Understanding Data from First Principles

Tilman Plehn

Universität Heidelberg

Bad Neuenahr, September 2018

Theory Summary

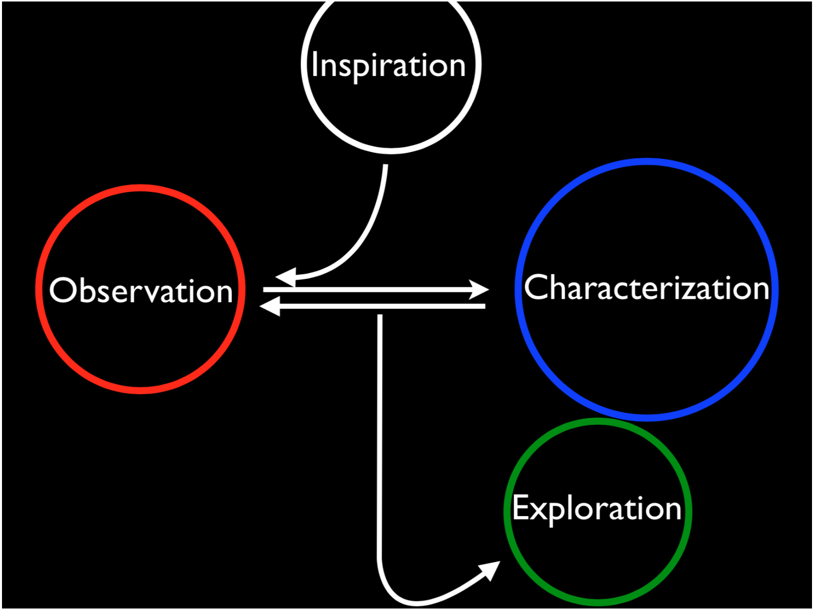
Tim M.P. Tait

University of California, Irvine

Top 2014

October 3 2014

How things used to look



An era of data

Inspiration milestones for the TeV scale

1964 Higgs boson [confirmed]

1967 model of leptons [Weinberg, confirmed]

1974 supersymmetry [Wess-Zumino]

1984 composite Higgs [Kaplan, Georgi, Dimopoulos]

1998 large extra dimensions [Arkani-Hamed, Dimoloulos, Dvali]

1999 small extra dimensions [Randall, Sundrum]

2000 little Higgs [Arkani-Hamed, Schmaltz]

$\Rightarrow$ brilliant people, but nothing beats data

From data to renormalizable Lagrangians

– LHC data is described by QFT [fundamental physics, Standard Model is not a model]

– perturbative $SU(3) \times SU(2) \times U(1)$ works [as does resummation]

– BSM physics exists [neutrino, matter-antimatter, coupling unification]

$\rightarrow$ bottom-up search for weakly coupled BSM physics [precision physics]

$\rightarrow$ bottom-up search for heavy BSM physics [EFT, precision physics]

$\Rightarrow$ once we see something, we build a new QFT description

An era of data

Inspiration milestones for the TeV scale

1964 Higgs boson [confirmed]

1967 model of leptons [Weinberg, confirmed]

1974 supersymmetry [Wess-Zumino]

1984 composite Higgs [Kaplan, Georgi, Dimopoulos]

1998 large extra dimensions [Arkani-Hamed, Dimoloulos, Dvali]

1999 small extra dimensions [Randall, Sundrum]

2000 little Higgs [Arkani-Hamed, Schmaltz]

$\Rightarrow$ brilliant people, but nothing beats data

Role of theory

– simulation tools [amazing progress]

– precision predictions [more amazing progress]

– analysis ideas [not so amazing progress]

– interpretation frameworks [hardly amazing progress]

Pushing back modelling

The way we compare data and theory [likelihood-free inference]

1- simulate events

hard process: perturbative field theory
QCD jets: resummed perturbative field theory
hadronization: QCD-inspired modelling
pile-up, detector,...: whatever works

2a compare simulated events with measured events

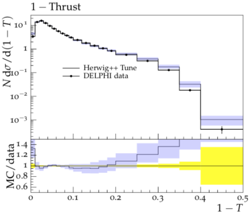
2b other interface, but the same: fiducial or unfolding,....

3- theory-inspired distibutions etc just illustration? What's the weakest link?

