Published by Institute of Physics Publishing for SISSA/ISAS
Received: *March 1, 2004*
Revised: *April 21, 2004*
Accepted: *May 4, 2004*

Effective K-factors for $gg \to H \to WW \to \ell\nu\ell\nu$ at the LHC

Giovanna Davatz, Guenther Dissertori, Michael Dittmar and Felicitas Pauss

Institute for Particle Physics (IPP), ETH Zurich
CH-8093 Zurich, Switzerland
E-mail: giovanna.davatz@cern.ch, guenther.dissertori@cern.ch, michael.dittmar@cern.ch, felicitas.pauss@cern.ch

Massimiliano Grazzini

Department of Physics, CERN, Theory Division
CH–1211 Geneva 23, Switzerland
E-mail: massimiliano.grazzini@cern.ch

Abstract: A simulation of the search for the Standard Model Higgs boson at the LHC, in the channel gg $\to$ H $\to$ WW $\to \ell\nu\ell\nu$, is described. Higher-order QCD corrections are taken into account by using a reweighting procedure, which allows us to combine event rates obtained with the PYTHIA Monte Carlo program with the most up-to-date theoretical predictions for the transverse-momentum spectra of the Higgs signal and its corresponding WW background. With this method the discovery potential for Higgs masses between 140 and 180 GeV is recalculated and the potential statistical significance of this channel is found to increase considerably. For a Higgs mass of 165 GeV a signal-to-background ratio of almost 2:1 can be obtained. A statistical significance of five standard deviations might already be achieved with an integrated luminosity close to $0.4\,\mathrm{fb}^{-1}$. Using this approach, an experimental effective K-factor of about 2.04 is obtained for the considered Higgs signature, which is only about 15% smaller than the theoretical inclusive K-factor.

Keywords: QCD, Higgs Physics, Hadronic Colliders.

JHEP05(2004)009

Contents

1. Introduction

The design of the ATLAS and CMS experiments at the LHC was guided by the requirement to have high sensitivity for discovering the Higgs boson within the full mass range between 100 GeV and approximately 1 TeV [1].

In recent years a large effort has gone into accurate calculations of many Higgs signal and background cross sections, which in most cases are now known to next-to-leading order (NLO) accuracy [2]. For the dominant Standard Model (SM) Higgs production mechanism, gluon-gluon fusion, and as far as the total cross section is concerned, even next-to-next-to-leading order (NNLO) [3] QCD corrections have been computed[1].

By contrast, and despite the recent considerable progress in improving Monte Carlo (MC) event generators, a complete MC program where the same higher-order QCD corrections are included does not exist yet.

QCD corrections to signal and background cross sections are thus either ignored or taken into account in a very naive approach. Normally, results obtained with a standard MC program are simply scaled with the so-called inclusive K-factor. Although the K-factor should not be considered as a physical quantity, this approach is believed to provide a reasonable simulation environment, which allows us to study the acceptance for many signatures. In the context of Higgs searches, an example is the decay of Higgs bosons into four leptons, $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$. This signature is not really sensitive to additional jet activities, and it is usually assumed that the search sensitivity depends mainly on signal and background cross sections, the particular detector model and the used selection criteria. Consequently, a simple scaling of signal and background with the inclusive K-factor, according to the most accurate theoretical prediction, should give reasonable results. This assumption has recently been confirmed in a more quantitative study for the decay $\mathrm{H} \to \mathrm{ZZ} \to 4\ell$ [4].

[1]More precisely, the NNLO calculation has been performed in the large m_{top} approximation.

However, the situation is different if, in addition to some particle identification, the event kinematics has to be exploited to separate the signal from backgrounds. A typical example is the proposed Higgs search in the mass range 155–180 GeV [5], where the identification of the decay $H \to W^+W^- \to \ell^+\nu\ell'^-\bar{\nu}$ requires a jet veto in order to remove $t\bar{t}$ events. In addition, other cuts are required, which exploit the spin correlations between the W bosons and the resulting transverse momentum (p_T) spectra of the charged leptons. These cuts are particularly sensitive when the Higgs mass is close to $2M_W$ and if the Higgs boson is produced with small transverse momentum p_T^H. Consequently, we cannot expect that the inclusive K-factors can be used directly and a more careful investigation is needed.

The effects of a jet veto on the K-factor have been studied in QCD perturbation theory up to NNLO in ref. [6]. These results show that the impact of higher-order QCD corrections is reduced if a jet veto is applied. This study demonstrated clearly that the simple K-factor scaling cannot be used in this case. It is therefore important to determine the effective "experimental" K-factor in combination with a detailed simulation of cuts, for both signal and background.

In this paper we reconsider the $gg \to H \to W^+W^- \to \ell^+\nu\ell'^-\bar{\nu}$ channel by using a reweighting procedure of events generated with the PYTHIA [7] MC program. The reweighting is performed according to the most up-to-date theoretical predictions for p_T spectra of the Higgs [8] and the non-resonant WW production, which is the most important background. This method allows us to include higher-order QCD corrections in combination with experimental selection criteria to a good approximation.

The paper is organized as follows. In section 2 the reweighting technique is introduced. In section 3 we use PYTHIA to determine the detection efficiency as a function of p_T^H, using selection criteria based on the ideas given in [5]. These selection criteria are relatively robust and it can be expected that only minor modifications are needed in more accurate detector simulations. Next, the p_T^H spectrum is reweighted so as to agree with the spectrum obtained in ref. [8]. A similar procedure is applied to the main background, the non-resonant WW production (section 5). Finally, in section 6, the effective experimental K-factor is calculated for the signal and the background and the possible statistical signal significance is obtained.

2. K-factors and the reweighting technique

Generally speaking, the number of events for a given integrated luminosity and a particular process, as computed at NLO, is given by

$$\frac{N}{\mathcal{L}} = \sigma_{NLO} = K_I \, \sigma_{LO} \tag{2.1}$$

where K_I is called the inclusive K-factor, which is thus defined as the ratio of the inclusive (total) NLO and LO cross sections. At LO, the produced particle (system) X (we have in mind X = H, WW, ...) has no transverse momentum, while in (N)NLO additional jets lead to a non-vanishing p_T spectrum. Suppose that a jet veto is applied. Depending on the jet-detection capabilities of the hypothetical experiment one can define an effective parton-level K-factor from the ratio of accepted events at (N)NLO and at LO (for this example

JHEP05(2004)009

all LO events would be accepted). This effective theoretical (parton-level) K-factor is in general smaller than the inclusive K-factor [6].

