

## Résumé

# Recherche de bosons de Higgs supplémentaires de haute masse se désintégrant en paire de taus dans l'expérience CMS au LHC à l'aide du *Machine Learning*

Malgré plusieurs décennies de prédictions expérimentalement vérifiées, le modèle standard (SM) souffre de lacunes que des théories allant au-delà (BSM, *Beyond SM*) tentent de combler. L'une d'entre elles, l'extension supersymétrique minimale du modèle standard (MSSM), prédit l'existence de cinq bosons de Higgs dont trois neutres,  $h$  correspondant au boson découvert en 2012 et  $H$  et  $A$  supplémentaires. Leur phénoménologie au LHC motive l'étude des événements avec une paire de leptons  $\tau$ .

Dans ce contexte, les événements récoltés par la collaboration CMS de 2016 à 2018 lors des collisions de protons du LHC avec une énergie dans le centre de masse de 13 TeV, correspondant à une luminosité intégrée de  $137 \text{ fb}^{-1}$ , sont exploités dans cette thèse.

Les jets sont des objets physiques omniprésents lors des collisions au LHC. Leur calibration à l'aide d'événements contenant un photon et un jet ( $\gamma + \text{jet}$ ) est présentée ainsi que les résultats obtenus pour l'année 2018. Cette étude, également réalisée pour l'année 2017-UL, est directement utilisée dans la calibration officielle de la collaboration CMS.

Les jets issus de quarks  $b$  sont utilisés afin d'augmenter la sensibilité de l'analyse des événements avec une paire de  $\tau$  aux bosons  $H$  et  $A$ . Aucun excès significatif par rapport aux bruits de fond attendus n'est observé. Des limites d'exclusion sur le produit de la section efficace de production de  $H$  et  $A$  avec leur rapport de branchement aux  $\tau$  sont données en fonction de leur masse pour les modes de production par fusion de gluon ou en association avec des quarks  $b$ . Ces limites sont comprises entre 10 pb à 110 GeV et  $10^{-3}$  pb à 3,2 TeV pour la fusion de gluons et entre 0,4 pb à 110 GeV et  $8 \times 10^{-4}$  pb à 3,2 TeV en association avec des quarks  $b$ . Dans le scénario  $M_h^{125}$ , ces limites se traduisent en une région d'exclusion dans le plan  $(m_A, \tan \beta)$ . Les valeurs de  $m_A$  inférieures à 600 GeV sont exclues. Cette limite passe à 2 TeV pour  $\tan \beta \gtrsim 50$ . Dans le scénario  $M_{H_1}^{125}(\text{CPV})$ , cette région est donnée dans le plan  $(m_{H^\pm}, \tan \beta)$ . Les valeurs de  $m_{H^\pm}$  inférieures à 400 GeV sont exclues. Lorsque  $\tan \beta \simeq 20$ , l'exclusion s'étend jusqu'à  $m_{H^\pm} \simeq 1,4 \text{ TeV}$ .

La reconstruction de la masse d'une paire de  $\tau$  est cruciale pour les analyses où un boson  $Z$  ou de Higgs se désintègre en  $\tau^+ \tau^-$ . Les neutrinos issus des désintégrations des  $\tau$ , invisibles dans le détecteur, compliquent cette tâche. Le *Machine Learning* apporte une solution. Le modèle obtenu dans cette thèse permet de reconstruire la masse d'une paire de  $\tau$  de 50 GeV à 800 GeV avec une résolution de 20 % à 50 GeV, 26 % à 250 GeV et 22 % à 800 GeV. Ce modèle est 60 fois plus rapide et propose une meilleure description du boson  $Z$  que l'algorithme SVFIT actuellement utilisé au sein de la collaboration CMS.

## Mots clés

Higgs, CMS, Jets, Calibration, Tau, Reconstruction de la masse, Réseau de neurones, Régression.



# Abstract

## Search for additional heavy Higgs bosons decaying to tau lepton pair in the CMS experiment at LHC with Machine Learning techniques

Despite decades of correct prediction after correct prediction, physicists know that the Standard Model (SM) does not show us the whole picture. Among the various extensions going beyond the SM (BSM) is the Minimal Supersymmetric extension of the SM (MSSM) in which there are three neutral Higgs bosons,  $h$  corresponding to the observed one discovered in 2012 and  $H$  and  $A$  being additional with respect to the SM. The MSSM phenomenology motivates the focus on events containing a di- $\tau$  pair, so is the MSSM  $H/A \rightarrow \tau\tau$  analysis the core of this thesis.

The search for  $H$  and  $A$  is performed in this thesis on the data collected with the CMS detector from 2016 to 2018 on proton collisions at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of  $137 \text{ fb}^{-1}$ .

Jets are complex physics objects obtained in the proton collisions occurring at the CERN LHC. Their calibration by the study of events containing a photon and a jet ( $\gamma + \text{jet}$ ) and the corresponding results on 2018 data are introduced. These results, as well as the ones obtained on 2017-UL data, are used in the official CMS jet calibration.

Indeed, jets originating from  $b$  quarks are used in the MSSM  $H/A \rightarrow \tau\tau$  analysis to enhance its sensitivity to neutral MSSM Higgs bosons produced in association with  $b$  quarks. No significant deviation above the expected background is observed. Model-independent limits are set on the product of the cross section and branching fraction for the production via gluon-fusion or in association with  $b$  quarks. These limits range from 10 pb at 110 GeV to  $10^{-3}$  pb at 3,2 TeV for gluon-fusion and from 0,4 pb at 110 GeV to  $8 \times 10^{-4}$  pb at 3,2 TeV for  $b$ -associated production. In the  $M_h^{125}$  scenario, these limits translate into an exclusion region in the  $(m_A, \tan\beta)$  plane. Values of  $m_A$  below 600 GeV are excluded and this goes up to 2 TeV for  $\tan\beta \gtrsim 50$ . In the  $M_{H_1}^{125}(\text{CPV})$  scenario, the region is given in the  $(m_{H^\pm}, \tan\beta)$  plane. Values of  $m_{H^\pm}$  below 400 GeV are excluded and this goes up to 1,4 TeV for  $\tan\beta \simeq 20$ .

Reconstruction of di- $\tau$  mass in a faster and more accurate way than the existing methods is crucial to test any theory involving Higgs boson and Z boson which are decaying to  $\tau^+\tau^-$ . However, it is an arduous task due to existence of neutrinos as decay product of each  $\tau$  lepton which are invisible to detectors at LHC. Deep neural networks (DNN) are applicable for this task. The reconstruction of the di- $\tau$  mass by a DNN is achieved in this thesis from 50 GeV to 800 GeV with a 20 % resolution at 50 GeV, 26 % at 250 GeV and 22 % at 800 GeV. This DNN is 60 times faster and better at describing the Z boson than the SVFIT algorithm currently used in CMS.

### Keywords

Higgs, CMS, Jets, Calibration, Tau, Mass reconstruction, Neural Network, Regression.

