

Kungliga Tekniska Högskolan (KTH)
Royal Institute of Technology

Master of Science in Engineering Physics

Prospective study for the development of an analysis to measure the tri-linear Higgs coupling using the Matrix Element Method with the ATLAS experiment at the HL-LHC

Florian Eble

Under the supervision of Jan Stark



May 2019



Abstract

The measurement of the Higgs boson self-coupling λ , also referred to as the Higgs boson trilinear coupling λ_{HHH} , is of great importance to experimentally reconstruct the Higgs potential and assess whether the Higgs boson discovered in 2012 at CERN is the one predicted by the Brout-Englert-Higgs mechanism. The Higgs boson potential was postulated *ad hoc* in the Standard Model (SM) and various extensions of the SM predict changes of the value of λ . The Higgs boson pair production, the most direct and stringent way to constrain λ , is a very rare process which has not yet been observed in the most recent data of the LHC (run 2). In this document, we present the first developments of an analysis to measure λ based on the Matrix Element Method (MEM) and assuming data from the HL-LHC. This method uses the likelihood to observe a sample of events given a theoretical model and a set of parameters. The MEM therefore makes it possible to determine which value of λ is the most likely to have produced a sample of di-Higgs event candidates. The MEM establishes a direct link between theoretical and experimental observables and is a statistically optimal method. The computation of the likelihoods was performed using the publicly available MadWeight and MoMEMta tools. Promising prospects were obtained using MoMEMta and a proof-of-principle of the analysis was established.

Acknowledgements

First of all, I would like to thank my supervisor Jan Stark for his constant availability, his valuable advice and his wise guidance throughout the Master Thesis. Jan's door was always open when I needed help. Jan shared his knowledge with passion and I have learned a lot with him. It was a great pleasure to work with such an involved physicist.

I would also like to thank:

- Nathan Readioff for providing detailed explanations and his code for the BDT training in his analysis for the measurement of the pair production and self-coupling of the Higgs boson.
- Annick Lleres for her work on the MEM for this analysis and her regular discussions about it.
- Benjamin Trocmé for organising group meetings where I could present my work and his support for extending my Master Thesis by one month to complete the work I had started.
- All the ATLAS team at LPSC for making me feeling comfortable at LPSC.
- Olivier Mattelaer for his discussions about the MEM and MadWeight, and his help in the process to understand the cause of the issue in the computation of the likelihoods with MadWeight.
- Sébastien Wertz for his advice on the Lua configuration files for future improvement of the analysis.
- the experts at CC-IN2P3 for the smooth running of the jobs.
- Luc Poggioli, responsible of the *Calcul ATLAS France* (CAF) group, for his advice on the efficient use of the ATLAS resources at CC-IN2P3.

Contents

List of Tables	v
List of Figures	vi
List of Abbreviations	ix
1 Introduction	1
2 Higgs boson pair production	4
2.1 Higgs self-coupling phenomenology	4
2.2 Signal final state and main background sources	4
2.3 Cross section calculations at different orders in perturbation theory	6
3 Description of the Matrix Element Method	8
3.1 Basic definitions	8
3.2 Likelihood definition	9
3.3 Phase-space parametrisation	11
3.4 Method of minimum negative log-likelihood	11
3.5 Comparison between cut-based analysis techniques and the Matrix Element Method	12
4 Description of the relevant Feynman diagrams for use in MadGraph5	13
4.1 Need for a specific UFO model	13
4.2 Existing UFO models	14
4.3 Creation and validation of a custom UFO model	14
5 Simulation of signal and background events	17
5.1 Event generation with MadEvent	17
5.2 Detector response	18
5.3 Pre-selection requirements	19
6 MadWeight: an automatic tool for the computation of the likelihoods	23
6.1 Introduction to MadWeight	23
6.2 Issues when running MadWeight on HH events: setup and code modifications . .	24
6.3 Matrix Element Analysis of pure signal ensembles	25
6.4 Matrix Element Analysis of mixed signal and non-resonant $b\bar{b}\gamma\gamma$ background ensembles	27
7 MoMEMta: a modular toolkit for the computation of the likelihoods	32
7.1 Introduction to MoMEMta	32

7.2	Technical issues encountered when running MoMEMta on HH events	33
7.3	Lua configuration files	34
7.4	Comparison between MadWeight and MoMEMta outputs	36
7.5	Matrix Element Analysis of pure signal ensembles	38
7.6	Matrix Element Analysis of a mixed signal and non-resonant $b\bar{b}\gamma\gamma$ background ensemble	42
7.7	Matrix Element Analysis of a mixed signal and $t\bar{t}H$ background ensemble	45
7.8	Matrix Element Analysis of a mixed signal, $b\bar{b}\gamma\gamma$ and $t\bar{t}H$ background ensemble	48
8	Conclusion and outlook	51
A	LO Run Card used for event generation	53
B	Dimension of the likelihood integral	59
B.1	Derivation of a general formula	59
B.2	Application to the signal and background processes	59
C	Distributions of the main kinematic variables for background events	61
C.1	Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events	61
C.2	Distributions of the main kinematic variables for the $t\bar{t}H$ background events	64
D	Lua configuration files	67
D.1	Lua configuration file for the signal and $b\bar{b}\gamma\gamma$ non-resonant background	67
D.2	Lua configuration file for the single Higgs boson $t\bar{t}H$ background	70
E	Event yield likelihood and extended likelihood	74
E.1	Event yield likelihood	74
E.2	Extended likelihood	74
F	Integration of the likelihood: convergence study	75
G	Performance of VEGAS and DIVONNE integration algorithms	80
	Bibliography	83

List of Tables

1.1	Table of expected constraints on λ	2
2.1	Table of most promising decay channels for a Higgs boson pair, with the corresponding branching ratio and approximate yield at 3000 fb^{-1} before any event selection is applied and assuming a total production cross-section of 40.8 fb . . .	5
2.2	Table of the main background sources	6
2.3	Table of the cross-section at $\sqrt{s} = 14 \text{ TeV}$ for Higgs boson pair production from gluon fusion at different orders using full theory or different approximations. . .	6
5.1	Expected number of events after pre-selection criteria are applied for signal and background processes.	19
G.1	Computation of the signal- $\kappa=1$ likelihood of several signal events, for which the integration with VEGAS seemed problematic, with VEGAS, DIVONNE and the basic Monte-Carlo with 15 million integrand evaluations.	81
G.2	Computation of the signal- $\kappa=1$ likelihood of several signal events, for which the integration with VEGAS seemed correct, with VEGAS, DIVONNE and the basic Monte-Carlo with 15 million integrand evaluations.	81

List of Figures

2.1	Feynman diagrams involved in the Higgs boson pair production at LO	4
2.2	Higgs boson pair production cross section at $\sqrt{s} = 14$ TeV as a function of $\lambda/\lambda^{\text{SM}}$	5
2.3	Examples of Feynman diagrams for the ttH and non-resonant $bb\gamma\gamma$ backgrounds.	7
3.1	Method of minimum negative log-likelihood method: graph of the negative log-likelihood, parabolic fit and measurement of a parameter with its associated statistical uncertainty.	12
4.1	Complete Feynman diagrams involved in the signal process at LO, including Higgs boson pair production and decay into a $b\bar{b}$ pair and a $\gamma\gamma$ pair.	14
4.2	Higgs boson pair production cross-section at LO and $\sqrt{s} = 14$ TeV as a function of κ	15
5.1	Distributions of the main kinematic variables for signal events after generation, after smearing, and after smearing and pre-selection requirements are applied. .	20
5.2	Distributions of the main kinematic variables for signal events after generation, after smearing, and after smearing and pre-selection requirements are applied. .	21
5.3	Distributions of the main kinematic variables for signal events after generation, after smearing, and after smearing and pre-selection requirements are applied. .	22
6.1	Negative logarithm of the event yield likelihood, the event kinematics likelihood from MadWeight and the extended likelihood for a sample of 16 signal events generated at $\kappa = 1$, and parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$	26
6.2	Histograms of the measured value of κ , the corresponding uncertainty and the pull value for signal events only, analysed with MadWeight using the signal Matrix Element.	27
6.3	Rejection power of the BDT and the LHR from MadWeight to classify signal events and $b\bar{b}\gamma\gamma$ continuum background events.	29
6.4	Distributions of main kinematic variables for the signal and $b\bar{b}\gamma\gamma$ background events, with and without cut on the LHR.	30
6.5	Distributions of main kinematic variables for the signal and $b\bar{b}\gamma\gamma$ background events, with and without cut on the LHR.	31
7.1	Comparison between MadWeight and MoMEMta working principles.	33
7.2	Comparison between MadWeight and MoMEMta outputs to the same events. . .	37
7.3	Negative logarithm of the event kinematics likelihood, the event yield likelihood and the extended likelihood as a function of κ for a set of 16 signal events without or with removing events having an event kinematics likelihood presenting fast variations.	39

7.4	Histograms of the measured value of κ , the corresponding statistical uncertainty and the pull value for signal events only, analysed with MoMEMta using the signal Matrix Element.	40
7.5	Histogram of the measured value of κ for a simple counting analysis.	41
7.6	Negative log-likelihood and histogram of the measured value of κ for a counting analysis in bins of m_{HH}	42
7.7	Rejection power of the BDT and the LHR from MoMEMta to classify signal events and $b\bar{b}\gamma\gamma$ continuum background events.	43
7.8	Negative logarithm of the event kinematics likelihood, the event yield likelihood and the extended likelihood as a function of κ for a sample of 16 signal events and 290 non-resonant $b\bar{b}\gamma\gamma$ background events.	44
7.9	Rejection power of the BDT and the LHR from MoMEMta to classify signal events and $t\bar{t}H$ background events.	47
7.10	Negative logarithm of the event kinematics likelihood, the event yield likelihood and the extended likelihood as a function of κ for a sample of 16 signal events and 84 $t\bar{t}H$ background events.	48
7.11	Negative logarithm of the event kinematics likelihood, the event yield likelihood and the extended likelihood as a function of κ for a sample of 16 signal events, 290 non-resonant $b\bar{b}\gamma\gamma$ background events and 84 $t\bar{t}H$ background events.	50
C.1	Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	61
C.2	Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	62
C.3	Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	63
C.4	Distributions of the main kinematic variables for the single Higgs boson $t\bar{t}H$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	64
C.5	Distributions of the main kinematic variables for the single Higgs boson $t\bar{t}H$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	65
C.6	Distributions of the main kinematic variables for the single Higgs boson $t\bar{t}H$ background events after generation, after smearing, and after smearing and pre-selection requirements are applied.	66
F.1	Event kinematics likelihood, event yield likelihood and extended likelihood as a function of κ for different sets of 16 signal events without or with removing events having an event kinematics likelihood presenting fast variations.	76
F.2	Correct integration and problematic integration of the likelihood of two signal events analysed with the signal hypothesis at $\kappa = 1$	77
F.3	Problematic integration of the likelihood for all values of κ for one signal event, with discontinuities in the likelihood curve as a function of κ	78
F.4	Problematic integration of the likelihood for all values of κ for one signal event, without discontinuities in the likelihood curve as a function of κ	79

List of Abbreviations

ATLAS A Thoroidal LHC ApparatuS, one of the two general purpose detectors at the LHC

BDT Boosted Decision Tree

Bkg Background

BSM Beyond Standard Model

CERN European Organisation for Nuclear Research. The abbreviation reflects the French name of the organisation, *Conseil Européen pour la Recherche Nucléaire*

CMS Compact Muon Solenoid, one of the two general purpose detectors at the LHC

HEFT Higgs Effective Field Theory

HL-LHC High Luminosity-Large Hadron Collider

HTL Heavy Top Limit

LAPP *Laboratoire d'Annecy de Physique des Particules*

LHC Large Hadron Collider

LHCO LHC Olympics

LHE Les Houches Event

LHR Likelihood Ratio

LO Leading Order

LPSC *Laboratoire de Physique Subatomique et de Cosmologie*

MC Monte Carlo

MEM Matrix Element Method

MET Missing E_T – Missing Transverse Energy

NLO Next-to-Leading Order

NNLO Next-to-Next-to-Leading Order

PDF Parton Distribution Function

SM Standard Model

UFO Universal FeynRules Output

Chapter 1

Introduction

Particle Physics is the science that studies the fundamental constituents of matter – elementary particles – their properties and their interactions. Particle physicists probe the smallest structures of Nature using complex scientific instruments called particle accelerators. Particle accelerators accelerate subatomic particles close to the speed of light and make them collide. Detectors surround the collision points to observe the particles produced in the collisions and the collected data are then analysed by physicists.

[CERN](#) (European Organisation for Nuclear Research, from the French name *Conseil Européen pour la Recherche Nucléaire*), founded in 1954 and headquartered near Geneva, hosts the Large Hadron Collider ([LHC](#)), a circular 27 kilometers circumference particle accelerator, as well as some of the major particle physics experiments in Europe. Proton beams going in opposite directions are accelerated until the protons acquire an energy of 6.5 GeV – 7 GeV after the future upgrade of the LHC into the [HL-LHC](#) – and are made to collide, in particular where the general-purpose detectors [ATLAS](#) and [CMS](#) are installed.

In 2012, CERN announced the discovery of a new particle, the Higgs boson, by experiments ATLAS and CMS. The existence of this particle was predicted in 1964 by Peter Higgs [\[1\]](#) to explain how gauge bosons can acquire non-zero masses as a result of spontaneous symmetry breaking within a gauge invariant model. Two other approaches to explain how mass could arise in gauge invariant theories were simultaneously published by Robert Brout and François Englert [\[2\]](#), and Gerald Guralnik, C. Richard Hagen and Tom Kibble [\[3\]](#). One challenge for the years to come is to measure precisely the couplings of this particle and to search for any deviations from the Higgs boson predicted by the Standard Model ([SM](#)). In particular, the measurement of the Higgs boson self-coupling λ (also referred to as the Higgs boson tri-linear coupling λ_{HHH}) is of great importance as it is expected to yield a deeper understanding of particle physics and cosmology. This measurement makes it possible to experimentally reconstruct the Higgs potential and check whether the Higgs boson discovered in 2012 at CERN is the one predicted by the Brout-Englert-Higgs mechanism. Furthermore, large deviations of the Higgs self-coupling from the Standard Model value could make possible the electroweak baryogenesis and help understand the asymmetry between matter and antimatter [\[4\]](#).

With the dataset obtained with the LHC up to now, the new particle discovered in 2012 is compatible with the Standard Model Higgs boson. Its couplings to the weak bosons (W and Z), to the third-generation quarks (t and b) and the third-generation charged lepton (τ) are measured with a precision of 10% to 25% by experiments ATLAS and CMS [\[5, 6, 7\]](#). Precision mea-

Analysis (using only the channel $b\bar{b}\gamma\gamma$ channel)	Expected constraint on λ
Cut-based analysis [9]	[-0.8 , 7.7] (at 95% CL)
Cut-based analysis using BDT [10]	[-0.1 , 2.4] (at 68% CL)
Theoretical best analysis (Neyman-Pearson lemma) [11]	[0.4 , 1.7] (at 68% CL)

Table 1.1: Table of expected constraints on λ

measurements of these quantities as well as the measurement of the tri-linear coupling are crucial because deviations from the Standard Model would indicate BSM phenomenology: in the SM the couplings of the Higgs boson to other particles are completely determined by their mass whereas in most BSM models the couplings are modified due to additional Feynman diagrams with virtual loops containing heavy particles. For that reason, more data are needed and the Large Hadron Collider will be upgraded into the High Luminosity Large Hadron Collider (HL-LHC): the instantaneous luminosity is expected to be five times larger the LHC nominal value, the energy of the beams will be raised to 7 GeV and critical components of the accelerator reaching the end of their lifetime will be replaced. The LHC will run until 2023, collecting 300 fb^{-1} of data over its whole lifetime. The installation of the HL-LHC will then take place from 2024 to 2026. During operation of the HL-LHC, approximately 250 fb^{-1} per year are expected to be collected, with the goal of collecting a total 3000 fb^{-1} a decade later, which is ten times higher than the expected value at the end of the LHC programme in 2023 [8].

The ATLAS group of *Laboratoire de Physique Subatomique et de Cosmologie* (LPSC) in Grenoble, France, is contributing to Higgs physics. In particular, part of the group works on prospects for the measurement of the Higgs boson self-coupling using the future data of the HL-LHC. After initial pessimistic studies using a simple cut-based analysis [9] developed at LAPP, further investigation at LPSC using Boosted Decision Trees (BDT) for background rejection and the di-Higgs mass m_{HH} in the extraction of λ yielded much better prospects for the measurement of λ [10]. The efforts for moving beyond simple cut-based analysis are strongly motivated by the theory paper [11] which quantifies, using the Neyman-Pearson lemma [12], the best accuracy that can theoretically be achieved. The expected constraints on λ from Refs. [9], [10] and [11] are presented in Tab. 1.1. In this context it was decided to study the performance of the Matrix Element Method (MEM) to measure the Higgs self-coupling and try to reach the narrowest uncertainty in Ref. [11].

Historically, the Matrix Element Method was first developed and used at the Tevatron experiments CDF and DØ to minimise the statistical uncertainty in the measurement of the $t\bar{t}$ events [13, 14, 15, 16, 17]. It has since provided the most precise measurement of the mass of the top quark with the DØ experiment [18]. This method is optimal by construction and establishes a direct link between theory and the experimental reconstruction of collisions in a detector [19]. In principle it can be used for any measurement and has never been applied to the measurement of the Higgs boson self-coupling. Two general implementations of the Matrix Element Method, MadWeight [20, 21] – fully automatic tool part of the MadGraph5_aMC@NLO framework [22] – and MoMEMta [23] – a very recent modular C++ toolkit for the MEM – are publicly available.

My work at LPSC has been to develop a first simple analysis based on the Matrix Element Method to determine λ and estimate the associated statistical uncertainties and expected significance. As a start, only the main sources of background are included. Signal and background

events are generated at LO at parton level using Monte Carlo simulations and are then smeared using a simple parametrised smearing function to simulate the response of the detector. Ultimately, the MEM analysis will have to be tested using fully simulated events. This analysis will serve as a benchmark for future more refined analyses.

The present document is my Master Thesis report as well as a technical document for physicists who will continue working on this analysis.

Chapter 2

Higgs boson pair production

2.1 Higgs self-coupling phenomenology

In proton-proton collisions, the dominant Higgs boson pair production mode is gluon fusion. Leading order (LO) Feynman diagrams for gluon fusion are shown in Fig. 2.1. These two diagrams contain a quark loop: a triangle and a box. These loops are dominated by the contribution of the top quark, since the heavier the particle, the stronger the Higgs boson couples to this particle. Only the triangle diagram involves the tri-linear Higgs coupling λ_{HHH} and it is impossible to disentangle the contributions of these two diagrams. The amplitudes of the two diagrams have opposite sign, which results in a destructive interference that reduces the total Higgs boson pair production cross-section.

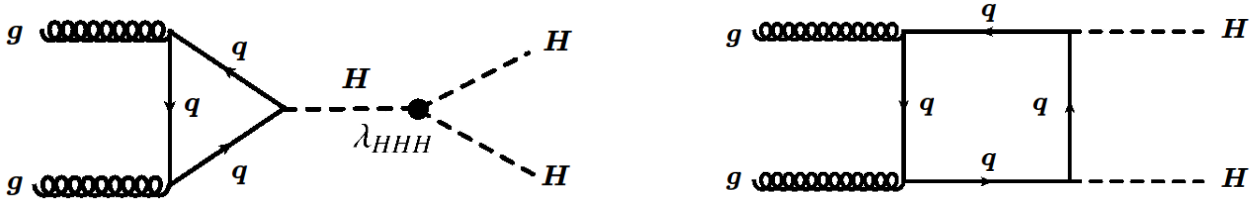


Figure 2.1: Feynman diagrams involved in the Higgs boson pair production at LO (from Ref. [24]).

The cross-section of the di-Higgs production from gluon fusion is displayed in Fig. 2.2. The strong dependence of the cross-section on the ratio of λ_{HHH} to its value in the Standard Model – $\lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$ – illustrates why a counting analysis is powerful to measure λ_{HHH} .

2.2 Signal final state and main background sources

The Higgs boson pair has various decay modes. The most promising ones are given in Tab. 2.1. The choice of the signal final state is a compromise between signal extraction quality and event quantity. The $H \rightarrow \gamma\gamma$ decay offers a very narrow mass peak and provides a very clean signal extraction. Other decay channels suffer from a broader mass peak and do not allow such a clean signal extraction. However the branching ratio of the $H \rightarrow \gamma\gamma$ decay is very small ($0.00227 \pm 1.73\%$ [28]) resulting in a very low expected yield for the $\gamma\gamma + \gamma\gamma$ final state. The

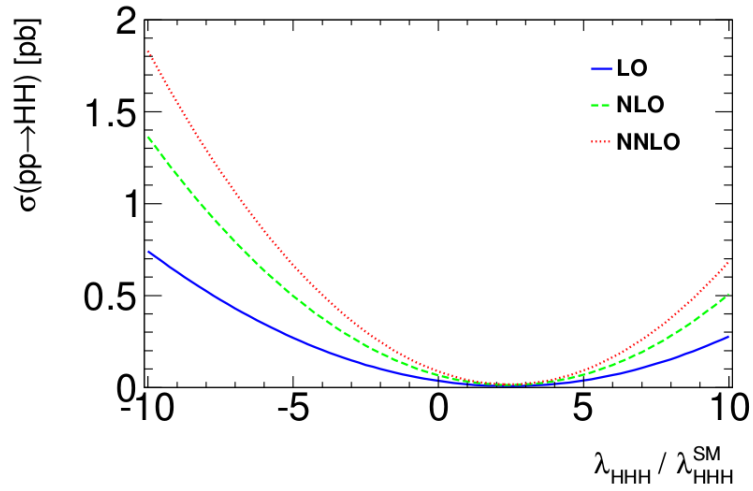


Figure 2.2: Higgs boson pair production cross section at $\sqrt{s} = 14$ TeV as a function of $\lambda/\lambda^{\text{SM}}$. The LO and NLO values are obtained with the HPAIR program [25]. References [26, 27] are used to obtain the NNLO cross-section. This figure was taken from Ref. [24].

Decay Channel	Branching Ratio (%)	Total Yield at 3000 fb ⁻¹
$b\bar{b} + b\bar{b}$	33	40 000
$b\bar{b} + W^+W^-$	25	31 000
$b\bar{b} + \tau\bar{\tau}$	7.3	8 900
$b\bar{b} + ZZ$	3.1	3 800
$\tau\bar{\tau} + W^+W^-$	2.7	3 300
$ZZ + W^+W^-$	1.1	1 300
$b\bar{b} + \gamma\gamma$	0.26	320
$\gamma\gamma + \gamma\gamma$	0.0010	1.2

Table 2.1: Table of most interesting possible decay channels for a Higgs boson pair, with the corresponding branching ratio and approximate yield at 3000 fb⁻¹ before any event selection is applied and assuming a total production cross-section of 40.8 fb (from Refs. [24], [26] and [27]).

branching ratio of the $H \rightarrow b\bar{b}$ decay is a lot larger ($0.587 \pm 0.65\%$ [28]). The $b\bar{b} + \gamma\gamma$ final state turns out to be a good compromise with an expected yield around 320 events at 3000 fb⁻¹. This document therefore examines the $b\bar{b} + \gamma\gamma$ channel.

Only the sources of background given in Tab 2.2 were considered in this analysis: the main non-resonant background, $b\bar{b}\gamma\gamma$, and the main single Higgs boson background, $t\bar{t}H$. Examples of Feynman diagrams for these two background processes are presented in Fig. 2.3. The non-resonant $b\bar{b}\gamma\gamma$ background has been used to develop the analysis at first and, unless stated otherwise, plots describing the analysis that include background events only include the non-resonant $b\bar{b}\gamma\gamma$ background. Expected yields for the main background sources will be detailed after pre-selection is exposed, in Tab. 5.1. It should be noted that, after pre-selection, the number of signal events relative to the number of background events is about 1%, which is very low.

Main Background Processes
non-resonant $bb\gamma\gamma$ production
$ttH(\rightarrow \gamma\gamma)$

Table 2.2: Table of the main background sources.

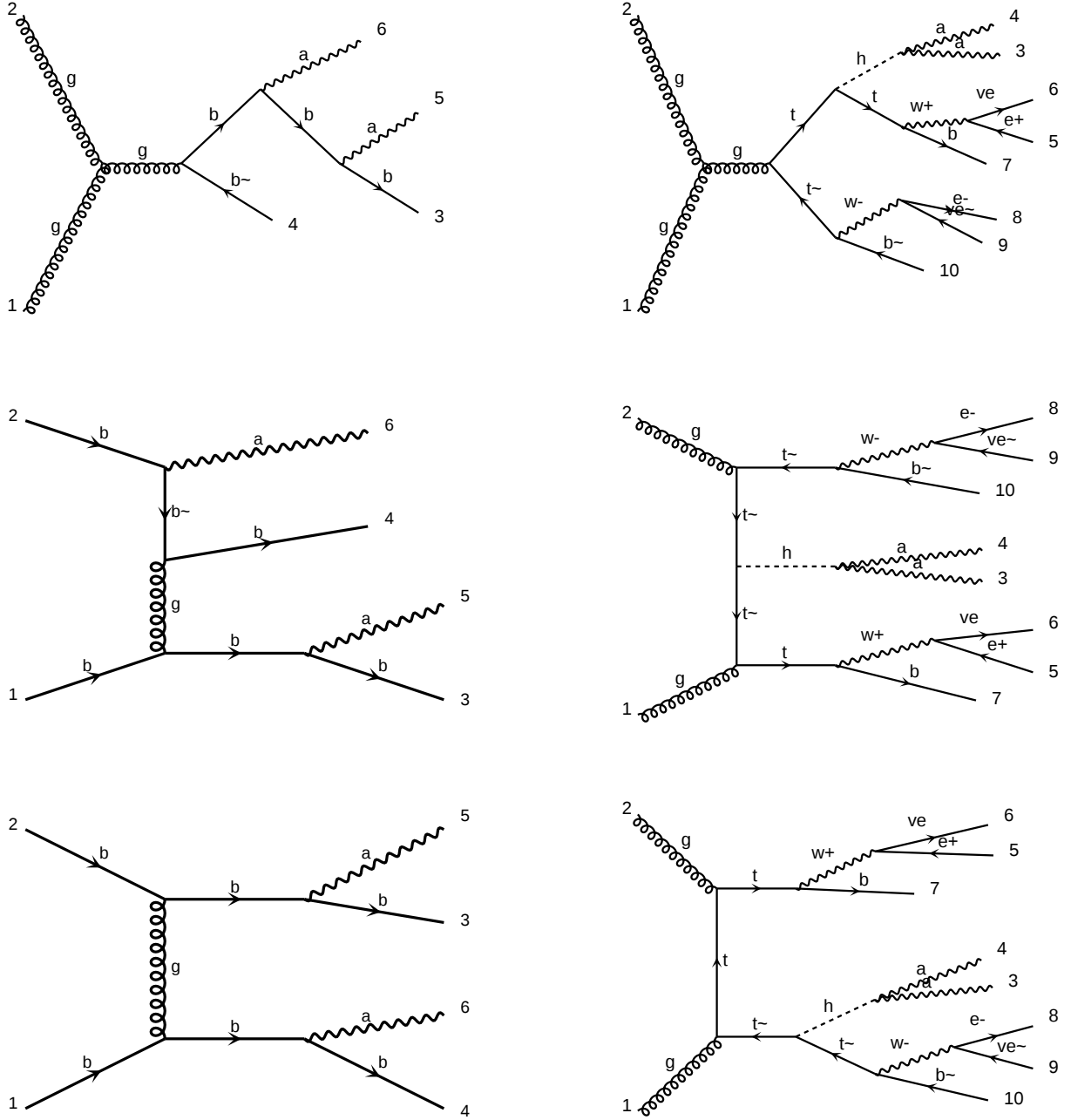
	σ_{LO} (fb)	σ_{NLO} (fb)	σ_{NNLO} (fb)
Basic HTL	$17.07^{+30.9\%}_{-22.2\%}$ [30, 31]	$31.93^{+17.6\%}_{-15.2\%}$ [30, 31]	$37.52^{+5.2\%}_{-7.6\%}$ [32]
B-i/proj HTL	$19.85^{+27.6\%}_{-20.5\%}$ [30, 31]	$38.32^{+18.1\%}_{-14.9\%}$ [30, 31]	$39.58^{+1.4\%}_{-4.7\%}$ [33]
FTapprox	$19.85^{+27.6\%}_{-20.5\%}$ [34]	$34.25^{+14.7\%}_{-13.2\%}$ [34]	$36.69^{+2.1\%}_{-4.9\%}$ [33]
Full Theory	$19.85^{+27.6\%}_{-20.5\%}$ [30, 31]	$32.88^{+13.5\%}_{-12.5\%}$ [29]	—

Table 2.3: Table of the cross-section at $\sqrt{s} = 14$ TeV for Higgs boson pair production from gluon fusion at different orders using full theory or different approximations.

2.3 Cross section calculations at different orders in perturbation theory

For the purpose of using automated tools for *e.g.* cross-section calculations and event generation, the loops involved in the Higgs boson pair production can be approximated by a point interaction with an effective coupling set to match the available loop calculations. The point-like interaction being only valid in the infinite mass limit for the particle involved in the loop, all kinematic variables of the Higgs boson should be smaller than twice the mass of the particle in the loop. Given the fact that the heavier the particle, the stronger the Higgs couples to it, we will only consider the top quark in the loops of the Higgs boson pair production diagrams. The infinite top mass approximation is referred to as Heavy Top Limit (HTL). The HTL is not completely fulfilled in the di-Higgs production at HL-LHC since most of the di-Higgs mass distribution is above 350 GeV (see Fig. 5.2). It is possible to correct for this since detailed calculations of differential cross sections taking into account the full mass dependence are available. The differential cross-sections must then be rescaled with the so-called K-factor $K = \sigma_{\text{finite mT}}/\sigma_{\text{infinite mT}}$ (the cross section $\sigma_{\text{finite mT}}$ being computed at NLO or NNLO). Figure 1 of Ref. [29] gives K-factors, as a function of the di-Higgs mass, for rescaling at different levels of precision (NLO HEFT, NLO FTapprox, NLO). Table 2.3 summarises the cross-section calculations at different orders using different approximations. The FTapprox calculation at NNLO is the most precise so far.

In this study no K-factor was used to correct for the HTL approximation. The ATLAS collaboration, *e.g.* for the cut-based analysis using BDT [10], simulates events at LO with parton shower and reweights the cross-section to FTapprox at NNLO.



(a) Examples of Feynman diagrams for the non-resonant $bb\gamma\gamma$ background.

(b) Examples of Feynman diagrams for the ttH background.

Figure 2.3: Examples of Feynman diagrams for the ttH and non-resonant $bb\gamma\gamma$ backgrounds.

Chapter 3

Description of the Matrix Element Method

Given a sample of events and different theoretical models or different parameters of a model, the Matrix Element Method ([MEM](#)) is a procedure which aims to determine which of the models or sets of model parameter values is the most likely to have produced these events. This method makes use of both theoretical information (via the Matrix Element of the considered process within the assumed model) and experimental information (via transfer functions describing the resolution of the detector). For an assumed model or assumed values for model parameters, this method provides the probability that the sample of events was produced, based on event-by-event calculations. By repeating this procedure for several hypotheses, the likelihood of the different theoretical models or the likelihood as a function of the model parameter(s) is obtained. The most plausible model or set of model parameter(s) is the one for which the likelihood is maximum.

In this study, we have used the MEM to analyse a sample of candidate di-Higgs events for different values of the ratio of λ_{HHH} to its value in the Standard Model, $\lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$. We will refer to this parameter as κ in the rest of this document. We will use the negative log-likelihood $-\ln(L)$ instead of the likelihood L , and thus look for its minimum value.

3.1 Basic definitions

Before getting into the details of the definition of the likelihood, we define the quantities that will be used in the rest of this section.

