

THE STANDARD MODEL HIGGS BOSON DISCOVERY POTENTIAL IN THE
DECAY CHANNEL $H \rightarrow ZZ^* \rightarrow 4\mu$ WITH THE CMS DETECTOR

By

ALEXEY ALEXANDROVICH DROZDETSKIY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

Copyright 2007
by
Alexey Alexandrovich Drozdetskiy

Happiness in my life is to do whatever I do in the best way I can for others,
and especially for one – for Maria.

ACKNOWLEDGMENTS

Friendship is unnecessary, like philosophy, like art... It has no survival value; rather it is one of those things that give value to survival.
– C. S. Lewis

This is probably one of the hardest parts of the dissertation. And not because I have hard time in remembering all the people with whom I have worked, who helped me and supported me, but because no “sensible” limits for dissertation size would ever be able to accommodate the list of all of them. It’s so for a very simple reason: for example many people from the University were working with me, we used something produced before by others, our families and close to us people were loving and patient enough to allow us to quench our thirst in research, etc... and we can continue to expand the list this way on and on. Some say it takes just a few people to connect any one of us with any one else through common work, friends, etc. So it’s clear what an impossible task it is to list on a few pages all the people.

And I would like to thank all, who helped, supported and worked with me, and influenced me knowingly or unknowingly: I remember all in my mind and my heart. And if paper is not large enough, they are. I’m grateful to all, including University of Florida (UFL, Gainesville, USA) professors, post docs, students, and not less - all the wonderful office managers and personal; members of the CMS collaboration (for countless work hours and out-of-work meetings and discussions, help and support); people from Novosibirsk State University (Russia) and Budker Institute (Novosibirsk, Russia); all teachers and friends in my life.

Still, there are a few people I would like to thank a little more than others: Guenakh Mitselmakher and Andrey Korytov - my supervisors, for all their experience and knowledge they passed onto me, both on physics and not less importantly on life in general. I thank them for always giving me the freedom to get my own grasp on a topic. And maybe even more importantly I would like to thank them for their constant kind and

friendly attitude I'm enjoyed for many years. These are wonderful people and I'm very happy to learned from and to have been with them.

Others from UFL HEP group with whom I was happy to work with and spend free time together: Darin Acosta (who also was my co-supervisor for some time), Rick Cavanaugh, Victor Barashko, Bobby Scurlock, Kostya Kotov, Yura Pakhotin, Sergo Jindariani, Yura Oksuzian and Sasha Sherstnev (and their families).

Some CMS collaboration colleagues inspired me: Paris Sphicas, Daniel Denegri, Salavat Abdullin, Sasha Nikitenko, and Bob Cousins... Thank you.

My “out-of-work-group” friends, some of whom I have known since I was 7 years old: Igor, Eldar, Olga, Katerina, and Dasha (and their families).

I can expand this list and continue forever...

I would like to return to the thought I wrote in the beginning: all people in my life are very important and I thank each of them from the bottom of my heart and will keep them there for always.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	9
LIST OF FIGURES	10
ABSTRACT	18
CHAPTER	
1 INTRODUCTION	1
2 THE STANDARD MODEL HIGGS BOSON	3
2.1 The Standard Model: Electroweak Symmetry Breaking.	3
2.2 The Standard Model Higgs Boson Mass: Current Experimental and Theoretical Constraints	5
2.2.1 Theoretical Constraints	5
2.2.2 Searches at Large Electron-Positron Collider (LEP)	7
2.2.3 Searches at Tevatron	7
2.2.4 Electroweak Precision Measurements	8
2.3 The Standard Model Higgs Boson at LHC	8
2.3.1 Production	8
2.3.2 Decays	10
2.4 The Standard Model Higgs Boson Searches at CMS Summary: Discovery Reach	11
2.5 Issues in the Standard Model	12
2.6 Supersymmetry Higgs	14
3 LARGE HADRON COLLIDER	15
4 COMPACT MUON SOLENOID (CMS) DETECTOR	18
4.1 Introduction	18
4.2 General Description	19
4.3 The Superconducting Magnet	21
4.4 Inner Tracking System	22
4.5 Electromagnetic Calorimeter (ECAL)	23
4.6 Hadron Calorimeter (HCAL)	24
4.7 Muon System	25
4.7.1 Barrel Muon System	25
4.7.2 Endcap Muon System	26
4.7.3 Resistive Plate Chambers	26
4.8 Trigger	26
4.8.1 Muon Trigger	26

4.8.1.1	CSC local trigger	27
5	CATHODE STRIP CHAMBERS (CSC)	28
5.1	Introduction	28
5.2	Chamber Mechanical Design	32
5.3	Electronic Design	36
6	FAST ALGORITHM FOR TRACK SEGMENT AND HIT RECONSTRUCTION IN CATHODE STRIP CHAMBERS	44
6.1	Introduction	44
6.2	Algorithm Principles	46
6.2.1	ProtoSegment Building	47
6.2.2	RecHits and Their Coordinates	48
6.2.3	RefinedSegment	51
6.3	Algorithm Validation at MTCC	52
6.3.1	ProtoSegments	52
6.3.2	Spatial Resolution	53
6.3.2.1	Resolution for high- p_T muons	55
6.3.2.2	Resolution for high- p_T muons with tabulated pedestals . .	60
6.3.2.3	Resolution for low- p_T , or highly-inclined, muons	62
6.3.2.4	Resolution for high- p_T muons with charge-dependent sigmas	63
6.3.3	RefinedSegments	63
6.4	Conclusions	65
6.5	Ratio Method	68
7	SEARCH STRATEGY FOR THE STANDARD MODEL HIGGS BOSON IN THE $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ DECAY CHANNEL USING $M(4\mu)$ -DEPENDENT CUTS.	73
7.1	Introduction	73
7.2	Physics Processes and Their Simulation	75
7.2.1	Signal: $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$	76
7.2.2	Background: $t\bar{t}$	77
7.2.3	Background: $(Z/\gamma^*)b\bar{b} \rightarrow 2\mu b\bar{b}$	77
7.2.4	Backgrounds: $q\bar{q} \rightarrow Z/\gamma^* Z/\gamma^* \rightarrow 4\mu$ and $q\bar{q} \rightarrow Z/\gamma^* Z/\gamma^* \rightarrow 2\mu 2\tau$. .	78
7.3	Trigger and Offline Muon Reconstruction	78
7.4	Higgs Search Strategy	79
7.4.1	Introductory Remarks on Significance	80
7.4.2	Optimization of the $M(4\mu)$ -Dependent Cuts	82
7.4.3	Systematic Errors	86
7.4.3.1	Uncertainties in the background	86
7.4.3.2	Significance with the background uncertainties included . .	91
7.4.4	Local Significance and Overall Statistical Fluctuation Probability . .	92
7.5	Summary	93
7.6	On the true significance of a local excess of events	94

8	CONCLUSIONS	109
APPENDIX		
A	RELATIVE CONTRIBUTIONS OF T- AND S-CHANNELS TO THE $ZZ^{(*)} \rightarrow 4\mu$ PROCESS	111
B	ZZ DISCOVERY WITH $< 1FB^{-1}$ WITH CMS IN $ZZ^{(*)} \rightarrow 4\mu$ MODE	113
C	MEASURING MUON RECONSTRUCTION EFFICIENCY FROM DATA	114
D	SENSITIVITY OF THE MUON ISOLATION CUT EFFICIENCY TO THE UNDERLYING EVENT UNCERTAINTY	116
E	GENETIC ALGORITHM FOR RECTANGULAR CUTS OPTIMIZATION	117
REFERENCES		118
BIOGRAPHICAL SKETCH		126

LIST OF TABLES

<u>Table</u>		<u>page</u>
2-1 The Standard Model elementary particles.	3	
7-1 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 7.2) calculated for $L = 30\text{fb}^{-1}$, and the number of simulated events.	76	
7-2 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 is normalized to the $Z \rightarrow 2\mu$ 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.	93	

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 Left: Divergent (top three) WW cross-section graphs and their cancellation (bottom two graphs). Right: Theoretical bounds on the SM Higgs mass as a function of the cut-off scale. It is assumed that the SM is a valid theory up to the scale Λ . The requirement of avoiding a Landau pole provides the upper bound and stability of the vacuum provides the lower bound [32].	6
2-2 Left: Expected and observed 95% CL cross section ratios for the combined CDF and D0 analyses, and the expected 95% CL ratios for the CDF and D0 experiments alone [34]. Right: The $\Delta\chi^2$ curve derived from high-Q ₂ precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs boson mass, assuming the SM to be the correct theory of nature [35].	7
2-3 Left: Typical diagrams for all relevant Higgs boson production mechanisms at leading order at the LHC.: (a) gluon fusion, (b) vector boson fusion, (c) Higgs-strahlung, (d) Higgs bremsstrahlung off top quarks. Right: Higgs production cross sections at the LHC for the various production mechanisms as a function of the Higgs mass. The full QCD-corrected results for the gluon fusion $gg \rightarrow H$, vector-boson fusion $qq \rightarrow VVqq \rightarrow Hqq$, vector-boson bremsstrahlung $q\bar{q} \rightarrow V^* \rightarrow HV$ and associated production $gg, q\bar{q} \rightarrow Ht\bar{t}$ are shown.	9
2-4 Left plot: total decay width (in GeV/c^2) of the SM Higgs boson as a function of its mass. Right plot: Branching ratios of the dominant decay modes of the SM Higgs particle. All relevant higher-order corrections are taken into account.	11
2-5 Left: The integrated luminosity needed for the 5σ discovery of the inclusive Higgs boson production $pp \rightarrow H + X$ with the Higgs boson decay modes $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$, and $H \rightarrow WW \rightarrow 2\ell 2\nu$ [8]. Right: The signal significance as a function of the Higgs boson mass for 30 fb^{-1} of the integrated luminosity for the different Higgs boson production and decay channels [8].	12
3-1 Accelerator complex at CERN.	16
4-1 An exploded view of the CMS detector.	20
5-1 Quater-view of the CMS Detectors. Cathode Strip Chambers of the Endcap Muon System are highlighted.	29
5-2 The ME2 cathode strip chambers. The outer ring consists of 36 ME2/2 chambers, each spanning 10° in ϕ ; while 18 20-degree ME2/1 chambers form the inner ring. The chambers overlap to provide contiguous coverage in ϕ	30

5-3	Left: A schematic view of a CMS cathode strip chamber made of seven trapezoidal panels. The panels form six gas gaps with planes of anode sensitive wires. The cutout in the top panel reveals anode wires and cathode strips. Only a few wires are shown to indicate their azimuthal direction. Strips of constant $\Delta\phi$ run lengthwise (radially). The largest CMS CSCs are 3.4 m long along the strip direction and up to 1.5 m wide along the wire direction. Right: A schematic view of a single gap illustrating the principle of a CSC operation. By interpolating charges induced on cathode strips by avalanche positive ions near a wire, one can obtain a precise localization of an avalanche along the wire direction.	31
5-4	Mechanical design of the CMS Cathode Strip Chambers (exploded view).	32
5-5	Left: CSC gas gain vs high voltage. Right: ME2/1 chamber singles rate vs high voltage (the overall sensitive area of all six planes in this chamber is ~ 9.5 m 2).	35
5-6	Left: If wires in ME1/1 chambers were not tilted, ionization electrons, as they drift towards anodes wires in strong magnetic field normal to the plane of the drawing, would be carried sideways by the Lorentz force. Direction and the size of the shift depend on whether electrons drift upward or downward and on how far away they are from wires to begin with. These sideway displacements would spread the charge over anode wires. Right: By tilting wires at the Lorentz angle α_L , one makes all ionization electrons arrive at the same point.	36
5-7	An overview of all custom-made CSC trigger and readout electronics. Functionality of various boards is described in the text.	37
5-8	Left: Muon signals as seen at the AFEB amplifier output. Right: A schematic event display showing anode signals in a six-plane chamber (small dark squares). The ALCT board FPGA logic is programmed to scan along the chamber and search for hits falling inside predefined patterns (gray cells) consistent with muons originating from the interaction point. For a pattern to be found, hits from at least four planes should be present in it.	39
5-9	Left: Basic functional diagram of a Cathode Front-End Board (functionality of the main blocks is explained in the text) and CFEB amplifier-shaper response on a δ -function input pulse. The overshoot is intended to compensate for the long tail present in muon hit signals resulting from a slow drift of positive ions away from anode wires. Right: Muon hit signals from six nearby strips. Four curves are actual oscilloscopes. Six lines of dots are digitized outputs.	40

5-10	Left: A simplified schematic illustrating the idea behind the comparator network capable of quickly localizing the muon hits within one half of a strip width. For each triad of nearby strips, comparators compare the central strip signal Q_c with a threshold, central-to-left $Q_c - Q_l$, central-to-right $Q_c - Q_r$, and right-to-left $Q_r - Q_l$. If $Q_c >$ threshold, $Q_c - Q_l > 0$, $Q_c - Q_r > 0$, $Q_r - Q_l > 0$ (as in this figure), the hit position is somewhere within the right half of the central strip. Right: CLCT (Cathode Local Charged Track) is a pattern of half-strip hits consistent with a muon track.	41
6-1	Left: The noise distribution for a three-sample-sum in the absence of a signal for the largest ME23/2 chambers. Right: The three-strip cluster charge distribution for the largest ME23/2 chambers.	53
6-2	Left: The noise distribution for a single-sample in the absence of a signal for the largest ME23/2 chambers. Right: The noise distribution for a two-sample sum in the absence of a signal for the largest ME23/2 chambers.	54
6-3	The 3rd-plane residuals for the five distinct $ x $ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5) for the middle part of the largest ME23/2 chambers (HV segment 3). The last plot (bottom right) shows a summary for a single-plane spatial resolution as a function of a hit position across a strip (hit coordinate is given in strip width units $ x $).	57
6-4	Summary of resolution results for the largest ME23/2 chambers. Left: the 3rd-plane residuals σ 's vs. five $ x $ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5). Right: overall six-plane resolution vs. $ x $ -part of a strip. Top plots show results in units of strip widths, the bottom—in absolute microns. [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.] . . .	58
6-5	Summary of resolution results for the top part of the ME1/1 chambers. Left: the 3rd-plane residuals σ 's vs. $ x/w $ -part of a strip. Right: overall six-plane resolution vs. $ x/w $ -part of a strip. Top plots show results in units of strip widths, the bottom—in absolute microns. [Notations are: black squares - HV segment 1.] . . .	59
6-6	The overall six-plane chamber resolution vs. five $ x $ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5) for dynamically measured pedestals (left) and calibrated pedestals (right). [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.] . . .	61
6-7	Left: Inclination angles for track segments reconstructed from di-strip CLCTs (ME1/1 chambers). Right: The spatial resolution per six-plane chamber for tracks reconstructed from di-strip CLCTs (ME1/1 chambers) for five $ x $ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5). [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.] . . .	61

6-8	Left: Probabilities for different numbers of charge clusters found in association with ME23/2 chamber ProtoSegments. Right: Layer charge clusters occupancy in association with ME23/2 chambers ProtoSegments.	64
6-9	Left: The χ^2/dof distributions for a linear fit for ME23/2 chamber ProtoSegments with six clusters (solid curve), five clusters (dashed curve), and four clusters (dashed-dotted curve). Right: Scatter plot of $\min(\chi_5^2/\text{dof})$ vs. χ_6^2/dof for six-cluster ME23/2 chamber ProtoSegments. Here, χ_6^2/dof is for a six-plane fit, $\min(\chi_5^2/\text{dof})$ is a minimum between 6 five-plane fits with one plane left out from the fit.	65
6-10	Left: Occupancy for $ x/w $ -coordinates of RefinedSegments. Right: Pools for all retained RecHits (i.e., RecHits associated with RefinedSegments.)	66
6-11	Left: Induced charge distribution calculated according to Gatti et al. [77] for large and ME1/1 chamber geometries. The Gatti parameter K_3 was taken to be 0.334 for large chambers and 0.379 for ME1/1 according to the empirical approximations in Ref. [78, 79]. Right: Ratio r versus a local coordinate x for large chambers calculated for variety of strip widths in the assumption of the Gatti charge distribution for large chambers (see plot on the left).	66
6-12	Left: Correction that must be added to the measured ratio r to obtain the hit position across a strip. Points correspond to the “theoretical” Gatti-based curve from Fig. 6-11, while the lines correspond to an empirical approximation described in the text as the 1st-order correction. Right: An occupancy distribution for the 1st-order corrected coordinate x^{1st}	69
6-13	Left: The second-order correction that must be added to x_1 to obtain the hit position across a strip x_2 Right: An occupancy distribution for the 2nd-order corrected coordinate x_2 (as expected, the distribution now is almost perfectly flat).	70
6-14	Sensitivity of the ratio method to electronic noise. The curves are given for $\delta Q_{\text{strip}}/Q_{\text{cluster}} = 0.01$, which corresponds to the nominal operation conditions. Pedestals are assumed to be tabulated. When pedestals are defined on the event-by-event basis, the noise contribution increases by a factor of 1.6. The resolution is given in units of strip width (left) and in microns (right)	71
6-15	Sensitivity of the ratio method to errors in electronic gain calibrations. The curves are given for $\delta \text{Gain}/\text{Gain} = 0.01$. The choice of this number is driven by the actual measurements of strip electronic gain variations between nearby strips (this is the only thing that matters), which gave 1.1%. The resolution is given in units of strip width (left) and in microns (right).	71
6-16	Left: Sensitivity of the ratio method to uncertainties in cross talks between strips. The curves are obtained for a 2% uncertainty in crosstalk, which corresponds closely to variations in crosstalk values between different chamber types (sizes). The resolution is given in units of strip width (left) and in microns (right).	72

7-1	$M(4\mu)$ distributions after generator-level cuts for $t\bar{t}$, $(Z/\gamma^*)b\bar{b}$, $Z/\gamma^*Z/\gamma^*$, and $m_H = 140 \text{ GeV}/c^2$ (log scale).	95
7-2	Same as Figure 7-1, but on a linear scale.	95
7-3	Enhancement to the signal samples' cross sections due to interference effects not accounted for at the generator level.	95
7-4	The $M(4\mu)$ -dependent NLO K-factor $K_{NLO}(M_{4\mu})$ for the $ZZ \rightarrow 4\mu$ process evaluated with MCFM [100].	95
7-5	Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the barrel region vs. p_T	96
7-6	Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the endcap region vs. momentum.	96
7-7	$M(4\mu)$ distribution for $m(H)=150 \text{ GeV}/c^2$ and the fit described in the text.	96
7-8	$M(4\mu)$ resolution vs. M_H	96
7-9	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.	97
7-10	Same as Figure 7-9, but for $N_B = 10$ events.	97
7-11	Dependence of the tracker-based muon isolation cut on the least isolated muon versus Higgs mass. The smooth dependence shown by the curve which follows general dependence of the cut variable was used for analysis-level cuts.	97
7-12	Dependence of the calorimeter-based cut on the least isolated muon versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.	97
7-13	Dependence of the p_T cut on the second-lowest- p_T muon versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.	98
7-14	Dependence of the $M(4\mu)$ window cuts versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.	98
7-15	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.	98
7-16	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.	98
7-17	Lowest muon p_T distribution, after analysis cuts were applied.	99

7-18 Maximum distance in XY-plane between muon impact point coordinates distribution, after analysis cuts were applied.	99
7-19 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).	100
7-20 $M(4\mu)$ invariant mass distribution for the three background subprocesses and a Higgs boson signal at $M_H = 150 \text{ GeV}/c^2$, after applying cuts on muon isolation and p_T	100
7-21 Expected excess significance S_{cL} with $L = 30 \text{ fb}^{-1}$ for different Higgs boson masses for $M(4\mu)$ -dependent (solid line) and independent cuts (dashed line). No systematic errors included.	101
7-22 Expected excess significance S_L with $L = 30 \text{ fb}^{-1}$ for different Higgs boson masses for $M(4\mu)$ -dependent (solid line) and independent cuts (dashed line). No systematic errors included.	101
7-23 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.	101
7-24 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).	101
7-25 Combined systematic error on the number of background events due to PDF and QCD scale uncertainties for the $\sigma(q\bar{q} \rightarrow ZZ \rightarrow 4\mu)$ process at NLO. Shown are: absolute cross section uncertainties (black squares); uncertainties relative to the experimentally measured $q\bar{q} \rightarrow Z \rightarrow 2\mu$ 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).	102
7-26 Top: the factors $K_{NLO}(M_{4\mu})$ in MCFM and EffNLO calculations versus $M(4\mu)$; bottom: the difference between them.	102
7-27 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 [4].	103

7-28 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	103
7-29 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).	104
7-30 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 7-29.)	104
7-31 Significance vs. Higgs mass in the absence of any systematic errors (open circles, dashed line) (see Figs. 7-21 and 7-22) 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 7-29); for the right no dK/K contribution is considered.	105
7-32 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 $Z \rightarrow 4\mu$. 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 7-29); for the right it is excluded.	105
7-33 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 $Z \rightarrow 4\mu$ (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. 7-29); for the right it is excluded.	106
7-34 The background pdf and an example of one pseudo-experiment with a statistical fluctuation appearing just like a signal.	108

7-35 Profile of the S_{cL} scan corresponding to 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 low-statistics searches.	108
7-36 S_{cL} cumulative probability density function.	108
7-37 Local significance “renormalisation” from an observed value to the true significance with a proper probabilistic interpretation.	108
A-1 Left: ZZ background: t- and s- channel diagrams. Right: S-channel contribution peak around Z^0 mass after pre-selection cuts for fully simulated events.	112
C-1 $Z \rightarrow 2\mu$ invariant mass peak: 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.	115

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE STANDARD MODEL HIGGS BOSON DISCOVERY POTENTIAL IN THE
DECAY CHANNEL $H \rightarrow ZZ^* \rightarrow 4\mu$ WITH THE CMS DETECTOR

By

Alexey Alexandrovich Drozdetskiy

August 2007

Chair: Guenakh Mitselmakher

CoChair: Andrey Korytov

Major Department: Physics

The Compact Muon Solenoid (CMS) is a general purpose detector at the Large Hadron Collider (LHC) currently under construction at CERN with start-up date in 2008. The putative Higgs boson is the most plausible solution for the mystery of mass in the Standard Model of elementary particles and its discovery is one of the prime goals for the LHC.

Higgs boson search. Our research presents an optimized strategy for a Higgs boson search in its four-muon decay channel, $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$, also known as a golden decay channel. The method automatically ensures an optimal signal-to-background ratio for any mass, at which the Higgs boson might appear. The most important theoretical and instrumental systematic errors are taken into account and our search was conducted in a broad range of possible Higgs boson masses.

Muon reconstruction. We developed an algorithm for a fast and efficient muon track segment reconstruction in Cathode Strip Chambers. Designed to be CPU-efficient, the algorithm is specifically targeted for High Level Trigger purposes. The segment-finding efficiency and the spatial resolution attainable with the proposed algorithm and the required CPU time were validated using the cosmic ray data taken by the CMS in 2006. The results of validation showed the efficiency and spatial resolution attainable with the

algorithm are well within the High Level Trigger requirements and the algorithm's timing performance is by far superior to all algorithms previously used in CMS.

CHAPTER 1

INTRODUCTION

If we knew what it was we were doing,
it would not be called research, would it?
– Albert Einstein

The Standard Model (SM) as a fundamental theory of elementary particles and their interactions has been so far very successful. But a very important question of this theory is still unanswered: does the Higgs boson exist? Our dissertation describes work done by the author, as a member of the Compact Muon Solenoid (CMS) collaboration.

One of the main goals of CMS is to discover Higgs boson. For this we will need good understanding of our detector and data, and we need to prepare data analysis tools to be ready to discover new physics we are looking for. The dissertation describes author's work in both directions.

First, it gives details on commissioning of Cathode Strip Chambers (CSCs), CMS endcap muon detectors, through analyzing data and validating new muon track hits and segments reconstruction algorithm, which would be suitable for High Level Trigger (HLT). Note that default algorithm used for this at the time of our analysis were just too slow for HLT.

Another major part of the dissertation – is our CMS the SM Higgs boson discovery potential in $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel with $M(4\mu)$ dependent cuts analysis using Monte Carlo (MC) data. Here we describe our analysis, which is much more realistic than previously existed ones on the topic:

- we used calibration from data techniques, as we would do with real data;
- full treatment of systematic errors was included and folded into signal-to-background significance calculations;
- dedicated effort was done to generate physics of signal and background processes properly, including use of dedicated matrix element generators, events re-weighting with dynamic Next-to-Leading Order (NLO) corrections;

- latest available at the moment full simulation and reconstruction validated software was used;
- CMS performance was optimized for all allowed SM Higgs boson masses;
- and as a result our analysis is part of the CMS Physics Technical Design Report, i.e. is an official CMS strategy for the SM Higgs boson discovery in $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel.

CHAPTER 2

THE STANDARD MODEL HIGGS BOSON

2.1 The Standard Model: Electroweak Symmetry Breaking.

The best known description of the fundamental particles and their interactions¹ is provided by a quantum field theory known as the Standard Model (SM). We will mention here some of the most important features of the model, not attempting by any means to compete with many already existing good descriptions of it.

The theory includes three generations of fermions (quarks and leptons, spin 1/2-particles) as well as interaction mediators (spin-1 bosons: γ , Z^0 , W^\pm and gluons - mediators of electromagnetic, weak and strong forces) (Tab. 2-1).

Table 2-1: The Standard Model elementary particles.

		Particles	Spin	Charge
Fermions	Leptons	$(\begin{smallmatrix} e \\ \nu_e \end{smallmatrix})$	1/2	$(\begin{smallmatrix} -1 \\ 0 \end{smallmatrix})$
	Quarks	$(\begin{smallmatrix} u \\ d \end{smallmatrix})$	1/2	$(\begin{smallmatrix} 2/3 \\ -1/3 \end{smallmatrix})$
Bosons	Gauge	γ	1	0
		Z^0, W^\pm	1	0, ± 1
		8 gluons	1	0
	[Higgs]	$[H^0]$	[0]	[0]

The description of the SM interactions is implemented by a gauge theory based on $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry. Gauge symmetry provides bosons, which serve as interaction mediators. Local gauge invariance makes the theory renormalizable and requires the gauge bosons to be massless. At the same time we know that Z^0, W^\pm bosons have masses.

The solution of the problem in the SM is the Higgs mechanism, which spontaneously breaks the gauge symmetry. Corresponding scalar potential added to the Lagrangian generates the vector boson and fermion masses in a gauge invariant way. This remnant

¹ gravitational effects not included

scalar field, the Higgs boson, is a part of the physical spectrum. This is the only missing piece of the SM that is still not confirmed experimentally.

In the 1960s, the theory of Quantum Electrodynamics (QED) was already very successful, and the theoretical community evolved Fermi's theory of weak interactions into a Yang-Mills theory [22] based on the symmetry group $SU(2)_L$. The problem with this approach was that gauge invariance does not allow masses for both the gauge bosons and the leptons. The observation of Peter Higgs was that the gauge invariance could be spontaneously broken with the addition of a doublet of complex scalar fields, ϕ , with Lagrangian

$$L_{Higgs} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \quad (2-1)$$

where the potential

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2-2)$$

is the key to spontaneous symmetry breaking [23, 24]. Having this mechanism for electroweak symmetry breaking, Glashow [25], Weinberg [26], and Salam [27] proposed a unified electroweak theory of the leptons. This theory has still a massless photon, allows for massive W bosons and leptons, predicts the massive, neutral, spin-1 Z boson, and predicts the massive, neutral scalar Higgs boson. The W^\pm and Z bosons were confirmed at the CERN SPS by the UA1 and UA2 collaborations [28, 29].

The Higgs boson itself acquires mass through self-coupling in the Higgs potential $V(\phi)$ (Eq. 2-2). The Higgs mass at the tree level is

$$m_H^2 = 2\lambda v^2 \quad (2-3)$$

where λ is an independent and unknown parameter. m_H can not be predicted by the SM.

At present the SM gives an excellent description of nature. It is a renormalizable quantum field gauge theory with massive fermions and vector gauge bosons. The only unknown parameter of the SM is the mass of the Higgs boson.

2.2 The Standard Model Higgs Boson Mass: Current Experimental and Theoretical Constraints

2.2.1 Theoretical Constraints

Besides allowing the introduction of Z/W-boson masses without breaking renormalizability and quark-lepton masses within the weak sector of SM with broken P-symmetry, the introduction of a new scalar particle was motivated by divergences in the scattering of a longitudinally polarized W bosons in the high energy limit [30, 31]. Without an additional interaction, the cross-section of that process (Fig. 2-1), would diverge and would violate unitarity bounds above $\sqrt{s} = 1.2$ TeV. An interaction with the Higgs boson cancels those divergences. This mechanism would work only if the Higgs boson is not too heavy, otherwise, the Higgs boson would not contribute enough to the scattering amplitude before unitarity is violated. Therefore, the Higgs boson mass should be less than 850 GeV [31].

Stronger bounds on the Higgs boson mass can be derived from the energy scale up to which the SM should be valid without the necessity of introducing new physics [30, 31]. If the Higgs boson is light, Higgs self-coupling strength λ is small and dominant loop contribution to the Higgs potential comes from top loops (λ_t is large due to the large mass of the top quark). If the SM is valid as an effective theory up to the scale Λ , then these loop contributions have to be summed up to this scale. For the Higgs mechanism to remain valid, the coupling λ must remain positive, otherwise no minimum exists in the potential and no stable spontaneous symmetry breaking occurs. This places a lower bound on λ , hence on m_H , depending on cut-off scale Λ (Fig. 2-1, right).

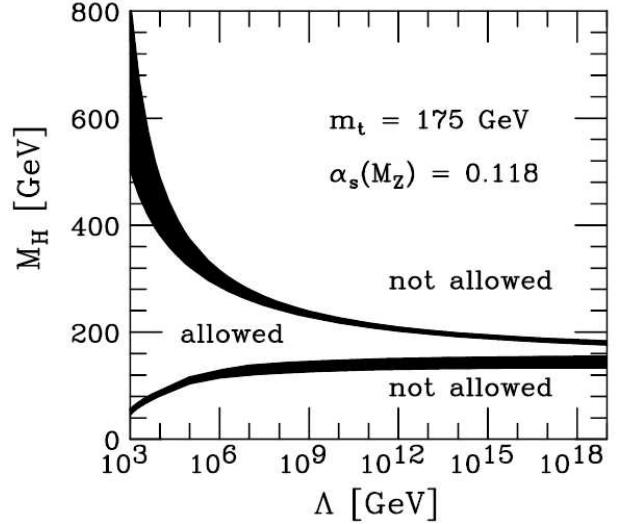
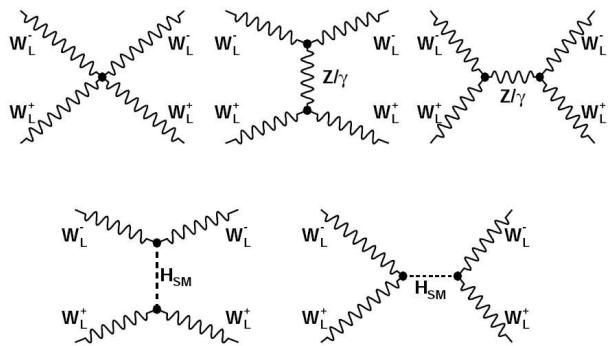


Figure 2-1: **Left:** Divergent (top three) WW cross-section graphs and their cancellation (bottom two graphs). **Right:** Theoretical bounds on the SM Higgs mass as a function of the cut-off scale. It is assumed that the SM is a valid theory up to the scale Λ . The requirement of avoiding a Landau pole provides the upper bound and stability of the vacuum provides the lower bound [32].

The energy dependence $\lambda(Q^2)$ can be derived from the renormalization group equations [30, 31]. If λ is large, the Higgs loop dominates over the top loop. Neglecting the graph with a top-quark loop, we can write λ as

$$\lambda(Q^2) = \frac{\lambda(v^2)}{1 - \frac{3\lambda(v^2)}{8\pi^2} \ln(Q^2/v^2)} \quad (2-4)$$

For a heavy Higgs boson, λ could grow to infinity (Landau-pole). Requiring that the self-coupling λ remains finite for arbitrary values of Q implies $\lambda(v) = 0$. Since $\lambda(v^2) = m_H^2/2v^2$, this would result in the non-interacting trivial theory. If, instead, λ is required to be finite only up to a scale Λ_{NP} , where the new physics enters, the mass bound can be written as

$$m_{H^0}^2 < \frac{8\pi^2 v^2}{3 \ln \Lambda_{NP}^2 / v^2} \quad (2-5)$$

One can read that if there is no new physics up to the Planck scale ($\Lambda \approx 10^{19}$ GeV), m_H must be $\sim 160 - 180$ GeV [32] (Fig. 2-1).

2.2.2 Searches at Large Electron-Positron Collider (LEP)

Direct searches for the Higgs boson have been conducted at the LEP experiments at CERN. No evidence for a signal was observed in a data from e+e- collisions up to the center of mass energies of 209 GeV. An experimental lower bound is set to $m_H > 114.4$ GeV at the 95% confidence level [33].

2.2.3 Searches at Tevatron

The search for the SM Higgs particle is continuing at the Tevatron for Higgs masses up to $\sim 130 \text{ GeV}/c^2$.

The most recent update was presented e.g. at the Moriond conference in March, 2007 and shows joint CDF/D0 sensitivity curves together with the SM prediction. The Tevatron experiments are close to the SM predictions and may add to competition during early LHC data taking (Fig.2-2).

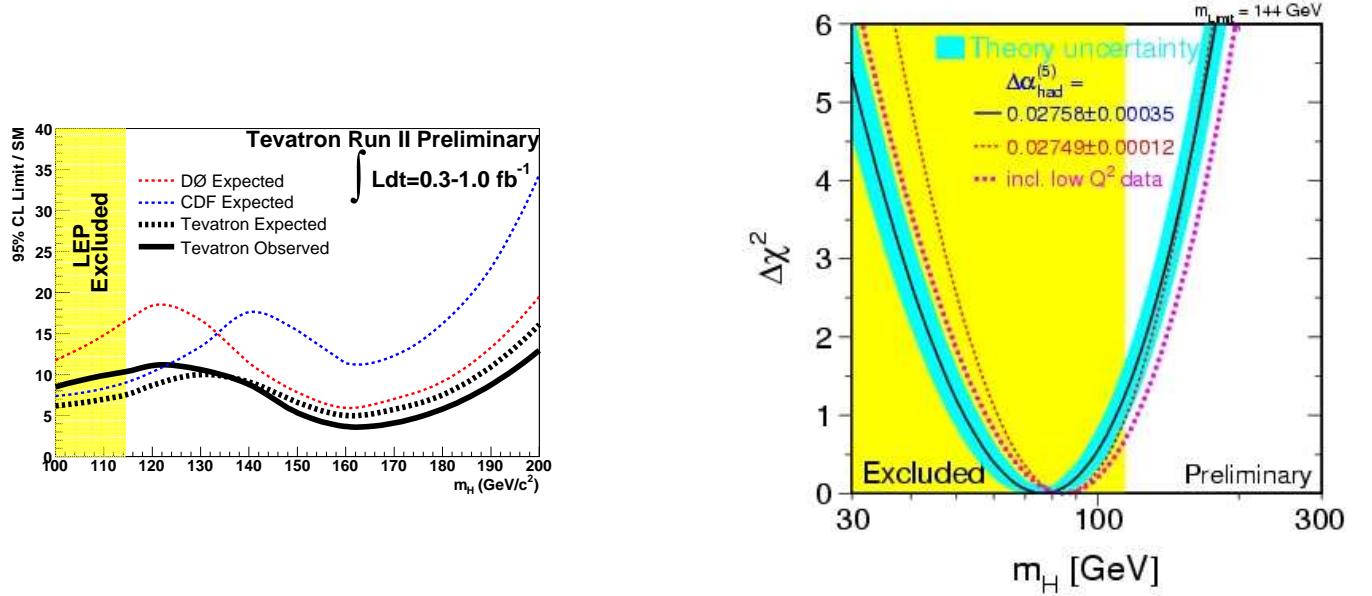


Figure 2-2: **Left:** Expected and observed 95% CL cross section ratios for the combined CDF and D0 analyses, and the expected 95% CL ratios for the CDF and D0 experiments alone [34]. **Right:** The $\Delta\chi^2$ curve derived from high- Q^2 precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs boson mass, assuming the SM to be the correct theory of nature [35].

