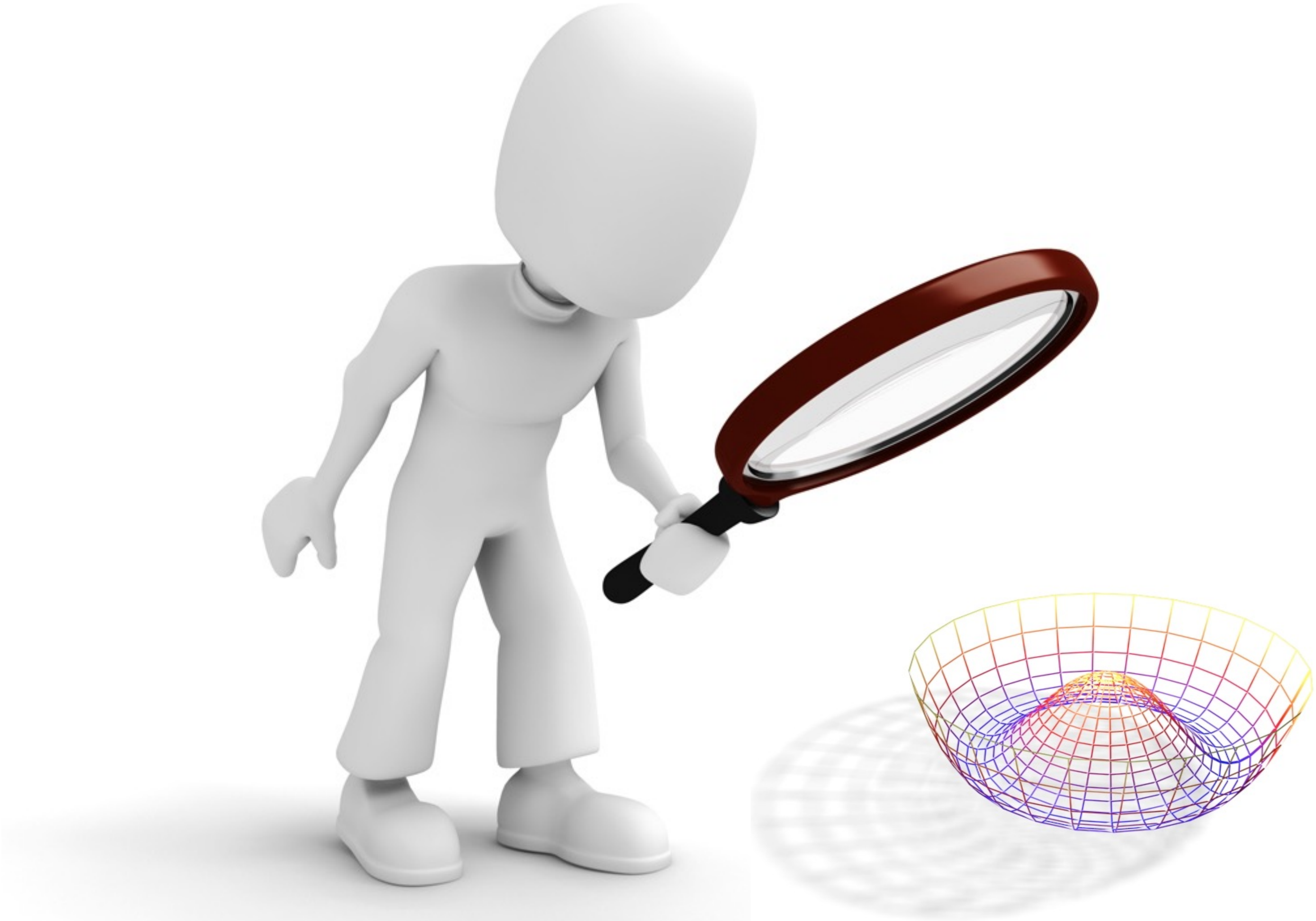
A PRACTICAL FRAMEWORK OF EFT FITS WITH PUBLISHED LIKELIHOODS



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

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

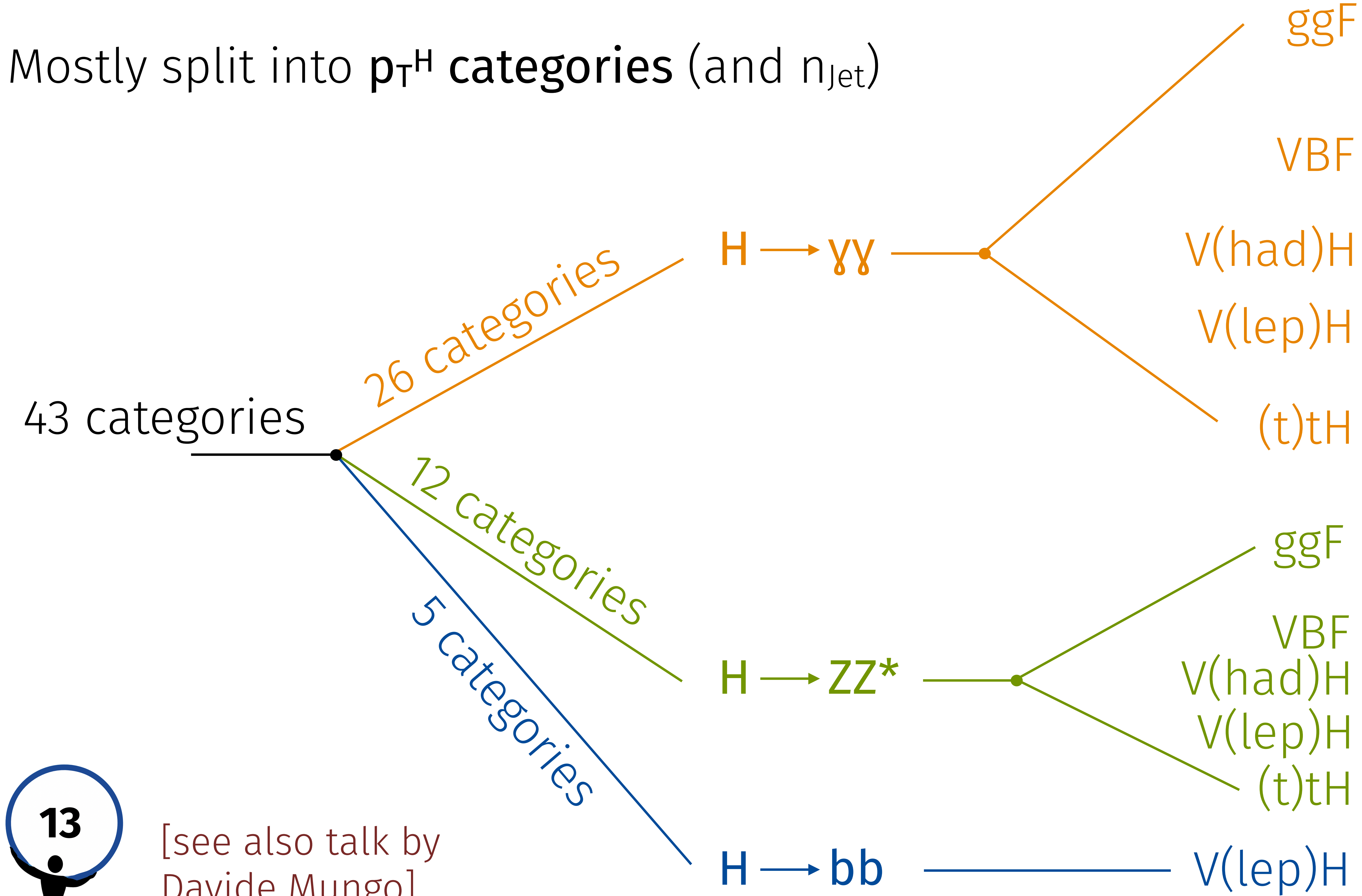
Example combined fits for EFTs

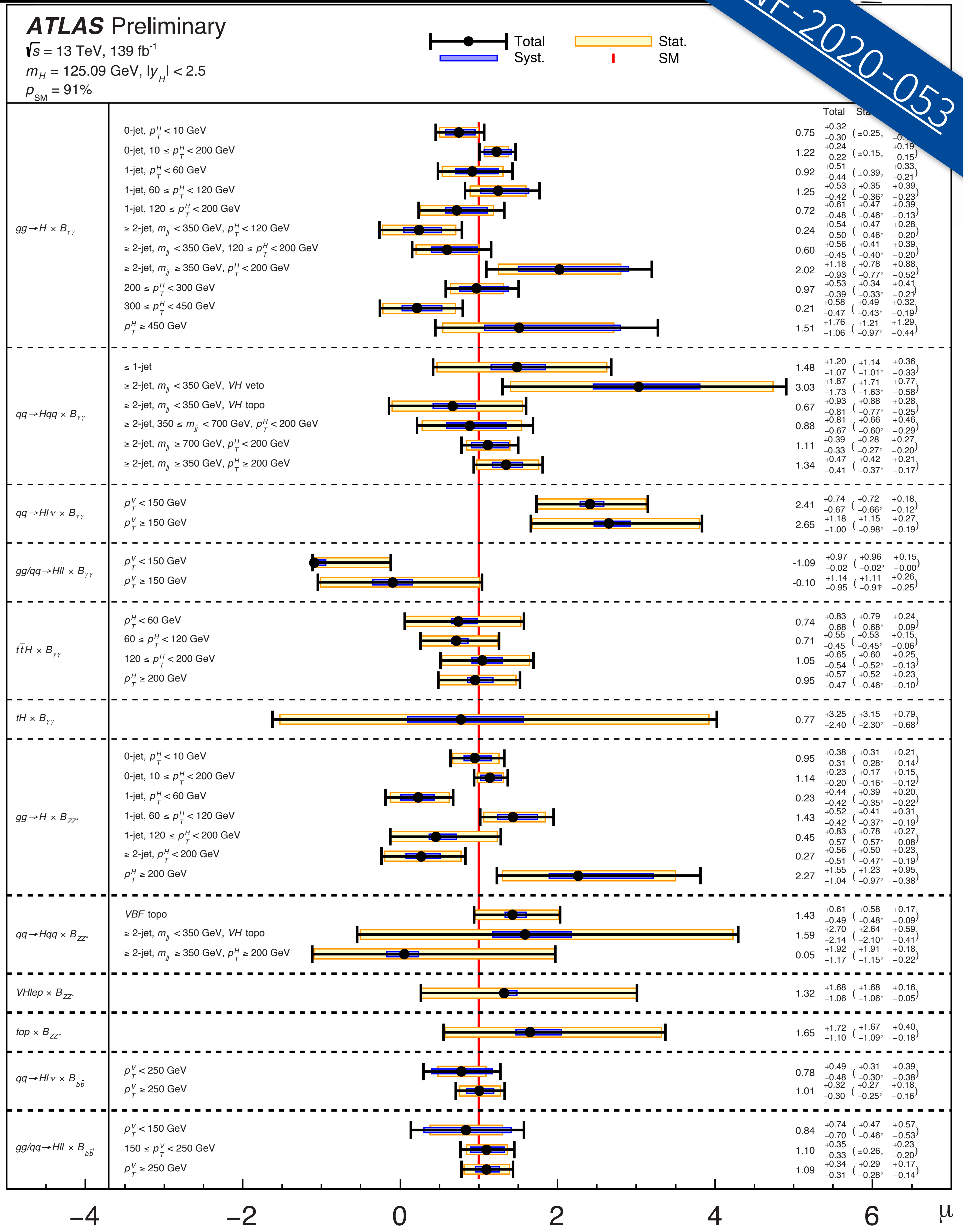
The STXS combination measurement

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]

A More Recent Example

ATLAS CONF Note

ATLAS-CONF-2023-052

26th August 2023

Interpretations of the ATLAS measurements of Higgs boson production and decay rates and differential cross-sections in pp collisions at $\sqrt{s} = 13$ TeV

2.3 Signal yield parametrization

In all analyses listed in Table 1, the likelihood function for each signal region k, with one or more bins r, is modeled as

$$L(\boldsymbol{N_k}|\boldsymbol{\mu},\boldsymbol{\theta}) \;=\; \prod_r \text{Poisson}\left(N_{k,r}|s_k(\boldsymbol{\mu},\boldsymbol{\theta})\cdot f_s^{k,r}(\boldsymbol{\theta}) \;+\; b_{k,r}(\boldsymbol{\theta})\right), \qquad (1)$$

where $N_{k,r}$ is the observed event count of bin r in region k, s_k is the expected signal count in region k, $f_s^{k,r}$ is the expected fraction of the signal in region k that is contained in bin r, and $b_{k,r}$ represents the expected event count from background processes. The ensemble of *parameters of interest* $\boldsymbol{\mu}$ describes the Higgs boson signal normalisation, while $\boldsymbol{\theta}$ represents the set of *nuisance parameters* taking into account the systematic uncertainties that originate from theoretical and experimental sources, as well as additional degrees of freedom without prior constraints such as background yields or normalisations in some of the input channels. The global likelihood function is then the product of the likelihood functions for each signal region k and of Gaussian or log-normal probability density functions that constrain the nuisance parameters.

Depending on the level of detail implemented in each analysis, the signal yield parameters $\boldsymbol{\mu}$ can be indexed by Higgs boson production process (i), decay mode (X), and fiducial phase space region defined at the particle level (k'). Analysis region k, defined at the reconstruction level, is typically chosen to match the particle-level region k' as closely as possible, in order to reduce the extrapolation uncertainty. As reconstruction-level selections do not generally correspond exactly to particle-level regions, multiple particle-level regions will contribute to the signal yield s_k.

Two distinct signal parametrization strategies are followed for the measurements listed in Table 1 and reported in Eq. (2) and Eq. (3). For those labeled as 'STXS', the signal yield for region k is modeled as a scale factor $\mu_k^{i,k',X}$ applied to the SM Higgs boson production cross-section times branching ratio, for each Higgs boson production process i and decay X, in a fiducial region k' defined at the particle level. Alternatively, for analyses labeled as 'differential', the signal yield is modeled as a cross-section $\sigma_{\text{fid.}}^{k',X}$ describing the sum of all production processes, separately for each Higgs boson decay mode X and fiducial region k' defined at the particle level. The corresponding parametrizations of the signal yield s_k in terms of the parameters of interest $\boldsymbol{\mu_k} = \{\mu_k^{i,k',X}\}$ and $\boldsymbol{\sigma}_{\text{fid.}} = \{\sigma_{\text{fid.}}^{k',X}\}$ and of the nuisance parameters $\boldsymbol{\theta}$ are:

$$s_k^{\text{STXS}}(\boldsymbol{\mu_k},\boldsymbol{\theta}) \;=\; \mathcal{L}\times\sum_{i,k',X}\mu_k^{i,k',X}\times(\sigma\times B)_{\text{SM,(N(N))NLO}}^{i,k',X}(\boldsymbol{\theta})\times\epsilon_{\text{STXS},k}^{i,k',X}(\boldsymbol{\theta}), \qquad (2)$$

$$s_k^{\text{diff.}}(\boldsymbol{\sigma}_{\text{fid.}},\boldsymbol{\theta}) \;=\; \mathcal{L}\times\sum_{k',X}\sigma_{\text{fid.}}^{k',X}\times\epsilon_{\text{diff.},k}^{k',X}(\boldsymbol{\theta}), \qquad (3)$$

where $\mathcal{L}$ is the integrated luminosity and $(\sigma\times B)_{\text{SM,(N(N))NLO}}^{i,k',X}$ is the calculation, at the highest available order, of the SM Higgs boson cross-section for the production process i in particle-level region k' multiplied by the SM Higgs boson branching ratio to the final state X. The factors $\epsilon_{\text{STXS},k}^{i,k',X}$ and $\epsilon_{\text{diff.},k}^{k',X}$ represent the products of acceptance times efficiency of the reconstruction-level region k for the particle-level fiducial phase space region k' and Higgs boson decay X (in production mode i for the STXS interpretation).

