CDF/../../../8356
Version 1.0
June 30, 2006

Validation of bhel0i and bhmu0i Dataset in Lepton + Jets

Florencia Canelli,[1] Brian Mohr [2]

[1] *Fermi National Accelerator Laboratory*
[2] *University of California, Los Angeles*

Abstract

We study the 0i dataset taken from September 5, 2005 to February 2006 with total integrated luminosity 275 pb^{-1}. We report the results of event selection in the lepton + jets channel using methods consistent with gen 5 Monte Carlo. We compare kinematic distributions of the 0i dataset with the 0h dataset. Finally, we compare some kinematic distributions to signal and background Monte Carlo. We find the 0i data agrees reasonably well with the 0h data and with the predicted distributions from Monte Carlo.

1 Introduction

We use run-dependent Monte Carlo to calibrate and test our mass analyses. The gen 5 Monte Carlo generaly only contains runs in the 0d dataset (data taken from the beginning of Run II to August 2004, 318 pb^{-1}). Since then, there have been two additional datasets, 0h (from December 2004 to September 5, 2005, 362 pb^{-1}) and 0i (from September 5, 2005 to February 2006, 275 pb^{-1}). A few gen 5 Monte Carlo samples were generated with the 0d+0h dataset run range for comparison to results with the nominal gen 5 Monte Carlo with the 0d run range. There are no gen 5 samples with the 0i run range. Since the 0d dataset, improvements have been made to the detector as well as the accelerator. Most importantly, new regions of phase space have been added to the CMX muon detectors (the keystone and miniskirt regions) and the average instantaneous luminosity per run has increased. One of the important effects of higher average instantaneous luminosity is that the average number of multiple interactions in an event increases. The gen 5 Monte Carlo was generated with no extra interactions (i.e. only a single event vertex) and does not have a good simulation for the keystone and miniskirt regions of the CMX detector. There are gen 5 Monte Carlo samples with one and two extra event vertices for systematic studies.

We used the gen 5 Monte Carlo when measuring the combined 0d and 0h datasets [1]. In this case, we reject otherwise tight muons in the keystone and miniskirt regions of the CMX detectors so that the data more closely matches the Monte Carlo. We also compared kinematic distributions in the 0h and 0d datasets to the Monte Carlo.

With the inclusion of the 0i data, we again wish to use the gen 5 Monte Carlo and show it is a reasonable estimation of the data. As with the 0h data, we reject events with an otherwise tight CMX muon in the keystone and miniskirt regions. We present tables and run/event lists for the events passing selection. We compare kinematic distributions in the 0i data to the 0h data, which was compared to the 0d data. We also compare the full dataset to the Monte Carlo and show the Monte Carlo does a reasonable job of modeling the data. In all the general kinematic distributions, the gen 5 Monte Carlo well represents the combined 0d+0h+0i dataset. Systematics due to the improperly modeled multiple interactions are treated elsewhere.

2 Selection Tables and Run/Event List in 0i Dataset

We use the event selection described in [2], consistent with the selection for gen 5 Monte Carlo and the 0d dataset. We apply an additional veto of CMX muons in the keystone and miniskirt regions. We use good run list version 13 [3]. Since the topNtuples used to generate these tables (version 6.1.2) contain overlapping runs, any identical run/event pairs are removed. We use version 06 of the Jet Energy Corrections. Tables 1 to 3 present the number of events in the 0i samples after various stages of selection for the CEM electrons, CMUP muons and CMX muons, respectively. Tables 4 to 7 present the run/event numbers of the events passing pretag (up to and including silicon good run) and b-tag (up to and including at least one b-tag) selection with either exactly 4, or 5 or more, jets for the bhel0i and bhmu0i datasets.

Table 1: Event selection table for CEM electrons

cut	0 jet	1 jet	2 jet	3 jet	4 jet	≥5 jet	total
INITIAL	275425	226084	198069	26097	3745	621	730041
OBSV	275425	226084	198069	26097	3745	621	730041
GOODRUN	249856	205266	179635	23683	3404	549	662393
TRIGGER	235752	184713	163566	20474	2839	446	607790
≥1 TLEP	235751	81300	8734	1237	195	41	327258
MET	150028	16197	2689	500	111	27	169552
ISO	150028	16197	2689	500	111	27	169552
!DILEP	149919	16082	2649	487	108	25	169270
!Z	148246	15312	2418	420	96	24	166516
!COSMIC	148246	15312	2418	420	96	24	166516
!DIFFZ	148153	15303	2415	419	96	24	166410
SI GOOD	142252	14701	2331	406	94	23	159807
TAGGABLE	0	9882	2016	380	89	23	12390
≥1 BTAG	0	219	119	39	32	10	419
≥2 BTAG	0	0	8	6	10	3	27
≥1 NTAG	0	56	23	8	2	2	91

Table 2: Event selection table for CMUP muons

cut	0 jet	1 jet	2 jet	3 jet	4 jet	≥5 jet	total
INITIAL	246632	78007	27825	5254	777	144	358639
OBSV	246632	78007	27825	5254	777	144	358639
GOODRUN	223917	70870	25278	4769	687	134	325655
TRIGGER	180807	25997	3166	446	81	15	210512
≥1 TLEP	104316	15194	1899	263	46	9	121727
MET	87491	8135	1209	177	34	5	97051
ISO	87491	8135	1209	177	34	5	97051
!DILEP	85239	7815	1160	165	34	5	94418
!Z	81414	7385	1101	158	33	4	90095
!COSMIC	78727	7370	1099	157	33	4	87390
!DIFFZ	76757	7289	1079	157	33	4	85319
SI GOOD	73510	7014	1029	148	31	4	81736
TAGGABLE	0	4688	900	139	29	4	5760
≥1 BTAG	0	114	51	18	13	2	198
≥2 BTAG	0	0	1	1	3	0	5
≥1 NTAG	0	23	11	3	2	1	40

