Unfolding with CUETP8M1 and MadGraph

2015 Data and MC samples

Patrick L.S. Connor Paolo Gunnellini Hannes Jung

Deutsches Elektronen-Synchrotron

20 September 2016

Introduction

Questions

1. are CUETP8M1 and MadGraph consistent?
2. are `RooUnfoldBayes` and `TUnfold` giving similar results?

Tests

1. matching with different ΔR limits
2. closure test and parameter tuning (number of iterations for `RooUnfoldBayes` and bias for `TUnfold`)
3. toy MC (with a Gaussian fit)

Application

1. to MC
2. to data
3. but still under study

Tests

- realised in parallel for CUETP8M1 and MadGraph
- from CMSSW 76X (2015 Data)
- working with inclusive jet (for comparison)

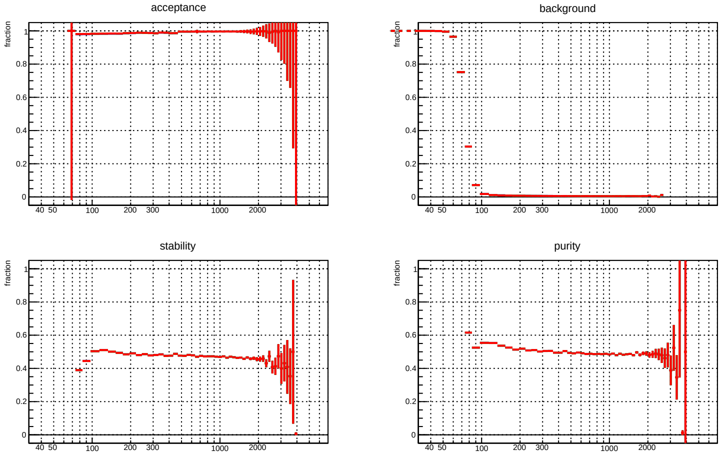
Matching

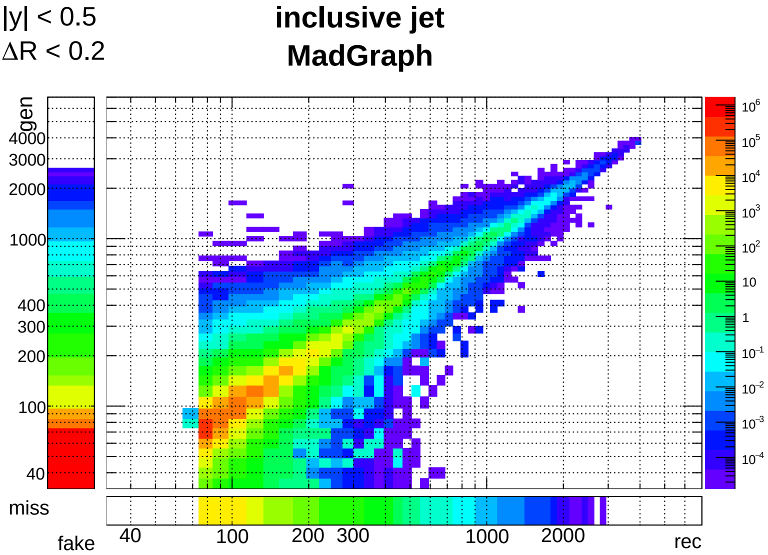
- $\Delta R = 0.2, 0.3, 0.4$
- no variation seen in ABPS
- in the RMs, only slight changes in the tails
- CUETP8M1 and MadGraph show the same
- plots are shown for MadGraph only
- in the next tests, 0.3 only will be considered

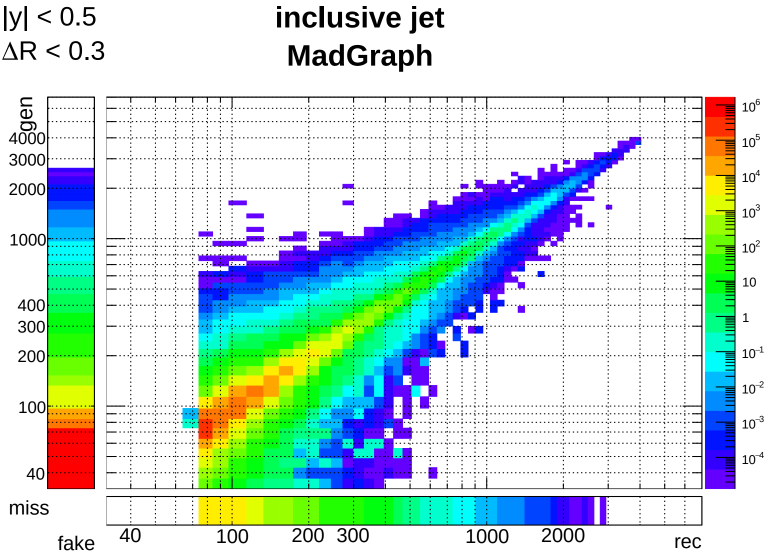
$|y| < 0.5$

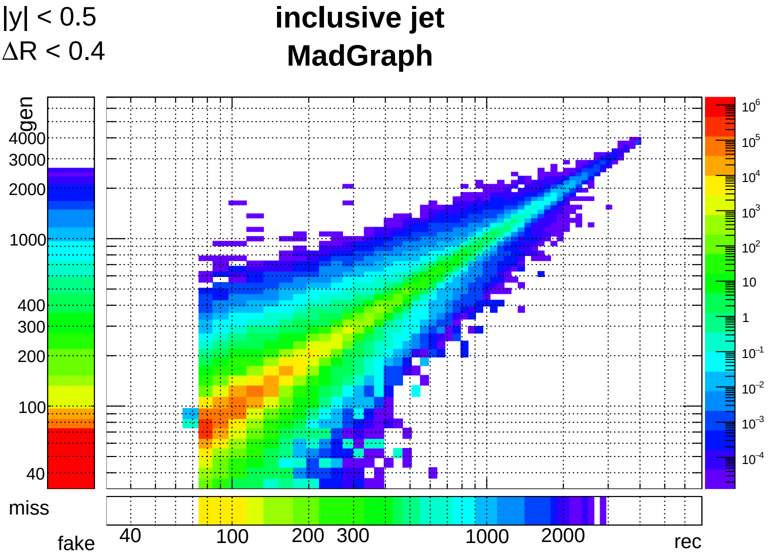
jet reconstruction

$\Delta R < 0.4$
$\Delta R < 0.3$
$\Delta R < 0.2$




|y| < 0.5
ΔR < 0.2
inclusive jet
MadGraph
gen
4000
3000
2000
1000
400
300
200
100
40
miss
fake
40
100
200
300
1000
2000
rec
10^6
10^5
10^4
10^3
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}


|y| < 0.5
ΔR < 0.3
inclusive jet
MadGraph
gen
4000
3000
2000
1000
400
300
200
100
40
miss
fake
40
100
200
300
1000
2000
rec
10^6
10^5
10^4
10^3
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}

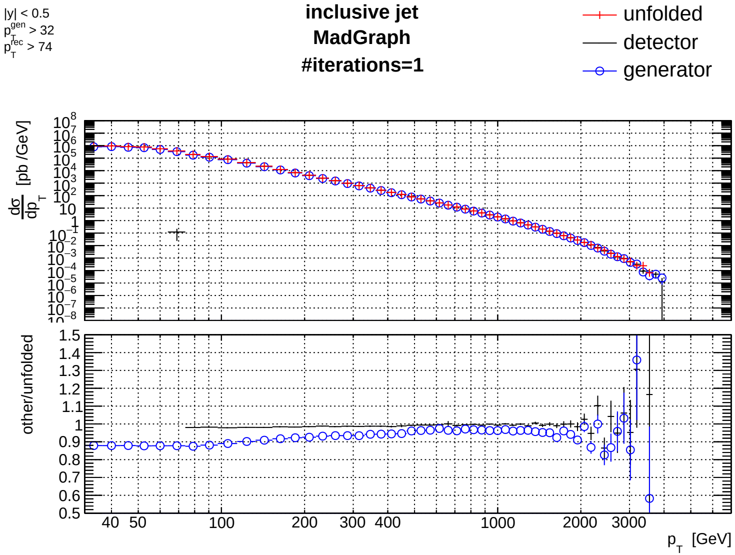

|y| < 0.5
ΔR < 0.4
inclusive jet
MadGraph
gen
4000
3000
2000
1000
400
300
200
100
40
miss
fake
40
100
200
300
1000
2000
rec
10^6
10^5
10^4
10^3
10^2
10
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}

