

Algorytm selekcjonujący przypadki z wymianą prądów neutralnych i naładowanych w eksperymencie MINOS

Tomasz Früboes

20 marca 2006

Eksperyment MINOS zaprojektowany został z myślą o badaniu fenomenu oscylacji neutrin. Oddziaływanie neutrina z materią może zachodzić zarówno przez wymianę nośników naładowanych ($W^\pm$) jak i neutralnych (Z^0). Określenie jakiego typu oddziaływanie zaszło w detektorze jest podstawą wielu metod badania zjawiska oscylacji neutrin. Stworzony i opisany został prosty algorytm selekcjonujący przypadki z wymianą bozonów $W^\pm$ lub Z^0 w dalekim detektorze eksperymentu MINOS.

1 Wstęp

Eksperyment MINOS (Main Injector Neutrino Oscillation Search) jest nowym eksperymentem działającym od 2005 r. w ośrodku Fermilab (Fermi National Accelerator Laboratory) pod Chicago ([1]). Jego zadaniem jest zbadanie zjawiska oscylacji zapachu neutrin odkrytego w eksperymencie Super-Kamiokande w 1998 roku.

Wykorzystywana wiązka neutrin, wytwarzana w projekcie Neutrinos at Main Injector (NuMI), składa się głównie z neutrin mionowych. Badane są one przez dwa detektory - bliski, znajdujący się w Fermilabie (odległy o kilkaset metrów od miejsca produkcji neutrin), oraz daleki, znajdujący się w kopalni w północnej Minnesocie (odległy o 735 km od Fermilabu).

Zadaniem bliskiego detektora jest określenie strumienia i rozkładu energii neutrin zanim efekty oscylacyjne będą widoczne. Dzięki pomiarom dokonowanym w bliskim detektorze można przewidzieć te wielkości (strumień i rozkład energii) dla detektora dalekiego. Pomiary w dalekim detektorze pozwolą wyznaczyć ubytek neutrin mionowych powstały na skutek oscylacji. Pozwoli to określić takie parametry modelu oscylacji jak różnice mas neutrin oraz kąty mieszania.

Neutrino mogą oddziaływać z materią poprzez wymianę neutralnego bozonu Z^0 lub jednego z dwóch naładowanych bozonów $W^\pm$ ([2]). Pierwszy typ oddziaływań nazywany jest również oddziaływaniem z wymianą prądów neutralnych (Neutral Current - NC), drugi zaś oddziaływaniem z wymianą prądów naładowanych (Charged Current - CC). W wyniku oddziaływania typu CC neutrino zamienia się w masywny lepton należący do tej samej rodziny ($\nu_e \rightarrow e$, $\nu_\mu \rightarrow \mu$, $\nu_\tau \rightarrow \tau$). Tak więc w wyniku oddziaływania typu CC neutrina mionowego obserwujemy mion i kaskadę hadronową, zaś w wyniku od-

działywania NC obserwowana jest tylko kaskada hadronowa (neutrino ulatuje poza detektor unosząc część energii).

Zliczanie liczby przypadków CC i/lub NC jest podstawą wielu testów mających zostać użytych przez eksperyment MINOS do badania oscylacji neutrin. Jednym z nich jest test porównujący stosunki ilości oddziaływań NC i CC w bliskim i dalekim detektorze ([3]). W przypadkach oddziaływania neutrin mionowych typu CC zawsze pojawia się mion. Wysoka zdolność penetrująca mionów powoduje, że długość przypadków CC będzie zwykle większa od długości przypadków NC (w których mion się nie pojawia). Umożliwia to sprawne odróżnianie przypadków CC od NC. Przypadki oddziaływania neutrin taonowych będą przeważnie przypadkami krótkimi (niezależnie od tego, czy oddziaływanie było typu NC czy CC), gdyż powstały taon szybko rozpada się na lżejsze cząstki. Stąd też przypadki oddziaływania neutrin taonowych będą zwykle kwalifikowane jako przypadki NC. Porównanie stosunków liczby przypadków $\frac{N_{CC}}{N_{NC}}$ dla bliskiego i dalekiego detektora może być zatem testem na wystąpienie oscylacji neutrin.

Opisywany algorytm powstał w ramach ćwiczenia z III pracowni fizyki cząstek elementarnych. Celem tego ćwiczenia było zapoznanie się z pewnymi aspektami fizyki neutrin, eksperymentem MINOS oraz pogłębienie znajomości środowiska analizy danych ROOT.

2 Algorytm selekcjonujący

2.1 Opis algorytmu

Algorytm został opracowany i przetestowany na próbce 134 tysięcy zdarzeń wysymulowanych dla dalekiego detektora. Wykorzystano gotowe pliki z danymi

w formacie TTree programu ROOT ([4]) wyprodukowane na potrzeby eksperymentu MINOS (pliki f210000XX_0000_L010185.sntp.R1_18_2.root, gdzie XX zmienia się od 01 do 30).

Algorytm selekcji wykorzystuje szereg prostych funkcji (ozn. `nazwaFunkcji()`), a także zmiennych niskiego poziomu (ozn. `nazwa.zmiennej`, [4, 5]):

- `evthdr.ph.sigcor` - wysokość sygnału dla całego zdarzenia (jest to wielkość proporcjonalna do energii widzianej w detektorze; w tej i następnych zmiennych PH odnosi się do wysokości zarejestrowanego sygnału - ang. Pulse Height)
- `evthdr.ntrack` - liczba torów zrekonstruowanych w zdarzeniu
- `evthdr.nshower` - liczba kaskad zrekonstruowanych w zdarzeniu
- `evthdr.planeall.n` - ilość płaszczyzn, w których zarejestrowano sygnał (długość zdarzenia)
- `evthdr.plane.n` - ilość płaszczyzn, w których zarejestrowano sygnał z wysokością pulsu co najmniej 2 fotoelektrony (długość zdarzenia)
- `PHfractionInLast6planes()` - stosunek wysokości sygnału w ostatnich sześciu płaszczyznach do całkowitej wysokości sygnału
- `trackPHfraction()` - stosunek całkowitej wysokości sygnału wiodącego toru, do całkowitej wysokości sygnału (całkowitej czyli wysumowanej po wszystkich paskach scyntylatora na podstawie których zrekonstruowany został dany tor; wiodący tor to najbardziej energetyczny tor zrekonstruowany w zdarzeniu)
- `trackLength()` - zasięg obliczony dla wiodącego toru (tj. ilość materiału wyrażona w $\frac{\text{g}}{\text{cm}^2}$ jaką pokonała cząstka od wierzchołka do końca toru)
- `trkMaxPh()` - całkowita wysokość pulsu wiodącego toru
- `shwMaxPh()` - całkowita wysokość pulsu wiodącej kaskady
- `averagePHperStrip()` - średnia wysokość sygnału w pasku scyntylatora
- `averagePHperPlane()` - średnia wysokość sygnału w płaszczyźnie
- `diffTrackShowerLengthsU()`, `diffTrackShowerLengthsV()` - różnica między długością wiodącego toru a długością wiodącej kaskady (w płaszczyźnie U lub V)

- `trkMomQp()` - stosunek ładunku cząstki (której tor został zrekonstruowany) do jej pędu
- `getEnergy()` - energia widoczna w zdarzeniu. Wartość energii jest proporcjonalna do wartości zmiennej `evthdr.ph.sigcor`. Współczynnik proporcjonalności obliczony został na podstawie proporcjonalności między zmiennymi `shw.ph.sigcor` a `shw.ph.gev` (wysokość pulsu przeliczona na GeV). Zastosowanie dedykowanej funkcji liczącej energię jest konieczne - zmienna `evthdr.ph.gev` obecnie nie jest wypełniana przez oprogramowanie odpowiedzialne za rekonstrukcje przypadków.

