Validation note for the CMS same-sign dilepton analysis: CMS_B2G_12_012

Daniele Barducci and Cédric Delaunay

LAPTh, Université Savoie Mont Blanc, CNRS, B.P. 110, F-74941 Annecy-le-Vieux, France

This note contains details of the implementation and validation of the CMS Search for top partners with charge 5/3 in the same-sign dilepton final state [1] within the MadAnalysis5 [2–4] framework. This search looks for a signal with two isolated same-sign leptons with \( p_T > 30 \text{ GeV} \) and \( H_T > 900 \text{ GeV} \). The implementation has been performed with MadAnalysis5 v1.2 using the standard version of the Delphes v3.2.0 [5] package.

I. SIMULATION DETAILS

The CMS collaboration provides several differential distributions, both in the published paper and in the additional material available at [6], together with a summary plot with the expected and observed 95% confidence level (CL) limit for the \( X_{5/3} \) QCD pair production cross section times branching fraction in a same-sign dilepton final state. A table with the number of signal events for \( m_{X_{5/3}} = 800 \text{ GeV} \) for the three same-sign categories \( ee, e\mu \) and \( \mu\mu \) separately is also provided. These category are however combined in a unique signal region when computing the exclusion CL. Finally, for a \( X_{5/3} \) pair decaying into a same-sign dilepton final state, the signal selection efficiency is claimed to be between 10% and 13% for \( m_{X_{5/3}} \in [750, 1000] \text{ GeV} \).

The signal process consists in the production of a \( X_{5/3} \) pair via QCD interactions. The \( X_{5/3} \) is assumed to couple only with the third generation of Standard Model (SM) quarks, and it can therefore only decay via charged current interactions through the process \( X_{5/3} \to W^+ t \).

The signal process has been generated with MG5\_aMC v2.3.0 [7] using the model file Rho resonances in composite Higgs model (including top partners) which has been implemented in the UFO format [8] through the Feynrules package [9], by the authors of [10]. The model is publicly available at the HEPMDB [11] web-page [1] and it has been modified with the inclusion of the QCD interactions for the extra quarks, adding the following lines in the vertices.py model file:

\[
V_{133} = \text{Vertex(name = 'V_{133}',}\n\text{particles = [ P.X53__tilde__, P.X53, P.g ],}\n\text{color = [ 'T(3,2,1)' ],}\n\text{lorentz = [ L.FFV1 ],}\n\text{couplings = {(0,0):C.GC_7)}}.
\]

The MG5\_aMC v2.3.0 syntax that has been used to generate the matrix element is:

\[
\text{generate p p > X53 X53~ / rho0 rho+ rho- @1}\n\text{add process p p > X53 X53~ j / rho0 rho+ rho- @2}\n\text{add process p p > X53 X53~ j j / rho0 rho+ rho- @3}
\]

and 12 signal samples, corresponding to \( m_{X_{5/3}} = 350, 450, 550, 600, 650, 700, 750, 800, 850, 900, 950 \) and 1000 GeV, have been generated.

Decays of unstable particles, including the \( X_{5/3} \) quark, parton showering and hadronisation have been performed with PYTHIA v6.4 [12] implemented in MG5\_aMC v2.3.0 through the pythia-pgs package. The MLM matching scheme [13, 14] was used to merge the 0, 1 and 2 jets samples. The samples have then been passed through Delphes v3.2.0 [5] for a fast detector simulation, using a MadAnalysis5
tuned card. In the run_card.dat the value ickkw has been fixed to the value of 1, while the value of the jet merging parameter xqcut has been tuned according to the ones reported in Tab. I. In the pythia_card.dat the flag showerkt=T has been added, while the value of QCUT has been fixed equal to xqcut.

Finally, the QCD pair production cross sections have been normalised to the next-to-next-to-leading-order (NNLO) predictions as reported on the CMS twiki web-page [6]. These values are indicated in Tab. I.

<table>
<thead>
<tr>
<th>X_{5/3} mass [GeV]</th>
<th>xqcut value</th>
<th>σ QCD [fb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>350</td>
<td>75</td>
<td>5.2×10^3</td>
</tr>
<tr>
<td>450</td>
<td>75</td>
<td>1.2×10^3</td>
</tr>
<tr>
<td>550</td>
<td>75</td>
<td>315.2</td>
</tr>
<tr>
<td>600</td>
<td>70</td>
<td>173.8</td>
</tr>
<tr>
<td>650</td>
<td>70</td>
<td>100.0</td>
</tr>
<tr>
<td>700</td>
<td>70</td>
<td>58.6</td>
</tr>
<tr>
<td>750</td>
<td>70</td>
<td>35.0</td>
</tr>
<tr>
<td>800</td>
<td>65</td>
<td>21.3</td>
</tr>
<tr>
<td>850</td>
<td>65</td>
<td>13.2</td>
</tr>
<tr>
<td>900</td>
<td>50</td>
<td>8.3</td>
</tr>
<tr>
<td>950</td>
<td>50</td>
<td>5.3</td>
</tr>
<tr>
<td>1000</td>
<td>50</td>
<td>3.3</td>
</tr>
</tbody>
</table>

TABLE I. xqcut values used in the signal simulation and NNLO cross sections adopted for the signal normalisation for the simulated X_{5/3} masses.

