Simple Analysis for LLPs —Truth-level Analysis Framework

Yvonne Ng @ DESY

2023-6-20

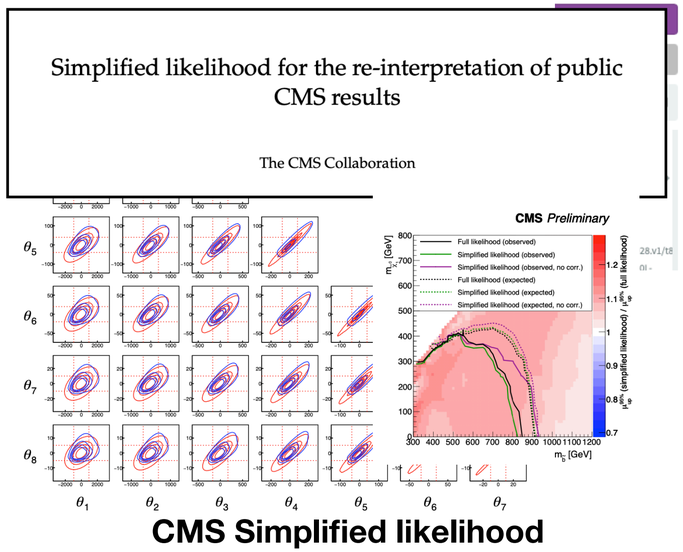
LLP13

A timeline on Reinterpretation

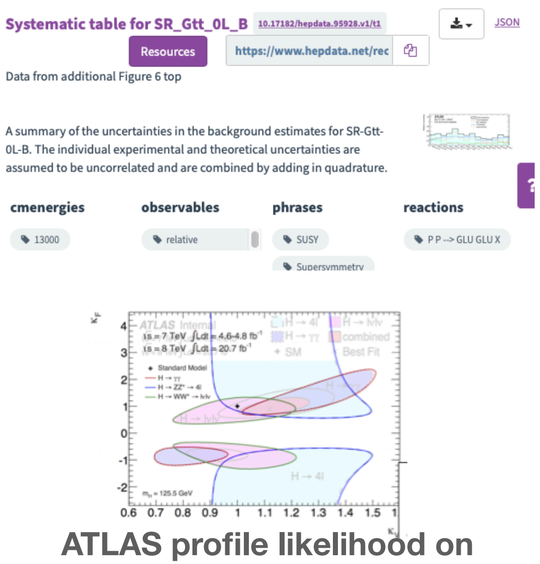
— Development of Tools and the Progression of Collaboration Standard

- 2000: 1st PHYSTAT
- 2010: introduction of Workspaces
- 2010: First proposition of Recast
- 2012: ATLAS profile likelihoods scans released for Higgs coupling
- 2017: CMS Simplified Likelihoods
- 2018: Release of analysis preservation and systematic reinterpretation with Recast within ATLAS

HEPData



CMS Simplified likelihood

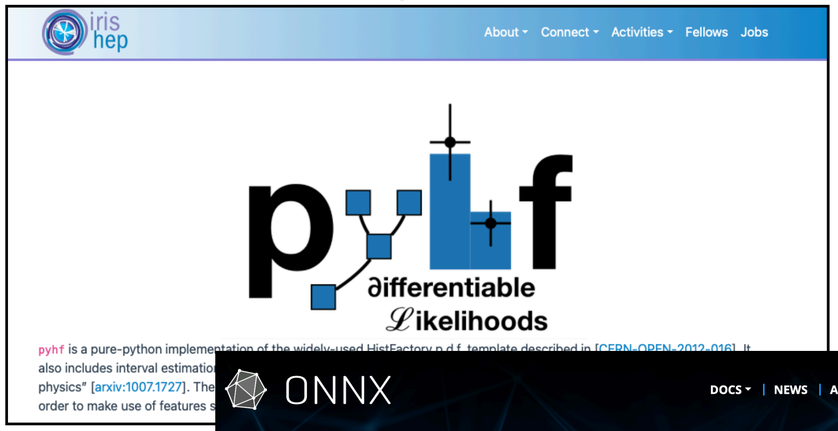


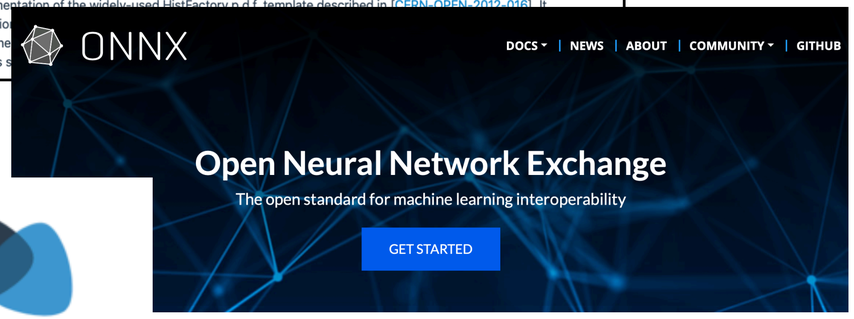
ATLAS profile likelihood on higgs coupling

A timeline on Reinterpretation

— Development of Tools and the Progression of Collaboration Standard

PyHF





Onnx

recast

Recast

***ATLAS development highlighted in blue**

- 2019: First full likelihood release in ATLAS for SUSY searches
  - Using json schema of Histfactory class of likelihood
- 2019: ATLAS publication of full statistical likelihoods PUB note released (JSON HF)
- 2020: First full likelihood release regularly in ATLAS for SUSY searches
- 2022: "Simple Analysis: Generator-level Analysis Framework" pub note released—Framework going public!
- 2022: Snowmass: "Reinterpretation and Long-Term Preservation of Data and Code"

Different Reinterpretation Tools

Less Precision

Level of Complexity

More information/Precision

Experiment specific

Simple Likelihood

Reinterpretation from likelihood

p(theory|data) ~ p(data|theory) * p(theory)

New Result!

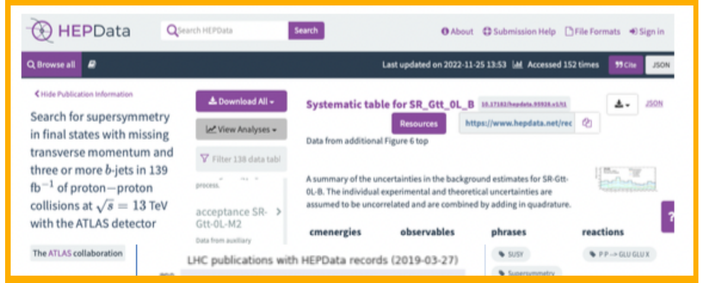
Theoretical Prior of New Model

Simple Likelihood Pubnote

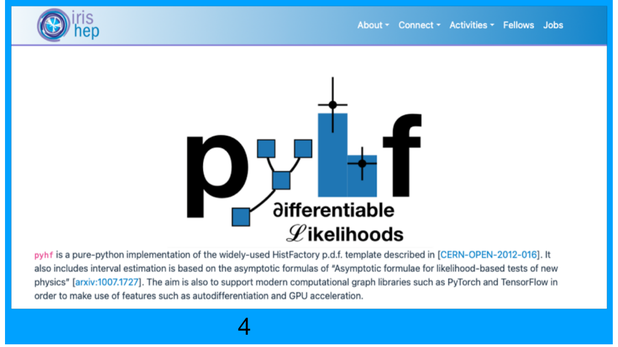
Existing Exp. Likelihood (What We Provide)

