Cut-based Optimization: Signal eff. ≥ 90% And Bkg eff. < 1%

Ánuar Sifuentes Name | 19. February 2026
ETP – Institut für experimentelle Teilchenphysik

Content

1. The Quest for εS ≥ 90%: Motivation and Goals
2. Building the Selection: Variables, Cuts and Scan Strategy
3. Did We Reach the Goal? Results and Comparison

Institut für Experimentelle Teilchenphysik

1. The Quest for εS ≥ 90%: Motivation and Goals

Motivation and Goals

IDEA: Reduce background contamination using classical cut-based selection as a first step toward training a Transformer-based classifier.

Goals:

- Signal efficiency **$\varepsilon_S \geq 90\%$**
- Background efficiency **$\varepsilon_B < 1\%$**

Strategy:

- Accept **$\varepsilon_B < 5\%$** as an intermediate milestone
- Scan signal efficiency thresholds from 70% to 100%

2. Building the Selection: Variables, Cuts and Scan Strategy

Variables used for cut optimization

- EVT_ThrustEmin_E
- EVT_ThrustEmin_Eneutral
- EVT_ThrustEmin_Nneutral
- EVT_ThrustEmax_E
- recoEmiss_e
- EVT_NTau23Pi
- EVT_ThrustEmin_NDV
- EVT_NVertex
- mDiTau_Vis
- EVT_ThrustEmin_NTau23PiCand

Four variable configurations were tested, each using a different subset of the available discriminating variables (see Apendix)

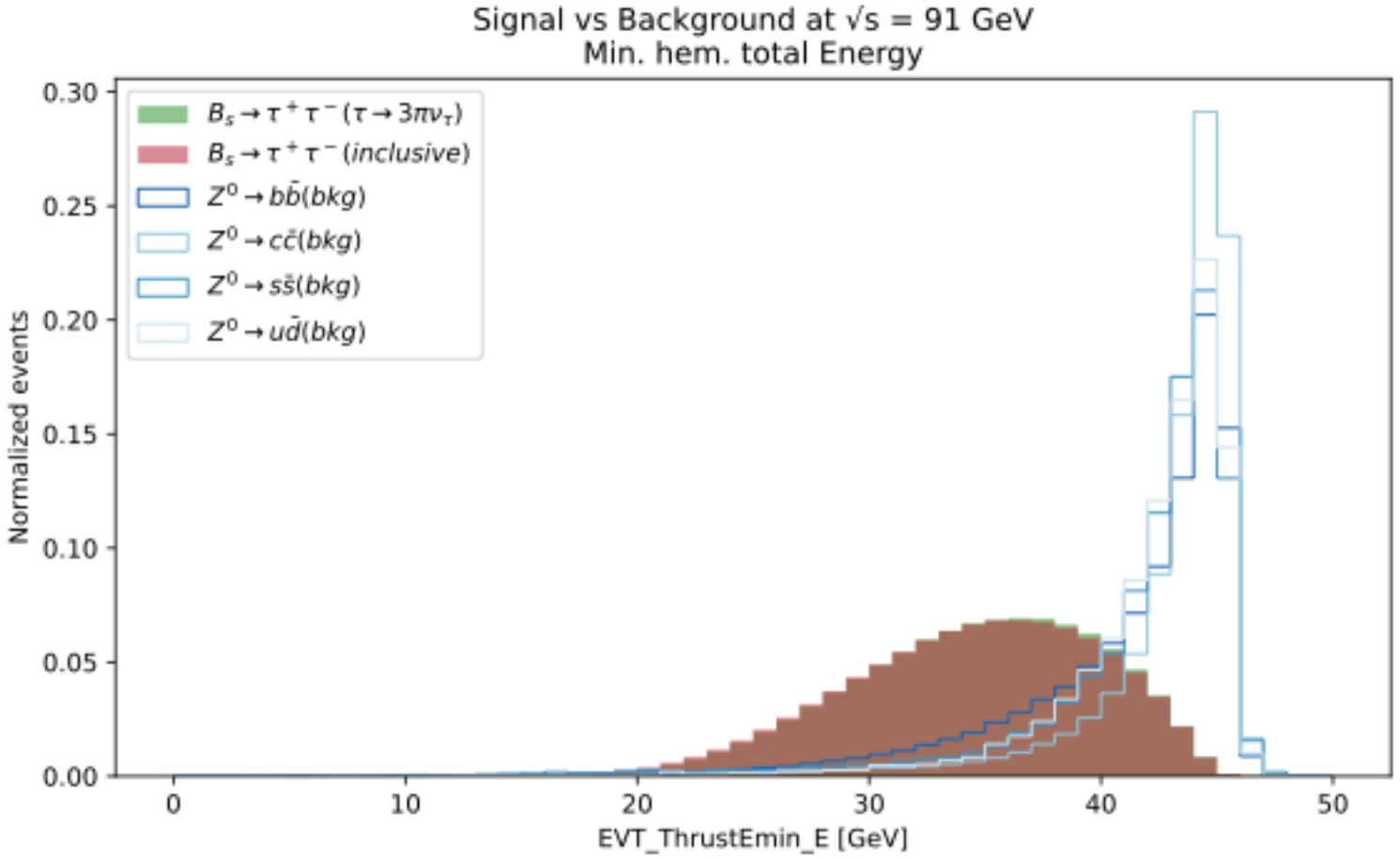
How was done the cut optimization?

