Data Driven
Background Determination

Isabell-A. Melzer-Pellmann

Dominant backgrounds

- ttbar (mainly for high N_{jet})
- W+Jets (mainly for low N_{jet})
- QCD multijet production

Method overview

- HCAL Noise suppression
- SM MET subtraction
- α_T jet balancing method
- χ^2(ttbar)
- Top Box Method
- Combinatorial reconstruction of ttbar / Topbox method
- Dilepton Background (in single lepton events)
- ttbar background estimation using b-tagging
- W polarisation method
- W background determination using Z control sample
- Suppression of signal contamination in background control regions
- Determination of correlation in ABCD method
- Fake electrons
- QCD background

HCAL Noise Suppression

Noise suppression based on precise hit timing

- Time resolution O(1ns) for high energies, worse for lower energies
- reject events which are out of time window around event time
- window needs to have broader shape for lower energies
- MET reduction by factor 5-10

SM MET subtraction

Largest backgrounds for lepton+jets+MET:

- ttbar+jets with t→bW(lν)
- W+jets with W→lν

Two components to MET: artificial and ν-induced

- Artificial MET:
  - jets, detector, beam related backgrounds,non-collision effects
  - ☛ Model MET with a pool of multi-jet QCD events based on:
    - N_{jet} above high pT (e.g. 50 GeV or higher)
    - J_T = SUM $|p_T Jet|$ with $p_T(jet)$>20 GeV
- ν-induced MET:
  - ☛ Model using charged l p_T spectra (l and ν spectra should be the same if W is not polarised in transverse plane, need small corrections as W is polarised in ttbar decays)

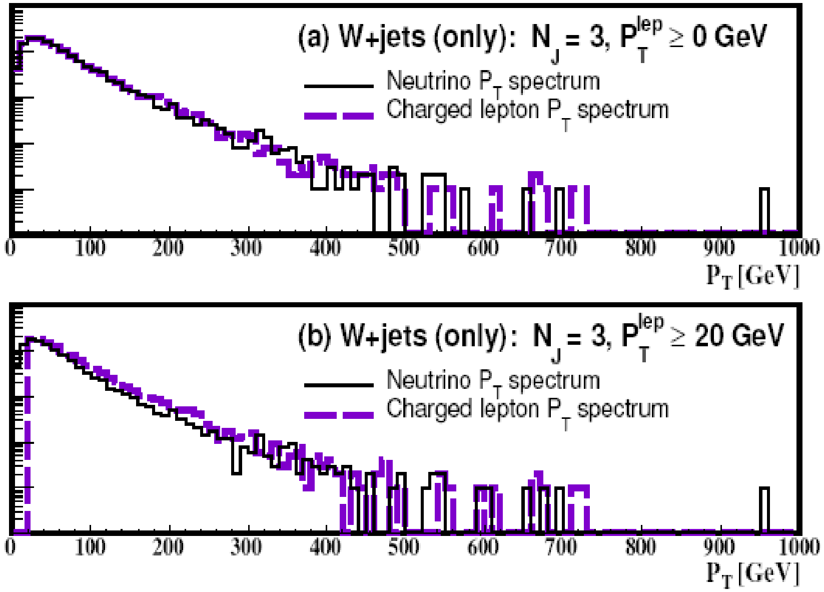
Assume that the two components interfere at random angle ϕ (0-π)

☛ Smear charged l pT with artificial MET prediction at angle ϕ

☛ Works better for events with large N_{jet}

SM MET subtraction (2)

