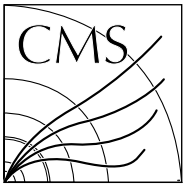
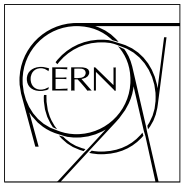


The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



14 June 2006

Search Strategy for the Standard Model Higgs Boson in the $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\mu$ Decay Channel using $M(4\mu)$-Dependent Cuts.

S. Abdullin[2], D. Acosta[3], P. Bartalini[3], R. Cavanaugh[3], A. Drozdetskiy[1,3], A. Korytov[3], G. Mitselmakher[3], Yu. Pakhotin[3], B. Scurlock[3], and A. Sherstnev[4]

CMS collaboration

Abstract

We present a strategy for a Higgs-boson search in its four-muon decay channel $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\mu$. The method is based on using optimized, smooth $M(4\mu)$-dependent cuts that automatically ensure an optimal signal-to-background ratio for any mass at which the Higgs boson might appear. The Higgs boson then manifests itself as a 4μ resonance-like peak over the continuum $M(4\mu)$ distribution and can be searched for using various statistical techniques. The most important theoretical and instrumental systematic errors as well as the fact that the search is conducted in a broad range of $M(4\mu)$ invariant masses (110-600 GeV/c^2) are taken into account.

[1] Corresponding author, e-mail: Alexey.Drozdetskiy@cern.ch

[2] Fermi National Laboratory, Chicago, IL, USA

[3] University of Florida, Gainesville, FL, USA

[4] Cavendish Laboratory, University of Cambridge, UK

1 Introduction

The $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ process is one of the cleanest channels (also known as a "gold-plated" channel) for discovering the Standard Model Higgs boson at the LHC. In this paper, we outline a complete analysis strategy for discovering the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ channel. The explored range of Higgs masses is 115-600 GeV/c^2.

The cuts, smooth functions of the four-muon invariant mass $M(4\mu)$, are such that at whatever unknown a priori mass the Higgs boson might appear, the signal-to-background ratio is already optimal to give the best chance of discovering it. This allows one to avoid a posteriori cut optimization. We give a direct comparison of results obtained with $M(4\mu)$-dependent (dynamic) and $M(4\mu)$-independent (flat) cuts.

The search for the Higgs-boson 4μ resonance-like peak can be done using the Log-Likelihood Ratio (LLR) [1] built for the entire $M(4\mu)$-distribution, binned or unbinned, or taking a straightforward counting experiment approach. We give a direct comparison of the two approaches.

A full treatment of the most important theoretical and instrumental systematic errors and their effect on the evaluation of the significance of the Higgs-boson observation are presented. To minimize systematic errors, new methods of reconstructing the most important corrections directly from data were developed. Among them are the muon reconstruction and isolation cut efficiencies. We also show that by using the measured $Z \rightarrow 2\mu$ cross section or an event count in the sidebands of the $M(4\mu)$ distributions, one can substantially reduce a number of theoretical and instrumental systematic errors.

In addition we verify by how much the local excess significance should be effectively degraded due to the fact that we look for a narrow resonance in a broad range of $M(4\mu)$ invariant masses.

The results are obtained with the official CMS detector simulation and reconstruction software [2, 3] and include pile-up events corresponding to an instantaneous luminosity of $2 \times 10^{33}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

The final results are presented in terms of the required integrated luminosity for observing the Standard Model Higgs boson at 5σ and 3σ significance levels and 95% CL exclusion limits. Also, we present the significance for a fixed value of an integrated luminosity equal to $30\,fb^{-1}$ and 95% CL exclusion contours in the (M_H, σ) plane for integrated luminosities of 3, 10, and 30 fb^{-1}.

Previous studies on the search for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ channel with CMS are described in [4]. Another ongoing study exploring the discovery potential with a different set of mass-independent cuts can be found elsewhere [5]. The results of the two parallel analyses using the $H \rightarrow 4e$ and $H \rightarrow 2e2\mu$ channels can be found in [6, 7].

2 Physics Processes and Their Simulation

The Higgs boson event samples for 18 Higgs boson mass points (see Table 1) and the three main background processes, $t\bar{t}$, $(Z/\gamma^*)b\bar{b}$, and $Z/\gamma^*Z/\gamma^*$, were simulated using the full CMS detector simulation and reconstruction software. Many other plausible background candidates, $b\bar{b}b\bar{b}$, $b\bar{b}c\bar{c}$, $c\bar{c}c\bar{c}$, single-top, $Zc\bar{c}$, $Wb\bar{b}$, $Wc\bar{c}$, fake, and π/K decay muons in QCD, were considered and found to be negligible.

To save CPU time, only events with at least $2\mu^+$ and $2\mu^-$ in the pseudorapidity range $|\eta| < 2.4$ and with $p_T > 3$ GeV/c were retained for further analysis. Muons outside these kinematical limits could not be reconstructed by CMS. Additional cuts were applied to di-muon invariant masses for the Higgs-boson samples ($m(Z) > 5$ GeV/c^2) and for $Z/\gamma^*Z/\gamma^*$ and $(Z/\gamma^*)b\bar{b}$ samples ($m(\mu^+\mu^-) > 5$ GeV/c^2). (The first $\mu^+\mu^-$ pair in $Z/\gamma^*Z/\gamma^*$ and $(Z/\gamma^*)b\bar{b}$ samples was defined as the one with its invariant mass closest to $m(Z^0)$, while the second $\mu^+\mu^-$ pair was made out of the two remaining highest p_T muons of opposite signs.) All analysis cuts on these observables, to be described below, are much more stringent than these generator-level preselection cuts. The expected numbers of surviving 4μ events for signal and backgrounds for an integrated luminosity of $L = 30\ fb^{-1}$ are given in Table 1. The $M(4\mu)$ distribution of events after these cuts is shown in Figures 1 and 2.

2.1 Signal: $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$

The Higgs-boson samples were generated with Pythia 6.225 [8] (LO gluon and weak-boson fusion, $gg \rightarrow H$ and $q\bar{q} \rightarrow q\bar{q}H$), interfaced via CMKIN [9] version 310 (PDF CTEQ5L). Only decay channels $Z \rightarrow 2l$ (where l stands for e, μ, and τ) were considered. $Z \rightarrow q\bar{q} \rightarrow 2l$ were not included in the simulation: being very similar to the $(Z/\gamma^*)b\bar{b}$ background, those events would be suppressed together with the $(Z/\gamma^*)b\bar{b}$ background by our

Table 1: The LO/NLO cross sections for various Higgs-boson masses and backgrounds, corresponding number of events with four muons surviving the generator level preselection cuts (see section 2) calculated for $L = 30 fb^{-1}$, and the number of simulated events.

Process	σ_{LO}, pb	σ_{NLO}, pb	4μ events at L=30	Simulated Events
$pp \to H\,(m_{\rm H} = 115)$	-	47.7	7.69	10000
$pp \to H\,(m_{\rm H} = 120)$	-	44.3	13.6	10000
$pp \to H\,(m_{\rm H} = 130)$	-	38.4	31.1	9000
$pp \to H\,(m_{\rm H} = 140)$	-	33.7	49.2	10000
$pp \to H\,(m_{\rm H} = 150)$	-	29.8	54.1	9000
$pp \to H\,(m_{\rm H} = 160)$	-	26.6	25.6	9000
$pp \to H\,(m_{\rm H} = 170)$	-	23.9	12.3	10000
$pp \to H\,(m_{\rm H} = 180)$	-	21.6	28.5	9000
$pp \to H\,(m_{\rm H} = 190)$	-	19.7	101	10000
$pp \to H\,(m_{\rm H} = 200)$	-	18.0	109	10000
$pp \to H\,(m_{\rm H} = 250)$	-	12.4	87.5	10000
$pp \to H\,(m_{\rm H} = 300)$	-	9.58	72.3	10000
$pp \to H\,(m_{\rm H} = 350)$	-	9.12	72.6	9000
$pp \to H\,(m_{\rm H} = 400)$	-	8.81	63.4	9000
$pp \to H\,(m_{\rm H} = 450)$	-	6.44	45.1	10000
$pp \to H\,(m_{\rm H} = 500)$	-	4.46	31.8	10000
$pp \to H\,(m_{\rm H} = 550)$	-	3.07	22.6	10000
$pp \to H\,(m_{\rm H} = 600)$	-	2.13	16.3	10000
$pp \to \mathrm{t\bar{t}}$	-	840	7000	92236
$pp \to (Z/\gamma^*)\mathrm{b\bar{b}} \to 2\mu b\bar{b}$	116	278	8694	124500
$pp \to \mathrm{Z}/\gamma^*\mathrm{Z}/\gamma^* \to 4\mu$	0.113	see text	2622	118000
$pp \to \mathrm{Z}/\gamma^*\mathrm{Z}/\gamma^* \to 2\mu 2\tau$	0.157	see text	48.8	10000

analysis cuts. QED radiation from the final-state muons is modeled with PHOTOS [10]. Events were re-weighted to correspond to the total NLO cross-section $\sigma(pp \to H) \cdot BR(H \to ZZ) \cdot BR(Z \to 2l)^2$, where $\sigma(pp \to H)$ and $BR(H \to ZZ)$ were taken from [11] and $BR(Z \to 2l) = 0.101$ [12].

There exists an additional enhancement to the cross section for $H \to 4\mu$ (in comparison to $\sigma(pp \to H) \cdot BR(H \to ZZ) \cdot BR(Z \to 2\mu)^2$) due to interference of permutations of muons originating from different Z's. The corrections are calculated with CompHEP and presented in Figure 3.

The $M(4\mu)$ distribution for $m_{\rm H} = 140$ GeV/c^2 after event generator-level cuts is shown in Figures 1 and 2. The low-mass tail is mostly due to events where muons did not come directly from Z-decays (e.g., via τ-decays) and to internal bremsstrahlung that also tends to move the 4-muon invariant mass off the peak.

2.2 Background: $\mathrm{t\bar{t}}$

The $\mathrm{t\bar{t}}$ sample was generated with Pythia 6.225 (LO $gg \to \mathrm{t\bar{t}}$ and $q\bar{q} \to \mathrm{t\bar{t}}$), interfaced via CMKIN version 110 (PDF CTEQ5L). Only the decay channels $t \to Wb \to l\nu b$ were considered. Events were re-weighted to correspond to the total NLO cross-section $\sigma(pp \to \mathrm{t\bar{t}}) \cdot BR(W \to l\nu)^2$, where $\sigma(pp \to \mathrm{t\bar{t}}) = 840$ pb was taken from [13] and the branching ratio $BR(W \to l\nu) = 0.320$ [12].

2.3 Background: $(Z/\gamma^*)\mathrm{b\bar{b}} \to 2\mu b\bar{b}$

The $(Z/\gamma^*)\mathrm{b\bar{b}} \to \mu^+\mu^- b\bar{b}$ samples were generated with CompHEP 4.2p1 [14] matrix element generator (PDF CTEQ5L, with QCD scales $\mu_R = \mu_F = M_Z$, b-quark mass $m_b = 4.85$ GeV/c^2, and a di-muon mass cut $m(\mu^+\mu^-) > 5$ GeV/c^2), interfaced to PYTHIA 6.225 for showering and hadronization. Included sub-processes were $q\bar{q}/gg \to \mathrm{Z}/\gamma^* b\bar{b} \to \mu^+\mu^- b\bar{b}$, where q could be any of the light quarks, (u, d, s, c) (initial states with b quarks were also considered at the generator level and found to be negligible). No restriction on b decays was applied. The corresponding CompHEP LO cross section was found to be 116 pb. To obtain the NLO cross section given in Table 1, we calculated the NLO K-factor using MCFM [15]: $K_{NLO} = 2.4 \pm 0.3$. The conditions for the MCFM NLO and LO calculations were as follows: CTEQ6, $\mu_R^2 = \mu_F^2 = \hat{s}$, $m_b = 0$, $M(Z^{res}) > 5$ GeV/c^2, $p_T(b) > 5$ GeV/c, $|\eta_b| < 10$, $M(b\bar{b}) > 10$ GeV/c^2.

2.4 Backgrounds: $q\bar{q} \to \mathrm{Z}/\gamma^*\mathrm{Z}/\gamma^* \to 4\mu$ and $q\bar{q} \to \mathrm{Z}/\gamma^*\mathrm{Z}/\gamma^* \to 2\mu 2\tau$

These two event samples were generated with CompHEP 4.2p1 matrix element generator (PDF CTEQ5L, with QCD scales $\mu_R = \mu_F = \hat{s}$, and the q quark could be u, d, s, c or b). The direct muons from Z/γ^*-decays were required to have $p_T > 3$ GeV/c and $|\eta(\mu)| < 2.5$. The direct τ's from Z/γ^* decays were required to have $p_T > 3$ GeV/c and decayed normally. Both t- and s-channel diagrams were included. The s-channel diagram, not available in PYTHIA, gives a large peak at $M(4\mu) = M_Z$, contributing about 10% to events with $120 < M(4\mu) < 180$ GeV/c^2, and can be safely neglected for higher 4μ invariant masses; see [16] for details. The interference between t- and s-channels was found to be always negligible. The CompHEP events were further interfaced to PYTHIA 6.225 for showering and hadronization. The CompHEP LO cross sections for the two sub-processes were 113 and 157 fb, respectively.

