

ePIC-Analysis

Common Physics Analysis Software for the EIC

Christopher Dilks

Duke University

for the ePIC Collaboration

A large, faint, light gray watermark of the Duke University logo is visible on the left side of the slide. It features a stylized 'D' with a 'U' inside, and the text 'Duke UNIVERSITY' below it.

Duke
UNIVERSITY

Research supported by the



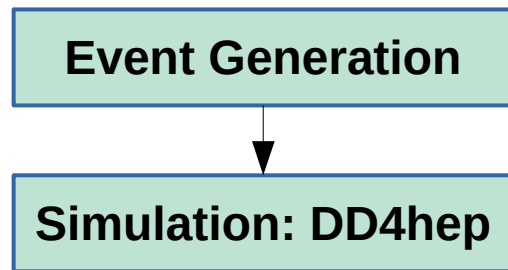
U.S. DEPARTMENT OF
ENERGY

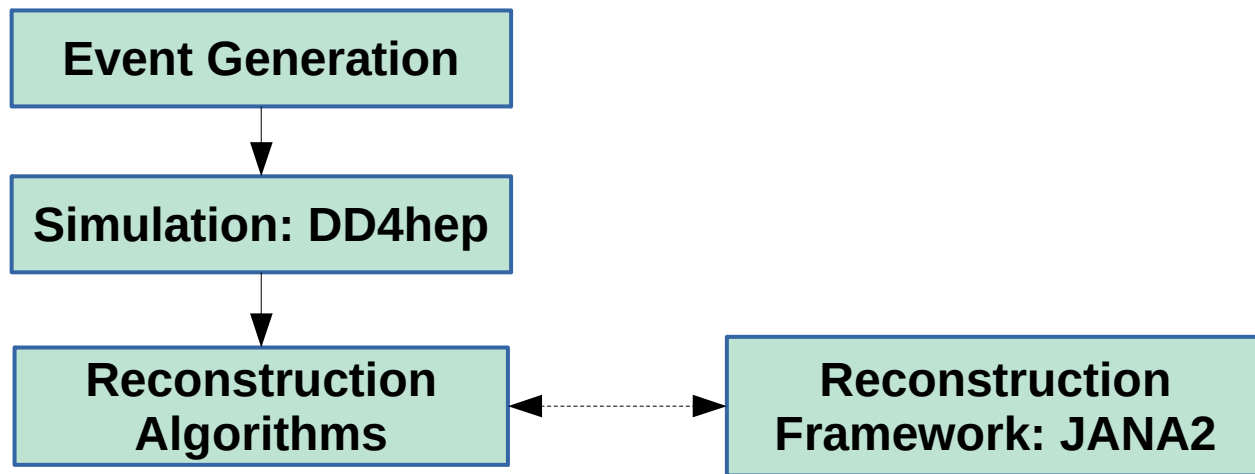
Office of
Science

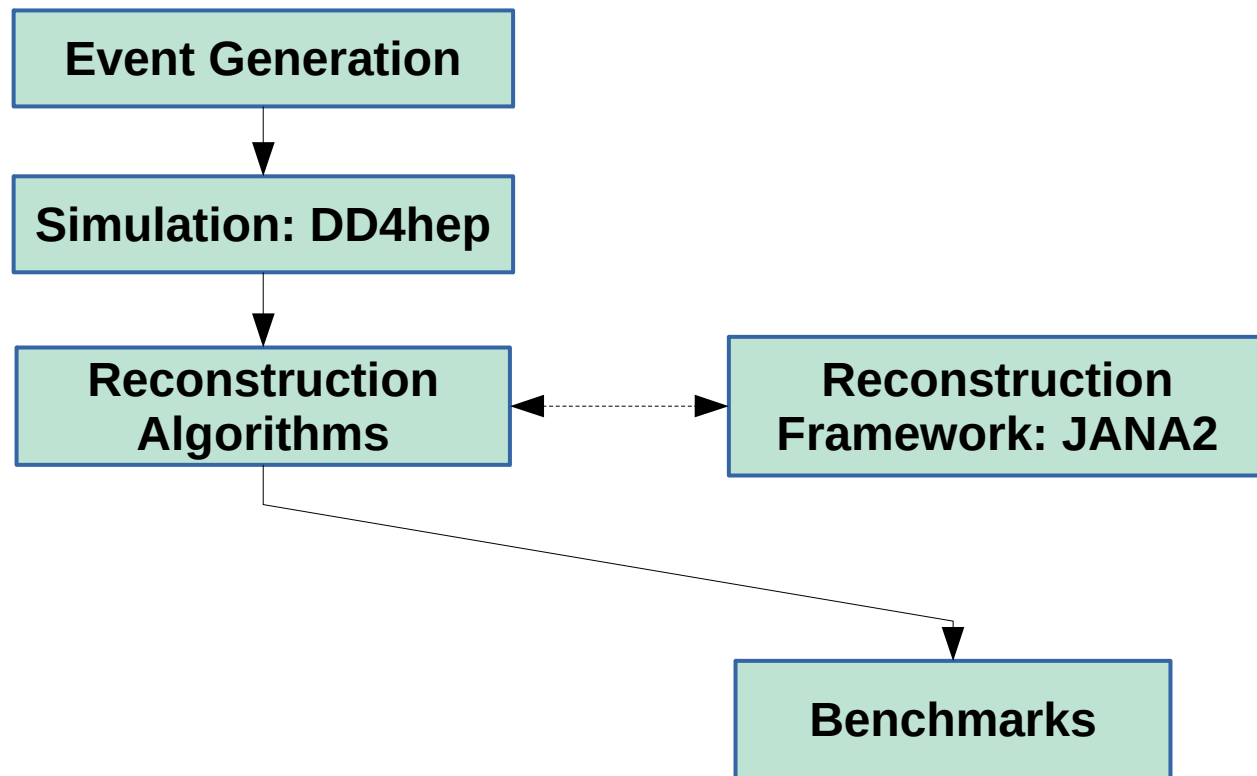
picture of EIC

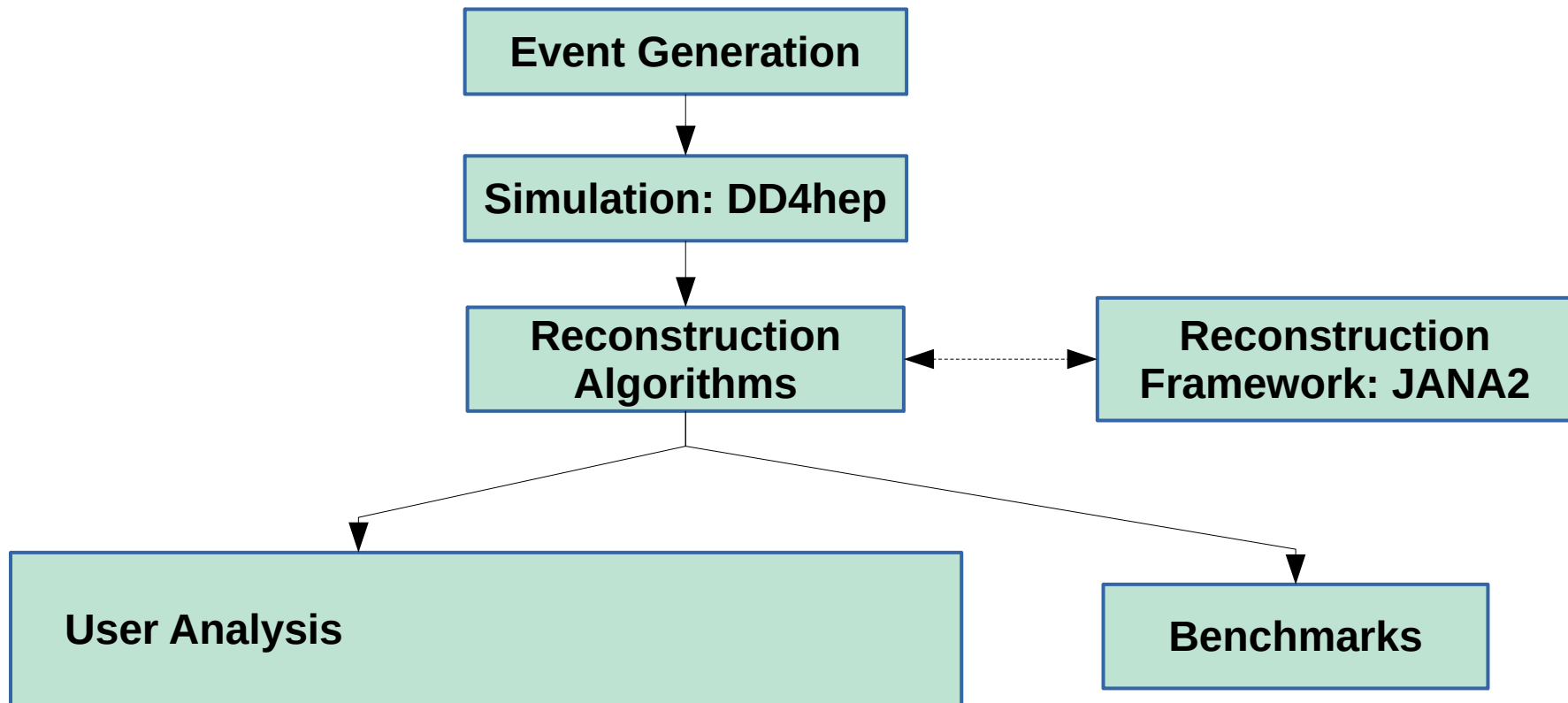
picture of ePIC

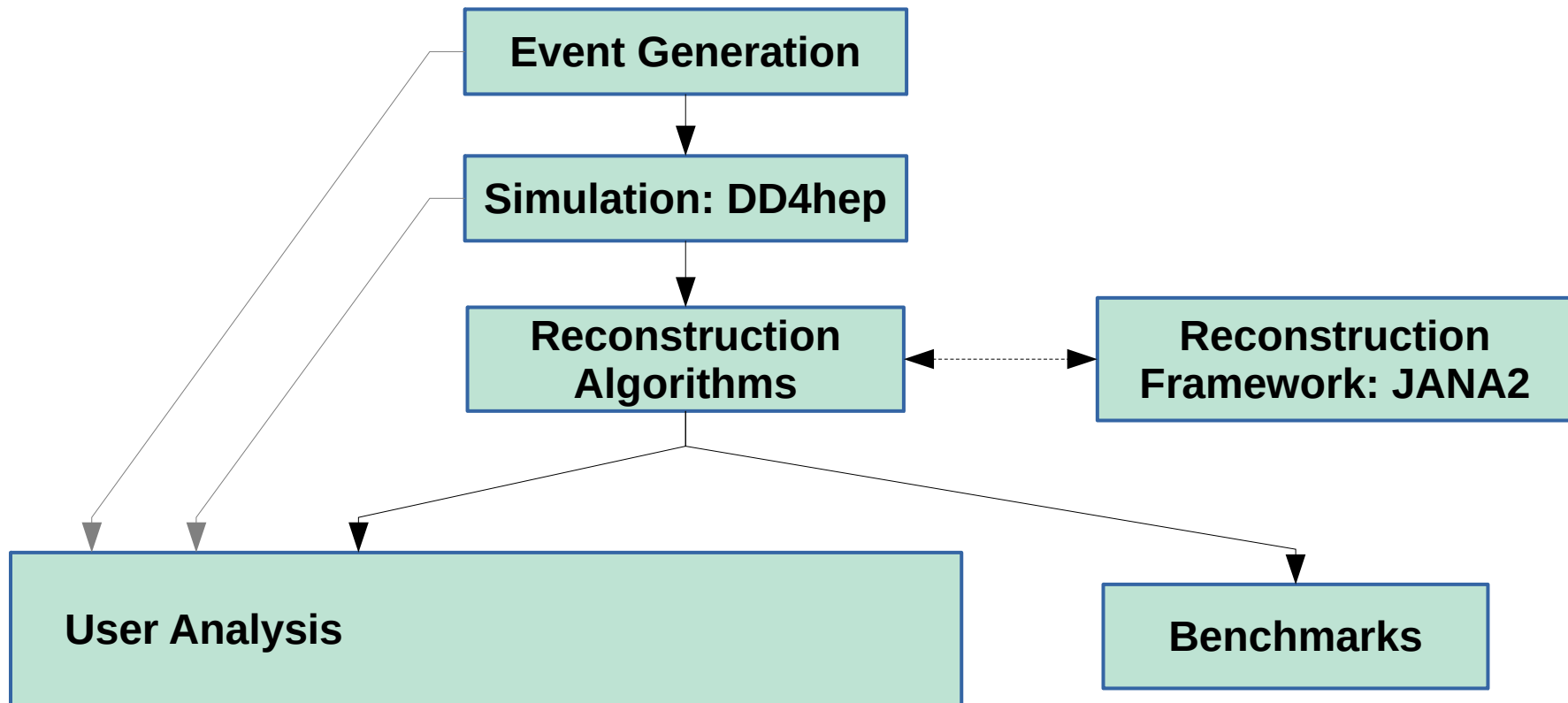
very brief, just to provide context
where this SW is applied

