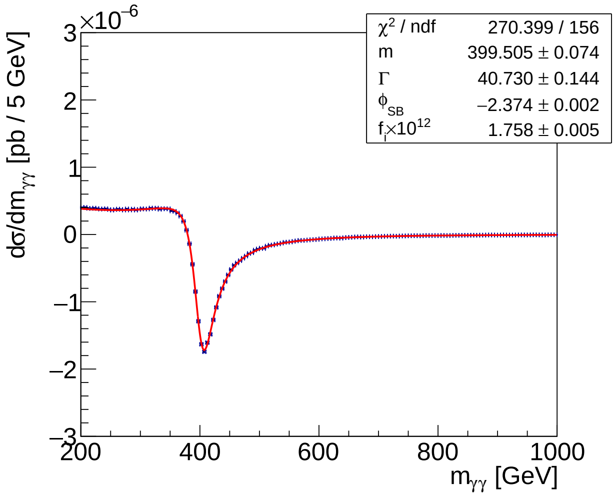

Figure 3.8: A fit of the interference lineshape.

peak-dip structure of the interference is captured well by the analytical form, extracting a relative phase of $\phi_{\mathcal{SB}} = -2.37$, the sine (cosine) of which is approximately -0.69 (-0.72). This result agrees well with the true value, $\cos\phi_{\mathcal{SB}} = \sin\phi_{\mathcal{SB}} = -0.7071$.

Note, however, an extracted normalization incompatible with Eq. (3.28) and the previous results of Figs. 3.6 and 3.7 (right). We find a suppression of the interference pattern by a factor of approximately $\sqrt{2/15}$, due to the reasons discussed previously.

3.2 A general parametrization

At the beginning of this chapter, we calculated a generic amplitude describing a resonant and background process (Eq. (3.5)), and found a simple expression for the differential cross section in terms of this amplitude (Eq. (3.16)). In this section, we will develop these equations to derive a general functional form capable of describing invariant mass distributions in a model-independent fashion. This general parametrization will then be studied using the benchmark physics model and the MC samples we have generated.

Let us begin by re-visiting the relevant equations:

$$|\mathcal{A}(q^2)|^2 = \frac{|\mathcal{S}(q^2)|^2}{(q^2-m^2)^2 + m^2\Gamma^2} + |\mathcal{B}(q^2)|^2 + \frac{2|\mathcal{S}(q^2)||\mathcal{B}(q^2)|}{(q^2-m^2)^2 + m^2\Gamma^2}\left\{ \left(q^2-m^2\right)\cos\phi_{\mathcal{SB}} + m\Gamma\sin\phi_{\mathcal{SB}} \right\}, \quad \text{(3.5 revisited)}$$

which is the absolute square of a general parton-level amplitude, and

$$\frac{d\sigma}{dq} \propto \frac{\mathcal{L}_{gg}(q^2)}{q}|\mathcal{A}(q^2)|^2, \quad \text{(3.16 revisited)}$$

the differential cross section for the $gg \to \gamma\gamma$ processes of interest. The differential cross section obtained by combining these equations can be expressed as a sum of the contributions from the individual components,

$$\frac{d\sigma}{dq} = \frac{d\sigma_{\mathcal{S}}}{dq} + \frac{d\sigma_{\mathcal{I}}}{dq} + \frac{d\sigma_{\mathcal{B}}}{dq}\,, \tag{3.36}$$

where the subscripts denote signal, interference, and background contributions to the cross section:[4]

$$\frac{d\sigma_{\mathcal{S}}}{dq} = \frac{\mathcal{L}_{gg}(q^2)}{q}\, f_{\mathrm{BW}} |\mathcal{S}(q^2)|^2, \tag{3.37}$$

$$\frac{d\sigma_{\mathcal{I}}}{dq} = \frac{\mathcal{L}_{gg}(q^2)}{q}\, 2 f_{\mathrm{BW}} |\mathcal{S}(q^2)||\mathcal{B}(q^2)| \Phi(\phi_{\mathcal{SB}}), \tag{3.38}$$

with f_{BW} denoting the Breit-Wigner distribution (Eq. (3.6)), and $\Phi(\phi_{\mathcal{SB}})$ a shorthand for the phase-related terms in the interference,

$$\Phi(\phi_{\mathcal{SB}}) = \left(q^2 - m^2\right) \cos\phi_c + m\Gamma \sin\phi_s\,, \tag{3.39}$$

where the now-independent "phases" ϕ_s and ϕ_c account for multiple helicity configurations. The background is again parametrized in an *ad hoc* manner,

$$\frac{d\sigma_{\mathcal{B}}}{dq} = \frac{\mathcal{L}_{gg}(q^2)}{q} |\mathcal{B}(q^2)|^2 \equiv F_{\mathcal{B}}(q^2). \tag{3.26 revisited}$$

There are two particular points of note. Firstly, because we only consider interfering background processes, the same parton luminosity function, $\mathcal{L}_{gg}(q^2)$, enters each of the three differential cross sections above. Secondly, the *ad hoc* description of the background differential cross section already encompasses information on the luminosity function. Thus, let us introduce a factor of 1 to the signal and interference contributions:

$$\frac{d\sigma_{\mathcal{S}}}{dq} = \frac{\mathcal{L}_{gg}(q^2)}{q}\, f_{\mathrm{BW}} |\mathcal{S}(q^2)|^2 \times \frac{|\mathcal{B}(q^2)|^2}{|\mathcal{B}(q^2)|^2} \tag{3.40}$$

$$= f_{\mathrm{BW}}\, \frac{|\mathcal{S}(q^2)|^2}{|\mathcal{B}(q^2)|^2} \times F_{\mathcal{B}}(q^2), \tag{3.41}$$

and

$$\frac{d\sigma_{\mathcal{I}}}{dq} = \frac{\mathcal{L}_{gg}(q^2)}{q}\, 2 f_{\mathrm{BW}} |\mathcal{S}(q^2)||\mathcal{B}(q^2)| \Phi(\phi_{\mathcal{SB}}) \times \frac{|\mathcal{B}(q^2)|^2}{|\mathcal{B}(q^2)|^2} \tag{3.42}$$

$$= 2 f_{\mathrm{BW}}\, \frac{|\mathcal{S}(q^2)|}{|\mathcal{B}(q^2)|}\, \Phi(\phi_{\mathcal{SB}}) \times F_{\mathcal{B}}(q^2). \tag{3.43}$$

[4] Proportionality factors have been absorbed into the amplitudes.

By doing so, Eq. (3.36) can be written fully as

$$\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\left[f_{\mathrm{BW}}R^2(q^2) + 2f_{\mathrm{BW}}R(q^2)\Phi(\phi_{\mathcal{SB}}) + 1\right], \tag{3.44}$$

where we have defined the ratio of signal to background amplitudes as

$$R(q^2) \equiv \frac{|\mathcal{S}(q^2)|}{|\mathcal{B}(q^2)|}. \tag{3.45}$$

A parametrization of the differential cross section in terms of this ratio removes the explicit appearance of the gluon luminosity function (or indeed, any parton luminosity in general), allowing us to sidestep the need to approximate its form. Additionally, Eq. (3.44) becomes generally applicable to the invariant mass distribution of any final state, as its description of a resonance and interference is completely relative to the background. This also means that it is able to describe invariant mass distributions not necessarily parametrized as a differential cross section. For example, experimental datasets are typically given in terms of a differential number of events,

$$\frac{dN}{dq}, \tag{3.46}$$

where the number of events is a simple scaling of the cross section by the integrated luminosity of the experiment,

$$N = \sigma \times \mathcal{L}_{\mathrm{int}}. \tag{3.47}$$

The *ad hoc* background description of such a dataset will thus already contain the appropriate factor of $\mathcal{L}_{\mathrm{int}}$ that scales Eq. (3.44) to describe a differential number of events.

3.2.1 Casting $R(q^2)$ as a parameter

The general template of Eq. (3.44) provides a description that is applicable to our summed ("BIS") histogram of background, interference, and signal MC events. However, there remains an ambiguity regarding the ratio, $R(q^2)$, and the exact form it should take. Owing to its definition in terms of the q-dependent signal and background amplitudes, it is generally an unknown, model-dependent function of the invariant mass. To maintain the model-independence of the functional form, a general approximation of the ratio is thus required.

As a starting point, we neglect the q-dependence to approximate the ratio as a parameter of the functional form, $R(q^2) \equiv R$. Under this assumption, fits of Eq. (3.44) to BIS histograms are 5-dimensional, in the parameters

$$\{m, \Gamma, R, \phi_c, \phi_s\}. \tag{3.48}$$

The background component of the functional form, $F_{\mathcal{B}}(q^2)$, is held completely fixed according to the description found in Fig. 3.7 (right). Fits were performed on his-

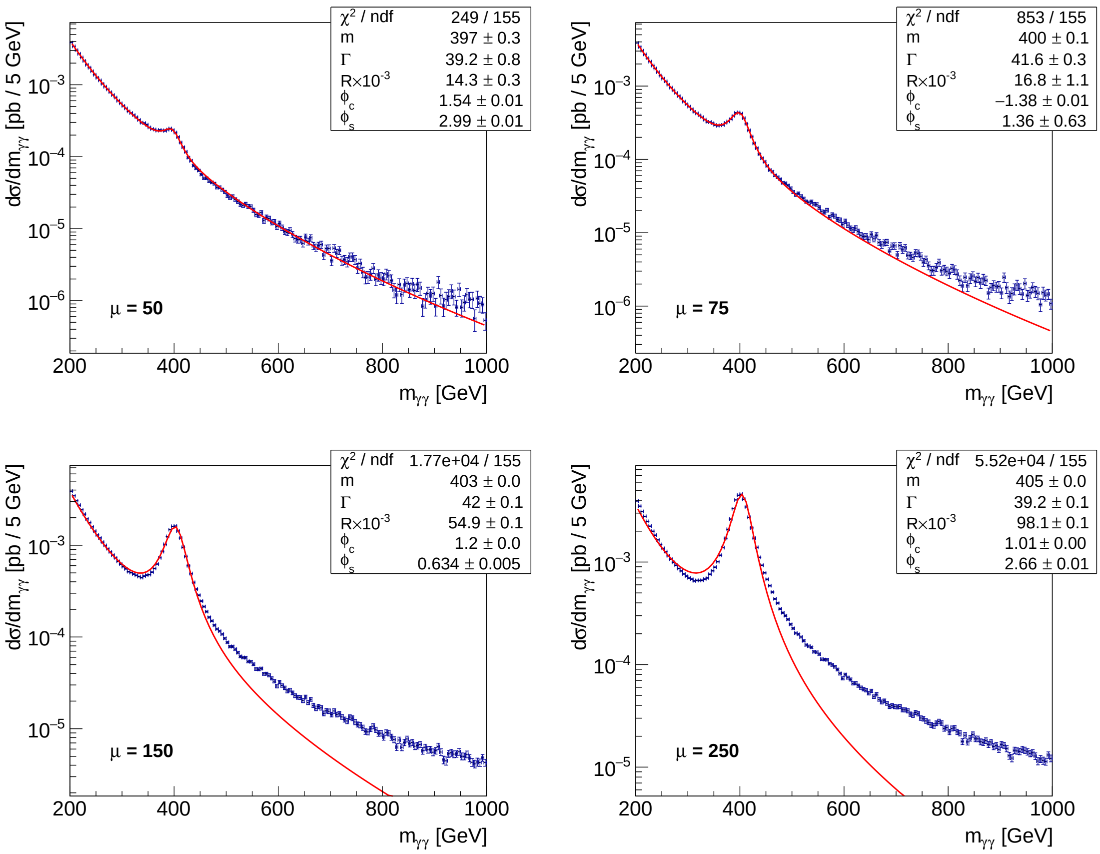

Figure 3.9: Fits over 200–1000 GeV of the summed invariant mass distributions, using the general functional form with the ratio, R, treated as a parameter. Various scaling factors, μ, have been applied to enhance signal and interference contributions to the histograms.

tograms with various scalings applied to the signal and interference contributions, corresponding to $\mu = 50$, 75, 150, and 250.

Results are presented in Fig 3.9. The data corresponding to $\mu = 50$ is fitted well, with only a slight underestimation of the data above approximately 800 GeV. The quality of fit decreases as the strength factor is increased, most noticeably in the high-q regions of the data: we see that the best-fit lineshapes converge to the fixed background contribution at high invariant masses. Additionally for the larger μ fits, the functional form is unable to fully capture the shape of the trough left of the signal peak. This suggests that the general functional form, under the assumption of a q-independent ratio parameter, is applicable only to a restricted invariant mass region about a peak; its form is not flexible enough to accommodate significant contributions of non-background events too far from the peak.

Upon inspection of best-fit parameter values extracted in the fits, we find m and Γ consistent with their true values of 400 GeV and 40 GeV respectively. Following from the definition of the phase-related parameters in Eq. (3.39), with a suppression of approximately $\sqrt{2/15}$ (see Sec. 3.1.2), we expect values of $\phi_c \approx 1.83$ and $\phi_s \approx -0.26$, but these are not reflected in our results. The values extracted for the phases also vary significantly for different μ. However, failing to correctly characterize the interference might not be too surprising, since one would expect to be sensitive to its effect only if a good parametrization of the data is found in the first place.

Let us, however, discuss the use of $\phi_{\mathcal{SB}}$ as a parameter, and the possible consequences of this from a computational point of view. Currently, fits are performed using ROOT's `TH1::Fit` method, which employs the MIGRAD algorithm to minimize an objective function (the χ^2). The minimization begins by calculating the first derivative of the objective function at a user-specified starting point, using this information to determine the direction in parameter space that decreases the χ^2. This process repeats until the fitter converges to the point of minimum χ^2. The important point of note is that finding the correct solution relies heavily on obtaining an accurate approximation of the first derivative [36, 37].

In setting up a fit, we are required to choose upper and lower limits for each free parameter in the fit. Because of the invariance of $\phi_{c/s}$ under the transformation[5]

$$\phi_{c/s} \to \phi_{c/s} + 2\pi k, \qquad k \in \mathbb{Z}, \tag{3.49}$$

we restricted the range of the phases to (approximately) $\phi_{c/s} \in [-\pi, \pi]$. Ideally, this would mean that the extracted $\phi_{c/s}$ corresponded to their true values within the range specified. However, a poorly chosen initial value can affect the calculation of the gradient of the χ^2, such that the minimization attempts to converge to a solution outside the specified range of values for the parameters. In such a scenario, it is possible for the algorithm to declare convergence of the fit without reaching the correct minimum in parameter space.

To avoid this issue, one can instead treat $\cos\phi_c$ or $\sin\phi_s$ as parameters. By doing this, the invariance of $\phi_{c/s}$ no longer plays a role, and we are simply left with parameters that take a range of physical values in $[-1, 1]$.

3.2.2 Taylor expansion of $R(q^2)$

In the previous section, we found that neglecting q-dependence in the ratio of amplitudes led to generally poor descriptions of BIS histograms over large invariant mass ranges, especially for those receiving large signal (and interference) contributions. However, the goal of an analysis is to extract the physical parameters of a resonance; it can be sufficient to consider only a limited invariant mass window, if accurate predictions of the parameters are obtainable.

Nevertheless, being able to parametrize a larger invariant mass window can help to constrain the parameters of the functional form, especially if there are limited number of events available in a dataset. Thus, it might be desirable to inject some q-dependence to the ratio of amplitudes in a model-independent fashion. Let us do this now by performing a Taylor series expansion for $R(q^2)$ in the variable q^2,

[5] $\phi_{c/s}$ refers to ϕ_c and ϕ_s.

calculated about the square of the resonance mass, m^2:

$$R(q^2) = R_{(0)} + R_{(1)}(q^2 - m^2) + \mathcal{O}\left((q^2 - m^2)^2\right). \tag{3.50}$$

By construction of the Taylor series, the expansion coefficients are, analytically,

$$R_{(0)} = R(m^2), \qquad R_{(1)} = \left.\frac{\partial R(q^2)}{\partial q^2}\right|_{q^2=m^2}, \tag{3.51}$$

and so on for higher order terms. However, in practice, these coefficients are treated as parameters of the functional form. Our results in the previous section are, of course, equivalent to the case in which we truncate Eq. (3.50) at the zeroth order. In principle, one would also perform a similar expansion for the relative phase, $\phi_{\mathcal{SB}}$, as it is generally a q-dependent quantity; we present such a consideration in App. A, but will neglect it for the studies presented in the main body of this manuscript.

The general functional form will thus be $4+j$-dimensional in its parameters, where $j \in \mathbb{Z}^+$ is the order past which we choose to truncate the expansion of $R(q^2)$. This is normally chosen at the order which satisfied, to some desired precision,

$$R_{(i+1)}(q^2 - m^2)^{i+1} \ll R_{(i)}(q^2 - m^2)^i \quad \forall \quad i \geq j. \tag{3.52}$$

However, a choice cannot be made without prior knowledge of the expansion coefficients, $R_{(i)}$, and the resonance mass, m. Furthermore, the expansion includes factors of $(q^2 - m^2)^i$, which is a function whose value increases quickly with q. Because of this, the problem also depends on the invariant mass range being considered, as a larger window will lead to the inclusion of increasingly higher order terms.

The two decisions we need to make in a given fitting problem, then, are:

1. The invariant mass range that the functional form is to fit, and

2. The order at which the Taylor expansion of $R(q^2)$ is to be truncated.

These points are related in that making a choice for one of these points fixes the "correct" choice for the other. Thus, there are two ways to proceed. The first is to decide on the invariant mass range to be considered in an analysis and begin with the zeroth order approximation, increasing the order of the expansion until Eq. (3.52) is found to hold sufficiently well. However, with no cause to discard any events a priori, one would simply choose to perform the fit over as large an invariant mass range as possible, with limitations coming only from the number of events measured in the experiment. In this scenario, the Taylor expansions are probably not ideal, as one would need to include very high order terms to accommodate the large fit window, leading to a functional form featuring an excessive number of free parameters. Instead, one could consider performing the expansion using a basis of functions that are slowly-varying or steeply-falling in the neighbourhood of $q^2 = m^2$;

for example,

$$R(q^2) = R_{(0)} + R_{(1)} \log\left(\frac{q^2}{m^2}\right) + \mathcal{O}\left(\left(\log\frac{q^2}{m^2}\right)^2\right). \tag{3.53}$$

The alternative, and arguably simpler, method is to immediately perform a truncation of the Taylor expansion at some arbitrary order. This yields a general functional form with a set of parameters fixed by the choice of truncation; its fit to a data distribution holds exactly at $q^2 = m^2$, but its description becomes less accurate as one moves away from the peak. While this means that only a subset of the data collected in an experiment will be relevant to a fit, the extraction of the physical parameters corresponding to a resonance should be unaffected. In addition, maintaining a consistent set of parameters between analyses allows for a simpler interpretation and comparison of results.

We choose to proceed using the second method and a truncation of the Taylor expansion at the first order, being the simplest choice that does not neglect the q-dependence of $R(q^2)$. Substitution into Eq. (3.44) yields:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Big[&f_{\mathrm{BW}}\left(R_{(0)}^2 + 2R_{(0)}R_{(1)}(q^2-m^2)\right)\\ &+2f_{\mathrm{BW}}\left(R_{(0)} + R_{(1)}(q^2-m^2)\right)\Phi(\phi_{\mathcal{SB}})\\ &+1+\mathcal{O}\left((q^2-m^2)^2\right)\Big].\end{aligned} \tag{3.54}$$

Note that as defined currently, $R_{(0)}$ has units of mass-squared, while $R_{(1)}$ is dimensionless. To more easily interpret fit results, we obtain parameters that are of the same canonical dimensions by scaling with factors of the mass,

$$R_{(0)} \to \frac{R_{(0)}}{m^2}, \qquad R_{(1)} \to R_{(1)}, \tag{3.55}$$

to cast both parameters as dimensionless quantities. Also writing the phase-related contributions to the interference explicitly in terms of $c_\phi \equiv \cos\phi_c$ and $s_\phi \equiv \sin\phi_s$, the functional form to first order in $R(q^2)$ becomes:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Bigg[&f_{\mathrm{BW}}\left(m^4R_{(0)}^2 + 2m^2R_{(0)}R_{(1)}(q^2-m^2)\right)\\ &+2f_{\mathrm{BW}}\left(m^2R_{(0)}c_\phi(q^2-m^2) + R_{(0)}m^3\Gamma s_\phi\right)\\ &+2f_{\mathrm{BW}}(q^2-m^2)\left(R_{(1)}c_\phi(q^2-m^2) + R_{(1)}m\Gamma s_\phi\right)\\ &+1+\mathcal{O}\left((q^2-m^2)^2\right)\Bigg].\end{aligned} \tag{3.56}$$

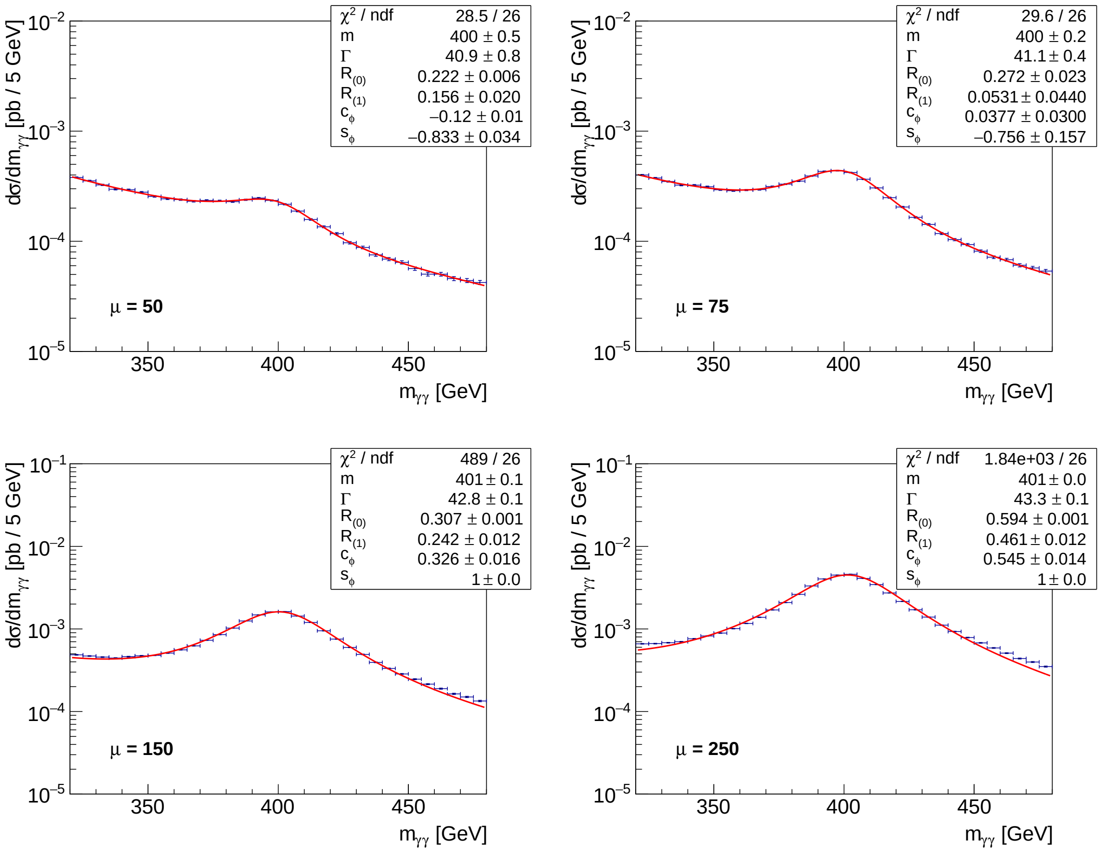

Figure 3.10: Fits of our functional form with $R(q^2)$ expanded to the first order. From top to bottom: Fitted histograms correspond to scaling factors $\mu = 50$, 75, 150, 250.

Note that $\mathcal{O}\left((q^2 - m^2)^2\right)$ refers specifically to terms from the Taylor expansion; the factor of $(q^2 - m^2)^2$ appearing explicitly in Eq. (3.56) arises from separate contributions of $(q^2 - m^2)$ from the squaring of the resonance propagator, and from the first order term of the Taylor expansion of $R(q^2)$.

We now present fits of Eq. (3.56) to BIS histograms, with scaling factors again chosen to be $\mu = 50$, 75, 150, and 250. Discarding the $\mathcal{O}\left((q^2 - m^2)^2\right)$ term, the functional form contains six parameters:

$$\left\{m, \Gamma, R_{(0)}, R_{(1)}, c_\phi, s_\phi\right\}. \tag{3.57}$$

Fit windows were restricted to ± 80 GeV about the peak at $m_{\gamma\gamma} = 400$ GeV, as we know that the first order approximation holds only in the region close to the peak. Fig. 3.10 shows our results. The arbitrarily chosen ± 80 GeV fit window yields a good parametrization of the data distribution for the $\mu = 50$ and $\mu = 75$ fits, returning a $\chi^2/\text{ndf} \approx 1$. However, for the larger μ, we find the fits to deviate slightly from the data near the edges of the window, resulting in a rapid increase of the χ^2 due to the very small uncertainties of the datapoints.

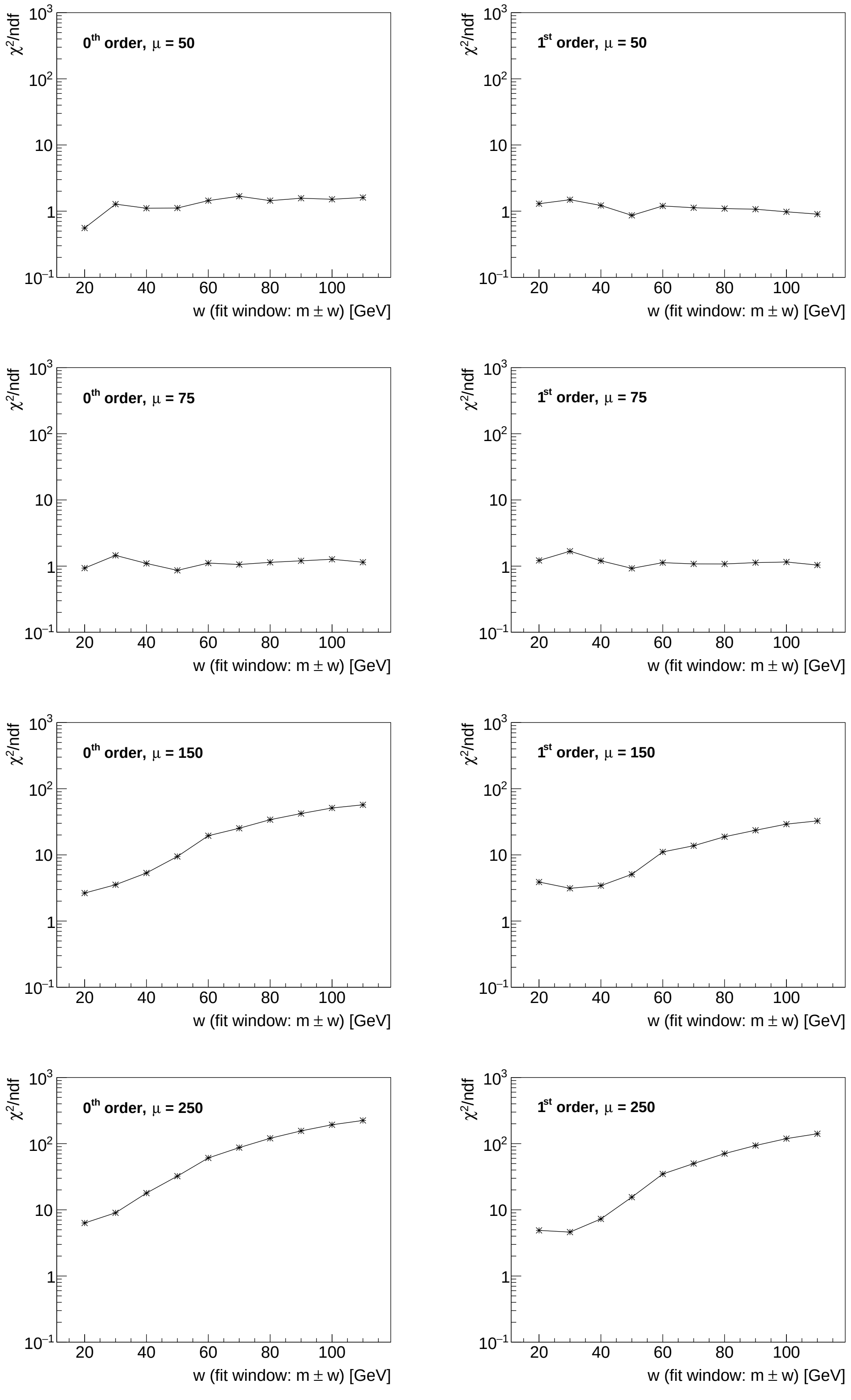

Figure 3.11: Visualization of the χ^2/ndf against fit window size for factors $\mu = 50$, 75, 150, 250. The left-hand column shows results using the zeroth order truncation, while the right-hand column includes the first order parameter.

This trend of decreasing fit quality with increasing μ is similar to our (zeroth order) results from the previous section, though inclusion of the first order parameter, $R_{(1)}$, has improved the fit qualities somewhat; the evolution of best-fit χ^2/ndf, for each μ and both the zeroth and first order truncations, is plotted against the fitted mass window in Fig. 3.11. For $\mu = 50$ and $\mu = 75$, there is little difference between the zeroth and first order results, with a stable $\chi^2/\mathrm{ndf} \approx 1$ found even for the largest window tested (± 110 GeV, centred about the 400 GeV peak). Differences between the two choices of truncation become more apparent in the $\mu = 150$ and $\mu = 250$ results: in the zeroth order case, one finds an approximately exponential increase in the χ^2/ndf up to a threshold window of about ± 60 GeV, past which the χ^2/ndf still increases exponentially, but at a lower rate. The same behaviour past ± 60 GeV is seen in the first order results, but below this threshold, one finds a slower, less-than-exponential deterioration in the fit quality; as expected, these results indicate that the inclusion of higher order Taylor coefficients extends the mass window that can be described by the general functional form.

We note that the $\mu = 50$ and $\mu = 75$ results of Fig. 3.10 show that the first order functional form is able to provide an exceptional description of the data in the 400 ± 80 GeV invariant mass window. Despite this, we are still unable to extract the expected phases, and we also see that their best-fit values differ between the two results. This is surprising, since a change in the scaling of the signal and interference contributions to the data should affect only the $R_{(i)}$ parameters in the functional form, as the shape of the interference has not been altered. We will investigate this problem of unphysical phases in the following section.

On a tangent, let us highlight the importance of keeping $F_{\mathcal{B}}(q^2)$ fixed, not only from a physical standpoint but also from a functional one. Without fixing this background contribution, one introduces additional degrees of freedom to Eq. (3.44) that can greatly increase the variety of lineshapes it can exhibit. Fig. 3.12 shows a result in which we assume the zeroth order truncation of $R(q^2)$ and allow a floating normalization of the background component in the functional form. Fits were repeated on the histogram corresponding to $\mu = 250$. In comparison to the previous zeroth and first order results, we find these fits to yield much lower χ^2/ndf in general, and also deteriorate at a slower rate as the fit window size increases. Of course, despite the apparently improved description of the data, we know that the parameters found in such a result do not reflect the physical truth.

3.3 An alternative parametrization

In the previous section, we presented a general functional form and applied it to selected points in the benchmark signal model. Despite finding accurate parametrizations of the MC histograms in localized regions about the signal peak, we were unable to extract expected values for the interference-related parameters. Let us first note that the general functional form to first order in $R(q^2)$ (Eq. (3.56)) can be

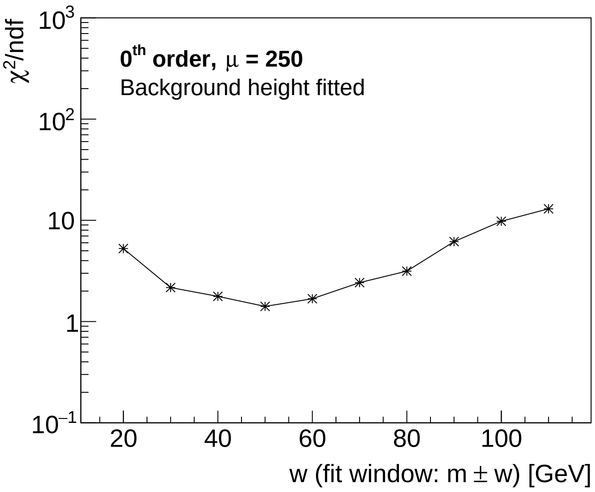

Figure 3.12: Fits to the sample with scaling factor $\mu = 250$ applied, with the height of the background component in the functional form allowed to vary. The ratio parameter is truncated at the zeroth order.

rewritten by grouping terms with like powers of $(q^2 - m^2)$,

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Big[f_{\mathrm{BW}}\Big\{&m^4 R_{(0)}^2 + 2R_{(0)}m^3\Gamma s_\phi \\ &+2(q^2-m^2)\left(m^2 R_{(0)}R_{(1)} + m^2 R_{(0)}c_\phi + R_{(1)}m\Gamma s_\phi\right) \\ &+2(q^2-m^2)^2\left(R_{(1)}c_\phi\right)\Big\} + 1\Big]\,. \end{aligned} \tag{3.58}$$

In this form, an alternative parametrization is obvious:

$$\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\left[f_{\mathrm{BW}}\left\{a_0 m^4 + 2a_2 m^2(q^2-m^2) + 2a_4(q^2-m^2)^2\right\} + 1\right], \tag{3.59}$$

where the a_i are dimensionless quantities, related to the physical parameters by

$$a_0 m^4 = m^4 R_{(0)}^2 + 2R_{(0)}m^3\Gamma s_\phi, \tag{3.60}$$

$$a_2 m^2 = m^2 R_{(0)}R_{(1)} + m^2 R_{(0)}c_\phi + R_{(1)}m\Gamma s_\phi, \tag{3.61}$$

$$a_4 = R_{(1)}c_\phi. \tag{3.62}$$

In this section, we will utilize the a_i description of the general functional form to study the physical parameters[6] by performing closure tests. For this, we make use of Asimov datasets [39], which perfectly represent the model generating the data. We select points in the a_i parametrization as input; the choice of a set of values for the parameters fixes the shape and normalization of the input distribution. The

[6]We henceforth refer to these (the $R_{(i)}$, c_ϕ and s_ϕ) as the “RC” (Ratio-(Co)sine) parameters, of the “RC parametrization” of the general functional form.