6

HEPData:
Storing data products from analyses for reinterpretation

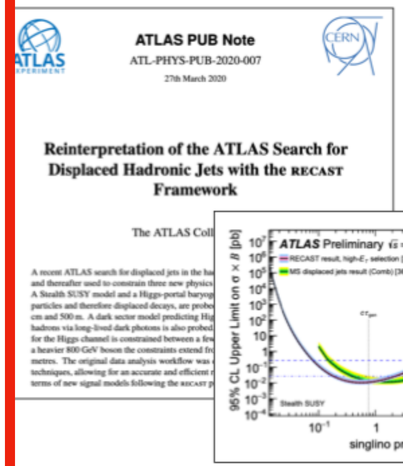


HistFactory Json:
Storing data products from analyses for reinterpretation



4

Recast:
Storing data products from analyses for reinterpretation

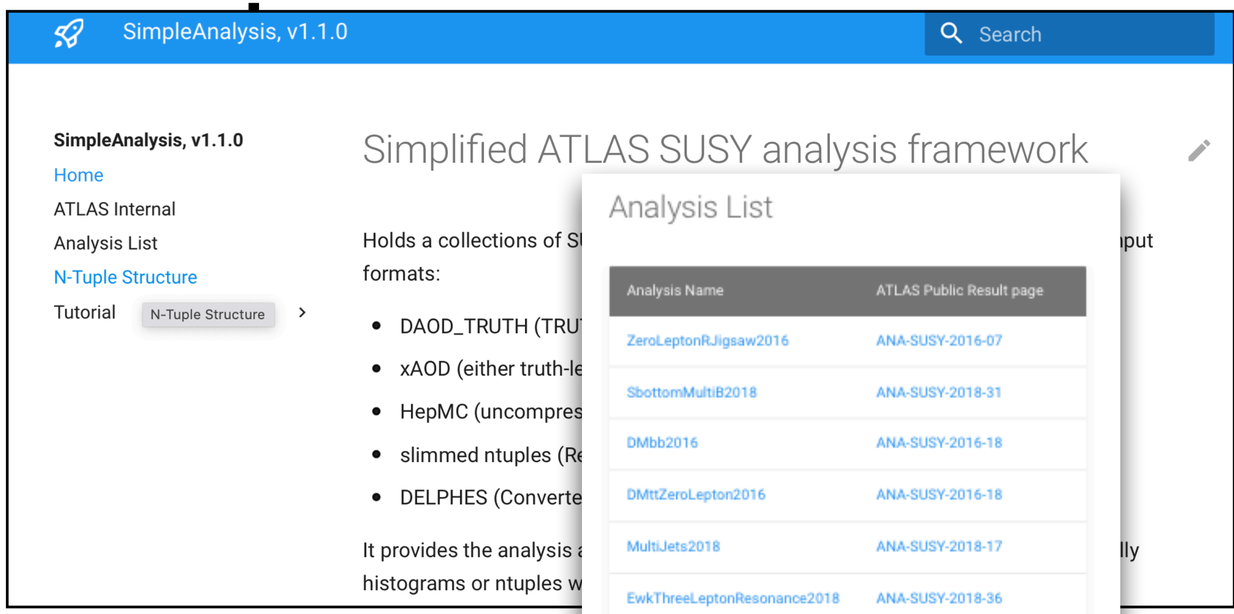


recast

preservation of the original analysis pipeline for collaboration-approvable results

Simple Analysis
—Experiment-independent middle

- Preserve **full analysis pipeline** for manipulation/maximal information
- **Open source**
  - Experiment non-specific, not tie to detector reconstruction
  - Approximation can be added
- Allow for **reinterpretation up to the truth level**
  - Detector effect can be added
- Non-experiment input format
  - **HEPMC, DAOD, ROOT-ntuple, Delphes** etc.
- For Public use: can be used outside of ATLAS
- Additional information for reinterpretation



SimpleAnalysis: [Webpage]

GitLab

Source Code

[SimpleAnalysis PubNote]

Available list of Analyses

Simple Analysis Framework

Generator input format

DAOD
(ATLAS format)

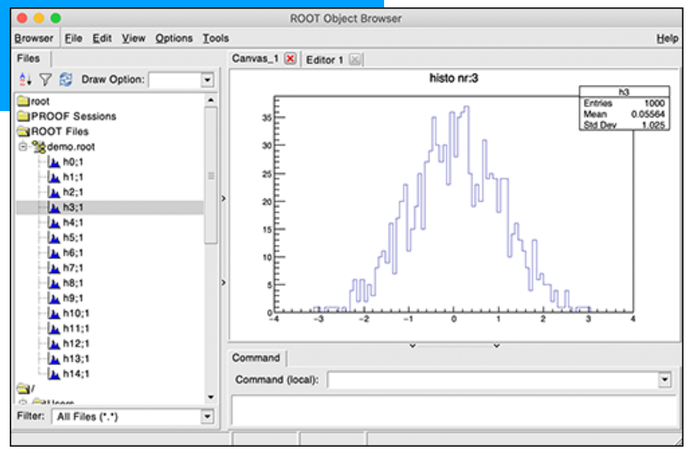
HEPMC
ROOT
DELPHES

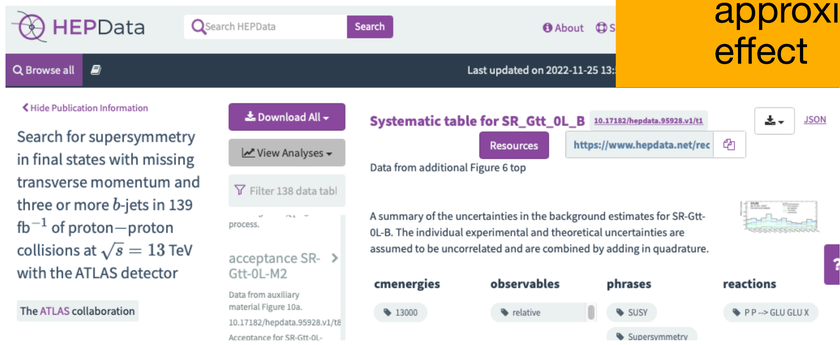
Simple Analysis Routine

Complex kinematic variables
Custom efficiency maps
Additional utilities for serializing NN and multi-variate analysis /BDT

HEP data publishing

- Truth level acceptance and selection efficiencies for HEPData publishing
- Root flat-ntuple files containing histograms
- Smearing available to approximate detector effect



