Interpolation Grids for Inclusive Z-Boson Production at NNLO perturbative QCD

CERN Summer Student Programme 2022

Christiane Mayer
March 27, 2023

under the supervision of
Alexander Huss and Klaus Rabbertz

Abstract

This project demonstrates the use of interpolation grids for NNLO predictions in perturbative QCD of the inclusive Z-Boson production cross section at the LHC. An automated workflow combining the parton level Monte Carlo event generator, NNLOJET, and the grid interpolation library, fastNLO, is used for this purpose. The closure of the approximate grids is verified by comparing results to a full Monte Carlo event generation at parton level and an analysis of the scale uncertainty, statistical uncertainty, and PDF dependence of the prediction is performed. The uncertainty due to the interpolation bias as well as the statistical uncertainties have been reduced below the desired 0.1% accuracy dictated by the experimental precision. The scale and PDF uncertainties due to the underlying theoretical models have the same order of magnitude. The resulting theoretical predictions are compatible with experimental measurements by the CMS collaboration within these uncertainties. In addition, prototypes for further automatisation in the Monte Carlo event generation are discussed.

1 Introduction

Measurements of the cross sections of different processes in proton-proton collisions at the LHC encode a lot of information about fundamental parameters of the standard model as well as the internal structure of the protons themselves. In order to extract this information, theoretical predictions need to be available with parameters that can be fitted to the experimental data. This report makes next-to-next-to leading order (NNLO) predictions in perturbative QCD for inclusive Z-boson production for a phenomenological study using data by the CMS experiment [1]. For a sensible comparison of the experimental and theoretical results, it is important that they both have roughly the same level of accuracy. With current experimental measurements having uncertainties around 1%, theoretical predictions need to reduce their uncertainties to the same level or better. This requires terms in the perturbative expansion up to higher orders which however increases the computation cost significantly. Therefore it is important to find approximations to enable faster repeated evaluations of predictions with different parameters to see how they will affect the overall result.
One such method proceeds through interpolation grids for a given process and observable that only requires a single in-depth calculation and can then be re-evaluated for different energy scales and parton distribution functions. This evaluation is much faster than the full Monte Carlo event generation with high numbers of events. For any such approximations it is of vital importance to perform closure tests to ensure that they provide sufficiently accurate results. However once they are verified to be a valid approximation, they can be used to quantify uncertainties in the theoretical predictions that would otherwise be intractable without running multiple full calculations.

2 Probing the Proton

At low energies and large distances the proton can be considered a point particle for most purposes. In order to probe the internal structure of the proton, high enough energies are needed to resolve the quarks and gluons inside the proton. At the LHC proton-proton collisions occur at energies up to 13.6 TeV, which is more than enough to enable reactions among the individual constituent partons in the protons.
The probability of a particle collision leading to any final state of particles is quantified by the process cross section. These cross sections can be measured by recording the number of events leading to the final state over a large amount of collisions. It is often useful to not only look at the total cross section for a process but at the differential cross section across an observable such as the transverse momentum p_T perpendicular to the beams or the rapidity of the particles since that can provide more insights into the momentum fraction each of the partons carries within the proton.
For theoretical predicitons, the interactions of partons are governed by the strong interaction described by QCD with the strong coupling constant α_S. Perturbative methods in QCD can only be used at high energies where the strong coupling constant becomes small, but at low energies, which are relevant for the description of the internal structure of protons, α_S becomes large and is thus non perturbative. In order to make predictions in perturbative QCD for high energy collisions, we need to provide some experimentally determined input parameters and make approximations for the parts of the calculation that are non-perturbative.
To this end, an energy scale μ_F is introduced as the cut-off energy for which the internal structure of the proton can no longer be described perturbatively. Below this energy, pre-determined parton distribution functions (PDFs) are supplied to the calculation to model the proton structure. There are multiple different sets of PDFs that each use different theoretical models and experimental input and they are continuously improved upon as more data and predictions with higher precision become available. These parton distribution functions provide a probability distribution for each of the partons (valence quarks, gluons, sea quarks) over the fraction x of the total proton momentum they carry. A second energy scale appearing in the calculation is the one at which the strong coupling constant α_S is evaluated. This energy scale needs to correspond to the relevant energy scale of the reaction which is characteristic for each process.
In a calculation of the full perturbative series to infinity, the final result would not depend of the chosen energy scales μ_R and μ_F. However, in practice we are limited in computational power, so only the first few orders in the perturbative series can be evaluated, which will introduce an uncertainty in the final result, since higher terms are missing. We cannot know what contribution the next order in the perturbative series would bring without calculating it, but we can use the scale dependence of the orders we have already computed to assign an uncertainty to the result, without going to the next order.

