Published for SISSA by Springer

Received: June 26, 2025
Revised: September 23, 2025
Accepted: October 20, 2025
Published: December 5, 2025

Theory uncertainties in the extraction of α_s from Drell-Yan at small transverse momentum

Thomas Cridge,[a] **Giulia Marinelli**[b] **and Frank J. Tackmann**[b]

[a] *Elementary Particle Physics, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium*
[b] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*

E-mail: thomas.cridge@uantwerpen.be, giulia.marinelli@desy.de, frank.tackmann@desy.de

Abstract: We perform a detailed pseudodata study to estimate the expected theory uncertainty in the extraction of the strong coupling constant, $\alpha_s(m_Z)$, from a fit to the measured Drell-Yan transverse momentum (q_T) spectrum at small $q_T \ll m_Z$. We consider two approaches to estimate the dominant perturbative uncertainties. We first discuss that the traditional approach based on varying unphysical scales is insufficient here because it cannot correctly account for bin-by-bin theory correlations in the q_T spectrum, which are critically important in this case. We then use this case as a nontrivial application of a new approach based on theory nuisance parameters (TNPs), which encodes the correct theory correlations by construction. Moreover, the TNPs can be profiled in the fit thereby allowing the data to constrain the theory uncertainties in a consistent manner. We furthermore discuss the interplay with nonperturbative effects in the peak region $q_T \lesssim 10\,\text{GeV}$, from where most of the α_s sensitivity originates. The associated nonperturbative uncertainties on α_s when fitting only the q_T spectrum are large. They can in principle be reduced by including additional constraints on the nonperturbative Collins-Soper kernel from lattice QCD calculations. We find that these improvements in the treatment of perturbative and nonperturbative uncertainties and their correlations will enable a competitive α_s extraction from Drell-Yan data at small q_T. We also discuss the implications of our findings, calling into question a recent α_s extraction from the Z q_T spectrum by the ATLAS experiment.

Keywords: Resummation, Specific QCD Phenomenology, The Strong Coupling

ArXiv ePrint: 2506.13874

https://doi.org/10.1007/JHEP12(2025)043

Contents

JHEP12(2025)043

1 Introduction

The strong coupling constant α_s, which parameterizes the strength of the strong interactions, is the least well known among the fundamental forces of nature. Its precise determination has thus been of great interest with many complementary determinations from different sources being pursued over the years [1]. The current world average of $\alpha_s(m_Z) = 0.1180(9)$ [2] has a relative uncertainty of 0.8%, with the most precise result stemming from the average of lattice QCD determinations $\alpha_s(m_Z) = 0.1183(7)$ (0.6%) [3].

One of the currently most precise determinations from experimental measurements at colliders comes from fits to hadronic event shapes in e^+e^- collisions [4], most notably a determination from the thrust distribution [5–7] yielding $\alpha_s(m_Z) = 0.1136(12)$ with a precision of about 1%, which however has shown a persistent tension with the world average, despite much theoretical scrutiny [7–15]. An independent determination of similarly high precision at or below the one-percent level from collider measurements is thus of great importance. A promising candidate for such a high-precision determination is the transverse-momentum (q_T) spectrum in Drell-Yan production, which has been measured to subpercent precision at the LHC [16–24] and for which theoretical predictions have been reaching the same N^{3}LL$'$ and (approximate) N^{4}LL resummation as for e^+e^- event shapes resulting in a perturbative precision at the few-percent level [25–32].

The sensitivity to α_s in the q_T spectrum arises solely from higher-order QCD corrections. A robust theoretical description of the q_T spectrum is thus mandatory to obtain a reliable extraction of $\alpha_s(m_Z)$. It requires in particular a reliable treatment of theory uncertainties and their correlations. The dominant α_s sensitivity is a subtle shape effect in the peak of the spectrum — a change in α_s effectively shifts the peak of the spectrum. Since the α_s dependence is intimately tied to the perturbative expansion itself, missing higher-order terms can closely mimic the effect of changing the value of α_s. Furthermore, nonperturbative effects induce a similar shape effect in the peak of the spectrum. Hence, for a reliable α_s determination it is vital to reliably distinguish the shape effects induced by changing α_s from those induced by missing higher-order corrections or by nonperturbative effects, so we must be able to correctly account for the latter two. This means it is critically important to account for the correct point-by-point (or bin-by-bin) theory correlations in the q_T spectrum, which are determined by the precise shape of the individual components of the theory uncertainty.

The traditional approach to estimate perturbative theory uncertainties based on scale variations suffers from well-known limitations: The estimated uncertainties are not very meaningful, and in particular they fundamentally lack a proper account of theory correlations. As discussed in section 2.2, the underlying reason is that the scales that are being varied are unphysical and not actual parameters of the calculation. Their variation cannot be interpreted like that of an ordinary parameter whose uncertainty is being propagated. One consequence is that scale variations do not provide the correct shape of the missing higher-order terms. Given the critical importance of accounting for the correct point-by-point theory correlations, the scale-variation approach is basically ill-suited for this application. We find that the perturbative uncertainty on $\alpha_s(m_Z)$ one might derive based on scale variations is indeed not reliable and quite arbitrary due to it being subject to uncontrolled correlation assumptions. For example, a recent extraction from Drell-Yan measurements at small q_T claimed a record

JHEP12(2025)043

precision of $\alpha_s(m_Z) = 0.1183(9)$ [33], as precise as the current world average. However, as discussed in detail in appendix A, it relies on a particular scale-variation recipe as well as other questionable assumptions, so the quoted uncertainty on $\alpha_s(m_Z)$ cannot be taken at face value.

The new approach of theory nuisance parameters (TNPs) introduced in ref. [34] overcomes the limitations of scale variations and provides meaningful theory uncertainties including in particular correct theory correlations. The purpose of our paper is thus twofold: Our first goal is to demonstrate the performance and robustness of TNPs in a real-world application that demands meaningful theory uncertainties and correlations. We find that the TNPs perform as advertised. Our second goal is to estimate the theory uncertainties in $\alpha_s(m_Z)$ we can expect to achieve. To do so we perform a detailed study using so-called Asimov fits to (unfluctuated) pseudodata. This is a standard procedure to estimate expected uncertainties in a clean and controlled environment that is unobscured by statistical fluctuations in the data and unbiased by subleading effects that are present in the real data and not (yet) accounted for in the theoretical description. In our case, it allows us to study in a theoretically fully controlled environment the dominant perturbative and nonperturbative uncertainties in the q_T spectrum and their propagation to and impact on $\alpha_s(m_Z)$. In particular, by profiling the TNPs in the fit and by incorporating lattice QCD constraints on nonperturbative parameters, we expect that it will be possible in the future to achieve sufficiently small theory uncertainties to allow for an extraction of $\alpha_s(m_Z)$ from the Drell-Yan q_T spectrum that is competitive with other measurements. It remains to be seen whether or not it will be possible to reach or beat the precision of lattice QCD determinations or the world average.

We use the binning and experimental uncertainties of the recent ATLAS 8 TeV inclusive measurement [19] as a representative example for our Asimov study. This choice is partially motivated by the fact that the same dataset was used in ref. [33]. Our exact numerical results of course depend on this choice and will change by including other datasets. However, our overall qualitative findings on the robustness (or lack thereof) and performance of the different theory uncertainty approaches hold in general.

The dominant perturbative and nonperturbative uncertainties we wish to address enter via the dominant leading-power resummed contribution to the small-q_T spectrum, which is therefore the focus of our investigations. In the appendix, we also briefly consider subdominant effects from power corrections suppressed by $\mathcal{O}(q_T^2/m_Z^2)$ and finite quark masses relevant for $q_T^2 \sim m_c^2, m_b^2$. While it is important to include these effects to obtain a percent-level or better description of the real data, we can drop them in our main study since we consistently drop them in both our pseudodata and our fitted theory model.

Another important source of uncertainty is due to the limited knowledge of parton distribution functions (PDFs), which are required inputs for such hadron collider determinations of $\alpha_s(m_Z)$. The PDFs themselves have a dependence on α_s [35–39], and this non-trivial PDF-α_s correlation must be propagated faithfully to the extracted output $\alpha_s(m_Z)$. Whilst the propagation of these uncertainties is in principle a more straightforward exercise, in practice it faces some nontrivial caveats. We therefore do not consider them here but will address them in a dedicated paper [40].

The remainder of this paper is organized as follows. In section 2, we discuss in more detail the general issue of theory uncertainties and their correlations for differential spectra,

JHEP12(2025)043

the limitations of scale variations, and the general ideas behind the TNP approach. In section 3, we summarize our theory and pseudodata setup. We then discuss in section 4 the treatment of perturbative uncertainties in the extraction of α_s and including nonperturbative uncertainties in section 5. We conclude in section 6. Appendix A provides a detailed discussion of the estimation of theory uncertainties in ref. [33] in comparison to our own analysis. In appendix B we collect some additional numerical results.

2 Theory uncertainties and correlations for differential spectra

In this section, we provide a brief general discussion of perturbative theory uncertainties for the case of a differential spectrum. The discussion here is largely adapted from sections 2 and 3 of ref. [34], to which we refer for a more in-depth discussion. In section 2.1, we highlight the importance of theory correlations. In section 2.2 we discuss the limitations of scale variations, and in section 2.3 we discuss the approach of theory nuisance parameters.

We consider a differential quantity $f(x)$ that depends on some variable x, and its perturbative expansion in a small parameter α,

$$f(x,\alpha) = f_0(x) + f_1(x)\,\alpha + f_2(x)\,\alpha^2 + f_3(x)\,\alpha^3 + \mathcal{O}(\alpha^4)\,. \tag{2.1}$$

We will generically use $f_n(x)$ to denote the expansion coefficients and specifically $\hat{f}_n(x)$ to denote their true values. Calculating the first few true coefficients, we obtain a theory prediction for $f(x)$ at leading order (LO), next-to-leading order (NLO), next-to-next-to-leading order (NNLO), and so on,

$$\begin{aligned} \text{LO:} \quad & f(x,\alpha) = \hat{f}_0(x)\,, \\ \text{NLO:} \quad & f(x,\alpha) = \hat{f}_0(x) + \hat{f}_1(x)\,\alpha\,, \\ \text{NNLO:} \quad & f(x,\alpha) = \hat{f}_0(x) + \hat{f}_1(x)\,\alpha + \hat{f}_2(x)\,\alpha^2\,. \end{aligned} \tag{2.2}$$

The perturbative theory uncertainty we wish to discuss in the following is due to the fact that the predictions are only approximations to the exact result because of truncating the series at a certain order.

2.1 Theory correlations

Theory correlations, namely the correlations in the theory uncertainties of different predictions, are in principle required whenever several predictions are used at the same time, for example when one performs a simultaneous interpretation of several measurements. A standard example is the interpretation of a differential spectrum, which requires point-by-point (or bin-by-bin) theory correlations across the spectrum.

To appreciate the potential importance of correlations, first consider the case of two observables f and g that have both a relative uncertainty $\delta_f = \delta_g = \delta$ with correlation ρ. The relative uncertainty of their ratio, $\delta_{f/g}$, as a function of ρ is given by

$$\delta_{f/g} = \delta\sqrt{2(1-\rho)}\,. \tag{2.3}$$

When ρ is close to 1, i.e., in the limit of strong correlation, $\delta_{f/g}$ is very sensitive to the precise value of ρ, as illustrated in table 1, because the square root becomes infinitely steep for $\rho \to 1$.

JHEP12(2025)043

ρ	99.5%	98%	95.5%	87.5%
$\delta_{f/g}/\delta$	0.1	0.2	0.3	0.5

Table 1. Dependence of the relative uncertainty of the ratio f/g on the correlation ρ, see text for details.

The different points (or bins) of a differential spectrum $f(x)$ are a priori separate observables. When obtaining a theory prediction for $f(x)$ we are actually obtaining predictions simultaneously for many separate (though closely related) observables $f(x_i)$. The question of how well the *shape* of $f(x)$ is known then corresponds to the question of how the theory uncertainties at different points in the spectrum are correlated. That is, $f \equiv f(x_i)$ and $g \equiv f(x_j)$ now correspond to the prediction of the spectrum at any two points x_i and x_j, and the shape of the spectrum is equivalent to the ratio $f(x_i)/f(x_j)$.

As an illustration, consider the case where we have the same relative uncertainty δ at all points. Let ρ_{ij} denote the correlation between the uncertainties of $f(x_i)$ and $f(x_j)$. In the extreme case where the uncertainties at all points are 100% correlated, $\rho_{ij} = 1$, the shape of the spectrum would be known exactly with vanishing uncertainty and δ would correspond to a pure overall normalization uncertainty. In the other extreme where the uncertainties are completely uncorrelated among all points, $\rho_{ij} = 0$, the shape would be completely unknown within an overall band of relative size δ. In practice, the theory predictions for neighboring points (or bins) are closely related, so we naturally expect their uncertainties to be strongly correlated. We are thus precisely in the limit of strong correlations, and so the precise point-by-point (or bin-by-bin) correlations are crucial to correctly account for the shape uncertainty. Just like for the ratio f/g in table 1, tiny differences in the point-by-point correlations can lead to dramatically different shape uncertainties.

It is important to realize that different quantities, including neighboring points or bins in a spectrum, do not by themselves have a notion of being correlated with each other. The only thing that can be correlated are their uncertainties. More precisely, the impact of a common source of uncertainty is fully correlated between different quantities that depend on that same source. This is fundamentally the only way a correlation can arise. For example, neighboring bins measured by strictly independent experiments will have uncorrelated experimental uncertainties, because they do not share a common source of uncertainty. Since two neighboring bins tend to be very similar observables, their theory predictions tend to involve common ingredients with a similar impact on both bins. This is the (only) reason why we expect their theory uncertainties to be strongly correlated.

More generally, when different quantities depend on several independent sources of uncertainty, the correlation of their total uncertainty depends on the relative size of the various fully correlated impacts from each source. In terms of covariance matrices, the total covariance matrix is the sum of several 100% correlated ones, which is in general not 100% correlated anymore. It then follows that to obtain the correct theory correlations for a differential spectrum we have to break down the total theory uncertainty into (mutually independent) components, where each component corresponds to a well-defined common

JHEP12(2025)043

source of uncertainty so it is fully correlated across the spectrum. The shapes of the components, reflecting the impact of their respective sources across the spectrum, can be different and their combination then determines the net point-by-point correlations of the total theory uncertainty. As we will discuss in section 2.3, theory nuisance parameters are constructed exactly to provide this breakdown. In contrast, as discussed next, scale variations do not provide the means for such a breakdown.

2.2 Limitations of scale variations

The series coefficients in eq. (2.1) depend on the precise way of performing the expansion, i.e., the perturbative scheme we use, which corresponds to the precise choice or definition of the expansion parameter α. We can define a new scheme by defining a different expansion parameter $\tilde{\alpha}$,

$$\tilde{\alpha}(\alpha) = \alpha\big[1 + b_0\,\alpha + b_1\,\alpha^2 + b_2\,\alpha^3 + \mathcal{O}(\alpha^4)\big]\,, \tag{2.4}$$

which is fully specified by the coefficients b_k. For scale variations in QCD, $\alpha \equiv \alpha_s(\mu_0)$ corresponds to the strong coupling at the chosen central scale μ_0, $\tilde{\alpha} \equiv \alpha_s(\mu)$ is the coupling at some other scale μ, and the b_k are given in terms of the QCD beta function coefficients β_k as

$$\begin{aligned}
b_0 &= \frac{\beta_0}{2\pi}\ln\frac{\mu_0}{\mu} && = 0.85\,L\,,\\
b_1 &= \frac{\beta_0^2}{4\pi^2}\ln^2\frac{\mu_0}{\mu} + \frac{\beta_1}{8\pi^2}\ln\frac{\mu_0}{\mu} && = 0.72\,L^2 + 0.34\,L\,,\\
b_2 &= \frac{\beta_0^3}{8\pi^3}\ln^3\frac{\mu_0}{\mu} + \frac{5\beta_0\beta_1}{32\pi^2}\ln^2\frac{\mu_0}{\mu} + \frac{\beta_2}{32\pi^3}\ln\frac{\mu_0}{\mu} && = 0.61\,L^3 + 0.72\,L^2 + 0.13\,L\,.
\end{aligned} \tag{2.5}$$

For illustration, in the second equalities on the right-hand side we plugged in $n_f = 5$ and used $L \equiv \ln(\mu_0/\mu)/\ln 2$, so the usual convention corresponds to varying $L = \pm 1$.

As long as the b_k in eq. (2.4) are $\mathcal{O}(1)$, $\tilde{\alpha}$ only differs from α by higher-order terms and should thus provide a similarly good expansion parameter. Using $\tilde{\alpha}$ to perform the perturbative expansion, we obtain different truncated predictions in terms of different expansion coefficients $\tilde{f}(x)$. For example, at NLO we have

$$\begin{aligned}
\text{NLO:}\quad \tilde{f}(x,\tilde{\alpha}) &= \hat{\tilde{f}}_0(x) + \hat{\tilde{f}}_1(x)\,\tilde{\alpha}\\
&= \hat{f}_0(x) + \hat{f}_1(x)\,\alpha + b_0\hat{f}_1(x)\,\alpha^2 + b_1\hat{f}_1(x)\,\alpha^3 + \mathcal{O}(\alpha^4).
\end{aligned} \tag{2.6}$$

In the second line we rewrote the predictions in terms of the original $\hat{f}_n(x)$ and α. As expected, to the order one is working, the predictions in different schemes agree but they differ by higher-order terms beyond the nominal working order.

To all orders, the different expansions must give exactly the same result, $f(x,\alpha) = \tilde{f}(x,\tilde{\alpha}) = f(x)$. The dependence on the scheme is thus an artifact of truncating the series at a finite order. In the scale-variation approach, the residual scheme differences are then exploited to provide an estimate of the theory uncertainty by taking $\Delta f(x,\alpha) = \tilde{f}(x,\tilde{\alpha}) - f(x,\alpha)$. For example, at NLO and NNLO this yields

$$\begin{aligned}
\text{NLO:}\quad & \Delta f(x,\alpha) = b_0\hat{f}_1(x)\,\alpha^2 + b_1\hat{f}_1(x)\,\alpha^3 + \mathcal{O}(\alpha^4)\,,\\
\text{NNLO:}\quad & \Delta f(x,\alpha) = \{2b_0[\hat{f}_2(x) - b_0\hat{f}_1(x)] + b_1\hat{f}_1(x)\}\alpha^3 + \mathcal{O}(\alpha^4)\,.
\end{aligned} \tag{2.7}$$

JHEP12(2025)043

The resulting uncertainty estimate is automatically of the correct one higher order, $\Delta f(x,\alpha) \sim \mathcal{O}(\alpha^{n+1})$ at N^{n}LO, which makes this approach very convenient to use in practice.

However, the scale-variation approach suffers from well-known limitations. The parameters b_k, or in case of QCD the single parameter L in eq. (2.5), are a priori arbitrary and contain no information on the actual missing higher-order terms. At NLO for example there is nothing that guarantees that $b_0\hat{f}_1(x)$ is a good estimate of $\hat{f}_2(x)$ in any way. In fact, typically $f_2(x)$ will have a (much) more complicated x dependence than $f_1(x)$ and will not just be proportional to it. As a result, eq. (2.7) does not provide a very reliable uncertainty estimate. Typical pitfalls leading to an underestimation are when a lower-order $\hat{f}_{\leq n}(x)$ happens to be accidentally small (e.g. in the vicinity of zero-crossings) or when $f_{n+1}(x)$ contains new structures not present in $f_{\leq n}(x)$.

While one might be able to avoid these pitfalls of underestimation, there is a more severe limitation of scale variations, which is fundamental and unavoidable: The b_k (or L) are unphysical parameters. That is, they are artificially introduced via the scheme change, but they are not actual parameters of $f(x)$ with an existing but unknown true value. While there might be a certain value of L for which $\Delta f(x_i,\alpha)$ coincides with the true value of the missing higher-order term(s) for a specific x_i, this value of L will be different for different x_i. There may also be no (reasonable) value of L for which this is the case. Therefore, the resulting variation $\Delta f(x,\alpha)$ in eq. (2.7) *cannot* be interpreted as a fully correlated uncertainty like the variation due to an ordinary parameter whose uncertainty is being propagated. In other words, eq. (2.7) does not provide a correct parameterization of the x dependence of the missing higher-order terms, which however would be necessary to obtain the correct theory correlation structure in x [34].

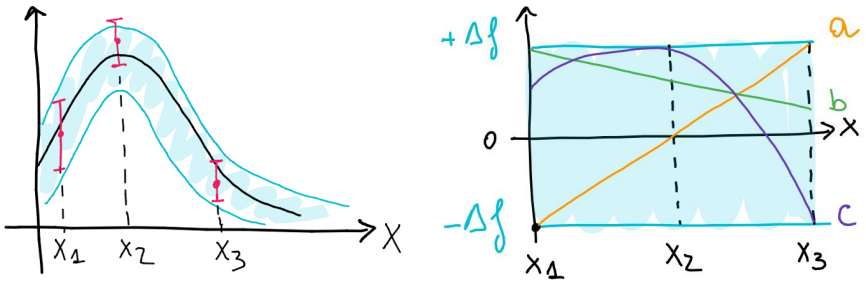
Oftentimes, including for our case of interest, predictions involve multiple scale parameters (e.g. due to making separate choices of the expansion parameter in different parts of the calculation). The standard approach to mitigate the danger of misestimation is by considering a variety of different variations of all relevant scales in some fashion to obtain a variety of different results $\tilde{f}_k(x)$. The uncertainty estimate is then obtained by taking the envelope over all variations[1]

$$\Delta f_k(x) = \tilde{f}_k(x) - f_{\text{default}}(x)\,, \qquad \Delta f(x) = \max_k \left|\Delta f_k(x)\right|\,. \tag{2.8}$$

The hope is that at any given x_i there will be at least one variation $\Delta f_k(x)$, not necessarily the same for all x_i, that gives a reasonable upper limit for the possible size of the missing higher-order terms. Taking the envelope then ensures that the total uncertainty does not scale with the number of considered variations, which is in principle arbitrary. Performing as many as possible different types of variations (as opposed to merely increasing the size of variation) will then generically improve the reliability of the estimate. This is particularly true in situations (including our case of interest) where individual $\Delta f_k(x)$ tend to cross through zero at some x_i and thus individually grossly underestimate the uncertainty around that point.

[1]Here we take the default result as the central value and the maximum absolute difference as its symmetric uncertainty. Other ways to construct the envelope are also used, further demonstrating the arbitrary nature of these prescriptions. However, the precise way of doing so is irrelevant for our discussion.

JHEP12(2025)043



	ρ_{12}	ρ_{13}	ρ_{23}
a	0	-1	0
b	$+1$	$+1$	$+1$
c	$+1$	-1	-1

Figure 1. Interpreting a differential spectrum by scanning scale variations. Left: The theory prediction (black line) with a theory uncertainty band (blue) from the envelope of scale variations compared to measurements (red) at three points $x_{1,2,3}$. Middle: The normalized uncertainty band (blue) with three possible scale variations (orange, green, violet) filling out the band. Right: The assumed correlations ρ_{ij} between the uncertainties at the points x_i and x_j resulting from a given scale variation. Note that for $\rho_{ij} = 0$ the uncertainty at either x_i or x_j must vanish.

However, even assuming that the final envelope $\Delta f(x)$ provides a reasonable theory uncertainty for the spectrum at any given x_i, the key question is how to propagate this uncertainty from the spectrum to the parameter of interest when performing an interpretation, which as discussed in section 2.1 requires the correct point-by-point correlations across the spectrum, i.e., the correlation between $\Delta f(x_i)$ and $\Delta f(x_j)$ for any two x_i and x_j. This is where we run afoul of the fundamental lack of correlations in the scale-variation approach.

A common method to "propagate" the scale-variation envelope from the spectrum to the parameters of interest is by scanning over scale variations. That is, the interpretation is repeated for all predictions $\tilde{f}_k(x)$ that fill out the envelope. The various different results obtained for the parameters of interest are then enveloped to give an estimate of the theory uncertainty on the parameters of interest. This method is however insufficient due to the fundamental limitations discussed above: Since the underlying scales are unphysical parameters, the resulting variations $\Delta f_k(x)$ do not correspond to well-defined fully correlated uncertainty components or sources. Rather they merely provide different ways to probe the possible size of the total theory uncertainty. This is exactly why we take their envelope instead of adding them in quadrature. In other words, by taking an envelope we explicitly acknowledge the fact that the individual $\Delta f_k(x)$, and the point-by-point correlation they would imply, are not particularly meaningful and that they do not correspond to well-defined uncertainty components. Otherwise we would combine them in quadrature. Therefore, scanning over scale variations does not provide the necessary breakdown of the total uncertainty into uncertainty components.

Instead, this scanning over enveloped variations amounts to trying out various correlation models for the same total uncertainty band, as illustrated in figure 1. We might then wonder what we can learn from this. First, it is important to realize that none of the trial variations actually provides a realistic correlation model, because as seen in figure 1, they are by construction always fully correlated, only allowing correlations $\rho_{ij} = \pm 1$ (or $\rho_{ij} = 0$ in case $\Delta f(x_{i,j}) = 0$). This is clearly problematic when one is sensitive to shape effects, where as discussed in section 2.1, the precise correlation near $\rho = 1$ completely determines the resulting

JHEP12(2025)043

shape uncertainty and thus the impact on the parameter of interest. Therefore any given variation with its implied correlation model can strongly underestimate (or overestimate) the resulting theory uncertainty on the parameter of interest. One might be "lucky" such that one trial variation sufficiently closely mimics the shape effect of the parameter of interest so as to yield a "conservative" uncertainty and avoid underestimation. However, this might just as well be very "unlucky" and yield a substantial overestimation.

