PROCEEDINGS OF SCIENCE

NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks

PoS(DIS2018)089

Tomáš Ježo[*†]
Physics Institute, Universität Zürich, Zürich, Switzerland
E-mail: tomas.jezo@physik.uzh.ch

Measurements of $t\bar{t}H$ production in the $H \to b\bar{b}$ channel depend in a critical way on the theoretical uncertainty associated with the irreducible QCD $t\bar{t}$ + b-jet background. We introduce a new $pp \to t\bar{t}b\bar{b}$ POWHEG generator in the 4F scheme based on POWHEG-BOX-RES and on OPENLOOPS for fast evaluation of the scattering amplitudes. We present predictions and uncertainties for $t\bar{t}$ + b-jet observables at the 13 TeV LHC. We also consider theoretical uncertainties related to the POWHEG matching method and to the parton shower (PS) modelling, with emphasis on $g \to b\bar{b}$ splittings. In general, matching and shower uncertainties turn out to be remarkably small. This is confirmed by a consistent comparison against SHERPA+OPENLOOPS.

XXVI International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2018)
16-20 April 2018
Kobe, Japan

*Speaker.

†Work supported by the Swiss National Foundation (SNF) under contracts BSCGI0-157722 and CRSII2-160814.

1. Introduction

At the Large Hadron Collider (LHC), searches for $t\bar{t}H$ production in the $H \to b\bar{b}$ channel are plagued by a large QCD background, which is dominated by $t\bar{t}b\bar{b}$ production. The availability of precise theoretical predictions for this multi-particle background process is of crucial importance for the sensitivity of $t\bar{t}H(b\bar{b})$ analyses. The process $pp \to t\bar{t}b\bar{b}$ is also very interesting on its own, as it provides a unique laboratory to explore the QCD dynamics of heavy-quark production and to test state-of-the-art Monte Carlo predictions in a nontrivial multi-scale environment.

As a result of its α_S^4 dependence, the leading-order (LO) $t\bar{t}b\bar{b}$ cross section is highly sensitive to variations of the renormalisation scale. The uncertainty corresponding to standard factor-two scale variations amounts to 70–80%, and the inclusion of next-to-leading order (NLO) QCD corrections [1, 2, 3], where the scale dependence reduces down to 20–30%, is mandatory. In Ref. [4] it was pointed out that matching and shower effects also play an unexpectedly important role in $t\bar{t} + b$-jet production. This is due to the fact that two hard b jets can arise from two hard jets each involving a collinear $g \to b\bar{b}$ splitting. In simulations of $pp \to t\bar{t}b\bar{b}$ at NLO and matched to PS (NLO+PS), such configurations result from the combination of a $g \to b\bar{b}$ splitting that is described at NLO accuracy through $t\bar{t}b\bar{b}$ matrix elements together with a second $g \to b\bar{b}$ splitting generated by the PS. The impact of this mechanism can have similar magnitude to the $t\bar{t}H(b\bar{b})$ signal, and a thorough understanding of the related matching and shower uncertainties is very important for $t\bar{t}H$ analyses.

A first assessment of NLO+PS uncertainties was presented in Ref. [5] through a tuned comparison of NLO+PS $t\bar{t}b\bar{b}$ simulations in POWHEL [6, 7], SHERPA+OPENLOOPS [4] and MADGRAPH5AMC@NLO [8]. On the one hand, this study has revealed significant differences between the two generators based on the MC@NLO matching method, i.e. SHERPA and MADGRAPH5AMC@NLO. On the other hand, in spite of the fact that SHERPA+OPENLOOPS and POWHEL implement different matching methods and different parton showers, the predictions of these two generators turned out to be quite consistent. However, due to the limitations related to the use of the five-flavour scheme in POWHEL—which have been overcome only very recently with the 4F upgrade of POWHEL [9]—the agreement between POWHEL and SHERPA+OPENLOOPS did not allow any firm conclusions to be drawn in the study of Ref. [5].

