Theory Uncertainties from Nuisance Parameters.

Frank Tackmann

Deutsches Elektronen-Synchrotron

SCET 2019 Workshop
San Diego, March 27, 2019

[arxiv:1904.sooon]

Theory Uncertainties and Correlations.

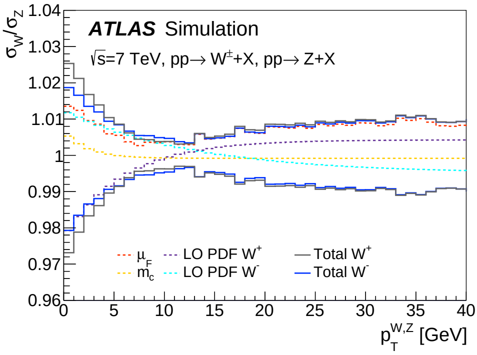
Reliable theory uncertainties are essential for any precision studies and interpretation of experimental measurements

- Especially when theory uncertainties $\gtrsim$ experimental uncertainties
- Correlations can have significant impact
  - In fact, whenever one combines more than a single measurement, one should ask how the theory uncertainties in the predictions for each measurement are correlated with each other
  - Correlations among different points in a resummed spectrum
  - Correlations between predictions for different Q, processes, observables, ...
- So far we have (mostly) been skirting the issue
  - However, experimentalists have to treat theory uncertainties like any other systematic uncertainty, and in absence of anything better they have to make something up based on naive scale variations
  - In likelihood fits, some (possibly enveloped) scale variation impact will get treated as a free nuisance parameter and floated in the fit

Example: Measurement of the W Mass.

Small $p_T^W < 40\,\mathrm{GeV}$ is the relevant region for m_W

- Needs very precise predictions for p_T^W spectrum
- $\simeq 2\%$ uncertainties in p_T^W translate into $\simeq 10\,\mathrm{MeV}$ uncertainty in m_W
- Direct theory predictions for p_T^W are insufficient



⇒ Strategy: Exploit precisely measured Z p_T spectrum to get best possible description for W

- Regardless how precisely $\mathrm{d}\sigma(W)/\mathrm{d}p_T$ can be calculated directly, one always wants to exploit Z data to maximize precision

Extrapolating from Z to W

$$\underbrace{\frac{\mathrm{d}\sigma(W)}{\mathrm{d}p_T}}_{\text{needed}} = \underbrace{\left[\frac{\mathrm{d}\sigma(Z)}{\mathrm{d}p_T}\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{\mathrm{d}\sigma(W)/\mathrm{d}p_T}{\mathrm{d}\sigma(Z)/\mathrm{d}p_T}\right]_{\text{theory}}}_{\substack{\text{calculate precisely}\\ \text{theory uncertainties cancel}}}$$

- Ratio is just a proxy
  - More generally: Combined fit to both processes
  - Tuning Pythia on Z and using it to predict W is one example of this

Extrapolating from Z to W

$$\underbrace{\frac{d\sigma(W)}{dp_T}}_{\text{needed}} = \underbrace{\left[\frac{d\sigma(Z)}{dp_T}\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{d\sigma(W)/dp_T}{d\sigma(Z)/dp_T}\right]_{\text{theory}}}_{\substack{\text{calculate precisely} \\ \text{theory uncertainties cancel}}}$$

- Ratio is just a proxy
  - More generally: Combined fit to both processes
  - Tuning Pythia on Z and using it to predict W is one example of this
- Crucial Caveat: Cancellation fundamentally relies on theory correlations
  - Take 10% theory uncertainty on $d\sigma(W)$ and $d\sigma(Z)$
    - $\rightarrow$ 99.5% correlation yields 1% uncertainty on their ratio
    - $\rightarrow$ 98.0% correlation yields 2% uncertainty on their ratio – 2$\times$ larger!
- One of many examples, this happens whenever experiments extrapolate from some control region or process to the signal region

Theory Correlations.

