Study D^0 reconstruction with D^0 and DIS samples

Rongrong Ma

02/04/2025

Simulation samples

ep @ 10x100	D0 sample	DIS sample	Scale factor for DIS sample
$Q^2 > 1$	0.985 M (LOCATION)	4.98 M (LOCATION)	~350
$Q^2 > 100$	0.985 M (LOCATION)	4.97 M (LOCATION)	~99

- Scale factors determined based on simulation statistics
  - $Q^2 > 1$ - D0:DIS ~ 1:1770
  - $Q^2 > 100$ - D0:DIS ~ 1:500
- D^0 signals in DIS sample are removed. A very small effect

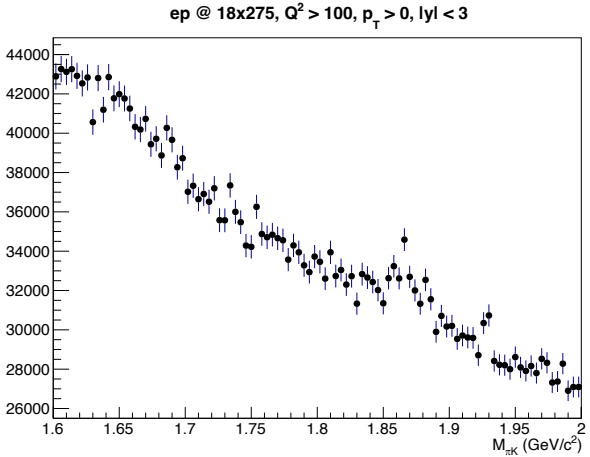
D^0 reconstruction

- Truth PID

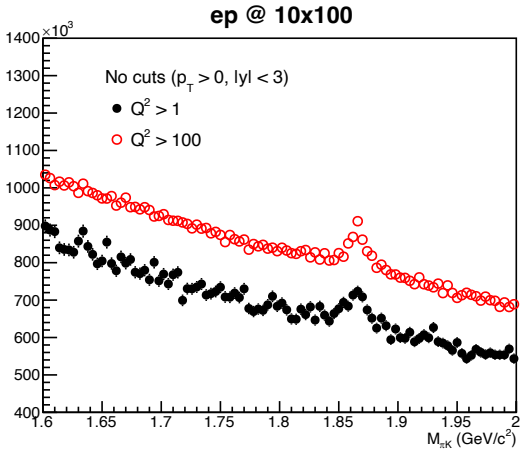
- Topological cuts
  - $DCA_{\pi} > 20\ \mu m$, $DCA_{K} > 20\ \mu m$
  - $DCA_{12} < 70\ \mu m$
  - $DCA_{D0} < 100\ \mu m$
  - Decay length > 50 μm
  - $\cos\theta > 0.95$