Algorytm składa się z serii warunków przetwarzanych w ustalonej kolejności. Każdy warunek składa się z jednego lub większej ilości cięć, tj. żądań aby dana zmienna lub funkcja przyjmowała określone wartości. W przypadku spełnienia któregoś z warunków (a więc spełnienia wszystkich cięć które zawiera dany warunek) zwrócona zostaje jedna z trzech wartości - CC, NC lub UNKNOWN. Pierwsze dwie z nich odpowiadają stwierdzeniu, że oddziaływanie było określonego typu. Wartość UNKNOWN odpowiada stwierdzeniu, że algorytm nie potrafi rozstrzygnąć z jakiego typu oddziaływaniem mieliśmy doczynienia. Warunki ułożone są hierarchicznie, tak więc zwrócenie określonej wartości (CC/NC/UNKNOWN) powoduje przerwanie działania algorytmu (dalsze warunki nie są przetwarzane).

Rys. 1 przedstawia rozkład energii zmierzonej w dalekim detektorze dla przypadków z zerem zrekonstruowanych torów dla oddziaływań typu CC i NC. Takie przypadki określane są przez algorytm zawsze jako typu NC.

Rys. 2 przedstawia rozkład energii zmierzonej w detektorze dla wszystkich przypadków. Zaznaczono trzy zakresy energii, które odpowiadają trzem różnym sposobom selekcjonowania danego przypadku. Poniżej, dla każdego zakresu energii, opisano warunki wykorzystywane w selekcji przypadków. Numeracja warunków odpowiada położeniu w hierarchii (im liczba wyższa tym mniej przypadków "dociera" do danego warunku).

I. Niskie energie - energia zmierzona w detektorze poniżej około 1.6 GeV (`evthdr.ph.sigcor` poniżej 20000)

1. `evthdr.plane.n >= 15` → CC (rys. 3.1)
2. `evthdr.plane.n <= 5` → NC (rys. 3.2)
3. (`evthdr.plane.n = (6 lub 7)`) i (`evthdr.planeall.n >= 10`) i (`PHfractionInLast6planes() > 0.6`) → NC (rys. 3.3-3.5)
4. (`PHfractionInLast6planes() < 0.6`) i (`trkMomQp() < 0.7`) i (`trackPHfraction() > 0.35`) → CC (rys. 3.6-3.8)

5. Pozostałe $\rightarrow$ UNKNOWN

II. Średnie energie - energia zmierzona w detektorze od około 1.6 GeV do 5.7 GeV (evthdr.ph.sigcor od 20000 do 70000)

1. evthdr.plane.n > 30 $\rightarrow$ CC (rys. 4.1)
2. (PHfractionInLast6planes() < 0.23) i (trackPHfraction() > 0.27) i (evthdr.plane.n > 15) $\rightarrow$ CC (rys. 4.2-4.4)
3. (PHfractionInLast6planes() > 0.22) i (PHfractionInLast6planes() < 0.34) i (shwMaxPh() < 40000) i (trkMaxPh() > 7000) $\rightarrow$ CC (rys. 4.5-4.7)
4. evthdr.plane.n < 8 $\rightarrow$ NC (rys. 5.8)
5. (evthdr.plane.n > 9) i (trkMaxPh() > 7000) i (trkMaxPh() < 16000) i (trackPHfraction() > 0.22) $\rightarrow$ UNKNOWN (rys. 5.9 - 5.11)
6. (averagePHperStrip() < 600) i (diffTrackShowerLengthsU() < 0 lub diffTrackShowerLengthsV() < 0) $\rightarrow$ NC (rys. 5.12 - 5.13)
7. Pozostałe $\rightarrow$ UNKNOWN

III. Wysokie energie - energia zmierzona w detektorze powyżej około 5.7 GeV (evthdr.ph.sigcor powyżej 70000)

1. trackLength() >= 2000 $\rightarrow$ CC (rys. 6.1)
2. evthdr.plane.n >= 50 $\rightarrow$ CC (rys. 6.2)
3. diffTrackShowerLengthsU() > 25 $\rightarrow$ CC (rys. 6.3)
4. diffTrackShowerLengthsV() > 25 $\rightarrow$ CC (rys. 6.4)
5. (getEnergy() < 10) i (averagePHperPlane() < 3000) i (PHfractionInLast6planes() < 0.15) i (evthdr.plane.n > 21) $\rightarrow$ CC (rys. 6.5-6.8)
6. (getEnergy() < 10) i (PHfractionInLast6planes() > 0.12) $\rightarrow$ NC (rys. 6.9-6.10)
7. Pozostałe $\rightarrow$ UNKNOWN

Do selekcji próbowano wykorzystać zmienne określające rozwartość kaskady hadronowej w płaszczyznach U lub V. Stworzono algorytm określający kąt rozwarcia kaskady. Zmienna ta niestety nie separowała przypadków CC i NC, nie została wykorzystana więc w algorytmie selekcji. W procesie tworzenia algorytmu natrafiono na ciekawy problem - okazało się, że wierzchołek kaskady (określany przez oprogramowanie rekonstrukcji przypadków) w większości przypadków umieszczany jest w środku kaskady. W takich przypadkach algorytm określał kąt rozwarcia na bliski 2π . Rozwiązaniem problemu było określenie współrzędnej z (czyli współrzędnej wzdłuż osi wiązki) wierzchołka kaskady równej współrzędnej z najwcześniej zapalonego paska.

2.2 Wydajność algorytmu

Podstawowymi parametrami określającymi dobroć algorytmu są czystość Pur i efektywność Eff , zdefiniowane jako

$$Pur = \frac{N_{true}}{N_{true} + N_{false}}$$
$$Eff = \frac{N_{true}}{N_{all}}$$

gdzie N_{true} oznacza liczbę poprawnie wyselekcjonowanych przypadków danego typu - NC lub CC, N_{false} liczbę przypadków określonych nieprawidłowo, N_{all} oznacza liczbę wszystkich przypadków danego typu.