Correlations only come from common sources of uncertainties

- Straightforward for unc. due to input parameters ($\alpha_s(m_Z)$, PDFs, ...)

What to do about perturbative theory uncertainties?

- ✗ Scale variations are not quantitatively reliable to begin with
- ✗ Moreover, they are inherently ill-suited for correlations
  - ✗ Scales are not physical parameters with an uncertainty that can be propagated, they simply specify a particular perturbative scheme
  - ✗ They are not the underlying source of uncertainty, i.e., they do not become better known at higher order
  - ✗ Taking an envelope is not a linear operation and so does not propagate
  - ✗ Trying to decide how to correlate scale variations (e.g. between processes) is really just a bandaid, but not addressing the real problem

Theory Correlations.

Correlations only come from common sources of uncertainties

- Straightforward for unc. due to input parameters ($\alpha_s(m_Z)$, PDFs, ...)

What to do about perturbative theory uncertainties?

✗ Scale variations are not quantitatively reliable to begin with

✗ Moreover, they are inherently ill-suited for correlations

  ✗ Scales are not physical parameters with an uncertainty that can be propagated, they simply specify a particular perturbative scheme

  ✗ They are not the underlying source of uncertainty, i.e., they do not become better known at higher order

  ✗ Taking an envelope is not a linear operation and so does not propagate

  ✗ Trying to decide how to correlate scale variations (e.g. between processes) is really just a bandaid, but not addressing the real problem

✗ Even the most sophisticated profile scale variations are insufficient

  ✗ The profile shapes are designed to turn off resummation and match to fixed-order, not to capture correlations in the spectrum

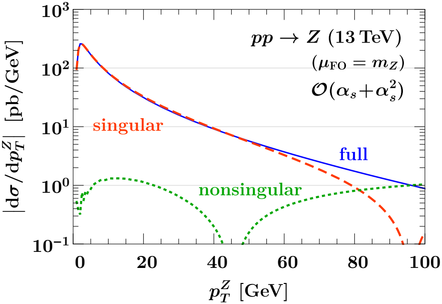
  ✗ See e.g. inconsistent uncertainties from spectrum vs. cumulant scales

Power Expansion.

Define scaling variable $\tau \equiv p_T^2/m_V^2, \mathcal{T}_0/m_V, \ldots$ and expand in powers of τ

$$
\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}\tau} = \delta(\tau) &+ \alpha_s\Big[\frac{\ln\tau}{\tau} + \frac{1}{\tau} + \delta(\tau) + f_1^{\mathrm{nons}}(\tau)\Big] \\
&+ \alpha_s^2\Big[\frac{\ln^3\tau}{\tau} + \frac{\ln^2\tau}{\tau} + \frac{\ln\tau}{\tau} + \frac{1}{\tau} + \delta(\tau) + f_2^{\mathrm{nons}}(\tau)\Big] \\
&+ \quad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \ddots + \ldots\Big] \\
= \quad &\mathrm{d}\sigma^{(0)}/\mathrm{d}\tau + \mathcal{O}(\tau)/\tau
\end{aligned}
$$

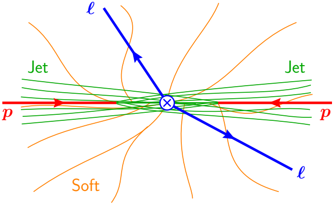
- For small $\tau \ll 1$
  - Logarithmic terms completely dominate perturbative series
  - Their all-order structure is actually simpler and more universal, which allows their resummation
  - Also holds the key for a rigorous treatment of theory correlations



Factorization and Resummation.

