Preapproval of MUO-24-001: Identification of soft muons using multivariate techniques at the CMS experiment in proton-proton collisions at $\sqrt{s}$ = 13.6 TeV

2024/06/03

Federica Maria Simone[1]
Dmytro Kovalskyi, Zhangqier Wang[2]
Priyanka Sadangi, Sanjay Kumar Swain[3]
Mia Liu, Jan Schulte, Yibo Zhong[4]

INFN Bari[1], MIT[2], NISER[3], Purdue University[4]

Overview

Aim: Machine Learning based MVA techniques are developed to improve soft muon identification in Run3

Documentation status:

- **CADI:** MUO-24-001
- **Analysis Note:** AN2023/172
- **Paper Draft:** MUO-24-001(v2)
- **Twiki:** SoftMuonMVA

CMS PAPER MUO-24-001

DRAFT
CMS Paper

The content of this note is intended for CMS internal use and distribution only

2024/05/24
Archive Hash: a13c0ff-D
Archive Date: 2024/05/24

Identification of soft muons using multivariate techniques at the CMS experiment in proton-proton collisions at $\sqrt{s} = 13.6\,\mathrm{TeV}$

The CMS Collaboration

Motivation

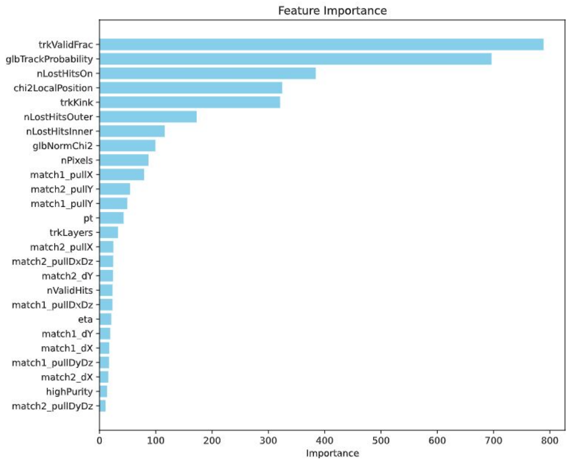
- Identifying soft muons can be challenging due to the significant contribution of “fake” muons.
  - Predominantly produced by the decay in flight of kaons and pions.
  - The contribution decreases exponentially with pT, becomes small by approximately 10 GeV.
  - Since those are real muons, separating them from signal is challenging
- The current soft muon MVA ID was initially developed using data from Run 2 collected in 2016. Because of the detector upgrade and running condition changes, it is not optimal for Run3, and we need to retrain a soft MVA ID for Run 3 analyses.
- The new ID benefits analyses involving soft muons (<10 GeV).
  - Focus on improving performance especially at very low pT (<4 GeV)

Sample and Models:

- Samples:
  - Inclusive Dilepton MinBias MC sample: enriched in both signal and background muons /InclusiveDileptonMinBias_TuneCP5Plus_13p6TeV_pythia8/Run3Summer22MiniAODv3-Pilot_124X_mcRun3_2022_realistic_v12-v5/MINIAODSIM
- Training Sample: preselection applied : “isGlobal || isTracker”
  - Background (decay in flight): muons from decays in flight of pions and kaons
  - Signal (genuine muons) : muons originating from heavy flavor decays (or J/Psi)
- Models:
  - A Deep Neural Networ (DNN) and two flavors of Boosted Decision Trees (BDTs), HGB, and XGBoost (XGB), are studied, XGB is selected in the end.
  - Input Variables:
    - DNN, HGB : 42 variables
    - XGB : 26 variables

Input Variables for XGB model

- Use variables related to quality of the inner and standalone muon tracks, as well as the matching between the two and the the global track
- 26 variables are used in the Xgb model.
- Feature Importance:

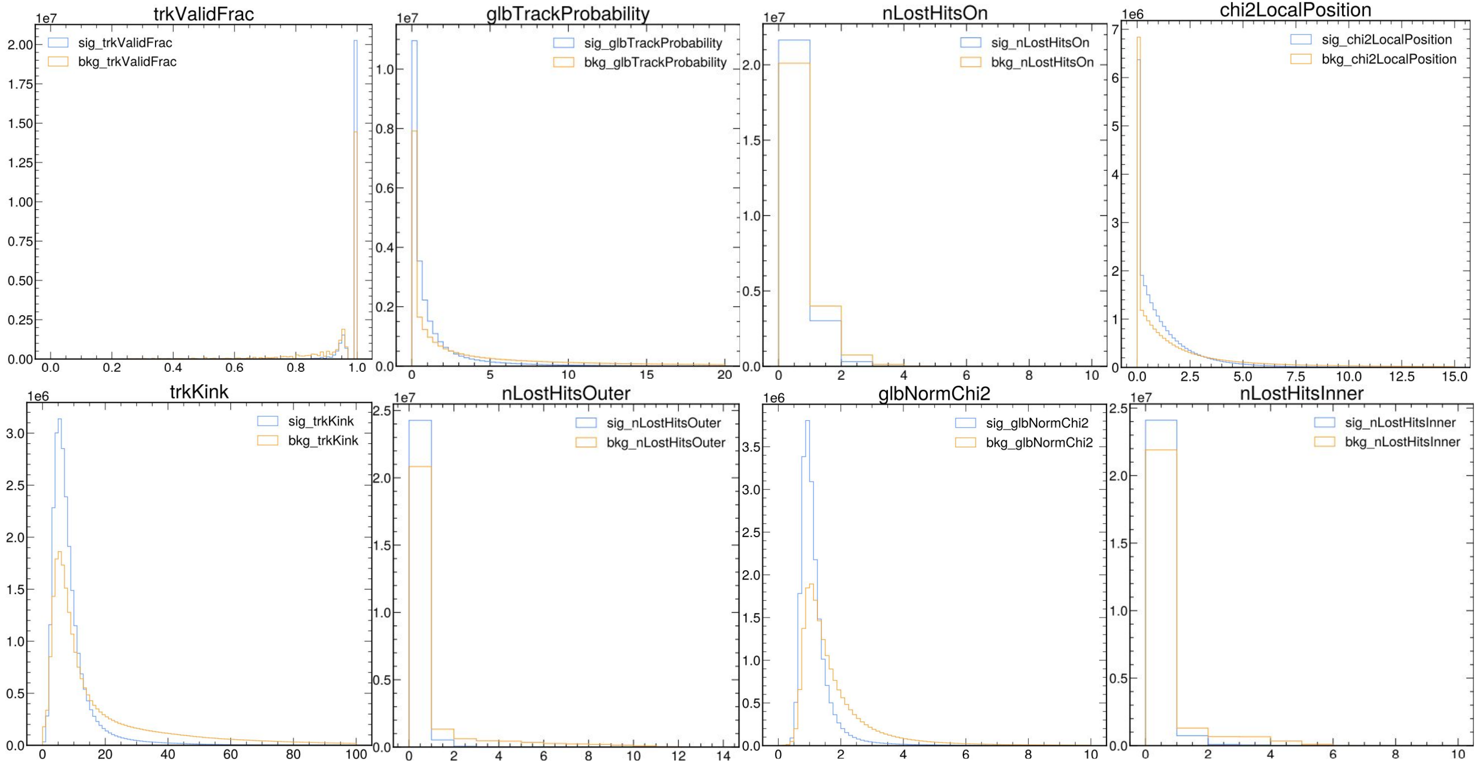


- Distributions for the most important features are shown on the next slide, but the others are in the backup.

