Wish list from phenomenologists: material needed for implementing and validating LHC new physics searches in fast simulation tools

G. Chalons, K. De Causmaecker, B. Dumont, B. Fuks, I.-W. Kim, S. Kraml, K. Sakurai, D. Schmeier, D. Sengupta, J. Tattersall (for the ATOM, CheckMATE and MadAnalysis 5 teams)

Version 1.0, date: September 30, 2014

Abstract
We detail the material that we need for a pertinent interpretation of ATLAS and CMS analyses based on event simulation. The document is intended to give guidelines for the experimental collaborations and may serve as a useful check list for the collaboration-internal approval process.

1. Analysis description

1.1 Clear description of all the cuts and their sequence

A clear and unambiguous definition of all the cuts and the sequence in which they are applied is essential for reproducing any analysis. This concerns not only the definition of the signal regions but also the pre-selection criteria. At present, the cuts are described well in most papers. Their sequence is however not always entirely clear. A negative example regarding both aspects is the CMS same-sign dileptons + MET analysis, CMS-SUS-13-013, for which the ordering of the dilepton selection as well as the jet selection criteria are unclear.

A systematic presentation of all cuts in tabulated form, strictly following the exact sequence in which they are applied, would be highly appreciated. If page limitations are an issue in the publication, this information could also be conveyed through auxiliary cut flow tables that detail every step of the analysis, see point 2.3 below.

1.2 Efficiencies for physics objects

Please provide efficiencies as function of p_T (and, where relevant, η) for all physics objects considered in the analysis: electrons, muons, taus, di-leptons, as well as for b-tagging, mis-tagging rates, *etc*. If an efficiency is not given explicitly in the paper, provide a clear reference where to find it (*e.g.* reference to the precise figure in another paper or a performance note, or even better, an analytical formula).

While this should be standard, there are still papers that miss out on this. For example, in CMS-SUS-2013-011 (stop search in the single-lepton channel, published in EPJC) the identification-only efficiencies for electron and muons are missing; they were however provided by the CMS collaboration upon request and are now available on the analysis twiki page.

1.3 Efficiencies for triggers, event cleaning, etc.

Please provide efficiencies for everything that we cannot directly reproduce in the fast simulation, in particular efficiencies for triggers and event cleaning.

Unfortunately many analyses give no or insufficient information about the trigger. For example, ATLAS-CONF-2013-035 (trileptons + MET) has no trigger information. One is left to choose to use the single lepton trigger from ATLAS-CONF-2012-104 or the di-lepton trigger from ATLAS-CONF-2012-049, neither of which is a good solution. The situation is similar for event cleaning: often cut-flow tables start with cleaned events, but no efficiencies are given for the cleaning procedure. We therefore ask that efficiencies for triggers and event cleaning be provided systematically in all analyses.

1.4 Digitized plots

We also note that digitizing performance plots is utterly time consuming and introduces unwarranted uncertainties, especially when reading numbers off log-scale plots. We therefore ask that efficiency curves always be provided in numerical form in addition to the paper plots. A prominent example are the electron efficiencies in ATLAS-CONF-2014-032: while the full two-dimensional binning at 8 TeV is highly valuable information in principle, extracting all these numbers by hand from the plots is a hugely time consuming task. Another example are the b-tagging efficiency curves presented in Fig. 2 of ATLAS-CONF-2014-046: an accurate extraction of the information from the log scale plots is hardly feasible. It would be highly valuable if, in addition to the png, eps and pdf versions of the figures, also the plot data were given in, *e.g.*, a Root or a simple text file.

1.5 Special kinematical variables

CMS is providing code modules for some kinematical variables, for example M_{T2}^{W} or the special procedure 'razor' analysis. This is extremely helpful and highly appreciated. Going a step further, we think that a central repository of all kinematical variables used by the collaborations would be beneficial. In addition, we note that for many variables the naming convention can often be ambiguous (for example, the different versions of asymmetric M_{T2}) and the analyses should always be crystal clear of which final states are used as input.

2. Validation material

2.1 Benchmark points

For the validation of an analysis implementation it is important to work with the exact same benchmark points as used in the experimental paper. This concerns not only the masses of mother and daughter particles, but also the properties of off-shell particles in production and decay processes, and mixings between different states (e.g. the precise gaugino-higgsino mixing in the SUSY case). All these can have important effects on cross sections and kinematic distributions.

A convenient way of unambiguously defining benchmark points is to make the SLHA files (including the full mass spectrum and decay tables) and/or parton-level MC event files available. ATLAS and CMS are already doing this for many of their papers. This is excellent practice and should become a standard procedure not only for journal publications but also for conference notes (ATLAS) and public analysis summaries (CMS).

