Ludwig-Maximilians-Universität
München

DESY Hamburg

Bachelor's Thesis

Precision study at NLO QCD for semileptonic Di-boson processes at future Linear Colliders

Marcel Niedermeier

supervised by

Dr.Jürgen Reuter

Prof. Dr. Gerhard Buchalla

December 18, 2018

Ludwig-Maximilians-Universität
München

DESY Hamburg

Bachelorarbeit

Eine Präzisionsstudie von NLO-QCD-Korrekturen zu semileptonischen Diboson-Prozessen an künftigen Linearbeschleunigern

Marcel Niedermeier

betreut durch

Dr.Jürgen Reuter

Prof. Dr. Gerhard Buchalla

December 18, 2018

Abstract

In this Bachelor's thesis a precision study of semi-leptonic di-boson processes at the future International Linear Collider (ILC) is presented. Results are given up to Next-to-leading order (NLO) in perturbative Quantum Chromodynamics (QCD), where the Monte Carlo (MC) event generator WHIZARD, which is briefly presented, with an interface to the library OpenLoops is used. The discussion includes a scan of the integrated cross section over a significant range of centre-of-mass energies as well as an analysis of selected differential observables.

Zusammenfassung

Die vorliegende Bachelorarbeit stellt eine Präzisionsstudie von semileptonischen Diboson-Prozessen am künftigen International Linear Collider (ILC) vor. Dafür wird der Monte-Carlo-Eventgenerators WHIZARD zusammen mit dem Verzeichnis OpenLoops verwendet; die Resultate werden bis zur ersten Ordnung in perturbativer Quantenchromodynamik (NLO-QCD) entwickelt. Diskutiert wird ein Scan des integrierten Wirkungsquerschnitts über ein repräsentatives Intervall variierender Schwerpunktsenergien sowie eine Analyse ausgewählter differentieller Observablen.

Contents

List of Figures

1 Introduction

A future linear accelerator and collider like the ILC [1] provides the cleanest possible environment to investigate lepton collisions and productions: there is no QCD initial state radiation, no parton distribution functions are needed and no hadronic remnants appear as a possible background in the detector. Therefore, given these excellent experimental conditions, in order to validate the Standard Model of particle physics (SM), experimentalists have to be provided with precise theoretical simulations of a given interaction, which can then subsequently be compared to the measured data. In the setup presented here, precision studies of top-quark production have already been published and can for example be found in references [2, 3]. Here, we provide a similar analysis for the two processes $e^+e^- \to \mu\mu jj$ and $e^+e^- \to \mu\nu jj$, which constitute the production of pairs of W and Z bosons with their semileptonic decays including their irreducible Standard Model (SM) backgrounds. In each of these cases, we will account for the radiation of two associated light-quark jets.

To our knowledge, none of these processes has been investigated up to NLO contributions which arise when the real and virtual QCD corrections to the tree-level calculation are taken into account as well. Since these calculations considerably rely on the savvy use of numerical integration tools, we will make use of the MC event generator WHIZARD [4, 5] in order to perform our calculations. An interface to the library Openloops [6] will permit us to use its pre-generated NLO matrix elements, which considerably reduces the computational effort [1].

The massive use of computational tools necessitates a careful presentation of the consecutive implementation steps of a given process up to NLO-QCD[2]. Here it will be as follows:

In the first section we give an overview of the phenomenology and the physics of the two processes in question in order to get a first glimpse at the results that could be expected. To do so, we will consider the tree-level Feynman diagrams and a representative sample of corrections that arise at NLO, which will lead us to important differences between the charged and the neutral process. Based on this, we will also conclude on some of the difficulties that result from a calculation at NLO, and crucially differentiates it from a seemingly similar calculation at leading

[1] The main reason for this is that WHIZARD's internal matrix element generator cannot calculate the necessary virtual matrix elements, hence the need for an external package.

[2] The usage of WHIZARD at NLO-QCD is in some aspects quite different from a usage at LO and has several consequences which will be detailed later.

order (LO). We will complete this section by an interlude which justifies the use and the advantages of a computational framework to evaluate our results.
The second section presents the MC event generator WHIZARD, gives an overview of its structure and explains which external packages have been included. Several important asides will summarise the principles of event generation, subtraction schemes and jet clustering in order to familiarise the reader with the working mechanisms of an NLO Monte Carlo event generator like WHIZARD.
Consequently, the central part of this thesis is the computational validation of our expectations. After explaining how the calculation has been set up, the computational results are organised in two main parts. Firstly, we will present the integrated cross sections over a range of centre-of-mass (CM) energies from 100 GeV to 3000 GeV, where we compare scans at LO and at NLO. Secondly, we analyse the cross section differential with respect to the most interesting observables to reveal important differences and similarities between the two processes, and to study the effects of the NLO QCD corrections.
Finally, in a brief conclusion we summarise the results that have been obtained in the context of the currently realisable computational precision and thus give a prospect on what could be refined in a future, even deeper precision study.

2 Physical background and Phenomenology

A brief phenomenological review will help us to have a sound understanding of the numerical results presented in section four.
The two processes that are investigated are

$$e^+e^- \to \mu\nu jj \tag{1}$$

and

$$e^+e^- \to \mu\mu jj, \tag{2}$$

where j is the standard notation for a jet. On a semantic level, j as a shorthand notation is a particle container and represents the partons that will form the two jets, in the present case we will allow the two light quark families and their respective antiquarks:

$$j = u, d, c, s \tag{3}$$

and

$$j = \overline{u}, \overline{d}, \overline{c}, \overline{s}. \tag{4}$$

Both processes constitute the pair production of two heavy electroweak gauge bosons:

$$e^+e^- \to W^+W^- \tag{5}$$

and

$$e^+e^- \to ZZ, \tag{6}$$

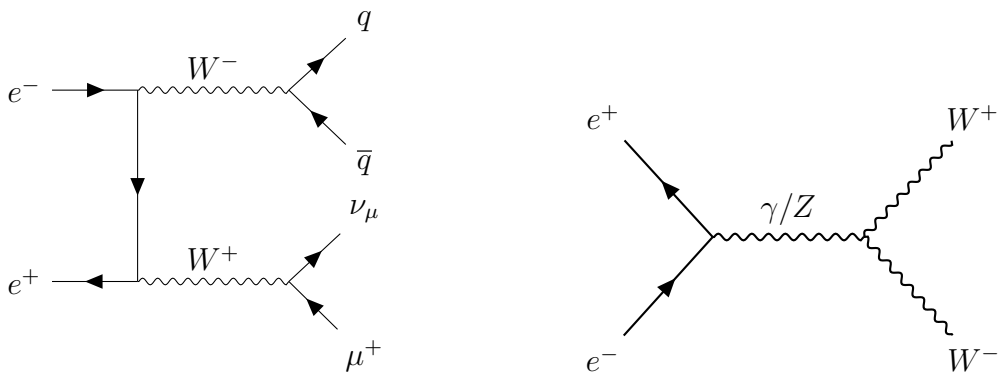
with subsequent semi-leptonic decays, including their irreducible SM backgrounds, e.g. photon radiation into a lepton or jet pair.
We will refer to the former process as *charged* and to the latter process as *neutral*, as will be exposed in the following two subsections. Moreover, we have decided to omit a flavour sum over different lepton flavours in the final state to reduce the computational effort and to highlight muon production processes.

2.1 Phenomenology of $e^+e^- \to \mu\nu jj$

The charged process rightfully bears its name: it describes an electron scattering where a muon/neutrino pair and a pair of two jets are produced, both of them being mediated by a charged off-shell W boson.
Here, we have made the following assumptions:

- We are only interested in the presence of a muon/neutrino pair and thus allow for a flavour sum over these particles and their antiparticles. Thus there are two different possibilities: $\mu^-, \overline{\nu}_\mu$ and μ^+, ν_μ.
- The jets originate at the partonic level from a quark/antiquark pair which subsequently hadronises in the moments after the interaction. Only the heavy top and bottom quarks are not considered as light (massless) quarks here, but we consequently allow a flavour sum over the first two quark families.

All of these possibilities are summarised in the corresponding Feynman diagrams at LO in figures 1a and 1b. The two W bosons are either produced by the exchange of a virtual fermionic current or are the result of the decay of a virtual Z boson or a photon. In the right panel of figure 1 we will obtain the same subsequent decay of the diboson pair as is shown in the left panel. Since the light quarks on which the jet production depends are assumed to be massless, cuts have to be defined to avoid divergencies due to soft and/or collinear emissions spoiling the integration.
We can expect the following progression of the integrated cross section: since the process is exclusively mediated by W bosons, the probability of interaction decreases rapidly for collision energies significantly lower than the production threshold of the W pair, which is roughly 160 GeV. The highest cross section can be expected shortly after the 2-W-resonance, when a larger phase space is made accessible to the different final state particles. At higher energies the cross section will decrease again.



(a) Diboson production via a virtual fermionic current.

(b) Diboson production via a virtual Z boson.

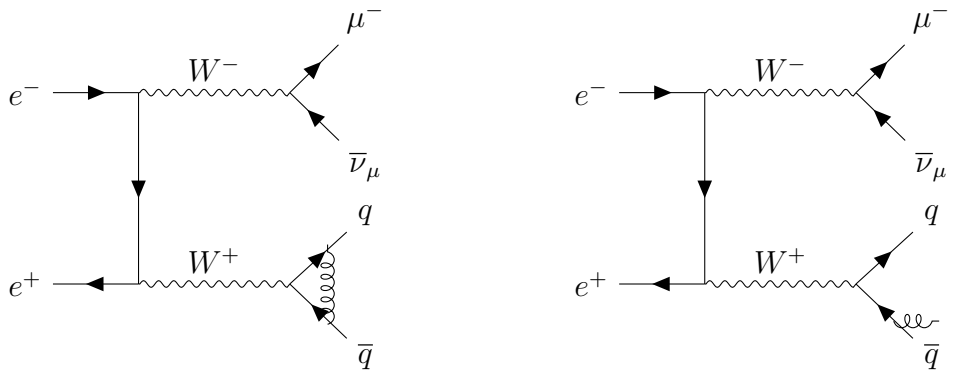
Figure 1: LO Feynman diagrams for the process $e^+e^- \to \mu\nu jj$

The NLO QCD corrections we consider for this process contribute in the following

way:

- Virtual gluon loops can occur in the hadronic decay of the W boson. Another possibility are self energy diagrams where a gluon loop occurs within a single external (anti-)quark line. Once the UV renormalisation has been done by the one-loop providers, those diagrams are no longer needed.
- Furthermore, real gluon emission is possible. Said gluons are susceptible to undergo a hadronisation process and will thus be detected as additional jets.
- The leptonically decaying W boson is subject to interactions of pure EW nature and is not influenced by the corrections we consider.

For these corrections, figures 4a and 4b show representative sample diagrams. We want to stress that, since flavour sums over the first and second generation quark flavours are present in the final state, many equivalent diagrams with different final states can be obtained from the diagrams shown here.



