Matching uncertainties in the prediction of the Higgs boson transverse momentum in the SM and beyond

Emanuele Bagnaschi

Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany
E-mail: emanuele.bagnaschi@desy.de

We present the results of our recent study [1] of the theoretical uncertainties that affect the predictions for the Higgs-boson transverse-momentum in gluon fusion when fixed- and all-order results are matched. Our investigation consists of a twofold analysis: first we present a detailed comparison of two recently introduced prescriptions for the determination of the matching scale [2, 3], then we apply the results of these methods to three widely used matching frameworks, namely the aMC@NLO and POWHEG Monte Carlo approaches and analytic resummation. The results of our study are applied to the production of the SM Higgs boson and of the neutral Higgs bosons of the Two-Higgs-Doublet Model in a variety of scenarios.

DESY 16-173
1. Introduction

After the discovery of the Higgs boson at the LHC [4, 5], a quest has started to characterize the newly found particle. Its properties have already been a topic of studies of many articles whose aim was to understand the compatibility with the Standard Model (SM) predictions. So far these studies have relied on the precise predictions available for production and decay rates. Only in the last year the experimental measurements for differential observables have been published [6, 7]. As already stressed by various authors [8, 9, 10, 11, 12, 13, 14, 15, 16] the Higgs-transverse momentum ($p_\perp$) distribution opens the possibility of probing the loop dynamic of the gluon fusion process. This observable is therefore sensible to modifications of the Yukawa couplings and/or to the presence of new states beyond the SM ones. The opportunity of exploiting the $p_\perp$ measurement is becoming more and more interesting as the LHC accumulates a new wealth of data during the second run of operations. At the same time, when considering models of new physics with enlarged Higgs sectors, an accurate description of the transverse momentum, which can show sensible deviations from the SM prediction, is required.

To properly describe the transverse momentum distribution, whose fixed-order prediction is logarithmically divergent in the limit $p_\perp \to 0$, one needs to resum the terms enhanced by powers of $\log(p_\perp/m_h)$ to all orders in $\alpha_s$, where $m_h$ is the mass of the scalar resonance. This resummation is usually performed either analytically or algorithmically. Due to the theoretical formalism on which it is based, the resummation procedure is strictly valid only in the limit of collinear emissions and therefore, for gluon fusion, in the limit of zero transverse momentum of the Higgs boson. Therefore, to properly describe the whole $p_\perp$-spectrum, a matching between the fixed order and the resummed results is needed. Particular care has to be taken to avoid any kind of double counting. This has been achieved in various frameworks, both analytic [17, 18, 19, 20] and numeric [21, 22]. Common to all the approaches is the introduction of a new unphysical scale, which we will subsequently denote as the matching scale ($\mu$), whose role is to define the transition region between the fixed- and the all-order results. The dependence of the matched result on this scale is of higher logarithmic order, however a careless choice can ruin the perturbative convergence of the result. This is especially true for those processes that are characterized by more than one scale, as the one that we are considering in our study, as we will see in the next sections.

2. The Higgs transverse momentum as a multiscale problem

Higgs production in gluon fusion has been originally studied in the so-called Heavy Quark Effective Field Theory (HQEFT) obtained in the limit where the top-quark mass is taken to be very large compared to Higgs boson mass. The use of the HQEFT has the advantage of reducing what is a one-loop LO process to a tree level one, with a sensible decrease of the complexity of the computation. Under this approximation, the total cross section has been computed up to $N^3$LO, while differential computations are available up to NNLO (see ref. [23, 24, 25] for a complete reference of all the results available). However, being this an effective description, it is valid only if we are not probing mass scales that are equal or larger than the top quark mass. Therefore, the HQEFT is a description limited to Higgs boson masses smaller than the top quark mass, for what concerns the total cross section, and to $p_\perp$ less than top mass for the transverse momentum
distribution. Moreover, it neglects completely the contribution coming from diagrams where the coupling of the Higgs to the gluons is mediated by a loop of light quarks. The latter are important for precision predictions in the SM and fundamental for Beyond-Standard-Model (BSM) Higgs boson production, where it can happen for specific regions of the BSM parameter space that the dominant contribution to the cross section is from bottom quark diagrams, oppositely to the SM. Therefore, to properly describe the Higgs boson transverse momentum, we need to perform the computation in the full theory, being it either the SM or a BSM model as the THDM or the MSSM.

In the case of the SM and the THDM, complete computations are available up to NLO. In the case of the MSSM, even the full NLO result is not known analytically.

Restricting ourselves, for simplicity to a THDM-like scenario, the description of the Higgs transverse momentum in gluon fusion is characterized by four mass scales: the Higgs boson mass; the top quark mass; the bottom quark mass; the $p_\perp$ of the radiated parton. All these physical scales and their non-trivial interplay have to be taken into account in the choice of the matching scale.

