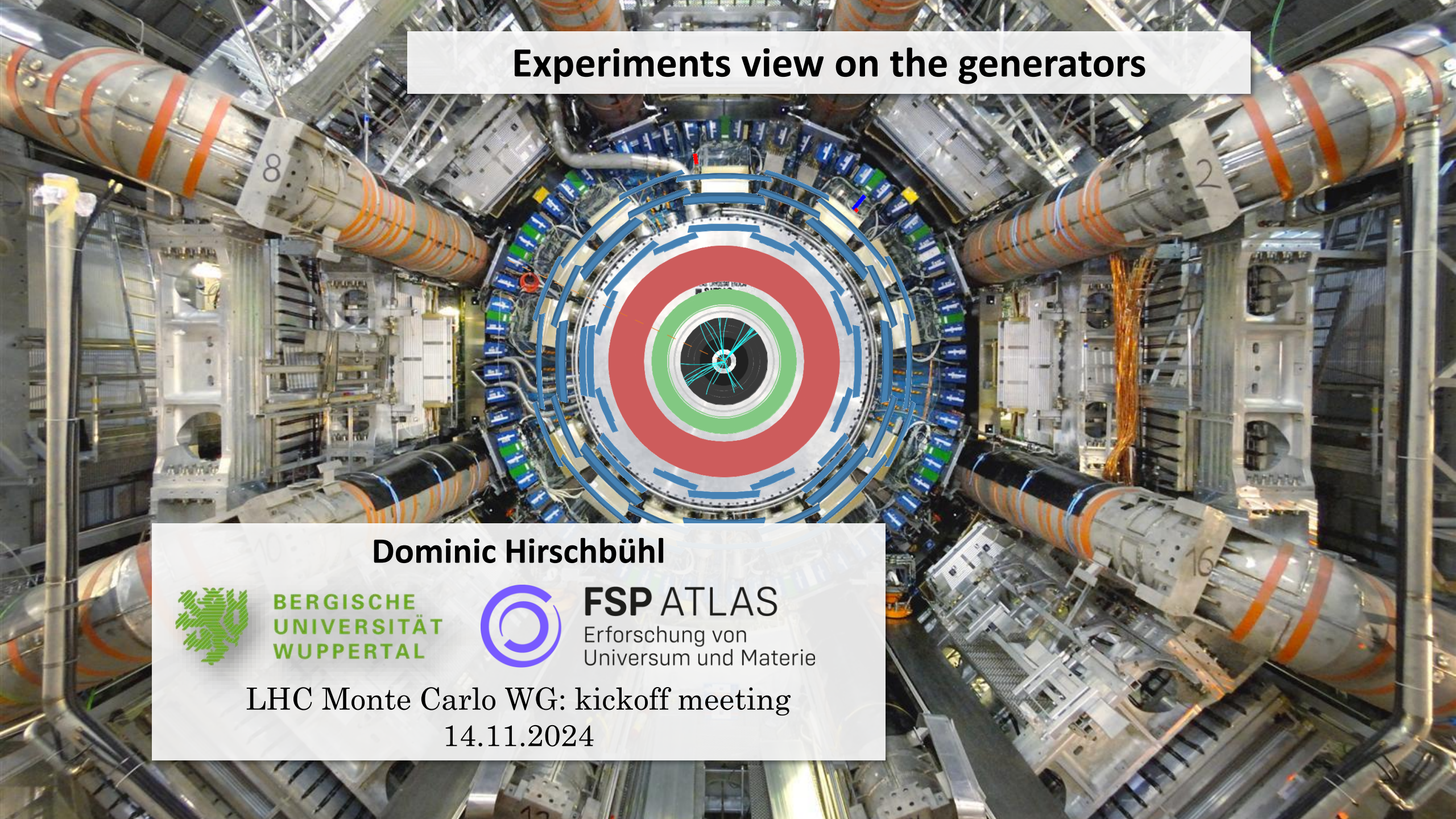
Experiments view on the generators



Dominic Hirschbühl

BERGISCHE UNIVERSITÄT WUPPERTAL

FSP ATLAS
Erforschung von Universum und Materie

LHC Monte Carlo WG: kickoff meeting
14.11.2024

Needs and expectations

Technical point of interests

Physics issues

Collaboration

Generator techniques

Technical points of interest

Improvements in calculation efficiency

Improvements in the calculation itself (in random order)

