Open Likelihoods Workshop

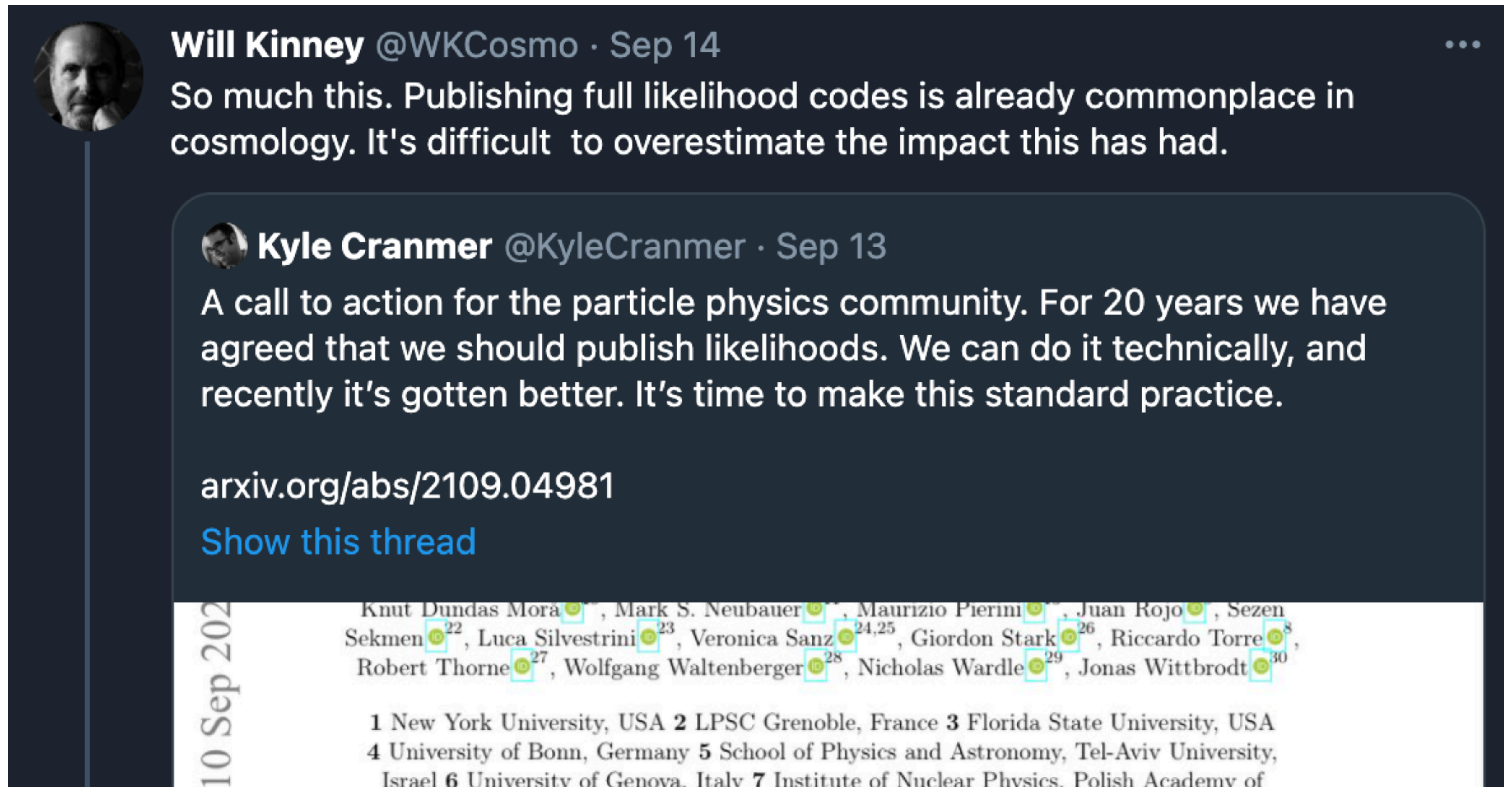
Lukas Heinrich, CERN

Introduction

Welcome to the public likelihoods workshop!

First of its kind in HEP.

Result of a lot of progress in the community over past few years




arXiv:2109.04981

SciPost Physics | Submission

09.04981v1 [hep-ph] 10 Sep 2021

Publishing statistical models: Getting the most out of particle physics experiments

Kyle Cranmer[1*], Sabine Kraml[2‡], Harrison B. Prosper[3§] (editors), Philip Bechtle[4], Florian U. Bernlochner[4], Itay M. Bloch[5], Enzo Canonero[6], Marcin Chrzaszcz[7], Andrea Coccaro[8], Jan Conrad[9], Glen Cowan[10], Matthew Feickert[11], Nahuel Ferreiro Iachellini[12,13] Andrew Fowlie[14], Lukas Heinrich[15], Alexander Held[1], Thomas Kuhr[13,16], Anders Kvellestad[17], Maeve Madigan[18], Farvah Mahmoudi[15,19], Knut Dundas Morå[20], Mark S. Neubauer[11], Maurizio Pierini[15], Juan Rojo[8], Sezen Sekmen[22], Luca Silvestrini[23], Veronica Sanz[24,25], Giordon Stark[26], Riccardo Torre[8], Robert Thorne[27], Wolfgang Waltenberger[28], Nicholas Wardle[29], Jonas Wittbrodt[30]

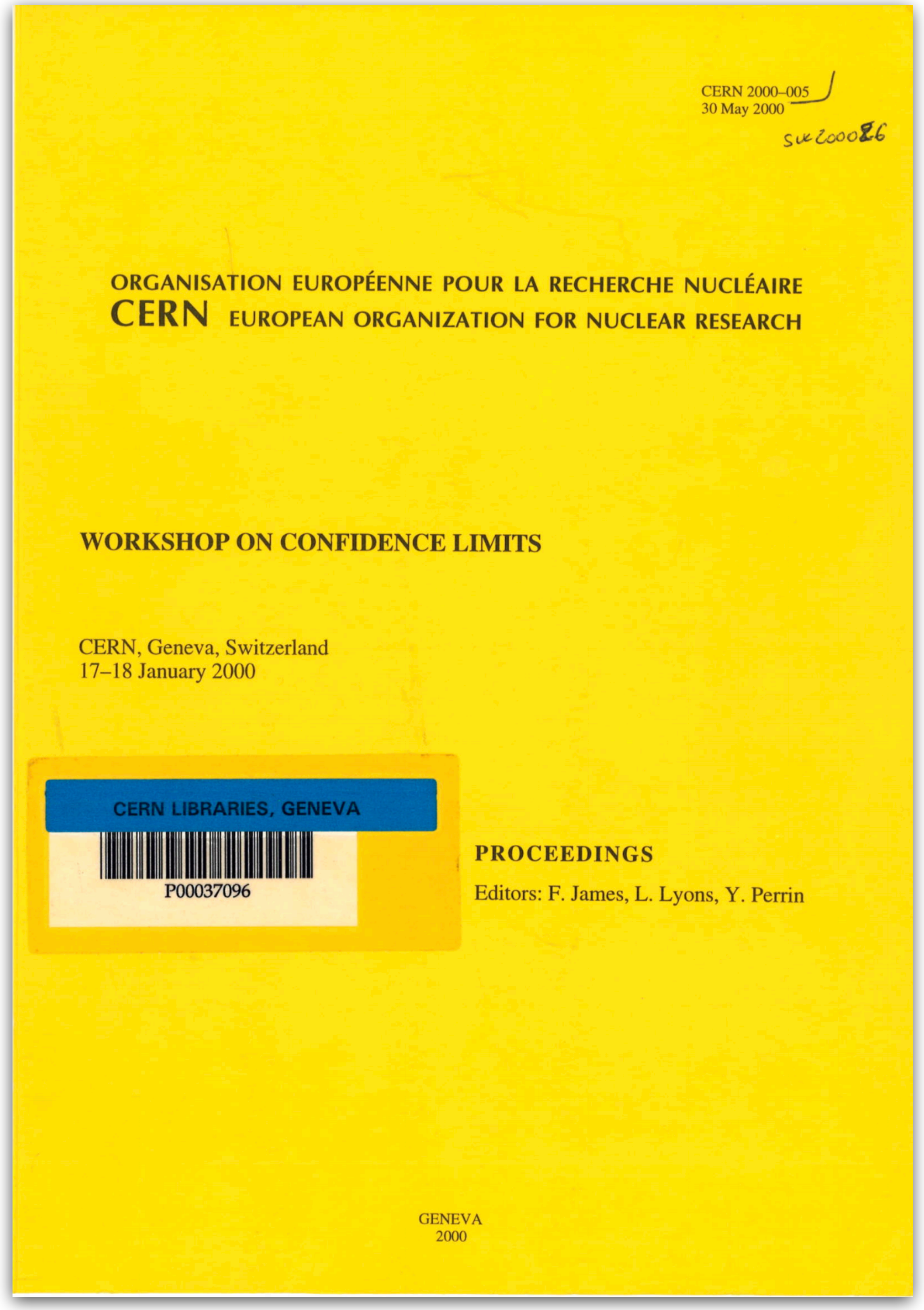
1 New York University, USA **2** LPSC Grenoble, France **3** Florida State University, USA **4** University of Bonn, Germany **5** School of Physics and Astronomy, Tel-Aviv University, Israel **6** University of Genova, Italy **7** Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland **8** INFN, Sezione di Genova, Italy **9** Oskar Klein Centre, Stockholm University, Sweden **10** Royal Holloway, University of London, UK **11** University of Illinois at Urbana-Champaign, USA **12** Max Planck Institute for Physics, Munich, Germany **13** Exzellenzcluster ORIGINS, Garching, Germany **14** Nanjing Normal University, Nanjing, PRC **15** CERN, Switzerland **16** Ludwig-Maximilians-Universität München, Germany **17** University of Oslo, Norway **18** DAMTP, University of Cambridge, UK **19** Lyon University, France **20** Columbia University 10027, USA **21** VU Amsterdam and Nikhef, The Netherlands **22** Kyungpook National University, Daegu, Korea **23** INFN, Sezione di Roma, Italy **24** University of Sussex, UK **25** IFIC, Universidad de Valencia-CSIC, Spain **26** SCIPP, UC Santa Cruz, CA, USA **27** University College London, UK **28** HEPHY and University of Vienna, Austria **29** Imperial College London, UK **30** Lund University, Sweden

