KIT-Fakultät für Physik
Institut für Experimentelle Teilchenphysik
Prof. Dr. Ulrich Husemann
Dr. Matthias Schröder
Daniela Schäfer, Michael Waßmer

Übungen zu Teilchenphysik I

Wintersemester 2017/18

Übungsblatt Nr. 10 Bearbeitung bis 11. Januar 2018

Search for New Physics at CMS

Data-Driven Background Estimation

Part I

This is the first part of a two parts exercise. Please keep the results of this part because the next one requires them as input. Setting up the code requires to copy one time a large file, cf. Section 2. **Please do this before the actual class to avoid delays.**

1 Introduction

Despite its success, the Standard Model (SM) of particle physics is still a very incomplete description of nature. First of all, gravity has so far not be included consistently into the SM framework, and although gravitational interactions are perfectly negligible at the subatomic level, this is certainly unacceptable for a fundamental theory. Even more, Dark Matter and Dark Energy, which together constitute approximately 95% of the total energy in the universe, lack candidates in the SM. These and other observations as well as theoretical considerations, such as the apparent need for fine tuning of radiative corrections to the Higgs-boson mass, demonstrate that the SM has to be understood as a low-energy approximation of some more fundamental theory. Numerous extensions to the SM have been proposed that address the mentioned shortcomings and often predict the existence of new particles. It is one of the primary missions of the physics programme at the Large Hadron Collider (LHC) to search for signals of any new physics processes.

1.1 Supersymmetry

A prominent example is Supersymmetry (SUSY), which postulates the invariance of the Lagrangian under transformations that relate fermionic and bosonic degrees of freedom (*supersymmetric transformations*) and generally requires the introduction of new particles as a consequence. The existence of these additional particles has many exciting consequences: Most strikingly, they can provide excellent candidates for Dark Matter. Moreover, the presence of supersymmetric partner particles in virtual loops can cancel the otherwise large corrections to the Higgs-boson mass; as a consequence, the theory becomes less dependent on the exact values of its parameters (it becomes *more natural*) — a feature typically considered desirable. The presence of supersymmetric particles in virtual loops can also lead to common gauge-coupling strengths when extrapolated to higher scales, which hints to some underlying unified theory. Furthermore, SUSY models often include mechanisms triggering electroweak symmetry breaking. It is worth pointing out that SUSY has been introduced for purely theoretical arguments since it is the only possible further symmetry that does not violate the gauge invariance of the SM[1]; it has not been constructed to solve the problems of the SM!

The simplest SUSY model that is not in contradiction with the measurements is the Minimal Supersymmetric Standard Model (MSSM). Here, a spin-0 partner-particle is introduced for the left- and right-handed chiral states of each SM fermions and a spin-1/2 partner-particle for each SM gauge boson[2]. The supersymmetric partner particles are denoted as the SM particles but with a tilde, e.g. $\tilde{e}_{L/R}$ for the electron partners, and they are named with a prepended 's' in case of fermions, e.g. selectron for the electron partner, and with an appended 'ino' in case of the vector bosons, e.g. Wino for the W boson. The physical particles are in general superpositions of the sfermions $\tilde{f}_{L/R}$, called $\tilde{f}_{1/2}$, and of the gauginos, called charginos $\tilde{\chi}^{\pm}_{1,2}$ and neutralinos $\tilde{\chi}^{0}_{1,2,3,4}$. In the MSSM, one additionally postulates conservation of a discrete quantum number *R-parity* that is computed from the spin of the considered particles. Its conservation has the phenomenological consequence that SUSY particles can only be produced in pairs and thus that the lightest supersymmetric particle (LSP) is stable. In realistic models, the LSPs are typically electrically neutral and only weakly interacting, such as neutrinos, and form the Dark Matter candidates.

