# **SBS Software and Computing**

**JLab Software Review** 

Seamus Riordan

seamus@anl.gov

November 27, 2018



#### Project Overview

- Super Bigbite is a broad program utilizing large acceptance and high luminosity
- Core measurements are elastic nucleon form factors at high Q<sup>2</sup>
- Total of 184 days of running approved
- Present projected start 2021







#### Detector Overview

#### • Several major new subsystems

| Event Rate                       | Up to ~5 kHz                 |
|----------------------------------|------------------------------|
| GEM Trackers                     | 100k strips (≤40% occupancy) |
| Hadronic calorimeter             | 288 FADC                     |
| Electromagnetic Calorimeter      | 1700 ADC                     |
| Scintillator coordinate detector | 2000 TDC                     |
| Gas Cherenkov                    | 550 TDC/ADC                  |
| Scintillator Timing Plane        | 360 TDC/ADC                  |

- Reuse of existing
  - Bigbite EM calorimetery (~200 PMTs)
  - HERMES RICH (2000 PMTs)



# Program and Kinematics

| Projected Year | Experiment | Production<br>Days | Luminosity<br>[Hz/cm²] | Peak GEM Rate<br>[kHz/cm <sup>2</sup> ] |
|----------------|------------|--------------------|------------------------|---|
| 2021           | GMn        | 25                 | 3×10 <sup>38</sup>     | 100                                     |
|                | GEn-Recoil | 5                  | 1×10 <sup>38</sup>     | 100                                     |
| 2022           | GEn        | 50                 | 6×10 <sup>36</sup>     | 100                                     |
| 2023           | GEp        | 45                 | 8×10 <sup>38</sup>     | 700                                     |
| 2024           | SIDIS      | 64                 | 4×10 <sup>36</sup>     | 50                                      |

- Form factor experiments benefit from relatively clean coincident kinematics, high trigger threshold
- GMn, GEp (cryo targets) have largest anticipated rates
- Q<sup>2</sup> = 12 GeV<sup>2</sup> GEp kinematic has highest rate and serves as ultimate benchmark for performance

#### Present Projected Timeline and Milestones

|           | Software activities  | Experiment running and analysis  |
|-----------|--|--|
|           | Completed – Full simulation interfaced to analysis<br>Completed – Decoders and channel analysis                                      |  |
| Fall 2018 | Began analysis of digitized simulated experiments  |  |
| Jan 2019  | Collection of online and offline analysis and displays   |  |
| 2019      | Neutron FF experiment simulated analysis<br>Goal – GMn tracking to 80% efficiency, 8 Hz  |  |
| 2020      | Jan - GMn, GEn ready for analysis<br>Proton FF experiment simulated analysis<br>Online, offline scripts finalized from commissioning | March - GMn installation begins  |
| 2021      | Goal – GEp tracking to 80% efficiency, 3 Hz<br>Jun - GEp ready for analysis<br>Start simulated analysis of SIDIS, TDIS               | Jan - GMn start of run<br>GMn analysis begins<br>Fall – GEn start of run |
| 2022      | Jan - SIDIS ready for analysis   | Fall - GEp start of run?   |
| 2023      |  | GEp analysis begins?<br>SIDIS start of run?                              |

# Analysis Readiness Checklist

Simulation of backgrounds and analysis expected to take place on JLab farm

Full analysis includes

- Version controlled repository of software for each experiment
- Organized passes for
  - Pool of simulated individual background events
  - Pool of simulated digitized events with background superimposed
- Use of tracks from reconstruction under full background and +50%, +100% conditions
- Calorimeter cluster reconstruction under superimposed background
- Cherenkov cluster reconstruction
- Integration of inter-detector information for
  - Identifying good track pathways
  - Probabalistic track identification
- Full event-by-event kinematic and vertex reconstruction including optics models
- Identification of
  - Random backgrounds
  - Contamination of inelastic and radiative events

# Collaboration Workforce

- Consistent and active workforce
- Weekly regular meetings

| Subproject                |                           |
|---------------------------|---------------------------|
| Subgroup and Organization | ANL                       |
| Front end decoding        | JLab, INFN, ANL, CMU, UVA |
| GEM Analysis and Tracking | JLab, UConn, UVA          |
| HCal                      | CMU                       |
| ECal                      | UConn                     |
| Coordinate Detector       | CNU                       |
| Timing Hodoscope          | Glasgow                   |
| GRINCH Cerenkov           | W&M                       |
| Bigbite Legacy            | ANL                       |