To follow this requirement, it was first proposed by Grazzini et al. [26] to split the complete squared matrix element into components that are characterized just by a single mass scale or by a specific combination thereof. Originally this split was into two parts, the top contribution and the bottom plus the interference terms. Here we follow more recent developments where the amplitude is divided into three terms [2, 3]. To achieve this, we rewrite the full amplitude as

$$|\mathcal{M}(t+b)|^2 = |\mathcal{M}(t)|^2 + |\mathcal{M}(b)|^2 + (|\mathcal{M}(t+b)|^2 - |\mathcal{M}(t)|^2 - |\mathcal{M}(b)|^2),$$

then compute separately the matched prediction for each contribution, with a different matching scale each, and finally sum all the three components together. We note that, while the decomposition given in eq. (2.1) is a trivial identity at the level of the total cross section, because the latter is independent on the matching scale, it yields a modified shape for the transverse momentum distribution with respect to the one obtained using a single scale for the full amplitude. Our master formula for the best prediction of the Higgs boson transverse momentum distribution is therefore given by

$$\frac{d\sigma}{d\mathbf{p}_\perp} = \frac{d\sigma_t}{d\mathbf{p}_\perp}|_{\mu_t} + \frac{d\sigma_b}{d\mathbf{p}_\perp}|_{\mu_b} + \frac{d\sigma_{int}}{d\mathbf{p}_\perp}|_{\mu_{int}},$$

where with $\mu_t$, $\mu_b$ and $\mu_{int}$ we have denoted respectively the matching scale for the top and bottom contributions and the interference term.

2.1 Matching scale determination

In our work we have compared two prescriptions that recently appeared in the literature\(^1\): the partonic analysis published in ref. [3] (BV) and the hadronic analysis in ref. [2] (HMW).

The BV prescription is based upon the observation that the resummation formalism relies on the factorization of the squared matrix element in the limit of soft and/or collinear emissions (see

\(^1\)In the context of SCET, though only in the HQEFT, a detailed study on the problem of scale determination has been published in ref. [30, 31, 32].
Figure 1: On the left, ratio (C) between the full squared matrix amplitude and the collinear limit of the gluon initiated subprocess for the top, bottom and interference terms as well as for the HQEFT computation, as a function of the Higgs transverse momentum; the dashed lines indicate a deviation of ±10% from one. On the right, ratio between the matched and the fixed order predictions at the hadronic level, for the top contribution and for different values of matching scale. The latter is here denoted by $Q_{\text{res}}$ because the results shown are obtained in the framework of AR.

also ref. [27]). In detail, the accuracy of the collinear approximation of the gluon fusion process, as a function of $p_\perp$, is evaluated at the partonic level, separately for each term in eq. (2.1). If the approximation is violated by a well-defined threshold, here chosen to be 10%, for a given value of $p_\perp$, then the latter is taken to be the matching scale to be used in the matched computation (see the left plot of fig. 1). The procedure is applied separately to the gluon-gluon and the quark-gluon subprocesses. The results are then averaged with a differential-weight that keeps into account that the two channels contribute with varying proportions at different $p_\perp$ due to the distinct behavior of the quark and gluon PDFs.

The HMW method follows from two principles: for $p_\perp\gtrsim m_H$ the spectrum is correctly described by fixed-order perturbation theory and therefore the latter should be the definitive prediction in this range; one would like to have an all-order result for a $p_\perp$-range as large as possible. Practically, this translates into the definition of a scale $Q_{\text{max}}$ as the maximum scale for which the resummed distribution is within the interval $[0,2] \cdot [d\sigma^NLO/dp_\perp]$ for $p_\perp \geq m_H$. The matching scale $\mu$ is then taken to be equal to half $Q_{\text{max}}$ (see the right plot of fig. 1).

We stress that, in both cases, the scales are independent on the Yukawa coupling of the quark to the Higgs.

2.2 Matching scale comparison

In fig. 2 we show the two scale sets as a function of the Higgs boson mass, for the scalar case. Due to the different assumptions on which the two procedures are based, it is not surprising that the numerical values are different. Indeed the BV prescription is sensitive to the behavior of the transverse spectrum in the low-$p_\perp$ region, while the HMW method is designed around the high-$p_\perp$ tail. What we observe is a moderate agreement for the top contribution (with only the BV scale showing a sensitivity to the $2 \cdot m_{\text{top}}$ threshold) and a very good agreement for the bottom one. Instead, for the interference term the two prescriptions can differ by a large amount and specifically this happens when the LO term is much smaller than the NLO one. Indeed, in this case we have that
the resummed contribution, being proportional\(^2\) to the LO term, is small and the Higgs transverse momentum will be given almost completely by the hard emission from the NLO term. Then the collinear approximation will fail for any value of \(p_\perp > 0\) and therefore the BV scale will vanish; on the contrary, because the matched curve is almost identical to the fixed order one, for every value of the resummation scale, the HMW scale will tend to be very large. As a general feature of both scale sets, we observe that, for heavy Higgs masses, the scales for all the three contributions are much smaller than the commonly used value of \(m_h/2\).