However, so far, SUSY particles have not yet been observed, and hence, they must have higher masses than the known SM particles (*SUSY is broken*). Theoretical arguments favour masses not much higher than the TeV range, and thus, SUSY particles could be directly produced in the proton-proton collisions at the LHC. In many SUSY models, gluino-gluino ($\tilde{g}\tilde{g}$), gluino-squark ($\tilde{g}\tilde{q}$), and squark-squark ($\tilde{q}\tilde{q}$)

[1] SUSY is the only non-trivial extension of the Poincaré group.

[2] The Higgs sector is a little more complicated since the invariance under supersymmetric transformation requires the introduction of a second Higgs doublet, resulting in five physical Higgs bosons, two of which are charged.

pair production can occur with particularly large cross sections compared to other SUSY-production channels. The squarks and gluinos will predominantly decay into coloured SM particles and pairs of stable LSPs.

A detailed review of SUSY can be found e.g. in [1, 2].

1.2 Search for Supersymmetry at CMS

In general, there are usually SM processes with the same signature (SM background) as the sought-for new physics processes. Therefore, one can only claim to observe a new process when the observed number of candidate events is significantly larger than the expected number of SM background events. Hence, precise knowledge of the expected SM background is crucial when searching for new-physics processes.

In Fig. 1, cross sections of different SM processes at the LHC are depicted. Expected SUSY cross sections are typically somewhere in the $< 1\,\mathrm{pb}$ regime, i.e. orders of magnitudes smaller than the SM background. The typical strategy to cope with this situation is to

- first, apply suitable event selection criteria to reduce the SM background compared to the expected signal;
- then, gain a precise understanding of the residual SM backgrounds to be sensitive to tiny excesses in the data.

A typical challenge of determining the residual background is the fact that the simulation is often associated with a rather large uncertainty in these cases because the signal regions are extreme phase-space regions, e.g. with a large number of jets where QCD predictions are difficult. Therefore, very often, backgrounds are determined from data (in a *data-driven* way).

Data-driven background estimation As a simple example, imagine that Z($\nu\nu$)+ jets events, i.e. Z+jets events with the Z boson decaying into two neutrinos, cf. Fig. 2 (left), were a background process to your search. In order to estimate their contribution in a data-driven way, one could take Z+jets events where the Z boson decays into two muons, cf. Fig. 2 (right). The Z-boson decay does not depend on the rest of the event, therefore one can expect that the kinematic properties of the hadronic part of the event is the same as for the Z($\nu\nu$) + jets events. Hence, in principle, one can apply the signal event selection to Z($\mu\mu$) + jets, omit the two muons in the event reconstruction, and use the selected events as prediction for the Z($\nu\nu$) + jets events. (Corrections have to be applied to account for the different

