Validation of ADL/CutLang analyses with SModelS EM-creator for reinterpretation studies

**Analysis Description Language workshop
[November 21-22, 2022]**

**Sezen Sekmen[1], Changgi Huh[1], Junghyun Lee[1], Gokhan Unel[2],
Wolfgang Waltenberger[3], Jan Mrozek[4]**

[1]Department of Physics, Kyungpook National University
**[2]The University of California, Irvine**
[3]Research institute of the Austrian Academy of Sciences in Vienna (HEPHY)
[4]Red Hat Software

Simplified model spectra (SMS)

- **Simplified model spectra (SMS) are defined by a set of particles and their production & decays.**
  - Particle masses, cross sections, branching ratios are free parameters.

- **SMSs are practical for comparing sensitivities of different analyses.**

- **SMSs can be combined to form full models.**
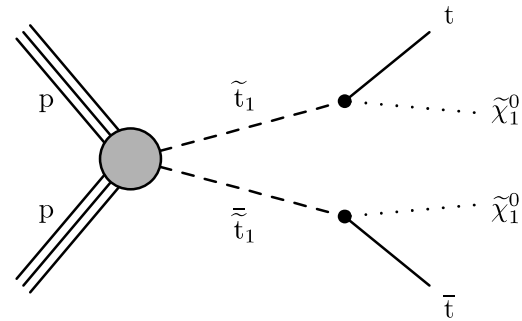
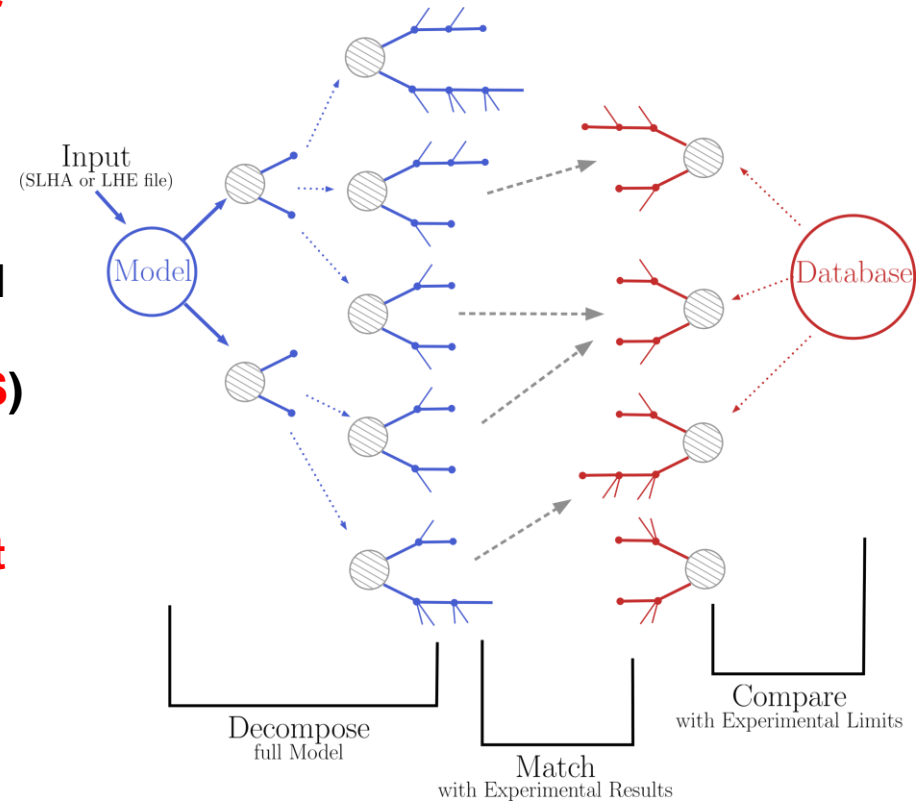So, **How do you interpret the LHC result of the Simplified model?**

SModelS

- **SModelS** is an automatic, **public tool for interpreting simplified-model results** from the LHC.

- It is based on a general procedure to **decompose** Beyond the Standard Model (**BSM**) **collider signatures** presenting a Z_2 **into** Simplified Model Spectrum (**SMS**) **topologies**.