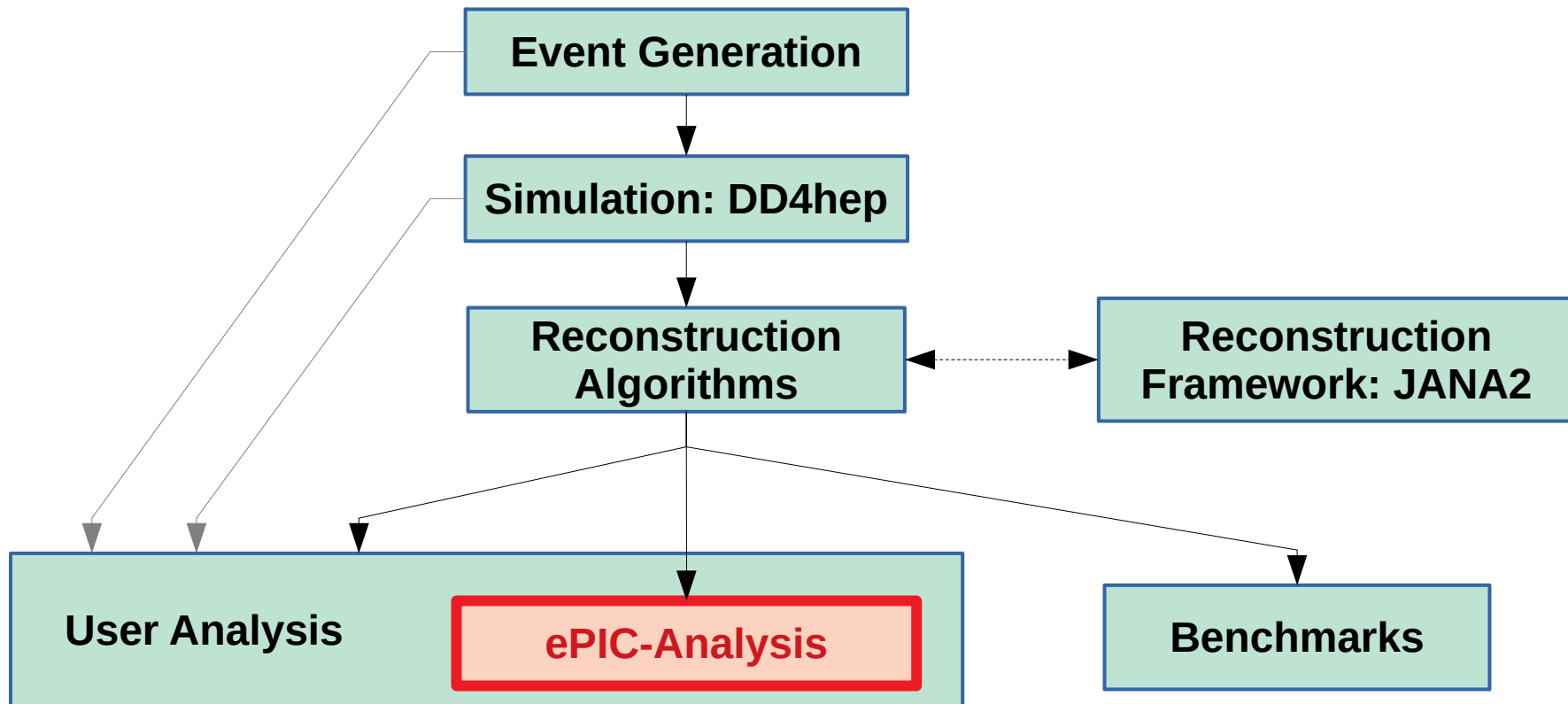


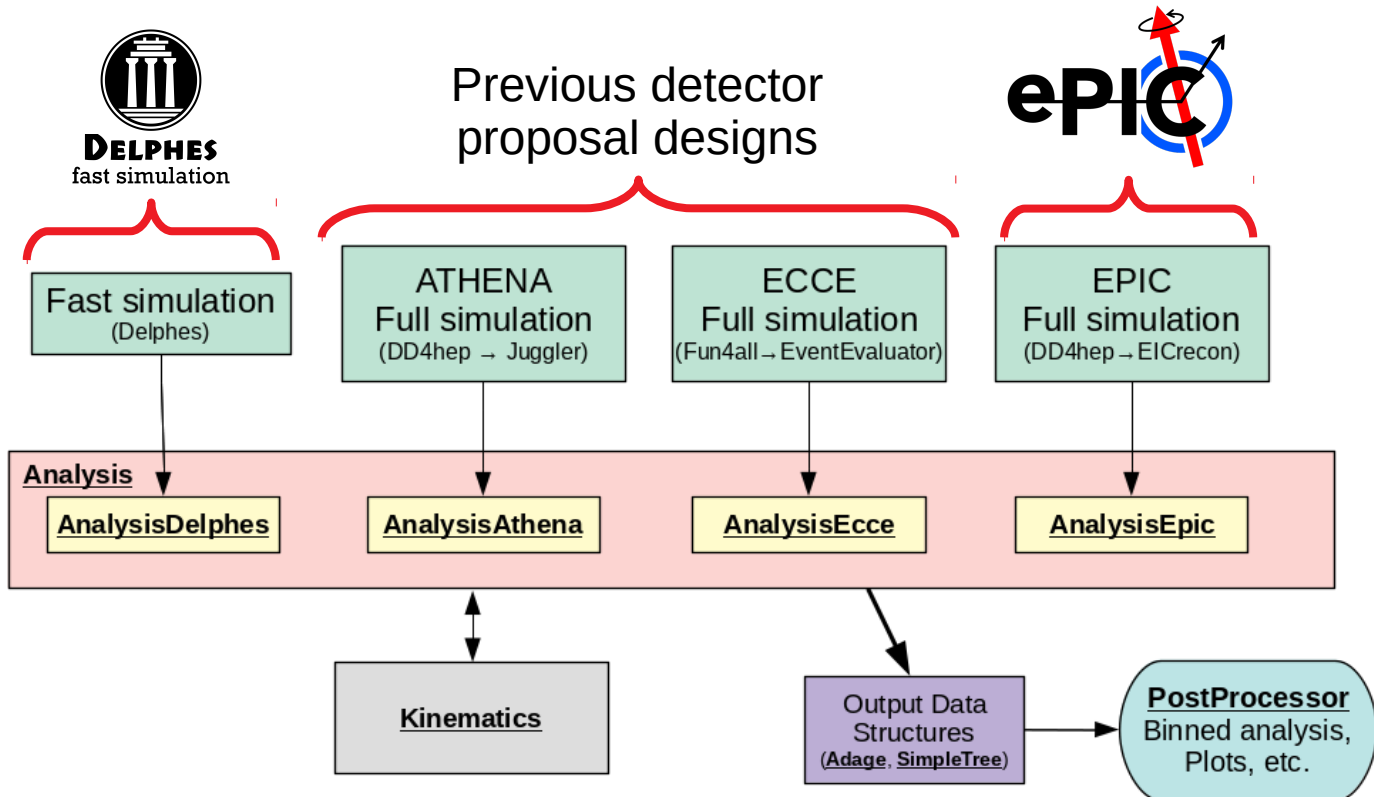






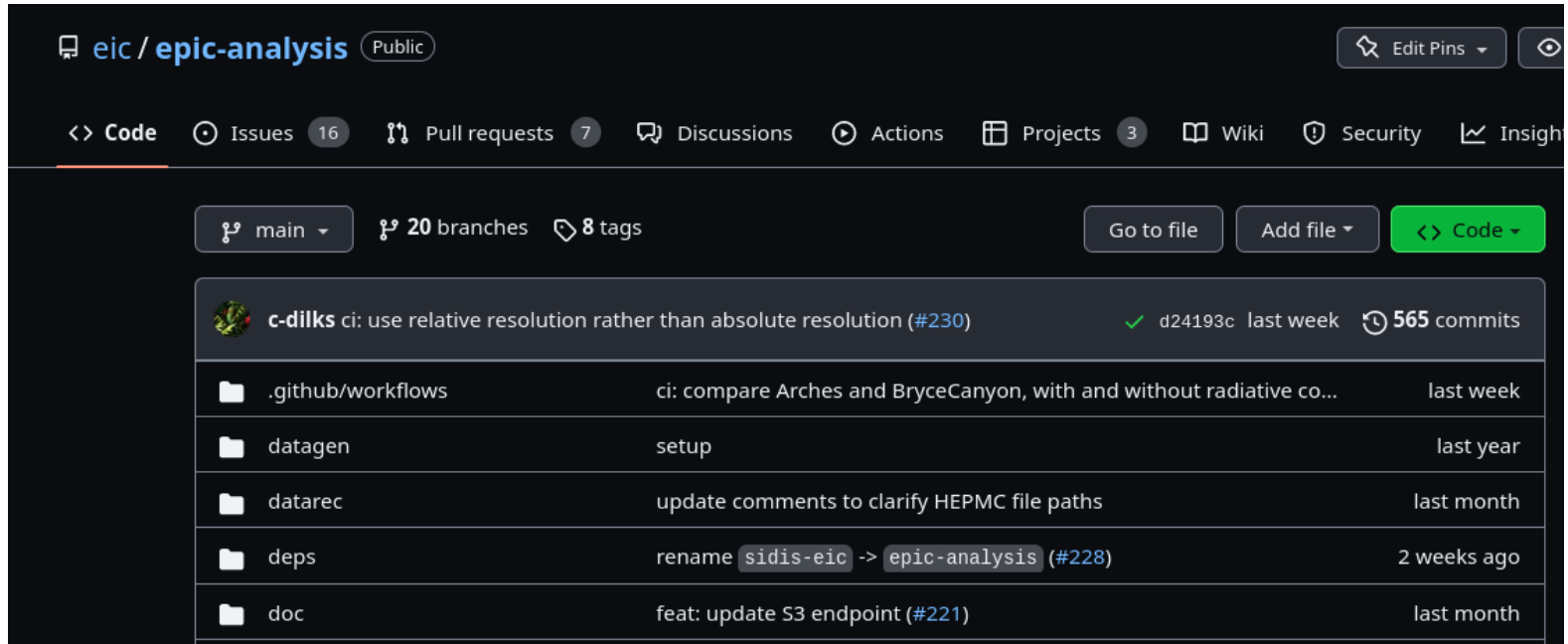






- TTrees
- Multidimensionally binned objects (Histograms, ...)