Event sample

Let us consider a sample of N events composed of n final state partons, all corresponding to the same final state F and being candidate signal events.

Let x_j^i be the measured four-vector of the particle j of the event i .

Let $\mathbf{x}^i = (x_1^i, \dots, x_n^i)$ be the measured four-vectors of the n final state particles of the event i .

Let $\mathbf{x} = (\mathbf{x}^1, \dots, \mathbf{x}^N)$ be the full description of the event sample.

Hypothesis

Let $\mathbf{h} = (h_1, \dots, h_{n_h})$ be a set of n_h assumed values for model parameters. This set of values is the hypothesis under which the likelihood is computed.

Process fraction

The likelihood must take into account that different processes (signal and background processes) can lead to the same final state F . Let f_p be the fraction of events from process p in the entire event sample, satisfying the relation $\sum_{p=1}^{n_p} f_p = 1$, where n_p is the number of processes. If unknown, these fractions are additional parameters in the hypothesis and are to be determined simultaneously with the other physics parameters during the minimisation of $-\ln L_{\text{sample}}$. In the case of the measurement of λ , the fractions of background processes can be efficiently constrained. In this study, they will be treated as known parameters.

Likelihoods

Let $L_{\text{sample}}(\mathbf{h}|\mathbf{x})$ be the likelihood of observing the event sample with the properties $\mathbf{x}$ for the hypothesis $\mathbf{h}$.

Let $L_{\text{event}}^i(\mathbf{h}|\mathbf{x}^i)$ be the likelihood of observing the event i with the measured four-vectors $\mathbf{x}^i$ for the hypothesis $\mathbf{h}$.

Let $L_{\text{process}}^p(\mathbf{h}|\mathbf{x}^i)$ be the likelihood of observing the event i with the measured four-vectors $\mathbf{x}^i$, for the process p and the hypothesis $\mathbf{h}$.

3.2 Likelihood definition

Sample likelihood

The likelihood which will be maximised is the likelihood of the entire sample. The sample likelihood is the product of the event likelihoods:

$$L_{\text{sample}}(\mathbf{h}|\mathbf{x}) = \prod_{i=1}^N L_{\text{event}}^i(\mathbf{h}|\mathbf{x}^i) \quad (3.1)$$

Event likelihood

The likelihood of one event is a linear combination of the likelihoods of this event for all the processes:

$$L_{\text{event}}^i(\mathbf{h}|\mathbf{x}^i) = \sum_{p=1}^{n_p} f_p L_{\text{process}}^p(\mathbf{h}|\mathbf{x}^i) \quad (3.2)$$

Likelihood of a process

In a collision between two partons a_1 and a_2 , the likelihood of the n final state partons of an event to be *produced* with the four-momenta $\mathbf{y} = (y_1, \dots, y_n)$ by the process p with the assumed parameters $\mathbf{h}$ is proportional to the differential cross-section of the corresponding process:

$$d\sigma_p(a_1 a_2 \rightarrow \mathbf{y}; \mathbf{h}) = \frac{(2\pi)^4 |\mathcal{M}_p(a_1 a_2 \rightarrow \mathbf{y}; \mathbf{h})|^2}{q_1 q_2 s} \delta \left(a_1 + a_2 - \sum_{j=1}^n y_j \right) d^{4n} y \quad (3.3)$$

- where
- $\mathcal{M}_p(a_1 a_2 \rightarrow \mathbf{y}; \mathbf{h})$ denotes the Matrix Element of $a_1 a_2 \rightarrow \mathbf{y}$ for the process p and the hypothesis $\mathbf{h}$, obtained from an analytical calculation,
 - q_1 and q_2 are the momentum fractions of partons a_1 and a_2 (also referred to as Bjorken x scaling variables),
 - s is the Mandelstam variable corresponding to the squared center-of-mass energy of the proton-proton collision,
 - $\delta\left(a_1 + a_2 - \sum_{j=1}^n y_j\right)$ is the energy-momentum conservation,
 - $d^{4n}y$ is an element of the phase space of the n final-state partons.

The differential cross-section of proton-proton collisions is the convolution of the cross-section of partons collisions, given in Eq. 3.3, with the Parton Distribution Functions (PDF), summed over all possible flavors a_1 and a_2 of the colliding partons:

$$d\sigma_p(pp \rightarrow \mathbf{y}; \mathbf{h}) = \int_{q_1, q_2} \sum_{a_1, a_2} f_{a_1}(q_1) f_{a_2}(q_2) d\sigma_p(a_1 a_2 \rightarrow \mathbf{y}; \mathbf{h}) dq_1 dq_2 \quad (3.4)$$

where $f_a(q)$ is the probability density to find a parton of flavor a and momentum fraction q in the proton.

The resolution of the detector is taken into account by convoluting this differential cross-section with the transfer function $W(\mathbf{x}^i, \mathbf{y})$, which describes the probability for a partonic final state $\mathbf{y}$ to be reconstructed as $\mathbf{x}^i$ in the detector. In the MEM, it is commonly assumed that the transfer function is "factorisable": it can be written as the product of single-particle resolution functions, which can each be written as the product of the resolution functions associated with the physical quantities that are measured. These transfer functions are obtained from a parametrised model of the resolution of the detector. The likelihood to *measure* the n final state partons of an event with the four-momenta $\mathbf{x}^i = (x_1^i, \dots, x_n^i)$, under the hypothesis that they have been produced by the process p with the assumed parameters $\mathbf{h}$ is proportional to the differential cross-section:

$$d\sigma_p(pp \rightarrow \mathbf{x}^i; \mathbf{h}, W) = \int_{\mathbf{y}} d\sigma_p(pp \rightarrow \mathbf{y}; \mathbf{h}) W(\mathbf{x}^i, \mathbf{y}) \quad (3.5)$$

In order to obtain the likelihood for the process p , this differential cross-section to observe a given reconstructed event needs to be normalised to the observable cross-section to produce the final state partons F , denoted $\sigma_p^{\text{obs}}(pp \rightarrow F)$:

$$L_{\text{process}}^p(\mathbf{h}|\mathbf{x}^i) d\mathbf{x}^i = \frac{d\sigma_p(pp \rightarrow \mathbf{x}^i; \mathbf{h}, W)}{\sigma_p^{\text{obs}}(pp \rightarrow F)} \quad (3.6)$$

The (differential) likelihood to observe the event described by the four-vectors $\mathbf{x}^i$ and composed of the final state partons F , in the hypothesis that they have been produced from a proton-proton collision by the process p with the assumed parameters $\mathbf{h}$ is therefore:

$$L_{\text{process}}^p(\mathbf{h}|\mathbf{x}^i) d\mathbf{x}^i = \frac{(2\pi)^4}{\sigma_p^{\text{obs}}(pp \rightarrow F) s} \int_{\mathbf{y}} \int_{q_1, q_2} \sum_{a_1, a_2} f_{a_1}(q_1) f_{a_2}(q_2) \frac{|\mathcal{M}_p(a_1 a_2 \rightarrow \mathbf{y}; \mathbf{h})|^2}{q_1 q_2} W(\mathbf{x}^i, \mathbf{y}) \delta\left(a_1 + a_2 - \sum_{j=1}^n y_j\right) dq_1 dq_2 d^{4n}y \quad (3.7)$$

3.3 Phase-space parametrisation

The likelihood is a convolution of the Matrix Element squared and the transfer function, which both present variations of several orders of magnitude in different regions of the phase-space. For example, the Matrix Element has narrow peaks corresponding to the resonance of the propagators. The numerical integration of the likelihood is thus a complicated task. Therefore, a general algorithm for MadWeight [20, 21] has been developed to provide a phase-space mapping optimised for integration by adaptive Monte-Carlo methods, such as VEGAS [35].

The optimised phase-space mapping is such that the variables corresponding to the strength of a narrow peak of the likelihood, either from the transfer function or the Matrix Element, are each mapped onto a single variable of integration. The corresponding set of variables and their associated change of variables is referred to as a block. There are two sorts of blocks: Main Blocks and Secondary Blocks. Main Blocks correspond to changes of variables that are designed to exploit the energy-momentum conservation: 4 variables are integrated out due to the δ function in Eq. 3.7. Integration variables that are not included in the Main Block are part of the Secondary Blocks. In some cases, no Secondary Blocks are needed.

In this study, we have used the Main Blocks A and F, and no Secondary Block was used. Let us give some details about these two blocks:

- Main Block A: This transformation integrates out the momentum fraction of the colliding partons q_1 and q_2 and the norm of the 3-momenta of two visible final state particles.

Example: $pp \rightarrow H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma)$
 $pp \rightarrow b\bar{b}\gamma\gamma$

- Main Block F: In this transformation, the three-momenta of two invisible final state particles (*i.e.* 6 variables) are replaced by the invariant masses of the mother particles of each missing particles (*i.e.* 2 variables).

Example: $pp \rightarrow (W^+ \rightarrow l^+\nu_l)(W^- \rightarrow l^-\bar{\nu}_l)$

3.4 Method of minimum negative log-likelihood

General formulation of the principle

The most plausible set of parameter values $\tilde{\mathbf{h}}$ for the production of the event sample is the value for which the sample likelihood $L_{\text{sample}}(\mathbf{h}|\mathbf{x})$ is maximum. Instead of working with the likelihood, we will use its negative logarithm, referred to as negative log-likelihood: $-\ln(L_{\text{sample}}(\mathbf{h}|\mathbf{x}))$.

The most plausible hypothesis $\tilde{\mathbf{h}}$ is thus the one that minimises the negative log-likelihood. The statistical uncertainty σ on the measurement of $\mathbf{h}$ is given by the values of $\mathbf{h}$ where the negative log-likelihood rises of 0.5 above its minimum [36]:

$$-\ln(L_{\text{sample}}(\tilde{\mathbf{h}} + \sigma|\mathbf{x})) = -\ln(L_{\text{sample}}(\tilde{\mathbf{h}}|\mathbf{x})) + 0.5.$$

Formulation in the particular case of this study

In our study, the parameter to be measured is $\kappa = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$: the most plausible value of κ is the one which minimises the negative log-likelihood $-\ln(L_{\text{sample}}(\kappa))$.

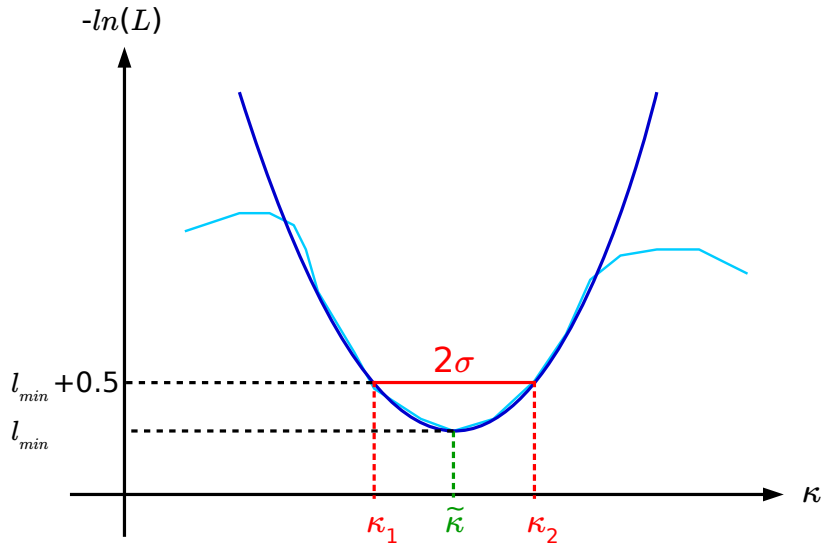


Figure 3.1: Method of minimum negative log-likelihood method: graph of the negative log-likelihood (cyan), parabolic fit (blue) and measurement of a parameter (green) with its associated statistical uncertainty ($\sigma = (\kappa_2 - \kappa_1)/2$, red).

In practice, the negative log-likelihood can be approximated by a second order polynomial in the region around its minimum, and the negative log-likelihood is fitted by a parabola. We determine the minimum of the negative log-likelihood and the statistical uncertainty using the parabolic fit: if p is the second order polynomial fit and κ_1, κ_2 are the two solutions of $p(\kappa) = \min(p) + 0.5$ ($\kappa_1 < \kappa_2$), then $\tilde{\kappa} = \min(p) \pm (\kappa_2 - \kappa_1)/2$. Figure 3.1 illustrates this procedure.

3.5 Comparison between cut-based analysis techniques and the Matrix Element Method

The MEM analysis is conceptually very different from cut-based analysis techniques. In the MEM, all the available information is used – all events in the dataset, all theoretical and all experimental information – making it an optimal analysis technique.

Unlike cut-based analyses which rely on the event yield after stringent event selection requirements, all the events available after soft pre-selection requirements (this will be explained later in Sec. 5.3) are analysed. Furthermore, the MEM analysis is optimal by construction since it uses all the available theoretical (Matrix Element) and experimental (measured three-vectors of the final-state particles and transfer function) information. In addition, the event yield can be taken into account in the MEM analysis by using the extended likelihood (see Appendix E; this was done in this study). Several unknown parameters, whether they are theoretical parameters (describing physics processes) or experimental parameters (describing the detector response), can be determined at the same time, thus reducing systematic uncertainties. Large gains are expected in comparison with cut-based analyses. This is a major advantage in analyses of very rare signals like $pp \rightarrow H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma)$: in the cut-based analysis of Ref. [10], only 6.5 signal events are expected *on average* after the BDT-based selection.

Chapter 4

Description of the relevant Feynman diagrams for use in the MadGraph5 framework: Universal FeynRules Output model

4.1 Need for a specific UFO model

As discussed in Sec. 2.1, the Matrix Element for the di-Higgs production at leading order is based on the interference of two diagrams. Figure 4.1 shows the complete Feynman diagrams for the signal, including Higgs boson pair production and its decay into a $b\bar{b}$ pair and a $\gamma\gamma$ pair. Thanks to the Higgs boson being a spin-0 particle, spin correlations do not need to be considered in the decay chains. The Feynman diagrams involve loops – gluon fusion to one Higgs boson, to two Higgs bosons and Higgs boson decay into $\gamma\gamma$ – from which infinities arise and that cannot be handled by automated means. By-hand computations and renormalisation techniques are thus needed to compute such Feynman diagrams.

For the purpose of using automated tools, these loops can be approximated by a point interaction with an effective coupling set to match the available loop calculations. As discussed in Sec. 2.3, the approximation is only valid in the infinite mass limit of the particles involved in the loop. Loops involved in the di-Higgs production have already been addressed in Sec. 2.3. In its decay into $\gamma\gamma$, the Higgs boson mainly couples to the W boson and top quark. The infinite mass limit is fulfilled for this process since $m_H < 2m_t$ and $m_H < 2m_W$. The computations of the form factors for the $H\gamma\gamma$, ggH and $ggHH$ effective vertices are detailed in Refs. [37], [38] and [39].

In the framework we have used to generate events and compute likelihoods – MadEvent [40] and MadWeight [20, 21] in the MadGraph_aMC@NLO framework [22] – the physics model is encoded in the so-called Universal FeynRules Output format (UFO) [41]. All the information of the model – particles, parameters, couplings, vertices – is stored in a set of Python files which describe all possible interactions within the model. Therefore we need a UFO model that includes the three effective vertices ggH , $ggHH$ and $H\gamma\gamma$ to be able to compute the Matrix Element for the signal process. Furthermore, in order to test different hypotheses for the tri-linear Higgs coupling λ_{HHH} , a multiplicative factor to the tri-linear Higgs coupling needs to be introduced in the model. This parameter will be referred to as κ in the following and is

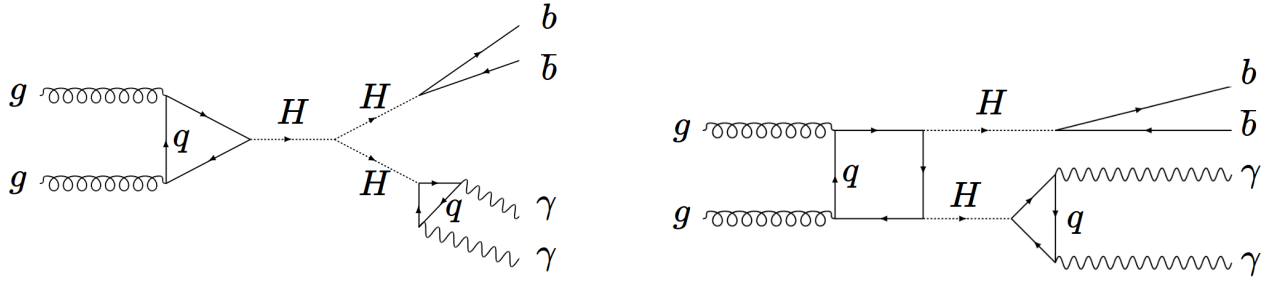


Figure 4.1: Complete Feynman diagrams involved in the signal process at LO, including Higgs boson pair production and decay into a $b\bar{b}$ pair and a $\gamma\gamma$ pair.

defined as:

$$\kappa = \frac{\lambda_{HHH}}{\lambda_{HHH}^{\text{SM}}}.$$

4.2 Existing UFO models

The Higgs Effective Field Theory (HEFT) [42] is a model in which the Higgs boson couples directly to gluons and photons. Fermionic and W boson loops involved in gluon fusion and photon fusion are approximated by a point interaction with an effective coupling, as described in Sec. 4.1. The ggH and $H\gamma\gamma$ effective vertices are implemented in the HEFT model but not the $ggHH$ vertex. Based on Ref. [39], the SM_EFT_FF.bt model includes form factors for the Higgs pair production in gluon–gluon collisions. This model includes the ggH and $ggHH$ effective vertices, but it does not include $H\gamma\gamma$ vertex.

There is therefore no model including all the effective vertices needed to compute the signal Matrix Element, but each of the necessary effective vertices is available in at least one of the two UFO models.

4.3 Creation and validation of a custom UFO model

The implementation of a new UFO model by hand without automatic tools is usually a tedious and error-prone task due to the complex description of the interaction between particles. The first step in the automatization of the implementation of a UFO model has been to start from the Lagrangian to automatically derive the Feynman Rules from it, and then to interface the resulting rules with the MC generator used. An interface with FeynRules allows one to export a new model into the UFO format [41].

It is usually not recommended to change a model by hand because the modifications should be consistent with the rest of the model. However, adding an effective field theory to the SM consists in adding a new term to the Lagrangian, which results in new Feynman Rules that do not interfere with the rest of the Lagrangian. It is therefore safe to add the effective vertex missing in a model based on the definition in the other model. We have thus added the effective vertex $H\gamma\gamma$, with its corresponding coupling, as described in the HEFT model into the model

SM_EFT_FF_bt. We have then added κ as a new external parameter to the UFO model and multiplied the coupling corresponding to the tri-linear Higgs vertex by this parameter. Therefore it becomes possible to compute the signal Matrix Element for any value of κ . To remove any ambiguity in the following, we will refer to this new model as **SM_EFT_FF_bt_haa_kappa**.

Validation of the UFO model

In order to check this new model, we have computed the cross section of the Higgs boson pair production at $\sqrt{s} = 14$ TeV and $\kappa = 1$, and the branching ratio of $HH \rightarrow b\bar{b}\gamma\gamma$. We have then compared these values to the expected numbers. We have also tested the dependence of the Higgs boson pair production cross-section as a function of κ .

The LO cross section of the production of Higgs boson pairs via gluon fusion $gg \rightarrow HH$ at 14 TeV is found to be $0.01718 \pm 1.615 \times 10^{-5}$ pb using MadEvent [40]. This is compatible with official ATLAS di-Higgs production samples, for which a cross-section of $0.01719 \pm 1.072 \times 10^{-5}$ pb is reported [43]. This is also in good agreement with the theoretical calculations in the HTL approximation (cf. Tab. 2.3): $17.07^{+30.9\%}_{-22.2\%}$ fb.

Unlike the ATLAS collaboration, which uses Pythia to decay the Higgs bosons, the decay of Higgs bosons is carried out using the custom **SM_EFT_FF_bt_haa_kappa** UFO model in this study. The cross-section of the full process $gg \rightarrow H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma)$ computed using MadEvent with the **SM_EFT_FF_bt_haa_kappa** UFO model is 0.05631 fb. The branching ratio for the decay $HH \rightarrow b\bar{b}\gamma\gamma$ is therefore 0.003278. Given that the most recent computations for the branching ratio are $BR(H \rightarrow b\bar{b}) = 0.5824 \pm 0.65\%$ and $BR(H \rightarrow \gamma\gamma) = 0.002270 \pm 1.73\%$ [28], the current best estimation of the branching ratio for $HH \rightarrow b\bar{b}\gamma\gamma$ is 0.002641. This is not too far from our prediction, and the production rate is anyway re-normalized using the cross-section at NNLO-FTapprox (cf. Tab. 2.3).

We have also tested the implementation of the κ parameter in the UFO model **SM_EFT_FF_bt_haa_kappa**. Figure 4.2 shows the Higgs boson pair production cross-section at LO and $\sqrt{s} = 14$ TeV as a function of κ . This plot compares very well to Fig. 2.2.

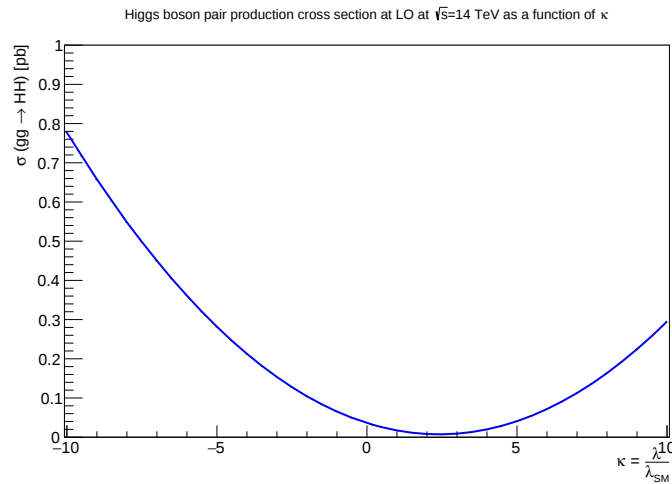


Figure 4.2: Higgs boson pair production cross-section at LO and $\sqrt{s} = 14$ TeV as a function of κ . The cross-section was computed using MadEvent [40] and the UFO model **SM_EFT_FF_bt_haa_kappa**.

Another test of the model is to generate events, draw the distribution of all kinematic variables and compare them to expected distributions. However, this is not easily doable since ATLAS official samples are different from our events and have different kinematic distributions. The ATLAS collaboration do use MadGraph5 as a starting point to produce event samples. But the output of MadGraph5 is the Higgs boson pair, and is subject to further processing: Pythia is used to add additional QCD radiation, to hadronise partons and to decay the Higgs bosons. Truth-level jets are then reconstructed by running a jet clustering algorithm on the truth-level event (after fragmentation). The truth-level jets are used for generator-level studies rather than the partons (which are used in this report). This results in considerable differences in some kinematic variables such as the invariant mass of the $b\bar{b}$ pair. The QCD radiation from Pythia leads to a non-zero transverse boost of the di-Higgs system.

Chapter 5

Simulation of signal and background events

5.1 Event generation with MadEvent

Signal and background events were generated at LO at parton level using the multi-purpose event generator MadEvent [40], part of the MadGraph5_aMC@NLO framework¹ [22]. MadGraph command lines produce a process-specific stand-alone code which allows the user to calculate cross sections and to produce unweighted events in the standard LHE output format (Les Houches Event). The PDF used for event generation is NNPDF23_lo_as_0130_qed.

The main command lines used to generate signal events and background events are given below:

- Signal event generation

```
MG5_aMC> import model SM_EFT_FF_bt_haa_kappa
MG5_aMC> generate p p > h h , h > b b~, h > a a
```

- $b\bar{b}\gamma\gamma$ continuum background event generation

```
MG5_aMC> import model SM_EFT_FF_bt_haa_kappa
MG5_aMC> define p = g u c d s b u~ c~ d~ s~ b~
MG5_aMC> define x = b b~
MG5_aMC> generate p p > x x a a
```

- $t\bar{t}H$ background event generation

```
MG5_aMC> import model SM_EFT_FF_bt_haa_kappa
MG5_aMC> define p = g u c d s b u~ c~ d~ s~ b~
MG5_aMC> generate p p > h t t~, h > a a , (t > w+ b, w+ > e+ ve) ,
(t~ > w- b~, w- > e- ve~)
```

For the $t\bar{t}H$ background process, the final-state is not exactly $b\bar{b}\gamma\gamma$: the Higgs boson decays into a pair of photon and each top quark decays into a b quark and a W boson. The W boson cannot decay into a photon and the branching ratio into a b quark and an up quark (resp. charm quark) is extremely low: 0.00004 ± 0.00010 (resp. 0.00080 ± 0.00005) (Tab. 7 of

¹Version 2.3.3 released on 25/10/2015 was used.

Ref. [44]). Therefore we can neglect this decay channel in this simplified study and consider that the b quarks can only be produced by the decay of the top quarks. The production of ttH events therefore consists in producing two top quarks and one Higgs boson, with the Higgs boson decaying into two photons and the two top quarks decaying in a b quark and a W boson. What the W boson decays into does not matter, so we have simulated its decay into an electron and a neutrino. We will ignore the W decay in the subsequent analysis.

Generator-level requirements

The production rate for the $b\bar{b}\gamma\gamma$ non resonant background becomes very large for low transverse momentum p_T of the final state particles. To not waste CPU in the generation of uninteresting background events, some requirements were applied during the background event generation. These requirements do not aim at enhancing the analysis and are not part of the analysis. The following requirements were applied:

- minimum p_T for the b quarks: $p_T(b) \geq 20.0$ GeV
- minimum p_T for the photons: $p_T(\gamma) \geq 25.0$ GeV
- minimum invariant $b\bar{b}$ mass: $m_{b\bar{b}} \geq 45.0$ GeV
- minimum invariant di-photon mass: $m_{\gamma\gamma} \geq 60.0$ GeV
- maximum invariant di-photon mass: $m_{\gamma\gamma} \leq 200.0$ GeV
- minimum distance between b quarks: $\Delta R \geq 0.4$
- minimum distance between photons: $\Delta R \geq 0.4$
- minimum distance between b quarks and photons: $\Delta R \geq 0.4$

No selection requirement was applied for the generation of the signal events and the ttH background events.

In order to set the collision energy to 14 TeV, at which the HL-LHC will operate, the energy of the colliding beams is set to 7000 GeV. The LO card used for background event generation can be found in Appendix A.

5.2 Detector response

After generation of the truth-level events, the response of the detector needs to be simulated to emulate real data events. For the purpose of this initial study of the MEM, the response of the detector was approximated using a gaussian for the energy of the final state objects (photons and b partons). No smearing is applied to the azimuthal angle ϕ and the pseudorapidity η . This approximation is commonly made in the MEM analysis [19]. We have applied a gaussian smearing of 1% on the energy of the generated photons and of 10% on the energy of the generated b quarks. We implemented this smearing in the conversion from the LHE format to the LHCO format, that MadWeight can read. No b-tagging efficiency, photon efficiency and fake-rate were implemented.

The response of the detector can later be refined to match the actual response of the ATLAS detector, with more realistic smearing, efficiencies and fake-rates. This was not the goal of this proof of principle of the MEM analysis for measuring λ . As explained in appendix B, at least 3 smearing functions need to be implemented for the integration not to have dimension 0.

5.3 Pre-selection requirements

As the computation of the likelihood in the MEM is very CPU consuming, there is no point in computing likelihoods for events that are at the same time very background-like and in a region of phase space that does not add a lot of sensitivity to the extraction of λ from the signal events. Therefore we have applied a set of pre-selection requirements to the background and signal samples that were generated. These requirements are the same as in Ref. [10] plus the requirements $123 < m_{\gamma\gamma} < 127$ GeV on the di-photon invariant mass. Specifically, the pre-selection requirements are:

- $\Delta R_{\gamma b} > 0.2$
- $p_T(\text{subleading } \gamma) > 30$ GeV
- $p_T(\text{leading } \gamma) > 43$ GeV
- $|\eta_\gamma| < 2.37$, and $|\eta_\gamma| < 1.37$ or $1.52 < |\eta_\gamma|$ (electromagnetic barrel-end cap crack region)
- $p_T(b \text{ jet}) > 35$ GeV
- $|\eta_b| < 2.5$
- $123 < m_{\gamma\gamma} < 127$ GeV

Table 5.1 shows the expected number of events after pre-selection criteria are applied for signal and main background processes. The numbers in this table have been taken from Ref. [10] (before the $m_{\gamma\gamma}$ requirement) and [45] (after the $m_{\gamma\gamma}$ requirement). Table 5.1 includes detector efficiencies, whereas our simulation does not.

Process	Events after preselection (not including $m_{\gamma\gamma}$)	Events after preselection (including $m_{\gamma\gamma}$)
$H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma), \kappa = 1$	20	16
Non-resonant $b\bar{b}\gamma\gamma$	10 100	326
Non-resonant $c\bar{c}\gamma\gamma$	630	21
Non-resonant $j\bar{j}\gamma\gamma$	690	22
Non-resonant $b\bar{b}j\gamma$	14 000	273
Non-resonant $c\bar{c}j\gamma$	480	10
Non-resonant $b\bar{b}jj$	3 600	77
$Z(\rightarrow b\bar{b})\gamma\gamma$	230	7
Total continuum background	29 730	736
$b\bar{b}H(\rightarrow \gamma\gamma)$	4	3
$Z(b\bar{b})H(\rightarrow \gamma\gamma)$	26	17
$ggH(\rightarrow \gamma\gamma)$	28	21
$ttH(\rightarrow \gamma\gamma)$	124	84
Total single Higgs boson background	182	125
Total backgrounds	29 912	861

Table 5.1: Expected number of events after pre-selection criteria are applied for signal and background processes. Figures in this table are normalised by cross-section and a luminosity of 3000 fb^{-1} . ATLAS detector efficiencies and fake-rates are taken into account.

The distributions of the main kinematic variables for signal events are presented in Figs. 5.1-5.3 after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green). The distribution of the main kinematic variables for background events are presented in Appendix C.

