

MC Validation for V+jets Production



OUTLINE

I. Introduction II. "Soft Physics" Tunes III. Testing ME-PS Matching IV. V+Heavy Flavor Quark(s) V. Conclusions





- What are the important features to get a reliable description of the V+jets phenomenology at Hadron Colliders?
- Need to get <u>both</u> hard (ME pQCD) and soft physics (resummation, UE,...) right, plus correct values for flavor ratios (V+HF-jets / V+LF-jets)
 - <u>Shapes:</u> different variables depend on different ingredients
 - Resummation $\rightarrow p_T(V) => JES$, bkgd kinematics, mE_T calibration,...
 - ME \rightarrow Multijet topologies: N_{iets}, p_T(jets), $\Delta \phi(j,j')$, $\Delta R(j,j')$,...
 - HF fractions: depend on heavy quarks PDFs, $\mathcal{P}(g \rightarrow QQ)$, FF, (exp) HF-jets tagging
 - <u>Normalizations:</u>
 - Need at least NLO (QCD)
 - Some NNLO (QCD) results are available, but might have to check against NLO (EW) in certain regions of phase space!!!
 - Might normalize directly on data in some cases
- Don't forget this is also an experimental challenge:
 - Have to rely on an accurate detector simulation (dead material, realistic and time dependent noise and dead/bad channels maps,...)
 - Have to rely on a good trigger simulation (or to measure trigger efficiencies in data)





because it occurs at $Q_{2nadry} \ll Q_{hard scatt}$, but it's still a perturbative process





- MI: multiple pp or ppbar collisions in a given bunch crossing
- MPI: multiple parton interactions in a given pp or ppbar collision
- Minimum Bias:
 - it's a trigger condition (not a physics process)
 - low p_T QCD, (SD single arm), DD are the contributing subprocesses
 - Note: UE in most hard collisions also fire this trigger term
- UE: interactions of spectator partons, including possible MPI
- Pile-Up:
 - Definition: signals overlapping in time in a sub-detetctor
 - Consequence: may add up events from different bunch crossings
 - It's a detector issue, not a physics process, nor a trigger condition!









UE Extrapolation from Tevatron to LHC

R. Field



UE is much harder at the LHC than at the TEVATRON!!!

But, large uncertainties in the extrapolation => will have to re-tune UE on LHC data!!!



Soft Physics Tunes:













MCFM

Matrix Elements: NLO Normalization

- Authors: J.M. Campbell, R.K. Ellis
- Programming Language: Fortran 77
- Set of hard-coded LO & NLO ME (SM, Higgs)
- Output: total and differential cross sections @ LO and NLO
- Parton level cuts possible
- N_{LO}(FS partons)<4; N_{NLO}(FS partons)<3
- Interface to LHAPDF
- Large nber of SM processes
- NLO ME have <u>massles</u> b and c quarks, LO can account for these masses
- Beyond NLO: Note that
 - σ_{NNLO} (p+p/pbar→t+tbar) are available ref₁: N. Kidonakis and R. Vogt, Phys.Rev.D68 (2003) 114104 ref₂: M. Cacciari et al., JHEP 0404 (2004) 068
 - σ_{NNLO} (p+p/pbar→W/Z) are available ref₁: R. Hamberg, W.L. van Neerven, T. Matsuura, Nucl.Phys.B359 (1991) 343; Err.-Ibid. B644 (2002) 403 ref₂: K. Melnikov and F. Petriello Phys. Rev. D74 (2006) 114017



Matrix Elements:



LO Matched to Parton Shower

Alpgen v2

- Authors: ML. Mangano et al.
- Programming Language: Fortran 77 (90)
- Set of hard-coded LO ME + Alpha algorithm (SM, Higgs, user-supplied NP)
- Parton Shower: interface to Pythia or Herwig
- MLM Matching between ME-PS (provided)
- LO Normalization, LL Shapes
- Recent CTEQ & MRST PDF files available
- Large nber of FS partons (and γ)
- Special decays:
 - Taus: Tauola interface to Pythia & Herwig
 - B/C-Hadrons: EvtGen interface to Pythia & Herwig



<u>W(→ev)+jets @ CDF Run II</u>













Matrix Elements:



LO Matched to Parton Shower

Sherpa

- •Authors: F. Krauss et al.
- Programming Language: C++
- « Generator » of LO ME (SM, Higgs, SUSY, ED) \rightarrow AMEGIC++
- Parton Shower LUND string \rightarrow APACIC++
- CKKW Matching between ME-PS
- LO Normalization, LL Shapes
- Interface to LHAPDF v3

. . .