3 Inclusive Z-Boson Production

The aim of this project is to reproduce the measurements by the CMS collaboration for the differential cross section over the absolute Z-boson rapidity y^Z for inclusive Z-boson production as shown in Figure 5 of Ref. [1]. Rapidity is closely related to the scattering angle θ and defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \tag{1}$$

The relevant measurements were taken during LHC run 2 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and correspond to an integrated luminosity of 35.9 fb^{-1}. Z-bosons can be produced in proton-proton collisions via the Drell-Yan process, where quarks from the two protons interact as shown in Figure 1. The Z-boson itself subsequently decays into a lepton pair with opposite charge that can be detected. The combined invariant mass of the two leptons will be equal to the invariant mass of the Z-boson and is therefore a good parameter for the relevant energy scales μ_R and μ_F for the calculation. The same final result of a lepton pair can also be produced by an exchanged photon instead of the Z-boson, so both cases are indistinguishable in the experiment and therefore also need to be taken into account in the theoretical predictions. However, by selecting only events with combined invariant mass of both leptons in a 15 GeV window around the Z-boson mass of 91.1876 GeV, it is possible to focus mainly on the contribution of Z-bosons to the overall process. Furthermore, only events are selected with leptons that have pseudo-rapidities $|\eta| < 2.4$ and transverse momenta $p_T > 25$ GeV. The differential cross section is recorded in 12 equally spaced observable bins in $|y^Z|$ ranging from 0 to 2.4. All these selections and binning need to be matched by the theoretical prediction.
The experimental study is for inclusive Z-boson production which means that events with or without jets are recorded as long as they have a lepton pair matching the selection criteria. The emission of jets is impossible at tree level, shown in Figure 1, which is the leading order contribution for this process. Jets become relevant only at NLO and higher where real emissions of gluons can lead to their formation.

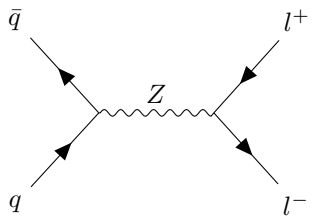


Figure 1: Leading order Drell-Yan process including a Z-boson

4 The Computation

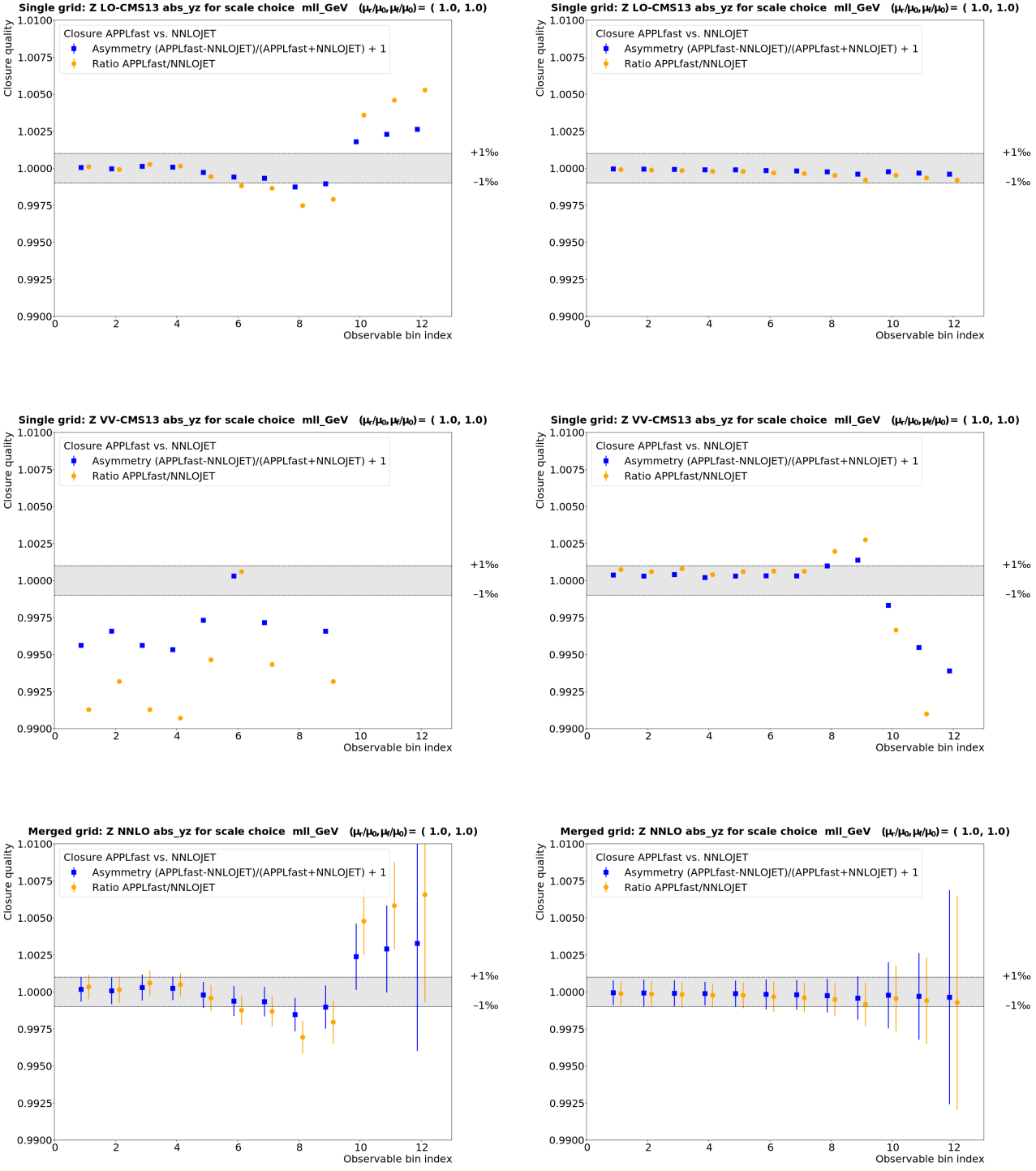
Each higher order in the perturbation series for the cross section becomes more and more challenging and thus more computational resources are needed. Therefore a primary focus when setting up the computation is to find the balance between accuracy of the result and the speed at which the computations can be performed. The simulation in this report uses a parton level Monte Carlo event generator, NNLOJET [2], to calculate total cross sections in each observable bin for a specific choice of scales and PDF sets. Instead of rerunning the full calculation for different parameters to find the uncertainties of this prediction, an interpolation grid is generated using fastNLO [3] that separates the contribution of the scales and the PDF sets from the computationally tedious part of the calculation. These coefficient grids only need to be re-weighted for different scales or PDF sets, which is much faster than re-doing the entire calculation. Both of these programs run in multiple steps to make computations more efficient and to check the quality of the approximations before the full calculation with high statistical precision is started. An automated workflow for all these steps is provided in the nnlo-law-analysis package, which is described in detail in [4].
All calculations are split up into the perturbative orders (LO, NLO, NNLO) and the higher orders are further split into different channels. The next-to-leading order is split into two channels: processes that include a virtual loop (V channel) and processes with a real emission (R channel) and at next-to-next-to-leading order, there are three channels for this specific process that feature two of these phenomena (RR, RV, VV channels). Each of these channels can be calculated separately and they only need to be added up to calculate the final cross section. While the total cross section of any physical process has to be a positive number, the individual contributions of each of these channels can be negative as long as the sum over all the contributions gives a positive number.
The precision of Monte Carlo methods increases with the number of samples, but using a uniform sample distribution over a large interval is almost always not the best strategy when there are distinct structures in the integrand. Before the event generation with large event numbers is started, it is beneficial to do faster