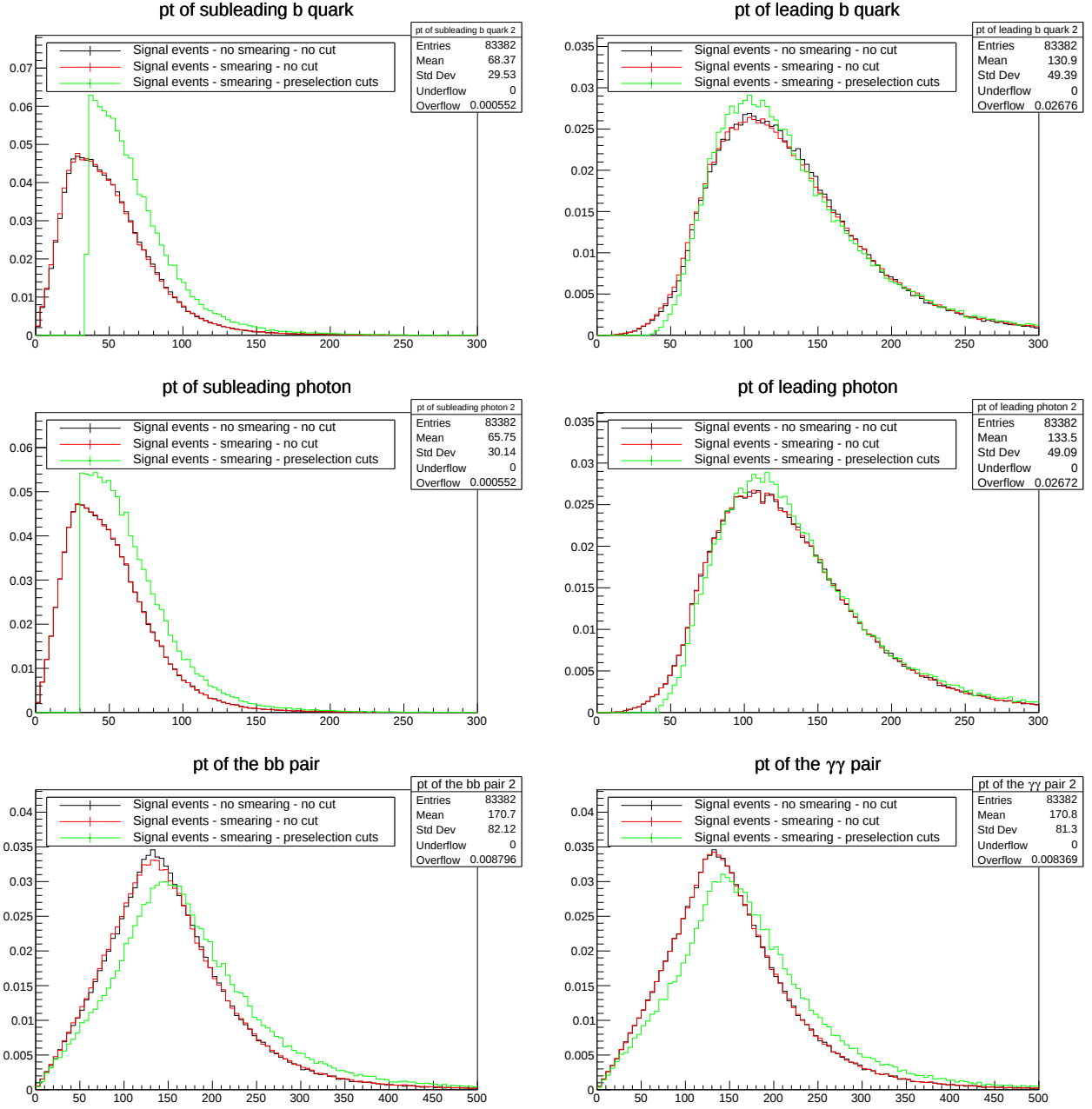


Figure 5.1: Distributions of the main kinematic variables for signal events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

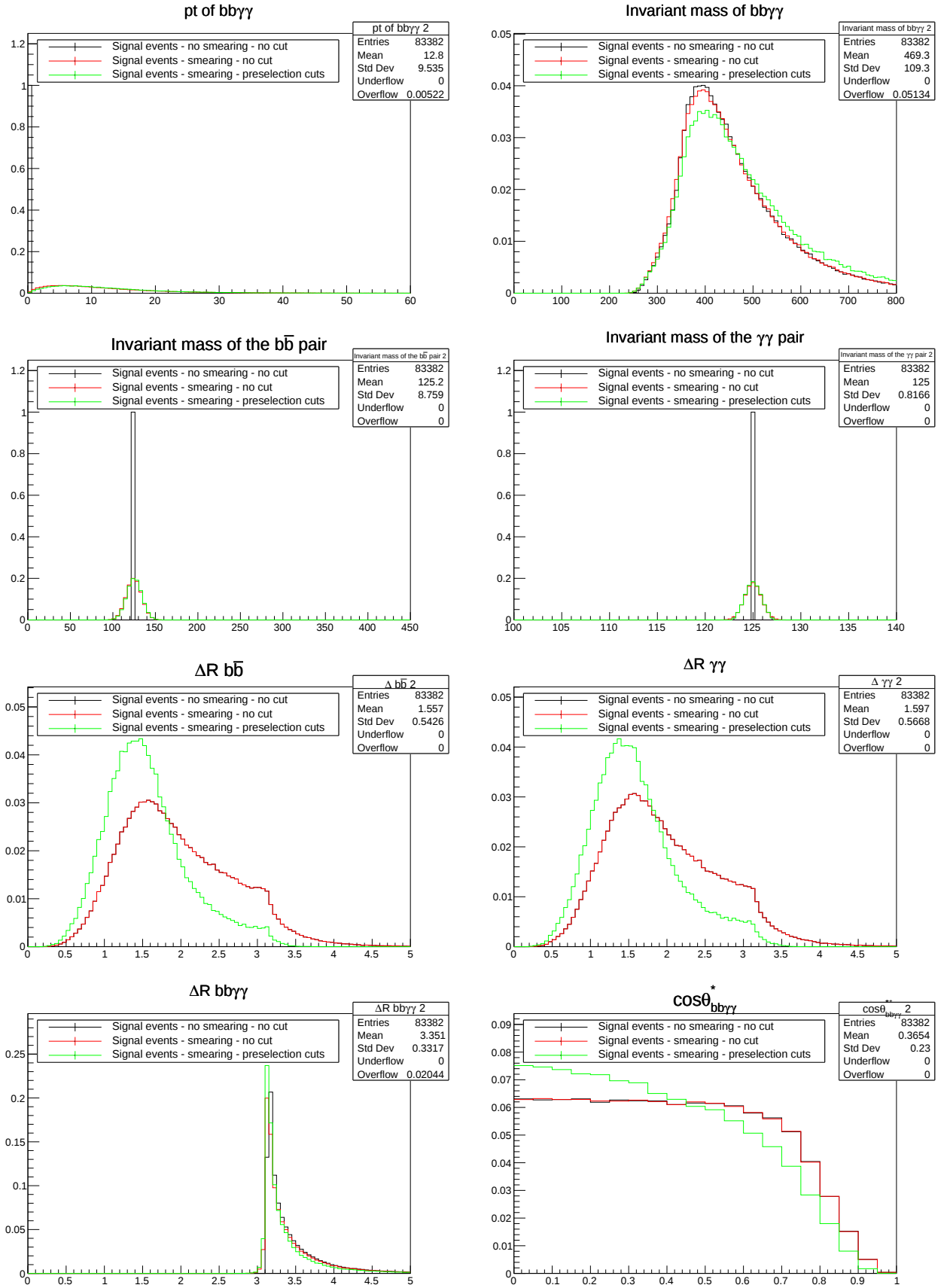


Figure 5.2: Distributions of the main kinematic variables for signal events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

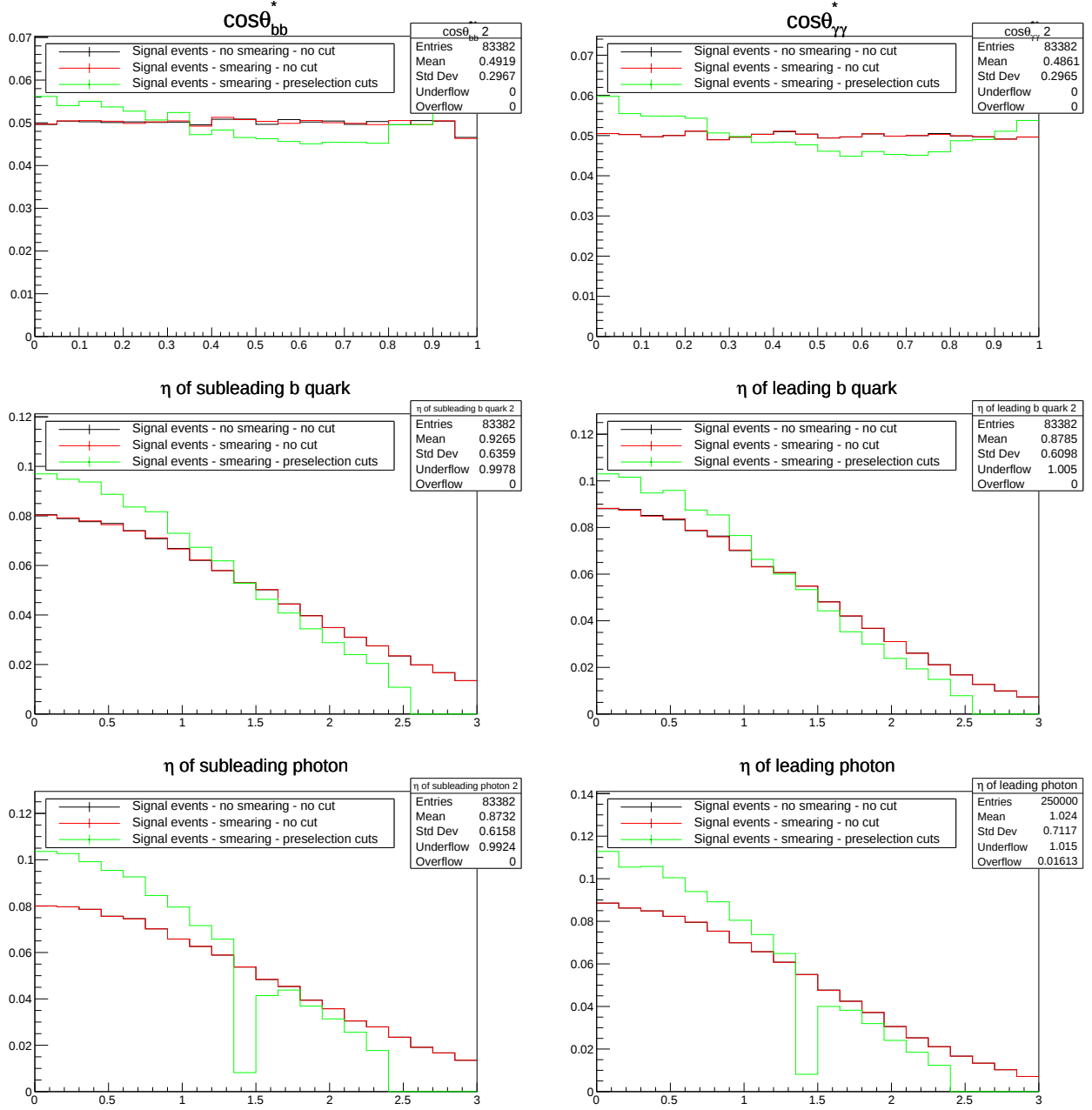


Figure 5.3: Distributions of the main kinematic variables for signal events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

Chapter 6

MadWeight: an automatic tool for the computation of the likelihoods

6.1 Introduction to MadWeight

MadWeight [20, 21] is a code that is integrated into the MadGraph5_aMC@NLO framework¹ [22]. MadWeight is dedicated to the computation of the per-event likelihood given a certain model and set of parameter values. It is a general implementation of the MEM: given a sample of events and different theoretical models or different values for the model parameters, MadWeight provides the necessary information for determining which model is the most likely to have produced these events. The likelihood is computed based on the method explained in Chapter 3 and the multi-dimensional integration is carried out by the VEGAS algorithm [35] (Fortran implementation) with default parameters. For the user, the procedure to compute the likelihood with MadWeight is quite straightforward and can be decomposed in the following steps:

Step 1: Define the model

The user can specify which model to use for the computation of the likelihoods. Models are defined in MG5_aMC/models. We have used our model SM_EFT_FF_bt_haa.kappa, described in Sec. 4.3.

Step 2: Define the process

The process for which the user would like to calculate the process likelihood L_{process}^p must be defined and can include decay chains. For the signal Matrix Element, the process was set to the di-Higgs production and decay into a $b\bar{b}$ and a $\gamma\gamma$ pairs. For the $b\bar{b}\gamma\gamma$ continuum background Matrix Element, it is set to the production of $b\bar{b}\gamma\gamma$ from colliding partons. The corresponding command lines in the MadGraph interface for the signal process are:

```
MG5_aMC> generate g g > h h , h > b b~, h > a a
```

and for the background:

```
MG5_aMC> define p = g u c d s b u~ c~ d~ s~ b~  
MG5_aMC> define x = b b~  
MG5_aMC> generate p p > x x a a
```

¹Version 2.3.3 released on 25/10/2015 was used.

Step 3: Define transfer functions

The transfer functions, which represent the resolution function used to describe the detector response, must be set. These transfer functions used in this study match the very simplified detector response that was simulated, cf. Sec. 5.2. User-defined transfer functions can be implemented in `MG5_aMC/Template/MadWeight/transfer_function/data`.

Step 4: Define hypothesis value for model parameter(s)

Hypotheses for the unknown parameters to be tested by the MEM have to be set. In the analysis, we have run MadWeight for values of κ ranging from -3 to +10 by steps of 0.5 (except for 0 for which an error occurs and which is replaced by 0.2 instead). This can be set in the template MadWeight card `MG5_aMC/Template/MadWeight/Cards/MadWeight_card.dat` or in the command line interface with (to set κ to 2 for instance):

```
MG5_aMC> set MW_parameter 13 2
```

Step 5: Define run information

Run information such as beam energy, choice of Parton Distribution Function (PDF), renormalisation and factorisation scales have to be set. They can be set by editing `MG5_aMC/Template/MadWeight/Cards/run_card.dat`.

In this study, the PDF used for the computation is `NNPDF23_lo_as_0130_qed`. It is the same as the one used for event generation.

Step 6: Define integration parameters

VEGAS integration parameters can be set by editing `MG5_aMC/Template/MadWeight/Cards/MadWeight_card.dat`. We have kept the default values:

number of points (by permutation) in MadWeight integration for survey: `MW_int_points=2000`

number of points (by permutation) in MadWeight integration for refine: `MW_int_refine=10000`

ratio between estimated integral value and estimated error on the integral: `precision=0.005`

6.2 Issues when running MadWeight on HH events: setup and code modifications

When first running MadWeight on HH events, we encountered some technical issues that we present here.

No MET in LHCO file

MadWeight expects to read a LHCO file always having a line with type 6, corresponding to MET [46]. MET is not used in the calculation of the per-event likelihood for the $b\bar{b}\gamma\gamma$ final state. We changed MadWeight code to by-pass this check and set all variables relative to MET to 0.

Integration dimension is 0

We first used MadWeight with the ready-to-use transfer function single gaussian, corresponding to a transfer function on photons energy only (not on b quarks energy). By doing so, the dimension of the likelihood integrand was 0, as explained in Appendix B. We therefore added a transfer function on the energy of the b quarks to raise the dimension of the integral to 2.

Parallel run mode

Running MadWeight in parallel mode resulted in occasional errors which we interpreted to come from undesirable delays resulting in an incorrect communication between the different parallel jobs. Parallel run mode was deactivated by setting `run_mode` to 0 in `MG5_aMC/input/mg5_configuration.txt`. Instead of running MadWeight in parallel (multi-core) mode, we have submitted large series of single-core batch jobs.

Other modifications (not to correct for running issues)

HTML opening was deactivated by setting `automatic_html_opening` to `False` in `MG5_aMC/input/mg5_configuration.txt`.

The MadWeight card in `MG5_aMC/Template/MadWeight/Cards/MadWeight_card.dat` was changed to set line with tag 11 to `sminputs` and line with tag 12 to 5 (to match with the definition of the parameter κ we introduced in the model `SM_EFT_FF_bt_haa_kappa`).

The energy of the two beams was set to 7000 GeV in the run card `MG5_aMC/Template/MadWeight/Cards/run_card.dat`.

The mass of the Higgs boson was set to 125 GeV in `MG5_aMC/models/SM_EFT_FF_bt_haa_kappa/parameters.py`. Events were generated with this value for the Higgs boson mass.

6.3 Matrix Element Analysis of pure signal ensembles

The pre-selected signal events generated at $\kappa=1$ (see Chapter 5) were analysed using MadWeight with the model `SM_EFT_FF_bt_haa_kappa`, the signal Matrix Element and different hypotheses for κ , as described in Sec. 6.1. The likelihoods of:

- 50 signal events were computed for $\kappa=-3, -2.5, -2, -1.5, -1, -0.5, 0.5, 1.5, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10$
- 2000 signal events were computed for $\kappa=0.2, 2$
- 12000 signal events were computed for $\kappa=1$

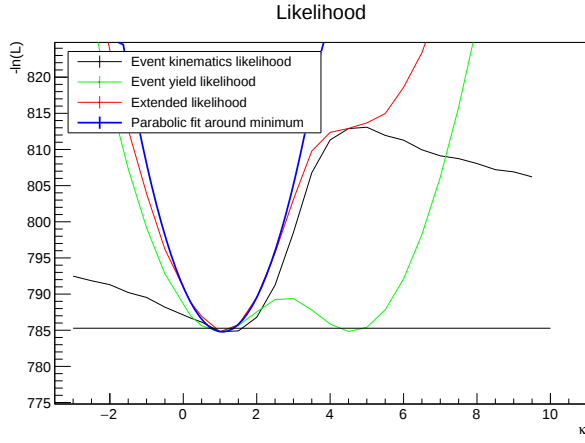
This allows us to perform the following two studies: scans of the likelihood as a function of κ for a given ensemble of simulated events and toy studies using large numbers of simulated events to quantify the precision of the measurement of κ .

Likelihood as a function of κ

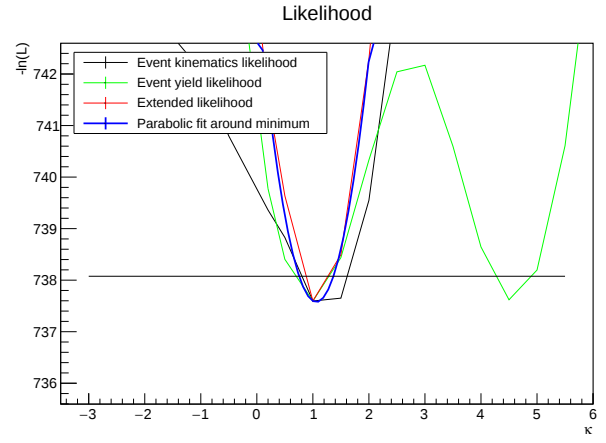
The likelihood of a sample of 16 signal events² was computed for all values of κ and the plot of the likelihood as a function of κ is drawn³. The corresponding graph is the black curve of Fig. 6.1, referred to as event kinematics likelihood since it is based on the Matrix Element Method. As expected the minimum of the likelihood is obtained for a value of κ very close to 1. The likelihood corresponding to a counting analysis, where only the event yield is used to extract λ , is represented in green. One interesting feature of the event kinematics likelihood is that it strongly rejects the region around $\kappa \approx 5$ that is allowed by the counting analysis. The two likelihoods are statistically independent and can be combined into an extended likelihood (red curve) that takes into account both event kinematics and event yield. The extended likelihood is then fitted using a parabola and just three values of κ close to the minimum, specifically

²16 events is the expected yield at $\kappa = 1$ after pre-selection, see Tab. 5.1.

³7 events were removed from the sample of 16 events and replaced by 7 other "non problematic" events due to numerical issues. This issue will be addressed later in Sec. 7.5 and with more details in Appendix F.



(a) Large view on the likelihoods.



(b) Zoom on the region close to the minimum.

Figure 6.1: Negative logarithm of the event yield likelihood (green), the event kinematics likelihood from MadWeight (black) and the extended likelihood (red) for a sample of 16 signal events generated at $\kappa = 1$, and parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$ (blue). An arbitrary offset has been applied to all likelihood curves such that they are on the same scale. The horizontal black line corresponds to 0.5 above the minimum of the parabola; its intersection with the fitting parabola gives the statistical uncertainty on the measurement of the minimum likelihood.

$\kappa = 0.2, 1$ and 2 . The measured value of κ as well as the corresponding uncertainty in the measurement of κ are determined from the fitted parabola. For this sample, $\kappa = 1.07 \pm 0.30$.

Ensemble test

By repeating this procedure for a collection of samples with a random number of events following a Poisson distribution with a mean yield of 16 events, the measured value of κ , denoted κ_{meas} , and the corresponding uncertainty κ_{unc} are obtained for each sample. For each sample the pull value w , defined as

$$w = \frac{\kappa_{\text{meas}} - \kappa_{\text{true}}}{\kappa_{\text{unc}}}$$

where κ_{true} is the value of κ at which events were generated, *i.e.* 1, is computed. The histograms of κ_{meas} , κ_{unc} and the pull w are given in Fig. 6.2.

If the implementation of the analysis is correct, then a normal distribution with average 0 and standard deviation 1 is expected for the pull. This is what we observe here. Therefore it seems that the Matrix Element Method applied to signal events only and using likelihoods computed with MadWeight provides, within the available statistics, an unbiased determination of κ .

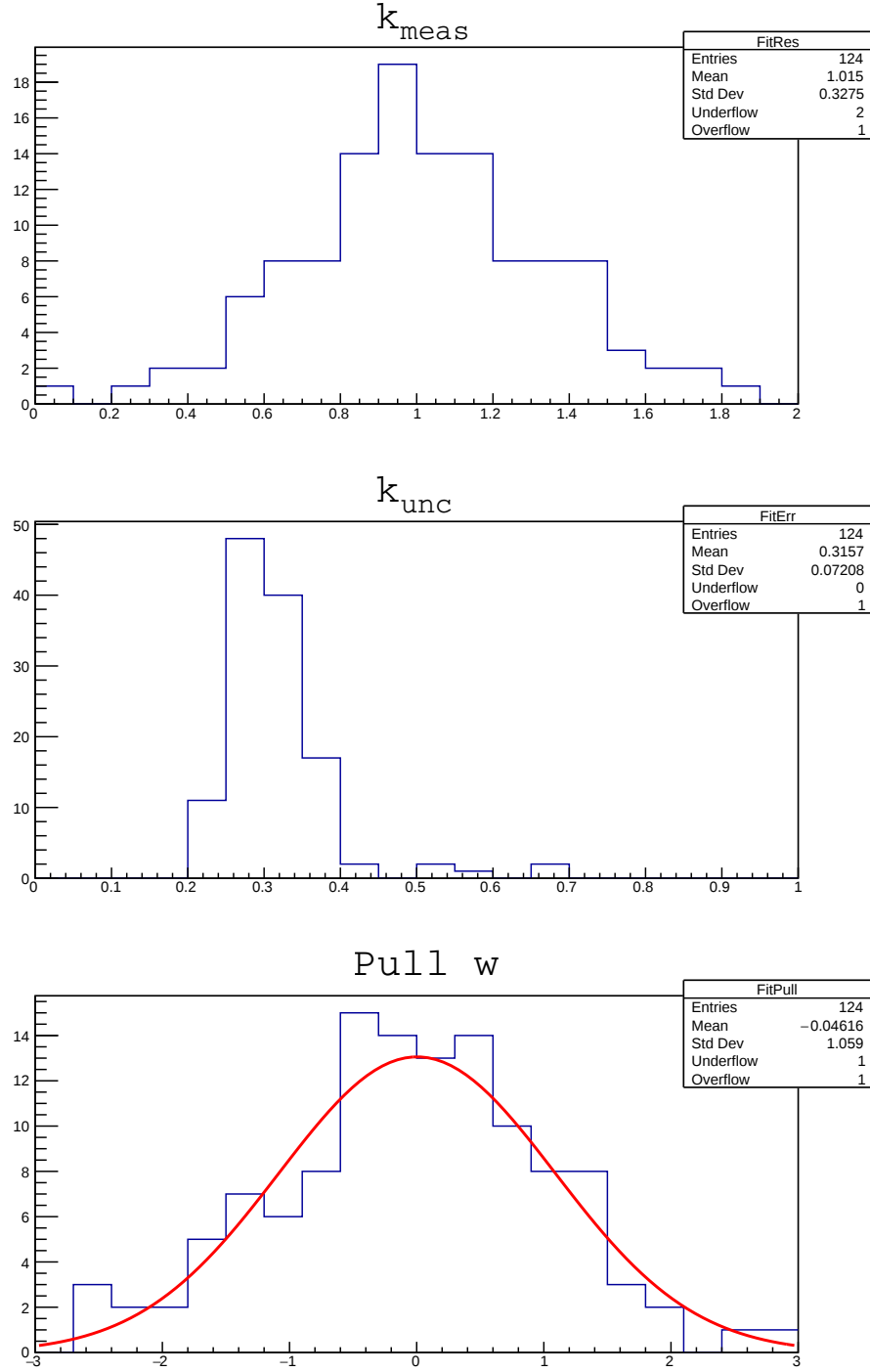


Figure 6.2: Histograms of the measured value of κ (κ_{meas} , top), of the corresponding uncertainty (κ_{unc} , middle) and of the pull value (w , bottom) for signal events only, analysed with MadWeight using the signal Matrix Element, obtained by a parabolic fit on the extended likelihood.

6.4 Matrix Element Analysis of mixed signal and non-resonant $b\bar{b}\gamma\gamma$ background ensembles

The next step in the development of the analysis is to add background events to simulated samples. In this section, we will only consider the $b\bar{b}\gamma\gamma$ continuum background and we will

refer to it as background. The addition of background to the analysis requires the computation of three new types of jobs: the likelihood for the signal hypothesis of background events, and the likelihood for the background hypothesis of signal and background events.

The pre-selected signal events generated at $\kappa = 1$ were analysed using MadWeight with the model `SM_EFT_FF_bt_haa_kappa` and the background Matrix Element. The likelihoods of 12000 events were computed. The generated and pre-selected background events (see Chapter 5) were analysed using MadWeight with the model `SM_EFT_FF_bt_haa_kappa`, the signal Matrix Element and different values of κ , as described in Sec. 6.1. The likelihoods of 15000 background events were computed for $\kappa=1$. These events were also analysed using MadWeight with the model `SM_EFT_FF_bt_haa_kappa` and the background Matrix Element. The likelihoods of 15000 events were computed.

The likelihood ratio (LHR) is the most powerful variable to discriminate between two hypotheses. As a result the likelihood ratio $L_{\text{signal}}/L_{\text{bkg}}$ should be the most powerful variable to classify signal and background events. In order to check this expectation, we have compared the rejection power based on the likelihood ratio to the one obtained with the most powerful cut-based analysis up to now, described in Ref. [10]. This analysis uses a BDT to classify signal and background events. The BDT is trained using fully simulated and reconstructed events to reject a cocktail of backgrounds (from different background processes). A direct comparison between the performance of this BDT and our LHR is thus not perfectly relevant. Therefore we have trained a BDT exactly as it is done in the analysis [10] (same BDT configuration, same training algorithm, same input variables except those irrelevant to our simplified study such as the number of initial electrons), but on the events we have generated, and to reject the $b\bar{b}\gamma\gamma$ continuum background only. The rejection powers of the two methods are presented in Fig. 6.3. The BDT clearly surpasses the LHR, which is not normal.

In order to further investigate this issue, we made a scatter plot of the LHR versus the BDT score, Fig. 6.3. It was found that some background events were strongly misclassified by the LHR (encircled in red). Plots of the distributions of the most discriminant kinematic variables (see Figs. 6.4 and 6.5) for the signal and background events, with and without cut on the LHR, revealed that these background events were indeed not correctly classified by the LHR: it is expected that the background events selected by a cut on the LHR should have kinematic features very close to the signal kinematics. However for most kinematic variables, the distribution of the background events returned by the LHR cut does not tend towards the signal distribution at all. As a control plot, the kinematic distributions of the signal events retained by the LHR cut were plotted and these distributions tend to be more signal-like (the selected signal events are further away from the background distribution than the full set of generated signal events), as expected.

Therefore it seems that something goes wrong in the computation of the likelihoods. We have searched for possible errors in the definition of the transfer functions, in the smearing and in our control plots, but we did not find any. This issue has been reported to the MadWeight authors and they agree that these results are problematic [47] and believe that this might be due to the Single-Diagram-Enhanced multi-channel integration technique. Information about this integration technique can be found in Ref. [48]. We have attempted to track down the cause of this issue in the MadWeight code. This turned out to be a difficult task. Even understanding the exact choice for the change of variables that is automatically done by MadWeight [20, 21],

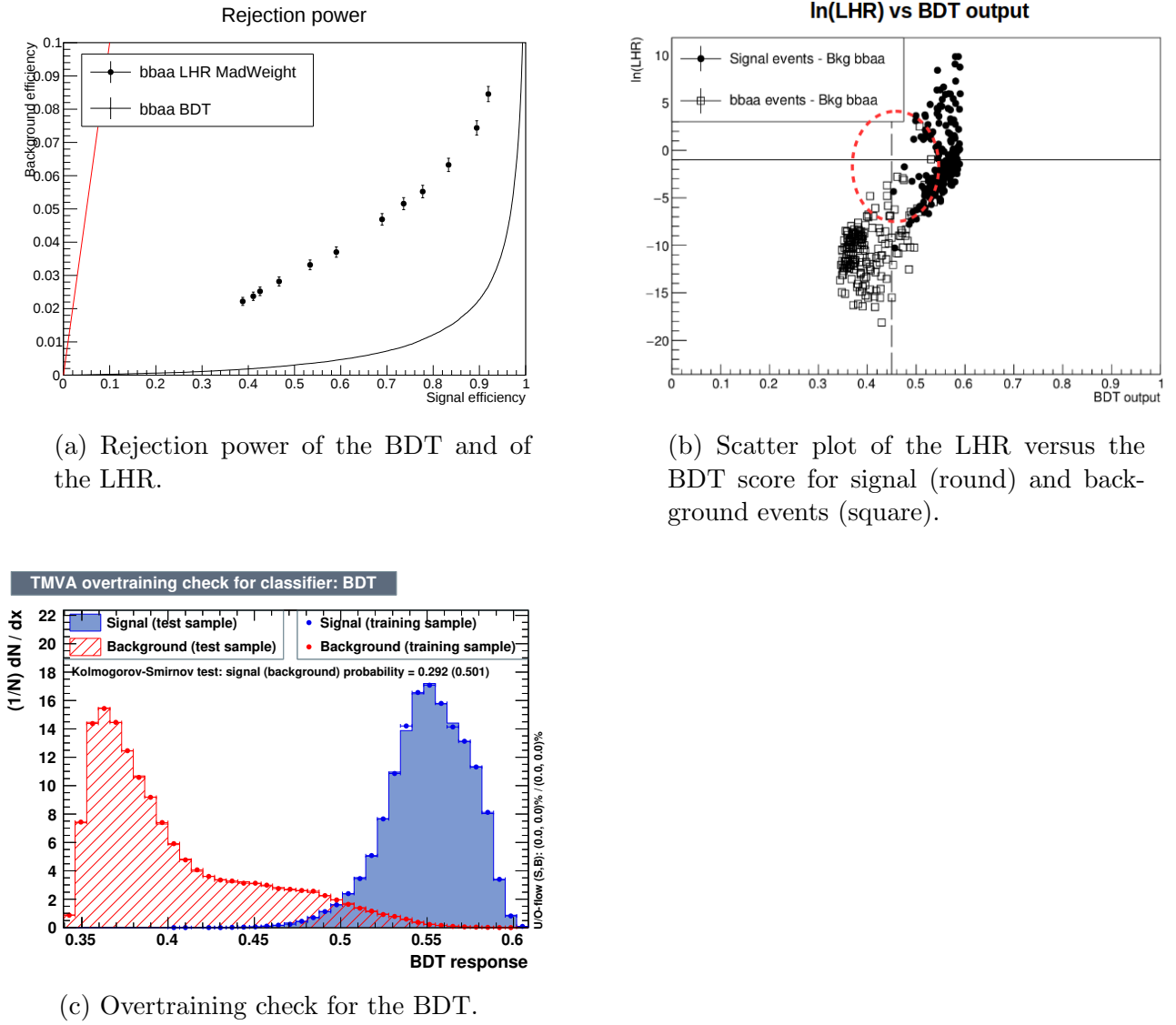


Figure 6.3: Top left: Rejection power of a BDT trained as in Ref. [10] and of the LHR from MadWeight to classify signal events and $b\bar{b}\gamma\gamma$ continuum background events. Top right: Scatter plot of the LHR versus the BDT score for signal (round) and background events (square). Bottom: Overtraining check for the BDT.

was a demanding task. We found out that the Main Block used was the Main Block A, but did not find which particles energy was transformed. Instead of pursuing this investigation, we have decided to test the new MoMEMta tool which is designed to be a modular toolkit to compute the likelihoods, with a higher level of control and the possibility to try out other multi-dimensional integration algorithms. MoMEMta reuses many of the key concepts that have been developed for MadWeight, including the available variable transformations (main and secondary blocks). The computation of the likelihoods using MoMEMta is the subject of the next chapter.