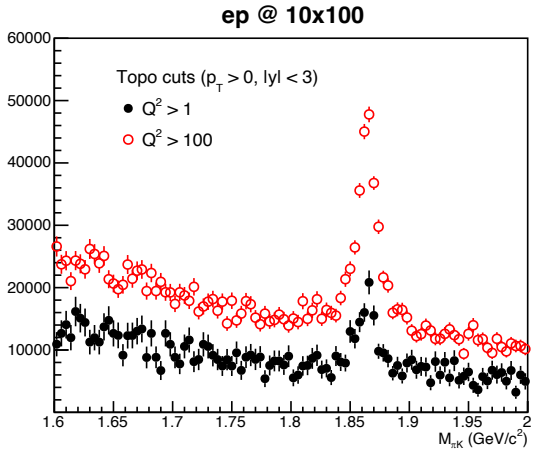
Inclusive D0+DIS sample

Before

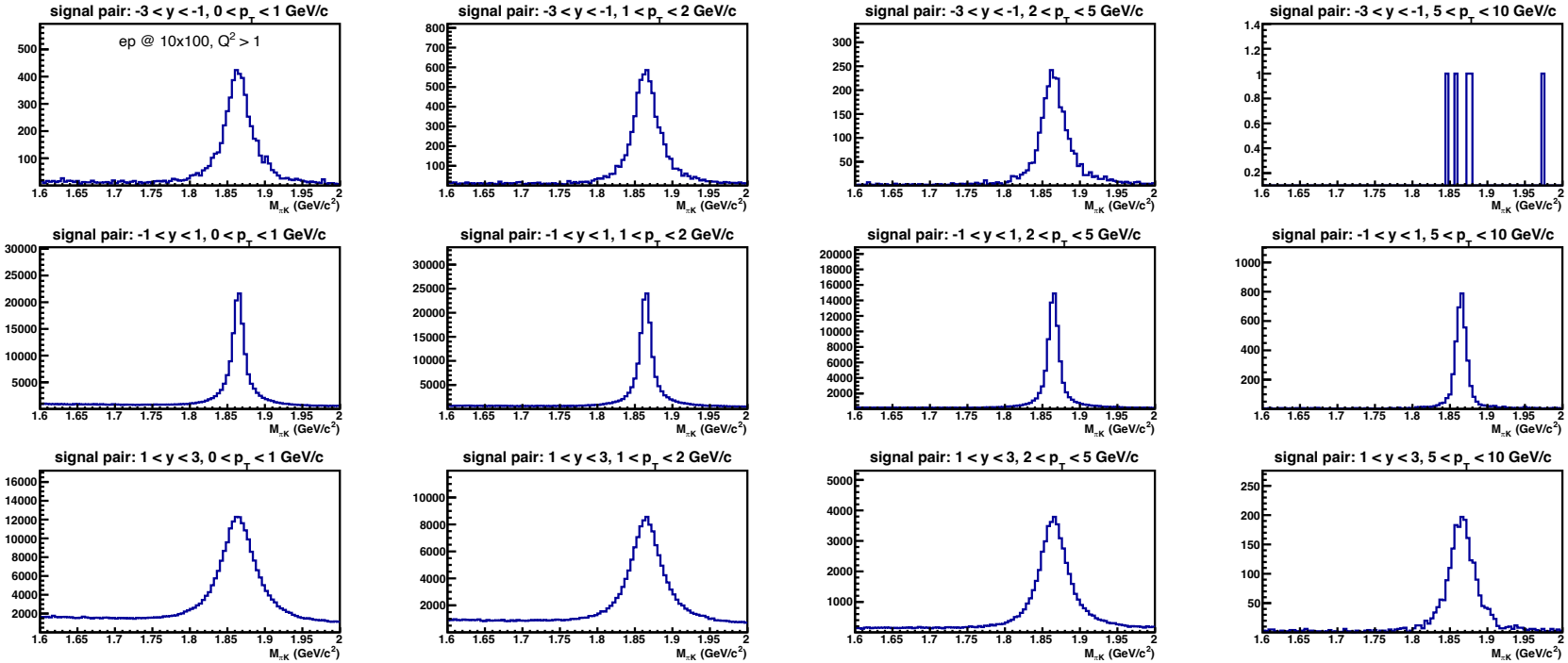


Now

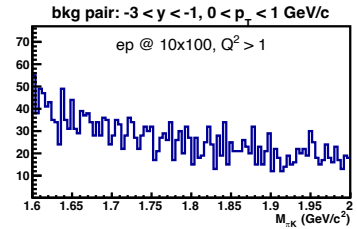


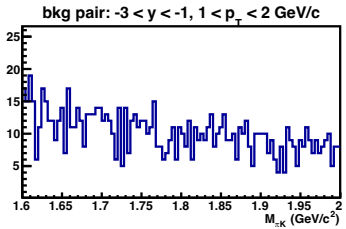


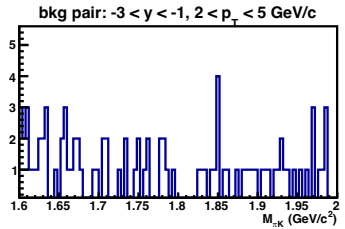
$Q^2 > 1$: D0 sample

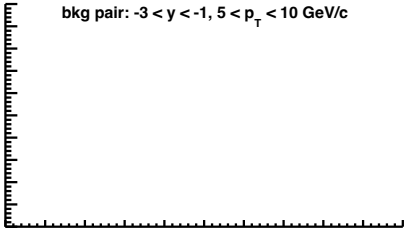


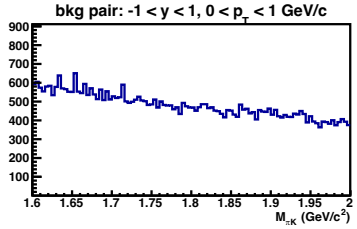
$Q^2 > 1$: DIS sample

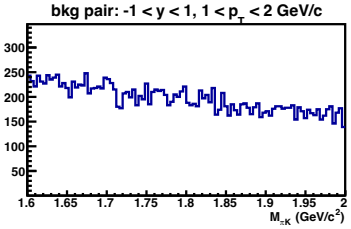