To account for contributions to all the NLO diagrams and to the NNLO gluon fusion process ($gg \to ZZ$, known to contribute $\approx 20\%$ with respect to the LO [17]), we re-weighted events with a $M(4\mu)$-dependent K-factor $K(M_{4\mu}) = K_{NLO}(M_{4\mu}) + 0.2$. The NLO K-factor $K_{NLO}(M_{4\mu})$ was obtained with MCFM [15] and is shown in Figure 4. All details on calculation of this $M(4\mu)$-dependent K-factor and the dynamical differences between NLO and LO are summarized elsewhere [18].

The $M(4\mu)$ distributions after generator-level cuts are shown in Figures 1 and 2. The peak at M_Z is due to the s-channel contribution. This peak sits on the shoulder of the enhancement around $M(4\mu) > 100$ GeV/c^2—this peak corresponds to one of the two Z's going on-shell in the t-channel. The next bump around $2M_Z$ is due to both Z's going on-shell.

3 Trigger and offline muon reconstruction

Muons have a very clean detection signature resulting in a high trigger efficiency [19]. The inclusive muon triggers based on the selection of a single muon with $p_T > 19$ GeV/c or di-muons with $p_T > 7$ GeV/c assures an efficiency of practically 100% for collecting events that have four high-p_T muons.

In order to minimize muon reconstruction systematic uncertainties, we select only those reconstructed muons that have a transverse momentum $p_T > 7$ GeV/c, if they are in the central pseudorapidity region ($|\eta| < 1.1$), or with total momentum $p > 13$ GeV/c, if they are in the endcaps ($|\eta| > 1.1$). Figures 5 and 6 show the efficiency turn-on curves - the choice of these cuts is obvious from the figures. These cuts do not affect the number of accepted signal events dramatically.

Also, we require that all four possible combinations of reconstructed di-muon masses satisfy $m(\mu^+\mu^-) > 12$ GeV/c^2. As in the previous case, this cut has very little effect on the Higgs-boson events and is primarily intended to suppress poorly simulated hadronic background contributions originating from charmonium and bottomonium di-muon decays.

The most important characteristic distinguishing the Higgs-boson decays from all backgrounds is the presence of a peak in the four-muon invariant mass distribution. Figure 7 shows such a distribution for M_H=150 GeV/c^2. A Gaussian fit of the peak gives σ=1.1 GeV/c^2. One can see a noticeable tail toward smaller masses - mostly due to internal bremsstrahlung radiation and events with intermediate τ-leptons ($ZZ \to 2\tau 2\mu \to 4\mu 4\nu$ and $ZZ \to 4\tau \to 4\mu 8\nu$). The four-muon mass detector resolution $\delta M(4\mu)$ as a function of $M(4\mu)$ is given in Figure 8.

4 Higgs Search Strategy

The strategy for searching the Higgs boson being explored in this note is as follows.

- First, given a distinct localization of the Higgs-boson signal as a resonance-like peak in the invariant mass of four muons, the cuts can be made $M(4\mu)$-dependent. The cut optimization is described in section 4.2. The results with flat, $M(4\mu)$-independent, cuts are also presented for comparison.
- Second, after the cuts are applied, we search for the 4μ resonance-like peak over the continuum background. We require the peak to be consistent with the Standard Model Higgs. We explicitly compare the *potential sensitivity* of the two approaches: Log-Likelihood Ratio that takes into account the entire shape of the $M(4\mu)$-distributions for the signal and background as well as a straightforward counting experiment (at this

stage, no systematic errors are yet included; thus, we call this part of the analysis a potential sensitivity). In future, more sophisticated statistical tools can and will be added.

- Third, the most important theoretical and instrumental systematic errors are evaluated (section 4.3.1). We propose and analyze the merits of a number of methods for obtaining various analysis corrections directly from data in order to minimize our reliance on Monte Carlo simulation (both physics and detector performance). By doing this, we significantly reduce systematic errors and uncertainties.

- Fourth, systematic uncertainties are included in the evaluation of the significance of observing the Higgs-boson signal (section 4.3.2). For the counting experiment approach, this can be done analytically in a straightforward way. For LLR and other more sophisticated statistical tools, this can be done only by running a large number of pseudo-experiments and would also require a knowledge of all correlations across the data, the $M(4\mu)$-spectrum in this case,—these correlations are not yet available.

- And finally, should an excess of events consistent with the Higgs boson be observed, one should be careful in probabilistic interpretation of a local significance. A considerable over-estimation is possible due to the fact that the range of masses open for searching a relatively narrow signal is very large. In section 4.4, we outline a straightforward methodology of evaluating the scope of this effect. To be able to do it right, well defined search assumptions must be set a priori.

4.1 Introductory remarks on significance

As discussed above, after applying cuts, we will be searching for a possible local excess of events in the $M(4\mu)$ invariant mass spectrum. This can be done by using a likelihood ratio of the probability of observing the data in the case of the *signal-plus-background* hypothesis, to the probability of observing the data in the presence of the *background-only* hypothesis: $Q = P(observables|s+b)/P(observables|b)$. For the purpose of this study, the final observable we use for the likelihood ratio is the four-muon invariant mass. In principle, the list of observables can be extended further, but this requires a substantially larger sample of Monte Carlo data to be able to take into account all correlations properly. The likelihood ratio is known to give the best statistical discriminating power between two hypotheses [1].

The log-likelihood ratio (LLR) can be built for a narrow region in the vicinity of the peak (counting experiment):

$$-2\ln(Q) = 2\,s - 2\,n\ln\left(1+\frac{s}{b}\right), \tag{1}$$

for the entire binned spectrum:

$$-2\ln(Q) = 2\,S - 2\sum_{bins}^{N} n_i \ln\left(1+\frac{s_i}{b_i}\right), \tag{2}$$

or the unbinned spectrum:

$$-2\ln(Q) = 2\,S - 2\sum_{events}^{N} \ln\left(1+\frac{pdf_s(m_k)}{pdf_b(m_k)}\right). \tag{3}$$

The significance estimator $S = \sqrt{2\ln(Q)}$, S_{cL} for a counting experiment and S_L for the entire spectrum, is known to follow very closely (even for cases with few data) the one-sided Gaussian probability that one associates with the true significance S:

$$P = \int_{S}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})\,dx. \tag{4}$$

Figures 9 and 10 show that the S_{cL} tracks the true significance S for even small numbers of background events. Note that other popular quick estimators, $S_1 = s/\sqrt{b}$, $S_2 = s/\sqrt{n_o}$, and $S_{12} = 2(\sqrt{n_o} - \sqrt{b})$ also shown in Figures 9 and 10, do not work as well for smaller event counts and large values of significance.

The S_L estimator, being sensitive to the full shape of the signal and background distributions, has a leading edge over the simple counting experiment S_{cL} estimator. We typically observe 5-10% difference. The S_{cL} estimator, being local, is the natural tool for optimization of $M(4\mu)$-dependent cuts. The convolution of systematic errors into evaluation of a significance of an event excess is also very transparent and can be done analytically.

By definition, the significance values we obtain using our Monte Carlo samples are actually mean values for the expected significance to be measured in real data, should the Higgs boson exist at a given mass. However, the actual significance to be observed may be higher or smaller. The spread would be ± 1 in the limit of infinite data and is somewhat larger for the typical numbers of signal and background events in our case.

Since the estimators S_L and S_{cL} are only *estimators*, one needs to re-evaluate their probabilistic properties. For the LLR approach, one needs to run a large number of pseudo-experiments in order to evaluate the true confidence levels CL_b and CL_{b+s} (probabilities) of observing S_L smaller than a particular threshold (or, equivalently, -2ln(Q) larger than a particular threshold), for the background-only and background-plus-signal hypotheses, respectively. For the counting experiment approach this can be done analytically: e.g., assuming the background-only hypothesis, the probability of observing S_{cL} larger than S_{cL}, corresponding to the number of observed events n_o, is

$$1 - CL_b = P = p(n \geq n_o|b) = \sum_{k=n_o}^{+\infty} \frac{b^k}{k!} e^{-b}. \quad (5)$$

For non-integer n_o (as is the case for weighted Monte Carlo events), we can use a smooth interpolation of this function between integers n.

Before we include the systematic errors, it is this probability P that we use in conjunctions with Eq. 4 to define the true counting experiment significance S_c. Including systematic errors into evaluation of significance is discussed in Section 4.3.2.

4.2 Optimization of the $M(4\mu)$-dependent cuts

By taking advantage of the fact that the Higgs-boson resonance $H \rightarrow 4\mu$ is relatively narrow, we use $M(4\mu)$-dependent cuts for its search. The analysis steps in this case are as follows:

- First, events with 4 muons ($2\mu^+2\mu^-$) satisfying p_T, p, and $M(\mu^+\mu^-)$ quality cuts as described in section 3 are selected. This ensures that muons are reliably reconstructed and removes a "contamination" originating from heavy quarkonia decays.
- Second, after reconstructing a four-muon invariant mass, the $M(4\mu)$-dependent cuts are applied. The cuts, being smooth functions of $M(4\mu)$, are optimized in such a way that they maximize the S_{cL} significance of the Higgs signal excess *at all Higgs boson mass points*.
- And, finally, the resulting $M(4\mu)$ distribution is analysed for the presence of a Higgs-boson resonance. The search can be done using any statistical technique. In this note, we explicitly compare the potential of the LLR built from the $M(4\mu)$ distributions and a straightforward counting experiment approach.

To perform the desired $M(4\mu)$-dependent cut optimization, we used a recently developed program GARCON [1] [21]. The counting experiment significance estimator S_{cL} is the natural tool for such optimization.

For cut optimization, we considered the following muon kinematic variables:

- Tracker-based and calorimeter-based isolation energy for all four muons, ordered by isolation energy.
- Transverse momentum p_T of all four muons, ordered by p_T.

[1] GARCON stands for Genetic Algorithm for Rectangular Cuts OptimizatioN. This program allows one, in an automatic way, to optimize cut values and then verify the stability of the results, checking effectively a large number of cut sets, which, in a straightforward approach, would take an astronomical amount of time. In this analysis, we optimized 18 M_H points using 16 cut variables, with the step on each one equal to 0.025 (40 steps) of a typical variable value, which makes $18 \cdot (2 \cdot 16)^{40}$ cut set values to try in a straightforward way. We run the optimization and verifi cation on $\approx 3 \cdot 10^5$ events. More details on GARCON are available on its web-page and in a dedicated paper [21]. This program, being relatively new, is already widely used for other CMS analyses [22].

- Two invariant masses of di-muons $m_1(\mu^+\mu^-)$ and $m_2(\mu^+\mu^-)$, where the first pair of muons is the one that gives the invariant mass closest to m_Z.
- Parameter describing mismatch of muon displaced vertices.

The first half of the available Monte Carlo simulated data was used for the cut optimization (no systematic errors were included at this stage). The results for the 18 Higgs mass points were then fit to obtain smooth $M(4\mu)$-dependent cuts. It was found that, given the level of the expected dominant backgrounds ($t\bar{t}$, $Zb\bar{b}$, ZZ), there are only three critical discriminating cuts:

- Muon isolation cut, both tracker- and calorimeter-based, on the worst isolated muon, or equivalently one common cut on all four muons (Figs. 11 and 12). This cut strongly suppresses $t\bar{t}$ and $Zb\bar{b}$ backgrounds. One can see that the cuts gets tighter as $M(4\mu)$ gets smaller as the role of $Zb\bar{b}$ and $\mathrm{t\bar{t}}$ increases (Fig. 1).
- p_T cut on the second lowest p_T muon, or equivalently one common cut on the three highest p_T muons (Fig. 13). This cut helps to further suppress $Zb\bar{b}$ background to the level well below ZZ and reduces the ZZ background at high four-muon invariant masses. This cut gets more stringent with increasing $M(4\mu)$, as the transverse momentum of muons from a high mass Higgs boson tend to be higher than those from ZZ background.
- And, of course, $M(4\mu)$ window being used for scanning over the background (Fig. 14). It roughly corresponds to the $\pm 2\sigma$ width, where σ is the Higgs boson peak width that includes the detector resolution and the Standard Model Higgs-boson width.

In the next step, we applied the three critical cuts to the second half of the available Monte Carlo events that were not used for the optimization of the cuts. We tried both the cuts as they came from the optimization with GARCON and, alternatively, the smooth cuts shown by the lines in Figures 11- 14. The results of these comparisons are shown in Figure 19. The observed stability of the results as we switch from the first half of the sample used for cut optimization to the second half, as well as from "the optimal" to the smooth cuts, ensures that the cut optimization did not pick peculiar phase space corners corresponding to statistical flukes. Note that some "optimal" value points are absent in Figures 11-14: they are out of the plot-ranges and, in some cases, pushed by GARCON to the extreme limits. The latter means that this particular cut parameter for this particular Higgs mass point is not effective, as often is the case, for example, for two isolation-parameter-based variables, due to their very high degree of correlation.

Figure 20 shows the $M(4\mu)$ invariant mass distribution for the three background subprocesses and a Higgs-boson signal at $M_H = 150$ GeV/c^2 after applying the three smooth $M(4\mu)$-dependent cuts. One can see that the $t\bar{t}$ and $Zb\bar{b}$ backgrounds are now suppressed well below the irreducible ZZ background.

