Updates pT Training: KNN

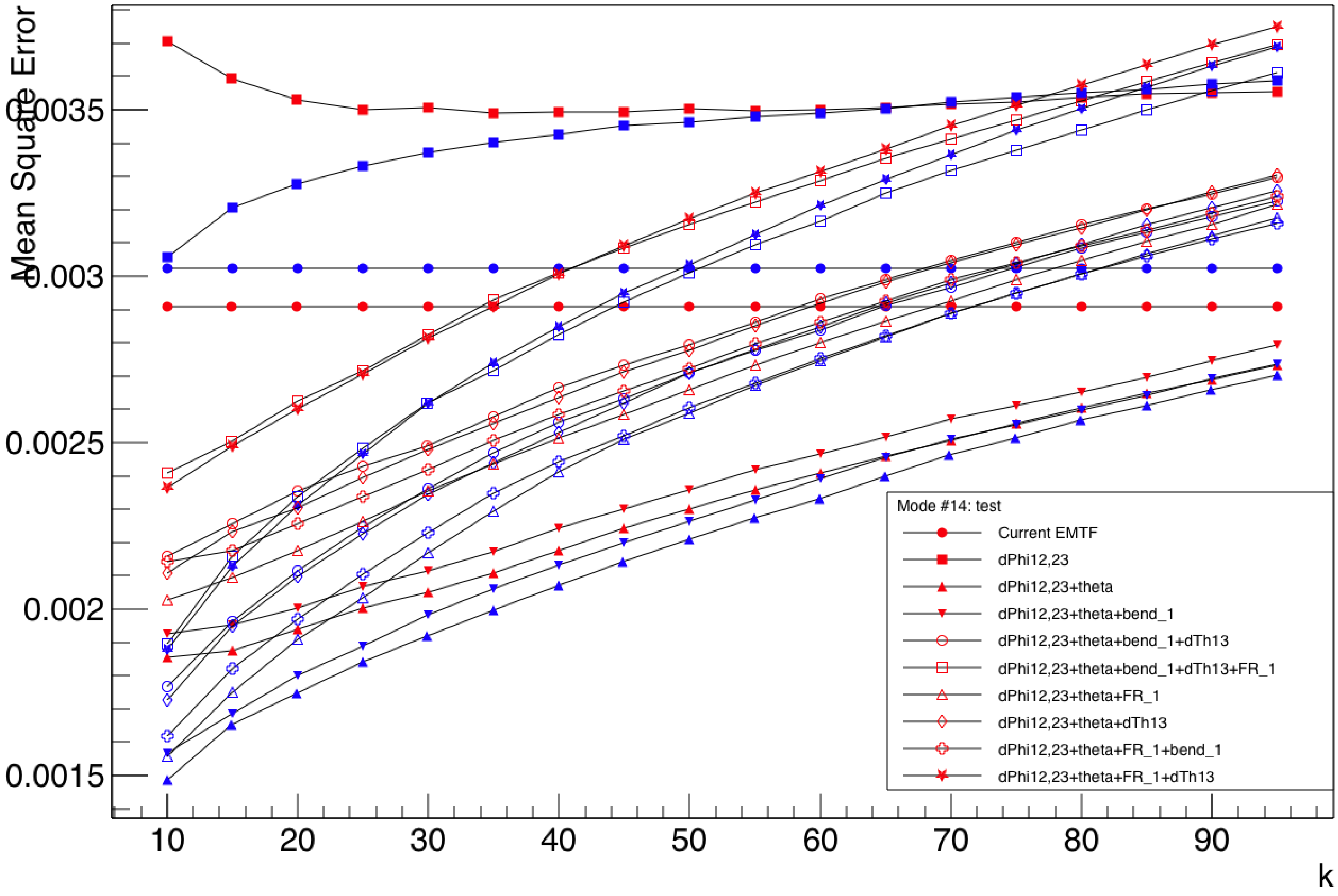
Wei Shi

weishi@rice.edu

EMTF Working Meeting

Mode 14: Station 1-2-3

- Add compare to current EMTF pT(LUT v1)
- Add input variables based on 2016 LUT(dPhi12, 23; theta; CLCT1; dTheta13; FR 1)(keep scale frac 0.8)
- Tune KNN scale frac so as performance is best
- Add input combinations of dPhis, ring number of station 1
- Divide up low/high pT
- Higher statistics[optional]
- Question: what's FR bit? What's bend_1?(integer)
- Weight 1/pT when doing Standard Deviation calculation?