3. Simulation setup

In our study of the theoretical uncertainties of the Higgs transverse momentum distribution we used three different codes: MoRe-SusHi, which implements the Analytic Resummation (AR) procedure; \texttt{gg\_H\_2HDM} from the POWHEG-BOX framework; \texttt{aMCSusHi}, based upon the aMC@NLO framework and the SusHi [28] amplitudes. The uncertainty band due the matching is determined using the following prescription: given the reference values \((\mu_t, \mu_b, \mu_{int})\) for the three matching scales discussed in the previous sections, we consider all the possible combinations generated by taking half and twice these values or the reference values themselves; for each combination we compute our prediction for the Higgs \(p_\perp\); finally we take the envelope for each \(p_\perp\) bin, i.e. we take the maximum and the minimum value among all the predictions. Only for AR, we follow ref. [2] and we apply an additional damping factor to the error band at large \(p_\perp\).

In our study, besides the SM, we considered various THDM-II scenarios. Each of the latter was chosen because it is characterized by a specific feature, e.g. a very large Yukawa coupling to one of the two quarks, whose impact on the \(p_\perp\) distribution we want to understand. We also considered both scalar and pseudoscalar productions. In this proceeding, due to the restricted space available,
Figure 3: Shapes of the transverse momentum distribution for a SM Higgs boson of mass $m_h = 125$ GeV. We show the curves from AR (black, solid), MC@NLO (red, dotted) and POWHEG (blue, dashed overlaid by points) for the BV scale set. For reference we also show the fixed order result (green, dash-dotted with open boxes). On the left, we plot the spectrum up to 400 GeV, on the right we show a zoom around the first 100 GeV. In the main frame we show the absolute distributions, in the middle inset the ratio of the central curves to the analytic resummation result and in the bottom inset the uncertainty bands, again normalized to the AR resummation value.

we report the results only for the SM Higgs boson using BV scales and for heavy scalar production in the large-$b$ scenario with the HMW scale set. See table 1 of ref. [1] for a list of all scenarios.

All the numerical results are computed for the LHC, with a center-of-mass energy of $\sqrt{s} = 13$ TeV, using the MSTW2008nlo68c1 PDF set and the associated value of $\alpha_s(M_Z) = 0.120179$. The renormalization and factorization scales are set to the Higgs boson mass; the quark masses are fixed at $m_{t_{op}} = 172.5$ GeV and $m_b = 4.75$ GeV respectively; the Yukawa couplings of the Higgs to quarks are renormalized in the On-Shell (OS) scheme. We used Pythia8 as our Parton Shower [29].

4. SM results

In fig. 3 we show the shape of the transverse momentum distribution (i.e. the integral of each curve is normalized to one) obtained using the BV scale set for a SM Higgs boson of $m_h = 125$ GeV. Qualitatively we observe that all the codes agree within their uncertainty bands, at least for $p_{\perp} < m_b$. More in detail, we observe that the two MC event generators, POWHEG and MC@NLO are in excellent agreement in the region $10 < p_{\perp}/\text{GeV} < 130$, while they differ by about 20% from the central AR prediction. The position of the peak is also slightly different between the MCs and AR. Turning now our attention to the high-$p_{\perp}$ tail, we see that: the AR prediction approaches the fNLO at the level of 5% above $p_{\perp} \simeq 130$ GeV; the transition to the NLO prediction in MC@NLO is around $p_{\perp} \simeq 180$ GeV; POWHEG, on the other hand, always remains 20% above the fixed order result. The latter is a general feature of the POWHEG matching that we observe and that will be analyzed more in detail in the THDM analysis that follows.
Concerning the uncertainty bands, for $p_\perp > 130$ GeV AR has no uncertainty band due to the artificial suppression given by the damping factor, as mentioned in section 3; on the other hand, the uncertainty for the MC@NLO prediction is of order $\pm 10\%$, while the width of the POWHEG band decreases uniformly from $\pm 20\%$ to $\pm 10\%$. In the intermediate region, all three codes show a bulgy structure with a maximum of $\pm 20\%$ and $\pm 35\%$ for AR and the MCs respectively and a minimum of just a few percent above the peak position. Finally, in the small-$p_\perp$ region, the AR uncertainty band grows to $100\%$, the POWHEG one to $40\%$, while MC@NLO shows only a moderate $\pm 15\%$ uncertainty.