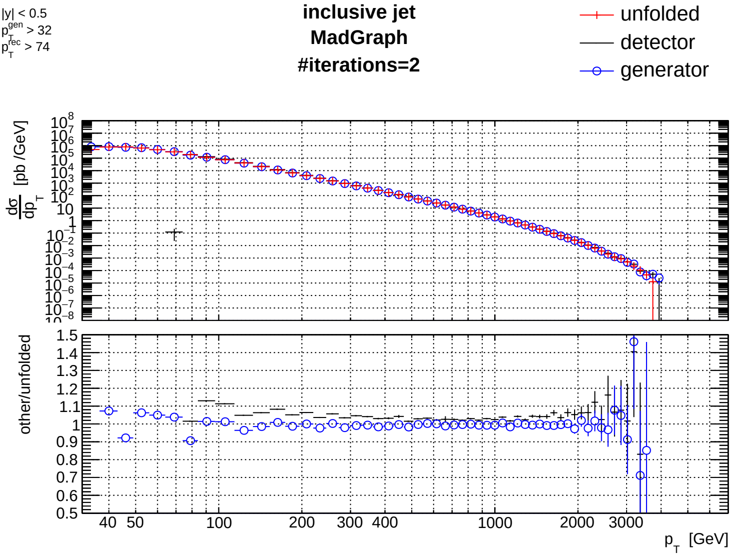
Closure test

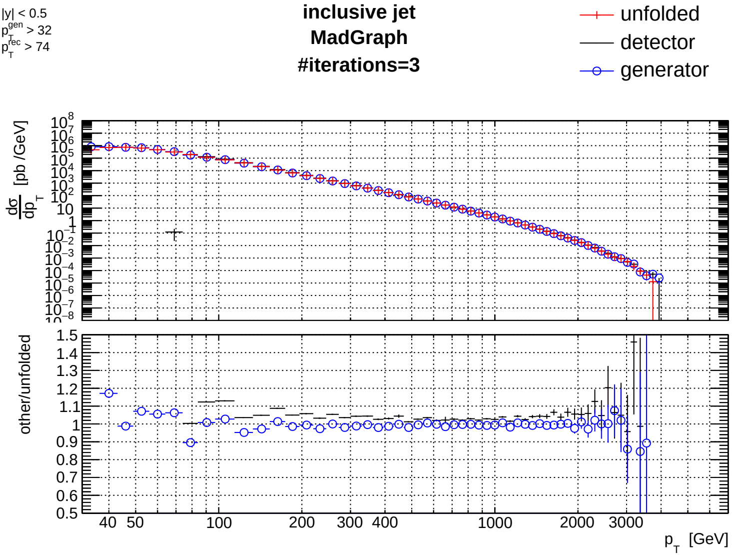
- Samples are divided in two equivalent parts: *odd* and *even*
- MadGraph-*odd* is unfolded with MadGraph-*even* or similar for CUETP8M1
- Unfolding is done either with RooUnfoldBayes or TUnfold
- MadGraph works better for both algorithms, though not 100% agreeing...
- CUETP8M1 gives strange results for both algorithms (sth related to the uncertainties... will be fixed later with the toy)

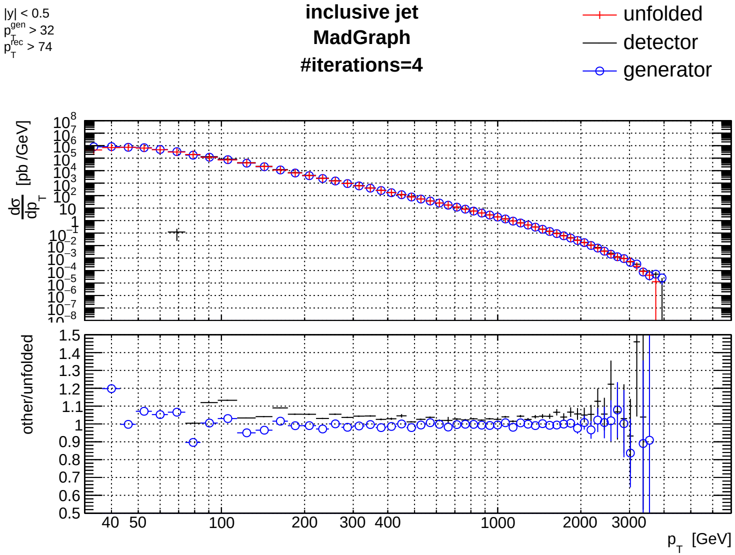
RooUnfoldBayes

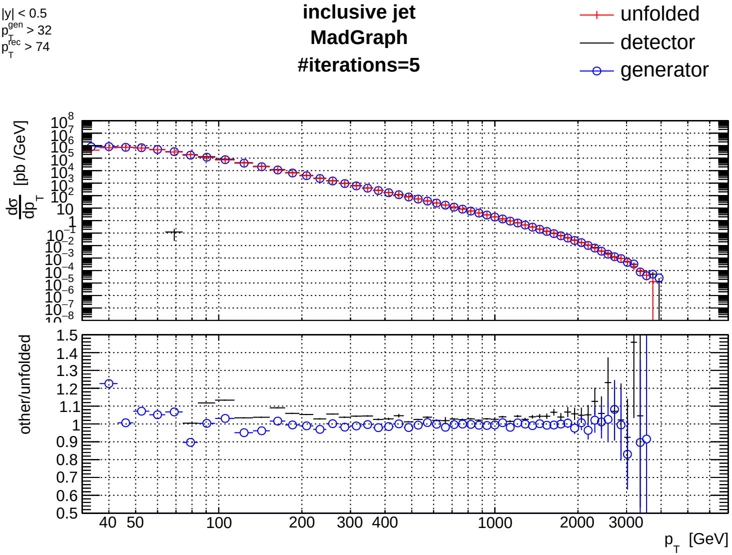
- MadGraph is given for each iteration
- transverse momentum in first rapidity bin
- at high values, unfolding work well, but at low value, spectrum is fluctuating along iterations

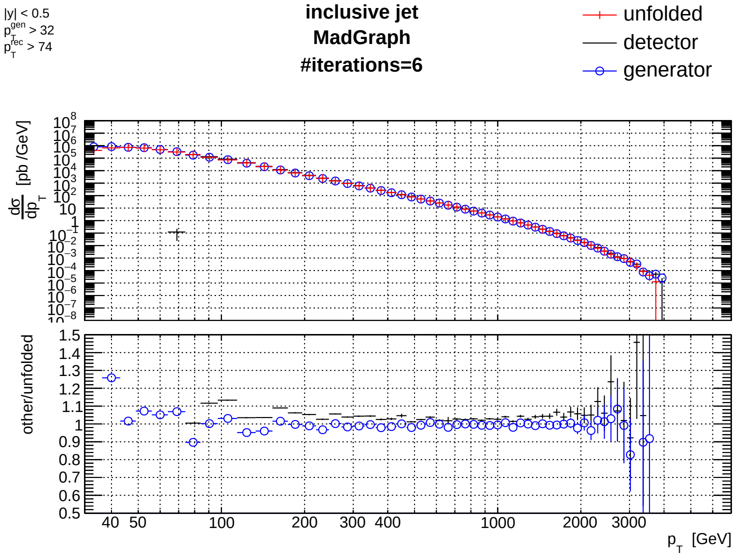

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=1
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

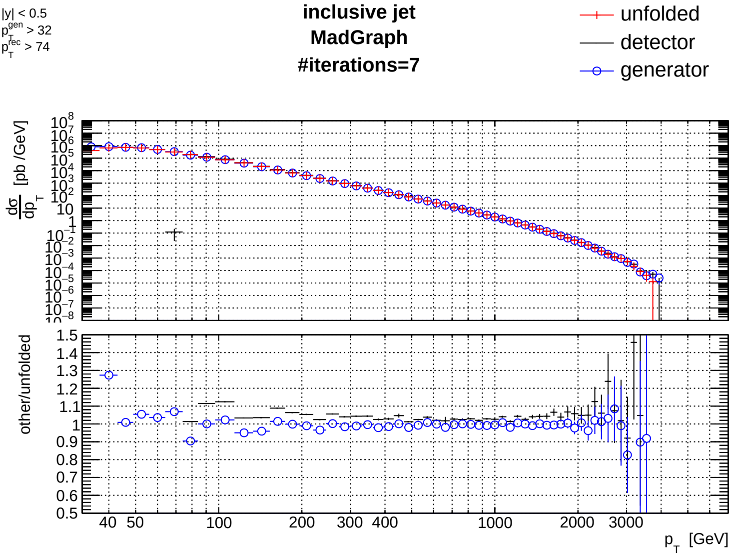

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=2
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

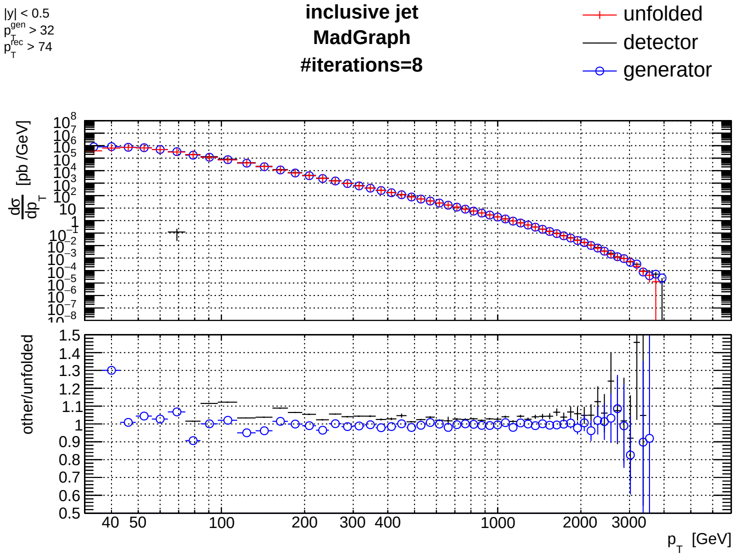

|y| < 0.5
p_T^{gen} > 32
p_T^{rec} > 74
inclusive jet
MadGraph
#iterations=3
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

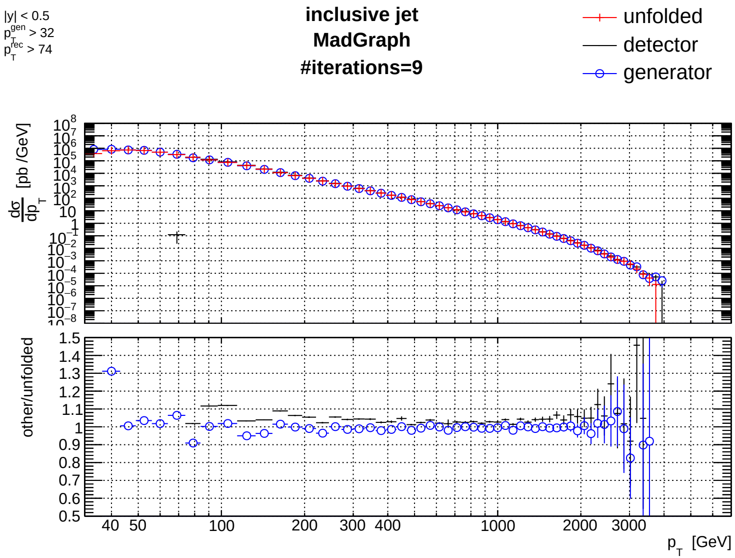

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=4
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

