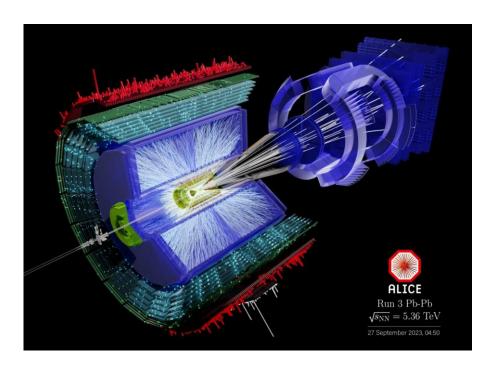




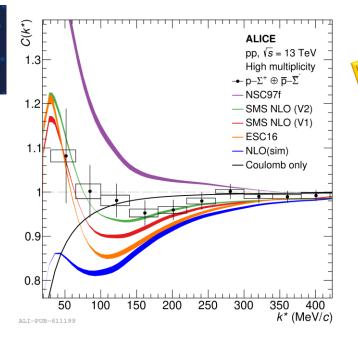


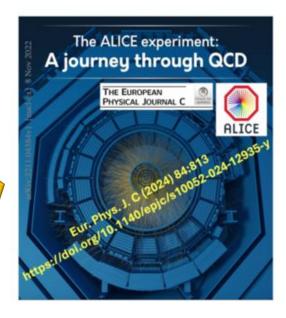
## 首届ALICE实验与重离子物理讲习班

The 1st ALICE Experiment and Heavy-Ion Physics School









### **ALICE Data Preparation**

Jian Liu (刘剑) University of Liverpool

3 - 10 November 2025 | Fudan University, Shanghai, China

# Why DPG matters?

**DPG:** Data Preparation Group

Useful link (the ALICE DPG TWiki):

https://twiki.cern.ch/twiki/bin/viewauth/ALICE/AliceDPG

#### **Activities**

TWiki > ALICE Web > AliceDPG (2025-07-22, CatalinLucianRistea)



#### ALICE Data Preparation Group (DPG) Home Page

#### Table Of Content

- ↓ Organization, coordination, meetings
- ↓ DPG software
- ↓ Raw data production
- ↓ Monte Carlo production
  - ↓ General information
  - ↓ Lists of Monte Carlo productions
  - ↓ Run3 specific information on Monte Carlo productions
- ↓ Known issues
- Known issues in Monte Carlo productions
- ↓ AOD production
- ↓ Quality Assurance / Quality Control
  - ↓ QA/QC organization, material and contacts
  - → Run Lists
  - ↓ Analysis QA/QC
- ↓ Analysis Tools
  - Event selection and properties (AOT events)
  - → Track selection and properties (AOT tracks)
- ↓ Useful links

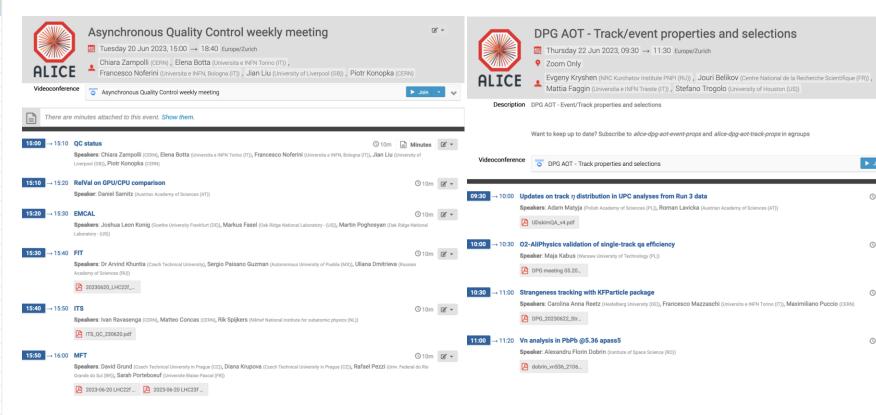
The Data Preparation group is responsible for **steering and coordinating** the **reconstruction of the data** collected by Alice and the preparation and the execution of the **Monte Carlo simulations**.

It also in charge of **organizing** the **Quality Assurance** of the reconstructed and simulated data (at the detector, tracking and particle identification levels) and of **developing** and improving the **Quality Assurance tools** for future runs.

The third area of responsibility of the DPG is in the preparation and maintenance of the **Analysis Objects and Tools**, which includes the production, maintenance, quality assurance and bookkeeping of the **AOD** (Analysis Object Data) files, as well as the coordination of the groups working on **event selections and properties** and **track selections and properties**.

### Main activities

Production	Description
.HC22t_apass3	LHC22t - apass3 of LHC22t - 13.6 TeV, EPN, O2-3522
.HC22r_apass3	LHC22r - apass3 of LHC22r - 13.6 TeV, CPU, O2-3522
HC22p_apass3	LHC22p - apass3 of LHC22p - 13.6 TeV, EPN, O2-3522
.HC22m_apass3_relval_epn3_nogpu	LHC22m - Release Validation 2 in view of apass3 of pp data from 2022, 13.6 TeV, EPN noGPU, O2-3564
.HC22q_apass3	LHC22q - apass3 of LHC22q - 13.6 TeV, CPU, O2-3522
.HC22m_apass3_relval_epn2_debug	LHC22m - Release Validation 2 (debug) in view of apass3 of pp data from 2022, 13.6 To EPN, O2-3540
.HC22m_apass3_relval_cpu2_debug	LHC22m - Release Validation 2 (debug) in view of apass3 of pp data from 2022, 13.6 To CPU, O2-3540
.HC22m_apass3_relval_epn2	LHC22m - Release Validation 2 in view of apass3 of pp data from 2022, 13.6 TeV, EPN, Q2-3540
.HC22m_apass3_relval_cpu2	LHC22m - Release Validation 2 in view of apass3 of pp data from 2022, 13.6 TeV, CPU, O2-3540
.HC22t_apass3_relval_cpu_bis	LHC22t - Release Validation in view of apass3 of pp data from 2022, 13.6 TeV, CPU, updated tag, O2-3540
.HC22t_apass3_relval_cpu	LHC22t - Release Validation in view of apass3 of pp data from 2022, 13.6 TeV, CPU, O2-3540
.HC22o_apass3_relval_cpu	LHC22o - Release Validation in view of apass3 of pp data from 2022, 13.6 TeV, CPU, O2-3540
.HC22m_apass3_relval_epn	LHC22m - Release Validation in view of apass3 of pp data from 2022, 13.6 TeV, EPN, O2-3540
.HC22m_apass3_relval_cpu	LHC22m - Release Validation in view of apass3 of pp data from 2022, 13.6 TeV, CPU, O2-3540
.HC22s_apass4_meanvtx	LHC22s - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 5.36 TeV Pb-Pb, O2-3523
.HC22o-test_apass2_meanvtx	LHC220-test - Calib pass to extract mean vertex calib after alignment and reco update, CTFs every 10 minutes - 13.6 TeV pp, O2-3523
.HC22t_apass2_meanvtx	LHC22t - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 13.6 TeV pp, O2-3523
.HC22r_apass2_meanvtx	LHC22r - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 13.6 TeV pp, O2-3523
.HC22q_apass2_meanvtx	LHC22q - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 13.6 TeV pp, O2-3523
.HC22p_apass2_meanvtx	LHC22p - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 13.6 TeV pp, O2-3523
.HC22o_apass2_meanvtx	LHC22o - Calib pass to extract mean vertex calib after alignment and reco update, 10 C every 10 minutes - 13.6 TeV pp, O2-3523
.HC22h_apass2_residuals2	LHC22h - Low rate runs residuals extraction (2), O2-3512
.HC22m_apass2_meanvtx	LHC22m - Calib pass to extract mean vertex calib after alignment and reco update, 10 devery 10 minutes - 13.6 TeV pp, O2-3523
.HC22t_apass2	LHC22t - apass2 of LHC22t on CPU - 13.6 TeV pp, O2-3464
.HC22q_apass2	LHC22q - apass2 of LHC22q on CPU - 13.6 TeV pp, O2-3464
.HC22t_apass2	LHC22t - apass2 of LHC22t on EPN - 13.6 TeV pp, O2-3464
.HC22j_apass2_residuals	LHC22j - Low rate runs residuals extraction, O2-3512
.HC22h_apass2_residuals	LHC22h - Low rate runs residuals extraction, O2-3512
.HC22f_apass2_residuals	LHC22f - Low rate runs residuals extraction, O2-3512
.HC22p_apass2	LHC22p - apass2 of LHC22p on CPU - 13.6 TeV pp, O2-3464
HC22r_apass2	LHC22r - apass2 of LHC22r on CPU - 13.6 TeV pp, O2-3464



https://indico.cern.ch/category/8532/

https://indico.cern.ch/category/8451/

③30m 🗷 🕶

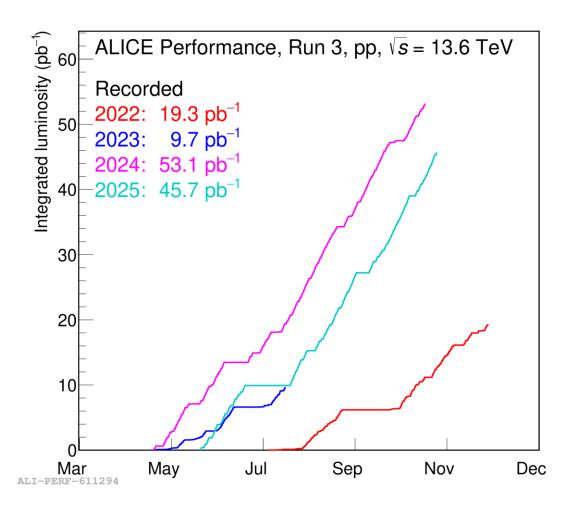
③30m 🗷 🕶

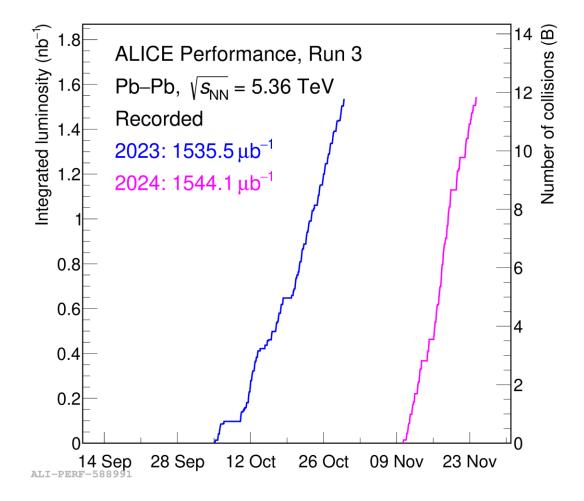
③30m 🗷 🕶

③20m 🗷 🕶

https://alimonitor.cern.ch/

## Where do we stand?

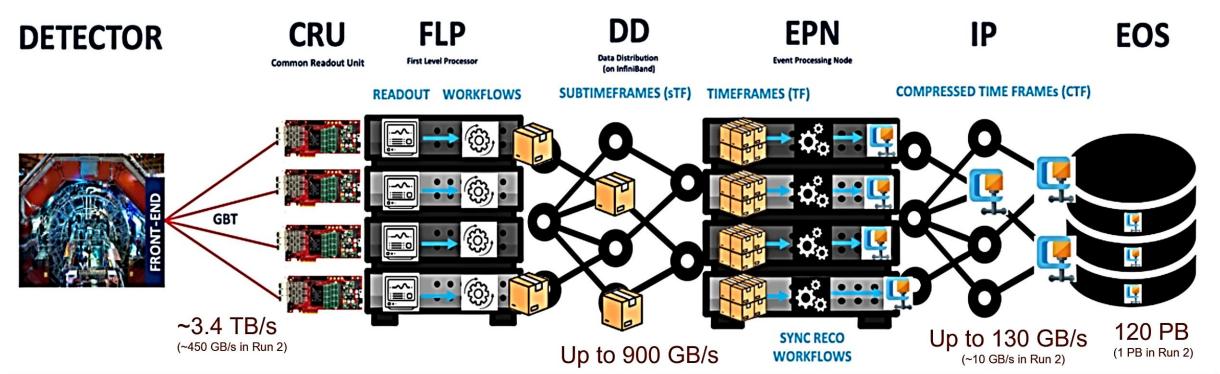




# Online/Offline (O<sup>2</sup>)

#### Targeting to record large minimum-bias sample

- All collisions stored for main detectors with continuous readout in combination with triggered detectors
- Extreme online data compression (~3.4 TB/s → ~0.1 TB/s) → GPUs
  to speed up online (and offline) processing
- First level processors (FLP)
  - Readout of detectors (3 TB/s) and raw data processing
  - 200 nodes in total
- Event processing nodes (EPN)
  - Synchronous/asynchronous event reconstruction
  - 280 nodes each with 8 GPUs (extended to 350 nodes)



