QCD background smearing function studies

CMS SUSY hadronic working group meeting

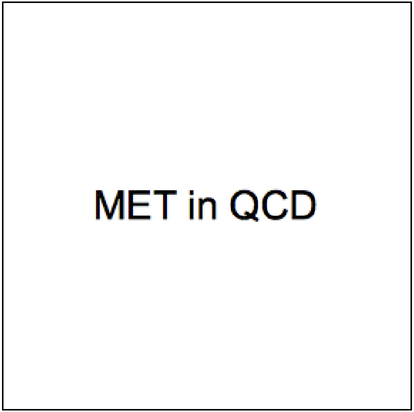
Matthias Schröder
long list of names

June 26th, 2009

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Motivation: Understanding the QCD background to $\not\!\!E_T$

- Missing transverse energy, $\not\!\!E_T$, is an important signature in the search for new physics
- Large $\not\!\!E_T$-background expected from QCD events in the all-hadronic channel:
  - Particles invisible to the calorimeter e.g. μ or ν
  - Mismeasurement due to intrinsic calorimeter resolution
  - Mismeasurement due to detector acceptance
  - ...



Understanding of QCD contribution to $\not\!\!E_T$ important

Outline

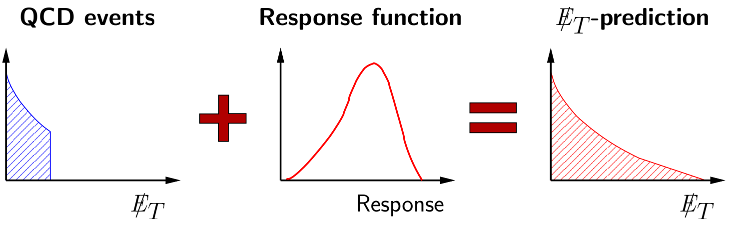
1.
   -
   -
   -
2.
   -
   -
3.
4.
5.
6.
7.

Different concepts to estimate the QCD $\not\!\!E_T$ contribution

- Estimation from MC simulation
- ABCD method
- Jet smearing method
- ...

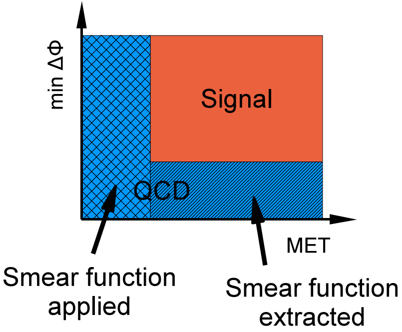
Concept of the jet smearing method

1. Selection of well measured QCD events
2. Smearing with resolution function $p_T^{\text{meas}}/p_T^{\text{true}}$



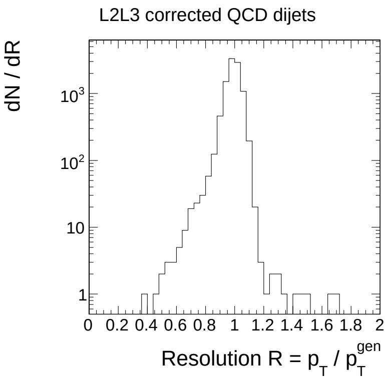
Challenges of the jet smearing method

- Application of resolution function to QCD events (UCSB)
  - Selection criteria
  - Acceptance effects
  - Normalization of smeared events
  - Effect of double smearing
- Determination of resolution function ⟶ this talk



Determination of the resolution function

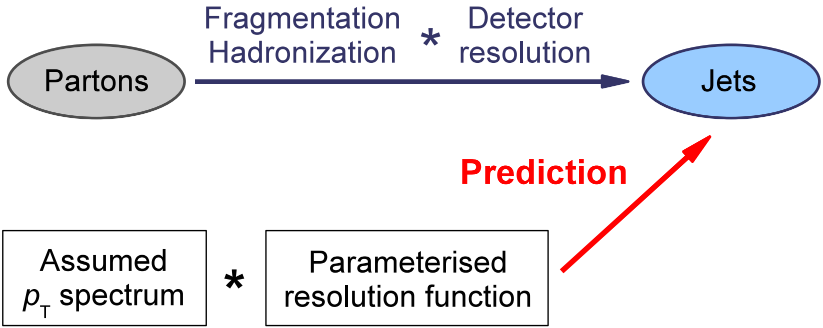
- True jet p_T from MC simulation
  - Depends completely on simulation
- γ-jet or Z-jet events
  - $p_T^{\gamma,Z}$ relatively well measured
  - Low statistics, jet-parton matching uncetrainty
- MET projection method
  - N-jets with $\not{E}_T$ parallel to one jet
  - Other jets assumed to be measured correctly
  - Might neglect mismeasurement of other jets



Resolution fit ⟶ this talk

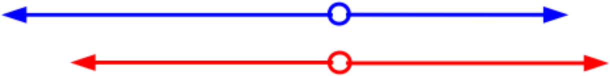
- N-jet events
- All jets assumed to be mismeasured by resolution function

