Characterization of Hadronic Showers in the Belle II Electromagnetic Calorimeter

Emanuele Zanusso

Supervisor: Prof. **Stefano Spataro**

Dipartimento di Fisica
Università degli Studi di Torino

January 14, 2026

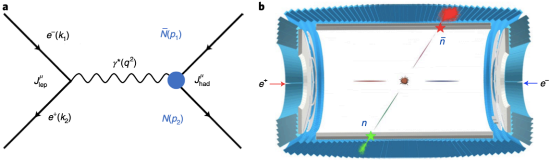
Outline

1. Anti-neutrons in physics experiments
2. Study via signal Monte Carlo sample
3. Study of Monte Carlo cocktail
4. Outlook

Anti-neutron in HEP experiments

The $\bar{n}$ plays a key role in several physics measurements, such as:

- The neutron e.m. form factor studies in $e^+ + e^- \rightarrow n + \bar{n}$ process



- Some decay channels studied at B-factories which involve $\bar{n}$
  1. The hyperons decay channel:

$$\bar{\Lambda}^0 \rightarrow \pi^0 + \bar{n}, \qquad \bar{\Sigma}^- \rightarrow \pi^- + \bar{n}, \qquad \bar{\Lambda}_c \rightarrow K^0_s + \pi^0 + \bar{n}$$

- Discrimination between other neutral particles (γ) and $\bar{n}$

Anti-neutrons in astrophysics

The $\bar{n}$ also plays a key role in several astrophysics measurements, such as:

- Studying $\bar{n}$ - anti-hyperon potential to improve the understanding of the equation of state of the neutron stars
- Investigating dark matter through anti-deuterons ($\bar{D}$) in cosmic rays, produced by dark matter annihilation or decay

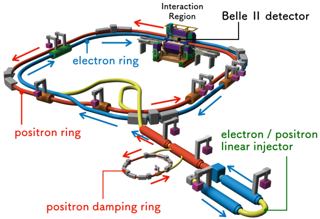
$$A_{d.m.} + B_{d.m.} \rightarrow hadrons\ (n, \bar{n}, p, \bar{p}\ \text{etc...})$$
$$X_{d.m.} \rightarrow \text{hadrons}\ (n, \bar{n}, p, \bar{p}\ \text{etc...})$$

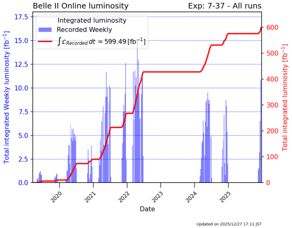
$\bar{D}$ is mainly produced through a coalescence mechanism $\bar{n} + \bar{p} \rightarrow \bar{D}$, where $\bar{p}$ and $\bar{n}$ are nearby in the phase-space

The Belle II experiment

SuperKEKB is an asymmetric $e^+ \, e^-$ collider (Tsukuba, Japan)

- 7 GeV electron beam (HER)
- 4 GeV positron beam (LER)
- Peak Luminosity $\sim 5.1 \times 10^{34} cm^{-2} s^{-1}$ 🏆
- Design Luminosity $\sim 8 \times 10^{35} cm^{-2} s^{-1}$
$\rightarrow$ x40 the Belle's one

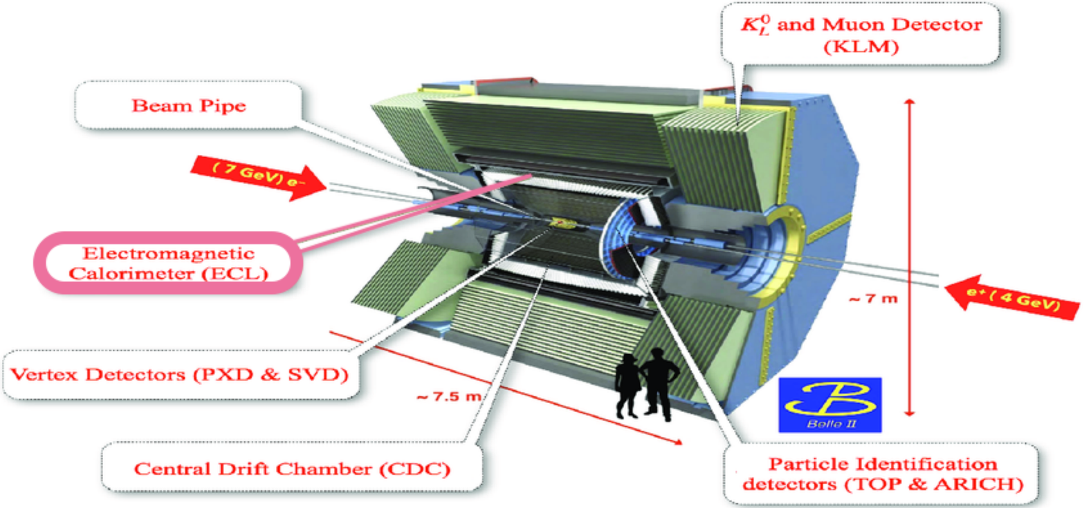




It operates mainly around $\Upsilon(4S)$ resonance (~ 10.58 GeV):

- This resonance decays almost exclusively into entangled couples of $B\,\bar{B} \rightarrow$ B-factory
- Several goals: flavour physics, BSM physics, heavy hadrons spectroscopy etc...