The definition of effective experimental K-factors is different and somewhat more complicated, because most experimental simulations are based on LO cross sections, supplemented with parton showering through the standard MC programs. Thanks to additional (mostly soft) jets, an approximate p_{T} spectrum of the particular final state is generated. However, this spectrum is at best an approximation, since only soft and collinear radiation from the primary parton subprocess can be generated correctly.[2]

On the other hand, if the (N)NLO calculation has been performed differentially as a function of a kinematical variable ξ, the above equation can then be rewritten as

$$\frac{N}{\mathcal{L}} = \int \frac{d\sigma_{\mathrm{NLO}}}{d\xi} d\xi = \int K(\xi) \frac{d\sigma_{\mathrm{MC-LO}}}{d\xi} d\xi \,, \tag{2.2}$$

where the integral goes over the complete possible range of ξ, and the ξ-dependent K-factor is defined as

$$K(\xi) = \frac{d\sigma_{\mathrm{NLO}}(\xi)/d\xi}{d\sigma_{\mathrm{MC-LO}}(\xi)/d\xi} \,. \tag{2.3}$$

In our case PYTHIA is employed as a leading-order MC program (MC-LO). It is now possible to study the effects of selection cuts that depend on ξ. In this paper we will be concerned with the case $\xi = p_{\mathrm{T}}$, p_{T} being the transverse momentum of the Higgs boson (signal) or of the WW pair (background). The efficiency can be calculated as a function of p_{T} using the PYTHIA program. After all cuts are applied we can also define an inclusive average efficiency to detect process X. The total number of accepted events within PYTHIA coincides with the sum of the events that were accepted differentially over the entire $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum.

However, if the p_{T}-dependent K-factor defined in eq. (2.3) is applied to correct (reweight) the spectrum, the total number of accepted events can only be obtained from the sum of the differentially accepted weighted events. This new number of accepted events at (N)NLO can now be compared with the one accepted in the unweighted PYTHIA simulation and their ratio defines the effective experimental K-factor K_{eff}.

3. Higgs signal selection

The PYTHIA MC program is used for the simulation of the signal and the different types of relevant backgrounds. The strategy to separate signal events of the type $\mathrm{pp} \to \mathrm{H} \to \mathrm{W^+W^-} \to \ell^+\nu\ell'^-\bar{\nu}$ from the various backgrounds is based on the ideas described in ref. [5]. The signal selection proceeds in two steps.

First, events that contain exactly two isolated and oppositely charged high-p_{T} leptons (electrons or muons) are selected. These leptons originate mainly from the decays of W bosons. They should not be back-to-back in the plane transverse to the beam and their invariant mass should be considerably smaller than the Z mass. Furthermore, a

[2] QCD radiation at large transverse momenta is strongly suppressed and can be accounted for either with matrix-element corrections [9, 10] or by matching the full NLO calculation to the parton shower [11].

substantial missing transverse momentum is required. Essentially, these criteria select only events that contain a pair of W's, these being either signal events or backgrounds from non-resonant WW production $q\bar{q} \to WW \to \ell^+\nu\ell'^-\bar{\nu}$, from $t\bar{t} \to WbWb \to \ell^+\nu\ell'^-\bar{\nu}bb$ and $Wtb \to WWb \to \ell^+\nu\ell'^-\bar{\nu}b$.

Following this preselection, the criteria for the second step further separate the Higgs signal events from backgrounds using: (1) the somewhat shorter rapidity plateau for signal events, (2) the jet activity in signal events reduced with respect to the background from $t\bar{t}$ production, (3) the effects of spin correlations and the mass of the resonant and non-resonant WW system, resulting in a small opening angle for the lepton-lepton system and a somewhat mass-dependent characteristic p_T spectrum of the charged leptons.

In detail the following cuts are applied:

1. The event should contain two leptons, electrons or muons, with opposite charge, each with a minimal p_T of 20 GeV and a pseudorapidity $|\eta|$ smaller than 2.

2. In order to have isolated leptons, it is required that the transverse energy sum from detectable particles (defined as "stable" charged or neutral particles with a p_T larger than 1 GeV), found inside a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2)} < 0.5$ around the lepton direction, be smaller than 10% of the lepton energy. Furthermore the invariant mass of all particles within the cone should be smaller than 2 GeV, and at most one additional detectable particle inside a cone of $\Delta R < 0.15$ is allowed.

3. The dilepton mass, $m_{\ell\ell}$, has to be smaller than 80 GeV.

4. The missing p_T of the event, required to balance the p_T vector sum of the two leptons, should be larger than 20 GeV.

5. The two leptons should not be back-to-back in the plane transverse to the beam direction. The opening angle between the two leptons in this plane is required to be smaller than 135°.

Dilepton events, originating from the decays of W and Z bosons, are selected with criteria 1 and 2. Lepton pairs that originate from the inclusive production of $Z \to \ell\ell(\gamma)$, including Z decays to τ leptons, are mostly removed with criteria 3–5.

Starting with this initial set of requirements, the following criteria exploit the differences between Higgs events and the so-called "irreducible" background from continuum production of $pp \to W^+W^-X$ events.

6. The opening angle ϕ between the two charged leptons in the plane transverse to the beam should be smaller than 45° and the invariant mass of the lepton pair should be smaller than 35 GeV.[3]

7. For jets, which are formed with a cone algorithm, a minimum transverse momentum of 20 GeV is required. Events with a jet of p_T^{jet} larger than a chosen value p_{Tmin}^{jet} and a pseudorapidity $|\eta^{jet}|$ of less than 4.5 are removed.

[3] A minimal angle (or mass) of 10° (10 GeV) might be needed in order to reject badly measured $\Upsilon \to e^+e^-(\mu^+\mu^-)$ decays. Such a cut would not change the signal efficiency in any significant way.