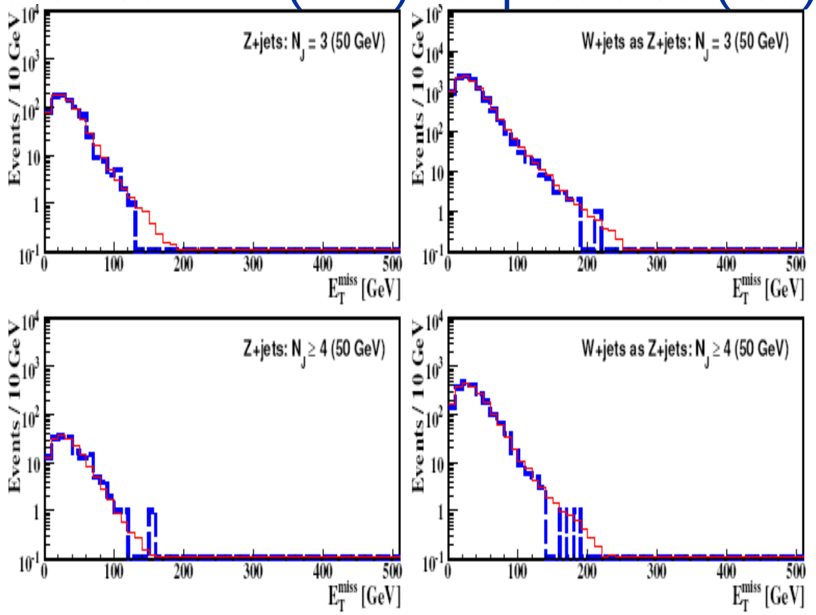
l – v comparison for W



Harder cut on charged lepton p_T makes v spectrum softer

☛ needs to be corrected

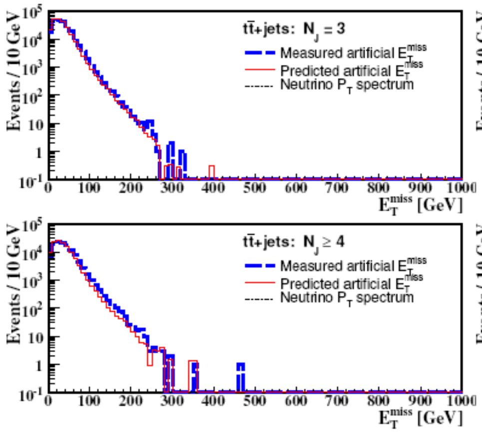
observed MET (blue) vs. predicted (red)



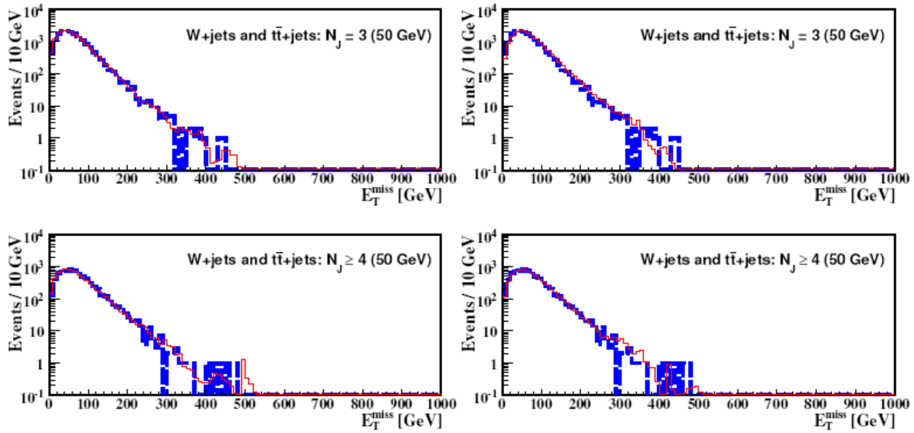
W+jets as Z+jets with v treated as 2nd mu using generator p_T

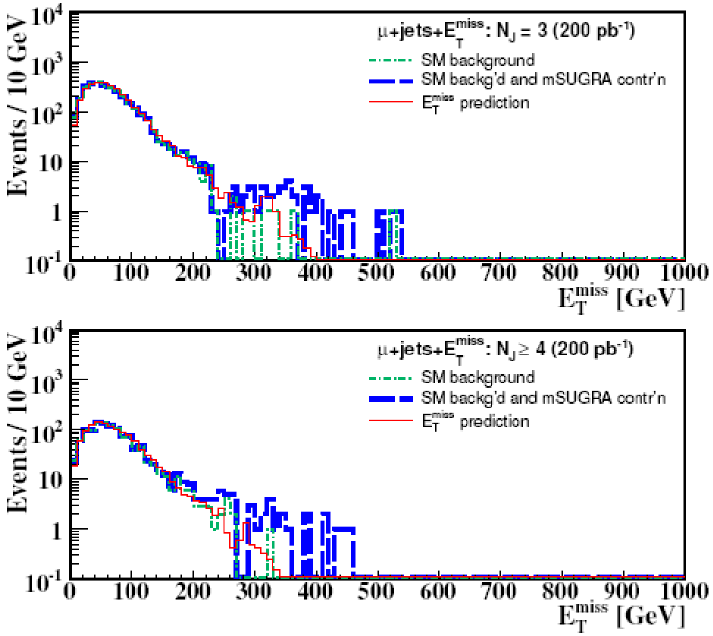
SM MET subtraction (3)

artificial MET



artificial + genuine MET





Prediction works best for models with soft lepton spectra

α_T jet-balancing method

Lepton isolation: key tool to reduce background from fake or heavy-flavor electrons and muons;

