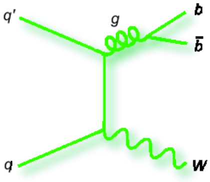
W/Z + Jets from Theory

The Story So Far

Stephen Mrenna

Computing Division
Fermilab

TeV4LHC@CERN 2005



Not-Top

More Interesting Than Top

Understanding W+Jets is Critically Important

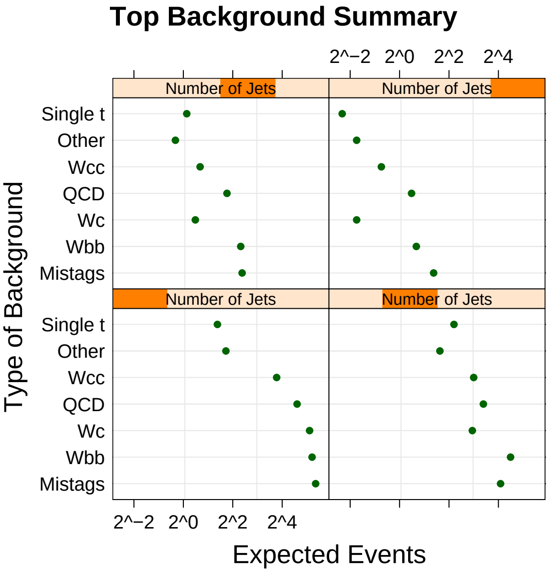
- Signature $Wb\bar{b}+X$ is common to unconfirmed Standard Model processes and many new physics processes
- we "know" that Standard Model top is there

$$\text{Top} \equiv \text{Data} - \text{Not-Top}$$

- As JES uncertainty is reduced (CDF m_t), understanding of Not-Top sets/limits understanding of Top
- Advanced (i.e. NN, DT) search techniques for single t exploit differences in many (11) kinematic variables
- Not-Top challenges our tools

Better tools $\Rightarrow$ more challenging questions

Not-Top Cocktail
CDF PRD, 162 ipb



Complicated Structure

$t\bar{t}$ contamination in Njets=3,4 (1.0,1,3)

work on Mistags,Wbb,QCD

QCD,Mistags reducible

trust basic properties of B,D hadron decays, e.g. K mesons

Mixing the Cocktail

Method 2

Monte Carlo ratio
$R = (W + b - jets)/(W + jets)$

- Common factors cancel

Measure $W + jets$ (no b-tag)

$\text{data}(W + b - jets) = R \times \text{data}(W + jets)$

Wcj/Wbb from Monte Carlo

- Several R's

Tools

- Tree−Level (MadGraph, Alpgen, etc.)
- Parton−shower (Pythia, Herwig, etc.)
- NLO−Level (MCFM, etc.)
- Combinations of these

Matrix Elements + Parton Showers

MLM Method

Parton shower and hadronization are essential for studying b-jets

- Parton shower W+Npartons but reject emissions that are too hard (i.e. each post-shower jet should have a pre-shower parton associated with it)
- Build up *inclusive* or *exclusive* samples (i.e. allow or disallow pure PS jets)
- $\delta R/R \sim 25 - 30\%$

Heavy Flavor (HF)

LEP, Run1 $\Rightarrow$ PS underestimates HF
PS inefficient in generating HF

- $P_{qq}(z) = \frac{1}{2}(z^2 + (1-z)^2)$

  no soft ($z \longrightarrow 0$) enhancement
  subleading log in PS

- Use ME with $b\bar{b}$ explicit

  Remove additional HF from PS

- R supplemented by phenomenological factor 1.5

Method 2 at Tree Level

Madevent (Stelzer and Maltoni)

X-Check

Graph	Cross Sect(fb)
Sum (Wbb)	8.934
Sum (Wjj)	1061.627
$ug \to e^+ \nu_e dg$	327.810
$u\bar{d} \to e^+ \nu_e gg$	257.060
$g\bar{d} \to e^+ \nu_e \bar{u}g$	137.300
$\bar{d}g \to e^+ \nu_e \bar{u}g$	48.591
$u\bar{u} \to e^+ \nu_e \bar{u}d$	47.425
$u\bar{d} \to e^+ \nu_e d\bar{d}$	36.644
$gu \to e^+ \nu_e dg$	34.445
$u\bar{d} \to e^+ \nu_e u\bar{u}$	29.816
...	...

