Mid Term Report 2009

Type of the project	Helmholtz-(Hochschul-) Nachwuchsgruppe
Support Number	VH-NG-503
Topic	Probing Electroweak Symmetry Breaking at LHC: Higgs Physics with the CMS Detector
Scientist in charge	Dr. Alexei Raspereza
Helmholtz Center	DESY
University Partner	University of Karlsruhe
Reference period	1.03.2009 – 31.03.2010

1. Introduction

The Young Investigator Group VH-NG-503 has started its work on March 1st 2009. During the past year, in addition to group leader, one PhD student and one Postdoc have joined the group. Furthermore, a member of the DESY CMS group in Zeuthen (Igor Marfin) has joined the activities of the group as a part time associate member. The activities of the group are centered at DESY Hamburg. As of March 2010 the following researches contributed to the group activities.

Members of the Group	*Position*	*Activities*
Dr. Alexei Raspereza	Group Leader	Validation of tracking and b-tagging software, analysis of the first collision data in the context of strange hadron spectroscopy, CMS MSSM Higgs analysis
Dr. Roberval Walsh	Postdoc	Commissioning of the Beam Condition Monitor (BCM1F) of the CMS detector, development and maintenance of the BCM readout software, performance studies of b-tagging algorithms
Agni Bethani	PhD student	CMS MSSM Higgs analysis, validation of the CMS tracking software with the first collision data
Associate members of the Group		
Igor Marfin	Guest scientist	Implementation of the method for determination of the b-

		tagging efficiency with top pair events
University Partner		
Prof. Dr. Thomas Mueller	University of Karlsruhe	

2. Commitments of the Young Investigator Group

The main focus of the group activities lies on commitments in the context of the CMS experiment. CMS, standing for Compact Muon Solenoid, is one of the four detectors installed on the Large Hadron Collider (LHC) at CERN [1]. As a multi-purpose apparatus, CMS will provide particle physicist with an experimental tool for the exploration of the Electroweak Symmetry Breaking mechanism and physics beyond the Standard Model. The Young Investigator Group covers several aspects of the detector operation, reconstruction of the data and physics analysis. This includes commissioning and operation of the Beam Condition Monitor of the CMS detector, validation of the reconstruction software, namely tracking and b-tagging algorithms, and analysis of experimental data in the context of searches for the neutral Higgs bosons of the Minimal Supersymmetric extension of the Standard Model (MSSM).

3. CMS Fast Beam Conditions Monitor

The CMS Fast Beam Conditions Monitor (BCM1F) is one of the subsystems of the CMS Beam Radiation Monitoring (BRM) [2]. The BMC1F detector monitors the beam-halo flux on a bunch-by-bunch basis. It consists of two planes installed around the beam pipe located at $z = \pm 1.8$ m from the interaction point. Each plane has four modules with 5x5 mm^2 radiation hard single crystal diamond sensors (sCVD) and radiation hard front-end electronics. The location of BCM1F near the CMS pixel detector is very important for the safe operation of that detector.

The data acquisition (DAQ) systems of BCM1F comprise analog-to-digital converters (ADCs) for signal sampling, scalers for counting rates and time-to-digital converters (TDCs) for time measurements. The BCM1F TDCs have time a resolution of 0.8 ns. The signals from the modules come from the detector through optical fibers to an optical receiver. Then they are fanned into the ADCs and discriminators, where thresholds are set, then into the scalers and the TDCs. The data is processed in a readout PC and sent to monitors in the CMS control room and to storage for offline analysis.

In the past year Dr. Roberval Walsh had been contributing to the activities related to BCM1F together with the BCM1F group at DESY-Zeuthen and the BRM group at CERN. Below is a summary of those contributions.

3.1 Software development for DAQ of the TDCs

Most of the current software used for DAQ of the BCM1F TDCs has been developed by Dr. Walsh. He also maintains the codes and implements new features. The developed software consists of the three main parts: readout, data storage and data publication. The data from BCM1F are sent via the CERN network to the BRM shift crew in the CMS control room. Because of the large amount of data produced by the TDCs, the data storage was carefully studied to minimize storage resources and avoid data loss due to computing dead-time or computer crashes. The software is currently in a very stable version.

