The Matrix Element Measurement Method

Frank Fiedler
Johannes Gutenberg-Universität Mainz

Pizza Seminar
November 25, 2011

further reading:

- NIM A624:203-218 (2010)
  - on the matrix element method in general and its application to the top mass measurement
- arXiv:1003.0521 [hep-ex]
  - on top mass measurements, includes a comparison of measurement techniques

On a Foggy Day

Suppose...

- you arrive at a conference venue on a very foggy day.

Can you...

- decipher the signs?

On a Foggy Day

Suppose...

- you arrive at a conference venue on a very foggy day.

Can you...

- decipher the signs?

HOTEL

On a Foggy Day

Suppose...

- you arrive at a conference venue on a very foggy day.

Can you...

- decipher the signs?

On a Foggy Day

Suppose...

- you arrive at a conference venue on a very foggy day.

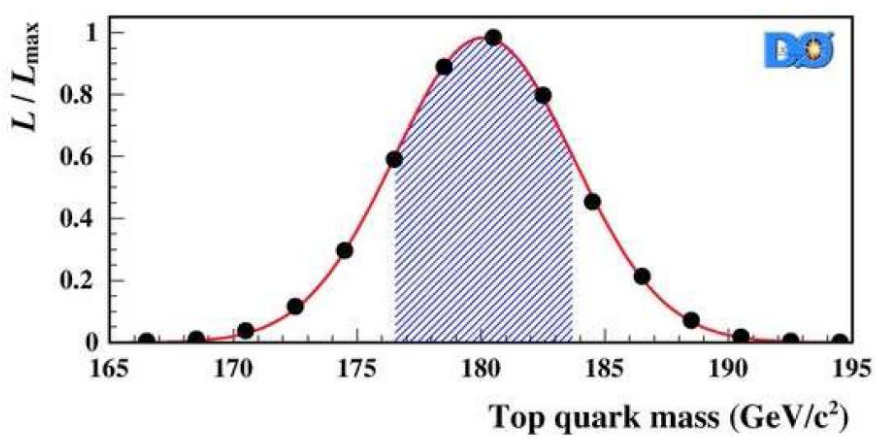
Can you...

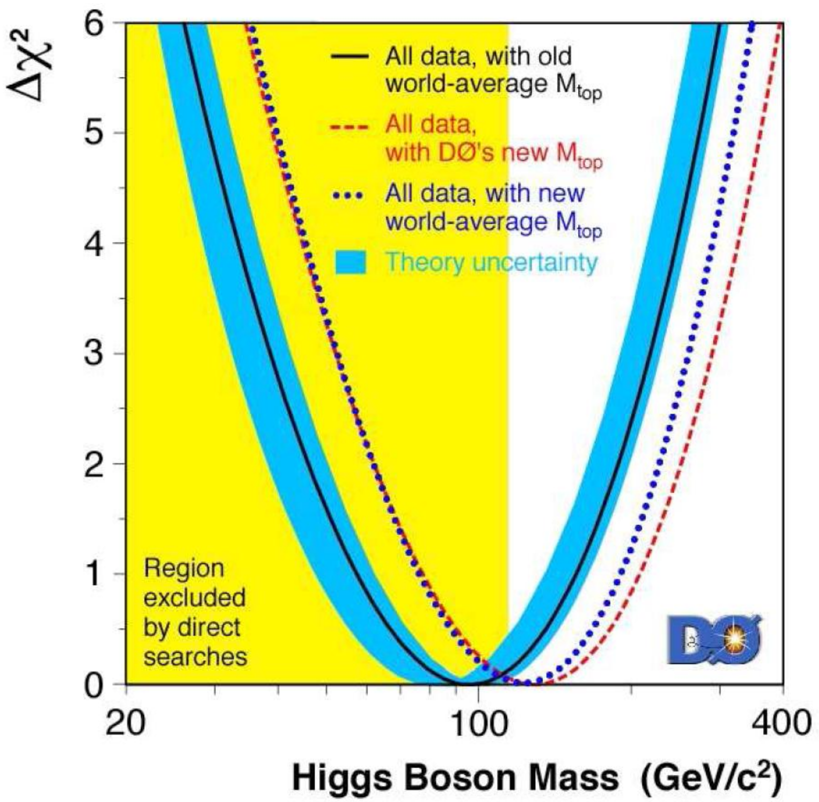
- decipher the signs?

The Top Quark Mass, 2004

Nature 429:638-642 (2004) / arXiv: hep-ex/0406031:

- “A precision measurement of the mass of the top quark”
- “The improvement in statistical uncertainty over our previous measurement [which was based on the same integrated luminosity] is equivalent to collecting a factor of 2.4 as much data.”





Overview

Prologue:

- on a foggy day
- measurement of the top quark mass in 2004

The method:

- how the matrix element measurement method works
- how this is different from 'normal' measurements

Applications of the method:

- measurement of the top quark mass, including technical issues
- calibration
- other measurements, very briefly

What do I mean by m_{top}?

- various different meanings of the same word

Wish list

Template Measurements

Imagine you want to measure some quantity, e.g. the mass m_X of a resonance X.

Typically:

- pick a final state

- trigger on it and select events
- in each selected event, ...
  - reconstruct final state particles
  - compute value of an ***estimator***
  - fill estimator into a histogram
- look at the histogram
- know that the detector isn't perfect

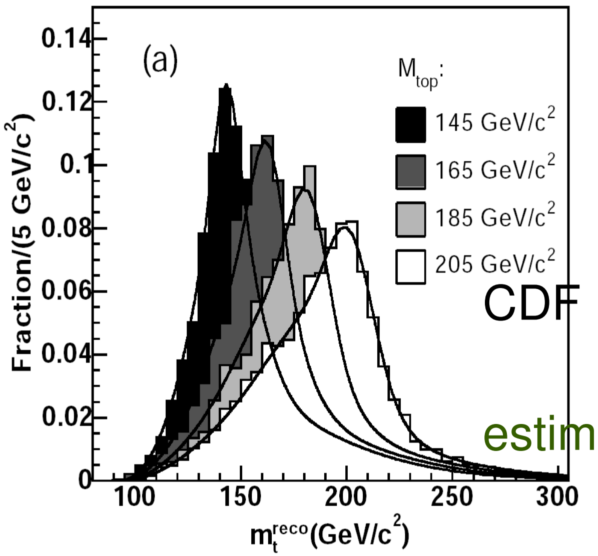
Template Measurements

Imagine you want to measure some quantity, e.g. the mass m_X of a resonance X.

Typically:

- pick a final state
- trigger on it and select events
- in each selected event, ...
  - reconstruct final state particles
  - compute value of an ***estimator***
  - fill estimator into a histogram
- look at the histogram
- know that the detector isn't perfect

- generate signal and background MC (signal for various assumed masses m_X)
- simulate trigger, select events
- in each selected event, ...
  - reconstruct final state particles
  - compute value of the estimator
  - fill estimator into another histogram: templates for various m_X



CDF measurement of m_{top}

estimator: m_{top}^{reco}

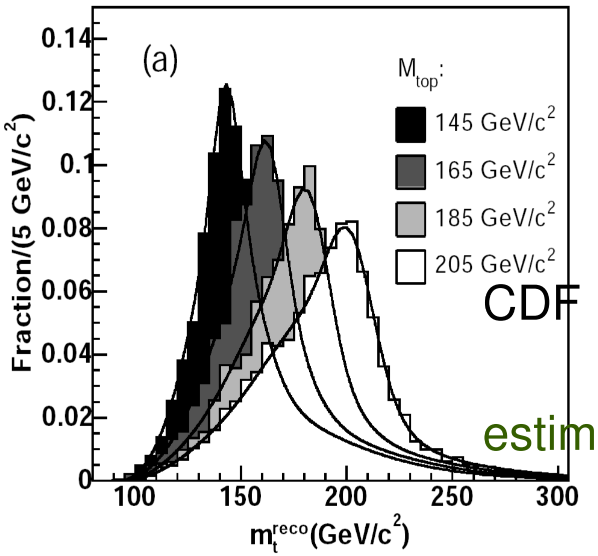
Template Measurements

Imagine you want to measure some quantity, e.g. the mass m_X of a resonance X.

Typically:

- pick a final state
- trigger on it and select events
- in each selected event, ...
  - reconstruct final state particles
  - compute value of an ***estimator***
  - fill estimator into a histogram
- look at the histogram
- know that the detector isn't perfect
- ***compare data histogram with MC histograms***
- extract the mass m_X from a fit
- (worry about systematics...)

- generate signal and background MC (signal for various assumed masses m_X)
- simulate trigger, select events
- in each selected event, ...
  - reconstruct final state particles
  - compute value of the estimator
  - fill estimator into another histogram: templates for various m_X



CDF measurement of m_{top}

estimator: m_{top}^{reco}

Measurement Strategies

Pros and cons of the **template measurement method**:

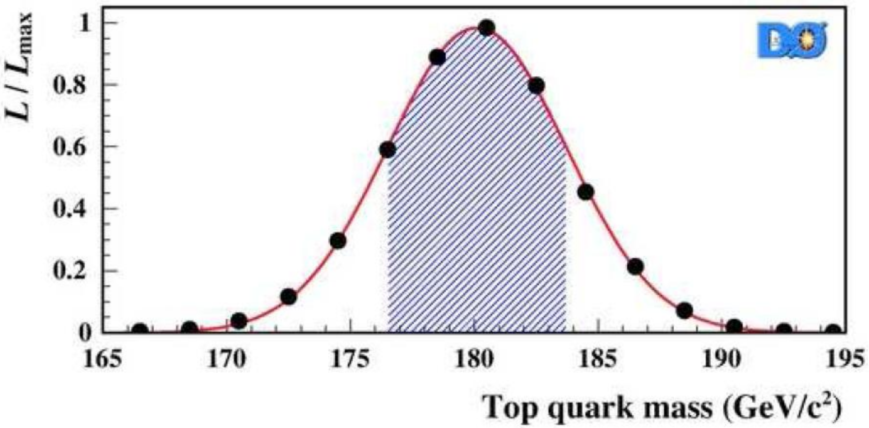
- straightforward to implement
- you see what you're going to get
- only one estimator per event: have to choose one set of reconstructed 4-momenta
- what estimator to take if full kinematic reconstruction of the event is impossible?
- all events enter with the same weight, but: ...
  - likelihood of an event to be signal different for different events
  - different amounts of mass information in different events

Next: **matrix element method**...