2.2.1 Prosty algorytm selekcji

Jako “wprawkę” zaimplementowano wyjątkowo prosty algorytm selekcji przypadków, w którym selekcja odbywa się tylko na podstawie liczby zrekonstruowanych torów. Przypadki z zerem zrekonstruowanych torów określone były jako przypadki typu NC. Gdy zrekonstruowany został jeden lub więcej tor, przypadek zostawał określony jako typu CC. Dodatkowo jeśli żaden tor w przypadku nie spełniał wymagań co do jakości (zmienna `trk.fit.pass` równa `false`) przyjmowano, że zdarzenie było typu NC. Czystość i efektywność (w funkcji energii mierzonej w detektorze) dla tego algorytmu przedstawiają rysunki 7 i 8.

Opisany powyżej prosty algorytm dobrze działa dla przypadków NC dla energii do około 2 GeV (czystość na poziomie 0.6-0.8, efektywność od 0.5 do 1). Nie jest to dziwne, gdyż dla takich energii dominują przypadki typu NC.

Dla energii większych niż 3 GeV dla przypadków NC prosty algorytm jest wysoce niezadowolający. Czystość i efektywność osiągają praktycznie stałą wartość równą odpowiednio 0.25 i 0.15.

Dla przypadków CC sytuacja jest odwrotna niż dla przypadków NC - dla energii powyżej 2 GeV prosty algorytm osiąga czystość i efektywność rzędu 90%. Wynika to z faktu, że dla przypadków typu CC wraz ze wzrostem energii częściej pojawia się daleko lecący mion, którego tor łatwo zrekonstruować.

2.2.2 Wyniki dla pełnego algorytmu selekcji

Efektywność (w funkcji energii) algorytmu selekcjonującego przypadki CC i NC przedstawia rysunek 9, czystość rysunek 10. Efektywność dla przypadków CC ma wartość około 0 dla najniższych energii (ze spadkiem energii widzianej w detektorze znacznie spada liczba przypadków typu CC, rys. 2). Efektywność wraz ze wzrostem energii stopniowo zwiększa się osiągając wartość około 0.85 dla energii 5 GeV i wyższych. Czystość dla przypadków CC w całym zakresie energii utrzymuje się na poziomie 0.95.

Dla przypadków NC algorytm sprawuje się gorzej. Przypadki typu NC są znacznie trudniejsze do selekcjonowania niż przypadki CC, ze względu na brak wyraźnej sygnatury, takiej jak pojawienie się mionu dla przypadków CC. Efektywność selekcji przypadków NC ma wartość około 1 dla najniższych energii i stopniowo spada by osiągnąć minimum około wartości 0.2 dla energii rzędu 5 GeV. Efektywność następnie gwałtownie wzrasta do poziomu 0.6-0.8 dla energii 6 GeV i wyższych. Skok ten odpowiada przejściu z rejonu mierzonych średnich energii (od 1.6 do 5.7 GeV) do rejonu wysokich energii (powyżej 5.7 GeV), czyli ze zmianą sposobu w jaki selekcjonowane są przypadki.

Czystość selekcji przypadków NC ma wartość około 0.8 dla najniższych energii i wraz ze wzrostem energii spada do wartości rzędu 0.5 dla energii 3 GeV i wyższych. Jest to wartość mała i oznacza, że w tym zakresie energii przypadek określony jako NC z prawdopodobieństwem 50% był rzeczywiście przypadkiem NC (i z takim samym prawdopodobieństwem był to przypadek CC). Z drugiej strony dla energii powyżej 3 GeV stosunek liczby przypadków NC i CC wynosi około 1:10 (rys. 11), więc uzyskanie czystości 50% oznacza polepszenie tego stosunku 10-krotnie.

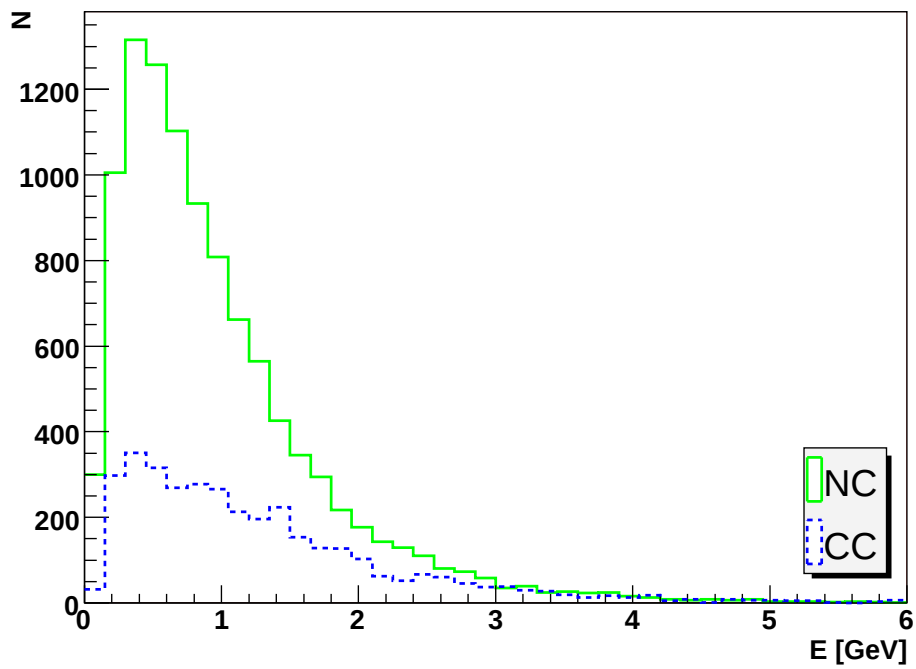
Duża część przypadków określana jest przez algorytm jako nieznanego typu (UNKNOWN). Stosunek liczby takich przypadków do liczby wszystkich przypadków przedstawia rysunek 12. Jest to wartość niemała - od 20% do 50% dla zmierzonej energii do 5 GeV, i od 5% do 10% dla energii powyżej 5 GeV.