Motivated by the above we present a new POWHEG generator for $pp \to t\bar{t}b\bar{b}$ in the 4F scheme in Ref. [10]. At variance with the POWHEL generator of Ref. [9], this new POWHEG generator is implemented in the POWHEG-BOX-RES framework [11] using OPENLOOPS, which guarantees a very fast evaluation of the required $2 \to 4$ and $2 \to 5$ matrix elements, and supports top-quark decays including spin-correlation effects. Moreover, in order to guarantee a consistent resummation of QCD radiation, the separation of the so-called singular and finite parts in POWHEG-BOX, is not restricted to initial-state radiation as in Ref. [9] but is applied also to final-state radiation.

In these proceedings we briefly review some of the findings of a much more extensive and complete study in Ref. [10], that this manuscript derives from. Specifically, we briefly discuss technical subtleties that arise when matching multi-scale processes like $pp \to t\bar{t}b\bar{b}$ to PS within the POWHEG framework and show a subset of our predictions for $t\bar{t} + b$-jet observables compared to those obtained using SHERPA. We also compare against the predictions of inclusive NLO+PS $t\bar{t}$ generators.

2. Parton shower matched simulations for $t\bar{t}$ + b-jets

In the following we briefly review some aspects of the POWHEG method [12, 13] pertaining to the separation of radiation into singular and finite parts. We pay particular attention to issues related to the multi-scale nature of the process at hand. In particular we point out that the treatment of the recoil associated with the real emissions can induce sizeable distortions of the underlying $t\bar{t}b\bar{b}$ cross section. This technical inconvenience restricts the domain of applicability of QCD factorisation in a way that can jeopardise the efficiency of event generation and can also lead to unphysical resummation effects.

The master formula for the description of NLO radiation in the POWHEG approach consists of two contributions due to splitting the real emission into singular and finite parts, respectively denoted by R_s to R_f. The splitting is achieved via a damping function F that fulfills $F \to 1$ and $F \to 0$ respectively, in the infrared and hard regions of the phase space. Only the singular contribution is used for the calculation of the Sudakov form factor for the generation of the hardest emission, and thus resummed. The default functional form of F in POWHEG-BOX [14, 15] is $F(\Phi) = h^2_\mathrm{damp}/(h^2_\mathrm{damp} + k^2_\mathrm{T}(\Phi))$. It smoothly shifts the weight of real radiation from R_s to R_f when the hardness of the emission, k_T, becomes of the order of the h_damp parameter or higher.

In addition to the well-known h_damp-dependent damping mechanism, POWHEG-BOX also implements a theta function of the form $\tilde{F}(\Phi) = \theta\,(h_\mathrm{bzd} - R(\Phi)/\mathcal{R}(\Phi))$[1] [16, 14], where $\mathcal{R}$ corresponds to the infrared (soft and collinear) approximation of the full matrix element. By default, the cut-off parameter h_bzd is set equal to 5. In this way, in the vicinity of IR singularities, where $R/\mathcal{R} \to 1$, radiative contributions are attributed to R_s and resummed. On the contrary, when the real emission matrix element largely exceeds the IR approximation, the resummation of the full R/B kernel is not well justified, and the corresponding events are attributed to the finite remnant through the theta function. In the standard POWHEG-BOX, and in Ref. [9], the damping function $\tilde{F}$ is applied only to initial-state radiation. However, in the present $t\bar{t}b\bar{b}$ generator we have extended it to all (massless or massive) final-state emitters, that have a singular region associated with it, in order to ensure a consistent resummation of QCD radiation off b-quarks.

The requirement $R(\Phi) < h_\mathrm{bzd}\,\mathcal{K}(\Phi_\mathrm{rad})\,B(\Phi_\mathrm{B})$ was originally introduced in order to avoid possible divergences of $R(\Phi)/B(\Phi_\mathrm{B})$ due to the so-called Born zeros, i.e. phase space regions where $B(\Phi_\mathrm{B}) \to 0$. Such divergences are not physical and cancel in the $\bar{B}/B$ ratio, but they can still lead to dramatic inefficiencies in the event generation.