### Heart beat and time frame

#### Heart Beat (HB)

issued in continuous & triggered modes to all detectors

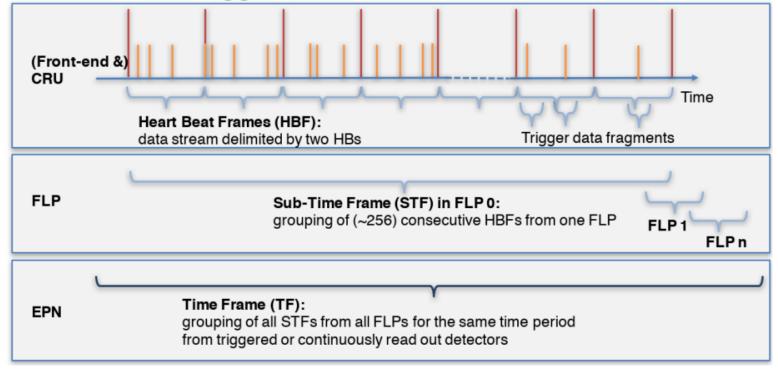
#### Physics trigger

can be sent to upgraded detectors will be sent to non-upgraded detectors

#### HBF and TF rates programmable Typical values:

- HB: 1 per orbit, 89.4 μs: ~10 kHz
- TF: 1 every ~20 ms: ~50 Hz
- → 1 TF = ~256 HBF

#### Continuous & Triggered read-out



- HB allows synchronization and TF sampling from detectors with continuous and triggered readouts
- Synchronized with LHC clock
- HBF is the smallest chunk of data which is inspected by CTP and can be dropped if the quality is bad

LHC revolution frequency is ≈ 11 245 Hz



Each proton bunch completes one full turn around the ring every  $1/11245 \approx 89.4 \mu s$ , know as an orbit.



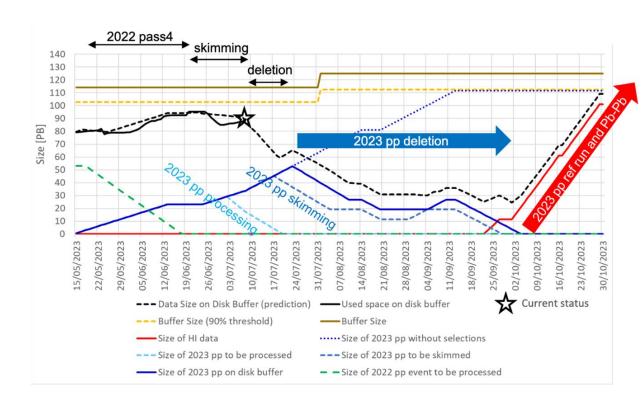
The ring is divided into 3 564 RF buckets



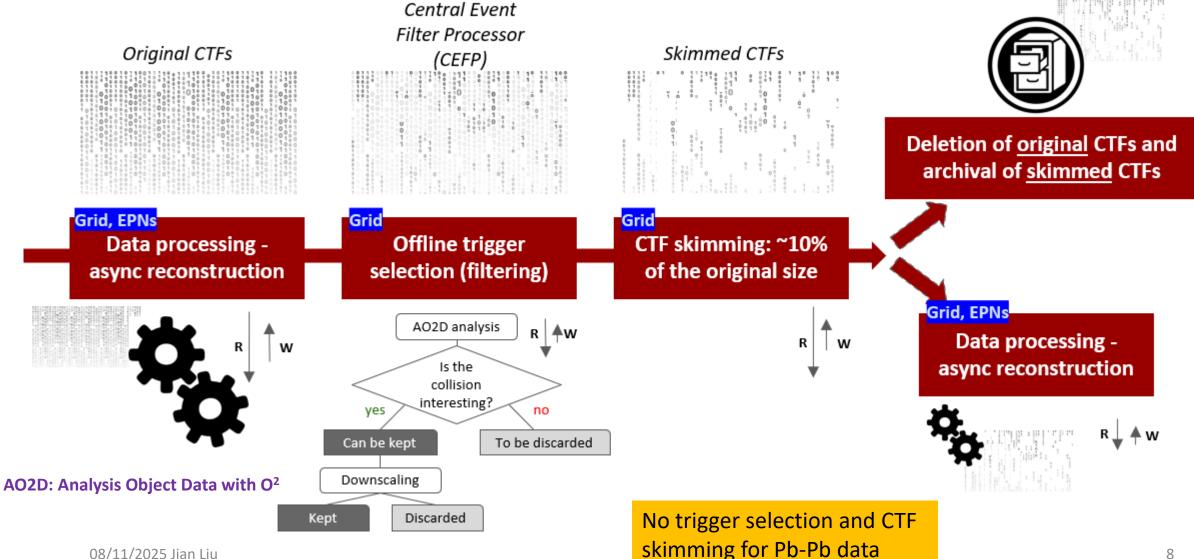
One bunching crossing (BC) is  $T_{crbit}/3564 = 24.95$  ns

# Challenges for pp data in Run 3

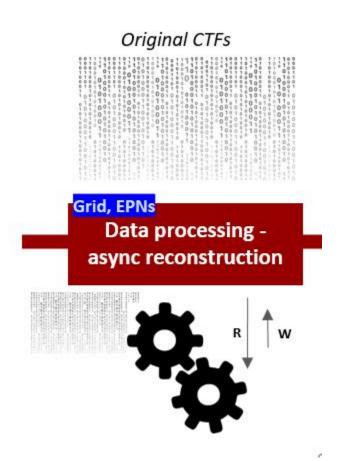
- Very large data sample collected in pp collisions
  - ~9 PB of data taken every week
    - Data are stored in the so-called CTF (Compressed Time Frame) files
    - Organized in 10 minutes long folders in alien
  - Disk buffer at P2 is ~135 PB
  - Considering ~21 weeks of pp data taking → ~190 PB from pp data taking
  - This alone exceeds the O2 buffer space, not including the fact that on the buffer we need to keep some old data
    - detector data
    - Pb-Pb from previous year
    - other data
- Solution: reduce the size of the CTFs by applying an offline trigger selection (OTS); CTFs are then re-written in the so called "skimmed CTFs"



# Asynchronous data processing for pp



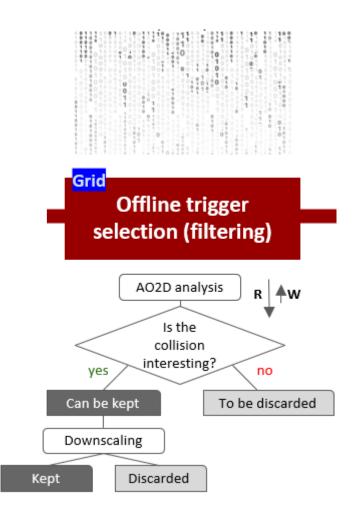
# pp data processing: C/APass



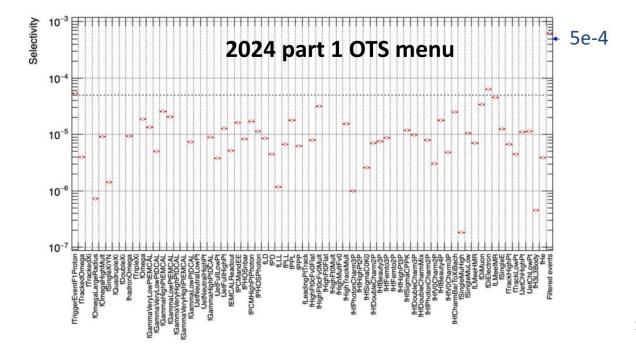
- 1. Calibration Pass (CPass): first reconstruction of the data, triggered automatically, runs on 10% of the CTFs, selected in a controlled-random way
- 2. Manual calibration to fix online calibrations, improve others, and QC to collect feedback from data quality
  - First assessment happens already in Online; async reconstruction allows to validate and possibly reject some data
- 3. APass1 (Asynchronous Pass): first full reconstruction of the data, triggered by completion of manual calibration and QC
- 4. After APass1, TPC PID calibration for analysis is extracted
  - Neural Network approach

NB: in case of major issues, a data pass may need to redo.

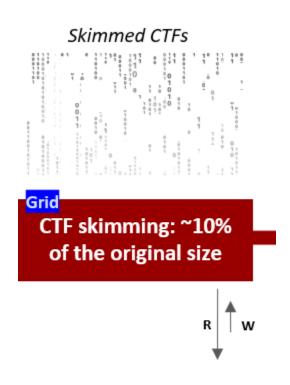
# pp data processing: offline trigger selection



- Offline Trigger Selection (OTS): analysis on AO2Ds from APass1 to identify collisions interesting for analysis
  - 1. Selected BCs are identified by different analyses
  - 2. Total selectivity budget is of the order of 5e<sup>-4</sup> and it is shared among the PWGs
    - Goal: final CTF compression of the order of 4%



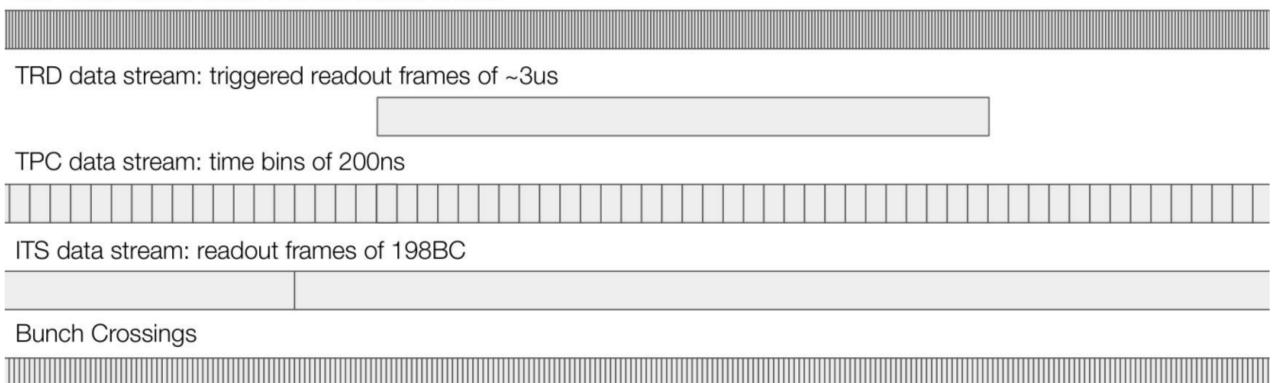
# pp data processing: CTF skimming



- 1. CTF skimming: following the decision of the OTS, CTFs are read again, and skimmed CTFs are written out selecting only the interesting BCs, with margin
  - 1. Margin needed in order to keep the full event information for all detectors (e.g. in TPC we keep 4000BC, a drift time)
  - 2. Extra margin of 1000BC added for safety
- 2. Original CTFs are then deleted (for good!)
  - 1. A 5 per mille MinBias sample is kept
    - 1. 2022: 1 pb<sup>-1</sup> was kept (larger sample)
    - 2. 2023: 0.5 pb<sup>-1</sup> was kept
- 3. Further reconstruction passes can happen only on the skimmed CTFs (and the MinBias sample)
- 4. Since skimmed CTFs are only a few percent of the size of the original ones, further processing is fast
- 5. apass1\_skimmed follows
  - 1. e.g. 24al
    - 1. apass1 output: 5 PB
    - 2. apass1\_skimmed: 230 TB  $\rightarrow$  <5%

