Theory Uncertainties and Correlations from Theory Nuisance Parameters.

Frank Tackmann

Deutsches Elektronen-Synchrotron

SCET workshop
Salamanca, April 16, 2024

European Research Council
Established by the European Commission

[arXiv:240x.yyyyy]

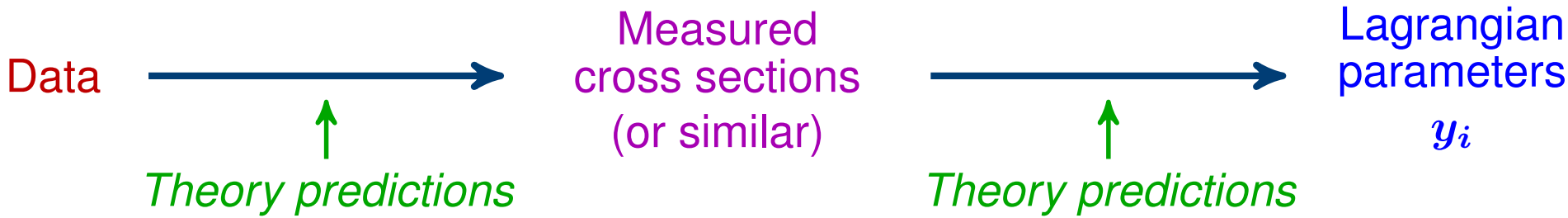
Classification of Systematic Uncertainties.

[Pekka Sinervo – PHYSTAT 2003]

- **The Good**:
  Your own calibrations → basically statistical
- **The Bad**:
  Using other peoples results, poorly modelled data or analysis technique, model assumptions, ...
- **The Ugly**:
  Different theoretical estimates, theory with limited number of terms, ...

Outline.

Theory Uncertainties

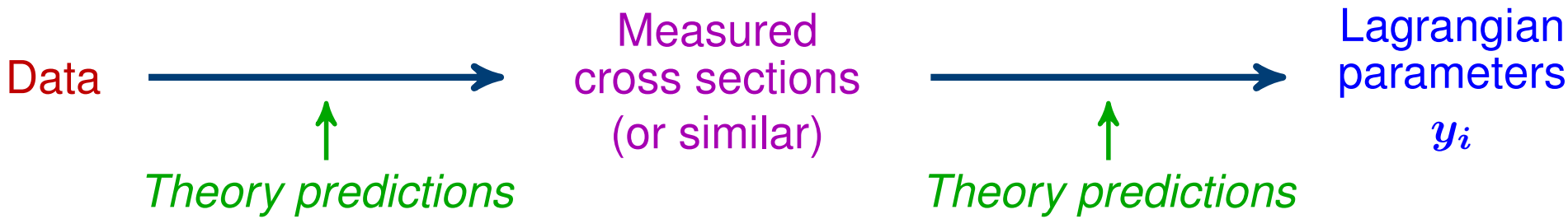


- We always compare some measured quantity to its theory prediction

$$f^{\text{measured}} = f^{\text{predicted}}(y_i) \qquad \Rightarrow \qquad y_i$$

- The *theory uncertainty* is due to the fact that the formula for $f(y_i)$ is not exact but derived in some approximation
  - Note: It is *not* the inexact knowledge of parameters needed in the (otherwise exact) formula

Theory Uncertainties



- We always compare some measured quantity to its theory prediction

$$f^{\text{measured}} = f^{\text{predicted}}(y_i) = f(y_i) \pm \Delta f \quad \Rightarrow \quad y_i \pm \Delta y_i$$

- The *theory uncertainty* is due to the fact that the formula for $f(y_i)$ is not exact but derived in some approximation
  - Note: It is *not* the inexact knowledge of parameters needed in the (otherwise exact) formula

$\Rightarrow$ The Challenge: How to account for the inexactness of the formula itself?

- Usually, we quote some "uncertainty" Δf and a resulting Δy_i, but their precise actual meaning is unclear
- The theory uncertainty is "ugly" compared to other systematics because a priori there is no auxiliary measurement to improve inexactness
- So let's change that ...

Theory Correlations.

Correlations can be crucial once several predictions are used in combination

- Prototype of many data-driven methods or any type of combined fit

$$\underbrace{f(y_i)}_{\text{wanted}} = \underbrace{\left[g(y_j)\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{f(y_i) \pm \Delta f}{g(y_j) \pm \Delta g}\right]_{\text{predicted}}}_{\text{theory uncertainties cancel}}$$

  - Cancellation of theory uncertainties is often assumed or taken for granted
  - But obviously relies crucially on precise correlation between Δf and Δg

Theory Correlations.

Correlations can be crucial once several predictions are used in combination

- Prototype of many data-driven methods or any type of combined fit

$$\underbrace{f(y_i)}_{\text{wanted}} = \underbrace{\left[g(y_j)\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{f(y_i) \pm \Delta f}{g(y_j) \pm \Delta g}\right]_{\text{predicted}}}_{\text{theory uncertainties cancel}}$$

  - Cancellation of theory uncertainties is often assumed or taken for granted
  - But obviously relies crucially on precise correlation between Δf and Δg
  - For example: Take a 10% uncertainty for both Δf and Δg, then

a correlation of	99.5%	98%	95.5%	87.5%
yields a reduction by a factor of	10	5	3.33	2
and an uncertainty on the ratio of	1%	2%	3%	5%

$\Rightarrow$ The Challenge: How to account for correlation between Δf and Δg?

  - Depends on the extent to which inexactness in f and g are related

Scale Variations.

The ugly past and present

Our Standard Estimation Method.

- We're expanding in a small quantity α and calculate first few corrections

$$f(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3\,\alpha^3 + \cdots$$

- To estimate inexactness due to unknown higher orders we make a variable transformation

$$\alpha = \alpha(\tilde{\alpha}) = \tilde{\alpha} + b_0\,\tilde{\alpha}^2 + \mathcal{O}(\tilde{\alpha}^3)$$

  and perform the expansion in $\tilde{\alpha}$ instead

$$f(\tilde{\alpha}) \equiv f(\alpha(\tilde{\alpha})) = f_0 + f_1\,\tilde{\alpha} + \left[f_2 + f_1\,b_0\right]\tilde{\alpha}^2 + \cdots$$

- And finally take the difference as the "uncertainty"

$$\Delta f = f(\alpha) - f(\tilde{\alpha}) = f_1\,b_0\,\alpha^2 + \mathcal{O}(\alpha^3)$$

Our Standard Estimation Method.

In summary

$$f(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + \mathcal{O}(\alpha^3)$$

$$f^{\text{predicted}} = f_0 + f_1\,\alpha \pm \Delta f$$

with $$\Delta f = f_1\,b_0\,\alpha^2 + \mathcal{O}(\alpha^3)$$

We effectively account for inexactness by approximating $f_2 \approx f_1\,b_0$

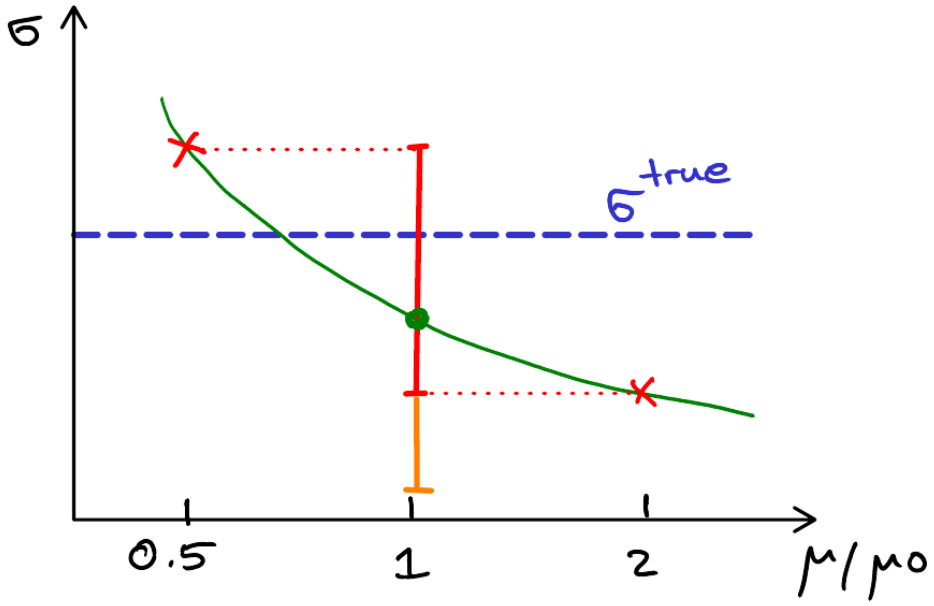
✓ Resulting estimated Δf is indeed $\mathcal{O}(\alpha^2)$

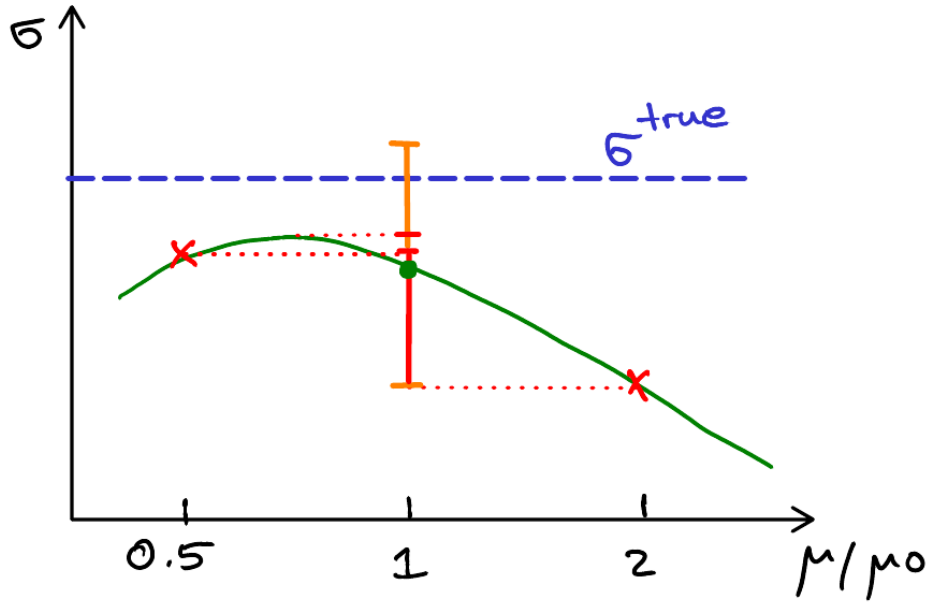
- Including α^2 term in the prediction we would get $\Delta f \sim \mathcal{O}(\alpha^3)$, and so on

✗ But nothing guarantees that this is a good approximation (often it is not)

- f_2 often has more complex internal structure than f_1
- But b_0 is just a (rather arbitrary) constant and usually the same for any f

Scale Variations.





$$\alpha \equiv \alpha_s(\mu_0)\,, \qquad \tilde{\alpha} \equiv \alpha_s(\mu)\,, \qquad b_0 \sim \beta_0 \ln \frac{\mu}{\mu_0}$$

Corresponds to a continuous choice of variable transformation

- μ (or b_0) is *not* an actual parameter with a true value that f depends on
  - No value for it might ever capture the true result (happens regularly)
  - At higher order, uncertainty reduces *only because* μ (or b_0) becomes less relevant and *not because* it somehow becomes better known

What About Correlations?

$$f(\alpha) = f_0 + f_1\,\alpha \pm \Delta f \qquad \text{with} \qquad \Delta f = f_1\,b_0\,\alpha^2 + \cdots$$
$$g(\alpha) = g_0 + g_1\,\alpha \pm \Delta g \qquad \text{with} \qquad \Delta g = g_1\,b_0\,\alpha^2 + \cdots$$

How are Δf and Δg correlated?