II. VALIDATION

The analysis selection requires isolated leptons, which are defined by computing the scalar sum of the \( p_T \) of all neutral and charged reconstructed particles within a cone of size \( \Delta R \) around the selected lepton. This sum is divided by the \( p_T \) of the lepton, which is considered isolated if the ratio is below 0.15 (0.2) in a cone \( \Delta R = 0.3 \) (0.4) for electrons (muons). A category of loose leptons is also defined, where the ratios are increased to 0.60 (0.40) for electrons (muons). Jets are reconstructed with FastJet [15], via an anti-\( k_T \) [16] algorithm, with distance parameter 0.5 and are required to have \( p_T > 30 \) GeV.

In the experimental analysis, jet substructure algorithms to tag boosted jets from top and W decays are also used. Since at the moment this is not a feature implementable in MadAnalysis5, we did not use any jet substructure technique for our analysis.

The signal region is then obtained by applying the following set of cuts:

- At least two isolated same-sign leptons as defined above. If more than one same-sign pair is present, the one with the highest leading lepton \( p_T \) is considered,

- Dilepton Z boson veto: \( M_{ee} < 76 \) GeV or \( M_{ee} > 106 \) GeV,

- Trilepton Z boson veto: \( M_{ll} < 76 \) GeV or \( M_{ll} > 106 \) GeV, where \( M_{ll} \) is the invariant mass of either one of the selected leptons and any other same flavour opposite sign loose lepton in the event, with \( p_T > 15 \) GeV,

- \( N_C \geq 7 \), where \( N_C \) is the number of the reconstructed constituents of the event (leptons and jets),

- \( H_T >900 \) GeV, where \( H_T \) is the scalar sum of the \( p_T \) of all the selected jets and leptons in the event.

We now compare the official CMS results with the ones of our simulation. However, since we did not use any jet substructure technique for our analysis, we refrain to compare our jet distributions with the official ones.
A. Published plots

After the signal selection the analysis report 6.8±2.1 background events and 9 observed events. A signal is therefore excluded if it predicts more than 10.1 events in the unique signal region of this search. We obtain a limit on the $X_{5/3}$ mass of 785 GeV, see Fig. 1 (right panel), in good agreement with the official CMS result of 800 GeV.

![Distribution of $H_T$ for all channels combined, after the full selection except for the $H_T$ requirement itself for $m_{X_{5/3}} = 800$ GeV (left) and number of signal events in function of the $X_{5/3}$ mass (right). The horizontal red line in the right plot represents the number of signal events excluded at 95% CL, as obtained with the module exclusion_CLs.py available within the MadAnalysis5 package.](image-url)

FIG. 1. Distribution of $H_T$ for all channels combined, after the full selection except for the $H_T$ requirement itself for $m_{X_{5/3}} = 800$ GeV (left) and number of signal events in function of the $X_{5/3}$ mass (right). The horizontal red line in the right plot represents the number of signal events excluded at 95% CL, as obtained with the module exclusion_CLs.py available within the MadAnalysis5 package.

<table>
<thead>
<tr>
<th>Signal Region</th>
<th>CMS official results</th>
<th>MA5 results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ee$</td>
<td>2.1</td>
<td>2.3</td>
</tr>
<tr>
<td>$\mu\mu$</td>
<td>2.8</td>
<td>2.1</td>
</tr>
<tr>
<td>$e\mu$</td>
<td>4.7</td>
<td>4.2</td>
</tr>
</tbody>
</table>

TABLE II. Signal events in the three selection channels for the official CMS results and our simulation for $m_{X_{5/3}} = 800$ GeV.
B. Unpublished plots

In this Section we compare other unpublished plots, publicly available in additional material at on the CMS twiki web-page [6].

FIG. 2. Distribution of $H_T$ for all channels combined after the same-sign selection, the $Z$ veto and a requirement of at least two jets, for $m_X=800$ GeV

FIG. 3. Distribution of the leading lepton $p_T$ for the $ee$ and $\mu\mu$ selection channels after the same-sign selection, the $Z$ veto and a requirement of at least two jets, for $m_X=600$ GeV
FIG. 4. Distribution of the leading lepton $p_T$ for the $\mu \mu$ selection channels and all channels combination after the same-sign selection, the $Z$ veto and a requirement of at least two jets, for $m_{X/3} = 600$ GeV.

FIG. 5. Distribution of $H_T$ for $ee$, $\mu \mu$, $e\mu$ selection channels and their combination after the same-sign selection, the $Z$ veto and a requirement of at least two jets, for $m_{X/3} = 600$ GeV.
ACKNOWLEDGMENTS

We thank the CMS collaboration, in particular Aram Avetisyan and Devdatta Majumder, for useful discussions during the completion of this work.