* kyle.cranmer@nyu.edu, ‡ sabine.kraml@lpsc.in2p3.fr, § harry@hep.fsu.edu

Introduction

Culmination of 20 year effort to change data practices in HEP

- now lot of momentum and even money to push further
- this talk: short review of motivation & background


CERN 2000–005
30 May 2000

ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

WORKSHOP ON CONFIDENCE LIMITS

CERN, Geneva, Switzerland
17–18 January 2000

CERN LIBRARIES, GENEVA
P00037096

PROCEEDINGS
Editors: F. James, L. Lyons, Y. Perrin

GENEVA
2000

2000


SciPost Physics — Submission

arXiv:2109.04981v1 [hep-ph] 10 Sep 2021

Publishing statistical models: Getting the most out o
physics experiments

Kyle Cranmer[1*], Sabine Kraml[2‡], Harrison B. Prosper[3§] (edit
Philip Bechtle[4], Florian U. Bernlochner[4], Itay M. Bloch[5], Enzo Canone
Chrzaszcz[7], Andrea Coccaro[8], Jan Conrad[9], Glen Cowan[10], Matthew
Nahuel Ferreiro Iachellini[12,13] Andrew Fowlie[14], Lukas Heinrich[15], Alexa
Thomas Kuhr[13,16], Anders Kvellestad[17], Maeve Madigan[18], Farvah Ma
Knut Dundas Morå[20], Mark S. Neubauer[11], Maurizio Pierini[15], Juan R
Sekmen[22], Luca Silvestrini[23], Veronica Sanz[24,25], Giordon Stark[26], Ric
Robert Thorne[27], Wolfgang Waltenberger[28], Nicholas Wardle[29], Jonas

1 New York University, USA 2 LPSC Grenoble, France 3 Florida State Uni
4 University of Bonn, Germany 5 School of Physics and Astronomy, Tel-Avi
Israel 6 University of Genova, Italy 7 Institute of Nuclear Physics, Polish
Sciences, Krakow, Poland 8 INFN, Sezione di Genova, Italy 9 Oskar Kle
Stockholm University, Sweden 10 Royal Holloway, University of London, UK
of Illinois at Urbana-Champaign, USA 12 Max Planck Institute for Physic
Germany 13 Exzellenzcluster ORIGINS, Garching, Germany 14 Nanjing Nor
Nanjing, PRC 15 CERN, Switzerland 16 Ludwig-Maximilians-Universität
Germany 17 University of Oslo, Norway 18 DAMTP, University of Cambridg
University, France 20 Columbia University 10027, USA 21 VU Amsterdam a
Netherlands 22 Kyungpook National University, Daegu, Korea 23 INFN, Sez
Italy 24 University of Sussex, UK 25 IFIC, Universidad de Valencia-CSIC, Sp
UC Santa Cruz, CA, USA 27 University College London, UK 28 HEPHY an
Vienna, Austria 29 Imperial College London, UK 30 Lund University,

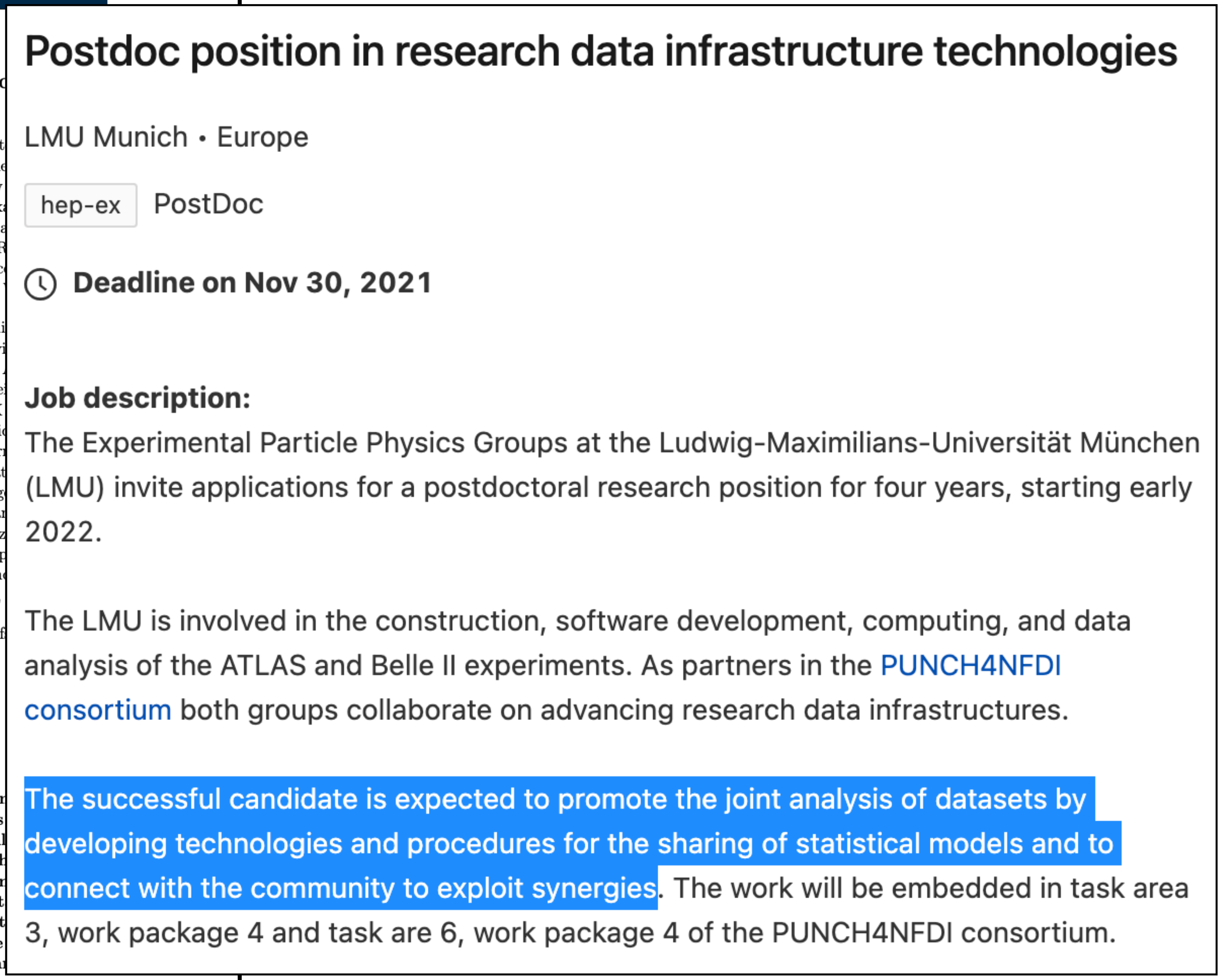
* kyle.cranmer@nyu.edu, ‡ sabine.kraml@lpsc.in2p3.fr, § harry@hep.f
(mailing list: open-likelihoods@cern.ch)

September 9, 2021

Abstract

The statistical models used to derive the results of experimental a
incredible scientific value and are essential information for analysis
and reuse. In this paper, we make the scientific case for systematical
the full statistical models and discuss the technical developments th
practical. By means of a variety of physics cases — including parton
functions, Higgs boson measurements, effective field theory interpret
searches for new physics, heavy flavor physics, direct dark matter det
averages, and beyond the Standard Model global fits — we illustrate
information on the statistical modelling can enhance the short- a
impact of experimental results.

1


Postdoc position in research data infrastructure technologies

LMU Munich • Europe

hep-ex PostDoc

Deadline on Nov 30, 2021

Job description:

The Experimental Particle Physics Groups at the Ludwig-Maximilians-Universität München (LMU) invite applications for a postdoctoral research position for four years, starting early 2022.

The LMU is involved in the construction, software development, computing, and data analysis of the ATLAS and Belle II experiments. As partners in the PUNCH4NFDI consortium both groups collaborate on advancing research data infrastructures.

The successful candidate is expected to promote the joint analysis of datasets by developing technologies and procedures for the sharing of statistical models and to connect with the community to exploit synergies. The work will be embedded in task area 3, work package 4 and task are 6, work package 4 of the PUNCH4NFDI consortium.