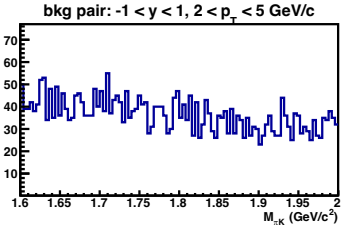


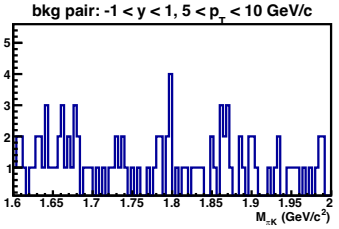


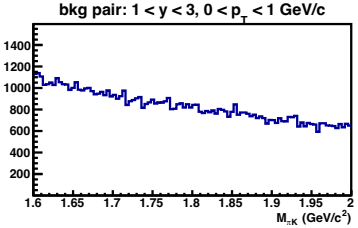


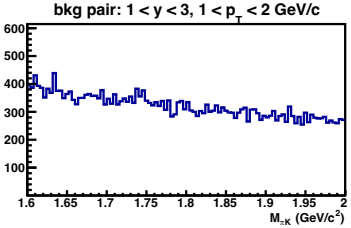


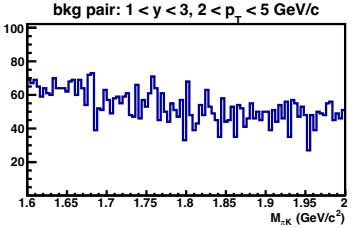


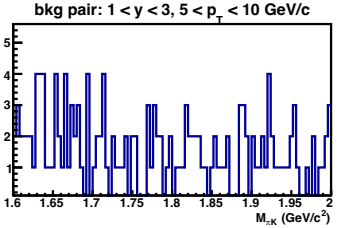




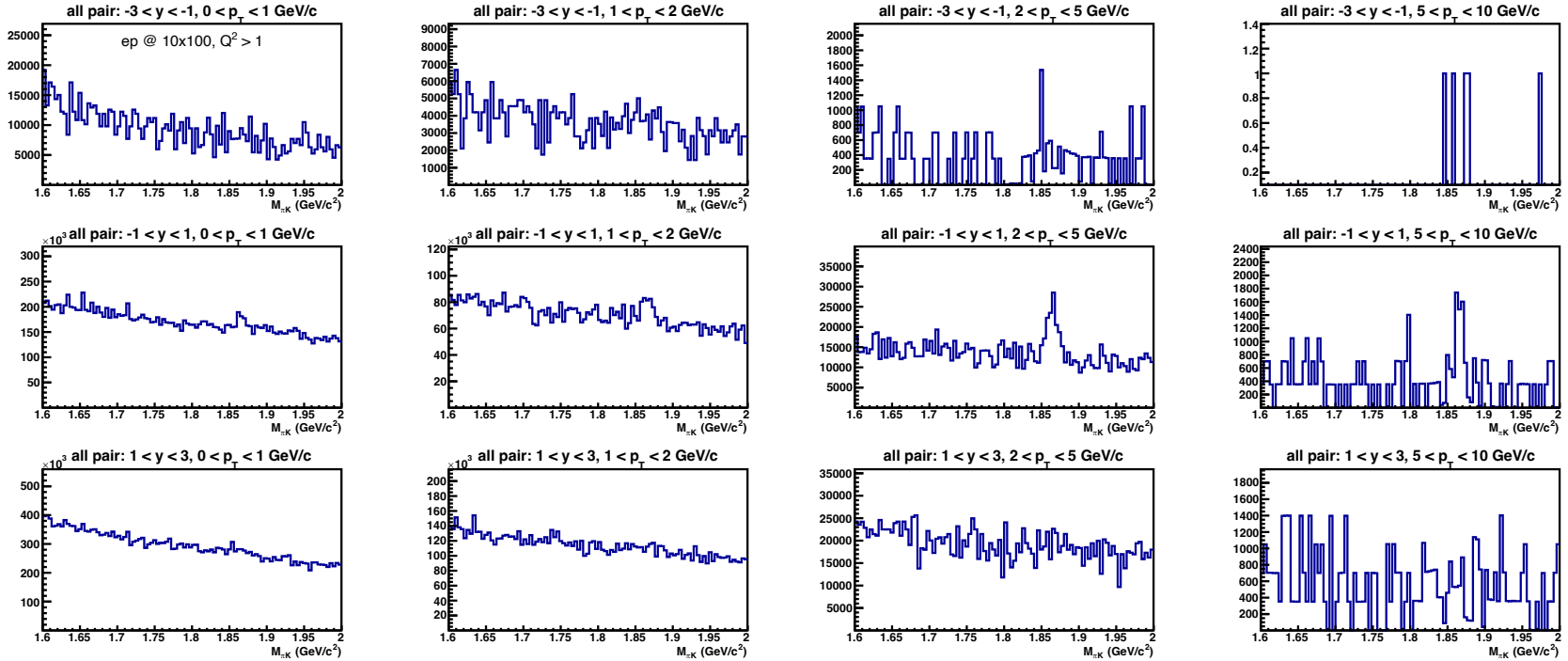






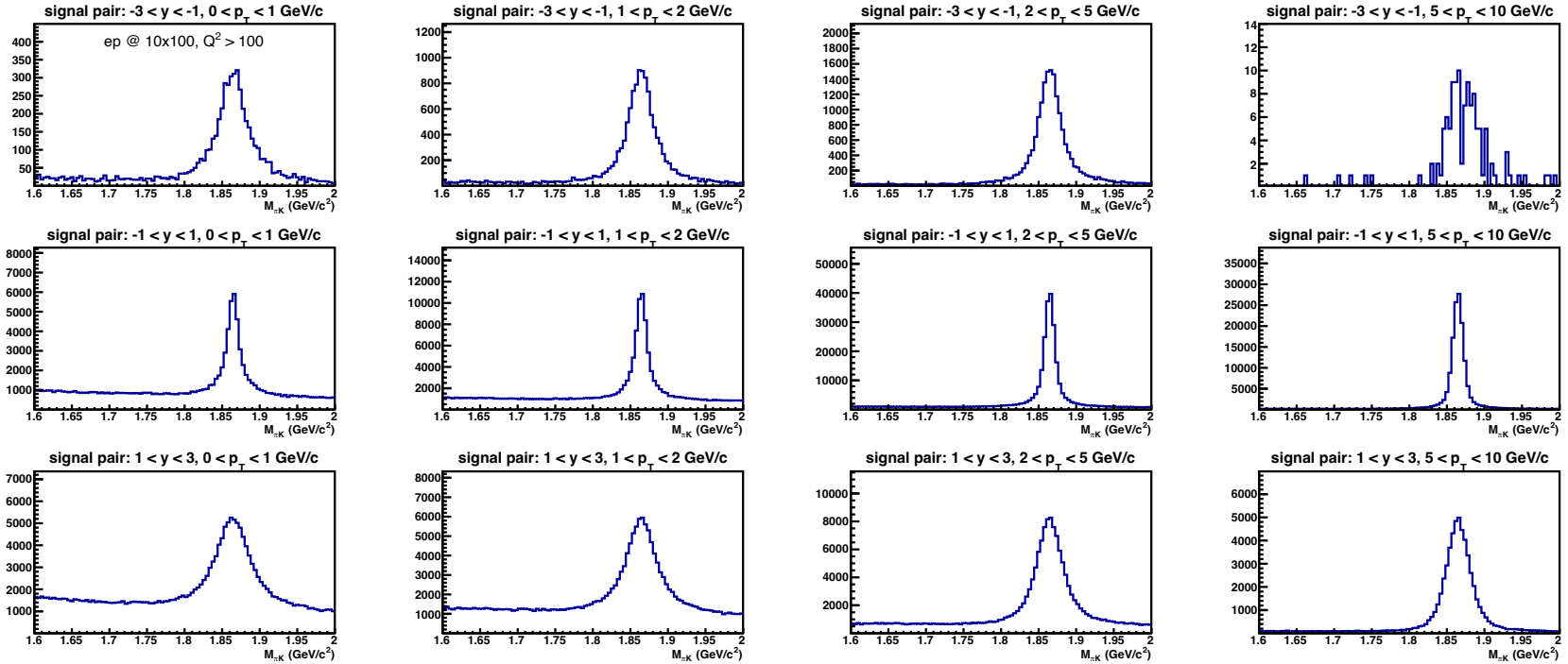


$Q^2 > 1$: D0 + scaled DIS

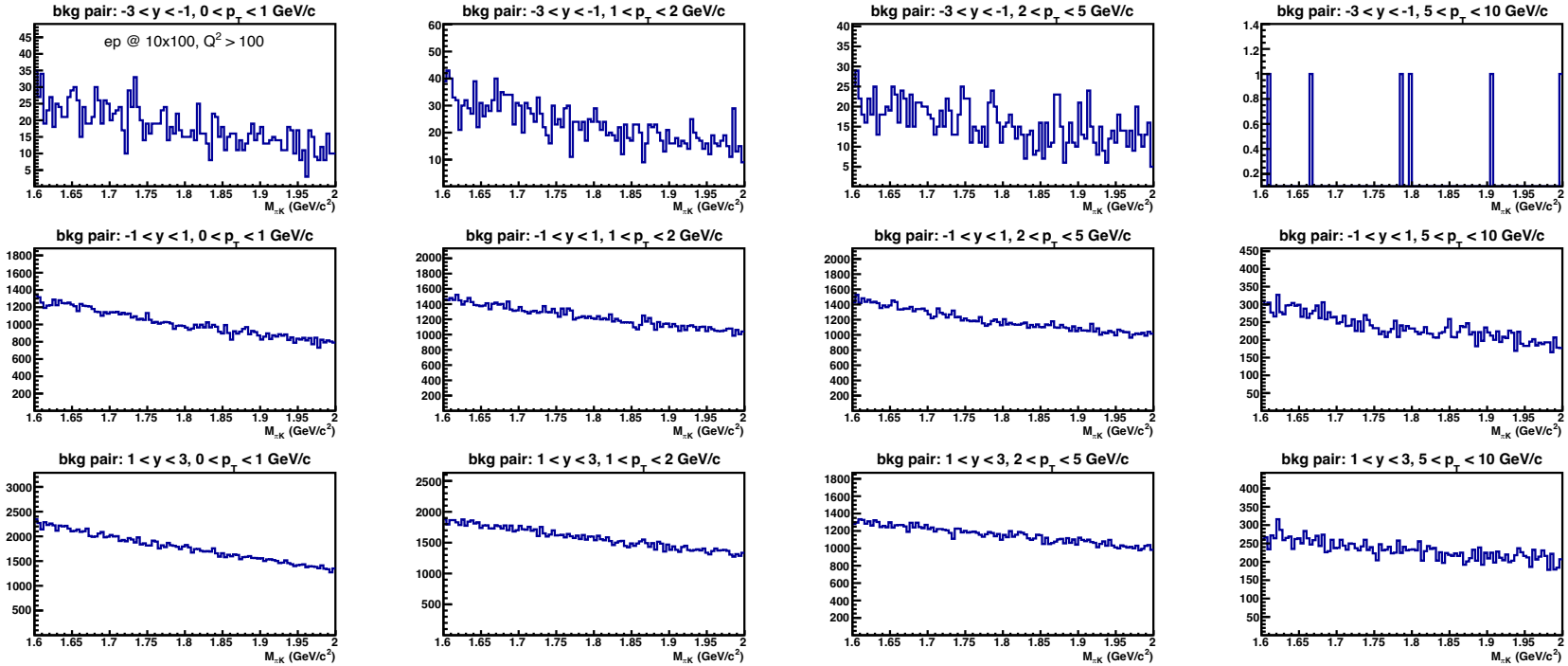