2.2.4 Electroweak Precision Measurements

An indirect measurement of m_H within the SM framework is possible using the precision measurements of the fundamental parameters, e.g., m_Z , Γ_{Z^0} , m_{W^\pm} . Such measurements have been performed by several experiments and a global fit to these electroweak observables with the Higgs boson mass as a free parameter sets limits on m_H [36, 35].

Figure 2-2 (right) shows the $\Delta\chi^2$ curve, derived from high-Q2 precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs boson mass, assuming the SM to be the correct theory of nature. The preferred value for its mass, corresponding to the minimum of the curve, is at 76 GeV, with an experimental uncertainty of +33 and -24 GeV (at 68% confidence level derived from $\Delta\chi^2 = 1$ for the black line, thus not taking the theoretical uncertainty shown as the blue band into account). This result is only affected a little by the low-Q2 results such as the NuTeV measurement. (While this is not a proof that the Standard-Model Higgs boson actually exists, it does serve as a guideline in what mass range to look for it. The precision electroweak measurements tell us that the mass of the Standard-Model Higgs boson is lower than about 144 GeV (one-sided 95% confidence level upper limit derived from $\Delta\chi^2 = 2.7$ for the blue band, thus including both the experimental and the theoretical uncertainty). This limit increases to 182 GeV when including the LEP-2 direct search limit of 114 GeV shown in yellow.)

2.3 The Standard Model Higgs Boson at LHC

The Higgs coupling depends on the mass of the particle. So, in general Higgs boson production comes through production of the heavy particles and it decays to heavy particles as well.

2.3.1 Production

The main production channels for the SM Higgs boson are (diagrams are shown at Fig. 2-3 (left)):

- a gluon fusion via top loop
- b vector boson (W/Z) fusion
- c W/Z associated production
- d $t\bar{t}$ fusion

The production cross section at LHC as a function of the Higgs boson mass is shown in Fig. 2-3.

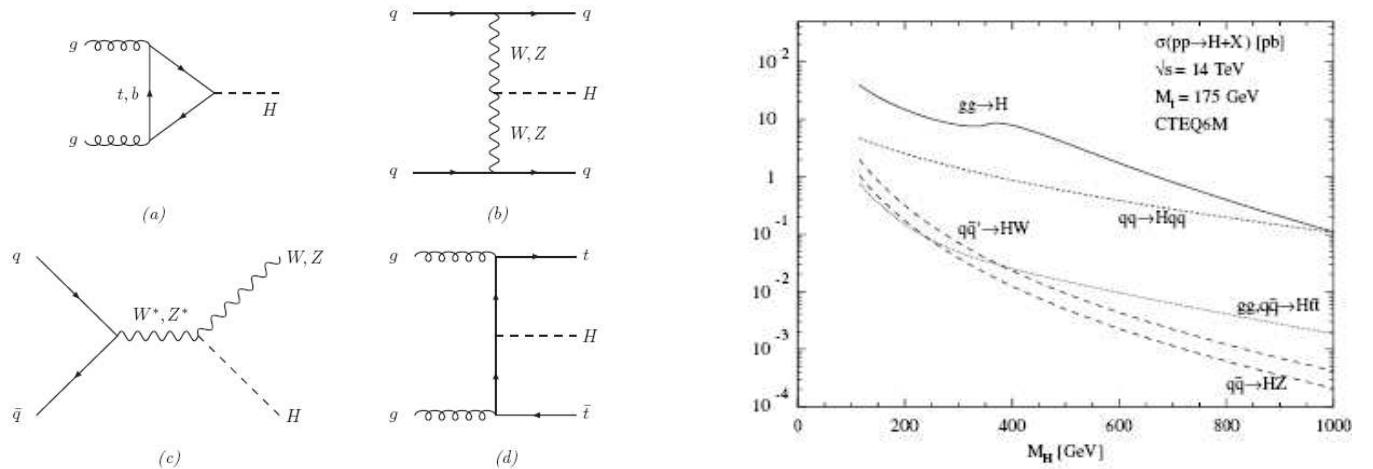


Figure 2-3: **Left:** Typical diagrams for all relevant Higgs boson production mechanisms at leading order at the LHC.: (a) gluon fusion, (b) vector boson fusion, (c) Higgs-strahlung, (d) Higgs bremsstrahlung off top quarks. **Right:** Higgs production cross sections at the LHC for the various production mechanisms as a function of the Higgs mass. The full QCD-corrected results for the gluon fusion $gg \rightarrow H$, vector-boson fusion $qq \rightarrow VVqq \rightarrow Hqq$, vector-boson bremsstrahlung $q\bar{q} \rightarrow V^* \rightarrow HV$ and associated production $gg, q\bar{q} \rightarrow Ht\bar{t}$ are shown.

Due to the large size of the top Yukawa couplings and the gluon densities, the gluon fusion, $pp \rightarrow gg \rightarrow H$, is the dominant Higgs boson production mechanism for the whole Higgs mass range.

In our Higgs boson search analysis, the NLO cross sections and branching ratios for the Higgs boson calculated with the programs HDECAY [37], HIGLU [38] are used, as well as the NLO cross sections for the background processes, when available². QCD NLO

² The same is true for results in section 2.4

corrections known for long time, they increase Higgs boson production cross section by $\sim 50 - 100\%$, NNLO QCD corrections are of the order of 20%.

2.3.2 Decays

The Higgs decay modes can be divided into two different mass ranges. For $M_H \leq 135\text{GeV}/c^2$ the Higgs boson mainly decays into $b\bar{b}$ and $\tau^+\tau^-$ pairs with branching ratios of about 85% and 8% respectively (see Fig. 2-4, right plot). Both $b\bar{b}$ and $\tau^+\tau^-$ Higgs boson decay modes are (almost) impossible for discovery at hadron collider due to overwhelming background level.

The decay modes into $c\bar{c}$ and gluon pairs, with the latter mediated by top and bottom quark loops, accumulate a branching ratio of up to about 10%, but do not play a relevant role at the LHC (they are even harder to discover than $b\bar{b}$ and $\tau^+\tau^-$ modes due to huge QCD background level).

The most important Higgs decays in this mass range at the LHC is the decay into photon pairs, which is mediated by W , top and bottom quark loops. It reaches a branching fraction of up to 2×10^{-3} . Its importance is in clean environment and controllable background, which makes it an important discovery channel for small Higgs boson masses.

For Higgs masses above $135\text{ GeV}/c^2$ the main decay modes are those into WW and ZZ pairs, where one of the vector bosons is off-shell below the corresponding kinematical threshold. These are very important discovery channels, in particular $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ is called “golden mode” for very clean (4 muons) final state, which provides low/medium integrated luminosity level for discovery of the Higgs boson for all possible Higgs boson masses in consideration ($\sim 115 - 600\text{GeV}/c^2$).

WW/ZZ decay modes dominate over the decay into $t\bar{t}$ pairs, the branching ratio of which does not exceed $\sim 20\%$ as can be seen from Fig. 2-4 (right plot).

The total decay width of the Higgs boson, shown in Fig. 2-4 (left plot), does not exceed about $1\text{ GeV}/c^2$ below the WW threshold. For very large Higgs masses the total

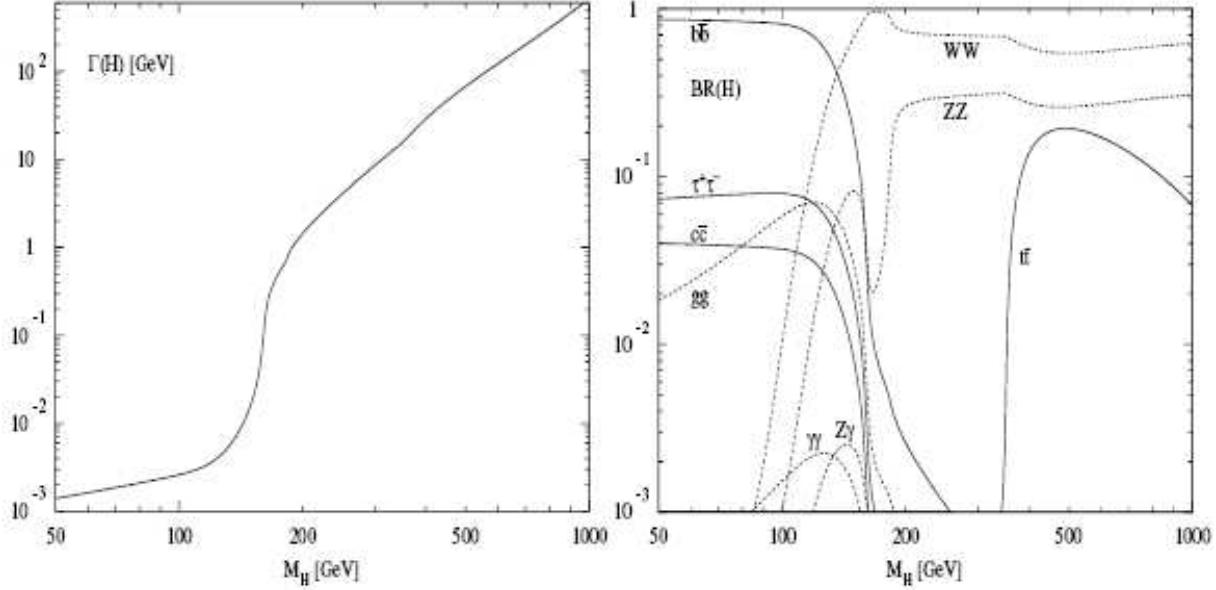


Figure 2-4: Left plot: total decay width (in GeV/c^2) of the SM Higgs boson as a function of its mass. Right plot: Branching ratios of the dominant decay modes of the SM Higgs particle. All relevant higher-order corrections are taken into account.

decay width grows up to the order of the Higgs mass itself so that the interpretation of the Higgs boson as a resonance becomes questionable.

2.4 The Standard Model Higgs Boson Searches at CMS Summary: Discovery Reach

Figure 2-5 (left) [8] shows the integrated luminosity needed for the 5σ discovery of the inclusive Higgs boson production $\text{pp} \rightarrow H + X$ with the Higgs boson decay modes $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$, and $H \rightarrow WW \rightarrow 2\ell 2\nu$ – three front runners among all possible decay channels. $H \rightarrow \gamma\gamma$ dominates discovery reach up to $m_H \sim 130\text{GeV}/c^2$, similarly it's $H \rightarrow WW \rightarrow 2\ell 2\nu$ dominates for $\sim 150 - 180$ GeV and $H \rightarrow ZZ \rightarrow 4\ell$ – in the rest of the possible Higgs boson masses.

Figure 2-5 (right) [8] shows the signal significance as a function of the Higgs boson mass for 30 fb^{-1} of the integrated luminosity for the different Higgs boson production and decay channels.

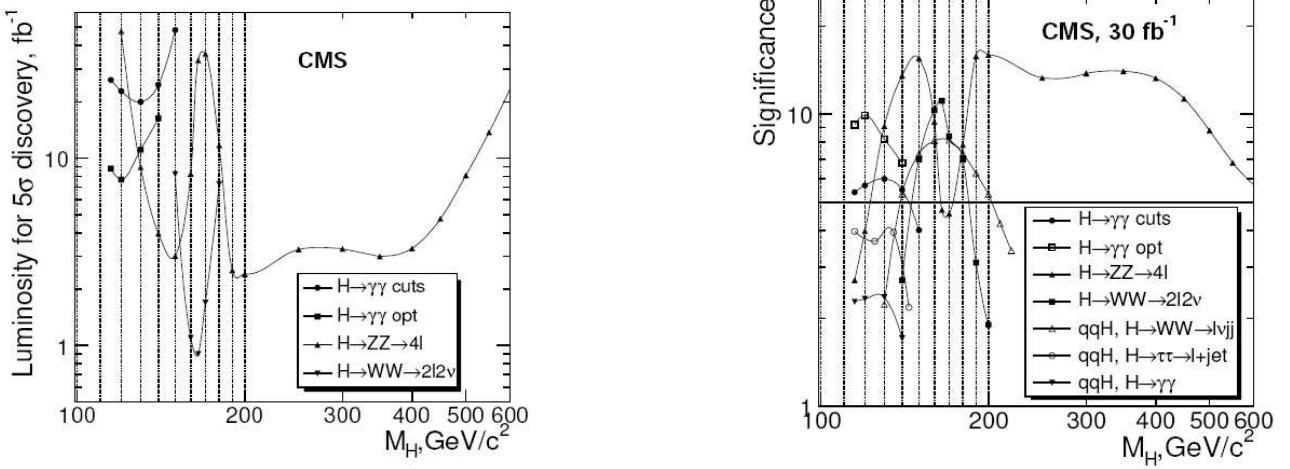


Figure 2-5: **Left:** The integrated luminosity needed for the 5σ discovery of the inclusive Higgs boson production $pp \rightarrow H + X$ with the Higgs boson decay modes $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ \rightarrow 4\ell$, and $H \rightarrow WW \rightarrow 2\ell 2\nu$ [8]. **Right:** The signal significance as a function of the Higgs boson mass for 30 fb^{-1} of the integrated luminosity for the different Higgs boson production and decay channels [8].

2.5 Issues in the Standard Model

In spite of the experimental success of the SM, it is believed that it does not describe Nature completely, but it is a low energy effective theory of a more fundamental theory. Following is an overview of some of the SM obstacles.

There is experimental evidence that the neutrinos are massive, contrary to the assumptions of the SM [39]. During the past years several neutrino experiments like the SuperKamiokande, K2K, SNO and Kamland [40, 41, 42, 43] have established the presence of neutrino oscillations. This is a sign of the neutrino masses: only massive particles have a time evolution and therefore can oscillate if mass differences between the various neutrino mass eigenstates exist.

The second obstacle comes from the creation of the Universe and the cosmological precision measurements. The lack of the antimatter in the visible Universe and the measured ratio $n_\gamma/n_b \approx 10^9$ [44] places a lower bound on the amount of CP violation, which is one of the three requirements for the creation of the matter-antimatter asymmetry [45]. The SM incorporates CP violation only by the CKM mechanism

[46]. The measured CP violation in the SM is smaller by at least eight orders of magnitude than the one needed to generate the cosmologically observed matter-antimatter asymmetry.

Third, another condition formulated by Andrei Sakharov requires baryon number violation. Currently, there is no experimental evidence of particle interactions where the conservation of baryon number is broken perturbatively: this would appear to suggest that all observed particle reactions have equal baryon number before and after.

Another observation that cannot be explained within the SM is the requirement for dark matter [47, 48]. According to the contemporary measurements, only 5% of the amount of the total energy is stored in an ordinary matter as known by the SM. For the remaining 95% of the energy of the Universe there is no explanation in the SM.

Furthermore, if the SM is valid up to an energy scale Λ , then the size of divergent loop contributions is $\Delta m_H^2 \sim \Lambda^2$. On the other hand loop corrections to the fermion masses are only subject to the logarithmic divergences, so that the overall correction is of the scale of mass itself and no fine-tuning problem emerges. For the Higgs boson this means that if the SM is valid up to the Planck scale of $\Lambda_P = 10^{19}$ GeV, then the natural scale of the Higgs boson mass is Λ_P while all other particles have natural mass scales below VEV v . This is the so-called hierarchy problem, which refers to the extremely large splitting of the weak scale and natural cut-off scale, the Planck scale. In order to achieve the necessary Higgs mass range of $m_H < 1$ TeV, an unnatural fine-tuning with the relative precision of $m_H/\Lambda_P > 10^{-16}$ has to be applied. This fine-tuning is not explained in the context of the SM but it can be solved by extensions of the SM.

The SM leaves unexplained why the strong and the electroweak gauge structure is $SU(3)_C \times SU(2)_L \times U(1)_Y$ with different gauge couplings and fermionic quantum numbers whose values are not predicted by the model. There have been many efforts to unify the gauge groups and to have only one coupling at the energy scale of Grand Unification [49, 50, 51].

There are more questions: why neutrino masses are so small; why there is no CP violation in QCD, etc.

2.6 Supersymmetry Higgs

Some of the problems mentioned above can be solved by introducing a new symmetry which relates bosons to fermions. Under such a symmetry, so called Supersymmetry (SUSY), every SM fermion has a bosonic partner, and every SM boson has a fermionic partner.

An additional Higgs doublet together with its supersymmetric partner must be introduced into the SUSY extension of the SM.

Generally broken SUSY has a huge parameter space and therefore has a very limited predictive power. An example of much more constrained version, with less free parameters, is the Minimal Supersymmetric Standard Model (MSSM).

In MSSM the eight real scalar degrees of freedom (that correspond to two complex Higgs doublets) are three Goldstone bosons absorbed by the W^\pm and Z, and five physical Higgs bosons: neutral CP-even h and H, neutral CP-odd A, and two charged H^\pm . The features of the Higgs particles greatly depend on choice of parameters for particular SUSY models.

MSSM predicts that mass of the lightest Higgs should have mass ≤ 135 GeV [52].

CHAPTER 3

LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) [53] is a 27 km circumference particle accelerator spanning the Swiss-French boarder. It is a project carried out by CERN (European Organization of Nuclear Research) [54], which is a large international collaboration consisting of many countries from all over the world. The LHC is currently under construction and has the first beams scheduled in 2008. When completed it will provide proton-proton collisions at the center-of-mass energy 14 TeV as well as heavy ion collisions.

The LHC consists of a number of accelerator units (Fig. 3-1 [53]). Two linear accelerators, LINAC2 and LINAC3, will be used for the initial acceleration of protons and lead ions respectively. The PS (Proton Synchrotron) will be used to provide a low energy beam (25 GeV) with the final bunch structure. The beams are pre-accelerated using the SPS (Super PS), and then are injected into the LHC near IP2/IP8 for the beam that circulates clock wise/counter-clock wise, where particles will be accelerated from 450 GeV to the nominal energy of 7 TeV (for proton beam).

Inside the LHC accelerator, the particles circulate in opposite directions in two separate beam pipes. The diameter of the beam is reduced by focusing magnets before the beam enters every experimental interaction point to achieve high luminosity collisions. The two beam pipes are surrounded by shells of superconducting coils creating the magnetic field which guides the beams to follow a circular path. Particles inside the beam are grouped into approximately cylindrical bunches with a diameter and a length of approximately 16μ and 8 cm at the interaction point respectively. The LHC filling scheme for proton beam will have a bunch separation of 25 ns.

The LHC will undergo a series of accelerator/detector commissioning stages and pilot runs at the beginning of its operation in 2008. Regular physics runs will start with low luminosity runs, planned luminosity is up to $2 * 10^{33} cm^{-1}s^{-1}$, with up to about $10fb^{-1}$ of

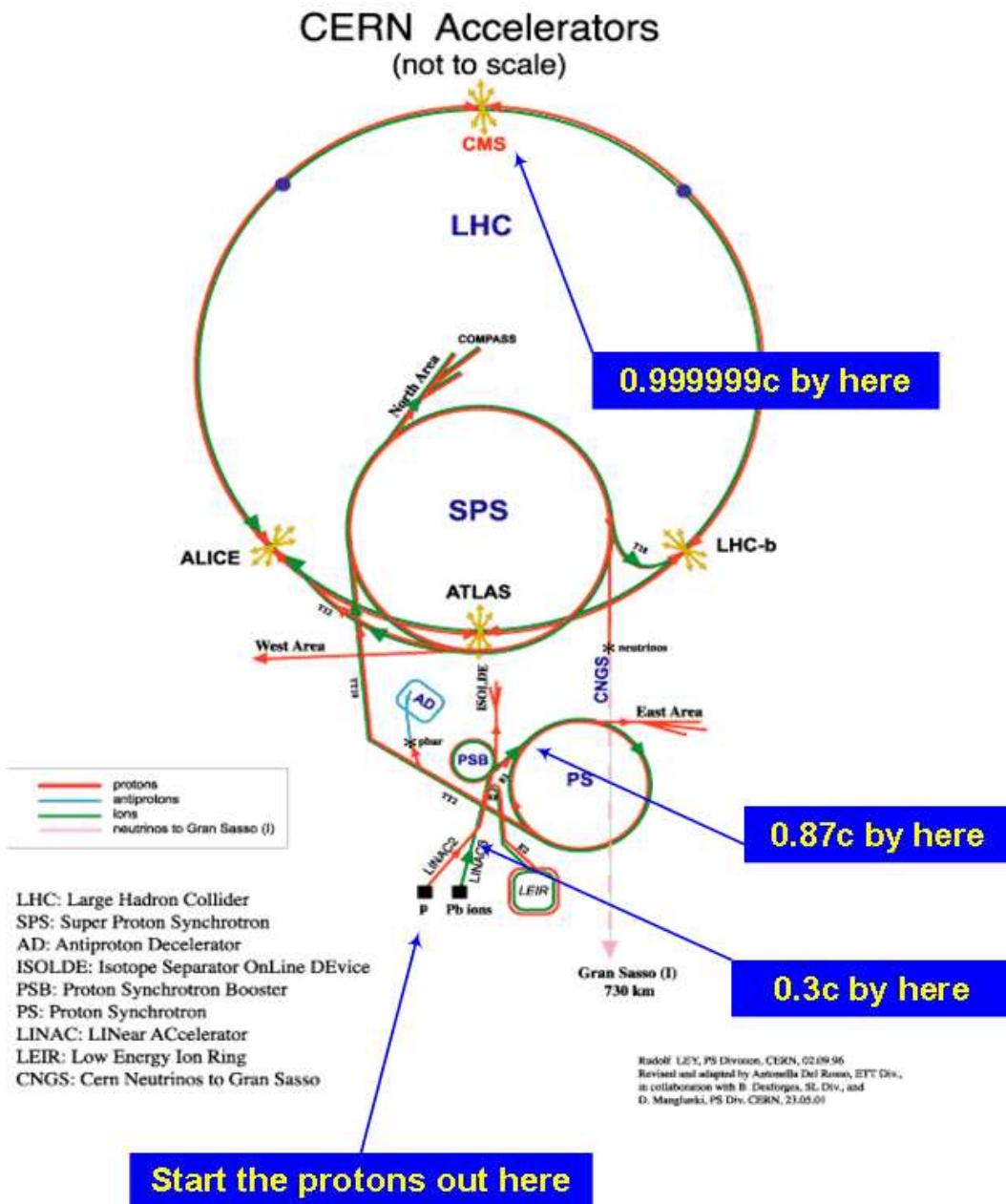


Figure 3-1: Accelerator complex at CERN.

integrated luminosity per year and will go up to $\sim 10^{34} cm^{-1}s^{-1}$ with about $100 fb^{-1}$ of integrated luminosity per year.

There are four main experiments at the LHC: two general purpose detectors, CMS (Compact Muon Solenoid) [55] and ATLAS [56] (A Toroidal Lhc ApparatuS), and two with dedicated detectors, ALICE [57] (A Large Ion Collider Experiment) and LHCb (the Large Hadron Collider Beauty experiment) [58] which will study heavy ion and B-physics respectively plus relatively small TOTEM experiment [59] for diffraction/forward physics situated next to CMS, which also would allow to have more precise luminosity measurements. The CMS detector is situated approximately 100 m underground at IP5 (Interaction Point in 5th LHC octant), and will be taking data from both the p-p and Pb-Pb collision runs.

CHAPTER 4

COMPACT MUON SOLENOID (CMS) DETECTOR

We give a short introduction into the CMS and its subsystems the following chapter (see e.g. [7, 8, 60] for more details).

4.1 Introduction

The total proton-proton cross-section $\sqrt{s} = 14$ TeV is roughly 100 mb. At design luminosity the general-purpose detectors will therefore observe an event rate of approximately 10^9 inelastic events/s. This leads to a number of experimental challenges. The online event selection process (“trigger”) must reduce the approximately 1 billion interactions/s to no more than about 100 events/s for storage and subsequent analysis. The short time between bunch crossings, 25 ns, has major implications for the design of the readout and trigger systems.

At the design luminosity, a mean of about 20 inelastic (hard-core scattering) collisions will be superimposed on the event of interest. This implies that around 1000 charged particles will emerge from the interaction region every 25 ns. The products of an interaction under study may be confused with those from other interactions in the same bunch crossing. This problem clearly becomes more severe when the response time of a detector element and its electronic signal is longer than 25 ns. The effect of this pile-up can be reduced by using high-granularity detectors with good time resolution, resulting in low occupancy. This requires a large number of detector channels. The resulting millions of detector electronic channels require very good synchronization.

The large flux of particles coming from the interaction region leads to high radiation levels, requiring radiation-hard detectors and front-end electronics.

The detector requirements for CMS [7, 8] to meet the goals of the LHC physics program can be summarized as follows:

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx 1\%$ at 100 GeV/c), and the ability to determine unambiguously the charge of muons with $p < 1$ TeV/c.

- Good charged particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ 's and b-jets, requiring pixel detectors close to the interaction region.
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution (1% at at 100 GeV/c), wide geometric coverage, measurement of the direction of photons and/or correct localization of the primary interaction vertex, π^0 rejection and efficient photon and lepton isolation at high luminosities.
- Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam direction counter clockwise looking from above. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. Pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$. Thus, the momentum and energy measured transverse to the beam direction, denoted by P_T and E_T , respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss} .

4.2 General Description

An important aspect driving the detector design and layout is the choice of the magnetic field configuration for the measurement of the momentum of muons. Large bending power is needed to measure precisely the momentum of charged particles. This forces a choice of superconducting technology for the magnets. The design configuration chosen by CMS [61] is discussed below.

At the heart of CMS sits a 13-m-long, 5.9 m inner diameter, 4 T superconducting solenoid (Fig. 4-1). In order to achieve good momentum resolution within a compact spectrometer without making stringent demands on muon-chamber resolution and alignment, a high magnetic field was chosen. The return field is large enough to saturate

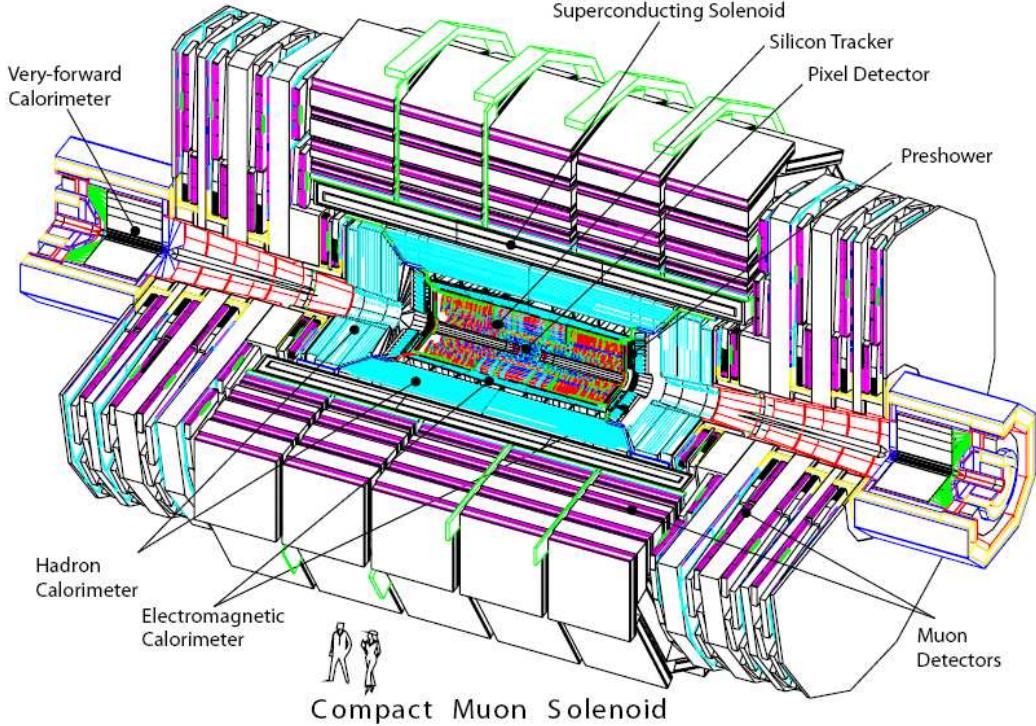


Figure 4-1: An exploded view of the CMS detector.

1.5 m of iron, allowing 4 muon “stations” to be integrated to ensure robustness and full geometric coverage. Each muon station consists of several layers of aluminium drift tubes (DT) in the barrel region and cathode strip chambers (CSCs) in the endcap region, complemented by resistive plate chambers (RPCs).

The bore of the magnet coil is also large enough to accommodate the inner tracker and the calorimetry inside. The tracking volume is given by a cylinder of length 5.8 m and diameter 2.6 m. In order to deal with high track multiplicities, CMS employs 10 layers of silicon microstrip detectors, which provide the required granularity and precision. In addition, 3 layers of silicon pixel detectors are placed close to the interaction region to improve the measurement of the impact parameter of charged-particle tracks, as well as the position of secondary vertices. The EM calorimeter (ECAL) uses lead tungstate ($PbWO_4$) crystals with coverage in pseudorapidity up to $|\eta| < 3.0$. The scintillation light is detected by silicon avalanche photodiodes (APDs) in the barrel region and vacuum phototriodes (VPTs) in the endcap region. A preshower system is

installed in front of the endcap ECAL for π^0 rejection. The ECAL is surrounded by a brass/scintillator sampling hadron calorimeter with coverage up to $|\eta| < 3.0$. The scintillation light is converted by wavelength-shifting (WLS) fibres embedded in the scintillator tiles and channeled to photodetectors via clear fibres. This light is detected by novel photodetectors (hybrid photodiodes, or HPDs) that can provide gain and operate in high axial magnetic fields. This central calorimetry is complemented by a “tailcatcher” in the barrel region-ensuring that hadronic showers are sampled with nearly 11 hadronic interaction lengths. Coverage up to a pseudorapidity of 5.0 is provided by an iron/quartz-fibre calorimeter. The Cerenkov light emitted in the quartz fibres is detected by photomultipliers. The forward calorimeters ensure full geometric coverage for the measurement of the transverse energy in the event. An even higher forward coverage is obtained with additional dedicated calorimeters (not shown in the figure) and with TOTEM.

The overall dimensions of the CMS detector are a length of 21.6 m, a diameter of 14.6 m and a total weight of 12 500 tons. The thickness of the detector in radiation lengths is greater than $25 X_0$ for the ECAL, and the thickness in interaction lengths varies from $7\text{--}11\lambda_I$ for HCAL depending on η .

4.3 The Superconducting Magnet

The compact design of CMS requires a very strong magnetic field in order to induce enough bending of the charged particle trajectories so that the momentum of the particles can be measured up to the highest momentum expected at the LHC. The basic design goal was to be able to reconstruct 1 TeV muons with $\sim 10\%$ P_T resolution which scales with $1/B$ where B is the strength of the magnetic field. In the 4 Tesla field, trajectories of charged particles with $pT > 0.7$ GeV reach the ECAL front surface (in the absence of tracker material), and muons with $pT > 4$ GeV extend through the muon chambers.

The superconducting magnet for CMS has been designed to reach a 4 T field in a free bore of 6m diameter and 12.5m length with a stored energy of 2.6 GJ at full current.

The flux is returned through a 10000 t yoke comprising of 5 wheels and 2 end caps composed of three disks each [62]. The distinctive feature of the 220-t cold mass is the four-layer winding made from a stabilized reinforced NbTi conductor. The ratio between stored energy and cold mass is critically high (11.6KJ/ kg), causing a large mechanical deformation (0.15%) during energization, well beyond the values of previous solenoidal detector magnets.

4.4 Inner Tracking System

The inner tracking system of CMS is designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from the LHC collisions as well as a precise reconstruction of secondary vertices. It surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. The CMS solenoid provides a homogeneous magnetic field of 4 Tesla over the full volume of the tracker. At the LHC design luminosity of $10^{34} cm^{-2}s^{-1}$ there will be on average about 1000 particles from more than 20 overlapping proton proton interactions traversing the tracker for each bunch crossing, i.e. every 25 ns. Therefore a detector technology featuring high granularity and read-out speed is required, such that the trajectories can be identified reliably and attributed to the correct bunch crossing. However, these features imply a high power density of the on detector electronics which in turn requires efficient cooling. This is in direct conflict with the aim to reduce the amount of material in the tracker to a minimum in order to limit multiple scattering, bremsstrahlung, photon conversion and nuclear interactions, and a compromise had to be found in this respect. The intense particle flux will also cause severe radiation damage to the tracking system. The main challenge in the design of the tracking system was to develop detector components able to operate in this harsh environment for an expected lifetime of 10 years. These requirements on granularity, speed and radiation hardness lead to a tracker design entirely based on silicon detector technology. The CMS tracker is composed of a pixel detector with three barrel layers at radii between 4.4 cm and 10.2 cm and a silicon strip tracker with 10 barrel detection

layers extending outwards to a radius of 1.1 m. Each system is completed by end caps which consist of 2 disks in the pixel detector and 9 disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudo-rapidity of $|\eta| < 2.5$. With about 200 m^2 of active silicon area the CMS tracker is the largest silicon tracker ever built [63].

For single charged particle of transverse momenta of 1, 10 and 100 GeV the expected resolution of transverse momentum depends on pseudorapidity [7, 8] and high momentum tracks (100 GeV) the transverse momentum resolution is around 1 - 2% up to $|\eta| \approx 1.6$, beyond which it degrades due to the reduced lever arm. At a transverse momentum of 100 GeV multiple scattering in the tracker material accounts for 20 to 30% of the transverse momentum resolution while at lower momentum it is dominated by multiple scattering. The transverse impact parameter resolution reaches $10\mu\text{m}$ for high P_T tracks, dominated by the resolution of the first pixel hit, while at lower momentum it is degraded by multiple scattering (similarly for the longitudinal impact parameter). For muons, the expected track reconstruction efficiency (a function of pseudo-rapidity) is about 99% over most of the acceptance. For $|\eta| \approx 0$ the efficiency decreases slightly due to gaps between the ladders of the pixel detector at $z \approx 0$. At high the efficiency drop is mainly due to the reduced coverage by the pixel forward disks. For pions and hadrons in general the efficiency is lower because of interactions with the material in the tracker.

4.5 Electromagnetic Calorimeter (ECAL)

The Electromagnetic Calorimeter of CMS is a hermetic homogeneous calorimeter made of 61,200 lead tungstate (PbWO_4) crystals mounted in the central “barrel” part, closed by 7,324 crystals in each of the two end-caps. A preshower detector is placed in front of the end-caps crystals. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) in the end-caps. The use of high density crystals has allowed the design of a calorimeter which is fast, has fine granularity and is radiation resistant, all important characteristics in the LHC environment. One of the

driving criteria in the design was the capability to detect the decay to two photons of the Higgs boson. This capability is enhanced by the good energy resolution provided by a homogeneous crystal calorimeter.

A typical energy resolution for central impact when the energy was reconstructed in a 3×3 matrix of crystals was found to be:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E/(GeV)}}\right)^2 + \left(\frac{125(MeV)}{E}\right)^2 + (0.30\%)^2 \quad (4-1)$$

The principal aim of the CMS Preshower detector is to identify neutral pions in the endcaps within a fiducial region $1.653 < |\eta| < 2.6$. It also helps the identification of electrons against minimum ionizing particles, and improves the position determination of electrons and photons with high granularity.