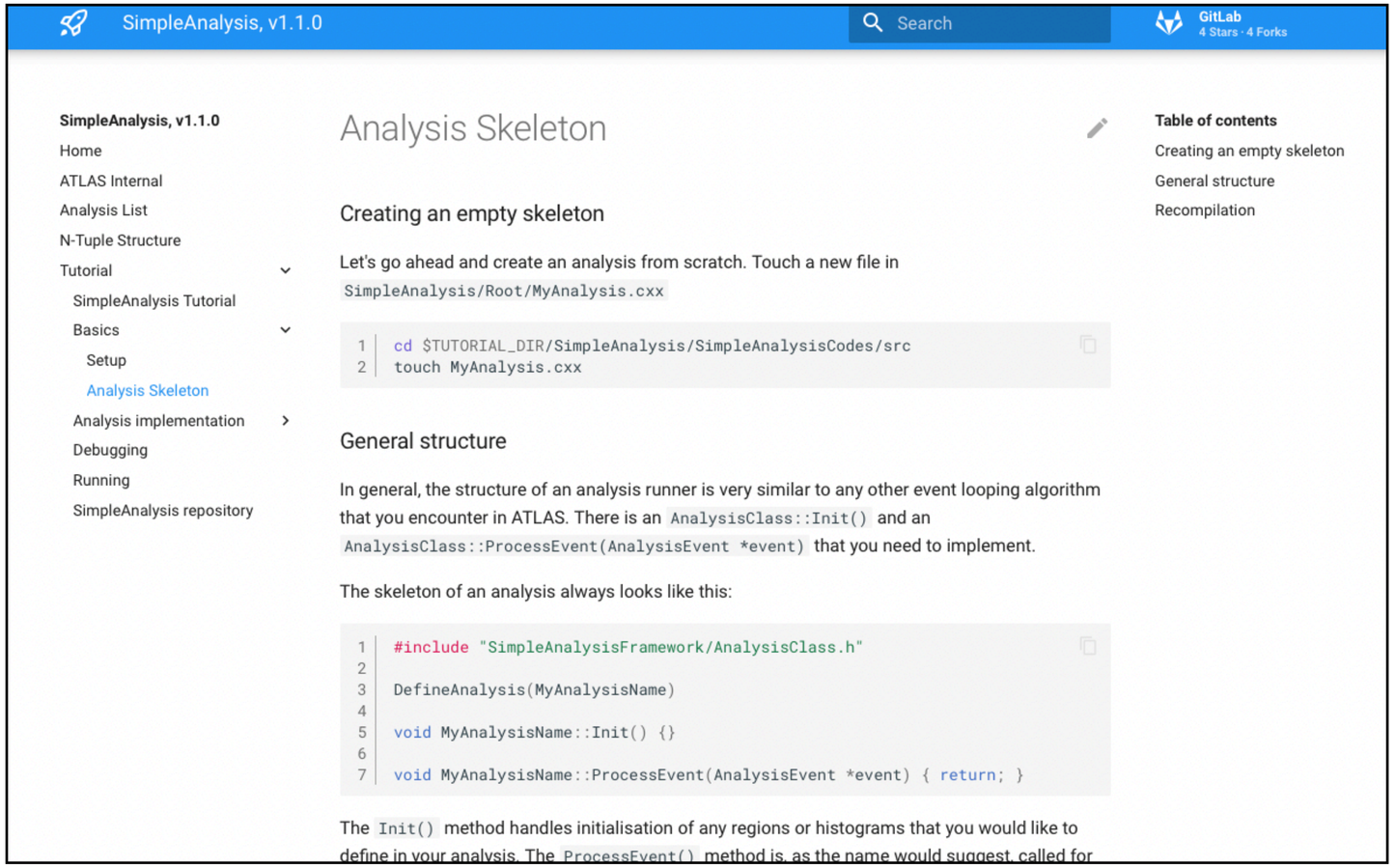


Simple Analysis
—Tutorial

- Tutorial contains:
  - Building analysis selection

    Analysis pipeline Preservation
  - Reinterpreting results from previous SA pipelines.

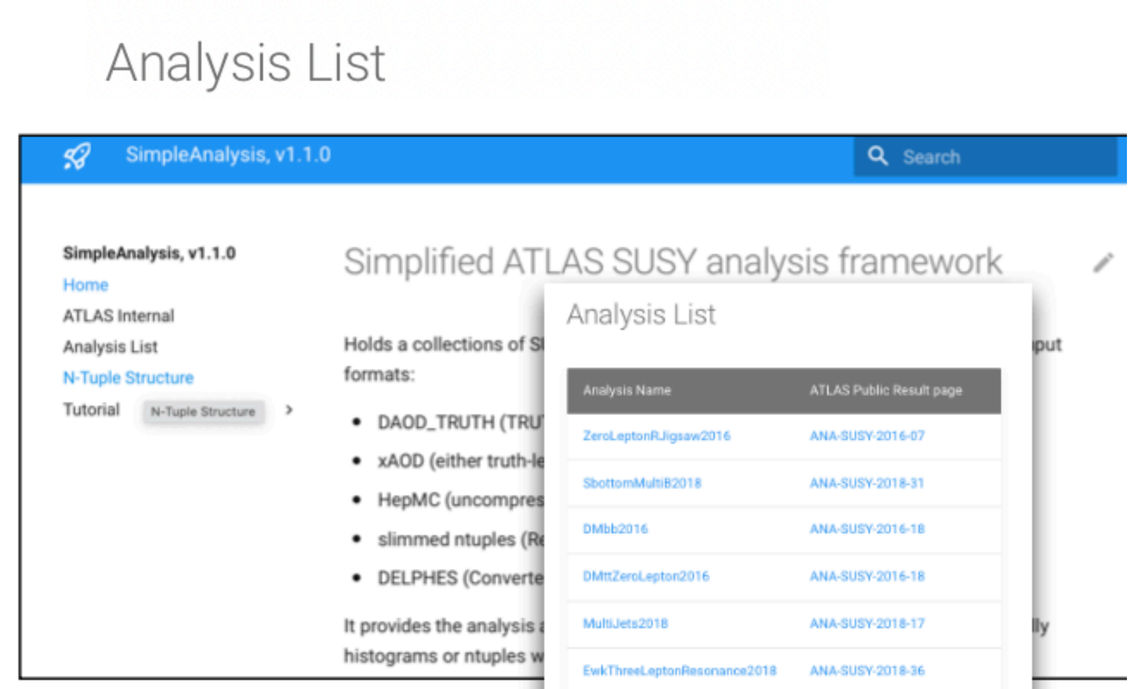
    Reinterpretation Effort
  - Welcome feedback!



Simple Analysis

—Current analysis list

- Most SUSY analyses on ATLAS publishe
  analysis selection at the level of generat
- Growing list can be found here.



SimpleAnalysis: Webpage

GitLab

Source Code

SimpleAnalysis PubNote

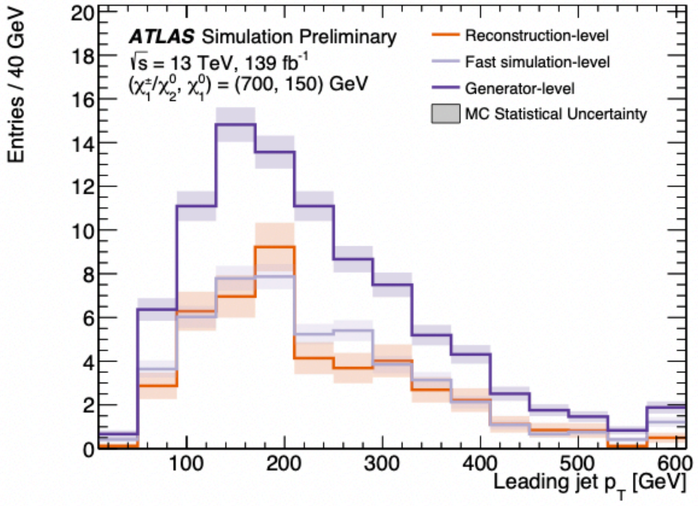
Available list of Analyses

EwkTwoLeptonRJ2016 ANA-SUSY-2017-03

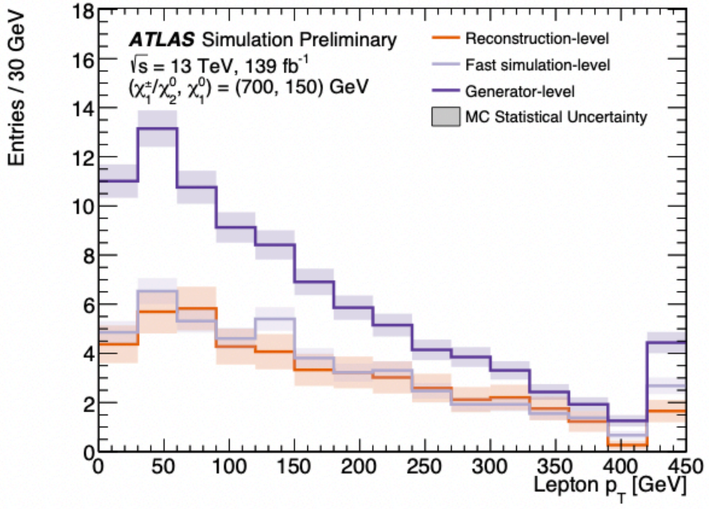
Simple Analysis

—Pub note

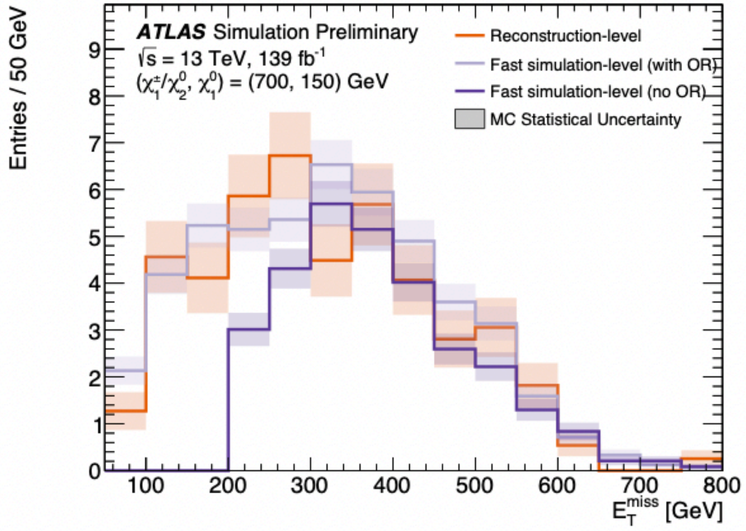
- Pub note described the tool and analysis pipelines for preservation
- Comparing the generator level analysis by SimpleAnalysis with the Monte Carlo samples at reconstruction level
- One lepton+ two b-jet analysis is used to demonstrate the validation