In summary, scale variations provide no means to correctly address the problem of theory correlations. Therefore, when theory correlations are important, or even critical as in cases like ours when looking for shape effects, they cannot be used to reliably propagate the theory uncertainties from the original spectrum to the parameters of interest. One might wonder why this fundamental lack of theory correlations in scale variations is then not discussed more often. The main reasons might be that theory correlations are often either not as important or not recognized to be important, or simply because there has not been an alternative so far. There are however other cases, some of which quite similar to ours, where the lack of theory correlations in scale variations has been well recognized to be problematic or limiting. To mitigate these problems, the best that can be done with scale variations is to impose an explicit and case-specific correlation model on the theory uncertainty, see e.g. refs. [7, 41–52]. In comparison, by scanning over scale variations one basically relies on some implicit ad hoc theory correlation model.

2.3 Theory nuisance parameters

The limitations of scale variations have led to a variety of alternative methods being explored over the years [53–58] including the TNP approach [34, 59–64]. Among the motivations for the latter is to allow one to properly account for theory correlations.

The fundamental source of the perturbative theory uncertainty is not the choice of scale or perturbative scheme. Rather, the actual sources of uncertainty in any particular scheme are its missing unknown $\hat{f}_n(x)$. To be able to correctly account for the theory uncertainty, we have to actually include the (leading) source(s) of uncertainty in our predictions. To do so, compared to eq. (2.2) we have to include (at least) the next term in the series, which contains the dependence on the unknown series coefficient $f_n(x)$. This is the starting point of the TNP approach [34]. To obtain a perturbative prediction for $f(x)$ at order $\mathrm{N}^{m+k}\mathrm{LO}$, we include the true values of the first m series coefficients and in addition the next $k \geq 1$ terms whose coefficients are unknown. We are typically in a regime where the next order is the dominant source of uncertainty, in which case $k = 1$, so the predictions in eq. (2.2) turn into

$$\begin{aligned}
\mathrm{N}^{0+1}\mathrm{LO}: &\quad f(x,\alpha,\theta_1) = \hat{f}_0(x) + f_1(x,\theta_1)\,\alpha\,, \\
\mathrm{N}^{1+1}\mathrm{LO}: &\quad f(x,\alpha,\theta_2) = \hat{f}_0(x) + \hat{f}_1(x)\,\alpha + f_2(x,\theta_2)\,\alpha^2\,, \\
\mathrm{N}^{2+1}\mathrm{LO}: &\quad f(x,\alpha,\theta_3) = \hat{f}_0(x) + \hat{f}_1(x)\,\alpha + \hat{f}_2(x)\,\alpha^2 + f_3(x,\theta_3)\,\alpha^3\,.
\end{aligned} \tag{2.9}$$

Here, the (considered to be) unknown series coefficients $f_n(x,\theta_n)$ are parameterized in terms of *theory nuisance parameters* θ_n, where for simplicity we let $\theta_n \equiv \{\theta_{n,i}\}$ denote an appropriate set of parameters $\theta_{n,i}$ that are by construction scalars [34].

JHEP12(2025)043

A key condition that the TNP parameterization $f_n(x, \theta_n)$ has to satisfy is that there must exist true values $\hat{\theta}_n$ for which the true value $\hat{f}_n(x)$ is reproduced,

$$\hat{f}_n(x) = f_n(x, \hat{\theta}_n)\,. \tag{2.10}$$

This condition implies that the TNPs are true parameters of the perturbative series. As for any other ordinary parameter, we can assign to them a best estimate with some uncertainty,

$$\theta_n = u_n \pm \Delta u_n\,, \tag{2.11}$$

which can come from theory considerations or experimental constraints or both. Without loss of generality we will assume that, prior to applying any experimental constraints, the θ_n are chosen such that based on theory information only we have $u_n = 0$ and $\Delta u_n \equiv \Delta\theta_n = 1$.[2]

The central prediction then corresponds to setting the θ_n to their central values. Note that in general we can have $f_n(x, \theta_n = 0) \neq 0$, so the central prediction at say $\mathrm{N}^{2+1}\mathrm{LO}$ does not necessarily coincide with the previous $\mathrm{N}^2\mathrm{LO}$ prediction without TNPs. This is not a bug but a feature, because the TNP parameterization allows us to correctly include contributions that are already known, e.g. from lower-order terms. On the other hand, eq. (2.10) implies that when evaluated at the true values $\hat{\theta}_n$, the $\mathrm{N}^{2+1}\mathrm{LO}$ result agrees with the $\mathrm{N}^3\mathrm{LO}$ result.

The condition in eq. (2.10) ensures that the θ_n are well-defined common parameters for the predictions at different x. Their unknown value therefore represents a common source of uncertainty whose impact is 100% correlated across all x. At the same time, different individual $\theta_{n,i}$ are a priori mutually independent, so they can be varied separately and the impacts of different $\theta_{n,i}$ are a priori fully uncorrelated. The theory uncertainty can thus be evaluated by propagating the uncertainties Δu_n through the interpretation to the parameters of interest using any desired standard method for error propagation and combination. In particular, they can be consistently treated and profiled in fits to data like any other nuisance parameters. In this case, their post-fit uncertainties can of course become correlated, which one has to properly account for when combining uncertainties.

Theory-based constraints inevitably involve some human prejudice (unless they come from calculations that are either exact or numerically approximate with a well-defined numerical uncertainty). As shown in ref. [34], it is nevertheless possible to obtain reliable and statistically meaningful theory constraints. When the TNPs are constrained by measurements, we can furthermore decide how much to rely on any given theory constraint by choosing how much to impose it as an additional constraint on the fit.

In summary, the TNP approach provides truly parametric theory uncertainties. As a result, it correctly captures the point-by-point correlations in x by providing the necessary breakdown of the total theory uncertainty into mutually independent uncertainty components. In essence, by requiring eq. (2.10), the TNP parameterization must encode the correct dependence of the missing higher order on the observable x, which is precisely what is needed to encode the correct correlation structure. The general principles and strategies for constructing suitable parameterizations are discussed in ref. [34]. The relevant TNPs for our case at hand are discussed in section 3.2.3.

[2] Starting from a theory constraint $u_n \pm \Delta u_n$ for θ_n, we can always reparameterize it as $\theta'_n = (\theta_n - u_n)/\Delta u_n$, such that the corresponding constraint on θ'_n is 0 ± 1.

JHEP12(2025)043

3 Theory description and pseudodata setup

In this section, we summarize the theoretical description of the q_T spectrum which we use for our study. After discussing the overall theory requirements in section 3.1, we discuss the relevant perturbative contributions in section 3.2 and the nonperturbative effects in section 3.3. In section 3.4, we discuss our Asimov fit setup.

3.1 Theory requirements

The transverse momentum (q_T) distribution of dilepton pairs from Drell-Yan production is one of the most precisely measured observables at hadron colliders, with the latest LHC measurements reaching subpercent precision over a broad range of q_T values. This high level of experimental precision puts high demands on the theoretical description. For a reliable (sub)percent-level extraction of $\alpha_s(m_Z)$ from the small-q_T spectrum, not only must the theory predictions be accurate to the (sub)percent level, but as already discussed the theory uncertainties and correlations must also be well understood and reliably quantified.

The differential cross section in q_T can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T} = \frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}q_T} + \frac{\mathrm{d}\sigma_{\text{nons}}}{\mathrm{d}q_T}\,. \tag{3.1}$$

Here, $\mathrm{d}\sigma^{(0)}$ denotes the leading-power contribution, which dominates the spectrum in the small-q_T region, $q_T \ll Q$, where $Q \equiv \sqrt{q^2} \sim m_Z$ is the dilepton invariant mass. The nonsingular term, $\mathrm{d}\sigma_{\text{nons}}$, contains all remaining contributions, which are power-suppressed by $\mathcal{O}(q_T^2/Q^2)$ and thus subdominant relative to $\mathrm{d}\sigma^{(0)}$.

The largest sensitivity to $\alpha_s(m_Z)$ comes from the small-q_T region. Since $\mathrm{d}\sigma^{(0)}$ dominates here, it is also the dominant source of perturbative and nonperturbative uncertainties and therefore the main focus of our study. At each order in α_s, $\mathrm{d}\sigma^{(0)}$ contains logarithms $\ln^n(q_T/Q)$, which must be resummed to obtain precise and perturbatively stable predictions. In parallel, it is affected by nonperturbative corrections of $\mathcal{O}(\Lambda_{\text{QCD}}^n/q_T^n)$, which shape the q_T spectrum at very low $q_T \lesssim 10\,\text{GeV}$. Their associated uncertainties therefore play an important role for extracting $\alpha_s(m_Z)$. A complete treatment of perturbative uncertainties and a proper parameterization of nonperturbative effects in the leading-power contribution are thus indispensable for a reliable α_s extraction at low q_T.

Another key ingredient is the treatment of parton distribution functions (PDFs). PDFs influence both the normalization and shape of the spectrum. They are extracted from global fits that themselves depend on assumptions about α_s. This introduces a nontrivial correlation between the value of $\alpha_s(m_Z)$ one extracts and the underlying PDFs used in the prediction [35–39]. Thus, a consistent extraction must account for PDF uncertainties and, ideally, include the interplay between PDFs and α_s in a consistent fashion. A more detailed study of PDF uncertainties in the extraction of $\alpha_s(m_Z)$ from the Z q_T spectrum will be presented in a forthcoming paper [40].

Beyond these leading contributions, achieving a (sub)percent-level theoretical description requires incorporating various subleading effects. One such effect is the nonsingular contribution in eq. (3.1). Similarly, quark mass corrections scaling as m_q^2/q_T^2, and QED and electroweak effects become relevant at this level of precision (see e.g. refs. [65–70]). Neglecting

JHEP12(2025)043

these subleading effects can cause a substantial bias in the extracted value of $\alpha_s(m_Z)$, so it is important to include them in a fit to real data. However, once they are included, we do not expect them to have a major influence on the uncertainty in $\alpha_s(m_Z)$. We can therefore consistently neglect them in our study as further discussed in section 3.4.

Summarizing, for our study we focus on the leading-power contribution, its resummed perturbative prediction and associated uncertainty as discussed in section 3.2, and its nonperturbative effects and associated uncertainty as discussed in section 3.3. Among the subleading effects, the nonsingular contribution is the largest. We therefore briefly study its impact in appendix B.1, showing that it can indeed be neglected for the purposes of our main analysis. In appendix A.3 we also briefly comment on the potential bias from neglecting quark masses.

3.2 Resummation of perturbative contributions at small q_T

We use the SCET resummation framework including TNPs of refs. [32, 34, 52]. In the following, we focus on the aspects relevant to our study and refer to these references for further details. The numerical results for our study are obtained with the help of SCETlib [71] and its implementation of q_T resummation with TNPs [25, 32, 34, 52, 72].

We denote the four-momentum of the vector boson by q^μ, its invariant mass by $Q \equiv \sqrt{q^2}$, its total rapidity by Y, and its transverse momentum by $q_T = |\vec{q}_T|$. We are interested in the leading-power cross section fully differential in Q, Y and $\vec{q}_T$, collectively denoted as $\mathrm{d}\sigma^{(0)}/\mathrm{d}^4q$. In our framework, its resummation is based on its factorization theorem [73–84], which is written as

$$\begin{aligned}\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}^4q} = \frac{1}{2E_{\mathrm{cm}}^2} L_{VV'}(q^2) \sum_{a,b} H_{VV'\,ab}(q^2,\mu) \\ \times \int \frac{\mathrm{d}^2\vec{b}_T}{(2\pi)^2}\, e^{\mathrm{i}\vec{b}_T\cdot\vec{q}_T}\, \tilde{B}_a(x_a,b_T,\mu,\nu/Q)\, \tilde{B}_b(x_b,b_T,\mu,\nu/Q)\tilde{S}(b_T,\mu,\nu)\,.\end{aligned} \tag{3.2}$$

Here, E_{cm} is the hadronic center-of-mass energy and $L_{VV'}(q^2)$ denotes the leptonic tensor associated with the final-state vector boson $VV' = \{\gamma\gamma, \gamma Z, Z\gamma, ZZ\}$, which is not affected by QCD corrections. The sum runs over parton flavors a, b, and the remaining ingredients are the hard function $H_{VV'\,ab}(q^2,\mu)$, the beam functions $\tilde{B}_{a,b}(x_{a,b}, b_T, \mu, \nu/Q)$, and the soft function $\tilde{S}(b_T,\mu,\nu)$. The variable $b_T = |\vec{b}_T|$ is the Fourier-conjugate of q_T and $x_{a,b} = (Q/E_{\mathrm{cm}})\exp^{\pm Y}$ encode the dependence on the rapidity and collider energy.

3.2.1 Resummation

Each of the perturbative ingredients in eq. (3.2) obeys a renormalization group evolution (RGE), which determines its scale and kinematic dependence to all orders in α_s. As a result, the factorization theorem predicts the full q_T dependence and, for fixed $x_{a,b}$, also the Q dependence. Although it does not predict the full functional form in $x_{a,b}$, it simplifies this dependence to the product of two one-dimensional beam functions. We now describe each ingredient in turn.

The **hard function** $H_{VV'\,ab}$ encodes virtual corrections to the underlying hard partonic interaction process $ab \to V$. It can be extracted as the infrared-finite part of the corresponding

JHEP12(2025)043

form factors. Explicit expressions for $H_{VV'ab}$ and $L_{VV'}$ in terms of the underlying SCET Wilson coefficients $C_q(q^2,\mu)$ are given in ref. [52]. The C_q's RGE equation is given by

$$\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\ln C_q(q^2,\mu) = \Gamma^q_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{-q^2-\mathrm{i}0}{\mu^2} + 2\gamma^q_C[\alpha_s(\mu)]\,, \tag{3.3}$$

where $\Gamma^q_{\text{cusp}}(\alpha_s)$ and $\gamma_C(\alpha_s)$ are the cusp and noncusp anomalous dimensions, respectively.

The **soft function** captures wide-angle soft radiation. It additionally depends on the rapidity renormalization scale ν. Its scale dependence is governed by the following RGE system

$$\begin{aligned}
\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\ln\tilde{S}(b_T,\mu,\nu) &= 4\Gamma^q_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{\mu}{\nu} + \tilde{\gamma}_S[\alpha_s(\mu)]\,,\\
\nu\frac{\mathrm{d}}{\mathrm{d}\nu}\ln\tilde{S}(b_T,\mu,\nu) &= \tilde{\gamma}_\nu(b_T,\mu)\,,\\
\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\tilde{\gamma}_\nu(b_T,\mu) &= -4\Gamma^q_{\text{cusp}}[\alpha_s(\mu)]\,.
\end{aligned} \tag{3.4}$$

The rapidity anomalous dimensions $\tilde{\gamma}_\nu(b_T,\mu)$ encodes the nontrivial b_T dependence, which in turn is governed by its own μ RGE.

The **beam functions** $\tilde{B}_i(x,b_T,\mu,\nu/Q)$ describe the extraction of a parton i carrying momentum fraction x at some transverse-momentum conjugate b_T from an unpolarized hadron. They evolve according to

$$\begin{aligned}
\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\ln\tilde{B}_q(x,b_T,\mu,\nu/Q) &= 2\Gamma^q_{\text{cusp}}[\alpha_s(\mu)]\ln\frac{\nu}{Q} + \tilde{\gamma}_B[\alpha_s(\mu)]\,,\\
\nu\frac{\mathrm{d}}{\mathrm{d}\nu}\ln\tilde{B}_q(x,b_T,\mu,\nu/\omega) &= -\frac{1}{2}\tilde{\gamma}_\nu(b_T,\mu)\,,\\
\mu\frac{\mathrm{d}}{\mathrm{d}\mu}\tilde{\gamma}_\nu(b_T,\mu) &= -4\Gamma^q_{\text{cusp}}[\alpha_s(\mu)]\,.
\end{aligned} \tag{3.5}$$

While their RGE fully determines their b_T and Q dependence, their dependence on x is not predicted and only arises via the boundary condition at canonical scales $\mu = b_0/b_T$ and $\nu = Q$, defined as

$$\tilde{b}_i(x,\alpha_s) \equiv \tilde{B}_i(x,b_T,\mu=b_0/b_T,\nu/Q=1)\,. \tag{3.6}$$

In the perturbative region $q_T \sim 1/b_T \gg \Lambda_{\text{QCD}}$, this boundary condition is computed in terms of standard PDFs at the canonical scale $\mu_f^{\text{can}} = b_0/b_T$,

$$\tilde{b}_{i,n}(x) = \sum_j \int\frac{\mathrm{d}z}{z}\,\tilde{I}_{ij,n}(z)\,f_j\left(\frac{x}{z},\mu_f^{\text{can}}\right)\,, \tag{3.7}$$

where $\tilde{I}_{ij,n}(z)$ are perturbative matching kernels. The x dependence of the beam function is thus given by the Mellin convolution of these kernels with the PDFs.

The resummation proceeds by evaluating each of the perturbative ingredients in eq. (3.2) in fixed order at their natural (canonical) μ and ν scales, where they are free of large logarithmic corrections. These are called the boundary conditions. Starting from their boundary conditions, the functions are then evolved to common overall μ and ν scales by

JHEP12(2025)043

solving their respective RGEs. In the combination of all RGEs and upon Fourier-transforming back to momentum space, the logarithms of q_T/Q are resummed to all orders in perturbation theory. The canonical scales in b_T space are given by

$$\mu_B^{\rm can} = \mu_S^{\rm can} = \nu_S^{\rm can} = \mu_0^{\rm can} = \mu_f^{\rm can} = \frac{b_0}{b_T}\,, \qquad \nu_B^{\rm can} = \mu_H^{\rm can} = Q\,, \tag{3.8}$$

where $\mu_{S,B}$ and $\nu_{S,B}$ are the scales of the soft and beam functions, μ_0 is the scale at which the boundary condition of the rapidity anomalous dimension is evaluated, μ_f is the scale at which the matching to PDFs in eq. (3.7) is performed, and $b_0 = 2e^{-\gamma_E} \approx 1.12292$ is a conventional factor.

Since we do not include nonsingular power corrections in our main study, we do not discuss here the procedure for switching off the resummation at large q_T to recover the fixed-order result. For that purpose, the canonical scales in eq. (3.8) are modified using hybrid profile scales [85] (see ref. [32] for details), defined as

$$\begin{aligned}
&\mu_H = Q\,,\\
&\mu_B = Q\, f_{\rm run}\left(\frac{q_T}{Q}, \frac{1}{Q}\mu_*\left(\frac{b_0}{b_T}, \mu_B^{\rm min}\right)\right), \qquad \nu_B = Q\,,\\
&\mu_S = Q\, f_{\rm run}\left(\frac{q_T}{Q}, \frac{1}{Q}\mu_*\left(\frac{b_0}{b_T}, \mu_S^{\rm min}\right)\right), \qquad \nu_S = Q\, f_{\rm run}\left(\frac{q_T}{Q}, \frac{1}{Q}\mu_*\left(\frac{b_0}{b_T}, \nu_S^{\rm min}\right)\right),\\
&\mu_0 = \mu_*\left(\frac{b_0}{b_T}, \mu_0^{\rm min}\right), \qquad \mu_f = Q\, f_{\rm run}\left(\frac{q_T}{Q}, \frac{1}{Q}\mu_*\left(\frac{b_0}{b_T}, \mu_f^{\rm min}\right)\right),
\end{aligned} \tag{3.9}$$

where the profile function $f_{\rm run}(x, y)$ ensures a smooth transition between the canonical scales in the resummation region, $x = q_T/Q \ll 1$, and the fixed-order region, $x = q_T/Q \sim 1$. The most relevant point for our purposes is that the transition only starts at a transition point x_1 (by default $x_1 = 0.3$) and below it $f_{\rm run}(x \le x_1, y) = y$, so the default scales in eq. (3.9) reduce to

$$\mu_{B,S,0,f} = \mu_*\left(\frac{b_0}{b_T}, \mu_{B,S,0,f}^{\rm min}\right), \quad \nu_S = \mu_*\left(\frac{b_0}{b_T}, \nu_S^{\rm min}\right), \quad \nu_B = \mu_H = Q \qquad (\text{for } q_T \le x_1 Q)\,. \tag{3.10}$$

Hence, in this "canonical" q_T region, the default scales we use only differ from their canonical values in eq. (3.8) by the μ_* function, which simply enforces a smooth lower bound on the scales to avoid the nonperturbative regime at large b_T, where the QCD Landau pole would be encountered, with

$$\mu_B^{\rm min} = \mu_S^{\rm min} = \mu_0^{\rm min} = 1\,\text{GeV}\,, \qquad \mu_f^{\rm min} = 1.4\,\text{GeV}\,, \qquad \nu_S^{\rm min} = 0\,. \tag{3.11}$$

3.2.2 Scale variations

We now briefly summarize the scale-variation setup in SCETlib, referring to ref. [32] for more details. The central scale choices are equivalent to the default scales in eqs. (3.9) and (3.10). The scale variations are performed by varying the scales around their central values with

JHEP12(2025)043

multiplicative factors. In the canonical region $q_T \leq x_1 Q$, these are

$$\begin{aligned}
\mu_H &= 2^{w_{\rm FO}}\, Q\,, \\
\mu_B &= 2^{w_{\rm FO}}\, f_{\rm vary}^{v_{\mu_B}}\, \mu_*\left(\frac{b_0}{b_T}, \frac{\mu_B^{\rm min}}{2^{w_{\rm FO}} f_{\rm vary}^{v_{\mu_B}}}\right), \qquad \nu_B = 2^{w_{\rm FO}}\, f_{\rm vary}^{v_{\nu_B}}\, Q\,, \\
\mu_S &= 2^{w_{\rm FO}}\, f_{\rm vary}^{v_{\mu_S}}\, \mu_*\left(\frac{b_0}{b_T}, \frac{\mu_S^{\rm min}}{2^{w_{\rm FO}} f_{\rm vary}^{v_{\mu_S}}}\right), \qquad \nu_S = 2^{w_{\rm FO}}\, f_{\rm vary}^{v_{\nu_S}}\, \mu_*\left(\frac{b_0}{b_T}, \frac{\nu_S^{\rm min}}{2^{w_{\rm FO}} f_{\rm vary}^{v_{\nu_S}}}\right), \\
\mu_f &= 2^{v_{\mu_f}}\, \mu_*\left(\frac{b_0}{b_T}, \frac{\mu_f^{\rm min}}{2^{v_{\mu_f}}}\right).
\end{aligned} \tag{3.12}$$

The same overall factors for each scale also appear inside its μ_* function, which ensures that the effective minimum scales at large b_T are unaffected by the variations. The boundary scale μ_0 is held fixed, as its variation is effectively captured already by those of $\nu_{B,S}$. The different variation factors are used to perform different classes of variations, which we describe in turn.

The **resummation uncertainty $\boldsymbol{\Delta_{\rm resum}}$** is obtained by varying the beam and soft scales via $v_{\mu_B}, v_{\nu_B}, v_{\mu_S}, v_{\nu_S}$ by taking $v_i = \{-1, 0, +1\}$ with $v_i = 0$ corresponding to the central scale. Here, $f_{\rm vary} \equiv f_{\rm vary}(q_T/Q)$ is a function of q_T/Q that approaches 2 for $q_T \to 0$ and 1 for $q_T/Q \to 1$, i.e., it smoothly turns off the variation outside the resummation region. We perform 36 variations of suitable combinations of the v_i and then take their maximum envelope as the resulting final uncertainty $\Delta_{\rm resum}$.

For the **fixed-order uncertainty $\boldsymbol{\Delta_{\rm FO}}$**, we simultaneously vary the scales of the hard, beam, and soft functions by an overall factor of 2 by taking $w_{\rm FO} = \{-1, 0, +1\}$. The final uncertainty is again taken as the maximum envelope of these two variations. This variation only probes the fixed-order boundary conditions without changing the resummed logarithms, hence its name.

The **uncertainty related to the DGLAP running of the PDFs $\boldsymbol{\Delta_f}$** is obtained by varying the PDF scale μ_f by taking $v_{\mu_f} = \{-1, 0, +1\}$ and taking the maximum envelope of the two variations.

Finally, the **matching uncertainty $\boldsymbol{\Delta_{\rm match}}$** is given by the maximum envelope of the variations of the transition points x_i inside the $f_{\rm run}$ function determining the transition from the canonical to the fixed-order region. There are three points (x_1, x_2, x_3) determining the start, midpoint, and endpoint of the transition. Their central values are $(0.3, 0.6, 0.9)$ and their variations are

$$(x_1, x_2, x_3) \in \{(0.4, 0.75, 1.1), (0.2, 0.45, 0.7), (0.4, 0.55, 0.7), (0.2, 0.65, 1.1)\}\,. \tag{3.13}$$

These different classes of variations are specifically chosen to probe separate aspects of the perturbative series. They are then considered as independent uncertainties and added in quadrature to obtain the total perturbative uncertainty,

$$\Delta_{\rm pert} = \sqrt{\Delta_{\rm FO}^2 + \Delta_f^2 + \Delta_{\rm match}^2 + \Delta_{\rm resum}^2}\,. \tag{3.14}$$

In particular, the μ_f variation is not enveloped with the other scale variations, but is instead treated as a separate component. The reason is that it probes the uncertainty due to DGLAP running, which can be considered as a separate source of perturbative uncertainty.