Leading-power spectrum factorizes into
hard, collinear, and soft contributions, e.g. for p_T

$$\frac{d\sigma^{(0)}}{d\vec{p}_T} = \sigma_0\, H(Q,\mu) \int d^2\vec{k}_a\, d^2\vec{k}_b\, d^2\vec{k}_s \times B_a(\vec{k}_a, Qe^{Y}, \mu, \nu)\, B_b(\vec{k}_b, Qe^{-Y}, \mu, \nu) \times S(\vec{k}_s, \mu, \nu)\, \delta(\vec{p}_T - \vec{k}_a - \vec{k}_b - \vec{k}_s)$$



- Each function is a renormalized object with an associated RGE
  - Structure depends on type of variable but is universal for all hard processes

$\Rightarrow$ Dependence on p_T and Q is fully determined to all orders by a coupled system of differential equations

- Their solution leads to resummed predictions
- Each resummation order (only) requires as ingredients anomalous dimensions and boundary conditions entering the RG solution

Simplest Example: Multiplicative RGE.

All-order RGE and its solution

$$\mu \frac{\mathrm{d}H(Q,\mu)}{\mathrm{d}\mu} = \gamma_H(Q,\mu)\, H(Q,\mu)$$

$$\Rightarrow \qquad H(Q,\mu) = H(Q) \times \exp\Bigl[\int_Q^\mu \frac{\mathrm{d}\mu'}{\mu'} \gamma_H(Q,\mu')\Bigr]$$

Necessary ingredients

- Boundary condition

$$H(Q) = 1 + \alpha_s(Q)\, h_1 + \alpha_s^2(Q)\, h_2 + \cdots$$

- *Anomalous dimension*

$$\gamma_H(Q,\mu) = \alpha_s(\mu)\bigl[\Gamma_0 + \alpha_s(\mu)\,\Gamma_1 + \cdots\bigr] \ln\frac{Q}{\mu} + \alpha_s(\mu)\bigl[\gamma_0 + \alpha_s(\mu)\,\gamma_1 + \cdots\bigr]$$

$\Rightarrow$ Resummation is determined by coefficients of three fixed-order series

- True regardless of how RGE is solved in more complicated cases

Theory Nuisance Parameters.

Perturbative series at leading power is determined to all orders by a coupled system of differential equations (RGEs)

→ Each resummation order only depends on a few semi-universal parameters

→ Unknown parameters at higher orders are the actual sources of perturbative theory uncertainty

	boundary conditions			anomalous dimensions			
order	h_n	s_n	b_n	γ_n^h	γ_n^s	Γ_n	β_n
LL	h_0	s_0	b_0	—	—	Γ_0	β_0
NLL′	h_1	s_1	b_1	γ_0^h	γ_0^s	Γ_1	β_1
NNLL′	h_2	s_2	b_2	γ_1^h	γ_1^s	Γ_2	β_2
$\mathrm{N^3LL'}$	h_3	s_3	b_3	γ_2^h	γ_2^s	Γ_3	β_3
$\mathrm{N^4LL'}$	h_4	s_4	b_4	γ_3^h	γ_3^s	Γ_4	β_4

- Basic Idea: Treat them as theory nuisance parameters
  - ✓ Vary them independently to estimate the theory uncertainties
  - ✓ Impact of each independent nuisance parameter is fully correlated across all kinematic regions and processes
  - ✓ Impact of different nuisance parameters is fully uncorrelated
- Price to Pay: Calculation becomes quite a bit more complex

Numerous Advantages.

Immediately get all benefits of parametric uncertainties

- ✓ Encode correct correlations
- ✓ Can be propagated straightforwardly
  - Including Monte Carlo, BDTs, neural networks, ...
- ✓ Can be consistently included in a fit and constrained by data
  - Even okay to use control measurements to reduce theory uncertainties
  - Due to central-limit theorem, total theory uncertainty becomes Gaussian

Numerous Advantages.