https://inspirehep.net/jobs/1955443

arXiv:2109.04981

2021

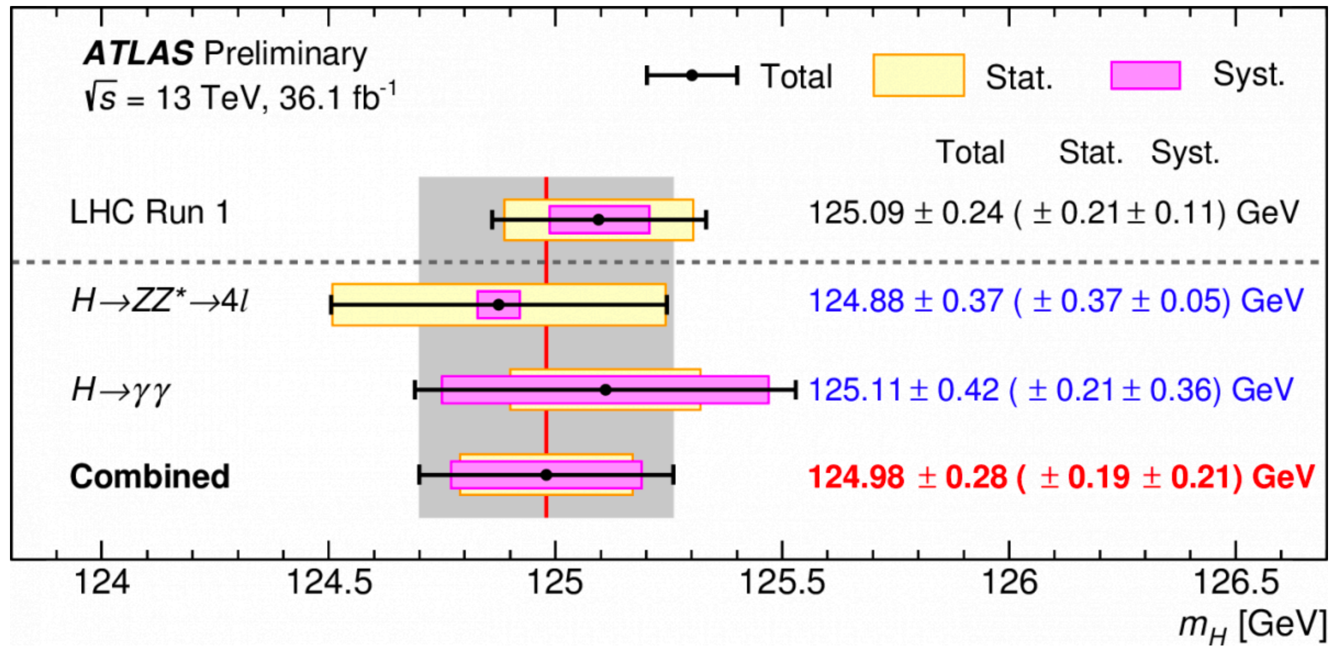
Introduction

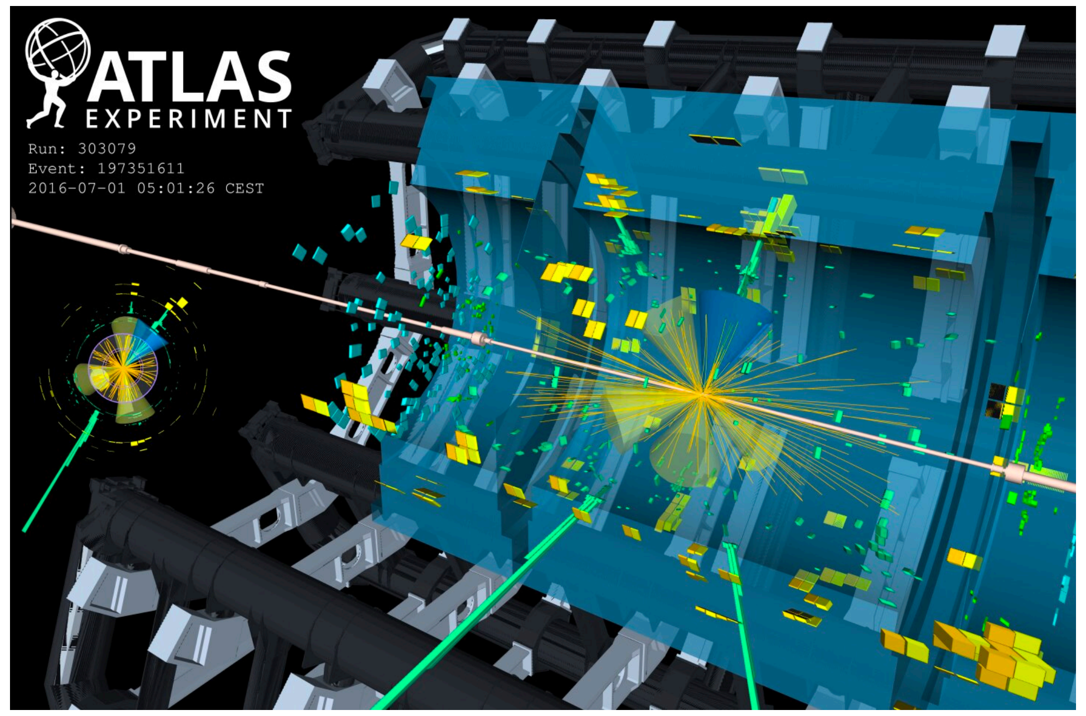
The likelihoods is the basis almost all experimental results.

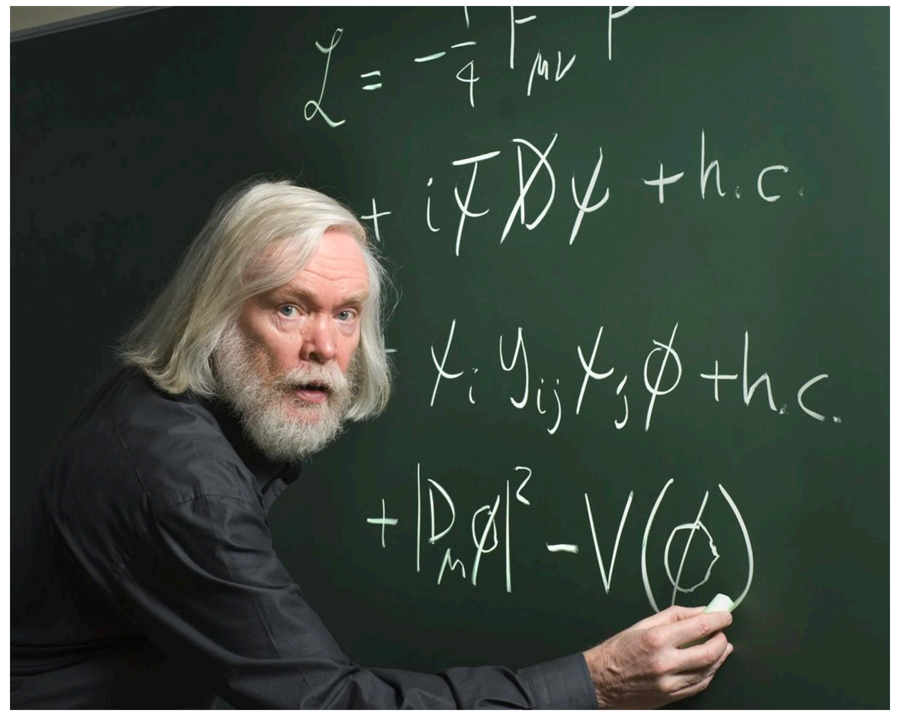
- **from that perspective it seems obvious that this should be a "data product"**
- **but both sociological & technical challenges**

$$p(\text{theory}|\text{data}) = \frac{p(\text{data}|\text{theory})}{p(\text{data})} p(\text{theory})$$

Posterior ↑ Evidence ↑ Prior ↑







results / insight **experimentalists** **theorists**

Big Picture Question: What to publish

We use term likelihood but term a bit ambiguous

Textbook definition: $L(\theta) = p(x\,|\,\theta)_{x \text{ fixed}}$

a pure function of the parameters θ with the data "baked it"

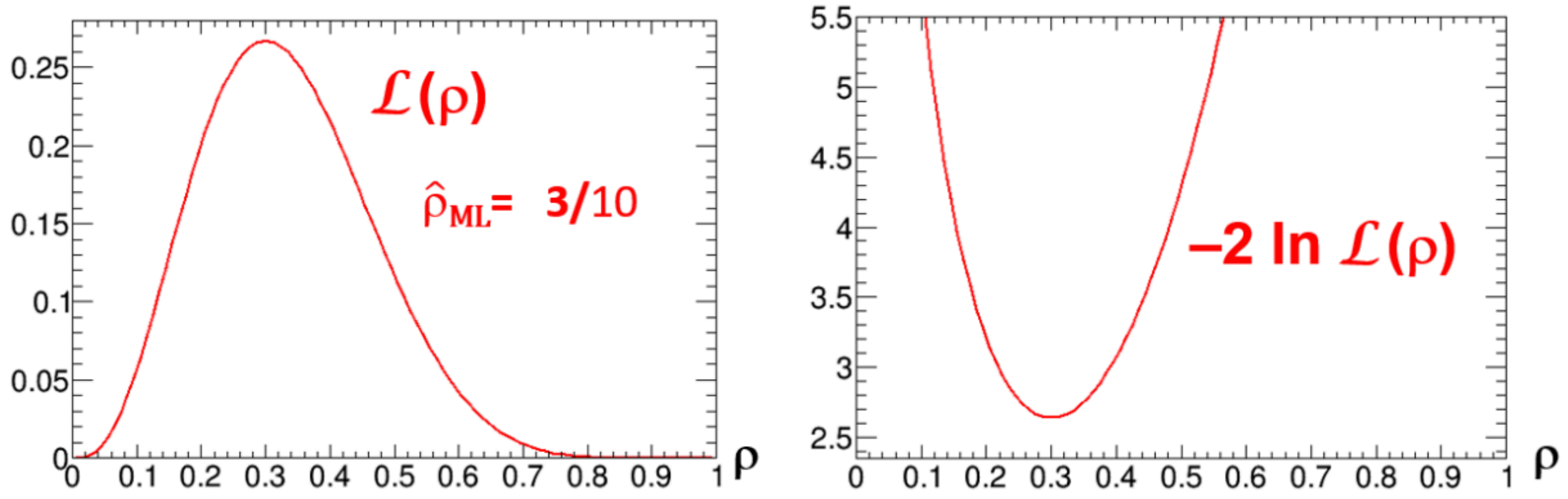


Figure 7: (left) Likelihood function $\mathcal{L}(\rho)$ for $n_{\mathrm{on}} = 3$ successes in $n_{\mathrm{tot}} = 10$ trials in the binomial model of Eqn. 8. (right) Looking ahead to Section 8, the plot of $-2\ln\mathcal{L}(\rho)$.

arXiv:1807.05996

At first glance seems ok
Likelihood Principle (LP) : all inference is only a function of likelihood

Likelihood principle

From Wikipedia, the free encyclopedia

In statistics, the **likelihood principle** is the proposition that, given a statistical model, all the evidence in a sample relevant to model parameters is contained in the likelihood function.