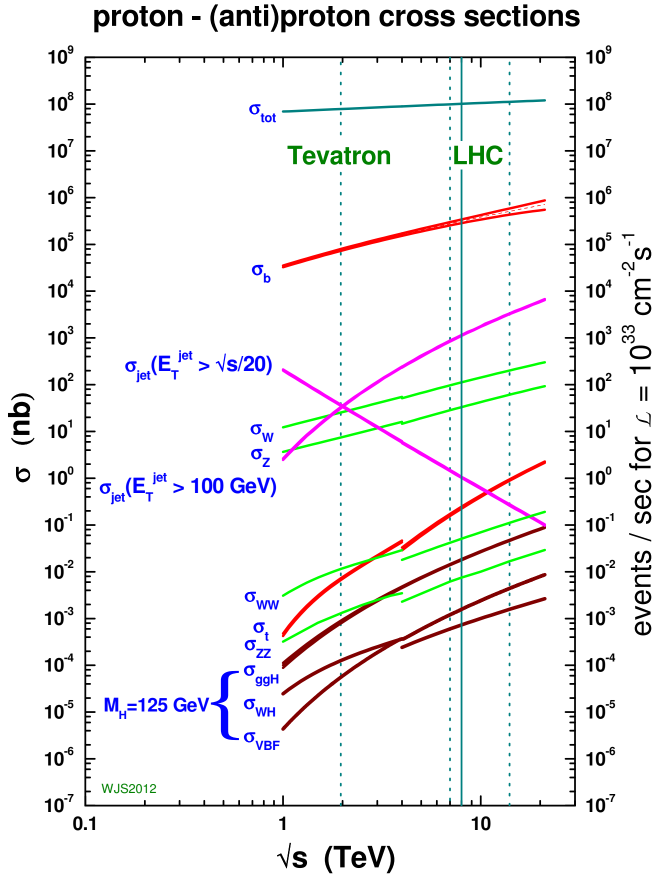

Figure 1: SM cross sections as a function of the pp/pp collision centre-of-mass energy. Taken from [3].

branching ratios of the Z boson and for muon acceptance and reconstruction efficiency effects.) The advantage of the method is that one does not have to rely on the simulation to estimate the difficult hadronic part of the events. One disadvantage of this method, however, is the limited statistical precision because the branching ratios $\mathcal{B}(\mathrm{Z} \to \mu\mu) < \mathcal{B}(\mathrm{Z} \to \nu\nu)$.

In this exercise, we will study a typical search for new physics beyond the SM, following an analysis performed by the CMS collaboration [4]. The search is a generic search for new physics, but the signal signature is motivated by the expectations from R-parity conserving SUSY and targets $\tilde{g}\tilde{g}$, $\tilde{g}\tilde{q}$, and $\tilde{q}\tilde{q}$ pair production. Thus, the signature we are looking for in the detector consists of several jets with high transverse momentum p_{T}, large missing transverse momentum due to the LSPs, and no leptons, as illustrated in Fig. 3. The detector signature of a candidate event, found in 2011, is depicted in Fig. 4. Accordingly, the sensitive variables of the analysis are:

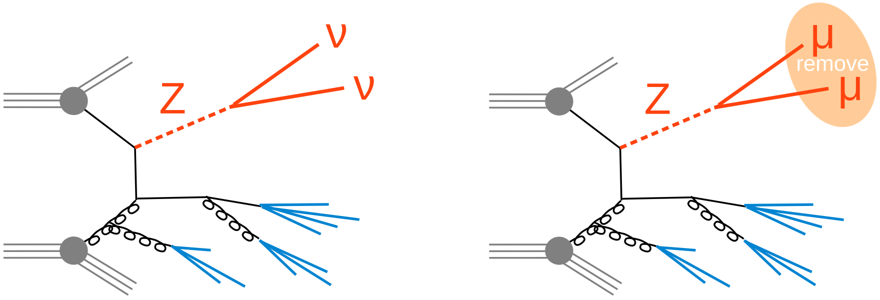

Figure 2: Data-driven estimation of $Z(\nu\nu)$ + jets background from $Z(\mu\mu)$ + jets events.

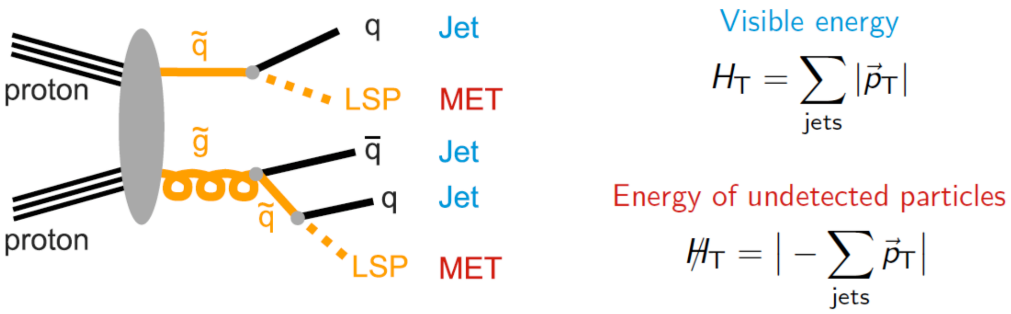
q̃
q
Jet
proton
LSP
MET
g̃
q̄
Jet
q
Jet
proton
q̃
LSP
MET
Visible energy
$H_T = \sum_{\text{jets}} |\vec{p}_T|$
Energy of undetected particles
$\not{H}_T = |-\sum_{\text{jets}} \vec{p}_T|$

Figure 3: Illustration of the signal process and the sensitive variables.