Concept of the resolution fit method



(see talk by C. Sander)

Details for dijet events



- In each event i, probability for a given dijet configuration p_T^{j1}, p_T^{j2} is

$$\mathcal{P}^i_{1,2} = \frac{1}{N} \int_0^\infty \mathrm{d}p_T^{\text{true}}\, f(p_T^{\text{true}}) \cdot r_b(p_T^{j1}/p_T^{\text{true}}) \cdot r_b(p_T^{j2}/p_T^{\text{true}})$$

  - f is the probability density function (pdf) of p_T^{true}
  - r_b is the parameterized resolution function
  - N is the normaliztion
- Likelihood $\tilde{\mathcal{L}} = \prod_{i=0}^{N_{\text{evt}}} \mathcal{P}^i_{1,2}$ maximal for correct response function r_b
- Minimization w.r.t. b of negative log-likelihood function

$$\mathcal{L} = -\sum_{i=0}^{N_{\text{evt}}} \ln(\mathcal{P}^i_{1,2})$$

Possible paramtrizations of the resolution function

Analytic function

- Smooth behaviour
- Small number of parameters
- Normalization simple in some cases
- Functional form difficult to determine

Step function

- Normalization simple
- Describes any distribution if appropriately binned
- Many parameters
- Discontinuities

Spline

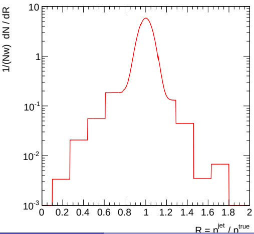
- Smooth behaviour
- Normalization difficult
- Technically complex

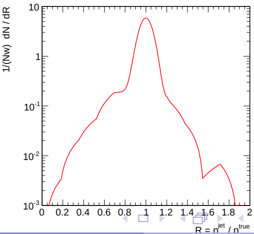
Details of the resolution parametrization

Superposition of Gaussian and interpolated step function

$$r_b(R) = c \cdot G(R; 1, \sigma) + (1 - c) \cdot S(R; b)$$

- Normalization c
- Central Gaussian G around 1 (assume calibrated jets)
- Step function S with $N = 8$ parameters to describe tails
  - Bin content $b_i \geq 0$ by construction — fit guided by penalty terms
  - Actual $S(R)$ is linear interpolation of adjacent bin contents
  - Only $N - 1$ parameters are fitted $\longrightarrow$ fixed scale





The fitting framework

- Extension of the *kalibri* tool from the University of Hamburg
- Originally developed for jet calibration via an **unbinned fit**
- Utilizes `LVMINI` by V. Blobel (reference)
  - Limited memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) algoritm
  - Support for $\mathcal{O}(10^4)$ parameters
- Highly configurable object-oriented framework
- Support for **different parametrizations** and data sets
- Support for **parameter limits**
- **Multiple threads** to exploit use of multi-core processors
- Automated production of **control plots**
- Validation via **Toy Monte Carlo**
  - Generation of jet 4-momenta according to specified p_T spectrum
  - Distribution of jet energy on several towers according to specified jet shape
  - Simulation of tower measurement according to specified resolution model

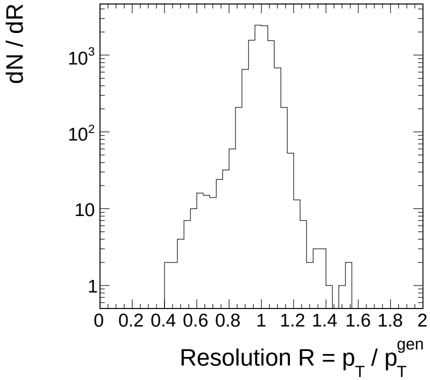
More information on *kalibri*

Input sample from ToyMC

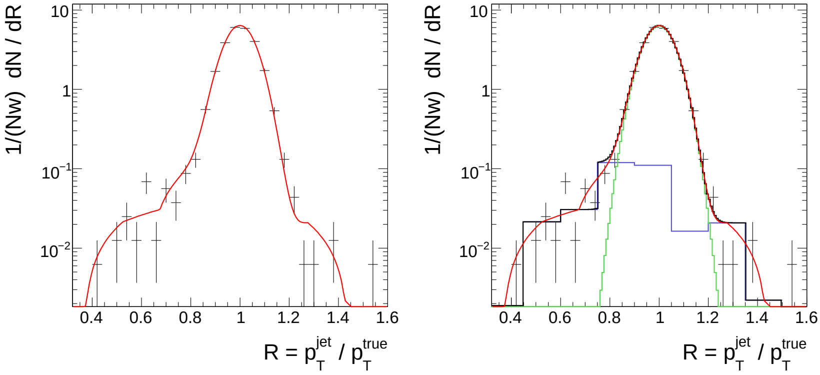
- Generation of dijet events
- Simulated resolution is sum of two Gaussians

  $$c \cdot G_0(1, \sigma_0) + (1 - c) \cdot G_1(\mu_1, \sigma_1)$$

  where
  - $c = 0.96$
  - $\sigma_0 = 0.06$
  - $\mu_1 = 0.9$
  - $\sigma_1 = 0.25$
- p_T^{true}-spectrum flat or falling



