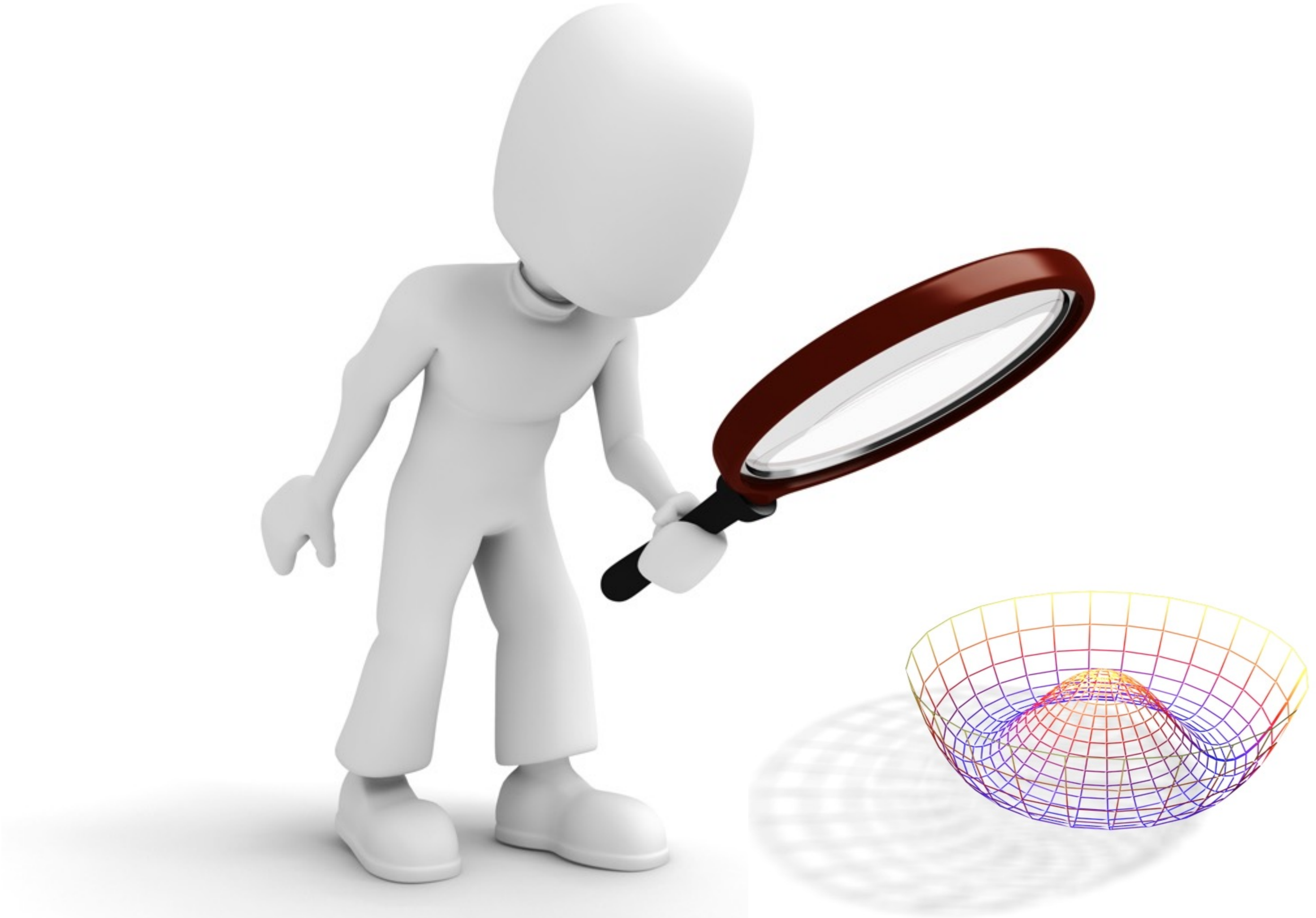
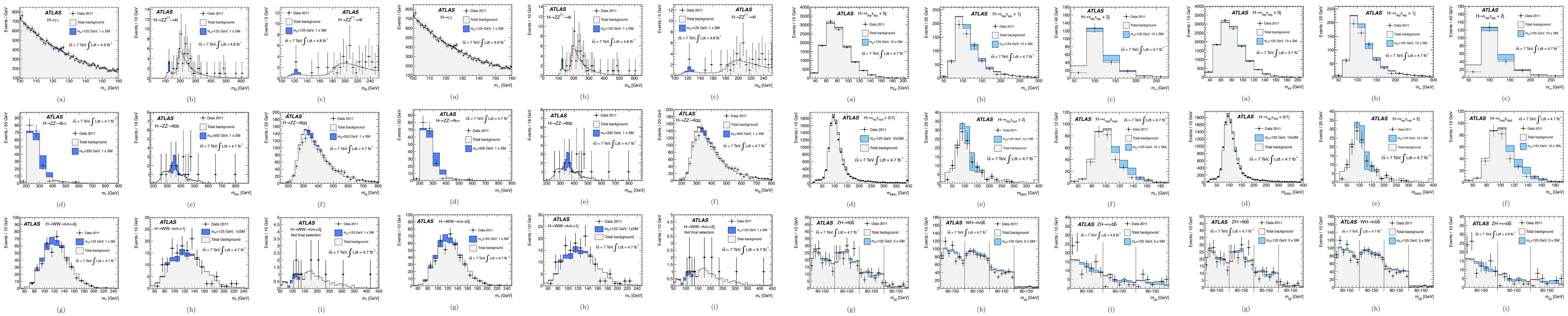
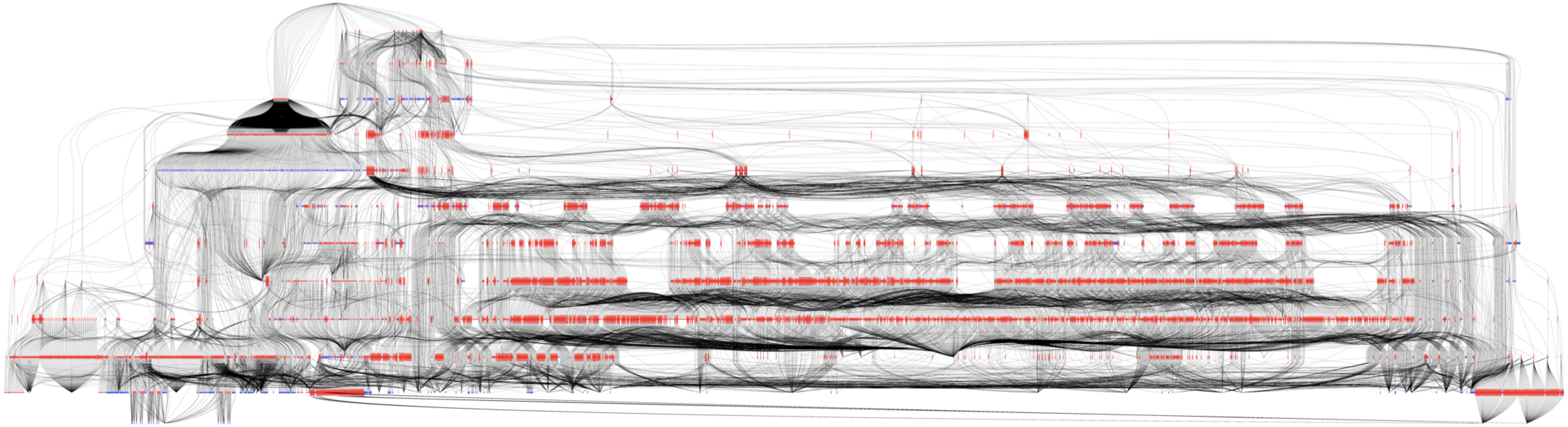
GRAND COMBINATIONS @ HL-LHC & FAIROS-HEP



@KyleCranmer
University of Wisconsin-Madison
Data Science Institute
Physics, Computer Science, Statistics

Combined fits for the Higgs discovery





$$\mathbf{f}_{\rm tot}(\mathcal{D}_{\rm sim},\mathcal{G}|\boldsymbol{\alpha}) = \prod_{c\in\text{channels}}\left[\text{Pois}(n_c|\nu_c(\boldsymbol{\alpha}))\prod_{e=1}^{n_c} f_c(x_{ce}|\boldsymbol{\alpha})\right]\cdot\prod_{p\in\mathbb{S}} f_p(a_p|\alpha_p)$$

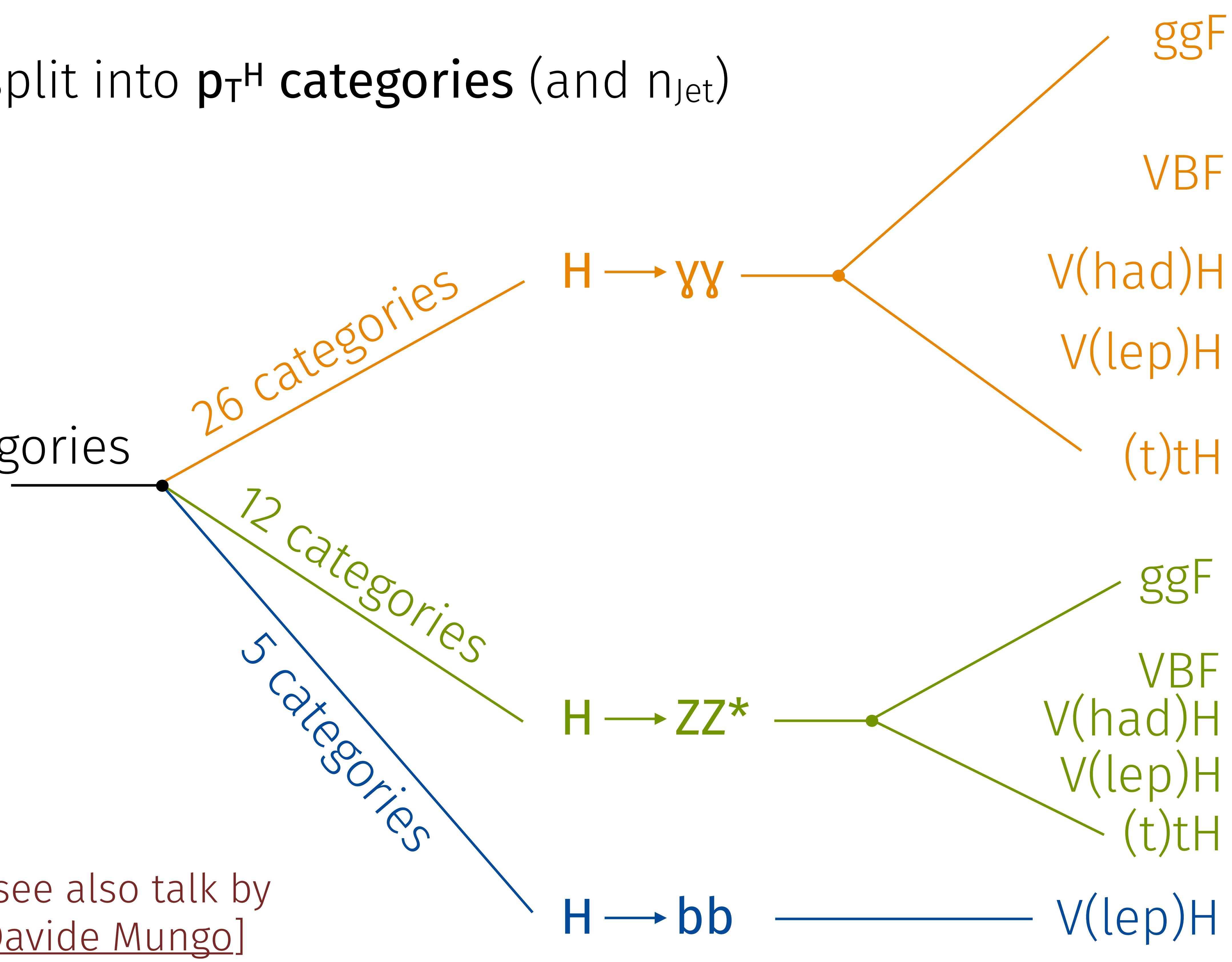
Combined fits for EFTs

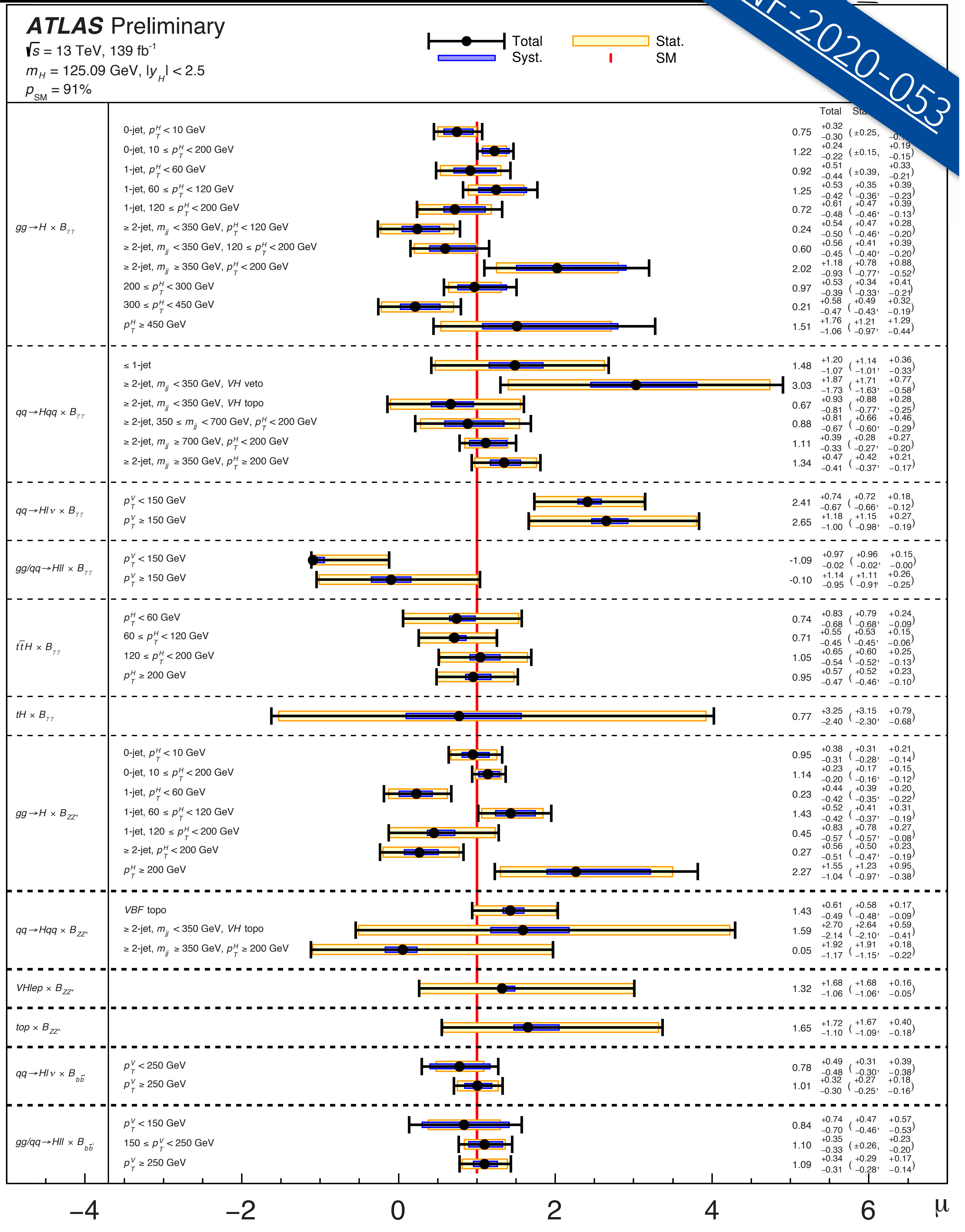
The STXS combination measurement

ATLAS-CONF-2020-053

Aim: EFT interpretation of the 139 fb^{-1} combination of $H \to ZZ^* \to 4\ell$, $H \to \gamma\gamma$ and $H \to bb$ merged stage-1.2 STXS measurement

Mostly split into **p_T^H categories** (and n_{Jet})





[see also talk by Davide Mungo]

Comments on Unfolding

- Fiducial and differential cross section measurements
  - → minimise model dependence
  - → relatively restricted sensitivity (hard to combine different channels)
  - → re-interpretable outside experiment

Saskia Falke, Higgs2020

- Desirable to have results at particle level, and distributions (STXS or fiducial distr.)

Unfolding is deceptively attractive

- It seems very convenient and to address what we want to know, but
- unfolding is a can of worms statistically and pushes many problems down stream
  - Combinations, correlated systematics, artifacts and bias introduced by the unfolding procedure. These will all turn into systematics in the final results.
- It is good for fast approximate answers, but
- I do not recommend it as a platform for the final "gold standard" results.