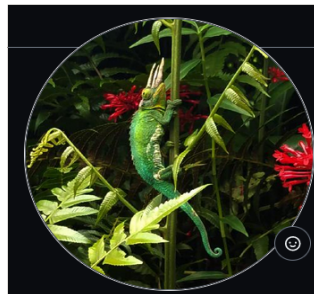
<https://github.com/eic/epic-analysis>



The screenshot shows the GitHub repository page for `eic/epic-analysis`. The repository is public and has 16 issues, 7 pull requests, 3 projects, and 8 tags. The main branch is selected, and there are 20 branches in total. The repository is currently checked out to commit `d24193c` from last week, with 565 commits in total.

Commit	Message	Time
<code>c-dilks</code>	ci: use relative resolution rather than absolute resolution (#230)	last week
	ci: compare Arches and BryceCanyon, with and without radiative co...	last week
	setup	last year
	update comments to clarify HEPMC file paths	last month
	rename <code>sidis-eic</code> -> <code>epic-analysis</code> (#228)	2 weeks ago
	feat: update S3 endpoint (#221)	last month

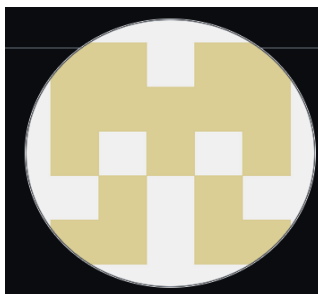
Thanks to Our Contributors!



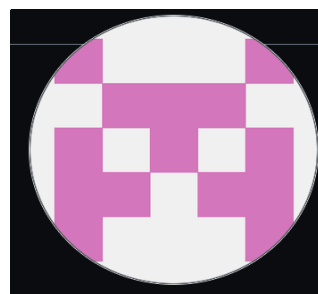
Christopher Dilks



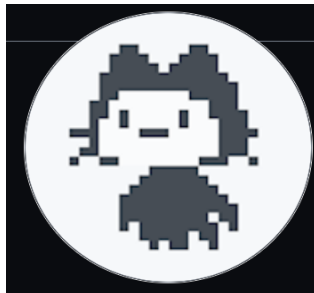
Connor Pecar



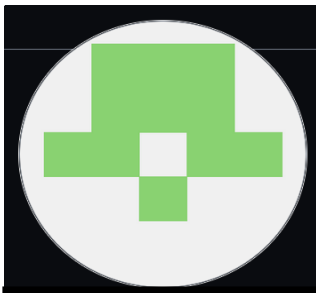
Duane Byer



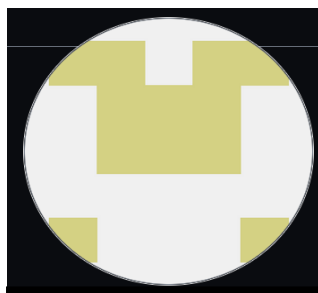
Gregory Matousek



Sanghwa Park



Matthew McEneaney



Ralf Seidl



Brian Page

TODO: add newer contributors: Dmitry K., Kevin A., others?

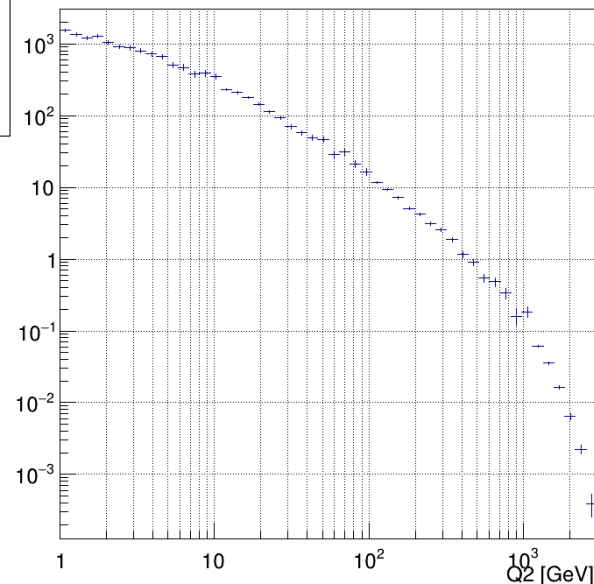
Q² Weighting

- ◆ Cross section falls rapidly with Q² → high Q² events are rare
 - Generate events in various bins of Q²
 - Re-weight them using the cross sections to combine their data
 - Populates statistics even at very high Q²
 - Allows for study of a broad range of Q², without having to wait for rare high Q² events
- ◆ ePIC-Analysis provides a common Q² weighting implementation

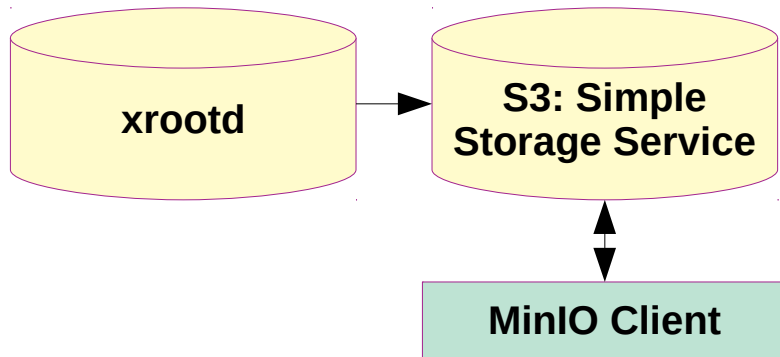
Q² Bins

- 1 – 10 GeV²
- 10 – 100 GeV²
- 100 – 1000 GeV²
- 1000 GeV² and above

Q² distribution, π⁺ tracks, p_T^{had}>0.10, W>3.00, x_F>0.00, 0.01<y<0.95, 0.20<z<0.90