Mean Square Error
0.0035
0.003
0.0025
0.002
0.0015
10
20
30
40
50
60
70
80
90
k
Mode #14: test
Current EMTF
dPhi12,23
dPhi12,23+theta
dPhi12,23+theta+bend_1
dPhi12,23+theta+bend_1+dTh13
dPhi12,23+theta+bend_1+dTh13+FR_1
dPhi12,23+theta+FR_1
dPhi12,23+theta+dTh13
dPhi12,23+theta+FR_1+bend_1
dPhi12,23+theta+FR_1+dTh13

BACK UP

Steps

- Find optimized input/target/k/scalfrac for KNN for other modes
- Select the better one method using for other modes: change inputs
- Include RPC hits and repeat 1-3 step
- Remove/truncate to fit in 29 bits
- Repeat for other modes
- Train charge assignment(classification)

Mode 14

- # of train events: 18746(4)
- # of test events: 18747(2)

[0]Current EMTF performance:[marker 20]

- EMTF Standard Deviation of train sample: 0.0030236
- EMTF Standard Deviation of test sample: 0.00290853

Euclidean distance

x: input variables(dimension d)
of train event
y: input variables of test event

$$R_{\text{rescaled}} = \left(\sum_{i=1}^{d} \frac{1}{w_i^2} |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

$$\langle t(i,V) \rangle = \frac{\sum_{j \in V} w_j t_j f(\mathrm{dis}(i,j))}{\sum_{j \in V} w_j f(\mathrm{dis}(i,j))}$$

[1]dPhis_2_inv_pT_(UseKernel)

- [marker 21]

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.0037048	0.00359	0.00352443	0.00349985	0.00350756	0.00349082	0.00349363	0.00349395	0.0035017
MSE(train)	0.00305859	0.0032069	0.00327449	0.00332954	0.00337016	0.00340326	0.00342645	0.00345231	0.00346173
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.00349819	0.00350048	0.00350693	0.00351632	0.00352492	0.00353682	0.00354709	0.00354961	0.00355508
MSE(train)	0.00348042	0.00348979	0.00350404	0.00352481	0.00353682	0.00354927	0.00356243	0.00357705	0.00358805

[2]dPhis_theta_2_inv_pT_UseKernel_ScalFrac_0p8

- [marker 22]

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00185647	0.00187403	0.00193868	0.00199863	0.00204975	0.00210894	0.00217745	0.0022427	0.00230163
MSE(train)	0.00148462	0.00164991	0.00174524	0.00183729	0.00191387	0.00199452	0.00207228	0.00213605	0.00220639
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.0023579	0.00240821	0.00245871	0.0025071	0.00255731	0.00260458	0.0026495	0.00269085	0.00273436
MSE(train)	0.00227102	0.00232995	0.00239768	0.00245809	0.00251172	0.00256187	0.00260985	0.00265738	0.00270102

[3]dPhis_theta_CLCT1_2_inv_pT_UseKernel_ScalFrac_0p8

- [marker 23]

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00192618	0.00195364	0.0020048	0.0020631	0.0021158	0.00217391	0.00224329	0.00229999	0.00235878
MSE(train)	0.00156352	0.00168727	0.00180032	0.00189026	0.00198354	0.00206121	0.00213061	0.00219875	0.00226366
k	55	60	65	70	75	80	85	90	95
MSE(test)	0.00241844	0.00246526	0.00251816	0.00257211	0.0026116	0.00265083	0.00269722	0.00274727	0.00279335
MSE(train)	0.00232797	0.00239232	0.00245496	0.00251	0.0025538	0.00259669	0.00264323	0.0026915	0.00273533

[4] dPhis_theta_CLCT1_dTh13_2_inv_pT_UseKernel_ScalFrac_0p8

- [marker 24]

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00215762	0.00225546	0.00235517	0.00242963	0.00248896	0.00257808	0.00266539	0.00273477	0.00279262
MSE(train)	0.00176634	0.00196423	0.00211612	0.00224619	0.00235997	0.00246641	0.00255935	0.00263202	0.00270892
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.00286146	0.00293098	0.00299153	0.00304792	0.00310173	0.00315473	0.00320266	0.0032476	0.00329641
MSE(train)	0.00277716	0.00283893	0.00291318	0.00296775	0.0030255	0.0030842	0.00313272	0.00317845	0.00322621

[5] dPhis_theta_CLCT1_dTh13_FR1_2_inv_pT_UseKernel_ScalFrac_0p8

- [marker 25]