(a) Leading jet p_{T}



(b) Leading lepton p_{T}



(f) Missing transverse momentum

LLP analysis in Simple Analysis

- Currently few LLP analyses are available in SA
  - Disappearing track
  - ..More to come!
  - R22 is work in progress that could use more person input

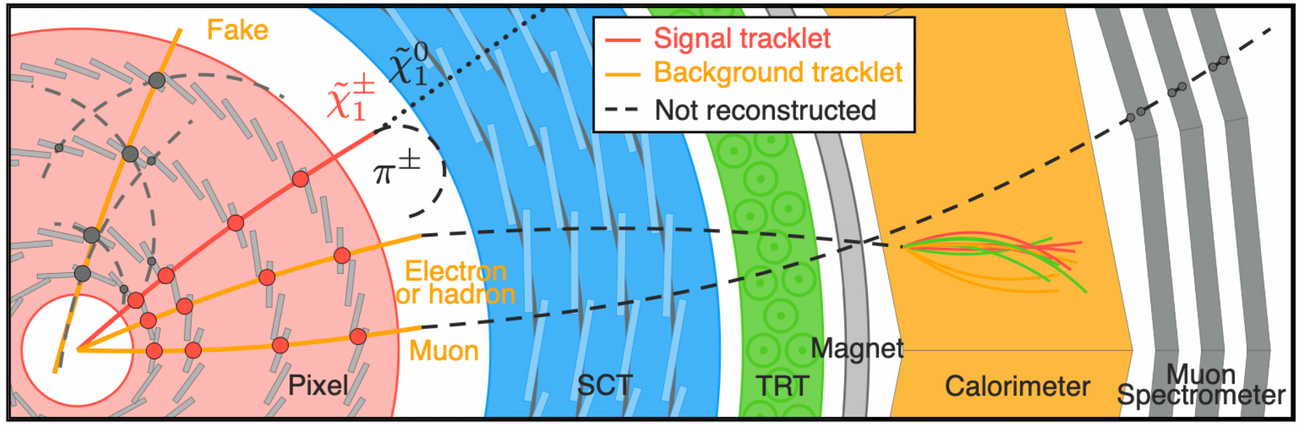
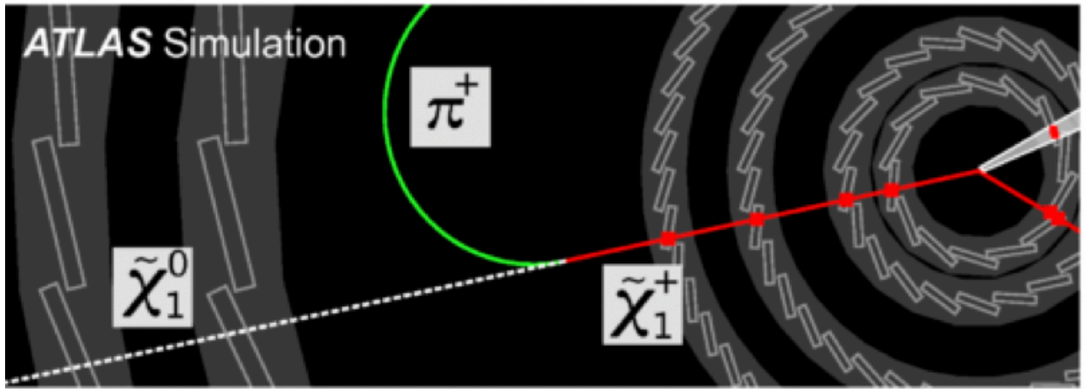


Image from Toshiaki Kaji



Future of Simple Analysis

-A full reusing pipeline(Mapyde)

- Fresh off the arrive today at 9am

REDUCE, REUSE, REINTERPRET: AN END-TO-END PIPELINE FOR RECYCLING PARTICLE PHYSICS RESULTS

A PREPRINT

Giordon Stark*[1], Camila Aristimuno Ots[1,2], and Mike Hance†[,1]

[1]University of California, Santa Cruz, Santa Cruz Institute for Particle Physics, 1156 High Street, Santa Cruz, CA 95064
[2]University of Southern California, Liquid Propulsion Laboratory, 854 Downey Way, Los Angeles, CA 90089

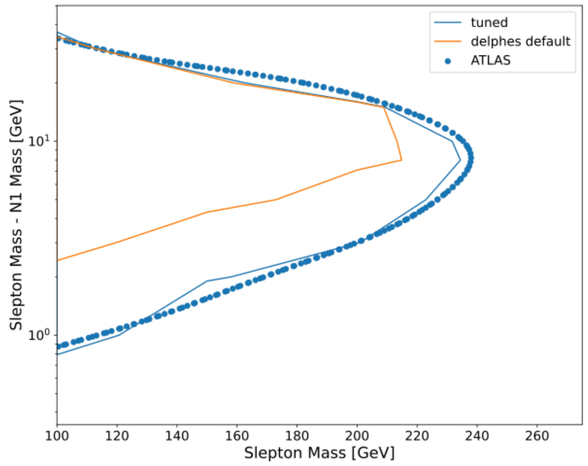
Mapyde

- Mayede: Performing a full pipeline of truth level analysis from event generation to statistical analysis!

Containing

- MADGRAPH5_AMC@NLO(MADGRAPH) [26, 27] (event generation)
- PYTHIA8 [28] (parton shower, hadronization, decays)
- DELPHES [29–31] (detector simulation)
- `SimpleAnalysis` [32, 33] (analysis description)
- `pyhf` [34, 35] (probability model fitting)