First-principle simulations [Marek Schönherr]

– QCD whenever possible
– be clear about input
– include error bars

⇒ keep track of modelling



Tune uncertainties
Tune performed by minimising

$$\chi^2(\vec{\mathbf{x}}) = \frac{1}{N}\sum_{i\in\mathcal{O}} w_b \frac{(\mathrm{MC}_i(\vec{\mathbf{x}}) - \mathrm{Data}_i)^2}{\sigma^2_{i,\mathrm{Data}}}$$

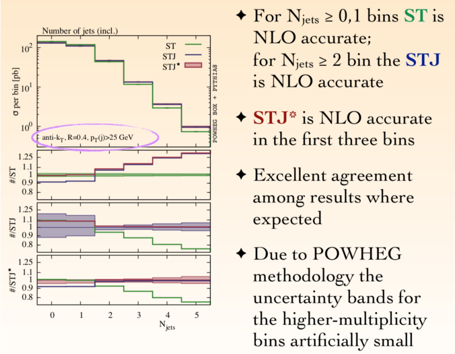
Define Eigentunes as set of eigenvectors in χ^2 potential, normalised to some predefined $\Delta\chi^2$ (same as PDF)

Jets everywhere

LHC physics is multi-jet physics [Rikkert Frederix]

– multi-jet simulation key to successful LHC simulations
– *t*-channel single top with forward jet
  jet structure relevant for $t\bar{t}$ rejection
– simulation of second jet important
  combination with parton shower crucial to describe data
– cool technology: map phase spaces using a neural net

⇒ deep learning the latest addition to our toolbox



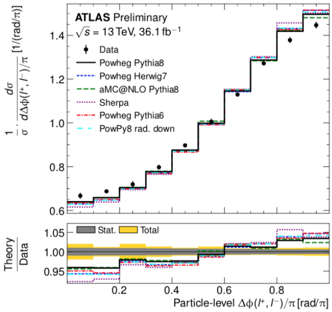
Top pairs

Precision predictions in top sector in great shape

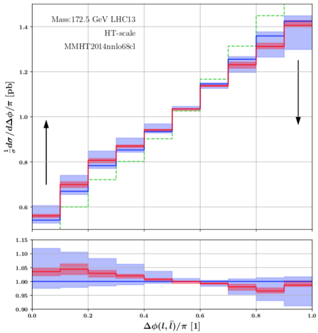
– NLO QCD also off-shell
– NNLO QCD distributions with decays [numerics limiting factor?]
– NLO e-w distributions feasible [more complex than NLO QCD]

Kinematic distributions to NNLO [Rene Poncelet]

– closing gap between of traditional fixed-order rates and MC
– even better: anomalous distributions



NWA @ NNLO predictions



Top pairs

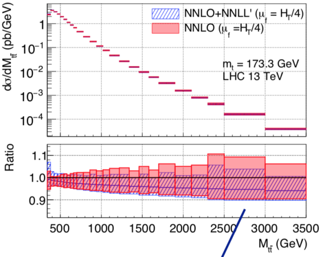
Precision predictions in top sector in great shape

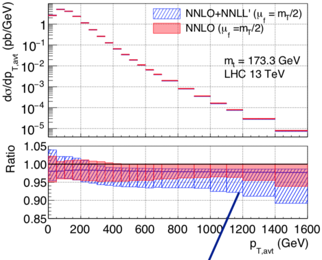
- NLO QCD also off-shell
- NNLO QCD distributions with decays [numerics limiting factor?]
- NLO e-w distributions feasible [more complex than NLO QCD]

Combining with soft resummation [Li Lin Yang]

NNLO+NNLL′

Czakon, Ferroglia, Heymes, Mitov, Pecjak, Scott, Wang, **LLY**: 1803.07623





Resummation reduces scale variation

Resummation softens the spectrum

Top pairs with jets

Pushing multi-leg fixed order [Stefano Pozzorini]

- $t\bar{t}b\bar{b}$ (most) interesting part of $t\bar{t}jj$ [top-Higgs overlap]
- phase space complex, hard process maybe $t\bar{t}$
- α_s^4 leading to large K-factor and scale dependence
- massive b-jet radiation? multi-scale process with cut-off?

$\Rightarrow$ a tale of logarithms?

