Invisible Higgs with Full SiD Simulation/Reconstruction

Chris Potter

University of Oregon

Introduction

- Assuming ILC $\sqrt{s} = 250$ GeV, $e^+e^- \to ZH$ with hadronic Z decay and invisible H decay.
- Fast SiD (Delphes) Simulation Study Previously Presented
  - Samples are fast Delphes SiD simulation with DSiD
    - 10ab^{-1} signal generated with Whizard 2.6.4 by me, inclusive Z decays
    - 250fb^{-1} backgrounds are Whizard 1.95 by Tim Barklow for DBD
  - Polarization scheme is 80% e^-, 30% e^+ polarized beams.
  - Luminosity sharing is 250/250 fb^{-1} (LR/RL).

Introduction

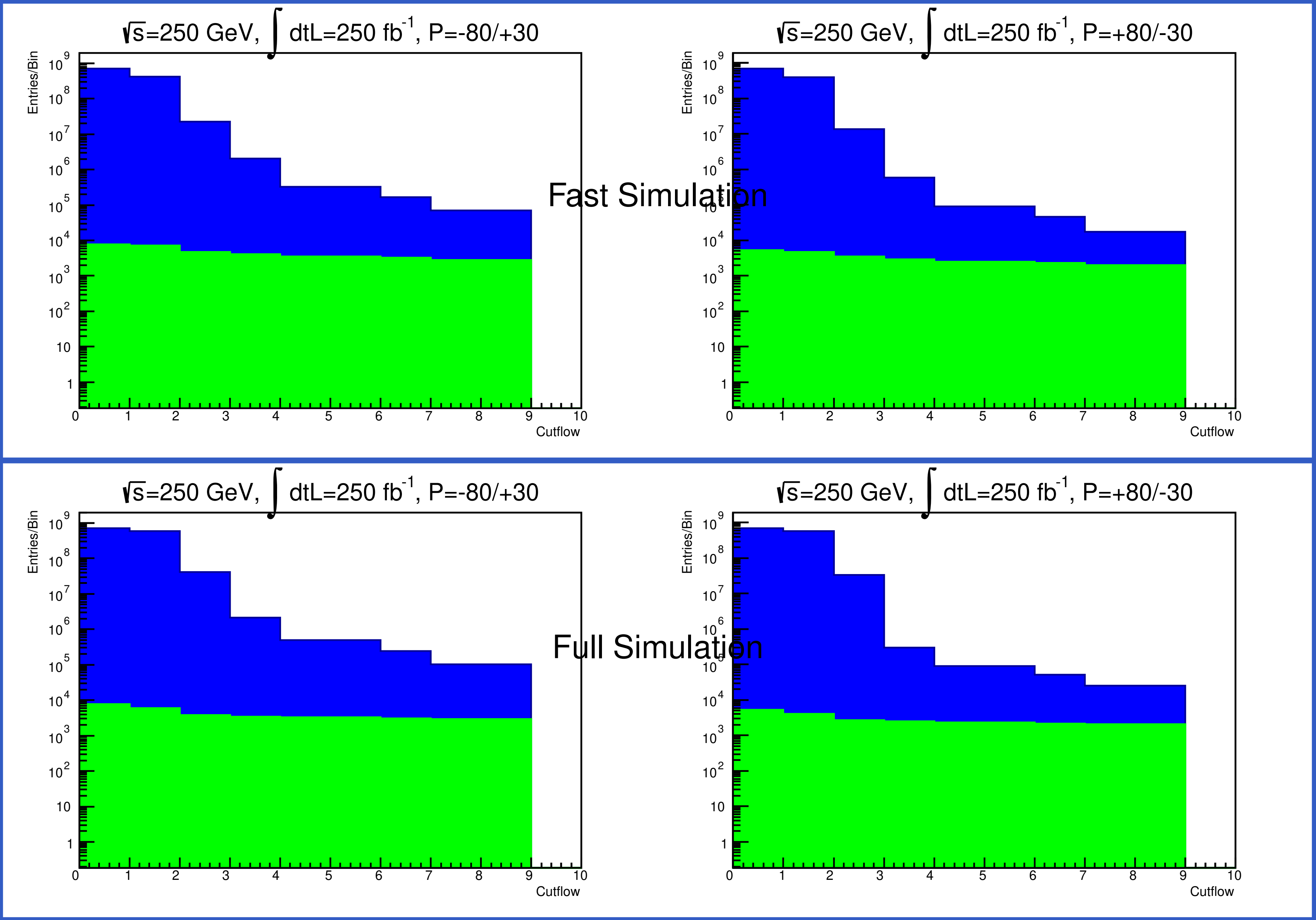
- Assuming ILC $\sqrt{s} = 250$ GeV, $e^+e^- \to ZH$ with hadronic Z decay and invisible H decay.
- Fast SiD (Delphes) Simulation Study Previously Presented
  - Samples are fast Delphes SiD simulation with DSiD
    - 10ab^{-1} signal generated with Whizard 2.6.4 by me, inclusive Z decays
    - 250fb^{-1} backgrounds are Whizard 1.95 by Tim Barklow for DBD
  - Polarization scheme is 80% e^-, 30% e^+ polarized beams.
  - Luminosity sharing is 250/250 fb^{-1} (LR/RL).
- Full SiD Simulation (ILCSoft/ddsim/Geant4) and Reconstruction Today
  - Same generator files used for Delphes study are used here.
    - 10ab^{-1} signal generated with Whizard 2.6.4 by me, inclusive Z decays
    - 250fb^{-1} backgrounds are Whizard 1.95 by Tim Barklow for DBD
  - ILCSoft v02-00-02, SiD o2_v3, Marlin with truth tracks and PandoraPFA.
  - Same polarization scheme and luminosity sharing as for the Delphes study.