Data Storage and Retrieval

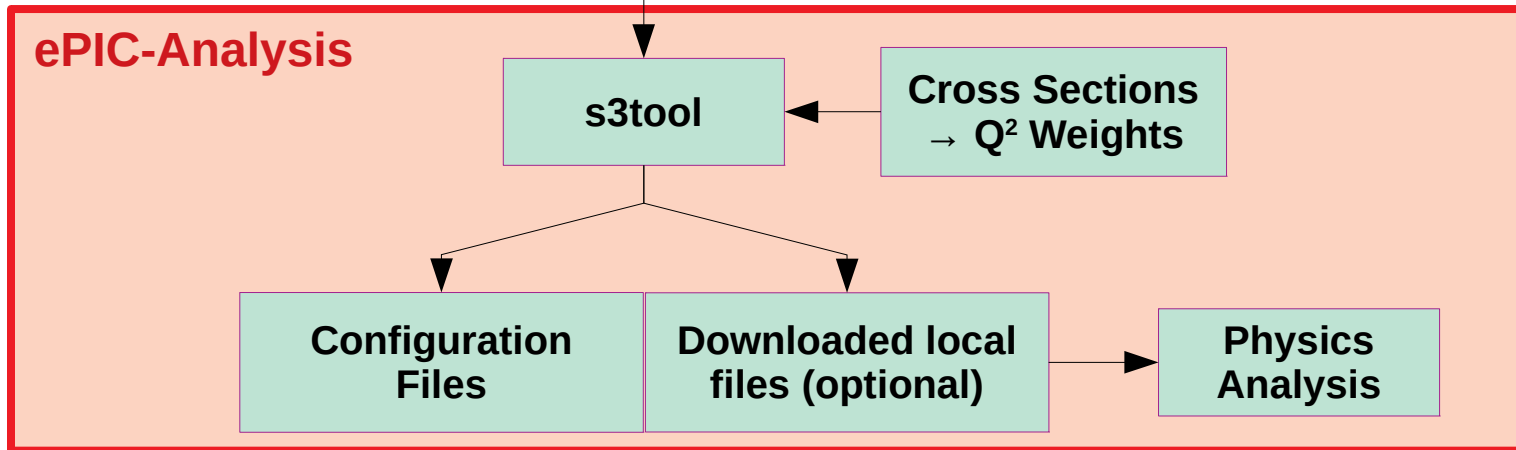


◆ S3: common cloud storage provider

- MinIO Client for read-access
- Streamable to ROOT:
 - `TFile::Open(s3_URL)`

◆ Tools in ePIC-Analysis

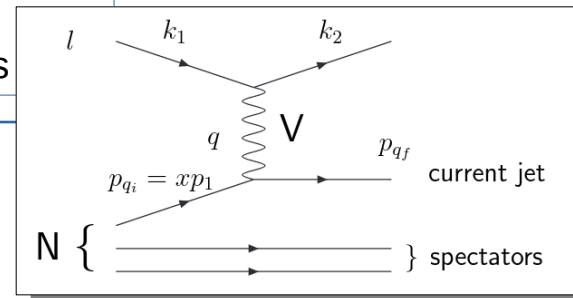
- Automated file retrieval (for streaming or downloading)
- Organization, Production versions
- Application of Q^2 weights



Kinematics Reconstruction

- ◆ Study SIDIS in a *particle collider* context
- ◆ Need to develop tools for accurate reconstruction of event kinematics

- | | |
|---|--|
| i) <i>Leptonic variables</i> | $q \equiv q_l = k_2 - k_1, \quad y_l = p_1 \cdot (k_1 - k_2) / p_1 \cdot k_1$ |
| ii) <i>Hadronic variables</i> [81] | $q \equiv q_h = p_2 - p_1, \quad y_l = p_1 \cdot (p_2 - p_1) / p_1 \cdot k_1$ |
| iii) <i>Jacquet-Blondel variables</i> [82] | $Q_{JB}^2 = (\vec{p}_{2,\perp})^2 / (1 - y_{JB}), \quad y_{JB} = \Sigma / (2E(k_1))$
$\Sigma = \sum_h (E_h - p_{h,z})$ |
| iv) <i>Mixed variables</i> [81] | $q = q_l, y_m = y_{JB}$ |
| v) <i>Double angle method</i> [83] | $Q_{DA}^2 = \frac{4E(k_2)^2 \cos^2(\theta(k_2)/2)}{\sin^2(\theta(k_2)/2) + \sin(\theta(k_2)/2) \cos(\theta(k_2)/2) \tan(\theta(p_2)/2)},$
$y_{DA} = 1 - \frac{\sin(\theta(k_2)/2)}{\sin(\theta(k_2)/2) + \cos(\theta(k_2)/2) \tan(\theta(p_2)/2)},$ |
| vi) <i>θ_y method</i> [84] | $Q_{\theta_y}^2 = 4E(k_2)^2 (1 - y_{JB}) \frac{1 + \cos(\theta(k_2))}{1 - \cos(\theta(k_2))}, \quad y_{\theta_y} = y_{JB}$ |
| vii) <i>Σ method</i> [85] | $Q_{\Sigma}^2 = \frac{(\vec{k}_{2,\perp})^2}{1 - y_{\Sigma}}, \quad y_{\Sigma} = \frac{\Sigma}{\Sigma + E(k_2)[1 - \cos(\theta(k_2))]}$ |
| viii) <i>$e\Sigma$ method</i> [85] | $Q_{e\Sigma}^2 = Q_l^2, \quad y_{e\Sigma} = \frac{Q_l^2}{s x_{\Sigma}}$ |

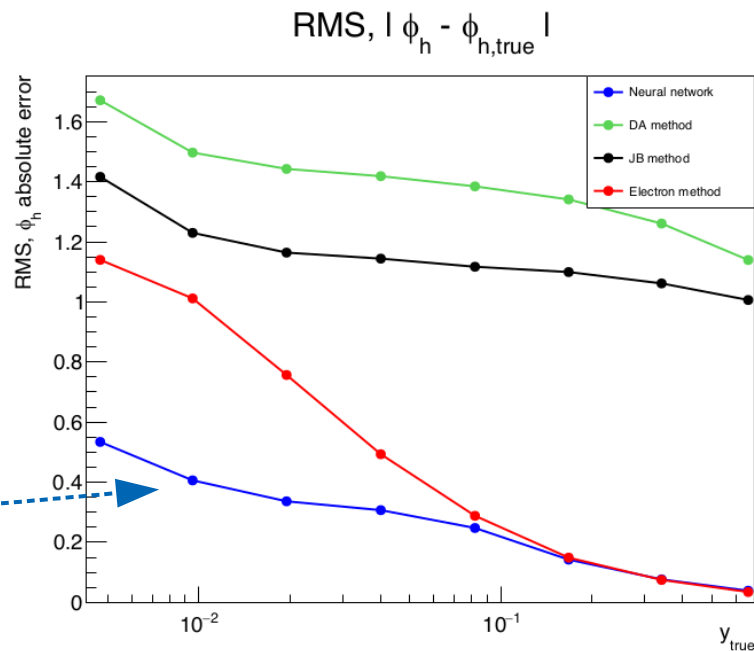
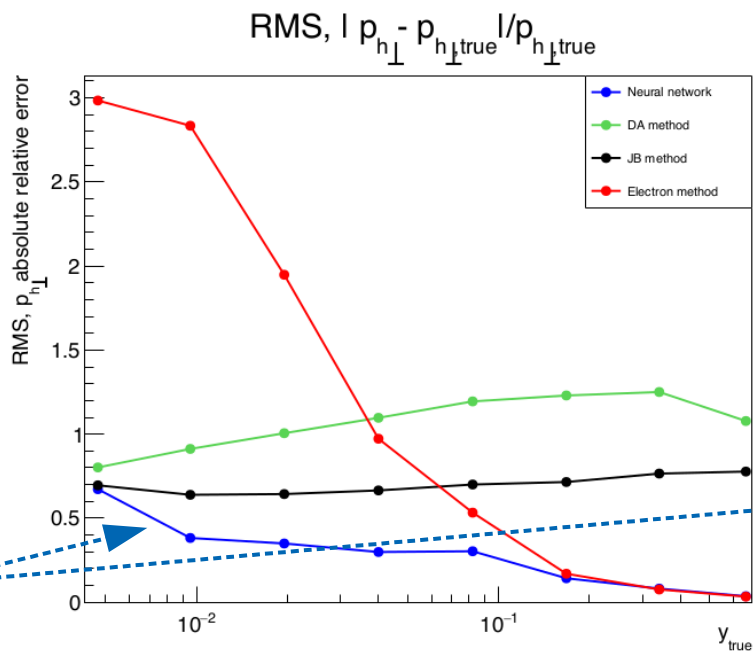


Prog.Part.Nucl.Phys. 69 (2013) 28-84, 1208.6087 [hep-ph]

Kinematics Reconstruction

- ◆ **Kinematics calculations performed in dedicated class(es)**
 - Used for both reconstructed and MC generated particles
 - Inputs: beams, scattered electron, hadronic final state, and observed particles (single hadrons for SIDIS, jets, etc.)
- ◆ **Calculations**
 - Inclusive variables (x , Q^2 , W , y , ...)
 - 6 methods: electron, J.B., double angle, mixed, sigma, eSigma
 - SIDIS variables (p , p_T , z , ϕ_h , ...)
 - Jet variables (z , p_T , j_\perp , ...)
 - In general uses Lorentz invariant calculations; boost to specific frames when needed
- ◆ **Future Plan**
 - Cross check with upstream calculations from the reconstruction framework and/or upstream our methods

Kinematics Reconstruction With Machine Learning



Particle Flow Network

AI for kinematics reconstruction shows promising results!

