Computing & Software, A.I. Report

C. Fanelli, D. Lawrence

ECCE Bi-Weekly Meeting
June 7 2021

EIC/ECCE Computing Resources

For more information:
[J. Lauret Computing Report at EICUG, 5/20/2021]

Storage

- Pledged resources

BNL	1 PB
JLab	1 PB

Availability:

BNL
- physical disk in place
- S3 access, but details in /cvmfs not worked out
- Authentication?

} *Jerome L.*

JLab
- physical disk installed w/ base OS
- Networking this week + zfs filesystem
- Possibly available next week (for onsite work)

Compute

- Pledged resources

BNL	2k cores - 4 months
JLab	2k cores - 4 months
OSG	4k cores

Availability:
now-ish

- ***All pledged resources are for "EIC" and not for specific proto-collaborations***
  - Fairly allocate between ECCE, ATHENA, CORE, …
    - UG formation of Computing Coordination Group (CCG)
      - Committee forming. More details on proto-collaboration membership
    - Start with administrative controls for quota management

	factor	units	err on value	err on total
Total ep σ	45	μb	0	0
Integrated luminosity	5	fb-1	0	0
Fraction of total σ/physics chan.	1.00E-02		1.00E-02	2.25E+00
Frac. signal events needed for simulation	2.00E-02		2.00E-02	2.25E+00
Num. physics generators to simulate	25		5	0.45
Num. detector configurations to simulate	2		1	1.125
Total number of events to simulate	**2.25** *Bevents*	billion events		*contingency* **3.4** *Bevents*
Event size	300	kB/event	150	0.34
Total number of events	2.25	billion events	3.4	1.02
Total storage	**0.675** *PB*	PB		*contingency* **1.08** *PB*

Cameron's SIDIS 1M event test

VERY PRELIMINARY

Estimate + 1σ
=
1.7PB

ECCE Compute Estimate

Disclaimer:
These numbers are preliminary and based on a very small sample test using events generated for sPHENIX and not necessarily ECCE physics

2 jobs run on OSG nodes

Default events from tutorial

https://www.phenix.bnl.gov/WWW/publish/phnxbld/sPHENIX/files/sPHENIX_G4Hits_sHijing_9-11fm_00000_00010.root

Total committed by [(BNL + JLab)*75% + OSG] / 2 = 4k cores x 4 months (x 75%) = **10 Mcore-hrs**

expect ~ ½ for ECCE

102 Events - avg. time : 2,644 sec (single thread)
2 Events - avg. time : 215 sec (single thread)

time/event: 24.3 sec
overhead: 166 sec

assume 2k events/job

for **2.25B** events:
15.2 Mcore-hrs for event simulation
51.9 kcore-hrs for overhead *(program startup and shutdown)*

for **1.0B** events:
6.8 Mcore-hrs for event simulation
23.1 kcore-hrs for overhead *(program startup and shutdown)*

Ongoing Activities and Next Meeting

- S3 integration
- Updated simulation timeline
- Top level production scripts
- Metadata

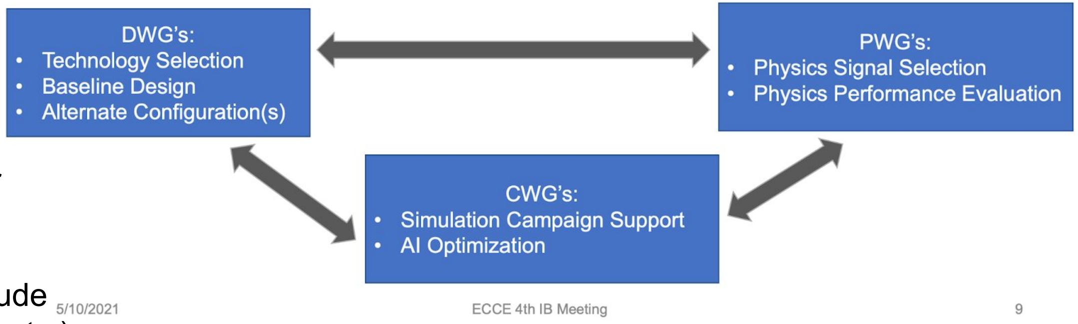
Software & Computing Bi-weekly meeting:

Wed. June 9 @ 13:00 EST

https://indico.bnl.gov/event/12088/

A.I. WG

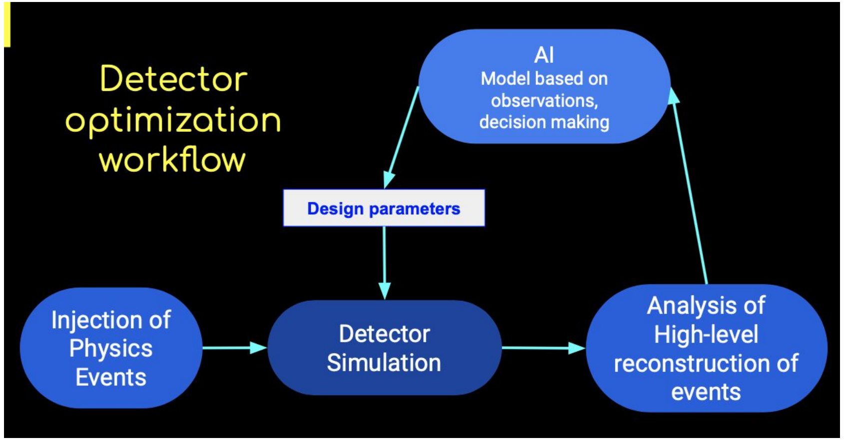
(convener: W. Phelps)



- During the proposal phase we will work with other working groups (physics and detector) to assist in **detector design** optimization
- In the future this scope could be expanded to include other AI applications as well (AI assisted tracking, etc.)

- Keeping in mind the "inner to outer" design process/strategy discussed at the 5th IB:

Optimize technology choices together with physics performance
- Formulate a dynamic timeline for decision making for the global simulation
- Start with the design of the inner layers, e.g., fix tracking and PID and then work outwards (radially and in polar angle), e.g., for PID it is important to have knowledge of the magnetic field and the tracking resolution and also minimizing material



- Suggested/Identified the following activities (discussion later):
  - Tracking (Brunel, MIT, Regina, work in progress)
  - PID --- DIRC (CNU, MIT), d-RICH (MIT d-RICH paper)
  - Calorimetry (CUA, MIT, Regina, work started within eRD1, link to presentation)
  - Far Forward --- ZDC (Duquense, JLab)

cfanelli@mit.edu, wphelps@jlab.org

A.I. WG Meeting

- A.I. WG Meeting on June 4. Link: https://indico.bnl.gov/event/12052/
  - Discussion on scope
    - Optimization does not mean necessarily “fine-tuning”. In a complex problem with, e.g., multiple design criteria (e.g., performance, cost, material) it helps identifying/approximate the best set of trade-off solutions (Pareto frontier) and decisions can be made based on that.
    - We want to use these algorithms to: (1) steer the design and suggest combinations of parameters that a “manual”/brute-force optimization will likely miss to identify; (2) further optimize some particular detector technology (see d-RICH paper, e.g., optics properties)
  - Provided examples (dRICH paper, Tracker using Fun4All). See next slides.

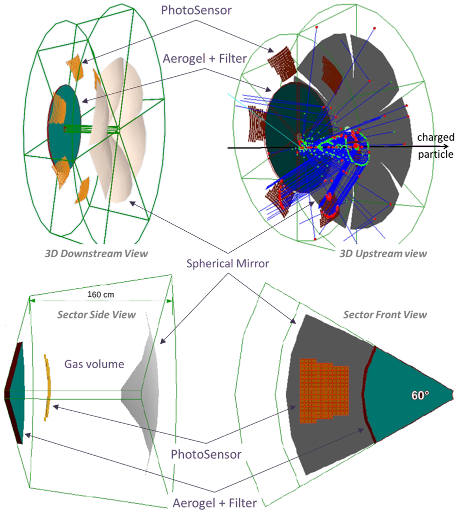
- References
  - CF’s lectures at AI4NP: https://github.com/cfteach/AI4NP_detector_opt
  - dRICH: https://iopscience.iop.org/article/10.1088/1748-0221/15/05/P05009/meta

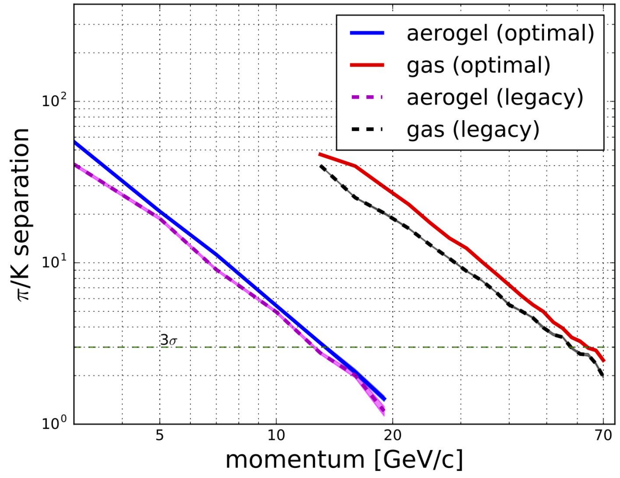
- Next Meeting this week (originally on Thu, but we will announce soon new day/time to not overlap with Detector WG)