warmup runs with less events to optimise the phase space sampling. The first step is a Vegas integration to find the optimal sampling distribution over all the dimensions of phase space. The higher orders in the calculation have more dimensions of phase space and while the Vegas integration in leading order converges after only a few iterations, higher order channels will sometimes find false peaks and will in general converge much slower. Therefore, plots of the phase space distribution after each iteration are provided to allow for a check of their quality and, if found to be insufficient, it is possible to rerun the warmup for specific channels building up on the previous results.
Another warmup run is used to initiate the tables for the grids by finding the minimum and maximum values of x and Q^2 for the process. Since the size of the table is set at the start, it is important to make sure it only spans over the accessed phase space to avoid having empty entries at the edges of the table, since that would only decrease the resolution of the table while increasing its size.
After the warmup runs have determined optimised limits, the full event generation can be run, which will provide a large number of predictions for the cross-section of each channel in each bin together with corresponding interpolations grids. Subsequently, the NNLOJET predictions and interpolation grids for each bin and channel are statistically combined using suitable weights to suppress inaccurate results or outliers. Finally, the automated setup adds up the partial results and generates a number of plots that allow for an intuitive visual check of the quality of the grids as well as the uncertainties of the overall cross section.

5 Grid Closure

Before the grids can be used to compare data and theory or make predictions, it must be ensured that their approximate results match with the full Monte Carlo event generation. For this we can calculate the asymmetry and the ratio between the cross sections calculated via the grids with the results as obtained by the initial NNLOJET calculation. It is important to check not only the final merged grid for closure but also each individual channel to make sure they are sensible as well. In order for the grids to be used for comparison with data, an interpolation bias of less than 0.1% and a numerical statistical uncertainty of at most 1% is essential.
In Figure 2 the interpolation quality is shown for two distributions of grid support nodes in x, either equidistant in $-\sqrt{-\log x}$ on the left or equidistant in $\log x$ on the right.[1] For a full calculation it is mandatory to check all closure plots to ensure a good interpolation quality. Here only the examples for a single grid of the LO and VV channels and for the fully merged grid at NNLO are presented. Even in the leading order the single grid shows significant systematic deviations in the cross section the higher transverse momentum bins in the $-\sqrt{-\log x}$ sampling while the $\log x$ remains within the desired closure of $\pm 0.1\%$ as shown by the blue band. The $\log x$ sampled grid shows deviations for a VV-channel single grid as well, while the results for the $-\sqrt{-\log x}$ sampling in the VV-channel are nowhere near good enough.
The merged grid that uses all channels and weights each observable bin of each table according to the statistical quality is within the uncertainty band for the $\log x$ sampling even with the discrepancies in NNLO single grids.
For for the $\sqrt{\log_{10}}$ sampling, however, the merged grid shows systematic deviations outside the desired accuracy for high y^Z. Since both grids use the exact same settings otherwise, it is safe to draw the conclusion that for the specific process of inclusive Z-boson production an x distance measure of $\log x$ as opposed to $-\sqrt{-\log x}$ leads to a significant improvement of grid quality.
The accuracy of these approximations can always be improved by using more support nodes to have a tighter net of grid points. However, this will increase the size of the grids substantially which consequently increases the amount of storage space as well as the runtime for all calculations done with the grids after. Therefore it is important to find a balance between the quality of the grids and their size: once the closure of the grids is met to the desired accuracy, there is no further benefit in making the tables more accurate since the remaining uncertainties from the theoretical model will be larger than any of the approximation uncertainty due to the grids as we will see in section 6. To find the right balance for this process, multiple grids with x distance measure $\log x$ but different number of support points per magnitude were generated to compare their closure with NNLOJET results in Figure 3. Three sets of grids were produced with 8, 12 and 16 bins per magnitude, respectively, and the plots here show the single grid closure for the RR and VV channels which usually exhibit the worst closure. The first single grid for the RR channel with 8 bins per magnitude shows only a slight deviation in some points for higher y^Z, while the single VV grid shows systematic

