GPU-based Online Reconstruction for SpinQuest Studies at Fermilab

The 11th Workshop of the APS Topical Group on Hadronic Physics
Mar 15th, 2025

Utsav Shrestha

Mississippi State University

Outline

- Introduction and Motivation
  - Nucleon Spin and Transverse Momentum-dependent Distributions (TMDs)
  - Sivers Functions and Transverse Single Spin Asymmetry (TSSA)
  - Transversity in SpinQuest
  - The E1039/SpinQuest Experimental Setup
- Graphics Processing Unit (GPU)-based Online Reconstruction (OR)
  - GPU OR Requirements
  - GPU Program Structure
  - Multi-threaded GPU Parallelization
  - Track Reconstruction, GPU Vertexing and Dimuon Reconstruction
  - CPU and GPU Comparisons
  - Online Display
- Summary and Outlook

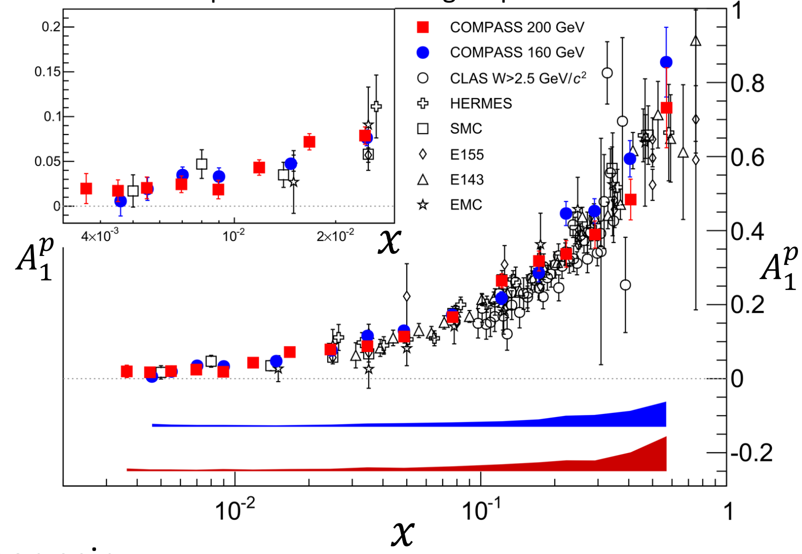
Nucleon Spin and Transverse Momentum-dependent Distributions (TMDs)

Nucleon Spin Decomposition:

- $S_N = \frac{1}{2} = \frac{1}{2}\Delta\Sigma + \Delta G + L_q + L_g$
  $\Delta\Sigma$: quarks and antiquarks spin
  ΔG: gluon spin contribution
  L_q, L_g: quark/guon angular momentum

- $\Delta\Sigma$ measurements show ≈ 30% contribution
  $\Delta\Sigma = 0.12 \pm 0.09 \pm 0.14$
  [EMC: Nucl. Phys. B328, 1 (1989)]
  $0.26 < \Delta\Sigma < 0.36$
  [COMPASS: Phys. Lett. B753, 18 (2016)]