For each interpretation based on a particular model (SMEFT, 2HDM, or MSSM) with a vector of model parameters $\boldsymbol{\alpha}$, the original signal parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}_{\text{fid.}}$ are replaced with expressions that parameterise the model predictions, e.g. $\sigma_{\text{fid.}}^{k',X}\to\sigma_{\text{fid.}}^{k',X}(\boldsymbol{\alpha})$, so that the likelihood of Eq. (1) is directly expressed in terms of the parameters $\boldsymbol{\alpha}$. Then, constraints on these parameters can be directly inferred from the modified likelihood expression. The model-specific reparametrizations of the signal parameters are detailed in Sections 3 and 4.

The acceptance factors ϵ_{STXS} and $\epsilon_{\text{diff.}}$, 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.

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: Event-by-event reweighing as a function of EFT parameters based on truth-level kinematics has some advantages.

- Removes some approximations & provides an avenue to consider new EFT operators after the fact

Publishing Statistical Models

The first PhyStat

It was 23 years ago!

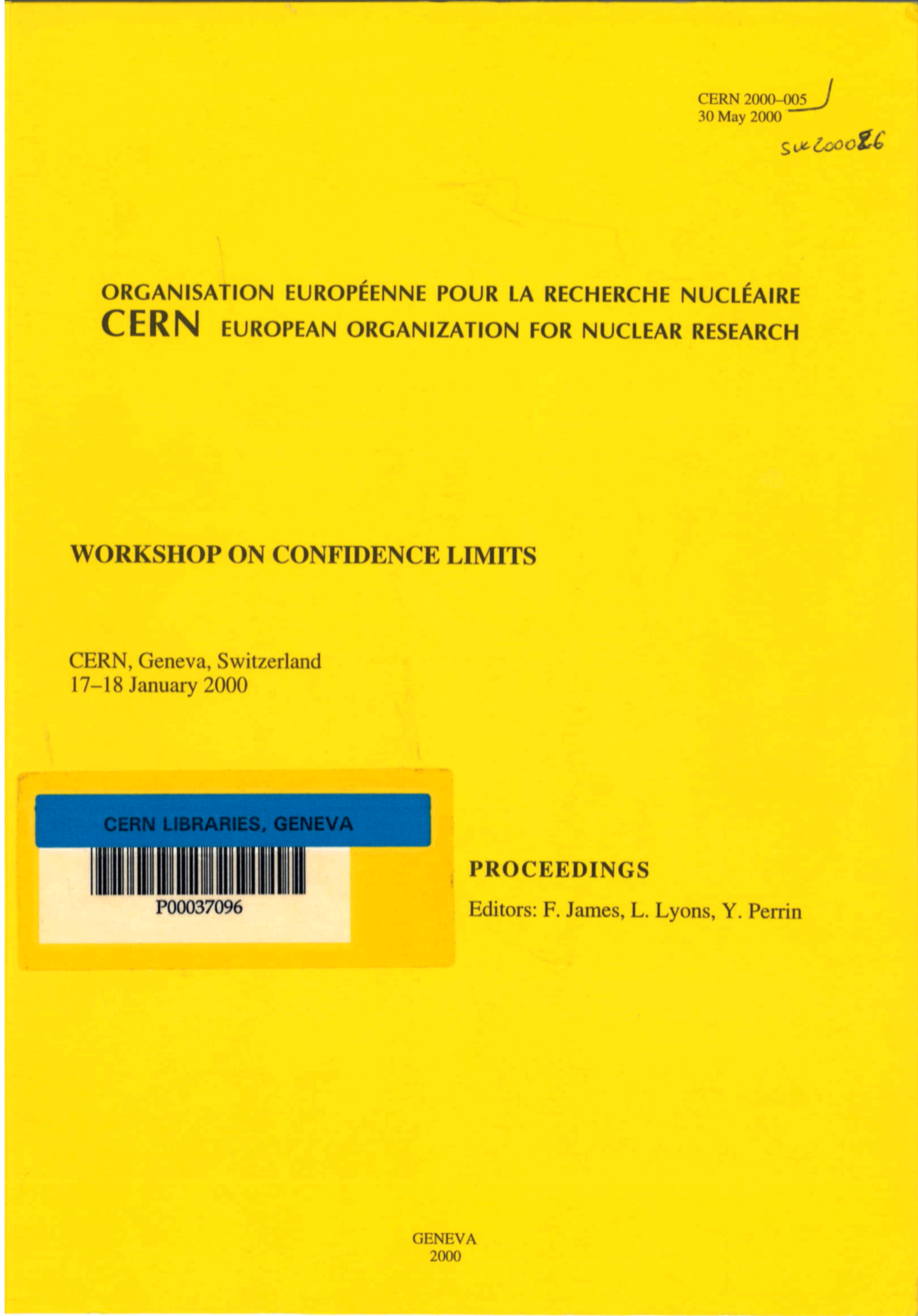
Massimo Corradi

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

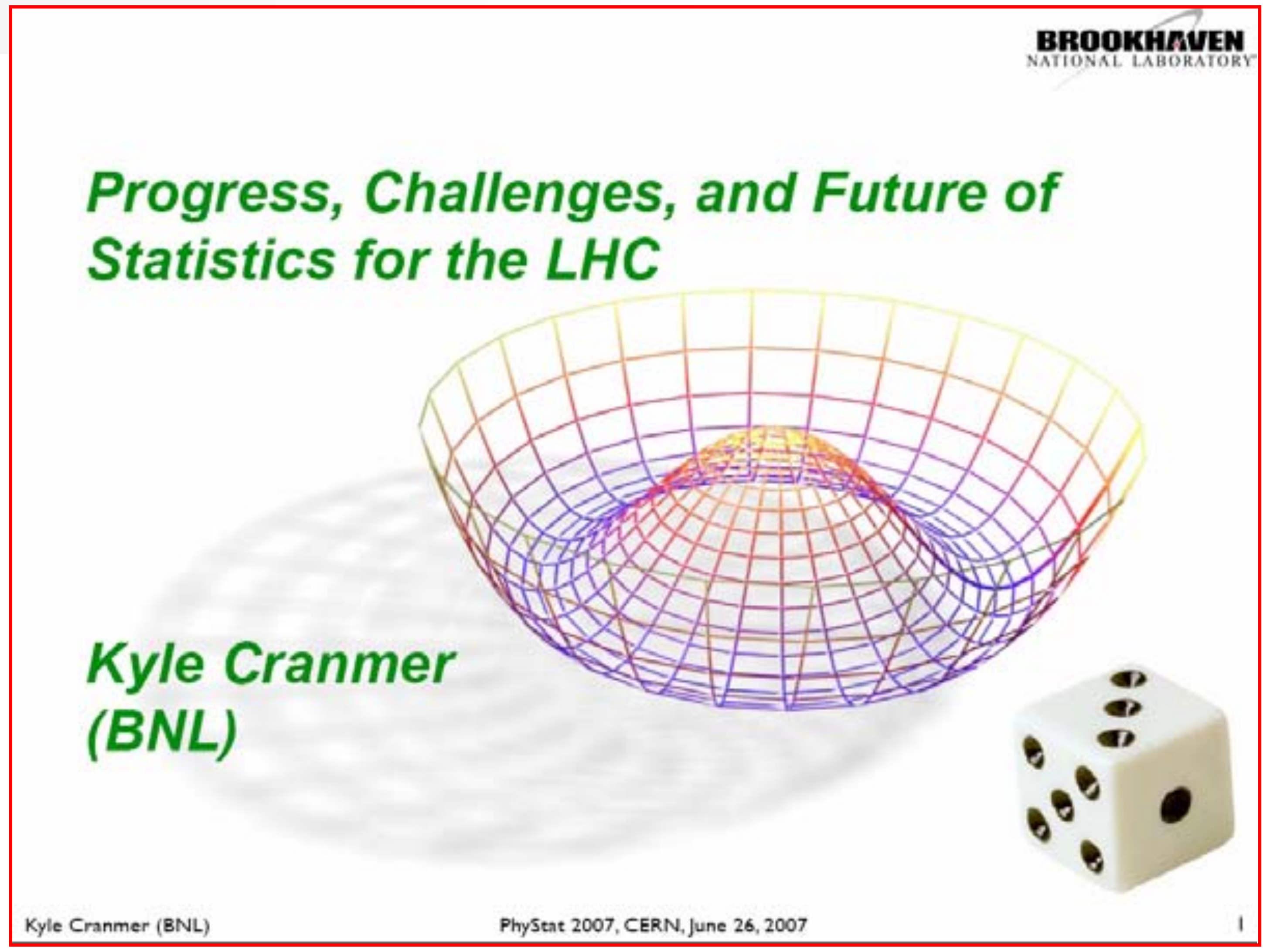
CERN, Geneva, Switzerland
17–18 January 2000