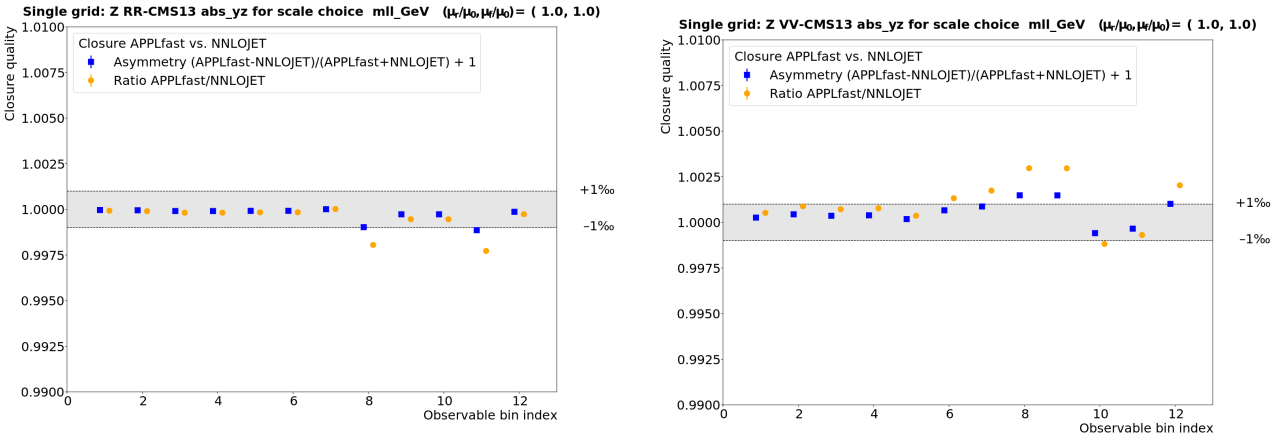
[1] An equidistant distribution in x cannot lead to reasonable results in the context of PDF interpolations.



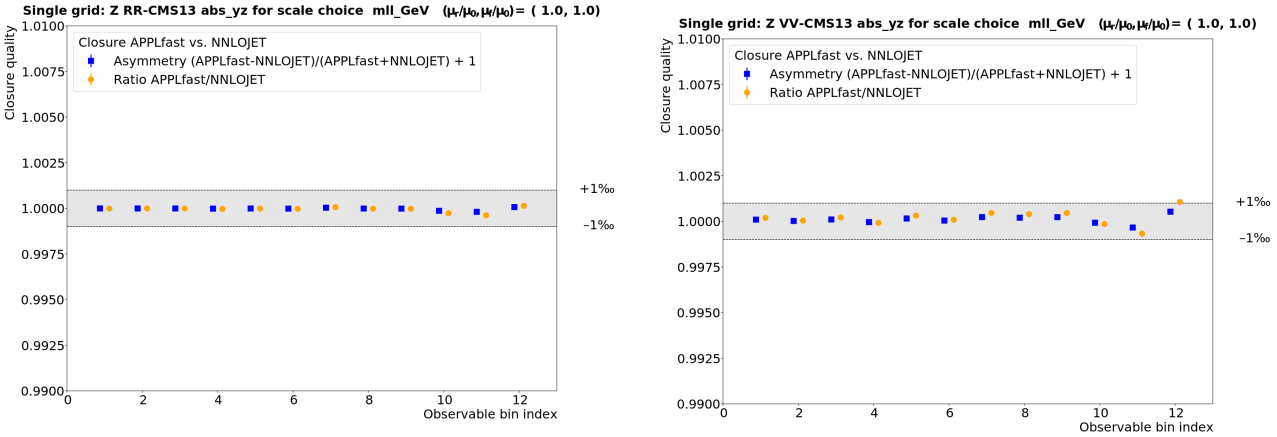
(a) x Distance Measure $-\sqrt{-\log x}$