Introduction

- Assuming ILC $\sqrt{s} = 250$ GeV, $e^+e^- \to ZH$ with hadronic Z decay and invisible H decay.
- Fast SiD (Delphes) Simulation Study Previously Presented
  - Samples are fast Delphes SiD simulation with DSiD
    - 10ab^{-1} signal generated with Whizard 2.6.4 by me, inclusive Z decays
    - 250fb^{-1} backgrounds are Whizard 1.95 by Tim Barklow for DBD
  - Polarization scheme is 80% e^-, 30% e^+ polarized beams.
  - Luminosity sharing is 250/250 fb^{-1} (LR/RL).
- Full SiD Simulation (ILCSoft/ddsim/Geant4) and Reconstruction Today
  - Same generator files used for Delphes study are used here.
    - 10ab^{-1} signal generated with Whizard 2.6.4 by me, inclusive Z decays
    - 250fb^{-1} backgrounds are Whizard 1.95 by Tim Barklow for DBD
  - ILCSoft v02-00-02, SiD o2_v3, Marlin with truth tracks and PandoraPFA.
  - Same polarization scheme and luminosity sharing as for the Delphes study.
- Problematic Issues with Reconstruction (can be addressed)
  - Unrealistic precision: hits fitted but MC truth is used to assign hits to tracks
  - Low efficiency: efficiency for trackfinding is somewhat lower then expected.
  - Particle ID unoptimized: PandoraPFA track/cluster matching unoptimized.

Cutflow: Fast vs. Full

Signal is green with BR($H \to$ inv.)=0.10, all background is blue.



Yields and Significance: Fast vs. Full

Fast Simulation

Requirement (Fast)	S(LR)	B(LR)	$\frac{S}{\sqrt{S+B}}$	S(RL)	B(RL)	$\frac{S}{\sqrt{S+B}}$
No Cut	2.79e+04	2.55e+09	0.552	1.89e+04	2.5e+09	0.378
$N_e + N_\mu = 0$	2.53e+04	1.48e+09	0.659	1.71e+04	1.42e+09	0.453
$\frac{N_{trk} \in [6,24]}{N_{pfo} \in [12,40]}$	1.73e+04	7.97e+07	1.94	1.23e+04	4.95e+07	1.75
$20 \leq p_T^{vis} \leq 60$ GeV	1.47e+04	7.24e+06	5.45	1.04e+04	2.11e+06	7.16
$75 \leq m_{vis} \leq 105$ GeV	1.25e+04	1.15e+06	11.6	8.89e+03	3.13e+05	15.7
$N_{jet} = 2$	1.25e+04	1.15e+06	11.6	8.89e+03	3.13e+05	15.7
$-0.9 \leq \cos\theta_{jj} \leq -0.2$	1.16e+04	5.91e+05	15	8.25e+03	1.55e+05	20.4
$110 \leq m_{recoil} \leq 140$	1.01e+04	2.39e+05	20.2	7.17e+03	5.5e+04	28.7