JHEP12(2025)043

3.2.3 TNPs for q_T resummation

We now briefly summarize the theory nuisance parameters for q_T resummation, referring to section 6 of ref. [34] for more details. The leading-power spectrum in eq. (3.2) is fully determined by the hard, beam and soft functions, $F = \{H, B, S\}$. Each of these function satisfies a coupled system of RGEs, given in eq. (3.3), eq. (3.4) and eq. (3.5). Their solution can be written schematically as

$$F(\alpha_s^{\text{can}}, L) = F(\alpha_s^{\text{can}}) \exp\left\{ \int_0^L \mathrm{d}L' \Big[\Gamma_{\text{cusp}}(\alpha_s(L'))L' + \gamma_F(\alpha_s(L')) \Big] \right\}, \tag{3.15}$$

where $L \equiv \ln(\mu/\mu_F^{\text{can}})$ and $\alpha_s^{\text{can}} \equiv \alpha_s(\mu_F^{\text{can}})$. It depends on the boundary condition $F(\alpha_s)$ (defined at strictly canonical scale for $L = 0$) as well as the cusp and noncusp anomalous dimensions, $\Gamma_{\text{cusp}}(\alpha_s)$ and $\gamma_F(\alpha_s)$. These perturbative series contain the complete perturbative information necessary for the leading-power spectrum. We can therefore treat each of them following the TNP approach of ref. [34] and as outlined in section 2.3, parameterizing their individual missing higher-order coefficients with TNPs. Since the q_T dependence is completely predicted by the factorization theorem in terms of these, this treatment correctly encodes the point-by-point correlations in q_T.

Denoting generically all boundary conditions as $F(\alpha_s)$ and all anomalous dimensions as $\gamma(\alpha_s)$, we write

$$\begin{aligned} F(\alpha_s) &= 1 + \sum_{n=1} F_n \left(\frac{\alpha_s}{4\pi}\right)^n, \\ \gamma(\alpha_s) &= \sum_{n=0} \gamma_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}, \end{aligned} \tag{3.16}$$

and parametrize each coefficient (for fixed $n_f = 5$) as

$$F_n(\theta_n^F) = 4^n C_n (n-1)!\, \theta_n^F, \tag{3.17}$$

$$\gamma_n(\theta_n^\gamma) = C_{n+1} 4^{n+1}\, \theta_n^\gamma. \tag{3.18}$$

Here, C_n is the relevant leading n-loop coefficient, which in our case is given by $C_n = C_F C_A^{n-1}$, and θ_n^F, θ_n^γ are the TNPs.

The generalized order counting for resummation including TNPs is denoted as N^{m+k}LL and defined in analogy to N^{m+k}LO in section 2.3 as follows: A given perturbative resummation order, N^nLL, is uniquely and well defined by including all underlying perturbative series in the RGE to a certain order in α_s, namely the boundary conditions to $(n-1)$ loops, the noncusp anomalous dimensions to n loops, and the cusp anomalous dimensions and QCD beta functions to $n+1$ loops. At N^{m+k}LL we use the complete resummation structure at N^nLL with $n = m + k$ with the true values for all coefficients relevant up to N^mLL and the TNP parameterization for the last k coefficients of each series. For example, eq. (3.15)

JHEP12(2025)043

at N^{3+1}LL takes the form

$$\text{N}^{3+1}\text{LL:}\qquad F(\alpha_s^{\text{can}},L)=\Bigg[1+\sum_{k=1}^{2}\hat{F}_k\left(\frac{\alpha_s^{\text{can}}}{4\pi}\right)^k+F_3(\theta_3^F)\left(\frac{\alpha_s^{\text{can}}}{4\pi}\right)^3\Bigg]\exp\Bigg\{\int_0^L \mathrm{d}L'\Bigg[\sum_{k=0}^{3}\hat{\Gamma}_k\Big[\frac{\alpha_s(L')}{4\pi}\Big]^{k+1}L'+\Gamma_4(\theta_4^\Gamma)\Big[\frac{\alpha_s(L')}{4\pi}\Big]^5 L' +\sum_{k=0}^{2}\hat{\gamma}_{F,k}\Big[\frac{\alpha_s(L')}{4\pi}\Big]^{k+1}+\gamma_{F,3}(\theta_3^{\gamma_F})\Big[\frac{\alpha_s(L')}{4\pi}\Big]^4\Bigg]\Bigg\}. \tag{3.19}$$

Furthermore, in the full resummed cross section, the product of all boundary conditions is expanded in α_s including cross terms up to $\mathcal{O}(\alpha_s^3)$ as relevant for N^{4}LL. Note that despite the fact that the highest k coefficients at N^{m+k}LL are treated as unknown, due to having the complete $(m+k)$th-order structure, this order still contains more perturbative information than N^{m}LL, so the perturbative accuracy at N^{m+k}LL is generically higher than at N^{m}LL.

For our analysis of the Drell-Yan q_T spectrum, we consider a minimal set of seven independent TNPs,

$$\theta_n^\gamma:\gamma\in\{\Gamma,\gamma_\mu,\gamma_\nu\}\,,\qquad \theta_n^F:F\in\{H,S,B_{qq},B_{qg}\}\,, \tag{3.20}$$

associated with the following independent perturbative series and thus sources of uncertainty: The cusp anomalous dimension (Γ_{cusp}), the noncusp anomalous dimensions (γ_μ, γ_ν), and the boundary conditions of the hard (H), soft (S) and beam (B_{qq}, B_{qg}) functions. Their detailed normalization conventions are discussed in ref. [34]. For the beam functions in particular, based on eq. (3.7), we adopt the parameterization

$$\tilde{I}_{ij,n}(z,\theta_n^{B_{ij}})=\frac{3}{2}\,\theta_n^{B_{ij}}\,\hat{\tilde{I}}_{ij,n}(z)\,, \tag{3.21}$$

which uses their known z dependence multiplied by an overall prefactor of 3/2 to be conservative. With this parameterization we effectively treat their shape as known and their overall normalization as unknown. This choice is a compromise. It means while we do not yet account for the uncertainty associated with their shape, we have the correct correlations in z and thereby in Y for the dominant overall normalization uncertainty we do consider. Since the dominant partonic channels for Z boson production are $ij=\{qq,qg\}$, we introduce only two TNPs for the beam boundary conditions. For simplicity, singlet contributions that appear at higher orders and only have a very minor impact are not considered separately. Specifically, we use a single effective TNP, $\theta_n^{B_{qq}}$, which varies all qq channels together,

$$\theta_n^{B_{qq}}\equiv\theta_n^{B_{qqV}}\equiv\theta_n^{B_{q\bar{q}V}}\equiv\theta_n^{B_{qqS}}\equiv\theta_n^{B_{qq\Delta S}}\,, \tag{3.22}$$

and a separate TNP, $\theta_n^{B_{qg}}$, for the qg channel.

As explained in section 2.3 and following ref. [34], each nuisance parameter θ_n is modelled as a Gaussian-distributed random variable which by default has zero mean,

$$\theta_n=0\pm\Delta\theta_n\,, \tag{3.23}$$

with $\Delta\theta_n$ setting the width of the distribution and hence the size of the variation. Unless otherwise stated, we take $\Delta\theta_n=1$ as our default theory constraint, corresponding to a 68%

JHEP12(2025)043

theory confidence level. We also explore the impact of relaxing the theory constraint in section 4.2.4. The default (pre-fit) result at $\mathrm{N}^{m+k}\mathrm{LL}$ is then given by the default resummed result as discussed in section 3.2.1 with all TNPs that enter it set to zero. Its uncertainty is obtained by varying each TNP by $\Delta\theta_n$ and adding all variations in quadrature (since by default all TNPs are considered independent). After fitting the TNPs to data, the post-fit result for the resummed spectrum is then given by using the post-fit central values and uncertainties of all TNPs.

Finally, it is important to note that the above minimal set of TNPs does not yet capture all potential sources of uncertainties. As discussed in ref. [34], additional sources are the full set of partonic channels in the beam functions as well as their functional dependence, the DGLAP splitting functions, singlet contributions to the hard function, and the QCD β function. These are currently not accounted for and are left for future improvements. Among these, the latter two are likely to be irrelevant while the uncertainties associated with the DGLAP splitting functions may prove to be the most relevant.

3.3 Nonperturbative effects

The q_T spectrum is affected by nonperturbative effects that scale as $\Lambda^2_{\mathrm{QCD}}/q_T^2$, and which are expected to become relevant at the few-percent level for $q_T \lesssim 10$ GeV [32]. In the literature a variety of phenomenological models have been used to incorporate nonperturbative effects, typically involving a small number of free parameters fitted to data. Since this region of moderately small q_T is also where the dominant sensitivity to $\alpha_s(m_Z)$ comes from, we require a correct parameterization of nonperturbative effects to ensure a reliable theoretical description in this region with correct point-by-point correlations of the nonperturbative uncertainties.

In this section, we briefly discuss the nonperturbative approach used in SCETlib [32], which is based on performing a systematic OPE expansion of nonperturbative effects. Next, we discuss the specific nonperturbative parameterization we employ for our study here.

3.3.1 General discussion

The nonperturbative effects enter via the beam and soft functions appearing in eq. (3.2). To discuss their nonperturbative contributions, it is convenient to combine their product in b_T space into conventional transverse-momentum dependent PDFs (TMD PDFs), defined as

$$\tilde{f}_i(x, b_T, \mu, \zeta) = \tilde{B}_i\left(x, b_T, \mu, \nu/\sqrt{\zeta}\right)\sqrt{\tilde{S}(b_T, \mu, \nu)}\,. \tag{3.24}$$

Here, ζ is the so-called Collin-Soper scale, which now takes over the role of the rapidity scale ν, whose dependence cancels on the right-hand side. Similarly, the combined rapidity RGE of the beam and soft functions gets replaced by the ζ evolution of the TMD PDF,

$$\tilde{f}_i(x, b_T, \mu, \zeta) = \tilde{f}_i(x, b_T, \mu, \zeta_0)\exp\left[\frac{1}{2}\tilde{\gamma}_\zeta(b_T, \mu)\ln\frac{\zeta}{\zeta_0}\right]. \tag{3.25}$$

The physical value of the final scale ζ in the resummed cross section is $\zeta = Q^2$, as seen in eq. (3.2). The boundary scale ζ_0 is equivalent to ν_S^2 and its canonical value is $\sqrt{\zeta_0} = b_0/b_T$, which eliminates the potentially large rapidity logarithms $\ln(\zeta_0 b_T^2/b_0^2)$ in the TMD PDF's

JHEP12(2025)043

boundary condition. The associated anomalous dimension is known as the Collins-Soper (CS) kernel, $\tilde{\gamma}_\zeta(b_T,\mu)$, which is simply related to the rapidity anomalous dimension,

$$\tilde{\gamma}_\zeta(b_T,\mu) = \frac{1}{2}\tilde{\gamma}_\nu(b_T,\mu)\,. \tag{3.26}$$

There are in total two distinct sources of nonperturbative contributions, namely to the TMD PDF boundary condition and the CS kernel, which can be written as [86],

$$\begin{aligned}\tilde{\gamma}_\zeta(b_T,\mu) &= \tilde{\gamma}_\zeta^{\text{pert}}(b^*(b_T),\mu) + \tilde{\gamma}_\zeta^{\text{np}}(b_T)\,,\\ \tilde{f}_i(x,b_T,\mu,\zeta_0) &= \tilde{f}_i^{\text{pert}}(x,b_T,\mu,\zeta_0)\tilde{f}_i^{\text{np}}(x,b_T,\zeta_0)\,.\end{aligned} \tag{3.27}$$

Here, the quantities on the left-hand side are defined in full QCD. The separation into perturbative and nonperturbative parts on the right-hand side is well defined, but a priori not unique. Namely, the exact definition of the nonperturbative pieces, $\tilde{f}^{\text{np}}$ and $\tilde{\gamma}_\zeta^{\text{np}}$, is implicitly determined by the exact definition of the perturbative pieces, $\tilde{f}^{\text{pert}}$ and $\tilde{\gamma}_\zeta^{\text{pert}}$, including all choices of boundary scales and cutoff prescriptions used to avoid the Landau pole, as discussed below.

The nonperturbative contribution to the CS kernel, $\tilde{\gamma}_\zeta^{\text{np}}(b_T)$, is universal, like the CS kernel (or rapidity anomalous dimension) itself, in that it only depends on whether the considered process is quark or gluon initiated (which for simplicity is suppressed in our notation). In particular, it is universal across processes such as Drell-Yan production (including Z and W bosons) and semi-inclusive deep inelastic scattering. It can also be accessed via lattice QCD calculations [87–92]. In principle, this allows one to obtain information on $\tilde{\gamma}_\zeta^{\text{np}}(b_T)$ by comparing to the lattice results, for which a proper treatment of massive quark corrections is however essential [66].

Unlike the CS kernel, the nonperturbative TMD PDF boundary condition, $\tilde{f}_i^{\text{np}}(x,b_T)$, is not universal to all quarks. Like the TMD PDF itself, in addition to b_T it also depends on the flavor i and Bjorken x of the interacting parton. It can be thought of as describing the intrinsic transverse momentum of the partons inside the proton.

At large $b_T \sim 1/\Lambda_{\text{QCD}}$, $\tilde{f}_i^{\text{np}}(x,b_T)$ and $\tilde{\gamma}_\zeta^{\text{np}}(b_T)$ are genuine nonperturbative functions of b_T. However, the relevant region of interest for our purposes is the spectrum for $q_T \gtrsim 1\,\text{GeV}$, which is determined by much smaller b_T, namely where $1/b_T \sim q_T \gg \Lambda_{\text{QCD}}$ is still in the perturbative domain. In this region, nonperturbative effects are not negligible, but we can systematically expand them in an operator product expansion (OPE) in powers of $\Lambda_{\text{QCD}} b_T \ll 1$ [75, 76, 93–97] (see also ref. [98]). The OPE of the final TMD PDF entering the factorized cross section takes the form [97]

$$\tilde{f}_i(x,b_T,\mu,Q) = \tilde{f}_i^{(0)}(x,b_T,\mu,Q)\left\{1 + b_T^2\left[\Lambda_{2,i}(x) + \lambda_2^\zeta \ln\frac{b_T Q}{b_0}\right] + \mathcal{O}(\Lambda_{\text{QCD}}^4 b_T^4)\right\}, \tag{3.28}$$

where the leading term $\tilde{f}_i^{(0)}$ corresponds to the (resummed) purely perturbative result at canonical boundary scales, and the higher-order terms arise from expanding the nonperturbative contributions to the CS kernel and TMD PDF boundary condition (at canonical $\zeta_0 = b_0^2/b_T^2$),

$$\begin{aligned}\tilde{\gamma}_\zeta^{\text{np}}(b_T) &= \lambda_2^\zeta\, b_T^2 + \mathcal{O}(\Lambda_{\text{QCD}}^4 b_T^4)\,,\\ \tilde{f}_i^{\text{np}}(x,b_T) &= 1 + \Lambda_{2,i}(x)\, b_T^2 + \mathcal{O}(\Lambda_{\text{QCD}}^4 b_T^4)\,.\end{aligned} \tag{3.29}$$

JHEP12(2025)043

Here, λ_2^ζ is a single number related to a gluon vacuum condensate [96], while $\Lambda_{2,i}(x)$ is in principle still a function of x and i. For the q_T spectrum, λ_2^ζ is the parametrically leading nonperturbative effect since in eq. (3.28) it is enhanced by the large logarithm $\ln(b_T Q)$.

3.3.2 Nonperturbative parameterization

In our case, the perturbative contributions, $\tilde{f}^{\text{pert}}$ and $\tilde{\gamma}_\zeta^{\text{pert}}$ in eq. (3.27) are given by the default resummed results discussed in section 3.2. In addition, the explicit b_T dependence of the perturbative CS kernel is evaluated at $b^*(b_T)$, as indicated in eq. (3.27), to regulate the nonperturbative large-b_T behaviour. Specifically, we use

$$b^*(b_T) = b_T \left(1 + \frac{b_T^6}{b_{\text{max}}^6}\right)^{-\frac{1}{6}}, \tag{3.30}$$

which smoothly interpolates between the perturbative small-b_T region and the nonperturbative large-b_T region, where $b_0/b_{\text{max}} = 1\,\text{GeV}$ acts as an upper cutoff. That is, $b^*(b_T \ll b_{\text{max}}) = b_T$ while $b^*(b_T \gg b_{\text{max}}) = b_{\text{max}}$ with a smooth transition in between. The advantage of using a higher power than the usual quadratic b^* prescription is that in this way, the perturbative result differs from the strictly canonical one only starting at $\mathcal{O}(b_T^6)$, which means it does not alter the OPE below $\mathcal{O}(b_T^6)$ [97].

To include the nonperturbative contributions, we have to choose a nonperturbative model. We require the model to reproduce the correct OPE in eq. (3.29) for $b_T \to 0$. The OPE does not constrain the behaviour for $b_T \to \infty$. While as mentioned before the precise functional form at large b_T is of limited relevance, to have some guidance we require the model to obey the asymptotic limits suggested in ref. [94],

$$\begin{aligned} \tilde{\gamma}_\zeta^{\text{np}}(b_T \to \infty) &\to -\text{const}\,, \\ \ln \tilde{f}^{\text{np}}(x, b_T \to \infty) &\to -\text{const} \times b_T\,. \end{aligned} \tag{3.31}$$

A convenient choice satisfying the requirements in both limits is as follows,

$$\begin{aligned} 2\tilde{\gamma}_\zeta^{\text{np}}(b_T) = \tilde{\gamma}_\nu^{\text{np}}(b_T) &= -\lambda_\infty \tanh\left(\frac{\lambda_2}{\lambda_\infty} b_T^2 + \frac{\lambda_4}{\lambda_\infty} b_T^4\right) \qquad (3.32) \\ &= \begin{cases} -\lambda_2 b_T^2 - \lambda_4 b_T^4 + \mathcal{O}(b_T^6) & 1/b_T \gg \Lambda_{\text{QCD}} \\ -\lambda_\infty & 1/b_T \ll \Lambda_{\text{QCD}}\,, \end{cases} \\ \ln \tilde{f}^{\text{np}}(x, b_T) &= -b_T \Lambda_\infty \tanh\left[\left(\frac{\Lambda_2}{\Lambda_\infty} + \frac{\Lambda_4}{\Lambda_\infty} b_T^2\right) b_T + \frac{1}{3}\frac{\Lambda_2^3}{\Lambda_\infty^3} b_T^3\right] \qquad (3.33) \\ &= \begin{cases} -\Lambda_2 b_T^2 - \Lambda_4 b_T^4 + \mathcal{O}(b_T^6) & 1/b_T \gg \Lambda_{\text{QCD}} \\ -\Lambda_\infty b_T & 1/b_T \ll \Lambda_{\text{QCD}}\,, \end{cases} \end{aligned}$$

where we exploited that $\tanh(x \to 0) \to x$ and $\tanh(x \to \infty) \to 1$. Each model depends on three parameters: $\lambda_{2,4}$ and $\Lambda_{2,4}$ determine the quadratic and quartic terms in the expansion around $b_T = 0$ and are in one-to-one correspondence with the quadratic and quartic OPE coefficients. The parameters λ_∞ and Λ_∞ set the asymptotic behaviour for $b_T \to \infty$.

We note that the above model for the TMD PDF does not yet include flavor or x dependence, which would be required for a more complete treatment. As we explain in

JHEP12(2025)043

section 3.4, this simplified setup is sufficient for Asimov fits since we can use the same model in both pseudodata and fitted theory model. However, it is likely not sufficient for a fit to real data, which in general requires flavor and x dependence [31, 99, 100]. As shown in ref. [32], the full flavor and x dependence of the nonperturbative TMD model can be approximated by an effective flavor-averaged model that only depends on the rapidity Y and a given resonant boson type and center-of-mass energy. However, this also implies that once several different rapidity bins are considered, the effective TMD parameters must at least depend on rapidity. The fits to the q_T spectrum performed in ref. [62] found indeed evidence for the presence of a nontrivial effective rapidity dependence of Λ_2.

For the purpose of our Asimov fits, we also need some representative values for the nonperturbative parameters to be used as true values in our Asimov pseudodata. As already mentioned above, the nonperturbative CS kernel can be constrained by lattice QCD data [89–91]. This requires properly accounting for quark flavor thresholds and quark mass effects as will be discussed in ref. [66]. Using ref. [66] to fit our nonperturbative parameterization of the CS kernel to the lattice QCD data, we obtain the following results,

$$\begin{aligned} \lambda_\infty &= 1.6853 \pm 0.5069\,, \\ \lambda_2 &= (0.0870 \pm 0.0332)\,\mathrm{GeV}^2\,, \\ \lambda_4 &= (0.0074 \pm 0.0066)\,\mathrm{GeV}^4\,, \end{aligned} \tag{3.34}$$

with the resulting correlation matrix

$$C = \begin{pmatrix} 1 & 0.5212 & -0.7249 \\ 0.5212 & 1 & -0.9135 \\ -0.7249 & -0.9135 & 1 \end{pmatrix}. \tag{3.35}$$

We will use these as representative values for the CS kernel parameters of our model. For the TMD parameters, due to a lack of robust constraints, we choose a generic size of $(0.5\,\mathrm{GeV})^n$ as representative central values. Specifically, we take

$$\Lambda_2 = 0.25\,\mathrm{GeV}^2\,, \qquad \Lambda_4 = 0.06\,\mathrm{GeV}^4\,, \qquad \Lambda_\infty = 1\,\mathrm{GeV}\,. \tag{3.36}$$

We have checked that the results of our Asimov fits do not depend on the precise central values chosen.

3.4 Asimov fit setup

In this section, we outline the methodology and rationale behind the use of Asimov fits, and provide all relevant details of our setup used to study the uncertainties in extracting $\alpha_s(m_Z)$ from the Drell-Yan q_T spectrum.

3.4.1 General discussion

To estimate the expected uncertainties in the extraction of $\alpha_s(m_Z)$ — or any other theory parameter of interest — it is common practice to perform Asimov fits. These fits serve as

JHEP12(2025)043

controlled tests of the fitting framework and allow a transparent study of the sensitivity to different theory ingredients.[3]

The fits we perform use a standard χ^2 minimization procedure, defined as

$$\chi^2 = \sum_{i,j}(y_i - \lambda_i)^T C_{ij}^{-1}(y_j - \lambda_j) + \sum_i \frac{(\theta_i - 0)^2}{\Delta\theta_i^2}\,, \tag{3.37}$$

where y_i represent the data, λ_i the corresponding theory model, and C_{ij} the covariance matrix encoding all experimental uncertainties and correlations between the data points. The λ_i implicitly depend upon $\alpha_s(m_Z)$ and the parameters of the underlying theory prediction, i.e. the TNPs and nonperturbative parameters discussed in sections 3.2.3 and 3.3. Within the fit, $\alpha_s(m_Z)$, the TNPs, and nonperturbative parameters are treated as continuous parameters. The last term imposes the Gaussian theory constraint on the TNPs, θ_i, about their central values of 0 with a prior uncertainty of $\Delta\theta_i$, as discussed in section 3.2.3.

What distinguishes an Asimov fit is that the data are replaced by pseudodata — i.e., the (unfluctuated) theory predictions computed at a fixed known value of the parameter of interest, in our case $\alpha_s(m_Z) = 0.118$. The covariance matrix remains that of the actual experimental measurement, ensuring that the uncertainties remain realistic. By replacing experimental data with theory-generated pseudodata, this setup offers multiple advantages:

- The results are unobscured by statistical fluctuations or potential outlier behaviour present in the real data.
- It avoids potential inconsistencies or biases due to subleading effects present in the real data but unaccounted for in the theory model.
- The impact of various theory components or uncertainties can be tested and estimated in isolation. In particular, subleading effects can be consistently dropped in both pseudodata and theory model.
- Since the theory model and pseudodata are matched (except for intentional differences in some tests), the minimum χ^2 should be close to zero, providing a clean validation test for the fitting procedure.

As discussed in section 3.1, we focus on the dominant sources of theory uncertainty, namely the perturbative resummation and nonperturbative modelling of the dominant leading-power contribution to the small-q_T spectrum. Subleading effects such as nonsingular power corrections, finite quark-mass effects, and QCD+electroweak corrections are known to impact the shape of the q_T spectrum at the few-percent level but can be neglected in our Asimov fits as long as we consistently neglect them in both the theory model and pseudodata. Indeed, being subleading effects, their associated uncertainty will also be subdominant with respect to the dominant sources of uncertainties.

One might worry that even subleading contributions, by altering the shape of the spectrum, could subtly affect the estimate of the dominant uncertainties. Among these, the

[3]The rationale behind the Asimov fits in this setup is in many ways similar to the closure tests performed in other contexts where pseudodata are used to test the quantification of uncertainties.