In the case of $pp \to t\bar{t}b\bar{b}$, such effects can arise from the interplay of soft and collinear enhancements due to NLO light-jet radiation and to the generation of the $b\bar{b}$ system in regions with $m_{b\bar{b}} \ll m_{t\bar{t}b\bar{b}}$ and/or $p_{\mathrm{T},b\bar{b}} \ll m_{t\bar{t}b\bar{b}}$. For example, let us consider a $gg \to t\bar{t}b\bar{b}g$ event with a gluon emission from the initial state. Its kinematics are generated starting from a $gg \to t\bar{t}b\bar{b}$ Born event through a mapping that creates the required gluon recoil by boosting the final state in the transverse direction, modifying the kinematics of the $b\bar{b}$ system as well. Because the $t\bar{t}b\bar{b}$ system has enough energy to absorb the recoil of gluon emissions much harder that the $b\bar{b}$ system this may lead to a very significant reduction of $p_{\mathrm{T},b\bar{b}}$. This violates the main assumption that justifies the POWHEG master formula, namely $R_\alpha(\Phi_\alpha)/B(\Phi_B) \sim \mathcal{R}_\alpha(\Phi_\alpha)/B(\Phi_B) = \mathcal{K}_\alpha(\Phi_\mathrm{rad})$, which requires a sufficiently hard $t\bar{t}b\bar{b}$ process as compared to the k_T of NLO radiation. In particular, due to

[1] The $\tilde{F}$ function multiplies F when combined.

the sensitivity of the Born amplitude to scales of the order $p_{\mathrm{T},b\bar{b}} \sim (E_{b\bar{b}}/E_{t\bar{t}b\bar{b}})\, p_{\mathrm{T},j} \ll p_{\mathrm{T},j}$, the factorisation condition is not fulfilled.

3. Setup

The predictions in this work are obtained using the same choices for the heavy-quark masses, flavour schemes, renormalisation and factorization scales and PDFs as in Ref. [10], which also correspond to the setup recommended in [5]. The POWHEG-BOX parameters h_{bzd} and h_{damp}, which control the resummation of NLO radiation, have been set to $h_{\mathrm{bzd}} = 2$ and $h_{\mathrm{damp}} = H_{\mathrm{T}}/2 = 1/2 \sum_{i=t,\bar{t},b,\bar{b}} E_{\mathrm{T},i}$. To account for the uncertainties associated with this choice we apply the independent variations $h_{\mathrm{bzd}} = 2,\ 5,\ 10$ and $h_{\mathrm{damp}} = H_{\mathrm{T}}/4, H_{\mathrm{T}}/2, H_{\mathrm{T}}, 1.5\, m_t$, varying both parameters one at a time. At LO+PS level we set the shower starting scale equal to $H_{\mathrm{T}}/2$ and vary it up and down by a factor of two in order to assess the related uncertainty. At NLO+PS, the shower starting scale is dictated by the kinematics of real emission matrix elements in the POWHEG method.

In order to assess uncertainties due to the parton-shower modelling of $g \to b\bar{b}$ splittings we vary the parameter `TimeShower:weightGluonToQuark`, which permits one to select out of various forms of the $g \to Q\bar{Q}$ splitting kernel in PYTHIA8 8, considering the options 2 and 4 [17]. In addition to the functional form of the heavy-quark splitting kernel we also vary the scale of α_{S} in the parton shower. To this end, we set `TimeShower:weightGluonToQuark` to 6 and 8, which corresponds to the options 2 and 4 with $\alpha_{\mathrm{S}}(p_T^2)$ replaced by $\alpha_{\mathrm{S}}(m^2_{b\bar{b}})$ in the heavy-quark splitting kernel. Moreover, using `TimeShower:renormMultFac`, we vary $\alpha_{\mathrm{S}}(p_{\mathrm{T}}^2) \to \alpha_{\mathrm{S}}(\xi p_{\mathrm{T}}^2)$ with prefactors $\xi = 0.1, 1, 10$. This latter variation is applied to all final-state QCD splittings, i.e. also splittings of type $g \to gg$, $q \to qg$, etc.

For the reconstruction of jets we use the anti-k_{T} [18] algorithm with $R = 0.4$. We select jets that fulfil $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$ both for the case of light jets and b-jets. At parton level, we define as b-jet a jet that contains at least a b-quark, such that jets that contain a $b\bar{b}$ pair arising from a collinear $g \to b\bar{b}$ splitting are also tagged as b-jets.

