Which additional information has to be provided for measurements?

Louie Corpe (UCL)
LHC EW WG General Meeting, Dec 2019
Global Fits and Reinterpretation of SM Measurements Session
• Analysis preservation is increasingly important discussion in HEP
  --> *new and ambitious goals of combinations, recasting, EFT fits etc...*

• Are we routinely storing enough information on HEPData to efficiently re-use the measurements we make at the LHC?
  --> *Not always! ...And when we do, the expts follow different conventions... Areas where small policy shifts can boost impact of analyses*

• Prompted by discussion on correlations: Dec 18, Feb 19, July 19
  Attempt recommendation in [note](#) to be agreed between LHC expts
  • Give *recommendations on conventions to follow* depending on what level of re-interpretation is needed

• Recommendations discussed in dedicated presentations to experiments
  • ATLAS (SM, HDBS, PMG groups) Oct-Dec 2019
  • CMS (Generator Group) Late Oct 2019
  • LHCb (QEE Group) Nov 2019
  • ALICE : planned discussion in Jan 2019
• Analysis preservation is increasingly important discussion in HEP
  --> new and ambitious goals of combinations, recasting, EFT fits etc...

• Are we routinely storing enough information on HEPData to efficiently re-use the measurements we make at the LHC?
  --> Not always! ...And when we do, the expts follow different conventions...
  Areas where small policy shifts can boost impact of analyses

• Prompted by discussion on correlations: Dec 18, Feb 19, July 19
  Attempt recommendation in note to be agreed between LHC expts
  • Give recommendations on conventions to follow depending on what level of re-interpretation is needed

• Recommendations discussed in dedicated presentations to experiments
  • ATLAS (SM, HDBS, PMG groups) Oct-Dec 2019
  • CMS (Generator Group) Late Oct 2019
  • LHCb (QEE Group) Nov 2019
  • ALICE : planned discussion in Jan 2019

I’ll reming the recommendations, and will summaries some of the feedback from experiments so far
What are the current practices?

And how/why do they need to change?
• Define fiducial volume of measurement (eg **Rivet Routine**) Often comes years later, or not at all! Documenting analysis logic w/ code snippet is vital for preservation

• Give results with uncertainties for in each bin. Separate stat vs syst uncertainties at minimum. Stat/syst not enough to model correlations if re-interp is to be trusted

• Use exact names given by CP groups for error sources/nuisance parameters (NPs) This is not always respected

**Existing recommendations**

**ATLAS SM measurement from 2017**, but still no rivet routine!

**CMS 13 TeV jet substructure measurements** in tt: with full uncertainty breakdown!
Existing recommendations

- If strong correlations...
  a) explicit covariance or correlation matrix OK only don’t intend to combine with other measurements
  b) give breakdown of signed(!) effect of each NP. Can then rebuild covariance matrix

- Prefer to use b) since a) implicitly symmetrizes, and information to correlate with other measurements is insufficient.

Recent CMS tt measurement which included covariance matrix directly

Recent ATLAS measurement which included covariance matrix directly
Existing recommendations

• Statistical correlations from bootstrap method, store either replicas or correlation matrix. Replicas best for future combinations but need make TH*DBootstrap code public

• Theory predictions not usually in HEPData
Exception: when theory predictions complicated/intensive to be produced
Policy from time when bkg calculations were cheap. **No longer the case!** A shift here can boost our impact

• *(For searches: exclusion limits + necessary info to emulate analysis (e.g. kinematic distributions, signal acceptances and selection efficiencies)* Rivet can now preserve searches w/ custom smearing+efficiency capabilities)
Updated HEPData Recommendations

Defines 3 scenarios for levels of information to provide on HEPData

Gives concrete recommendations for the format of objects which are to be stored

Hopefully can be used by experiments to update internal recommendations (process has started within ATLAS!)
3 Scenarios for re-interpretation

• Identify different levels of recommendations, depending on the analysis type and how re-interpretable it needs to be:

  Scenario A - Maximum Re-interpretability
  Scenario B - Approximate Re-interpretability
  Scenario C - Minimum Requirements for Analysis Preservation

Best case - aims to provide maximal information for reinterpretations. Should be gold standard for precision measurements
• Identify different levels of recommendations, depending on the analysis type and how re-interpretable it needs to be:

Scenario A - Maximum Re-interpretability
Scenario B - Approximate Re-interpretability
Scenario C - Minimum Requirements for Analysis Preservation

Best case - aims to provide maximal information for reinterpretations. Should be gold standard for precision measurements

Closest to current situation. Plenty of information published. Not necessarily enough for strict combinations... but good enough for many analyses (especially searches)
3 Scenarios for re-interpretation

• Identify different levels of recommendations, depending on the analysis type and how re-interpretable it needs to be:

  Scenario A - Maximum Re-interpretability
  Scenario B - Approximate Re-interpretability
  Scenario C - Minimum Requirements for Analysis Preservation

  Best case - aims to provide maximal information for reinterpretations. Should be gold standard for precision measurements

  Closest to current situation. Plenty of information published. Not necessarily enough for strict combinations... but good enough for many analyses (especially searches)

  Bare minimum for a search to be re-interpretable
A - Maximum Re-interpretability

- **For precision analyses**: for future combinations, measurements of SM parameters, PDF fitting... Enough info for exact combination

- **Phase space definition**: Rivet analysis must be provided
  - If results only at detector level, Rivet analysis should still be provided, with adequate smearing and efficiency tables

- **Stat correlations**: Bootstrap Replicas attached to HEPData entry

- **Syst correlations uncertainty breakdown**, = effect of each NP on each bin
  - -> can convert to cov matrix + correlate w/ other measurements

- **Background**: include SM prediction from latest MC generators w/ breakdown of theory uncertainty if possible

- If likelihood fit used: **post-fit values of the NPs in each bin**
B - Approximate Re-interpretability

• **For standard measurements or searches** to be re-interpreted approximately. E.g generator tuning, and recasting of searches

• **Phase space definition**: Rivet analysis must be provided
  - If results only at detector level, Rivet analysis should still be provided, with adequate smearing and efficiency tables

• **Statistical correlations**: correlation matrices. *Can’t infer corrs between analyses, but OK if re-interpreting result in isolation*

• **Systematic correlations**: uncertainty breakdown, = effect of each NP on each bin -> cov matrix + correlate w/ other measurements
  - OR, cov matrix for each distribution: e.g. for simplified likelihoods

• **Background**: include **SM prediction from latest MC generators** w/ breakdown of theory uncertainty if possible
C - Bare Minimum

- **Minimum amount of info** for result to be re-used meaningfully. e.g. if only rough estimate of MC/data agreement or sensitivity to new models needed.

- **Phase Space Definition**: Ideally, Rivet routine... if not...
  - detailed description of the region of interest
  - per-object efficiency tables
  - explicit definitions of each variable used in the selection,
  - cutflows of the effect of each selection on well-defined signals

- **Statistical correlations**: omitted if negligible bin migrations. Stat error per bin still needed (assumed uncorrelated between bins)

- **Systematic correlations**: uncert breakdown or explicit cov matrices

- **Background**: SM bkg prediction of MC generators, w/ breakdown of theory uncertainty if possible
Feedback from Experiments
so far...
• Possibility to include directly the Likelihood in the `pyhf` framework. Potentially more powerful than uncertainty breakdown?
• Should we be considering “industry-standard” data types like Pandas?
• Relationship of HEPData+Rivet vs RECAST?
• Will the TH*DBootstrap classes (currently ATLAS-only) be made public (yes, ATLAS Pub note to accompany public git repo in the new year)
• Are SM predictions always necessary (eg LHCb?)
• Analysis contacts uploading HEPData entries need to ensure that the uncertainty breakdowns are reviewed by the collaboration before being made public -> include in supplemental/aux material
• Could we include BDT xml files on HEPData in the case of ML techniques?
• Give links to examples of HEPdata routines where things have been done well (show best practice!)
Summary

• What additional info is needed to achieve our global fit/re-interp goals?

• LHC EW WG is reviewing recommendations for HEPData to answer these questions. Recommendations document in preparation...

• Recent developments are excellent opportunity to review status and see what we can do better
  --> maximise impact of our measurements and searches!

• In particular, we should be more diligent about:
  • preserving analysis logic in Rivet routines in a timely manner
  • preserving correlations between bins (stat, syst)
  • including state of the art SM predictions in HEPData

• Next steps: experiments to try to use the recommendations doc to review their internal practices: so that everyone is on the same page.
  • ATLAS have already started this process.
Backup