DESY 13-122
MIT–CTP 4479

Jet p_T Resummation in Higgs Production at NNLL′+NNLO

Iain W. Stewart,[1] Frank J. Tackmann,[2] Jonathan R. Walsh,[3] and Saba Zuberi[3]

[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *Theory Group, Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany*
[3] *Ernest Orlando Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720*

(Dated: July 5, 2013)

We present predictions for Higgs production via gluon fusion with a p_T veto on jets and with the resummation of jet-veto logarithms at NNLL′+NNLO order. These results incorporate explicit $\mathcal{O}(\alpha_s^2)$ calculations of soft and beam functions, which include the dominant dependence on the jet radius R. In particular the NNLL′ order accounts for the correct boundary conditions for the N^{3}LL resummation, for which the only unknown ingredients are higher-order anomalous dimensions. We use scale variations in a factorization theorem in both rapidity and virtuality space to estimate the perturbative uncertainties, accounting for both higher fixed-order corrections as well as higher-order towers of jet-p_T logarithms. This formalism also predicts the correlations in the theory uncertainty between the exclusive 0-jet and inclusive 1-jet bins. At the values of R used experimentally, there are important corrections due to jet algorithm clustering that include logarithms of R. Although we do not sum logarithms of R, we do include an explicit contribution in our uncertainty estimate to account for higher-order jet clustering logarithms. Precision predictions for this $H+0$-jet cross section and its theoretical uncertainty are an integral part of Higgs analyses that employ jet binning.

arXiv:1307.1808v1 [hep-ph] 6 Jul 2013

I. INTRODUCTION

After the discovery of a Higgs boson [1, 2], a central objective of the LHC physics program is to measure the properties of the new particle by exploiting all accessible production and decay channels. The $gg \to H \to WW$ channel is very sensitive to the Higgs coupling to W gauge bosons. The $gg \to H \to \tau\tau$ channel provides direct sensitivity to the Higgs couplings to fermions and is the only measurable channel that gives direct access to the Higgs couplings in the leptonic sector of the Standard Model. In both these channels the experimental analyses separate the data into jet bins to take advantage of the fact that the signal over background ratio, as well as the dominant background contributions, strongly depend on the number of jets in the final state. Of particular importance is the 0-jet bin, where any hard jets are vetoed, as it contains the largest signal cross section.

Extracting the Higgs couplings from the measured exclusive 0-jet cross section requires precise theoretical predictions. Any type of jet veto introduces a veto scale, $k^{\rm cut}$. For a tight jet veto, $k^{\rm cut} \ll m_H$, large Sudakov logarithms of the veto scale, $\alpha_s^n \ln^m(k^{\rm cut}/m_H)$, appear in the perturbative series and must be resummed to all orders to obtain a meaningful perturbative prediction. For $k^{\rm cut} \sim m_H$, fixed-order perturbation theory can safely be applied, and the cross section with arbitrary cuts has been calculated at fixed next-to-next-to-leading order (NNLO) [3–6]. In the transition region between these two limits, both the veto logarithms and nonlogarithmic fixed-order corrections are numerically important, and a complete description including both types of perturbative corrections must be used to obtain the best possible theoretical precision. For earlier theoretical work on analytic resummation for Higgs jet vetoes see for example [7–14].

In principle, there are several different ways to implement a veto on additional emissions due to initial-state and final-state radiation in a given process. A "global jet veto" corresponds to a restriction applied to the sum of all radiation, for example through a global event shape such as beam thrust [15] [or equivalently $(N=0)$-jettiness [16]] or $E_T = \sum|p_T|$, and allows for precise resummed predictions [7, 15–18].

The current experimental analyses use a jet clustering algorithm (the anti-k$_{\rm T}$ algorithm [19] with a jet radius $R = 0.4$ for ATLAS and $R = 0.5$ for CMS) to identify jets. The jet veto is then implemented by requiring $p_T^{\rm jet} < p_T^{\rm cut}$ for any jets with $|\eta^{\rm jet}| < \eta^{\rm cut}$ (while jets at larger pseudorapidities are unrestricted). The typical experimental ranges are $p_T^{\rm cut} \sim 25 - 30\,{\rm GeV}$ for $\eta^{\rm cut} \sim 4.5 - 5$ (with the high value of $\eta^{\rm cut}$ having a small effect on the cross section). In contrast to a global jet veto, this procedure corresponds to a "local jet veto", since the restriction on final state radiation is applied separately to each individual local cluster of emissions.

For a cut on either $E_T < p_T^{\rm cut}$ or $p_T^{\rm jet} < p_T^{\rm cut}$, the jet-veto scale is set by p_T and Sudakov double logarithms of the ratio $p_T^{\rm cut}/m_H$ arise. The leading correction to the 0-jet cross section for Higgs production via gluon fusion has the form

$$\sigma_0(p_T^{\rm cut}) = \sigma_{\rm LO}\Big(1 - \frac{\alpha_s C_A}{\pi}\, 2\ln^2\frac{p_T^{\rm cut}}{m_H} + \dots\Big)\,, \qquad (1)$$

where $\sigma_{\rm LO}$ denotes the lowest-order cross section. The hierarchy between $p_T^{\rm cut}$ and m_H implies that resummation of logarithms of $p_T^{\rm cut}/m_H$ should be performed. For the $p_T^{\rm jet} \le p_T^{\rm cut}$ veto, the resummation of $p_T^{\rm cut}$-logarithms up to NNLL has been presented in Refs. [8–10].

In this paper, we calculate the resummed $H+0$-jet cross section from gluon fusion using the framework of soft-collinear effective theory (SCET) [20–24], where the cross section is factorized into calculable pieces and the

resummation is performed by renormalization group evolution (RGE) in both virtuality and rapidity space. We determine the cross section at NNLL$'_{p_T}$+NNLO order, where we use the subscript p_T to explicitly denote the fact that the resummation order only counts logarithms of $p_T^{\rm cut}/m_H$ (and not R^2). The primed order counting is described for example in Ref. [7]. It includes the NNLL resummation and in addition the full $\mathcal{O}(\alpha_s^2)$ dependence of the functions appearing in the factorization theorem (including in our case the $\mathcal{O}(\alpha_s^2)$ effects from jet clustering). These corrections incorporate the dominant NNLO corrections at small $p_T^{\rm cut}$ into the resummed result. They are formally part of the N^{3}LL resummation for which they provide the correct RGE boundary conditions. The missing ingredients for a complete N^{3}LL resummation are the unknown three-loop non-cusp and four-loop cusp anomalous dimensions. We also include the "nonsingular" $\mathcal{O}(\alpha_s^2)$ corrections that vanish as $p_T^{\rm cut} \to 0$, which are not part of the resummation. Thus our results incorporate the complete NNLO cross section for all values of $p_T^{\rm cut}$, including the total NNLO cross section in the limit of large $p_T^{\rm cut}$. This allows us to also obtain resummed predictions for the exclusive 0-jet event fraction (or efficiency) and the inclusive 1-jet cross section with a cut $p_T^{\rm jet} \geq p_T^{\rm cut}$ on the leading jet.

In our analysis, we place a particular emphasis on a careful estimate of the remaining perturbative uncertainties in our predictions. The different contributions to the uncertainty are estimated by appropriate variations of the different scales in virtuality and rapidity space appearing in the factorization theorem. This allows us to distinguish between and account for both resummation and fixed-order uncertainties. This formalism then automatically determines the correlations in the perturbative uncertainties between the total inclusive, exclusive 0-jet, and inclusive 1-jet cross sections.

The 0-jet cross section defined by $p_T^{\rm jet} \leq p_T^{\rm cut}$ has a complex dependence on the jet algorithm, whose effect is to introduce a nontrivial dependence on the jet radius R,

$$\ln \frac{\sigma_0^{(n)}(p_T^{\rm cut})}{\sigma_{\rm LO}} \supset \Big(\frac{\alpha_s}{4\pi}\Big)^n \ln \frac{m_H}{p_T^{\rm cut}}\, C_n(R)\,. \qquad (2)$$

For small R^2 the numerically most relevant terms contain logarithms of R^2 and at $\mathcal{O}(\alpha_s^n)$ are of the form [11]

$$C_n(R) \sim \ln^{n-1} R^2 + \ln^{n-2} R^2 + \ldots \ln R^2 + \mathcal{O}(R^2)\,. \quad (3)$$

They also contain subleading power corrections of $\mathcal{O}(R^2)$. The jet clustering effects start at $\mathcal{O}(\alpha_s^2)$ ($n = 2$) relative to $\sigma_{\rm LO}$. They were first calculated in Ref. [10] and at present are the only clustering corrections that are known. They turn out to have a sizeable effect on the cross section for jet radii $R = 0.4$ and 0.5. The large clustering effects for such small values of R imply that the logarithms of R^2 should be formally treated as being of similar size as the logarithms of $p_T^{\rm cut}/m_H$ and hence should also be resummed. In particular, as Eq. (2) shows, counting $\ln R^2 \sim \ln(p_T^{\rm cut}/m_H)$ implies that there are NLL terms from clustering at each order in α_s. However, the clustering coefficients $C_{n>2}(R)$ are unknown, and in principle a separate fixed-order calculation is required to obtain each one. This renders the resummation of the clustering logarithms intractable at present. In our analysis, we incorporate the known $\mathcal{O}(\alpha_s^2)$ clustering effects, calculate the $\mathcal{O}(\alpha_s^2)$ clustering effects that involve $\ln R^2$ without a $\ln(m_H/p_T^{\rm cut})$, and include an explicit contribution in our uncertainty estimate for unknown higher-order clustering terms.

The paper is laid out as follows: In Sec. II, we overview how the cross section is computed using SCET and give a summary of the results for each part of the cross section. In Sec. III, we discuss how the perturbative uncertainties are estimated through scale variation and how its various components are combined to estimate the total perturbative uncertainty in the 0-jet cross section, the 0-jet fraction, and the inclusive 1-jet cross section. In Sec. IV, we present the results of our numerical analysis and our predictions for the LHC for these cross sections. We conclude in Sec. V.

II. FACTORIZATION WITH A JET ALGORITHM AT SMALL R

The factorization of the $pp \to H$ + 0-jet cross section with a jet algorithm has been discussed in Refs. [8, 11]. For the case of the p_T veto on jets in Higgs production via gluon fusion, the factorized cross section is given by

$$\begin{aligned}
&\sigma_0(p_T^{\rm cut}) \\
&\quad = \sigma_B H_{gg}(m_t, m_H, \mu) \int {\rm d}Y\, B_g(m_H, p_T^{\rm cut}, R, x_a, \mu, \nu) \\
&\qquad \times B_g(m_H, p_T^{\rm cut}, R, x_b, \mu, \nu)\, S_{gg}(p_T^{\rm cut}, R, \mu, \nu) \\
&\qquad + \sigma_0^{R\rm sub}(p_T^{\rm cut}, R) + \sigma_0^{\rm ns}(p_T^{\rm cut}, R, \mu_{\rm ns})\,, \qquad (4)
\end{aligned}$$

where

$$x_{a,b} = \frac{m_H}{E_{\rm cm}}\, e^{\pm Y}\,, \quad \sigma_B = \frac{\sqrt{2} G_F\, m_H^2}{576\pi E_{\rm cm}^2}\,. \qquad (5)$$

The first term in Eq. (4) provides the leading contribution to the cross section at small $p_T^{\rm cut}$, and contains all the singular logarithmic terms $\alpha_s^i \ln^j(p_T^{\rm cut}/m_H)$. It is factorized into hard, beam, and soft functions, which are discussed below. For instance, the leading double logarithm in Eq. (1) is split up as

$$\ln^2 \frac{p_T^{\rm cut}}{m_H} = \ln^2 \frac{m_H}{\mu} + 2 \ln \frac{p_T^{\rm cut}}{\mu} \ln \frac{\nu}{m_H} + \ln \frac{p_T^{\rm cut}}{\mu} \ln \frac{\mu\, p_T^{\rm cut}}{\nu^2}\,, \qquad (6)$$

where the three terms on the right-hand side are contained in the hard, beam, and soft functions, respectively. In Eqs. (4) and (6) μ is the usual renormalization/factorization scale in virtuality, while ν denotes the corresponding scale in rapidity [25, 26]. Hence, we can

already see that both invariant mass and rapidity running will be needed to resum the $\ln(p_T^{\rm cut}/m_H)$ terms with renormalization group methods.

The resummation at NNLL$'$ requires determining the functions H_{gg}, B_g, and S_{gg} to $\mathcal{O}(\alpha_s^2)$, as well as identifying their anomalous dimensions to $\mathcal{O}(\alpha_s^2)$, and their cusp anomalous dimensions to $\mathcal{O}(\alpha_s^3)$. We present our new two-loop results for the soft and beam functions in Secs. II B and II C below, leaving the details of the calculations to a separate publication.

The second term in Eq. (4), $\sigma_0^{R\rm sub}(p_T^{\rm cut}, R)$, contains $\mathcal{O}(R^2)$ contributions whose all-orders soft-collinear factorization is challenging and not known at present. In the $R^2 \ll 1$ regime, these corrections can formally be treated as subleading power corrections. Numerically, they are indeed very small for the values $R \simeq 0.4$–0.5 which are of interest. (As explained in Ref. [11], counting $R \sim 1$, they would significantly complicate the soft-collinear factorization already at leading order in the power counting.) As shown in Ref. [9], their contribution to the NNLL series is obtained by multiplying them with the same evolution factor as the singular terms, and we will follow this same approach here. Their contribution to the resummed cross section is discussed in Sec. II D.

The last term in Eq. (4), $\sigma_0^{\rm ns}(p_T^{\rm cut}, R, \mu_{\rm ns})$, contains $\mathcal{O}(p_T^{\rm cut}/m_H)$ nonsingular corrections, which vanish for $p_T^{\rm cut} \to 0$ but become important at large $p_T^{\rm cut}$. These terms are added to the NNLL$'$ result and are required to reproduce the complete NNLO cross section and achieve the full NNLL$'$+NNLO accuracy. Our extraction and analysis of these terms is discussed in Sec. II E.

A. Hard Function

The hard function, H_{gg}, in Eq. (4) is determined by matching QCD onto the gluon fusion operator O_{ggH} in SCET. As discussed in detail in Ref. [7], this matching can be performed as a two-step matching [27–30] or a one-step matching. Since parametrically $m_H/m_t \simeq 1$, we employ the one-step matching, which also makes it easy to include the m_t dependence of the ggH form factor in the matching coefficient $C_{ggH}(m_t, m_H, \mu)$.

The hard matching coefficient satisfies the RG equation

$$\frac{\rm d}{{\rm d}\ln\mu} \ln\bigl[C_{ggH}(m_t, m_H, \mu)\bigr] = \gamma_H^g(m_H, \mu)\,, \qquad (7)$$

where the anomalous dimension has the structure

$$\gamma_H^g(m_H, \mu) = \Gamma^g_{\rm cusp}[\alpha_s(\mu)] \ln\frac{-m_H^2 - {\rm i}0}{\mu^2} + \gamma_H^g[\alpha_s(\mu)]\,. \qquad (8)$$

The solution of Eq. (7) yields the RGE of the matching coefficient from an initial scale μ_H to some final scale μ,

$$C_{ggH}(m_t, m_H, \mu) \qquad (9)$$
$$= C_{ggH}(m_t, m_H, \mu_H) \exp\biggl[\int_{\mu_H}^{\mu} \frac{{\rm d}\mu'}{\mu'}\, \gamma_H^g(m_H, \mu')\biggr]\,.$$

The hard function is then given by the absolute value squared of the RG evolved coefficient,

$$H_{gg}(m_t, m_H, \mu) = \bigl|C_{ggH}(m_t, m_H, \mu)\bigr|^2\,. \qquad (10)$$

For the resummation at NNLL$'$, we require the NNLO result for C_{ggH}, the two-loop result for the non-cusp hard anomalous dimension γ_H^g, and the three-loop result for the gluon cusp anomalous dimension $\Gamma^g_{\rm cusp}$ [31]. These results as well as the explicit NNLL expression for the evolution factor can be found in App. B of Ref. [7].

To all orders in perturbation theory the matching coefficient contains logarithms of the ratio $(-m_H^2 - {\rm i}0)/\mu_H^2$. Choosing a real value for the starting scale $\mu_H \sim m_H$ leaves large Sudakov double logarithms $\ln^2(-1 - {\rm i}0) = -\pi^2$, leading to a poorly convergent perturbative expansion of the hard function at this scale. Since these terms are associated with the logarithms in the matching coefficient, they can be summed through its RGE by using an imaginary starting scale $\mu_H \simeq -{\rm i}m_H$ [32–34]. In this way, the double logarithms are fully minimized, leading to a much better perturbative convergence [27, 35]. To illustrate this numerically, for $m_H = 125\,{\rm GeV}$ we find

$$\begin{aligned}
\frac{H_{gg}(\mu_H = m_H)}{H_{gg}^{(0)}(\mu_H = m_H)} &= 1 + 0.815 + 0.356 + \cdots\,,\\
\frac{H_{gg}(\mu_H = -{\rm i}m_H)}{H_{gg}^{(0)}(\mu_H = -{\rm i}m_H)} &= 1 + 0.274 + 0.042 + \cdots\,,
\end{aligned} \qquad (11)$$

where $H_{gg}^{(0)}$ is the lowest-order result in each case, and the second and third numbers on the right-hand side give the NLO and NNLO corrections, respectively. The substantial improvement in convergence also implies reduced perturbative uncertainties. We therefore use the imaginary hard scale as the default choice in our numerical results.