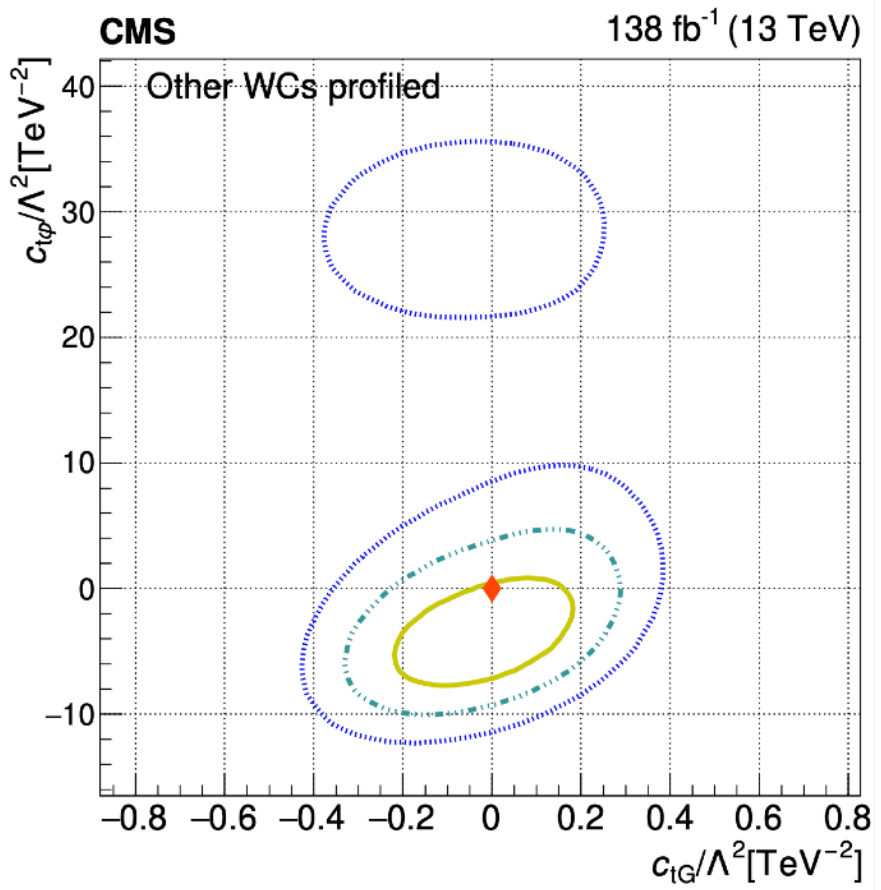
Likelihood scans vs. full statistical models

Combining detector-level analyses

- EFT analyses cannot always be summarised in a covariance matrix
  - Most likely possible if only linear terms are taken into account
- Combinations **must be done by at the likelihood level**

$$\mathcal{L} = \prod \mathcal{P}\left(x_i | n_{\rm bkg}(\theta) + n_{\rm Signal}(C_i, \theta)\right) P(\theta)$$

  - All necessary **changes must be implemented at detector level**



- In the near future, detector-level analyses will be available for combination also outside the collaborations
  - **CMS is planning to release the analysis likelihoods** along with instructions to evaluate them
  - Ideas to provide profiled likelihood ratio parametrised by a neural network (see CHEP 2023 contribution)

Sergio Sánchez Cruz 5

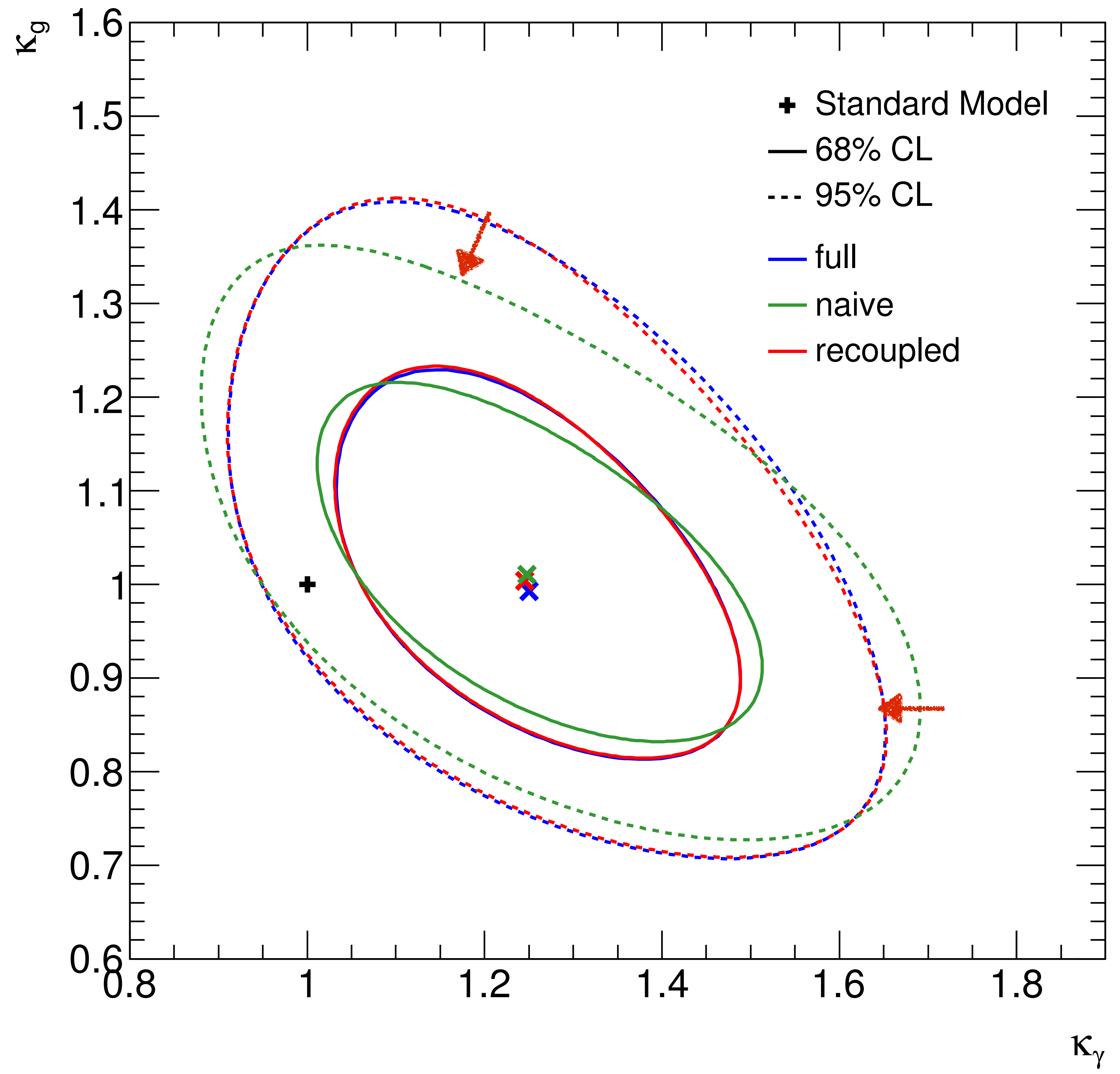
Warning: the term "likelihood" is used to describe both full statistical model and the likelihood function for observed data.

Sharing profile likelihood isn't a good approach to combinations, you **double count** constraint terms and have **inconsistent profiling**.

- Studied extensively in

https://inspirehep.net/literature/1275827

See also section 4.2.3 in arxiv:2109.04981



My message

Top Level Message: We should publish the full statistical model (aka “likelihood”) for measurements that constrain EFT coefficients

- Lots of progress in publishing statistical models recently in BSM searches

Second Level Message: There are a few ways to describe the dependence on EFT parameters. We can and should separate the specification and implementation.

- First define a **specification** for one or more of these choices that removes all ambiguity. This allows multiple groups to **implement** the specification.

Third Level Message: In addition to publishing statistical models, RECAST-like infrastructure would allow us to consider new EFT operators and update / improve background modeling after publishing

- This infrastructure is being used in BSM searches already

Likelihood Publishing + RECAST =

Message 1:
Publishing Statistical Models

The first PhyStat

It was 24 years ago!

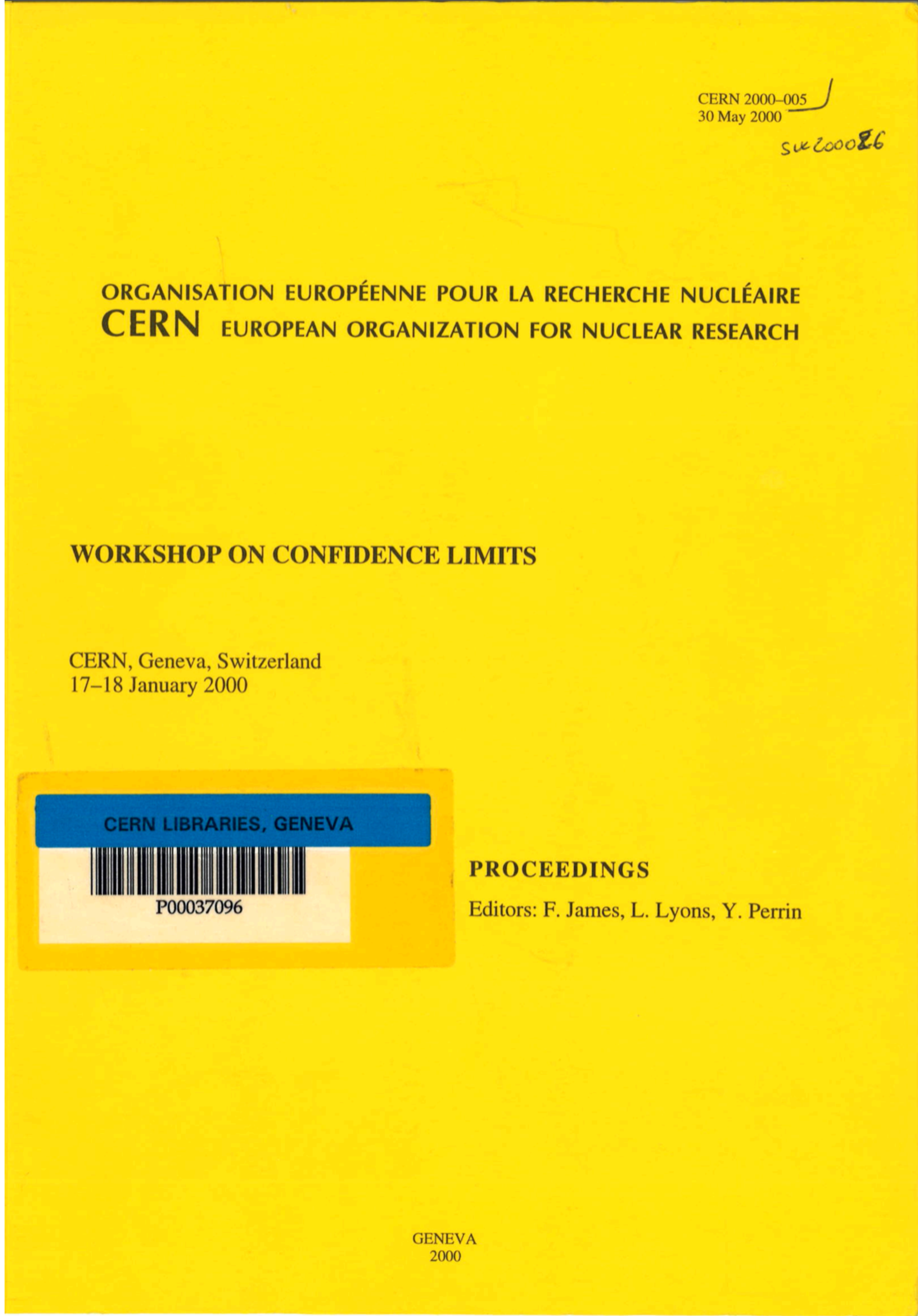
Massimo Corradi

It seems to me that there is a general consensus that what is really meaningful for an experiment is *likelihood*, and almost everybody would agree on the prescription that experiments should give their likelihood function for these kinds of results. Does everybody agree on this statement, to publish likelihoods?

Louis Lyons

Any disagreement ? Carried unanimously. That's actually quite an achievement for this Workshop.

https://cds.cern.ch/record/411537?ln=en


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

PROCEEDINGS

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

GENEVA
2000

SciPost Physics [arxiv:2109.04981] Submission

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]

Concepts

While it may seem overly technical, these subtle distinctions are very important.

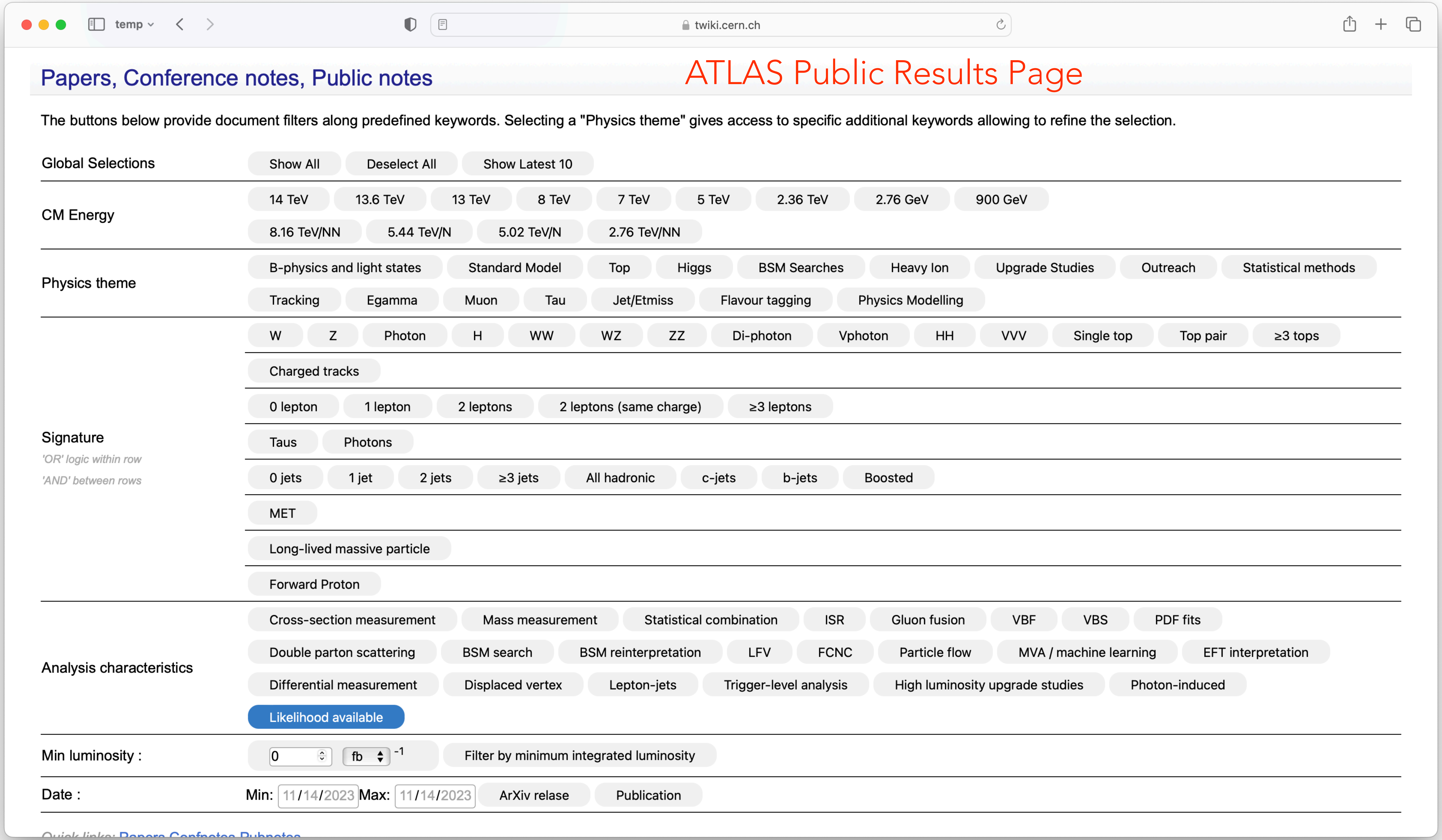
We overcame decades of stagnation when we focused on **declarative specification** for **closed-world** models and moved to standard approaches to **serialization** (e.g. ROOT binary to JSON/ yaml)

- breakthrough with pyhf

Glossary of terms

- **Statistical model**: This is a synonym for the probability model $p(x, y|\mu, \theta)$ as in Eq. (7) that includes dependence on the data x and y, the parameters of interest μ and nuisance parameters θ, access to the individual terms and the ability to generate pseudo- (or synthetic-) data (i.e., "toy Monte Carlo").
- **Likelihood**: The value of the statistical model for a given *fixed* dataset as a function of the parameters, e.g., $L(\mu, \theta)$ in Eq. (7).
- **Constraint term**: A term in the full statistical model that relates auxiliary data y to a particular nuisance parameter θ.
- **Observed data** the n, x, and y of Eq. (7) needed to construct the likelihood.
- **Open-world**: An approach to statistical modelling that allows users to define and implement custom components in the statistical model.
- **Closed-world**: An approach to statistical modelling that requires users to work with a finite set of modelling components.
- **Declarative specification**: An unambiguous specification (e.g., of a statistical model) that is independent of implementation. Often there exists a reference implementation of a specification, but in the declarative approach there may be multiple implementations that are conceptually and mathematically equivalent.
- **Serialization**: The process of writing a data structure (e.g., a statistical model) in memory to a file in a way that can be read back into memory. Loading the serialized object typically requires access to compatible software libraries present at the time of serialization.

It's a reality



It's a reality

https://www.hepdata.net/search/?q=analysis:HistFactory

Find all papers which include specific types of **analysis**.

- analysis:rivet (Rivet analysis)
- analysis:MadAnalysis (MadAnalysis 5 analysis)
- analysis:HistFactory (likelihoods in HistFactory format)

HEPData — About · Submission Help · File Formats · Sign in

analysis:HistFactory | Search | Reset search | Advanced | JSON

Max results · Sort by · Reverse order — Showing **10** of **28** results

Date: 2019 … 2023

Collaboration: ATLAS 28

Subject_areas: hep-ex 28

Phrases:
- Proton-Proton Scattering 3
- Cross Section 2
- SUSY 2
- Supersymmetry 2
- Top 2

Next 5 · Show All

Reactions:
- P P --> CHARGINO+ CHARGINO- 1
- P P --> CHARGINO+ NEUTRALINO 1
- P P --> CHARGINO+- NEUTRALINO 1

« ‹ 1 2 › »

HistFactory HistFactory HistFactory HistFactory **Search for flavour-changing neutral-current couplings between the top quark and the photon with the ATLAS detector at $\sqrt{s} = 13$ TeV**

The ATLAS collaboration Aad, Georges ; Abbott, Braden Keim ; Abbott, Dale ; *et al.*

Phys.Lett.B 842 (2023) 137379, 2023.

Inspire Record 2077557 · DOI 10.17182/hepdata.129959

This letter documents a search for flavour-changing neutral currents (FCNCs), which are strongly suppressed in the Standard Model, in events with a photon and a top quark with the ATLAS detector. The analysis uses data collected in pp collisions at $\sqrt{s} = 13$ TeV during Run 2 of the LHC, corresponding to an integrated luminosity of 139 fb^{-1}. Both FCNC top-quark production and decay are considered. The final state consists of a charged lepton, missing transverse momentum, a b-tagged jet, on…

0 data tables match query

HistFactory HistFactory **Measurement of the $t\bar{t}t\bar{t}$ production cross section in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector**

The ATLAS collaboration Aad, Georges ; Abbott, Braden Keim ; Abbott, Dale ; *et al.*

JHEP 11 (2021) 118, 2021.