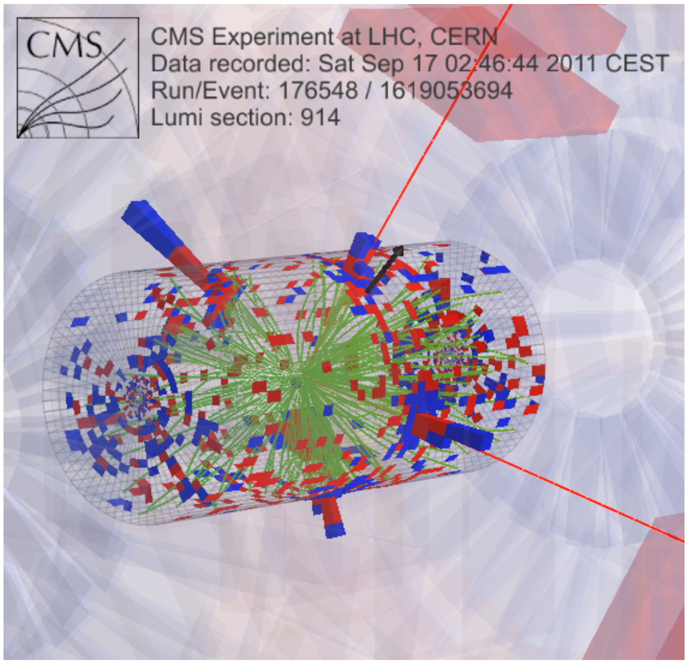
CMS
CMS Experiment at LHC, CERN
Data recorded: Sat Sep 17 02:46:44 2011 CEST
Run/Event: 176548 / 1619053694
Lumi section: 914

Figure 4: Candidate event found in the 2011 CMS data.

- $\mathrm{N_{jets}}$, the number of jets;
- H_T, the scalar sum of the p_T of all jets, a measure of the visible energy; and
- $\not{H}_\mathrm{T}$, the magnitude of the negative vectorial sum of the jet p_T, a measure of the missing transverse momentum, i.e. the energy of the undetected particles.

The SM background is determined almost entirely from data, which is one of the key features of the analysis.

The focus of the exercise will be on understanding (some of) the data-driven background determination methods. (If you want to learn about the other methods, have a look at the original publication cited above.) The exercise is performed with a series of simple Root-level C++ scripts and plain C++ classes. Some of them are incomplete, and you will be instructed on how to complete them. You will also be asked some questions meant to gauge your understanding of the topics being discussed. The data you will use have been prepared in a reduced Root ntuple format suitable for easy, interactive access and a fast turn-around time. The events store kinematic and other properties of higher-level objects, such as muons and jets, that have been reconstructed from the detector signals. In case of data, events have been preselected according to certain quality criteria that reject e.g. events affected by detector noise or malfunctioning detector components.

2 Preparation: Setting up the Code

The code and input files for this exercise are provided in

```
/home/staff/mschrode/TP1_WS17/susy/susy_ex_part1.tar.gz
```

Copy the archive to your working area. **Please do this before the actual class because the archive is rather large and copying will take a while.** Extract the archive:

```
tar -zxvf susy_ex_part1.tar.gz
```

This will create the directory `susy_ex`, which contains all code and input data needed for this exercise. You then need to adjust the path to the input data, which is stored in

```
susy_ex/data/ntuples
```

To do so, modify the line `susy_ex/Utils/Sample.h` that says

```
// Path to ntuple directory. Adapt to your environment.
TString Sample::path_ = "<path>/susy_ex/data/ntuples/";
```

and set `<path>` to the path of your `susy_ex` installation. (You can go to `susy_ex` and type `pwd` to obtain the path.) Note that the last character of the `TString` must be a `/`.