JHEP12(2025)043

nonsingular terms are the most relevant, as they contribute up to $\mathcal{O}(5\%)$ of the spectrum for $q_T \lesssim 30\,\text{GeV}$ [52]. For this reason, in appendix B.1, we explicitly verify that including nonsingular terms up to $\mathcal{O}(\alpha_s^2)$ has a negligible effect on the estimated uncertainty for $\alpha_s(m_Z)$.

Nevertheless, we emphasize that in fits to real data, these subleading effects must of course be included in the fitted theory model, as they are present in the real data. While they may not substantially affect our estimate of the dominant uncertainties, they can significantly bias the central value of $\alpha_s(m_Z)$ and degrade the fit quality if not properly accounted for.

3.4.2 Fit setup

We perform our study using pseudodata based on the inclusive Z-boson q_T spectrum at $E_{\text{cm}} = 8$ TeV in the dilepton mass window $80\,\text{GeV} \leq Q \leq 100\,\text{GeV}$ as measured by ATLAS in ref. [19]. Specifically, we include the first nine q_T bins covering the region $q_T \in [0, 29]\,\text{GeV}$, for each of eight bins in rapidity $Y \in [0, 3.6]$, yielding 72 data points in total. We use the full experimental covariance matrix, incorporating all bin-to-bin correlations, as provided by ATLAS. On top of this, we include a fully correlated 1.8% luminosity uncertainty. These are the data also used in ref. [33].

For our theory predictions, both for pseudodata and for the fitted theory model, we employ the SCETlib framework as discussed in sections 3.2 and 3.3. All pseudodata are generated at a fixed value of $\alpha_s(m_Z) = 0.118$, and thus all fits must give a similar value for the fitted $\alpha_s(m_Z)$. When using scale variations, we use the resummation at N^{4}LL and follow the variation scheme defined in SCETlib [32] as described in section 3.2.2.

When using the TNP approach, we use the SCETlib implementation of ref. [34] as described in section 3.2.3. Our nominal theory model is given by N^{3+1}LL while for the pseudodata we use either the default N^{3+1}LL result with zero as true values for all TNPs or alternatively N^{4}LL, which is equivalent to N^{3+1}LL with all TNPs set to their true values. As we will see, the advantage of using N^{3+1}LL is that the perturbative precision is sufficiently high compared to the experimental precision to obtain nontrivial results. At the same time the true values of all TNPs are still known, which allows us to perform more nontrivial tests by fitting to N^{4}LL, simulating a fit to real data.

For the nonperturbative modelling, we always use the same nonperturbative model for both the theory model and pseudodata, and whenever the nonperturbative parameters are fixed they are set to their central values given in section 3.3.

Regarding the PDF dependence, we always use the MSHT20aN3LO PDF set [59], and account for the interplay between $\alpha_s(m_Z)$ and the PDFs in the q_T spectrum. Since α_s and the PDFs are strongly correlated, we ensure that the PDF set entering the prediction is consistent with the value of $\alpha_s(m_Z)$ being tested. For this purpose, we perform a linear interpolation between the available PDF sets for $\alpha_s(m_Z) = 0.118$ and $\alpha_s(m_Z) = 0.117$ when $\alpha_s(m_Z) < 0.118$ and between the PDF sets for $\alpha_s(m_Z) = 0.118$ and $\alpha_s(m_Z) = 0.119$ when $\alpha_s(m_Z) > 0.118$. This interpolation is only used for the PDFs, while the dependence on $\alpha_s(m_Z)$ in the perturbative prediction itself is always treated exactly.

For minimization, we use the standard Minuit fitting tool [101–103].

JHEP12(2025)043

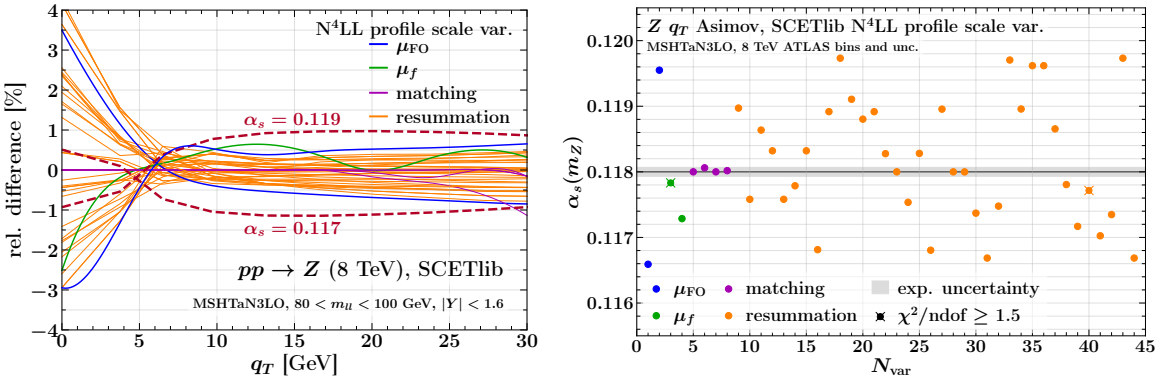


Figure 2. Left panel: Relative differences in the q_T spectrum of individual scale variations at N^{4}LL. Different colored lines show different classes of variations as defined in section 3.2.2. The red dashed lines show the relative difference from varying $\alpha_s(m_Z)$ for comparison. Right panel: The resulting $\alpha_s(m_Z)$ values when performing the Asimov fit to the central N^{4}LL prediction for each scale variation. Different colors correspond to the same classes of variations as on the left. The grey band shows the fit uncertainty for comparison.

4 Perturbative uncertainties

In this section, we present our results obtained using both scale variation and TNP-based approaches to estimate the expected perturbative uncertainty in the extraction of $\alpha_s(m_Z)$ from the Z q_T spectrum. We describe and highlight two distinct methods that can be applied with TNPs: scanning and profiling. Subsequently, we explore the profiling of TNPs in more detail by performing fits against higher-order pseudodata and by modifying the prior theory constraint $\Delta\theta_n$ on the TNPs.

4.1 Scale variations

To estimate the uncertainty due to missing higher-order corrections — referred to as the perturbative uncertainty — it is common practice to consider scale variations. A key argument for this procedure is that the envelope of various scale variations can provide a measure of the overall perturbative uncertainty in the spectrum. By taking the envelope, one can mitigate the arbitrariness in choosing the specific number and type of scale variations to be considered. As a result, it is a widely-used method for estimating perturbative uncertainties in differential spectra, which may often provide a reasonable estimate of the approximate size of uncertainties. The resulting uncertainty band then has limited sensitivity to the individual scale variation shapes. This is sufficient for the uncertainties at the level of the q_T spectrum where we only require the rough size of the uncertainty band but not the detailed correlations within it. The inadequacy of scale variations manifests itself when interpreting the spectrum, i.e., when the perturbative uncertainty needs to be propagated from the spectrum to the parameter of interest, in our case $\alpha_s(m_Z)$.

One may ask why is it not sufficient to propagate each scale variation to a fit of $\alpha_s(m_Z)$ and then take the envelope at the level of the extracted $\alpha_s(m_Z)$ value? Any specific scale choice would then correspond to a fully correlated or anti-correlated assumption, as explained in section 2.2. In taking the envelope of scale variations at the level of the q_T spectrum we

JHEP12(2025)043

explicitly acknowledge that their individual shapes are not physical, and assume complete ignorance regarding the correlations between different bins in the spectrum. However, the propagation of the uncertainty from the q_T spectrum to $\alpha_s(m_Z)$ is crucially sensitive to the precise point-by-point correlations and the individual theory shapes. Accordingly, any uncertainty resulting from such a scale-variation scanning and enveloping procedure for $\alpha_s(m_Z)$ should not be considered as a trustworthy or meaningful uncertainty but merely as an indication of the possible impact of missing higher-order corrections, i.e., as a means to investigate the possible theory sensitivity (which could still very well be over- or underestimated).

Despite these known limitations of the scale variation approach, it is still interesting in this context to use our pseudodata setup to see what happens if scale variations are utilised. For our purposes they will serve to demonstrate that correlations and the precise shape of the theory uncertainties are indeed critical for the extraction of $\alpha_s(m_Z)$.

In the left panel of figure 2, we show the breakdown of all scale variations for the q_T spectrum at N^{4}LL from SCETlib as discussed in section 3.2.1. Overlaid on these variations are the central predictions for $\alpha_s(m_Z) = 0.117$ and $\alpha_s(m_Z) = 0.119$, illustrating that the sensitivity to $\alpha_s(m_Z)$ in the q_T spectrum is primarily a shape effect. To illustrate the scale-variation scanning in practice, we use the Asimov fit setup described in section 3.4. Our pseudodata corresponds to the N^{4}LL central scale prediction at $\alpha_s(m_Z) = 0.118$. We then repeat the fit of $\alpha_s(m_Z)$ for each scale variation, each representing a different theory model. The right panel of figure 2 shows the results, where each point represents the fitted value of $\alpha_s(m_Z)$ for one of the scale variations on the left. Points marked with a cross indicate cases where $\chi^2/\mathrm{n_{dof}} \geq 1.5$, signifying a very poor fit to the pseudodata. The fitted $\alpha_s(m_Z)$ values strongly depend on the choice of scale variation. While some variations have almost no impact on $\alpha_s(m_Z)$, yielding values very close to the nominal value of 0.118, others yield very different values. This demonstrates that as anticipated the impact on $\alpha_s(m_Z)$ strongly depends on the precise shape of any given variation. As a result, the precise choice of scale variations directly dictates the size of the would-be perturbative uncertainty, making this approach unsuitable for a reliable uncertainty estimate.

Nevertheless, for reference we can obtain a scale-variation based estimate of the perturbative uncertainty in $\alpha_s(m_Z)$ by considering different options to envelope the fitted results. The envelopes of each different class of scale variations described in section 3.2.2 yield

$$\begin{aligned} \Delta_{\mathrm{FO}} &= 1.55 \times 10^{-3}\,, & \Delta_f &= 0.71 \times 10^{-3}\,, \\ \Delta_{\mathrm{resum}} &= 1.73 \times 10^{-3}\,, & \Delta_{\mathrm{match}} &= 0.06 \times 10^{-3}\,. \end{aligned} \tag{4.1}$$

We can then either sum these in quadrature or take the naive total envelope, obtaining

$$\begin{aligned} \text{sum of envelopes:} \quad & \Delta_{\mathrm{pert}} = \sqrt{\Delta_{\mathrm{FO}}^2 + \Delta_f^2 + \Delta_{\mathrm{resum}}^2 + \Delta_{\mathrm{match}}^2} = 2.43 \times 10^{-3}\,, \\ \text{total envelope:} \quad & \Delta_{\mathrm{pert}} = 1.73 \times 10^{-3}\,. \end{aligned} \tag{4.2}$$

Regardless of the approach, we find that the perturbative uncertainty estimated via scale variations at the currently highest known order, N^{4}LL, is large — much larger than our desired level of precision for $\alpha_s(m_Z)$

JHEP12(2025)043

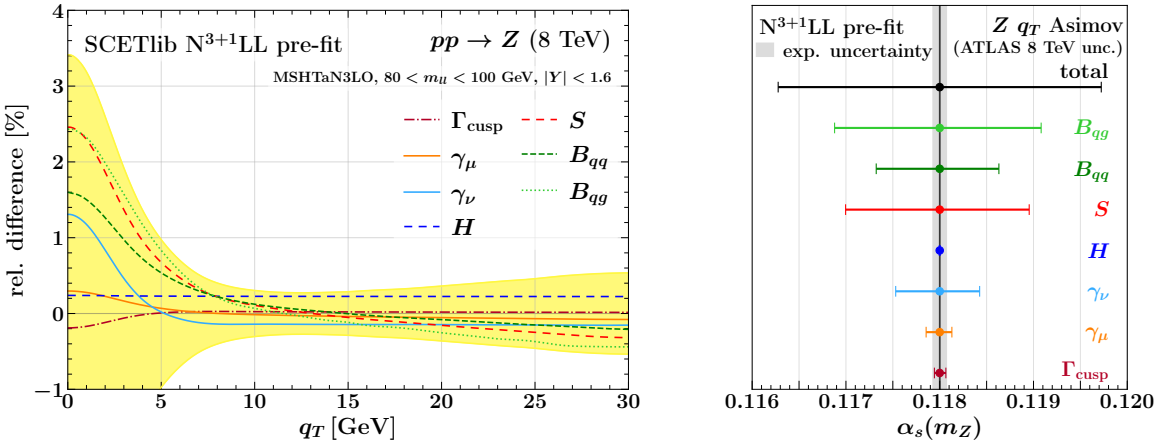


Figure 3. Left panel: Relative uncertainties in the q_T spectrum with TNPs at N^{3+1}LL. The different lines show the impact of varying the corresponding TNP by $+1$ or -1, corresponding to 68% theory CL. The yellow band shows their sum in quadrature. Right panel: Corresponding uncertainty on $\alpha_s(m_Z)$ from performing the Asimov fit to the central N^{3+1}LL prediction for each TNP variation. The grey band shows the fit uncertainty for comparison.

4.2 Theory nuisance parameters

As explained in section 2, to properly account for the point-by-point correlations across the q_T spectrum, the uncertainty must be decomposed into well-defined components corresponding to the underlying sources of uncertainty. The TNP approach has been developed to do just that. In this section, we explore different methods using TNPs to estimate the perturbative uncertainty in the extraction of $\alpha_s(m_Z)$. Our setup is as described in section 3.4, with the default theory model at N^{3+1}LL and the Asimov pseudodata at N^{3+1}LL and later also at N^{4}LL.

4.2.1 Scanning

The first method we consider is scanning, which follows an analogous procedure to the one used for scale variations. However, the crucial difference is that we have now decomposed the uncertainty into its distinct sources, parameterized in terms of well-defined theory nuisance parameters, which correctly capture and propagate the bin-by-bin correlations in the q_T spectrum.

The left panel of figure 3 presents the breakdown of the individual TNP variations for the q_T spectrum at N^{3+1}LL accuracy. Each TNP is varied by $\Delta\theta_n = 1$. Since the resulting variations in the spectrum are nearly symmetric, for clarity, we display only the up variations for the anomalous dimensions, hard, and soft functions, and the down variations for the beam functions. Each TNP corresponds to an independent source of uncertainty and is thus varied independently, capturing the correct correlations across the q_T spectrum. We can now sum in quadrature the seven independent TNP variations to obtain the total uncertainty band, depicted in yellow.

To perform the scanning, we use the central N^{3+1}LL prediction as Asimov data and repeat the fit for N^{3+1}LL theory models with each TNP varied separately up and down by $\Delta\theta_n = 1$. The right panel of figure 3 shows the results, where the error bars represent the difference

JHEP12(2025)043

between the fitted $\alpha_s(m_Z)$ and the nominal expected value of 0.118, i.e., the uncertainty on the fitted $\alpha_s(m_Z)$ due to a given TNP. As for the spectrum, we can now sum the individual uncertainties in quadrature to obtain the total uncertainty on $\alpha_s(m_Z)$, which yields

$$\Delta_{\text{pert}} = 1.75 \times 10^{-3}\,. \tag{4.3}$$

Although this estimate accounts for the correct theory correlations in the q_T spectrum, it does not reach our desired precision on $\alpha_s(m_Z)$.

The scanning approach however does not fully exploit the high precision available in the data, which could in principle impose strict constraints on the parameters of the theory model. As each parameter is separately varied but otherwise held fixed in the fit, the fit must compensate the resulting change in the theory model as best it can solely into a change of $\alpha_s(m_Z)$. However there is no feedback whether a given variation of the theory model is actually favored or disfavored by the data. This is a key limitation of the scanning approach. Furthermore, since we explore the parameter space separately in each direction rather than all together, the global minimum of the multi-dimensional parameter space is not explored, which means the quality of the minimum and the precision obtained can be suboptimal.

Rather than relying on scanning — and its inherent limitations — we would like to be able to fully explore the parameter space and allow the fit to select the regions where the theory model best matches the data while avoiding regions that are inconsistent with it. This challenge can be effectively addressed by profiling the TNPs in the fit as we discuss next.

4.2.2 Profiling

A more advanced approach to estimate systematic uncertainties is by profiling them. In this approach, the nuisance parameters associated with the systematic uncertainties are simultaneously fitted along with the parameters of interest, rather than scanning them one-by-one. In this way the full parameter space is explored, avoiding the limitations of the scanning approach and enabling the data to constrain variations within these uncertainties to agree with the measurements. This therefore enables a reduction in the size of the associated uncertainties for the parameters of interest. This is a standard approach used for experimental systematic uncertainties (see e.g. refs. [104, 105]) as well as for PDF uncertainties [106–108], though several caveats must be considered in general. In particular, one must be careful in the application to PDF uncertainties — we leave this discussion to a future dedicated paper [40].

We now apply this approach to the nuisance parameters encoding the perturbative uncertainties, i.e. the TNPs, by fitting them together with $\alpha_s(m_Z)$. This allows the fit to find the optimal balance between adjusting one or more TNPs, and consequently the theory model, or adjusting $\alpha_s(m_Z)$ to find the best description of the data. We stress that (in contrast to scale variations) fitting the TNPs is theoretically consistent and allowed because the TNPs are proper parameters, which encode the correct theory correlations. As a result, this method enables constraining the TNPs by the experimental data, thereby (potentially) reducing the theory uncertainties. The allowed range of variation for the TNPs is also constrained by the prior theory constraint, as explained in sections 3.2.3 and 3.4, for which we use $\Delta\theta_n = 1$ by default.

JHEP12(2025)043

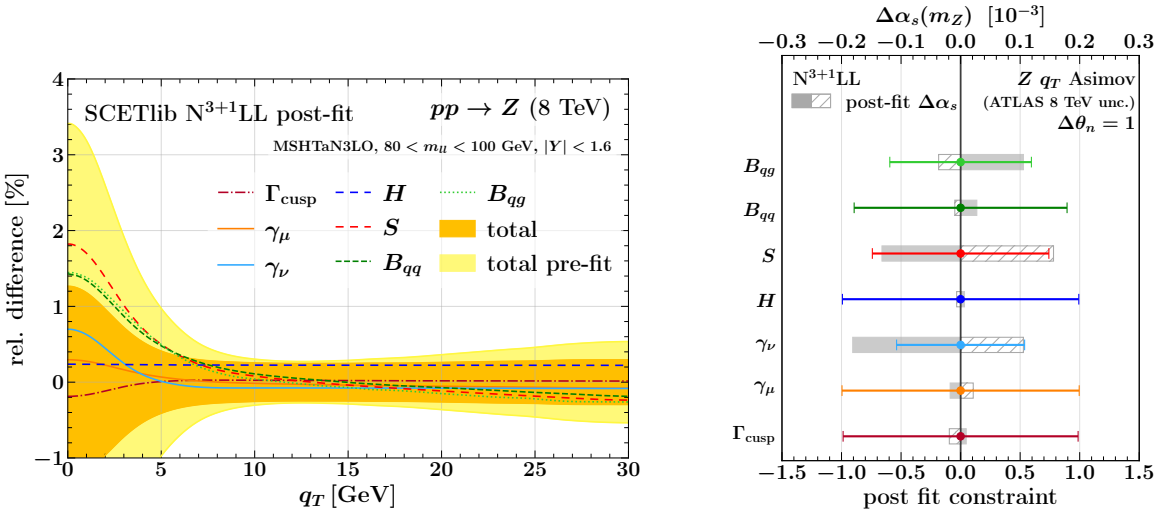


Figure 4. Left panel: Relative uncertainties in the q_T spectrum at N^{3+1}LL before and after profiling the TNPs. The different lines show the post-fit relative impact of each TNP and the orange band the total post-fit uncertainty. The yellow band shows the pre-fit uncertainty corresponding to the yellow band in figure 3. Right panel: Post-fit constraints on the TNPs (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) TNP variations.

To start, we perform an Asimov fit with both pseudodata and theory model given by the N^{3+1}LL prediction. The resulting perturbative uncertainty on $\alpha_s(m_Z)$ is[4]

$$\Delta_{\text{pert}} = 0.45 \times 10^{-3}\,. \tag{4.4}$$

The left panel of figure 4 presents the post-fit prediction for the q_T spectrum with a breakdown of the post-fit uncertainties, showing the upward variation for each TNP, except for the beam functions which display the downward variation. As a general comment, which applies to this and all analogous plots in the following, note that the individual TNP post-fit variations do not necessarily have to lie within the total post-fit uncertainty, as there are now nontrivial correlations between the TNPs imposed by the data. In particular, if two TNPs are negatively correlated, their combined uncertainty on the spectrum and $\alpha_s(m_Z)$ can be smaller than their individual impacts.

After profiling the TNPs, the uncertainties for both $\alpha_s(m_Z)$ and the q_T spectrum are significantly smaller compared to simply scanning the TNPs and the corresponding pre-fit uncertainties. However, focusing solely on the final uncertainty after profiling can also be misleading and may lead to incorrect conclusions. An important aspect when profiling systematic uncertainties is to examine to what extent the data actually pulls and constraints the nuisance parameters in order to ensure that the constrained post-fit theory model is still adequate and allowed. In our case, we have to ensure that the theory model is not overconstrained to the point that neglected yet higher-order terms become relevant for

[4]This value is the total fit uncertainty, which in principle includes the experimental uncertainties as well. However, the pure fit uncertainty when only fitting $\alpha_s(m_Z)$ is tiny, 0.06×10^{-3}, so we simply refer to this uncertainty as the perturbative one, regardless of the use of profiling.

JHEP12(2025)043

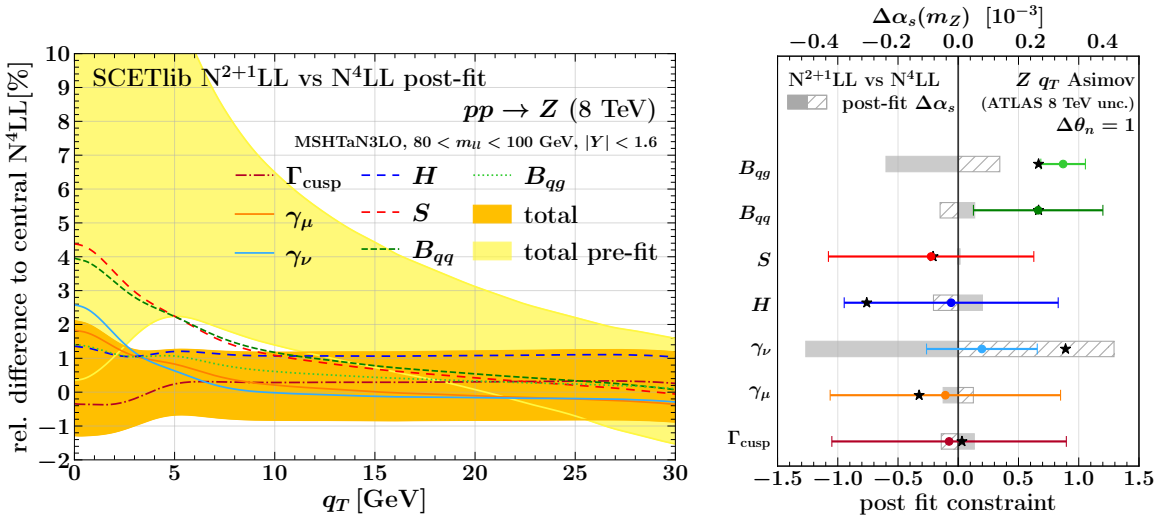


Figure 5. Left panel: Uncertainties in the q_T spectrum at N^{2+1}LL relative to the N^{4}LL, before (yellow band) and after (orange band) profiling the TNPs. The different lines show the post-fit relative impact of each TNP. Right panel: Post-fit constraints on the TNPs (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) TNP variations. The stars indicate the true values of the TNPs.

determining the theory uncertainty [34]. The right panel of figure 4 presents a pull/impact plot for all TNPs. The error bars refer to the bottom x-axis and show the post-fit constraints for each TNP. The solid (dashed) grey bars refer to the top x-axis and show the corresponding impact on $\alpha_s(m_Z)$ due to each TNPs downward (upward) post-fit uncertainty.[5] The post-fit central value for all TNPs is 0, in perfect agreement with their values in the pseudodata, as should be the case here. The post-fit constraints on the individual TNPs are overall not strong. The individual impacts provide further insight into which TNPs are most important (after profiling) for determining $\alpha_s(m_Z)$. Here, γ_ν, S, and B_{qg} have the biggest post-fit impact on $\alpha_s(m_Z)$ and they are also the ones that are somewhat constrained. This indicates that these parameters have the strongest correlation with $\alpha_s(m_Z)$. Of course, to study the exact correlations between all parameters in detail, one has to examine the post-fit covariance matrix.

4.2.3 Profiling against N^{4}LL

To further explore the profiling with TNPs, we next perform more general Asimov fits where the pseudodata is generated from the highest N^{4}LL central prediction, still evaluated at $\alpha_s(m_Z) = 0.118$, while using different orders for the theory model, specifically N^{2+1}LL and N^{3+1}LL. Since the N^{4}LL result contains the true values of the TNPs, this setup mimics the fit to real data, which contain nature's all-order result, and provides a realistic test of whether the TNPs are correctly constrained and whether we can recover their true values at this order.