Other possible cuts such as invariant masses of the muon pairs, impact point parameters, kinematical cuts on other muons, and isolation parameters on other muons do not significantly help to improve the results further, see Figures 15-18. The cuts on these observables may still be useful and play a role of "safeguards" to suppress possible unaccounted-for backgrounds related to the beam halo, detector mis-performance, etc.

Additional variables that may help to discriminate H from the dominant ZZ background have been studied: $p_T(4\mu)$, number of jets and their E_T, etc. However, these variables are driven by the NLO production processes, while our samples were generated at the Leading Order by Pythia and CompHEP. Therefore, any conclusions that we might derive from these samples would not be reliable. Some angular distributions built from muons also have some differences originating from the underlying spin structures, but they are not sufficiently discriminating to be used for cuts and may be strongly affected by the NLO diagrams.

Figures 21 and 22 show the significances S_{cL} and S_L at $L = 30\ fb^{-1}$ for the expected excesses of events for different Higgs-boson masses. To emphasize the gain in the sensitivity achievable with $M(4\mu)$-dependent cuts, the results for flat cuts, optimized for $M_H = 150$ GeV/c^2, are also superimposed. As expected, one universal set of cuts cannot deliver the optimal performance for the full range of possible Higgs masses. The gain in significance can be easily translated into probabilistic terms. For example, the Higgs boson with $M_H = 500$ GeV/c^2 is right at the 5σ-discovery threshold for an integrated luminosity $L = 30\ fb^{-1}$ (Fig. 21). The difference in the average expected significances, 5.3 and 4.6, means in this case that the chances of discovering the Higgs boson with $M_H = 500$ GeV/c^2 at $L = 30\ fb^{-1}$ are $< 40\%$ for the flat cuts and $> 60\%$ for the $M(4\mu)$-dependent cuts.

Figure 23 shows the same results as in Figure 21 in terms of the luminosity needed for observing an excess of events over the expected background in the presence of the Standard Model Higgs boson at 5σ significance.

Figure 24 gives an idea of how the experimental exclusion limits will map onto the plane of cross section vs. Higgs-boson mass for a few different integrated luminosities.

4.3 Systematic Errors

The analysis of the systematic errors can be divided into two distinct stages. First, one needs to understand the uncertainties in predicting the background. Second, these uncertainties in the background have to be included in the evaluation of the significance of an excess of events, should it be observed.

4.3.1 Uncertainties in the background

Uncertainties in the signal are not very important for establishing an excess of events over the background. It is the uncertainties in the background that are of main concern. After applying the analysis cuts as described earlier, ZZ production is the dominant irreducible background, with all other processes giving much smaller contributions. This reduces the analysis of systematic errors to the $ZZ \to 4\mu$ process. The main uncertainties are as follows:

- PDF and QCD scale uncertainties;
- NLO and NNLO contributions vs LO;
- Integrated luminosity;
- Trigger efficiency;
- Muon reconstruction efficiency;
- Muon isolation cut efficiency;
- Four-muon mass $M(4\mu)$ resolution;
- Four-muon mass $M(4\mu)$ absolute scale.

One can try to evaluate/guess the theoretical and detector performance related uncertainties starting from the first principles. However, the credibility of the detector performance systematic errors estimated this way is always shaky, especially during the earlier stages of the detector operation when the changes in the system are frequent and hard to monitor; and they must be timely incorporated into the detector Monte Carlo simulation.

Therefore, we developed methods to evaluate various corrections, such as muon reconstruction efficiency, muon isolation cut efficiency, $M(4\mu)$ resolution, and absolute scale, directly from data in order to minimize our relying on the Monte Carlo simulation, and, thus, significantly reducing the associated systematic errors.

Moreover, throughout this analysis, we estimate the background around a particular $M(4\mu)$ area (signal region) in reference to a *measured* control sample. Note that this completely eliminates uncertainties associated with measuring the luminosity and reduces the sensitivity to PDF and QCD-scales. For the control sample, we use either the inclusive $Z \to 2\mu$ process or sidebands of the $M(4\mu)$ spectrum itself. When we refer to a control sample, we will use the factor ρ, defined so that $b = \rho \cdot B$, where b is the expected number of background events in the signal window and B is the measured number of events in the control sample.

The PDF and QCD scale uncertainties in the $ZZ \to 4\mu$ production cross section were studied at the NLO level using MCFM [15]. Systematic errors associated with PDFs were estimated by giving $\pm 1\sigma$ variations to the 20 CTEQ6M parameters. By varying independently the renormalization and factorization scales by a factor of two up and down from their default values $\mu_R = \mu_F = 2M_Z$, we found the sensitivity of the ZZ cross section to the QCD scale uncertainties. All details of these studies can be found in Ref. [23]. Figure 25 shows these PDF and QCD scale uncertainties, added in quadrature, versus $M(4\mu)$. The three curves correspond to (a) the absolute predictions (relatively flat, $\delta b/b \approx 6\%$); (b) the prediction normalized to the measured $Z \to 2\mu$ cross section (note that the $Z \to 2\mu$ cross section can be measured with instrumental systematic errors, not including luminosity, of less than 2% (CDF results, Phys. Rev. Letter, 94 (2005) 091803)) ($\delta\rho/\rho \approx 1\%$ for $M(4\mu)$ close to M_Z and then steadily increasing toward larger four-muon invariant masses); and (c) prediction normalized to sidebands of the

$M(4\mu)$ distribution itself in the range from 100-600 GeV/c^2 (($\delta\rho/\rho$ is at its minimum when the signal window is at the place where most of the events are).

Beyond-Leading-Order correction uncertainties were estimated as follows. The $M(4\mu)$-dependent K-factor $K(M_{4\mu})$ for the $ZZ \to 4\mu$ process was evaluated with two very different programs: MCFM [15] and EffNLO [18]. The latter is a package smoothly splicing together MadGraph [20] (NLO $pp \to 4\mu + jet$) and Pythia (LO $pp \to 4\mu$ + ISR-jets). The relative difference in $K(M_{4\mu})$ is shown in Figure 26. The NNLO diagrams include *new processes* (we define a process as new if it has a distinctly different initial state and, therefore, variations of QCD scales do not necessarily give a feel for its relative importance): $gg \to ZZ \to 4\mu$ (box-diagram), contributing about $(20 \pm 8)\%$ to the LO cross section [17] (note that this contribution was calculated without virtual photons in the propagators) and $qq \to 4\mu + qq$ via Z-bremsstrahlung (not yet calculated) or via vector-boson fusion (implemented in Pythia, very small). Since the nature of all these differences/variations is not well understood, we present the final results with and without these uncertainties included. Certainly, more theoretical work in this area is needed. All other higher-level diagrams can be considered as corrections to the distinct LO, NLO, and NNLO processes discussed above. Omission of these higher-order corrections would manifest itself as a sensitivity of the calculated cross sections to the QCD scale variations discussed earlier.

Luminosity measurement uncertainties are expected to be 10% at the time of an integrated lumunosity of $1\ fb^{-1}$, 5% at $L = 10\ fb^{-1}$, and 3% for larger luminosities. When we estimate the ZZ background events in the signal region via the measured number of events in the control samples, the luminosity uncertainties largely cancel out.

Muon trigger efficiency, being very close to 100% due to the presence of four muons, does not have substantial systematic errors.

The muon reconstruction efficiency can be measured directly from data with an uncertainty of better that 1%. The method uses a data sample based on single-muon HLT (HLT stands for High Level Trigger, the final stage of online filtering after which the data are recorded on tape) trigger with $p_T > 19$ GeV/c. This sample will contain inclusive W, Z, and other processes in the approximate ratio W:Z:others = 10:1:small [24]. By counting the number of $Z \to 2\mu$ events in the resonance peak of the invariant mass distributions built from the HLT muon and all other tracks, the HLT muon and all other standalone muons and the HLT muon and all other globally reconstructed muons, one can evaluate the efficiency of finding globally-reconstructed muons with better than 1% precision. Such a measurement will automatically account for the real detector performance, including intermittent and smooth variations in time. All details can be found elsewhere [25]. The four-muon efficiency therefore will be known with an absolute error of better than 4%. When we deduce the expected $ZZ \to 4\mu$ events from the measured $Z \to 2\mu$ cross section, this uncertainty partially cancels out and becomes 2%. This efficiency remains fairly flat vs $M(4\mu)$, which makes this error completely negligible if sidebands are used for evaluating the number of expected background events in the signal region.

The muon isolation cut is very important as it allows us to suppress otherwise overwhelming $t\bar{t}$ and $Zb\bar{b}$ backgrounds well below the ZZ background. As we apply this cut, we also cut ZZ (and Higgs) events by $\approx 15 - 30\%$. This cut is very sensitive to the underlying event physics, which, unfortunately, is not very well understood and has substantial uncertainties. As in the case of the muon efficiency, we developed a scheme for evaluating the muon isolation cut efficiency directly from data. Again, we appeal to the inclusive $Z \to 2\mu$ sample. The Z events have very similar underlying event activity as ZZ events. We show that, by using random directions in Z events and evaluating the energy flow in isolation cones around them, one can predict the 4-muon ZZ event losses due to the muon isolation cut with a systematic error of less than 2% [26] (Fig. 27).

The uncertainty on the muon p_T resolution directly propagates into the four-muon invariant mass $M(4\mu)$ reconstruction. This almost does not affect the background distribution. However, the $M(4\mu)$ distribution width drives the width of the $M(4\mu)$ window that we use for evaluating the signal excess significance at low Higgs-boson masses. Fortunately, even making a mistake in the $M(4\mu)$ distribution width by as much as 25% has only a tiny effect on evaluating a significance of an excess of events (Fig. 28). Also, the muon p_T resolution is fairly easy to measure from data using the measured J/ψ and Z peak widths with a precision much better than needed.

The uncertainty on the muon p_T scale can be similarly calibrated from data using the measured J/ψ and Z peaks. The effect of these uncertainties on the number of background events in a signal window appears only on steep slopes of the $M(4\mu)$ distribution. For the steepest part of the $M(4\mu)$ distribution in the $180 - 200$ GeV/c^2 range, we obtain $\delta b/b \approx 0.1\delta M_{4\mu}$, where $\delta M_{4\mu}$ is in GeV/c^2. This implies that to be able to neglect this effect, one needs to know the momentum scale with a precision of 0.1 GeV/c at $p_T \approx 50$ GeV/c. This can be easily achieved with just a few hundred $Z \to 2\mu$ events.

Figures 29 and 30 summarize all the systematic errors on the expected number of events in the $ZZ \to 4\mu$ background for the two methods: via referencing to the total measured $Z \to 2\mu$ cross section and via referencing to the event count in the sidebands of the $M(4\mu)$ spectrum itself.

4.3.2 Significance with the background uncertainties included

If the background has uncertainties, which we will express in terms of a probability density function $f(b)$, the probability of observing at least n_o events becomes

$$P = p(n \geq n_o|b) * f(b) = \int_0^{+\infty} p(n \geq n_o|b)\, f(b)\, db, \tag{6}$$

which can be again converted into true significance S_c using Eq. 4.

We will use a log-normal form of a probability density function for the absolute systematic errors for expected number of background events b_0 with a relative uncertainty $\delta = \Delta b/b_0$:

$$f(b) = \frac{1}{\sqrt{2\pi}\ln(k)} \exp\left(-\frac{\ln^2(b/b_0)}{2\ln^2(k)}\right) \frac{1}{b}. \tag{7}$$

In this equation, $k = 1 + \delta$ and δ is the sum in quadrature of all the uncertainties. For relatively small errors, this form of equation gives a Gaussian distribution with average b_0 and $\sigma = \delta \cdot b_0$. One advantage of using the log-normal presentation is that it does not have a tail spilling over into $b < 0$. Also, and maybe more importantly, this equation give an intuitively correct representation for very large uncertainties. For example, such a statement as "we estimate that the background is b_0 with a factor of 2 uncertainty" probably implies that we assume that the chances for the true background to be somewhere between $b_0/2$ and $2b_0$ are about 68%, while the chances of being larger than $2b_0$ or smaller than $b_0/2$ are approximately equal—Eq. 7 does just that for any value of k, small or large (k would be equal 2 in this case).

The statistical part of the probability density function $f(b)$ for the background in the signal region, estimated from the observed event count in a control sample B ($b = \rho B$), can be obtained using Bayes' theorem and a flat prior:

$$f(b) = \frac{1}{\rho} \frac{(b/\rho)^B e^{-b/\rho}}{\Gamma(B+1)}. \tag{8}$$

The full probability density function $f(b)$ for the background, estimated using sidebands when there are uncertainties on the factor ρ, can be easily obtained by a convolution of the two equations shown above.

Figure 31 gives three curves: the significance vs Higgs mass in the absence of any systematic errors (both for the plain S_{cL} estimator and the true significance S_c) as well as the significance that includes all uncertainties in the background when it is estimated from the measured $Z \to 2\mu$ cross section. All three curves correspond to the total integrated luminosity of 30 fb^{-1}. Figure 32 shows curves (with and without systematic errors) for the required luminosity for 5σ-discovery, 3σ-evidence, and 95% CL exclusion limit for the Standard Model Higgs boson.

The comparison between two ways of normalization, to the $Z \to 2\mu$ process and the $ZZ \to 4\mu$ sidebands, is made in terms of the luminosity required for 5σ-discovery (Fig. 33). The difference is not dramatic. The true benefit of using two approaches to estimating background from data is in their complementarity.