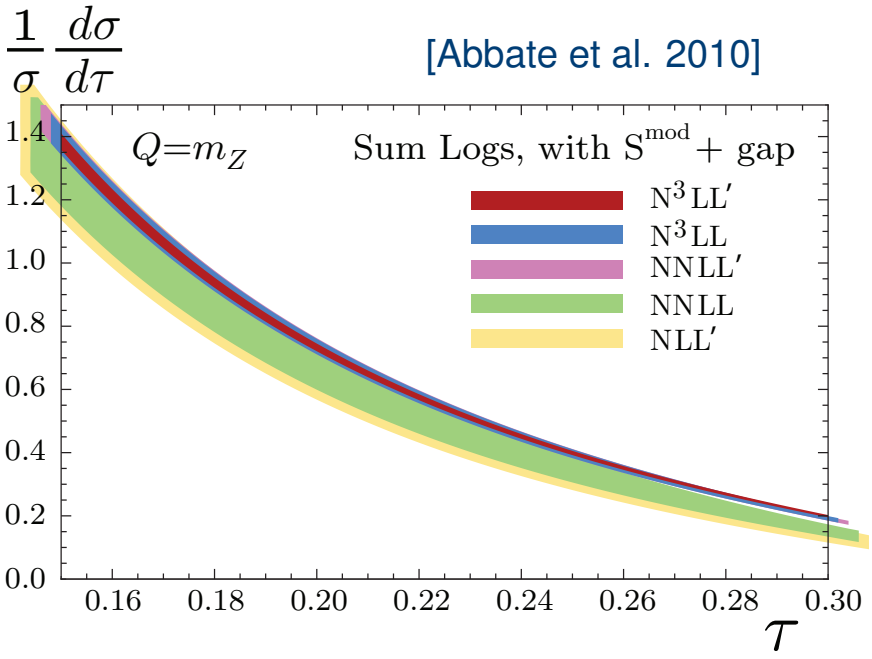
- We don't know – the method simply does not tell us
  - Correlations require a common uncertain parameter
    (or more generally a common source of uncertainty)
  - b_0 (or μ) is not a common or uncertain parameter, we just made it up
  - ✗ A priori, scale variations *do not* imply correct correlations

- Best we can do is *assume* some theoretically motivated but still ad hoc correlation model that we impose on Δf and Δg

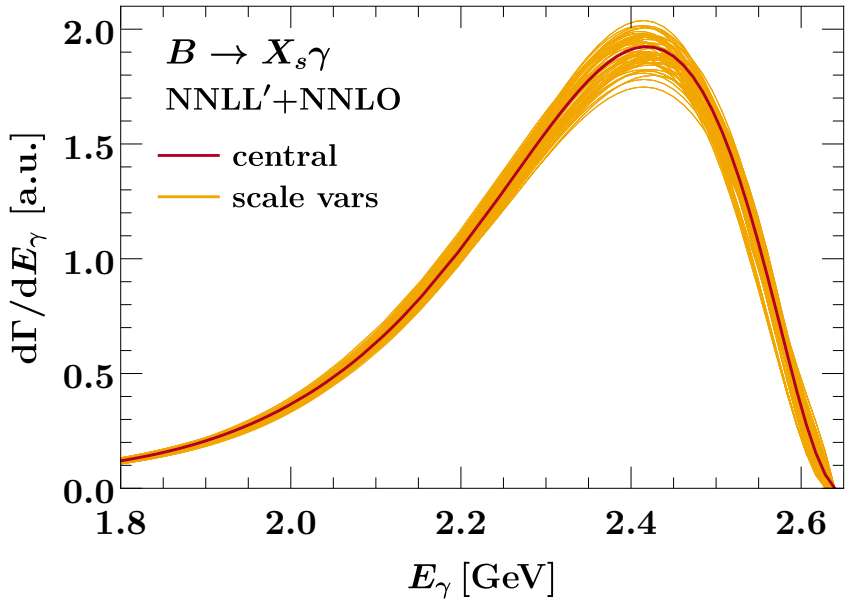
$\Rightarrow$ Probably the most severe shortcoming of scale variations

Scale Variations for Differential Spectrum.

[Abbate et al. 2010]



[Bernlochner et al. 2020]



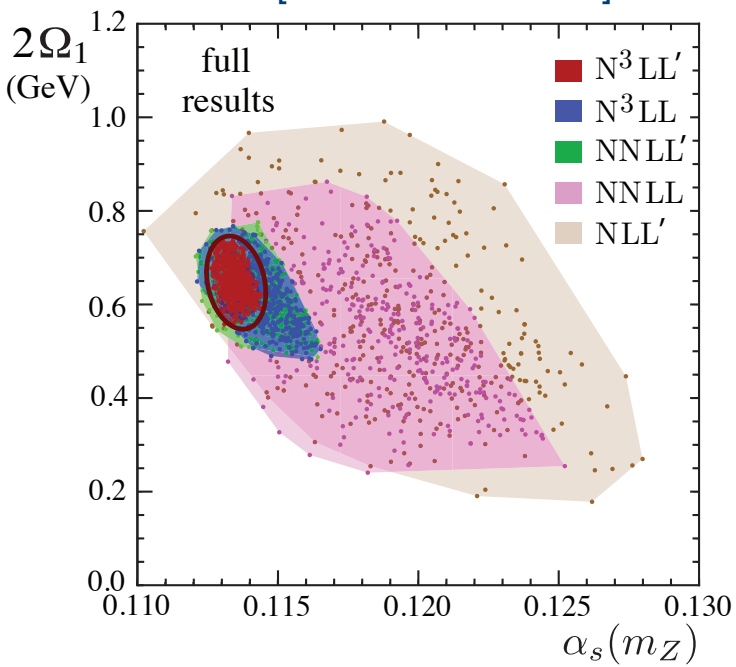
Now $f(\alpha; x)$ is some differential spectrum in x, e.g. τ, E_γ, p_T^Z, ...

- Its $\Delta f(x)$ comes from envelope of various scale variations
  - Take previous $f(\alpha) \equiv f(\alpha; x_1)$ and $g(\alpha) = f(\alpha; x_2)$ to be spectrum at different points in x
  - We don't know their correct correlation
  - ✗ A priori, scale variations *do not* imply correct shape uncertainties

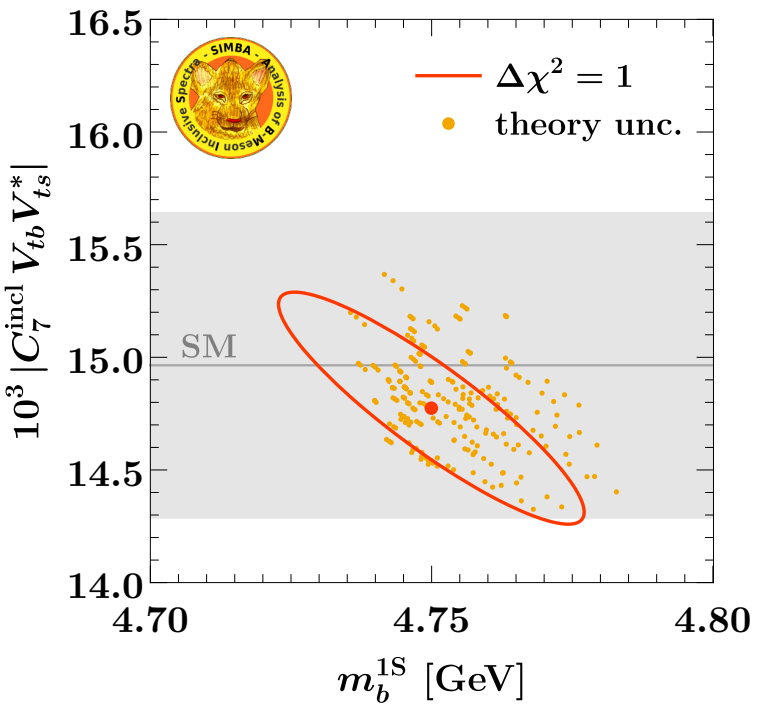
$\Rightarrow$ How to interpret and propagate this envelope?

Propagating Scale Variation Envelope.

[Abbate et al. 2010]



[Bernlochner et al. 2020]



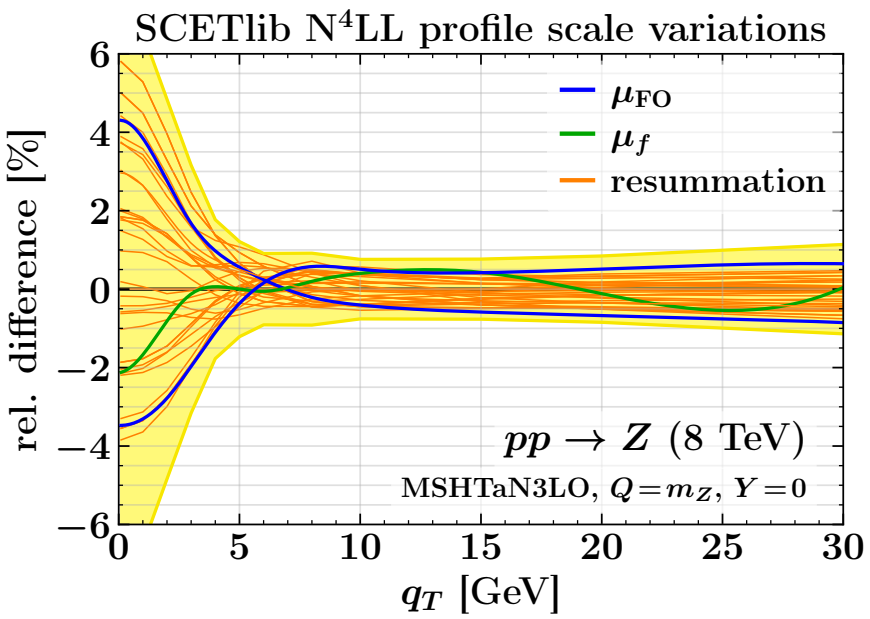
- Repeat fit for each individual scale variation, take envelope of the result
  - Each variation provides a different trial correlation model for $\Delta f(x)$
    - → To reduce chance of accidental underestimation use as many variations as possible to “fill out” the envelope