# CTF skimming

TOF data stream: stream of hits sorted in time

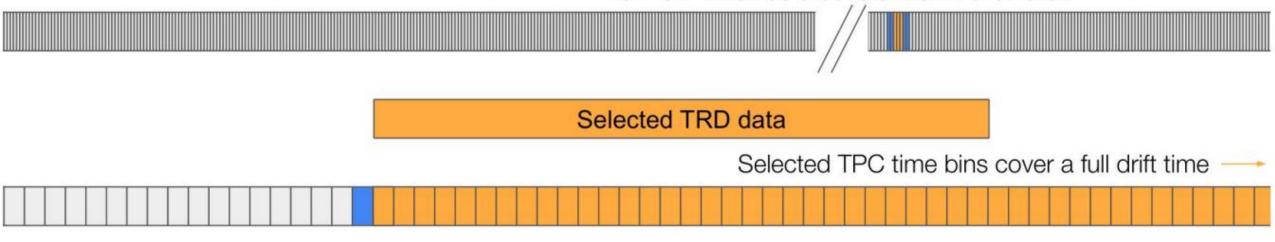


- The CTF reader has been modified to read, for each detectors, only the data related to the selected BC
  - These data will contain pile-up for most detectors due to their ROF being larger than the average distance between two collisions

#### Pictorial view of the data stream of 4 barrel detectors

# CTF skimming

TOF raw time has a constant shift of 1/3 orbit



#### Selected ITS ROF



Selected BC + tolerance due to finite collision time resolution

When selecting an event, we do not define only a selected BC, but a window of selected BCs for two reasons

- The collision time resolution is finite, and we allow for a 4-sigma window around the measured collision time
- The found collision is associated at analysis level to the closest T0 signal, that might be few tens of BCs away

# Data productions

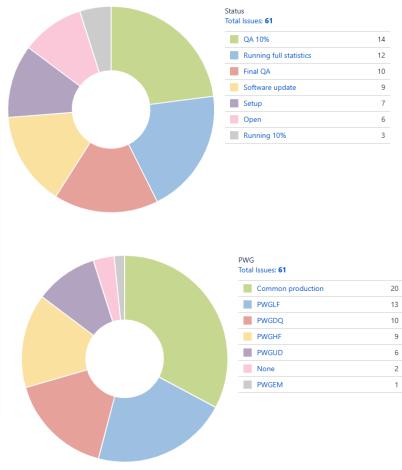
https://alimonitor.cern.ch/production/raw.jsp

						RAW Production	Cycles									
								_					<b>-</b>		Processing re	
						Raw data		Reco	nstructed ESDs	1		Output	Timing		Software versions	>>
Production	Description	Col.	Status	Run Range	Runs	Chunks Size	Chunks	%		%	Events	Size	Running	Saving		Err
LHC25ai_apass1_skimmed	LHC25ai - apass1_skimmed of LHC25ai - 13.6 TeV, CPU, O2-6261	<i>⊗</i> рр		566017 - 566123	7	430,850 3.838 PB	105,760	25%	16.61 TB	1%	1,382,086,015	16.61 TB	15y 58d	1y 101d	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	c9-
LHC25am_apass1	LHC25am - apass1 of LHC25am - 13.6 TeV, CPU, O2-6439	<i></i> ⊘ рр	Running	566493 - 566579	8	613,016 <sup>5.463</sup> PB	529,341	86%	415.9 TB	8%	78,796,262,564	415.9 TB	330y 24d	22y 120d	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	
LHC25ai_skimming	LHC25ai - CTF skimming and re-writing of LHC25ai - 13.6 TeV, O2-6261	<i></i> ⊘ рр	Running	566017 - 566123	7	430,850 <sup>3.838</sup> PB	0	0%	185.6 TB	4%	0	185.6 TB	13y 177d		O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	3-2
LHC25ah_apass1_sampled	LHC25ah - apass1_sampled of LHC25ah - 13.6 TeV, CPU, O2-6261	<i>6</i> рр	Running	564704 - 565805	98	6,640,043 <sup>59.15</sup> PB	33,227	1%	26.01 TB	8%	4,998,082,090	26.01 TB	20y 234d	1y 321d	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	:9- 3-2
LHC24aq_apass1_muon_matching4	LHC24aq - apass1 - pp 13.6 TeV, Reconstruction of one 2024 ppRef run for muon alignment studies, replay updated tag, 02-6245	<i></i> рр	Running	559408 - 559408	1	167,156 <sup>1.492</sup> PB	2,406	1%	2.437 TB	11%	469,096,270	2.437 TB	242d 6:39	4d 13:03	O2PDPSuite::daily-20251026-0000	)-1
LHC24ar_apass3	LHC24ar - Pb-Pb 5.36 TeV, apass3 of 5.36 TeV Pb-Pb 2024 data, CPU, plus helper EPN, O2-6321	🔗 РЬРЬ	Running	559631 - 559903	6	127,851 <sup>1.126</sup> PB	108,733	85%	87.14 TB	8%	456,324,479	87.14 TB	107y 318d	5y 215d	O2PDPSuite::async-async-2024-pbpb-apass3-v slc9-alidist-async-2024-pbpb-apass3-v2	
LHC24ar_apass3	LHC24ar - Pb-Pb 5.36 TeV, apass3 of 5.36 TeV Pb-Pb 2024 data, EPN-Mi100, O2-6321	🔗 РЬРЬ	Running	559574 - 559595	2	5,104 <sup>42.09</sup> TB	1,982	39%	1.602 TB	9%	8,009,332	1.602 TB	1y 135d	9d 19:26	O2PDPSuite::async-async-2024-pbpb-apass3-v slc9-alidist-async-2024-pbpb-apass3-v2	
LHC24ar_apass3	LHC24ar - Pb-Pb 5.36 TeV, apass3 of 5.36 TeV Pb-Pb 2024 data, CPU, Polaris, O2-6321	🔗 РЬРЬ	Running	559575 - 559901	5	164,515 <sup>1.463</sup> PB	112,718	69%	91.28 TB	8%	476,400,532	91.28 TB	212y 206d	8y 183d	O2PDPSuite::async-async-2024-pbpb-apass3-v slc9-alidist-async-2024-pbpb-apass3-v2	
LHC25ac_apass1_sampled	LHC25ac - apass1_sampled of LHC25ac - 13.6 TeV, CPU, O2-6014	<i></i> ⊘ рр	Running	563182 - 563998	53	3,792,025 33.79 PB	18,948	0%	15.23 TB	8%	2,886,886,293	15.23 TB	12y 81d	265d 14:07	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	
LHC24ar_apass3_test_mi100	LHC24ar - Pb-Pb 5.36 TeV, test apass3 of 5.36 TeV Pb-Pb 2024 data, production tag, EPN-Mi100, O2- 6321		Running	559827 - 559827	1	62,312 <sup>571.2</sup> TB	1,889	3%	1.57 TB	9%	8,024,751	1.57 TB	1y 115d	8d 11:13	O2PDPSuite::async-async-2024-pbpb-apass3-v slc9-alidist-async-2024-pbpb-apass3-v2	
LHC24ar_apass3_test_mi50	LHC24ar - Pb-Pb 5.36 TeV, test apass3 of 5.36 TeV Pb-Pb 2024 data, production tag, EPN-Mi50, O2- 6321		Running	559827 - 559827	1	62,312 <sup>571.2</sup> TB	61,654	99%	50.75 TB	8%	253,110,112	50.75 TB	41y 320d	177d 11:46	O2PDPSuite::async-async-2024-pbpb-apass3-v slc9-alidist-async-2024-pbpb-apass3-v2	
LHC25am_apass1_tpc_nn_cpu	LHC25am - apass1 of LHC25am - for TPC NN, 13.6 TeV, CPU, O2-6360	<i></i> рр	Running	566696 - 566697	2	10,184 90.62 TB	10,148	100%	9.024 TB	9%	1,727,605,940	9.024 TB	8y 221d	14d 14:53	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3	
LHC25am_apass1_tpc_nn	LHC25am - apass1 of LHC25am - for TPC NN, 13.6 TeV, EPN, O2-6360	<b>⊘</b> рр	Running	566696 - 566697	3	15,176 <sup>134.4</sup> TB	590	4%	457.8 GB	8%	87,652,859	457.8 GB	133d 10:43	1d 14:11	O2PDPSuite::async-async-2025-pp-apass1-v3-slc alidist-async-2025-pp-apass1-v3- O2PDPSuite::daily-20251010-0000	-2,

# MC productions

#### https://alimonitor.cern.ch/MC/

Create Production »		MonteCarlo production reques	ts (show details)				
	Production tag and anchoring			Request tracking		Current status	Events
	Files		- Any - 🗸				
Tag	Jira tickets Git AliEn Anchor prod.	Pass	Collision Energy Rt (GeV) list	n Comment st	PWG	Production Reques	ted Gene
LHC25k4a_2		apass4	Pb-Pb 5,360	1 - HF c, b enriched CRMode2 embedded (1signal-2bkg embedd pattern) in A			
LHC25k4a_1	😑 🚖 LHC23zzh	apass4	Pb-Pb 5,360	1 - HF c, b enriched CRMode2 embedded (1signal-2bkg embedd pattern) in A			
LHC25k3	🖨 🚖 LHC24ap	apass1	p-p 5,360	1 - Prompt charmonia to e+e- in pp at, (2024 - ppRef anchoring), O2-6433			
LHC25k2		apass4	Pb-Pb 5,360	1 - Prompt charmonia to e+e- in PbPb at, (2023 - pass5 anchoring), O2-6430			
LHC25j1b		apass2	Pb-Pb 5,360	1 - General-purpose, test for ITS perPrimaryVertexProcessing anchored to			
LHC25h5_pp_residuals_yesdist		apass1	p-p 13,600	2 - General Purpose, anchored to apass1 of 2025 data, low-B data, TPC co			1
LHC25h5_pp_residuals_nodist		apass1	p-p 13,600	2 - General Purpose, anchored to apass1 of 2025 data, low-B data, TPC co			2
LHC25h5_nodist_yescorr_v2_oldtag	☐ ☐ LHC24ar	apass2	Pb-Pb 5,360	1 - General Purpose, anchored to apass2 of 2024 data, TPC correction map			
LHC25h5_dist_v2_oldtag		apass2	Pb-Pb 5,360	1 - General Purpose, anchored to apass2 of 2024 data, TPC correction map			
LHC25j9		apass7	p-p 13,600	1 - Non-prompt psi(2S) -> 3/psi pi pi in pp at, (2022 - pass7 anchoring)			
LHC25j7		apass4	Pb-Pb 5,360	1 - Prompt Jpsi and psi(2S) in Pb-Pb collisions at forward rapidity (202			
LHC25h5_nodist_yescorr_v2		apass2	Pb-Pb 5,360	1 - General Purpose, anchored to apass2 of 2024 data, TPC correction map			
LHC25h5_dist_v2		apass2	Pb-Pb 5,360	1 - General Purpose, anchored to apass2 of 2024 data, TPC correction map			
LHC25j5		apass5	Pb-Pb 5,360	1 - UD Simulation of forward dimuons anchored to apass5 of Pb-Pb data pe			1
LHC25b8a_v8	⊟ ⊟ LHC24ar	apass2	Pb-Pb 5,360	1 - High occupancy studies anchored to PbPb 2024 apass2, no distortions,			
LHC25j4b		apass1	p-p 13,600	1 - General-purpose, test for ITS deltaROF anchored to apass1 of pp 2025			2
LHC25j4a		apass1	p-p 13,600	1 - General-purpose, test for ITS deltaROF anchored to apass1 of pp 2025			2
LHC25b4b6	⊟ ⊟ LHC24a, LHC24q	apass1	p-p 5,360	1 - General-purpose, anchored to apass1 of pp ref 2024 data, prod-v13, L			88
	<u> </u>		F 0.00				