In figure 5, we consider the case where the theory model is given by the N^{2+1}LL prediction. The left panel shows the pre-fit and post-fit uncertainties for the q_T spectrum analogous

[5]The impacts are evaluated by repeating the fit while fixing each TNP one at a time to its post-fit value plus/minus its post-fit uncertainty. The difference of the obtained $\alpha_s(m_Z)$ to the result of the nominal fit gives the impact on $\alpha_s(m_Z)$.

JHEP12(2025)043

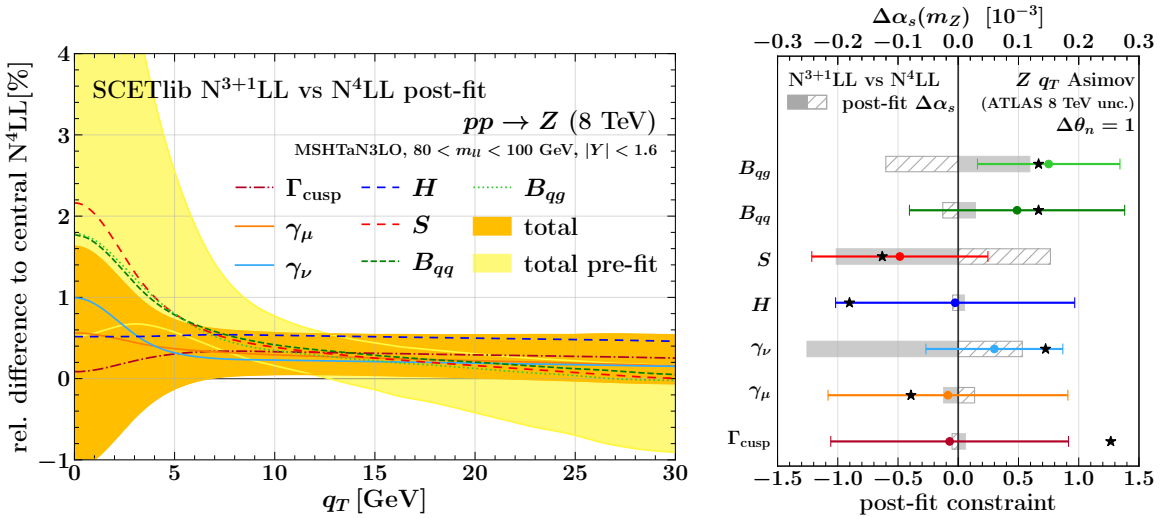


Figure 6. Left panel: Uncertainties in the q_T spectrum at N^{3+1}LL relative to the N^{4}LL, before (yellow band) and after (orange band) profiling the TNPs. The different lines show the post-fit relative impact of each TNP. Right panel: Post-fit constraints on the TNPs (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) TNP variations. The stars indicate the true values of the TNPs.

to the left panel of figure 4 but now relative to the true N^{4}LL result. The profiling again leads to a significant reduction in the uncertainty. Furthermore, we see that the post-fit prediction is now pulled toward the true result, with a significant adjustment of the shape of the spectrum. Whilst the shape of the pre-fit prediction differs from the true shape, the post-fit prediction closely reproduces it.

The right panel of figure 5 shows the corresponding pull plot for the TNPs (analogous to the right panel of figure 4) with the true values of the TNPs marked by stars. The fitted TNPs at N^{2+1}LL now exhibit strong pulls toward their true values, showing that, as for the q_T spectrum, the fit is moving the theory model in the right direction also at the level of individual TNPs. This is exactly what we would like to happen in this situation where the data has a higher precision than the pre-fit theory constraint. However, some of the TNPs, such as γ_ν, B_{qq}, and in particular B_{qg}, are quite strongly constrained now. A tight post-fit constraint implies that the corresponding parameter at the next higher-order becomes relevant. In other words, we are starting to "overfit" the N^{2+1}LL theory model. The precision of the data is starting to exceed the inherent precision of the theory model, suggesting that it is becoming insufficient and using a higher-order theory-model might be warranted.

In figure 6 we repeat the same exercise with the N^{3+1}LL theory model. The post-fit uncertainties for the q_T spectrum are very similar to those in figure 4 when profiling N^{3+1}LL against itself, except that now the central value is shifted toward the true result. The same holds at the level of individual TNPs. Their post-fit uncertainties are very similar to before, while their central values are shifted toward their true values, which is again as we would like it to be. Compared to the N^{2+1}LL case, the TNPs are now less constrained, as expected, with no indications yet of overfitting the theory model.

JHEP12(2025)043

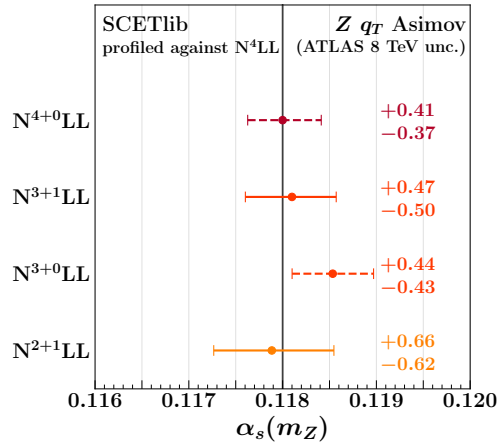


Figure 7. Perturbative uncertainties on $\alpha_s(m_Z)$ when profiling the TNPs at different orders against the central N^{4}LL prediction. The numerically displayed uncertainties are in units of 10^{-3}.

In figure 7, we present the results for $\alpha_s(m_Z)$ when profiling the TNPs at various orders against the N^{4}LL pseudodata. The uncertainty on $\alpha_s(m_Z)$ is only slightly larger at N^{2+1}LL than at N^{3+1}LL, which highlights the fact that after profiling it is largely driven by the constraints imposed by the measurement on the theory model. In addition, we also include results at N^{3+0}LL and N^{4+0}LL — these orders serve as approximations of N^{3+1}LL and N^{4+1}LL, respectively, see appendix B.2 for their definition and a more detailed discussion. While these approximations lack the correct formal structure of their corresponding higher orders, we include them here for completeness. Note that although N^{4+0}LL is technically our currently highest available order, we prefer to use N^{3+1}LL as our nominal theory model, because it has the correct theory correlation structure. Furthermore, its theory precision is sufficiently high compared to the measurement precision and after profiling it yields uncertainties comparable to N^{4+0}LL.

At both N^{2+1}LL and N^{3+1}LL we observe a small bias in the fitted central value of $\alpha_s(m_Z)$. First note that when fitting the lower-order theory model against the higher-order pseudodata, we intentionally introduce a mismatch. Without profiling, this naturally leads to a potential bias in $\alpha_s(m_Z)$ as it potentially absorbs some of the deficiencies of the theory model. This type of bias is actually what is estimated and supposed to be covered by the scanning approach (for either TNPs or scale variations). Indeed, fitting the central N^{2+1}LL and N^{3+1}LL models against N^{4}LL and evaluating the uncertainty by scanning over TNPs, we find $\alpha_s(m_Z) = 0.1211 \pm 0.0044$ and $\alpha_s(m_Z) = 0.1195 \pm 0.0018$, respectively. The biases compared to the correct value of 0.118 are thus indeed covered by their respective scanning uncertainties.

When profiling the TNPs, the fit is allowed to improve the theory model by moving the TNPs toward their true values, which reduces this type of bias. Nonetheless, in doing so the fit must balance the theory and data constraint. The true values of some of the parameters, shown as stars in the pull plots, can always happen to be of order the size of the theory constraint away from 0. As a result, the fit picks up penalties as it tries to move the parameters to their true values, which induces a residual bias in $\alpha_s(m_Z)$ due to the imposed

JHEP12(2025)043

theory constraint. Importantly, we observe that this type of residual bias is much smaller and practically negligible compared to the output uncertainties.

To summarize, our profiling results highlight the full power of the TNP approach: Our Asimov tests demonstrate that when profiling the TNPs, they are properly constrained by the data and pulled toward their true values, leaving only a negligible residual bias. The high precision of the data thus allows us to consistently reduce the theory uncertainties and to achieve the sensitivity required for an extraction of $\alpha_s(m_Z)$ at the (sub)percent level.

We have to caution, however, since as discussed in section 3.2.3, with the minimal set of TNPs we use here, there still remain neglected sources of uncertainty, which could still enlarge the perturbative uncertainty on $\alpha_s(m_Z)$. Whilst some of these can be anticipated to be negligible, such as the singlet contributions to the hard function given the observed minor impact of the hard function on $\alpha_s(m_Z)$, others may yield nonnegligible additional contributions. In particular, given the notable impact of the already included beam function uncertainties, the impact of higher order structures in the beam functions are such an example. Another example are the perturbative uncertainties due to PDF evolution, which are not yet accounted for by the TNPs, but which could be of relevance (as also indicated by the nontrivial size of Δ_f in the scale-variation approach). The same caveats of course also apply to our earlier scanning results.

4.2.4 Dependence on pre-fit theory constraints

The final aspect we would like to discuss concerns the pre-fit theory constraints imposed on the TNPs, as discussed in sections 3.2.3 and 3.4. So far, our default choice has been $\Delta\theta_n = 1$. As the adopted theory constraint ultimately depends on some theoretical judgement, we might question whether it is overly restrictive and to what extent the uncertainties we find depend on it. These issues have already been discussed in ref. [34]. Here, we provide a practical application allowing us to explicitly answer these questions.

When the theory constraints on TNPs and the precision of the data are of comparable size — such that neither clearly dominates — both central value and uncertainty of the parameters of interest will in principle depend on the theory constraint. As we have seen, the profiling already essentially eliminates the bias in the central value. One key message to emphasize is that profiling also substantially reduces the dependence of the uncertainties on the theory constraint.[6] In contrast, during scanning, the uncertainty on $\alpha_s(m_Z)$ directly depends on the chosen theory constraint.

To demonstrate this, we can relax the theory constraint by increasing $\Delta\theta_n$. This effectively reduces the weight of the theory constraint and allows the data to play a larger role in constraining the TNPs and thus determining the final uncertainty. We again use the $\mathrm{N^{3+1}LL}$ prediction as theory model and profile it against the $\mathrm{N^4LL}$ pseudodata and now vary the pre-fit constraints by factors of 2, setting $\Delta\theta_n = 1, 2, 4$.

In figure 8, we show the resulting pull plots. As the pre-fit uncertainty on the TNPs is increased, also the post-fit uncertainty on the TNPs increases but much less so for some TNPs. Overall, the post-fit values become more and more compatible within uncertainties

[6] This was also discussed in ref. [59] in the context of the theory nuisance parameters employed in their approximate $\mathrm{N^3LO}$ PDF fit.

JHEP12(2025)043

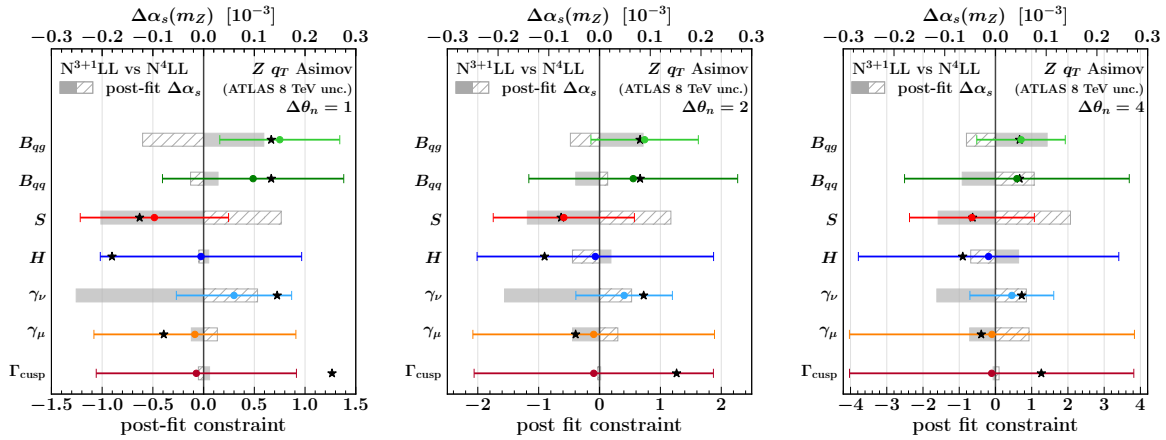


Figure 8. Post-fit constraints on TNPs at N^{3+1}LL profiled against the N^{4}LL central prediction using different pre-fit theory constraints, left panel: $\Delta\theta_n = 1$, central panel: $\Delta\theta_n = 2$ and right panel: $\Delta\theta_n = 4$. The stars indicate the true values of the TNPs.

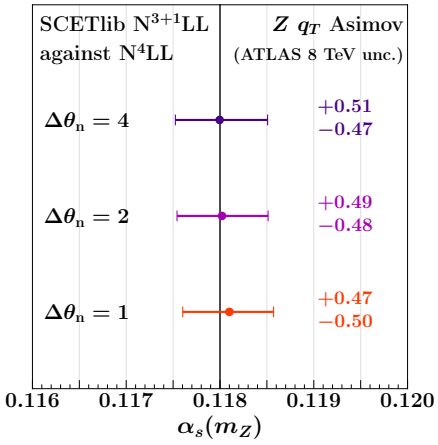


Figure 9. Perturbative uncertainties on $\alpha_s(m_Z)$ when profiling TNPs at N^{3+1}LL against the N^{4}LL central prediction using different pre-fit theory constraints $\Delta\theta_n = 1, 2, 4$. The displayed uncertainties are in units of 10^{-3}.

with the TNPs' true values, because it becomes easier for the fit to move the TNPs to their true values. With $\Delta\theta_n = 4$, the constraints on certain TNPs become stronger relative to the pre-fit constraint and the less constrained TNPs. This is expected since with $\Delta\theta_n = 4$ the TNPs become essentially unconstrained, which makes the differing sensitivities to the various TNPs more clear as the data acts to constrain the most relevant TNPs. When relaxing the theory constraint, the TNPs having the strongest impact on $\alpha_s(m_Z)$ remain the same, namely γ_ν, S, and B_{qg}. However, their individual impacts reduce while the impact of the other TNPs increases.

In figure 9, we show the results for $\alpha_s(m_Z)$ for all three cases. Interestingly, despite the shifts in the impacts of individual TNPs, the total uncertainty on $\alpha_s(m_Z)$ remains essentially unchanged. This shows that the total uncertainty after profiling is essentially driven by the constraints imposed by the data. The precise theory constraint does not affect the total

JHEP12(2025)043

uncertainty but only the relative distribution of the uncertainties among the TNPs. This is the case in a situation like ours here where the uncertainties of the data are $\lesssim$ the pre-fit theory uncertainties, i.e., in a situation where without profiling we would be limited by theory uncertainties. In the opposite situation, where the data uncertainties are larger than the pre-fit theory uncertainties, the theory uncertainty will depend on the theory constraint also after profiling simply because the data will be too weak to constrain the theory uncertainties. However, in this situation the theory uncertainties are subdominant to begin with, and whilst the theory uncertainty itself will depend more on the adopted theory constraint, the total uncertainty will not. Therefore, as discussed before, the only situation to be avoided is that the inherent precision of the theory model, i.e., the order to which TNPs are included is too low compared to the data precision such that it becomes overconstrained. In this case, the theory model can be improved by including additional TNP orders.

As already discussed in the context of figure 7, there is a minor bias in the central value of $\alpha_s(m_Z)$ due to the $\Delta\theta_n = 1$ theory constraint. Now, as the theory constraint is weakened, the true values of the TNPs lie well within it, so the associated penalties are reduced. As a result, the bias further reduces for $\Delta\theta_n = 2$ and disappears completely for $\Delta\theta_n = 4$.

Based on these results, we are confident in our default choice of $\Delta\theta_n = 1$. In case one is worried about the associated theoretical prejudice, it can always be relaxed. Even in the extreme case of $\Delta\theta_n = 4$, which here amounts to practically removing the theory constraint, the data are still able to sufficiently constrain the TNPs on their own. Most importantly, the conclusions of our study do not depend on the precise theory constraint.

5 Nonperturbative effects

In this section we investigate another major source of uncertainty: q_T-dependent nonperturbative effects as discussed in section 3.3. We consider the nonperturbative contributions from both the CS kernel and the TMD PDF boundary condition, with two parameters for each, which we find to be the minimal necessary set.

Throughout this section, we use the central N^{4}LL prediction at $\alpha_s(m_Z) = 0.118$ as pseudodata, while the theory model is given by the N^{3+1}LL prediction. We adopt the nonperturbative models for the CS kernel and TMD PDF discussed in section 3.3.2. Their parameters $\lambda_{2,4}$ and $\Lambda_{2,4}$ directly correspond to the leading quadratic and quartic OPE coefficients, which are expected to be the dominant effects. In the pseudodata, they are fixed to their central values, while in the theory model they are now treated as additional fit parameters. The parameters Λ_∞ and λ_∞, which control the large b_T behaviour, are always kept fixed as the fit has practically no sensitivity to them. This confirms the expectation that the dominant nonperturbative sensitivity indeed comes from the small-b_T behaviour, where the OPE applies and determines the b_T dependence of nonperturbative contributions.

Our nominal result thus corresponds to profiling the TNPs at N^{3+1}LL along with $\lambda_{2,4}$ for the CS kernel and $\Lambda_{2,4}$ for the TMD PDFs. For $\lambda_{2,4}$ we impose the representative lattice constraints in eqs. (3.34) and (3.35) as prior Gaussian constraints. As we will see, including the lattice constraints for $\lambda_{2,4}$ helps guide the exploration of parameter space, allowing the data to determine the TMD parameters $\Lambda_{2,4}$, which contribute to the uncertainty but for which we otherwise lack constraints from first-principles calculations.

JHEP12(2025)043

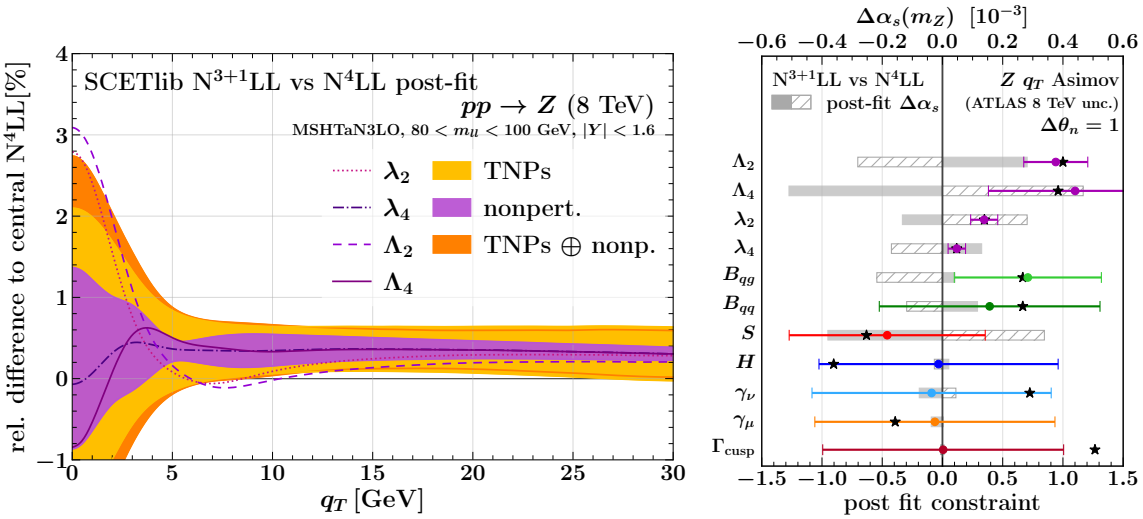


Figure 10. Left panel: Uncertainties in the q_T spectrum at N^{3+1}LL relative to the N^{4}LL, due to the nonperturbative parameters (purple band), the TNPs (orange band) and their sum in quadrature band (in dark orange). The different lines show the post-fit relative impact of each nonperturbative parameter in the spectrum. Right panel: Post-fit constraints on the TNPs and nonperturbative parameters (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) variations of each parameter. The stars indicate the true value of the parameters. The post-fit constraints for Λ_2 and λ_2 (Λ_4 and λ_4) are normalized to $0.25\,\text{GeV}^2$ ($0.06\,\text{GeV}^4$).

In figure 10, we show the result of our nominal fit including nonperturbative parameters. The left panel displays the breakdown of the post-fit uncertainties in the q_T spectrum. The uncertainty due to the nonperturbative parameters is shown in violet, the perturbative uncertainty due to TNPs is shown in orange, and their combined uncertainty is represented by the dark orange band, which corresponds to the total post-fit uncertainty for this Asimov test.[7] We do not show a pre-fit band here, since we do not have pre-fit constraints for $\Lambda_{2,4}$. The perturbative pre-fit uncertainty corresponds to the one shown in figure 6. Overlaid on the plot are the post-fit relative impacts of each nonperturbative parameter. A strong negative correlation emerges between λ_2 and Λ_2, as evident from their significant impact on the spectrum despite the relatively small size of the overall nonperturbative uncertainty. In the right panel of the same figure, we show the constraints on all fit parameters, along with their individual impacts on $\alpha_s(m_Z)$. Compared to the purely perturbative case discussed in section 4.2.3, the post-fit constraints on the TNPs are still consistent with their true values (indicated by stars), but some are now not as strongly constrained by the data. Due to the additional fit parameters introduced, the data has additional freedom now to determine which parameters are most relevant in reducing the uncertainty on the parameter of interest. It is clear that some nonperturbative parameters, especially $\Lambda_{2,4}$, have a large impact on $\alpha_s(m_Z)$.

[7]When profiling multiple sources of uncertainty, the breakdown of the total uncertainty into individual components is somewhat ambiguous. Here, we obtain the individual bands on the q_T spectrum by considering the relevant sub-matrices of the total post-fit covariance matrix. One benefit of the Asimov approach however is that we can also consider different sources in isolation, as was done for the TNP component in section 4.2.

JHEP12(2025)043

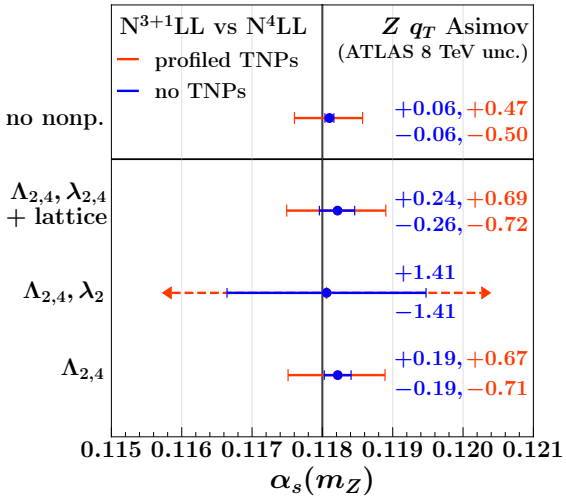


Figure 11. Uncertainty on $\alpha_s(m_Z)$ from various Asimov tests. Red bars include TNP profiling (i.e., fitting $\alpha_s(m_Z)$ together with all theory nuisance parameters). Below, we show cases excluding TNP profiling, in blue, with different sets of fitted parameters. From top to bottom: fit only $\alpha_s(m_Z)$, fit $\alpha_s(m_Z)$, $\lambda_{2,4}$, $\Lambda_{2,4}$ including lattice QCD constraints for CS kernel parameters, fit $\alpha_s(m_Z)$, λ_2 and $\Lambda_{2,4}$, fit $\alpha_s(m_Z)$ and $\Lambda_{2,4}$. All the uncertainties are in units of 10^{-3}.

Next, we explore the relationship between the nonperturbative effects and the uncertainty on $\alpha_s(m_Z)$, and how the TNPs behave in this context. Figure 11 presents the results of several Asimov fits as discussed next. For each case, we display two uncertainty bars. The outer red bars show the total uncertainty of the nominal fit simultaneously profiling TNPs and nonperturbative parameters. The inner blue bars show the uncertainty associated with fitting $\alpha_s(m_Z)$ alongside the nonperturbative parameters only. They are obtained by repeating the fit fixing the TNPs to their nominal post-fit central values.

At the top of the plot, we show the baseline result previously discussed in section 4.2.3, with no nonperturbative parameter included in the fit. In this case, the inner blue uncertainty is the uncertainty due to only fitting $\alpha_s(m_Z)$ with all the other theory parameters fixed, yielding an uncertainty of $\Delta_{\text{fit}} = 0.06 \times 10^{-3}$. Including the TNPs in the fit increases the uncertainty to $\Delta_{\text{pert}} = ^{+0.47}_{-0.50} \times 10^{-3}$. (As before, we ignore the very small contribution from Δ_{fit} and simply refer to Δ_{pert} here and similarly for Δ_{nonp} below.)

We now investigate what happens when also fitting the nonperturbative parameters. Starting from the bottom of the plot, the first test we perform involves fitting only the $\Lambda_{2,4}$ parameters of the TMD PDF, while keeping the parameters of the CS kernel, $\lambda_{2,4}$, fixed at their central values. This yields a total uncertainty of $\Delta_{\text{pert}\oplus\text{nonp}} = ^{+0.67}_{-0.71} \times 10^{-3}$.

