The Top Quark Mass

Alexander Mitov

C. N. Yang Institute for Theoretical Physics
SUNY Stony Brook

Top mass: why we care?

We need to know the top mass because it is “portable”:

$$M_t = 173.1 \pm 1.3 \text{ GeV}/c^2$$

Current best measurement
CDF+D0: 0903.2503

Places where the top mass is crucial:

- Higgs mass

Precision Electroweak Measurements and Constraints on the Standard Model arXiv:0811.4682v1 [hep-ex]

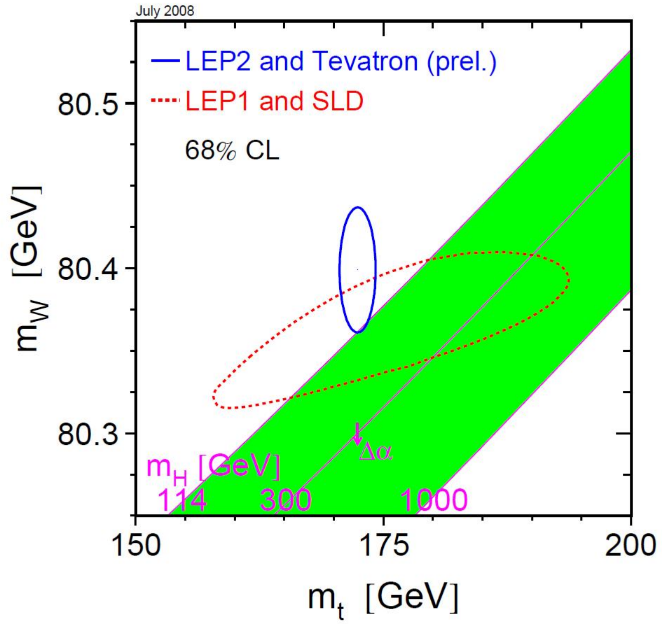
➢ Lower limit from direct searches:

ALEPH, DELPHI, L3, and OPAL Collaboration ‘03

M_H > 114 GeV; recent exclusion of 160-170 GeV range from Tevatron

➢ Indirect constraints from LEP + M_{top} + M_W

$M_H = 84 {}^{+34}_{-26}$ GeV



Top mass: why we care?

Places where the top mass is crucial:

- Higgs-inflation

Bezrukov, Shaposhnikov '07-'08
De Simone, Hertzbergy, Wilczek'08

Assume non-minimal coupling to gravity:

$$\mathcal{L}_h = -|\partial H|^2 + \mu^2 H^\dagger H - \lambda(H^\dagger H)^2 + \xi H^\dagger H \, \mathcal{R}$$

Then: Higgs = inflaton provided:

1) $10^3 < \xi < 10^4$

2) $m_h > 125.7\,\mathrm{GeV} + 3.8\,\mathrm{GeV}\left(\frac{m_t - 171\,\mathrm{GeV}}{2\,\mathrm{GeV}}\right) - 1.4\,\mathrm{GeV}\left(\frac{\alpha_s(m_Z) - 0.1176}{0.0020}\right) \pm \delta$