Yields and Significance: Fast vs. Full

Fast Simulation

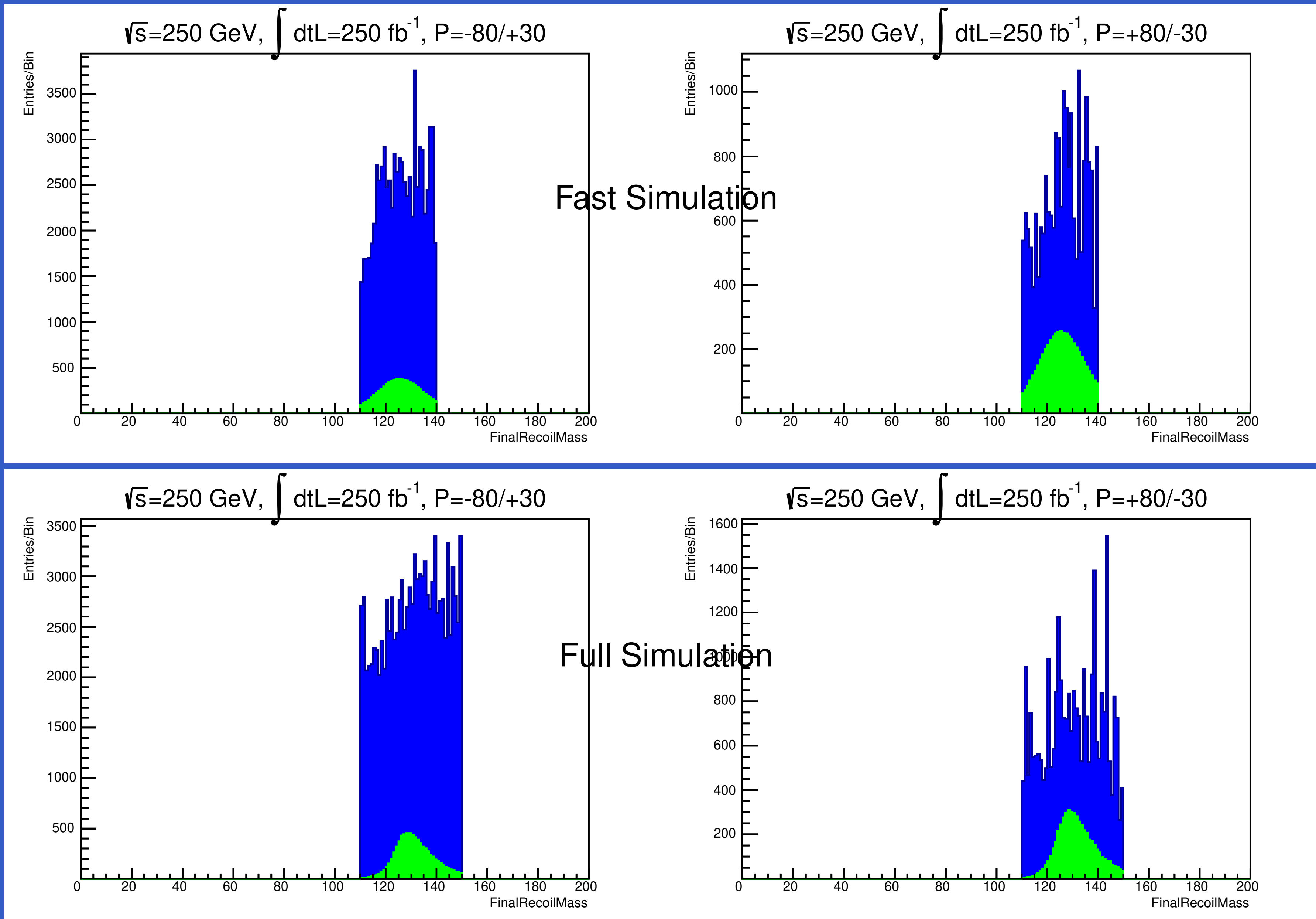
Requirement (Fast)	S(LR)	B(LR)	$\frac{S}{\sqrt{S+B}}$	S(RL)	B(RL)	$\frac{S}{\sqrt{S+B}}$
No Cut	2.79e+04	2.55e+09	0.552	1.89e+04	2.5e+09	0.378
$N_e + N_\mu = 0$	2.53e+04	1.48e+09	0.659	1.71e+04	1.42e+09	0.453
$\frac{N_{trk} \in [6,24]}{N_{pfo} \in [12,40]}$	1.73e+04	7.97e+07	1.94	1.23e+04	4.95e+07	1.75
$20 \leq p_T^{vis} \leq 60$ GeV	1.47e+04	7.24e+06	5.45	1.04e+04	2.11e+06	7.16
$75 \leq m_{vis} \leq 105$ GeV	1.25e+04	1.15e+06	11.6	8.89e+03	3.13e+05	15.7
$N_{jet} = 2$	1.25e+04	1.15e+06	11.6	8.89e+03	3.13e+05	15.7
$-0.9 \leq \cos\theta_{jj} \leq -0.2$	1.16e+04	5.91e+05	15	8.25e+03	1.55e+05	20.4
$110 \leq m_{recoil} \leq 140$	1.01e+04	2.39e+05	20.2	7.17e+03	5.5e+04	28.7

Full Simulation

Requirement (Full)	S(LR)	B(LR)	$\frac{S}{\sqrt{S+B}}$	S(RL)	B(RL)	$\frac{S}{\sqrt{S+B}}$
$20 \leq p_T^{vis} \leq 70$ GeV	1.25e+04	7.71e+06	4.48	8.84e+03	1.07e+06	8.53
$75 \leq m_{vis} \leq 105$ GeV	1.16e+04	1.79e+06	8.63	8.21e+03	3.14e+05	14.5
$N_{jet} = 2$	1.16e+04	1.79e+06	8.63	8.21e+03	3.14e+05	14.5
$-0.9 \leq \cos\theta_{jj} \leq -0.2$	1.08e+04	8.68e+05	11.5	7.65e+03	1.78e+05	17.7
$110 \leq m_{recoil} \leq 150$	1.03e+04	3.6e+05	17	7.33e+03	8.39e+04	24.2