Longitudinal double spin asymmetry in Deep Inelastic Scattering on proton

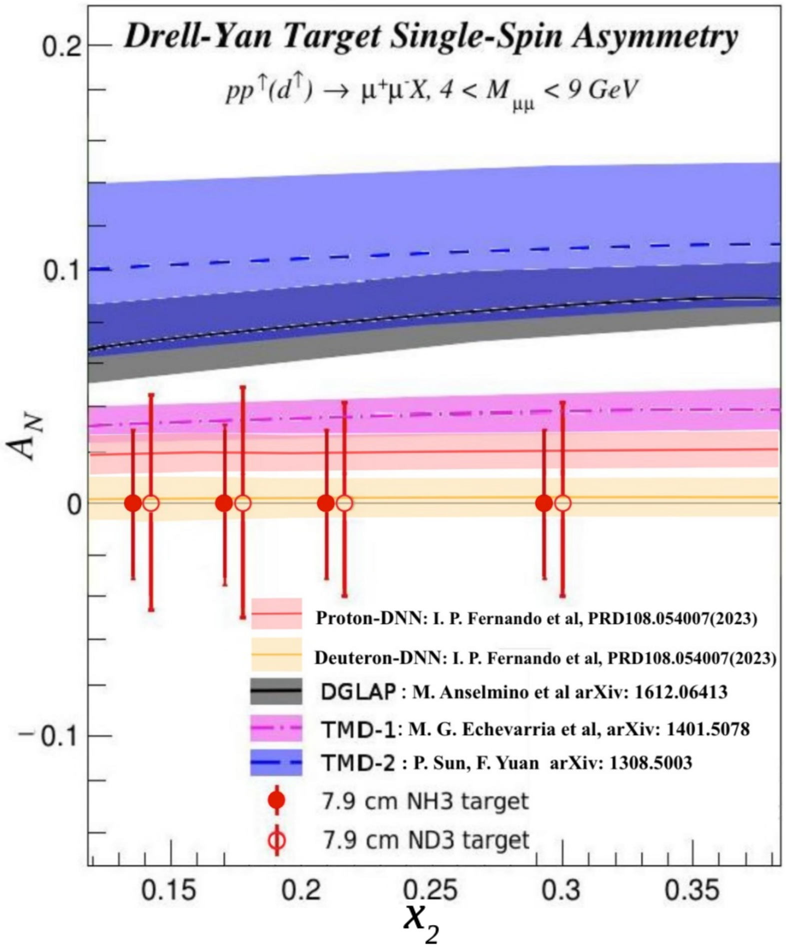


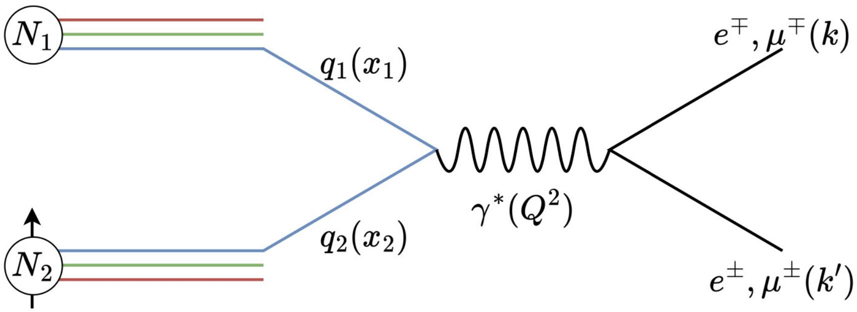
- $\Delta G + L_q + L_g$ contributes more than half of the nucleon spin

<table>
<tr><td colspan="2" rowspan="2">Quark TMDs</td><td colspan="3">Nucleon Polarization</td></tr>
<tr><td>Unpolarized (U)</td><td>Longitudinally (L)</td><td>Transversely (T)</td></tr>
<tr><td rowspan="3">Quark Polarization</td><td>U</td><td>$f_1(x, k_T)$
unpolarized TMD</td><td>-</td><td>$\boldsymbol{f_{1T}^{\perp}(x, k_T)}$
Sivers Function</td></tr>
<tr><td>L</td><td>-</td><td>$g_1(x, k_T)$
helicity</td><td>$g_{1T}(x, k_T)$
worm-gear</td></tr>
<tr><td>T</td><td>$h_1^{\perp}(x, k_T)$
Boer-Mulders Function</td><td>$h_{1L}^{\perp}(x, k_T)$
worm-gear</td><td>$\boldsymbol{h_1(x, k_T)}$ Transversity
$h_{1T}^{\perp}(x, k_T)$ Pretzelosity</td></tr>
</table>

Sivers Functions and TSSA

- Drell-Yan (DY) is a clean probe for TMDs, without fragmentation function uncertainties



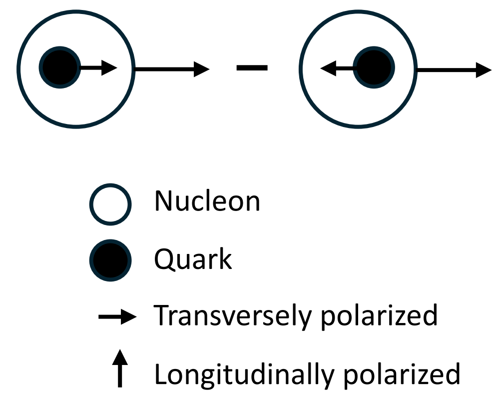


x_1: parton momentum fraction in the beam proton
x_2: parton momentum fraction in the target nucleon
$x_F \equiv x_1 - x_2$: Feynman variable

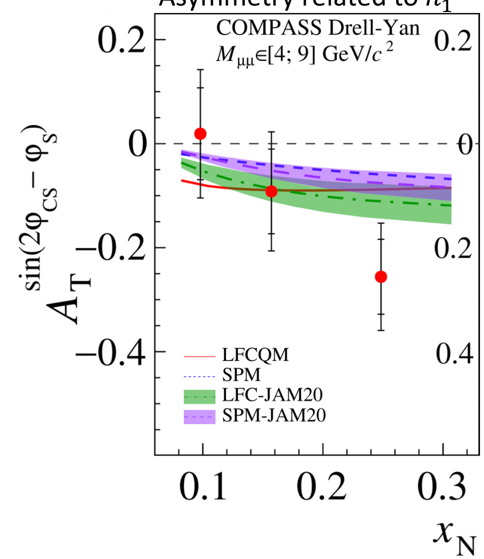
- SpinQuest uses a 120 GeV unpolarized proton beam on a transversely polarized NH_3/ND_3 target
- The Sivers Functions parametrize transverse momentum of quarks in transversely polarized nucleon
- SpinQuest will provide the first measurement of the sea quark Sivers function using polarized DY, probing spin structure at small x
- SpinQuest also provides a unique opportunity to study the poorly known gluon Sivers functions using TSSA in J/ψ production

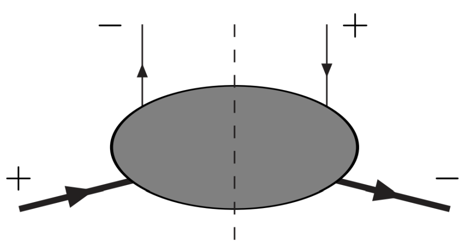
Transversity in SpinQuest

- Transversity is chirally-odd, suppressed in Deep Inelastic Scattering as the handbag diagram cannot flip quark chirality
- Accessing it requires two chirality flips — achievable in polarized Drell-Yan or semi-inclusive leptoproduction
- SpinQuest uses transversely polarized targets in the Drell-Yan process, offering a clean, model-independent extraction of the sea quark transversity distribution, $h_1(x)$
- Avoids fragmentation function uncertainties and final state interactions, unlike Semi-Inclusive Deep Inelastic Scattering
- Absence of gluon counterpart in spin-1/2 hadrons, means it evolves as a non-singlet quantity without gluon contributions