$90 < M_{jj} < 110$ GeV, standard jets

$R \times 1.5 = 1.3\%$ (MLM $= 1.4\%$)

- $\langle R \rangle$ roughly the same

Many different topologies

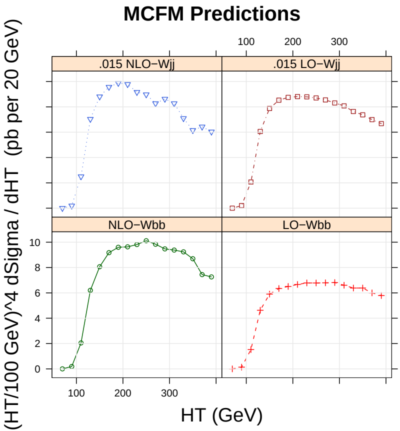
Dominant ones not $q\bar{q}$

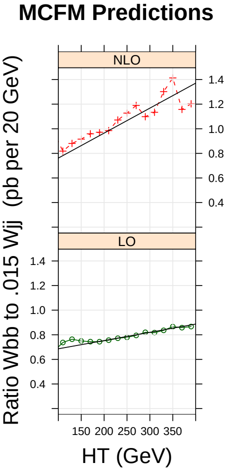
- again, no $z \to 0$ enhancement

Different topologies parton shower and hadronize differently

Many effects have to be modelled well to have a reliable prediction

MCFM Campbell and Ellis (see also Campbell & Huston)

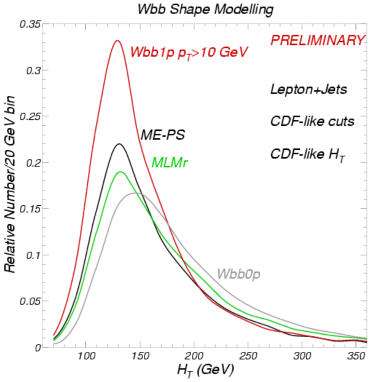




Significant change in normalization and shapes LO ⇒ NLO

Matrix Element-Parton Shower Matching

SM, PR *JHEP 0405:040,2004*, SM, JH, JC *in progress*



Testing Different Predictions

- Matching scheme needed to make inclusive predictions with hard emissions
- Pseudoshower Method (ME-PS) reweights matrix elements to look like parton showers where they should. Motivated by Catani et al., but more flexible and tuned to Pythia, Herwig, etc.

Review of Matching

Pseudo-Shower Method

1. Generate $W + N$ parton events, applying a cut ${p_T}^2_{cut}$ on shower p_T^2 (p_T^2 for ISR, $z(1-z)m^2$ for FSR)
2. Form a p_T^2-ordered parton shower history
3. Reweight with $\alpha_s(p_T^2)$ for each emission
4. Add parton shower and keep if no emission harder than $p^2_{T\,cut}$: (save this event)
5. Remove softest of N partons, fix up kinematics, add parton shower and keep if no emission harder than $p^2_{T\,softest}$
6. Continue until no partons remain, or an emission is too hard
7. If not rejected, use the saved event