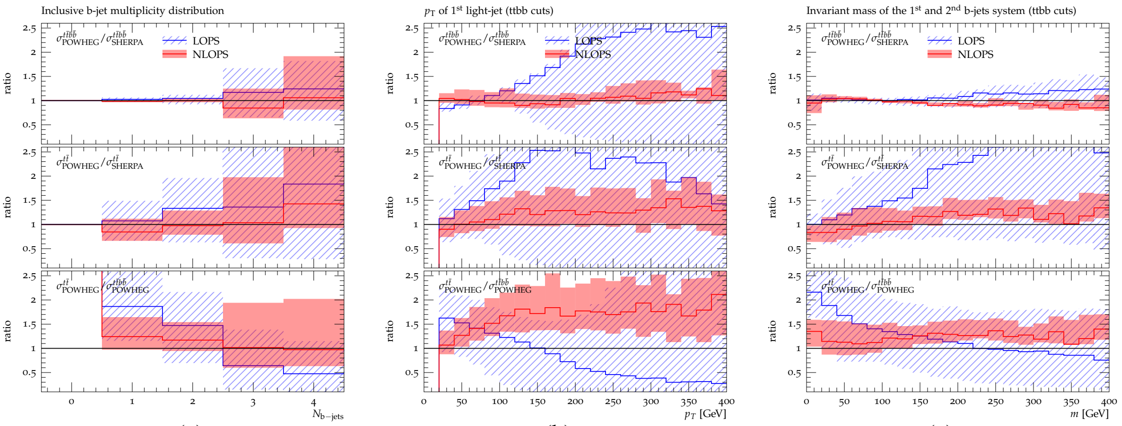
We categorise events according to the number N_b of b-jets that do not arise from top decays and fulfil the acceptance cuts. For the analysis of cross sections and distributions we consider an inclusive selection with $N_b \geq 1$ and a more exclusive one with $N_b \geq 2$. We refer to them as ttb and ttbb selections, respectively.

4. Results

In Fig. 1 we compare $t\bar{t} + b$-jet predictions based on POWHEG+PYTHIA8 and SHERPA. This comparison is done both for (N)LO+PS $pp \to t\bar{t}b\bar{b}$ generators in the 4F scheme and for corresponding generators of inclusive $t\bar{t}$ production in the 5F scheme[2].

Input parameters, QCD scales and matching parameters are chosen as coherently as possible across all generators. In particular, the parameter $h_{\mathrm{damp}} = H_{\mathrm{T}}/2$ in POWHEG is identified with the resummation scale μ_Q in the SMC@NLO framework of SHERPA. Instead, for what concerns the parton showers we simply use standard settings, i.e. we do not try to improve the agreement between generators by tuning the PYTHIA8 and SHERPA showers.

[2] In the case of POWHEG we use `hvq` [19].



(a) (b) (c)

Figure 1: Predictions for $pp \to t\bar{t}+b$-jets at $\sqrt{s} = 13$ TeV: distributions in the inclusive number of additional b-jets (a), the first light jet (b) with ttbb cuts, and in the invariant mass (c) with ttbb cuts. The various ratio plots compare $t\bar{t} + b$-jet observables as described in LO+PS (blue) and NLO+PS (red) simulations based on $pp \to t\bar{t}b\bar{b}$ or $pp \to t\bar{t}$ matrix elements in POWHEG+PYTHIA8 or SHERPA. In the ratios shown in the upper and middle frame POWHEG predictions are normalised to SHERPA ones for the case of $pp \to t\bar{t}b\bar{b}$ and $pp \to t\bar{t}$ simulations, respectively. The third frame displays the ratio of $t\bar{t}$ to $t\bar{t}b\bar{b}$ POWHEG predictions. For all ratios the numerator and denominator are evaluated at the same order, and uncertainties are applied only to the numerator. They correspond to the combination in quadrature of h_{damp} and h_{bzd} variations with the uncertainties due to the modelling of $g \to b\bar{b}$ splittings and the choice of α_S and the shower starting scale in PYTHIA8.