Finally, Table 2 summarizes the most important results for the $M(4\mu)$-dependent cuts that we presented in this note.

4.4 Local significance and overall statistical fluctuation probability

In searching for a new phenomena in a wide range of parameter phase space (in our case, we search for a narrow resonance in a very broad range of invariant masses), one inevitably encounters a well-known problem of overestimating the overall significance of a "local discovery." The scale of the effect can be quite large (see Appendix, Sec. 7 for all details) and one must exercise a caution in evaluating probabilistic interpretation of observing an excess of events at a particular mass. In the case study we consider in the Appendix, we show that observing a 2σ excess would be basically guaranteed, an observation of a "local 3σ" excess would be hardly of any significance

Table 2: Summary of the results: number of signal and background events in a window used for a counting experiment with the $M(4\mu)$-dependent cuts; systematic error on the background normalized to the $Z \to 2\mu$ process ($\delta K_{NLO}/K_{NLO}$ is not included); three different significances without systematic errors included: the S_L estimator for the Log Likelihood Ratio (LLR) built for the full $M(4\mu)$ spectrum, S_{cL} LLR estimator built for a counting experiment approach, and the S_c true significance for the counting experiment approach; the final result for S_c, now including all systematic errors.

Mass (GeV/c^2)	Signal s	Bkgd b	Syst Error $\delta b/b$	S_L (no syst)	S_{cL} (no syst)	S_c (no syst)	S_c (with syst)
115	2.13	0.92	3%	1.98	1.75	1.54	1.54
120	4.00	1.15	4%	2.88	2.72	2.57	2.56
130	12.45	2.06	3%	5.97	5.64	5.54	5.52
140	23.22	2.65	3%	9.09	8.45	8.39	8.35
150	28.09	2.42	4%	10.84	9.92	9.87	9.81
160	14.25	3.01	4%	6.04	5.64	5.55	5.53
170	6.32	3.63	4%	3.00	2.73	2.61	2.60
180	14.54	7.10	4%	4.83	4.38	4.30	4.26
190	54.95	17.00	4%	10.85	9.89	9.85	9.59
200	62.78	19.93	4%	11.32	10.48	10.45	10.11
250	54.48	21.62	5%	9.83	9.09	9.05	8.61
300	40.43	13.40	6%	9.26	8.30	8.25	7.90
350	40.68	10.75	7%	9.50	8.93	8.88	8.47
400	33.59	8.00	8%	9.07	8.36	8.31	7.95
450	24.07	5.68	8%	7.57	7.10	7.03	6.81
500	16.65	5.58	9%	5.81	5.31	5.23	5.05
550	12.10	5.70	9%	4.44	4.04	3.92	3.80
600	8.72	5.04	9%	3.58	3.20	3.10	3.00

($\approx 15\%$ chance), and the "significance of a local 5σ-discovery" would actually correspond to a true statistical significance of $\approx 3.8\sigma$. In the same Appendix, we also discuss possible ways to reduce the scale of the effect.

5 Summary

Discovery of the Standard Model Higgs boson in the "gold-plated" decay mode $H \to ZZ^{(*)} \to 4\mu$ was analyzed in the context of the CMS Detector. The explored range of Higgs-boson masses was 115-600 GeV/c^2. The Monte Carlo samples for signal and background were generated to represent the NLO cross sections, including $M(4\mu)$-dependent K-factors. To simulate the detector response and reconstruct physics objects, the full CMS Detector simulation and reconstruction software was used. We explored the Higgs boson discovery potential for different analysis variations, including the use of $M(4\mu)$-dependent and flat cuts, Log Likelihood Ratio based on the full $M(4\mu)$ spectrum, and a straightforward counting experiment approach.

A full treatment of the most important theoretical and instrumental systematic errors and their effect on evaluation of significance of the Higgs-boson observation using mass-dependent cuts and a counting experiment approach were presented. To minimize systematic errors, a number of methods of reconstructing the necessary corrections directly from data were developed.

We showed that at $\approx 2\ fb^{-1}$ of integrated luminosity, we would be able to start excluding the SM Higgs boson at 95% CL for M_H in vicinity of 200 GeV/c^2. By the time we reach $\approx 30\ fb^{-1}$, we would exclude the Standard Model Higgs in its four-muon decay mode in the mass range $M_H = 120 - 600$ GeV/c^2, if indeed it does not exist.

The discoveries at the level of "5σ" local significance could be already possible at $\approx 10\ fb^{-1}$ for M_H in the range 140-150 and 190-400 GeV/c^2. By the time we reach $\approx 30\ fb^{-1}$, the discovery range would open up to 130-160 and 180-500 GeV/c^2. An observation of the Higgs boson with the mass $M_H \approx 170$ GeV/c^2 or ≈ 600 GeV/c^2 in the $H \to ZZ^{(*)} \to 4\mu$ decay channel would require an integrated luminosity of the order of $100\ fb^{-1}$.

6 Acknowledgments

We would like to thank M. Aldaya, P. Arce, R. Breedon, J. Caballero, R. Cousins, B. Cruz, T. Ferguson, U. Gasparini, P. Garcia, J. Hernandez, I. Josa, P. Moisenz, E.R. Morales, N. Neumeister, A. Nikitenko, and G. Quast for

their active participation in the analysis discussions.

References

[1] J. Neyman and E. Pearson, *On The Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference*, Part I, Biometrika, Vol. 20A, No.1/2 (July 1928), pp.175-240;
J. Neyman, and E. Pearson, *On the Problem of the most Efficient Tests of Statistical Hypotheses*, Philosophical Transactions of the Royal Society of London, Vol. 231 (1933), pp. 289-337;
S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Annals of Mathematical Statistics, 9:60-62, 1938.

[2] OSCAR, *Object-oriented Simulation for CMS Analysis and Reconstruction*, http://cmsdoc.cern.ch/OSCAR

[3] ORCA, *Object-oriented Reconstruction for CMS Analysis*, http://cmsdoc.cern.ch/ORCA

[4] S. Abdullin et al., *"Summary of the CMS Potential for the Higgs Boson Discovery"*, CMS Note 2003/033;
M. Sani, *Search for the Standard Model Higgs Boson in four-Muon Final State with CMS*, CMS CR-2004/035, proceeding of *Physics at LHC*, Vienna, Austria, July 2004;
V. Bartsch, *Simulation of Silicon Sensors and Study of the Higgs Decay* $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\mu$ *for CMS (LHC)*, Ph.D. thesis, IEKP-KA/2003-26 University of Karlsruhe (2003).

[5] M. Aldaya et al., *"Search for the Standard Model Higgs boson in the* $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\mu$ *decay channel using a mass-independent analysis"*, CMS Note 2006/106.

[6] S. Baffioni et al., *Search for the Standard Model Higgs Boson in the Four-Electron Final State with CMS*, CMS Note 2006/115.

[7] D. Futyan and D. Giordano, *Search for the Standard Model Higgs Boson in the Two-Electron and Two-Muon Final State with CMS*, The CMS Physics Technical Design Report, Volume 2, in preparation.

[8] T. Sjostrand, L. Lonnblad, and S. Mrenna, *PYTHIA 6.2 Physics and Manual, report LU-TP-01-21*, Aug 2001, arXiv:hep-ph/0108264.

[9] CMKIN, *CMS Interface for Event Generators, http://cmsdoc.cern.ch/cmsoo/projects/CMKIN*

[10] E. Barberio and Z. Was, Computer Phys. Commun. 79 (1994) 291;
E. Barberio, J. van Eijk, and Z. Was, Computer Phys. Commun. 66 (1991) 115.

[11] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension*, arXiv:hep-ph/9704448;
M. Spira, *HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections*, arXiv:hep-ph/9510347.

[12] Particle Data Group, *Review of Particle Properties*, Phys. Lett. B 562 (2004) 1.

[13] F. Maltoni, *Theoretical Issues and Aims at the Tevatron and LHC*, proceeding of *Hadron Collider Physics Symposium*, Les Diableretes, Switzerland, July 2005, printed version: Springer Proceedings in Physics, Vol. 108, 2006, ISBN: 3-540-32840-8.

[14] CompHEP collaboration: *"CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33"*, hep-ph/9908288, http://theory.sinp.msu.ru/parser/parser.php?file=/phpcms/comphep

[15] J. M. Campbell, arXiv:hep-ph/0105226, *W/Z + B anti-B / jets at NLO using the Monte Carlo MCFM*.

[16] S. Abdullin et al., *"Relative Contributions of t- and s-Channels to the* $\mathrm{ZZ} \to 4\mu$ *Process"*, CMS Note 2006/057.

[17] C. Zecher, T. Matsuura, and J.J. van der Bij, *"Leptonic Signals from off-shell Z Boson Pairs at Hadron Colliders"*, hep-ph/9404295.

[18] P. Bartalini et al., *"NLO vs. LO: kinematical differences for signal and background in the* $\mathrm{H} \to \mathrm{ZZ}^{(*)} \to 4\mu$ *analysis"*, CMS Note in preparation.

[19] CMS Collaboration, *"The CMS Physics Technical Design Report, Volume 1"*, CERN/LHCC, 2006.

[20] F. Maltoni and T. Stelzer, *MadEvent: Automatic Event Generation with MadGraph*, arXiv:hep-ph/0208156, JHEP 0302 (2003) 027.

[21] S. Abdullin et al., *"GARCON: Genetic Algorithm for Rectangular Cuts OptimizatioN. User's manual for version 2.0"*, hep-ph/0605143, http://drozdets.home.cern.ch/drozdets/home/genetic/

[22] D. Acosta et al., *"CMS Discovery Potential for mSUGRA in Same Sign Di-muon Events with Jets and Large Missing Transverse Energy in pp collisions at* $\sqrt{s} = 14$ *TeV"*, CMS Note in preparation;
D. Acosta et al., *"CMS Discovery Potential for mSUGRA in Single Muon Events with Jets and Large Missing Transverse Energy in pp Collisions at* $\sqrt{(s)} = 14$ *TeV"*, CMS Note in preparation;
V. Abramov et al., *"Selection of single top events with the CMS detector at LHC"*, CMS Note 2006/084;
W.de Boer et al., *"Trilepton final state from neutralino-chargino production in mSUGRA"*, CMS Note in preparation.

[23] S. Abdullin et al., *Study of PDF and QCD scale uncertainties in* $\mathrm{pp} \to \mathrm{ZZ} \to 4\mu$ *events at the LHC*, CMS Note 2006/068.

[24] CMS Collaboration, *The Trigger and Data Acquisition project*, v.2, p.308, 2002.

[25] D. Acosta et al., *Measuring Muon Reconstruction Efficiency from Data*, CMS Note 2006/060.

[26] S. Abdullin et al., *Sensitivity of the Muon Isolation Cut Efficiency to the Underlying Event Uncertainties*, CMS Note 2006/033.

[27] E. L. Lehmann, *"A Theory of some Multiple Decision Problems"*, The Annals of Mathematical Statistics, Vol. 28 1, 1, (1957).

[28] R. O'Neill and G. B. Wetherill, *"The Present State of Multiple Comparison Methods"*, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 33 2, 218, (1971).

[29] J. Hauser, *"Search for New Physics in Tails of Distributions"*, private communications, paper in preparation.

[30] B. Knuteson, *"A Quasi-Model-Independent Search for New High* p_T *Physics at D0"*, thesis, 2000.

[31] B. Abbott et al., *"Search for New Physics in e mu X Data at D0 Using Sleuth: A Quasi-Model-Independent Search Strategy for New Physics"*, D0 Collaboration, Phys.Rev. D62 (2000) 092004, hep-ex/0006011,
B. Abbott et al., *"A Quasi-Model-Independent Search for New Physics at Large Transverse Momentum"*, D0 Collaboration, Phys.Rev. D64 (2001) 012004, hep-ex/0011067.

[32] C. Finley et al., *"On the evidence for clustering in the arrival directions of AGASA's ultrahigh energy cosmic rays"*, Astroparticle Physics 21 (2004) 359-567.

[33] S.I. Bityukov and N.V. Krasnikov, *"On the observability of a signal above background"*, NIM A 452 (2000) 518-524

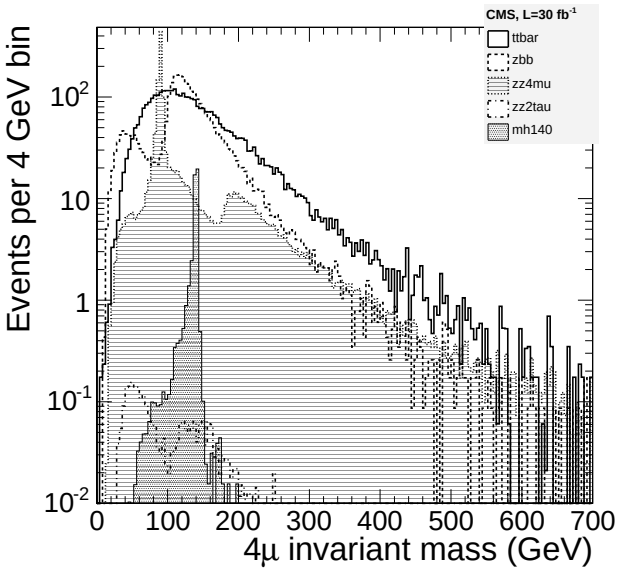


Figure 1: $M(4\mu)$ distributions after generator-level cuts for $\mathrm{t\bar{t}}$, $(Z/\gamma^*)\mathrm{b\bar{b}}$, $\mathrm{Z}/\gamma^*\mathrm{Z}/\gamma^*$, and $m_{\mathrm{H}} = 140\ \mathrm{GeV}/c^2$ (log scale).