3.2 Data analysis

The BCM1F detector was fully operational during the LHC runs in November and December 2009. With the data accumulated during that period Dr. Walsh started to perform analyses using the TDCs information for better understanding of the detector and perform calibrations. A main source of uncertainties in the time measurements is due to the lack of information about the length of the cables from the detector to the electronics. An analysis on an event-by-event basis is in progress to determine systematic delays in time between modules in the same plane. In addition, an analysis is ongoing to perform calibration using the relative time of modules located at opposite sides of the CMS detector (the time-of-flight between the planes of BCM1F is 12 ns). The time distribution for the individual modules of one bunch crossing the CMS detector is shown in Figure 1.

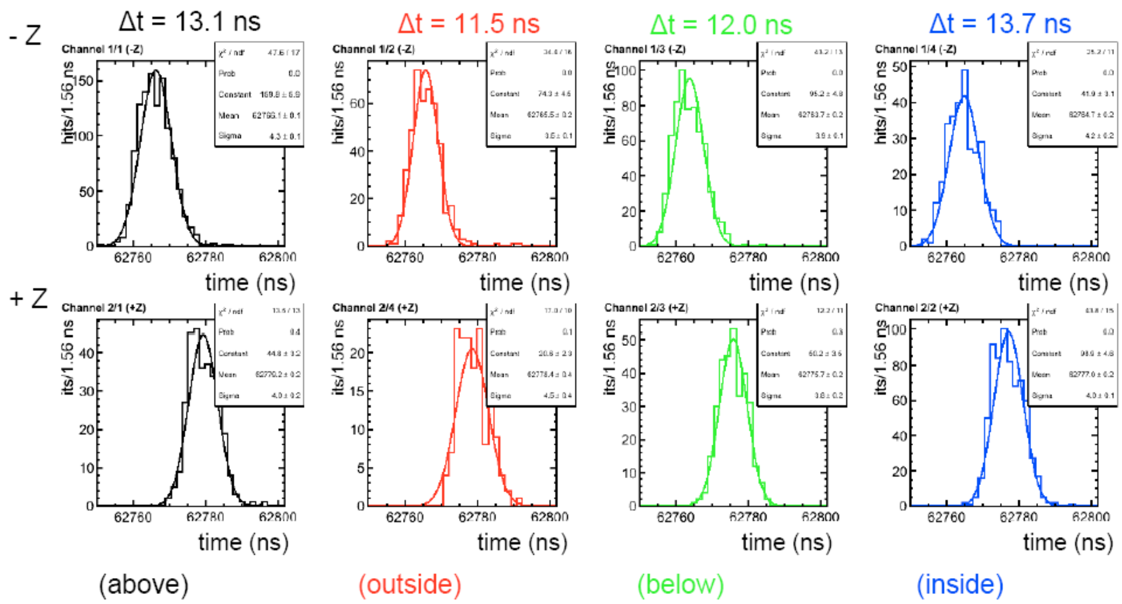

Figure 1: The time distribution for the the BCM1F individual modules of one bunch crossing the CMS Detector. The plots on top (bottom) show the response of the modules in the -Z (+Z) plane. The time difference between opposite modules is also shown.

Using the excellent time resolution of the TDCs and the privileged location of BCM1F and knowing that the time between two bunches in the LHC is 24.95ns, the analysis has been performed to extract the LHC bunch number scheme. The code has been developed to provide on-line information of the bunch number scheme obtained using BCM1F to the BRM shift crew in the control room.

The first results of the analysis of the data from 2009 runs have been presented by Dr. Roberval Walsh at the Deutsche Physikalische Gessellschaft (DPG) meeting held in Bonn, Germany, on March 15^{th} – 19^{th} 2010.

4. Physics Analysis: Search for Neutral Supersymmetric Higgs Bosons

One of the goals of the physics program at LHC is the exploration of the Higgs mechanism as the most favored scenario of the Electroweak Symmetry Breaking. In the Standard Model, the Electroweak Symmetry Breaking is realized by introducing one doublet of the scalar complex fields with non-zero vacuum expectation value, which breaks SU(2) symmetry and generate masses of weak bosons and fermions. The Higgs mechanism in the

Standard Model predicts one more physical state, Higgs boson – a scalar particle, appearing as a remnant of the Electroweak Symmetry Breaking mechanism. A more complex structure of the Higgs sector is expected in Supersymmetric models, which are proposed in order to overcome fundamental inconsistency of the Standard Model, known as hierarchy problem. Minimal Supersymmetric Standard Model (MSSM) postulates two Higgs doublets, which give rise to five physical states, two neutral CP-even Higgs bosons, *h* and *H,* one CP-odd Higgs boson *A* and two charged Higgs bosons *H*□. These particles will be target of searches at LHC.