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00240845	0.00250239	0.00262402	0.00271693	0.00282451	0.0029276	0.00300992	0.00308327	0.00315439
MSE(train)	0.00189398	0.00215861	0.00233636	0.00248231	0.00261598	0.00271606	0.00282472	0.00292125	0.00301088
k	55	60	65	70	75	80	85	90	95
MSE(test)	0.003222	0.00328881	0.00335467	0.00341367	0.00346904	0.00352862	0.00358405	0.00364054	0.00369461
MSE(train)	0.00309586	0.00316695	0.0032485	0.00331647	0.00337904	0.00344032	0.0034992	0.00355591	0.00361142

[6]dPhis_theta_FR1_2_inv_pT_UseKernel_ScalFrac_0p8

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00202694	0.00209336	0.00217506	0.00226277	0.00235517	0.00243539	0.00251424	0.00258543	0.00265756
MSE(train)	0.00155848	0.00175156	0.00190838	0.00203467	0.00216866	0.00229516	0.0024117	0.00250946	0.00258935
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.00273167	0.00279954	0.00286366	0.0029271	0.00298995	0.00304883	0.00310641	0.00315711	0.0032147
MSE(train)	0.00267291	0.00274749	0.00281933	0.00288944	0.00295091	0.00300845	0.00306747	0.00312258	0.00317649

[7]dPhis_theta_dTh13_inv_pT_UseKernel_ScalFrac_0p8

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00210994	0.00223347	0.00230528	0.00239404	0.00248048	0.00255741	0.00263471	0.00271158	0.00277691
MSE(train)	0.00172766	0.00194963	0.00209856	0.00222702	0.00234427	0.00243869	0.00253076	0.00261852	0.00270938
k	55	60	65	70	75	80	85	90	95
MSE(test)	0.00285207	0.0029194	0.00298436	0.00304194	0.00309561	0.00314478	0.00319992	0.00325235	0.00330314
MSE(train)	0.00277898	0.00284816	0.00291774	0.0029795	0.00303838	0.00309468	0.00315397	0.00320661	0.00325678

[8]dPhis_theta_FR1_CLCT1_2_inv_pT_UseKernel_ScalFrac_0p8

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00214129	0.00217582	0.0022562	0.00233732	0.00241946	0.00250759	0.00258447	0.00265593	0.00272222
MSE(train)	0.00161878	0.00182203	0.00196928	0.00210648	0.00222855	0.00234753	0.00244176	0.00251897	0.00260381
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.00279667	0.00286159	0.00292602	0.00298902	0.00304153	0.00308975	0.00313726	0.00319022	0.00324027
MSE(train)	0.00267993	0.00275365	0.00282001	0.0028889	0.00294967	0.00300643	0.00306054	0.00311075	0.00316041

[9]dPhis_theta_FR1_dTh13_2_inv_pT_UseKernel_ScalFrac_0p8

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00236445	0.00249065	0.00260032	0.00270773	0.00281433	0.00290992	0.00300767	0.00309268	0.00317089
MSE(train)	0.00187939	0.00212703	0.00231127	0.00246732	0.00261875	0.00273922	0.00284721	0.00294866	0.00303545
k	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**
MSE(test)	0.0032484	0.00331412	0.00338188	0.00345107	0.00351225	0.00357396	0.00363427	0.00369474	0.00375044
MSE(train)	0.00312394	0.0032137	0.00329128	0.00336388	0.0034386	0.00350354	0.00356629	0.00363012	0.00369025

BDT parameters(mode 15) from Andrew B.(check)

For mode 15

- 400 trees
- Depths: 5
- 1/pT weight(0-120 GeV), Log2(pT) is better target than 1/pT
- For very high pT, unweighted events better(>120 GeV)
- Input variables: FR bits bring significant improvement at low and high pT
- In addition to track theta, FR 1, and dPhi 1-2, 2-3, and 3-4(LUT v1), add combinations of dPhis, and ring number of station 1
- https://indico.cern.ch/event/608207/contributions/2451751/subcontributions/218758/attachments/1402616/2142649/2017_01_26_Mode_15_BDT.pdf