4.6 Hadron Calorimeter (HCAL)

The hadron calorimeters are particularly important for the measurement of hadron jets and neutrinos or exotic particles resulting in apparent missing transverse energy [61].

Hadronic showers have a large lateral shower size, so the degree of granularity provided by the ECAL is not required for the HCAL. However, moderate transverse granularity, hermeticity and wide rapidity coverage are essential in measuring the total missing ET of an event. The HCAL consists of three main parts: Hadron Barrel (HB) and Hadron Endcap (HE), which extends to $\eta = 3.0$, and Hadron Forward (HF) calorimeters, which are located around the beam pipe outside the muon system at $|z| = 10.9$ m to complete the coverage up to $\eta = 5.3$. Due to the restricted space available between the ECAL and the magnetic coil, the HB is only 1 m in thickness whereas the HE is 1.8 m thick; additional scintillation layers (Hadron Outer calorimeter, HO) are installed just outside the magnet, using the coil as an absorber, to provide a combined depth of at least 11 absorption lengths.

For gauging the performance of the HCAL, it is usual to look at the jet energy resolution and the missing transverse energy resolution. The granularity of the sampling in

the 3 parts of the HCAL has been chosen such that the jet energy resolution, as a function of E_T , is similar in all 3 parts. The resolution of the missing transverse energy (E_T^{miss}) in QCD dijet events with pile-up is given by $\sigma(E_T^{miss}) \approx 1.0\sqrt{\Sigma E_T}$ if energy clustering corrections are not made, while the average E_T^{miss} is given by $\langle E_T^{miss} \rangle \approx 1.25\sqrt{\Sigma E_T}$ [7].

4.7 Muon System

The muon system is a tracking device in the outermost region of the CMS detector, where only muons and neutrinos would have passed through the calorimeters without depositing a large fraction of their energy. The challenge of the muon system is to provide fast recognition, efficient reconstruction and online background suppression, which are required for triggering purposes and also, good measurement for higher P_T muons. There are four layers of muon stations in the barrel and the endcaps interleaved with the iron return yokes, each providing track segments reconstructed from a few distributed hits. These will later be combined with information from the inner silicon tracking system to form complete muon tracks.

The muon system employs a combination of three different detectors: Drift Tubes (DT) in the barrel and Cathode Strip Chambers (CSC) in the endcaps, both assisted by Resistive Parallel plate Chambers (RPC).

The DT and CSC detectors are used to obtain a precise position measurement; background rejection is ensured by the highly redundant design, where each of the four stations contains 12 DT layers for the barrel and 6 CSC layers in the endcap. The RPCs, which have a very fast response, provide a time resolution of the order of 1 ns, and are dedicated for the triggers.

4.7.1 Barrel Muon System

The CMS barrel muon detector is made of four stations forming concentric cylinders around the beam line: three of them consist of 60 drift chambers each, the fourth, of 70. The total number of sensitive wires is about 195,000. The choice of a drift chamber as the

tracking detector for the barrel muon system was possible due to the low expected rate and the relatively low intensity of the local magnetic field.

Within a single station, the DT layers provide a final spacial resolution of $100\ \mu\text{m}$.

4.7.2 Endcap Muon System

The endcap muon system is described in details in the following chapter.

A final spacial resolution within a single station for CSCs is $80\text{-}150\ \mu\text{m}$ depending on the station.

The CSCs are chosen for the endcaps for their performance in the presence of a high intensity varying magnetic field, and their ability to handle high particle rates.

4.7.3 Resistive Plate Chambers

Resistive Plate Chambers (RPC) are gaseous parallel-plate detectors that combine good spatial resolution with a time resolution comparable to that of scintillators. An RPC is capable of tagging the time of an ionizing event in much shorter time than the 25 ns between two consecutive bunch crossings (BX). A total of six layers of RPC chambers will be embedded in the barrel iron yoke, two located in each of the first and second muon stations and one in each of the two last stations.

4.8 Trigger

The CMS muon trigger and CSC muon trigger in particular will be used in the following chapters. We will give a short overview here, more details are available in [64, 65, 7, 8, 60].

4.8.1 Muon Trigger

All three muon systems (DT, CSC, RPC) take part in the trigger. The DT chambers provide local trigger information: ϕ and η -projections for track segments. The CSCs deliver 3-dimensional track segments. All chamber types also identify the bunch crossing at which event took place. The Regional Muon Trigger consists of the DT and CSC Track Finders, which join segments to complete tracks and assign physical parameters to them. In addition, the RPC trigger chambers provide their own track candidates based on

regional hit patterns. The Global Muon Trigger then combines the information from the three subdetectors, achieving an improved momentum resolution and efficiency compared to the stand-alone systems. The initial rapidity coverage of the muon trigger is $|\eta| \leq 2.1$ at the startup of LHC. The design coverage is $|\eta| \leq 2.4$.

4.8.1.1 CSC local trigger

Muon track segments, also called Local Charged Tracks (LCT), consisting of positions, angles and bunch crossing information are first determined separately in the nearly orthogonal anode and cathode views. The trigger electronics determines the centre of gravity of the charge with a resolution of half a strip width, between 1.5 and 8 mm, depending on the radius. By demanding that at least four layers are hit, the position of a muon can be determined with a resolution of 0.15 strip widths. The best two LCTs of each chamber are transmitted to the regional CSC trigger, the CSC Track Finder, which joins segments to complete tracks. The CSC trigger hardware consist of several different electronics boards, in particular, ALCT (for anode LCT) and CLCT (for cathode LCT). Trigger motherboard (TMB) looks for coincidence between two (ALCT, CLCT), if found, information transmitted to Muon Port Cards (MPC). Up to three best candidates then sent from MPC to CSC track finder. If then global trigger provides accept signal, anode and cathode information is passed to Data Acquisition system (DAQ) and then stored for further offline reconstruction and analysis.

CHAPTER 5

CATHODE STRIP CHAMBERS (CSC)

We give a short introduction into the CSC in the following chapter (see e.g. [65, 7, 8, 60] for more details).

5.1 Introduction

At the time of the LHC start-up, the CMS Endcap Muon System [65, 7, 8] will consist of 468 Cathode Strip Chambers (CSCs) arranged in groups as follows: 72 ME1/1, 72 ME1/2, 72 ME1/3, 36 ME2/1, 72 ME2/2, 36 ME3/1, 72 ME3/2, 36 ME4/1 (Figs. 5-1 and 5-2). The de-scoped 72 ME4/2 chambers will not be available during early years of CMS operation. The chambers are trapezoidal and cover either 10 or 20 degrees in ϕ ; all chambers, except for the ME1/3 ring, overlap and provide contiguous ϕ -coverage. A muon in a pseudorapidity range $1.2 < |\eta| < 2.4$ crosses 3 or 4 CSCs. In the endcap barrel overlapping range $0.9 < |\eta| < 1.2$, muons are detected by both the barrel Drift Tubes (DTs) and endcap CSCs. For pseudorapidity $|\eta| < 2.0$, muons are also detected by Resistive Plate Chambers (RPCs).

CSCs are multi-wire proportional chambers comprised of six anode wire planes interleaved between seven cathode panels (Fig. 5-3). Wires run azimuthally and define track's radial coordinate. Strips are milled on cathode panels and run lengthwise at constant $\Delta\phi$ width. Following the original CSC idea [66], we obtain the muon coordinate along wires (ϕ -coordinate in CMS coordinate system) by interpolating charges induced on strips (Fig. 5-3 (right)). The largest chambers, ME2/2 and ME3/2, are about $3.4 \times 1.5 \text{ m}^2$ in size. The overall area covered by the sensitive planes of all chambers is about 5000 m^2 , the total gas volume is $> 50 \text{ m}^3$, and the total number of wires is about 2 million. There are about 9,000 high voltage channels in the system. The number of cathode strip readout channels with 12-bit signal digitization is about 220,000, and the number of anode readout channels is about 180,000.

CSCs provide functions of a precision muon measurement and muon trigger in one device. They can operate at high rates and in large and non-uniform magnetic fields.

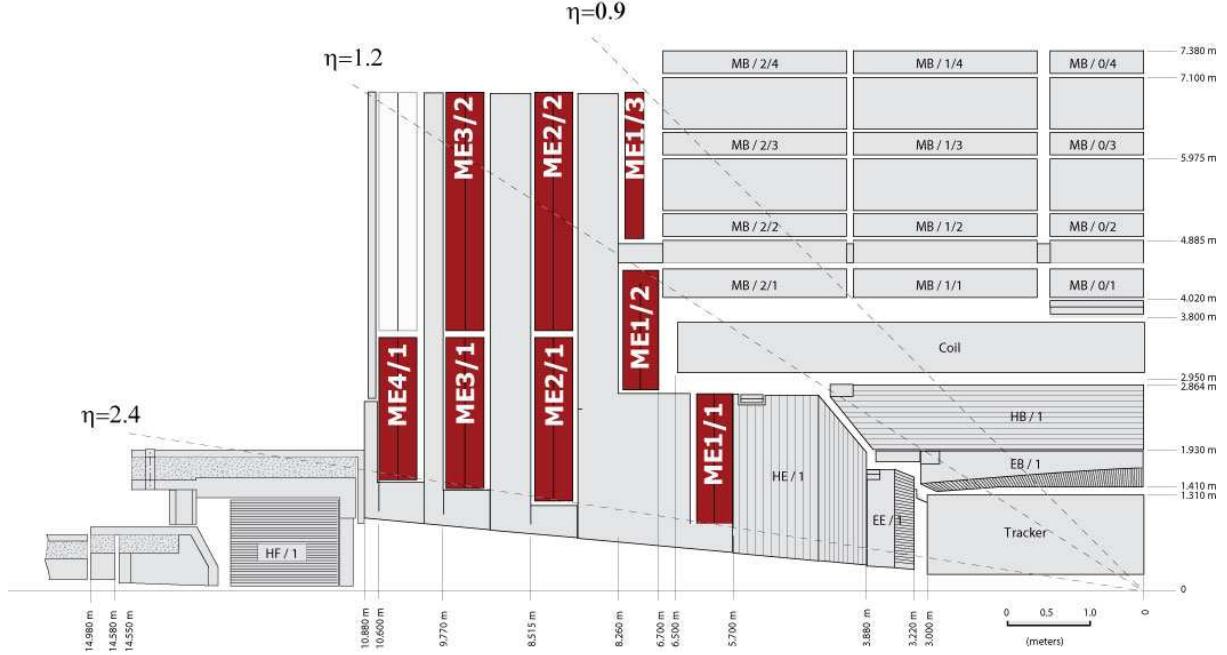


Figure 5-1: Quater-view of the CMS Detectors. Cathode Strip Chambers of the Endcap Muon System are highlighted.

They do not require tight gas, temperature or pressure control. Additionally, the radial fan-shaped strip pattern, natural for measurements in the endcap region, can be easily arranged on cathodes.

The requirements for the CMS cathode strip chamber performance are as follows:

- Reliable and low maintenance operation for at least 10 years at the full LHC luminosity, i.e. at estimated random hit rates up to 1 kHz/cm^2 .
- At least 99% efficiency per chamber for finding track stubs at the 1st Level trigger.
- At least 92% probability per chamber of identifying correct bunch crossings at the 1st Level trigger. With such an efficiency per chamber and 3-4 CSCs on a muon track path, a plain majority rule ensures that the reconstructed muons will be assigned to a correct bunch crossing number in more than 99% of cases.
- About 2 mm resolution in r_ϕ at the 1st Level trigger.

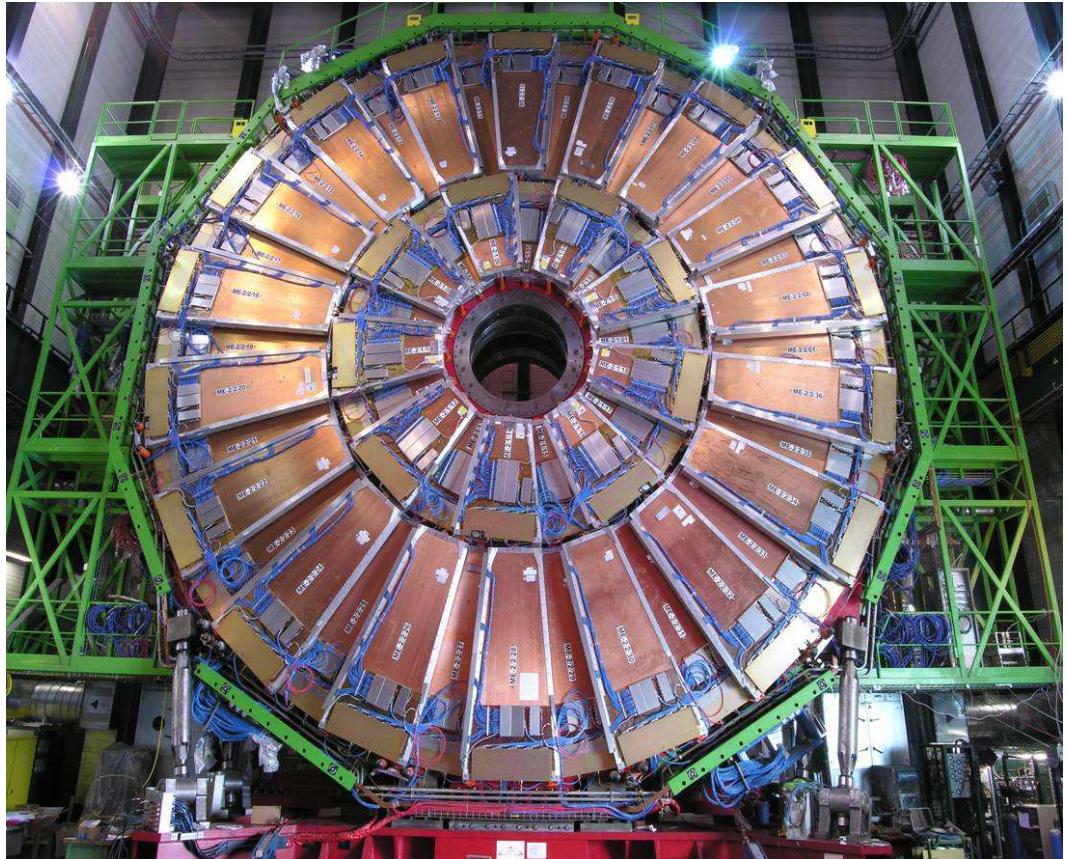


Figure 5-2: The ME2 cathode strip chambers. The outer ring consists of 36 ME2/2 chambers, each spanning 10° in ϕ ; while 18 20-degree ME2/1 chambers form the inner ring. The chambers overlap to provide contiguous coverage in ϕ .

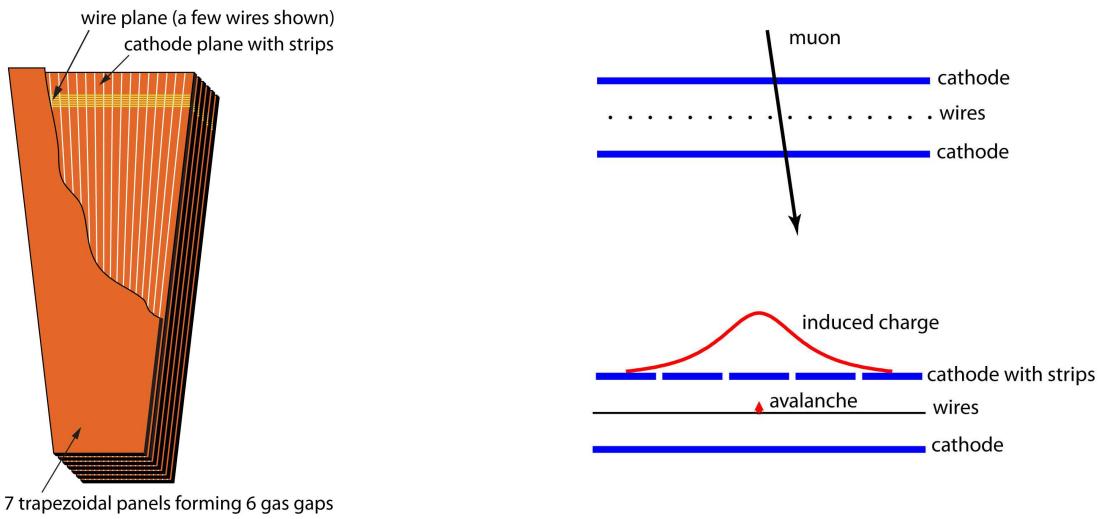


Figure 5-3: **Left:** A schematic view of a CMS cathode strip chamber made of seven trapezoidal panels. The panels form six gas gaps with planes of anode sensitive wires. The cutout in the top panel reveals anode wires and cathode strips. Only a few wires are shown to indicate their azimuthal direction. Strips of constant $\Delta\phi$ run lengthwise (radially). The largest CMS CSCs are 3.4 m long along the strip direction and up to 1.5 m wide along the wire direction. **Right:** A schematic view of a single gap illustrating the principle of a CSC operation. By interpolating charges induced on cathode strips by avalanche positive ions near a wire, one can obtain a precise localization of an avalanche along the wire direction.

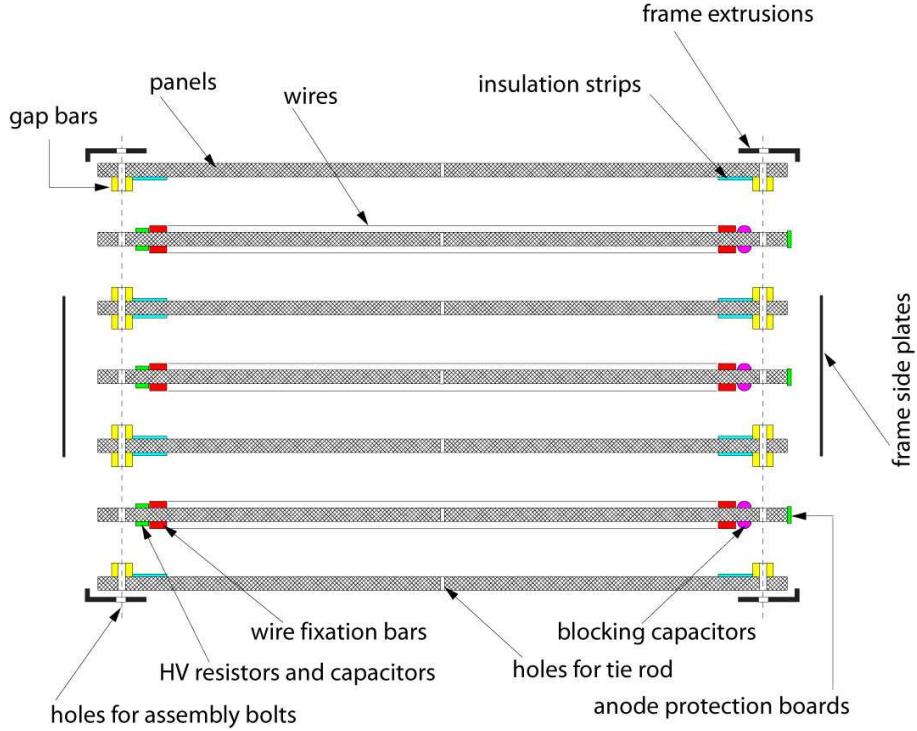


Figure 5-4: Mechanical design of the CMS Cathode Strip Chambers (exploded view).

- About $75 \mu\text{m}$ off-line spatial resolution in r_ϕ for ME1/1 and ME1/2 chambers and about $150 \mu\text{m}$ for all others.

5.2 Chamber Mechanical Design

The larger 396 chambers and the smallest 72 ME1/1 chambers have somewhat different mechanical designs. Below, we describe the design of the larger chambers on an example of ME2/2 and, then, at the end of this section, summarize the ME1/1-specific features that distinguish them from all other chambers.

The largest chambers, 144 ME2/2 and ME3/2, are 3.4 m long and 1.5 m wide (Fig. 5-4). Seven 16.2-mm thick trapezoidal panels are the basis for the mechanical structure. The panels are made of 12.7-mm thick polycarbonate honeycomb core with two 1.6-mm FR4 skins commercially glued on each side. FR4 is fire-retardant fiberglass/epoxy material widely used for printed circuit boards. The FR4 skins are copper-clad on their outer surfaces—the copper layers serve the role of cathodes.

FR4 cathode gap bars are glued to both sides of each other panel (panels 1, 3, 5, 7 in Fig. 5-4) so that when all panels are stacked together, these cathode bars define six gas gaps of 9.5 mm. To provide an additional support, there are four spacers placed between panels along the chamber centerline. When all seven panels are put together, the entire stack is tighten down with bolts along the chamber perimeter (through holes in the cathode gap bars) and in 4 points along the chamber centerline (through holes in the spacers). Such arrangement ensures that panels have no more than 60 cm of unsupported length. Measurements showed that most of the panels were flat within the required $\pm 300 \mu\text{m}$ on such spans. This specification arises from the desire to keep gas gain variations within a factor of two.

One side of six panels has a milled pattern of 80 strips. Strips, being radial, have a varying pitch from 8.4 mm at the bottom to 16.0 mm at the top. The gap between strips is about 0.5 mm. The precision of milling is better than $50 \mu\text{m}$ (rms). Milling is done with a cutter tilted at 45 degrees to make groove edges smoother (otherwise, sharp edges and burrs might provoke sparking and discharges).

Anode wires are wound on three so-called “anode” panels (panels 2, 4, 6 in Fig. 5-4); these panels do not have gap bars. A special winding machine was designed to wind wires directly on a panel rotating around its long axis at a speed of about 5 turns per minute. This allows for winding one panel (about 1000 wires on each side) in less than 4 hours. The wire spacing of about 3.2 mm is defined by combs, threaded rods running the full panel length and attached to the panel edges during winding. Gold-plated tungsten wires, 50 μm in diameter, are stretched at 250-gram tension (about 70% of the elastic limit) and run their full length up to 1.2 m without any intermediate supports. The electrostatic stability limit for the longest wires is above 6 kV (the nominal operational point is 3.6 kV). Based on the measurements during production, the wire tension non-uniformity does not exceed $\pm 10\%$, while wire spacing variations are less than $\pm 150 \mu\text{m}$. Wires found to fall out of these specs were replaced.

After winding, the wires are first glued and then soldered to anode bars of 4.75 mm height (half of the gas gap). The anode bars are made of copper-clad FR4 and carry an electric artwork. An automated soldering machine allowed for soldering at a speed of 3.5 s per joint. Groups of sixteen wires make one anode readout channel of about 5 cm width. High Voltage (HV) is distributed to the wire groups on one end and signals are read out on the other end via 1 nF blocking capacitors.

Each wire plane is sub-divided by spacer bars into 5 independent HV segments, which allows us to independently regulate or turn off HV on any of the five sections. In places where the spacer bars are inserted (and prior to their installation), eight wires are removed. Two gold-plated 200 μm guard wires inserted in place of the first and eighth removed thin wires to eliminated edge effects. The very first and the last wires in each plane are also thick. If the edge thin wires are left unguarded, the electric field on them would be much larger than for the rest of wires, which would provoke discharges. All in all, such plane segmentation, by virtue of introducing intermediate panel supports and individual HV control over smaller wire plane sections, makes the overall chamber performance very robust.

After stacking the panels and tightening bolts (with o-rings), continuous beads of sealant RTV are applied along the outer seams between panels and gap bars. The o-rings around bolts and the RTV seal make chambers hermetic. Should a need arise, chambers can be opened, serviced, and resealed. The gas enters into one of the outer gas gaps via an inlet in a cathode gap bar, then flows from one plane to another in zigzag manner via special holes in panels, and finally exits from the last gas gap via an outlet in a gap bar. The leak rate was measured during production and upon installation of chambers and was required to be less than 1% of the chamber volume per day at overpressure of 7.5 mbar (<2 cc/min for the largest chambers whose gas volume is about 200 liters).

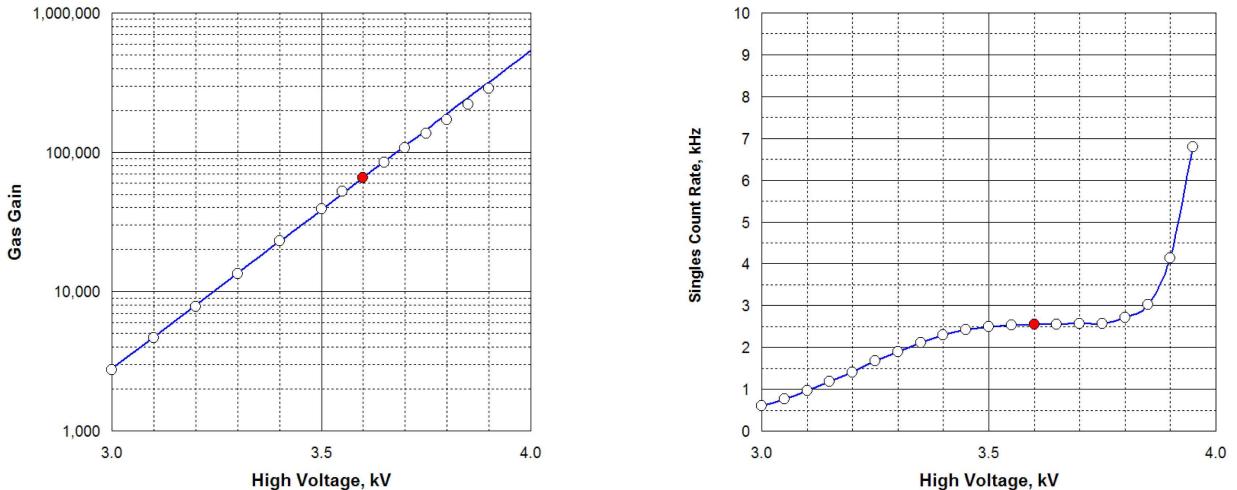


Figure 5-5: **Left:** CSC gas gain vs high voltage. **Right:** ME2/1 chamber singles rate vs high voltage (the overall sensitive area of all six planes in this chamber is $\sim 9.5 \text{ m}^2$).

Side plates made of 3.2-mm thick Al extrusions are attached along the chamber perimeter. They stiffen the chamber and interconnect top and bottom copper skins to form a complete Faraday cage.

The nominal gas mixture is $\text{Ar} + \text{CO}_2 + \text{CF}_4 = 40\% + 50\% + 10\%$. The CO_2 component is a non-flammable quencher needed to achieve large gas gains, while the main function of CF_4 is to prevent polymerization on wires. The detailed discussion of the gas optimization for CMS CSCs can be found elsewhere [67].

Figure 5-5 (left) shows the chamber gas gain vs. high voltage. The nominal operational HV point is chosen to be 3.6 kV, which corresponds to gas gain of the order 7×10^4 . Taking into account that a minimum ionizing particle (mip) leaves behind in a gas gap about 100 electrons, the total charge in an avalanche per mip is about 1 pC. As will be shown below, at this operational point, the cathode and anode electronics have a very high efficiency and an adequate signal-to-noise ratio. The operational range of chambers extends to 3.9 kV. Typically, we start seeing a sharp rise in rate of spurious pulses at about 3.9-4.0 kV—see Fig. 5-5 (right).

The 72 smallest ME1/1 chambers have variations in the mechanical design with respect to all other chambers. First, the gas gap is 7 mm, wire diameter is $30 \mu\text{m}$, and

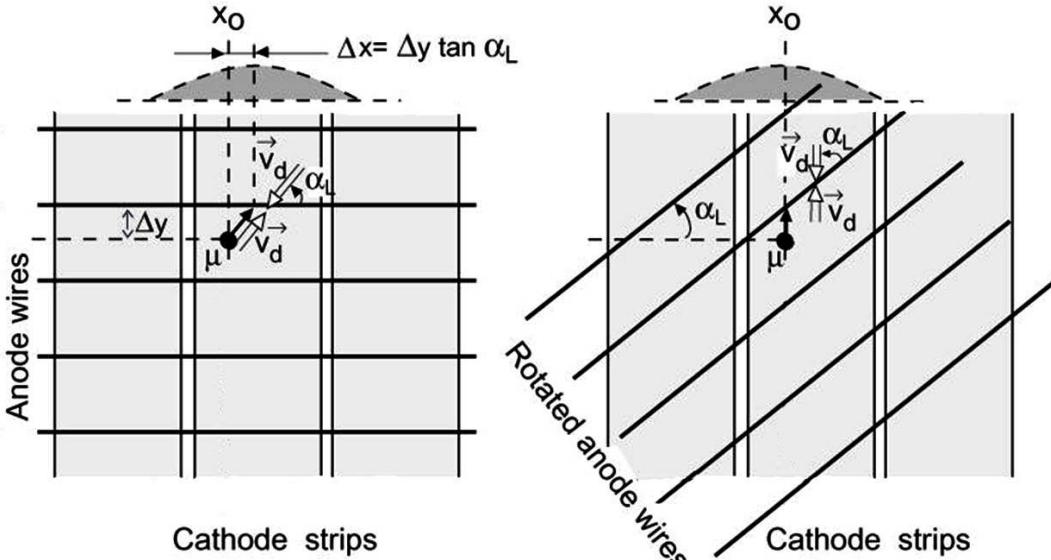


Figure 5-6: **Left:** If wires in ME1/1 chambers were not tilted, ionization electrons, as they drift towards anodes wires in strong magnetic field normal to the plane of the drawing, would be carried sideways by the Lorentz force. Direction and the size of the shift depend on whether electrons drift upward or downward and on how far away they are from wires to begin with. These sideway displacements would spread the charge over anode wires. **Right:** By tilting wires at the Lorentz angle α_L , one makes all ionization electrons arrive at the same point.

wire spacing is 2.5 mm. As a consequence, the nominal high voltage for these chambers is somewhat lower, 2.9 kV. Secondly, and more importantly, anode wires in ME1/1 chambers are not azimuthal, but rather rotated by an angle $\alpha_L = 29^\circ$ as shown in Fig. 5-6. Unlike all other CSCs, the ME1/1 chambers are inside the CMS solenoid and see its strong and uniform 4 Tesla axial field. If the wires were not tilted, the ionization electrons drifting toward wires would be carried sideways at the Lorentz angle and become spread along wires and across strips. The wire tilt compensates for the Lorentz angle so that electrons drift parallel to strips and the precise measurement of the r_ϕ -coordinate remains possible.

5.3 Electronic Design

Figure 5-7 shows a schematic layout of the custom trigger and readout electronic boards developed for the Endcap Muon System.

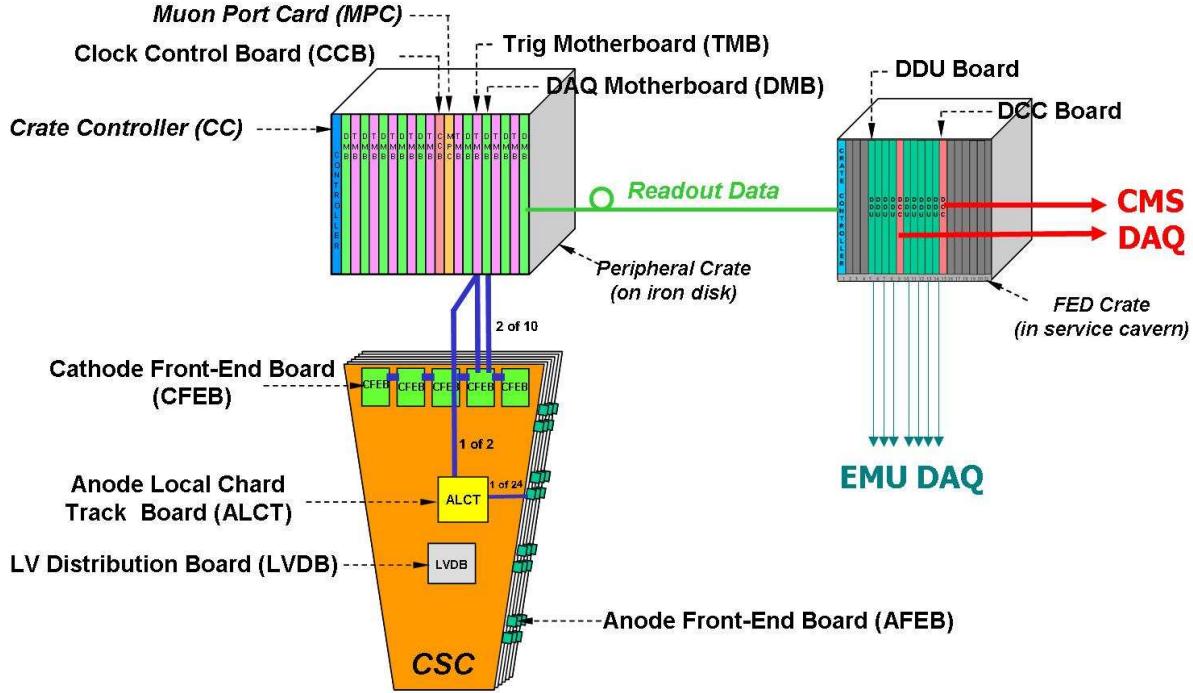


Figure 5-7: An overview of all custom-made CSC trigger and readout electronics. Functionality of various boards is described in the text.

An anode front-end board (AFEB) has one 16-channel amplifier-discriminator ASIC. The amplifier has a 30 ns shaper (semi-Gaussian with a 2-exponent tail cancellation designed to suppress the slow signal component associated with a drift of positive ions away from anode wires), about 7 mV/fC sensitivity, and 1.4 fC noise at a typical 180 pF wire group capacitance for the largest chambers. With the 30-ns shaping time, AFEB sees about 12% of the total avalanche charge, i.e. about 130 fC on average. A typical chamber signal as seen at the output of this amplifier is shown in Fig. 5-8 (left). The constant-fraction discriminator has a threshold nominally set at 20 fC (input equivalent charge) and its slewing time is less than 3 ns for 60-600 fC signal range. Depending on a chamber size, there are 12 to 42 AFEBs per chamber. All details on the AFEB design and performance can be found elsewhere [67].

Every 25 ns in sync with the LHC collisions, all AFEB outputs, 40-ns long step pulses, are sampled by an FPGA-based Anode Local Charged Track (ALCT) board, one

board per chamber. The recorded yes/no information is stored in FIFO. Upon receiving a CMS-wide "Level-1 trigger accept" command (L1A), the recorded information from a proper time window is extracted and reported to DAQ. The latency of the L1A command with respect to the time of a collision is 3.2 μ s. The width of a record with raw hits transmitted to DAQ can be set to be as large as 32 bits (1 bit per 25 ns), i.e. 800 ns.

The ALCT board has another important function. Based on the information from all anode channels, the FPGA code constantly (every 25 ns) searches for patterns of hits in six planes that would be consistent with muon tracks originating from the interaction point. For a pattern to be valid, we require that hits from at least 4 planes were present in the pattern. Figure 5-8 (right) illustrates how patterns are identified among spurious single-plane hits. Due to a large neutron-induced photon background, we expect a substantial rate of such single-plane hits. However, these hits, being completely uncorrelated, would not typically line up to form track-like patterns. Found patterns, Anode Local Charged Tracks, are trigger primitives. They are transmitted further upstream to the muon Level-1 trigger electronics that builds muon track candidates from these primitives. The time it takes to form an anode track trigger primitive is 225 ns (including drift time). The ALCT board can find up to 2 such patterns per each bunch crossing, which is well adequate for the expected chamber track occupancy at the nominal LHC luminosity.

One cathode front-end board serves (6 planes) \times (16 strips) = 96 channels and has six parallel chains of the following chips (Fig. 5-9 (left)): 16-channel amplifier-shaper ASIC, 16-channel Switched Capacitor Array ASIC, 12-bit 1-channel ADC, and 16-channel Comparator ASIC. There are 4 to 5 CFEBs per chamber.