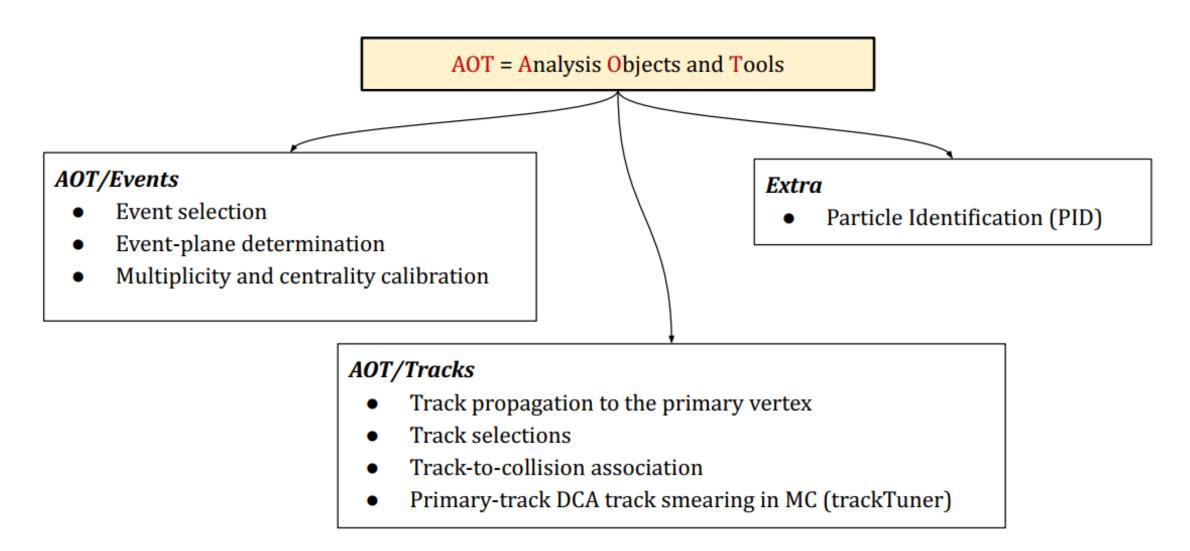


- General-purpose MC productions: managed by the DPG (one per collision system and reco pass)
- PWG-specific productions: requested by PWGs and processed after approval of the Physics Board
  - no PB approval needed if expected CPU time < 1 day at 10k cores</li>

# Two-tag mechanism in MC

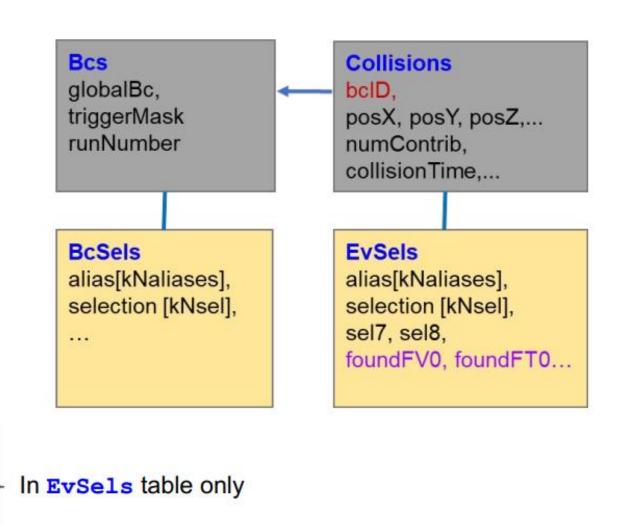
- Finished and consolidated the 2-tag software approach of O2DPG:
  - 1 tag used for generation, GEANT and workflow logic (updated)
    - Stable tags for simulation released ~every 4 weeks
    - Naming scheme O2PDPSuite::MC-prod-2025-vXX
    - Stable MC Software Releases can be found at: https://aliceo2group.github.io/simulation/docs/mc-software-releases/#mc-software-releases
  - 1 tag used for reconstruction (fixed)
- 2-tag approach is converging and showing its benefits for software management, some bug fixes needed though

# Analysis objects and tools



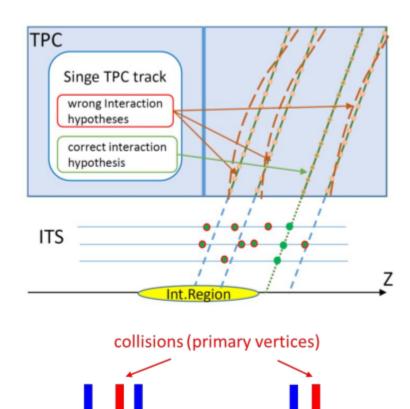
### AOT – event selection

- EvSels table joinable with Collisions table. To be used in analyses based on loops over Collisions (primary vertices), i.e. majority of ALICE analyses.
- BcSels table joinable with BCs table. To be used in analyses based on loops over BCs table such as muon arm UPCs, luminosity monitoring etc.
- Main contents:
  - aliases [kNaliases]: fired trigger aliases (trigger classes)
  - selection [kNsel]: decisions on single selection criteria
  - sel7, sel8 (historical names): selection decisions = logical AND of several selection criteria
  - foundFV0, foundFT0, foundBC: indices to FV0, FT0 and BC entries matched to current collision



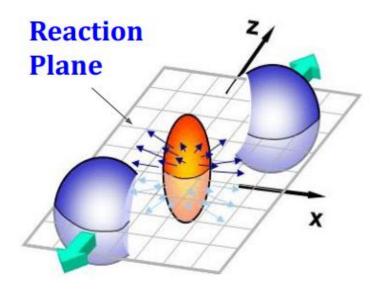
# AOT – event selection challenges in Run 3

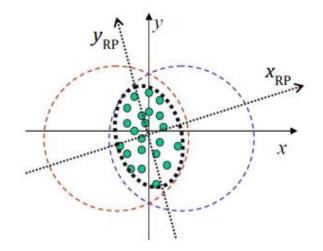
- TPC: tracks are drifting towards endcaps (~100 μs drift time)
  - z-time ambiguity for TPC-only tracks
- ITS integration time ~5 μs in pp (15 μs in Pb-Pb)
  - Several overlapping events (in 650 kHz INEL pp runs)
  - No precise timestamp
- z-time ambiguity for TPC tracks can be resolved via:
  - ITS-TPC matching  $\rightarrow \sim 100$  ns resolution
  - $\circ$  TOF matching → precise timing (resolution < 1 ns).
- Collision time uncertainty depends on
  - time resolution of single tracks
  - number of contributors
- Event selection challenge: most probable collision bc is not precise and might be shifted wrt bc with corresponding FIT signals
  - Might be a problem in high-rate environment
  - (e.g. typical distance between collisions in high-rate pp ~40 bcs)
- Solution: search for FIT info in neighbouring bcs and provide foundBC, foundFT0, foundFV0 indices + flags, trigger aliases and decisions corresponding to foundBC



19

# AOT – event plane





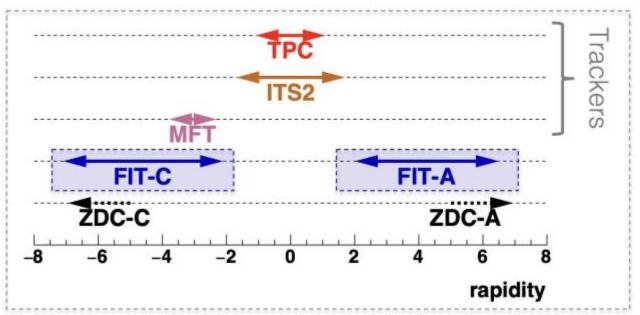
- In heavy-ion collisions:
  - colliding Pb nuclei can overlap partially
  - geometrical anisotropy → momentum anisotropy
  - origin of azimuthal anisotropy
- In anisotropic flow analyses:
  - Reaction plane → plane formed by the impact parameter vector b and the collision axis
  - In practice we cannot find RP but can only approximate and this approximation is known as Event Plane
  - Event Flow Vector " $Q_n$ " and Event Plane Angle  $\psi_n$  from the  $\underline{n}^{th}$  harmonics are defined as:

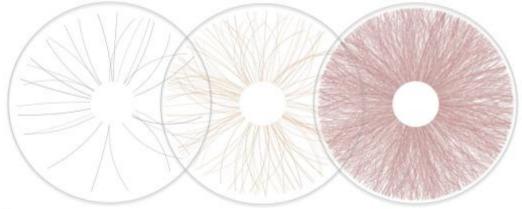
$$Q_{n,x} = \sum_{i} \omega_{i} \cos(n\varphi_{i}) \quad Q_{n,y} = \sum_{i} \omega_{i} \sin(n\varphi_{i})$$
  
$$\Psi_{n} = (1/n) \arctan(Q_{n,y}/Q_{n,x})$$

where  $\omega_i$  is weights and  $\phi_i$  is particle's azimuthal angle

# AOT- multiplicity/centrality

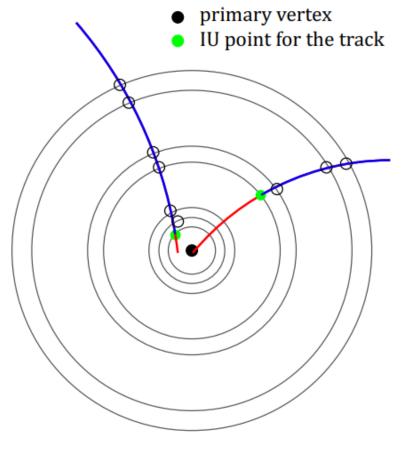
- Many analyses need the event multiplicity/centrality:
  - Study an observable as a function of multiplicity/centrality
  - Other tasks depends on multiplicity/centrality selection (e.g. PID, Q-vector, ...)