Table 3: Event selection table for CMX muons

cut	0 jet	1 jet	2 jet	3 jet	4 jet	≥5 jet	total
INITIAL	246632	78007	27825	5254	777	144	358639
OBSV	246632	78007	27825	5254	777	144	358639
GOODRUN	223917	70870	25278	4769	687	134	325655
TRIGGER	180807	25997	3166	446	81	15	210512
≥1 TLEP	63702	8965	1041	156	30	3	73897
MET	53295	4643	689	116	24	3	58770
ISO	53295	4643	689	116	24	3	58770
!DILEP	52175	4472	666	113	23	3	57452
!Z	50228	4255	623	104	23	3	55236
!COSMIC	49912	4248	623	103	22	3	54911
!DIFFZ	44987	4097	588	94	22	3	49791
SI GOOD	43230	3914	559	93	20	3	47819
TAGGABLE	0	2596	476	89	20	3	3184
≥1 BTAG	0	40	18	9	6	0	73
≥2 BTAG	0	0	1	3	2	0	6
≥1 NTAG	0	12	3	1	2	0	18

Table 4: Run/event list for bhel0i 4 jet pretag events

203819	2576851	205964	2067507	209402	611693
203819	3894430	205995	3435934	209437	4347132
203819	7609959	206033	7372720	209437	5160294
203874	837983	206186	8131916	209439	3255646
204077	613751	206186	10746480	209465	2369336
204204	1864246	206282	92075	209486	3028699
204257	11760484	206326	6735620	209489	210241
204264	4038	206326	12347673	209489	3063417
204378	624826	206695	4955884	209761	500847
204402	411283	206951	3695717	209761	606104
204402	3924353	206982	13255	209762	1086037
204447	2908437	207078	574202	209762	2492314
204549	1102743	207079	955759	209773	5803429
204550	7646460	207079	2229794	209789	58144
204593	87032	207079	18867699	209819	10301925
204640	291513	207080	688663	209849	1183422
204640	2237983	207115	26425	209982	49226
204643	57292	207151	473247	209986	5290222
204677	1222941	207214	6058942	210008	3097811
204721	921915	207421	5194571	210008	8610161
204746	255259	207488	2477561	210914	18377850
204746	748687	207488	3747131	211135	899577
204748	9152703	209085	3697233	211316	1736264
204750	3902489	209123	3454791	211337	1995298
205009	6313728	209263	245370	211337	12798551
205310	58338	209263	2427716	211396	6588704
205314	1059962	209316	7020668	211397	1059823
205322	1355625	209341	2335530	211441	199145
205627	3601712	209373	13924681	211449	2826198
205640	7076930	209373	17855621	211882	1660038
205719	235097	209399	1972113	211883	2030757
				211903	23828610

Table 5: Run/event list for bhel0i pretag ≥ 5 jet events

204257	897055	205074	5154575	207114	1654886
204257	4912478	205991	1000375	209265	6003601
204266	49918	206285	14542849	209402	8508323
204447	3183448	206326	6536298	209402	15204780
204679	261144	206436	55019	209507	3054439
204714	930162	206537	500843	209907	2580088
204827	660355	206669	1804953	212133	1174035
204827	5634289	206697	1283235		

Table 6: Run/event list for bhmu0i pretag 4 jet events

203826	2201856	205966	5016023	207115	3014994
204204	101652	205966	20056884	207115	12009584
204339	2453479	205991	4192529	207151	479744
204339	2816209	205991	6324276	207153	5682815
204378	783371	205994	18703	207512	342227
204402	2939085	206033	1568581	209021	797869
204469	621060	206285	14834398	209053	1834688
204589	8218	206326	4815862	209191	6933350
204650	2484068	206326	13525374	209318	2796669
204748	550951	206830	2739255	209341	12604512
204750	4333547	206951	1711212	209537	8035665
205018	3180321	206989	5933858	209761	416428
205018	3690938	207079	1847362	209770	847655
205548	94885	207079	2920112	209819	2062462
205558	988503	207079	5293129	209986	10591065
205627	1724773	207079	5948787	211312	1977991
205891	340346	207114	1570102	211452	213251

Table 7: Run/event list for bhmu0i pretag ≥ 5 jet events

204111	52744	209507	5688680	211265	6950139
205151	4020488	209862	1339964	211292	1019493
209489	1389909				

Table 8: Run/event list for bhel0i b-tag 4 jet events

203819	2576851	206186	8131916	209437	5160294
204077	613751	206326	6735620	209465	2369336
204402	411283	207079	2229794	209789	58144
204549	1102743	207079	18867699	209982	49226
204593	87032	207080	688663	210008	3097811
204640	291513	207488	2477561	211316	1736264
205009	6313728	207488	3747131	211337	1995298
205314	1059962	209263	2427716	211337	12798551
205719	235097	209316	7020668	211449	2826198
205964	2067507	209373	13924681	211882	1660038
206033	7372720	209399	1972113		

Table 9: Run/event list for bhel0i b-tag ≥ 5 jet events

204447	3183448	206669	1804953	209402	8508323
204679	261144	206697	1283235	209507	3054439
204714	930162	209265	6003601	212133	1174035
205991	1000375				

Table 10: Run/event list for bhmu0i b-tag 4 jet events

203826	2201856	205994	18703	207153	5682815
204204	101652	206326	4815862	209053	1834688
204339	2453479	206989	5933858	209318	2796669
204650	2484068	207079	1847362	209819	2062462
204750	4333547	207114	1570102	211312	1977991
205018	3690938	207115	12009584	211452	213251
205966	5016023				

Table 11: Run/event list for bhmu0i b-tag ≥ 5 jet events

205151	4020488
209507	5688680

3 Kinematic Comparison of 0i Dataset

We first compare distributions of some kinematic variables between the 0h and 0i dataset. The variables are jet multiplicity, lepton p_T, missing E_T, jet p_T and b-tagged jet p_T. Selection is made up to, but not including, the good run silicon requirement (except for the b-tagged jet p_T). In the jet mulitiplicity plot, no additional selection is made. In the lepton p_T, missing E_T and jet p_T plots, we require jet multiplicity of ≥ 1 jets. In the jet p_T plot, we include all jets in the event. In the b-tagged jet p_T plot, we further require the event to pass the silicon good run list, to have a taggable jet and to have a b-tagged jet. We include all b-tagged jets in the plot. These plots were made using version *1* invfb (gen 6) of topNtuple and version 05b of the Jet Energy Corrections for both the 0h and 0i datasets.