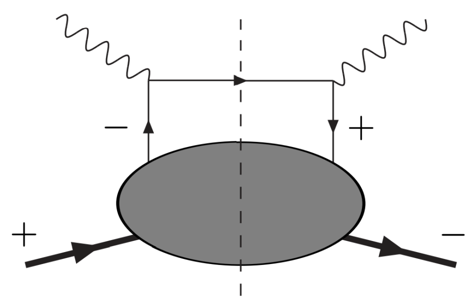


DY Twist-2 Transverse Spin Asymmetry related to h_1



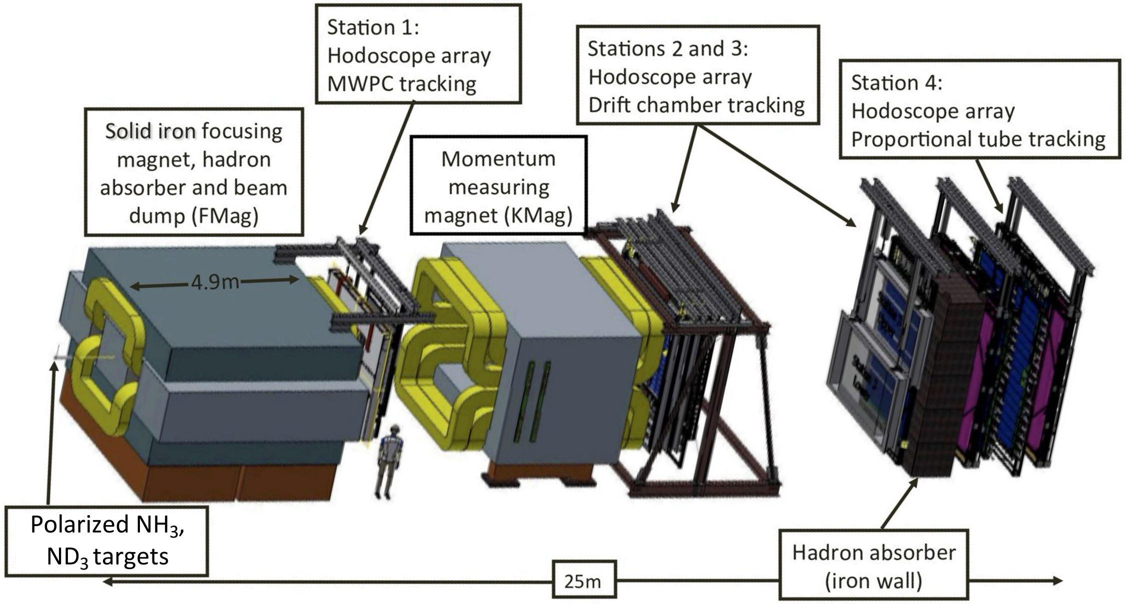


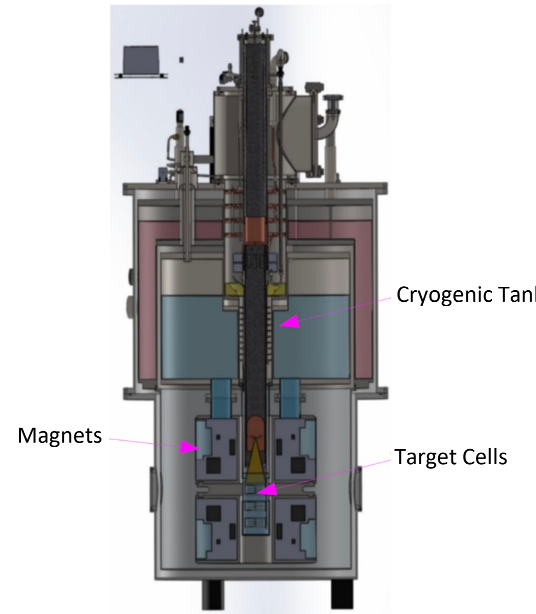
Chirally-odd distribution, $h_1(x)$



A handbag diagram forbidden by chirality conservation

The E1039/SpinQuest Experimental Setup



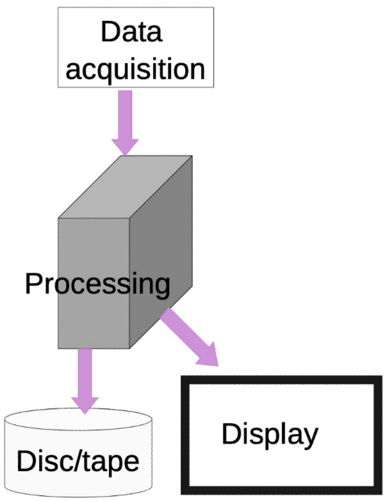


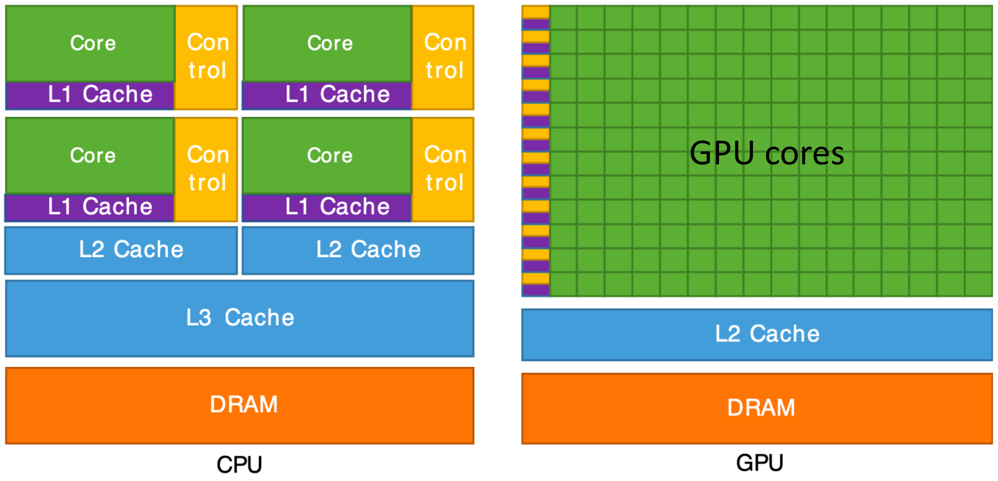
- ➢ Beam (*p*, 120 GeV) is delivered in 4.4-sec long "spills" every 60 sec
- ➢ Polarized targets: Ammonia (NH_3) and deuterated ammonia (ND_3)

More details on the Target System is on V. Ansari's talk: Tomorrow @ 9:50 AM

GPU OR Requirements

- Real-time reconstruction and monitoring of SpinQuest data with ultra-fast analysis program using GPU:
  - Multithreading is pivotal to achieve the required processing speed