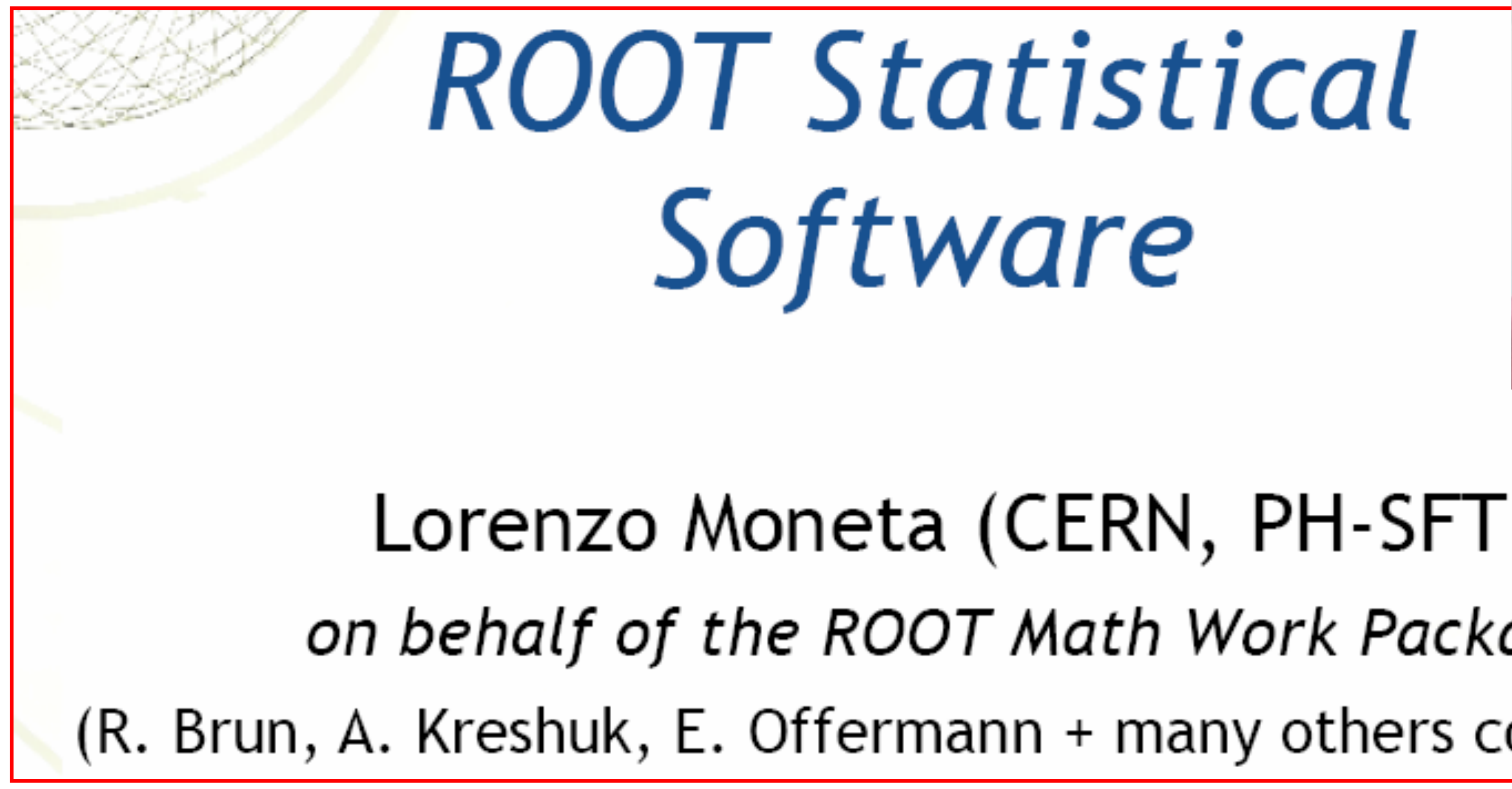
PROCEEDINGS

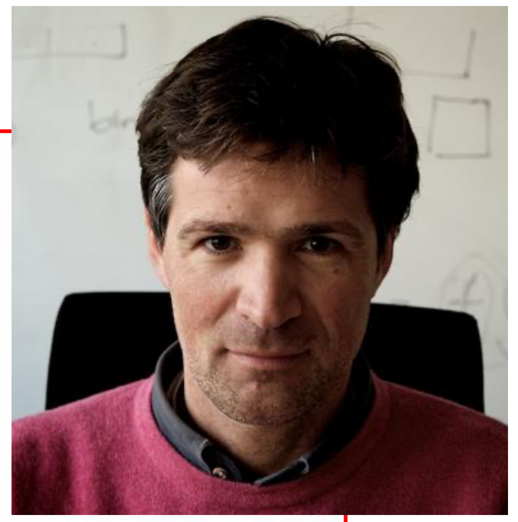
Editors: F. James, L. Lyons, Y. Perrin

GENEVA
2000

PhyStat 2007



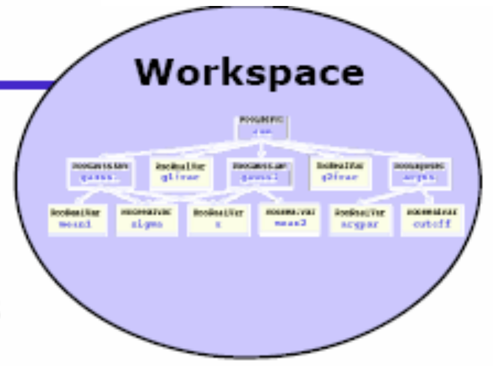




Statistics software for the LHC

Wouter Verkerke

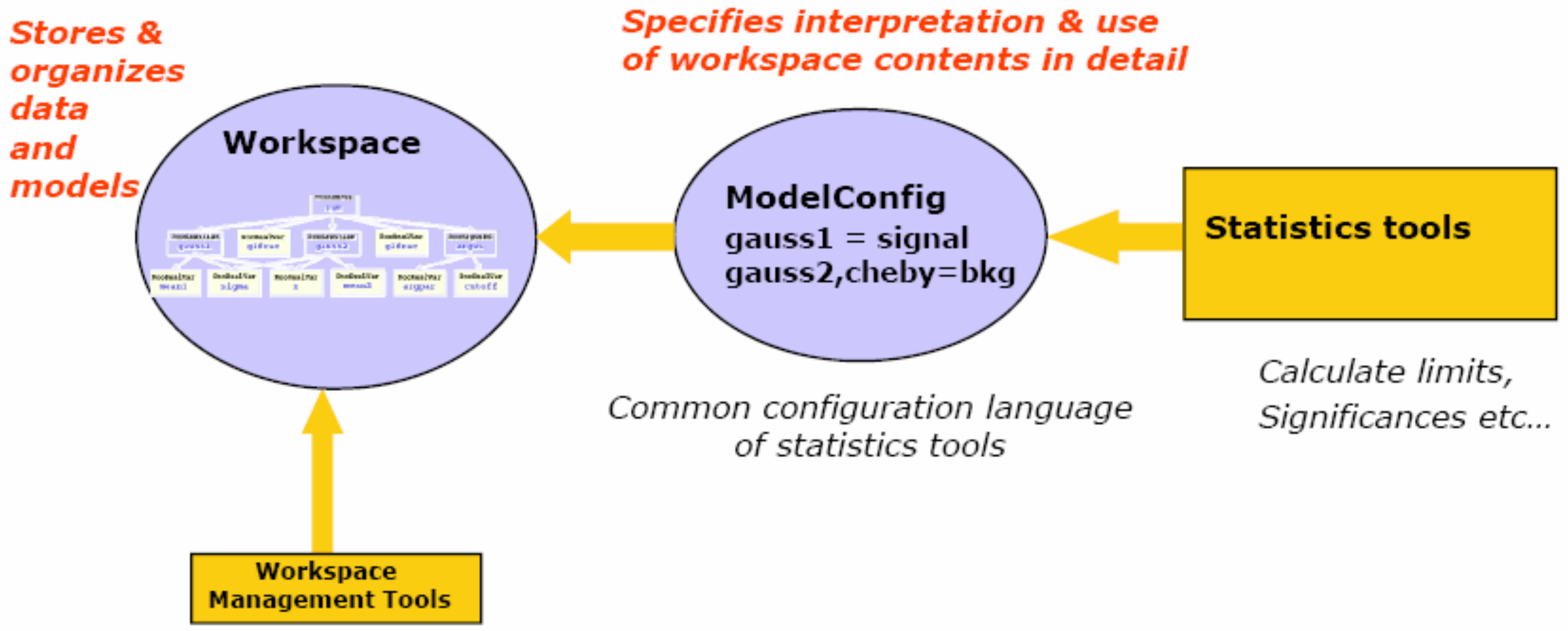
The Workspace as publication



- Now have functional RooWorkspace class that can contain
  - Probability density functions and its components
  - (Multiple) Datasets
  - Supporting interpretation information (RooModelConfig)
  - Can be stored in file with regular ROOT persistence
- **Ultimate publication of analysis...**
  - Full likelihood available for Bayesian analysis
  - Probability density function available for Frequentist analysis
  - Information can be easily extracted, combined etc...
  - Common format for sharing, combining of various physics results

Framework design & RooFit adaptations

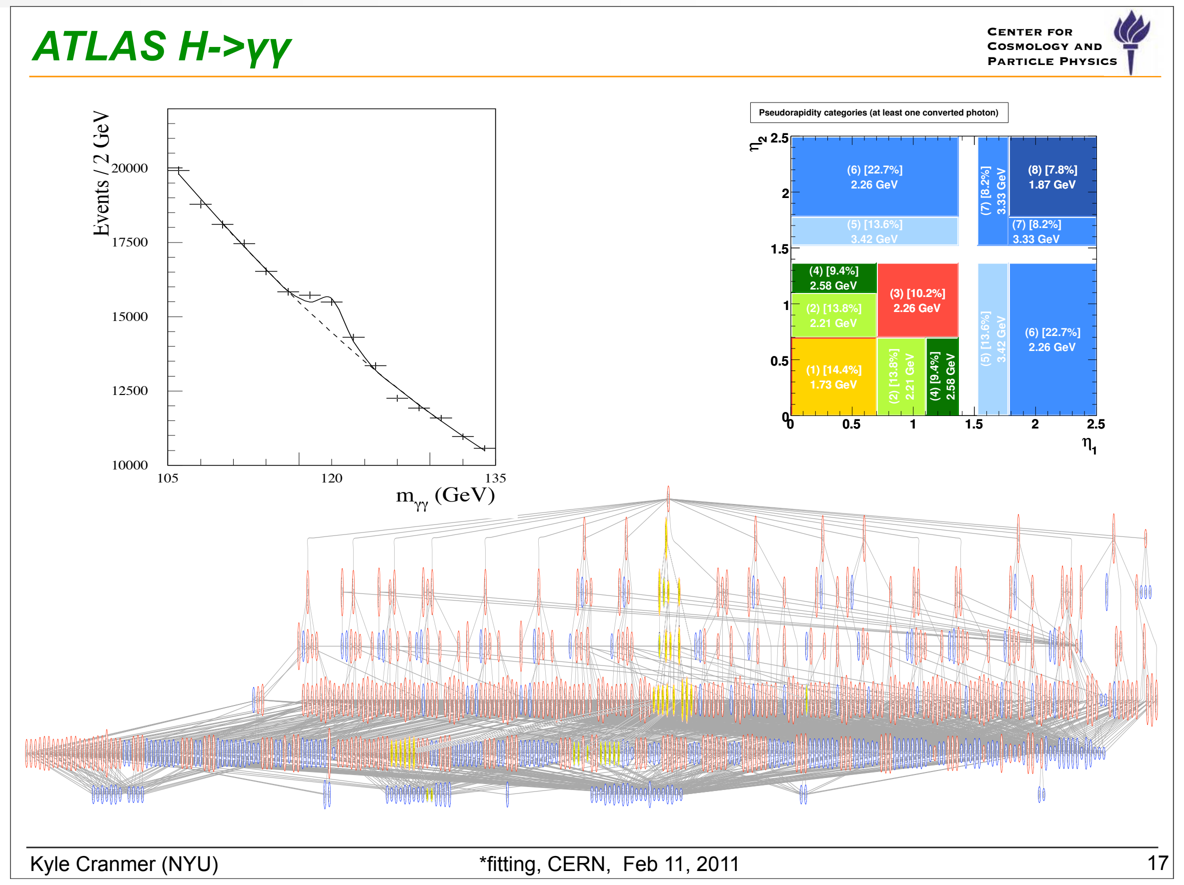
- Have had more meetings last 3 months to review RooFit lessons from BaBar
  - Kyle, Amir Farbin (ex-Babar), Frank Wrinklmeyer (ex-Babar), WV
  - Design for *WorkSpace* and *ModelConfig* concept in RooFit to interface with statistics tools

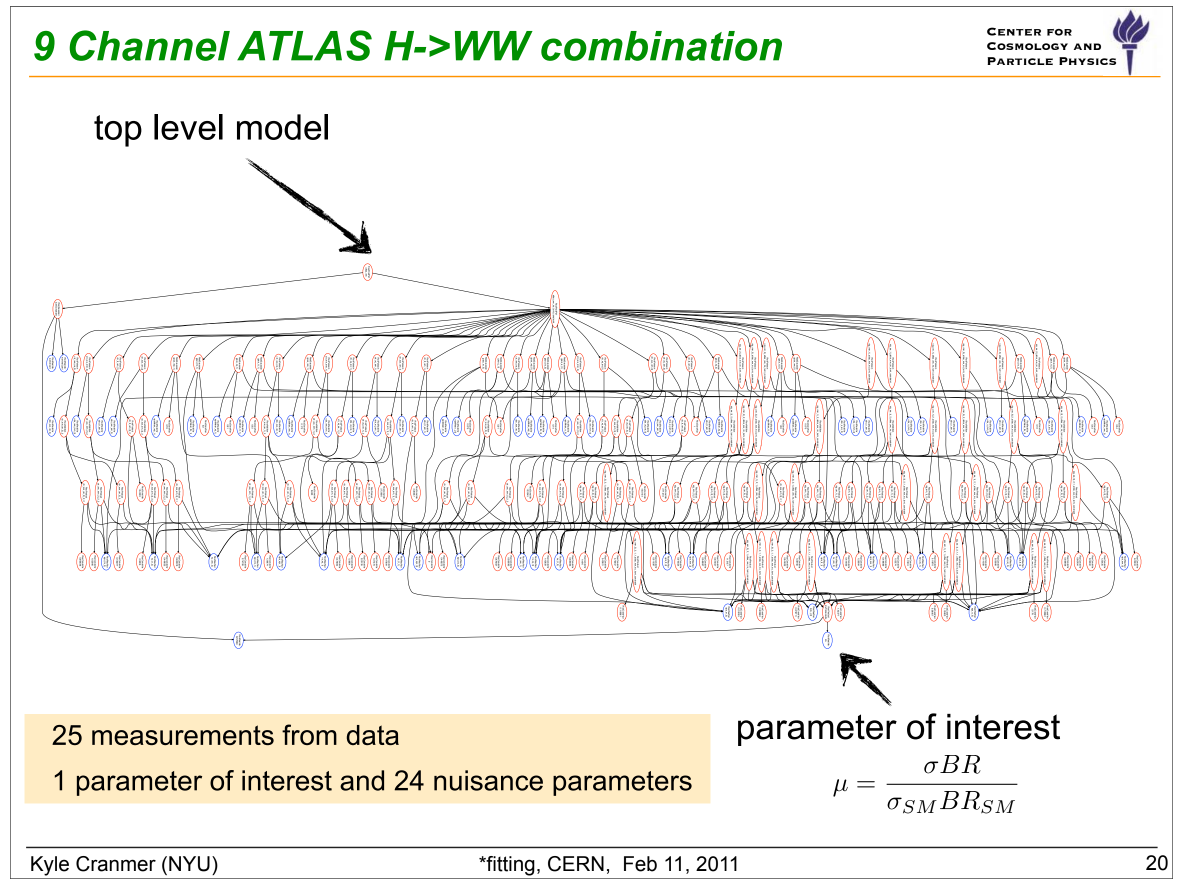


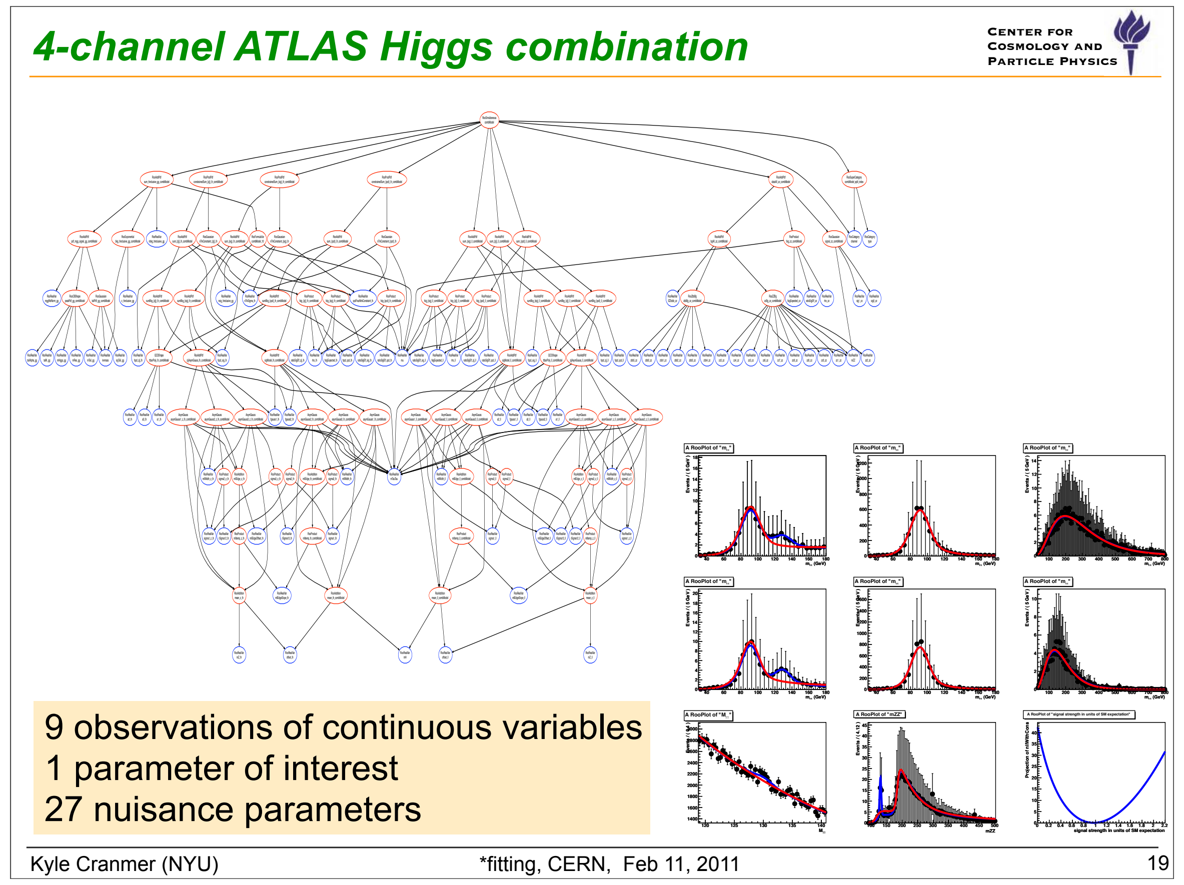
Bob Cousins Slide: http://indico.cern.ch/conferenceDisplay.py?confId=100458