(b) x Distance Measure $\log x$

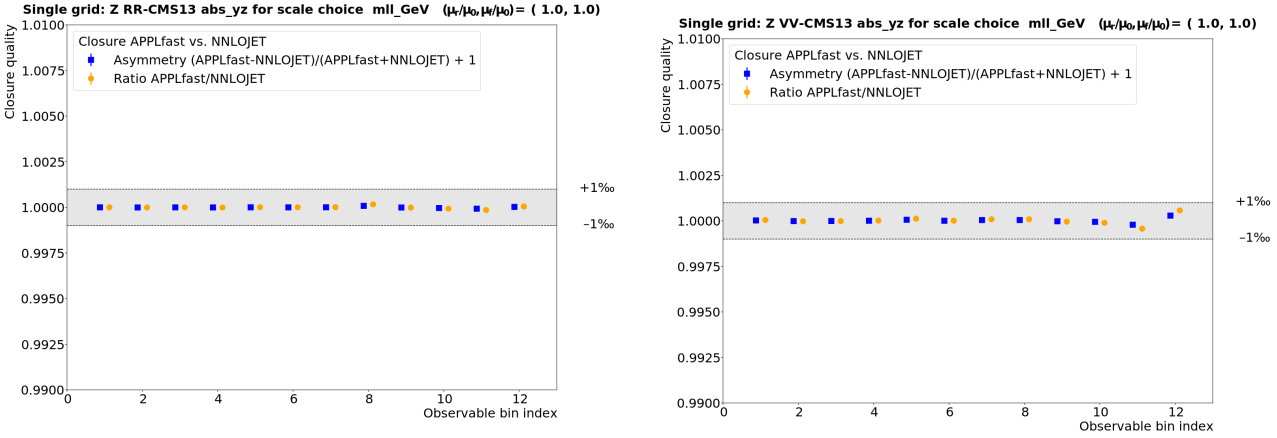
Figure 2: Comparison of the single grid closure for LO (top row) and VV (middle row) and of the NNLO merged grid closure (bottom row) for the two different x distance measures $-\sqrt{-\log x}$ (left) and $\log x$ (right)



(a) 8 bins per magnitude



(b) 12 bins per magnitude



(c) 16 bins per magnitude

Figure 3: Comparison of the grid closure of 8 (top row), 12 (middle row), and 16 (bottom row) support nodes per magnitude in x. Single grid RR closure is shown to the left and the single grid VV closure to the right.

support nodes per magnitude	table size [MB]
8	50
12	100
16	180

Table 1: Increasing Table Size

deviations outside of the desired uncertainty band.
While increasing the number of bins per magnitude to 12 does not get rid of the systematic deviations entirely, it does reduce them to within the desired $\pm 0.1\%$ band. Further increasing the number of bins per magnitude reduces the few remaining deviating points even more. However, as shown in Table 1, this slight increase in grid quality comes at an 80% increase in the size of the table. For the purpose of this report this further increase in quality is unnecessary and would only cost more storage space and computation time. Overall, the set of grids with 12 bins per magnitude are able to match the calculations from NNLOJET within 0.1% and can therefore safely be used for calculations requiring accuracy at this level.

6 Computational Uncertainties

Once the closure of the grids to the desired accuracy is fulfilled, the uncertainties of the theoretical prediction can be estimated.
Being a perturbative calculation there is a fundamental source of uncertainty associated with the truncation of the series at finite order. One way of estimating this uncertainty without calculating the next order in the perturbation is to look at the uncertainty due to the scales. The relevant energy scale for μ_R and μ_F was set to the invariant mass of the two final state leptons. The standard way of estimating the scale uncertainty of the final result is to use a 6 point method of varying both scales by factors of 2. Figure 4 shows the scale uncertainty bands for each of the perturbation orders.
It can be seen that, while the NLO correction is substantial for $|y^Z| > 1.3$, the NNLO correction is small everywhere. More importantly, the scale uncertainty as a proxy for missing higher orders successively decreases from LO to NLO to NNLO.
A more detailed split into how much each channel contributes to the total cross section in each rapidity bin can be seen in Figure 5. At the lower rapidity regions, the leading order is responsible for most of the total cross section since the real and virtual contributions mostly cancel each other in the higher order corrections. While the virtual contributions stay mostly constant over all the rapidity bins, the real contributions become more significant in the higher rapidity region, leading to an important contribution to the total cross section as can even be seen in Figure 4. At leading order, two quarks annihilate to create a Z-boson that will therefore have momentum strictly in the beam direction. This corresponds to a very restrictive kinematic configuration that is strongly correlated to high-x PDFs at large rapidities $|y^Z|$. This restriction is lifted at higher orders through the real emissions of additional particles which thus exhibits sizeable effects at high $|y^Z|$.
Since only a finite number of events were generated and used for the calculation of the final cross section, statistical uncertainties will always remain. Figure 6 shows the percentual uncertainty due the limited statistical precision for each of the channels. The double real channel contributes most to the total uncertainty with the real-virtual coming second. It would be possible to further increase the number of events in those channels to decrease their statistical uncertainty, but since the overall uncertainty is still below 1%, the current result is considered to be good enough for the purpose at hand.
The final source of uncertainty to this cross section comes from the uncertainty regarding our understanding of the internal structure of the proton. Figure 7 shows the effect of using different PDF sets on the same interpolation grid along with their respective uncertainties. Each of the PDF sets has a built in method for calculating their respective uncertainties such as having multiple PDF sets with slightly changing parameters within their confidence interval that can be used in the calculation to see how those changes would affect the final cross section with respect to the default CT14 set. All three PDF sets show the same trend in the total cross section, but especially in the lower rapidity region they predict a slightly larger cross section. Such deviations could be used to draw conclusions about which PDF sets are more consistent with the data or how others could be further constrained.