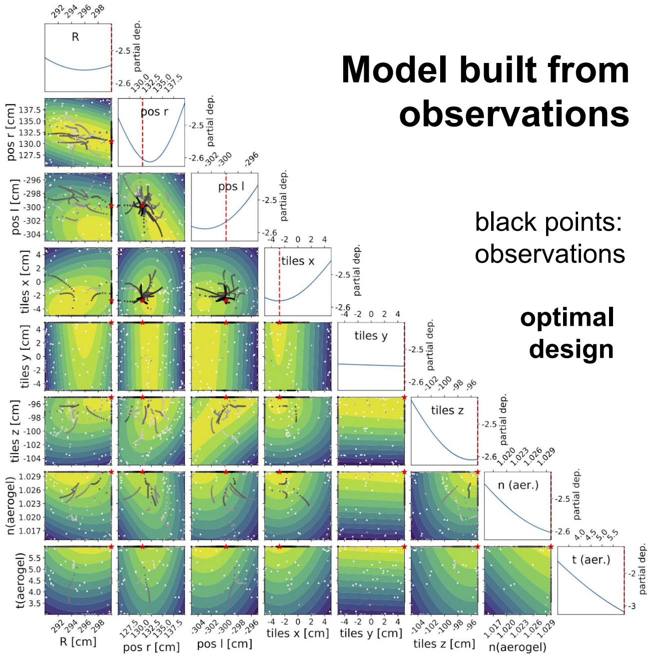
dRICH Example

https://iopscience.iop.org/article/10.1088/1748-0221/15/05/P05009/meta

parameter	description	range [units]	tolerance [units]
R	mirror radius	[290,300] [cm]	100 [μm]
pos r	radial position of mirror center	[125,140] [cm]	100 [μm]
pos l	longitudinal position of mirror center	[-305,-295] [cm]	100 [μm]
tiles x	shift along x of tiles center	[-5,5] [cm]	100 [μm]
tiles y	shift along y of tiles center	[-5,5] [cm]	100 [μm]
tiles z	shift along z of tiles center	[-105,-95] [cm]	100 [μm]
$n_{aerogel}$	aerogel refractive index	[1.015,1.030]	0.2%
$t_{aerogel}$	aerogel thickness	[3.0,6.0] [cm]	1 [mm]







Model built from observations

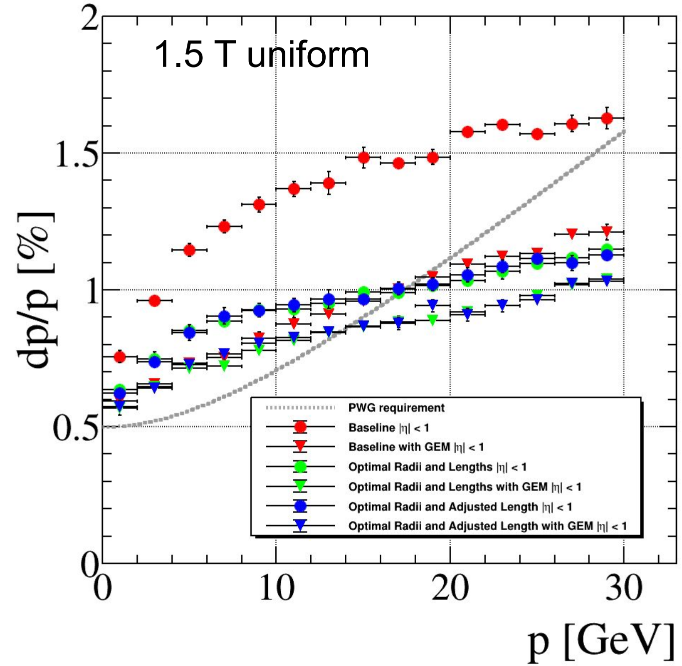
black points: observations

optimal design

Si Tracker Exercise

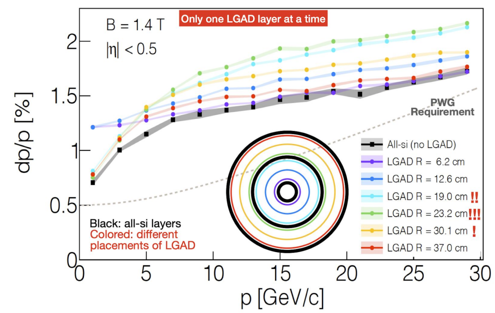
K. Suresh, U. of Regina & CF

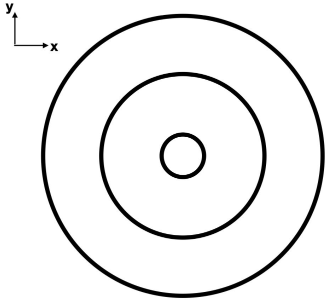
https://github.com/reynier0611/g4lblvtx/ (private repo)
Fun4All_G4_simplified_v2_DIRC_barrel_GEM.C