Asimov histogram can then be constructed by setting the content of each bin to the value of the input function evaluated at the central invariant mass value of that bin. Uncertainties mimic those for our MC samples,

$$\Delta_{\text{bin}} = \underbrace{\sqrt{N_{\text{tot}} \frac{f(q_{\text{bin}})}{\sum_i f(q_i)}}}_{\text{counting uncertainty}} \times \underbrace{\frac{\sum_i f(q_i)}{N_{\text{tot}}}}_{\text{scaling factor}}, \tag{3.63}$$

where N_{tot} is the total number of events we wish to simulate within the invariant mass window constructed, and $f(q_{\text{bin}})$ is the input function evaluated at the central q value of the corresponding bin. The sum is performed over all bins of the Asimov. The first term is a contribution arising from regular counting statistics,

$$\Delta N = \sqrt{N}, \tag{3.64}$$

for N observed events, while the second scales the uncertainty to account for the fact that the Asimov is normalized to the input distribution, by virtue of its construction.

We begin with the trivial case of using the a_i parameters to generate Asimov histograms, then fitting to these the a_i parametrization. This is a simple closure test with the goal of verifying that we are able to correctly extract the inputs to the Asimovs.

The functional form tested (Eq. (3.59)) contains five parameters: m, Γ, and the three a_i. Ten thousand sets of input values for the a_i were generated en masse using ROOT's `TRandom2` random number generator, with $m = 400$ GeV and $\Gamma = 40$ GeV chosen in all cases. Asimov histograms were created over the 300–500 GeV invariant mass range for each set of inputs. Bin uncertainties were calculated using Eq. (3.63), assuming $N_{\text{tot}} = 1\,000\,000$ events within the invariant mass range constructed. Fits were then performed on the Asimovs; due to the gradient-descent nature of ROOT's fitting algorithm, initial values of $a_i = 0.5$ were used to ensure impartial results across fits.

We present our results in Fig. 3.13, which shows a collation of the ratio of best-fit to input a_i values found by the fits. A pronounced, δ-function-like spike is found at 1, indicating that input a_i values were correctly recovered in every fit, as expected.

This toy closure test was then repeated using the RC parametrization. To ensure that a_i inputs were physical, we generated random values for the RC parameters, then used Eqs. (3.60)–(3.62) for calculation of the inputs; otherwise, the construction of Asimovs follows the procedure described previously. To the Asimovs, the six parameters m, Γ, $R_{(0)}$, $R_{(1)}$, c_ϕ, and s_ϕ of Eq. (3.58) were fitted, with initial values of 0.5 again specified for each parameter (bar m and Γ).

The ratios of best-fit to input values are presented in Fig. 3.14. These results are a stark constrast to those for the a_i parameters: instead of a δ-like peak, we see that many fits have converged to other solutions, although the highest point of these distributions corresponds still to a ratio of 1. These results indicate that the physical parametrization of the general functional form can still yield unexpected

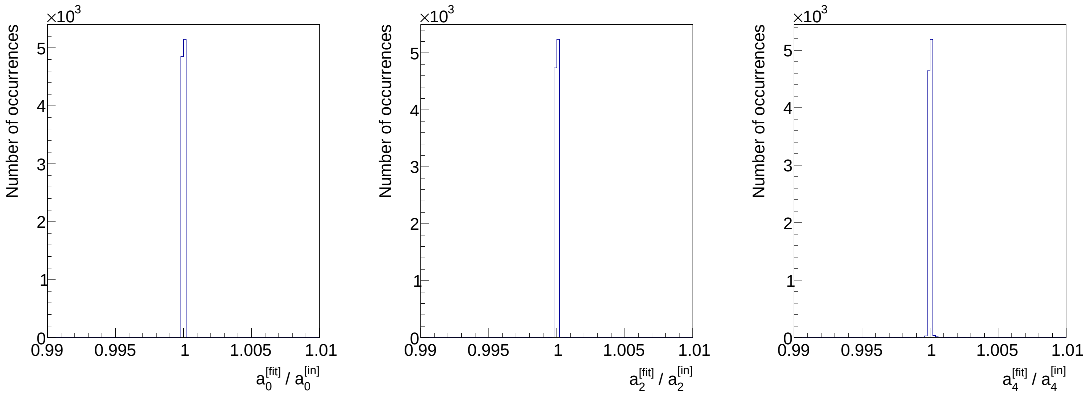

Figure 3.13: A collation of the ratio of extracted best-fit to input parameter values for each of the a_i parameters, from performing 10 000 closure tests using Asimov datasets constructed from randomly generated input points in its space.

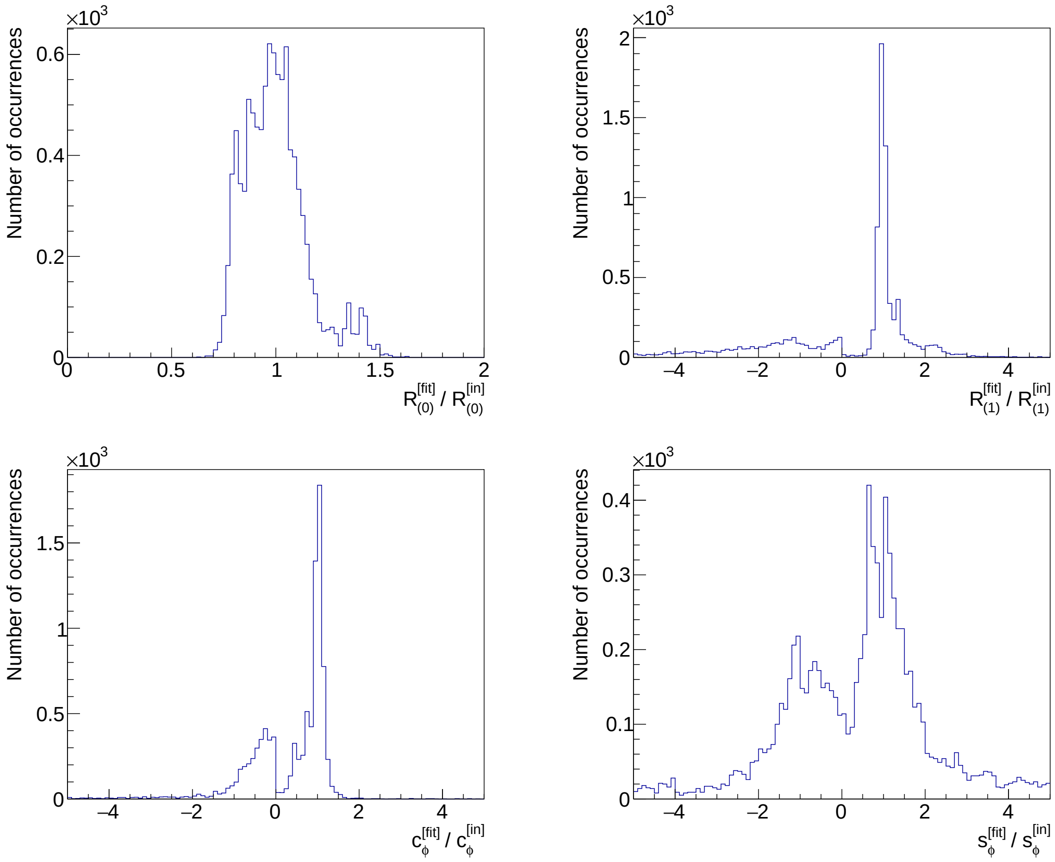

Figure 3.14: A collation of the ratio of extracted best-fit to input parameter values for each of the RC parameters, from performing 10 000 closure tests using Asimov datasets constructed from randomly generated input points in its space.

results, even in the ideal scenario of Asimov datasets. Thus, degeneracies within the space of the RC parameters may be unavoidable. In addition, note that the clustering of solutions about 1 might indicate flat structure in parameter space, which can affect the gradient-descent approach to finding a minimum if convergence is declared prematurely.

3.3.1 Analysis of RC degeneracy

In hindsight, the results of Fig. 3.14 are not surprising—after all, finding solutions for a fit in terms of the RC parameters is essentially equivalent to first finding appropriate values for the a_i, then solving Eqs. (3.60)–(3.62) for the RC. The problem of finding four unknowns with only three equations immediately indicates non-unique solutions.

One might suggest, then, that fitting in terms of the a_i parameters of Eq. (3.59) would lead to simpler and more stable results. However, the a_i parameters can only serve as numerical aids, since they do not have clear physical interpretations; hence, after finding best-fit values for the a_i, one is still faced with the task of finding the corresponding RC parameters.

This is a non-trivial exercise, which we now demonstrate for the special case of a helicity correction factor of 1, such that the c_ϕ and s_ϕ parameters can be expressed using the trigonometric identity:

$$s_\phi = \pm\sqrt{1 - c_\phi^2}\,. \tag{3.65}$$

In this regime, it might be possible to identify unique solutions for the physical parameters, since the problem reduces to solving a system of three equations (given by Eqs. (3.60)–(3.62), assuming known values for m, Γ, and the a_i from a prior fit) for three unknowns (the $R_{(0)}$, $R_{(1)}$, and c_ϕ).

We begin by rearranging Eq. (3.60) to find an equation for c_ϕ in terms of m, Γ, $R_{(0)}$, and a_0,

$$a_0 - R_{(0)}^2 = 2R_{(0)}\frac{\Gamma}{m}\lambda\sqrt{1 - c_\phi^2}\,, \tag{3.66}$$

$$\implies \lambda\sqrt{1 - c_\phi^2} = \frac{m\left(a_0 - R_{(0)}^2\right)}{2R_{(0)}\Gamma}\,, \tag{3.67}$$

$$\implies c_\phi = \sqrt{1 - \frac{m^2\left(a_0 - R_{(0)}^2\right)^2}{4R_{(0)}^2\Gamma^2}}\,, \tag{3.68}$$

where we have denoted the sign ambiguity of s_ϕ as λ. Substituting into Eq. (3.62), we find

$$R_{(1)} = \frac{a_4}{c_\phi} = a_4 \left(\sqrt{1 - \frac{m^2 \left(a_0 - R_{(0)}^2\right)^2}{4R_{(0)}^2 \Gamma^2}} \right)^{-1} . \tag{3.69}$$

Using Eqs. (3.66), (3.68) and (3.69) in Eq. (3.61) then yields an equation for $R_{(1)}$ in terms of m, Γ, $R_{(0)}$, and all three of the a_i parameters:

$$R_{(1)} = \frac{1}{a_2} \left(R_{(0)} a_4 + \frac{a_4^2 \left(a_0 + R_{(0)}^2\right)}{2R_{(0)} \left(1 - \frac{m^2 (a_0 - R_{(0)}^2)^2}{4R_{(0)}^2 \Gamma^2}\right)} \right), \tag{3.70}$$

with a similar expression for c_ϕ easily following from Eq. (3.62):

$$c_\phi = a_2 \left(R_{(0)} + \frac{a_4 \left(a_0 + R_{(0)}^2\right)}{2R_{(0)} \left(1 - \frac{m^2 (a_0 - R_{(0)}^2)^2}{4R_{(0)}^2 \Gamma^2}\right)} \right)^{-1} . \tag{3.71}$$

Thus, given values for $R_{(0)}$, m, Γ, a_0, a_2, and a_4, we can calculate the (unique) values of $R_{(1)}$ and c_ϕ, and also fix λ by noting that the sign of the right hand side of Eq. (3.66) is determined by the sign of λ.

The final exercise is then to find a relationship that yields an $R_{(0)}$ value(s) given only the m, Γ, and a_i. We can immediately see the possibility of multiple $R_{(0)}$ solutions from Eq. (3.60), which can be written as a quadratic equation,

$$b_1 R_{(0)}^2 + b_2 R_{(0)} + b_3 = 0, \tag{3.72}$$

with

$$b_1 = -1 - \frac{a_4}{R_{(1)}^2}, \tag{3.73}$$

$$b_2 = \frac{2a_2}{R_{(1)}}, \tag{3.74}$$

$$b_3 = -a_0. \tag{3.75}$$

However, this does not yet account for the $R_{(0)}$ dependence of $R_{(1)}$, and upon substitution of Eq. (3.70), we instead find the equation

$$F\left(R_{(0)}; m, \Gamma, a_0, a_2, a_4\right) = A_1^2 - A_2 = 0, \tag{3.76}$$

where

$$A_1 = -2a_0a_4 + a_0^2\left(\frac{m}{\Gamma}\right)^2 - 4R_{(0)}^2 - 2a_4R_{(0)}^2 - 2a_0R_{(0)}^2\left(\frac{m}{\Gamma}\right)^2 + R_{(0)}^4\left(\frac{m}{\Gamma}\right)^2, \quad (3.77)$$

and

$$A_2 = -4a_2^2\left(a_0\,\frac{m}{\Gamma} + R_{(0)}\left(2 - R_{(0)}\,\frac{m}{\Gamma}\right)\right)\left(a_0\,\frac{m}{\Gamma} - R_{(0)}\left(2 + R_{(0)}\,\frac{m}{\Gamma}\right)\right). \quad (3.78)$$

$F(R_{(0)}; m, \Gamma, a_0, a_2, a_4)$ is a polynomial of order eight; once the values of m, Γ, and the a_i are specified, its roots yield solutions for $R_{(0)}$. However, note that this is an even function of $R_{(0)}$, and since the physical domain of solutions require $R_{(0)} \geq 0$ by definition, the problem reduces to a quartic one. Thus, each point in a_i parameter space can translate to as many as four distinct, but equivalent, points in RC parameter space.

Let us provide an example using the following set of arbitrarily chosen RC parameter values:

$$\{m,\, \Gamma,\, R_{(0)},\, R_{(1)},\, c_\phi\} = \{400,\, 40,\, 0.2,\, 0.1,\, 0.5\}, \quad (3.79)$$

corresponding to

$$\{m,\, \Gamma,\, a_0,\, a_2,\, a_4\} = \{400,\, 40,\, 0.075,\, 0.13,\, 0.05\}. \quad (3.80)$$

Fig. 3.15 (leftmost) shows the behaviour of Eq. (3.76) for this set of a_i values. In addition to the input $R_{(0)} = 0.2$, we find a second solution of $R_{(0)} \approx 0.19$, and also another one at $R_{(0)} \approx 0.39$, though this final solution is not strictly physical as the function does not meet the axis at this point. Recall that $a_0 - R_{(0)}^2$ yields the correct choice for λ, so in this example, the two physical solutions correspond to $\lambda = 1$, while the third (approximate) solution corresponds to $\lambda = -1$.

The other two plots in Fig. 3.15 depict Eqs. (3.70) and (3.71), showing the solutions for $R_{(1)}$ and c_ϕ once a value for $R_{(0)}$ is fixed. The most striking feature in these plots are the asymptotes near the $R_{(0)}$ solutions, a behaviour that seems to persist also for other sets of input parameters; thus, even small variations in the extraction of $R_{(0)}$ can potentially lead to very different $R_{(1)}$ and c_ϕ results. While we have derived this result assuming the special case of Eq. (3.65), one can imagine a generalization to explain the extraction of unexpected and sporadic relative phases in the MC fits of previous sections.

A numerical verification of these results was performed using an Asimov dataset of 1 000 000 events over the 300–500 GeV invariant mass window. All three of the exact ($\lambda = 1$) and approximate ($\lambda = -1$) solutions predicted analytically were found (Fig. 3.16), through variation of initial parameter values for the $\lambda = 1$ case. The possibility of numerical instabilities, due to divergences in the analytical descriptions of $R_{(1)}$ and c_ϕ, has not affected the convergence to expected solutions in these fits, presumably because of the high-statistic Asimov datasets used. One-dimensional

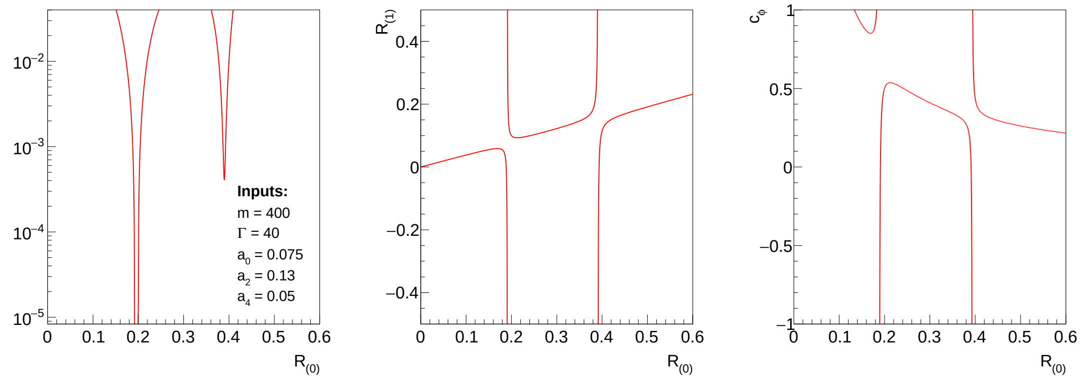

Figure 3.15: Left: a plot that shows two distinct $R_{(0)}$ solutions (zeroes) for one set of a_i parameter values. Centre (right): the solution for $R_{(1)}$ (c_ϕ) once a value for $R_{(0)}$ is fixed.

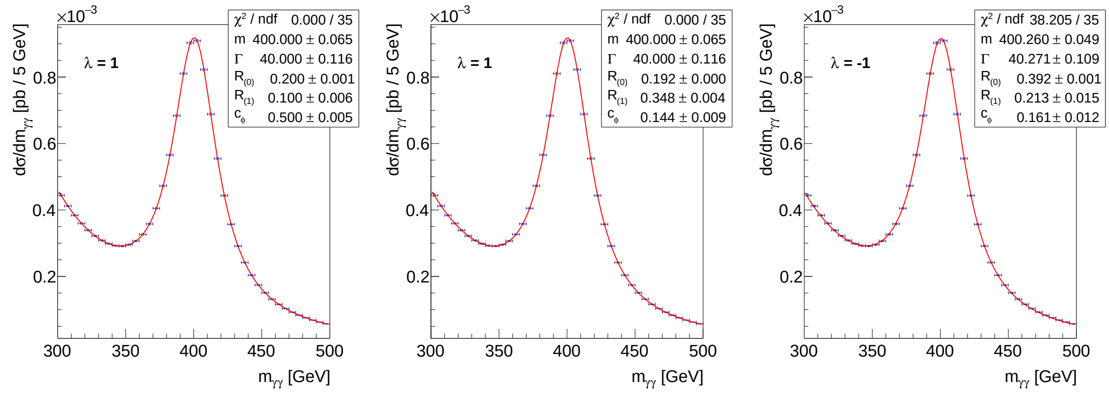

Figure 3.16: Two equivalent and one approximate solution in RC parameter space corresponding to $\{m,\, \Gamma,\, a_0,\, a_2,\, a_4\} = \{400,\, 40,\, 0.075,\, 0.13,\, 0.05\}$ in a_i space.

scans over the parameters, performed for $\lambda = 1$, also distinctly reveal the multiple minima: Fig. 3.17 shows the variation in the χ^2/ndf as $R_{(0)}$, $R_{(1)}$, and c_ϕ are each held fixed over a range of values.

3.4 Summary

We have presented a functional form for the general characterization of resonances and their interference with SM backgrounds, by encapsulating model-dependence within a Taylor expansion of a quantity, $R(q^2)$, defined as the ratio of an assumed model and interfering SM background amplitudes. This was tested for a Higgs-like scalar resonance in the $\gamma\gamma$ final state, using fully simulated MC event samples generated with MadGraph5_aMC@NLO. The general functional form was found to provide a good description of the invariant mass distribution in a localized window

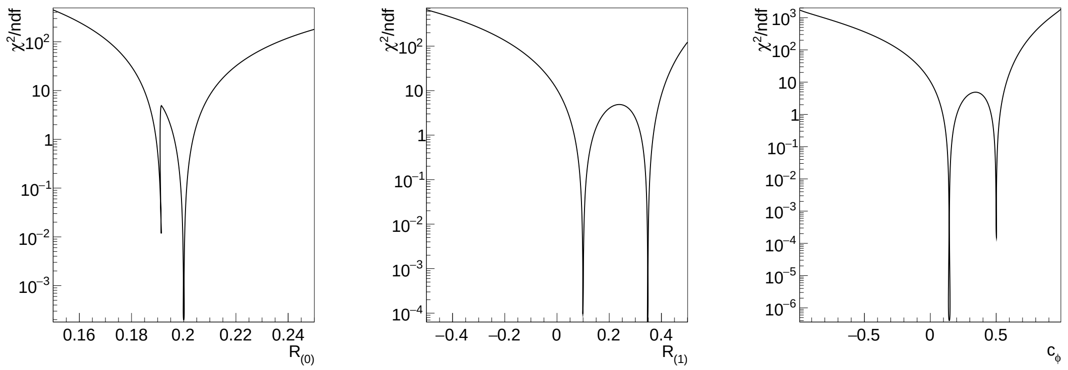

Figure 3.17: Scans of the $R_{(0)}$, $R_{(1)}$, and c_ϕ parameter spaces, for $\lambda = 1$.

about the signal peak, with a stricter restriction placed on the size of this window as the size of the signal increased.

However, the extraction of unexpected relative phases in our fits prompted us to investigate the functional form using an equivalent parametrization in terms of unphysical a_i parameters. Analytical and numerical studies involving this alternative basis of parameters, performed in the specific case that removed the correction factors of helicity selection rules, revealed degeneracies in the space of physical parameters which were largely unavoidable. In addition, the analytical descriptions for $R_{(1)}$ and c_ϕ in this special case indicated a sensitivity to very small changes in the value of $R_{(0)}$. Thus, it might prove difficult to extract the physically correct $R_{(1)}$, c_ϕ, and s_ϕ from a fit, at least with the method involving ROOT's fitting algorithm, which strives to return a single best-fit value, with corresponding uncertainty, for each parameter.

Chapter 4

Global Optimization

In this chapter, we continue our study of the physical parametrization of the general functional form,

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Big[f_{\mathrm{BW}}\Big\{&m^4R_{(0)}^2 + 2R_{(0)}m^3\Gamma s_\phi \\ &+2(q^2-m^2)\left(m^2R_{(0)}R_{(1)} + m^2R_{(0)}c_\phi + R_{(1)}m\Gamma s_\phi\right) \\ &+2(q^2-m^2)^2\left(R_{(1)}c_\phi\right)\Big\} \\ +1\Big]&. \end{aligned} \qquad \text{(3.58 revisited)}$$

However, with a high possibility of multiple solutions in the space of the RC parameters, the gradient-descent fitting approach implemented by ROOT's `TH1::Fit` function is not ideal. We thus turn to the attractive alternative of global optimization techniques, which entail a large-scale sampling of parameter space and reporting of the quality of fit at each point sampled. With this approach, one ideally would not need to worry about finding solutions corresponding to local extrema, and the existence of any degeneracies in parameter space are immediately revealed. The need to specify arbitrary starting points, as gradient-descent methods would require, can also be avoided.

We begin the chapter with a brief overview of some statistical concepts, and describe in more detail the particular sampling algorithm we will use to obtain our results. We will use this to visualize the degeneracies between the RC parameters of the general functional form, especially when one does not assume the special case of $s_\phi = \pm\sqrt{1-c_\phi^2}$ but instead treats the two parameters independently. Nevertheless, a simple final result in the form of parameter values and uncertainties can still be desirable; to this end, we explore possible modifications to the RC parameters in an attempt to lift their degeneracies. A practical method of applying the general functional form to more realistic scenarios will also be investigated; in particular, we propose a procedure with which a suitable fit window can be determined when no prior knowledge of the data is assumed.

4.1 Statistical background

4.1.1 The likelihood function

Consider a sample of events $\boldsymbol{x} = \{x_1, x_2, \cdots, x_N\}$ drawn from a continuous random variable, X, which is distributed according to the probability distribution function (PDF) of some statistical model $M(X; \boldsymbol{\Theta})$. The *likelihood function* for this sample is defined as the PDF of the generating model, but with the model parameters, $\boldsymbol{\Theta}$, acting as the variables:

$$\mathcal{L}(\boldsymbol{\Theta}; X) \equiv M(X;\, \boldsymbol{\Theta}). \tag{4.1}$$

Our studies are performed using histograms of binned events. Bins can thus be considered as independent counting experiments following a Poisson distribution; in the asymptotic limit of a large number of events, this tends towards a Gaussian:

$$G(n; N, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(n-N)^2}{2\sigma^2}\right], \tag{4.2}$$

where n is a random variable representing the number of events, and the parameters N and σ are respectively the mean and standard deviation of the Gaussian corresponding to the particular bin. To define the likelihood function, the interpretation of these quantities are reversed: n becomes the number of observed events in a bin, and σ its uncertainty. The parameter of interest is N, which represents the number of events predicted by some model; in the fitting scenario of a functional form to a binned invariant mass distribution, we have, for the i^{th} bin,

$$N_i \approx F(q_i; \boldsymbol{\theta}), \tag{4.3}$$

where q_i denotes the central invariant mass value of the bin, and $F(q_i; \boldsymbol{\theta})$ the fitted functional form, with parameters $\boldsymbol{\theta}$. Thus,

$$\mathcal{L}_i(\boldsymbol{\theta}; q_i, n_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(n_i - F(q_i; \boldsymbol{\theta}))^2}{2\sigma_i^2}\right]. \tag{4.4}$$

The full likelihood function for a binned fitting procedure is simply the product of the likelihood for each bin:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{\text{all bins}} \mathcal{L}_i(\boldsymbol{\theta}; q_i, n_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\prod_j \sigma_j} \exp\left[-\frac{1}{2}\sum_i \left(\frac{n_i - F(q_i; \boldsymbol{\theta})}{\sigma_i}\right)^2\right]. \tag{4.5}$$

4.1.2 Maximum likelihood estimation

Maximum likelihood estimation (MLE) provides a frequentist approach to the problem of estimating the parameters of a model given some observed dataset. As the name of the method suggests, this is achieved through maximization of a likelihood function, $\mathcal{L}(\boldsymbol{\theta})$. Values for the parameters that maximize the likelihood are known as the maximum likelihood (ML) estimators, $\hat{\boldsymbol{\theta}}$, for which the observed data is most likely, assuming correctness of the model.

Under the MLE method, constant contributions to the likelihood can be discarded, as they do not affect the location of its maximum. It is also equivalent to instead maximize a monotonically increasing function of the likelihood, or introduce negative factors to convert the problem to one of minimization. Particularly for the likelihood function of Eq. (4.5), note that

$$\lambda(\boldsymbol{\theta}) \equiv -2\log\mathcal{L}(\boldsymbol{\theta}) = \sum_i \left(\frac{n_i - F(q_i;\boldsymbol{\theta})}{\sigma_i}\right)^2 , \tag{4.6}$$

where the constant pre-factor of $(\sqrt{2\pi}\prod_j \sigma_j)^{-1}$ has been discarded as it does not affect the optimization. Thus, MLE is essentially equivalent to the exercise of minimizing the χ^2.

However, the problem of parameter estimation involves not only finding the ML estimators, $\hat{\boldsymbol{\theta}}$, but also the uncertainties associated with them. Consider the single-parameter case and assume a Gaussian distribution of the likelihood in parameter space, with mean $\hat{\theta}$ and standard deviation σ; thus,

$$\lambda(\theta) = -2\log\mathcal{L}(\theta) = \frac{\left(\theta - \hat{\theta}\right)^2}{\sigma^2} . \tag{4.7}$$

Confidence intervals for the parameter, θ, can then be constructed using the likelihood ratio test,

$$\Lambda(\theta) \equiv \frac{\mathcal{L}(\theta)}{\mathcal{L}(\hat{\theta})} , \tag{4.8}$$

which assesses the goodness-of-fit of a point in parameter space to that at the ML estimator. For example, the 1σ interval corresponds to $[\hat{\theta}-\sigma,\, \hat{\theta}+\sigma]$, whence

$$-2\log\Lambda(\hat{\theta}\pm\sigma) = \lambda(\hat{\theta}\pm\sigma) - \lambda(\hat{\theta}) = 1. \tag{4.9}$$

Thus, one can identify all points in parameter space that satisfy

$$\lambda(\theta) < \lambda(\hat{\theta}) + 1, \tag{4.10}$$

to lie within the 1σ confidence interval for that parameter; this can be interpreted as the interval that, with an approximately 68% probability (for the 1σ case), contains the true value for θ.[1]

Despite being derived under a Gaussian assumption for θ, the confidence intervals found according to Eq. (4.10) are good approximations even in non-Gaussian cases [40]. A generalization to higher n-dimensional parameter spaces can also be

[1] Note that this is a statement on the confidence interval found, and not the true value of the parameter.

made [41, 39]:

$$\lambda(\theta) < \lambda(\hat{\theta}) + \mathcal{Q}_{\chi^2_n}(p), \tag{4.11}$$

where $\mathcal{Q}_{\chi^2_n}$ is the inverse cumulative distribution function (ICDF) of a χ^2_n distribution with n degrees of freedom, and p the probability of enclosing the true parameter value(s). One can verify that setting $n = 1$ and $p \approx 0.68$ indeed recovers Eq. (4.10).

4.1.3 Bayesian inference

Bayesian inference is a statistical method built upon Bayes' theorem; given data $\boldsymbol{X}$, and a hypothesis model M, with parameters $\boldsymbol{\Theta}$, the theorem states:

$$\mathrm{P}\left(\boldsymbol{\Theta}|\boldsymbol{X}, M\right) = \frac{\mathrm{P}\left(\boldsymbol{X}|\boldsymbol{\Theta}, M\right)\mathrm{P}\left(\boldsymbol{\Theta}|M\right)}{\mathrm{P}\left(\boldsymbol{X}|M\right)}, \tag{4.12}$$

where $\mathrm{P}(\boldsymbol{\Theta}|\boldsymbol{X}, M) \equiv \mathcal{P}(\boldsymbol{\Theta})$ is the *posterior probability distribution* of the parameters; $\mathrm{P}(\boldsymbol{X}|\boldsymbol{\Theta}, M) \equiv \mathcal{L}(\boldsymbol{\Theta})$ is the likelihood function; $\mathrm{P}(\boldsymbol{\Theta}|M) \equiv \pi(\boldsymbol{\Theta})$ are the *prior distributions* of the parameters; and $\mathrm{P}(\boldsymbol{X}|M) \equiv \mathcal{Z}$ is the *Bayesian evidence*, which normalizes the posterior distribution:

$$\mathcal{Z} = \int_{\boldsymbol{\Theta}_{\text{all}}} \mathcal{L}(\boldsymbol{\Theta})\pi(\boldsymbol{\Theta})\, d^D\boldsymbol{\Theta}\,, \tag{4.13}$$

where $\boldsymbol{\Theta}_{\text{all}}$ indicates a definite integral over the entire space of the D parameters.

Bayesian inference can be applied to problems of parameter estimation through maximization of the posterior probability distribution, $\mathcal{P}(\boldsymbol{\Theta})$, in *maximum a posteriori* (MAP) estimation. This is the Bayesian counterpart to the MLE method, and reduces to the MLE in the special case of uniform priors; note that the normalizing $\mathcal{Z}$ factor of Eq. (4.13) has integrated out all dependence on the model parameters $\boldsymbol{\Theta}$, and so can be neglected as it will not affect the maximization of $\mathcal{P}(\boldsymbol{\Theta})$. The evidence plays a more important role when one wishes to compare two competing models, M_0 and M_1, by computing the ratio of their posterior probabilities for a given dataset,

$$\frac{\mathrm{P}(M_1|\boldsymbol{D})}{\mathrm{P}(M_0|\boldsymbol{D})} = \frac{\mathrm{P}(\boldsymbol{D}|M_1)\mathrm{P}(M_1)}{\mathrm{P}(\boldsymbol{D}|M_0)\mathrm{P}(M_0)} = \frac{\mathcal{Z}_1\pi(M_1)}{\mathcal{Z}_0\pi(M_0)}\,, \tag{4.14}$$

also known as the *Bayes factor*. As there would usually be no *a priori* preference for either model, $\pi(M_1)/\pi(M_0) = 1$, so the test reduces to a comparison of the Bayesian evidence for each model. A larger evidence indicates a better agreement of the corresponding model with the observed data.