Inspire Record 1869695 · DOI 10.17182/hepdata.105039

A measurement of four-top-quark production using proton-proton collision data at a centre-of-mass energy of 13 TeV collected by the ATLAS detector at the Large Hadron Collider corresponding to an integrated luminosity of 139 fb^{-1} is presented. Events are selected if they contain a single lepton (electron or muon) or an opposite-sign lepton pair, in association with multiple jets. The events are categorised according to the number of jets and how likely these are to contain b-hadrons. A…

0 data tables match query

HistFactory HistFactory **Observation of single-top-quark production in association with a photon using the ATLAS detector**

Clemens Lange @clelange · Apr 16

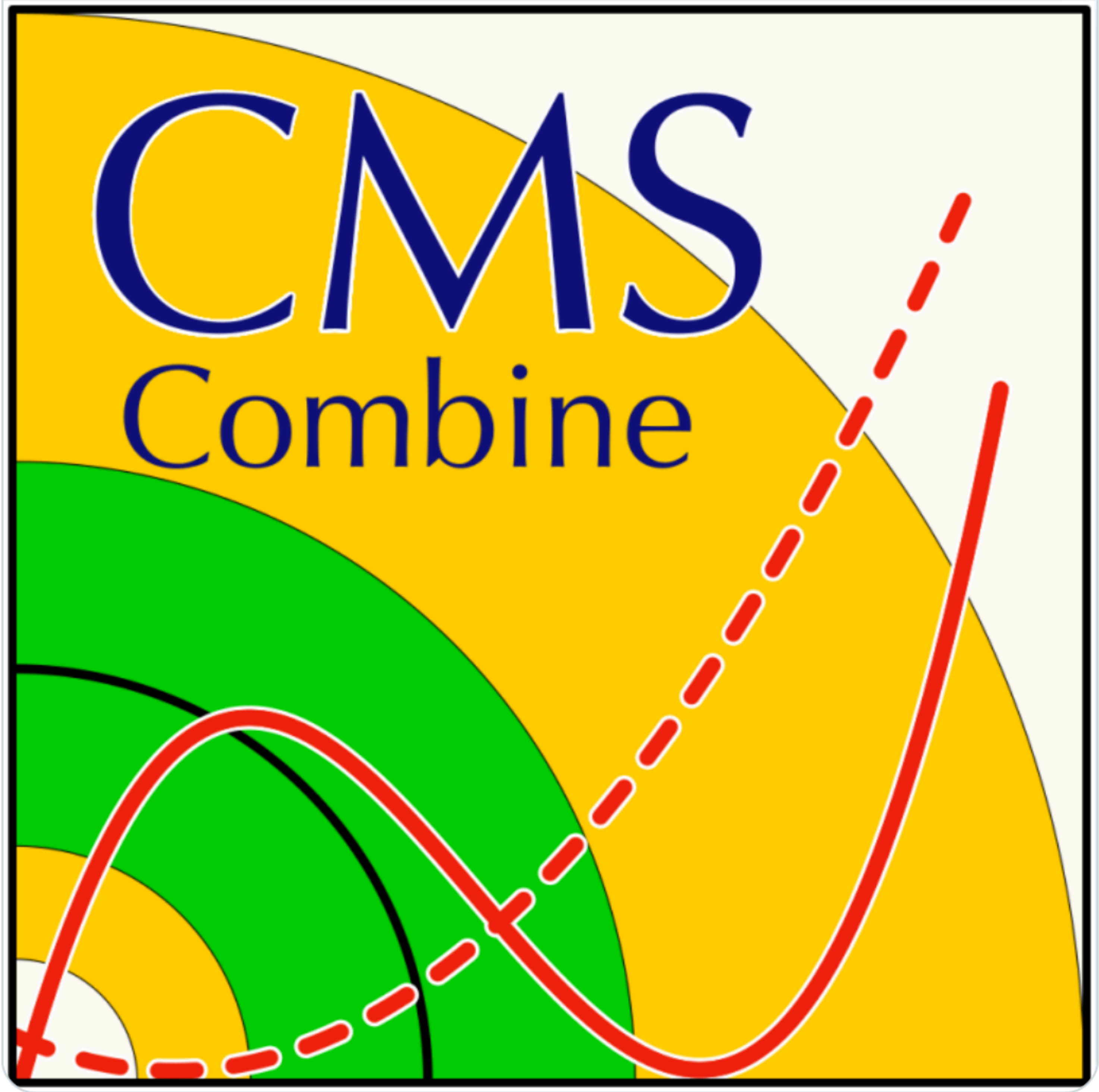
Finally out: the first full statistical model of a **CMS** physics analysis (the 2012 #HiggsBoson discovery), together with one of the most important pieces of software used and developed in the collaboration, the "**Combine**" tool. Correct link to the briefing: cms.cern/news/cms-commi...

CMS Experiment CERN @CMSExperiment · Apr 16

More than #OpenData, CMS continues to move towards #OpenScience

The statistical analysis tool used to find the #Higgs boson, along with the full statistical model and the data used to make the ...