2.2 Exact configuration of Monte Carlo tools

Different versions of Monte Carlo (MC) tools or even just slightly different settings can produce quite different results. It is easy to waste a lot of time trying to validate an analysis if one does not have the exact right MC setting.

Therefore, the exact versions of the MC tools used should always be clearly documented. The optimum would actually be if run cards and input scripts for MadGraph, Pythia, Herwig, etc. were made available on the analysis twiki page or on `HepData`. Again, this concerns not only journal publications but also for conference notes and public analysis summaries.

2.3 Detailed cut flows

For us, cut-flow tables are the most useful and easiest form of results to validate our analysis with. Both collaborations have made huge steps forward in routinely providing cut-flows but we believe further improvements can still be made including,

- Detailed cut flows for each signal region in the analysis, if possible for a variety of different benchmark points.

- The more steps that are provided in cut-flows, the better. This allows us to precisely determine where our simulations differ from the experimental collaborations and provide more realistic tunings. In contrast, grouping *e.g.* various jet selection criteria into one step in the cut-flow results severaly reduces the amount of useful information.
- The cut-flows should precisely reflect the ordering in which the cuts are applied in the analysis, see also item 1.1 above. (While this may seem obvious, it is presently not always the case.)
- Cut-flows are only useful if they contain reasonable statistics so that the dominant uncertainty is not the statistical error associated with the MC sample size. For example, numerous cut-flows have only $\sim$10 events in the final signal which means a precise validation is difficult.
- Cut-flows have now become invaluable for SUSY studies but unfortunately this appears to be the only analyses where they are applied routinely. We would appreciate cut-flows for ALL analyses (including Standard Model measurements) as these can often be re-casted into a search for models of new physics.

2.4 Kinematic distributions

Histograms of kinematic distributions after specific cuts provide extra cross-checks for our simulations. Especially useful are kinematical distributions of simulated signal samples so that we can validate that these are reproduced correctly. Moreover, distributions of general variables like the hadronic activity, the number of jets, of b-tagged jets, of leptons and their respective pT histograms after the preselection are extremely useful for validation purposes. Again, these should preferably be provided (also) in numerical form.

3. Statistical interpretation

We would like to combine signal regions as done by the experiments themselves. Moreover, we would like to combine the limits from different analyses to obtain more pertinent limits. Fully documented errors are a pre-requisite for this task. Unfortunately at the current time we have to assume that errors are either fully correlated or fully un-correlated if we wish to do combinations. The problem is that assuming uncorrelated errors only improves bounds marginally while assuming fully correlated errors significantly overestimates bounds.

We therefore ask that all signal regions have the backgrounds broken down into the individual sources. For each source we then need the statistical and systematic components of the error. With this information we are then able to estimate a error correlation matrix that enables a combination of signal regions. A possible alternative (or even better in addition) would be for the collaborations to provide a final likelihood model for each analysis; this could be done through small standalone code modules that take the number of BSM events expected in each signal region and return the likelihood for the particular model point.

Conclusions

It is important for the legacy of the LHC that the analyses carried out in the search for new physics can be interpreted in the contexts of all kinds of theoretical models, now and in the future. This clearly requires a community-wide effort including the development of recasting tools based on fast simulation.

Non-collaboration members however do not have access to the experimental data, nor the MC event set simulated with an official collaboration detector simulation, nor collaboration-internal documentation of analyses and detector performance. Therefore, the implementation and validation of ATLAS and CMS analyses for re-interpretation of the experimental results in general contexts is a tedious task, even more so as the information given in the experimental papers is often incomplete.

Following the guidelines given in this document would considerably improve the current situation and make the deployment of phenomenological tools for the interpretation of the LHC search results (and their use in physics studies) much more efficient. This would also give useful feedback to the experimental collaborations, e.g., about the coverage in parameter space of their analyses and the impact of their results beyond simplified or constrained models. Finally, providing a clear and complete documentation will also assure that analyses will remain reproducible for the experimental collaborations themselves.

Acknowledgements

This document originated from the mini-workshop on recasting ATLAS and CMS new physics searches, held 8–12 Sep 2014 at LPSC Grenoble. We thank Marie-Helene Genest and Tommaso Lari from ATLAS as well as Samuel Bein and Keith Ulmer from CMS for their valuable participation in the meeting. Financial support by the "Investissements davenir, Labex ENIGMASS" is gratefully acknowledged.