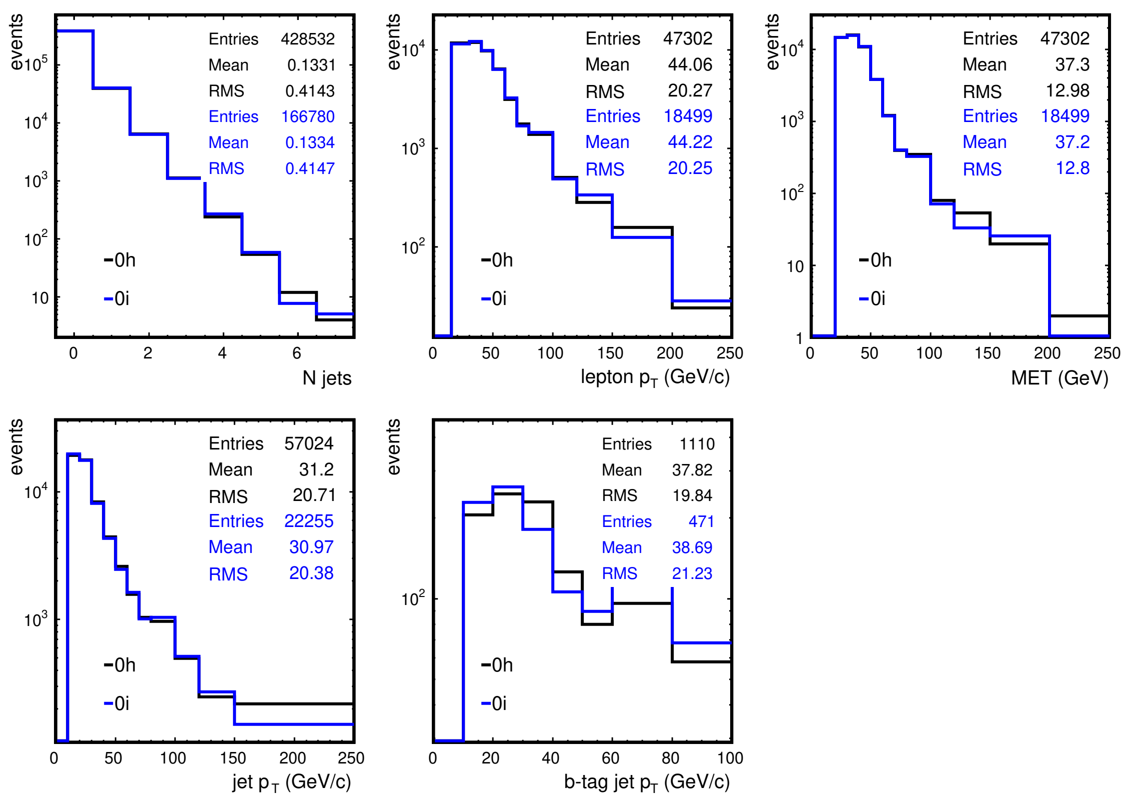

Figure 1: Comparison of distributions of select kinematic variables between the bhel0h and bhel0i datasets (electrons). Clockwise from top left: jet multiplicity, lepton p_T, missing E_T, b-tagged jet p_T, and jet p_T.

Figure 1 shows the shape comparison for electrons. The distributions are scaled to have the same number of entries. We see good agreement in these variables throughout the distributions. Figure 2 shows the ratio of the bin content for the two datasets. This ratio is calculated after the bins have been scaled, so we expect the ratio to be unity for each bin if there is good agreement. The bins with high statistics lie very close to unity, while the lower statistics bins fluctuate about unity, so we have good agreement in these variables between the two datasets.