C. Pecar, 2nd Workshop on AI for the EIC (Oct. 2022)

See also M. Diefenthaler, et al., Eur.Phys.J.C 82 (2022) 11, 1064

- TTrees: SIDIS, Jets, etc.
- Adage
 - Arbitrary multi-dimensional binning for any object
 - Primary object: a set of histograms for coverage and resolution

Multidimensional Binning

- ◆ Problem: The need for multidimensional analysis caused deeply nested for loops to spread throughout epic-analysis
- Not maintainable and not generalized
 - Very susceptible to bugs

```
for (auto z_bin : z_bins) {
  for (auto y_bin : y_bins) {

    action_before_x_Q2_subloop( z_bin, y_bin );

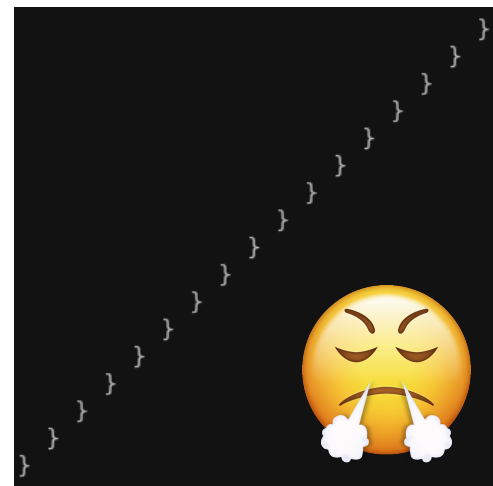
    for (auto Q2_bin : Q2_bins) {
      for (auto x_bin : x_bins) {

        action_for_each_bin( z_bin, y_bin, Q2_bin, x_bin );

      }
    }

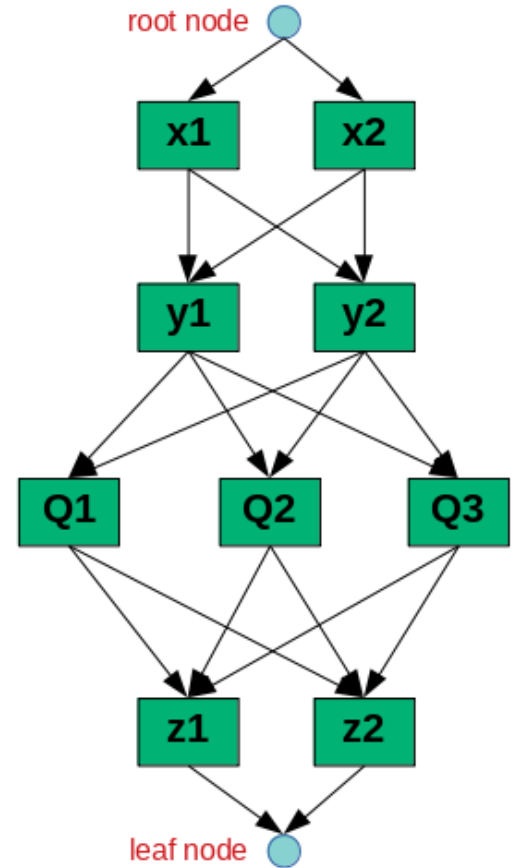
    action_after_x_Q2_subloop( z_bin, y_bin );

  }
}
```

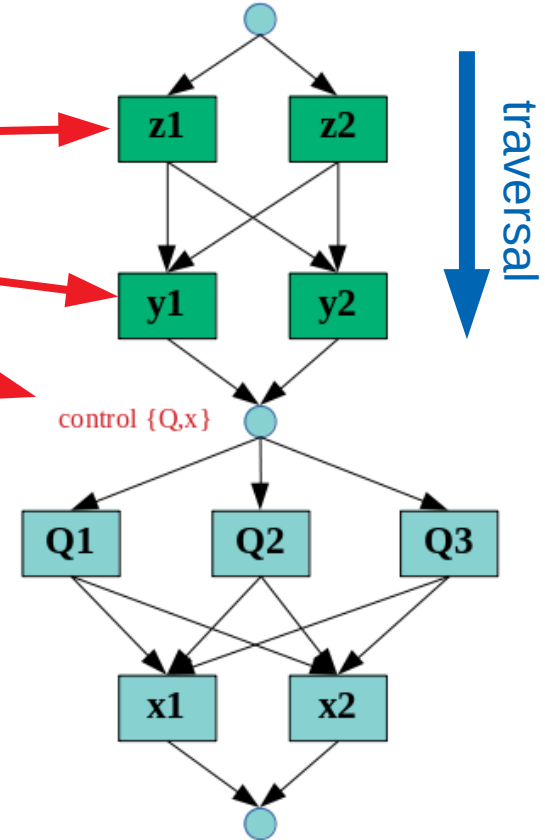


- ◆ Generalize multidimensional binning implementation with a Directed Acyclic Graph (DAG)
 - Fully connected layers of 1D bins
 - One path from root node to leaf node == 1 multidimensional bin
- ◆ Store 1st order functions as additional “control nodes”
 - Executable during depth-first traversal
 - For any multidimensional bin
 - For any lower-dimensional subset of bins
 - Control node stores two 1st ordered functions, executed when:
 - Traversing toward leaf
 - Backtracking toward root