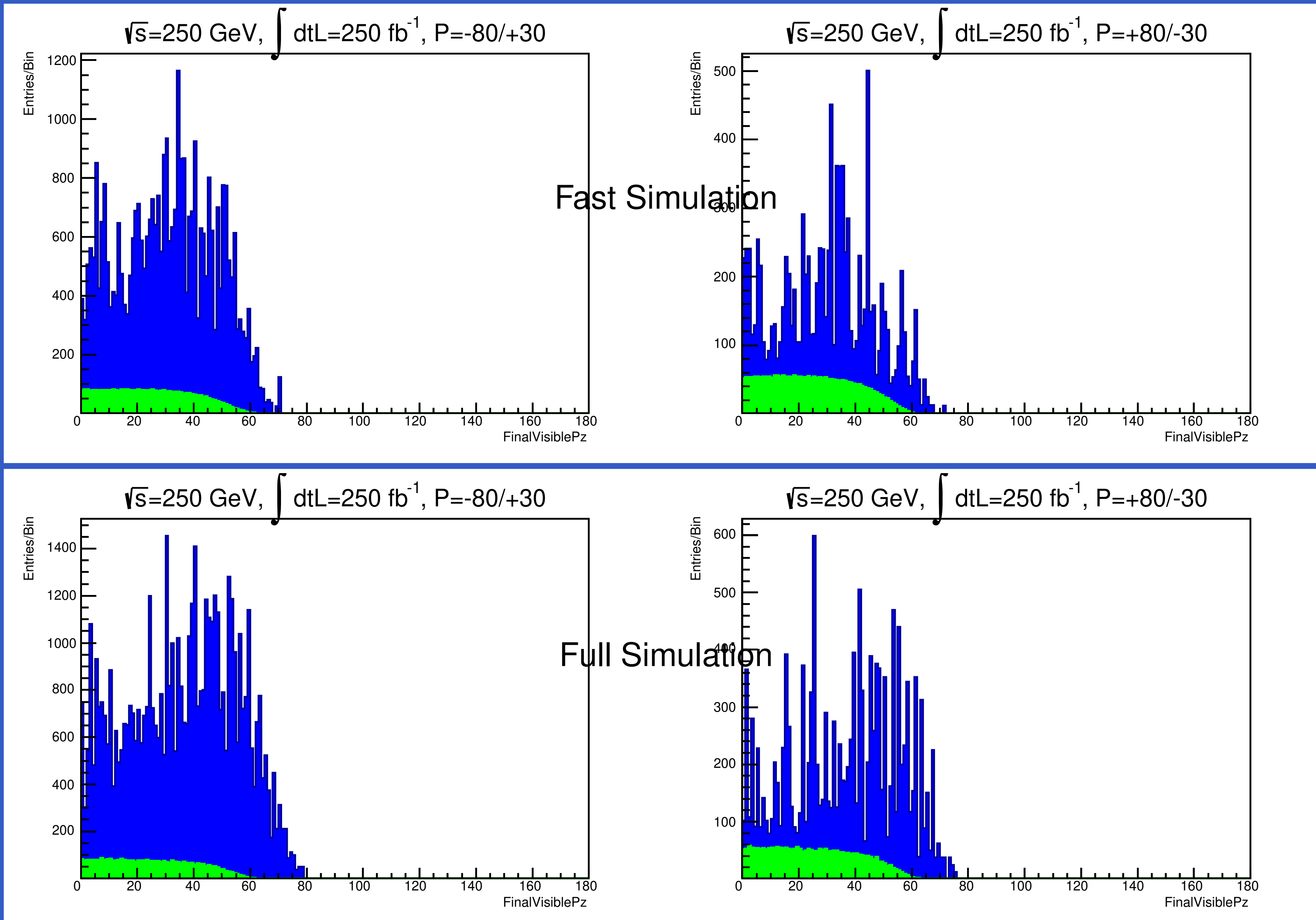
Recoil Mass: Fast vs. Full (After All Cuts)

Signal is green with BR($H \to$ inv.)=0.10, all background is blue.



Visible Longitudinal Momentum p_z^{vis} (AAC)

Signal is green with BR($H \rightarrow$ inv.)=0.10, all background is blue.



Recall: Event Shape Variables

a simple mean value will be a bad estimator. Mostly, one uses either a truncated mean (like eliminating the 20% of highest individual measurements), or resorts to a full maximum likelihood treatment.

Ample discussion on optimizing detectors and readout for ionization sampling, and a vast amount of literature is also found in review papers, e.g. [Lehraus83], [Allison91]. [Blum93] discusses in detail the collection and analysis of ionization samples in drift chambers.

Jet Chamber. A drift chamber made of multiple cells of moderate size, named thus because of the optimal two-track resolution as needed in jets; → Drift Chamber.