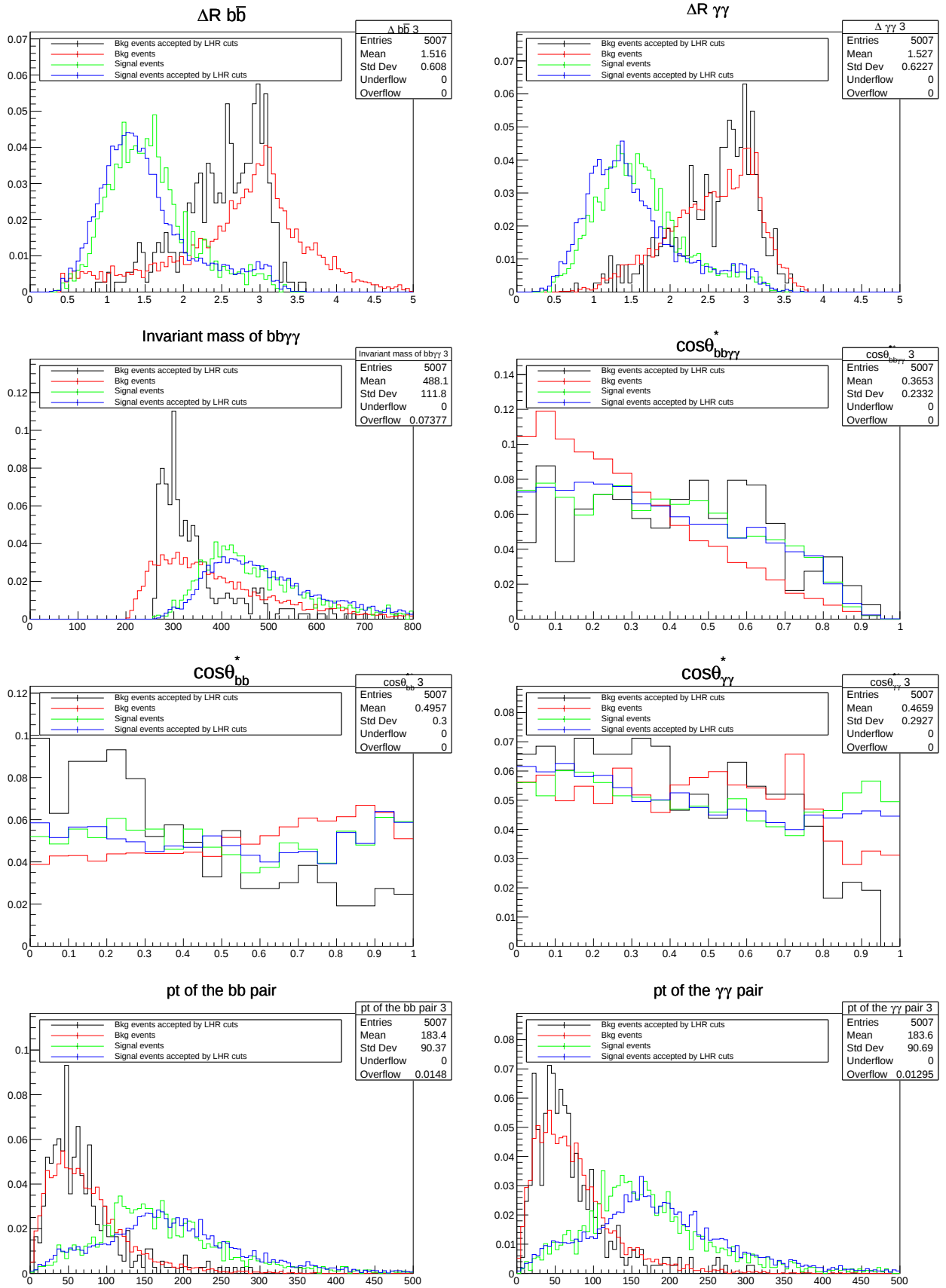


Figure 6.4: Distributions of main kinematic variables for the signal and $b\bar{b}\gamma\gamma$ background events, with and without cut on the LHR. The efficiency of the LHR cut is 42% for signal events and 2.5% for background events.

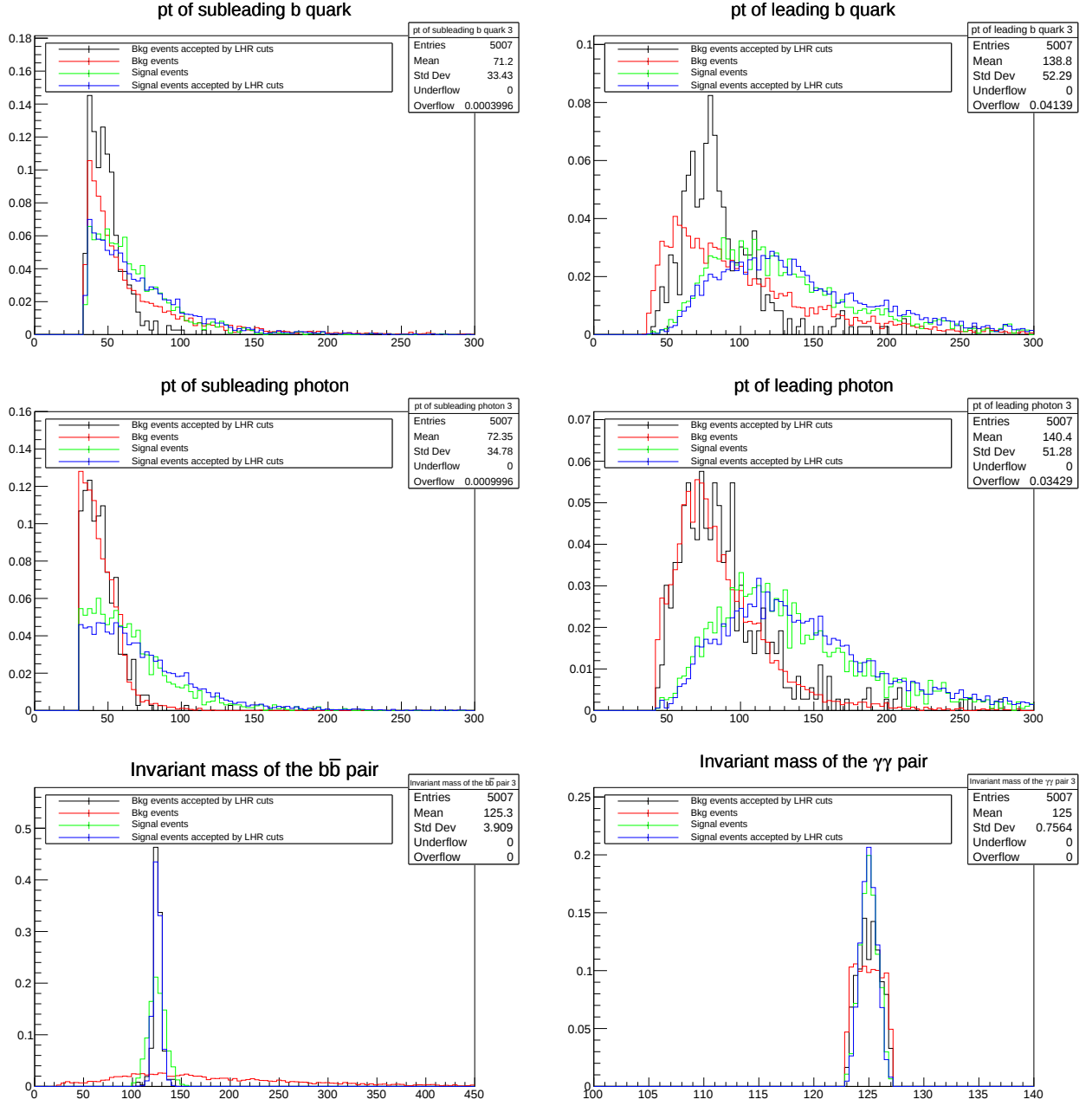


Figure 6.5: Distributions of main kinematic variables for the signal and $b\bar{b}\gamma\gamma$ background events, with and without cut on the LHR. The efficiency of the LHR cut is 42% for signal events and 2.5% for background events.

Chapter 7

MoMEMta: a modular toolkit for the computation of the likelihoods

7.1 Introduction to MoMEMta

Similarly to MadWeight, MoMEMta [23] is a software package dedicated to the computation of the per-event likelihood given a certain model and set of parameters. We have used the version 1.0.0, released on 22/05/2018. MoMEMta is a C++ framework designed in a modular way in order for the user to configure the MEM computation at all levels. Unlike MadWeight which is fully automatic, MoMEMta allows the user to configure the different steps in the computation of the likelihood via a configuration file written in the Lua scripting language. This modularity is achieved by linking modules together, the output of a module being an input of another module, in order to define the MEM integrand. Each module corresponds to an elementary operation in the MEM computation: transfer functions, change of variables for the phase-space parametrisation, Matrix Element, etc. MoMEMta comes with a set of useful modules, but the user can define custom modules. The multi-dimensional integration is carried out by one of the algorithms available in the CUBA library [49] (Vegas, Suave, Cuhre, Divonne). The user can freely choose one algorithm. Figure 7.1 compares the two different approaches of MadWeight and MoMEMta. The different tints of grey roughly indicate how transparent the computation is for the user: the lighter, the more transparent.

For the user, the procedure to compute the likelihood with MoMEMta is not as straightforward as with MadWeight, since the user must take care of the Lua file. It can be decomposed in the following steps:

Step 1: Export the Matrix Element from MadWeight to MoMEMta

The Matrix Element of the process must be available in a suitable format for use in MoMEMta. For this purpose, a MoMEMta - MadGraph Matrix Element Exporter (MoMEMta-MaGMEE) is provided to convert the Fortran implementation of the Matrix Element used in MadWeight to a suitable C++ code that can be used by MoMEMta. The shared library resulting from the compilation of the code can then be indicated in the Lua configuration file.

Step 2: Write a C++ script to run MoMEMta

All the computations are driven by a C++ file which takes care of the:

- **event**: four-vectors of the particles in the event are stored in MoMEMta objects.

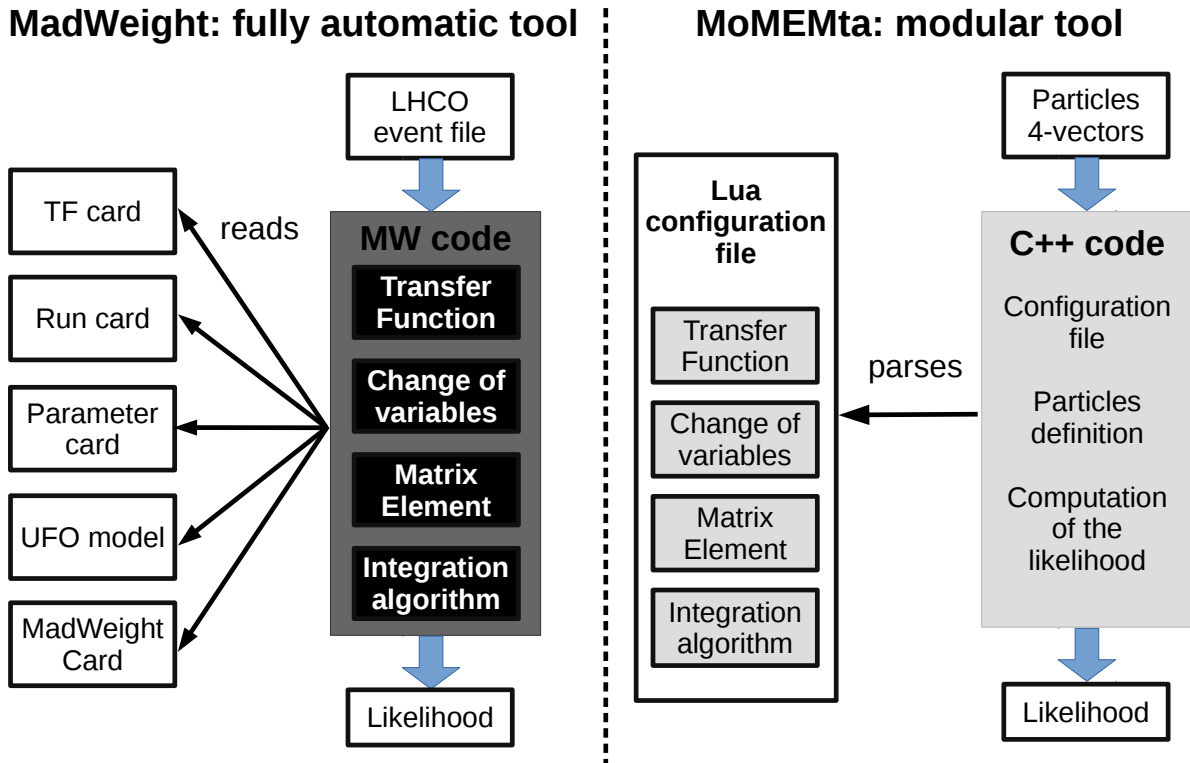


Figure 7.1: Comparison between MadWeight and MoMEMta working principles. The lighter the color, the more transparent the computation for the user.

- **MEM configuration:** the Lua configuration file to be parsed has to be indicated.
- **computation of the event likelihood:** the likelihood of the provided event is computed accordingly to the MEM configuration.

Step 3: Write the Lua configuration file

All the steps in the computation of the event likelihood have to be defined in the Lua file. This includes:

- Defining the final state particles.
- Defining the transfer functions.
- Defining the change of variables for the phase-space parametrisation (which Main Block and Secondary Block(s) to use [21]).
- Defining the jacobians resulting from the change of variables and the transfer functions.
- Defining the Matrix Element (that was created in step 1).
- Defining the integrand as the output of the last module.
- Choosing the multi-dimensional integration algorithm.

7.2 Technical issues encountered when running MoMEMta on HH events

Two major technical issues were encountered when first running MoMEMta and two other modifications were needed to match with the PDF used during the generation.

The first issue came from the fact that MoMEMta tagged our four-vectors as non-physical input, leading MoMEMta to stop before computing the likelihood. This issue was traced back to having different representations of the floating point numbers in MoMEMta – which uses the `ROOT::Math::LorentzVector<ROOT::Math::PxPyPzE4D<double>>` class – and our code – which uses the `TLorentzVector` class. As a result, computed from the energy and the three-momentum, the photon mass was nearly 0 but negative for MoMEMta, instead of nearly 0 but positive when using the `TLorentzVector` class. Using the `ROOT::Math::LorentzVector<ROOT::Math::PxPyPzE4D<double>>` type, we multiplied the momentum of the photons by the long double $(1 - 1/2^n)$, where n decreases from 53¹, until the mass of the photon turns out to be positive. This was always achieved with $n = 53$, meaning that only the least significant bit needed to be changed.

The second issue was an error that occurs when calling the signal Matrix Element. This was traced back to the fact that the MoMEMta - MadGraph Matrix Element Exporter (MoMEMta-MaGMEE) was unable to fully export the signal Matrix Element, resulting in an incomplete signal Matrix Element. In the model `SM_EFT_FF_bt_haa_kappa`, the Lorentz structure in the definition of the ggH and $ggHH$ vertices, representing the spin tensors associated with the vertex, are set to external and are defined in the `UFO_DIR/Fortran` directory [50]. It seems that the MoMEMta - MadGraph Matrix Element Exporter does not take into account this way to define the spin tensors. Therefore the Fortran routines defined in `UFO_DIR/Fortran` were not encoded in the C++ code of the signal Matrix Element. In order to correct for this problem, we have translated by hand the Fortran routines to C++ and added them to the code of the signal Matrix Element. We checked this tedious and error-prone task by running the Matrix Element code of MadWeight and MoMEMta with exactly the same input event, the same random seed, the same PDF and the same parameters. The translation of the code from Fortran to C++ was successful as we obtained exactly the same outputs.

During the comparison between MadWeight and MoMEMta, we noted that the default PDF and the strong coupling constant α_S were not adequate in MoMEMta. The PDF used for event generation is `NNPDF23_lo_as_0130_qed`, and requires α_S to be equal to 0.130, whereas default PDF in MoMEMta is `CT10nlo`. Setting α_S automatically to 0.130 when using `NNPDF23_lo_as_0130_qed` is implemented in MadWeight, but is not implemented in MoMEMta. We therefore set the Lua configuration file to choose the PDF `NNPDF23_lo_as_0130_qed` and changed the value of α_S to 0.130 in the C++ code relative to the model parameters.

7.3 Lua configuration files

The Lua configuration file is the key feature in the computation of the likelihood with MoMEMta. It defines all the steps in the computation of the likelihood and allows the user to have full control and insight into the computation. The Lua configuration file consists of different modules that are linked together, the output of a module being an input of another module. Each module performs a specific operation in the definition of the likelihood integrand, such as transfer functions or change of variables, and the integrand is defined as the output of the final module. The Lua configuration files we have used are given in Appendix D.

¹53 is the size of the significand in bits for `long double` when GCC compiler is used with recent x86 CPUs.

Signal process

The Main Block A was used for the signal process. Information about the Main Block A can be found in Sec. 3.3 and Ref. [21]. We decided to integrate over the energy of one photon and one b -parton and compute the four-vectors of the other particles (the two gluons, the other photon and the other b -parton) by conservation of the energy-momentum. As a technical consequence, the evaluation of the transfer function is done differently for the photon and the b -parton depending on whether their energy is integrated out in the Main Block A or not. This is why we have used the module `GaussianTransferFunctionOnEnergyEvaluator` for the photon and the b -parton whose four-vectors are computed by energy-momentum conservation, and the module `GaussianTransferFunctionOnEnergy` for the other photon and the b -parton. As explained in Appendix B, the dimension of the likelihood integral is 2. The above procedure provides exactly 2 degrees of freedom (the energies of one photon and one b quark)². The Matrix Element we have used is exactly the same as the one used with MadWeight: $gg \rightarrow H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma)$. It has been imported from MadWeight with the MoMEMta - MadGraph Matrix Element Exporter (MoMEMta-MaGMEE) and by-hand completion (cf. technical issues detailed in Sec. 7.2). The likelihood was integrated with VEGAS (with default parameters: 25000 evaluations of the integrand at each iteration until a precision of 0.05 or 500000 evaluations of the integrand is reached) for signal events and $b\bar{b}\gamma\gamma$ background events, and DIVONNE (with default parameters) for $t\bar{t}H$ background events.

$b\bar{b}\gamma\gamma$ background

The Lua configuration file for the non-resonant $b\bar{b}\gamma\gamma$ background is the same as the one for the signal process, except for the Matrix Element and the capability to change κ . The Matrix Element we have used is exactly the same as the one used with MadWeight: $pp \rightarrow b\bar{b}\gamma\gamma$. It has been successfully imported from MadWeight with the MoMEMta - MadGraph Matrix Element Exporter (MoMEMta-MaGMEE). The likelihood was integrated with VEGAS (with default parameters: 25000 evaluations of the integrand at each iteration until a precision of 0.05 or 500000 evaluations of the integrand is reached) for signal events and $b\bar{b}\gamma\gamma$ background events, and DIVONNE (with default parameters) for $t\bar{t}H$ background events.

$t\bar{t}H$ background

The final-state produced in $t\bar{t}H$ events is more complex than $b\bar{b}\gamma\gamma$, but only the $b\bar{b}\gamma\gamma$ final-state objects are used as input to the MEM. As previously discussed in Sec. 5 the top quark decays into a b quark and a W boson. The W boson decays into a b quark with a negligible fraction and cannot decay into a photon. For this simplified study, we have therefore assumed that the b quarks do not originate from the decay of the W bosons. As a result, the observed b -partons must come from the decay of the top quark and not the weak boson.

$t\bar{t}H$ background events input to the MEM thus consist of the two b quarks coming from the decay of the top quarks and the two photons coming from the decay of the Higgs boson, the W bosons being discarded from the events.

The $t\bar{t}H$ background hypothesis is based on the process $pp \rightarrow H(\rightarrow \gamma\gamma)t(\rightarrow bW^+)\bar{t}(\rightarrow \bar{b}W^-)$. The corresponding Matrix Element was successfully imported from MadWeight with the MoMEMta - MadGraph Matrix Element Exporter (MoMEMta-MaGMEE). In the Lua file, the W bosons are treated as massive invisible particles with the Main Block F. Information about the Main Block F can be found in Sec. 3.3 and Ref. [21]. Important features of this Lua file

²It would probably be much more optimal to replace the energies of the photon and of the b quark by the mass of the Higgs bosons to exploit directly the Breit-Wigner of the Higgs boson. This can be achieved using the `SecondaryBlockCD` module and was not tried due to lack of time.

are: the introduction of a permutator module for the two b quarks, the four **GaussianTransferFunctionOnEnergy** modules taking care of the transfer functions on the energy of the b -partons and the photons, the **BreitWignerGenerator** module for the decay of the top quarks and the **BlockF** module. As explained in Appendix B, the dimension of the likelihood integral is 8. The above procedure provides exactly 8 degrees of freedom (the energy of the two photons and the two b quarks, the invariant mass of the two top quarks, and the energy of two colliding partons)^{3 4}. The likelihood was integrated with DIVONNE with default parameters.

7.4 Comparison between MadWeight and MoMEMta outputs

Computation of the likelihood with MoMEMta

As a first step, we were interested in comparing the event likelihood of the same set of events computed by MadWeight and MoMEMta. Therefore we have analysed with MoMEMta (VEGAS algorithm with default parameters) a subset of the events analysed with MadWeight and compared the event likelihood of the two subsets.

The pre-selected signal events generated at $\kappa=1$ (see Chapter 5) were analysed using MoMEMta with the signal hypothesis (*i.e.* the configuration file presented in Appendix D.1). The likelihoods of 40 and 2000 signal events were respectively computed for $\kappa=-3, -2.5, -2, -1.5, -1, -0.5, 0.5, 1.5, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10$, and $\kappa=0.2, 1, 2$.

These events were also analysed using MoMEMta with the background hypothesis (*i.e.* the background configuration file, which differs from the configuration file presented in Appendix D.1 by a few lines [Matrix Element and lines relative to varying κ]). The likelihoods of 2000 events were computed.

The generated and pre-selected background events (see Chapter 5) were analysed using MoMEMta with the signal hypothesis. The likelihoods of 400 and 2000 background events were respectively computed for $\kappa=-3, -2.5, -1.5, -0.5, 0.2, 0.5, 1.5, 2.5, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10$ and $\kappa=-2, -1, 1, 2, 3$.

These events were also analysed using MoMEMta with the background hypothesis. The likelihoods of 2000 events were computed.

Comparison between MadWeight and MoMEMta

The likelihood for the signal- $\kappa=1$ hypothesis and the likelihood for the background hypothesis of the same set of background and signal events computed using MadWeight and MoMEMta are compared in the $\ln L_{\text{MW}}$ versus $\ln L_{\text{MoMEMta}}$ scatter plots of Fig. 7.2. In each scatter plot, a dot corresponds to one event. 1000 or 2000 events are drawn in the scatter plots, depending on how many points were needed to clearly see patterns in the scatter plot.

The outputs of the two codes are often very different. For the background hypothesis – whether signal or background events are considered – the likelihood computed using MadWeight is al-

³MoMEMta considers that this Lua file corresponds to 9 degrees of freedom due to the permutation of the two b quarks, treated as an additional degree of freedom.

⁴It would probably be much more optimal to replace the energy of one of the two photons by the mass of the Higgs boson to exploit directly the Breit-Wigner of the Higgs boson. This can be achieved using the **SecondaryBlockCD** module and was not tried due to lack of time.

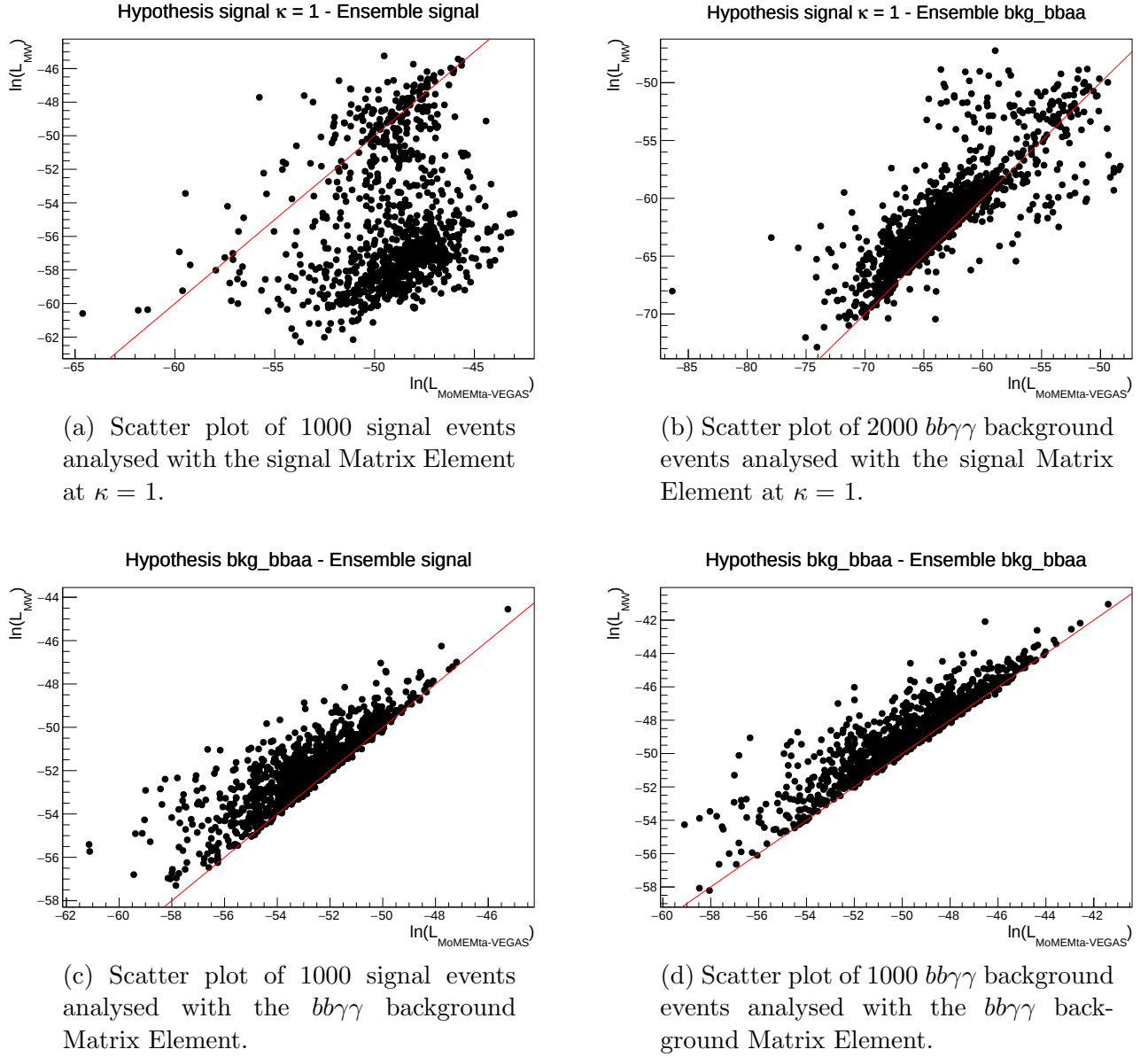


Figure 7.2: Comparison between MadWeight and MoMEMta outputs. The likelihood for the signal— $\kappa = 1$ hypothesis and the likelihood for the background hypothesis of the same set of background and signal events were computed using MadWeight and MoMEMta. The event likelihoods of the two codes are compared in a $\ln L_{MW}$ versus $\ln L_{MoMEMta}$ scatter plot. The red line is the $\ln L_{MW} = \ln L_{MoMEMta}$ line.

ways larger than the likelihood computed by MoMEMta and the events are located along the $\ln L_{MW} = \ln L_{MoMEMta}$ line. It is as if a probability was missing in the definition of the likelihood in MadWeight. A possible scenario would be a missing transfer function in this likelihood: for the events for which the smearing had changed the energy very little, the output of MadWeight and MoMEMta would be the same, but the events for which the smearing had changed the energy significantly, MoMEMta would give a significantly smaller likelihood than MadWeight. The origin of the differences has not been investigated but it should be possible to test several hypotheses easily ⁵.

⁵Given that the Lua file configures all the steps in the computation of the likelihood with MoMEMta, it is

Background events analysed with the signal hypothesis seem to be affected by the same problem as signal and background events analysed with the background hypothesis. But for some events, the likelihood computed using MadWeight is smaller than the likelihood computed by MoMEMta. Also it seems that there is a population of events at high L_{MW} and L_{MoMEMta} that are less concentrated along the $\ln L_{\text{MW}} = \ln L_{\text{MoMEMta}}$ line.

Among all cases, the signal events analysed using MoMEMta with the signal hypothesis show the largest deviations from MadWeight. It seems that there are two populations of events: a small one at high L_{MW} which is located on the $y = x$ line and another, much larger in terms of number of events, at low L_{MW} but higher L_{MoMEMta} that is located far from the $y = x$ line. This behaviour is very different from the three previous cases. The likelihood of signal events for the signal hypothesis computed with MoMEMta is therefore often a lot larger than the one computed with MadWeight.

The two effects described above should result in a significant improvement in the rejection power, as signal events are better classified as signal-like with MoMEMta than with MadWeight. This will be addressed in Sec. 7.6.

7.5 Matrix Element Analysis of pure signal ensembles

As detailed in Sec. 7.4, the likelihoods for the signal Matrix Element and several hypotheses for κ were computed for the signal events. This allows us to perform the following two studies: scan of the likelihoods as a function of κ for a given ensemble of simulated events and toy studies using large numbers of simulated events to quantify the precision of the measurement of κ .

Likelihood as a function of κ

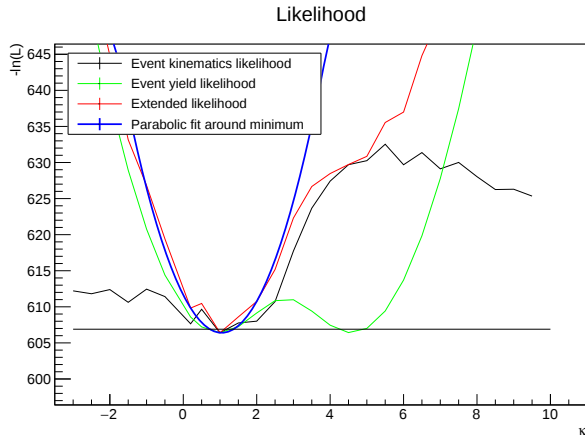
Based on Eq. 3.1, the likelihood of a sample of 16 signal events ⁶ was computed for all values of κ and the plot of the likelihood as a function of κ is drawn. The corresponding graph is the black curve of Fig. 7.3, referred to as event kinematics likelihood since it is based on the Matrix Element Method. The likelihood corresponding to a counting analysis, where only the event yield is used to extract λ , is represented in green (see Appendix E and Eq. E.1 for details). One interesting feature of the event kinematics likelihood is that it strongly rejects the region around $\kappa \approx 5$ that is allowed by the counting analysis. The two likelihoods are statistically independent and can be combined into an extended likelihood (red curve) that takes into account both event kinematics and event yield (see Appendix E and Eq. E.3 for details). The extended likelihood is then fitted using a parabola and just three values of κ close to the minimum, specifically $\kappa = 0.2, 1$ and 2 . The measured value of κ as well as the corresponding uncertainty in the measurement of κ are determined from the fitted parabola. For this sample, the estimation of κ is consistent with the value used for event generation ($\kappa = 1$): $\kappa = 1.29 \pm 0.33$.