For pT>30 GeV: standard recommendations from V+jets Cross PAG

For pT<30 GeV: $\mathrm{TrkIso}_{abs} = \sum_{\Delta R<0.3} p_T^{track}$ < 3GeV (e) or < 5 GeV (mu)

α_T: from all-hadronic analysis to reproduce kinematics in di-jet event

Idea: construct two pseudo-jets, which balance each other in H_T, where pseudo-jet H_T (scalar sum of p_T of all jets in pseudo-jet)

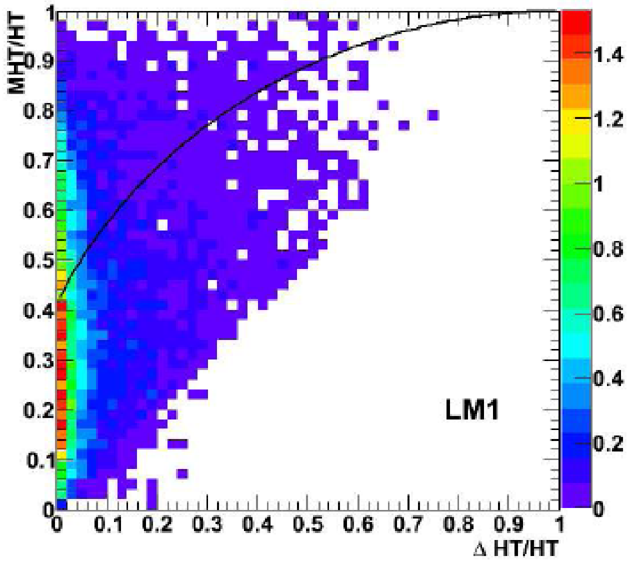
Jets are combined in pseudo-jets by minimizing $\Delta H_T = |H_T1 - H_T2|$

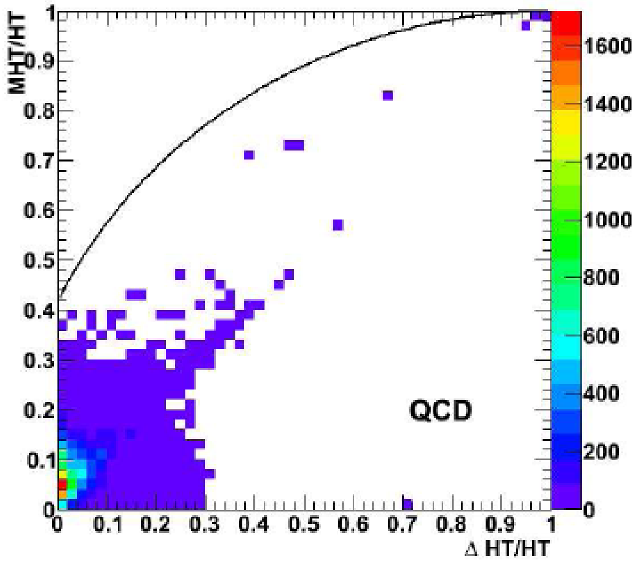
$$\alpha_T = \frac{1}{2}\frac{H_T - \Delta H_T}{M_T} = \frac{1}{2}\frac{H_T - \Delta H_T}{\sqrt{H_T^2 - MH_T^2}}$$

Completely balanced system: $\Delta H_T = 0$

Leptonic α_T: similar to all-hadronic, just including lepton in variables

α_T jet-balancing method (2)