(a) NLO correction due to a virtual gluon loop

(b) NLO correction due to a real gluon emission

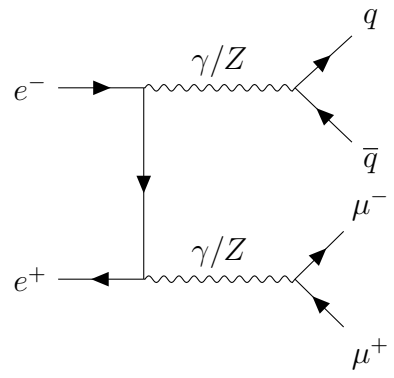
Figure 2: NLO Feynman diagrams for the process $e^+e^- \to \mu\nu jj$

2.2 Phenomenology of $e^+e^- \to \mu\mu jj$

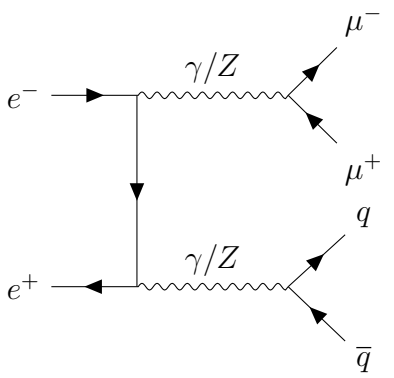
The neutral process is in many ways similar to the charged process, as it describes an equivalent scattering experiment where a muon/antimuon pair is created with an associated 2-jet system. The important difference, however, is that the mediating vector particle can now either be a Z boson or a photon. This has several consequences:

- We can now omit the flavour sum over particles and antiparticles in the leptonically decaying boson, the muon/neutrino system is replaced by muon/anti-muon pair.
- A strong photonic background is present in the low dimuon and dijet invariant mass range. These background processes are of minor interest to us, and we suppress them as much as possible by choosing appropriate phase space cuts.

As the Feynman diagrams at LO in figures 3a and 3b show, both final state systems can originate from a photon, which, especially in the range of low dimuon and dijet invariant masses, is the preferred vector particle. So when this process is simulated, the integration below the Z-production resonance is considerably altered by said photonic background, yielding much higher cross sections than those of the weakly mediated process. Hence, suitable phase space cuts are needed to restrict both final state systems to lie in a reasonable region around the Z boson mass, allowing to select the signal process.



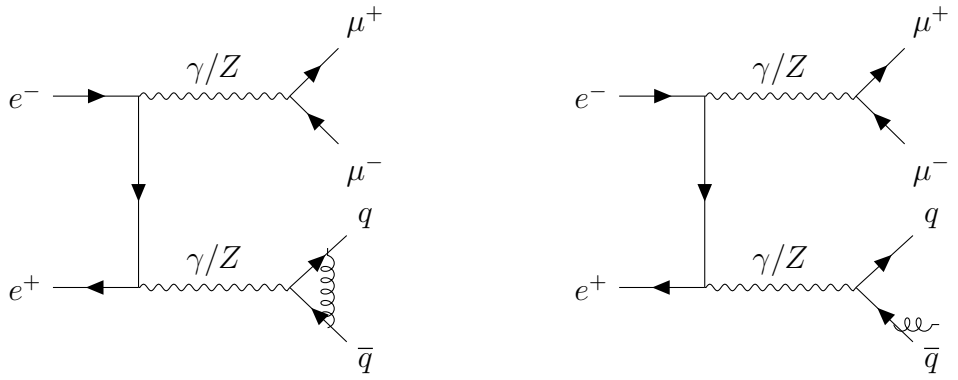
(a) Diboson production via a virtual fermionic current.



(b) In complete analogy to figure 3a, charge conjugated diagram.

Figure 3: LO Feynman diagrams $e^+e^- \to \mu\mu jj$

The corrections to the neutral process at NLO QCD are very similar to those already shown for the charged process. EW corrections have not been considered in the analysis, and so we are left with real and virtual gluon loops. These couple neither to the leptonic system nor to the vector bosons, therefore we obtain the same NLO corrections in the jet system, as representative Feynman diagrams at NLO in figures 4a and 4b illustrate. Again, we would like to remind here that more equivalent diagrams can be obtained by writing down the flavour sum explicitly.



(a) NLO correction due to a virtual gluon loop

(b) NLO correction due to a real gluon emission

Figure 4: NLO Feynman diagrams $e^+e^- \to \mu\mu jj$

2.3 A temporary conclusion: why we need WHIZARD

So far, we have defined two physical processes that can be considered within the SM of particle physics. Before we proceed with the computational aspects of this analysis, we want to make the reader aware of the reasons why such heavy use of computational tools is necessary[3].

The two processes $e^+e^- \to \mu\mu jj$ and $e^+e^- \to \mu\nu jj$ are two scattering processes in a lepton collider setup involving the additional production of two jets. Central to a sound understanding of these processes is the integrated cross section σ which we can associate to any CM collision energy and which represents the statistical occurrence of this process. The integrated cross section, as insinuates its name, is obtained by integrating the differential cross section over all possible angles in space. On the other hand, the differential cross section results from evaluating the corresponding matrix elements in the chosen physics model.

In the SM, we evaluate these matrix elements by constructing the relevant Feynman diagrams which stand for an expansion of the S-matrix and yield, according to the associated Feynman rules, the integral that we need to evaluate. These integrals can become arbitrarily complex for more and more final state particles or in higher order perturbation theory due to more (and more) complicated Feynman diagrams. For the least possible number of vertices we obtain the Leading Order, by including the

[3]The content of this subsection can be found in as much detail as the reader wishes in any textbook on Quantum Field Theory (QFT). Here, we have made use of [8] and [9] as our standard references on QFT and particle physics.

diagrams with the next higher order in the coupling constant, such as the 1-Loop-diagrams, one can carry out calculations up to the Next-to-Leading-Order, and so forth.

Specialised computational tools such as WHIZARD and others help us to analyse in principle arbitrarily complicated physical processes in an efficient manner due to the automation of matrix element generation, phase space integration and event generation as well as convenience tasks like CM energy scans. Calculations can be set up to exactly match the experimental conditions that they are to be compared with, cuts can be applied to highlight different phase space regions or to filter undesired background processes and different physics models can be tested, to mention only a few advantages.
It is natural to also include calculations at higher orders to extend the applicability of WHIZARD and to increase the precision of the results it yields. This leads to the following problem: at NLO, we have to consider loop and real emission diagrams. Loop diagrams potentially contain ultraviolet (UV) and infrared (IR) divergencies when performing the loop integrals:

- UV divergencies occur when internal momenta become arbitrarily large. This is dealt with in renormalisation theory and absorbed in the constants of the model.

- IR-divergencies are the result of arbitrarily low momenta or arbitrarily collinear momenta. They only exist for massless (or quasi massless) particles and concern internal loops and final state radiation. The latter can be removed by appropriate cuts, whereas the former motivates dimensional regularisation.

Dimensional regularisation is a regularisation procedure to render either phase space or loop integrations finite by deforming the integrations analytically to non-integer $4 \pm 2\epsilon$ dimensions. The KLN-theorem [10, 11] indeed ensures that soft and collinear singularities cancel between virtual and real diagrams order by order in perturbation theory. Since dimensional regularisation cannot be included in tools based on MC integration which can only generate and treat vectors with a whole number of entries - see for example in reference [12] -, workarounds had to be created, which led to the famous subtraction schemes, e.g. the Frixione-Kunszt-Signer (FKS) subtraction [13, 14]. We will come back to this in the next section.

Explaining the complexity of numerical calculations at NLO using MC event generators in all gory details is beyond the scope of this thesis. More information can be found in the references.

3 Computational environment: WHIZARD

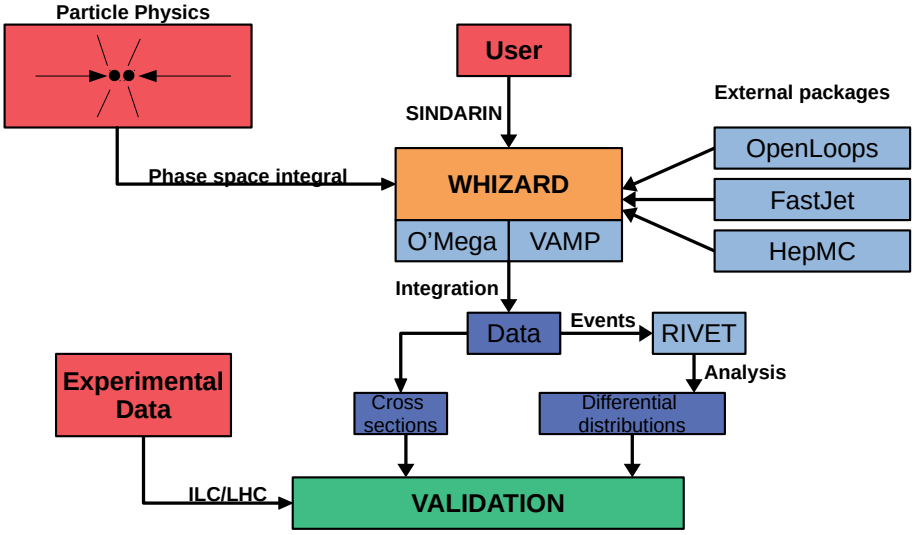


Figure 5: Workflow in WHIZARD

This graphic representation in figure 5 displays the workflow of how WHIZARD has been used in the present thesis. We will present WHIZARD by delving directly into its application in order to highlight the close connection between the physics established in the first chapter and the tools we use to process it. After briefly commenting on each step, several foci will go into more detail and expose the internal working mechanisms of WHIZARD that are important for our study.

3.1 Workflow

As explained in the second chapter, the interest in a certain physical process in a given physics model ultimately leads to a possibly large number of Feynman diagrams that need to be evaluated and, depending on the exact setup of the calculation, rendered finite. WHIZARD provides a core program that gathers MC integration and event generator functionalities in one easily accessible software. In our case, the main functions of WHIZARD that have been used are on the one hand the pure evaluation of phase space integrals and the subsequent computation of cross sections, and on the other hand the generation of events which yields an analysis of the differential distribution of the observables of interest. We will come back

to this in more detail in the first focus of this section. For up-to-date information on WHIZARD's functionalities as well as current and future developments, the reader may consult the manual [15] on the corresponding web page and references [16, 17, 18].
The two main components of WHIZARD, as represented in 5, are the packages O'Mega [19] and VAMP [20]. O'Mega is used to generate the requested matrix elements, but can do so only at LO. Suitable phase space parametrisations are generated by WHIZARD, which are then used by VAMP to perform adaptive multichannel integrations. Given the setup of our analysis, we also need to interface several external packages: OpenLoops [6] is a publicly accessible library which provides us with the necessary matrix elements at NLO and thus extends the applicability of WHIZARD considerably[4]. HepMC [23] is an event format in High Energy Physics (HEP) and FastJet [24] is responsible for jet clustering, to be detailed in Focus 3. WHIZARD can in principle deal with processes of an arbitrary complexity, e.g. a high number of final state particles. Of course limitations are imposed by the available computing power, but the user can influence the behaviour of WHIZARD for example by adjusting the desired precision of the integrations she wants to perform.