See very nice talk by R. Cruz , 05/21/2021

Placement of LGAD layer inside the tracking volume

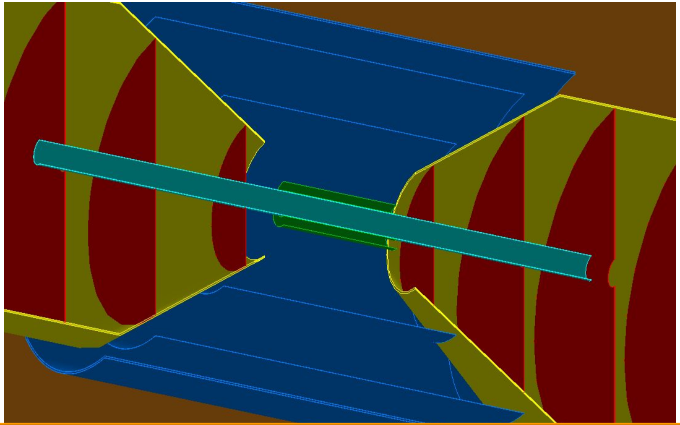




All-silicon tracker layers

Barrel layer	radius [cm]
1	3.30
2	5.70
3	21.00
4	22.68
5	39.30
6	43.23

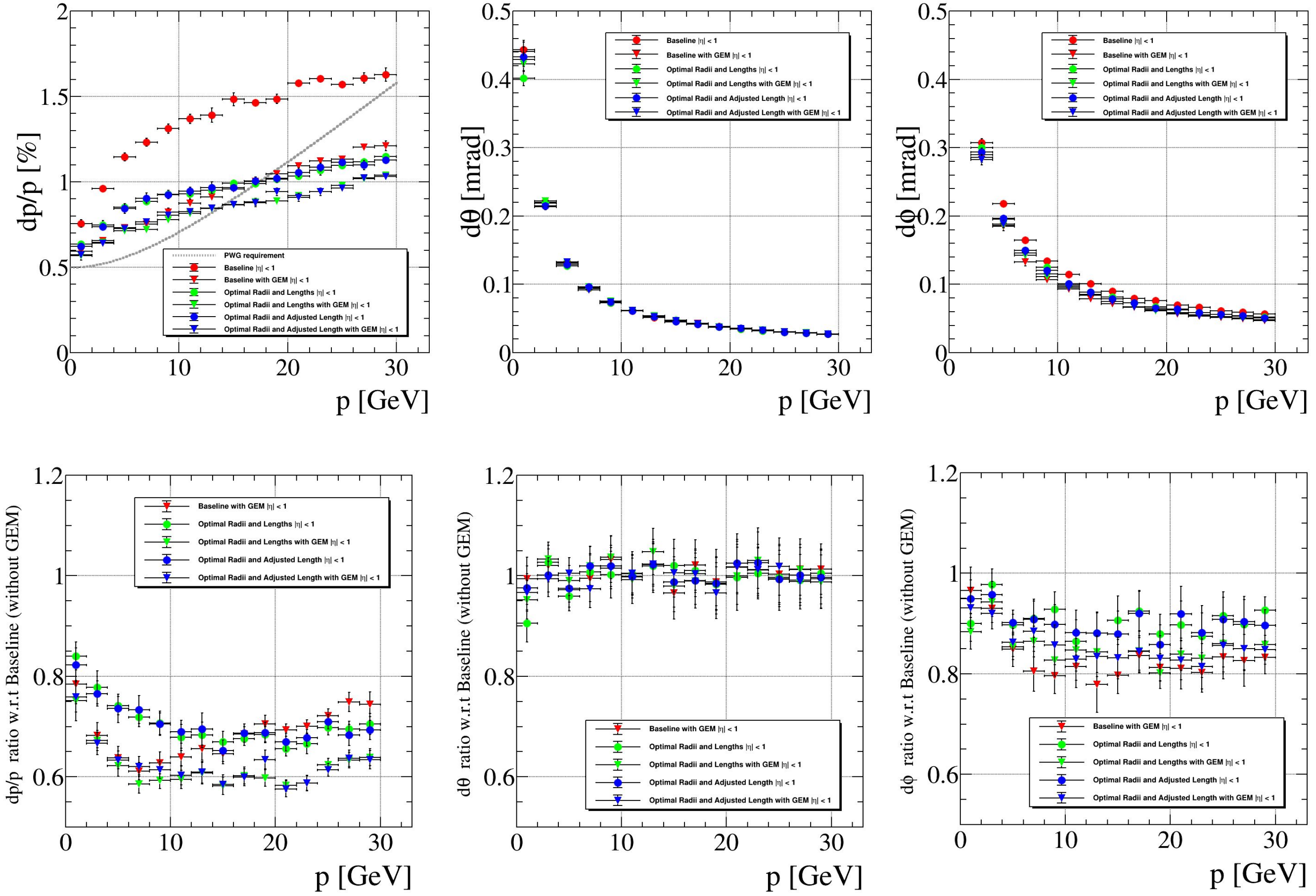
Baseline

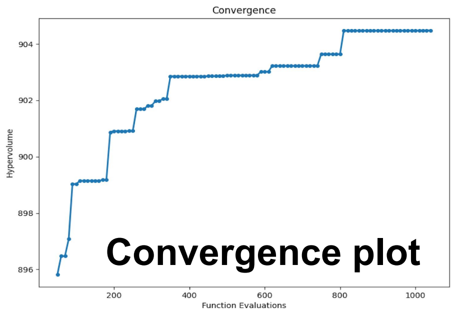


Optimised Barrel radii = [25.00, 41.84, 49.90, 50.45] cm

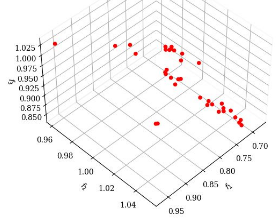
Compared to June 4, and following some suggestions received on adjusting the lengths of the layers, this plot has been updated for |η|<1, 0<P<30 GeV/c, and lengths of the layers increased (see Figure on the right).

