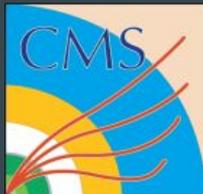


CMS Policies and Practises on providing information for reinterpretations

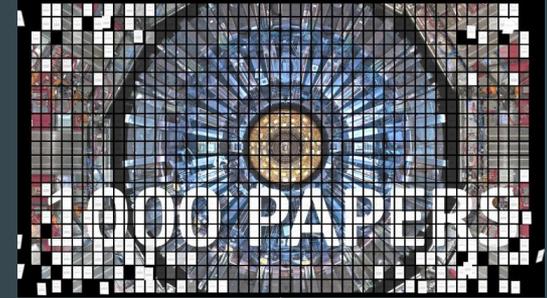


José Ruiz
on behalf of CMS Collaboration

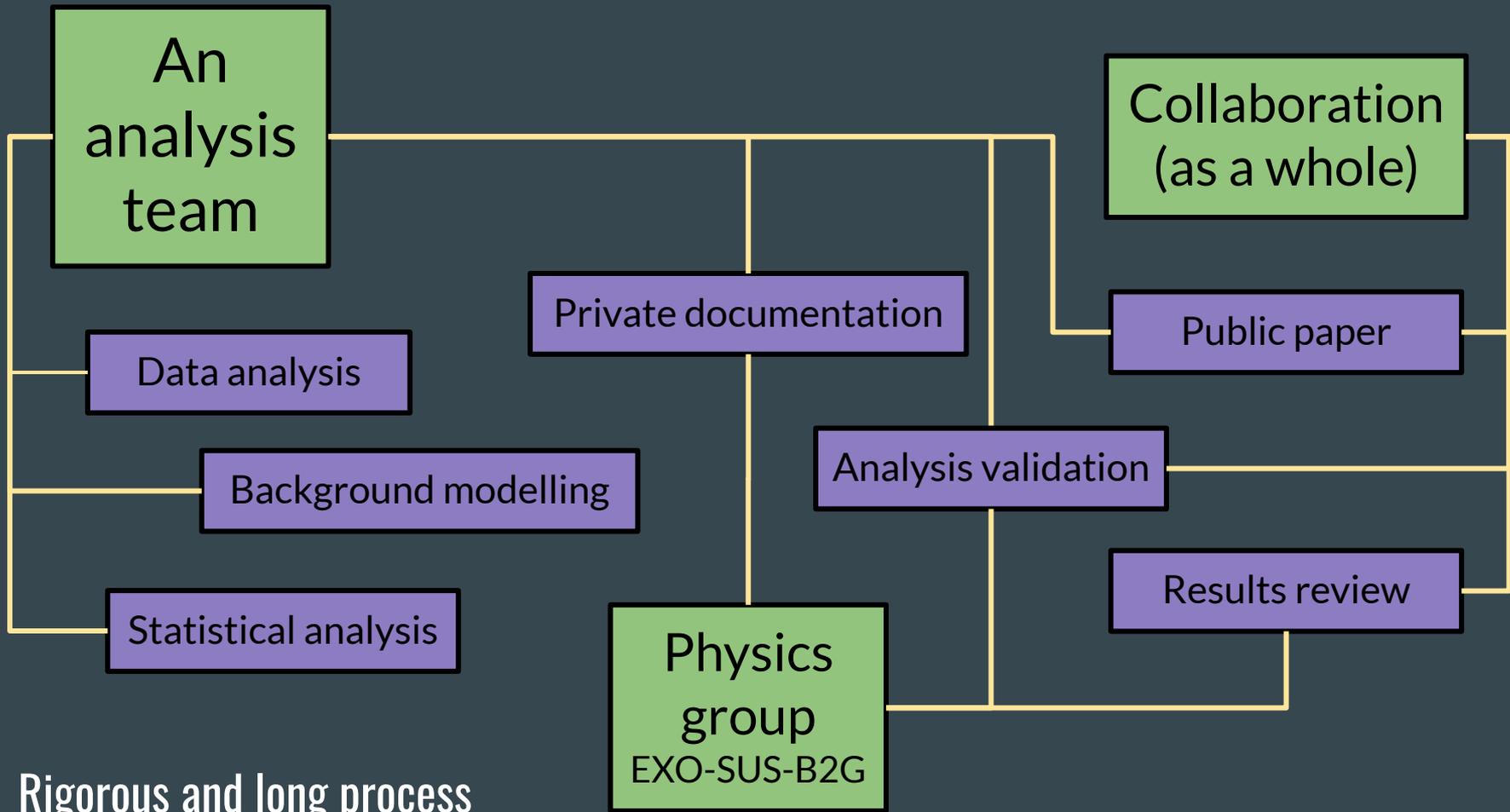


**This talk will focus on BSM
searches at CMS**

How to get away with a search



An
analysis
team