4.1.4 MultiNest and the nested sampling algorithm

To employ either the MLE or MAP methods for estimating model parameters, one requires a way of sampling parameter space. Many techniques exist for this purpose, such as the Metropolis-Hastings or Hamiltonian sampling algorithms, of the Markov

Chain Monte Carlo (MCMC) class of methods, which have been used extensively in astrophysical and other physics applications [42–44]. The differential evolution algorithm [45–47], which traverses a problem space in a survival-of-the-fittest (set of parameter values) manner, has also found success as a robust and computationally efficient alternative to MCMCs [48–50].

For our work, we choose to use MultiNest [51], which implements a variant of the nested sampling (NS) algorithm specifically for the study of multimodal distributions [52–54]. Thus, we expect it to perform well in identifying the multiple solutions of the general functional form.

At its core, the NS algorithm is a Bayesian inference tool that seeks to calculate the Bayesian evidence (Eq. (4.13)). This is a multidimensional integral over the parameter space of a problem, and can be computationally expensive; the NS algorithm simplifies the calculation by introducing the *prior volume*,

$$X(\ell) = \int_{\mathcal{L}(\boldsymbol{\Theta})>\ell} \pi(\boldsymbol{\Theta})\, d^D\boldsymbol{\Theta}, \tag{4.15}$$

which is an integral of the parameter priors over the region bounded by the $\mathcal{L}(\boldsymbol{\Theta}) = \ell$ iso-likelihood contour. Note that since prior distributions are PDFs (hence normalized to unity), $X(\ell) \in [0, 1]$, with $X(0) = 1$. Using this, one can convert the multidimensional evidence integral to a 1-dimensional one:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)\, dX, \tag{4.16}$$

where $\mathcal{L}(X) = X^{-1}(\ell)$ is a monotonically decreasing function of X, and represents the value of the likelihood at the boundary of a given prior volume. Assuming Riemann integrability of $\mathcal{L}(X)$, Eq. (4.16) can be approximated as a weighted sum,

$$\mathcal{Z} \approx \sum_{i=1}^{M} \mathcal{L}(X_i) w_i, \tag{4.17}$$

with the trapezoidal rule yielding the weights as $w_i = \frac{1}{2}(X_{i-1} - X_{i+1})$. Thus, to calculate the evidence, one requires values for the $\mathcal{L}(X_i)$ and X_i, for $i = 0, 1, \cdots, M$, satisfying $0 < X_M < \cdots < X_i < \cdots < X_0 = 1$.

The sum of Eq. (4.17) is computed in MultiNest as follows: One begins with the full prior volume (corresponding to $X_0 = X(0) = 1$) and from this, draws a sample of N_{live} random ("live") points. The point with the lowest likelihood in this sample, $\mathcal{L}_0$, is then discarded and replaced with a point of likelihood $\mathcal{L}_1 > \mathcal{L}_0$. To improve the efficiency of finding a point that meets this requirement, a variant of the clustered nested sampling [54–56] method is employed, which uses the X-means clustering algorithm [57] to accommodate the possibility of a multimodal distribution. This generates a D-dimensional ellipsoid (or a cluster of ellipsoids) using a covariance matrix calculated from the live points; the acceptance rate of a sampled point increases if it is drawn from this hyper-ellipsoid(s), instead of the full prior volume.

In this way, the prior volume enclosing the N_{live} points is reduced every iteration, $X_i < X_{i-1}$; the algorithm traverses nested shells of the likelihood, hence its nomenclature. For $i > 0$, a probabilistic approximation for X_i is made:

$$X_i \approx t_i X_{i-1}, \tag{4.18}$$

where t_i is a random variable following a PDF describing the largest out of N_{live} samples drawn from a uniform distribution across $[0, 1]$,

$$t_i \sim \mathrm{P}(t) = N_{\text{live}}\, t^{N_{\text{live}}-1}\,. \tag{4.19}$$

However, since the t_i are independent for each iteration, one can write

$$X_i \approx t_i t_{i-1} \cdots t_1 t_0 X_0 = t_i t_{i-1} \cdots t_1 t_0, \tag{4.20}$$

and so one can utilize the expectation and standard deviation of the t_i,

$$E[\log t] = -\frac{1}{N_{\text{live}}}\,, \quad \sigma[\log t] = \frac{1}{N_{\text{live}}}\,, \tag{4.21}$$

to obtain:

$$\log X_i \approx -\frac{(i \pm \sqrt{i})}{N_{\text{live}}} \implies X_i \approx \exp\left(-\frac{(i \pm \sqrt{i})}{N_{\text{live}}}\right). \tag{4.22}$$

Thus, an approximation for the evidence improves with every subsequent iteration; the algorithm terminates once a desired accuracy is achieved.

The NS algorithm thus provides an efficient solution to the problem of sampling parameter space and identifying regions of high likelihood, despite its original goal of calculating the Bayesian evidence. Indeed, MultiNest can be re-purposed for frequentist analyses using the MLE method, provided that one increases the number of live points to the order of 10^3 to 10^4 (in contrast to $N_{\text{live}} \sim 10^2$ required for accurate evidence approximations), and imposes a stricter stopping criterion to allow for a more thorough exploration of high likelihood regions in parameter space [58].

4.2 Fit of RC parameters using MultiNest

Let us return to the toy problem from Sec. 3.3.1, in which we fitted the general functional form (Eq. (3.58)) to an Asimov histogram constructed using these input RC parameter values:

$$\{m,\, \Gamma,\, R_{(0)},\, R_{(1)},\, c_\phi\} = \{400,\, 40,\, 0.2,\, 0.1,\, 0.5\}, \tag{3.79 (revisited)}$$

with input $s_\phi \approx 0.866$ fixed using its relation with c_ϕ in the case of a single helicity configuration.

We repeat this toy study to illustrate the capabilities of MultiNest. The analysis is performed in a frequentist approach by defining uniform priors for each fit parameter. To investigate degeneracies between the RC parameters in the general helicity configuration case, we return to using the six-parameter description of Eq. (3.58) with s_ϕ distinct from c_ϕ. The MultiNest configuration of `nlive` $= 5000$ live points, with tolerance `tol` $= 0.001$ defining the convergence condition, was found to provide a good balance between computational cost and an adequate exploration of parameter space. The MultiNest output is processed and plotted using `pippi` [59] as 2-dimensional *profile likelihood ratios*,

$$\Lambda(\theta_1, \theta_2) = \frac{\mathcal{L}(\theta_1, \theta_2; \hat{\boldsymbol{\Theta}}|_{\theta_1,\theta_2})}{\mathcal{L}(\hat{\boldsymbol{\Theta}})}\,, \tag{4.23}$$

where θ_1 and θ_2 are the two parameters relevant to a given plot, with $\hat{\boldsymbol{\Theta}}|_{\theta_1,\theta_2}$ denoting values for the remaining parameters such that the likelihood is maximized at each point in the θ_1-θ_2 plane. $\mathcal{L}(\hat{\boldsymbol{\Theta}})$ denotes the maximum likelihood in the full parameter space.

The result of the fit is presented in Fig. 4.1 as three separate contour plots over the m-Γ, $R_{(0)}$-$R_{(1)}$, and c_ϕ-s_ϕ planes in parameter space. Confidence regions corresponding to $p = 0.683$ (1σ) and $p = 0.954$ (2σ) have been outlined with white, defined according to Eq. (4.11) with $n = 2$. The Gaussian-like peak in the m-Γ plane is expected; of more interest are the large, curving degeneracies found in the $R_{(0)}$-$R_{(1)}$ and c_ϕ-s_ϕ plots. Removing the restriction on s_ϕ has extended the multiple but unique solutions previously predicted, to a single connected region in parameter space. These results also verify the robustness of MultiNest as a sampling algorithm for MLE analyses, even in the case of highly non-trivial distributions.

4.3 The $a_{i,j}$ parametrization

The RC parameter space of the general functional form has been found to exhibit complex degenerate behaviour, largely owing to the fact that its four parameters can be transformed into an alternative basis of only three a_i parameters. This is not a problem in particular, as long as one understands that these degeneracies exist and hence use appropriate tools to perform an analysis. However, the ideal scenario in which one reports a set of best-fit parameter values and uncertainties is still desirable for its much simpler communication of results.

Motivated by this, let us explore possible means of lifting the RC degeneracies. Recall the definition of $R(q^2)$ as the ratio of signal (excluding a factor of the propagator) to background amplitudes (Eq. (3.45)). By doing this, a general functional form was derived, in which a description of the background was assumed to be obtainable from a fit to data or otherwise. While this remains true, any physical background will generally be a steeply decreasing function of q. We exploit this fact

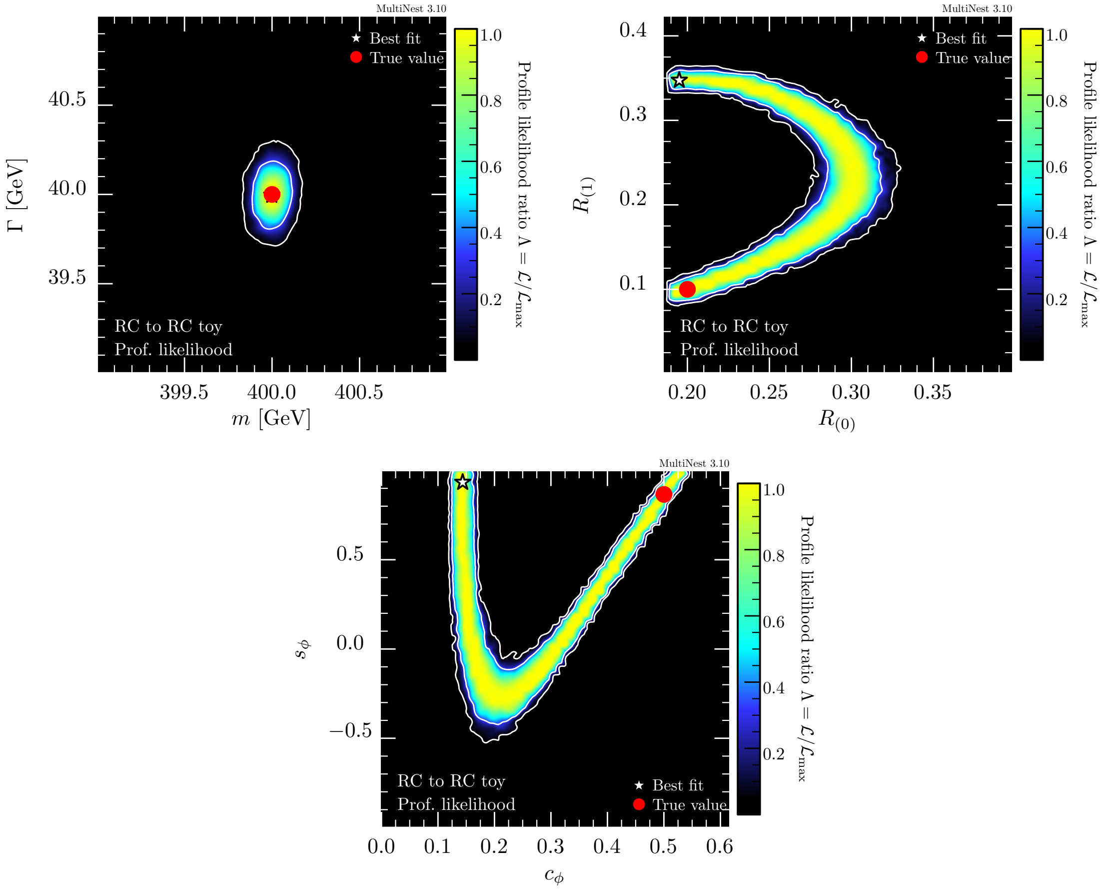

Figure 4.1: A MULTINEST fit of the six RC parameters to a toy Asimov constructed from the same general parametrization. The white star and red circle respectively mark the best-fit and input points. The contours correspond to 68% (1σ) and 95% (2σ) confidence regions.

analytically to modify the parton-level amplitude:

$$\mathcal{A}(q^2) = \frac{\mathcal{S}(q^2)}{q^2 - m^2 + im\Gamma} + \left(\frac{m^2}{q^2}\right)^a \frac{\mathcal{B}(q^2)}{m^2}\,, \tag{4.24}$$

where the q^{-2a} function, for $a \in \mathbb{R}^+$, serves to parametrize the explicit q-dependence of the background amplitude. Indeed, we find q^{-5} to provide a good approximation of the MADGRAPH5_AMC@NLO $gg \to \gamma\gamma$ background invariant mass distribution, corresponding to $a \approx 1.3$ in the amplitude. If one can approximate the dominant kinematic behaviour of the signal and background respectively as a Breit-Wigner and the q^{-2a} function, then the ratio of amplitudes, $R(q^2)$, will be a relatively flat function of q^2, with a quickly converging Taylor expansion. Of course, this will only be true in the neighbourhood of q^2 about m^2, as the signal also receives the same PDF and flux contributions whence the fact of steeply falling backgrounds arise.

The culmination of re-defining the background amplitude according to Eq. (4.24) is a corresponding redefinition of $R(q^2)$, which amounts to a change of its expansion coefficients in the functional form,

$$R_{(0)} \to \left(\frac{q^2}{m^2}\right)^a R_{(0)}, \quad R_{(1)} \to \left(\frac{q^2}{m^2}\right)^a R_{(1)}. \tag{4.25}$$

Hence, the q^{-2a} function can also be interpreted as one that parametrizes the q-dependence of $R(q^2)$, even though it was motivated by the fact of steeply falling backgrounds. It is thus acceptable to consider values for the parameter a not necessarily derived from background-only considerations.

Explicitly, the general functional form under this regime is

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2) \Bigg[f_{\mathrm{BW}} \Bigg\{ & m^4 \left(\frac{q^2}{m^2}\right)^{2a} a_{0,4} + m^4 \left(\frac{q^2}{m^2}\right)^{a} a_{0,2} \\ & +2(q^2-m^2)\left(m^2 \left(\frac{q^2}{m^2}\right)^{2a} a_{2,4} + m^2 \left(\frac{q^2}{m^2}\right)^{a} a_{2,2} \right) \\ & +2(q^2-m^2)^2 \left(\left(\frac{q^2}{m^2}\right)^{a} a_{4,2} \right) \Bigg\} \\ +1 \Bigg], \end{aligned} \tag{4.26}$$

where the $a_{i,j}$ parameters are:

$$a_{0,4} = R_{(0)}^2, \tag{4.27}$$

$$a_{0,2} = 2R_{(0)}\frac{\Gamma}{m}s_\phi, \tag{4.28}$$

$$a_{2,4} = R_{(0)}R_{(1)}, \tag{4.29}$$

$$a_{2,2} = R_{(0)}c_\phi + R_{(1)}\frac{\Gamma}{m}s_\phi, \tag{4.30}$$

$$a_{4,2} = R_{(1)}c_\phi. \tag{4.31}$$

Using Eqs. (4.27)–(4.29) and (4.31), expressions for the RC parameters can be found:

$$R_{(0)} = \sqrt{a_{0,4}}\,, \tag{4.32}$$

$$R_{(1)} = \frac{a_{2,4}}{\sqrt{a_{0,4}}}\,, \tag{4.33}$$

$$c_\phi = \frac{a_{4,2}\sqrt{a_{0,4}}}{a_{2,4}}\,, \tag{4.34}$$

$$s_\phi = \frac{m}{2\Gamma}\frac{a_{0,2}}{\sqrt{a_{0,4}}}\,, \tag{4.35}$$

where the positive solution for $R_{(0)}$ is the correct one. Eq. (4.30) yields an identity,

$$\Delta \equiv a_{2,2} - \frac{a_{0,4}a_{4,2}}{a_{2,4}} - \frac{a_{2,4}a_{0,2}}{2a_{0,4}} = 0. \tag{4.36}$$

Of course, this can never be exactly satisfied in a numerical procedure, but it is sufficient for Δ to be statistically compatible with zero; for example, an $X\sigma$ compatibility, assuming no correlations between the $a_{i,j}$, would require

$$|\Delta| \leq \sqrt{\sum_{a_{i,j}} \left(\left| \frac{\partial \Delta}{\partial a_{i,j}} \right|^2 \delta a_{i,j}^2 \right)}, \tag{4.37}$$

where $\delta a_{i,j}$ denotes the $X\sigma$ level uncertainty for the corresponding $a_{i,j}$.

In this way, we have obtained a 5-parameter description of the general functional form (the $a_{i,j}$) that can be transformed into a basis of only four parameters (the RC). Note that the parameter a is not to be considered on the same level as the other parameters, as its value is not to be determined from a fit. Instead, one assumes a (positive real) value for it prior to the fit. In the limiting case of $a = 0$, one recovers the a_i parameters previously studied,

$$a_{0,4} + a_{0,2} \stackrel{a=0}{=} a_0, \tag{4.38}$$

$$a_{2,4} + a_{2,2} \stackrel{a=0}{=} a_2, \tag{4.39}$$

$$a_{4,2} \stackrel{a=0}{=} a_4, \tag{4.40}$$

but for any other choice of a, the $a_{i,j}$ can be distinguished; thus, much in the same way that multiple RC solutions can correspond to a single point in a_i space, it might be possible to find (degenerate) solutions for the $a_{i,j}$, but which all map to a unique solution in RC space.

4.3.1 MultiNest $a_{i,j}$ toy closure tests

Let us perform the most trivial test for the proposed $a_{i,j}$ parametrization, as we did for the RC in Sec. 4.2. We construct a toy Asimov histogram using a chosen input point in $a_{i,j}$ (or equivalently, RC) parameter space, then perform a fit of the $a_{i,j}$ using MultiNest, with the goal of visualizing its parameter space in the ideal scenario that at least one exact solution (the input point) is known to exist. This is done for six different a values: 1, 1.5, 2, 2.5, 3, and 3.5, to test whether its choice affects a set of results.

We choose the particular set of input parameters

$$\{m,\, \Gamma,\, R_{(0)},\, R_{(1)},\, c_\phi,\, s_\phi\} = \{400,\, 40,\, 0.2,\, 0.1,\, 0.5,\, 0.866\}\,, \tag{4.41}$$

with corresponding input $a_{i,j}$ obtainable from Eqs. (4.27)–(4.31),

$$\{a_{0,4},\, a_{0,2},\, a_{2,4},\, a_{2,2},\, a_{4,2}\} = \{0.04,\, 0.03464,\, 0.02,\, 0.10866,\, 0.05\}\,. \tag{4.42}$$

Asimov histograms are constructed over the 300–500 GeV invariant mass range, with Eq. (3.63) providing uncertainty estimates assuming 1 000 000 events in this region. We perform the fits in terms of m, Γ, and the five $a_{i,j}$ parameters; Eqs. (4.32)–(4.35) then provide the transformations back to RC space.

For brevity, we show only the $a = 1.5$ result here, with results for the other a values delegated to App. B. Profile likelihood plots are presented in Fig 4.2, with the 1σ and 2σ contours shown, and including the transformed result of RC parameters. We again find a well-defined high likelihood peak in the m-Γ plane, agreeing well with the inputs chosen. For the $a_{i,j}$ parameters, linear correlations between the parameters are seen. The 1σ region meets or crosses the zero value in all cases; in particular for $a_{0,4}$ and $a_{2,4}$, whose reciprocal enters the expressions for $R_{(1)}$, c_ϕ, and s_ϕ (Eqs. (4.33)–(4.35)), this leads to a very large variation in the RC parameters. A larger choice for a yields more constrained regions in $a_{i,j}$ space, but does not improve the complicated result in RC space.

However, there are yet additional constraints that can be imposed on these results. We note that many of the high likelihood $a_{i,j}$ points are unphysical, corresponding to c_ϕ or s_ϕ values outside the range $[-1, 1]$. We can hope to better constrain the $a_{i,j}$ result by rejecting all unphysical solutions, but this will obviously not improve the situation for RC parameters, as the high likelihood region found for both c_ϕ and s_ϕ currently enclose the entire range of physical values.

We can additionally take into consideration the identity of Eq. (4.36). Fig. 4.3 shows a 1-dimensional profile likelihood plot for the quantity, Δ, which we expect to be approximately zero for a consistent transformation between the $a_{i,j}$ and RC parameters. This is clearly not the case, as we see a large variation in its value. The compatibility condition of Eq. (4.37) does not hold here, since the $a_{i,j}$ are clearly correlated; we instead make the arbitrary choice of restricting Δ to be in $[-1, 1]$. Fig. 4.4 shows the result after imposing these restrictions on Δ, c_ϕ, and s_ϕ. Indeed, we find that the size of the contours between the $a_{i,j}$ parameters have been greatly reduced. However, their degeneracies remain, and importantly, we still find a poorly constrained solution for c_ϕ and s_ϕ.

Recall that the introduction of a q^{-2a} function was supposed to reduce the q-dependence of $R(q^2)$, leading to a more quickly-convergent Taylor expansion in q^2. Currently, we assume a truncation at the first order; however, the zeroth order might prove sufficient for a given problem. Setting $R_{(1)} = 0$ in Eqs. (4.27)–(4.31) yields $a_{2,4} = a_{4,2} = 0$, and

$$a_{2,2} = R_{(0)} c_\phi, \tag{4.43}$$

with no change to $a_{0,4}$ or $a_{0,2}$. The equation for c_ϕ is then

$$c_\phi = \frac{a_{2,2}}{\sqrt{a_{0,4}}}\,, \tag{4.44}$$

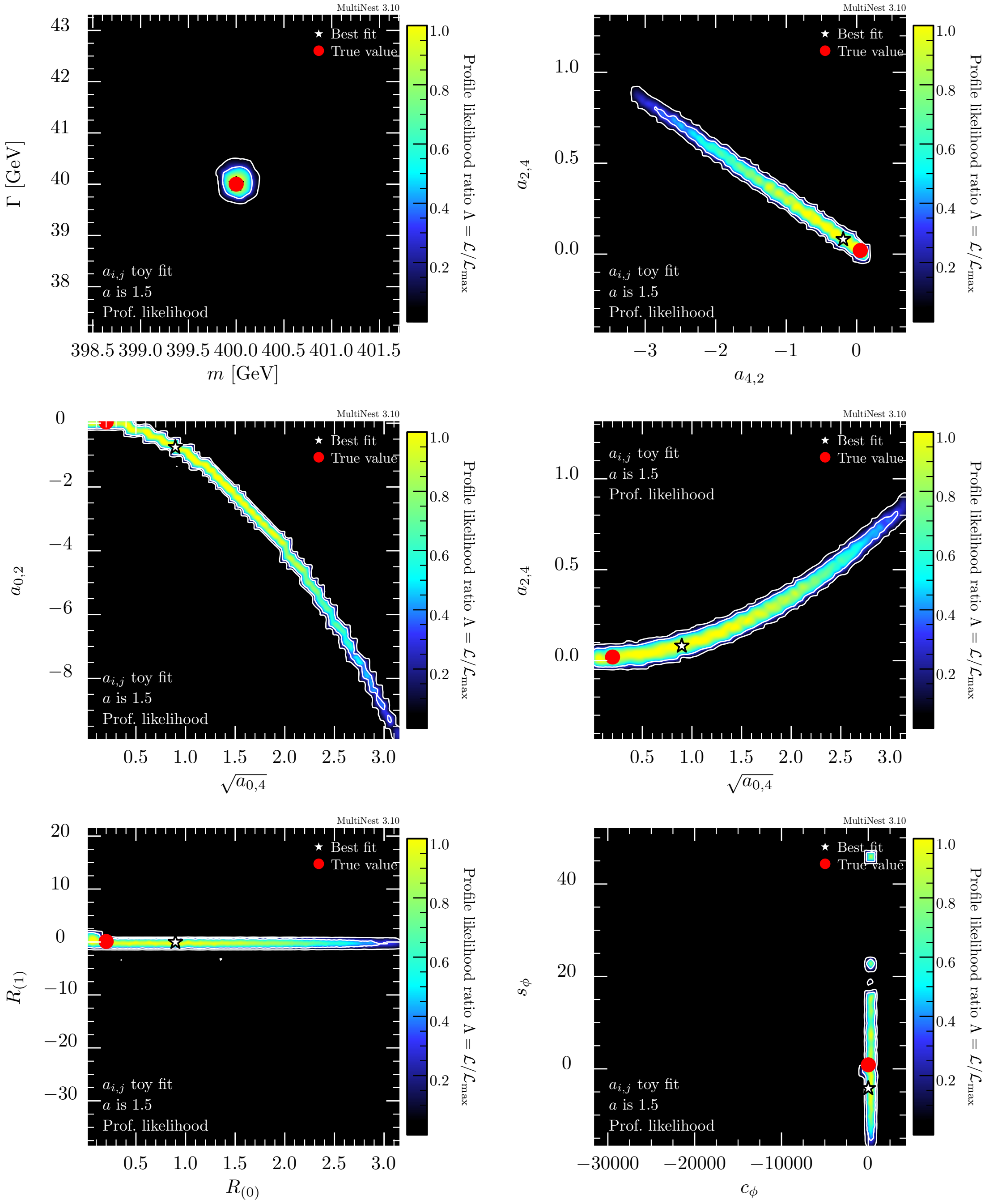

Figure 4.2: Profile likelihood contours of the $a_{i,j}$ parameters fitted to an Asimov toy constructed from an input point in its space, with $a = 1.5$. The bottom two plots show the result in terms of the RC parameters.

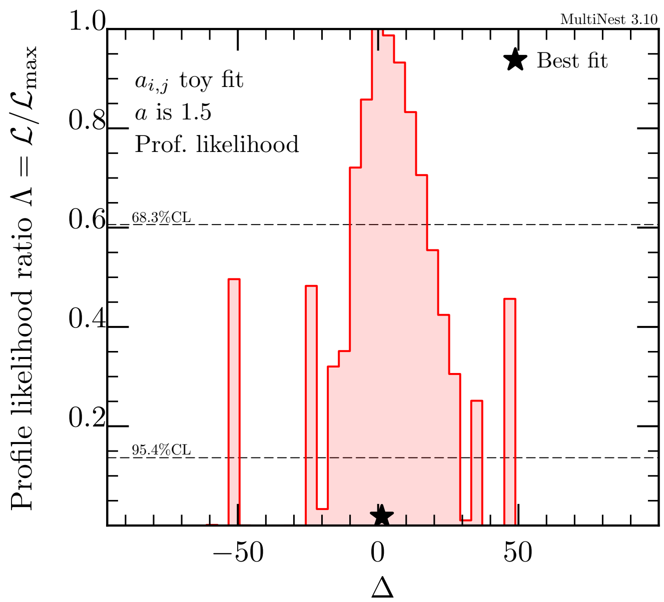

Figure 4.3: One-dimensional profile likelihood ratio plot for the quantity, Δ, when $a = 1.5$. For the correct relationship between the $a_{i,j}$ and RC parameters, one requires this to adopt a value of approximately zero.

with $R_{(0)}$ and s_ϕ as before.

One can check whether a problem can be reduced to the zeroth order by looking at the $a_{4,2}$-$a_{2,4}$ first order result: a high likelihood at the point $(0, 0)$ indicates the existence of a good solution at the zeroth order. Note that it is not sufficient to only require $R_{(1)} = 0$, as this can be achieved when $a_{2,4} = 0$ but $a_{4,2} \neq 0$. For our $a = 1.5$ result, we find the zero to lie very close to the boundary of, but still within, the 1σ region in $a_{4,2}$-$a_{2,4}$ space. Thus, while including the first order parameters will return a better fit, we still anticipate an acceptable fit under the zeroth order assumption, with the benefit of possibly finding a cleaner result between its parameters.

The result of the zeroth order fit, with $a = 1.5$, to the (first order) Asimov histogram is shown in Fig. 4.5. The best fit corresponds to $\chi^2/\text{ndf} \approx 0.1$, indicating that a good agreement with the toy can indeed be found using only the zeroth order parameters. The solutions at this order are highly constrained to small elliptical regions, although some linear correlation between the parameters can still be seen.

However, a problem with fitting in terms of the $a_{i,j}$ is highlighted by these results: it is difficult to ensure that, after transforming to the RC parameters, physical values for c_ϕ and s_ϕ are found. In particular, the current zeroth order result finds s_ϕ values that are unphysical. Note that this does not indicate the non-existence of physical solutions that could provide a good fit of the data, only that any such solution is less preferred than the one found.

Combined with the fact that correlations between the $a_{i,j}$ seem unavoidable, our initial motivation for its use—to find "nice" solutions representable as a value with some uncertainty—has not been met. Thus, if one were to utilize Eq. (4.26) in a

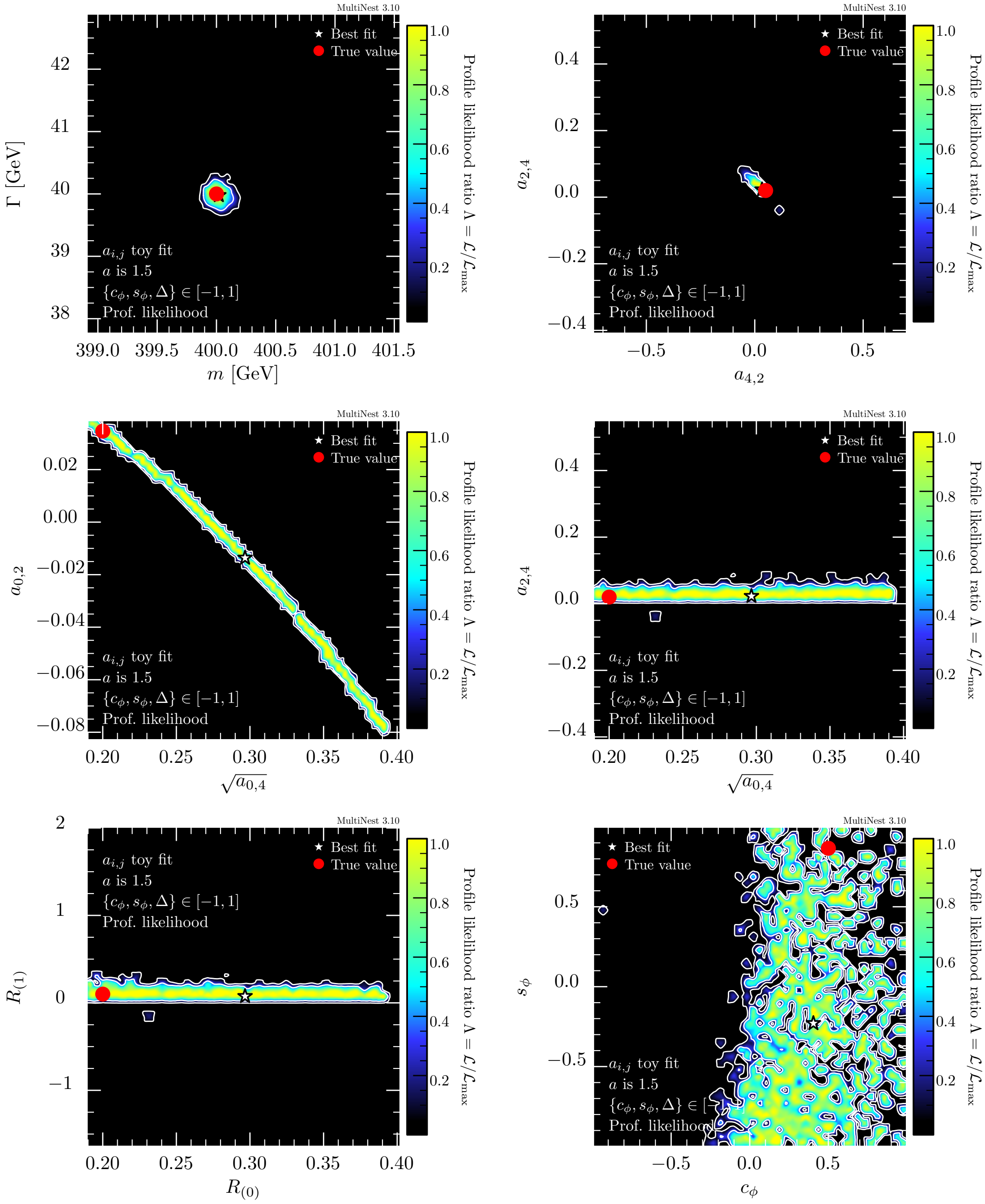

Figure 4.4: The $a_{i,j}$ fit result for $a = 1.5$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded. The fit was performed in terms of the $a_{i,j}$ parameters, leading to the somewhat poor visualization of RC parameter space.