The front-end amplifier-shaper ASIC has 100 ns shaping time and 0.85 mV/fC sensitivity over the liner range up to 1 V. The equivalent noise level at \sim 300 pF strip capacitance is typically 1.5 fC. The shaping is based on semi-Gaussian transfer function with an overshoot designed to compensate for the $1/t$ signal tail due to slow drift of

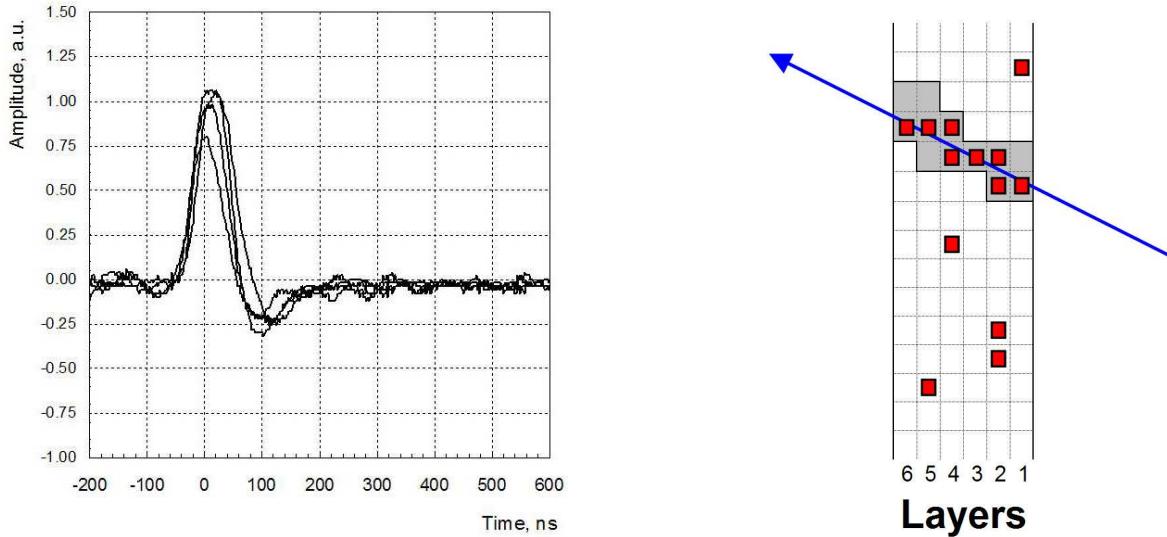


Figure 5-8: **Left:** Muon signals as seen at the AFEB amplifier output. **Right:** A schematic event display showing anode signals in a six-plane chamber (small dark squares). The ALCT board FPGA logic is programmed to scan along the chamber and search for hits falling inside predefined patterns (gray cells) consistent with muons originating from the interaction point. For a pattern to be found, hits from at least four planes should be present in it.

positive ions. After convolution with the current pulse produced in a chamber by a muon, the amplifier-shaper signal peaks at around 150 ns and has no tails—see Fig. 5-9 (right). With the 100-ns shaping time, CFEB sees about 8% of the total avalanche charge, i.e. about 100 fC on average.

The output from this chip is split in two pathways. One leads to the Switched Capacitor Array (SCA) ASIC. The SCA chip samples a waveform of each strip signal every 50 ns in sync with the LHC clock and stores this analog information on its capacitors. The depth of such analog memory is 96 capacitor cells per each channel, or $96 \times 50 \text{ ns} = 4.8 \mu\text{s}$. Upon receiving the L1A command in $3.2 \mu\text{s}$ after collision, 8 or 16 consecutive samples from the proper time range in the SCA line of capacitors are retrieved and one-by-one digitized by the 12-bit flash ADCs. The digital information is passed on to the data acquisition system via an intermediate digital data buffer. For the digitization and subsequent readout by DAQ to happen, the L1A signal must be in

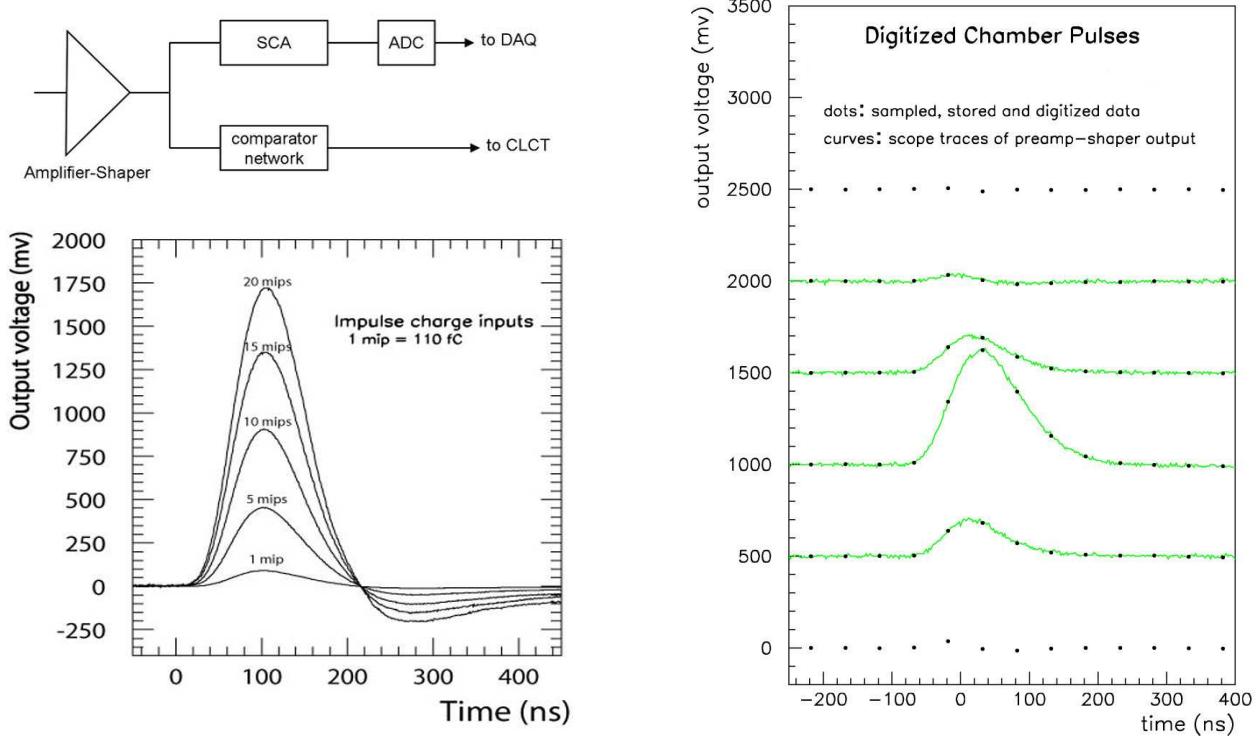


Figure 5-9: **Left:** Basic functional diagram of a Cathode Front-End Board (functionality of the main blocks is explained in the text) and CFEB amplifier-shaper response on a δ -function input pulse. The overshoot is intended to compensate for the long tail present in muon hit signals resulting from a slow drift of positive ions away from anode wires. **Right:** Muon hit signals from six nearby strips. Four curves are actual oscilloscopes. Six lines of dots are digitized outputs.

a coincidence with the Cathode Local Charged Track primitive decision to be described next.

The second amplifier-shaper output goes to the Comparator ASIC. This chip compares signals on triplets of nearby strips at the time when signals reach the maximum amplitude. By means of such comparisons, the Comparator ASIC can identify muon hit location within one half of a strip width, independently of a signal amplitude, an induced charge shape (as long as it is “bell”-like), and a strip width itself [68]. Figure 5-10 (left) illustrates the basic idea behind the Comparator ASIC algorithm.

Comparator half-strip hits are sent to the Trigger Mother Board (TMB). Similarly to the ALCT board, the TMB searches for patterns of half-strip comparator hits that would be consistent with muon tracks of interest—see Fig. 5-10 (right). Up to two Cathode

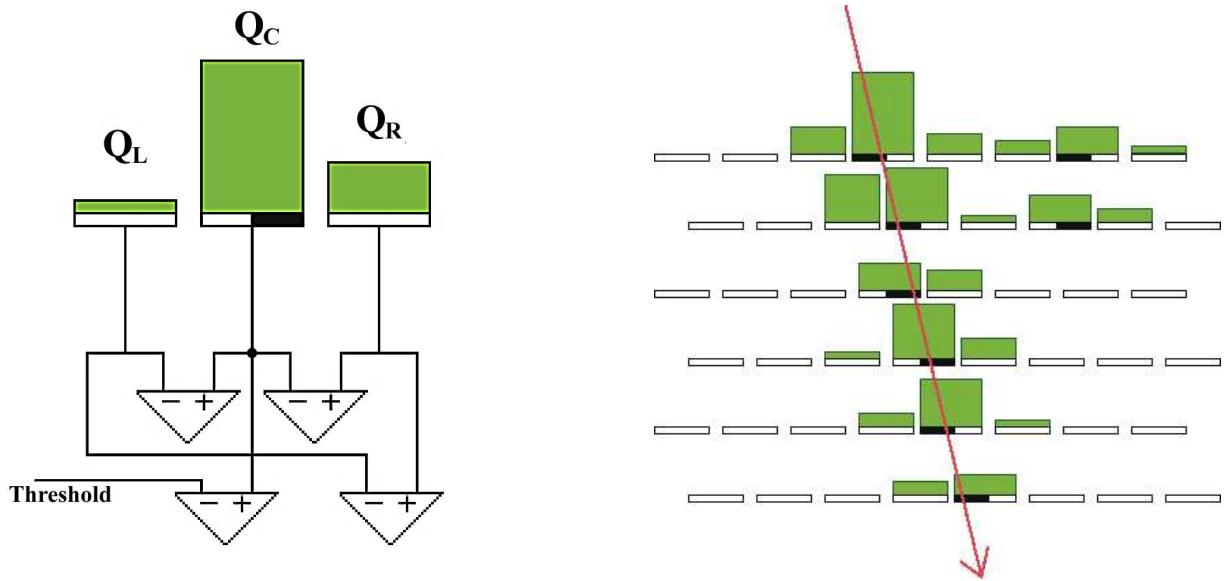


Figure 5-10: **Left:** A simplified schematic illustrating the idea behind the comparator network capable of quickly localizing the muon hits within one half of a strip width. For each triad of nearby strips, comparators compare the central strip signal Q_c with a threshold, central-to-left $Q_c - Q_l$, central-to-right $Q_c - Q_r$, and right-to-left $Q_r - Q_l$. If $Q_c >$ threshold, $Q_c - Q_l > 0$, $Q_c - Q_r > 0$, $Q_r - Q_l > 0$ (as in this figure), the hit position is somewhere within the right half of the central strip. **Right:** CLCT (Cathode Local Charged Track) is a pattern of half-strip hits consistent with a muon track.

Local Charged Tracks (CLCTs) per bunch crossing can be found. As in the ALCT pattern search, for a CLCT pattern to be found, half-strip hits from at least four planes should be present in it. There is one TMB per chamber. Unlike ALCT boards, TMBs are not on chambers, but rather in peripheral crates mounted along the outer rim of the endcap iron disks.

The TMB also matches ALCT and CLCT patterns found within a chamber to make correlated 2d-LCTs ($2d\text{-LCT} = \text{ALCT} \times \text{CLCT}$), up to two per bunch crossing. These 2d-LCTs are sent to the Muon Port Card (MPC), which serves 9 chambers covering either 60° - or 30° -sectors in ϕ . For each bunch crossing, MPC performs a preliminary sorting of all received correlated 2d-LCTs and finds 3 best candidates—these are then sent further upstream to the muon L1-trigger electronics.

All raw data are collected by DAQ Mother Boards (DMB). There is one DMB for each chamber; DMBs are located in peripheral crates as well. The data consists of raw anode and comparator hits within a time windows up to 32 bunch crossing, ALCT and CLCT decisions in the same windows, and digitized strip signal waveform (8 or 16 50-ns time samples). Status of the various electronic boards is also a part of the event record. The data collected by DMB is passed to DDU (Detector-Dependent Unit), then further to DCC (Data Concentration Card), and finally to the CMS Filter Farm to be processed by the CMS High-Level Trigger software. Event size per chamber is about 5 kBytes.

It is important to note that the CSC readout is intrinsically zero-suppressed. The anode raw data in a particular chamber are passed upstream only if there was an ALCT pattern in coincidence with the L1A signal. Likewise, the cathode information, comparator hits and digitized strip signal waveforms, are passed upstream to DAQ only if there was a similar $\text{CLCT} \times \text{L1A}$ coincidence. The coincidence window is programmable, but nominally set at 75 ns, i.e. ± 1 bunch crossing.

At the ultimate LHC luminosity, on average, we expect to find track stubs in 2 chambers per each L1A signal. With the maximum CMS L1A rate of 100 kHz, the data flow rate from CSCs to HLT is estimated to be around 1 GB/sec.

Operation of peripheral VME crates is supported by Clock-Control Board (CCB) and custom Crate Controller (CC). CCB, as its name implies, distributes LHC clock and all CMS control commands (like L1A signals).

The High Voltage system is custom-made and provides channel-by-channel regulated voltage up to 4.0 kV with about 10 V precision. Currents of less than $10 \mu\text{A}$ can be measured with a precision of 100 nA, while the precision for larger currents is about 1%. The system can provide more than $100 \mu\text{A}$ current for individual channels as long as the average consumption does not exceed $40 \mu\text{A}$ per channel. The maximum expected current at the ultimate LHC luminosity for the most loaded HV segment is $<10 \mu\text{A}$.

CHAPTER 6

FAST ALGORITHM FOR TRACK SEGMENT AND HIT RECONSTRUCTION IN CATHODE STRIP CHAMBERS

6.1 Introduction

Muon reconstruction in the High Level Trigger (HLT) starts from finding local track segments in muon chambers (for the endcap muon system, in six-plane cathode strip chambers). The found local track segments in all muon chambers and associated with them reconstructed hits, RecHits, are then used for a standalone muon reconstruction. An event is accepted by the L2 trigger (the first stage of the HLT event selection), if there is at least one muon with transverse momentum $p_T > 19 \text{ GeV}/c$ or if there are at least two muons with $p_T > 7 \text{ GeV}/c$ each (these are luminosity dependent threshold taken from [64] for $2 * 10^{33} \text{ cm}^{-2} \text{ sec}^{-1}$ luminosity). The following stages of HLT involve calorimeter (L2.5—muon isolation) and reconstruction of tracks in the tracker (L3—refined momentum and muon isolation). More details on the CMS High Level Trigger can be found elsewhere [64].

This chapter deals with the track segment reconstruction. The four main criteria qualifying a track segment finding performance are as follows:

- CPU time per event and operational code robustness;
- efficiency of finding segments associated with muons of interest;
- spatial resolution for localizing such segments;
- rate of finding secondary segments not directly associated with a muon of interest.

CPU time per event is derived from the overall HLT Event Filter Farm size (4000 CPUs) and the maximum L1 trigger rate of 100 kHz. Given these numbers, one has only 40 ms per event for the whole HLT processing, which includes the L2 stage (standalone muon) with 100% weight, L2.5 (calorimeter isolation) with $\sim 30\%$ weight, and L3 (tracks in the tracker) with $\sim 20\%$ weight. The weights come from the fact that each

next level trigger sees only a smaller fraction of the incoming events. The local segment reconstruction within a single chamber is among the simplest operations required from HLT and, therefore, only a very small fraction of the overall 40-ms budget can be allotted for that. A typical L1-trigger event is expected to have on average three CSCs with track stubs in them. At the time of writing this chapter, the HLT algorithms have not been optimized for time performance and, therefore, it is not possible to set a quantitative specification on the maximum CPU allotment for a local segment reconstruction, but, clearly, it is hardly going to be much more generous than a few ms per CSC (existing algorithm were measured to take about 15 ms per event). **So, our goal was to develop and validate with real data a new, fast, reliable, efficient, and precise algorithm for CSCs which would be suitable for HLT needs.**

Efficiency of finding muon track segments is desired to be at least 99% over the sensitive chamber area. This will allow for reconstruction of muon tracks in the whole system with high efficiency and small systematic uncertainties.

Spatial resolution per segment at HLT is not required to be much better than ~ 0.5 mm. This can be illustrated by the following simple considerations:

- ▷ As far as the standalone muon reconstruction is concerned (L2-part of HLT), the error in a sagitta for the ME1-ME2-ME3 stations due to multiple scattering in 60-cm iron disks separating them for a muon with $p_T=20$ GeV at $\eta=2.0$ ($p=75$ GeV) is about 1 mm. The error for sagitta based on IP-ME1-ME2 (IP—interaction point) for the same muons is larger than 3 mm.
- ▷ For matching a found standalone muon to tracks in the Si tracker (or wise versa, matching Si tracks to muon stubs in the muon chambers), one needs to make an extrapolation through the calorimeters. The extrapolation error due to multiple scattering for the same benchmark muons ($p_T=20$ GeV at $\eta=2.0$) is about 4 mm.
- ▷ Once the tracker hits are included in the muon reconstruction, the muon momentum resolution is then completely defined by the tracker up to 200 GeV in p_T [69]. Therefore, below the highest expected p_T threshold of 19 GeV, the final muon momentum measurement at HLT is insensitive to the muon chamber resolution at all.

Rate of finding associated secondary segments not directly associated with a muon of interest is another very important measure of merit for segment finding performance

both for HLT and offline muon reconstruction. Secondary segments can originate from physical secondary tracks (e.g., jet punchthrough and electromagnetic showers associated with high-energy muon bremsstrahlung). Combinatorial fakes resulting from all possible pairings of anode and cathode hits as one forms 2d-hits and/or segments are yet another source of secondary segments. A single secondary segment, simply by an incorrect association with a physical soft muon, can result in a dramatic mis-measurement of a muon momentum and promoting it to a much higher p_T . At the trigger level, this leads to larger rates of fake high- p_T muons and, even worse, to flattening of the trigger rate vs. muon p_T threshold, thus, resulting in inability to control muon trigger rates. In offline data analyses, one would have to cope with the associated high rates of fake high p_T muons with poorly understood systematic uncertainties.

Development and testing of the proposed algorithm was performed in new CMS software framework (CMSSW), which is being developed for the future use by CMS detector for taking and analyzing data.

Details of our analysis and supporting studies on the new algorithm are being prepared for publication and can be also found in the author's talks at the internal CMS collaboration meetings of different levels [21].

6.2 Algorithm Principles

Suggested by the University of Florida group, the fast track segment finding algorithm described in this chapter has become known as the UF algorithm. In the interests of brevity and to distinguish the new algorithm from the others (SK, TC, DF, ST, see for example [70] for details), we will further refer to it under this nickname. The UF algorithm is built on the following principles.

- First, instead of starting from pairing anode and cathode strip charge clusters in individual planes to form 2d-RecHits and then trying to build 2d-segments of them, the UF algorithm starts from finding 1d-segments, Anode- and CathodeSegments, and pairs them to build 2d-ProtoSegments. This order saves substantially CPU time needed to reconstruct complex events with more than just one hit per plane.

- Second, having found 2d-ProtoSegments, the UF algorithm then takes only those strip charge clusters that fall within a close proximity of the found segments and reconstructs their precise coordinates. To speed up calculations, the reconstruction is done without accessing any calibrations and/or databases. By default, we use the ratio-method [71]. The ratio method requires no iterations, is very fast, and delivers a spatial resolution substantially better than what is required for the HLT purposes.
- And finally, the last step is to refine parameters of the found 2d-ProtoSegments and remove “bad” RecHits. To do that, we make a linear fit of all RecHits originally associated with each 2d-ProtoSegment. If fit’s χ^2 is not satisfactory, some RecHits are allowed to be pruned. This is also a linear non-iterative procedure taking very little time. The final linear fit defines the RefinedSegment parameters.

In short the key principles of the UF algorithm are: first, 1d-segments; second, 2d-segments; then, refine RecHits; and, at last, pruning of “bad” RecHits and refinement of 2d-segment parameters. Certainly, additional layers of sophistication can be added to each of the three steps of the UF algorithm, provided they help improve the performance and keep the overall computational time in check.

6.2.1 ProtoSegment Building

To define 2d-ProtoSegments, in the simplest implementation of the UF algorithm, we just use ALCT (Anode Local Charged Track) and CLCT (Cathode Local Charged Track) trigger primitives for AnodeSegment and CathodeSegment. These trigger primitives (patterns of raw hits consistent with muons originating from the interaction point) are already found by the “near-frontend” electronics and reported in the data stream together with all raw data. All details on electronic architecture and functions can be found in Muon Technical Design Report [65]. Up to two ALCTs (ALCT0 and ALCT1) and two CLCTs (CLCT0 and CLCT1) can be found per each bunch crossing. If more than one primitive is found in either projection, then we combine them combinatorially to form 2d-ProtoSegments, of which we can have 1, 2 or 4 per chambers. The two coordinates of a 2d-ProtoSegment are ALCT’s KeyWireGroup and CLCT’s KeyHalfStrip. All in all, building ProtoSegments this way requires virtually no CPU usage.

It is important to note that whenever two ALCTs and/or two CLCTs are reported by the frontend electronics, they are always the very best two muons candidates from all possible hit combinations. From physics side, the chances of finding two prompt muons within one chamber are very small. So, in fact, limiting ourselves to the best two muon segment candidates is probably more of an advantage than a limitation as long as the efficiency of finding the segment corresponding to the muon of interest remains high (see below).

Also worthwhile stressing is that this approach does not introduce any potential inefficiency as compared to any other algorithm that starts from raw data. By design, the CSC readout in CMS is intrinsically zero-suppressed and one does not get any raw data in the readout stream unless there was a trigger primitive found [65].

Efficiencies of finding trigger primitives were extensively studied in past on chamber prototypes operating in a muon beam or cosmic rays, in a muon beam with a superimposed flux of random hits [72], and with high energy muons accompanied by bremsstrahlung radiation due to muons passing through an iron slab in front of a chamber [73]. More recently, these efficiencies were studied *in situ* with 18 chambers installed in CMS and operating in cosmic rays [74]. In all these cases, the efficiency of finding trigger primitives was measured to be higher than 99%. For example, the latter studies gave an efficiency of $99.9 \pm 0.05\%$ for finding $\text{ALCT} \times \text{CLCT}$ 2d-patterns for muons passing through the chamber sensitive volume.

6.2.2 RecHits and Their Coordinates

As was mentioned earlier, the default UF algorithm does not use any calibration constants and does not access any databases. All information on the internal chamber geometry needed for the local reconstruction is taken from the chamber drawings.

Below are a few notations and definitions to be used further in the chapter:

- Noise associated with a single time sample (first time sample in CFEB readout) is σ_1 . Three samples added together have a spread of σ_3 .

- The averages of the first two SCA samples in CFEB readout, measured on event-by-event and channel-by-channel basis, define pedestals for the current event. Below in the section on the validation of the algorithm, we show that using tabulated pedestals allows one to achieve an even better resolution, but at an additional CPU cost.
- CFEBCluster is defined as a 3×3 matrix = (3 strips) \times (3 SCA samples) with the central sample being a local maximum. For a cluster to be identified as such, the amplitude of the central strip is required to be larger than $k\sigma_1$ (k is a configurable parameter; by default, $k=10$). For each strip in a cluster, three time samples are added together (pedestal subtracted). This way the 9-sample information in a cluster is reduced to three charges: Q_l , Q_c , Q_r (left, central, right). The cluster charge is $Q = Q_l + Q_c + Q_r$.
- Throughout this chapter, noise, pedestals, and charges are expressed in terms of ADC counts.

For each ProtoSegment present in the chamber, the following is performed.

Starting from ProtoSegment's coordinates (KeyHalfStrip, KeyWireGroup), the nearest CFEBCluster in each plane within $\pm n$ strips around the KeyHalfStrip is identified (n is a configurable parameter; by default, $n=5$). At the moment, clusters from different planes are not required to line up in time. Strip signal charges Q_l , Q_c , Q_r for such clusters are calculated and used to build the ratio r defined as follows:

$$r = \frac{1}{2} \frac{Q_r - Q_l}{Q_c - \min(Q_r, Q_l)} \quad (6-1)$$

This ratio changes monotonically, but not linearly from -0.5 to 0 to 0.5 as the true hit coordinate in strip width units changes from -0.5 to 0 to 0.5 (in points -0.5, 0, and 0.5, they coincide). Cluster's local coordinate in strip width units is then calculated via a function $x = f(r, w)$ that corrects for non-linearity between r and x (see Appendix I for details). The strip width parameter w in this function is derived from the chamber geometry and ProtoSegment's coordinates (KeyHalfStrip, KeyWireGroup). Note that x , being measured in strip width units, can be thought of as a ϕ -coordinate of a hit in units of strip $\Delta\phi$ -width.

The x -coordinate is then given an offset, according to the number of its central strip and taking into account the ± 0.25 staggering of strip patterns in odd/even planes in the large chambers.

It is important to note that, being built from charge differences, the ratio method is intrinsically not very sensitive to cross talks variations. Another important feature is that for hits close to strip edges, the ratio r basically becomes the ratio of two large charges on nearby strips and has nearly no sensitivity to charge fluctuations on the third strip with a very small share of the induced charge (and, thus, very little useful information)—for example, this is not the case for the center-of-gravity method.

By default, the UF algorithm assigns errors to reconstructed RecHits x -coordinates according to the tabulated functions $\sigma_x(CSCtype, HVsegment, |x|)$ for five ranges of $|x|$. The values for this function were obtained directly from MTCC data. More details are given further below where we discuss the algorithm validation.

The default option does not attempt to correct the errors in x for the charge in a cluster. To take into account the charge, a good understanding of charge-dependent and charge-independent contributions is needed. Note that at nominal gas gain, the errors for x -coordinates in the area between strips, where the resolution is the best, are already dominated by charge-independent contributions. Therefore, a charge-dependent assignment of errors may improve performance somewhat, but the effect is not expected to be dramatic. These studies are in progress. If they show that taking into account the charge-dependent component gives better performance without a risk for the algorithm’s robustness, the found corrections can be easily included.

If CFEB data are not present, then the x -coordinate is given by the center of the nearest half-strip with a comparator response. Again, as in the case with CFEBClusters, the search is done in the range of $\pm n$ strips around the KeyHalfStrip. The error in this case is defined as $\sigma_x = 0.5/\sqrt{12} = 0.144$.

RecHits are assigned local y -coordinates according to the center-of-gravity of anode hits in the corresponding plane and falling within the envelope of the corresponding AnodeSegment pattern. To calculate RecHit's y -coordinate in ME1/1 chambers with tilted wires, one also needs to take into account RecHit's x -coordinate.

6.2.3 RefinedSegment

To refine the x -coordinate of the original ProtoSegment, we make a standard non-iterative weighed linear fit of x -coordinates of all RecHits associated with the segment. The returned parameters are an intercept $x_{segment}$ at local coordinate $z=0$ (center of the middle chamber panel), with an estimated error on it, slope dx/dz with an estimated error, and $\chi^2/\text{d.o.f.}$. The local $x_{segment}$ -coordinate is directly related to the ϕ -coordinate of the CMS coordinate system.

If $\chi^2/\text{d.o.f.} > \chi^2_{cut}$, we re-fit the line by leaving out one plane at a time from the fit. The best fit gives new parameters for the RefinedSegment. Another round of such pruning is allowed (with the same criteria on χ^2) as long as the number of remaining hits does not fall below four, i.e. the final refined segment must have at least four RecHits associated with it.

For a 2d-RefinedSegment, the local y -coordinate is improved (in comparison to what one gets by taking plain KeyWireGroup) by taking a simple center of gravity of actual hits belonging to the AnodeSegment pattern. The local y -coordinate can be easily converted into the CMS global coordinates r and η .

Note that the refinement of 2d-segments means just that, refinement; i.e. the earlier found 2d-segments do not get removed, nor are new segments added in this process.

6.3 Algorithm Validation at MTCC

During August-October 2006, a slice of the CMS Detector was used to take cosmic ray data during the first activation of the CMS solenoid. This important CMS milestone was named Magnet Test and Cosmic Challenge (MTCC). Details on the MTCC scope and the obtained results can be found elsewhere [75]. This section describes the performance of the UF algorithm as obtained using the MTCC data.

Measurements of the processing time for the default UF algorithm were done on an Intel P4 2.8 GHz Dual Xeon Server with the CMSSW version 1.2.0 as the default framework. To benchmark the algorithm performance we used the CMS Global Run No. 4188. On average, the UF algorithm was found to take ~ 0.36 ms per matched LCT. This time includes finding segments, reconstruction/selection of uncorrupted RecHits associated with these segments, and refining segment parameters.

6.3.1 ProtoSegments

As was mentioned earlier, in the current version of the UF algorithm, the ProtoSegment is defined as a matched pair of trigger primitive patterns, ALCT \times CLCT. Efficiency of finding such correlated LCTs and their properties were extensively studied elsewhere [74]. Here we just summarize the main results from those studies:

- For muons going through a fully sensitive chamber area, the efficiency of finding a correlated LCT was measured in data to be $99.9 \pm 0.05\%$.
- The highest quality correlated LCT found in a chamber was near the muon position predicted from external chambers; the distribution spread of less than 2 cm (RMS) was consistent with multiple scattering of muons.

Therefore, the MTCC data confirm the earlier results showing that ALCT and CLCT patterns provide a very robust method of identifying muons. In contrast to the earlier results obtained in beam tests with only a small portion of a chamber illuminated by muons, the MTCC studies were done with cosmic rays probing nearly the entire area of 8

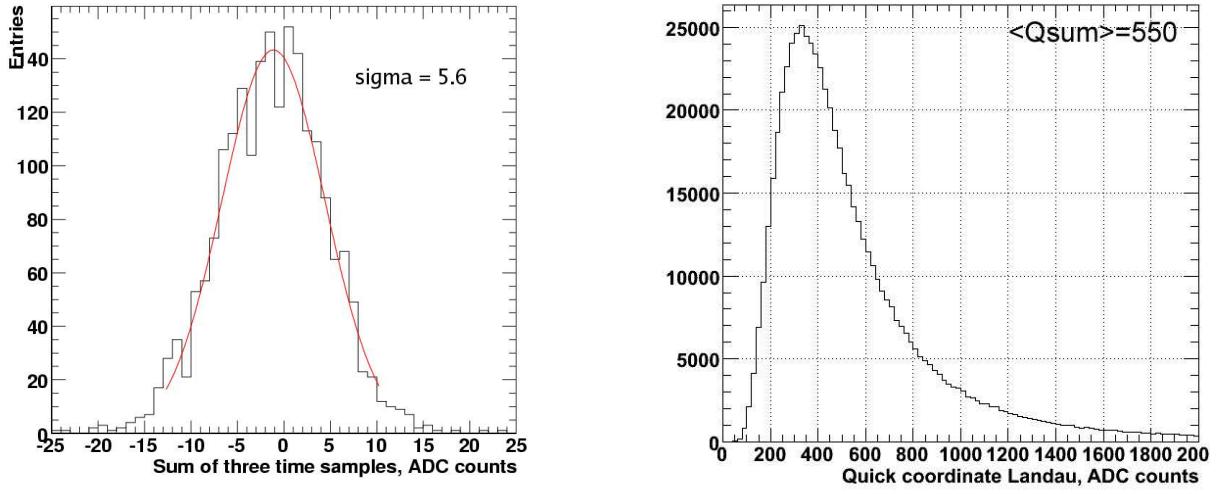


Figure 6-1: **Left:** The noise distribution for a three-sample-sum in the absence of a signal for the largest ME23/2 chambers. **Right:** The three-strip cluster charge distribution for the largest ME23/2 chambers.

large chambers operated *in situ* together with the rest of the CMS detectors participating in the MTCC.

6.3.2 Spatial Resolution

This sub-section summarizes results of the studies of the spatial resolution attainable with the ratio method in the context of the CMS Cathode Strip Chambers.

A signal-to-noise ratio is among the key parameters affecting the chamber performance. As described earlier, a strip charge in the UF algorithm is defined as a three-time-sample sum. Figure 6-1 (left) shows the three-sample sum distribution in the absence of a signal—the Gaussian sigma of this distribution is $\sigma_3 \sim 5.5$ ADC counts. Figure 6-1 (right) shows the Landau distribution for 3-strip charge clusters. The average charge is around 550 ADC counts. Therefore, the MTCC data were taken with the gas gain corresponding to signal-to-noise ratio of approximately 100:1.

Figure 6-2 (left) shows the single-sample noise level. It has a spread $\sigma_1 \sim 3.0$ ADC counts. It is worthwhile noting that $\sigma_3 > \sqrt{3}\sigma_1$, which clearly demonstrates the expected noise correlations between time samples.

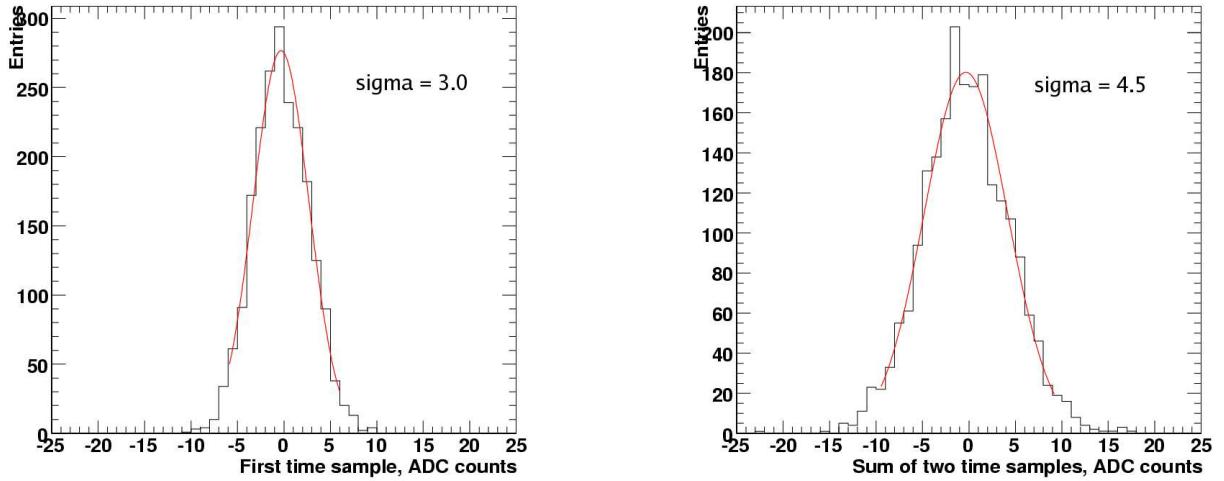


Figure 6-2: **Left:** The noise distribution for a single-sample in the absence of a signal for the largest ME23/2 chambers. **Right:** The noise distribution for a two-sample sum in the absence of a signal for the largest ME23/2 chambers.

Figure 6-2 (right) shows the two-sample sum distribution in an absence of a signal—the Gaussian sigma of this distribution is $\sigma_2 \sim 4.6$ ADC counts. The two-sample sum is of special interest for further considerations. The UF algorithm assumes two modes of operation. First, default, mode does not use any calibration constants, including pedestals. The pedestals for individual channels are evaluated on event-by-event basis from the first two samples, where signal is not present: $ped = (sample1 + sample2)/2$. Then, such dynamically defined pedestal is subtracted from all three samples used to form a three-sample-sum charge. This procedure contributes an additional error in determination of the charge: $(\sigma_2/2) \times 3 = 6.9$ ADC counts. Combined with $\sigma_3 = 5.5$, the total error in charge measurements becomes 8.8 ADC counts, which is 1.6 times worse than 5.5 ADC counts one could have, if tabulated (calibrated) pedestals were used. This is a substantial penalty for not using calibrated pedestals. Therefore, we single out pedestals from all calibration constants and implement a second mode of the UF algorithm that does use pre-defined pedestals. UF algorithm takes ~ 0.45 ms per matched LCT with this option ON.