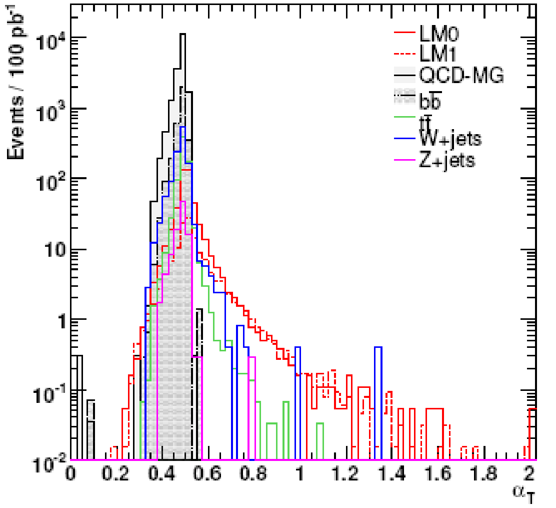


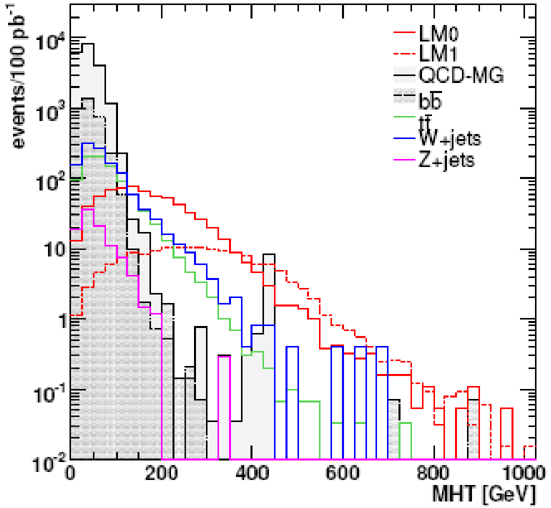


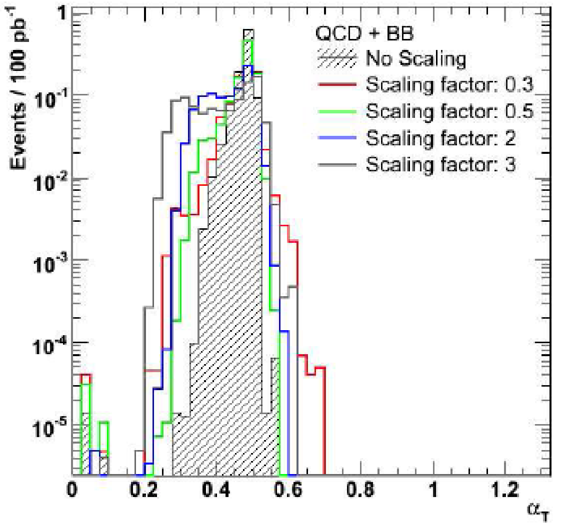
H_T>350 GeV
curve for
α_T=0.55

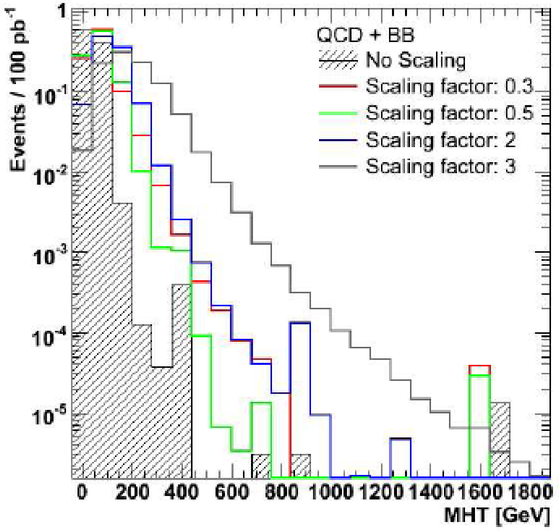
Good separation power

α_T jet-balancing method (3)









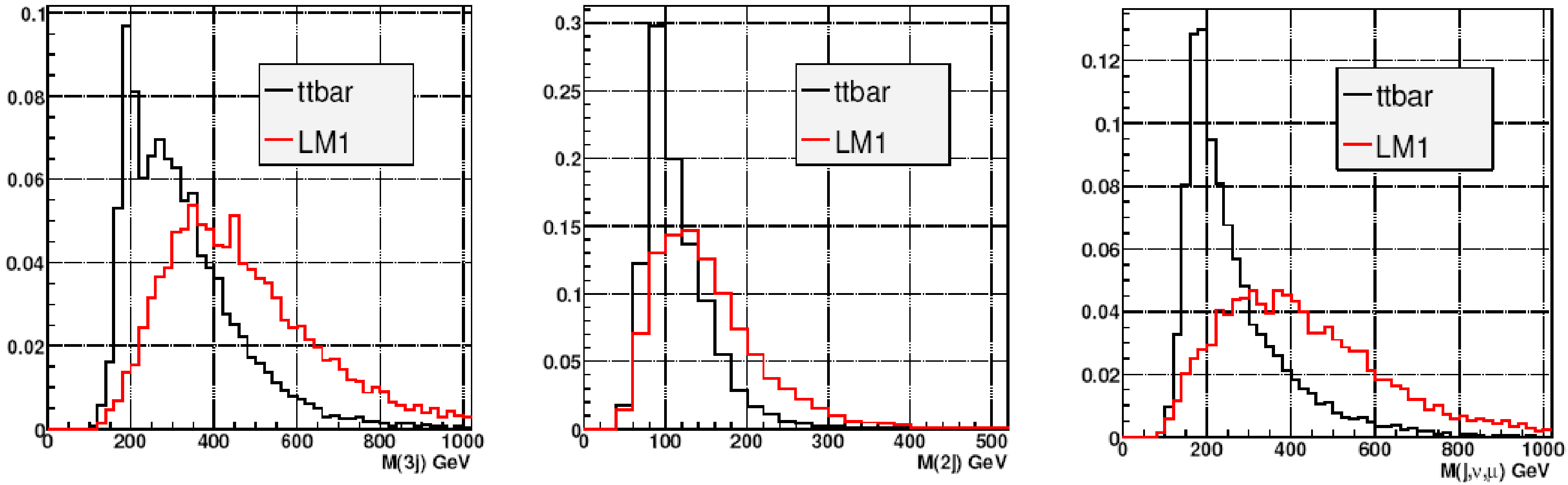
Robustness of α_T
for bad jet calibration
(MHT not so good)