Show more



Using published likelihoods

Just a few lines of code to download the statistical model, re-run fit, make diagnostic plots

```
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/1935437?view=true", "bottom-squarks")

# construct a workspace from a background-only model and a signal hypothesis
bkg_only_workspace = pyhf.Workspace(json.load(open("bottom-squarks/RegionC/BkgOnly.json")))
patchset = pyhf.PatchSet(json.load(open("bottom-squarks/RegionC/patchset.json")))
workspace = patchset.apply(bkg_only_workspace, "sbottom_600_280_150")

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

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

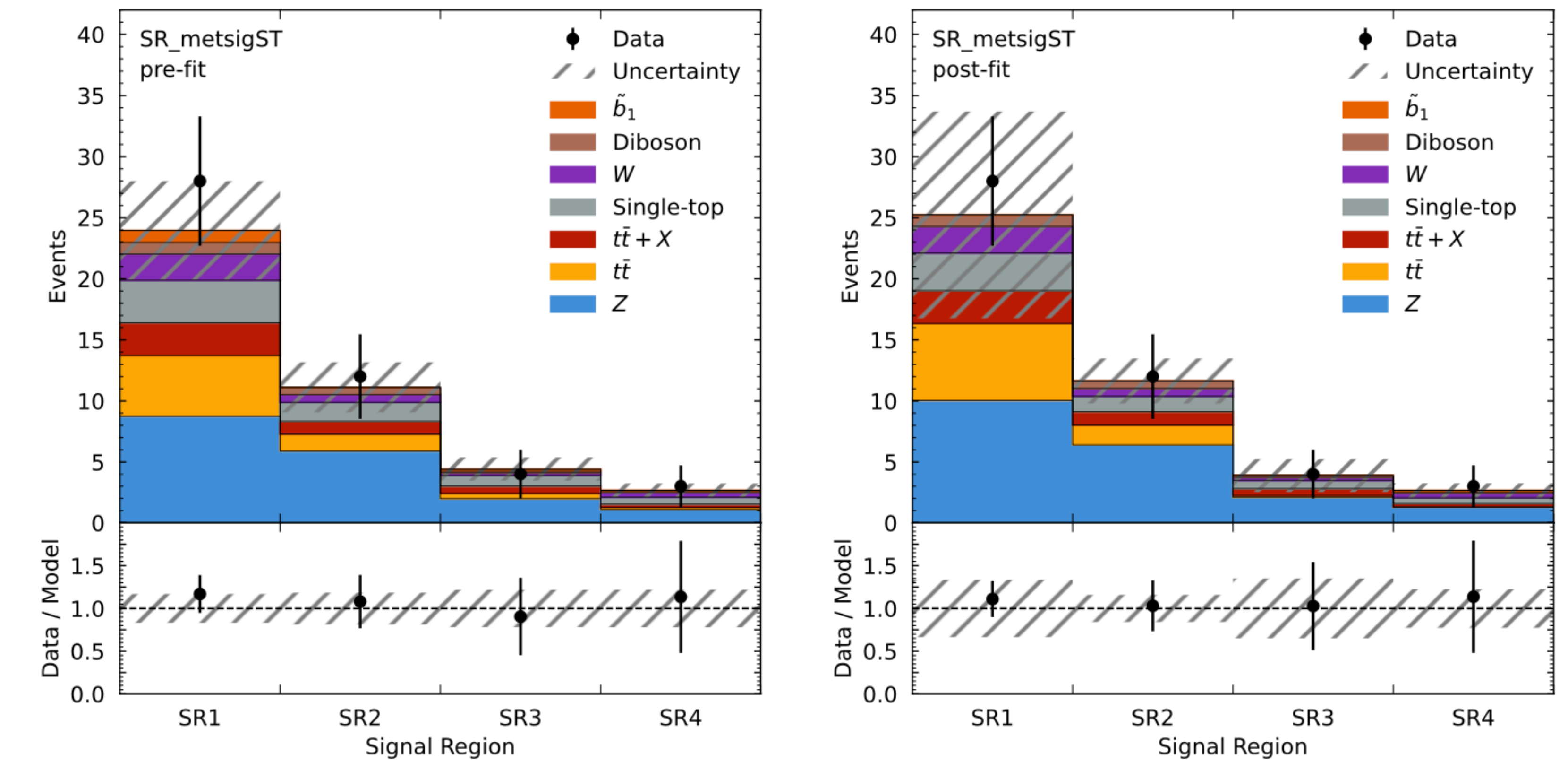
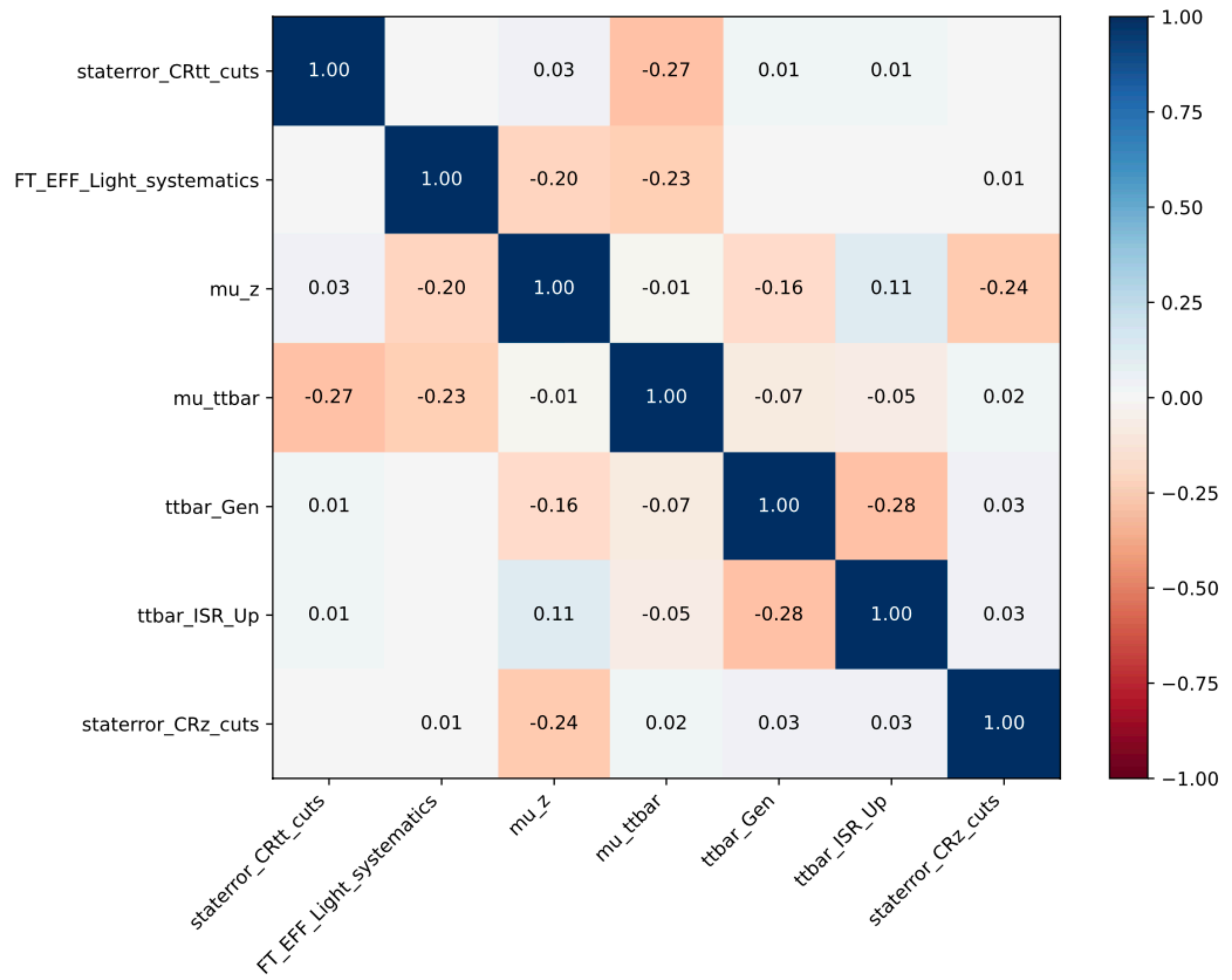


Figure 3: Pre-fit (left) and post-fit (right) visualizations of a selected signal hypothesis for four signal regions of the ATLAS search [41] of a bottom-squark of mass 600 GeV with a second-lightest neutralino of mass 280 GeV and lightest supersymmetric particle of mass 150 GeV generated from the full statistical models published in Ref. [20] using code from Ref. [40].



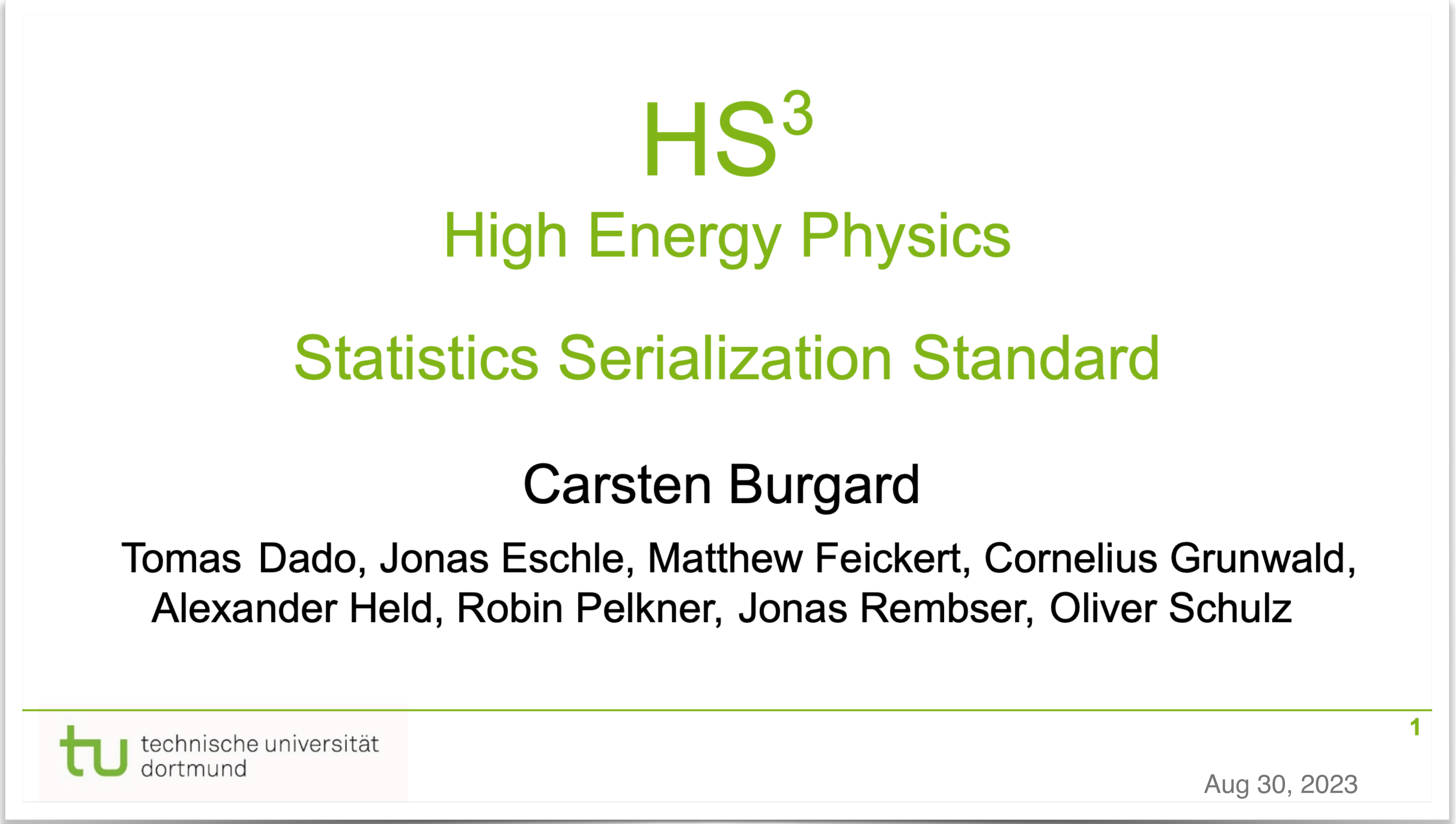
Browse and interact with published statistical models

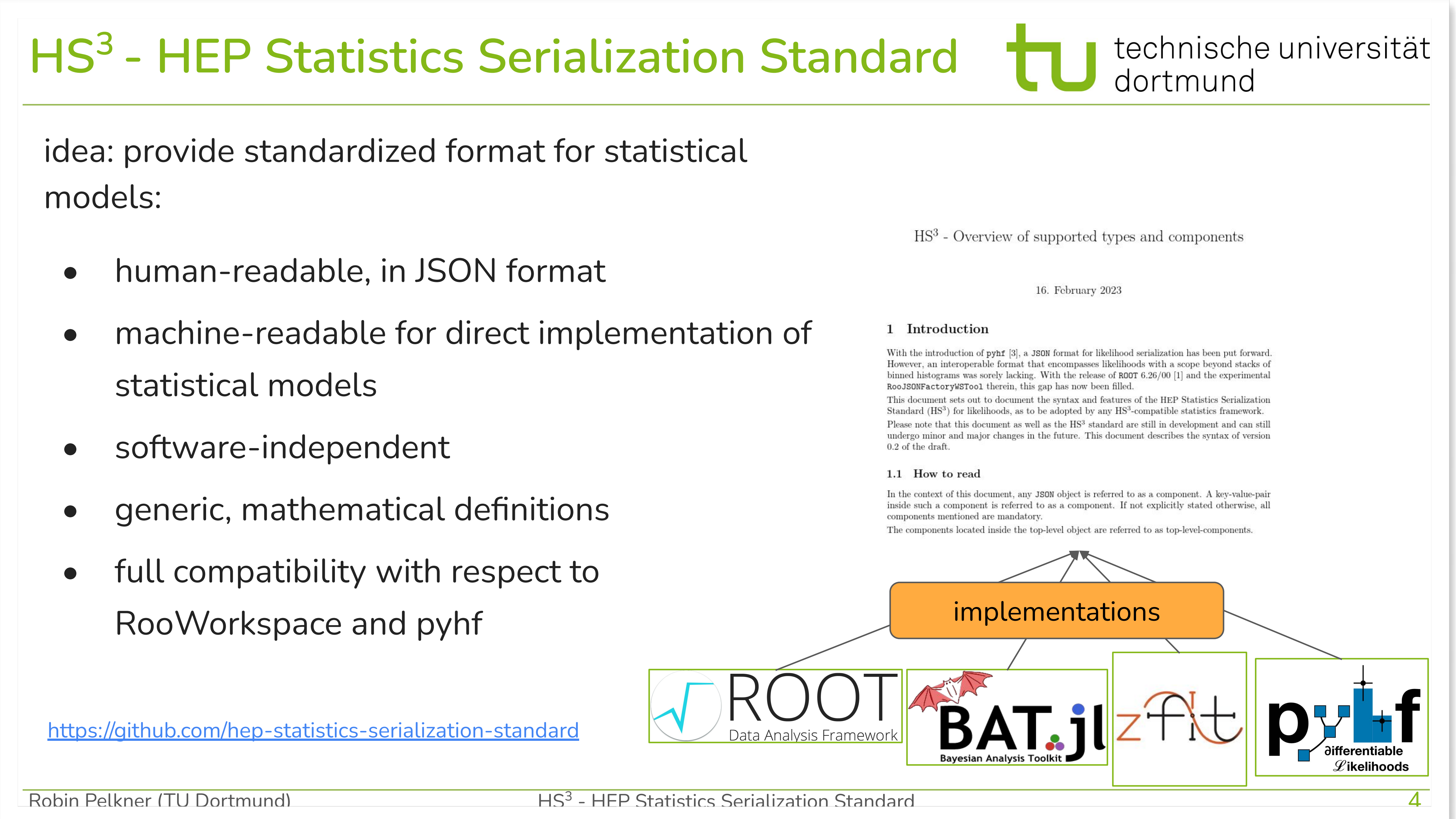
http://hepexplorer.net. Built by FAIROS-HEP

The HS3 Effort

There is now an effort to create a common serialization standard for pyhf, RooFit, BAT, zfit, etc. models

- Key idea: separate **specification** from **implementation**





Talk at Reinterpretation Forum [link]
https://indico.cern.ch/event/1264371/contributions/5338176/
https://videos.cern.ch/record/2296062
https://github.com/hep-statistics-serialization-standard

Message 2:
EFT-Specific Model Specification

The HistFactory **specification**

Also applies to Combine's binned-templates

The HistFactory **specification** is pure math with two main implementations (original C++ version in ROOT/RooFit and newer python version pyhf)

- Widely used and has *almost* everything needed for EFT

HistFactory Template: at a glance

$$f\,(\text{data}|\text{parameters}) = f\,(\vec{n}, \vec{a}|\vec{\eta}, \vec{\chi}) = \prod_{c\,\in\,\text{channels}} \prod_{b\,\in\,\text{bins}_c} \text{Pois}\,(n_{cb}|\nu_{cb}\,(\vec{\eta}, \vec{\chi})) \prod_{\chi\,\in\,\vec{\chi}} c_\chi\,(a_\chi|\chi)$$

$\vec{n}$: events, $\vec{a}$: auxiliary data, $\vec{\eta}$: unconstrained pars, $\vec{\chi}$: constrained pars

$$\nu_{cb}(\vec{\eta}, \vec{\chi}) = \sum_{s\,\in\,\text{samples}} \underbrace{\left(\sum_{\kappa\,\in\,\vec{\kappa}} \kappa_{scb}(\vec{\eta}, \vec{\chi})\right)}_{\text{multiplicative}} \left(\nu^0_{scb}(\vec{\eta}, \vec{\chi}) + \underbrace{\sum_{\Delta\,\in\,\vec{\Delta}} \Delta_{scb}(\vec{\eta}, \vec{\chi})}_{\text{additive}}\right)$$

Use: Multiple disjoint channels (or regions) of binned distributions with multiple samples contributing to each with additional (possibly shared) systematics between sample estimates

Main pieces:

- Main Poisson p.d.f. for simultaneous measurement of multiple channels
- Event rates $\nu_{cb}(\vec{\eta}, \vec{\chi})$ (nominal rate ν^0_{scb} with rate modifiers)
  - encode systematic uncertainties (e.g. normalization, shape)
- Constraint p.d.f. (+ data) for "auxiliary measurements"

4

The HistFactory **specification**

Also applies to Combine's binned-templates

… but the HistFactory **specification** is not natural for describing interference effects encountered in EFTs.

- We can create / **extend the specification** to handle EFT parameter dependence

HistFactory Template: at a glance

$$f(\text{data}|\text{parameters}) = f(\vec{n}, \vec{a}|\vec{\eta}, \vec{\chi}) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}_c} \text{Pois}(n_{cb}|\nu_{cb}(\vec{\eta}, \vec{\chi})) \prod_{\chi \in \vec{\chi}} c_\chi(a_\chi|\chi)$$

$\vec{n}$: events, $\vec{a}$: auxiliary data, $\vec{\eta}$: unconstrained pars, $\vec{\chi}$: constrained pars

$$\nu_{cb}(\vec{\eta}, \vec{\chi}) = \sum_{s \in \text{samples}} \underbrace{\left(\sum_{\kappa \in \vec{\kappa}} \kappa_{scb}(\vec{\eta}, \vec{\chi})\right)}_{\text{multiplicative}} \left(\nu^0_{scb}(\vec{\eta}, \vec{\chi}) + \underbrace{\sum_{\Delta \in \vec{\Delta}} \Delta_{scb}(\vec{\eta}, \vec{\chi})}_{\text{additive}}\right)$$

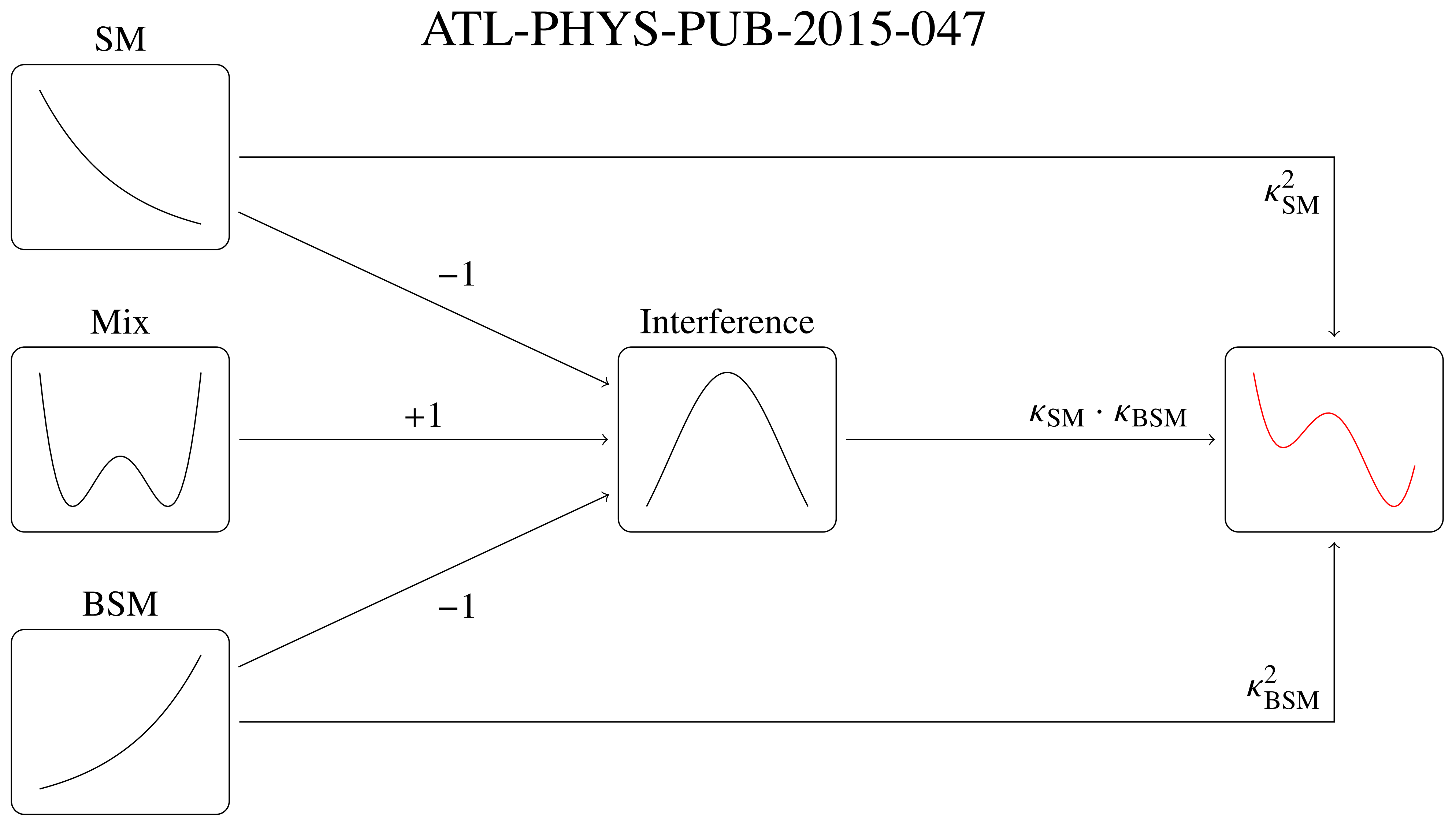
Use: Multiple disjoint channels (or regions) of binned distributions with multiple samples contributing to each with additional (possibly shared) systematics between sample estimates

Main pieces:

- Main Poisson p.d.f. for simultaneous measurement of multiple channels
- Event rates $\nu_{cb}(\vec{\eta}, \vec{\chi})$ (nominal rate ν^0_{scb} with rate modifiers)
  - encode systematic uncertainties (e.g. normalization, shape)
- Constraint p.d.f. (+ data) for "auxiliary measurements"

4

EFT "morphing" trick



As one changes the parameters of the EFT, the distributions $p(x|\alpha)$ change due to interference.

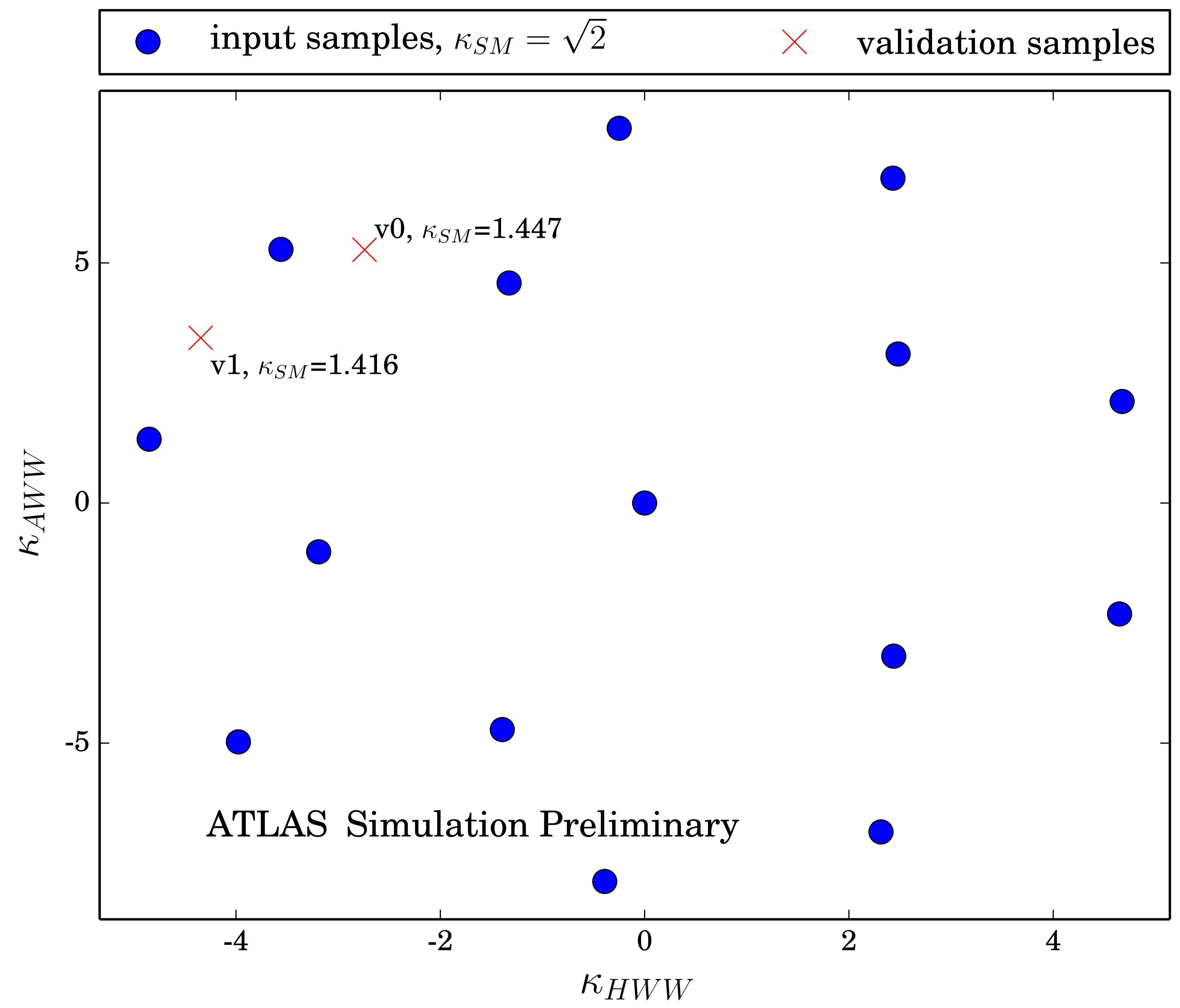
But there is a trick:

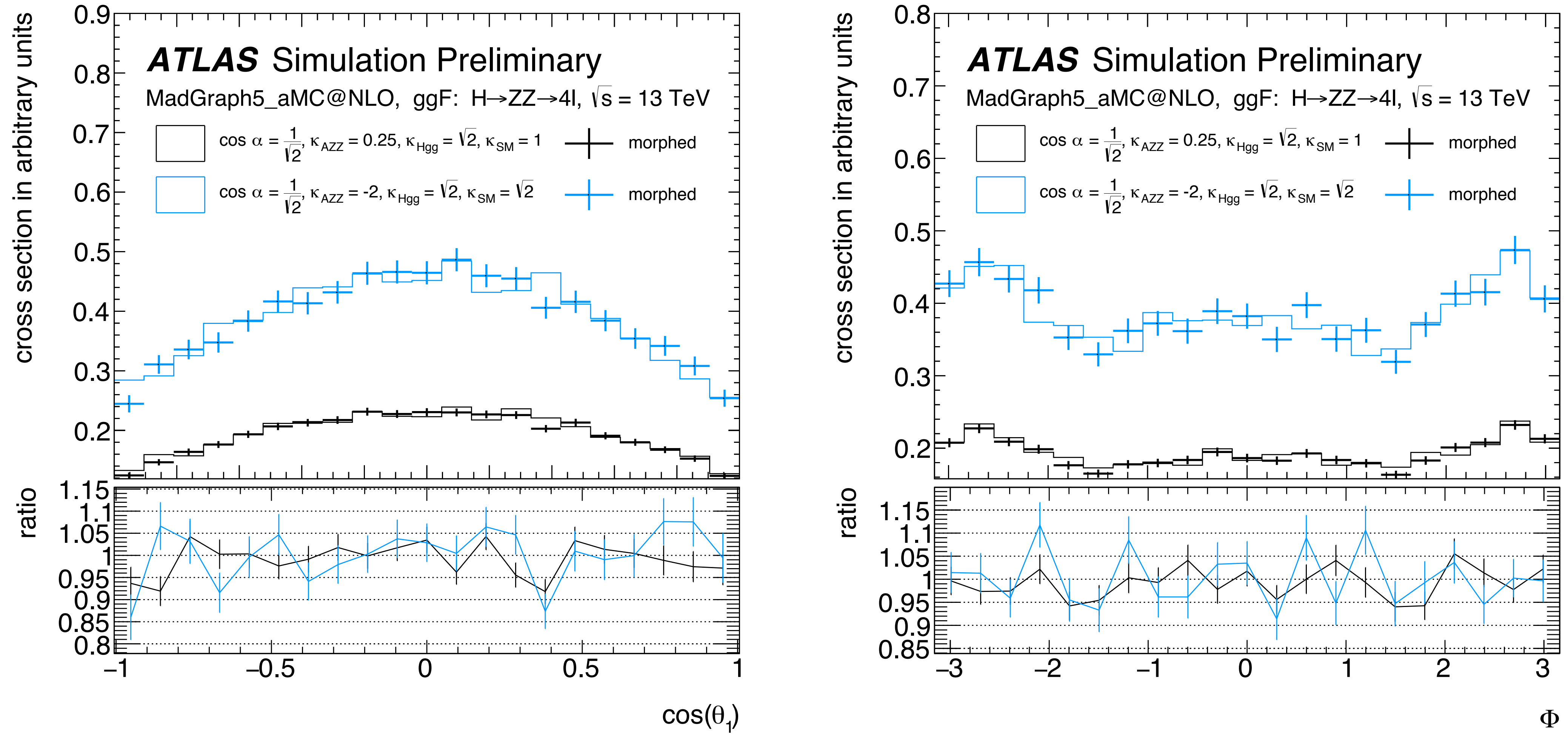
Simple example:

$$|g_1 M_{SM} + g_2 M_{BSM}|^2 = g_1^2 |M_{SM}|^2 + 2 g_1 g_2 Re\left[M^*_{SM} M_{BSM}\right] + g_2^2 |M_{BSM}|^2$$

3-d vector space, distribution for any point in this space is linear mixture of distribution for 3 basis samples!

(real examples need more basis samples)





Physical Positive Probabilities

EFT "morphing" trick

production / decay

$$d\sigma \propto \left| \left(\mathcal{M}^p_{\text{SM}} + \sum_i \frac{f_i}{\Lambda^2} \mathcal{M}^p_i \right) \left(\mathcal{M}^d_{\text{SM}} + \sum_j \frac{f_j}{\Lambda^2} \mathcal{M}^d_j \right) \right|^2$$

Express EFT as a mixture:

$$p(x|\alpha) = \sum_c w_c(\alpha) p_c(x)$$

Fully differential cross-section

$w_c(\alpha)$ are polynomials, $p_c(x)$ are physical distributions!

Can truncate to $\mathcal{O}(\Lambda^{-n})$ if desired

Process	Number of components for n operators					
	$\mathcal{O}(\Lambda^0)$	$\mathcal{O}(\Lambda^{-2})$	$\mathcal{O}(\Lambda^{-4})$	$\mathcal{O}(\Lambda^{-6})$	$\mathcal{O}(\Lambda^{-8})$	Σ
hV / WBF production	1	n	$\frac{n(n+1)}{2}$			$\frac{(n+1)(n+2)}{2}$
$h \to VV$ decay	1	n	$\frac{n(n+1)}{2}$			$\frac{(n+1)(n+2)}{2}$
Production + decay	1	n	$\frac{n(n+1)}{2}$	$\binom{n+2}{3}$	$\binom{n+3}{4}$	$\binom{n+4}{4}$

Table 1: Number of components c as given in Eq. (6) for different processes, sorted by their suppression by the EFT cutoff scale Λ.

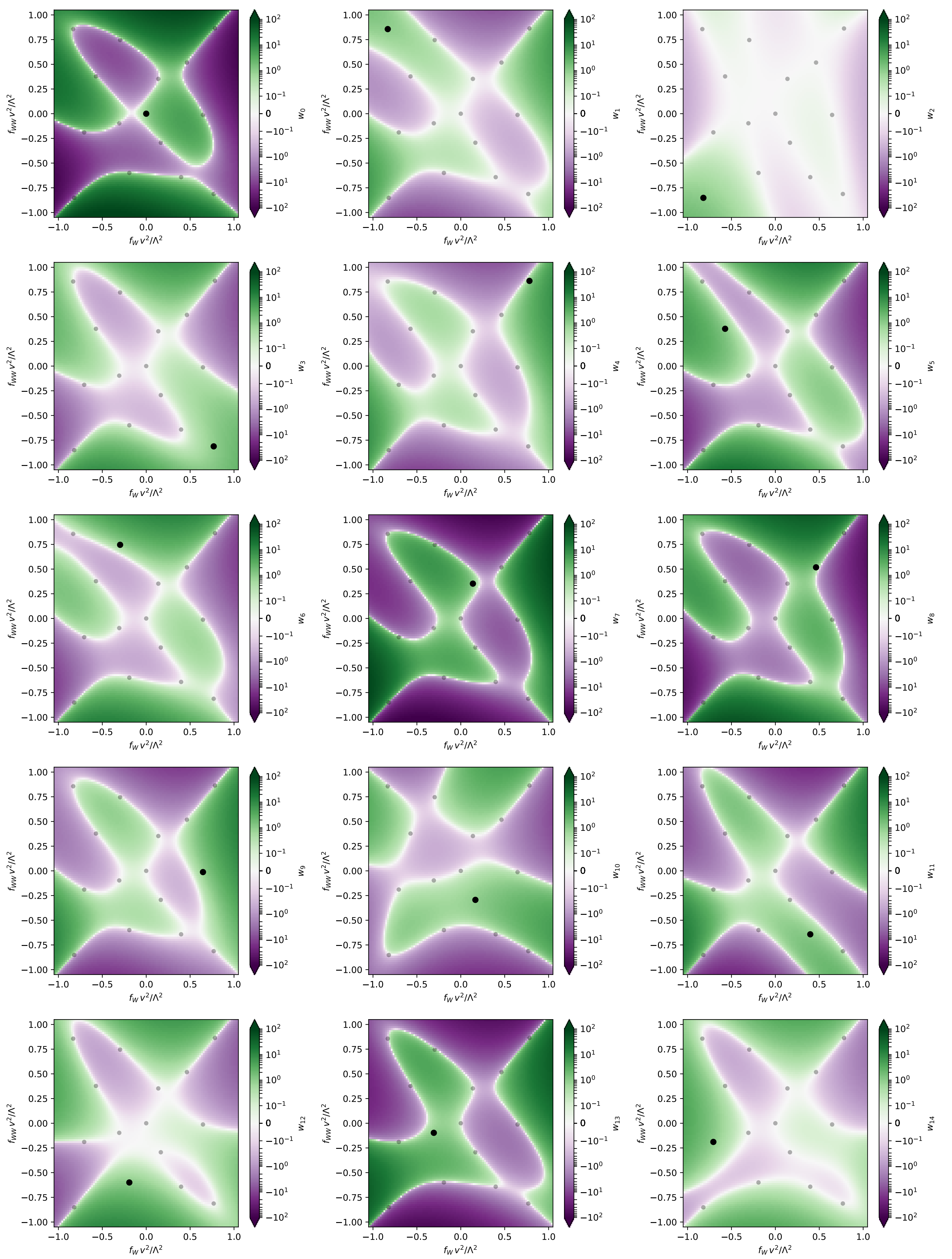


Figure 13: Morphing weights $w_i(\boldsymbol{\theta})$ for basis points distributed over the full relevant parameter space.

For 2 BSM operators affecting VBF Higgs production and decay, we need a 15-D vector space

For 5 BSM operators we need 126-D vector space

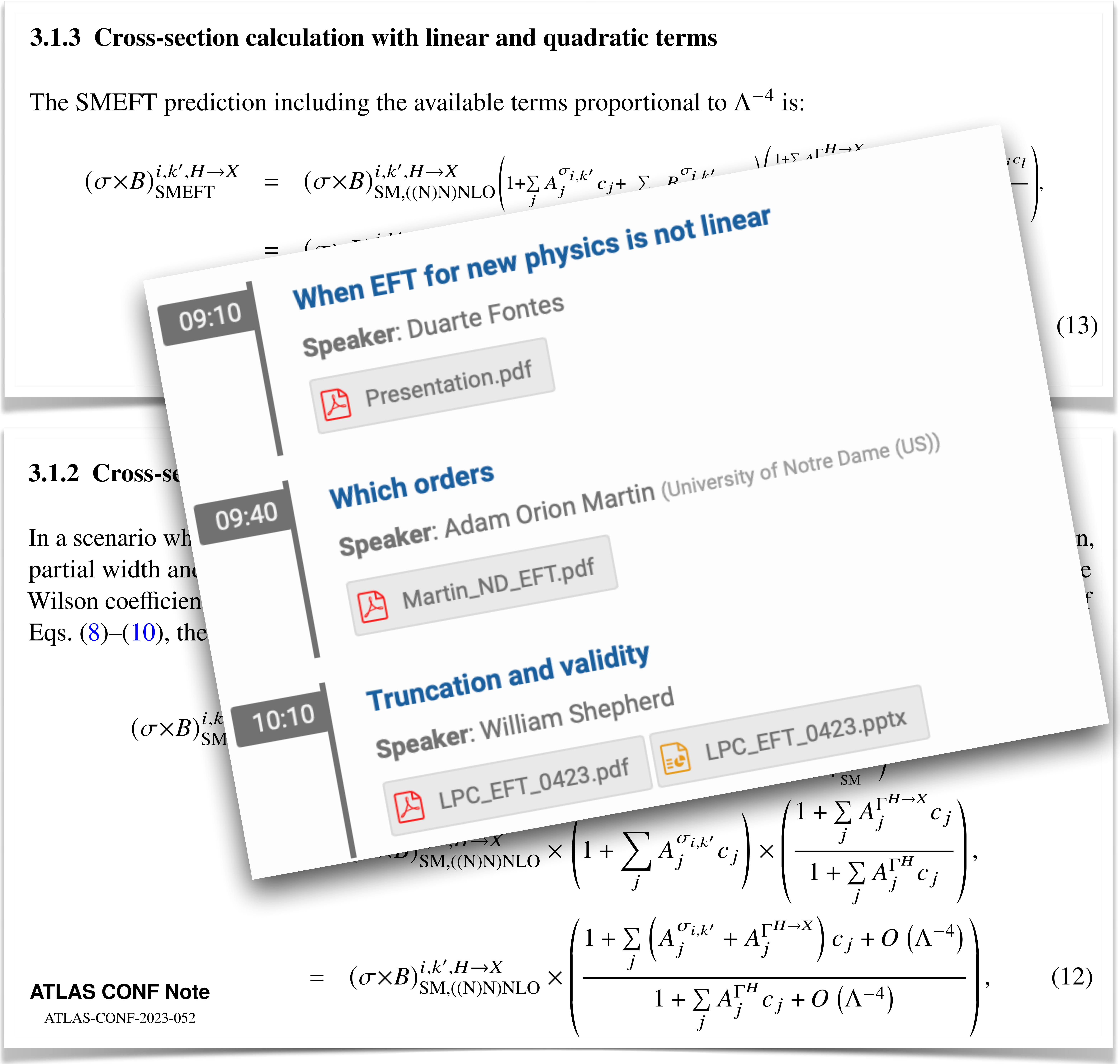
This is implemented in MadMiner

Other descriptions

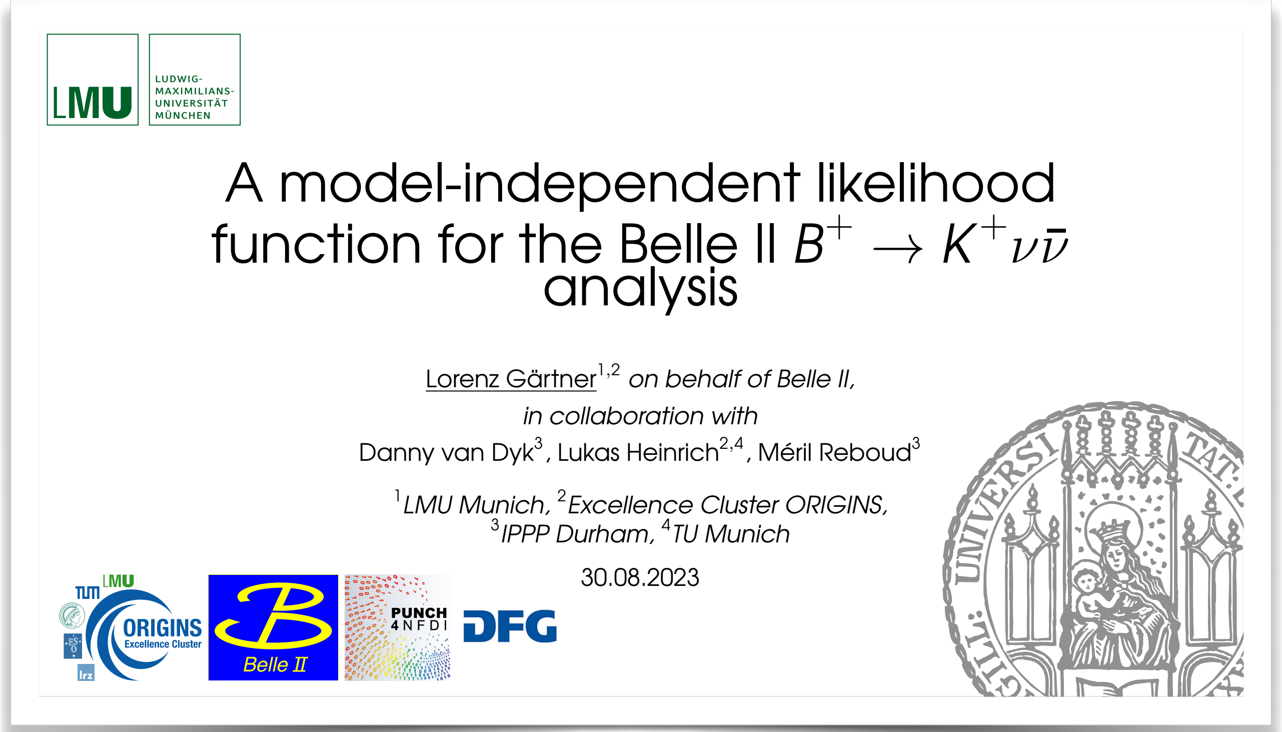
Same idea, different in details

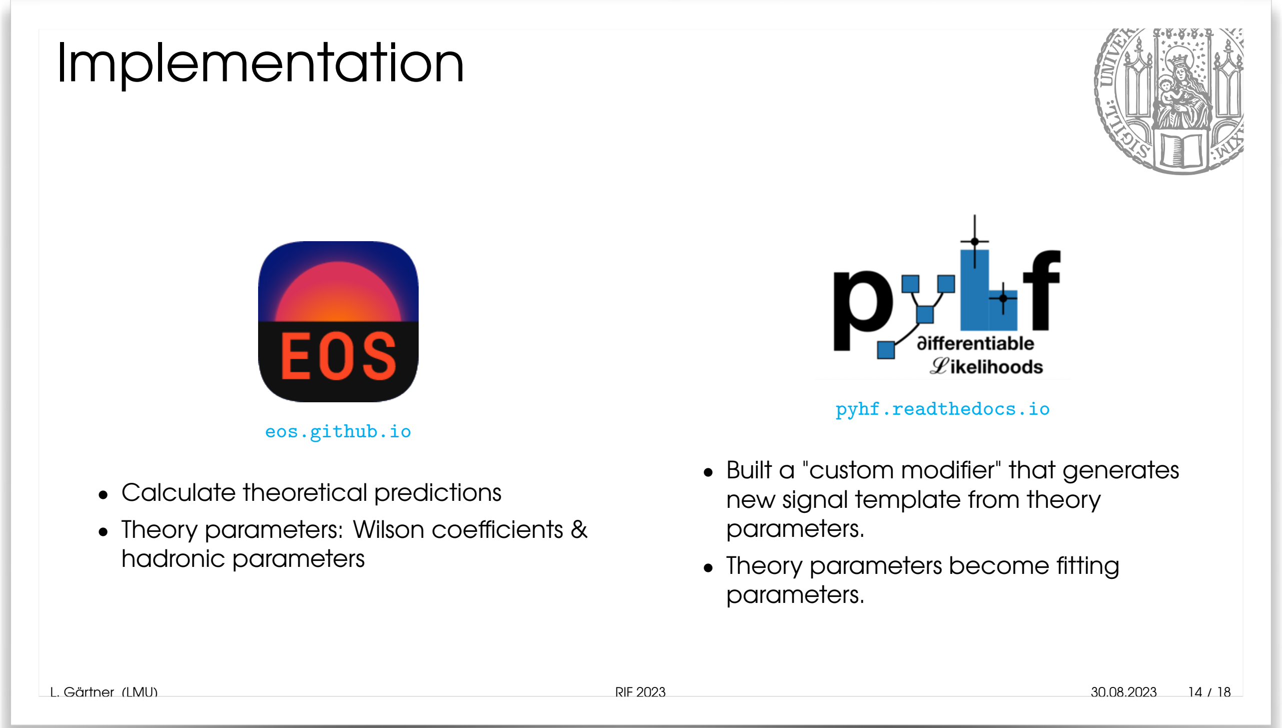
Here are two concrete examples for describing how the (truth-level) fiducial cross section in phase space region k' depends on the EFT coefficients $\alpha = \{c_j\}$

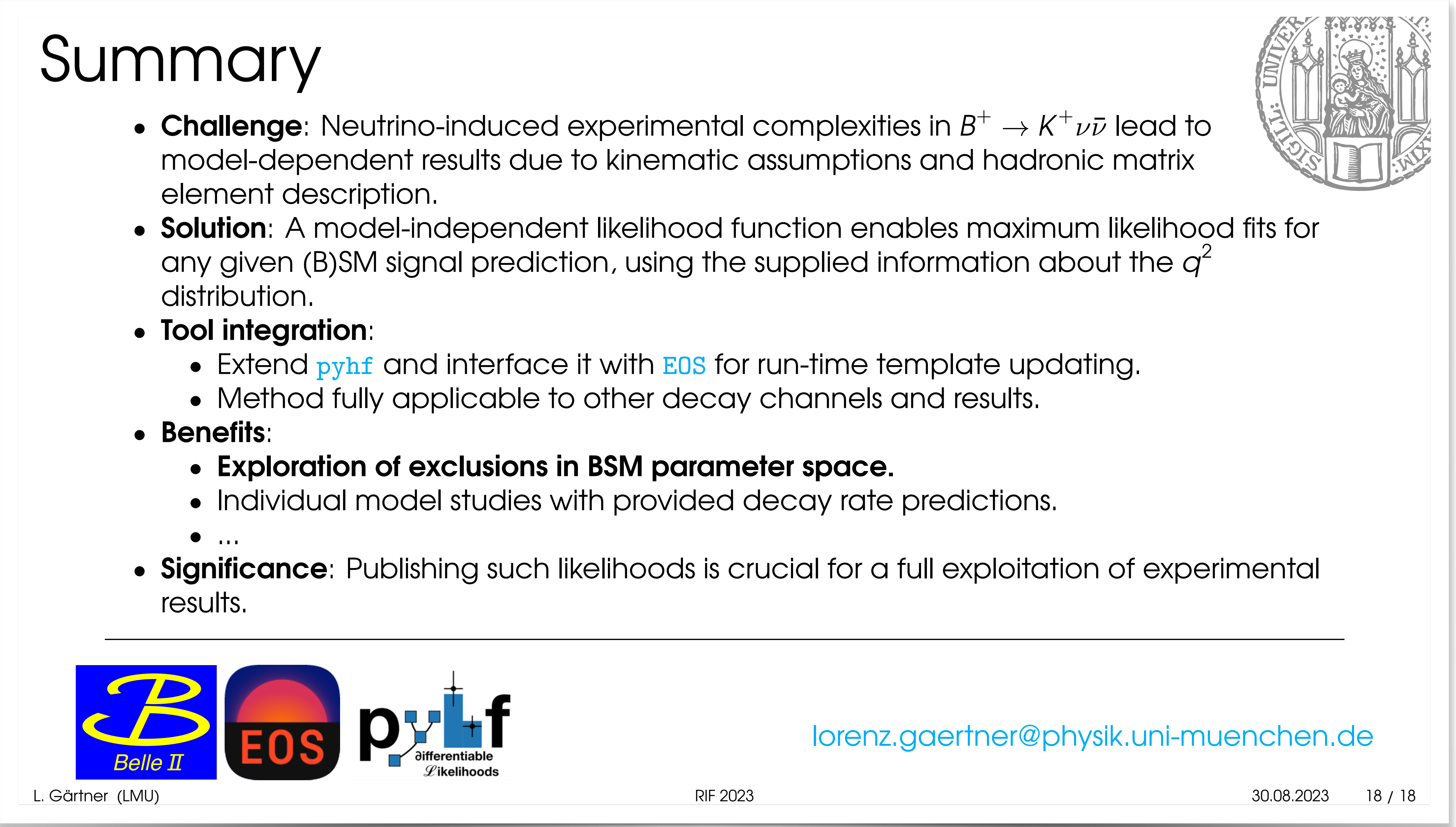
- Can extend to fully differential cross-section $\left.\frac{d\sigma(\alpha)}{dz}\right|_{z_i}$ where z_i is the truth-level kinematics
- Used in MadMiner for Simulation-based inference



Extending template specifications for EFT fits



