Recent updates in dielectron analysis with MPD experiment

Sudhir Pandurang Rode*, Itzhak Tserruya

*LHEP, JINR, Russia and Weizmann Institute of Science, Rehovot, Israel.

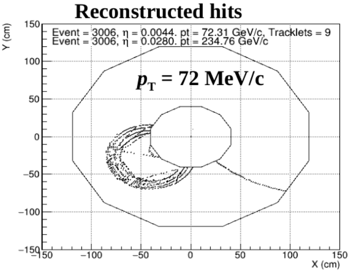
April 25, 2024

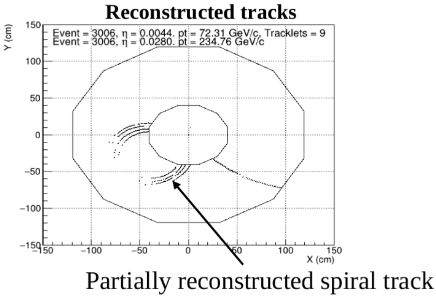
XIII MPD Collaboration meeting

Content

- Quick recap of the analysis so far
- The "Lost" electrons.
- Use of Machine learning approach for eID
  - Details of the classifiers
  - Training of the MC sample
  - Performance validation
  - Implementation in the dilepton analysis
- Next steps

Quick recap





- With current track reconstruction algorithm, low p_{T} tracks are not reconstructed properly even though full hit information is available in the detector for tracks that enter the TPC ($p_{\mathrm{T}} >\approx 30$ MeV/c).
- Question is, in an ideal detector, what would be the maximum possible benefit in the combinatorial background (CB) reduction, if we were to detect these tracks.
- As per our principle study, potentially, there is about 5-8 factor improvement possible in CB rejection.

Quick recap: Analysis strategy

⇒ Three electron pools:

→ Pool-1 for fully reconstructed tracks[1] in fiducial area ($|\eta| < 0.3$)

→ Pool-2 for fully reconstructed tracks in veto area $0.3 < |\eta| < 1.0$.

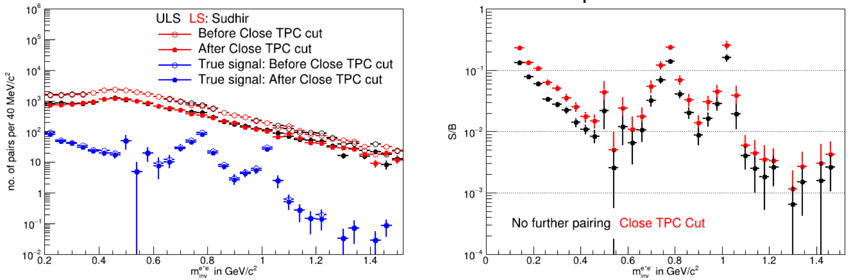
→ Pool-3 with tracks reconstructed in the TPC only.

- Step 1 - No further pairing (NFP): Tracks belonging to fully reconstructed π^0 Dalitz are tagged and not used for further pairing.
- Step 2 - Close TPC cut (CTC): Track from Pool-1 in an event is paired with tracks from Pool-3 in the same event and both tracks are removed as a potential Dalitz pair if they have $M_{\text{inv}} < 80$ MeV$/c^2$ and opening angle < 10 degrees (this cut is opening angle dependent).
- Step 3 - Rest of the tracks with $p_{\text{T}} > 200$ MeV from Pool-1 are paired among themselves to build ULS and LS pair spectra.

[1] TOF matched tracks identified in the TPC and TOF

Quick recap: Dielectron cocktail[3]

Request 25 $\rightarrow$ 36M events



Mass region: 0.2 to 1.5 GeV/c $\rightarrow$

Steps	Sig	LS	S/B	[2]BFE $= \frac{S^2}{S+2B}$
Before CTC	644.5	26285.2	0.024	7.8
After CTC	575.9	13317.7	0.043	12.2

- Due to limited satistics, signal is not U-L, but it is true reconstructed di-electron pairs.
- Close TPC cut approach improves S/B ratio by $\approx 75-80\%$ $\rightarrow$ CB rejection by factor 2.
- Still significant improvement possible by improving the recognition of low p_T tracks.

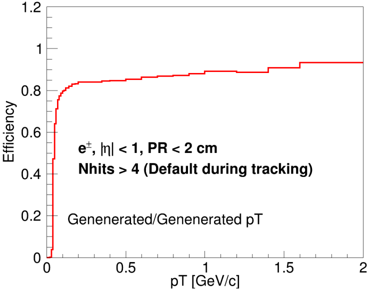
[2]Background free equivalent - signal with same relative error as in background free situation
[3]TPC+TOF analysis

Quick recap

Total reconstructed tracks after close TPC cut:	**1.69268e+06**
Below: Only Conversion and π^0 Dalitz sources are considered --	
a. Track has Partner with pT < 35 MeV ($\|\eta\|$ < 2.5):	**419595 (~25%)**
b. Track has Partner inside TPC i.e. 35 < pT < 100 MeV ($\|\eta\|$ < 2.5):	**580428 (~34%)**
c. Track has Partner with pT > 110 MeV ($\|\eta\|$ < 2.5):	**266075 (~16%)**
Track is hadron:	**102041 (~6%)**
Rest (Signal (η, etc), conversion, π^0 Dalitz whose partner outside TPC, ...)	**324536 (~19%)**

- Information (a.) is not available and therefore, it is lost.
- (b.) is recoverable upon improvement in reconstruction of turning tracks → requires expert to look into algorithm.
- In principle, (c.), is recoverable too, at least partially. In this presentation, we look into two possibilities:
  - The track has not been reconstructed at all → no trace → "Lost" electrons.
  - Improve the efficiency as tracks may not have satistfied one of the selection cuts → ML approach.

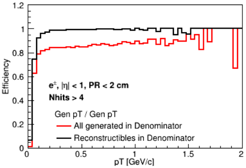
Section I - The "Lost" electrons[4]

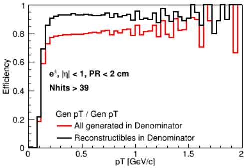


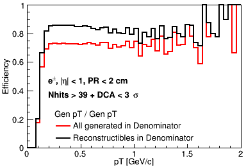
- This ratio of reconstructed electron tracks with Nhits in the TPC $>$ 4 to the all generated tracks should be close to 1.
- Some electrons are "lost" as they do not leave any MC points in the TPC $\rightarrow$ not "reconstructible".
- Effect propagates through different selection cuts and gives significantly less efficiency than what we should achieve.

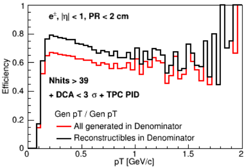
[4] These numbers/analysis are using MPDROOT version: request 25 version of MPDROOT (commit b95c9cb8 on `https://git.jinr.ru/nica/mpdroot/-/commits/massprod`

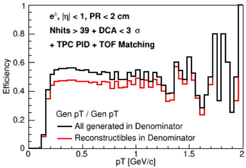
Step by step demonstration: reconstruction efficiency

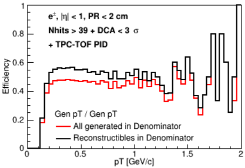


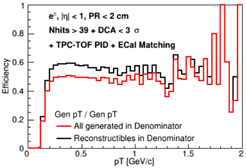


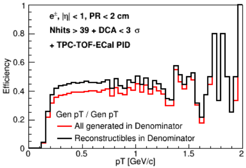






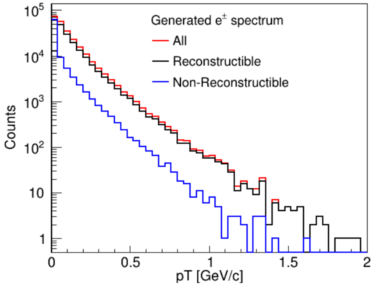


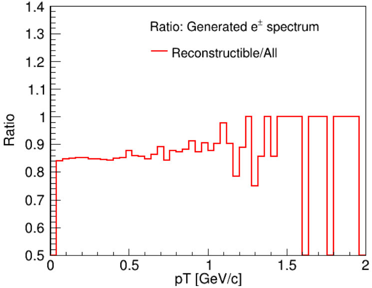




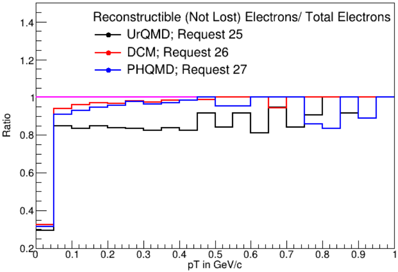
- Request 25.
- All generated electrons in the denominator.
- Only electrons having MC points in the TPC in the denominator.

Primary $e^{\pm}$ within $|\eta| < 1.0$: Lost electrons

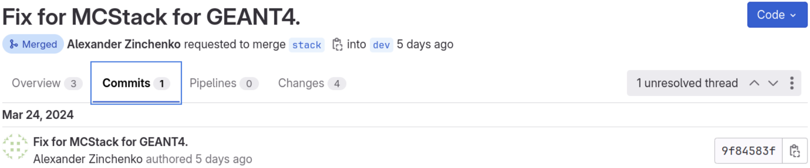




- ≈15% of electrons do not leave MC points in the TPC hence not reconstruction of those electrons.
- This problem is not observed in other productions, only in Request 25 (use external pythia8 decayer).
- After reporting the problem, Alexander Zinchenko has fixed the issue in the MC track GEANT4 settings → next slides.