Si Tracker Exercise |η|<1





Pareto front



Improvement compared to baseline (without GEM) by changing radii

1-<R(dp/p)>	1-<R(dθ)>	1-<R(dφ)>
30±2%	**2±2%**	**11±2%**

Backups

Current estimate for 2021 ECCE storage requirement

- sPHENIX mock data challenge: 100M pp events -> 130TB
- Estimate 1B events needed for proposal development -> 1.3PB
- Include contingency for larger eA event size, etc … -> 2PB

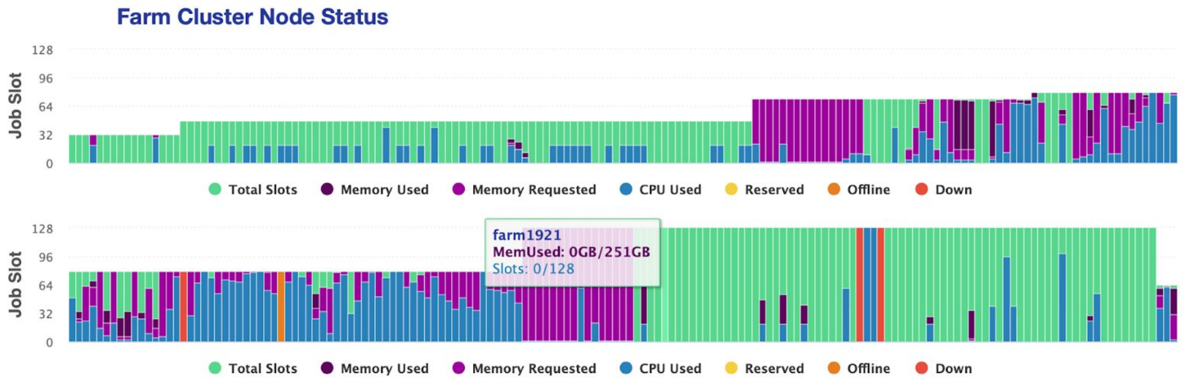
1M event test sets

	pythia6 - general	**pythia6 - SIDIS**	**pythia6 - HF & Jets**
DST	159GB	220GB	177GB
Evaluator	59GB	78GB	60GB
Total	**218GB**	**298GB**	**237GB**