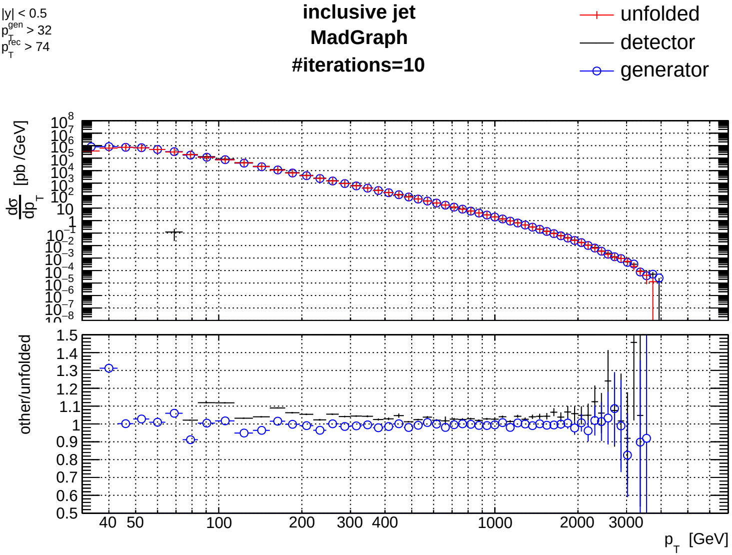

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=5
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

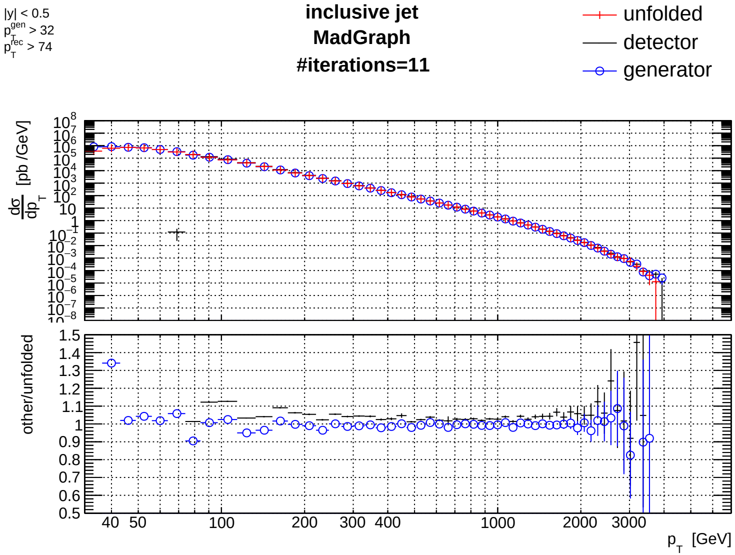

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=6
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]
40 50 100 200 300 400 1000 2000 3000

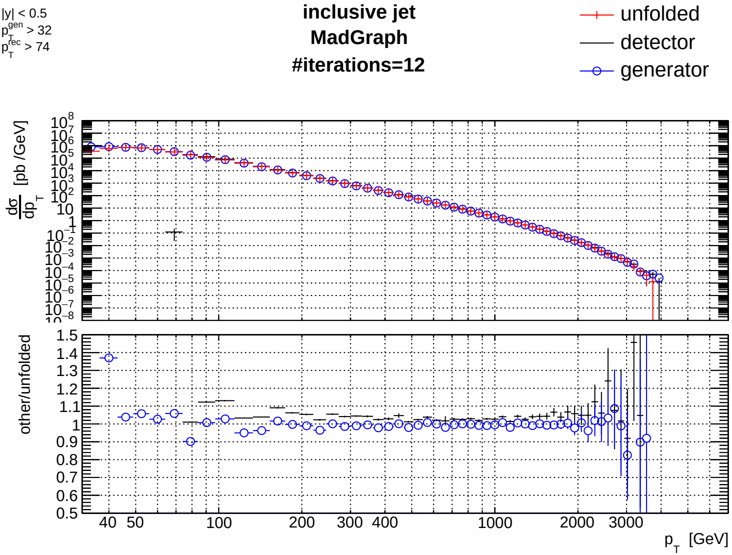

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=7
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

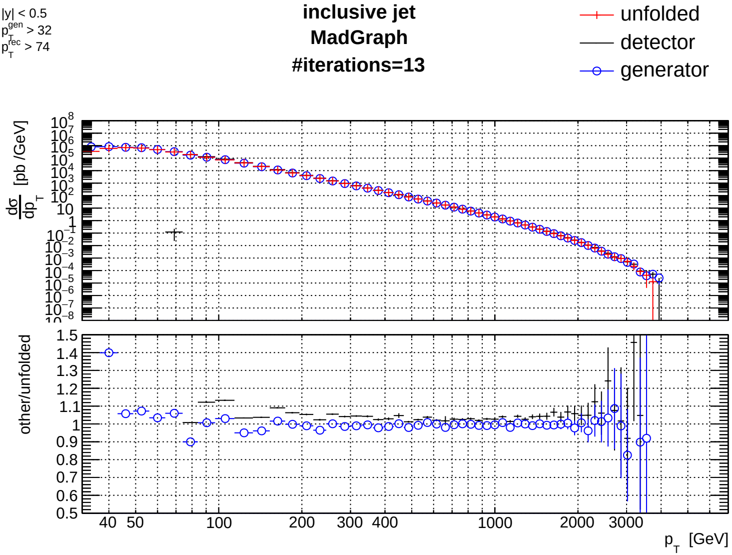

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=8
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

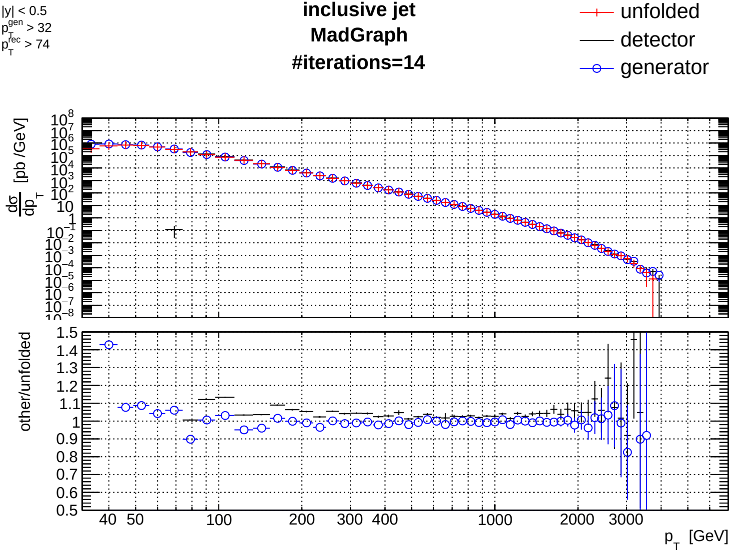

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=9
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

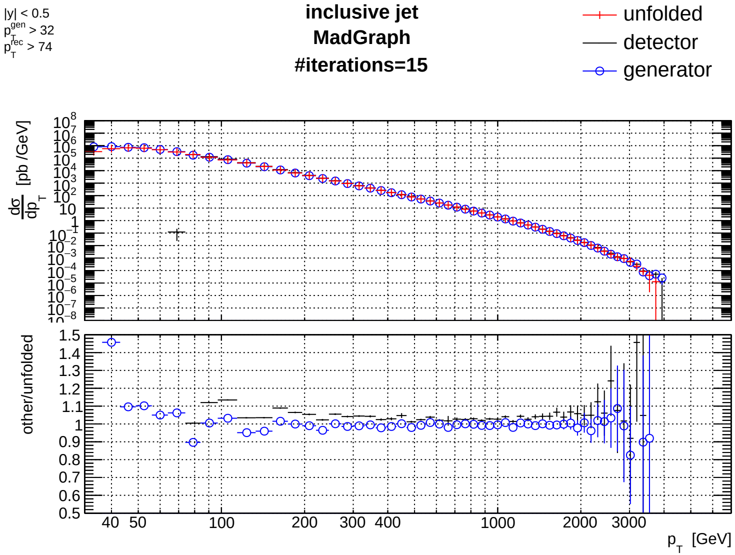

|y| < 0.5
p_T^{gen} > 32
p_T^{rec} > 74
inclusive jet
MadGraph
#iterations=10
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

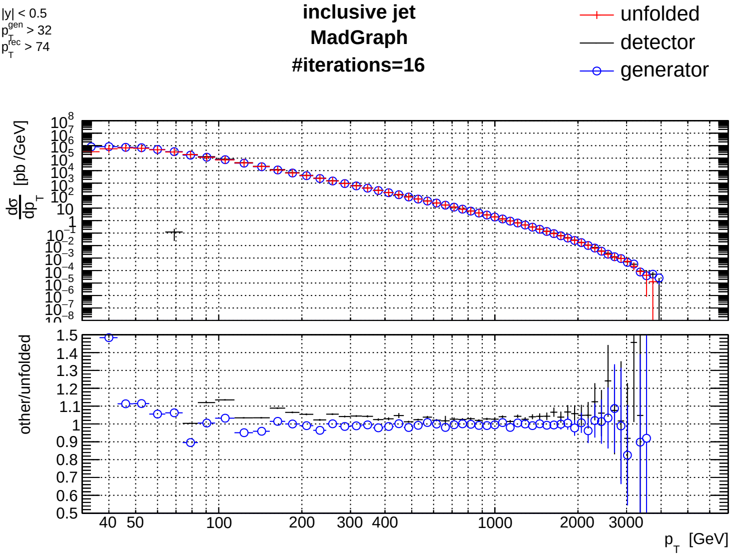

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=11
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]
40 50 100 200 300 400 1000 2000 3000