To talk to WHIZARD, the user needs to learn the scripting language SINDARIN which has been specifically designed for this purpose, see sections 4 and 5 of [15]. WHIZARD can thus either be steered by a .sin file or even interactively in the console, but for complex processes the use of steering scripts has obvious advantages. A brief tutorial on how SINDARIN has been used to implement a particle physics process in WHIZARD is given in the last section of this chapter.
Once the user has successfully installed WHIZARD, linked all the necessary packages and set up a working SINDARIN, she hopefully can harvest the desired output. Here, the harvest consists of two parts: the integrated cross section of the process as a function of the collision energy and a given number of events. Whereas the integrated cross sections are immediately usable for our purposes, the events have to be fed in an analysis tool first. Even though some analysis functionalities are also implemented in WHIZARD, in order to obtain a coherent analysis, i.e. an evaluation of the process observables, at NLO, we will use the software RIVET [25].
Last but not least, a future validation can compare the integrated cross sections and differential distributions given through the analysis to actual experimental data, and thus find either limitations imposed by the computational environment, or, more

[4]There exist other possibilities, e.g. the package RECOLA [21, 22]

significantly, failures in the physics model that has been employed.

3.2 Focus 1: Generating events in WHIZARD

Let us start this focus by stating what an event actually is: in a particle detector, huge numbers of beam bunches are crossed every second and interact in accordance with the underlying physics model. An event describes the result of one of these interactions. If we define the physical processes we would like to investigate, different events possess the same initial and final states [5], but can have different kinematics, e.g. different distributions of momenta between the final states.
In an experimental setup, different events occur successively and thus without any correlation. What we are interested in is therefore a possibility to simulate these events, which will enable us to obtain the statistics inherent to a certain process. These can ultimately be compared to a sufficiently large sample of experimental data to verify whether said predictions match the "actual behaviour".

How is event generation handled by WHIZARD? To fully understand how events are obtained from a phase space integral we would have to expose how a MC integration is set up in WHIZARD, which lies beyond the scope of this thesis [6]. An important fact is that due to the generation of random numbers we obtain a non-uniform grid of the phase space which is used to replace the integral with a weighted sum. The reason to use a non-uniform grid is that it renders possible a representation of the singularities of the propagators of resonant particles as well as the radiation patterns associated to QED and QCD. This grid is then reused in the generation of events that will therefore carry the weight resulting from the previous binning. Physically this is undesirable: as stated above, we are only interested in uncorrelated events of equal weight, hence when weighted events are generated they have to undergo an unweighting procedure.

For our purposes, we have made the following decisions:

- Due the fact that unweighted events cannot be generated for a fixed-order NLO calculation (phase space regions may have positive or negative weights), we

[5] To be more precise, equivalent initial and final state, in order to allow flavour sums.

[6] Further information may be found in references [2, 26]; for an introduction to MC methods see for example [27].

will generate only weighted events. For more clarity and resemblance between the steering scripts for event generation we dispose of a simulation at LO and of three separate simulations for the Born, real and virtual integrals at NLO, respectively.

- Before the events are analysed with RIVET, we set the event normalisation to the total NLO cross section.

- Finally, RIVET rescales the histograms by dividing by the respective sum of event weights to obtain histograms that are normalised to the cross section.

The commented analysis that has been written using RIVET can be found in the appendix.

3.3 Focus 2: Subtraction schemes in WHIZARD

An important ingredient of an NLO calculation is the subtraction scheme used to render the phase space integrals finite, as has been mentioned in section 2. A brief reminder: the need for a subtraction scheme is given by the fact that dimensional regularisation as an instrument to deal with IR divergencies is not an option for MC event generators.

The squared matrix element at LO is given by the square of the absolute value of the sum of the matrix elements representing the different possibilities to mediate the process:

$$|M_{LO}|^2 = |\sum_i M_{LO,i}|^2, \tag{7}$$

where each $M_{LO,i}$ thus results from a different Feynman diagram [7].
Similarly, we obtain a sum of different contributions in the matrix element at NLO [8]:

$$|M_{NLO}|^2 = |M^B_{NLO}|^2 + |M^R_{NLO}|^2 + 2Re(M^B_{NLO} M^{V*}_{NLO}). \tag{8}$$

This should be read as follows: The squared matrix element at NLO is made up of three components:

[7] Again, any QFT textbook like [8] will motivate this.
[8] Compare with reference [12].

- First comes the so called *Born contribution*, which has the same Feynman diagrams as the matrix element at LO. Nevertheless, the *Born contribution* can slightly differ from the LO contribution when evaluated due to additional NLO settings, e.g. scale variations that are taken into account or parameters like coupling constants, masses and decay widths.
- The second matrix element is referred to as *real contribution* for it represents real emissions (in this work only in the final state), e.g. gluons that can form a jet.
- Finally, the *virtual contributions* account for internal loops.

Note that modern tools like WHIZARD do not use sums over Feynman diagrams, but recursive methods to calculate amplitudes much more efficiently. Subtraction schemes enter the stage when this equation is integrated to obtain the cross section at NLO. Integrating the NLO matrix element gives

$$\sigma_{NLO} = \int d\mathrm{PS}\, |M_{NLO}|^2 = \sigma^B + \int d\mathrm{PS_{n+1}}\, |M^R_{NLO}|^2 + \int d\mathrm{PS_n}\, 2Re(M^B_{NLO} M^{V*}_{NLO}) \tag{9}$$

where dPS is a (very) compact notation for the Lorentz invariant phase space measure including the flux factor. The integrals in this expression are not yet necessarily finite, but can be rendered so if divergent terms are taken from the real phase space and added in the integral of the virtual matrix element:

$$\begin{aligned} \sigma_{NLO} = \sigma^B + \int d\mathrm{PS_{n+1}}\, (|M^R_{NLO}|^2 - d\sigma^{\mathrm{S}}(\mathrm{PS_{n+1}})) \\ + \int d\mathrm{PS_n}\, 2Re(M^B_{NLO} M^{V*}_{NLO}) + \int d\mathrm{PS_n}\, d\mathrm{PS_{rad}}\, d\sigma^{\mathrm{S}}(\mathrm{PS_{n+1}}). \end{aligned} \tag{10}$$

The first integral in this expression is finite by construction, the second line is finite by KLN. The subtraction terms are chosen in the way that the integration over the one-particle radiation phase space variables can be carried out analytically, such that these integrated subtraction terms can be integrated numerically with the virtual terms over the Born phase space.

As the term being subtracted from the real matrix element and added back to the virtuals is not uniquely defined, there are several different subtraction schemes. WHIZARD uses the Frixione-Kunszt-Signer- (FKS) subtraction scheme, which is implemented in the program for arbitrary processes[9].

[9] Another possibility would be the Catani-Seymour-Algorithm [28], for example.

3.4 Focus 3: Jet-Clustering in WHIZARD with FastJet

This study is designed to account for the presence of jets in the two scattering processes we are considering. By jet, one understands either a bundle of quarks and gluons clustered together in a catchment area defined by a certain jet measure, or (after fragmentation and hadronic decays) calorimetric energy deposits by electromagnetic showers in the (electromagnetic or hadronic) calorimeter of the detector. The first is a theoretical concept which is used in the study of this thesis [24, 29].
WHIZARD makes use of the C++ library FastJet - which has to be specified during the installation and configuration of WHIZARD - as a provider for several jet algorithms which are responsible for the clustering of selected partons to form a jet. Further explanations of the working mechanisms of jet-clustering can be found for example in [30].
When one uses jet clustering in WHIZARD, there are two main parameters that have to be defined in the SINDARIN file: the jet algorithm which is given by a statement like

```
jet_algorithm = kt_algorithm,
```

here showing as an example the commonly used `kt_algorithm`, and the corresponding jet-parameter:

```
jet_r = 0.4.
```

The latter roughly defines the area which is taken into account to cluster the jets and can be adjusted if, for example, jets tend to be emitted in the same direction.

3.5 Focus 4: Using WHIZARD - speaking SINDARIN

In order to use WHIZARD, the user has to write a SINDARIN file that represents the physical process to be analysed as well as all the necessary settings and parameters. WHIZARD supports multiple physics models; in our case only the SM of particle physics is relevant and will be considered as a default setting from now on.
The most important part of a SINDARIN is of course the physical process itself,

intuitively defined by a statement like

```
process ee2lljj = e1, E1 => lep, lep, jet, jet,
```

which in this case stands for the process

$$e^+e^- \to lljj\,.$$

The aliases `lep` and `jet`, representing leptons and jets, can be defined at the beginning of the file, so the user has the possibility to implement a flavour sum in the final state if she wishes to consider multiple generations of equivalent final state particles. `e1` and `E1` simply denote the electron and the positron, respectively.
In addition to some technical settings like the integration mode or the running of the strong coupling constant, the user can define her personal parameters in the physics model that is used. A convenient choice is to set the light quark masses to zero, being insignificant in the context of the collision energies we are working with. A convenient by-product of this a considerable simplification of the numerical calculation.
Once the technical parameters have been set, one can adjust beam options and jet clustering. Possible beam options are for example beam polarisation, the correction of beamstrahlung effects and initial or final state radiation (ISR/FSR), described for example in [31]. Final state jets are dealt with by specifying a jet algorithm and a jet parameter to cluster the jets. Further important setting include the application of phase-space cuts in order to select the signal process and define a finite cross section. Finally, the user has to decide whether she wants to perform the calculations at LO or at NLO. For this, she also has to set the desired precision of the integration which is realised by the repetition and the refinement of a pre-set number of iterations. If an integration at NLO is done, it is important to specify explicitly how WHIZARD can access the necessary matrix elements, since O'Mega itself can only evaluate tree-level Feynman diagrams. Here, the matrix elements provided by OpenLoops can simply be loaded during the integration.
To make use of WHIZARD's event generation, the necessary settings specifying for example the number of events and the desired event format have to be implemented after the integration. Optionally, some analysis statements can be defined if the user wishes to use WHIZARD's internal analysis. This implies the definition of the relevant observables after a `simulate` statement at the desired precision, which will

afterwards be filled in histograms. Alternatively, the events can be fed in a external analysis tool like RIVET.

Since the physical as well as the computational background have now been motivated, in the next section we will proceed with the numerical evaluation of our calculations and present the results that we have obtained.

4 Numerical results

The numerical study of the processes $e^+e^- \to \mu\mu jj$ and $e^+e^- \to \mu\nu jj$ comprises two main parts: important first results are the scans of the integrated cross sections of the charged and the neutral process, respectively, since they describe their signal rates in the full range of CM energies of future lepton colliders. In tables 1 and 2, we will furthermore present selected values of the integrated cross section at some selected CM energies, representing different choices for a linear collider setup or production thresholds for vector bosons.
Secondly, we will focus on the differential distribution of certain observables, i.e. analyse how the cross section varies with respect to a variation of energies or transverse momenta of jets and leptons or other variables.
This will give us a thorough overview of what the SM predicts for these two lepton collisions and provide experimentalists with the necessary MC data that can be used to compare the theoretical predictions to future experimental data.