7 Automated Quality Checks

While the workflow used in this project already mostly automated the process of creating the interpolation grids, it still required quality checks of the interpolation grids by looking at plots and adjusting parameters by hand if the desired accuracy is not met. Not only does this increase the time it takes to create grids, but it also needs the user to have sufficient experience

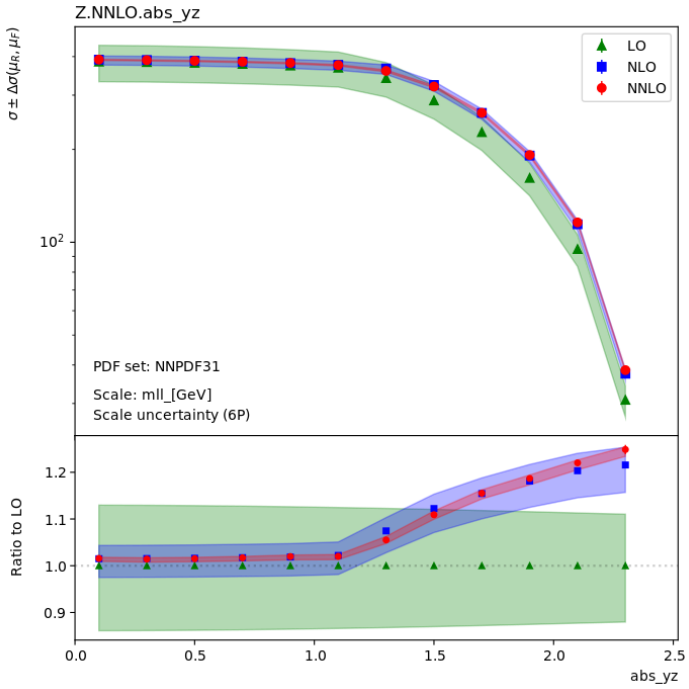


Figure 4: Scale uncertainties for all orders of the calculation.

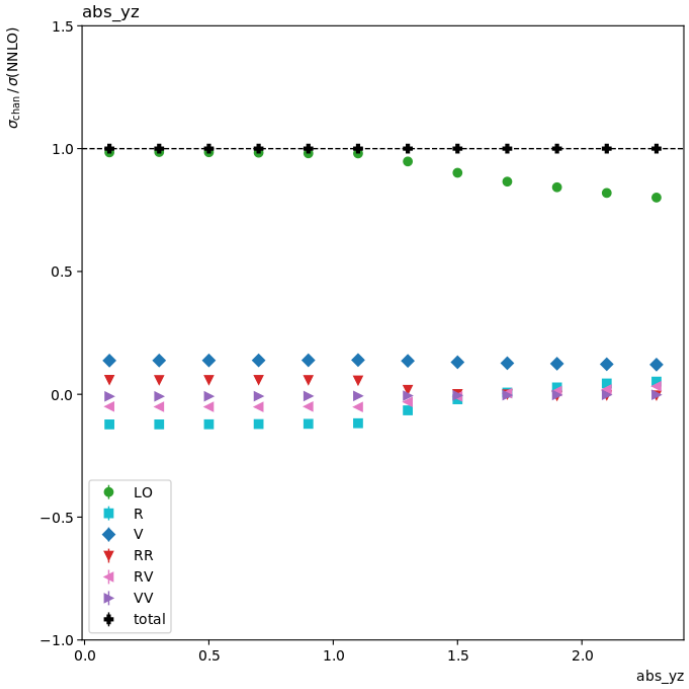


Figure 5: Contribution of each channel to the total cross section at NNLO.