- Profiling of generator code
  - Collaborations have quite some experience here
- Usage of GPUs
  - Experiments all investigating GPU usage, and GPUs are more available
  - Landscape of LHC computing hardware may change
- Optimized parallelization
- Many points have been discussed in the last “**Event generators' and N(n)LO codes' acceleration”** meeting

Improvements in the way the event generation is done

- Reducing negative events weight fraction
- Generation in interchangeable steps
  - Separate ME / Shower / Hadronisation where possible

Generation in interchangeable steps

Output Format	Les Houches Events		HepMC / EVNT	HepMC / EVNT
Generator	**Matrix Element**	**Parton Shower**	**Stable Particles**	**Afterburner**
Beam Halo Generator			✓	
Cavern Background Generator			✓	
Cosmic Ray Generator			✓	
EPOS	Only Minimum Bias	✓	✓	
EvtGen				✓
Herwig	$2 \to 2$ LO and NLO	✓	✓	
Hijing	Only Minimum Bias	✓	✓	
Hto4l	✓			
Hydjet	Only Minimum Bias	✓	✓	
MadGraph5_aMC@NLO	✓			
ParticleGun			✓	
Photos				✓
Powheg Box	✓			
Prophecy4f	✓			
Protos	✓			
Pythia	Only $2 \to 2$	✓	✓	
Pythia 8B			✓	
QGSJet	Only $2 \to 2$ jets	✓	✓	
Sherpa	✓	✓	✓	
STARlight		✓	✓	
SuperChic	✓			
Tauola				✓
Other Generators (via LHEF)	✓			
Other Generators (via HepMC)			✓	

arXiv:2404.06335

Stolen from C. Gütschow: talk

Future event generation workflows

- Approach 1: produce parton-level samples centrally with input from the MC developers, provide them in a shared space for all experiments
  - experiments run their preferred shower setup (✔)
  - allows for affordable plug & play between different models (✔)
  - lowers cost threshold for reproducing larger setups after some time if need be (✔)
  - requires more storage for parton-level events (✘)
  - new infrastructure needs to be set up and maintained (✘)

More interchangeable steps means more sophisticated systematic recipes and more flexibility in the (re-)generation of events

Technical points of interest

- Our production versions often lag development versions
- New advances are sometimes very hard to deploy or not well supported
  - MC event generators are running within our Athena framework using dedicated interface class
  - Athena framework is quite different from standalone running
    → relies on LCG layers, fixed versions of gcc, python etc.
  - Inclusion in our framework ensures consistency of general Parameters (EW parameters, particle masses, PDF sets, ett.)
- Problems to feedback patches to the authors
  → would be nice to have common infrastructure and software management

POWHEG+PYTHIA8	POWHEG BOX v2	$t\bar{t}$ NLO	5FS	PYTHIA 8.230	POWHEG $h_{\text{damp}} = 1.5 m_{\text{top}}$ $p_{\text{T}}^{\text{hard}} = 0$ globalRecoil recoilToColoured=ON	A14	nom.
POWHEG+PYTHIA8 h_{damp}	POWHEG BOX v2	$t\bar{t}$ NLO	5FS	PYTHIA 8.230	POWHEG $h_{\text{damp}} = 3 m_{\text{top}}$	A14	syst.
POWHEG+PYTHIA8 $p_{\text{T}}^{\text{hard}}$	POWHEG BOX v2	$t\bar{t}$ NLO	5FS	PYTHIA 8.230	POWHEG $p_{\text{T}}^{\text{hard}} = 1$	A14	syst.
POWHEG+PYTHIA8 RecoilToTop	POWHEG BOX v2	$t\bar{t}$ NLO	5FS	PYTHIA 8.230	POWHEG recoilToTop	A14	syst.
POWHEG+HERWIG 7	POWHEG BOX v2	$t\bar{t}$ NLO	5FS	HERWIG 7.1.3	POWHEG	H7.1-Default	syst.