Integration of the likelihood: convergence study

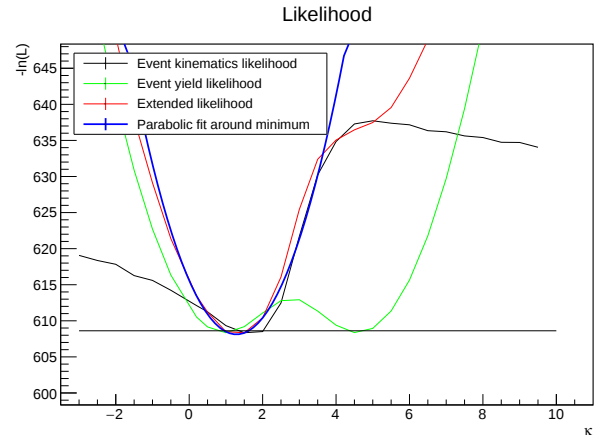
When first making the plots of the event kinematics likelihood, very fast variations or discon-

possible to voluntarily make mistakes in the configuration file, e.g. forgetting a transfer function, and to note what would be the effect of such a modification in the definition of the Matrix Element.

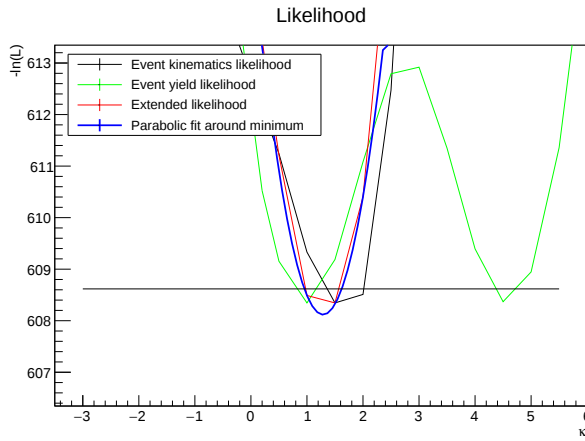
⁶16 events is the expected yield at $\kappa = 1$ after pre-selection, see Tab. 5.1.



(a) Likelihood scan for a set of 16 signal events without removing problematic events.



(b) Likelihood scan for a set of 16 signal events after removal of problematic events: 13 events out of the 16 events in Fig. 7.3a were kept; 3 problematic events were replaced by non-problematic ones.



(c) Zoom on the minimum region of the likelihood in Fig. 7.3b.

Figure 7.3: Negative logarithm of the event kinematics likelihood (black), the event yield likelihood (green) and the extended likelihood (red) as a function of κ for a set of 16 signal events without (a) or with (b,c) removing events with an event kinematics likelihood that presents fast variations (referred to as problematic events in the subcaptions). An arbitrary offset has been applied to all likelihood curves such that they are on the same scale. The parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$ is represented in blue. The horizontal black line corresponds to 0.5 above the minimum of the parabola; its intersection with the fitting parabola gives the statistical uncertainty on the measurement of the minimum likelihood.

tinuities were observed in the likelihood scans, while the likelihood scans were expected to be smooth. This was traced back to an imperfect integration of the likelihood for a subset of the events, and these problematic events were removed from the sample and replaced by unproblematic events. Figure 7.3a (resp. Figures 7.3b and 7.3c) shows the likelihood before (resp. after) this event substitution. Details about this issue are given in Appendix F.

Ensemble tests

By repeating the procedure of computing the extended likelihood for a collection of samples with a random number of events following a Poisson distribution with a mean yield of 16 events⁷, the measured value of κ , denoted κ_{meas} , and the measured uncertainty κ_{unc} are obtained for each sample. For each sample the pull value w , defined as

$$w = \frac{\kappa_{\text{meas}} - \kappa_{\text{true}}}{\kappa_{\text{unc}}}$$

where κ_{true} is the value of κ at which events were generated, *i.e.* 1, is computed. The histograms of κ_{meas} , κ_{unc} and the pull w are given in Fig. 7.4.

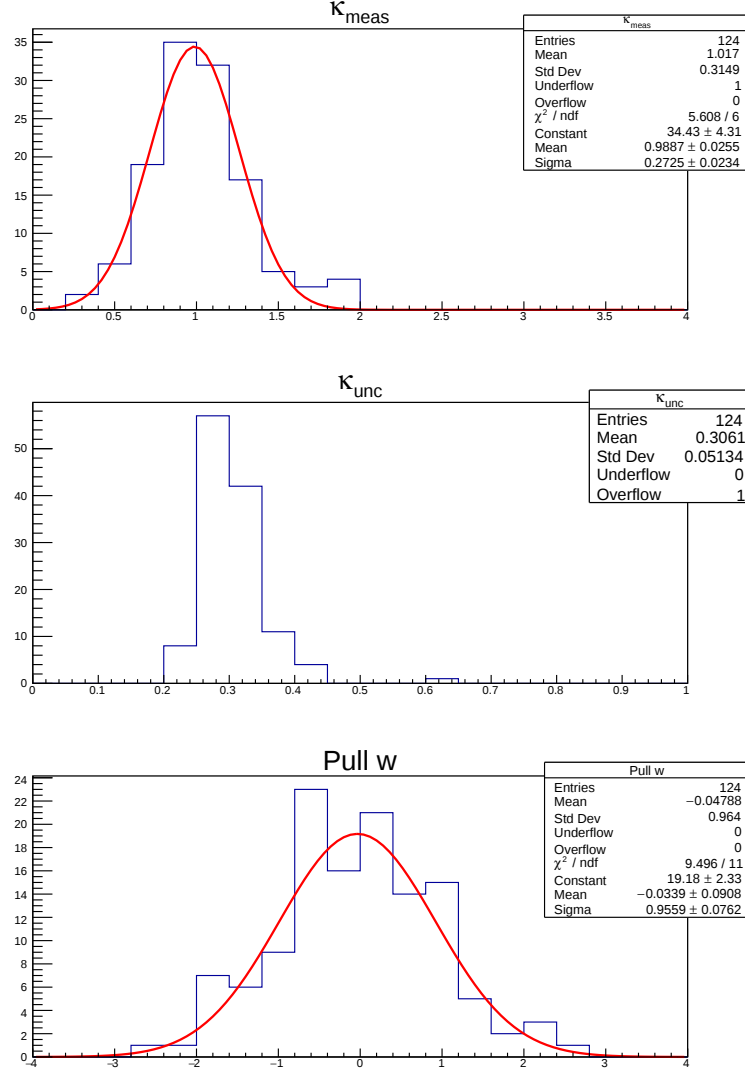


Figure 7.4: Histograms of the measured value of κ (κ_{meas} , top), of the corresponding statistical uncertainty (κ_{unc} , middle) and of the pull value (w , bottom) for signal events only, analysed with MoMEMta using the signal Matrix Element and the Lua configuration file presented in Appendix D.1.

⁷16 events is the expected yield at $\kappa = 1$ after pre-selection, see Tab. 5.1.

In this ensemble test, no removal of problematic events is used. These problematic events are expected to cause a small contribution to the width of the distribution of κ_{meas} . If the implementation of the analysis is correct, then a normal distribution with average 0 and standard deviation 1 is expected for the pull. This is what we observe here. Therefore it seems that the Matrix Element Method applied to signal events only and using likelihoods computed with MoMEMta provides, within the available statistics, an unbiased determination of κ .

These ensemble tests were compared with ensemble tests performed with a counting analysis. Two counting analyses were studied: a simple counting experiment and a counting by bin of the di-Higgs mass m_{HH} . Due to the shape of the di-Higgs production cross section as a function of κ (cf. Fig. 2.2) a counting experiment suffers from a two-fold ambiguity in the extraction of κ . For the purpose of this ensemble study, the solution at a lower value of κ (*i.e.* the correct solution), has always been plotted. The goal of this ensemble study is to illustrate the statistical power of the counting experiment in the absence of the two-fold ambiguity.

In the case of the simple counting analysis, one pseudo-experiment consists in generating just one random number, namely the event count, according to a Poisson distribution with a mean yield of 16 events⁸. The analysis of a pseudo-experiment consists in finding the value of κ corresponding to this event yield. By repeating this procedure for a large number of pseudo-experiments, the histogram of the measured value of κ can be drawn, see Fig. 7.5. The standard deviation on the measurements, *i.e.* the statistical uncertainty in the measurement, is 0.401 ± 0.003 . The spikes in the distribution are due to the fact that the event count in a given pseudo-experiment can only be an integer number.

In the case of the counting analysis in bins of m_{HH} , three bins were defined: $m_{HH} < 350$ GeV, $350 \text{ GeV} < m_{HH} < 480$ GeV and $m_{HH} > 480$ GeV. One pseudo-experiment consists in generating a random number for the signal event count in each bin according to a Poisson distribution (with a parameter dependent on the bin). In the analysis of a pseudo-experiment, the likelihood of observing the generated random number of events is computed for several values of κ (from -3 to 10 by step of 0.5) for each bin. The product of the likelihoods of the three bins is computed. The measured value of κ is obtained by fitting the likelihood product

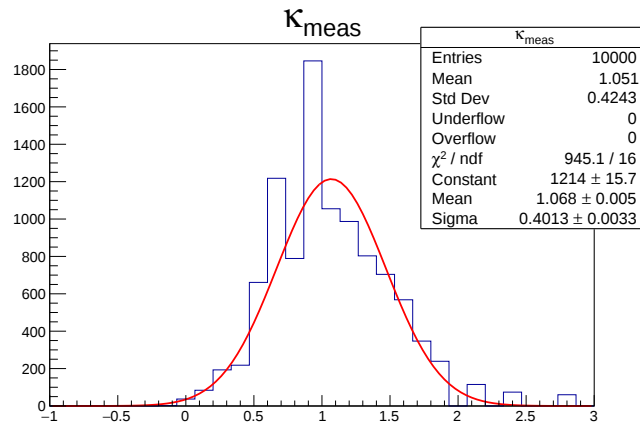


Figure 7.5: Histogram of the measured value of κ for a simple counting analysis.

⁸16 events is the expected yield at $\kappa = 1$ after pre-selection, see Tab. 5.1.

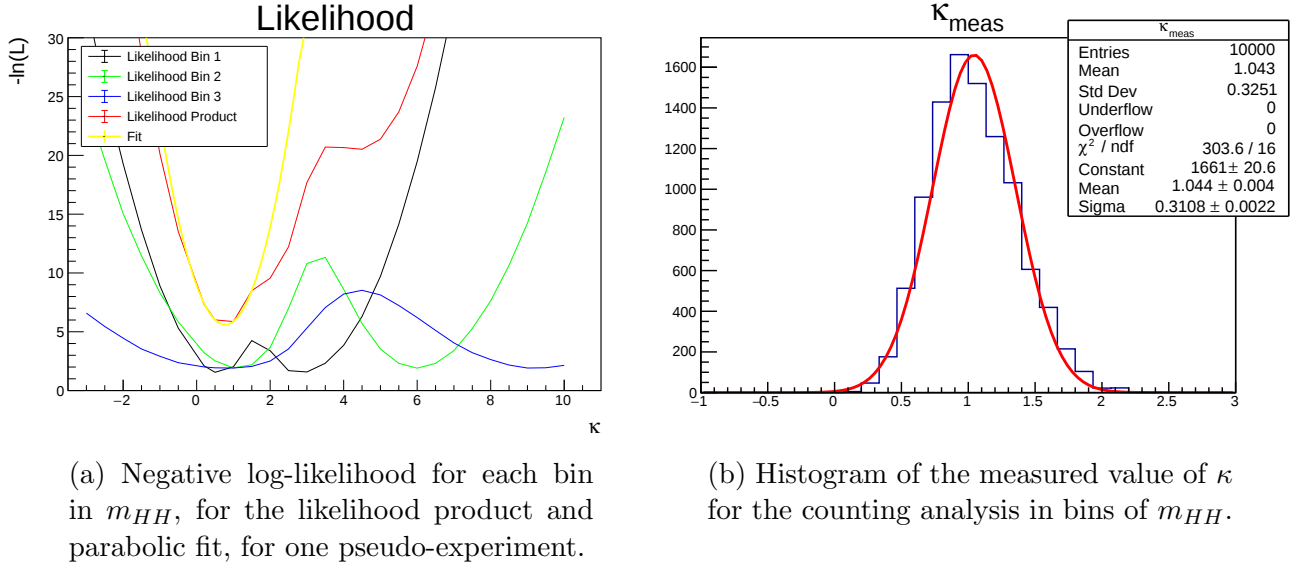


Figure 7.6: Negative log-likelihood and histogram of the measured value of κ for a counting analysis in bins of m_{HH} .

with a parabola in the minimum region (only points at $\kappa=0.5, 1$ and 1.5 are used). By repeating this procedure for a large number of pseudo-experiments, the histogram of the measured value of κ can be drawn, see Fig. 7.6. The standard deviation on the measurements is 0.311 ± 0.002 . This is much lower than in the simple counting analysis. This shows that the di-Higgs mass m_{HH} observable carries a large fraction of the kinematics information that allows to distinguish between different values of κ .

In the case of the MEM analysis (with extended likelihood), the standard deviation associated with the measurement of κ is 0.273 ± 0.023 (cf. Fig. 7.4). This is lower than the standard deviations obtained for the two counting experiments. Even though the combination of the event yield and the m_{HH} observable is a very good tool for the extraction of κ , significant improvements can be obtained via the full event kinematics with the MEM. The use of the MEM is expected to improve our ability to discriminate between signal and background. This will be quantified in the remainder of this chapter.

7.6 Matrix Element Analysis of a mixed signal and non-resonant $b\bar{b}\gamma\gamma$ background ensemble

In this section, only the non-resonant $b\bar{b}\gamma\gamma$ background is considered. Studying this background was motivated by the fact that it is the dominant background (see Tab. 5.1). As detailed in Sec. 7.4, the likelihoods of $b\bar{b}\gamma\gamma$ background events for the $b\bar{b}\gamma\gamma$ background Matrix Element and for the signal Matrix Element for several hypothesis for κ were computed.

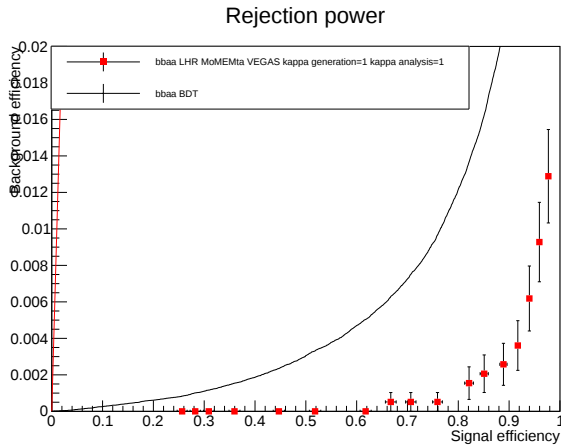
Integration of the likelihood: convergence study

As detailed in Appendix F, the likelihoods of some signal events analysed using the signal Matrix Element were not perfectly integrated: the integration stopped before a stable value was reached. When the evolution of the estimate of the likelihood integral as a function of

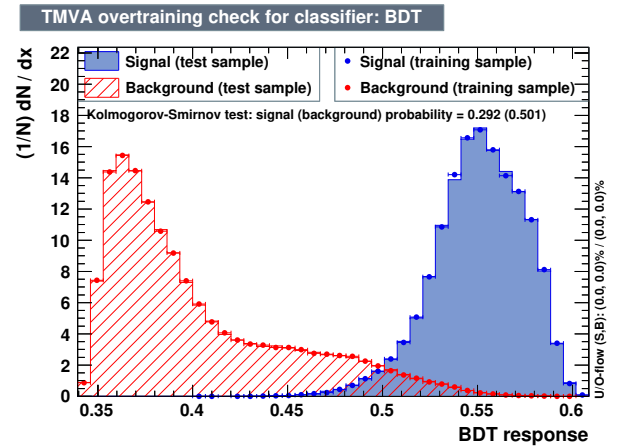
the number of iterations is very different for different κ hypotheses, the likelihood scan as a function of κ shows discontinuities. Such events are removed from the analysis. For other events concerned by this imperfect integration, all likelihood integrals for the different hypotheses are affected in a similar way. Such events do not introduce discontinuities in the likelihood scan as a function of κ . As discussed in Appendix G, their likelihood is underestimated. Approximately half of the signal events are affected. The same issue affected $b\bar{b}\gamma\gamma$ background events analysed with the signal hypothesis, albeit a smaller fraction of the events (approximately 10% of $b\bar{b}\gamma\gamma$ background events). A future resolution of this numerical issue is therefore expected to improve the plots given in this section, since the signal likelihood of signal events should be larger while (almost) all other likelihoods are correctly estimated.

Rejection power

The likelihood ratio $L_{\text{signal}}/L_{\text{bkg}}$ (LHR) is the most powerful variable to classify signal and background events. In order to check this expectation, we have compared the rejection power based on the likelihood ratio from MoMEMta to the rejection power obtained with the most powerful cut-based analysis up to now, described in Ref. [10]. This cut-based analysis uses a BDT to classify signal and background events. The BDT is trained using fully simulated and reconstructed events to reject a cocktail of backgrounds (from different background processes). A direct comparison between the performance of this BDT and our LHR is thus not perfectly relevant. Therefore a BDT was trained exactly as it is done in the analysis [10] (same BDT configuration, same training algorithm, same input variables except those irrelevant to our simplified study such as the number of initial electrons), but on the events we have generated and to reject the $b\bar{b}\gamma\gamma$ continuum background only. The rejection powers of the two methods are presented in Fig. 7.7. The rejection of the LHR from MoMEMta is clearly better than the BDT. The improvement with respect to MadWeight is consistent with the differences observed between MadWeight and MoMEMta in Sec. 7.4. This plot proves that the MEM is a very promising method for the di-Higgs analysis.



(a) Rejection power of the BDT and of the LHR.



(b) Overtraining check for the BDT.

Figure 7.7: Left: Rejection power of a BDT trained as in Ref. [10] and of the LHR from MoMEMta to classify signal events and $b\bar{b}\gamma\gamma$ continuum background events. Right: Overtraining check for the BDT.

Likelihood as a function of κ

326 non-resonant $b\bar{b}\gamma\gamma$ background events are expected after the pre-selection requirements are applied (see Tab. 5.1). The likelihoods of only 290⁹ non-resonant $b\bar{b}\gamma\gamma$ background events were computed for the signal hypothesis. The likelihood of a sample of 16 signal events and 290 background events can therefore be drawn. Based on Eqs. 3.1 and 3.2, the event likelihood and the sample likelihood are computed according to the following expressions:

$$L_{\text{event } i} = \frac{16}{306} \times L_{\text{event } i}^{\text{signal}} + \frac{290}{306} \times L_{\text{event } i}^{\text{bkg } b\bar{b}\gamma\gamma}$$

$$L_{\text{sample}} = \prod_{i=1}^{306} L_{\text{event } i} \quad (7.1)$$

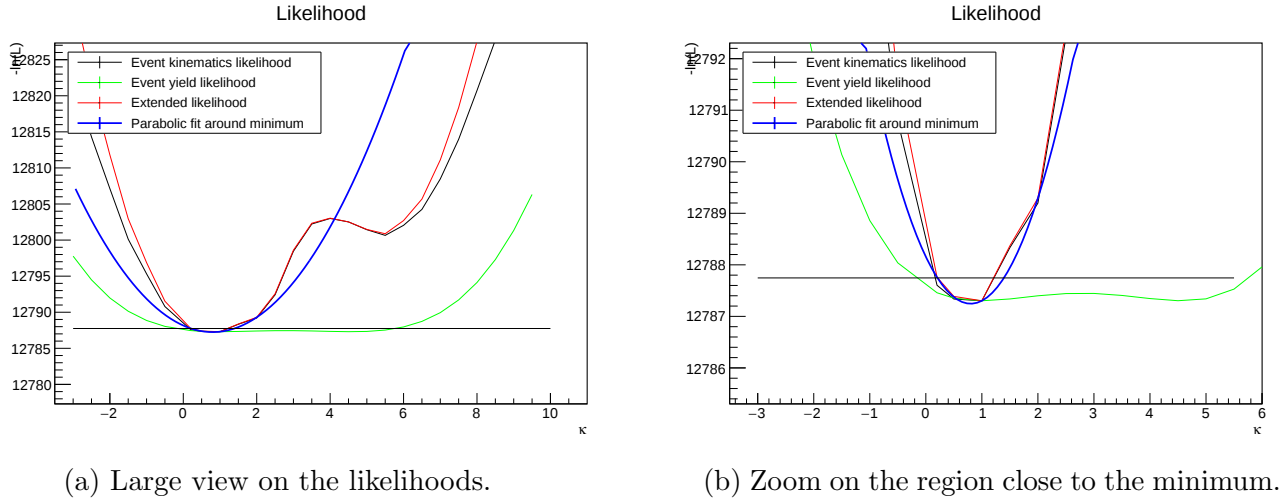


Figure 7.8: Negative logarithm of the event kinematics likelihood (black), the event yield likelihood (green) and the extended likelihood (red) as a function of κ for a sample of 16 signal events generated at $\kappa = 1$ and 290 non-resonant $b\bar{b}\gamma\gamma$ background events, and parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$ (blue). An arbitrary offset has been applied to all likelihood curves such that they are on the same scale. The horizontal black line corresponds to 0.5 above the minimum of the parabola; its intersection with the fitting parabola gives the statistical uncertainty on the measurement of κ .

The corresponding graph is the black curve of Fig. 7.8, referred to as event kinematics likelihood since it is based on the Matrix Element Method. The likelihood corresponding to a counting analysis, where only the event yield is used to extract λ , is represented in green (see Appendix E and Eq. E.1 for details). One interesting feature of the event kinematics likelihood is that it strongly rejects the region around $\kappa \approx 5$ that is allowed by the counting analysis. The two likelihoods are statistically independent and can be combined into an extended likelihood (red curve) that takes into account both event kinematics and event yield (see Appendix E and Eq. E.3 for details). The extended likelihood is then fitted using a parabola and just three values of κ close to the minimum, specifically $\kappa = 0.2, 1$ and 2 . The measured value of κ as well as the corresponding uncertainty in the measurement of κ are determined from the fitted parabola. For this sample, $\kappa = 0.80 \pm 0.59$. This measurement is consistent with the simulated

⁹The intent was to include at least 326 events, but various jobs ran into CPU time limit.

value of κ : 1.

This plot is important since it illustrates the measurement of the tri-linear Higgs coupling λ that will be performed with the data of the HL-LHC. The statistical uncertainty is much lower than the previous analyses [9] and [10] and is comparable to the predictions of [11]. However this plot does not show the outcome of an exhaustive and refined analysis, taking into account all background sources and detector efficiency.

Ensemble tests

No pull plot including background events was drawn due to the CPU needed to compute the likelihood for the signal hypothesis of thousands of $b\bar{b}\gamma\gamma$ events. It takes on average approximately 1 hour 20 minutes to compute the likelihood of one non-resonant $b\bar{b}\gamma\gamma$ background event.

7.7 Matrix Element Analysis of a mixed signal and ttH background ensemble

In this section, only the single Higgs boson ttH background is considered. Studying this background was motivated by the fact that it is the dominant single Higgs background (see Tab. 5.1). The likelihoods of ttH background events for the ttH background Matrix Element and for the signal Matrix Element for several hypothesis for κ were computed. After studying the performance of VEGAS and DIVONNE integration algorithms on $b\bar{b}\gamma\gamma$ background events (see Appendix G), we have decided to switch to DIVONNE for computing the likelihood integral.

The pre-selected signal events generated at $\kappa=1$ (see Chapter 5) were analysed using MoMEMta with the ttH background hypothesis, *i.e.* the configuration file presented in Appendix D.2. The likelihoods of 2000 signal events were computed.

The generated and pre-selected ttH background events (see Chapter 5) were analysed using MoMEMta with the signal hypothesis, *i.e.* the configuration file presented in Appendix D.1. The likelihoods of 160 and 2000 background events were respectively computed for $\kappa=-3, -2.5, -2, -1.5, -1, -0.5, 0.2, 0.5, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10$ and $\kappa=1$. These events were also analysed using MoMEMta with the ttH background hypothesis. The likelihoods of 2000 events were computed.

Integration of the likelihood

As detailed in the paragraph "Integration of the likelihood: convergence study" of Sec. 7.6, the signal events responsible for discontinuities in the scan of the likelihood were removed from the analysis, and due to the signal likelihood of some signal events being underestimated, the results obtained are suboptimal.

The likelihood of ttH background events analysed with the signal hypothesis turns out to be zero for approximately 85% of the events. Having events with extremely small likelihood is expected as ttH background events have very different kinematics in comparison with the signal events (compare Figs. 5.1-5.3 and C.4-C.6). In particular, major discrepancies can be found in the distributions of $\Delta R(b\bar{b}\gamma\gamma)$ and m_{bb} : ttH background events with low $\Delta R(b\bar{b}\gamma\gamma)$ or high m_{bb} are not signal-like at all. We have verified that events with a zero likelihood indeed have a

kinematics very different from the signal.

The likelihood of signal events analysed with the ttH background hypothesis is zero very often too. Approximately 30% of signal events have a null likelihood under the signal hypothesis (at $\kappa = 1$).

Approximately 8% of ttH background events also have a zero likelihood under the ttH background hypothesis. This is a problem since the events are generated and analysed with the same Matrix Element. So an exact zero likelihood is impossible. We believe that this is due to an imperfect integration by DIVONNE. The fact that the very narrow $H \rightarrow \gamma\gamma$ peak (the natural width of the Higgs boson is 4 MeV [28]) needs to be found by the numerical integration algorithm makes this task particularly difficult. Due to the lack of time, this has not been elucidated.

Given the good results with the $b\bar{b}\gamma\gamma$ background and the consistent explanation of null likelihood of ttH background events analysed with the signal hypothesis, we have good confidence in the signal Matrix Elements and the meaning of these zero likelihoods. However the ttH background events which have a zero likelihood under the ttH background are not physical, and these events are removed from the analysis: 6 ttH events out of 84 were removed and replaced by non-problematic events.

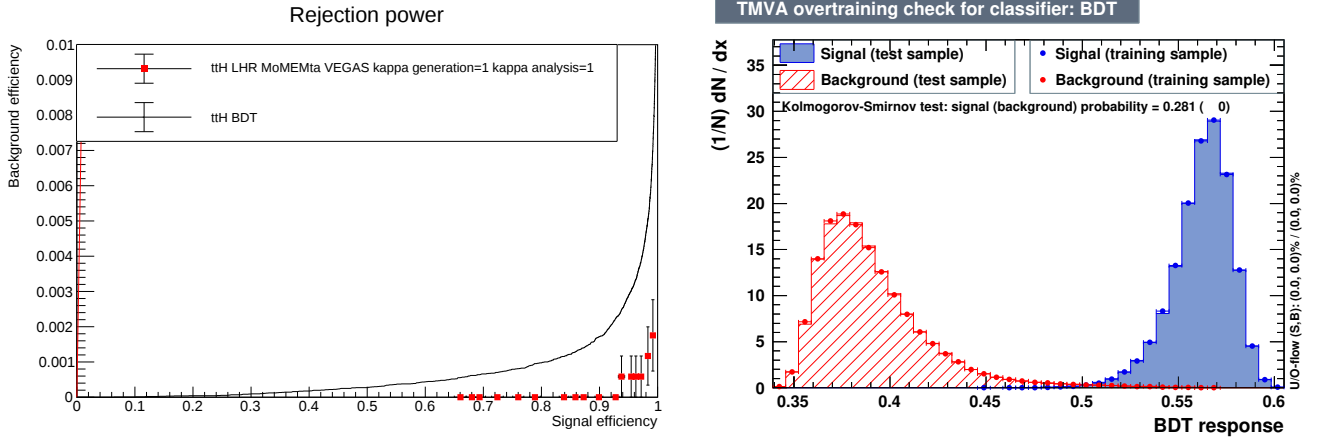
Rejection power

The likelihood ratio $L_{\text{signal}}/L_{\text{bkg}}$ (LHR) is the most powerful variable to classify signal and background events. In order to check this expectation, we have compared the rejection power based on the likelihood ratio from MoMEMta to the rejection power obtained with the most powerful cut-based analysis up to now, described in Ref. [10]. This cut-based analysis uses a BDT to classify signal and background events. The BDT is trained using fully simulated and reconstructed events to reject a cocktail of backgrounds (from different background processes). A direct comparison between the performance of this BDT and our LHR is thus not perfectly relevant. Therefore a BDT was trained exactly as it is done in the analysis [10] (same BDT configuration, same training algorithm, same input variables except those irrelevant to our simplified study such as the number of initial electrons), but on the events we have generated and to reject the ttH background only.

In order to be able to compute the likelihood ratio, the likelihood under the ttH background hypothesis must not be zero. As discussed in the above paragraph "Integration of the likelihood", signal events can have a genuine zero likelihood, but not ttH background events. Therefore, the zero likelihood of signal events was replaced by 10^{-45} , which is at least 10 orders of magnitude lower than the lowest likelihood under the signal hypothesis, and ttH events with zero likelihood were removed. After this operation, 1980 signal events and 1710 ttH events are available. The rejection powers of the BDT and the LHR are presented in Fig. 7.9. The rejection of the LHR is better than the one of the BDT.

Likelihood as a function of κ

For the reasons given in the paragraph "Integration of the likelihood" of Sec. 7.7, the ttH events having a zero likelihood for the ttH hypothesis were removed. All other events with zero likelihoods were kept with their likelihood unchanged. 84 ttH background events are expected



(a) Rejection power of the BDT and of the LHR.

(b) Overtraining check for the BDT.

Figure 7.9: Left: Rejection power of a BDT trained as in Ref. [10] and of the LHR from MoMEMta to classify signal events and ttH background events. 1980 signal events and 1710 ttH events were considered for making this plot. Right: Overtraining check for the BDT.

after the pre-selection requirements are applied (see Tab. 5.1). The likelihood of a sample of 16 signal events and 84 ttH background events was drawn. Based on Eqs. 3.1 and 3.2, the event likelihood and the sample likelihood are computed according to the following expressions:

$$\begin{aligned}
 L_{\text{event } i} &= \frac{16}{100} \times L_{\text{event } i}^{\text{signal}} + \frac{84}{100} \times L_{\text{event } i}^{\text{bkg } ttH} \\
 L_{\text{sample}} &= \prod_{i=1}^{100} L_{\text{event } i}
 \end{aligned} \tag{7.2}$$

The corresponding graph is the black curve of Fig. 7.10, referred to as event kinematics likelihood since it is based on the Matrix Element Method. The likelihood corresponding to a counting analysis, where only the event yield is used to extract λ , is represented in green (see Appendix E and Eq. E.1 for details). One interesting feature of the event kinematics likelihood is that it strongly rejects the region around $\kappa \approx 5$ that is allowed by the counting analysis. The two likelihoods are statistically independent and can be combined into an extended likelihood (red curve) that takes into account both event kinematics and event yield (see Appendix E and Eq. E.3 for details). The extended likelihood is then fitted using a parabola and just three values of κ close to the minimum, specifically $\kappa = 0.2, 1$ and 2 . The measured value of κ as well as the corresponding uncertainty in the measurement of κ are determined from the fitted parabola. For this sample, $\kappa = 1.13 \pm 0.33$, which is consistent with the simulated value of κ : 1. The statistical uncertainty on the measurement of λ is the same as the one for the analysis of a pure signal event ensemble: the ttH background does not significantly degrade the statistical uncertainty.

Ensemble tests

No pull plot including background events was drawn due to the CPU needed to compute the likelihood for the signal hypothesis of thousands of ttH events.