Natural scale choice for inclusive $\sigma_{t\bar{t}b\bar{b}}$

- for $m_b = m_t$ the natural choice is $\mu_R = m_t$
- natural generalisation $\mu_R = \sqrt{m_b m_t} \Rightarrow$ good convergence for $1 \leq m_t/m_b \leq 36$

m_b [GeV]	4.75	15.7	52.1	172.5
$\sqrt{m_b m_t}$ [GeV]	28.7	52.1	94.8	172.5
$K(N_b \geq 0)$	1.14	1.24	1.32	1.35

$\Rightarrow$ $\langle \mu_{R,\text{def}} \rangle \simeq 66\,\text{GeV}$ should be reduced by factor 2–3 to match $\sqrt{m_b m_t} = 28.7\,\text{GeV}$

Top pairs with stuff

Combining strong and electroweak corrections [Marco Zaro]

– tree-level Feynman diagrams with $\alpha\alpha_s^2$, $\alpha^2\alpha_s$, α^3 [remember V+jets]

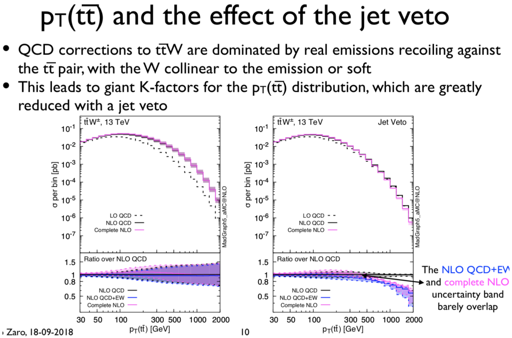
– QCD and ew corrections mixing classes

– t-W scattering at NLO mediated by y_t

– 'hard process' phase-space-dependent, p_{tt} formally safe, but especially bad

– large K-factor controlled by jet veto

$\Rightarrow$ reproducing data with MC is dangerous, what's our story?



Top mass from first principles

Top mass as link from Lagrangian to data [Paolo Nason, Andre Hoang]

– new physics inspiration:

electroweak precision data
vacuum stability
hierarchy problem

impact from future LHC measurement?

– actually, field theory question
related to large-n_F loop diagrams [renormalons]

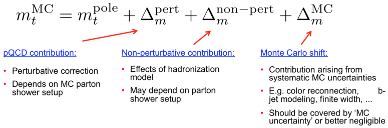
Tick the correct statements:

- ☐ Direct top mass measurements measure the Pole Mass.
- ☐ Direct top mass measurements measure the Monte Carlo Mass.
- ☐ Direct top mass measurements measure the Monte Carlo Mass. but you can pretend that it is the pole mass, just inflate the error a bit.
- ☐ The top is the only SM particle with more than one mass.
- ☐ You should use only leptons to avoid hadronization uncertainty.
- ☐ You should use at least NLO calculations to measure the pole mass.
- ☐ The top pole mass has renormalons, you should stay away from it.
- The MC mass differs from the pole mass by
  ☐ terms of order $m\alpha_s$; ☐ terms of order Λ_{QCD}; ☐ terms of order $\alpha_s\Gamma_t$.
- The Pole Mass renormalon ambiguity is
  ☐ ≈ 1GeV; ☐ ≈ 250 MeV; ☐ ≈ 200 MeV; ☐ ≈ 110 MeV.

Top mass from first principles

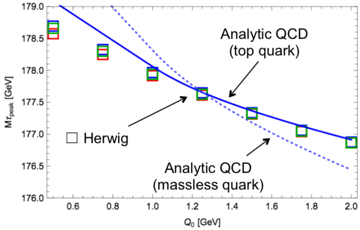
Top mass as link from Lagrangian to data [Paolo Nason, Andre Hoang]

– new physics inspiration:
  electroweak precision data
  vacuum stability
  hierarchy problem
  impact from future LHC measurement?
– actually, field theory question



– progress from cross talk: QCD vs Monte Carlo
  pole mass vs shower cut-off controllable

⇒ pushing back modelling, again!



Top Yukawa measurement

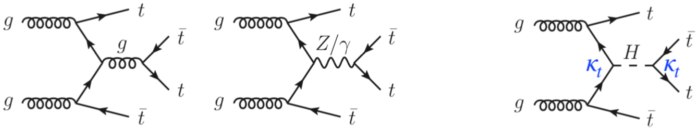
Width-independent top Yukawa measurement [Qing-Hong Cao]