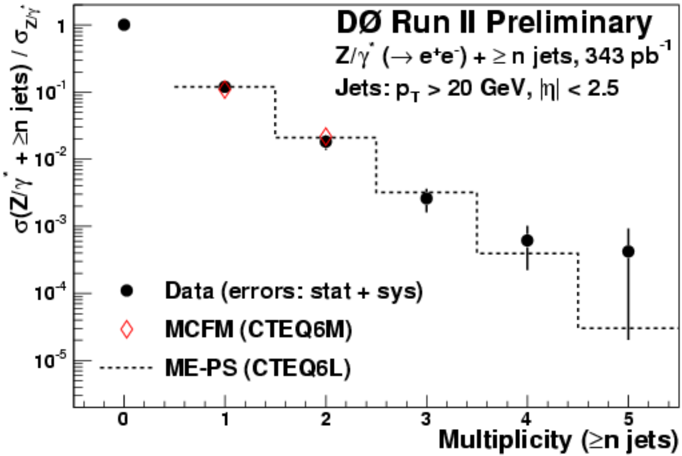
Why it works

- For each N, PS does not add any jet harder than $p^2_{T\,cut}$
- Can safely add different N samples with no double-counting
  - Apply looser rejection on highest N
- Pseudo-showers assure correct PS limit, while retaining hard emissions
  - Interpolates between limits

$b\bar{b}$ Modifications

- Apply no cuts on $b\bar{b}$ pair in ME
  - Efficient generation of HF
  - "exact" kinematics
- When $b\bar{b}$ pair is removed from PS history, skip the pseudoshower
  - ME entirely (no Sudakov)
- Use $\alpha_s\left(\frac{1}{4}m^2\right)$ for weight

Cross check on Run2 data

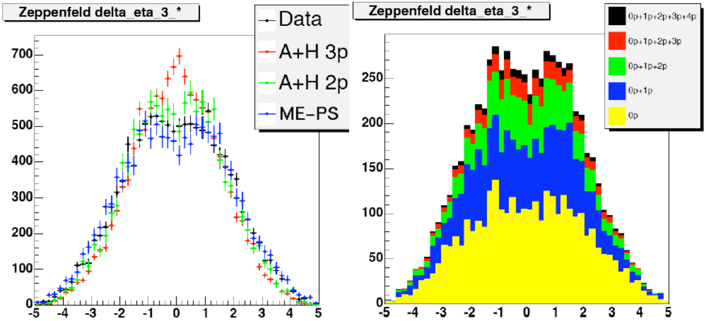


Includes up to $Zjjj$, $j = q, g$

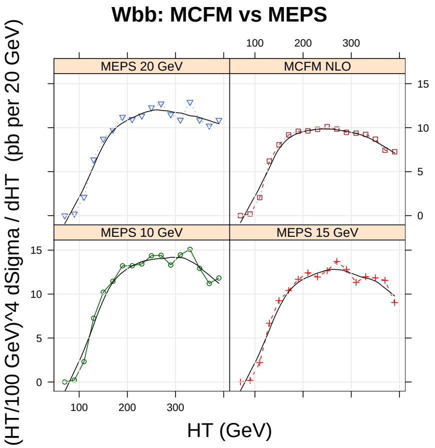
Kinematic comparisons with Run2 data

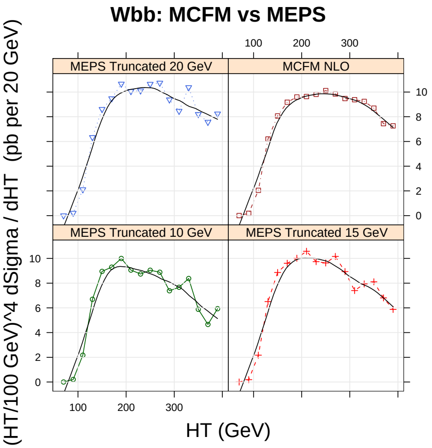
Tag jets > 8 GeV/c; 3rd jet > 8 GeV/c; $\Delta\eta>1$