- Fluctuations in background overwhelm signal

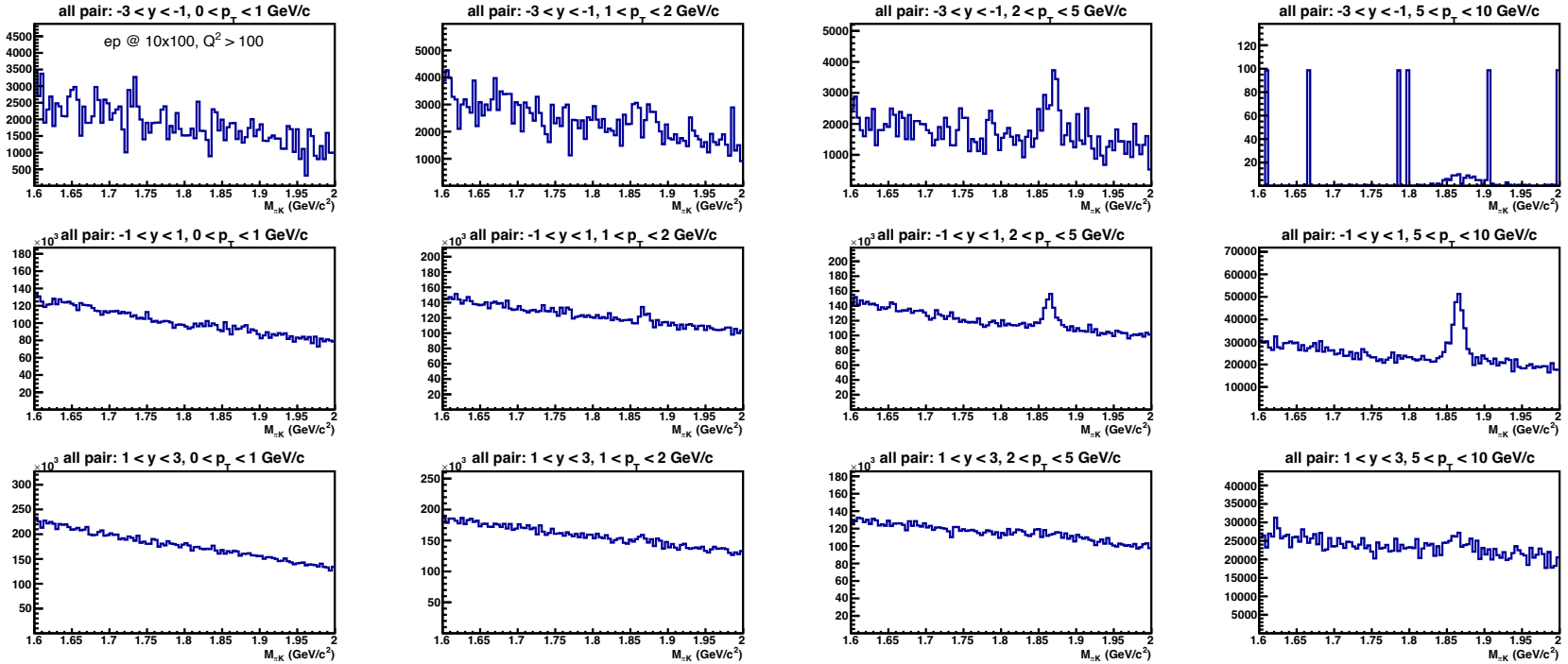
$Q^2 > 100$: D0 sample



$Q^2 > 100$: DIS sample



$Q^2 > 100$: D0 + scaled DIS



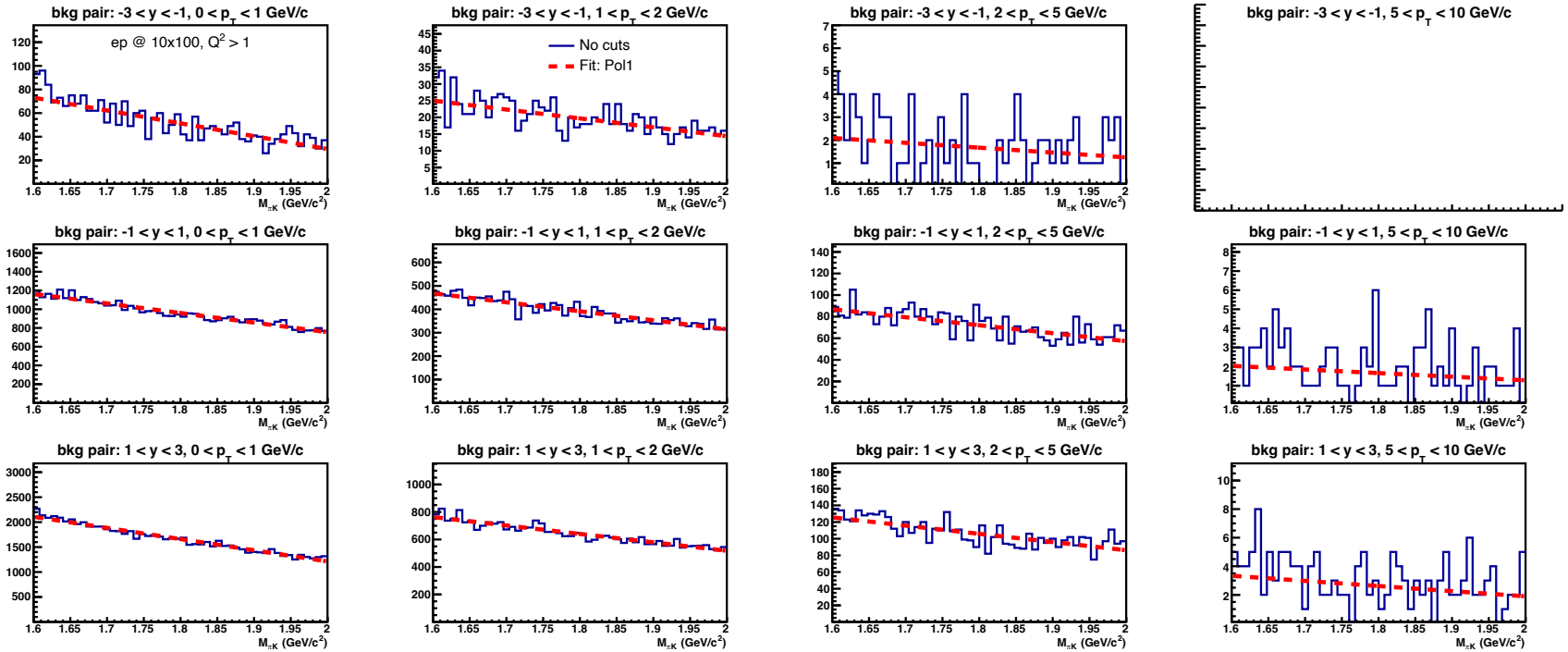
- Fluctuations in background overwhelm signal

How to overcome fluctuations?

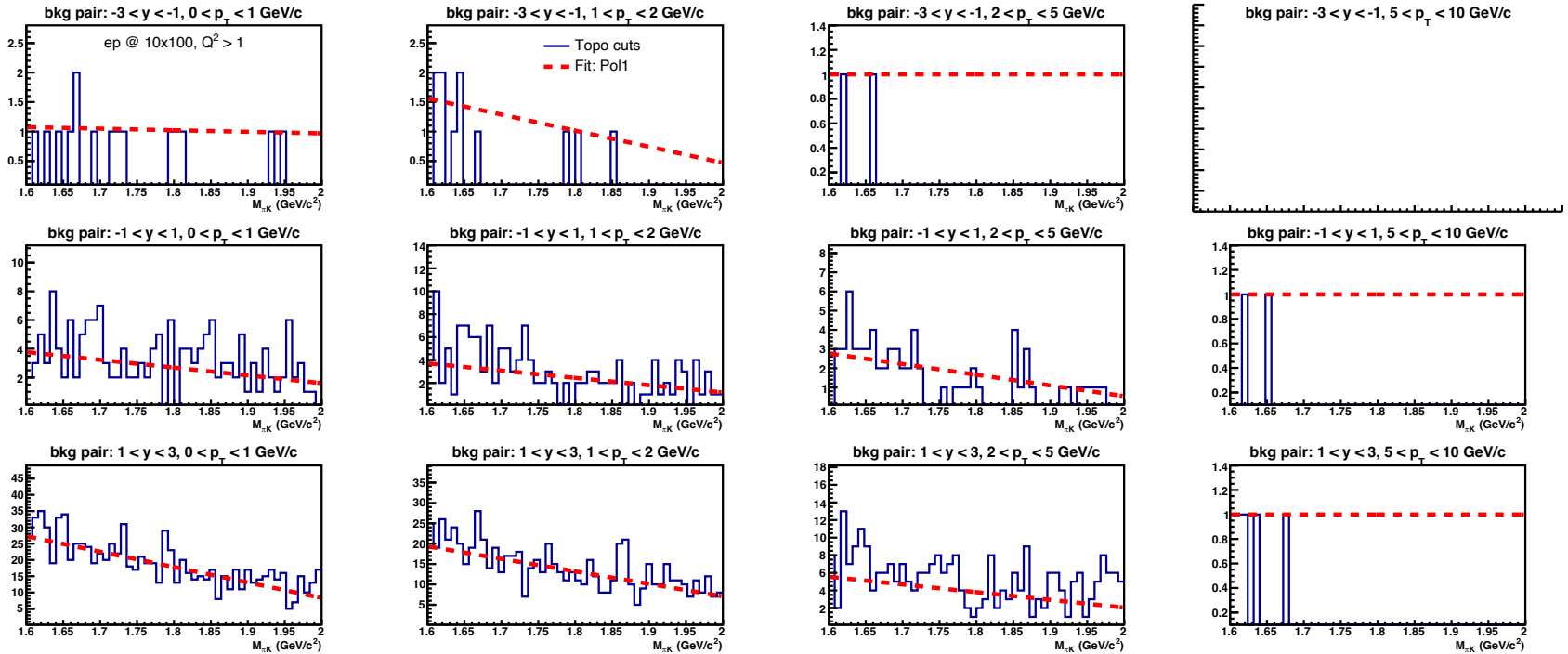
- Fluctuations in background are artificially amplified when scaling DIS samples