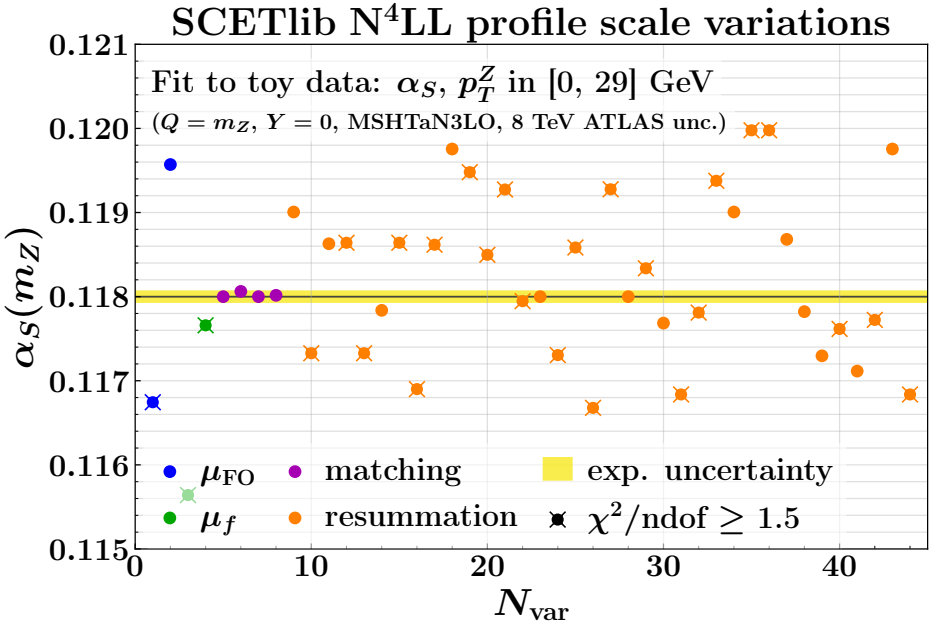
✓ Maintains shape of each variation allowing for possible anticorrelations

✗ Still uses scale as if it were a parameter that could imply correlation

✗ Each trial $\Delta f(x)$ is 100% (anti)correlated across x

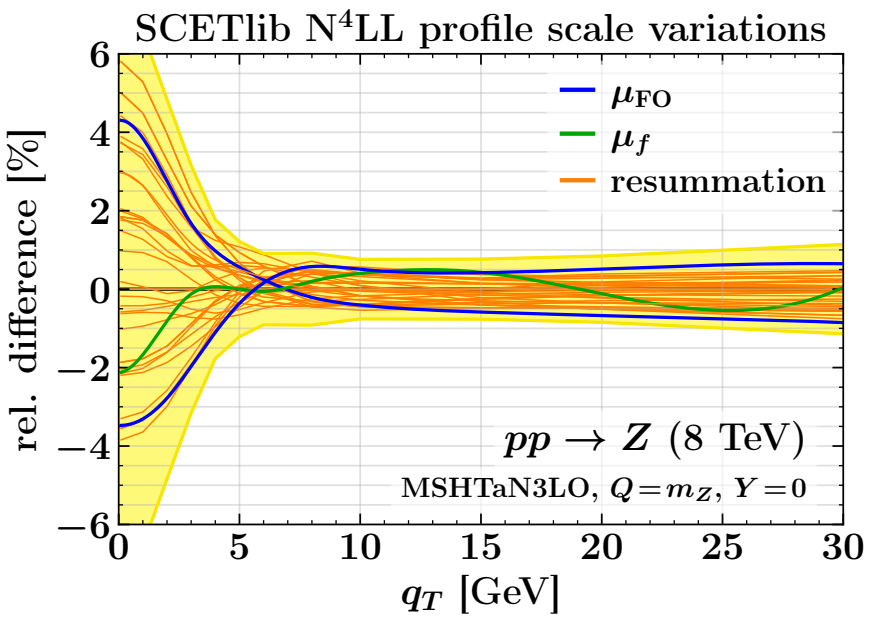
α_s from Z p_T Spectrum [sneak preview from Giulia's talk next].

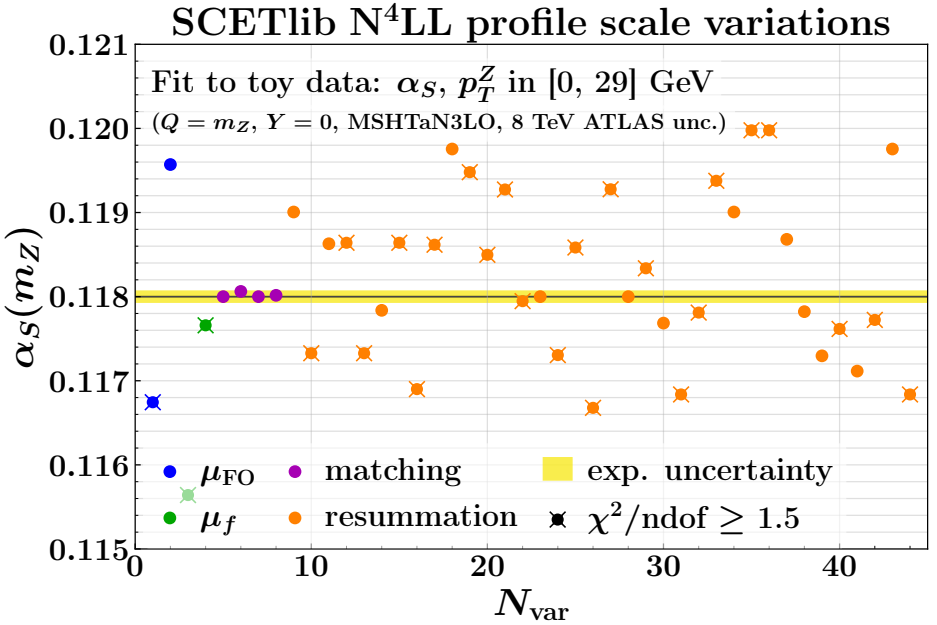




- Theory correlations in $q_T \equiv p_T^Z$, i.e. shape of theory variation within the band, strongly impacts the result (as expected ...)
- Fit should be able to decide whether to allow or constrain some theory excursion vs. changing α_s to compensate

α_s from Z p_T Spectrum [sneak preview from Giulia's talk next].





- Theory correlations in $q_T \equiv p_T^Z$, i.e. shape of theory variation within the band, strongly impacts the result (as expected ...)
- Fit should be able to decide whether to allow or constrain some theory excursion vs. changing α_s to compensate

$\Rightarrow$ Upshot: Scale variations are just not sufficient for this purpose

- Especially when theory uncertainties become dominant or limiting unc.

Theory Nuisance Parameters.

The non-ugly future

What We Should be Doing.

$$f(\alpha) = f_0 + f_1\,\alpha \underbrace{+ f_2\,\alpha^2 + f_3\,\alpha^3 + \mathcal{O}(\alpha^4)}_{\text{source of the theory uncertainty}}$$

Identify the actual source of uncertainty: unknown $f_2, f_3, \ldots$

- $f(\alpha)$ is only a function of α $\Rightarrow$ f_n are just numbers
  - Simplest case
- $f(\alpha) = f(\alpha; x)$ $\Rightarrow$ $f_n(x)$ are functions
  - If dominant/leading x dependence is known $\rightarrow$ can be reduced to previous
- $f(\alpha) = f(\alpha; x_1, x_2, \ldots)$ $\Rightarrow$ $f_n(x_1, x_2, \ldots)$ are N-dim. functions
  - Need honest/realistic estimate of an unknown function ...

What We Should be Doing.

N^{1+1}LO: $f^{\text{predicted}}(\alpha) = f_0 + f_1\,\alpha + f_2(\theta_2)\,\alpha^2$

Parametrize the (leading) source of uncertainty $f_2 \equiv f_2(\theta_2)$

- In terms of unknown but well-defined parameters θ_n, which are the *theory nuisance parameters (TNPs)*
  - Simplest: Use f_2 itself: $f_2(\theta_2) \equiv \theta_2$
  - Better: Account for known internal structure of f_2 (color, partonic channels, ...)
- Include the full parametrized next term in the prediction $+ f_2(\theta_2)\,\alpha^2$
  - Sufficient to include the next term, we always assume that expansion converges, so f_3 is not yet relevant
    ($\rightarrow$ We assume a renormalon-free series)
  - When f_2 becomes known (or tightly constrained), need to include f_3

    N^{2+1}LO: $f^{\text{predicted}}(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3(\theta_3)\,\alpha^3$

Finally: Vary all θ_i to account for correctly correlated theory uncertainty

Application to Drell-Yan p_T Spectrum.

Step 1: Expand the spectrum in powers of p_T ($\alpha \equiv \alpha_s$, $x \equiv p_T$)

$$f(\alpha; x) \equiv \frac{\mathrm{d}\sigma}{\mathrm{d}p_T}(\alpha_s) = \underbrace{\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}p_T}}_{\mathcal{O}(1)} + \underbrace{\frac{\mathrm{d}\sigma^{(1)}}{\mathrm{d}p_T}}_{\mathcal{O}(p_T/m_Z)} + \underbrace{\frac{\mathrm{d}\sigma^{(2)}}{\mathrm{d}p_T}}_{\mathcal{O}(p_T^2/m_Z^2)} + \cdots$$

- Leading power: p_T dependence is known and organized by resummation

$$\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}p_T} = f_0\,\delta(p_T) + \sum_n \alpha_s^n \left\{ f_n\,\delta(p_T) + \sum_m f_{nm} \left[\frac{\ln^m(p_T/m_Z)}{p_T} \right]_+ \right\}$$

  - $\mathrm{d}\sigma^{(1)}/\mathrm{d}p_T$: Solely from kinematic effects on top of $\mathrm{d}\sigma^{(0)}/\mathrm{d}p_T$
    [Ebert, Michel, Stewart, FT 2020]

- Higher-power (nonsingular) terms $\mathrm{d}\sigma^{(2)}/\mathrm{d}p_T$: Small correction

Application to Drell-Yan p_T Spectrum.

Step 1: Expand the spectrum in powers of p_T ($\alpha \equiv \alpha_s$, $x \equiv p_T$)

$$f(\alpha; x) \equiv \frac{\mathrm{d}\sigma}{\mathrm{d}p_T}(\alpha_s) = \underbrace{\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}p_T}}_{\mathcal{O}(1)} + \underbrace{\frac{\mathrm{d}\sigma^{(1)}}{\mathrm{d}p_T}}_{\mathcal{O}(p_T/m_Z)} + \underbrace{\frac{\mathrm{d}\sigma^{(2)}}{\mathrm{d}p_T}}_{\mathcal{O}(p_T^2/m_Z^2)} + \cdots$$

- Leading power: p_T dependence is known and organized by resummation

$$\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}p_T} = \Big[H \times B_a \otimes B_b \otimes S\Big](\alpha_s; L \equiv \ln p_T/m_Z)$$

  - $\mathrm{d}\sigma^{(1)}/\mathrm{d}p_T$: Solely from kinematic effects on top of $\mathrm{d}\sigma^{(0)}/\mathrm{d}p_T$
    [Ebert, Michel, Stewart, FT 2020]

- Higher-power (nonsingular) terms $\mathrm{d}\sigma^{(2)}/\mathrm{d}p_T$: Small correction

Application to Drell-Yan p_T Spectrum.

