WARWICK
THE UNIVERSITY OF WARWICK

Analysis-driven optimisation of Pandora

* Shout out to Dom for being a great support – thank you for the help!!

Isobel Mawby DUNE UK 24/8/21

Introduction

WARWICK
THE UNIVERSITY OF WARWICK

An optimised traditional nue/numu selection procedure

- **Traditional?** In the sense that an event is assessed on the basis of its leading lepton (should it exist)
- **Optimised?** Selection procedure is tuned to maximise the sensitivity to δ_{CP}

Analysis Workflow

Reconstructed far detector MC event

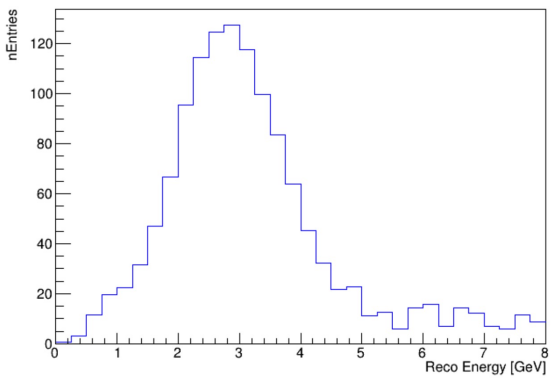
Is CC nue selected?

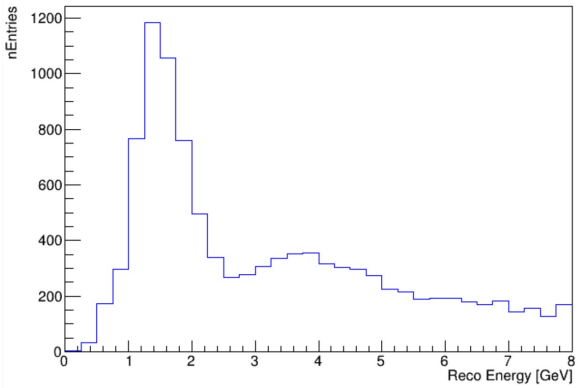
no

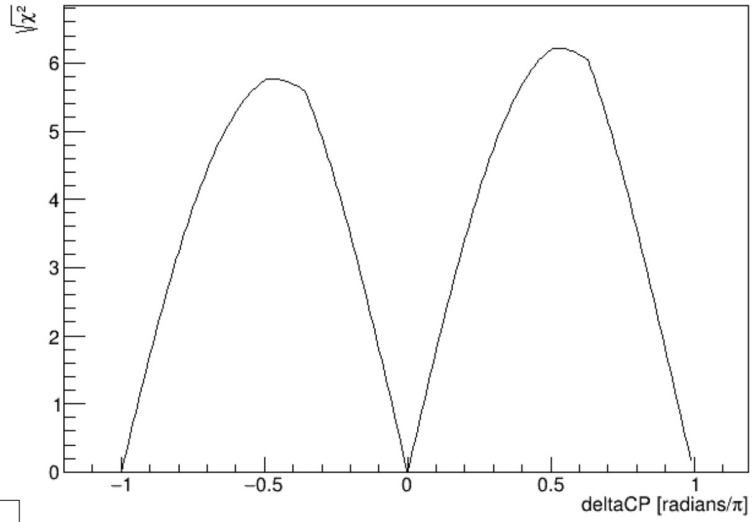
Is CC numu selected?

yes

yes







$$\sqrt{\chi^2} = \min(\sqrt{\chi^2}_{\delta_{CP}=0}, \sqrt{\chi^2}_{\delta_{CP}=\pi})$$

$$\chi^2 = 2\sum_i (E_i - O_i + O_i \log\frac{O_i}{E_i})$$

$\delta_{CP} = 0, \pi$ $\delta_{CP} \neq 0, \pi$

*For 200 values of δ_{CP}

Is this a nue event?

