Outcome of Director's Review Regarding Software/Analysis

- **Recommendation:** "The collaboration is strongly encouraged to develop an end-to-end realistic simulation and reconstruction to further optimize cost and physics reach."
- **Finding:** "Having a [sic] functional simulation and reconstruction routines as soon as possible should be a high priority in the software effort."
- **Finding:** "Early exploration of the tools available at Jefferson Lab that can handle the data at the expected scale of SoLID will be crucial in minimizing the [sic] false starts in software development."
- **Recommendation:** "Compare the resource levels you have assumed in some key areas (particularly in software, [...]) to make sure the estimates align with other similar projects or there is a good reason they do not."

Developing SoLID Software Specifications

- Need specifications for (at least) each of
  - Simulation
  - Digitization
  - Databases
  - File formats
  - Reconstruction Framework
  - Calibrations
  - Physics Analysis
- Goal: Write up design document within next few months
  - High-level overview of end-to-end simulation/reconstruction chain
  - Detailed specs for items above (as much as is realistic at this point)
  - Have (early) draft ready for May collaboration meeting

Questions to Ask: Simulations, Digitization

- Simulation
  - Package: GEMC vs. `remoll_solid`?
  - Which physics/background event generators?
  - (experts fill in more details)
- Digitization
  - Integrated into simulation vs. standalone
  - Trigger simulation?
  - Digitize at hardware level?
  - Support CODA output?

Databases, File Formats

- Databases
  - Organization: Geometry, Mapping, Configuration, Calibrations, more/less? ...
  - Indexing: run numbers, event ranges, time ranges, "variations"?
  - Contents: must be suitable for all of simulations, digitization, reconstruction, analysis
  - Storage, low-level format: flat files, XML, SQL
  - Engines: MySQL, CCDB, ...
- Data File Format(s)
  - EVIO, ROOT, (many other possibilities) ...
  - Support more than one format at any stage?
  - Inclusion of metadata, database parameters, ...
  - Detailed detector-level format specs

Reconstruction Framework I

- Online vs. offline
  - Data flow
  - Desired online results
  - Level-3 trigger/farm
- Desired user experience
  - Interface: Command line, shell scripts, interpreter, GUI
  - Configuration: compiled vs. runtime-configurable
  - Data interface: low-level (e.g. C-structures, method calls) vs. high-level (e.g. ROOT's tree variables)
  - Condition testing: at code level, pre-defined configurable, arbitrary expressions
  - Input formats: How to switch between MC and data?
  - Output: pre-written engine(s), or up to the user?
  - Flexibility: Amount of work involved for adding a new detector system
  - Auto-configuration for any particular features, e.g. from input files?

Reconstruction Framework II

- Technical Requirements
  - Multi-threading: must!
  - Support for distributed computing?
  - Output event ordering? Sync at special events?
  - Output data complexity (scalars, arrays, matrices, structures, objects)? Nesting depth?
  - Propagate MC truth data to output? Which?
  - Support substitution of any input data with MC truth data? Which?
  - Level of modularity
  - Performance: Minimum analysis rate? Maximum memory?
  - Software dependencies (e.g. allow/require `boost`, ROOT?)
- How to configure SoLID geometries efficiently?
  - PVDIS: fixed number of fully independent sectors
  - SIDIS: variable number of overlapping sectors
  - ...

Level of Physics Analysis Provided

- Calibrated detector data
- Tracks/vertices/4-vectors
- PID, particle hypothesis likelihoods?
- Kinematics for typical reactions? Which?
- Reaction identification?
- Path for users to modify and extend provided methods

Calibrations, Alignment, Monitoring

- Calibration/alignment needs
- Precision requirements, esp. for tracker alignment
- Devise set of MC studies for effects of misalignments
- Specify “calibration loop”
- Required turnaround time?
- Outline of online monitoring scheme

Developing the Specs

- Find authors, make timeline for spec document
- Evaluate any packages in depth? If so, which, how?
- Presentations? Comparison charts?
- Decision process?

Once Specs Are in Place

- Combine existing standalone simulations into single one
- Develop missing digitizations
- Identify/develop required analysis modules
- Import/develop reconstruction algorithms
  - Tracking
  - Calorimeter clustering
  - (PID)
- Develop rough timeline for completion of tasks
- Find manpower, delegate responsibilities