Coordinating a simplified model effort

SLAC-CERN:

Jay Wacker (jgwacker@stanford.edu),
Tim Cohen (timcohen@slac.stanford.edu),
Kiel Howe (howek@stanford.edu),
Maurizio Pierini (maurizio.pierini@cern.ch),
Till Eifert (till.eifert@cern.ch),
Zach Marshall(zach.marshall@cern.ch),
Chris Rogan

LBL-DESY:

Michele Papucci (mpapucci@lbl.gov),
Andreas Weiler (andreas.weiler@cern.ch),
Kazuki Sakurai (kazuki.sakurai@desy.de),
Lisa Zeune (lisa.zeune@desy.de)
...

SMODELS:

Wolfgang Waltenberger (wolfgang.waltenberger@gmail.com),
Sabine Kraml (sabine.kraml@lpsc.in2p3.fr)
Suchita Kulkarni (suchita.kulkarni@lpsc.in2p3.fr),
Sezen Sekmen (sezen.sekmen@cern.ch),
Andre Lessa (lessa@if.usp.br)...

RWTH Aachen:

Jory Sonneveld (sonneveld@physik.rwth-aachen.de),
(Lennart Oymanns),
Michael Krämer (mkraemer@physik.rwth-aachen.de),
Lisa Edelhäuser (ledelhaeuser@physik.rwth-aachen.de) ...

Stockholm:

Oscar Stal (oscar.stal@fysik.su.se) , ...

MadAnalysis5 (Strasbourg-CERN):

Benjamin Fuks (fuks@cern.ch),
Eric Conte (eric.conte@iphc.cnrs.fr)

Motivation

Simplified Models have become one of the standard methods to search for physics beyond the Standard Model. Simplified Models reduce full models with dozens of particles and hundreds of parameters down to a handful of particles and parameters, allowing for the commonalities between different models to become apparent. This prevents experimental collaborations from duplicating effort by performing a multiple analyses for the same signature. At the same time, Simplified Models allow the explorations of physically allowed models that may not have been proposed. Simplified Models also allow the experimentally relevant parameters to be searched plotted, often illustrating where existing analyses begin to lose sensitivity at boundaries of parameter space.

The challenge with Simplified Models is that any full model decomposes into many different Simplified Models. The spectrum of Simplified Models sometimes requires several different particles with many different parameters. Existing experimental analyses are often sensitive to these alternate Simplified Models, so there is not necessarily any additional work for the experiments to do in the limit setting process, it is only a matter of reinterpreting existing results. While the experimental collaborations can do this, it is often beyond their computing resources and the effort required in reinterpreting results could be spent in performing new analyses.

Simplified Model reanalysis is hard to do correctly, requiring both attention to detail, but also significant computing resources. Currently there are at least 6 groups performing Simplified Model reanalysis.

There are multiple reasons that coordination is necessary. First, experimental analyses often present results on related, but different Simplified Models. This is mostly because there has not been prioritization and coordination between experimental collaborations and with the theory community. Next, there is no standard format to present the results from Simplified Model results. Finally, there are too many Simplified Models for every relevant experimental analysis to present efficiencies for every Simplified Model.

We therefore feel that there needs to be more coordination of efforts between theoretical groups and between experimental efforts. By increasing coordination, the maximal gain in coverage of Simplified Models can be achieved and results can be thoroughly vetted.

There are several levels of coordination needed. Not all levels of coordination are necessary to make a substantial improvement over the current state of Simplified Models

1.) Definitions of Simplified Models
2.) Prioritization of Simplified Models
3.) Format for Presentation of Efficiencies
4.) Location to Store Existing Efficiencies
5.) Format for Analysis Level Cuts and Reanalysis Method
6.) Providing Citable References for Future Use

Definitions of Simplified Models

Many Simplified Models are near standardized now, nevertheless, there can be differences in the parameterization of Simplified Models. The naming and parameterization scheme should be standardized so that there is no confusion on which Simplified Models have been analyzed.

Prioritization of Simplified Models

Simplified Models have well-defined notion of prioritization. For instance, direct decays of a heavy particle to the lightest particle should be studied before multiple stage cascade decays. This is because the simpler the topology, the more likely that that topology will appear in a wide variety of models. Additionally, the more complicated the topology is, the more parameters are necessary to define it and therefore it is more challenging to exhaustively study these complicated topologies.

Format for Presentation of Efficiencies

If an analysis of a Simplified Model is performed by either an experimental collaboration or a theoretical reanalysis, then the format should be standardized so that it can be easily reused in analyzing topologies appearing in full models.

Location to Store Efficiencies

There should be a location to store efficiencies so that they can be downloaded. It is also a natural location to coordinate efforts on new Simplified Models.

Format for Analysis Level Cuts and Reanalysis Method

This is translating existing experimental analyses into a language that can be used in analyzing new Simplified Models. This does not necessarily need to be coordinated with the experiments since this is simply transcribing their stated analysis level cuts into a standardized format, although any coordination with the experiments would be appreciated. Also include the N_expected and N_observed and the error in search regions.

From Benj: we have a python interface to implement analysis cuts in a very human-readable fashion. Translation into C++ is already there (link to the MadAnalysis 5 core). It could be straightforwardly generalized to be compatible with the formats needed by the other codes involved (assuming they use a modern programming language).

Likelihood Discussion

For more complicated analyses such as ones that combine together multiple control regions, or ones that use more complicated use of likelihood tests, there needs to be a thorough vetting of the procedure used to reanalyze the experimental results. Some results may be too complicated or provide too little information to reanalyze.

Providing Citable Credit

For each topology reanalysis the full list of citations should be provided including

- The relevant experimental analyses
- The Simplified Model reanalysis papers
- The article or references that that transcribe the cuts for the first time
- The relevant citations for the tools used to generate the plots

These can be provided in a bibtex format to simplify the citation procedure.

Additionally, Digital Object Identifiers (DOIs) should be considered to come as an additional citable analysis.

In addition to papers that can and should be written on reanalyses of specific topologies with specific experimental analyses, there should also be regular updates combining the global information together.   This can be done annually or after each major run of the LHC.

From SMODELS + Aachen team:
Following are the suggestions:
1) (Strong) connection with experimental groups should be made
2) The scope and aim of the current draft should be well defined and possibly expanded
3) In order to coordinate the entire effort, a meeting/workshop at CERN can called for