- Improvement of memory management which is much more “rigid” on GPU compared to Central Processing Unit (CPU):
  - Memory must be pre-allocated on GPU (input + ouput)
- Input data are copied from CPU to GPU

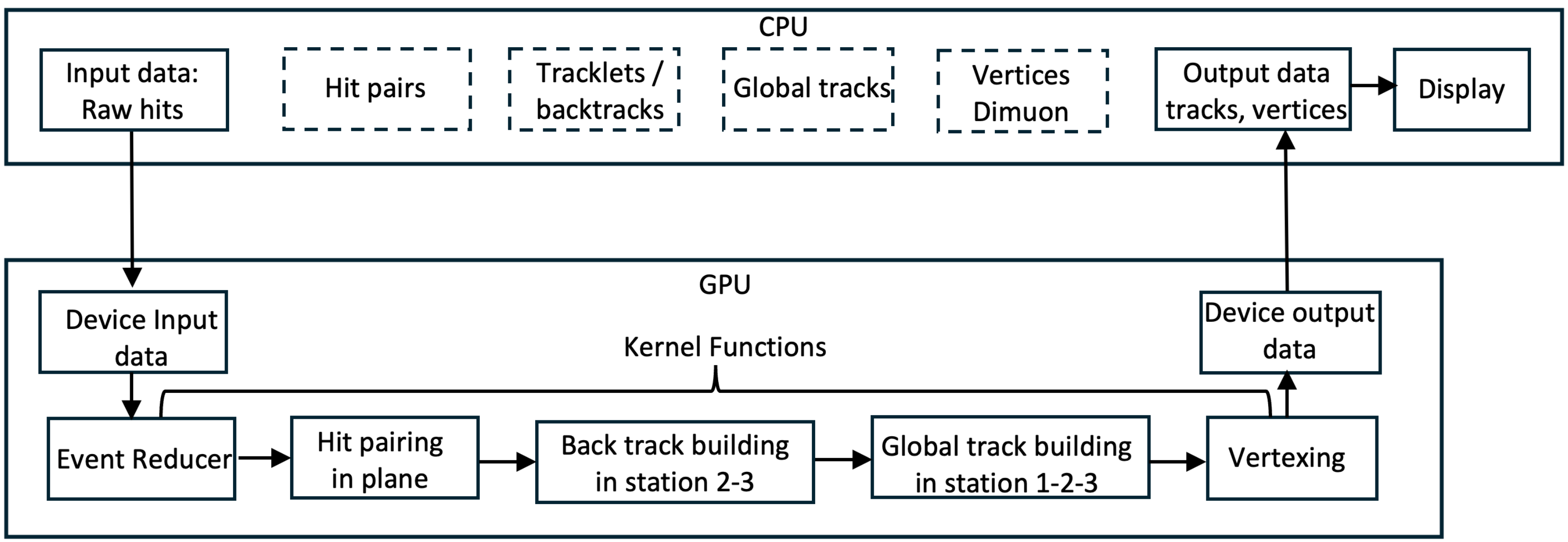




Lambda Vector Workstation
- Ubuntu 22.04, Lambda Stack
- 2x NVIDIA RTX 4090
- 16,384 CUDA cores per GPU
- Alma Linux 9 Docker
- NVIDIA Docker 2, CUDA Toolkit 12.6
- E1039 Core package

GPU Program Structure

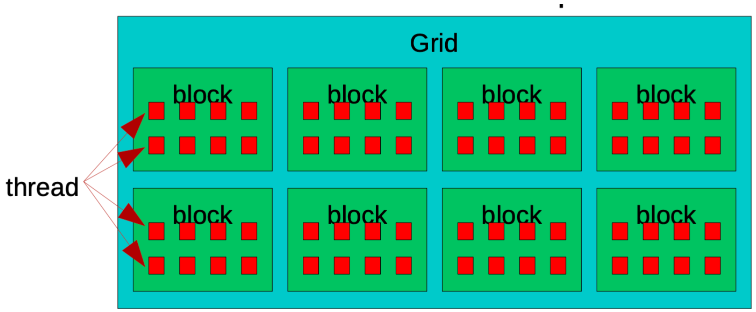
- **Input**: Decoded DST file: Contains hits, formatted for E1039/SpinQuest
- **Memory Management**: Declares structures and allocates memory
- **Data Processing / Kernel Functions**: Event reduction, global track reconstruction, vertexing, etc
- **Output**: DST file and histograms, information of reconstructed muon tracks and dimuons