- **SModelS** method **provides a way to cast BSM predictions for the LHC** in a model independent framework, which can be directly confronted with the relevant experimental constraints.

Input (SLHA or LHE file)

Model

Database

Decompose full Model

Match with Experimental Results

Compare with Experimental Limits

- Refer to official documentation fo SModelS : https://smodels.readthedocs.io/en/stable/

- SmodelS github : https://github.com/SModelS/smodels

Motivation of automation system

- **SModelS requires SMS efficiency maps: event selection efficiencies versus SMS mass parameters.**
  - **It calculates limits from these efficiency maps.**

- **We need to generate many mass points to draw the limit plots for validation.**
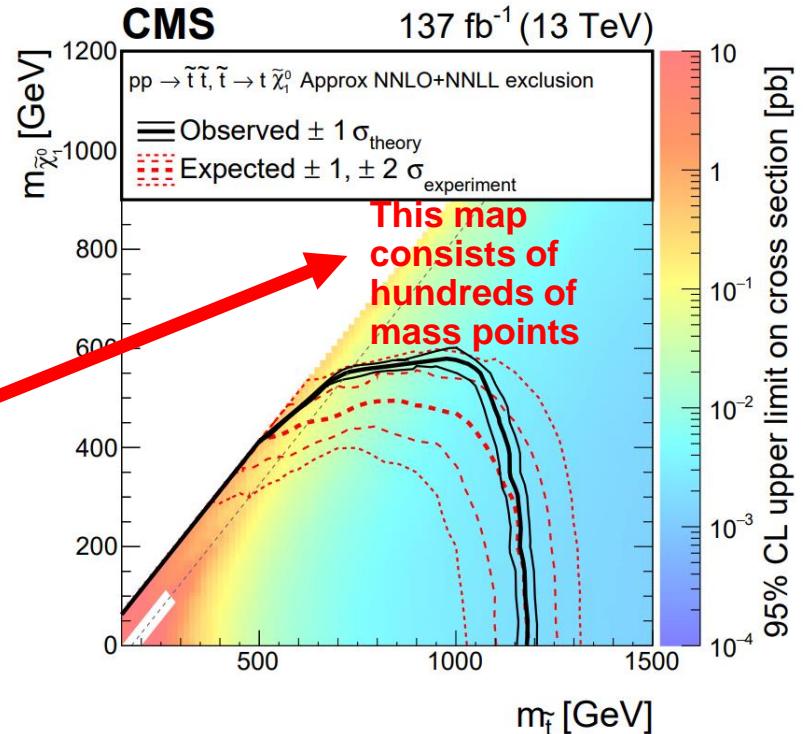
- **That would be a very repetitive work**

This map consists of hundreds of mass points

Figure : The 95% CL upper limit on the production cross sections of the T2tt
Ref : CMS-SUS-19-006 (cern.ch)

WE NEED AN " AUTOMATION SYSTEM "

EM-creator

- **EM-creator (Efficiency Map creator) was developed by Wolfgang W. to produce efficiency maps that can be input to SModelS.**
  - It can be used to validate analysis implementations by comparing to experimental results.