ttbar background using χ^2(ttbar)

Construct x^2 by choosing the four highest pt jets and takeing the lowest x^2 permutation as the combination of jets assigned to hadronic W, b-jet of hadronic top, b-jet of leptonic top:

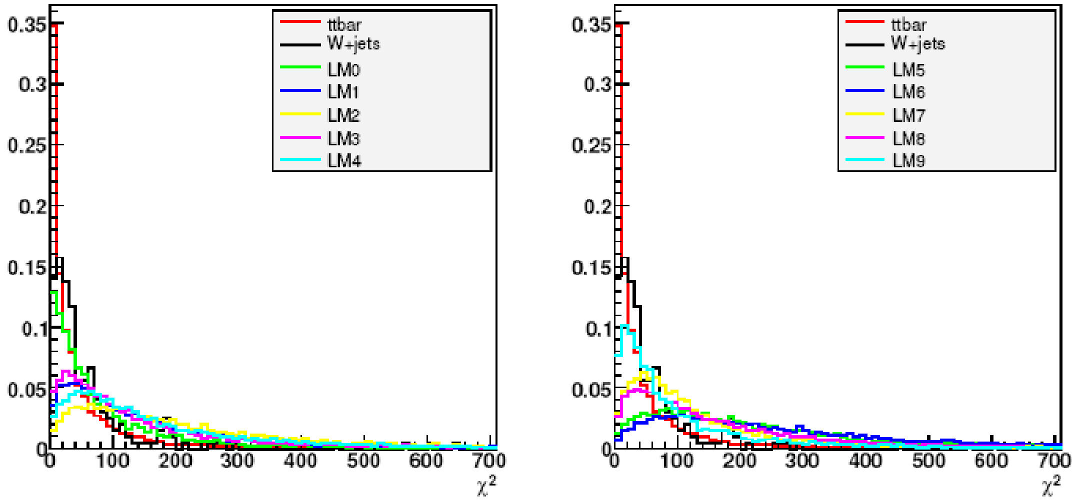
$$\chi^2(t\bar{t}) = \frac{(M_{j_1 j_2} - M_W)^2}{\sigma^2_{jj}} + \frac{(M_{j_1 j_2 j_3} - M_t)^2}{\sigma^2_{jjj}} + \frac{(M_{W_{\ell\nu} j_4} - M_t)^2}{\sigma^2_{\mu\nu j}}$$

$$\sigma^2_{jj} = 10.5, \sigma^2_{jjj} = 19.3, \sigma^2_{\mu\nu j} = 21.2$$



Masses for the lowest x^2

ttbar background using χ^2(ttbar) (2)



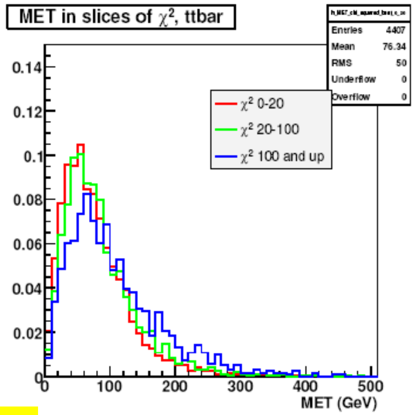
$x^2 < 20$:
50% ttbar
30% W+jets
2-24% SUSY