- Fit background distributions from DIS, and use the fit function to sample “scaled” background

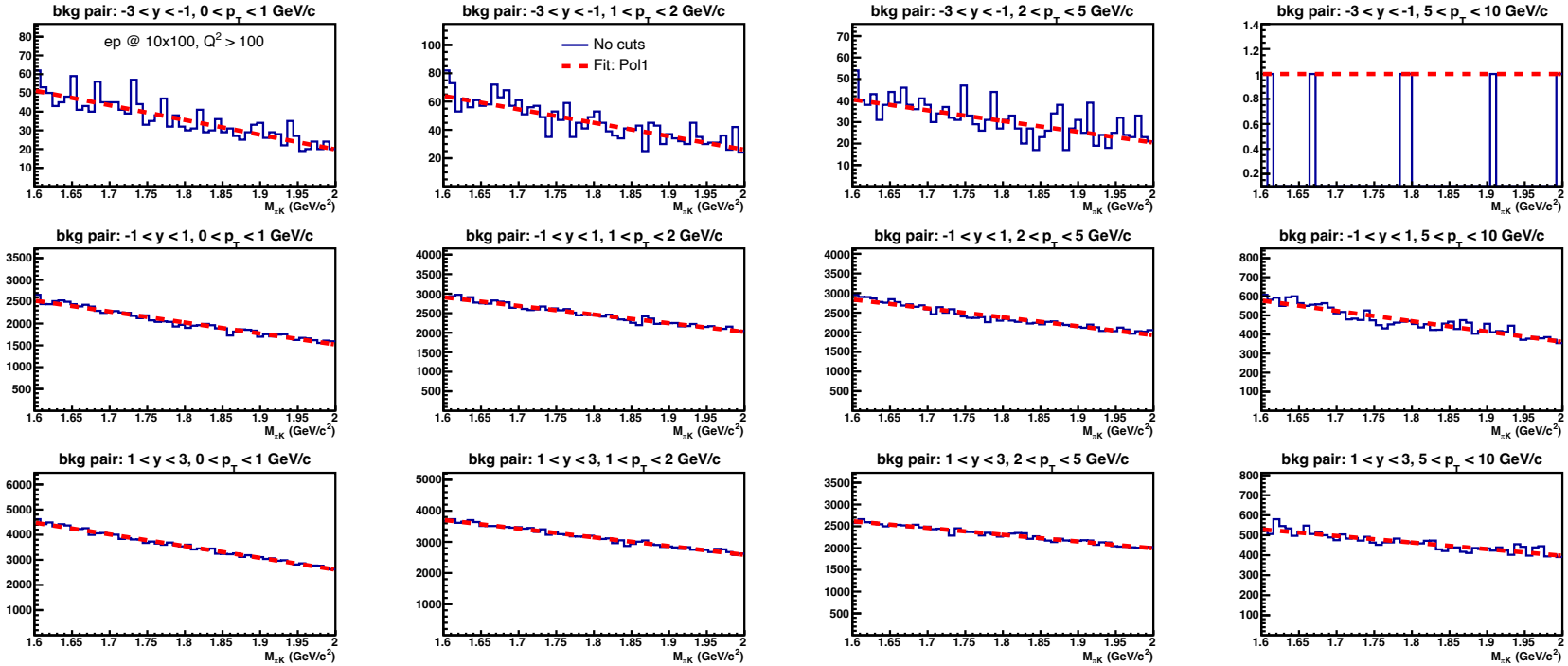
$Q^2 > 1$: fit DIS sample



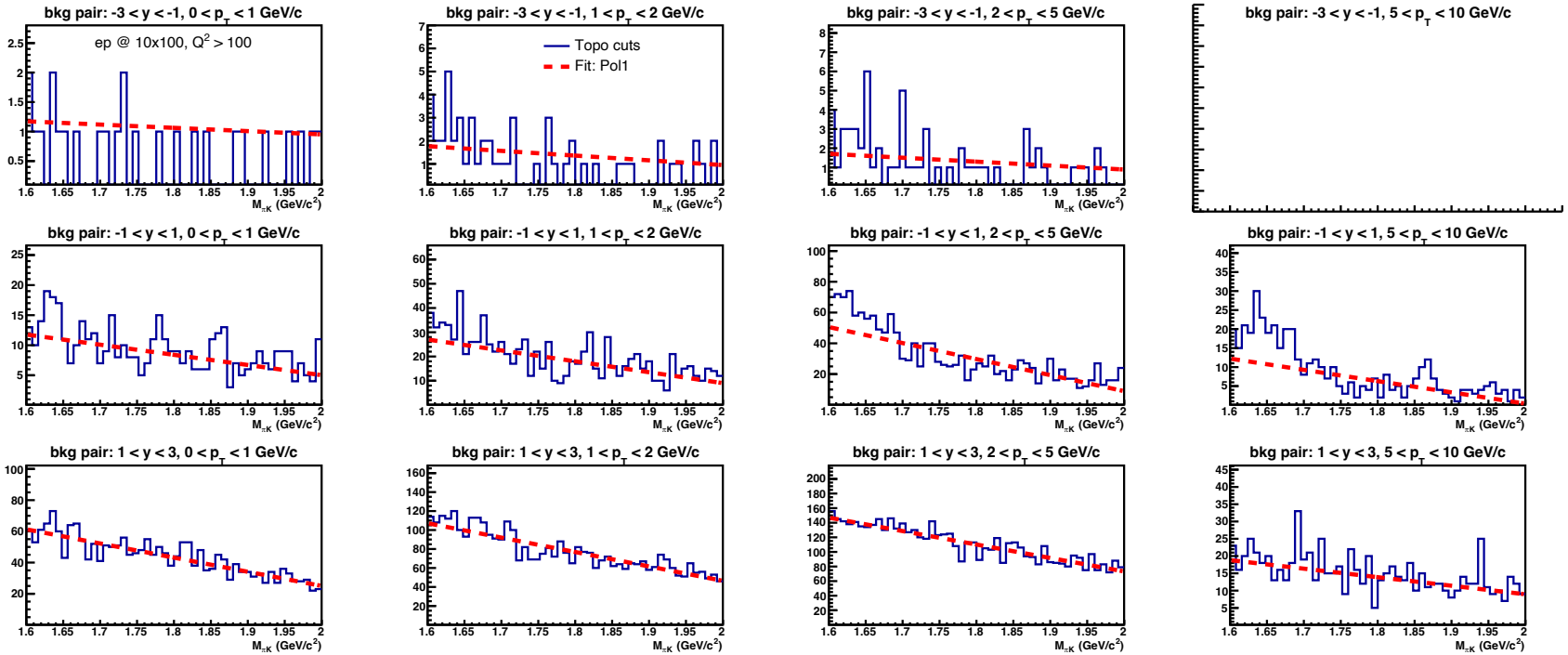
$Q^2 > 1$: fit DIS sample with topo cuts



