Top mass extraction from dilepton events with emphasis on the theoretical uncertainties

Alexander Mitov

Cavendish Laboratory

UNIVERSITY OF CAMBRIDGE

Based on:

Frederix, Frixione, Mitov; to appear.

Why the top mass?

- ✓ Knowing the top mass has important implications beyond immediate collider physics
  - ✓ Higgs inflation
  - ✓ Vacuum stability in SM and beyond
  - ✓…
- ✓ How well do we know the top mass?
  - ➢ m_{top} is not an observable; cannot be measured directly.
  - ➢ It is extracted indirectly, through the sensitivity of observables to m_{top}

$$\sigma^{\mathrm{exp}}(\{Q\}) = \sigma^{\mathrm{th}}(m_t, \{Q\})$$

  - ✓ The implication: the "determined" value of m_{top} is as sensitive to theoretical modeling as it is to the measurement itself
- ✓ The measured mass is close to the pole mass (top decays …)
- ✓ Lots of activity (past and ongoing). A big up-to-date review:

Juste, Mantry, Mitov, Penin, Skands, Varnes, Vos, Wimpenny '13

The message I'd like to convey: the problem is not "academic"

Example: look at the spread across current measurements

arXiv:1403.4427

- Current World Average: m_{top}= 173.34±0.76 GeV
- New CMS (l+j): m_{top}= 172.04 ± 0.19 (stat.+JSF) ± 0.75 (syst.) GeV. TOP-14-001

✓ Comparable uncertainties; rather different central values!

- This is possible in the context of my discussion: different theory systematics.

To me, the problem of m_{top} extraction should turn from "more precise determination" to better understanding of the theory systematics and their size.

In order to properly understand and estimate the theory systematics we propose a particular observable

$$pp \to t\bar{t} + X$$

$$t \to W + b + X$$

$$W \to \ell + \nu_\ell$$

These are ttbar dilepton events, subject to standard cuts:

$$|\eta_\ell| \leq 2.4 \ , \ |\eta_b| \leq 2.4 \ ,$$

$$p_{T,\ell} \geq 20 \text{ GeV} \ , \ p_{T,b} \geq 30 \text{ GeV}$$

- Construct the distributions from leptons only
- Require b-jets [anti-k_T, R=0.5] within the detector (i.e. integrate over)

The definition of the observable possesses several important properties:

- It is inclusive of hadronic radiation, which makes it well-defined to all perturbative orders in the strong coupling,
- It does not require the reconstruction of the t and/or $\bar{t}$ quarks (indeed we do not even speak of t quark),
- Due to its inclusiveness, the observable is as little sensitive as possible to modelling of hadronic radiation. This feature increases the reliability of the theoretical calculations.

✓ The top mass is extracted from the ***shapes, not normalizations***, of the following distributions:

kinematic distribution
$p_T(\ell^+)$
$p_T(\ell^+\ell^-)$
$M(\ell^+\ell^-)$
$E(\ell^+) + E(\ell^-)$ ← Studied before by: Biswas, Melnikov, Schulze '10
$p_T(\ell^+) + p_T(\ell^-)$

✓ Working with distributions directly is cumbersome.