4.1 Setup of the calculation

The workflow of the calculations can be summarised in a few words: We write separate SINDARIN files for the scan at LO and NLO as well as for the event generation at LO and NLO, the latter being followed by a subsequent analysis in RIVET. To obtain a clear and interpretable result, we have chosen the following settings to evaluate the phase-space integrals (the curious reader may refer to the appendix to study the SINDARIN and the RIVET analysis scripts where these have been implemented):

- All light quark and lepton masses, being insignificant compared to collision energies in the range of $10^2 - 10^3$ GeV, are set to zero.
- The resulting jets are clustered by FastJet, using the `kt_algorithm` and a `jet-parameter` of `R=0.4`.
- Several mild phase-space cuts have been implemented in order to render the calculations finite: phase-space integrals are only evaluated if there are at least two jets having a pseudorapidity[10] $|\eta|$ of less than 4. The same holds true for

[10] A selection cut on the absolute value of the pseudorapidity is necessary to account for the

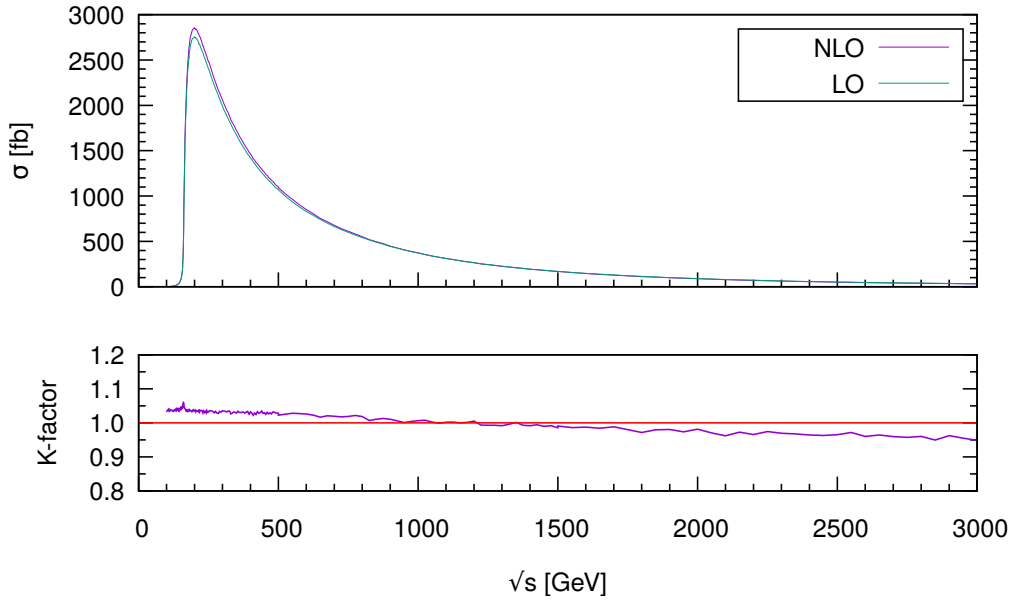


Figure 6: Integrated cross sections for the process $e^+e^- \to \mu\nu jj$ as a function of $\sqrt{s}$; corresponding *K-factor*

the produced leptons which we require furthermore to have an invariant mass of more than 10 GeV. In addition to that, we require each jet, in particular a third jet due to real gluon emissions, to have a transverse momentum of more than 20 GeV. This protects the integration against a potentially soft or collinear jet at NLO. If any of the cut-conditions is violated, the corresponding matrix element is vetoed.

- For the neutral process $e^+e^- \to \mu\mu jj$, we define the following additional cut: The invariant masses of both the lepton- and the jet-system must lie in the energy range $80 - 100$ GeV, thus centred around the on-shell Z boson mass. This ensures that the photonic background is sufficiently reduced.

- We approximate the *CKM matrix* as unity.

At this point, we must define an important quantity that will be of great use to us in both the analysis of the neutral and the charged process: the comparison of data at LO and NLO consequently leads to a quotient, the *K-factor*, which describes the

detector coverage: final state particles can only be detected if they have an technically measurable angle with the beam axis, corresponding to an upper bound of the pseudorapidity η as given in the text.

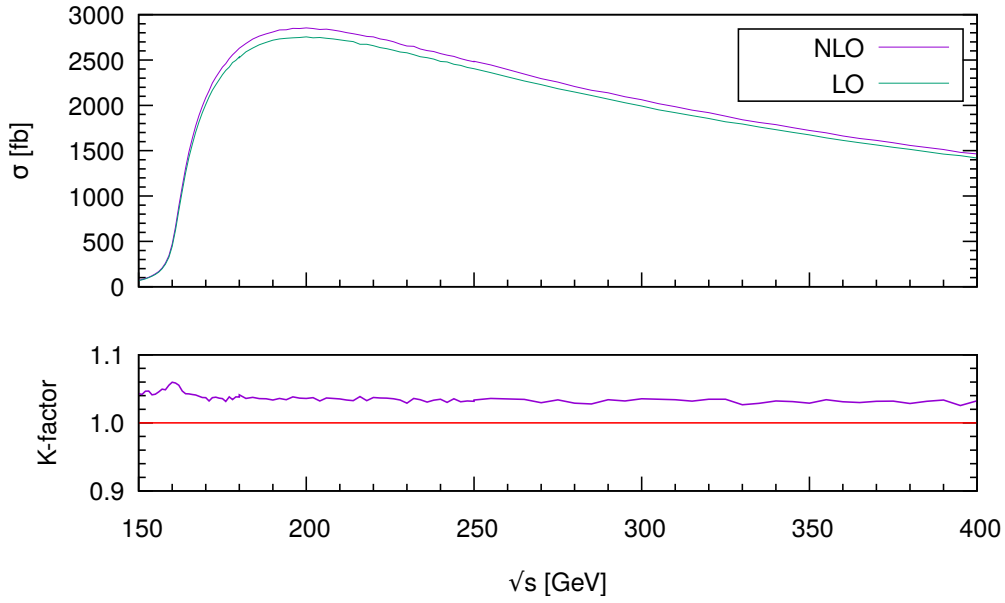


Figure 7: Integrated cross sections for the process $e^+e^- \to \mu\nu jj$ as a function of $\sqrt{s}$ in a reduced range centred around the peak region; corresponding *K-factor*

ratio of the two integrated cross-sections:

$$K^{NLO} = \frac{\sigma^{NLO}}{\sigma^{LO}}.$$

A similar definition will be applied in the case of the differential distributions. Studying the behaviour of the *K-factor* thus permits us to tell where the results at NLO yield important corrections to the results at LO, and where the corrections tend to be negligible. The main advantage of a Monte Carlo simulation is that the *K-factor* can be studied differentially in any possible observable.

4.2 Integrated cross sections and scans

The following section shows the numerical results that we have obtained for the integrated cross sections as a function of the CM collision energy, thus giving us information about the signal rates of the two processes in different linear collider setups. We have decided to show the plots with renormalisation scales of m_W and m_Z, respectively. Theoretical higher order uncertainties due to scale variations by a factor of two around the central scales are marked in tables 1 and 2; the corresponding scans can be found in the appendix.

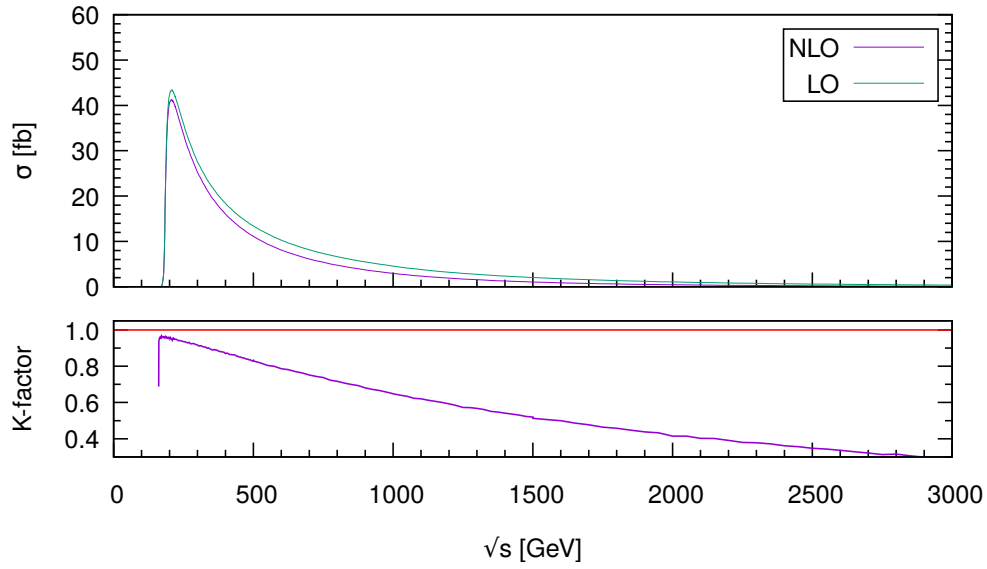


Figure 8: Integrated cross sections for the process $e^+e^- \to \mu\mu jj$ as a function of $\sqrt{s}$, corresponding *K-factor*

4.2.1 Integrated cross section of $e^+e^- \to \mu\nu jj$

Figure 6 shows how the integrated cross section of the charged process changes with $\sqrt{s}$, figure 7 displays a section of the previous scan by showing only low CM energies up to 400 GeV. The integrated cross section peaks slightly above the kinematic threshold for on-shell WW production at 200 GeV with 2.85 pb at LO, and then falls off quadratically with energy to half a pb at roughly 850 GeV. At full CLIC energy of 3000 GeV it drops below 100 fb.

In the lower diagram of figure 6, we can observe how the *K-factor* associated to the NLO QCD corrections to the muon/neutrino production process varies.
Globally, it is striking that the NLO QCD corrections are contained in a range of $\pm 5\%$ around the LO predictions. For low collision energies, the results at NLO take slightly bigger values than the results at LO. Both predictions are almost equal in the range of roughly $950 - 1200$ GeV. Above 1500 GeV, the NLO results are decreasing faster than the results at LO and can be found in the region between 1% and 5% below LO.

Table 1 shows several values of the integrated cross sections at LO and NLO,

accounting for small errors caused by a different choice of the renormalisation scale.

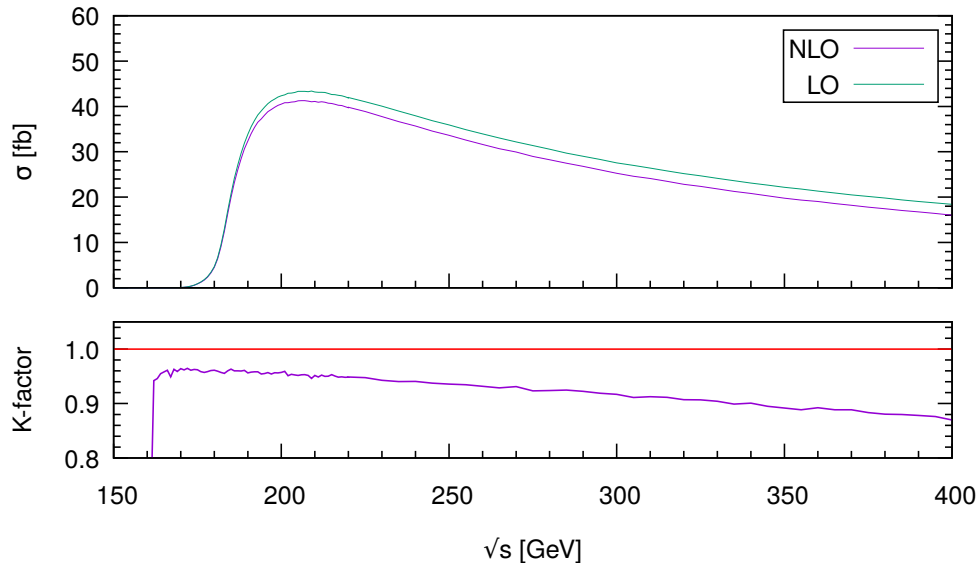


Figure 9: Integrated cross sections for the process $e^+e^- \to \mu\nu jj$ as a function of $\sqrt{s}$ in a reduced range centred around the peak region; corresponding *K-factor*

4.2.2 Integrated cross section of $e^+e^- \to \mu\mu jj$

In comparison to the charged process, figures 8 and 9 depict the integrated cross section of the neutral process.