The Belle II experiment

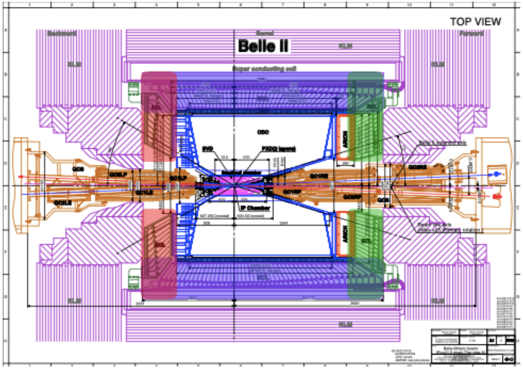


The Electromagnetic CaLorimeter

The ECL plays a central role in this thesis

- Array of **CsI(Tl)** crystals (8376 $6x6x30cm^3$ crystals in total)
- It covers barrel and end-cap regions ($12° \leq \theta \leq 155°$)
- Energy resolution of 4% @100 MeV and 1.6% @8 GeV

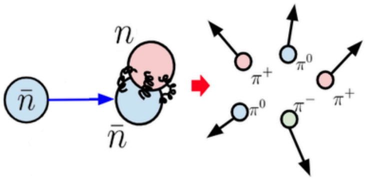
Barrel
FWD endcap
BWD endcap

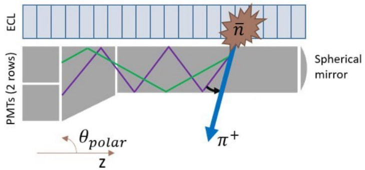


Anti-neutron interactions in physics

The $\bar{n}$ interacts primarily via strong nuclear force, producing hadronic showers
It can annihilate with nucleons in the ECL, producing light mesons (mainly pions)

- π^0 decays into $\gamma\gamma$, producing electromagnetic showers that are fully contained in the ECL
- $\pi^\pm$ undergo hadronic interactions, which are not fully contained in the ECL $\rightarrow$ both the forward (KLM) and backward (TOP) directions are involved



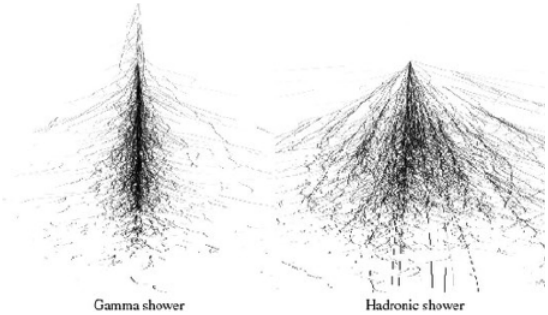


Electromagnetic and hadronic showers

Different processes occur for electromagnetic and hadronic showers:

1. Bremsstrahlung and pair production process (e^+, e^-, γ)
2. Interactions of hadrons with the material ($p, n, pions$, ...)

- About the 95% of the hadronic shower is contained within a cylinder of radius λ_{had} ($\sim$ 44.12 cm in CsI(Tl))
- About the 90% of the e.m. shower is contained within a cylinder of radius R_M ($\sim$ 3.6 cm in CsI(Tl))



The MANTRA project (PRIN2022)

Measuring **A**nti-**N**eutron: **T**agging and **R**econstruction **A**lgorithm

- A general method to measure the $E_{\bar{n}}$ up to 10 GeV, by combining information from:
  1. A detector with high time resolution (TOP)
  2. An electromagnetic calorimeter (**ECL**)
  3. A muon system (KLM)

- These features are common in modern general-purpose collider experiments such as **Belle II** and BESIII, which do not have a hadronic calorimeter
- For MANTRA project, only signals from ECL and TOP are taken into account. In this thesis only ECL signals are studied

Are $\bar{n}$ hadronic showers correctly simulated in the Belle II software?