JHEP05(2004)009

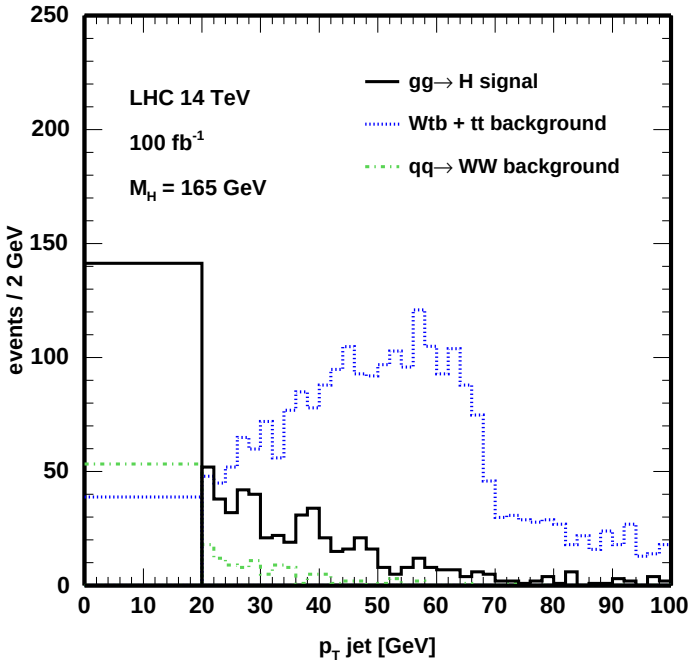

Figure 1: Number of accepted signal and background events as a function of the $p_{\mathrm{T}}^{\mathrm{jet}}$ of the leading jet. The simulated Higgs mass is 165 GeV. All cuts except the jet veto are applied. Events without a reconstructed jet are evenly distributed over the first 10 bins, since only those jets are counted which have a reconstructed $p_{\mathrm{T}}^{\mathrm{jet}}$ larger than 20 GeV.

8. Finally, the p_{T} spectrum of the two charged leptons is exploited. For this, the two leptons are classified according to their p_{T} into p_{Tmin}^{ℓ} and p_{Tmax}^{ℓ}. It is found that the p_{Tmax}^{ℓ} and p_{Tmin}^{ℓ} distributions show a jacobian-peak like structure for the signal, which also depends on the simulated Higgs mass. In the case of a Higgs mass close to 165 GeV, p_{Tmax}^{ℓ} should be between 35 and 50 GeV, while the p_{Tmin}^{ℓ} should be larger than 25 GeV. For $M_{\mathrm{H}} = 140$ GeV the p_{T} of the leptons should be larger than 20 GeV, whereas for $M_{\mathrm{H}} = 180$ GeV, the lepton p_{Tmax}^{ℓ} has to be larger than 45 GeV and p_{Tmin}^{ℓ} larger than 25 GeV.

Figure 1 shows the $p_{\mathrm{T}}^{\mathrm{jet}}$ of the hardest (leading) reconstructed jet for signal and background events from PYTHIA after applying the cuts 1 to 6 and cut 8. The events with no reconstructed jet are equally distributed over the 10 bins between 0 and 20 GeV. As can be seen from figure 1, the particular choice of the cut value for the jet transverse momentum does not seem to be very critical for the observation of a Higgs signal, but it needs to be studied in detail if a precision cross section measurement is envisaged, or if the search is to be extended to much lower values of the Higgs mass with smaller signal-to-background ratios.

Applying all selection criteria, including the optimized lepton p_{T} cuts, an accepted cross section of 15.9 fb for the Higgs signal with a mass of 165 GeV can be expected, above a background of 12.3 fb. In principle, to estimate the complete signal rate, the contribution from vector boson fusion, which, for $M_{\mathrm{H}} = 165$ GeV, is about 0.7 fb, should be included. For the purpose of this paper, this contribution will be neglected in the following.

JHEP05(2004)009

Process	$\sigma_{\text{LO PYTHIA}} \times \text{BR}^2$ [pb]	Number of events, $\mathcal{L} = 5\,\text{fb}^{-1}$		
		All cuts except cuts on p_{T} lepton	$35 < p^{\ell}_{\text{Tmax}} < 50$	$25 < p^{\ell}_{\text{Tmin}}$
$gg \to H \to WW$	1.06	176	110	80
$qq \to WW$	7.38	243	83	30
$t\bar{t}$	52	47	15	5
Wtb	5.2	87	46	26

Table 1: Cross sections obtained with PYTHIA for signal and backgrounds, and expected number of events after applying various cuts, for $M_{\text{H}} = 165\,\text{GeV}$ and an integrated luminosity of 5 fb^{-1}. In all cases the leptonic branching ratios of both W bosons, $\text{W} \to \ell\nu$ with $\ell = \text{e}, \mu, \tau$ are included.

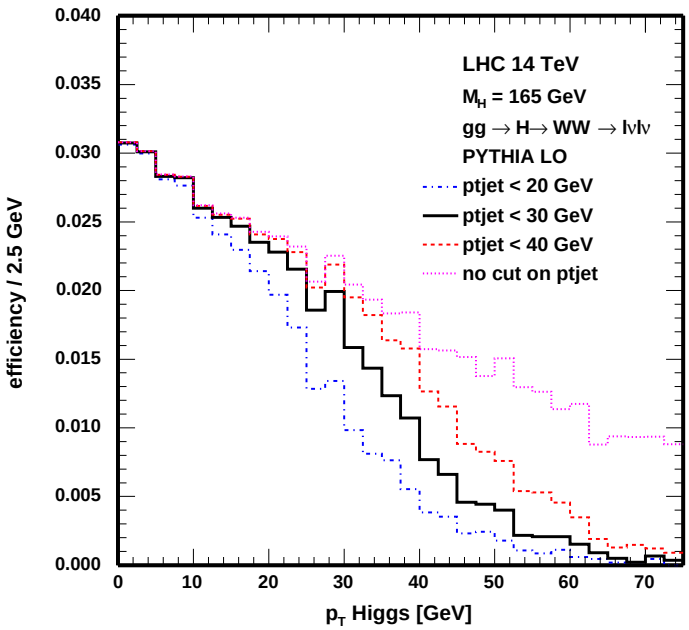

Figure 2: Signal selection efficiency as a function of the Higgs transverse momentum, for a Higgs mass of 165 GeV and three different jet veto cuts. For completeness, the efficiency curve for all cuts, excluding the jet veto, is also shown.

The LO cross sections indicate that for Higgs masses near 165 GeV a statistically significant signal should be observable with a luminosity of slightly more than $1\,\text{fb}^{-1}$. The relevant PYTHIA cross sections for the signal and backgrounds as well as some event rates expected for a luminosity of $5\,\text{fb}^{-1}$ are given in table 1.