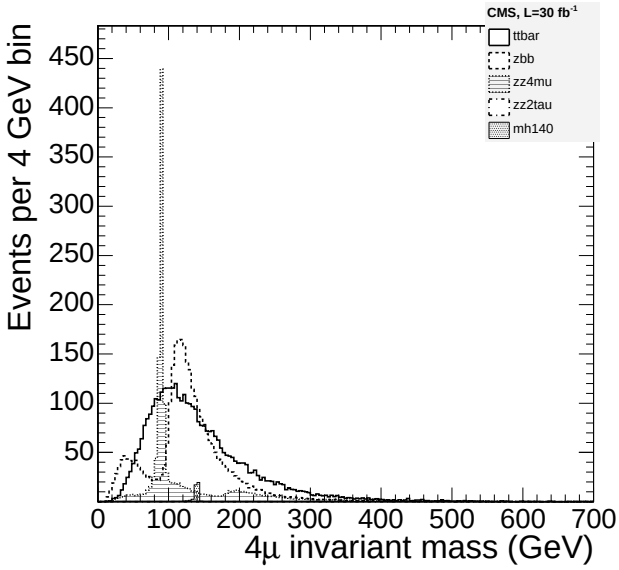


Figure 2: Same as Figure 1, but on a linear scale.

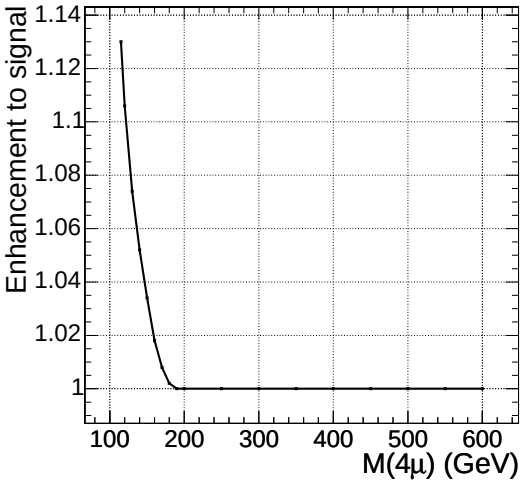


Figure 3: Enhancement to the signal samples' cross sections due to interference effects not accounted for at the generator level.

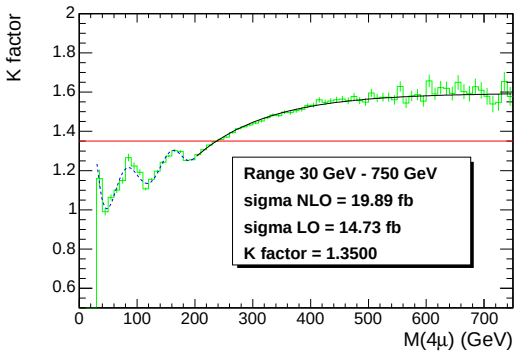


Figure 4: The $M(4\mu)$-dependent NLO K-factor $K_{NLO}(M_{4\mu})$ for the $ZZ \to 4\mu$ process evaluated with MCFM [15].

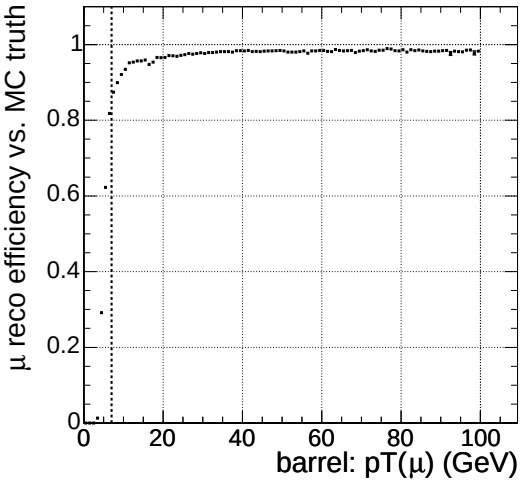


Figure 5: Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the barrel region vs. p_T.

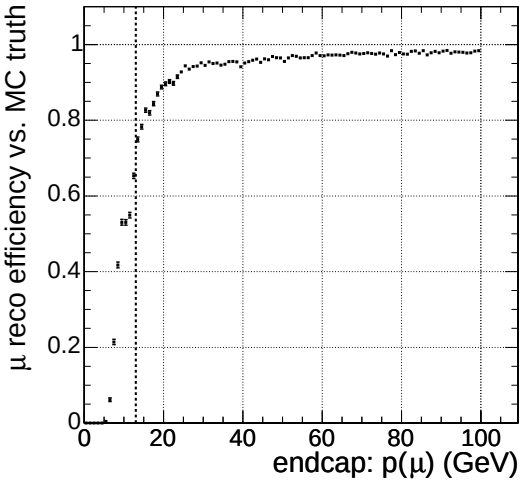


Figure 6: Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the endcap region vs. momentum.

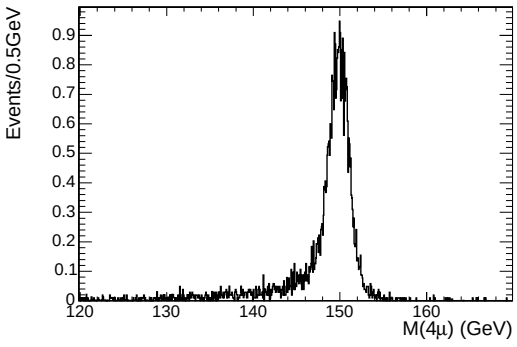


Figure 7: $M(4\mu)$ distribution for m(H)=150 GeV$/c^2$ and the fit described in the text.

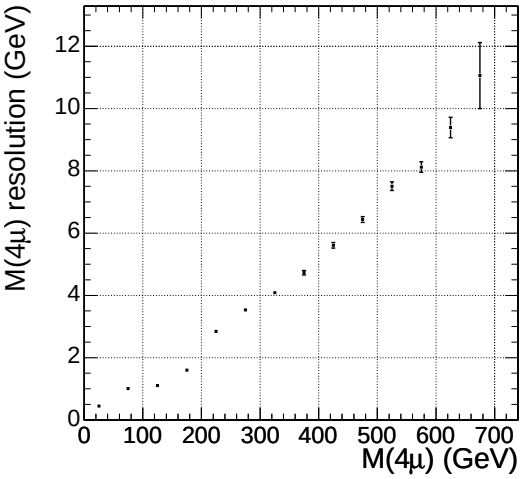


Figure 8: $M(4\mu)$ resolution vs. M_H.

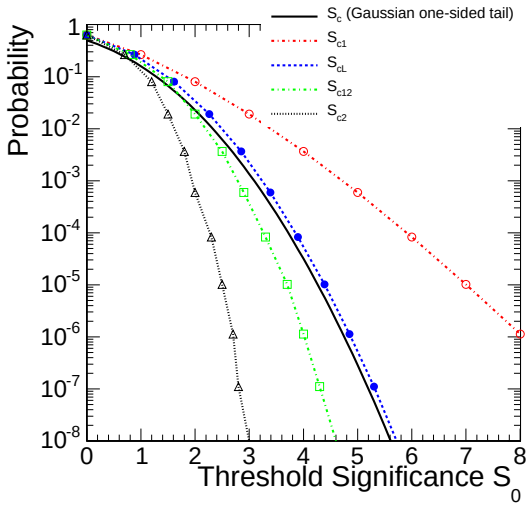


Figure 9: Comparison of different significance estimators for 1 background event: probability of measuring significance $S > S_0$, background only case, $N_B = 1$ event. Equations for different significance estimators are given in the text.

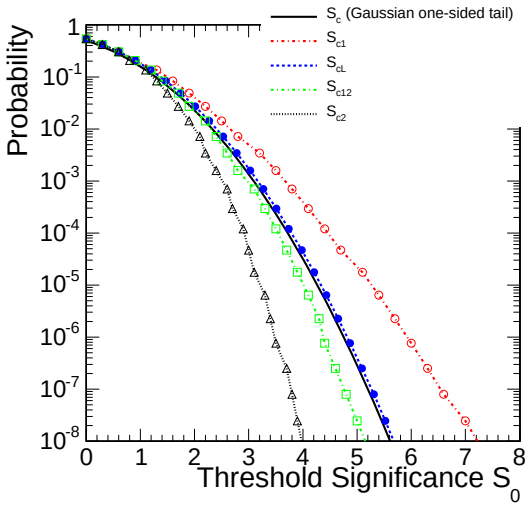


Figure 10: Same as Figure 9, but for $N_B = 10$ events.

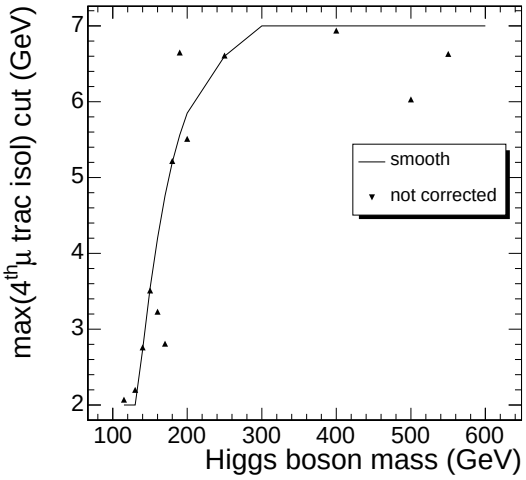


Figure 11: Dependence of the tracker-based muon isolation cut on the least isolated muon versus Higgs mass. Smooth dependance (curve) which follows general dependance of the cut variable was used for analysis-level cuts.

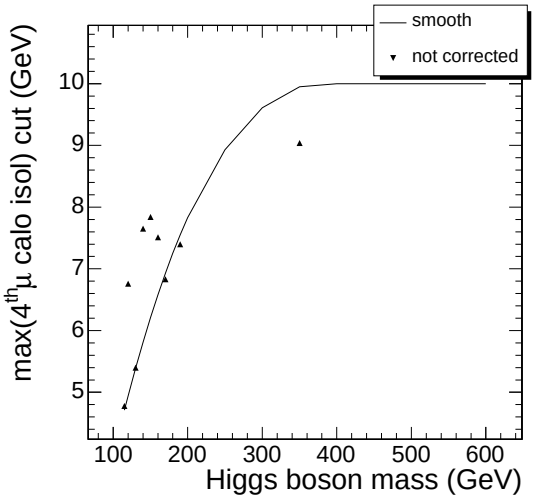


Figure 12: Dependence of the calorimeter-based cut on the least isolated muon versus Higgs mass. Smooth curve has the same meaning as for Figure 11.

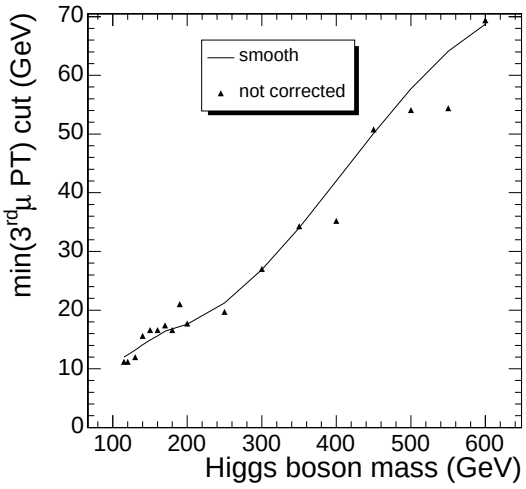


Figure 13: Dependence of the p_T cut on the second-lowest-p_T muon versus Higgs mass. Smooth curve has the same meaning as for Figure 11.

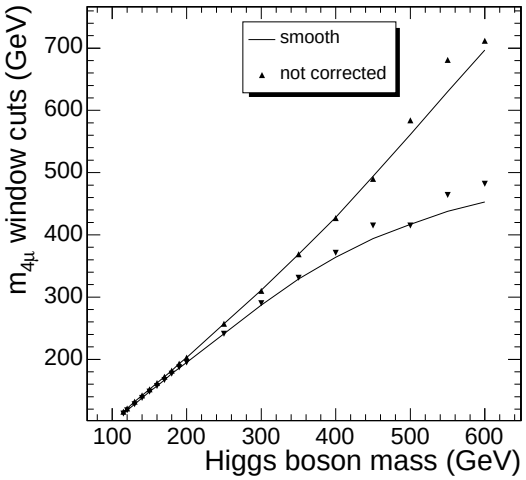


Figure 14: Dependence of the $M(4\mu)$ window cuts versus Higgs mass. Smooth curve has the same meaning as for Figure 11.