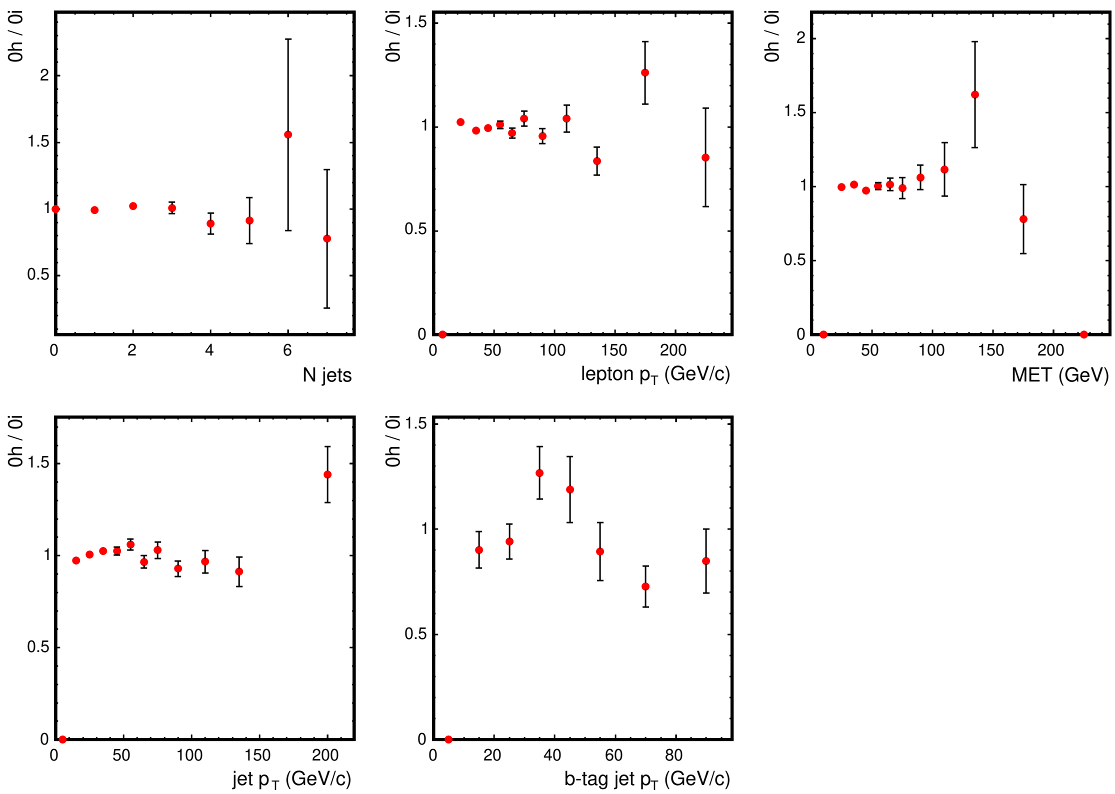

Figure 2: Ratio of distributions of select kinematic variables between the bhel0h and bhel0i datasets (electrons). Clockwise from top left: jet multiplicity, lepton p_T, missing E_T, b-tagged jet p_T, and jet p_T.

Figure 3 shows the shape comparison for muons. The distributions are scaled to have the same number of entries. We see good agreement in these variables throughout the distributions. Figure 4 shows the ratio of the bin content for the two datasets. As with the electrons, this ratio is calculated after the bins have been scaled. Again, we see good agreement between the two datasets.

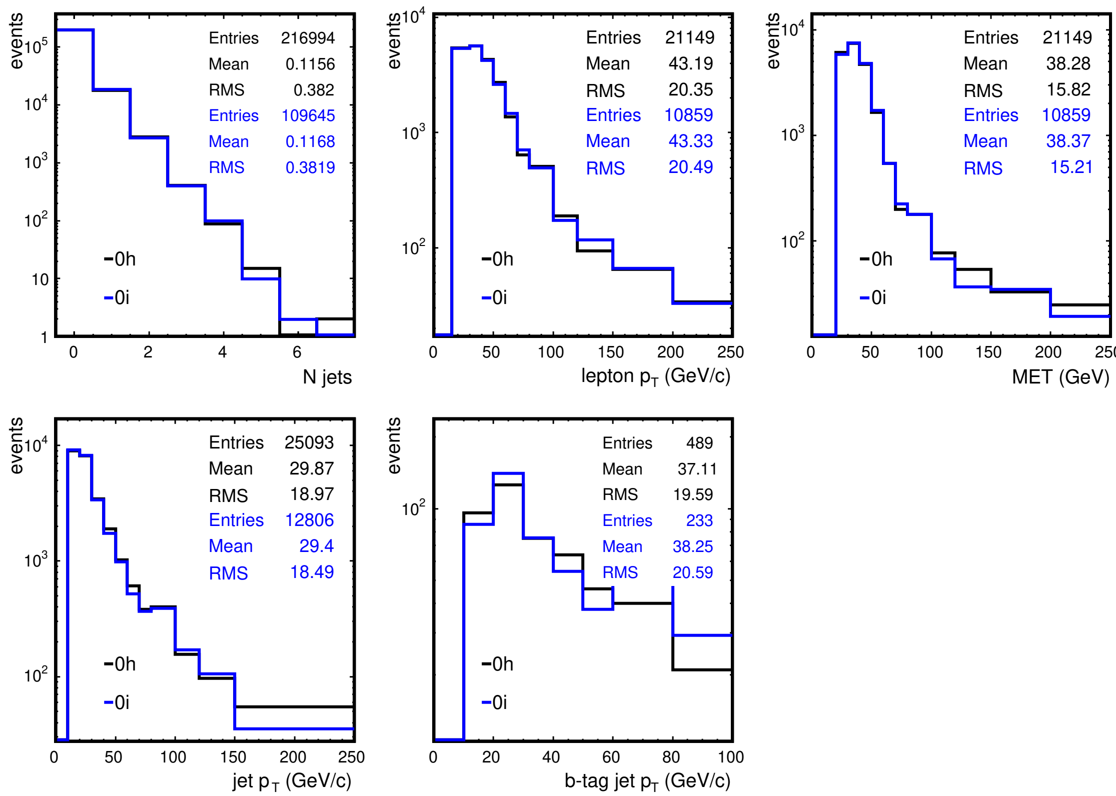

Figure 3: Comparison of distributions of select kinematic variables between the bhmu0h and bhmu0i datasets (muons). Clockwise from top left: jet multiplicity, lepton p_T, missing E_T, b-tagged jet p_T, and jet p_T.