As we have restricted the masses of the muon/antimuon- and the jet system to lie between 80 GeV and 100 GeV, the results that we obtain for the cross section begin at 160 GeV. After a very sharp rise, σ peaks at roughly 205 GeV (so ca. 25 GeV above the kinematical threshold) with a value of 41 fb. For higher CM energies, it tends to decrease marginally faster than the cross section of the charged process, dropping below 10 fb at roughly 530 GeV. The total cross section further falls off linearly with $\sqrt{s}$ and drops below 1 fb above 1500 GeV.

Here, the *K-factor* shown in the lower panel of figure 8 does manifest a different behaviour than the *K-factor* we have studied in the previous section. It has its highest value of 0.97 right above the peak position, which is itself shifted compared to the LO peak by about -5%. For higher energies, the NLO cross section is constantly and increasingly lower than the LO cross section, yielding a *K-factor* of less than 1.

$\sqrt{s}$	σ_{LO}[pb]	σ_{NLO}[pb]	*K-factor*
160	0.4446	$0.4711^{+0.36\%}_{-0.62\%}$	$1.060^{+0.28\%}_{-0.66\%}$
200	2.755	$2.854^{+0.21\%}_{-0.81\%}$	$1.036^{+0.19\%}_{-0.77\%}$
250	2.405	$2.481^{+0.64\%}_{-0.12\%}$	$1.032^{+0.58\%}_{-0.19\%}$
500	1.070	$1.101^{+0.45\%}_{-0.73\%}$	$1.028^{+0.58\%}_{-0.69\%}$
1000	0.3710	$0.3734^{+0.00\%}_{-0.73\%}$	$1.006^{+0.00\%}_{-0.70\%}$
1500	0.1694	$0.1670^{+0.47\%}_{-0.05\%}$	$0.9860^{+0.50\%}_{-0.10\%}$

Table 1: Integrated cross sections for $e^+e^- \to \mu\nu jj$ at different values of $\sqrt{s}$

$\sqrt{s}$	σ_{LO}[fb]	σ_{NLO}[fb]	*K-factor*
180	4.707	$4.524^{+0.35\%}_{-0.53\%}$	$0.9611^{+0.35\%}_{-0.53\%}$
200	42.34	$40.52^{+0.14\%}_{-0.55\%}$	$0.9570^{+0.15\%}_{-0.55\%}$
250	35.94	$33.63^{+0.56\%}_{-1.02\%}$	$0.9357^{+0.57\%}_{-1.01\%}$
500	13.40	$11.16^{+1.70\%}_{-2.90\%}$	$0.8328^{+1.71\%}_{-2.85\%}$
1000	4.555	$2.954^{+5.45\%}_{-7.11\%}$	$0.649^{+5.38\%}_{-7.81\%}$
1500	2.038	$1.056^{+7.90\%}_{-11.4\%}$	$0.5181^{+7.87\%}_{-13.7\%}$

Table 2: Integrated cross sections for $e^+e^- \to \mu\mu jj$ at different values of $\sqrt{s}$

At roughly 1600 GeV, the NLO results account for a correction of 50% and ultimately diminish the LO prediction for σ by 70% for the very high collision energies of more than 2800 GeV.

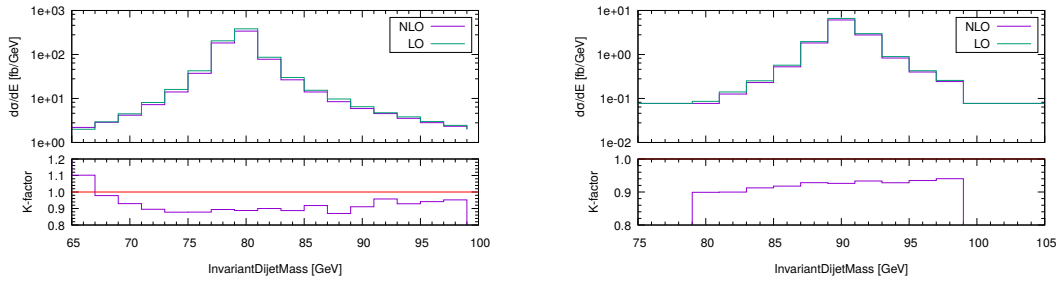
At this point we should stress that the NLO corrections for the neutral process are much more dominant compared to the charged process, even though one would intuitively expect them to be in a similar range. At the time when this thesis was written, this fact was not yet fully understood and deserves further investigation.

More exact values for chosen energies can be found in table 2, showing a similar table with results for the integrated cross sections as well as possibles corrections to it.

4.3 Differential distributions

We conclude our study of the processes $e^+e^- \to \mu\nu jj$ and $e^+e^- \to \mu\mu jj$ by presenting the differential distributions of the integrated cross section with respect to chosen observables.
Calculations have been done exclusively at a CM energy of $\sqrt{s} = 250$ GeV, representing the official setup for the first stage of the ILC. The setup in this section is the same as in the previous section, no further assumptions have been made. The physics analysis has been performed using the program RIVET[11].
Every plot contains the corresponding histograms at LO and at NLO, and the ratio of both datasets is again given by a *K-factor*. It should be stressed that both processes are, topologically speaking, very similar. Hence, we can expect that differences between the differential distribution result almost exclusively from the different kinematics of the final state particles due to different intermediate vector bosons. The photonic background of the neutral process mainly plays a role for low collision energies but is reliably reduced due to the phase space cuts that we have applied.

4.3.1 Reconstruction of the intermediate bosons



(a) Reconstructed invariant dijet mass of the charged process

(b) Reconstructed invariant dijet mass of the neutral process

Figure 10: Reconstructed invariant dijet masses of intermediate bosons

Figure 10 shows the reconstructed invariant dijet mass of $e^+e^- \to \mu\nu jj$ in the left panel and the reconstructed invariant dijet mass of $e^+e^- \to \mu\mu jj$ in the right panel. As expected, one immediately recognises the Breit-Wigner shaped mass distributions

[11] The analysis script that has been written for this purpose can be found in the appendix and may serve as a reference similar projects.

of the W and Z boson. The peaks of the masses are at the bin positions around the input values

$$m_W = 80.385 \pm 0.015\,\text{GeV}$$

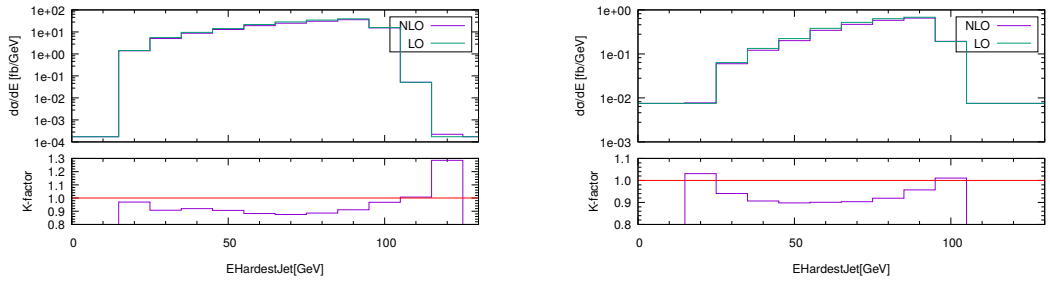
for the W boson and

$$m_Z = 91.1876 \pm 0.0021\,\text{GeV}$$

for the Z boson (taken from [32]), which matches the peaks of the histograms lying in the regions $79 - 81$ GeV and $89 - 91$ GeV.

At NLO, one observes a very similar behaviour in both histograms. Apart from a slight rise of about 10% at 66 GeV in the histogram of the reconstructed mass of the W boson, the NLO calculations yield a relatively constant correction between -5% and -10% over the whole relevant energy range. This confirms our expectations: at NLO, we can anticipate the total cross section to be less sensitive to a variation of the mass of the dijet-system due to the presence of a third jet, resulting from real gluon emissions.

4.3.2 Comparison of the jet systems



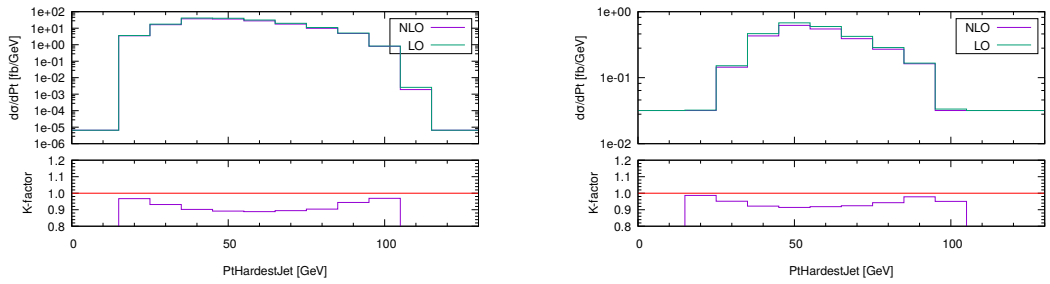
(a) Energy of the hardest jet in the charged process

(b) Energy of the hardest jet in the neutral process

Figure 11: Comparison of the energy distribution of the two hardest jets for both processes

Another possibility to compare the processes $e^+e^- \to \mu\nu jj$ and $e^+e^- \to \mu\mu jj$ is to study the behaviour of the hardest jets. Their energies are shown in figure 11. As the phenomenological background in the second chapter motivates, only the jet-systems are subject to corrections due to QCD. We can therefore expect to witness the same

corrections in both processes, which is in accord with the differential distributions we obtain: for low energies, the NLO results match the LO within a deviation of $\pm 3\%$, then drop at about -10% below LO between 30 GeV and 80 GeV, which can be attributed to the presence of a third jet. Only for high energies, we again observe a slight increase of the *K-factor*.
It should also be stressed that we can access more phase space in the case of the charged process since the mediating W boson has about 10 GeV less mass than the Z boson of the neutral process and thus a higher kinetic energy. This is manifest in the differential distributions in the high energy range: in the case of the mediating Z boson, the cross section is sensitive to variations of the jet energy up to about 100 GeV whereas in the case of a charged current, variations can be observed until 120 GeV.