3 Sample Definition and Composition

In this section, we will learn how to access the data and simulated events and how to apply the event selection criteria (*cuts*). We will investigate the data and compare it to the expected properties of the SM backgrounds. We will also get a feeling of how hypothetical SUSY signals look like.

3.1 Event reconstruction

At CMS, the recorded and simulated events are reconstructed using a 'particle-flow' algorithm. This algorithm reconstructs individual particles in each event, namely charged hadrons, photons, neutral hadrons, muons, and electrons, using the information from the tracker, the calorimeters, and the muon system. These particles are then clustered into jets using the anti-k_T algorithm; here, a size parameter of 0.5 has been used. Corrections are applied to account for the variation of the jet response with p_T and η. Contributions of particles from pile-up collisions are mitigated by discarding those charged particles that originate from vertices other than the primary vertex.

3.2 Event selection

Candidate events are required to pass a *baseline selection* defined as follows:

1. no well-reconstructed and isolated lepton (electron or muon) with $p_\mathrm{T} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$ are present (*lepton veto*);

2. $\mathrm{N_{jets}} \geq 3$, where only jets with $p_\mathrm{T} > 50\,\mathrm{GeV}$ and $|\eta| < 2.5$ are considered;

3. $H_\mathrm{T} > 500\,\mathrm{GeV}$, where again only jets with $p_\mathrm{T} > 50\,\mathrm{GeV}$ and $|\eta| < 2.5$ are considered;

4. $\not{H}_\mathrm{T} > 200\,\mathrm{GeV}$, where jets with $p_\mathrm{T} > 30\,\mathrm{GeV}$ and $|\eta| < 5.0$ (sic!) are considered; and

5. $\Delta\phi(\mathrm{jet}_{1,2},\vec{\not{H}}_\mathrm{T}) > 0.5$ and $\Delta\phi(\mathrm{jet}_3,\not{H}_\mathrm{T}) > 0.3$, where jet_n refers to the n-th jet ordered in p_T.

Later on, the selected events are distributed in 36 exclusive search bins that are defined on top of this baseline selection by even tighter $\mathrm{N_{jets}}$, H_T, and $\not{H}_\mathrm{T}$ selection criteria. But for now, we will stick with the baseline selection.

Consider the topology of the selected events.

- **Question 3.2.1** Why does one expect the jets in signal events to have relatively large p_T? Likewise, why are we looking for events with relatively large $\not{H}_\mathrm{T}$?
- **Question 3.2.2** Which SM processes result in the same final state as our signal process? Given these background processes, can you motivate the baseline selection? Which of the background processes do you expect to dominate at large $\mathrm{N_{jets}}$, at large H_T, and which at large $\not{H}_\mathrm{T}$? Don't do any calculation, just use your intuition.

In the following, we will perform several exercises that will help us to better understand the properties of the processes we are investigating and to check whether the answers to Question 3.2.2 are correct. It will also give us some introduction to the technical aspects of how to use the ntuples and perform an analysis with Root.

Before you start, make sure you have executed the initial setup of the code as described above.

3.3 Kinematic properties of the SM-background and signal processes

We will now investigate the SM background processes using simulated (Monte Carlo, MC) events and compare them to the expectations for new-physics signal processes. There are ntuples available for several different SM as well as potential SUSY-signal processes. They are listed in Table 1 together with their cross-sections and with the total number of simulated events. Beware that certain kinematic cuts have been applied at the generation of the MC samples, as stated in the comment field of Table 1. This is considered with the quoted cross sections, which are therefore termed *fiducial* cross sections.

- **Question 3.3.1** What do the different background events look like? Draw example Feynman diagrams.