3) $m_h \lesssim 190\,\mathrm{GeV}$

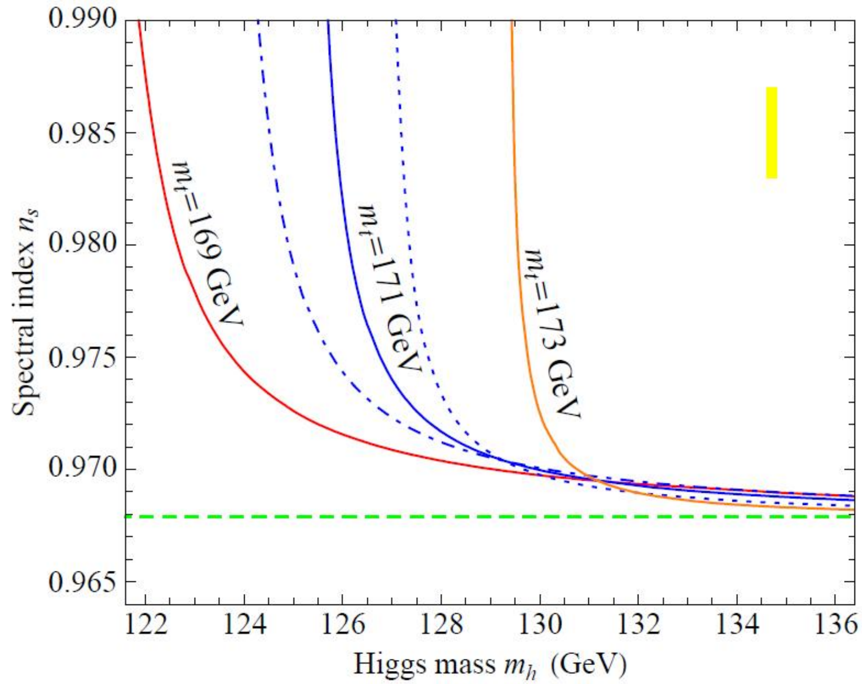
- Theory remains perturbative at high energy,
- Consistent inflation; consistent with WMAP!

Top mass: why we care?

- Higgs-inflation

Bezrukov, Shaposhnikov '07-'08
De Simone, Hertzbergy, Wilczek'08

Provided it works ☺
the model is very predictive!



De Simone, Hertzbergy, Wilczek arXiv:0812.4946v2

Figure 1: The spectral index n_s as a function of the Higgs mass m_h for a range of light Higgs masses. The 3 curves correspond to 3 different values of the top mass: $m_t = 169\,\text{GeV}$ (red curve), $m_t = 171\,\text{GeV}$ (blue curve), and $m_t = 173\,\text{GeV}$ (orange curve). The solid curves are for $\alpha_s(m_Z) = 0.1176$, while for $m_t = 171\,\text{GeV}$ (blue curve) we have have also indicated the 2-sigma spread in $\alpha_s(m_Z) = 0.1176 \pm 0.0020$, where the dotted (dot-dashed) curve corresponds to smaller (larger) α_s. The horizontal dashed green curve, with $n_s \simeq 0.968$, is the classical result. The yellow rectangle indicates the expected accuracy of PLANCK in measuring n_s ($\Delta n_s \approx 0.004$) and the LHC in measuring m_h ($\Delta m_h \approx 0.2\,\text{GeV}$). In this plot we have set $N_e = 60$.

Top mass: why we care?

So, to summarize, the top mass is needed:

- with numerical precision,
- with confidence about its <u>definition.</u>

Recall: mass is not observable; it is a formal parameter and is thus sensitive to its formal definition.

Unless we have a reasonable control over both mass definition and mass value, we cannot be confident we are doing a good job!

How to measure the top mass?

At the LHC the top mass measurement can be done with "confidence"

Here is the idea:

- Find an observable sensitive to the value of the top mass;
- Fix all other parameters and fit the data by tuning the mass.
  (of course, we hope for data with sufficient statistics ☺)
- If beyond LO we become sensitive to the definition of the mass, too.

Example 1: the total top-pair cross-section.

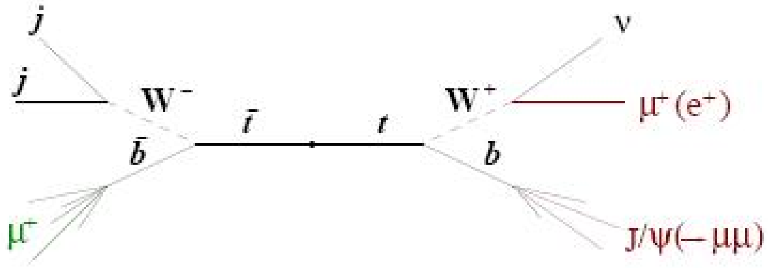
- It allows extraction of the mass with ~ 4% accuracy.