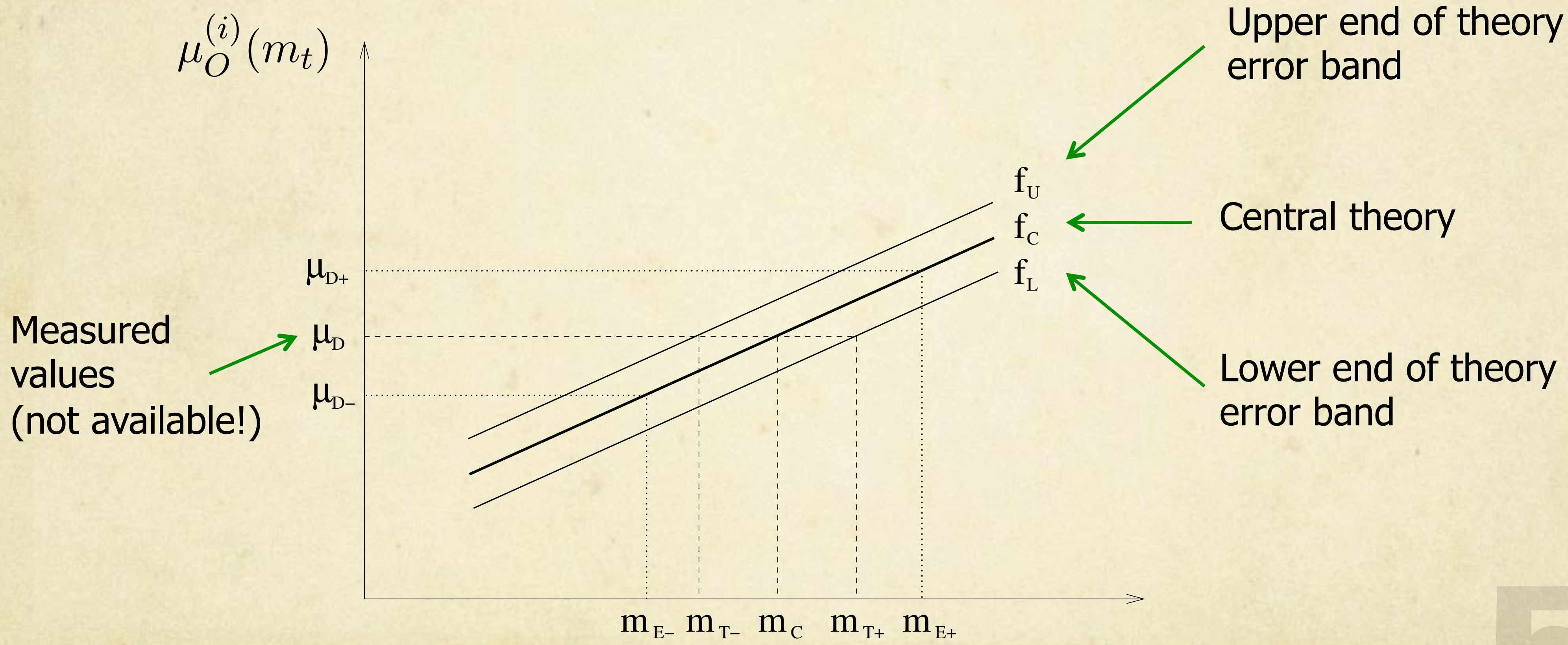
✓ Instead, utilize the first 4 moments of each distribution

$$\sigma = \int d\sigma \qquad \mu_O^{(i)} = \frac{1}{\sigma}\int d\sigma\, O^{\,i} \qquad \mu_O^{(0)} = 1\,, \qquad \mu_O^{(1)} = \langle O \rangle$$

Note: both are subject to cuts (or no cuts); we tried both.

- Here is how it all works:

1) Compute the dependence of the moments $\mu_O^{(i)}(m_t)$ on the top mass
2) Measure the moment
3) Invert 1) and 2) to get the top mass (would be the pole mass, since this is what we use)