Immediately get all benefits of parametric uncertainties

- ✓ Encode correct correlations
- ✓ Can be propagated straightforwardly
  - Including Monte Carlo, BDTs, neural networks, ...
- ✓ Can be consistently included in a fit and constrained by data
  - Even okay to use control measurements to reduce theory uncertainties
  - Due to central-limit theorem, total theory uncertainty becomes Gaussian

Additional theory benefits compared to scale variations

- Uncertainties can be evaluated in one space and propagated to another (Fourier conjugate, cumulant, spectrum)
- Can do partial orders and fully exploit all known higher-order information
  - Can account for new structures appearing at higher order
- Fully factorizes the uncertainties
  - Can study perturbative convergence at level of individual building blocks
  - Much safer against accidental underestimates due to multiple parameters

How to Vary What.

- Level 1: At given order vary parameters around their known values

$$c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\, c_2 + \cdots\big] \;\to\; c_0 + \alpha_s(\mu)(c_1 + \tilde{\theta}_1)$$

  - Simpler but perhaps less robust

- Level 2: Implement the full next order in terms of unknown parameters

$$c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\, c_2 + \cdots\big] \;\to\; c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\, \theta_2\big]$$

  - More involved, but also more robust, allowing for maximal precision

- In general, can have combination of both

How to Vary What.

- Level 1: At given order vary parameters around their known values

$$c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\,c_2 + \cdots\big] \;\to\; c_0 + \alpha_s(\mu)(c_1 + \tilde{\theta}_1)$$

  - Simpler but perhaps less robust

- Level 2: Implement the full next order in terms of unknown parameters

$$c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\,c_2 + \cdots\big] \;\to\; c_0 + \alpha_s(\mu)\big[c_1 + \alpha_s(\mu)\,\theta_2\big]$$

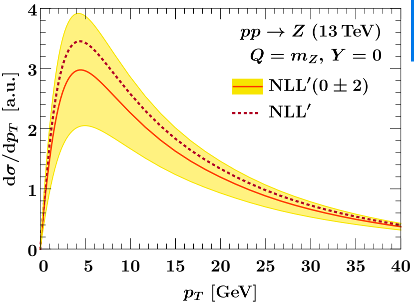
  - More involved, but also more robust, allowing for maximal precision

- In general, can have combination of both

Note: Some parameters are actually functions of additional variables

- E.g. beam function constants, auxiliary dependences (jet radius, ...)
- In principle, one needs to parametrize an unknown function
  - Can e.g. expand/parametrize in terms of appropriate functional basis
  - Compared to scale variations, choices are now explicit and testable

$Z\ p_T$ Spectrum.



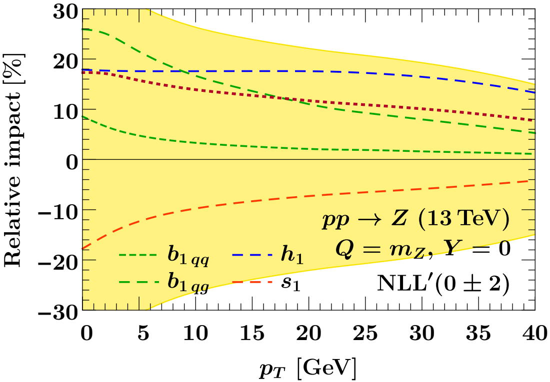
For illustration use

- Level 1: $\tilde{\theta}_i = (0 \pm 0.25) \times c_i$
- Level 2: $\theta_i = (0 \pm 2) \times c_i$

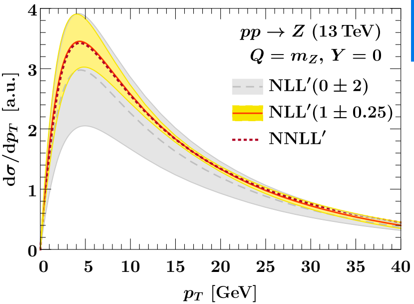
  (with the true values for c_i)

Relative impact of different nuisance parameters

- h_1
- b_1: $q \to q,\ g \to q$
- s_1



$Z\ p_T$ Spectrum.