1. Inspect signal and background distributions per variable
2. Define cut ranges where separation is most evident
3. Sample cut values uniformly within each range
4. Evaluate all possible cut combinations (combinatorial scan)

   At least 10^5 combinations evaluated per configuration

5. Save combinations satisfying **$\varepsilon S \geq$ target AND $\varepsilon B < 5\%$**
6. Scan target signal efficiency from 100% down to 70%
7. Save the cut that **maximizes significance $S/\sqrt{B}$**

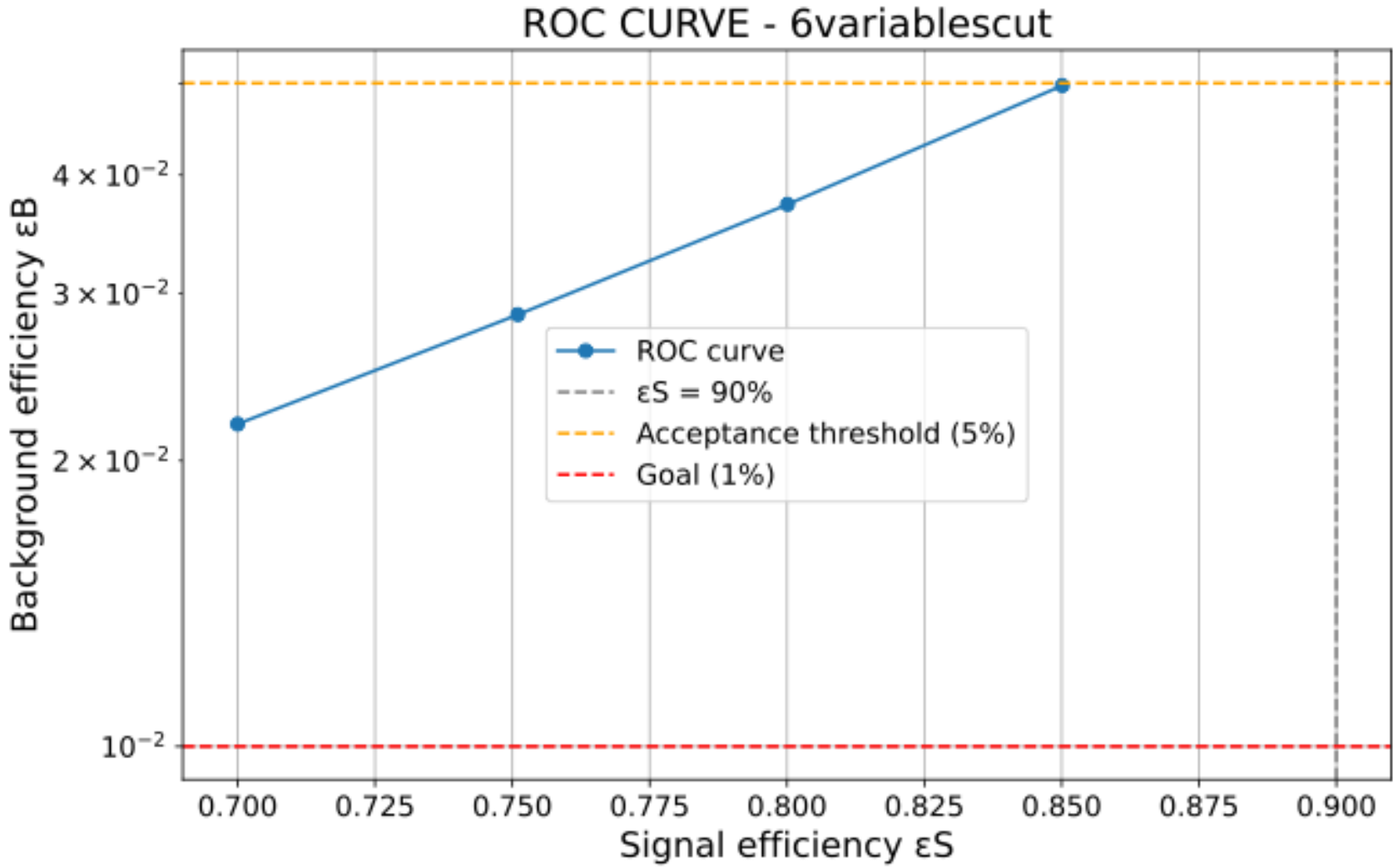
Example of possible range



EVT_ThrustEmin_E_range
= np.linspace(40, 47, 8)

