Integrating Analysis

MOLLER Experiment Readiness Review

July 30, 2025

Sanghwa Park

(JLab)

Jefferson Lab

U.S. DEPARTMENT OF ENERGY | Office of Science

JSA

Charges

- 14. What is the simulation and data analysis software status for the experiment? **Has readiness for expedient analysis of the data been demonstrated?** What is the projected timeline for the first publication? Please provide, if possible, a documented track record from previous experiments.

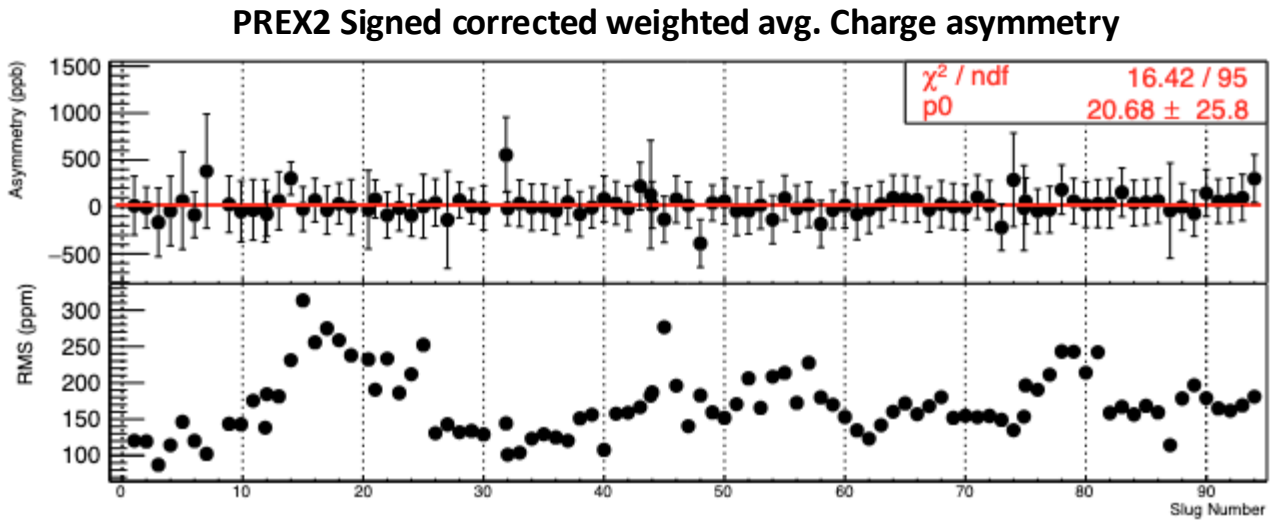
- 15. **What do you plan to have for prompt monitoring of the performance of the experimental apparatus?**

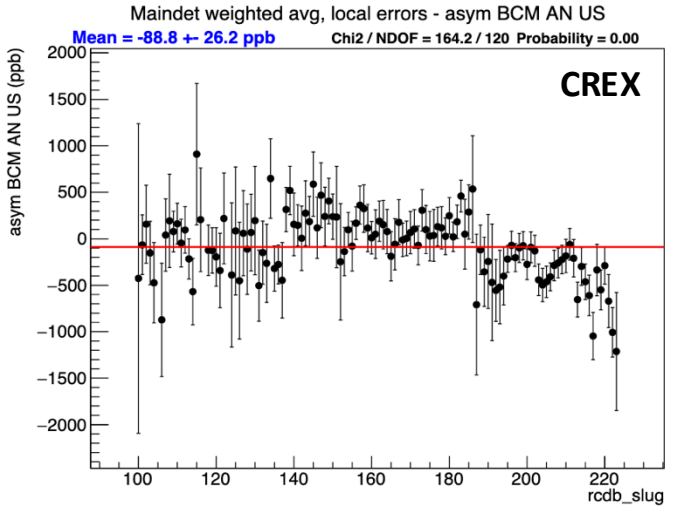
Online feedback system

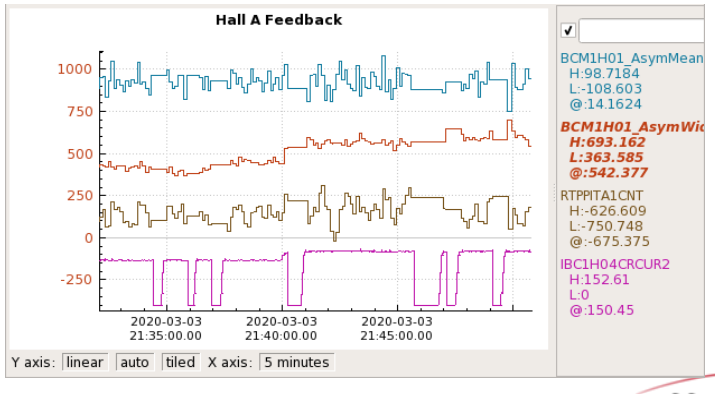
- An independent real-time process
- Online analyzer calculates helicity-correlated (HC) charge asymmetries and position differences
- Determines modified values for INJ control devices (RTP cell, helicity magnets, …) to correct the HC beam parameters to zero and communicates these via EPIC
- Feedback algorithm used in PREX/CREX and Qweak exists and is operational
- Algorithm supports different time-scales for charge and position feedback
- During MOLLER, we will run the charge feedback on both Hall A and C beams

Feedback during PREX2/CREX

- Automated charge feedback cycle in 7.5s intervals
- Also ran feedback on Hall C beam during PREX2/CREX
- Slow, manual corrections done on the position differen