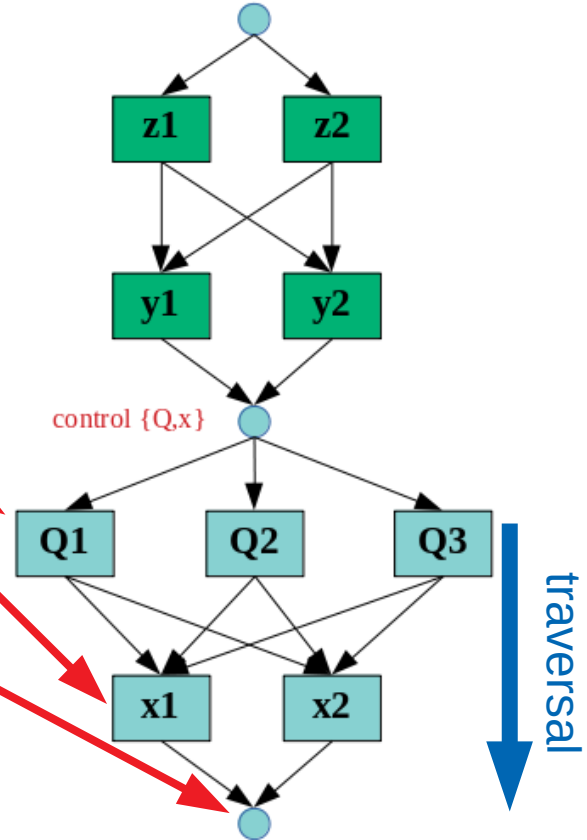
<https://github.com/c-dilks/adage>



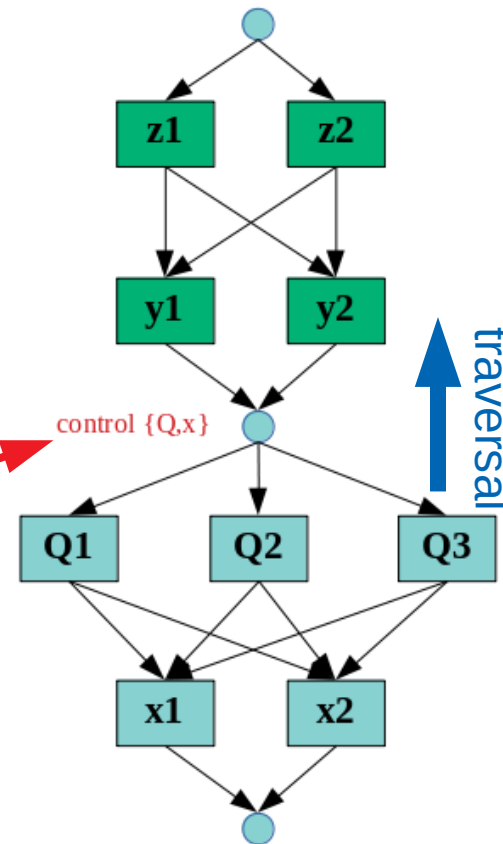
```
for (auto z_bin : z_bins) {  
  for (auto y_bin : y_bins) {  
  
    action_before_x_Q2_subloop( z_bin, y_bin );  
  
    for (auto Q2_bin : Q2_bins) {  
      for (auto x_bin : x_bins) {  
  
        action_for_each_bin( z_bin, y_bin, Q2_bin, x_bin );  
  
      }  
    }  
  
    action_after_x_Q2_subloop( z_bin, y_bin );  
  
  }  
}
```



```
for (auto z_bin : z_bins) {  
  for (auto y_bin : y_bins) {  
  
    action_before_x_Q2_subloop( z_bin, y_bin );  
  
    for (auto Q2_bin : Q2_bins) {  
      for (auto x_bin : x_bins) {  
  
        action_for_each_bin( z_bin, y_bin, Q2_bin, x_bin );  
  
      }  
    }  
  
    action_after_x_Q2_subloop( z_bin, y_bin );  
  
  }  
}
```



```
for (auto z_bin : z_bins) {  
  for (auto y_bin : y_bins) {  
  
    action_before_x_Q2_subloop( z_bin, y_bin );  
  
    for (auto Q2_bin : Q2_bins) {  
      for (auto x_bin : x_bins) {  
  
        action_for_each_bin( z_bin, y_bin, Q2_bin, x_bin );  
  
      }  
    }  
  
    action_after_x_Q2_subloop( z_bin, y_bin );  
  
  }  
}
```



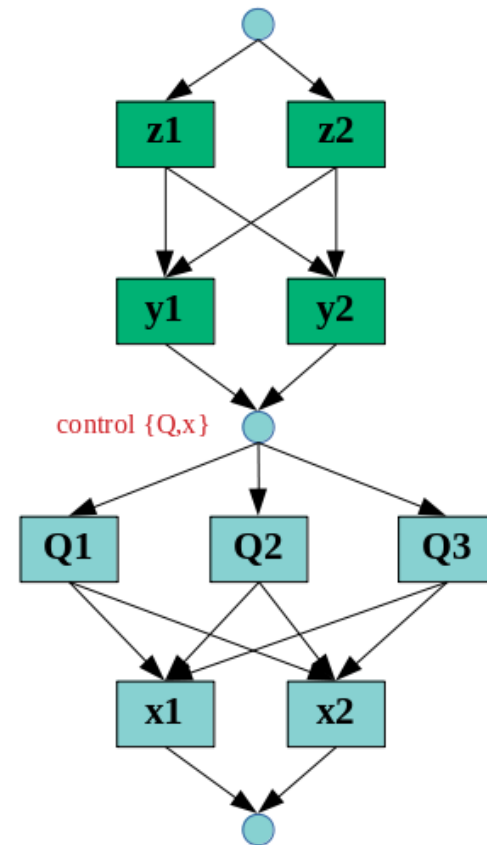
In Practice:

```
// define bins
...

// define lambdas
action_before_x_Q2_subloop = ... ;
action_after_x_Q2_subloop  = ... ;
action_for_each_bin        = ... ;

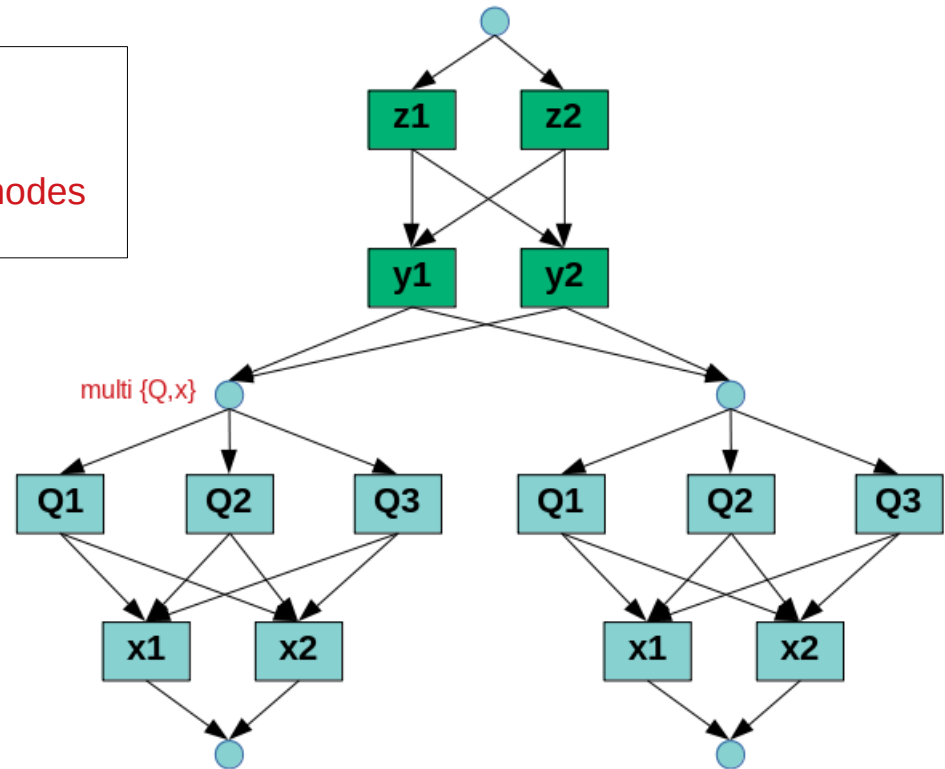
// attach lambdas to the DAG
Adage->BeforeSubloop( {"x","q2"}, action_before_x_Q2_subloop );
Adage->AfterSubloop(  {"x","q2"}, action_after_x_Q2_subloop  );
Adage->Payload( action_for_each_bin );

// run
Adage->Execute();
```



◆ Additional Support

- Conditional execution of subloops
- Repeated subloops, with different control nodes



- ◆ Runs for every “git commit” (on a pull request)
 - Could be triggered by upstream repositories

- ◆ Job matrices for:
 - Data sources [ePIC full simulation, Delphes fast simulation, previous designs]
 - ePIC runs include radcor and no-radcor versions
 - Reconstruction method [electron, DA, JB, ...]

- ◆ Build tests, Valgrind, etc.