“pt”,
"eta",
"trkKink",
"glbTrackProbability",
"chi2LocalPosition",
"glbNormChi2",
"trkValidFrac",
"match1_dX",
"match1_pullX",
"match1_pullDxDz",
"match1_dY",
"match1_pullY",
"match1_pullDyDz",
"match2_dX",
"match2_pullX",
"match2_pullDxDz",
"match2_dY",
"match2_pullY",
"match2_pullDyDz",
"nPixels",
"nValidHits",
"nLostHitsInner",
"nLostHitsOn",
"nLostHitsOuter",
"trkLayers",
"highPurity"

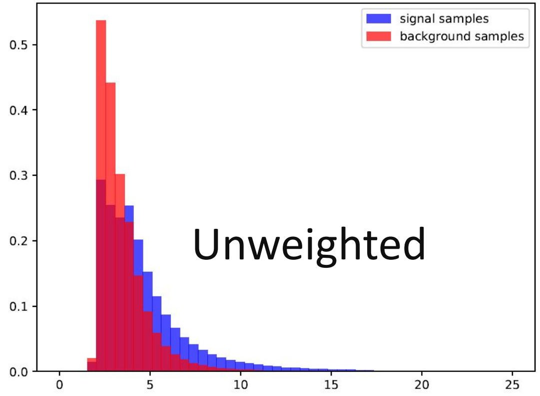
Input feature distributions

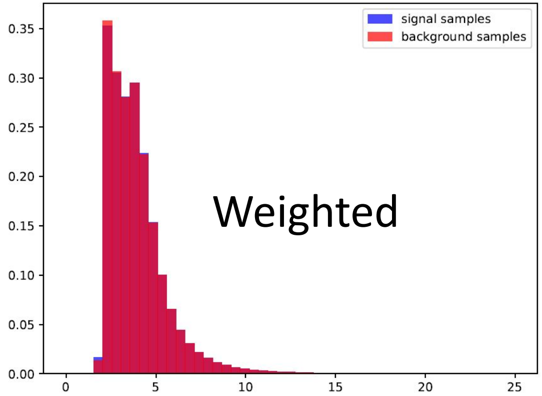
- ❖ Shown are the 8 most sensitive inputs
- ❖ Background has real muons -> Differences not huge, rely on correlated differences in many input variables



p_T distribution reweighting

- The bkg and signal have different dependence on p_T.
  - As we want this ID to be general purpose, we don't want the model to associate higher p_T with signal
- We reweight the samples to have matching p_T distributions to erase the impact.
- Assign a weight value to each muon based on (p_T,η)
- SF(p_T,η) = signal(p_T,η)/background(p_T,η)
- No direct weight functionality in training tools. Workaround is to randomly drop muons from the training samples
- Generate random value x between [0,1]
  - If SF>1: drop from signal sample when x>1/SF
  - If SF<1: drop from background sample when x>SF
- N.B.: Still found it to be useful to keep p_T as an input feature due to its correlations with other variables





Model and Training

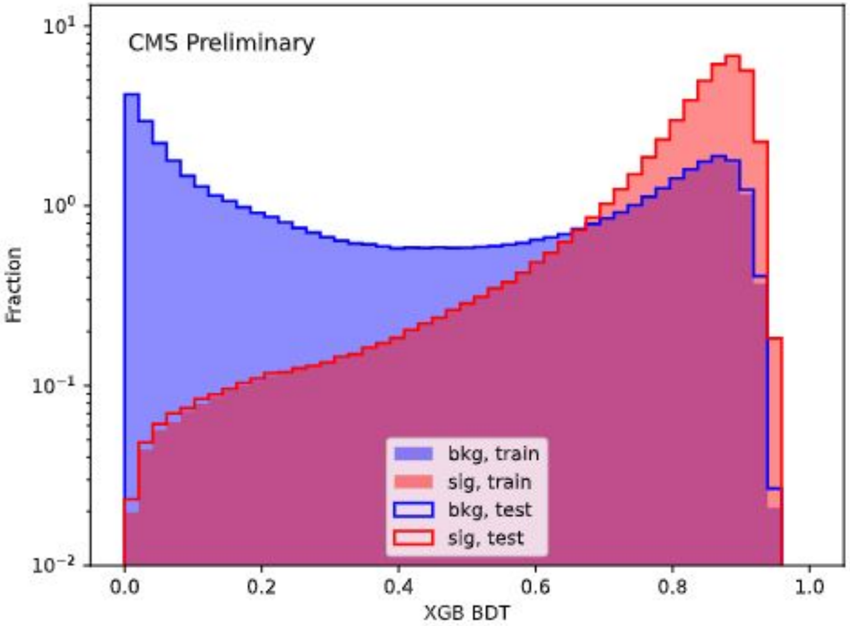
- Described here is the final model, information for the other two models we tested are in the backup
- Final model is a gradient boosted decision tree implement in XGBoost
- For the training, Scikit-learn is used.
- 10M muons are used for the training and another 5M are used for the validation to avoid overtraining.
- The parameters is tuned manually to achieve the best performance (AUCROC).
- Most of the parameters take the default values except for the following:

parameter	setting
eta	0.01
max_depth	10
silent	1
eval_metric	auc
subsample	0.6
alpha	8.0
gamma	2.0
lambda	1.0
min_child_weight	0.00001
colsample_bytree	1.0

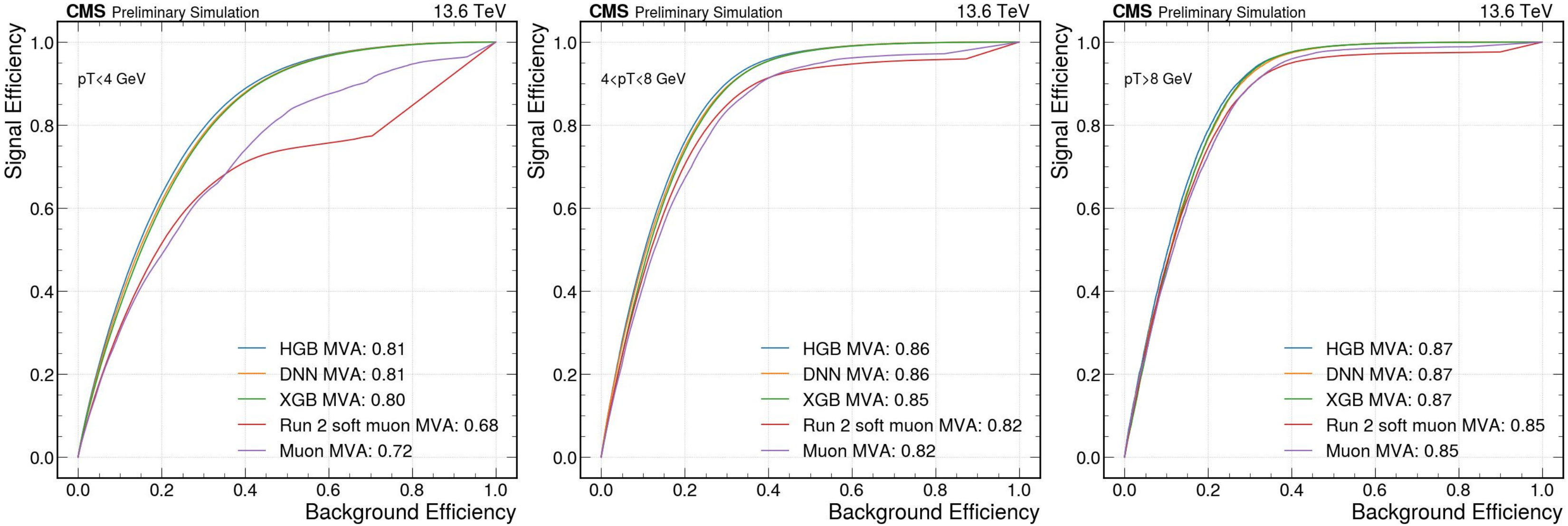
Check for overtraining

❖ To check for signs of overtraining, compare the distribution of the classifier score in the training and validation samples

❖ The model has nearly identical performance in both samples
-> No hint of overtraing