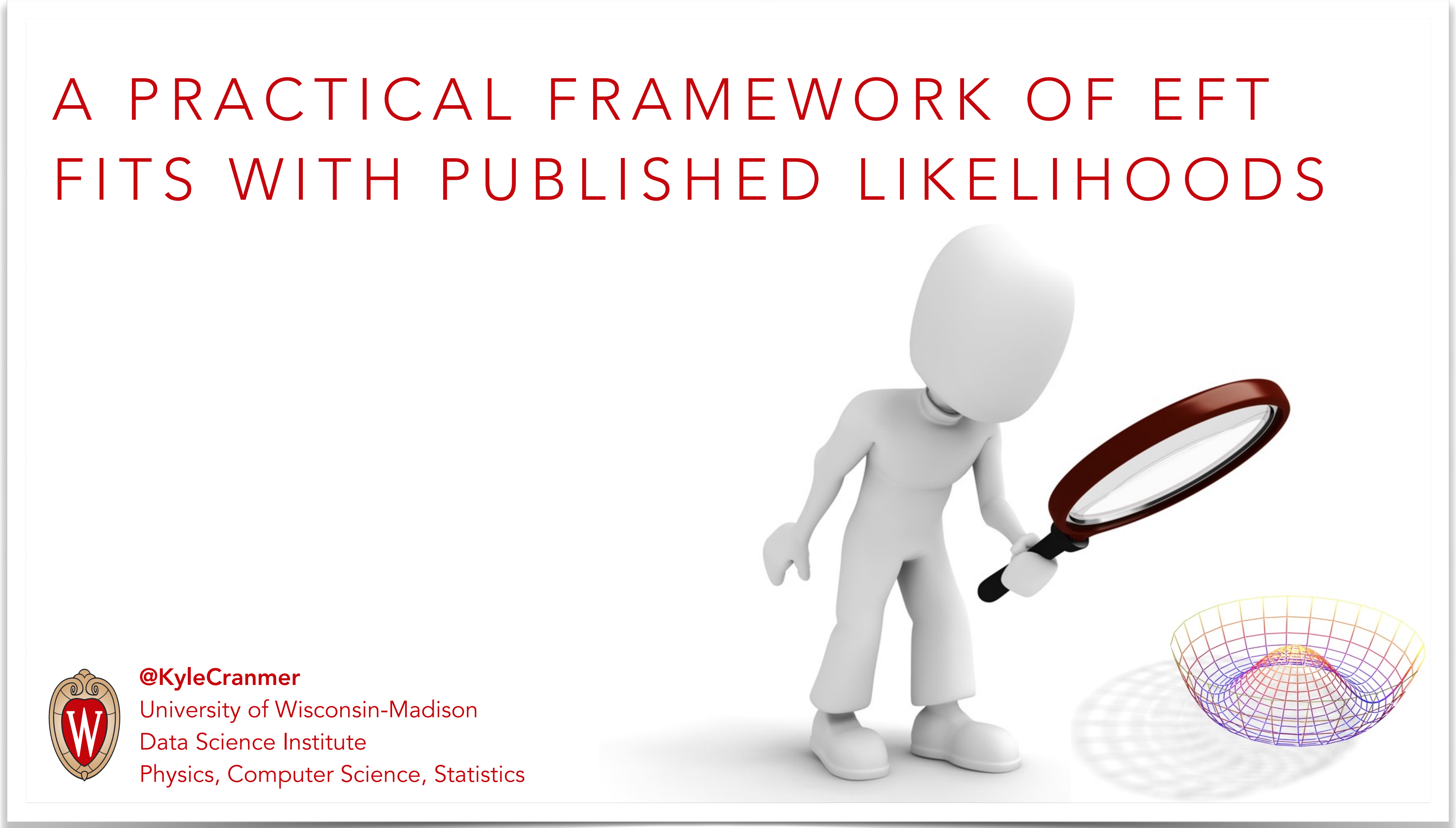




This was done by Belle II

- Subject of my talk at LHC EFT WG, with some more details about "on-the-fly" reweighting

Focus of LHC EFT effort should be to converge on the specification(s).



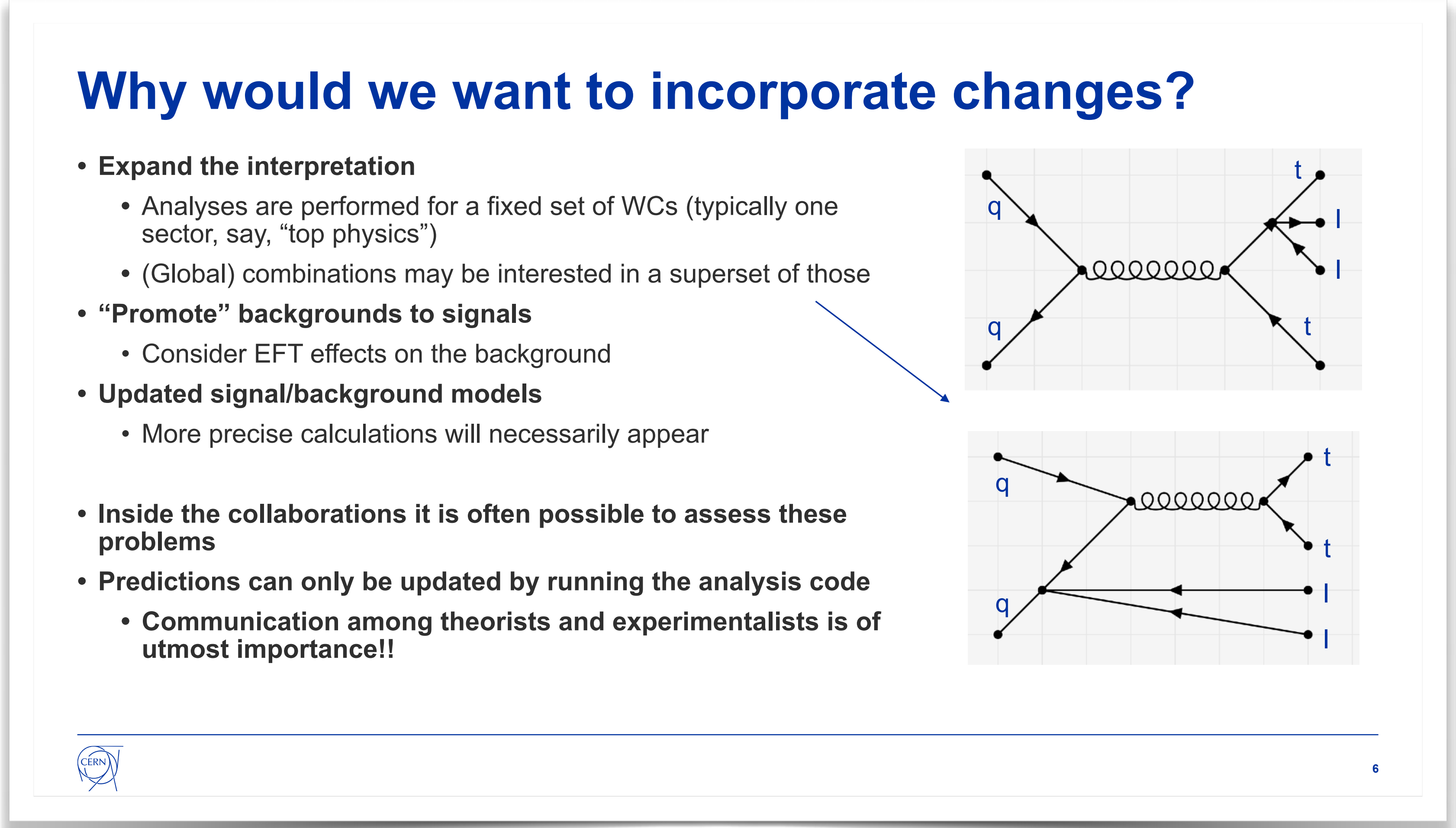
https://conference.ippp.dur.ac.uk/event/1178/contributions/6443/
https://indico.cern.ch/event/1296757/timetable/#3-area-4-a-practical-framework

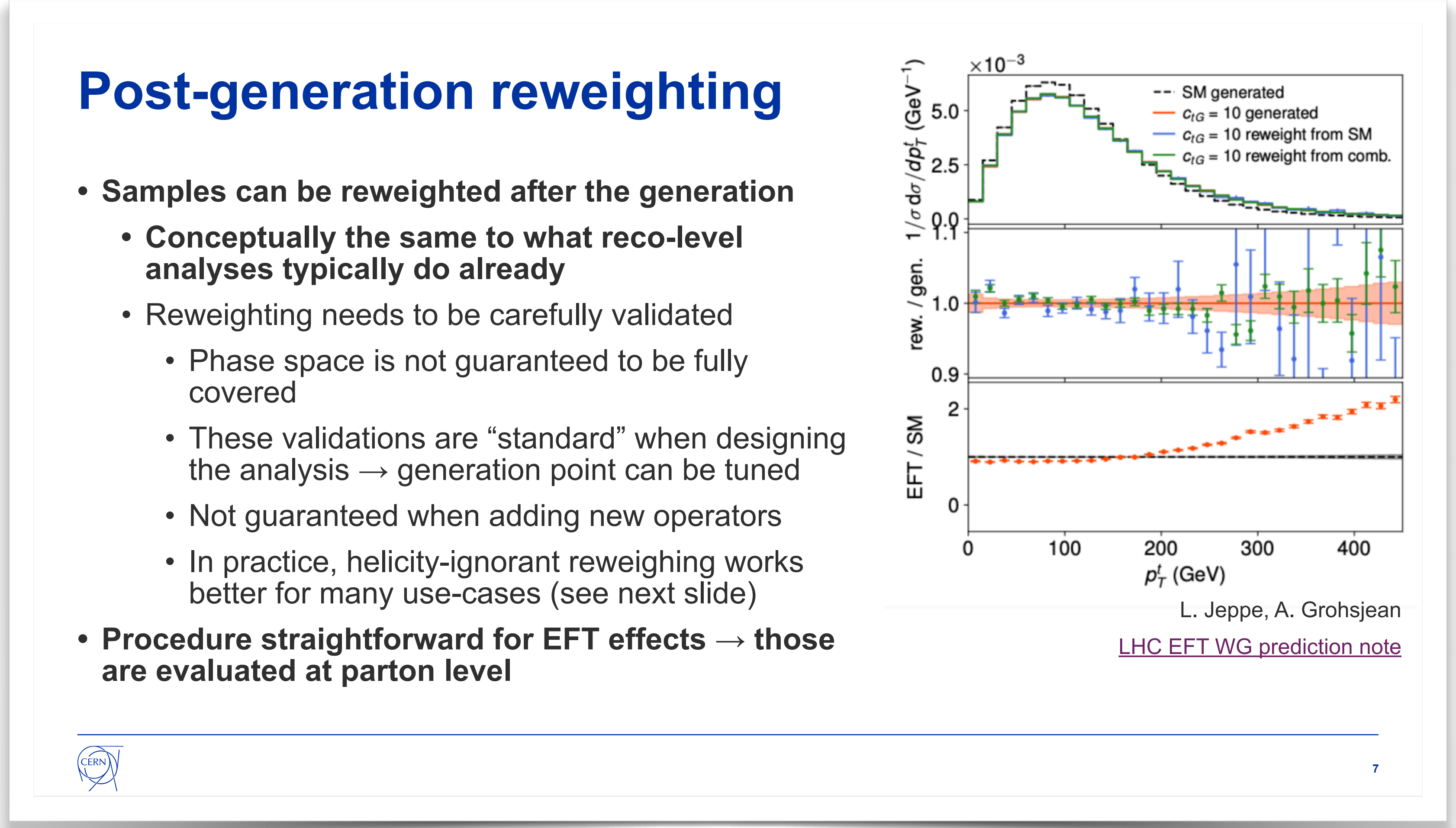
Message 3:
Extensible EFT reinterpretation with RECAST infrastructure

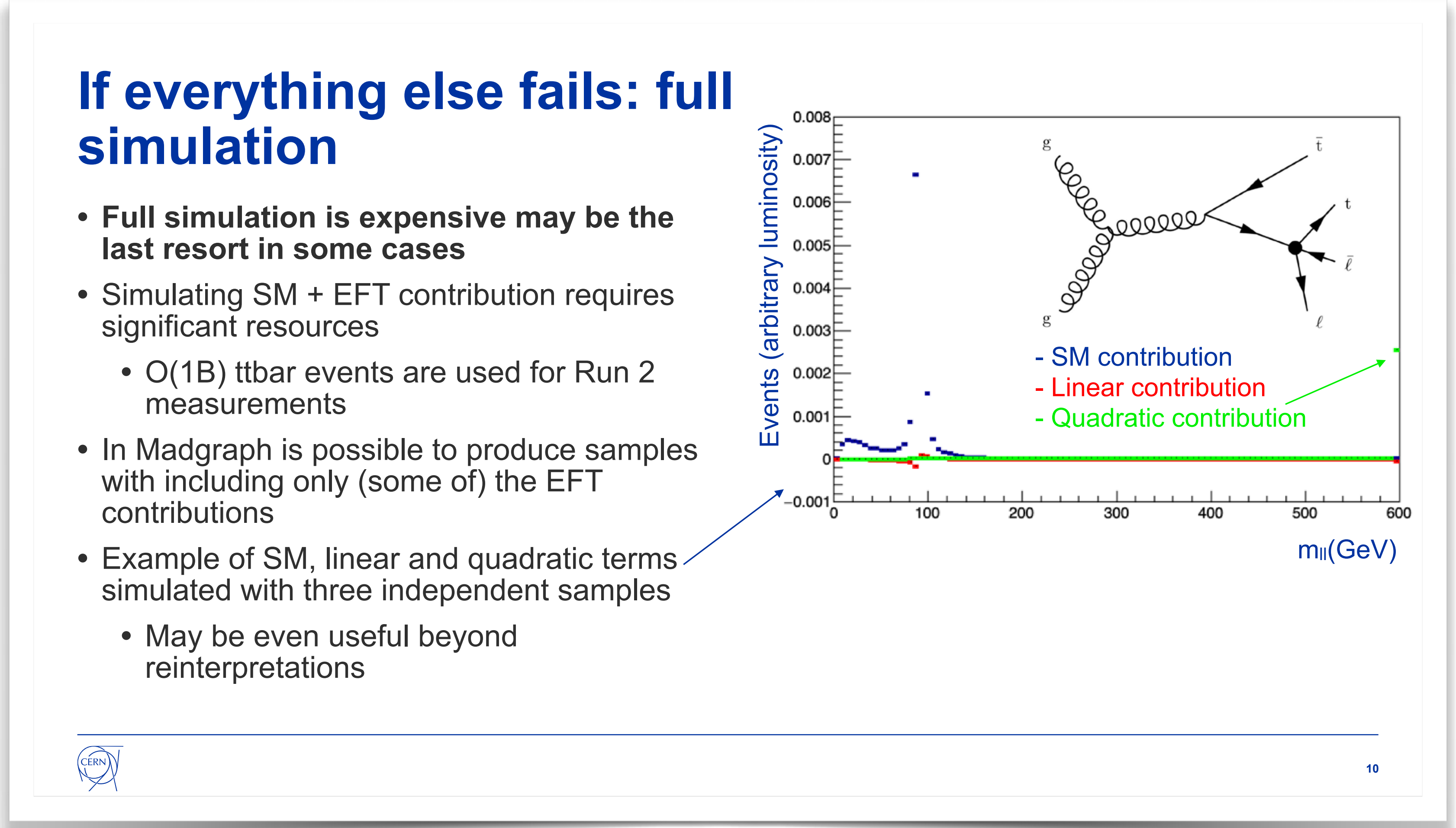
Extending the EFT reinterpretation

This morning, Sergio covered motivations and strategies for incorporating changes at a lower level (new EFT operators, changes to sig/bkg models, etc.) keeping the analysis strategy (event selection, observables, etc.) fixed

- Update signal / background components & export a new statistical model







In BSM context, this is often referred to as "recasting" and we have built infrastructure to do this

Preservation & Reinterpretation

First results using the RECAST reinterpretation framework and publishing full statistical likelihoods (using pyhf) in 2019

- Recent pMSSM effort uses 10 searches, 19 parameter pMSSM, and 20,000 parameter points

recast reana

ATLAS PUB Note

ATL-PHYS-PUB-2019-029

5th August 2019

Reproducing searches for new physics with the ATLAS experiment through publication of full statistical likelihoods

The ATLAS Collaboration

The ATLAS Collaboration is starting to publicly provide likelihoods associated with statistical fits used in searches for new physics on HEPData. These likelihoods adhere to a specification first defined by the HistFactory p.d.f. template. This note introduces a JSON schema that fully describes the HistFactory statistical model and is sufficient to reproduce key results from published ATLAS analyses. This is per-se independent of its implementation in ROOT and it can be used to run statistical analysis outside of the ROOT and RooStats/RooFit framework. The first of these likelihoods published on HEPData is from a search for bottom-squark pair production. Using two independent implementations of the model, one in ROOT and one in pure Python, the limits on the bottom-squark mass are reproduced, underscoring the implementation independence and long-term viability of the archived data.

ATL-PHYS-PUB-2019-029
05 August 2019

ATLAS PUB Note

ATL-PHYS-PUB-2019-032

11th August 2019

RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model of a dark Higgs boson decaying to two b-quarks

The ATLAS Collaboration

The reinterpretation of a search for dark matter produced in association with a Higgs boson decaying to b-quarks performed with RECAST, a software framework designed to facilitate the reinterpretation of existing searches for new physics, is presented. Reinterpretation using RECAST is enabled through the sustainable preservation of the original data analysis as re-executable declarative workflows using modern cloud technologies and integrated with the wider CERN Analysis Preservation efforts. The reinterpretation targets a model predicting dark matter production in association with a hypothetical dark Higgs boson decaying into b-quarks where the mass of the dark Higgs boson m_s is a free parameter, necessitating a faithful reinterpretation of the analysis. The dataset has an integrated luminosity of 79.8 fb^{-1} and was recorded with the ATLAS detector at the Large Hadron Collider at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV. Constraints on the parameter space of the dark Higgs model for a fixed choice of dark matter mass m_χ = 200 GeV exclude model configurations with a mediator mass up to 3.2 TeV.

ATL-PHYS-PUB-2019-032
12 August 2019

ENERGY FRONTIERS

Reports from the Large Hadron Collider experiments

ATLAS

Electroweak SUSY after LHC Run 2

Supersymmetry (SUSY) provides elegant solutions to many of the problems of the Standard Model (SM) by introducing new boson/fermion partners for each SM fermion/boson, and by extending the Higgs sector. If SUSY is realised in nature at the TeV scale, it would accommodate a light Higgs boson without excessive fine-tuning. It could furthermore provide a viable dark-matter candidate, and be a key ingredient to the unification of the electroweak and strong forces at high energy. The SUSY partners of the SM bosons can mix to form what are called charginos and neutralinos, collectively referred to as electroweakinos.

Electroweakinos would be produced only through the electroweak interaction, where their production cross sections in proton–proton collisions are orders of magnitude smaller than strongly produced squarks and gluinos (the supersymmetric partners of quarks and gluons). Therefore, while extensive searches using the Run 1 (7–8 TeV) and Run 2 (13 TeV) LHC datasets have turned up null results, the corresponding chargino/neutralino exclusion limits remain substantially weaker than those for strongly interacting SUSY particles.

The ATLAS collaboration has recently released a comprehensive analysis of the electroweak SUSY landscape based on its Run 2 searches. Each individual search targeted specific chargino/neutralino production mechanisms and subsequent decay modes. The analyses were originally interpreted in so-called "simplified models", where only one production mechanism is considered, and only one possible decay. However, if SUSY is realised in nature, its particles will have many possible production and decay modes, with rates depending on the SUSY parameters. The new ATLAS analysis brings these pieces together by reinterpreting 10 searches in the phenomenological Minimal Supersymmetric Standard Model (pMSSM), which includes a range of SUSY particles, production mechanisms and decay modes governed by 19 SUSY parameters. The results provide a global picture of ATLAS's sensitivity to electroweak SUSY and, importantly, reveals the gaps that remain to be explored.

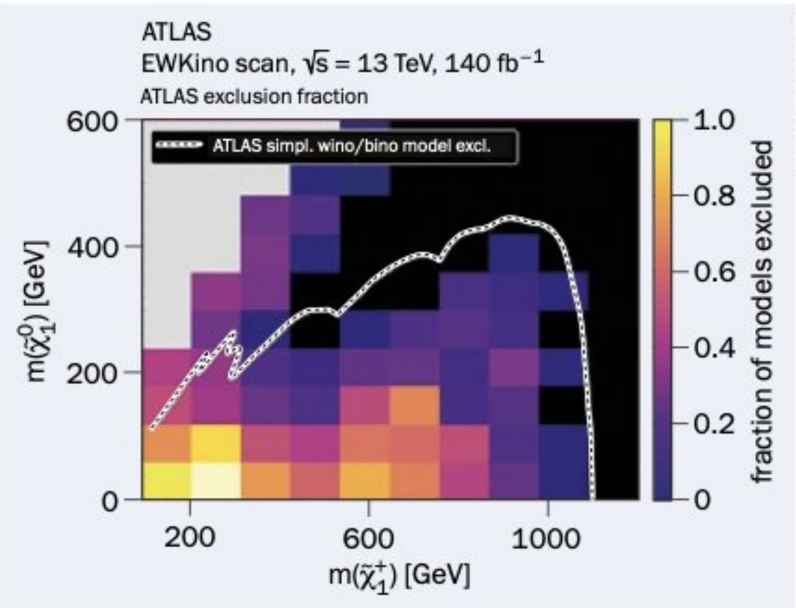


Fig. 1. *The fraction of pMSSM models excluded by ATLAS in the plane of the lightest chargino mass (x-axis) versus the lightest neutralino mass (y-axis). The dashed line shows the exclusion of simplified SUSY models reported by individual searches in this mass plane.*

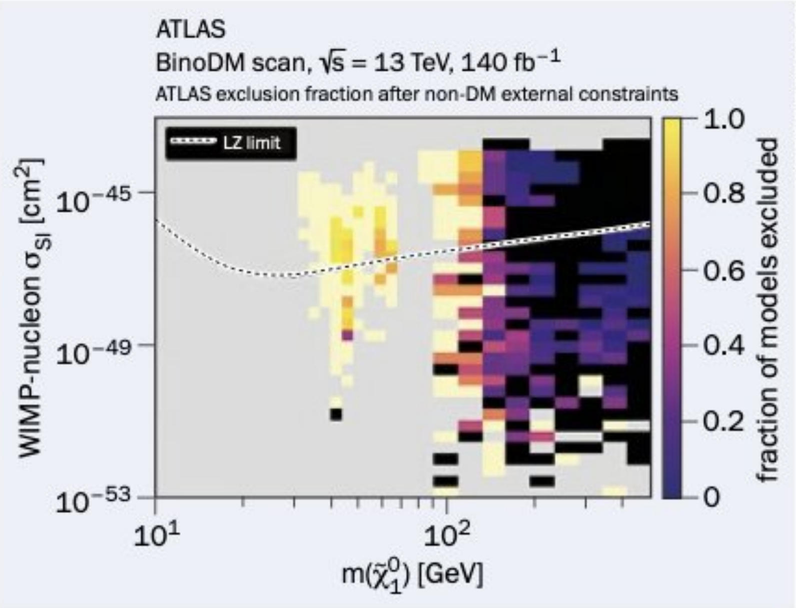


Fig. 2. *The fraction of pMSSM models excluded by ATLAS in the plane of the lightest neutralino (i.e. the dark-matter candidate) mass (x-axis) versus the spin-independent WIMP–nucleon scattering cross-section (y-axis). The dashed line shows the upper limits from the LZ experiment.*

The 19-dimensional pMSSM parameter space was randomly sampled to produce a set of 20,000 SUSY model points. The 10 selected ATLAS searches were then performed on each model point to determine whether it is excluded with at least 95% confidence level. This involved simulating datasets for each SUSY model, and re-running the corresponding analyses and statistical fits. An extensive suite of reinterpretation tools was employed to achieve this, including preserved likelihoods and RECAST – a framework for preserving analysis workflows and re-applying them to new signal models.