Reinterpretation Results from Mapyde



ATLAS Slepton result

Wino-bino reinterpretation from Mayede against ATLAS results

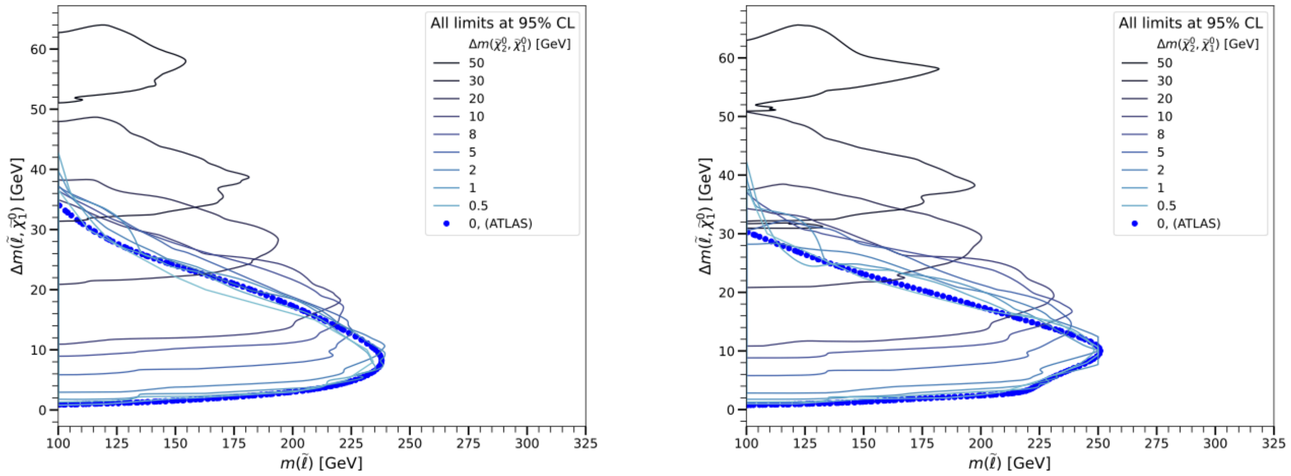
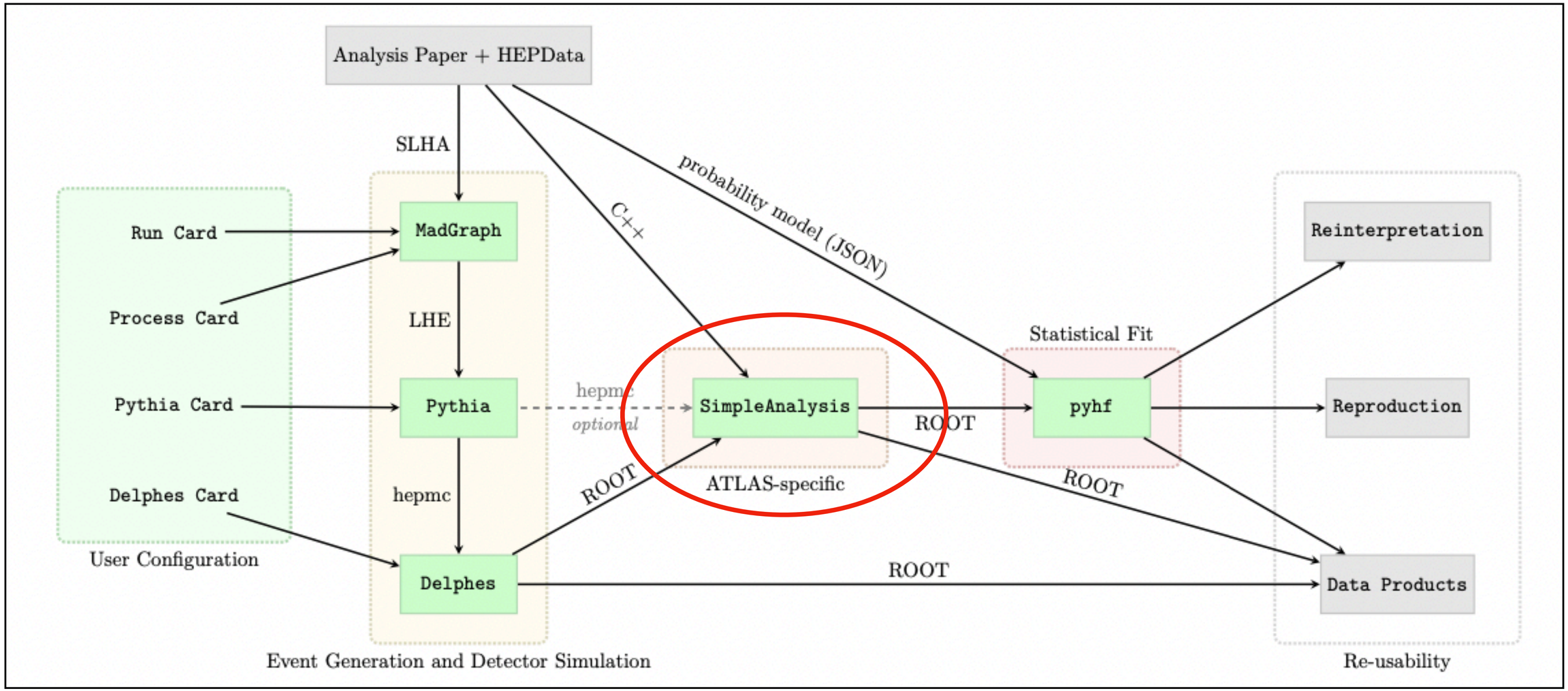


Figure 6: Expected (left) and observed (right) constraints on the slepton-wino-bino model. The model is parameterized by the slepton mass, the slepton-$\widetilde{\chi}_1^0$ mass splitting, and the $\widetilde{\chi}_2^0$-$\widetilde{\chi}_1^0$ splitting, and compared against the slepton-bino results from Ref. [40].

Simple Analysis pub note summary



Summary

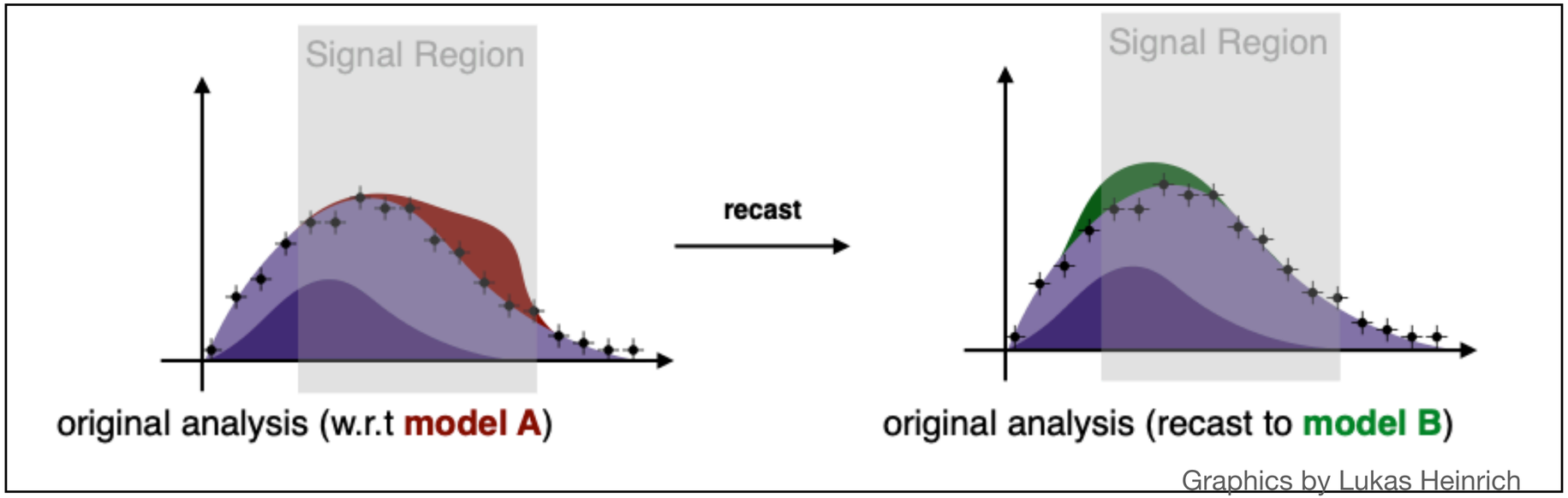
- Simple Analysis is fully **public,** available for theorists/experimentalists alike.
- Generator level pipeline preservation allow for analysis detail to be preserved.
- Reinterpretation made more straight forward.
- Reconstruction/detector effect can be added later on.
- Validation has been made on the generator/reco level kinematics plot.
- Mapyde, a new pipe that that would allow for experimental independent full reinterpretation is available today!

Back up

What is re-interpretation?

—And its importance in the BSM context

- Reinterpretation: The reusing of experimental results for alternative physics hypothesis
- **Advantages**: 1. Reuse of data/workflow/optimization 2. Less labor intensive 3. Time saving