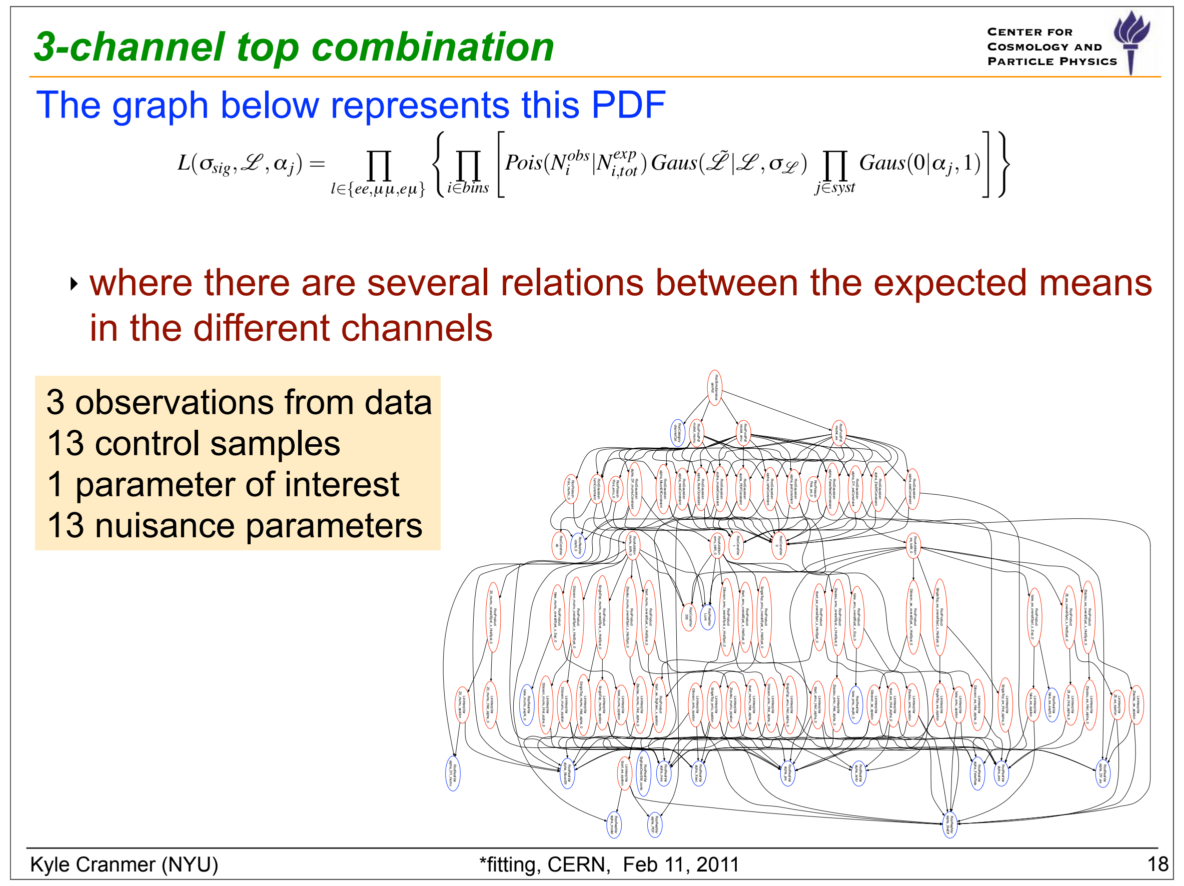
Early LHC examples (2011)

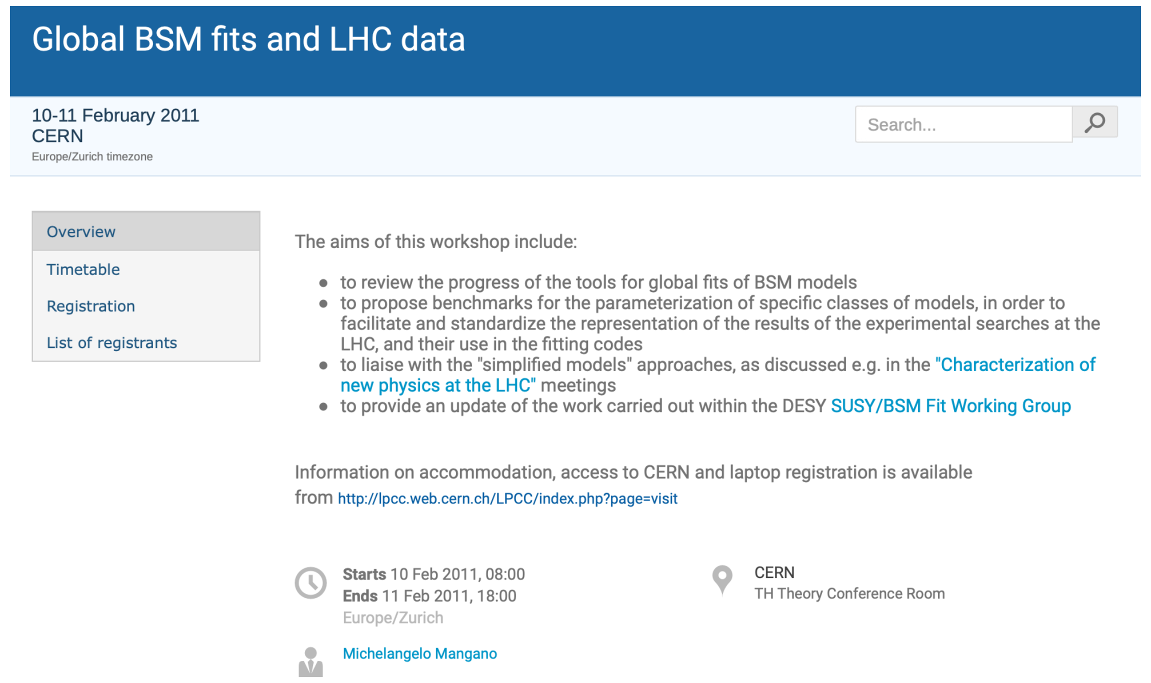
https://indico.cern.ch/event/118137/overview

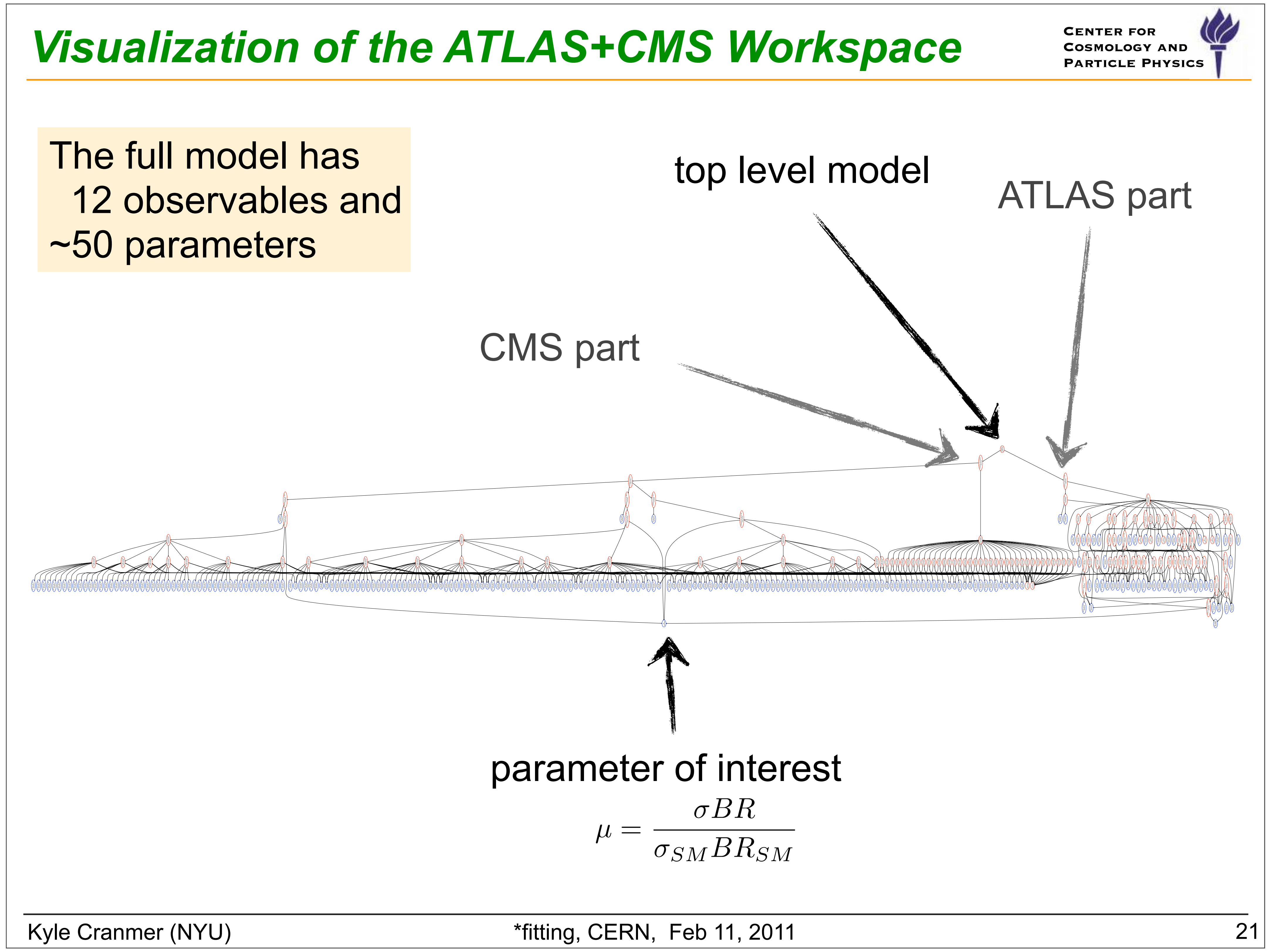




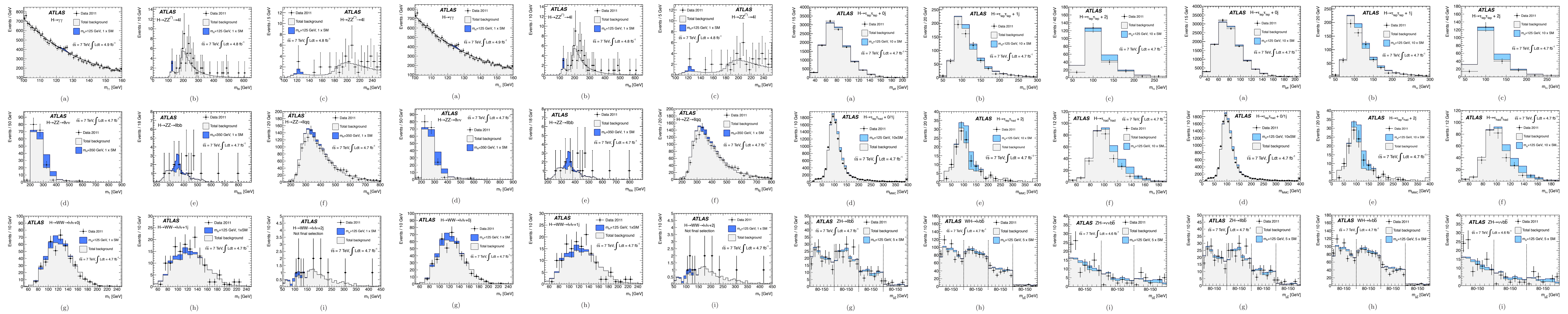


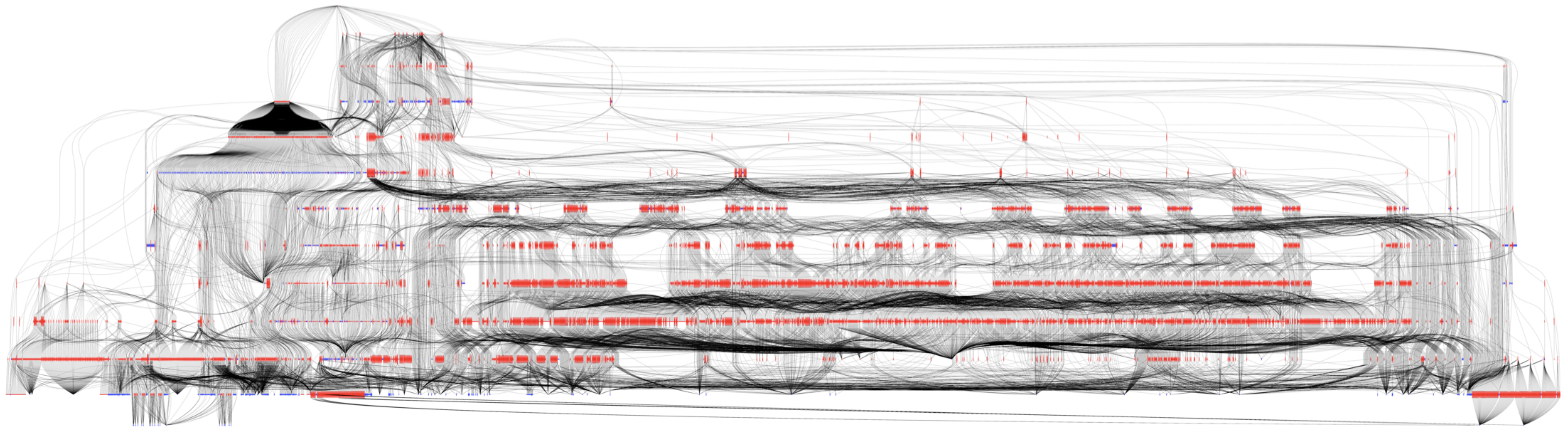






Combined fits for the Higgs discovery





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

Searches for New Physics: Les Houches Recommendations for the Presentation of LHC Results

S. Kraml[1], B.C. Allanach[2], M. Mangano[3], H.B. Prosper[4], S. Sekmen[3,4] (editors), C. Balazs[5], A. Barr[6], P. Bechtle[7], G. Belanger[8], A. Belyaev[9,10], K. Benslama[11], M. Campanelli[12], K. Cranmer[13], A. De Roeck[3], M.J. Dolan[14], T. Eifert[15], J.R. Ellis[16,3], M. Felcini[17], B. Fuks[18], D. Guadagnoli[8,19], J.F. Gunion[20], S. Heinemeyer[17], J. Hewett[15], A. Ismail[15], M. Kadastik[21], M. Krämer[22], J. Lykken[23] F. Mahmoudi[3,24], S.P. Martin[25,26,27], T. Rizzo[15], T. Robens[28], M. Tytgat[29], A. Weiler[30]

Why public likelihoods

- The statistical model of an experimental analysis provides the complete mathematical description of that analysis

  p(o|α) relating the observed quantities *o* to the parameters **α**

- Given the likelihood, all the standard statistical approaches are available for extracting information from it

- Essential information for any detailed interpretation of experimental results

  = determining the compatibility of the observations with theoretical predictions

Les Houches Recommandations (2012)

3b: When feasible, provide a mathematical description of the final likelihood function in which experimental data and parameters are clearly distinguished, either in the publication or the auxiliary information. Limits of validity should always be clearly specified.

3c: Additionally provide a digitized implementation of the likelihood that is consistent with the mathematical description.

arXiv:1203.2489

https://indico.cern.ch/event/957797/contributions/4026032/

Theorist rejoice

Now: full likelihoods !!

ATL-PHYS-PUB-2019-029 (05 Aug 2019)

- Plain-text serialisation of HistFactory workspaces, JSON format
  - Provides background estimates, changes under systematic variations, and observed data counts at the same fidelity as used in the experiment.

	Description	Modification	Constraint Term c_χ	Input
constrained	Uncorrelated Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Pois}\left(r_b = \sigma_b^{-2} \vert \rho_b = \sigma_b^{-2}\gamma_b\right)$	σ_b
	Correlated Shape	$\Delta_{scb}(\alpha) = f_p\left(\alpha \vert \Delta_{scb,\alpha=-1}, \Delta_{scb,\alpha=1}\right)$	$\text{Gaus}\,(a = 0 \vert \alpha, \sigma = 1)$	$\Delta_{scb,\alpha=\pm1}$
	Normalisation Unc.	$\kappa_{scb}(\alpha) = g_p\left(\alpha \vert \kappa_{scb,\alpha=-1}, \kappa_{scb,\alpha=1}\right)$	$\text{Gaus}\,(a = 0 \vert \alpha, \sigma = 1)$	$\kappa_{scb,\alpha=\pm1}$
	MC Stat. Uncertainty	$\kappa_{scb}(\gamma_b) = \gamma_b$	$\prod_b \text{Gaus}\left(a_{\gamma_b} = 1 \vert \gamma_b, \delta_b\right)$	$\delta_b^2 = \sum_s \delta_{sb}^2$
	Luminosity	$\kappa_{scb}(\lambda) = \lambda$	$\text{Gaus}\,(l = \lambda_0 \vert \lambda, \sigma_\lambda)$	$\lambda_0, \sigma_\lambda$
free	Normalisation	$\kappa_{scb}(\mu_b) = \mu_b$		
	Data-driven Shape	$\kappa_{scb}(\gamma_b) = \gamma_b$		

Rate modifications defined in HistFactory for bin *b*, sample *s*, channel *c*.