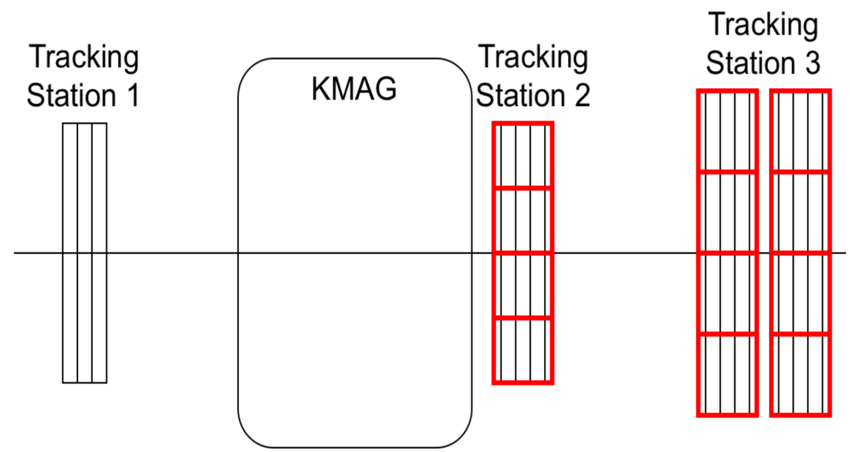


GPU Parallelization: Per-event Multithreading

- Optimizing Processing Speed with Multithreading
  - Track candidate search is parallelized across multiple threads, each covering a defined region of the detector acceptance (32 threads in total)
  - Workload is evenly distributed across threads to maximize GPU utilization and efficiency

- GPU Workload Distribution
  - Grid dimension: Parallel execution of multiple events
  - Block dimension: Intra-event parallelism for efficient track reconstruction





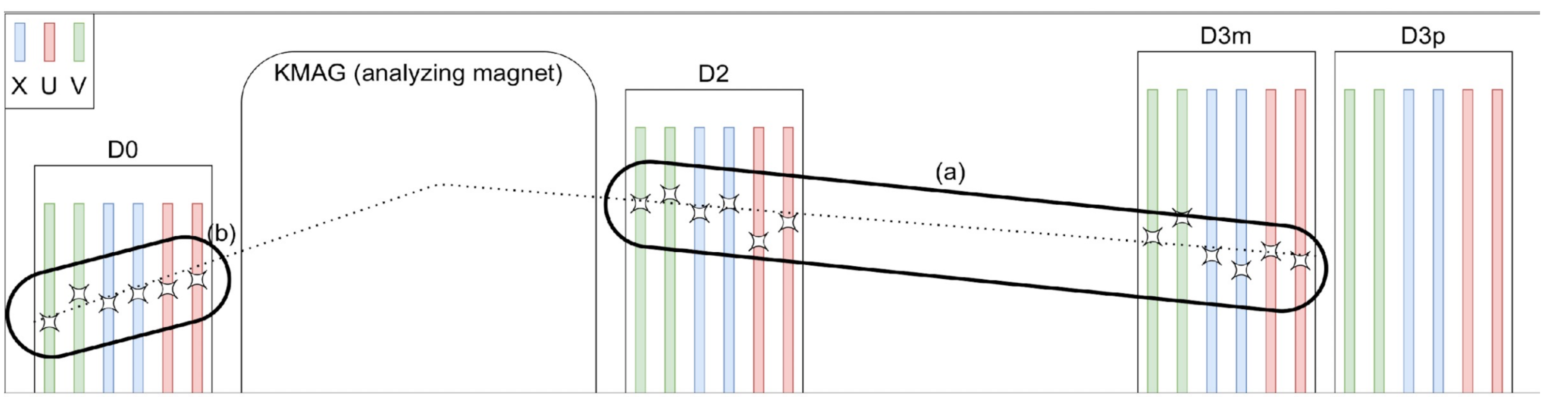
Track Reconstruction

- Reconstruct straight tracks from station station 2 Drift Chambers (D2) to station 3 DC top/bottom (D3p/D3m),
- Associate hits with station 1 DC (D0) to straight tracks,
- Combining D2-D3p/m and D0 track segments -> momentum