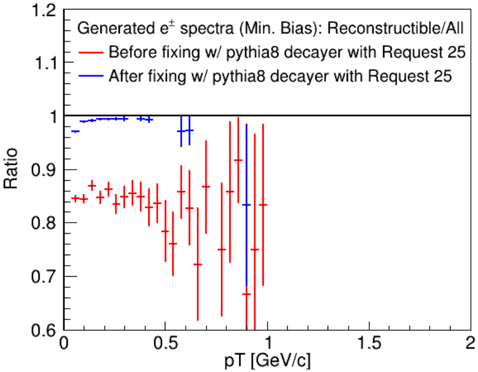


Problems with MCStack



- The main problem was that with GEANT4 previous input settings the MCStack was not used for handling decay products, while it was used to put secondary particles in Pythia decayer.
- This may have affected the cascade decays, i.e., for example, π^0-mesons from omegas. If the input setting "stackPopper" in g4Config.C is added, the stack starts to be used.
- In addition, due to usage of some internal variable to pass some information (which was overwritten by GEANT4), the particle with some index in the event (number 11) was lost.

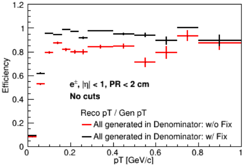
Primary $e^{\pm}$ within $|\eta| < 1.0$ - $\approx$ 6-8K Min. Bias UrQMD BiBi events

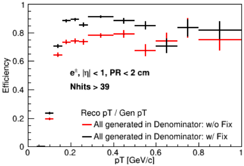


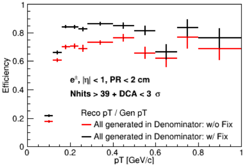
- For this study, the files in the commit shown in previous slides were added to the request 25 version of MPDROOT (commit b95c9cb8 on `https://git.jinr.ru/nica/mpdroot/-/commits/massprod`).
- I have also updated the beam pipe geometry (air $\rightarrow$ vacuum).
- Before fix: 6242 events.
- After fix: 7649 events.

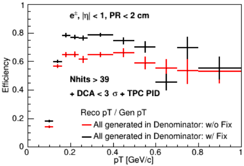
- With new updates in the MCStack and GEANT4 settings, the issue of lost electrons due to external pythia decayer seems to have vanished.
- Before and after fix scenario - MPDROOT version (one used for Request 25).
- Results with latest versions also show similar improvement (see the back up).

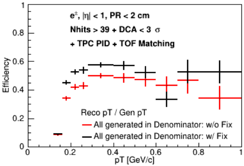
Primary $e^{\pm}$ within $|\eta| < 1.0$ - $\approx$ 6-8K Min. Bias UrQMD BiBi events

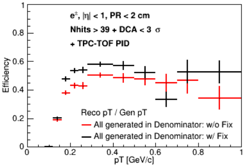


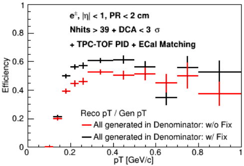


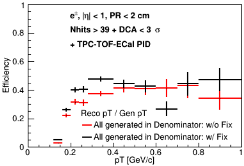












- Significant improvement in the single electron efficiency.
- This is expected to impact the di-electron analysis.
- Therefore, we would like to request for a new production for dielectrons to assess that.

Improvement in the reconstruction efficiency

- For $0.2 < p_{\mathrm{T}} < 2$ GeV/c.

Cuts	Efficiency±Error	Improvement±Error
Nhits > 4	0.820±0.008	
-	0.959±0.004	1.17±0.01
Nhits > 39	0.755±0.009	
-	0.882±0.006	1.17±0.01
+ DCA cut	0.712±0.010	
-	0.837±0.007	1.18±0.02
+ TPC PID	0.633±0.010	
-	0.753±0.009	1.19±0.02
+ TOF Matching	0.456±0.011	
-	0.541±0.010	1.19±0.03
+ TOF PID	0.462±0.011	
-	0.547±0.010	1.18±0.03
+ ECAL Matching	0.484±0.011	
-	0.576±0.010	1.19±0.03
+ ECAL PID	0.355±0.010	
-	0.427±0.010	1.20±0.04

Conclusions: Section I

- The issue of lost electrons in the TPC during Geant transport seems to have been fixed.
- The effect of this on the electron reconstruction and PID efficiency ($\approx$ 20%) is seen (Request 25 MPDROOT version).
- Similar effect is also seen with latest versions of MPDROOT[5]
- This is expected to have an effect on the dielectron analysis.
- If the agreement is reached, a new production for dielectrons with this fix is requested.

Thanks to Alexander Zinchenko for the discussions and fixing this issue.

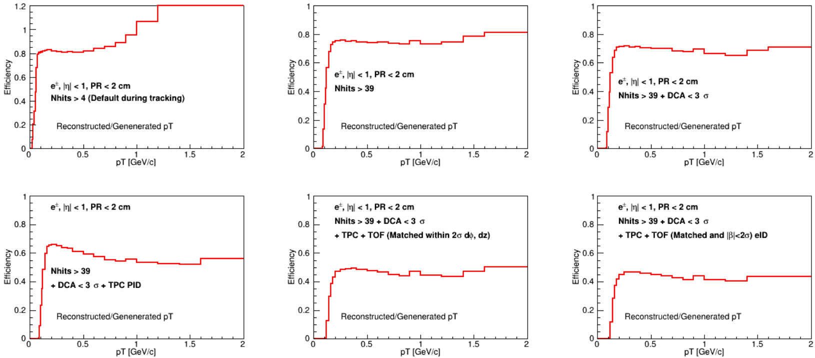
[5] the results are in the back-up, however, there is some inconsistency related to the conversions which need to be cross-checked.

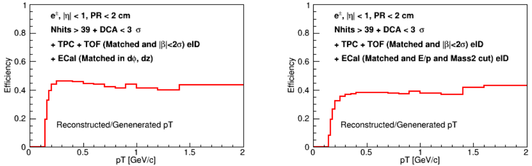
Section II: Machine learning approach for eID

Total reconstructed tracks after close TPC cut:	**1.69268e+06**
Below: Only Conversion and π^0 Dalitz sources are considered --	
a. Track has Partner with pT < 35 MeV ($\|\eta\|$ < 2.5):	**419595 (~25%)**
b. Track has Partner inside TPC i.e. 35 < pT < 100 MeV ($\|\eta\|$ < 2.5):	**580428 (~34%)**
c. Track has Partner with pT > 110 MeV ($\|\eta\|$ < 2.5):	**266075 (~16%)**
Track is hadron:	**102041 (~6%)**
Rest (Signal (η, etc), conversion, π^0 Dalitz whose partner outside TPC, ...)	**324536 (~19%)**

- In principle, (c.), is recoverable too.
  - Improve the efficiency as tracks may not have satistfied one of the selection cuts → ML approach.
- This is not only for the tracks in (c.) but tracks in (b.) as well. The improvement in the efficiency can help in enhancing the S/B, signal significance and background free equivalent signal.

Step-by-step efficiency using selection cuts





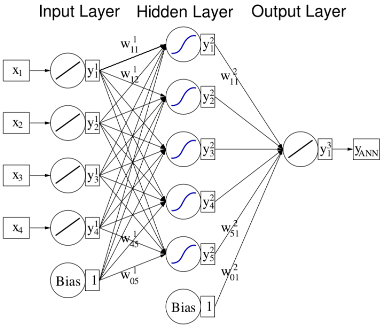
- Significant drop in efficiency due to 1D cuts.
- Improvement in the efficiency → better S/B, signal significance and background free equivalent signal.

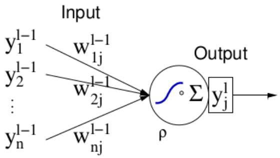
Machine Learning[6]

- The Machine learning can help in increasing the electron identification efficiency.
- Various algorithms are available, such as, neural networks, decision trees etc.
- TMVA package from cern ROOT library is utilized.
- It is user friendly and good starting point for the beginners.
- After initial study, Multi-Layer Perceptron (MLP) and Boosted Decision Tree (BDT) are employed for this study so far.

[6] **Disclaimer:** I am not a Machine learning expert. So my understanding about the topic may not be entirely true.

Neural Network: Multi-Layer Perceptrons (MLP)





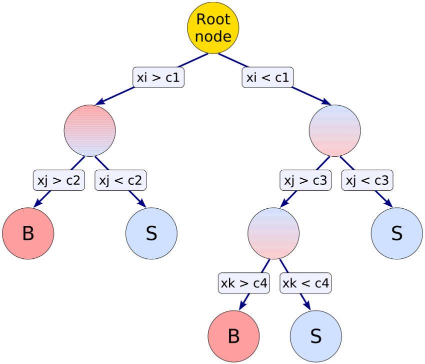
$$\kappa:\ (y_1^{(\ell)},..,y_n^{(\ell)}|w_{0j}^{(\ell)},..,w_{nj}^{(\ell)}) \rightarrow \begin{cases} w_{0j}^{(\ell)} + \sum_{i=1}^{n} y_i^{(\ell)} w_{ij}^{(\ell)} & \textit{Sum}, \\ w_{0j}^{(\ell)} + \sum_{i=1}^{n} \left(y_i^{(\ell)} w_{ij}^{(\ell)}\right)^2 & \textit{Sum of squares}, \\ w_{0j}^{(\ell)} + \sum_{i=1}^{n} |y_i^{(\ell)} w_{ij}^{(\ell)}| & \textit{Sum of absolutes}, \end{cases}$$

$$\alpha:\ x \rightarrow \begin{cases} x & \textit{Linear}, \\ \dfrac{1}{1+e^{-kx}} & \textit{Sigmoid}, \\ \dfrac{e^x - e^{-x}}{e^x + e^{-x}} & \textit{Tanh}, \\ e^{-x^2/2} & \textit{Radial}. \end{cases}$$

- Can be used to solve complex non-linear problems.
- Works well with both small and large input data.
- Helps to obtain quick predictions after the training and the same accuracy ratio with large as well as small data.
- In TMVA, all neural networks are feed-forward Multilayer Perceptrons.
- Training method: Back-Propagation (BP).