6.3.2.1 Resolution for high- p_T muons

The best chamber resolution matters only for high- p_T muons whose track stubs are nearly perpendicular to cathode strips. Softer muons undergo large multiple scattering and a good detector resolution is less important. To evaluate the chamber resolution performance in this context, we select events according to the following criteria:

- We require only one ALCT and only one CLCT per chamber. This allowed us to extract the spatial resolution parameters intrinsic to the chambers themselves and not obscured by presence of showers.
- CLCT patterns are required to be half-strip patterns only (di-strip patterns are invoked for finding highly inclines muons in the absence of half-strip patterns).
- Also, to stay away from edge effects, we require that the ProtoSegment’s KeyWireGroup is at least one wire group away from chamber edges and dead areas separating HV segments, and KeyHalfStrip is at least 2 strips away from the chamber sides.
- We require six charge clusters to be associated with a ProtoSegment, which gets rid of incomplete tracks.
- The resolution is evaluated via residuals between a hit coordinate in the 3rd plane and the track coordinate in this plane as predicted from a straight line fit of hits in the remaining 5 planes (1, 2, -, 4, 5, 6). These residuals give a conservative estimate of the chamber resolution as they include track prediction errors. We do not attempt to correct for this effect (it can be easily estimated to result in an extra factor as large as 1.2 for a single-plane resolution in the “good” strip areas).
- We expect a few percent of δ -electrons per plane. The fraction of dramatically corrupted 5-plane fits due to δ -electrons in at least one plane is consequently is 5 times larger. Therefore, to make sure that the reference prediction is not dramatically compromised, we cut events whose 5-plane fit $\chi^2/\text{dof} = \chi^2/3 > 10$. This removes about 25% of tracks. Since typical δ -electrons do not have enough energy to penetrate from plane to plane, this cut does not bias the measurement in the test plane number 3. It rather insures that the reference prediction for the hit position in the 3rd plane is self-consistent.

The resolution per plane is analyzed and reconstructed for five regions across a strip width: 0-0.1 (strip center), 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5 (strip edges). To obtain

the final resolution values for these five strip regions, we perform the procedure described above iteratively. The convergence is reached in just a few iterations.

This analysis is performed for each chamber HV segment separately and, therefore, at the end, the chamber resolution is tabulated as $\sigma(CSCtype, HVsegment, |x|)$, where there are 5 distinct ranges for $|x|$: 0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5.

Figure 6-3 shows residuals for these five distinct ranges for the middle part (HV segment 3) of the largest ME23/2 chambers. The last plot shows a summary for a single-plane resolution σ_i vs. i -th $|x|$ -part of a strip ($i=1$ for $0 < |x| < 0.1$, $i=2$ for $0.1 < |x| < 0.2$, ..., $i=5$ for $0.4 < |x| < 0.5$).

These σ_i 's allow one to evaluate the overall six-plane (full-chamber) resolution as follows:

$$\frac{1}{\sigma_{csc,i}^2} = \frac{3}{\sigma_i^2} + \frac{3}{\sigma_{6-i}^2} \quad (6-2)$$

Figure 6-4 shows single-plane and full-chamber resolutions for all HV segments for ME23/2 chambers. For completeness, the results are shown both in units of strip widths and microns. Figure 6-5 show similar results for another distinct chamber type: ME1/1.

Clearly, the obtained resolutions surpass by far the HLT goal of 0.5 mm or so. To remind, the presented results do not use any calibration constants. Nor do we correct for internal plane mis-alignment, which, from the FAST site measurements, is estimated to contribute about 50 μ m RMS per plane [76].

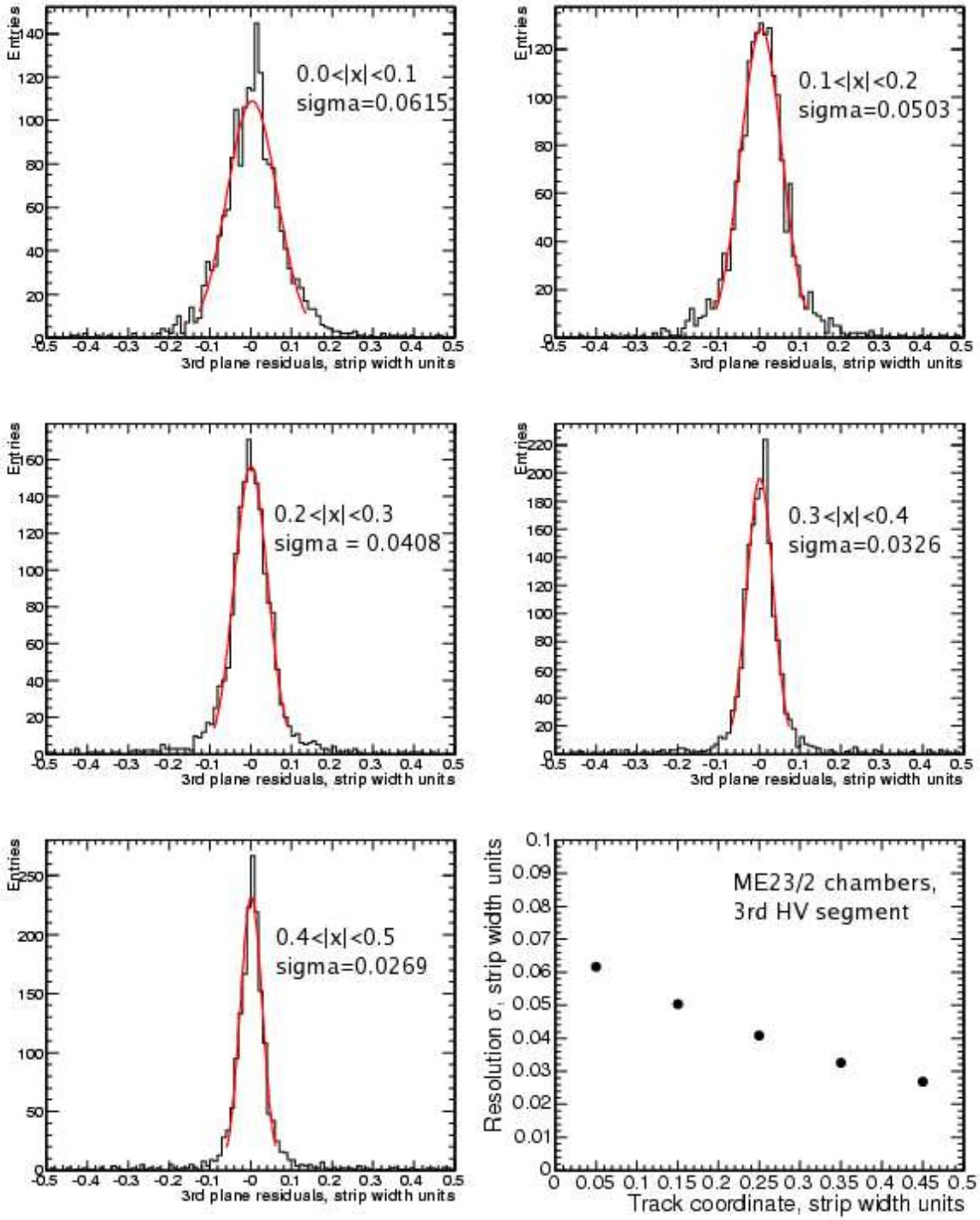


Figure 6-3: The 3rd-plane residuals for the five distinct $|x|$ strip part ranges ($0-0.1$, $0.1-0.2$, $0.2-0.3$, $0.3-0.4$, and $0.4-0.5$) for the middle part of the largest ME23/2 chambers (HV segment 3). The last plot (bottom right) shows a summary for a single-plane spatial resolution as a function of a hit position across a strip (hit coordinate is given in strip width units $|x|$).

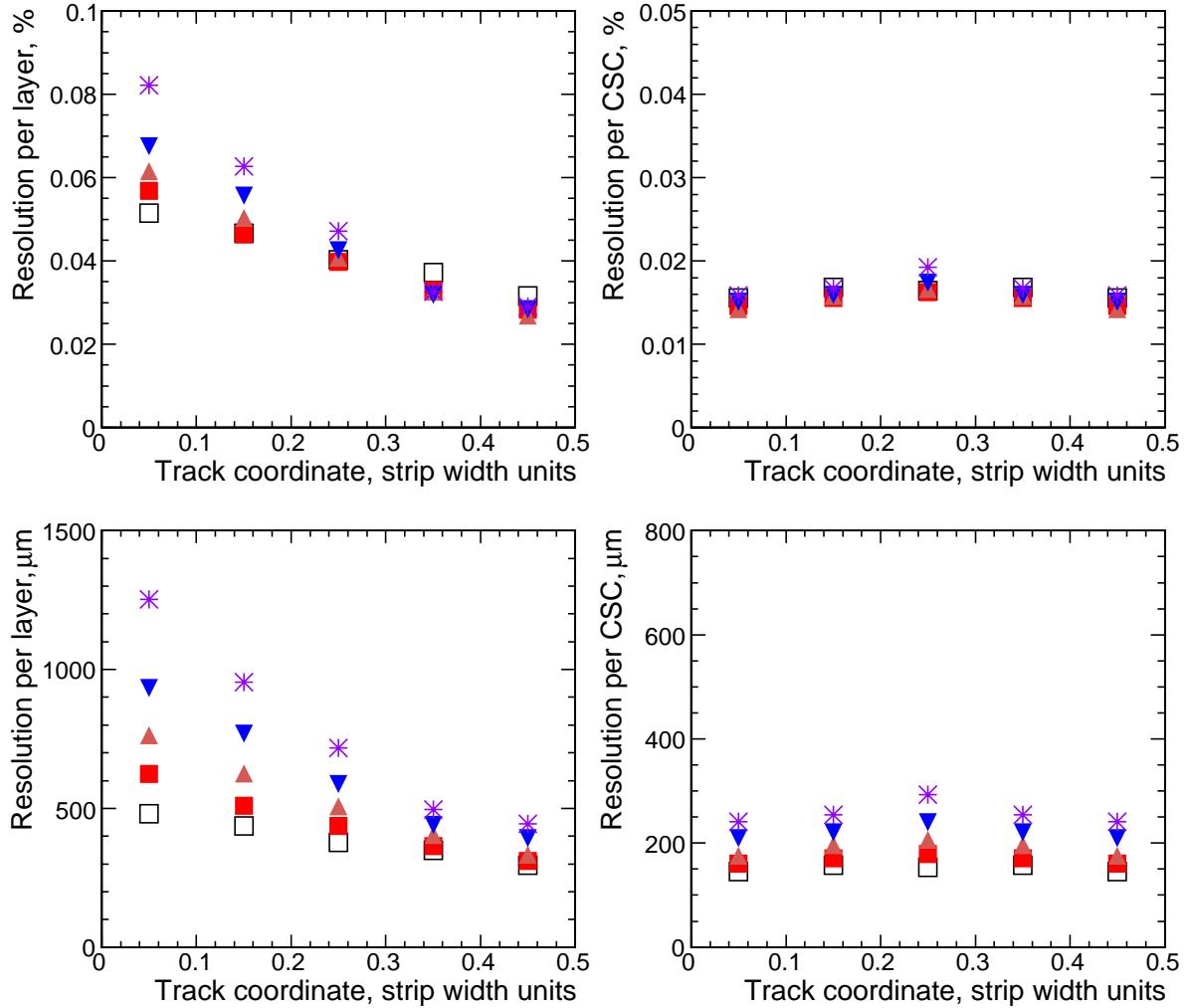


Figure 6-4: Summary of resolution results for the largest ME23/2 chambers. Left: the 3rd-plane residuals σ 's vs. five $|x|$ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5). Right: overall six-plane resolution vs. $|x|$ -part of a strip. Top plots show results in units of strip widths, the bottom—in absolute microns. [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.]

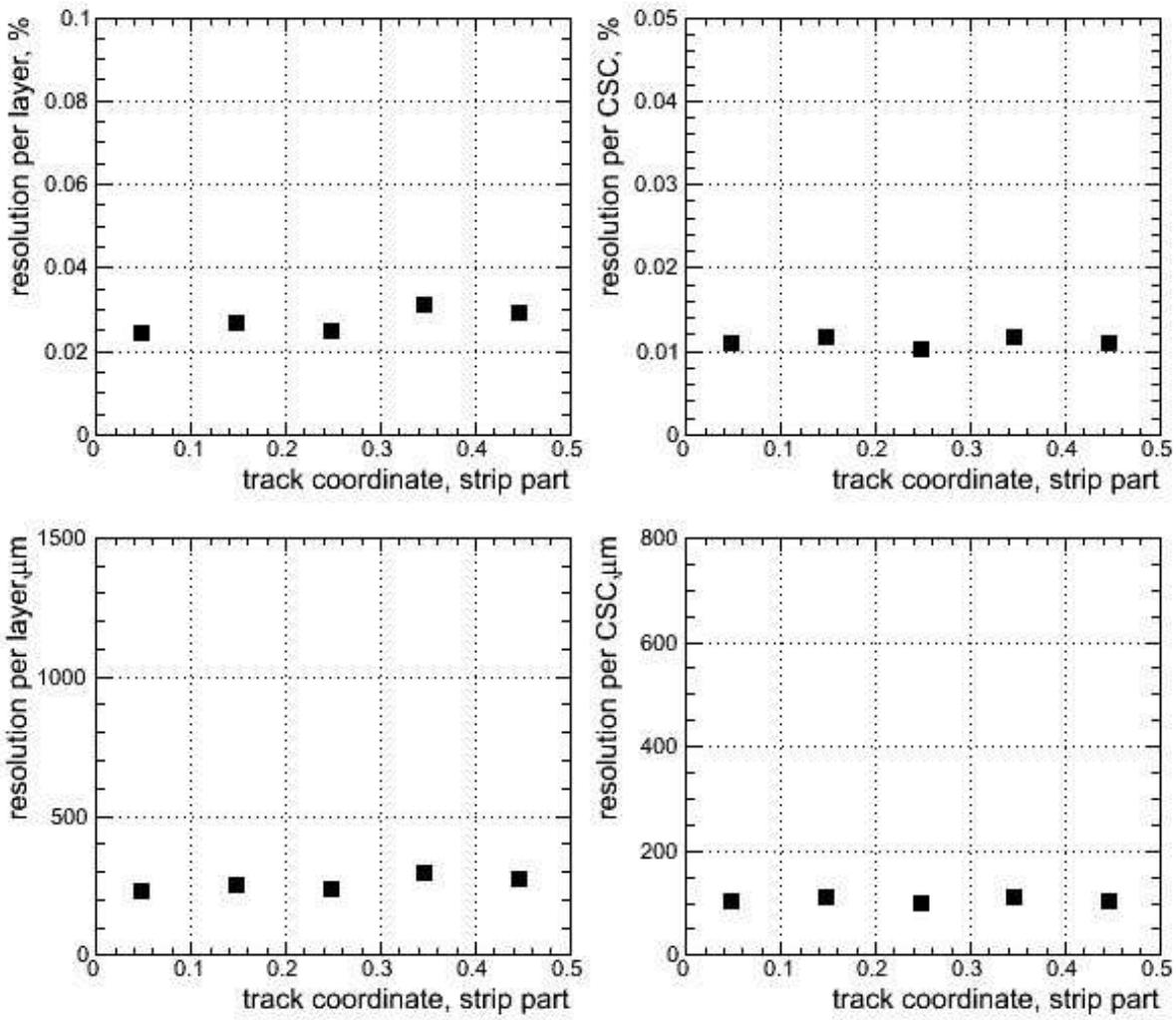


Figure 6-5: Summary of resolution results for the top part of the ME1/1 chambers. Left: the 3rd-plane residuals σ 's vs. $|x/w|$ -part of a strip. Right: overall six-plane resolution vs. $|x/w|$ -part of a strip. Top plots show results in units of strip widths, the bottom—in absolute microns. [Notations are: black squares - HV segment 1.]

6.3.2.2 Resolution for high- p_T muons with tabulated pedestals

Measuring pedestals dynamically (on event-by-event basis) allows one to not use calibrations and would help at very high rates when pedestals are expected to float. However, as it was mentioned earlier, these benefits come with a penalty—the electronic noise results in errors in the measured pedestals, which makes a significant contribution to the final errors in charge measurements.

To evaluate the effect quantitatively, we measured pedestals directly from data and used them in the analysis. The 3rd-plane residuals obtained with tabulated pedestals are then used to evaluate the full six-plane resolution for ME23/2 chambers. The results are shown in Figure 6-6. The left plot is for dynamically measured pedestals, the right plot is for tabulated pedestals. The gain in resolution is quite tempting. As long as the rate of hits in chambers is not too high, one can certainly take advantage of this option, especially in the offline analysis.

As far as the HLT is concerned, the improvement in resolution, however impressive it might be, is not critical (the resolution with dynamic pedestals is already better than needed). Nevertheless, the option of using tabulated pedestals certainly can be used for the HLT as well, as long as the CPU penalty (yet to be measured) due to the need to manage about 200K array of constants is acceptable.

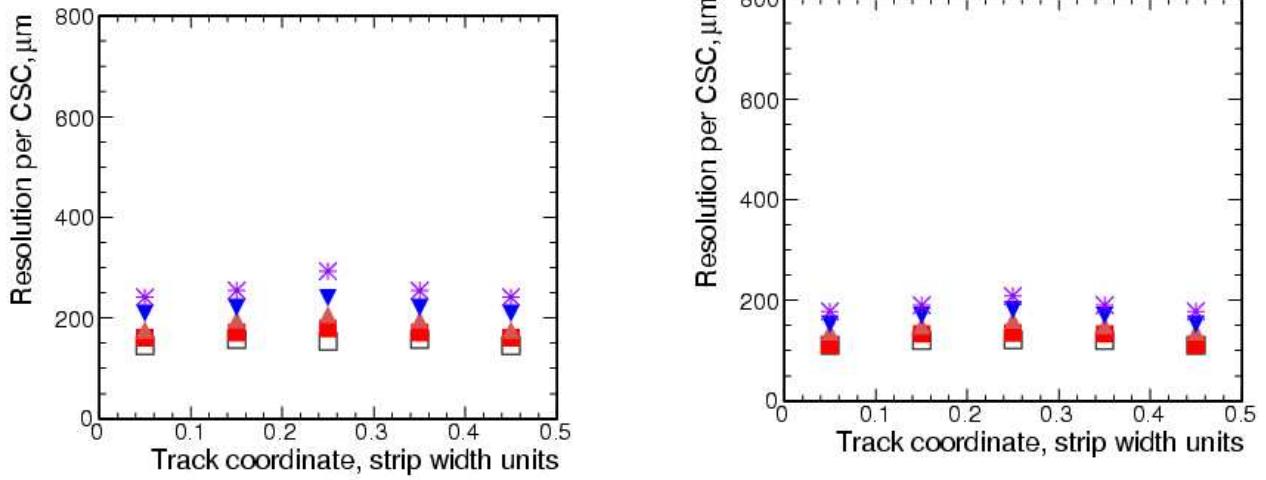


Figure 6-6: The overall six-plane chamber resolution vs. five $|x|$ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5) for dynamically measured pedestals (left) and calibrated pedestals (right). [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.]

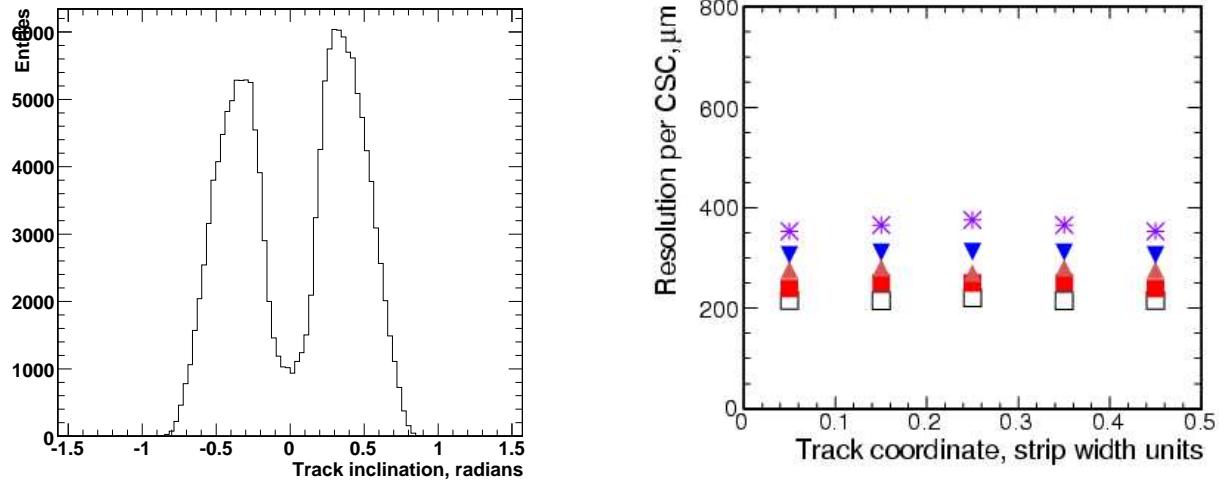


Figure 6-7: **Left:** Inclination angles for track segments reconstructed from di-strip CLCTs (ME1/1 chambers). **Right:** The spatial resolution per six-plane chamber for tracks reconstructed from di-strip CLCTs (ME1/1 chambers) for five $|x|$ strip part ranges (0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5). [Notations are: black empty squares - HV segment 1; red squares - HV segment 2; triangles up - HV segment 3; triangles down - HV segment 4; stars - HV segment 5.]

6.3.2.3 Resolution for low- p_T , or highly-inclined, muons

Very low p_T tracks, will have a noticeable angle (due to magnetic field and multiple scattering) of inclination dx/dz . For example, muons with $p_T \sim 3$ GeV/c at $\eta=1.6$ (just barely making into ME1/1 chambers), will be swept sideways by magnetic field B and would go through ME1/1 chambers at an angle of $\alpha_B \sim 0.35$ radians. For other stations and for muons with larger P_T and/or at higher pseudo-rapidities, the angle α_B is smaller. In the current version of CLCT-finding firmware, such “highly-inclined” muons are captured by CLCT di-strip patterns. The resolution for such segments is going to be worse due to smearing of the ionization charge along anode wires. But the requirements on the resolution for soft muons are also less stringent due to much larger multiple scattering (in comparison to benchmark muons of $p_T=20$ GeV/c that we have used so far).

To evaluate the spatial resolution attainable with the UF algorithm for “highly-inclined” muons, we selected ProtoSegments based on $LCT = (\text{di-strip CLCT}) \times ALCT$ in ME23/2 chambers. This is a conservative estimate as the effect of worsening in ME1/1 chambers will be smaller. The results are shown in Figure 6-7. The left plot shows angles for track segments reconstructed from di-strip CLCTs. One can see that the range is much wider than expected for a few GeV p_T muons. The right plot is the resolution per six-plane chamber—clearly, the resolution is still much better than a few-millimeter resolution needed for such soft muons.

6.3.2.4 Resolution for high- p_T muons with charge-dependent sigmas

An additional layer of sophistication can be added to the RecHit reconstruction by adjusting the errors on RecHit coordinates $\sigma(CSCtype, HV\ segment, |x|)$ for the actual charges in clusters. Before implementing such an adjustment, one must perform detailed studies of the relative interplay of different sources of errors in the hit reconstruction. To name a few, among these contributions are electronic noise (improvement in resolution with charge as $\sim 1/Q$), δ -electrons (degradation of resolution for too large charges), and a number of constant terms (place-to-place variations and event-to-event fluctuations, in induced charge shape, mis-calibrations, mechanical plane misalignment, etc.).

Given that the spatial resolution that we obtain without any of these corrections is already much better than what is needed for HLT, these corrections, once properly tuned, would make much more sense for offline analysis than for HLT. Of course, they can be used at HLT as well.

6.3.3 RefinedSegments

As was outlined in section on the algorithm principles, a 2d-RefinedSegment is basically a 2d-ProtoSegment, whose parameters have been refined using the precise coordinates of RecHits. The last touch in building RefinedSegments is pruning of seemingly “bad” RecHits. This is done by identifying and throwing away one-two “bad” RecHits that make dramatically bad contributions to a linear fit χ^2 . We derive the χ^2 -cut criteria from MTCC data as follows.

We select ProtoSegments based on LCT=(half-strip CLCT) \times ALCT in ME23/2 chambers and suppress complicated events with showers by requiring that there is only one ProtoSegment per chamber. To stay away from edge effects, we require that the ProtoSegment’s KeyWireGroup is at least a unit away from the chamber edges and the borders separating HV segments, and KeyHalfStrip is at least 2 strips away from the chamber sides. The chamber ME2/3/28 is excluded from further analysis as it has one HV segment switched off.

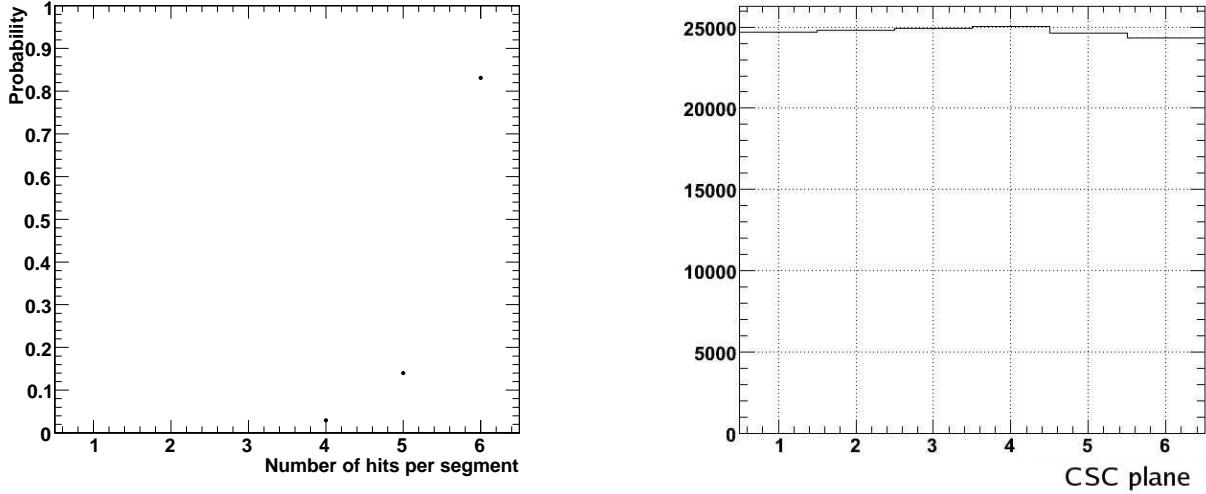


Figure 6-8: **Left:** Probabilities for different numbers of charge clusters found in association with ME23/2 chamber ProtoSegments. **Right:** Layer charge clusters occupancy in association with ME23/2 chambers ProtoSegments.

Figure 6-8 (left) shows how many charge clusters are found on such ProtoSegments. Note that by the UF algorithm design, only one cluster per plane can be associated with a given ProtoSegment. Figure 6-8 (right) shows an occupancy of a particular plane from 1st to 6th with charge clusters.

Figure 6-9 (left) shows χ^2/dof distribution for a linear fit for events with six, five, and four clusters associated with a ProtoSegment. The distributions have a long tails due to δ -electrons and other possible sources of hit corruption (e.g., showers).

Figure 6-9 (right) shows a scatter plot of $\min(\chi_5^2/\text{dof})$ vs. χ_6^2/dof for events with six-cluster ProtoSegments. Here, χ_6^2/dof is obtained for a six-plane fit, $\min(\chi_5^2/\text{dof})$ is a minimum between 6 five-plane fits with one plane dropped from the fit. The choice of cuts for pruning bad hits is shown by dashed lines. Probability of having 6-hit, 5-hit, and 4-hit RefinedSegments is then 78%, 13%, 9% correspondingly. We do not prune bad hits in segments with four remaining hits.

Figure 6-10 (left) shows that the occupancy for $|x|$ -coordinates for RefinedSegments remains flat, which demonstrate that the pruning of hits does not bias segments.

Figure 6-10 (right) shows pools for all retained RecHits (i.e., RecHits associated with

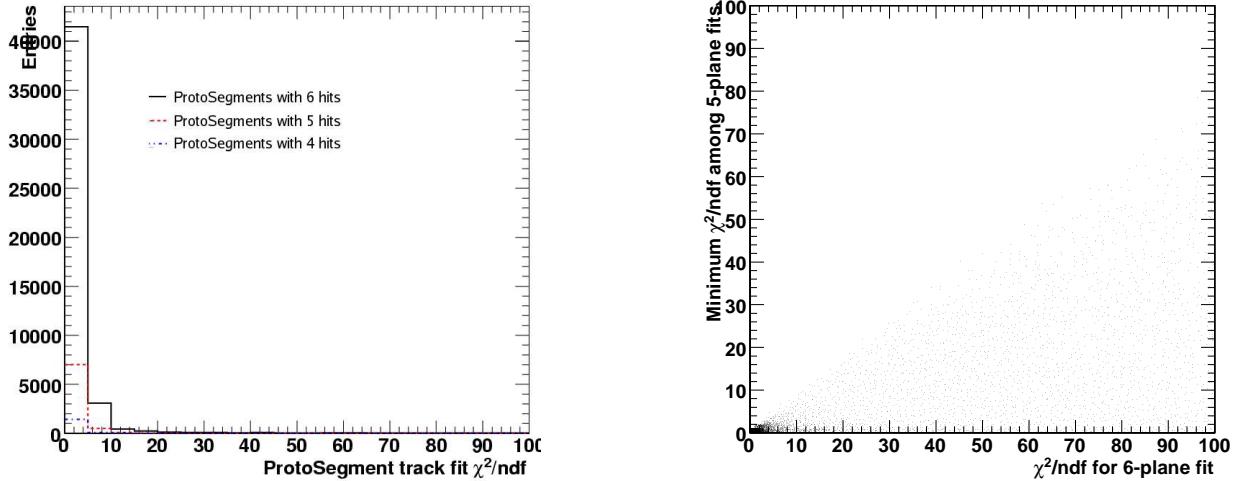


Figure 6-9: **Left:** The χ^2/dof distributions for a linear fit for ME23/2 chamber ProtoSegments with six clusters (solid curve), five clusters (dashed curve), and four clusters (dashed-dotted curve). **Right:** Scatter plot of $\min(\chi_5^2/\text{dof})$ vs. χ_6^2/dof for six-cluster ME23/2 chamber ProtoSegments. Here, χ_6^2/dof is for a six-plane fit, $\min(\chi_5^2/\text{dof})$ is a minimum between 6 five-plane fits with one plane left out from the fit.

RefinedSegments. One can see an obvious reduction in tails in comparison to distributions shown in Fig. 6-3.

The final optimization of the χ^2 cuts can be done when a good reference for a whole segment is available (e.g., in detector Monte Carlo simulation, provided that Monte Carlo is shown to reproduce data at the adequate level of details). In principle, more sophisticated additional criteria for pruning can be further employed: e.g., one can take into account the charge cluster shape and number of anode hits associated with a RecHit. These were studied in the past and shown to have some, albeit very limited, discriminating power against “bad” hits. However, their use for the HLT purposes is hardly justifiable.

6.4 Conclusions

A new fast algorithm for reconstructing track segments in Cathode Strip Chambers is proposed. The algorithm was validated with the real cosmic ray data taken with 36 CSC chambers operated as a part of CMS-wide MTCC test program in the second half

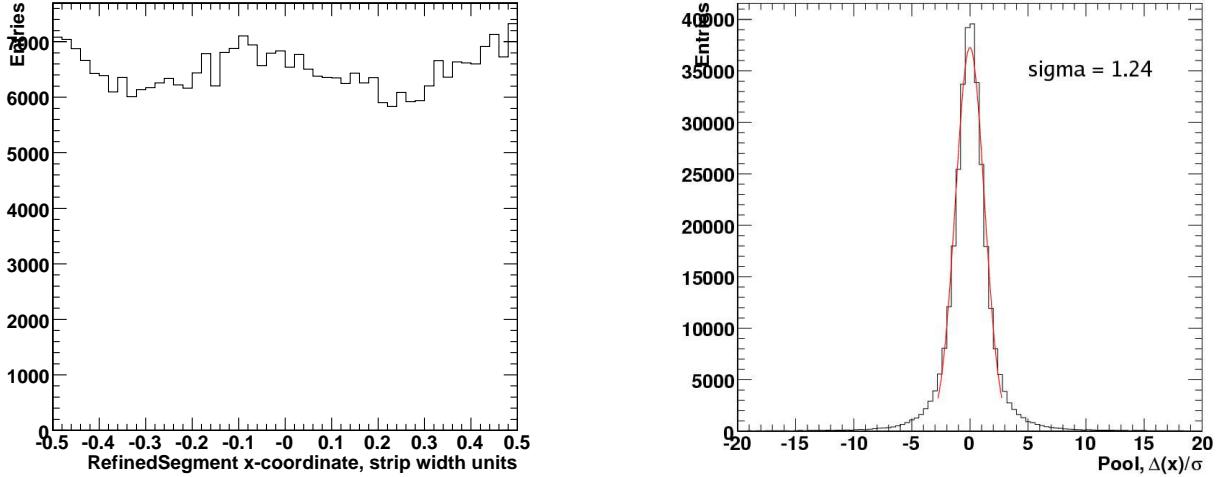


Figure 6-10: **Left:** Occupancy for $|x/w|$ -coordinates of RefinedSegments. **Right:** Pools for all retained RecHits (i.e., RecHits associated with RefinedSegments.)

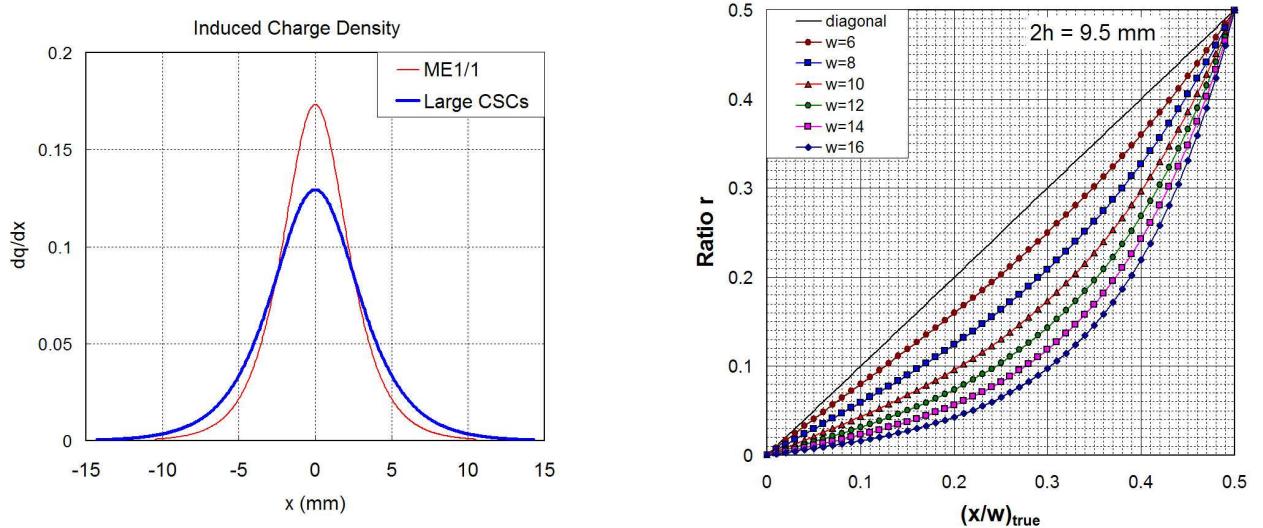


Figure 6-11: **Left:** Induced charge distribution calculated according to Gatti et al. [77] for large and ME1/1 chamber geometries. The Gatti parameter K_3 was taken to be 0.334 for large chambers and 0.379 for ME1/1 according to the empirical approximations in Ref. [78, 79]. **Right:** Ratio r versus a local coordinate x for large chambers calculated for variety of strip widths in the assumption of the Gatti charge distribution for large chambers (see plot on the left).

of 2006. The algorithm proved to provide high speed, high efficiency, and good spatial precision—all well within the High Level Trigger requirements.