3. Did We Reach the Goal? Results and Comparison

“6variablescut”-ROC Curve

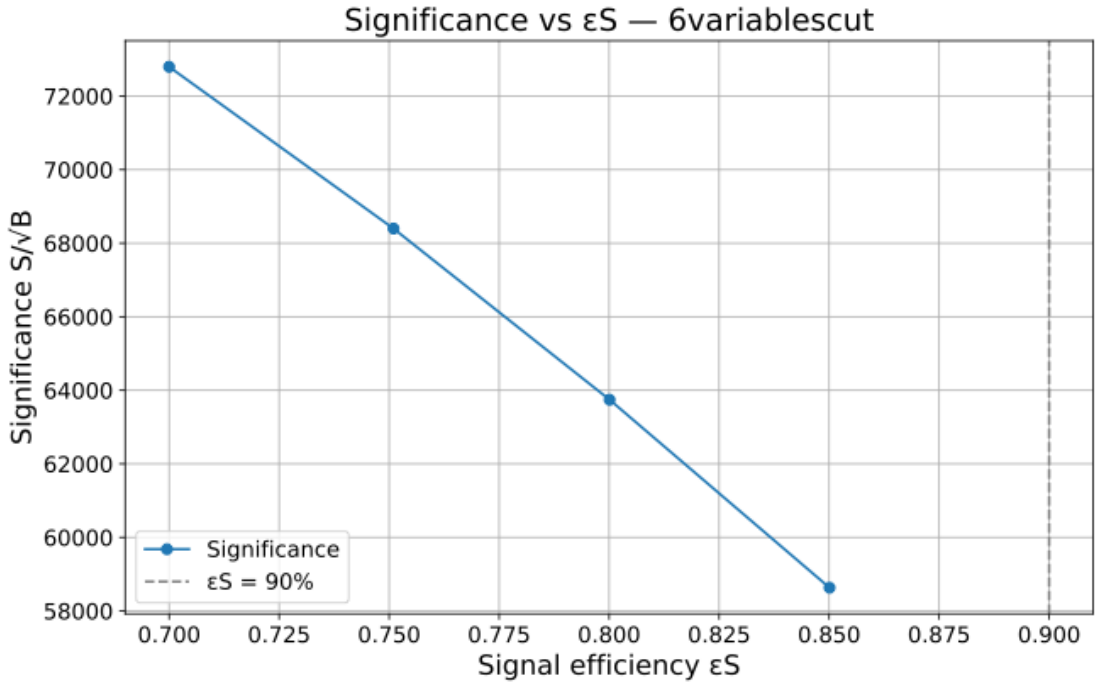


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.00% and εB = 2.19%)

(highest signal eff:
εS = 85.01% and εB = 4.97%)

“6variablescut”-Significance plots

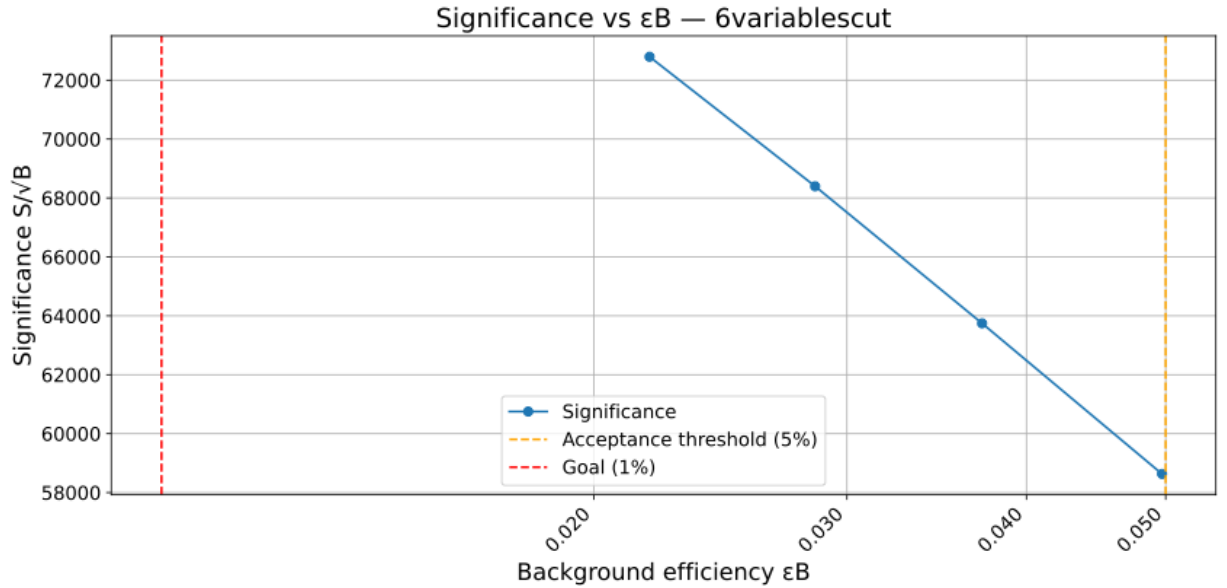


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.00% and εB = 2.19%)