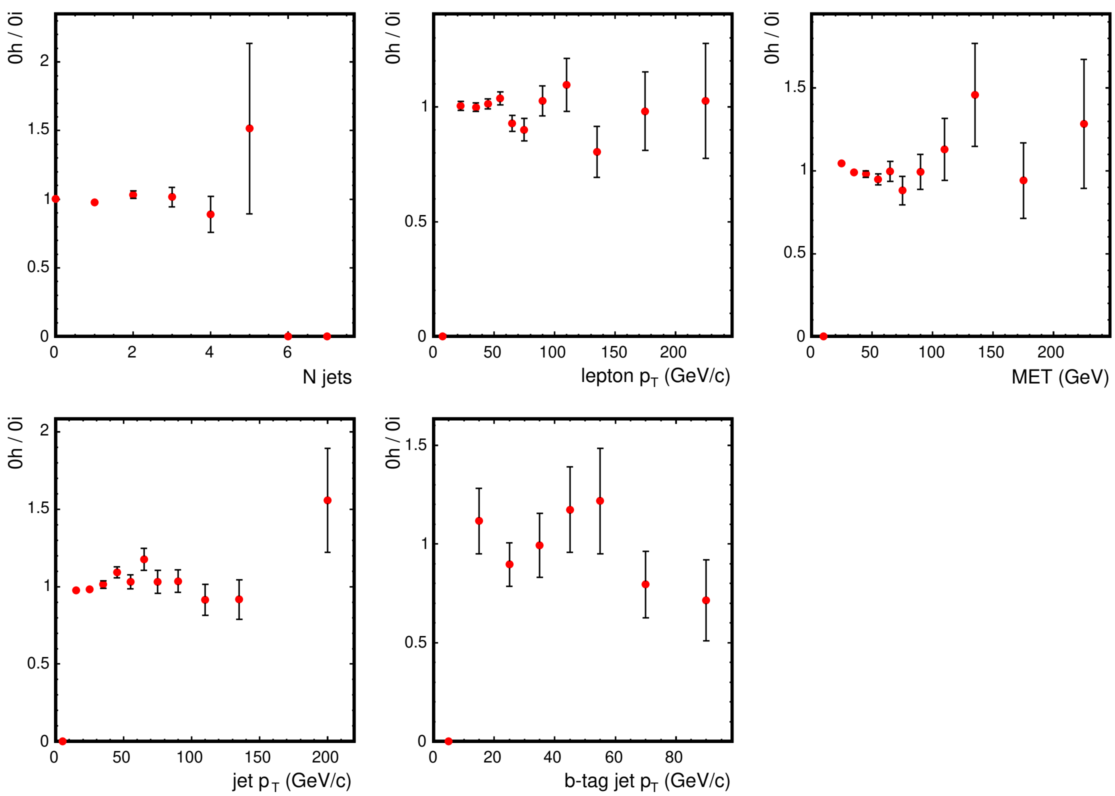

Figure 4: Ratio of distributions of select kinematic variables between the bhmu0h and bhmu0i datasets (muons). Clockwise from top left: jet multiplicity, lepton p_T, missing E_T, b-tagged jet p_T, and jet p_T.

4 Comparison to Monte Carlo

We now compare data to Monte Carlo in the pretag and b-tag selection. In each case, we require exactly four tight jets in the events and the events to pass the QCD veto. Pretag selection includes selection up to, and including, the silicon good run list. B-tag selection requires at least one tagged jet. We use version 6.1.2 of the top ntuple for data, version 5.3.3_nt for the Monte Carlo and gen 5 Monte Carlo. We use Jet Energy Corrections version 06 for the 0i data, 05b for the 0h+0i data and 05 for the Monte Carlo. The signal Monte Carlo is $t\bar{t}$ at M_{top} = 172.5 GeV/c^2. The non-W (QCD) sample is from the non-isolated lepton sideband of the data. In the pretag plots, we consider non-W (QCD) and W+jets backgrounds only. We use [4] to determine the sample composition: non-W is 4%, W+jets is 60% and $t\bar{t}$ is 36%. This composition is determined in the cross-section measurement with the combined 0d+0h datasets. We assume the relative contribution is the same for the 0d+0h+0i dataset. In the b-tag plots, we use Table 12 to determine the event composition. This table was originially derived from the cross-section prediction for the 0d dataset [5]. It was scaled by luminosity and used for the combined 0d+0h datasets, and we now scale by luminosity for the combined 0d+0h+0i datasets. This b-tag composition estimate did not include the QCD veto. We estimated the effects of the QCD veto by applying it to the Monte Carlo for everything but the non-W (QCD) contribution. The effect of the QCD veto on the non-W contribution was estimated by assuming the cut affects the pretag ≥ 4 jet selection [6] in the same way as the b-tag exactly 4 jet selection. In both pretag and b-tag, we scale the Monte Carlo to the data by the number of entries.

Figure 5 shows the individual jet p_T of the four jets in the pretag selected events, where the jets are sorted by decreasing p_T. We see good agreement between data and Monte Carlo. Figure 6 shows the missing E_T, lepton p_T, b-tagged jet p_T and jet multiplicity for the pretag selected events. We see good agreement in missing E_T and jet multiplicity and relatively good agreement in b-tagged jet p_T. The lepton p_T distribution is acceptable, but we see a bit of a mismatch in the low p_T bin. This may be due to improper modeling of the non-W (QCD) background, since this is the most significant bin in which this background contributes. Figure 7 shows the difference in the angle between the momentum vectors of the leading and second jet, the leading jet and lepton, and the difference in the angle between the transverse vectors of the leading jet and missing E_T. We see good agreement between data and Monte Carlo in each case.

Table 12: Expected event contribution for b-tag sample with 955 pb^{-1} total integrated luminosity. Assumes $t\bar{t}$ cross-section of 6.1 pb.

Sample	Expected Events
Background	
$W + 4p$ (Mistags)	6.16 ± 1.28
Non-W (QCD)	5.26 ± 2.61
Wbb	4.70 ± 2.16
$Wc\bar{c}$	2.24 ± 1.07
Wc	1.40 ± 0.63
single top	1.14 ± 0.27
EW (WW,WZ,ZZ)	1.08 ± 0.24
Total Background	21.97±3.84
Signal	
$t\bar{t}$ (6.1 pb)	110.63 ± 12.58
Observed	
Data	166