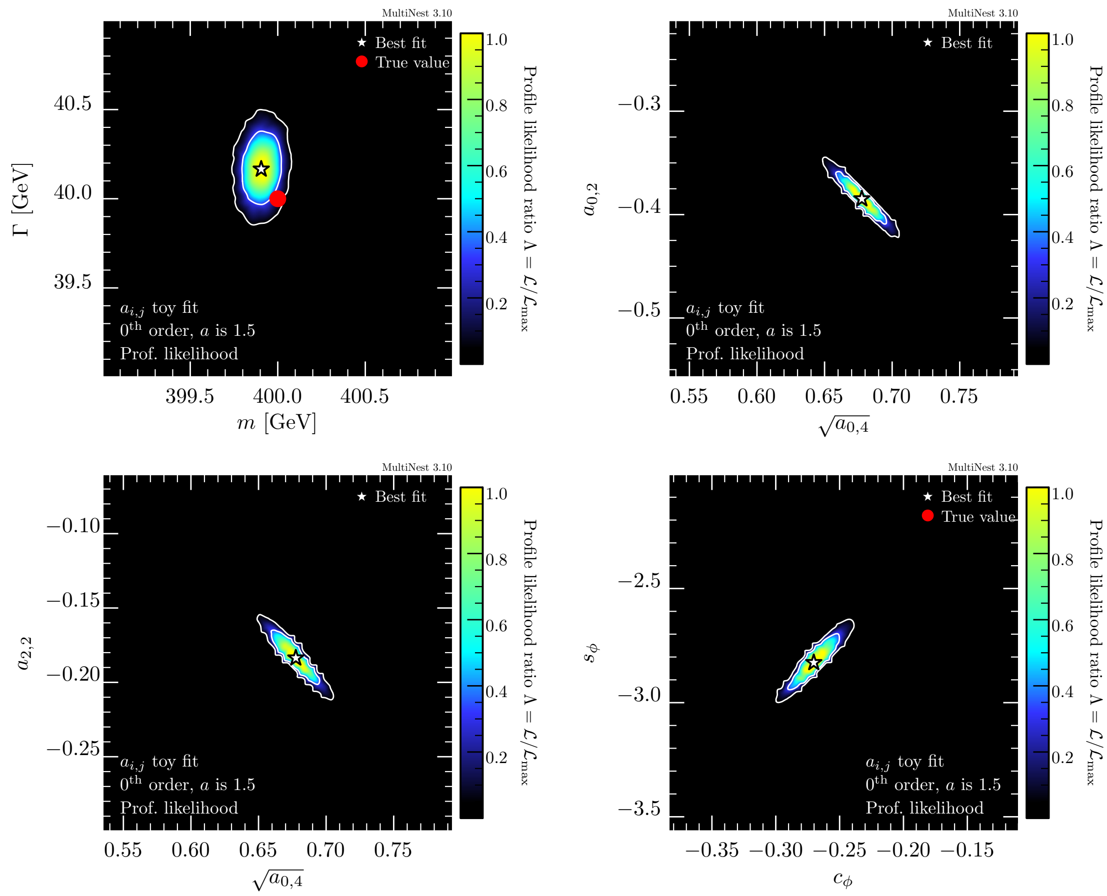

Figure 4.5: The result of a zeroth order $a_{i,j}$ fit to the Asimov constructed from first order inputs, with $a = 1.5$.

fit, it would be simpler to directly fit the RC parameters (Eqs. (4.32)–(4.35)), which would also ensure adequate sampling of its parameter space.

In proceeding, we will return to the general functional form and set of RC parameters as defined prior to this section, to avoid the complication of needing to choose a value for the a parameter prior to a fit.

4.4 Fit window tests

Let us return to the original problem posed in Ch. 3, which sought to test the viability of the general functional form in describing the physics of a benchmark model. We have now verified MultiNest to be a suitable tool for obtaining results capable of visualizing the complicated correlations between the parameters of the general functional form; this has been shown for the ideal scenario in which the fitted data corresponded exactly to at least one set of values for the RC parameters. However, for a distribution derived from a general physics model, the RC parameters provide only an approximation: a good description can indeed be found in an invariant mass window restricted about the signal peak, but we know that this deteriorates with the size of the window.

Our goal in this section is to determine, in a more statistically rigorous manner, the threshold invariant mass window for which an accurate description of the data still holds. Beyond showing that a good fit exists, we also wish to verify that one can correctly extract the parameters of the physics model, given a fit result in terms of the RC parameters. Although our results will be obtained using only the benchmark model presented in Sec. 3.1, we will again make use of Asimov toys to generate a variety of physical distributions to test. Due to the wide range of behaviour these Asimovs will exhibit, we assume our results to generalize well to other physics models.

4.4.1 Methodology and results

Our testing procedure is as follows. We first obtained Asimov toys using the analytical description of the benchmark physics model (PM), which was derived in Sec. 3.1. This description contains five parameters:

$$\{m,\, \Gamma,\, c_\phi,\, s_\phi,\, f_s\}. \tag{4.45}$$

By choosing different input values for these parameters when constructing the Asimovs, a large variety of lineshapes could be tested.

Ten input mass points were chosen, with values of $m = 400$ GeV to $m = 1300$ GeV in increments of 100 GeV. Three choices were made for the input width: $\Gamma/m = 3\%$, 5%, and 10%. For the three remaining parameters of the PM functional form, two different sets of inputs will be used:

$$\{c_\phi, s_\phi, f_s\} = \begin{cases} \{0.7,\, 0.3,\, 4{\times}10^{-16}\} & \text{“set 1”}, \\ \{-0.8,\, 0.1,\, 4{\times}10^{-18}\} & \text{“set 2”}. \end{cases} \tag{4.46}$$

Notably, the height of the signal, f_s, is two orders of magnitude smaller in the set 2 inputs. The Asimovs created using the distributions defined by these inputs[2] are shown in Fig. 4.6, for the choice of 2 GeV bin widths and 10 000 000 events across the 100–1600 GeV invariant mass range visualized.

To each Asimov constructed, multiple fits of the general functional form were performed, over invariant mass windows of different sizes centred about the corresponding mass input to that Asimov: window widths were chosen to be $w = 20$ GeV to $w = 200$ GeV, in increments of 20 GeV, such that fit windows were given by $m \pm w$. However, in a blind search, the true mass of the resonance is of course unknown a priori; while the general parameters were obtained by means of a Taylor expansion about the (square of the) resonance mass, it is unrealistic to expect that one can always choose a fit window exactly centred about this point. Thus, we also performed additional fits in which the central invariant mass value of the fit window

[2]For $\Gamma/m = 5\%$ only; see App. C.1 for the 3% and 10% Asimovs.

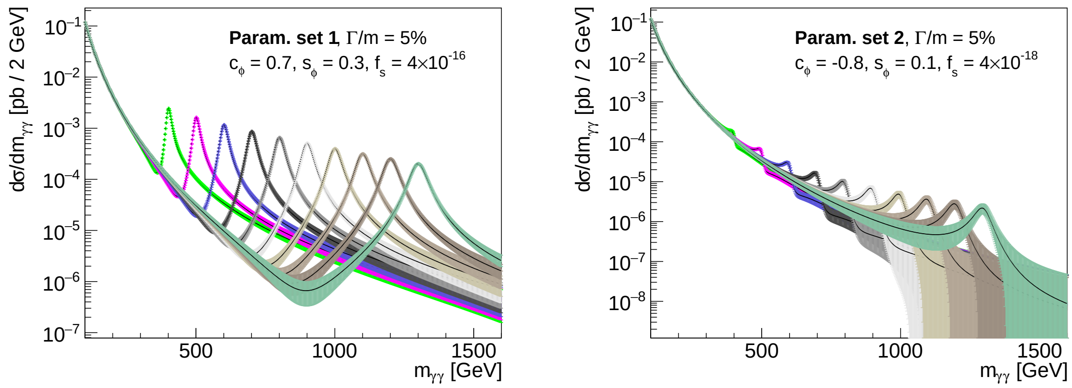

Figure 4.6: Visualization of Asimov datasets, generated from the PM functional form at various mass points with $\Gamma/m = 5\%$, for the two sets of input parameters.

was shifted to the left or right of its true value, to test whether a good fit could still be obtained. These were chosen to have a central invariant mass value of $m \pm w/2$, for fit windows $(m \pm w/2) \pm w$.

To identify an approximate threshold where the general functional form still fits the input toy histograms well, we note that the $-2 \log \mathcal{L}$ quantity of our fits follows a χ^2_ν distribution (Eq. (4.6)), with the number of degrees of freedom, ν, equal to the difference between the number of datapoints (histogram bins) and the number of fitted parameters. We can thus calculate, for example, the 1σ cut-off by evaluating the ICDF of a χ^2_ν distribution, $\mathcal{Q}_{\chi^2_n}(p)$, at $p \approx 0.68$. If the best-fit χ^2 of a given result falls below this value, then the general functional form is able to provide a good description (within 1σ) of the toy, for that choice of fit window. We present our results in terms of the ratio of best-fit to cut-off χ^2,

$$\frac{\chi^2_{\rm bf}}{\chi^2_{1\sigma}} \,. \tag{4.47}$$

Thus, an exact fit will find a result of zero, with the 1σ threshold at one; values larger than this indicates increasing deterioration of the fit quality.

The results corresponding to input $\Gamma/m = 5\%$ for both sets of input parameters are presented in Fig. 4.7 as colour plots, with the ratio of chi-squares on the colour axis (the 3% and 10% results are left to App. C.2). We find a trend of decreasing fit quality as the window size increases, and also as the input mass decreases. This effect is more pronounced in the set 1 results, with a fit window of $m \pm 40$ GeV being the largest choice suitable for fitting a 400 GeV mass. In contrast, the ratio of chi-squares found in the set 2 results remain small across the entire range of windows tested. In general, this shows that a larger signal will restrict the fit to smaller windows.

Comparing with the results obtained when the fit window is shifted, we find that while there are small distortions to the exact value of the χ^2 ratios, the general

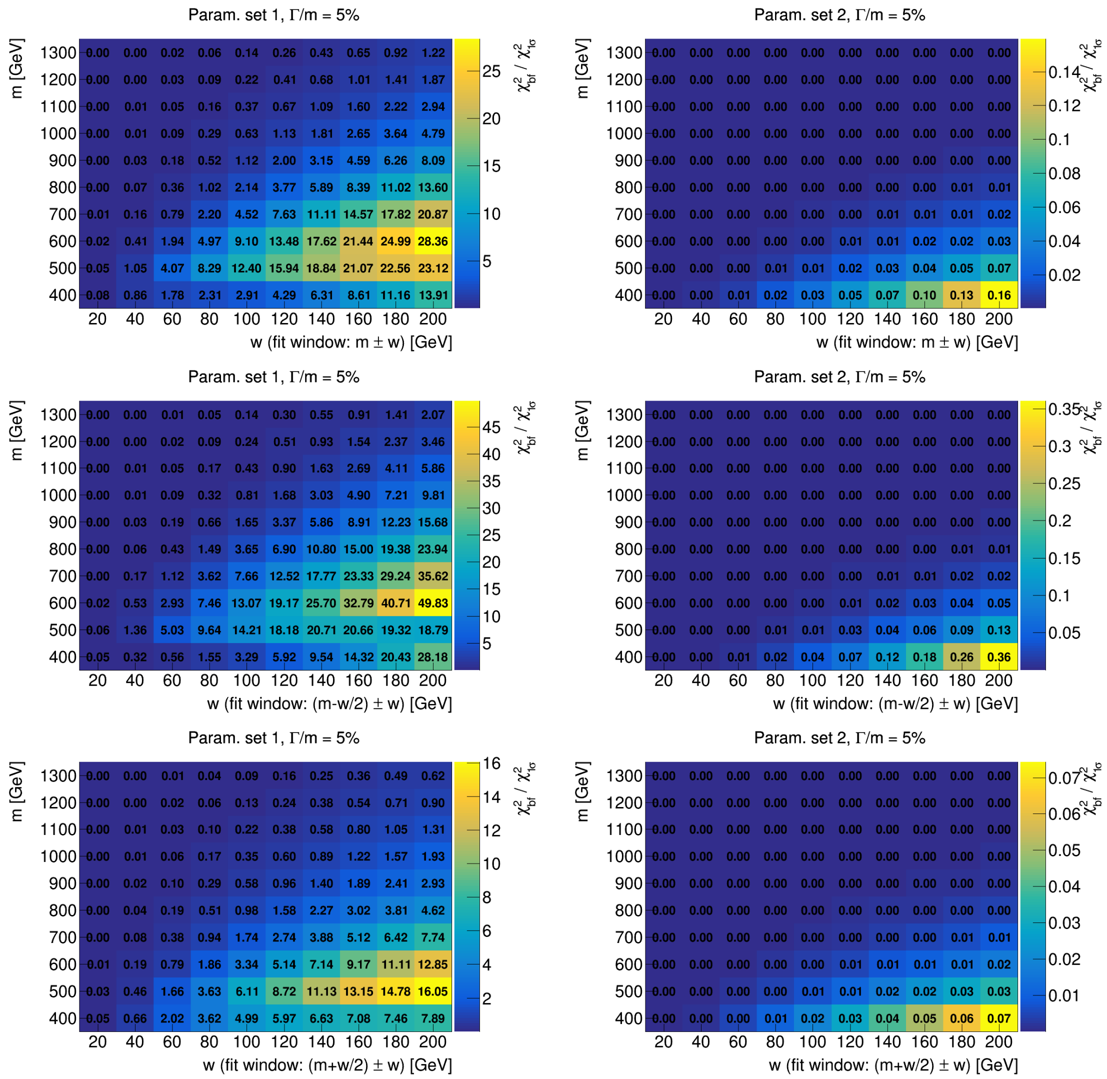

Figure 4.7: Colour plots showing the quality of fits relative to the 1σ threshold of a χ^2_ν distribution. Results correspond to input widths of $\Gamma/m = 5\%$. Left (right): results using set 1 (2) of the input PM parameters. From top to bottom, results correspond to a fit window: centred about the input mass; shifted to the left by $w/2$; shifted to the right by $w/2$.

conclusions remain unchanged. Thus, only the size of a fit window is important to a fit of the general functional form, with results being mostly insensitive to the exact mass range that this window covers.

These results have been presented in an easily interpretable manner for many different input points in the PM parameter space, and are generally consistent with the conclusions previously reached in Sec. 3.2.2, where results were obtained using MC samples. However, there we were unable to adequately answer the question of whether expected parameter values could be extracted by the fit, since the fitting tool used was ill-suited for the problem.

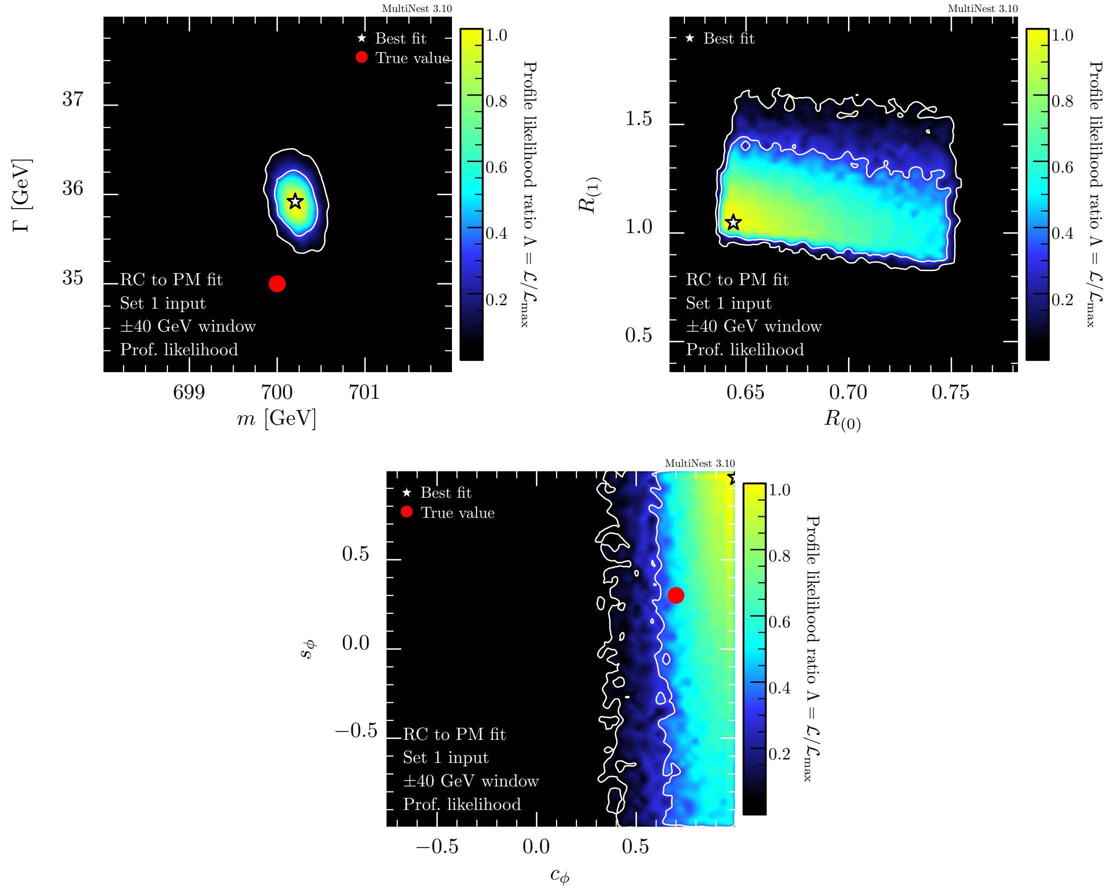

Figure 4.8: Profile likelihood contours for the fit result corresponding to input $m = 700$ GeV, $\Gamma/m = 5\%$, set 1 input PM parameters, and a fit window of $m \pm 40$ GeV. The white star and red circle respectively mark the best-fit and true (input PM) points.

Now, armed with the MultiNest sampling technique, let us study this problem again using a particular instance from the array of toy fits. We visualize the profile likelihood contours for the result corresponding to input $m = 700$ GeV and $\Gamma/m = 5\%$, with $w = 40$ GeV centred about the true mass. Fig. 4.8 shows the result for the set 1 inputs, and Fig. 4.9 the set 2.

For the RC parameters that have an analogous PM counterpart (m, Γ, c_ϕ, s_ϕ), expected input values have been marked using a red circle. We find these to agree very well with the high likelihood regions found in the set 2 results. For set 1, the true phase does not coincide with the best-fit point, but is nevertheless contained within the 1σ likelihood region. However, the true point in the m-Γ plane does not fall inside even the contour at the 2σ level. This initially appears to be a contradiction to the good fit reported by Fig. 4.7 for this set of inputs. However, we simply note that the parameters appearing in the general functional form cannot always be interpreted as exact counterparts to those appearing in a physics model, since they are only an approximation; what matters is that the lineshape of the physical distribution can be adequately characterized by the RC parameters.

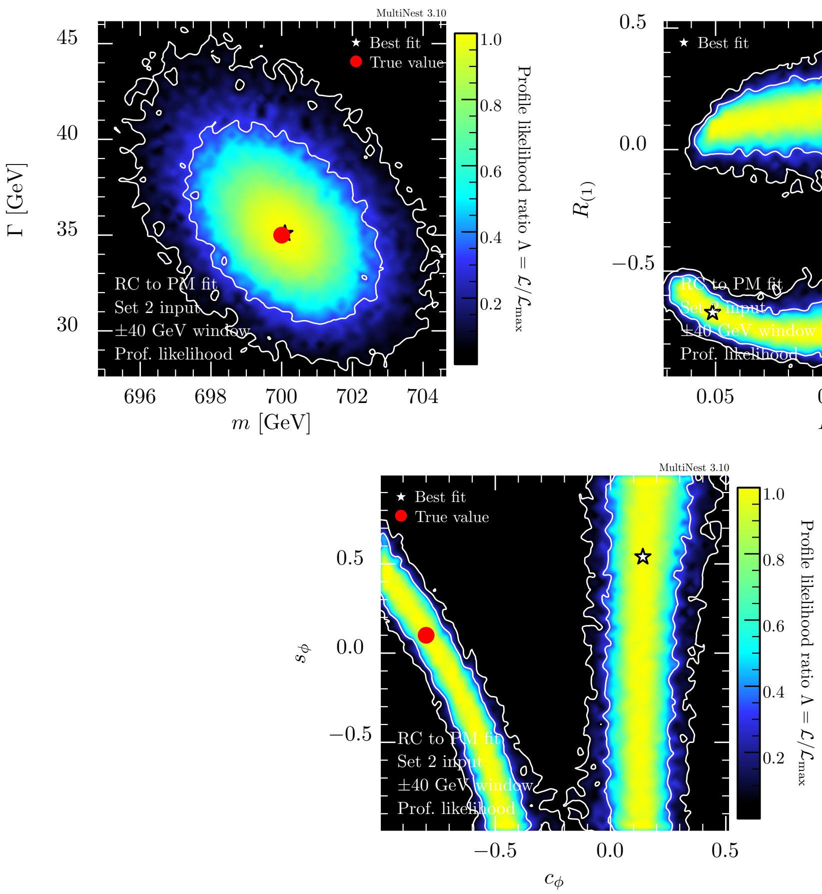

Figure 4.9: Profile likelihood contours for the fit result corresponding to input $m = 700$ GeV, $\Gamma/m = 5\%$, set 2 input PM parameters, and a fit window of $m \pm 40$ GeV. The white star and red circle respectively mark the best-fit and true (input PM) points.

Still, it might be desirable to explicitly confirm that the true input PM values can be extracted from these distorted results in RC space. We will do this by performing a closure fit of the physics model to a dataset representative of RC fit results.

4.4.2 Constructing histograms from fit results

Let us first explore methods for obtaining a histogram from the fit result of the general functional form. There are several plausible ways in which this can be done; the simplest is perhaps to consider the lineshape of only the best-fit result, and construct an Asimov histogram from this distribution under the same conditions as those used to create the initial PM toy (defined over the same invariant mass range, and using the same calculation of bin uncertainties).

However, this neglects the information available to us from the MultiNest sampling of RC parameter space. Let us instead describe a procedure that makes use of this: our aim is to fully utilize the MultiNest sampling output to construct a "results" histogram, using the corresponding functional form of those results. We

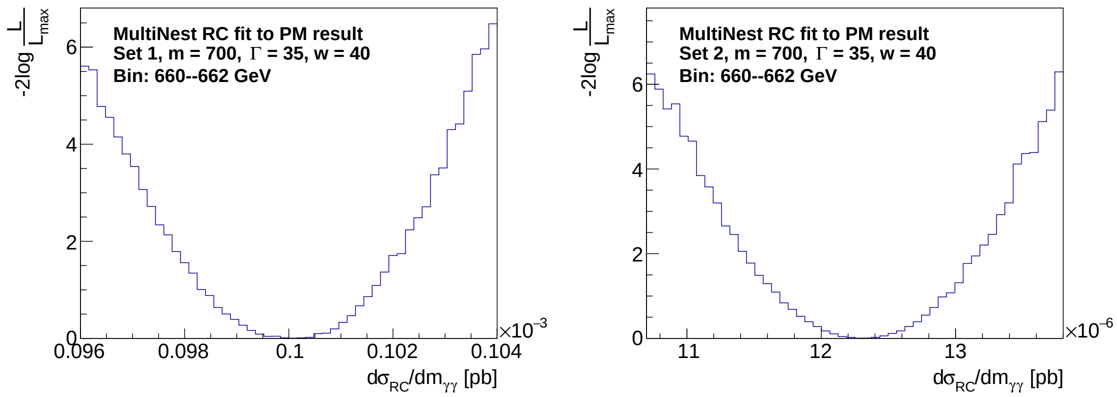

Figure 4.10: The distribution of the profile likelihood ratio, obtained from the MultiNest fit output, against the value of general functional form in the first bin of the constructed results histogram. Left (right): obtained using the set 1 (set 2) fit result. The parabolic distributions indicate a good approximation for Gaussian bin likelihoods.

define this histogram over an invariant mass range equal to the fit window of the initial fit, with the same binning as that of the original fitted dataset.

We then make the assumption of an independent Gaussian random variable for each bin. The validity of this assumption is checked for the MultiNest outputs corresponding to the results of Figs. 4.8 and 4.9. The likelihood distribution of the RC functional form in the first fitted bin is presented in Fig. 4.10. Indeed, we see an approximately Gaussian distribution for both results. Assuming this to be true in general, Eq. (4.10) can then be used to define a range of values, for each bin in a results histogram, corresponding to a 1σ uncertainty interval. Alternatively, one can note that the functional form generating this dataset is 6-dimensional; one would thus prescribe the condition of Eq. (4.11) instead, with $n = 6$ degrees of freedom. This will, of course, lead to larger bin uncertainties compared to the 1-dimensional assumption.

We test both methods of defining the 1σ interval by performing a simple toy closure test of the PM functional form. An Asimov toy is first created using an input PM distribution corresponding to the parameters:

$$\{m,\, \Gamma,\, c_\phi,\, s_\phi,\, f_s\} = \{400,\, 40,\, 0.5,\, 0.866,\, 0.4{\times}10^{-16}\}. \tag{4.48}$$

The PM functional form is then fitted to this toy histogram, and the result is used to construct results histograms using the two methods. A second fit of the PM functional form is then performed on the two results histograms; the results of these fits are compared to the initial fit to see which method provides a more accurate uncertainty definition.

We find that the second method, using Eq. (4.11) for determining likelihood cut-offs (with $n = 5$ for the PM functional form), more accurately reflects the uncertainties in the initial toy. The supporting figures are presented in App. C.3.

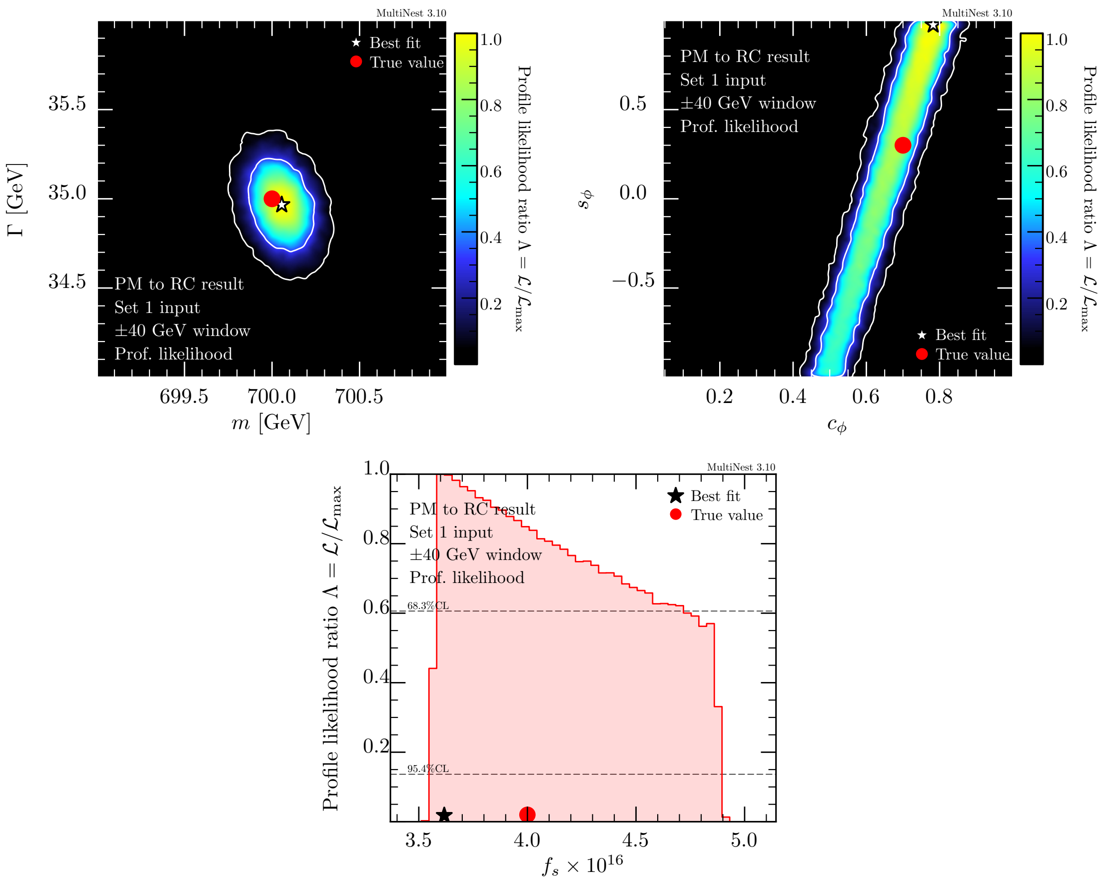

Figure 4.11: Closure fit of the PM functional form to the RC results histogram corresponding to $m = 700$ GeV, $\Gamma/m = 5\%$, set 1 input PM parameters, and a fit window of $m \pm 40$ GeV. The white star and red circle respectively mark the best-fit and true (input PM) points.

4.4.3 Physics model closure fit

Using the method described, we construct a histogram from the RC fit result for the set 1 inputs, with $m = 700$ GeV, $\Gamma/m = 5\%$, and $m \pm 40$ GeV fit window. The fit of the PM functional form to this histogram is presented in Fig. 4.11. Indeed, despite the disagreeing result of m and Γ in RC parameter space, we find that their expected input values can still be extracted correctly if the physics model parameter space is instead considered.

4.5 Significance of a discovery

So far, our results have been presented in terms of the profile likelihood ratio (Eq. (4.23)), suitable for the purpose of parameter estimation using the MLE method. However, this result makes sense only if one assumes the existence of a signal within the invariant mass window fitted; the profile likelihood simply presents a result that is relative to the point of best fit, regardless of the fit quality at that point. A sensible procedure for an analysis thus requires a method for first determining whether or not the existence of a signal is likely.

To do this, one can perform a test to determine whether any region of the observed invariant mass distribution deviates significantly from the expected background. The relevant test statistic is the likelihood ratio defined as

$$\Lambda_0(\boldsymbol{\theta}) \equiv \frac{\mathcal{L}_0}{\mathcal{L}(\boldsymbol{\theta})}\,, \tag{4.49}$$

or equivalently

$$q_0 = \begin{cases} -2\log\Lambda_0 & \text{for } 0 < \Lambda_0 \leq 1, \\ 0 & \text{for } \Lambda_0 > 1, \end{cases} \tag{4.50}$$

where $\mathcal{L}_0$ is the likelihood under a background-only assumption, and $\mathcal{L}(\boldsymbol{\theta})$ the likelihood of a signal-inclusive model at some point, $\boldsymbol{\theta}$, in its parameter space. Smaller (larger) values for Λ_0 (q_0) indicate an increasing disagreement between the observed data and the background-only assumption. This prescribes the procedure for a null (background-only) hypothesis test: its rejection, to some statistical degree, implies the alternative hypothesis (discovery of a signal).

A measure for quantifying this incompatibility is the p-value,

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)\, dq_0, \tag{4.51}$$

where $f(q_0|0)$ is the conditional PDF of q_0 under the null assumption, and $q_{0,\mathrm{obs}}$ is the observed q_0 value. The p-value is the probability of additional observations of the data being incompatible, to an equal or greater degree, with the background-only assumption; a small p-value can thus be interpreted as a low probability of obtaining the observed data if the background hypothesis were true, indicating a preference for signal-inclusive models.

To obtain the p-value, $f(q_0|0)$ is required. While its exact form will not be knowable, one can make a Gaussian approximation for Λ_0, analogous to Eq. (4.7). Assuming the validity of this approximation, the conditional PDF for q_0 can then be written as [39]

$$f(q_0|0) = \frac{\delta(q_0)}{2} + \frac{1}{2\sqrt{2\pi q_0}}\,\exp\left(-\frac{q_0}{2}\right), \tag{4.52}$$

which is a combination of a delta function at $q_0 = 0$, and a χ^2 distribution with one degree of freedom. One then finds the p-value to be

$$p_0 = 1 - F(q_{0,\mathrm{obs}}|0), \tag{4.53}$$

where $F(q_{0,\text{obs}}|0)$ is the cumulative distribution function (CDF) of $f(q_0|0)$,

$$F(q_{0,\text{obs}}|0) = \int_{-\infty}^{q_{0,\text{obs}}} f(q_0|0)\, dq_0, \tag{4.54}$$

which in this case is equivalent to the CDF of a standard Gaussian evaluated at $\sqrt{q_{0,\text{obs}}}$,

$$F(q_{0,\text{obs}}|0) \equiv F_G\left(\sqrt{q_{0,\text{obs}}}\right). \tag{4.55}$$

Commonly in particle physics, a result is also presented in terms of the (local) significance,

$$\mathcal{Z} = \mathcal{Q}_G(1-p), \tag{4.56}$$

where $\mathcal{Q}_G(p)$ is the inverse CDF of a standard Gaussian. Given a probability p, the significance corresponds to the number of standard deviations above the mean of the Gaussian such that an upper-tail probability of p is found. For the p_0 according to Eq. (4.53), a simple expression for the significance is found:

$$\mathcal{Z}_0 = \sqrt{q_0}\,. \tag{4.57}$$

Typically, a claim of discovery in particle physics accompanies a significance[3] of $\mathcal{Z}_0 = 5$ standard deviations, corresponding to a p-value of $p_0 \approx 2.87\times10^{-7}$.