The Matrix Element Method

The matrix element method to measure a set of parameters $\vec{\alpha}$

- pick a reaction that depends on $\vec{\alpha}$
- select events
- compute the likelihood L_{sample} to observe the sample of selected events as a function of assumed values of the parameters $\vec{\alpha}$ you want to measure
- fit -ln L_{sample} and determine $\vec{\alpha}$
- done :-)



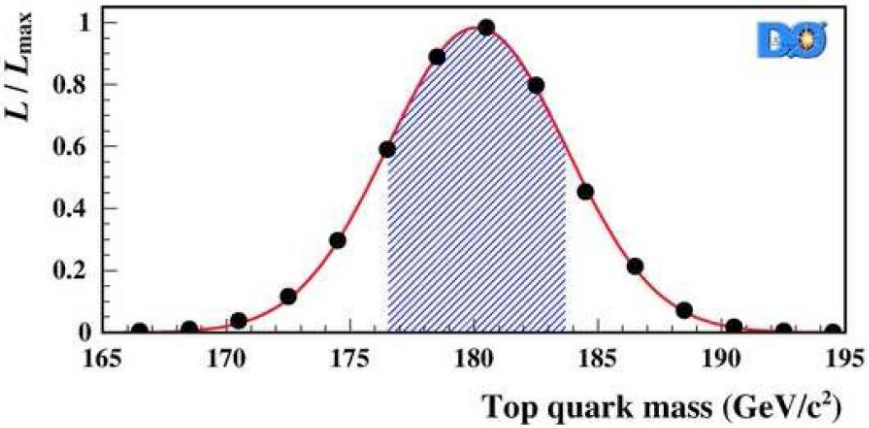
The Matrix Element Method

The matrix element method to measure a set of parameters $\vec{\alpha}$

- pick a reaction that depends on $\vec{\alpha}$
- select events

the fine print:

- compute the likelihood L_{sample} to observe the sample of selected events as a function of assumed values of the parameters $\vec{\alpha}$ you want to measure
- fit -ln L_{sample} and determine $\vec{\alpha}$
- done



The Matrix Element Method

How to compute L_{sample}:

- events are independent

$$L_{\text{sample}}(\vec{\alpha}\quad) = \prod_{i=1}^{N} L_{\text{evt}}(x_i,\ \vec{\alpha}\quad)$$

The Matrix Element Method

How to compute L_{sample}:

- events are independent

$$L_{\text{sample}}(\vec{\alpha},\ \vec{f}) = \prod_{i=1}^{N} L_{\text{evt}}(x_i,\ \vec{\alpha},\ \vec{f})$$

- each event can arise from different processes

$$L_{\text{evt}}(x_i,\ \vec{\alpha},\ \vec{f}) = \sum_{\text{processes}} f_P L_P(x_i,\ \vec{\alpha})$$

The Matrix Element Method

How to compute L_{sample}:

- events are independent

$$L_{\text{sample}}(\vec{\alpha},\, \vec{f}) = \prod_{i=1}^{N} L_{\text{evt}}(x_i,\, \vec{\alpha},\, \vec{f})$$

- each event can arise from different processes

$$L_{\text{evt}}(x_i,\, \vec{\alpha},\, \vec{f}) = \sum_{\text{processes}} f_P L_P(x_i,\, \vec{\alpha})$$

- where the likelihoods for an event x_i to arise from any process sums up to one

$$\sum_{\text{processes}} f_P = 1$$

The Matrix Element Method

How to get the likelihood L_P for an event x_i to have arisen from process P:

- the **detector *W*** has smeared the **original event *y*** so that we **measure x_i** :

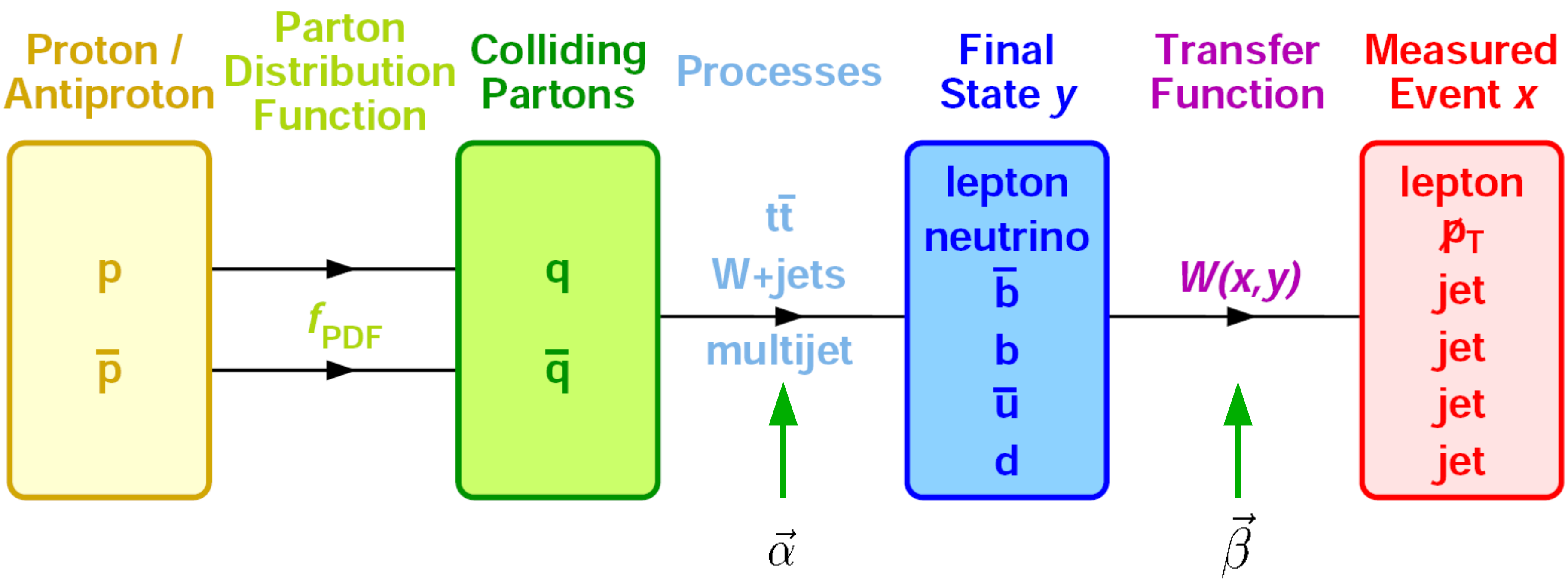
$$L_P(x_i) = \int_y L_P(y) W(x_i,\, y) \mathrm{d}y$$

- we can compute the likelihood to produce an **event *y***
  (likelihood is proportional to the differential cross section):

$$L_P(y) \sim \mathrm{d}\sigma_P(y) \sim |\mathcal{M}(pp \xrightarrow{P} y)|^2 \, \mathrm{d}\phi$$

The Matrix Element Method

Example: ttbar production at the Tevatron:



- only the matrix element $|\mathcal{M}(y)|^2$ depends on the parameters $\vec{\alpha}$ (e.g.: top mass)
- simultaneously measure parameters $\vec{\beta}$ describing the detector response *W* (e.g.: jet energy scale) => multi-parameter fit of -ln $L_{\rm sample}(\vec{\alpha},\ \vec{\beta},\ \vec{f})$

Measurement Strategies

Pros and cons of the **template measurement method**:

- straightforward to implement
- you see what you're going to get
- only one estimator per event: have to choose one set of reconstructed 4-momenta
- what estimator to take if full kinematic reconstruction of the event is impossible?
- all events enter with the same weight, but: ...
  - likelihood of an event to be signal different for different events
  - different amounts of mass information in different events

Pros and cons of the **matrix element method**:

- hard to compute integrals
- hard to debug
- make optimal use of kinematic information
- natural to use even if full kinematic reconstruction impossible
- each event is assigned an optimal weight

Overview

Prologue:

- on a foggy day
- measurement of the top quark mass in 2004

The method:

- how the matrix element measurement method works
- how this is different from 'normal' measurements

Applications of the method:

- measurement of the top quark mass, including technical issues
- calibration
- other measurements, very briefly

What do I mean by m_{top}?

- various different meanings of the same word

Wish list

Measurement of the Top Quark Mass

... as an example of how to implement such a measurement

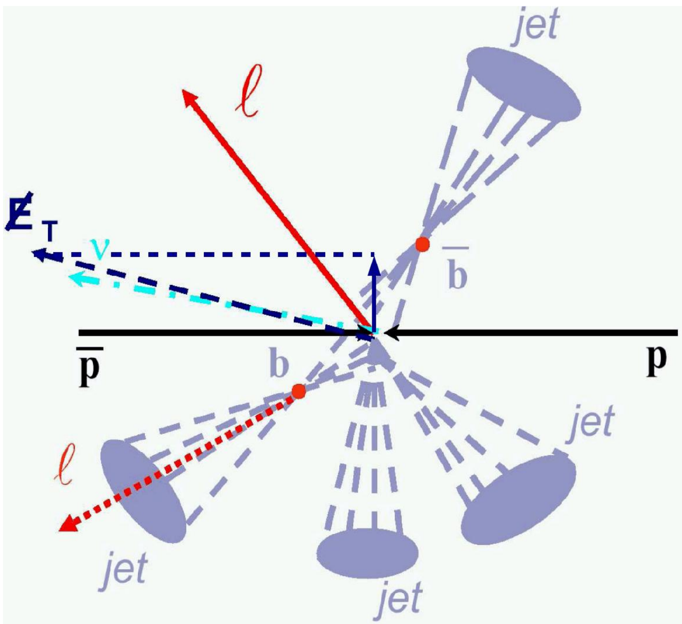
Implementations of the matrix element method:

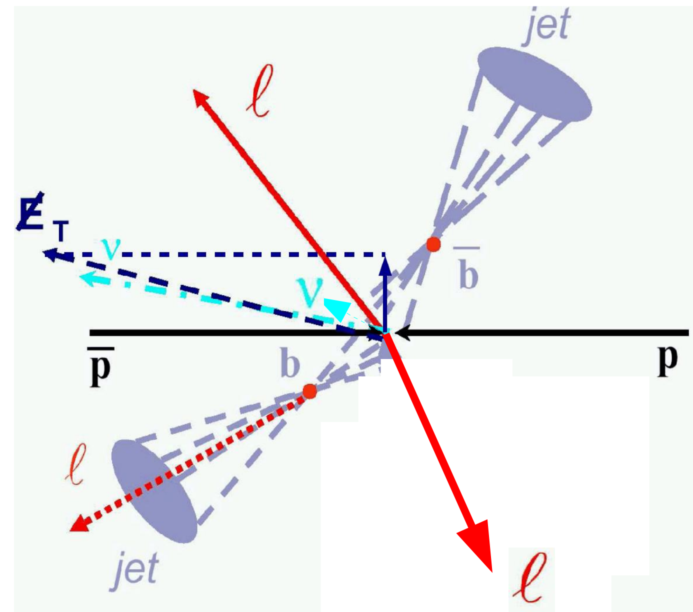
- shown here: results with "CERN-Mainz-Munich-Hamburg" code
- the competition:
  - MadWeight
  - program for original D0 matrix element m_{top} measurement (Nature)
  - CDF programs
  - new implementations for ATLAS, others I don't know of...