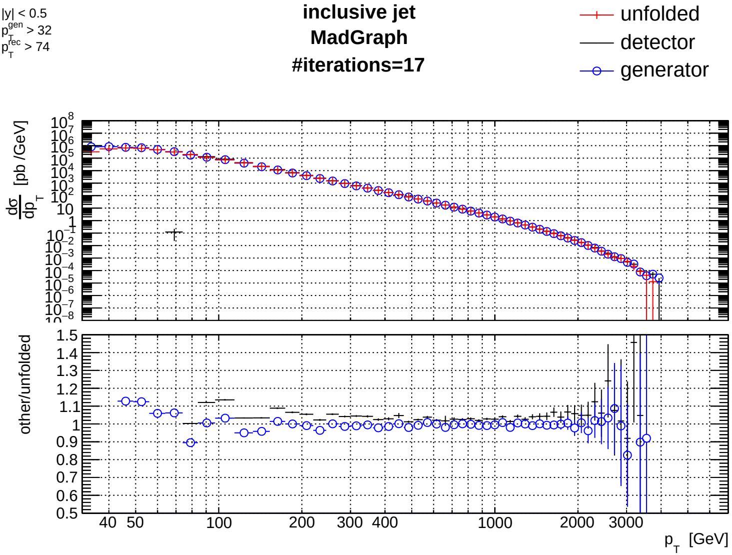

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=12
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

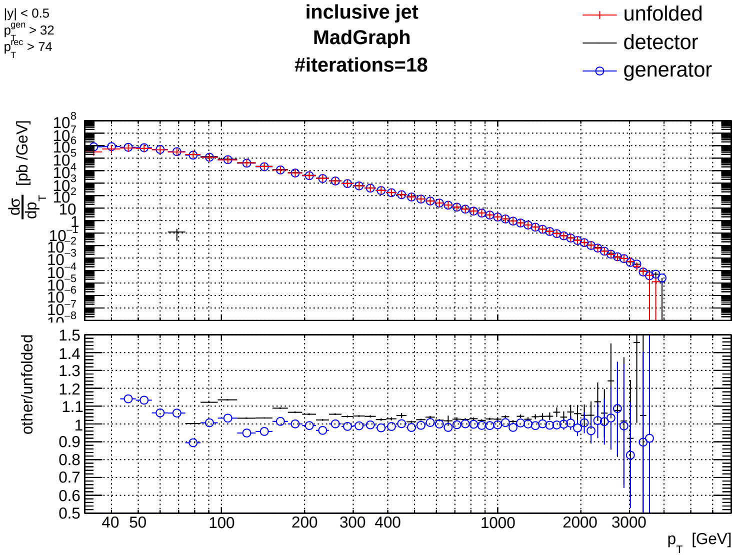

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=13
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]
40 50 100 200 300 400 1000 2000 3000

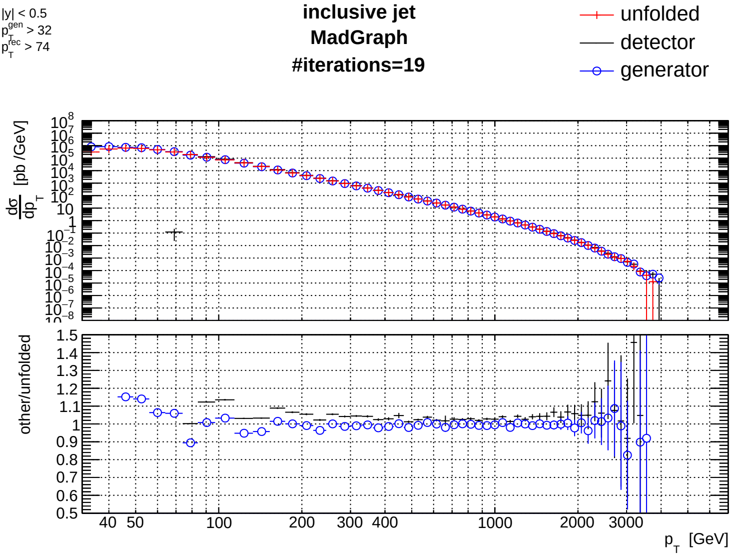

|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=14
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=15
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=16
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

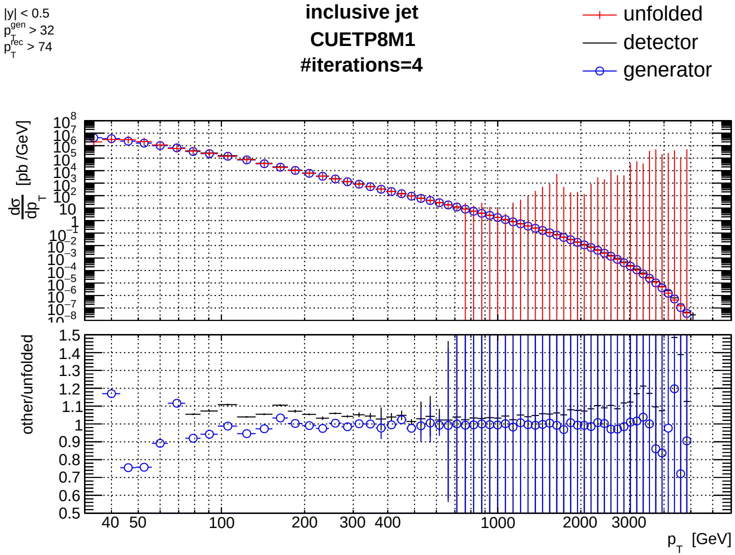

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
MadGraph
#iterations=17
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=18
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
p_T^gen > 32
p_T^rec > 74
inclusive jet
MadGraph
#iterations=19
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

CUETP8M1

- central values are consistent with MadGraph's
- however uncertainties at high pt are too large (but this will be cured with the toy MC)
- only the 4-iteration case is shown

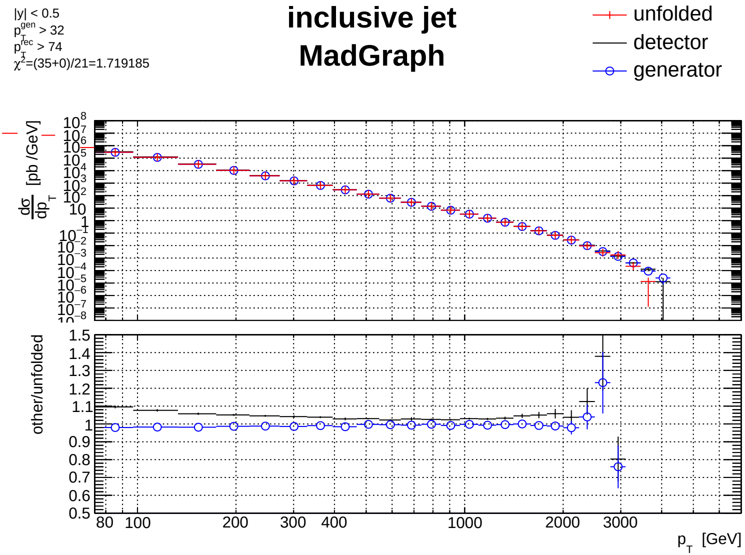

|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
inclusive jet
CUETP8M1
#iterations=4
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

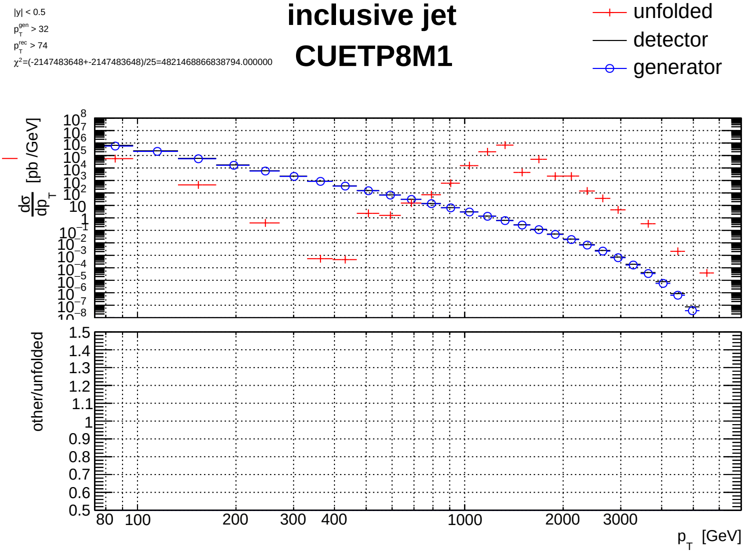
TUnfold

- MadGraph works out-of-the-box
- but CUETP8M1 fails (pretty sure it is also related to large uncertainties)

Unfolding with CUETP8M1 and MadGraph

Patrick Connor

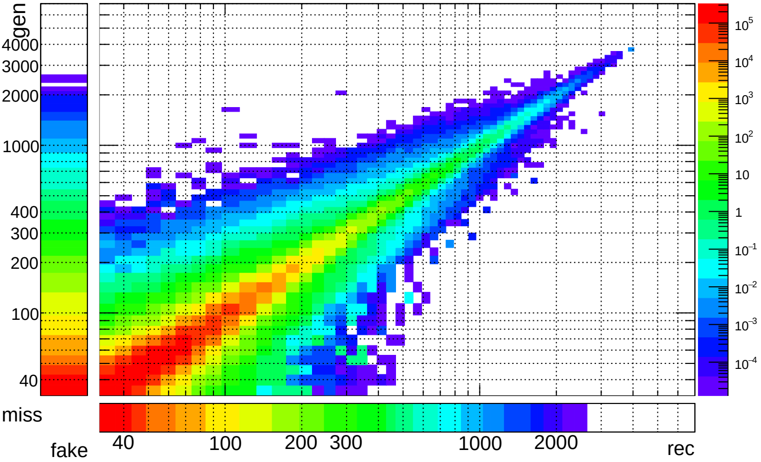



|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 74$
χ^2=(-2147483648+-2147483648)/25=4821468866838794.000000
inclusive jet
CUETP8M1
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