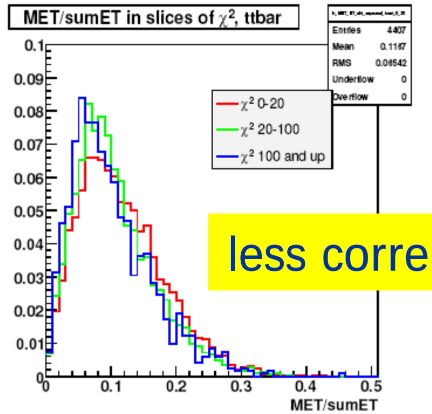
Try to mix this with 2nd variable exploiting the ABCD method

χ^2 vs MET/SUM(ET) ABCD method

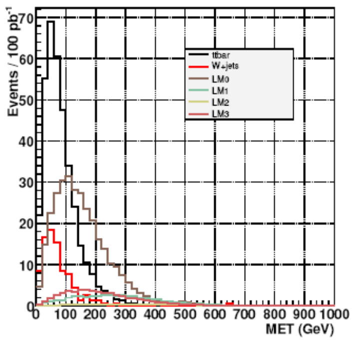
Strong correlation between x2 and MET, better (3 times smaller) for MET/SUM(ET)

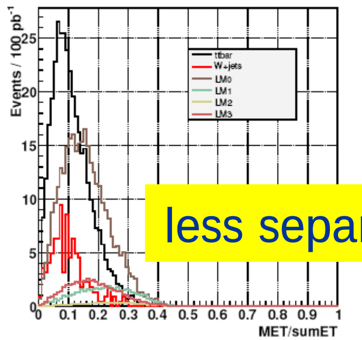
$$\sum E_T = \not{E}_T + E_T^{\text{jets}}$$



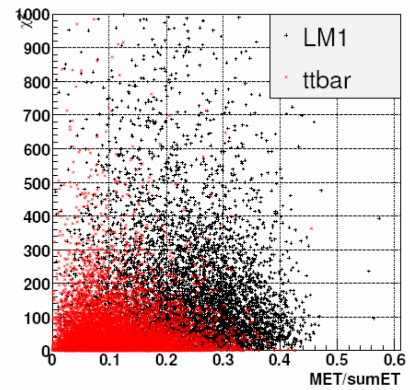


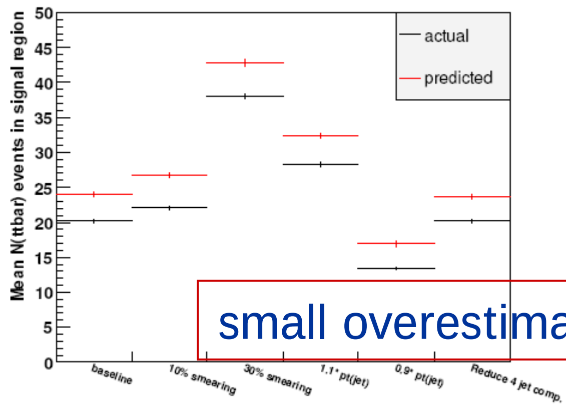
less correlation





less separation

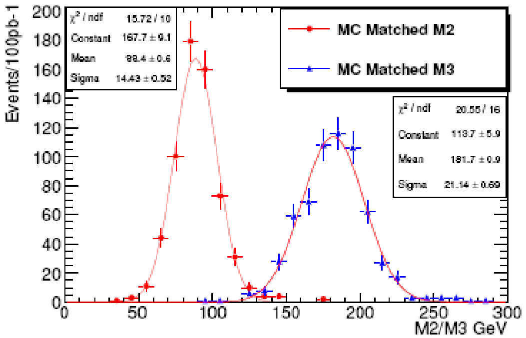




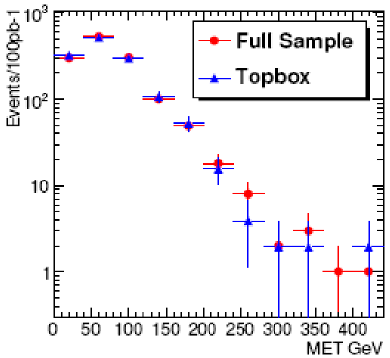
small overestimation of BG

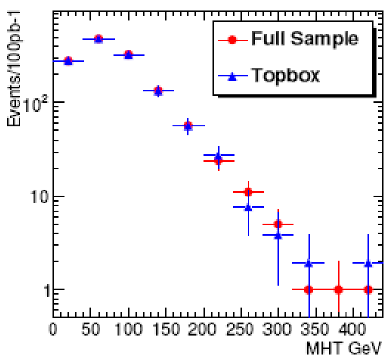
ttbar background from TopBox

Calculate 2-jets (M2) and 3-jets (M3) invariant mass combinations
Mi=SUM (4-Momentum(Jet) with Mass(Jet) set to 0)