WARWICK
THE UNIVERSITY OF WARWICK

Reject if the reconstructed neutrino vertex is outside the DUNE fiducial volume

↓

Assign all showers **pandrizzle** scores, choosing the **pandrizzliest** shower to be the electron candidate

↓

Reject event if the candidate electron pandrizzle score falls **below** a cut value

Electron/gamma separation is difficult! So to increase purity…

Assign all tracks **pandizzle** scores, choosing the **pandizzliest** track to be the muon candidate

↓

Reject event if the candidate muon pandizzle score falls **above** a cut value

*credit to Dom Brailsford for development

Pandrizzle*: a BDTG that assigns a score to a shower in the interval is [-1, 1] reflecting how electron-like it is

Pandizzle*: a BDTG that assigns a score to a track in the interval [-1, 1] reflecting how muon-like it is

Tune cuts such that the **deltaCP sensitivity coverage** is optimised

Is this a numu event?

WARWICK
THE UNIVERSITY OF WARWICK

Reject if the reconstructed neutrino vertex is outside the DUNE fiducial volume

↓

Assign all tracks **pandizzle** scores, choosing the **pandizzliest** track to be the muon candidate

↓

Reject event if the candidate muon pandizzle score falls **below** a cut value

*credit to Dom Brailsford for development

Pandrizzle*: a BDTG that assigns a score to a shower in the interval is [-1, 1] reflecting how electron-like it is

Pandizzle*: a BDTG that assigns a score to a track in the interval [-1, 1] reflecting how muon-like it is

Tune cuts such that the selection efficiency*purity is optimised

Selection Metrics

WARWICK
THE UNIVERSITY OF WARWICK

$$\text{efficiency} = \frac{\#\ \text{signal selected}}{\text{total signal}}$$

$$\text{purity} = \frac{\#\ \text{flavour signal selected}}{\text{total selected}}$$

$$\text{BG rejection} = \frac{\#\ \text{background rejected}}{\text{total background}}$$

Signal:

- CC nue(numu) interaction on liquid argon with a true neutrino vertex inside the DUNE FD fiducial volume

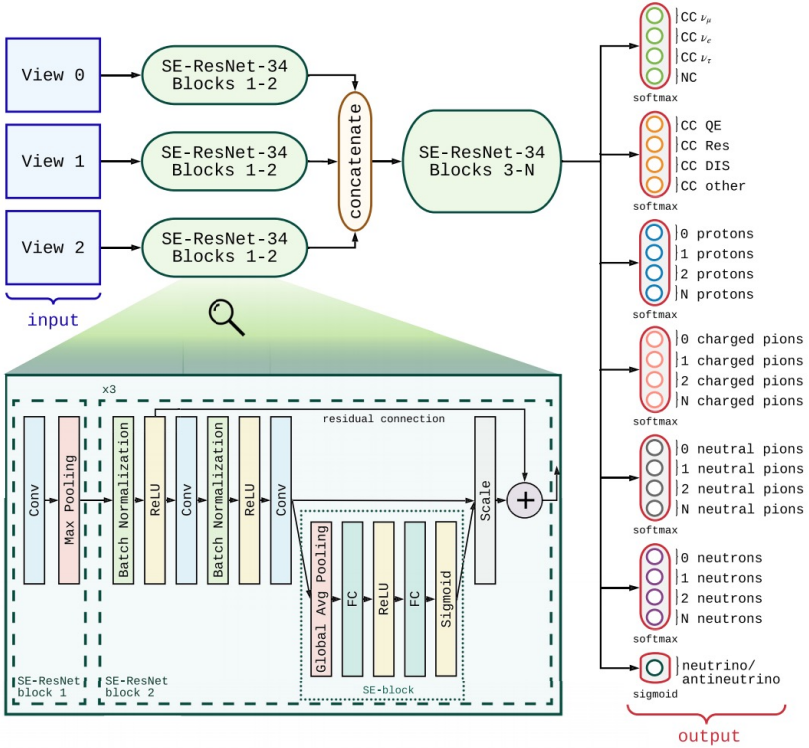
Flavour signal:

- CC nue(numu) or anue(anumu) interaction on liquid argon with a true neutrino vertex inside the DUNE FD fiducial volume

	Nue Efficiency	Nue Purity	Nue BG	Numu Efficiency	Numu Purity	Numu BG
‘izzle selection’	60.0%	67.1%	98.6%	88.3%	87.2%	94.4%