As already mentioned in the introduction, the analysis mainly selects signal events with low p^{H}_{T}. The efficiency to detect a Higgs boson with these selection criteria, defined as the ratio of all accepted over all generated events, can be studied as a function of the generated transverse momentum of the Higgs. The results for different jet veto cuts ($p^{\text{jet}}_{\text{Tmin}}$ = 20, 30 and 40 GeV) are shown in figure 2. As expected, Higgs events with large p^{H}_{T} are almost always rejected with the proposed criteria, and the efficiency drops quite quickly as p^{H}_{T} reaches the value of the jet veto $p^{\text{jet}}_{\text{Tmin}}$.

JHEP05(2004)009

4. Signal and background $p_{\rm T}$ spectra

At the lowest order in QCD perturbation theory, the Higgs boson is produced with vanishing transverse momentum. Thus, in order to generate a non-vanishing $p_{\rm T}^{\rm H}$, NLO corrections should be considered. At transverse momenta $p_{\rm T}^{\rm H}$ of the order of $M_{\rm H}$ the perturbative expansion is reliable, being controlled by a small expansion parameter $\alpha_{\rm S}(M_{\rm H}^2)$. NNLO corrections have been computed, in the large-$m_{\rm top}$ approximation, first numerically [12] and later analytically [13, 14].[4]

When $p_{\rm T}^{\rm H}$ is much smaller than $M_{\rm H}$ the convergence of the fixed-order expansion is spoiled, as the coefficients of the perturbative series in $\alpha_{\rm S}(M_{\rm H}^2)$ are enhanced by powers of large logarithmic terms, $\ln^m(M_{\rm H}/p_{\rm T}^{\rm H})$. These terms must be resummed to all orders to give a perturbative prediction valid down to small $p_{\rm T}^{\rm H}$.

In ref. [8] the resummation of these logarithmic contributions is performed up to next-to-next-to-leading logarithmic accuracy (NNLL) and matched to the fixed-order (NNLO) result valid at large $p_{\rm T}^{\rm H}$, in order to avoid double counting. We thus obtain a prediction that is always as good as the fixed-order prediction, but much better in the small-$p_{\rm T}^{\rm H}$ region. Note that the resummation of logarithmically enhanced contributions is approximately performed by standard parton shower MC programs, which should thus account for the shape of the distribution in the small-$p_{\rm T}^{\rm H}$ region.

In the following we will use results obtained with the numerical program of ref. [8] at NNLL+NNLO accuracy. The formalism of ref. [8] implements a unitarity constraint, such that the integral of the distribution is the total NNLO cross section [3]. At variance with the calculation of ref. [8], the $p_{\rm T}^{\rm H}$ spectrum is here obtained using the MRST2002 NNLO [15] parton distributions and $\alpha_{\rm S}$ computed in the three-loop approximation.

The expected Higgs $p_{\rm T}^{\rm H}$ spectra for $M_{\rm H} = 165\,\text{GeV}$ from PYTHIA and from the resummed calculation are shown in figure 3. It can be seen that PYTHIA provides a softer $p_{\rm T}^{\rm H}$ spectrum and differs from the perturbative calculation over the whole range of $p_{\rm T}^{\rm H}$. The ratio between the two spectra can be used to define the $p_{\rm T}^{\rm H}$-dependent K-factor, $K(p_{\rm T}^{\rm H})$ (see section 2).

This $p_{\rm T}^{\rm H}$-dependent K-factor rises from approximately 1 at small $p_{\rm T}^{\rm H}$ to 3 at a $p_{\rm T}^{\rm H}$ around 50 GeV, and then decreases again to about 2.2 at a $p_{\rm T}^{\rm H}$ of 200 GeV, as shown in figure 4. Note that at relatively large transverse momenta, the PYTHIA event generator is supplemented with hard matrix-element corrections [10], thus explaining the approximately flat K-factor at large $p_{\rm T}^{\rm H}$, but the normalization is still fixed to LO.

The $p_{\rm T}^{\rm H}$-dependent K-factor can be used to apply a weight to events generated with PYTHIA. The idea of the reweighting procedure is based on the assumption that the kinematics of Higgs events for a particular $p_{\rm T}^{\rm H}$ is reasonably well described by PYTHIA and that the efficiency of the cuts is computed correctly. Since the $p_{\rm T}$ spectrum is generated by multiple radiation from the incoming partons, the rapid variation of the K-factor for $p_{\rm T}^{\rm H} \lesssim 40\,\text{GeV}$ in figure 4 could suggest an improper treatment of the effect of a jet veto in PYTHIA. In order to check the reliability of our reweighting procedure we have compared

[4] Contrary to the convention adopted in refs. [12, 13, 14], we use here the classification of perturbative orders based on the total cross section: the $p_{\rm T}$ spectrum starts at NLO.

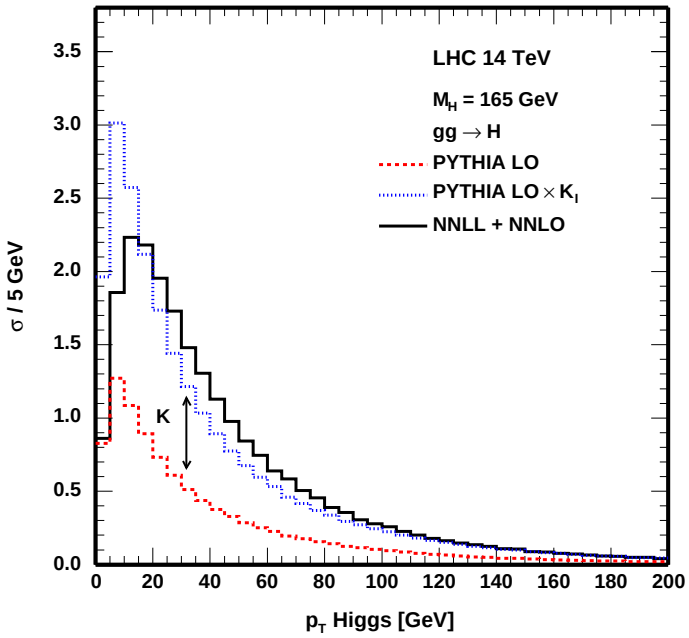

Figure 3: The Higgs production cross section for gg $\rightarrow$ H, as a function of the Higgs transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$, for a Higgs mass of 165 GeV, obtained with PYTHIA and with the NNLL+NNLO calculation. The spectrum from PYTHIA rescaled with the inclusive K-factor is also shown for comparison.