The considered SUSY signal processes LM6 and LM9 are expected in a particular type of the MSSM where the SUSY breaking is mediated by gravity (termed: mSUGRA or cMSSM). In these models, among others, a unification of the particle

Table 1: Simulated background and SUSY signal (LM6 and LM9) processes together with their fiducial cross sections σ at $\sqrt{s} = 8\,\text{TeV}$ centre-of-mass energy and generated total number of events. The 'comment' column lists cuts applied at sample generation. The ID refers to an identifier used across this exercise.

ID	process	σ [pb]	MC events	comment
11	$\mathrm{W}(l\nu) + \text{jets}$	30.08	6 619 654	for $H_\mathrm{T} > 400\,\text{GeV}$
12	$\mathrm{t\bar{t}} + \text{jets}$	127.06	1 925 324	only semi- and dileptonic decays
13	$\mathrm{Z}(\nu\nu) + \text{jets}$	6.26	51 637	for $H_\mathrm{T} > 400\,\text{GeV}$
14	QCD	8630.0	44 443 155	for $H_\mathrm{T} > 500\,\text{GeV}$
21	LM6	0.502	51 282	$m_0 = 85\,\text{GeV}$, $m_{1/2} = 400\,\text{GeV}$: $m_{\tilde{g}} > m_{\tilde{q}}$
22	LM9	9.287	51 282	$m_0 = 1450\,\text{GeV}$, $m_{1/2} = 175\,\text{GeV}$: $m_{\tilde{q}} > m_{\tilde{g}}$

masses at very high energies is assumed, from which the masses at the electroweak scale are determined by RGE evolution. They are phenomenologically interesting because they depend only on very few parameters, and thus, it is relatively easy to probe the entire phase space. Therefore, they have been widely used as benchmark models before and during the first LHC run from 2010 to 2012. However, the most interesting parameter space has been excluded by now — among others, by this analysis!

The parameters m_0 and $m_{1/2}$ define the common mass of the scalars (sleptons, squarks, and Higgs bosons) and the gauginos and higgsinos at the high scale, respecitvely. The chosen parameter points LM6 and LM9 refer to 'low-mass' benchmark scenarios, which represent phase-space regions favoured by the pre-LHC data [5]. In case of the LM9, the squark masses at the electroweak scale are larger than the gluino masses. Therefore, signal events are dominated by gluino pair-production. Since gluinos decay via intermediate squarks into quarks, cf. Fig. 3, one expects a particularly large number of jets and large H_T in this case. In case of LM6, the squark pair-production dominates, and since squarks can decay directly to quarks, cf. Fig. 3, there are typically fewer jets but large $\not{H}_\mathrm{T}$.

In addition to the cuts at MC generation, the following preselection has been applied when writing the ntuples (Note that the number of events quoted in Tab. 1 refers to the total number of generated events before preseletion!):

- $H_\mathrm{T} > 300\,\text{GeV}$ and $\mathrm{N_{jets}} \geq 2$; and in addition
- $\not{H}_\mathrm{T} > 100\,\text{GeV}$ in case of QCD.

- **Question 3.3.2** Can you imagine why the generator-level and preselection cuts have been applied?

In order to analyse the simulated events, go to the `General` directory in your working area and execute the script `general1.C` with ROOT by typing

```
root -l -b -q general1.C+\(id,nEvts=-1\)
```

Note the `+` after the name of the script in the above command. This will tell ROOT to compile the script before execution. The argument `id` should be replaced with the values listed in the first column of Table 1, and `nEvts` is the number of processed events (`nEvts`< 0 means all events). We will start with analysing the $\mathrm{W}(l\nu) + \text{jets}$ sample, so set `id` to 11. Also, to save some time, run over 1 million events for example:

```
root -l -b -q general1.C+\(11,1E6\)
```

This will produce an output ROOT file `General_WJets.root` that contains various control distributions. Browse the file and investigate its content. You can conveniently plot the distributions with the script `plotSample.C`:

```
root -l -b -q plotSample.C+\(id\)
```

The argument `id` refers again to the sample, i.e. has to be set to 11 for the $\mathrm{W}(l\nu)$+jets sample. The script also stores the plots as png files in the current directory.

