Report

Readiness Review for pass1 processing of the HallB/CLAS12 RG-B Winter 2020 data set

Monday, January 25, 2021

Review committee:
Nathan Baltzell
Marco Mirazita
Cole Smith
Stepan Stepanyan (chair)
Larry Weinstein

The review of the readiness of RG-B to process the January 2020 data set took place on January 22, 2021, in a remote setting over the bluejeans. The meeting agenda and presentations can be found on the review page:
https://clasweb.jlab.org/wiki/index.php/Run_Group_B - tab=Pass-1_review__28Winter_29

This data is the third and the last data set from the run group that will be processed to provide the full available statistics for the ongoing analysis. We should note that this will not be the final processing of this data set. The final round of data processing will happen after the ongoing work on software improvements and the tracking alignment will be complete.

The review committee thanks the RG-B team for preparing the presentations, providing ancillary information, and patiently answering our questions during the review. Below are our answers to the Review Charge questions and some comments and recommendations. The RG-B team has done a great job with calibrations and understanding the data set. The group can proceed with the processing of data after addressing the two recommendations.

Review Charge:

Charge #1: *Is the quality of detector calibration and alignment adequate to achieve the performance specifications foreseen for CLAS12 or achievable at the current time, given the "state of the art" calibration, alignment and reconstruction algorithms?*

No – While most of the detectors are calibrated close to the achievable performance at the current time, there is still room for improvements (see below).

Comments:

1. The forward DC has not been calibrated for this run period. Calibration constants from RG-K have been used instead, resulting to a larger spread in means and wider widths

of residuals. This is a re-occurring situation with DC calibration and should be fixed for the final processing (pass2) of any data set. We believe, no matter how small changes in the calibration constants are, all the detectors must be calibrated using the run(s) of the given run period.
At the closeout, we have been informed that DC was calibrated overnight using one of RG-B runs. The review committee will appreciate to see a new timelines of residuals (mean and widths).

2. The energy loss calibrations of FTOF and CTOF have not been presented but were discussed. We think it is important for RG-B to have the energy calibration and the energy clustering in the reconstruction software for these detectors to be complete for the final, pass2, processing of the data. The energy loss is an important measurement to identify deuterons in the studies of coherent scattering reactions.
3. While we understand the RICH reconstruction is still somewhat ad hoc, the presented Cherenkov angle vs. momentum distributions for negatively and positively charged tracks had visible differences. For positively charged tracks, points for K+ and p lined up above calculated lines, while for K-, albeit with small statistics, points were more lined up with the expected line. Understand the difference in RICH response for negatives and positives is important. It was not clear if the difference was just because of the calibrations of different tiles (13 and 14) or there is something more fundamental.
4. Average number of photoelectrons for LTCC sectors 3 and 5 have been presented that were within expectations. However, it is important to look at the photoelectron distribution over the acceptance region of detectors, similar to what was presented for HTCC – distribution of Np.e. in (XY) plane. Such distributions for LTCC will show fiducial regions and acceptance holes.

Recommendation:

In the first presentation, two beam energies were quoted for the run period, 10.2 GeV and 10.4 GeV. The run was at a fixed beam energy of 10.4 GeV. It was a wrong BSY energy reading in EPICS at the beginning of the run period that ended up in RCDB. The beam energy must be corrected to 10.4 GeV in RCDB for the initial period of the run.

Charge #2: *(a) Is data quality as a function of run number or time for the data set that is proposed for pass1 cooking stable and understood? (b) Is reconstruction efficiency consistent with expectations and reproducible by appropriate MC simulations?*

(a) Yes – timelines for most of the parameters have been shown and discussed. Most of the monitored parameters are within expected tolerances, DC residuals being the biggest outliers.
(b) Yes – Reconstruction efficiency of the forward detector has been studied using the spring 2019 luminosity scan data. Analysis included the beam background merged low luminosity run. It has been demonstrated that merging the background reproduces the efficiency loss (similar to RG-A).

Comments:

As it was presented at the closeout, DC has been calibrated for one run of the RG-B winter data set. We suggest to run another pass0 and present timelines of the residuals with the new calibration constants before pass1 starts.

Recommendation:

There was a discrepancy between the number of electrons in sector 4 shown in the sampling fraction plots (Yordanka) and in timelines (Electrons/trigger). The number of entries in the sampling fraction plots was not presented, but visually it looked like the number of electrons in sector 4 was much higher than in other sectors. The timelines showed the sector 4 had the lowest number of electrons/trigger. This discrepancy must be understood before pass1 starts.

Charge #3: *(a) Are analysis plans for the data set developed at adequate levels? (b) Is the list of planned skims defined and tested running analysis trains on preliminary data? (c) Are preliminary analysis results for the main reaction channels and observable available and consistent with expectations? (d) Is all ancillary information (helicity, Faraday Cup, …) available and understood?*

(a) Yes – The analysis plans for the leading candidates for the first publications from RG-B, nDVCS for example, are well developed using the processed data.
(b) Yes – The skim list was presented, new skim algorithms (wagons) have been developed. The output of skims will be ~15% of the DST data volume.
(c) Yes – The preliminary analysis results for many physics reactions are encouraging.
(d) Yes – All ancillary information is available and has been checked,

Comments:

- Run group developed a new optimized skims for physics analysis. There should be a unified plan for redoing skims with the latest wagons for the whole RG-B data set.
- Define the main physics reactions that likely to produce publications from the pass1 processed data

Recommendations:

None.

Charge #4: *Are data processing tools that will be used adequate for the proposed processing task? Is the data management plan (staging area, tape destination, directory structure, logs, …) defined and appropriate given the available resources? Is the estimate of resources needed to complete the task sound?*

Yes – Data processing tools, the data management plan, and the required resources are adequate for processing of this data set.

Comment:

- With current farm load, and without competition from other CLAS12 run groups for data processing resources, the whole data set will be processed in about 1.5 – 2 month time frame. The skim output, about 6 TB, will be available for physics analysis either on volatile or on cache disks.
- Plans for storing skims for physics analysis – volatile vs. cache.

Recommendations:

None.

Charge #5: *What are the plans for monitoring the quality of the cooking output and identify/correct failures?*

Yes – Standard monitoring tools are used.

Comment:

The monitoring of normalized yields and beam spin asymmetries have been shown for pass0.

Recommendations:

None

Charge #6: *Is the manpower adequate for the proposed data processing?*

Yes – The available manpower is adequate for processing of the RG-B winter 2020 data set.

Recommendations:

None.