Measurement of the Top Quark Mass

... as an example of how to implement such a measurement

decay channels: “lepton+jets” “dilepton”



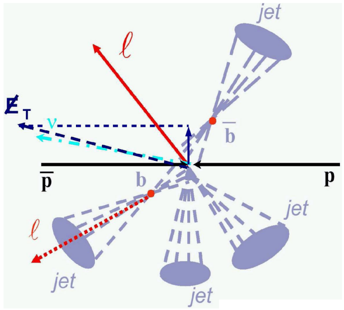


- relatively clean (main bkg: *W*+jets)
- full kinematic reconstruction possible
- 24 possible assignments of jets to final-state (anti-)quarks

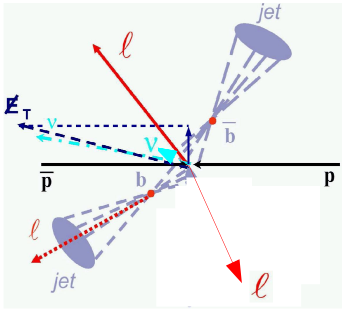
- very clean (main bkg: *Z*+jets, *WW*+jets)
- full kinematic reconstruction impossible
- 2 possible assignments of jets to final-state (anti-)b-quarks

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*

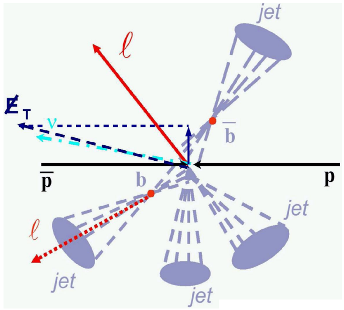


$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

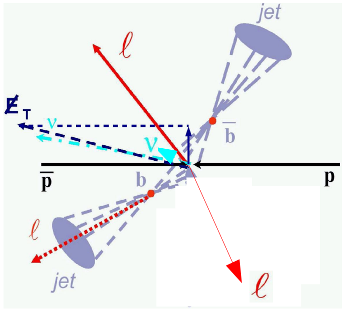
- 6 final-state particles, known masses
  => 18-dimensional MC integration over possible final states *y*

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



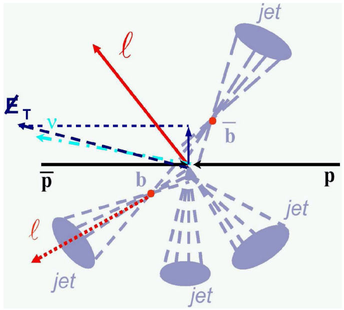
$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

- 6 final-state particles, known masses
  => 18-dimensional MC integration over possible final states *y*

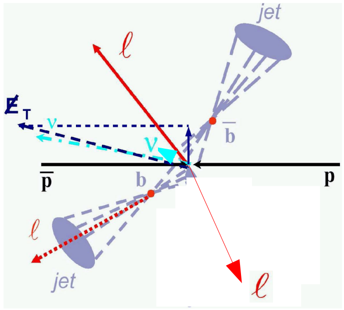
- switch on brain
  0. perhaps assume p_T(ttbar)=0
  1. know we have narrow resonances in the decay chain
  2. consider only final states *y* that can have led to measured event *x*
  => reduce number of dimensions

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

1. Narrow resonances in the decay chain
for ttbar events:

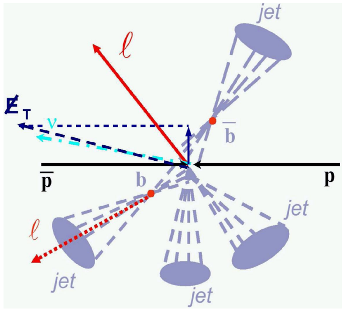
- **top quark with hadronic W decay:**
  - W resonance is narrow in comparison with jet energy resolution
  - top resonance is narrow in comparison with jet energy resolution

- **top quark with leptonic W decay:**
  - W resonance is narrow in comparison with unknown ν p_z
  - top resonance is narrow in comparison with jet energy, ν p_T resolution

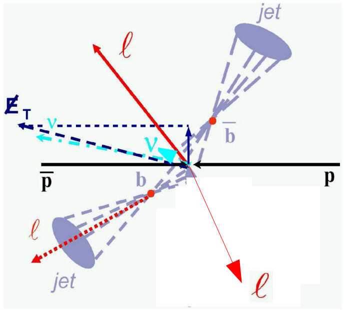
=> only final states y with appropriate masses contribute significantly to integral

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

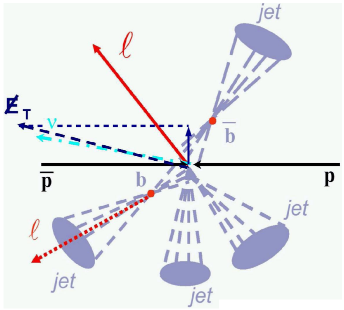
2. Consider only final states *y* that can have led to measured event *x*
for ttbar events:

- (anti-)quark => jet; electron => electron; muon => muon
- electrons: energy and direction well-measured (compare: ν, q)
- muons: direction well-measured (compare: ν, q)
  integrate over q/p_T
- quarks: direction well-measured (compare: energy)
  integrate over E_q

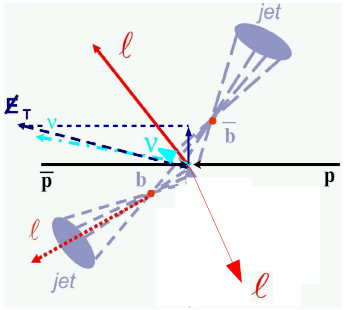
- assignment final-state particles ↔ measured objects unknown

Signal Likelihood Calculation

"lepton+jets"
event x



"dilepton"
event x



$$L_P(x_i) = \int_y L_P(y) W(x_i,\, y) \mathrm{d}y$$

Choice of variables for integration over final states y:

- minimization of dimension of MC integration
- easy computation of 4-momenta from values of integration variables (quadratic equation)

Integration variables (lepton+jets case):

- $m_{t_{had}}{}^2$, $m_{t_{lep}}{}^2$, $m_{W_{had}}{}^2$
- E_u, $(q/p_T)_\mu$
- $(p^z)_{b\nu}$

=> 5-6 dimensions

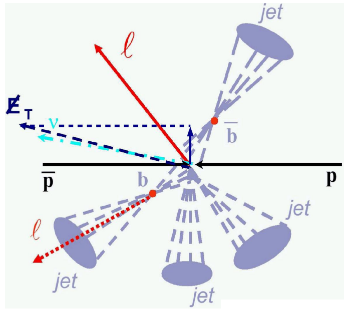
Integration variables (dilepton case):

- $m_{t_1}{}^2$, $m_{t_2}{}^2$
- E_{b_1}, E_{b_2}, $(q/p_T)_\mu$
- $(p^x)_{\nu_1}$-$(p^x)_{\nu_2}$, $(p^y)_{\nu_1}$-$(p^y)_{\nu_2}$

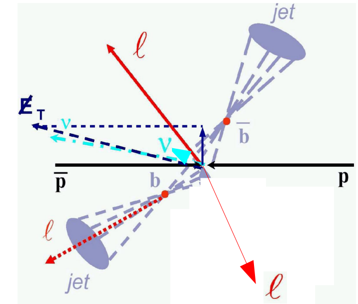
=> 6-8 dimensions

Signal Likelihood Calculation

“lepton+jets” event x



“dilepton” event x



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

Naïve estimate of computation time...

for one lepton+jets event *x*:

- **per quark-jet assignment:**
  - $\mathcal{O}(10^4\text{-}10^5)$ calls => 2-5% accuracy
  - takes 2-5 s

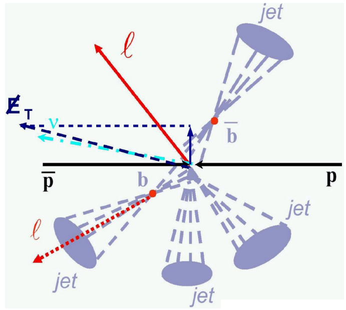
dedicated implementation of $|\mathcal{M}|^2$ for
g g → t tbar → b u dbar bbar l ν
q qbar → t tbar → b u dbar bbar l ν

detector parametrization optimized for speed

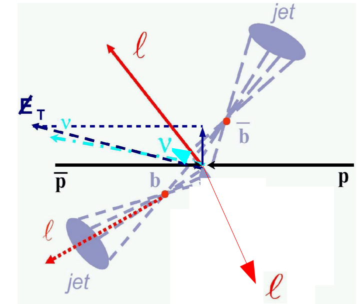
no more than $\mathcal{O}$(ms) per quark-jet assignment, please... :-)

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

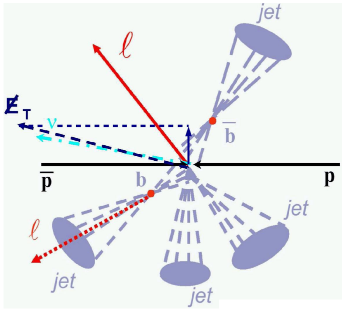
Naïve estimate of computation time...

for one lepton+jets event *x*:

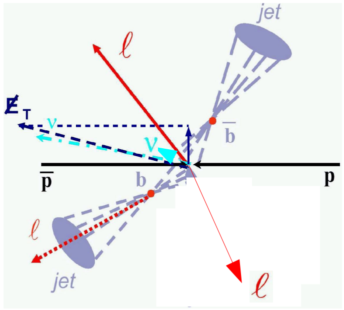
- **per quark-jet assignment:**
  - $\mathcal{O}(10^4\text{-}10^5)$ calls => 2-5% accuracy
  - takes 2-5 s
- **for all assignments and one top quark mass assumption:**
  - 24 · (2-5) s = (1-2) min
- **for $\mathcal{O}(10)$ m_{top} assumptions:**
  - ~10-20 CPU-minutes per event

Signal Likelihood Calculation

"lepton+jets" event *x*



"dilepton" event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

Naïve estimate of computation time...
for one lepton+jets event *x*:

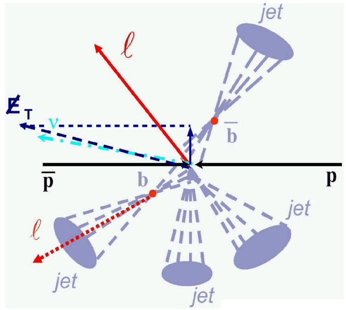
- **per quark-jet assignment:**
  - $\mathcal{O}(10^4\text{-}10^5)$ calls => 2-5% accuracy
  - takes 2-5 s
- **for all assignments and one top quark mass assumption:**
  - 24 · (2-5) s = (1-2) min
- **for $\mathcal{O}(10)$ m_{top} assumptions:**
  - ~10-20 CPU-minutes per event

Naïve estimate of computation time...
for one dilepton event *x*:

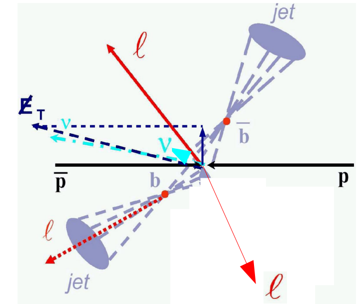
- **per quark-jet assignment:**
  - $\mathcal{O}(10^5\text{-}10^6)$ calls => 2-5% accuracy
  - takes $\mathcal{O}(10\text{-}30)$ s
- **for both assignments and one top quark mass assumption:**
  - 2 · 30 s = 1 min
- **for $\mathcal{O}(10)$ m_{top} assumptions:**
  - ~10 CPU-minutes per event

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

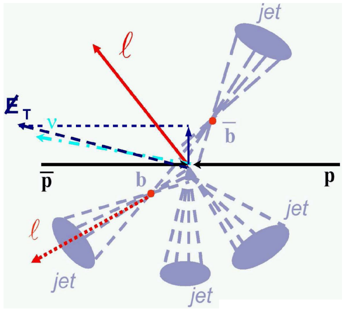
Naïve estimate of computation time...

for 1000 data events:	~ 10 CPU days
for 3000 MC events:	~ 30 CPU days
• but you need those for say 5 true top mass values:	~150 CPU days
for background likelihood computation	~factor 2
for systematic uncertainties:	~factor 2

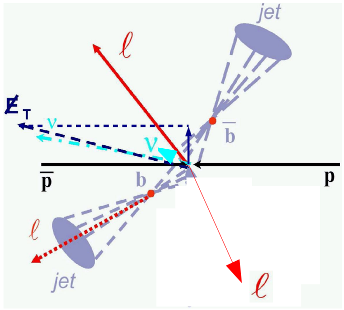
=> 2 CPU years for a top mass measurement

Signal Likelihood Calculation

"lepton+jets"
event *x*



"dilepton"
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

Naïve estimate of computation time...

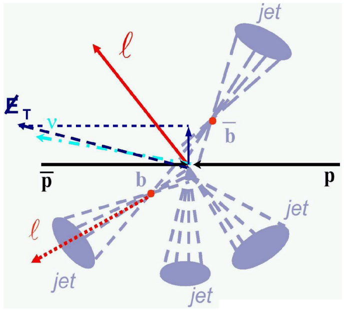
for 1000 data events:	~ 10 CPU days
for 3000 MC events:	~ 30 CPU days
• but you need those for say 5 true top mass values:	~150 CPU days
for background likelihood computation	~factor 2
for systematic uncertainties:	~factor 2

=> 2 CPU years for a top mass measurement

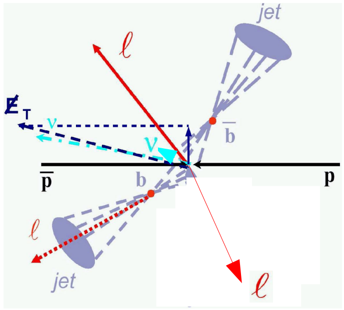
- usually measure >1 parameter (top mass, jet energy scale, b-jet energy scale)
  => factor 5-10 for each additional parameter
- isr jets easily make things worse

Signal Likelihood Calculation

“lepton+jets”
event *x*



“dilepton”
event *x*



$$L_P(x_i) = \int_y L_P(y) W(x_i, y) \mathrm{d}y$$

CPU years for a top mass measurement

=> optimize:

- quick look at quark-jet assignments to see which ones really contribute before starting lengthy calculations
- avoid recomputation of values you know
- ...

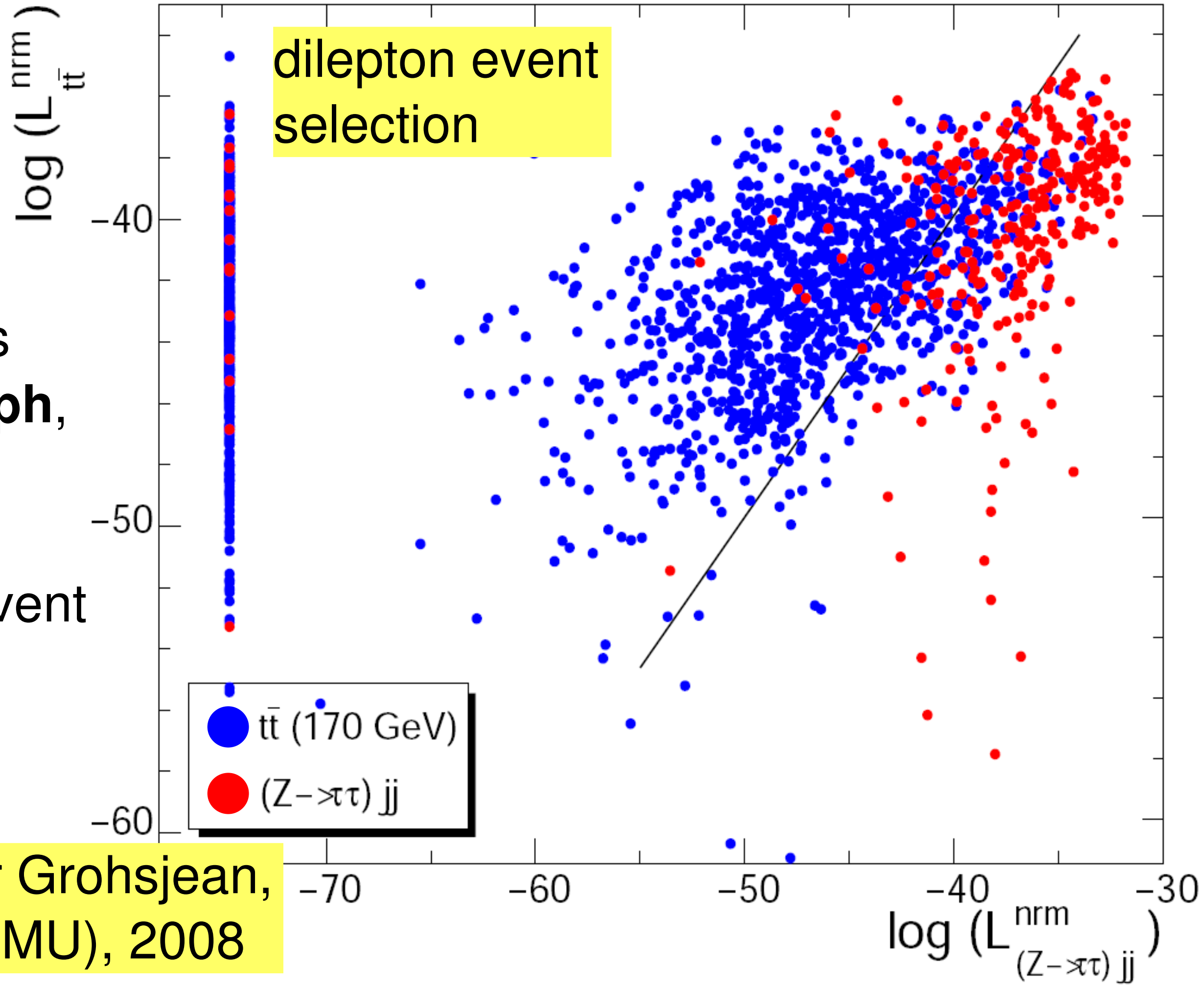
Background Likelihood Calculation

signal likelihood:

- simple process, dedicated implementation of LO matrix element
- calculation for several assumed values of m_{top}

background likelihood:

- many processes
  e.g. for *W*+jets, *Z*+jets
  => take implementation
  from existing programs
  (interfaces to **MadGraph**, **Vecbos**, **Alpgen**)
  => slow
- only need one value per event



PhD thesis Alexander Grohsjean, Munich University (LMU), 2008

Testing the Method

Approximations:

- LO matrix element
- not all possible processes considered
- parametrized detector resolution (no Geant)

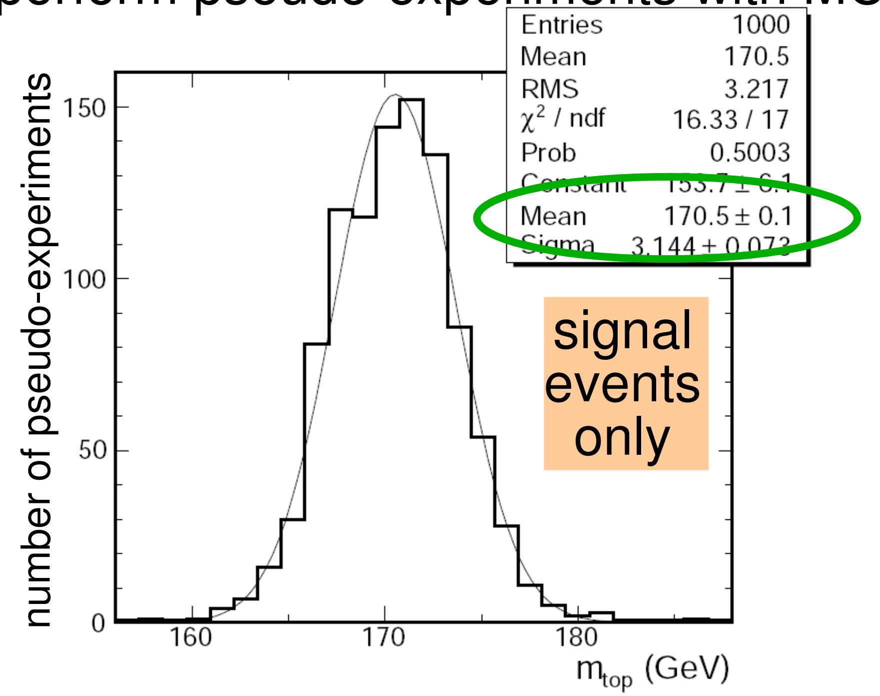
=> need to **test & calibrate** the method with **full Monte Carlo** simulation!