In addition to the TMD PDF, we also have to account for the nonperturbative uncertainties due to the CS kernel. We find that including in addition $\lambda_{2,4}$ in the fit without any further constraints on them does not yield a stable fit. Including only λ_2 the full fit including TNPs is barely able to converge and does not yield sensible uncertainties. We are only able to obtain a trustworthy result when including $\Lambda_{2,4}$ together with λ_2 while fixing the TNPs. It yields a very large nonperturbative uncertainty of $\Delta_{\text{nonp}} = \pm 1.41 \times 10^{-3}$. This results

JHEP12(2025)043

suggests that without additional constraints, the Drell-Yan data may not provide sufficient sensitivity to reliably extract $\alpha_s(m_Z)$ together with all relevant nonperturbative parameters.

Fortunately, as discussed in section 3.3, the CS kernel can be constrained using lattice QCD data. Incorporating these constraints — via both the central values from eq. (3.34) and the associated covariance matrix in eq. (3.35) — enables us to include both $\lambda_{2,4}$ in the fit, alongside the TMD PDF parameters $\Lambda_{2,4}$. This yields our nominal fit already shown in figure 10. The $\alpha_s(m_Z)$ uncertainties resulting from this fit are

$$\begin{aligned} \Delta_{\text{nonp}} &= ^{+0.24}_{-0.26} \times 10^{-3}\,, \\ \Delta_{\text{pert}\oplus\text{nonp}} &= ^{+0.69}_{-0.72} \times 10^{-3}\,. \end{aligned} \tag{5.1}$$

While this outcome happens to match closely the case where only the TMD parameters were varied, this is somewhat of a coincidence as it depends on the relative size of the constraints provided by the data and lattice results. Nevertheless, it is very encouraging as it demonstrates that it is possible to consistently account for all relevant nonperturbative parameters in this scenario.

Given the sensitivity of the low-q_T region to nonperturbative effects, and the resulting dependence of $\alpha_s(m_Z)$ on the nonperturbative parameters, one could consider performing fits using only data above a certain q_T threshold, such as $q_T > 5\,\text{GeV}$ or $10\,\text{GeV}$. This would reduce the dependence on nonperturbative effects, potentially allowing for a cleaner determination of $\alpha_s(m_Z)$ at the expense of increased fit uncertainties. A detailed investigation of this approach is left to future work.

6 Conclusions

We have presented an analysis of the expected theory uncertainties in the determination of the strong coupling constant from the Z-boson transverse momentum (q_T) spectrum. The Z q_T spectrum is a benchmark observable of the LHC precision programme and has been measured to subpercent precision. This offers the potential for a high-precision determination of the strong coupling constant, with a percent level change of $\alpha_s(m_Z)$ inducing a similar level change in the shape of the q_T spectrum.

A competitive determination of $\alpha_s(m_Z)$ depends crucially on the control of the theoretical uncertainties. However, to determine whether a 1% theory effect causes a 0.5% or 2% effect on $\alpha_s(m_Z)$, a reliable, precise description of the *shape* of the theory uncertainties is required. Therefore, a robust treatment of theory correlations across the relevant region of q_T is critical.

As the dependence on α_s arises out of the perturbative expansion, one major source of theory uncertainty is the perturbative uncertainty from missing higher orders. Furthermore, with the largest α_s sensitivity originating from the region around 5 GeV, nonperturbative effects are also important. In this work we focus largely upon the perturbative uncertainty, in addition to analysing the nonperturbative effects. To do so, we utilise an Asimov fit setup, taking the binning, experimental uncertainty and correlations of the ATLAS 8 TeV inclusive Z q_T measurement [19] as an example. This allows us to focus our efforts on the dominant expected theory uncertainties on $\alpha_s(m_Z)$ and analyse them in a fully controlled setting without having to worry about statistical fluctuations or possible biases due to unaccounted effects in the data.

JHEP12(2025)043

First, considering the perturbative uncertainty, we find that the traditional approach of varying unphysical scales is insufficient. In particular, in this context it suffers from fundamental limitations as it offers only partial information on the theory uncertainty itself and no insight on its correlations. As a result, the uncertainty one might obtain on $\alpha_s(m_Z)$ from this approach strongly depends on the specific choices and variations performed, demonstrating that it cannot be relied upon to propagate the uncertainty. Our main focus is therefore on exploring the new theory nuisance parameter (TNP) approach to assess this vital source of uncertainty. A key advantage of this approach is that the TNPs encode the higher-order structure of the theory prediction. As such, they are true parameters of the perturbative series and their variation correctly encapsulates the point-by-point correlations of the theory uncertainty. This allows for the first time for a consistent profiling approach to perturbative theory uncertainties, enabling the fit to the experimental data to constrain the theory uncertainty. We utilise our Asimov fit setup to explore this possibility in this context, also serving to demonstrate the benefits of the novel TNP approach in general. We find that profiling the theory uncertainties is vital to reach the level of precision desired on $\alpha_s(m_Z)$.

Secondly, we explore the nonperturbative uncertainty in the extraction of $\alpha_s(m_Z)$. We utilise a nonperturbative parameterization based on a systematic OPE expansion of nonperturbative effects including both the nonperturbative effects of the Collins-Soper kernel and of the TMD PDF, accounting for the quadratic and quartic OPE coefficients of each. We find that reaching a competitive precision on $\alpha_s(m_Z)$ necessitates the incorporation of nonperturbative constraints on the Collins-Soper kernel from lattice QCD.[8]

Whilst these constitute two of the largest sources of uncertainty for the determination of $\alpha_s(m_Z)$, there are a variety of further effects relevant at the desired level of precision. PDF uncertainties are also large, and whilst they can in principle be treated in a straightforward fashion, there are various subtleties which we leave to future work. The TNPs used to account for the perturbative uncertainty are also not yet complete: the x dependence of the beam functions currently utilises the known shape, whilst uncertainties arising from the PDF evolution within the resummation are yet to be incorporated. At the same time, a variety of effects that could be neglected in the Asimov test case need to be included in a fit to real data. These include nonsingular power corrections, heavy quark mass effects, QED/EW corrections, and the x and flavor dependence of nonperturbative TMD effects.

Overall, our analysis demonstrates the crucial role of theory uncertainties and their correlations in the extraction of $\alpha_s(m_Z)$. Reliably propagating these from the q_T spectrum to the strong coupling is a complex task, which nonetheless can be achieved thanks to the TNP approach. Furthermore, the TNP approach allows for a data-driven reduction of perturbative uncertainties. Meanwhile, the nonperturbative uncertainty can be reduced with the help of lattice QCD data. Together, this facilitates the reduction of these key uncertainties to a level which we anticipate will allow for an extraction of $\alpha_s(m_Z)$ from the Drell-Yan q_T spectrum that is competitive with other measurements.

[8] At least when only using the experimentally most precise Drell-Yan data at the Z pole.

JHEP12(2025)043

Acknowledgments

We thank Peter Ploessl and Bahman Dehnadi for providing results with quark-mass effects from ref. [66], and Johannes Michel and Iain Stewart for useful discussions and comments on the manuscript. We also thank our experimental colleagues for many helpful discussions. Last but not least, we thank Johannes Michel, Georgios Billis, and Markus Ebert for their many contributions to the SCETlib code base. 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) and from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — EXC 2121 "Quantum Universe" — 390833306. TC acknowledges funding by Research Foundation-Flanders (FWO) project number: 12E1323N.

A Comments on recent α_s extraction by ATLAS

In this appendix, we comment on the treatment of theory uncertainties in ref. [33], which performed a fit to the Drell-Yan q_T spectrum measured in ref. [19] to extract $\alpha_s(m_Z)$. In the next three subsections we discuss the three uncertainties of perturbative origin that according to our estimates are dominant.[9] In appendix A.4, we comment on the nonperturbative model used in ref. [33].

In table 2, we summarize our estimates of the expected size of perturbative uncertainties based on fits to Asimov data and compare them to the values quoted in ref. [33]. For scale variations, we use the quadrature sum of envelopes here. Using instead the naive total envelope (1.73×10^{-3}) leads to a slightly smaller total perturbative uncertainty of $\pm 2.30 \times 10^{-3}$. Our uncertainty estimates are based on fits to Asimov data, so while the precise numbers will differ in a fit to the real data their typical size will be the same. For the uncertainties due to scale variations, we have explicitly checked in preliminary fits to real data that the size of the obtained uncertainty is indeed very similar. We thus have reason to believe that the total theory uncertainty quoted in ref. [33] is underestimated by a factor of 4 to 5. This would make it by far the dominant uncertainty and as a result also the *total* uncertainty of $\pm 0.9 \times 10^{-3}$ quoted in ref. [33] would be underestimated by about a factor of 3.

A.1 Perturbative uncertainties from scale variations

The analysis in ref. [33] uses the method of enveloped scale variations discussed in section 2.2 to estimate the perturbative theory uncertainty on the resummed q_T spectrum. This uncertainty is propagated to the fitted $\alpha_s(m_Z)$ by scanning over individual scale variations. This yields maximum deviations in $\alpha_s(m_Z)$ of $+0.23 \times 10^{-3}$ and -0.61×10^{-3}, which are symmetrized to the quoted $\pm 0.42 \times 10^{-3}$ scale-variation uncertainty on $\alpha_s(m_Z)$.

As we have argued in the main text, extracting $\alpha_s(m_Z)$ from the small-q_T spectrum crucially relies on shape effects and is therefore extremely sensitive to the theory uncertainty on the shape of the spectrum and the corresponding bin-by-bin theory correlations. As discussed in detail in section 2.2, these are not correctly captured by scale variations, which

[9] Additional perturbative uncertainties are due to QED ISR and the matching to fixed order. They are estimated in ref. [33] to not be significant, which we also believe to be the case, so we do not discuss them here.

JHEP12(2025)043

Perturbative uncertainty	Absolute uncertainty on $\alpha_s(m_Z)$ in units of 10^{-3}	
	Ref. [33]	Our estimate of expected size
Scale variations	± 0.42	± 2.43
N^{4}LL* approximation	± 0.04	
N^{4}LL$'$ approximation		± 0.75
Flavor/quark masses	$+0.40\ -0.29$	± 1.32
Total	$+0.58\ -0.51$	± 2.87

Table 2. Dominant uncertainties of perturbative origin on $\alpha_s(m_Z)$ in units of 10^{-3}. The table compares the values quoted in ref. [33] with our estimate of their expected size based on fits to Asimov data. See text for more details.

means the resulting variations in the fitted $\alpha_s(m_Z)$ cannot be regarded as a trustworthy uncertainty estimate. Scanning over scale variations in section 4.1 we find significantly larger maximum variations in the fitted $\alpha_s(m_Z)$, as seen in the right panel of figure 2. The envelope over all scale variations yields 1.73×10^{-3}, while adding in quadrature the envelopes of different classes of scale variations yields 2.43×10^{-3}, see eqs. (4.1) and (4.2).

Our scale-variation study is performed at N^{4}LL, which is the currently highest known (essentially) complete perturbative order. It yields uncertainties of $0.5 - 1\%$ in the range $5\,\text{GeV} < q_T < 30\,\text{GeV}$ and growing up to $3 - 4\%$ for $0 < q_T < 5\,\text{GeV}$, as seen in the left panel of figure 2. Unfortunately, from the limited information given in ref. [33] it is not clear what the exact size of the scale-variation based uncertainties on the q_T spectrum are at their highest used order (denoted N^{4}LL* in the next subsection). From our own runs of the public version of DYTURBO (v1.4.2) we find spectrum uncertainties of $0.5 - 1\%$ for $q_T < 30\,\text{GeV}$ and growing to $2 - 3\%$ for $q_T \to 0$.[10] Therefore, the envelope of scale variations used in ref. [33] at the level of the q_T spectrum should be of similar size to ours here. This means, for similar-size scale variations in the q_T spectrum we obtain up to five times larger variations in $\alpha_s(m_Z)$.

One difference which could (at least partially) contribute to this large discrepancy is that the scale-variation recipes used in the SCETLIB and DYTURBO resummation frameworks differ. Whilst both use the conventional factor of two for varying scales up and down and avoid simultaneous variations that would compound to factors of four in any scale ratios, the specific types of scale variations differ substantially and are more limited in DYTURBO. In particular, the low scale $\sim q_T \sim 1/b_T$ is varied as part of the scale variation setup in SCETLIB, while it is kept fixed in DYTURBO. We stress that there are no compelling theoretical arguments for keeping this low scale fixed when estimating uncertainties using scale variations.[11]

[10] Similarly, in ref. [30], numerical results for the q_T spectrum at this N^{4}LL* order are given for $E_{\text{cm}} = 13\,\text{TeV}$ and including fiducial lepton cuts, in which case the spectrum uncertainties are also close to around 1%.

[11] This low scale is genuine to the resummation problem, as discussed in the following subsection, and therefore must effectively appear in any resummation framework. It is also kept variable in the original CSS framework [75] on which DYTURBO is based. In the context of an ongoing benchmarking study of various q_T resummation frameworks within the LHC electroweak working group, the need for varying this low scale has been widely recognized by now and, as far as we can tell, has been gradually adopted also by most other resummation frameworks.

JHEP12(2025)043

As discussed in section 2, in a situation like this where theory correlations are critically important, for the same overall uncertainty in the spectrum, the resulting uncertainty on $\alpha_s(m_Z)$ entirely depends on the specific scale variations that happen to be included in the scanning. A similar breakdown into individual scale variations as in our figure 2 of the results in ref. [33] might shed some more light on this. Ultimately, however, the large discrepancies we observe between different resummation frameworks using different scale variation recipes likely just reflect the fact that the scale-variation method cannot be relied upon for propagating the perturbative uncertainty from the q_T spectrum to $\alpha_s(m_Z)$.

Nevertheless, if scale variations are the method of choice, one must be aware of their limitations and at minimum it must be ensured that a given scale-variation recipe does not underestimate the resulting uncertainty on the parameter of interest. For this purpose, we have clear evidence that the type of scale variations performed in SCETlib yield a more realistic uncertainty for $\alpha_s(m_Z)$. This is because when scanning over TNPs, which do encode the correct theory correlations, spectrum uncertainties of comparable (or even somewhat smaller) size translate into an uncertainty on $\alpha_s(m_Z)$ of 1.75×10^{-3} (see section 4.2.1 and figure 3), which is similar to what we find with scale variations.

We also stress that the much reduced perturbative uncertainty on $\alpha_s(m_Z)$ that we find after profiling the TNPs *cannot* be taken to justify the much smaller scale-variation based uncertainties quoted in ref. [33], because the TNP profiling allows the data to improve the overall precision of the theory description (and for example also induces nontrivial post-fit correlations between different TNPs). In other words, the resulting theory description obtained as an output after profiling the TNPs has an intrinsically higher precision than what can be ascribed to a traditional input theory prediction based on scale variations.

We finally note that another possible difference is that we are fitting only $\alpha_s(m_Z)$ when scanning over scale variations in section 4.1, whereas in ref. [33] also the PDFs and some nonperturbative parameters are included in the fit. In principle, it could happen that other fit parameters absorb some of the scale variations such that their impact on $\alpha_s(m_Z)$ is reduced. For this to be the case, the effect of a particular scale variation would have to match the shape change induced by adjusting another fit parameter better than that induced by adjusting $\alpha_s(m_Z)$, which would make the sensitivity to the precise shape of the theory uncertainty even more relevant. By including nonperturbative parameters or PDFs in our fit while scanning over TNPs (as they provide the correct theory uncertainty shapes), we find that this is not the case, i.e., including these additional fit parameters does not alter the impact of the perturbative uncertainties on $\alpha_s(m_Z)$.

A.2 Different N^{4}LL countings and approximation uncertainties

Different nomenclatures for resummation orders involving Sudakov double-logarithms appear in the literature. A common nomenclature, which we also use here, is such that at N^{n}LL one includes the n-loop evolution (i.e. n-loop noncusp and $(n+1)$-loop cusp anomalous dimensions) together with $(n-1)$-loop fixed-order boundary conditions of the hard, beam, and soft functions in our notation. At N^{n}LL$'$ order the boundary conditions are included at one order higher, i.e., at n-loop order and as they appear at N^{n+1}LL. As a result, the

JHEP12(2025)043

	boundary conditions		anomalous dimensions	
resummation order	h	$\tilde{b}$, $\tilde{s}$	$\gamma^\mu, \gamma^\nu, P_{ij}$	$\Gamma_{\text{cusp}}, \beta$
$\text{N}^3\text{LL}'$	$\alpha_s^3(Q)$	$\alpha_s^3(b_0/b_T)$	α_s^3	α_s^4
N^4LL	$\alpha_s^3(Q)$	$\alpha_s^3(b_0/b_T)$	α_s^4	α_s^5
$\text{N}^4\text{LL}^\star$	$\alpha_s^4(Q)$	$\alpha_s^3(b_0/b_T) + \alpha_s^4(Q)$	α_s^4	α_s^5
$\text{N}^4\text{LL}'$	$\alpha_s^4(Q)$	$\alpha_s^4(b_0/b_T)$	α_s^4	α_s^5
N^5LL	$\alpha_s^4(Q)$	$\alpha_s^4(b_0/b_T)$	α_s^5	α_s^6

Table 3. Order counting conventions. Here, $\text{N}^4\text{LL}^\star$ refers to what is called N^4LL in refs. [30, 33]. In practice, the first three rows have a comparable perturbative precision, as have the last two rows. See the text for further details.

perturbative precision at $\text{N}^n\text{LL}'$ is often close to N^{n+1}LL and thus much higher than N^nLL. This is summarized in table 3.

The N^nLL nomenclature used in refs. [30, 33] also includes n-loop boundary terms, and for clarity we will refer to it as N^nLL^* here. According to ref. [30], this N^4LL^* supposedly provides a reliable approximation of $\text{N}^4\text{LL}'$, with the implied higher perturbative precision compared to standard N^4LL. However, there is a crucial difference in how the boundary terms are included at N^nLL^*, making it rather different from $\text{N}^4\text{LL}'$ and with a perturbative precision more comparable to standard N^4LL, as indicated in table 3 and as we will discuss now.

Let us first consider the soft function. As discussed in section 3.2, the full soft function $\tilde{S}(b_T, \mu, \nu)$ is a perturbative series in $\alpha_s(\mu)$ and contains Sudakov logarithms $\ln^n(\mu b_T/b_0)$ and $\ln^n(\nu b_T/b_0)$. Its resummation proceeds by evaluating it at its canonical scales $\mu_S^{\text{can}} = \nu_S^{\text{can}} = b_0/b_T$ [see eqs. (3.15) and (3.8)], which defines its boundary condition,

$$\tilde{s}[\alpha_s(b_0/b_T)] \equiv \tilde{S}(b_T, \mu = b_0/b_T, \nu = b_0/b_T) = \sum_{n=0} \tilde{s}_n \left[\frac{\alpha_s(b_0/b_T)}{4\pi}\right]^n , \tag{A.1}$$

where the $\tilde{s}_n$ are the so-defined n-loop soft boundary coefficients. The boundary series in eq. (A.1) is by construction free of large logarithms and can thus be evaluated at fixed order in α_s. It is the starting point of the evolution from the low scale $1/b_T \sim q_T$ to the high scale Q, which resums the logarithms of $Qb_T \sim Q/q_T$.

The analogous beam boundary condition was already defined in eq. (3.6),

$$\tilde{b}_i[x, \alpha_s(b_0/b_T)] \equiv \tilde{B}_i(x, b_T, \mu = b_0/b_T, \nu/Q = 1) = \sum_{n=0} \tilde{b}_{i,n}(x) \left[\frac{\alpha_s(b_0/b_T)}{4\pi}\right]^n , \tag{A.2}$$

where as shown in eq. (3.7) the beam boundary coefficients $\tilde{b}_{i,n}(x)$ are computed in terms of PDFs evaluated at $\mu_f = b_0/b_T$.

Note that the beam and soft boundary conditions in eqs. (A.1) and (A.2) *genuinely* appear as a low-scale perturbative series in $\alpha_s(b_0/b_T)$. The appearance of a perturbative series at the low scale $\mu \sim b_0/b_T \sim q_T$ is a fundamental and universal feature of q_T resummation and not specific to a particular resummation framework. (It is also the reason why the beam and soft functions become nonperturbative at sufficiently small $q_T \sim 1/b_T$, see section 3.3.)

JHEP12(2025)043

By contrast, the canonical scale of the hard function is $\mu_H^{\text{can}} = Q$, so its boundary condition is defined as

$$\tilde{h}[\alpha_s(Q)] \equiv H(Q^2, \mu = Q) = \sum_{n=0} \tilde{h}_n \left[\frac{\alpha_s(Q)}{4\pi}\right]^n , \tag{A.3}$$

and thus it genuinely appears as a high-scale perturbative series in $\alpha_s(Q)$.

We can now state the difference between N^nLL and $\text{N}^n\text{LL}'$ more explicitly: At N^nLL, the hard, beam, and soft boundary series in eqs. (A.1), (A.2), and (A.3) are included in eq. (3.15) up to fixed $\mathcal{O}(\alpha_s^{n-1})$ with their respective canonical α_s. At $\text{N}^n\text{LL}'$, they are instead included up to $\mathcal{O}(\alpha_s^n)$ *exactly* as they would be at N^{n+1}LL. Thus, the only difference between $\text{N}^n\text{LL}'$ and the full N^{n+1}LL is that the latter further includes the anomalous dimensions in the evolution factor in eq. (3.15) to one order higher, see table 3.

Toward higher orders in q_T resummation, the contributions from boundary terms become the dominant effect while the anomalous dimensions become less important. The $\text{N}^n\text{LL}'$ result is therefore numerically close to the full N^{n+1}LL result. For example, the SCETlib results at $\text{N}^3\text{LL}'$ and N^4LL are almost indistinguishable in both central value and scale variation uncertainties [32].[12] Furthermore, as the beam and soft boundary coefficients come with a low-scale α_s, they are much more important than the hard function. We have seen this explicitly in section 4.2.

In DYTurbo, following refs. [109, 110], the low-scale boundary series in the CSS formalism [75], which corresponds to our soft and beam boundary series, is rewritten to superficially appear as high-scale perturbative series as follows (taking the soft function as an example):

$$\begin{aligned}
\tilde{s}[\alpha_s(b_0/b_T)] &= \tilde{s}[\alpha_s(Q)] \times \frac{\tilde{s}[\alpha_s(b_0/b_T)]}{\tilde{s}[\alpha_s(Q)]} = \tilde{s}[\alpha_s(Q)] \times \exp\left\{\ln \frac{\tilde{s}[\alpha_s(b_0/b_T)]}{\tilde{s}[\alpha_s(Q)]}\right\} \\
&= \tilde{s}[\alpha_s(Q)] \times \exp\left\{-\int_{\alpha_s(b_0/b_T)}^{\alpha_s(Q)} \mathrm{d}\alpha_s \, \frac{\mathrm{d}\ln \tilde{s}(\alpha_s)}{\mathrm{d}\alpha_s}\right\} \\
&= \tilde{s}[\alpha_s(Q)] \times \exp\left\{-\int_{b_0/b_T}^{Q} \mathrm{d}\ln\mu \, \frac{\mathrm{d}\alpha_s}{\mathrm{d}\ln\mu} \, \frac{\mathrm{d}\ln \tilde{s}(\alpha_s)}{\mathrm{d}\alpha_s}\right\} .
\end{aligned} \tag{A.4}$$

The first line is clearly trivial, and in the remaining lines the exponential is further cast into a form that resembles the exponential evolution factor in eq. (3.15). As long as $\tilde{s}(\alpha_s)$ is consistently included to the same order in α_s everywhere it appears, eq. (A.4) is a strict identity, i.e., there is no formal or numerical difference between the left-hand side and the various versions on the right-hand side.[13] This is the case at N^nLL and $\text{N}^n\text{LL}'$.

However, in the order-counting adopted at N^nLL^*, the last line of eq. (A.4) is used with the factor $\tilde{s}[\alpha_s(Q)]$ considered as a fundamental high-scale boundary series, and akin to the treatment of the hard boundary series at $\text{N}^n\text{LL}'$, it is included to n-loop order (i.e. $\tilde{s}_n$). At the same time, the exponential factor is considered to be part of the overall evolution factor, which is included to n-loop order (akin to the evolution order used at $\text{N}^n\text{LL}'$). As a result,

[12] This applies when using common PDFs. Switching from NNLO to approximate N^3LO PDFs, as also formally required for the PDF evolution when switching from $\text{N}^3\text{LL}'$ to N^4LL, does make a noticeable difference.

[13] For the last line in eq. (A.4), this also requires the running of $\alpha_s(\mu)$ to be treated in the same way on both sides of the equation. Otherwise there will be formally higher-order differences which are however irrelevant for our discussion.

JHEP12(2025)043

the $\tilde{s}(\alpha_s)$ boundary series in the exponent is only included to $(n-1)$-loop order (i.e. to $\tilde{s}_{n-1}$). This is because the running of α_s starts at one loop, so the $(n-1)$-loop boundary coefficient $\tilde{s}_{n-1}$ only appears as an n-loop contribution as $\beta_0\,\tilde{s}_{n-1}$.