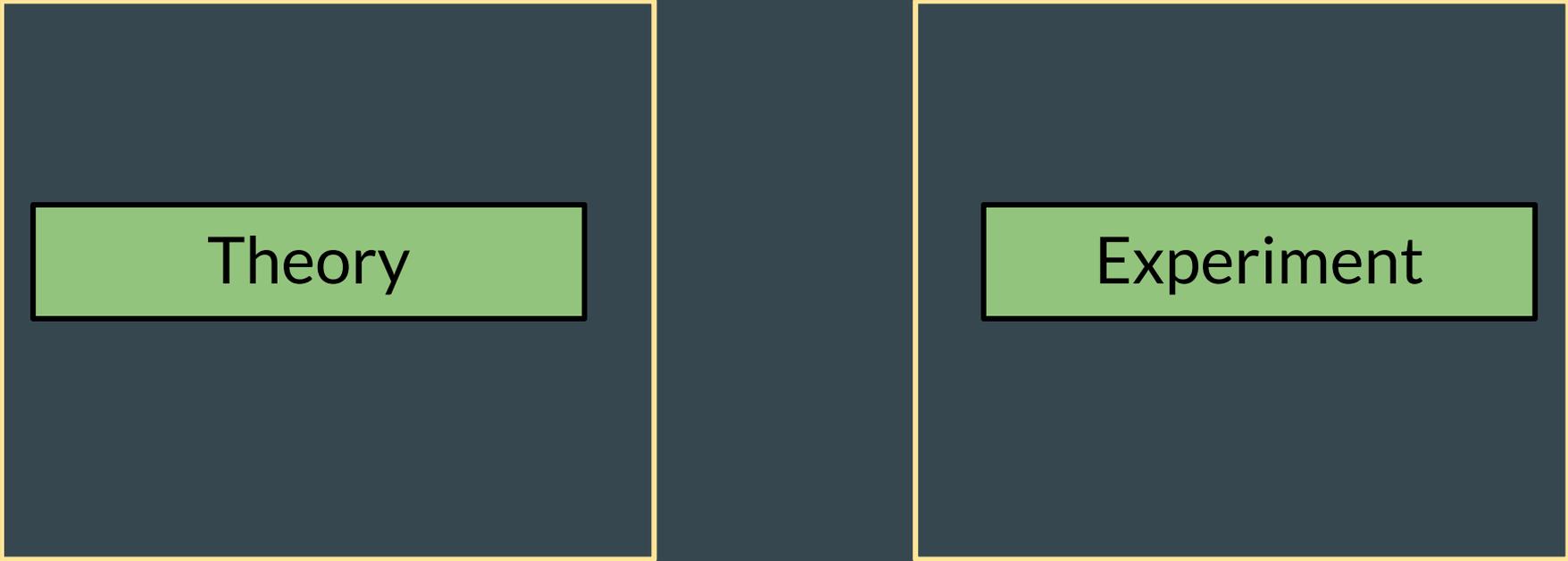
Rigorous and long process

Simplistic description

Depending on many factors it can take roughly one or more years to make and publish and analysis.

Analysis teams as well as the coordination areas are doing a continuous effort to get results in shorter timescales as well as in the most usable way by the community.