- ME–PS decomposition



$$\eta_3^* = \eta_3 - \frac{\eta_1+\eta_2}{2}$$, A+H ≡ Alpgen+Herwig

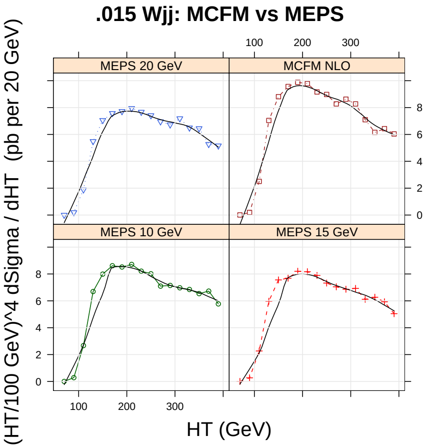




Matched Datasets have a systematically larger rate and different shape

Truncated Datasets contain only $Wb\bar{b} + Wb\bar{b}j$

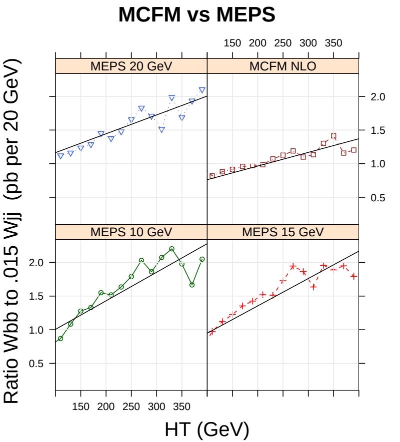
HO topologies modify shape

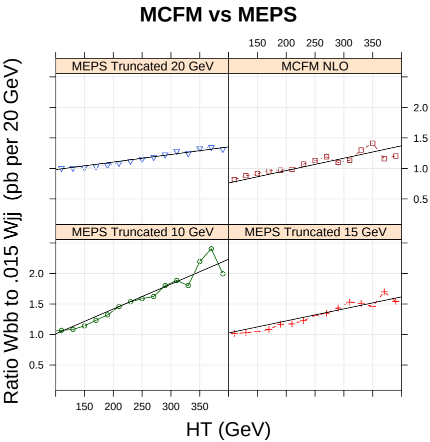


Wjj Matched Datasets have less variation with cutoff

Matched normalization here is smaller (no skipped Sudakov)

Stiffer shape (HO topologies)





Matched Datasets have consistently steeper slopes (note: MCFM steeper than LO)

Truncated Datasets contain only $Wb\bar{b} + Wb\bar{b}j$

Slopes more consistent with MCFM

Conclusions

We need to understand Not-Top

- MCFM and Matched ME-PS predictions allow us to study methods for determining the ratio $R = Wbb/Wjj$
- MCFM already indicated a stiffer dR/dH_T spectrum than "standard" matching methods

  Campbell and Huston, confirmed here

- Pseudo-shower predictions are significantly stiffer than MCFM

  Topologies up to $Wb\bar{b}jjj$ are included and affect the dR/dH_T tail

- Many questions remain
  - Which distributions are the most important for testing different predictions?
  - Is there a kinematic difference between the different components of Not-Top? Can we discriminate Wbb, Wjj and Wcj?

Extra Slides

Improved Search for Single Top Quark Production at DØin Run II

http://www-d0.fnal.gov/Run2Physics/top/public/winter05/singletop/

95% Confidence Level Expected/Measured Upper Limits
(after final selections, with systematics, using Bayesian statistics)

		s-channel	t-channel
Cut-Based	Electron	11.4/10.8	15.1/17.5
	Muon	13.0/15.2	18.1/13.0
	Combined	9.8/10.6	12.4/11.3
Decision Trees	Electron	6.9/7.9	9.3/13.8
	Muon	7.3/14.8	10.9/7.9
	Combined	4.5/8.3	6.4/8.1
Neural Networks	Electron	7.0/7.3	8.8/7.5
	Muon	7.0/8.7	9.5/7.4
	Combined	4.5/6.4	5.8/5.0