X: vertical wires

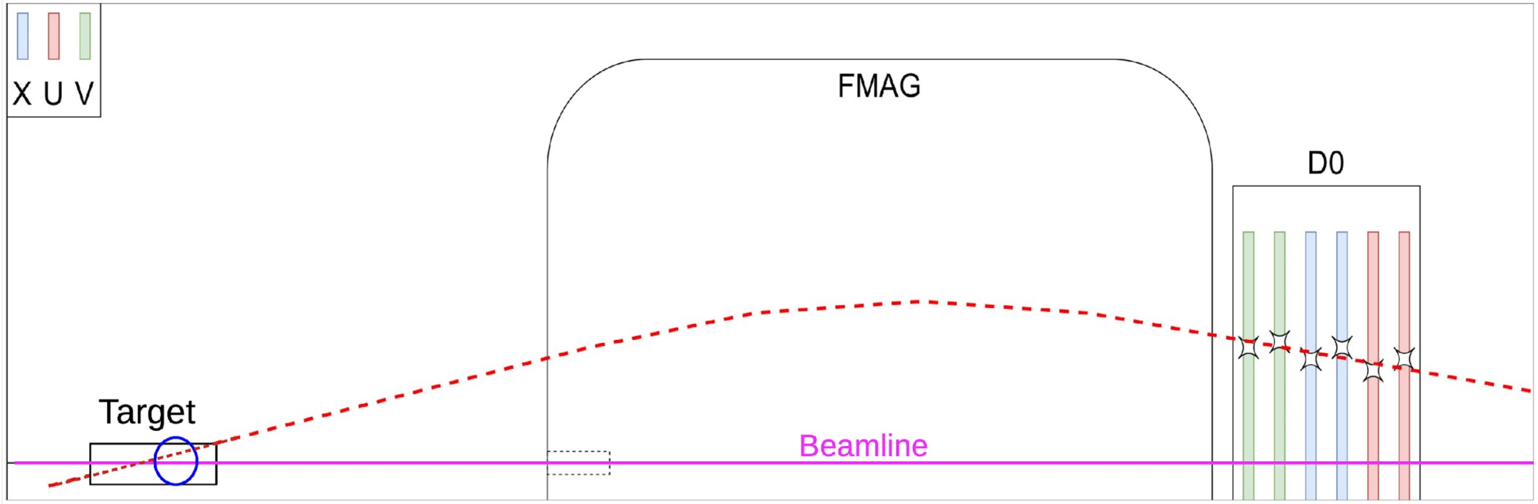
U: wires at +14 degrees with respect to X wires

V: wires at -14 degrees with respect to X wires



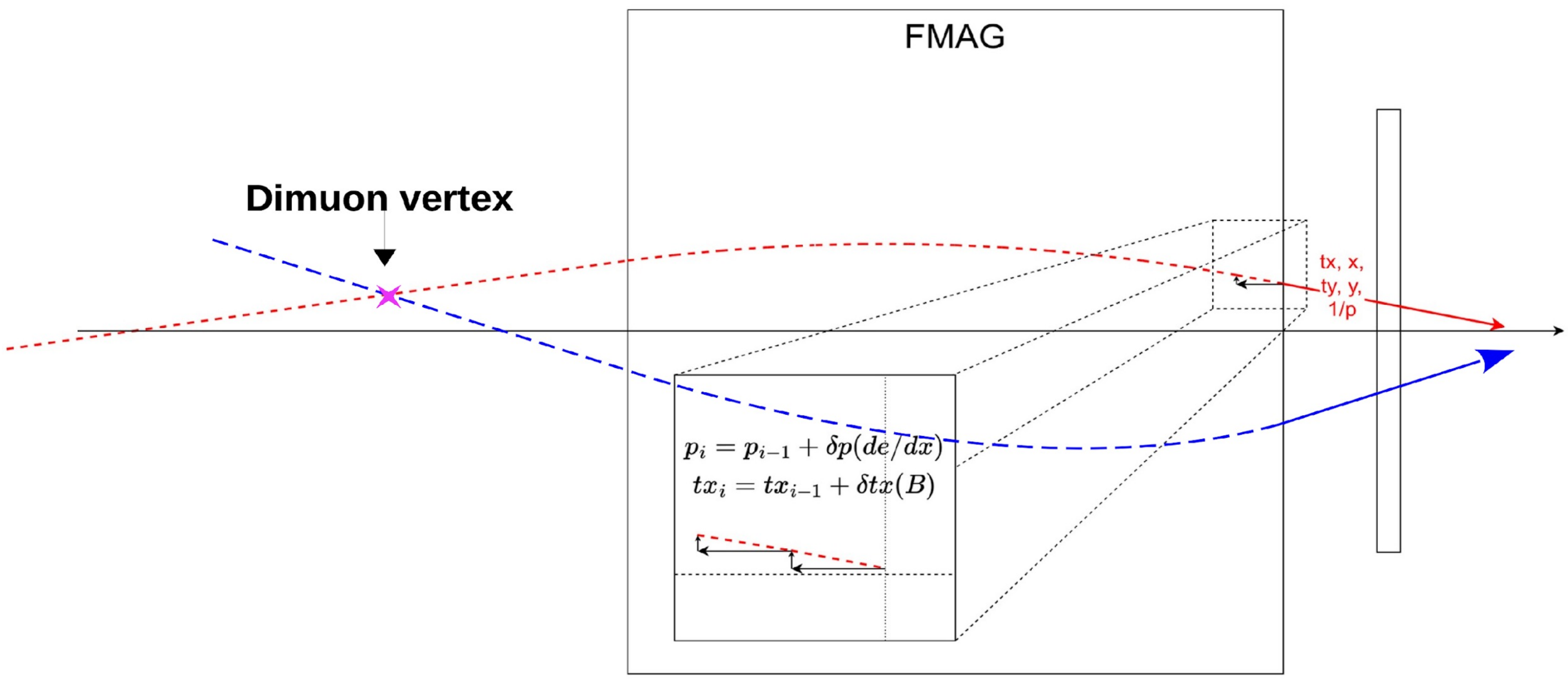
GPU Vertexing

- Back-swimming the reconstructed track through Fmag,
- Extrapolate the track to the target plane,
- Use of distance of closed approach (DOCA) from the beam line to estimate the vertex position