Step 2: Identify and parametrize all independent sources of uncertainties

$$\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}p_T} = \Big[H \times B_a \otimes B_b \otimes S\Big](\alpha_s; L \equiv \ln p_T/m_Z)$$

$$F(\alpha_s, L) = F(\alpha_s)\, \exp \int_0^L \mathrm{d}L' \left\{\Gamma[\alpha_s(L')]\, L' + \gamma_F[\alpha_s(L')]\right\}$$

- Problem is reduced to several scalar series with unknown numbers
  - Missing boundary conditions

$$F(\alpha_s) = F_0 + \alpha_s\, F_1 + \alpha_s^2\, F_2(\theta) + \cdots$$

  - Missing anomalous dimensions (plus beta function and splitting functions)

$$\gamma_F(\alpha_s) = \alpha_s\, \gamma_{F0} + \alpha_s^2\, \gamma_{F1} + \alpha_s^3\, \gamma_{F2}(\theta) + \cdots$$
$$\Gamma(\alpha_s) = \alpha_s\, \Gamma_0 + \alpha_s^2\, \Gamma_1 + \alpha_s^3\, \Gamma_2 + \alpha_s^4\, \Gamma_3(\theta) + \cdots$$

  - We will call this N^{2+1}LL, and analogously for N^{3+1}LL, etc.

$\Rightarrow$ Remaining task: How exactly to define and vary the θ_i?

Systematic Uncertainties via NPs in Fits.

Standard method to include systematic unc. is via nuisance parameters

ML fits: $$L(y, \theta_i) = P(d|y, \theta_i) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(u_i - \theta_i)^2}{2\sigma_i^2}\right]$$

χ^2 fits: $$\chi^2(y, \theta_i) = \sum_d \frac{[d(\theta_i) - f_d^{\text{predicted}}(y)]^2}{\sigma_d^2} + \sum_i \frac{(u_i - \theta_i)^2}{\sigma_i^2}$$

- $d =$ measurement with statistical uncertainty σ_d
- $y =$ parameter(s) of interest
- $\theta_i =$ nuisance parameters

- Auxiliary (real or imagined) measurements provide constraint on θ_i
  - Usually assumed to be independent and Gaussian distributed
  - $u_i =$ best estimate of θ_i (from an actual measurement or our best guess)
  - $\sigma_i =$ uncertainty on u_i (the estimated "systematic uncertainty")

Theory Uncertainties via TNPs.

As before, except that now prediction depents on θ_i (ignoring exp. NPs now)

ML fits: $$L(y, \theta_i) = P(d|y, \theta_i) \times \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(u_i - \theta_i)^2}{2\sigma_i^2}\right]$$

χ^2 fits: $$\chi^2(y, \theta_i) = \sum_d \frac{[d - f_d^{\text{predicted}}(y, \theta_i)]^2}{\sigma_d^2} + \sum_i \frac{(u_i - \theta_i)^2}{\sigma_i^2}$$

Important to distinguish what we need to estimate

- We *do not* need a precise estimate of the true value of θ_i (or f_2)
  - Often our best estimate for θ_i will be $u_i = 0$
  - We can (and will) have $f_2(\theta_2 = 0) \neq 0$
  - If we obtain nontrivial information on θ_i we can include it via $u_i \neq 0$
- We *do* need to estimate σ_i of u_i (the systematic "theory uncertainty")
  - i.e., how is θ_i allowed to vary
  - For this purpose, it is sufficient to understand the *generic size* of f_2, e.g. we can consider the most dominant contributions or limits

TNP Parameterization.

cross sections, boundary conditions: $$F(\alpha_s) = 1 + \sum_{n=1} \left(\frac{\alpha_s}{4\pi}\right)^n F_n$$

anomalous dimensions: $$\gamma(\alpha_s) = \sum_{n=0} \left(\frac{\alpha_s}{4\pi}\right)^{n+1} \gamma_n$$

- Parametrize F_n and γ_n as

$$F_n(\theta_n) = 4C_r(4C_A)^{n-1}(n-1)!\,\theta_n(n_f)$$

$$\gamma_n(\theta_n) = 2C_r(4C_A)^n\,\theta_n(n_f)$$

  - $C_r =$ leading color factor
  - $C_A^{n-1} =$ leading n-loop color factor in large-N_c limit

Key questions

- What is the typical size of the θ_n? How are they distributed?
- In statistics terms: QCD has a population of θ_n – how is it distributed?

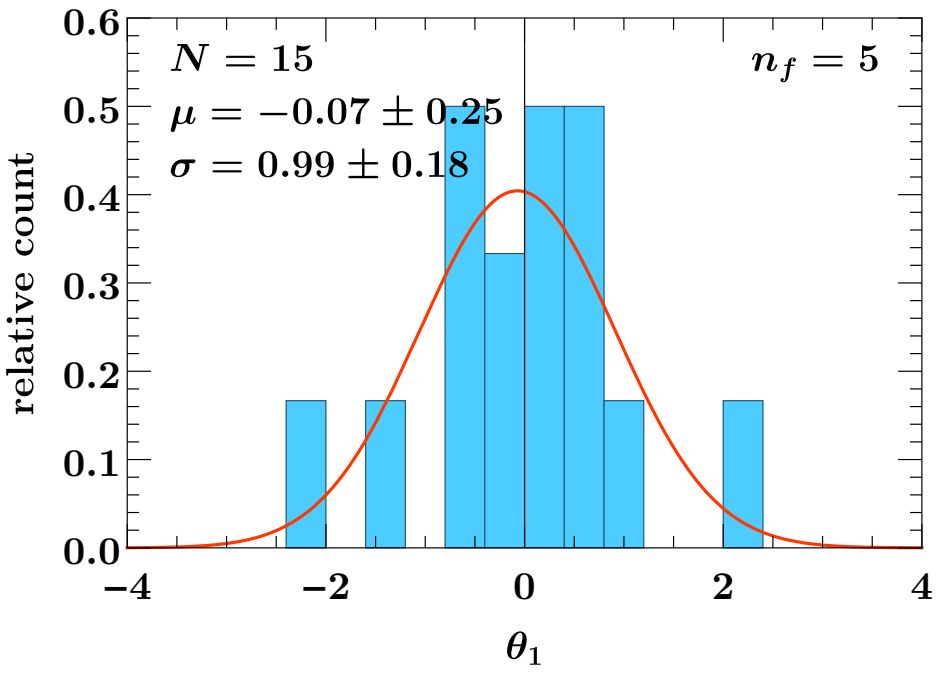
Fineprint: Need a normalization convention to fix ambiguity $F \to F^k$ and $\gamma \to \pm k\gamma$

$\to$ F^k should correspond to $|\mathcal{A}_k|^2$ for k external legs

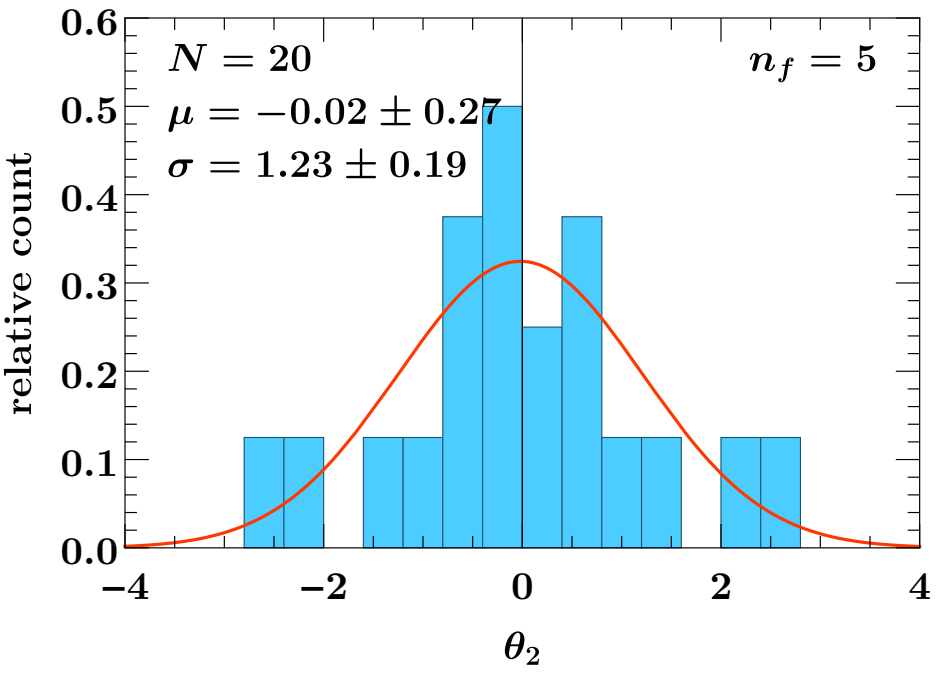
$\to$ γ_n analogously with sign convention $\theta_0 \geq 0$

Distribution of TNPs: Boundary Conditions.

1 loop: θ_1



2 loop: θ_2



$$F_n(\theta_n) = 4C_r(4C_A)^{n-1}(n-1)!\,\theta_n(n_f)$$

Estimate unknown θ_n population by sample distribution of known series

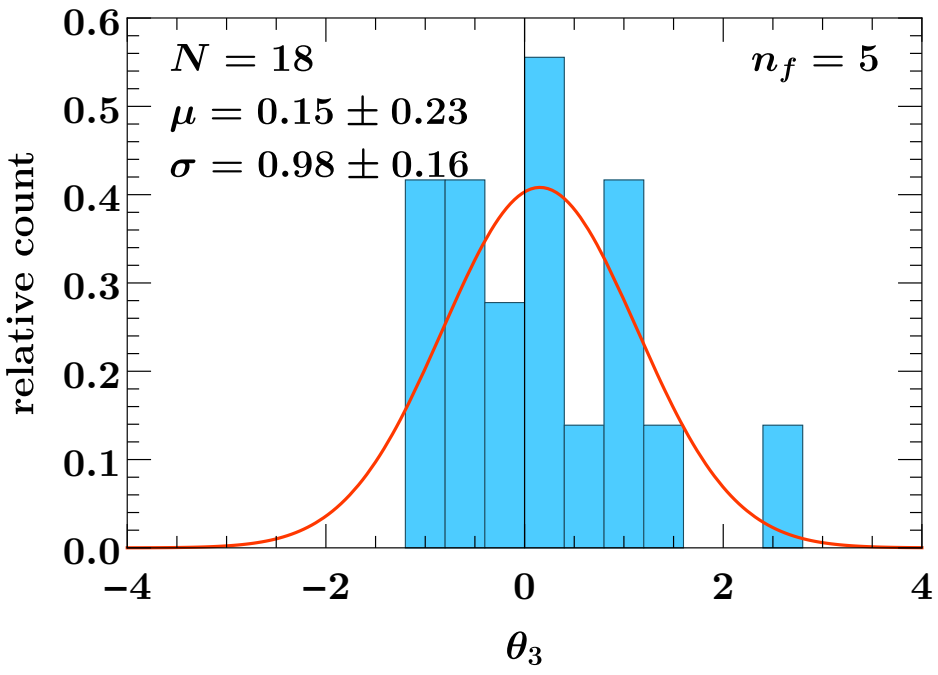
- ✓ Good fit to a Gaussian with mean 0 and variance 1
- ✓ Provides well-defined and reliable estimate: $u_i = 0$ with $\sigma_i = 1$