HEPData

Resources

gz File

Archive of full likelihoods in the HistFactory JSON format described in ATL-PHYS-PUB-2019-029 Provided are 3 statiscal models labeled RegionA RegionB and RegionC respectively each in their own sub-directory. For each model the background-only model is found i the file named 'BkgOnly.json' For each model a set of patches for various signal points is provided

Download

  - Usage: RooFit, **pyhf**
  - Target: long-term data/analysis preservation, reinterpretation purposes

So far available for 4/12 SUSY analyses with 139 fb^{-1}

SUSY-2018-31 (1908.03122)	multi-b sbottom: 2b+2H(bb)
SUSY-2018-04 (1911.06660)	stau search, 2 hadr. taus
SUSY-2019-08 (1909.09226)	1 lept. + H(bb), EW-ino
SUSY-2018-06 (1912.08479)	3 lept. EW-ino

https://indico.cern.ch/event/957797/contributions/4026032/

Theorist rejoice

Reinterpretation Forum Report 2020

"…. In fact, many of the data products discussed here, such as signal/background yields and correlations, are used by the various external reinterpretation packages to construct likelihoods. Whilst extremely useful, the likelihoods constructed from these products are however always only an approximation to the true underlying experimental likelihood. The reinterpretation workflow can be greatly facilitated and rendered much more precise if the original likelihood of the analysis is published in full. We strongly encourage the movement towards the publication of full experimental likelihoods wherever possible."

"ATLAS has recently started to do this using a JSON serialisation of the likelihood […] The provision of this full likelihood information is much appreciated and we hope that it will become a standard, as it **greatly improves the quality of any reinterpretation**."

Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2
arXiv:2003.07868, SciPost Phys. 9, 022 (2020)

S. Kraml - Feedback on use of public likelihoods - 24 Sep 2020 7

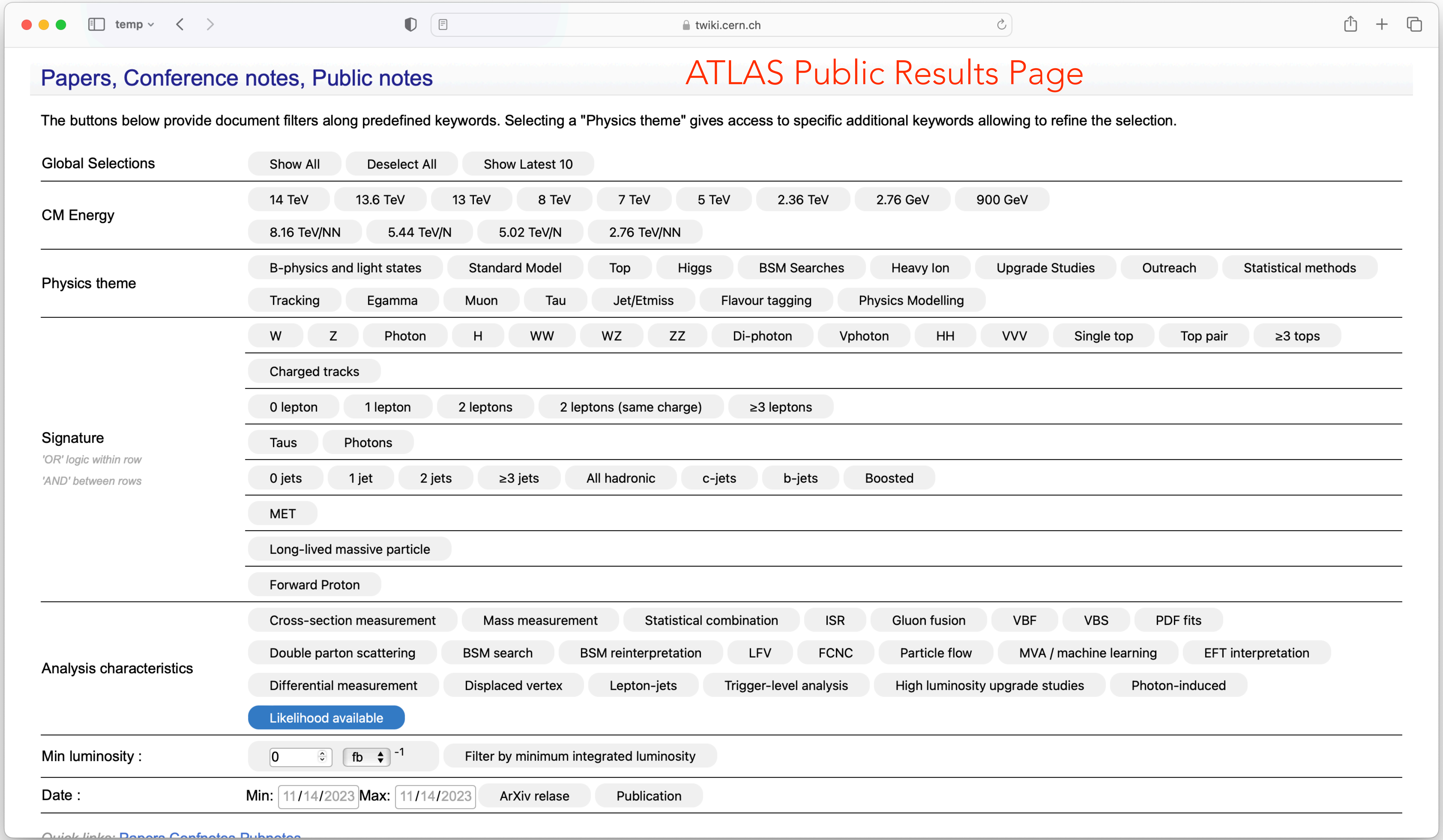
https://indico.cern.ch/event/957797/contributions/4026032/

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]

It's a reality

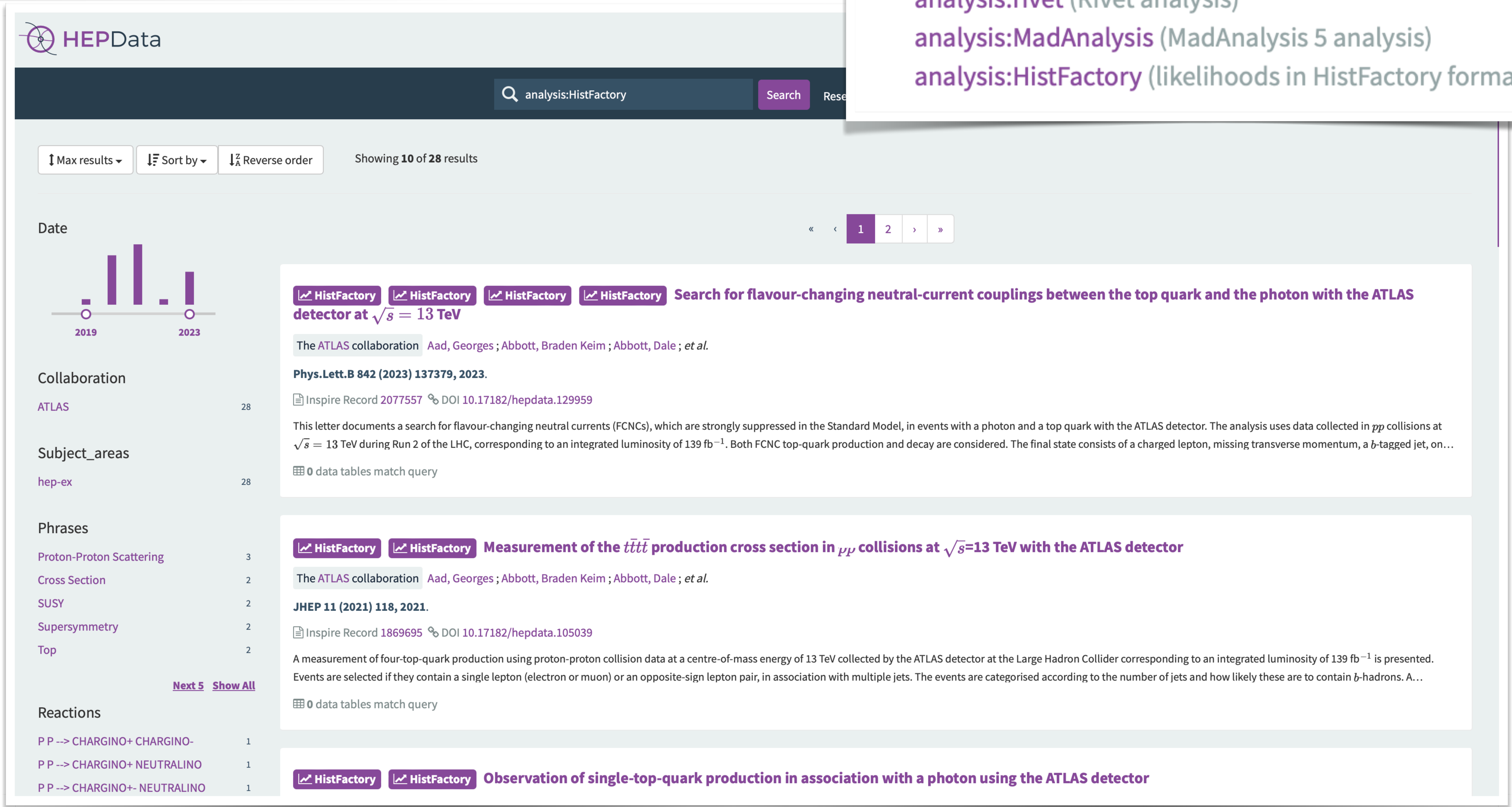


It's a reality

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)



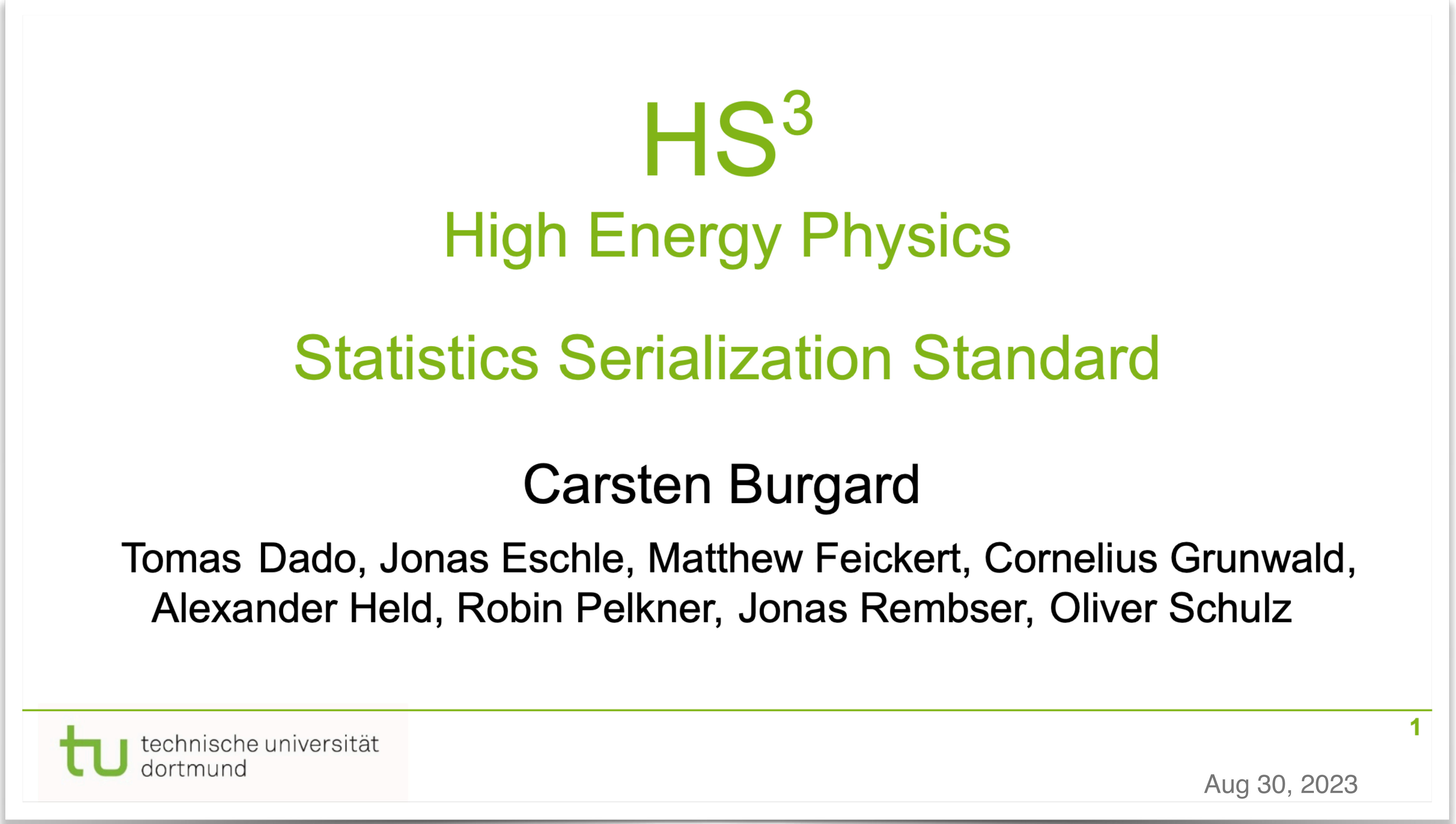
Browse and interact with published statistical models