The results show that, while electroweakino masses have been excluded up to 1 TeV in simplified models, the coverage with regard to the pMSSM is not exhaustive. Numerous scenarios remain viable, including mass regions nominally covered by previous searches (inside the dashed line in figure 1). The pMSSM models may evade detection due to smaller production cross-sections and decay probabilities compared to simplified models. Scenarios with small mass-splittings between the lightest and next-to-lightest neutralino can reproduce the dark-matter relic density, but are particularly elusive at the LHC. The decays in these models produce challenging event features with low-momentum particles that are difficult to reconstruct and separate from SM events.

Beyond ATLAS, experiments such as LZ aim at detecting relic dark-matter particles through their scattering by target nuclei. This provides a complementary probe to ATLAS searches for dark matter produced in the LHC collisions. Figure 2 shows the LZ sensitivity to the pMSSM models considered by ATLAS, compared to the sensitivity of its SUSY searches. ATLAS is particularly sensitive to the region where the dark-matter candidate is around half the Z/Higgs-boson mass, causing enhanced dark-matter annihilation that could have reduced the otherwise overabundant dark-matter relic density to the observed value.

The new ATLAS results demonstrate the breadth and depth of its search programme for supersymmetry, while uncovering its gaps. Supersymmetry may still be hiding in the data, and several scenarios have been identified that will be targeted, benefiting from the incoming Run 3 data.

Further reading

ATLAS Collab. 2024 arXiv:2402.01392.

A RECAST-like service for EFTs

Consider the case where ATLAS and CMS publish statistical models parametrized for some subset of operators in a specified EFT basis.

- **Sometime later** one wants to reinterpret the analysis for **a different set of operators** keeping the same event selection, breakdown of signal and control regions, observables, binning, etc.

RECAST is a framework for reinterpretations like this for BSM searches

- In general, this requires running new signal through the full MC simulation + reco + analysis chain. ATLAS is actually doing this with preserved analysis workflows!

In most cases for EFTs we can simply **reweight the existing fully simulated SM events** (doesn't require running more simulation, reconstruction, etc.)

- The service could calculate the coefficients for the mini-database based on truth-level kinematics and export a new statistical model that implements the specification as describe above.

FAIROS-HEP

What is FAIROS-HEP?

Recently, the US National Science Foundation funded a new Research Coordination Network project titled "FAIROS-HEP".

FAIROS = **F**indable. **A**ccessible. **I**nteroperable. **R**eusable. **O**pen **S**cience.

The FAIROS-HEP project aims to connect groups of researchers thinking about FAIR data in HEP and other experts in this field to envision a more **cohesive infrastructure** around **data and publications** in HEP.

- By focusing on FAIR data practices and how data and software can be linked to physics results, we hope to build a network of researchers thinking about how we can create a "living publication" to preserve and extend physics results.
- The project includes **some funding for building infrastructure** as well as future **workshops connecting groups**.

FAIROS-HEP Continues a Legacy of Contributions

DASPOS (2012-2016)

- https://daspos.crc.nd.edu/
- Contributions to **RECAST** led to **REANA** as a spinoff project now led by CERN
- Supported REANA Common Workflow Language

DIANA-HEP (2015-2021)

- https://diana-hep.org/
- Contributions to **REANA**, **RECAST**, launched **pyhf likelihood publishing**, early work in **simulation-based inference**, Active Learning for reinterpretation
- Supported GitHub -> Zenodo DOI minting

IRIS-HEP (2018-?)

- https://iris-hep.org/
- Major contributions to likelihood publishing, **HEPData** integration,

SCAILFIN (2018-2021)

- https://scailfin.github.io/
- Contributions to REANA (Slurm and HPC backends, applications built on top of REANA), Active Learning for reinterpretation

FAIROS-HEP (2022-2025)

- https://fairos-hep.org/
- **Continue the legacy of contributions, help coordinate the ecosystem**

Conclusion

Top Level Message: We should publish the full statistical model (aka "likelihood") for measurements that constrain EFT coefficients

- Lots of progress in publishing statistical models recently in BSM searches

Second Level Message: There are a few ways to describe the dependence on EFT parameters. We can and should separate the specification and implementation.

- First define a **specification** for one or more of these choices that removes all ambiguity. This allows multiple groups to **implement** the specification.

Third Level Message: In addition to publishing statistical models, RECAST-like infrastructure would allow us to consider new EFT operators and update / improve background modeling after publishing

- This infrastructure is being used in BSM searches already

Bonus: The FAIROS-HEP project has funds to support (travel to) workshops to coordinate the design of this infrastructure.

PhyStat workshop on SBI in Munich

PHYSTAT-SBI 2024 - Simulation Based Inference in Fundamental Physics

15–17 May 2024
Max Planck Institute for Physics
Europe/Zurich timezone

Enter your search term

- Overview
- Call for Abstracts
- Timetable
- My Conference
  - My Contributions
- Registration
- Participant List
- Videoconference
- Practical Information

Fueled by the recent advances of Machine Learning in the last decade, a new breed of techniques have been developed to tackle statistical inference problems for "likelihood-free" cases, where it is possible to sample from the data-generating process (i.e. via stochastic simulators) but a closed form evaluation of the density is intractable.

This group of methods is known as "simulation-based inference" (SBI) or "likelihood-free inference" (LFI) and will be the dedicated topic of this PHYSTAT Workshop taking place from May 15th - May 17th 2024 at the Max-Planck Institute for Physics (MPP) in Garching near Munich.

PHYSTAT https://phystat.github.io/Website/) is a long-running workshop series that brings together statisticians, machine learning researchers and physicists to discuss shared topics and foster collaboration among the research communities.

Confirmed Speakers:

- Kyle Cranmer (U Wisconsin-Madison)
- Antoine Wehenkel (Apple)
- Gilles Louppe (U Liège)
- Laurence Levasseur (U Montréal)
- Ann Lee (Carnegie Mellon)
- Julia Linhart (INRIA)
- Noemi Montel (U of Amsterdam)
- Jakub Tomczak (Eindhoven)
- Christoph Weniger (GRAPPA)
- Alexander Held (U Wisconsin-Madison)

https://indico.cern.ch/event/1355601/overview

Backup & Referees

Related talks

Talk at Higgs 2020

- https://indico.cern.ch/event/900384/contributions/3796227/
- Unfolding vs. simplified likelihoods vs. full statistical models
- STXS vs. Fully differential
- Simulation-based inference

Talk at LHC EFT WG in 2023

- https://indico.cern.ch/event/1296757/
- Full statistical models and on-the-fly reweighting

A brief idea

Journal of Brief Ideas   Home   New idea   Trending ideas   All ideas   About   Search   Contact

Recasting through reweighting

By Kyle Cranmer, Lukas Heinrich

physics   reinterpretation   lhc