The MANTRA project

Anti-neutrons cannot be reconstructed by sub-detectors.
The measurement of the energy is a two-step process:

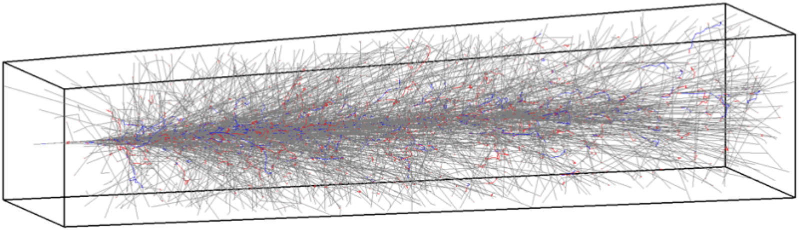
1. $\bar{n}$ identification via its induced ECL clusters (study of the shower shape)
2. Combine the signals from TOP and ECL to reconstruct the $\bar{n}$ energy, in cases of backscattering or pre-showering
   - If only π^0 are produced ($\sim 5\%$), the energy is all contained in the calorimeter, the shower is fully reconstructed
   - Otherwise ($\sim 95\%$): the products may escape the crystals
     $\rightarrow$ the goal is to complement the calorimeter information with that from the adjacent detectors

Preliminary concept

Several channels can be selected to look at $\bar{n}$ annihilations, such as:

- $e^+ + e^- \rightarrow p + \bar{n} + \pi^- + (\gamma_{ISR})$ (Mine)
- $\bar{\Lambda}_c \rightarrow K_s^0 + \pi^0 + \bar{n}$
- $\Lambda(\rightarrow\ p + \pi^-) + \bar{\Lambda}(\rightarrow \bar{n} + \pi^0)$

Several variables can be used to validate the showers shape for $\bar{n}$ identification, such as **Energy**, **Zernike Moments**, **Lateral momentum** and **Second moment**



Analysis outline

1. Study of the selected signal channel $e^+ + e^- \to p + \bar{n} + \pi^- + (\gamma_{ISR})$
   (a) Recoil identification from the system $p + \pi^-$ (with and without ISR)
   (b) Study of the kinematic recoil variables (momentum, angles, energy, etc...)
   (c) Study of the effect of 1C kinematic fit over the recoil mass
   (d) Study of ECL shower shape variables

2. Study of MC cocktail sample:
   (a) Recoil identification from the system $p + \pi^-$ (with and without ISR)
   (b) Study of the kinematic recoil variables (momentum, angles, energy, etc...)

3. Study of real data sample:
   (a) Recoil identification from the system $p + \pi^-$ (with and without ISR)
   (b) Constraint with 1C kinematic fit over the recoil mass
   (c) Examine Data/MC agreement in ECL cluster shapes from $\bar{n}$ channel

Analysis outline (1)

- The analyzed channel is:

$$e^+ + e^- \to p + \bar{n} + \pi^- + (\gamma_{ISR})$$

The reconstructed particles are (cuts and selections in the next slide):

(a) $p + \pi^- + (\gamma_{ISR})$ which compose the recoil system

(b) Neutral clusters associated to $\bar{n}$ candidates list used to compare its variables with those of the recoil

rowNo	decay tree	decay final state	iDcyTr	nEtr	nCEtr
1	$vpho \to \pi^- \bar{n} p$	$\pi^- \bar{n} p$	0	35291	35291
2	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p$	$\pi^- \bar{n} p\gamma^I\gamma^I$	2	22971	58262
3	$e^+e^- \to vpho\gamma^I, vpho \to \pi^- \bar{n} p$	$\pi^- \bar{n} p\gamma^I$	1	18735	76997
4	$vpho \to \pi^- \bar{n} p\gamma^F$	$\pi^- \bar{n} p\gamma^F$	3	10005	87002
5	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^I\gamma^F$	6	5274	92276
6	$e^+e^- \to vpho\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^F$	4	4621	96897
7	$vpho \to \pi^- \bar{n} p\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^F\gamma^F$	7	1503	98400
8	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^I\gamma^F\gamma^F$	8	700	99100
9	$e^+e^- \to vpho\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^F\gamma^F$	5	597	99697
10	$vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^F\gamma^F\gamma^F$	9	167	99864
11	$e^+e^- \to vpho\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^F\gamma^F\gamma^F$	12	63	99927
12	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^I\gamma^F\gamma^F\gamma^F$	10	61	99988
13	$e^+e^- \to vpho\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^F\gamma^F\gamma^F\gamma^F$	11	4	99992
14	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^I\gamma^F\gamma^F\gamma^F\gamma^F$	15	4	99996
15	$vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F$	14	2	99998
16	$vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F\gamma^F$	13	1	99999
17	$e^+e^- \to vpho\gamma^I\gamma^I, vpho \to \pi^- \bar{n} p\gamma^F\gamma^F\gamma^F\gamma^F\gamma^F$	$\pi^- \bar{n} p\gamma^I\gamma^I\gamma^F\gamma^F\gamma^F\gamma^F\gamma^F$	16	1	100000

- **100k events**. The reconstruction efficiency is:

$$\epsilon = \frac{n^\circ \text{ of reconstructed candidates}}{n^\circ \text{ of generated events}} \sim 22\%(18\%)$$