Communicating two worlds

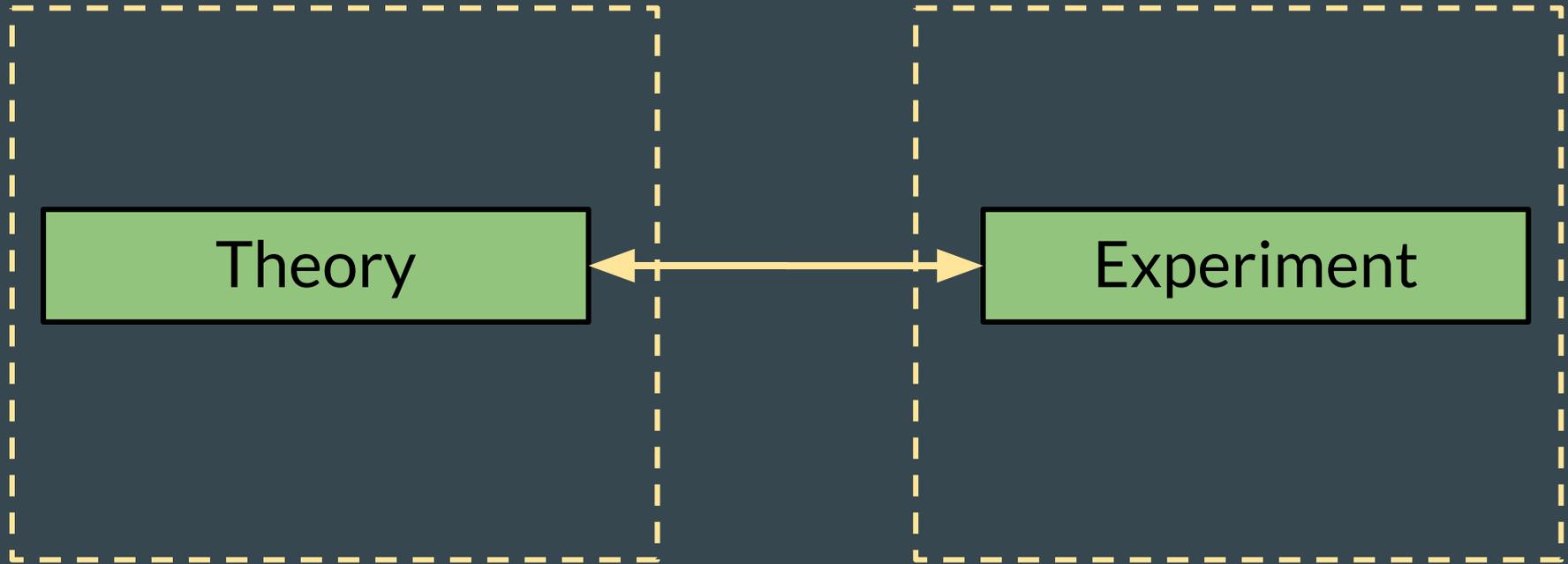


The diagram consists of two large, empty rectangular boxes with yellow borders, positioned side-by-side. Each box contains a smaller, horizontal rectangular box with a black border and a light green fill. The left green box contains the word 'Theory' and the right green box contains the word 'Experiment'. The overall background is a dark blue-grey color.

Theory

Experiment

Communicating two worlds



CMS policies

What is a CMS result (paper):

1. Strategy and procedures description.
2. Figures, tables, numbers:
 - a. Histograms
 - b. Estimations (e.g. the trigger efficiency is at least 99% for signal like events)
 - c. Total counts
 - d. Efficiencies.
 - e. Covariance matrix.
 - f. Acceptance maps.
 - g. Limits
3. Statistical interpretation

What is not (usually) in a CMS public result:

1. Montecarlo samples
2. Data
3. Data analysis code
4. Technical recipes/specifications
5. Internal review details

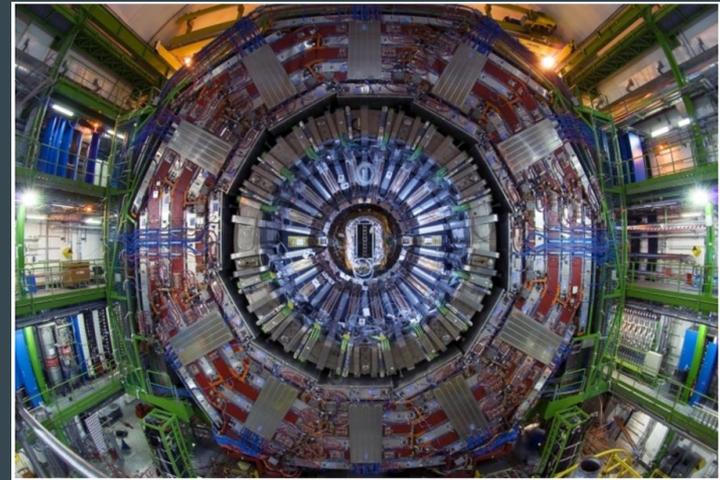
From a CMS paper (experiments in general)

- Difficult to reproduce. Some important details are often private or it is not feasible to redo them outside the collaboration.
- Numbers are fetchable, histograms not so much: using histograms for reinterpretations imply to postprocess them and convert them to numbers, if they are delivered only as figures.
- Statistical analysis and interpretation: typically the statistical procedure can have details that are very technical and the statistic routines could be difficult to use without guidance.
- For new signals a whole set of approximations have to be done as detector effects are not really possible to calculate from outside the collaboration.