Hint: compare to the current best value from the Tevatron ~ 0.8%

It is not all bad news: we are confident about what we measure ☺

Example 2: "J/Psi final state"

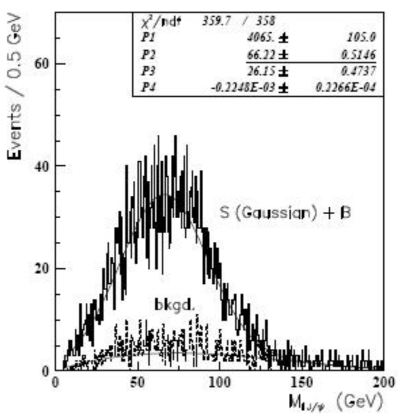
Jet measurements are hard at the LHC; check out the lepton signal

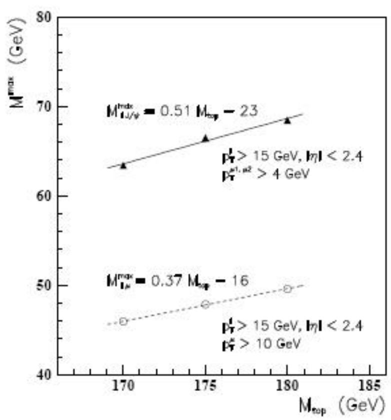


Proposed by: A. Kharchilava '99

R. Chierici, A. Dierlamm CMS NOTE 2006/058
Corcella, Mangano, Seymour '00

Idea: - study the invariant mass distribution of $M_{J/\psi-\ell}$ in top decay
- explore the strong correlation between peak position and M_{top}





- ✓ Experimentally very clean signal
- ✓ Low branching ratio ~ 10^{-5} , but
- ✓ Compensated by large top rates
- ✓ ~ 1000 events/year at LHC (14 TeV)
- ✓ Accuracy ≤ 1 GeV achievable.

The Tevatron

Finally, let's turn our attention to the Tevatron

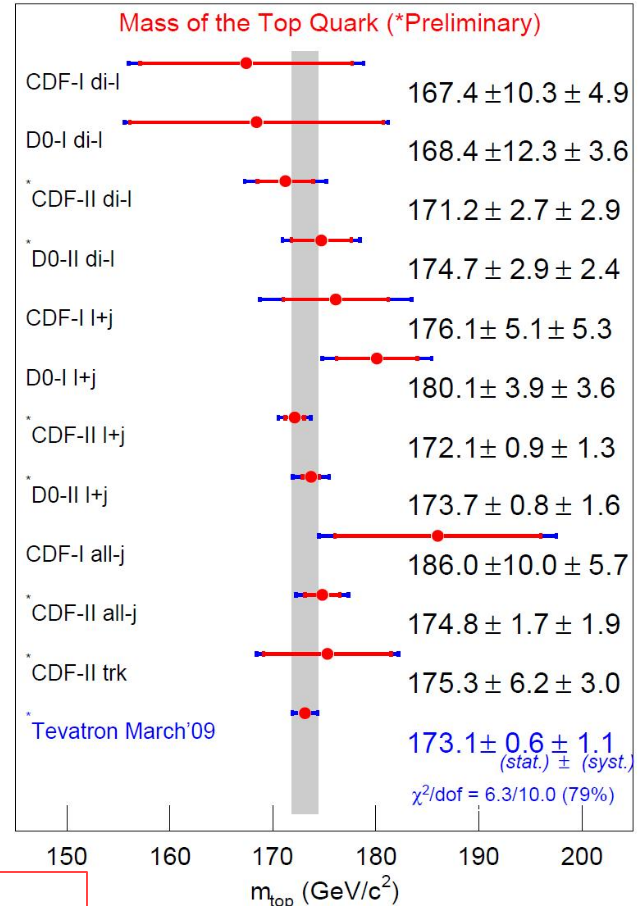
(after all, that's where all top data is ☺)

The Tevatron: the latest numbers

A combination of 11 measurements:

CDF+D0: 0903.2503

	Run I published					Run II preliminary					
	CDF			DØ		CDF				DØ	
	all-j	l+j	di-l	l+j	di-l	l+j	di-l	all-j	trk	l+j	di-l
$\int \mathcal{L}\, dt$	0.1	0.1	0.1	0.1	0.1	3.2	1.9	2.9	1.9	3.6	3.6
Result	186.00	176.10	167.40	180.10	168.40	172.14	171.15	174.80	175.30	173.75	174.66
iJES	0.00	0.00	0.00	0.00	0.00	0.74	0.00	1.64	0.00	0.47	0.00
aJES	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	1.32
bJES	0.60	0.60	0.80	0.71	0.71	0.38	0.40	0.21	0.00	0.07	0.26
cJES	3.00	2.70	2.60	2.00	2.00	0.32	1.73	0.49	0.60	0.00	0.00
dJES	0.30	0.70	0.60	0.00	0.00	0.08	0.09	0.08	0.00	0.84	1.46
rJES	4.00	3.35	2.65	2.53	1.12	0.40	1.90	0.21	0.10	0.00	0.00
lepPt	0.00	0.00	0.00	0.00	0.00	0.18	0.10	0.00	1.10	0.18	0.32
Signal	1.80	2.60	2.80	1.11	1.80	0.34	0.78	0.23	1.60	0.45	0.65
MC	0.80	0.10	0.60	0.00	0.00	0.51	0.90	0.31	0.60	0.58	1.00
UN/MI	0.00	0.00	0.00	1.30	1.30	0.00	0.00	0.00	0.00	0.00	0.00
BG	1.70	1.30	0.30	1.00	1.10	0.50	0.38	0.35	1.60	0.08	0.08
Fit	0.60	0.00	0.70	0.58	1.14	0.16	0.60	0.67	1.40	0.21	0.51
CR	0.00	0.00	0.00	0.00	0.00	0.41	0.40	0.41	0.40	0.40	0.40
MHI	0.00	0.00	0.00	0.00	0.00	0.09	0.20	0.17	0.70	0.05	0.00
Syst.	5.71	5.28	4.85	3.89	3.63	1.35	2.98	1.99	3.11	1.60	2.43
Stat.	10.00	5.10	10.30	3.60	12.30	0.94	2.67	1.70	6.20	0.83	2.92
Total	11.51	7.34	11.39	5.30	12.83	1.64	4.00	2.61	6.94	1.80	3.80



Parameter	Value (GeV/c^2)
$M_t^{\text{all-j}}$	175.1 ± 2.6
$M_t^{\text{l+j}}$	172.7 ± 1.3
$M_t^{\text{di-l}}$	171.4 ± 2.7

Signal includes:
Theory and pdf uncertainties.
Quite small??

The Tevatron

- "Best" channel: lepton + jet.

- They have relatively few top-pair events:

  For example the latest published sample in the (lepton+jet) includes ~ 220 events!

- This is not exactly big statistics (in the usual sense);

  naively, $1/\sqrt{N} \sim 7\% >>$ than the 0.8% from the Tevatron.

So, how is that possible?

Matrix element methods

Some statistics:

The answer involves not-so-popular statistical methods:

REF: PDG '08 - Statistics
- Probability

- Frequentist statistics (the usual one):

  probability is interpreted as the frequency of the outcome of a repeatable experiment.

- Bayesian statistics:

  the interpretation of probability is more general and includes degree of belief (called subjective probability). One can then speak of a probability density function (p.d.f.) for a parameter, which expresses one's state of knowledge about where its true value lies.

Bayes' theorem

$$P(\mathrm{theory}|\mathrm{data}) \propto P(\mathrm{data}|\mathrm{theory})P(\mathrm{theory})$$

Interpretation: *the prior degree of belief is updated by the data from the experiment*

Proof:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

More Statistics

REF: PDG '08 - Statistics

p.d.f. := probability density function

In Bayesian statistics, all knowledge about θ is summarized by the posterior p.d.f. p(θ|x), which gives the degree of belief for θ to take on values in a certain region given the data x.

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{L(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int L(\boldsymbol{x}|\boldsymbol{\theta}')\pi(\boldsymbol{\theta}')\, d\boldsymbol{\theta}'}$$

L(x|θ) - the likelihood function, i.e., the joint p.d.f. for the data given a certain value of θ,

π(θ) - the prior p.d.f. for θ.

Bayesian statistics supplies no unique rule for determining π(θ); this reflects the experimenter's subjective degree of belief about θ before the measurement was carried out

The method of maximum likelihood

How to get L?