Jet Variables. In many collisions, observable secondary particles are produced in highly collimated form, called particle *jets*. This is a consequence of the *hadronization* of partons (quarks or gluons) produced in hard collisions. Jets for a given initial parton can vary widely in shape, particle content, and energy spectrum; there is, of course, also substantial blurring due to instrumental effects: the finite resolution and granularity of detectors (calorimeter cells and muon measurements), and escaping neutrinos.

The earliest evidence for jets was in e^+e^- collisions (SLAC and DESY), producing secondary hadrons; subsequently, they were also observed in hadronic collisions (e.g. UA experiments and ISR at CERN). Frequently, two main jets are observed which dominate the energy balance of the collision; in hadronic collider events, the balance is observed only laterally, due to the difficulty of observing at large (absolute) rapidity, and due to the structure function (→), which leaves the hard quark encounter with a longitudinal boost. Often, the main jets are accompanied by one or more broader jet(s), interpreted as radiated gluons. The following scalar jet variables were used in the early jet studies, describing mostly a two- or three-jet situation from e^+e^- events.

a) Sphericity $\equiv (3/2)\min\left(\sum \boldsymbol{p}_\mathrm{T}^2/\sum \boldsymbol{p}^2\right)$
p_T is the transverse momentum perpendicular to a unit vector $\boldsymbol{n}$, the sums are over all particles of the reaction, and the minimum is formed with respect to $\boldsymbol{n}$.

b) Thrust $\equiv 2\max\left(\sum_1 \boldsymbol{p}_\mathrm{L}/\sum|\boldsymbol{p}|\right)$
p_L is the longitudinal momentum along a unit vector $\boldsymbol{n}$. Summation is over all particles for $|p|$, over those with $\boldsymbol{p}\cdot\boldsymbol{n} > 0$ for p_L; the maximum is formed with respect to $\boldsymbol{n}$. For the hadronic system in deep inelastic scattering, a *current thrust* has been proposed (→ [Webber95]) which is the thrust evaluated in the *Breit frame*, defined to be the frame in which the momentum transfer q is spacelike and along the longitudinal ($z-$) axis (→ Deep Inelastic Scattering Variables).

c) Spherocity $\equiv (4/\pi\min(\sum|p_\mathrm{T}|/\sum|p|)^2$
p_T is the transverse momentum perpendicular to a unit vector $\boldsymbol{n}$, the sums are over all particles, and the minimization is with respect to $\boldsymbol{n}$

d) Triplicity $\equiv \max(\boldsymbol{n}_1\sum_1 \boldsymbol{p} + \boldsymbol{n}_2\sum_2 \boldsymbol{p} + \boldsymbol{n}_3\sum_3 \boldsymbol{p})/\sum|\boldsymbol{p}|$
Here, a general classification into three classes must take place; in each class i, $\boldsymbol{n}_i$ describes the axis and particles are associated to the class for which $\boldsymbol{p}\cdot\boldsymbol{n}$ is largest. The maximum must be found over all possible $\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3$.

e) Planarity $\equiv \max(\sum p_{\mathrm{T}1}^2 - \sum p_{\mathrm{T}2}^2)/\sum p_\mathrm{T}^2$
Here, $p_{\mathrm{T}1}$ and $p_{\mathrm{T}2}$ are defined to be axes of a Cartesian coordinate system whose third axis is p_L. The variable indicates how well a reaction satisfies the assumption of being in a plane. The maximum is found with respect to the plane orientation. Maximization gives the same result if $\sum p_{\mathrm{T}1}^2$ is maximized. The solution is therefore given by the principal axes (obtained in "principal component analysis"), and the planarity is the complement of a two-dimensional equivalent to sphericity (called *circularity*). If p_L is along x, then the direction for $p_{\mathrm{T}1}$ is obtained by rotating in the y-z plane through an angle α given by

$$\tan(2\alpha) = \left(2\sum(p_y p_z)\right) / \left(\sum p_y^2 - \sum p_z^2\right)$$
$$p_{\mathrm{T}1} = p_y \cos\alpha + p_z \sin\alpha$$
$$p_{\mathrm{T}2} = p_z \cos\alpha - p_y \sin\alpha\,.$$