Modes

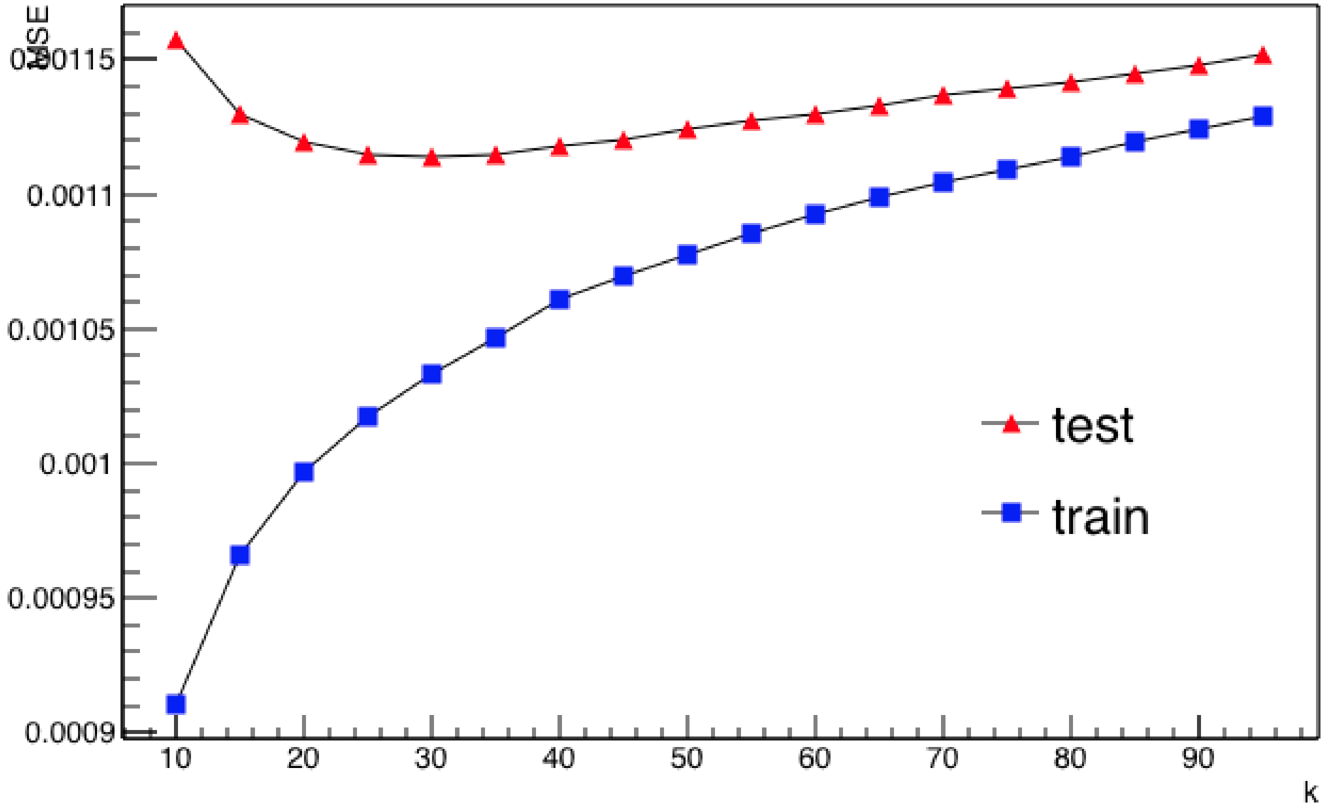
Mode #	Definiition in code	Stations
15	1+2+4+8	1,2,3,4
14	2+4+8	1,2,3
13	1+4+8	1,2,4
12	4+8	1,2
11	1+2+8	1,3,4
10	2+8	1,3
9	1+8	1,4
8	8	1
7	1+2+4	2,3,4
6	2+4	2,3
5	1+4	2,4
4	4	2
3	1+2	3,4
2	2	3
1	1	4

Version 0->1 changes

- Phi->phi_int
- Theta->theta_int
- Why this?
- Double_t dPhi12 = acos(cos((phi2 - phi1)*(PI/180.)));
- dPhi12 *= (sin((phi2 - phi1)*(PI/180.)) / max(BIT, abs(sin((phi2 - phi1)*(PI/180.)))));
- dPhi12 *= (180./PI);