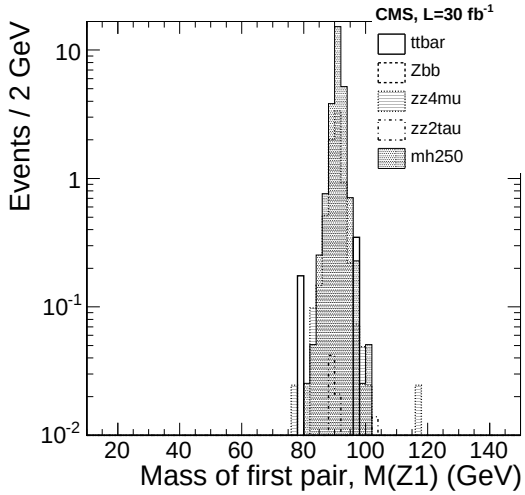


Figure 15: First muon pair invariant mass distribution, M(Z1) (invariant mass of two opposite sign muons closest to Z^0-mass), after analysis cuts were applied.

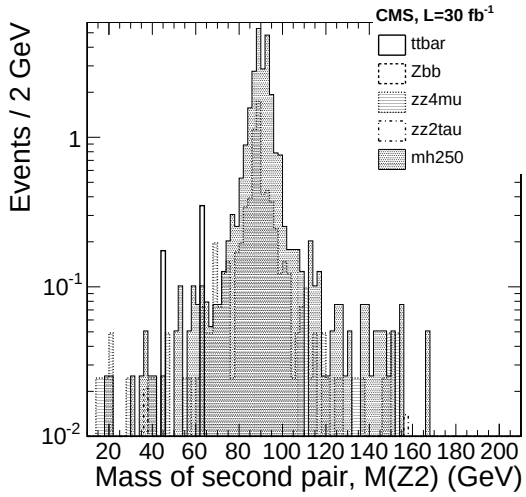


Figure 16: Second muon pair invariant mass distribution, M(Z2) (invariant mass of two other highest p_T opposite sign muons after muons for Z1 selected), after analysis cuts were applied.

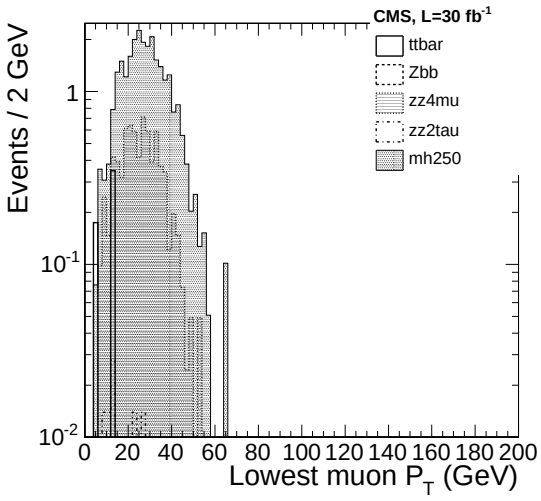


Figure 17: Lowest muon p_T distribution, after analysis cuts were applied.

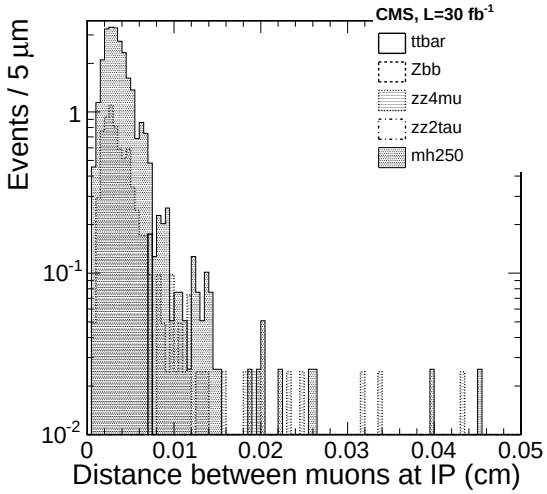


Figure 18: Maximum distance in XY-plane between muon impact point coordinates distribution, after analysis cuts were applied.

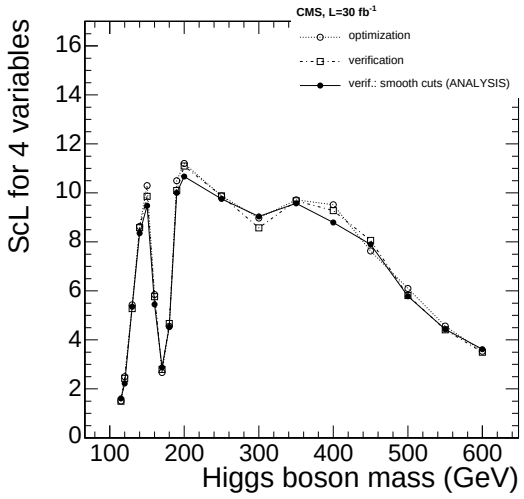


Figure 19: S_{cL} vs. Higgs-boson mass: optimized cut values and GARCON optimization step when first half of the simulated data are used (empty circles markers and dotted curve); optimized cut values and GARCON verification step when second half of the simulated data are used (empty squares markers and dashed curve); smoothed cut values used with full analysis program when second half of the simulated data are used (filled circles markers solid curve).

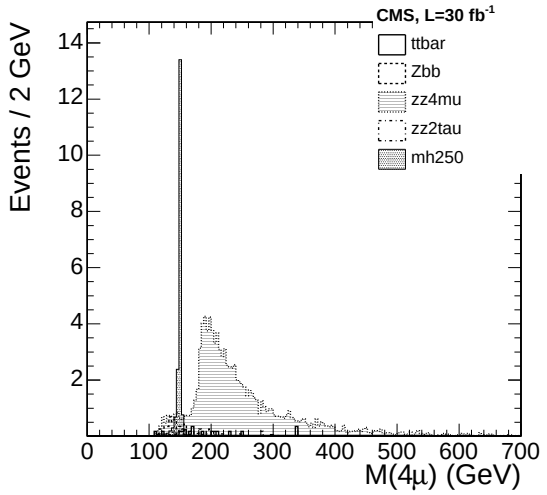


Figure 20: $M(4\mu)$ invariant mass distribution for the three background subprocesses and a Higgs-boson signal at $M_H = 150$ GeV$/c^2$, after applying cuts on muon isolation and p_T.

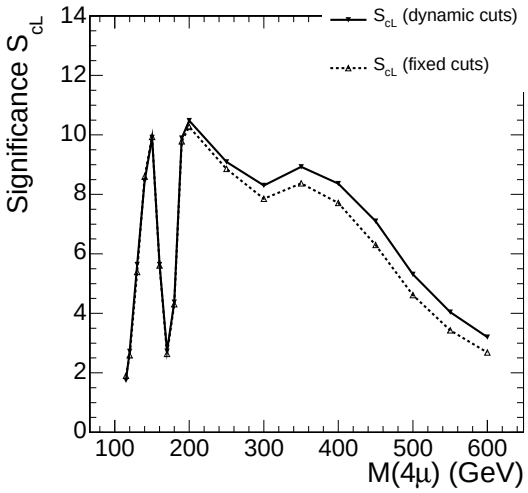


Figure 21: Expected excess significance S_{cL} with $L = 30\ fb^{-1}$ for different Higgs-boson masses for $M(4\mu)$-dependent (solid line) and independent cuts (dashed line). No systematic errors included.

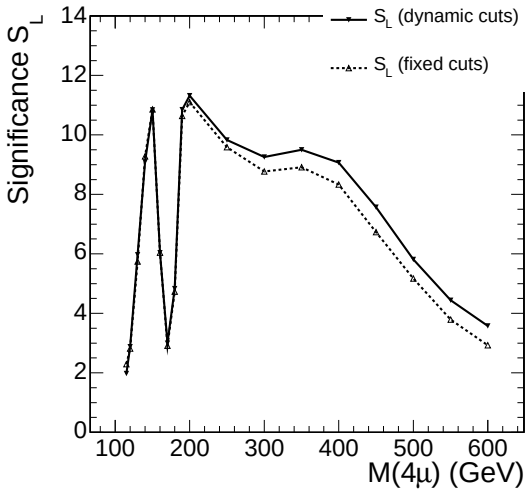


Figure 22: Expected excess significance S_L with $L = 30\ fb^{-1}$ for different Higgs-boson masses for $M(4\mu)$-dependent (solid line) and independent cuts (dashed line). No systematic errors included.

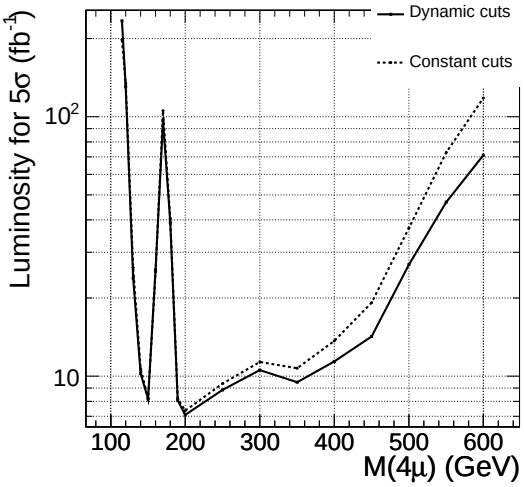


Figure 23: Luminosity required to reach a 5σ event excess for different Higgs-boson masses for $M(4\mu)$-dependent (solid line) and independent cuts (dashed line). No systematic errors included.

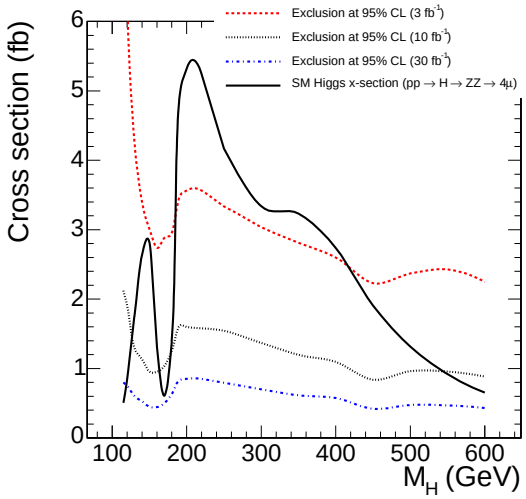


Figure 24: The 95% CL exclusion contours for the SM Higgs hypothesis in the (cross section; mass) plane at integrated luminosities of 3 (upper dashed line), 10 (middle dashed line), and 30 (lower dotted-dashed line) fb^{-1}. The Standard Model (SM) Higgs-boson cross section for $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ is also shown (solid black line).

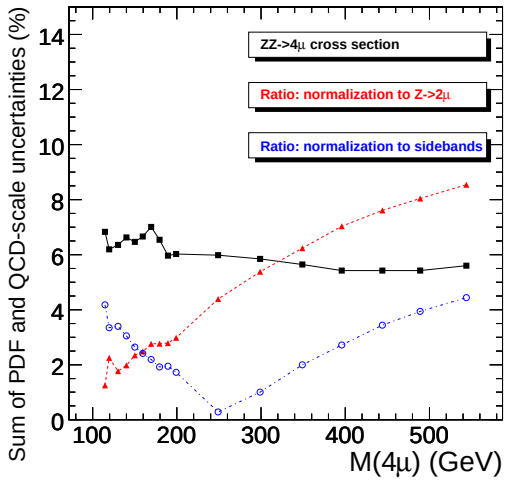


Figure 25: Combined systematic error on the number of background events due to PDF and QCD scale uncertainties for the $\sigma(q\bar{q} \to ZZ \to 4\mu)$ process at NLO. Shown are: absolute cross section uncertainties (black squares); uncertainties relative to the experimentally measured $q\bar{q} \to Z \to 2\mu$ cross section (open circles); uncertainties relative to the experimentally measured number of $M(4\mu)$ sideband events (triangles; side bands are defined as the full $M(4\mu)$ range 100 - 600 GeV$/c^2$, excluding the "signal" region).

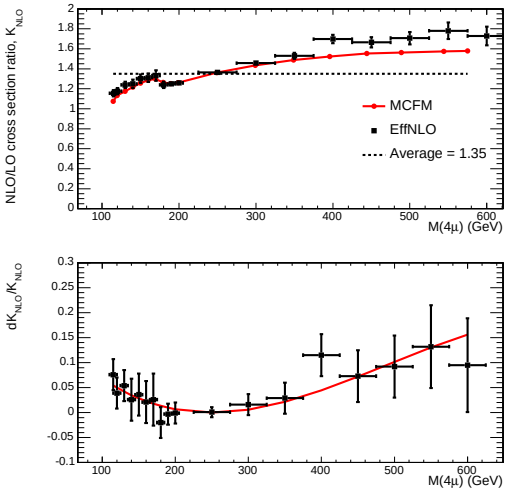


Figure 26: Top: the factors $K_{NLO}(M_{4\mu})$ in MCFM and EffNLO calculations versus $M(4\mu)$; bottom: the difference between them.

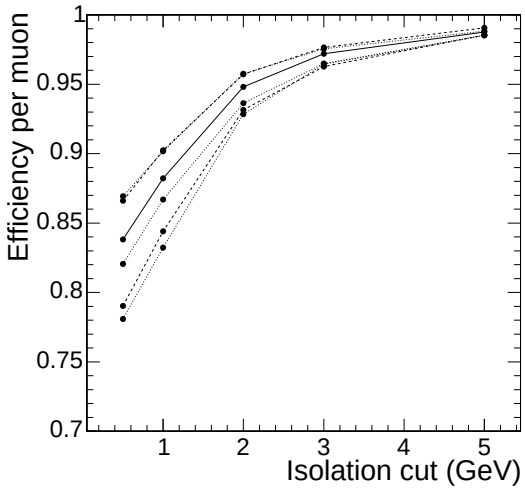


Figure 27: Muon isolation cut efficiency for random cone direction for Z-inclusive (dashed lines) and for ZZ (solid lines) events. The middle lines are for the default Pythia multi-parton interactions (MPI). The upper and lower lines correspond to uncertainties in modeling the MPI physics [26].