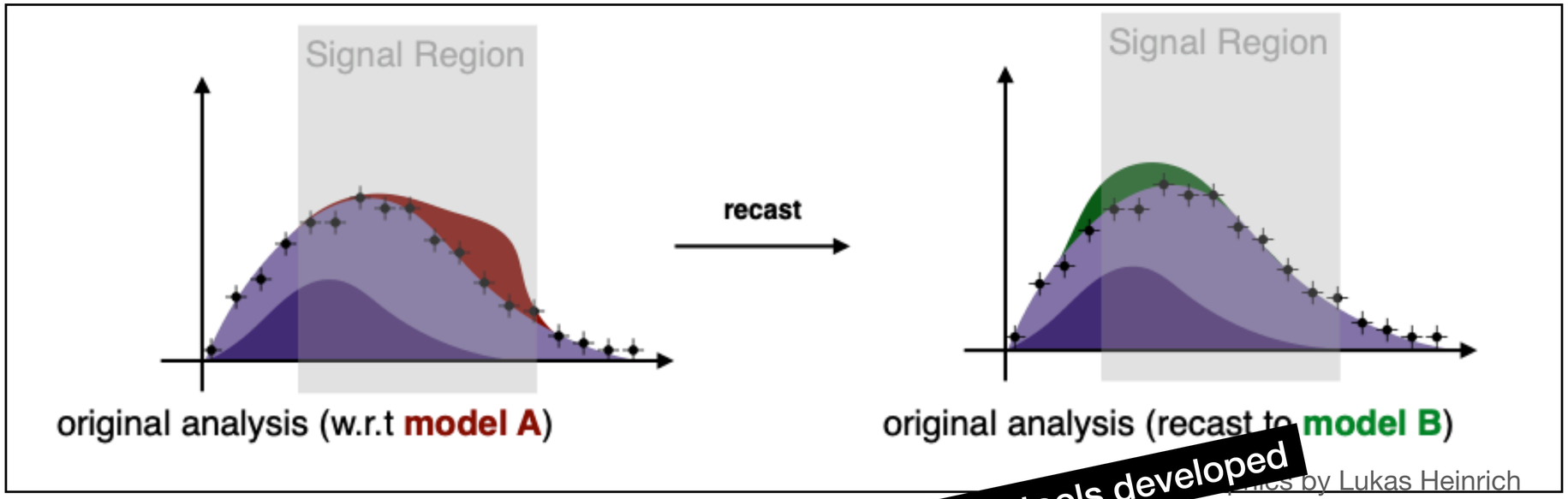


- **Two types of Reinterpretation**
- > 1. Re-interpretaion (Single data-> Many theory)
  Abundance of theoretical ideas vs experimental results available
- > 2. Combination (Many data -> Single theory)
  “Combination of results (E.G. Higgs combination)

What is re-interpretation?

—And its importance in the BSM context

- Reinterpretation: The reusing of experimental results for alternative physics hypothesis
- **Advantages**: 1. Reuse of data/workflow/optimization 2. Less labor intensive 3. Time saving



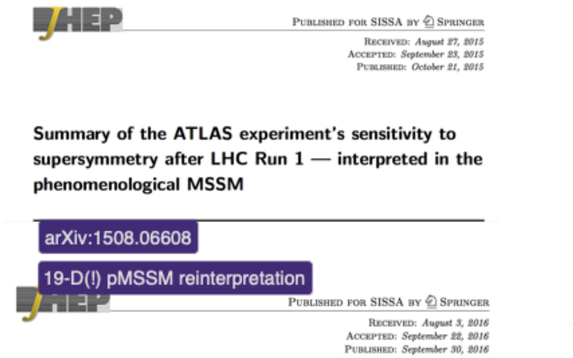
Slides by Lukas Heinrich

- **Two types of Reinterpretation**
- > 1. Re-interpretaion (Single theory)
  Abundance of theoretical ideas vs experimental results available
- > 2. Combination (Many data -> Single theory)
  “Combination of results (E.G. Higgs combination)

Both performed on ATLAS using various tools developed

Available reinterpretation tools

Within Experiment


JHEP
Published for SISSA by Springer
Received: August 27, 2015
Accepted: September 23, 2015
Published: October 21, 2015

Summary of the ATLAS experiment's sensitivity to supersymmetry after LHC Run 1 — interpreted in the phenomenological MSSM

arXiv:1508.06608

19-D(!) pMSSM reinterpretation

JHEP
Published for SISSA by Springer
Received: August 3, 2016
Accepted: September 22, 2016
Published: September 30, 2016

Dark matter interpretations of ATLAS searches for the electroweak production of supersymmetric particles in $\sqrt{s} = 8$ TeV proton-proton collisions

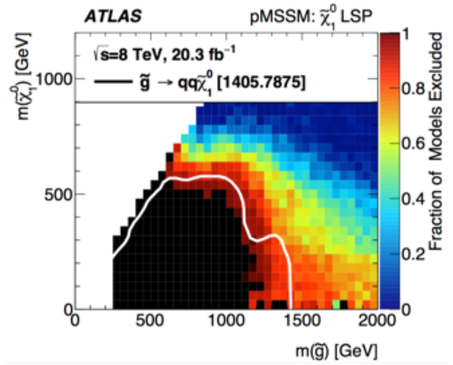
5-D scan of E

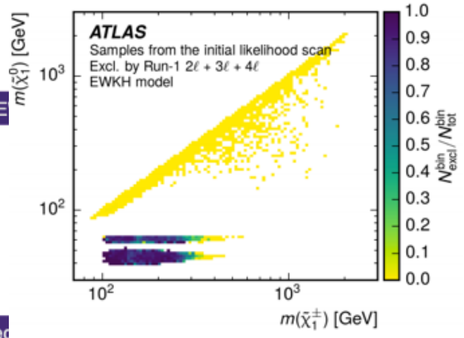
A re-interpretation of $\sqrt{s} = 8$ TeV ATLAS results on electroweak supersymmetry production to explore general gauge mediated models

The ATLAS Collaboration

3-D recast for General Gauge Me

ATLAS-CONF-2016-033





Outside of Experiment

Unfolded Reinterpretation(Without detector effect) Tools:

MADAnalysis 5

Other Homegrown HepMC based Analysis

MadAnalysis

Rivet

Folded Reinterpretation(With Detector Effect) Tools:

CHECKMATE

Homegrown Toy Detector Approximation

Delphes based Analysis

Experiments preserve analyses in full fidelity internally. Re-interpret on-demand for models that are deemed sufficiently promising.

Tools used by ATLAS for reinterpretation

—Introducing the players

Precision Level	Preserved Item	Tools/Service used by ATLAS
0	Data product	HEPData
1	Likelihood	Simple likelihood JSON HistFactory
2	Truth Level Pipeline	SimpleAnalysis/Rivet
10	Full Reco Pipeline	Recast
10+	Preservation of ML Model/optimization	ONNX

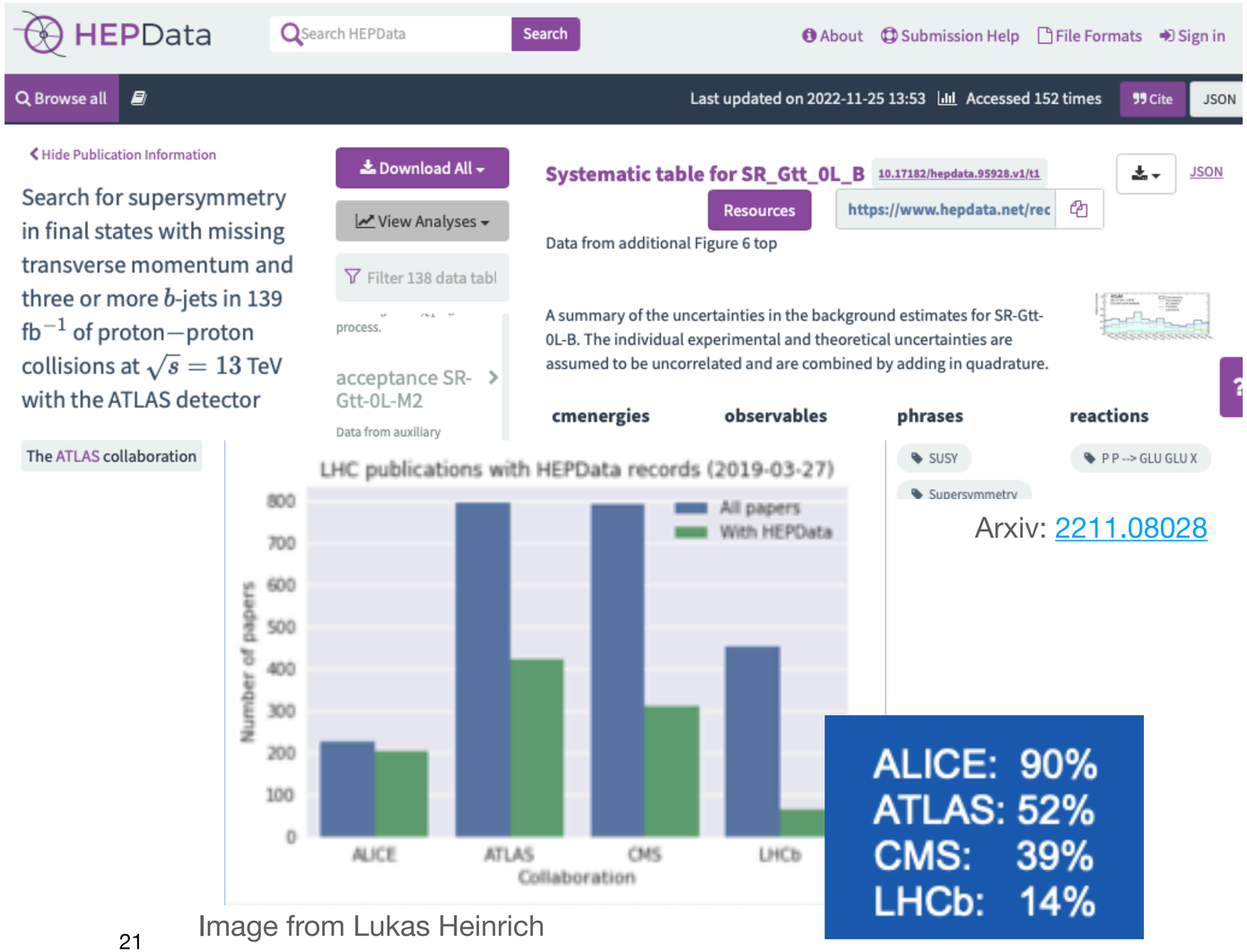
++Increased level of Information/ Precision

Increase precision enable better reinterpretation from original analysis!