Euer default $t\bar{t}$ sample for run 2 was setup 2017, we have:

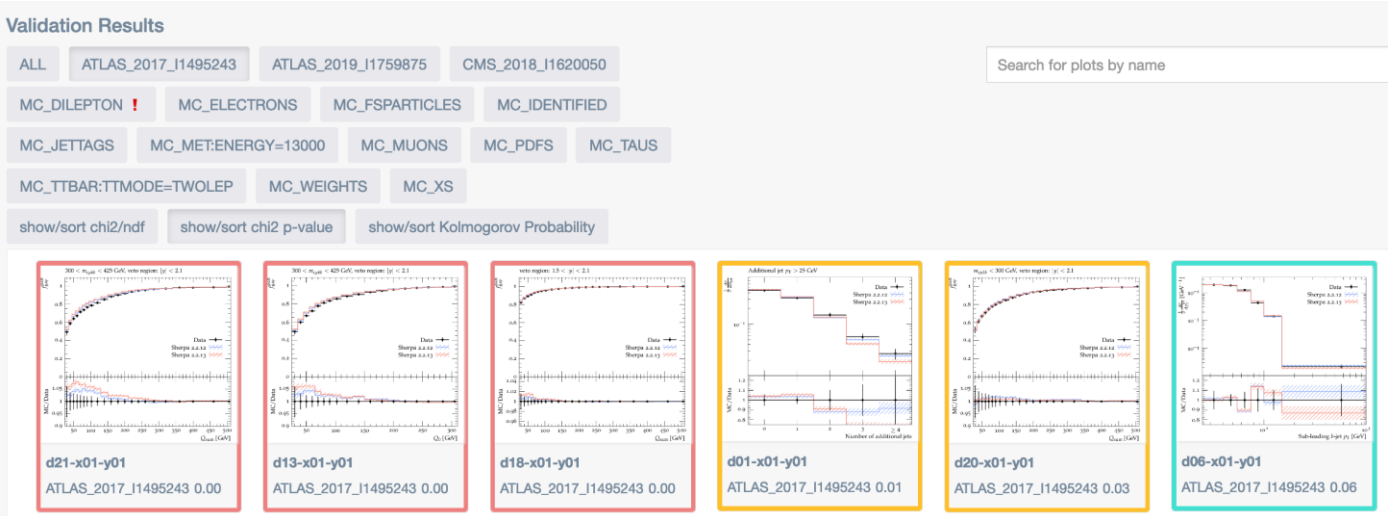
- ~100B LHE events
- > 5B showered and simulated events