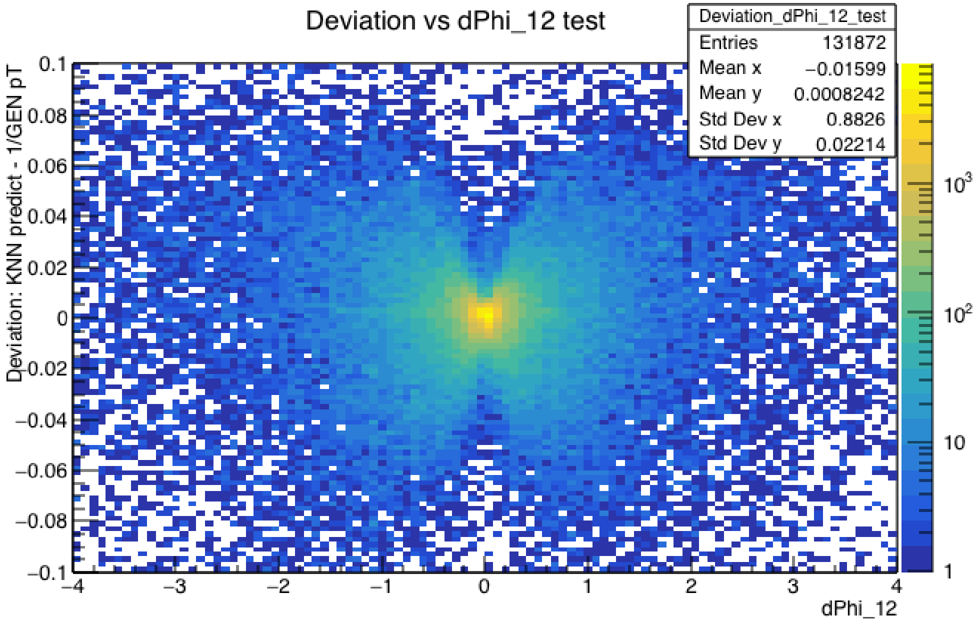
"nkNN=30:ScaleFrac=0.8:SigmaFact=1.0:Kernel=Gaus:UseKernel=F:UseWeight=T:!Trim"