4.3x smaller than sPHENIX MDC numbers

Jefferson Lab Computing Resources

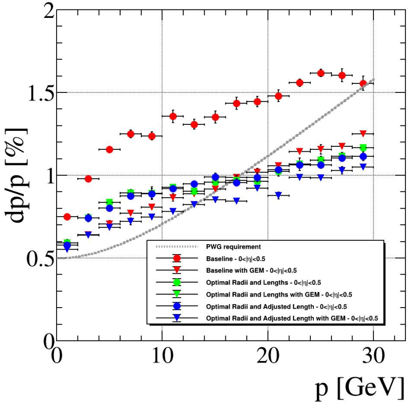
- The scientific computing cluster known as the “farm” has 25k cores
  - EIC Projects are allocated 10%
- 1PB for EIC use

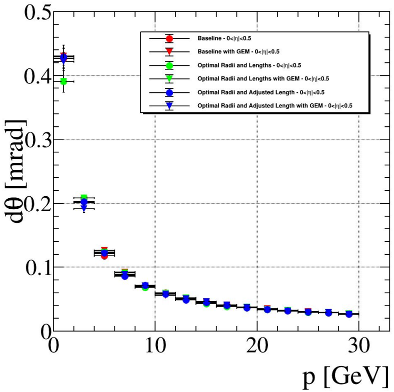


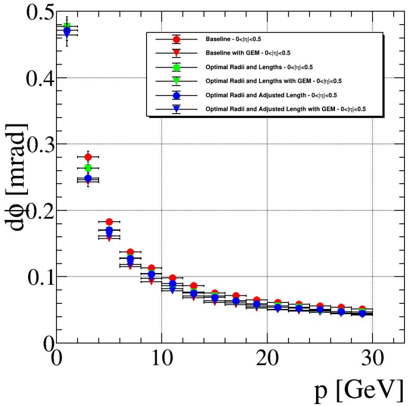
The “Farm” status on an unusually quiet day (scicomp.jlab.org)

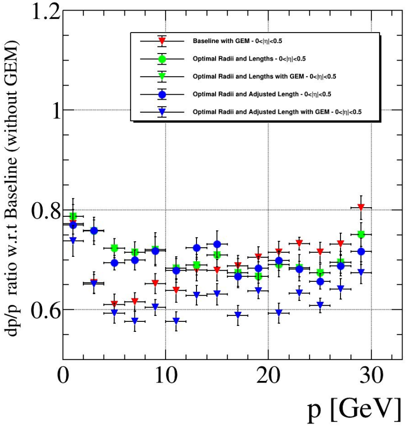
- Batch use as well as interactive use supported with
  - Nodes with up to two 32 core AMD Epyc Processors (128 threads), 256GB Ram, 1TB SSD local storage
  - 3 Nodes with 4 Titan RTX Cards (24 GB Memory)
  - GPU nodes also available through jupyterhub.jlab.org
  - Additional GPU nodes arriving soon