B. Soft Function

The soft function describes the soft radiation across the entire event. It is defined as a forward scattering matrix element of soft Wilson lines along the two incoming beam directions, with the jet-veto measurement on the final state,

$$S_{gg}(p_T^{\rm cut}, R, \mu, \nu) = \langle 0|Y_{n_b}\, Y_{n_a}^\dagger\, \mathcal{M}^{\rm jet}(p_T^{\rm cut}, R)\, Y_{n_a}\, Y_{n_b}^\dagger|0\rangle\,. \qquad (12)$$

Here, the measurement function $\mathcal{M}^{\rm jet}(p_T^{\rm cut}, R)$ acts on the soft final state by clustering it into jets of radius R and requiring that all these jets have $p_T < p_T^{\rm cut}$. This local veto on individual jets can be divided into a global veto and a local correction from the jet algorithm clustering, consequently dividing the soft function into a global

term and a jet algorithm correction[1],

$$S_{gg}(p_T^{\rm cut}, R, \mu, \nu) = S_{gg}^G(p_T^{\rm cut}, \mu, \nu) + \Delta S_{gg}^{\rm jet}(p_T^{\rm cut}, R, \mu, \nu)\,. \tag{13}$$

This isolates the jet algorithm effects into $\Delta S_{gg}^{\rm jet}$, which makes them easier to compute and analyze their resummation properties. Note that these jet algorithm corrections are defined relative to the chosen global veto, while the full soft function on the left-hand side is uniquely defined by specifying the jet-veto measurement, $\mathcal{M}^{\rm jet}(p_T^{\rm cut}, R)$. At $\mathcal{O}(\alpha_s^2)$, where the clustering corrections are first nonzero, the two-particle phase space constraints of the anti-k_T algorithm are identical to other k_T-type jet algorithms, which include k_T and Cambridge-Aachen [36–39]. This is also true for the jet algorithm effects in the beam function, and thus our calculation does not distinguish between these jet algorithms at the order to which we work.

The soft and beam functions separately contain rapidity divergences. When they are combined in the cross section, the rapidity divergences cancel, leaving large "rapidity logarithms" $\ln(p_T^{\rm cut}/m_H)$ at fixed order. We employ the rapidity renormalization group [25, 26], which allows one to apply standard effective theory and RG methods to regulate and renormalize the rapidity divergences and perform the resummation of the associated rapidity logarithms. It introduces an arbitrary rapidity renormalization scale ν, whose role in the rapidity RGE is the same as that of the usual renormalization scale μ in the standard virtuality RGE.

In our case, the soft function is multiplicatively renormalized in both μ and ν,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\ln\mu} \ln S_{gg}(p_T^{\rm cut}, R, \mu, \nu) &= \gamma_S^g(\mu, \nu)\,, \\
\frac{\mathrm{d}}{\mathrm{d}\ln\nu} \ln S_{gg}(p_T^{\rm cut}, R, \mu, \nu) &= \gamma_\nu^g(p_T^{\rm cut}, R, \mu)\,.
\end{aligned} \tag{14}$$

The anomalous dimensions have the general structure [11]

$$\begin{aligned}
\gamma_S^g(\mu, \nu) &= 4\Gamma^g_{\rm cusp}[\alpha_s(\mu)] \ln\frac{\mu}{\nu} + \gamma_S^g[\alpha_s(\mu)]\,, \\
\gamma_\nu^g(p_T^{\rm cut}, R, \mu) &= -4\eta_\Gamma^g(p_T^{\rm cut}, \mu) + \gamma_\nu^g[\alpha_s(p_T^{\rm cut}), R]\,,
\end{aligned} \tag{15}$$

where

$$\eta_\Gamma^g(\mu_0, \mu) = \int_{\mu_0}^{\mu} \frac{\mathrm{d}\mu'}{\mu'}\, \Gamma^g_{\rm cusp}[\alpha_s(\mu')] = \Gamma^g_{\rm cusp} \ln\frac{\mu}{\mu_0} + \cdots \tag{16}$$

[1] Technically, this division into global and clustering contributions is affected by the fact that non-Abelian exponentiation occurs for the soft function, and only specifies how the genuinely new terms at each perturbative order are divided. Since the first nontrivial clustering correction only arises at $\mathcal{O}(\alpha_s^2)$, Eq. (13) holds for the soft function through NNLO. The exponentiation of lower-order results will mix global and clustering contributions at higher orders in the soft function.

sums an all-orders set of terms in the anomalous dimension that are determined by the RG consistency. (They are required to ensure the exact path independence of the evolution in the two-dimensional μ-ν space [26].) The RGE of the soft function is obtained by solving Eq. (14). Evolving first in rapidity and then in virtuality, we have

$$\begin{aligned}
&S_{gg}(p_T^{\rm cut}, R, \mu, \nu) \\
&\quad = S_{gg}(p_T^{\rm cut}, R, \mu_S, \nu_S) \exp\biggl[\ln\frac{\nu}{\nu_S}\, \gamma_\nu^g(p_T^{\rm cut}, R, \mu_S)\biggr] \\
&\qquad \times \exp\biggl[\int_{\mu_S}^{\mu} \frac{\mathrm{d}\mu'}{\mu'}\, \gamma_S^g(\mu', \nu)\biggr]\,.
\end{aligned} \tag{17}$$

We have calculated the complete soft function to $\mathcal{O}(\alpha_s^2)$, which to our knowledge is the first two-loop calculation employing the rapidity renormalization. Our result for the perturbative soft function through $\mathcal{O}(\alpha_s^2)$ is

$$\begin{aligned}
&S_{gg}(p_T^{\rm cut}, R, \mu_S, \nu_S) = \\
&\quad 1 + \frac{\alpha_s(\mu_S)}{4\pi}\Bigl[2\Gamma_0^g L_S^\mu (L_S^\mu - 2L_S^\nu) - \frac{\pi^2}{3} C_A\Bigr] \\
&\qquad + \frac{\alpha_s^2(\mu_S)}{(4\pi)^2}\biggl\{\frac{1}{2}\Bigl[2\Gamma_0^g L_S^\mu (L_S^\mu - 2L_S^\nu) - \frac{\pi^2}{3} C_A\Bigr]^2 \\
&\qquad + 2\beta_0 L_S^\mu \Bigl[2\Gamma_0^g L_S^\mu \Bigl(\frac{1}{3} L_S^\mu - L_S^\nu\Bigr) - \frac{\pi^2}{3} C_A\Bigr] \\
&\qquad + 2\Gamma_1^g L_S^\mu (L_S^\mu - 2L_S^\nu) \\
&\qquad + \gamma_{S\,1}^g L_S^\mu + \gamma_{\nu\,1}^g(R)\, L_S^\nu + s_2(R)\biggr\}\,,
\end{aligned} \tag{18}$$

where we abbreviated

$$L_S^\mu \equiv \ln\frac{\mu_S}{p_T^{\rm cut}}\,, \qquad L_S^\nu \equiv \ln\frac{\nu_S}{p_T^{\rm cut}}\,. \tag{19}$$

Hence, the natural soft scales for which the large logarithms in the soft function are minimized are $\mu_S \sim p_T^{\rm cut}$ and $\nu_S \sim p_T^{\rm cut}$.

In Eq. (18) and in the following, the β function and anomalous dimensions are expanded as

$$\begin{aligned}
\beta(\alpha_s) &= -2\alpha_s \sum_{n=0}^{\infty} \beta_n \Bigl(\frac{\alpha_s}{4\pi}\Bigr)^{n+1}\,, \\
\gamma(\alpha_s) &= \sum_{n=0}^{\infty} \gamma_n \Bigl(\frac{\alpha_s}{4\pi}\Bigr)^{n+1}\,,
\end{aligned} \tag{20}$$

where the coefficients needed in Eq. (18) are

$$\begin{aligned}
\beta_0 &= \frac{11}{3} C_A - \frac{4}{3} T_F\, n_f\,, \\
\beta_1 &= \frac{34}{3} C_A^2 - \Bigl(\frac{20}{3} C_A + 4C_F\Bigr) T_F\, n_f\,, \\
\Gamma_0^g &= 4C_A\,, \\
\Gamma_1^g &= 4C_A\Bigl[\Bigl(\frac{67}{9} - \frac{\pi^2}{3}\Bigr) C_A - \frac{20}{9} T_F\, n_f\Bigr]\,,
\end{aligned} \tag{21}$$

and $C_A = 3$, $C_F = 4/3$, $T_F = 1/2$, and $n_f = 5$ is the number of light quark flavors. The coefficients β_2 and Γ_2^g are also used in the NNLL resummation.

At one loop, the non-cusp soft and rapidity anomalous dimensions vanish,

$$\gamma_{S\,0}^g = 0\,, \qquad \gamma_{\nu\,0}^g(R) = 0\,. \tag{22}$$

The dependence on the jet algorithm starts to enter at two loops through the two-loop ν anomalous dimension, $\gamma_{\nu\,1}^g(R)$, which determines the coefficient of the single logarithm of $\ln(\nu/p_T^{\rm cut})$, as well as the nonlogarithmic two-loop soft constant, $s_2(R)$. For the two-loop coefficients of the non-cusp anomalous dimensions we find

$$\begin{aligned}
\gamma_{S\,1}^g &= 8C_A\biggl[\Bigl(\frac{52}{9} - 4(1+\pi^2)\ln 2 + 11\zeta_3\Bigr)C_A \\
&\quad + \Bigl(\frac{2}{9} + \frac{7\pi^2}{12} - \frac{20}{3}\ln 2\Bigr)\beta_0\biggr] \\
&= 16C_A^2\,(-3.83)\,, \\
\gamma_{\nu\,1}^g(R) &= -16C_A\biggl[\Bigl(\frac{17}{9} - (1+\pi^2)\ln 2 + \zeta_3\Bigr)C_A \\
&\quad + \Bigl(\frac{4}{9} + \frac{\pi^2}{12} - \frac{5}{3}\ln 2\Bigr)\beta_0\biggr] + C_2(R) \\
&= 16C_A^2\,(4.16) + C_2(R)\,.
\end{aligned} \tag{23}$$

Here, $C_2(R)$ is the clustering correction due to the jet algorithm, and was computed earlier in Ref. [11]. It is given by

$$\begin{aligned}
C_2(R) &= 2C_A\Bigl[\Bigl(1 - \frac{8\pi^2}{3}\Bigr)C_A + \Bigl(\frac{23}{3} - 8\ln 2\Bigr)\beta_0\Bigr]\ln R^2 \\
&\quad + 15.62C_A^2 - 9.17C_A\beta_0 + C_2^{R\rm sub}(R) \\
&= 16C_A^2\,\bigl(-2.49\ln R^2 - 0.49\bigr) + \mathcal{O}(R^2)\,,
\end{aligned} \tag{24}$$

where $C_2^{R\rm sub}(R) \sim \mathcal{O}(R^2)$ contains all subleading power corrections in R^2. Note that we define the clustering effects in $C_2(R)$ relative to the global E_T veto. A different choice, such as the p_T of the Higgs used in Ref. [10], would give a different R-independent constant in $C_2(R)$. Nevertheless, the full result for $\gamma_{\nu 1}^g(R)$ is independent of this choice and our final NNLL cross section agrees with that of Ref. [10].

For the two-loop soft function constant $s_2(R)$, which is not determined from RGE constraints, we find

$$\begin{aligned}
s_2(R) &= C_A\biggl[\Bigl(\frac{19}{3} - 10\ln 2 + 8\zeta_3\Bigr)C_A \\
&\quad + \Bigl(-\frac{163}{9} + \frac{58}{3}\ln 2 + 8\ln^2 2\Bigr)\beta_0\biggr]\ln R^2 \\
&\quad - 18.68C_A^2 - 3.25C_A\beta_0 + s_2^{R\rm sub}(R) \\
&= 16C_A^2\,\bigl(0.43\ln R^2 - 1.69\bigr) + \mathcal{O}(R^2)\,,
\end{aligned} \tag{25}$$

where $s_2^{R\rm sub}(R) \sim R^2$. This result for $s_2(R)$ is new and also constitutes the first calculation of the $p_T^{\rm cut}$ independent clustering terms in the soft function.

The terms not proportional to $\ln R^2$ in $C_2(R)$ and $s_2(R)$ involve complicated phase-space integrals, which are computed numerically. The contributions of $\gamma_{\nu\,1}^g(R)$ and $s_2(R)$ to the fixed NNLO cross section including their full R dependence are shown in Fig. 1.

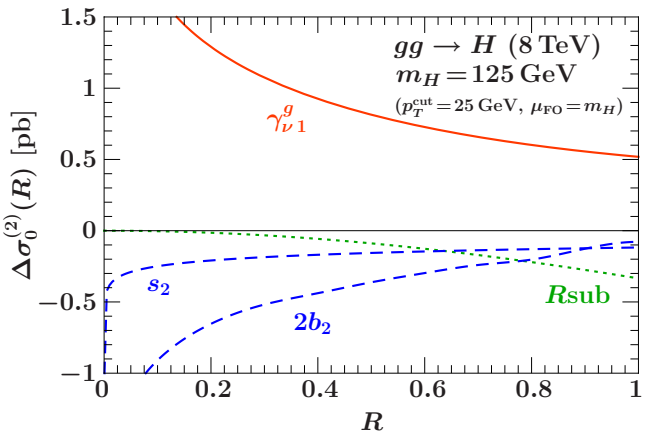

FIG. 1: Jet-algorithm dependent $\mathcal{O}(\alpha_s^2)$ contributions to the fixed NNLO cross section from different sources, for $\mu_{\rm FO} = m_H$ and $p_T^{\rm cut} = 25\,\text{GeV}$. The ν anomalous dimension coefficient $\gamma_{\nu\,1}^g$ is given in Eq. (23), the $\mathcal{O}(\alpha_s^2)$ soft function constant terms in Eq. (25), the beam function constant terms in Eq. (39) and the following paragraph, and the clustering effects on uncorrelated emissions in Eq. (40).

As mentioned above, the jet algorithm corrections in the soft function start at $\mathcal{O}(\alpha_s^2)$. They have the all-order structure

$$\begin{aligned}
&\Delta S_{gg}^{\rm jet}(p_T^{\rm cut}, R, \mu_S, \nu_S) \\
&\quad = \sum_{n\geq 2}\frac{\alpha_s^n(\mu_S)}{(4\pi)^n}\Bigl[C_n(R)\ln\frac{\nu_S}{p_T^{\rm cut}} + \Delta s_n(R)\Bigr]\,,
\end{aligned} \tag{26}$$

where $C_n(R)$ and $\Delta s_n(R)$ contain up to $n-1$ powers of $\ln R^2$. The $C_n(R)$ in the soft function are the same as in Eq. (2) for the cross section. The beam functions contain an equivalent set of terms $\sim \alpha_s^n C_n(R)\ln(m_H/\nu_B)$. In the fixed-order cross section (i.e. for $\nu_B = \nu_S = \nu$) they combine with the soft function terms to give the total clustering correction $\sim \alpha_s^n C_n(R)\ln(m_H/p_T^{\rm cut})$ in Eq. (2). For $R^2 \sim p_T^{\rm cut}/m_H$, the leading $\ln^{n-1} R^2$ terms in $C_n(R)$ formally count as NLL in the exponent of the cross section. Similarly, the leading $\ln^{n-1} R^2$ terms in $\Delta s_n(R)$, as well as the $\ln^{n-2} R^2$ terms in $C_n(R)$, formally count as NNLL. The anomalous dimension $\gamma_\nu^g(R)$ includes the $C_n(R)$, so its perturbative series explicitly contains the $\ln R^2$ terms, which means that the NLL and higher logarithmic series from $\ln R^2$ clustering corrections are not resummed here. A formalism for this resummation is not currently known. Since these clustering corrections are numerically large at $\mathcal{O}(\alpha_s^2)$, we perform an estimate of the potential size of the higher-order clustering effects as part of our uncertainty analysis.

C. Beam Function

The beam function is defined as the forward proton matrix element of collinear gluon fields. It provides a combined description of collinear initial-state radiation from the incoming gluons together with their extraction from the colliding protons via the nonperturbative parton distribution functions (PDFs)[15].