3 Podsumowanie

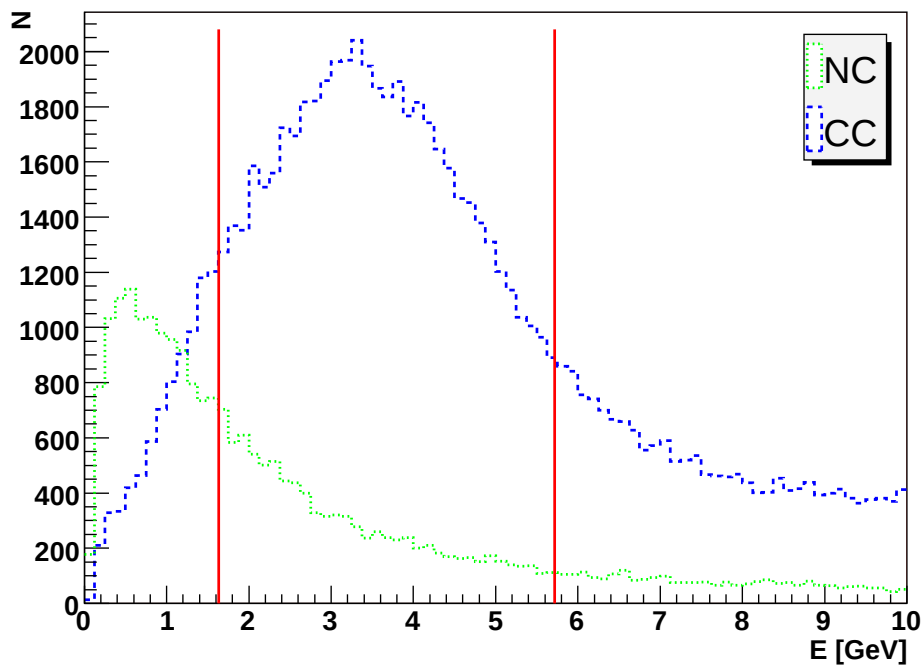
Stworzony został algorytm selekcjonujący przypadki oddziaływań neutrin mionowych z wymianą prądów neutralnych (NC) i naładowanych (CC) w dalekim detektorze eksperymentu MINOS. Charakteryzuje się on wysoką czystością i dobrą efektywnością dla przypadków CC w całym zakresie energii. Czystość i efektywność dla przypadków NC są zadowalające w zakresie energii do 2 GeV. W pozostałym zakresie energii algorytm dla przypadków NC wymaga dalszych studiów.

Literatura

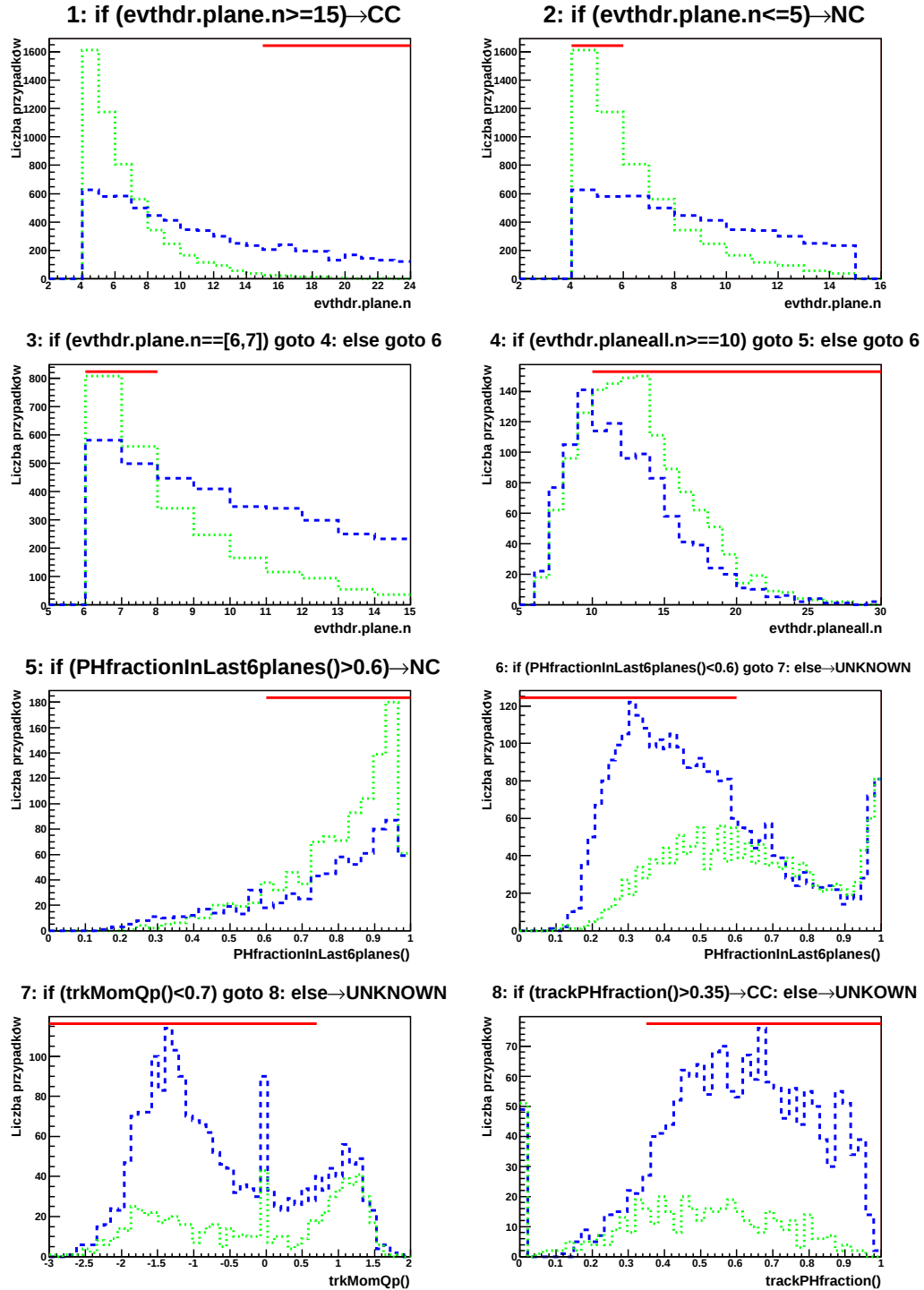
- [1] The Minos Collaboration: *The Minos Detectors Technical Design Report*, NuMI-L-337
- [2] Donald H. Perkins: *Wstęp do fizyki wysokich energii*, Wydawnictwo Naukowe PWN 2004
- [3] David S. Ayres: *The MINOS long-baseline experiment at Fermilab*, NuMI-NOTE-COMP-0227
- [4] Roy Lee: *Users Guide to the CandEventSR Reconstruction Tree Version 2.00*, NuMI-NOTE-COMP-0860
- [5] *Dokumentacja oprogramowania rekonstrukcji eksperymentu MINOS*,
http://www-numi.fnal.gov/offline_software/srt_public_context/WebDocs/doxygen/loon/html/