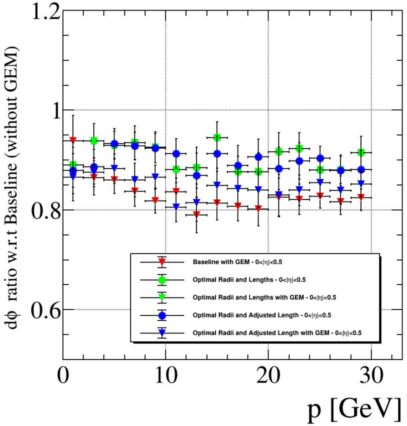
Si Tracker Exercise |η|<0.5

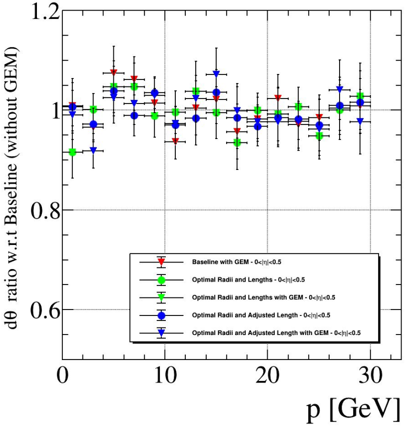




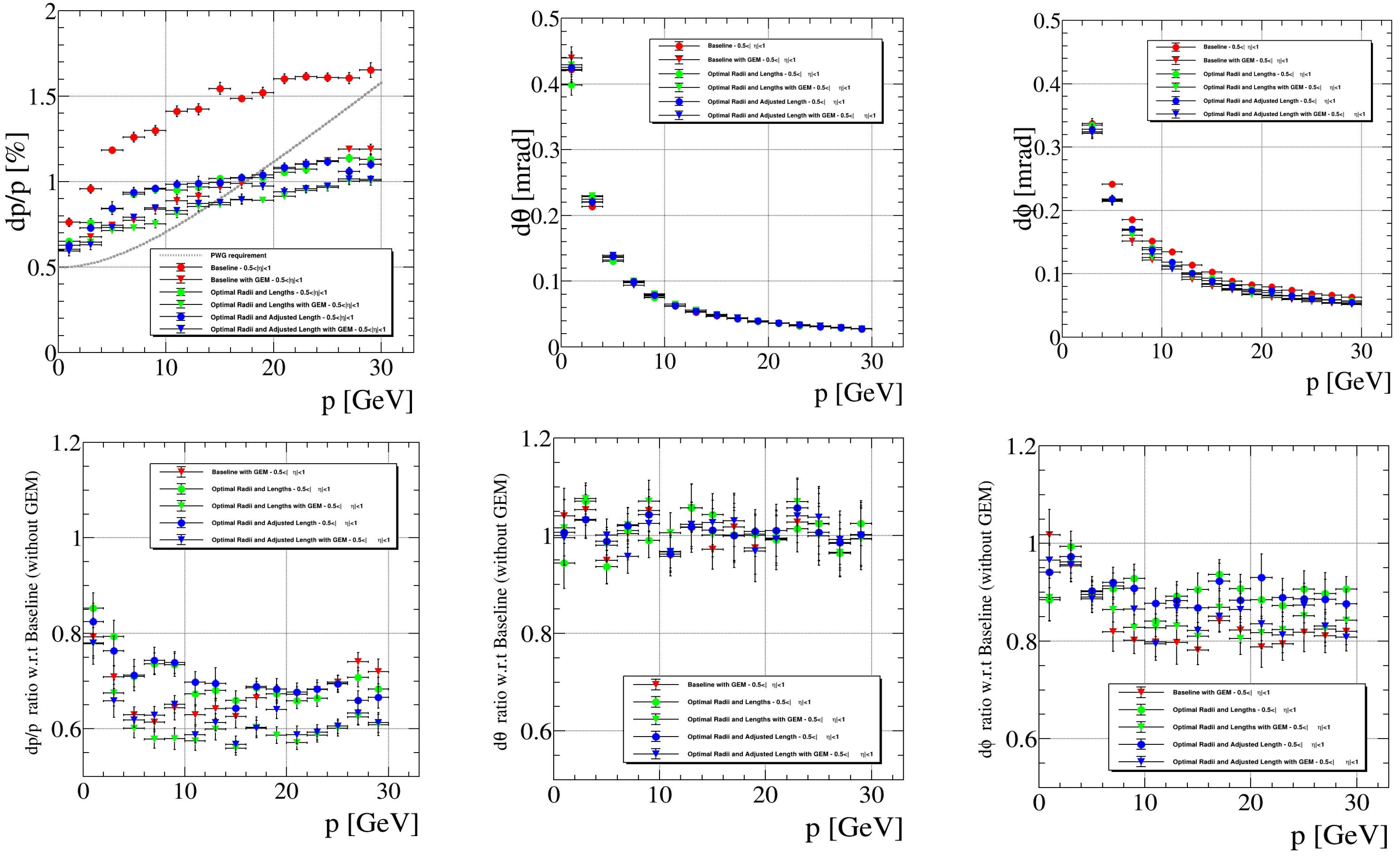




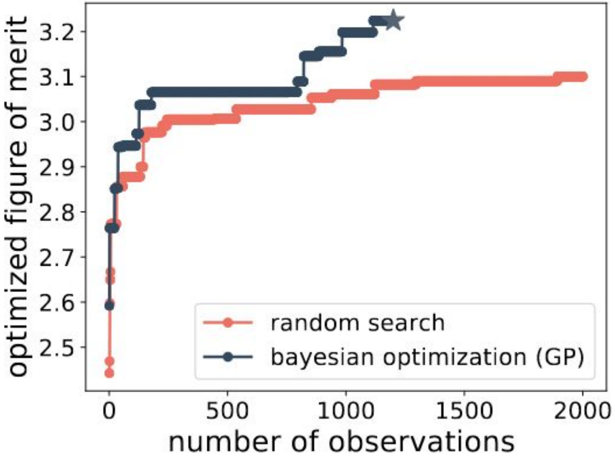


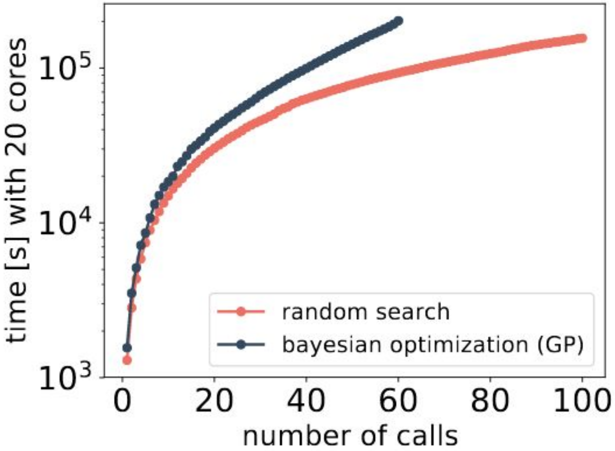


Si Tracker Exercise 0.5<|η|<1.0



Comparison with Random Search





Each call:
400 tracks generated/core
20 cores

1 design point ~ 10 mins/CPU

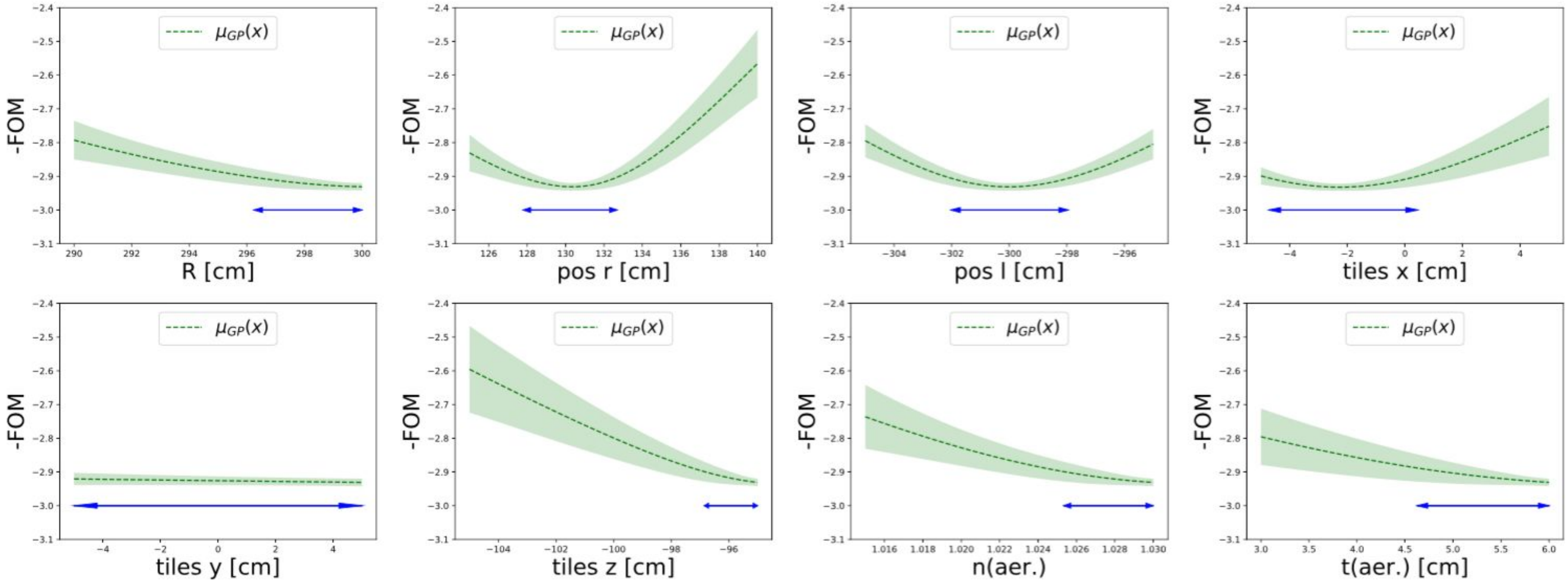
Budget: 100 calls

- BO with GP scales cubically with number of observations.
- Bayesian optimization methods are more promising because they offer principled approaches to weighting the importance of each dimension.
- For this 8D problem - even with 50 cores, RS looks unfeasible due to the curse of dimensionality.
  - Recall that the probability of finding the target with RS is $1-(1-v/V)^T$, where T is trials, v/V is the volume of target relative to the unit hypercube

Bergstra, Bengio, “Random search for hyper-parameter optimization”, J. Mach. Learn. Res.13 (Feb) (2012) 281–305.

Tolerance Regions

- BO provides a model of how the FoM depends on the parameters, hence it is possible to use the posterio to define a tolerance on the parameters (regions ensuring improved PID, see previous slide).



- Larger than the construction tolerances on each parameter.
  Notice a small lateral shift of the tiles has negligible impact on the PID capability.