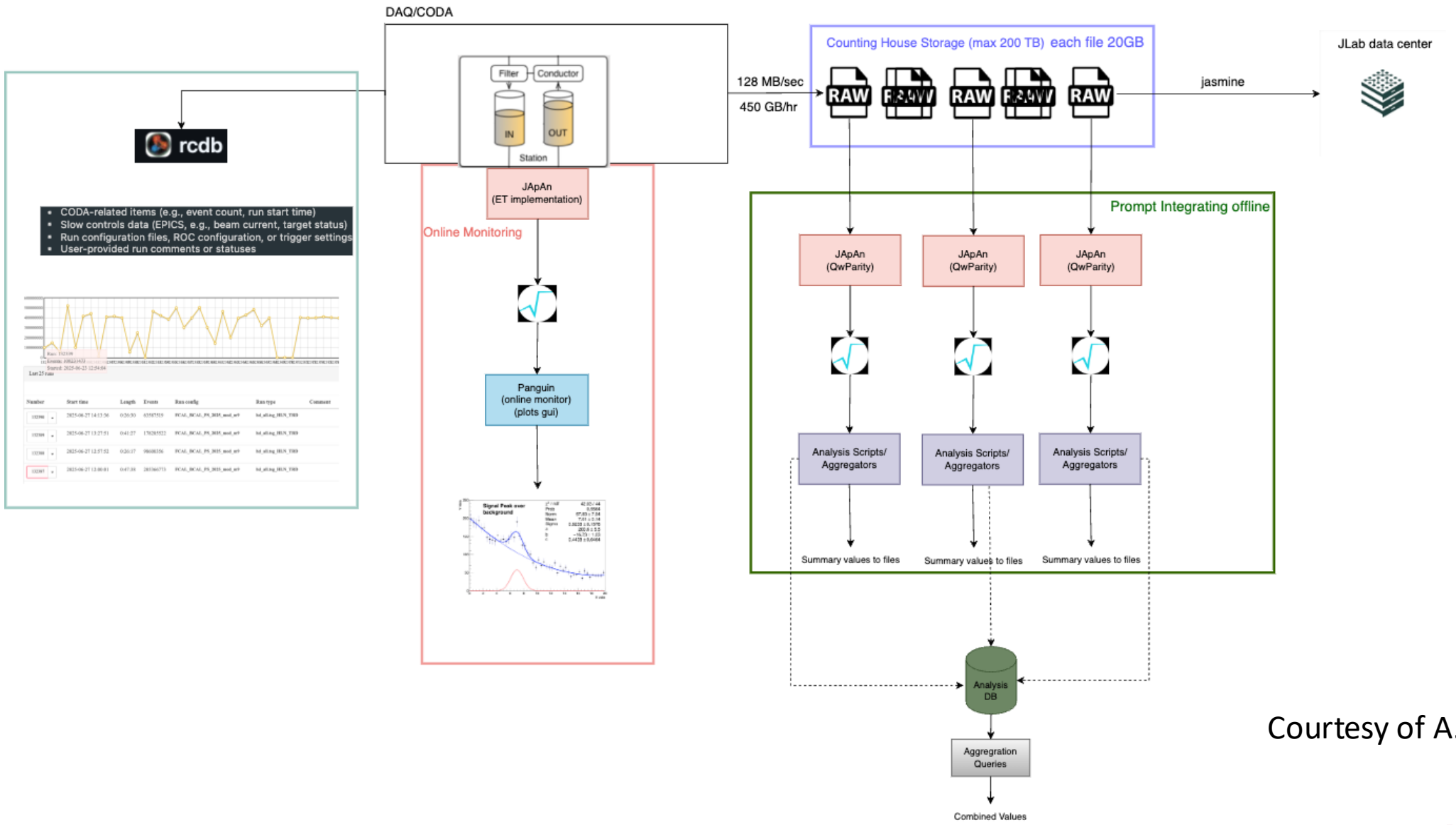






Jefferson Lab

Data flow, online monitoring and prompt analysis

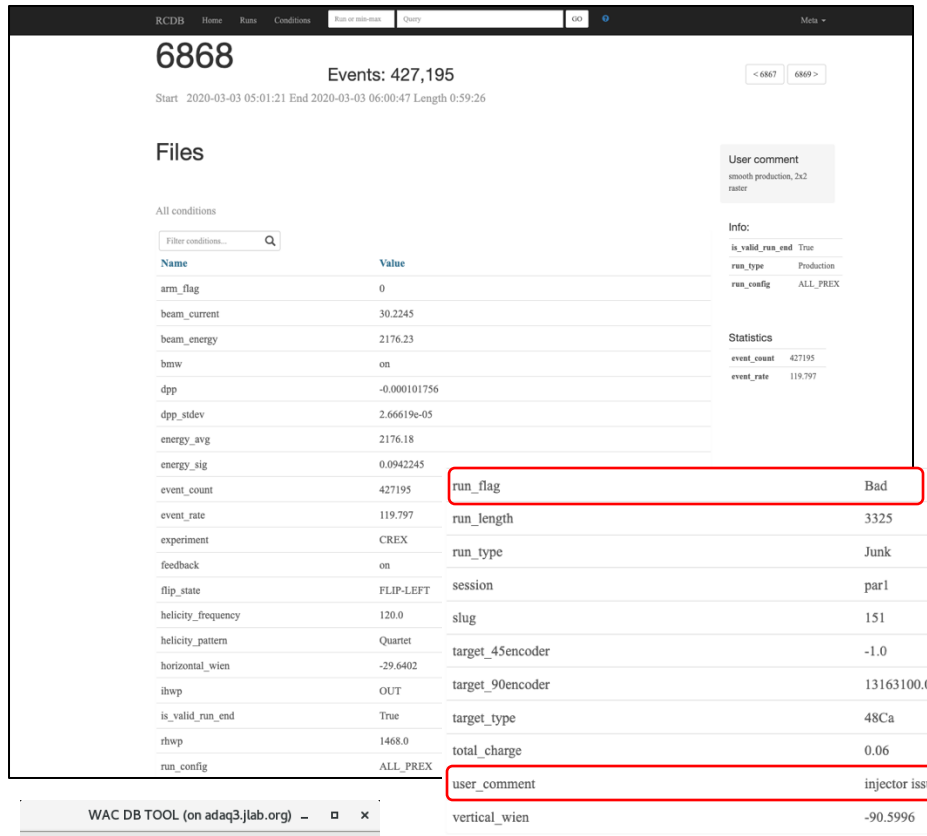


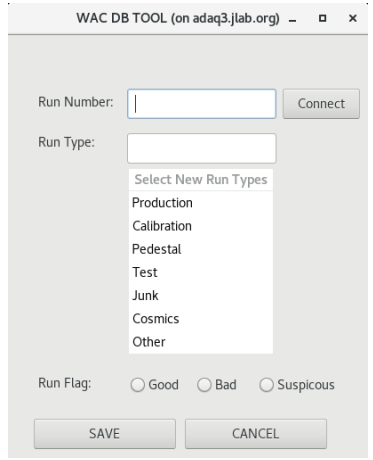
Courtesy of A. Panta

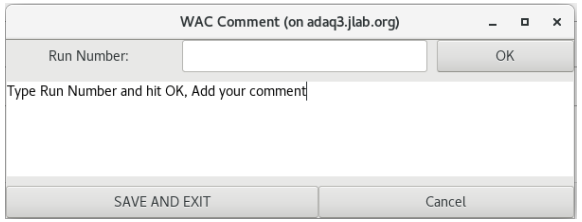
Jefferson Lab

Run Control Database (RCDB)