DUNE CVN

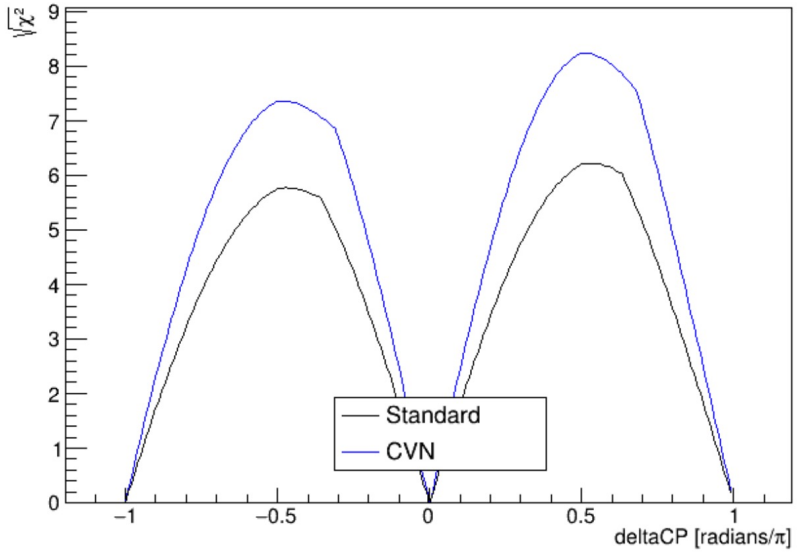
- Details documented in https://journals.aps.org/prd/pdf/10.1103/PhysRevD.102.092003



- The DUNE CVN is a **convolutional neural network**
- Has access to the **whole** image
- For (a)nue selection: P(nue) > 0.85
- For (a)numu selection: P(numu) > 0.5

DUNE CVN Comparison

	Nue Efficiency	Nue Purity	Nue BG	Numu Efficiency	Numu Purity	Numu BG
CVN	82.7%	90.9%	99.6%	88.2%	96.6%	98.7%
‘izzle selection’	60.0%	67.1%	98.6%	88.3%	87.2%	94.4%



The CVN is really, really good! How can we get closer, where are our losses?

Cheating Studies

WARWICK
THE UNIVERSITY OF WARWICK

- Perform cheating studies to understand what is limiting our sensitivity
- **Cheating Studies** – When working with MC data we can slot in the truth for a particular reconstruction task
  e.g. Set the neutrino vertex by accessing the MC neutrino vertex instead of using Pandora algorithms
- Begin with **'vague'** cheating hypotheses and use results to motivate future studies and pinpoint areas that can be worked on

Pandora Reconstruction → Track/shower characterisation → Selection → Energy Estimation

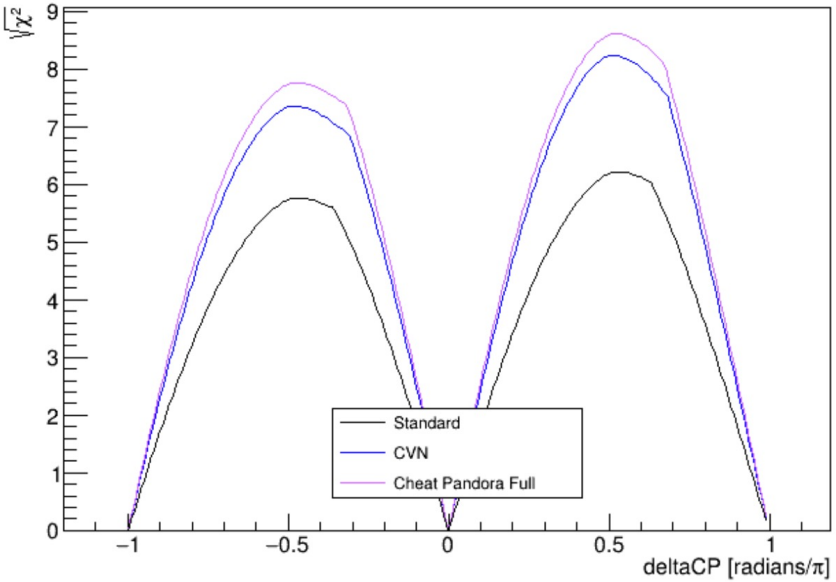
- Due to time constraints we'll focus on the Pandora stage, but I promise it's the most interesting

Cheating Test: Cheat Pandora Full

WARWICK
THE UNIVERSITY OF WARWICK

- **What happens if the pandora reconstruction is perfect?**
- This perfects the:
  - Clustering of all particles (with **EM showers folded back**)
  - 3D space points
  - Particle vertex placement
  - Particle track/shower ID
  - Neutrino vertex placement
  - Particle hierarchy construction