- A bit of add'l overhead at startup, but very powerful & self-contained package
- Limitation in the nber of FS partons???
- Reproduces MC@NLO shapes in many cases
- NP: Can generate consistently signal and background
- Special decay module: \rightarrow HADRONS++
 - Taus: self-contained, spin-corr. uses helicity amplitudes



B,

LO Matched to Parton Shower

Matrix Elements:

<u>*Y**/Z(→ee)+jets @ D0 Run II</u>

 $\int \mathcal{L}dt = 0.95 \, fb^{-1}$

- Trigger: e and di-e
- 2 OS e, $p_T(e)$ >25 GeV, $|\eta(e)|$ <1.1 and $|\eta(e)|$ <2.5,
- e-likelihood, E/p cut,...
- Use evts within $70 < M_{ee} < 100 \text{ GeV}$
- Jets: midpoint cone w/ split-merge
 - ΔR=0.5
 - •
- MC: Sherpa v1.0.6
 - PDF: CTEQ6L1
 - ME: Z+0j, Z+1j, Z+2j, Z+3j * CKKW Matching



Matrix Elements: LO Matched to Parton Shower

Sherpa







CDF & D0 Run I





Matrix Elements:



LO Matched to Parton Shower

MC@NLO

- Authors: S. Frixione, B.R. Webber
- Programming Language: Fortran 77
- Set of hard-coded NLO ME (SM, Higgs)
- Parton Shower: Herwig (cluster model)
- Special Matching between ME and Herwig PS
- Small nber of partons in FS (do not use for multijet topologies)
- NLO Normalization, NLL Shapes (up-to 1st emission, rest is PS)
- Reproduces PS in soft/collinear regions and ME in hard/wide angle emissions





MC@NLO





(w/ s-quark PDF evolved from fixed target expt to Tevatron Q²)

• Prospect: w/ more \mathcal{L}_{int} will constrain s-quark PDF







W+bb @ CDF Run II



- Measure: $\sigma(W+bb) = 0.90\pm0.32 \text{ pb} (E_T(j)>20 \text{ GeV}, |\eta|<2.0)$
- Theory: $\sigma(W+bb) = 0.74 \pm 0.18 \text{ pb} (\text{ALPGEN: LO QCD})$
- NB: no anomalies « à la super b-jets » found in CDF & D0 Run II



V+HF-jets



DØ

100

 L_{xy}/σ_{xy}

Data

50



(a)

15

Phys. Rev. Lett. 94, 161801 (2005)

(b)

DØ

$$\int \mathcal{L}dt = 0.18 \, fb^{-1}$$

• Events Sel.:

- $pT(e/\mu) > 20 \text{ GeV}$
- $80 < M_{ee} < 100 \text{ GeV}$
- $65 < M_{uu} < 115 \text{ GeV}$

• Jets:

• $|\eta| < 2.5$

• Extract b-jet fraction using $p_{T}(jets) > 20 \text{ GeV}$

• First measurement of this kind (at the time)

Entries / 10 GeV/c Entries / -Expectation -Expectation Charm+Mistag Charm 15 Bkgd+Mistag Bkgd 10 10 5 100 50 150 $R = \frac{\sigma[Z(\to \ell^+ \ell^-) + b + X]}{\sigma[Z(\to \ell^+ \ell^-) + jets]} = 0.021 \pm 0.004(stat)^{+0.002}_{-0.003}(syst)$

Data

• Conclusion: in agreement wrt NLO QCD using MCFM: $R = 0.018 \pm 0.004$

• Prospect: window to check b-quark PDF (not enough stat. to constrain)

• $\Delta R=0.5$



least 1 tagged jet



 $M_{\rm g}$ (GeV/c²)

UE and hadronization corrections:

 $\int \mathcal{L}dt = 1.5 \, fb^{-1}$

Sec. Vtx mass fit in events with at

- derived using Pythia (Tune A)
- sytematics=|Pythia-Herwig|

	CDF Run II Preliminary Data	PYTHIA	MCFM NLO	MCFM NLO
				+ue +had
$\sigma(Z^0 + b \operatorname{jet})$	$0.94 \pm 0.15 \pm 0.15 \; \mathrm{pb}$	0.84	0.51	0.56
$\sigma(Z^0 + b \operatorname{jet}) / \sigma(Z^0)$	$0.00369 \pm 0.00057 \pm 0.00055$	0.0035	0.0021	0.0023
$\sigma(Z^0 + b \operatorname{jet}) / \sigma(Z^0 + \operatorname{jet})$	$0.0235 \pm 0.0036 \pm 0.0045$	0.0218	0.0188	0.0177

• Conclusion: $\sigma(Z+b)$ is 2σ away from MCFM (NLO QCD)



V+HF-jets



<u> Di-b-jets @ CDF Run II</u>

$$\int \mathcal{L}dt = 0.26 \, fb^{-1}$$

Note that $\Delta \phi(b_1, b_2)$ is dominated by:

- \bullet the flavor excitation and gluon splitting sub-processes at small $\Delta \varphi$
- \bullet the flavor creation sub-process at large $\Delta \varphi$







- Soft Physics Tunes:
 - Possibility to tune Soft Physics w/ PS generators
 - UE and Min. Bias collision model => good starting point for the LHC but will have to be re-done
 - Tune $p_T(V)$ using (artificially) large intrinsic k_T , OR
 - Use analytical q_T resummation to predict $p_T(V)$, then reweight using: [MC truth $p_T(V)$] / [Resummed $p_T(V)$], OR
 - Measure $p_T(V)$ in data, then reweight
- Multijet Topologies:
 - Tools:
 - Alpgen+PS * MLM Matching
 - Sherpa * CKKW Matching
 - Also CompHep, Madgraph,... but no public matching code
 - Results:
 - Good agreement in shapes wrt to Tevatron Run II data, provided NLO norm.





- V+Heavy Flavor:
 - Measure V+HF / V+jets (many systematics cancel in these ratios)
 - Good agreement* wrt NLO QCD



Prospects



A Specific b-jet Energy Scale

• Will (slightly) improve both central values and resolutions

• In place at CDF: Signal: Z→bb $\int \mathcal{L}dt = 0.58 fb^{-1}$





• Only partly available at D0: Correct only for the calorimeter response difference wrt LF-jet: O(+5 %) relative

Signal: γ+b



Additional gains:

- Stat: b-tagging @ trigger LVL at D0
- Energy Flow (or simply « TrackCal » jets at D0)





BACK-UP





Use LO a	Parameter	Tune DW	Tune D6	Tune QW	Tune QK	NLO Structure Function!	
with $\Lambda = 192$ MeV!	PDF	CTEQ5L	CTEQ6L	CTEQ6.1	CTEQ6.1		
	MSTP(2)	1	1	1	1		
UE Parameters	MSTP(33)	0	0	0	1		
	PARP(31)	1.0	1.0	1.0	1.8 🛰		
	MSTP(81)	1	1	1	1	K-factor (T. Sigstrand)	
	MSTP(82)	4	4	4	4	(1. Sjostrand)	
	PARP(82)	1.9 GeV	1.8 GeV	1.1 GeV	1.9 GeV		
	PARP(83)	0.5	0.5	0.5	0.5	Tune A energy dependence!	
	PARP(84)	0.4	0.4	0.4	0.4		
	PARP(85)	1.0	1.0	1.0	1.0		
	PARP(86)	1.0	1.0	1.0	1.0		
ISR Parameter	PARP(89)	1.8 TeV	1.8 TeV	1.8 TeV	1.8 TeV		
	PARP(90)	0.25	0.25	0.25	0.25		
	PARP(62)	1.25	1.25	1.25	1.25		
	PARP(64)	0.2	0.2	0.2	0.2		
	PARP(67)	2.5	2.5	2.5	2.5		
	MSTP(91)	1	1	1	1		
	PARP(91)	2.1	2.1	2.1	2.1		
Intrinsic KT	PARP(93)	15.0	15.0	15.0	15.0		