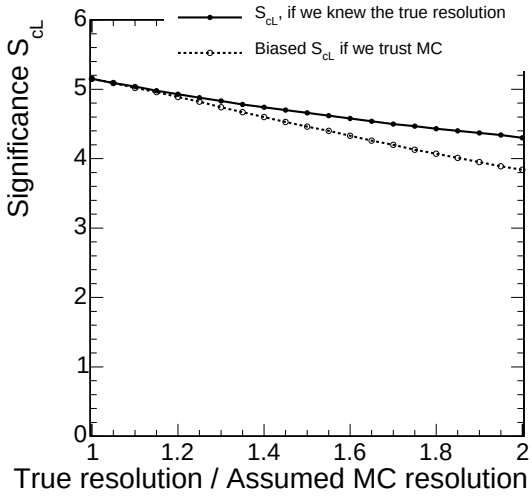


Figure 28: An example of a possible bias in evaluating the significance of an event excess due to a non-optimal choice of the signal window width for the case when the true detector resolution is worse than one assumes. The bias, being in the "conservative" direction, is very small even for errors in the $M(4\mu)$ resolution as large as 25% or worse

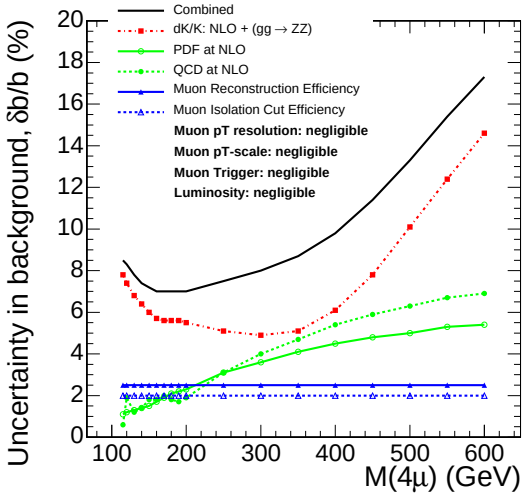


Figure 29: Uncertainties in the number of $ZZ \rightarrow 4\mu$ background events in the signal region window at different $M(4\mu)$. The window size is $\pm 2\sigma$ of the expected experimental Higgs resonance width. The event count is referenced to the number of $Z \rightarrow 2\mu$ events. Shown are: combined uncertainty (upper solid line), uncertainty on NLO k-factor (next dashed line with squares), PDF and QCD scale uncertainties at NLO (next two lines with circles: dashed and solid, respectively); muon reconstruction and muon cut efficiencies (the last two lines with triangles: solid and dashed, respectively).

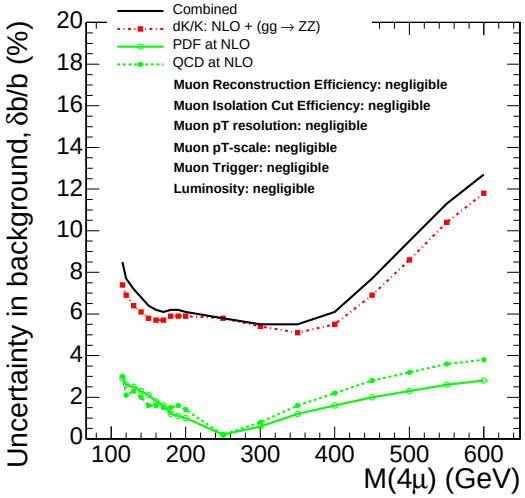
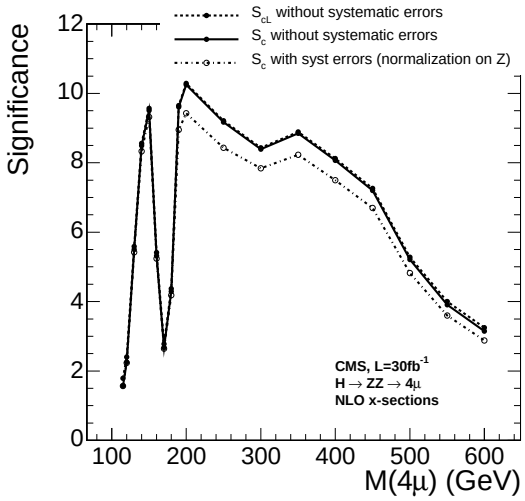


Figure 30: Uncertainties in the number of $ZZ \rightarrow 4\mu$ background events in the signal region window at different $M(4\mu)$. The window size is $\pm 2\sigma$ of the expected experimental Higgs resonance width. The event count in signal region, b, is calculated from the number of $ZZ \rightarrow 4\mu$ events in the range 100-700 GeV$/c^2$ (excluding the signal region window), B, via $b = \rho \cdot B$. (Same notations as Figure 29.)



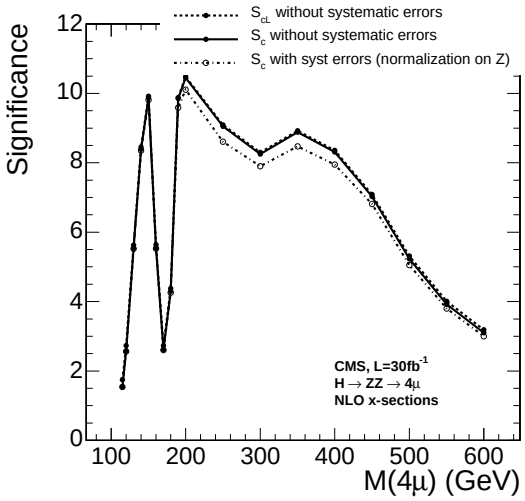


Figure 31: Significance vs. Higgs mass in the absence of any systematic errors (open circles, dashed) (see Figs. 21 and 22) and significance that includes all uncertainties in the background when it is estimated via referencing to the measured $Z \rightarrow 2\mu$ cross section (closed circles). Curves correspond to the total integrated luminosity of 30 fb^{-1}. For the left plot systematic uncertainty takes into account dK/K contribution (see Figure 29); for the right it is excluded.

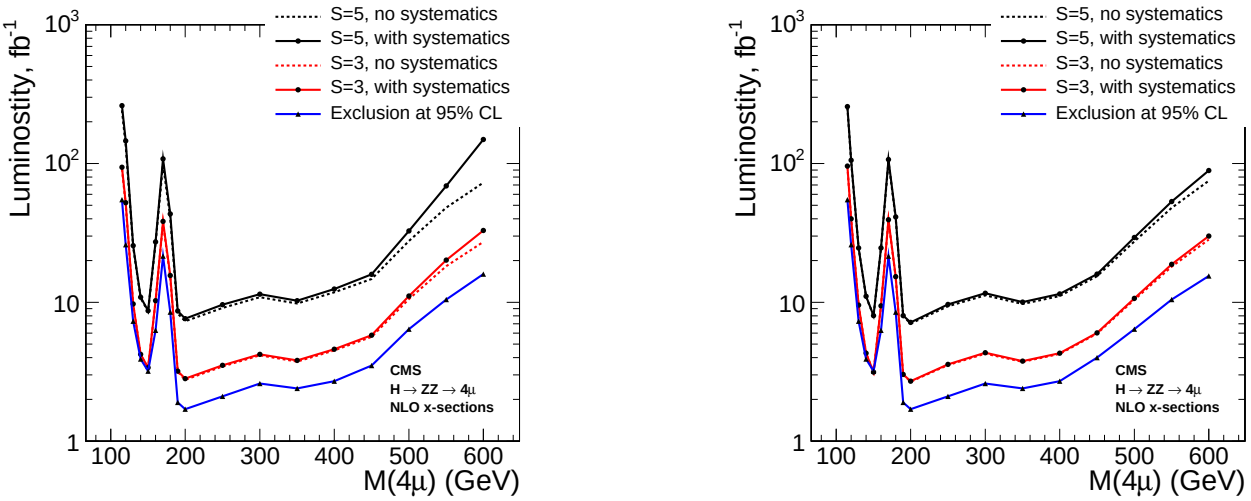


Figure 32: Integrated luminosity needed for 95%CL exclusion (triangles, lower line), 3σ (circles, two middle lines), and 5σ (circles, two upper lines) discovery versus Higgs boson mass. Curves for: no systematic errors on background; with systematic errors on background when it is estimated via referencing to the measured $Z \to 4\mu$. systematic errors when the background is estimated $M(4\mu)$ spectrum sidebands. For the left plot systematic uncertainty takes into account the dK/K contribution (see Figure 29); for the right it is excluded.

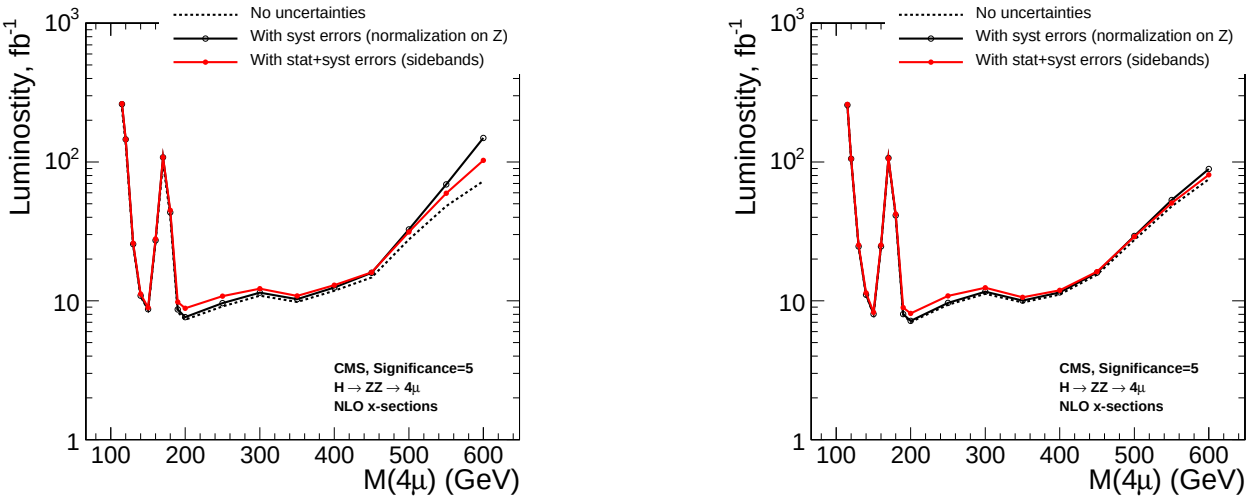


Figure 33: Integrated luminosity needed for a 5σ discovery of the Higgs boson versus its mass for: no systematic errors on background (dashed line), with systematic errors on background when it is estimated via referencing to the measured $Z \to 4\mu$ (empty circles) and with systematic and statistical errors when the background is estimated from the $M(4\mu)$ spectrum sidebands (filled circles). For the left plot systematic uncertainty takes into account the dK/K contribution (see Fig. 29); for the right it is excluded.

7 Appendix. Re-evaluating Significance of a Local Event Excess in the $H \to ZZ^{(*)} \to 4\mu$ Search

7.1 Introduction

In searching for a new phenomenon in a wide range of parameter phase space (e.g., in a search for a narrow resonance in a very broad range of invariant masses), one inevitably encounters a problem of overestimating significance of a "local discovery." The reason for an overestimation can be illustrated by an example of searching for a one-bin wide resonance in a histogram of, say, 1000 bins, representing a perfectly well-known background. The probability of a particular bin fluctuating up to give a 3σ excess or higher is $p = 0.14\%$. With 1000 bins being checked for an excess, the chance of finding at least one bin with such an excess is almost guaranteed, $1 - 0.9986^{1000} = 0.75$, or 75%, which makes a 3σ-excess in one of the bins completely insignificant. Note that this conclusion does not depend on the number of expected background events per bin (or, equivalently, the background spectrum shape), as long as one knows it.

This is a well-known and quite obvious fact described in papers on statistics dating back at least to 1957; see for example [27, 28]. The issue of overestimation of a significance when trying multiple hypotheses has been discussed in recent papers from the HEP community as well: see, for example, papers from CDF [29] and D0 [30] and a whole the software package "Sleuth" [31]; the issue is also discussed in astrophysics and high energy cosmic rays observations (see for example [32]), etc. The famous "Feynman trap" (the question about probability to observe a license plate with a particular number after it was seen) can be mentioned here as well.

However, even when the effect is acknowledged, it is not uncommon that people just shrug off the overestimation and argue that this is exactly the reason why the threshold of a discovery is raised to the 5σ level so that one would not have to worry about such things. This, however, does not relieve one from evaluating the scope of the effect quantitatively in a particular search and presenting it openly in a publications reporting an "$X\sigma$"-excess.

In this section, we estimate the scope of the effect (and discuss ways to reduce it) in searching for the Higgs-boson resonance in its $H \to ZZ^{(*)} \to 4\mu$ decay channel.