Like the soft function, the beam function is multiplicatively renormalized in both μ and ν,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\ln\mu}\ln B_g(m_H, p_T^{\mathrm{cut}}, R, x, \mu, \nu) &= \gamma_B^g(m_H, \mu, \nu)\,, \\ \frac{\mathrm{d}}{\mathrm{d}\ln\nu}\ln B_g(m_H, p_T^{\mathrm{cut}}, R, x, \mu, \nu) &= -\frac{1}{2}\gamma_\nu^g(p_T^{\mathrm{cut}}, R, \mu)\,.\end{aligned} \tag{27}$$

The anomalous dimensions can be determined from those of the hard and soft functions using the consistency of the factorization theorem. The ν anomalous dimension, γ_ν^g, is the same as in Eq. (15). The μ anomalous dimension is given by

$$\begin{aligned}\gamma_B^g(m_H, \mu, \nu) &= 2\Gamma^g_{\mathrm{cusp}}[\alpha_s(\mu)]\ln\frac{\nu}{m_H} + \gamma_B^g[\alpha_s(\mu)]\,, \\ \gamma_B^g(\alpha_s) &= -\gamma_H^g(\alpha_s) - \frac{1}{2}\gamma_S^g(\alpha_s)\,,\end{aligned} \tag{28}$$

with the resulting one-loop and two-loop coefficients

$$\begin{aligned}\gamma^g_{B\,0} &= 2\beta_0\,, \\ \gamma^g_{B\,1} &= 2\beta_1 + 8C_A\biggl[\Bigl(-\frac{5}{4} + 2(1+\pi^2)\ln 2 - 6\zeta_3\Bigr)C_A \\ &\quad + \Bigl(\frac{5}{24} - \frac{\pi^2}{3} + \frac{10}{3}\ln 2\Bigr)\beta_0\biggr]\,.\end{aligned} \tag{29}$$

The RGE of the beam function follows from solving Eq. (27), and is analogous to that of the soft function,

$$\begin{aligned}&B_g(m_H, p_T^{\mathrm{cut}}, R, x, \mu, \nu) \\ &\quad = B_g(m_H, p_T^{\mathrm{cut}}, R, x, \mu_B, \nu_B) \\ &\qquad \times \exp\biggl[\frac{1}{2}\ln\frac{\nu_B}{\nu}\,\gamma_\nu^g(p_T^{\mathrm{cut}}, R, \mu_B)\biggr] \\ &\qquad \times \exp\biggl[\int_{\mu_B}^{\mu}\frac{\mathrm{d}\mu'}{\mu'}\,\gamma_B^g(m_H, \mu', \nu)\biggr]\,.\end{aligned} \tag{30}$$

Note that in contrast to the PDF evolution, the evolution of the beam function does not change its value of x. This is a general feature of beam functions and is due to the fact that their evolution describes the initial-state radiation from an incoming parton that is not confined to the proton anymore, while the PDF evolution is frozen out at the beam scale μ_B [15, 40].

At the beam scale, the gluon beam function can be computed as a convolution between perturbative matching kernels, $\mathcal{I}_{gj}(m_H, p_T^{\mathrm{cut}}, z, \mu_B, \nu_B)$, and the standard quark and gluon PDFs, $f_j(x, \mu_B)$,

$$\begin{aligned}&B_g(m_H, p_T^{\mathrm{cut}}, R, x, \mu_B, \nu_B) \\ &\quad = \sum_j \int_x^1 \frac{\mathrm{d}z}{z}\,\mathcal{I}_{gj}(m_H, p_T^{\mathrm{cut}}, R, z, \mu_B, \nu_B)\, f_j\Bigl(\frac{x}{z}, \mu_B\Bigr)\,.\end{aligned} \tag{31}$$

We expand the matching kernels $\mathcal{I}_{gj}$ to $\mathcal{O}(\alpha_s^2)$ as (suppressing the arguments for brevity)

$$\mathcal{I}_{gj} = \delta_{gj}\delta(1-z) + \frac{\alpha_s(\mu_B)}{4\pi}\,\mathcal{I}_{gj}^{(1)} + \frac{\alpha_s^2(\mu_B)}{(4\pi)^2}\,\mathcal{I}_{gj}^{(2)} + \mathcal{O}(\alpha_s^3)\,. \tag{32}$$

The $\mathcal{O}(\alpha_s)$ coefficients are common to several observables, and we agree with the calculation of $\mathcal{I}_{gg}^{(1)}$ using the rapidity regulator in Ref. [26]. We find,

$$\begin{aligned}\mathcal{I}_{gg}^{(1)}(m_H, p_T^{\mathrm{cut}}, z, \mu_B, \nu_B) &= 4C_A L_B^\mu \\ &\quad \times \bigl[2L_B^\nu\delta(1-z) - P_{gg}(z)\bigr]\,, \\ \mathcal{I}_{gq}^{(1)}(m_H, p_T^{\mathrm{cut}}, z, \mu_B, \nu_B) &= 2C_F\bigl[-2L_B^\mu P_{gq}(z) + I_{gq}^{(1)}(z)\bigr]\,, \\ I_{gq}^{(1)}(z) &= z\,,\end{aligned} \tag{33}$$

where we abbreviated

$$L_B^\mu \equiv \ln\frac{\mu_B}{p_T^{\mathrm{cut}}}\,, \qquad L_B^\nu \equiv \ln\frac{\nu_B}{m_H}\,. \tag{34}$$

The natural scales for the beam function are thus $\mu_B \sim p_T^{\mathrm{cut}}$ and $\nu_B \sim m_H$. Our results for $\mathcal{I}_{gj}^{(1)}$ agree with Ref. [8], after taking into account the different rapidity regularization.

The $\mathcal{O}(\alpha_s^2)$ kernel for the gg contribution is given by

$$\begin{aligned}&\mathcal{I}_{gg}^{(2)}(m_H, p_T^{\mathrm{cut}}, R, z, \mu_B, \nu_B) \\ &\quad = 32C_A^2(L_B^\mu)^2 L_B^\nu\bigl[L_B^\nu\delta(1-z) - P_{gg}(z)\bigr] \\ &\qquad + 4C_A\beta_0(L_B^\mu)^2\bigl[2L_B^\nu\delta(1-z) - P_{gg}(z)\bigr] \\ &\qquad + 8(L_B^\mu)^2\bigl[C_A^2(P_{gg}\otimes P_{gg})(z) \\ &\qquad\quad + 2C_F T_F n_f(P_{gq}\otimes P_{qq})(z)\bigr] \\ &\qquad - 8L_B^\mu\bigl[P_{gg}^{(1)}(z) + 2C_F T_F n_f(I_{gq}^{(1)}\otimes P_{qq})(z)\bigr] \\ &\qquad + \Bigl[L_B^\mu\bigl(2\Gamma_1^g L_B^\nu + \gamma^g_{B\,1}\bigr) - \frac{1}{2}\gamma^g_{\nu\,1}(R)L_B^\nu\Bigr]\delta(1-z) \\ &\qquad + I_{gg}^{(2)}(z, R)\,.\end{aligned} \tag{35}$$

The $\mathcal{O}(\alpha_s^2)$ kernel for the gq contribution is given by

$$\begin{aligned}&\mathcal{I}_{gq}^{(2)}(m_H, p_T^{\mathrm{cut}}, R, z, \mu_B, \nu_B) \\ &\quad = 16C_A C_F L_B^\mu L_B^\nu\bigl[-2L_B^\mu P_{gq}(z) + I_{gq}^{(1)}(z)\bigr] \\ &\qquad + 8C_F\beta_0 L_B^\mu\bigl[-L_B^\mu P_{gq}(z) + I_{gq}^{(1)}(z)\bigr] \\ &\qquad + 8C_F(L_B^\mu)^2\bigl[C_A(P_{gg}\otimes P_{gq})(z) \\ &\qquad\quad + C_F(P_{gq}\otimes P_{qq})(z)\bigr] \\ &\qquad - 8L_B^\mu\bigl[P_{gq}^{(1)}(z) + C_F^2(I_{gq}^{(1)}\otimes P_{qq})(z)\bigr] + I_{gq}^{(2)}(z, R)\,.\end{aligned} \tag{36}$$

The convolutions $(g\otimes h)(z)$ are defined as

$$(g\otimes h)(z) \equiv \int_z^1 \frac{\mathrm{d}\xi}{\xi}\,g\Bigl(\frac{z}{\xi}\Bigr)h(\xi)\,. \tag{37}$$

The various splitting functions $P_{ij}(z)$ and convolutions between them are given in App. B2 of Ref. [7]. The

additional convolutions we need are

$$(I_{gq}^{(1)} \otimes P_{qg})(z) = 1 + z - 2z^2 + 2z\ln z\,, \qquad (38)$$
$$(I_{gq}^{(1)} \otimes P_{qq})(z) = 1 + \frac{z}{2} - z\ln z + 2z\ln(1-z)\,.$$

The terms involving logarithms of μ and ν in the $\mathcal{I}_{gj}$ kernels are fully determined by renormalization group (RG) constraints. The nonlogarithmic terms $I_{gi}^{(2)}(z,R)$ require the full two-loop calculation of the beam functions. Note that the full two-loop qq contribution to the beam function for the transverse momentum of the vector boson has been computed recently in Ref. [41]. At two loops, the $p_T^{\rm jet}$ beam function needed here is different and requires a separate calculation. Like the soft function, it receives both global and jet clustering contributions. In particular, we can calculate directly the leading clustering corrections proportional to $\ln R^2$, and determine the contribution from the remaining terms numerically, giving

$$I_{gg}^{(2)}(z,R) = \frac{C_A}{2}\biggl[\Bigl(1-\frac{8\pi^2}{3}\Bigr)C_A + \Bigl(\frac{23}{3} - 8\ln 2\Bigr)\beta_0\biggr] \times P_{gg}(z)\ln R^2 + I_{gg}^{(2,c)}(z) + I_{gg}^{(2,R\rm sub)}(z,R)\,,$$
$$I_{gq}^{(2)}(z,R) = 2C_F^2\Bigl(3 - \frac{\pi^2}{3} - 3\ln 2\Bigr)P_{gq}(z)\ln R^2 + I_{gq}^{(2,c)}(z) + I_{gq}^{(2,R\rm sub)}(z,R)\,. \qquad (39)$$

Here, $I_{gg}^{(2,c)}(z)$ denotes the constant R independent terms, while $I_{gg}^{(2,R\rm sub)}(z,R)$ are the $\mathcal{O}(R^2)$ suppressed contributions. Their explicit form is not known at present. We extract their total contribution after convolution with the PDFs numerically from the fixed-order cross section as explained in Sec. II E below. This is sufficient for practical purposes, since their effect is found to be numerically small compared to the $\ln R^2$ terms for $R \sim 0.4$–0.5. The total contribution (from both beam functions) of the full $I_{gg}^{(2)}(z,R)$ and $I_{gq}^{(2)}(z,R)$ to the fixed NNLO cross section is shown by the blue dashed line in Fig. 1 that is labeled as $2b_2$.

D. $\mathcal{O}(R^2)$ Corrections From Uncorrelated Emissions

Starting at $\mathcal{O}(\alpha_s^2)$, the clustering effects from the jet algorithm includes contributions that scale as powers of R^2 in the small R limit. Clustering effects from correlated emissions in the soft or collinear sectors are included in the subleading $\mathcal{O}(R^2)$ corrections in the soft and beam functions. On the other hand, the clustering of uncorrelated emissions from the soft and collinear beam sectors inhibits the factorization of the jet-veto measurement into independent soft and collinear measurements at $\mathcal{O}(R^2)$. The all-order factorization of the cross section at this level is therefore not known at present.[2]

The full contribution from clustering of uncorrelated emissions to the fixed NNLO cross section is [10]

$$\sigma_0^{(2)}(p_T^{\rm cut}) \supset \sigma_{\rm LO}\Bigl(\frac{\alpha_s C_A}{\pi}\Bigr)^2 \ln\frac{m_H}{p_T^{\rm cut}}\Bigl(-\frac{\pi^2}{3}R^2 + \frac{R^4}{4}\Bigr)\,. \qquad (40)$$

It is shown by the green dotted line in Fig. 1 for $p_T^{\rm cut} = 25\,\text{GeV}$. As one can see, at the R values of interest it is numerically very small compared to the corresponding $\ln R^2$ enhanced clustering corrections contained in $\gamma_{\nu 1}^g(R)$, and can thus safely be treated as a power correction.

As argued in Refs. [9, 10], the above $\mathcal{O}(\alpha_s^2)$ coefficient determines the complete NNLL series coming from this contribution [i.e. no new coefficients appear at $\mathcal{O}(\alpha_s^3 L^2)$ or higher]. Therefore, we can include this correction in the resummed cross section at NNLL by multiplying it with the total evolution factor as follows,

$$\sigma_0^{R\rm sub}(p_T^{\rm cut},R) = \frac{\alpha_s^2(\sqrt{\mu_B\mu_S})}{\pi^2}C_A^2\ln\frac{m_H}{p_T^{\rm cut}}\Bigl(-\frac{\pi^2}{3}R^2 + \frac{R^4}{4}\Bigr) \times [F^{(0)}U_0](\mu_H,\mu_B,\mu_S,\nu_B,\nu_S)\,. \qquad (41)$$

Here, $F^{(0)}$ denotes the leading fixed-order contributions from the hard, beam, and soft functions, and U_0 is their combined NNLL evolution factor [given explicitly in Eq. (57) below]. Since these corrections come from soft or collinear emissions we choose to evaluate the argument of the α_s^2 in the prefactor at the geometric mean of the beam and soft scales.

In Ref. [8] this coefficient is absorbed into the two-loop rapidity anomalous dimension, which amounts to

[2] The statement in Ref. [8] that soft-collinear mixing is absent at leading power for $R \sim 1$ relies on a power counting for collinear rapidities (y_c) and soft rapidities (y_s) where $y_c \gg y_s \sim \mathcal{O}(1)$ such that $y_c - y_s \gg R \sim 1$. For typical values of $p_T = 25\,\text{GeV}$ and $Q = 125\,\text{GeV}$ there is a legitimate power expansion in $\lambda = p_T/Q = 0.2 \ll 1$. But this gives $y_c \simeq \ln(1/\lambda) = 1.6$, which does not clearly satisfy $y_c \gg y_s \sim 1$. Indeed, physically, emissions at fixed p_T tend to be uniform in rapidity rather than having a rapidity gap between soft and collinear regions. (The analogous statement using light-cone variables is $e^{-R} \gg e^{y_s - y_c} = \sqrt{(k_s^-/k_s^+)(p_c^+/p_c^-)} \sim \mathcal{O}(1)\times\lambda$. For $R = 1$, this corresponds to counting $0.37 \gg \mathcal{O}(1)\times\lambda = \mathcal{O}(1)\times 0.2$.) As discussed in detail in Ref. [11], the contribution from clustering a soft and a collinear emission is $\sim R^2$, so the only way to expand it to zero is $R^2 \ll 1$.

The fact that soft and collinear modes in SCET-II are only distinguished by their rapidity does not automatically imply that their rapidities are parametrically widely separated as $y_s \ll y_c$, since in practice amplitudes from each of these modes are integrated over all rapidities and we must worry about contributions from overlapping regions. If there is a double counting for infrared singularities from the overlap region then this is removed by 0-bin subtractions [42], but in general these subtractions do not suffice to remove finite contributions from the overlap region. Thus a proof of factorization at $\mathcal{O}(R^2)$, including also soft-collinear mixing contributions, will require additional arguments to all orders in α_s, and remains an interesting open question.

writing this contribution as $A\exp(\alpha_s^2)$, instead of $A(1+\alpha_s^2)$ as in Ref. [9]. Since this contribution first appears at $\mathcal{O}(\alpha_s^2)$, either form gives the same NNLL contribution and the difference is higher order, meaning the results of Refs. [9, 10] do not determine which is the correct all-order structure beyond NNLL.

E. Nonsingular Contributions

In fixed-order perturbation theory, the cross section at $\mu_f = \mu_r = \mu_{\rm FO}$ has the all-order structure

$$\sigma_0^{\rm FO}(p_T^{\rm cut},\mu_{\rm FO}) = \sigma_0^{\rm s}(p_T^{\rm cut},\mu_{\rm FO}) + \sigma_0^{\rm ns}(p_T^{\rm cut},\mu_{\rm FO})\,, \quad (42)$$
$$\sigma_0^{\rm s}(p_T^{\rm cut},\mu_{\rm FO}) = \sum_m \sum_{n\le 2m} c_{mn}(\mu_{\rm FO})\,\alpha_s^m(\mu_{\rm FO})\ln^n\frac{p_T^{\rm cut}}{m_H}\,.$$

Here, the singular cross section, $\sigma_0^{\rm s}$, contains all terms that are nonzero for $p_T^{\rm cut}\to 0$ and which are contained in the resummed cross section. The nonsingular cross section, $\sigma_0^{\rm ns}$, scales as $\mathcal{O}(p_T^{\rm cut}/m_H)$ and vanishes for $p_T^{\rm cut}\to 0$. To reproduce the full fixed-order cross section we have to include the nonsingular terms, in particular when going to large $p_T^{\rm cut}$ where they become important.