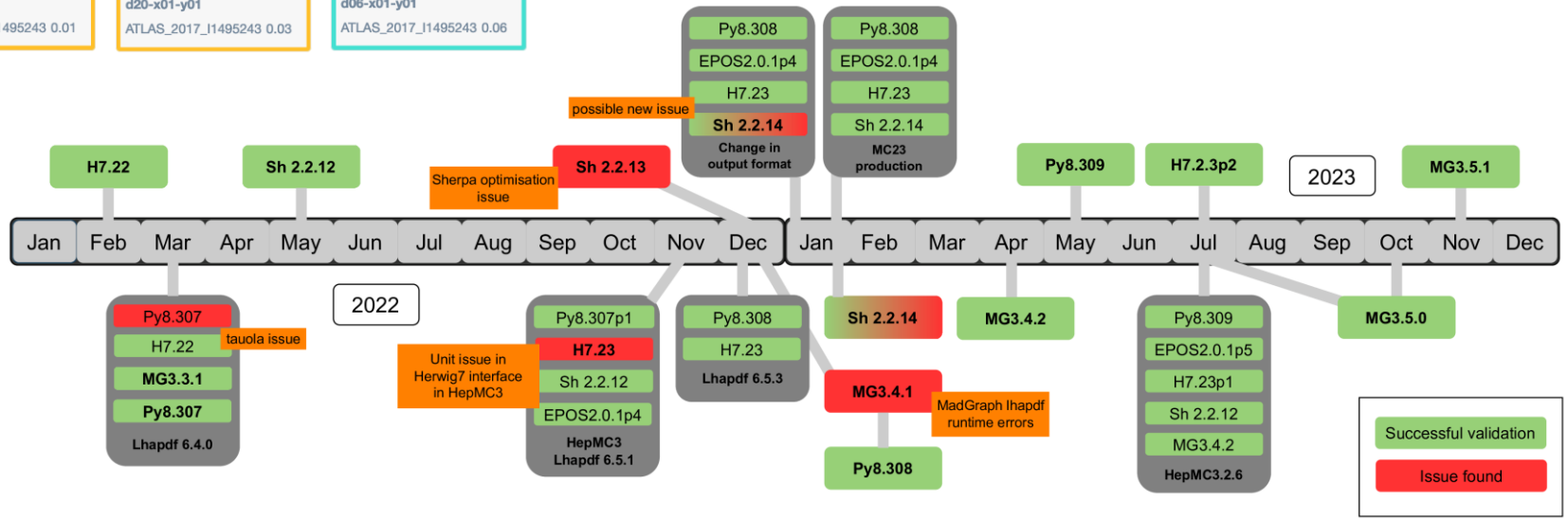
https://indico.cern.ch/event/1444046/contributions/6188473/attachments/2964686/5215638/LHCTopWG_111124_ttplusHF.pdf

Technical points of interest

Regular validations of all new generator versions using a RIVET based framework

[ATL-PHYS-PUB-2024-013]

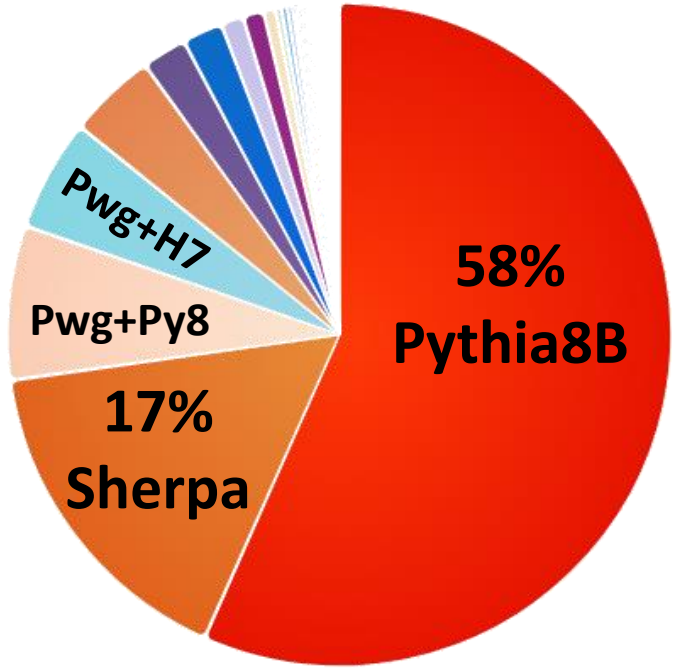




Generator techniques

- **Efficient slicing**
  - Low statistics for very high cross-section kinematic regions
  - Higher statistics in the tails of distributions
- **Improving filter efficiency** (high pT, heavy flavor, etc)
  - Ensuring we generate what we want
  - Major missing piece: ‘flavor enhancement’
- **New phase-space sampling techniques that avoid biases**
  - Particularly for populating unusual kinematic regions
- **Integrating systematic uncertainties as weights**
  - Current systematic model comes with up to 7 alternative samples