Performance



- Test on muons selected from Inclusive Dilepton MC sample.
  - size for the three p_T windows: ~2M, ~1M, and ~ 0.2 M
- All Run3 MVAs are better than the Run 2 MVA, especially in the low pT range.
- The Run3 models have similar performance.
  - We chose the XGB model as it uses fewer inputs and was easier to integrate in CMSSW

Working Points

- ❖ In Run3, 4 working points are recommended.
- ❖ Medium working point is chosen to have ~same background rejection as the Run 2 MVA working point available in MiniAOD
- ❖ Other working points chosen so that the background rate scales by roughly a factor of 2 between them
  - ➢

name	classifier threshold
soft	> 0.47
loose	> 0.63
medium	> 0.74
tight	> 0.83

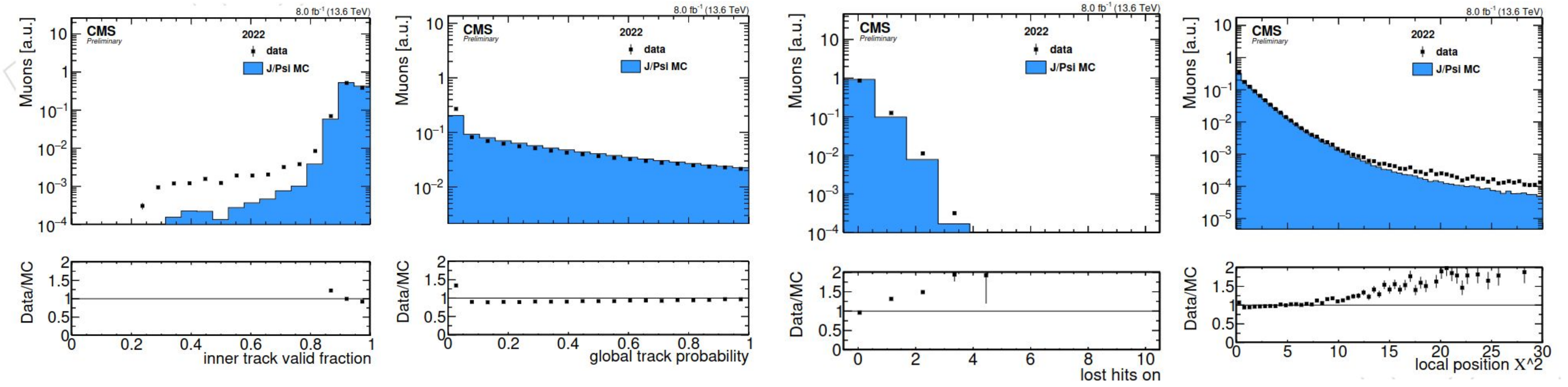
Validation in data

- ❖ Run 3 soft muon MVA shows good performance in MC
- ❖ Validated in Run 3 collision data in three ways:
  - ➢ Data/MC comparisons of input variables and classifier scores
  - ➢ Efficiency to select signal muons is measured on J/Psi events
  - ➢ Background rate measured separately for $\pi \to \mu$ and $K \to \mu$ decays
- ❖ Full 2022 + 2023 datasets used

Data/MC comparisons

variable	selection
	Tag selection
p_T	$> 8\,\text{GeV}$
$\lvert\eta\rvert$	< 2.4
ID	isTightMuon
trigger	HLT_Mu8 OR HLT_Mu15 OR HLT_Mu17 OR HLT_Mu19 OR HLT_Mu20 OR HLT_IsoMu24
	Probe selection
p_T	$> 2\,\text{GeV}$
$\lvert\eta\rvert$	< 2.4
dZ	< 0.5
ID	isTrackerMuon OR isGlobalMuon
	Pair selection
charge	opposite sign
$\Delta R(\text{tag}, \text{probe})$	> 0.3
$m_{\mu\mu}(\text{tag}, \text{probe})$	$2.50 - 3.70\,\text{GeV}$

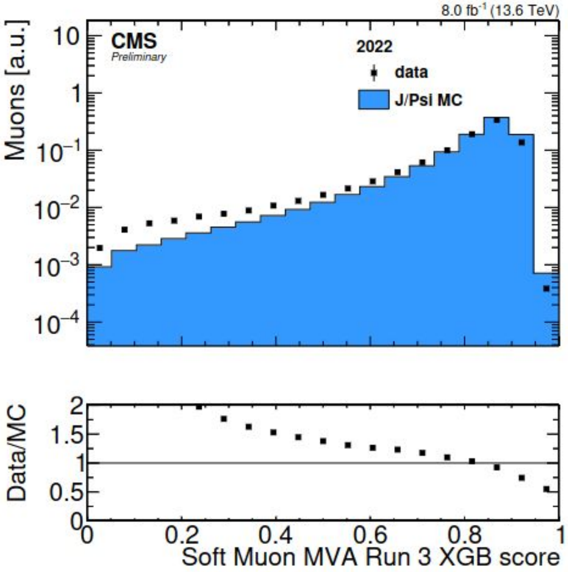
- Use J/Psi[1] events selected for efficiency measurement
  - Selection conditions are listed:
  - We are comparing to J/Psi MC, and backgrounds are not subtracted from DATA.
  - We know that pT spectrum in data is somewhat harder than in the J/Psi sample we use
- Generally, we observe quite good agreement in the modeling of all input variables
  - Shown here is 2022, very similar performance in other data taking periods
- Data/MC comparisons not intended for publication

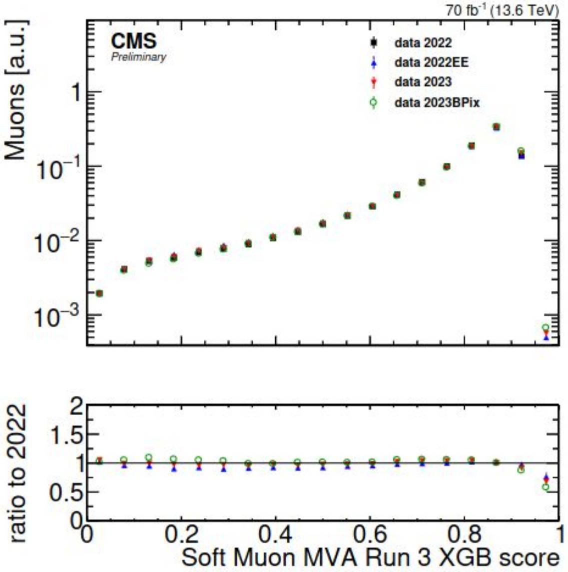


[1] /JpsiTo2Mu_JpsiPt8_TuneCP5_13p6TeV_pythia8/

Data/MC comparisons

- Classifier score is overall quite well modeled by MC
- Data performance expected to be very similar to simulation
- Classifier score distribution is the same for different data taking eras -> Stable performance throughout Run 3, no need for individual trainings for different detector conditons