An important feature of the NNLL′ (NLL′) resummed result is that by construction its fixed-order expansion to NNLO (NLO) in terms of $\alpha_s(\mu_{\rm FO})$ can be obtained by simply setting all scales equal to $\mu_{\rm FO}$. And this also precisely reproduces the fixed-order singular contributions. Hence, we can determine the nonsingular corrections by subtracting the latter from the full fixed-order cross section,

$$\sigma_0^{\rm ns}(p_T^{\rm cut},R,\mu_{\rm FO}) = \sigma_0^{\rm FO}(p_T^{\rm cut},R,\mu_{\rm FO}) - \sigma_0^{\rm resum'}(p_T^{\rm cut},R,\mu_i=\nu_i=\mu_{\rm FO})\,. \quad (43)$$

At NLO, this procedure is straightforward since the one-loop hard, beam, and soft functions required at NLL′ are completely known, while $\sigma_0^{\rm NLO}(p_T^{\rm cut},\mu_{\rm FO})$ is easily obtained numerically e.g. from MCFM.

At NNLO, we obtain the full fixed-order cross section by subtracting the NLO $gg\to H+j$ cross section for a leading jet with $p_T^{\rm jet}>p_T^{\rm cut}$, obtained using MCFM [5, 43], from the total NNLO cross section [44–46]. For the resummed NNLL′ cross section we include all available contributions through $\mathcal{O}(\alpha_s^2)$ summarized in the previous subsections, including the $\sigma_0^{R\rm sub}$ terms in Eq. (41). The only missing pieces at two loops are the unknown $I_{gj}^{(2,c)}(z)$ and $I_{gj}^{(2,R\rm sub)}(z,R)$ terms in the beam function, which when integrated against the PDFs give a $p_T^{\rm cut}$ independent contribution determined by a constant $b_2^{(c+R\rm sub)}(R)$. Hence, we have

$$\sigma_0^{\rm ns,NNLO}(p_T^{\rm cut},R,\mu_{\rm FO}) + \sigma_{\rm LO}(\mu_{\rm FO})\frac{\alpha_s^2(\mu_{\rm FO})}{(4\pi)^2}2b_2^{(c+R\rm sub)}(R) = [\sigma_{\ge 0}^{\rm NNLO}(\mu_{\rm FO}) - \sigma_{\ge 1}^{\rm NLO}(p_T^{\rm cut},R,\mu_{\rm FO})] - \sigma_0^{\rm NNLL'}(p_T^{\rm cut},R,\mu_i=\nu_i=\mu_{\rm FO})\,. \quad (44)$$

Here, the right-hand side is obtained numerically and then fit with a set of functions suitable to describe the $p_T^{\rm cut}$ dependence of $\sigma_0^{\rm ns}(p_T^{\rm cut},R,\mu_{\rm FO})$. Since the latter vanishes for $p_T^{\rm cut}\to 0$, this fit also allows us to determine the numerical value of $2b_2^{(c+R\rm sub)}(R)$ from the intercept at $p_T^{\rm cut}=0$. Note that since there are large numerical cancellations between the full and singular results on the right-hand side, the remaining nonsingular data has large statistical fluctuations for $p_T^{\rm cut}\to 0$. Ensuring a stable fit result therefore required the use of very high statistics from MCFM as well as a careful validation of the fitting procedure.

Note also that the scale μ_B at which the $b_2^{(c+R\rm sub)}$ contribution is evaluated in the beam function is relevant at NNLL′ (i.e. it contributes to the subset of N^{3}LL effects that are supposed to be included at NNLL′). In the numerical determination above the PDFs are evaluated at a fixed $\mu_B=\mu_{\rm FO}$. To account for this we rescale it by the PDF dependence of the LO cross section, as indicated in Eq. (44). Since we perform the nonsingular fit at different values of $\mu_{\rm FO}$, we are able to check that this captures the PDF scale dependence to very good approximation.

At large $p_T^{\rm cut}$, the distinction between singular and nonsingular contributions becomes meaningless since both are of similar size and there are nontrivial cancellations between them (as can be seen in Fig. 4 below). When using the imaginary scale setting in the hard function, it modifies the cross section at all values of $p_T^{\rm cut}$. Therefore, it is important to implement an analogous improvement for the nonsingular contributions, since otherwise these cancellations would be spoiled. The final expression for the nonsingular cross section entering in Eq. (4) is given by

$$\sigma_0^{\rm ns}(p_T^{\rm cut},R,\mu_{\rm ns}) = \Big\{\sigma_0^{\rm ns(1)}(p_T^{\rm cut},R,\mu_{\rm ns})\Big[1-\frac{\alpha_s(\mu_{\rm ns})}{2\pi}C_A\pi^2\Big] + \sigma_0^{\rm ns(2)}(p_T^{\rm cut},R,\mu_{\rm ns})\Big\}U_H(-{\rm i}\mu_{\rm ns},\mu_{\rm ns})\,. \quad (45)$$

Here, $\sigma_0^{\rm ns(i)}(p_T^{\rm cut},R,\mu_{\rm ns})$ are the $\mathcal{O}(\alpha_s^i)$ nonsingular terms obtained numerically for given values of R and $\mu_{\rm ns}$, and $U_H(-{\rm i}\mu_{\rm ns},\mu_{\rm ns})$ is the evolution factor of the hard function. The latter is used to apply the analogous resummation of π^2 terms to the nonsingular cross section as was induced by the hard function in the singular terms.

The NLO and NNLO nonsingular contributions for $R=0.4$ and $\mu_{\rm ns}=m_H$ are shown in Fig. 2 for both real (left panel) and imaginary (right panel) scale setting. We observe that the latter substantially improves the perturbative convergence also in the nonsingular terms at all values of $p_T^{\rm cut}$. This is not unexpected from the point of view of the power expansion in SCET. For $p_T^{\rm cut}\ll m_H$ and at subleading order in the SCET power counting, the nonsingular terms would arise from a combination of leading and subleading hard, beam, and soft functions, and many of the hard functions in these contributions can be expected to require an imaginary hard scale.

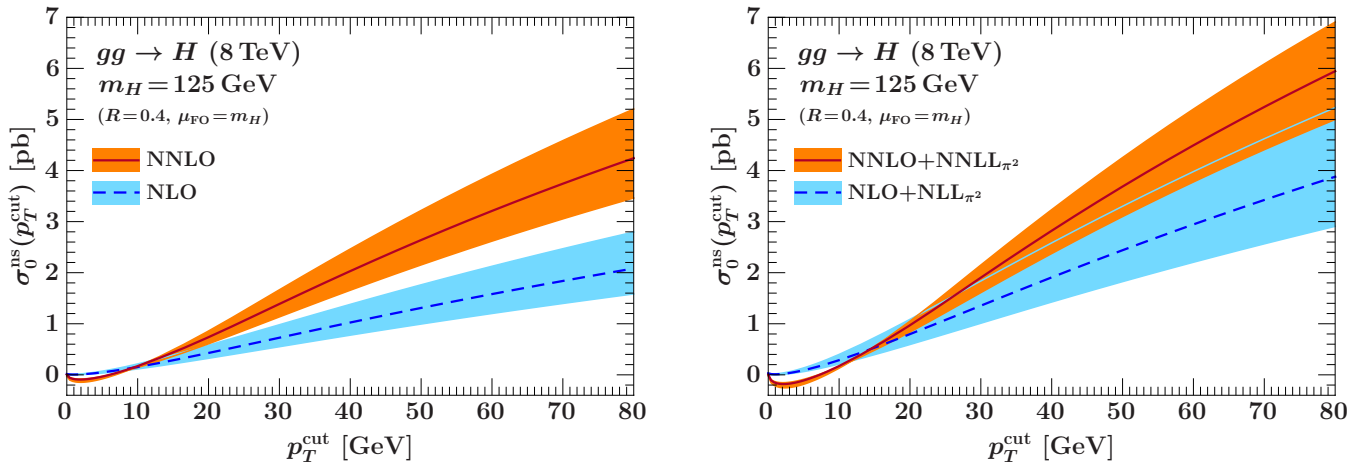

FIG. 2: The nonsingular cross section at $\mu_{\rm ns} = m_H$ at NLO (blue, dashed) and NNLO (orange, solid) for $R = 0.4$. We compare the pure fixed-order nonsingular terms (on the left) with the nonsingular terms that include π^2 summation (on the right). The latter shows a substantially improved perturbative convergence from NLO to NNLO.

III. RESUMMATION AND PERTURBATIVE UNCERTAINTIES

A critical aspect of precision cross section predictions is the theoretical control of perturbative uncertainties. Ultimately, the formal perturbative accuracy in the predictions is only meaningful together with a robust understanding and estimate of theoretical uncertainties.

The categorization of the data into jet bins is used in the experimental analyses to optimize the control of backgrounds and experimental systematic effects. In the end, the information from all measured categories flows together, thereby maximizing the use of the available data. In this context, vetoing jets in the 0-jet cross section amounts to dividing the total inclusive cross section, $\sigma_{\rm tot} \equiv \sigma_{\geq 0}$, into an exclusive 0-jet bin equivalent to $\sigma_0(p_T^{\rm cut})$ and the remaining inclusive 1-jet bin,

$$\sigma_{\geq 0} = \sigma_0(p_T^{\rm cut}) + \sigma_{\geq 1}(p_T^{\rm cut})\,. \tag{46}$$

Therefore, a complete theoretical description of this binning procedure is needed. This requires a framework, which, in addition to the resummation of $\sigma_0(p_T^{\rm cut})$ at small $p_T^{\rm cut}$, provides a valid description of the cross section at all values of $p_T^{\rm cut}$ as well as the correlations between the perturbative uncertainties in the jet bins and the total cross section.

As we discuss in detail in this section, the framework we use for resummation and fixed-order matching, based on SCET and profile functions, is well-suited for this task. It provides us with direct theoretical handles to reliably assess the perturbative uncertainties and allows us to predict the required correlations by utilizing common underlying theory parameters in the scales μ_H, μ_B, μ_S, ν_B, and ν_S. These are varied to obtain the uncertainty estimates.

In Sec. III A we give an overview of perturbative uncertainties for jet bins, and establish the necessary notation. As the jet-veto cut is increased our resummed results smoothly reproduce the fixed-order cross section and its standard uncertainties by using profile functions, which are discussed in Sec. III B. In Sec. III C we explain how variations of the hard, soft, and beam scales in the effective theory determine the fixed-order and jet-binning uncertainties. Finally, in Sec. III D we discuss our estimate for the additional uncertainty from clustering effects at higher orders in perturbation theory. Note that we will not discuss additional parametric uncertainties from input parameters such as PDFs or $\alpha_s(m_Z)$. These have to be estimated separately and included with the usual uncertainty propagation.

A. Perturbative Uncertainties in Jet Binning

A convenient way to describe the uncertainties involved in the jet binning is in terms of fully correlated and fully anticorrelated components [12, 47], which amounts to parametrizing the covariance matrix for $\{\sigma_0, \sigma_{\geq 1}\}$ as

$$C(\{\sigma_0, \sigma_{\geq 1}\}) = \begin{pmatrix} (\Delta_0^{\rm y})^2 & \Delta_0^{\rm y}\,\Delta_{\geq 1}^{\rm y} \\ \Delta_0^{\rm y}\,\Delta_{\geq 1}^{\rm y} & (\Delta_{\geq 1}^{\rm y})^2 \end{pmatrix} + \begin{pmatrix} \Delta_{\rm cut}^2 & -\Delta_{\rm cut}^2 \\ -\Delta_{\rm cut}^2 & \Delta_{\rm cut}^2 \end{pmatrix}. \tag{47}$$

The first correlated component, denoted with a superscript "y", can be interpreted as an overall yield uncertainty shared among all bins. The second anticorrelated component can be interpreted as a migration uncertainty between the two bins, which is introduced by the binning cut and drops out in their sum. The total uncertainty for each bin is given by

$$\begin{aligned} \Delta_{\geq 0} &= \Delta_0^{\rm y} + \Delta_{\geq 1}^{\rm y} \equiv \Delta_{\geq 0}^{\rm y}\,, \\ \Delta_0^2 &= (\Delta_0^{\rm y})^2 + \Delta_{\rm cut}^2\,, \\ \Delta_{\geq 1}^2 &= (\Delta_{\geq 1}^{\rm y})^2 + \Delta_{\rm cut}^2\,. \end{aligned} \tag{48}$$

Equation (47) is a completely generic parametrization of a 2×2 symmetric matrix. This choice of parameters is convenient because of the above physical interpretation. An additional advantage is that the uncertainties are described in terms of two independent components, which are fully correlated or anticorrelated between the different observables, so that the experimental implementation is straightforward (e.g. in a profile likelihood fit, the yield and migration uncertainties can each be implemented by an independent nuisance parameter).

To estimate each uncertainty component in our resummation framework we make the following identifications:

$$\Delta_i^{\rm y} \equiv \Delta_{\mu i}\,, \qquad \Delta_{\rm cut} \equiv \Delta_{\rm resum}\,. \tag{49}$$

Here, $\Delta_{\mu i}$ corresponds to the uncertainties in the cross section that reproduce the fixed-order uncertainty in the total cross section and probe the nonlogarithmic contributions at finite $p_T^{\rm cut}$. This makes it natural to identify these with the yield uncertainties. The resummation uncertainty, $\Delta_{\rm resum}$, corresponds to the intrinsic uncertainty in the resummed logarithmic series. The logarithms $\ln(p_T^{\rm cut}/m_H)$ are directly caused by the binning cut and at small $p_T^{\rm cut}$ are the dominant veto-dependent effect, which cancels between σ_0 and $\sigma_{\geq 1}$. Hence, higher-order logarithms are the primary source of uncertainty in the division of the cross section into bins and we can therefore identify $\Delta_{\rm resum}$ with the migration uncertainty. Furthermore, $\Delta_{\rm resum}$ vanishes at large $p_T^{\rm cut}$ where the resummation of logarithms becomes unimportant. This is consistent with the fact that in this limit migration effects become irrelevant since $\sigma_{\geq 1}$ becomes numerically much smaller than $\sigma_0(p_T^{\rm cut})$. Our procedure to estimate $\Delta_{\mu i}$ and $\Delta_{\rm resum}$ through scale variations in the resummed cross section is discussed in the following sections.

With these identifications, the full covariance matrix for $\{\sigma_{\geq 0},\sigma_0,\sigma_{\geq 1}\}$ is given by

$$C\big(\{\sigma_{\geq 0},\sigma_0,\sigma_{\geq 1}\}\big) = C_\mu + C_{\rm resum}\,, \tag{50}$$

where

$$C_\mu = \begin{pmatrix} \Delta_{\rm tot}^2 & \Delta_{\rm tot}\Delta_{\mu 0} & \Delta_{\rm tot}\Delta_{\mu\geq 1} \\ \Delta_{\rm tot}\Delta_{\mu 0} & \Delta_{\mu 0}^2 & \Delta_{\mu 0}\Delta_{\mu\geq 1} \\ \Delta_{\rm tot}\Delta_{\mu\geq 1} & \Delta_{\mu 0}\Delta_{\mu\geq 1} & \Delta_{\mu\geq 1}^2 \end{pmatrix},$$
$$C_{\rm resum} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Delta_{\rm resum}^2 & -\Delta_{\rm resum}^2 \\ 0 & -\Delta_{\rm resum}^2 & \Delta_{\rm resum}^2 \end{pmatrix}, \tag{51}$$

and we can easily read off the uncertainties in the different cross sections

$$\begin{aligned} \Delta_{\rm tot} \equiv \Delta_{\mu\geq 0} &= \Delta_{\mu 0} + \Delta_{\mu\geq 1}\,, \\ \Delta_0^2(p_T^{\rm cut}) &= \Delta_{\rm resum}^2 + \Delta_{\mu 0}^2\,, \\ \Delta_{\geq 1}^2(p_T^{\rm cut}) &= \Delta_{\rm resum}^2 + (\Delta_{\rm tot} - \Delta_{\mu 0})^2\,. \end{aligned} \tag{52}$$

The uncertainties in other observables follow by standard uncertainty propagation. For example, for the 0-jet efficiency, $\epsilon_0(p_T^{\rm cut}) \equiv \sigma_0(p_T^{\rm cut})/\sigma_{\geq 0}$, we have

$$\frac{\Delta_{\epsilon_0}^2(p_T^{\rm cut})}{\epsilon_0^2(p_T^{\rm cut})} = \frac{\Delta_0^2(p_T^{\rm cut})}{\sigma_0^2(p_T^{\rm cut})} + \frac{\Delta_{\rm tot}^2}{\sigma_{\rm tot}^2} - 2\frac{\Delta_{\rm tot}\Delta_{\mu 0}}{\sigma_{\geq 0}\sigma_0(p_T^{\rm cut})}\,. \tag{53}$$

Through the last term the correlation between $\Delta_{\rm tot}$ and $\Delta_{\mu 0}$ reduces the relative uncertainty in the 0-jet efficiency, which will be noticeable in our numerical analysis. In particular, in the limit of large $p_T^{\rm cut}$ where $\epsilon_0 \to 1$ the uncertainty Δ_{ϵ_0} will go to zero as it should.