	Parameter	Tune DWT	ATLAS	Tune D6T	Tune QWT	Tune QKT	NLO Structure Function
Use LO α_s with $\Lambda = 192$ MeV!	PDF	CTEQ5L	CTEQ5L	CTEQ6L	CTEQ6.1	CTEQ6.1	
UE Parameters	MSTP(2)	1	1	1	1	1	
	MSTP(33)	0	0	1	1	1	
	PARP(31)	1.0	1.0	1.0	1.0	1.8 🛰	K-factor
	MSTP(81)	1	1	1	1	1	
	MSTP(82)	4	4	4	4	4	(1. Sjostrand)
	PARP(82)	1.9409 GeV	1.8 GeV	1.8387 GeV	1.1237 GeV	1.9409 GeV	
	PARP(83)	0.5	0.5	0.5	0.5	0.5	Tune A energy dependence
ISR Parameter	PARP(84)	0.4	0.5	0.4	0.4	0.4	
	PARP(85)	1.0	0.33	1.0	1.0	1.0	
	PARP(86)	1.0	0.66	1.0	1.0	1.0	
	PARP(89)	1.96 TeV	1.0 TeV	1.96 TeV	1.96 TeV	1.96 TeV	
	PARP(90)	0.16	0.16	0.16	0.16	0.16	
	PARP(62)	1.25	1.0	1.25	1.25	1.25	
	PARP(64)	0.2	1.0	0.2	0.2	0.2	
	PARP(67)	2.5	1.0	2.5	2.5	2.5	
	MSTP(91)	1	1	1	1	1	
	PARP(91)	2.1	1.0	2.1	2.1	2.1	
Intrinsic KT	PARP(93)	15.0	5.0	15.0	15.0	15.0	

Soft Physics Tunes: q_T **Resummation**



p_T(V) from Analytical Resum. vs D0 Run II



40 < M_{ee} < 200 GeV 0 < q_T < 260 GeV |Y| < 3







<u>γ+b @ CDF Run II</u>







Particle identification







- There are 2 classes of Event Generators:
 - 1. The « Parton Shower » Generators (aka MC)
 - 2. The « Matrix Element » Generators (aka FOME)

Basic « How To »

<u>PS</u>

- Start from $2 \rightarrow N$ (LO) ME, w/ N up to 2/3
- Radiate IS and FS gluons
- Start parton showers from those
 - note that each branching iteratively decreases Q(node i)< Q(hard scatter)
 - ... downto Q(node f)~ Λ_{QCD}
 - => it creates a natural link between the hard scatter and the hadronization scale
- Underlying event (including MPI)
- Decay of unstable particles
- •...
- SHAPES: accurate at Leading Log (LL)
- RATES: purely LO normalization is retained







- There are 2 classes of Event Generators:
 - 1. The « Parton Shower » (aka MC) Generators
 - 2. The « Matrix Element » (aka FOME) Generators

<u>PS</u>

Advantages:

• Generate inclusive samples (generate in one shot, directly comparable to data)

Drawbacks:

Does not account for quantum interferences between compatible transition amplitudes
Additional jets are mostly produced in special corners of the available phase space (due to the soft and collinear approximation)

Examples:

Pythia, Herwig, Isajet,... Ariadne (Dipole Showers)

<u>ME</u>

• Have to generate exclusive samples (generate in many steps, then mix and merge the samples)

Advantages:

Drawbacks:

Explicitely accounts for quantum interferences between compatible transition amplitudes
Additional jets are correctly produced over the available phase space, but not in some corners (due to singularities of the ME in the soft and collinear regions)

Examples:

Alpgen, Sherpa, Comphep, Vecbos,... mc@nlo,...