CPU usage in the last 6 months



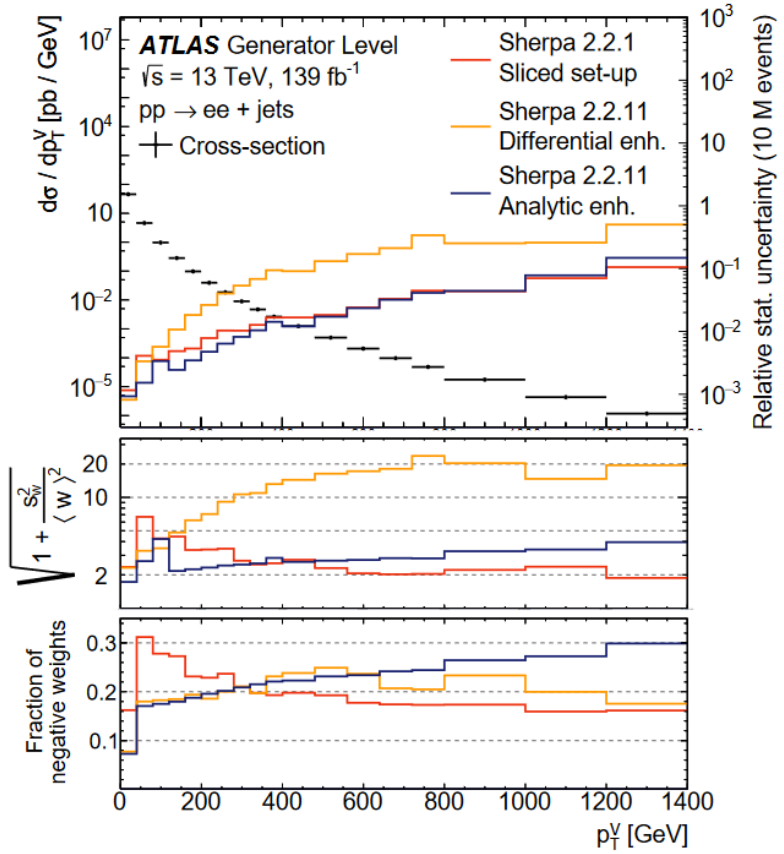
This is only a snapshot and not representative of the whole MC production.

Treatment of negative weights

- Negative weights are a statistics killer
  - Statistical power of a sample with negative weight fraction ε is reduced by 1/(1-2ε)2
  - ε=25% → 4x larger sample is needed for the same statistical power

- If the negative weight fraction is >30%,
  samples are hardly usable

- Various techniques have been proposed for improving this
  - How can we ensure a widely-deployed solution?
  - Should we focus the community on one solution
    to avoid divergence?



Physics

Modelling of specific processes

- Studies currently done with or without interactions with the authors

- ATLAS can provide quite some computing power and also person power through authorship qualifiers

Systematic uncertainties

Title	ATLAS link	Publication	Category	Dominant uncertainty grouped systematics
Search for same-charge top-quark pair production	TOPQ-2021-14	arxiv:2409.14982	search	ME&PS of ttW
Measurement of top-quark pair production in association with charm quarks	TOPQ-2021-26	arxiv:2409.11305	fiducial xs	tt modeling 9% of 17%
Search for ttH/A->tttt production	EXOT-2022-13	arxiv:2408.17164	search	tt+>=1b modelling (11% of 14%)
tW production	TOPQ-2018-07	Phys. Rev. D 110 (2024) 072010	incl. xs	Jet flavour composition
Measurement of tt production with additional b-jets	TOPQ-2019-03	arxiv:2407.13473	fiducial xs 3b	b-tagging
			fiducial xs 4b	jet

In many (almost all) analyses modelling uncertaintines are dominiant → would be nice to have a common set of recommendations for specific generator setups (NLO+PS / Mulitjet matched/merged etc)