1. *Fixed Order*

In a pure fixed-order prediction, there is no way to fully disentangle the two uncertainty components. Using a common fixed-order scale variation for all observables amounts to setting $\Delta_{\rm cut} = 0$ and setting $\Delta_i^{\rm y} \equiv \Delta_i^{\rm FO}$. However, as demonstrated in detail in Refs. [12, 47], at small values of $p_T^{\rm cut}$, as soon as the logarithmic corrections become sizable, migration effects are important and cannot be neglected. Doing so can lead to a significant underestimate of uncertainties. A more reliable fixed-order estimate is obtained by explicitly taking into account $\Delta_{\rm cut}$ by using instead

$$\Delta_0^{\rm y} = \Delta_{\geq 0}^{\rm FO} \equiv \Delta_{\rm tot}\,, \qquad \Delta_{\rm cut} = \Delta_{\geq 1}^{\rm FO}\,, \tag{54}$$

where $\Delta_{\geq i}^{\rm FO}$ are the fixed-order uncertainties in the inclusive cross sections. (As explained in Ref. [12], this choice is motivated by the fact that the perturbative series in $\sigma_{\geq 1}$ starts as $\alpha_s \ln^2(p_T^{\rm cut}/m_H)$ and its fixed-order scale variation therefore directly estimates the size of the $p_T^{\rm cut}$ logarithms. An alternative prescription proposed in Ref. [10] yields very similar results for $\Delta_0(p_T^{\rm cut})$.)

With the choice in Eq. (54) the uncertainties in the pure fixed-order prediction are described by

$$C_{\rm ST}\big(\{\sigma_{\geq 0},\sigma_0,\sigma_{\geq 1}\}\big) = \begin{pmatrix} \Delta_{\rm tot}^2 & \Delta_{\rm tot}^2 & 0 \\ \Delta_{\rm tot}^2 & \Delta_{\rm tot}^2 + (\Delta_{\geq 1}^{\rm FO})^2 & -(\Delta_{\geq 1}^{\rm FO})^2 \\ 0 & -(\Delta_{\geq 1}^{\rm FO})^2 & (\Delta_{\geq 1}^{\rm FO})^2 \end{pmatrix}. \tag{55}$$

These are the default fixed-order Higgs jet-binning uncertainties used by the experiments, and also what we will use when comparing our results to fixed order in Sec. IV.

B. Resummation and Matching to Fixed Order with Profile Scales

In the effective field theory framework, the resummation is performed by RGE running. First, we evaluate each of the hard, beam, and soft functions appearing in the factorized cross section at their natural virtuality scales μ_i and rapidity scales ν_i. Next, we evolve them all to arbitrary, common scales: μ for invariant mass and ν for rapidity. This resums the logarithms of the invariant mass ratios μ_i/μ_j and rapidity ratios ν_i/ν_j. As

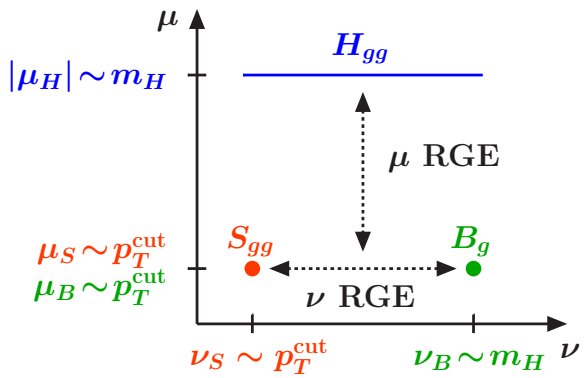

FIG. 3: Combined renormalization group evolution in virtuality and rapidity. The hard, beam, and soft functions are evolved in the virtuality scale μ, where the characteristic scales are $\mu_H \sim m_H$ and $\mu_B \sim \mu_S \sim p_T^{\rm cut}$. Additionally, rapidity logarithms are summed by evolving the beam and soft functions in the rapidity scale ν, with characteristic scales $\nu_B \sim m_H$ and $\nu_S \sim p_T^{\rm cut}$.

we saw in Sec. II, the beam and soft functions evolve in both virtuality and rapidity space, while the hard function only evolves in virtuality. The evolution together with the natural scales is illustrated in Fig. 3. Finally, the evolved functions are combined together in the cross section at the common scales (ν, μ), which is a point in the plane shown in this figure.

The resummed cross section is explicitly independent of the arbitrary scales μ and ν at each order in resummed perturbation theory, which means we are free to pick any convenient values. Taking $\mu = \mu_B$ and $\nu = \nu_S$, and combining all the ingredients detailed in Sec. II, the cross section in Eq. (4) takes the form

$$\begin{aligned}
&\sigma_0(p_T^{\rm cut}) \\
&\quad = \sigma_B H_{gg}(m_t, m_H, \mu_H) \int \mathrm{d}Y\, B_g(m_H, p_T^{\rm cut}, R, x_a, \mu_B, \nu_B) \\
&\qquad \times B_g(m_H, p_T^{\rm cut}, R, x_b, \mu_B, \nu_B)\, S_{gg}(p_T^{\rm cut}, R, \mu_S, \nu_S) \\
&\qquad \times U_0(p_T^{\rm cut}, R; \mu_H, \mu_B, \mu_S, \nu_B, \nu_S) \\
&\qquad + \sigma_0^{R\rm sub}(p_T^{\rm cut}, R) + \sigma_0^{\rm ns}(p_T^{\rm cut}, R, \mu_{\rm ns})\,, \qquad (56)
\end{aligned}$$

where the combined renormalization group evolution factor U_0 is given by

$$\begin{aligned}
&U_0(p_T^{\rm cut}, R; \mu_H, \mu_B, \mu_S, \nu_B, \nu_S) \\
&\quad = \left| \exp\left[\int_{\mu_H}^{\mu_B} \frac{\mathrm{d}\mu'}{\mu'}\, \gamma_H^g(m_H, \mu') \right] \right|^2 \\
&\qquad \times \exp\left[\int_{\mu_S}^{\mu_B} \frac{\mathrm{d}\mu'}{\mu'}\, \gamma_S^g(\mu', \nu_S) \right] \\
&\qquad \times \exp\left[\ln \frac{\nu_B}{\nu_S}\, \gamma_\nu^g(p_T^{\rm cut}, R, \mu_B) \right]. \qquad (57)
\end{aligned}$$

Next, we discuss how to choose numerical values for the scales $\mu_H, \mu_B, \mu_S, \nu_B$, and ν_S as a function of $p_T^{\rm cut}$, which are referred to as profile scales [48, 49]. For this purpose we can distinguish three different regimes according to the relative importance of the singular and nonsingular cross section contributions. In Fig. 4, the singular and nonsingular terms are plotted against the total fixed-order cross section at $\mathcal{O}(\alpha_s^2)$.

In the resummation region at low values of $p_T^{\rm cut}$, the singular contributions dominate and must be resummed, while the nonsingular contributions are perturbative power corrections. To resum the logarithms, the scales should parametrically follow their canonical values dictated by the RGE,

$$\begin{aligned}
&\mu_H \sim -\mathrm{i} m_H\,, && \mu_B \sim \mu_S \sim p_T^{\rm cut}\,, \\
&\nu_B \sim m_H\,, && \nu_S \sim p_T^{\rm cut}\,. \qquad (58)
\end{aligned}$$

At large $p_T^{\rm cut} \gtrsim m_H/2$, the singular and nonsingular contributions are equally important, and fixed-order perturbation theory should be used. In this fixed-order region it is essential that the resummation is turned off to ensure that the correct fixed-order cross section is obtained. The reason is that there are important cancellations between singular and nonsingular terms, which are spoiled if the resummation is kept on too long. In this region, all virtuality scales must approach a common fixed-order scale and the rapidity scales must be equal,

$$|\mu_H| = \mu_B = \mu_S = \mu_{\rm ns} = \mu_{\rm FO}\,, \qquad \nu_B = \nu_S\,. \qquad (59)$$

Finally, in the transition between the resummation and fixed-order regions typically both the logarithmic resummation as well as the fixed-order corrections are important. To obtain a proper description of this transition region, which in our case also includes the experimental range of interest, we have to use profiles that incorporate the constraints imposed by the resummation toward small $p_T^{\rm cut}$ and the fixed-order matching toward large $p_T^{\rm cut}$, together with a smooth interpolation between these two regimes. There is a growing body of literature on the construction of appropriate profiles in a variety of contexts [7, 13, 14, 48–56].

For the central profiles we take

$$\begin{aligned}
&\mu_H = -\mathrm{i}\mu_{\rm FO}\,, && \mu_{\rm ns} = \mu_{\rm FO}\,, \\
&\nu_B = \mu_{\rm FO}\,, \\
&\mu_B = \mu_S = \nu_S = \mu_{\rm FO}\, f_{\rm run}(p_T^{\rm cut}/m_H)\,. \qquad (60)
\end{aligned}$$

That is, we take fixed values for μ_H, $\mu_{\rm ns}$, and ν_B, while μ_B, μ_S, and ν_S are constructed in terms of the common profile function

$$f_{\rm run}(x) = \begin{cases}
x_0\left[1 + (x/x_0)^2/4\right] & x \le 2x_0\,, \\
x & 2x_0 \le x \le x_1\,, \\
x + \dfrac{(2 - x_2 - x_3)(x - x_1)^2}{2(x_2 - x_1)(x_3 - x_1)} & x_1 \le x \le x_2\,, \\
1 - \dfrac{(2 - x_1 - x_2)(x - x_3)^2}{2(x_3 - x_1)(x_3 - x_2)} & x_2 \le x \le x_3\,, \\
1 & x_3 \le x\,.
\end{cases} \qquad (61)$$

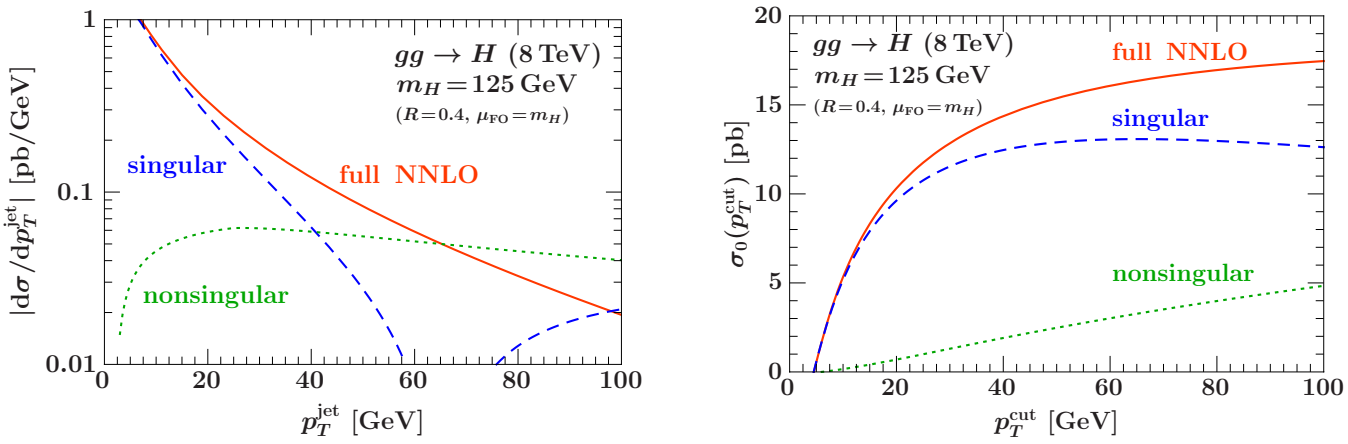

FIG. 4: Singular and nonsingular contributions to the fixed NNLO cross section (using $R = 0.4$ and $\mu_{\rm FO} = m_H$). Left: The magnitude of the contributions differential in $p_T^{\rm jet}$. Right: The corresponding contributions to the integrated cross section as a function of $p_T^{\rm cut}$. The resummation, transition, and fixed-order regions are clearly visible as the relative importance of the singular and nonsingular terms changes with $p_T^{\rm jet}$ and $p_T^{\rm cut}$.

The first regime, $x \leq 2x_0$, is the nonperturbative region and the scales $\mu_{B,S}$ and ν_S asymptote as $x \to 0$ to a fixed scale $x_0 \mu_{\rm FO} \gtrsim \Lambda_{\rm QCD}$. This ensures that factors of $\alpha_s(\mu_i)$ that enter from solving perturbatively defined anomalous dimension equations, never become nonperturbative. The second regime has the canonical scaling for resummation. The third and fourth have quadratic scaling (of positive and negative second derivative, respectively) and simply provide a smooth transition to the final (constant) region where all scales are equal and resummation is turned off. This profile function and its first derivative are both continuous.

For the overall scale parameter we have $\mu_{\rm FO} \sim m_H$ and for our central result we will use $\mu_{\rm FO} = m_H$ in Eq. (60). In Eq. (61) the parameters x_i mark the boundary between the different regimes, and their values are chosen by considering the importance of the singular versus nonsingular contributions plotted in Fig. 4. The singular and nonsingular contributions become comparable near $p_T^{\rm cut} = 40\,\text{GeV}$ so the profile must transition towards the fixed-order result beyond this value. For our central profiles we choose

$$x_0 = 2.5\,\text{GeV}/\mu_{\rm FO}\,, \quad \{x_1, x_2, x_3\} = \{0.15,\, 0.4,\, 0.65\}\,. \tag{62}$$

For $\mu_{\rm FO} = m_H = 125\,\text{GeV}$ the $\{x_1, x_2, x_3\}$ values correspond to $\{19, 50, 81\}\,\text{GeV}$. The resulting central profile scales are shown in Fig. 5, so we see that the transition occurs roughly between 30–65 GeV. In the next subsection, we discuss in detail the profile scale variations that we use to evaluate perturbative uncertainties.

Note that in the transition from small to large $p_T^{\rm cut}$, we are essentially forced to keep the hard scale at its imaginary value $\mu_H = -im_H$. In principle, one could contemplate rotating it to the real axis as a function of $p_T^{\rm cut}$ to turn off the resulting resummation of large π^2 terms in the hard virtual corrections. However, this

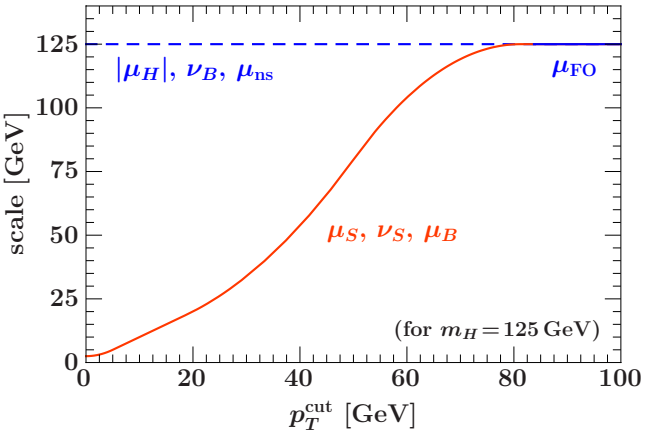

FIG. 5: The central profile scale for the low scales μ_B, μ_S, ν_S as a function of $p_T^{\rm cut}$, together with the central value for the high scales $|\mu_H|, \nu_B$.

would inevitably lead to an unphysical result of a decreasing cross section with increasing $p_T^{\rm cut}$. What this means is that the significantly improved perturbative stability observed in the small p_T region also directly translates into an improved convergence in the fixed-order cross section at large $p_T^{\rm cut}$, simply because a large part of the total cross section comes from the small p_T region. Furthermore, as we have seen in Fig. 2, the imaginary scale also translates into an improved convergence of the nonsingular contributions themselves. The total cross section for $\mu_H = -im_H$ increases by about 7% compared to the NNLO cross section evaluated at $\mu_{\rm FO} = m_H/2$. This increase is quite consistent with the expected increase in the total cross section at N^{3}LO from the recent estimate in Ref. [57].