- **EM-creator was adapted to work with ADL/CutLang by Jan M., Gökhan Ü. and Wolfgang W.**
  - Efficiencies are calculated with CutLang using analyses written in ADL.
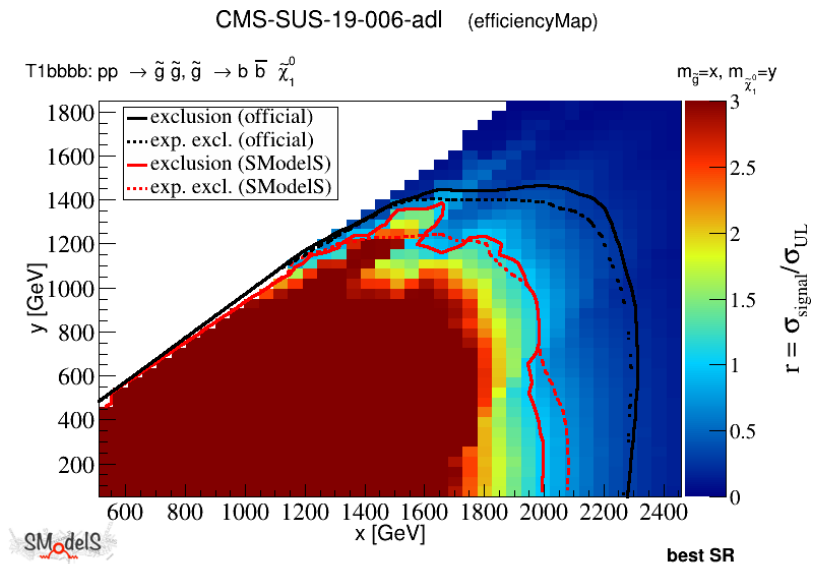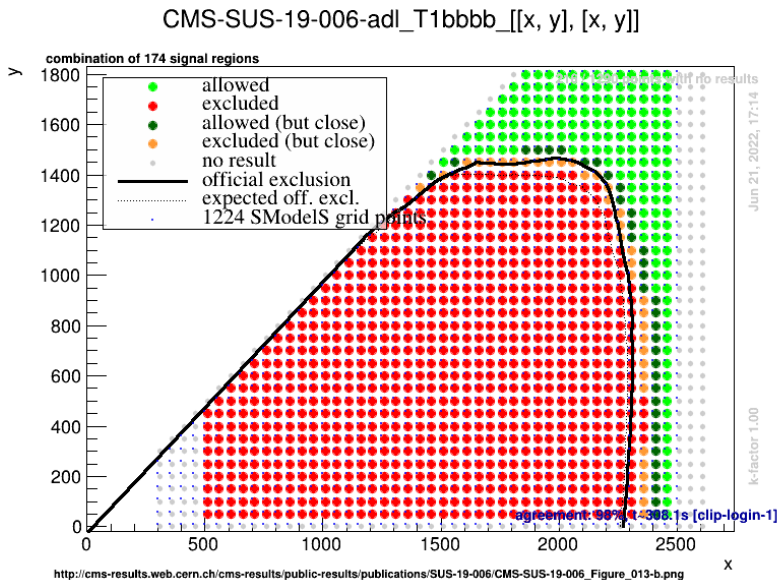
EM-creator workflow