Decision Tree: Boosted Decision Tree (BDT)

- A decision tree takes a set of input features and splits input data recursively based on those features.
- Nodes: Place where Data is split.
- Leaves: Represent a class label or probability.
- Each split at a node is chosen to maximize information gain or minimize entropy.
- The splits are created recursively $\rightarrow$ the process is repeated until some stop condition is met.



- Boosting is a method of combining many weak learners (trees) into a strong classifier.
- Adaptive boost is used in this work.

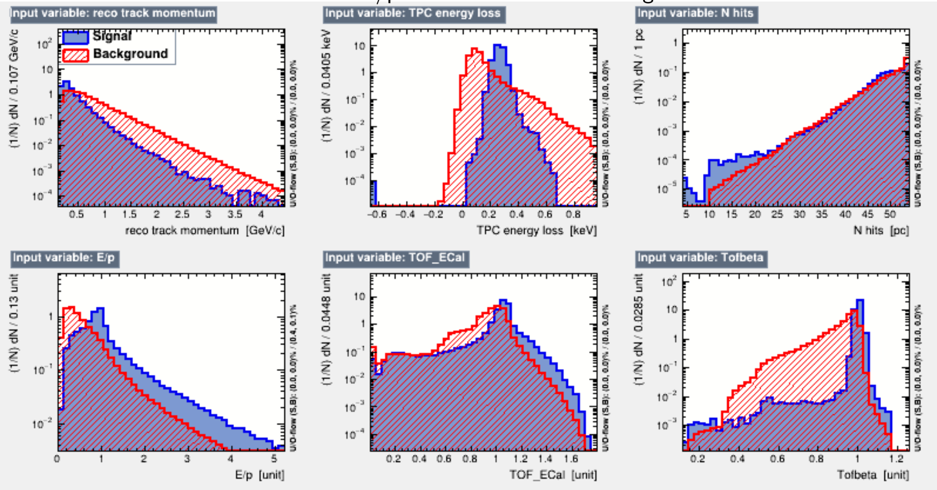
Details

- All charged tracks with DCA $< 3\sigma$ and matched in TOF ($< 2\sigma$ of dϕ and dz) and ECal ($< 3\sigma$ of dϕ and dz) $\rightarrow e^{\pm}$ (Signal) and Rest (Background).
- Three samples:
  - Sample 1: Training.
  - Sample 2: Overtraining test.
  - Sample 3: Performance validation.
- Sample 1 and 2 are of equal size with actual proportion of Signal (284K) and Background (47M) each, respectively.
- The Kolmogorov Smirnov test provides a p-value[7] equal to the statistical probability that two samples are drawn from the same distribution.

[7] The smaller the p, the greater the overtraining. Since the training and testing samples will never be identical, a very small degree of overtraining may be unavoidable. As a rule of thumb, it is recommended to try to reduce overtraining if $p < 0.01$, especially if the separation is visibly poorer for the testing samples than for the training samples.

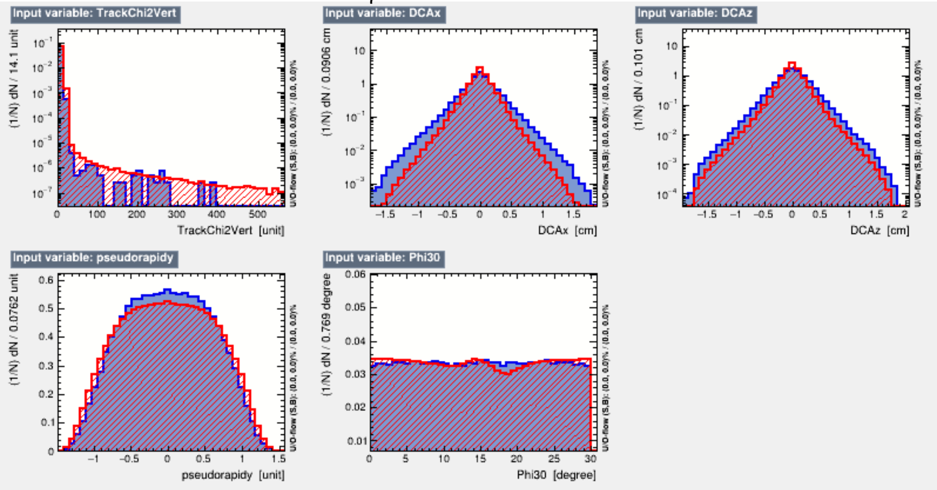
Input variables

- Momentum
- dEdX
- No of Hits
- E/p
- Time of flight in the ECal
- Time of flight in the TOF

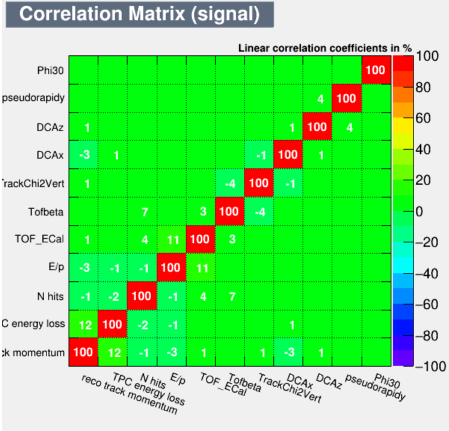


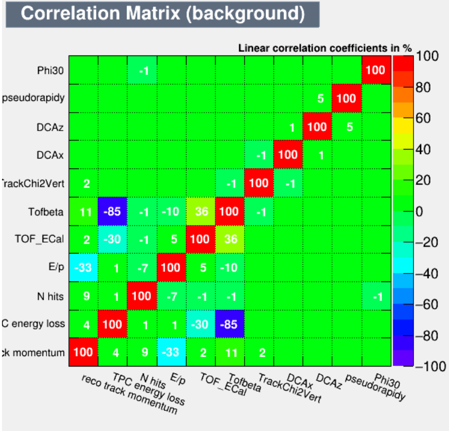
Input variables

- Track chi2 to vertex
- DCAx
- DCAz
- η
- Azimuthal angle, ϕ



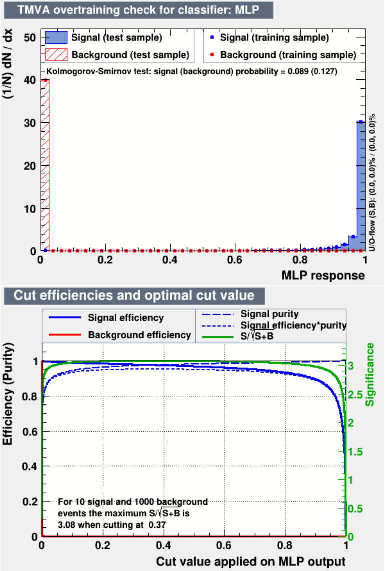
Correlation matrices: $e^{\pm}$ (Signal) and Rest (Bkg)

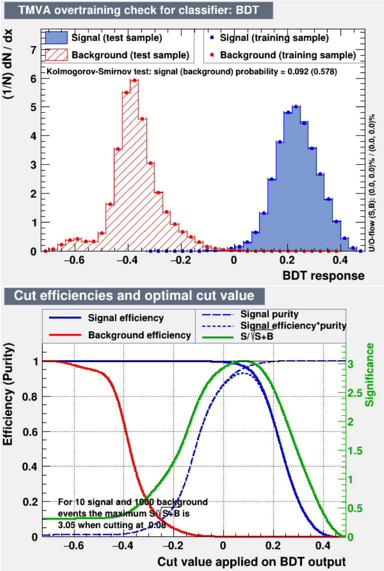




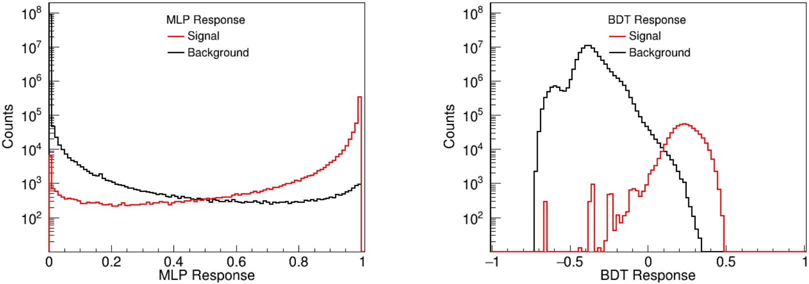
- Almost all variables for signal are independent.
- In case of background, there is correlation among some variables, for instance, dEdx and Tofbeta.

Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest (Bkg)



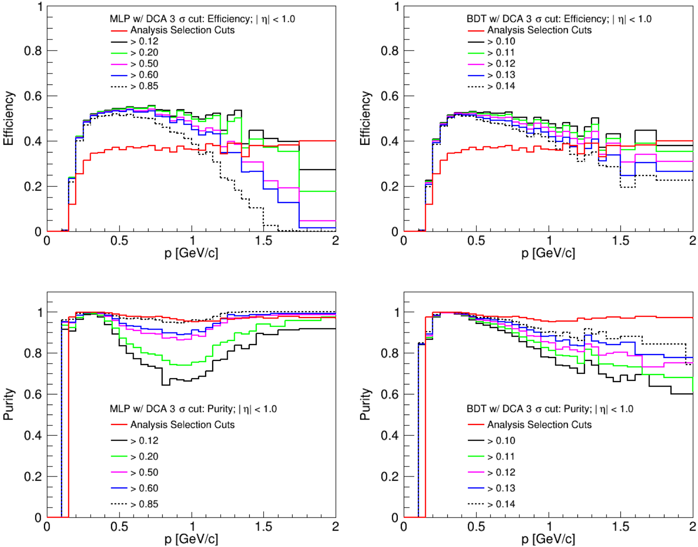


Performance validation using test sample



- Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest (Bkg).
- Response for actual proportion of signal and background in the test sample.
- Clear separation between signal and background by both classifiers.

Efficiency and Purity: $e^{\pm}$

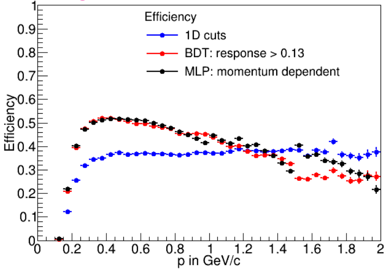


- Denominator: All generated $e^{\pm}$ tracks (PR < 2 cm).
- Numerator: $+$ Response cut.
- Purity: All $e^{\pm}$ to charged tracks with $\mathrm{DCA} < 3\sigma$ matched in TOF and ECAL within Response cut.
- With momentum dependent selection of response, purity as good as 1D cuts (analysis selection cuts) and better efficiency can be achieved.

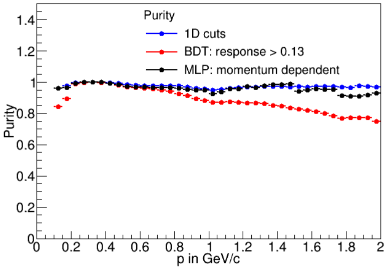
Implementation of Machine learning results in pair analysis: $\approx$ 21M events

Efficiencies and Purity: $\approx$ 21M events

Total single electron reconstruction efficiency



Electron purity



- MLP is performing better at higher momenta.
- Significant improvement in the efficiency.
- Purity with MLP matches with the 1D cuts.
- BDT: response > 0.13.
- MLP: momentum dependent, for p < 1.0, response > 0.85, $1.0 <$ p < 1.15, response > 0.7, $1.15 <$ p < 1.25, response > 0.6, $1.25 <$ p < 1.5, response > 0.5, $1.5 <$ p < 1.75, response > 0.2 upto p > 1.75, response $> 0.12 \rightarrow$ smoothening required.

Analysis strategy (slightly updated) - Reminder

$\Rightarrow$ Three electron pools:

$\rightarrow$ Pool-1 for fully reconstructed tracks[8] in fiducial area ($|\eta| < 0.3$)

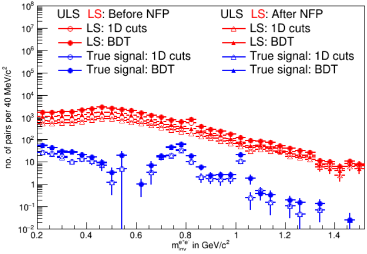
$\rightarrow$ Pool-2 for fully reconstructed tracks in veto area $0.3 < |\eta| < 1.0$.

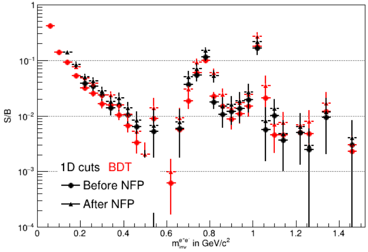
$\rightarrow$ Pool-3 with tracks not matched/identified in the TOF.

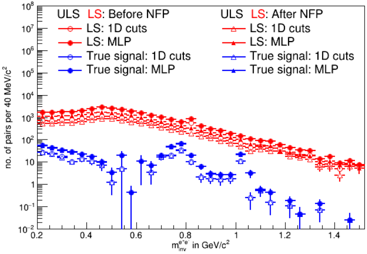
- Step 1 - No further pairing (NFP): Tracks belonging to fully reconstructed π^0 Dalitz are tagged and not used for further pairing.
- Step 2 - Close TPC cut (CTC): Track from Pool-1 in an event is paired with tracks from Pool-3 in the same event and both tracks are removed as a potential Dalitz pair if they have $M_{\rm inv} < 80$ MeV$/c^2$ and opening angle < 10 degrees (No opening angle dependent selection).
- Step 3 - Rest of the tracks with $p_{\rm T} > 200$ MeV from Pool-1 are paired among themselves to build ULS and LS pair spectra.

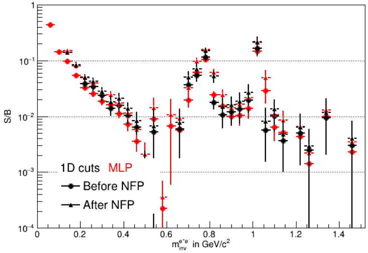
[8] TOF and ECal matched tracks identified in the TPC, TOF and ECal

Cocktail after No further pairing (NFP) using BDT & MLP (Fid. < 0.3)

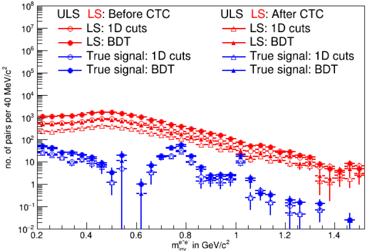


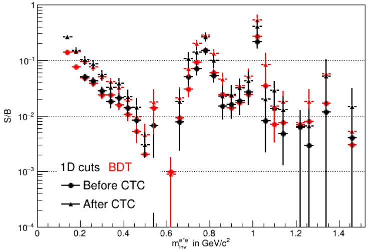


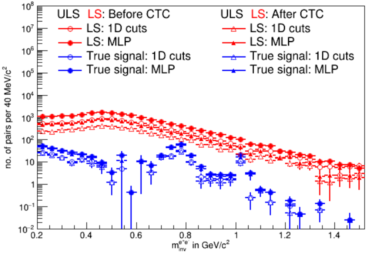


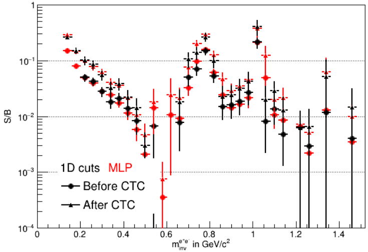


Cocktail after Close TPC Cut (CTC)[9] using BDT & MLP (Fid. < 0.3)









[9] Here, along with TPC only, tracks matched in ECal but not in the TOF are also included.

Comparison of results using 1D cuts, BDT and MLP

Following values are estimated in the invariant mass between 0.2 to 1.5 GeV/c $\rightarrow$

1) Fiducial region is $|\eta| < 0.3$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	12425±111	9124±96	4355±66	29733±172	18502±136	8250±91	29865±173	18510±136	8223±91
B	12580±112	9236±96	4411±66	29026±170	17884±134	8093±90	29105±171	17765±133	7991±89
U-B	-156±158	-112±136	-55±94	706±242	619±191	157±128	760±243	746±190	233±127
(U-B)/B	-0.012±-0.000	-0.012±-0.000	-0.013±-0.000	0.024±0.000	0.035±0.001	0.019±0.001	0.026±0.001	0.042±0.001	0.029±0.001
BFE	1.0±1.0	0.7±0.8	0.3±0.6	8.5±2.9	10.5±3.2	1.5±1.2	9.8±3.1	15.3±3.9	3.3±1.8
S	219	214	188	420	405	356	453	439	386
S/B	**0.017**	**0.023**	**0.043**	**0.014**	**0.023**	**0.044**	**0.016**	**0.025**	**0.048**
BFE	**1.9**	**2.5**	**3.9**	**3.0**	**4.5**	**7.7**	**3.5**	**5.4**	**9.1**

Comparison of results using 1D cuts, BDT and MLP

Following values are estimated in the invariant mass between 0.2 to 1.5 GeV/c $\rightarrow$

1) Fiducial region is $|\eta| < 0.3$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	12425±111	9124±96	4355±66	29733±172	18502±136	8250±91	29865±173	18510±136	8223±91
B	12580±112	9236±96	4411±66	29026±170	17884±134	8093±90	29105±171	17765±133	7991±89
U-B	-156±158	-112±136	-55±94	706±242	619±191	157±128	760±243	746±190	233±127
(U-B)/B	-0.012±-0.000	-0.012±-0.000	-0.013±-0.000	0.024±0.000	0.035±0.001	0.019±0.001	0.026±0.001	0.042±0.001	0.029±0.001
BFE	1.0±1.0	0.7±0.8	0.3±0.6	8.5±2.9	10.5±3.2	1.5±1.2	9.8±3.1	15.3±3.9	3.3±1.8
S	219	214	188	420	405	356	453	439	386
S/B	**0.017**	**0.023**	**0.043**	**0.014**	**0.023**	**0.044**	**0.016**	**0.025**	**0.048**
BFE	**1.9**	**2.5**	**3.9**	**3.0**	**4.5**	**7.7**	**3.5**	**5.4**	**9.1**