7.2 $H \to ZZ^{(*)} \to 4\mu$ Search Strategy

After applying optimal cuts on muon p_T, muon isolation, and di-muon masses, the expected $M(4\mu)$-spectrum for background at L=30 $fb^{-}1$ is shown in Figure 34. The average total number of background events above $M(4\mu) = 100$ GeV/c^2 is 225. The range used in the study is 110-600 GeV/c^2.

The search for the Higgs boson is performed by sliding a priori defined $M(4\mu)$ mass window (approximately $\pm 2\sigma(m_{4\mu})$ of the experimental signal widths, Figure 14) over the a priori defined range of 110-600 GeV/c^2 and looking for a local "significant" excess of events over the expected background using the S_{cL} estimator. Note that we already showed that the S_{cL} in this case follows very closely the true Poissonian significance. In these studies, we define signal as $s = n_o - b$, where n_o is the number of observed events and b is the expected background:

$$S_{cL} = sign(n_o - b)\sqrt{2n_o \ln(n_o/b) - 2(n_o - b)}. \quad (9)$$

The signed S_{cL} takes negative values for the cases when a lack of events is observed (the more significant the deficit, the larger the $|S_{cL}|$ is).

Although we use the counting experiment as the case study, the problem is of a very general nature. For example, the analysis using LLR built for the entire $M(4\mu)$ spectrum would suffer from the same problem as well, the actual scope of course being dependent on the details of the analysis.

One of the ways to reduce the scale of the effect is to employ various techniques for checking the consistency of an excess events with respect to the expected properties of the signal (e.g., its cross section, which was allowed to be free in these studies—this, however, is not expected to change results dramatically). In addition, the $M(4\mu)$-distribution shape can be validated for compatibility with the signal-plus-background hypothesis. Also, one can analyze the consistency of various observables, especially those not used in the analysis, with the expectations. Another method is to decrease a priori the "phase space" of possible parameters open for search. For example, the Higgs-boson mass in these studies was a free parameter that effectively had a flat prior probability. In principle, the Higgs boson mass prior probability could be forced to be consistent with the precision electroweak measurements. In addition, one can choose to use the early data to get an early indication/hint on the possible excess, form a

hypothesis on the Higgs-boson mass, discard the initial data, and evaluate significance of the Higgs-boson presence of now a well-defined mass using the new data. The optimal strategy is clearly yet to be defined. A note analyzing all these possibilities and discussing their pros and cons is in preparation.

7.3 True Significance Evaluation

The analysis is based on performing $\approx 10^8$ pseudo-experiments. Each pseudo-experiment is an ensemble of N randomly generated events with $M(4\mu)$ probability density function (pdf) as given by the expected background (Fig. 34). The smooth $M(4\mu)$ pdf is a fit of the background $M(4\mu)$ distribution as obtained in the $H \to ZZ^{(*)} \to 4\mu$ analysis based on the full detector simulation. The number of events N per pseudo-experiment is played out according to the Poisson distribution.

For each pseudo-experiment, we slide the mass window (see Fig. 14) with step 0.4 GeV/c^2 within a priori defined mass range (110-600 GeV/c^2) and identify the $M(4\mu)$ point M_{max} at which S_{cL} reaches its maximum S_{max}.

An example of a pseudo-experiment and results of the Higgs-boson search corresponding to it are shown in Figures 35 and 36. This particular example was picked from the many pseudo-experiments where a 5σ-discovery claim would be possible.

After performing 10^8 pseudo-experiments, we obtain the differential probability density function for S_{max} and its corresponding cumulative probability function $P(S_{max} > S)$. Both distributions are shown in Figures 37 and 38. The dashed (lower) line in Figure 38 shows the probability associated with the integral of a one-sided Gaussian tail above $S * \sigma$:

$$P = \int_S^{+\infty} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)\, dx = 0.5\,(1 - erf(S/\sqrt{2})). \tag{10}$$

One can see that for larger S_{cL}, the curve $P(S_{max} > S)$ is substantially higher than its naive interpretation would imply. Next, the real probability $P(S_{max} > S)$ reconstructed in these studies can be used to convert the observed S_{max} to its true significance S_{true} by using the inverse of the function shown above. The result of such renormalization is presented in Figure 39.

Recalling the example given in the first paragraph of the introduction section, it is clear that the conclusions on renormalization of a probability to observe an excess of events with a given "local significance" should not depend on the shape of the background or the cumulative number of events in the background spectrum (as the number of background events decreases, the structure of S_{max} pdf becomes more and more discrete corresponding to a smaller range of possible observed number of events). To illustrate this independence from the background shape explicitly, we performed another round of similar studies with pseudo-experiments, but now using a flat pdf for background with an average number of events of 4 GeV$^{-1}c^2$. An example of a pseudo-experiment with an accidental "5σ" excess at $M(4\mu) = 220$ GeV/c^2 and summary results of the S_{cL} scans for 10^8 pseudo-experiments are presented in Figures 40 and 41. As expected, one can see that the results for backgrounds with flat and "realistic" shape probability density functions are nearly identical.

7.4 Summary

For the case of the $H \to ZZ^{(*)} \to 4\mu$ search in the 110-600 GeV/c^2 range, the difference between the true probability to observe a local excess of events and a probability naively reconstructed from a "local significance" exceeds a factor of 200 (in the limit of small probabilities). This makes an observation of a "local 3σ" excess completely insignificant and dilutes the "significance of a local 5σ-discovery" to the true statistical significance of $\approx 3.8\sigma$. These conclusions depend only on the range of search and the width of a sliding window (or, equivalently, the experimental Higgs-boson resonance peak width); they remain valid for a background of any shape and level.

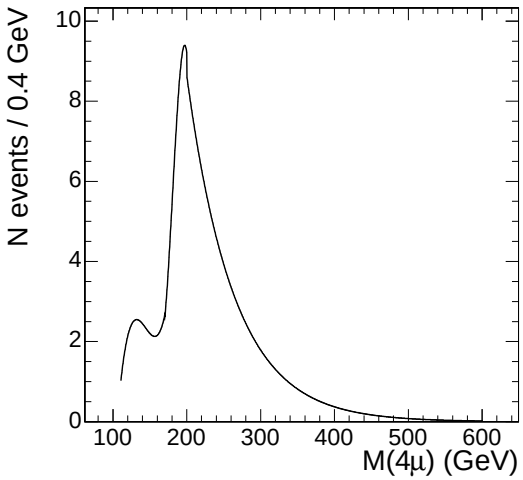


Figure 34: Background profile after applying all analysis cuts, except for the mass window cut.

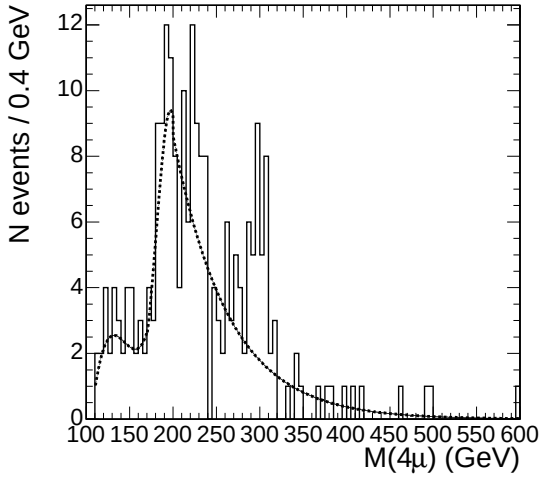


Figure 35: A pseudo-experiment example. N events is per 0.4 GeV$/c^2$ step superimposed on a 100 bin histogram for better visibility.

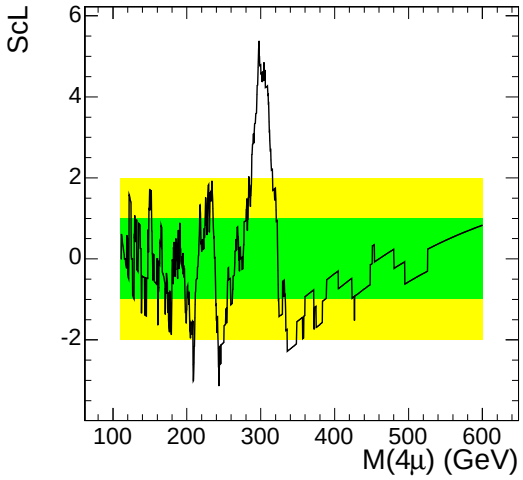


Figure 36: S_{cL} profile for the pseudo-experiment example shown on the left. Green (inner) and yellow (outer) bands denote $\pm 1\sigma$ and $\pm 2\sigma$ intervals. Spikes that can be seen are due to events coming in or dropping off the trial-window, a feature of searches with few data.

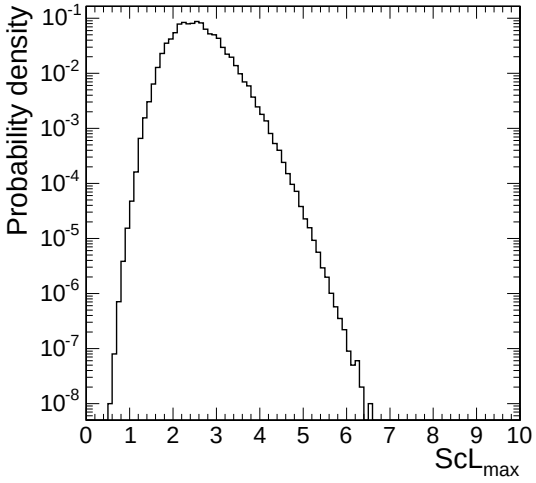


Figure 37: S_{cL} differential probability density function.

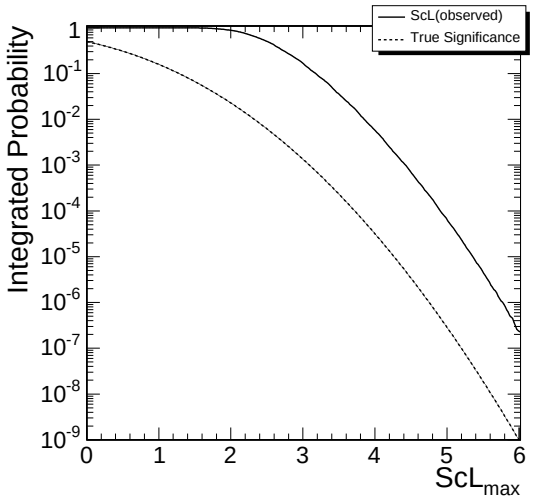


Figure 38: S_{cL} cumulative probability density function.

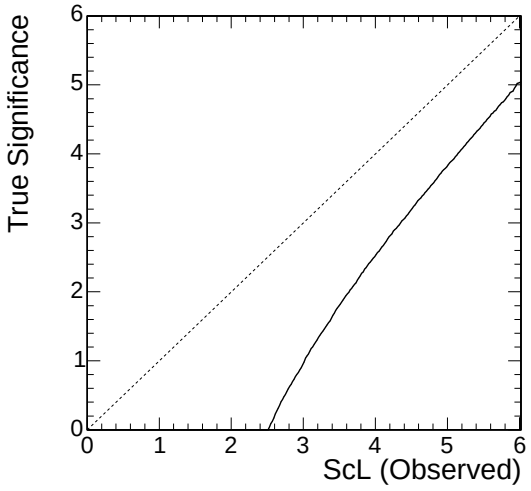


Figure 39: Local significance renormalization.

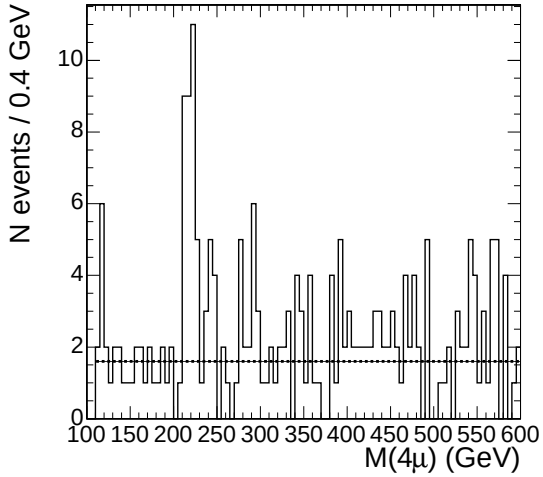


Figure 40: Pseudo-experiment example (flat background). N events is per 0.4 GeV/c^2 step superimposed on a 100 bin histogram for better visibility.

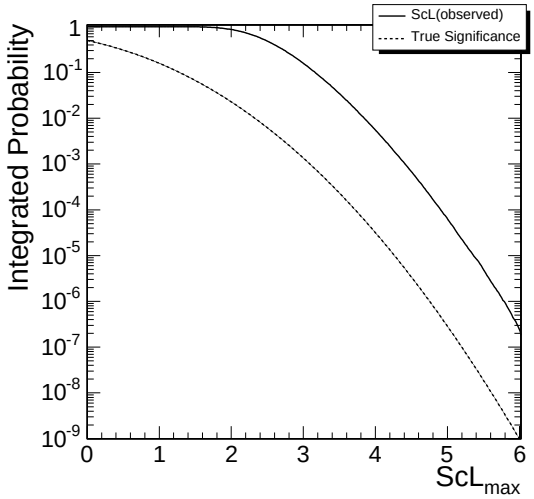


Figure 41: S_{cL} cumulative probability density function (flat background).