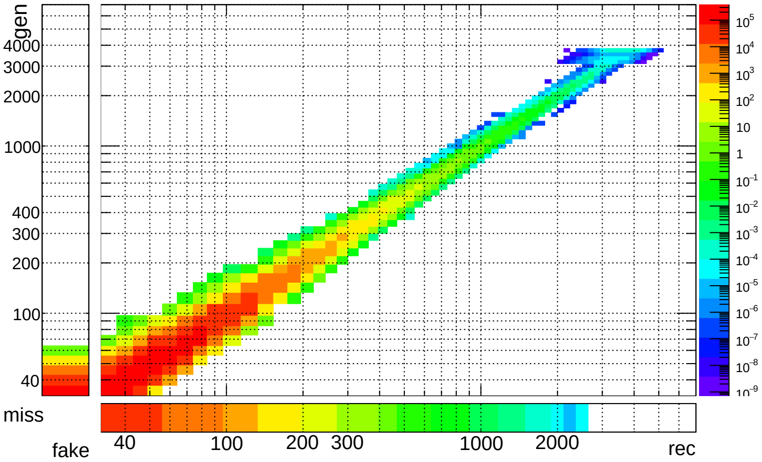
Tests

- still doing closure test with distinct samples
- Gaussian fit is used for test (but this neglects the deviations to lower reconstructed values)
- fake-subtraction is included in the procedure
- able to recover unfolding with CUETP8M1 but Gaussian fit is to sufficient for a proper unfolding
- note: this regularisation method is normally not necessary for TUnfold (which includes its own regularisation method) but I still give it a try

inclusive jet original

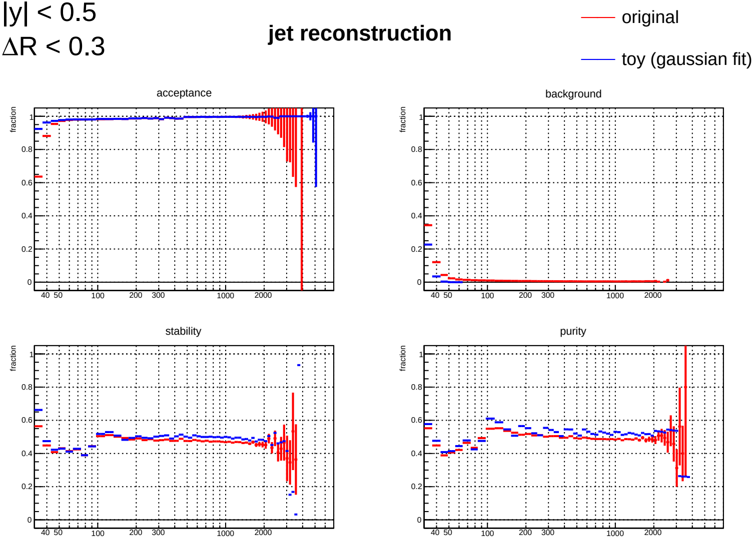


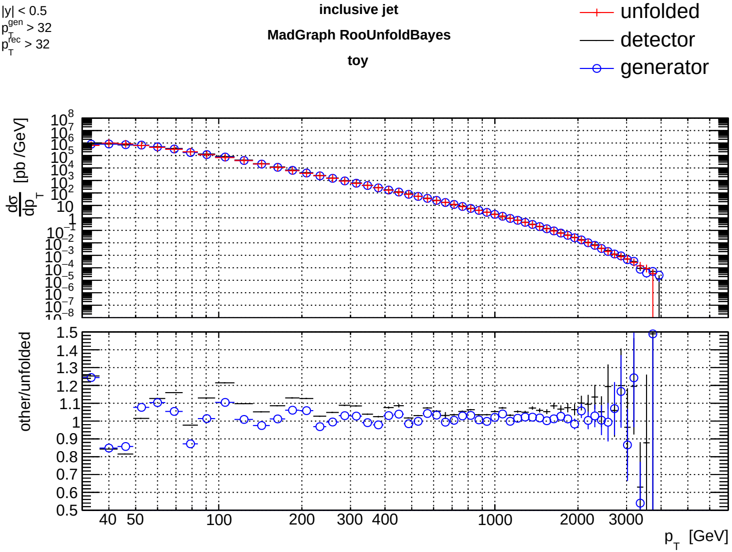
inclusive jet
toy (gaussian fit)

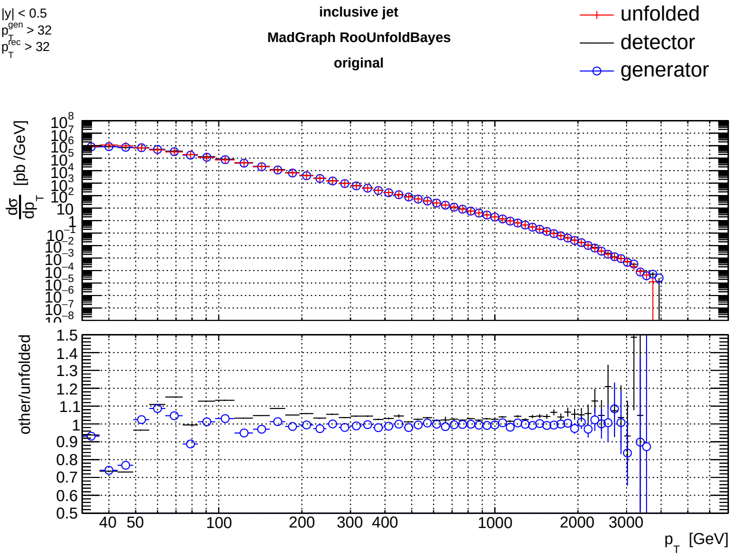


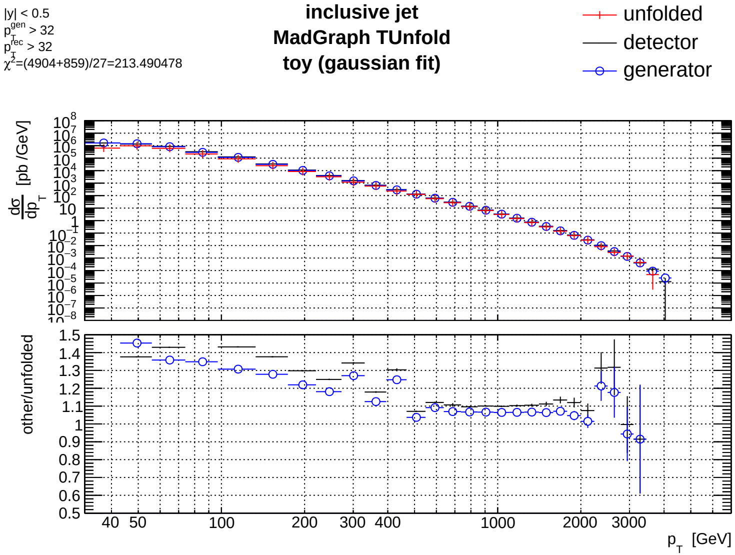
$|y| < 0.5$
$\Delta R < 0.3$

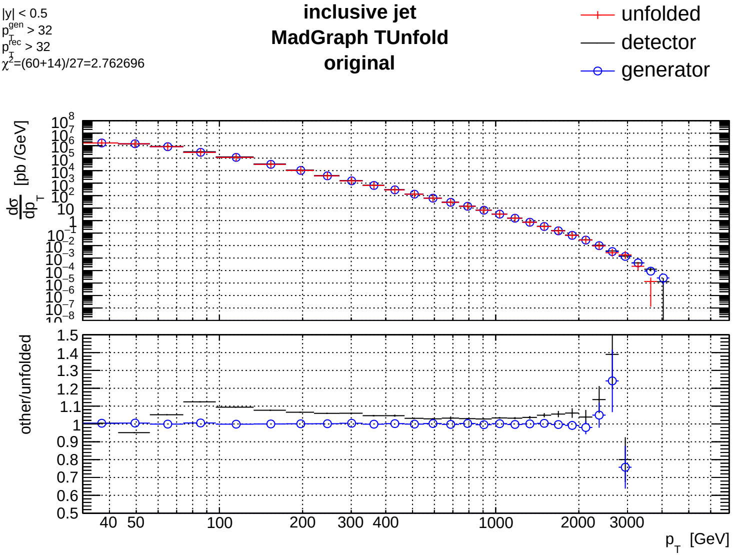
jet reconstruction



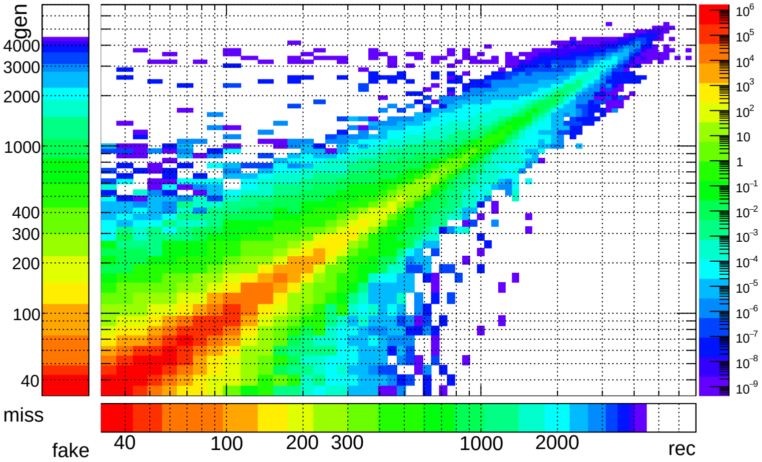

|y| < 0.5
p_T^gen > 32
p_T^rec > 32
inclusive jet
MadGraph RooUnfoldBayes
toy
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