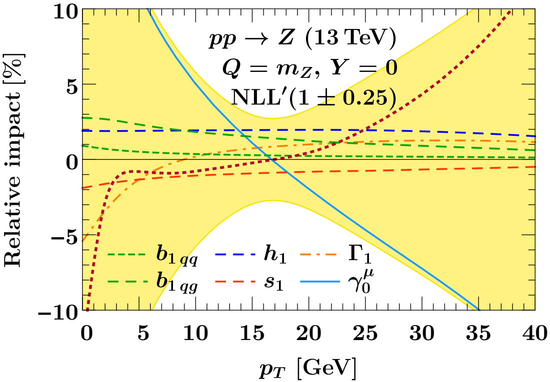
For illustration use

- Level 1: $\tilde{\theta}_i = (0 \pm 0.25) \times c_i$
- Level 2: $\theta_i = (0 \pm 2) \times c_i$

  (with the true values for c_i)

Relative impact of different nuisance parameters

- h_1
- γ_0^μ
- b_1: $q \to q$, $g \to q$
- Γ_1
- s_1

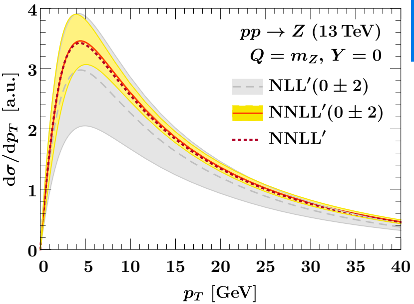


$Z\ p_T$ Spectrum.

For illustration use

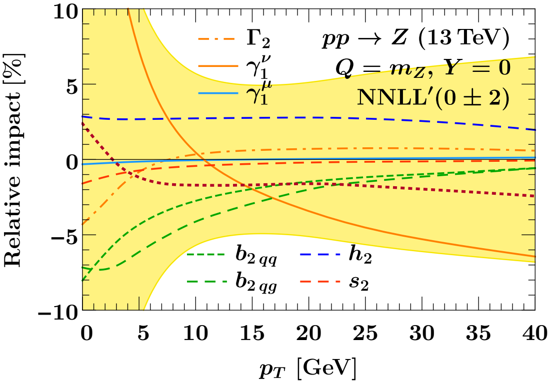
- Level 1: $\tilde{\theta}_i = (0 \pm 0.25) \times c_i$
- Level 2: $\theta_i = (0 \pm 2) \times c_i$

  (with the true values for c_i)