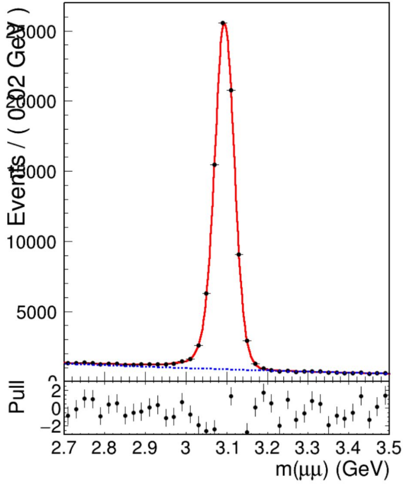


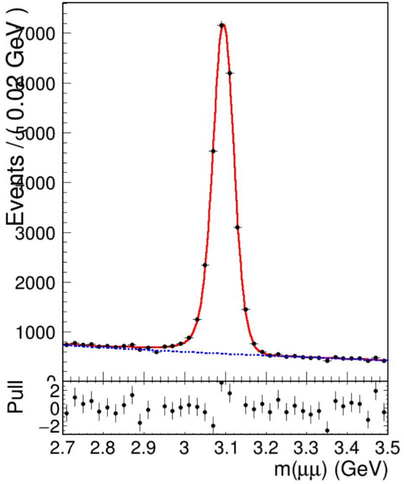


Efficiency Measurement

- Efficiency is measured on 2022 + 2023 data using Tag & Probe on $J/\psi \to \mu^+\mu^-$ events
- POG spark tool is used with settings close to what is used for POG SFs for low-p_T muons
- Different probe selection to match selection used for training of the model

variable	selection
Tag selection	
p_T	$> 8\,\text{GeV}$
$\|\eta\|$	< 2.4
ID	isTightMuon
trigger	HLT_Mu8 OR HLT_Mu15 OR HLT_Mu17 OR HLT_Mu19 OR HLT_Mu20 OR HLT_IsoMu24
Probe selection	
p_T	$> 2\,\text{GeV}$
$\|\eta\|$	< 2.4
dZ	< 0.5
ID	isTrackerMuon OR isGlobalMuon
Pair selection	
charge	opposite sign
$\Delta R(\text{tag, probe})$	> 0.3
$m_{\mu\mu}(\text{tag, probe})$	$2.50 - 3.70\,\text{GeV}$

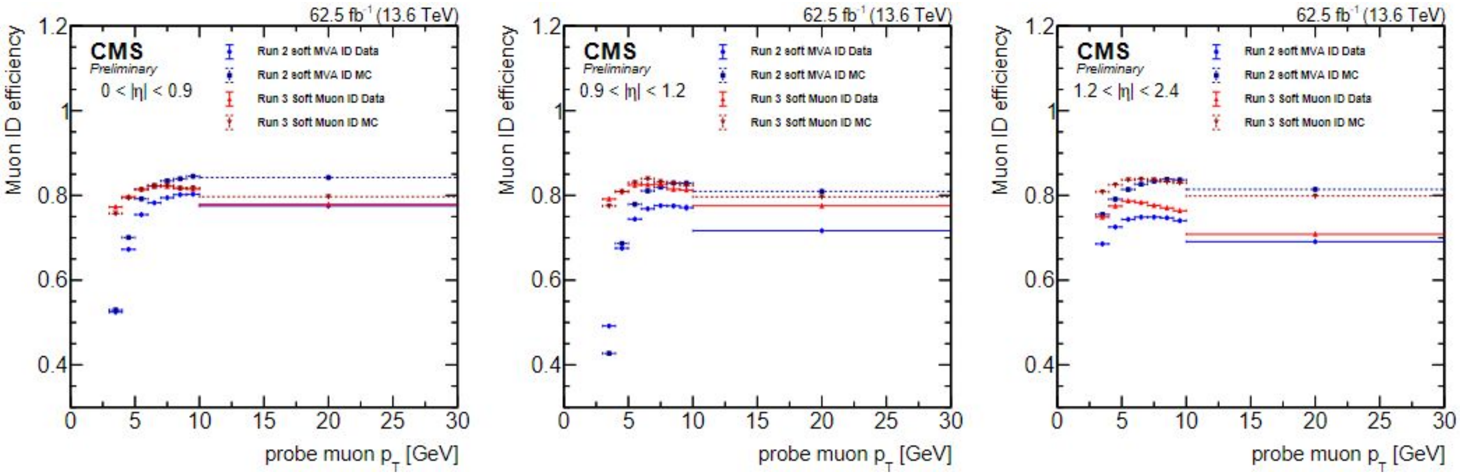




- Fits done to the passing probe and passing + failing probe distributions to extract efficiency
- Signal shape based on MC template
- Background modeled with CMS shape

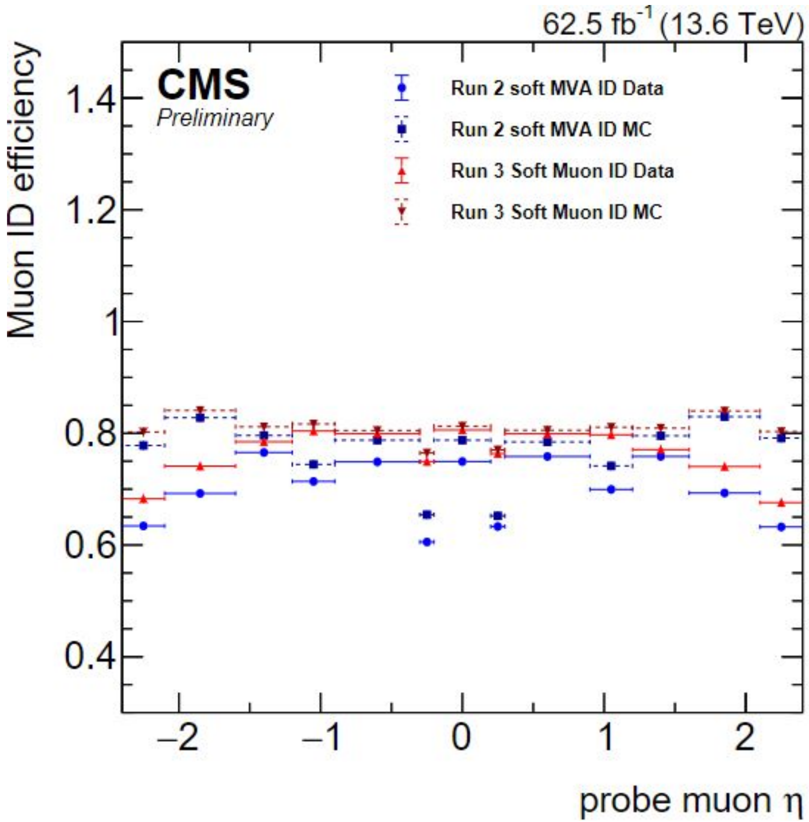
Efficiency Measurement

- Compare Run 2 and Run 3 MVA using medium WP for Run 3 MVA (~same background rejection)
- Run3 is better than Run2 at low p_T range in all η bins.
- Run 3 MVA shows some efficiency drop towards higher p_T
  - Still better than Run 2
  - p_T range not very relevant for analyses using this ID
- Much better Data/MC agreement for Run 3 MVA, except at high η due to CSC out of synch issues



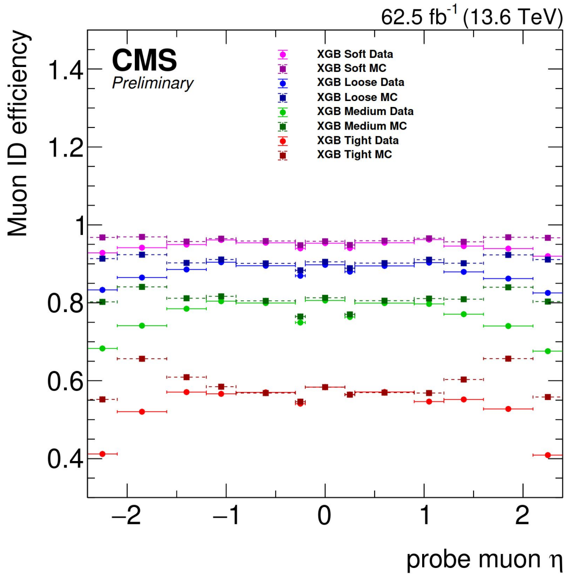
Efficiency Measurement

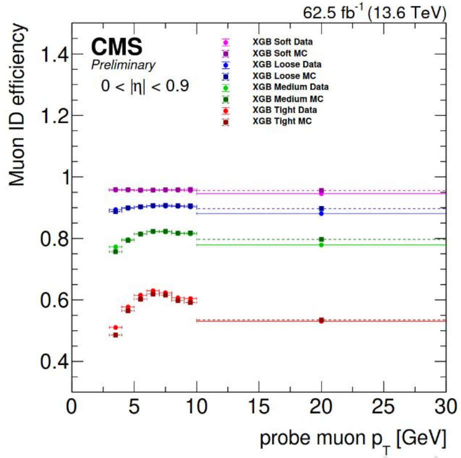
- Efficiency integrated over p_T as a function of η
- Confirms observations from previous slide