– off-shell $t\bar{t}H$, $H \to t\bar{t}$

– avoid loop-induced production with its model dependence

– go for seriously off-shell

$\Rightarrow$ cancellation feature of Standard Model

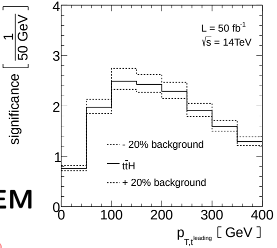


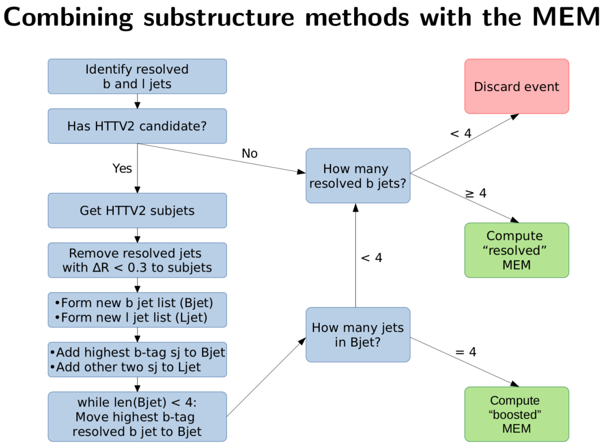
$$\sigma(t\bar{t}t\bar{t}) = \sigma^{\rm SM}(t\bar{t}t\bar{t})_{g/Z/\gamma} + \kappa_t^2 \sigma^{\rm SM}_{\rm int} + \kappa_t^4 \sigma^{\rm SM}(t\bar{t}t\bar{t})_H$$

Boosted MEM

Combining MEM with top tagging [Maren Meinhard]

– search for $t\bar{t}H$ combinatorics-limited [calling for MEM]
– signal-background significance from boosted regime
⇒ never stop trying new things!



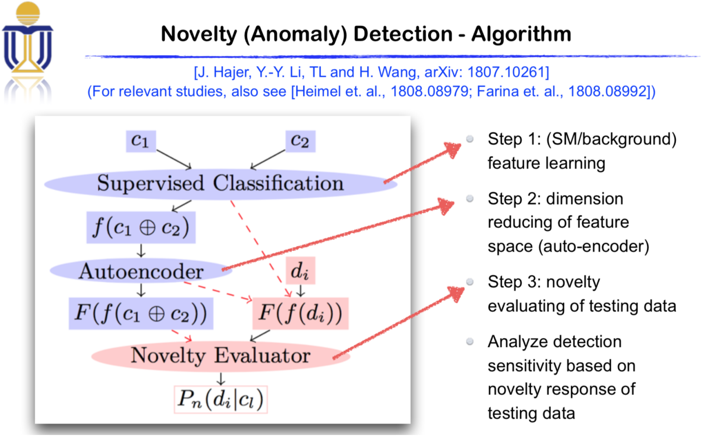


Machine learning

So what if we really do not know what we are searching for? [Tao Liu]

– yes, we have been doing ML forever, but we are not cutting edge anymore
– established ML answer: auto-encoder
– application: search for anomalous kinematics

⇒ just a diagnosis tool, don't be Chefarzt about it

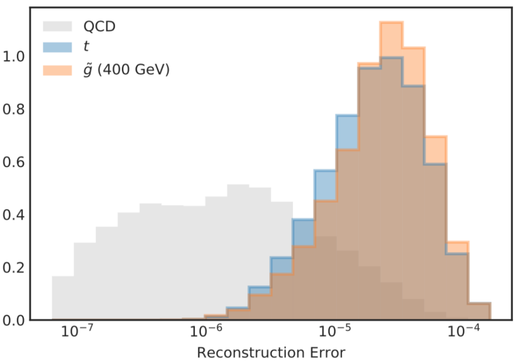


More machine learning

If events are too hard, use jets [David Shih; coordinated with Heimel, Kasieczka, TP, Thompson]

– lots of training data in all phase space regions
– check for jets which do not look like QCD
– uncertainties the next big challenge
– de-correlate/control the jet mass for control regions
$\Rightarrow$ training on data, searching in data means no modelling, yet again

Can use reconstruction error as an anomaly threshold.



Sub-jet physics

Top EFT

Bottom-up analysis in QFT framework: effective theory

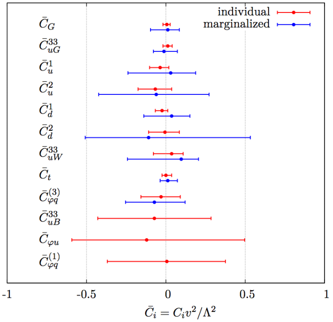
– 59 operators, assuming good-taste symmetries
– affecting rates (coupling modifiers) and kinematics (Lorentz structures)
– best framework for global analysis, leading to SMEFT
– renormalization no problem, at least in top sector
– for details, check endless Higgs-related discussions

⇒ where are the ATLAS/CMS results??

Constraining top quark effective theory in the LHC Run II era

The TopFitter Collaboration

Andy Buckley, Christoph Englert, James Ferrando, David J. Miller, Liam Moore, Michael Russell, and Chris D. White



Top-bottom link

Understanding EFT analyses [Seth Moortgat]

– upgrading $t\bar{t}b\bar{b}$ from background to signal

– categorize EFT effects in terms of physics

– flavor-EFT fit established and very constraining

– $SU(2)$ doublet means t_L talks to bottom sector

$\Rightarrow$ combining machine learning and EFT!