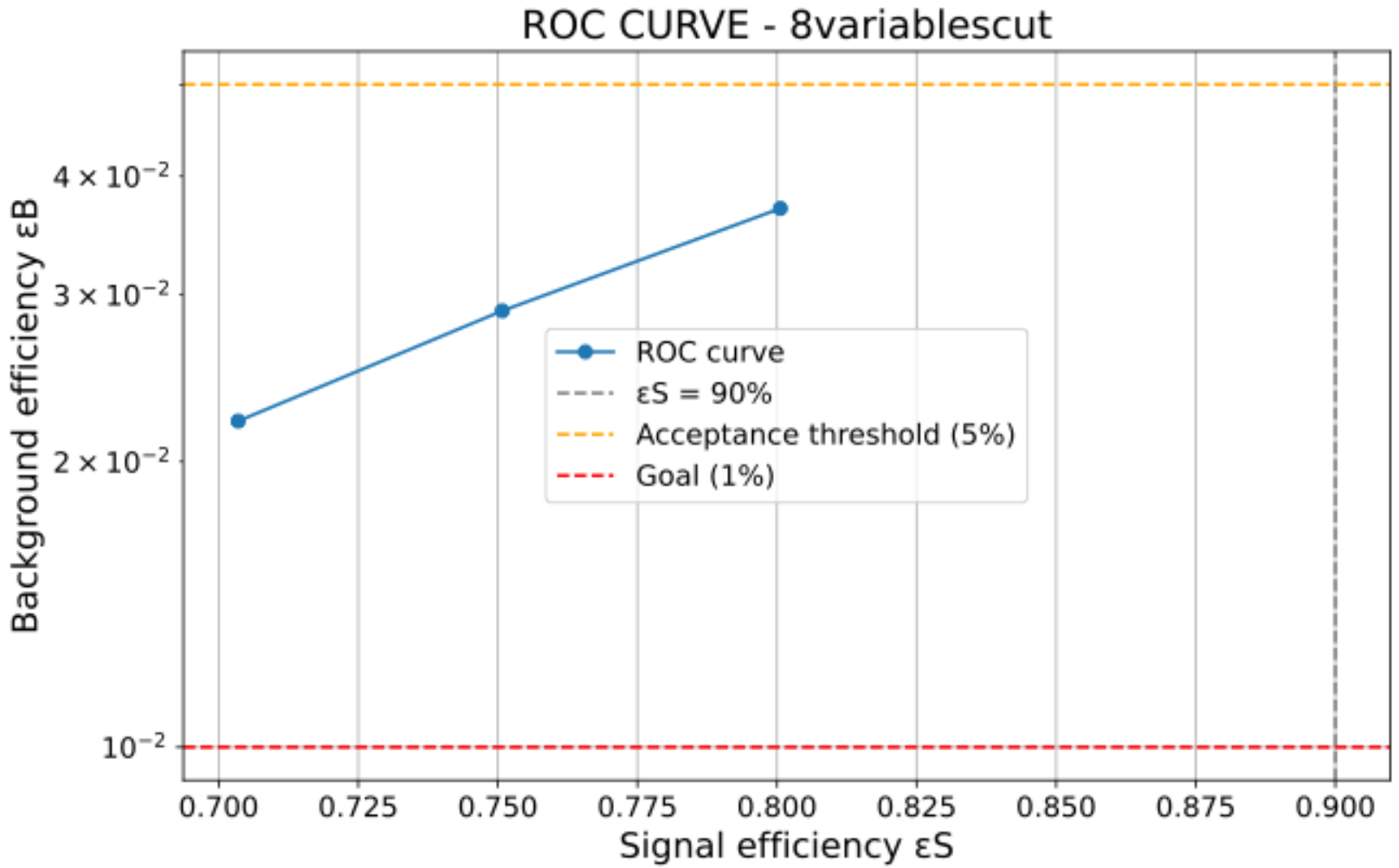
(highest signal eff:
εS = 85.01% and εB = 4.97%)

Significance vs
Bkg efficiency



Significance vs
Signal efficiency

“8variablescut”-ROC Curve

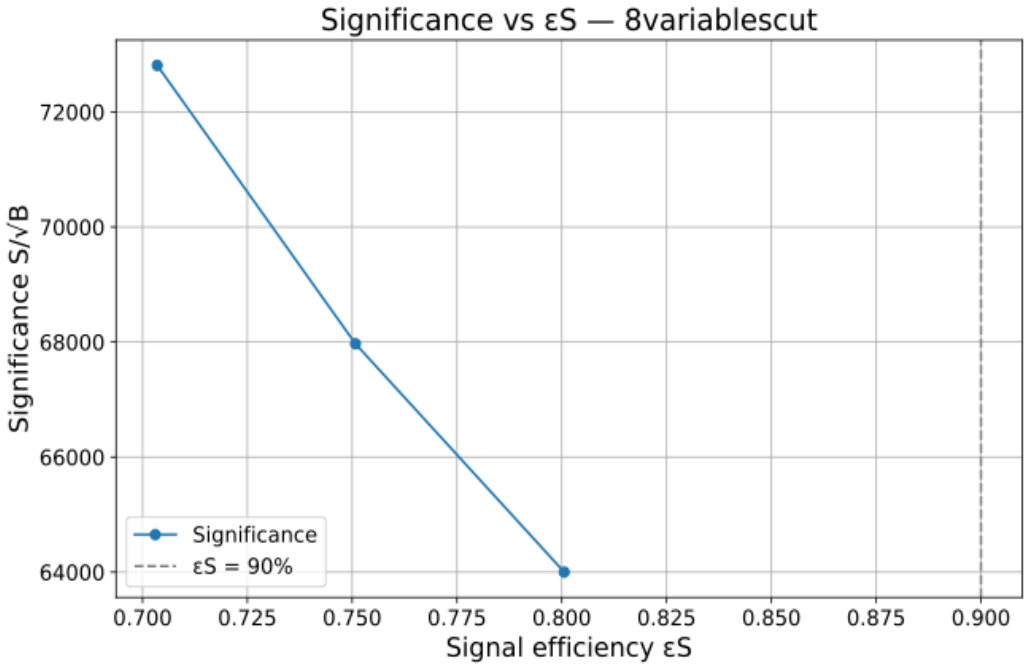


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.35% and εB = 2.21%)

(highest signal eff:
εS = 80.06% and εB = 3.70%)