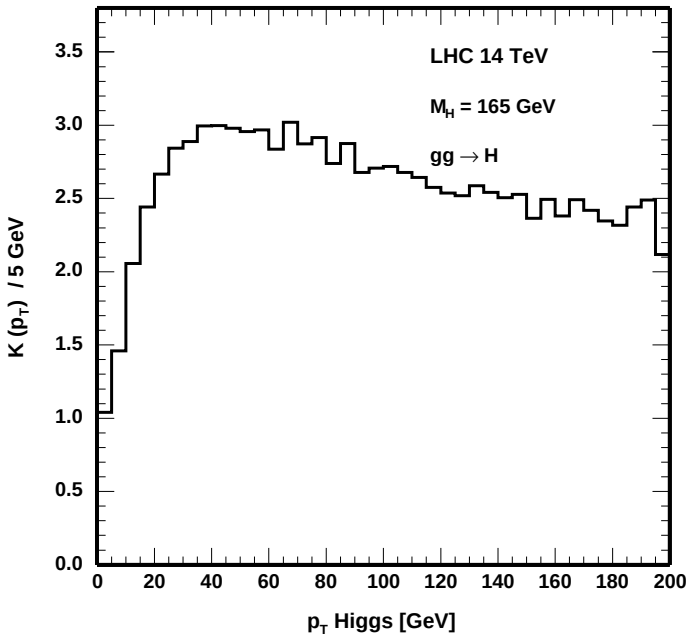

Figure 4: The $p_{\mathrm{T}}^{\mathrm{H}}$ dependence of the K-factor, as defined in section 2.

the efficiency of a jet veto with the one obtained with HERWIG [16], which is known to provide a better description of the $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum in the small-$p_{\mathrm{T}}^{\mathrm{H}}$ region [17]. When $p_{\mathrm{T}}^{\mathrm{H}} \lesssim 40$ GeV, the efficiencies differ by less than 5%, thus confirming the validity of our approach.

JHEP05(2004)009

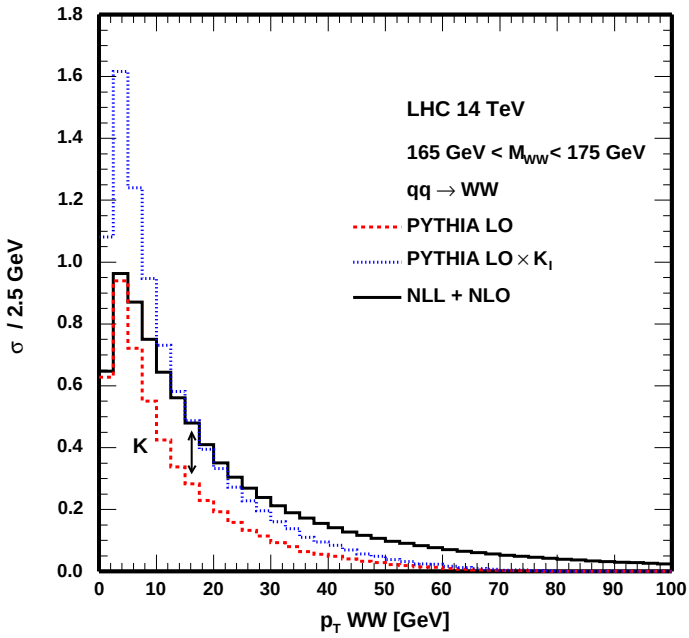

Figure 5: The p_{T} spectrum of the non-resonant WW system with a mass of $170 \pm 5\,\mathrm{GeV}$, obtained from PYTHIA and from the NLL+NLO calculation. The spectrum from PYTHIA rescaled with the inclusive K-factor is also shown for comparison.

Consequently, it is possible to obtain an approximation for the (N)NLO distributions of the kinematic observables used to select the final state by simply reweighting each PYTHIA event in such a way that the new $p_{\mathrm{T}}^{\mathrm{H}}$ spectrum matches the one from the QCD calculation. Of course, the $p_{\mathrm{T}}^{\mathrm{H}}$-dependent signal efficiency is not altered by the reweighting.

A similar procedure is applied for the main background, the continuum production of WW pairs. Here transverse momentum spectra obtained with PYTHIA are reweighted according to QCD predictions at next-to-leading-logarithmic (NLL) accuracy, which are matched to the perturbative NLO result [18, 19], valid at large transverse momenta $p_{\mathrm{T}}^{\mathrm{WW}}$ of the WW pair. For this calculation[5] MRST2002 NLO densities and a running α_{S} in the two-loop approximation are used, so that the integral of the spectrum is fixed to the total NLO cross section [21].

In order to compare the p_{T}-dependent WW spectrum from PYTHIA with the one from the higher-order calculation, the dependence on the mass of the WW system has to be taken into account. This is done for three different mass intervals, $170 \pm 5\,\mathrm{GeV}$, $200 \pm 5\,\mathrm{GeV}$ and $250 \pm 5\,\mathrm{GeV}$, which cover the mass range where the WW events are a potential background for the Higgs signal, using the selection criteria described above. The expected $p_{\mathrm{T}}^{\mathrm{WW}}$ spectra from the two calculations in the WW mass range of $170 \pm 5\,\mathrm{GeV}$ are shown in figure 5. The corresponding K-factors, as a function of $p_{\mathrm{T}}^{\mathrm{WW}}$ for the three different WW mass intervals are shown in figure 6.

[5]The NLL resummed WW cross section is computed according to the formalism of refs. [20, 8]. The NLO result used for the matching is obtained with the MCFM package [18]. More details on these results will be given elsewhere.

