Thursday, August 17 2023

Report

Readiness Review for Pass-2 processing of the Hall-B/CLAS12 RG-K 2018 data set

Review committee:
Nathan Baltzell
Marco Battaglieri (chair)
Larry Weinstein
Daniel Carman (temporarily replacing Marco Mirazita)

The review of the readiness of RG-K to process a second pass of the 2018 data set with the latest improved reconstruction software available took place on August 16 on Zoom. The meeting agenda and presentations can be found on the review page:
https://clasweb.jlab.org/wiki/index.php/Run_Group_K#tab=Pass2_Review_Documents

The review committee would like to thank the RG-K team for preparing the presentations and addressing the reviewer's questions. The Committee was nicely impressed by RG-K's preparation for this review.

We believe the RG-K team satisfactorily addressed all the charges of this review **suggesting to start RG-K 2018 data cooking as soon as possible.**

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 data set calibration 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. CVT alignment followed the successful approach used by RG-A and RG-B. RG-K presented a detailed comparison between the pass-1 and pass-2 results showing a significant improvement in term of statistics. RG-

K pass-2 will utilize both the AI-supported tracking and de-noising (despite the limited gain per track estimated to be ~2%).

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-K timelines presented at the review demonstrate good stability (within the specs) of all CLAS12 subsystems as a function of time (or, equivalently, run number). A detailed list reporting run classification (gold and silver) as well as special running conditions (empty target, low luminosity, …) has been provided. A list of problematic runs with known issues (wrong helicity) has been compiled.

COMMENTS:

The fictitious change of sign of the helicity noticed for a few production runs should be fixed during data processing correcting the DB entries.

RECOMMENDATIONS:

None

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 and validated for the CLAS12 subdetectors. The DC map has been updated for cable swaps and validated for one run. RG-K will use the same tables generated for RG-A F18.

COMMENTS:

Track reconstruction efficiency obtained by adding data background to Monte Carlo pseudodata does not always align with what is measured (typically overestimating the tracking efficiency). This has been tracked back to a difference between the kinematics of data and pseudodata and higher-order corrections not yet implemented in GEMC (e.g. effect of acquisition thresholds).

RECOMMENDATIONS:

None

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:

RG-K already published a paper based on pass-1 reconstruction of the same data set. Considering the significant experience gained by the RG-K group in performing high-level physics analyses, the Committee does not identify any significant issues.

COMMENTS:

None

RECOMMENDATIONS:

None

Charge #5: Are the 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 processing time per event available and resources needed to complete the task sound?

FINDINGS:

RG-K presented estimates of the necessary disk space and computer time needed by pass-2 cooking. The cooked data set size (estimated to be less than 200 TB) is compatible with the current disk resources allocated to CLAS12. The processing time has been estimated to be about 12 days if the full computing resources allocated to Hall B will be used. No other RGs are expected to reconstruct data while RG-K will run pass-2 reconstruction.

COMMENTS:

None

RECOMMENDATIONS:

None

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-K presented some results used to monitor the cooking output This, together with high-level physics analyses, will be used to monitor the pass-2 data.

COMMENTS:

None

RECOMMENDATIONS:

None

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

FINDINGS:

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

COMMENTS:

None

RECOMMENDATIONS:

None