- Detectors used for multiplicity/centrality:
  - $\circ$  FT0: -3.3 <  $\eta$  < -2.1, 3.5 <  $\eta$  < 4.9
  - $\circ$  FV0: 2.2 <  $\eta$  < 5.0
  - $\circ$  FDD: -6.9 <  $\eta$  < -4.9, 4.7 <  $\eta$  < 6.3
  - Central barrel detectors → number of tracks used to fit the primary vertex (N\_PV<sub>tracks</sub>)

# AOT - track propagation to the PV



IU tracktrack propagated to PV

- In AO2Ds: TracksIU
  - IU: Innermost Update point
  - Track-parametrization at the IU written in AO2Ds
     → not the same radius for all tracks ↔ track.x()
  - Table written in the AO2Ds
- In analysis: Tracks
  - Track-parametrization after the propagation to the distance of closest-approach (DCA) to the primary vertex
  - In the workflow: dca<sub>XY</sub> and dca<sub>Z</sub> calculated as well
  - on-the-fly by the track-propagation workflow

#### processStandard

- track parameters
- dca<sub>xy</sub>, dca<sub>z</sub> values

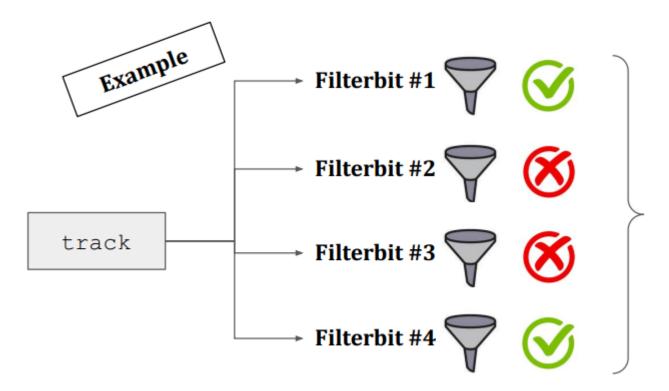
process functions available for analysis

#### processCovariance

- track parameters and covariance matrix
- dca<sub>XY</sub>, dca<sub>Z</sub> values and uncertainties
- → much more resources consumed!

### AOT – track selector

- For each (<u>propagated!</u>) track the following check is done: does it satisfy the selections defined in the *i*-th predefined set (filterbit)?
- Track-by-track filling of aod::TrackSelection table, according to the responses

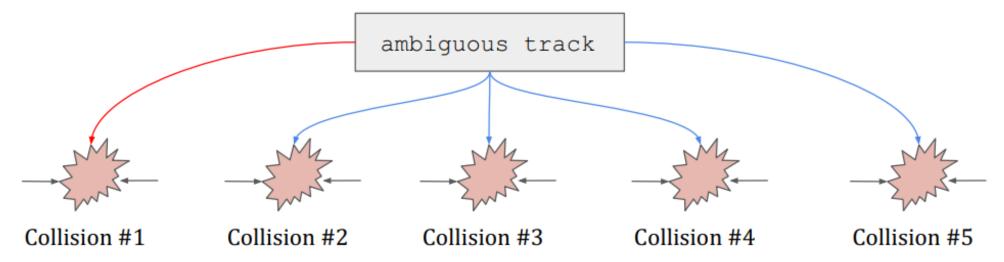


- This table contains for each track the flag for each filterbit
- To have a flag for each single cut: use the aod::TrackSelectionExtension table (not filled by default!)

aod::TrackSelection

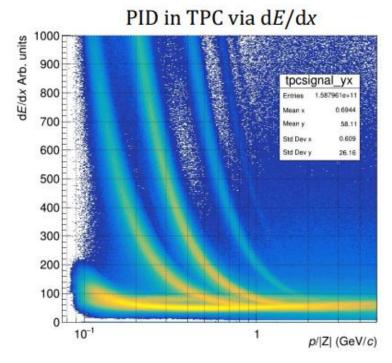
	FB #1	FB #2	FB #3	FB #4
track0	<b>⊗</b>	8	8	8
track1	8	<b>⊗</b>	8	8
track2	<b>⊗</b>	<b>⊗</b>	<b>⊗</b>	8
track3	<b>⊗</b>	<b>⊗</b>	<b>⊗</b>	<b>⊗</b>
track4	8	8	<b>⊗</b>	8

### AOT - track-to-collision associator



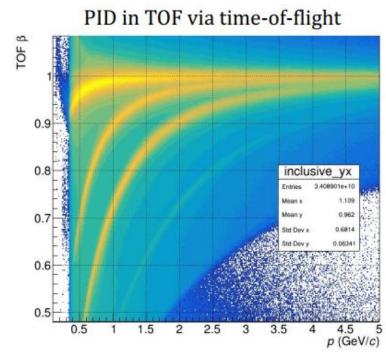
- Continuous readout → ambiguous tracks: tracks with more than 1 collision possible
- By <u>default</u>, in the <u>AO2D</u> the track.collisionId() is that of the <u>first compatible coll</u>ision in time
- track-to-collision-associator: duplication of the track to each collisions compatible in time
  - candidates!
    - → duplicates removable with analysis selections (topology)
  - 🙁 single-track analyses: signal duplication
    - → duplicates removed with specific analysis selections? To be studied in analysis

### AOT - PID in TPC and TOF



- Different particle species have separate distributions of dE/dx (TPC) and time-of-flight (TOF)
- In analysis: n<sub>g</sub> values are used

Tools/parameterizations provided by TPC, TOF experts (not by DPG!)



#### TPC

- Calibrations provided via Bethe-Block (BB) parametrizations and via Neural Network (use "NN" when available, list here)
  - In MC the dE/dx observed in data is reproduced by random-sampling the dE/dx parametrizations from data → "tune-on-data"
    - o  $n_{\sigma}$  centered at 0 and with  $\sigma=1$

# Asynchronous QC

- First validation of asynchronous reconstruction processing
  - Crucial for the assessment of the data quality and reconstruction performance
  - Detector-level tasks integrated and running steadily
- Same software as used at P2, but configuration modified to fit the async reconstruction
  - Including checks and new features
- Possibility to run post-processing tasks including trending
- Mainly run-by-run checks, including comparison of reconstruction passes...
  - Feedback should result in flagging the data
- ...and data vs MC
  - To validate MC settings, anchoring
- Weekly meetings on Tue at 3 pm (ZOOM meeting) and Friday 3 pm (minutes-only virtual meeting); daily meeting during Pb-Pb and light ion runs at 15:30
  - PWGs are warmly invited to attend and also welcome to report at async-QC meeting with material aimed at defining the data quality
  - Differential studies (e.g. vs IR, vs eta...) more suited to AOT meetings on Thu morning at 9:30 am

# Analysis QC

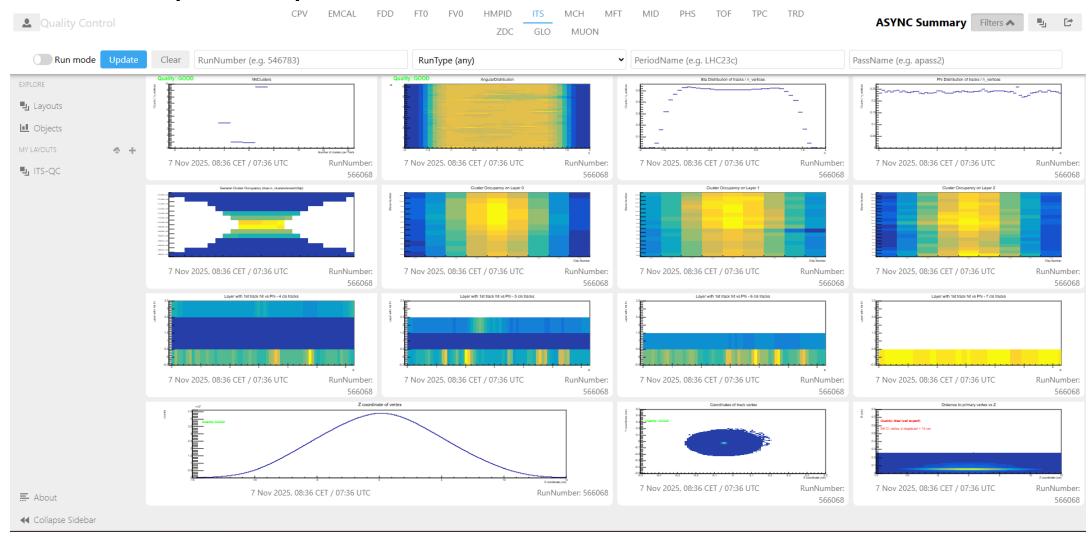
- After reconstruction, some basic analysis are run centrally in a children job
- They use the same sw as the reconstruction
- Very useful for QC
- Every PWG can add its own analysis and report at the aQC meeting (Tue 3 pm)

#### 

Most importantly the tasks from companies need to be put in a list called tasks. This will be translated into the common command-line

#### Analysis QC manual:

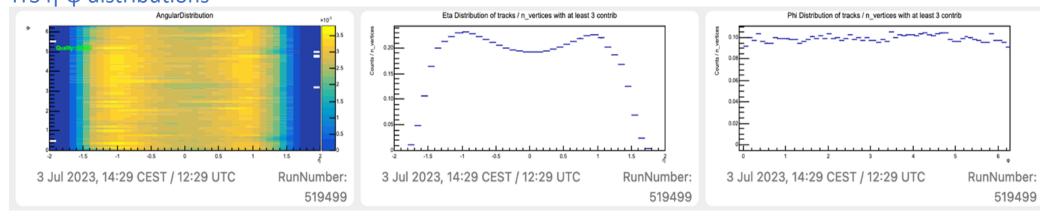
# QC GUI (QCG)



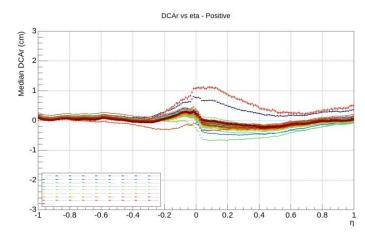
Example layout: <a href="https://ali-qcg.cern.ch/?page=layoutShow&layoutId=67c062a100b1fb4fb47985a9&tab=ITS">https://ali-qcg.cern.ch/?page=layoutShow&layoutId=67c062a100b1fb4fb47985a9&tab=ITS</a>

## QC from detectors

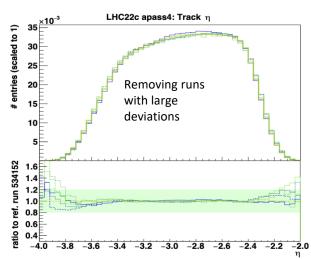
#### ITS $\eta$ - $\varphi$ distributions