JHEP05(2004)009

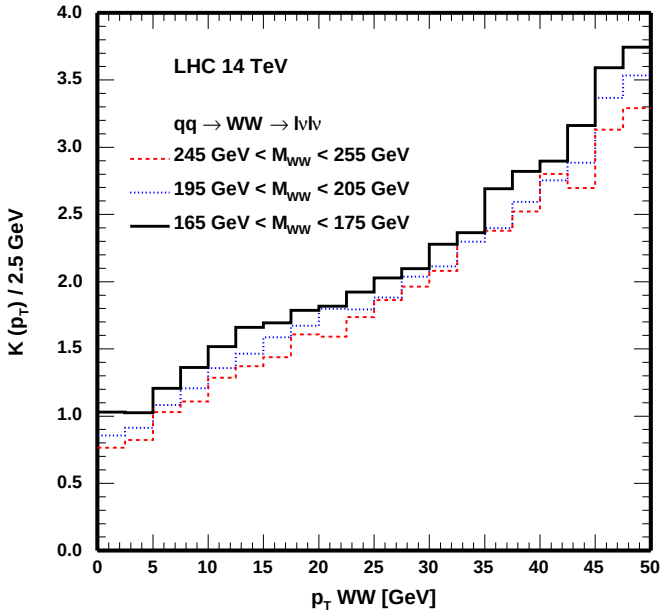

Figure 6: The p_{T} dependence of the K-factor for the non-resonant WW system and three different WW mass intervals.

The difference between the $p_{\mathrm{T}}^{\mathrm{WW}}$ spectrum in PYTHIA and the one calculated in NLL+NLO QCD is particularly large for large transverse momenta. This is because, contrary to the Higgs signal, for WW production no hard matrix-element corrections are applied in PYTHIA, and thus the corresponding spectrum falls rather quickly as $p_{\mathrm{T}}^{\mathrm{WW}}$ increases.

For the analysis described in this paper only the events with relatively small p_{T} are relevant and the corresponding event weights for the non-resonant WW production are found to increase from about 1 at small transverse momentum to almost 4 at a transverse momentum of 50 GeV, slightly depending on the mass of the WW system. However, since most of the relevant WW continuum background comes from events with an invariant mass around threshold and relatively low transverse momentum, we take as an approximate weighting factor for the WW events the one obtained for the mass range 170 ± 5 GeV. As can be seen from figure 6, this will slightly overestimate the WW background.

For the $\mathrm{t\bar{t}}$ and Wtb background, the p_{T} spectrum is taken from PYTHIA, while the corresponding cross sections are both simply scaled by a constant K-factor of 1.5, so that the rescaled $\mathrm{t\bar{t}}$ cross section from PYTHIA matches the inclusive NLO $\mathrm{t\bar{t}}$ cross section [22, 23]. Although this is certainly a crude approximation, it is applied only to the less important $\mathrm{t\bar{t}}$ and Wtb backgrounds. Moreover, as we will discuss in the next section, the accuracy with which the background is known is less important with respect to the one of the signal, as far as the statistical significance is concerned.

5. Results

The effective experimental K-factor can be computed from the sum of the accepted cross section over all the p_{T} bins. The numbers for a Higgs mass of 165 GeV are given in table 2,

JHEP05(2004)009

	$M_{\mathrm{H}} = 165\,\mathrm{GeV}$			
$p_{\mathrm{T}}^{\mathrm{H}}$ [GeV]	$\sigma_{\mathrm{NNLO+NNLL}}$ [pb]	σ_{PYTHIA} [pb]	K	ϵ [%]
0–5	0.861	0.828	1.040	3.0
5–10	1.856	1.272	1.460	2.8
10–15	2.233	1.086	2.057	2.6
15–20	2.180	0.892	2.443	2.4
20–25	1.954	0.733	2.667	2.2
25–30	1.729	0.608	2.842	1.9
30–35	1.481	0.513	2.889	1.5
35–40	1.306	0.436	2.995	1.2
40–45	1.129	0.377	2.997	0.7
45–50	0.976	0.327	2.980	0.4
50–55	0.843	0.285	2.958	0.3
55–60	0.746	0.251	2.968	0.2
60–65	0.637	0.225	2.836	0.1
65–70	0.585	0.194	3.020	0.0
70–80	0.960	0.332	2.892	0.0
80–90	0.744	0.265	2.808	0.0
90–100	0.584	0.217	2.691	0.0
100–200	2.276	0.809	2.813	0.0
Total	23.08	9.74		
	$K_{\mathrm{I}} = 2.37$	$K_{\mathrm{eff}} = 2.04$		

Table 2: Higgs production cross sections as a function of the Higgs transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$, obtained in NNLL+NNLO QCD and with PYTHIA, for a Higgs mass of 165 GeV. The third and fourth columns list the p_{T}-dependent K-factor and the signal selection efficiency, respectively.

both for the NNLL+NNLO and for the PYTHIA prediction. The signal efficiency after the selection described in section 3 is also given in table 2. The efficiency vanishes for $p_{\mathrm{T}}^{\mathrm{H}}$ above 65 GeV. Therefore, although the K-factor for the bin $65 < p_{\mathrm{T}}^{\mathrm{H}} < 70\,\mathrm{GeV}$ is about 3, it will not contribute when computing the effective experimental K-factor K_{eff}.

From the integration over all $p_{\mathrm{T}}^{\mathrm{H}}$ bins, the inclusive K-factor with respect to PYTHIA, without any selection cuts, is found to be $K_{\mathrm{I}} = 2.37$. This is roughly 15% larger than $K_{\mathrm{eff}} = 2.04$, which is obtained after all cuts are applied, including the jet veto of 30 GeV. This means that the number of accepted reweighted events is a factor of 2.04 larger than in the unweighted case. Similar numbers are obtained for other Higgs masses. The results for Higgs masses of 140 GeV and 180 GeV are given in table 3. The estimated effective K-factor for the WW background, integrating over the entire WW mass spectrum and using the p_{T}-dependent weighting factor determined for the WW mass interval of 165–175 GeV, is found to be 1.36. Considering only the WW mass interval from 165–175 GeV, the effective K-factor would be 1.44, which is about 18% lower than the inclusive K-factor for this WW mass interval.