- Successfully deployed and used for PREX2/CREX
- Store run conditions (slow control data, DAQ configuration and parameters, experiment setting, user comments)
- Can easily add new conditions and modify the values
- Provides initial run QA and classification (IHWP status, slug number)
- Incorporated into the aggregation scripts during PREX2/CREX
- Deployment for a new experiment is straightforward

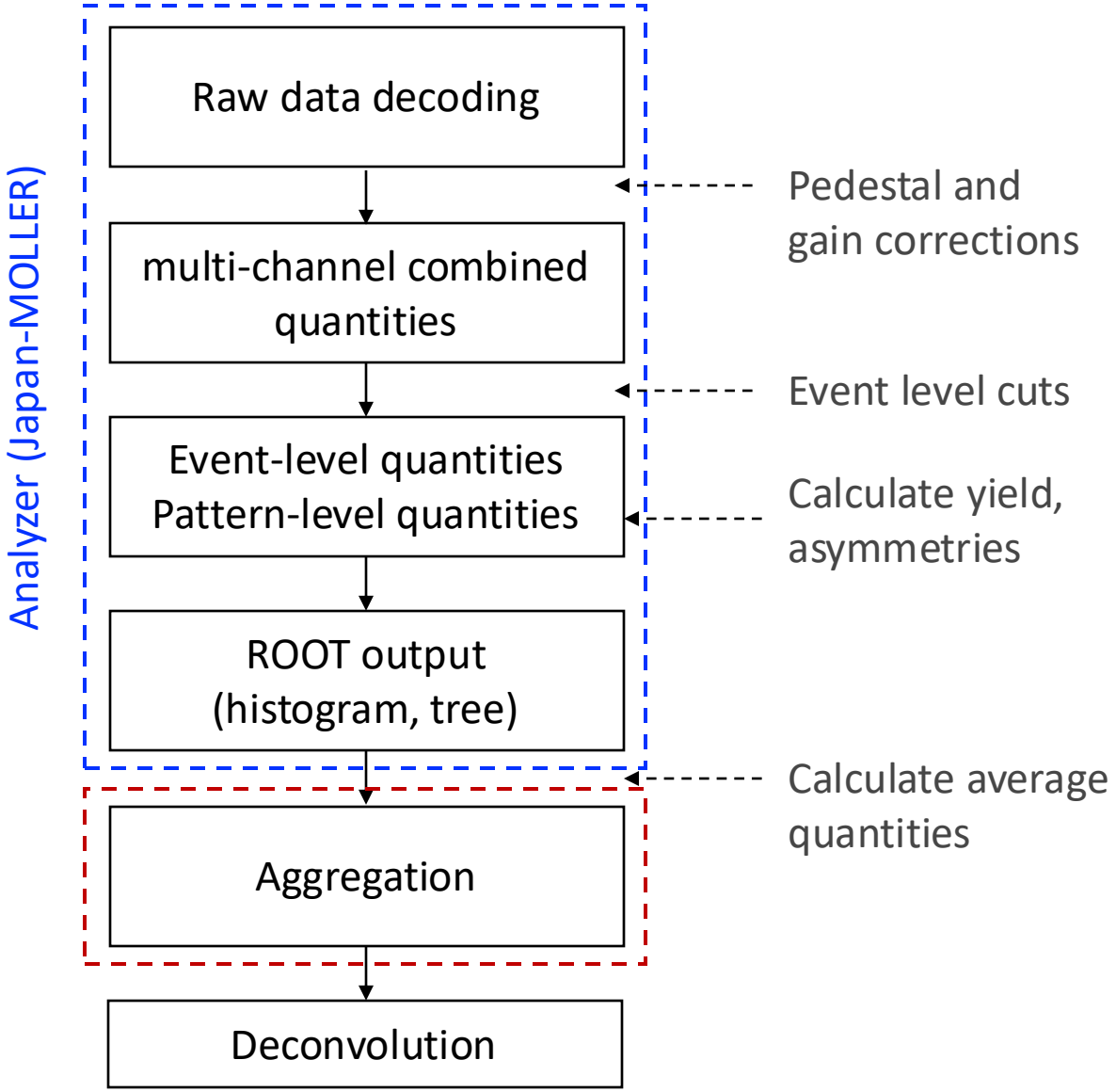






Jefferson Lab

Prompt and full offline analysis process



- Raw data is processed by the analyzer (https://github.com/JeffersonLab/japan-MOLLER)
- Calculates the yields and asymmetries and perform correlation analysis
- Analyze dithering (BMOD) data
- Output ROOT file is produced with histograms and TTrees
- Aggregator combines the outputs for different time periods (run, slug, pitt, grand summary) with calculated average quantities
- Deconvolution analysis tool has been developed and tested with simulation

The software for integrating analysis already exists and has been tested

Jefferson Lab

Prompt analysis and aggregation

Prompt analysis:

- Runs on data files shortly after the data is taken
- Parallel processing of analysis jobs: a run is split into segments, allows a quick check on the regression analysis

Aggregation:

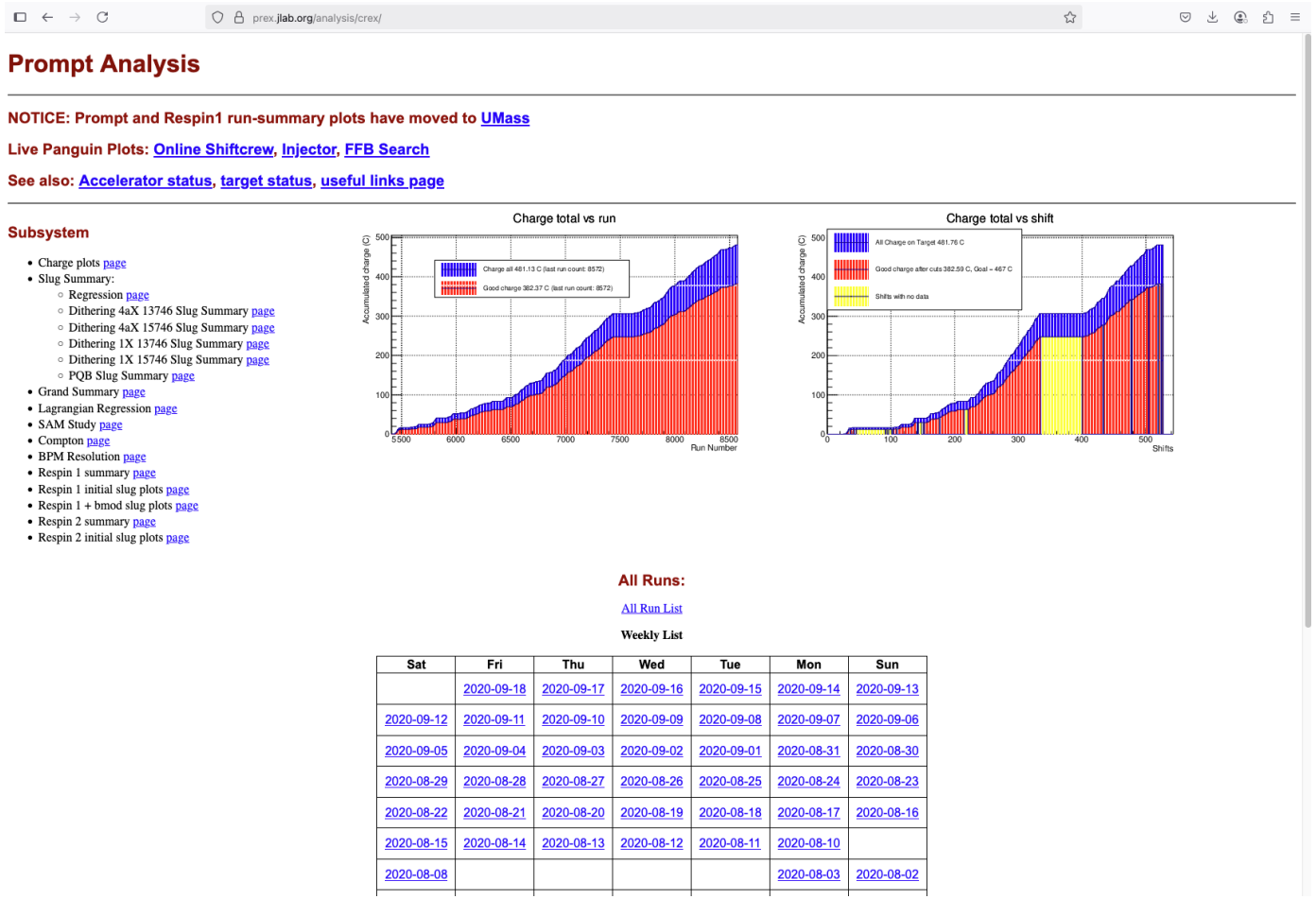
- Prompt analysis will produce several data output files
- Aggregator will collect and combine them to display the results for different time periods (minirun, run, day, slug, pitt, grand summary)
- Previous parity experiments:
  - QWeak: MySQL based analysis database
  - PREX2 and CREX: ROOT files

Prompt analysis during PREX2/CREX

- A run was split into miniruns with each minirun containing about 9000 multiplets (~5 min of data with good events)
- Miniruns with the same IHWP status were combined into ”slug” (~8h of data)
- Used ROOT file format for aggregation
  - Only small number of detectors and monitors involved, output size was not a concern
  - More careful thoughts need to be given for MOLLER
- Data quality was monitored and discussed, led by a Weekly Analysis Coordinator
  - Initial data QA done during experiment running
  - Great training opportunity for students
- Summary plots were posted to analysis webpage

PREX2 and CREX

- Prompt analysis public page for PREX2 and CREX: https://prex.jlab.org/analysis/



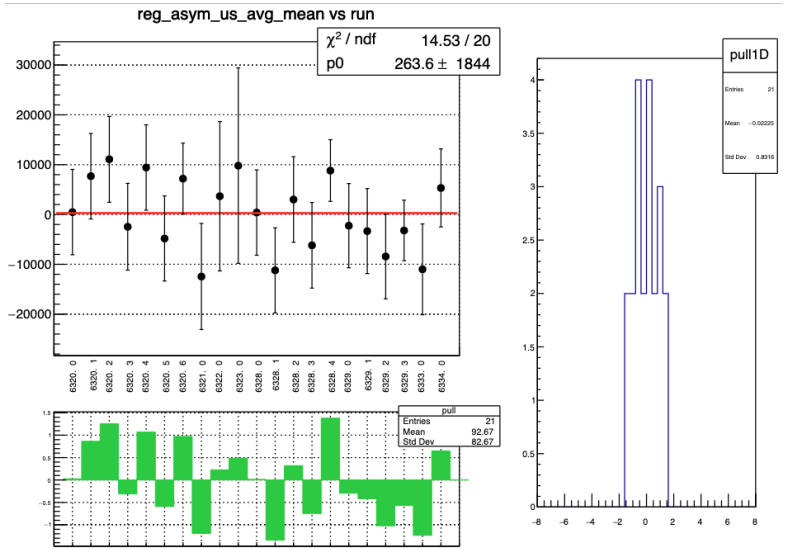
- Automated scripts to run the analysis and post the summary plots into the public webpage