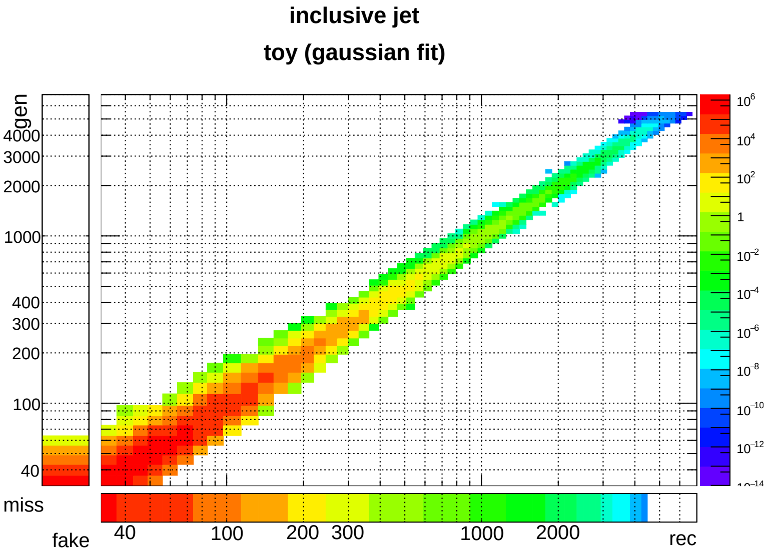

|y| < 0.5
p_T^gen > 32
p_T^rec > 32
inclusive jet
MadGraph RooUnfoldBayes
original
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 32$
χ^2=(4904+859)/27=213.490478
inclusive jet
MadGraph TUnfold
toy (gaussian fit)
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
p_T^gen > 32
p_T^rec > 32
χ²=(60+14)/27=2.762696
inclusive jet
MadGraph TUnfold
original
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

inclusive jet original



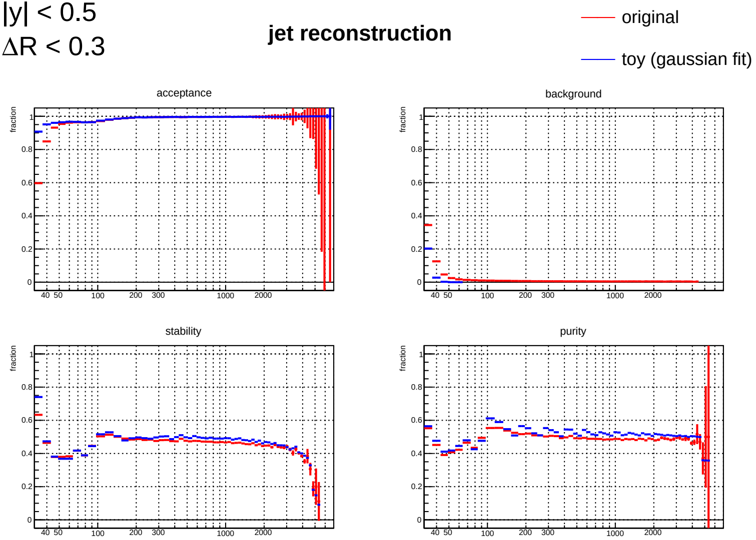

inclusive jet
toy (gaussian fit)
gen
4000
3000
2000
1000
400
300
200
100
40
miss
fake
40
100
200
300
1000
2000
rec
10^{6}
10^{4}
10^{2}
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}

Unfolding with CUETP8M1 and MadGraph

Patrick Connor

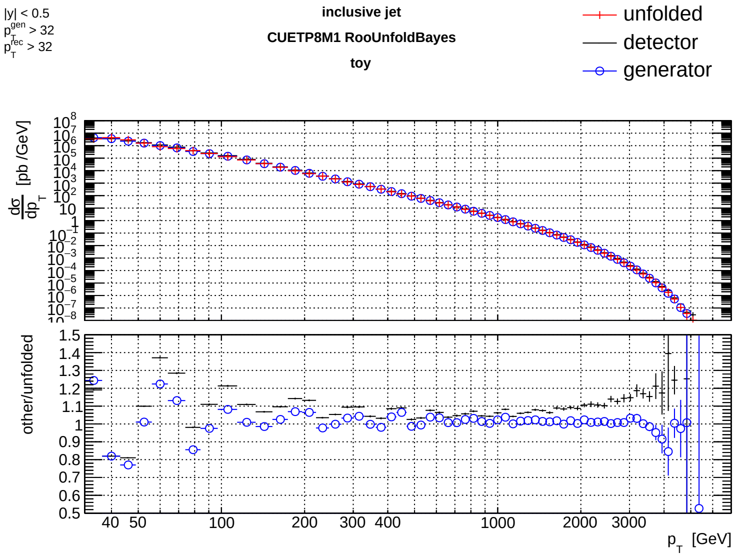
$|y| < 0.5$
$\Delta R < 0.3$

jet reconstruction



Unfolding with CUETP8M1 and MadGraph

Patrick Connor



Unfolding with CUETP8M1 and MadGraph

Patrick Connor

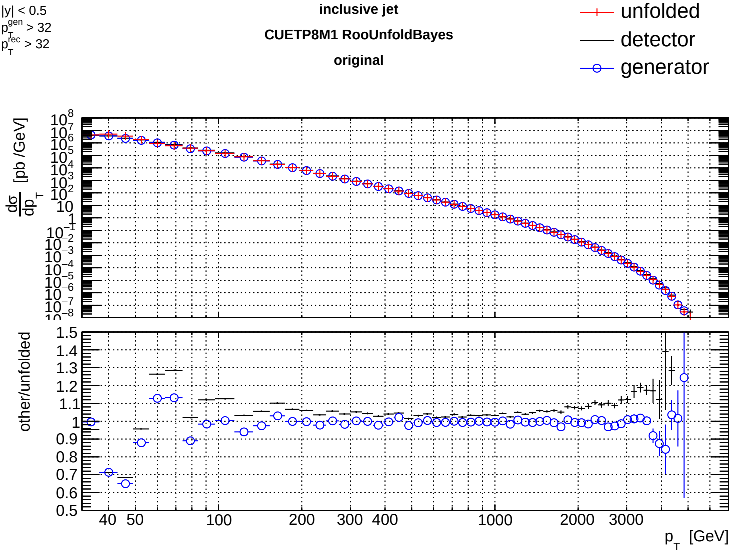
Introduction

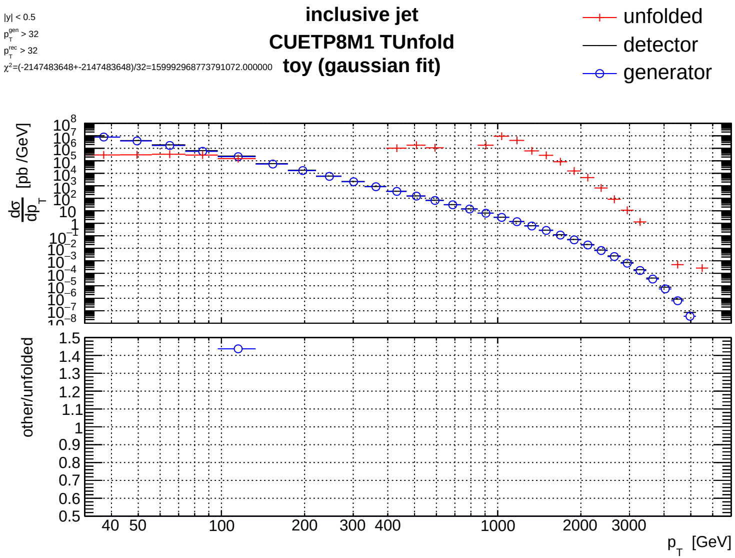
Tests
- Matching
- MadGraph
- Closure test
  - RooUnfoldBayes
  - TUnfold
- Toy MC
  - MadGraph
  - **CUETP8M1**

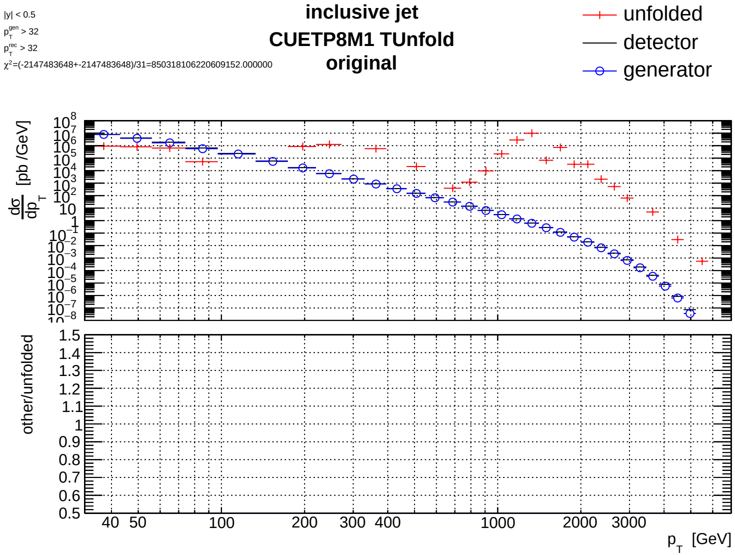
Application
- MC
- Data

Summary and early conclusions

Back-up




|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 32$
χ^2=(-2147483648+-2147483648)/32=159992968773791072.000000
inclusive jet
CUETP8M1 TUnfold
toy (gaussian fit)
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]



|y| < 0.5
$p_T^{gen} > 32$
$p_T^{rec} > 32$
χ^2=(-2147483648+-2147483648)/31=850318106220609152.000000
inclusive jet
CUETP8M1 TUnfold
original
unfolded
detector
generator
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]