Alpgen & the « MLM » Matching

- Generate a parton level configuration based on LO ME, w/ N_{part} hard partons
- Apply kinematical cuts on those configs: p_{Tpart}^{min} and $\Delta R_{part-part}$
- Perform PS (no showers veto, no Sudakov form factor reweighting)
- Cluster the partons using a jet reco algo: $E_{T_j}^{min}$, ΔR_j
- Match parton to parton jets:
 - for each ME parton select a parton jet based on $min(\Delta R_{part-j})$
 - if this min($\Delta \mathbf{R}_{part-j}$) < $\Delta \mathbf{R}_{j}$ the parton is matched
 - a parton jet can be matched only to a single ME parton
 - Exclusive matching:
 - Keep the event only if each parton is matched to a jet & $N_{part} = N_{jets}$
 - Inclusive matching:
 - Keep the event only if each parton is matched to a jet & N_{part}<N_{jets}
- ickkw option: reweight events by $\alpha_{S}(k_{T}{}^{2})/\alpha_{S}(Q_{HS}{}^{2})$ calculation at each node of the PS
- Public code available in Alpgen v2 (LHA interface to Herwig and Pythia)
 . REF:
- . Alpgen Doc: M.L. Mangano, http://m.home.cern.ch/m/mlm/www/alpgen/alpdoc.pdf
- . hep-ph/0602031, "Matching Parton Showers and Matrix Elements" by S. Hoeche, F. Krauss, N. Lavesson, L. Lönnblad, M.L. Mangano, A. Schaelicke, S. Schumann





Sherpa & the « CKKW » Matching

- . Applies to e⁺e⁻ -> jets+X (slightly modified for hadron collisions)
- . Try to separate the contributions of ME and PS in the phase space using a $k_{\rm T}$ cluster paramater $y_{\rm ini}$
- . ME and PS are matched w/:
 - modified ME (Sudakov FF and $\alpha_{S}(k_{T}^{2})/\alpha_{S}(Q^{2})$ reweighting)
 - modified showers (vetos to cancel y_{ini} dependance at NLL accuracy)
- . Reminder k_T clustering (Durham algo):

$$\mathbf{y}_{ij} = \frac{2 \times \min(\mathbf{E}_i, \mathbf{E}_j)^2 \times (1 - \cos \theta_{ij})}{\mathbf{Q}^2} = \frac{\mathbf{k}_T^2}{\mathbf{Q}^2}$$

- if $y_{ij} > y_{cut}$ the objects i and j are resolved
- elseif $y_{ij} < y_{cut}$ objects i and j are clustered => 4- p_{ij} = 4- p_i + 4- p_j
- Note: Process independent procedure!!!

. REF:

S. Catani, F. Krauss, B. Webber, R. Kuhn, JHEP11 (2001) 063 "QDC Matrix elements + parton showers"





S. Höche

CDF Run I



Sherpa Tuned UE





S. Höche

CDF Run II

Sherpa x K-factor







Pythia's Variant of the « CKKW » Matching

- Start w/ parton level configuration based on LO ME
- Feed Pythia w/ this using the full flavor and color flow
- Perform a k_T clustering to determine "nodal values"
- Events are reweighted by $\alpha_{s}(k_{T})/\alpha_{s}(M_{Z})$ factor for each cluster $e^{+}e^{-} \rightarrow \gamma^{*}/Z \rightarrow 2p: 0 \alpha_{s}$ reweighting,

3p: $1 \alpha_s$ reweighting,...

```
np: (n-2) \alpha_s reweighting
```

- The clustering yields a parton shower (PS) history where each line is weighted by a Sudakov form factor
- Primary partons are showered in Pythia from Q_{max} @ nodal value down to $Q = \Lambda_{QCD}$
- In the PS each emission w/ $k_T > k_T min$ is vetoed!!!
- To account for unknown higher order contributions (mainly leading logs), the events w/ largest nber of partons is not vetoed
- . REF: S. Mrenna and P. Richardson, JHEP05 (2004) 040

"Matching matrix elements and parton showers with HERWIG and PYTHIA"



II. The FOME Generators



<u> Madgraph+Pythia</u>

• Patriot samples provided by S. Mrenna for CDF and D0 Run II











- **<u>1. Final State Showers (forward evolution)</u>**
- Sudakov Form Factor:

$$\Delta a \to bc(t_0, t) = \exp\left(-\int_{t_0}^t \frac{dt'}{t'} \int dz \frac{\alpha_s(k_T^2)}{2\pi} P_a \to bc(z)\right)$$

• Interpretation: $1-\mathcal{P}(\text{for a to split into b+c between } t_0 \text{ and } t)$

1

2. Initial State Showers (backward evolution)

- In principle equivalent to FS showers
 - but both end fixed => quite different in practice
- DGLAP Equations:
 - Start at HS Q²
 - Evolve backwards
 - Weight w/ PDFs at each x and t