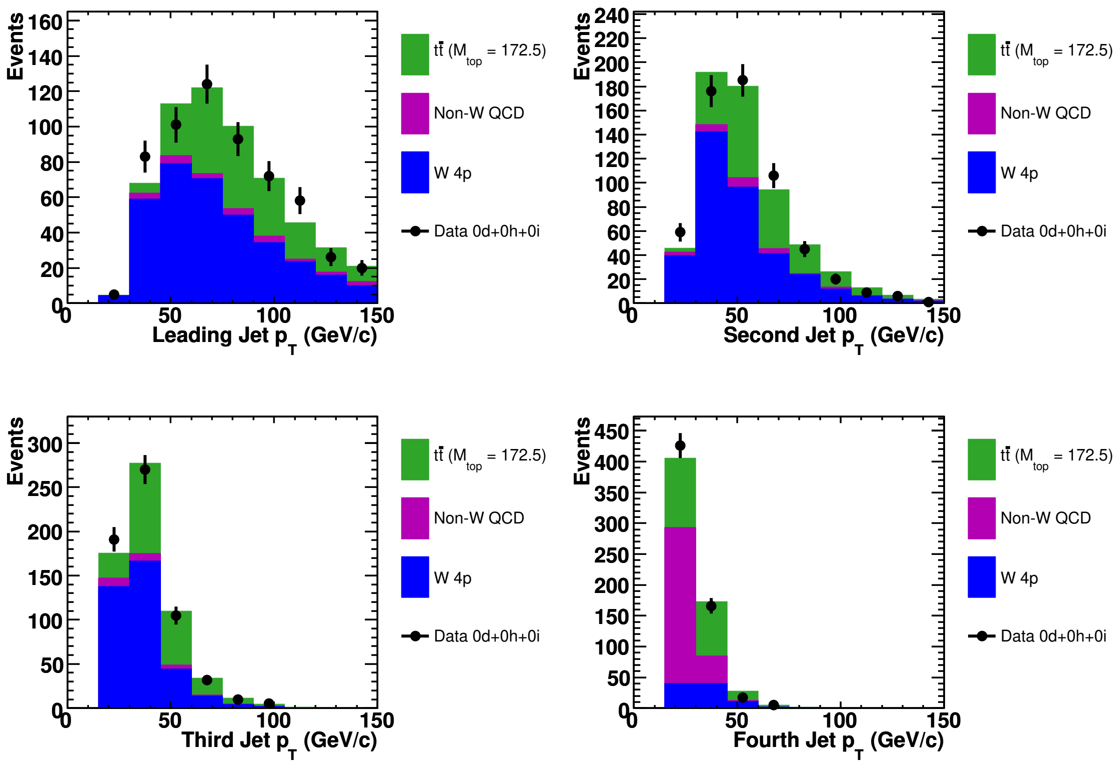

Figure 5: Jet p_T of four jets in pretag sample, sorted by decreasing p_T.

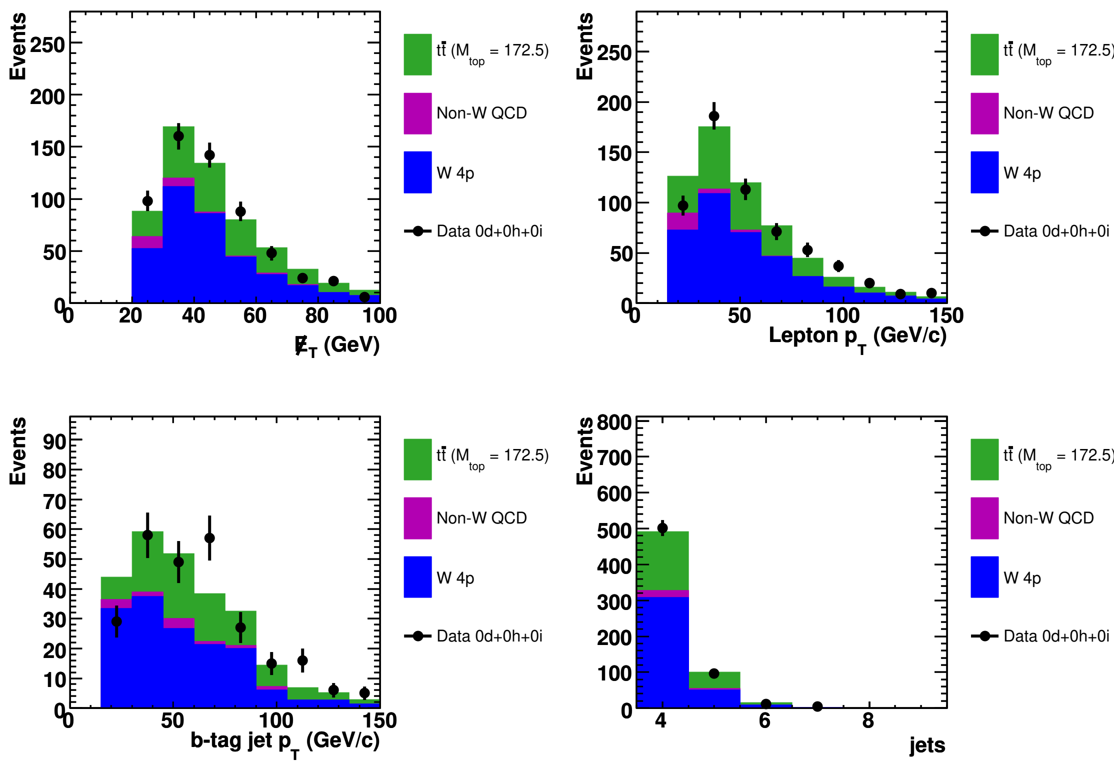

Figure 6: Clockwise from top-left: Missing E_T, lepton p_T, b-tagged jet p_T and jet multiplicity for pretag selection.