4.5.1 Toy scan over m and Γ

Let us demonstrate the procedure for a background-only hypothesis test, using an Asimov dataset containing a signal at $m = 400$ GeV, with a width $\Gamma = 10$ GeV. The general functional form is used as the alternative to the null hypothesis.

We performed a scan over the m and Γ parameters by partitioning the m-Γ plane into a uniform grid of 100×100 pixels. Independent fits using MultiNest were performed for each pixel by constraining the uniform priors of the m and Γ parameters according to the boundaries of the partition. The maximum likelihood for each fit was extracted to calculate the corresponding significance, using Eqs. (4.49), (4.50) and (4.57).

Note that while we know the general functional form to provide a good approximation only in a restricted region about a signal peak, a large fit window can still be used during the m-Γ scan. This is because parameter estimation is not the purpose of the scan; we are only interested in whether a better alternative to the background-only description exists. Thus, it is not particularly necessary to precisely capture the lineshape of a resonance in this step, as long as it is possible to infer invariant mass regions of interest. Of course, if such a region is found to exist, the scan should be repeated to more accurately approximate the discovery significance.

[3] To be more precise, one requires this for the *global* significance, which takes into account the possibility for random fluctuations in any region of the data (the "look-elsewhere effect" (LEE)). Refer to [22, 60, 61] for a more detailed summary and study of this effect.

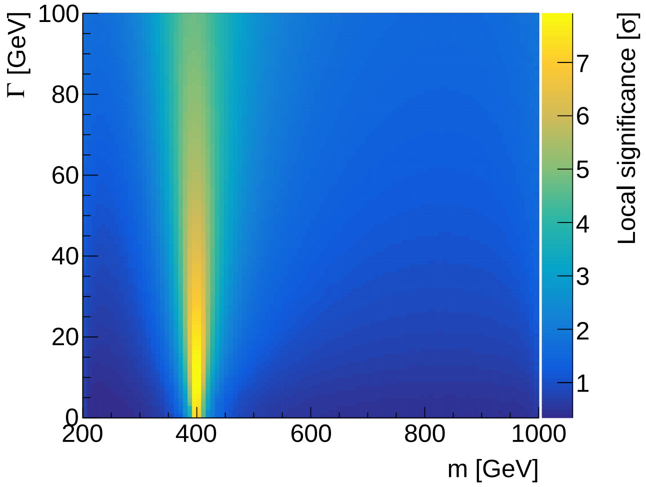

Figure 4.12: Local significance for rejecting the background-only hypothesis, for an Asimov dataset generated at input mass $m = 400$ GeV and width $\Gamma = 10$ GeV.

Our toy results are visualized in Fig. 4.12. A distinct region of high local significance can be identified at mass and width values expected from the input physics. In a blind search, such a result can be used to presuppose the existence and approximate locale of a resonance; one could then proceed with fits over this invariant mass region for the purpose of estimating the parameters of the general functional form.

4.6 Summary

In this chapter, we introduced the nested sampling algorithm and its MultiNest implementation as a means for performing global optimization problems involving the general functional form. Despite the large and non-trivial degeneracies between the RC parameters of the general functional form, MultiNest has proven to be robust and effective in sampling the problem space for the purpose of conducting frequentist tests under the MLE method.

The simplicity of results presented as a set of parameter values and their associated uncertainties is still attractive; we explored avenues for lifting the RC degeneracies by introducing the set of $a_{i,j}$ parameters, obtained through a modification at the amplitude level. However, we found the degeneracies between these $a_{i,j}$, and also the RC parameters, to persist.

We thus returned to the general functional form, to first order in the approximation for $R(q^2)$, as derived in the previous chapter. By performing fits to many toy Asimov histograms generated using the benchmark Higgs characterization model, we found the general template to provide a good description of data close to a peak, with fit quality deteriorating as the fitted mass window increases, and also as the size of the signal increases. Results were mostly unaffected by an overall shift of the fit window; despite being defined using a Taylor expansion about the true mass of a resonance, one does not need to know this value precisely before fitting with the general functional form.

We propose the procedure of an analysis, given a sample of events, as follows:

1. Perform a scan over the m-Γ plane of the general functional form, and produce a plot of discovery significance. The remaining parameters are to be profiled over, to find the maximum likelihood at each point of the scan.

2. Identify any invariant mass regions of high local significance; if any exist, perform fits to the data over these regions, using fit windows of increasing size.

3. Extract the parameters of the general functional form using the largest possible (or sufficiently large) fit window that still yields a good fit to the data. Such a window can be found by considering the χ^2 of a fit relative to the 1σ threshold of a χ^2_ν distribution, for an appropriate number of degrees of freedom, ν.

Chapter 5

Detector Effects

The studies so far have been performed using ideal datasets that exhibit characteristics entirely due to an assumed physics model. To these truth-level distributions, a direct fit of the general functional form then revealed insight into the model parameters behind the data. However, such a procedure neglects any effects due to realistic particle reconstruction in a detector, which would result in the observation not of the truth, but instead of a reconstructed distribution with features that are a distortion of the truth.

Due to these detector-based effects, the standard procedure of an analysis is to model the shape of a signal contribution to an invariant mass distribution as a convolution [4]:

$$F_{\mathcal{S}}^{\text{reco}}(q) = (F_{\mathcal{S}}^{\text{truth}} * \text{DR})(q) = \int_{-\infty}^{\infty} F_{\mathcal{S}}^{\text{truth}}(Q)\,\text{DR}(q-Q)\,dQ, \tag{5.1}$$

where $F_{\mathcal{S}}^{\text{truth}}(q)$ is the truth-level signal description predicted by an assumed physics model, and $\text{DR}(q)$ denotes a *detector resolution* function, which can be parametrized by careful construction of the detector in Monte Carlo studies.

While Eq. (5.1) is written specifically to describe a reconstructed signal, the convolution is also valid for the interference and background components. However, a direct parametrization of the reconstructed background is typically found instead, either by extrapolation from Monte Carlo studies or with a functional form fit to observed data. Current procedures also often neglect interference, so that the convolution is only necessary for the signal component; data is then described with

$$N_{\mathcal{S}}\left|F_{\mathcal{S}}^{\text{reco}}(q)\right|_{\text{norm}} + N_{\mathcal{B}}\left|F_{\mathcal{B}}^{\text{reco}}(q)\right|_{\text{norm}}, \tag{5.2}$$

where $F_{\mathcal{S}}^{\text{reco}}(q)$ and $F_{\mathcal{B}}^{\text{reco}}(q)$ are treated as PDFs of the reconstructed signal and background invariant mass lineshapes (and thus normalized to one), and assuming a dataset in terms of a differential number of events, the coefficients $N_{\mathcal{S}}$ and $N_{\mathcal{B}}$ represent the number of signal and background events respectively, with values to be extracted from a fit.

A description involving the general functional form is not so straightforward, as it contains terms associated with interference; an analogous $N_{\mathcal{I}}|F_{\mathcal{I}}^{\text{reco}}(q)|$ term cannot simply be added to Eq. (5.2), since the normalization of the interference will depend on both the signal and the background. Furthermore, the $R_{(i)}$ parameters are loosely derived from the ratio of signal to background amplitudes; their correct extraction thus relies on knowing the background at the truth-level.

In this chapter, MadGraph5_aMC@NLO [29] and Pythia8 [32] will be used to generate events that will then be passed through the Delphes 3.4.1 [33] detector simulation to yield pseudo-realistic samples of reconstructed events. The scalar Higgs-like signal decaying to two photons of the HC model [26] is again assumed. We will begin by describing Delphes 3.4.1 in more detail, particularly for the case of photon reconstruction, before verifying that it produces a diphoton invariant mass distribution well described by Eq. (5.1). We then seek a general description of reconstructed distributions by convolving the general functional form with a detector resolution function, and confirm that results consistent with those at the truth-level can be extracted.

5.1 Delphes 3 fast detector simulation

Monte Carlo simulations play an integral role in collider physics, providing a means for refining search strategies prior to an experiment in order to optimize the uptime of the apparatus. This entails an extensive and detailed simulation of detector effects, usually achieved via a recreation of the detector using the Geant4 toolkit [62–64]. Naturally, such a simulation is computationally expensive; thus, when stringent levels of accuracy are unnecessary (for example, in phenomenological studies), one might seek a cheaper alternative.

Delphes is a framework designed for this purpose. Its generic detector set-up comprises of, in order: an inner tracking volume, electromagnetic (ECAL) and hadron (HCAL) calorimeters, and a muon identification system. These are cylindrically centred about the beam axis, with the volume of the two calorimeters exactly overlapping. The basic operation of the Delphes detector response is to read, as input, the stable (or sufficiently long-lived) particles from an event generator such as Pythia8 [32], and perform a smearing of their momentum vectors to simulate the limited resolution of a detector.

The simulation first propagates the long-lived particles through the inner tracking volume, which contains a uniform magnetic field aligned parallel to the beam axis. In this region, electrically charged particles follow a helical trajectory. Delphes reconstructs the tracks of these particles with some user-defined efficiency, and assumes a perfect angular resolution with smearing only of the transverse component of particle momenta. Neutral particles such as photons simply trace a straight-line path to the calorimeters.

After passing through the inner volume, particles arrive at the calorimeters, which are segmented in the pseudorapidity, η, and azimuthal angle, ϕ. A fixed fraction of their energy is then deposited in the calorimeter cell they arrive at. By default, electrons and photons deposit all their energy in the ECAL, while neutral and charged hadrons deposit in the HCAL. The η-dependent energy resolution, σ, is calculated independently for both calorimeters:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S(\eta)}{\sqrt{E}}\right)^2 + \left(\frac{N(\eta)}{E}\right)^2 + C(\eta)^2, \tag{5.3}$$

where E is the energy, and $S(\eta)$, $N(\eta)$, and $C(\eta)$ respectively denote the stochastic, noise, and constant contributions to the resolution. Using this, the total energy in a given tower of calorimeter cells is then smeared using a log-normal distribution, $\log\mathcal{N}(m, s)$, with mean m and variance s, according to

$$E_{\text{tower}} = \sum_{\text{particles}} \log\mathcal{N}\left(f_{\text{E}}E, \sigma_{\text{E}}(E,\eta)\right) + \log\mathcal{N}\left(f_{\text{H}}E, \sigma_{\text{H}}(E,\eta)\right), \tag{5.4}$$

where the sum is performed over all particles that deposited some fraction of their energy in the tower, either in the ECAL (with fraction f_E) or HCAL (f_H).

True photons in the ECAL, with their energies smeared according to Eq. (5.4), along with electrons whose tracks were not reconstructed in the inner detector, are reconstructed as photons by DELPHES. To ensure that these are not part of a jet, an isolation variable is also defined for each reconstructed particle:

$$I(P) = \frac{1}{p_T(P)} \sum_{i \neq P}^{\Delta R < R,\, p_T(i) > p_T^{\min}} p_T(i), \tag{5.5}$$

where $p_T(P)$ denotes the transverse momentum of particle P, and

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < R, \tag{5.6}$$

defines a conical region of radius R around the particle. $I(P)$ is thus a measure of any additional transverse momenta in the vicinity of a particle, relative to its own p_T. A large value for $I(P)$ indicates the significant presence of other particles, while $I(P) = 0$ indicates perfect isolation.

To simulate pile-up effects that occur in the high luminosity bunch-crossings of the LHC, DELPHES also makes use of a pre-generated sample of low-q^2 QCD events, placed randomly along the beam axis. The average contamination density of these interactions, ρ, is estimated using the Jet Area method with the FASTJET package [65–67]. Pile-up subtraction results in a correction to the isolation variable,

$$I(P) \to I(P) - \frac{\rho\pi R^2}{p_T(P)}\,. \tag{5.7}$$

A particle is considered to be isolated if $I(P) < I_{\min}$. The default values used by DELPHES are $p_T^{\min} = 0.5$ GeV, $R = 0.5$, and $I_{\min} = 0.12$.

The result of the DELPHES simulation is an output in the ROOT data format. To reconstruct the diphoton invariant mass distribution, we run an analysis macro to identify all events containing at least two photons. The most energetic pair of photons in each event is considered to contain the prompt diphotons originating from the resonance decay. However, this will inadvertently include photons not produced by the signal, and additionally also the untracked electron pairs misidentified as photons.

In contrast to the photons produced in a hard scattering event, these "soft" photons will typically have smaller transverse momenta. To reject such events, we will impose kinematic cuts on the lower end of the p_T spectrum, requiring

$$p_T > p_T^{\text{cut}} , \tag{5.8}$$

for some choice of p_T^{cut}.

5.2 Double-Sided Crystal Ball

In this section, we test the analytical approximation of Eq. (5.1) against the reconstructed diphoton invariant mass distributions simulated by Delphes 3.4.1. The true physics model functional form (see Sec. 3.1), with all parameters fixed from a fit to truth events, will be used as the truth-level description. The detector resolution function will be parametrized as a Double-Sided Crystal Ball (DSCB) function, which comprises of a Gaussian core with power law tails:

$$\text{DSCB}(q) = N_{\text{DSCB}} \begin{cases} \exp\left(-\frac{t^2}{2}\right) & \text{for } -\alpha_{\text{low}} \le t \le \alpha_{\text{high}}, \\ \dfrac{\exp\left(-\frac{\alpha_{\text{low}}^2}{2}\right)}{\left[\frac{\alpha_{\text{low}}}{n_{\text{low}}}\left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t\right)\right]^{n_{\text{low}}}} & \text{for } t < -\alpha_{\text{low}}, \\ \dfrac{\exp\left(-\frac{\alpha_{\text{high}}^2}{2}\right)}{\left[\frac{\alpha_{\text{high}}}{n_{\text{high}}}\left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t\right)\right]^{n_{\text{high}}}} & \text{for } t > \alpha_{\text{high}}, \end{cases} \tag{5.9}$$

where

$$t = \frac{q - \mu_{\text{DSCB}}}{\sigma_{\text{DSCB}}} . \tag{5.10}$$

The parameters μ_{DSCB} and σ_{DSCB} are the mean and standard deviation of the Gaussian core, while the n_{low} and n_{high} exponents dictate the distribution of the power law tails. The points at which the Gaussian changes into the tail distributions is determined by the α_{low} and α_{high} parameters. N_{DSCB} is a factor that normalizes the DSCB to unity.

5.2.1 Narrow-width signal modelling

In the scenario that the width of a signal is small, the truth-level description is highly peaked at the resonance mass:

$$F_{\text{NWA}}^{\text{truth}} \sim \delta(q - m). \tag{5.11}$$

Thus, if the description of Eq. (5.1) can be assumed, it follows that the reconstructed distribution will be representative of the detector resolution function,

$$F_{\text{NWA}}^{\text{reco}} \approx \text{DR}(q). \tag{5.12}$$

To verify the DSCB function for parametrizing the Delphes 3.4.1 detector resolution, we test its description of reconstructed signals generated under the narrow width approximation (NWA). The signal samples were generated using MadGraph5_aMC@NLO at six mass points, m = 200, 250, 300, 400, 600, and 800 GeV, with widths set to $\Gamma = 4$ MeV. In extracting the reconstructed photons from these samples, we imposed the particular selection criterion of $p_T > 60$ GeV.

ROOT's histogram fitting method was used to perform the seven-dimensional fit of the DSCB parameters:

$$\{\mu_{\text{DSCB}}, \sigma_{\text{DSCB}}, \alpha_{\text{low}}, \alpha_{\text{high}}, n_{\text{low}}, n_{\text{high}}, N\}, \tag{5.13}$$

where note that N is simply an overall parameter and does not necessarily correspond to the N_{DSCB} normalization factor in Eq. (5.9). Fig. 5.1 shows the results. A good fit to each of the reconstructed distributions is found. However, the best-fit parameter values are not consistent between the fits; most notably, the width of the Gaussian core, σ_{DSCB}, increases for the signals at higher mass points, indicating a finer resolution of less energetic photons. The large uncertainties associated with n_{low} and n_{high} arise due to their large correlations with the other parameters.

To generalize our results to yield a parametrization applicable at any mass point, we assume a linear dependency between each DSCB parameter and the invariant mass. Using the fit results at the six mass points generated, best-fit values and uncertainties are plotted against the mass; a linear relationship is then extrapolated through a fit of the straight-line function:

$$Aq + B. \tag{5.14}$$

Fig. 5.2 shows the results. We find the linear approximation to hold sufficiently well. These results prescribe a q-dependent function for each DSCB parameter; evaluating at a given q defines an approximation of the Delphes 3.4.1 detector resolution at that invariant mass.

5.2.2 Large-width modelling with PM

The method of extracting DSCB parameters by extrapolation of NWA fit results allows us to predict the reconstructed lineshape of a signal at any mass point. We now demonstrate the convolution of Eq. (5.1) for describing a large-width signal generated at the mass point m = 400 GeV, with width Γ = 40 GeV. The same $p_T > 60$ GeV condition is again imposed during the extraction of photons.

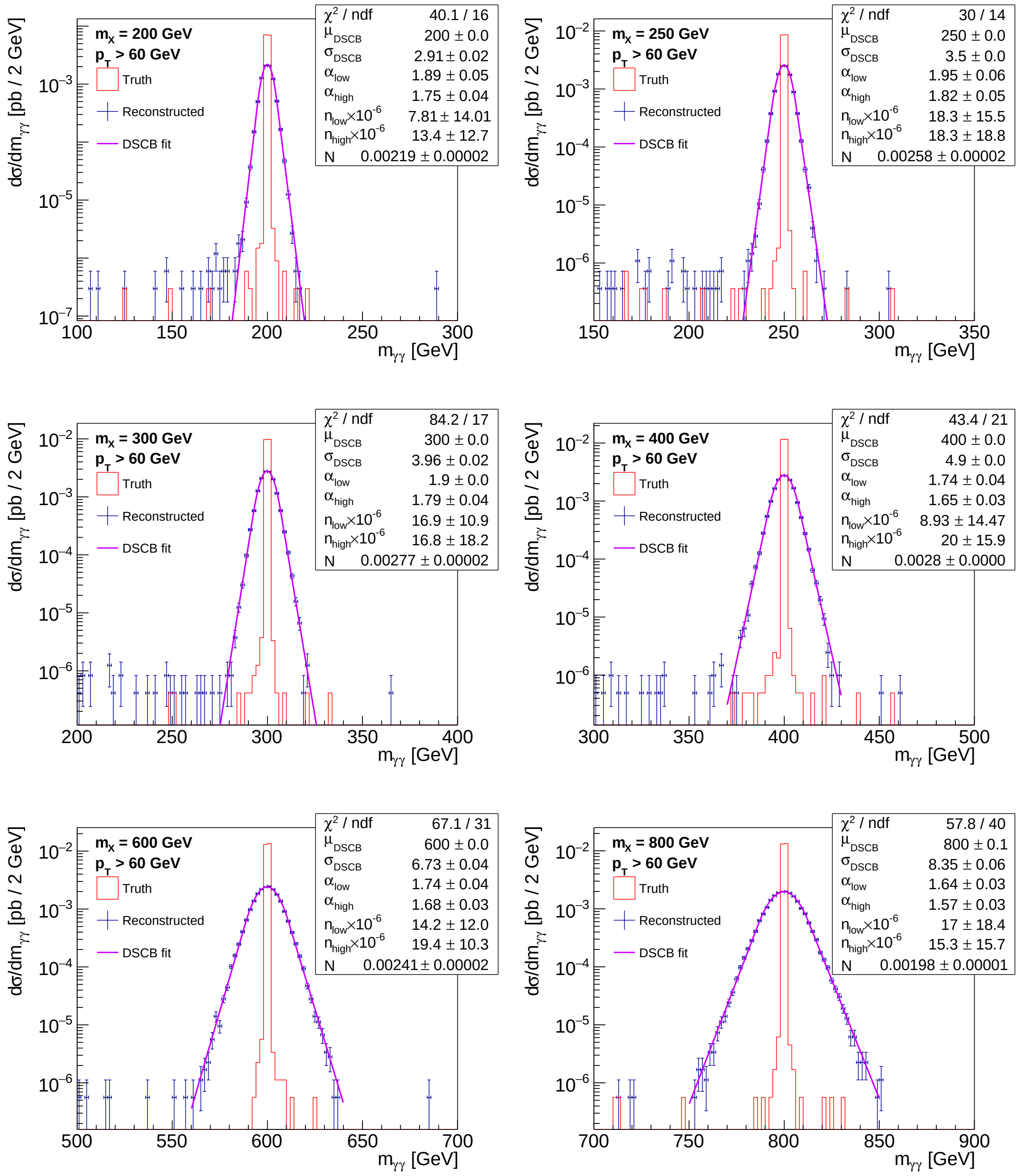

Figure 5.1: Fits of the DSCB function to reconstructed NWA signal samples, generated at six different mass points. A $p_T > 60$ GeV cut has been imposed on the reconstructed photons.

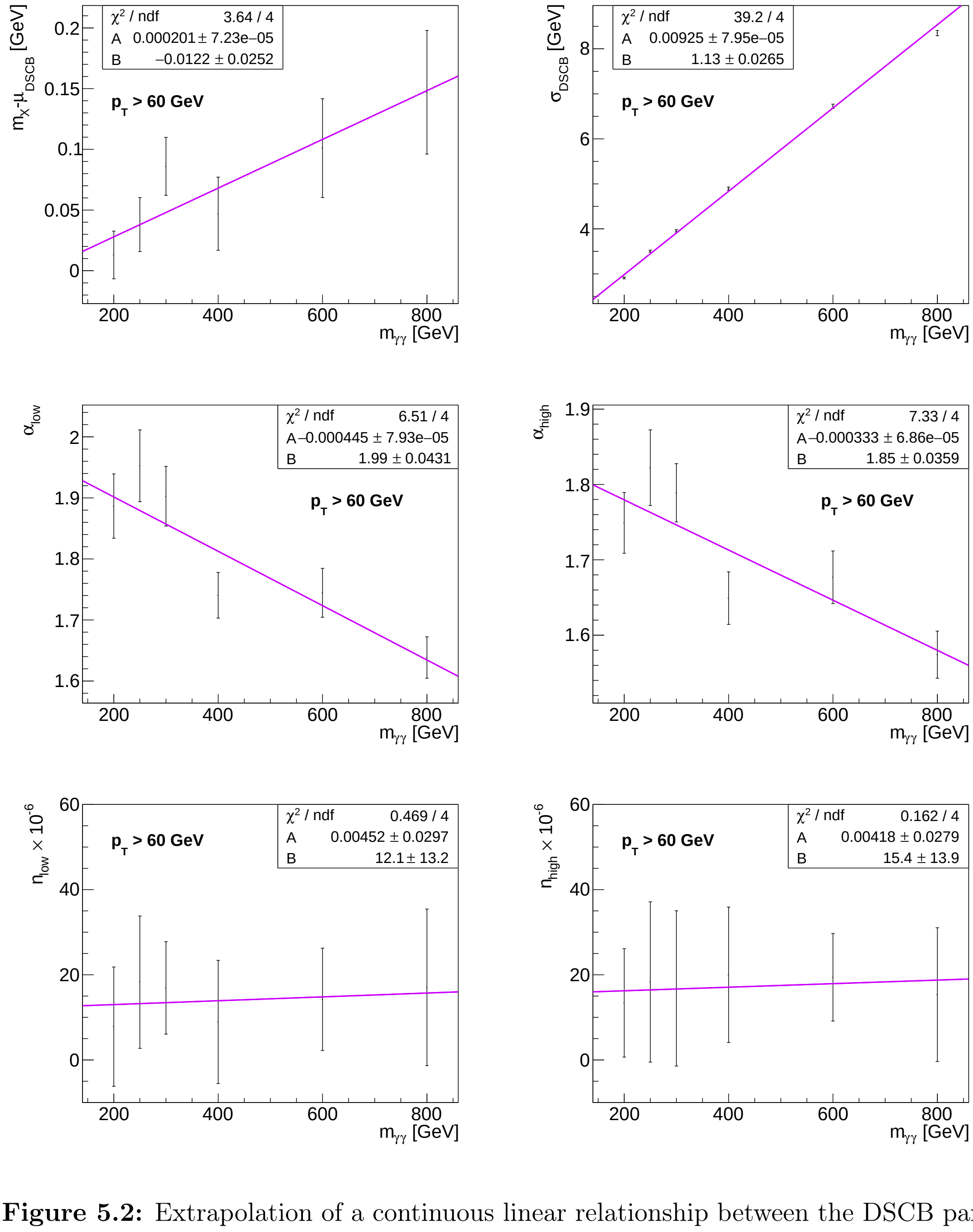

Figure 5.2: Extrapolation of a continuous linear relationship between the DSCB parameters and the diphoton invariant mass, using the NWA samples generated at six discrete mass points.

Recall the analytical description of the signal model found previously:

$$\frac{d\sigma_{\mathcal{S}}}{dq} \propto \mathcal{L}_{gg}(q^2) \times q^7 \times f_{\mathrm{BW}} \times |A(\tau)|^2. \qquad \text{(3.23 revisited)}$$

A definite set of values for its parameters (m, Γ, and normalization f_s) was first fixed by fitting the MC sample at the truth-level. This distribution was then convolved with the DSCB, with all its parameters fixed according to the results extrapolated from the NWA fits:

$$\begin{aligned}
\mu_{\mathrm{DSCB}} &= 0, &(5.15)\\
\sigma_{\mathrm{DSCB}}(q=m) &= \sigma^A_{\mathrm{DSCB}} m + \sigma^B_{\mathrm{DSCB}}, &(5.16)\\
\alpha_{\mathrm{low}}(q=m) &= \alpha^A_{\mathrm{low}} m + \alpha^B_{\mathrm{low}}, &(5.17)\\
\alpha_{\mathrm{high}}(q=m) &= \alpha^A_{\mathrm{high}} m + \alpha^B_{\mathrm{high}}, &(5.18)\\
n_{\mathrm{low}}(q=m) &= n^A_{\mathrm{low}} m + n^B_{\mathrm{low}}, &(5.19)\\
n_{\mathrm{high}}(q=m) &= n^A_{\mathrm{high}} m + n^B_{\mathrm{high}}, &(5.20)
\end{aligned}$$

where quantities with superscript A's and B's denote the linear coefficients found for the corresponding parameter in Fig. 5.2, and m denotes the mass of the large-width resonance.

We similarly perform a fit of the reconstructed interference and background distributions. The results are presented in Fig. 5.3. The single free parameter in these fits is an overall normalization,

$$N = N_{\mathrm{DSCB}} \times \varepsilon, \tag{5.21}$$

which receives a contribution from the normalization of the DSCB, as well as the selection efficiency, ε, of the reconstructed events. The efficiency represents the fraction of reconstructed photons retained after the p_T selection cuts have been applied. Overall, a good description of the reconstructed lineshapes is found for all the components. However, we find a small discrepancy between the values of N found. Since a common DSCB parametrization was used in the fits (thus corresponding to the same N_{DSCB}), Eq. (5.21) implies that there is a discrepancy in the efficiencies of each component.

There is also a clear fall-off in the data at low invariant masses, especially visible in the $m_{\gamma\gamma} \lesssim 200$ GeV region of the signal, and $m_{\gamma\gamma} \lesssim 300$ GeV in the background. This arises due to the p_T cut imposed on the reconstructed photons; the effect will ultimately be model-dependent, and cannot be predicted using the simple description provided by the convolution with a DSCB. While it is possible to optimize cuts to extend the region of good fit to lower $m_{\gamma\gamma}$ in the current study involving Monte Carlo samples, a baseline trigger-level selection criteria will be dependent upon the characteristics of the detector, and cannot be chosen arbitrarily. In a realistic search procedure, it might thus be necessary to restrict the search range to

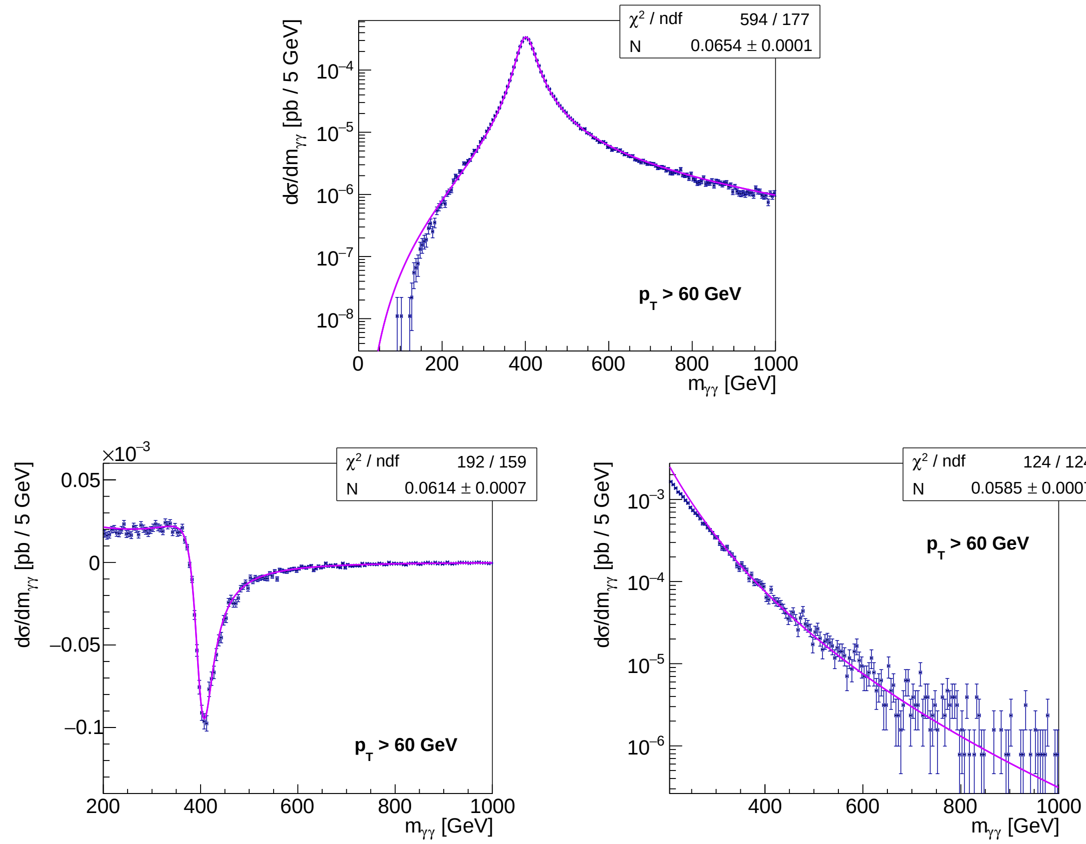

Figure 5.3: A fit of the reconstructed signal, interference, and background distributions using a convolution of their respective truth-level lineshapes and a DSCB detector resolution function. Fits were performed over 0–1000 GeV (signal), 200–1000 GeV (interference), and 300–1000 GeV (background). The single free parameter in each fit is N, an overall normalization factor.

above some invariant mass threshold, to avoid a mis-characterization of the signal lineshape.

5.3 Discrepancy in selection efficiencies

The results found in the previous section have shown that the convolution description of Eq. (5.1) provides a good approximation of the reconstructed signal lineshape, and holds also for reconstructed interference and background distributions, which is a necessary extension since a detector cannot realistically distinguish diphoton pairs based on their production method. However, photons produced in different processes will display different p_T distributions; thus, the imposition of a hard cut on minimum photon p_T will reject differing fractions of signal, interference, and background events from the reconstructed record, reflected in the discrepant normalization parameters and the reduced number of events at low invariant masses.

Table 5.1: Normalizations extracted when different p_T cuts are applied.

p_T^{cut} [GeV]	N_s (± 0.0001)	N_i (± 0.0008)	N_b (± 0.0008)
0	0.0669	0.0661	0.0666
30	0.0664	0.0658	0.0648
45	0.0662	0.0652	0.0622
60	0.0654	0.0635	0.0585

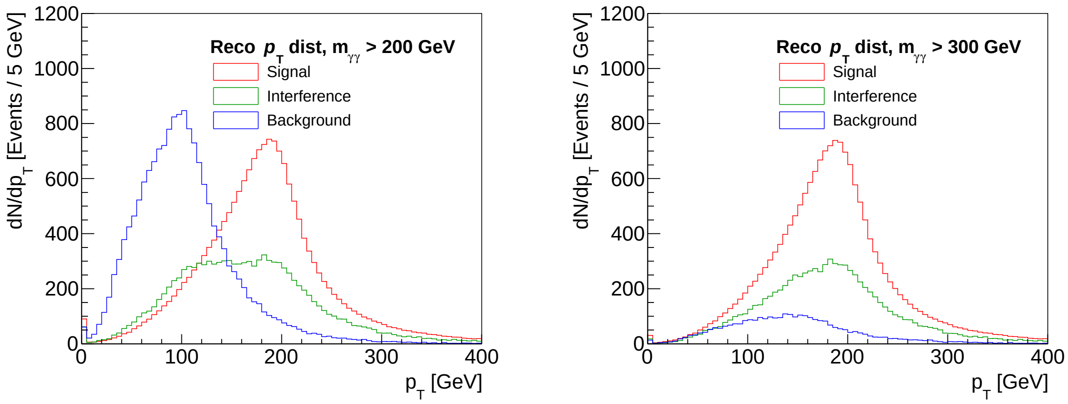

Figure 5.4: Reconstructed signal, interference, and background photon p_T distributions. Left (right): including only the diphoton pairs with an invariant mass above 200 GeV (300 GeV). As the minimum $m_{\gamma\gamma}$ threshold increases, fewer low p_T photons remain.