6.5 Ratio Method

After charges Q_l , Q_c , Q_r on left, central, and right strips are measured, their ratio is built as follows:

$$r = \frac{1}{2} \frac{Q_r - Q_l}{Q_c - \min(Q_r, Q_l)} \quad (6-3)$$

Figure 6-11 (right) shows ratio r as a function of a local coordinate x calculated for the induced charged distribution according to Gatti et al. [77] (same Figure, left). The local coordinate x is assumed to be in strip units, $x=0$ corresponds to the strip center, and $x = \pm 0.5$ means right/left strip edges. This ratio r is a monotonic, but not linear function of a hit position across a strip x

We find the conversion function from r to x in two steps.

The first correction is an approximate inversion of the “theoretical” function $r(x, w)$. Figure 6-12 (left) shows a correction that one needs to add to r to obtain the coordinate x . The points correspond to the “theoretical” Gatti function. We find that the 1st-order correction can be parameterized quite well with the following empirical function (w is a strip width in cm):

$$g(r, w) = \frac{r(0.5 - r)}{a/w^b + c|r|}, \quad (6-4)$$

where $a=0.27$ (0.11), $b=2.7$ (2.9), $c=1.25$ (1.25) for large (and ME1/1) chambers.

After applying this correction, we get the 1st-order corrected coordinate $x_1 = r + g(r, w)$ is expected to be within 1% of the true coordinate x —see Fig. 6-12 (left). This is already sufficient for the HLT purposes. However, this correction is purely theoretical and must be checked against the reality.

Figure 6-12 (right) shows the experimental occupancy distribution dN/dx_1 . It has an obvious “wave”, which is a manifestation of the fact that the induced charge does not quite follow the “theoretical” Gatti. Curiously enough, the shape of the wave looks very similar for all parts of all chambers, i.e. it is very similar for different strip widths. This

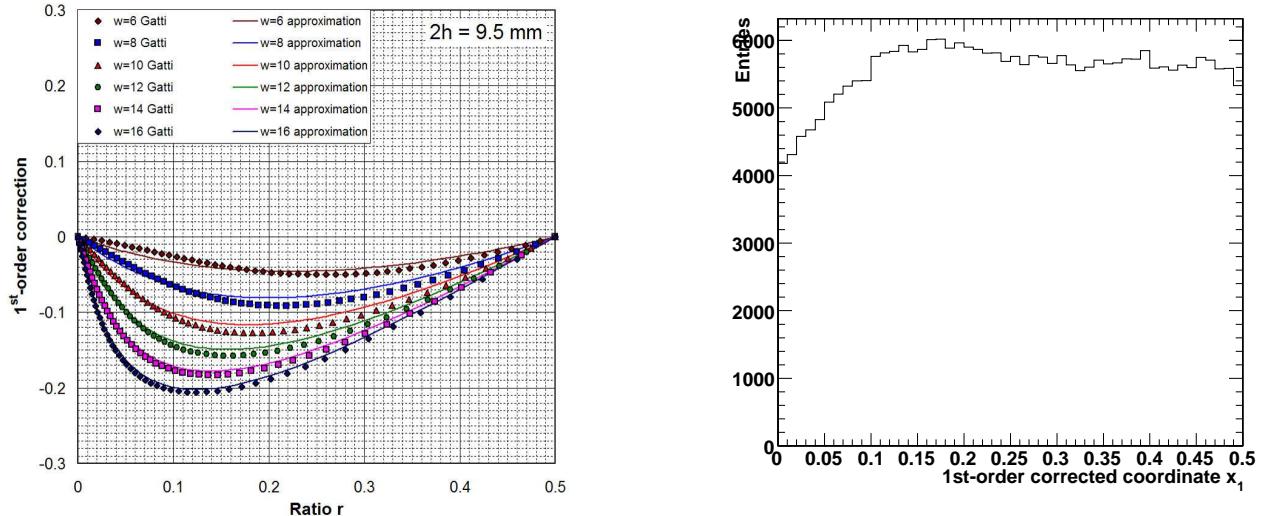


Figure 6-12: **Left:** Correction that must be added to the measured ratio r to obtain the hit position across a strip. Points correspond to the “theoretical” Gatti-based curve from Fig. 6-11, while the lines correspond to an empirical approximation described in the text as the 1st-order correction. **Right:** An occupancy distribution for the 1st-order corrected coordinate x^{1st} .

allows us to introduce a second-order empirical correction in a strip-width independent manner. This correction can be derived directly from the shape of the dN/dx_1 occupancy distribution and can be parameterized as follows:

$$x_2 = x_1 + c_1 e^{b_1 x_1} + c_2 e^{b_2 x_1} + c_0, \quad (6-5)$$

where $b_1 = -11$, $b_2 = -6.5$, $c_1 = 0.070273$, $c_2 = -0.072769$, and $c_0 = -(c_1 + c_2)$. The second-order correction $x_2 - x_1$ is actually quite modest (see Fig. 6-13). However, it does make the occupancy plots for x_2 almost perfectly flat.

The sensitivity of the ratio method to the typical electronics noise, calibration errors, and cross-talk uncertainties are shown in Figs. 6-14, 6-15, 6-16. The typical values used to make these plots are the results of pre-installation testing of all 396 large chambers at so-called Final Assembly and System Tests sites [76].

The typical noise is $\sim 1\%$ of the average cathode cluster charge, when chambers operate at the nominal gas gain. In addition to the FAST site measurements, the noise

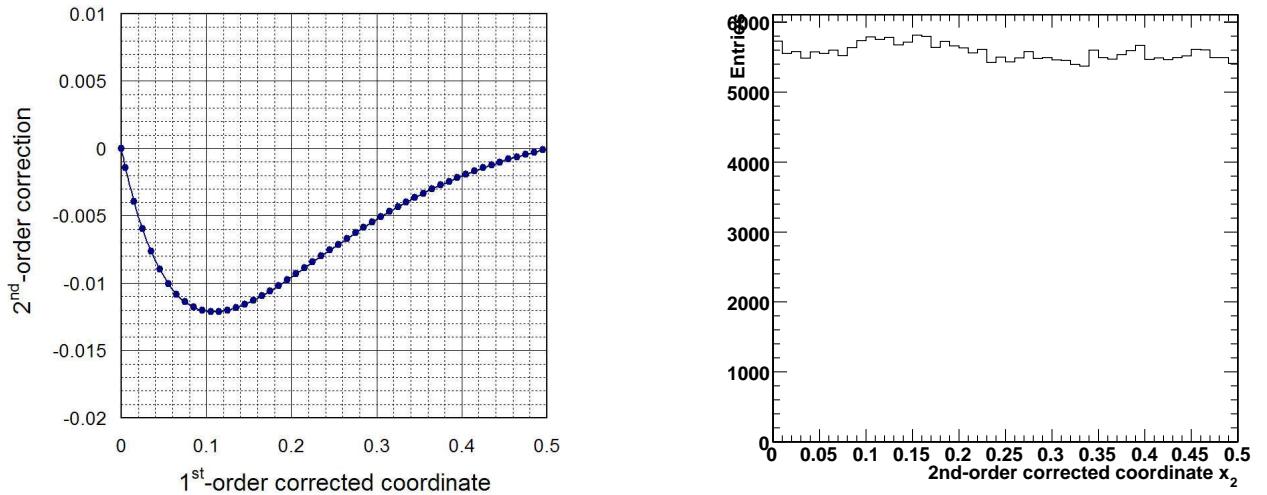


Figure 6-13: **Left:** The second-order correction that must be added to x_1 to obtain the hit position across a strip x_2 **Right:** An occupancy distribution for the 2nd-order corrected coordinate x_2 (as expected, the distribution now is almost perfectly flat).

levels were again measured in situ during data taking for all installed chambers and found to remain unchanged. If one does not use calibration constants, the spread of electronic channel gains between nearby strips is found to be $\sim 1\%$. The differences in crosstalks between strips for chambers of the same type are very small; the 2% number used for making Fig. 6-16 represents the RMS value for all crosstalks collected in one histogram regardless of the chamber type.

One can see that channel gain calibrations and crosstalks are the least of our concerns. The noise contribution to the resolution between strips is also very small. These figures allow one to gauge how electronics performance may contribute to the CSC spatial resolution when the ratio method is used.

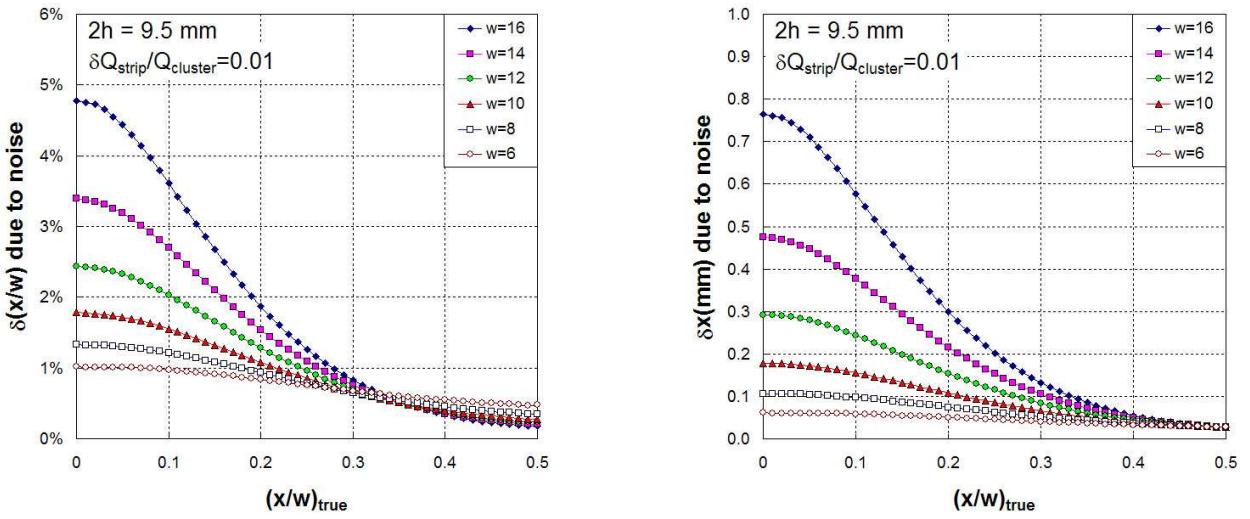


Figure 6-14: Sensitivity of the ratio method to electronic noise. The curves are given for $\delta Q_{strip}/Q_{cluster} = 0.01$, which corresponds to the nominal operation conditions. Pedestals are assumed to be tabulated. When pedestals are defined on the event-by-event basis, the noise contribution increases by a factor of 1.6. The resolution is given in units of strip width (left) and in microns (right)

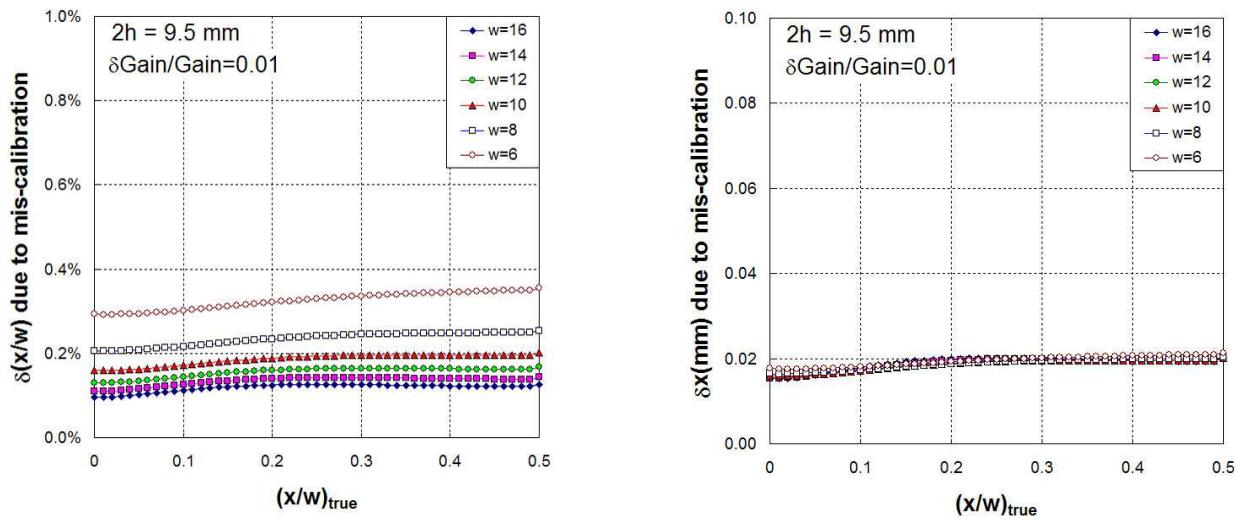


Figure 6-15: Sensitivity of the ratio method to errors in electronic gain calibrations. The curves are given for $\delta Gain/Gain = 0.01$. The choice of this number is driven by the actual measurements of strip electronic gain variations between nearby strips (this is the only thing that matters), which gave 1.1%. The resolution is given in units of strip width (left) and in microns (right).

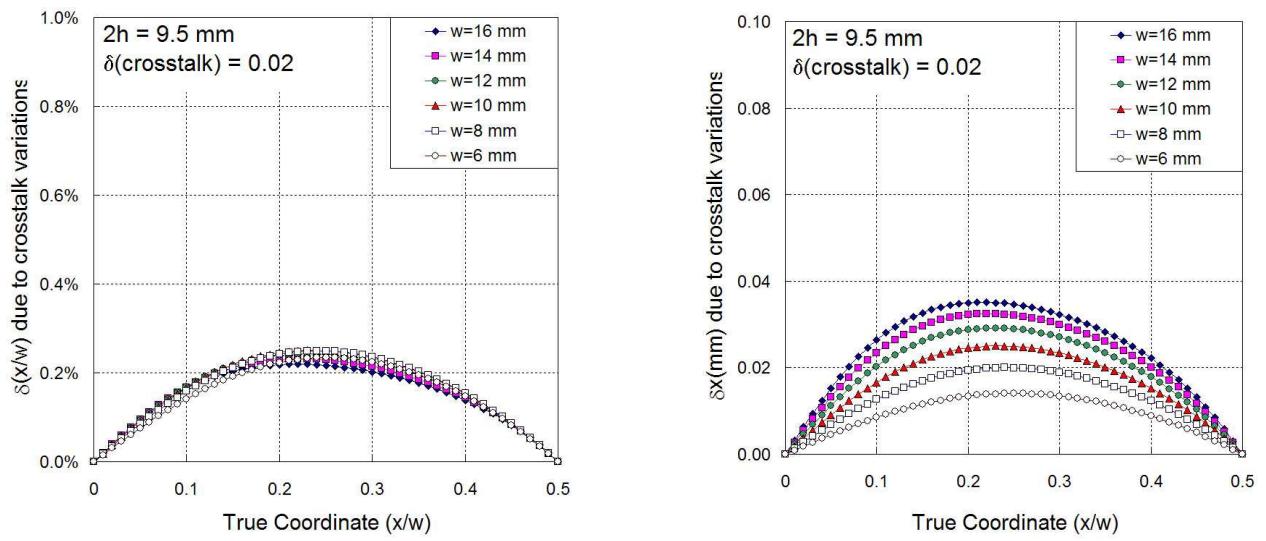


Figure 6-16: **Left:** Sensitivity of the ratio method to uncertainties in cross talks between strips. The curves are obtained for a 2% uncertainty in crosstalk, which corresponds closely to variations in crosstalk values between different chamber types (sizes). The resolution is given in units of strip width (left) and in microns (right).

CHAPTER 7

SEARCH STRATEGY FOR THE STANDARD MODEL HIGGS BOSON IN THE $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ DECAY CHANNEL USING $M(4\mu)$ -DEPENDENT CUTS.

7.1 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 chapter, 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) [80, 81, 82] 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 will 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 full CMS detector simulation and reconstruction software [83, 84] and include pile-up events corresponding to an instantaneous luminosity of $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$.

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 [85, 86, 87]. Another ongoing study exploring the discovery potential with a different set of mass-independent cuts can be found elsewhere [88]. The results of the two parallel analyses using the $H \rightarrow 4e$ and $H \rightarrow 2e2\mu$ channels can be found in [89, 90].

In short, we believe our $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ analysis is more realistic than previously existed ones on the topic:

- we used calibration from data techniques, as we would do with real data;
- full treatment of systematic errors was included and folded into signal-to-background significance calculations;
- dedicated effort was done to generate physics of signal and background processes properly, including use of dedicated matrix element generators, events re-weighting with dynamic Next-to-Leading Order (NLO) corrections;
- latest available at the moment full simulation and reconstruction validated software was used;
- CMS performance was optimized for all allowed SM Higgs boson masses;

- and as a result our analysis is part of the CMS Physics Technical Design Report [7, 8], i.e. is an official CMS strategy for the SM Higgs boson discovery in $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel.

Details of our analysis and supporting studies can be found in a list of refereed papers [1, 2, 3, 4, 5, 6, 7, 8, 9, 10], in all of which the author of this thesis was either the leading author or one of the leading co-authors. Results of our work presented at 10 international conferences [11, 12, 13, 14, 15, 16, 17, 18, 19, 20]; the author also gave more than 70 talks at the internal CMS collaboration meetings of different levels [21].

7.2 Physics Processes and Their Simulation

The Higgs boson event samples for 18 Higgs boson mass points (see Table 7-1) and the three main background processes, $t\bar{t}$, $(Z/\gamma^*)bb$, 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 7-1. The $M(4\mu)$ distribution of events after these cuts is shown in Figures 7-1 and 7-2.

Table 7-1: 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 7.2) calculated for $L = 30\text{fb}^{-1}$, and the number of simulated events.

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

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

The Higgs boson samples were generated with Pythia 6.225 [91] (LO gluon and weak-boson fusion), interfaced via CMKIN [92] 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 analysis cuts. QED radiation from the final-state muons is modeled with PHOTOS [93, 94]. Events were re-weighted to correspond to the total NLO cross-section $\sigma(pp \rightarrow H) \cdot BR(H \rightarrow ZZ) \cdot BR(Z \rightarrow 2l)^2$, where $\sigma(pp \rightarrow H)$ and $BR(H \rightarrow ZZ)$ were taken from [95, 96] and $BR(Z \rightarrow 2l) = 0.101$ [97].

There exists an additional enhancement to the cross section for $H \rightarrow 4\mu$ (in comparison to $\sigma(pp \rightarrow H) \cdot BR(H \rightarrow ZZ) \cdot BR(Z \rightarrow 2\mu)^2$) due to interference of permutations of muons originating from different Z's. The corrections are calculated with CompHEP and presented in Figure 7-3.

The $M(4\mu)$ distribution for $m_H = 140$ GeV/ c^2 after event generator-level cuts is shown in Figures 7-1 and 7-2. 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.

7.2.2 Background: $t\bar{t}$

The $t\bar{t}$ sample was generated with Pythia 6.225 (LO $gg \rightarrow t\bar{t}$ and $q\bar{q} \rightarrow t\bar{t}$), interfaced via CMKIN version 110 (PDF CTEQ5L). Only the decay channels $t \rightarrow Wb \rightarrow l\nu b$ were considered. Events were re-weighted to correspond to the total NLO cross-section $\sigma(pp \rightarrow t\bar{t}) \cdot BR(W \rightarrow l\nu)^2$, where $\sigma(pp \rightarrow t\bar{t}) = 840$ pb was taken from [98] and the branching ratio $BR(W \rightarrow l\nu) = 0.320$ [97].

7.2.3 Background: $(Z/\gamma^*)b\bar{b} \rightarrow 2\mu b\bar{b}$

The $(Z/\gamma^*)b\bar{b} \rightarrow \mu^+\mu^-b\bar{b}$ samples were generated with CompHEP 4.2p1 [99] 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 $q\bar{q}/gg \rightarrow Z/\gamma^*b\bar{b} \rightarrow \mu^+\mu^-b\bar{b}$, 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 7-1, we calculated the NLO K-factor using MCFM [100]: $K_{NLO} = 2.4 \pm 0.3$. 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 .

7.2.4 Backgrounds: $q\bar{q} \rightarrow Z/\gamma^* Z/\gamma^* \rightarrow 4\mu$ and $q\bar{q} \rightarrow Z/\gamma^* Z/\gamma^* \rightarrow 2\mu 2\tau$

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 $p_T > 3 \text{ GeV}/c$ and $|\eta(\mu)| < 2.5$. The direct τ 's from Z/γ^* decays were required to have $p_T > 3 \text{ GeV}/c$ 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 $120 < M(4\mu) < 180 \text{ GeV}/c^2$, and can be safely neglected for higher 4μ invariant masses; see [2] 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 ($gg \rightarrow ZZ$, known to contribute $\approx 20\%$ with respect to the LO [101]), 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 [100] and is shown in Figure 7-4. All details on calculation of this $M(4\mu)$ -dependent K-factor and the dynamical differences between NLO and LO are summarized elsewhere [6].

The $M(4\mu)$ distributions after generator-level cuts are shown in Figures 7-1 and 7-2. The peak at M_Z is due to the s-channel contribution. This peak sits on the shoulder of the enhancement around $M(4\mu) > 100 \text{ GeV}/c^2$ —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.

7.3 Trigger and Offline Muon Reconstruction

Muons have a very clean detection signature resulting in a high trigger efficiency [7]. The inclusive muon triggers based on the selection of a single muon with $p_T > 19 \text{ GeV}/c$

or di-muons with $p_T > 7 \text{ GeV}/c$ 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 $p_T > 7 \text{ GeV}/c$ if they are in the central pseudorapidity region ($|\eta| < 1.1$), or with total momentum $p > 13 \text{ GeV}/c$ if they are in the endcaps ($|\eta| > 1.1$). Figures 7-5 and 7-6 show the efficiency turn-on curves - the choice of these cuts is obvious from the figures. These cuts do not dramatically affect the number of accepted signal events.

We require that all four possible combinations of reconstructed di-muon masses satisfy $m(\mu^+\mu^-) > 12 \text{ GeV}/c^2$. 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 7-7 shows such a distribution for $M_H=150 \text{ GeV}/c^2$. A Gaussian fit of the peak gives $\sigma=1.1 \text{ GeV}/c^2$. One can see a noticeable tail toward smaller masses - mostly due to internal bremsstrahlung and events with intermediate τ -leptons ($ZZ \rightarrow 2\tau 2\mu \rightarrow 4\mu 4\nu$ and $ZZ \rightarrow 4\tau \rightarrow 4\mu 8\nu$). The four-muon mass detector resolution $\delta M(4\mu)$ as a function of $M(4\mu)$ is given in Figure 7-8.

7.4 Higgs Search Strategy

The strategy for searching the Higgs boson involves successive steps:

- 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 7.4.2. 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 the future, more sophisticated statistical tools can and will be added.

- Third, the most important theoretical and instrumental systematic errors are evaluated (section 7.4.3.1). 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 7.4.3.2). 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.
- 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 7.4.4, 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.

7.4.1 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: $Q = P(\text{observables}|s + b)/P(\text{observables}|b)$. 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 [80].

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

$$-2\ln(Q) = 2s - 2n \ln \left(1 + \frac{s}{b} \right), \quad (7-1)$$

for the entire binned spectrum:

$$-2\ln(Q) = 2S - 2 \sum_{bins}^N n_i \ln \left(1 + \frac{s_i}{b_i} \right), \quad (7-2)$$

or the unbinned spectrum:

$$-2\ln(Q) = 2S - 2 \sum_{events}^N \ln \left(1 + \frac{pdf_s(m_k)}{pdf_b(m_k)} \right). \quad (7-3)$$

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 :

$$P = \int_S^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx. \quad (7-4)$$

Figures 7-9 and 7-10 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 7-9 and 7-10, 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, $-2\ln(Q)$ 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

$$1 - CL_b = P = p(n \geq n_o | b) = \sum_{k=n_o}^{+\infty} \frac{b^k}{k!} e^{-b}. \quad (7-5)$$

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. 7–4 to define the true counting experiment significance S_c . Including systematic errors into evaluation of significance is discussed in Section 7.4.3.2.

7.4.2 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 7.3 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 analysis, 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 GARCON¹ [102]. 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 verification on $\approx 3 \cdot 10^5$ events. More details on GARCON are available on its web-page and in a dedicated paper [102]. This program, being relatively new, is already widely used for other CMS analyses [103, 104, 105].

- 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. 7-11 and 7-12). 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 $t\bar{t}$ increases (Fig. 7-1).
- p_T cut on the second lowest p_T muon, or equivalently one common cut on the three highest p_T muons (Fig. 7-13). 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. 7-14). 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 7-11- 7-14. The results of these comparisons are shown in Figure 7-19. 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 7-11-7-14: 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 7-20 shows the $M(4\mu)$ invariant mass distribution for the three background subprocesses and a Higgs boson signal at $M_H = 150 \text{ GeV}/c^2$ 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 7-15-7-18. 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 7-21 and 7-22 show the significances S_{cL} and S_L at $L = 30 \text{ fb}^{-1}$ 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 $M_H = 150 \text{ GeV}/c^2$, 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 $M_H = 500 \text{ GeV}/c^2$ is right at the 5σ -discovery threshold for an integrated luminosity $L = 30 \text{ fb}^{-1}$ (Fig. 7-21). 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 $M_H = 500 \text{ GeV}/c^2$ at $L = 30 \text{ fb}^{-1}$ are $< 40\%$ for the flat cuts and $> 60\%$ for the $M(4\mu)$ -dependent cuts.

Figure 7-23 shows the same results as in Figure 7-21 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 7-24 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.

7.4.3 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.

7.4.3.1 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 $ZZ \rightarrow 4\mu$ 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 $Z \rightarrow 2\mu$ 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. Particular examples are given below in overview and all the details are available in referenced below papers we published. In general, selection criteria for $Z \rightarrow 2\mu$ events were selected as close as possible to selection cuts described above in this analysis and control samples for ZZ – are events on the left and right side of the signal window.

The PDF and QCD scale uncertainties in the $ZZ \rightarrow 4\mu$ production cross section were studied at the NLO level using MCFM [100]. 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. [5]. Figure 7-25 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 $Z \rightarrow 2\mu$ cross section (note that the $Z \rightarrow 2\mu$ 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 $ZZ \rightarrow 4\mu$ process was evaluated with two very different programs: MCFM [100] and EffNLO [6]. The latter is a package smoothly splicing together MadGraph [106] (NLO $pp \rightarrow 4\mu + jet$) and Pythia (LO $pp \rightarrow 4\mu + ISR-jets$). The relative difference in $K(M_{4\mu})$ is shown in Figure 7-26. 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): $gg \rightarrow ZZ \rightarrow 4\mu$ (box-diagram), contributing about $(20 \pm 8)\%$ to the LO cross section [101] (note that this contribution was calculated without virtual photons in the propagators) and $qq \rightarrow 4\mu + qq$ 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 luminosity of 1 fb^{-1} , 5% at $L = 10 \text{ fb}^{-1}$, 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.

The 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 than 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 $p_T > 19 \text{ GeV}/c$. This sample will contain inclusive W, Z, and other processes in the approximate ratio W:Z:others = 10:1:small [107]. By counting the number of $Z \rightarrow 2\mu$ 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 [3], see also Sec. C. The four-muon efficiency therefore will be known with an absolute error of better than 4%. When we deduce the expected $ZZ \rightarrow 4\mu$ events from the measured $Z \rightarrow 2\mu$ 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 $Z \rightarrow 2\mu$ 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% [4] (Fig. 7-27).

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. 7-28). 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 $Z \rightarrow 2\mu$ events.

Figures 7-29 and 7-30 summarize all the systematic errors on the expected number of events in the $ZZ \rightarrow 4\mu$ background for the two methods: via referencing to the total

measured $Z \rightarrow 2\mu$ cross section and via referencing to the event count in the sidebands of the $M(4\mu)$ spectrum itself.

7.4.3.2 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

$$P = p(n \geq n_o | b) * f(b) = \int_0^{+\infty} p(n \geq n_o | b) f(b) db, \quad (7-6)$$

which can be again converted into true significance S_c using Eq. 7-4.

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 $\delta = \Delta b / b_0$:

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

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. 7-7 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:

$$f(b) = \frac{1}{\rho} \frac{(b/\rho)^B e^{-b/\rho}}{\Gamma(B+1)}. \quad (7-8)$$

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 7-31 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 $Z \rightarrow 2\mu$ cross section. All three curves correspond to the total integrated luminosity of 30 fb^{-1} . Figure 7-32 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 $Z \rightarrow 2\mu$ process and the $ZZ \rightarrow 4\mu$ sidebands, is made in terms of the luminosity required for 5σ -discovery (Fig. 7-33). The difference is not dramatic. The true benefit of using two approaches to estimating background from data is in their complementarity.

Finally, Table 7-2 summarizes the most important results for the $M(4\mu)$ -dependent cuts that we presented in this chapter.

7.4.4 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. 7.6 for all details) and one must exercise a caution in evaluating probabilistic interpretation

Table 7-2: 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 is normalized to the $Z \rightarrow 2\mu$ 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

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 ($\approx 15\%$ 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.

7.5 Summary

Discovery of the Standard Model Higgs boson in the “gold-plated” decay mode $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ 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 $\approx 2 \text{ fb}^{-1}$ of integrated luminosity, we would be able to start excluding the SM Higgs boson at 95% CL for M_H in vicinity of $200 \text{ GeV}/c^2$. By the time we reach $\approx 30 \text{ fb}^{-1}$, we would exclude the Standard Model Higgs in its four-muon decay mode in the mass range $M_H = 120 - 600 \text{ GeV}/c^2$, if indeed it does not exist.

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

7.6 On the true significance of a local excess of events

In searching for new phenomena in a wide range of possible signal hypotheses (e.g., a narrow resonance of unknown mass over a broad range background), special care must be exercised in evaluating the true significance of observing a local excess of events. In past, this fact was given substantial scrutiny by statisticians (e.g., [108, 109]) and physicists alike (e.g., [110, 111, 112, 113, 114]). The purpose of this Appendix is to quantify a

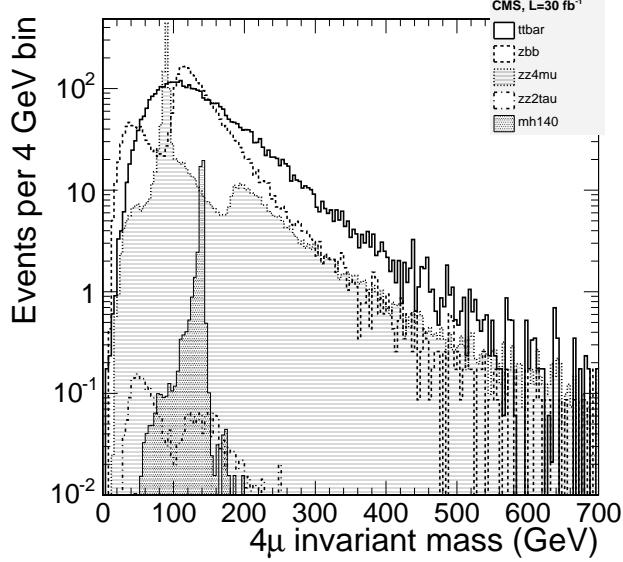


Figure 7-1: $M(4\mu)$ distributions after generator-level cuts for $t\bar{t}$, $(Z/\gamma^*)b\bar{b}$, $Z/\gamma^*Z/\gamma^*$, and $m_H = 140 \text{ GeV}/c^2$ (log scale).

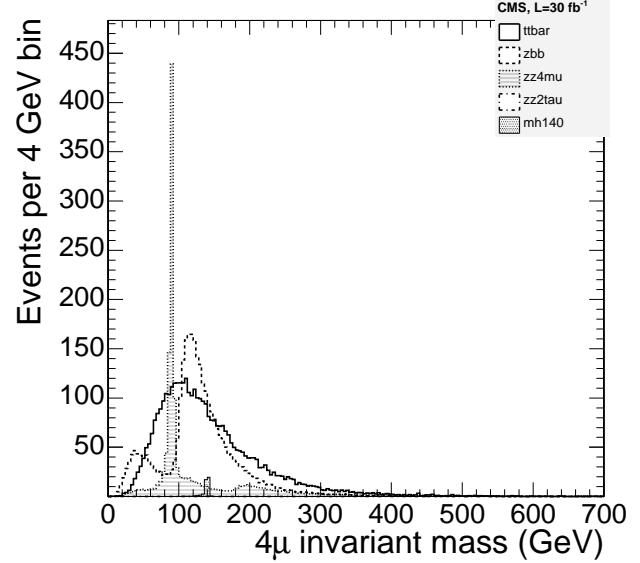


Figure 7-2: Same as Figure 7-1, but on a linear scale.

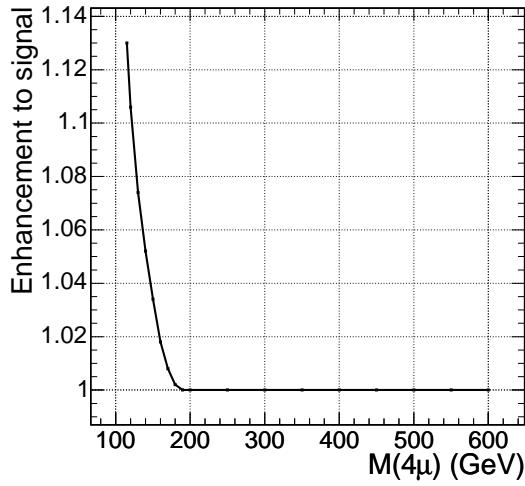


Figure 7-3: Enhancement to the signal samples' cross sections due to interference effects not accounted for at the generator level.

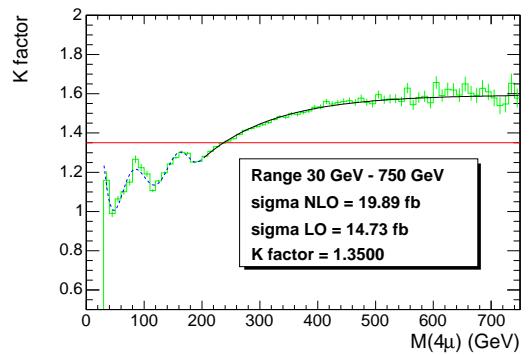


Figure 7-4: The $M(4\mu)$ -dependent NLO K-factor $K_{NLO}(M_{4\mu})$ for the $ZZ \rightarrow 4\mu$ process evaluated with MCFM [100].

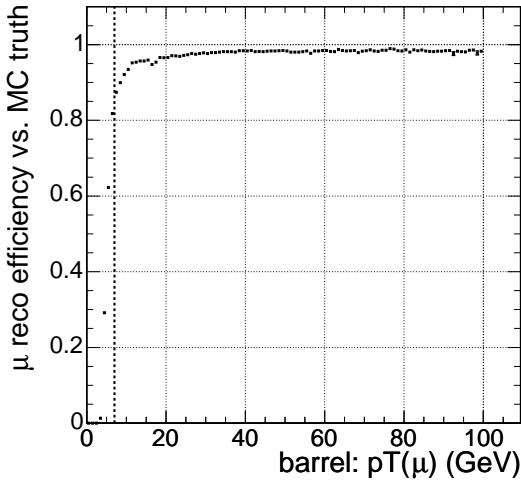


Figure 7-5: Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the barrel region vs. p_T .