The ratios in the upper frames show POWHEG $pp \to t\bar{t}b\bar{b}$ predictions normalised to corresponding SHERPA predictions at LO+PS and NLO+PS accuracy. The bands describe the combination in quadrature of all matching and shower uncertainties in POWHEG+PYTHIA8 (referred to shower uncertainties in the following), while only nominal SHERPA predictions are considered in the ratios. At LO+PS, for observables that are inclusive with respect to jet radiation we find deviations between 10–40% and comparably large shower uncertainties. In contrast, in the jet-p_T distributions the predictions of PYTHIA8 are far above the ones of SHERPA, with differences that can reach a factor 2.5 in the tails. These differences are perfectly consistent with LO+PS shower uncertainties, which are dominated by variations of the PYTHIA8 starting scale.

Moving to NLO+PS reduces the direct dependence on the PS. At the same time, differences between the POWHEG and SMC@NLO matching methods come into play. In practice, at NLO+PS we observe a drastic reduction of shower uncertainties, especially in the light-jet p_T-distribution. Also the differences between POWHEG and SHERPA become very small at NLO+PS. The ttb and ttbb cross sections agree at the percent level, and differential b-jet observables deviate by more than 5% only in the tail of the $m_{b_1b_2}$ distribution. Even the light-jet spectra in the ttbb phase space deviate by less than 10–20% up to high p_T, in spite of the limited formal accuracy (LO+PS) of such observables. In the light of these results, NLO+PS theoretical uncertainties related to the matching scheme and the PS seem to be well under control in $pp \to t\bar{t}b\bar{b}$ and are clearly subleading as compared to QCD scale uncertainties shown on Figs. 7 and 8 in Ref. [10].

In the central frames we compare (N)LO+PS generators of inclusive $t\bar{t}$ production based on POWHEG+PYTHIA8 and SHERPA. In this case, the $g \to b\bar{b}$ final-state splittings that give rise to

$t\bar{t}$ + b-jet signatures are entirely controlled by the PS. At LO+PS, the parent gluon that splits into $b\bar{b}$ is also generated by the PS. Nevertheless, the ttb and ttbb LO+PS cross sections predicted by POWHEG and SHERPA deviate by less than 30%–40%. Instead, as expected, the shapes of $t\bar{t}$ + b-jet observables vary very strongly, and in all considered light-jet and b-jet distributions PYTHIA8 results exceed SHERPA ones by a factor of two and even more. This excess is well consistent with the estimated LO+PS shower uncertainties. At NLO+PS, only $g \to b\bar{b}$ splittings are controlled by the PS, while the emission of their parent gluon is dictated by LO matrix elements. Consequently, we observe a drastic reduction of shower uncertainties as compared to LO+PS. The differences between POWHEG and SHERPA are also largely reduced at NLO, nevertheless they remain quite significant in various distributions.

To provide a more complete picture of the uncertainties of inclusive $t\bar{t}$ simulations, in the lower frames we compare POWHEG+PYTHIA8 generators of inclusive $t\bar{t}$ production and $t\bar{t}b\bar{b}$ production. Shower uncertainties are shown only for the $t\bar{t}$ generator. At LO+PS, the $t\bar{t}$ generator is strongly sensitive to the modelling of $pp \to t\bar{t}g$ through initial-state gluon radiation in PYTHIA8. As a result, the $t\bar{t}$ generator overestimates the ttb and ttbb cross sections by about 90% and 50%, respectively. This excess is strongly sensitive to the shower starting scale, and in the p_{T}-distributions it is confined to the regions below 100–200 GeV, while the tails are strongly suppressed. Also the $m_{b_1b_2}$ distribution features a strong shape difference as compared to LO+PS $t\bar{t}b\bar{b}$ predictions.

Such differences decrease significantly at NLO+PS. The ttb and ttbb cross sections predicted by the $t\bar{t}$ generator overshoot $t\bar{t}b\bar{b}$ results by only 15–20%, and b-jet observables also feature an improved agreement with $t\bar{t}b\bar{b}$ predictions. Nevertheless, in b-jet observables we find quite significant shape differences, especially for the $m_{b_1b_2}$ distribution, and shower uncertainties remain far above the ones of the $t\bar{t}b\bar{b}$ generator (see upper frame). As for the light-jet spectra, $t\bar{t}$ predictions turn out to lie above $t\bar{t}b\bar{b}$ ones by about a factor of two in the tails.