However, rewriting the boundary series as in eq. (A.4) does not change the fundamental fact that it is a low-scale series. This fact is merely hidden on the right-hand side. Reversing the steps in eq. (A.4) with the N^{n}LL* counting applied to the last line on the right-hand side, it results in effectively using the following n-loop boundary series on the left-hand side,

$$\tilde{s}^{(n)}[\alpha_s(Q)] \times \frac{\tilde{s}^{(n-1)}[\alpha_s(b_0/b_T)]}{\tilde{s}^{(n-1)}[\alpha_s(Q)]} = \sum_{k=0}^{n-1} \tilde{s}_k \left[\frac{\alpha_s(b_0/b_T)}{4\pi}\right]^k + \tilde{s}_n \left[\frac{\alpha_s(Q)}{4\pi}\right]^n , \tag{A.5}$$

where the superscripts denote the order in α_s to which each series is kept. Hence, due to the mismatch in the treatment of $\tilde{s}(\alpha_s)$ between prefactor and exponent, the n-loop coefficient $\tilde{s}_n$ remains multiplied by the high-scale $\alpha_s^n(Q)$, as indicated in table 3, while its correction to the low scale is postponed to N^{n+1}LL*. By contrast, it is directly included at its correct low scale at both N^{n}LL$'$ and full N^{n+1}LL.

The analogous discussion applies to the beam boundary series, which is rewritten in the same fashion to appear as a high-scale series, including also PDFs evaluated at the high scale, multiplied by a corresponding correction factor that also includes the evolution of the PDFs back down to their natural low scale. At N^{n}LL*, the beam boundary terms are included at their correct low scale only to $\tilde{b}_{i,n-1}$, while $\tilde{b}_{i,n}$ is included in terms of α_s and PDFs at the high scale.

As a side remark, we comment on the counting of towers of logarithms $\alpha_s^n L^m$ in the perturbative series, with some choice of the large logarithm L, which traditionally is often used to define resummation orders. When counting logarithms in the cross section (formally counting $\alpha_s L^2 \sim 1$), N^{n}LL* and N^{n}LL$'$ capture the same complete logarithmic towers and one more than N^{n}LL. When counting logarithms in the exponent (formally counting $\alpha_s L \sim 1$), all three orders capture the same complete logarithmic towers. This does not mean that the orders are equivalent for all purposes. The three orders differ by including different subsets of higher logarithmic towers, which are well defined and numerically relevant. This simply means that their relevant distinction is not (or not easily) visible from a traditional counting of logarithms. Instead, here we define the resummation order (or formal accuracy) at the level of the RGE and its ingredients, which allows us to make a formal distinction between the orders as shown in table 3.

This rather peculiar treatment of the low-scale boundary conditions at N^{n}LL* has several consequences. Since $\alpha_s(b_0/b_T)/\alpha_s(Q)$ can become $\gtrsim 2$ at small $1/b_T \sim q_T$, and this mismatch is raised to the nth power, the above-mentioned dominant effect of the n-loop soft and beam boundary coefficients is not yet present at N^{n}LL*. For example, the effect of $\tilde{s}_4$ and $\tilde{b}_{i,4}$ can be easily suppressed by an order of magnitude or more at N^{4}LL*, which is indeed what we find below. Their full impact then only appears at the next order in the disguise of a large evolution effect. Moreover, since they are only included as a high-scale boundary condition like the hard function, also their effect on the shape of the spectrum, which is particularly relevant for extracting α_s, is not yet present. For these reasons, the N^{4}LL* result does not provide a tangible improvement, formally or practically, compared to N^{4}LL. This

JHEP12(2025)043

is confirmed by the comparable scale variations we observed at the spectrum level in the previous subsection. In particular, and regardless of being approximate or exact, it does not provide a reliable approximation of N^{4}LL$'$.

The analysis in ref. [33] is performed at the so-called approximate N^{4}LL ("N^{4}LLa") order of ref. [30], meaning at approximate N^{4}LL*. There are two types of approximations involved: First, the 4-loop evolution is formally approximate, because the required 4-loop splitting functions and 5-loop cusp anomalous dimensions are only known approximately. These approximations are also present at N^{4}LL. They are expected to cause more minor effects and are not discussed here further.[14] The second type of approximation comes from the fact that the 4-loop beam and soft coefficients, $\tilde{s}_4$ and $\tilde{b}_{i,4}$, are completely unknown.

In refs. [30, 33] the additional uncertainty due to the 5-loop cusp anomalous dimension and the unknown 4-loop beam and soft coefficients are estimated to be $0.1 - 0.2\%$ for the q_T spectrum with a resulting negligible $\pm 0.04 \times 10^{-3}$ effect on $\alpha_s(m_Z)$. Since at N^{4}LL* the latter are effectively treated like a hard-function contribution, this is consistent with the rather negligible impact of the hard function and cusp anomalous dimension we observe. However, as explained above, this does not provide a realistic estimate of the actual possible impact of the missing 4-loop beam and soft coefficients.

The dominant effect of the beam and soft functions is clearly evident in our TNP studies, where as expected they cause by far the largest uncertainties. From our TNP scanning at N^{3+1}LL, we know that the 3-loop beam and soft functions each cause an effect of order $\sim 1 \times 10^{-3}$ on $\alpha_s(m_Z)$. We might expect the corresponding uncertainties due to the 4-loop beam and soft functions to be between two and a few times smaller than that, but certainly not 30 to 40 times smaller. We can obtain a more realistic estimate of the impact of the missing 4-loop beam and soft functions from their associated TNPs at N^{4+0}LL (see appendix B.2).[15] From the analogous TNP scanning at this order we find a combined perturbative uncertainty on $\alpha_s(m_Z)$ due to the missing 4-loop beam and soft functions of 0.75×10^{-3}, as given in table 2, so indeed a reduction by roughly a factor of 2 compared to N^{3+1}LL in line with expectation. Their full effect would appear at N^{5}LL* and already by itself can be twice as large as the quoted scale-variation uncertainty at N^{4}LL*, providing further evidence that the latter is strongly underestimated.

A.3 Effects due to finite bottom and charm quark masses

The masses of charm and bottom quarks, m_c and m_b, are usually neglected, which is an appropriate description in the limit $m_{b,c} \ll q_T$. On the other hand, for $q_T \sim m_{b,c}$, which in the case of m_b is right in the peak region of the q_T spectrum, they cause nontrivial effects at the few-percent level. We thus expect them to be important at our level of precision.

For $q_T \sim m_b$, the mass effects genuinely change the resummation structure [65]. They can be classified into primary and secondary mass effects. Primary mass effects are due to the quark masses of the incoming quarks in the $b\bar{b} \to Z/\gamma^*$ and $c\bar{c} \to Z/\gamma^*$ hard processes.

[14] Although the available approximations for 4-loop splitting functions are increasingly accurate for phenomenology [111], their effect could in principle be of some relevance and deserves further study.

[15] Ideally, we would use N^{4+1}LL, which we have not yet available, but N^{4+0}LL provides a sufficient approximation for our purposes here.

JHEP12(2025)043

Secondary mass effects are due to the quark masses in virtual charm and bottom quark loops as well as final-state gluons splitting into massive quarks.

It is important to note that switching from a massless 5-flavor description to a massless 4-flavor description at some threshold scale $\mu_b \sim m_b$ is not sufficient to correctly account for bottom mass effects, since in the region $q_T \sim m_b$ neither description is actually valid. A correct description of the bottom threshold actually requires the correct treatment of the bottom mass effects. The same holds for the charm threshold, which due to the smaller charm mass is less relevant in practice.

Ref. [33] uses a massless 5-flavor description by default. The uncertainties due to neglecting quark-mass effects are estimated by changing to a variable-flavor massless description in various parts of the calculation and by varying the associated threshold scales. From the above discussion, this can only serve as a rough indication of the expected order of magnitude of mass effects but it may or may not reflect the actual size of the effects. Another estimate comes from including the effect of final-state gluon splitting into massive quark pairs. Unfortunately, ref. [33] provides no further information than that, so it is unclear to us to what extent this provides a complete description of secondary mass effects. The final uncertainty on $\alpha_s(m_Z)$ due to mass effects is taken as the envelope of all observed differences yielding $^{+0.40}_{-0.29} \times 10^{-3}$.

A full treatment of quark mass effects in the resummed q_T spectrum at NNLL′ based on ref. [65] has been implemented in SCETlib and will be presented elsewhere [66]. Here, we use it to estimate the bias induced in the extracted $\alpha_s(m_Z)$ from using a massless description. To do so, we generate Asimov data including bottom and charm mass effects and fit it with a 5-flavor massless theory model at the same resummation order. We find a bias on $\alpha_s(m_Z)$ of 1.32×10^{-3}, as given in table 2. The effect is entirely driven by the bottom quark mass. This rather strong sensitivity to mass effects in $\alpha_s(m_Z)$ is not unexpected, given that the dominant sensitivity to $\alpha_s(m_Z)$ arises from the peak region of the spectrum $q_T \sim 5\,\text{GeV}$, which is precisely where bottom mass effects are important. And as we have seen before, order percent shape changes in the q_T spectrum can easily cause changes of order $\sim 1 \times 10^{-3}$ in $\alpha_s(m_Z)$.

A.4 Nonperturbative model

In ref. [33], nonperturbative effects are incorporated via the following phenomenological model

$$\begin{aligned} S_{\text{NP}}(b_T) &= \exp\left[-g_j(b_T) - g_K(b_T)\ln\frac{Q^2}{Q_0^2}\right], \\ g_j(b_T) &= \frac{g\,b_T^2}{\sqrt{1+\lambda\,b_T^2}} + \text{sign}(q)\left[1-\exp(-|q|b_T^4)\right], \\ g_K(b_T) &= g_0\left\{1-\exp\left[-\frac{1}{g_0}\,\frac{C_F\alpha_s(b_0/b_*)}{\pi}\,\frac{b_T^2}{b_{\text{lim}}^2}\right]\right\}, \end{aligned} \tag{A.6}$$

which must account for the product $\tilde{f}_i^{\text{np}}(x,b_T)\tilde{f}_j^{\text{np}}(x,b_T)$ in our notation of section 3.3, and where $g_j(b_T)$ and $g_K(b_T)$ encode the nonperturbative models for the TMD PDF and CS kernel, respectively.

This model originates from ref. [94], and while it has traditionally been used in the past, it has several limitations that are potentially important at the high level of precision

JHEP12(2025)043

required here. First, the model does not account for any flavor or x dependence in the nonperturbative TMD parameters, which in general cannot be neglected [31, 99, 100]. As discussed in section 3.3.2, since the fit to real data includes multiple rapidity bins, the nonperturbative TMD model should at minimum allow for an effective rapidity dependence.

Secondly, the model does not reproduce the OPE expansion in eq. (3.28). While it has quadratic and quartic TMD parameters, g and q (equivalent to our Λ_2 and Λ_4), which are included as fit parameters, it does not have parameters to encode the genuine nonperturbative contributions to the CS kernel. In particular, the $\mathcal{O}(b_T^2)$ piece of $g_K(b_T)$, only cancels the $\mathcal{O}(b_T^2)$ contribution induced by the quadratic b_* prescription used in the perturbative part of the CS kernel. This leaves the genuine nonperturbative OPE contribution due to λ_2^ζ unaccounted for, which as discussed in section 3.3 is the parametrically *leading* nonperturbative effect.

Furthermore, the canonical $\ln(b_T Q)$ multiplying the CS kernel is replaced by a fixed $\ln(Q/Q_0)$. This amounts to shifting a corresponding $\ln(b_T Q_0)$ dependence into the nonperturbative TMD, which however is not accounted for in $g_j(b_T)$. This effectively neglects a nontrivial b_T (and thus q_T) dependence, which could easily spoil the point-by-point correlations of the nonperturbative uncertainties.

In summary, the model does not provide a correct parameterization of nonperturbative effects in the important region of moderately small q_T or $1/b_T$. The effect of this on the extraction of $\alpha_s(m_Z)$ is a priori unclear and needs further study. For example, as indicated by our results in section 5, including λ_2 as a free parameter in the fit can significantly enlarge the fit uncertainties. Whilst the model has several additional parameters (λ, g_0, b_{lim}, Q_0), which are being varied to estimate additional nonperturbative uncertainties, they do not assess the impact of these limitations. Ref. [33] also performed a fit excluding the region of $q_T \leq 5\,\text{GeV}$, where nonperturbative effects are expected to have their dominant impact. This provides a potentially useful test, which however is only necessary and not sufficient. It would invalidate the model if it gave inconsistent results, but the reverse is not true — if it gives consistent results this does not validate the model.

B Additional results

B.1 Nonsingular power corrections

In this appendix, we briefly investigate the impact of including the nonsingular contribution in eq. (3.1), which has been neglected in our main analysis. It is suppressed by $\mathcal{O}(q_T^2/Q^2)$ relative to the leading-power resummed contribution $\mathrm{d}\sigma^{(0)}$, which captures the leading behaviour of the q_T spectrum in the $q_T \to 0$ limit. Since the nonsingular is subdominant for $q_T \ll Q$, it is typically evaluated at fixed order in α_s at the hard scale $\mu \sim Q$ as the difference

$$\frac{\mathrm{d}\sigma_{\text{nons}}}{\mathrm{d}q_T} = \left[\frac{\mathrm{d}\sigma_{\text{FO}}}{\mathrm{d}q_T} - \frac{\mathrm{d}\sigma_{\text{FO}}^{(0)}}{\mathrm{d}q_T}\right]_{q_T>0}, \tag{B.1}$$

where $\mathrm{d}\sigma_{\text{FO}}$ is the full fixed-order result and $\mathrm{d}\sigma_{\text{FO}}^{(0)}$ is the fixed-order expansion of the leading-power resummed contribution.

The full $\mathcal{O}(\alpha_s)$ term, called LO_1, is implemented in SCETlib, while the full $\mathcal{O}(\alpha_s^2)$ contribution, referred to as NLO_1, is obtained numerically using DYturbo [112]. We are

JHEP12(2025)043

Asimov data	theory model	$\alpha_s(m_Z)\,[10^{-3}]$
$\mathrm{N^4LL{+}NLO_1}$	$\mathrm{N^{3+1}LL}$	116.34 ± 0.41
$\mathrm{N^4LL{+}NLO_1}$	$\mathrm{N^{3+1}LL{+}LO_1}$	118.81 ± 0.47
$\mathrm{N^4LL{+}NLO_1}$	$\mathrm{N^{3+1}LL{+}NLO_1}$	$118.09^{+0.46}_{-0.49}$

Table 4. Results for $\alpha_s(m_Z)$ in units of 10^{-3} and its perturbative uncertainty from Asimov tests including different nonsingular contributions.

primarily interested in checking whether the nonsingular can influence the uncertainty in $\alpha_s(m_Z)$ due to the dominant leading-power uncertainties. We therefore add the nonsingular terms directly to the resummed contribution, without introducing additional nuisance parameters for them. For computational efficiency, here we include the nonsingular terms only up to $\mathrm{NLO_1}$. We follow the setup described in section 3.4, with one minor modification: The difference in eq. (B.1) is subject to large numerical cancellations for $q_T \to 0$. To avoid the associated numerical instabilities we begin the q_T binning at 0.5 GeV.

We perform a set of Asimov tests following the procedure in section 4.2.3. For simplicity, we do not fit the nonperturbative parameters here, since doing so does not change our conclusions. The nominal perturbative uncertainty for this case obtained in section 4.2.3 is $\Delta_{\mathrm{pert}} = ^{+0.47}_{-0.50} \times 10^{-3}$. The Asimov data are now defined as the $\mathrm{N^4LL{+}NLO_1}$ prediction at $\alpha_s(m_Z) = 0.118$, which we fit with three different theory models that differ in the nonsingular terms they include. The results are shown in table 4. In the case where no nonsingular term is included (top row), we find an uncertainty $\Delta_{\mathrm{pert}} = 0.41 \times 10^{-3}$, slightly smaller than the nominal one. Including the $\mathrm{LO_1}$ and $\mathrm{NLO_1}$ terms (middle and bottom rows), the uncertainty increases again marginally and closely reproduces the nominal one. We conclude that under the assumptions and setup of our pseudodata analysis, the nonsingular contributions do not affect the uncertainty of $\alpha_s(m_Z)$. They can therefore be safely neglected for our purposes. However, as also seen in table 4, neglecting the nonsingular corrections in the theory model induces a substantial bias in the central value of $\alpha_s(m_Z)$, which will necessitate their inclusion in the fit to real data.

B.2 Approximate TNP orders

Here, we briefly discuss an approximate TNP implementation, following ref. [34], that may be useful in cases where incorporating the full next-order structure is not feasible for practical reasons. As an approximation to the $\mathrm{N^{m+1}LO}$ implementation, one can consider reusing the structure of the existing $\mathrm{N^mLO}$ prediction, absorbing the uncertainty term into the highest known order. For example, compared to $\mathrm{N^{1+1}LO}$ and $\mathrm{N^{2+1}LO}$ in eq. (2.9), we have

$$\begin{aligned} \mathrm{N^{1+0}LO}: &\qquad f(x,\alpha,\theta_2) = \hat{f}_0(x) + \left[\hat{f}_1(x) + \alpha_0 f_2(x,\theta_2)\right]\alpha\,, \\ \mathrm{N^{2+0}LO}: &\qquad f(x,\alpha,\theta_3) = \hat{f}_0(x) + \hat{f}_1(x)\,\alpha + \left[\hat{f}_2(x) + \alpha_0 f_3(x,\theta_3)\right]\alpha^2\,. \end{aligned} \tag{B.2}$$

This approximation is expected to roughly preserve the overall size of the uncertainty due to θ_2 or θ_3, while its shape (and thus the correlation) is approximated based on the lower-order structure.

JHEP12(2025)043

Here, α_0 is a chosen fixed number, which does not depend on a renormalization scale and does not participate in counting powers of α_s. For a single perturbative series with no further internal structure, this approximation is somewhat academic. If α_s is evaluated at a fixed central scale and α_0 is chosen to agree with that value, then the central $\mathrm{N}^{m+0}\mathrm{LO}$ result would reproduce the central $\mathrm{N}^{m+1}\mathrm{LO}$ result.

However, as soon as the prediction in question is more complicated, e.g., α_s is evaluated at a dynamic scale or the total prediction involves the combination of several perturbative series expanded against each other, this approximation becomes nontrivial. This is the case with resummation, for which the $\mathrm{N}^{m+0}\mathrm{LL}$ result is defined by absorbing the highest TNP-parameterized terms of all anomalous dimensions and boundary conditions into their respective previous order terms analogous to eq. (B.2). For example, compared to $\mathrm{N}^{3+1}\mathrm{LL}$ in eq. (3.19), we now have

$$\begin{aligned}\mathrm{N}^{3+0}\mathrm{LL}: \quad F(\alpha_s^{\mathrm{can}}, L) = \Bigg[1 + \hat{F}_1 \frac{\alpha_s^{\mathrm{can}}}{4\pi} + \Big(\hat{F}_2 + a_0\, F_3(\theta_3^F)\Big)\left(\frac{\alpha_s^{\mathrm{can}}}{4\pi}\right)^2\Bigg] \exp\Bigg\{\int_0^L \mathrm{d}L' \Bigg[\\ \sum_{k=0}^{2} \hat{\Gamma}_k \Big[\frac{\alpha_s(L')}{4\pi}\Big]^{k+1} L' + \Big(\hat{\Gamma}_3 + a_0\, \Gamma_4(\theta_4^\Gamma)\Big)\Big[\frac{\alpha_s(L')}{4\pi}\Big]^4 L' \\ + \sum_{k=0}^{1} \hat{\gamma}_{F,k} \Big[\frac{\alpha_s(L')}{4\pi}\Big]^{k+1} + \Big(\hat{\gamma}_{F,2} + a_0\, \gamma_{F,3}(\theta_3^{\gamma_F})\Big)\Big[\frac{\alpha_s(L')}{4\pi}\Big]^3\Bigg]\Bigg\}, \end{aligned} \tag{B.3}$$

where $a_0 \equiv \alpha_0/(4\pi)$ and α_s^{can} is again evaluated at the canonical scale of F, which is usually a dynamic scale. At a more generalized boundary scale, as used in the actual resummation, the full μ-dependent boundary condition as predicted by the RGE in terms of $\hat{F}_1$, $\hat{F}_2$, $\hat{\Gamma}_0$, $\hat{\Gamma}_1$, $\hat{\gamma}_{F,0}$, $\hat{\gamma}_{F,1}$ is used with $\hat{F}_2$ replaced by $\hat{F}_2 + a_0\, F_3(\theta_3^F)$ as above. The products of boundary conditions in the full cross section are expanded keeping terms up to $\mathcal{O}(\alpha_s^2)$ (without counting powers of α_0). Hence, at $\mathrm{N}^{3+0}\mathrm{LL}$ the TNP-dependent uncertainty terms associated with the missing $\mathrm{N}^4\mathrm{LL}$ are absorbed into the $\mathrm{N}^3\mathrm{LL}$ structure, while at $\mathrm{N}^{3+1}\mathrm{LL}$ they appear in the correct $\mathrm{N}^4\mathrm{LL}$ structure. As emphasized before, using the full $\mathrm{N}^{m+1}\mathrm{LL}$ result should always be preferred if possible. However, in case it is unavailable or impractical, the $\mathrm{N}^{m+0}\mathrm{LL}$ result can serve as a viable approximation for it.

In figure 7, we already included the results for $\alpha_s(m_Z)$ at $\mathrm{N}^{3+0}\mathrm{LL}$ and $\mathrm{N}^{4+0}\mathrm{LL}$. Here, we provide the corresponding plots for the q_T spectrum and TNP pull plots. The Asimov fit follows the same setup as in section 4.2.3. The results at $\mathrm{N}^{3+0}\mathrm{LL}$ are shown in figure 12. In the left panel, we observe a significant reduction in uncertainty from the pre-fit to the post-fit bands. This reduction is particularly strong due to the high degree of correlation in the fit. Such a reduction may be suspicious, and indeed, the TNP pull plot in the right panel provides insights into the reasons for this behaviour. Some TNPs, particularly the beam functions, become strongly constrained by the data. When the constraints on these parameters become this tight, it suggests that the considered order may not be sufficient to correctly account for the remaining theory uncertainties. This is an example of overfitting the theory model, as previously discussed in section 4.2.3. Regarding the TNPs' impacts on $\alpha_s(m_Z)$, we find that γ_ν, S and B_{qg} have the strongest impact on α_s. This is consistent with what we observed in the $\mathrm{N}^{3+1}\mathrm{LL}$ case. In contrast, at $\mathrm{N}^{2+1}\mathrm{LL}$, the pattern of TNP constraints and pulls was quite different.

JHEP12(2025)043

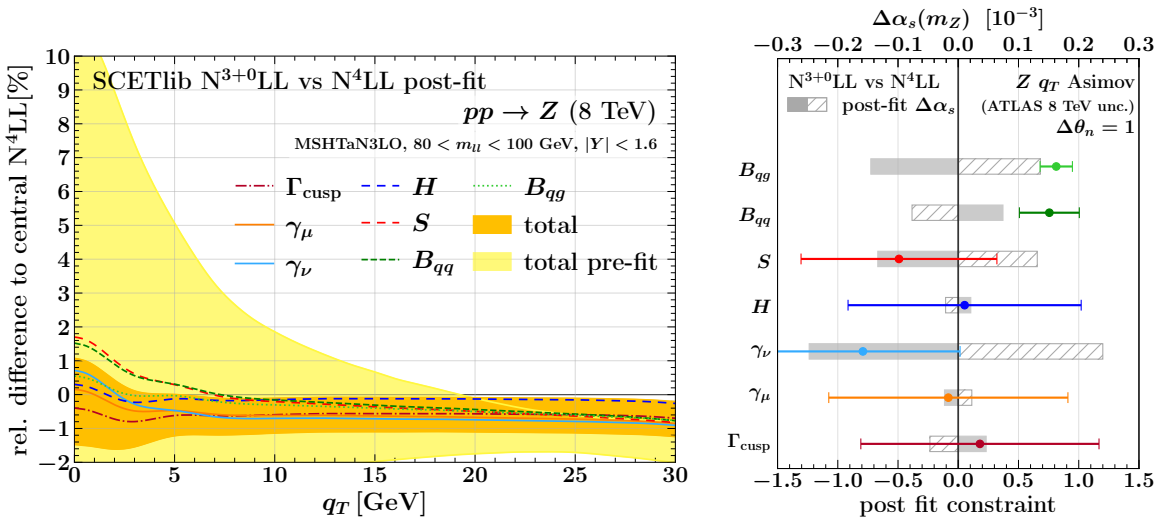


Figure 12. Left panel: Uncertainties in the q_T spectrum at N^{3+0}LL relative to the N^{4}LL, before (yellow band) and after (orange band) profiling the TNPs. The different lines show the post-fit relative impact of each TNP. Right panel: Post-fit constraints on the TNPs (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) TNP variations.