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{L(\boldsymbol{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int L(\boldsymbol{x}|\boldsymbol{\theta}')\pi(\boldsymbol{\theta}')\,d\boldsymbol{\theta}'}$$

REF: PDG '08 - Statistics

The method of maximum likelihood

Suppose we have a set of N measured quantities $x = (x_1, \ldots, x_N)$ described by a joint p.d.f. $f(x; \theta)$, where $\theta = (\theta_1, \ldots, \theta_n)$ is set of n parameters whose values are unknown.

The likelihood function is given by the p.d.f. evaluated with the data x, but viewed as a function of the parameters, i.e., $L(\theta) = f(x; \theta)$.

If the measurements x_i are statistically independent and each follow the p.d.f. $f(x; \theta)$, then the joint p.d.f. for x factorizes and the likelihood function is:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} f(x_i; \boldsymbol{\theta})$$

Then:

$$\frac{\partial \ln L}{\partial \theta_i} = 0$$

Gives the maximum likelihood estimators, i.e. $\theta = (\theta_1, \ldots, \theta_n)$

Hint: *θ – is to be m_{top}*

Back to top physics at the Tevatron

References:

Kondo et al: late 80's mid 90's

Dalitz and Goldstein: 90's

See also Adam Gibson, PhD Thesis, '06

FERMILAB-PUB-08-242-E

NOTE: in the following I'll consider only the (lepton+jet) mode!

Experimentalist study events with:

1 lepton + (exactly) 4jets + large missing E_T

At least one jet is required to be tagged as b-jet.

Back to top physics at the Tevatron

Here I follow *arXiv:hep-ph/9802249v1*

They both generate and test LO events only (I'll comment later) :

$$t \to W^+ b \quad \text{and} \quad \bar{t} \to W^- \bar{b}$$

$$\bar{q} + q \to \bar{t} + t$$

$$(a) \quad W^+ \to l^+ \nu_l \quad \text{or} \quad (b) \quad W^+ \to u\bar{d} \,\text{or}\, c\bar{s}$$

$$(a) \quad W^- \to l^- \bar{\nu}_l \quad \text{or} \quad (b) \quad W^- \to d\bar{u} \,\text{or}\, s\bar{c}.$$

Note: an event is defined in 11 dimensional space of kinematical parameters x_i

Example event from the Tevatron:

event		p_x	p_y	p_z	$E(GeV)$
40758	e^+	-94.313	-50.113	48.523	117.306
	jet j1	86.267	26.685	-21.881	92.913
	j2	-26.220	74.310	23.996	82.373
	j3	46.052	47.417	43.659	79.217
	j4	30.613	-22.003	76.790	85.545

Back to top physics at the Tevatron

Here I follow *arXiv:hep-ph/9802249v1*

Step 1:

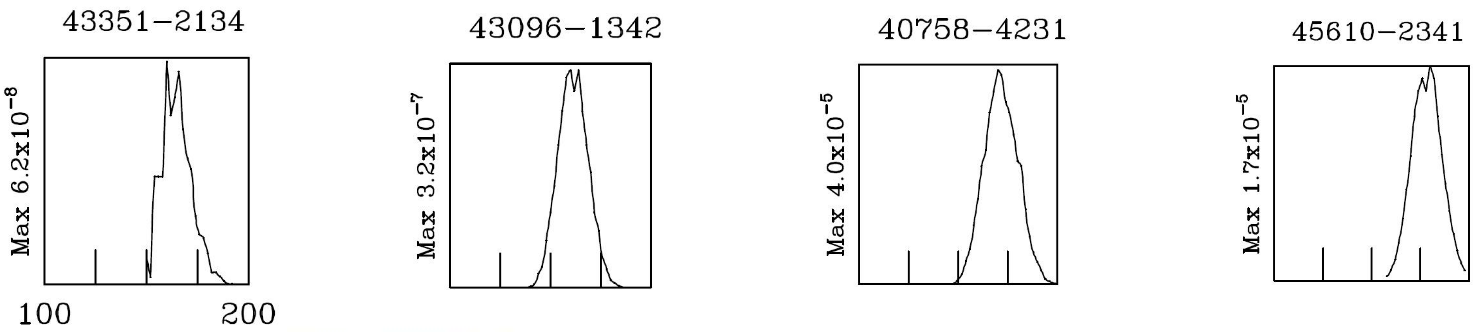
Take the measured configuration of momenta for the final leptons and jets in a single event i and evaluate the probability

$$P_i(m) = P(\text{configuration event } i|m)$$

that these production and decay processes could produce the observed configuration if the top quark mass were m.