C. Yield and Resummation Uncertainties Via Profile Scale Variations

To evaluate the perturbative uncertainties in our predictions we vary the profile scales about the central profiles defined in the previous section. We consider several types of variation in turn, and discuss how they are used to determine the yield and resummation uncertainties that appear in the matrices C_μ and $C_{\rm resum}$ in Eq. (51).

The first type of variation is a collective variation of all of the scales up or down by a factor of 2. This is accomplished by taking $\mu_{\rm FO} = 2m_H$ or $\mu_{\rm FO} = m_H/2$ in Eq. (60). At large $p_T^{\rm cut}$, where all scales become equal to $\mu_{\rm FO}$, this variation becomes equivalent to the usual scale variation in the fixed-order cross section. Indeed, in the limit of very large $p_T^{\rm cut}$ it reproduces the fixed-order scale variation of the total cross section.[3] When varying $\mu_{\rm FO}$, all scale ratios are kept fixed, so this does not change any of the arguments inside the logarithms $\ln(\mu_H/\mu_{B,S})$ and $\ln(\nu_B/\nu_S)$ that sum up the large $\ln(m_H/p_T^{\rm cut})$ terms. Hence, this variation is clearly identified as contributing to the yield uncertainties.

A second type of variation is to the profile shape. The values $\{x_1, x_2, x_3\}$ determine the boundaries between the different scaling regions of the low-scale profiles as a function of $p_T^{\rm cut}$. We account for the ambiguity in this shape by using four different choices for $\{x_1, x_2, x_3\}$ to provide a variation away from the central scale choice $\{x_1, x_2, x_3\} = \{0.15, 0.4, 0.65\}$:

$$\begin{aligned}\{x_1, x_2, x_3\} \; : &\{0.1, 0.3, 0.5\}\,, \{0.2, 0.5, 0.8\}\,, \\ &\{0.04, 0.4, 0.8\}\,, \{0.2, 0.35, 0.5\}\,. \end{aligned} \tag{63}$$

These changes to the profile have an impact on the uncertainty from varying $\mu_{\rm FO}$ since they determine the transition between the region where the resummation is active and where the fixed-order prediction is used and hence the extent of the fixed-order region. They also vary the logarithms $\ln(\mu_H/\mu_{B,S})$, and hence have some impact on uncertainties that would be associated to resummation. In practice, with $\mu_{\rm FO} = m_H$ the effect of varying the x_i in the central profile is smaller than the other resummation uncertainties (discussed below), whereas when varying $\mu_{\rm FO}$ up and down there is a noticeable impact on the yield uncertainties. Therefore we will group this variation with the yield uncertainty, and use each of the five profiles specified by $\{x_1, x_2, x_3\}$ together with each of the three values of $\mu_{\rm FO}$. This set of profile variations is plotted in the left panel of Fig. 6. We still note that the range of cross section values obtained from changing $\mu_{\rm FO}$ with a fixed profile is significantly larger than the range from changing the profile via $x_{1,2,3}$ for a fixed $\mu_{\rm FO}$, and hence the $\mu_{\rm FO}$ variation is the more important variation by far.

The total yield uncertainty for the 0-jet cross section is thus defined as the maximum absolute deviation from the central scale over all 14 variations,

$$\Delta_{\mu 0}(p_T^{\rm cut}) = \max_{v_i \in V_\mu} \left| \sigma_0^{v_i}(p_T^{\rm cut}) - \sigma_0^{\rm central}(p_T^{\rm cut}) \right| . \tag{64}$$

where V_μ is the set of variations. To determine the total uncertainty in the fixed-order cross section we make use of the fact that $\lim_{p_T^{\rm cut}\to\infty} \Delta_{\mu 0}(p_T^{\rm cut}) = \Delta_{\rm tot}$, and in practice we extract $\Delta_{\rm tot}$ for $p_T^{\rm cut} = 600\,\text{GeV}$. Together this determines the two parameters occurring in the yield covariance matrix C_μ.

Resummation uncertainties are estimated through variations of the beam and soft scales, while keeping $\mu_{\rm FO} = m_H$ at its central value. The variations of the beam and soft scales are performed with a multiplicative variation factor $f_{\rm vary}(p_T^{\rm cut})$. For a generic beam or soft scale μ_i or ν_i, the up and down variations are performed via the variations

$$\begin{aligned} \mu_i^{\rm up}(p_T^{\rm cut}) &= \mu_i^{\rm central}(p_T^{\rm cut}) \times f_{\rm vary}(p_T^{\rm cut}/m_H)\,, \\ \mu_i^{\rm down}(p_T^{\rm cut}) &= \mu_i^{\rm central}(p_T^{\rm cut}) \,/\, f_{\rm vary}(p_T^{\rm cut}/m_H)\,, \\ \nu_i^{\rm up}(p_T^{\rm cut}) &= \nu_i^{\rm central}(p_T^{\rm cut}) \times f_{\rm vary}(p_T^{\rm cut}/m_H)\,, \\ \nu_i^{\rm down}(p_T^{\rm cut}) &= \nu_i^{\rm central}(p_T^{\rm cut}) \,/\, f_{\rm vary}(p_T^{\rm cut}/m_H)\,. \end{aligned} \tag{65}$$

The variation factor is defined by

$$f_{\rm vary}(x) = \begin{cases} 2(1 - x^2/x_3^2) & 0 \le x \le x_3/2\,, \\ 1 + 2(1 - x/x_3)^2 & x_3/2 \le x \le x_3\,, \\ 1 & x_3 \le x\,. \end{cases} \tag{66}$$

It is designed to smoothly turn off these variations, since they must turn off when the resummation is turned off at high $p_T^{\rm cut}$ values. These variations for μ_B, μ_S, ν_B, and ν_S are plotted in the right panel of Fig. 6.

The resummation uncertainty is a combination of a set of up, down, and central values for the μ_B, μ_S, ν_B, and ν_S scales. The dependence on each of these scales cancels between RG evolution and the fixed-order contributions at the order we are working, while the remaining residual dependence probes the higher-order contributions in resummed perturbation theory.

The purpose of an individual scale variation is to vary the argument of the logarithms it appears in by a factor in order to probe the potential size of higher-order logarithms of that scale. For our profiles the variation factor above is 1/2 or 2 for $p_T^{\rm cut} \to 0$ and goes towards 1 for $p_T^{\rm cut} \to x_3 m_H$ where the resummation is turned off. Certain combinations of scale variations are undesirable as they double the variations of the logarithms, for example $\{\nu_B^{\rm up}, \nu_S^{\rm down}\}$ gives a factor of 4 variation for the logarithm of ν_B/ν_S. To avoid varying the scales in logarithms outside of the desired factor of 2 range, we consider all the ratios of beam and soft scales that appear in

[3] For $p_T^{\rm cut} > x_3 m_H$ and real $\mu_H = \mu_{\rm FO}$ we exactly reproduce the fixed-order cross section scale variation for equal factorization and renormalization scales. If these two scales are varied independently they give essentially the same final result since the renormalization scale variation dominates by far.

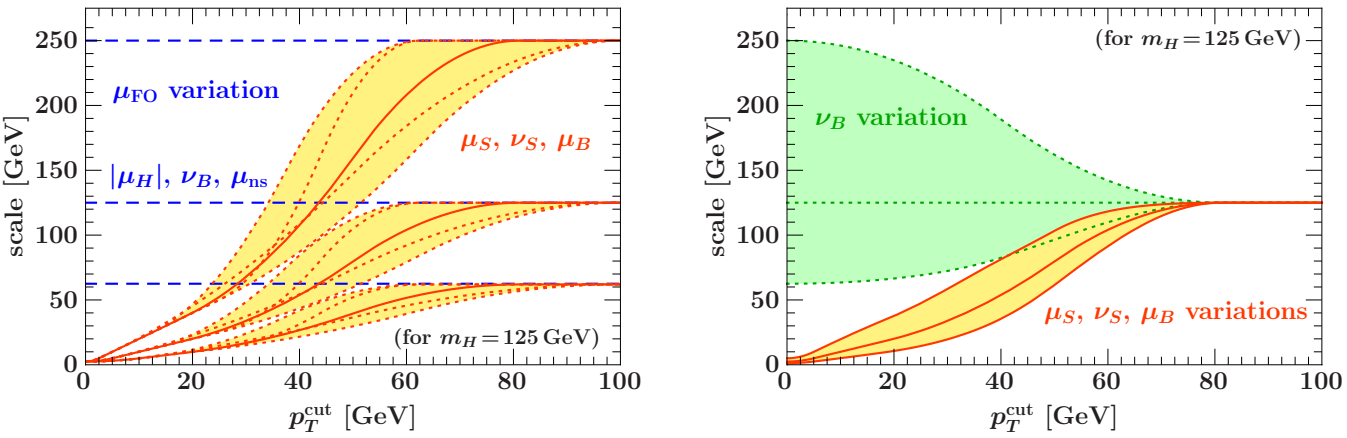

FIG. 6: The variations of the central profiles as described in the text. On the left, the variations are shown that contribute to the yield uncertainty, where all scales are collectively multiplied by a factor 2 or 1/2, for all four profile shapes. The central profile shape is shown with thick lines, while the other profile shapes are shown with dotted lines, and we shade between the shapes. On the right, the variations of μ_B, μ_S, and ν_S (solid lines, yellow shading) and ν_B (dotted lines, green shading) are shown which contribute to the resummation uncertainty. Combinations of variations of these scales make up the set of variations that we perform to asses the uncertainties in our prediction.

the factorization,

$$\frac{\mu_S}{\mu_B} \sim \frac{\mu_S}{\nu_S} \sim 1\,, \qquad \frac{\nu_B}{\nu_S} \sim \frac{m_H}{p_T^{\rm cut}}\,. \tag{67}$$

All of these scalings are respected by the central profiles. We then constrain the variations about the central profiles to not violate any of these scaling relations by more than a factor of 2 (as would happen for instance by varying μ_B up and μ_S down). We make one additional constraint on the variations by considering the evolution factor U_0 in Eq. (57). The summation of rapidity logarithms contains the factor

$$\exp\left[\ln\Big(\frac{\nu_B}{\nu_S}\Big)\,\gamma_\nu^g(p_T^{\rm cut}, R, \mu_B)\right]. \tag{68}$$

This is a unique combination as it features a large logarithm of ν_B/ν_S multiplying a rapidity anomalous dimension that depends on μ_B. A simultaneous variation of μ_B down with either ν_S down or ν_B up gives sensitivity to small scales $\alpha_s(\mu_B)$, and the effect is effectively doubled by the $\ln(\nu_S/\nu_B)$ variation, leading us to eliminate these two combinations from the set of scale variations we consider.

With these restrictions, there are 35 remaining (of an original possible 80) profile scale variations of μ_B, μ_S, ν_B, and ν_S away from their central profile which probe the resummation uncertainty. We note that without separately varying μ_B and μ_S, and without explicit variations of the ν_B and ν_S scales there would be only a single up/down variation and a significant reduction in the resummation uncertainty. Exploring a much larger space for the scale variations is crucial to reliably estimate the uncertainty from the summation of logarithms. Note that at small R the large $\ln R^2$ effects appear through the rapidity RGE, so it is important to vary the rapidity scales to probe the effect of these terms on the $p_T^{\rm cut}$ resummation. For the final resummation uncertainty we use

$$\Delta_{\rm resum}(p_T^{\rm cut}) = \max_{v_i \in V_{\rm resum}} \left|\sigma_0^{v_i}(p_T^{\rm cut}) - \sigma_0^{\rm central}(p_T^{\rm cut})\right|, \tag{69}$$

where $V_{\rm resum}$ is the above set of 35 resummation scale variations. This uncertainty determines the covariance matrix $C_{\rm resum}$, and together with C_μ gives the full covariance matrix.

D. Uncertainties from Clustering Effects

The purpose of the profile scale variations is to estimate the effect of uncalculated higher-order terms in the cross section. This includes the higher-order corrections in the perturbative series of the various anomalous dimensions, which would be needed for the resummation at N^{3}LL. While this is effective for the logarithms of $p_T^{\rm cut}/m_H$, which are being resummed, the clustering effects generate an all-orders series of logarithms of $p_T^{\rm cut}/m_H$ and logarithms of R^2. In particular, as explained at the end of Sec. II B, the $\ln R^2$ terms appear as an unresummed series of large logarithms in the rapidity anomalous dimension. The effect of these terms on the resummed cross section is not necessarily well estimated from scale variation of the lowest order term alone.

The new clustering effects (those not determined from soft function exponentiation) arising at $\mathcal{O}(\alpha_s^n)$ depend on a coefficient $C_n(R)$, whose small R limit has the form in Eq. (3). The term with the most factors of $\ln R^2$ at $\mathcal{O}(\alpha_s^n)$ gives a contribution to the cross section of the

order	matching (H_{gg}, B_g, S_{gg})	nonsingular	$\gamma^g_{H,B,S}$	γ^g_ν	$\Gamma^g_{\rm cusp}$	β	PDF	$\alpha_s(m_Z)$
NLL$_{p_T}$	LO	-	1-loop	1-loop	2-loop	2-loop	LO	0.13939
NLL$'_{p_T}$+NLO	NLO	NLO	1-loop	1-loop	2-loop	2-loop	NLO	0.12018
NNLL$'_{p_T}$+NNLO	NNLO	NNLO	2-loop	2-loop	3-loop	3-loop	NNLO	0.11707

TABLE I: Perturbative ingredients entering at each order in resummed perturbation theory.

form

$$\ln\frac{\sigma_0^{(n)}(p_T^{\rm cut})}{\sigma_{\rm LO}} \supset C_{n,n-1}\left[\frac{\alpha_s(p_T^{\rm cut})C_A}{\pi}\ln R^2\right]^{n-1} \times\left[\frac{\alpha_s(p_T^{\rm cut})C_A}{\pi}\ln\frac{m_H}{p_T^{\rm cut}}\right], \qquad (70)$$

where only the lowest $\mathcal{O}(\alpha_s^2)$ clustering coefficient $C_{2,1} = -2.49$ is known [see Eq. (24)]. Note that $\ln R^2$ dependent terms with more powers of $\ln(m_H/p_T^{\rm cut})$ are determined by exponentiation through the rapidity RGE [i.e. the terms in Eq. (70) arise as higher-order corrections in $\gamma_\nu^g(R)$].

Until a calculation of any of the higher-order clustering coefficients exists, the best we can do is to estimate their effect on the cross section. To derive an uncertainty estimate from higher-order clustering effects, we use the ansatz $C_{3,2} = \pm C_{2,1}$ and add the corresponding $\mathcal{O}(\alpha_s^3)$ term to the rapidity anomalous dimension. We have chosen the above way of factoring out color factors and defining the higher-order clustering coefficients, such that $C_{2,1}$ is roughly an $\mathcal{O}(1)$ number and the higher-order corrections scale with a power of

$$\frac{\alpha_s(p_T^{\rm cut})C_A}{\pi}\ln R^2\,. \qquad (71)$$

In this way, taking $C_{3,2} = \pm C_{2,1}$ leads to a reasonable estimate of the potential size of the higher-order clustering corrections. For example, for $R = 0.4$, $p_T^{\rm cut} = 25\,\text{GeV}$, this factor is -0.25, so taking $C_{3,2} = \pm C_{2,1}$ the $\mathcal{O}(\alpha_s^3)$ clustering term would give a 25% correction to the $\mathcal{O}(\alpha_s^2)$ clustering term. This leads to a clustering uncertainty which is not negligible but fortunately does not dominate the uncertainty. Numerical results for different parameters of phenomenological interest are given in the next section.

IV. PREDICTIONS FOR THE LHC

In this section we present our predictions for the exclusive 0-jet cross section σ_0, the inclusive 1-jet cross section $\sigma_{\geq 1}$, and the exclusive 0-jet fraction ϵ_0. In analyzing our results we will consider varying: the perturbative order (NLL$_{p_T}$, NLL$'_{p_T}$+NLO, and NNLL$'_{p_T}$+NNLO), the choice of jet radius R, and the choice of $p_T^{\rm cut}$. The Higgs mass dependence may also be examined, but we will fix $m_H = 125\,\text{GeV}$. The order of the hard, beam, and soft functions, nonsingular corrections, and anomalous dimensions entering at each order in the resummed cross section are given in Table I. We use the MSTW 2008 PDFs [58] with their $\alpha_s(m_Z)$ at the relevant order as shown in Table I.[4]

We start with a summary of our main results. In Table II we give our predictions for each of $\sigma_{\geq 0}$, σ_0, $\sigma_{\geq 1}$, and ϵ_0 using $p_T^{\rm cut} \in \{25, 30\}\,\text{GeV}$ and $R \in \{0.4, 0.5, 0.7\}$. The uncertainties are determined by the covariance matrix in Eq. (51). The basic parameters in the matrix are the resummation uncertainty $\Delta_{\rm resum}$ and the fixed-order uncertainties $\Delta_{\rm tot}$, $\Delta_{\mu 0}$, and $\Delta_{\mu\geq 1} = \Delta_{\rm tot} - \Delta_{\mu 0}$. The values of these uncertainties for two examples are

	$p_T^{\rm cut} = 25\,\text{GeV}$, $R = 0.4$	$p_T^{\rm cut} = 30\,\text{GeV}$, $R = 0.5$
$\Delta_{\rm tot}$:	1.49 pb	1.49 pb
$\Delta_{\rm resum}$:	0.86 pb	0.52 pb
$\Delta_{\mu 0}$:	0.87 pb	0.70 pb
$\Delta_{\mu\geq 1}$:	0.62 pb	0.79 pb

(72)

which can be compared to total uncertainties quoted in Table II. In Eq. (72) the reduction in resummation uncertainties at larger R and $p_T^{\rm cut}$ is to be expected, and is mainly driven by the increase in $p_T^{\rm cut}$. This is also the main reason for the reduced uncertainties with increasing $p_T^{\rm cut}$ in σ_0 and ϵ_0 at NNLL$'_{p_T}$+NNLO, seen in Table II.