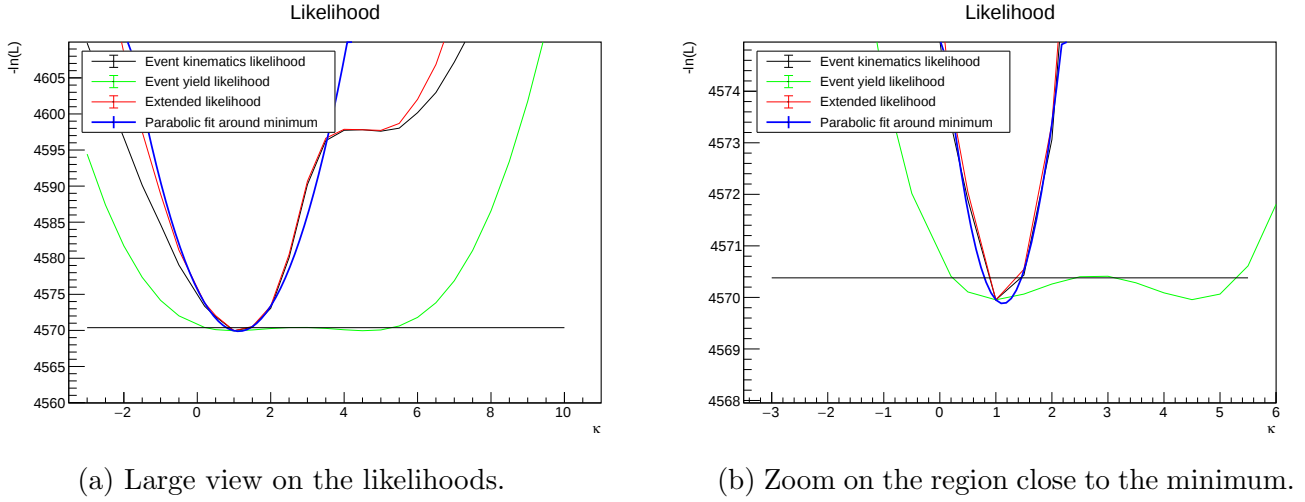


Figure 7.10: Negative logarithm of the event kinematics likelihood (black), the event yield likelihood (green) and the extended likelihood (red) as a function of κ for a sample of 16 signal events generated at $\kappa = 1$ and 84 $t\bar{t}H$ background events, and parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$ (blue). An arbitrary offset has been applied to all likelihood curves such that they are on the same scale. The horizontal black line corresponds to 0.5 above the minimum of the parabola; its intersection with the fitting parabola gives the statistical uncertainty on the measurement of κ .

7.8 Matrix Element Analysis of a mixed signal, $b\bar{b}\gamma\gamma$ and $t\bar{t}H$ background ensemble

In this section, the non-resonant $b\bar{b}\gamma\gamma$ background and the single Higgs boson $t\bar{t}H$ backgrounds are considered. These two backgrounds are the dominant non-resonant and single Higgs backgrounds (see Tab. 5.1).

The generated and pre-selected $t\bar{t}H$ background events (see Chapter 5) were analysed using MoMEMta with the $b\bar{b}\gamma\gamma$ background hypothesis (*i.e.* the configuration file presented in Appendix D.1, except for the integration algorithm, DIVONNE was used instead of VEGAS). The likelihood of 2000 $t\bar{t}H$ events were computed.

The generated and pre-selected $b\bar{b}\gamma\gamma$ background events (see Chapter 5) were analysed using MoMEMta with the $t\bar{t}H$ background hypothesis (*i.e.* the configuration file presented in Appendix D.2). The likelihood of 2000 $b\bar{b}\gamma\gamma$ events were computed.

Integration of the likelihood

As detailed in the paragraph "Integration of the likelihood: convergence study" of Sec. 7.6, the signal events responsible for discontinuities in the scan of the likelihood were removed from the analysis, and due to the signal likelihood of some signal events being underestimated, the results obtained are suboptimal.

As detailed in the paragraph "Integration of the likelihood" of Sec. 7.7, $t\bar{t}H$ background events having a null likelihood for the $t\bar{t}H$ background hypothesis were removed from the analysis.

The likelihood of ttH background events analysed with the $b\bar{b}\gamma\gamma$ hypothesis turns out to be zero for approximately 75% of the events. It is very likely that the kinematics of individual ttH events are excluded from the most likely kinematics of $b\bar{b}\gamma\gamma$ background events. This was not checked.

The likelihood of $b\bar{b}\gamma\gamma$ background events analysed with the ttH hypothesis turns out to be zero for approximately 1% of the events.

These zero likelihoods are interpreted as genuine null likelihood due to the fact the kinematics of $b\bar{b}\gamma\gamma$ and ttH events are different. The events with the null likelihoods were thus kept for the analysis.

Likelihood as a function of κ

326 non-resonant $b\bar{b}\gamma\gamma$ and 84 ttH background events are expected after the pre-selection requirements are applied (see Tab. 5.1). The likelihood of a sample of 16 signal events, 290¹⁰ non-resonant $b\bar{b}\gamma\gamma$ background events and 84 ttH background events was drawn. Based on Eqs. 3.1 and 3.2, the event likelihood and the sample likelihood are computed according to the following expressions:

$$L_{\text{event } i} = \frac{16}{390} \times L_{\text{event } i}^{\text{signal}} + \frac{290}{390} \times L_{\text{event } i}^{\text{bkg } b\bar{b}\gamma\gamma} + \frac{84}{390} \times L_{\text{event } i}^{\text{bkg } ttH}$$

$$L_{\text{sample}} = \prod_{i=1}^{390} L_{\text{event } i} \quad (7.3)$$

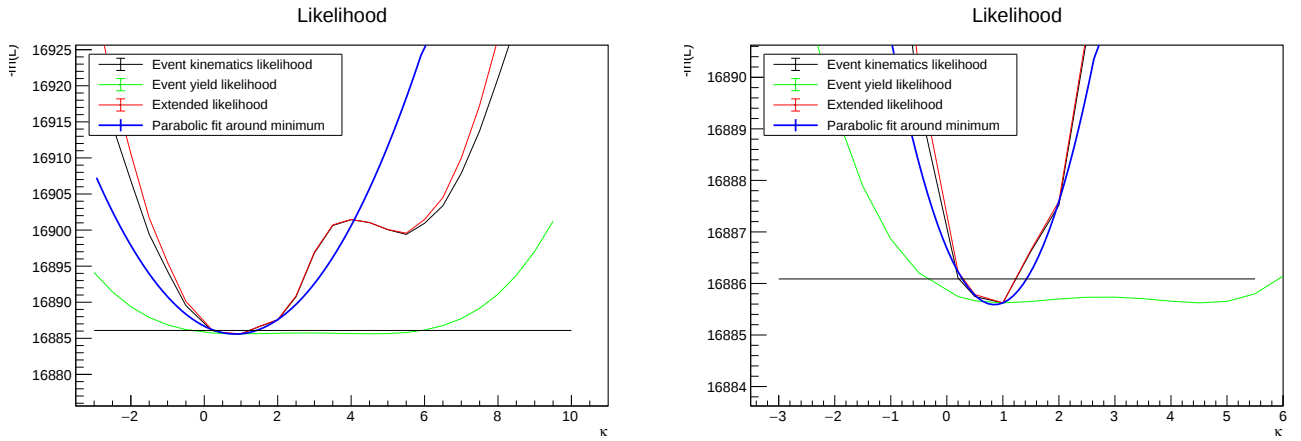
The corresponding graph is the black curve of Fig. 7.11, referred to as event kinematics likelihood since it is based on the Matrix Element Method. The likelihood corresponding to a counting analysis, where only the event yield is used to extract λ , is represented in green (see Appendix E and Eq. E.1 for details). One interesting feature of the event kinematics likelihood is that it strongly rejects the region around $\kappa \approx 5$ that is allowed by the counting analysis. The two likelihoods are statistically independent and can be combined into an extended likelihood (red curve) that takes into account both event kinematics and event yield (see Appendix E and Eq. E.3 for details). The extended likelihood is then fitted using a parabola and just three values of κ close to the minimum, specifically $\kappa = 0.2, 1$ and 2 . The measured value of κ as well as the corresponding uncertainty in the measurement of κ are determined from the fitted parabola. For this sample, $\kappa = 0.85 \pm 0.57$, which is consistent with the simulated value of κ : 1. The statistical uncertainty on the measurement of λ is the same as the one for the analysis of a mixed signal and non-resonant $b\bar{b}\gamma\gamma$ background ensemble: the ttH background does not significantly degrade the statistical uncertainty.

The statistical uncertainty is much lower than the previous analyses [9] and [10] and is of the same order as the predictions of [11]. However this plot does not show the outcome of an exhaustive and refined analysis, taking into account all background sources and detector efficiency.

Ensemble tests

No pull plot including background events was drawn due to the CPU needed to compute the likelihood for the signal and ttH hypotheses of thousands of $b\bar{b}\gamma\gamma$ and ttH events.

¹⁰The intent was to include at least 326 events, but various jobs ran into CPU time limit.



(a) Large view on the likelihoods.

(b) Zoom on the region close to the minimum.

Figure 7.11: Negative logarithm of the event kinematics likelihood (black), the event yield likelihood (green) and the extended likelihood (red) as a function of κ for a sample of 16 signal events generated at $\kappa = 1$, 290 non-resonant $b\bar{b}\gamma\gamma$ background events and 84 $t\bar{t}H$ background events, and parabolic fit of the extended likelihood using the points at $\kappa = 0.2, 1, 2$ (blue). An arbitrary offset has been applied to all likelihood curves so that they are on the same scale. The horizontal black line corresponds to 0.5 above the minimum of the parabola; its intersection with the fitting parabola gives the statistical uncertainty on the measurement of κ .

Chapter 8

Conclusion and outlook

In this report, we have demonstrated that the Matrix Element Method can be successfully used for the measurement of the tri-linear Higgs boson coupling by implementing this technique in a simplified case: no detector efficiency, not all background processes and events generated at parton level were used. Despite these simplifying assumptions, this analysis includes all the main conceptual aspects of the MEM analysis.

As expected from the definition of this method, the MEM is a powerful technique and provides very promising prospects for the measurement of the tri-linear Higgs boson coupling. The results we have obtained using the publicly available tool MoMEMta are very encouraging. The improvement from the use of the MEM over counting experiments and counting experiments in bins of m_{HH} (like e.g. in Ref. [10]) for pure signal samples has been discussed in Sec. 7.5. Even more importantly, the substantial improvement in signal versus background discrimination power from the use of the MEM instead of the BDT from Ref. [10] has been shown in Secs. 7.6 and 7.7. In addition to the inherent improvement in the rejection power, the MEM benefits from the fact that no strict cuts for background *rejection* from the event sample are needed; instead all events (after a loose preselection) are exploited in the analysis (cf. discussion in Sec. 3.5).

Only the main non-resonant and single-Higgs background processes were considered in this analysis, accounting for slightly less than half of the expected total number of background events (see Tab. 5.1). The addition of the other background processes will increase the statistical uncertainty. This increase depends on the kinematics of the processes and the results obtained in this study cannot be extrapolated. If the degradation of the statistical uncertainty due to the additional backgrounds is the same as the degradation due to the backgrounds that were simulated in this study, then a statistical uncertainty of the order of 0.8 is expected.

The next steps in this analysis are to include all background processes, detector efficiency and a more realistic detector simulation. For the detector simulation, a parametrised simulation like in Ref. [10] (the latest ATLAS study) or in the Delphes program [51] will be used. Large simulated samples – including all backgrounds – of the fully simulated events are expected to be available at the time of the HL-LHC. When these more realistic simulations are used in the event generation, more realistic transfer functions will have to be implemented for the MEM analysis. For practical reasons, the speed of the MEM calculations, in particular for the signal hypothesis, will have to be improved. In the present study, the CPU time per event for the MEM calculation using the signal hypothesis was around 1h 15. At least an order of magnitude

in execution speed can be gained by explicitly exploiting the Higgs mass peaks in the Lua file (cf. discussion in Sec. 7.3) and by manually optimising the code for the evaluation of the signal Matrix Element that was automatically generated by MadEvent [52]¹. In this code, a large number of unused helicity amplitudes are calculated in every single call.

¹This optimised code for the signal Matrix Element is available. You can contact A. Lleres at annick.lleres@lpsc.in2p3.fr.

Appendix A

LO Run Card used for event generation

The parameter card used for background event generation is the following:

```
*****
#                               MadGraph5_aMC@NLO                               *
#                                                                           *
#                               run_card.dat MadEvent                           *
#                                                                           *
# This file is used to set the parameters of the run.                       *
#                                                                           *
# Some notation/conventions:                                                 *
#                                                                           *
# Lines starting with a '#' are info or comments                           *
#                                                                           *
# mind the format:  value      = variable      ! comment                    *
*****
#
*****
# Running parameters
*****
#
*****
# Tag name for the run (one word)                                           *
*****
tag_1      = run_tag ! name of the run
*****
# Run to generate the grid pack                                             *
*****
False      = gridpack !True = setting up the grid pack
*****
# Number of events and rnd seed                                           *
# Warning: Do not generate more than 1M events in a single run             *
# If you want to run Pythia, avoid more than 50k events in a run.          *
*****
75 = nevents ! Number of unweighted events requested
0  = iseed   ! rnd seed (0=assigned automatically=default))
*****
# Collider type and energy                                                 *
# lpp: 0=No PDF, 1=proton, -1=antiproton, 2=photon from proton,            *
```

```

#                                     3=photon from electron      *
#####
1      = lpp1      ! beam 1 type
1      = lpp2      ! beam 2 type
7000   = ebeam1    ! beam 1 total energy in GeV
7000   = ebeam2    ! beam 2 total energy in GeV
#####
# Beam polarization from -100 (left-handed) to 100 (right-handed)  *
#####
0.0    = polbeam1 ! beam polarization for beam 1
0.0    = polbeam2 ! beam polarization for beam 2
#####
# PDF CHOICE: this automatically fixes also alpha_s and its evol.  *
#####
nn23lo1 = pdlabel      ! PDF set
230000  = lhaid        ! if pdlabel=lhapdf, this is the lhapdf number
#####
# Renormalization and factorization scales                        *
#####
False = fixed_ren_scale ! if .true. use fixed ren scale
False  = fixed_fac_scale ! if .true. use fixed fac scale
91.188 = scale          ! fixed ren scale
91.188 = dsqrt_q2fact1  ! fixed fact scale for pdf1
91.188 = dsqrt_q2fact2  ! fixed fact scale for pdf2
-1 = dynamical_scale_choice ! Choose one of the preselected dynamical choices
1.0 = scalefact         ! scale factor for event-by-event scales
#####
# Time of flight information. (-1 means not run)
#####
-1.0 = time_of_flight ! threshold below which info is not written
#####
# Matching - Warning! ickkw > 1 is still beta
#####
0 = ickkw          ! 0 no matching, 1 MLM, 2 CKKW matching
1 = highestmult    ! for ickkw=2, highest mult group
1 = ktscheme       ! for ickkw=1, 1 Durham kT, 2 Pythia pTE
1.0 = alpsfact     ! scale factor for QCD emission vx
False = chcluster  ! cluster only according to channel diag
True = pdfwgt      ! for ickkw=1, perform pdf reweighting
4 = asrwgtflavor   ! highest quark flavor for a_s reweight
True = clusinfo     ! include clustering tag in output
3.0 = lhe_version  ! Change the way clustering information pass to shower.
#####
#####
#
#####
# Automatic ptj and mjj cuts if xqcut > 0
# (turn off for VBF and single top processes)
#####
True = auto_ptj_mjj ! Automatic setting of ptj and mjj
#####

```

```

#
#*****
# BW cutoff (M+/-bw cutoff*Gamma)
#*****
15.0 = bw cutoff      ! (M+/-bw cutoff*Gamma)
#*****
# Apply pt/E/eta/dr/mij/kt_durham cuts on decay products or not
# (note that etmiss/ptll/ptheavy/ht/sorted cuts always apply)
#*****
False = cut_decays    ! Cut decay products
#*****
# Number of helicities to sum per event (0 = all helicities)
# 0 gives more stable result, but longer run time (needed for
# long decay chains e.g.).
# Use >=2 if most helicities contribute, e.g. pure QCD.
#*****
0 = nhel              ! Number of helicities used per event
#*****
# Standard Cuts
#*****
#
#*****
# Minimum and maximum pt's (for max, -1 means no cut) *
#*****
20.0 = ptj            ! minimum pt for the jets
20.0 = ptb            ! minimum pt for the b
25.0 = pta            ! minimum pt for the photons
0.0 = ptl            ! minimum pt for the charged leptons
0.0 = misset          ! minimum missing Et (sum of neutrino's momenta)
0.0 = ptheavy         ! minimum pt for one heavy final state
-1.0 = ptjmax         ! maximum pt for the jets
-1.0 = ptbmax         ! maximum pt for the b
-1.0 = ptamax         ! maximum pt for the photons
-1.0 = ptlmax         ! maximum pt for the charged leptons
-1.0 = missetMax      ! maximum missing Et (sum of neutrino's momenta)
#*****
# Minimum and maximum E's (in the center of mass frame) *
#*****
0.0 = ej             ! minimum E for the jets
0.0 = eb             ! minimum E for the b
0.0 = ea             ! minimum E for the photons
0.0 = el             ! minimum E for the charged leptons
-1.0 = ejmax         ! maximum E for the jets
-1.0 = ebmax         ! maximum E for the b
-1.0 = eamax         ! maximum E for the photons
-1.0 = elmax         ! maximum E for the charged leptons
#*****
# Maximum and minimum absolute rapidity (for max, -1 means no cut) *
#*****
-1.0 = etaj          ! max rap for the jets
-1.0 = etab          ! max rap for the b

```

```

2.7 = etaa      ! max rap for the photons
-1.0 = etal     ! max rap for the charged leptons
0.0 = etajmin   ! min rap for the jets
0.0 = etabmin   ! min rap for the b
0.0 = etaamin   ! min rap for the photons
0.0 = etalmin   ! main rap for the charged leptons
*****
# Minimum and maximum DeltaR distance                                     *
*****
0.4 = drjj      ! min distance between jets
0.4 = drbb      ! min distance between b's
0.0 = drll      ! min distance between leptons
0.4 = draa      ! min distance between gammas
0.4 = drbj      ! min distance between b and jet
0.4 = draj      ! min distance between gamma and jet
0.0 = drjl      ! min distance between jet and lepton
0.4 = drab      ! min distance between gamma and b
0.0 = drbl      ! min distance between b and lepton
0.0 = dral      ! min distance between gamma and lepton
-1.0 = drjjmax  ! max distance between jets
-1.0 = drbbmax  ! max distance between b's
-1.0 = drllmax  ! max distance between leptons
-1.0 = draamax  ! max distance between gammas
-1.0 = drbjmax  ! max distance between b and jet
-1.0 = drajmax  ! max distance between gamma and jet
-1.0 = drjlmax  ! max distance between jet and lepton
-1.0 = drabmax  ! max distance between gamma and b
-1.0 = drblmax  ! max distance between b and lepton
-1.0 = dralmax  ! max distance between gamma and lepton
*****
# Minimum and maximum invariant mass for pairs                           *
# WARNING: for four lepton final state mlll cut require to have         *
#           different lepton masses for each flavor!                     *
*****
25.0 = mmjj      ! min invariant mass of a jet pair
45.0 = mmbb      ! min invariant mass of a b pair
60.0 = mmaa      ! min invariant mass of gamma gamma pair
0.0 = mlll       ! min invariant mass of l+l- (same flavour) lepton pair
-1.0 = mmjjmax   ! max invariant mass of a jet pair
-1.0 = mmbbmax   ! max invariant mass of a b pair
200.0 = mmaamax  ! max invariant mass of gamma gamma pair
-1.0 = mlllmax   ! max invariant mass of l+l- (same flavour) lepton pair
*****
# Minimum and maximum invariant mass for all letpons                     *
*****
0.0 = mnnl       ! min invariant mass for all letpons (l+- and vl)
-1.0 = mnnlmax   ! max invariant mass for all letpons (l+- and vl)
*****
# Minimum and maximum pt for 4-momenta sum of leptons                   *
*****
0.0 = ptllmin    ! Minimum pt for 4-momenta sum of leptons(l and vl)

```

```

-1.0 = ptllmax ! Maximum pt for 4-momenta sum of leptons(l and vl)
#*****
# Inclusive cuts *
#*****
0.0 = xptj ! minimum pt for at least one jet
0.0 = xptb ! minimum pt for at least one b
0.0 = xpta ! minimum pt for at least one photon
0.0 = xptl ! minimum pt for at least one charged lepton
#*****
# Control the pt's of the jets sorted by pt *
#*****
0.0 = ptj1min ! minimum pt for the leading jet in pt
0.0 = ptj2min ! minimum pt for the second jet in pt
0.0 = ptj3min ! minimum pt for the third jet in pt
0.0 = ptj4min ! minimum pt for the fourth jet in pt
-1.0 = ptj1max ! maximum pt for the leading jet in pt
-1.0 = ptj2max ! maximum pt for the second jet in pt
-1.0 = ptj3max ! maximum pt for the third jet in pt
-1.0 = ptj4max ! maximum pt for the fourth jet in pt
0 = cutuse ! reject event if fails any (0) / all (1) jet pt cuts
#*****
# Control the pt's of leptons sorted by pt *
#*****
0.0 = ptl1min ! minimum pt for the leading lepton in pt
0.0 = ptl2min ! minimum pt for the second lepton in pt
0.0 = ptl3min ! minimum pt for the third lepton in pt
0.0 = ptl4min ! minimum pt for the fourth lepton in pt
-1.0 = ptl1max ! maximum pt for the leading lepton in pt
-1.0 = ptl2max ! maximum pt for the second lepton in pt
-1.0 = ptl3max ! maximum pt for the third lepton in pt
-1.0 = ptl4max ! maximum pt for the fourth lepton in pt
#*****
# Control the Ht(k)=Sum of k leading jets *
#*****
0.0 = htjmin ! minimum jet HT=Sum(jet pt)
-1.0 = htjmax ! maximum jet HT=Sum(jet pt)
0.0 = ihtmin !inclusive Ht for all partons (including b)
-1.0 = ihtmax !inclusive Ht for all partons (including b)
0.0 = ht2min ! minimum Ht for the two leading jets
0.0 = ht3min ! minimum Ht for the three leading jets
0.0 = ht4min ! minimum Ht for the four leading jets
-1.0 = ht2max ! maximum Ht for the two leading jets
-1.0 = ht3max ! maximum Ht for the three leading jets
-1.0 = ht4max ! maximum Ht for the four leading jets
#*****
# Photon-isolation cuts, according to hep-ph/9801442 *
# When ptgmin=0, all the other parameters are ignored *
# When ptgmin>0, pta and draj are not going to be used *
#*****
0.0 = ptgmin ! Min photon transverse momentum
0.0 = R0gamma ! Radius of isolation code

```

```

1.0 = xn ! n parameter of eq.(3.4) in hep-ph/9801442
1.0 = epsgamma ! epsilon_gamma parameter of eq.(3.4) in hep-ph/9801442
True = isoEM ! isolate photons from EM energy (photons and leptons)
*****
# WBF cuts *
*****
0.0 = xetamin ! minimum rapidity for two jets in the WBF case
0.0 = deltaeta ! minimum rapidity for two jets in the WBF case
*****
# KT DURHAM CUT *
*****
-1.0 = ktdurham
0.0 = dparameter
*****
# maximal pdg code for quark to be considered as a light jet *
# (otherwise b cuts are applied) *
*****
4 = maxjetflavor ! Maximum jet pdg code
*****
# Jet measure cuts *
*****
0.0 = xqcut ! minimum kt jet measure between partons
*****
#
*****
# Store info for systematics studies *
# WARNING: If use_syst is T, matched Pythia output is *
#           meaningful ONLY if plotted taking matchscale *
#           reweighting into account! *
*****
True = use_syst ! Enable systematics studies
#
*****
# Parameter of the systematics study
# will be used by SysCalc (if installed)
*****
#
0.5 1 2 = sys_scalefact # factorization/renormalization scale factor
None = sys_alpsfact # \alpha_s emission scale factors
30 50 = sys_matchscale # variation of merging scale
# PDF sets and number of members (0 or none for all members).
Ct10nlo.LHgrid = sys_pdf # matching scales
# MSTW2008nlo68cl.LHgrid 1 = sys_pdf

```

Appendix B

Dimension of the likelihood integral

B.1 Derivation of a general formula

The dimension of the likelihood integral, given in Eq. 3.7, is a matter of counting the degrees of freedom that remain after the energy-momentum conservation and the transfer functions have been applied.

Let us consider the general case of a final state with N partons, among which there are N_v visible particles. Each of these partons is described by a four-vector (η, ϕ, p_T, m) . The two colliding partons have a well-known momentum (because the beam direction is known) but the fraction of the proton energy they carry – commonly referred to as the Bjorken x scaling variable – is unknown. So the dimension of the phase-space before any conservation law and a transfer function are taken into account is $4 \times N + 2$.

The conservation of the energy-momentum removes 4 degrees of freedom. The space-space dimension decreases to $4 \times (N - 1) + 2$.

In the MEM, all final state particles are identified and treated as on-shell particles, meaning that their invariant mass is determined. This decreases the number of degrees of freedom by N . The phase-space dimension reduces to $3 \times N - 2$.

The transfer functions of physical quantities for which the resolution of the detector is assumed to be perfect are Dirac functions. Suppose that N_ϕ , N_η and N_E non Dirac transfer functions are applied on the final state objects respectively on the measurement of ϕ , η and the energy. This means that $N_v - N_\phi$, $N_v - N_\eta$ and $N_v - N_E$ kinematic variables are known exactly, and the degrees of freedom relative to these variables vanish. **The phase-space dimension after all conservation laws and transfer functions is $3 \times (N - N_v) + N_\phi + N_\eta + N_E - 2$.**

B.2 Application to the signal and background processes

Let us consider the $b\bar{b}\gamma\gamma$ final-state and calculate the dimension of the integral for the signal, $b\bar{b}\gamma\gamma$ and $t\bar{t}H$ backgrounds hypotheses. Let us remind that a transfer function on the energy of the photons and b quarks is applied.

• **Integral dimension for the signal and non-resonant $b\bar{b}\gamma\gamma$ hypotheses**

Since these two processes result in the same final state, the dimension of the integral is the same. The final state is composed of $N = 4$ particles and all particles are visible: $N_v = 4$. Four transfer functions (on the b quarks and photons energy) are applied: $N_\phi + N_\eta + N_E = 4$. Therefore the dimension of the likelihood integral is 2.

• **Integral dimension for the $t\bar{t}H$ hypothesis**

In the way the Lua configuration file is written, the final state is composed of $N = 6$ particles: $\gamma\gamma W^+ W^- b\bar{b}$. The 2 W bosons are invisible for the MEM: $N_v = 4$. Four transfer functions (on the b quarks and photons energy) are applied: $N_\phi + N_\eta + N_E = 4$. Therefore the dimension of the likelihood integral is 8.

Appendix C

Distributions of the main kinematic variables for background events

C.1 Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events

All the distributions are normalised to unit surface.