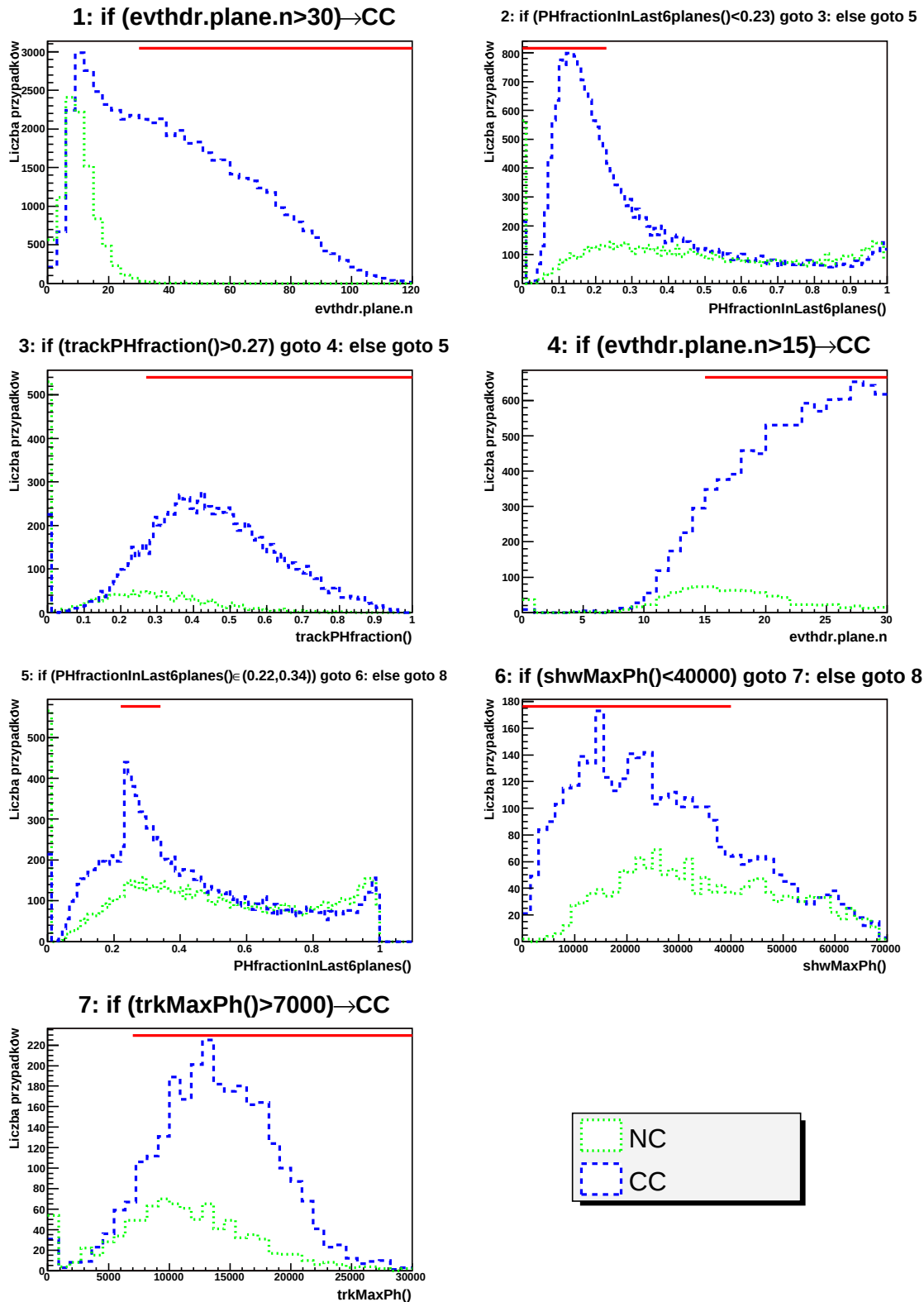
Rysunek 1: Rozkład energii mierzonej w detektorze dla przypadków NC i CC w przypadku zera zrekonstruowanych torów.



Rysunek 2: Rozkład energii mierzonej w detektorze dla przypadków NC i CC. Linie określają trzy różne zakresy energii. Każdy przypadek jest różnie przetwarzany w zależności od przedziału energii do którego należy.

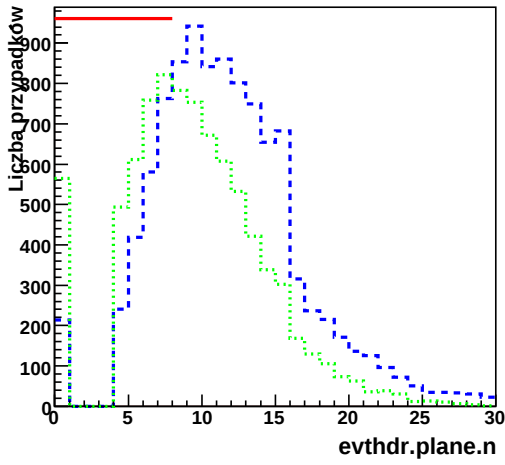


Rysunek 3: Rozkłady zmiennych stosowanych w selekcji przypadków z niską energią widzianą w detektorze (do około 1.6 GeV). Cienką, zieloną linią zaznaczono rozkład dla przypadków typu NC, linią niebieską (nieco grubszą) dla przypadków typu CC. Nad każdym z rysunków znajduje się fragment kodu odpowiedzialny za selekcję na podstawie danej zmiennej. Strzałka (→) oznacza zwrócenie danej wartości, fraza goto N oznacza skok do rysunku N w przypadku spełnienia bieżącego cięcia, fraza else goto N oznacza skok do rysunku N w przypadku niespełnienia bieżącego cięcia. Czerwoną linią na górze każdego rysunku zaznaczono obszar, z którego przypadki spełniają dane cięcie.

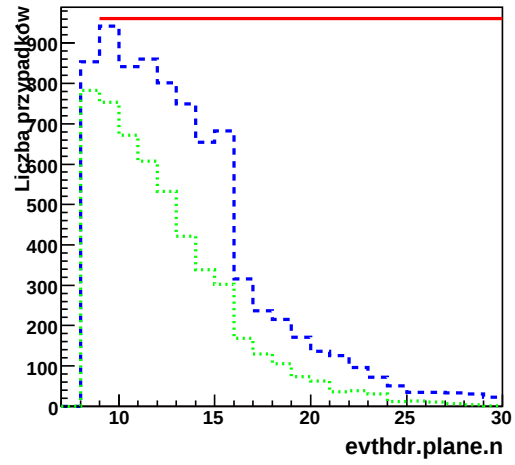


Rysunek 4: Rozkłady zmiennych stosowanych w selekcji przypadków ze średnią energią widzianą w detektorze (od około 1.6 GeV do około 5.6 GeV). Ciąg dalszy na następnej stronie.