Applied selections and cuts

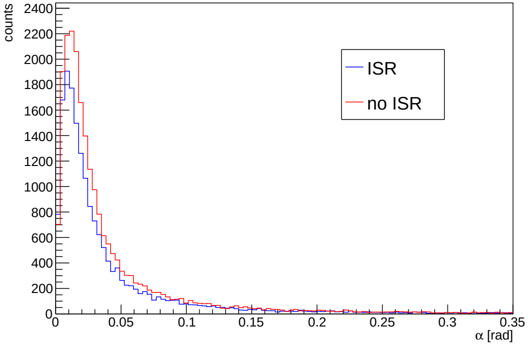
(a) **proton**: standard PID selection, with tracks required to originate from the IP

(b) **pion**: standard PID selection

(c) **anti-neutron**: neutral clusters from ECL

(d) 0 GeV < recoil mass < 2 GeV

(e) α < 0.35 rad (~ 20 deg)



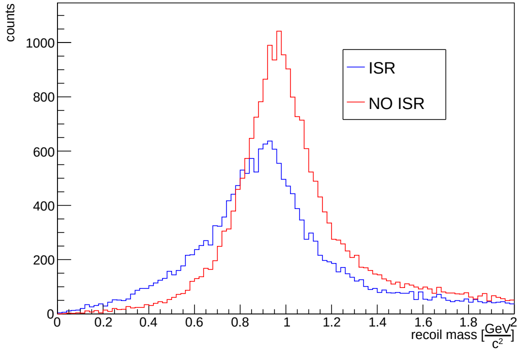
Where α is the angle between the recoil vector direction and the closest $\bar{n}$ cluster

The recoil mass

- The recoil mass is well reconstructed in both ISR (p, π^-, γ_{ISR}) and no ISR (p, π^-) cases

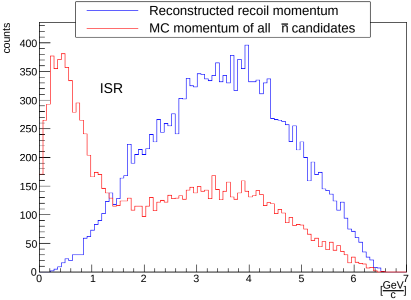
  $\rightarrow$ Variables associated with the $\bar{n}$ candidate clusters can be compared with the reconstructed recoil variables (p, θ)

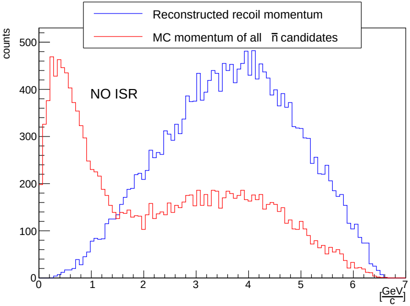
- Since there is more than one γ_{ISR} per event, the distribution shows a higher number of entries in the left tail



The recoil and the $\bar{n}$ momentum

The list of all $\bar{n}$ candidates shows two structures in the MC momentum distribution:

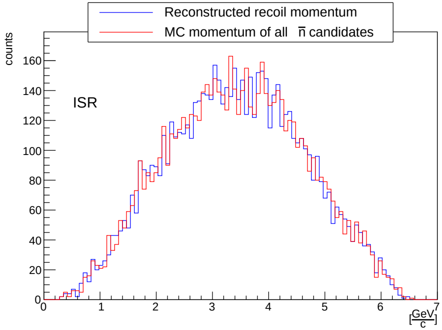


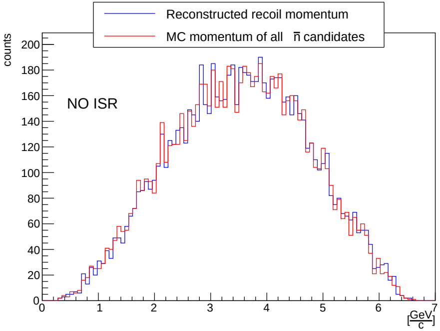


- The $\bar{n}$ candidate list shows a discrepancy with the recoil momentum
  (a) Could cluster split-off and/or pre-showering introduce the left peak?
  (b) Is there a problem in the MC association?

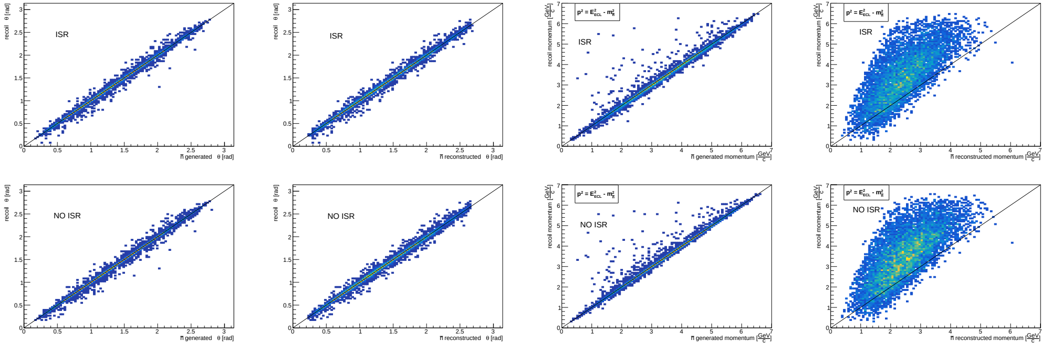
The recoil and the $\bar{n}$ momentum

- MC truth can be imposed by selecting only MC associated $\bar{n}$ candidates from the $\bar{n}$ list ($\bar{n}$ MC truth ID)
- A good correspondence can be observed between the two distributions, in both the ISR and NO ISR cases