Discuss the distributions!

In this example, events are selected if no isolated lepton has been reconstructed or identified, i.e. if the baseline-selection step 1 in the list above is fulfilled. Do the shapes of the distributions meet your expectations?

- **Question 3.3.3** What is the reason for the lower cut-off in the $\mathrm{N}_{\mathrm{jets}}$ and H_{T} distributions?
- **Question 3.3.4** The $\not{H}_{\mathrm{T}}$ distributions drops towards very low $\not{H}_{\mathrm{T}}$; what is the reason for this?

Now open the two scripts we have just run, i.e. `general1.C` and `plotSample.C`, in your favourite editor. Familiarise yourself with the code.

- How is the event content accessed? How is the content read from file?
- Where are the histograms declared and filled? Where are they drawn?
- How is the lepton veto performed?

Make sure you understand what is being done because the following exercises will build up on this!

We still want to better understand the motivation for the baseline selection cuts. Therefore, we will compare the relevant kinematic distributions for the different processes. As you have noticed, so far only $\mathrm{N_{jets}}$, H_T, and $\not{H}_\mathrm{T}$ are calculated in `general1.C`. However, the baseline selection also includes the $\Delta\phi$ cuts (step 5), and thus, we also want to investigate the $\Delta\phi$ distributions. Please implement the code to compute $\Delta\phi$ to `general1.C` below the line

```
//>>> PLACE DELTA PHI COMPUTATION HERE
```

following the instructions and hints given there.

Now we have the full set of relevant kinematic distributions at hand, and we can compare them for the different processes. Thus, execute

```
root -l -b -q general1.C+\(id\)
```

again and again and again for the background samples `11` to `14` and the signal samples `21` to `22` in Table 1. You can automate this using simple `for` loops, e.g. in a bash shell do

```
for id in {11..14} 21 22; do root -l -b -q general1.C+\(${id}\); done
```

We will now compare the shapes of the different kinematic distributions of the different processes. This can be done with the prepared script `plotSampleComparison.C` by executing

```
root -l -b -q plotSampleComparison.C+
```

- **Question 3.3.5** How do the processes differ?
  - Explain the differences of the $\mathrm{N_{jets}}$ and H_T distributions of QCD and $\mathrm{Z}(\nu\nu) + \mathrm{jets}$ events.
  - Explain the different $\not{H}_\mathrm{T}$ distributions of the QCD and the other processes. Hint: what is the primary source of $\not{H}_\mathrm{T}$ in the $\mathrm{Z}(\nu\nu) + \mathrm{jets}$, $\mathrm{W}(l\nu) + \mathrm{jets}$ and $\mathrm{t\bar{t}} + \mathrm{jets}$ events? Where does $\not{H}_\mathrm{T}$ stem from in QCD events?
  - Explain the behaviour of the $\Delta\phi$ distributions. Compare again QCD with the other processes and focus on $\Delta\phi(\mathrm{jet}_1, \vec{\not{H}}_\mathrm{T})$ and $\Delta\phi(\mathrm{jet}_2, \vec{\not{H}}_\mathrm{T})$.

Now, modify the script `plotSampleComparison.C` to also plot the signal samples.

- **Question 3.3.6** What is the motivation for the different baseline-selection cuts? Which cuts aim at rejecting QCD events?

3.4 Expected event yields (simulation)

We will now use the simulated events to estimate the number of SM-background events after the full baseline selection. In the script `general2.C`, the corresponding cuts, i.e. the selection steps 1 to 5, are already implemented using the auxiliary class `Selection` in `susy_ex/Utils/Selection.h`. Also note that we do not have to compute the selection variables $\mathrm{N}_{\mathrm{jets}}$, H_{T}, $\not{H}_{\mathrm{T}}$, and $\Delta\phi$ ourselves, they are in fact already present in the event content! Execute

```
root -l -b -q general2.C+\(id\)
```

where again we start with the $\mathrm{W}(l\nu)+\mathrm{jets}$ sample, i.e. with `id` set to 11. The script will produce the output file `General_WJets-Yields.root` that contains the known kinematic distributions as well as a new histogram `hYields` that stores the number of simulated events passing the selection. The first bin contains the number of events after the baseline selection. The further bins store the number of events after tighter cuts on H_{T}, $\not{H}_{\mathrm{T}}$, and $\mathrm{N}_{\mathrm{jets}}$. Have a look at the code to identify the additional requirements.