A likelihood function arises from a probability density function considered as a function of its distributional parameterization argument. For example, consider a model which gives the

Profile likelihood

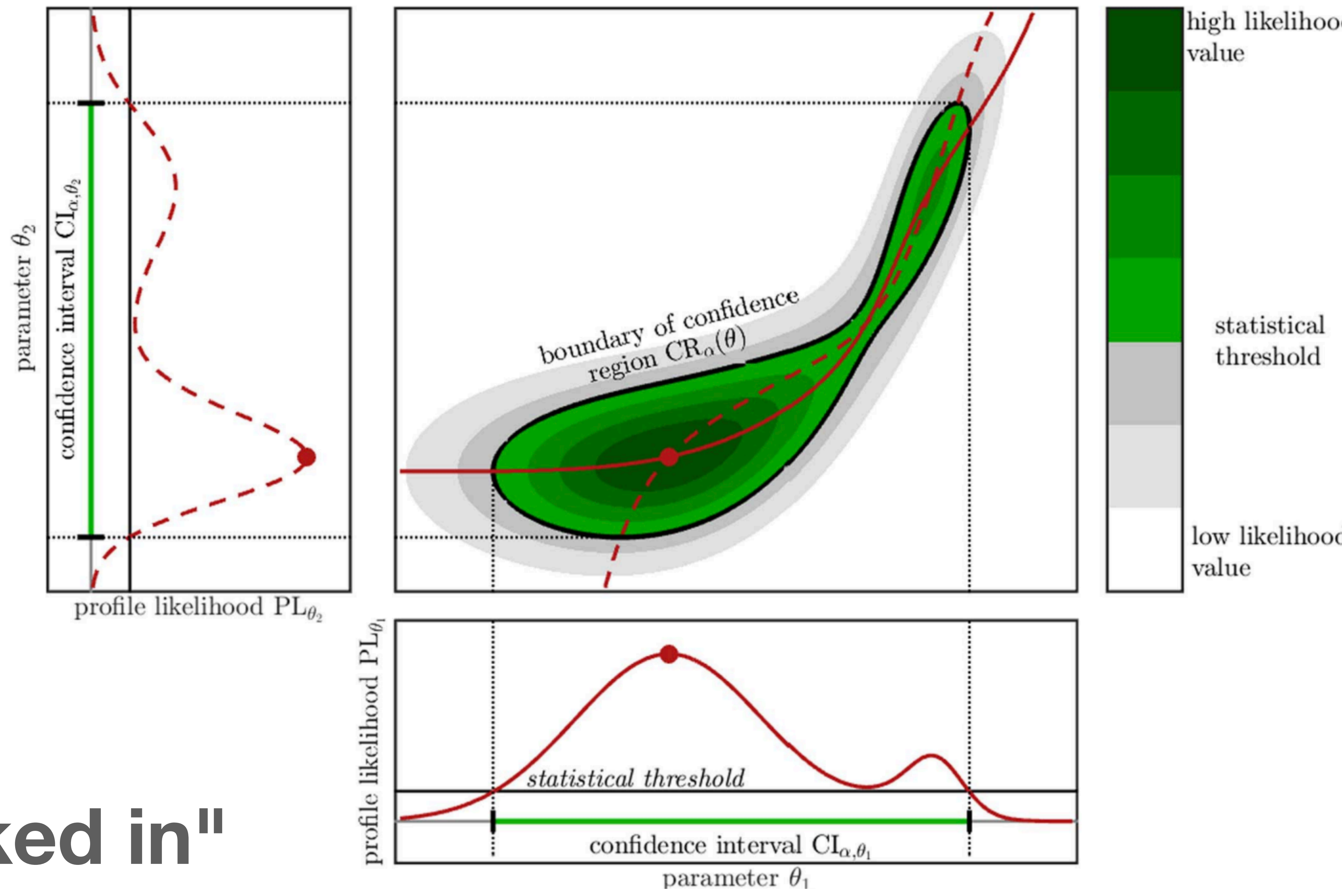
HEP analyses have many systematic uncertainties.

- **We usually compute the "profile likelihood" as a function of just the parameters of interest**

$$L(\theta) = L(\mu, \nu)$$

$$\downarrow$$

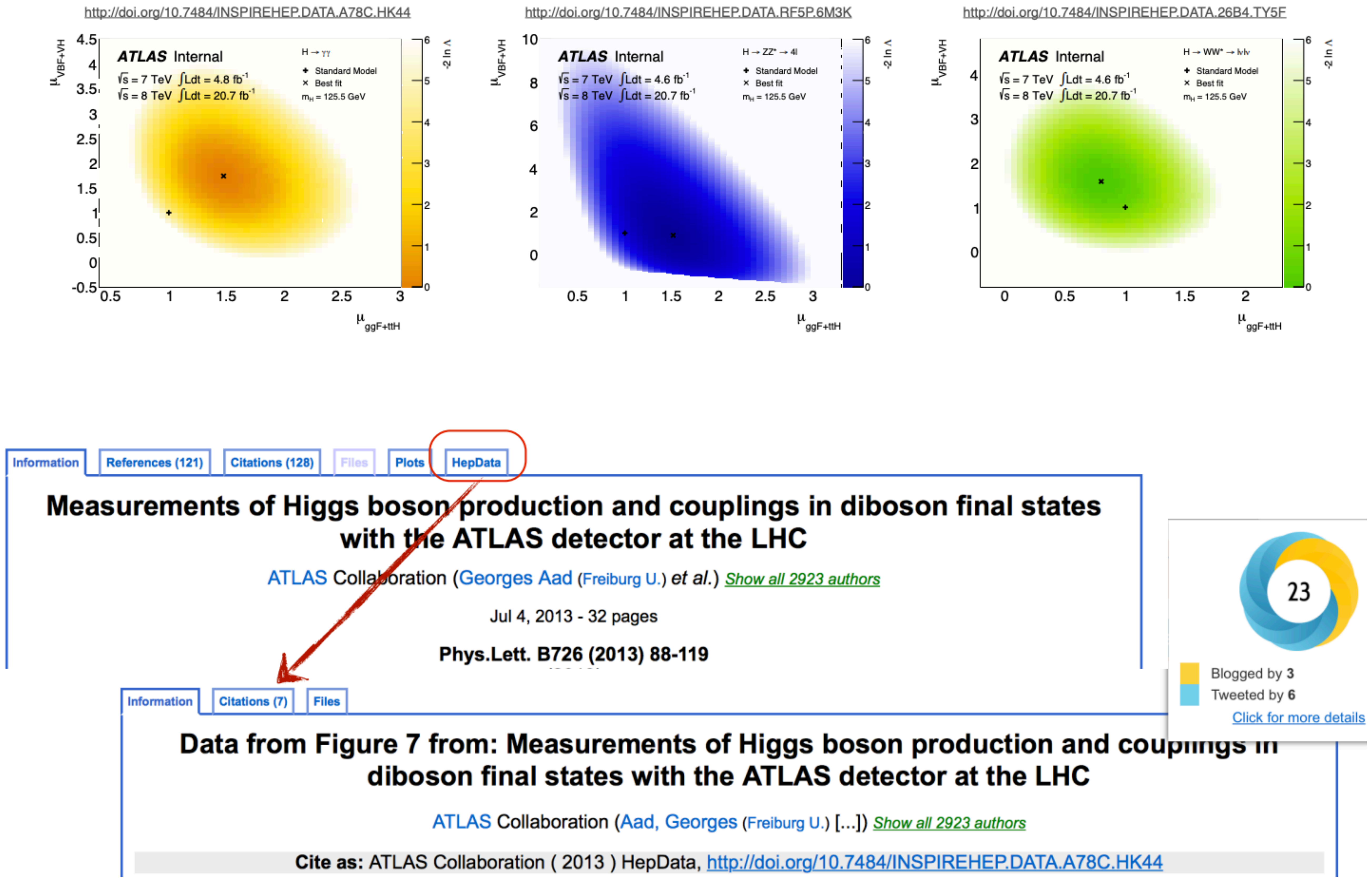
$$L_{\text{profile}}(\mu) = L(\mu, \hat{\hat{\nu}}(x))$$



- **To simplify, maybe we just publish this instead?**
- **but now not only data but also systematics "baked in"**

First Breakthrough

- **ATLAS started publishing profile likelihood scans in 2012**
- **but did not become common practice also because publishing profile likelihood has limitations what you can do with it**



Issues with Publishing Likelihood Scans

Publishing Likelihood Scans has disadvantages

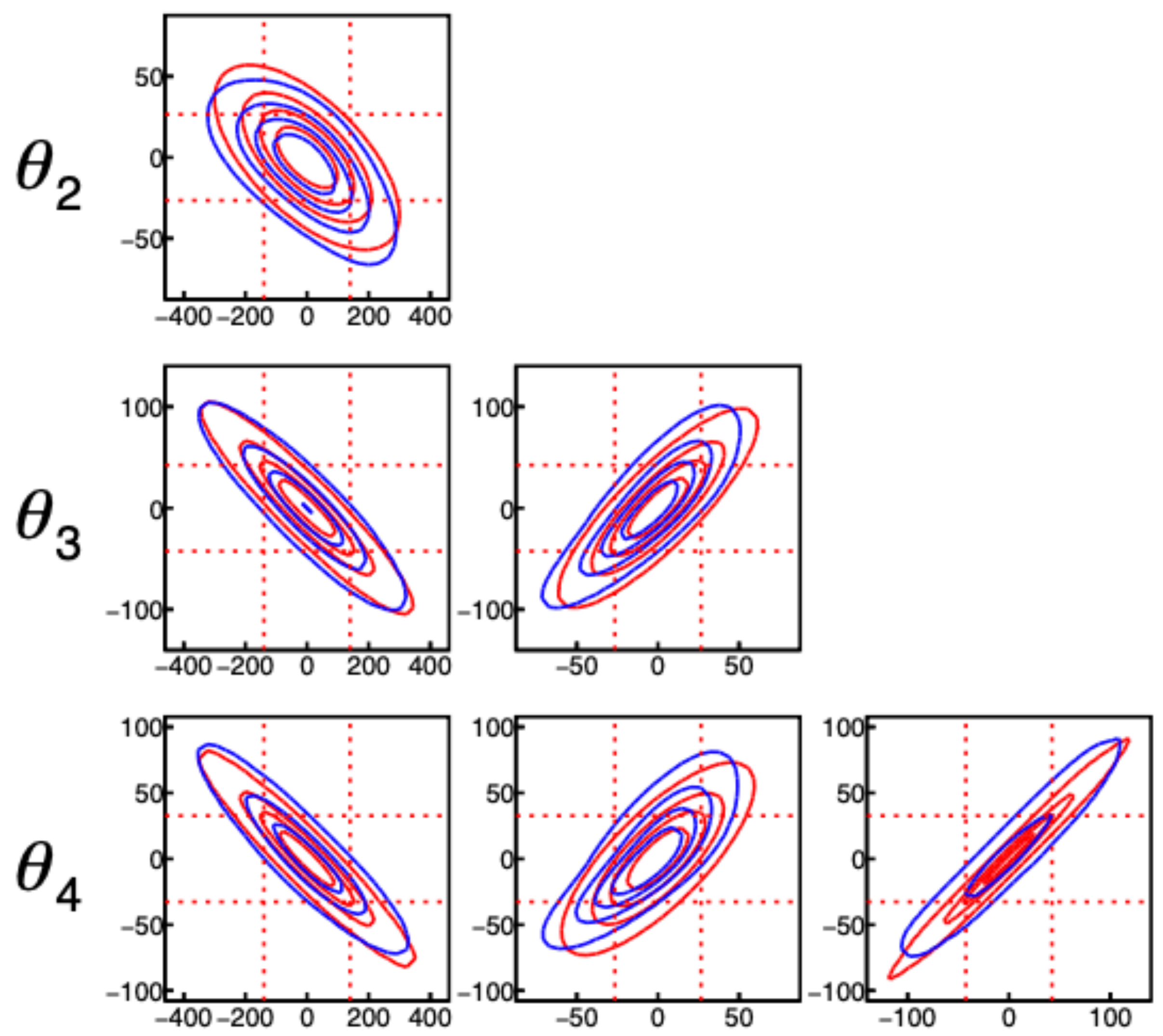
- data is baked in
  - cannot evaluate likelihood on new data
  - cannot sample from the model