We will discuss additional aspects of Table II and associated figures for σ_0, $\sigma_{\geq 1}$, and ϵ_0 in the following subsections. In Figs. 7, 8, and 9 we will show predictions at different orders and compare our most accurate prediction to the NNLO result. In Eq. (73) we will estimate the uncertainty from higher-order clustering terms. Then in Fig. 10 we will plot various correlation coefficients as a function of $p_T^{\rm cut}$, and in Table III give correlation coefficients for two different values of R. In App. A, we will discuss in more detail the impact of the π^2 summation on our analysis.

[4] At NLL, the α_s running order required by the LO PDFs and the resummation differ. In this case, we use the pragmatic solution of including the required 2-loop beta function coefficients in the RGE evolution kernels, but use the 1-loop running required by the LO PDFs to obtain the numerical value of α_s at a given scale. This mismatch does not happen at the higher orders.

	$\sigma_{\geq 0}$ [pb]	$\sigma_0(p_T^{\rm cut})$ [pb]	$\sigma_{\geq 1}(p_T^{\rm cut})$ [pb]	$\epsilon_0(p_T^{\rm cut})$
NLL$'_{p_T}$+NLO				
$p_T^{\rm cut} = 25\,{\rm GeV}$	20.46 ± 3.37 (16.5%)	11.19 ± 1.98 (17.7%)	9.27 ± 2.76 (29.7%)	0.547 ± 0.086 (15.8%)
$p_T^{\rm cut} = 30\,{\rm GeV}$	20.46 ± 3.37 (16.5%)	12.70 ± 2.07 (16.3%)	7.76 ± 2.67 (34.5%)	0.621 ± 0.090 (14.5%)
NNLL$'_{p_T}$+NNLO ($R = 0.4$)				
$p_T^{\rm cut} = 25\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	12.67 ± 1.22 (9.6%)	9.01 ± 1.06 (11.8%)	0.584 ± 0.040 (6.8%)
$p_T^{\rm cut} = 30\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	14.09 ± 0.96 (6.8%)	7.60 ± 0.93 (12.3%)	0.650 ± 0.028 (4.4%)
NNLL$'_{p_T}$+NNLO ($R = 0.5$)				
$p_T^{\rm cut} = 25\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	12.40 ± 1.12 (9.0%)	9.28 ± 1.03 (11.1%)	0.572 ± 0.036 (6.2%)
$p_T^{\rm cut} = 30\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	13.85 ± 0.87 (6.3%)	7.83 ± 0.94 (12.0%)	0.639 ± 0.026 (4.1%)
NNLL$'_{p_T}$+NNLO ($R = 0.7$)				
$p_T^{\rm cut} = 25\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	11.97 ± 1.05 (8.8%)	9.71 ± 0.97 (10.0%)	0.552 ± 0.032 (5.7%)
$p_T^{\rm cut} = 30\,{\rm GeV}$	21.68 ± 1.49 (6.9%)	13.48 ± 0.83 (6.2%)	8.20 ± 0.92 (11.2%)	0.622 ± 0.024 (3.8%)

TABLE II: Predictions for various cross sections with complex scale setting $\mu_H = -i\mu_{\rm FO}$ and $\mu_{\rm FO} = m_H$ as the central scale choice, and with the total combined perturbative uncertainties. For convenience we also show the equivalent percent uncertainty in brackets after each result.

A. The 0-Jet Cross Section

The fundamental quantity measured by experiments that needs to be calculated theoretically is $\sigma_0(p_T^{\rm cut}, R)$, the fiducial cross section in the 0-jet bin. For this reason the predictions discussed here for the 0-jet cross section at NNLL$'_{p_T}$+NNLO are our main results. The purpose of the resummation is to improve the precision and accuracy of the fixed-order cross section when $p_T^{\rm cut} \ll m_H$, so it is natural to compare the resummed result to the NNLO cross section. For NNLO we use the central scale $\mu_{\rm FO} = m_H$ throughout. In addition, to verify the validity of our uncertainty analysis it is important to study the convergence of the resummation by studying different orders in the resummed perturbation theory. We make these comparisons in Fig. 7 using $R = 0.4$. From the top left panel one sees that there is indeed a substantial reduction of uncertainties when increasing the accuracy of the resummation/matching, with higher orders falling inside the uncertainty bands of the lower order results, as desired. From the top right panel one sees that the NNLL$'_{p_T}$+NNLO prediction has noticeably smaller uncertainties than the NNLO prediction. This is expected for smaller $p_T^{\rm cut}$, but even remains true for larger $p_T^{\rm cut}$ due to the π^2 summation that is present in the NNLL$'_{p_T}$+NNLO result, but not in the NNLO result. (The corresponding comparisons for $R = 0.5$ are quite similar, yielding the same conclusions.)

The bottom left panel shows percent uncertainties for the two highest order resummation results, and also breaks them down into the contributions from the resummation uncertainty $\Delta_{\rm resum}$ and the total uncertainty from combining yield and resummation uncertainties in quadrature. For large $p_T^{\rm cut}$ the yield uncertainties dominate at both NLL$'_{p_T}$+NLO and NNLL$'_{p_T}$+NNLO, since the resummation is not important in this region. For both of these orders the resummation uncertainty starts to have a relevant impact for $p_T^{\rm cut} \lesssim 40\,{\rm GeV}$.

In the bottom right panel of Fig. 7 we show the percent uncertainties relative to the central curve for the NNLL$'_{p_T}$+NNLO and NNLO cross sections. In this figure the size of the improvement is clear. For instance, for $R = 0.4$ and $p_T^{\rm cut} = 25\,{\rm GeV}$ the uncertainty decreases from 20% at NNLO to 9.6% at NNLL$'_{p_T}$+NNLO. Similar improvements by roughly a factor of 2 are observed for $p_T^{\rm cut} = 30\,{\rm GeV}$ and for $R = 0.5$. Jet binning is a key aspect of the experimental $H \to WW$ and $H \to \tau\tau$ analyses, which will therefore directly benefit from this substantial improvement in the theoretical uncertainties.

The clustering effects provide an additional uncertainty. Using the procedure discussed in Sec. III D, the relative uncertainty from clustering, $\Delta_0^{\rm clus}(p_T^{\rm cut})/\sigma_0(p_T^{\rm cut})$, is

$$\begin{array}{lcc} (\Delta_0^{\rm clus}/\sigma_0)(p_T^{\rm cut}) & p_T^{\rm cut} = 25\,{\rm GeV} & p_T^{\rm cut} = 30\,{\rm GeV} \\ R = 0.4: & 3.6\% & 2.9\% \\ R = 0.5: & 2.1\% & 1.7\% \\ R = 0.7: & 0.5\% & 0.4\% \end{array} \qquad (73)$$

Since our method of estimating these uncertainties is likely to be improved in the future by calculations or a better understanding of clustering effects, we have not included them in the plots or in our numbers in Table II. These clustering uncertainties are small compared to the perturbative uncertainties discussed above and shown in Table II, but are nonnegligible, so we will quote them as an additional uncertainty on each 0-jet cross section. One should interpret these with care since they come from a rough estimate of the higher-order clustering coefficient which could easily be twice as large or one-half as large. As representative final results we quote the following val-

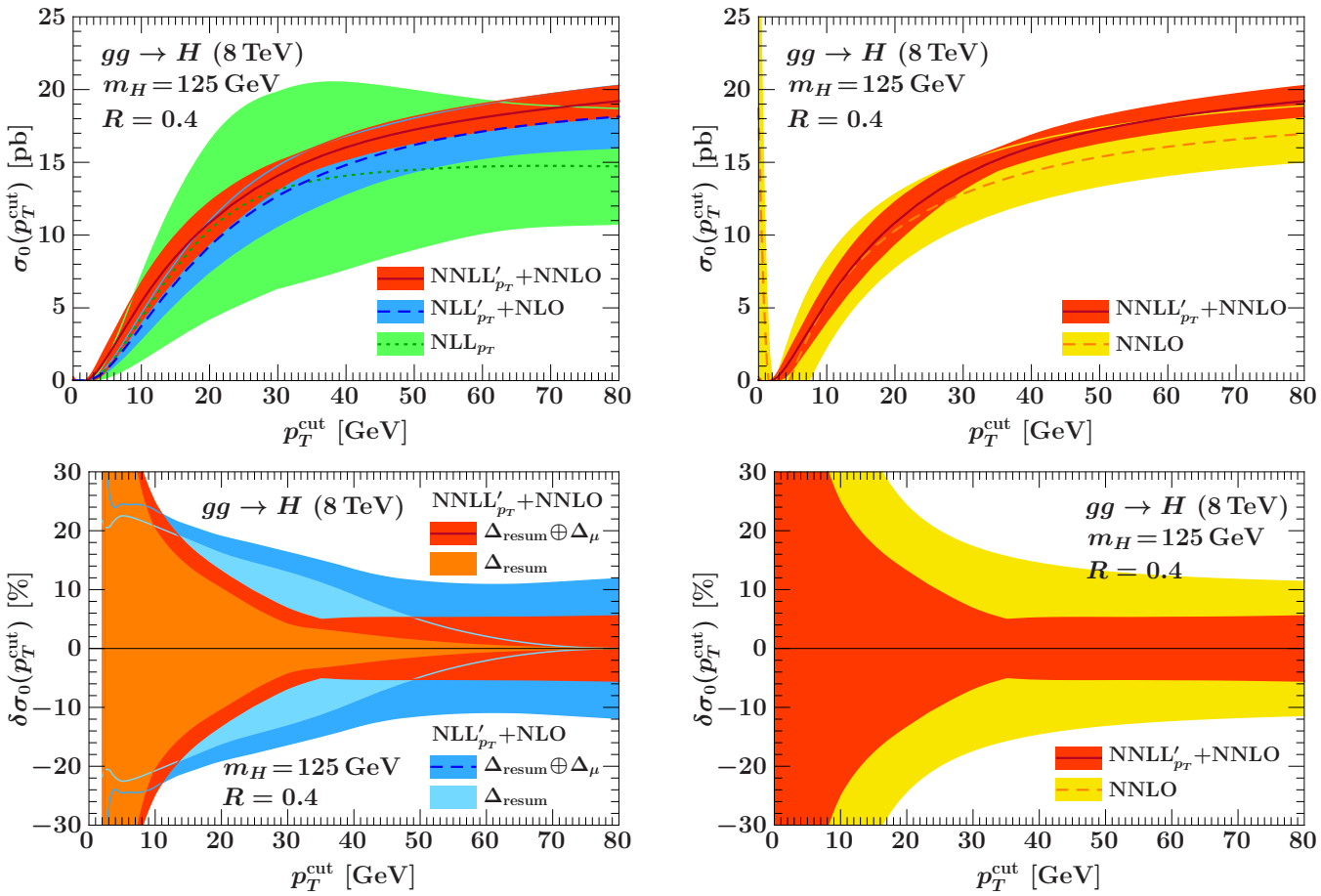

FIG. 7: The 0-jet cross section for $R = 0.4$ and $m_H = 125\,\text{GeV}$. On the left we show the NLL$_{p_T}$, NLL$'_{p_T}$+NLO, and NNLL$'_{p_T}$+NNLO predictions. A good convergence and reduction of uncertainties at successively higher orders is observed. On the right we compare our best prediction at NNLL$'_{p_T}$+NNLO to the fixed NNLO prediction. The lower plots show the relative uncertainty in percent for each prediction. On the lower left the lighter inside bands show the contribution from $\Delta_{\rm resum}$ only, while the darker outer bands show the total uncertainty from adding $\Delta_{\rm resum}$ and Δ_μ in quadrature.

ues for $\sigma_0(p_T^{\rm cut}, R)$ with both theoretical uncertainties:

$$\sigma_0(25\,\text{GeV}, 0.4) = 12.67 \pm 1.22_{\rm pert} \pm 0.46_{\rm clust}\,\text{pb}\,,$$
$$\sigma_0(30\,\text{GeV}, 0.5) = 13.85 \pm 0.87_{\rm pert} \pm 0.24_{\rm clust}\,\text{pb}\,. \quad (74)$$

It is interesting to compare our results and uncertainties for σ_0 to the NNLL+NNLO results presented earlier in Ref. [9]. Our results build on their results in a few ways. In particular, our RG approach includes π^2 resummation, our results are quoted as NNLL$'$ because they go beyond NNLL by including the complete NNLO singular terms in the fixed-order matching (which are the correct boundary conditions for the N^{3}LL resummation), and finally we use a factorization based approach to uncertainties, which also makes predictions for the correlations between the different jet bins.

Comparing σ_0 at $p_T^{\rm cut} = 25\,\text{GeV}$ and $R = 0.4$ our central values agree with those in Ref. [9], and are well within each other's uncertainties. Our perturbative uncertainty of 9.6% is a bit smaller than the 13.3% uncertainty for σ_0 of Ref. [9] which seems reasonable given the above mentioned additions. One important ingredient in this comparison is the inclusion of the π^2 resummation which improves the convergence of our results and decreases our uncertainty. On the other hand, in Ref. [9] the central scale is chosen to be $\mu_{\rm FO} = m_H/2$ which also works in the same direction, decreasing the uncertainty relative to the choice $\mu_{\rm FO} = m_H$. For the total cross section Ref. [9] has a 7.4% uncertainty, whereas we have 6.9% uncertainty using $\mu_{\rm FO} = m_H$ and including π^2 resummation (see Table II). From Table IV in appendix App. A we see that our perturbative uncertainty for $\sigma_0(25\,\text{GeV}, 0.4)$ would increase to 12.8% if the π^2 resummation were turned off (while still taking the central $\mu_{\rm FO} = m_H$), and that at this level the uncertainty would become comparable to that of Ref. [9]. For $p_T^{\rm cut} = 30\,\text{GeV}$ and $R = 0.5$ our central values remain perfectly compatible with Ref. [9], and the uncertainties follow a pattern similar to the case above.

B. The Inclusive 1-Jet Cross Section

The inclusive 1-jet cross section contains the same jet-veto logarithms as the exclusive 0-jet cross section,

$$\sigma_{\geq 1}(p_T^{\rm cut}) = \sigma_{\geq 0} - \sigma_0(p_T^{\rm cut})\,. \tag{75}$$

Here, $p_T^{\rm cut}$ in $\sigma_{\geq 1}(p_T^{\rm cut})$ is now the lower limit on the p_T of the leading jet in this inclusive cross section. Since our resummation framework consistently includes both $\sigma_0(p_T^{\rm cut})$ and $\sigma_{\geq 0}$, we can determine a resummed prediction for $\sigma_{\geq 1}(p_T^{\rm cut})$ from their difference. A nontrivial ingredient in this prediction is determining its perturbative uncertainty via the theory covariance matrix determined in Sec. III.

In Fig. 8, we show the convergence of the resummed and matched predictions at different orders, as well as the comparison to the fixed-order cross section. The total cross section used to obtain $\sigma_{\geq 1}(p_T^{\rm cut})$ is evaluated with an accuracy equal to the fixed-order matching results contained in $\sigma_0(p_T^{\rm cut})$. This is required to enforce $\sigma_{\geq 1}(p_T^{\rm cut} \to \infty) \to 0$. For this reason in the left panel of Fig. 8 the NLL_{p_T} distribution (whose matching does not even include the full tree-level matrix element for the $H+1$ jet rate) is lower than the higher-order distributions. The NNLL'_{p_T}+NNLO distribution is well contained within the NLL'_{p_T}+NLO uncertainty band, with the expected improvement in accuracy. Note that when including the resummation, $\sigma_{\geq 1}(p_T^{\rm cut})$ approaches the total cross section as $p_T^{\rm cut} \to 0$, whereas it would diverge at fixed order.