Following this procedure, the next step consists in calculating the luminosity requirements for the observation of a Higgs signal with a statistical significance of five standard

JHEP05(2004)009

	$M_{\rm H} = 140\,{\rm GeV}$			$M_{\rm H} = 180\,{\rm GeV}$		
$p_{\rm T}^{\rm H}$ [GeV]	$\sigma_{\rm NNLO+NNLL}$ [pb]	$\sigma_{\rm PYTHIA}$ [pb]	ϵ [%]	$\sigma_{\rm NNLO+NNLL}$ [pb]	$\sigma_{\rm PYTHIA}$ [pb]	ϵ [%]
0–5	1.342	1.230	2.2	0.680	0.674	0.7
5–10	2.727	1.805	2.1	1.420	1.052	0.8
10–15	3.430	1.495	2.1	1.809	0.918	0.8
15–20	3.162	1.208	2.1	1.762	0.759	0.8
20–25	2.816	0.979	1.9	1.642	0.633	0.8
25–30	2.436	0.800	1.7	1.421	0.526	0.8
30–35	2.094	0.668	1.6	1.250	0.440	0.8
35–40	1.781	0.567	1.3	1.081	0.379	0.6
40–45	1.370	0.483	1.0	0.974	0.330	0.5
45–50	1.445	0.411	0.7	0.820	0.287	0.3
50–60	2.088	0.675	0.5	1.367	0.473	0.2
60–70	1.578	0.516	0.2	1.067	0.369	0.1
70–80	1.198	0.407	0.1	0.829	0.297	0.1
80–90	0.934	0.316	0.0	0.664	0.239	0.0
90–100	0.724	0.257	0.0	0.530	0.194	0.0
100–200	2.665	1.003	0.0	2.064	0.830	0.0
Total	31.79	12.82		19.38	8.40	
	K_I=2.48	$K_{\rm eff}$= 2.25		K_I=2.30	$K_{\rm eff}$=2.03	

Table 3: Higgs production cross sections as a function of the Higgs transverse momentum $p_{\rm T}^{\rm H}$, obtained in NNLL+NNLO QCD and with PYTHIA, for a Higgs mass of 140 GeV (left) and 180 GeV (right). The signal selection efficiency is also given.

deviations. In order to calculate the potential statistical significance of a signal it is usually assumed that the background is accurately known. With this assumption and requiring a significance of five standard deviations, the number of signal events S has to be equal to $5\times\sqrt{B}$,[6] B being the number of background events. Thus, ignoring systematic uncertainties, the accurate knowledge of the signal cross section is more important than the absolute value for the background.

However, once systematic uncertainties are considered, the accurate knowledge of the background becomes relevant, especially if S/B is smaller than 1. In addition, uncertainties from the luminosity and the parton distribution functions, as well as from the experimental efficiency need to be considered in detail. Most of these uncertainties can only be determined accurately once the first LHC data are obtained.

The transverse momentum spectra of the two leptons, $p_{\rm Tmax}^{\ell}$ and $p_{\rm Tmin}^{\ell}$, after the reweighting and with all cuts applied, are shown in figure 7 for $M_{\rm H} = 165\,{\rm GeV}$.

The accepted events for signal and backgrounds for $M_{\rm H} = 140$, 165, 180 GeV and the corresponding statistical significance for $5\,{\rm fb}^{-1}$ are given in table 4. We see that signal-to-background ratios between about 1:2 and 2:1 can be obtained in the mass range under

[6]For small event numbers, this has to be replaced with a probability calculation based on Poisson statistics.

JHEP05(2004)009

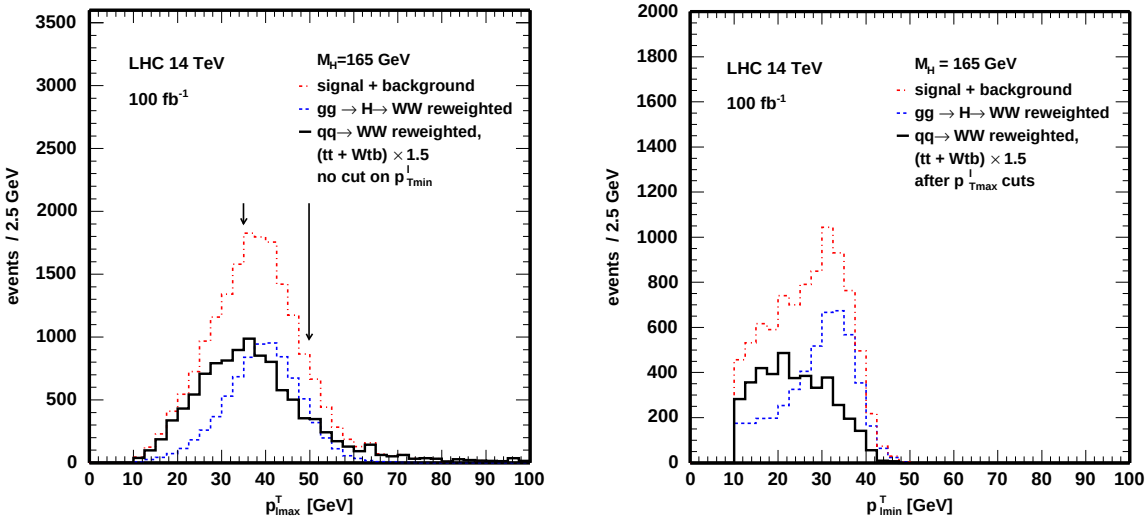

Figure 7: Transverse momentum spectra of the leading lepton (left) and the lepton with the smaller p_{T} (right) from $\mathrm{gg} \to \mathrm{H} \to \mathrm{WW} \to \ell^+\nu\ell'^-\bar{\nu}$ and from the considered backgrounds, as obtained from PYTHIA with event reweighting. The expected background from non-resonant W-pair production is reweighted using the p_{T}-dependent weighting factor, while the ones from $\mathrm{t\bar{t}}$ and Wtb are simply scaled by a factor of 1.5. The p^{ℓ}_{Tmin} spectrum in the right plot is obtained after all cuts, including the cut on p^{ℓ}_{Tmax}, as indicated by the arrows in the left plot.

M_{H} [GeV]	S	WW	Wtb	$\mathrm{t\bar{t}}$	S/B	$S/\sqrt{B}$
140	106	158	87	34	0.38	6
165	162	44	40	7	1.78	17
180	48	23	17	7	1.02	7

Table 4: Number of signal and background events and corresponding statistical significance for an integrated luminosity of 5 fb^{-1}.

consideration. In the case of $M_{\mathrm{H}} = 165\,\mathrm{GeV}$, a statistical significance of five standard deviations can already be achieved with an integrated luminosity of about $0.4\,\mathrm{fb}^{-1}$.

In order not to affect the discovery potential in any significant way, the systematic uncertainties on the background have to be controlled to better than about 10–20%. This may well be achievable knowing that (1) the shape of various distributions can be calculated with good accuracy, and that (2) kinematic regions with small signal contributions can be isolated and used to normalize the potential backgrounds with systematic accuracies of perhaps 5–10%.