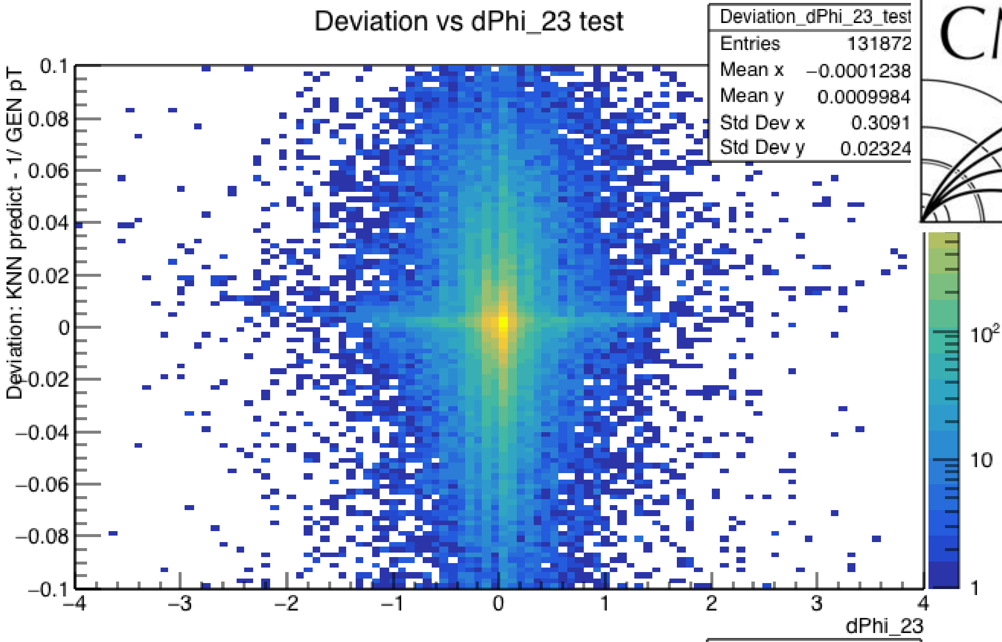
MSE = Deviation^2/N

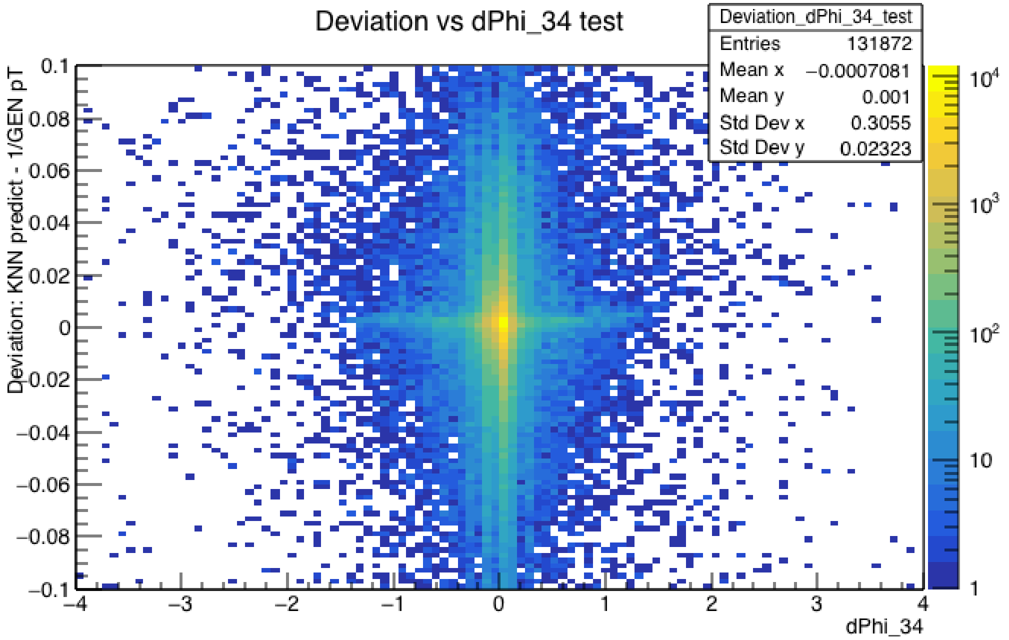


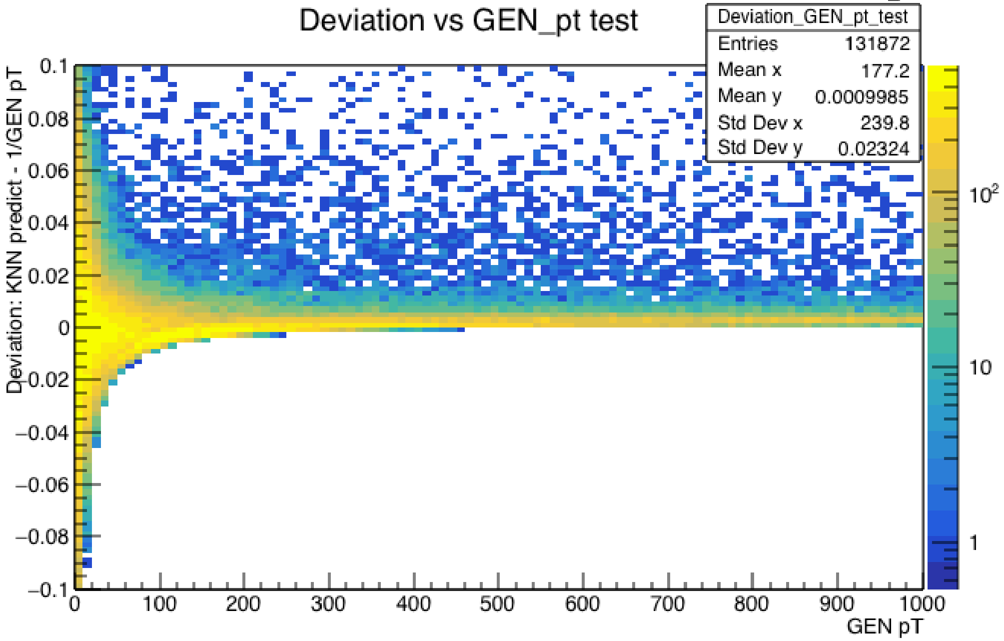
Mode: 15
Input variables: dphi12, dphi23, dphi34
Target variable: 1/Gen pT

/afs/cern.ch/work/a/abrinke1/public/EMTF/Analyzer/ntuples/EMTF_MC_NTuple_SingleMu_noRPC_300k.root