Fineprint: Strong n_f dependence, for $n_f \to 0$ variance increases to $\sigma_i \sim n$

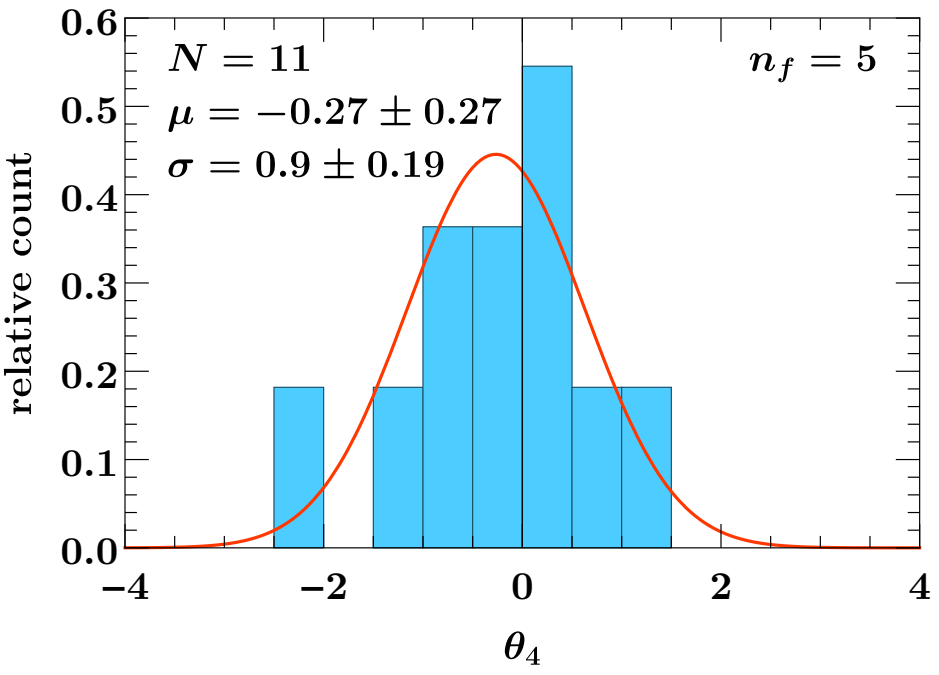
→ Nontrivial correlation between different n_f, can be estimated from sample

Distribution of TNPs: Boundary Conditions.

3 loop: θ_3



4 loop: θ_4



$$F_n(\theta_n) = 4C_r(4C_A)^{n-1}(n-1)!\,\theta_n(n_f)$$

Estimate unknown θ_n population by sample distribution of known series

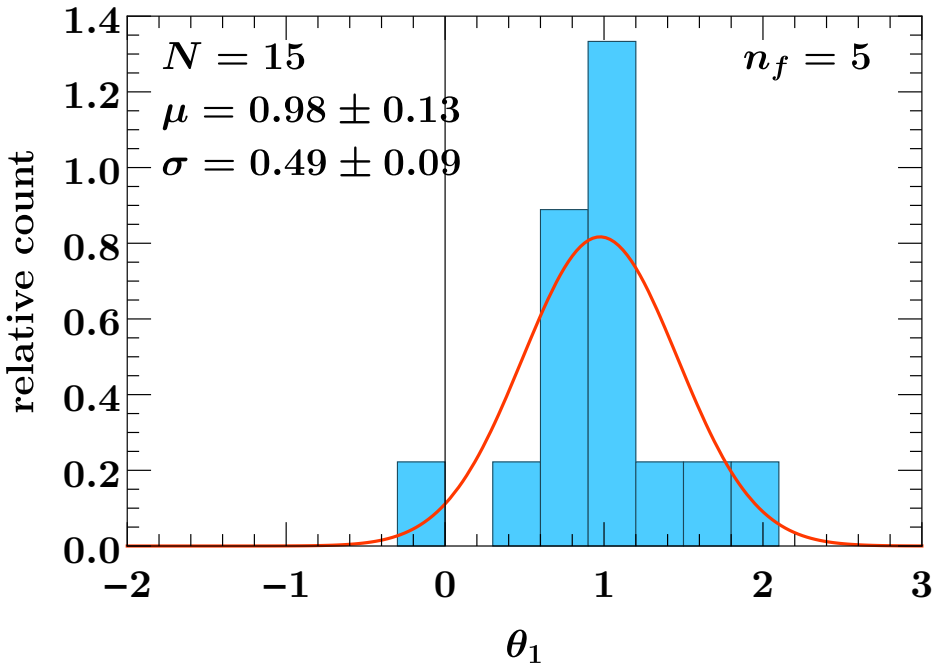
- ✓ Good fit to a Gaussian with mean 0 and variance 1
- ✓ Provides well-defined and reliable estimate: $u_i = 0$ with $\sigma_i = 1$

Fineprint: Strong n_f dependence, for $n_f \to 0$ variance increases to $\sigma_i \sim n$

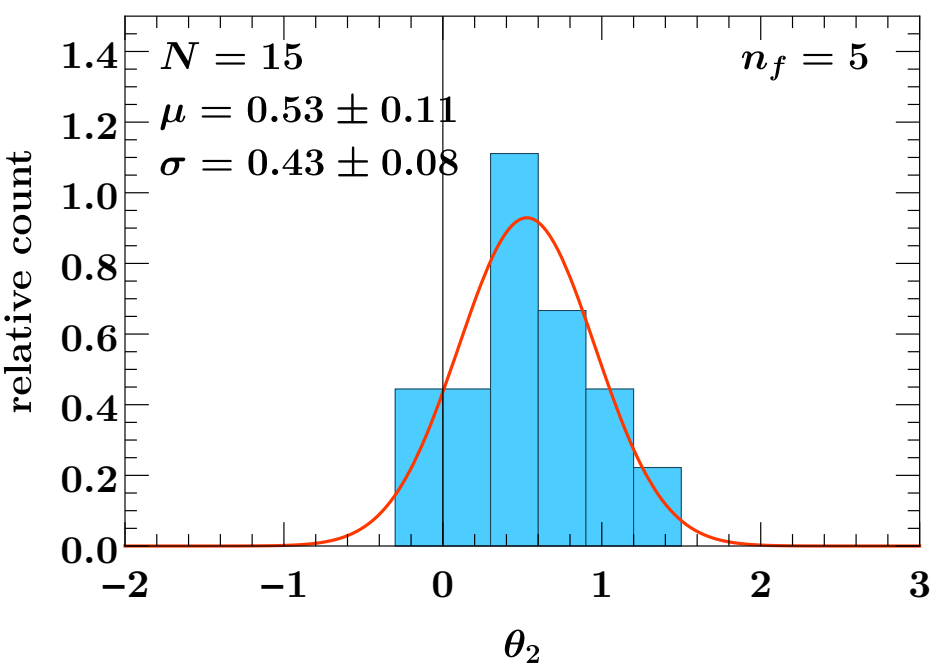
→ Nontrivial correlation between different n_f, can be estimated from sample

Distribution of TNPs: Anomalous Dimensions.

2 loop: θ_1



3 loop: θ_2



$$\gamma_n(\theta_n) = 2C_r(4C_A)^n\,\theta_n(n_f)$$

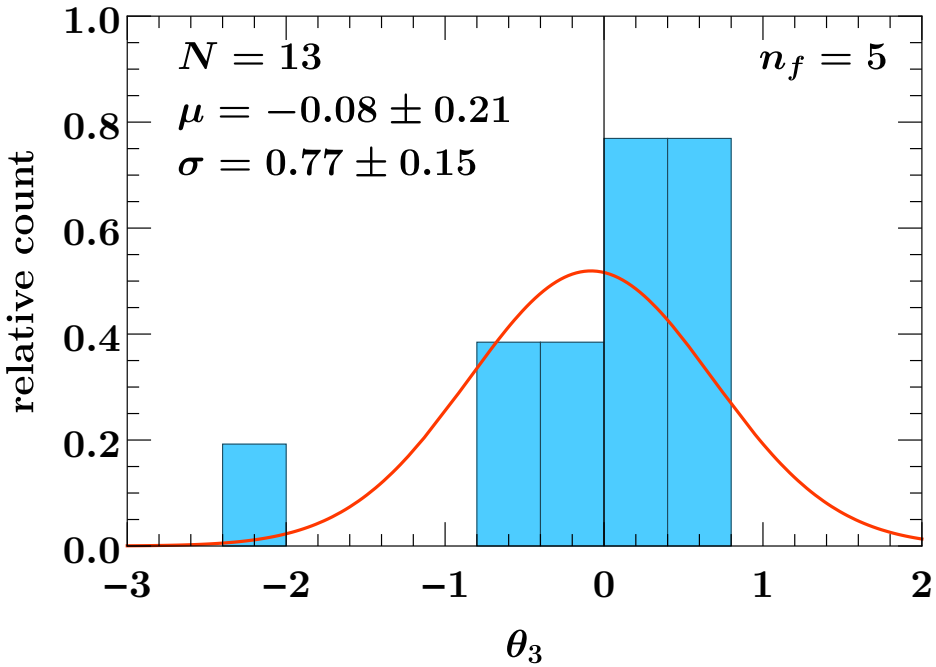
Estimate unknown θ_n population by sample distribution of known series

✓ Good fit to a Gaussian with variance ~ 0.5 and mean $\neq 0$

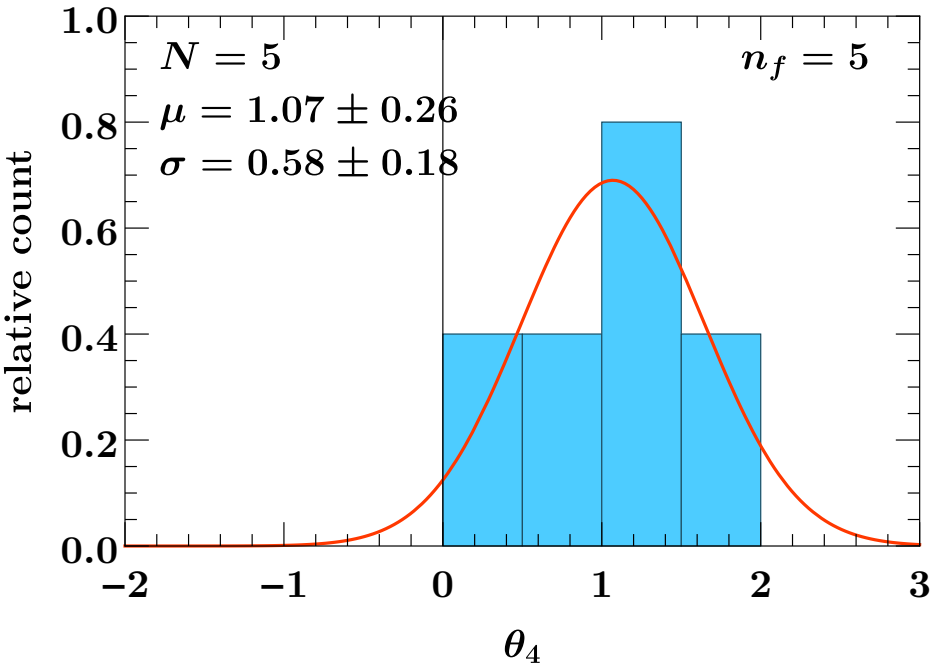
- In the following will use $u_i = 0$ with $\sigma_i = 1$
  (just for simplicity, as we will mostly care about θ_3)

Distribution of TNPs: Anomalous Dimensions.