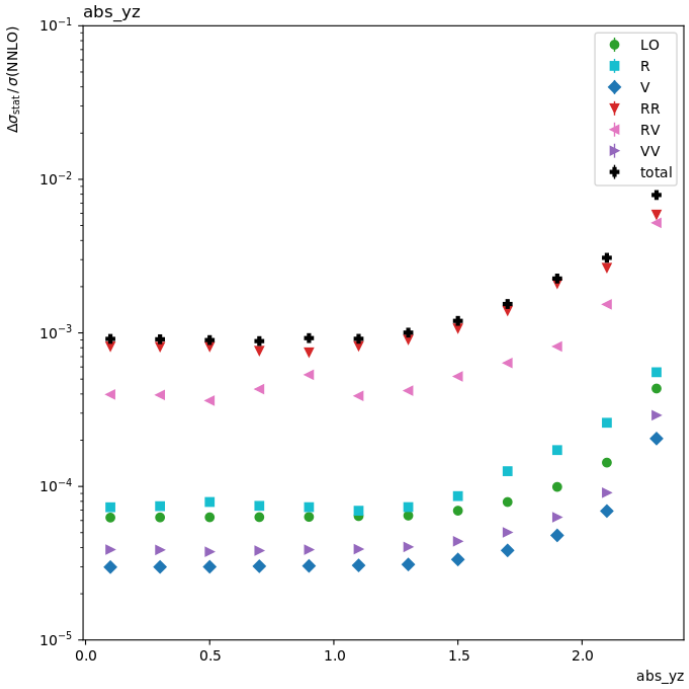


Figure 6: Percentage of statistical uncertainties contributed by each channel to the NNLO result.

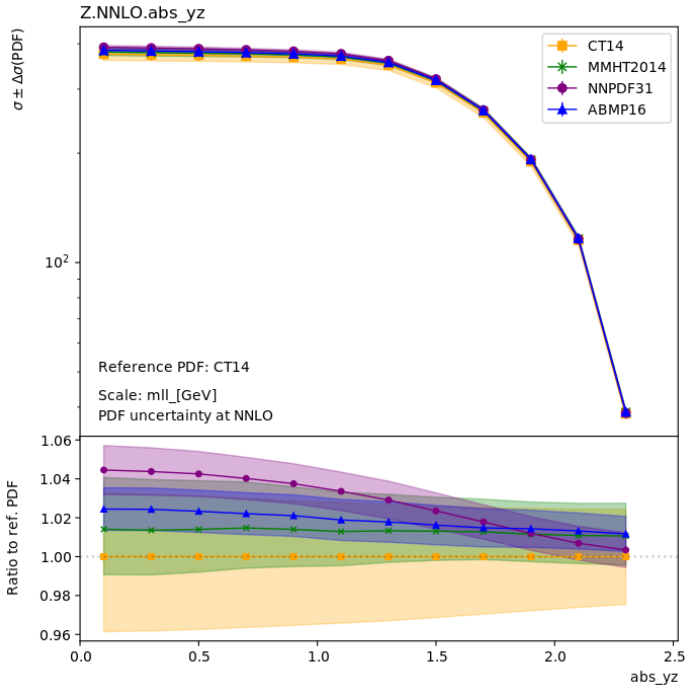


Figure 7: Total cross section at NNLO evaluated for the three PDF sets of MMHT2014, NNPDF31, ABMP16 in comparison to CT14

in making judgements based on the plots about what parameter adjustments are sensible.
As a side project I have worked on prototypes for quality checks for NNLOJET that will automatically adapt the parameters to reach the desired quality, but not waste computation time and storage space beyond that. The framework for this workflow was provided by the Luigi python package [5] that allows for multiple tasks to be defined with complex dependencies among each other.
The first thing I worked on was an automated check of the quality of the Vegas warmup. The number of iterations for optimising the phase space sampling and the number of events for each iteration is given in a runcard where all of the process parameters and observables are defined as well. After these iterations finish, plots are produced to check if the sampling in each dimension converges and if not, the runcard has to be edited to increase the number of events or iterations until the sampling is at an adequate level. The individual warmup runs are not too computationally challenging, but the manual quality checks require human intervention and judgement which ideally should be avoided.
I wrote a prototype for an automated quality check that reads in the bin edges in phase space for each iteration and calculates their differences from each iteration to the next. Summing over the absolute values of each of these differences allows for one number to be assigned to each iteration as a measure of similarity to the previous one, i.e. how far the phase space sampling has converged with each iteration. It is then possible to set a threshold for this parameter after which the warmup is deemed good enough to be used for the proper calculation with higher numbers of events. If the threshold is not met after the initially assigned number of iterations, a re-warmup is performed that reuses the existing results, but does more iterations with a higher number of events. This process will be repeated for each of the channels until either the threshold is met or a preset number of reruns is performed to avoid an infinite loop. Both, the threshold and the maximum number of reruns need to be set manually in advance and especially finding a sensible value for the threshold is far from a trivial task, since it will be different for each process, channel, and number of bins etc.
The NNLOJET log files provide more parameters that can be used to determine whether the warmup is good enough: the total integrated cross section with its standard deviation is given as well as a χ^2 value for each iteration. More quality checks can be defined this way that are satisfied if the ratio of the standard deviation to the total integral is smaller than a given threshold and the χ^2 value is smaller than yet another threshold.
The last optimisation that can be made automatically is to allocate the number of events for each channel that leads to the lowest statistical uncertainties in the final combined cross section. From the log files we can find a percentage error on the total cross sections from each channel as described above but also the time it took the program to generate the specified number of events. Since the final cross section will be a sum over all the individual channels C, the errors are propagated as:

$$\delta\sigma_{tot} = \sqrt{\sum_C \delta\sigma_C^2}.$$

We can also assume that the errors scale as statistical uncertainties with respect to the number of events N_C in each channel as:

$$\delta\sigma_C \propto \frac{1}{\sqrt{N_C}}.$$

If a testrun with N_C as the initial number of events in channel C is performed to find the uncertainty and the average calculation time T_C per event in any given channel, then it is a simple optimisation problem to find the best numbers of additional events $\overline{N_C}$ for each channel to reach the lowest possible statistical uncertainty for a given maximum computation time T_{tot}.:

$$\overline{N_C} = \sqrt{\frac{\delta_C^2 N_C}{T_C}} \frac{(T_{tot} + \sum_i T_i N_i)}{\sum_i \sqrt{\delta_i^2 T_i N_i}} - N_C. \qquad (2)$$

Implementing this into the workflow reduces the need for a manual check of the log files and avoids use of unoptimised preset event numbers. Alternatively, it would be possible to set a target accuracy for statistical uncertainty and a maximum runtime to reduce unnecessary high runtimes when the target accuracy is already met.
The quality checks and optimisation steps discussed in this section, have been developed as a prototype Luigi workflow, but have not yet been generalised for more processes than the inclusive Z-boson production and need more careful testing before they can be used with confidence generically.

8 Conclusion

This report demonstrates how theoretical predictions of inclusive Z-boson production cross sections for the LHC at NNLO accuracy in perturbative QCD can be

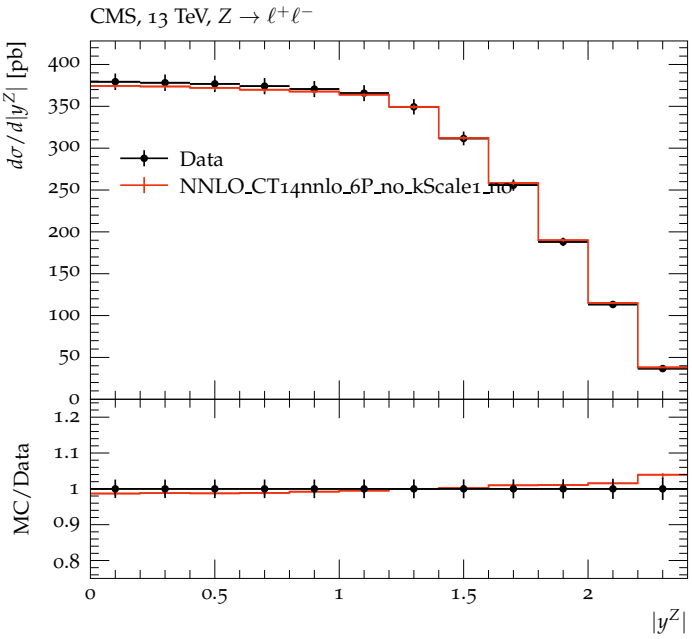


Figure 8: Comparison of theoretical predictions at NNLO using the CT14 PDF set to the experimental data of CMS

made. The comparison of the theoretical results to the experimental measurement by the CMS collaboration is shown in Figure 8. They agree very well with each other with only some small systematic trends of the experimental cross section being larger than the theoretical prediction in the lower rapidity region while still remaining well within the uncertainty bands. Other PDF sets even predict a slightly larger cross section in this region than the CT14 PDF set.

Approximate methods such as the interpolation grids used in this report are an indispensable tool for the repeated calculations of cross sections at NNLO accuracy that are needed to quantify the uncertainty on the theory predictions. It is demonstrated here that the interpolation grids are a good approximation at the current level of accuracy needed. With the option of repeating the cross section calculation for different PDF sets with much less computation cost, it is possible to use these interpolation grids to further constrain the PDF sets in the future.

References

[1] Albert M Sirunyan et al. "Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 12 (2019), p. 061. DOI: 10.1007/JHEP12(2019)061. arXiv: 1909.04133 [hep-ex].

[2] T. Gehrmann et al. *Jet cross sections and transverse momentum distributions with NNLOJET*. 2018. DOI: 10.48550/ARXIV.1801.06415. URL: https://arxiv.org/abs/1801.06415.

[3] T. KLUGE, K. RABBERTZ, and M. WOBISCH. "FAST pQCD CALCULATIONS FOR PDF FITS". In: *Deep Inelastic Scattering DIS 2006*. WORLD SCIENTIFIC, Jan. 2007. DOI: 10.1142/9789812706706_0110. URL: https://doi.org/10.1142%5C%2F9789812706706_0110.

[4] Miguel Santos Correa. "Automated Production of Interpolation Grids at NNLO for the Triple-Differential Z+Jet Cross Section Measurement at the LHC". MA thesis. Karlsruhe Institute of Technology (KIT), 2018.

[5] *Luigi Package*. URL: https://luigi.readthedocs.io/en/stable/.