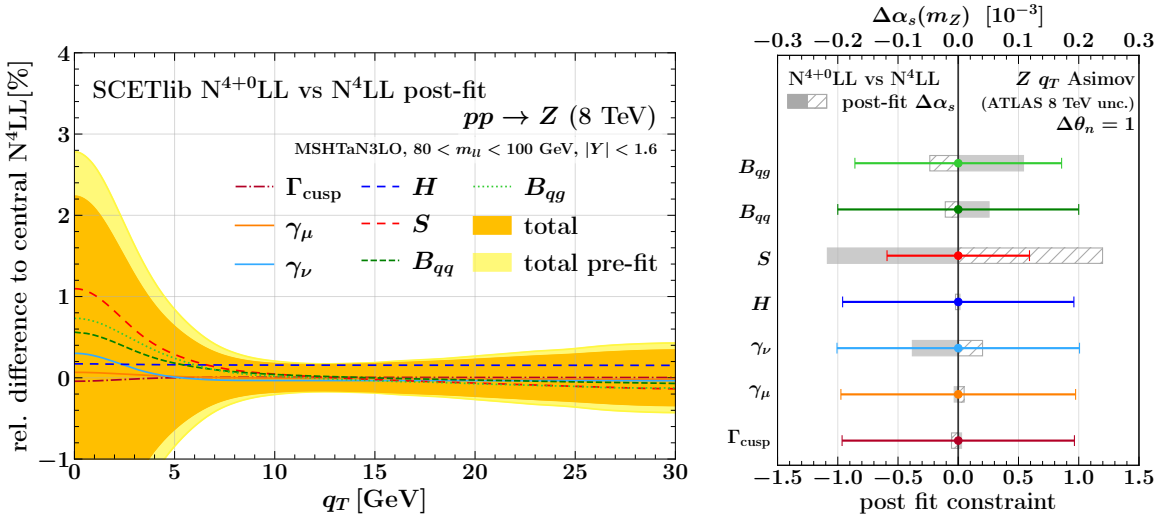


Figure 13. Left panel: Uncertainties in the q_T spectrum at N^{4+0}LL relative to the N^{4}LL, before (yellow band) and after (orange band) profiling the TNPs. The different lines show the post-fit relative impact of each TNP. Right panel: Post-fit constraints on the TNPs (error bars) and their impact on $\alpha_s(m_Z)$, with the solid (dashed) grey band showing the impact of the post-fit downward (upward) TNP variations.

JHEP12(2025)043

The results for the $N^{4+0}LL$ case are shown in figure 13. By construction, the central $N^{4+0}LL$ prediction is identical to that of N^4LL. The reduction in uncertainty across the q_T spectrum is less pronounced compared to the $N^{3+0}LL$ case, and similarly the constraints on the TNPs are no longer as tight as in the $N^{3+0}LL$ case.

Both $N^{3+0}LL$ and $N^{4+0}LL$ may serve as viable alternatives to their respective $N^{m+1}LL$ counterparts. While the difference in the resummation structure may only have a limited effect on the overall theory uncertainty, they can impact the theory correlations, which may not be fully captured. For this reason, whenever possible, it is preferable to use the full $N^{m+1}LL$ prescription.

Data Availability Statement. This article has no associated data or the data will not be deposited.

Code Availability Statement. This article has no associated code or the code will not be deposited.

References

[1] D. d'Enterria et al., *The strong coupling constant: state of the art and the decade ahead*, *J. Phys. G* **51** (2024) 090501 [arXiv:2203.08271] [INSPIRE].

[2] Particle Data Group collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001 [INSPIRE].

[3] Flavour Lattice Averaging Group (FLAG) collaboration, *FLAG Review 2024*, arXiv:2411.04268 [INSPIRE].

[4] S. Kluth, *Tests of Quantum Chromo Dynamics at e+ e- Colliders*, *Rept. Prog. Phys.* **69** (2006) 1771 [hep-ex/0603011] [INSPIRE].

[5] R. Abbate et al., *Thrust at N^3LL with Power Corrections and a Precision Global Fit for $\alpha_s(mZ)$*, *Phys. Rev. D* **83** (2011) 074021 [arXiv:1006.3080] [INSPIRE].

[6] R. Abbate et al., *Precision Thrust Cumulant Moments at N^3LL*, *Phys. Rev. D* **86** (2012) 094002 [arXiv:1204.5746] [INSPIRE].

[7] M.A. Benitez et al., *On Determining $\alpha_s(m_Z)$ from Dijets in e^+e^- Thrust*, *JHEP* **07** (2025) 249 [arXiv:2412.15164] [INSPIRE].

[8] T. Becher and M.D. Schwartz, *A precise determination of α_s from LEP thrust data using effective field theory*, *JHEP* **07** (2008) 034 [arXiv:0803.0342] [INSPIRE].

[9] A.H. Hoang, D.W. Kolodrubetz, V. Mateu and I.W. Stewart, *Precise determination of α_s from the C-parameter distribution*, *Phys. Rev. D* **91** (2015) 094018 [arXiv:1501.04111] [INSPIRE].

[10] F. Caola et al., *Linear power corrections to e^+e^- shape variables in the three-jet region*, *JHEP* **12** (2022) 062 [arXiv:2204.02247] [INSPIRE].

[11] P. Nason and G. Zanderighi, *Fits of α_s using power corrections in the three-jet region*, *JHEP* **06** (2023) 058 [arXiv:2301.03607] [INSPIRE].

JHEP12(2025)043

[12] P. Nason and G. Zanderighi, *Fits of α_s from event-shapes in the three-jet region: extension to all energies*, *JHEP* **06** (2025) 200 [arXiv:2501.18173] [INSPIRE].

[13] G. Bell et al., *Effects of renormalon scheme and perturbative scale choices on determinations of the strong coupling from e+e- event shapes*, *Phys. Rev. D* **109** (2024) 094008 [arXiv:2311.03990] [INSPIRE].

[14] U.G. Aglietti, G. Ferrera, W.-L. Ju and J. Miao, *Thrust distribution in electron-positron annihilation at full NNNLL+NNLO (and beyond) in QCD*, *Phys. Rev. Lett.* **134** (2025) 251904 [arXiv:2502.01570] [INSPIRE].

[15] M.A. Benitez et al., *A Precise Determination of α_s from the Heavy Jet Mass Distribution*, arXiv:2502.12253 [INSPIRE].

[16] ATLAS collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **09** (2014) 145 [arXiv:1406.3660] [INSPIRE].

[17] ATLAS collaboration, *Measurement of the transverse momentum and ϕ^*_η distributions of Drell-Yan lepton pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 291 [arXiv:1512.02192] [INSPIRE].

[18] ATLAS collaboration, *Measurement of the transverse momentum distribution of Drell-Yan lepton pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 616 [arXiv:1912.02844] [INSPIRE].

[19] ATLAS collaboration, *A precise measurement of the Z-boson double-differential transverse momentum and rapidity distributions in the full phase space of the decay leptons with the ATLAS experiment at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J. C* **84** (2024) 315 [arXiv:2309.09318] [INSPIRE].

[20] CMS collaboration, *Measurement of the Rapidity and Transverse Momentum Distributions of Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. D* **85** (2012) 032002 [arXiv:1110.4973] [INSPIRE].

[21] CMS collaboration, *Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **02** (2017) 096 [arXiv:1606.05864] [INSPIRE].

[22] CMS collaboration, *Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **12** (2019) 061 [arXiv:1909.04133] [INSPIRE].

[23] LHCb collaboration, *Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **01** (2016) 155 [arXiv:1511.08039] [INSPIRE].

[24] LHCb collaboration, *Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **09** (2016) 136 [arXiv:1607.06495] [INSPIRE].

[25] G. Billis et al., *Higgs p_T Spectrum and Total Cross Section with Fiducial Cuts at Third Resummed and Fixed Order in QCD*, *Phys. Rev. Lett.* **127** (2021) 072001 [arXiv:2102.08039] [INSPIRE].

[26] W.-L. Ju and M. Schönherr, *The q_T and $\Delta\phi$ spectra in W and Z production at the LHC at N^{3}LL'+N^{2}LO*, *JHEP* **10** (2021) 088 [arXiv:2106.11260] [INSPIRE].

[27] E. Re, L. Rottoli and P. Torrielli, *Fiducial Higgs and Drell-Yan distributions at N^{3}LL'+NNLO with RadISH*, *JHEP* **09** (2021) 108 [arXiv:2104.07509] [INSPIRE].

JHEP12(2025)043

[28] X. Chen et al., *Third-Order Fiducial Predictions for Drell-Yan Production at the LHC*, *Phys. Rev. Lett.* **128** (2022) 252001 [arXiv:2203.01565] [INSPIRE].

[29] T. Neumann and J. Campbell, *Fiducial Drell-Yan production at the LHC improved by transverse-momentum resummation at N4LLp+N3LO*, *Phys. Rev. D* **107** (2023) L011506 [arXiv:2207.07056] [INSPIRE].

[30] S. Camarda, L. Cieri and G. Ferrera, *Drell-Yan lepton-pair production: q_T resummation at N4LL accuracy*, *Phys. Lett. B* **845** (2023) 138125 [arXiv:2303.12781] [INSPIRE].

[31] V. Moos, I. Scimemi, A. Vladimirov and P. Zurita, *Extraction of unpolarized transverse momentum distributions from the fit of Drell-Yan data at N^4LL*, *JHEP* **05** (2024) 036 [arXiv:2305.07473] [INSPIRE].

[32] G. Billis, J.K.L. Michel and F.J. Tackmann, *Drell-Yan transverse-momentum spectra at N^3LL' and approximate N^4LL with SCETlib*, *JHEP* **02** (2025) 170 [arXiv:2411.16004] [INSPIRE].

[33] ATLAS collaboration, *A precise determination of the strong-coupling constant from the recoil of Z bosons with the ATLAS experiment at $\sqrt{s} = 8$ TeV*, arXiv:2309.12986 [INSPIRE].

[34] F.J. Tackmann, *Beyond scale variations: perturbative theory uncertainties from nuisance parameters*, *JHEP* **08** (2025) 098 [arXiv:2411.18606] [INSPIRE].

[35] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [arXiv:1912.10053] [INSPIRE].

[36] S. Forte and Z. Kassabov, *Why α_s cannot be determined from hadronic processes without simultaneously determining the parton distributions*, *Eur. Phys. J. C* **80** (2020) 182 [arXiv:2001.04986] [INSPIRE].

[37] NNPDF collaboration, *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022) 428 [arXiv:2109.02653] [INSPIRE].

[38] T. Cridge, L.A. Harland-Lang, A.D. Martin and R.S. Thorne, *An investigation of the α_S and heavy quark mass dependence in the MSHT20 global PDF analysis*, *Eur. Phys. J. C* **81** (2021) 744 [arXiv:2106.10289] [INSPIRE].

[39] T. Cridge, L.A. Harland-Lang and R.S. Thorne, *A first determination of the strong coupling α_S at approximate N^3LO order in a global PDF fit*, *Eur. Phys. J. C* **84** (2024) 1009 [arXiv:2404.02964] [INSPIRE].

[40] T. Cridge, G. Marinelli and F.J. Tackmann, *The Treatment of PDF Uncertainties in the Determination of Precise Parameters*, to appear.

[41] C.F. Berger et al., *Higgs Production with a Central Jet Veto at NNLL+NNLO*, *JHEP* **04** (2011) 092 [arXiv:1012.4480] [INSPIRE].

[42] I.W. Stewart and F.J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, *Phys. Rev. D* **85** (2012) 034011 [arXiv:1107.2117] [INSPIRE].

[43] A. Banfi, G.P. Salam and G. Zanderighi, *NLL+NNLO predictions for jet-veto efficiencies in Higgs-boson and Drell-Yan production*, *JHEP* **06** (2012) 159 [arXiv:1203.5773] [INSPIRE].

[44] S. Gangal and F.J. Tackmann, *Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion*, *Phys. Rev. D* **87** (2013) 093008 [arXiv:1302.5437] [INSPIRE].

[45] I.W. Stewart, F.J. Tackmann, J.R. Walsh and S. Zuberi, *Jet p_T resummation in Higgs production at NNLL'+NNLO*, *Phys. Rev. D* **89** (2014) 054001 [arXiv:1307.1808] [INSPIRE].

JHEP12(2025)043

[46] LHC Higgs Cross Section Working Group collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, *CERN Yellow Rep. Monogr.* **2** (2017) 1 [arXiv:1610.07922] [INSPIRE].

[47] J.R. Andersen et al., *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*, arXiv:1803.07977 [INSPIRE].

[48] J.M. Lindert et al., *Precise predictions for V+ jets dark matter backgrounds*, *Eur. Phys. J. C* **77** (2017) 829 [arXiv:1705.04664] [INSPIRE].

[49] L.A. Harland-Lang and R.S. Thorne, *On the Consistent Use of Scale Variations in PDF Fits and Predictions*, *Eur. Phys. J. C* **79** (2019) 225 [arXiv:1811.08434] [INSPIRE].

[50] NNPDF collaboration, *Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies*, *Eur. Phys. J. C* **79** (2019) 931 [arXiv:1906.10698] [INSPIRE].

[51] W. Bizon et al., *The transverse momentum spectrum of weak gauge bosons at N^3LL + NNLO*, *Eur. Phys. J. C* **79** (2019) 868 [arXiv:1905.05171] [INSPIRE].

[52] M.A. Ebert, J.K.L. Michel, I.W. Stewart and F.J. Tackmann, *Drell-Yan q_T resummation of fiducial power corrections at N^3LL*, *JHEP* **04** (2021) 102 [arXiv:2006.11382] [INSPIRE].

[53] M. Cacciari and N. Houdeau, *Meaningful characterisation of perturbative theoretical uncertainties*, *JHEP* **09** (2011) 039 [arXiv:1105.5152] [INSPIRE].

[54] E. Bagnaschi, M. Cacciari, A. Guffanti and L. Jenniches, *An extensive survey of the estimation of uncertainties from missing higher orders in perturbative calculations*, *JHEP* **02** (2015) 133 [arXiv:1409.5036] [INSPIRE].

[55] M. Bonvini, *Probabilistic definition of the perturbative theoretical uncertainty from missing higher orders*, *Eur. Phys. J. C* **80** (2020) 989 [arXiv:2006.16293] [INSPIRE].

[56] C. Duhr, A. Huss, A. Mazeliauskas and R. Szafron, *An analysis of Bayesian estimates for missing higher orders in perturbative calculations*, *JHEP* **09** (2021) 122 [arXiv:2106.04585] [INSPIRE].

[57] A. David and G. Passarino, *How well can we guess theoretical uncertainties?*, *Phys. Lett. B* **726** (2013) 266 [arXiv:1307.1843] [INSPIRE].

[58] A. Ghosh et al., *Statistical patterns of theory uncertainties*, *SciPost Phys. Core* **6** (2023) 045 [arXiv:2210.15167] [INSPIRE].

[59] J. McGowan, T. Cridge, L.A. Harland-Lang and R.S. Thorne, *Approximate N^3LO parton distribution functions with theoretical uncertainties: MSHT20aN^3LO PDFs*, *Eur. Phys. J. C* **83** (2023) 185 [*Erratum ibid.* **83** (2023) 302] [arXiv:2207.04739] [INSPIRE].

[60] B. Dehnadi, I. Novikov and F.J. Tackmann, *The photon energy spectrum in $B \to X_s\gamma$ at N^3LL'*, *JHEP* **07** (2023) 214 [arXiv:2211.07663] [INSPIRE].

[61] P. Cal et al., *Jet veto resummation for STXS H+1-jet bins at aNNLL$'$+NNLO*, *JHEP* **03** (2025) 155 [arXiv:2408.13301] [INSPIRE].

[62] CMS collaboration, *High-precision measurement of the W boson mass with the CMS experiment at the LHC*, arXiv:2412.13872 [INSPIRE].

[63] M.A. Lim and R. Poncelet, *Robust estimates of theoretical uncertainties at fixed-order in perturbation theory*, arXiv:2412.14910 [INSPIRE].

[64] T. Clark, S. Gangal and J.R. Gaunt, *The Drell-Yan process at NNLL$'$+NNLO using rapidity dependent jet vetoes*, *JHEP* **08** (2025) 060 [arXiv:2504.06353] [INSPIRE].

JHEP12(2025)043

[65] P. Pietrulewicz, D. Samitz, A. Spiering and F.J. Tackmann, *Factorization and Resummation for Massive Quark Effects in Exclusive Drell-Yan*, *JHEP* **08** (2017) 114 [arXiv:1703.09702] [INSPIRE].

[66] B. Dehnadi, P. Ploessl and F.J. Tackmann, *Flavor thresholds and quark-mass effects in the Collins-Soper kernel*, (2025).

[67] L. Cieri, G. Ferrera and G.F.R. Sborlini, *Combining QED and QCD transverse-momentum resummation for Z boson production at hadron colliders*, *JHEP* **08** (2018) 165 [arXiv:1805.11948] [INSPIRE].

[68] G. Billis, F.J. Tackmann and J. Talbert, *Higher-Order Sudakov Resummation in Coupled Gauge Theories*, *JHEP* **03** (2020) 182 [arXiv:1907.02971] [INSPIRE].

[69] A. Autieri, L. Cieri, G. Ferrera and G.F.R. Sborlini, *Combining QED and QCD transverse-momentum resummation for W and Z boson production at hadron colliders*, *JHEP* **07** (2023) 104 [arXiv:2302.05403] [INSPIRE].

[70] L. Buonocore, L. Rottoli and P. Torrielli, *Resummation of combined QCD-electroweak effects in Drell Yan lepton-pair production*, *JHEP* **07** (2024) 193 [arXiv:2404.15112] [INSPIRE].

[71] M.A. Ebert et al., *SCETlib: A C++ Package for Numerical Calculations in QCD and Soft-Collinear Effective Theory*, DESY-17-099 (2018).

[72] G. Billis, M.A. Ebert, J.K.L. Michel and F.J. Tackmann, *A toolbox for q_T and 0-jettiness subtractions at N^3LO*, *Eur. Phys. J. Plus* **136** (2021) 214 [arXiv:1909.00811] [INSPIRE].

[73] J.C. Collins and D.E. Soper, *Back-To-Back Jets in QCD*, *Nucl. Phys. B* **193** (1981) 381 [*Erratum ibid.* **213** (1983) 545] [INSPIRE].

[74] J.C. Collins and D.E. Soper, *Back-To-Back Jets: Fourier Transform from B to K-Transverse*, *Nucl. Phys. B* **197** (1982) 446 [INSPIRE].

[75] J.C. Collins, D.E. Soper and G.F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, *Nucl. Phys. B* **250** (1985) 199 [INSPIRE].

[76] J. Collins, *Foundations of Perturbative QCD*, Cambridge University Press (2011) [DOI:10.1017/9781009401845] [INSPIRE].

[77] C.W. Bauer, S. Fleming and M.E. Luke, *Summing Sudakov logarithms in $B \to X_s\gamma$ in effective field theory*, *Phys. Rev. D* **63** (2000) 014006 [hep-ph/0005275] [INSPIRE].

[78] C.W. Bauer, S. Fleming, D. Pirjol and I.W. Stewart, *An effective field theory for collinear and soft gluons: Heavy to light decays*, *Phys. Rev. D* **63** (2001) 114020 [hep-ph/0011336] [INSPIRE].

[79] C.W. Bauer, D. Pirjol and I.W. Stewart, *Soft collinear factorization in effective field theory*, *Phys. Rev. D* **65** (2002) 054022 [hep-ph/0109045] [INSPIRE].

[80] C.W. Bauer et al., *Hard scattering factorization from effective field theory*, *Phys. Rev. D* **66** (2002) 014017 [hep-ph/0202088] [INSPIRE].

[81] T. Becher and M. Neubert, *Drell-Yan Production at Small q_T, Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J. C* **71** (2011) 1665 [arXiv:1007.4005] [INSPIRE].

[82] M.G. Echevarria, A. Idilbi and I. Scimemi, *Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone*, *JHEP* **07** (2012) 002 [arXiv:1111.4996] [INSPIRE].

[83] J.-Y. Chiu, A. Jain, D. Neill and I.Z. Rothstein, *A Formalism for the Systematic Treatment of Rapidity Logarithms in Quantum Field Theory*, *JHEP* **05** (2012) 084 [arXiv:1202.0814] [INSPIRE].

JHEP12(2025)043

[84] Y. Li, D. Neill and H.X. Zhu, *An exponential regulator for rapidity divergences*, *Nucl. Phys. B* **960** (2020) 115193 [arXiv:1604.00392] [INSPIRE].

[85] G. Lustermans, J.K.L. Michel, F.J. Tackmann and W.J. Waalewijn, *Joint two-dimensional resummation in q_T and 0-jettiness at NNLL*, *JHEP* **03** (2019) 124 [arXiv:1901.03331] [INSPIRE].

[86] R. Boussarie et al., *TMD Handbook*, arXiv:2304.03302 [INSPIRE].

[87] M. Schlemmer et al., *Determination of the Collins-Soper Kernel from Lattice QCD*, *JHEP* **08** (2021) 004 [arXiv:2103.16991] [INSPIRE].

[88] Lattice Parton (LPC) collaboration, *Nonperturbative determination of the Collins-Soper kernel from quasitransverse-momentum-dependent wave functions*, *Phys. Rev. D* **106** (2022) 034509 [arXiv:2204.00200] [INSPIRE].

[89] A. Avkhadiev, P.E. Shanahan, M.L. Wagman and Y. Zhao, *Collins-Soper kernel from lattice QCD at the physical pion mass*, *Phys. Rev. D* **108** (2023) 114505 [arXiv:2307.12359] [INSPIRE].

[90] Lattice Parton (LPC) collaboration, *Lattice calculation of the intrinsic soft function and the Collins-Soper kernel*, *JHEP* **08** (2023) 172 [arXiv:2306.06488] [INSPIRE].

[91] H.-T. Shu et al., *Universality of the Collins-Soper kernel in lattice calculations*, *Phys. Rev. D* **108** (2023) 074519 [arXiv:2302.06502] [INSPIRE].

[92] A. Avkhadiev, P.E. Shanahan, M.L. Wagman and Y. Zhao, *Determination of the Collins-Soper Kernel from Lattice QCD*, *Phys. Rev. Lett.* **132** (2024) 231901 [arXiv:2402.06725] [INSPIRE].

[93] J.C. Collins and D.E. Soper, *Parton Distribution and Decay Functions*, *Nucl. Phys. B* **194** (1982) 445 [INSPIRE].

[94] J. Collins and T. Rogers, *Understanding the large-distance behavior of transverse-momentum-dependent parton densities and the Collins-Soper evolution kernel*, *Phys. Rev. D* **91** (2015) 074020 [arXiv:1412.3820] [INSPIRE].

[95] J. Collins et al., *Relating Transverse Momentum Dependent and Collinear Factorization Theorems in a Generalized Formalism*, *Phys. Rev. D* **94** (2016) 034014 [arXiv:1605.00671] [INSPIRE].

[96] A.A. Vladimirov, *Self-contained definition of the Collins-Soper kernel*, *Phys. Rev. Lett.* **125** (2020) 192002 [arXiv:2003.02288] [INSPIRE].

[97] M.A. Ebert, J.K.L. Michel, I.W. Stewart and Z. Sun, *Disentangling long and short distances in momentum-space TMDs*, *JHEP* **07** (2022) 129 [arXiv:2201.07237] [INSPIRE].

[98] S. Ferrario Ravasio, G. Limatola and P. Nason, *Infrared renormalons in kinematic distributions for hadron collider processes*, *JHEP* **06** (2021) 018 [arXiv:2011.14114] [INSPIRE].

[99] A. Bacchetta et al., *Effect of Flavor-Dependent Partonic Transverse Momentum on the Determination of the W Boson Mass in Hadronic Collisions*, *Phys. Lett. B* **788** (2019) 542 [arXiv:1807.02101] [INSPIRE].

[100] MAP (Multi-dimensional Analyses of Partonic distributions) collaboration, *Flavor dependence of unpolarized quark transverse momentum distributions from a global fit*, *JHEP* **08** (2024) 232 [arXiv:2405.13833] [INSPIRE].

[101] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343 [INSPIRE].

JHEP12(2025)043

[102] W.C. Davidon, *Variable metric method for minimization*, ANL-5990-Rev. 2 (1966) [INSPIRE].

[103] F. James and M. Winkler, *MINUIT User's Guide* [INSPIRE].

[104] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [*Erratum ibid.* **73** (2013) 2501] [arXiv:1007.1727] [INSPIRE].

[105] R.D. Cousins and L. Wasserman, *PHYSTAT Informal Review: Marginalizing versus Profiling of Nuisance Parameters*, arXiv:2404.17180 [INSPIRE].

[106] H. Paukkunen and P. Zurita, *PDF reweighting in the Hessian matrix approach*, *JHEP* **12** (2014) 100 [arXiv:1402.6623] [INSPIRE].

[107] S. Camarda et al., *QCD analysis of W- and Z-boson production at Tevatron*, arXiv:1503.05221.

[108] C. Willis et al., *New method for reducing parton distribution function uncertainties in the high-mass Drell-Yan spectrum*, *Phys. Rev. D* **99** (2019) 054004 [arXiv:1809.09481] [INSPIRE].

[109] S. Catani, D. de Florian and M. Grazzini, *Universality of nonleading logarithmic contributions in transverse momentum distributions*, *Nucl. Phys. B* **596** (2001) 299 [hep-ph/0008184] [INSPIRE].

[110] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*, *Nucl. Phys. B* **737** (2006) 73 [hep-ph/0508068] [INSPIRE].

[111] A. Cooper-Sarkar et al., *A Benchmarking of QCD Evolution at Approximate N^3LO*, arXiv:2406.16188 [INSPIRE].

[112] S. Camarda et al., *DYTurbo: Fast predictions for Drell-Yan processes*, *Eur. Phys. J. C* **80** (2020) 251 [*Erratum ibid.* **80** (2020) 440] [arXiv:1910.07049] [INSPIRE].

JHEP12(2025)043