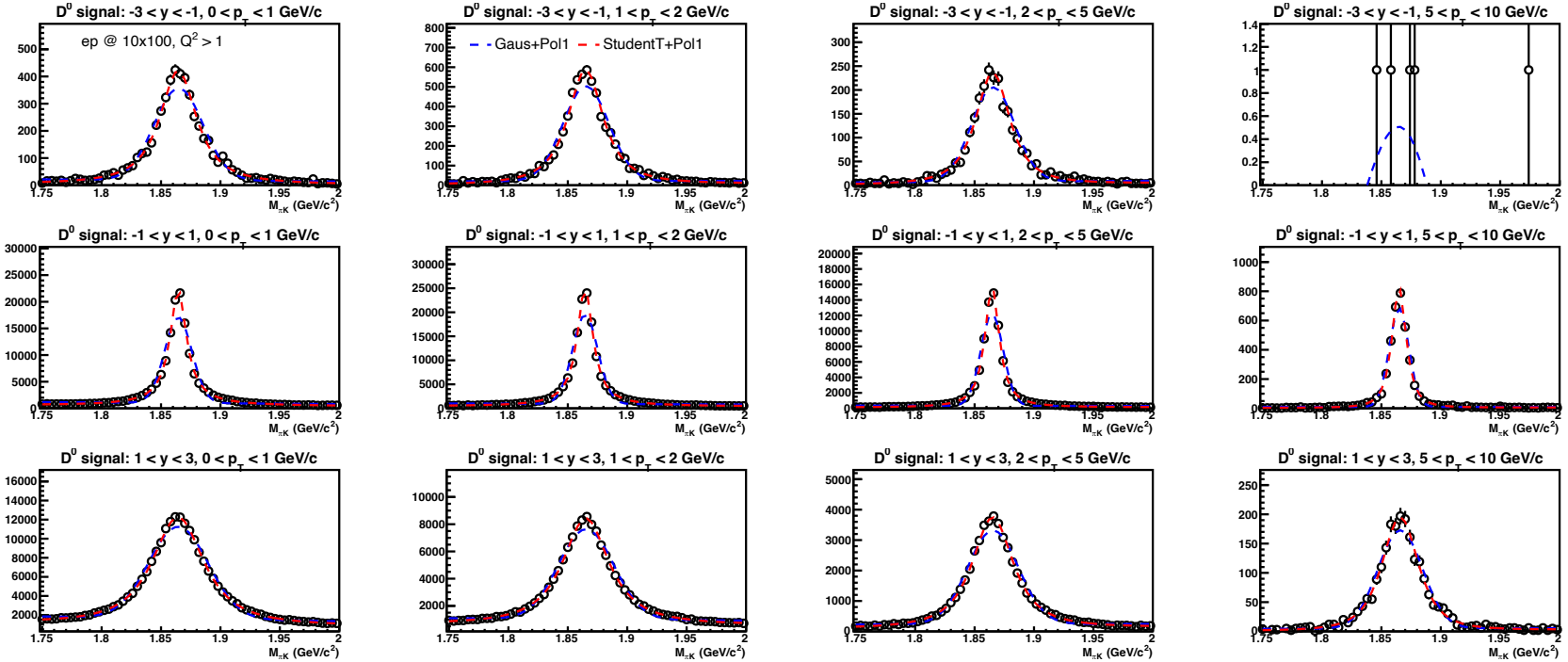
$Q^2 > 100$: fit DIS sample



$Q^2 > 100$: fit DIS sample with topo cuts

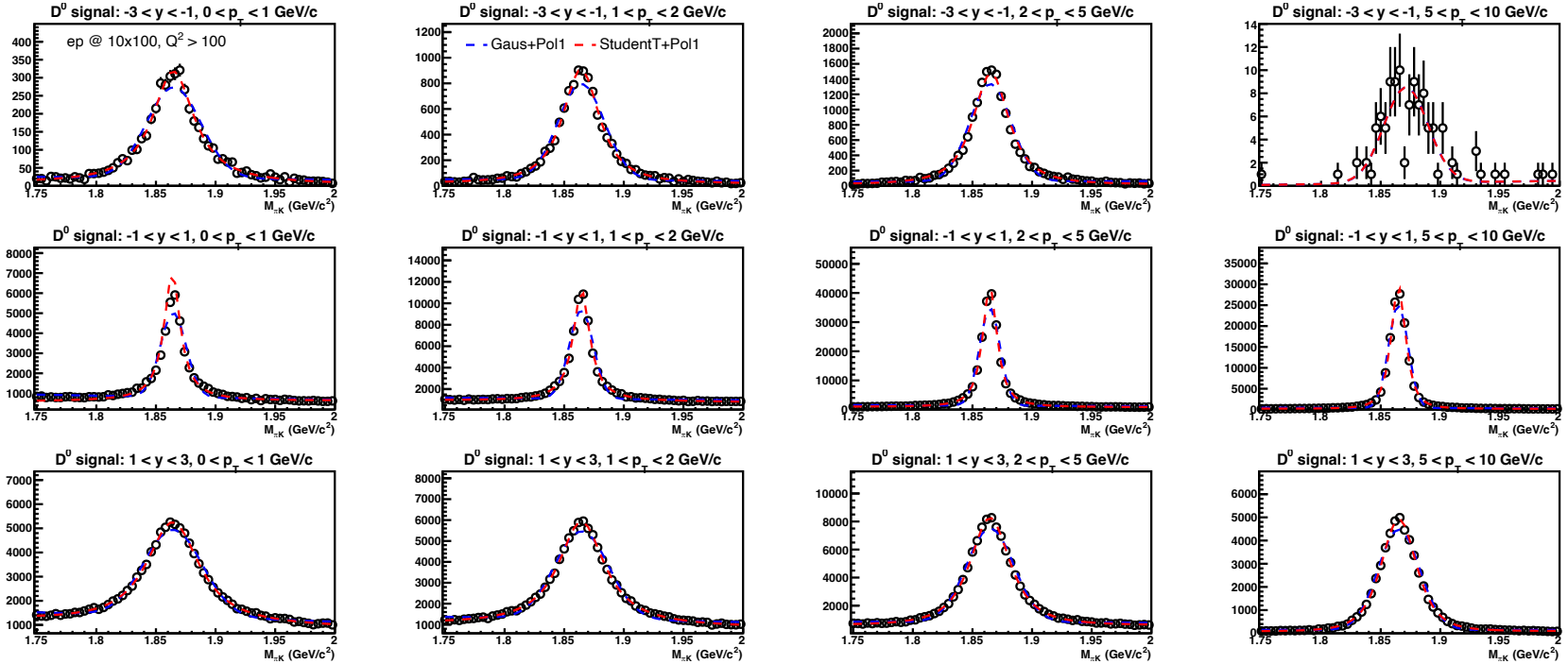


$Q^2 > 1$: fit D0 sample