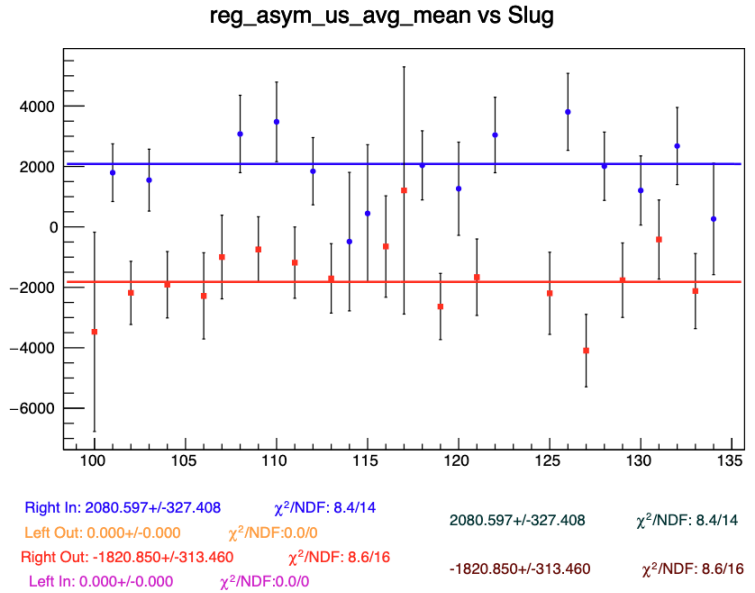
  - Live status
  - Charge accumulation tracker
  - Summary:
    - PDF, csv, ROOT file
    - Mini run, slug, daily, grand summary
    - Regression and dithering analysis
    - Compton polarimetry analysis
    - Other misc studies (SAM, BPM, ..)
    - Full offline analysis

PREX2 and CREX

Slug Summary

- slug100: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug101: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug102: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug103: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug104: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug105: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug106: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug107: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug108: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-108.txt grand_100-108_linear.txt
- slug109: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug110: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug111: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug112: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug113: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug114: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-114.txt grand_100-114_linear.txt
- slug115: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-115.txt grand_100-115_linear.txt
- slug116: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug117: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug118: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug119: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug120: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug121: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug122: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug123: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug124: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug125: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-125.txt grand_100-125_linear.txt
- slug126: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-126.txt grand_100-126_linear.txt
- slug127: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug128: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-128.txt grand_100-128_linear.txt
- slug129: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-129.txt grand_100-129_linear.txt
- slug130: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug131: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's | Extras: grand_100-131.txt grand_100-131_linear.txt
- slug132: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug133: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug134: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's
- slug135: All - Run List - Total Q/Slug Q - Minirun Summary (Text file, csv) - Grand (Signed, rootfile, textfile, csv) | Dithering: Slopes - Day 1 set α's & δ's





Prompt analysis for MOLLER

- Data rate is expected to be around 130 MB/s
- A run will be split into segments: file-split size of 20 GB (~150 seconds)
- Prompt analysis will run using CH computers (plan to purchase more)
- Full offline analysis will use the JLab farm resources
- MOLLER will have much larger and complex dataset:
  - ~400 channels (detectors and beam parameters) and their correlations
  - The total raw data volume will be at the petabyte scale
  - Creating ROOT TTrees of all elements would also result in petabyte scale output
- Several parallel data reduction choices
  - Average yields and asymmetries for all data elements over short period and store in a database
  - Generate ROOT TTrees for only the key elements for all helicity patterns
  - Keep the full ROOT TTrees for a small subset of data files
- Development and testing of analysis and monitoring tools is ongoing with mock data and simulation

Integrating analysis tasks and workforce

- Ongoing tasks:
  - Output data structure and optimization
  - Aggregation software
- Using mock data and simulation for testing of throughput and developing analysis tools
  - Mock-data generator within japan allows benchmarking of analyzer performance with a model data-stream similar to that expected from the full DAQ
  - Currently adding further complexity to the mock-data to develop the tools which will be needed by shift crews and experts

- Analysis working group:
  - Weekly meeting, Monday 11:30am
  - https://moller.jlab.org/wiki/index.php/Analysis_Software
  - Peoplepower: S. Chatterjee, P. King, O. Hansen, B. Waidyawansa, A. Panta, A. Sen, J. Shirk, D. VanGilder, W. Deconinck, Michael, K. Paschke, P. Gautam, D. McNulty, D. Spayde, H. Liu, R. Conway, I. Akinyemi, S. Sarker, J. Mammei, S. Regmi, B. Uduwara, Xiang, Lasitha, J. Poudel, J. Roche, ...

Summary

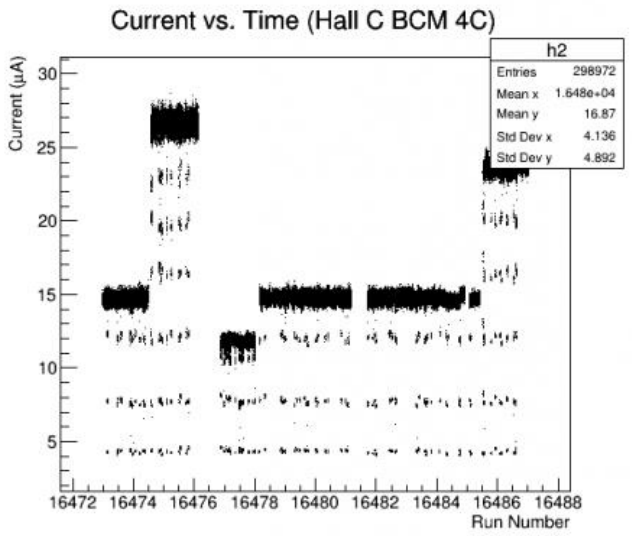
14. What is the simulation and data analysis software status for the experiment? **Has readiness for expedient analysis of the data been demonstrated?** What is the projected timeline for the first publication? Please provide, if possible, a documented track record from previous experiments.

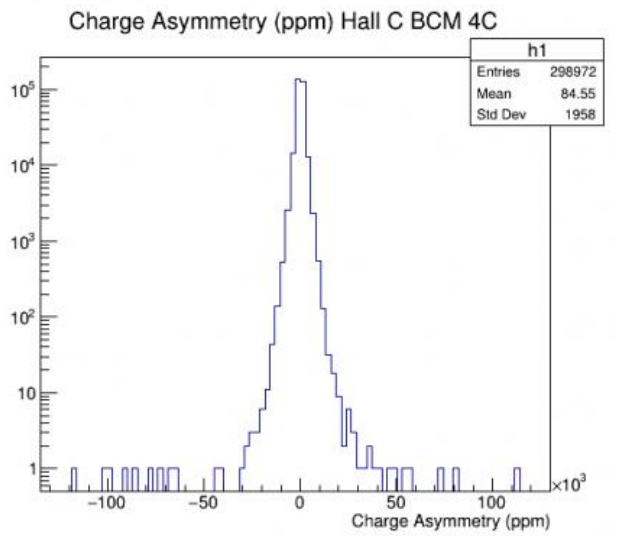
15. **What do you plan to have for prompt monitoring of the performance of the experimental apparatus?**