→ Excludes some statistical techniques beyond LP
- Frequentist test statistic distributions
- Bayesian prior/posterior predictive distributions

- in profile likelihoods, nuisance parameter are fixed
  - cannot statistically combine multiple profile likelihoods targeting the same parameters of interest if they share nuisance parameters

- structure of the likelihood is opaque: cannot "patch" the likelihood for reinterpretation

Another Simplified Approach

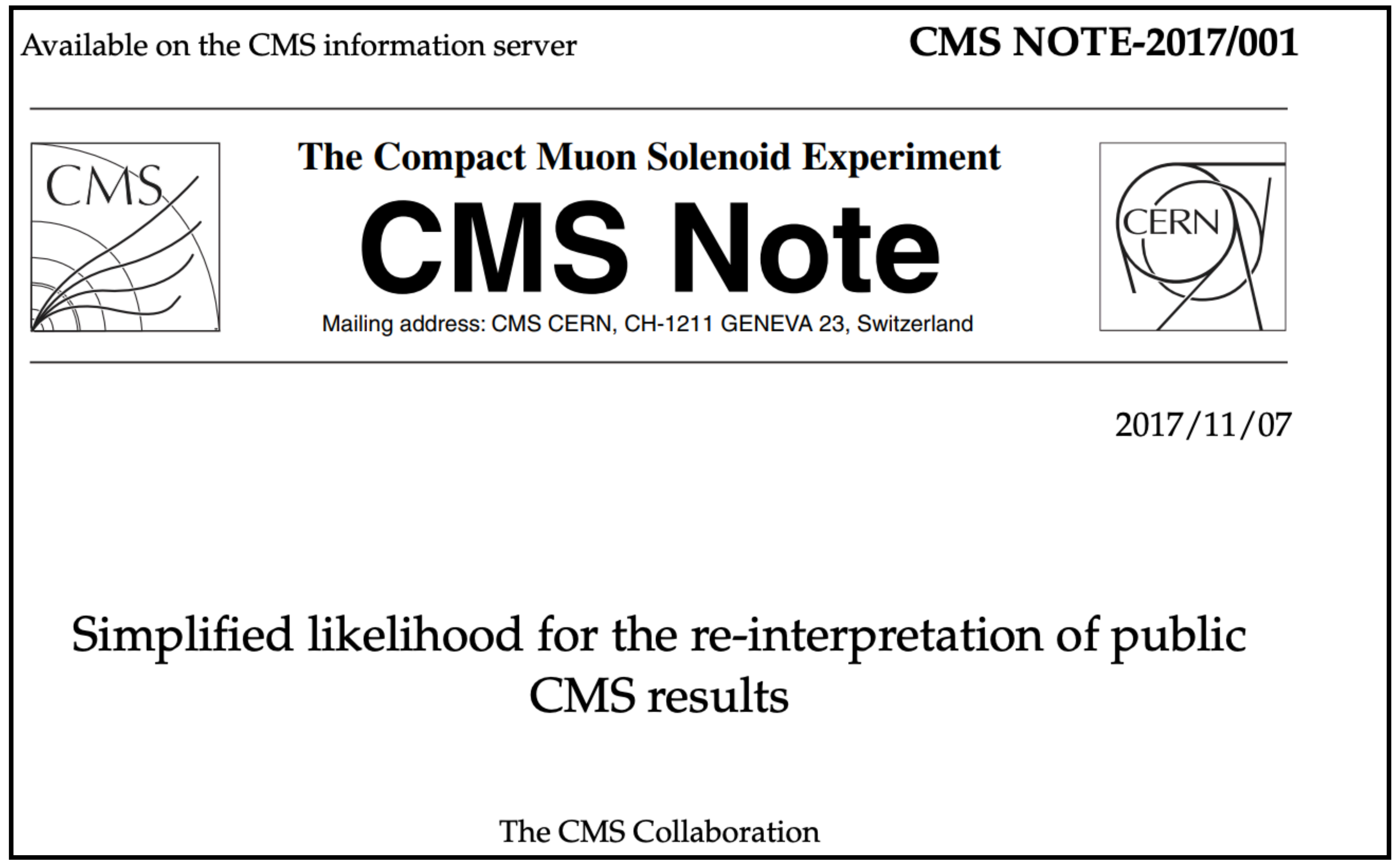


Instead of "full precision" profile likelihood another approach is use a simplified model

- 2017 CMS has started to publish"simplified likelihoods"
- Not a ilkelilhod scan: in principle have a fully defined but simplified statistical model

$$\mathcal{L}_S(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu \cdot s_i + b_i + \theta_i)^{n_i} e^{-(\mu \cdot s_i + b_i + \theta_i)}}{n_i!} \cdot \exp\left(-\frac{1}{2}\boldsymbol{\theta}^T \mathbf{V}^{-1} \boldsymbol{\theta}\right),$$

- solves some issues with Likelihood scans but loses precision & still not combinable
  - simplified L'hood derived in analyis-specific setting
  - not the likelihood the experiment uses for its results



Available on the CMS information server CMS NOTE-2017/001

The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

2017/11/07

Simplified likelihood for the re-interpretation of public CMS results

The CMS Collaboration

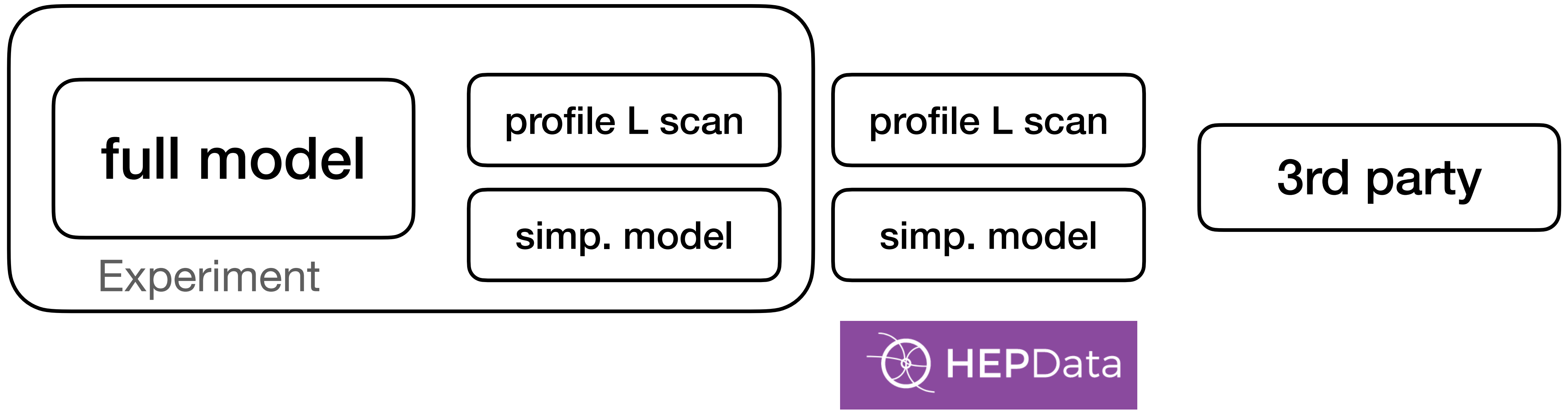
So what gives

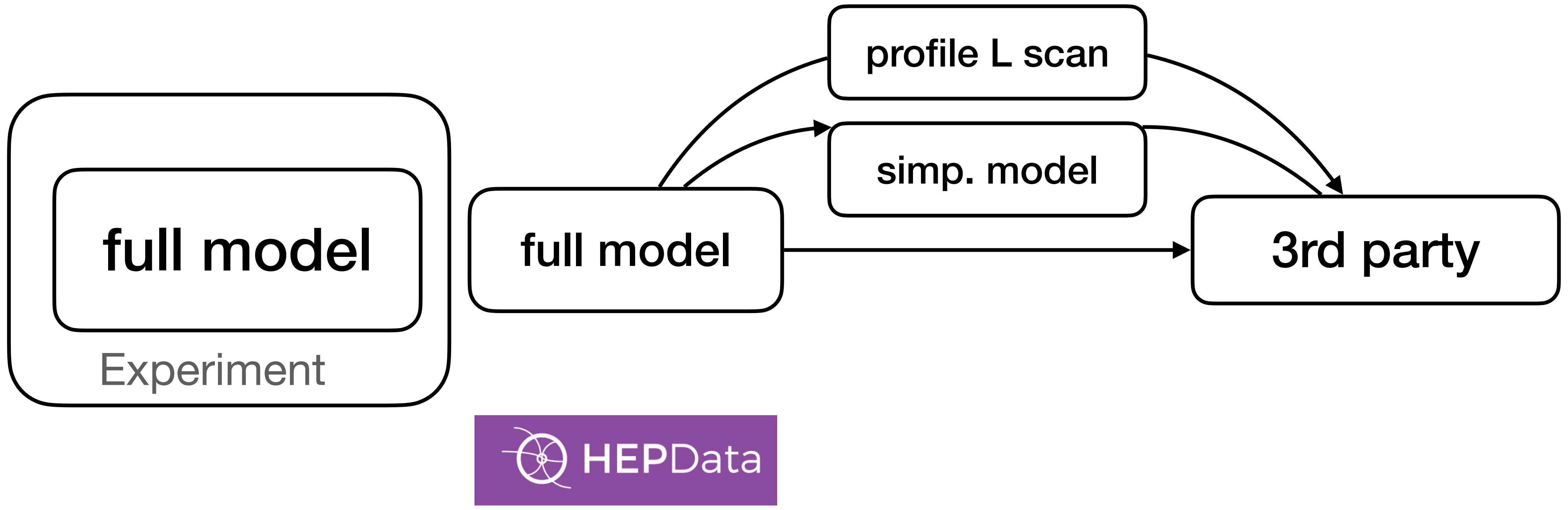
Solution to most of our problems is to take a leap and

publish the full statistical model as it is used internally

Lots of advantages:

- no loss of information: this is the likelihood that was used by original analyzers
- full set of systematics: this can be the basis of a **real combination**
- full structure apparent: can be a **tool for reinteretation**
- simplified approaches still possible but developed / carried out in public





How do we do it?