5. Summary

Searches for $t\bar{t}H$ production in the $H \to b\bar{b}$ channel call for a precise theoretical description of the irreducible QCD $t\bar{t}$ + b-jet background. We have presented a new $pp \to t\bar{t}b\bar{b}$ POWHEG generator in the 4F scheme based on OPENLOOPS and POWHEG-BOX-RES, the latter requiring a gentle modification in order to overcome subtle technical issues that arise in matching multi-scale processes to PS using the POWHEG method. At NLO+PS, the matching uncertainties due to h_{damp} and h_{bzd} and shower uncertainties due to the modelling of $g \to b\bar{b}$ and the choice of α_{S} in PYTHIA8 turn out to be rather small and clearly subleading with respect to QCD scale variations. We find that with the help of parton-shower tuning, the inclusive NLO+PS $t\bar{t}$ generators may potentially be amenable to a reasonable description of inclusive $t\bar{t}$ + b-jet observables. However, it should be clear that NLO+PS $t\bar{t}b\bar{b}$ generators are mandatory in order to achieve an acceptable level of shower systematics.

6. Acknowledgments

I am grateful to Silvia Ferrario Ravasio and Katie Whitfield for a careful reading of the manuscript.

References

[1] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD corrections to $pp \to t\bar{t}b\bar{b} + X$ at the LHC*, *Phys. Rev. Lett.* **103** (2009) 012002, [arXiv:0905.0110].

[2] G. Bevilacqua, M. Czakon, C. G. Papadopoulos, R. Pittau, and M. Worek, *Assault on the NLO Wishlist: $pp \to t\bar{t}b\bar{b}$*, *JHEP* **09** (2009) 109, [arXiv:0907.4723].

[3] A. Bredenstein, A. Denner, S. Dittmaier, and S. Pozzorini, *NLO QCD Corrections to Top Anti-Top Bottom Anti-Bottom Production at the LHC: 2. full hadronic results*, *JHEP* **03** (2010) 021, [arXiv:1001.4006].

[4] F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini, and F. Siegert, *NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks*, *Phys. Lett.* **B734** (2014) 210–214, [arXiv:1309.5912].

[5] **LHC Higgs Cross Section Working Group** Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922.

[6] M. V. Garzelli, A. Kardos, and Z. Trocsanyi, *$t\bar{t}b\bar{b}$ hadroproduction at NLO accuracy matched with parton shower*, *PoS* **EPS-HEP2013** (2013) 253.

[7] M. V. Garzelli, A. Kardos, and Z. Trocsanyi, *Hadroproduction of $t\bar{t}b\bar{b}$ final states at LHC: predictions at NLO accuracy matched with Parton Shower*, *JHEP* **03** (2015) 083, [arXiv:1408.0266].

[8] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **1407** (2014) 079, [arXiv:1405.0301].

[9] G. Bevilacqua, M. V. Garzelli, and A. Kardos, *$t\bar{t}b\bar{b}$ hadroproduction with massive bottom quarks with PowHel*, arXiv:1709.06915.

[10] T. Ježo, J. M. Lindert, N. Moretti, and S. Pozzorini, *New NLOPS predictions for $\boldsymbol{t\bar{t}+b}$ -jet production at the LHC*, *Eur. Phys. J.* **C78** (2018), no. 6 502, [arXiv:1802.00426].

[11] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, *JHEP* **12** (2015) 065, [arXiv:1509.09071].

[12] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146].

[13] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092].

[14] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581].

[15] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, *JHEP* **0904** (2009) 002, [arXiv:0812.0578].

[16] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO vector-boson production matched with shower in POWHEG*, *JHEP* **0807** (2008) 060, [arXiv:0805.4802].

[17] See http://home.thep.lu.se/~torbjorn/pythia82html/TimelikeShowers.html.

[18] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_T jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189].

[19] S. Frixione, P. Nason, and G. Ridolfi, *A Positive-Weight Next-to-Leading-Order Monte Carlo for Heavy Flavour Hadroproduction*, *JHEP* **09** (2007) 126, [arXiv:0707.3088].