By comparing the number of events after the baseline selection with the total number of events given in Table 1, we can compute the total selection efficiency

$$\epsilon = \frac{\text{number of MC events after cuts}}{\text{total number of MC events}} \,.$$

Together with the cross section, we can then compute the expected number of events in data for each process:

$$N_{\mathrm{data}} = \epsilon \cdot \sigma \cdot L \,.$$

- **Question 3.4.1** How many $\mathrm{W}(l\nu)+\mathrm{jets}$ events are expected in $1\,\mathrm{fb}^{-1}$ of data?

Let us now determine the expected event yields also for the other background processes. We do not have to compute the cross-section normalisation ourselves every time. Instead, we can use the weight already stored in the ntuples. It is returned by the `Event::weight()` function and includes the cross-section normalisation[3]. Adapt the `general2.C` script such that the histogram entries are weighted by the event weights. For this, you just need to modify the line which says

```
// Apply an event weight
const float weight = 1.;
```

Then, run the script `general2.C` for the SM-background samples and signal samples again. Afterwards, you can use the script `plotDataVsMC.C`,

[3] In addition, the weight factor also contains a correction to properly describe the impact of pile-up collisions.

```
root -l -b -q plotDataVsMC.C+
```

which plots the H_{T}, $\not\!\!H_{\mathrm{T}}$, and $\mathrm{N}_{\mathrm{jets}}$ distributions with the background processes stacked. The script also conveniently prints the event yields after the baseline and the additional selections.

- **Question 3.4.2** Discuss the result. Which background dominates in which phase-space region? Does this match your initial expectation (Question 3.2.2)?

3.5 Data

We will now study how these distributions look like for real pp collision data. The script `general2.C` also takes `id=1` as an argument to run over data:

```
root -l -b -q general2.C+\(1\)
```

Compare the $\mathrm{N}_{\mathrm{jets}}$, H_{T}, and $\not\!\!H_{\mathrm{T}}$ distributions in data with the sum of the background distributions obtained from simulation. For this, you can use again the script `plotDataVsMC.C` with the argument `true` to superimpose the data on the background stack:

```
root -l -b -q plotDataVsMC.C+\(true\)
```

- **Question 3.5.1** Do you observe any deviations of the data from the SM background expectation? What can you say about the existence of any new-physics processes?
- **Question 3.5.2** Which uncertainty is represented by the error bars? Are there any further uncertainties to be considered?
- **Question 3.5.3** In the analysis, the background yields are not taken from the simulation but, as explained above, obtained using data-driven methods. In the following exercise, we will discuss one of these background-prediction methods in detail. Why do we not use the simulated background yields but measure them from data?

References

[1] H. Baer and X. Tata, “Weak Scale Supersymmetry: From Superfields to Scattering Events”. Cambridge University Press, 2006. ISBN 0-521-85786-4.

[2] S. P. Martin, “A Supersymmetry primer”, `arXiv:hep-ph/9709356`.

[3] 2011. W.J. Stirling, private communication, and `http://www.hep.phy.cam.ac.uk/~wjs/plots/plots.html`.

[4] The CMS Collaboration, “Search for new physics in the multijet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s}$= 8 TeV”, *JHEP* **06** (2014) 055, `arXiv:1402.4770`. `doi:10.1007/JHEP06(2014)055`.

[5] The CMS Collaboration, “CMS technical design report, volume II: Physics performance”, *J. Phys.* **G34** (2007), no. 6, 995–1579. `doi:10.1088/0954-3899/34/6/S01`.