Level 0: HEPData

—Storing of experimental data product from analyses for reinterpretation

- Previously: Table-based webpage (40 year-old):
- New digital library : Paper
  - Database: Include Efficiency Map, yields, uncertainties, outflow tables
  - Reinterpretation tools: Rivet and SimpleAnalysis code, preservation of ML
  - Webservice
- Made an ATLAS requirement for publication
- Centralized tool across other HEP experiments



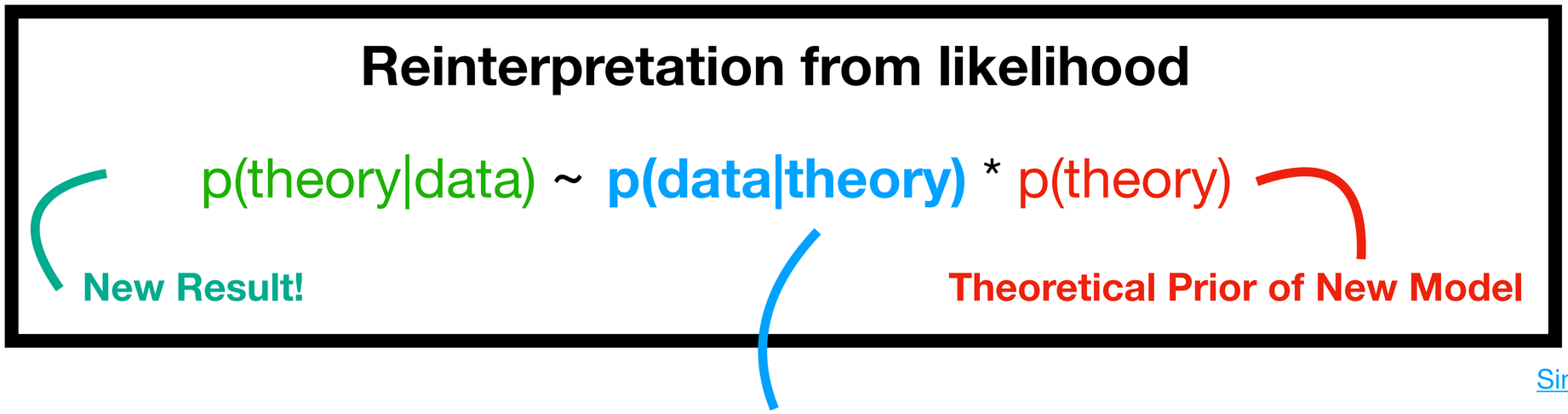
Arxiv: 2211.08028

Image from Lukas Heinrich

Level 1: Statistical Model (Simple Likelihood)

—Saving the most of our experiment results in one variable

- Previously, often HEPData information is used to construct approximate likelihood
- **Likelihood: Great bang for the buck**: encapsulate information about Limits, data/MC plots, yield take, and systematics. ->Best object that can be provided!
- ATLAS started providing them on HEPData since 2019 (SUSY)



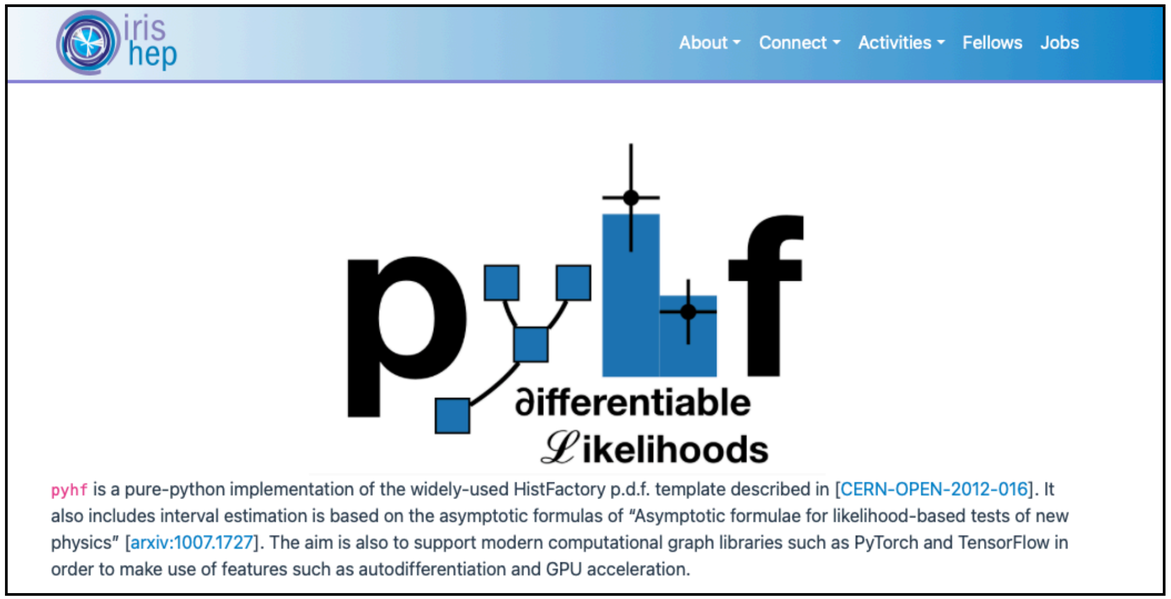
Simple Likelihood Pubnote

Level 1: HistFactory JSON

—pyhf as a pure-Python implementation of HistFactory with JSon

See "Publishing statistical models discussion" on Tuesday"

- PyHF: HistFactory in full python
- Increasingly used by ATLAS analyses (24 analyses published-up-to-date)
- Results in easy to read json format: Database
- Reinterpretation:
  - 1. Publishes statistical model (Likelihood) in json file format
  - 2. New version of SModel and MadAnalysis5 allow for automatic reinterpretation from pyHF json file



