COMPUTING METHODS IN HEP Home exam Spring 2016

(To be returned by 24.00 on Friday 29.5.2016)

One of the main goals of the LHC collider is to discover the Higgs boson. The CMS experiment was designed to be especially sensitive to a SM Higgs boson in the $H \rightarrow \gamma\gamma$ decay. Your task is to study the discovery potential of a SM like Higgs boson with a mass $m_H = 750$ GeV/c^2 and width 45 GeV in this decay channel. Assume a cross section of 6.2 fb.

The main backgrounds for this channel are the γ^*/Z, $\gamma\gamma$ from QCD box and born diagrams, QCD multi-jet events with $\hat{p}_T > 50$ GeV/c and γ+jets events. You can assume (with a good precision) that the total background shape is similar to the shape of the γ+jets background and that the total background cross-section is 4 times that of the γ+jets background. Furthermore, you can assume that the γ+jets background consists only of the $f\bar{f} \rightarrow \gamma\gamma$ and of the $gg \rightarrow \gamma\gamma$ processes.

Assume, that the LHC is running with a center-of-mass energy of 13 TeV and that the collected amount of data is 100 fb^{-1}.

1. Create generator-level datasets for

   a) signal

   b) background

   The events are triggered with a HLT_DoublePhoton60 trigger with a symmetric threshold of 60 GeV for both photons: $|\eta| < 2.5$ and $p_T > 60$ GeV/c. Fill the datasets with events passing the trigger. What is the trigger efficiency? (6 points)

2. Analyze your data.

   a) Identify the signal photons. Take as the photon candidates the photons in the event. Apply the following operations and standard selection criteria to each of the photon candidates:

      – Simulate the γ's measurement error by applying 1% Gaussian smearing to their momenta and 2 mrad Gaussian smearing to their angles θ and ϕ.

      – Selection of the two signal photons: Require that the photon candidate with the highest p_T satisfies $p_T > 200$ GeV/c and that the photon candidate with the second highest p_T satisfies $p_T > 150$ GeV/c.

      Report the number of events passing the selection. (6 points)

   b) Reconstruct the invariant mass $M_{\gamma\gamma}$. Produce a histogram of $M_{\gamma\gamma}$ with correct normalization (i.e. that each entry in the histogram corresponds to a correct amount of cross-section in fb). Plot two invariant mass distributions in the same figure: the background only, and the sum of signal and background. (6 points)

   c) Fit the signal+background and background only $M_{\gamma\gamma}$ histograms with appropriate functions over an appropriate mass range.

      Choose a reasonable mass window and calculate the number of signal and background events for integrated luminosity of 100 fb^{-1}. Estimate the statistical significance of the signal peak for these two cases with the naive expression $N_S/\sqrt{N_B}$, where N_S and N_B are the number of signal and background events, respectively. (6 points)

3. Reflect on your simulation and your results. If the data were real data, you have to present your work to the world-wide scientific community. Reason why the scientific community should believe in your results. Can you convince yourself of the finding? Why/why not? How would you improve your study to make it more realistic? (6 points)

Please make a gzipped tar-ball all your files (except the datasets) and return it by email to: Sami.Lehti@helsinki.fi. Include in your answer a written description how you proceed and where you argument any non-trivial choice you make.