• Spokespeople are responsible for experimental analysis

| Institution        | Collaborators (FTE/yr)                              |
|--------------------|---|
| ANL                | Riordan (0.2)                                       |
| CMU                | Quinn, Cornejo (0.5)                                |
| CNU                | Monaghan, Brash,<br>students                        |
| Glasgow            | Annand, Hamilton,<br>Montgomery                     |
| INFN               | Cisbani, Musico                                     |
| JLab               | Camsonne (0.2),<br>Hansen (0.1),<br>Future Postdoc? |
| UConn              | Puckett (0.2), Fuchey (0.5),<br>Student (0.5)       |
| UVA                | Liyanage, Gnanvo, Future<br>postdoc?, Di, Jian      |
| W&M                | Averett, Ayerbe-Gayoso                              |
| Professor/Staff Po | stdoc Student                                       |

## Goals and Status

- Complete simulation of all experiments
  - Robust simulation developed (more in following)
  - Completed integration to analysis framework in 2018
- Decoding, basic analysis of CODA data in Hall A analyzer
  - Decoders available for all subsystems (including MPD, FADC)
  - Output and low-level analysis of channel-level data completed
- Analysis of experiments
  - Clustering, tracking (more in following), inter-detector association underway
  - Optics models, spin transport, etc to be finalized
  - Simulated analysis of experimental observables to follow
- Online and Event Displays
  - Repository started for subsystems, individual groups are responsible
  - Software exists for legacy Bigbite systems (including optics)
  - Will continue to develop through assembly/commissioning period up to runs

# Simulation and Analysis Integration

- Spent significant effort to integrate simulation and analysis
- Full set of elastic, choice of inelastic and random background generators
- Have full chain to take Geant4 hits and produce digitized channels
  - Methodology similar to developed GEM digitization libraries
- Hall A analyzer has decoder to interface to simulation ROOT output
- Production of single virtual event rate of 0.1 Hz with fully correlated GEp background and digitization
- Now starting test of full analysis chain



# **GEM Tracking Software**

- Primary deployed algorithm using recursive TreeSearch (Raw combinatorics also employed for some analyses)
- GEMs provide six time samples over 25ns bins with jitter
- Hits are differentiated by fitting to spatial and temporal components
- Require amplitude matching between x-y components to obtain full 3D
- General restrictions are placed on search areas based on other detector knowledge]
- Basic multithreading implemented



# GEM Tracking – GMn

- Since 2016
  - Improved GEM response and validation based on data from constructed GEMs
  - Observe larger and wider background response
- Event reconstruction at
  - 70% tracking efficiency (2020 goal 80%)
  - 3 Hz (2020 goal 8Hz)
- Continuing to evaluate better separation of broad ADC clusters



# GEM Tracking – GEp (goal by mid 2021)

- GEp rates roughly factor 5 higher than GMn
  - Goal: 80% efficiency, 3 Hz by mid 2021
- Significant postprocessing will be required Reevaluating TreeSearch
- GMn will be critical to understanding high rate data
- Have postdoc focusing on this effort





#### Tape Storage Requirements

| Projected<br>Year | Experiment | Production<br>Days | Data Rate<br>[MB/s] | Raw Data<br>Volume<br>[TB] | Processed<br>Volume<br>[TB] | Simulation<br>Volume<br>[TB] |
|-------------------|------------|--------------------|---------------------|----------------------------|-----------------------------|------------------------------|
| 2021              | GMn        | 25                 | 125                 | 300                        | 200                         | 30                           |
|                   | GEn-Recoil | 5                  | 125                 | 50                         | 40                          | 5                            |
| 2022              | GEn        | 50                 | 150                 | 600                        | 500                         | 60                           |
| 2023              | GEp        | 45                 | 425                 | 1700                       | 1300                        | 300                          |
| 2024              | SIDIS      | 64                 | 320                 | 1300                       | 900                         | 300                          |

- GEM hits represent bulk of data
- Data rates scale approximately linearly with background

# **Computing Requirements**

| Projected<br>Period | Experiment | Analysis Rate<br>[ms/evt] | Simulation<br>kCore-hours | Analysis<br>kCore-hours |
|---------------------|------------|---------------------------|---------------------------|-------------------------|
| 2021-2023           | GMn        | 135                       | 150                       | 700                     |
|                     | GEn-Recoil | 135                       | 40                        | 100                     |
| 2022-2024           | GEn        | 135                       | 300                       | 1200                    |
| 2023-2025           | GEp        | 330                       | 1100                      | 2700                    |
| 2024-2026           | SIDIS      | 66                        | 2200                      | 1100                    |

- Assumes three passes over production data
- Simulation rate for total digitized with background 0.1 Hz for GEp full background
- Simulated analysis at similar rate as production

#### Summary

- SBS collaboration has significant resources dedicated to software
- SBS runs projected to start in two years (January 2021)
- Simulation well developed and recently interfaced to analysis
- Decoding and analysis infrastructure in place within Hall A analyzer
- Tracking remains primary challenge for software
  - Goals to have analysis ready approximately 1 year before each experiment