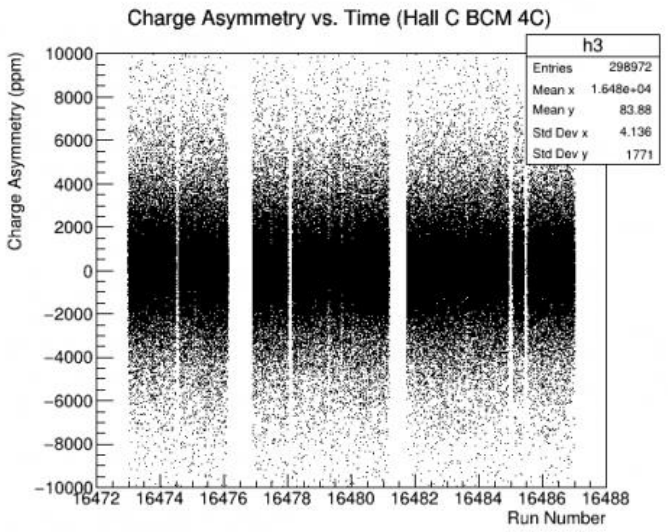
- Prompt analysis will be performed shortly after data is taken. Parallel processing of segmented data files will allow a quick analysis turnout and prompt monitoring of the data quality
- The software for integrating analysis already exists and has been tested
- Aggregation software is being developed with expected completion of Summer 2026
- Online feedback system is used to minimize helicity-correlated beam asymmetries. The system has been successfully used for previous parity experiments and other Hall A/C experiments until recently. No further development is required.

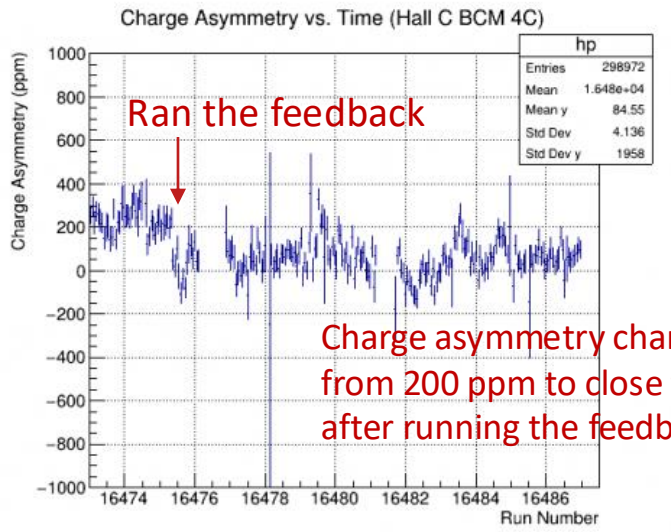
Backup

Charge feedback during NPS in Hall C









- The feedback system has been used in Hall A and C even after PREX2 and CREX experiments
- Controlled by the Hall A Parity DAQ for both Hall A and Hall C beams
- Feedback is not running constantly, but being turned on briefly to bring the average close to zero

Analysis software overview

Analysis software is not in the project scope, but impacts decisions such as the computing hardware needed for the online analysis.
Major development started last year, but building on prior experiments

Integrating analysis framework

- Building on the “japan” (Just Another Parity Analyzer) framework used by PREX/CREX and Qweak
- Mock-data generator within japan allows benchmarking of analyzer performance with a model data-stream similar to that expected from the full DAQ
- Currently adding further complexity to the mock-data to develop the tools which will be needed by shift crews and experts

Counting analysis framework

- Starting with a version of the Hall A “Podd”-based analysis package used by SBS
- The FADC decoding classes are in use in the counting test-stands in the Test Lab and at William and Mary
- The GEM decoding classes have been tested with data from the GEM test stand
- Continuing to adapt detector array and track-finding classes for the MOLLER detectors and geometry

Integration analysis process in japan-MOLLER

1. Decode raw data and apply pedestals and gain corrections
2. Form multi-channel combined quantities, such as beam positions
3. Apply event-level cuts. Use recent history of events (the “event ring”) to apply beam trip cuts and stability cuts
4. Use data exchange classes to calculate new event-level composite quantities (e.g., position corrected single-event yields, event-level correlation matrices)
5. Collect the events for a complete pattern; calculate the yield and asymmetries for all quantities. Blinding is applied during the asymmetry calculation.
6. Use data exchange classes to calculate new pattern-level composite quantities (e.g., position corrected asymmetries, asymmetry correlation matrices).
7. ROOT histogram and tree outputs can be selected at event-level, pattern-level, or minirun-level
8. If there have been sufficient patterns accumulated to complete a minirun, calculate and store the average quantities and rms.
9. At the end of a file/run calculate and store the average quantities and rms

Data rate, file segment sizes, and data set size

- The standard data rate for the integrating DAQ will be about 130 MB/s
  - The data rate is independent of beam being on or off;
  - One hour of data would be about 470 GB
- A convenient size for the data file segments might be 20 GB → ~150 seconds
  - Hour-long runs would be split into about 23 segments

- The total data footprint for the entire experiment is 14PB
  - Production + Compton raw data: 9.5PB (main production DAQ, 8PB; Compton, 1.5PB)
  - Production + Compton analysis results: 2PB
  - Tracking raw data & analysis results: 1.5PB (raw data, 1.25PB; analysis results, 0.25 PB)
  - Simulation: 1PB