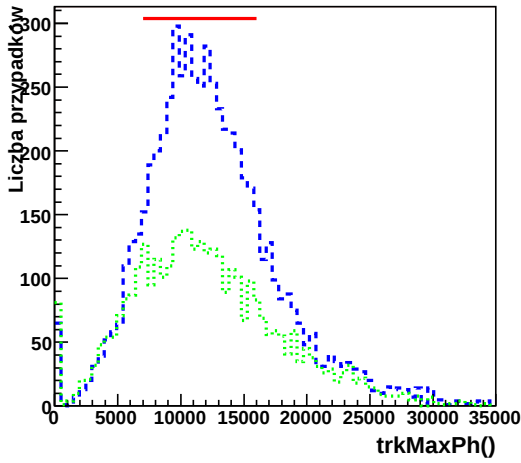
8: if (evthdr.plane.n<8)→NC



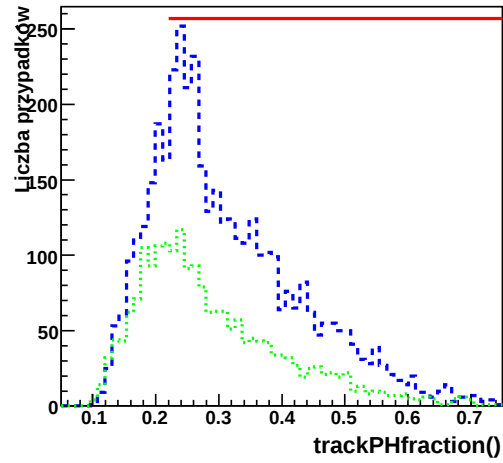
9: if (evthdr.plane.n>9) goto 10: else goto 12



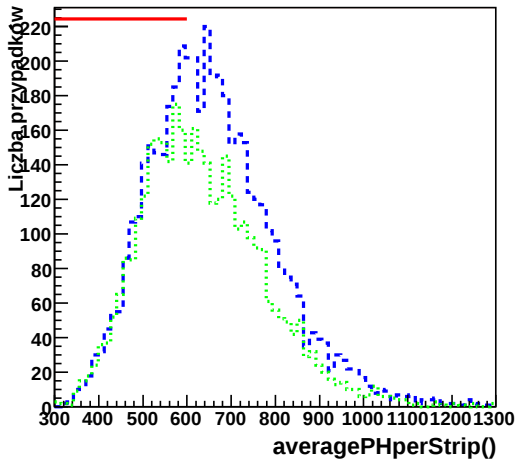
10: if (trkMaxPh()∈(7000,16000)) goto 11: else goto 12



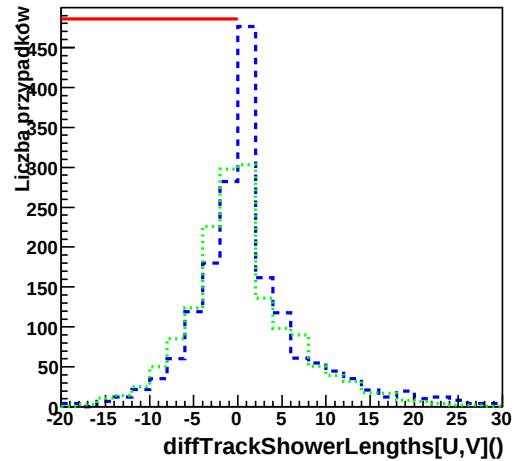
11: if (trackPHfraction()>0.22)→UNKNOWN: else goto 12



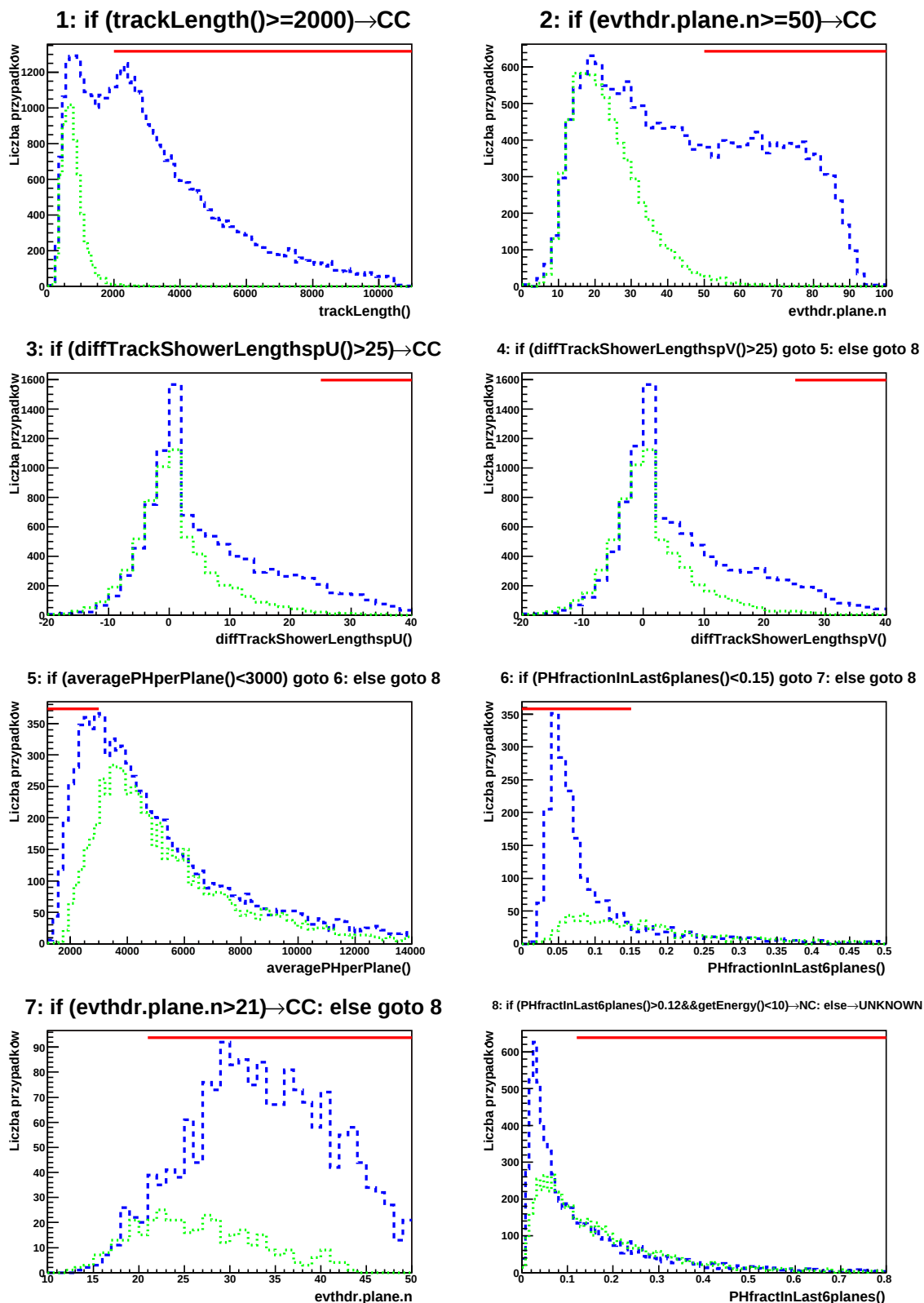
12: if (averagePHperStrip()<600) goto 13: else→UNKNOWN