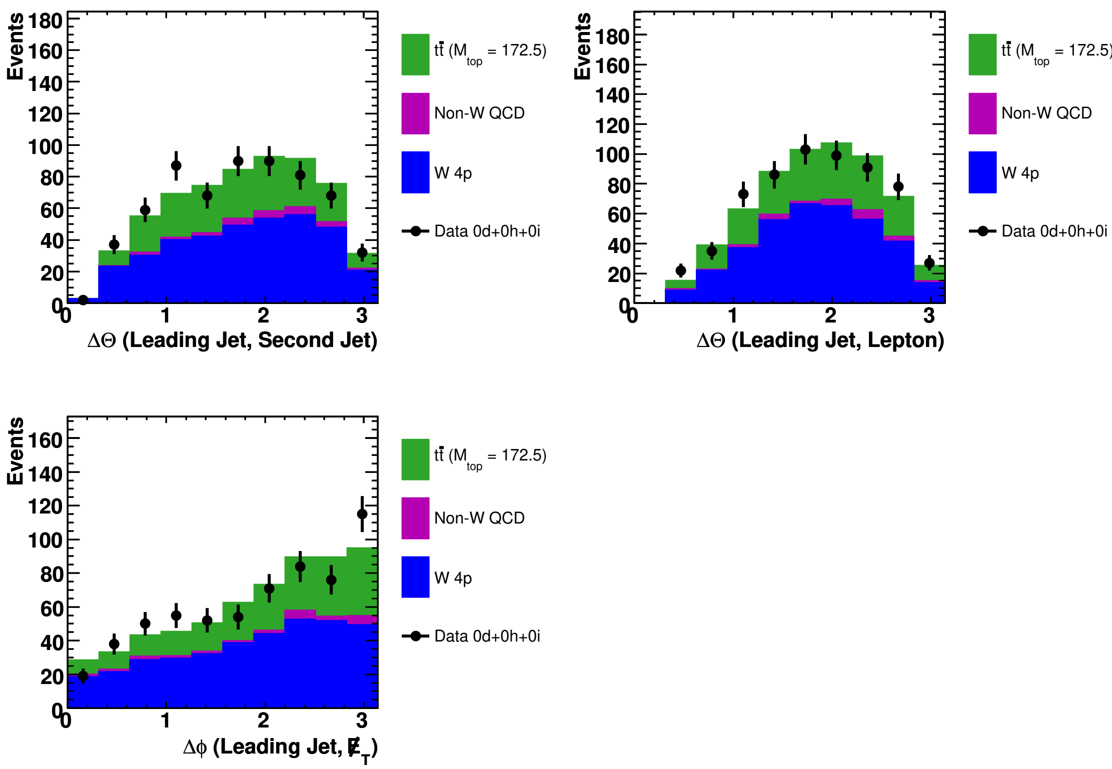

Figure 7: Angle between leading jet and: second jet, lepton, and missing E_T, in the b-tag sample.

Figures 8 to 10 show similar distributions with b-tag selection. Again, we see good agreement for the jet p_T distributions and angular distributions. The lepton p_T distribution looks better than the pretag case, but we still have a bit of an deficit in the first bin. Also, the missing E_T distribution appears slightly shifted to higer values in the data.

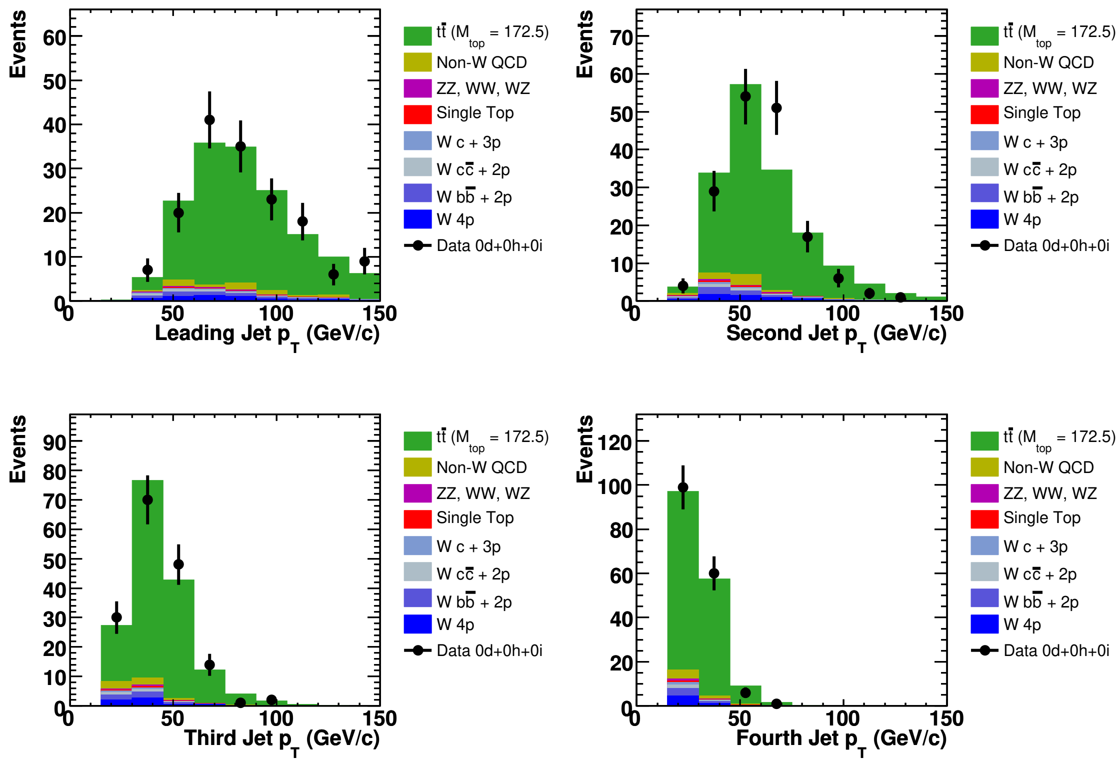

Figure 8: Jet p_T of four jets in b-tag sample, sorted by decreasing p_T.