4. Learning the effective operators

Case study: can a NN learn to distinguish between operators with left-handed top-quark currents (t_L) and right-handed top-quark currents (t_R)? If so, can we use this to improve limits on the Wilson coefficients?

A shallow neural network was constructed that combines 18 kinematical variables to predict one of ***three output classes***. From the network outputs, two discriminators are built; one to distinguish between SM and EFT in general, and another one to distinguish t_L from t_R.

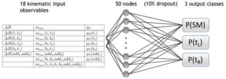
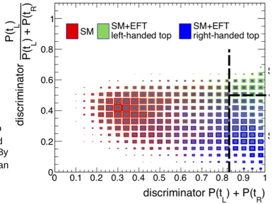


Figure (right): the x-axis represents the SM vs EFT discriminator, whereas the y-axis represents the t_L vs t_R discriminator. SM contributions are located to the left, whereas the EFT contributions are located to the upper right (t_L) and lower right (t_R) corners. By combining limits in dedicated signal regions, we can improve the limits/confidence intervals!

Scenario 1: observation of the SM

Scenario 2: observation of an EFT signal



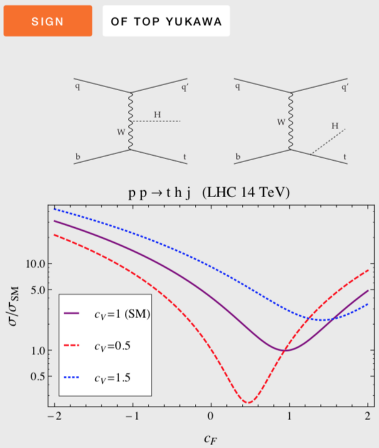
Top-Higgs link

Single-top plus Higgs production [Roberto Franceschini]

– tough analysis, like single top, but with less rate

– correlation $t\bar{t}H$ and WWH couplings [orthogonal to $H \to \gamma\gamma$]

$\Rightarrow$ in case you find single top too easy



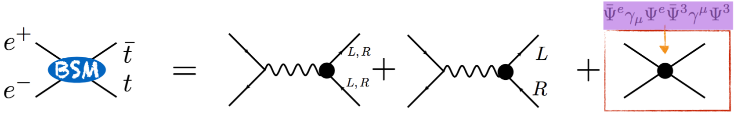
Best of EFT worlds

CLIC: precision and energy [Francesco Riva]

- EFT sensitive to v^2/Λ^2 in rate [LEP1]
- same EFT sensitive to p^2/Λ^2 in tails [LHC]
- learn from LHC for the future: $e^+e^- \to t\bar{t}$ at high energies

$\Rightarrow$ in our era of data it's too far away for me...

$\Rightarrow$ ...but: shows that EFT analyses only need first and last bins



$$\sim c_{(\Psi)^4} \frac{E^2}{M^2}$$

- maximal E-growth
- interference

What will we learn?

CLIC: $\sigma \sim 20\,\text{fb}$ @ $3\,\text{TeV}$ $\quad 2\text{ab}^{-1}$ $\longrightarrow$ $\delta\sigma/\sigma|_{stat} \sim 1\%$

- $M \sim 30 \frac{\sqrt{c_{(\Psi)^4}}}{g}\,\text{TeV}$

Even more data in the future

Data-driven analyses with full theory

– what limits our first-principle simulations?
– which precision calculations do we need next?
– which analysis ideas get us closer to data?
– how do we interpret searches without models?

– ATLAS/CMS: who is the audience for your papers?
– how can someone understand modern analyses?

$\Rightarrow$ we need to work on our story; sorry, Tim, but BSM inspiration is broken

Even more data in the future

Data-driven analyses with full theory

– what limits our first-principle simulations?
– which precision calculations do we need next?
– which analysis ideas get us closer to data?
– how do we interpret searches without models?

– ATLAS/CMS: who is the audience for your papers?
– how can someone understand modern analyses?

$\Rightarrow$ we need to work on our story; sorry, Tim, but BSM inspiration is broken

All 'field in crisis' talk is bullshit, but HL-LHC has a problem [Jan Kieseler]

1- particle physics has plenty of data
2- our standard theory works, if anything, too well
3- at least I have a big goal: dark matter
4- dead new physics models are good news

– but we should fight for the best students
let young people with new ideas run
become inventive again ourselves
avoid death by boredom