4.1 MSSM Higgs Boson Analyses within CMS

A number of analyses have been developed by the CMS Collaboration to search for the supersymmetric neutral Higgs bosons. These analyses exploit gluon-fusion and b-quark associated production of the neutral Higgs bosons and their decays to tau leptons. One of the key parameters, affecting Higgs boson phenomenology in the supersymmetric models, is tan□ - the ratio of vacuum expectation values of the two Higgs doublets. The initial studies showed that early observation of the supersymmetric Higgs bosons is possible in the case of the high□ scenario in which the cross sections of the Higgs production mechanisms are enhanced. These studies exploited the following tau decay modes:

- muon + hadronic tau-jet [3] ;
- electron + hadronic tau-jet [3];
- electron + muon [4].

4.2 Search for MSSM Higgs Bosons Decaying to Tau Leptons in Dimuon Channel

To improve the combined sensitivity to the neutral MSSM Higgs bosons, the Young Investigator Group has developed a novel analysis, searching for the MSSM Higgs boson decays to a pair of tau leptons in the channel, characterized by the presence of two muons and missing transverse energy. In this channel both tau leptons decay leptonically to muon and two neutrinos, □□□and □□. The analysis is challenged by the large background resulting from the direct muonic decays of *Z* bosons produced in association with jets. The developed analysis exploits kinematic characteristics of events, such as impact parameter of muons, missing transverse energy, pseudorapidity of the dimuon system *etc*, to distinguish between signal and main background processes. The likelihood technique, combining the most discriminating variables into a single discriminant, is implemented to enable efficient suppression of background while keeping high selection efficiency of signal events. The prospects of the Higgs boson searches with early LHC data have been studied with the simulated samples of signal and background events. At the time when these studies had been performed, the initial running of the LHC machine was planned at 10 TeV centre-of-mass energy. Therefore, the analysis has been performed, assuming 200 pb^{-1} of data collected at the centre-of-mass energy 10 TeV. Figure 2 shows the

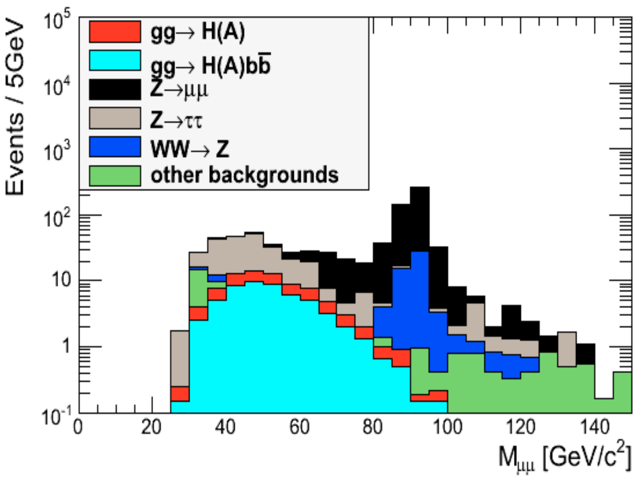

Figure 2: The dimuon invariant mass spectrum of the selected events in the MSSM Higgs boson analysis. The Higgs boson mass hypothesis $m_A = m_H = 115$ GeV and tan□=40.