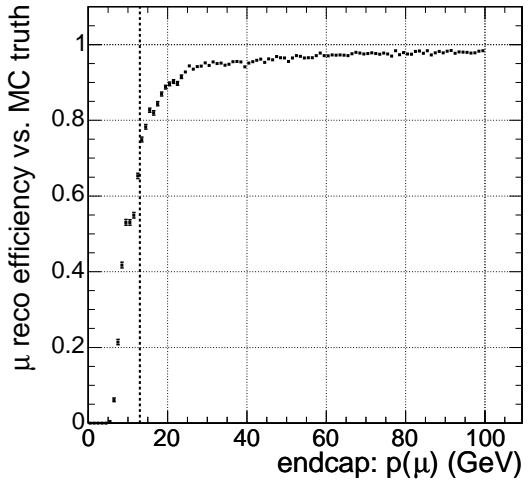


Figure 7-6: Global muon reconstruction efficiency calculated from matching reconstructed and true Monte Carlo muons in the endcap region vs. momentum.

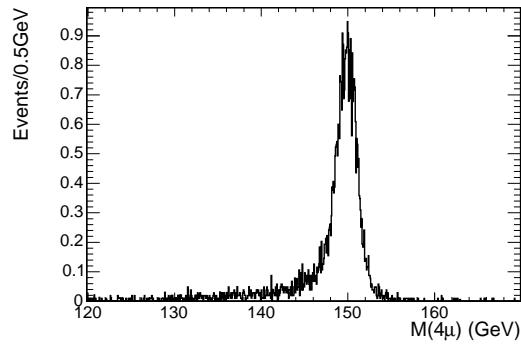


Figure 7-7: $M(4\mu)$ distribution for $m(H)=150 \text{ GeV}/c^2$ and the fit described in the text.

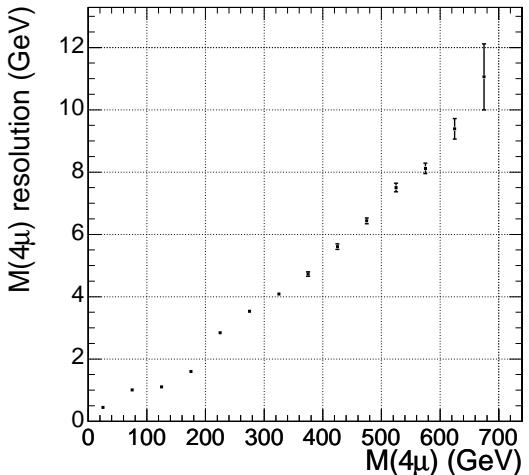


Figure 7-8: $M(4\mu)$ resolution vs. M_H .

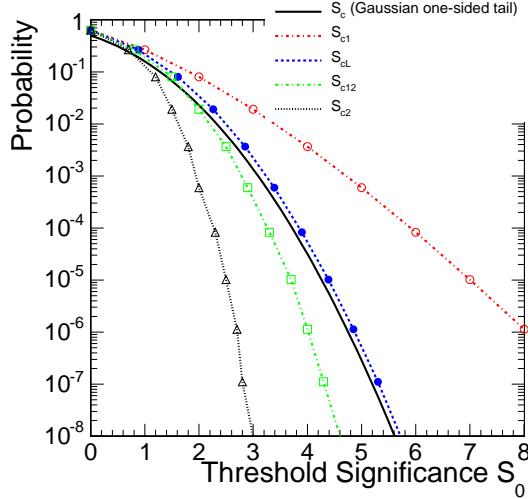


Figure 7-9: Comparison of different significance estimators for 1 background event: probability of measuring significance $S > S_0$, background only case, $N_B = 1$ event. Equations for different significance estimators are given in the text.

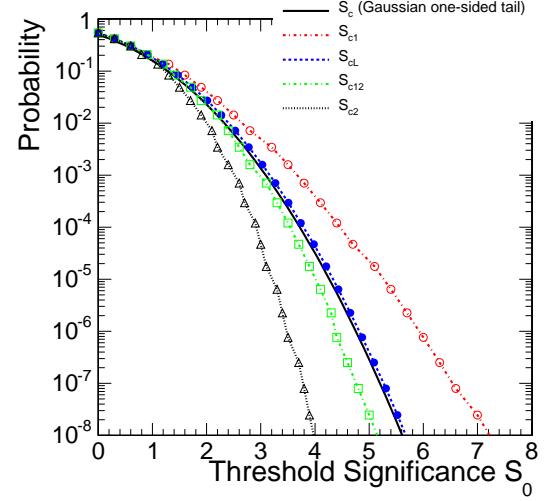


Figure 7-10: Same as Figure 7-9, but for $N_B = 10$ events.

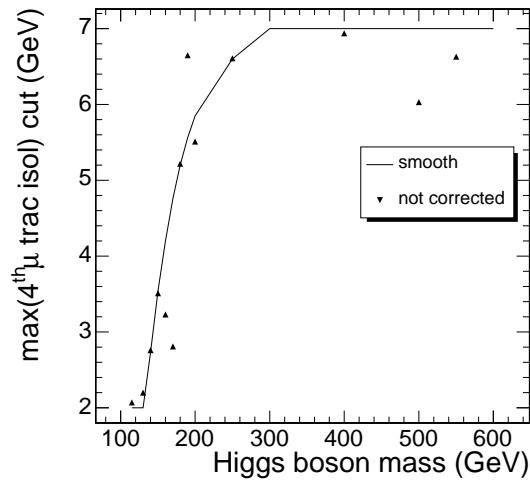


Figure 7-11: Dependence of the tracker-based muon isolation cut on the least isolated muon versus Higgs mass. The smooth dependence shown by the curve which follows general dependence of the cut variable was used for analysis-level cuts.

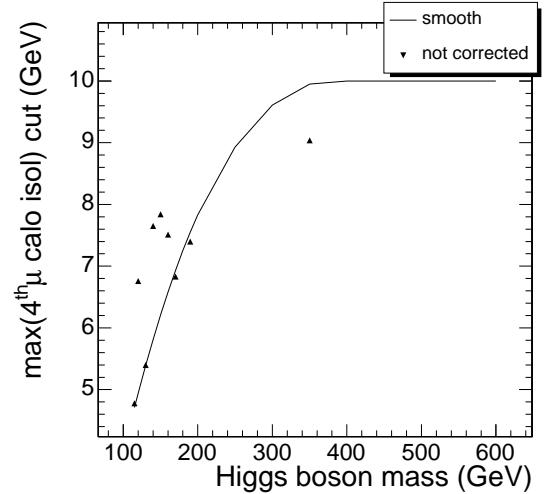


Figure 7-12: Dependence of the calorimeter-based cut on the least isolated muon versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.

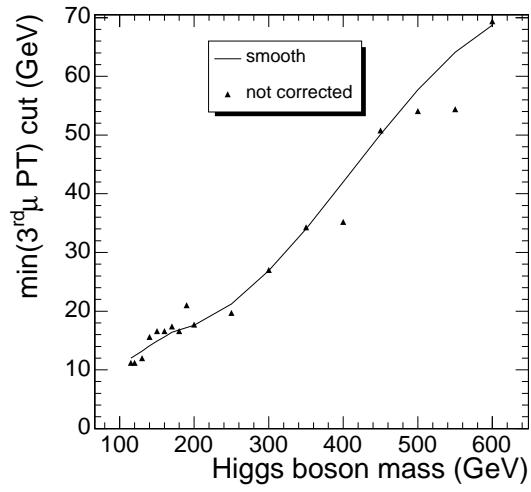


Figure 7-13: Dependence of the p_T cut on the second-lowest- p_T muon versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.

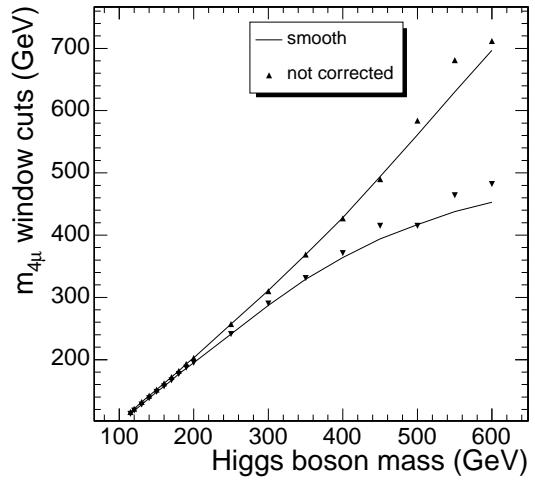


Figure 7-14: Dependence of the $M(4\mu)$ window cuts versus Higgs mass. The smooth curve has the same meaning as for Figure 7-11.

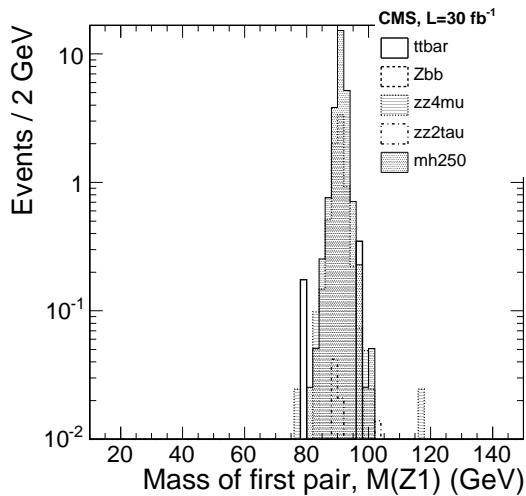


Figure 7-15: First muon pair invariant mass distribution, $M(Z_1)$ (invariant mass of two opposite sign muons closest to Z^0 -mass), after analysis cuts were applied.

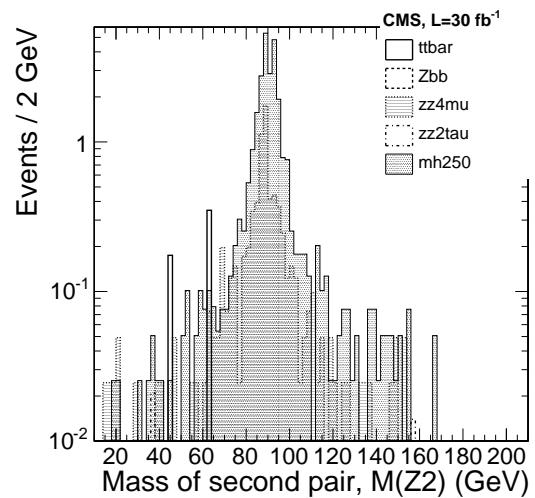


Figure 7-16: Second muon pair invariant mass distribution, $M(Z_2)$ (invariant mass of two other highest p_T opposite sign muons after muons for Z_1 selected), after analysis cuts were applied.

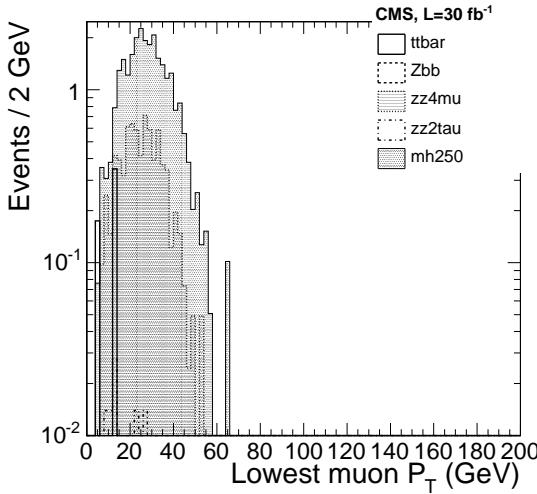


Figure 7-17: Lowest muon p_T distribution, after analysis cuts were applied.

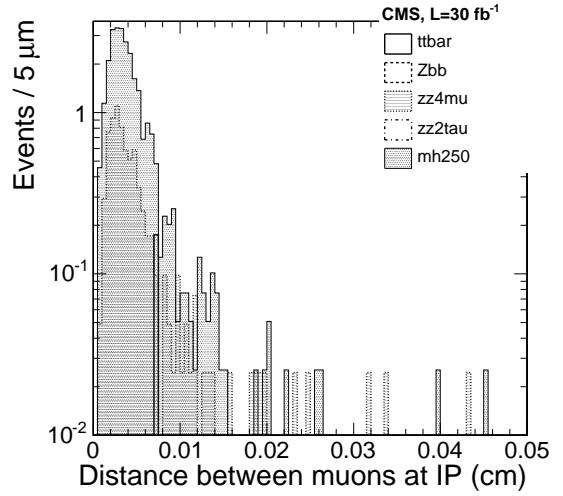


Figure 7-18: Maximum distance in XY-plane between muon impact point coordinates distribution, after analysis cuts were applied.

possible scope of this effect on an example of a search for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel. As the case study, we chose a counting experiment approach widely used in this volume.

The dashed line in Fig. 7-34 shows the expected 4μ invariant mass distribution for background at $L = 30 fb^{-1}$ after applying all the $m_{4\mu}$ -dependent analysis cuts. Using this distribution, we played out $\sim 10^8$ pseudo-experiments; an example is shown in Fig. 7-34. For each pseudo-experiment, we slid a signal region window across the spectrum looking for a local event excess over the expectation. The size of the window $\Delta m = w(m_{4\mu})$ was optimised and fixed a priori (about $\pm 2\sigma$) to give close to the best significance for a resonance with a width corresponding to the experimental SM Higgs boson width $\sigma(m_{4\mu})$. The step of probing different values of $m_{4\mu}$ was “infinitesimally” small ($0.05 GeV/c^2$) in comparison to the Higgs boson width of more than $1 GeV/c^2$. The scanning was performed in a priori defined range of $115-600 GeV/c^2$.

We used a significance estimator $S_{cL} = sign(s) \sqrt{2 n_o \ln(1 + s/b) - 2s}$, where b is the expected number of background events, n_o is the number of observed events, and the

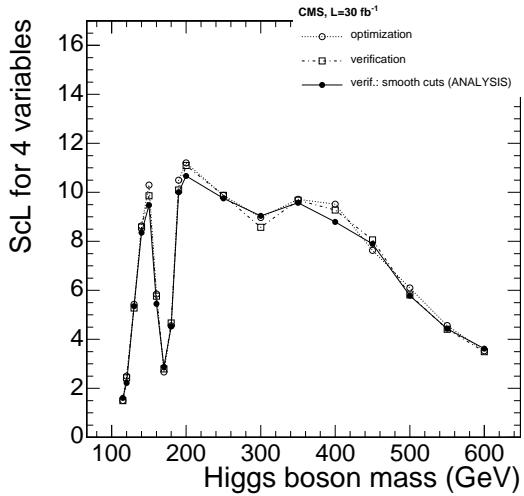


Figure 7-19: S_{cL} vs. Higgs boson mass: optimized cut values and GARCON optimization step when first half of the simulated data are used (empty circles markers and dotted curve); optimized cut values and GARCON verification step when second half of the simulated data are used (empty squares markers 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).

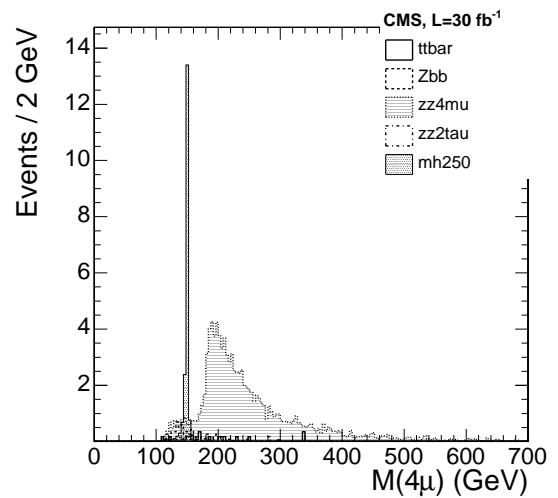


Figure 7-20: $M(4\mu)$ invariant mass distribution for the three background subprocesses and a Higgs boson signal at $M_H = 150 \text{ GeV}/c^2$, after applying cuts on muon isolation and p_T .

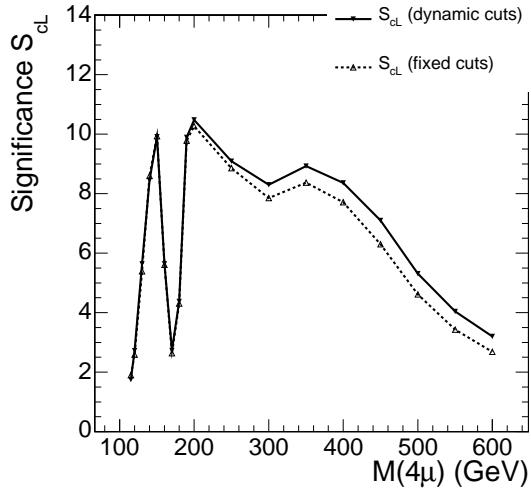


Figure 7-21: Expected excess significance S_{cL} with $L = 30 \text{ fb}^{-1}$ for different Higgs boson masses for $M(4\mu)$ -dependent (solid line) and independent cuts (dashed line). No systematic errors included.

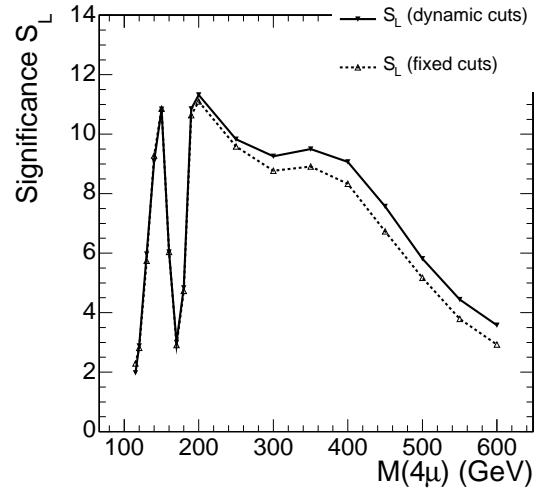


Figure 7-22: Expected excess significance S_L with $L = 30 \text{ fb}^{-1}$ for different Higgs boson masses for $M(4\mu)$ -dependent (solid line) and independent cuts (dashed line). No systematic errors included.

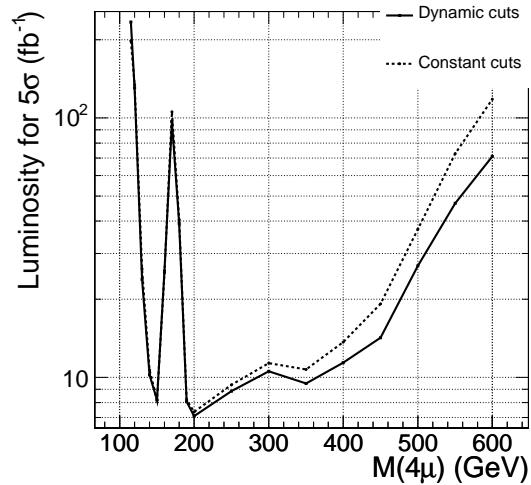


Figure 7-23: Luminosity required to reach a 5σ event excess for different Higgs boson masses for $M(4\mu)$ -dependent (solid line) and independent cuts (dashed line). No systematic errors included.

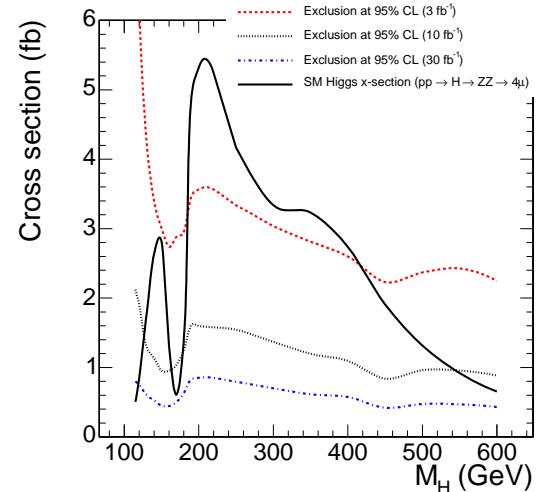


Figure 7-24: The 95% CL exclusion contours for the SM Higgs hypothesis in the (cross section; mass) plane at integrated luminosities of 3 (upper dashed line), 10 (middle dashed line), and 30 (lower dotted-dashed line) fb^{-1} . The Standard Model (SM) Higgs boson cross section for $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ is also shown (solid black line).

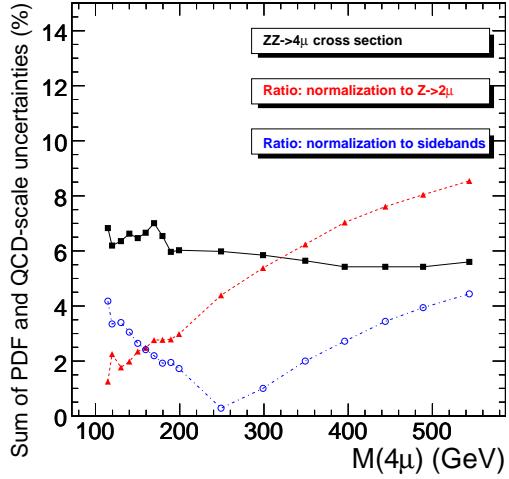


Figure 7-25: Combined systematic error on the number of background events due to PDF and QCD scale uncertainties for the $\sigma(q\bar{q} \rightarrow ZZ \rightarrow 4\mu)$ process at NLO. Shown are: absolute cross section uncertainties (black squares); uncertainties relative to the experimentally measured $q\bar{q} \rightarrow Z \rightarrow 2\mu$ 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).

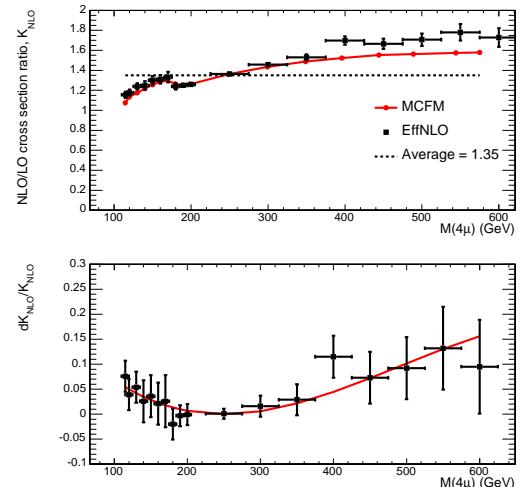


Figure 7-26: Top: the factors $K_{NLO}(M_{4\mu})$ in MCFM and EffNLO calculations versus $M(4\mu)$; bottom: the difference between them.

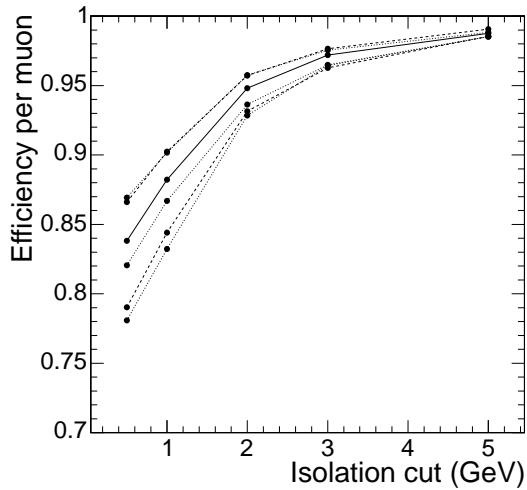


Figure 7-27: Muon isolation cut efficiency for random cone direction for Z -inclusive (dashed lines) and for ZZ (solid lines) events. The middle lines are for the default Pythia multi-parton interactions (MPI). The upper and lower lines correspond to uncertainties in modeling the MPI physics [4].

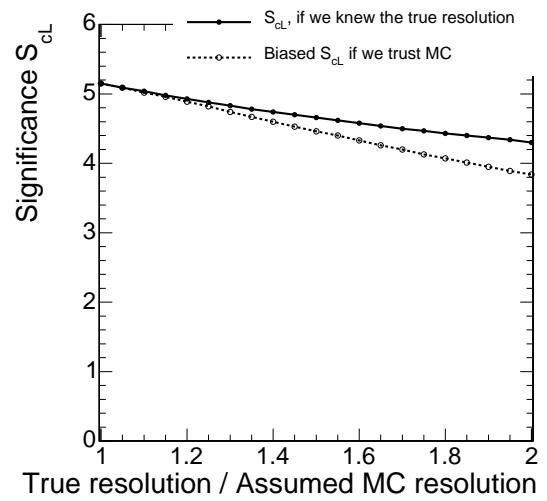


Figure 7-28: An example of a possible bias in evaluating the significance of an event excess due to a non-optimal choice of the signal window width for the case when the true detector resolution is worse than one assumes. The bias, being in the “conservative” direction, is very small even for errors in the $M(4\mu)$ resolution as large as 25% or worse

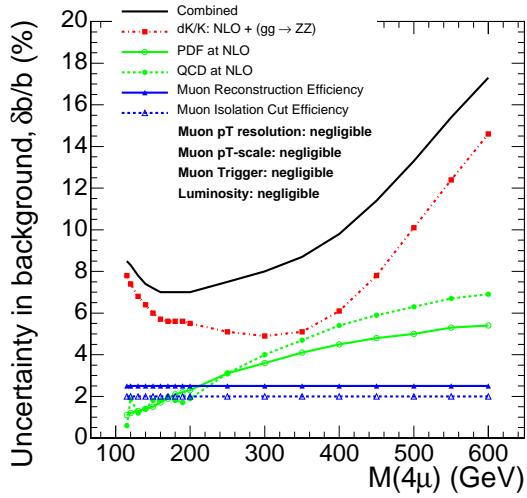


Figure 7-29: Uncertainties in the number of $ZZ \rightarrow 4\mu$ background events in the signal region window at different $M(4\mu)$. The window size is $\pm 2\sigma$ of the expected experimental Higgs resonance width. The event count is referenced to the number of $Z \rightarrow 2\mu$ events. Shown 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).

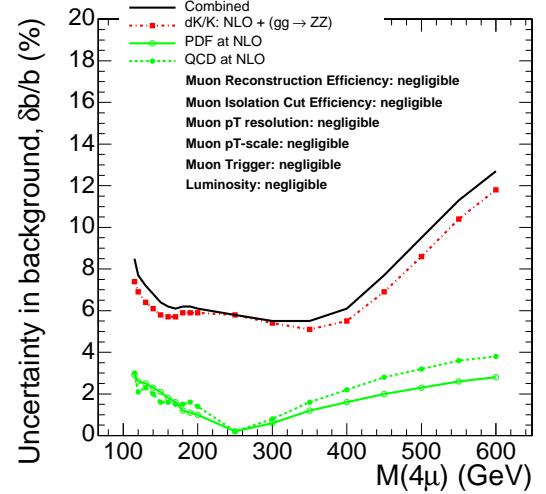


Figure 7-30: Uncertainties in the number of $ZZ \rightarrow 4\mu$ background events in the signal region window at different $M(4\mu)$. The window size is $\pm 2\sigma$ of the expected experimental Higgs resonance width. The event count in signal region, b , is calculated from the number of $ZZ \rightarrow 4\mu$ events in the range 100-700 GeV/c^2 (excluding the signal region window), B , via $b = \rho \cdot B$. (Same notations as Figure 7-29.)

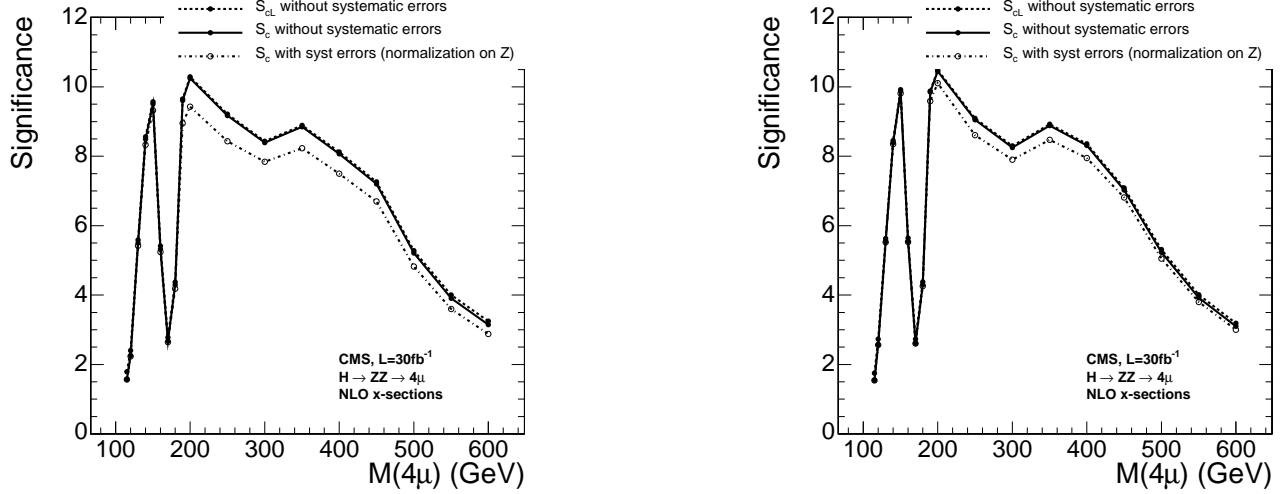


Figure 7-31: Significance vs. Higgs mass in the absence of any systematic errors (open circles, dashed line) (see Figs. 7-21 and 7-22) and significance that includes all uncertainties in the background when it is estimated via referencing to the measured $Z \rightarrow 2\mu$ cross section (closed circles). 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 7-29); for the right no dK/K contribution is considered.

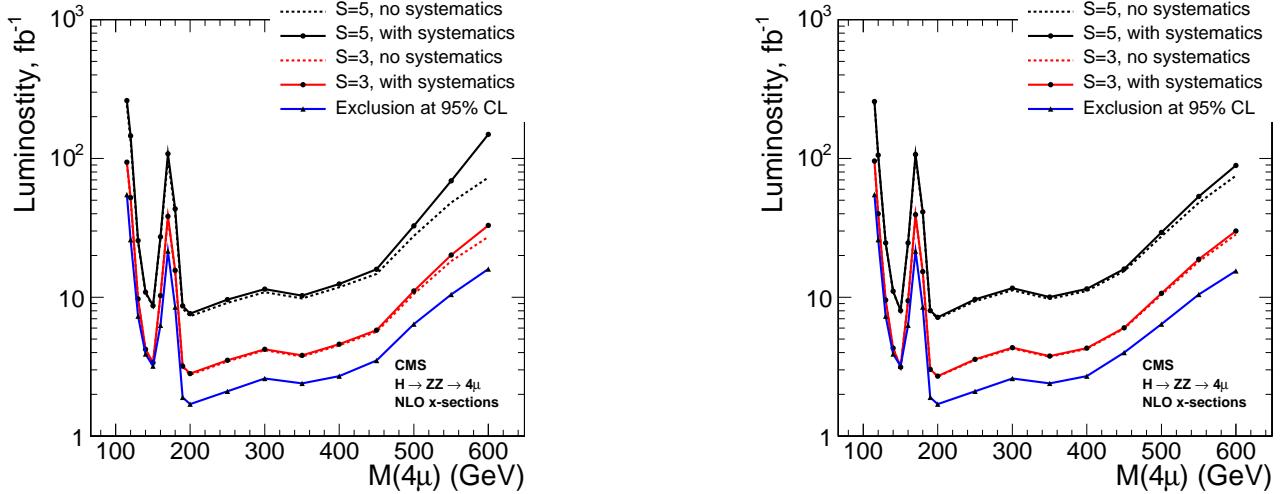


Figure 7-32: Integrated luminosity needed for 95%CL exclusion (triangles, lower line), 3σ (circles, two middle lines), and 5σ (circles, two upper lines) discovery versus Higgs boson mass. Curves for: no systematic errors on background; with systematic errors on background when it is estimated via referencing to the measured $Z \rightarrow 4\mu$. 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 7-29); for the right it is excluded.

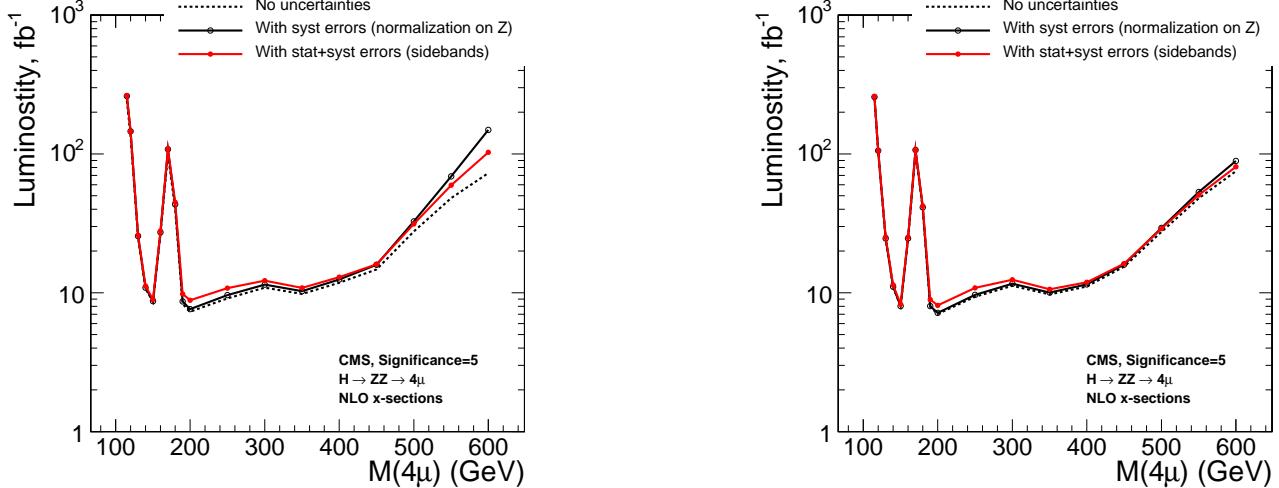


Figure 7-33: Integrated luminosity needed for a 5σ discovery of the Higgs boson versus its mass for: no systematic errors on background (dashed line), with systematic errors on background when it is estimated via referencing to the measured $Z \rightarrow 4\mu$ (empty circles) and with systematic and 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. 7-29); for the right it is excluded.

signal is defined as $s = n_o - b$). This estimator, based on the Log-Likelihood Ratio, is known to follow very closely the true Poisson significance, only slightly over-estimating it in the limit of small statistics [1]. Figure 7-35 presents the results of such a scan for the pseudo-experiment shown in Fig. 7-34. The maximum value of S_{cL} , S_{max} , and the corresponding mass of a “Higgs boson candidate” obtained in each pseudo-experiment were retained for further statistical studies.

After performing 10^8 pseudo-experiments, the differential probability density function for S_{max} and its corresponding cumulative probability function $P(S_{max} > S)$ (Fig. 7-36) were calculated. From Fig. 7-36, one can see that the frequency of observing some large values of S_{cL} (solid line) is much higher than its naive interpretation might imply (dashed line). If desired, the actual probability can be converted to the true significance. The result of such “renormalisation” is presented in Fig. 7-37. One can clearly see that the required de-rating of significance is not negligible; in fact, it is larger than the effect of

including all theoretical and instrumental systematic errors for this channel. More details on the various aspects of these studies can be found in [1].

There are ways of reducing the effect. A more detailed analysis of the shape of the $m_{4\mu}$ distribution will help somewhat. Using the predicted number of signal events $s = s_{theory}$ in the significance estimator to begin with and, then, for validating the statistical consistency of an excess $n_o - b$ with the expectation s_{theory} will reduce the effect further. One can also use a non-flat prior on the Higgs mass as it comes out from the precision electroweak measurements. Whether one will be able to bring the effect to a negligible level by using all these additional constraints on the signal hypotheses is yet to be seen. The purpose of this Appendix is not to give the final quantitative answer, but rather to assert that these studies must become an integral part of all future search analyses when multiple signal hypotheses are tried.