Clear advantages, but didn't happen for a long time.

Few challenges:

- Conceptual: finding a way to publish "any" model is hard
  easier to first solve slightly simpler/constrained problem space

- Format: need to find a software-agnostic format if possible
  text-based formats like JSON are a good idea

- Tooling: need to have toos & provide training for users beyond experiments
  this workshop!

Close World vs Open World

It's difficult to find a solution / data format to publish "any model"

- there are very few first-principles constraints as to what a valid $p(x \mid \theta)$
- in full generality boils down to essentially preserving abitrary programs

But if we're willing to constrain ourselves to subset of valid $p(x \mid \theta)$ it's easier

- "closed world" of finite number of building blocks from which complicated models can be built
- successful in LHC physics internally. Focus on this, solve general problem later

Closed World of HistFactory

Typical Binned Likelihood Model: HistFactory

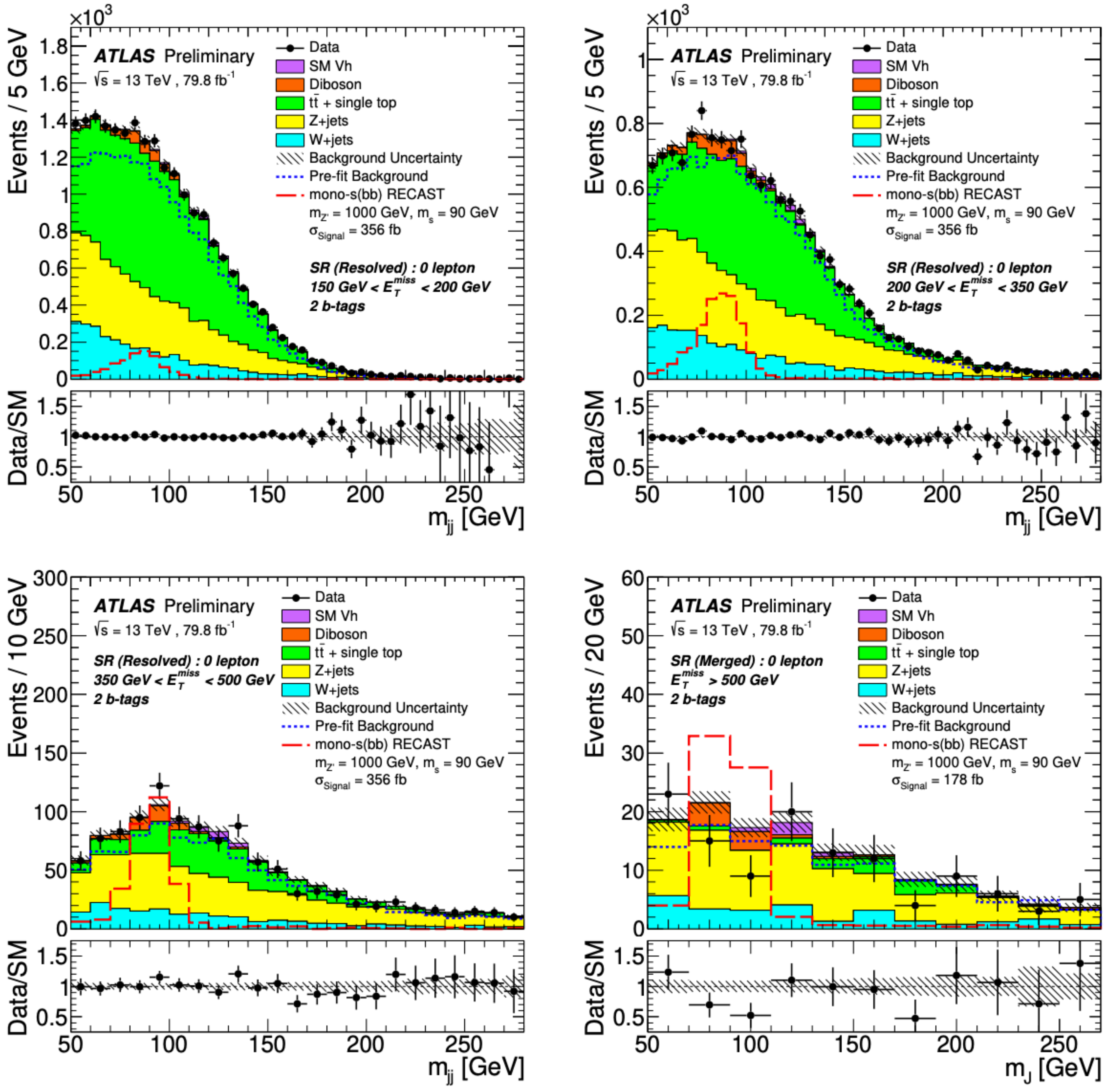
Models multile disjoint phase-space regions in a joint measrement

multiple regions

expected rates

$$p(x|\theta) = \prod_c \mathrm{Pois}(n_c(x)|\lambda_c(\theta))$$

binned data



Examle: 4 regions

Closed World of HistFactory

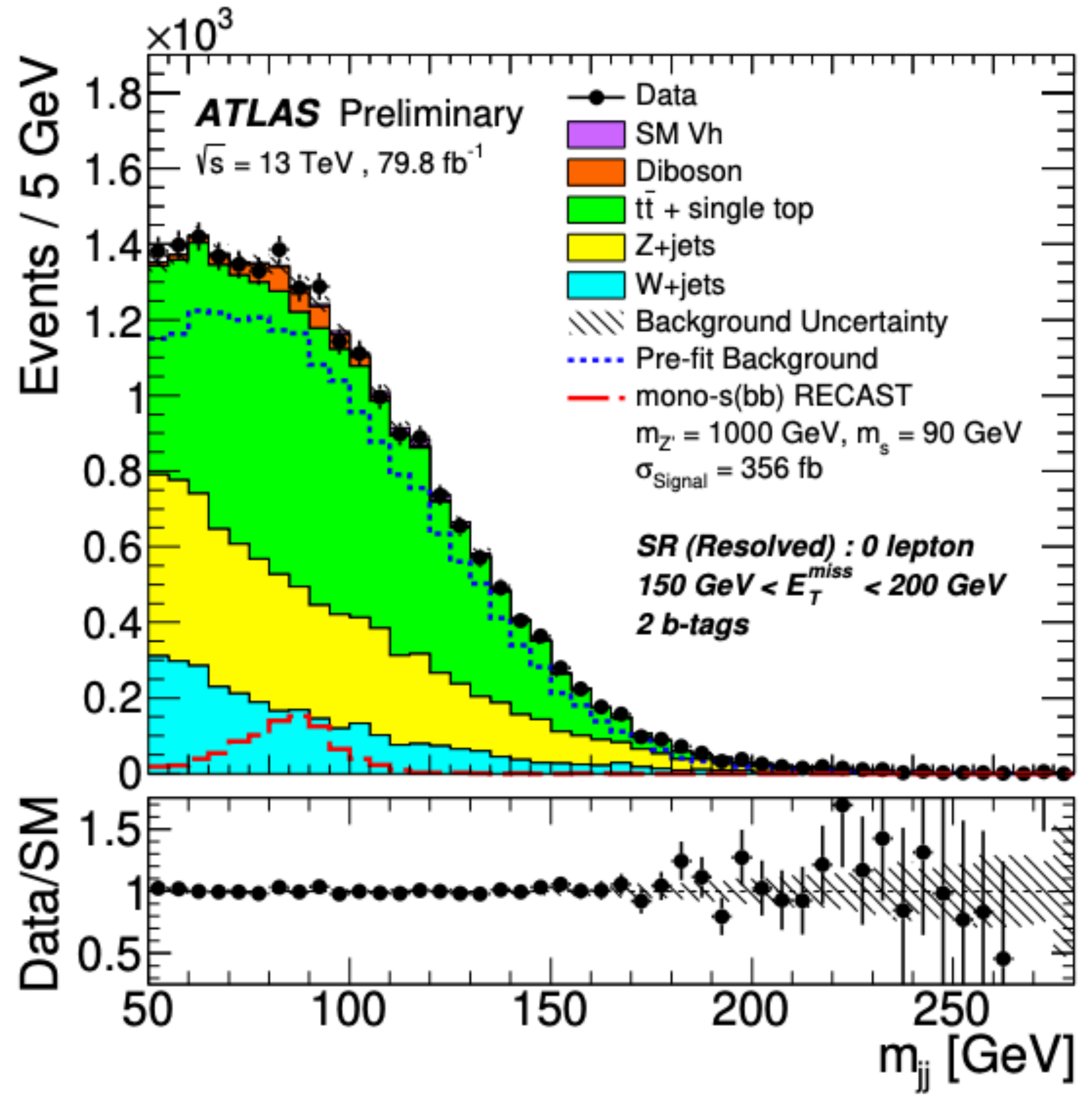
Typical Binned Likelihood Model: HistFactory