Number of events, ADL analysis file, mass range, SMS topology, …

Efficiency Map creator

| Madgraph & Pythia | Delphes | CutLang |
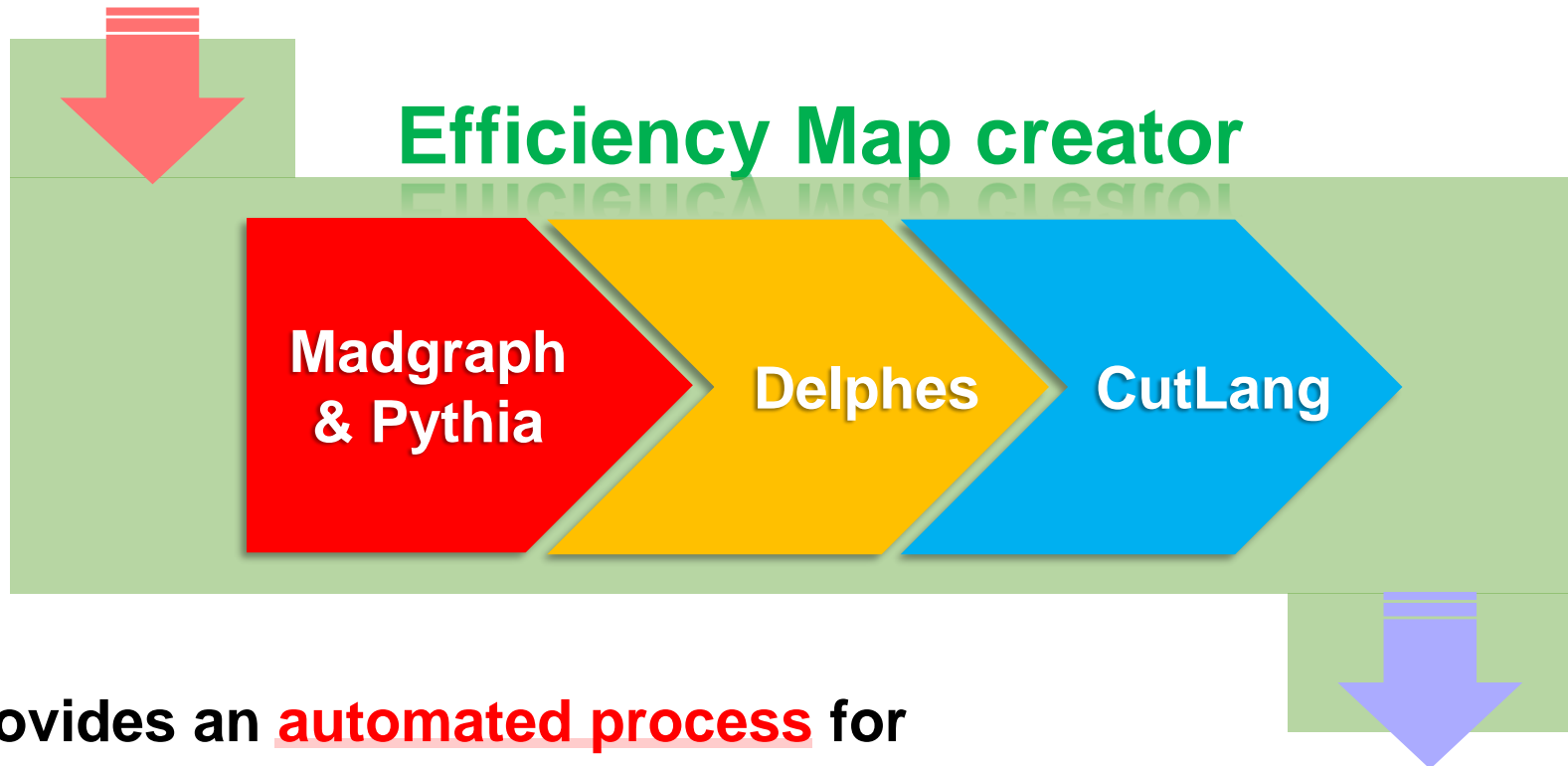
Provides an automated process for analysis through simple command line

Selected number of events / efficiencies in all regions & bins

Run EM-creator

- **Working environment**

CMSSW_12_3_0_pre6 in KNU Tier3 system

- **Command example**

Number of events/point

Range for mass parameters, (loweset m, highest m, Δ m)

`$ ./bake.py –n 10000 –m "[(250,2201,50),(10,2001,25)]" –analyses "cms_sus_19_005, cms_sus_19_006" – –cutlang –a –t T2ttoff –b`

Select Analysis (Relevant [.ADL] file needs to exist)

Select analysis program and run it

Topology

Make efficiency data for selected events
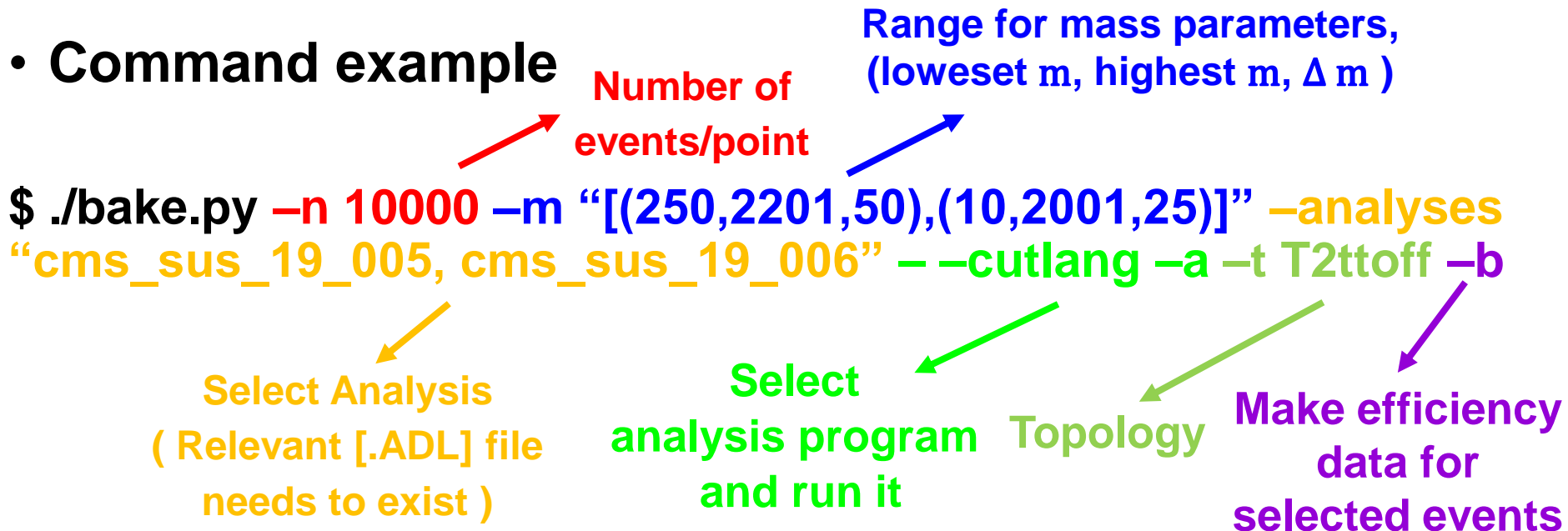
Operating log (1250, 1085) mass point

