

Pass1 readiness reviews – RG-A(2), -B, -K

S. Stepanyan (JLAB)

CLAS Collaboration meeting, July 21-24

Initial thoughts on Pass1

- The first pass1 readiness review took place in February, 2020
- The review committee: C. Smith, M. Mirazita, N. Baltzell, L. Weinstein, and S. Stepanyan(chair)
- We knew that:
 - the software was not fully ready, namely tracking in the central detector was not ready, so it was left out of the review, and some other parts of the software and calibrations have room for improvements
 - Magnetic fields were (are) not finalized
 - Monitoring tools were still in development
- Nevertheless, we had to process some data to produce the first physics results.
- The state of the software and calibrations were good enough for [as we thought] some physics reactions in RG-A experiments that will only use the CLAS12 forward detector

Readiness Review for “PASS1” cooking of CLAS12 data

Processing of CLAS12 data is a computing intense process that requires detailed planning and preparatory work to guarantee the quality of the output data. For this reason, a “readiness review” is requested to authorize the start of the reconstruction data processing of any new data set. The charge for this review is outlined below.

The review committee is requested to answer the charge questions based on the material presented by the Run-Group and report its findings, comments and recommendations to the CLAS Coordinating Committee.

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?

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

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

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?

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

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

First Reviews: RG-A fall 2018 “inbending” (L. Elouadrhiri)

- The review of “inbending” data took place on February 27 at JLAB - <https://clas12-docdb.jlab.org/cgi-bin/DocDB/private/DisplayMeeting?sessionid=361>
- About half a day of presentations (9), closeout right at noon.
- Overall RG-A was well prepared and addressed all the charge items. There have been few recommendations mainly on monitoring tools that were just getting developed
- Follow-up short meeting on March 9 cleared the committee’s recommendations
- A small part of the run period at the beginning was left out from the review with agreement that will be calibrated shortly, but it seems took much longer to get that portion calibrated than expected

This first review set the stage for other readiness reviews. RG-A team together with the software and CalCom groups did the most of the heavy lifting in preparing the whole calibration, monitoring and the data processing framework

RG-A fall 2018 “outbending”

- The review of “outbending” data took place on April 22, 2020, in remote setting over the bluejeans: <https://clas12-docdb.jlab.org/cgi-bin/DocDB/private/DisplayMeeting?conferenceid=2>
- About half a day of presentations, no recommendations, data set was ready for pass1

RG-B - spring 2019 data set (S. Niccolai)

- The review took place on May 8, 2020, remotely over bluejeans:
https://clasweb.jlab.org/wiki/index.php/Run_Group_B#tab=Pass-1_review
- Group was well prepared, presentations covered all charge items, committee had only four recommendations. Two of recommendations were on calibrations important for the RG-B physics reactions
- Run group addresses all the recommendations on June 10 and got approval for pass1

RG-K – winter 2018 data set (A. D'Angelo)

- The review took place on June 6, 2020, remotely over bluejeans:
https://clasweb.jlab.org/wiki/index.php/Run_Group_K#tab=Pass1_Review_Documents
- As other group, RG-K was well prepared, presentations covered all charge items, committee had only one recommendation unrelated to data quality but on the choice of large volume skims.
- The recommendation was addressed quickly and RG-K is now processing the data

Pass1 reviews, some remarks

- Reviews were short but productive. Full reports from the committee were sent to CCC and RGs within few days
- Typically materials were available for the reviewers almost at the last minute – difficult to catch everything. We may want to address this for the next review. A longer review and/or pass1 readiness note can be considered, reviewers need more than an evening to go through materials
- Unless RGs will diligently check every detail to make sure the full dataset is well calibrated (down to each element), a few hour long review with short presentations will not be able to catch everything.
- Nevertheless, considering these are the first steps in the calibration and processing of the huge amount of data from a complex detector like CLAS12, we did reasonably well. Yes, as always, there is a room for improvements
- Some of the charge items should be revised (e.g. 2nd sentence of charge #2 or 1st and 3rd sentences of charge #3). These items are in the present change since the envisioned “pass1” should have been very limited in scope, but since the scope widened, see the first bullet, they are far beyond the pass1 readiness review scope. The reconstruction efficiency is complicated, reaction dependent, and so the physics analysis. It is hard to judge if the analysis results are consistent with expectations based on a small fraction of pass0 data

More remarks

- At the beginning we agreed for the limited scope for so called pass1 data processing – started with the review and processing of a small fraction of RG-A for the first physics publications using only FD. It was expected to have the remaining calibration and software issues resolved by the end of the first round of RG-A data processing and then go with real PASS1
- But, we ended up with the processing of data from three RGs, and some analyses now aim to publish results that will include the whole CLAS12 detector – Pandora's box is open
- Now we should decide:
 - what should we do with the physics analyses that include not fully ready and reviewed systems (e.g. CVT). My personal preference – develop ad hoc corrections, review and approve for use where physics results are not compromised by using these corrections. Too much data will be processed by the time major expected improvements are done. It will most likely take another year (optimistic) or more to process them again. It is unwise to hold physics analysis that will not be compromised without these improvements
 - what we will do next, after major missing parts and improvements are ready (e.g. CVT tracking):
 - a) re-process the data that we already processed (a year or more long effort) or
 - b) continue processing the remaining data sets (e.g. from RG-A and B) to add more statistics, albite with better quality – this one is my personal preference. If we still have data left to be processed we should do so and then come back to reprocess.