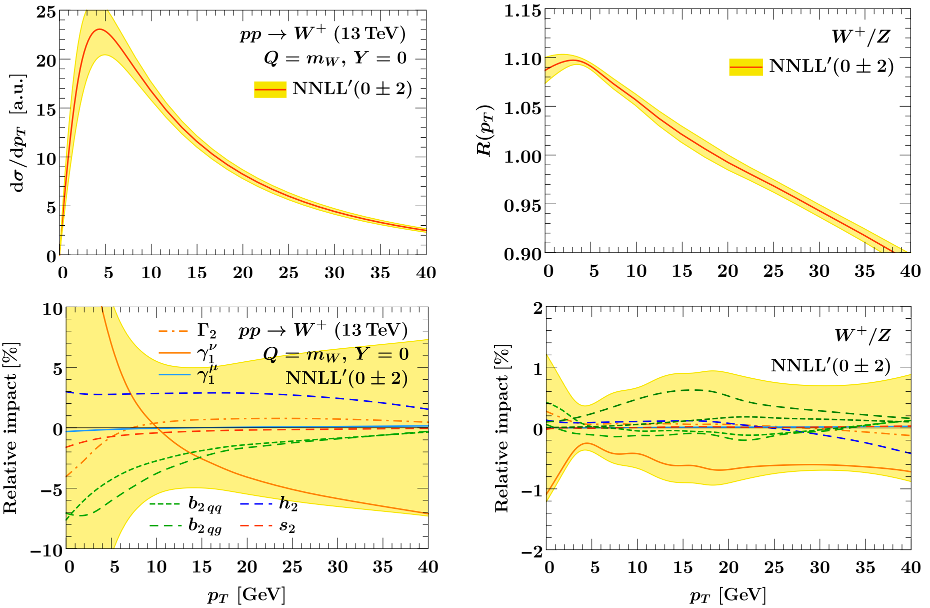


Relative impact of different nuisance parameters

- h_2
- γ_1^μ
- b_2: $q \to q$, $g \to q$
- Γ_2
- γ_1^ν
- s_2

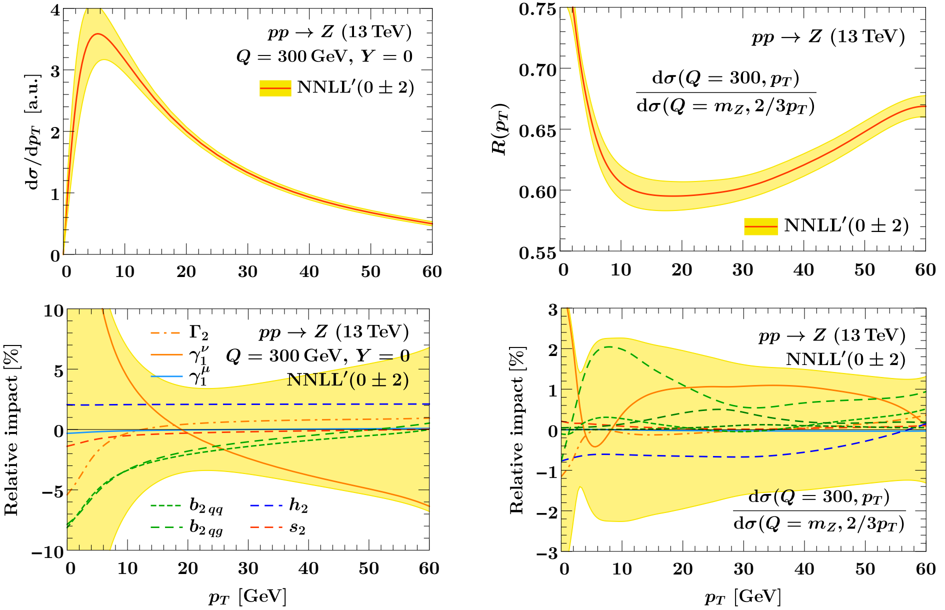


W vs. Z.



$pp \to W^+$ (13 TeV)
$Q = m_W,\ Y = 0$
NNLL$'(0 \pm 2)$
$\mathrm{d}\sigma/\mathrm{d}p_T$ [a.u.]
W^+/Z
$R(p_T)$
Γ_2
γ_1^ν
γ_1^μ
$b_{2\,qq}$
$b_{2\,qg}$
h_2
s_2
Relative impact [%]
p_T [GeV]

Drell-Yan at High Q vs. Z Pole.



Summary.

A theory prediction without an uncertainty is about as useful as a measurement without an uncertainty

- Uncertainties need to be reliable (small is not good enough ...)

Theory nuisance parameters overcome many problems of scale variations

- Allow to rigorously quantify pert. theory uncertainties and correlations
- Encode correct correlations
  - Between different p_T values, Q values, partonic channels, hard processes
  - Between different variables ($\vec{p}_T$, p_T^{jet}, $\mathcal{T}_0$, τ, C, ...),
  - Multi-differential cases, cases with auxiliary measurements, ...
- Can be propagated straightforwardly
  - Including Monte Carlo, BDTs, neural networks, ...
  - Crucial for consistent treatment of theory uncertainties by experiments

$\Rightarrow$ A plethora of applications to explore ...