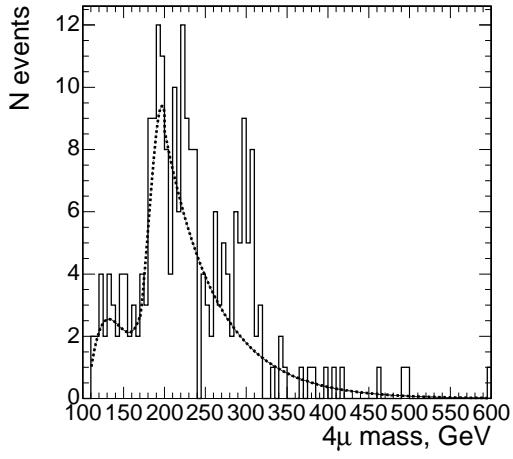


Figure 7-34: The background pdf and an example of one pseudo-experiment with a statistical fluctuation appearing just like a signal.

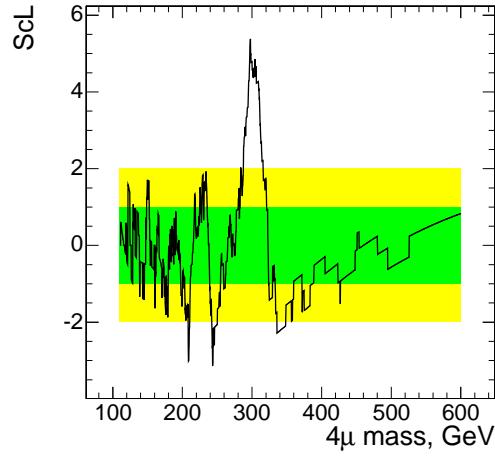


Figure 7-35: Profile of the S_{cL} scan corresponding to the pseudo-experiment example shown on the left. Green (inner) and yellow (outer) bands denote $\pm 1\sigma$ and $\pm 2\sigma$ intervals. Spikes that can be seen are due to events coming in or dropping off the trial-window, a feature of low-statistics searches.

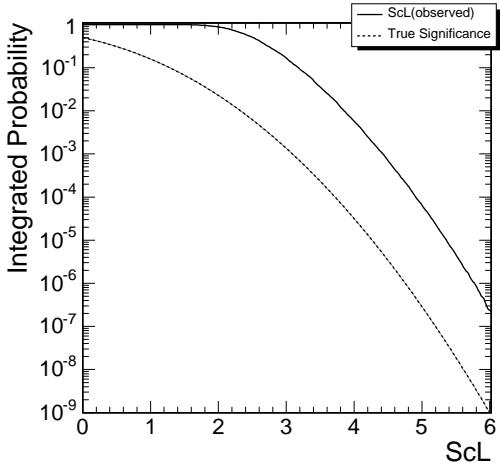


Figure 7-36: S_{cL} cumulative probability density function.

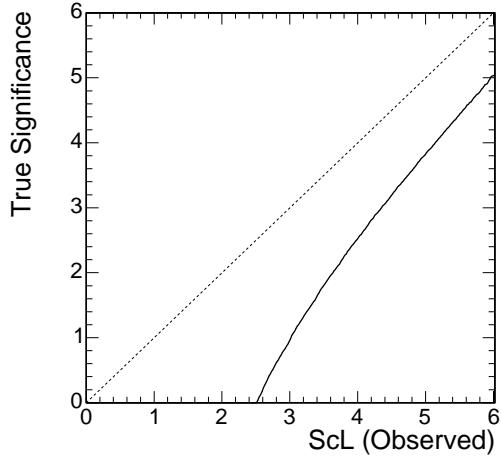


Figure 7-37: Local significance “renormalisation” from an observed value to the true significance with a proper probabilistic interpretation.

CHAPTER 8 CONCLUSIONS

The LHC will enable production of the SM Higgs boson in the entire range of its allowed possible mass (from $114.4\text{GeV}/c^2$ to $\sim 1\text{TeV}/c^2$).

An early discovery of the Higgs boson is one of the most important goals of the CMS experiment. This will require a good understanding of both the data and the physics of the background processes. The research work presented in the dissertation contributes to both the better understanding of the CMS detector and the development of methods which could lead to the discovery of the Higgs boson. The following is the summary of the main results:

Fast algorithm for track segment and hit reconstruction in cathode strip chambers

A new fast algorithm for reconstructing track segments in Cathode Strip Chambers is proposed. The algorithm was validated with the real cosmic ray data taken with 36 CSC chambers operated as a part of CMS-wide MTCC test program in the second half of 2006. The algorithm proved to provide high speed, high efficiency, and good spatial precision—all well within the High Level Trigger requirements.

Search strategy for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel using $M(4\mu)$ -dependent cuts

Discovery of the Standard Model Higgs boson in the “gold-plated” decay mode $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ was analyzed in the context of the CMS Detector. The explored range of Higgs boson masses was $115\text{-}600\text{ GeV}/c^2$. 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 $\approx 2 \text{ fb}^{-1}$ of integrated luminosity, we would be able to start excluding the SM Higgs boson at 95% CL for M_H in vicinity of $200 \text{ GeV}/c^2$. By the time we reach $\approx 30 \text{ fb}^{-1}$, we would exclude the Standard Model Higgs in its four-muon decay mode in the mass range $M_H = 120 - 600 \text{ GeV}/c^2$, if indeed it does not exist.

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

Supporting studies

Many additional topics in “support” of main analyses were studied – results for which are partly outlined in the set of following appendices.

APPENDIX A
RELATIVE CONTRIBUTIONS OF T- AND S-CHANNELS TO THE $ZZ^{(*)} \rightarrow 4\mu$
PROCESS

(A. Drozdetskiy et al., “Relative Contributions of t- and s-Channels to the $ZZ^{(*)} \rightarrow 4\mu$ Process”,
CMS Note 2006/057)

The $q\bar{q} \rightarrow ZZ \rightarrow 4\mu$ process is the main irreducible background in searches for the Higgs boson via its $H \rightarrow ZZ \rightarrow 4\mu$ decay mode. PYTHIA [91], an event generator commonly used for simulation of this process at the LHC, unfortunately is missing the s-channel contribution (Fig. A-1, left).

In our detailed study we show that the s-channel subprocess and its interference with the t-channel cannot be neglected in the context of the backgrounds for the $H \rightarrow 4\mu$ analysis in the area of interest, $M_{4\mu} > 115$ GeV if one aims to simulate the ZZ-background with a precision of 10% or better. This contribution remains non-negligible after all analysis cuts (Fig. A-1, right). A spectacular peak appearing at the 2-muon invariant mass $m(4\mu) \sim m_{Z^0}$ due to presence of the s-channel may prove to become a standard candle for monitoring the level of the ZZ-background.

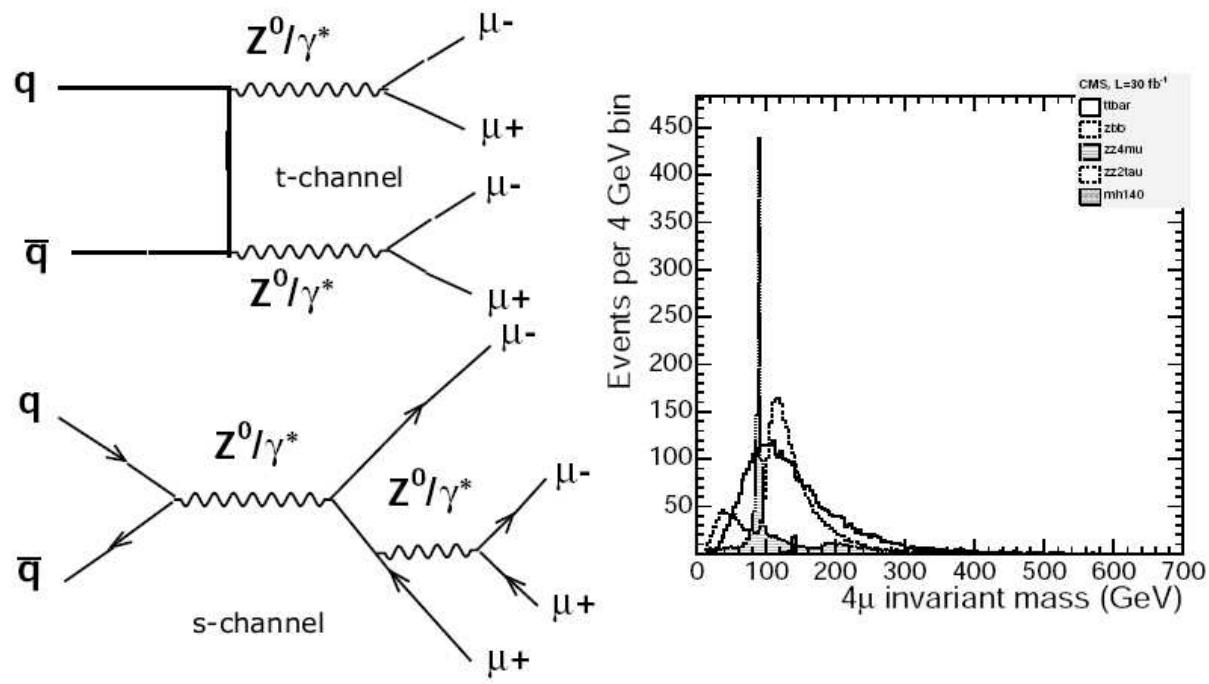


Figure A-1: **Left:** ZZ background: t- and s- channel diagrams. **Right:** S-channel contribution peak around Z^0 mass after pre-selection cuts for fully simulated events.

APPENDIX B
ZZ DISCOVERY WITH $< 1\text{fb}^{-1}$ WITH CMS IN $\text{ZZ}^{(*)} \rightarrow 4\mu$ MODE

(A. Drozdetskiy et al., “*Observation of the ZZ di-boson production in the ZZ^(*) → 4μ channel*”,

CMS Note in CMS approval process.)

We show that the observation of the first few $ZZ \rightarrow 4\mu$ events with significance in excess of 5σ should be expected by the time CMS integrates luminosity of $0.5\text{-}1\text{ fb}^{-1}$. The current estimate of the number of background events is $\leq 0.0173^{Zbb+tt} \pm 0.00998^{Zbb} \pm 0.0116^{tt}$ and is limited by the available MC $t\bar{t}$ sample statistics. A methodology of calculating significance in circumstances when no or very few Monte Carlo background events survive analysis cuts is described. We also demonstrate that a control sample of $4l$ -combinations of wrong flavor and/or charge combination provides a powerful tool for cross checking that the observed events are not due to some unaccounted backgrounds.

APPENDIX C
MEASURING MUON RECONSTRUCTION EFFICIENCY FROM DATA

(A. Drozdetskiy et al., “Measuring Muon Reconstruction Efficiency from Data”,
CMS Note 2006/060)

A method of measuring the global muon reconstruction efficiency ϵ directly from data was studied. With the data corresponding to an integrated luminosity $L = 10 \text{ fb}^{-1}$, the precision of measuring ϵ for muons in the P_T range of $10 - 100 \text{ GeV}$ is expected to be better than 1%, potentially much better. The method largely alleviates uncertainties associated with our ability to monitor and reproduce in Monte Carlo simulation all of details of the underlying detector performance.

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 $p_T > 19 \text{ GeV}/c$. This sample will contain inclusive W, Z, and other processes in the approximate ratio W:Z:others = 10:1:small [107]. By counting the number of $Z \rightarrow 2\mu$ 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 (Fig. C-1), one can evaluate the efficiency of finding globally-reconstructed muons with better than 1% precision. The four-muon efficiency therefore will be known with an absolute error of better than 4%. 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.

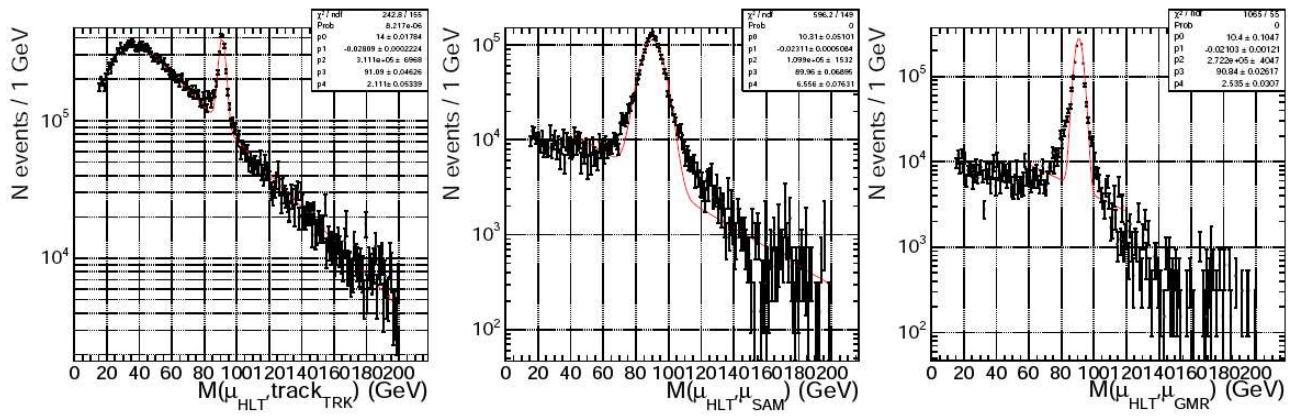


Figure C-1: $Z \rightarrow 2\mu$ invariant mass peak: built from the HLT muon and all other tracks, the HLT muon and all other standalone muons and the HLT muon and all other globally reconstructed muons.

APPENDIX D
SENSITIVITY OF THE MUON ISOLATION CUT EFFICIENCY TO THE
UNDERLYING EVENT UNCERTAINTY

(A. Drozdetskiy et al., “Sensitivity of the Muon Isolation Cut Efficiency to the Underlying Event Uncertainties”,
CMS Note 2006/033.)

The isolation cut efficiency per muon due to uncertainties in the considered underlying event (UE) models vary as much as $\pm 5\%$ (the efficiency itself and its uncertainty strongly depend on how tight the isolation cut is). The 4-muon isolation cut efficiency per event for $ZZ \rightarrow 4\mu$ background is measured to be $(78 \pm 6)\%$.

To decrease these large uncertainties to a negligible level with respect to other systematic uncertainties, we propose to calibrate the isolation cut efficiency from data using Z-inclusive events ($Z \rightarrow 2\mu$) and the random-cone technique. We show that this indeed significantly decreases the uncertainties associated with a poor understanding of the UE physics. There might be $\sim 2\%$ systematic shift in the 4-muon isolation cut efficiencies obtained this way. In principle, one could correct for this shift, but it does not appear to be necessary as this uncertainty is already smaller than other systematic and statistical errors.

APPENDIX E

GENETIC ALGORITHM FOR RECTANGULAR CUTS OPTIMIZATION

(A. Drozdetskiy 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/>)

Typically HEP analysis has quite a few selection criteria (cuts) to optimize for example a significance of the “signal” over “background” events: transverse energy/momenta cuts, missing transverse energy, angular correlations, isolation and impact parameters, etc. In such cases simple scan over multi-dimensional cuts space (especially when done on top of a scan over theoretical predictions parameters space like for SUSY e.g.) leads to CPU time demand varying from days to many years... One of the alternative methods, which solves the issue is to employ a Genetic Algorithm (GA), see e.g. [115, 116, 117].

We wrote a code, GARCON [102], which automatically performs an optimization and results stability verification effectively trying $\sim 10^{50}$ cut set parameters/values permutations for millions of input events in hours time. Examples of analyses are presented in the Physics TDR [8] and recent papers [1, 103, 104, 105].

In comparison to other automatized optimization methods GARCON output is transparent to user: it just says what rectangular cut values are optimal and recommended in an analysis. Interpretation of these cut values is absolutely the same as with eye-balling cuts when one selects a set of rectangular cut values for each variable in a “classical” way by eye.

All-in-all it is a simple yet powerful ready-to-use tool with flexible and transparent optimization and verification parameters setup. It is publicly available along with a paper on it [102] consisting of an example case study and user’s manual.

REFERENCES

- [1] A. Drozdetskiy *et al.*, Search strategy for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel using $M(4\mu)$ -dependent cuts, 2006, 2006/122 CMS Note.
- [2] A. Drozdetskiy *et al.*, Relative contributions of t- and s-channels to the $ZZ \rightarrow 4\mu$ process, 2006, 2006/057 CMS Note.
- [3] A. Drozdetskiy *et al.*, Measuring muon reconstruction efficiency from data, 2006, 2006/060 CMS Note.
- [4] A. Drozdetskiy *et al.*, Sensitivity of the muon isolation cut efficiency to the underlying event uncertainties, 2006, 2006/033 CMS Note.
- [5] P. Bartalini *et al.*, Study of PDF and QCD scale uncertainties in $pp \rightarrow ZZ \rightarrow 4\mu$ events at the LHC, 2006, 2006/068 CMS Note.
- [6] P. Bartalini *et al.*, NLO vs. LO: kinematical differences for signal and background in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ analysis, 2006, 2006/130 CMS Note.
- [7] M. Della Negra, A. Petrilli, A. Herve, L. Foa, *et al.*, CMS physics: technical design report, volume I, 2006, (CMS collaboration), in particular see section 9.3.4.
- [8] M. Della Negra, A. Petrilli, A. Ball, L. Foa, *et al.*, CMS physics technical design report, volume II, 2007, (CMS collaboration), in particular see sections 3.1, Appendix A.2, Appendix D.
- [9] V. Buescher *et al.*, Tevatron-for-LHC report: preparations for discoveries, 2006, hep-ph/0608322.
- [10] C. Buttar *et al.*, Les houches physics at TeV colliders 2005, standard model and Higgs working group: summary report, 2005, contributed to Les Houches workshop on physics at TeV colliders, Les Houches, France. hep-ph/0604120.
- [11] A. Drozdetskiy *et al.*, SM Higgs boson: recent developments at CMS and ATLAS, 2007, (CMS and ATLAS collaboration), talk given at The XLIIInd Rencontres de Moriond on Electroweak Interactions, La Thuile, Italy, 2007 (will be published in the conference proceedings).
- [12] A. Drozdetskiy *et al.*, CMS Detector sensitivity to the Standard Model Higgs boson in $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel, 2007, (CMS collaboration), talk given at American Physical April 2007 meeting.
- [13] A. Drozdetskiy and S. Abdullin, GARCON: genetic algorithm for rectangular cuts optimizatioN, 2007, talk given at The XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Amsterdam, Netherlands, 2007.

- [14] A. Drozdetskiy *et al.*, CMS detector sensitivity to the Standard Model Higgs boson in $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel, 2007, (CMS collaboration), talk given at The Hadron Collider Physics Symposium 2007, Elba, Italy (will be published in the conference proceedings).
- [15] S. Abdullin *et al.*, Search strategy for the standard model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel using M(4mu)-dependent cuts, 2007, (CMS collaboration), talk given at The Physics at LHC, Cracow, Poland, 2006 (will be published in proceedings).
- [16] A. Drozdetskiy *et al.*, Sensitivity of the muon isolation cut efficiency to underlying event uncertainty, 2006, (CMS collaboration), talk given at The 2nd HERA - LHC workshop, CERN, Geneva, Switzerland (published in proceedings).
- [17] A. Drozdetskiy *et al.*, MC@NLO vs PYTHIA for $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$, 2005, (CMS collaboration), talk given at The Les Houches 2005: Physics at TeV Colliders, Les Houches, France.
- [18] A. Drozdetskiy *et al.*, Effect of UE on isolation, 2005, (CMS collaboration), talk given at The Les Houches 2005: Physics at TeV Colliders, Les Houches, France (published in proceedings).
- [19] A. Drozdetskiy *et al.*, The discovery potential of supersymmetry at CMS within the mSUGRA model using same-sign di-muons., 2005, (CMS collaboration), talk given at The 40th Rencontres de Moriond on QCD and High Energy Hadronic Interactions, La Thuile, Aosta Valley, Italy (published in proceedings, hep-ex/0505034).
- [20] A. Drozdetskiy *et al.*, First detailed study on the CMS SUSY discovery potential with two same sign muons in the mSUGRA model, 2004, (CMS collaboration), talk given at The Physics at LHC, Vienna, Austria (Published in Czech.J.Phys.55:B249-B256,2005).
- [21] A. Drozdetskiy *et al.*, internal CMS collaboration meetings of different levels, 2002-2007, see: <http://indicosearch.cern.ch/> and <http://agenda.cern.ch/search.php>.
- [22] C.-N. Yang *et al.*, Conservation of isotopic spin and isotopic gauge invariance. Phys. Rev., 96:191-195, 1954.
- [23] P. Higgs *et al.*, Broken symmetries, massless particles and gauge fields. Phys. Lett., 12:132-133, 1964.
- [24] P. Higgs *et al.*, Broken symmetries and the masses of gauge bosons. Phys. Rev. Lett., 13:508-509, 1964.
- [25] S. Glashow, Partial symmetries of weak interactions. Nucl. Phys., 22:579-588, 1961.
- [26] S. Weinberg, Phys. Rev. Lett. 19, 1264 - 1266 (1967).

- [27] A. Salam, Elementary particle theory,. Almqvist and Wiksells, Stockholm, 1968.
- [28] G. Arnison *et al.*, Experimental observation of lepton pairs of invariant mass around 95- gev/c^2 at the cern sps collider. Phys. Lett., B126:398-410, 1983.
- [29] G. Arnison *et al.*, Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ -gev. Phys. Lett., B122:103-116, 1983.
- [30] J. F. Gunion *et al.*, The Higgs hunter's guide, Westview Press, 2001.
- [31] M. Spira *et al.*, Electroweak symmetry breaking and Higgs physics, hep-ph/9803257.
- [32] T. Hambye *et al.*, Phys. Rev. D 55, 7255 (1997), arXiv:hepph/9610272.
- [33] R. Barate *et al.*, LEP working group for Higgs boson searches, Phys. Lett. B 565, 61 (2003), arXiv:hep-ex/0306033.
- [34] Tevatron new phenomena & Higgs working group, combined D0 and CDF upper limits on the Standard Model Higgs boson production, http://tevnphwg.fnal.gov/results/d0conf_5227/, CDF Note 8384, D0 Note 5227.
- [35] LEP Electroweak Working Group. LEPEWWG/2005-01, arXiv:hep-ex/0511027, <http://lepewwg.web.cern.ch/LEPEWWG/> (Status of March 2007).
- [36] M. Grünwald, Precision tests of the Standard Model, hep-ex/0511018.
- [37] A. Djouadi, J. Kalinowski, and M. Spira, HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension, Comput. Phys. Commun., vol. 108 (1998), 56-74, hep-ph/9704448.
- [38] M. Spira, HIGLU: a program for the calculation of the total Higgs production cross section at hadron colliders via gluon fusion including QCD corrections, (1995), hep-ph/9510347.
- [39] M. Gonzalez-Garcia *et al.*, Rev. Mod. Phys. 75, 345 (2003), hepph/0202058.
- [40] Y. Fukuda *et al.*, Nucl. Instrum. Meth. A 501, 418 (2003).
- [41] A. Suzuki *et al.*, Nucl. Instrum. Meth. A 453, 165 (2000), hep-ex/0004024.
- [42] A. Suzuki *et al.*, Nucl. Instrum. Meth. A 453, 165 (2000), hep-ex/0004024.
- [43] A. Piepke *et al.*, Nucl. Phys. Proc. Suppl. 91, 99 (2001).
- [44] E. Kolb *et al.*, The Early universe, Addison-Wesley, Redwood City, USA (1990), Frontiers in physics, 69.
- [45] A. Sakharov, Pisma Zh. Eksp. Teor. Fiz. 5, 32 (1967) (JETP Lett. 5, 24 (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991)).

- [46] N. Cabibbo *et al.*, Phys. Rev. Lett. 10, 531 (1963).
- [47] C. Bennett *et al.*, Astrophys. J. Suppl. 148, 1 (2003), astro-ph/0302207.
- [48] D. Spergel *et al.*, Astrophys. J. Suppl. 148, 175 (2003), astro-ph/0302209.
- [49] E. Witten, Nucl. Phys. B 188, 513 (1981).
- [50] S. Dimopoulos *et al.*, Nucl. Phys. B 193, 150 (1981).
- [51] S. Dimopoulos *et al.*, Phys. Rev. D 24, 1681 (1981).
- [52] W. Yao *et al.*, (Particle Data Group Collaboration), J. Phys. **G33**, 1 (2006).
- [53] LHC Project, <http://lhcb.web.cern.ch/lhc/>,
http://lhcb-machine-outreach.web.cern.ch/lhc-machine-outreach/lhc_in_pictures.htm.
- [54] CERN, <http://www.cern.ch/>.
- [55] CMS Collaboration, <http://cms.cern.ch/>.
- [56] ATLAS Collaboration, <http://atlas.web.cern.ch/Atlas/index.html>.
- [57] ALICE Collaboration, <http://aliceinfo.cern.ch/>.
- [58] LHCb collaboration, <http://lhcb.web.cern.ch/lhcb/>.
- [59] TOTEM collaboration, <http://totem.web.cern.ch/Totem/>.
- [60] CMS Collaboration, The compact muon solenoid detector at LHC, in preparation, 2007.
- [61] CMS Collaboration, The Compact Muon Solenoid Technical Proposal, CERN/LHCC 94-38 (1994) LHCC/P1.
- [62] CMS Magnet Collaboration, The magnet project: Technical Design report, CERN, volume CERN/LHCC97-10, CMS TDR 1, 1997; A. Herve', et al., Status of the Construction of CMS magnet, IEEE Trans. Appl. Superconduct, Vol 14, No 2, pp. 524-547, June 2004; A. Herve', The CMS Detector Magnet, IEEE Trans on Appl. Superconductor, Vol. 10, No1, pp. 389-394, 2000; F. Kircher, et al., Final Design of the CMS Solenoid Cold Mass, IEEE Trans on Appl. Superconduct., Vol 10, n1, 407-410, March 2000.
- [63] CMS Collaboration, The Tracker Project Technical Design Report, CERN/LHCC 98-006, CMS TDR 5, April 1998; CMS Collaboration, Addendum to the CMS Tracker TDR, CERN/LHCC 2000-016, CMS TDR 5 Addendum 1, April 2000.
- [64] CMS Collaboration, CMS DAQ and HLT technical design report, CERN/LHCC 2002-26, 2002.

- [65] CMS Collaboration, The muon project. Technical design report, CERN/LHCC, 97-32, 1997.
- [66] G. Charpak *et al.*, High-accuracy, two-dimensional read-out in multiwire proportional chambers, Nucl. Instrum. Meth. 113:381-385, 1973.
- [67] C. Anderson *et al.*, Effect of gas composition on the performance of cathode strip chambers for the CMS Endcap Muon System, CMS Note 2004/033.
- [68] M. Baarmand *et al.*, Spatial resolution attainable with cathode strip chambers at the trigger level, Nucl. Instrum. Meth. A425:92-105, 1999.
- [69] On 'Global Muon Momentum Resolution' see for example: CMS Collaboration, Physics Technical Design Report, vol.1, CERN/LHCC, 2006-001, 2006.
- [70] On 'muon (TC) segment reconstruction algorithm' see for example: CMS Collaboration, Physics Technical Design Report, vol.1, CERN/LHCC, 2006-001, 2006.
- [71] J. Chiba *et al.*, Study of position resolution for cathode readout MWPC with measurement of induced charge distribution Nucl. Instrum. Meth. 206 (1983), p.451.
- [72] D. Acosta *et al.*, Large CMS cathode strip chambers: design and performance, Nucl. Instrum. Meth. 453 (2000) 182-187.
- [73] M. Baarmand *et al.*, Spatial resolution attainable with cathode strip chambers at the trigger level, Nucl. Instrum. Meth. A425:92-105, 1999.
- [74] Y. Pakhotin *et al.*, Performance of endcap muon system of the compact muon solenoid, American Physical Society Conference, April 14-17, 2007, Jacksonville, Florida, USA.
- [75] CMS collaboration, Summary report on magnet test and cosmic challenge in preparation, 2007.
- [76] V. Barashko *et al.*, Performance validation tests of the cathode strip chambers for CMS muon system, in Nuclear Science Symposium Conference Record, 2005 IEEE, Vo. 2, p.827-829, 23-29 Oct. 2005.
- [77] E. Gatti *et al.*, Optimum geometry for strip cathodes or grids in mwpc for avalanche localization along the anode wires, Nucl. Instrum. Meth. 163 (1979) 83.
- [78] E. Mathieson and J. Gordon, Cathode charge distributions in multiwire chambers: I. Measurement and theory, Nucl. Instrum. Meth. 227 (1984) 267-276.
- [79] E. Mathieson and J. Gordon, Cathode charge distributions in multiwire chambers: II. Approximate and empirical formulae, Nucl. Instrum. Meth. 227 (1984) 277-282.

- [80] 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.
- [81] 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.
- [82] S. Wilks, The large-sample distribution of the likelihood ratio for testing composite hypotheses, *Annals of Mathematical Statistics*, 9:60-62, 1938.
- [83] CMS collaboration, Object-oriented simulation for CMS analysis and reconstruction, <http://cmsdoc.cern.ch/OSCAR>.
- [84] CMS collaboration, Object-oriented reconstruction for CMS analysis, <http://cmsdoc.cern.ch/ORCA>.
- [85] S. Abdullin *et al.*, Summary of the CMS Potential for the Higgs boson discovery, CMS Note 2003/033.
- [86] M. Sani *et al.*, 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.
- [87] V. Bartsch, Simulation of silicon sensors and study of the Higgs decay $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ for CMS (LHC), Ph.D. thesis, IEKP-KA/2003-26 University of Karlsruhe (2003).
- [88] M. Aldaya *et al.*, Search for the Standard Model Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ decay channel using a mass-independent analysis, CMS Note 2006/106.
- [89] S. Baffioni *et al.*, Search for the Standard Model Higgs Boson in the four-electron final state with CMS, CMS Note 2006/115.
- [90] D. Futyan *et al.*, Search for the Standard Model Higgs boson in the two-electron and two-muon final state with CMS, CMS Note 2006/136.
- [91] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, and E. Norrbin, *Comput. Phys. Commun.* **135**, 238 (2001).
- [92] CMS collaboration, CMS interface for event generators, CMKIN, <http://cmsdoc.cern.ch/cmsoo/projects/CMKIN>.
- [93] E. Barberio *et al.*, *Computer Phys. Commun.* 79 (1994) 291.
- [94] E. Barberio *et al.*, *Computer Phys. Commun.* 66 (1991) 115.
- [95] M. Spira *et al.*, HDECAY: a program for Higgs Boson decays in the Standard Model and its supersymmetric extension, [hep-ph/9704448](https://arxiv.org/abs/hep-ph/9704448).

- [96] M. Spira *et al.*, HIGLU: a program for the calculation of the total Higgs production cross section at hadron colliders via gluon fusion including QCD corrections, hep-ph/9510347.
- [97] Particle Data Group, Review of Particle Properties, Phys. Lett. B 562 (2004) 1.
- [98] F. Maltoni, Theoretical issues and aims at the Tevatron and LHC, proceeding of The Hadron Collider Physics Symposium, Les Diableretes, Switzerland, July 2005, printed version: Springer Proceedings in Physics, Vol. 108, 2006, ISBN: 3-540-32840-8.
- [99] CompHEP collaboration, CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33, year =.
- [100] J. Campbell, W/Z + B anti-B / jets at NLO using the Monte Carlo MCFM, hep-ph/0105226.
- [101] C. Zecher *et al.*, Leptonic signals from off-shell Z boson pairs at hadron colliders, hep-ph/9404295.
- [102] A. Drozdetskiy *et al.*, GARCON: genetic algorithm for rectangular cuts optimization. User's manual for version 2.0, 2006, hep-ph/0605143, <http://drozdets.home.cern.ch/drozdets/home/genetic/>.
- [103] D. Acosta *et al.*, Potential to discover supersymmetry in events with muons, jets and missing energy in pp collisions at $\sqrt{s} = 14$ TeV with the CMS detector, CMS Note 2006/134.
- [104] V. Abramov *et al.*, Selection of single top events with the CMS detector at LHC, CMS Note 2006/084.
- [105] W. Boer *et al.*, Trilepton final state from neutralino-chargino production in mSUGRA, CMS Note 2006/113 (or just PTDR).
- [106] F. Maltoni *et al.*, MadEvent: automatic event generation with MadGraph, JHEP 0302 (2003) 027, arXiv:hep-ph/0208156.
- [107] CMS Collaboration, The Trigger and Data Acquisition project, v.2, p.308, 2002.
- [108] E. Lehmann, A Theory of some multiple decision problems, The Annals of Mathematical Statistics, Vol. 28 1, 1, (1957).
- [109] R. O'Neill *et al.*, The present state of multiple comparison methods, Journal of the Royal Statistical Society. Series B (Methodological), Vol. 33 2, 218, (1971).
- [110] J. Hauser, Search for new physics in tails of distributions, private communications, paper in preparation.

- [111] B. Knuteson *et al.*, A quasi-model-independent search for new high p_T physics at D0, thesis, 2000.
- [112] 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.
- [113] 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.
- [114] 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.
- [115] J. Holland, Adaptation in natural and artificial systems. The University of Michigan Press, Ann Arbor, 1975.
- [116] D. Goldberg, Genetic algorithms in search, optimization and machine learning. Addison Wesley, 1989.
- [117] S. Abdullin, Genetic algorithm for SUSY trigger optimization in CMS detector at LHC, NIM A 502 (2003) 693-695; Genetic algorithm for SUSY trigger optimization, talk given at the IV Conference LHC Days in Split, October 8-12, 2002.

BIOGRAPHICAL SKETCH

Alexey A. Drozdetskiy was born on March 31, 1978, in the city of Novosibirsk in Russia. After graduating with honors from high school in 1995, he continued his education at the Novosibirsk State University (NSU).

In 1999 he received his bachelor of science degree with honors, and in 2001 he received his master's degree with honors from the same university. While studying in the NSU, he worked in the Budker Institute in the Spherical Neutral Detector (SND) group on the VEPP2M, e^+e^- -collider. His work was related to developing new detectors for X-rays as well as studying a possible upgrade for the SND detector for forward physics at VEPP2M.

During his period of study at NSU he was awarded 4 years with International Soros Student Stipend in Physics, 2 years with Special NSU stipend and 8 times with student conferences diplomas.

In the winter of 2002, Alexey became a graduate student in the Department of Physics, University of Florida, Gainesville, FL. In the same year he joined CMS collaboration and worked with professors Guenakh Mitselmakher, Andrey Korytov and Darin Acosta on various aspects of experimental particle physics, particularly on the $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ CMS discovery potential (as well as CMS discovery potential for SUSY in final state with muons and E_T^{miss}). Since 2002 Alexey has been participating in commissioning of the Endcap Muon System (EMU) of the CMS detector (starting from electronics/detector tests at UFL site and up to the most recent performance studies and fast High Level Trigger algorithm development for the EMU, validated with real cosmic muons data, taken by CMS detector slice in fall of 2006).

In the period of his study at UF, Alexey was awarded with the UF Presidential Recognition, UF International Student Academic Award, and 5 Certificates of Achievements. He is co-author of more than 10 refereed publications, and results of the studies he

performed together with colleagues were presented at more than 10 international conferences.

He is currently based in Geneva, Switzerland, close to LHC and CMS experiment (CERN international laboratory), and gives free Yoga lessons to the community on voluntary basis.

In 2007, Alexey A. Drozdetskiy graduated from the University of Florida with the degree of doctor of philosophy.