(a) Transverse momentum of the hardest jet in the charged process

(b) Transverse momentum of the hardest jet in the neutral process

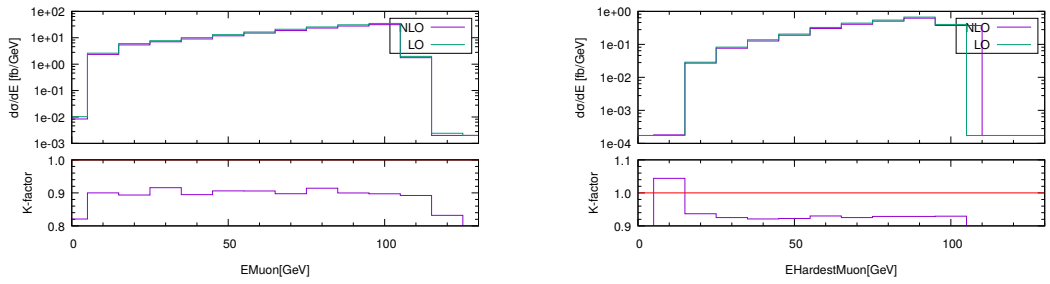
Figure 12: Comparison of the transverse momenta of the two hardest jets for both processes

A similar reasoning applies to a comparison of the transverse momenta of the hardest jets, illustrated in figure 12. The *K-factor* yields significant corrections of about -10% for energies centred around the peak of the distributions but approaches the LO for both high and low transverse momenta; about 10 GeV of additional transverse momentum are accessible for the hardest jet of the charged process.

A physical explanation for this behaviour could be that for low as well as for high transverse momenta, real corrections become less significant and no third jet appears. In the first case, both jets propagate close to the beam axis and not much phase

space is available, hence no or only few real gluon emissions occur. In the second case, the majority of energy available in the dijet system is used to form the transverse momentum of the hardest jet, allowing at most for soft gluon emissions since we also require the presence of a second jet. The overall trend of the corrections is very similar in both the charged and the neutral process. The most significant NLO corrections occur in the range of $40 - 70$ GeV - being close to the Jacobian peak we can conclude that the energy of the hardest jet is composed of about half of the invariant W or Z boson mass, as well as a recoil due to the lepton system and the remaining collision energy, which leads to the observed p_T-peak. Real radiations at NLO require some additional energy and are thus responsible for a negative NLO correction of about -10%.

4.3.3 Comparison of the lepton systems



(a) Energy of the (only) muon in the charged process

(b) Energy of the hardest muon in the neutral process

Figure 13: Comparison of the energies of the muons produced in both processes

Finally, we have a look at each of the two lepton systems of $e^+e^- \to \mu\nu jj$ and $e^+e^- \to \mu\mu jj$, resulting each from second W or Z boson.
As we consider only NLO corrections due to QCD, no additional vertices in the leptonic systems are taken into account, hence all possible differences between the leptonic observables at LO and NLO are the result of corrections occurring in the dijet system.
The left panel of figure 13 displays the differential distribution of the muon stemming from the W boson in the charged process, in the right panel we show the hardest muon produced in the neutral process. Again, we make the observation that in the charged process, more phase space is available to the leptons, the energy of the muon

may take a value up to 20 GeV higher than the energy of the hardest muon in the neutral process. In both cases, similar to the other NLO corrections we have studied, the *K-factor* yields mostly a correction of about -10%. Only for low muon energies in the neutral process we can observe an increase of 4% whereas in the charged process, NLO effects of -15% occur at the boundaries of the accessible energy range.

The overall trend of the NLO corrections resembles in a good correspondence the NLO corrections that we have studied for the dijet system. We can therefore conclude that the NLO corrections for the leptonic observables are a mirror image of the NLO corrections for the jet observables due to the recoil against the jet system.

5 Conclusion

This thesis had several goals: First, we wanted to present how a precision study at Next-to-leading order works, how it is set up and what difficulties it involves. Thus, a considerable part of this work presents the computational environment WHIZARD and how it is used at NLO, hoping that this might be an inspiration and even in a certain sense a manual for similar projects.

Then comes of course the actual evaluation of our results. Here, scans over a large range of centre of mass energies are presented to obtain a precise determination of the integrated cross section in different collider setups. Furthermore, differential distributions provide important information on the kinematical details of the process that helps to distinguish signal and background processes and to find possible deviations from the Standard Model.

A comparison with experimental data would thus be the next step in this analysis and permit the validation of our results to assess the correctness of the numerical implementation and use of the Standard Model. LEP2 data of electron-positron collisions exists up to a centre of mass energy of 209 GeV and has given rise to successful comparisons with numerical simulations. Nevertheless, only colliders like the aforementioned ILC provide us with the required precision for a full validation of our computations by producing (more significant) datasets for collision energies that have not been reached by LEP2.

In order to realistically simulate the environment of a high-energy lepton collider, beamstrahlung - the classical electromagnetic radiation from the collimated beam bunches - has to be taken into account. This is easily possible with WHIZARD, but for simplicity and to focus on the NLO corrections has not been done in this thesis. To even further deepen this study, one should also take into account electroweak corrections. There is initial state radiation, which is implemented in WHIZARD in terms of a structure function resumming soft-collinear photons to all orders and hard-collinear photons up to third order. This reduces the total available energy of the collider for a certain fraction of the events (radiative return). For the sake of simplicity, these effects have not been considered in this thesis. Secondly, there are virtual electroweak corrections together with real QED emissions that have to be properly matched to the ISR structure function. The automation of NLO EW

corrections in WHIZARD is work in progress and was not yet available to be included in this thesis.
As fixed-order NLO events are not positive, only weighted events could be generated for the differential distributions. For QCD NLO processes at lepton colliders, WHIZARD offers the possibility to match the (first) hard emissions from the QCD NLO fixed-order calculation with the leading-logarithmic parton-shower approximation of collinear (and soft) emissions to arrive in positive-weight (and possibly unweighted) events. This thesis focused only on the fixed-order part of the NLO QCD calculation. Lastly, a cross check with other MC generators could have been performed to confirm our numerical results, but this was beyond the timeframe of this thesis.

Appendices

A Acknowledgements

A significant part of this thesis originates in the DESY Summer School 2018 in Hamburg where I had the opportunity to work with the theory group. I would thus like to thank the organisers of the Summer Student's programme, Olaf Behnke and Rainer Gehrke, for providing this excellent opportunity. My supervisor during the Summer School as well as for this thesis is Dr. Jürgen Reuter, who deserves my special thanks at this point. Due to the choice of subject which goes beyond a normal bachelor student's curriculum, I have not only learned a lot of important methods of theoretical physics but could also strengthen my decision to pursue a career in the field. Last but not least, I want to thank Pascal Stienemeier and Vincent Rothe, currently PhD-students in Dr. Reuter's group, for providing a great support and helpful insight - your advice often showed me the right ideas on how to tackle the problems that have arisen during this project.

B Further results

B.1 Scan with different renormalisation scale settings

In the fourth section, the running of the strong coupling constant was given by

$$\alpha_s = \alpha_s(m_Z),\ \alpha_s = \alpha_s(m_W).$$

In addition to that, we show the dependence of the integrated cross sections of different renormalisation scales, namely

$$\alpha_s = \alpha_s(m_Z),\ \alpha_s = \alpha_s(0.5m_Z),\ \alpha_s = \alpha_s(2m_Z)$$

and

$$\alpha_s = \alpha_s(m_W),\ \alpha_s = \alpha_s(0.5m_W),\ \alpha_s = \alpha_s(2m_W)$$

in figures 14 and 15.

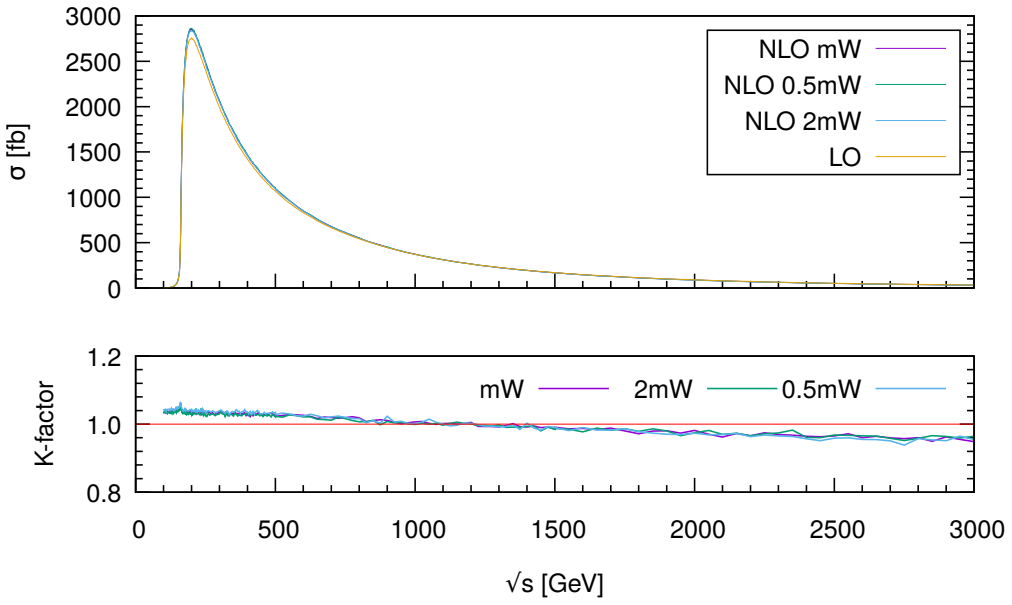


Figure 14: Integrated cross sections for the process $e^+e^- \to \mu\nu jj$ as a function of $\sqrt{s}$; corresponding *K-factor*

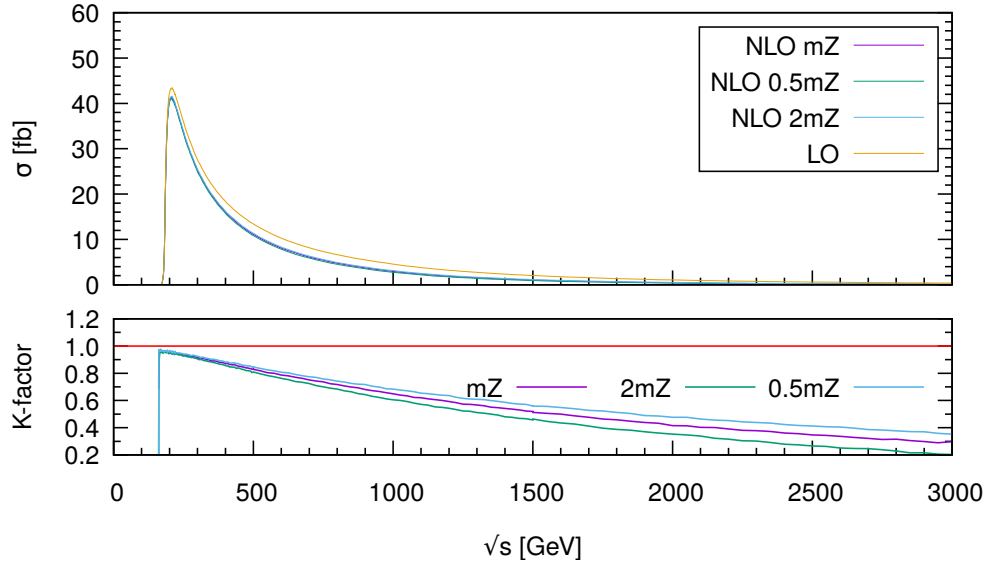
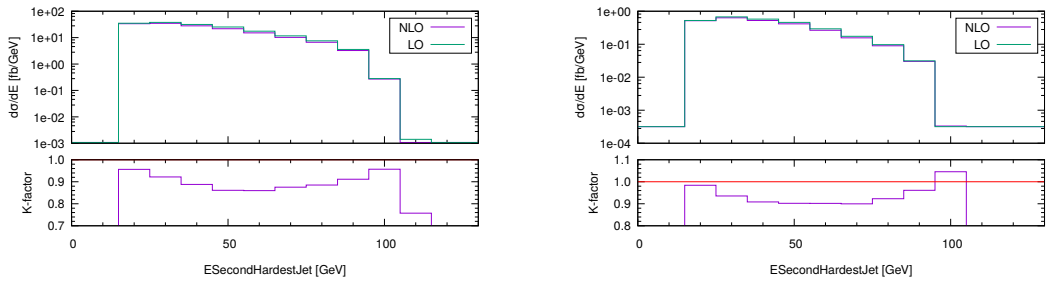