6. Summary

We have performed a simulation of the SM Higgs boson search at the LHC in the channel $\mathrm{gg} \to \mathrm{H} \to \mathrm{WW} \to l\nu l\nu$. QCD corrections have been included by using a reweighting procedure, allowing us to combine event rates obtained with PYTHIA with the most up-to-date theoretical predictions for the transverse momentum spectra for the $\mathrm{gg} \to \mathrm{H}$ signal and its main WW background.

JHEP05(2004)009

The reweighting method has been used to compare the experimental sensitivity to find the Higgs estimated with PYTHIA with the one obtained by taking higher-order QCD corrections into account. In particular, the effect of a jet veto on the weighted and unweighted events has been investigated. Using this procedure and a Higgs mass of 165 GeV, the effective experimental K-factor is only about 15% smaller than the inclusive K-factor. From these results the Higgs discovery potential for the channel $pp \to H \to W^+W^- \to \ell^+\nu\ell'^-\bar{\nu}$ is found to be significantly increased. Consequently, signals with a statistical significance of five standard deviations should be observable for a SM Higgs boson with masses between 140 and 180 GeV after the first few fb^{-1} of integrated luminosity.

The reweighting technique proposed in this paper can be applied to other final states and the results should be particularly accurate for hard scattering processes with little additional jet activity.

Acknowledgments

We thank Stefano Catani, Daniel de Florian, Stefano Frixione and Andre Holzner for useful comments and discussions. One of us (MG) would like to thank John Campbell for his help in the use of MCFM.

References

[1] *ATLAS: technical proposal for a general-purpose PP experiment at the large hadron collider at CERN*, CERN/LHCC/94–43;
CMS, the compact muon solenoid: technical proposal, CERN/LHCC/94–38.

[2] D. Cavalli et al., *The Higgs working group: summary report*, hep-ph/0203056.

[3] S. Catani, D. de Florian and M. Grazzini, *Higgs production in hadron collisions: soft and virtual QCD corrections at NNLO*, *J. High Energy Phys.* **05** (2001) 025 [hep-ph/0102227];
R.V. Harlander and W.B. Kilgore, *Soft and virtual corrections to $pp \to H + X$ at NNLO*, *Phys. Rev.* **D 64** (2001) 013015 [hep-ph/0102241];
R.V. Harlander and W.B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, *Phys. Rev. Lett.* **88** (2002) 201801 [hep-ph/0201206];
C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, *Nucl. Phys.* **B 646** (2002) 220 [hep-ph/0207004];
V. Ravindran, J. Smith and W.L. van Neerven, *NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions*, *Nucl. Phys.* **B 665** (2003) 325 [hep-ph/0302135].

[4] K. Cranmer, B. Mellado, W. Quayle and S.L. Wu, *Application of K factors in the $H \to ZZ* \to 4l$ analysis at the LHC*, hep-ph/0307242.

[5] M. Dittmar and H.K. Dreiner, *How to find a Higgs boson with a mass between 155-GeV to 180-GeV at the LHC*, *Phys. Rev.* **D 55** (1997) 167 [hep-ph/9608317].

[6] S. Catani, D. de Florian and M. Grazzini, *Direct Higgs production and jet veto at the Tevatron and the LHC in NNLO QCD*, *J. High Energy Phys.* **01** (2002) 015 [hep-ph/0111164].

[7] T. Sjostrand et al., *High-energy-physics event generation with PYTHIA 6.1*, *Comput. Phys. Commun.* **135** (2001) 238 [hep-ph/0010017].

[8] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *The q(t) spectrum of the Higgs boson at the LHC in QCD perturbation theory*, *Phys. Lett.* **B 564** (2003) 65 [hep-ph/0302104].

[9] M.H. Seymour, *Matrix element corrections to parton shower algorithms*, *Comp. Phys. Commun.* **90** (1995) 95–101 [hep-ph/9410414].

[10] G. Miu and T. Sjostrand, *W production in an improved parton shower approach*, *Phys. Lett.* **B 449** (1999) 313 [hep-ph/9812455].

[11] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *J. High Energy Phys.* **06** (2002) 029 [hep-ph/0204244].

[12] D. de Florian, M. Grazzini and Z. Kunszt, *Higgs production with large transverse momentum in hadronic collisions at next-to-leading order*, *Phys. Rev. Lett.* **82** (1999) 5209 [hep-ph/9902483].

[13] V. Ravindran, J. Smith and W.L. Van Neerven, *Next-to-leading order QCD corrections to differential distributions of Higgs boson production in hadron hadron collisions*, *Nucl. Phys.* **B 634** (2002) 247 [hep-ph/0201114].

[14] C.J. Glosser and C.R. Schmidt, *Next-to-leading corrections to the Higgs boson transverse momentum spectrum in gluon fusion*, *J. High Energy Phys.* **12** (2002) 016 [hep-ph/0209248].

[15] A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, *Uncertainties of predictions from parton distributions, I. Experimental errors*, *Eur. Phys. J.* **C 28** (2003) 455 [hep-ph/0211080].

[16] G. Corcella et al., *Herwig 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *J. High Energy Phys.* **01** (2001) 010 [hep-ph/0011363].

[17] M. Dobbs et al., *The QCD/SM working group: summary report*, hep-ph/0403100.

[18] J.M. Campbell and R.K. Ellis, *An update on vector boson pair production at hadron colliders*, *Phys. Rev.* **D 60** (1999) 113006 [hep-ph/9905386].

[19] L.J. Dixon, Z. Kunszt and A. Signer, *Vector boson pair production in hadronic collisions at $O(\alpha_s)$: lepton correlations and anomalous couplings*, *Phys. Rev.* **D 60** (1999) 114037 [hep-ph/9907305].

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

[21] J. Ohnemus, *An order alpha-s calculation of hadronic W^-W^+ production*, *Phys. Rev.* **D 44** (1991) 1403;
S. Frixione, *A next-to-leading order calculation of the cross-section for the production of W^-W^+ pairs in hadronic collisions*, *Nucl. Phys.* **B 410** (1993) 280.

[22] P. Nason, S. Dawson and R.K. Ellis, *The total cross-section for the production of heavy quarks in hadronic collisions*, *Nucl. Phys.* **B 303** (1988) 607.

[23] Details about the NLO calculation for the $t\bar{t}$ cross section and further references can be found in the proceedings of the workshop on *Standard model physics at the LHC*, M. Beneke et al., CERN 2000-004, Geneva 2000, pp. 427–432.

JHEP05(2004)009