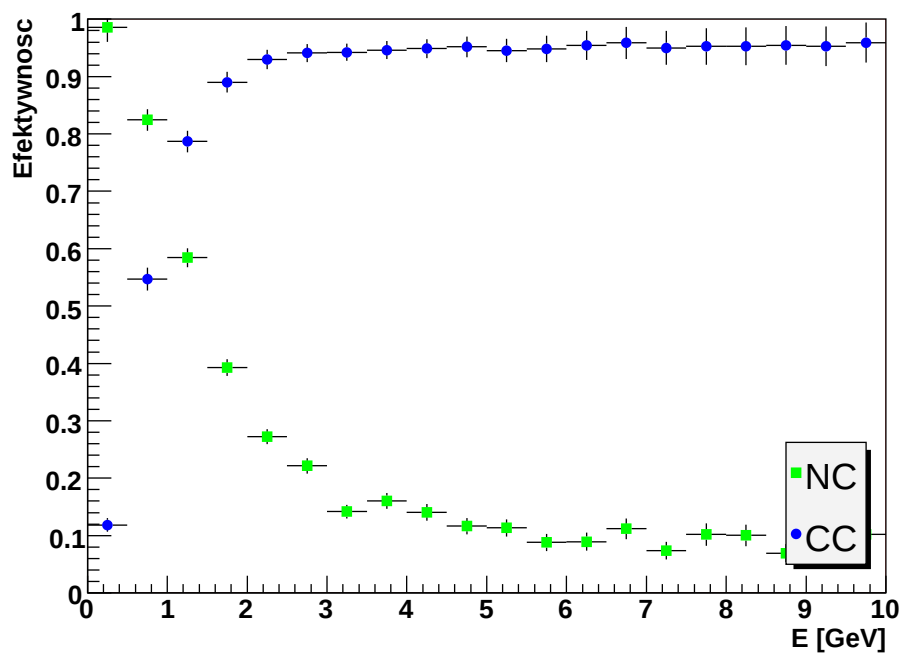
13: if (diffTrackShowerLengths[U,V](<0)→NC: else→UNKOWN



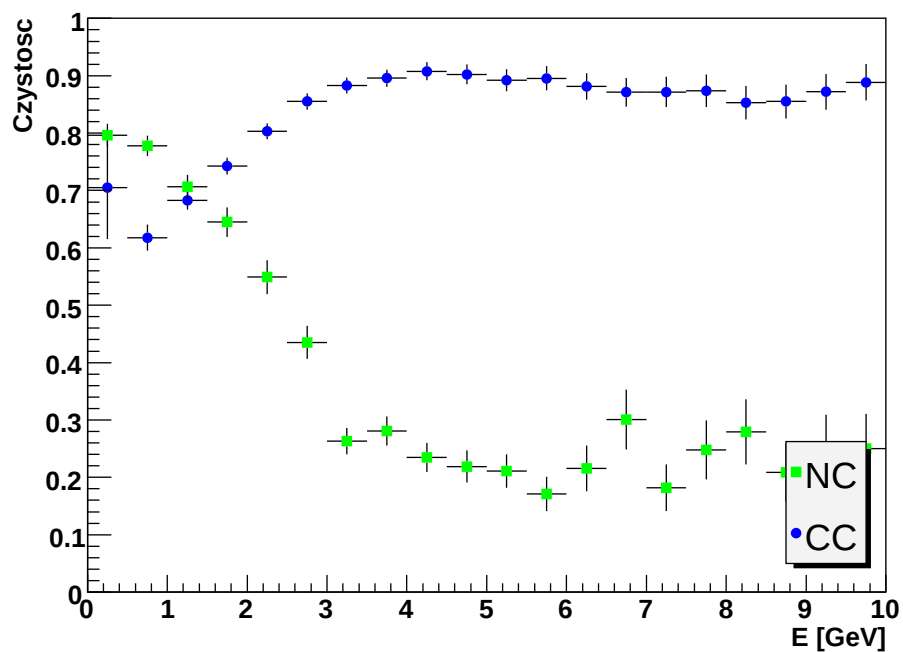
Rysunek 5: Rozkłady zmiennych stosowanych w selekcji przypadków z średnią energią widzianą w detektorze (od około 1.6 GeV do około 5.6 GeV) - ciąg dalszy.



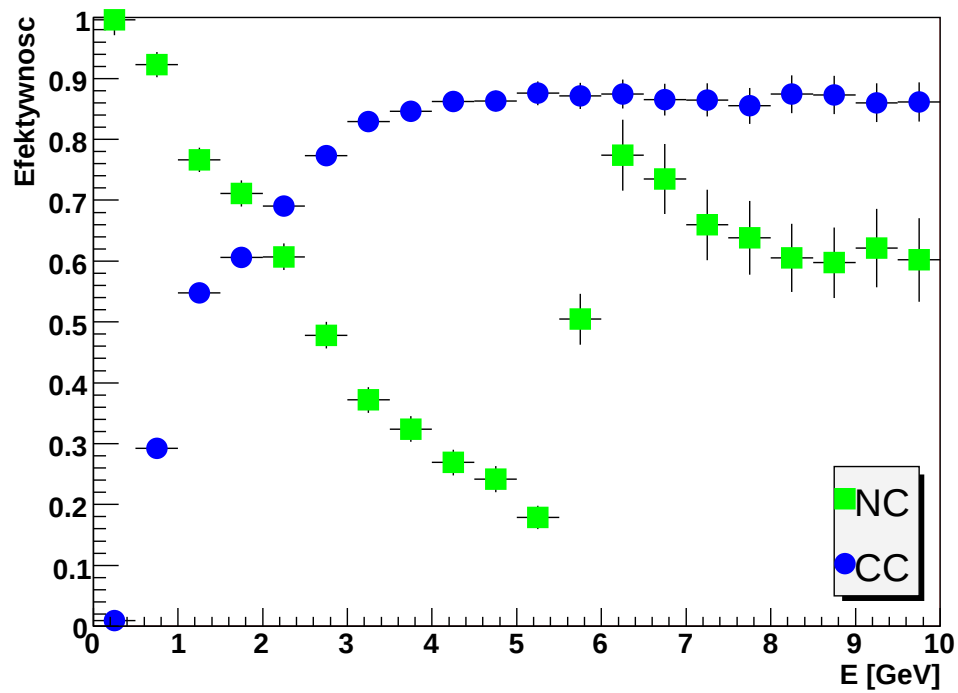
Rysunek 6: Rozkłady zmiennych stosowanych w selekcji przypadków z wysoką energią widzianą w detektorze (od około 5.6 GeV).