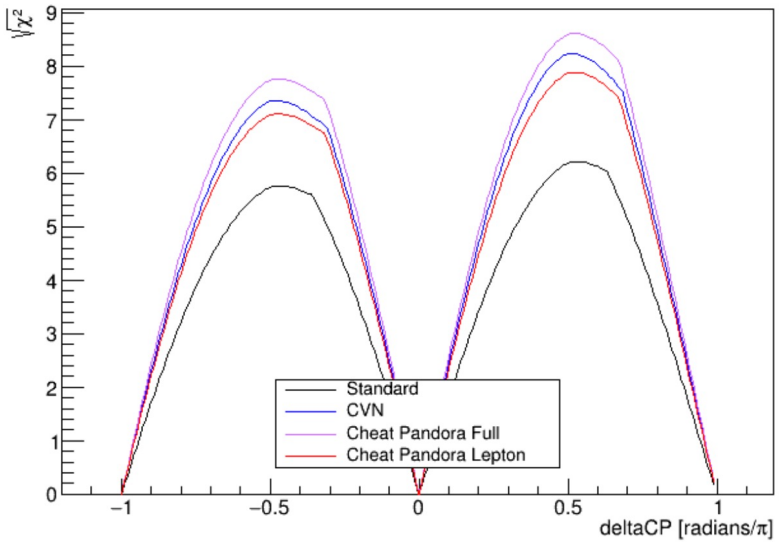
This cheating configuration results in large gains, motivating further Pandora cheating



Cheating Test: Cheat Pandora Lepton

WARWICK
THE UNIVERSITY OF WARWICK

- More specifically, **what happens if the pandora reconstruction is perfect for our leading lepton hierarchy** in CC nue and numu events?
- This means that the clustering, 3D space points, vertex placement, hierarchy construction and track/shower IDs are all correct for the particles in the **leading lepton hierarchy**
- The reconstruction, including the neutrino vertex placement, **will proceed as normal**

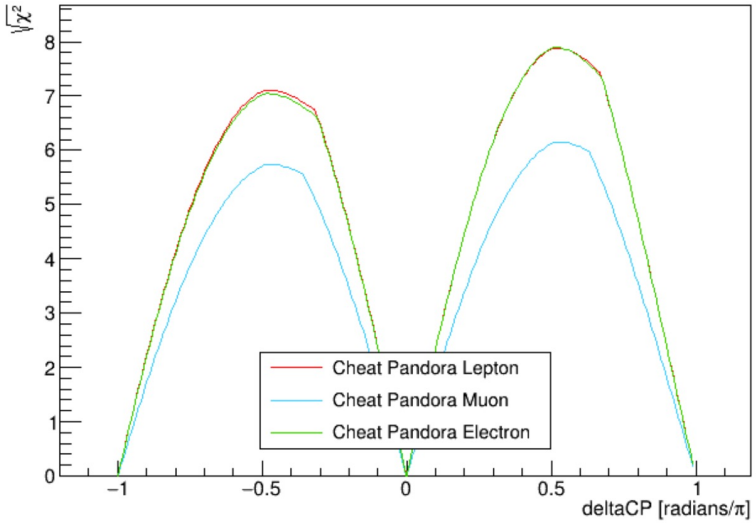


This suggests that we might be able to obtain significant improvements by focusing on the leading lepton improvements alone – very promising!

Cheating Test: Cheat Pandora Electron/Muon

WARWICK
THE UNIVERSITY OF WARWICK

- The nue appearance dominates the sensitivity, so we would expect the electron reconstruction to be more important here
- However, the pandizzle score is used to reject numu events and thus the muon reconstruction is also important
- So we can ask **how important is the reconstruction of each lepton?**