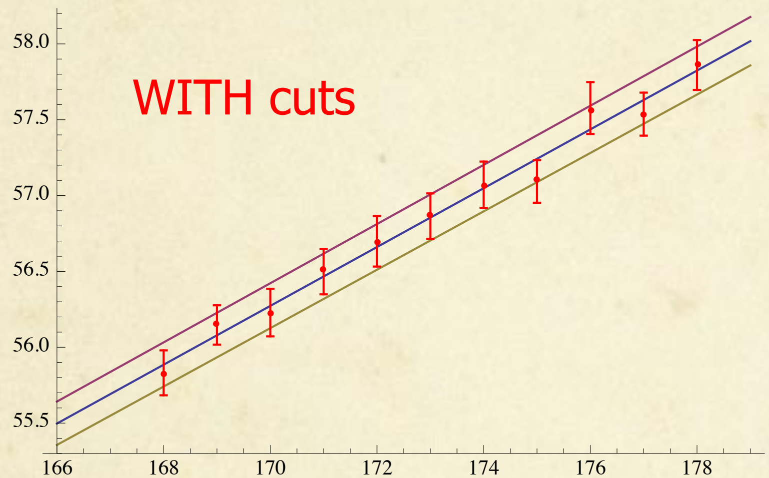


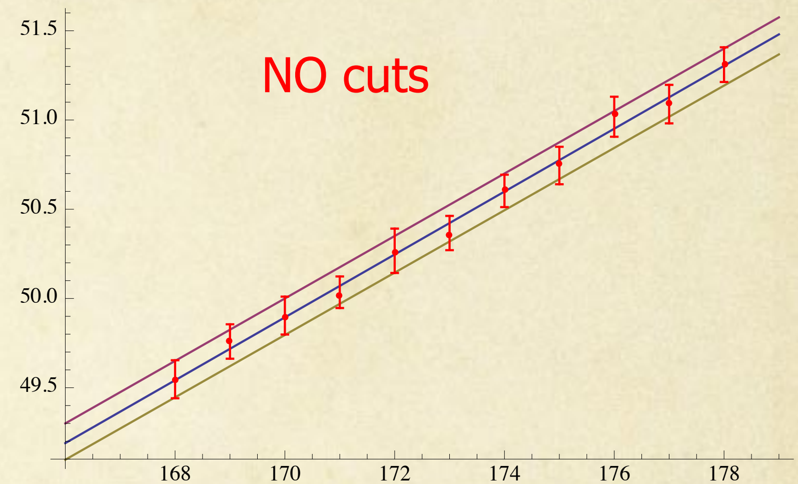
How to compute the theory error band for $\mu_O^{(i)}(m_t)$?

- Compute $\mu_O^{(i)}(m_t)$ for a finite number of m_t values: $m_t = (168, 169, \ldots, 178)$ GeV

  Then get best straight line fit (works well in this range).

Example:
- Single lepton P_T
- Subject to cuts





✓ Errors: pdf and scale variation; restricted independent variation

$$0.5 \leq \xi_F, \xi_R \leq 2 \qquad \xi_{F,R} = \mu_{F,R}/\hat{\mu} \text{ and } \hat{\mu} \text{ is a reference scale}$$

✓ There are statistical fluctuation (from MC even generation) No issue for lower moments 1M events; 30% pass the cuts.

Theory systematics

- We access them by computing the observables in many different ways.
- For a fair (albeit biased) comparison across setups and moments we use pseudodata (PD) generated by us
- Compare the systematics by comparing the top mass "extracted" by each setup from PD.

6 Setups:

label	fixer order accuracy	parton shower/fixed order	spin correlations
1	LO	PS	-
2	LO	PS	MS
3	NLO	PS	-
4	NLO	PS	MS
5	NLO	FO	-
6	LO	FO	-

3 F,R Scales:

$$\hat{\mu}^{(1)} = \frac{1}{2}\sum_i m_{T,i}\ ,\ i \in (t,\bar{t})\,,$$

$$\hat{\mu}^{(2)} = \frac{1}{2}\sum_i m_{T,i}\ ,\ i \in \ \text{final state}\,,$$

$$\hat{\mu}^{(3)} = m_t\,,$$

All is computed with aMC@NLO (with Herwig)

Theory systematics: impact of shower effects

obs.	$m_t^{(3)} - m_t^{(5)}$	$m_t^{(3)} - m_t^{\rm pd}$	$m_t^{(1)} - m_t^{(6)}$	$m_t^{(1)} - m_t^{\rm p}$
1	$-0.35^{+1.14}_{-1.16}$	$+0.12$	$-2.17^{+1.50}_{-1.80}$	-0.67
2	$-4.74^{+1.98}_{-3.10}$	$+11.14$	$-9.09^{+0.76}_{-0.71}$	$+14.19$
3	$+1.52^{+2.03}_{-1.80}$	-8.61	$+3.79^{+3.30}_{-4.02}$	-6.43
4	$+0.15^{+2.81}_{-2.91}$	-0.23	$-1.79^{+3.08}_{-3.75}$	-1.47
5	$-0.30^{+1.09}_{-1.21}$	$+0.03$	$-2.13^{+1.51}_{-1.81}$	-0.67