4 loop: θ_3



5 loop: θ_4



$$\gamma_n(\theta_n) = 2C_r(4C_A)^n\,\theta_n(n_f)$$

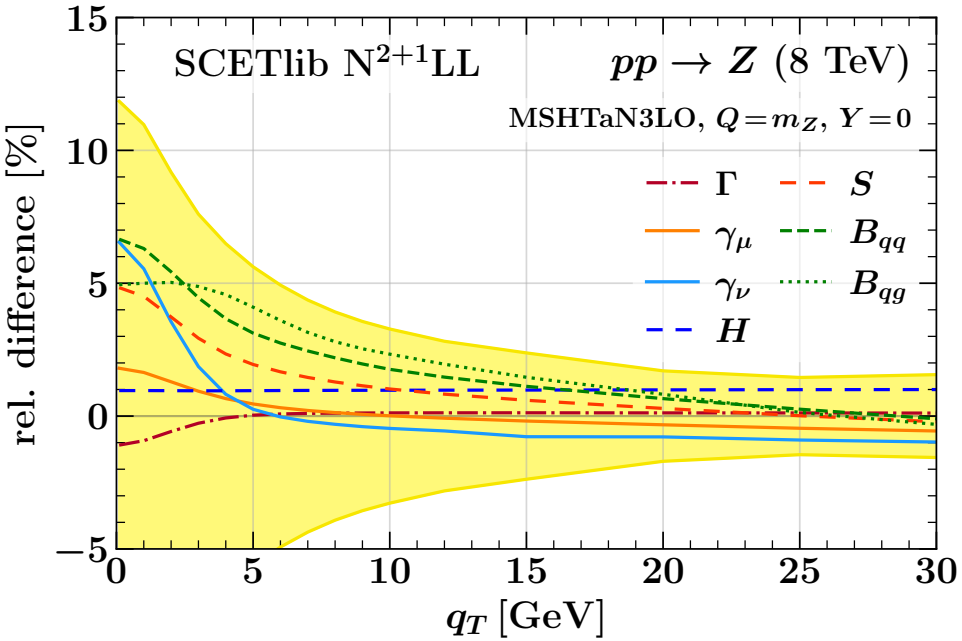
Estimate unknown θ_n population by sample distribution of known series

✓ Good fit to a Gaussian with variance ~ 0.5 and mean $\neq 0$

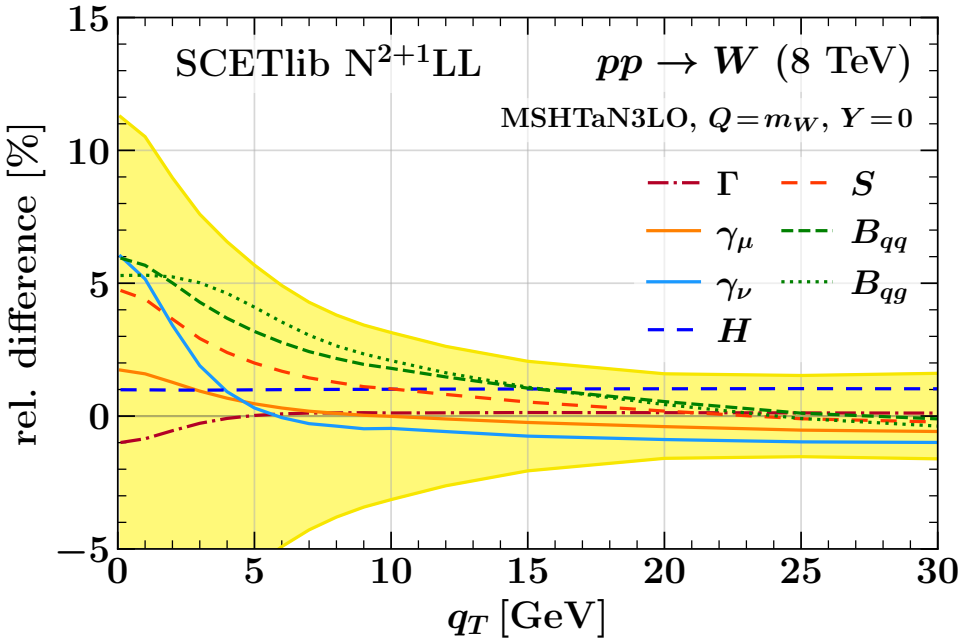
- In the following will use $u_i = 0$ with $\sigma_i = 1$
  (just for simplicity, as we will mostly care about θ_3)

Application to Drell-Yan p_T Spectrum.

relative impact for Z



relative impact for W



Step 3: Have fun with TNPs ($\to$ see Giulia's talk next for application to α_s fit)

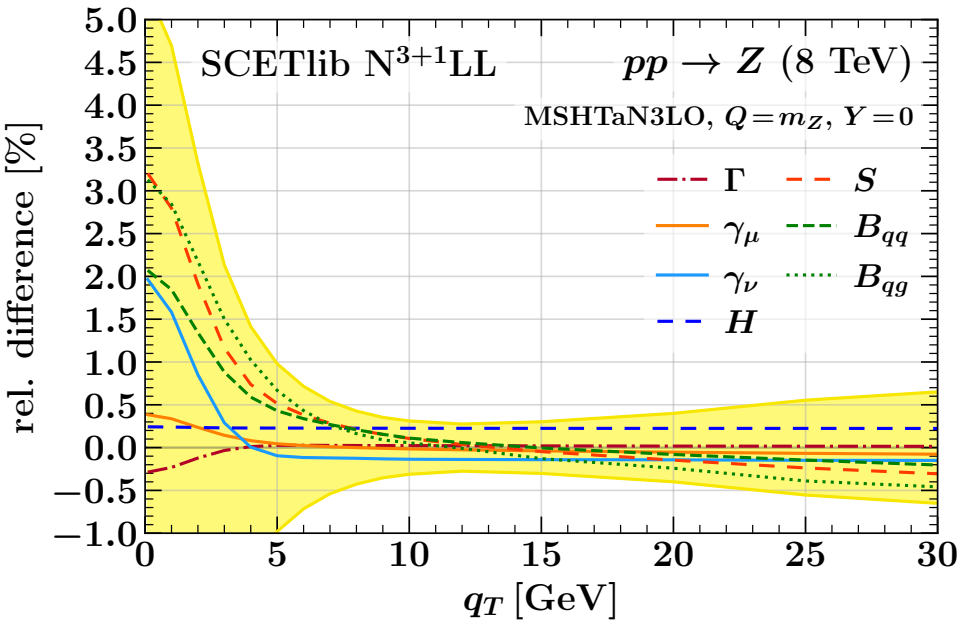
- Vary each θ_i, add in quadrature to get total unc. for spectrum

  Some disclaimers:
  - Beam boundary conditions B_{qj}: Using $f_n = (0 \pm 1.5) \times f_n^{\text{true}}$
  - DGLAP splitting functions not varied here (count as noncusp anom. dims.)
  - Hard boundary conditions H: No singlet corrections (enter for Z but not W)

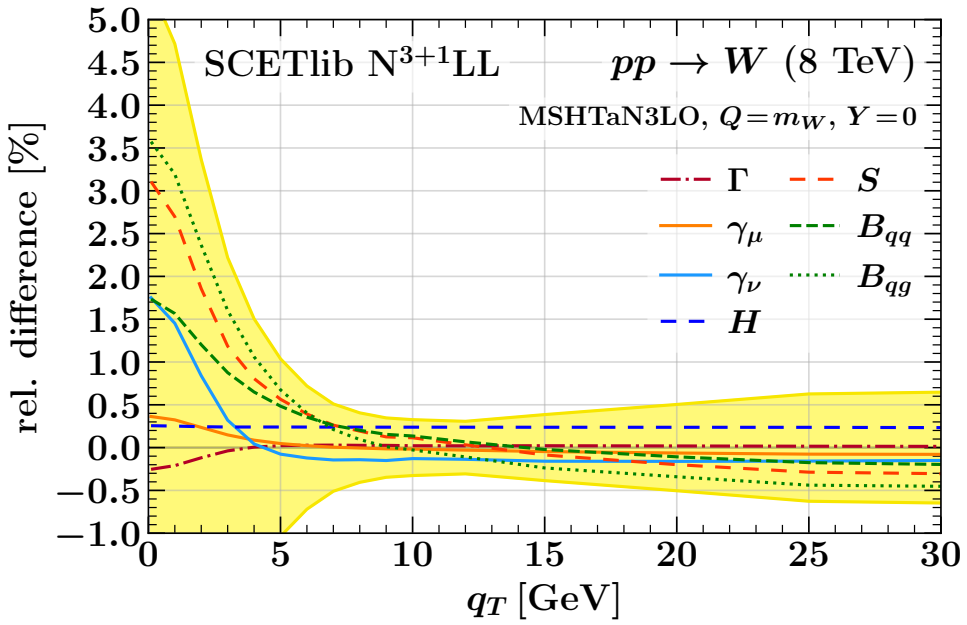
✓ Correlations in p_T and Q and between W and Z are correctly captured

Application to Drell-Yan p_T Spectrum.

relative impact for Z



relative impact for W



Step 3: Have fun with TNPs ($\to$ see Giulia's talk next for application to α_s fit)