$\bar{n}$ vs recoil vector correlation

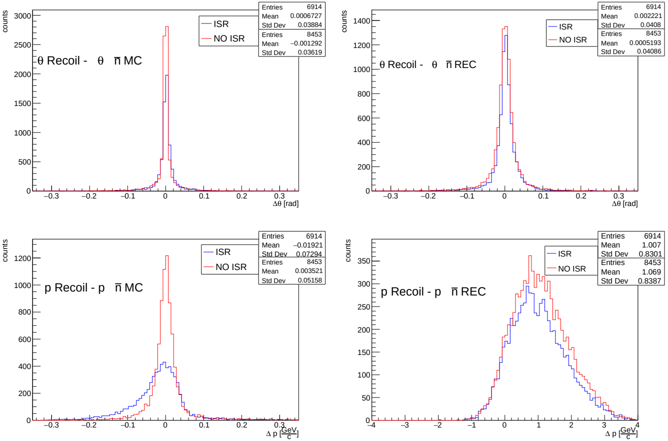


- Poor resolution is observed in reconstructed $\bar{n}$ momentum ($p^2_{rec} = E^2_{ECL} - m^2_{\bar{n}}$), in both the ISR and NO ISR cases

$\bar{n}$ vs recoil vector residuals

- Good correlation is observed at the generator level in both the momentum and θ distributions

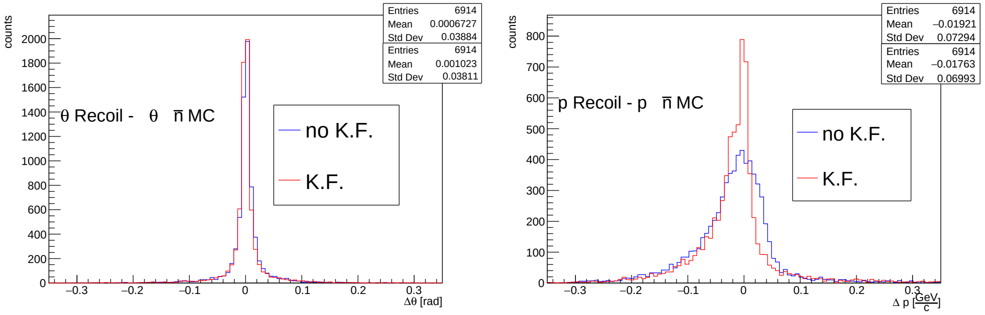
- Besides exhibiting poor resolution, the reconstructed momentum of the $\bar{n}$ is underestimated $\rightarrow$ missing energy in the shower



Kinematic Fit over the recoil mass

A 1C kinematic fit can possibly be used to add a constraint and improve the recoil resolution:

- No significant differences can be seen in θ_{recoil} vs MC $\theta_{\bar{n}}$
- An improvement can be observed in in p_{recoil} vs MC $p_{\bar{n}}$



$\bar{n}$ ECL cluster variables

The aim is to validate several $\bar{n}$ shower shapes variables, such as:

- **Energy** of the cluster, **E1E9** and **E9E21**

- **Lateral momentum** and **Second moment**:

$$C_{LM} = \frac{\sum_{i=2}^{n} E_i r_i^2}{E_0 r_0^2 + E_1 r_0^2 + \sum_{i=2}^{n} E_i r_i^2}, \text{ with } r_0 \simeq 5 \text{ cm}$$

$$C_{SM} = \frac{\sum_{i=0}^{n} E_i r_i^2}{\sum_{i=0}^{n} E_i}$$

10	11	12	13	14
25	2	3	4	15
24	9	1	5	16
23	8	7	6	17
22	21	20	19	18

$\bar{n}$ ECL cluster variables

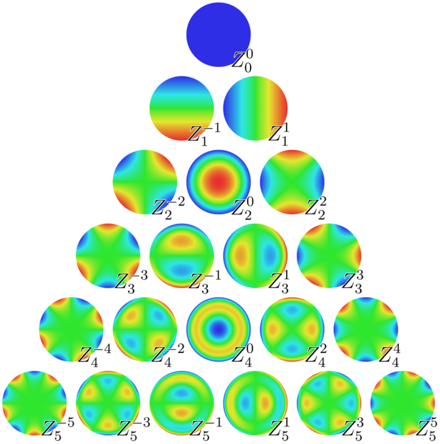
Zernike Moments:

$$Z_n^m(\rho,\phi) = R_n^m(\rho)\,\cos(m\,\phi)$$

$$Z_n^{-m}(\rho,\phi) = R_n^m(\rho)\,\sin(m\,\phi)$$

Where:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k\,(n-k)!}{k!\left(\frac{n+m}{2}-k\right)!\left(\frac{n-m}{2}-k\right)!}\,\rho^{n-2\,k}$$