Figure 15: Integrated cross sections for the process $e^+e^- \to \mu\nu jj$ as a function of $\sqrt{s}$; corresponding *K-factor*

B.2 Further differential distributions

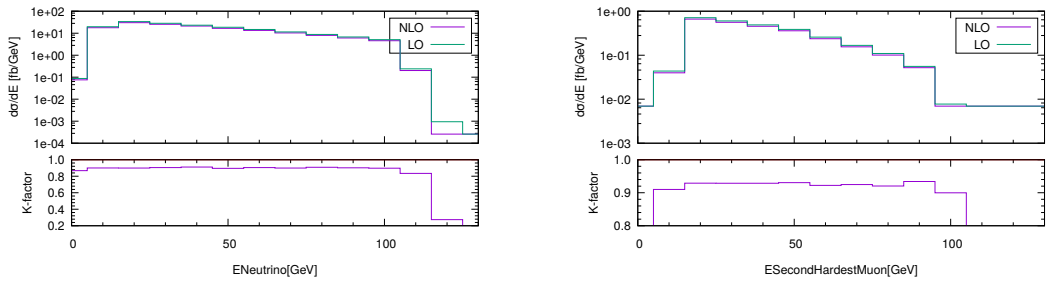
Here, we show further distributions of observables that we have chosen not to present in the main part of this thesis. They may complement the study and provide further insight.



(a) Energy of the second hardest jet in the charged process

(b) Energy of the second hardest jet in the neutral process

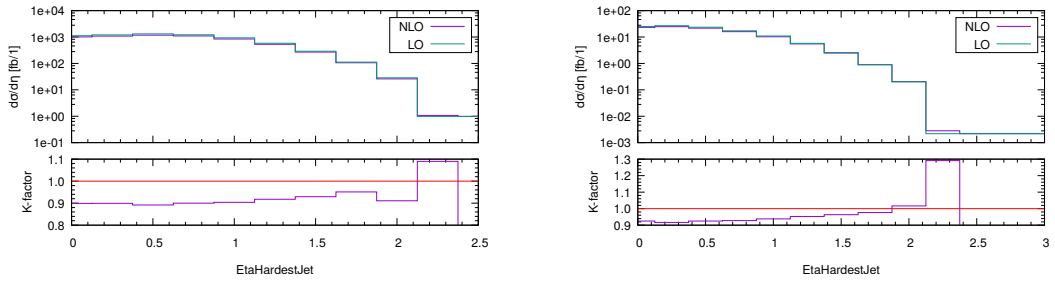
Figure 16: Comparison of the energies of the second hardest jets produced in both processes, respectively



(a) Energy of the neutrino in the charged process

(b) Energy of the second hardest muon in the neutral process

Figure 17: Comparison of the energies of the neutrino produced in the charged process and the second hardest muon produced in the neutral process



(a) Pseudorapidity of the hardest jet in the charged process

(b) Pseudorapidity of the hardest jet in the neutral process

Figure 18: Comparison of the pseudorapidities of the hardest jets produced in both processes

C Steering scripts for WHIZARD and RIVET

To illustrate the numerical predictions of Section 4 and to propose a guide for anyone wishing to reproduce them, this appendix presents the scripts that have been used to steer WHIZARD and, in a second step, to perform the analysis within the framework of RIVET.

C.1 Relevant SINDARIN scripts

The following SINDARIN illustrates the implementation of the process $e^+e^- \to \mu\mu jj$. An integration at NLO QCD is performed using the parameters and settings in listing 1, but only the Born cross section is evaluated. Steering scripts to evaluate the real and virtual contributions at NLO are constructed in complete analogy to the script presented here. To highlight the different parts of a SINDARIN file and their respective contents, the complete file has been split up into three different sections: the model definition and general parameter settings, jet settings and cuts and the process definition, integration and event simulation.

Listing 1: WHIZARD SINDARIN file for the NLO process - model definition and input parameters

```
# set model
model = SM (Complex_Mass_Scheme)
# set aliases
alias lep1 = e2:E2
alias jet1= u:U:d:D:c:C:s:S:g

# set parameters
ms = 0
mc = 0
me = 0
mmu = 0

alpha_power = 4
alphas_power = 0
alphas_nf = 4
?alphas_is_fixed = false
?alphas_from_mz = true
?alphas_from_lambda_qcd = false

?use_vamp_equivalences = false
$loop_me_method = "openloops"
$method = "openloops"
$integration_method = "vamp2"
$rng_method = "rng_stream"
openmp_num_threads = 1
?omega_openmp = false
```

Before the actual process is defined, information about the physics model and its important parameters if they are changed from their default values have to be given. One also has to specify the integration mode: WHIZARD allows to use a serial and a parallelised version of its multi-channel adaptive integrator, called VAMP and VAMP2. Here, the parallelised version VAMP2 has been used. As physics model, the SM in the Complex Mass Scheme has been used. After the definition of an alias to prepare the definition of the actual process, all external fermions have been set massless. The CKM matrix is considered to be the unit matrix within the SM, for a complete CKM matrix the model `SM_CKM` can be chosen. The user may of course choose the parameters that she needs to describe her experimental or numerical setup.
Furthermore, the general settings include specifications about, in this case, the strong coupling constant. It is also important to specify the integration methods and the origin of the NLO matrix elements. As method for the NLO matrix elements, the external library OpenLoops has been specified.

Listing 2: WHIZARD SINDARIN file for the NLO process - cuts and jet settings

```
sqrts = 250 GeV

# beam options
beams = e1, E1

# jets, cluster, subevents
jet_algorithm = kt_algorithm
jet_r = 0.4

cuts =
     let subevt @clustered_jets = cluster [jet1] in
     let subevt @selected_jets = select if abs(Eta) < 4 [@clustered_jets] in
     let subevt @sorted_jets = sort by Pt [@selected_jets] in
     let subevt @first_jet = extract index -1 [@sorted_jets] in
     let subevt @second_jet = extract index -2 [@sorted_jets] in
     count [@selected_jets] >= 2
     and all 80 GeV < M < 100 GeV [@first_jet, @second_jet]
     and all 80 GeV < M < 100 GeV [lep1, lep1]
     and all Pt > 20 GeV [@selected_jets]
     and all abs(Eta) < 4 [lep1]

# set scale
scale = mZ
```

After the model selection and settings of input parameters and methods for matrix elements and integrations, the user now sets the actual physics setup. This involves defining a collision energy and beam properties [12]. A very important part of each

[12]Here one could also account for example for beam polarisation, crossing angles or Initial State Radiation and thus very well approximate a given experimental environment of interest.

SINDARIN are the cuts that define the fiducial phase space in which the process is being studied. The jet clustering is already defined here by specifying a jet-algorithm and the jet-parameter. Finally, a renormalisation scale is set.

Listing 3: WHIZARD SINDARIN file for the NLO process - process definition and integration and event simulation

```
# process setup (@NLO, Born)
  process ee2lljj_NLO = e1, E1 => lep1, lep1, jet1, jet1
  {nlo_calculation = born}

# Integration/event generation
?fixed_order_nlo_events = true
integrate (ee2lljj_NLO){iterations = 5:50000:"gw", 3:100000}
n_events = 1000000
checkpoint = n_events / 50

?unweighted = false
?negative_weights = true
sample_format = hepmc
?hepmc_output_cross_section = true

# set sample normalisation
$sample_normalization = "sigma"

simulate(ee2lljj_NLO)
```

In the final part, we set the physical process that we want to analyse and use the particle containers that have been defined in the very beginning. Finally, there are SINDARIN commands to perform process integration and simulation, for which several technical parameters like detailed adaptation and iterations, event file format and event weight specifications have been set. Output is written in the `HepMc` format.

If the user wants to perform a scan over different CM energies, this can be implemented with different step sizes by the following statement:

Listing 4: WHIZARD SINDARIN file for the NLO process - Scan over CM energies

```
scan sqrts =(
     (150.0 GeV => 160.0 GeV /+ 2 GeV),
     (160.0 GeV => 220.0 GeV /+ 1 GeV),
     (220.0 GeV => 500.0 GeV /+ 5 GeV),
     (500.0 GeV => 1500.0 GeV /+ 25 GeV),
     (1500.0 GeV => 3000.0 GeV /+ 50 GeV))
     {
     integrate (ee2lljj_NLO) {iterations = 5:50000:"gw",
         3:100000 mult_call_real = 8 }
     }
```

C.2 Analysis in RIVET

According to the workflow described in Section 2, once the events are generated the analysis is done with the HEP programme RIVET. It reads the generated events in the `HepMc` format and offers a broad range of analysis tools. The user can either analyse her results with an existing analysis or, in case none of the established analyses matches the desired conditions, write her own analysis from scratch. This is what has been done here, where we show as an example the NLO analysis for the process $e^+e^- \to \mu\mu jj$. Analyses at LO and for the process $e^+e^- \to \mu\nu jj$ are similar to the analysis presented here.

A single RIVET analysis is a C++ class and consists (mainly) of three methods, namely `init()`, `analyze()` and `finalize()`. We show here for the sake of brevity an abbreviated version of our script that does not include every histogram that has been studied but which is otherwise semantically identical to the full analysis.