Improving communication of results

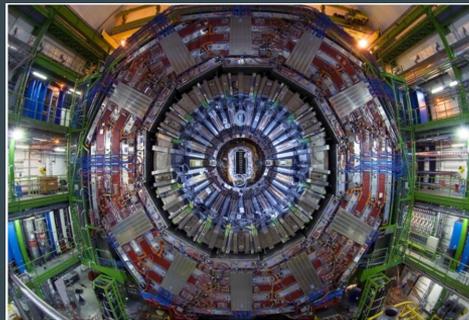
Different approaches:

1. Detector unfolding.
2. Supplemental material.
3. Open data.
4. Providing likelihoods.



Detector unfolding

Events



Unfolded
events

RIVET toolkit (Robust Independent Validation of Experiment and Theory)

rivet is hosted by Hepforge, IPPP Durham

- Rivet home
 - Contur
 - Professor
 - YODA
 - MCplots
 - AGILE
- Downloads
- Analyses
 - Standard analyses
 - Analysis changelog
 - Writing an analysis
- Analysis coverage & wishlists
 - General
 - No searches/HI
 - Searches
 - Heavy ion
- Documentation
 - Manual & talk links
 - Getting started / tutorials
 - Rivet via Docker
 - Changelog
 - Doxygen code/API docs
- Source code
- Contact

Rivet analysis coverage

Rivet analyses exist for 976/5812 papers = 17%. 183 priority analyses required.

Total number of Inspire papers scanned = 10551, at 2020-11-26

Breakdown by identified experiment (in development):

Key	ALICE	ATLAS	CMS	LHCb	Forward	HERA	$e^+e^- (\geq 12 \text{ GeV})$	$e^+e^- (\leq 12 \text{ GeV})$	Tevatron	RHIC	SPS	Other
Rivet wanted (total):	236	287	435	202	43	519	733	617	1238	464	61	1
Rivet REALLY wanted:	36	38	79	8	0	13	1	1	6	1	0	0
Rivet provided:	26/262 = 10%	167/454 = 37%	89/524 = 17%	16/218 = 7%	8/51 = 16%	9/528 = 2%	193/926 = 21%	381/998 = 38%	59/1297 = 5%	8/472 = 2%	5/66 = 8%	18/19 = 95%

Show greylist

Show blacklist

ALICE

ATLAS

CMS

LHCb

Forward

HERA

$e^+e^- (\geq 12 \text{ GeV})$

$e^+e^- (\leq 12 \text{ GeV})$

Tevatron

RHIC

SPS

Provides:
Analysis code.
Validation framework.
Analysis results.

CMS and RIVET

- Typically an effort pursued mainly by generators group.
- Specially interesting for the collaboration in order to validate Montecarlo generators, as it provides unfolded event.
- It would require some non-negligible amount of work to implement each analysis and making it available in RIVET.
- It has been difficult to establish a policy to enforce such implementation for every CMS analysis. Coordination has recommended it to the collaboration and the analysis teams are aware of the importance and very often try to make the effort to implement it. Technical difficulties usually overcome this effort.

HEPData - Supplemental material

About Submission Help Sign in

CMS Search Advanced

Max results Sort by Reverse order Showing 25 of 441 results

Date



« < 1 2 > »

Measurement of the Z boson differential production cross section using its invisible decay mode ($Z\nu\bar{\nu}$) in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS collaboration Sirunyan, Albert M ; Tumasyan, Armen ; Adam, Wolfgang ; *et al.*

CMS-SMP-18-003, 2020.

Inspire Record 1837084 DOI 10.17182/hepdata.96028

Measurements of the total and differential fiducial cross sections for the Z boson decaying into two neutrinos are presented at the LHC in proton-proton collisions at a center-of-mass energy of 13 TeV. The data were collected by the CMS detector in 2016 and correspond to an integrated luminosity of 35.9 fb^{-1} . In these measurements, events are selected containing an imbalan...