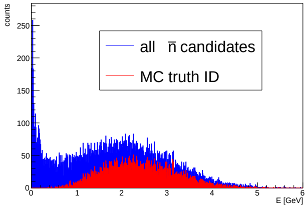


$\bar{n}$ ECL cluster variables

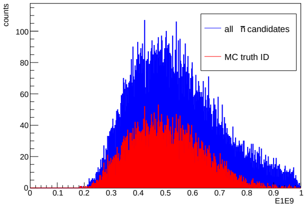
all $\bar{n}$ candidates

$\bar{n}$ MC truth ID

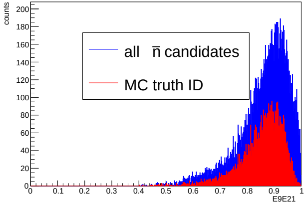
Cluster Energy



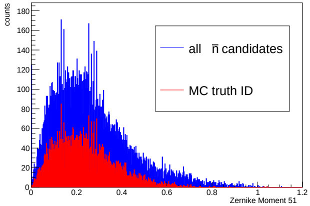
Cluster E1E9



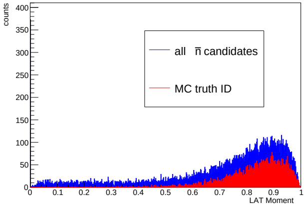
Cluster E9E21



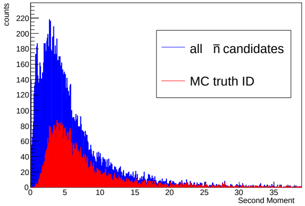
Cluster Zernike Moment 51



Cluster LAT Moment



Cluster Second Moment

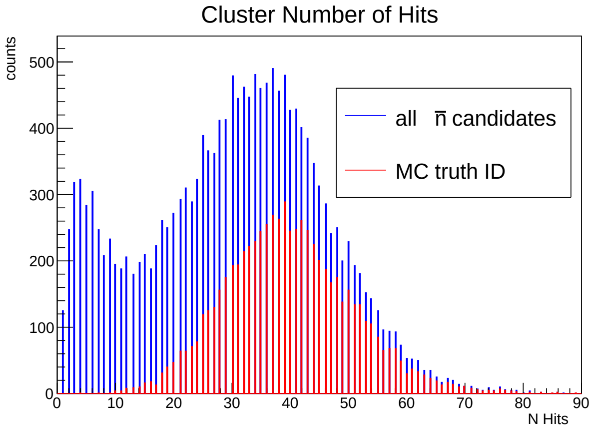


$\bar{n}$ ECL cluster variables

$\bar{n}$ clusters mainly involve 15 or more crystals

1. Could secondary clusters (split-off) be mis-identified as $\bar{n}$ primary cluster, during reconstruction?
2. Could it be a wrong association due to pre-showering in the TOP detector?

→ further studies can be addressed with a particle gun generator



Summary (1)

- Channel $e^+ + e^- \to p + \bar{n} + \pi^- + (\gamma_{ISR})$ has been studied and the recoil system is correctly reconstructed from the secondary background

- Reconstructed $\bar{n}$ variables, such as momentum and shower-shape variables, should be further investigated:
  (a) Cluster split-offs with the ECL and pre-showering occurring within the TOP may affect the shower shape distributions
  (b) A wrong Monte Carlo association may be present as a consequence

  $\to$ This could be studied by testing the ECL response with a $\bar{n}$ particle gun generator

- 1C kinematic fit can be possibly adopted during Data/MC comparison, in order to improve the recoil momentum resolution

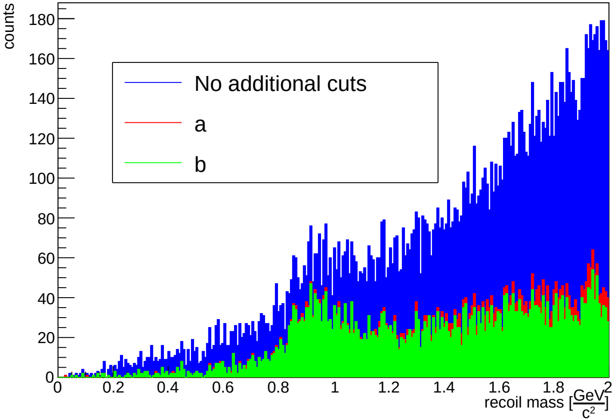
Analysis outline (2)