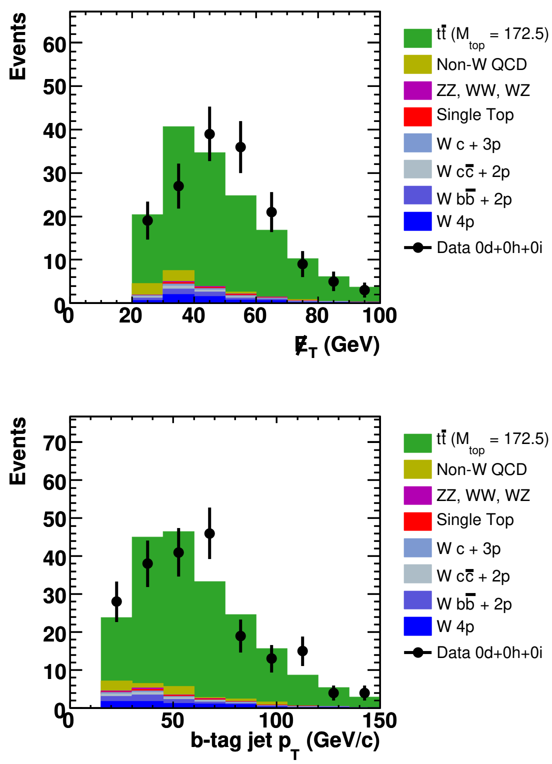

Figure 9: Clockwise from top-left: Missing E_T, lepton p_T, b-tagged jet p_T and jet multiplicity for b-tag selection.

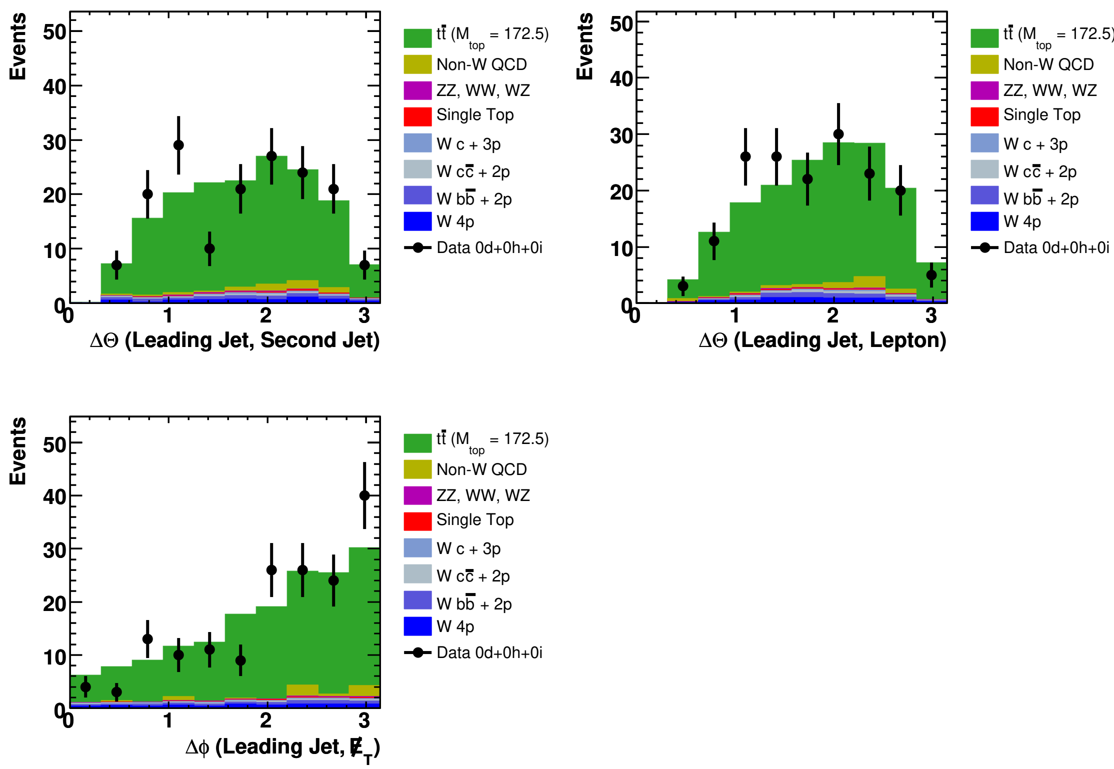

Figure 10: Angle between leading jet and: second jet, lepton, and missing E_T, in the b-tag sample.

5 Conclusion

We present the selection tables and run/event lists for the 0i dataset. Comparisons between the 0i and 0h dataset for some important kinematic variables show no deviations in the 0i dataset. The 0i dataset closely matches the 0h dataset and thus also matches the 0d dataset. Comparisons of the combined 0d+0h+0i dataset with the gen 5 Monte Carlo find no significant deviations in the data from Monte Carlo. We thus find the gen 5 Monte Carlo to be a reasonable simulation to use for preliminary mass measurements with the combined 0d+0h+0i datasets. Extra systematics for the unmodeled multiple interactions are treated elsewhere.

References

[1] CDF Note 8043, *Validation of Gen6 Data for MEAT and FlaME Analyses*, F. Canelli, B. Mohr, *et al.* 1

[2] CDF Note 7372, *Preliminary Event Selection and $t\bar{t}$ Signal Acceptance for the Winter 2005 Top Lepton + Jets Samples*, R. Erbacher, J. Nielsen, *et al.* 2

[3] CDF Note 5613, *The Good Run List*, H. Frisch, R. St. Denis, T. Vaiciulis, W. Badgett. 2

[4] CDF Note 8092, *Measurement of the Cross-Section for top pair production using event kinematics*, K. Lannon, R. Hughes, *et al.* 12

[5] CDF Note 7562, *Preliminary Background Estimate of the SECVTX-tagged Lepton + Jets Sample*, H. Bachacou, J. Guimaraes da Costa, *et al.* 12

[6] CDF Note 7563, *Measurement of top pair production cross section in Lepton Plus Jets Channel using neural networks*, K. Lannon, R. Hughes, *et al.* 12