Results for (1250, 1085) mass point

- **Mass point (1250, 1085)'s CutLang result**
  - **histoOut-CMS-SUS-19-006_CutLang.root**
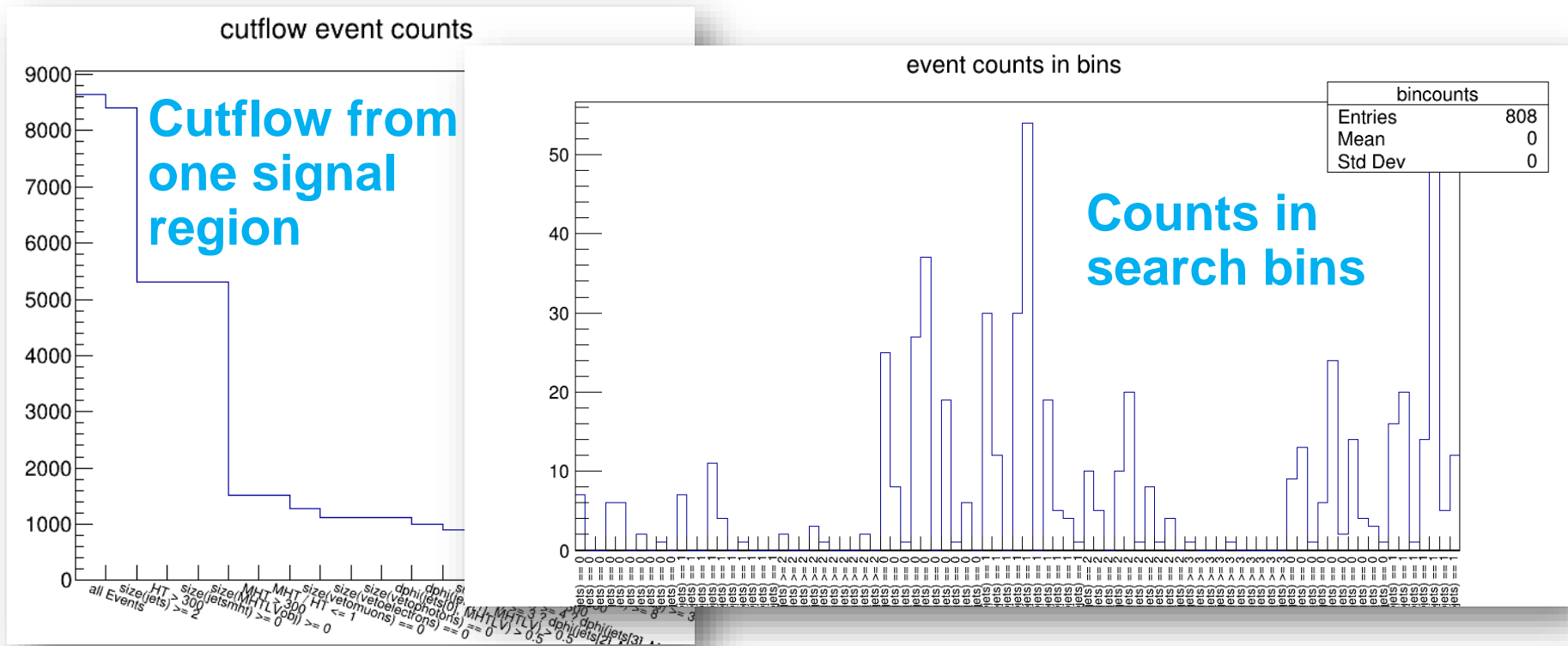
Cutflow from one signal region

