Friday, November 3, 2023

Report

Readiness Review for Pass2 processing of the Hall-B/CLAS12 RG-B Fall-2019/Winter-2020 data set

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

The review of the readiness of RG-B to process a second pass of the Fall-2019/Winter-2020 dataset with the latest improved reconstruction software available took place on November 1st on Zoom. The meeting agenda and presentations can be found on the review page:
https://clasweb.jlab.org/wiki/index.php/Run_Group_B#Pass-2_review_.28Fall19.2FW20.29

The review committee would like to thank the RG-B team for preparing the presentations and addressing the reviewer's questions.

We believe the RG-B team addressed all the charges of this review, suggesting starting **RG-B Fall-19/Winter-20 data cooking as soon as recommendations are cleared.**

Details about the charges and responses are reported below.

Review Charges

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?

FINDINGS:

The RG presented the status of dataset calibration for both run periods (Fall-19 and Winter-20) showing that the calibration constants that will be used to reconstruct data from the CLAS12 are well within the requested limits. The calibration constants appear to be stable over the whole run,

and the results are consistent. The CALCOM cleared calibration. The CVT shows very good results after alignment.

COMMENTS:

None

RECOMMENDATIONS:

None

Charge #2 Is data quality as a function of run number or time for the data set proposed for cooking stable and understood? Have runs been classified in terms of type (empty target, calibration, special, production, …) and quality (golden run, known issues, …), and is a detailed list available? Based on validation studies, have all CLAS12 subsystem performances been understood and issues identified?

FINDINGS:

The RG-B timelines presented at the review demonstrate good stability (within the specs) of all CLAS12 subsystems as a function of time (or, equivalently, run numbers). Sudden changes in the monitored quantities were explained by considering the different experimental conditions of the different data sub-sets (e.g. low energy runs, in-bending, out-bending, …) Some of the runs show behavior outside specifications: RG-B intends to process all runs (including outliers) postponing the decision of using them or not to the specific physics analysis they will be used for. A list of problematic runs with known issues has been compiled.

COMMENTS:

The Review Committee understands the rationale for including all possible runs in the Pass2 processing but the RG-B team should provide a clear strategy to address known issues in some of the runs that will be cooked (e.g. flagging the run quality or defining a post-processing procedure). We also noticed that LTCC timelines are out of specs because they were not updated. It would be good to have them correctly reported to avoid confusion in the future.

RECOMMENDATIONS:

The helicity shows a different sign definition between low energy (4GeV) and high energy (11 GeV) runs. RGB should define a clear strategy (keep the different signs and correct when physics analysis will be performed or correct with the existing post-processing procedure) and document it to avoid confusion in the future.

Charge #3: Has a Hardware (HW) status table (i.e., bad channel table) been compiled for use in the data and MC reconstructions? Has the efficiency versus beam current been studied? How does it compare to MC simulations with the merged background? Are the DAQ translation tables correcting for all known cable swaps? At what stage(s) in the software?

FINDINGS:

HW status tables have been defined for the CLAS12 subdetectors. DC map has been updated for cable swaps with a similar procedure used in the RG-B Spring 19 data set.

COMMENTS:

None

RECOMMENDATIONS:

Before proceeding with the Pass2 data processing, status tables for Fall-19 and Winter-20 data sets need to be fully validated.

Charge #4: Are analysis plans for the data set developed at adequate levels? Is the list of planned skims defined and tested running the analysis trains on preliminary data? Is all ancillary information helicity, Faraday Cup, …) available and understood?

FINDINGS:

Comparison with Pass1 results (where available) shows a clear improvement in statistics (e.g. pDVCS yield improved by ~60%) and in precision (e.g. a more reliable neutron identification in nDVCS channel). Comparison between Pass1 and Pass2 was only possible for data sets with a consistent experimental setup.

The yield increase and the good performance of the well-aligned Central Detector resulted in a significant increase in the physics reach of the RG-B data set.

COMMENTS:

We noticed that, despite the background inclusion in simulations, the slope in the reconstruction efficiency as a function of the beam current remains slightly different than in the data (as indicated by the luminosity scan). This is a known issue that needs to be addressed before any precise determination of cross-sections or other absolute observables. The same issue is valid for other RGs too.

RECOMMENDATIONS:

None

Charge #5: Are the data processing tools that will be used adequately 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 processing time per event available and resources needed to complete the task sound?

FINDINGS:

RG-B presented estimates of the necessary disk space needed by pass-2 cooking. The cooked dataset size (estimated to be DST: 26 TB + 40 TB, and SKIMS: 8 TB + 9 TB) is compatible with the current disk resources allocated to CLAS12.

COMMENTS:

The farm is currently busy with RGA Pass2 reconstruction. If RG-B reconstruction will start before RG-A ends, the CCC should define the best strategy to run in sequence or in parallel for the two analyses.

RECOMMENDATIONS:

RG-B should provide a detailed estimate of the time necessary for the two data sets to be cooked. The estimate should be discussed and agreed upon with Hall-B experts.

Charge #6: Have the tools for monitoring the quality of the cooking output and identifying/correcting failures been defined and ready to be used?

FINDINGS:

RG-B presented some results used to monitor the cooking output This, together with high-level physics analyses, will be used to monitor the pass2 data.

COMMENTS:

None

RECOMMENDATIONS:

None

Charge #7: Is the person-power identified and in place for the proposed data processing?

FINDINGS:

Personnel, including the coordinator, chef, calibrators, and physics analyzers were found to be adequate.

COMMENTS:

None

RECOMMENDATIONS:

None