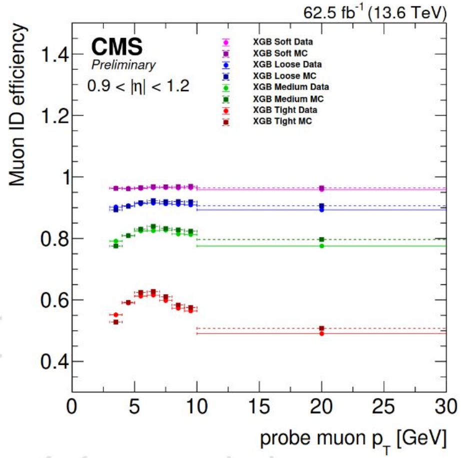


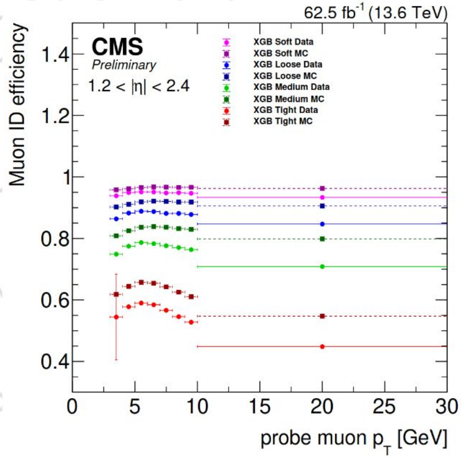
Efficiency Measurement

- Efficiency for all four WPs of the Run 3 soft muon MVA
- Efficiency reduces by ~factors of 2 going from one WP to the next
- Good Data/MC agreement for all WPs for central muons
- Effect of CSC issues on data efficiency visible in the endcaps, more pronounced for tighter WPs



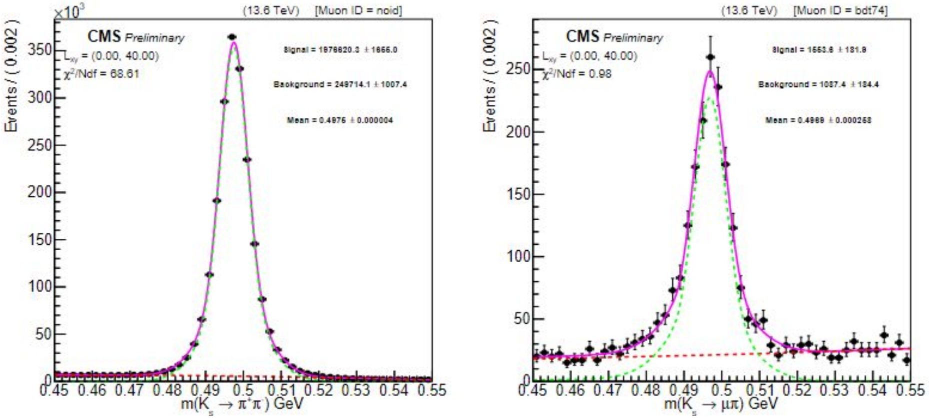






Background Rate

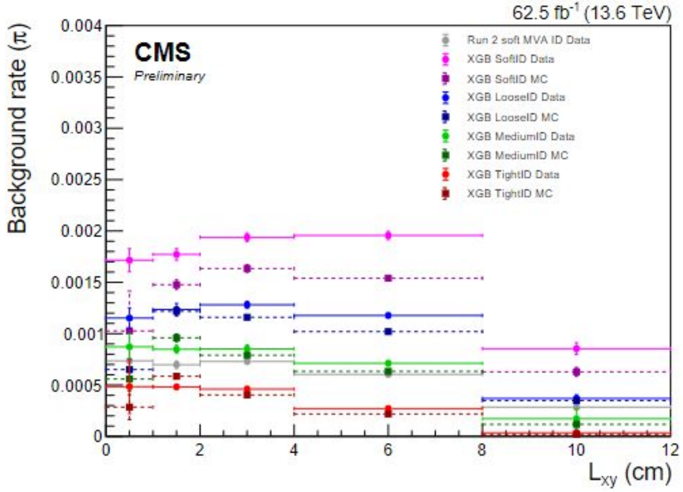
- Background rate is measured in data and simulation for muons from π and K decays, separately.
- MC samples used :
  - DY samples
    - DYto2L-2Jets_MLL-10to50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
    - DYto2L-2Jets_MLL-50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
    - DYJetsToLL_M-50_TuneCP5_13p6TeV-madgraphMLM-pythia8
    - DYto2E_M-50_NNPDF31_TuneCP5_13p6TeV-powheg-pythia8
  - WtoLNu-2Jets_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
  - TTTo2J1L1Nu_CP5_13p6TeV_powheg-pythia8
- Background rate for π → μν decays:
  - We perform a fit to ππ and μπ invariant mass distribution to extract the event yield of $K^0_S \to \mu\pi$ and $K^0_S \to \pi^+\pi^-$
  - A Crystall Ball added to a Gaussian centered at the same mean and a 2nd order Bernstein polynomial function are used to fit signal and combinatorial background component.
  - We distinguish the cases of all and those where one is reconstructed as a $K^0_S \to \pi^+\pi^-$ng a working point of the soft muon MVA. The background rate is defined as the ratio of number of μ to the total number of π.

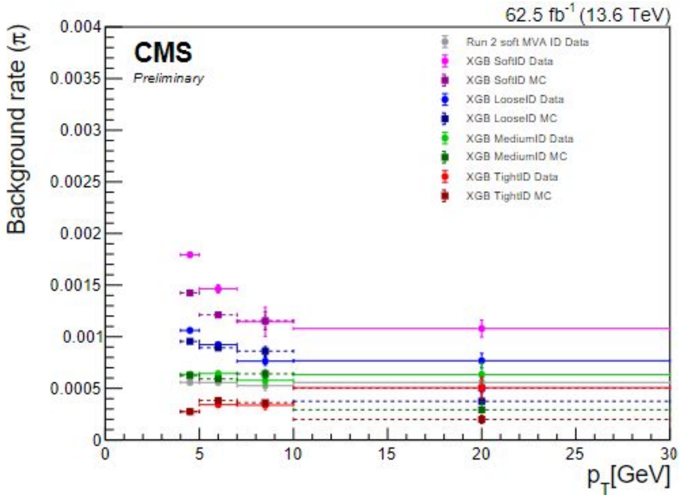


e.g. 2022 DATA Fitting before (left) and after (right) passing the Medium ID.

Background Rate

- $\pi \rightarrow \mu\nu$ Background rate is plotted in a function of pion flight length L_{xy} and muon p_T .