Recasting refers to reinterpreting the results of searches for new particles or standard model measurements in the context of different theoretical models [1]. The fundamental task is to replace the original hypothesis $p_0(x)$ with a new hypothesis $p_1(x)$, where x is some observed quantity. The effect of the detector response and analysis cuts can be encoded in a *folding* operator $\int, W(x|z)dz$ acting on the truth-level distribution $p(z)$. By keeping the analysis fixed, $W(x|z)$ does not change, thus recasting amounts to:

$$p_0(x) = \int, p_0(z)W(x|z)dz \quad \Longrightarrow \quad p_1(x) = \int, p_1(z)W(x|z)dz$$

There are two primary approaches:

- **folding**: Samples from $p_1(z)$ are run through a detector simulation and analysis chain to estimate $p_1(x)$ [2]. This is common when z is high-dimensional, $p_0(z)$ and $p_1(z)$ are very different, or $W(x|z)$ is sensitive to experimental details.
- **unfolding**: An alternate theory $p_1(z)$ is compared directly to an unfolded distribution $\hat{p}(z)$ obtained from applying an approximate inverse operation to the observed data. Typically, unfolding is restricted to low-dimensional x, z and Gaussian uncertainties.

We point out a third option

- **reweighting**: Reweight pre-folded events $(x_i, z_i) \sim, p_0(x, z)$ by the factor $r(z_i) = p_1(z_i)/p_0(z_i)$, as in

$$p_1(x) = \int p_1(z)W(x|z)dz = \int, p_0(z) \underbrace{\frac{p_1(z)}{p_0(z)}}_{\text{reweighting}} W(x|z)dz$$

This approach does not require simulating new events or the approximations used in unfolding. Note, sample variance becomes a problem if $r(z_i) \gg, 1$.

Sign in with ORCID

Authors

Kyle Cranmer, Lukas Heinrich

Metadata

DOI https:/doi.org/10.5281/zenodo.1013926

Published: 14 Oct, 2017

https://beta.briefideas.org/ideas/8106c030eba22dd3a8d268940d5e42d8

Exceprts from
Simulation-Based Inference

Conclusion

Likelihood fits in the data space are the gold standard for statistical inference

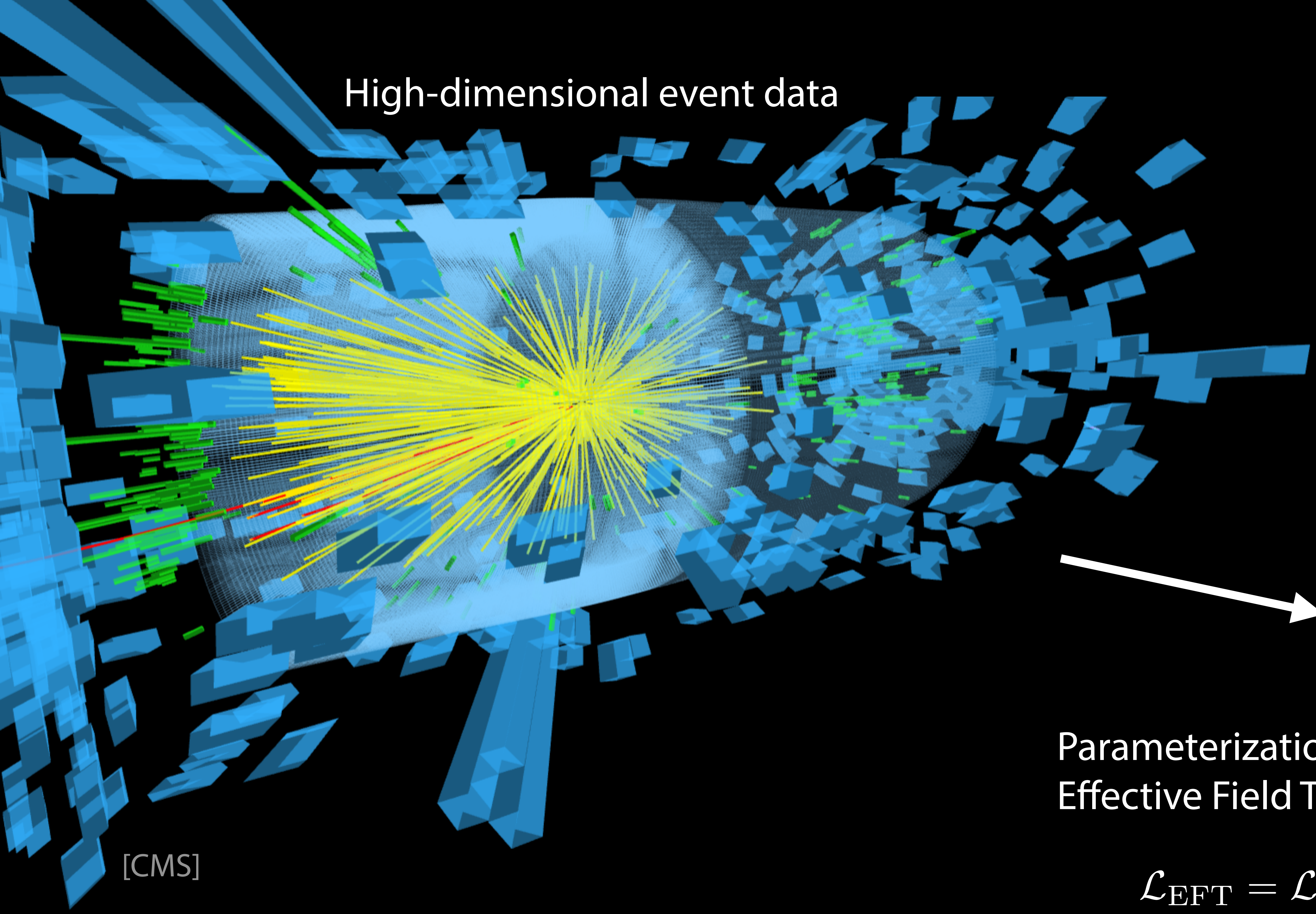
- RECAST and likelihood publishing are technical solutions that address model dependence and the theory-experiment interface
- STXS a good step, but more differential information can lead to large gain in sensitivity

Properties we want

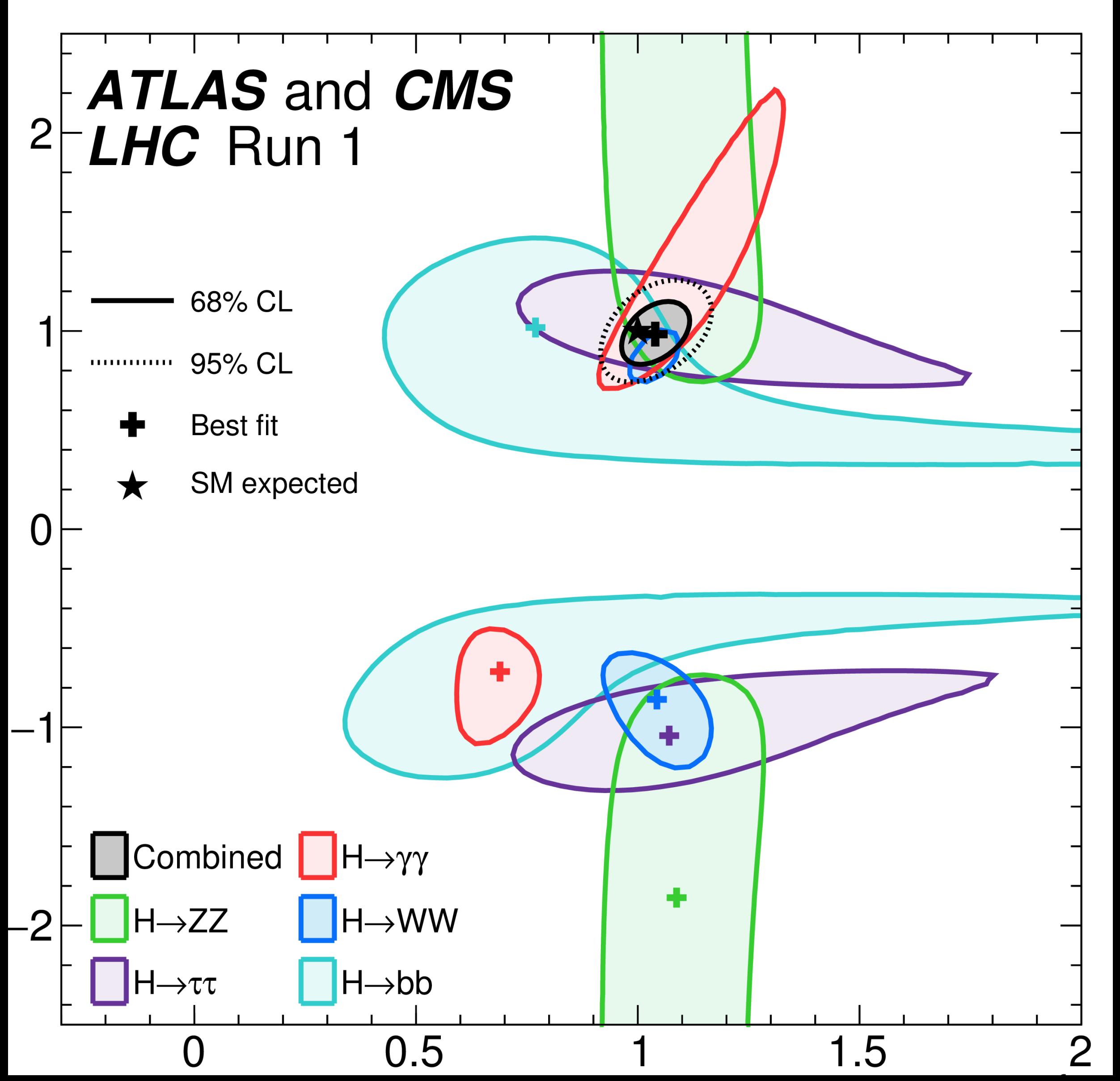
- Ability to be fully differential
- Exploit highest fidelity simulation (QCD, detector simulation) without approximations that introduce additional systematic errors
- Clear statistical motivation and compatibility with traditional combined analyses
- Scalability in terms of channels and parameters

The approach I presented (implemented in MadMiner) achieves these goals

High-dimensional event data



[CMS]



[ATLAS, CMS 1606.02266]

Precision constraints on new physics

Parameterization e.g. in Effective Field Theory:

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{f_i}{\Lambda^2} \mathcal{O}_i + \ldots$$

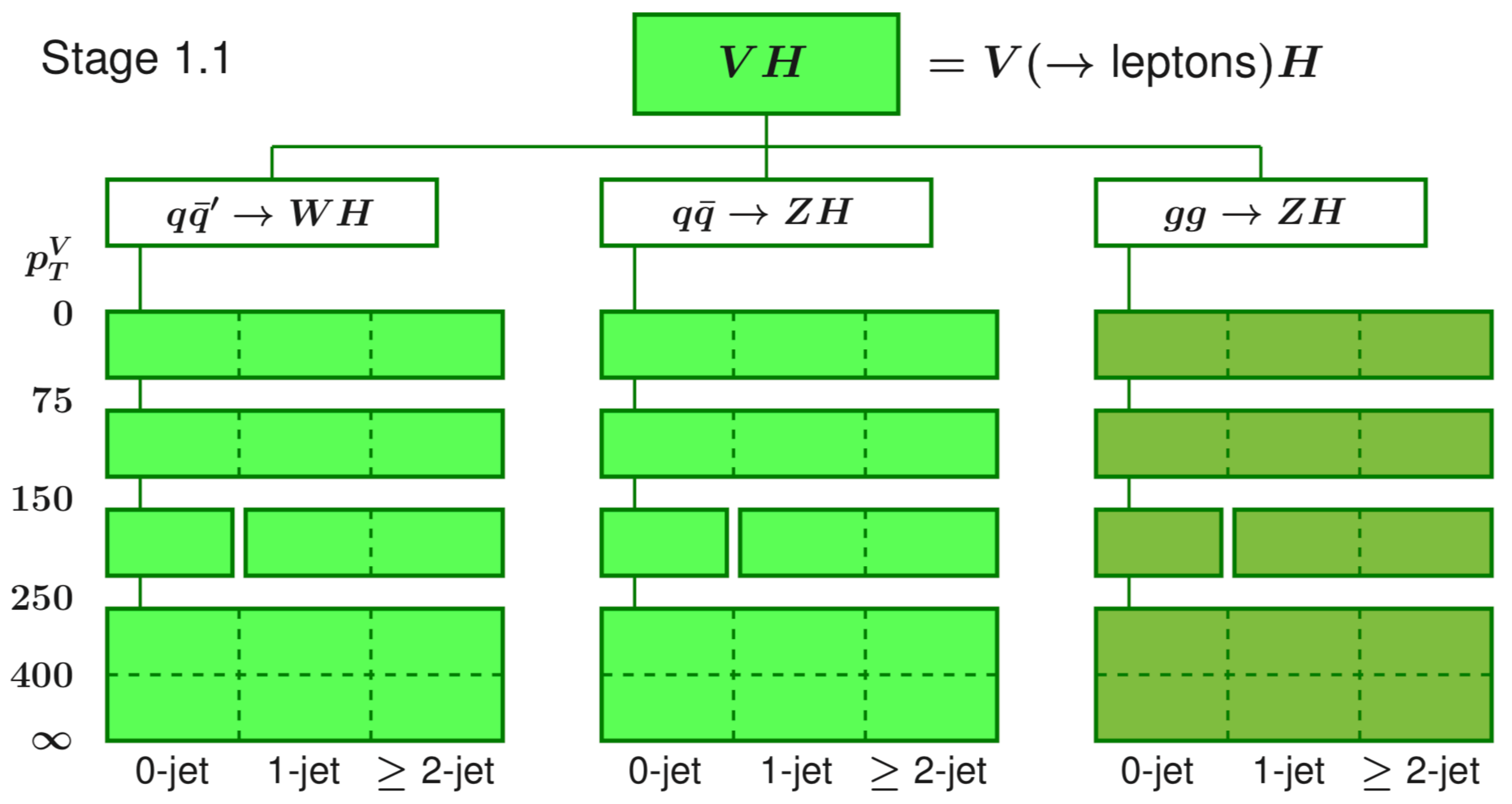
systematic expansion of new physics around Standard Model

10s to 100s "universal" parameters to measure

Benchmarking STXS in WH

[JB, S. Dawson, S. Homiller, F. Kling, T. Plehn 1908.06980]

- Simplified Template Cross-Sections (STXS) define observable bins that are supposed to capture as much information on NP as possible
  [N. Berger et al. 1906.02754; HXSWG YR4]



- Let's check! How much information on

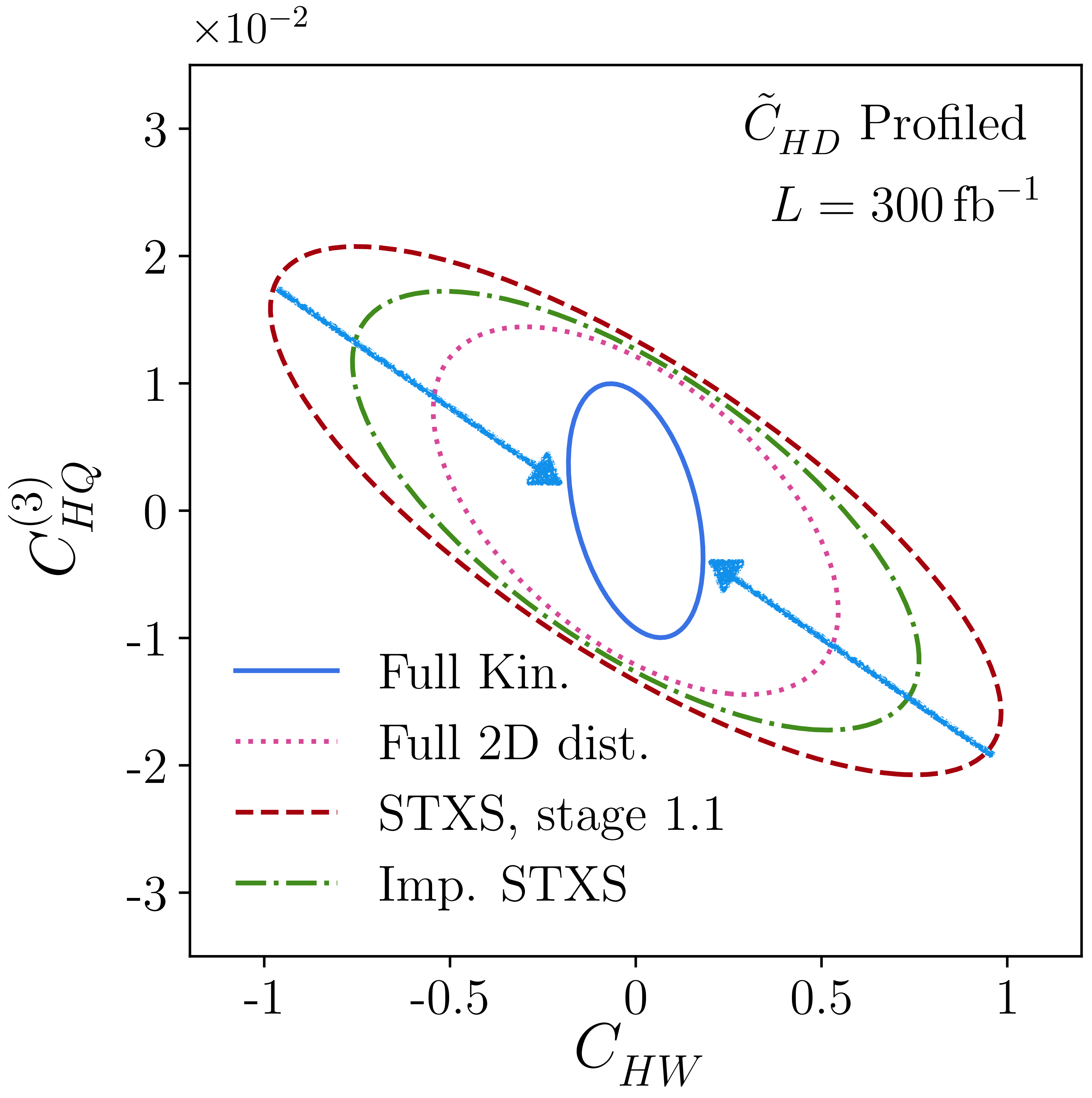
$$\tilde{\mathcal{O}}_{HD} = \mathcal{O}_{H\Box} - \frac{\mathcal{O}_{HD}}{4} = (\phi^\dagger\phi)\Box(\phi^\dagger\phi) - \frac{1}{4}(\phi^\dagger D^\mu\phi)^*(\phi^\dagger D_\mu\phi)$$

$$\mathcal{O}_{HW} = \phi^\dagger\phi W^a_{\mu\nu}W^{\mu\nu a}$$

$$\mathcal{O}^{(3)}_{Hq} = (\phi^\dagger i \overleftrightarrow{D}^a_\mu \phi)(\overline{Q}_L\sigma^a\gamma^\mu Q_L) \; ,$$

can we extract from $pp \to WH \to \ell\nu\, b\bar{b}$?

- Results: STXS are indeed sensitive to operators, adding a few more bins improve them, but a multivariate analysis is still stronger



An incomplete wrap-up of simulation-based inference methods

Method	Approximations	Upfront cost	Eval
Summary statistics:			
Likelihood for summary stats (standard histograms)	Reduction to summary stats	Fast	Fast
Approximate Bayesian Computation	Reduction to summary stats	Depends	Depends
Matrix elements:			
Matrix Element Method	Transfer fns	Fast	Slow
Optimal Observables	Transfer fns, optimal only locally	Fast	Slow
Neural networks:			
Neural likelihood	NN	Needs many samples	Fast
Neural posterior	NN	Needs many samples	Fast
Neural likelihood ratio	NN	Needs many samples	Fast
Neural networks + matrix elements:			
Neural likelihood (ratio) + gold mining (RASCAL etc)	NN	Needs less samples	Fast
Neural optimal observables (SALLY)	NN, optimal only locally	Needs less samples	Fast

Simulation-based inference references

Opinionated review
K. Cranmer, JB, G. Louppe:
"The frontier of simulation-based inference"
[1911.01429]

Do It Yourself (for LHC physics)
JB, F. Kling, I. Espejo, K. Cranmer:
"MadMiner: Machine learning—based inference for particle physics"
[CSBS, 1907.10621, https://github.com/diana-hep/madminer]

LHC HXSWG YR4 STXS
JB, S. Dawson, S. Homiller, F. Kling, T. Plehn:
"Benchmarking simplified template cross sections in WH production"
[JHEP, 1908.06980]

Use in Astro: Strong lensing
JB, S. Mishra-Sharma, J. Hermans, G. Louppe, K. Cranmer
"Mining for Dark Matter Substructure: Inferring subhalo population properties from strong lenses with machine learning"
[ApJ, 1909.02005]

Original works

JB, K. Cranmer, G. Louppe, J. Pavez:
"A guide to constraining Effective Field Theories with machine learning"
[PRD, 1805.00020]

JB, G. Louppe, J. Pavez, K. Cranmer:
"Mining gold from implicit models to improve likelihood-free inference"
[PNAS, 1805.12244]

Follow-up with incremental improvements
M. Stoye, JB, K. Cranmer, G. Louppe, J. Pavez:
"Likelihood-free inference with an improved cross-entropy estimator"
[NeurIPS workshop, 1808.00973]

PhyStat workshop on SBI in Munich

PHYSTAT-SBI 2024 - Simulation Based Inference in Fundamental Physics

15–17 May 2024
Max Planck Institute for Physics
Europe/Zurich timezone

Enter your search term

- Overview
- Call for Abstracts
- Timetable
- My Conference
  - My Contributions
- Registration
- Participant List
- Videoconference
- Practical Information

Fueled by the recent advances of Machine Learning in the last decade, a new breed of techniques have been developed to tackle statistical inference problems for "likelihood-free" cases, where it is possible to sample from the data-generating process (i.e. via stochastic simulators) but a closed form evaluation of the density is intractable.

This group of methods is known as "simulation-based inference" (SBI) or "likelihood-free inference" (LFI) and will be the dedicated topic of this PHYSTAT Workshop taking place from May 15th - May 17th 2024 at the Max-Planck Institute for Physics (MPP) in Garching near Munich.

PHYSTAT https://phystat.github.io/Website/) is a long-running workshop series that brings together statisticians, machine learning researchers and physicists to discuss shared topics and foster collaboration among the research communities.

Confirmed Speakers:

- Kyle Cranmer (U Wisconsin-Madison)
- Antoine Wehenkel (Apple)
- Gilles Louppe (U Liège)
- Laurence Levasseur (U Montréal)
- Ann Lee (Carnegie Mellon)
- Julia Linhart (INRIA)
- Noemi Montel (U of Amsterdam)