- Study of cocktail using the following MC sample:
  $q\bar{q}$ cocktail with 341 *M* events (Luminosity: 215 fb^{-1}) generated with Pythia8
- Only the *p* and π^- are used to build the recoil vector (NO ISR for the moment) $\rightarrow$ Can the signal be correctly reconstructed among the other background channels?
- Following selections have been applied:
  1. PID selection on *p* and π^- from the MC truth
  2. 0 GeV < recoil mass < 2 GeV
  3. The best candidate is selected via α (closest candidate)
- Same strategy as before:
  (a) Identify the signal peak near the $\bar{n}$ mass (~ 0.939 GeV) adding further cuts
  (b) Study the shower shape variables in a selected recoil mass region, replacing the MC truth selections with data real ones
  (c) Compare it with data (Data/MC agreement)

Recoil mass distribution - MC selections

- The following cuts are applied to enhance the signal:
  (a) No additional charged tracks in the event
  (b) No additional charged tracks in the event and alpha<0.35 rad (angular cut on the closest candidate)

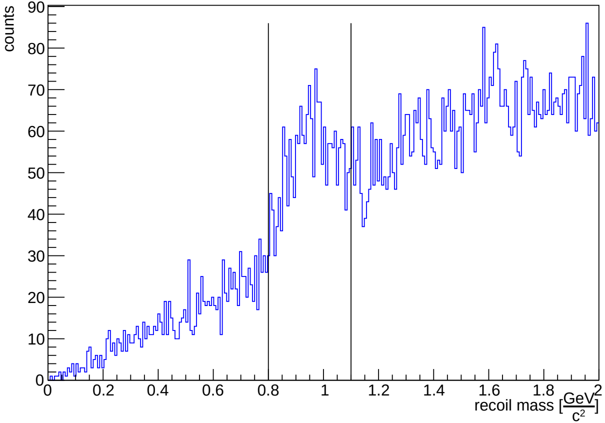
- A peak can be observed above the $\bar{n}$ mass $\rightarrow$ the signal is highlighted in the MC cocktail



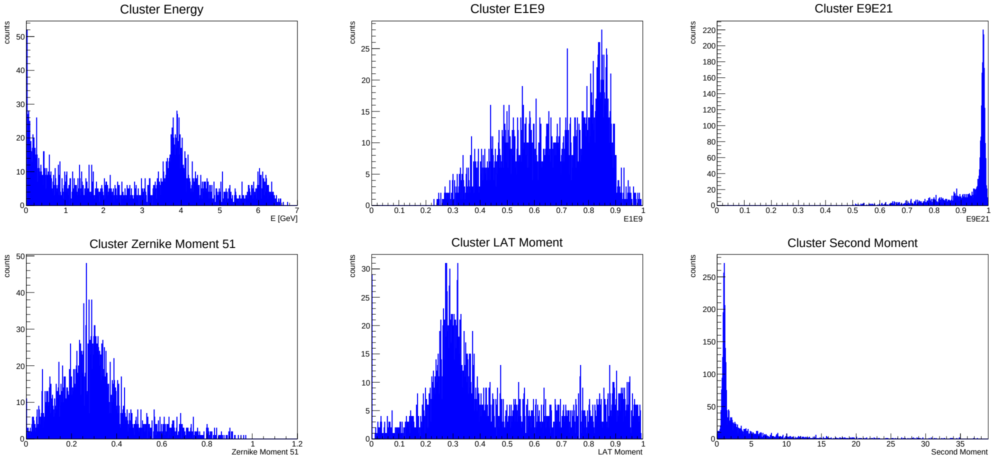
Recoil mass distribution - Real cuts

- ”Real cuts” can be applied instead of MC truth selections, such as:
  (a) **proton**: standard PID selection, with tracks required to originate from the IP
  (b) **pion**: standard PID selection

- To select the signal, a recoil mass cut in the range (0.8-1.1) GeV can be applied to study the shower shapes variables



$\bar{n}$ ECL cluster variables

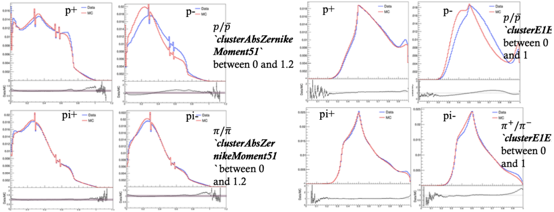


Background must be further suppressed → sideband method for background subtraction

Summary (2)

- Study of $q\bar{q}$ cocktail has been performed
- The recoil mass is correctly reconstructed using both Monte Carlo based selections and real cuts, in the presence of other background channels
- A further selection on the recoil mass has been applied in order to study the shower shape variables and then validate it in a Data/MC agreement

Analysis of a $\Lambda \to p + \pi^{-}$ $(\bar{\Lambda} \to \bar{p} + \pi^{+})$ sample shows that:



Could the Data/MC agreement observed for $\bar{p}$ also be expected for $\bar{n}$?