2) Fiducial region is $|\eta| < 0.7$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	69353±263	52668±229	23152±152	159160±399	102969±321	42979±207	157799±397	101350±318	42065±205
B	68820±262	52105±228	22590±150	157019±396	101008±318	41661±204	155322±394	98945±315	40583±201
U-B	533±372	563±324	562±214	2141±562	1961±452	1318±291	2477±560	2405±448	1482±287
(U-B)/B	0.008±0.000	0.011±0.000	0.025±0.000	0.014±0.000	0.019±0.000	0.032±0.001	0.016±0.000	0.024±0.000	0.037±0.001
BFE	2.1±1.4	3.0±1.7	6.9±2.6	14.5±3.8	18.8±4.3	20.5±4.5	19.6±4.4	28.9±5.4	26.6±5.2
S	1288	1266	1123	2482	2417	2056	2568	2494	2122
S/B	**0.019**	**0.024**	**0.050**	**0.016**	**0.024**	**0.049**	**0.017**	**0.025**	**0.052**
BFE	**11.9**	**15.2**	**27.3**	**19.5**	**28.6**	**49.5**	**21.1**	**31.0**	**54.1**

Comparison of results using 1D cuts, BDT and MLP

Following values are estimated in the invariant mass between 0.2 to 1.5 GeV/c $\rightarrow$

1) Fiducial region is $|\eta| < 0.3$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	12425±111	9124±96	4355±66	29733±172	18502±136	8250±91	29865±173	18510±136	8223±91
B	12580±112	9236±96	4411±66	29026±170	17884±134	8093±90	29105±171	17765±133	7991±89
U-B	-156±158	-112±136	-55±94	706±242	619±191	157±128	760±243	746±190	233±127
(U-B)/B	-0.012±-0.000	-0.012±-0.000	-0.013±-0.000	0.024±0.000	0.035±0.001	0.019±0.001	0.026±0.001	0.042±0.001	0.029±0.001
BFE	1.0±1.0	0.7±0.8	0.3±0.6	8.5±2.9	10.5±3.2	1.5±1.2	9.8±3.1	15.3±3.9	3.3±1.8
S	219	214	188	420	405	356	453	439	386
S/B	**0.017**	**0.023**	**0.043**	**0.014**	**0.023**	**0.044**	**0.016**	**0.025**	**0.048**
BFE	**1.9**	**2.5**	**3.9**	**3.0**	**4.5**	**7.7**	**3.5**	**5.4**	**9.1**

2) Fiducial region is $|\eta| < 0.7$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	69353±263	52668±229	23152±152	159160±399	102969±321	42979±207	157799±397	101350±318	42065±205
B	68820±262	52105±228	22590±150	157019±396	101008±318	41661±204	155322±394	98945±315	40583±201
U-B	533±372	563±324	562±214	2141±562	1961±452	1318±291	2477±560	2405±448	1482±287
(U-B)/B	0.008±0.000	0.011±0.000	0.025±0.000	0.014±0.000	0.019±0.000	0.032±0.001	0.016±0.000	0.024±0.000	0.037±0.001
BFE	2.1±1.4	3.0±1.7	6.9±2.6	14.5±3.8	18.8±4.3	20.5±4.5	19.6±4.4	28.9±5.4	26.6±5.2
S	1288	1266	1123	2482	2417	2056	2568	2494	2122
S/B	**0.019**	**0.024**	**0.050**	**0.016**	**0.024**	**0.049**	**0.017**	**0.025**	**0.052**
BFE	**11.9**	**15.2**	**27.3**	**19.5**	**28.6**	**49.5**	**21.1**	**31.0**	**54.1**

- At no further pairing step, S/B ratio remains similar for all three cases.
- Background free equivalent signal seems to have improved.
- After Close TPC cut, hint of improvement in S/B ratio using MLP and BDT.

Conclusions and Next steps: Section II

- Machine learning seems to be improving the PID efficiency.
- Enhancement in the background free equivalent signal, keeping S/B unchanged after no further pairing.
- Hint of improvement in the S/B after close TPC cut.
- Extend training to TPC only as well as TPC + ECal samples to further improve the S/B and significance.
- Optimise response cut for best efficiency and purity.
- Momentum differential training of the MC sample.

Thanks to Igor Rufanov for the discussions.

BACK-UP

MLP and BDT

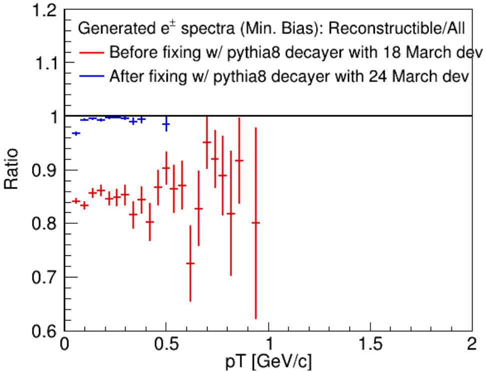
MLP	Training cycles	Hidden layers	Neuron activation function type	Neuron input function type
	600	1 (N+5)	tanh	sum

BDT	NTrees	BoostType	AdaBoostBeta	Max Depth
	850	AdaBoost	0.5	3

Details

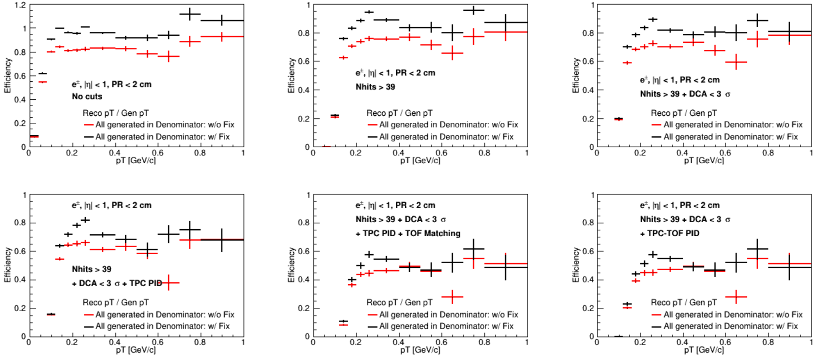
- For this study, the files in the commit shown in these slides were added to the March 24, 2024 version of MPDROOT (commit 9f84583f on `https://git.jinr.ru/nica/mpdroot/-/commit/9f84583fe2c2544d3bcad1739bf0fbf6104e5dc9`).
- And this version is used to get these results.
- I have also updated the beam pipe geometry (air $\rightarrow$ vacuum).
- For before fix scenario: March 18, 2024 version of MPDROOT (commit aa3dfb40 on `https://git.jinr.ru/nica/mpdroot/-/commit/aa3dfb40011f81336696432leb8be754cb06621a`).
- Before fix: 6264 events.
- After fix: 7715 events.

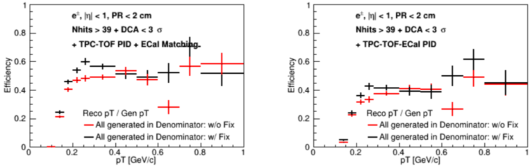
Primary $e^{\pm}$ within $|\eta| < 1.0$ - $\approx$ 6-8K Min. Bias UrQMD BiBi events



- With new updates in the MCStack and GEANT4 settings, new issue of lost electrons due to external pythia decayer seem to have vanished.

Primary $e^{\pm}$ within $|\eta| < 1.0$ - $\approx$ 6-8K Min. Bias UrQMD BiBi events





- This provides improvement in the single electron efficiency.
- This will give big boost to the di-electron analysis $\rightarrow$ CB rejection.
- Therefore, we would like to request for a new production for dielectrons.

Improvement in the reconstruction efficiency

- For $0.2 < p_{\mathrm{T}} < 2$ GeV/c.

Cuts	Efficiency±Error	Improvement±Error
Nhits > 4	0.824± 0.008	
-	0.967± 0.004	1.17± 0.01
Nhits > 39	0.749± 0.009	
-	0.889± 0.007	1.19 ±0.01
+ DCA cut	0.708± 0.009	
-	0.833± 0.008	1.18± 0.02
+ TPC PID	0.629± 0.009	
-	0.743± 0.010	1.18 ±0.02
+ TOF Matching	0.454± 0.010	
-	0.528± 0.011	1.16 ±0.03
+ TOF PID	0.460± 0.010	
-	0.533± 0.011	1.16± 0.03
+ ECAL Matching	0.483± 0.010	
-	0.555± 0.011	1.15 ±0.03
+ ECAL PID	0.358± 0.010	
-	0.406± 0.011	1.14± 0.04

Effect on multiplicities of electron sources

- Average multiplicities for $p_{\mathrm{T}} > 200$ MeV/c per 100 events ($|\eta| < 1.0$ and produced within 2 cm)

	Cuts	Average Multiplicity±Error	Improvement±Error
Before Fix	π^0-Dalitz	11.12±0.38	
After Fix	π^0-Dalitz	13.60±0.47	1.22±0.06
Before Fix	η-Dalitz	1.28±0.13	
After Fix	η-Dalitz	1.15±0.14	0.90±0.14
Before Fix	ρ^0	0.03±0.02	
After Fix	ρ^0	0.1±0.04	3.33±2.59
Before Fix	ω	0.23±0.06	
After Fix	ω	0.14±0.05	0.61±0.27
Before Fix	γ	2.58±0.18	
After Fix	γ	2.17±0.19	0.84±0.1

Contribution from conversions before beam pipe despite using updated geometry $\rightarrow$ needs cross-check.