Testing the Method

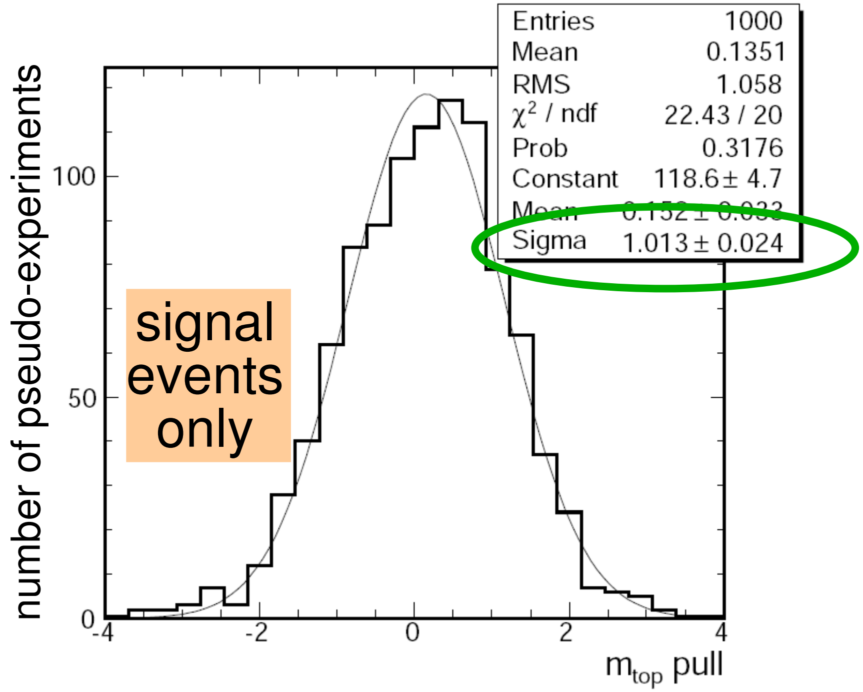
Approximations:
- LO matrix element
- not all possible processes considered
- parametrized detector resolution (no Geant)

=> need to **test & calibrate** the method with **full Monte Carlo** simulation!

=> perform pseudo-experiments with MC events:



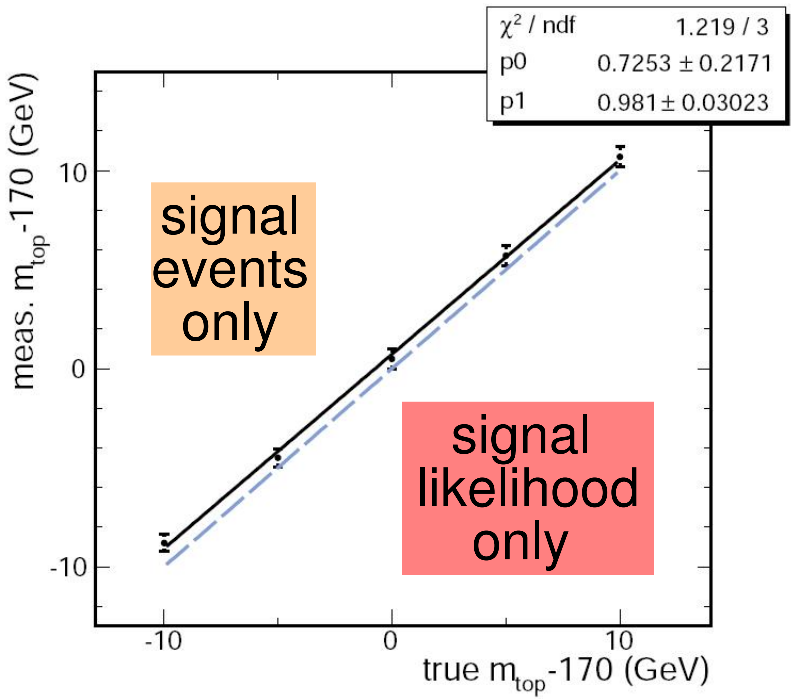
reproduce true m_{top}

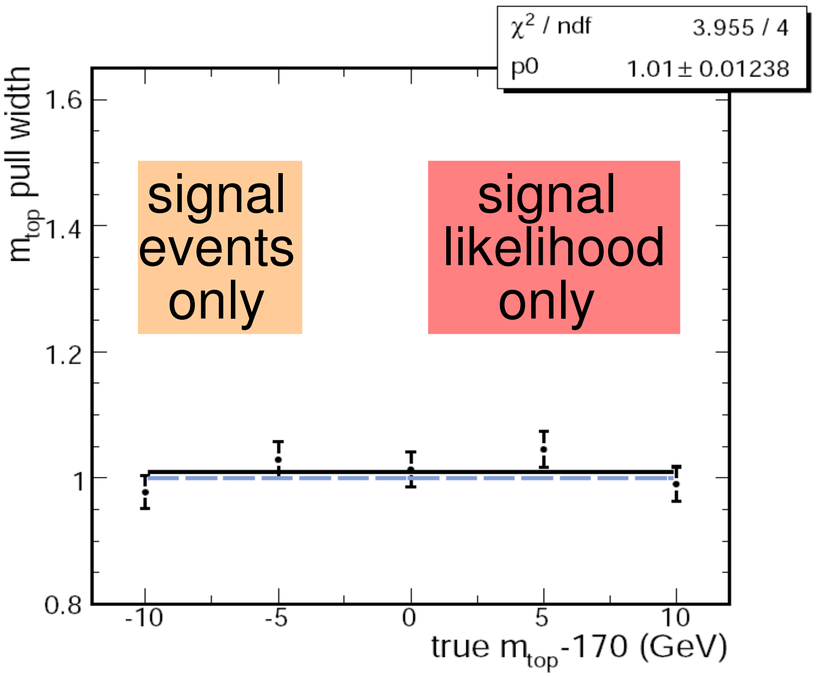


fitted uncertainty correct

Testing the Method

Test of the method: with smeared parton-level events





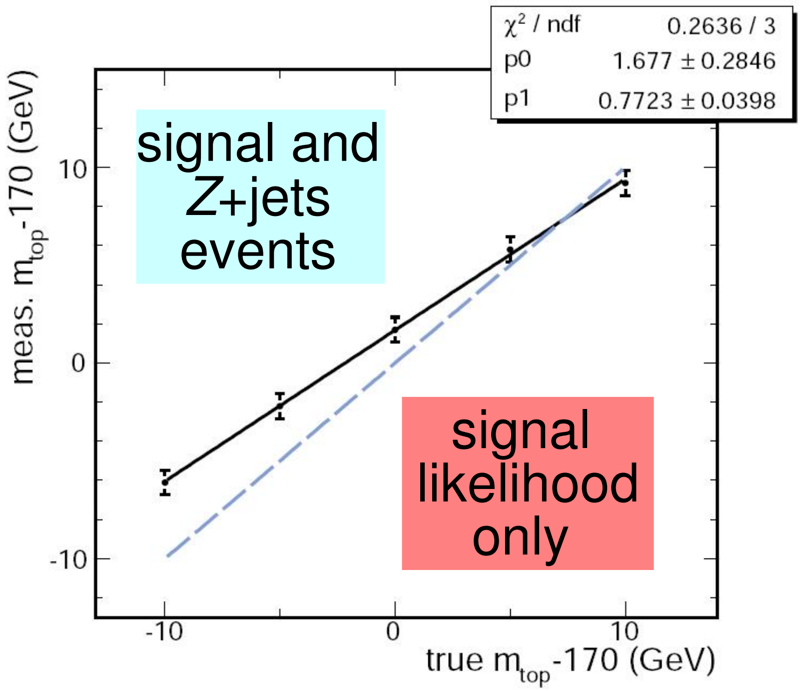
reproduce true m_{top}

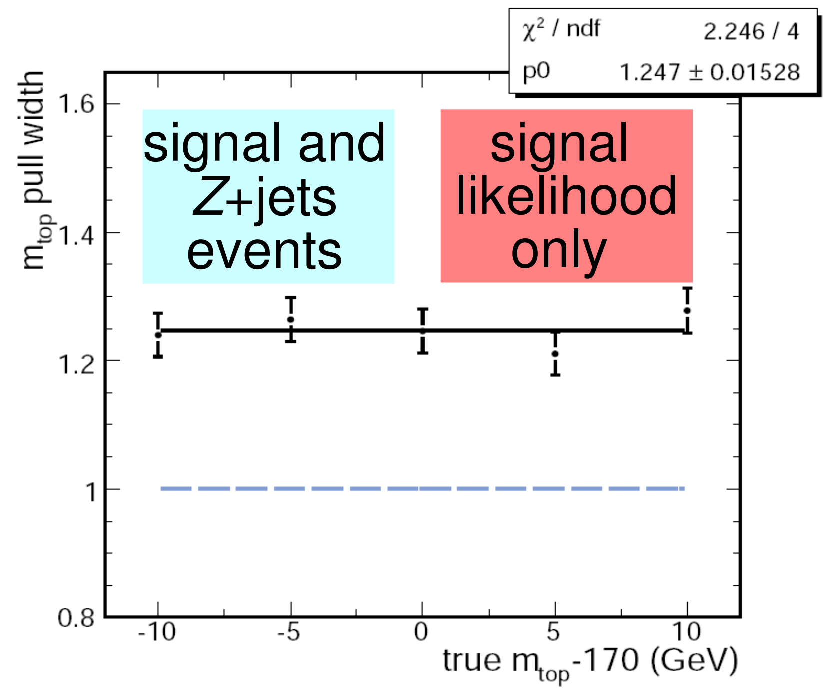
fitted uncertainty correct

for any input value of m_{top}

Testing the Method

Test of the method: with smeared parton-level events





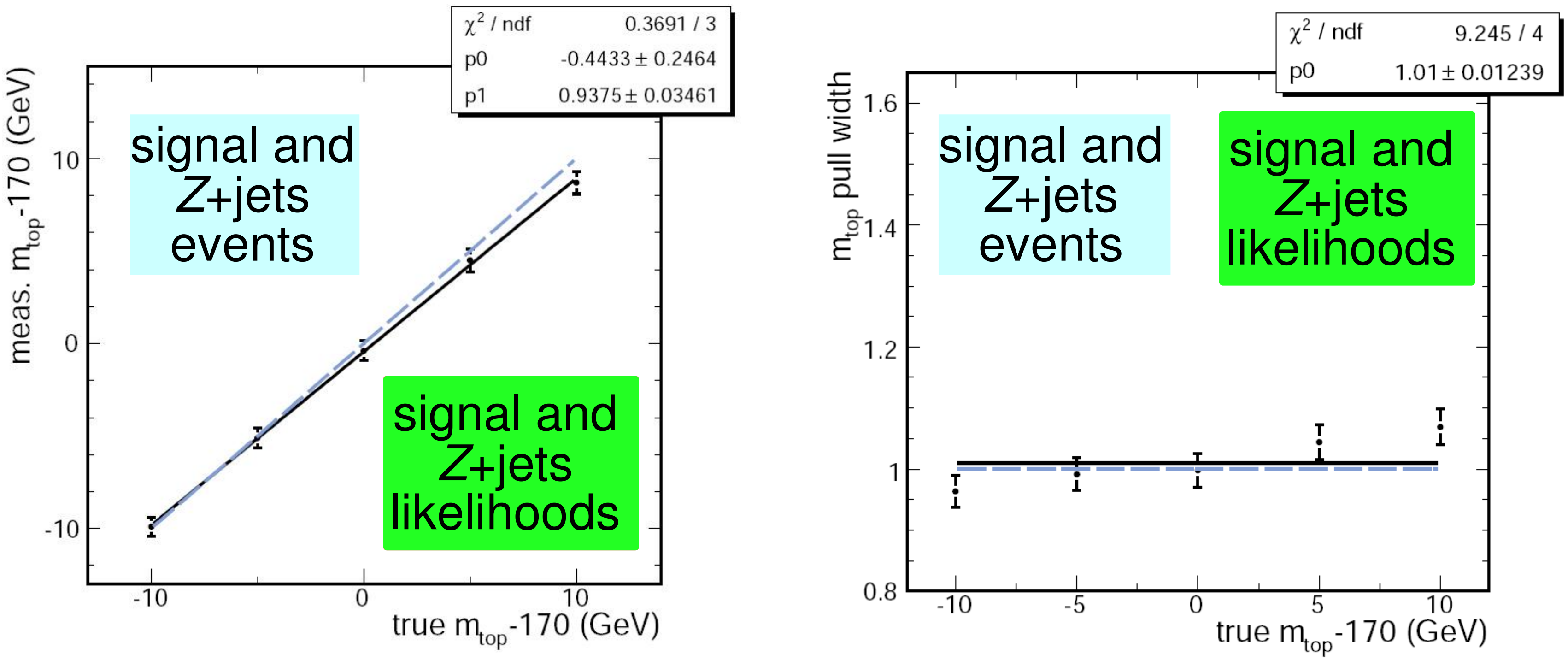
no longer reproduce true m_{top}

fitted uncertainty no longer correct

for any input value of m_{top}

Testing the Method

Test of the method: with smeared parton-level events



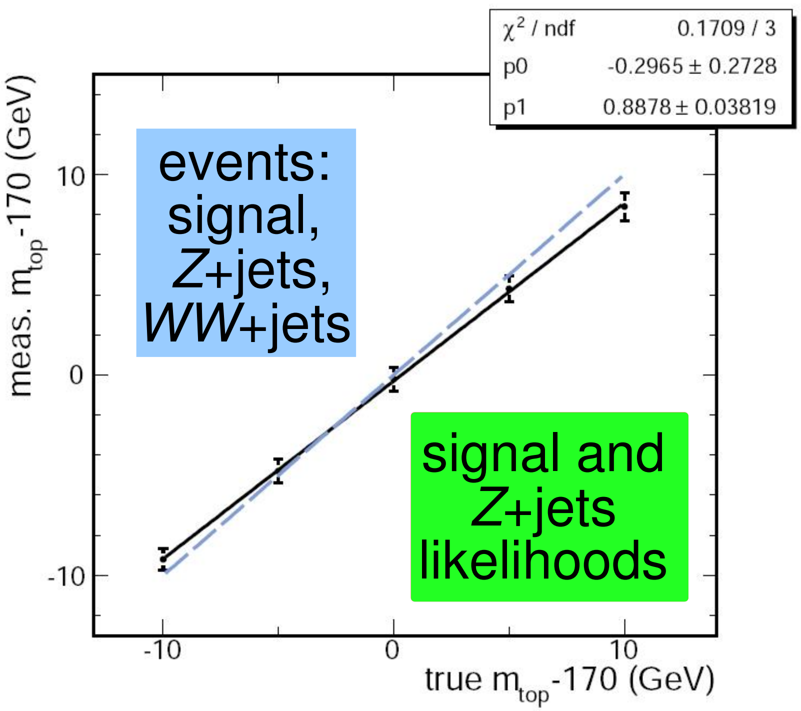
expect to reproduce true m_{top}

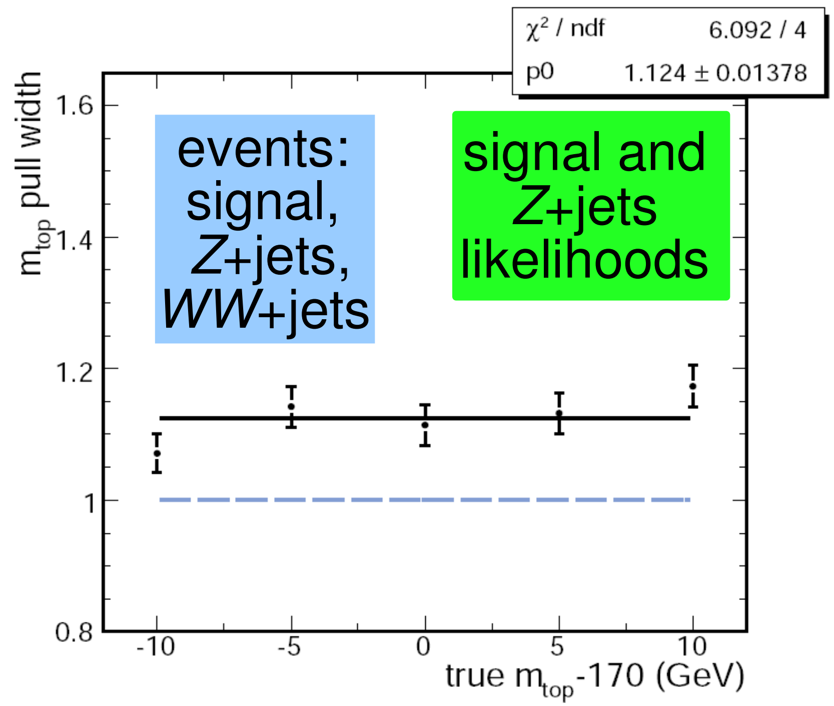
fitted uncertainty correct

for any input value of m_{top}

Testing the Method

Test of the method: with smeared parton-level events





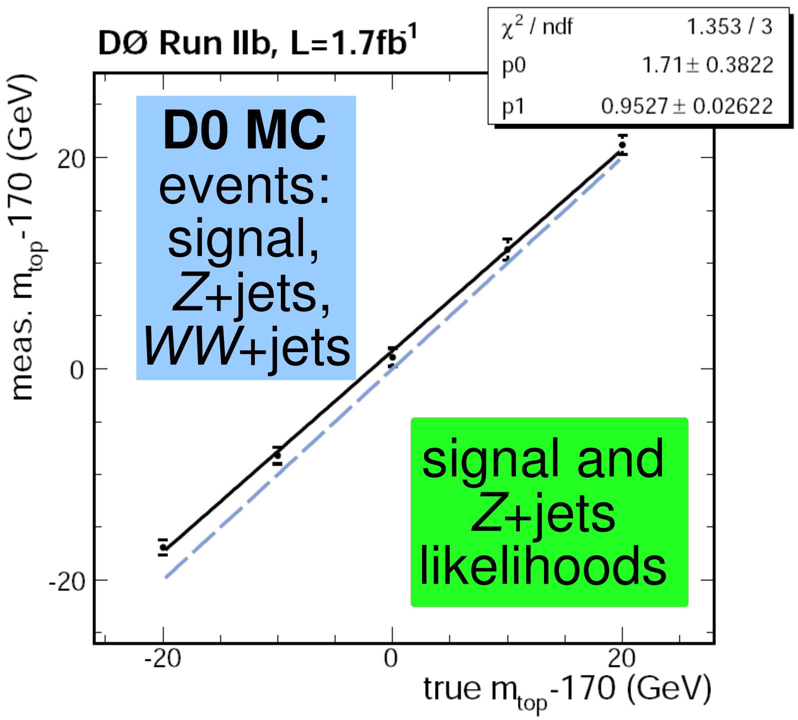
fitted m_{top} needs to be calibrated