- Jakub Tomczak (Eindhoven)
- Christoph Weniger (GRAPPA)
- Alexander Held (U Wisconsin-Madison)

https://indico.cern.ch/event/1355601/overview

On-the-fly Event-by-Event Reweighting

Excerpts from

https://indico.cern.ch/event/1296757/

Abstract

Recently there has been rapid increase in the number of full statistical models (or "likelihoods") published by the experiments.

- Most are based on the HistFactory (pyhf) format and published in HEPData.
- This allows theorists and others to reproduce and combine measurements with the same gold standard as the internal experimental results.
- However, these are mainly from SUSY and exotics searches and
- working with EFTs is more complicated because quantum interference effects lead to changes in the signal template (via the dependence of the differential cross-sections and phase-space dependent selection efficiency on the EFT parameters).

In this talk I will propose a simple, lightweight framework that would extend current likelihood publishing to overcome these challenges and enable 'exact' EFT fits (i.e. with the same level of detail as the internal experimental fits and combinations).

https://indico.cern.ch/event/1296757/timetable/

Scope of this talk

The focus of this talk is about a practical statistical framework for doing EFT fits

- Emphasis is on statistical correctness, not optimality of observables, etc.
- Fit distributions in the data space (no unfolding)
  - Focusing on binned template fits with full systematic uncertainty treatment
- With some user-defined observables x (probably 1-D or 2-D)
  - This talk is **not** about what is a good observable
- Independent of which EFT operators, which basis, how many parameters, etc.

The framework lends itself well to publishing the full statistical model so that groups outside experiments can re-do fits, perform combinations, etc.

- So it addresses many of the motivations for unfolding, but its cleaner statistically

Morphing histograms vs. event-by-event reweighting

Morphing histograms (or fiducial cross-sections estimated with MC) has some subtle issues:

- Statistical fluctuations for bin probability (or fiducial cross-section) can lead to unphysical negative probabilities when morphing to a new value of α
- Efficiency and acceptance aren't constant for all events in a given bin of the observable x, so there is some (mild) approximation

The acceptance factors ϵ_{STXS} and $\epsilon_{\mathrm{diff.}}$, as well as the signal shape factors f_s, are derived under the assumption of SM Higgs boson kinematics. For interpretations of the measurements in physics models that significantly alter kinematic distributions, additional correction factors may be needed to account for changes in the acceptance and signal shape as a function of BSM model parameters. These are discussed when applicable in Sections 3 and 4.

ATLAS CONF Note

ATLAS-CONF-2023-052

Morphing histograms vs. event-by-event reweighting

Morphing histograms (or fiducial cross-sections estimated with MC) has some subtle issues:

- Statistical fluctuations for bin probability (or fiducial cross-section) can lead to unphysical negative probabilities when morphing to a new value of α
- Efficiency and acceptance aren't constant for all events in a given bin of the observable x, so there is some (mild) approximation

However, event-by-event reweighing based on morphing avoids these issues

- The event weights are always positive
- The weights are for a specific event (that either passes or fails selection criteria), so there is no approximation due to averaging efficiencies / acceptances for different types of events.

Idea 1: a model that builds histograms on-the-fly

For any fully simulated event with observable x_i and MC truth record z_i that was generated from EFT with parameters α_0 (e.g. the SM), we can reweight to a new EFT parameter point α with

$$w_i(\alpha) = \frac{d\sigma(\alpha)/dz}{d\sigma(\alpha_0)/dz}\Big|_{z_i}$$

- Similar to what we do with PDF reweighing.
- Kinematics don't change! Efficiency and acceptance are already included by selection on reconstructed quantities on event-by-event basis.
- The α-dependence of differential cross-sections can be computed using "morphing" equations or closely related approaches

Idea: For each value of α fill a signal histogram with set of weighted events $\{x_i, w_i(\alpha)\}$

- Can do this on-the-fly while doing the fit.
- It captures the α-dependence of efficiency and acceptance

Details: how to build histograms on-the-fly

Idea: For each value of α fill a signal histogram with set of weighted events $\{x_i, w_i(\alpha)\}$

- Can do this on-the-fly while doing the fit
- It captures the α-dependence of efficiency and acceptance

Details: To do this, the statistical model would need to maintain a **tiny database** that includes information for a set of simulated events:

- Store x_i (observed value of observable) and the coefficients needed to reweight event to a new point α. For example:
    - The differential cross-section (at truth-level) for set of basis points as implemented in **MadMiner**
    - The fully differential versions of the coefficients $A_j^{i,k'}$ in ATLAS-CONF-2023-052

It may be a bit slow, but its very flexible and avoids the problems mentioned above.

Idea 2: RECAST-like service for EFTs

Consider the case where ATLAS and CMS publish statistical models parametrized for some subset of operators in a specified EFT basis.

- **Sometime later** one wants to reinterpret the analysis for **a different set of operators** keeping the same event selection, breakdown of signal and control regions, observables, binning, etc.

RECAST is a framework for reinterpretations like this for BSM searches

- In general, this requires running new signal through the full MC simulation + reco + analysis chain. ATLAS is actually doing this with preserved analysis workflows!

But for EFTs we can simply to **reweight the existing fully simulated SM events** (doesn't require running more simulation, reconstruction, etc.)

- The service could calculate the coefficients for the mini-database based on truth-level kinematics and export a new statistical model that implements the statistical model for those operators as describe above.

Conclusion

Recently there has been rapid increase in the number of full statistical models (or "likelihoods") published by the experiments — mainly for BSM searches and their reinterpretation.

- **Ironically, it's not being used much for EFTs. This should change!**
- It would allow theorists and others to reproduce and combine measurements with the same gold standard as the internal experimental results.

We will need to define new **specifications** for components of statistical models that describe the details for how distributions of observables depend on EFT parameters including interference effects

- This is already very mature, but we should make the specifications concrete and then **implement** them in public tools
- Approaches based on **event-by-event reweighting** and **on-the-fly creation of histograms** have some nice properties and should be explored

Finally, we have all the ingredients needed to create a **RECAST-like service for EFTs** that would allow us to reweight fully simulated samples of events to **new EFT scenarios at some point in the future**