Outlook

The following topics will be faced from today:

1. Study the cluster split-off and pre-showering effects using a particle gun, evaluating whether they can affect the Monte Carlo association

2. Study the complete cocktail production and attempt to further suppress the background

3. Study of the Data/MC agreement of the shower-shape variables using the evaluated real selections

Thank you for your attention

Emanuele Zanusso

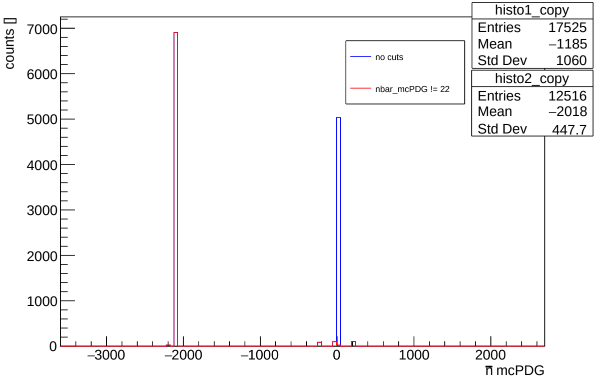
Supervisor: Prof. **Stefano Spataro**

Dipartimento di Fisica
Università degli Studi di Torino

January 14, 2026

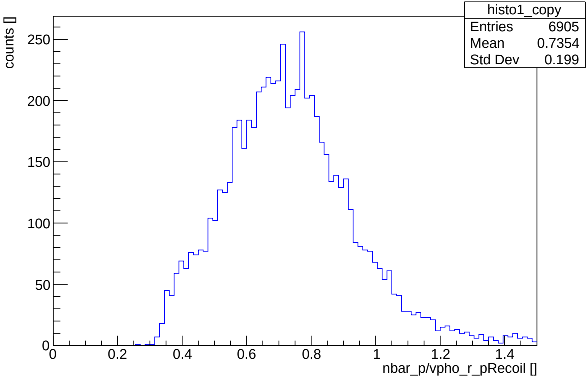
$\bar{n}$ mcPDG I

γ's are mis-identified as $\bar{n}$ in reconstruction:



$p_{\bar{n}}$/pRecoil I

$\bar{n}$ is underrated in the most of cases (annihilation process + loss of energy)



Recommended PID selections I

PID probability defined as:

protonID [source]

proton identification probability defined as $\mathcal{L}_p/(\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d)$, using info from all available detectors

pionID [source]

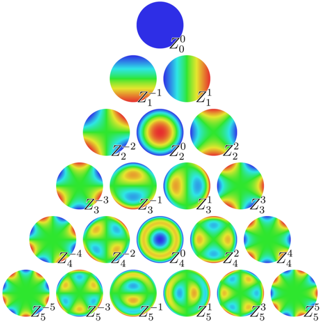
pion identification probability defined as $\mathcal{L}_\pi/(\mathcal{L}_e + \mathcal{L}_\mu + \mathcal{L}_\pi + \mathcal{L}_K + \mathcal{L}_p + \mathcal{L}_d)$, using info from all available detectors

```
ma.fillParticleList(f"p+:{lista}", "protonID > 0.9 and dr < 1 and abs(dz) < 3", path=main)
ma.fillParticleList(f"pi-:{lista}", "pionID > 0.1", path=main)
ma.fillParticleList(f"anti-n0:{lista}", "", path=main)

ma.reconstructDecay(f"vpho:list_rec -> p+:{lista} pi-:{lista}",cut="thetaInECLAcceptance",path=main
ma.reconstructDecay(f"vpho:gen -> vpho:list_rec anti-n0:{lista}",cut="",path=main)
```

Zernike Moments I

- Zernike polynomials are widely used as basis functions of image moments



Ci sono polinomi di Zernike pari e dispari. Quelli pari sono definiti come:

$$Z_n^m(\rho,\varphi) = R_n^m(\rho)\,\cos(m\,\varphi)$$

e quelli dispari come:

$$Z_n^{-m}(\rho,\varphi) = R_n^m(\rho)\,\sin(m\,\varphi)$$

dove m e n sono numeri interi non negativi con $n \geq m$, ϕ è l'angolo azimutale, ρ è la distanza radiale $0 \leq \rho \leq 1$ e $R^m{}_n$ sono i polinomi radiali definiti di seguito. I polinomi di Zernike hanno la proprietà di essere limitati a un intervallo da -1 a +1, cioè $|Z_n^m(\rho,\varphi)| \leq 1$. I polinomi radiali $R^m{}_n$ sono definiti come:

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k\,(n-k)!}{k!\,\left(\frac{n+m}{2}-k\right)!\,\left(\frac{n-m}{2}-k\right)!}\;\rho^{n-2\,k}$$

per $n - m$ pari e sono identicamente zero per $n - m$ dispari.

cluster LAT & cluster Second Moment I