In this section, we further investigate this discrepancy between the efficiencies of the signal, interference and background components. To verify that this effect arises from the selection criteria we imposed on the photons, we will begin by repeating the procedure of the previous section using a range of different p_T cuts. Tab. 5.1 shows the best-fit normalizations found for each choice of cut. Indeed, in the case where the reconstructed photons are extracted simply as the leading and sub-leading pair in an event with no additional constraints on their p_T, the three normalizations agree within reported uncertainties. A discrepancy is apparent only when a p_T cut is introduced to the analysis, becoming more pronounced as the cut threshold increases. The background normalization, N_b, experiences the largest rate of decrease as the minimum p_T increases, while the signal normalization, N_s, varies the least.

The reconstructed photon p_T distributions of each component are presented in Fig. 5.4 (left). We find that more photons of the background lie at the lower end of the p_T spectrum. A cut will thus reject a greater fraction of events from the background sample, in comparison to the interference and signal; this is the reason behind the discrepancy seen in their respective invariant mass distributions. However, note that these distributions consider all of the photons that result in an event with invariant mass $m_{\gamma\gamma} > 200$ GeV. We previously noted that it might be necessary to prescribe a search range above some minimum invariant mass threshold, due to possible distortions in the reconstructed lineshapes at low invariant mass regions.

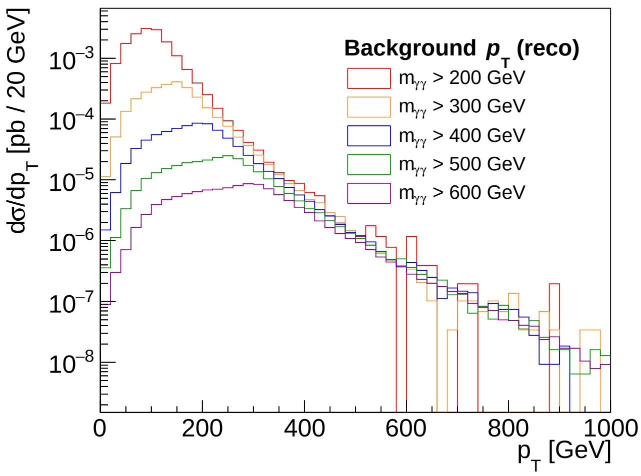

Figure 5.5: Reconstructed background p_T distributions, with various minimum $m_{\gamma\gamma}$ constraints imposed on the photons.

Table 5.2: The fraction of reconstructed background photons with a p_T less than 60 GeV ($f^{\gamma,\text{bck}}_{p_T<60}$), as a percentage of the total number of photons in the sample, for various minimum diphoton invariant mass conditions ($m^{\text{min}}_{\gamma\gamma}$).

$m^{\text{min}}_{\gamma\gamma}$ [GeV]	$f^{\gamma,\text{bck}}_{p_T<60}$ [%]
200	17.18
300	6.83
400	3.33
500	1.75
600	0.97

Fig. 5.4 (right) shows the p_T distributions when a $m_{\gamma\gamma} > 300$ GeV condition is imposed instead. Compared to the left-hand plot, very few low p_T photons remain.

Fig. 5.5 further compares the p_T distributions of background photons for increasing cut-off values of the diphoton invariant mass. Tab. 5.2 shows the corresponding fraction of photons which would be rejected if a $p_T > 60$ GeV cut were imposed. For $m_{\gamma\gamma} > 200$ GeV, almost a fifth of the background photons are rejected by the $p_T > 60$ GeV cut. The result in Fig. 5.3 corresponds to $m_{\gamma\gamma} > 300$ GeV, with approximately 7% of the events removed. The situation improves as the minimum cut-off increases; if one can restrict an analysis to the invariant mass range above 600 GeV, for example, then fewer than 1% of the background events in our current MC sample will be rejected.

However, note that the fitted normalizations differ only slightly between the samples, agreeing to one significant figure (two decimal places) in the results of Fig. 5.3. Realistically, this discrepancy will probably be sub-dominant in comparison to other sources of uncertainty; their differences are pronounced in the current study due only to the very precise uncertainties reported by the fits, consequent of the large number of MC events generated. Experimental datasets will typically contain fewer events, leading to larger uncertainties in N values. Furthermore, background distributions are often falling functions of the invariant mass, leading to fewer events, and thus larger uncertainties, at higher invariant masses.

Thus, we suggest that the approximation of equal efficiencies will probably prove valid in realistic searches. Distortions in the results of fits to reconstructed data will be inevitable, but mostly negligible; to minimize these effects, one could also choose to begin a fit at a higher $m_{\gamma\gamma}$ point, where data-driven constraints on the background will be less strict. Regardless, the imposition of necessary detector-based selection criteria will place a lower bound on the invariant mass range that a search can be performed over. If it is imperative to accurately parametrize data at low $m_{\gamma\gamma}$, it would be preferable to instead seek a better functional description of detector effects; a possible solution might be to perform the convolution over all relevant physical variables in addition to the invariant mass, but we will not explore this idea further.

5.4 General functional form (reconstructed)

We now proceed to apply the convolution method of incorporating detector effects to the model-independent functional form, which can be written in the form:

$$\left[\frac{d\sigma}{dq}\right]_{\mathrm{tr}} = F_{\mathcal{B}}^{\mathrm{tr}}(q)\left\{S + I + 1\right\}, \tag{5.22}$$

where S and I represent the signal and interference terms in the functional form respectively, and $F_{\mathcal{B}}^{\mathrm{tr}}(q)$ describes the truth-level background distribution. Following from Eq. (5.1) and the previous discussion, a reconstructed distribution can be described with

$$\left[\frac{d\sigma}{dq}\right]_{\mathrm{reco}} = N_s\left[\left(F_{\mathcal{B}}^{\mathrm{tr}} \times S\right) * \mathrm{DSCB}\right] + N_i\left[\left(F_{\mathcal{B}}^{\mathrm{tr}} \times I\right) * \mathrm{DSCB}\right] + N_b\left[F_{\mathcal{B}}^{\mathrm{tr}} * \mathrm{DSCB}\right], \tag{5.23}$$

where the pre-factors are given by Eq. (5.21), so the DSCB is not explicitly normalized to unity. However, one typically finds a function that directly describes the reconstructed background,

$$F_{\mathcal{B}}^{\mathrm{reco}} \equiv N_b\left[F_{\mathcal{B}}^{\mathrm{tr}} * \mathrm{DSCB}\right]. \tag{5.24}$$

This can be simplified because the background is often a smoothly varying function, such that the convolution with the (un-normalized) DSCB simply introduces its appropriate normalization factor,

$$F_{\mathcal{B}}^{\mathrm{reco}} \approx N_b\left[\frac{F_{\mathcal{B}}^{\mathrm{tr}}}{N_{\mathrm{DSCB}}}\right]. \tag{5.25}$$

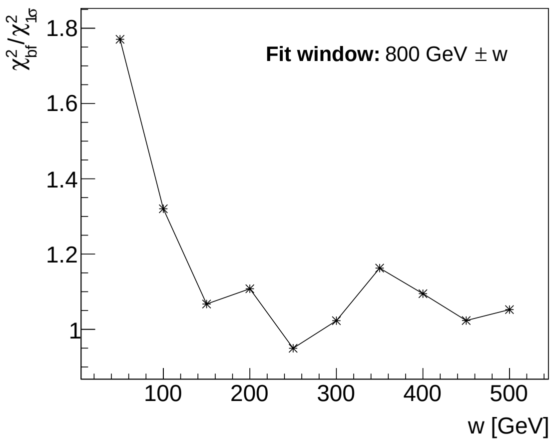

Figure 5.6: The variation of $\chi^2_{\mathrm{bf}}/\chi^2_{1\sigma}$ with the fit window used, for fits of the convolved template functional form to reconstructed MadGraph5_aMC@NLO events.

Eq. (5.23) can thus be rewritten as

$$\left[\frac{d\sigma}{dq}\right]_{\mathrm{reco}} = \frac{\varepsilon_s N_{\mathrm{DSCB}}}{\varepsilon_b}\left[(F_{\mathcal{B}}^{\mathrm{reco}} \times S) * \mathrm{DSCB}\right] + \frac{\varepsilon_i N_{\mathrm{DSCB}}}{\varepsilon_b}\left[(F_{\mathcal{B}}^{\mathrm{reco}} \times I) * \mathrm{DSCB}\right] + F_{\mathcal{B}}^{\mathrm{reco}}\,. \tag{5.26}$$

In the previous section, we discussed the validity of assuming equal efficiencies, $\varepsilon_s \approx \varepsilon_i \approx \varepsilon_b$. Under this approximation, the coefficients of the signal and interference terms in Eq. (5.26) reduce simply to N_{DSCB}, whose value can easily be computed for a given DSCB function. Hence, the convolution of the general functional form with a detector resolution function introduces no additional free parameters, and a fit to reconstructed events can be performed using the same set of parameters as those in the truth scenario: m, Γ, $R_{(0)}$, $R_{(1)}$, c_ϕ, and s_ϕ.

5.4.1 Fit to reconstructed MC

Eq. (5.26) was tested as follows. First, a description of the background component $F_{\mathcal{B}}^{\mathrm{reco}}$ was found by fitting a sample of reconstructed background-only events. We then noted that the assumption made in going from Eqs. (5.24) to (5.25) allowed us to find the appropriate normalization factor for the DSCB by performing a second fit to the reconstructed background, using the function defined as

$$N \times (F_{\mathcal{B}}^{\mathrm{reco}} * \mathrm{DSCB}). \tag{5.27}$$

The free parameter in this fit yields the value for N_{DSCB}:

$$N \equiv N_{\mathrm{DSCB}}. \tag{5.28}$$

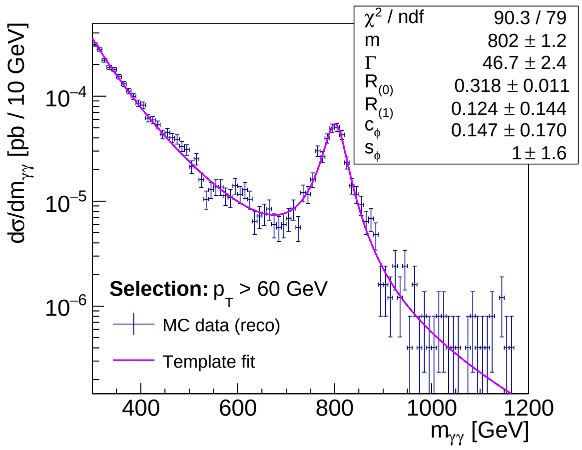

Figure 5.7: A fit of the convolved template functional form to the reconstructed MadGraph5_aMC@NLO sample over the 300–1200 GeV mass range.

With this and a description for $F_{\mathcal{B}}^{\text{reco}}$, we attempted fits of Eq. (5.26) to a sample of reconstructed events. Following the discussion of the previous section, we chose to generate a new MC sample with a more realistic number of events, to improve the approximation of equal selection efficiencies between the three components. The signal was generated at $m = 800$ GeV, with $\Gamma = 40$ GeV. A cut of $p_T > 60$ GeV was again imposed during the extraction of reconstructed photons. Preliminary fits were then performed using ROOT and various fit windows. Fig. 5.6 shows the ratio of best-fit (χ^2_{bf}) to 1σ cut-off $(\chi^2_{1\sigma})$ chi-squares against the windows tested. Unlike our previous results at the truth-level, a monotonically increasing ratio with the fit window is not seen. This is due to the stochastic nature of MC samples, in contrast to Asimovs that perfectly capture the underlying theory. To prevent random fluctuations from yielding volatile results, a sufficiently large fit window will thus be required; in the particular case of Fig. 5.6, the ratio of chi-squares reaches a minimum at $w = 250$ GeV. However, the distribution effectively plateaus past $w \approx 150$ GeV and does not worsen significantly for larger w, suggesting a resonance small enough that the entirety of its effect on the invariant mass distribution can be described. Indeed, Fig. 5.7 shows that the general template is able to fit the reconstructed distribution well over a large 300–1200 GeV mass range. For samples with larger signals, one would expect to instead see a relationship with a clear minimum in w.

An estimation of the parameters of the functional form was then performed using MultiNest, over a restricted window of 650–950 GeV about the signal peak. We additionally performed a fit to the truth-level distribution over the same window. The results from these two minimizations are compared in Fig. 5.8, with the reconstructed contours in colour and the truth in grey. We find a good agreement in the m-Γ contour between the two sets of results, with largely overlapping 1σ contours. For the other plots, we see two distinct 1σ regions at the truth-level, enveloped by

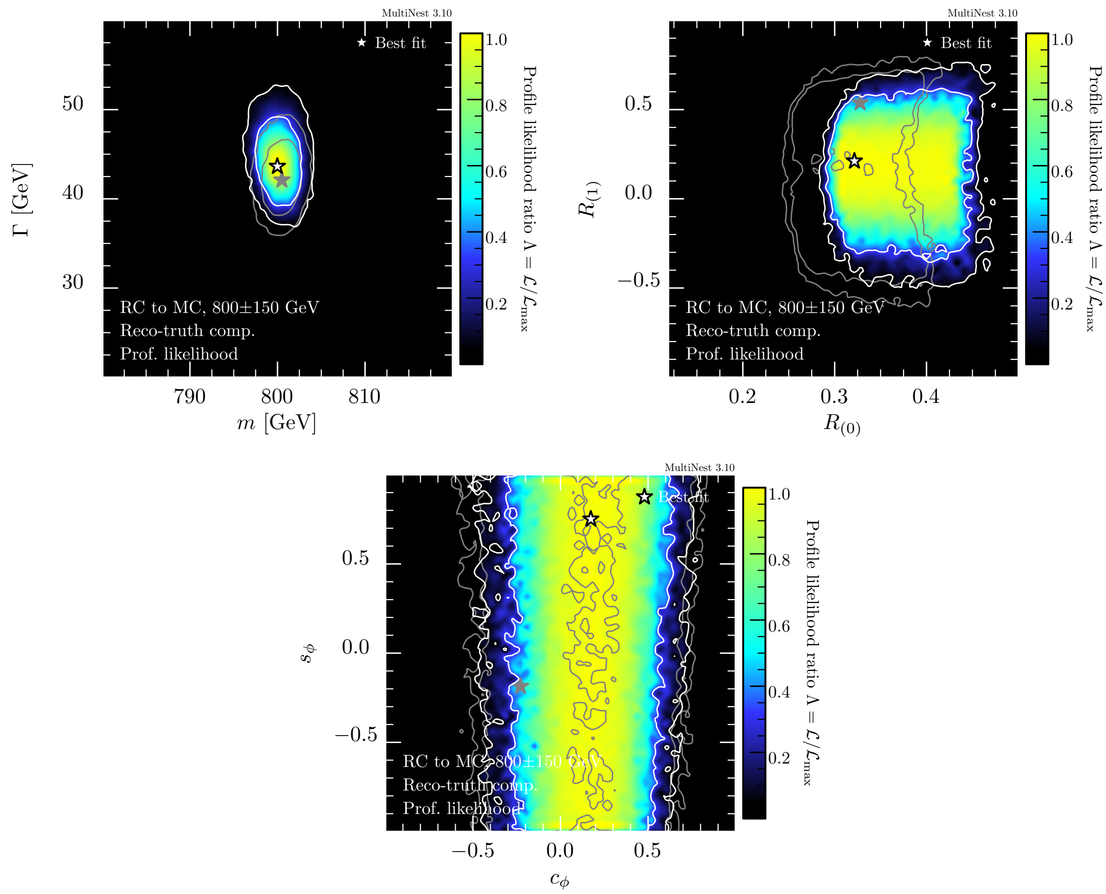

Figure 5.8: Two dimensional profile likelihood plots for the parameters of our general functional form, extracted from a global fit to reconstructed events over a 650–950 GeV window. The grey contours show 1 and 2σ boundaries for a fit to the truth-level distribution.

the 2σ contour. However, the reconstructed fit is unable to resolve this feature, and instead finds a single high likelihood region in the $R_{(0)}$-$R_{(1)}$ and c_ϕ-s_ϕ planes. Nevertheless, there is a large overlap between the truth and reconstructed results at the 2σ level.

An option previously suggested for reducing the distortions in reconstructed fit results was to shift the fit window to higher invariant masses. This was tested by performing another fit using a 700–1000 GeV window. The results are presented in Fig. 5.9. With the shift to higher invariant masses, the reconstructed fit is now able to distinguish two regions of high likelihood in the $R_{(0)}$-$R_{(1)}$ and c_ϕ-s_ϕ profile likelihood plots. The contours we find in the c_ϕ-s_ϕ plane, in particular, closely reflects the truth-level contours. The $R_{(1)}$ values between results also coincide closely, and while there is a slight preference for larger $R_{(0)}$ values in the reconstructed fit, the 1σ intervals largely overlap.

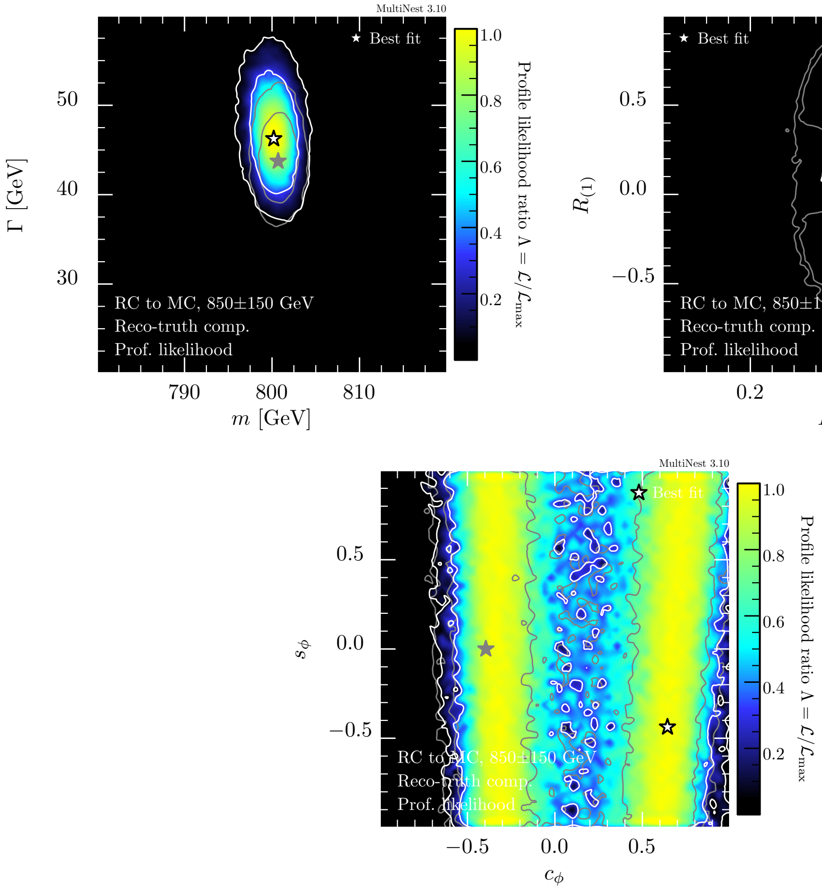

Figure 5.9: Two dimensional profile likelihood plots for the parameters of our general functional form, extracted from a global fit to reconstructed events over a 700–1000 GeV window. The grey contours show 1 and 2σ boundaries for a fit to the truth-level distribution.

5.5 Summary

In addition to underlying physical processes, experimental datasets will experience a distortion due to limitations of the detectors used to observe relevant decay events. This results in a smearing of the invariant mass distributions, which can be described as the convolution between a truth-level lineshape and an *ad hoc* detector resolution function, commonly parametrized with a DSCB. In this chapter, we verified the validity of such a treatment using our MC samples generated with MadGraph5_aMC@NLO and the Delphes 3.4.1 detector simulation. In general, the convolution was found to describe our reconstructed samples well; however, there were distortions in the low $m_{\gamma\gamma}$ regions of the data that were not predicted. These effects arise from the imposition of a p_T cut on reconstructed photons, introduced to mimic LHC detector trigger requirements. Thus, there will necessarily be a lower bound on the invariant mass in a resonance search, if detector effects are to be characterized as a simple convolution.

Additionally, the kinematic selection criteria introduced a small discrepancy to the overall normalizations of the reconstructed signal, interference and background invariant mass distributions. This was dependent on the number of events rejected from each component by the selection criteria, and hence inherently model-dependent. However, the discrepancy was found to be negligible when statistics are limited, as is typically the case for more realistic datasets. The approximation of equal normalizations was tested using a convolved form of our template functional form against a MC sample generated with a more realistic number of events. Results were compared with those from a fit performed directly on the truth-level distribution. Similar contours were observed in their profile likelihood plots, supporting our approximation. A closer agreement was found when the fit window was slightly shifted to a higher invariant mass range, indicating that normalization differences can indeed be better ignored as uncertainties in the data grow, and constraints on the reconstructed lineshape loosen.

Chapter 6

Conclusion

The top-down approach employed in the presentation of LHC resonance search results, while constituting a straightforward procedure in the scenario where there is strong motivation to believe in a particular resonance model *a priori* (for example, when seeking to better constrain the Higgs or other results predicted by the SM), becomes less ideal for reporting the results of a blind search for BSM physics. In the latter scenario, it is desirable for the reported results to be sufficiently generalizable to the plethora of existing BSM models, such that a given theorist is able to relate them to their pet model of interest.

In this thesis, we presented a general, model-independent functional form for this purpose:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Big[f_{\mathrm{BW}}\Big\{&m^4 R_{(0)}^2 + 2R_{(0)}m^3\Gamma s_\phi \\ &+2(q^2-m^2)\left(m^2 R_{(0)}R_{(1)} + m^2 R_{(0)}c_\phi + R_{(1)}m\Gamma s_\phi\right) \\ &+2(q^2-m^2)^2\left(R_{(1)}c_\phi\right)\Big\} \\ +1\Big]&, \qquad (3.58 \text{ revisited})\end{aligned}$$

which was derived using Quantum Field Theory arguments, and includes a general characterization of possible interference patterns arising between the signal and relevant SM backgrounds. The key motivation behind such a functional form is to introduce a bottom-up approach for analysis of experimental data: instead of selecting a benchmark BSM model and extracting a dataset to maximize the predicted signal, we propose a procedure that minimizes any model-dependent manipulation of the data, with results to be reported in terms of the parameters of the general functional form, which simply represents the lineshape of the data distribution observed. A theorist could then compute the same set of parameters for their particular model, and verify their result against those reported experimentally.

The procedure that we propose for an analysis is:

1. Perform a scan over the m-Γ plane of the general functional form, and produce a plot of the discovery significance. The remaining parameters of the functional form are to be profiled over, to yield the maximum significance at each point of the scan.

2. If any highly significant regions exist, repeat the fits of the general parameters to the data over these regions, using invariant mass windows of increasing size. The purpose of this is to identify an appropriate fit window to the data, since it is understood that the general parametrization is only a good approximation in the neighbourhood of the prospective signal peak.

3. Extract the parameters of the general functional form using the largest possible (or sufficiently large) fit window that still yields a good fit to the data. These results are in the form of 2-dimensional profile likelihood plots over the general parameter space, to appropriately visualize their non-trivial correlations.

In this, we have assumed that only a single resonance contributes to the invariant mass range considered; if there are multiple prospective signals, they must contribute to sufficiently separated regions of the data, such that independent analyses can be conducted for each. The general parametrization described also assumes that the background is produced in the same partonic channel as the signal (and hence generates an interference pattern). In the case of additional backgrounds yielding the same final state but induced by a different partonic process from the signal (an example being the $\gamma\gamma$ final state, which can be produced by gg, qg, or $q\bar{q}$ fusion in pp collisions), we assume that a suitable procedure for subtracting such backgrounds from the data exists. For this, we rely on the correctness of our understanding of SM physics, and thus of the non-interfering background contributions predicted by theoretical simulations.

In a practical application of the procedure to experimental datasets, one would perform the fits using a convolution of the general functional form with a detector resolution function:

$$\left[\frac{d\sigma}{dq}\right]_{\text{reco}} = \frac{\varepsilon_s N_{\text{DSCB}}}{\varepsilon_b}\left[(F_{\mathcal{B}}^{\text{reco}} \times S) * \text{DSCB}\right] + \frac{\varepsilon_i N_{\text{DSCB}}}{\varepsilon_b}\left[(F_{\mathcal{B}}^{\text{reco}} \times I) * \text{DSCB}\right] + F_{\mathcal{B}}^{\text{reco}}\,. \tag{5.26 revisited}$$

where S and I respectively denote the signal and interference terms in the functional form, and the description (and normalization) of the detector resolution function, commonly parametrized as a Double-Sided Crystal Ball, is to be found prior to the fit through stringent Monte Carlo simulations of the physical detector. The ε are selection efficiencies of the various components; however, the approximation of equal efficiencies holds quite well, as their discrepancies will generally be dwarfed by other sources of experimental uncertainty. This convolution description of detector smearing effects on the invariant mass distribution is in line with the methods currently employed by the LHC detector collaborations.

Appendices

Appendix A

Taylor expansion of $\phi_{\mathcal{SB}}$

Here, we assume the case of a single helicity description to retain only one relative phase in the general functional form. A Taylor expansion of $\phi_{\mathcal{SB}}$ in q^2 can then be written as:

$$\phi_{\mathcal{SB}}(q^2) = \phi_{\mathcal{SB}}^{(0)} + \phi_{\mathcal{SB}}^{(1)} \frac{(q^2 - m^2)}{m^2} + \mathcal{O}\left((q^2 - m^2)^2\right), \tag{A.1}$$

where

$$\phi_{\mathcal{SB}}^{(0)} = \phi_{\mathcal{SB}}(m^2), \qquad \phi_{\mathcal{SB}}^{(1)} = \left.\frac{\partial \phi_{\mathcal{SB}}(q^2)}{\partial(q^2/m^2)}\right|_{q^2=m^2}. \tag{A.2}$$

Substitution into the general functional form yields:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2) \Big[& f_{\mathrm{BW}} \Big(m^4 R_{(0)}^2 + 2R_{(0)} m^3 \Gamma \sin \phi_{\mathcal{SB}}^{(0)} \\ & + 2\left(q^2 - m^2\right)^2 \left(R_{(1)} \cos \phi_{\mathcal{SB}}^{(0)} - R_{(0)} \phi_{\mathcal{SB}}^{(1)} \sin \phi_{\mathcal{SB}}^{(0)} \right)\Big) \\ & + 2 f_{\mathrm{BW}} (q^2 - m^2) \Big(m^2 R_{(0)} \cos \phi_{\mathcal{SB}}^{(0)} + m^2 R_{(0)} R_{(1)} \\ & + R_{(1)} m\Gamma \sin \phi_{\mathcal{SB}}^{(0)} + R_{(0)} \phi_{\mathcal{SB}}^{(1)} m\Gamma \cos \phi_{\mathcal{SB}}^{(0)} \Big) \\ & + 1 + \mathcal{O}\left((q^2 - m^2)^2\right) \Big]. \end{aligned} \tag{A.3}$$

However, in the form of Eq. (A.3), one loses the invariance of the phase under the transformations $\phi_{\mathcal{SB}} \to \phi_{\mathcal{SB}} + 2\pi k$, $k \in \mathbb{Z}$, owing to the linear dependence on $\phi_{\mathcal{SB}}^{(1)}$. To avoid this issue, one could expand in terms of the cosine of the phase,

$$\cos \phi_{\mathcal{SB}} = c_{\phi(0)} + c_{\phi(1)} \frac{(q^2 - m^2)}{m^2} + \mathcal{O}\left((q^2 - m^2)^2\right), \tag{A.4}$$

with

$$\sin \phi_{\mathcal{SB}} = \lambda \sqrt{1 - \cos^2 \phi_{\mathcal{SB}}}, \qquad \lambda = \pm 1, \tag{A.5}$$

where

$$\sqrt{1 - \cos^2 \phi_{\mathcal{SB}}} = \sqrt{1 - c_{\phi(0)}^2} - \frac{c_{\phi(0)} c_{\phi(1)} (q^2 - m^2)}{m^2 \sqrt{1 - c_{\phi(0)}^2}}, \tag{A.6}$$

or alternatively, designate $\sin \phi_{\mathcal{SB}}$ as the parameter of interest:

$$\sin \phi_{\mathcal{SB}} = s_{\phi(0)} + s_{\phi(1)} \frac{(q^2 - m^2)}{m^2} + \mathcal{O}\left((q^2 - m^2)^2\right), \tag{A.7}$$

$$\cos \phi_{\mathcal{SB}} = \lambda \sqrt{1 - \sin^2 \phi_{\mathcal{SB}}}, \qquad \lambda = \pm 1, \tag{A.8}$$

$$\sqrt{1 - \sin^2 \phi_{\mathcal{SB}}} = \sqrt{1 - s_{\phi(0)}^2} - \frac{s_{\phi(0)} s_{\phi(1)} (q^2 - m^2)}{m^2 \sqrt{1 - s_{\phi(0)}^2}}. \tag{A.9}$$

These expansions yield a functional form:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Bigg[f_{\rm BW}\Big(& m^4 R_{(0)}^2 + 2R_{(0)} m^3 \Gamma \lambda \sqrt{1-c_{\phi(0)}^2} \\ & +2\left(q^2-m^2\right)^2 \left(R_{(1)} c_{\phi(0)} + R_{(0)} c_{\phi(1)}\right)\Big) \\ & +2 f_{\rm BW}(q^2-m^2)\Bigg(m^2 R_{(0)} c_{\phi(0)} + m^2 R_{(0)} R_{(1)} \\ & +R_{(1)} m\Gamma\lambda\sqrt{1-c_{\phi(0)}^2} - R_{(0)} m\Gamma\lambda \frac{c_{\phi(0)} c_{\phi(1)}}{\sqrt{1-c_{\phi(0)}^2}}\Bigg) \\ & + 1 + \mathcal{O}\left((q^2-m^2)^2\right)\Bigg],\end{aligned} \tag{A.10}$$

or equivalently, in terms of the sine:

$$\begin{aligned}\frac{d\sigma}{dq} = F_{\mathcal{B}}(q^2)\Bigg[f_{\rm BW}\Bigg(& m^4 R_{(0)}^2 + 2R_{(0)} m^3 \Gamma s_{\phi(0)} \\ & +2\left(q^2-m^2\right)^2 \left(R_{(1)}\lambda\sqrt{1-s_{\phi(0)}^2} - R_{(0)}\lambda\frac{s_{\phi(0)} s_{\phi(1)}}{\sqrt{1-s_{\phi(0)}^2}}\right)\Bigg) \\ & +2 f_{\rm BW}(q^2-m^2)\Big(m^2 R_{(0)}\lambda\sqrt{1-s_{\phi(0)}^2} + m^2 R_{(0)} R_{(1)} \\ & + R_{(1)} m\Gamma s_{\phi(0)} + R_{(0)} m\Gamma s_{\phi(1)}\Big) \\ & + 1 + \mathcal{O}\left((q^2-m^2)^2\right)\Bigg].\end{aligned} \tag{A.11}$$

Appendix B

Toy $a_{i,j}$ fit results, various a

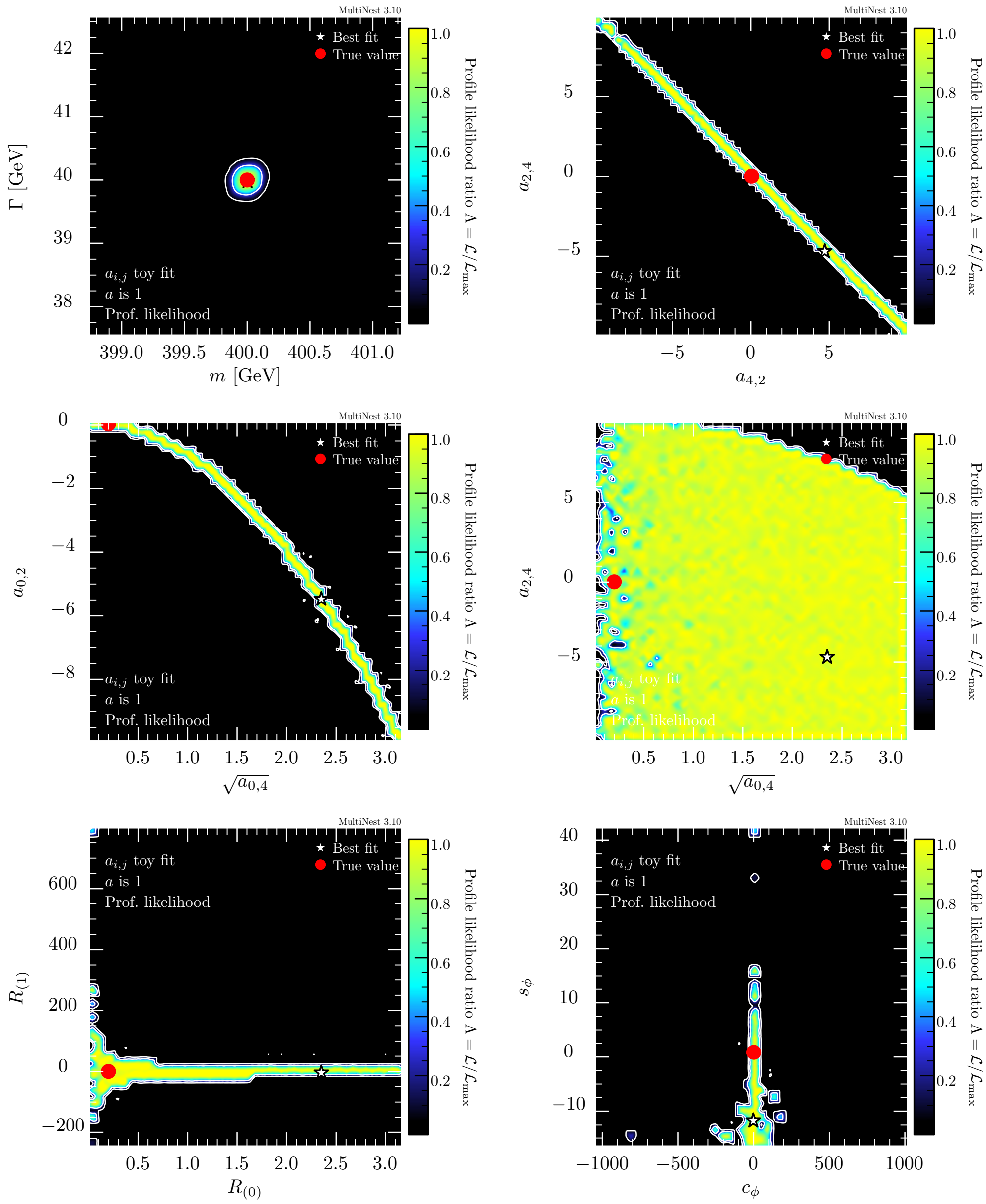

Figure B.1: Profile likelihood contours of the $a_{i,j}$ parameters fitte to an Asimov toy constructed from an input point in its space, with $a = 1$. The bottom two plots show the result in terms of the RC parameters.