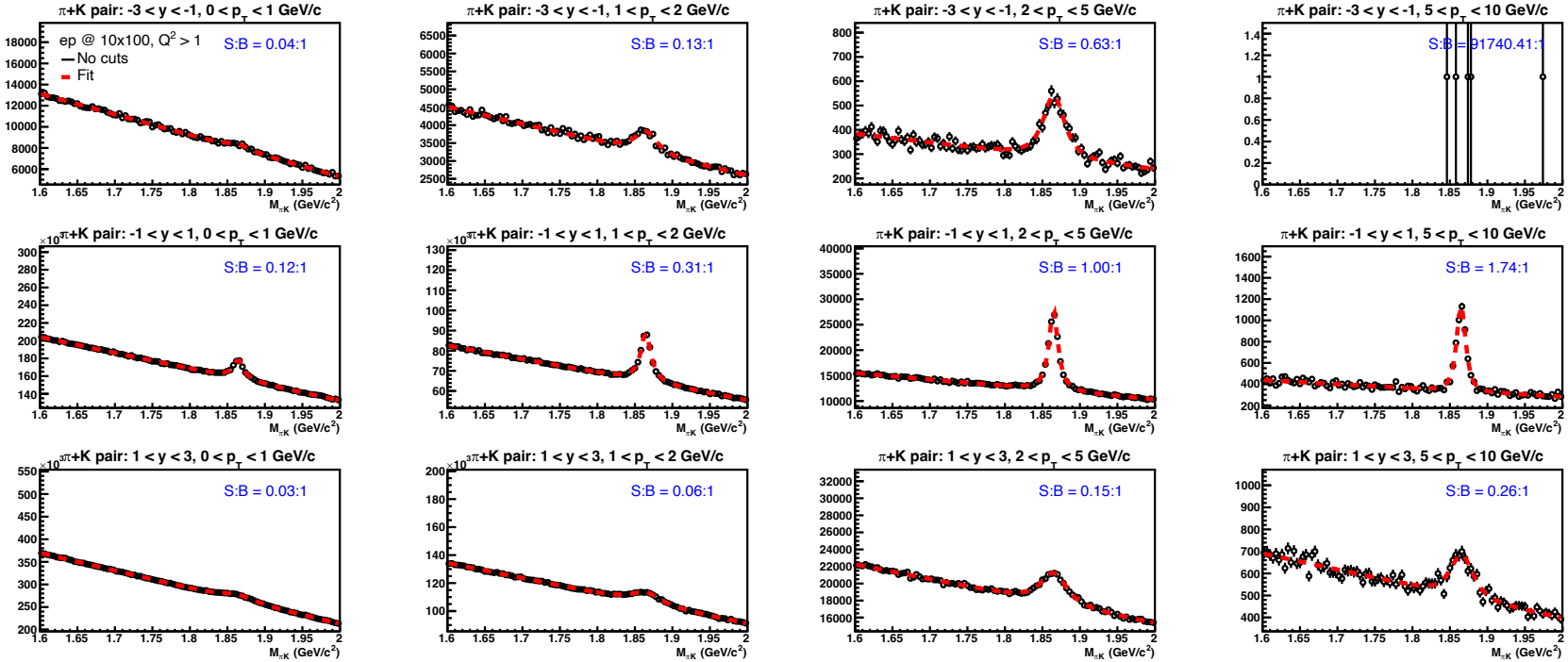
- Student-T function describes signal shape better

$Q^2 > 100$: fit D0 sample



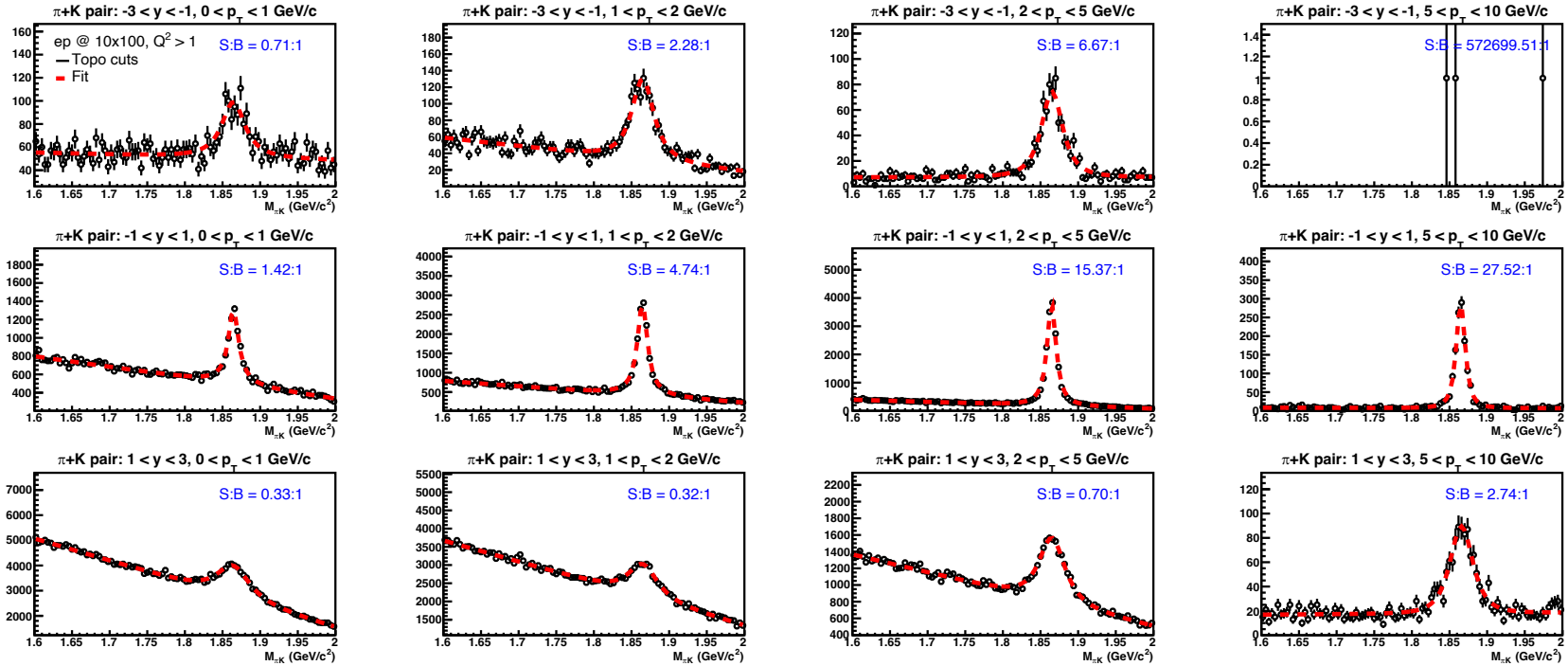
- Student-T function describes signal shape better