NLO

LO

label	kinematic distribution
1	$p_T(\ell^+)$
2	$p_T(\ell^+\ell^-)$
3	$M(\ell^+\ell^-)$
4	$E(\ell^+) + E(\ell^-)$
5	$p_T(\ell^+) + p_T(\ell^-)$

label	fixer order accuracy	parton shower/fixed order	spin correlations
1	LO	PS	-
2	LO	PS	MS
3	NLO	PS	-
4	NLO	PS	MS
5	NLO	FO	-
6	LO	FO	-

- Setups 2,3 are anomalous (More later).
- Clearly big impact of NLO corrections (shower matters more at LO).

NOTE: proper PS study would require Pythia etc. Not done here.

Theory systematics: impact of NLO vs LO effects

obs.	$m_t^{(4)} - m_t^{(2)}$	$m_t^{(4)} - m_t^{\rm pd}$	$m_t^{(3)} - m_t^{(1)}$	$m_t^{(3)} - m_t^{\rm pd}$	$m_t^{(5)} - m_t^{(6)}$	$m_t^{(5)} - m_t^{\rm pd}$
1	$+1.16^{+1.43}_{-1.60}$	$+0.41$	$+0.79^{+1.43}_{-1.60}$	$+0.12$	$-1.03^{+1.22}_{-1.43}$	$+0.47$
2	$-2.79^{+1.27}_{-1.65}$	-1.18	$-3.05^{+1.35}_{-1.64}$	$+11.14$	$-7.41^{+1.64}_{-2.72}$	$+15.87$
3	$-0.73^{+3.21}_{-3.45}$	$+0.84$	$-2.18^{+3.03}_{-3.30}$	-8.61	$+0.09^{+2.42}_{-2.91}$	-10.13
4	$+1.74^{+3.27}_{-3.78}$	$+0.16$	$+1.23^{+3.10}_{-3.61}$	-0.23	$-0.70^{+2.79}_{-3.09}$	-0.38
5	$+0.99^{+1.42}_{-1.72}$	$+0.25$	$+0.70^{+1.40}_{-1.72}$	$+0.03$	$-1.13^{+1.23}_{-1.33}$	$+0.33$

PS+MS

PS

-

label	kinematic distribution
1	$p_T(\ell^+)$
2	$p_T(\ell^+\ell^-)$
3	$M(\ell^+\ell^-)$
4	$E(\ell^+) + E(\ell^-)$
5	$p_T(\ell^+) + p_T(\ell^-)$

label	fixer order accuracy	parton shower/fixed order	spin correlations
1	LO	PS	-
2	LO	PS	MS
3	NLO	PS	-
4	NLO	PS	MS
5	NLO	FO	-
6	LO	FO	-

- Setups 2,3 are anomalous (More later).
- Clearly big impact of NLO corrections.

Theory systematics: impact of Spin-Correlations effects

obs.	$m_t^{(4)} - m_t^{(3)}$	$m_t^{(4)} - m_t^{\rm pd}$	$m_t^{(2)} - m_t^{(1)}$	$m_t^{(2)} - m_t^{\rm pd}$
1	$+0.29^{+1.17}_{-1.14}$	$+0.41$	$-0.08^{+1.66}_{-1.96}$	-0.75
2	$-12.32^{+1.62}_{-2.13}$	-1.18	$-12.58^{+0.90}_{-0.94}$	$+1.60$
3	$+9.45^{+2.36}_{-2.16}$	$+0.84$	$+8.00^{+3.74}_{-4.26}$	$+1.57$
4	$+0.39^{+2.93}_{-3.16}$	$+0.16$	$-0.11^{+3.42}_{-4.16}$	-1.58
5	$+0.22^{+1.12}_{-1.28}$	$+0.25$	$-0.06^{+1.65}_{-2.07}$	-0.73

NLO+PS

LO+PS

label	kinematic distribution
1	$p_T(\ell^+)$
2	$p_T(\ell^+\ell^-)$
3	$M(\ell^+\ell^-)$
4	$E(\ell^+)+E(\ell^-)$
5	$p_T(\ell^+)+p_T(\ell^-)$

label	fixer order accuracy	parton shower/fixed order	spin correlations
1	LO	PS	-
2	LO	PS	MS
3	NLO	PS	-
4	NLO	PS	MS
5	NLO	FO	-
6	LO	FO	-

- NOTE setups 2,3 Huge dependence on spin correlations
- NLO corrections make a difference.