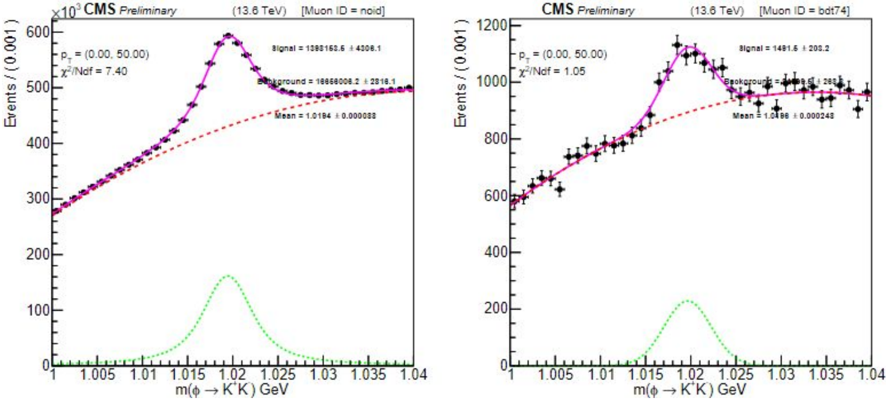


Data taking period	Soft XGB ID	Loose XGB ID	Medium XGB ID	Tight XGB ID
2022	1.25 ± 0.08	1.14 ± 0.09	1.11 ± 0.15	1.17 ± 0.15
2023	1.22 ± 0.07	1.10 ± 0.13	1.05 ± 0.16	1.07 ± 0.13

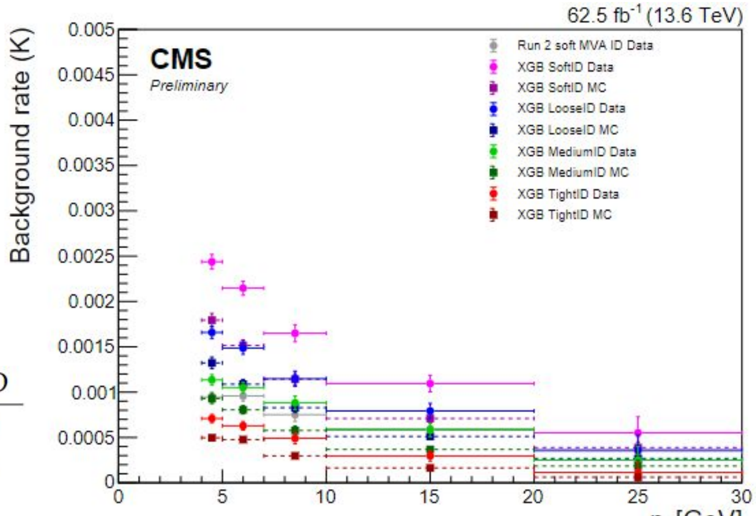
DATA/MC ratios for the $\pi \rightarrow \mu$ fake rate

Background Rate

- Background rate for K → μν decays:
  - We perform a fit to K+K− and μK invariant mass distribution to extract the event yield of Φ → K+K− and Φ → μK events.
  - A Voigtian function and a 2nd order Bernstein polynomial function are used to parameterize the signal and background components, respectively.
  - Calculate the ratio of number of muons and total number of kaons.
  - Background rates for muons from Kaons of different working points of the Run 3 softmuon MVA for 2022-2023 data and MC are plotted as a function of muon pT.



e.g. 2022 data before (left) and after(right) medium ID cut.



All 2022 and 2023 DATA and MC

Data taking period	Soft XGB ID	Loose XGB ID	Medium XGB ID	Tight XGB ID
2022	1.08 ± 0.24	0.88 ± 0.20	0.93 ± 0.24	0.94 ± 0.30
2023	1.61 ± 0.45	1.15 ± 0.40	0.96 ± 0.41	0.66 ± 0.66

DATA/MC ratios for the K → μ fake rate

Summary

- Multivariate techniques based on Machine Learning were developed for the Run3 softMuon identification: A XGBoost model was selected in the end.
- The new Run3 ID has a general better performance than the Run2 ID especially for the low p_T region.
- Four working points, soft, loose, medium, and tight are recommended.
- Background rate was studied for the Run3 ID.
- Run3 ID has higher efficiency than the Run2 ID when back rejections are similar.
- The documentation is in place.
- We would like to kindly ask for pre-approval.

BackUp

Model Details of DNN and HGB

DNN: 8 hidden layers , 256 nodes, dropout=0.2,

Forward: Linear, Activation: LeakyReLU

Loss: $$\text{FocalLoss} = -\sum_{i=1}^{i=n} \alpha_i (i - p_i)^\gamma \log_b(p_i)$$

HGB: sklearn

max_iter=1200, max_depth=10, other parameters take default values if any.

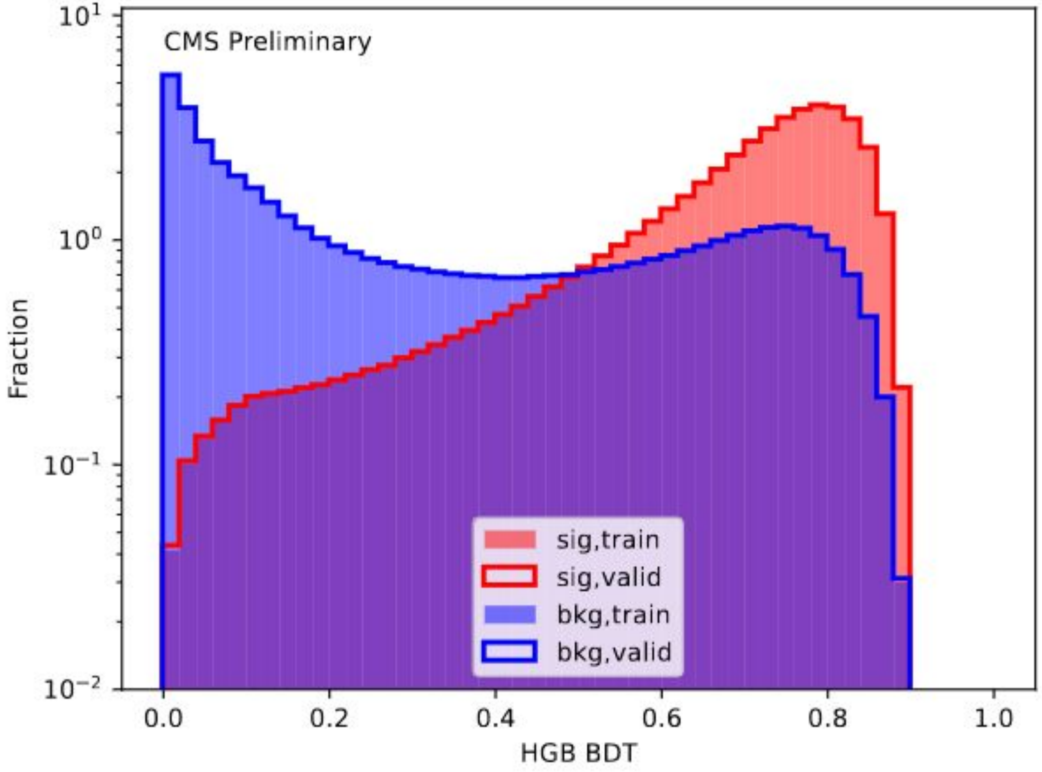
Xgboost: the same set up used in the Bmm study:

parameter	setting
eta	0.01
max_depth	10
silent	1
eval_metric	auc
subsample	0.6
alpha	8.0
gamma	2.0
lambda	1.0
min_child_weight	0.00001
colsample_bytree	1.0

Training for DNN and HGB

- After reweighting, balanced Dataset (35M signal + 35M background), is used for DNN and HGB, and 5/7 is used for training and 2/7 for validation
- For the training, we use pytorch and Scikit-learn
- In the training of all the models, we tune the parameters manually to achieve the best performance.
- Reweight :
  - HGB: used the same reweight methods as XGB
  - DNN: Define weight=1/SF for signal, weight=1 for background. and Multiply the weight in the loss function.