“8variablescut”-Significance plots



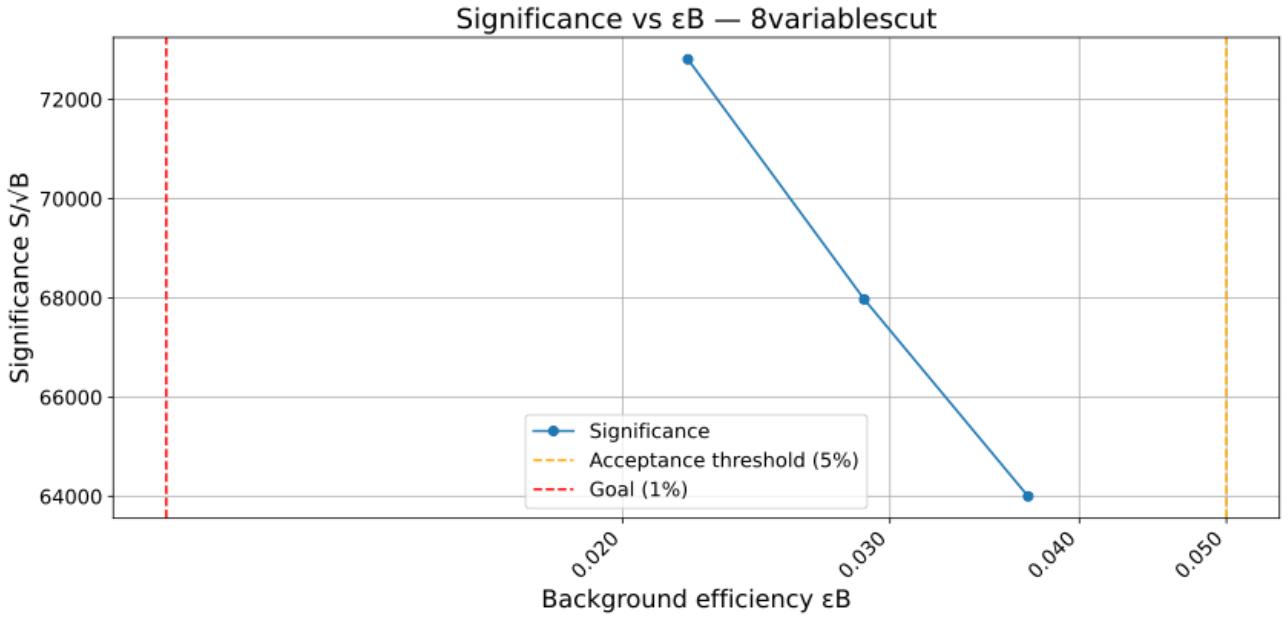
No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.35% and εB = 2.21%)

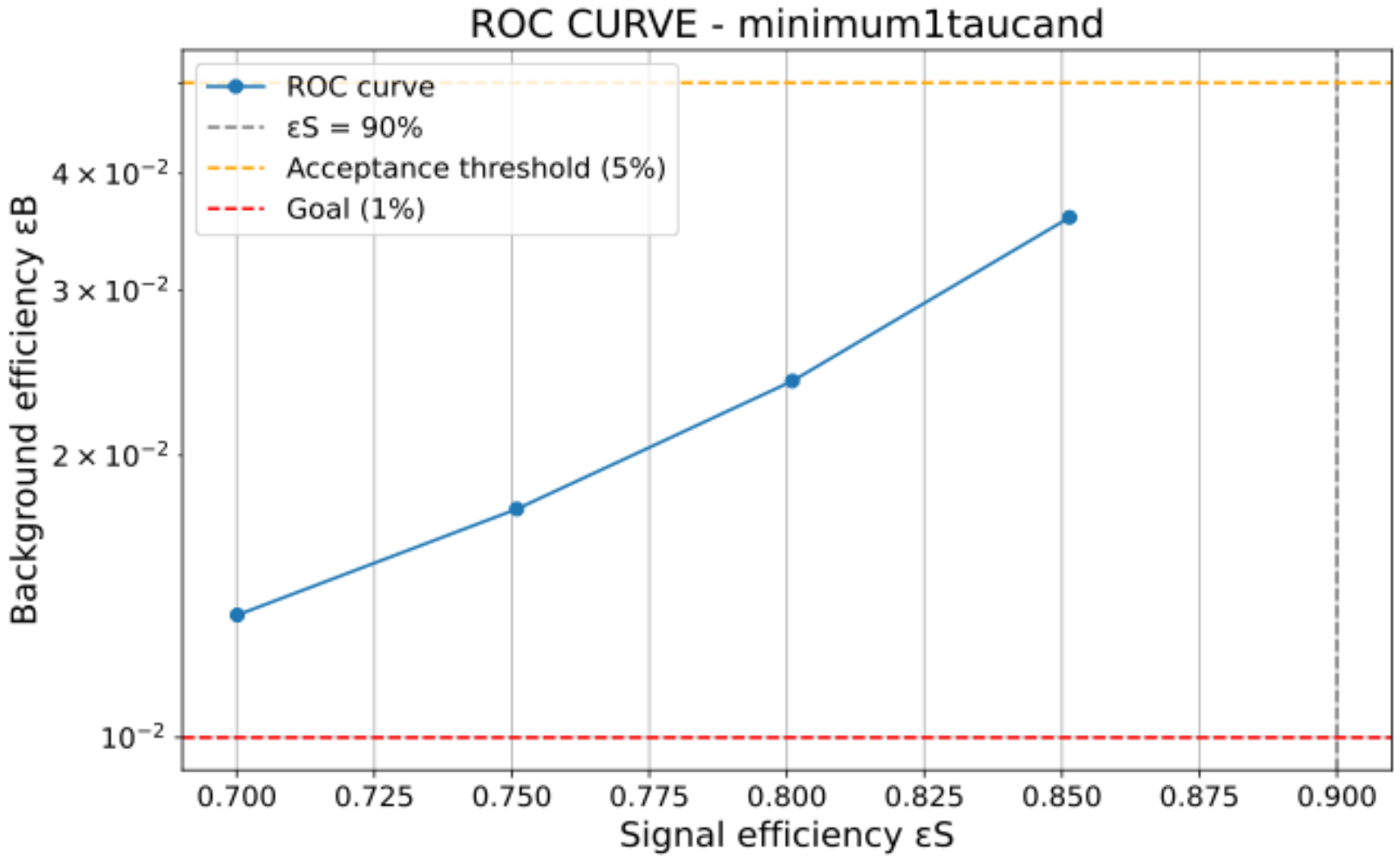
(highest signal eff:
εS = 80.06% and εB = 3.70%)