Mock-data and analysis testing

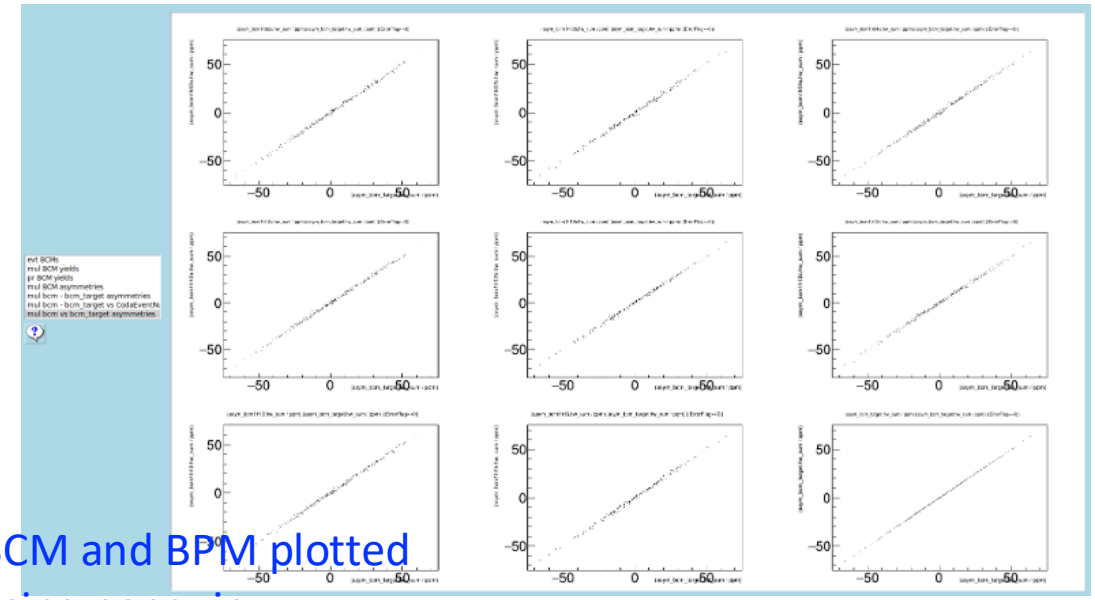
- Mock-data generator within “japan” creates realistic data files with time-dependent and randomized generation of beam parameters and detector signals, including correlations

- First model of the integrating ADC data-structure has been added; needs to be updated as we become familiar with the actual data stream
  - Recently we have added detector rates for each thin-quartz tile from simulation
  - Near future goal: add asymmetries and beam parameter correlations for each tile

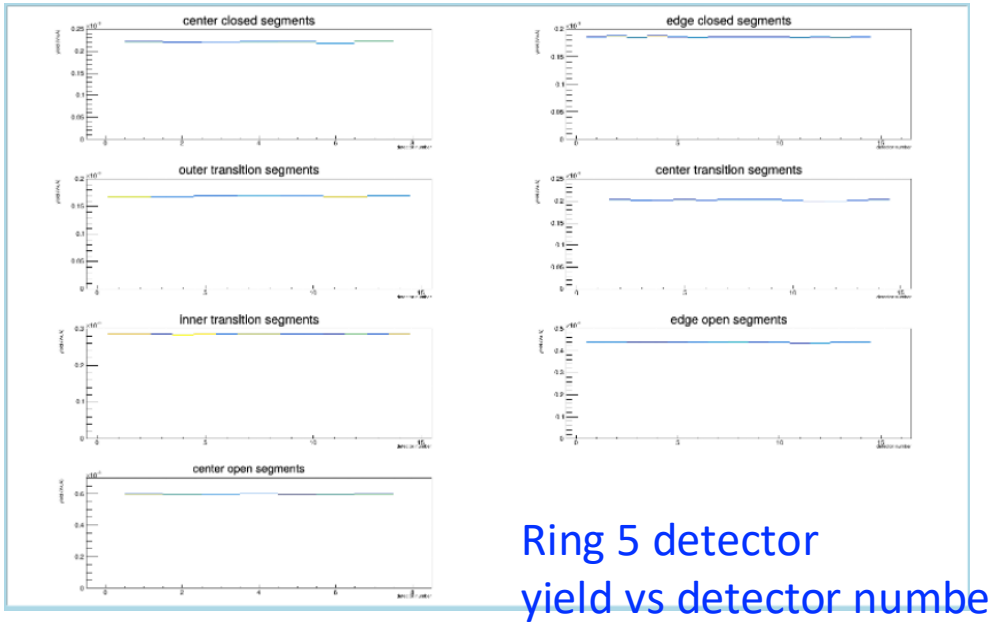
- Analysis of these data files allows testing of throughput and processing algorithms

Analysis using mock-data

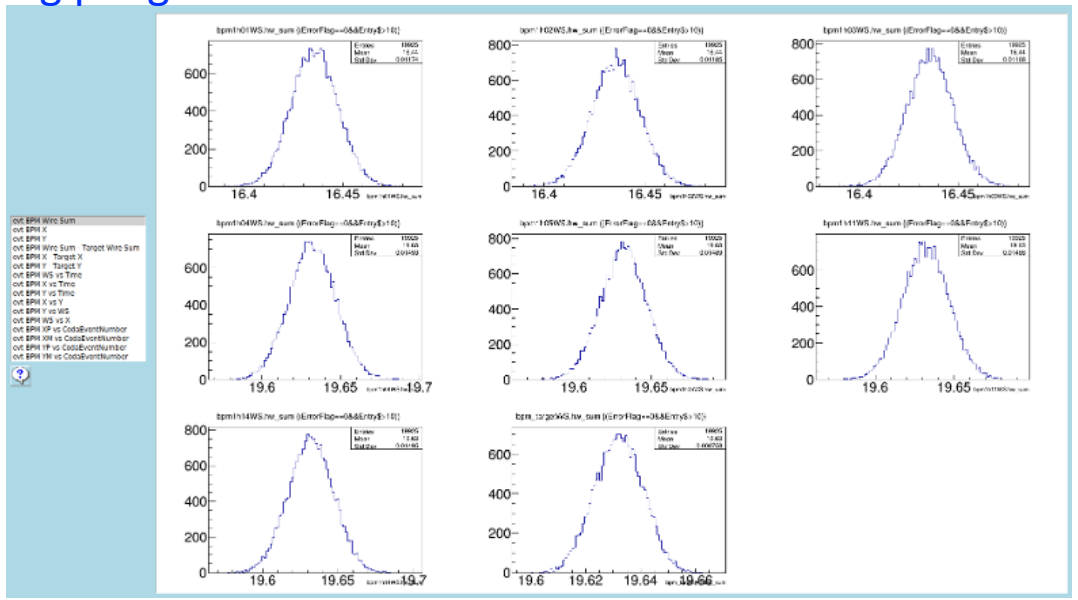
Donald VanGilder (Ohio Univ.)