Single Top

New Physics Warm-Up

- current state of single-Top is where we will be at the LHC with a few quality fb^{-1}
- the size of other NP signals
- it is a playground for new analysis techniques
- it challenges our tools

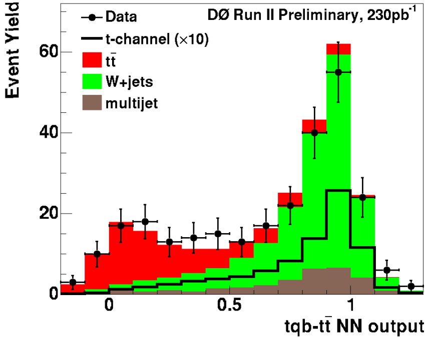
Not specific to NN analyses: may be more sensitive

Many (11) Kinematic Variables

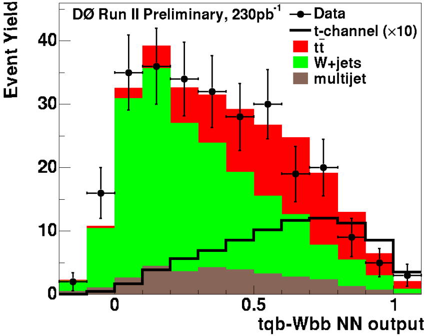
	Signal-Background Pairs			
	tb		tqb	
	Wbb	$t\bar{t}$	Wbb	$t\bar{t}$
Individual object kinematics				
$p_T(\text{jet1}_{\text{tagged}})$	√	√	√	—
$p_T(\text{jet1}_{\text{untagged}})$	—	—	√	√
$p_T(\text{jet2}_{\text{untagged}})$	—	—	—	√
$p_T(\text{jet1}_{\text{nonbest}})$	√	√	—	—
$p_T(\text{jet2}_{\text{nonbest}})$	√	√	—	—
Global event kinematics				
$M_T(\text{jet1},\text{jet2})$	√	—	—	—
$p_T(\text{jet1},\text{jet2})$	√	—	√	—
$M(\text{alljets})$	√	√	√	√
$H_T(\text{alljets})$	—	—	√	—
$M(\text{alljets} - \text{jet1}_{\text{tagged}})$	—	—	—	√
$H(\text{alljets} - \text{jet1}_{\text{tagged}})$	—	√	—	√
$H_T(\text{alljets} - \text{jet1}_{\text{tagged}})$	—	—	—	√
$p_T(\text{alljets} - \text{jet1}_{\text{tagged}})$	—	√	—	√
$M(\text{alljets} - \text{jet}_{\text{best}})$	—	√	—	—
$H(\text{alljets} - \text{jet}_{\text{best}})$	—	√	—	—
$H_T(\text{alljets} - \text{jet}_{\text{best}})$	—	√	—	—
$M(\text{top}_{\text{tagged}}) = M(W,\text{jet1}_{\text{tagged}})$	√	√	√	√
$M(\text{top}_{\text{best}}) = M(W,\text{jet}_{\text{best}})$	√	—	—	—
$\sqrt{\hat{s}}$	√	—	√	√
Angular variables				
$\Delta R(\text{jet1},\text{jet2})$	√	—	√	—
$Q(\text{lepton}) \times \eta(\text{jet1}_{\text{untagged}})$	—	—	√	√
$\cos(\text{lepton}, Q(\text{lepton})\times z)_{\text{top}_{\text{best}}}$	√	—	—	—
$\cos(\text{lepton},\text{jet1}_{\text{untagged}})_{\text{top}_{\text{tagged}}}$	—	—	√	—
$\cos(\text{alljets},\text{jet1}_{\text{tagged}})_{\text{alljets}}$	—	—	√	√
$\cos(\text{alljets},\text{jet}_{\text{nonbest}})_{\text{all jets}}$	—	√	—	—

Network Outputs

$t\bar{t}$ Training



$Wb\bar{b}$ Training



- How do we convince ourselves of a signal?
- How can we improve upon the search?