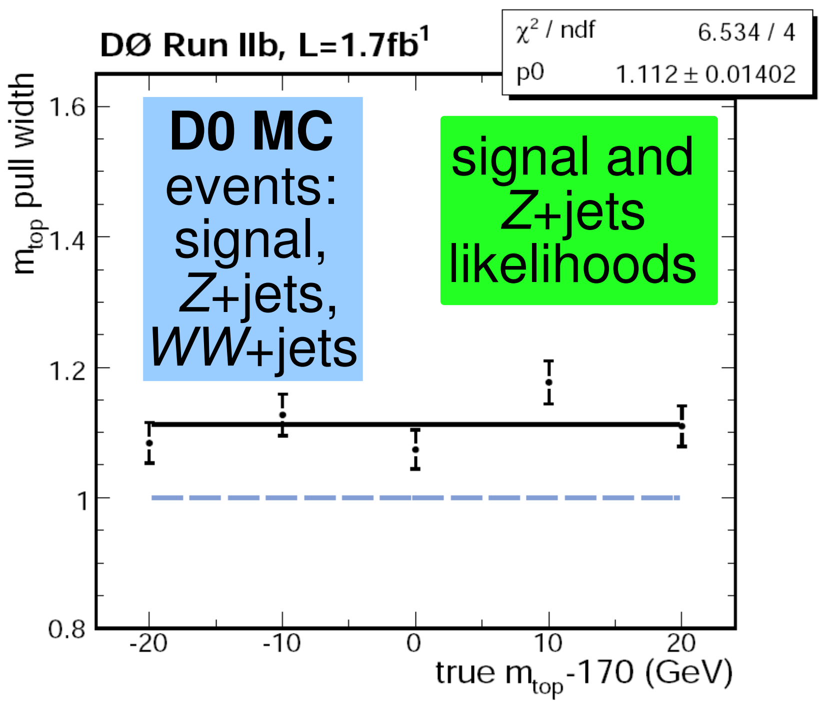
fitted uncertainty needs to be adjusted

for any input value of m_{top}

Calibrating the Method

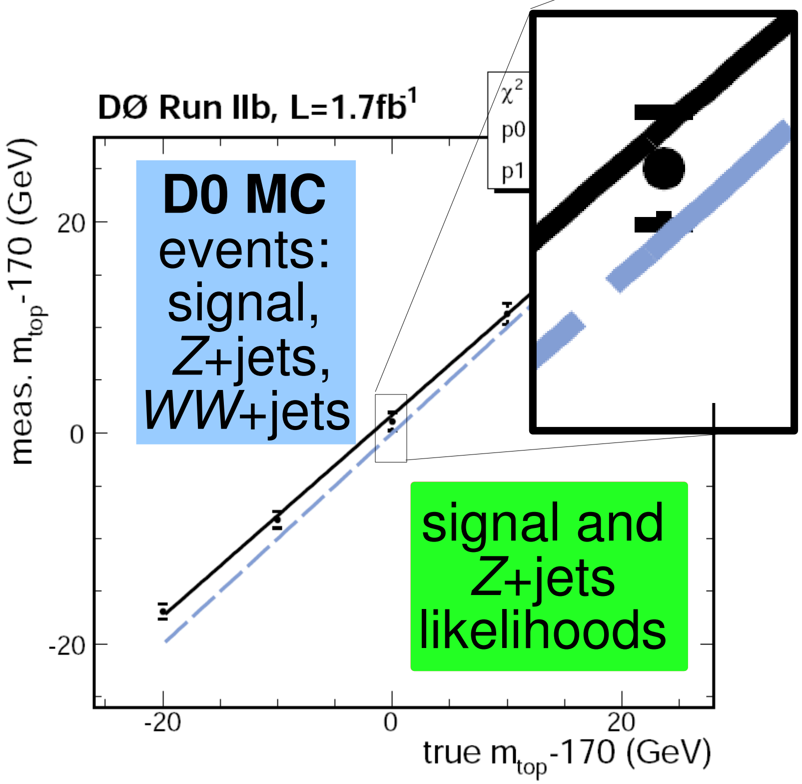
Calibration of the method: with fully simulated events





Calibrating the Method

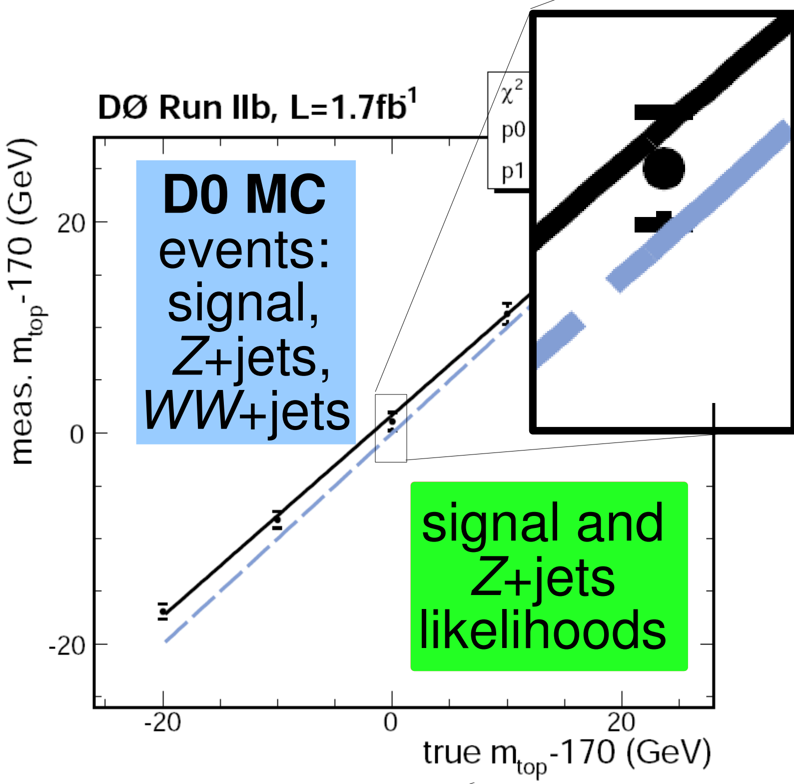
Calibration of the method:



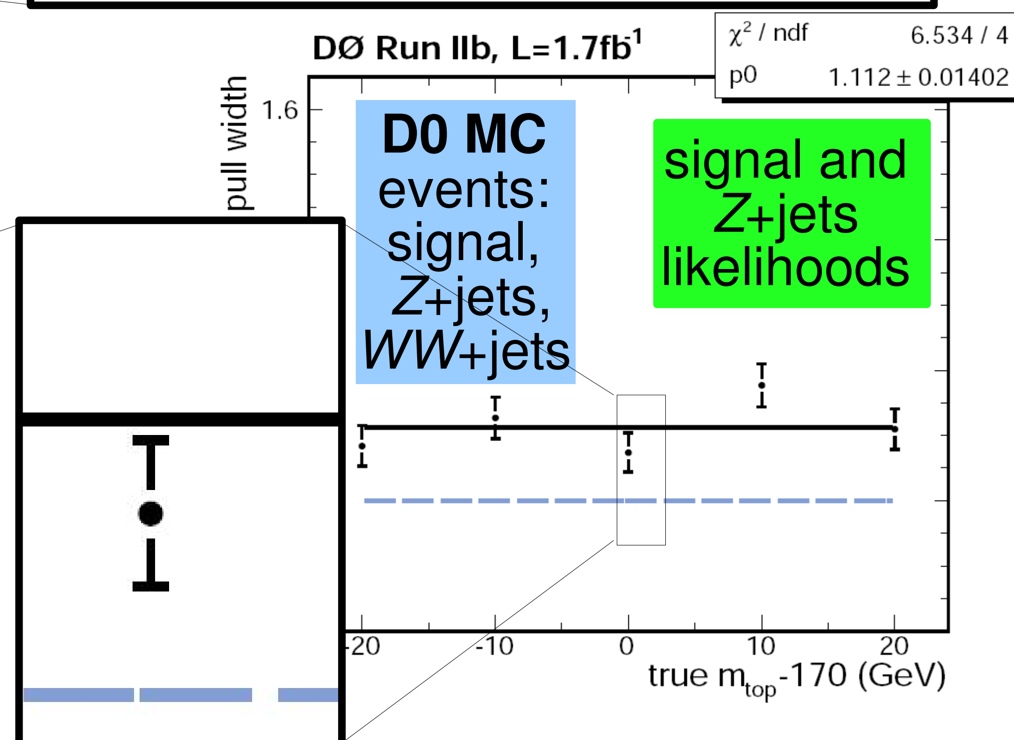
- **size of error bar**
  => systematic uncertainty from calibration procedure
- **deviation from perfect curve**
  => not a problem if MC used for calibration is correct
- **systematic errors**
  => vary MC parameters

Calibrating the Method

Calibration of the method:



- **size of error bar**
  => systematic uncertainty from calibration procedure
- **deviation from perfect curve**
  => not a problem if MC used for calibration is correct
- **systematic errors**
  => vary MC parameters

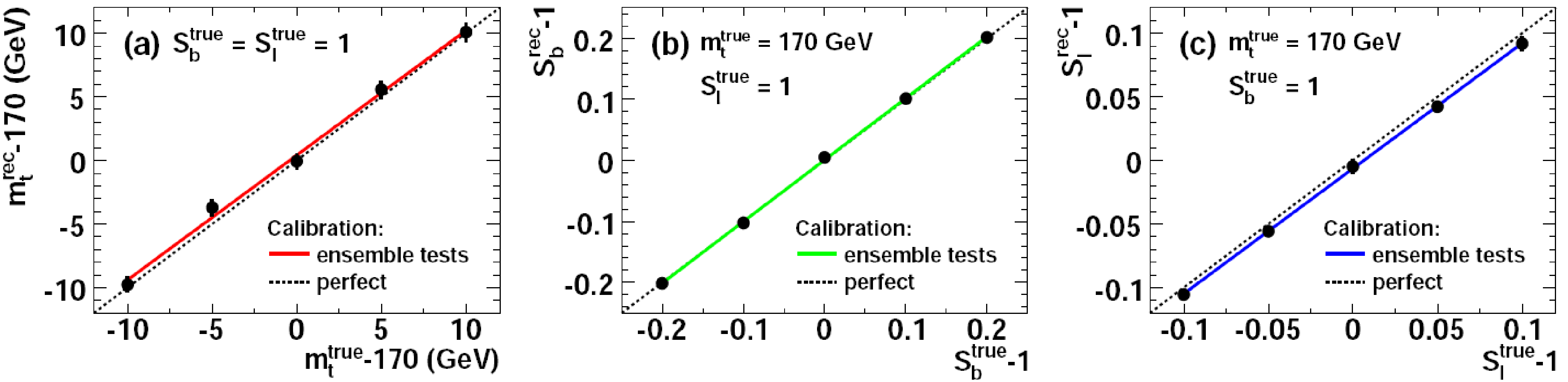


- **deviation from 1.0**
  => scale fitted statistical uncertainty accordingly

More Measurements

All of this also works for...