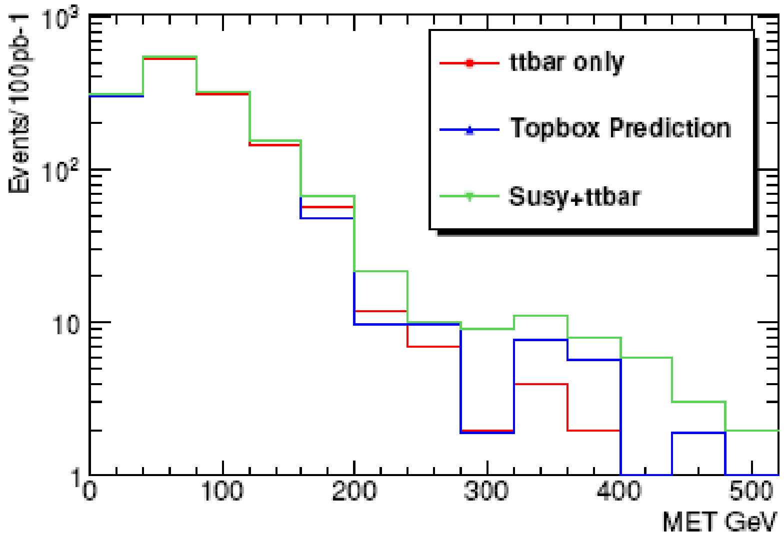


Choose MC based cuts:
70<M2<110 GeV
150<M3<210 GeV





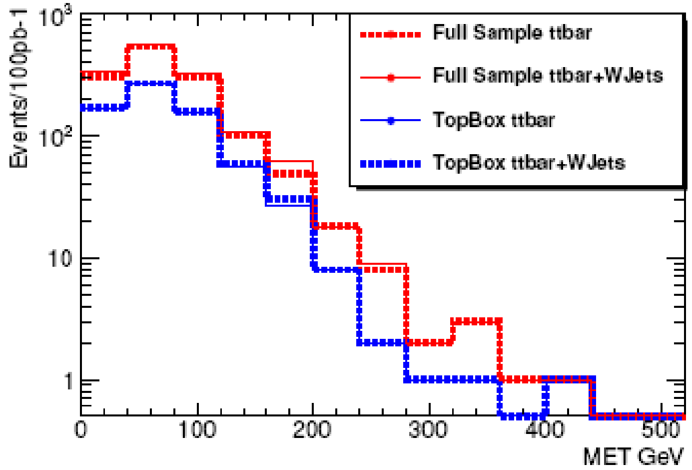
TopBox and full ttbar sample look similar

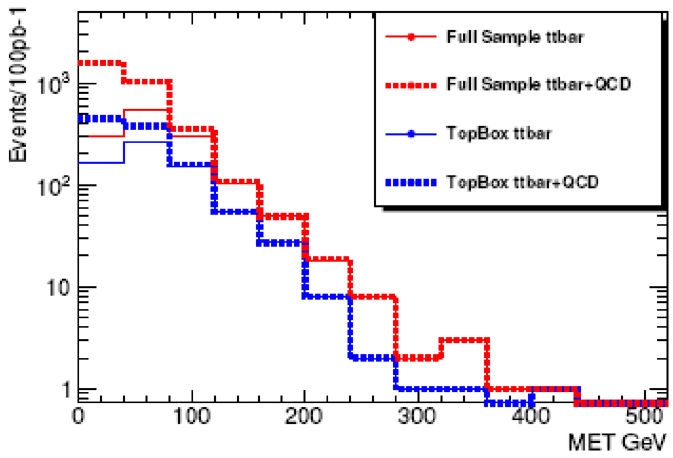


TopBox prediction with signal contamination

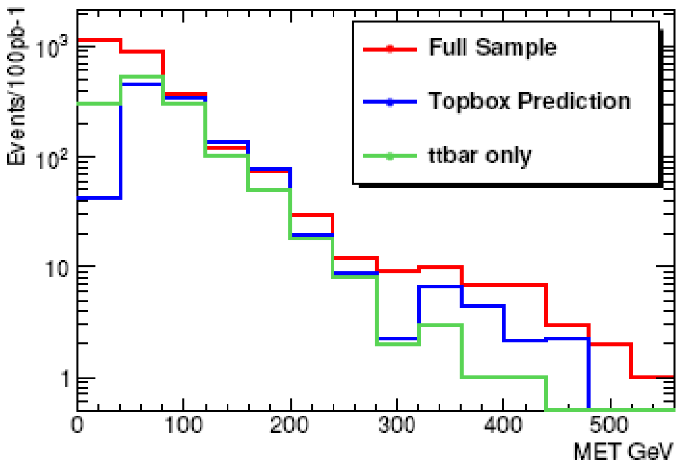
ttbar background from TopBox (2)