spectrum of invariant mass of the two muons for the sample of selected Monte Carlo events. The signal sample corresponds to the Higgs boson mass hypothesis $m_A = m_H$ = 115 GeV. The statistical significance of the signal has been evaluated using the dimuon invariant mass spectrum as the final discriminant. Data-driven methods have been developed to evaluate contribution of the main background processes, namely *Z*□□□ , *Z*□□□□and QCD events, to the final sample of selected events. These methods provide estimates of the systematic error on the background expectation. Taking these errors into account, we demonstrated that the studied channel can contribute to the observation of the signal from the neutral MSSM Higgs bosons with the statistical significance of more than 2.1(3.7) standard deviation for tan□ values greater than 30(40) and Higgs boson masses below 120 GeV. The analysis has been presented to the CMS *Higgs Physics Analysis Group* and included into official combination of the MSSM Higgs boson searches with the CMS detector. According to the final decision, in the years 2010-2011 LHC will be operated at center of mass energy 7 TeV. The analysis therefore has been re-optimized for this center of mass energy and is currently being repeated with the new Monte Carlo samples of the signal and background processes generated at center of mass energy 7 TeV. The corresponding CMS Analysis Note, documenting the details of the analysis, is being currently prepared by the group members Agni Bethani and Alexei Raspereza.

It should be noted that with only slight modifications, the analysis can be used to study *Z* boson production in association with jets, followed by *Z*□□□ decay with both tau leptons decaying to muons.

5. B-Tagging Related Studies

Identification of the heavy flavor jets will be a key ingredient of many physics analyses at LHC, including among others searches for supersymmetric Higgs bosons produced in association with b- quarks. The Young Investigator Group has joined activities of the *B-tagging Physics Object Group* of the CMS Collaboration and taken part in preparation for physics analyses employing b-tagging algorithms. The main fields of contribution are:

- study of the b fragmentation related systematic effects on b-tagging performance;
- development of the data-driven methods for the determination of the b-tagging efficiency and mistag rate.

5.1 Study of the b Fragmentation Related Systematic Effects on Performance of B-tagging Algorithms

Good control over systematic effects in identification of heavy flavor jets is prerequisite for reliable estimation of the signal selection efficiency and background rejection rate for the analyses relying upon b-tagging. One of the sources of the systematic effects on the performance of b-tagging algorithm is associated with an uncertainty on b fragmentation models. The studies of the relevant systematics effects have been performed using as an input results of the b fragmentation measurements at LEP and SLD experiments [5-8]. A number of models have been tested and parameters of these models have been varied within the experimental errors reported by LEP and SLD experiments. The effect of fragmentation model uncertainty on b-tagging efficiency has been studied using simulated samples of top quarks pairs and QCD dijet events. In order to enable comparison between various fragmentation models without regenerating Monte Carlo samples, the jet re-weighting procedure has been implemented, which assigns each b-jet with a weight given by the ratio of the probed fragmentation function to the fragmentation function used in generation of reference sample . The dedicated studies have shown that relative uncertainty on the b-tagging efficiency ranges from 1.2% for the simple algorithm based on the charged track impact parameters [9] to 4% for a more sophisticated combined secondary vertex algorithm [10] which exploits kinematic variables highly sensitive to b fragmentation, such as secondary vertex decay length and the ratio of the energy sum of all tracks contributing to the secondary vertex to the total energy of the charged tracks in a jet.

In addition, we studied impact of b fragmentation on the data driven determination of b-tagging efficiency using so called p_T^{Rel} method [11]. In this method, one of the jets in the selected sample of QCD dijet events is tagged using information about prompt muon, while another jet is probed for its compatibility with the b-quark hypothesis. This technique exploits distribution of the transverse momentum of the muon with respect to the jet axis in the tagged jet (this variable is referred to as p_T^{Rel}). The fraction of b- c- and uds-jets is determined by fitting the p_T^{Rel} distribution with Monte Carlo templates for b, c and light flavor jets. The method assumes adequate simulation of the p_T^{Rel} variable in Monte Carlo. It was found that shape of the p_T^{Rel} distribution in the sample of b-jets weakly depends on the b fragmentation, leading to a negligible error (less than 0.5%) in the determination of the b-tagging efficiency with the p_T^{Rel} method.

The software developed in the course of these studies has been delivered to the *B-Tagging Physics Object Group* and became a part of the official CMS software.