Results from Toy MC with flat spectrum

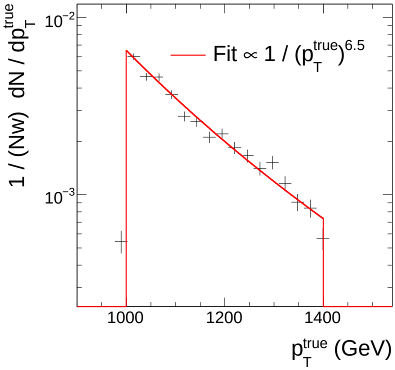


- Correct spectrum assumed during fit

Results from Toy MC with flat spectrum

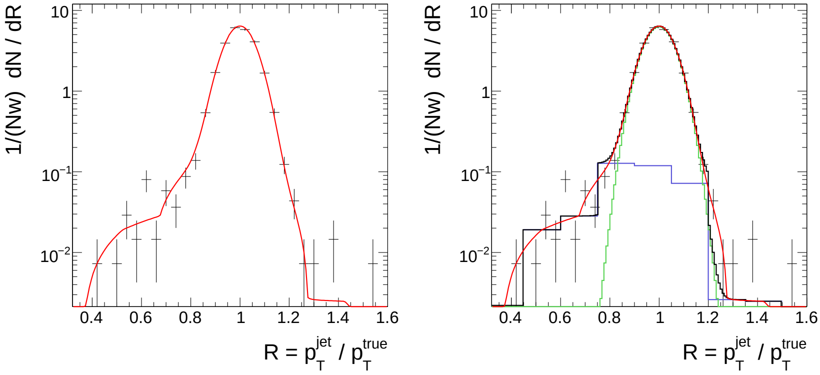
Parameter	Scale	Start value	Fitted value
c	1	0.96	0.951435
σ	0.1	0.6	0.602761
0	0.01	1	0.85526
1	0.1	1	0.970787
2	0.1	1	1.38628
3	0.5	1	1.08298
4	0.5	1	1
5	0.5	1	0.147644
6	0.1	1	0.941145
7	0.01	1	0.999452

Results from Toy MC with falling spectrum



- Correct spectrum assumed during fit

Results from Toy MC with falling spectrum



- Correct spectrum assumed during fit

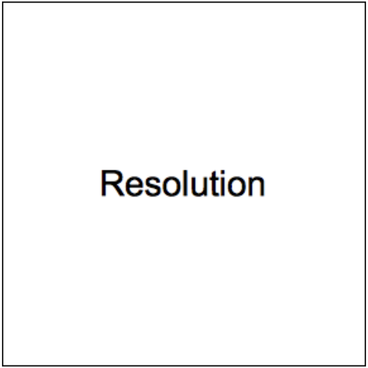
Results from Toy MC with falling spectrum

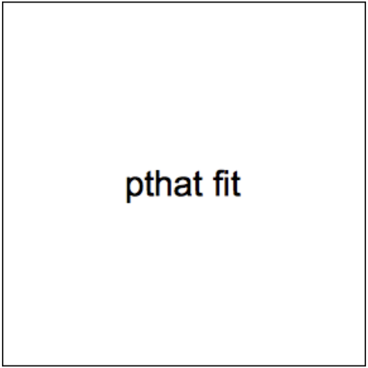
Parameter	Scale	Start value	Fitted value
c	1	0.96	0.944033
σ	0.1	0.6	0.599104
0	0.01	1	0.928335
1	0.1	1	0.806321
2	0.1	1	1.19273
3	0.5	1	1.07337
4	0.5	1	1
5	0.5	1	0.604311
6	0.1	1	0.10949
7	0.01	1	1.04362

Systematics

- Influence of variation of spectrum
- Influence of cuts on pt

Application on Summer08 QCDDiJets





- Event selection in example $\hat{p}_T$-bin
  - $x < \hat{p}_T < X$ GeV
  - Jet - GenJet matching criteria $\Delta R < 0.25$ (avoid bias, see the other talk)
  - No weights
- p_T^{true}-spectrum from fit on $\hat{p}_T$

Systematics

- Influence of variation of spectrum
- Influence of cuts on pt

Summary

Outlook

- Fit over whole pthat range with energy dependent function?
- Fit of spectrum?
- Resolution from fit to L2L3 corrected jets
- p_T-binning of resolution function: bias due to cut on measurement
- Splitting of resolution into b/c, calorimeter

Backup

Normalization

$$N = \int \int \int \mathrm{d}p_T^{\text{true}} \mathrm{d}p_T^{j1} \mathrm{d}p_T^{j1}\, f(p_T^{\text{true}}) \cdot r_b(p_T^{j1}/p_T^{\text{true}}) \cdot r_b(p_T^{j2}/p_T^{\text{true}})$$