$Q^2 > 1$: fit D0+DIS sample

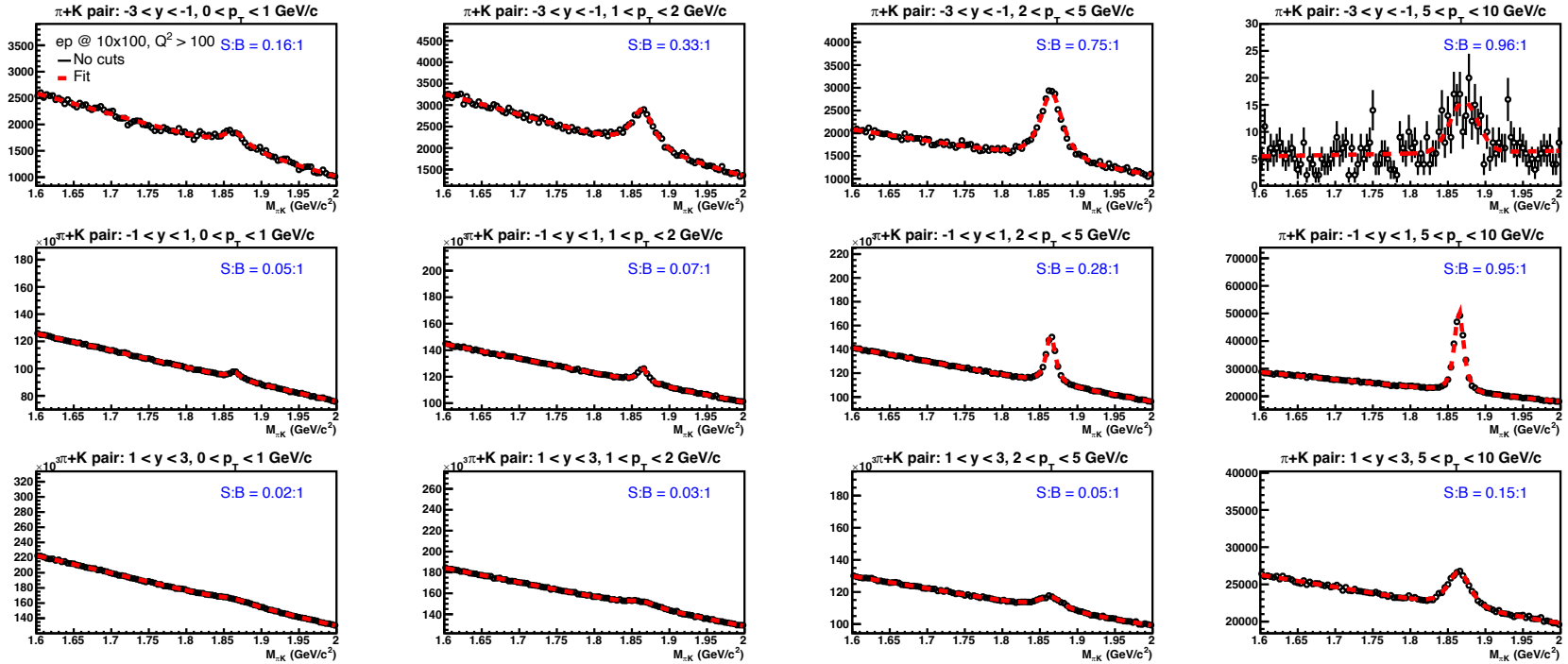


- S/B ratio within 2σ of signal peak

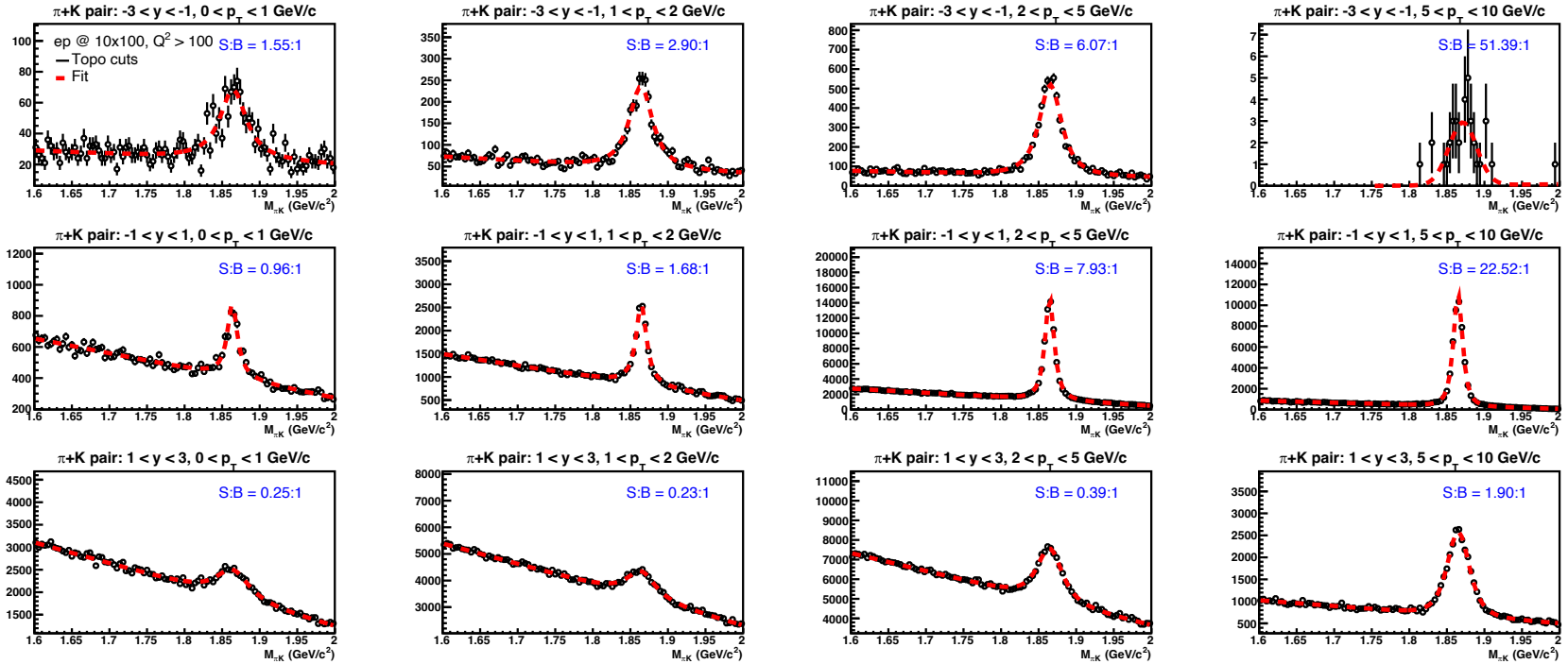
$Q^2 > 1$: fit D0+DIS sample with topo cuts



$Q^2 > 100$: fit D0+DIS sample



$Q^2 > 100$: fit D0+DIS sample with topo cuts



Summary

- D0 shape can be better fit with the student-T function
- Directly scaling DIS samples to mimic background introduces large fluctuations.
  - A fitting & sampling procedure is used to suppress the fluctuations.
  - This procedure does not work for machine learning since the correlations are lost. Producing a DIS sample with sufficient statistics for individual p_T-eta bin will be computationally very expensive.
- With straight topological cuts, good S/B ratios can be achieved in most p_T-eta bins. Maybe we do not need to apply machine learning in individual bins.