11 data tables

Figure 2 The measured and predicted inclusive fiducial cross sections in fb. The experimental measurement includes both statistical and systematics uncertainties. The theoretical prediction includes both the QCD scale and PDF uncertainties.

Table 2 Experimental uncertainties affecting transfer factors in the analysis that is used to estimate the W background in the signal region (SR). The number of W boson events are denoted as W_{SR} for the SR and in analogy as $W_{\mu\nu}$ ($W_{e\nu}$) for the single-muon (single-electron) control region (CR).

Table 3 Uncertainties assigned to the simulation based processes in SR and CRs.

More...

Collaboration [Reset](#)

CMS 441

TOTEM 2

ATLAS 1

Subject_areas

hep-ex 434

nucl-ex 62

Phenomenology-HEP 1

Phrases

Proton-Proton Scattering 226

Provides:

**Analysis results in a more
suitable format for recasting.**

**Common framework for
different sources.**

CMS and HEPData

- Has become an official and necessary step in order to make an analysis public. No analysis should be made public, at least in BSM, without the accompanying HEPData entry.
- All official plots and tables in the paper have to be added to the HEPData entry.
- Lightweight work, compared to other tools: It usually takes a reasonable amount of work to “translate” results into HEPData format. Additionally there is always an expert which helps preparing the HEPData entry and validating the process.
- Beyond tables and figures: CMS also usually provides efficiencies, acceptance maps, covariance/correlation matrices, and simplified likelihoods.
- Automatization code: [Link](#)

Open data

The screenshot shows the Open Data CERN website interface. At the top left is the logo "opendata CERN". A search bar contains the text "Search" and a magnifying glass icon. On the top right are links for "Help" and "About".

Below the search bar, there is a filter section on the left. It includes a "CMS x" tag, a checkbox for "include on-demand datasets", and a "Filter by type" section. The filter section lists various categories with their respective counts:

Category	Count
Dataset	1167
Collision	148
Derived	136
Simulated	883
Documentation	42
About	3
Activities	13
Authors	3
Guide	18
Help	2
Policy	2
Report	1
Environment	31
Condition	9
VM	9
Validation	13
News	8
Software	38
Analysis	14
Framework	3

On the right side of the page, there are sorting and display options: "Sort by: Best match asc." and "Display: detailed 20 results". Below this, it states "Found 3994 results.".

The main content area displays three articles:

- Getting Started with CMS 2010 Open Data**
To analyse CMS data collected in 2010, you need version 4.2.8 of CMSSW, supported only on Scientific Linux 5. If you are unfamiliar with Linux, take a look at this short introduction to Linux or try this...
Buttons: Documentation Guide CMS Getting Started
- Getting Started with CMS 2011 Open Data**
To analyse CMS data collected in 2011 and 2012, you need version 5.3.32 of CMSSW, supported only on Scientific Linux 6. If you are unfamiliar with Linux, take a look at this short introduction to Linux or try this...
Buttons: Documentation Guide CMS Getting Started
- CMS Guide to generate configuration files for event production**
Guide to produce CMSSW python configuration files to simulate, generate and/or reconstruct collision events.

Provides:
Full reproducibility.
Datasets.
Montecarlo events.
Code.

Providing public likelihoods

- Since 2017 a proposal for simplified likelihoods for public release has been made by the collaboration: [Link](#).
- So far there is no official requirement to publish full likelihoods for CMS analyses.
- However, there is a motivation to move along that direction and use the simplified likelihoods approach.
- Some steps have been taken towards having a common recipe to publish likelihoods and some internal discussion has taken place moreover the last year.
- More work is needed to overcome the technical difficulties to get CMS analyses up to this point.

Beyond publishing

CMS is also interested to have a fluent and continuous communication with the theory community:

1. Theory talks at working groups: Invited talks given by a theorist at the plenary meeting of the different physics groups.
2. CMS theorists: Theorists officially linked with CMS that can look at private information and advice analysis teams.
3. Working groups: Special working groups for specific topics which are thought to need a more intricate development with theory colleagues, e.g. EFT forum.

Conclusions

1. Several policies have been implemented within CMS to assure a better communication of analysis results, mainly thinking towards an easier and more precise way reusing CMS results.
2. Also practises have been set and tools have been put in place in order to have a more broad application within the analysis teams.
3. A culture of sharing the analysis results and “making the effort” has been spread twos the collaborations in the last years.
4. More works need to be done, specially to achieve full likelihood publishing.

Thank you!