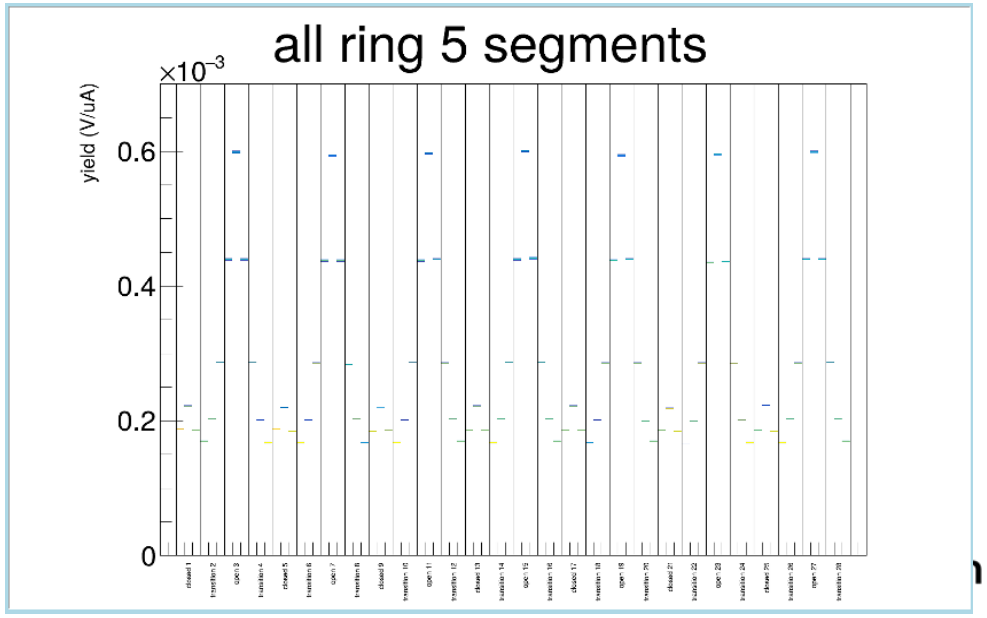


BCM and BPM plotted using panguin

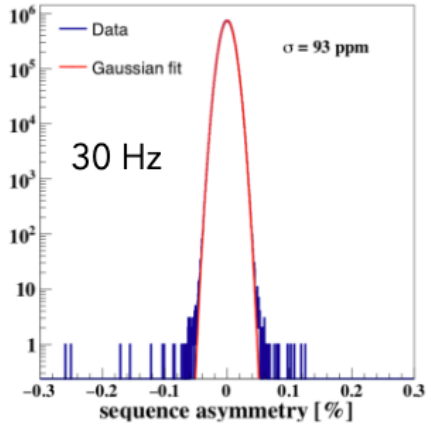


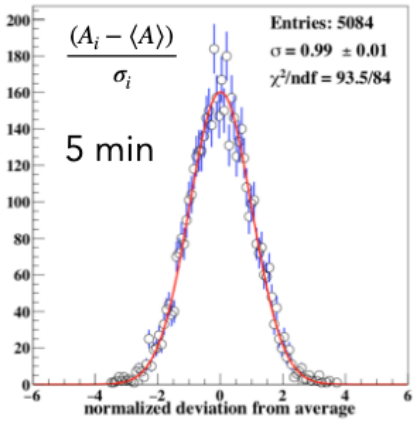
Ring 5 detector yield vs detector number

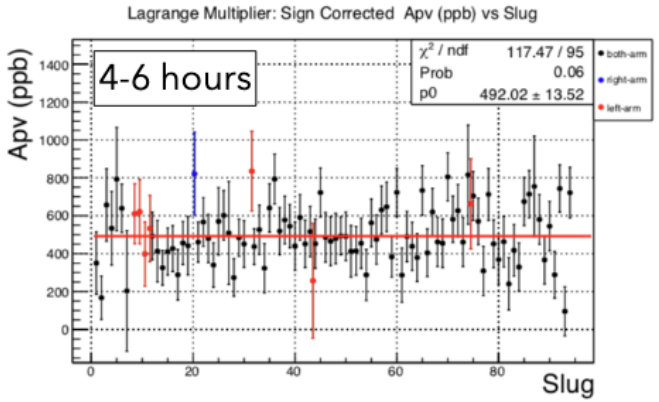


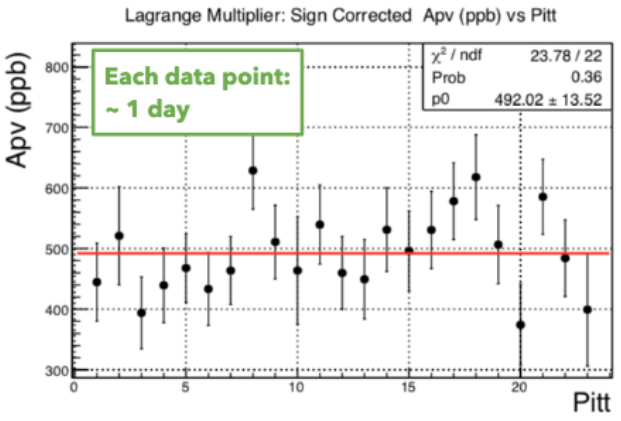


Aggregation and slow reversals









PREX data aggregated →

(noise) (systematics)

		~20PAC days PREX-2 (achieved)	~40PAC days CREX (achieved)	344PAC days MOLLER (required)	14PAC days Cumulative Helicity Correlation (Run1)
Aq	Intensity asymmetry	25 ppb	−88 ppb	10 ppb	< 40 ppb
ΔE/E	Energy asymmetry	0.8 ± 1 ppb	0.1 ± 1.0ppb	< 1.4 ppb	< 6 ppb
Dx	position differences	2.2 ± 4 nm	-5.2 ± 3.6nm	0.6 nm	$< 4 \times 10^{-9}$ m
Δθ	angle differences	$< 0.6 \pm 0.6$ nrad	-0.26 ± 0.16nrad	0.12 nrad	$< 0.5 \times 10^{-9}$ radian
A_σ	size asymmetry (quoted)	$< 3 \times 10^{-5}$	$< 3 \times 10^{-5}$	$< 10^{-5}$	$< 10^{-5}$

Did we achieve HCBA goals? We can't tell without aggregating to achieve precision

- Aggregated results must be tested for statistical consistency, unexpected correlations, etc.
- Systematic tests are more precise with longer integrate times, but then reduce the number if independent periods to compare…
- Studies with aggregated data are an important part of the MOLLER data evaluation (**S. Park**)