Counts in search bins

Partial efficiency from CutLang results

- **Cms_sus_19_006_T2ttoff_mass_1250_1085.embaked**

Finally, using the results of CutLang :
The em-creator outputs embaked file,
which contains partial efficiencies as above

EM-creator workflow

Number of events,
ADL analysis file, mass range
SMS topology, …

Delphes
MC result (.root)

Result
(.embaked)

Madgraph & Pythia → **Delphes** → **CutLang** → **Result : Partial efficiency**

Madgraph & pythia result (hepmc.tar.gz)

**CutLang analysis result (.root)**
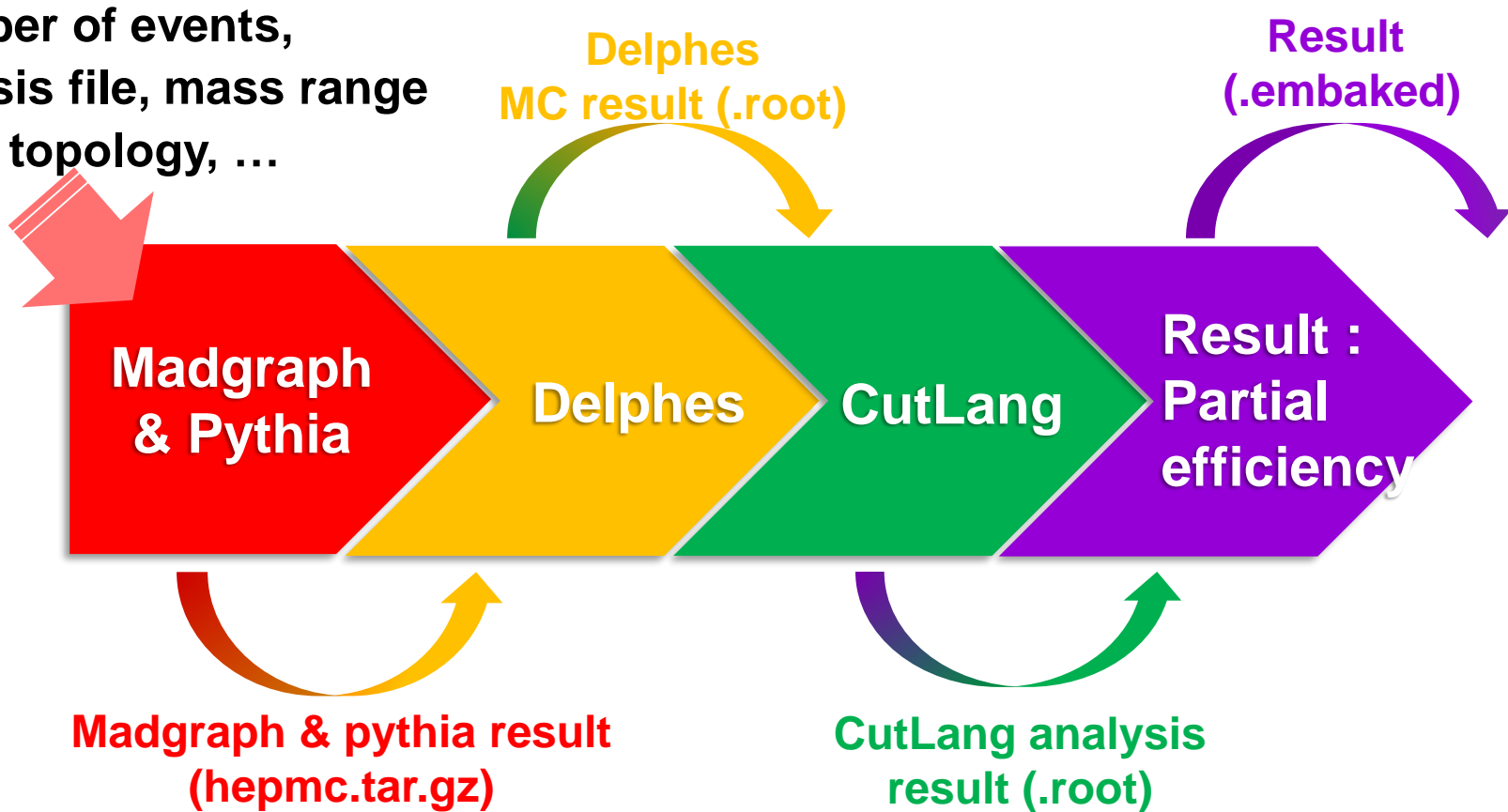
- **This setup is highly object-oriented and factorizable. Each component can be run individually.**