Classifier score for HGB



Input Variables for DNN and HGB

"eta",

"charge",

"chargeProduct",

"isGlobal",

"isTracker",

"isStandalone",

"isPF",

"trkKink",

"glbTrackProbability",

"chi2LocalPosition",

"glbNormChi2",

"trkValidFrac",

"chi2LocalMomentum",

"trkRelChi2",

"staRelChi2",

"trkNormChi2",

"staNormChi2",

"nStations",

"match1_dX",

"match1_pullX",

"match1_pullDxDz",

"match1_dY",

"match1_pullY",

"match1_pullDyDz",

"match2_dX",

"match2_pullX",

"match2_pullDxDz",

"match2_dY",

"match2_pullY",

"match2_pullDyDz",

"nPixels",

"segmentComp",

"nValidHits",

"nLostHitsInner",

"nLostHitsOn",

"nLostHitsOuter",

"trkLayers",

"trkLostLayersInner",

"trkLostLayersOn",

"trkLostLayersOuter",

"muonStationsWithValidHits",

"highPurity"

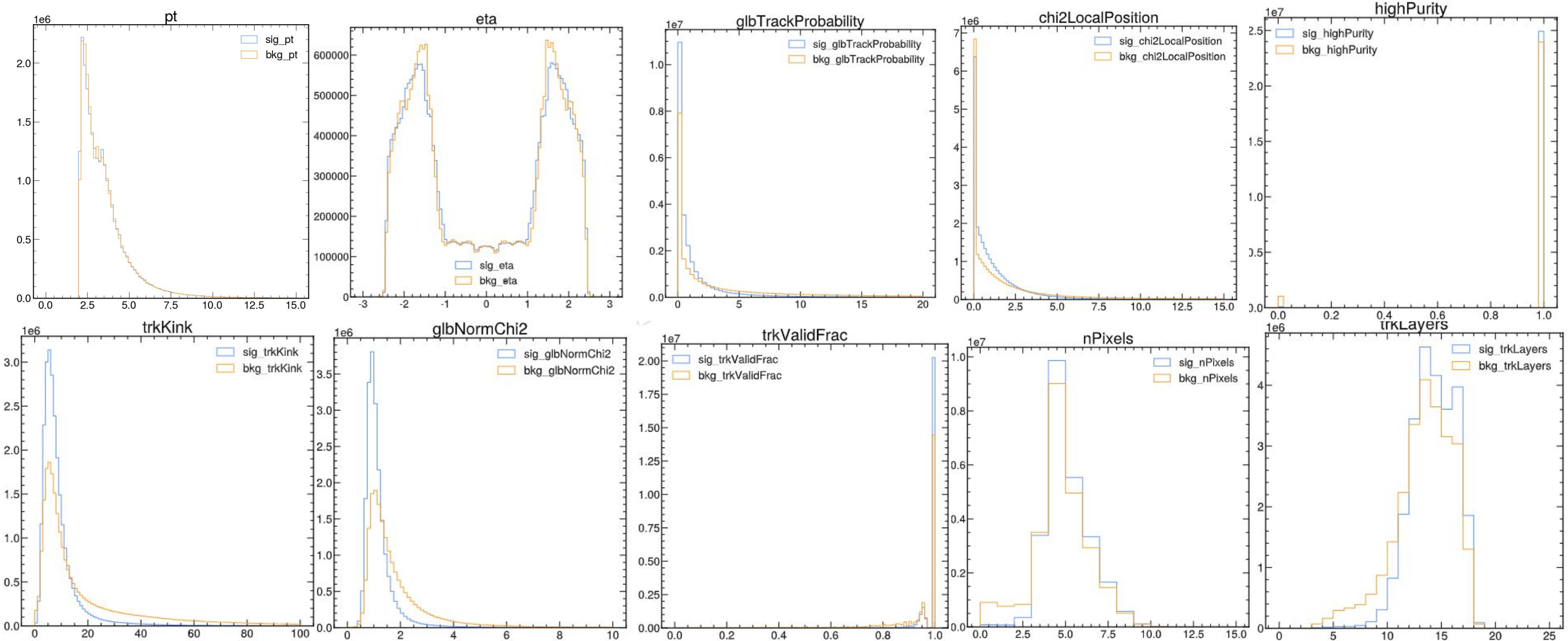
Run 2 soft MVA

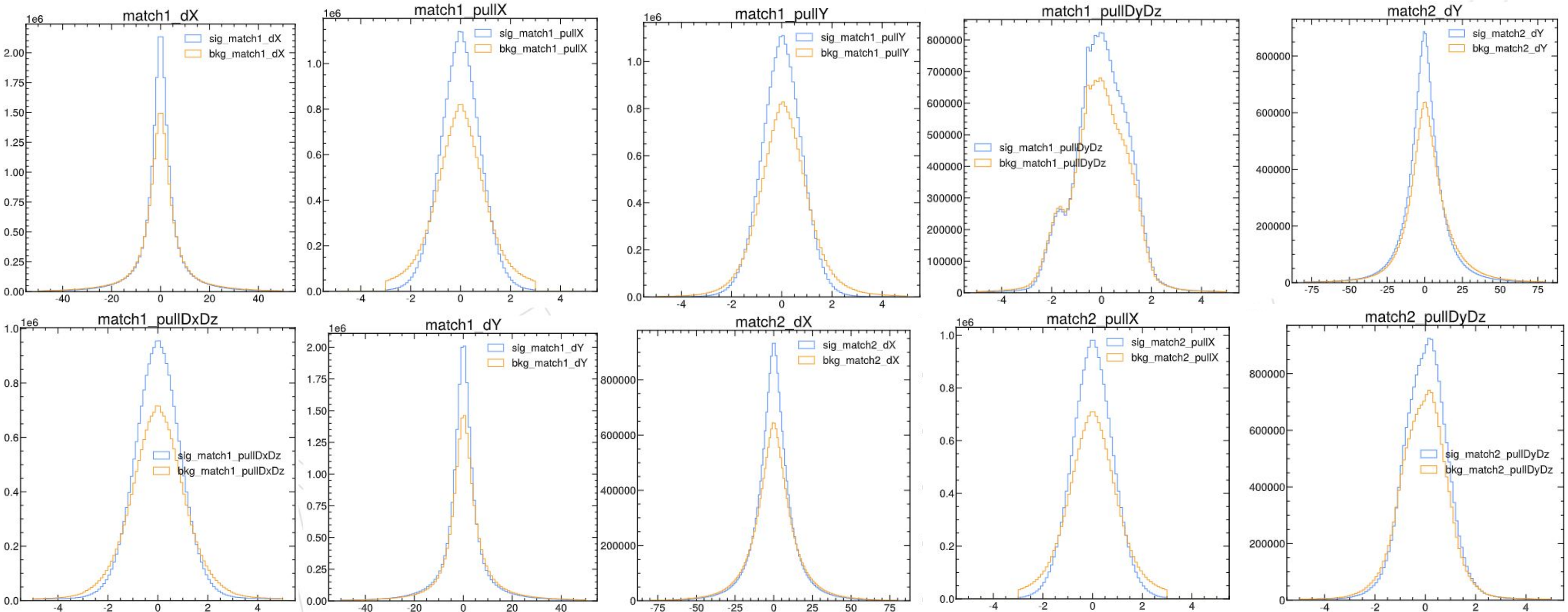
- Run 2 version of the soft MVA was trained on 2016 data
- Training dataset used global muons with pT > 4 GeV, |eta| < 1.4 where the inner track passes highPurity
- BDT trained with TMVA