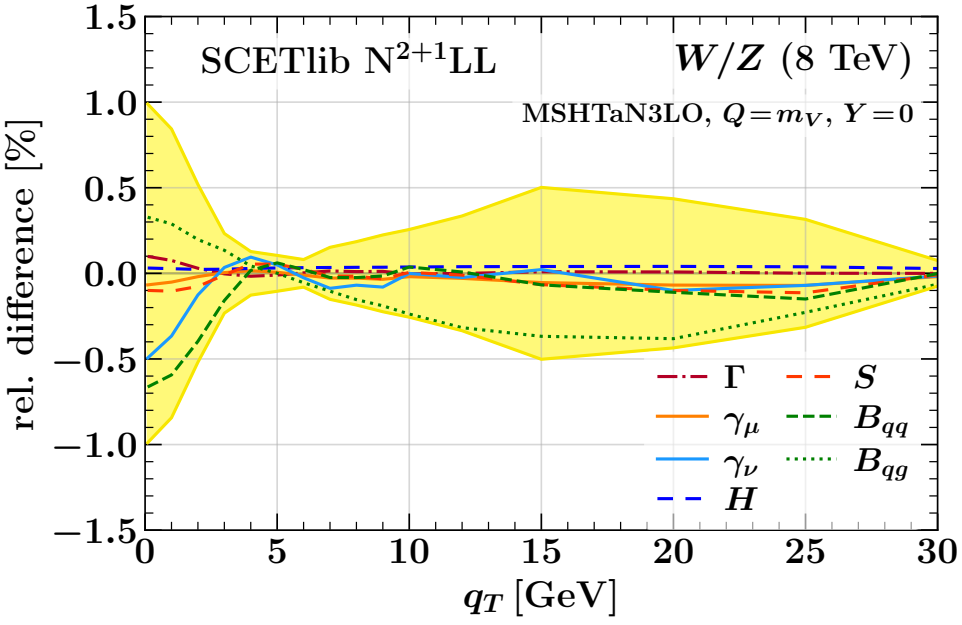
- Vary each θ_i, add in quadrature to get total unc. for spectrum

  Some disclaimers:
  - Beam boundary conditions B_{qj}: Using $f_n = (0 \pm 1.5) \times f_n^{\text{true}}$
  - DGLAP splitting functions not varied here (count as noncusp anom. dims.)
  - Hard boundary conditions H: No singlet corrections (enter for Z but not W)

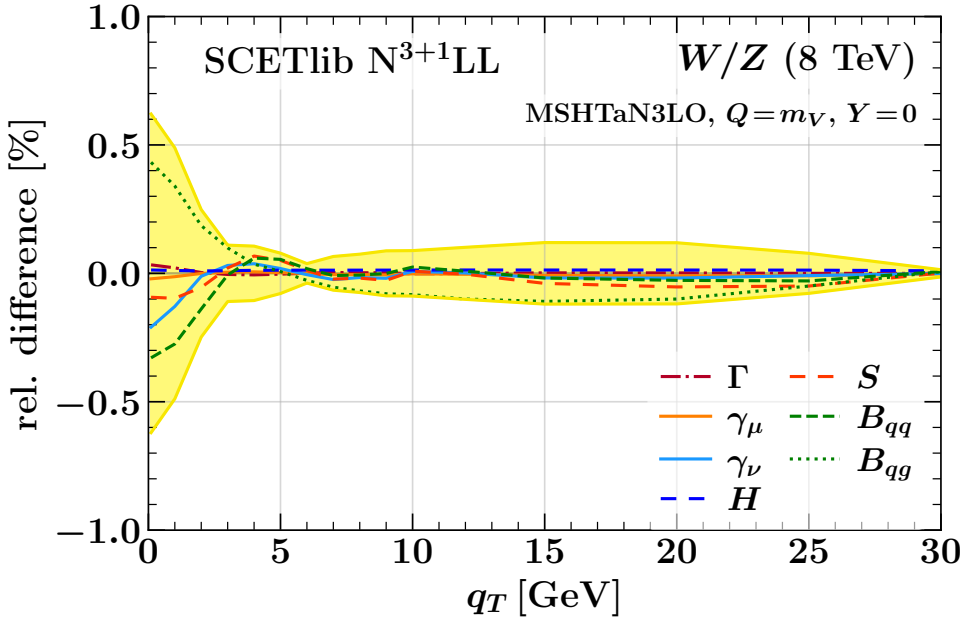
✓ Correlations in p_T and Q and between W and Z are correctly captured

Application to Drell-Yan p_T Spectrum.

relative impact for W/Z



relative impact for W/Z



Step 3: Have fun with TNPs ($\rightarrow$ see Giulia's talk next for application to α_s fit)

- Vary each θ_i, add in quadrature to get total unc. for spectrum or ratio or ...

  Some disclaimers:
  - Beam boundary conditions B_{qj}: Using $f_n = (0 \pm 1.5) \times f_n^{\text{true}}$
  - DGLAP splitting functions not varied here (count as noncusp anom. dims.)
  - Hard boundary conditions H: No singlet corrections (enter for Z but not W)

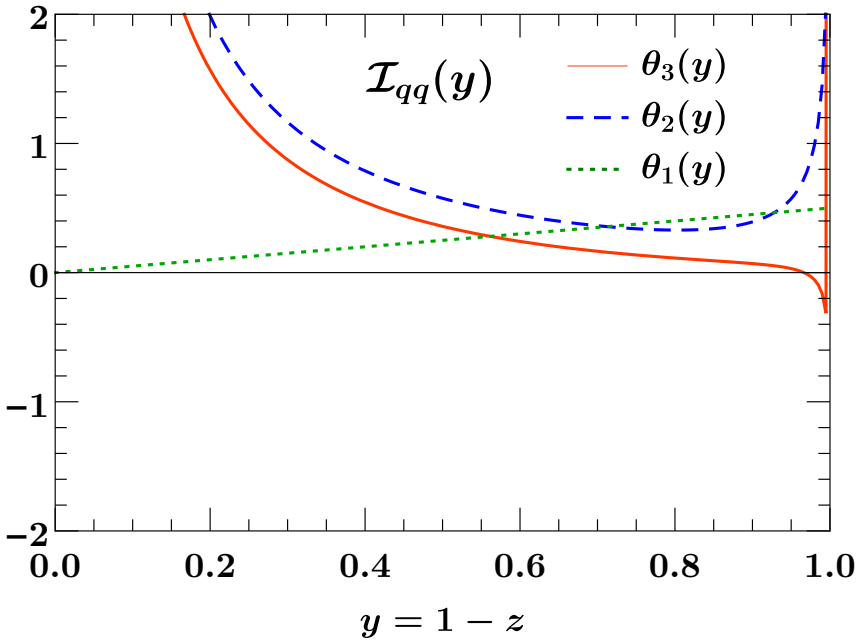
✓ Correlations in p_T and Q and between W and Z are correctly captured

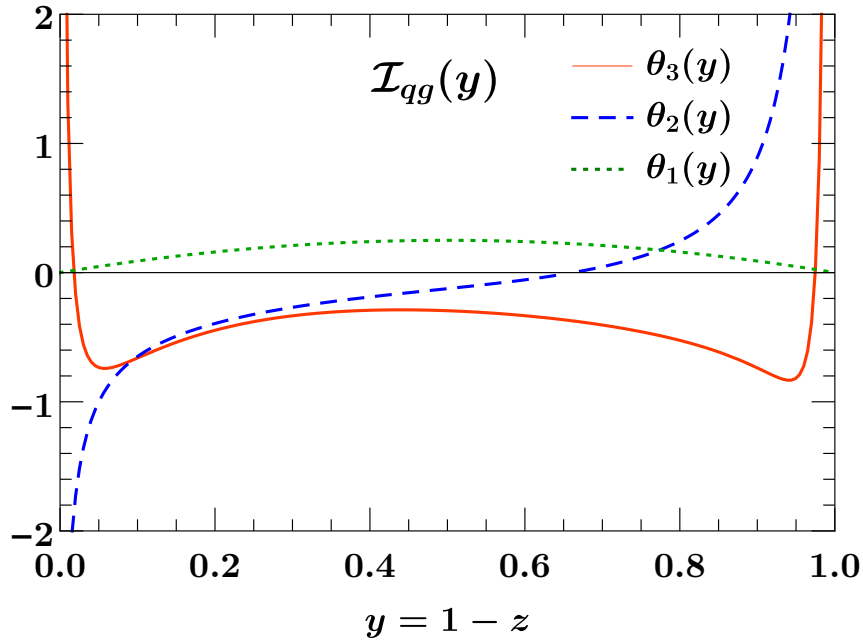
Functional TNPs.

What do we do when TNPs are genuine functions: $\theta_i \equiv \theta_i(x)$

- General strategy: Sit down and think ...
  - Recall: We don't need a precise estimate
  - Relax: We can never do worse than scale variations
- The associated uncertainty has to reflect our lack of knowledge
  - Normalization determines the overall uncertainty
  - The extent to which we control the shape in x is the extent to which we control the correlations across x
- Parametrize x dependence based on what we know
  - Exploit known functional dependence
  - Expand in known limits
  - Impose constraints on the integral
- Worst-case scenarios
  - Expand in suitable orthogonal basis, truncate at required resolution in x
  - Use overall $\theta \times$ lower-order shape

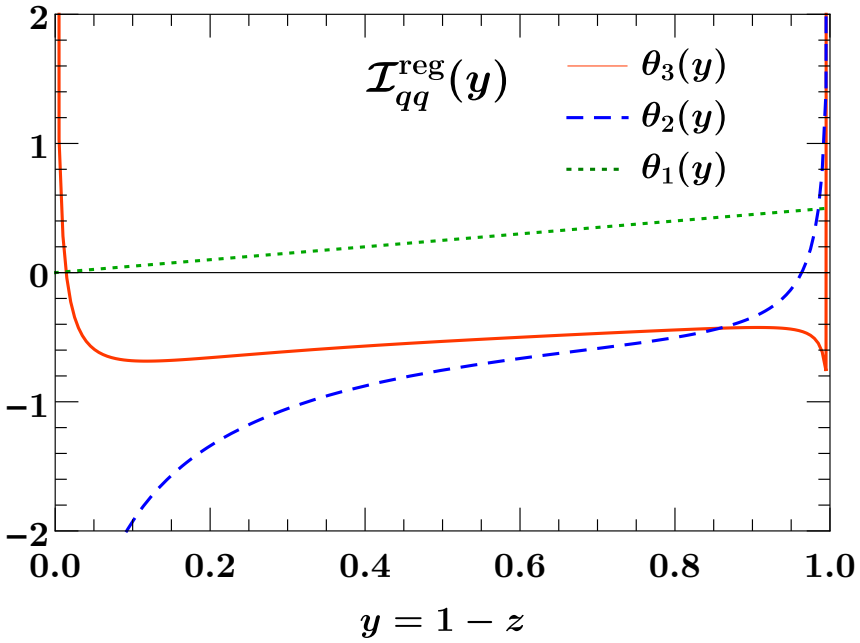
Beam Functions.