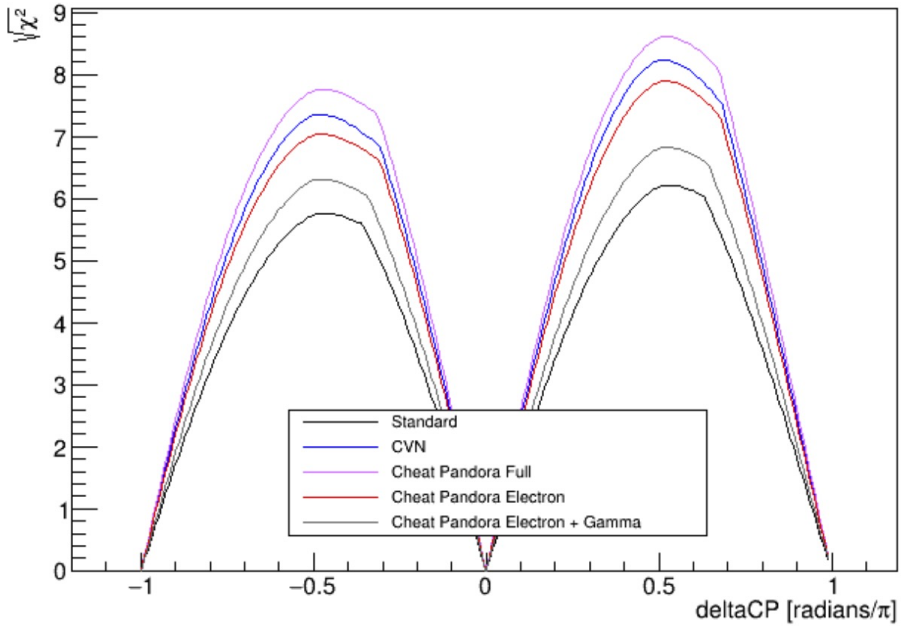


- This suggests that we might be able to obtain **significant improvements by focusing on the leading electron**
- **But,** perhaps pandrizzle is just learning the difference between well-reconstructed and poorly-reconstructed showers?

Cheating Test: Cheat Pandora Electron + Gamma

WARWICK
THE UNIVERSITY OF WARWICK

- Is pandrizzle really distinguishing between electrons and gammas or is pandrizzle simply learning the difference between well and less-well reconstructed showers?
- We can ask **what happens if we cheat the reconstruction of leading electron hierarchies in CC nue events and gammas in all events?**

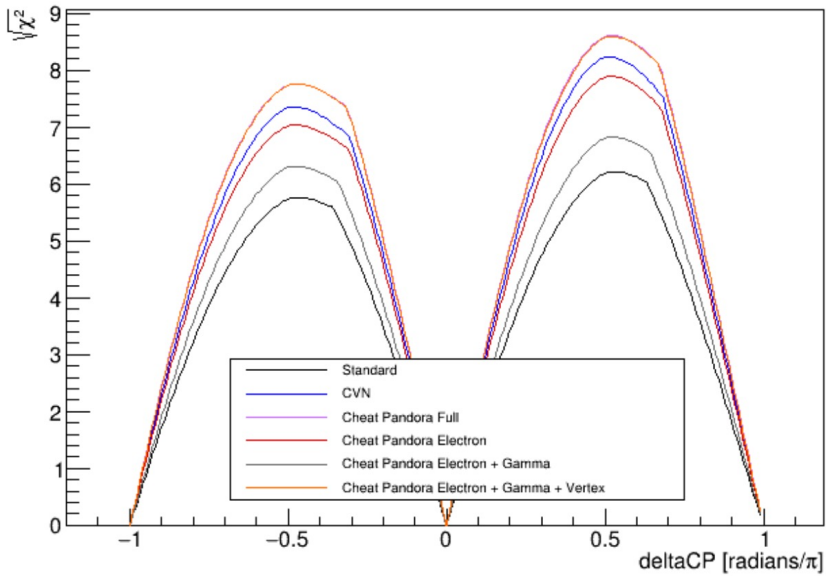


- As feared, it seems that pandrizzle was picking up on the topological differences of well reconstructed electron showers and impure gamma showers
- But why, if the electron appearance dominates the sensitivity, is this cheating configuration so far below the cheat pandora full configuration?