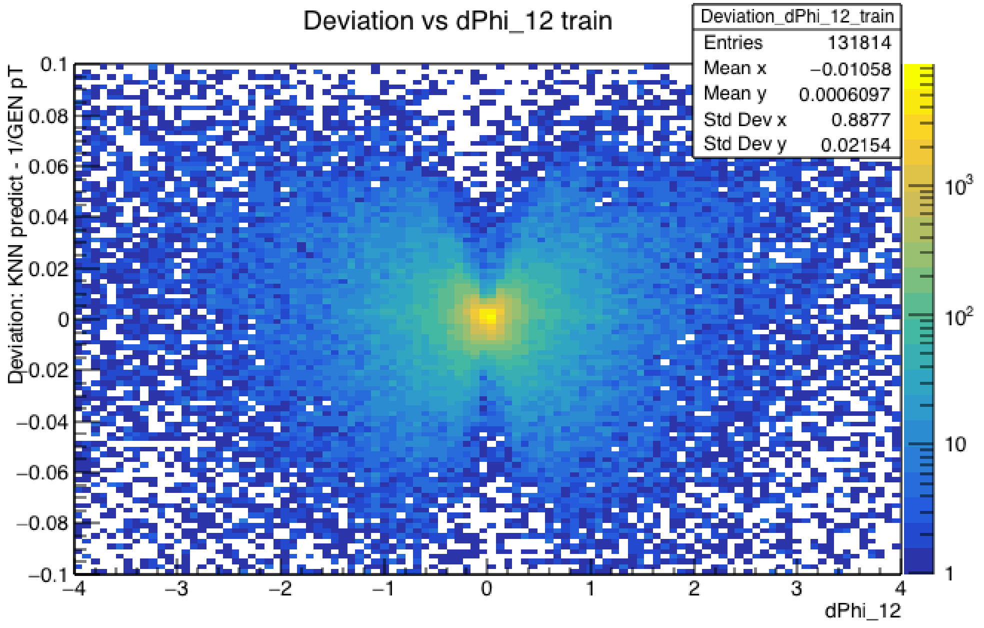


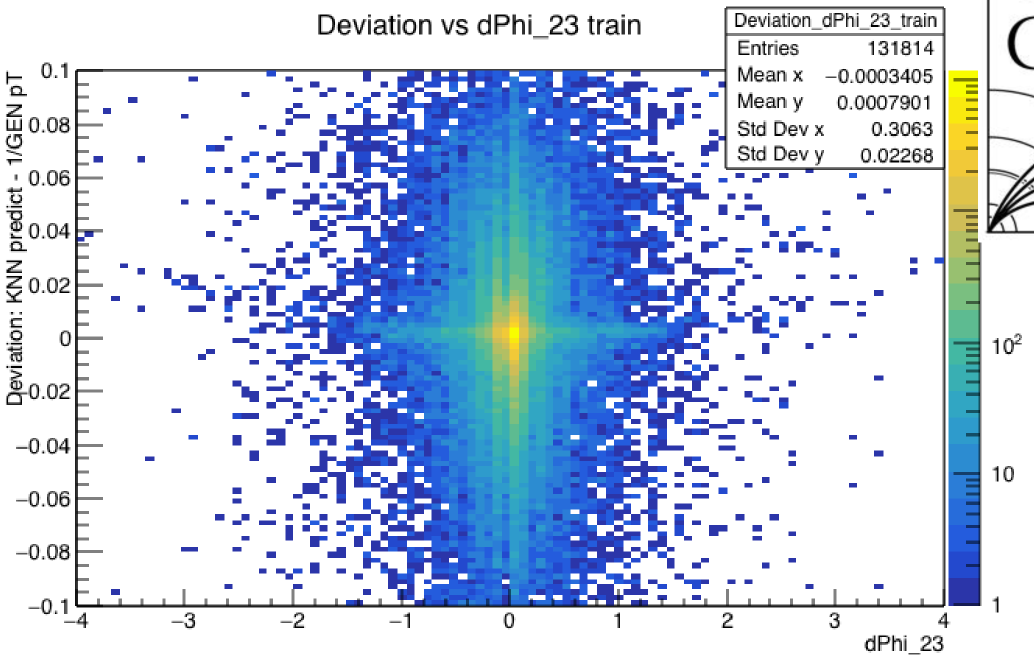


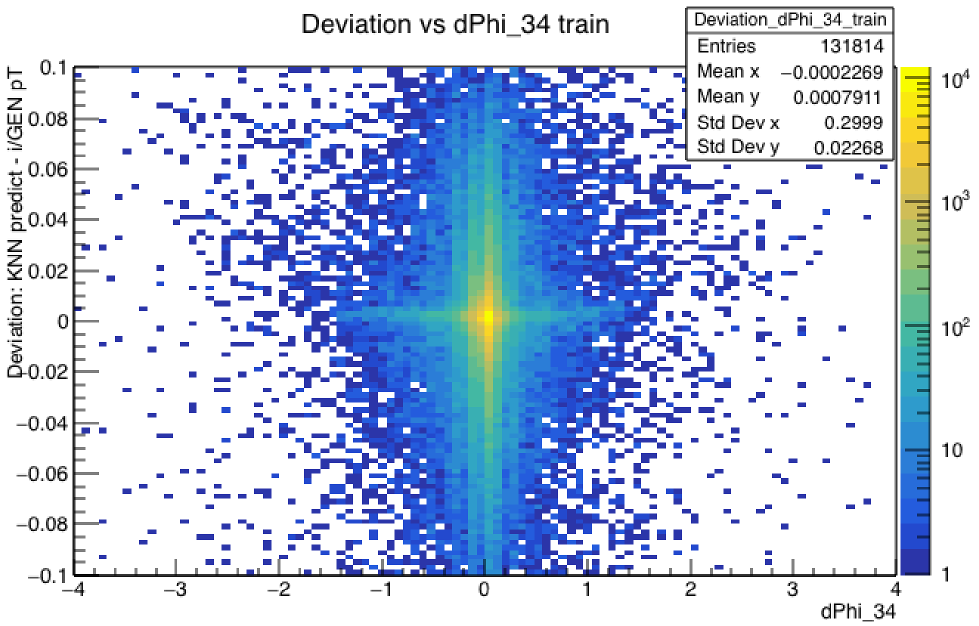


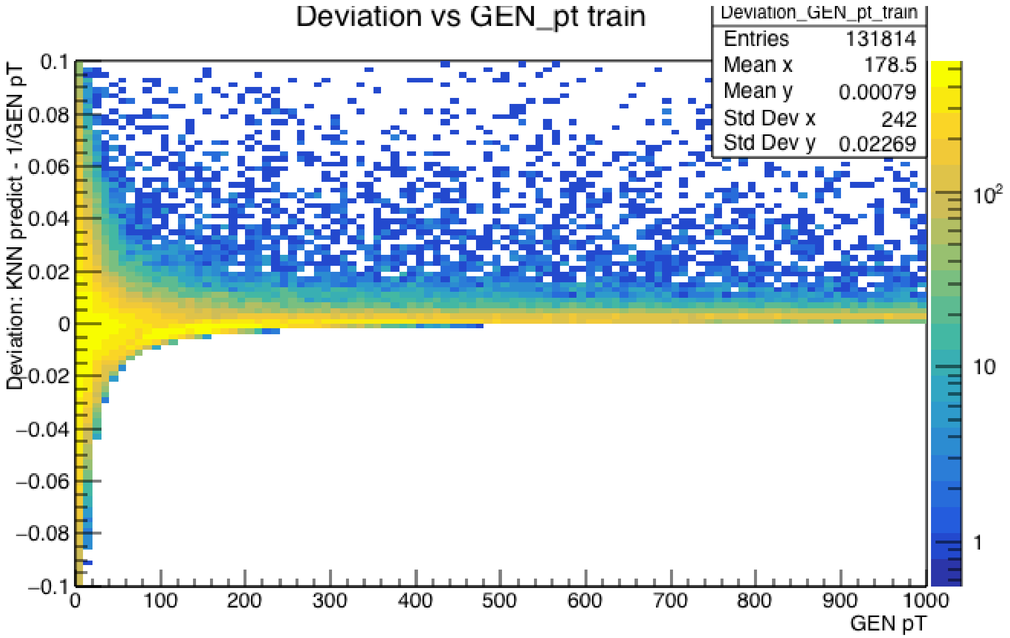
Mode: 15
Input variables: dphi12, dphi23, dphi34
Target variable: 1/Gen pT

Last plot shows predicted pT tends to be smaller than GEN pT, for large pT









Mode: 15
Input variables: dphi12, dphi23, dphi34
Target variable: 1/Gen pT

KNN Regression

- For a test event, the algorithm finds the k-nearest neighbours using the input variables, where each training event contains a regression value. The predicted regression value for the test event is the weighted average of the regression values of the k-nearest neighbours

"nkNN=95:ScaleFrac=0.8:SigmaFact=1.0:Kernel=Gaus:UseKernel=F:UseWeight=T:!Trim"

- Test with 131872 events, Train with 131814 events, mode 15, dphi12,23,34 as input, 1/Gen pT as target

k	10	15	20	25	30	35	40	45	50
MSE(test)	0.00115753	0.00112952	0.00111974	0.0011147	0.00111398	0.00111496	0.00111774	0.0011205	0.00112462
MSE(train)	0.000910486	0.00096604	0.000997173	0.00101723	0.00103356	0.00104684	0.00106053	0.00106954	0.00107752
k	55	60	65	70	75	80	85	90	95
MSE(test)	0.00112702	0.00112989	0.00113308	0.00113655	0.00113934	0.00114193	0.00114449	0.00114828	0.00115224
MSE(train)	0.00108521	0.00109236	0.00109916	0.00110448	0.00110943	0.00111416	0.0011192	0.00112404	0.00112887

Comments

- What pT range are we caring about? Depends on L1 endcap muon trigger threshold?
- Maybe we don't need to choose the best model, instead, we can use weighted models for final pT LUT
- Maybe we can use different model for different input variables, some model may perform better than others in certain parameter region(eta)

BDT good/bad

- Good: little tuning required(simple)

- Bad:

1. will ignore non-discriminating variables as for each node splitting only the best discriminating variable is used
2. theoretically best performance on a given problem is generally inferior to other techniques like neural networks.

- See TMVA tutorial