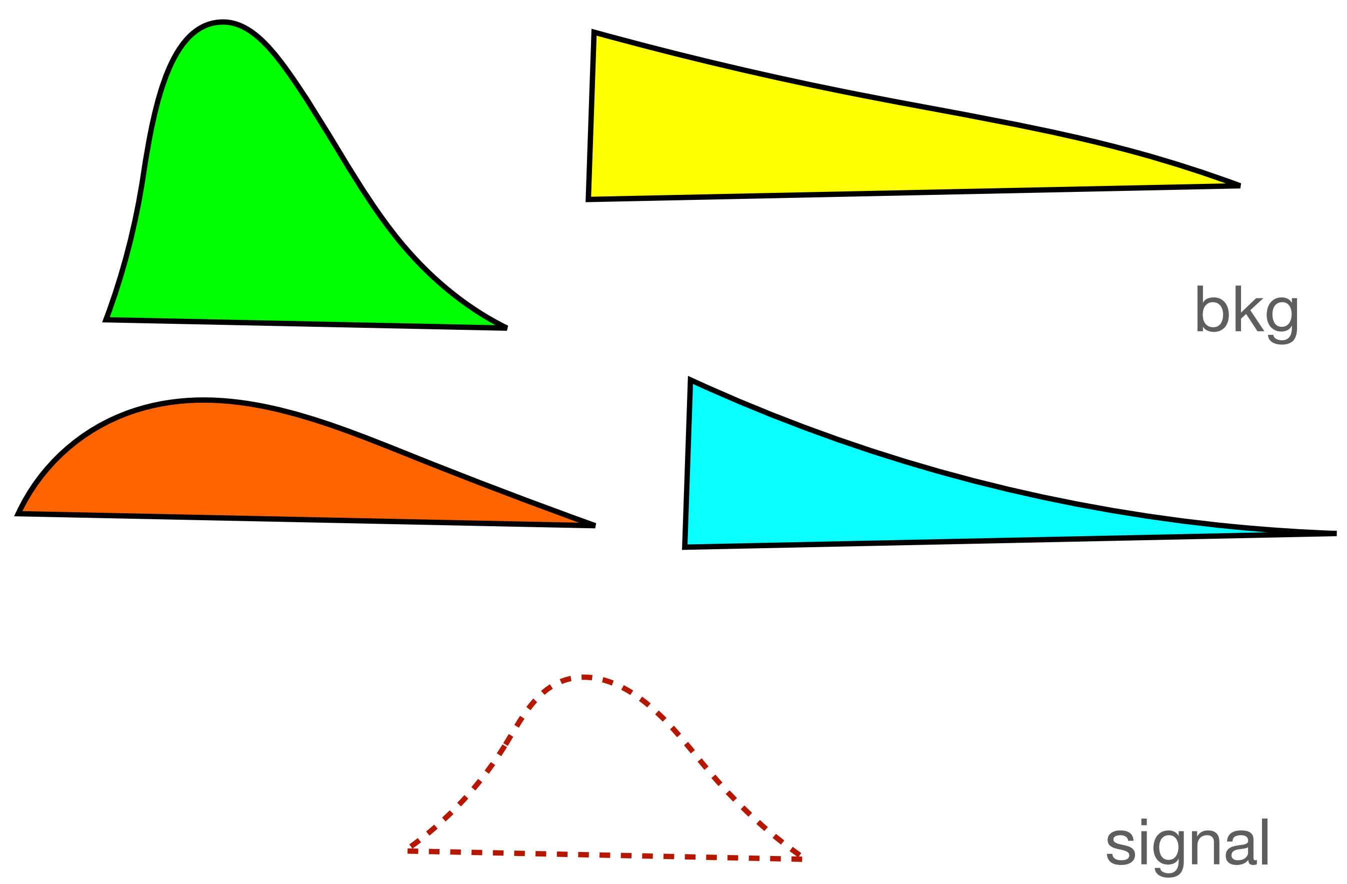
The expected rate of each region is the sum of multiple parametrized components

$$p(x|\theta) = \prod_c \mathrm{Pois}(n_c(x)|\lambda_c(\theta))$$

$$\mathrm{Pois}(n_c|\lambda(\theta)),\ \lambda(\theta) = \sum_{s\in\mathrm{samples}} \lambda_s(\theta)$$

uncertainties & normalization on components from a fixed set of "rate modifiers"





JSON Likelihoods

Models like HistFactory are easily described with "declarative languages"

- don't need full complexity of a program, just describe the building blocks

Original HistFactory: XML+ROOT-based

- good integration into remaining ROOT stats ecosystem

"pyhf" format: JSON (convertible to ROOT):

- this helps provide language-independent data products
- this is what ATLAS releases across groups (Top, SUSY, Exotics...)


```
{
    "channels": [{
        "name": "singlechannel",
        "samples": [{
                "name": "sig",
                "data": [12.0, 11.0],
                "modifiers": [{ "name": "mu", "data": null, "type": "normfactor" }]
            },
            {
                "name": "bkg",
                "data": [50.0, 52.0],
                "modifiers": [{ "name": "uncorr_bkguncrt", "data": [3.0, 7.0], "type": "shapesys" }]
            }
        ]
    }],
    "data": {
        "single
    },
    "toplvl": {
        "measu
            "c
            "na
        }]
    }
}
```


Published Statistical Models

Updating list of HEPData entries for publications using `HistFactory` JSON statistical models:

- Search for bottom-squark pair production with the ATLAS detector in final states containing Higgs bosons, b-jets and missing transverse mom
  2019. URL: https://doi.org/10.17182/hepdata.89408, doi:10.17182/hepdata.89408.
- Search for chargino-neutralino production with mass splittings near the electroweak scale in three-lepton final states in $\sqrt s$ = 13 TeV pp
  with the ATLAS detector. 2019. URL: https://doi.org/10.17182/hepdata.91127, doi:10.17182/hepdata.91127.
- Search for direct stau production in events with two hadronic τ-leptons in $\sqrt s = 13$ TeV pp collisions with the ATLAS detector. 2019. UR
  https://doi.org/10.17182/hepdata.92006, doi:10.17182/hepdata.92006.
- Search for direct production of electroweakinos in final states with one lepton, missing transverse momentum and a Higgs boson decaying into t
  jets in (pp) collisions at $\sqrt s=13$ TeV with the ATLAS detector. 2020. URL: https://doi.org/10.17182/hepdata.90607.v2,
  doi:10.17182/hepdata.90607.v2.
- Search for displaced leptons in $\sqrt s = 13$ TeV pp collisions with the ATLAS detector. 2020. URL: https://doi.org/10.17182/hepdata.9879
  doi:10.17182/hepdata.98796.
- Search for squarks and gluinos in final states with same-sign leptons and jets using 139 fb$^-1$ of data collected with the ATLAS detector. 202
  https://doi.org/10.17182/hepdata.91214.v3, doi:10.17182/hepdata.91214.v3.
- Search for trilepton resonances from chargino and neutralino pair production in $\sqrt s$ = 13 TeV pp collisions with the ATLAS detector. 202
  https://doi.org/10.17182/hepdata.99806, doi:10.17182/hepdata.99806.
- Search for squarks and gluinos in final states with jets and missing transverse momentum using 139 fb$^-1$ of $\sqrt s$ =13 TeV pp collision
  with the ATLAS detector. 2021. URL: https://doi.org/10.17182/hepdata.95664, doi:10.17182/hepdata.95664.

Beyond HistFactory

HistFactory is used in ATLAS, LHCb, BellIe-II, ...

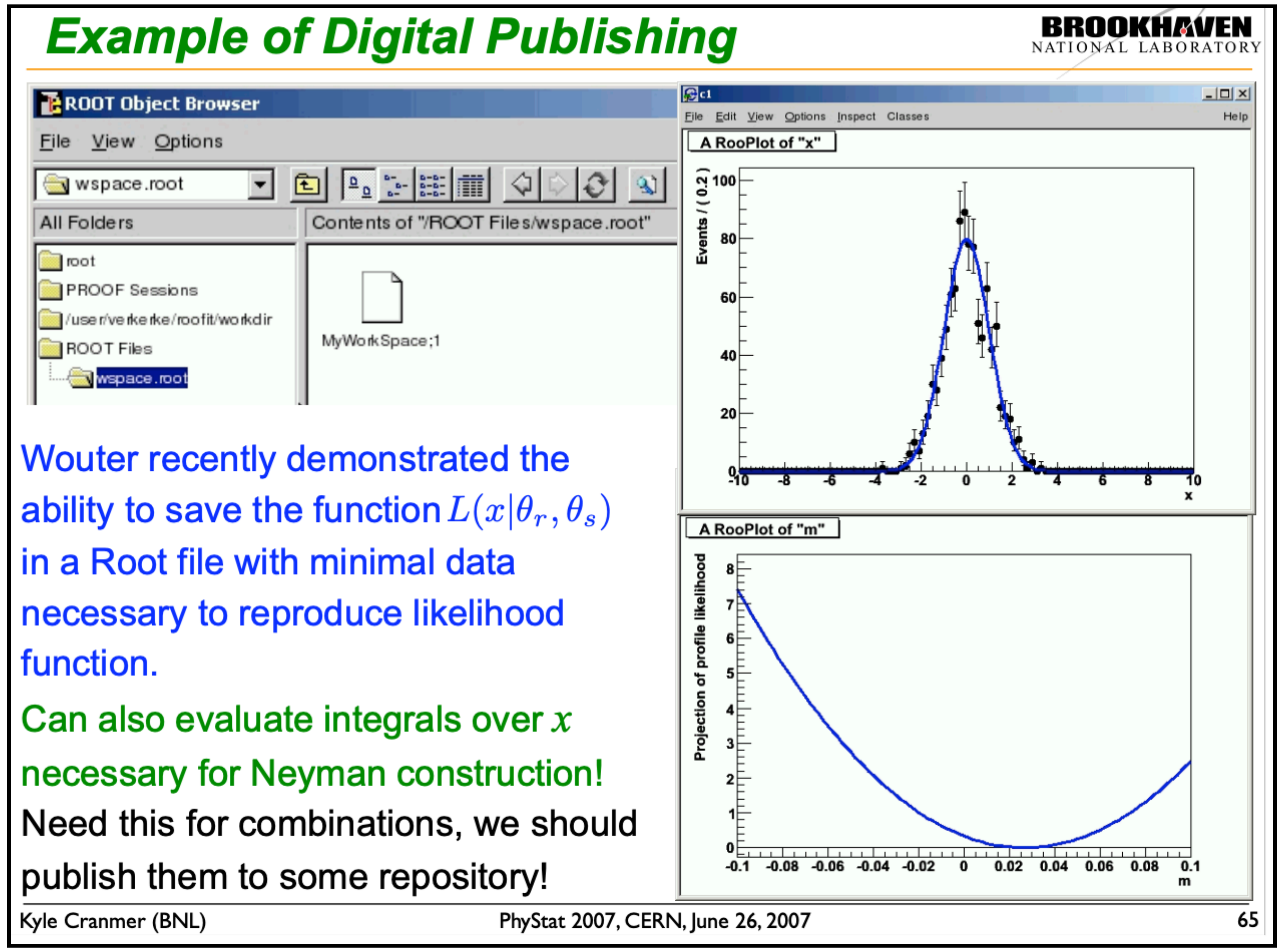
- CMS uses a similar model: "Combine" (originally from Higgs Group)
- would be nice to explore publishing those as well

JSON and Format have limits. ROOT is able to express a much wider set of probability models at the cost of being tied to ROOT.