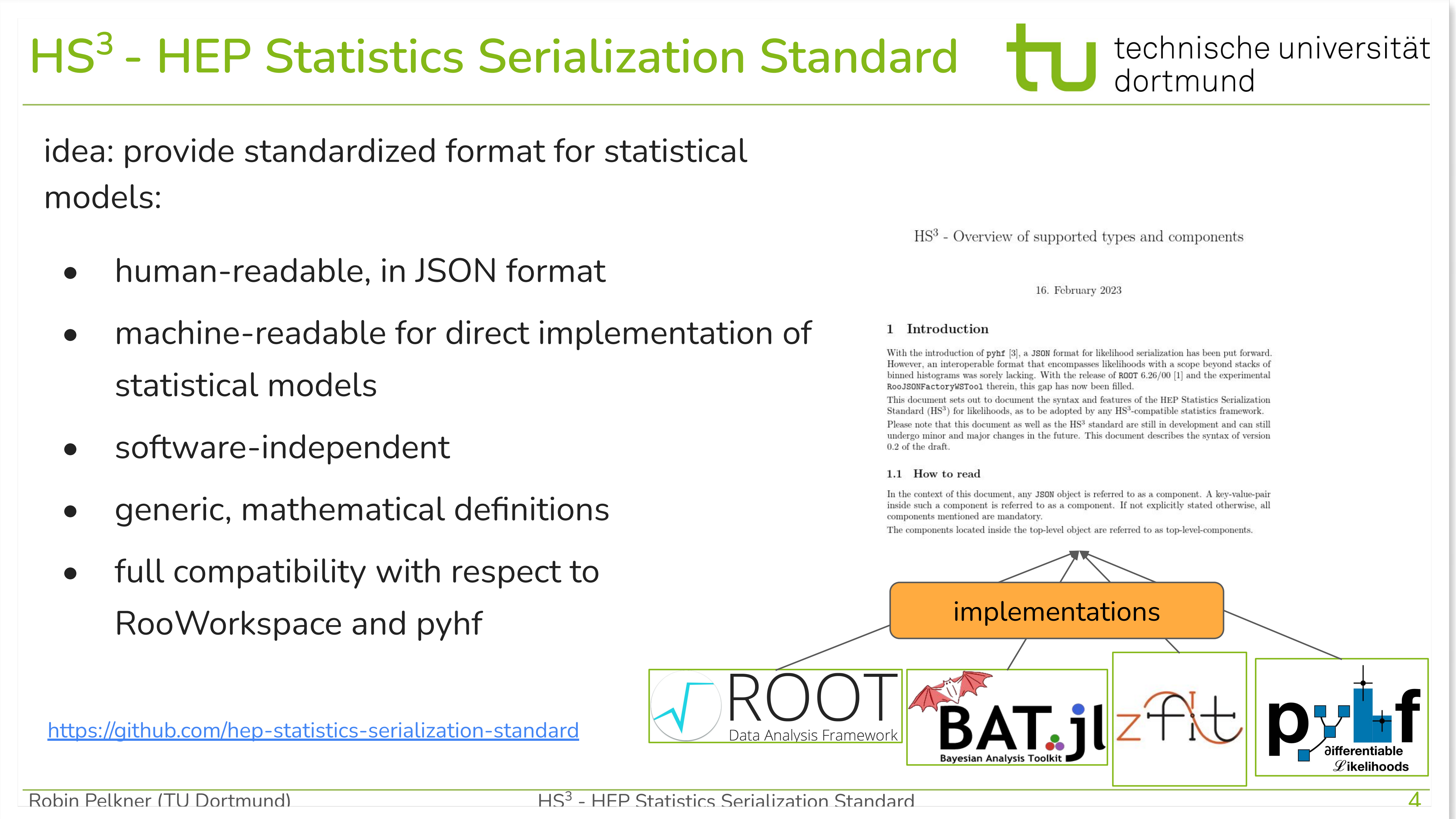
http://hepexplorer.net

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

EFT-Specific Model Specification

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

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

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

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

… 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

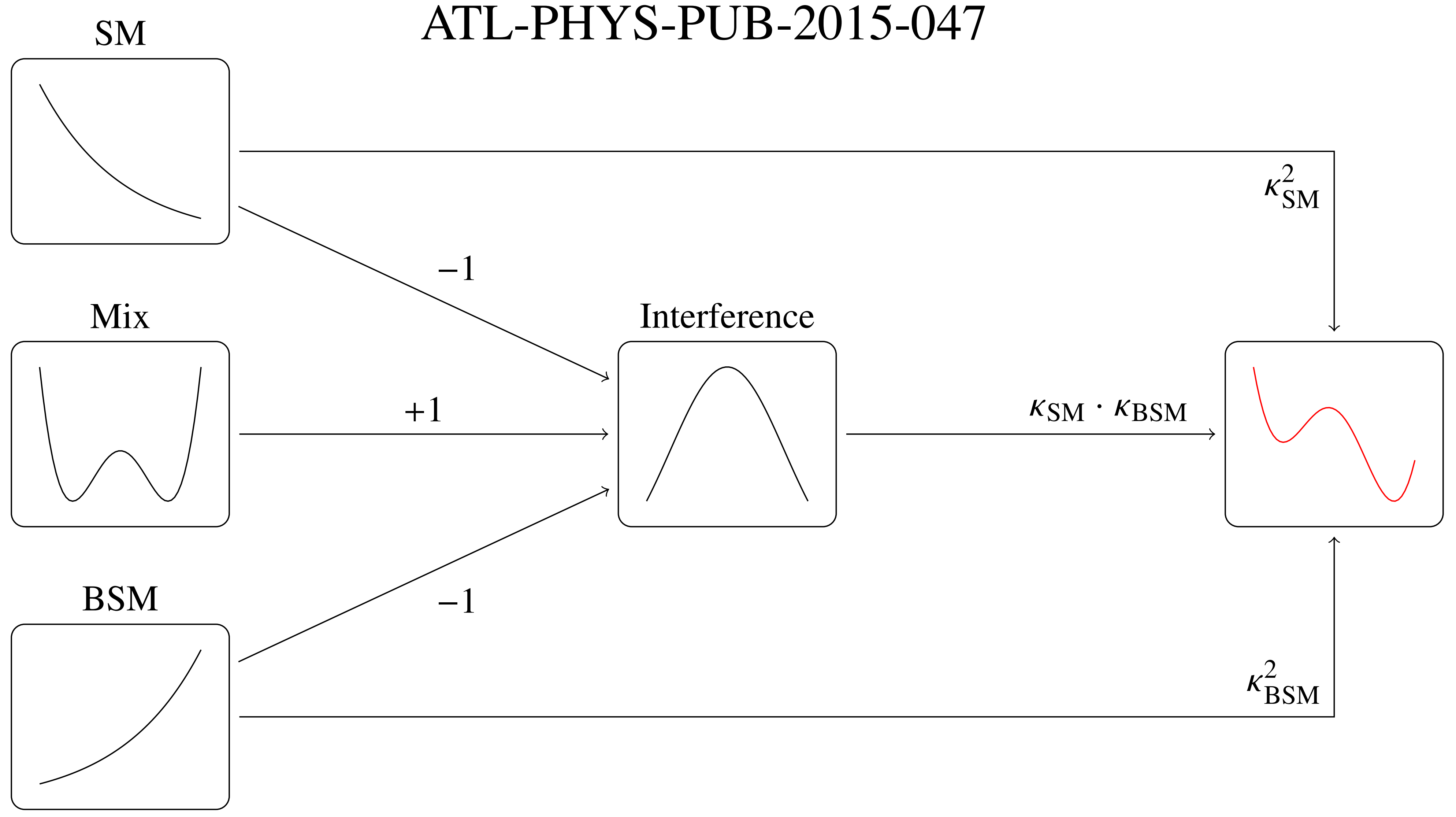
Main pieces:

- 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

ATL-PHYS-PUB-2015-047



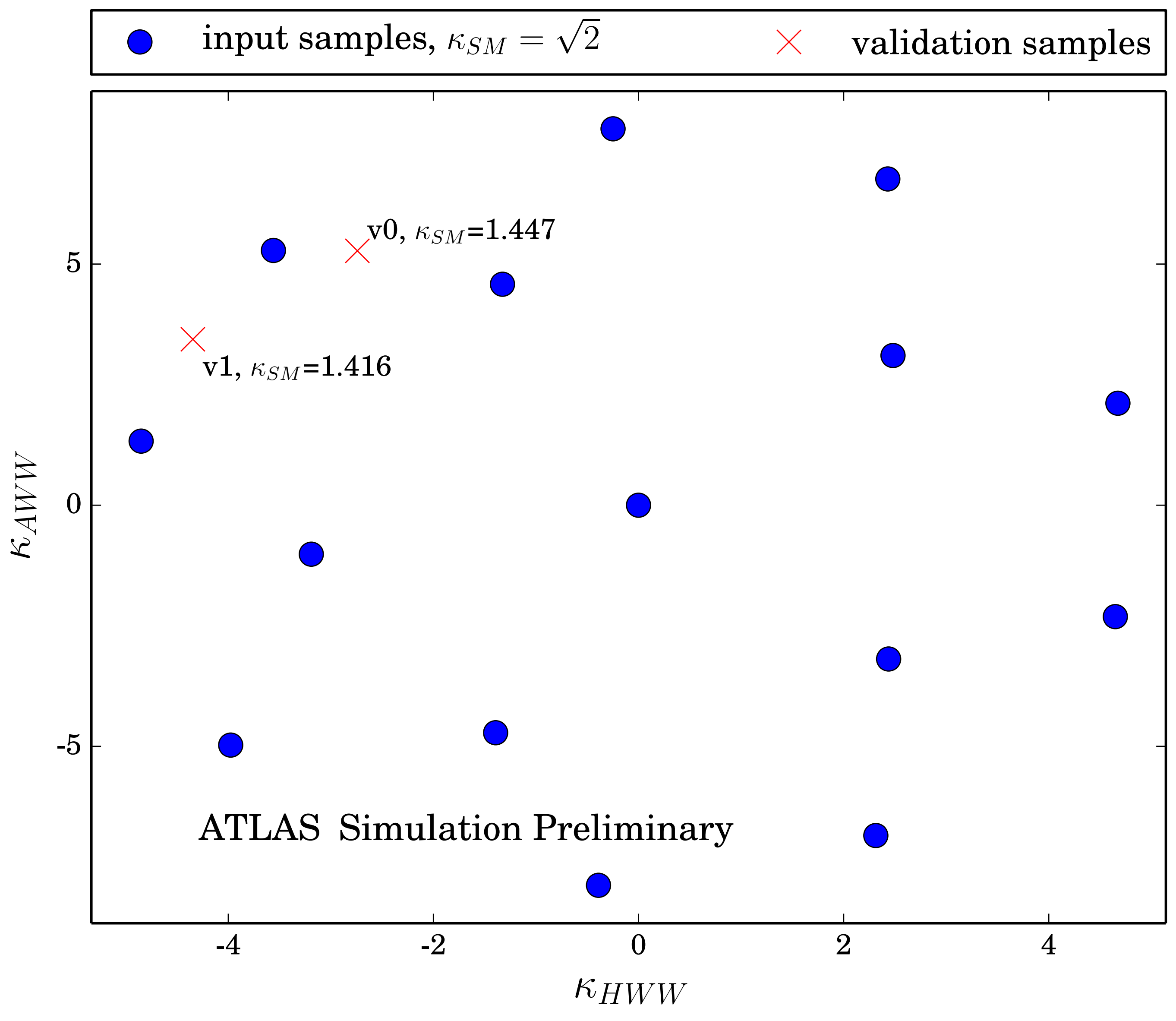
As one changes the parameters of the EFT, the distributions $p(x|\alpha)$ change due to interference.

But there is a trick:

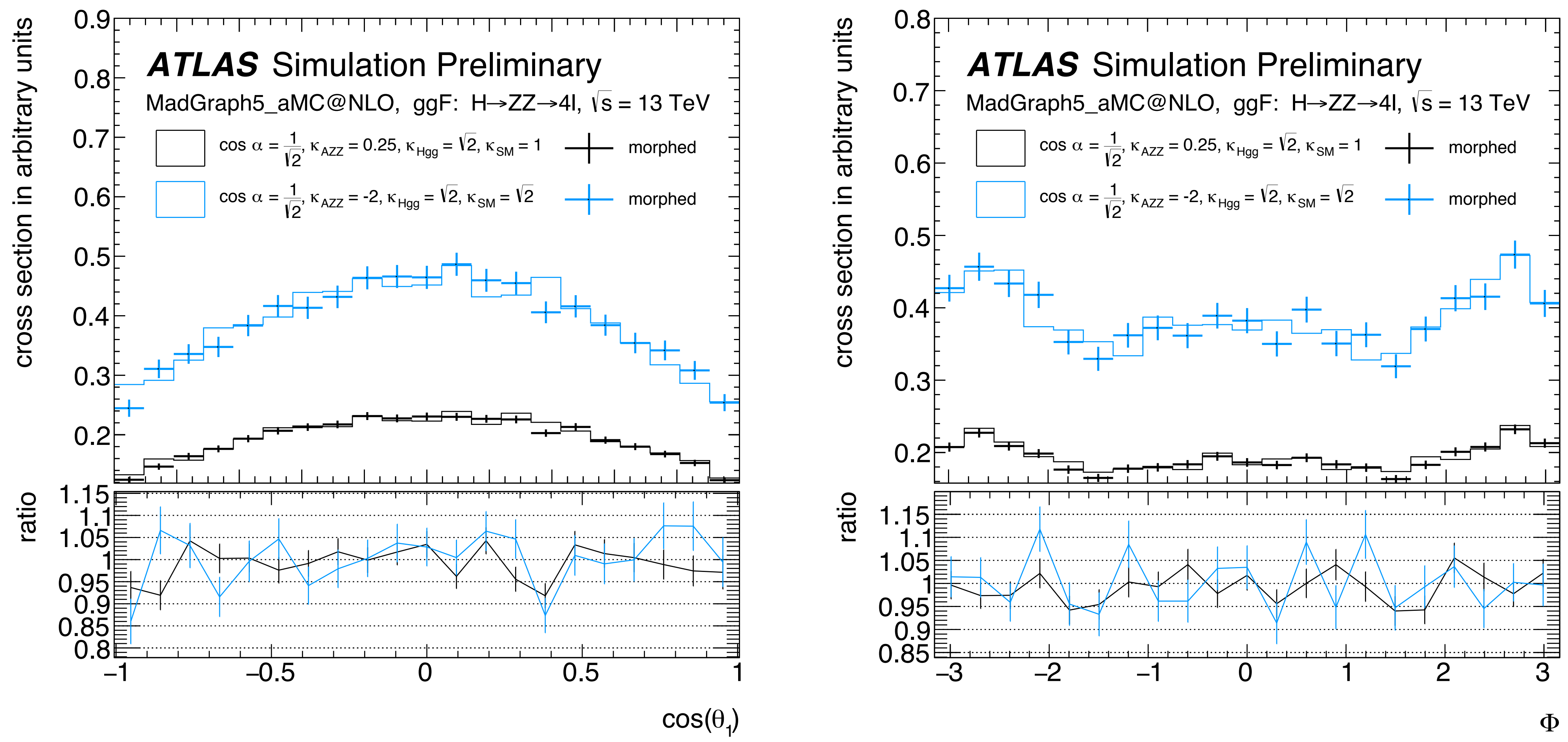
Simple example:

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

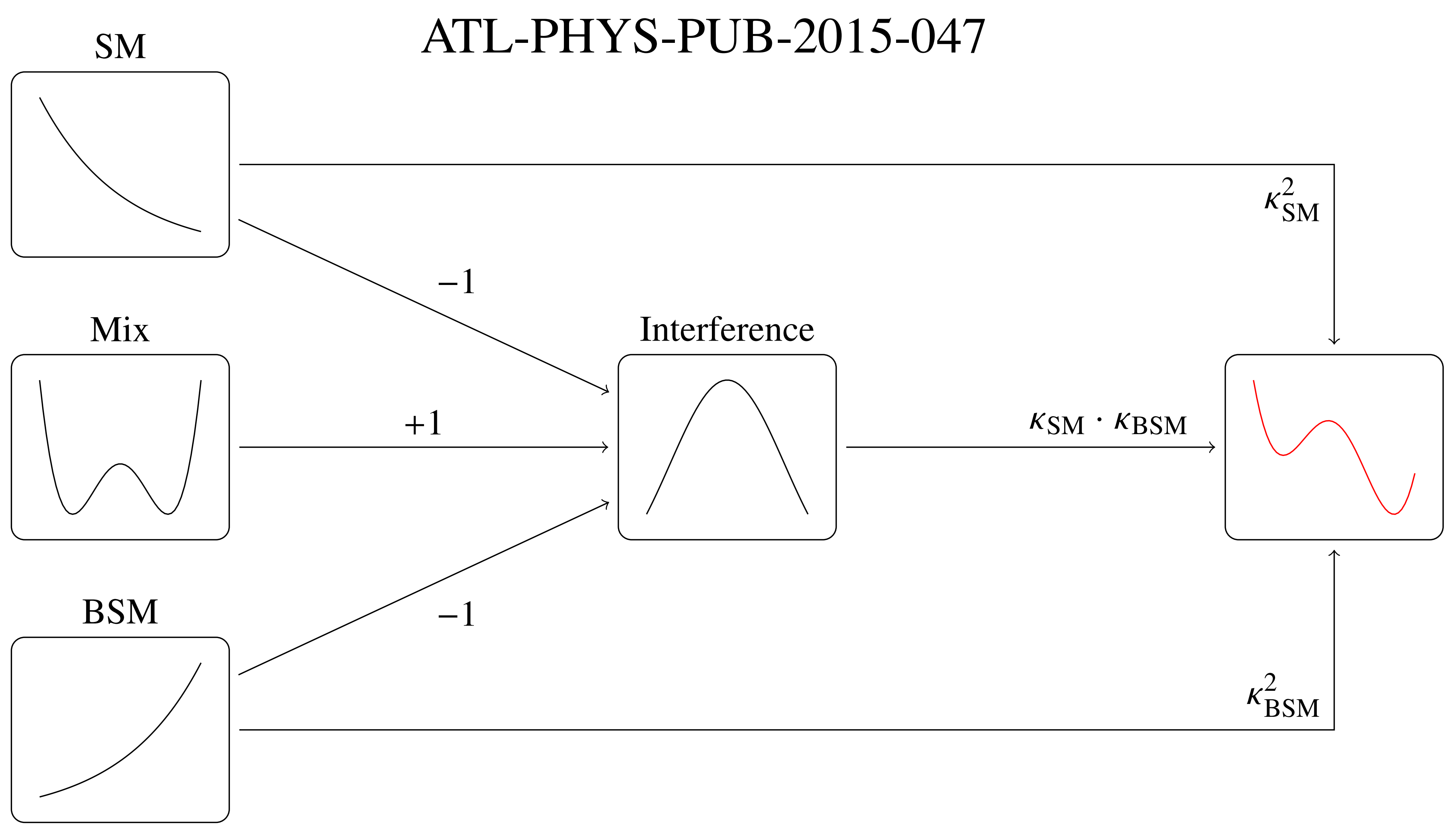
3-d vector space, distribution for any point in this space is linear mixture of distribution for 3 basis samples!



(real examples need more basis samples)



EFT “morphing” trick



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

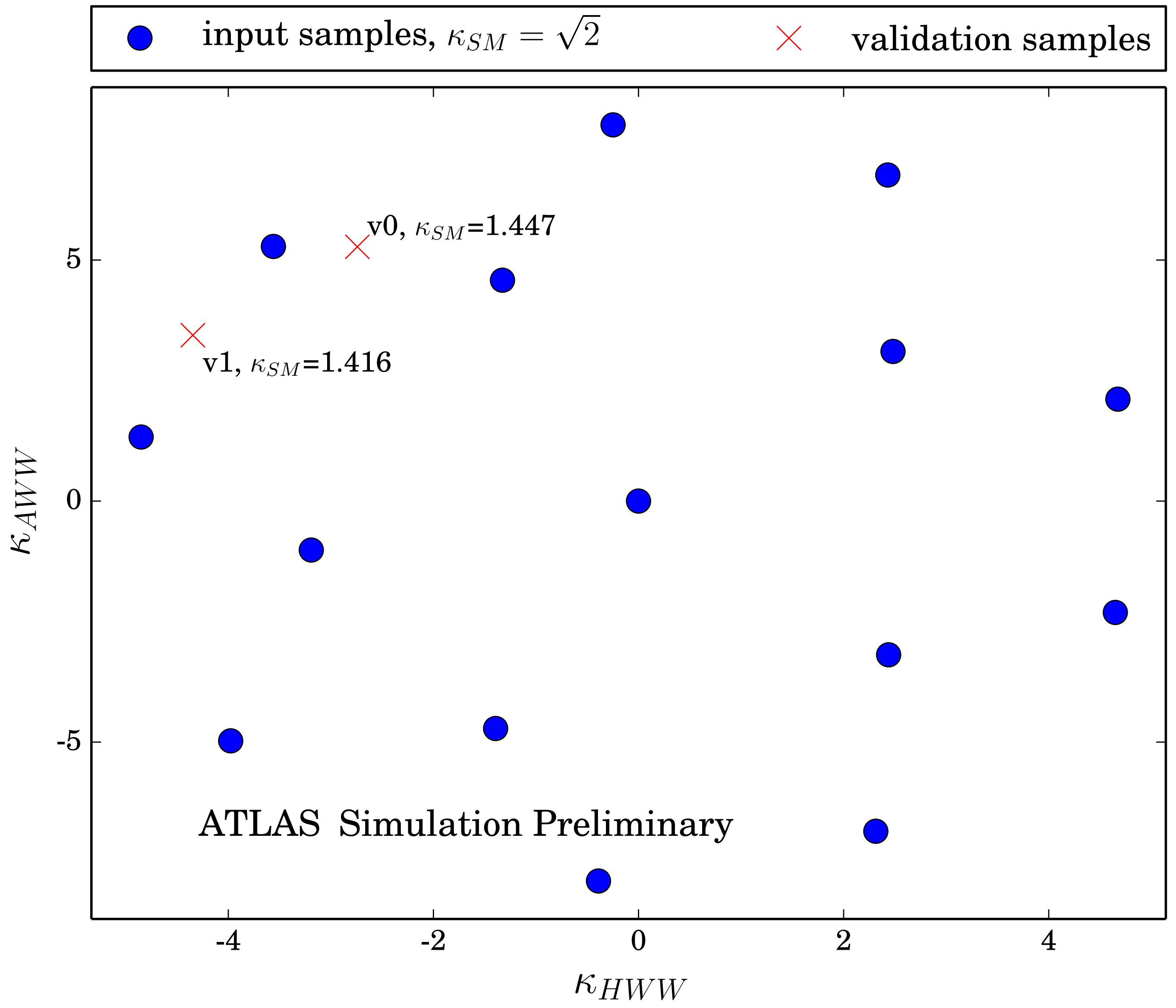
But there is a trick:

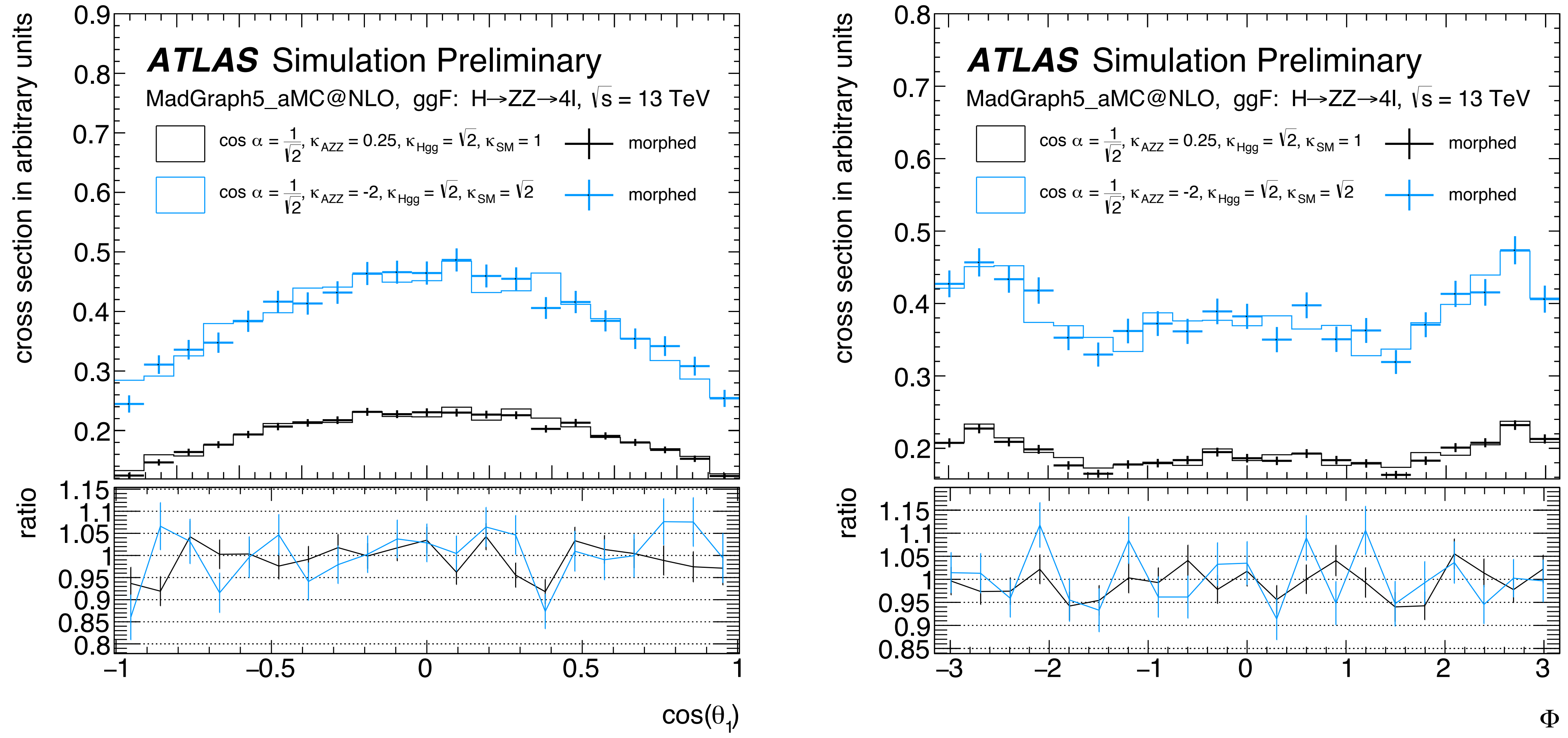
Simple example:

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

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

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

Express EFT as a mixture:

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

$w_c(\alpha)$ 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.

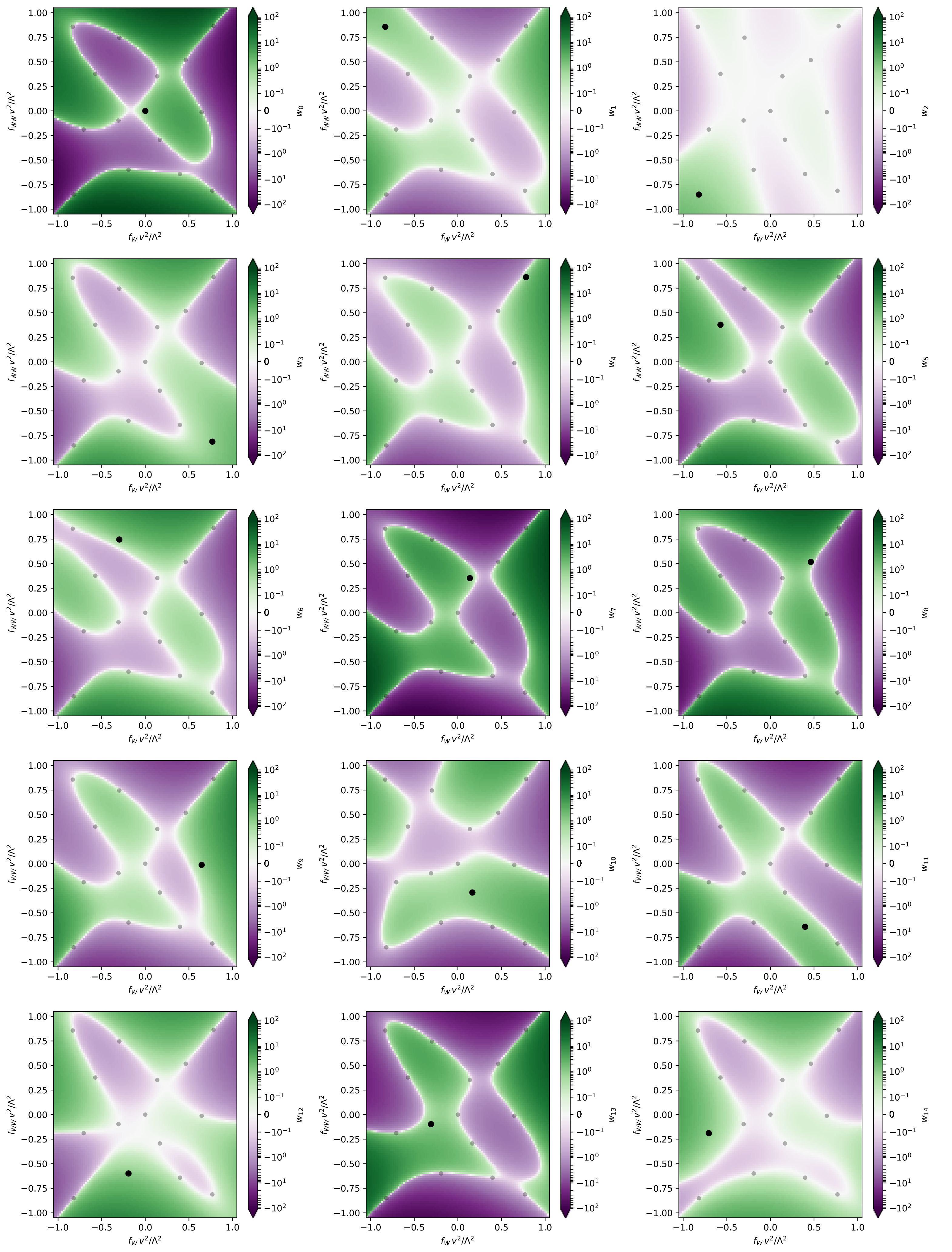


Figure 13: Morphing weights $w_i(\boldsymbol{\theta})$ for basis points distributed over the full relevant parameter space.

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

EFT "morphing" trick

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

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.