W+jets and QCD jets contribution:





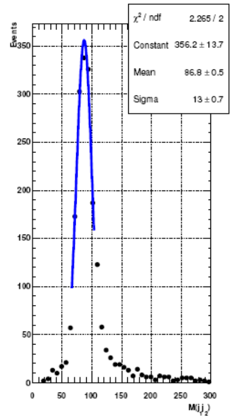
Slight overestimation with Topbox prediction



ttbar background with χ^2 sorting and sideband subtraction

Again, use χ^2:

$$\chi^2(t\bar{t}) = \frac{(M_{j_1 j_2} - M_W)^2}{\sigma_{jj}^2} + \frac{(M_{j_1 j_2 j_3} - M_t)^2}{\sigma_{jjj}^2} + \frac{(M_{W_{\ell\nu} j_4} - M_t)^2}{\sigma_{\mu\nu j}^2}$$



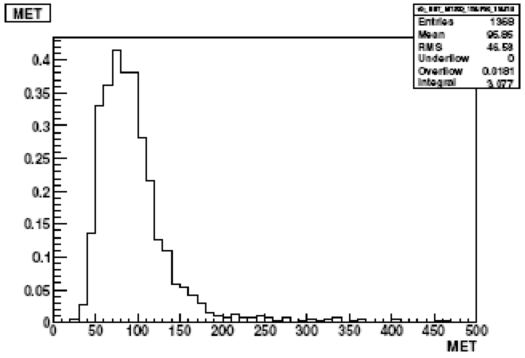
Then use hadronic W mass for sideband subtraction (reduce SUSY signal in ttbar sample, as SUSY is relatively flat):

- signal region: 60<M<120 GeV
- sideband: 120<M<180 GeV

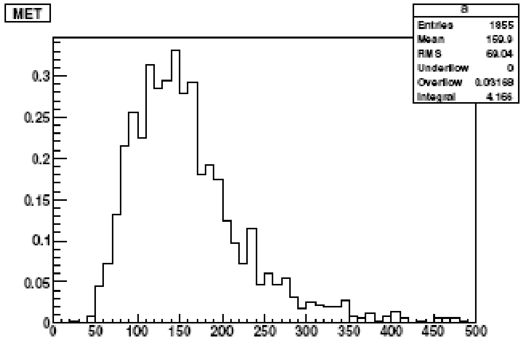
ttbar background tail also in sideband region

- subtraction will affect overall number of ttbar events
- only shape considered
- normalize sideband subtracted MET shape in low MET (50<MET< 100 GeV) region where ttbar (and W+jets) backgrounds dominate

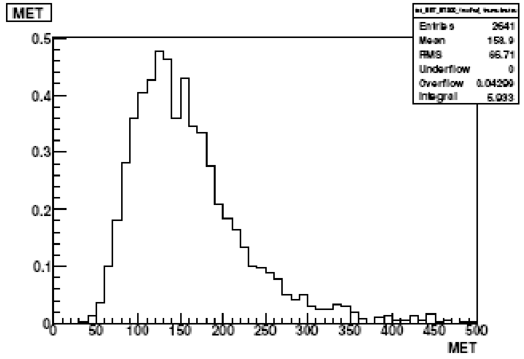
Dilepton background from ttbar



(a) MET from events with $1\nu_\mu$: $MET^{1\nu_\mu}_{True}$



(b) MET from events with $2\nu_\mu$ and $1\nu_\mu 1\nu_e$: $MET^{2\nu_\mu+1\nu_{mu}1\nu_e}_{True}$



(c) MET from events with $1\nu_\mu 1\nu_\tau$: $MET^{1\nu_\mu 1\nu_\tau}_{True}$

more next time...

Formulae collection

Important formulas

MET for muons:

$$MET = |p_{miss}| = \left| -\left(\sum_{calotowers} p_T - \sum_{muons} p_T^{deposit} - \sum_{muons} p_T \right) \right|$$

estimated CAL energy deposits of muons are removed and replaced by the measured pT from the global track fit

$$H_T = \sum_{\text{4 highest pT jets}} |p_T|$$

$$M_{eff} = \sum_{\text{4 highest pT jets}} |p_T| + p_T^{\text{single lepton with pT>10GeV}}$$

$$\sum E_T = \not{E}_T + jets$$