- ◆ Production of several 1D and 2D plots, in multidimensional binnings
 - Coverage
 - Resolution

Continuous Integration

Semi-Inclusive Deep Inelastic Scattering (SIDIS) Cuts

$$W > 3 \text{ GeV}$$

$$0.01 < y < 0.95$$

$$0.2 < z < 0.9$$

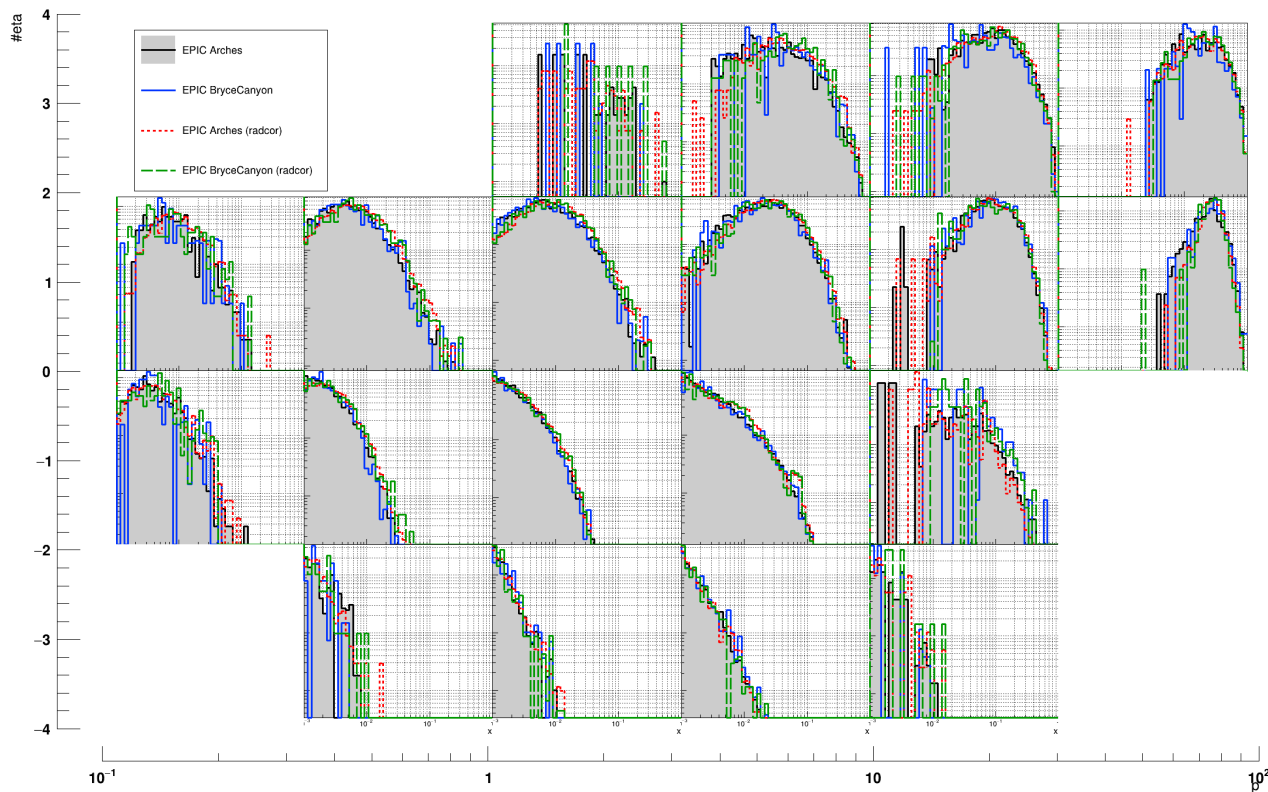
$$x_F > 0$$

$$p_T(\text{lab}) > 0.1 \text{ GeV}$$

**Focusing on beam energy of
18x275 GeV**

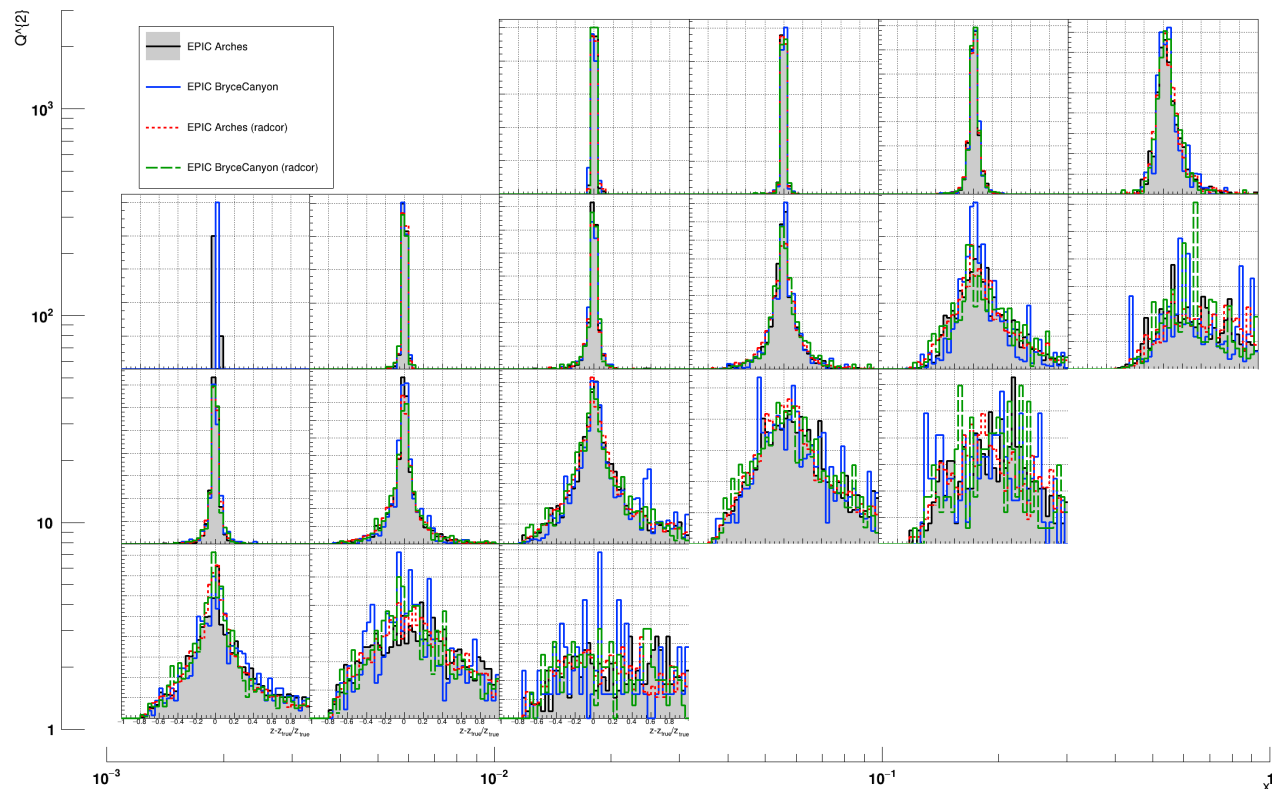
- ◆ Testing all available reconstruction methods
- ◆ Histograms in bins of
 - (x, Q^2)
 - (η, p)

Comparison of two different ePIC design options



- ◆ x distributions in bins of (x, Q^2)
- ◆ Comparisons:
 - With and without radiative corrections
 - ePIC designs “Arches” and “BryceCanyon”

Comparison of two different ePIC design options



- ◆ z resolutions in bins of (x, Q^2)
- ◆ Same comparisons
- ◆ Electron method, so resolution is not so good at low Q^2 and high x

Dihadrons

Jets

Synergy with DIS group

Cross-checking and upstreaming of algorithms to reconstruction framework