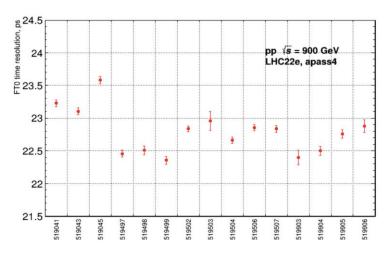
#### **TPC DCAr**



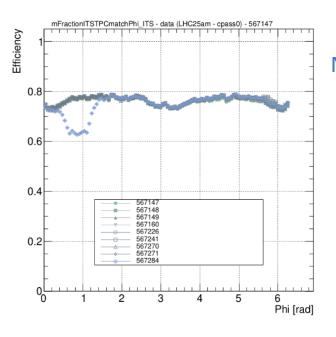
#### MFT tracks vs n



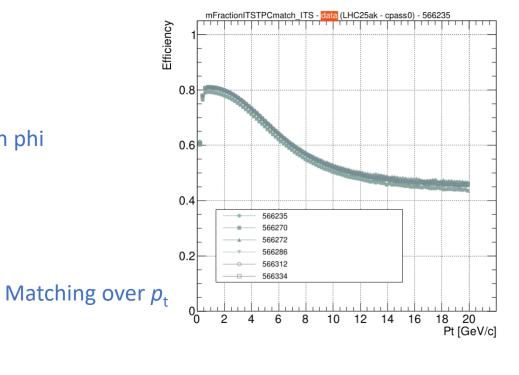
#### FT0 resolution

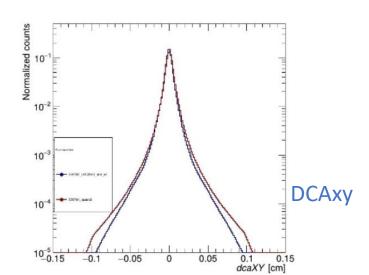


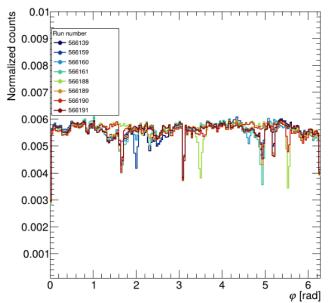
## QC from AOT



Matching in phi





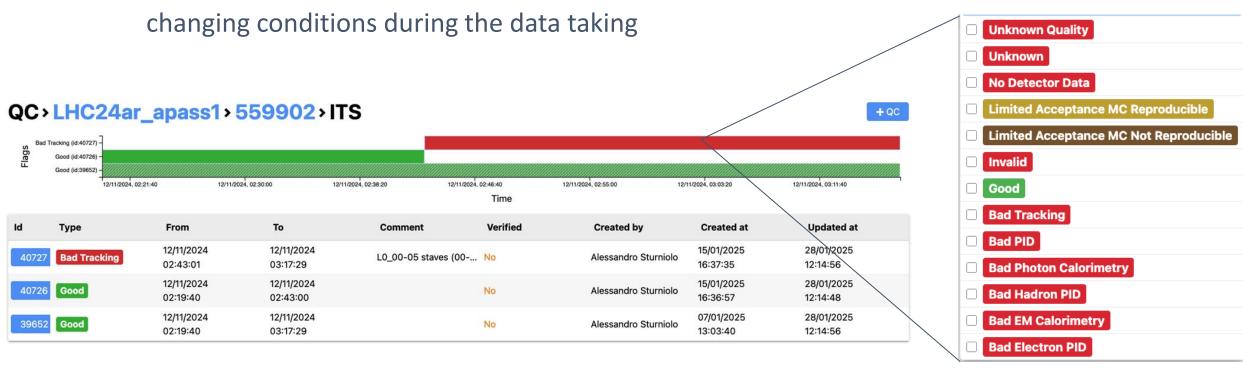


Tracks in phi

## The Run Condition Table (RCT)

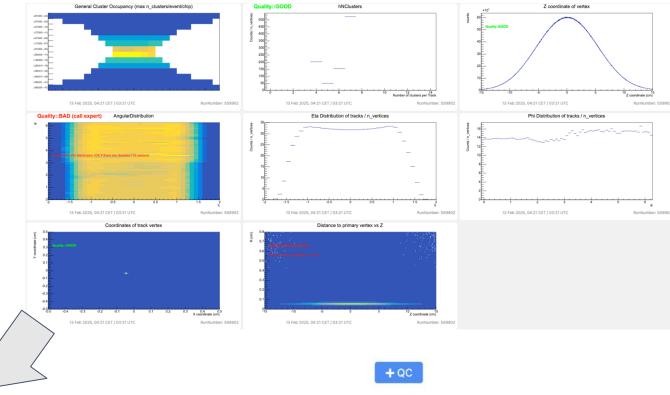
RCT in Bookkeeping: <a href="https://ali-bookkeeping.cern.ch/?page=lhc-period-overview">https://ali-bookkeeping.cern.ch/?page=lhc-period-overview</a>

- Database that contains the quality flags of each detector participating in the data taking
- There is one table for each data taking period and data processing pass (including online)
- The table contains a set of quality flags for each run and each detector
- The quality flags are time-dependent:
  - A single run can have one or multiple flags for each detector, to reflect possibly

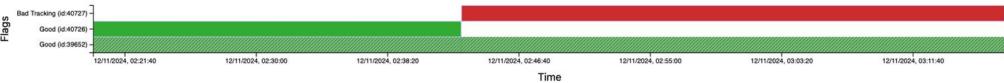


## From QC to RCT flags

- The time-dependent RCT flags are mainly set from the analysis of the QC plots
- Other sources of information can be used as well (ex. DCS)



#### QC>LHC24ar\_apass1>559902>ITS

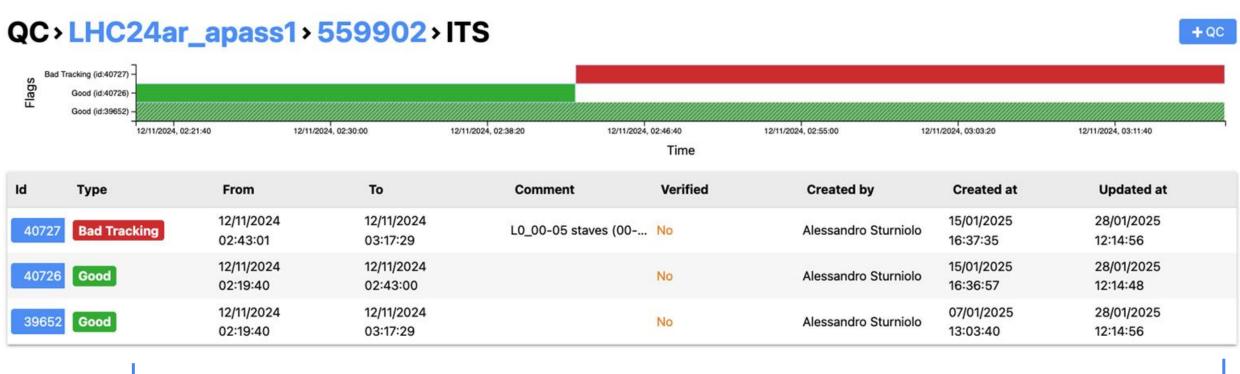


ld	Туре	From	То	Comment	Verified	Created by	Created at	Updated at
40727	Bad Tracking	12/11/2024 02:43:01	12/11/2024 03:17:29	L0_00-05 staves (00	No	Alessandro Sturniolo	15/01/2025 16:37:35	28/01/2025 12:14:56
40726	Good	12/11/2024 02:19:40	12/11/2024 02:43:00		No	Alessandro Sturniolo	15/01/2025 16:36:57	28/01/2025 12:14:48
39652	Good	12/11/2024 02:19:40	12/11/2024 03:17:29		No	Alessandro Sturniolo	07/01/2025 13:03:40	28/01/2025 12:14:56

- The official DPG DataSets are based on the detector quality flags set in the <u>RCT</u>
- The DPG DataSets include:
  - o CentralBarrelTracking (CBT): requiring FT0, ITS, TPC each with quality status Good or Limited Acceptance MC Reproducible
  - o **CBT\_hadronPID**: FT0, ITS, TPC, TOF each with quality status Good or Limited Acceptance MC Reproducible
  - o CBT\_electronPID: FT0, ITS, TPC TOF, TRD each with quality status Good or Limited Acceptance MC Reproducible
  - o CBT\_calo: FT0, ITS, TPC, EMC each with quality status Good or Limited Acceptance MC Reproducible
  - CBT\_muon: FT0, ITS, MCH, MID with quality status Good or Limited Acceptance MC Reproducible, TPC Good or Limited Acceptance MC Reproducible or BadPID
  - CBT\_muon\_global: FTO, ITS, MCH, MID, MFT with quality status Good or Limited Acceptance MC Reproducible, TPC Good or Limited Acceptance MC Reproducible or BadPID
- For official analyses, only the CBT\* DataSets must be used wherever available!
- The DataSets provide a granularity of a **full run**, hence they do not exploit the full potential of the information stored in the RCT

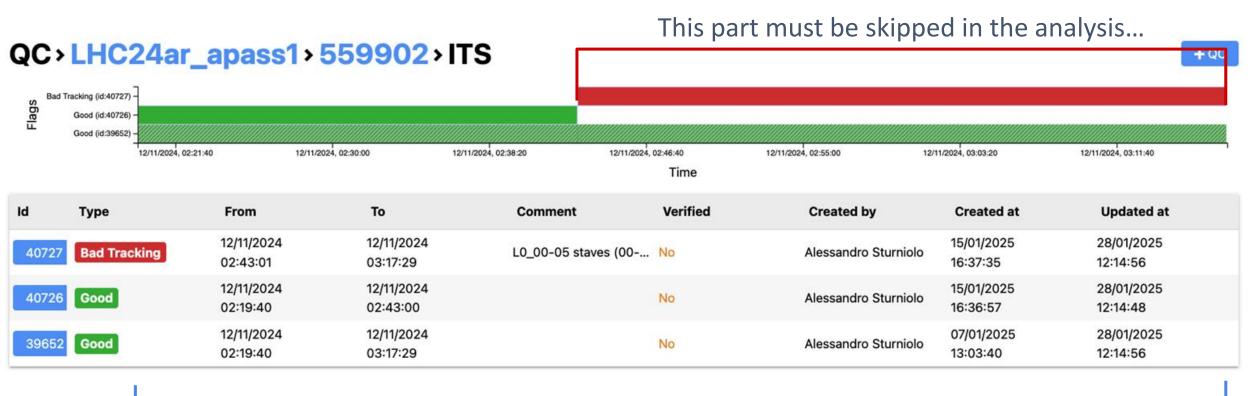
Runlists: <a href="https://twiki.cern.ch/twiki/bin/view/ALICE/AliDPGRunLists\_3">https://twiki.cern.ch/twiki/bin/view/ALICE/AliDPGRunLists\_3</a>
Runlists definitions: <a href="https://twiki.cern.ch/twiki/bin/view/ALICE/AliDPGRunListsDef\_3">https://twiki.cern.ch/twiki/bin/view/ALICE/AliDPGRunListsDef\_3</a>

 The DataSets contain all runs for which at least part of the data is Good or "Limited Acceptance MC Reproducible"



This run is included in all the CBT\* DataSets

 The DataSets contain all runs for which at least part of the data is Good or "Limited Acceptance MC Reproducible"



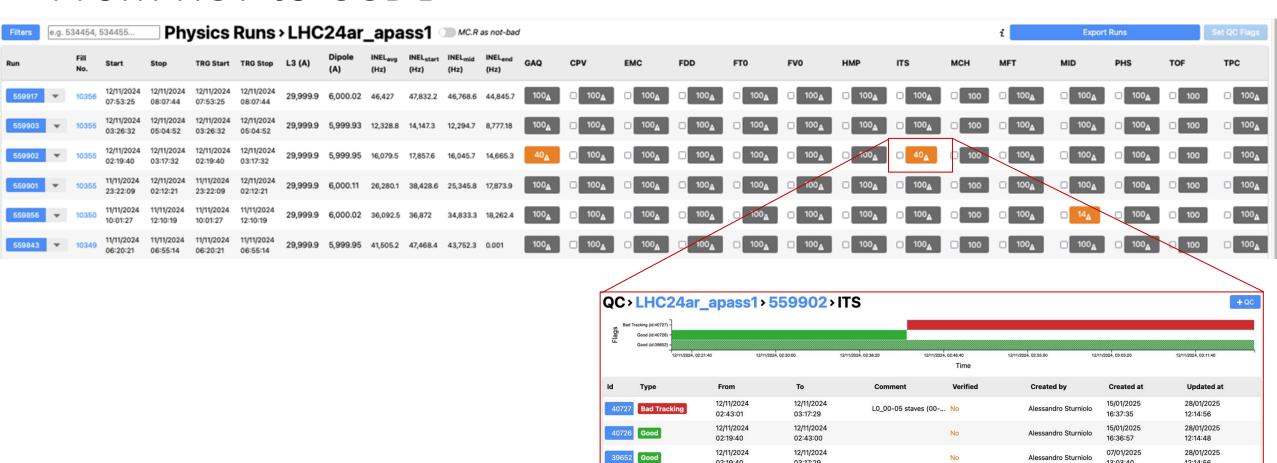
This run is included in all the CBT\* DataSets