EM-creator operation using HTCondor

- **The em-creator is very easy to divide and execute works for each mass point**

- **Suitable for use with HTCondor.
  We set up EM-creator in the KNU T3 HTCondor system.**

- **Example)**

  - **Apply a mass range like –m "[(250,2201,50),(10,2001,25)]" to em-creator 261 mass points come out**

  - **One can submit 261 jobs for each mass point.**
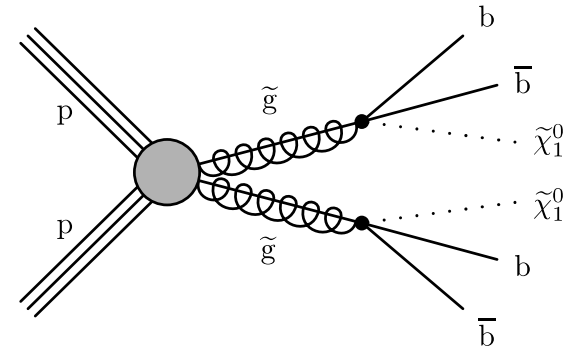
Result of em-creator with HTCondor
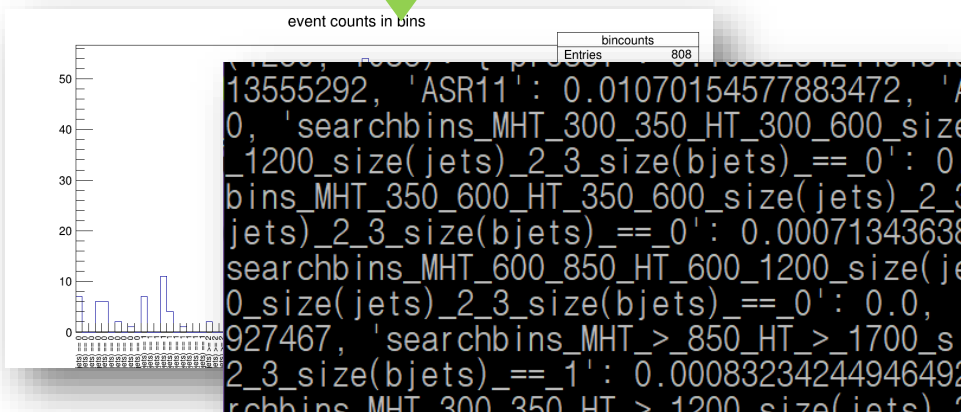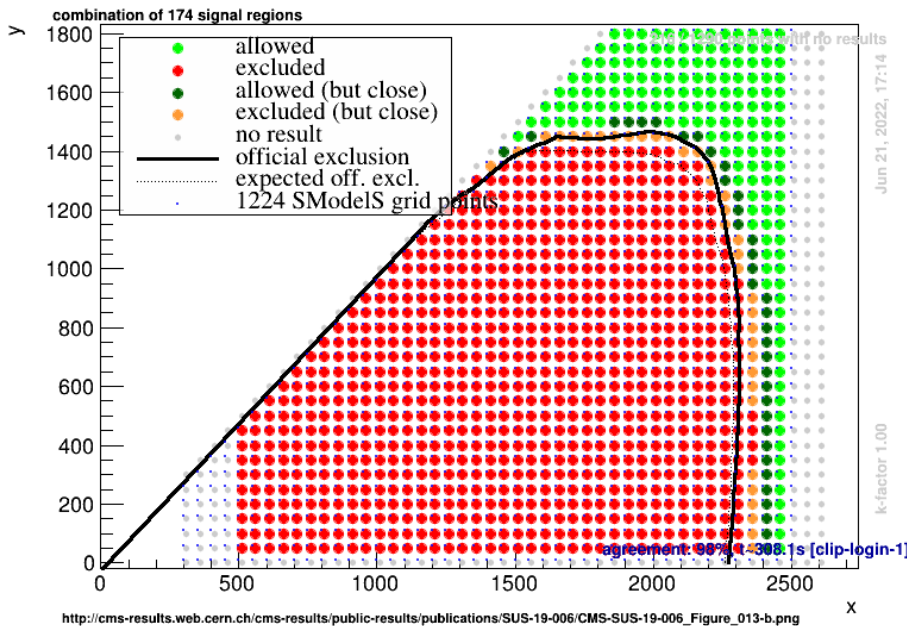