Cheating Test: Cheat Pandora Electron + Gamma + Nu Vertex

WARWICK
THE UNIVERSITY OF WARWICK

- One of two variables that are vital in electron/gamma separation is the gap between the nu vertex and the shower start position
- So, we can ask **what happens if we cheat the reconstruction of leading electron hierarchies in CC nue events and the primary gammas in all events as well as neutrino vertex placement?**



- As expected, the achieved sensitivity lies on top of that obtained in the cheat pandora full configuration
- So now, we must investigate the reasons for the vertex misplacement

Conclusion

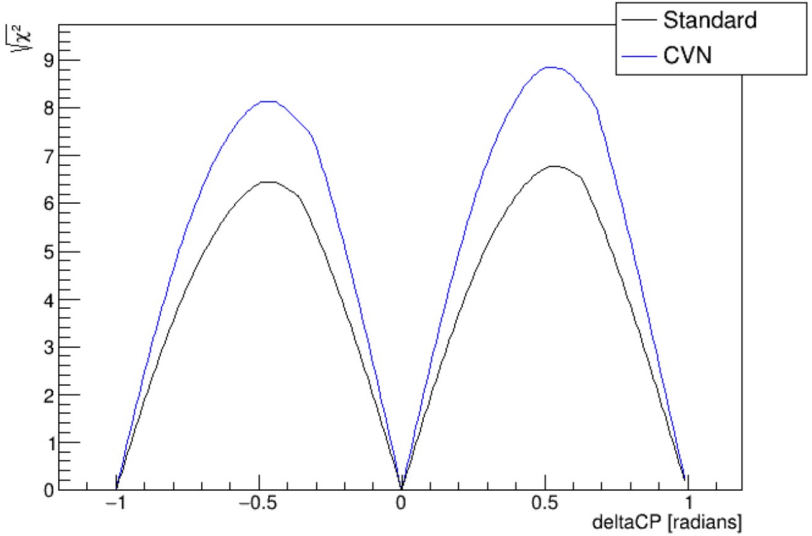
WARWICK
THE UNIVERSITY OF WARWICK

- To progress:
  - Probe deeper into the electron/gamma reconstruction i.e. what is important? The shower start? Completeness? Efficiency? Purity? vertex placement? etc…
  - Investigate the reasons for the vertex misplacement

BACKUP

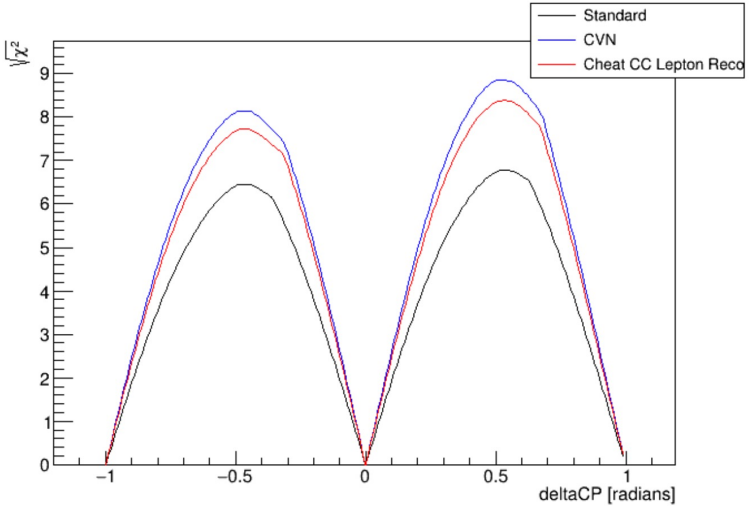
RHC DUNE CVN Comparison

	Anue Efficiency	Anue Purity	Anue BG	Anumu Efficiency	Anumu Purity	Anumu BG
CVN	86.6%	85.2%	99.6%	94.1%	96.8%	98.7%
'izzle selection'	53.4%	72.1%	99.5%	91.1%	89.3%	95.2%



RHC Cheating Test: Cheat Pandora Lepton

	Anue Efficiency	Anue Purity	Anue BG	Anumu Efficiency	Anumu Purity	Anumu BG
Standard	53.4%	72.1%	99.5%	91.1%	89.3%	95.2%
Cheat Lepton	67.4%	79.3%	99.5%	93.8%	91.3%	96.0%
CVN	86.6%	85.2%	99.6%	94.1%	96.8%	98.7%



Tuning the nue cuts

- A 2D histogram is filled with the pandizzle and pandrizzle scores of all events with a reconstructed inside the DUNE fiducial volume
- Apply a test pandizzle and pandrizzle cut position to obtain the selection sample at all deltaCP values and consequently the corresponding deltaCP sensitivity plot
- Investigate the entire pandizzle-pandrizzle phase space, choosing the cuts that optimise the **deltaCP sensitivity coverage**