 The DataSets contain all runs for which at least part of the data is Good or "Limited Acceptance MC Reproducible"

#### QC > LHC24ar\_apass1 > 560012 > ITS Limited Acceptance MC Reproducible (id:40725) -Flags Good (id:39810) -15/11/2024, 03:00:00 15/11/2024, 03:30:20 15/11/2024, 03:30:20 15/11/2024, 03:00:00 15/11/2024, 03:00 15/11/2024 Time Verified Updated at ld Type To Created by Created at From Comment 15/11/2024 15/11/2024 15/01/2025 15/01/2025 Limited Acceptance Alessandro Sturniolo L5 18 stave went off No 02:55:30 09:38:32 16:35:20 16:35:20 15/11/2024 15/11/2024 09/01/2025 09/01/2025 Good No Alessandro Sturniolo 02:55:30 16:41:07 09:38:32 16:41:07

The "Limited Acceptance MC Reproducible" data are also included in the analysis, since the issues are expected to be reproduced in the MC

#### From RCT to CCDB



02:19:40

03:17:29

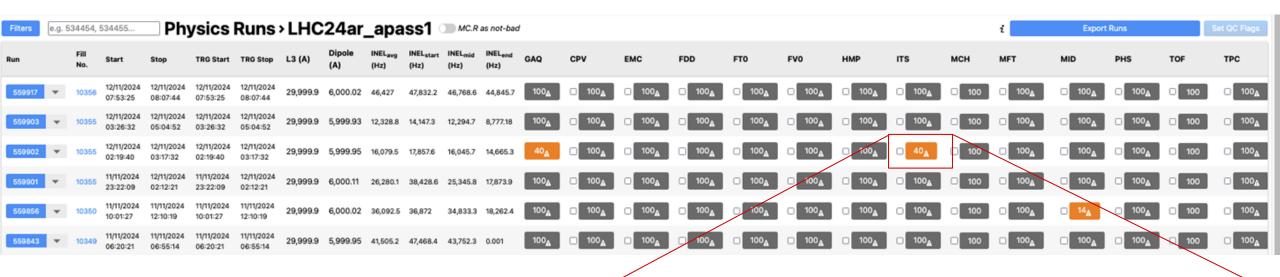
12:14:56

No

Alessandro Sturniolo

13:03:40

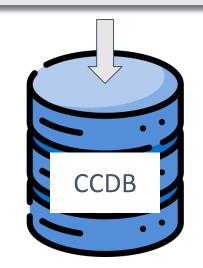
#### From RCT to CCDB

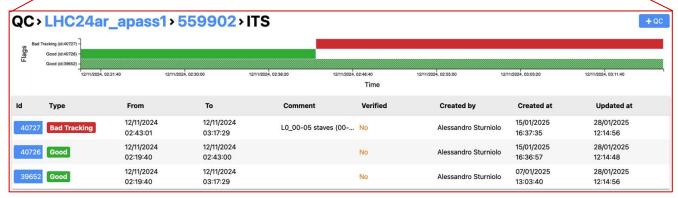


#### RCT CCDB Object

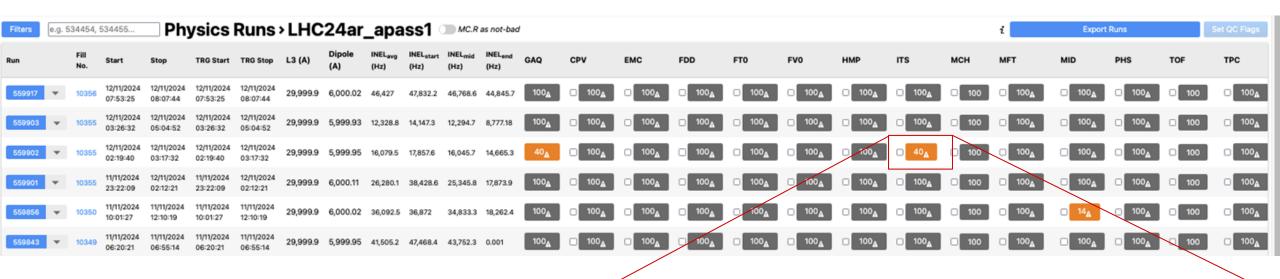
ITS Good from ... to ...

ITS Bad Tracking from ... to ...





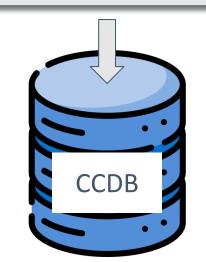
#### From RCT to CCDB



#### **RCT CCDB Object**

ITS Good from ... to ...

ITS Bad Tracking from ... to ...





CCDB objects for all 2023 and 2024 periods available since mid-March.

For 2022 the RCT is only partly filled at the moment.

#### From CCDB to Physics Analysis

- The RCT CCDB objects have been integrated in the EventSelection code of O2Physics
- New column added in bcsels (joinable with bcs) and evsels (joinable with collisions):
  - see <u>Common/DataModel/EventSelection.h</u>
  - DECLARE\_SOA\_BITMAP\_COLUMN(Rct, rct, 32); //! Bitmask of RCT flags
- For every bc and collision, rct column is filled with the 32-bit mask taken from CCDB map:
  - mapRCT = ccdb->getSpecific<std::map<uint64\_t, uint32\_t>>("Users/j/jian/RCT", ts, mdata);
  - mapRCT contains a map of timestamps and corresponding RCT masks
    - typically one entry per run if no detector quality changes
    - new entry for each change of rct state
    - provides fine-grained quality selection on top of Runlists
  - rct mask set to 0 if CCDB object not found (no FATAL)
  - Offcial CCDB path available: RCT/Flags/RunFlags
  - The initial test path, Users/j/jian/RCT, will also be maintained for a while

#### Full list of RCT selection flags

```
// RCT selection flags
enum RCTSelectionFlags {
  kCPVBad = 0.
  kEMCBad.
  kEMCLimAccMCRepr,
  kFDDBad.
  kFT0Bad,
  kFV0Bad.
  kHMPBad,
  kITSBad.
  kITSLimAccMCRepr,
  kMCHBad.
  kMCHLimAccMCRepr,
  kMFTBad,
  kMFTLimAccMCRepr,
  kMIDBad,
  kMIDLimAccMCRepr,
  kPHSBad,
  kTOFBad,
  kTOFLimAccMCRepr,
  kTPCBadTracking,
  kTPCBadPID,
  kTPCLimAccMCRepr,
  kTRDBad.
  kZDCBad,
  kNRCTSelectionFlags
};
```

 The list of flags is available in Common/CCDB/RCTSelectionFlags.h

#### Typical use cases:

- Collisions selection:
  - Copy full mask (e.g. for filtering): uint32\_t rct = collision.rct\_raw();
  - Access single bits: bool isTOFBad = collision.rct\_bit(kTOFBad);
  - Use helper function: see next slides
- BunchCrossing selection:
  - Copy full mask (e.g. for filtering): uint32\_t rct = bc.rct\_raw();
  - Access single bits: bool isTOFBad = bc.rct\_bit(kTOFBad);

- Helper code and classes to interpret the RCT quality bits and select collisions/BCs
  - Enum to provide definition of RCT bits as set in the dynamic tables
  - RCTFlagsChecker Helper class to check a given subset of bits in the RCT columns
  - See Common/CCDB/RCTSelectionFlags.h
- The checker class can be initialized in two ways:
  - 1. With a pre-defined bits selection, identified with a label:

#### For example:

```
RCTFlagsChecker myChecker("CBT hadronPID");
```

Six labels, corresponding to the official Runlists, are currently implemented.

Setting checkZDC to true includes also ZDC flags in the bits selection (for Pb-Pb).

Setting the last parameter to true allows to reject all events where the detectors relevant for the specified Runlist are flagged as LimitedAcceptance

- Helper code and classes to interpret the RCT quality bits and select collisions/BCs
  - Enum to provide definition of RCT bits as set in the dynamic tables
  - RCTFlagsChecker Helper class to check a given subset of bits in the RCT columns
  - See Common/CCDB/RCTSelectionFlags.h
- The checker class can be initialized in two ways:
  - 2. With a custom list of bits, provided as a vector or an initializer list:

```
RCTFlagsChecker(std::initializer_list<QualitySelectionFlags> bitsToCheck)
RCTFlagsChecker(const std::vector<QualitySelectionFlags> bitsToCheck)
```

#### For example:

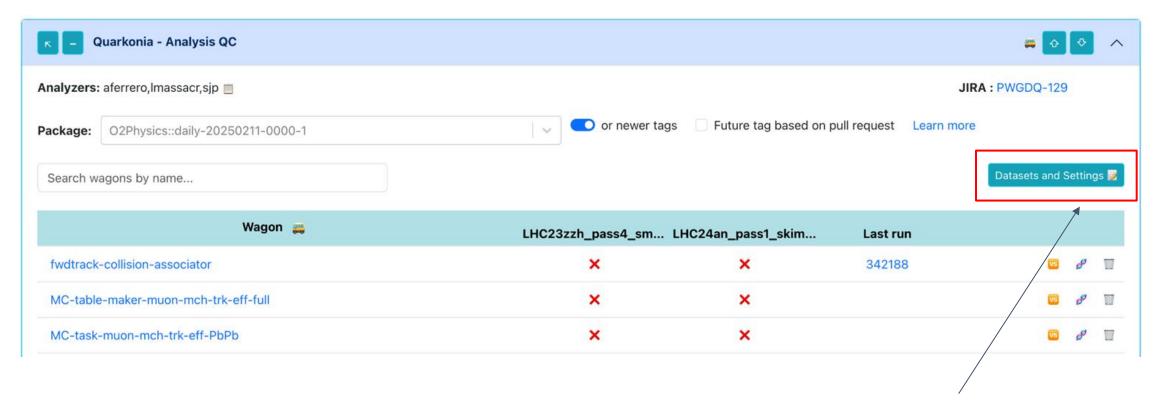
```
RCTFlagsChecker myChecker{kFT0Bad, kMFTBad};
```

- Helper code and classes to interpret the RCT quality bits and select collisions/BCs
  - Enum to provide definition of RCT bits as set in the dynamic tables
  - RCTFlagsChecker Helper class to check a given subset of bits in the RCT columns
  - See Common/CCDB/RCTSelectionFlags.h
- Collisions can then be checked easily:

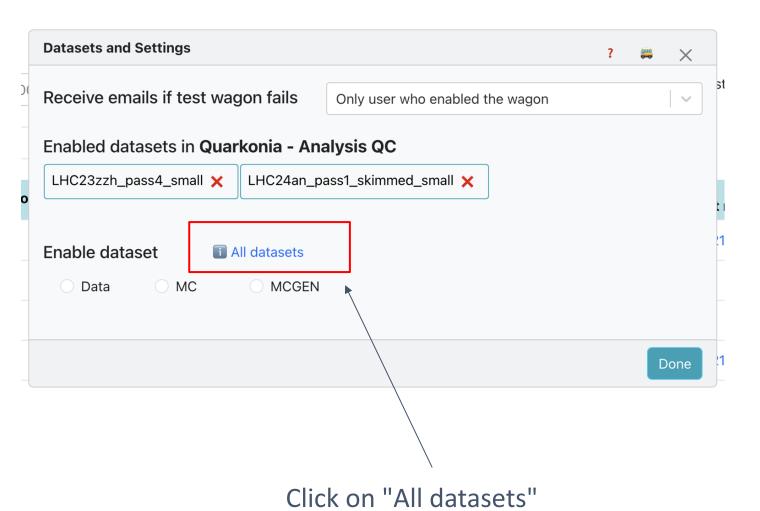
Putting all things together:

```
#include "Framework/runDataProcessing.h"
#include "Framework/AnalysisTask.h"
#include "Common/DataModel/EventSelection.h" 
RCTSelectionFlags.h is already included in EventSelection.h
using namespace 02;
using namespace o2::framework;
using namespace o2::aod::rctsel;
struct myExampleTask {
  // initialization with runlist label
  RCTFlagsChecker myChecker{ "CBT hadronPID" };
  void init(InitContext const&)
    // override initialization with the init() method
   myChecker.init("CBT hadronPID", true);
  void process(soa::Join<o2::aod::Collisions, o2::aod::EvSels>::iterator const& collision)
    // perform the check on the current collision
    if (myChecker(*collision)) {
      // process this collsion
```