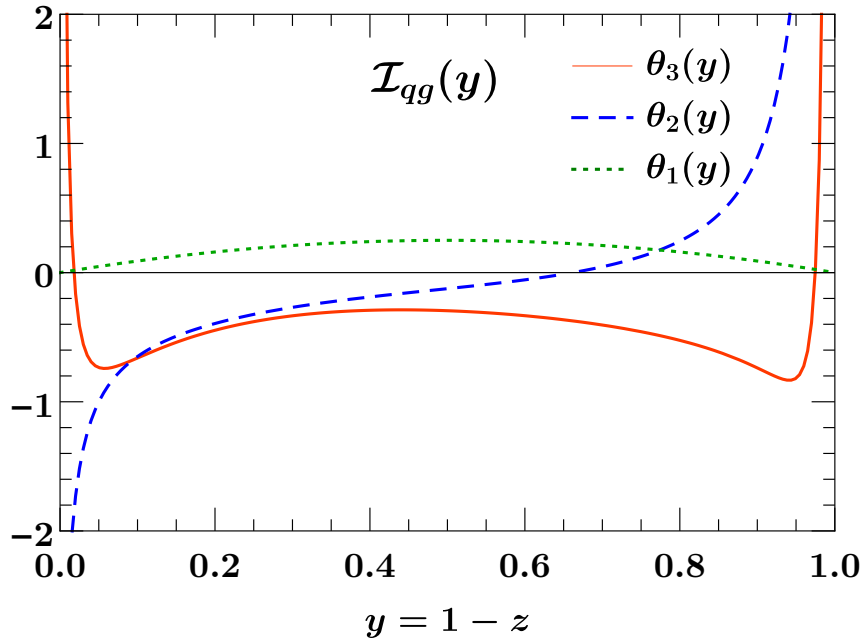




- With default normalization, bulk region is already $\lesssim 1$

Beam Functions.

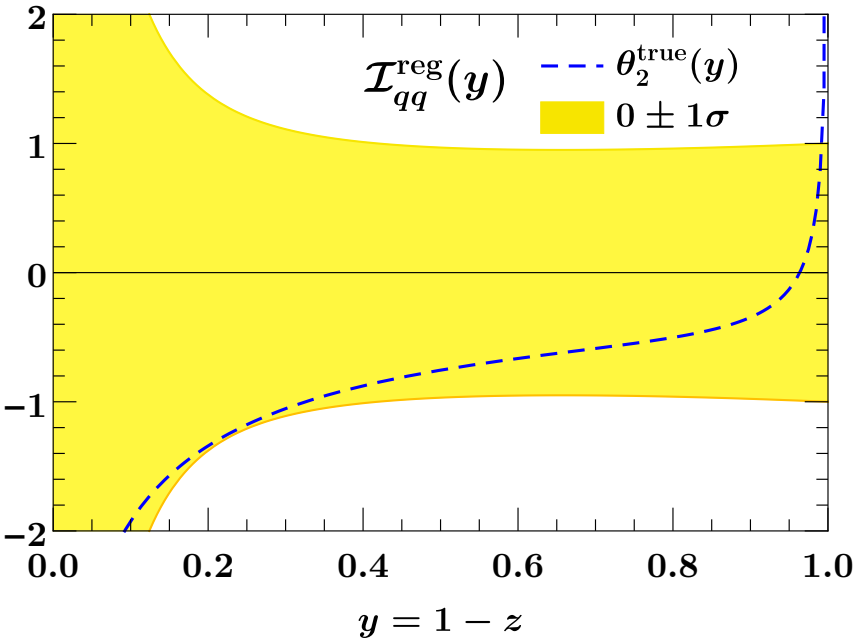


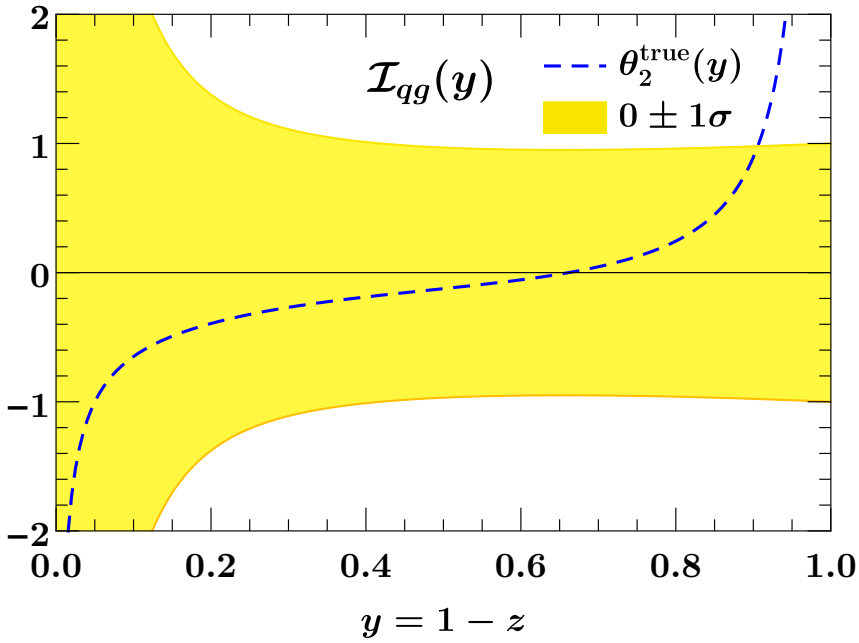


- With default normalization, bulk region is already $\lesssim 1$
- Singular limit $y = 1 - z \to 0$ is known to all orders in terms of γ_ν
  [Billis, Ebert, Michel, FT 2019]
- To account for remaining $y \to 0$ limit, use NLP expansion

$$\theta_n(y) = \theta_{n0} + \sum_{k=1} \theta_{nk} \frac{1}{k!} \ln^k y$$

Beam Functions.



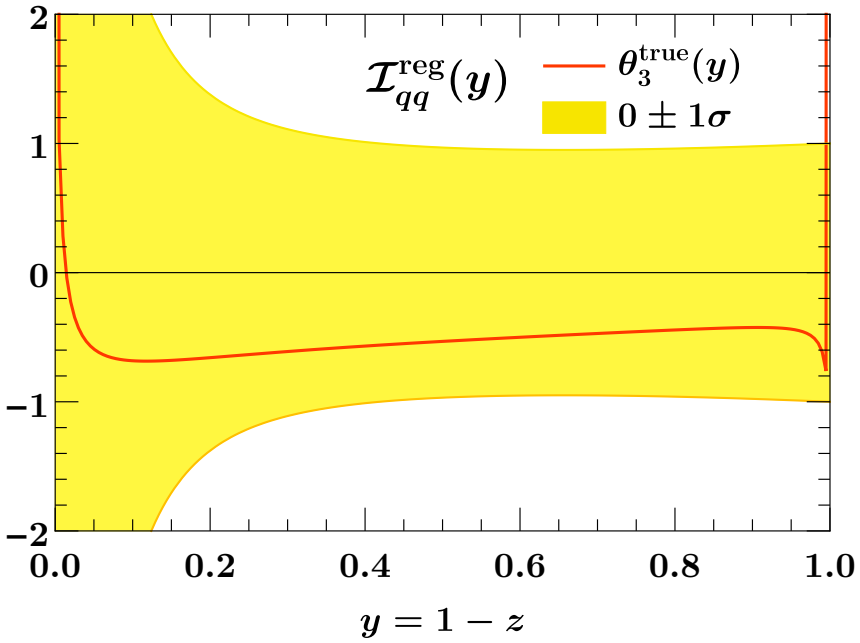


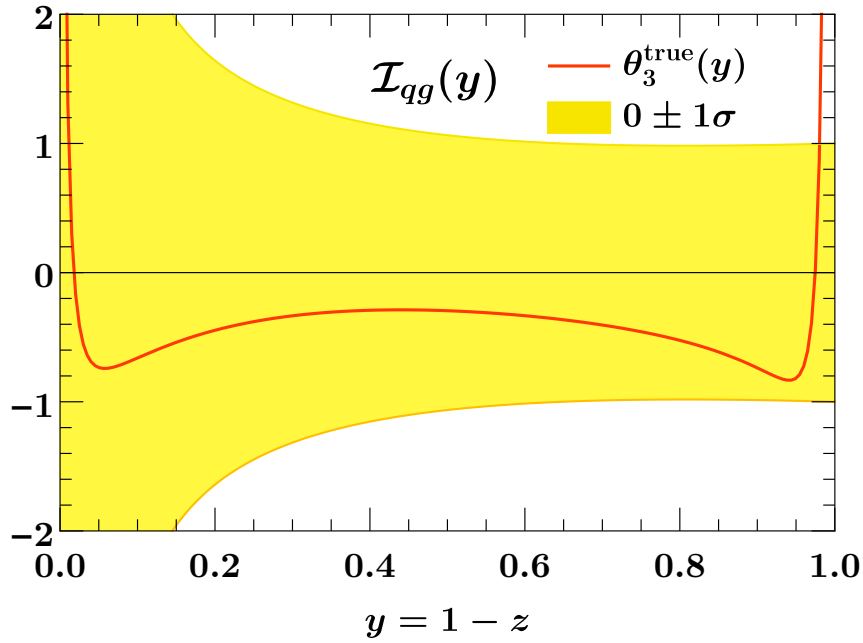
- With default normalization, bulk region is already $\lesssim 1$
- Singular limit $y = 1 - z \to 0$ is known to all orders in terms of γ_ν
  [Billis, Ebert, Michel, FT 2019]
- To account for remaining $y \to 0$ limit, use NLP expansion

$$\theta_n(y) = \theta_{n0} + \sum_{k=1} \theta_{nk} \frac{1}{k!} \ln^k y$$

- Vary $\theta_{nk} = 0 \pm 1$ with constraint that total integral is 0 ± 1

Beam Functions.





- With default normalization, bulk region is already $\lesssim 1$
- Singular limit $y = 1 - z \to 0$ is known to all orders in terms of γ_ν
  [Billis, Ebert, Michel, FT 2019]
- To account for remaining $y \to 0$ limit, use NLP expansion

$$\theta_n(y) = \theta_{n0} + \sum_{k=1} \theta_{nk} \frac{1}{k!} \ln^k y$$

- Vary $\theta_{nk} = 0 \pm 1$ with constraint that total integral is 0 ± 1

Other Examples.

- Splitting functions $P_{ij}(z)$ [See e.g. MSHT aN3LO Cridge, Harland-Lang, Thorne 2022]
  - Exploit $z \to 0$ and $z \to 1$ limits
  - Constrain moments
- Nonsingular power corrections
  - Parametrize NLP structure
  - NLP factorization extremely useful even to just provide constraints on log structure without knowing all the ingredients
- Mass-dependent functions $S(m, k_T)$
  - Consider $m \ll k_T$ and $m \gg k_T$ limits
- More complicated hard functions (diboson, V+jet, ...)

Summary.

Interpretation of LHC precision measurements requires theory predictions with reliable uncertainties and in particular correct correlations

Scale variations

- Neither particularly reliable nor can they do correlations
  - → One cannot rely on them for shape uncertainties which unfortunately is exactly what is often done
- Become insufficient as theory unc. $\gg$ experimental unc.

Theory nuisance parameters overcome many limitations of scale variations

- Allow for truly parametric, no-longer ugly theory uncertainties
  - ✓ Encode correct correlations
  - ✓ Can be consistently propagated everywhere (fits, MCs, neural networks, ...)
  - ✓ Can even be constrained by measurements (auxiliary and/or primary)
- Bonus: Can fully benefit from known partial higher-order information
- Price to pay: Obviously not as “easy and cheap” as scale variations

Additional Slides

Acknowledgments.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101002090 COLORFREE)