Listing 5: RIVET - init()

```
#include "Rivet/Analysis.hh"
#include "Rivet/Projections/FinalState.hh"
#include "Rivet/Projections/FastJets.hh"
#include "Rivet/Projections/IdentifiedFinalState.hh"
#include "Rivet/Projections/VetoedFinalState.hh"
#include <cmath>

namespace Rivet {
  class EE2LLJJ_ANALYSIS_NLO : public Analysis{
    #include "../NLOHisto1D.cc"

  public:
    // constructor
   DEFAULT_RIVET_ANALYSIS_CTOR(EE2LLJJ_ANALYSIS_NLO);

    double mjj;
    double mll;
    double js;
    double jFull;

    void init() {

      // identify final state
      const IdentifiedFinalState mu(PID::MUON);
      addProjection(mu, "mu");
      const IdentifiedFinalState mubar(-PID::MUON);
      addProjection(mubar, "mubar");

      // veto the final states that don't go into the jets
      VetoedFinalState vfs(Cuts::abseta<4);
      vfs.addVetoOnThisFinalState(mu);
      vfs.addVetoOnThisFinalState(mubar);

      // FastJet settings
      const double R=0.4;
      fastjet::JetDefinition ee(fastjet::kt_algorithm, R);
      FastJets jets(vfs, ee);
```

```
    addProjection(jets, "Jets");

    // book histograms
    // transverse momentum
    _hist_HardestJetPt = bookNLOHisto1D("HardestJetPt", 25, 0., 250.);

    // total energy
    _hist_HardestJetE = bookNLOHisto1D("HardestJetE", 25, 0., 250.);

    // angle with beam axis
    _hist_HardestJetTheta =
        bookNLOHisto1D("HardestJetTheta", 15, 0., 1.5);

    // pseudorapidity
    _hist_HardestJetEta = bookNLOHisto1D("HardestJetEta", 20, 0., 5.);

    // invariant dijet mass
    _hist_InvariantDijetMass = bookNLOHisto1D("InvariantDijetMass", 20, 70., 110.);

    eventCounter = 0;
    vetoCounter = 0;
  }
```

We first have to identify the final state of the interaction in question, which RIVET does automatically using the right projection. Knowing that we want to apply jet clustering methods, the muons are subject to a veto, but besides that every coloured particle in the final state can possibly be clustered in a jet. At NLO this also leaves us with the possibility of clustering a real gluon emission.
All the histograms that we want to use do have to be defined in the `init()` method.

Listing 6: RIVET - analyze()

```
// perform the per-event analysis
void analyze(const Event& event){

    // declare weight
    double weight = event.weight();

    const FastJets& fastjets = applyProjection<FastJets>(event, "Jets");
    const PseudoJets pseudo_jets = fastjets.pseudoJetsByPt();

    // initialize object jets
    Jets jets;
    foreach (const PseudoJet & pseudo_jet, pseudo_jets){jets.push_back (pseudo_jet); }

    //initialize objects particleVector mu, mubar and m
    ParticleVector mu_pt =applyProjection<IdentifiedFinalState>(event, "mu").particlesByPt();
    ParticleVector mubar_pt =applyProjection<IdentifiedFinalState>(event, "mubar").particlesByPt();
    ParticleVector m;
    foreach (const Particle & p, mu_pt) { m.push_back(p); }
    foreach (const Particle & p, mubar_pt) { m.push_back(p); }
    sort(m.begin(),m.end(),cmpMomByPt);

    // fill histograms, apply cuts
    if (jets.size() >= 2 &&
    80 < (jets[0].momentum()+jets[1].momentum()).mass() &&
    (jets[0].momentum()+jets[1].momentum()).mass() < 100 &&
    80 < (m[0].momentum()+m[1].momentum()).mass() &&
    (m[0].momentum()+m[1].momentum()).mass() < 100 &&
    jets[0].pt() > 20 && jets[1].pt() > 20)
      {
  // Jets
```

```
    // transverse momentum
    _hist_HardestJetPt->fill(jets[0].pt(), weight);

    // total energy
    _hist_HardestJetE->fill(jets[0].E(), weight);

    // angle with beam axis
    _hist_HardestJetTheta->fill(abs(std::cos(jets[0].theta())), weight);

    // pseudorapidity
    _hist_HardestJetEta->fill(jets[0].abseta(), weight);

    // invariant dijet mass
    const double M_jj = (jets[0].momentum() + jets[1].momentum()).mass();
    _hist_InvariantDijetMass->fill(M_jj, weight);

    eventCounter++;
    mll = (m[0].momentum()+m[1].momentum()).mass();
    mjj = (jets[0].momentum()+jets[1].momentum()).mass();
    js = jets.size();

    }

    else {
      js = jets.size();
      vetoCounter++;
      vetoEvent;
    }

}
```

In the `analyze()`-method the actual physics analysis is performed. We initialise the jet-object `jets` and `particlevectors` for each of the partonic muons, into which the detected final states are then projected. Through a sorting by the transverse momentum and the merging of the sorted muon-vectors we have access to the hardest and second hardest jet and muon, respectively, by addressing the first or the second entry of the corresponding vector object.
Physical cuts are simply implemented by an if-condition since in the analysis by RIVET, each event is treated separately and can thus either pass the if-condition or be filtered out without being correlated with any other event. Nevertheless, the cuts have been constructed to let all of the events pass, we utilise the same cuts as in WHIZARD.
The histograms are then filled with the respective properties of the muons and jets in every single event.

Listing 7: RIVET - finalize()

```
    // finish analysis
    void finalize(){

      // scale histograms to cross section
      // factor 1000 to scale from pb to fb

      double scale_factor = crossSection()*1000/ sumOfWeights();
```

```
      scale(_hist_HardestJetPt, scale_factor);
      scale(_hist_HardestJetE, scale_factor);
      scale(_hist_HardestJetEta, scale_factor);
      scale(_hist_HardestJetTheta, scale_factor);
      scale(_hist_InvariantDijetMass, scale_factor);

      cout << "Number of total events: " << eventCounter << endl;
      cout << "Numer of vetoed events: " << vetoCounter << endl;
      cout << "jets.size() = " << js << endl;

    }
    Histo1DPtr _hist_HardestJetPt
    Histo1DPtr _hist_HardestJetE
    Histo1DPtr _hist_HardestJetEta
    Histo1DPtr _hist_HardestJetTheta
    Histo1DPtr _hist_InvariantDijetMass

    private:
      int vetoCounter, eventCounter;

  };

  DECLARE_RIVET_PLUGIN(EE2LLJJ_ANALYSIS_NLO);
}
```

To finish the analysis, the histograms have to be scaled to account for the proper normalisation of the total NLO cross section, as explained in Section 3. Finally, the analysis has to be declared within the RIVET framework to be accessible to the user.

References

[1] Howard Baer, Tim Barklow, Keisuke Fujii, Yuanning Gao, Andre Hoang, Shinya Kanemura, Jenny List, Heather E. Logan, Andrei Nomerotski, Maxim Perelstein, et al. The International Linear Collider Technical Design Report - Volume 2: Physics. 2013.

[2] Bijan Chokoufe. *Scrutinizing the Top Quark at Lepton Colliders with Higher Orders*. Dr., Universität Hamburg, Hamburg, 2017. Universität Hamburg, Diss., 2017.

[3] Bijan Chokoufé Nejad, Wolfgang Kilian, Jonas M. Lindert, Stefano Pozzorini, Jürgen Reuter, and Christian Weiss. NLO QCD predictions for off-shell $t\bar{t}$ and $t\bar{t}H$ production and decay at a linear collider. *JHEP*, 12:075, 2016.

[4] Wolfgang Kilian, Thorsten Ohl, and Jurgen Reuter. WHIZARD: Simulating Multi-Particle Processes at LHC and ILC. *Eur. Phys. J.*, C71:1742, 2011.

[5] Mauro Moretti, Thorsten Ohl, and Jurgen Reuter. O'Mega: An Optimizing matrix element generator. pages 1981–2009, 2001.

[6] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012.

[7] Joshua Ellis. TikZ-Feynman: Feynman diagrams with TikZ. *Comput. Phys. Commun.*, 210:103–123, 2017.

[8] Stephen J. Blundell Tom Lancaster. *Quantum Field Theory for the Gifted Amateur*. 2014.

[9] David Griffiths. *Introduction to Elementary Particles*. 2010.

[10] T. Kinoshita. Mass singularities of Feynman amplitudes. *J. Math. Phys.*, 3:650–677, 1962.

[11] T. D. Lee and M. Nauenberg. Degenerate systems and mass singularities. *Phys. Rev.*, 133:B1549–B1562, Mar 1964.

[12] Christian Weiss. *Top quark physics as a prime application of automated higher-order corrections*. Dr., Universität Hamburg, Hamburg, 2017. Universität Hamburg, Diss., 2017.

[13] S. Frixione, Z. Kunszt, and A. Signer. Three jet cross-sections to next-to-leading order. *Nucl. Phys.*, B467:399–442, 1996.

[14] Rikkert Frederix, Stefano Frixione, Fabio Maltoni, and Tim Stelzer. Automation of next-to-leading order computations in QCD: The FKS subtraction. *JHEP*, 10:003, 2009.

[15] Juergen Reuter et al Wolfgang Kilian, Thorsten Ohl. Whizard 2.6: A generic monte-carlo integration and event generation package for multi-particle processes; manual, 2018.

[16] Wolfgang Kilian, S. Brass, T. Ohl, J. Reuter, V. Rothe, P. Stienemeier, and M. Utsch. New Developments in WHIZARD Version 2.6. 2018.

[17] C. Weiss, Bijan Chokoufe Nejad, Wolfgang Kilian, and Jurgen Reuter. Automated NLO QCD Corrections with WHIZARD. *PoS*, EPS-HEP2015:466, 2015.

[18] J. Reuter, F. Bach, B. Chokoufe, W. Kilian, T. Ohl, M. Sekulla, and C. Weiss. Modern Particle Physics Event Generation with WHIZARD. *J. Phys. Conf. Ser.*, 608(1):012063, 2015.

[19] Wolfgang Kilian Thorsten Ohl, Juergen Reuter. O'mega: Optimal monte-carlo event generation amplitudes, 2018.

[20] Thorsten Ohl. Vegas revisited: Adaptive Monte Carlo integration beyond factorization. *Comput. Phys. Commun.*, 120:13–19, 1999.

[21] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati. Recursive generation of one-loop amplitudes in the Standard Model. *JHEP*, 04:037, 2013.

[22] Stefano Actis, Ansgar Denner, Lars Hofer, Jean-Nicolas Lang, Andreas Scharf, and Sandro Uccirati. RECOLA: REcursive Computation of One-Loop Amplitudes. *Comput. Phys. Commun.*, 214:140–173, 2017.

[23] Lynn Garren et al Matt Dobbs, Jørgen Beck Hansen. Hepmc 2: a c++ event record for monte carlo generators, 2018.

[24] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[25] Andy Buckley, Jonathan Butterworth, Leif Lonnblad, David Grellscheid, Hendrik Hoeth, James Monk, Holger Schulz, and Frank Siegert. Rivet user manual. *Comput. Phys. Commun.*, 184:2803–2819, 2013.

[26] Pascal Stienemeier. Precise predictions for same-sign vector boson scattering at the lhc. Master's thesis, Universität Hamburg, DESY Theory group, 2018.

[27] Stefan Weinzierl. Introduction to Monte Carlo methods. 2000.

[28] S. Catani and M. H. Seymour. A General algorithm for calculating jet cross-sections in NLO QCD. *Nucl. Phys.*, B485:291–419, 1997. [Erratum: Nucl. Phys.B510,503(1998)].

[29] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett.*, B641:57–61, 2006.

[30] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[31] W. Kluge. Initial State Radiation: A Success story. *Nucl. Phys. Proc. Suppl.*, 181-182:280–285, 2008.

[32] M. et al. Tanabashi. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

Declaration of authorship

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, den