5. THDM results

The left plot of fig. 4 shows the same curves as the SM case discussed in the previous section. Differently from the SM, however, the discrepancies between the three frameworks are more marked. Concerning the central curves, we observe that POWHEG produces a spectrum that is significantly harder, being over $50\%$ above the fNLO result for $p_\perp > 200$ GeV. On the other hand, in the intermediate region between $10 \lesssim p_\perp/\text{GeV} \lesssim 130$, POWHEG and AR agree within $10\%$ while the MC@NLO curve is substantially larger. At smaller transverse momentum, for $p_\perp < 30$ GeV, the two MCs have a much better agreement.

The behavior of the uncertainty bands is also quite different in the various frameworks: the MC@NLO band blows up to $\mathcal{O}(100\%)$ around $p_\perp \simeq 125$ GeV; the POWHEG band remains very small all over the whole $p_\perp$ range.

To understand the origin of these differences, we investigated the dependence of the MC predictions on the prescription for the shower scale $Q_{sh}$. The latter is the scale that it is passed to the PS to be used as an upper bound for the $p_\perp$ of the emitted radiation. Concerning POWHEG, we
notice that by restricting, for the class of events\textsuperscript{3} that describe the high-$p_{\perp}$ tail, the shower scale to be at most the matching scale (either BV or HMW), we recover the fixed order result in the same way as AR and MC@NLO do (as it can be seen from the purple curve on the right plot of fig. 4). Moreover, the shape of the uncertainty band is also changed, showing now a bulge between 50 GeV and 100 GeV. Relatively to MC@NLO, we first recall that by default the shower scale is extracted from a probability distribution dependent on the LO kinematic and centered around the matching scale \cite{34}: if we replace the default distribution with the $\delta$-function $\delta(Q_{sh} - \mu)$, we observe a significant change for both the central prediction and uncertainty bands, as it can be seen by the yellow curve in the right plot of fig. 4. These observations lead us to conclude that for this specific scenario there is a high sensitivity not only to the numerical values of the scales but also to the specific details of the matching procedure.

### 6. Conclusions

In this talk we have presented the results of our recent study \cite{1} of the theoretical uncertainties intrinsic to the matching procedure between fixed- and all-order results in the computation of the transverse momentum distribution of the Higgs boson in gluon fusion. Specifically for this process, which involves different mass scales, even the choice of the central values for the matching scales has become a matter of debate. In this context, we performed a thorough analysis of the predictions obtained using three different matching frameworks (analytic resummation, POWHEG and MC@NLO) and two different prescriptions for the determination of the matching scales (BV or HMW). Our comparison was twofold: first we addressed the issue of the determination of the central value for the matching scale for the top and bottom contributions and for the interference term, by providing a qualitative and quantitative comparison of the BV and HMW approaches; then we compared the results for the shape of the $p_{\perp}$ distribution obtained with different scale-choices and frameworks.

We have found that the prediction of the Higgs transverse momentum is affected by uncertainties\textsuperscript{4} up to several tens of percent, depending on the scenario, the $p_{\perp}$ value and the framework under consideration. In the low-$p_{\perp}$ region we find reasonable agreement between the different codes, although AR usually shows a much softer spectrum. However, in all the three frameworks and especially for AR, the error bands grow in this region, therefore providing compatibility between the different results. In the intermediate region, we find non-trivial differences between the three frameworks, which are more pronounced in the bottom dominated scenarios. In the latter case, we also find a large dependence on the specific details of the matching formulation inside each framework. In the large-$p_{\perp}$ tail, where technically all the codes have LO accuracy \textsuperscript{5}, we find that POWHEG systematically predicts a harder spectrum than MC@NLO and AR. The latter are instead softer and compatible with the fNLO result. We identified one source of this difference in the

\textsuperscript{3}These are the remnant events. By default setting of the POWHEG-BOX their shower scale is set to the transverse momentum of the emitted parton. See ref. \cite{33} for a detailed description of the POWHEG-BOX.

\textsuperscript{4}In this study we have limited ourselves to the matching uncertainty, however the latter should always be combined with the fixed-order perturbative uncertainty, usually estimated through the variation of the renormalization and factorization scale.

\textsuperscript{5}In the SM case, new developments are available which provide NLO-QCD accuracy in the description of the Higgs $p_{\perp}$, see refs. \cite{35, 36, 37, 38, 39}.
prescription used to define the allowed phase space for radiation emission by the PS. Restricting the phase space, as it happens already in the MC@NLO framework, allows also POWHEG to approach the fNLO at high-$p_{\perp}$, as we have shown with a dedicated analysis.

References