Problems with MCStack

Fix for MCStack for GEANT4.

Code

Merged **Alexander Zinchenko** requested to merge `stack` into `dev` 5 days ago

Overview 3 | **Commits 1** | Pipelines 0 | Changes 4 — 1 unresolved thread

Mar 24, 2024

Fix for MCStack for GEANT4.
Alexander Zinchenko authored 5 days ago — `9f84583f`

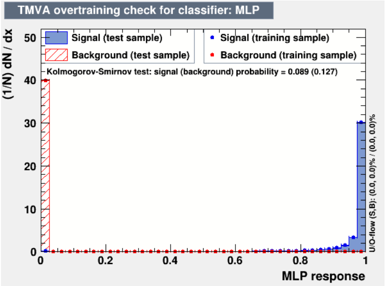
gconfig/MpdDecayConfig.txt +1 −0 View file @ 9f84583f

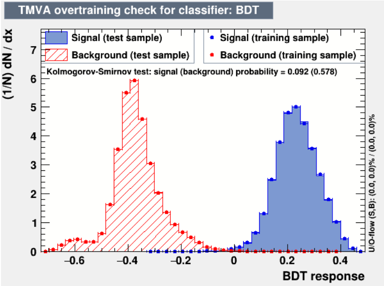
```diff
@@ -6,6 +6,7 @@
 # Spaces are channel separators.
 # ":*" means inclusive decay modes, i.e. affected only the channels
 # explicitly mentioned
+# If mother particle differs from anti-particle, the particle should be used
 #113:* 113:11:-11:x1
 #223:* 223:11:-11:x1 223:111:11:-11:x1
 #333:* 333:11:-11:x1 333:221:11:-11:x1
```

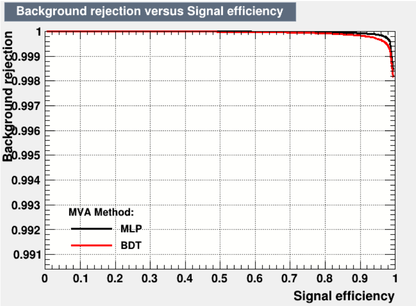
gconfig/g4Config.C 100755 → 100644 +2 −1 View file @ 9f84583f

```diff
@@ -28,7 +28,8 @@ void Config()
   //TG4RunConfiguration* runConfiguration
           //= new TG4RunConfiguration("geomRoot", "FTFP_BERT", "stepLimiter+specialCuts");
   TG4RunConfiguration* runConfiguration
-          = new TG4RunConfiguration("geomRoot", "FTFP_BERT+optical");
+     //AZ-240324          = new TG4RunConfiguration("geomRoot", "FTFP_BERT+optical");
+     = new TG4RunConfiguration("geomRoot", "FTFP_BERT+optical", "stepLimiter+stackPopper"); //AZ-240324
 
 /// Create the G4 VMC
   TGeant4* geant4 = new TGeant4("TGeant4", "The Geant4 Monte Carlo", runConfiguration);
```

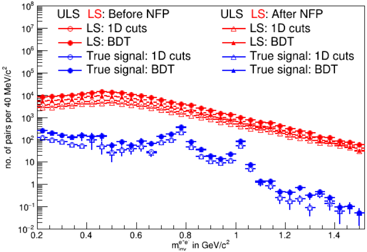
Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest (Bkg)

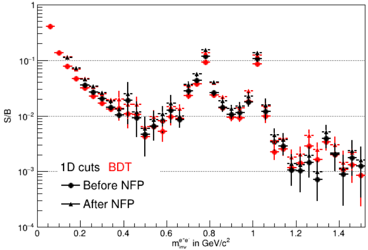


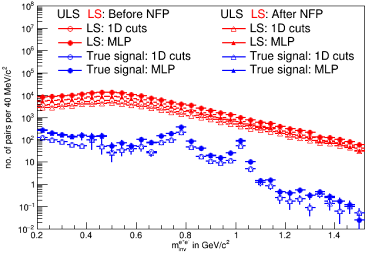


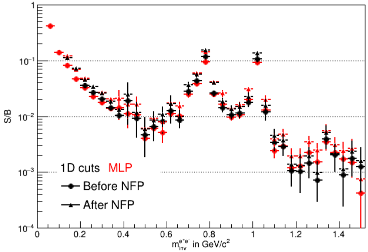


Cocktail after No further pairing (NFP) using BDT & MLP (Fid. < 0.7)

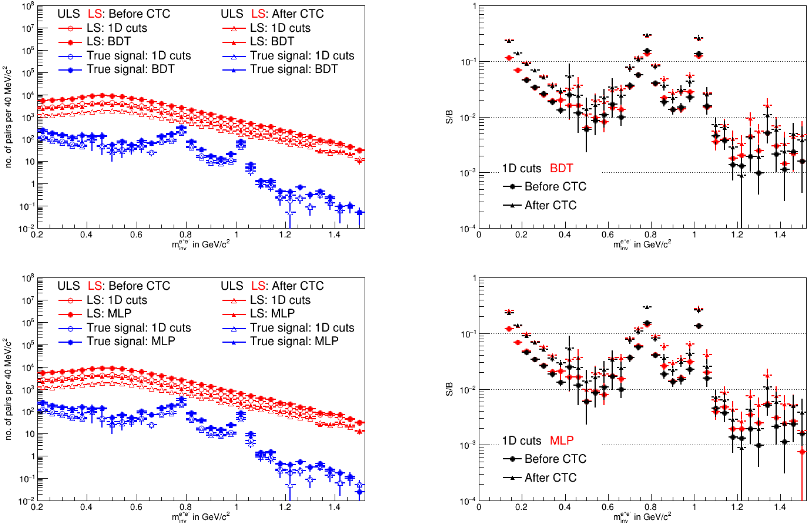






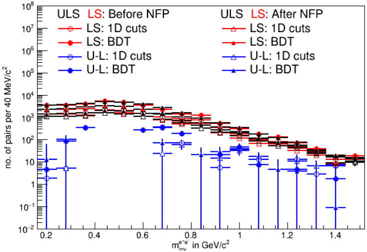


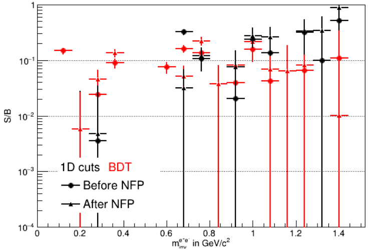
Cocktail after Close TPC Cut (CTC)[10] using BDT & MLP (Fid. < 0.7)

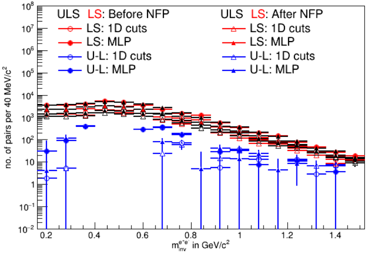


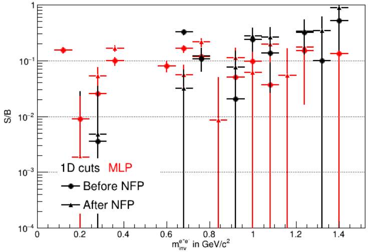
[10] Here, along with TPC only, tracks matched in ECal but not in the TOF are also included.

Cocktail after No further pairing (NFP) using BDT & MLP (Fid. < 0.3)

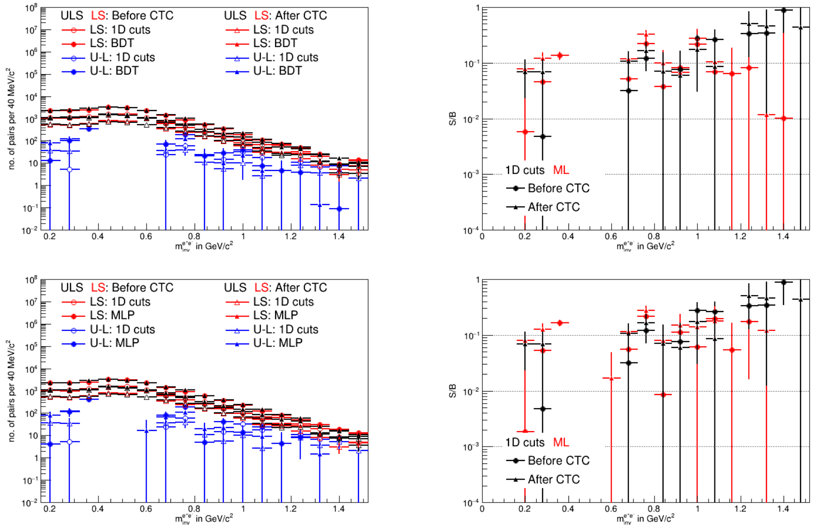






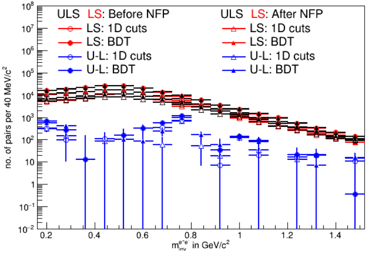


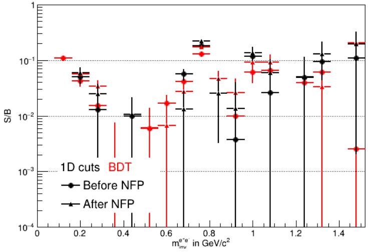
Cocktail after Close TPC Cut (CTC)[11] using BDT & MLP (Fid. < 0.3)

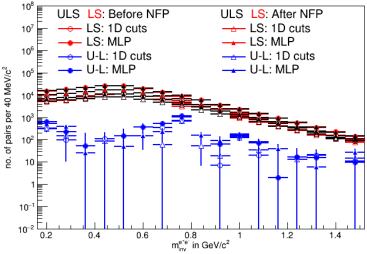


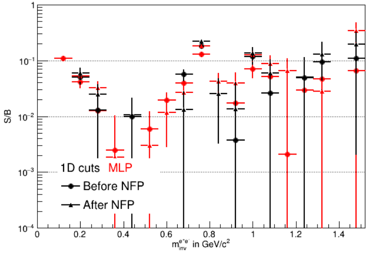
[11] Here, along with TPC only, tracks matched in ECal but not in the TOF are also included.