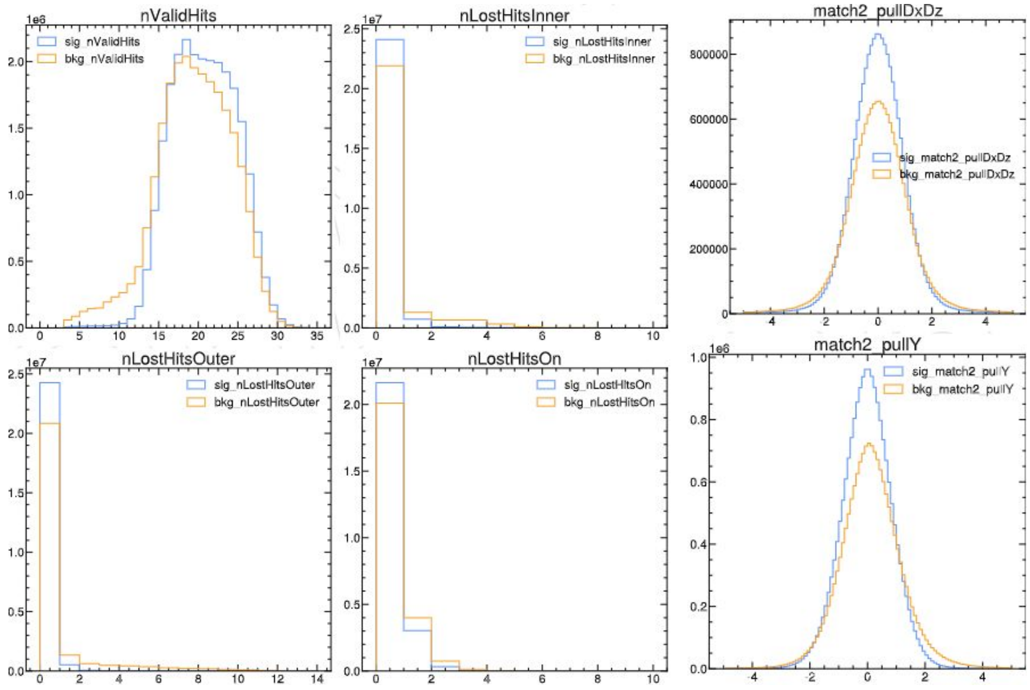
Table 4: Settings for BDT training in TMVA for the Run 2 soft muon MVA.

parameter	setting
nCuts	-1
MinNodeSize	1.5%
boostType	RealAdaBoost
adaBeta	0.3
muonFactor	2
maxDepth	12

Variables distribution in the training sample




match1_dX
sig_match1_dX
bkg_match1_dX
match1_pullX
sig_match1_pullX
bkg_match1_pullX
match1_pullY
sig_match1_pullY
bkg_match1_pullY
match1_pullDyDz
sig_match1_pullDyDz
bkg_match1_pullDyDz
match2_dY
sig_match2_dY
bkg_match2_dY
match1_pullDxDz
sig_match1_pullDxDz
bkg_match1_pullDxDz
match1_dY
sig_match1_dY
bkg_match1_dY
match2_dX
sig_match2_dX
bkg_match2_dX
match2_pullX
sig_match2_pullX
bkg_match2_pullX
match2_pullDyDz
sig_match2_pullDyDz
bkg_match2_pullDyDz



nValidHits
sig_nValidHits
bkg_nValidHits
nLostHitsInner
sig_nLostHitsInner
bkg_nLostHitsInner
match2_pullDxDz
sig_match2_pullDxDz
bkg_match2_pullDxDz
nLostHitsOuter
sig_nLostHitsOuter
bkg_nLostHitsOuter
nLostHitsOn
sig_nLostHitsOn
bkg_nLostHitsOn
match2_pullY
sig_match2_pullY
bkg_match2_pullY

Samples for J/Psi validation

	2022
MC	/JpsiTo2Mu_JpsiPt8_TuneCP5_13p6TeV_pythia8/Run3Summer22MiniAODv3-MUO_POG_124X_mcRun3_2022_realistic_v12-v2/MINIAODSIM
Data	/SingleMuon/Run2022B-22Sep2023-v1/MINIAOD /SingleMuon/Run2022C-22Sep2023-v1/MINIAOD /Muon/Run2022C-22Sep2023-v1/MINIAOD /Muon/Run2022D-22Sep2023-v1/MINIAOD
	2022EE
MC	/Jpsito2Mu_JpsiPT8_TuneCP5_13p6TeV_pythia8/Run3Summer22EEMiniAODv3-MUO_POG_124X_mcRun3_2022_realistic_postEE_v1-v2/MINIAODSIM
Data	/Muon/Run2022E-22Sep2023-v1/MINIAOD /Muon/Run2022F-PromptReco-v1/MINIAOD /Muon/Run2022G-PromptReco-v1/MINIAOD
	2023
MC	/Jpsito2Mu_JpsiPT8_TuneCP5_13p6TeV_pythia8/Run3Summer23MiniAODv4-MUO_POG_130X_mcRun3_2023_realistic_v14-v2/MINIAODSIM
Data	/Muon0/Run2023B-22Sep2023-v1/MINIAOD /Muon1/Run2023B-22Sep2023-v1/MINIAOD /Muon0/Run2023C-22Sep2023_v1-v1/MINIAOD /Muon0/Run2023C-22Sep2023_v2-v1/MINIAOD /Muon0/Run2023C-22Sep2023_v3-v1/MINIAOD /Muon0/Run2023C-22Sep2023_v4-v1/MINIAOD /Muon1/Run2023C-22Sep2023_v1-v1/MINIAOD /Muon1/Run2023C-22Sep2023_v2-v1/MINIAOD /Muon1/Run2023C-22Sep2023_v3-v1/MINIAOD /Muon1/Run2023C-22Sep2023_v4-v1/MINIAOD
	2023BPix
MC	/Jpsito2Mu_JpsiPT8_TuneCP5_13p6TeV_pythia8/Run3Summer23BPixMiniAODv4-MUO_POG_130X_mcRun3_2023_realistic_postBPix_v2-v2/MINIAODSIM
Data	/Muon0/Run2023D-22Sep2023_v1-v1/MINIAOD /Muon0/Run2023D-22Sep2023_v2-v1/MINIAOD /Muon1/Run2023D-22Sep2023_v1-v1/MINIAOD /Muon1/Run2023D-22Sep2023_v2-v1/MINIAOD

Datasets for background rate measurement

- EGamma(2022) EGamma0(2023) EGamma1(2023)
- ParkingDoubleElectronLowMass
- DY samples
  - DYto2L-2Jets_MLL-10to50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
  - DYto2L-2Jets_MLL-50_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
  - DYJetsToLL_M-50_TuneCP5_13p6TeV-madgraphMLM-pythia8
  - DYto2E_M-50_NNPDF31_TuneCP5_13p6TeV-powheg-pythia8
- WtoLNu-2Jets_TuneCP5_13p6TeV_amcatnloFXFX-pythia8
- TTTo2J1L1Nu_CP5_13p6TeV_powheg-pythia8

Event selection for π -> μ background rate

- $p_{\mathrm{T}}^{\pi_1} > 1\,\mathrm{GeV}$
- $p_{\mathrm{T}}^{\pi_2} > 4\,\mathrm{GeV}$
- pion track is highPurity
- $m_{\mathrm{K}^0_\mathrm{S}} \in [0.45, 0.55]$
- flight length significance $\ell_{xy}/\sigma(\ell_{xy}) > 3$
- Vertex probability greater than 0.001
- vertex displacement in XY plane wrt Beam Spot less than 8
- cosine of pointing angle in XY wrt BS greater than 0.999
- impact parameter significance of the candidate trajectory in 3D wrt PV less than 3
- 2D impact parameter significance for Track 1 and 2 wrt Beam Spot greater than 5
- kinematic vertex fit χ^2/dof of the two track less than 3
- Fire the HLT_Electron30 trigger

Event selection for K -> μ background rate

- $p_{\mathrm{T}}^{K} > 3\,\mathrm{GeV}$
- kaon track is highPurity
- $m_{\phi} \in [1.00, 1.04]$
- vertex displacement in XY plane wrt Beam Spot less than 4
- vertex probability greater than 0.3
- impact parameter significance of the candidate trajectory in 3D wrt PV less than 1
- distance of closest approach of tracks less than 0.004

DATA/MC comparison