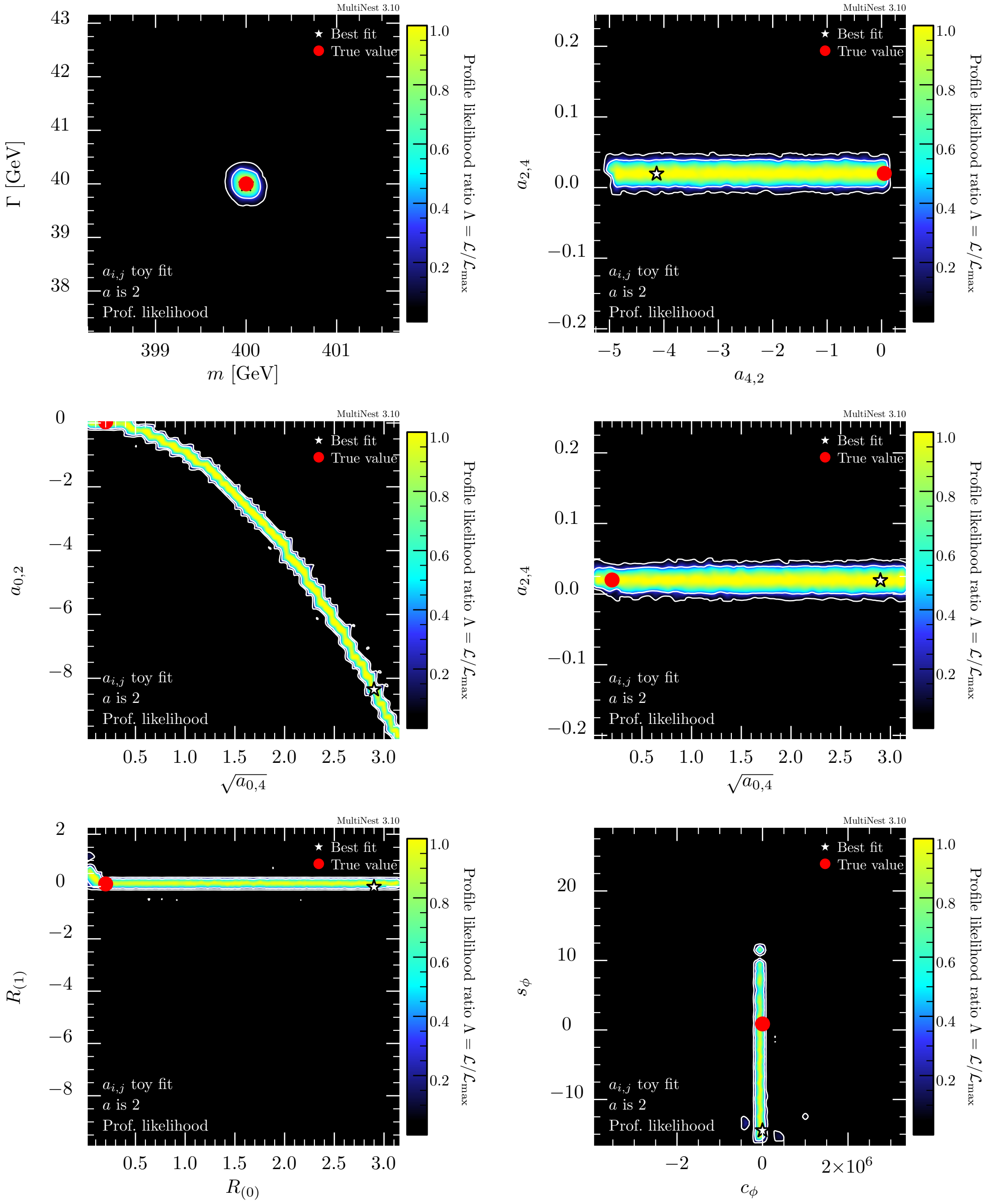

Figure B.2: Profile likelihood contours of the $a_{i,j}$ parameters fitted to an Asimov toy constructed from an input point in its space, with $a = 2$. The bottom two plots show the result in terms of the RC parameters.

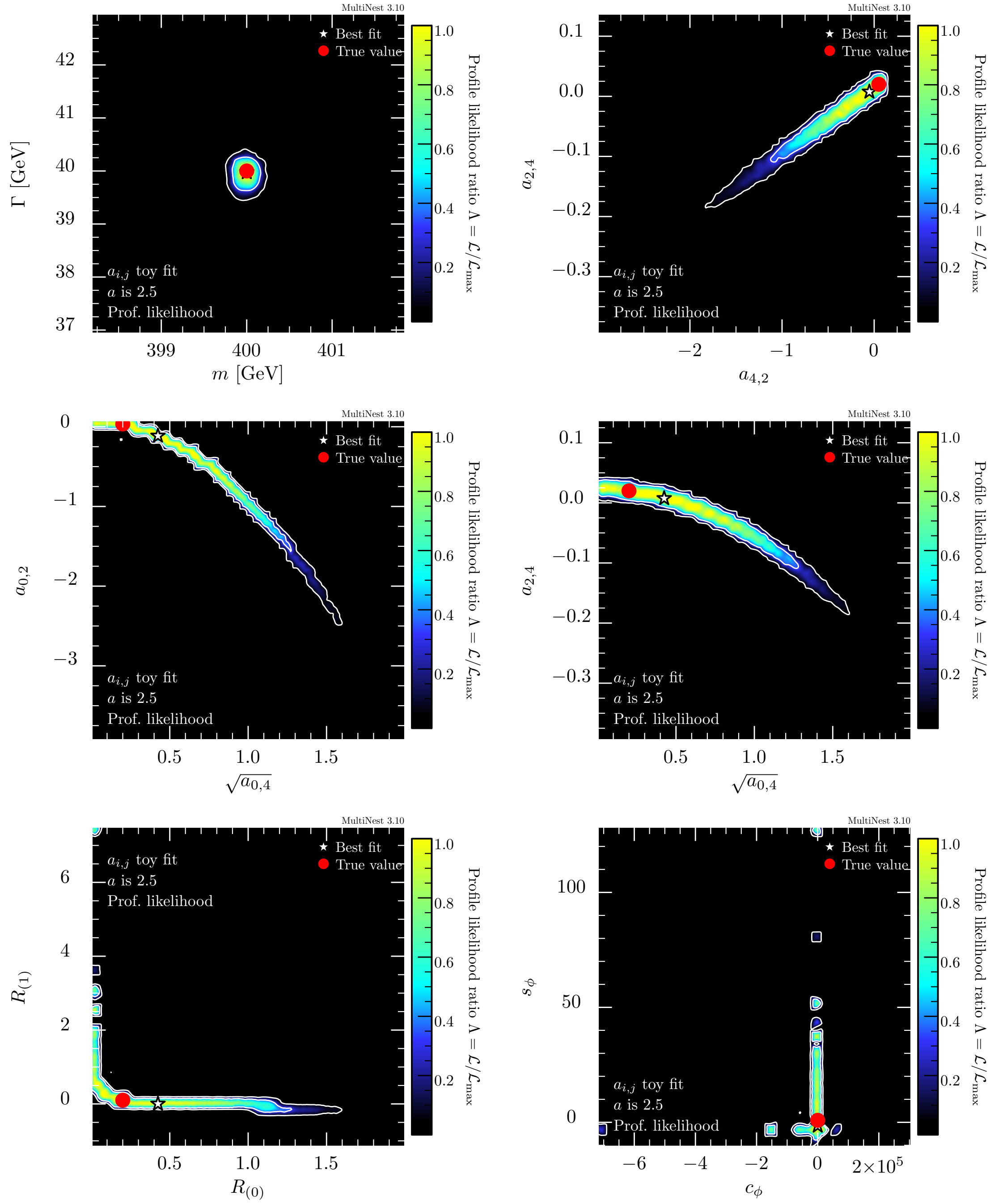

Figure B.3: Profile likelihood contours of the $a_{i,j}$ parameters fitted to an Asimov toy constructed from an input point in its space, with $a = 2.5$. The bottom two plots show the result in terms of the RC parameters.

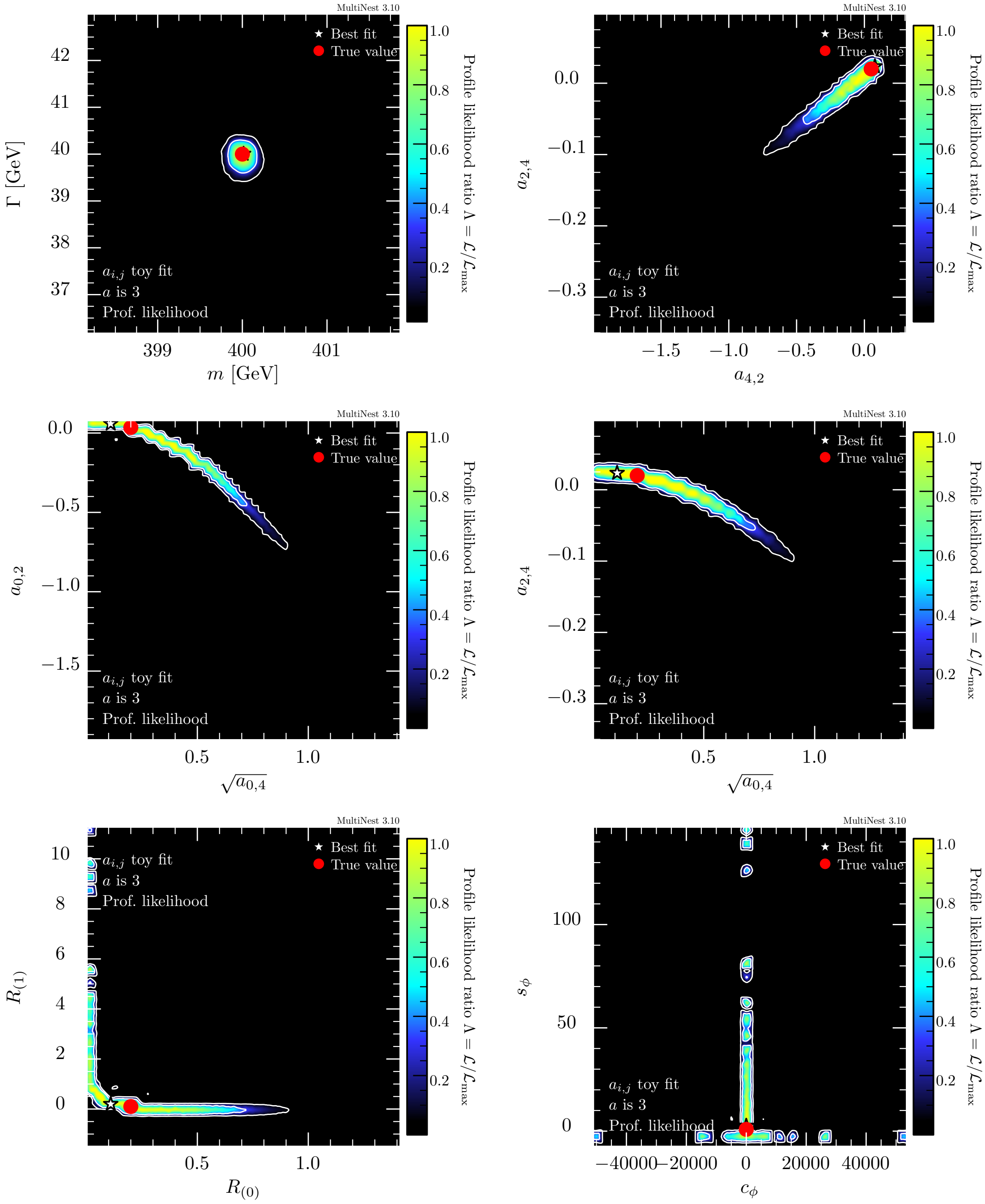

Figure B.4: Profile likelihood contours of the $a_{i,j}$ parameters fitted to an Asimov toy constructed from an input point in its space, with $a = 3$. The bottom two plots show the result in terms of the RC parameters.

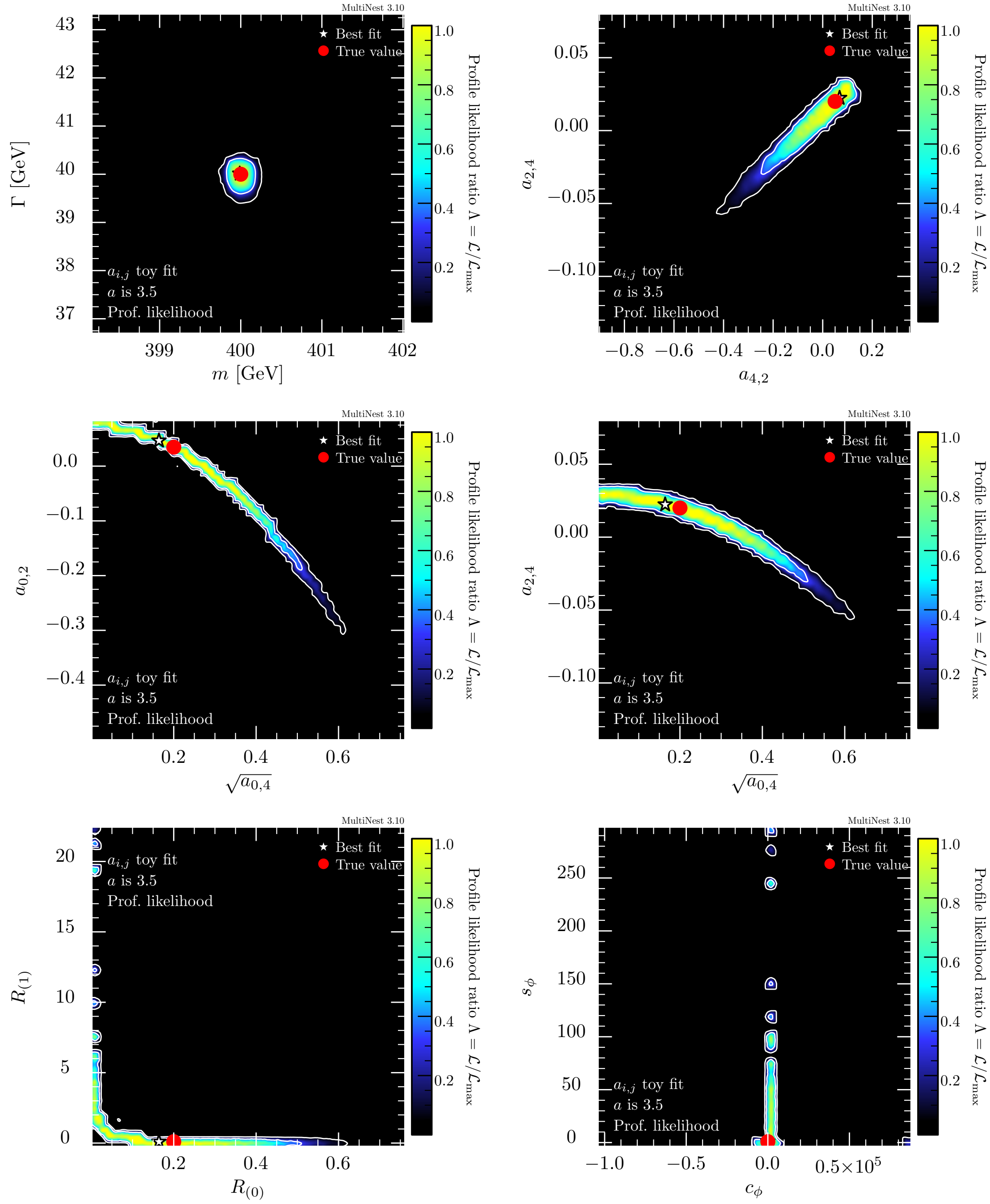

Figure B.5: Profile likelihood contours of the $a_{i,j}$ parameters fitted to an Asimov toy constructed from an input point in its space, with $a = 3.5$. The bottom two plots show the result in terms of the RC parameters.

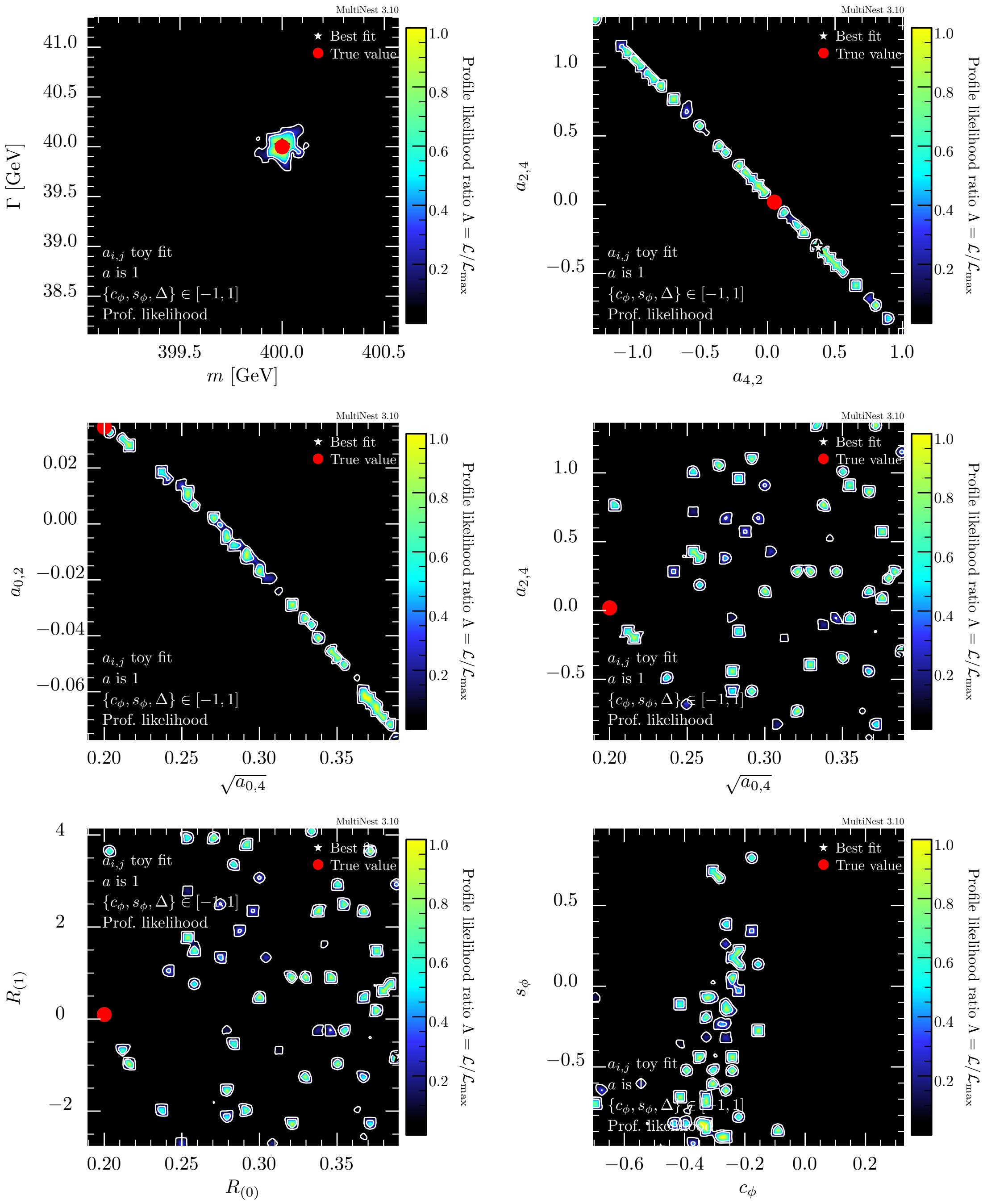

Figure B.6: The $a_{i,j}$ fit result for $a = 1$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded.

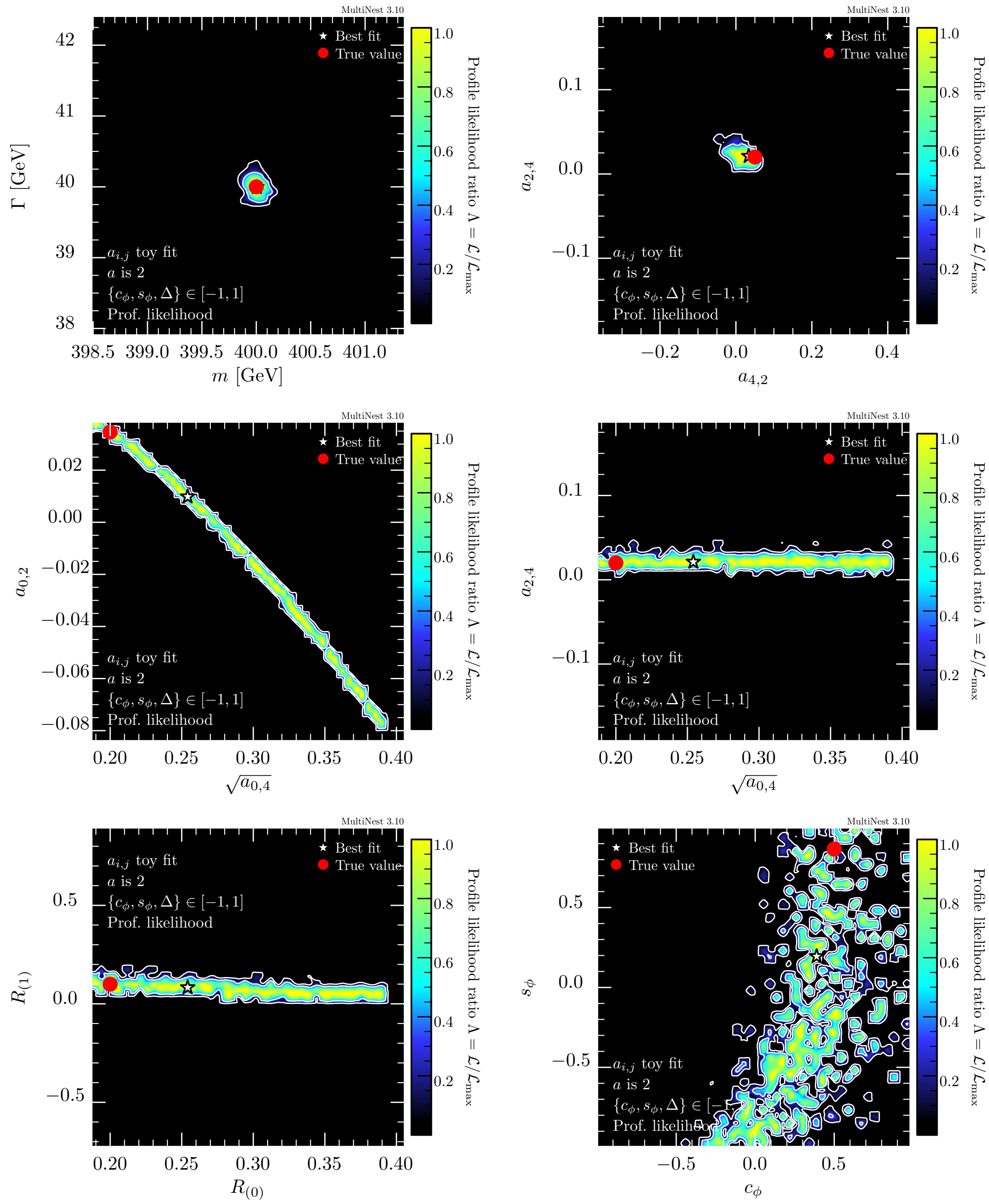

Figure B.7: The $a_{i,j}$ fit result for $a = 2$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded.

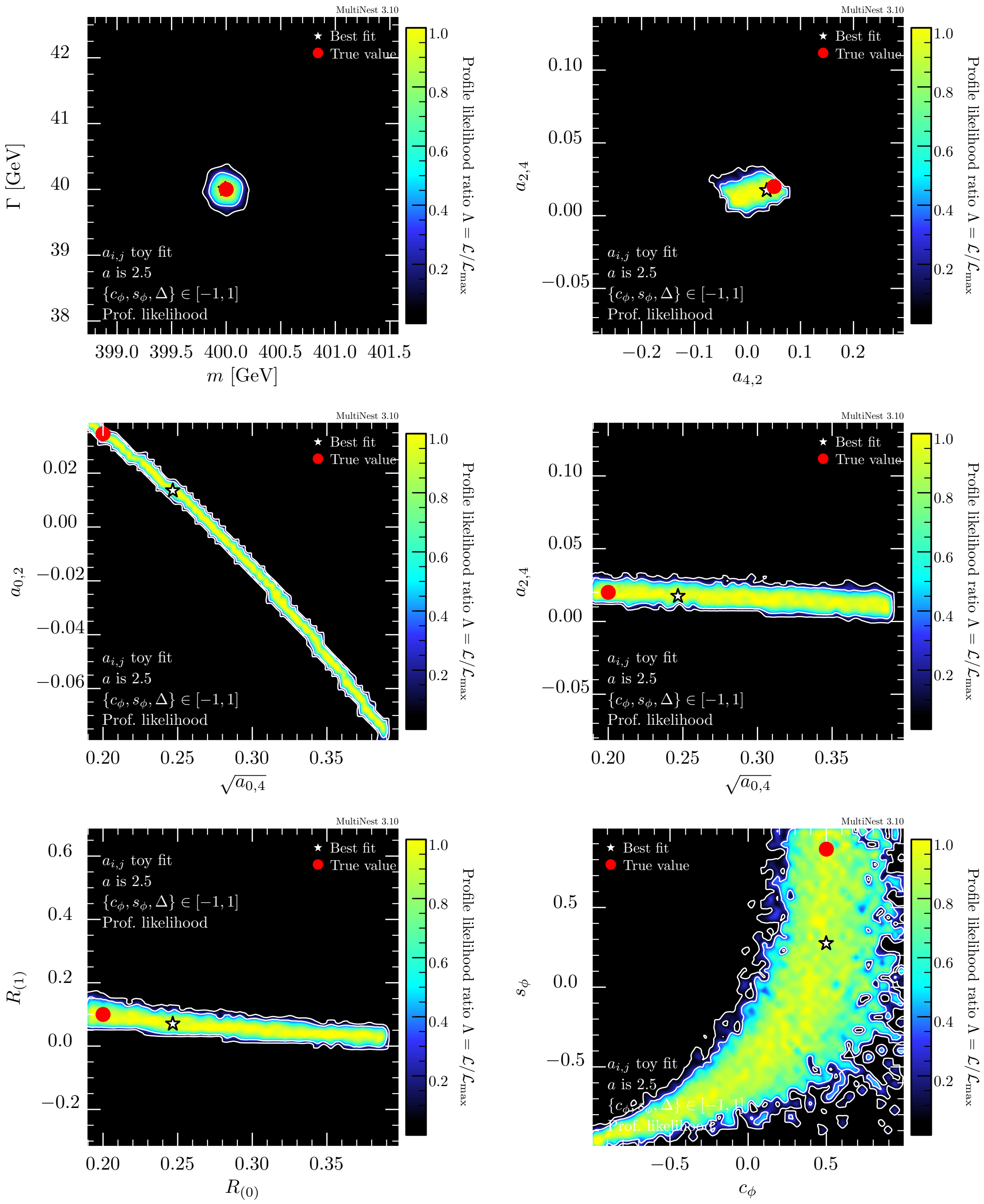

Figure B.8: The $a_{i,j}$ fit result for $a = 2.5$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded.

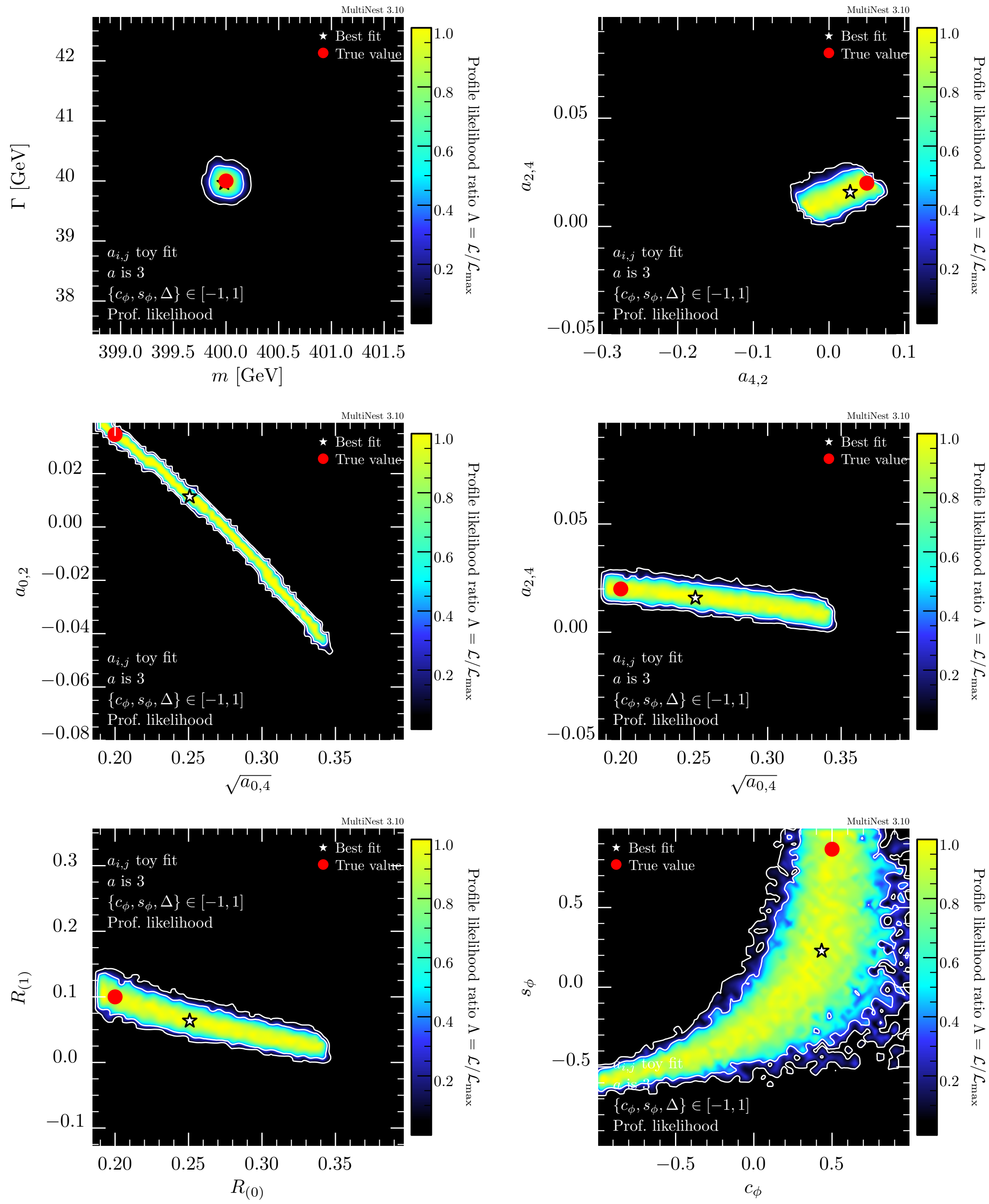

Figure B.9: The $a_{i,j}$ fit result for $a = 3$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded.