Cocktail after No further pairing (NFP) using BDT & MLP (Fid. < 0.7)

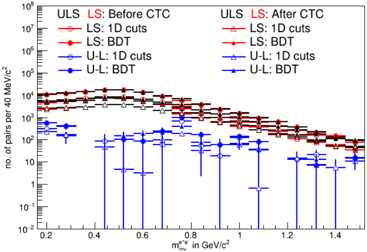


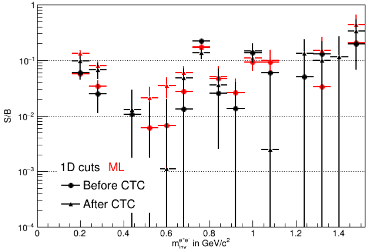


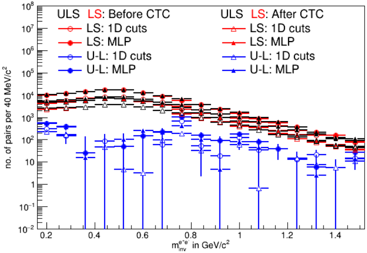


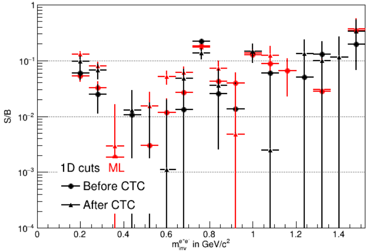


Cocktail after Close TPC Cut (CTC)[12] using BDT & MLP (Fid. < 0.7)









[12] Here, along with TPC only, tracks matched in ECal but not in the TOF are also included.

Comparison of results using 1D cuts, BDT and MLP

Following values are estimated in the invariant mass between 0.2 to 1.5 GeV/c $\rightarrow$

1) Fiducial region is $|\eta| < 0.3$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	12412±111	9150±96	4379±66	29733±172	18750±137	8394±92	29873±173	18770±137	8373±92
B	12568±112	9261±96	4428±67	29026±170	18121±135	8232±91	29114±171	18028±134	8141±90
U-B	-157±158	-111±136	-49±94	706±242	629±192	163±129	759±243	742±192	231±129
(U-B)/B	-0.012±-0.000	-0.012±-0.000	-0.011±-0.000	0.024±0.000	0.035±0.001	0.020±0.001	0.026±0.001	0.041±0.001	0.028±0.001
BFE	1.0±1.0	0.7±0.8	0.3±0.5	8.5±2.9	10.7±3.3	1.6±1.3	9.8±3.1	15.0±3.9	3.2±1.8
S	219	214	188	420	406	357	453	440	387
S/B	**0.017**	**0.023**	**0.043**	**0.014**	**0.022**	**0.043**	**0.016**	**0.024**	**0.047**
BFE	**1.9**	**2.4**	**3.9**	**3.0**	**4.5**	**7.6**	**3.5**	**5.3**	**9.0**

2) Fiducial region is $|\eta| < 0.7$.

	1D cuts			BDT			MLP		
	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC	Before NFP	After NFP	After CTC
U	69388±263	52925±230	23324±153	159160±399	104337±323	43753±209	157735±397	102707±320	42818±207
B	68851±262	52368±229	22769±151	157019±396	102367±320	42392±206	155256±394	100276±317	41306±203
U-B	537±372	558±324	555±215	2141±562	1970±455	1361±294	2479±559	2431±451	1511±290
(U-B)/B	0.008±0.000	0.011±0.000	0.024±0.000	0.014±0.000	0.019±0.000	0.032±0.001	0.016±0.000	0.024±0.000	0.037±0.001
BFE	2.1±1.4	3.0±1.7	6.7±2.6	14.5±3.8	18.8±4.3	21.5±4.6	19.6±4.4	29.1±5.4	27.1±5.2
S	1288	1266	1123	2482	2420	2058	2567	2497	2123
S/B	**0.019**	**0.024**	**0.049**	**0.016**	**0.024**	**0.049**	**0.017**	**0.025**	**0.051**
BFE	**11.9**	**15.1**	**27.0**	**19.5**	**28.3**	**48.8**	**21.0**	**30.7**	**53.2**

- At no further pairing step, S/B ratio remains similar for all three cases.
- Background free equivalent signal seems to have improved.
- After Close TPC cut, hint of improvement in S/B ratio using MLP and BDT.

Request 25 $\rightarrow$ 11M events

$\rightarrow$ **Fully reconstructed tracks: Pool 1**

- $|Vz| < 100$ cm.
- DCA x,y,z $< 3\sigma$.
- Nhits > 39
- TPC nSigma -2 to 2 sigma at p $=$ 0 and -1 to 2 sigma for p $>$ 800 MeV/c2.
- TOF nSigma -2 to 2 sigma
- TOF matching -2 to 2 sigma
- Limiting the eta acceptance of the reconstructed track to 0.3

$\rightarrow$ **Cuts on Partner: Pool 2**

- Same as Pool 1 except in $0.3 < \eta < 1.0$

$\rightarrow$ **Cuts on Partner for Close TPC Cut: Pool 3**

- $|\eta| < 2.5$, Nhits < 10
- DCA $<$ 3.5 sigma
- $|$TPC nSigma$| < 2$ sigma, Those tracks who DO NOT Matched in TOF within 2 Sigma (TPC ONLY).

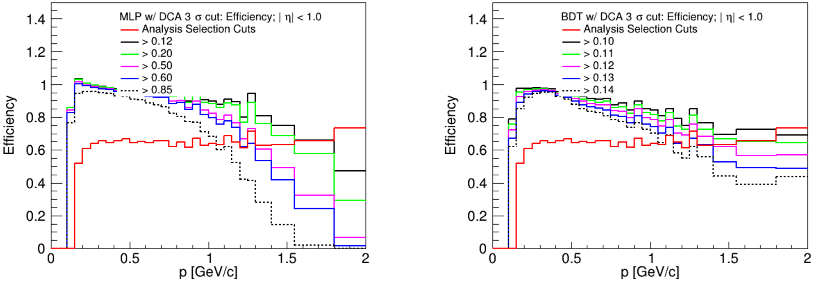
Analysis Selection Cuts vs Machine Learning

Steps	1D Cuts	Machine Learning
Denominator OR Input Sample	DCA $< 3\sigma$ Tracks matched in TOF and ECAL	DCA $< 3\sigma$ Tracks matched in TOF and ECAL
Numerator/Step 2	1D cuts	Train the model and test

$$\text{Efficiency in ML} = \frac{\text{No of primary e}^{\pm}\text{s after response cut}}{\text{No of e}^{\pm}\text{s in the input sample with DCA}{<}3\sigma + |\eta|{<}1.0 + \text{PR}{<}2.0\text{ cm}}$$

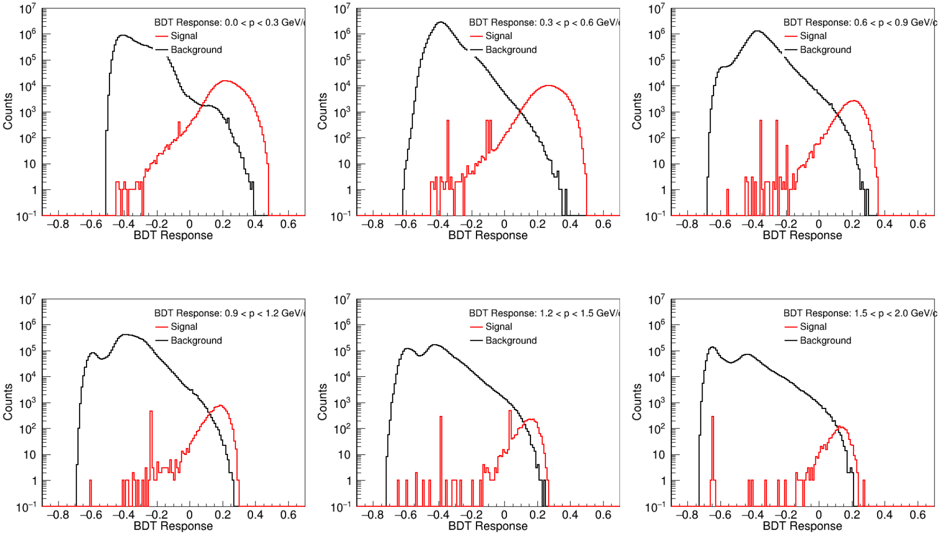
$$\text{Efficiency in 1D cuts} = \frac{\text{No of primary e}^{\pm}\text{s after selection cuts}}{\text{No of e}^{\pm}\text{s in the input sample with DCA}{<}3\sigma + |\eta|{<}1.0 + \text{PR}{<}2.0\text{ cm}}$$