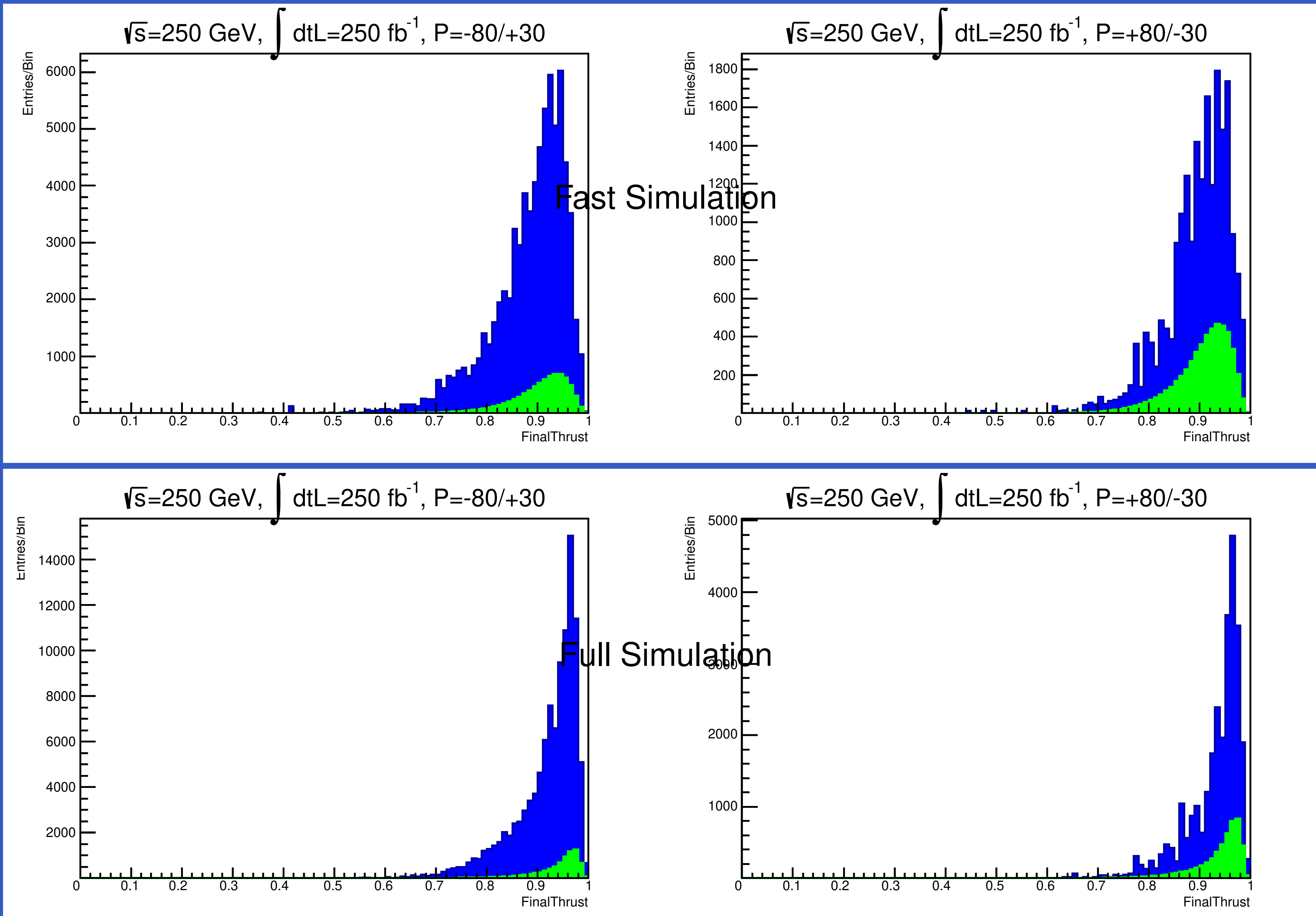
As maxima and minima differ by 90° in α, solving the above equation for an α in the range, say, from 0 to $\pi/2$, will still require deciding whether a maximum or minimum has been found. Obviously, $p_{\mathrm{T}1}^2 + p_{\mathrm{T}2}^2 = p_y^2 + p_z^2 = p_\mathrm{T}^2$.

The variables a) to c) all describe two-jet situations, with the jets back to back ($p_\mathrm{L} = 0$). d) has been used for a general three-jet situation. A detailed discussion of these variables can be found in [Brandt79]. The planarity e) describes a four-jet situation, if pairs of jets are correlated, e.g. a hard-scattered and a spectator system. The scattered system may have p_L different from zero, but will be in a plane with

Particle Detector Briefbook, Bock and Vasilescu, Springer-Verlag (1998)