Application

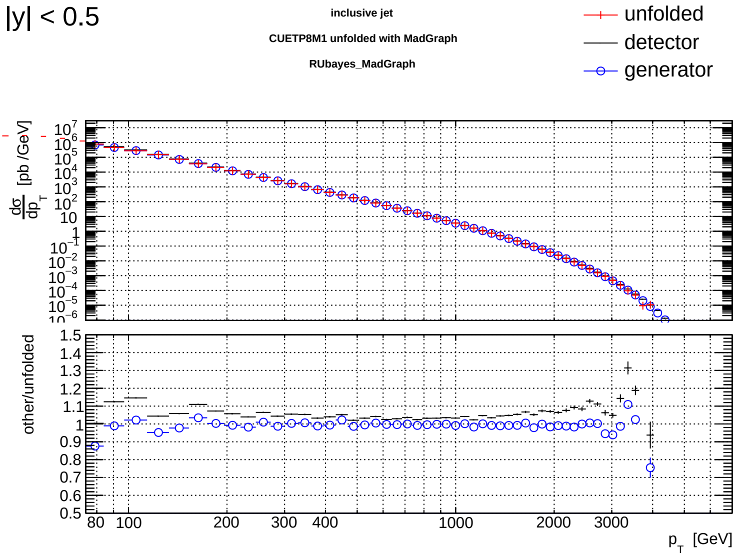
MC

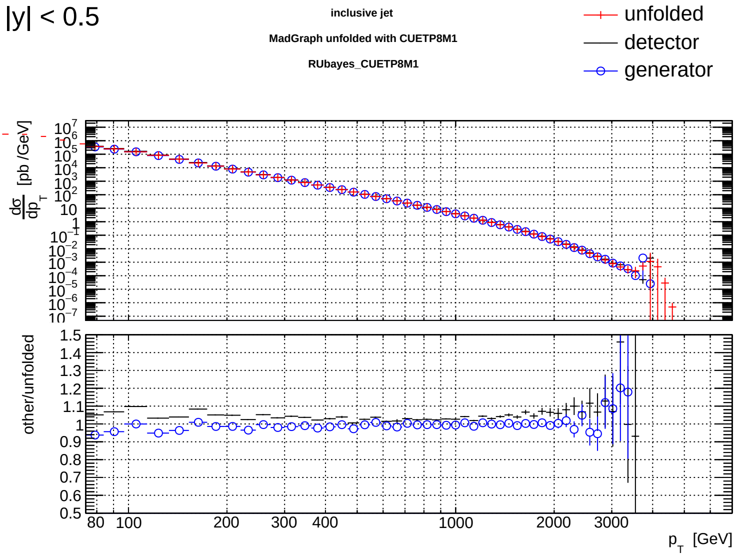
- unfolding CUETP8M1 with MadGraph or inversely
- trying both RooUnfoldBayes and TUnfold
- toy MC has not been tried yet

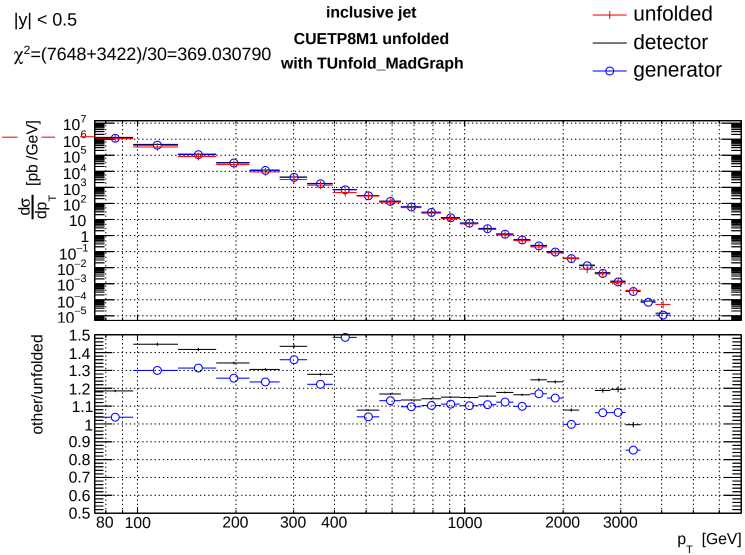
data

- unfolding data in the for cases

No fine tuning yet! Preliminary results!

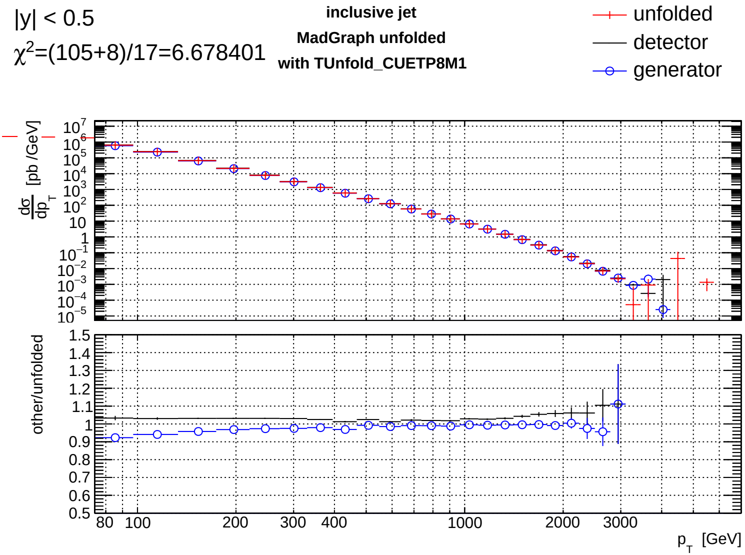

|y| < 0.5
inclusive jet
CUETP8M1 unfolded with MadGraph
RUbayes_MadGraph
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
inclusive jet
MadGraph unfolded with CUETP8M1
RUbayes_CUETP8M1
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
χ²=(7648+3422)/30=369.030790
inclusive jet
CUETP8M1 unfolded
with TUnfold_MadGraph
unfolded
detector
generator
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

Unfolding with CUETP8M1 and MadGraph

Patrick Connor



Unfolding with CUETP8M1 and MadGraph

Patrick Connor

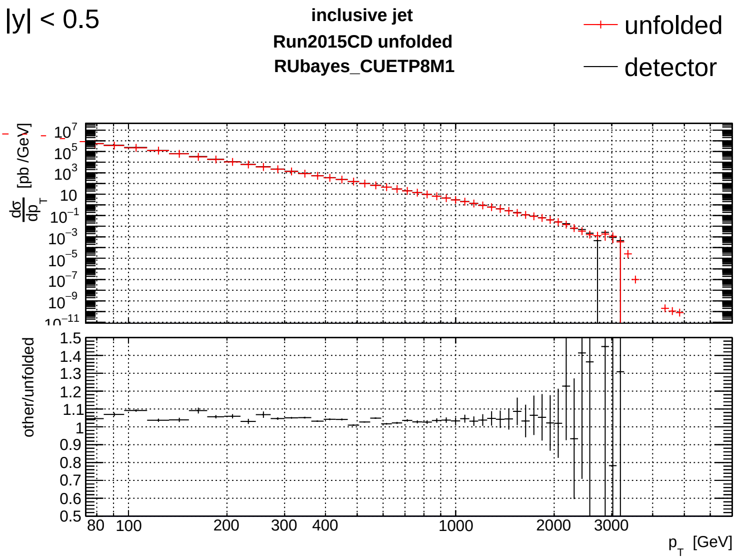
Introduction

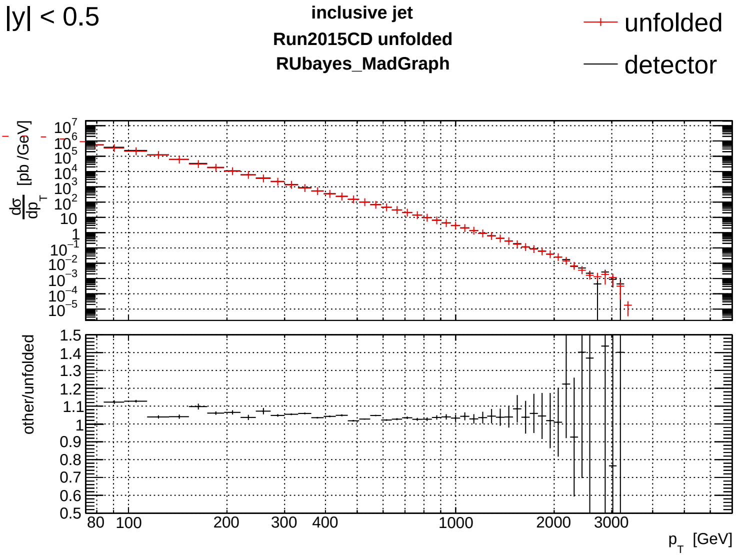
Tests
- Matching
- MadGraph
- Closure test
  - RooUnfoldBayes
  - TUnfold
- Toy MC
  - MadGraph
  - CUETP8M1