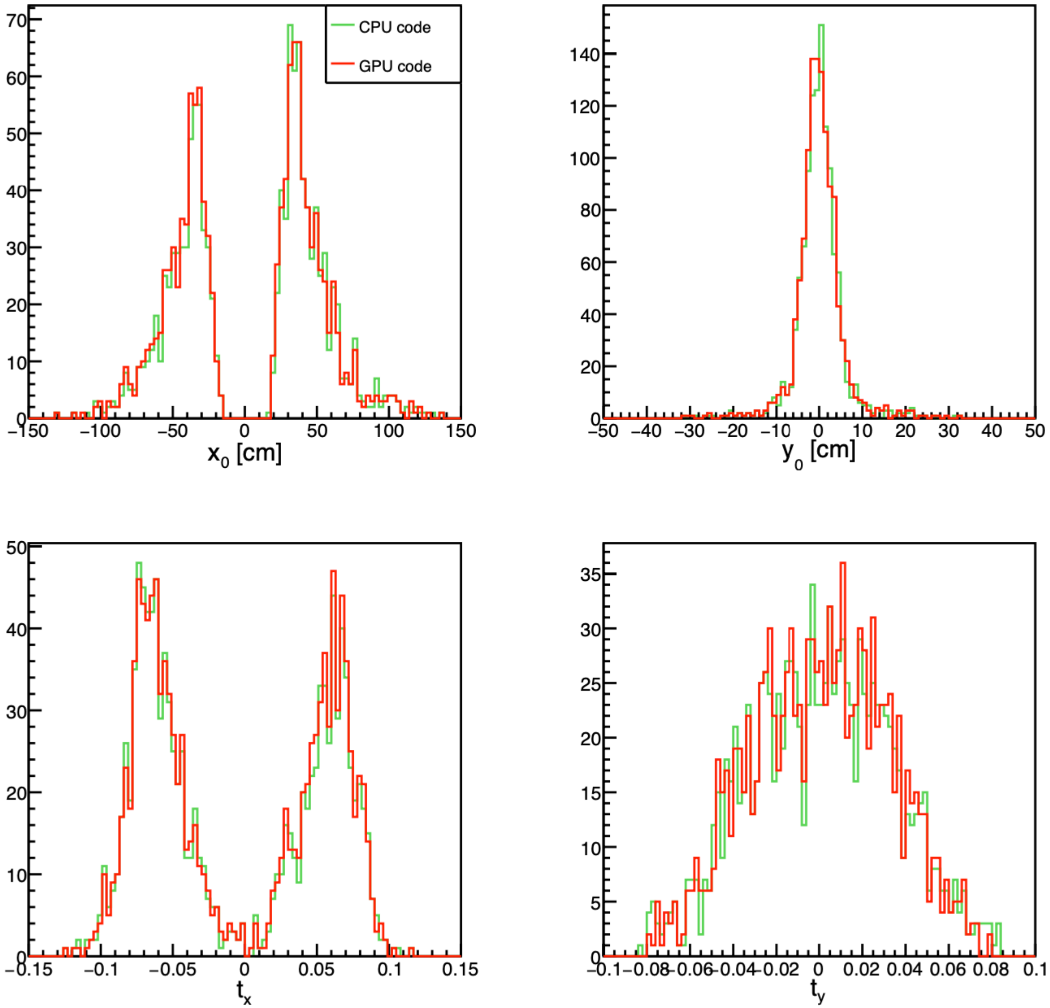


GPU Dimuon Reconstruction

- Implementation of the dimuon reconstruction is based on:
  - Pairing tracks of different charges,
  - Evaluating the vertex as the DOCA between the two muons tracks,
  - Reconstructing the momentum of each track as the momentum at the position of the dimuon vertex, not at the single-track original vertex



Track Reconstruction: GPU OR vs. CPU Ktracker



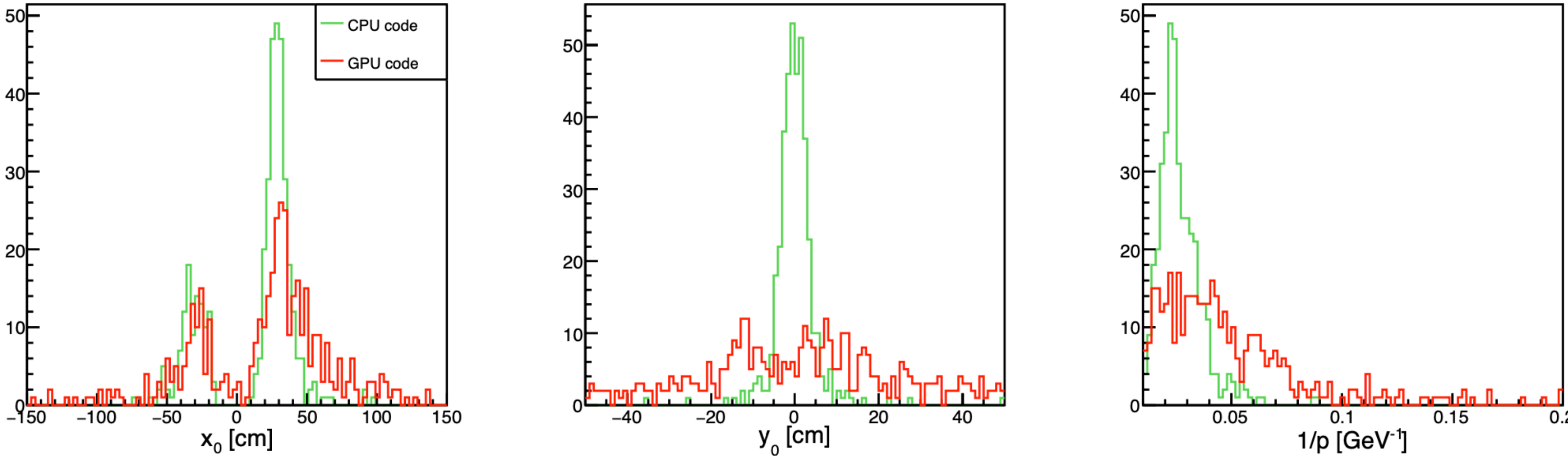
Pure Monte Carlo Dimuons:

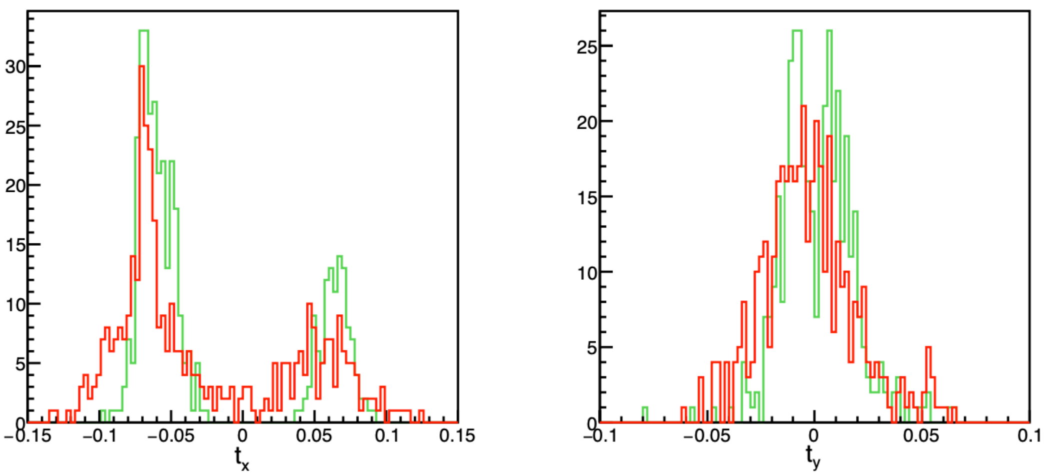
x_0, y_0 : track position at origin

t_x, t_y : track slope

p : momentum

Track Reconstruction: GPU OR vs. CPU Ktracker

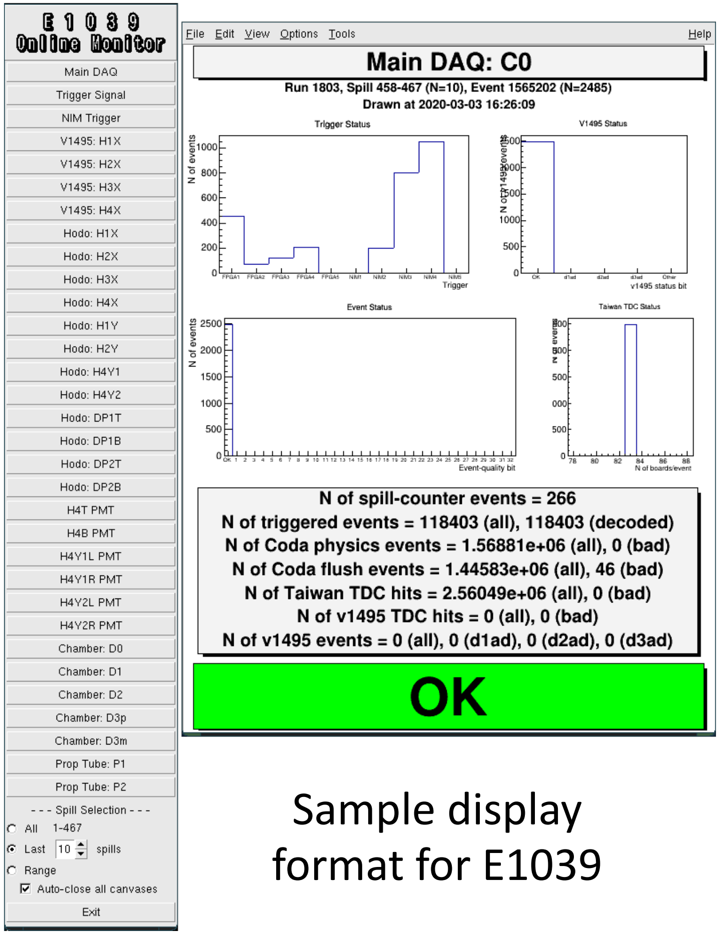




E1039 Commissioning Data
Run: 6155, Spill: 1941907
x_0, y_0 : track position at origin
t_x, t_y : track slope
p : momentum

GPU Online Display

- Copy the histogram’s arrays produced by the GPU OR to CPU
- Produce an output ROOT file per spill
- A new OnlMonClient module is being developed as an integral part of the E1039 online monitoring display
- The module will read, draw and update the ROOT file per spill



Sample display format for E1039

Summary and Outlook

- The SpinQuest experiment will provide crucial insight on the question of the nucleon spin puzzle
  - First measurement of the antiquarks Sivers function using polarized Drell-Yan
  - Unique opportunity for clean, model-independent extraction of sea quark transversity

- GPU online reconstruction program is close to deployment
  - GPU offers significant performance improvement compared to CPUs
  - GPU and CPU track reconstruction results compare reasonably well
  - Optimization of the code for SpinQuest real data processing is near completion
  - Online display plans are in progress

This work is supported in part by the U.S. DOE award #: DE-FG02-07ER41528

Thank you!