Physics

Tuning

- ATLAS was a bit sleeping on this topic in the last few years
- Growing interest but no concrete plans yet
- Interesting topic in the view of this group would be
  - To do some common tuning of different generators to reduce systematic uncertainties
  - Studies how universal current tunes are, i.e. in the view of NLO-merged setups
  - Do we need a consistent tuning of systematic variations, e.g. recoil-to-top vs. r_b

Collaboration

Technical aspects

Process specific questions

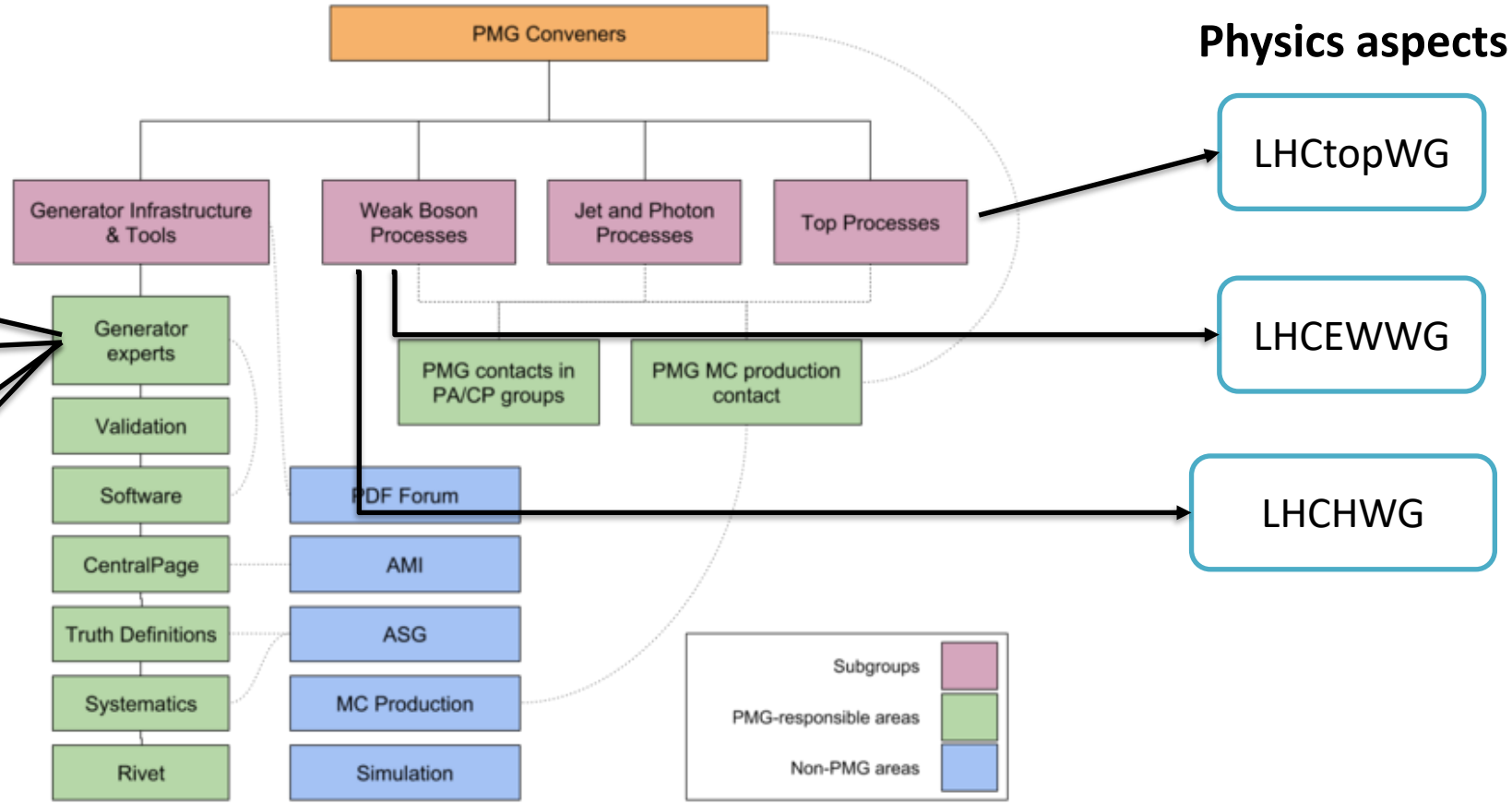
Pythia8
Two MCI

Herwig7

Sherpa
Experts part of ATLAS

Madgraph
Two MCI

Powheg
One MCI



Physics aspects

LHCtopWG

LHCEWWG

LHCHWG

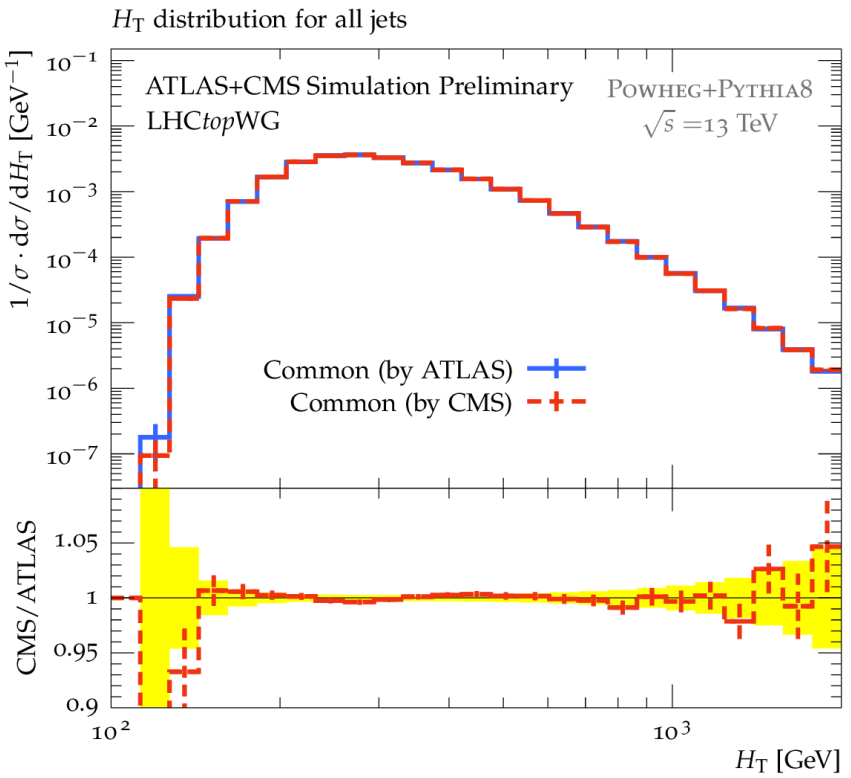
Regular invitations to the GIT meeting of generator experts

MCI: *Monte Carlo Interactions*

Sharing events

Could save 50% of CPU and eventually disk space when sharing generated events

- Several practical advantages
  - Could at least overlay identical theory lines
  - Or have a common alternative sample
  - Or ATLAS alternative is CMS baseline etc.
- Could also save “integration time” of software
- First step done within the LHCtopWG:
  - Common Powheg+Pythia8 $t\bar{t}$ sample based on **shared LHE files**
  - Common Sherpa $t\bar{t}$ sample based on **identical config files**



The main problem here is to agree on common parameters

→ MC community might propose default settings,
especially for shower tunes

Summary

Technical point of interests

- Computing efficiency
- Generation in interchangeable steps
- Speed-up of new codes into our framework

Generator techniques

- Efficient slicing, improving filter efficiencies
- New phase-space sampling techniques that avoid biases
- Integrating systematic uncertainties as weights

Physics

- Developing common set of systematic uncertainty recommendations for specific generator setups
- Tuning → universality of current tunes, consistent tuning of systematic uncertainties

Collaboration

- Tightening the interaction with the MC authors
- Sharing events / setups with other experiments