"Best" Theory Predictions (NLO+PS+MS): choice of scale and Moment

$m_t^{\rm pd} = 174.32$ GeV

[...] = χ^2 per d.o.f.

$\hat{\mu}^{(1)} = \frac{1}{2}\sum_i m_{T,i}\ ,\ i \in (t,\bar{t})\,,$

$\hat{\mu}^{(2)} = \frac{1}{2}\sum_i m_{T,i}\ ,\ i \in$ final state$\,,$

$\hat{\mu}^{(3)} = m_t\,,$

All 5 observables NLO+PS+MS

scale	$i=1$	$i=1\oplus 2$	$i=1\oplus 2\oplus 3$
1	$174.48^{+0.73}_{-0.77}[5.0]$	$174.55^{+0.72}_{-0.76}[5.0]$	$174.56^{+0.71}_{-0.76}[5.1]$
2	$174.73^{+0.77}_{-0.80}[4.3]$	$174.74^{+0.76}_{-0.79}[4.3]$	$174.91^{+0.75}_{-0.79}[4.1]$
3	$172.54^{+1.03}_{-1.07}[1.6]$	$172.46^{+0.99}_{-1.05}[1.6]$	$172.22^{+0.95}_{-1.04}[1.4]$
$1\oplus 2\oplus 3$	$174.16^{+0.81}_{-0.85}$	$174.17^{+0.80}_{-0.84}$	$174.17^{+0.78}_{-0.84}$

label	kinematic distribution
1	$p_T(\ell^+)$
2	$p_T(\ell^+\ell^-)$
3	$M(\ell^+\ell^-)$
4	$E(\ell^+)+E(\ell^-)$
5	$p_T(\ell^+)+p_T(\ell^-)$

Observables 1,4,5 NLO+PS+MS

scale	$i=1$	$i=1\oplus 2$	$i=1\oplus 2\oplus 3$
1	$174.67^{+0.75}_{-0.77}[3.0]$	$174.67^{+0.75}_{-0.77}[3.0]$	$174.61^{+0.74}_{-0.77}[3.2]$
2	$174.81^{+0.83}_{-0.80}[6.2]$	$174.80^{+0.82}_{-0.80}[6.2]$	$174.85^{+0.82}_{-0.80}[6.1]$
3	$172.63^{+1.85}_{-1.16}[0.2]$	$172.64^{+1.82}_{-1.15}[0.2]$	$172.58^{+1.81}_{-1.15}[0.2]$
$1\oplus 2\oplus 3$	$174.44^{+0.92}_{-0.87}$	$174.44^{+0.92}_{-0.87}$	$174.43^{+0.91}_{-0.87}$

Observable 1 NLO+PS+MS

scale	$i=1$	$i=1\oplus 2$	$i=1\oplus 2\oplus 3$
1	$174.73^{+0.80}_{-0.79}[0.2]$	$174.73^{+0.80}_{-0.79}[0.2]$	$174.72^{+0.80}_{-0.79}[0.2]$
2	$174.78^{+0.90}_{-0.90}[0.6]$	$174.78^{+0.90}_{-0.90}[0.6]$	$174.78^{+0.90}_{-0.90}[0.6]$
3	$172.73^{+2.0}_{-1.2}[0.5]$	$172.73^{+1.96}_{-1.19}[0.5]$	$172.73^{+1.96}_{-1.19}[0.5]$
$1\oplus 2\oplus 3$	$174.46^{+0.99}_{-0.92}$	$174.46^{+0.99}_{-0.92}$	$174.45^{+0.99}_{-0.92}$