Significance vs
Bkg efficiency

Significance vs
Signal efficiency



“minimum1TauCand”-ROC Curve

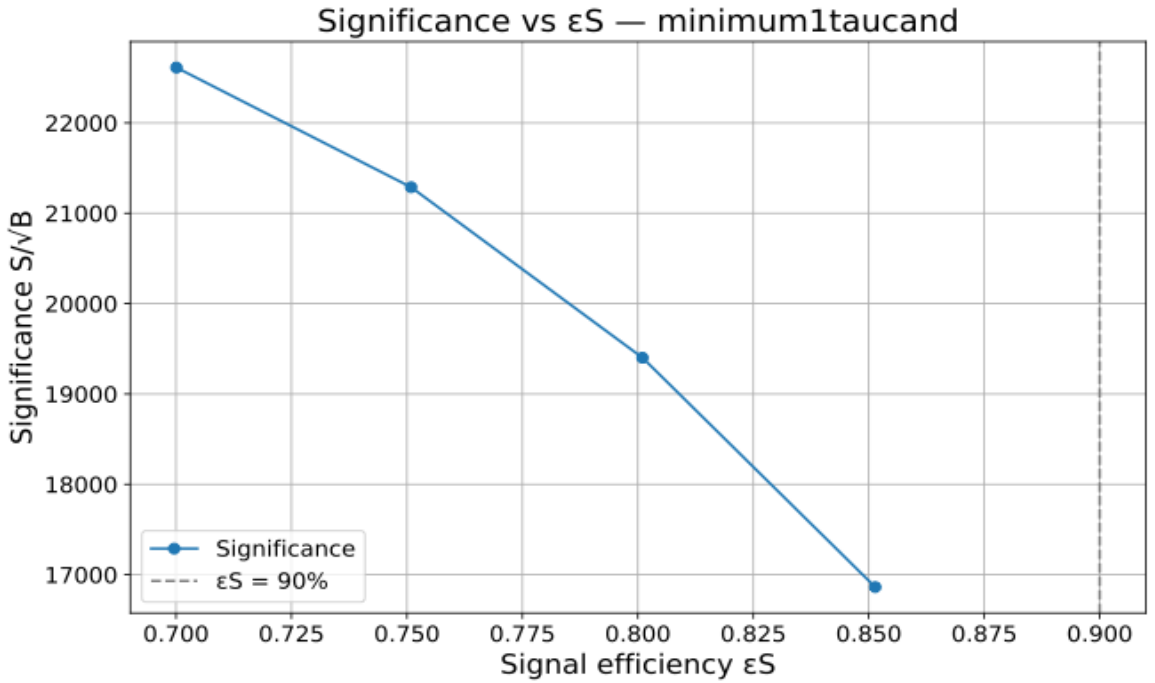


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.01% and εB = 1.35%)

(highest signal eff:
εS = 85.14% and εB = 3.59%)

“minimum1TauCand”-Significance plots

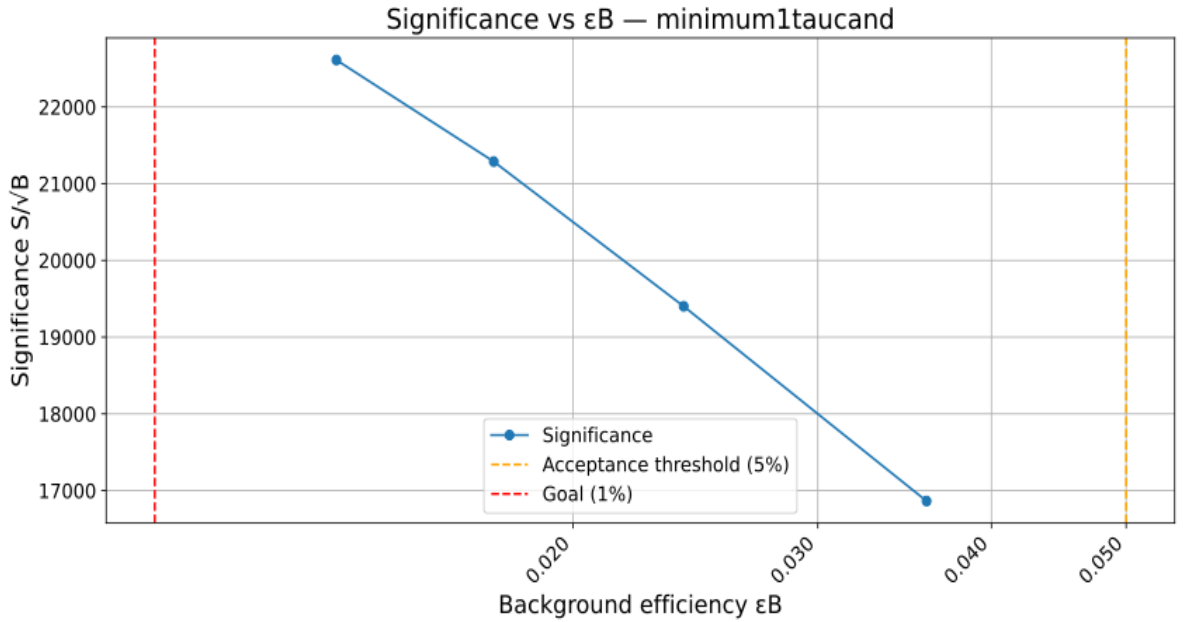


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 70.01% and εB = 1.35%)