Efficiency: Primary $e^{\pm}$

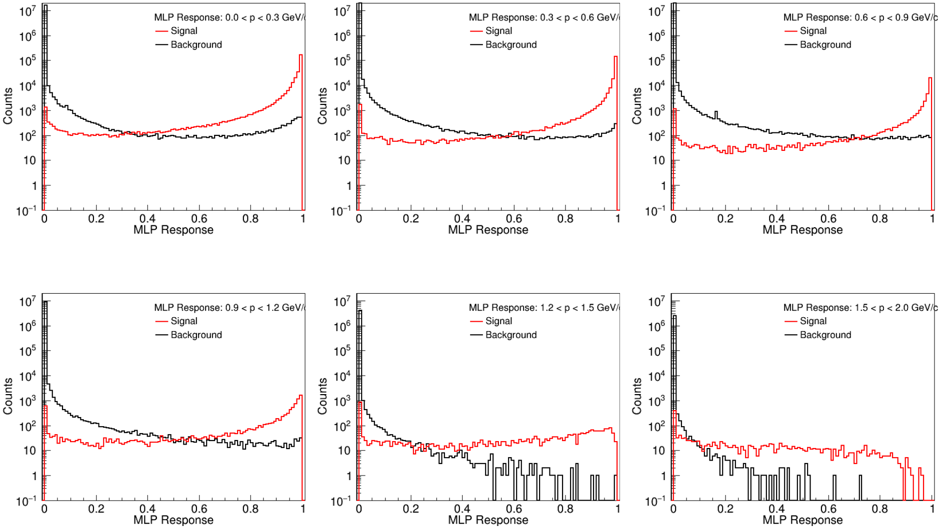


- Denominator: All $e^{\pm}$ tracks (PR $<$ 2 cm) with $\mathrm{DCA} < 3\sigma$ and matched in TOF and ECAL.
- Numerator: $+$ Response cut
- Denominator is same in both 1D cuts and machine learning.
- Benefit is that the inefficiency due to cuts on Nhits, TPC, TOF and ECAL is reduced with negligible comprise on the purity.
- However, the conversion contribution is more here because the Positron efficiency has increased.

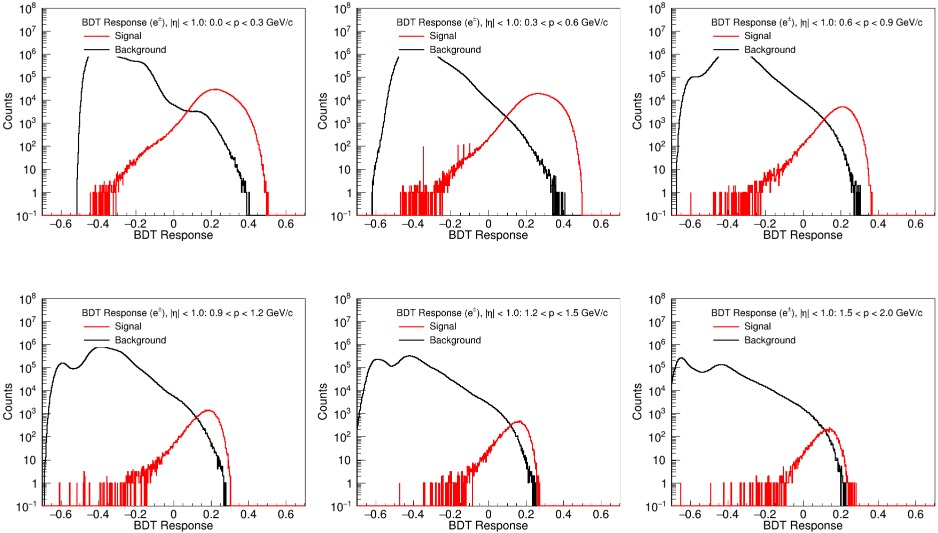
p dependent BDT Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest



p dependent MLP Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest



p dependent BDT Response with Prior DCA 3σ cut; All $e^{\pm}$ (Signal) and Rest



Problems with MCStack

simulation/generators/mpdGen/MpdDecayerPyt8.cxx +2 −1 View file @ 9f84583f

```diff
@@ -88,7 +88,8 @@ void MpdDecayerPyt8::Init()
     fBranch                 = channel.bRatio();

     // Random number generator
-    fRandom = new TRandom(0); // time-dependent seed
+    // AZ-220324 fRandom = new TRandom(0); // time-dependent seed
+    fRandom = new TRandom(1); // AZ-220324 time-independent seed
     // fRandom = gRandom;

     // fPythia8->Pythia8()->particleData.isResonance(113,kTRUE);
```

simulation/mcStack/MpdStack.cxx +23 −17 View file @ 9f84583f

```diff
@@ -133,31 +133,32 @@ void MpdStack::PushTrack(Int_t toBeDone, Int_t parentId, Int_t pdgCode, Double_t

      for (Int_t j = 0; j < npart; ++j) {
         TParticle tpart = *((TParticle *)fDecays->UncheckedAt(j));
-        // tpart.Print();
+        // tpart.Print(); //AZ-210324
         vDecays.push_back(tpart);
      }

      for (Int_t j = 0; j < npart; ++j) { // skip mother particle
         if (j == 0) {
-           particle->SetStatusCode(11); // decayed particle
-           continue;                    // skip mother particle
+           // AZ-200324 particle->SetStatusCode(11); // decayed particle
+           particle->SetDaughter(0, 11); // AZ-200324 decayed particle
+           continue;                     // skip mother particle
         }
         // TParticle *part = (TParticle*) fDecays->UncheckedAt(j);
         TParticle *part = &vDecays[j];
         Int_t      toDo = 1;
-        if (TString(gMC->GetName()) == "TGeant4") toDo = 0;
-        // PushTrack(toBeDone, trackId, part->GetPdgCode(),
+        // AZ-210324 if (TString(gMC->GetName()) == "TGeant4") toDo = 0;
+        //  PushTrack(toBeDone, trackId, part->GetPdgCode(),
         PushTrack(toDo, trackId, part->GetPdgCode(), part->Px(), part->Py(), part->Pz(), part->Energy(), part-
>Vx(),
                   part->Vy(), part->Vz(), time, polx, poly, polz, proc, ntr, weight, is);
      }
-     if (parentId < 0) fNPrimaries += (npart - 1); // treat decay products as primaries (dirty trick)
+     // AZ-220324 if (parentId < 0) fNPrimaries += (npart - 1); // treat decay products as primaries (dirty trick)
   } else {
      // --> Push particle on the stack if toBeDone is set
      if (toBeDone == 1) fStack.push(particle);
   }

   // --> Push particle on the stack if toBeDone is set
-  // AZ if (toBeDone == 1) fStack.push(particle);
+  // AZ if (toBeDone == 1) fStack.push(particle); //
 }

```

simulation/mcStack/MpdStack.cxx +23 −17 View file @ 9f84583f

```diff
 void MpdStack::PushTrack(Int_t toBeDone, Int_t parentId, Int_t pdgCode, Double_t px, Double_t py, Double_t pz,
@@ -226,18 +227,19 @@ void MpdStack::PushTrack(Int_t toBeDone, Int_t parentId, Int_t pdgCode, Double_t

      for (Int_t j = 0; j < npart; ++j) {
         if (j == 0) {
-           particle->SetStatusCode(11); // decayed particle
-           continue;                    // skip mother particle
+           // AZ-200324 particle->SetStatusCode(11); // decayed particle
+           particle->SetDaughter(0, 11); // AZ-200324 decayed particle
+           continue;                     // skip mother particle
         }
         // TParticle *part = (TParticle*) fDecays->UncheckedAt(j);
         TParticle *part = &vDecays[j];
         Int_t      toDo = 1;
-        if (TString(gMC->GetName()) == "TGeant4") toDo = 0;
-        // PushTrack(toBeDone, trackId, part->GetPdgCode(),
+        // AZ-210324 if (TString(gMC->GetName()) == "TGeant4") toDo = 0;
+        //  PushTrack(toBeDone, trackId, part->GetPdgCode(),
         PushTrack(toDo, trackId, part->GetPdgCode(), part->Px(), part->Py(), part->Pz(), part->Energy(), part-
>Vx(),
                   part->Vy(), part->Vz(), time, polx, poly, polz, proc, ntr, weight, is);
      }
-     if (parentId < 0) fNPrimaries += (npart - 1); // treat decay products as primaries (dirty trick)
+     // AZ-220324 if (parentId < 0) fNPrimaries += (npart - 1); // treat decay products as primaries (dirty trick)
   } else {
      // --> Push particle on the stack if toBeDone is set
      if (toBeDone == 1) fStack.push(particle);
@@ -270,6 +272,7 @@ TParticle *MpdStack::PopNextTrack(Int_t &iTrack)
   fCurrentTrack = thisParticle->GetStatusCode();
   iTrack        = fCurrentTrack;

+  if (thisParticle->GetDaughter(0) == 11) return (TParticle *)NULL; // AZ-220324
   return thisParticle;
 }
 // -------------------------------------------------------------------------
@@ -298,7 +301,8 @@ TParticle *MpdStack::PopPrimaryForTracking(Int_t iPrim)
    Fatal("MpdStack::PopPrimaryForTracking", "Not a primary track");
  }
```