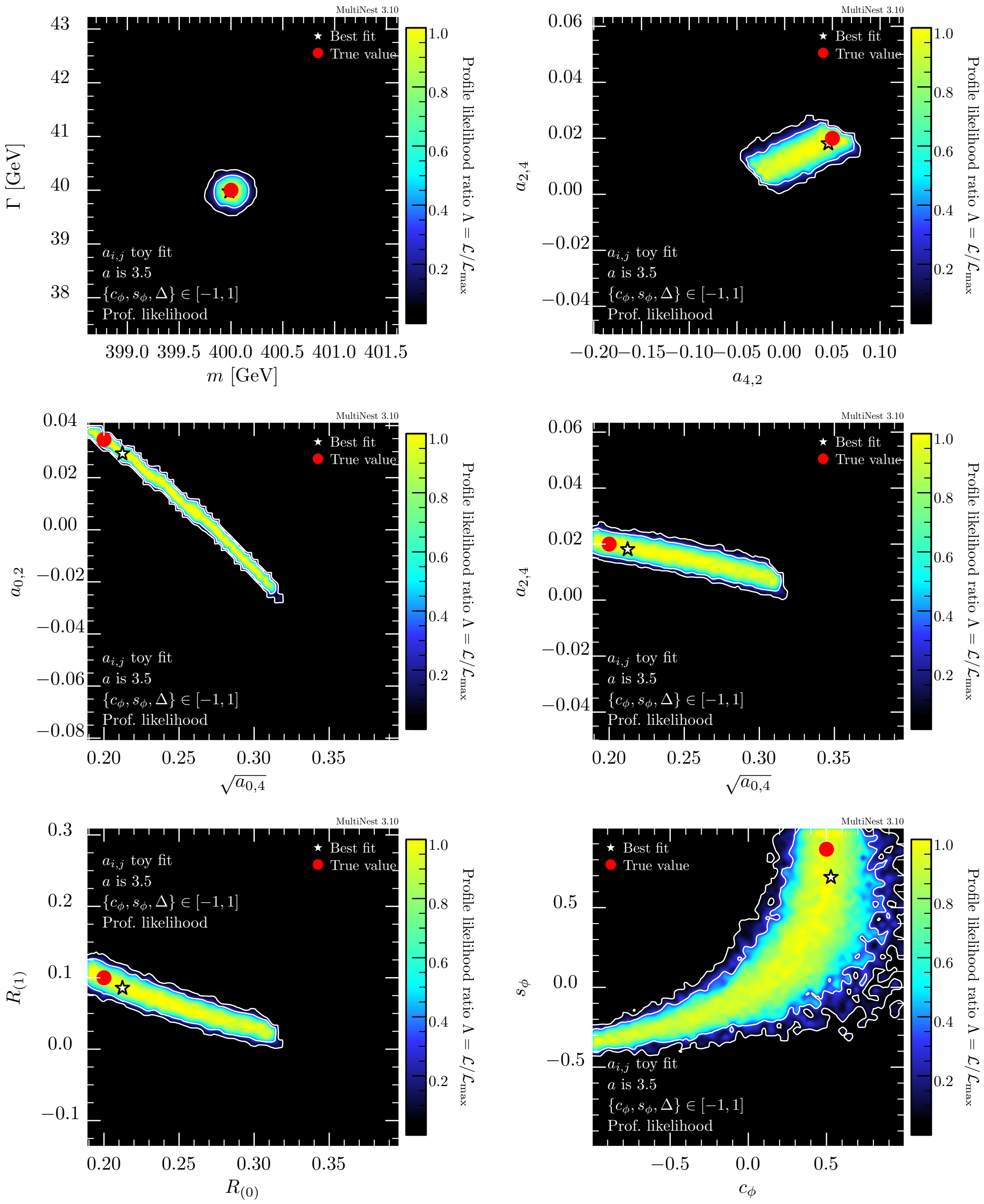

Figure B.10: The $a_{i,j}$ fit result for $a = 3.5$, after all c_ϕ, s_ϕ, and Δ points outside of $[-1, 1]$ have been discarded.

Appendix C

Fit window tests

C.1 Asimov toys of input distributions

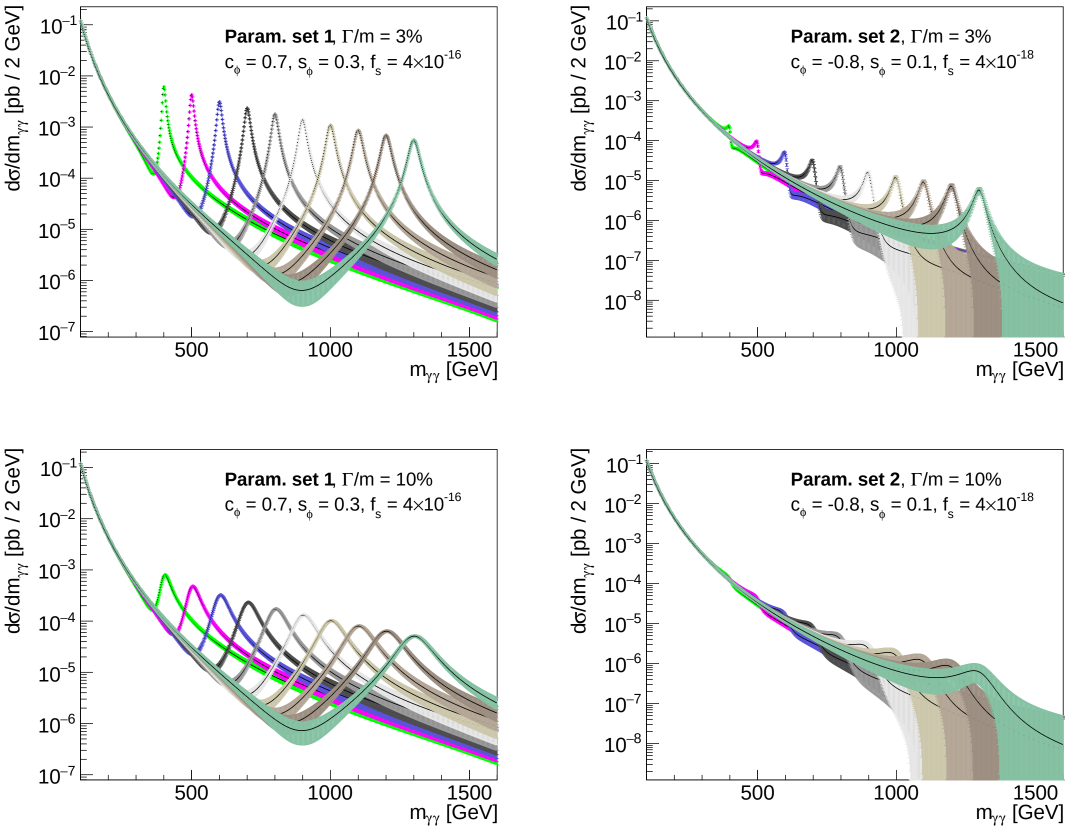

Figure C.1: Visualization of Asimov datasets, generated from the PM functional form at various mass points, with $\Gamma/m = 3\%$ and 10%, for the two sets of input parameters.

C.2 Results for $\Gamma/m = 3\%$ and 10%

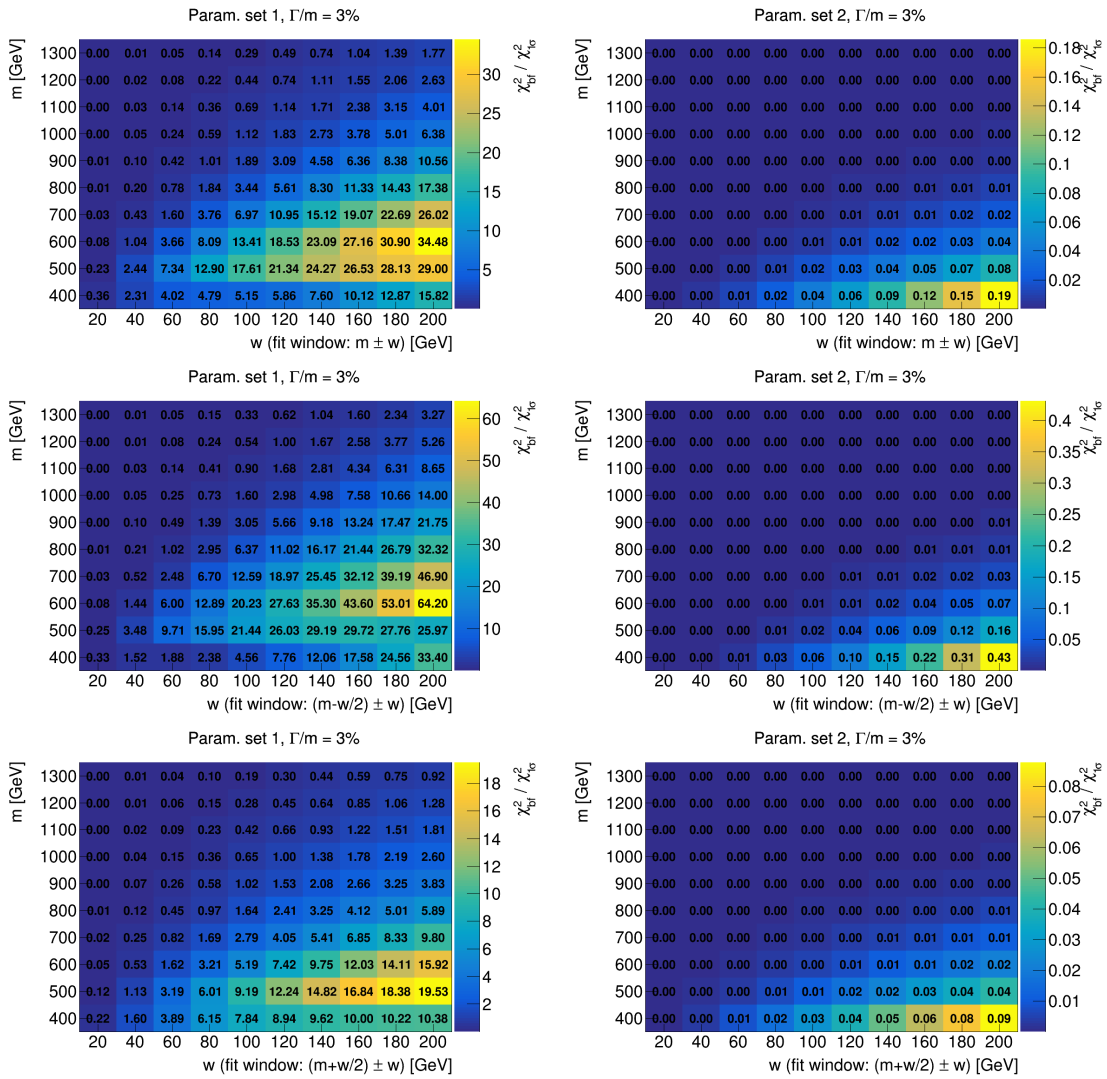

Figure C.2: Colour plots showing the quality of fits relative to the 1σ threshold of a χ^2_ν distribution. Results correspond to input widths of $\Gamma/m = 3\%$. Left (right): results using set 1 (2) of the input PM parameters. From top to bottom, results correspond to a fit window: centred about the input mass; shifted to the left by $w/2$; shifted to the right by $w/2$.

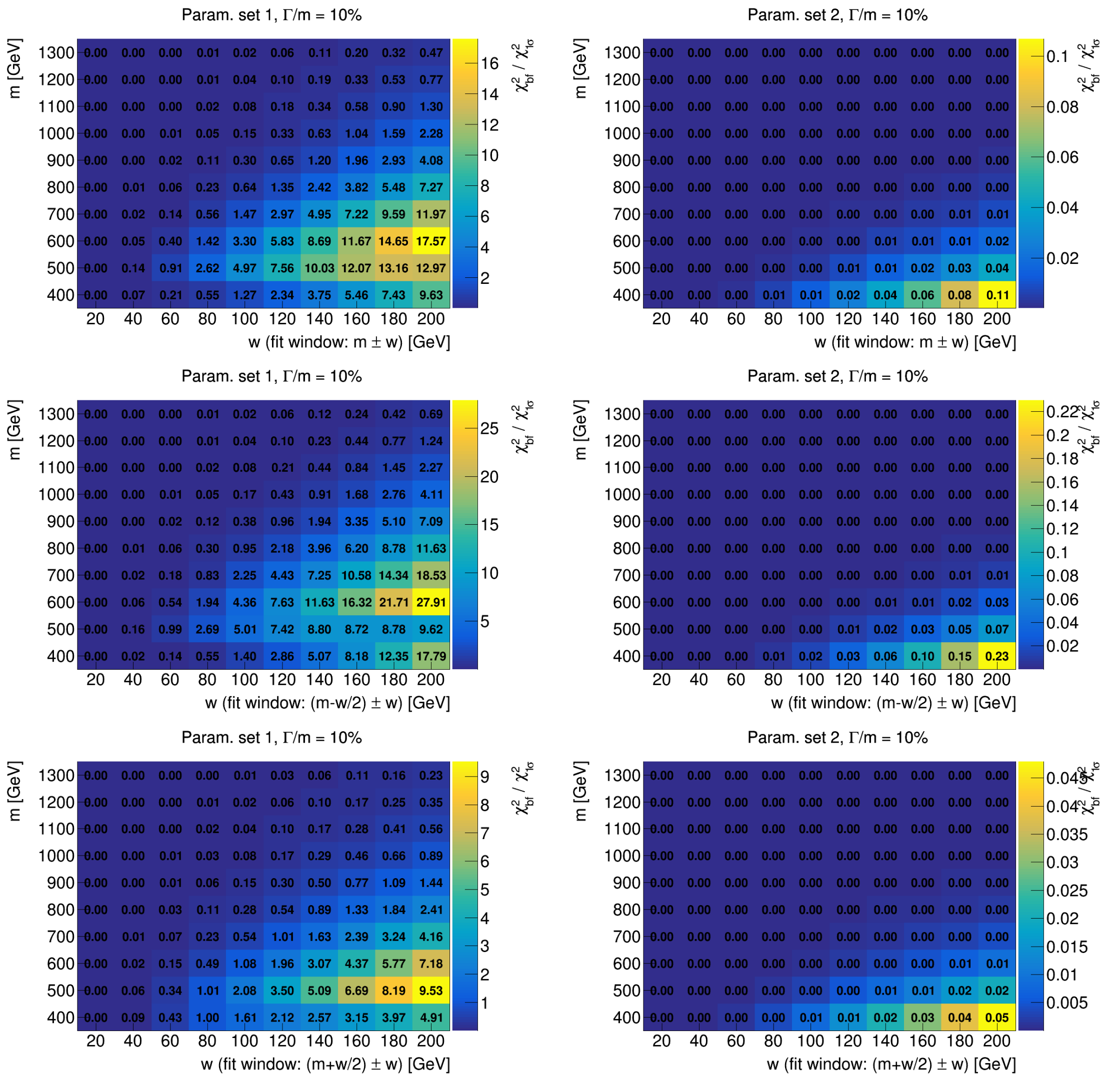

Figure C.3: Colour plots showing the quality of fits relative to the 1σ threshold of a χ^2_ν distribution. Results correspond to input widths of $\Gamma/m = 10\%$. Left (right): results using set 1 (2) of the input PM parameters. From top to bottom, results correspond to a fit window: centred about the input mass; shifted to the left by $w/2$; shifted to the right by $w/2$.

C.3 PM fit to PM toy results comparison

The result of fitting the PM to an Asimov generated using a point in its space as input was used to construct two histograms with their bin uncertainties defined differently: "method 1" assumes a 1-dimensional Gaussian random variable for each bin, while "method 2" assumes a 5-dimensional χ^2 distribution (corresponding to the five PM parameters). The fit of the PM to these two histograms is compared to the original fit to the Asimov.

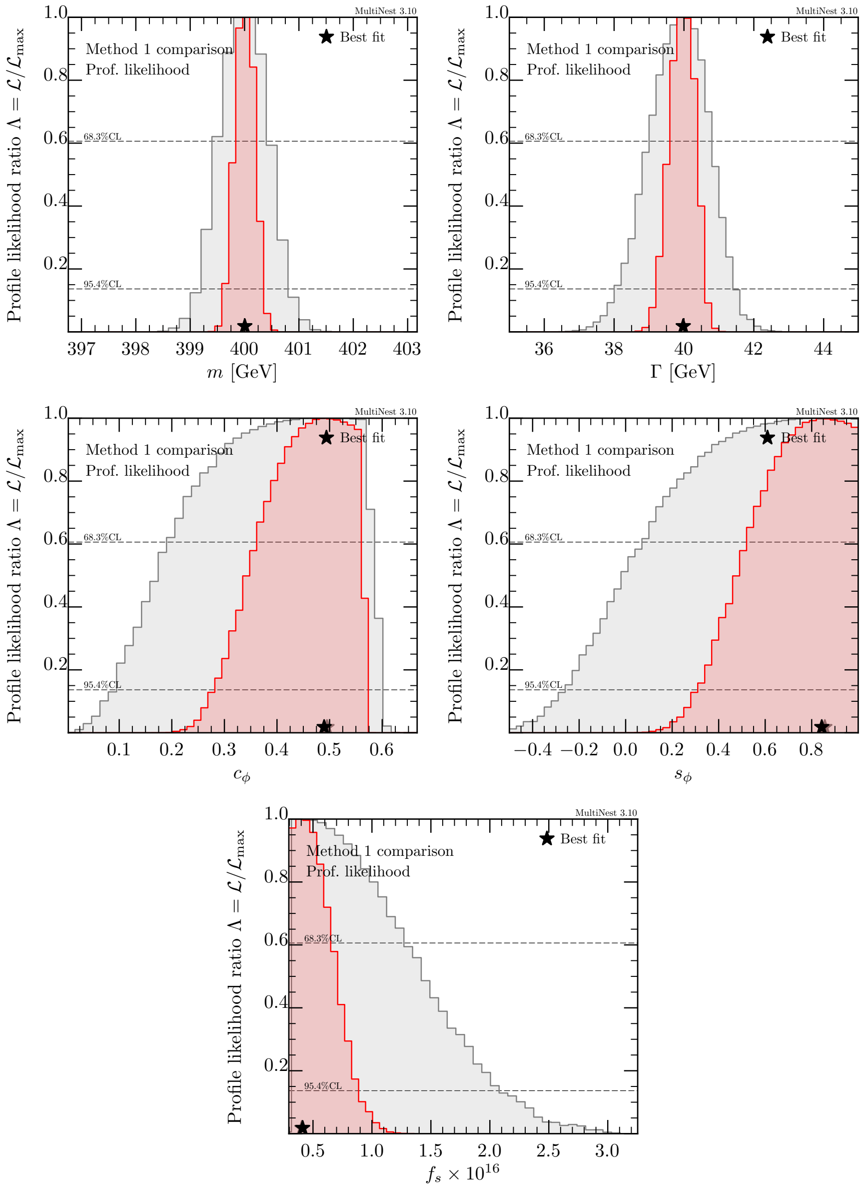

Figure C.4: Grey: fit to toy PM Asimov. Red: fit to results histogram constructed assuming "method 1". The fit to the results histogram finds more constrained confidence intervals for the parameters, suggesting an underestimate of bin uncertainties.

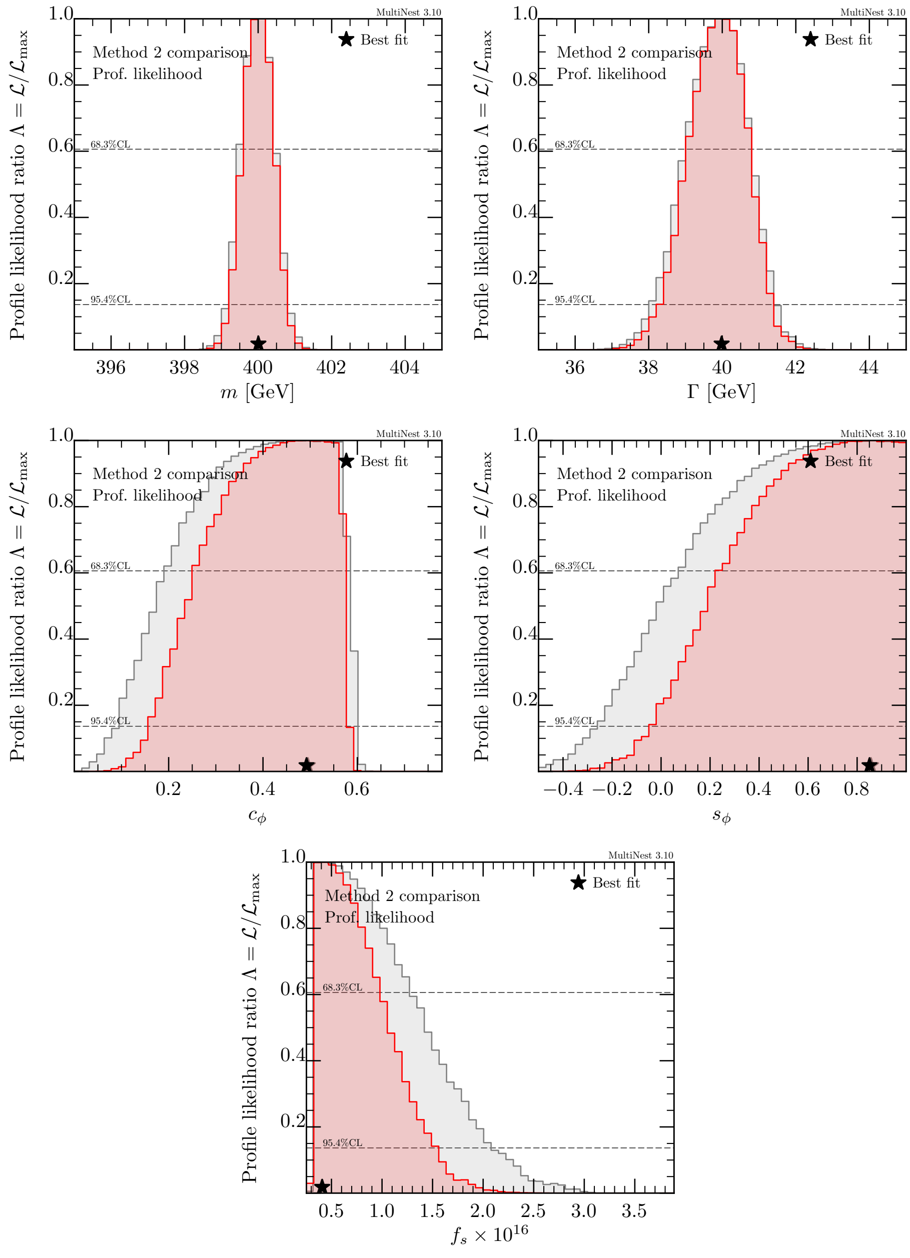

Figure C.5: Grey: fit to toy PM Asimov. Red: fit to results histogram constructed assuming "method 2". A better agreement between the two sets of results is found.

Bibliography

[1] ATLAS collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) 1 [1207.7214].

[2] CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30 [1207.7235].

[3] ATLAS collaboration, *Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Physics Letters B **775** (2017) 105 [1707.04147].

[4] ATLAS collaboration, *Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector*, Journal of High Energy Physics **09** (2016) 001 [1606.03833].

[5] CMS collaboration, *Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s} = 8$ and 13 TeV*, Physical Review Letters **117** (2016) 051802 [1606.04093].

[6] A. Djouadi, J. Ellis and J. Quevillon, *Interference effects in the decays of spin-zero resonances into $\gamma\gamma$ and $t\bar{t}$*, Journal of High Energy Physics **2016** (2016) 105 [1605.00542].

[7] A. Djouadi, J. Ellis, A. Popov and J. Quevillon, *Interference effects in $t\bar{t}$ production at the LHC as a window on new physics*, Journal of High Energy Physics **03** (2019) 119 [1901.03417].

[8] G. 't Hooft, *Why do We Need Local Gauge Invariance in Theories with Vector Particles? An Introduction*, NATO Advanced Study Institutes Series B **59** (1980) 101.

[9] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Physical Review Letters **32** (1974) 438.

[10] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356v7.

[11] ATLAS collaboration, *Search for supersymmetry in events with four or more leptons in $\sqrt{s} = 13$ TeV pp collisions with ATLAS*, Physical Review D **98** (2018) 032009 [1804.03602].

[12] CMS collaboration, *Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV*, Physical Review D **96** (2017) 032003 [1704.07781].

[13] CMS collaboration, *Search for supersymmetry in the all-hadronic final state using top quark tagging in pp collisions at $\sqrt{s} = 13$ TeV*, Physical Review D **96** (2017) 012004 [1701.01954].

[14] A. Cherchiglia, D. Stöckinger and H. Stöckinger-Kim, *Muon g-2 in the 2HDM: maximum results and detailed phenomenology*, *Physical Review D* **98** (2018) 035001 [1711.11567].

[15] A. Arhrib, R. Benbrik, S. J. D. King, B. Manaut, S. Moretti and C. S. Un, *Phenomenology of 2HDM with vectorlike quarks*, *Physical Review D* **97** (2018) 095015 [1607.08517].

[16] L. Evans and P. Bryant, *LHC Machine*, *Journal of Instrumentation* **3** (2008) S08001.

[17] ATLAS collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) S08003.

[18] CMS collaboration, *The CMS Experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08004.

[19] ALICE collaboration, *The ALICE experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08002.

[20] LHCb collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008) S08005.

[21] LHC New Physics Working Group collaboration, *Simplified Models for LHC New Physics Searches*, *Journal of Physics G: Nuclear and Particle Physics* **39** (2012) 105005 [1105.2838].

[22] Y. C. Yap, *Search for New Physics with Two Photons in the Final State with the ATLAS Detector*, Ph.D. thesis, Paris Diderot University, 2017.

[23] V. Bertone, S. Carrazza and J. Rojo, *APFEL: A PDF evolution library with QED corrections*, *Computer Physics Communications* **185** (2014) 1647 [1310.1394].

[24] S. Carrazza, A. Ferrara, D. Palazzo and J. Rojo, *APFEL Web: a web-based application for the graphical visualization of parton distribution functions*, *Journal of Physics G: Nuclear and Particle Physics* **42** (2015) 057001 [1410.5456].

[25] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Physical Review D* **82** (2010) 074024 [1007.2241].

[26] P. Artoisenet, P. de Aquino, F. Demartin, R. Frederix, S. Frixione, F. Maltoni et al., *A framework for Higgs characterisation*, *Journal of High Energy Physics* **2013** (2013) 43 [1306.6464].

[27] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr and B. Fuks, *FeynRules 2.0 – A complete toolbox for tree-level phenomenology*, *Computer Physics Communications* **185** (2014) 2250 [1310.1921].

[28] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer and T. Reiter, *UFO – The Universal FeynRules Output*, *Computer Physics Communications* **183** (2012) 1201 [1108.2040].

[29] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *Journal of High Energy Physics* **2014** (2014) 79 [1405.0301].

[30] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, *Journal of High Energy Physics* **06** (2011) 128 [1106.0522].

[31] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *Journal of High Energy Physics* **06** (2002) 029 [hep-ph/0204244].

[32] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An Introduction to PYTHIA 8.2*, *Computer Physics Communications* **191** (2015) 159 [1410.3012].

[33] DELPHES 3 collaboration, *DELPHES 3: a modular framework for fast simulation of a generic collider experiment*, *Journal of High Energy Physics* **2014** (2014) 57 [1307.6346].

[34] I. Antcheva, M. Ballintijn, B. Bellenot, M. Biskup, R. Brun, N. Buncic et al., *ROOT – A C++ framework for petabyte data storage, statistical analysis and visualization*, *Computer Physics Communications* **180** (2009) 2499 [1508.07749].

[35] A. Vautrin, *Study of interference for the search of new particles decaying into two photons*, summer research report, Laboratoire de Physique Nucléaire et de Hautes Énergies and École Nationale Supérieure des Mines de Paris, 2016–2017.

[36] F. James and M. Winkler, MINUIT *User's Guide* (2004).

[37] F. James, *The Interpretation of Errors* (2004).

[38] W. Verkerke and D. Kirkby, *RooFit Users Manual v2.91* (2008).

[39] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *The European Physical Journal C* **71** (2011) 1554 [1007.1727].

[40] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, *Transactions of the American Mathematical Society* **54** (1943) 426.

[41] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Annals of Mathematical Statistics* **9** (1938) 60.

[42] L. Tierney, *Markov Chains for Exploring Posterior Distributions*, *The Annals of Statistics* **22** (1994) 1701.

[43] S. Chib and E. Greenberg, *Understanding the Metropolis-Hastings Algorithm*, *The American Statistician* **49** (1995) 327.

[44] S. Sharma, *Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy*, *Annual Review of Astronomy and Astrophysics* **55** (2017) 213 [1706.01629].

[45] R. Storn and K. Price, *Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces*, *Journal of Global Optimization* **11** (1997) 341.

[46] J. Brest, S. Greiner, B. Boskovic, M. Mernik and V. Zumer, *Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems*, *IEEE Transactions on Evolutionary Computation* **10** (2006) 646.

[47] J. Brest, B. Bošković, S. Greiner, V. Žumer and M. S. Maučec, *Performance comparison of self-adaptive and adaptive differential evolution algorithms*, *Soft Computing* **11** (2007) 617.

[48] GAMBIT collaboration, *Comparison of statistical sampling methods with ScannerBit, the GAMBIT scanning module*, *The European Physical Journal C* **77** (2017) 761 [1705.07959].

[49] GAMBIT collaboration, *A global fit of the MSSM with GAMBIT*, *The European Physical Journal C* **77** (2017) 879 [1705.07917].

[50] GAMBIT collaboration, *Status of the scalar singlet dark matter model*, *The European Physical Journal C* **77** (2017) 568 [1705.07931].

[51] F. Feroz, M. P. Hobson and M. Bridges, *MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics*, *Monthly Notices of the Royal Astronomical Society* **398** (2009) 1601 [0809.3437].

[52] J. Skilling, *Nested Sampling*, *AIP Conference Proceedings* **735** (2004) 395.

[53] J. Skilling, *Nested Sampling for General Bayesian Computation*, *Bayesian Analysis* **1** (2006) 833.

[54] F. Feroz and M. P. Hobson, *Multimodal nested sampling: an efficient and robust alternative to MCMC methods for astronomical data analysis*, *Monthly Notices of the Royal Astronomical Society* **384** (2008) 449 [0704.3704].

[55] R. Shaw, M. Bridges and M. P. Hobson, *Clustered nested sampling: Efficient Bayesian inference for cosmology*, *Monthly Notices of the Royal Astronomical Society* **378** (2007) 1365 [astro-ph/0701867].

[56] P. Mukherjee, D. Parkinson and A. R. Liddle, *A nested sampling algorithm for cosmological model selection*, *The Astrophysical Journal Letters* **638** (2006) L51 [astro-ph/0508461].

[57] D. Pelleg and A. W. Moore, *X-means: Extending k-means with efficient estimation of the number of clusters*, in *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML '00, (San Francisco, CA, USA), pp. 727–734, Morgan Kaufmann Publishers Inc., 2000, http://dl.acm.org/citation.cfm?id=645529.657808.

[58] F. Feroz, K. Cranmer, M. Hobson, R. Ruiz de Austri and R. Trotta, *Challenges of Profile Likelihood Evaluation in Multi-Dimensional SUSY Scans*, *Journal of High Energy Physics* **06** (2011) 042 [1101.3296].

[59] P. Scott, *Pippi — Painless parsing, post-processing and plotting of posterior and likelihood samples*, *The European Physical Journal Plus* **127** (2012) 138 [1206.2245].

[60] E. Gross and O. Vitells, *Trial factors for the look elsewhere effect in high energy physics*, *The European Physical Journal C* **70** (2010) 525 [1005.1891].

[61] O. Vitells and E. Gross, *Estimating the significance of a signal in a multi-dimensional search*, *Astroparticle Physics* **35** (2011) 230 [1105.4355].

[62] J. Allison et al., *Recent developments in Geant4*, *Nuclear Instruments and Methods in Physics Research Section A* **835** (2016) 186.

[63] J. Allison et al., *Geant4 developments and applications*, *IEEE Transactions on Nuclear Science* **53** (2006) 270.

[64] GEANT4 collaboration, *GEANT4: A Simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A* **506** (2003) 250.

[65] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *The European Physical Journal C* **72** (2012) 1896 [1111.6097].

[66] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, *Journal of High Energy Physics* **04** (2008) 005 [0802.1188].

[67] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Physics Letters B* **659** (2008) 119 [0707.1378].