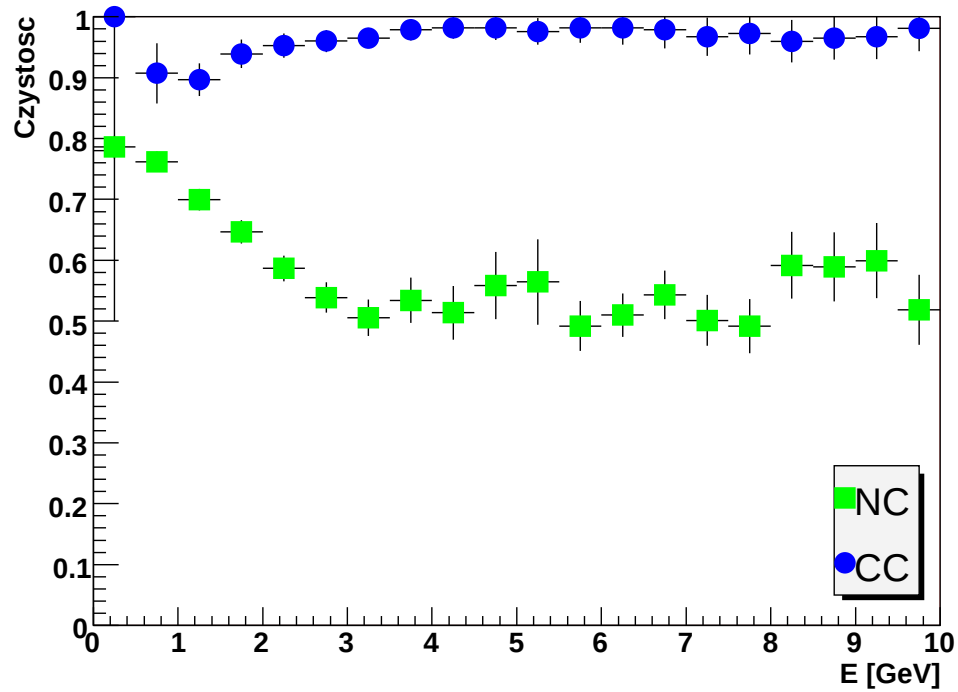
Rysunek 7: Efektywność prostego algorytmu selekcyjującego (selekcja na podstawie istnienia bądź nie zrekonstruowanych torów).



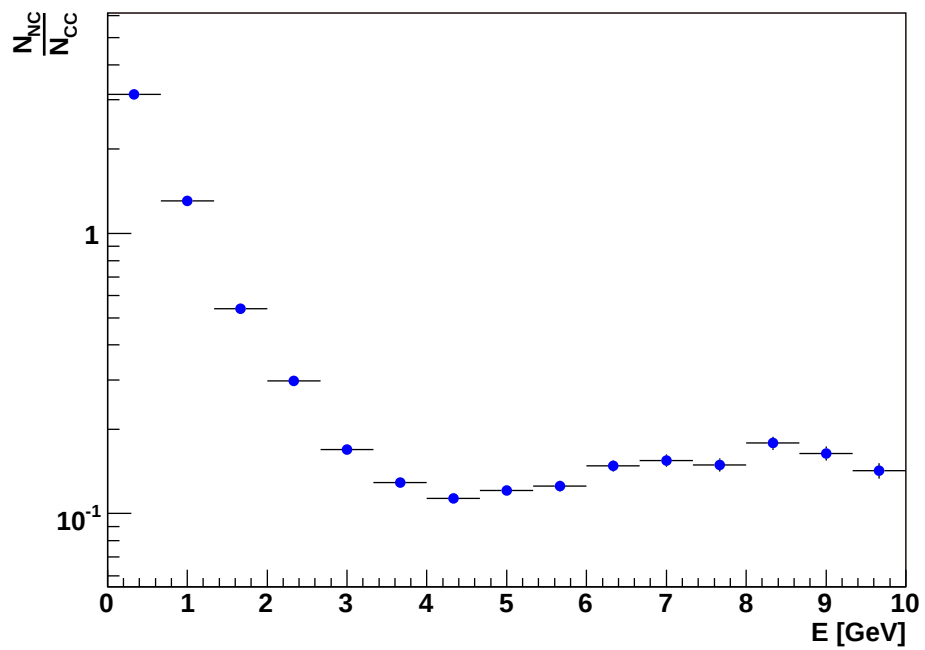
Rysunek 8: Czystość prostego algorytmu selekcyjującego (selekcja na podstawie istnienia bądź nie zrekonstruowanych torów).



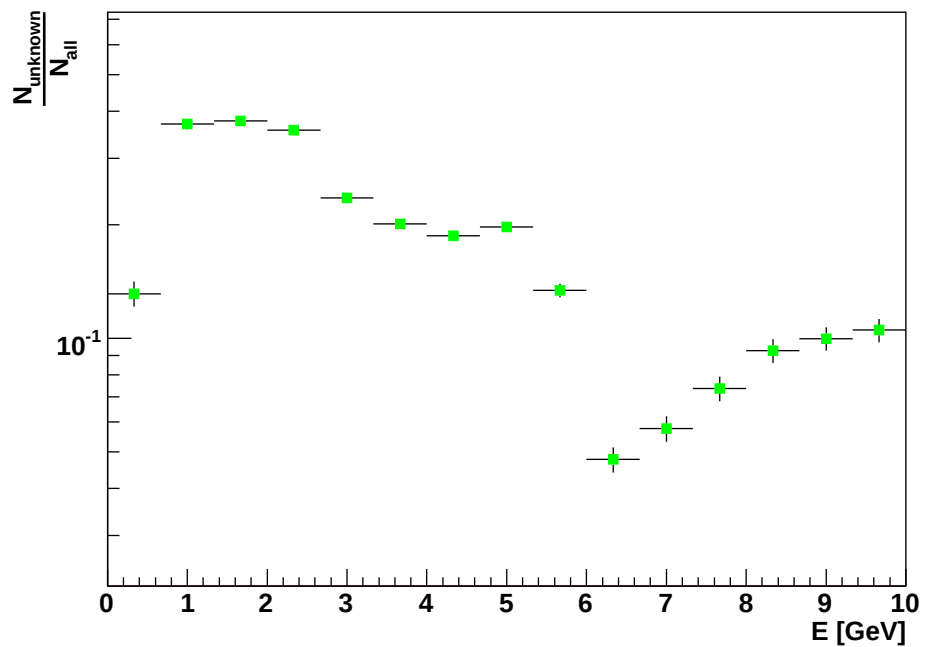
Rysunek 9: Efektywność algorytmu selekcjonującego przypadki typu NC i CC.



Rysunek 10: Czystość algorytmu selekcjonującego przypadki typu NC i CC.



Rysunek 11: Stosunek liczby przypadków NC do liczby przypadków CC (przed cięciami).



Rysunek 12: Stosunek liczby przypadków określonych jako nieznanego typu (UNKNOWN) do liczby wszystkich przypadków.