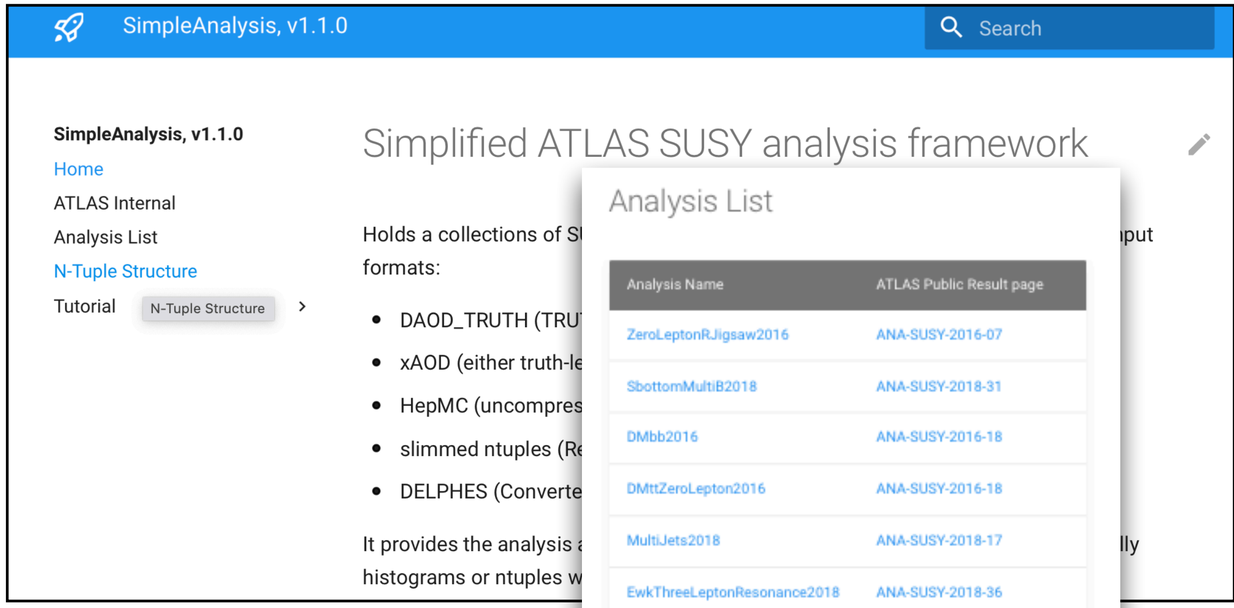
GitLab

Source Code

Level 2: SimpleAnalysis
—Truth level pipeline preservation

See tutorial by Giordon Stark Wednesday

- Preserves the analysis code pipeline
- Allow for reinterpretation up to the truth level
  - Detector effect can be added: public simulation like Delphes or internally by ATLAS fast simulation
- Experimentalist friendly input format (HepMC, DAOD, ROOT n-tuple, DELPHES)
- Large list of ATLAS SUSY analyses available:
- For **public use**: can be used external to ATLAS



SimpleAnalysis: Webpage

GitLab

Source Code

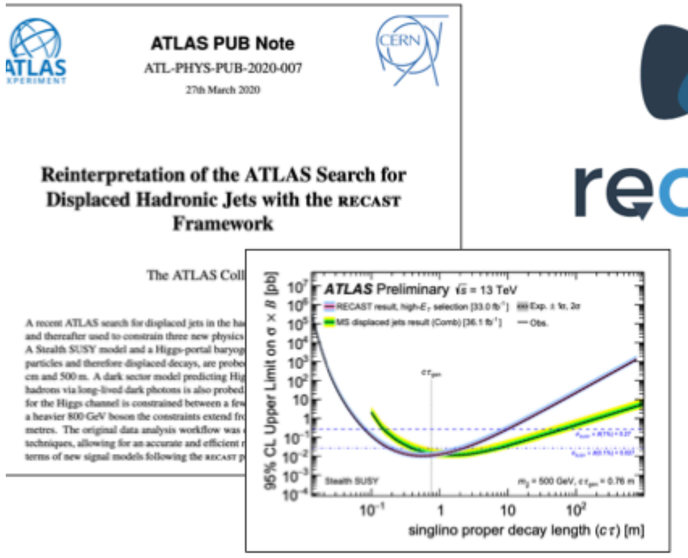
SimpleAnalysis PubNote

Available list of Analyses

Level 10: Recast

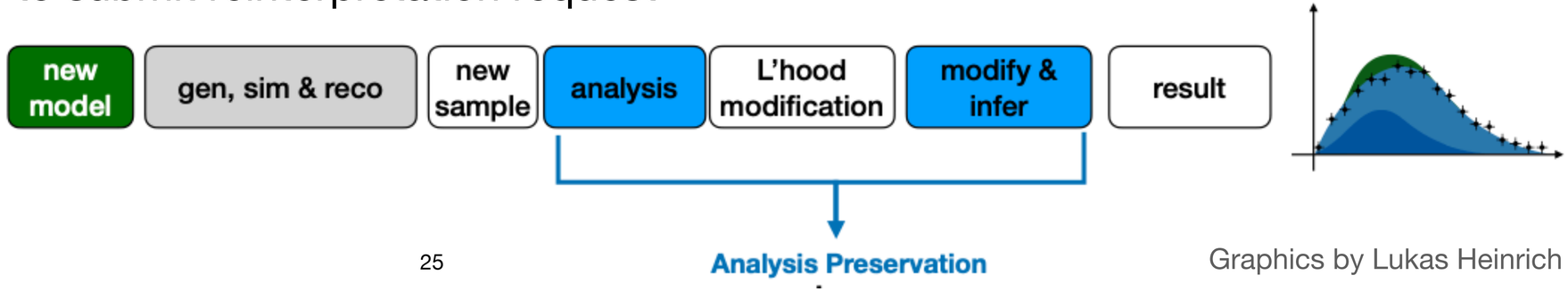
—Full Analysis pipeline preservation

- Preservation of original analysis pipeline for re-interpretaion
- Contain full detector simulation information + original code/ workflow
- Reinterpretation with full workflow + simulation effect
- Made a requirement for publication in Exotics, HDBS and SUSY
- Implementation Internal to ATLAS: some reinterpretation results published with the tool
- Software fully open source!
- Future: Allow for an interface to submit reinterpretation request online



recast

preservation of the original analysis pipeline for collaboration-approvable results



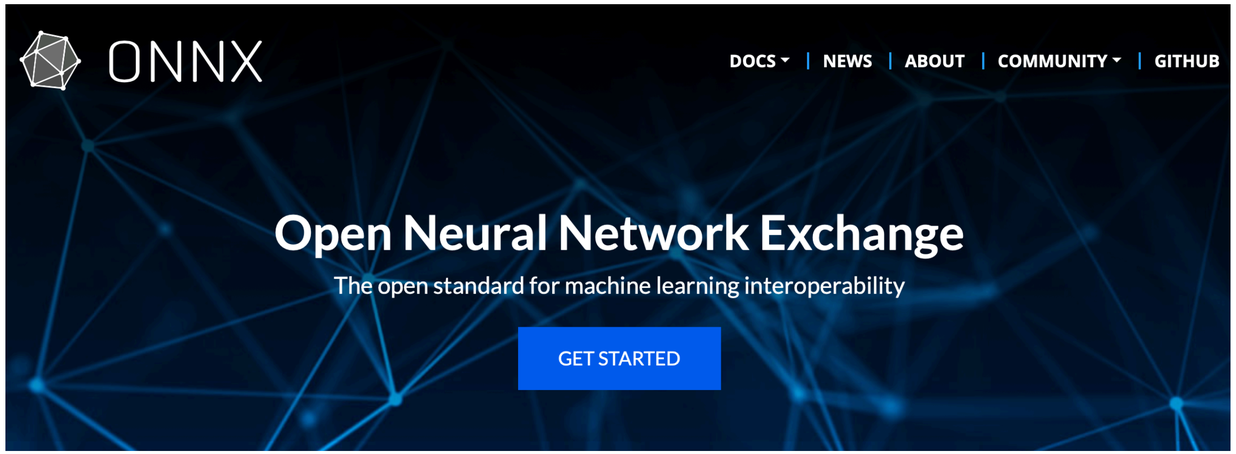
Graphics by Lukas Heinrich

Level 10+: ONNX/NN

See talk by Dan Guest later today

—Beyond usual pipeline preservation: Reuse of ML training results

- ML development significantly improve analyses sensitivity. Theorists would like to make use of the BDTs/NN.
- ONNX(Open Neural Network Exchange)
- BDT preservation -> convert to standalone c++ in SimpleAnalysis
- Non-BDT NN -> Preservation by Serializing through Onnx
- Can be preserved with SimpleAnalysis to be reuse
- 2 analyses already have ML encapsulated in the workflow



3 b-jets+

Search for supersymmetry in final states with missing transverse momentum and three or more b-jets in 139 fb^{-1} of proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector

The ATLAS Collaboration

ANA-SUSY-2018-30 SA with Onnx

One Lepton MultiJets

Search for R-parity-violating supersymmetry in a final state containing leptons and many jets with the ATLAS experiment using $\sqrt{s}$ = 13 TeV proton–proton collision data

The ATLAS Collaboration

ANA-SUSY-2019-04 SA with Onnx

Two existing analyses already contain Onnx

GitLab

Source Code