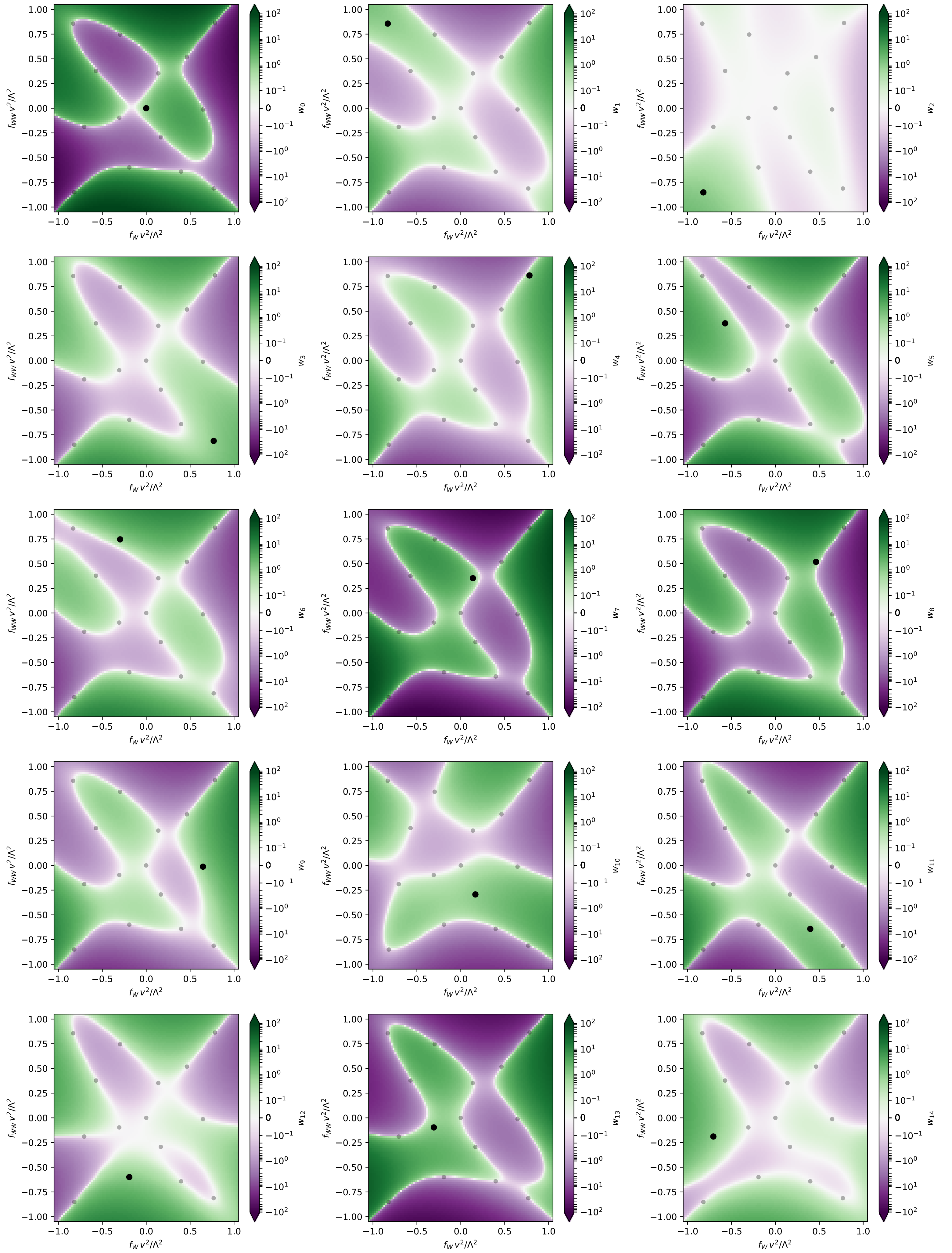


Figure 13: Morphing weights $w_i(\boldsymbol{\theta})$ for basis points distributed over the full relevant parameter space.

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

3.1.3 Cross-section calculation with linear and quadratic terms

The SMEFT prediction including the available terms proportional to Λ^{-4} is:

$$\begin{aligned}(\sigma\times B)^{i,k',H\to X}_{\text{SMEFT}} &= (\sigma\times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}}\left(1+\sum_j A_j^{\sigma_{i,k'}}c_j+\sum_{j,l\geq j}B_{jl}^{\sigma_{i,k'}}c_jc_l\right)\left(\frac{1+\sum_j A_j^{\Gamma^{H\to X}}c_j+\sum_{j,l\geq j}B_{jl}^{\Gamma^{H\to X}}c_jc_l}{1+\sum_j A_j^{\Gamma^H}c_j+\sum_{j,l\geq j}B_{jl}^{\Gamma^H}c_jc_l}\right),\\ &= (\sigma\times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}}\cdot\\ &\left(\frac{1+\sum_j\left(A_j^{\sigma_{i,k'}}+A_j^{\Gamma^{H\to X}}\right)c_j+\sum_{j,l\geq j}\left(B_{jl}^{\sigma_{i,k'}}+B_{jl}^{\Gamma^{H\to X}}+A_j^{\sigma_{i,k'}}A_l^{\Gamma^{H\to X}}+A_l^{\sigma_{i,k'}}A_j^{\Gamma^{H\to X}}\right)c_jc_l+O(\Lambda^{-6})}{1+\sum_j\left(A_j^{\Gamma^H}\right)c_j+\sum_{j,l\geq j}\left(B_{jl}^{\Gamma^H}\right)c_jc_l+O(\Lambda^{-6})}\right)\end{aligned} \quad (13)$$

3.1.2 Cross-section calculation with linear terms

In a scenario where Λ^{-4}-suppressed contributions are ignored, the predicted deviation of the cross-section, partial width and total width from their SM values can each be explicitly linearised as a function of the Wilson coefficients $\boldsymbol{c}$. Ignoring all Λ^{-4}-suppressed BSM terms in Eq. (7), and using the parametrisation of Eqs. (8)–(10), the expression for the cross-section times branching ratio reduces to

$$\begin{aligned}(\sigma\times B)^{i,k',H\to X}_{\text{SMEFT}} &= (\sigma\times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}}\times\left(1+\frac{\sigma^{i,k'}_{\text{int,(N)LO}}}{\sigma^{i,k'}_{\text{SM,(N)LO}}}\right)\times\left(\frac{1+\frac{\Gamma^{H\to X}_{\text{int}}}{\Gamma^{H\to X}_{\text{SM}}}}{1+\frac{\Gamma^{H}_{\text{int}}}{\Gamma^{H}_{\text{SM}}}}\right)\\ &= (\sigma\times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}}\times\left(1+\sum_j A_j^{\sigma_{i,k'}}c_j\right)\times\left(\frac{1+\sum_j A_j^{\Gamma^{H\to X}}c_j}{1+\sum_j A_j^{\Gamma^H}c_j}\right),\\ &= (\sigma\times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}}\times\left(\frac{1+\sum_j\left(A_j^{\sigma_{i,k'}}+A_j^{\Gamma^{H\to X}}\right)c_j+O\left(\Lambda^{-4}\right)}{1+\sum_j A_j^{\Gamma^H}c_j+O\left(\Lambda^{-4}\right)}\right),\end{aligned} \quad (12)$$

ATLAS CONF Note

ATLAS-CONF-2023-052

Event-by-Event Reweighting

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 $\epsilon_{\text{diff.}}$, 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**

Backup

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

3 Interpretations based on SM Effective Field Theory

3.1 Methodology of Effective Field Theory interpretations

The Standard Model Effective Field Theory provides an elegant language to encode the modifications of the Higgs boson properties induced by a wide class of BSM theories. Within the mathematical language of the SMEFT, the effects of BSM dynamics at high energies $\Lambda \gg v$, *i.e.* well above the electroweak scale $v = 246$ GeV, can be parametrised at low energies, $E \ll \Lambda$, in terms of higher-dimensional operators built up from the Standard Model fields and respecting its symmetries such as gauge invariance. This yields an effective Lagrangian:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{i}^{N_{d=6}} \frac{c_i}{\Lambda^2} O_i^{(6)} + \sum_{j}^{N_{d=8}} \frac{b_j}{\Lambda^4} O_j^{(8)} + \ldots, \tag{4}$$

where $\mathcal{L}_{\text{SM}}$ is the SM Lagrangian, $O_i^{(6)}$ and $O_j^{(8)}$ represent a complete set of operators of mass-dimensions $d = 6$ and $d = 8$, and c_i, b_j are the corresponding dimensionless Wilson coefficients. Operators with $d = 5$ and $d = 7$ violate lepton and/or baryon number conservation and are not considered in this study. The effective theory expansion in Eq. (4) is robust, fully general, and can be systematically matched to explicit UV-complete BSM scenarios.

The cross-section predictions for a specific process, calculated as described above, are estimated as the sum of three terms:

$$\sigma_{\text{SMEFT}} = \sigma_{\text{SM}} + \sigma_{\text{int}} + \sigma_{\text{BSM}}, \tag{5}$$

where σ_{SM} is the SM cross-section, σ_{int} describes the interference between the SMEFT operators (BSM processes) and SM operators, and σ_{BSM} is the cross-section involving exclusively SMEFT operators. When considering only $d = 6$ SMEFT operators, it follows from Eq. (4) that σ_{int} consists of terms involving a single $d = 6$ SMEFT operator, suppressing each term by a factor Λ^{-2}, and that σ_{BSM} contains terms involving products of two $d = 6$ SMEFT operators, suppressing each term by a factor Λ^{-4}. For this reason, the impact of the σ_{BSM} term is generally expected to be small, though its impact may still be non-negligible in certain regions of phase space, e. g. when energy scales are of order Λ.

The predictions are further modified by the impact of SMEFT operators on Higgs boson decay branching ratios. Since the Higgs boson is a narrow, scalar particle, and only on-shell production is considered in this analysis, its production cross-section and decay width factorise. The impact of SMEFT operators on production and decay therefore also factorises and can be derived independently. Thus, the cross-section for a given Higgs boson production process i in particle-level region k' and for a given decay mode $H \to X$ is

$$(\sigma \times B)^{i,k',H\to X}_{\text{SMEFT}} = \sigma^{i,k'}_{\text{SMEFT}} \times B^{H\to X}_{\text{SMEFT}} = \left(\sigma^{i,k'}_{\text{SM}} + \sigma^{i,k'}_{\text{int}} + \sigma^{i,k'}_{\text{BSM}}\right) \times \left(\frac{\Gamma^{H\to X}_{\text{SM}} + \Gamma^{H\to X}_{\text{int}} + \Gamma^{H\to X}_{\text{BSM}}}{\Gamma^{H}_{\text{SM}} + \Gamma^{H}_{\text{int}} + \Gamma^{H}_{\text{BSM}}}\right).$$

The factorised SMEFT prediction is calculated with ratios as in Eq. (6) to utilise the SM prediction at the highest available order:

$$(\sigma \times B)^{i,k',H\to X}_{\text{SMEFT}} = (\sigma \times B)^{i,k',H\to X}_{\text{SM,(N(N))NLO}} \left(1 + \frac{\sigma^{i,k'}_{\text{int,(N)LO}}}{\sigma^{i,k'}_{\text{SM,(N)LO}}} + \frac{\sigma^{i,k'}_{\text{BSM,(N)LO}}}{\sigma^{i,k'}_{\text{SM,(N)LO}}}\right) \left(\frac{1 + \frac{\Gamma^{H\to X}_{\text{int}}}{\Gamma^{H\to X}_{\text{SM}}} + \frac{\Gamma^{H\to X}_{\text{BSM}}}{\Gamma^{H\to X}_{\text{SM}}}}{1 + \frac{\Gamma^{H}_{\text{int}}}{\Gamma^{H}_{\text{SM}}} + \frac{\Gamma^{H}_{\text{BSM}}}{\Gamma^{H}_{\text{SM}}}}\right), \tag{7}$$

where the ratios $\sigma_{\text{int}}/\sigma_{\text{SM}}$ and $\Gamma_{\text{int}}/\Gamma_{\text{SM}}$ have a linear dependence on SMEFT operators and are suppressed by a factor Λ^{-2}, and the ratios $\sigma_{\text{BSM}}/\sigma_{\text{SM}}$ and $\Gamma_{\text{BSM}}/\Gamma_{\text{SM}}$ have a quadratic dependence on SMEFT operators and are suppressed by a factor Λ^{-4}. In the analysis, these ratios are parametrised as

$$\frac{\sigma^{i,k'}_{\text{int}}}{\sigma^{i,k'}_{\text{SM}}} = \sum_j A_j^{\sigma_{i,k'}} c_j \qquad \frac{\sigma^{i,k'}_{\text{BSM}}}{\sigma^{i,k'}_{\text{SM}}} = \sum_{j,l\geq j} B_{jl}^{\sigma_{i,k'}} c_j c_l \tag{8}$$

$$\frac{\Gamma^{H\to X}_{\text{int}}}{\Gamma^{H\to X}_{\text{SM}}} = \sum_j A_j^{\Gamma^{H\to X}} c_j \qquad \frac{\Gamma^{H\to X}_{\text{BSM}}}{\Gamma^{H\to X}_{\text{SM}}} = \sum_{j,l\geq j} B_{jl}^{\Gamma^{H\to X}} c_j c_l \tag{9}$$

$$\frac{\Gamma^{H}_{\text{int}}}{\Gamma^{H}_{\text{SM}}} = \sum_j A_j^{\Gamma^{H}} c_j \qquad \frac{\Gamma^{H}_{\text{BSM}}}{\Gamma^{H}_{\text{SM}}} = \sum_{j,l\geq j} B_{jl}^{\Gamma^{H}} c_j c_l, \tag{10}$$

with

$$A_j^{\Gamma^H} = \frac{\sum_X \Gamma^{H\to X}_{\text{SM}} A_j^{\Gamma^{H\to X}}}{\sum_X \Gamma^{H\to X}_{\text{SM}}} \qquad B_{jl}^{\Gamma^H} = \frac{\sum_X \Gamma^{H\to X}_{\text{SM}} B_{jl}^{\Gamma^{H\to X}}}{\sum_X \Gamma^{H\to X}_{\text{SM}}}. \tag{11}$$

In Eq. (11) all Higgs boson decay modes X with up to four final-state particles are included in the sum. All $A_j^{\sigma_{i,k'}}$, $A_j^{\Gamma^{H\to X}}$, $B_{jl}^{\sigma_{i,k'}}$ and $B_{jl}^{\Gamma^{H\to X}}$ coefficients are constant factors obtained from simulation that express the sensitivity of the process to the operators O_j and O_l that correspond to the Wilson coefficients c_j and c_l,

3.1.2 Cross-section calculation with linear terms

In a scenario where Λ^{-4}-suppressed contributions are ignored, the predicted deviation of the cross-section, partial width and total width from their SM values can each be explicitly linearised as a function of the Wilson coefficients $\boldsymbol{c}$. Ignoring all Λ^{-4}-suppressed BSM terms in Eq. (7), and using the parametrisation of Eqs. (8)–(10), the expression for the cross-section times branching ratio reduces to

$$\begin{aligned}
(\sigma \times B)^{i,k',H\to X}_{\text{SMEFT}} &= (\sigma \times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}} \times \left(1 + \frac{\sigma^{i,k'}_{\text{int,(N)LO}}}{\sigma^{i,k'}_{\text{SM,(N)LO}}}\right) \times \left(\frac{1 + \frac{\Gamma^{H\to X}_{\text{int}}}{\Gamma^{H\to X}_{\text{SM}}}}{1 + \frac{\Gamma^{H}_{\text{int}}}{\Gamma^{H}_{\text{SM}}}}\right) \\
&= (\sigma \times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}} \times \left(1 + \sum_j A_j^{\sigma_{i,k'}} c_j\right) \times \left(\frac{1 + \sum_j A_j^{\Gamma^{H\to X}} c_j}{1 + \sum_j A_j^{\Gamma^H} c_j}\right), \\
&= (\sigma \times B)^{i,k',H\to X}_{\text{SM,((N)N)NLO}} \times \left(\frac{1 + \sum_j \left(A_j^{\sigma_{i,k'}} + A_j^{\Gamma^{H\to X}}\right) c_j + O\left(\Lambda^{-4}\right)}{1 + \sum_j A_j^{\Gamma^H} c_j + O\left(\Lambda^{-4}\right)}\right),
\end{aligned} \tag{12}$$

where all higher order terms in the expansion are suppressed by power Λ^{-4} or beyond.