Hint: calculated as from LO QCD

Examples of $P_i(m)$ for few Tevatron events:



Back to top physics at the Tevatron

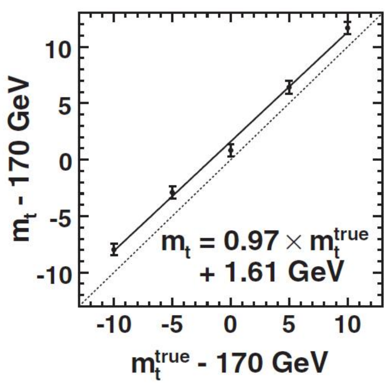
Step 2:
apply Bayes' Theorem:

$$P(m|\text{data set }\{i\}) = \prod_{i=1}^{N} P(\text{event } i|m) \cdot \Phi(m)$$

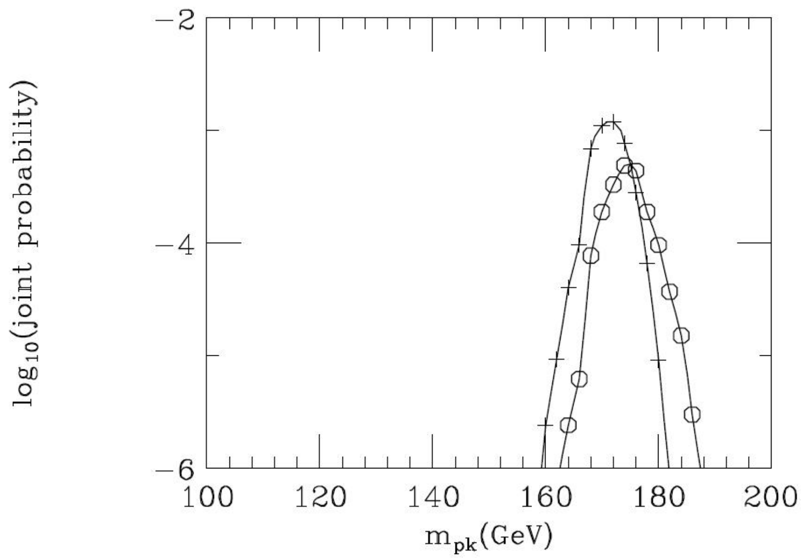
a priori probability that the top mass is m

In practice, what one does is:

- ✓ Construct $P(m) = \prod_{i=1}^{N} P_i(m)$
- ✓ Infer m_{top} from its extremum:



FERMILAB-PUB-08-242-E



arXiv:hep-ph/9802249v1

Conclusions

- Top mass is a fundamental parameter; used in many places outside top physics
- Precise value and definition are required for that!
- At the Tevatron the statistics is small for standard analyses:
  Bayesian approach developed and applied (pretty solid ☺)
- Open questions (that bother me):
- All this assumes we know exactly the distributions (calibrations).

  But that is not so. NLO brings 50% corrections => that is large uncertainty.
  How does that affect the extraction?
- I couldn't find this addressed in any paper.
- For theorists: even if the above is implemented,
  we do not have top-pair production and decay at NLO! And it must be fast!

Top quark: pdf's

Czakon, AM in progress

Comparison of central values for:

- m_{top}=172.4 GeV
- μ=m
- correct exact hard matching coefficients.

$\alpha_s(M_Z)$:

CTEQ 6.6: 0.118
MRST 2006 nnlo: 0.119
MSTW 2008 nnlo: 0.117
MSTW 2008 nlo: 0.120

MRST 2006 nnlo

NLO = 890 pb
NLO+NLL = 918 pb

MSTW 2008 nnlo

NLO = 857 pb
NLO+NLL = 885 pb

5.4 %

1.6 %

CTEQ 6.6

NLO = 844 pb
NLO+NLL = 871 pb

MSTW 2008 nlo

NLO = 906 pb
NLO+NLL = 935 pb

7.3 %