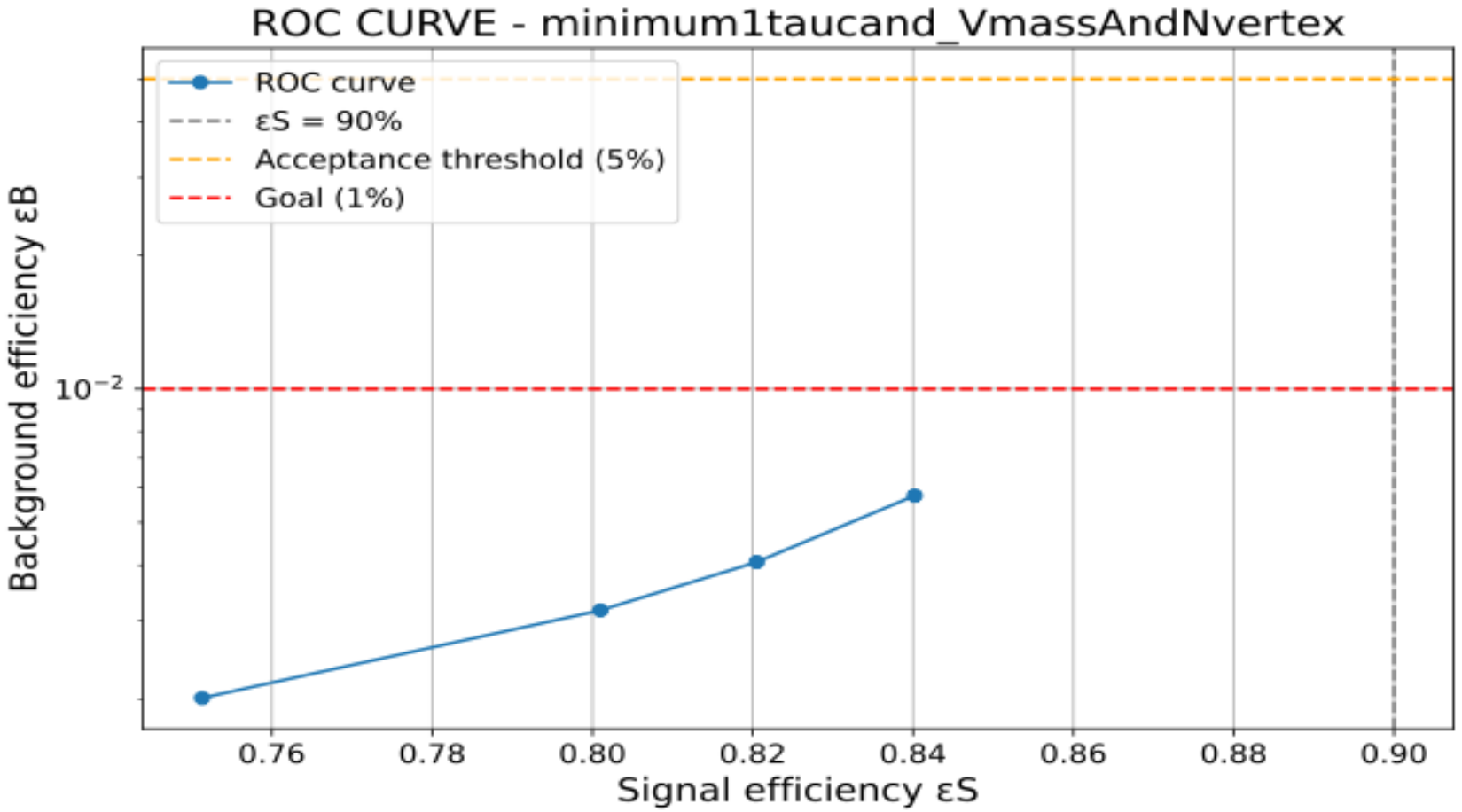
(highest signal eff:
εS = 85.14% and εB = 3.59%)

Significance vs
Bkg efficiency



Significance vs
Signal efficiency

“minimum1TauCand_VmassAndNVertex”- ROC Curve

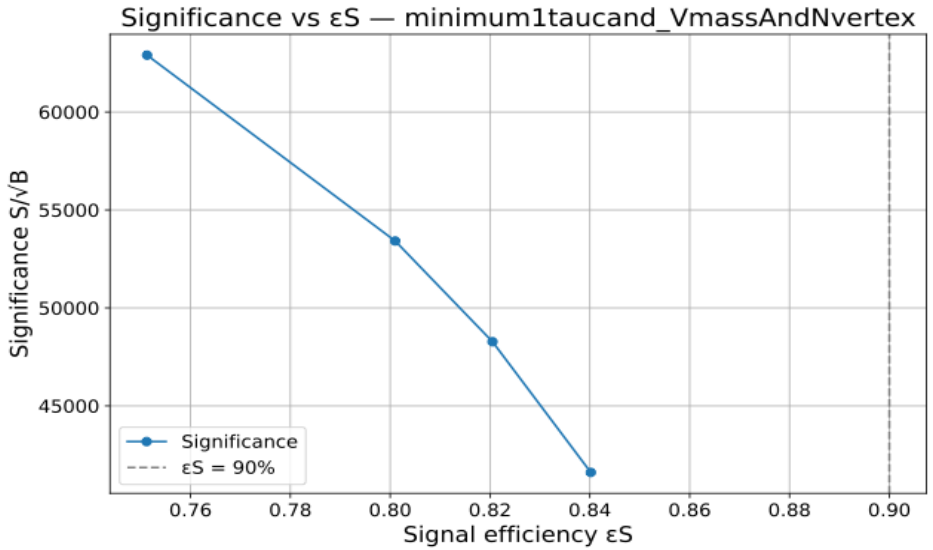


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 75.13% and εB = 0.20%)