5.2 Determination of B-Tagging Efficiency with Top Pair Events

Within the Standard Model, top quarks are expected to decay with almost 100% branching fraction to a *W* boson accompanied by a b-quark. The dedicated selection of the top pair semileptonic decays has been developed and the procedure of determination of the b-tagging efficiency and mistag rate based on proposal in [11] has been implemented within the CMS software framework by Igor Marfin. In the semileptonic channel, given the b-quark identification efficiencies and udsc-quark mistag rates, the number of events with n_b tagged b-jets, n_c tagged c-jets and n_l tagged light flavor jets can be predicted. By imposing a consistency between predicted number of events with one, two or more tagged jets to the actual number of observed events with that particular combination, the b-tagging efficiency and udsc-mistag rates can be measured. The following log-likelihood function is minimized to measure the tagging efficiency and mistag rate:

$$L=-2\sum log(P(N_n^{meas},N_n^{pred})),$$

where N_n^{meas} is the measured number of events with n tagged jets, N_n^{pred} is the predicted number of events with n tagged jets and $P(N_n^{meas},N_n^{pred})$ is the Poisson distribution. N_n^{pred} depends on the b-tagging efficiency, udsc-jet mistag rate and the total number of the expected top pair semileptonic events in the final selected sample. By minimizing the log-likelihood function, the constraints on the b-tagging efficiency and udsc-jet mistag rate can be imposed.

It was found that the method results in a large correlation between measured b-tagging efficiency and udsc-mistag rate. Nonetheless, the method provides an independent and valuable cross-check for the b-tagging calibration based on conventional P_T^{Rel} and *System8* methods [11].

6. Physics Studies and Validation of Tracking Software with First LHC Collision Data

CMS recorded the first LHC collisions at center of mass energies 900 GeV and 2.36 TeV in December 2009. These data have been used to calibrate detector and study its performance, verify reconstruction software and perform physics studies at 900 GeV and the highest collision energy (2.36 TeV) attained so far. To study the performance of the CMS tracking system, a number of hadronic resonances, such as K^0s and □□ were reconstructed, using the excellent CMS tracking devices. The Young Investigator Group contributed to the effort of tracking studies and performed the reconstruction of K*(892)□ and □□ resonances. The K*(892)□ resonance is reconstructed in the K*(892)□□$K^0{}_S$□□□decay channel. Events were required to contain a reconstructed primary vertex consisting of more than two tracks and having a fit probability greater than 0.5%. The K*(892)□ state is a strong resonance with an ext extremely short lifetime, hence both $K^0{}_S$ and □□□candidates are expected to originate from the primary vertex. The $K^0{}_S$ candidate is reconstructed as a pair of oppositely charged tracks, compatible with the common neutral vertex, which is well separated from the primary vertex. Furthermore, the invariant mass of the neutral vertex, assuming that both tracks are charged pions, must be consistent with the PDG value of the $K^0{}_S$ mass. A charged track is regarded as a pion candidate from the K*(892)□□$K^0{}_S$□□□decay if it is not already assigned to the $K^0{}_S$ and belongs to the primary vertex. Figure 3 shows an invariant mass distributions of the K*(892)□□$K^0{}_S$□□ candidates in collision data collected by CMS. Mass spectrum is fitted with the superposition of signal and background distributions. A relativistic Breit-Wigner formula was used to fit the signal distribution and to determine the yield and the mass of the reconstructed state.

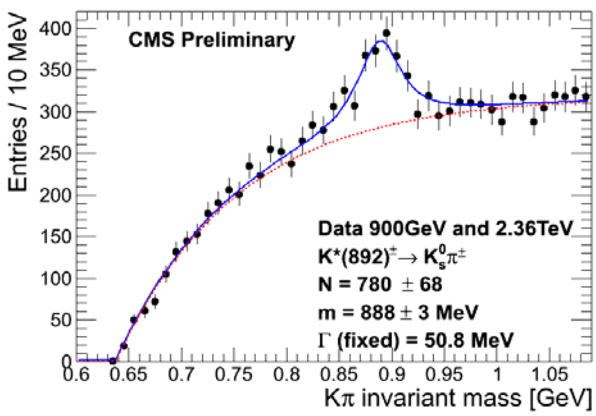

Figure 3: The $K^0{}_S$□$^{\pm}$ invariant mass distributions in the proton-proton collision data.

The fit of the signal distribution is performed with the normalization and mass as free parameters. The width is fixed to the K*(892)□ PDG value of 50.8 MeV. The fitted mass value of 888□3 MeV is consistent within statistical errors with the K*(892)□□PDG value of 891.66□0.26 MeV. The number of reconstructed K*(892)□□candidates in collision data is 780□68. We also studied kinematic distributions of the the reconstructed K*(892)□□candidates. The sideband background subtraction technique was exploited to reconstruct transverse momentum and pseudorapidity distribution of K*(892)□□candidates. The shapes of the kinematic distributions observed data are in good agreement with Monte Carlo predictions.