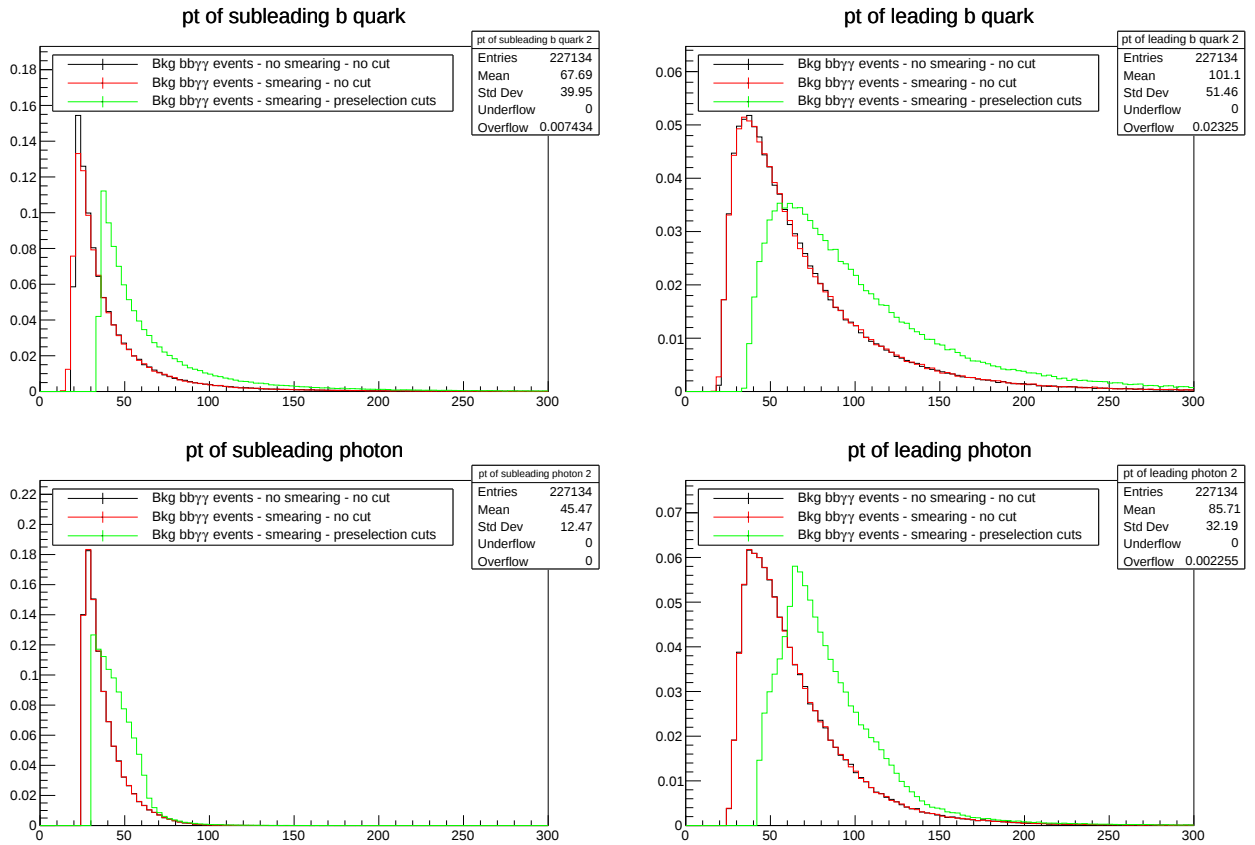


Figure C.1: Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

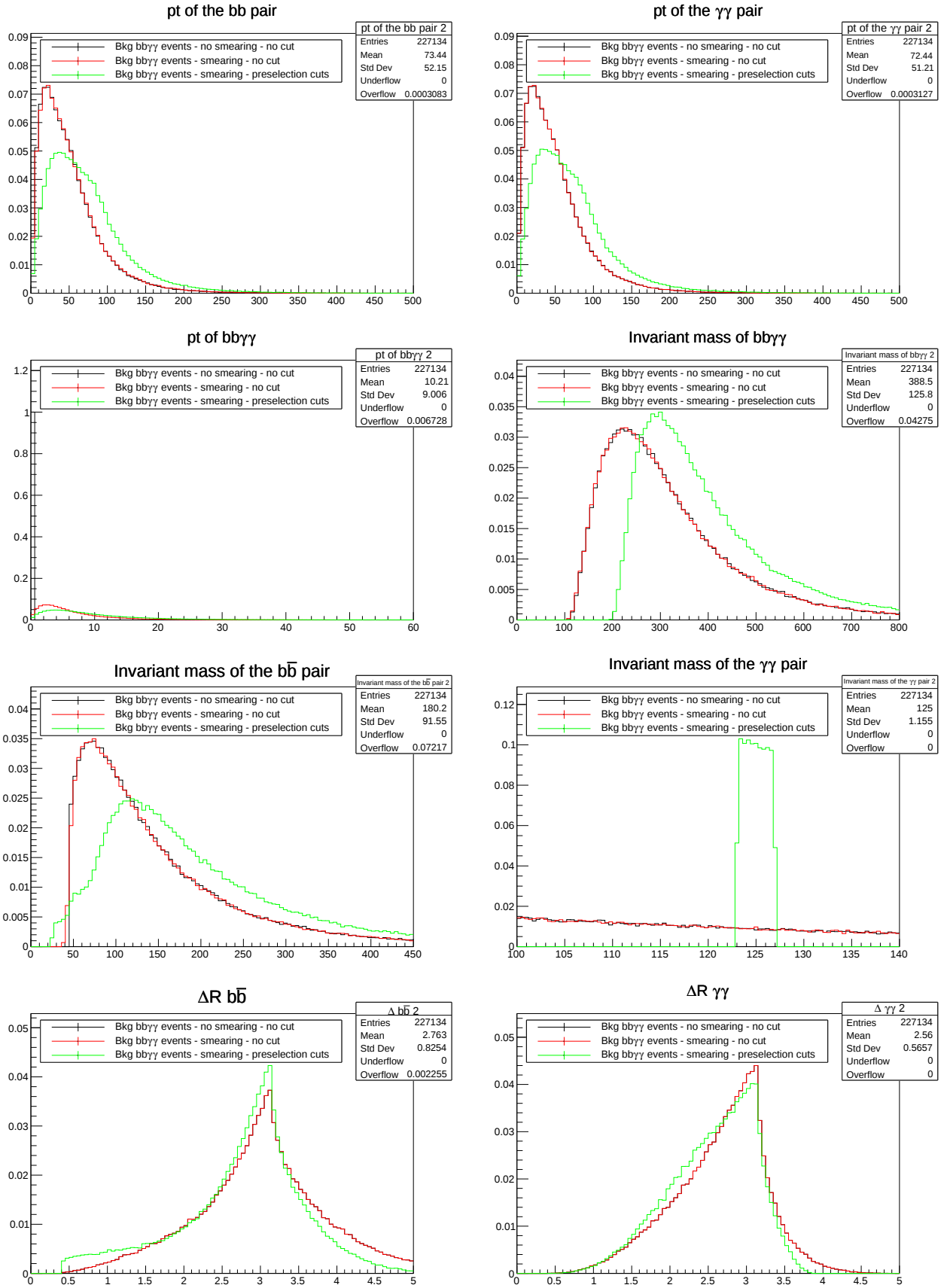


Figure C.2: Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

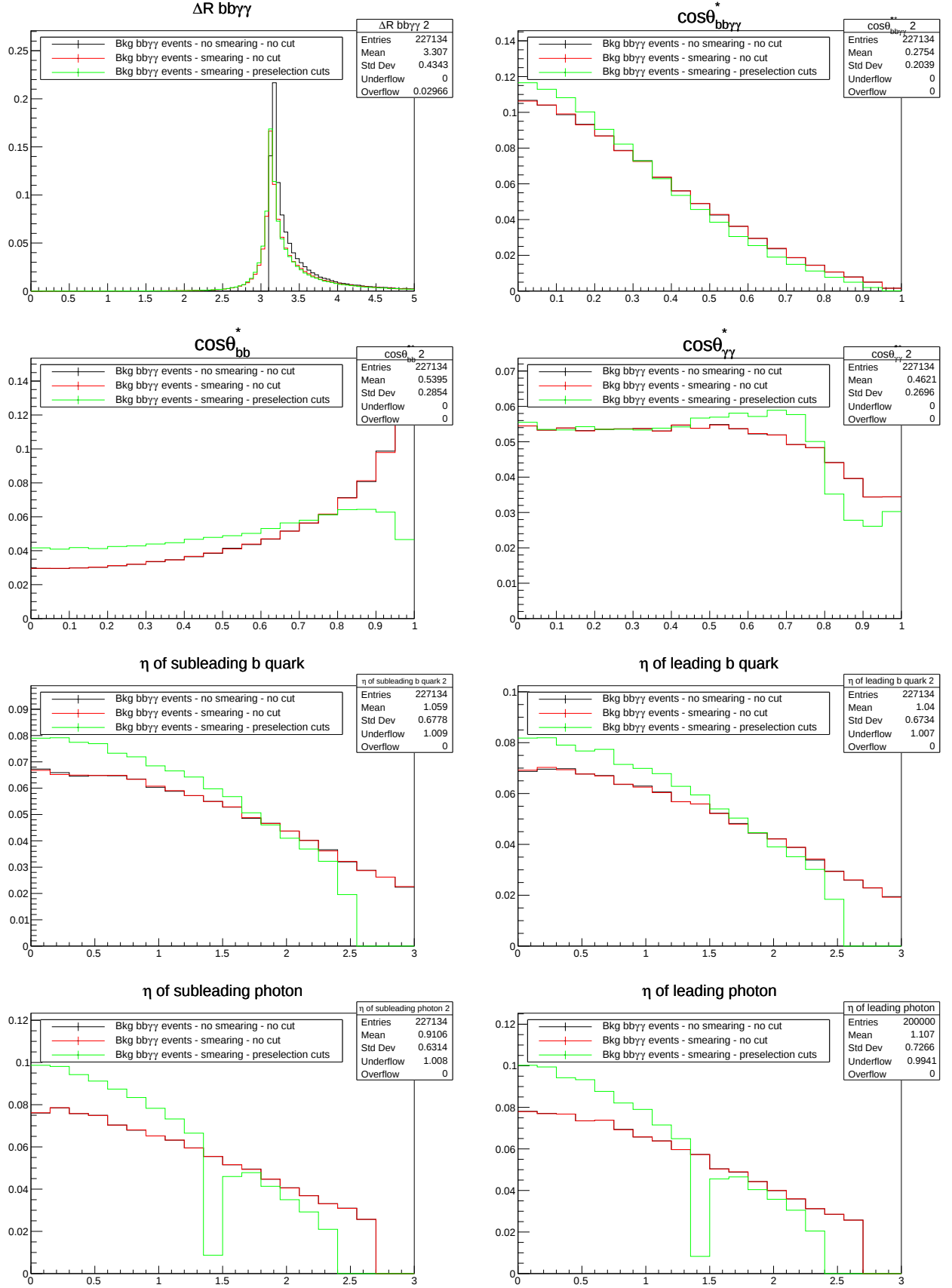


Figure C.3: Distributions of the main kinematic variables for the non-resonant $b\bar{b}\gamma\gamma$ background events after generation (black), after smearing (red), and after smearing and pre-selection requirements are applied (green).

C.2 Distributions of the main kinematic variables for the $t\bar{t}H$ background events

All the distributions are normalised to unit surface.

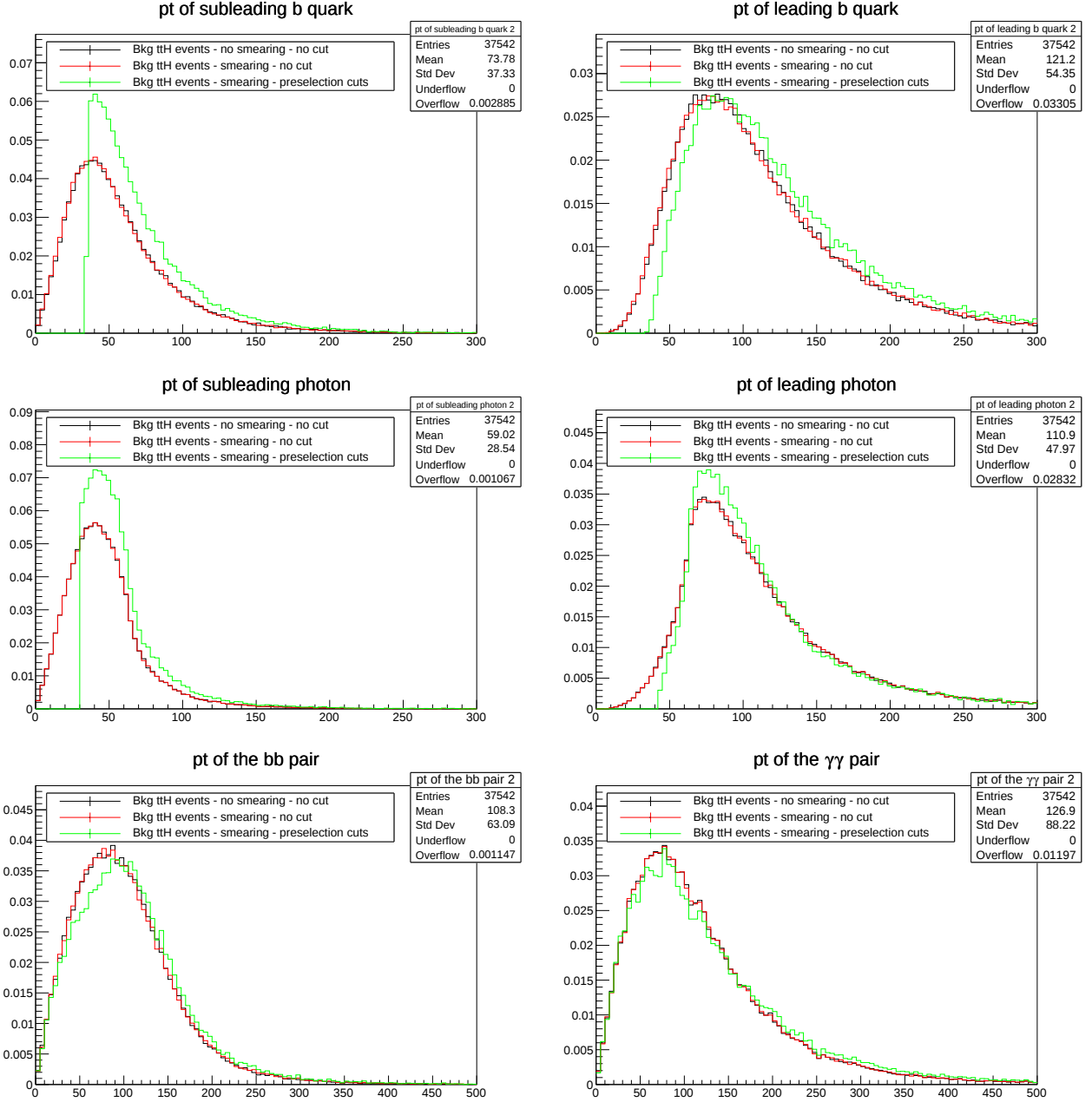


Figure C.4: Distributions of the main kinematic variables for the single Higgs boson $t\bar{t}H$ background events after generation (black), after smearing (red), and after smearing and preselection requirements are applied (green).

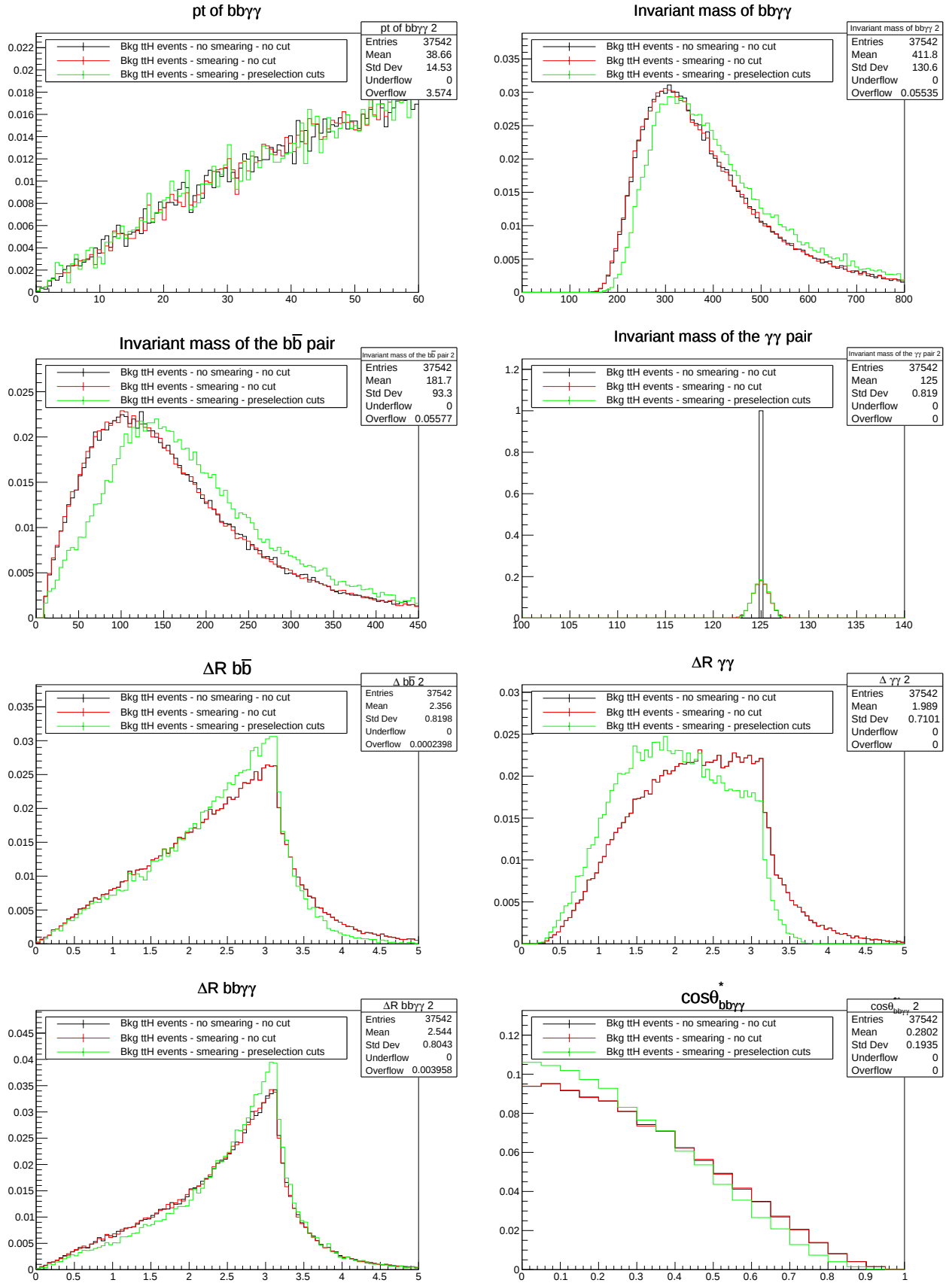


Figure C.5: Distributions of the main kinematic variables for the single Higgs boson ttH background events after generation (black), after smearing (red), and after smearing and preselection requirements are applied (green).

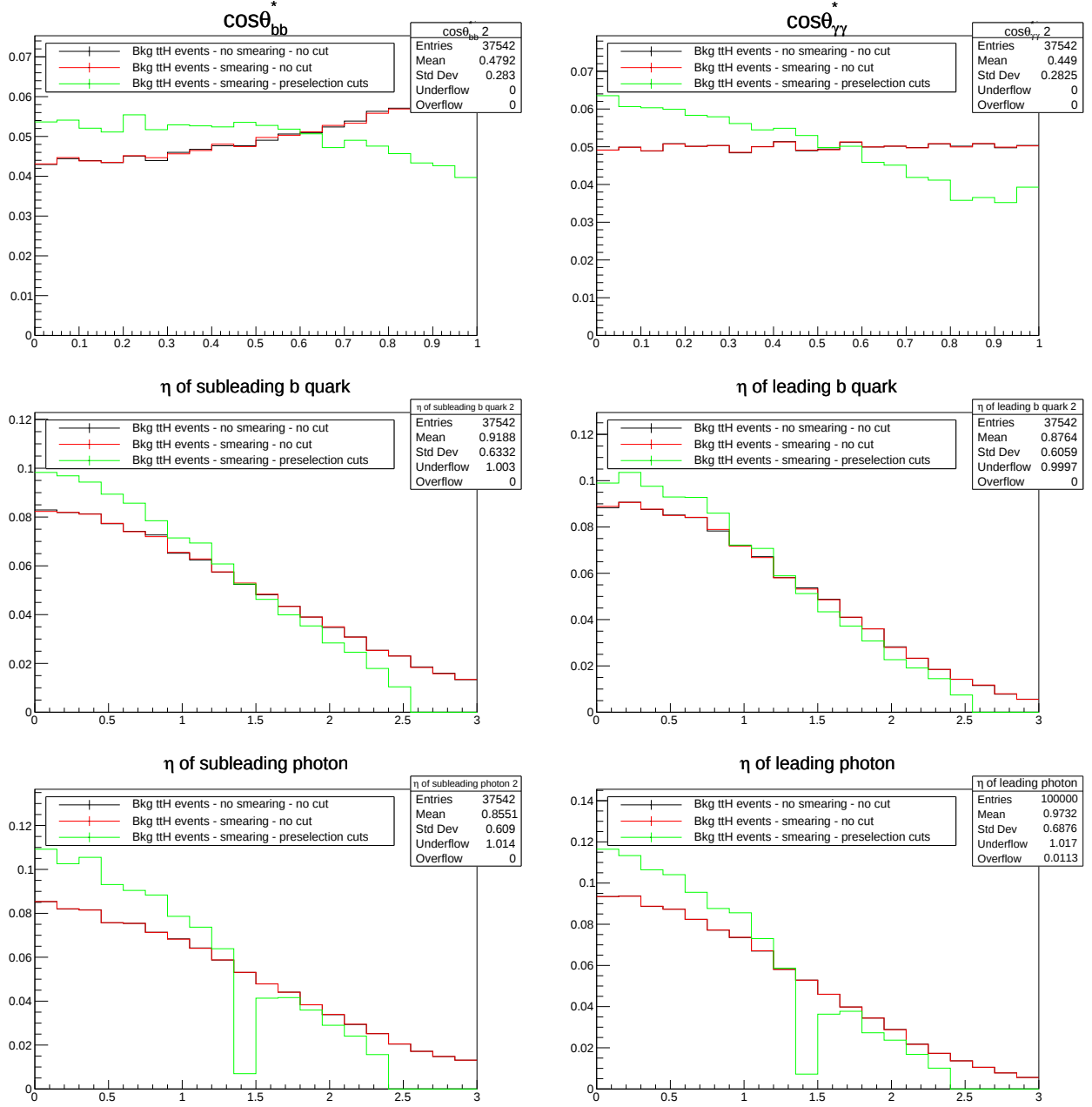


Figure C.6: Distributions of the main kinematic variables for the single Higgs boson ttH background events after generation (black), after smearing (red), and after smearing and preselection requirements are applied (green).

Appendix D

Lua configuration files

D.1 Lua configuration file for the signal and $b\bar{b}\gamma\gamma$ non-resonant background

The Lua configuration file presented in this section is the one used for the signal Matrix Element. The configuration file is very close to the Lua file for the non-resonant $b\bar{b}\gamma\gamma$ background. Differences between the two Lua files are indicated with the comments `-- bkg ME` and `-- signal ME`.

```
load_modules("../bbyy/MatrixElement/build/libme_gg2HH2bbyy.so") -- signal ME
-- load_modules("../bbyy/MatrixElement/build/libme_pp2bbyy.so") -- bkg ME

-- Register inputs
local y1 = declare_input("y1")
local y2 = declare_input("y2")
local b1 = declare_input("b1")
local b2 = declare_input("b2")

-- Define parameters
parameters = {
    energy = 14000.,
    H_mass = 125.,
    H_width = 0.005753088,
    kappa = 1., -- Default value, will be overridden
}

-- Define multi-dimensional integration algorithm
cuba = {
    algorithm = "vegas",
    relative_accuracy = 0.005,
    max_eval = 2000000,
    n_start = 100000,
    verbosity = 3,
}

-- Gaussian transfer function on energy of one photon and one b parton
GaussianTransferFunctionOnEnergy.tf_p1 = {
    ps_point = add_dimension(),
    reco_particle = y1.reco_p4,
```

```

    sigma = 0.01,
}
y1.set_gen_p4("tf_p1::output")

GaussianTransferFunctionOnEnergy.tf_p3 = {
    ps_point = add_dimension(),
    reco_particle = b1.reco_p4,
    sigma = 0.10,
}
b1.set_gen_p4("tf_p3::output")

-- Define phase-space
StandardPhaseSpace.phaseSpaceOut = {
    particles = { y1.gen_p4, b1.gen_p4 } -- only on visible particles
}

-- Change of variables: Main Block A
BlockA.blocka = {
    p1 = y2.reco_p4,
    p2 = b2.reco_p4,
    branches = { y1.gen_p4, b1.gen_p4 },
}

-- Loop over block solutions

Looper.looper = {
    solutions = "blocka::solutions",
    path = Path("initialState", "bbyy", "integrand", "tf_p2", "tf_p4")
}

-- Gaussian transfer function on energy of the other photon and b parton
GaussianTransferFunctionOnEnergyEvaluator.tf_p2 = {
    reco_particle = y2.reco_p4,
    gen_particle = 'looper::particles/1',
    sigma = 0.01,
}

GaussianTransferFunctionOnEnergyEvaluator.tf_p4 = {
    reco_particle = b2.reco_p4,
    gen_particle = 'looper::particles/2',
    sigma = 0.10,
}

-- Block A produce solutions with two particles
full_inputs = { y1.gen_p4, 'looper::particles/1', b1.gen_p4,
    'looper::particles/2' }

BuildInitialState.initialState = {
    do_transverse_boost = false,
    particles = full_inputs,
}

```

```

-- Jacobians resulting from TF and change of variables
jacobians = {'tf_p1::TF_times_jacobian', 'tf_p2::TF', 'tf_p3::TF_times_jacobian',
'tf_p4::TF', 'phaseSpaceOut::phase_space', 'looper::jacobian'}

-- Define Matrix Element
MatrixElement.bbyy = {
  pdf = 'NNPDF23_lo_as_0130_qed',
  pdf_scale = parameter('H_mass'),

  matrix_element = 'gg2HH2bbyy_SM_EFT_FF_bt_haa_kappa_P1_Sigma_SM_EFT_FF_bt
_haa_kappa_gg_aabbx', -- signal ME
  -- matrix_element = 'pp2bbyy_SM_EFT_FF_bt_haa_kappa_P1_Sigma_SM_EFT_FF_bt
_haa_kappa_gg_aabbx', -- bkg ME
  matrix_element_parameters = {
    card = '../bbyy/MatrixElement/Cards/param_card.dat' },

  override_parameters = { mdl_kappa = parameter('kappa') },

  initialState = 'initialState::partons',

  particles = {
    inputs = full_inputs,
    ids = {
      {
        pdg_id = 22,
        me_index = 1,
      },{
        pdg_id = 22,
        me_index = 2,
      },{
        pdg_id = 5,
        me_index = 3,
      },{
        pdg_id = -5,
        me_index = 4,
      },
    },
    jacobians = jacobians
  }
}

DoubleLooperSummer.integrand = { input = "bbyy::output" }

-- End of loop

integrand("integrand::sum") -- Define the integrand

```

D.2 Lua configuration file for the single Higgs boson ttH background

The Lua configuration file presented in this section is the one used for the single Higgs boson ttH Matrix Element.

```
-- LUA file for pp -> ttH, H->aa, t->b(visible)W(invisible)
-- Numbering of particles:
-- 1: y      -- 4: b      -- 34: t
-- 2: y      -- 5: W-     -- 56: t~
-- 3: W+     -- 6: b~
```

```
load_modules("../bbyy/MatrixElement/build/libme_tth.so")
```

```
-- Register inputs
local y1 = declare_input("y1")
local y2 = declare_input("y2")
local b1 = declare_input("b1")
local b2 = declare_input("b2")
```

```
-- Insert a Permutator module
add_reco_permutations(b1, b2)
```

```
parameters = {
    energy = 14000.,
    H_mass = 125.,
    H_width = 0.005753088,
    t_mass = 173.,
    t_width = 1.4915,
}
```

```
cuba = {
    algorithm='divonne',
    relative_accuracy = 0.005,
    verbosity = 3,
}
```

```
BreitWignerGenerator.flatter_s34 = {
    ps_point = add_dimension(),
    mass = parameter('t_mass'),
    width = parameter('t_width')
}
```

```
BreitWignerGenerator.flatter_s56 = {
    ps_point = add_dimension(),
    mass = parameter('t_mass'),
    width = parameter('t_width')
}
```

```

GaussianTransferFunctionOnEnergy.tf_p1 = {
    ps_point = add_dimension(),
    reco_particle = y1.reco_p4,
    sigma = 0.01,
}
y1.set_gen_p4("tf_p1::output")

GaussianTransferFunctionOnEnergy.tf_p2 = {
    ps_point = add_dimension(),
    reco_particle = y2.reco_p4,
    sigma = 0.01,
}
y2.set_gen_p4("tf_p2::output")

GaussianTransferFunctionOnEnergy.tf_p3 = {
    ps_point = add_dimension(),
    reco_particle = b1.reco_p4,
    sigma = 0.10,
}
b1.set_gen_p4("tf_p3::output")

GaussianTransferFunctionOnEnergy.tf_p4 = {
    ps_point = add_dimension(),
    reco_particle = b2.reco_p4,
    sigma = 0.10,
}
b2.set_gen_p4("tf_p4::output")

StandardPhaseSpace.phaseSpaceOut = {
    particles = { y1.gen_p4, y2.gen_p4, b1.gen_p4, b2.gen_p4 } -- only on visible particles
}

BlockF.blockf = {
    p3 = b1.gen_p4,
    p4 = b2.gen_p4,

    s13 = 'flatter_s34::s',
    s24 = 'flatter_s56::s',

    q1 = add_dimension(),
    q2 = add_dimension(),

    branches = { y1.gen_p4, y2.gen_p4 },

    m1=79.824360,
    m2=79.824360,
}

-- Loop over block solutions

Looper.looper = {

```

```

    solutions = "blockf::solutions",
    path = Path("initialState", "tth", "integrand")
}

-- Block F produce solutions with two particles
full_inputs = { y1.gen_p4, y2.gen_p4, 'looper::particles/1', b1.gen_p4,
'looper::particles/2', b2.gen_p4 }

BuildInitialState.initialState = {
    do_transverse_boost = true,
    particles = full_inputs,
}

jacobians = {'tf_p1::TF_times_jacobian', 'tf_p2::TF_times_jacobian',
'tf_p3::TF_times_jacobian', 'tf_p4::TF_times_jacobian', 'phaseSpaceOut::phase_space',
'flatter_s34::jacobian', 'flatter_s56::jacobian', 'looper::jacobian'}

MatrixElement.tth = {
    pdf = 'NNPDF23_lo_as_0130_qed',
    pdf_scale = parameter('H_mass'),

    matrix_element = 'tth_SM_EFT_FF_bt_haa_kappa_P1_Sigma_SM_EFT_FF_bt_haa_kappa_gg_aawpbwmbx',
    matrix_element_parameters = {
        card = '../bbyy/MatrixElement/Cards/param_card.dat',
    },

    initialState = 'initialState::partons',

    particles = {
        inputs = full_inputs,
        ids = {
            -- y1
            {
                pdg_id = 22,
                me_index = 1,
            },
            -- y2
            {
                pdg_id = 22,
                me_index = 2,
            },
            -- W+
            {
                pdg_id = 24,
                me_index = 3,
            },
            -- b1
            {
                pdg_id = 5,
                me_index = 4,
            },
        },
    },
}

```

```
-- W-
{
  pdg_id = -24,
  me_index = 5,
},
-- b2
{
  pdg_id = -5,
  me_index = 6,
},
}

jacobians = jacobians
}

DoubleLooperSummer.integrand = { input = "tth::output" }

-- End of loop
integrand("integrand::sum")
```

Appendix E

Event yield likelihood and extended likelihood

Given the strong dependence of the di-Higgs cross-section production on κ (see Fig. 2.2), the number of events observed after a given set of event selection requirements is applied is a valuable piece of information for the extraction of κ . This chapter provides mathematical details about the event yield likelihood and its combination with the event kinematics likelihood into the so-called extended likelihood.

E.1 Event yield likelihood

The event sample is composed of candidate di-Higgs events that passed the selection requirements given in Sec. 5.3. Some of these events are true di-Higgs events and the rest are background processes. In a cut-based analysis, knowing the expected number of events N_{bkg} corresponding to all background processes, the expected number of events $N_{\text{sig}}(\kappa)$ corresponding to the signal process at different values of κ , and the observed total number of events N_{obs} makes it possible to extract the most plausible value of κ . Assuming that the number of observed events follows a Poisson distribution, the event yield likelihood L_{yield} is the Poisson probability with parameter $N_{\text{bkg}} + N_{\text{sig}}(\kappa)$ to observe N_{obs} events:

$$L_{\text{yield}}(\kappa) = e^{-(N_{\text{bkg}} + N_{\text{sig}}(\kappa))} \frac{(N_{\text{bkg}} + N_{\text{sig}}(\kappa))^{N_{\text{obs}}}}{N_{\text{obs}}!} \quad (\text{E.1})$$

E.2 Extended likelihood

Starting with the unnumbered equation before Eq. (8) of Ref. [53] and using the conventions presented in Sec. 3.1, the extended likelihood, taking both event kinematics and event yield into account, is the product of the event kinematics likelihood and the event yield likelihood:

$$L_{\text{extended}} = \underbrace{\left(\prod_{i=1}^N L_{\text{event}}^i(\kappa | \mathbf{x}^i) \right)}_{\text{kinematics likelihood } L_{\text{sample}}(\kappa | \mathbf{x})} \times \underbrace{e^{-(N_{\text{bkg}} + N_{\text{sig}}(\kappa))} \frac{(N_{\text{bkg}} + N_{\text{sig}}(\kappa))^{N_{\text{obs}}}}{N_{\text{obs}}!}}_{\text{event yield likelihood } L_{\text{yield}}(\kappa)} \quad (\text{E.2})$$

The negative logarithm of the extended likelihood reads:

$$-\ln(L_{\text{extended}}) = -\ln(L_{\text{sample}}(\kappa | \mathbf{x})) + (N_{\text{bkg}} + N_{\text{sig}}(\kappa)) - N_{\text{obs}} \ln(N_{\text{bkg}} + N_{\text{sig}}(\kappa)) + cst \quad (\text{E.3})$$

where cst is a constant term can be freely removed since it does not depend on κ .

Appendix F

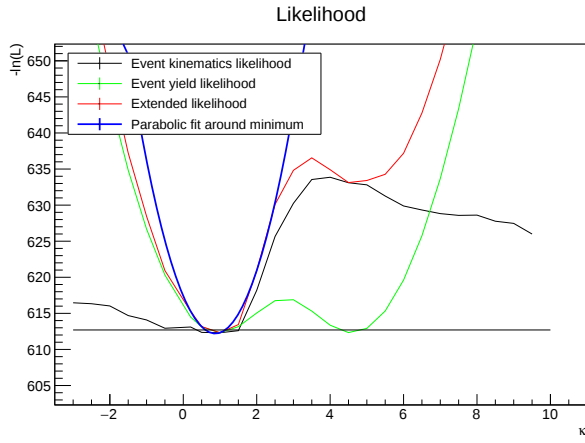
Integration of the likelihood: convergence study

When making the scans of the event kinematics likelihood, very fast variations or discontinuities were observed (especially for the signal hypothesis), whereas the likelihood was expected to be smooth. This is shown on the left-hand side plots of Fig. F.1. This has been traced back to an imperfect numerical integration. This imperfection is easy to spot in a plot of the evolution of the estimate of the integral with the number of integrand evaluations (see Fig. F.2): instead of having uncertainties on the estimate becoming smaller and smaller, and estimates that stay stable at the real value of the integral, the uncertainties are very small from the beginning of the integration, and the estimate of the integral increases nearly linearly with the integrand evaluations and does not reach at stable value. When this happens for an event, the integration for every value of κ is affected, see Fig. F.3.

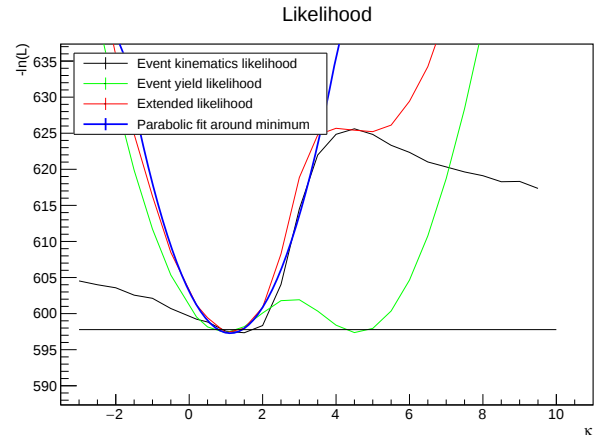
For some events affected by this numerical issue, the way the estimate of the integral increases is the same for all values of κ . This results in an underestimated value of the likelihood, but the fractional underestimate is roughly the same for all values of κ . The scan of the likelihood as a function of κ is therefore smooth. However, for other events, the computations for different values of κ are differently affected by this numerical issue. Just as in the first case, the likelihood for the signal hypothesis is underestimated, but unlike it, it is not underestimated by the same ratio for every value of κ . The likelihood as a function of κ is not smooth and presents fast variations at values of κ differently affected by this numerical issue. The two cases are presented in Figs. F.3 and F.4.