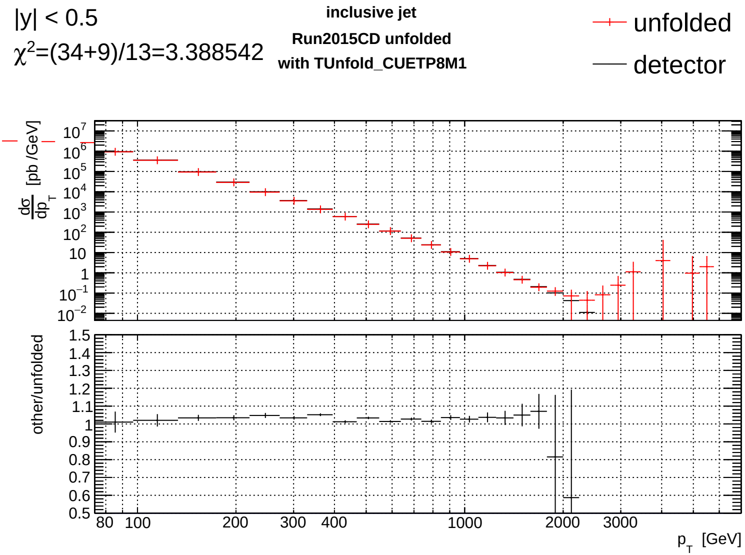
Application
- MC
- **Data**

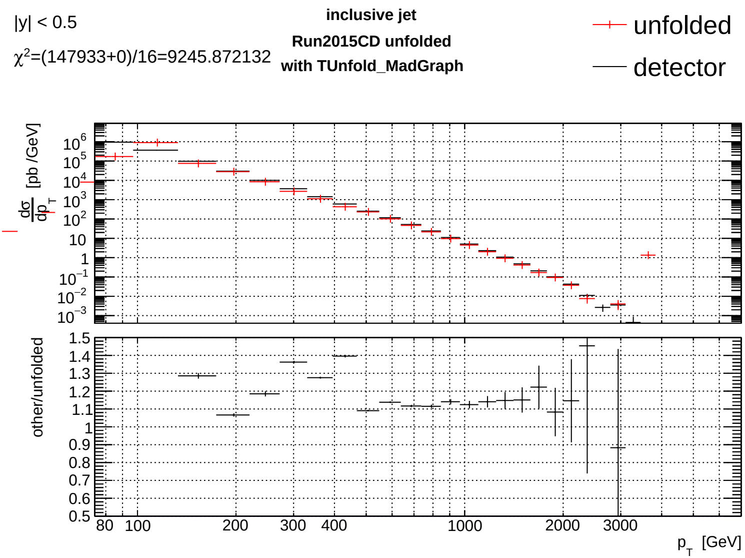
Summary and early conclusions

Back-up




|y| < 0.5
inclusive jet
Run2015CD unfolded
RUbayes_MadGraph
unfolded
detector
$\frac{d\sigma}{dp_T}$ [pb /GeV]
other/unfolded
p_T [GeV]


|y| < 0.5
χ²=(34+9)/13=3.388542
inclusive jet
Run2015CD unfolded
with TUnfold_CUETP8M1
unfolded
detector
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

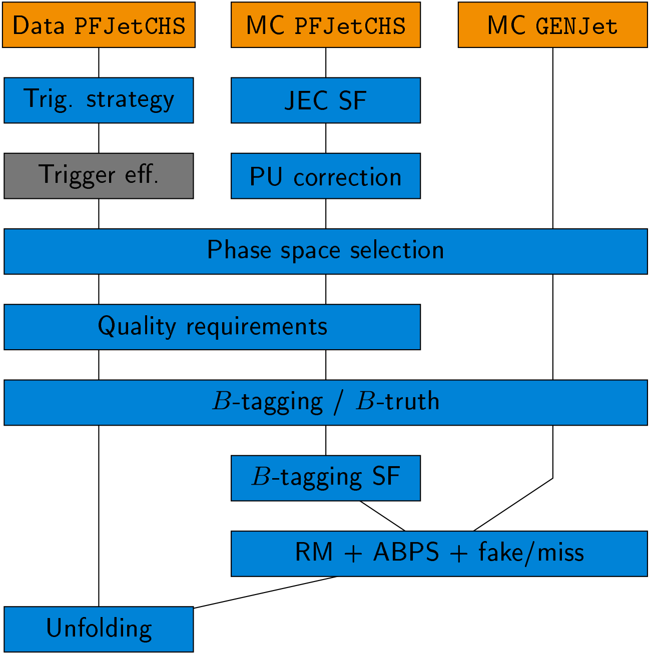

|y| < 0.5
χ²=(147933+0)/16=9245.872132
inclusive jet
Run2015CD unfolded
with TUnfold_MadGraph
unfolded
detector
dσ/dp_T [pb /GeV]
other/unfolded
p_T [GeV]

Summary and early conclusions

- still some tests to perform
- but soon we could start the AN/PAS

Back-up

- Ntuple processing
- Previous analysis
- b tagging at CMS
- Trigger strategy
- An event display

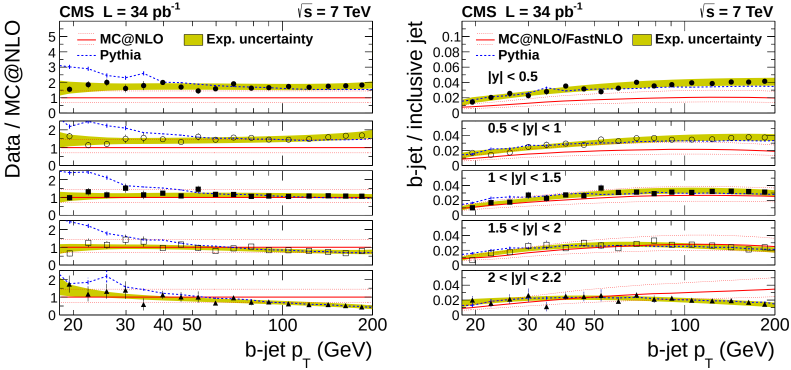


from blue (done) to gray (not started)

Previous analysis

Measurement of the **double differential cross section of b-inclusive production in transverse momentum and rapidity** at $13\,\mathrm{TeV}$

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_\perp\,\mathrm{d}y}(pp \to bX)$$



Datasets and phase space

Samples

data Run2015CD (full luminosity: $\int \mathcal{L}\,\mathrm{d}t = 2.26\,\mathrm{fb}^{-1}$)

MC CUETP8M1[1] and MadGraph

Selection

- $p_\perp^{\text{jet}} > 74\,\mathrm{GeV}$ (only at detector level)
- MET < 0.3 (id.)
- $|y^{\text{jet}}| < 4.7$
- anti-$k_\perp$ with $R = 0.4$

In the analysis, the $p_\perp$ spectrum is presented in 5 rapidity bins with following edges: $(0; 0,5; 1.0; 1,5; 2,0; 2,4)$

[1] except the slice from 15 to 30 GeV, where the PU is clearly overestimated

b tagging at CMS I

Jet Probability

TWiki[2]: "This is a more sophisticated algorithm, also exploiting the long lifetime of B hadrons. Its b tag discriminator is equal to the negative logarithm of the confidence level that all the tracks in the jet are consistent with originating from the primary vertex. This confidence level is calculated from the signed impact parameter significances of all good tracks. It reads the resolution function on these from a database (DB). Indeed, we have two versions of this tagger: JetProbabilityBJetTags and JetBProbabilityBJetTags - the latter uses only the four most displaced tracks, matching the typical reconstructed multiplicity of a B decay vertex."

b tagging at CMS II

Combined Secondary Vertex

B-tagging variable resulting of a MVA combining

1. **Track-Counting**: reject secondary vertices whose tracks are to close to the primary vertex
2. **Simple-Secondary-Vertex-Mass**: reject other meson candidates than B-mesons
3. **Soft-Lepton-Tag**: look for a non-isolated lepton in the jet

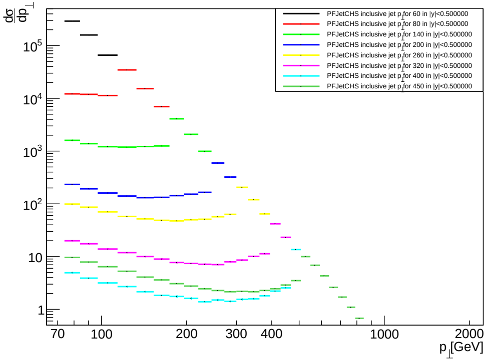
TWiki: " This sophisticated and complex tag exploits all known variables, which can distinguish b from non-b jets. Its goal is to provide optimal b tag performance, by combining information about impact parameter significance, the secondary vertex and jet kinematics. (Currently lepton information is not included). The variables are combined using a likelihood ratio technique to compute the b tag discriminator. A variant of this tagger combines the variables using the Multivariant Analysis (MVA) tool."

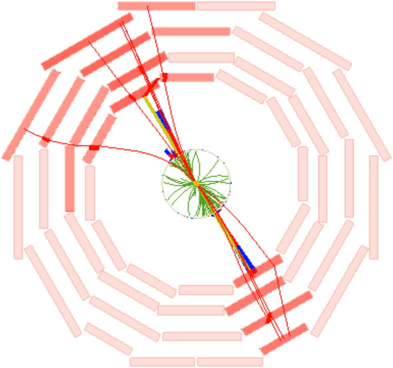
[2] b-Tagging Offline Guide as of 30 December 2015

Trigger strategy

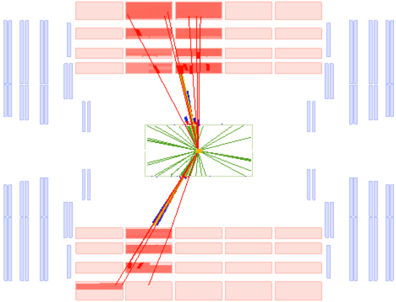
Turn-ons are defined such as to have 100% of efficiency in the central rapidity bin.

threshold	turn-on
60	74
80	97
140	174
200	272
260	300
320	362
400	468
450	507





Back-up



- Run 258706, LS 691, event 1173677191
- $p_T^{\text{lead}} = 2227.35\,\text{GeV}$
- remarkable: many muons
- a more systematic study would be of interest, but system often crashing (DAS issue?)