- **The directory corresponding to each mass point has CutLang result and efficiency result files.**

- **We can validate numerous models very effectively and conveniently through EM-creator!**
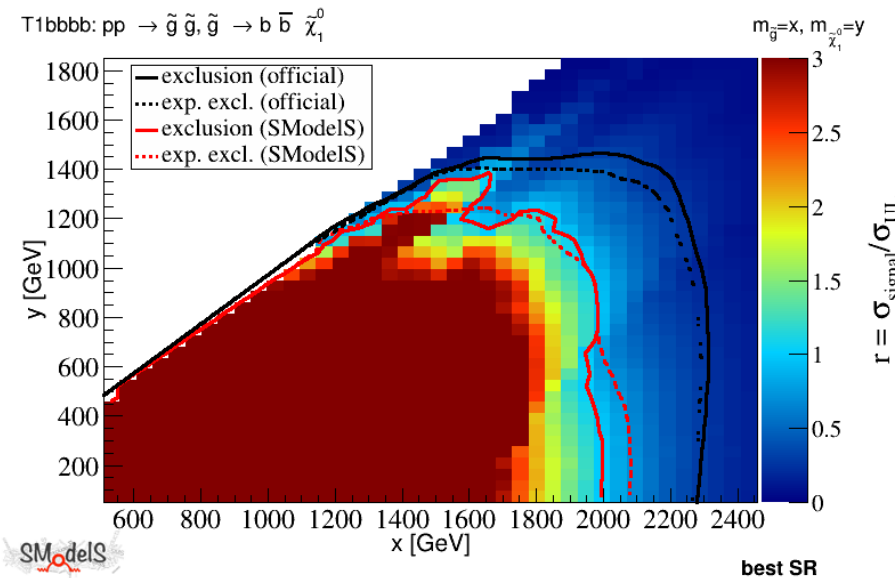
Final goal

- **Efficiency maps created for T1bbbb for CMS-SUS-19-006 by Wolfgang at the Vienna system. We will achieve this at KNU.**

CMS-SUS-19-006-adl_T1bbbb_[[x, y], [x, y]]

CMS-SUS-19-006-adl (efficiencyMap)

Summary

- **SmodelS EM-creator was adapted to ADL/CutLang** by Jan, Gökhan and Wolfgang.

- We **successfully installed** and **tested** the **EM-creator** in KNU T3 HTCondor system.

- We set up **each job to run a single point**. (One job for 10K evts from MG to CL took ~40 minutes)

- We **adapted the scripts to start from any stage in the chain**, e.g. only run the CutLang and efficiency bake.

- We will perform one complete run using CMS-SUS-19-006 to produce efficiencies.

Next steps

- **Coordinate with the SModelS team to set up a system to calculate limits from the em-baked efficiencies.**

- **Perform more robustness checks in our setup.**

- **Start an organized large scale validation of several analyses.**

- **Move towards publishing the EM-creator for public use.**