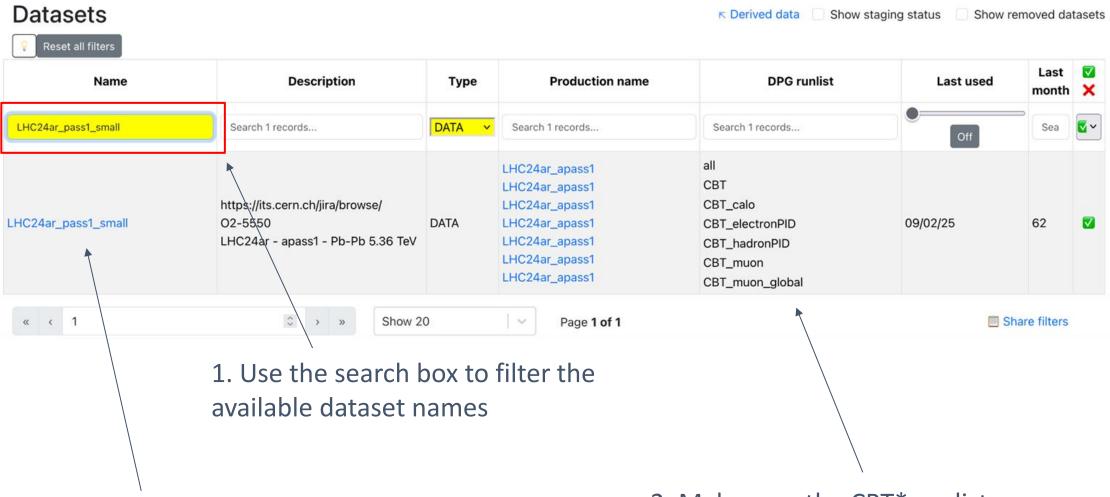
Open your analysis on HyperLoop (<a href="https://alimonitor.cern.ch/hyperloop/user">https://alimonitor.cern.ch/hyperloop/user</a>)



Click here to add a new Dataset

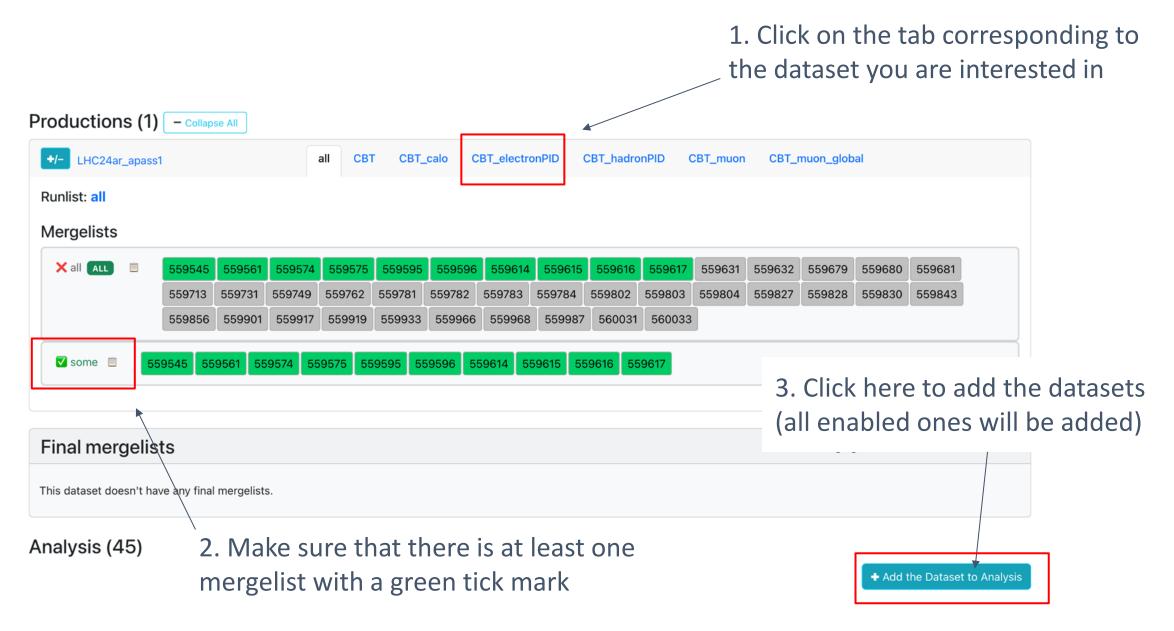


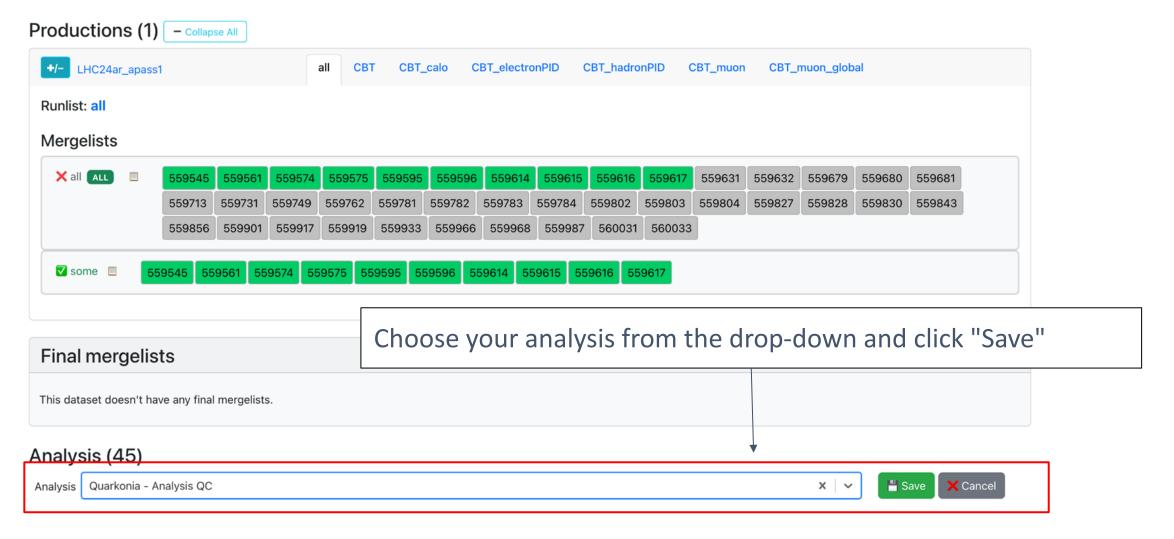
47

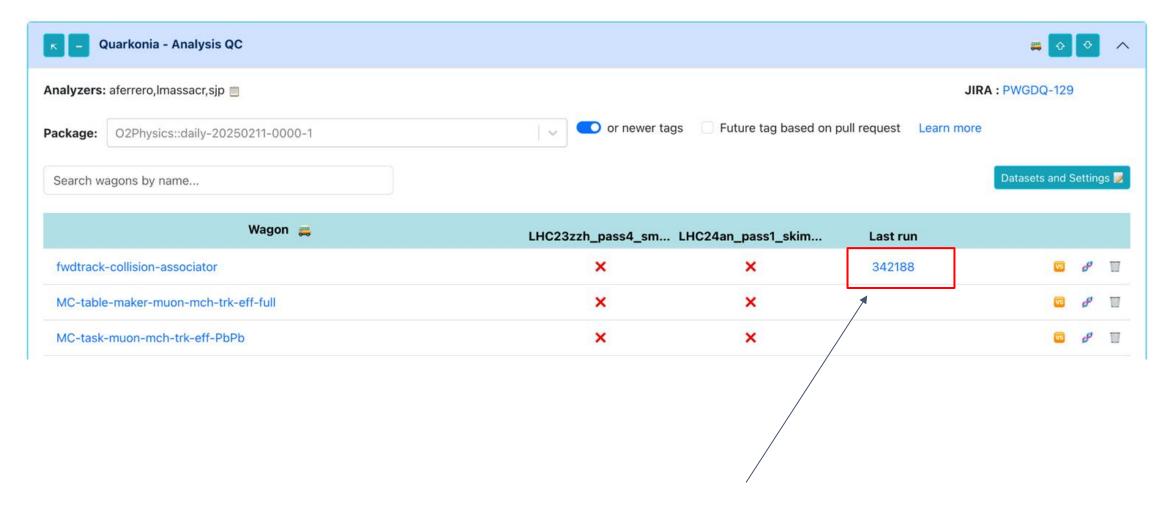


3. Click on the Dataset name

2. Make sure the CBT\* runlist you are interested in is listed here



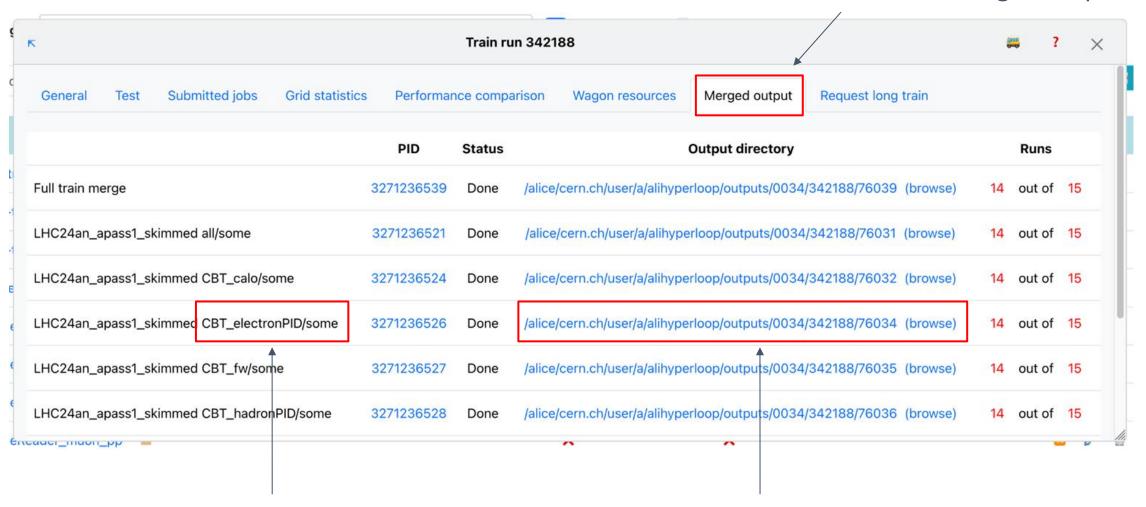




Once completed, you can access the merged outputs bly clicking on the train run ID

The jobs will generate several merged outputs, one for each of the enabled Datasets

Click on the "Merged output" tab



The Dataset name corresponding to a given merged output is visible on the left

Here is the link to the corresponding MonaLisa repository

# Enjoy your analysis!!

08/11/2025 Jian Liu 53

# Backup