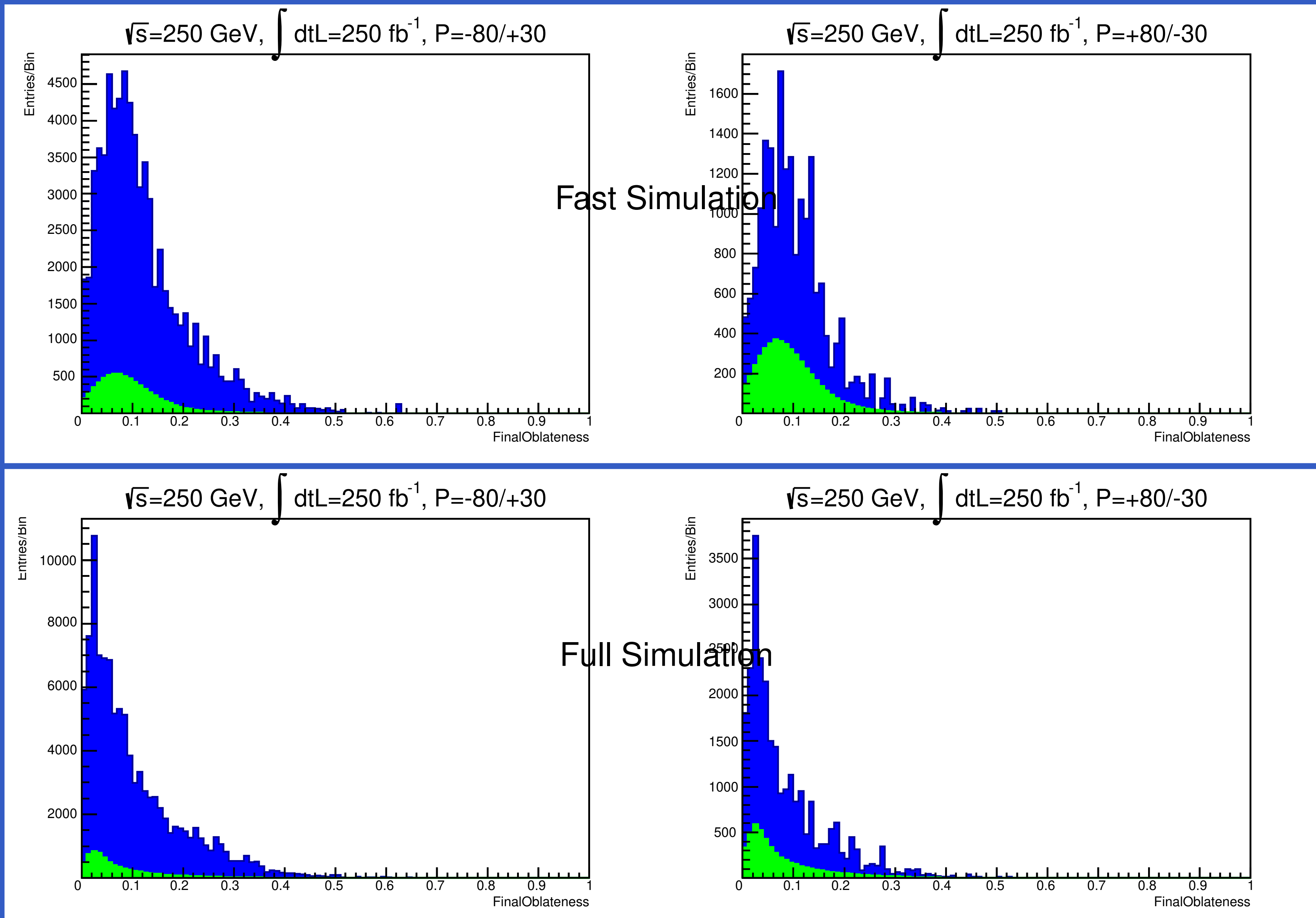
Boost to Z Frame, Thrust: Fast vs. Full (AAC)

Signal is green with BR($H \to$ inv.)=0.10, all background is blue.