- simultaneous measurement of *top mass* and *jet energy scales*
  => reduce systematics

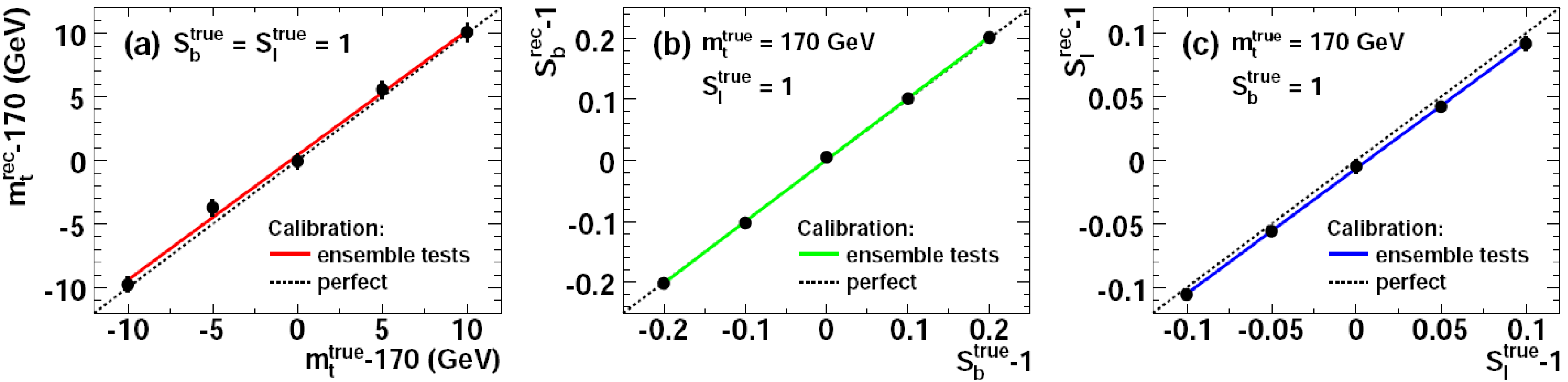


top mass *energy scale for b jets* *energy scale for light jets*

More Measurements

All of this also works for...

- simultaneous measurement of *top mass* and *jet energy scales*
  => reduce systematics



top mass *energy scale for b jets* *energy scale for light jets*

All of this also works for...

- measurement of m_H once a signal has been established

All of this also works in principle to...

- look for new physics, but ***only if*** you know exactly what you're looking for

Overview

Prologue:

- on a foggy day
- measurement of the top quark mass in 2004

The method:

- how the matrix element measurement method works
- how this is different from 'normal' measurements

Applications of the method:

- measurement of the top quark mass, including technical issues
- calibration
- other measurements, very briefly

What do I mean by m_{top}?

- various different meanings of the same word

Wish list

Interpretation of Measurements

What do I mean by *top mass*?

depending on the context, ...

- ... the mass m_t^{PDG} of a particle **(look it up in the PDG book)**
- ... a parameter m_t^{MC} of an event generator **(Pythia parameter)**
- ... the mass m_t^{true} that a top quark has in an event ("final state") y **(m_t^{true} is within a few σ_t^{MC} of m_t^{MC} ...)**
- ... the reconstructed mass m_t^{reco} computed from objects in the reconstructed event x **(estimator in the template method)**
- ... an integration variable, e.g. $m_{t_{had}}{}^2$ **(to integrate over final states y)**
- ... the current value of the integration variable **(in an integration step)**
- ... the result of a measurement.

Interpretation of Measurements

What do I mean by *top mass*?

depending on the context, ...

- ... the mass m_t^{PDG} of a particle
- ... a parameter m_t^{MC} of an event generator
- ... the mass m_t^{true} that a top quark has in an event ("final state") y
- ... the reconstructed mass m_t^{reco} computed from objects in the reconstructed event x
- ... an integration variable, e.g. $m_{t_{had}}^2$
- ... the current value of the integration variable
- ... the result of a measurement.

hope the generator reproduces nature

calibrate the measurement with generated events

Interpretation of Measurements

Validity of measurements with the matrix element method

Strictly speaking:

- measure a parameter of an event generator
- may use result...
  - to constrain generator parameters
  - to cross-check model

- if new physics is discovered
  - generator model proven wrong
  - measurement is “wrong”

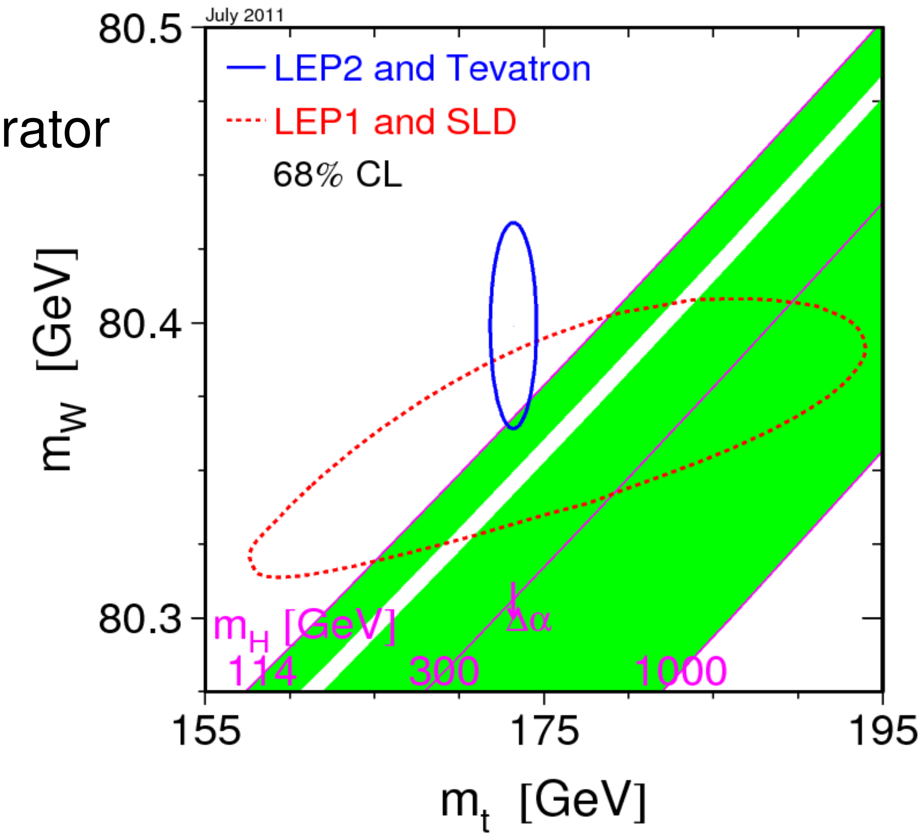
Interpretation of Measurements

Validity of measurements with the matrix element method

Strictly speaking:

- measure a parameter of an event generator
- may use result...
  - to constrain generator parameters
  - to cross-check model

- if new physics is discovered
  - generator model proven wrong
  - measurement is “wrong”

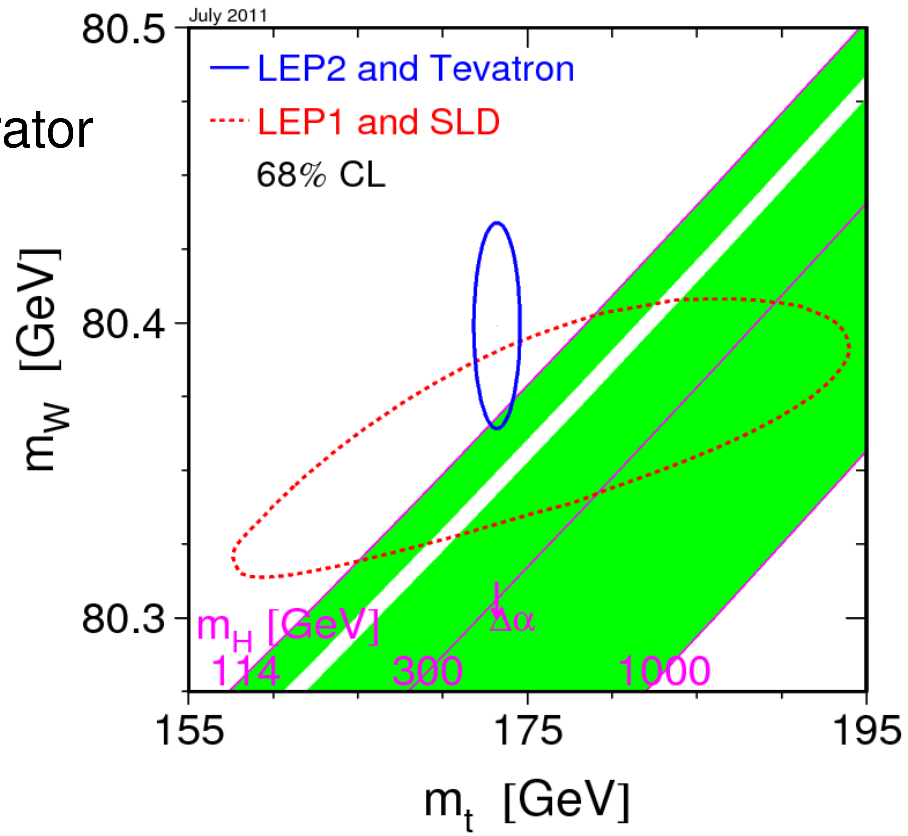


Interpretation of Measurements

Validity of measurements with the matrix element method

Strictly speaking:

- measure a parameter of an event generator
- may use result...
  - to constrain generator parameters
  - to cross-check model
- if new physics is discovered
  - generator model proven wrong
  - measurement is "wrong"
- true also of template measurement
- matrix element method tends to
  - ... have higher statistical sensitivity
  - ... be more affected by new physics



Overview

Prologue:

- on a foggy day
- measurement of the top quark mass in 2004

The method:

- how the matrix element measurement method works
- how this is different from 'normal' measurements

Applications of the method:

- measurement of the top quark mass, including technical issues
- calibration
- other measurements, very briefly

What do I mean by m_{top}?

- various different meanings of the same word

Wish list

Wish List

Current work in Mainz:

- ttbar resonance search at ATLAS:
  - use the matrix element method to reconstruct m_{ttbar} on an event by event basis

Malene Thyssen

- Higgs search and measurements at ATLAS:
  - $H \to WW \to l\,\nu\, l\,\nu$ channel
  - Higgs mass, spin and CP

Wish list:

- event generator that ...
  generates events y in the phase space region where a measured event x happens to be