The reconstruction of the □□ particle is particularly challenging since both a charged particle and a neutral particle decay have to be detected in the tracker. The DESY group has performed an independent analysis which cross-checked the official CMS result. The 2009 CMS collision dataset has been analyzed to reconstruct a sample of 50□9 charged □□baryons with a mass resolution of 4.0□0.8 MeV and a central mass value of 1322.8□0.8 MeV consistent with the PDG value.

The analysis of the K*(892)□□state performed by the members of the Young Investigator Group has been included in the ***Physics Analysis Summary*** [12] prepared by the

CMS *Tracking Group* to describe performance of the CMS tracking system in the first collision data.

7. Cooperation with the University Partner (Karlsruhe University and Karlsruhe Institute of Technology)

The Young Investigator Group has close ties to the CMS Group of the Karlsruhe Institute of Technology (KIT). The common working meetings between the members of the Young Investigator Group and members of the CMS KIT group are held every month to digest the progress of physics analysis and development of the analysis tools. As a result of close cooperation between two groups, data-driven methods for evaluation of the *Z+jets* and QCD backgrounds for the MSSM Higgs boson searches via Higgs boson decays to tau leptons have been developed. Currently both groups closely cooperate on the development of the methods to study performance of the b-tagging algorithms based on the reconstruction of secondary vertices resulting from the *B* and *D* meson decays. These methods will be used in the analysis of the data collected at center of mass energy of 7 TeV for validation and calibration of the b-tagging algorithms using information about secondary vertices in jets.

8. Analysis Notes, Publications and Presentations

For the reference period the Young Investigator Group has contributed to the following papers:

1) *„Tracking and Vertexing Results from First Collisions“* CMS Physics Analysis Summary TRK-10-001;
2) „*Search for the neutral MSSM Higgs Bosons via decays to tau leptons in the Dimuon Channel*“, CMS Analysis Note 2010/203 (in preparation);
3) *„Performance of the CMS Tracking Detector in First Collisions at LHC“,* paper submitted to Nuclear Instruments and Methods;

The following presentations have been given by the members of the group:

1) „*First Results from CMS Fast Beam Condition Monitor BCM1F*“, R. Walsh, talk given at annual DPG Meeting in Bonn, March 2010;
2) *„Search for the MSSM Higgs Boson Decaying to Tau Leptons in the Dimuon Channel“*, A. Raspereza, talk given at the meeting of the M-Tau-Tau Working Group of the Helmholtz Alliance, Karlsruhe Feb 17th 2010;
3) *„First Collisions with the CMS Detector at LHC“,* A. Raspereza, talk given at DESY Research Seminar, Hamburg March 8th 2010, Zeuthen March 9th.

9. Financial Status

The financial status report is attached as a separate MS Excel document.

References

[1] CMS Collaboration, R. Adolphi *et al.*, ***JINST* 3** S08004 [SPIRES].
[2] A. Bell et al., **Nucl. Instrum. Meth. A614**: **433-438, 2010.**
[3] M. V. Acosta, D. Colling and A. Nikitenko, CMS Analysis Note, AN-2009/143.
[4] G. Cerati and S. Malvezzi, CMS Analysis Note, AN-2009/171.
[5] ALEPH Collaboration, A. Heister *et al.,* **Phys. Lett. B512: 30-48, 2001.**
[6] G. Barker *et al.,* DELPHI-2002-069 CONF 603.
[7] SLD Collaboration, Kenji Abe et al., **Phys. Rev. Lett. 84: 4300-4304, 2000.**
[8] OPAL Collaboration, G. Abbiendi et al., **Eur.Phys.J.C29:463-478,2003.**
[9] A. Rizzi, F. Palla, G. Segneri, CMS Note 2006/019.
[10] T. Speer et al., CMS Analysis Note AN-2009/085.

[11] V. E. Bazterra, T. Speer, CMS Conference Report, CMS CR – 2009/032.
[12] K. Burkett *et al.,* CMS Physics Analysis Summary TRK-10-001.