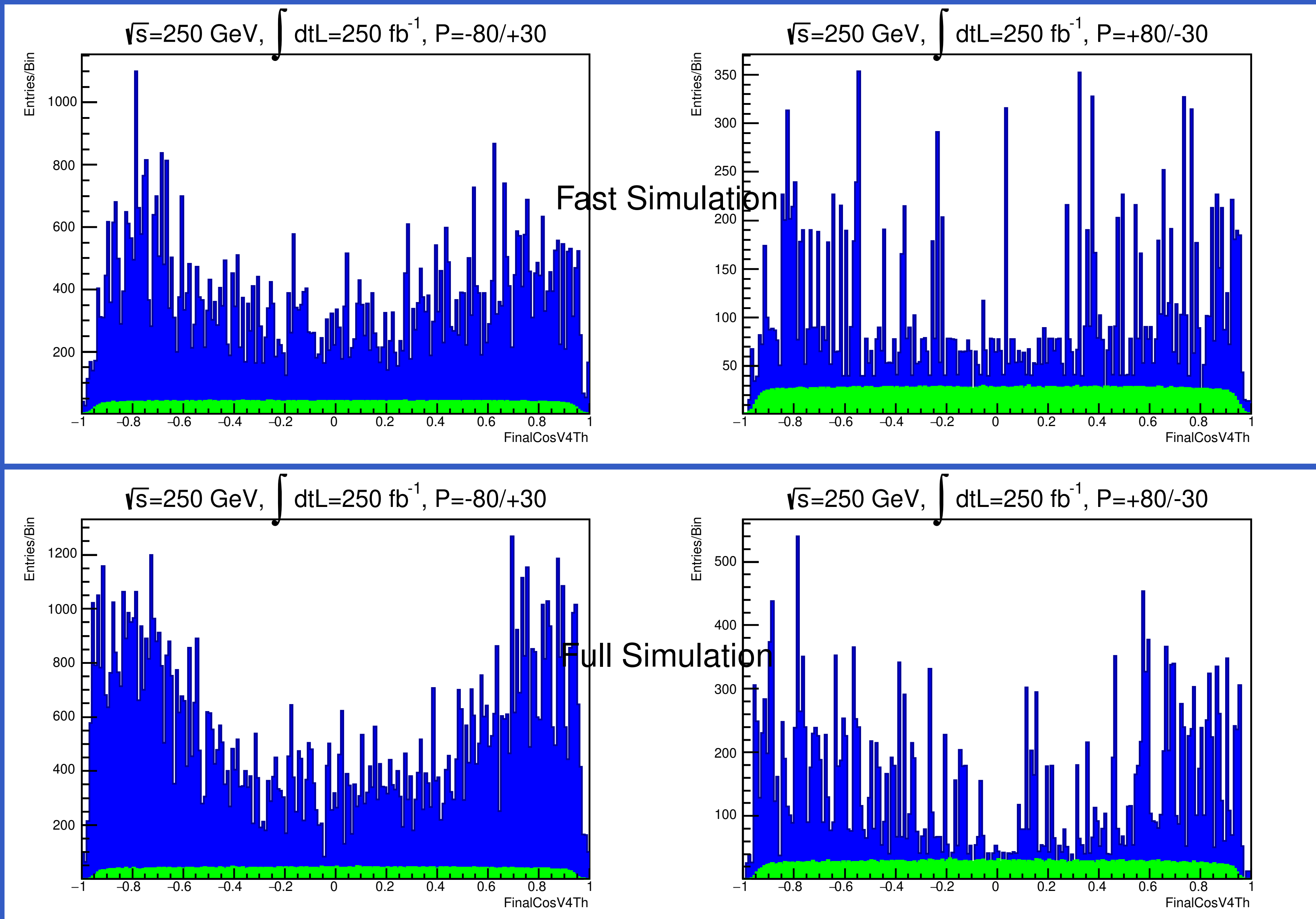
Boost to Z Frame, Oblateness: Fast vs. Full (AAC)

Signal is green with BR($H \to$ inv.)=0.10, all background is blue.



Boost to Z Frame, $\cos(\vec{p}_Z, \hat{n}_{thr})$: Fast vs. Full (AAC)

Signal is green with BR($H \to$ inv.)=0.10, all background is blue.



Conclusions

- Fast/full SiD simulation appear to be cross-validated, yielding reasonably similar results.
- The background composition from this full simulation study is broadly the same as from the fast study.
- Cuts values can be optimized for signal significance with Root TMVA. A cut on Z candidate p_z^{vis} may help to reduce 3 fermion backgrounds like $e\gamma \rightarrow eZ, \nu W$.

Conclusions

- Fast/full SiD simulation appear to be cross-validated, yielding reasonably similar results.
- The background composition from this full simulation study is broadly the same as from the fast study.
- Cuts values can be optimized for signal significance with Root TMVA. A cut on Z candidate p_z^{vis} may help to reduce 3 fermion backgrounds like $e\gamma \to eZ, \nu W$.
- After cut optimization, an MVA (simultaneous BDT, NN, etc) of eventshape variables in the Z candidate frame can reduce non-Z backgrounds like WW and bad-Z backgrounds like ZZ.
- Eventshape variables to consider:
  - Thrust measures how well confined to a single axis an event is.
  - Sphericity measures how isotropic an event is.
  - Planarity, oblateness measure how well confined to a plane an event is.
  - Triplicity measure how well confined to three axes an event is.
  - Fox-Wolfram moment measure similar properties.
  - These are all anti/correlated, but the nature of correlation is different in signal/background. MVA exploit this.

Conclusions

- Fast/full SiD simulation appear to be cross-validated, yielding reasonably similar results.
- The background composition from this full simulation study is broadly the same as from the fast study.
- Cuts values can be optimized for signal significance with Root TMVA. A cut on Z candidate p_z^{vis} may help to reduce 3 fermion backgrounds like $e\gamma \to eZ, \nu W$.
- After cut optimization, an MVA (simultaneous BDT, NN, etc) of eventshape variables in the Z candidate frame can reduce non-Z backgrounds like WW and bad-Z backgrounds like ZZ.
- Eventshape variables to consider:
  - Thrust measures how well confined to a single axis an event is.
  - Sphericity measures how isotropic an event is.
  - Planarity, oblateness measure how well confined to a plane an event is.
  - Triplicity measure how well confined to three axes an event is.
  - Fox-Wolfram moment measure similar properties.
  - These are all anti/correlated, but the nature of correlation is different in signal/background. MVA exploit this.
- Additional DBD generator files available for testing/training MVA and modeling background.
- A quick look at the leptonic channel yields signal significances consistent with Tokyo/ILD.