(highest signal eff:
εS = 84.02% and εB = 0.57%)

“minimum1TauCand_VmassAndNVertex”- Significance plots

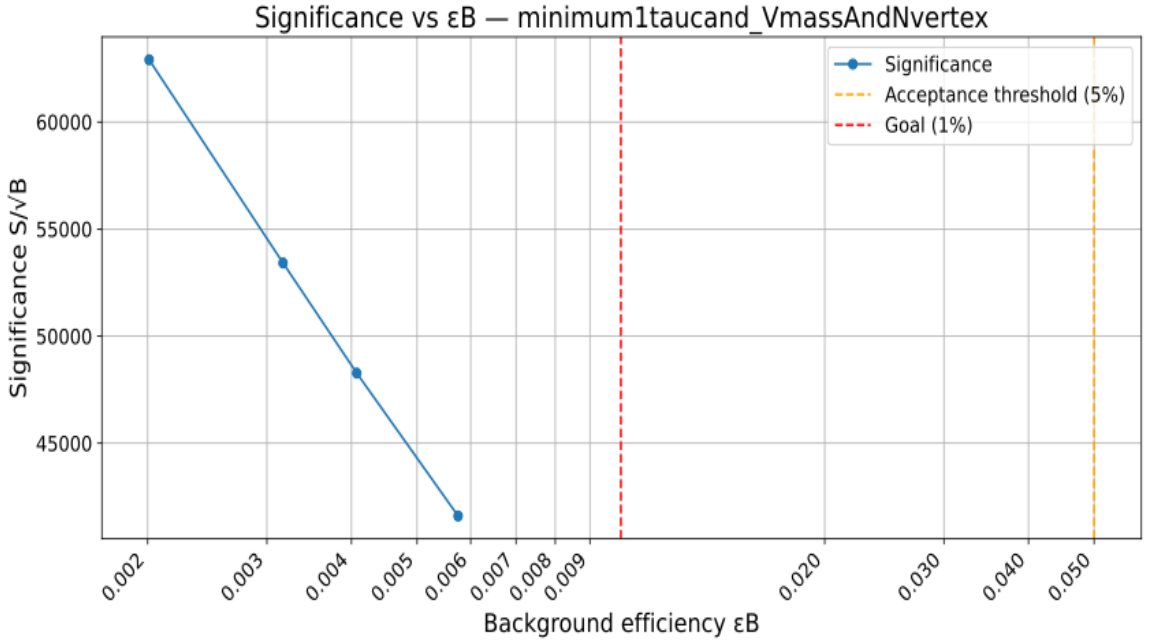


No cuts satisfied
εS ≥ 90% AND εB < 1%

(best significance:
εS = 75.13% and εB = 0.20%)

(highest signal eff:
εS = 84.02% and εB = 0.57%)

Significance vs
Bkg efficiency



Significance vs
Signal efficiency

Best cuts found from “minimum1TauCand_VmassAndNVertex”

- EVT_ThrustEmin_E < 47.00 GeV
- EVT_ThrustEmin_Eneutral < 13.00 GeV
- EVT_ThrustEmin_Nneutral < 23.00
- EVT_ThrustEmax_E < 55.00 GeV
- recoEmiss_e > 2.00 GeV
- EVT_NTau23Pi > 1.00
- EVT_ThrustEmin_NDV > 1.00
- EVT_Nvertex > 1.00
- mDiTau_Vis > 2.00
- EVT_ThrustEmin_NTau23PiCand > 1.00

εS = 84.02%
εB = 0.57%
S/√B: 41600

Summary + Practical Next Steps

Status:

- The goal of **εS ≥ 90%** AND **εB < 1%** was not achieved in any configuration
- Best results obtained **εS = 84.02%; εB = 0.57%**

  from (minimum1TauCand_VmassAndNVertex)
- This represents a promising baseline for further optimization

Next Steps:

- Apply the best cut configuration as a preselection
- Train a Transformer-based classifier on the preselected sample
- Evaluate additional background rejection with respect to this baseline

Thank you for your attention!

Appendix

	6variablecut	8variablecut	minimum1TauCand	min1TauCan_VmassAndNVertex
ThrustEmin_E	✅	✅	✅	✅
ThrustEmin_Eneutral	✅	✅	✅	✅
ThrustEmin_Nneutral	✅	✅	✅	✅
ThrustEmax_E	✅	✅	✅	✅
recoEmiss_e	✅	✅	✅	✅
NTau23Pi	✅ (fixed ≥ 1)	✅	✅	✅
ThrustEmin_NDV	❌	✅	✅	✅
NVertex	❌	❌	❌	✅
mDiTau_Vis	❌	❌	❌	✅
ThrustEmin_NTau23 PiCand	❌	✅ (fixed ≥ 2)	✅ (fixed ≥ 1)	✅ (fixed ≥ 1)