- "ROOT Workspace" serialization since 2007 available
- basis of e.g. Higgs Combinations

Maybe now that we've escaped the "potential well" for publishing likelihoods we can revisit publishing the ROOT Workspaces

- a different tradeoff but very valuable



Tooling

LHC Experiments are now releasing unprecedentent level of detail`

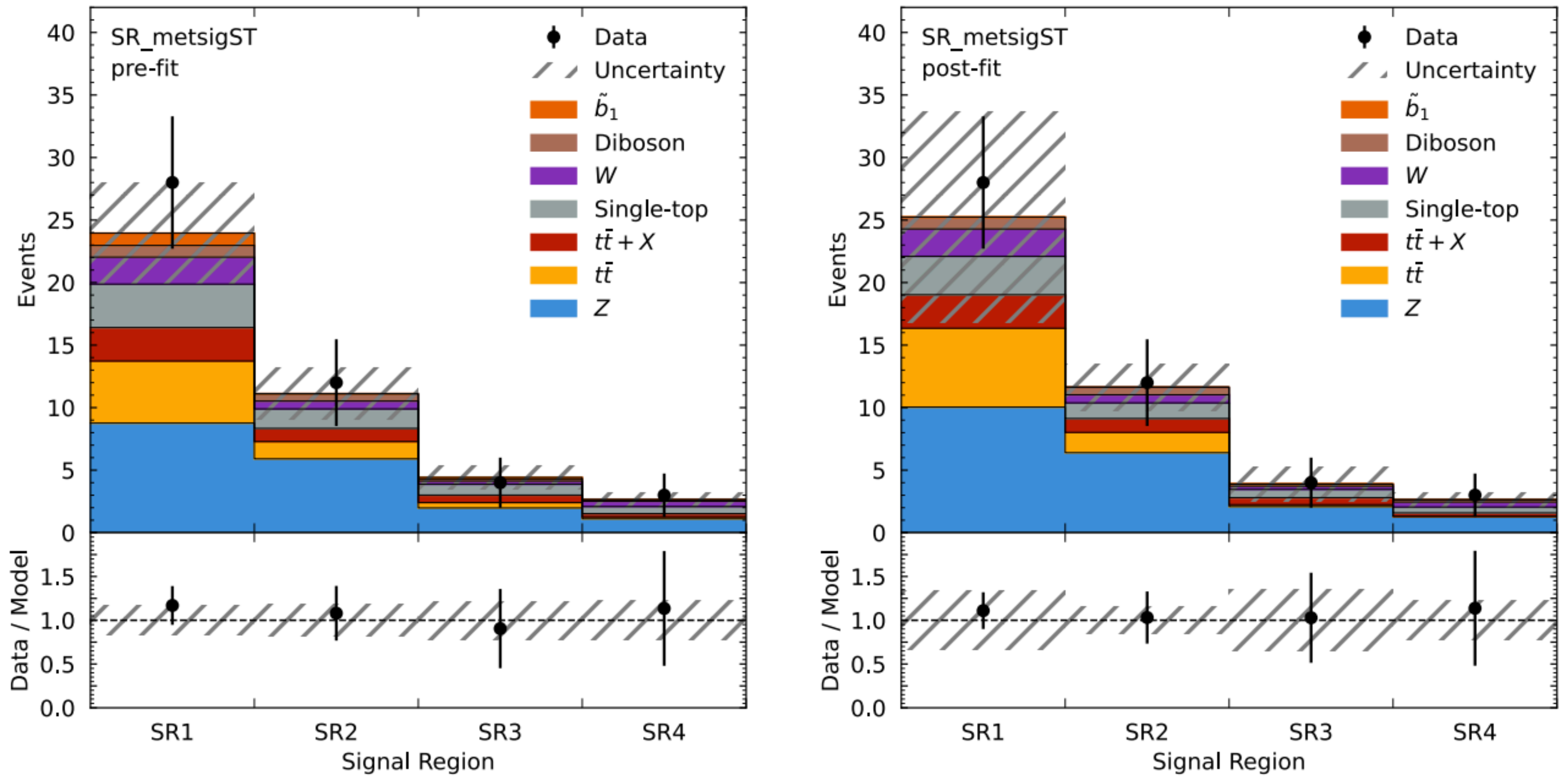
- requires good communication about conventions and terms
  - Example: "how do you define a expected limit"

- tools that both theorists and experimentalists share
  - should easily be integratable in their software
  - language-agnostic & easy python API of pyhf probably helped adoption

```
import json
import cabinetry
import pyhf
from cabinetry.model_utils import prediction
from pyhf.contrib.utils import download

# download the ATLAS bottom-squarks analysis probability models from HEPData
download("https://www.hepdata.net/record/resource/193543?...

# construct a workspace from a background-only model and
bkg_only_workspace = pyhf.Workspace(json.load(open("botto
patchset = pyhf.PatchSet(json.load(open("bottom-squarks/R
workspace = patchset.apply(bkg_only_workspace, "sbottom_6

# construct the probability model and observations
model, data = cabinetry.model_utils.model_and_data(worksp

# produce visualizations of the pre-fit model and observe
prefit_model = prediction(model)
cabinetry.visualize.data_mc(prefit_model, data)

# fit the model to the observed data
fit_results = cabinetry.fit.fit(model, data)

# produce visualizations of the post-fit model and observ
postfit_model = prediction(model, fit_results=fit_results
cabinetry.visualize.data_mc(postfit_model, data)
```

cabinetry

pyhf ∂ifferentiable ℒikelihoods



A SModelS interface for pyhf likelihoods

Gaël Alguero[a], Sabine Kraml[a], Wolfgang Waltenberger[b,c]

[a] *Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 Avenue des Martyrs, F-38026 Grenoble, France*
[b] *Institut für Hochenergiephysik, Österreichische Akademie der Wissenschaften, Nikolsdorfer Gasse 18, 1050 Wien, Austria*
[c] *University of Vienna, Faculty of Physics, Boltzmanngasse 5, A-1090 Wien, Austria*

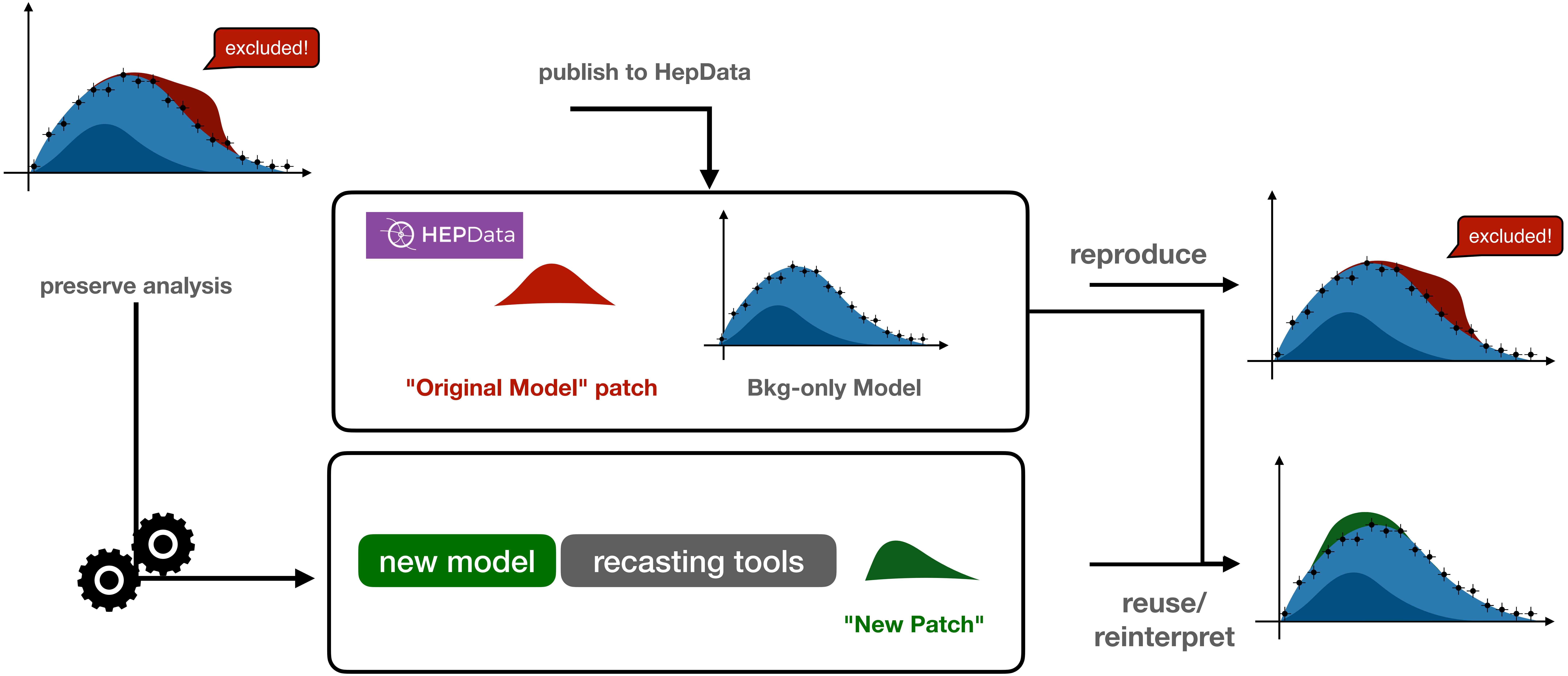
Abstract

SModelS is an automatized tool enabling the fast interpretation of simplified model results from the LHC within any model of new physics respecting a $\mathbb{Z}_2$ symmetry. We here present a new version of SModelS, which can use the full likelihoods now provided by ATLAS in the form of pyhf JSON files. This much improves the statistical evaluation and therefore also the limit setting on new physics scenarios.

For Reinterpretation: "Patching" as a first class operation

Byproduct of JSON tooling: manipulating the models easy to reason about

- **Patching: remove some parts, add new parts to the model**
- **natural connection to reinteretation. "recasting = producing new patches"**



Summary

Likelihood publishing has been a long-time coming

- **many intermediate achievements paving way for full model release**
- **now it's happening & we can try to build a community around this new practice**
- **best argument to experiments to continue publishing likelihoods is to demonstrate that you use & cite them**

Have a great workshop!