Theory systematics: Predictions

observable; setup	$i = 1$	$i = 1 \oplus 2$	$i = 1 \oplus 2 \oplus 3$
all; LO+PS	$187.90^{+0.6}_{-0.6}[428.3]$	$187.71^{+0.60}_{-0.60}[424.2]$	$187.83^{+0.58}_{-0.60}[442.8]$
all; LO+PS+MS	$175.98^{+0.63}_{-0.69}[16.9]$	$176.05^{+0.63}_{-0.68}[17.8]$	$176.12^{+0.61}_{-0.68}[18.9]$
all; NLO+PS	$175.43^{+0.74}_{-0.80}[29.2]$	$176.20^{+0.73}_{-0.79}[30.1]$	$175.67^{+0.73}_{-0.76}[31.2]$
all; $\mathrm{NLO_{FO}}$	$174.41^{+0.72}_{-0.73}[96.6]$	$174.82^{+0.71}_{-0.73}[93.1]$	$175.44^{+0.70}_{-0.68}[94.8]$
all; $\mathrm{LO_{FO}}$	$197.31^{+0.42}_{-0.35}[2496.1]$	$197.19^{+0.42}_{-0.35}[2505.6]$	$197.48^{+0.36}_{-0.35}[3005.6]$
1,4,5; LO+PS	$173.68^{+1.08}_{-1.31}[0.8]$	$173.68^{+1.08}_{-1.31}[0.9]$	$173.75^{+1.08}_{-1.31}[0.9]$
1,4,5; LO+PS+MS	$173.61^{+1.10}_{-1.34}[1.0]$	$173.63^{+1.10}_{-1.34}[1.0]$	$173.62^{+1.10}_{-1.34}[1.0]$
1,4,5; NLO+PS	$174.40^{+0.75}_{-0.81}[3.5]$	$174.43^{+0.75}_{-0.81}[3.5]$	$174.60^{+0.75}_{-0.79}[3.2]$
1,4,5; $\mathrm{NLO_{FO}}$	$174.73^{+0.72}_{-0.74}[5.5]$	$174.72^{+0.71}_{-0.74}[5.6]$	$175.18^{+0.64}_{-0.71}[4.6]$
1,4,5; $\mathrm{LO_{FO}}$	$175.84^{+0.90}_{-1.05}[1.2]$	$175.75^{+0.89}_{-1.05}[1.2]$	$175.82^{+0.89}_{-1.04}[1.2]$

$$m_t^{\mathrm{pd}} = 174.32\ \mathrm{GeV}$$

[...] = χ^2 per d.o.f.

label	kinematic distribution
1	$p_T(\ell^+)$
2	$p_T(\ell^+\ell^-)$
3	$M(\ell^+\ell^-)$
4	$E(\ell^+) + E(\ell^-)$
5	$p_T(\ell^+) + p_T(\ell^-)$

Conclusions

- ✓ New developments have resurrected the interest in knowing m_{top} precisely
  - ✓ Vacuum Stability in SM
  - ✓ Higgs Inflation
- ✓ There are many dedicated hadron collider measurements.
  They return consistent values around m_{top} = 173 GeV
  and uncertainty (mostly on the measurement!) of below 1 GeV.
- ✓ Questions remain: can there be a significant additional theoretical systematics O(1 GeV) ?
- ✓ This is not an abstract problem: m_{top} is not an observable and so is a theoretically defined concept.
- ✓ Proposed an approach, with emphasis on control over theory systematics.
  - ➢ NLO vs LO: O(1 GeV);
  - ➢ Shower effects much smaller at NLO than at LO.
  - ➢ Spin correlations crucial, but depend on the observable.
  - ➢ Awaiting the measurement: O(100k) events exist!
  - ➢ Adding higher moments is not a game changer
  - ➢ Unlikely to be able to use the data to tell which scale choice is 'right'.
  - ➢ Future improvements, notably NNLO, will likely also play an important role.
  - ➢ In some cases the differences are so big that the measurements will easily tell us which way of computing things is right and which is not!