Right-hand side plots of Fig. F.1 present the likelihood as a function of κ when these problematic events are removed and replaced by unproblematic events. The likelihood curve is now smooth. For the two sets of events presented in this plot, the estimation of κ is consistent with the simulated value of κ : $\kappa = 1.12 \pm 0.33$ for the first sample and $\kappa = 1.29 \pm 0.33$ for the second.

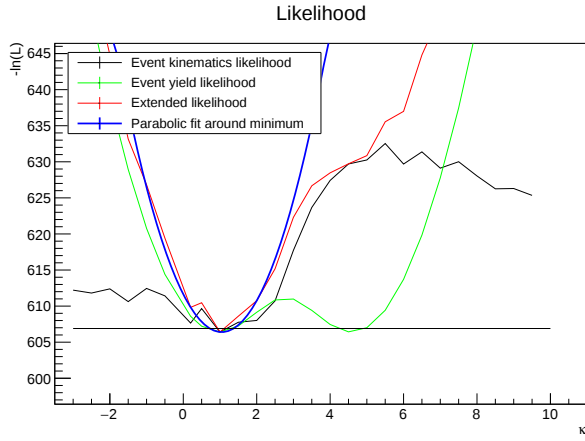
Due to the lack of time to correct for this numerical issue, no other integration algorithm or integration parameters was used to compute again the incorrect likelihoods and the events affected (which do not introduce discontinuities in the likelihood scan) were not removed from the samples used to draw the plots of the rest of this document. It seems that the phase-space region where the integrand takes large value was not successfully found by the VEGAS algorithm. One important hypothesis for the space-space mapping of MoMEMta and for VEGAS to work correctly is that, in the case of a very sharp integrand, the strength of each narrow peak in the integrand is associated with a single variable that in turn can be mapped onto one



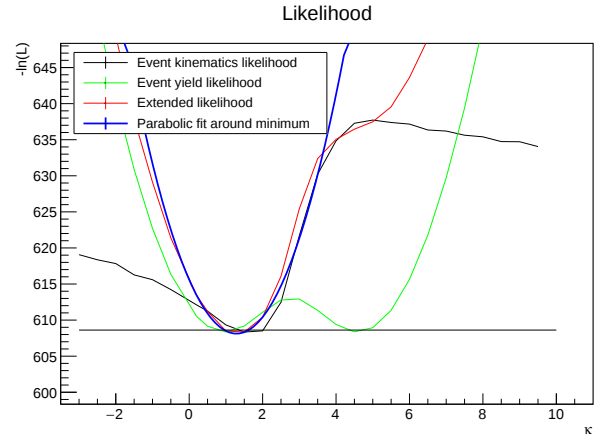
(a) Likelihoods for a first set of 16 events without removing problematic events.



(b) Likelihoods for a first set of 16 events with removing problematic events: 11 events out of the 16 ones of the left figure were kept and 5 problematic events were replaced by non-problematic ones.



(c) Likelihoods for a second and different set of 16 signal events without removing problematic events.



(d) Likelihoods for a second and different set of 16 signal events with removing problematic events: 13 events out of the 16 ones of the left figure were kept and 3 problematic events were replaced by non-problematic ones.

Figure F.1: Event kinematics likelihood (black), event yield likelihood (green) and extended likelihood (red) as a function of κ for different sets of 16 signal events (each row corresponding to a different set) without (left) or with (right) removing events having an event kinematics likelihood presenting fast variations (referred to as problematic events in the subcaptions). An arbitrary offset has been applied to all likelihood curves such that they are on the same scale. The black horizontal line is the line at 0.5 above the minimum and yields the measured uncertainty.

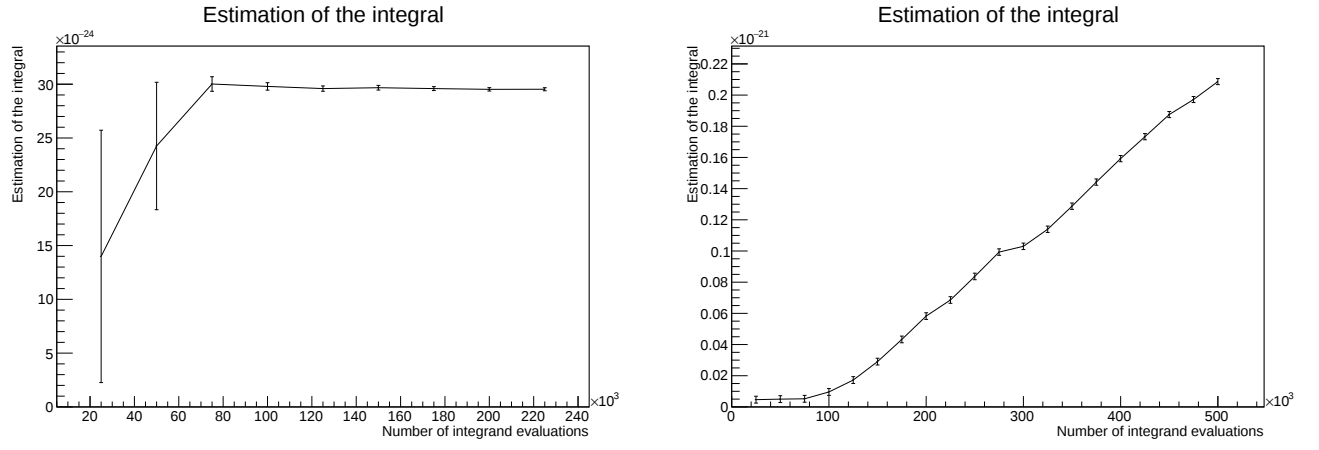
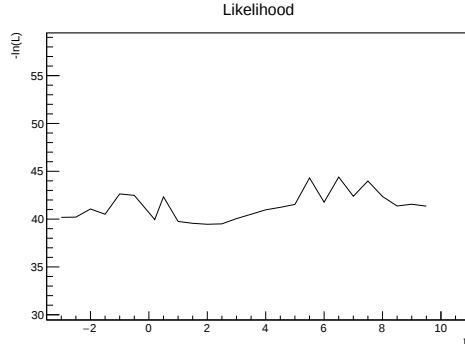
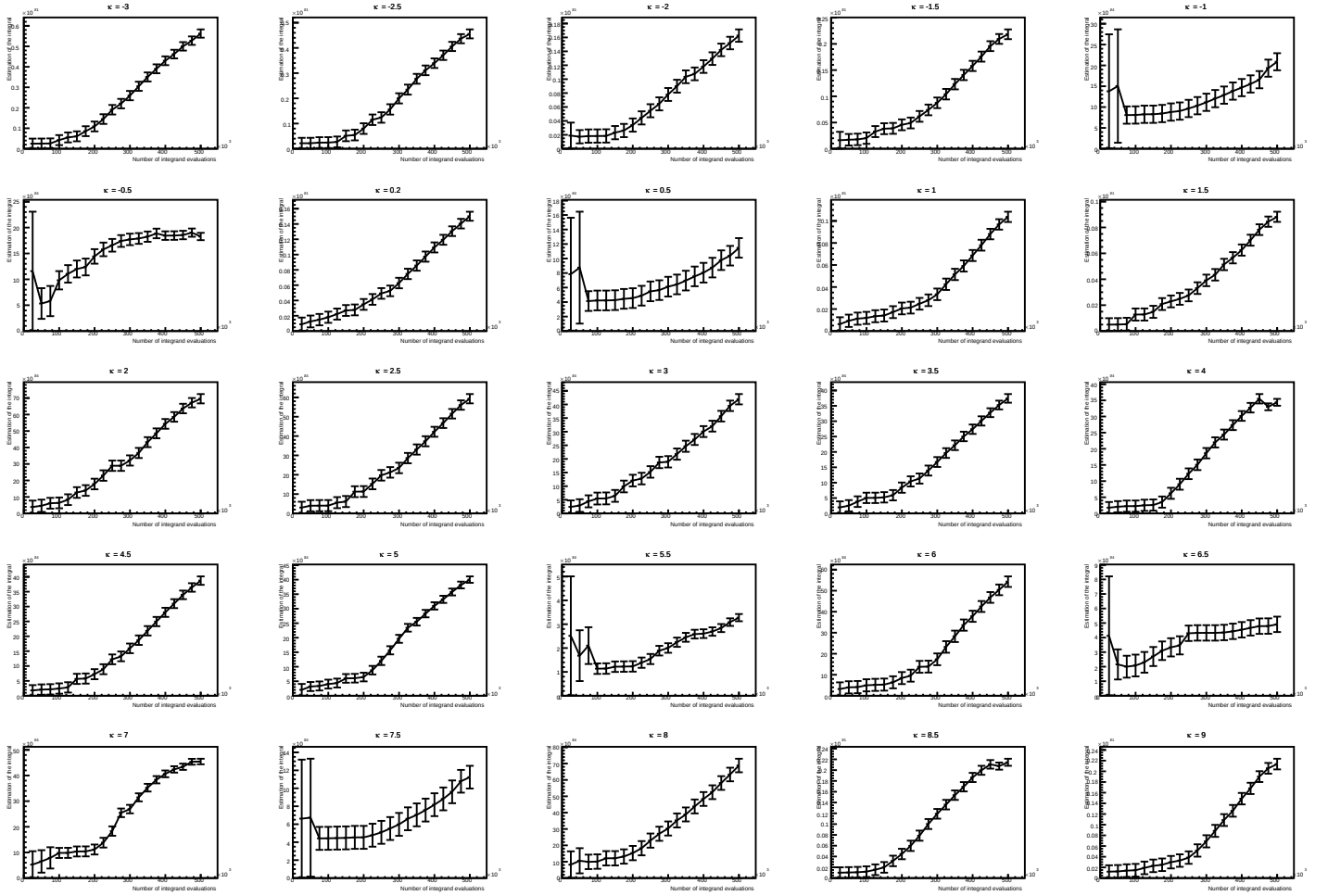


Figure F.2: Correct integration (left) and problematic integration (right) of the likelihood of two signal events analysed with the signal hypothesis at $\kappa = 1$.

variable of integration [21]. It might be possible that this condition is not entirely fulfilled for some events.

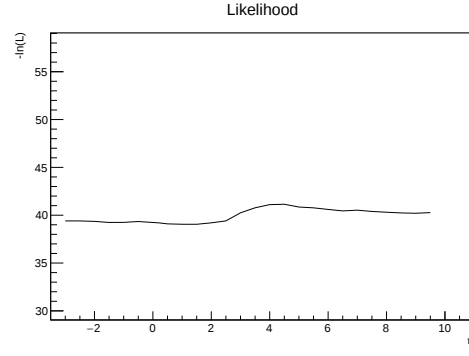


(a) Likelihood as a function of κ for the signal event number 36.

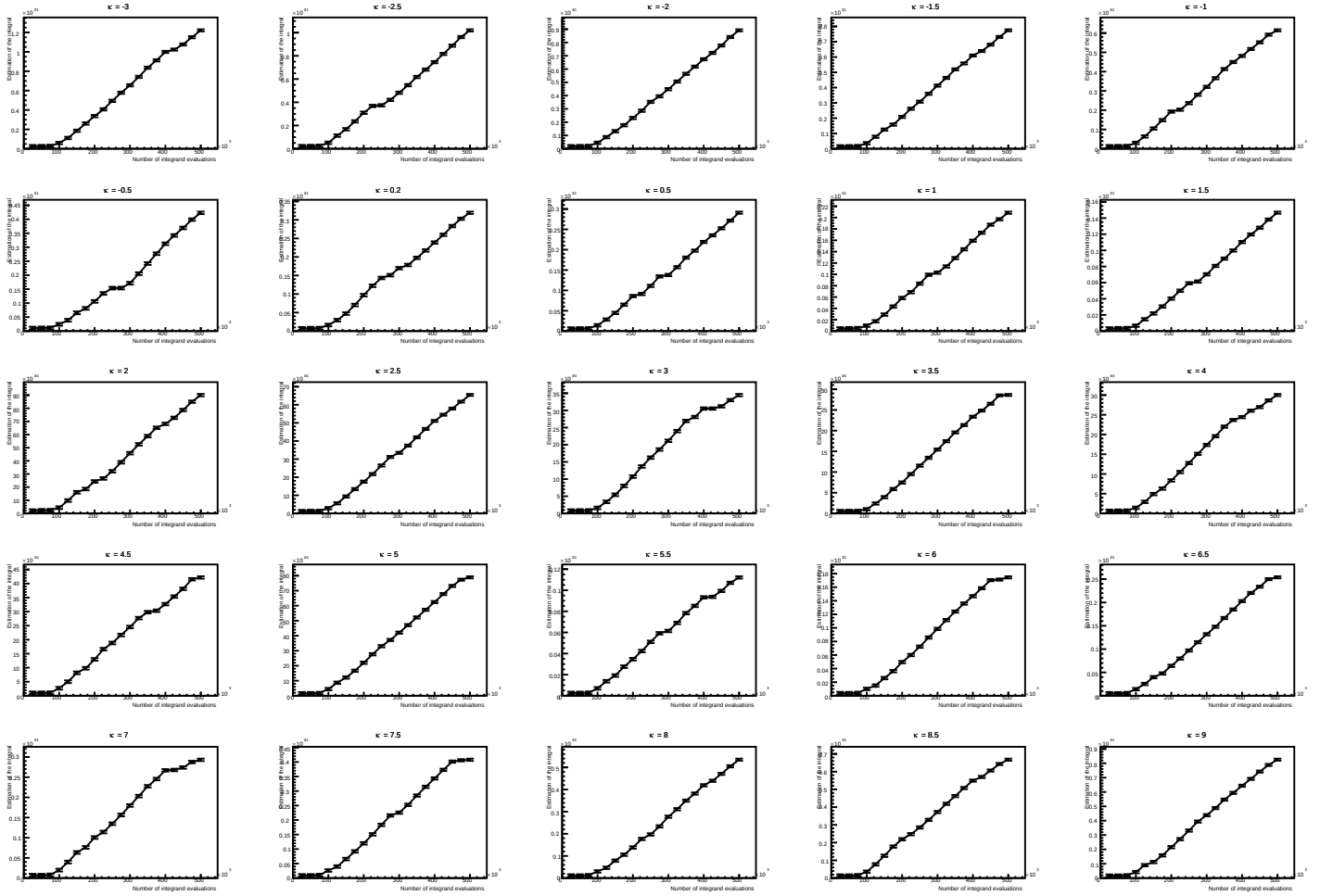


(b) Numerical integration of the likelihood of the signal event number 36 for all values of κ : each plot represents the estimate of the integral as a function of the number of integrand evaluations for one value of κ .

Figure F.3: Problematic integration of the likelihood for all values of κ for one signal event, with discontinuities in the likelihood curve as a function of κ . The computations for the different values of κ are affected differently by this numerical issue resulting in fast variations in the likelihood curve as a function of κ .



(a) Likelihood as a function of κ for the signal event number 32.



(b) Numerical integration of the likelihood of the signal event number 32 for all values of κ : each plot represents the estimate of the integral as a function of the number of integrand evaluations for one value of κ .

Figure F.4: Problematic integration of the likelihood for all values of κ for one signal event, without discontinuities in the likelihood curve as a function of κ . The computations for the different values of κ are affected similarly by this numerical issue resulting in a smooth likelihood curve as a function of κ .

Appendix G

Performance of VEGAS and DIVONNE integration algorithms

Given the apparent convergence issues with VEGAS discussed in Appendix F, we got interested in testing the DIVONNE algorithm, due to its conceptual differences with VEGAS. We could expect DIVONNE not to face the same issue as VEGAS and maybe to provide better results. Before trying out DIVONNE, VEGAS was run with different parameters, but this did not solve the issue. In order to assert that some results obtained with VEGAS were underestimated and to assess whether DIVONNE was doing better or not, we used the basic Monte-Carlo (MC) algorithm without any importance sampling and compared the three algorithms: provided that it is possible to integrate the integrand long enough, the basic Monte-Carlo algorithm must yield the true value of the integral. No additional code was used for the integration with the basic Monte-Carlo method: it was emulated by VEGAS with a very large number of initial evaluations of the integrand and a maximum number of evaluations of the integrand equal to it (to limit VEGAS to just one iteration); as VEGAS starts from a uniform grid, phase-space points are randomly chosen in all the phase-space.

A set of events for which the integration with VEGAS was "problematic" (*i.e.* events for which convergence plot shows that the final value was not stable or that the error on the estimate of the integral was constantly huge) as well as a set of non-problematic events were selected. These two sets were integrated using DIVONNE (with default parameters) and the basic Monte-Carlo method.

Table G.1 summarises the computation of the signal likelihood of a set of 10 problematic signal events. The events can be divided into four groups:

- Those for which the estimate of the likelihood of the 3 algorithms are compatible.
- Those for which the estimate of the likelihood of the 3 algorithms are of the same order of magnitude.
- Those for which DIVONNE and the basic Monte-Carlo provides the same estimate, which is different from the estimate of VEGAS.
- Those for which the 3 estimates are different.

For the events of the first and second group, it seems a coincidence that the result of VEGAS was close to the results of DIVONNE and the basic MC. The large number of events in the third group proves that DIVONNE generally provides an estimate of the integral that is close to the true value. Furthermore, the results of VEGAS are underestimated, as we expected

Event number	VEGAS	DIVONNE	BASIC MC
1	1.01×10^{-21} $\pm 5.00 \times 10^{-23}$	1.08×10^{-21} $\pm 1.38 \times 10^{-24}$	1.17×10^{-21} $\pm 3.52 \times 10^{-22}$
2	4.19×10^{-22} $\pm 1.98 \times 10^{-24}$	6.74×10^{-22} $\pm 3.24 \times 10^{-24}$	6.35×10^{-22} $\pm 2.47 \times 10^{-22}$
3	3.29×10^{-22} $\pm 1.57 \times 10^{-24}$	5.08×10^{-22} $\pm 1.41 \times 10^{-24}$	3.89×10^{-22} $\pm 1.27 \times 10^{-22}$
4	2.53×10^{-22} $\pm 1.07 \times 10^{-24}$	8.15×10^{-22} $\pm 2.55 \times 10^{-24}$	4.42×10^{-22} $\pm 1.24 \times 10^{-22}$
5	2.09×10^{-22} $\pm 1.93 \times 10^{-24}$	9.18×10^{-22} $\pm 2.39 \times 10^{-24}$	9.39×10^{-22} $\pm 3.43 \times 10^{-22}$
6	4.18×10^{-21} $\pm 1.77 \times 10^{-22}$	1.15×10^{-20} $\pm 2.14 \times 10^{-22}$	2.37×10^{-20} $\pm 8.97 \times 10^{-21}$
7	2.55×10^{-23} $\pm 1.16 \times 10^{-24}$	3.28×10^{-21} $\pm 1.27 \times 10^{-23}$	1.75×10^{-21} $\pm 6.48 \times 10^{-22}$
8	5.73×10^{-23} $\pm 1.18 \times 10^{-24}$	6.78×10^{-22} $\pm 8.54 \times 10^{-25}$	3.80×10^{-22} $\pm 1.33 \times 10^{-22}$
9	6.81×10^{-21} $\pm 6.86 \times 10^{-23}$	9.49×10^{-21} $\pm 2.74 \times 10^{-23}$	1.41×10^{-20} $\pm 3.48 \times 10^{-21}$
10	1.03×10^{-22} $\pm 4.67 \times 10^{-24}$	1.07×10^{-23} $\pm 7.37 \times 10^{-25}$	1.38×10^{-21} $\pm 1.09 \times 10^{-21}$

Table G.1: Computation of the signal- $\kappa=1$ likelihood of several signal events, for which the integration with VEGAS seemed problematic, with VEGAS, DIVONNE and the basic Monte-Carlo with 15 million integrand evaluations.

Event number	VEGAS	DIVONNE	BASIC MC
1	1.52×10^{-19} $\pm 1.46 \times 10^{-21}$	1.60×10^{-19} $\pm 2.79 \times 10^{-22}$	1.63×10^{-19} $\pm 1.33 \times 10^{-20}$
2	2.53×10^{-20} $\pm 2.03 \times 10^{-22}$	2.62×10^{-20} $\pm 5.53 \times 10^{-23}$	2.32×10^{-20} $\pm 2.47 \times 10^{-21}$
3	6.82×10^{-23} $\pm 3.23 \times 10^{-25}$	6.83×10^{-23} $\pm 3.13 \times 10^{-25}$	9.02×10^{-23} $\pm 3.46 \times 10^{-23}$

Table G.2: Computation of the signal- $\kappa=1$ likelihood of several signal events, for which the integration with VEGAS seemed correct, with VEGAS, DIVONNE and the basic Monte-Carlo with 15 million integrand evaluations.

based on the convergence plots. For the event 10, it seems that DIVONNE did not find the correct value at all, since the results of VEGAS and the basic MC are different from the result of DIVONNE and are larger.

Table G.2 summarises the computation of the signal likelihood of a set of 3 non-problematic signal events. The results of the 3 algorithms are always compatible.

As a preliminary conclusion based on a small number of events and sizable statistical uncertainties in the basic MC results, VEGAS often underestimates the signal likelihood of signal events while DIVONNE almost always succeeds in integrating these likelihoods.

Bibliography

- [1] P. W. Higgs, “*Broken Symmetries and the Masses of Gauge Bosons*”, [Phys. Rev. Lett. 13, 508 \(1964\)](#) .
- [2] F. Englert and R. Brout, “*Broken Symmetry and the Mass of Gauge Vector Mesons*”, [Phys. Rev. Lett. 13, 321 \(1964\)](#) .
- [3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “*Global Conservation Laws and Massless Particles*”, [Phys. Rev. Lett. 13, 585 \(1964\)](#) .
- [4] A. Noble and M. Perelstein, “*Higgs self-coupling as a probe of electroweak phase transition*”, [Phys. Rev. D 78, 063518 \(2008\)](#) , [arXiv:0711.3018 \[hep-ph\]](#).
- [5] The CMS Collaboration, “*Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV*”, Submitted to: Eur. Phys. J., [arXiv:1809.10733 \[hep-ex\]](#).
- [6] The ATLAS Collaboration, “*Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment*”, ATLAS-CONF-2018-031 (2018).
- [7] The ATLAS Collaboration, “*Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*”, [Journal of High Energy Physics 2016 8, 45 \(2016\)](#) .
- [8] CERN, “[The High-Luminosity LHC \(HL-LHC\) Project](#)”, May 2016. Accessed: Feb. 2019.
- [9] The ATLAS Collaboration, “*Study of the double Higgs production channel $H(\rightarrow b\bar{b})H(\rightarrow \gamma\gamma)$ with the ATLAS experiment at the HL-LHC*”, [ATL-PHYS-PUB-2017-001](#), Jan. 2017.
- [10] The ATLAS Collaboration, “*Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC*”, [ATL-PHYS-PUB-2018-053](#), Dec. 2018.
- [11] M. Reichert, A. Eichhorn, H. Gies, P. Pawlowski, T. Plehn, and M. Scherer, “*Probing baryogenesis through the Higgs boson self-coupling*”, [Phys. Rev. D 97, 075008 \(2018\)](#) .
- [12] J. Neyman and E. Pearson, “*On the problem of the most efficient tests of statistical hypotheses*”, [Phil. Trans. R. Soc. Lond. A 231 694-706, 289 \(1933\)](#) .
- [13] K. Kondo, “*Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum I. Method and Toy Models*”, [J. Phys. Soc. Jpn. 57, 4126 \(1988\)](#) .

- [14] K. Kondo, “*Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum II. Mass Spectra for $2 \rightarrow 2$ Processes*”, *J. Phys. Soc. Jpn.* **60**, 836 (1991) .
- [15] R. Dalitz and G. Goldstein, “*Decay and polarization properties of the top quark*”, *Phys. Rev. D* **45**, 1531 (1992) .
- [16] V. M. Abazov *et al.*, “*A precision measurement of the mass of the top quark*”, *Nature* **429**, 638 (2004) , [arXiv:0406031 \[hep-ex\]](#).
- [17] V. M. Abazov *et al.*, “*Helicity of the W boson in lepton + jets $t\bar{t}$ events*”, *Phys. Lett. B* **617**, 1 (2005) , [arXiv:0404040 \[hep-ex\]](#).
- [18] V. M. Abazov *et al.*, “*Measurement of the Top Quark Mass Using the Matrix Element Technique in Dilepton Final States*”, *Phys. Rev. D* **94** 3, 032004 (2016) , [arXiv:1606.02814 \[hep-ex\]](#).
- [19] F. Fiedler, A. Grohsjean, P. Haefner, and P. Schieferdecker, “*The Matrix Element Method and its Application in Measurements of the Top Quark Mass*”, *Nucl. Instrum. Meth. A* **624**, 203-218 (2010) , [arXiv:1003.1316v2 \[hep-ex\]](#).
- [20] P. Artoisenet and O. Mattelaer, “*MadWeight: Automatic event reweighting with matrix elements*”, in *Proceedings, 2nd International Workshop on Prospects for charged Higgs discovery at colliders (CHARGED 2008): Uppsala, Sweden, September 16-19, 2008, PoS CHARGED2008*, 025 (2008) .
- [21] P. Artoisenet, V. Lemaître, F. Maltoni, and O. Mattelaer, “*Automation of the matrix element reweighting method*”, *JHEP* **12**, 068 (2010) , [arXiv:1007.3300v2 \[hep-ph\]](#).
- [22] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “*The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*”, *JHEP* **07**, 079 (2014) , [arXiv:1405.0301 \[hep-ph\]](#).
- [23] S. Brochet, C. Delaere, B. François, V. Lemaître, A. Mertens, A. Saggio, M. Vidal Marono, and S. Wertz, “*MoMEMta, a modular toolkit for the Matrix Element Method at the LHC*”, *Eur. Phys. J. C* **79** 2, 126 (2019) , [arXiv:1805.08555 \[hep-ph\]](#).
- [24] The ATLAS Collaboration, “*Prospects for measuring Higgs pair production in the channel $H(\rightarrow \gamma\gamma)H(\rightarrow b\bar{b})$ using the ATLAS detector at the HL-LHC*”, *ATL-PHYS-PUB-2014-019*, Oct. 2014.
- [25] M. Spira, “HPAIR.” [Web Site of M. Spira](#), Accessed: March 2019.
- [26] D. de Florian and J. Mazzitelli, “*Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*”, *Phys. Rev. Lett.* **111**, 201801 (2013) , [arXiv:1309.6594 \[hep-ph\]](#).
- [27] J. Grigo, K. Melnikov, and M. Steinhauser, “*Virtual corrections to Higgs boson pair production in the large top quark mass limit*”, *Nucl. Phys. B* **888**, 17 (2014) , [arXiv:1408.2422 \[hep-ph\]](#).
- [28] CERN, “*SM Higgs Branching Ratios and Total Decay Widths (update in CERN Report4 2016)*”, (2016). Accessed: Feb. 2019.

- [29] S. Borowka, N. Greiner, G. Heinrich, S. Jones, M. Kerner, J. Schlenk, U. Schubert, and T. Zirke, “*Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence*”, *Phys. Rev. Lett.* **117** 1, 012001 (2016) , [arXiv:1604.06447 \[hep-ph\]](#).
- [30] E. W. N. Glover and J. J. van der Bij, “*Higgs boson pair production via gluon fusion*”, *Nucl. Phys. B* **309**, 282 (1988) .
- [31] S. Dawson, S. Dittmaier, and M. Spira, “*Neutral Higgs boson pair production at hadron colliders: QCD corrections*”, *Phys. Rev. D* **58**, 115012 (1998) , [arXiv:9805244 \[hep-ph\]](#).
- [32] D. de Florian, M. Grazzini, C. Hanga, S. Kallweit, J. M. Lindert, P. Maierhöfer, J. Mazzitelli, and D. Rathlev, “*Differential Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD*”, *JHEP* **09**, 151 (2016) , [arXiv:1606.09519 \[hep-ph\]](#).
- [33] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzitelli, “*Higgs boson pair production at NNLO with top quark mass effects*”, *JHEP* **05**, 059 (2018) , [arXiv:1803.02463 \[hep-ph\]](#).
- [34] F. Maltoni, E. Vryonidou, and M. Zaro, “*Top-quark mass effects in double and triple Higgs production in gluon-gluon fusion at NLO*”, *JHEP* **11**, 079 (2014) , [arXiv:1408.6542 \[hep-ph\]](#).
- [35] G. Lepage, “*A New Algorithm for Adaptive Multidimensional Integration*”, *J. Comput. Phys.* **27**, 192 (1978) .
- [36] S. Brandt, *Data Analysis: Statistical and Computational Methods for Scientists and Engineers*. Springer International Publishing, Feb. 2014, pp. 187–199.
- [37] M. A. Shifman, A. I. Vainshtein, M. B. Voloshin, and V. I. Zakharov, “*Low-Energy Theorems for Higgs Boson Couplings to Photons*”, *Sov. J. Nucl. Phys.*, **30**, 711.
- [38] R. Frederix, “[Higgs Effective couplings to gluons \(and photons\)](#)”, Oct. 2007. Accessed: Feb. 2019.
- [39] T. Plehn, M. Spira, and P. Zerwas, “*Pair production of neutral Higgs particles in gluon-gluon collisions*”, *Nucl. Phys. B* **479**, 46 (1996) , [arXiv:9603205 \[hep-ph\]](#).
- [40] F. Maltoni and T. Stelzer, “*MadEvent: Automatic event generation with MadGraph*”, *JHEP* **02**, 027 (2003) , [arXiv:0208156 \[hep-ph\]](#).
- [41] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, and T. Reiter, “*UFO - The Universal FeynRules Output*”, *Comput. Phys. Commun.* **183**, 1201 (2012) , [arXiv:1108.2040 \[hep-ph\]](#).
- [42] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, “*MadGraph/MadEvent v4: The New Web Generation*”, *JHEP* **09**, 028 (2007) , [arXiv:0706.2334 \[hep-ph\]](#).
- [43] J. Stark private communication, Feb. 2019.
- [44] J. Letts and P. Mattig, “*Direct determination of the CKM matrix from decays of W bosons and top quarks at high-energy e^+e^- colliders*”, *Eur. Phys. J. C* **21**, 211 (2001) , [arXiv:0102004 \[hep-ph\]](#).

- [45] J. Stark, private communication, Feb. 2019.
- [46] J. Thaler, “[How to Read LHC Olympics Data Files](#)”, Dec. 2006. Accessed: Feb. 2019.
- [47] O. Mattelaer, private communication, Dec. 2018.
- [48] F. Maltoni and T. Stelzer, “*MadEvent: Automatic event generation with MadGraph*”, [JHEP 02, 027 \(2003\)](#) , [arXiv:0208156 \[hep-ph\]](#).
- [49] T. Hahn, “*CUBA: A Library for multidimensional numerical integration*”, [Comput. Phys. Commun. 168, 78 \(2005\)](#) , [arXiv:0404043 \[hep-ph\]](#).
- [50] O. Mattelaer, “[Form Factors](#)”, Nov. 2017. Accessed: Feb. 2019.
- [51] The DELPHES 3 collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi, “*DELPHES 3: a modular framework for fast simulation of a generic collider experiment*”, [Journal of High Energy Physics 2014 2, 57 \(2014\)](#) .
- [52] A. Lleres, private communication, March 2019.
- [53] R. J. Barlow, “*Extended maximum likelihood*”, [Nucl. Instrum. Meth. A 297, 496 \(1990\)](#) .