In the right panel of Fig. 8 we compare the fixed-order result for $\sigma_{\geq 1}(p_T^{\rm cut})$ with the result obtained from Eq. (75) using our NNLL'_{p_T}+NNLO 0-jet distribution. (We label the NNLL'_{p_T}+NNLO prediction as such to be consistent with our predictions for other observables, although in terms of the fixed-order contributions it is not beyond the NLO result for $H+\geq 1$ jet denoted as NLO_1 in the figure.) The resummed prediction for $\sigma_{\geq 1}(p_T^{\rm cut})$ is larger than the NLO_1 result due to the summation of π^2 terms in $\sigma_{\geq 0}$ and $\sigma_0(p_T^{\rm cut})$ in Eq. (75). Without this π^2 summation, the resummed $\sigma_{\geq 1}(p_T^{\rm cut})$ would give a slightly lower rate than at fixed order. For $R = 0.4$ and $p_T^{\rm cut} = 25\,\mathrm{GeV}$ the fixed-order uncertainty is 20%. It is reduced to 11.8% at NNLL'_{p_T}+NNLO (see Table II). This reduction is similar to what was observed for σ_0, as is the mild dependence on R. On the other hand, increasing $p_T^{\rm cut}$ to 30 GeV does not really change the relative uncertainty for $\sigma_{\geq 1}$, unlike for σ_0. Note the importance of the theory correlations here, since we can see from Eq. (72) that the yield uncertainty $\Delta_{\mu \geq 1}$ alone behaves in the opposite fashion.

Our resummed results for the inclusive 1-jet cross section $\sigma_{\geq 1}(p_T^{\rm cut}, R)$ provide improved predictions compared to the accuracy of its NLO result, but should be used together with the appropriate theory uncertainty correlations determined here. As representative final results for $\sigma_{\geq 1}(p_T^{\rm cut}, R)$, where we also include the uncertainty from clustering estimated as in Eq. (74), we quote

$$\begin{aligned}\sigma_{\geq 1}(25\,\mathrm{GeV}, 0.4) &= 9.01 \pm 1.06_{\rm pert} \pm 0.46_{\rm clust}\,\mathrm{pb}\,,\\ \sigma_{\geq 1}(30\,\mathrm{GeV}, 0.5) &= 7.83 \pm 0.94_{\rm pert} \pm 0.24_{\rm clust}\,\mathrm{pb}\,.\end{aligned} \tag{76}$$

Note that the clustering uncertainties have a larger relative size here (5.1% and 3.0%) since $\sigma_{\geq 1}$ is numerically smaller than σ_0.

Recently, the $gg \to Hg$ contribution to the $H+\geq 1$-jet cross section has been calculated at NNLO [59]. This calculation includes all $\mathcal{O}(\alpha_s^3)$ corrections which include logarithms of $p_T^{\rm cut}/m_H$, π^2 terms, and nonsingular contributions. Our resummed calculation captures all of the logarithms of $p_T^{\rm cut}/m_H$ except for the single logarithms (which would require $\mathrm{N^3LL}$ resummation) as well as the π^2 terms at $\mathcal{O}(\alpha_s^3)$, but does not include any nonsingular contributions. In contrast, the fixed-order calculation does not include the resummation of the $p_T^{\rm cut}$ logarithms or π^2 terms beyond $\mathcal{O}(\alpha_s^3)$. The different theoretical ingredients in these two calculations makes a comparison between them interesting. In fact, for phenomenologically relevant parameters the $gg \to Hg$ NNLO calculation finds a K-factor relative to NLO that is quantitatively similar to the increase over the NLO cross section that we observe between the two central curves in the right panel of Fig. 8. As mentioned above, in our case the resummation of the $p_T^{\rm cut}$ logarithms lowers the 1-jet inclusive cross section relative to fixed NLO, but including also the π^2 summation raises it above. Although the purely virtual π^2 terms from the hard function cancel out in Eq. (75), there are real-virtual cross terms involving π^2 factors in $\sigma_{\geq 0}$ that are not canceled. This suggests that these π^2 terms may play an important role in determining the magnitude of the NNLO K-factor. (In contrast, the π^2 terms that can be determined from imaginary scale setting in the exclusive H+1-jet cross section are known to not play a dominant role at NLO [50].)

C. The 0-Jet Efficiency

Another observable that can be predicted using our results is the 0-jet efficiency,

$$\epsilon_0(p_T^{\rm cut}) = \frac{\sigma_0(p_T^{\rm cut})}{\sigma_{\geq 0}}\,. \tag{77}$$

Once again it is important to account for the correlations in theoretical uncertainties when computing the uncertainty in this observable according to Eq. (53). In Fig. 9, we plot $\epsilon_0(p_T^{\rm cut})$ and its uncertainty as a function of $p_T^{\rm cut}$ for $R = 0.4$ and we give explicit numbers in Table II. At NLL'_{p_T}+NLO the relative uncertainties for σ_0 and ϵ_0 are similar, but this is no longer the case at NNLL'_{p_T}+NNLO. With the decreased uncertainties that occur at this order, a more significant amount of the uncertainties in the numerator and denominator of Eq. (77) become positively correlated and cancel. As a result, our 0-jet efficiency at NNLL'_{p_T}+NNLO has smaller relative uncertainties than

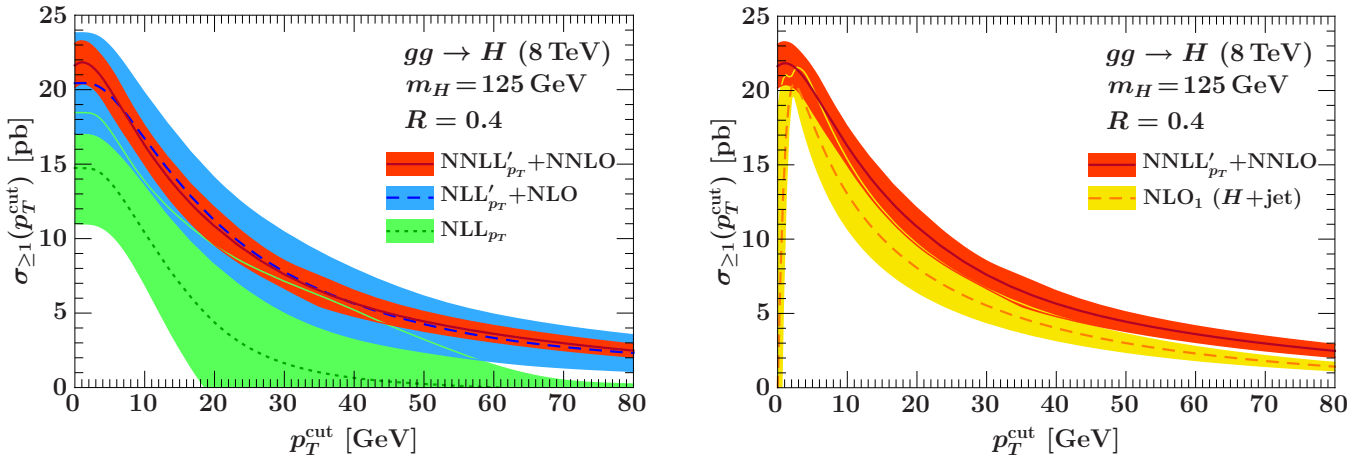

FIG. 8: The inclusive 1-jet cross section for $R = 0.4$ and $m_H = 125\,\text{GeV}$. On the left we show the different orders of our resummed predictions, and on the right we compare our best prediction to that derived from the fixed NNLO cross section. As in the 0-jet cross section, we observe a good convergence and reduction in uncertainties at successively higher orders of accuracy.

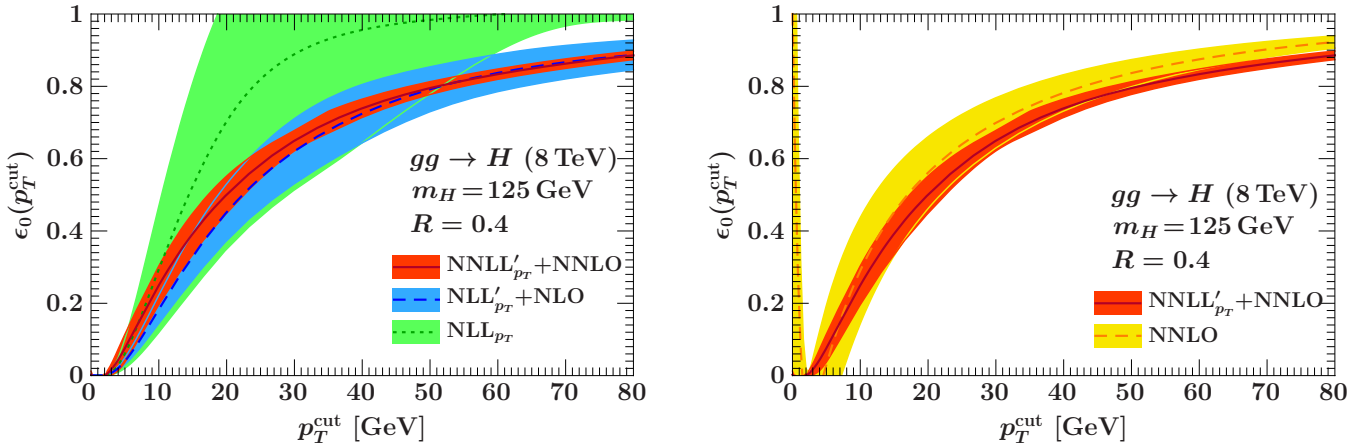

FIG. 9: The 0-jet efficiency for $R = 0.4$ and $m_H = 125\,\text{GeV}$. On the left we show the different orders in our resummed predictions, and on the right we compare our best prediction to that derived from the fixed NNLO cross section. Because the efficiency is the ratio of the 0-jet and total cross sections, the correlated fixed-order scale uncertainty in each quantity reduces the uncertainty in the 0-jet efficiency, making it relatively more accurate than the cross section itself.

our 0-jet cross section. This is reflected in both the numbers in Table II and in the results shown in Fig. 9.

In the left panel of Fig. 9 we show results for the efficiency at different orders. The results at NNLL′$_{p_T}$+NNLO are within the uncertainty band of the lower order NLL′$_{p_T}$+NLO results, and again display an improved level of precision. In the right panel of Fig. 9 we see that the comparison of $\epsilon_0(p_T^{\rm cut})$ between NNLL′$_{p_T}$+NNLO and pure NNLO follows a similar pattern of improvement to what we have already observed for the 0-jet and inclusive 1-jet cross sections.

Since the 0-jet efficiency is the more fundamental quantity in the framework of Ref. [9], it makes sense to extend the comparison made in Sec. IV A to this observable, again taking $R = 0.4$ and $p_T^{\rm cut} = 25\,\text{GeV}$. At NNLL+NNLO Ref. [9] has a 11.5% perturbative uncertainty for ϵ_0, which in their framework is assumed to be independent from the uncertainty in the total cross section. Thus, their uncertainty for σ_0 is always larger than that for ϵ_0. This 11.5% uncertainty for their ϵ_0 is close to the 9.6% uncertainty for our σ_0, but larger than the 6.8% uncertainty for our ϵ_0. For the analysis of Ref. [9] there is no corresponding cancellation of uncertainties between the numerator and denominator of Eq. (77), and hence the same cancellation that we observe does not occur.

D. Correlations

When evaluating the perturbative uncertainties via the profile scale variations as discussed in Sec. III C, the correlations in the total perturbative uncertainties between

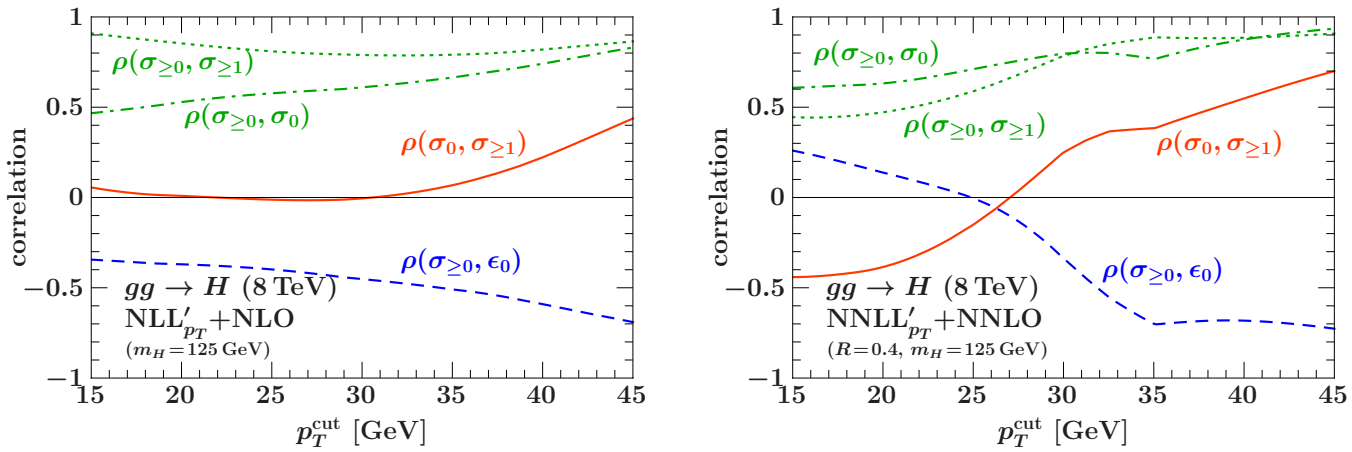

FIG. 10: Predicted correlation coefficients for the total perturbative uncertainties within the resummed predictions between different observables at NLL′$_{p_T}$+NLO (left) and NNLL′$_{p_T}$+NNLO (right, with $R = 0.4$). Since the correlations result from the interplay between the relative sizes of the C_μ and $C_{\rm resum}$ components, the changes between orders is not unexpected.

$p_T^{\rm cut} = 30$ GeV	$\sigma_{\geq 0}$	$\sigma_0(p_T^{\rm cut})$	$\sigma_{\geq 1}(p_T^{\rm cut})$	$\epsilon_0(p_T^{\rm cut})$
$R = 0.4$				
$\sigma_{\geq 0}$	1	0.80	0.78	−0.34
$\sigma_0(p_T^{\rm cut})$		1	0.25	0.30
$\sigma_{\geq 1}(p_T^{\rm cut})$			1	−0.85
$\epsilon_0(p_T^{\rm cut})$				1
$R = 0.5$				
$\sigma_{\geq 0}$	1	0.81	0.84	−0.44
$\sigma_0(p_T^{\rm cut})$		1	0.35	0.18
$\sigma_{\geq 1}(p_T^{\rm cut})$			1	−0.86
$\epsilon_0(p_T^{\rm cut})$				1

TABLE III: Correlations in the perturbative uncertainties between different observables at $p_T^{\rm cut} = 30$ GeV for $R = 0.4$ and $R = 0.5$.

the different observables are automatically predicted by the resulting total covariance matrix $C_\mu + C_{\rm resum}$. In previous subsections we have highlighted a few cases where these correlations are important for determining uncertainties, and in this section we discuss them in more detail.

As an example, in Table III we give the correlation coefficients obtained at $p_T^{\rm cut} = 30$ GeV for both $R = 0.4$ and $R = 0.5$. One observes that they have a fairly mild dependence on R. On the other hand, since the correlations arise from the interplay between the relative size of the anticorrelated component C_μ and correlated component $C_{\rm resum}$, they can have a much stronger dependence on $p_T^{\rm cut}$. Similarly, the correlation matrix can also change by a large amount between perturbative orders because uncertainties are decreased by going to higher order, and therefore the relative importance of C_μ and $C_{\rm resum}$ can change. These two features are illustrated in Fig. 10. The $p_T^{\rm cut}$ dependence is strongest in the correlations between the inclusive cross section, $\sigma_{\geq 0}$, and the exclusive 0-jet observables σ_0 (solid orange lines) and ϵ_0 (blue dashed lines). The reason for this is that the 0-jet observables receive contributions from $C_{\rm resum}$, whose importance relative to C_μ depends on $p_T^{\rm cut}$, while $\sigma_{\geq 0}$ has no contribution from $C_{\rm resum}$. We also see that at NNLL′$_{p_T}$+NNLO the correlation between σ_0 and $\sigma_{\geq 1}$ decreases toward smaller $p_T^{\rm cut}$ and turns negative below $\lesssim 30$ GeV, because the anticorrelated migration uncertainties from $C_{\rm resum}$ start dominating over their common correlated yield uncertainty in C_μ. This anti-correlation is not so evident in the resummed result at NLL′$_{p_T}$+NLO since C_μ plays a bigger role at this order. Finally, we observe that in the large $p_T^{\rm cut}$ regime, where the resummation turns off and the C_μ contributions dominate, the correlations between the 0-jet efficiency and the total cross section in our formalism approaches -1, as it must. For large $p_T^{\rm cut}$ the correlations between any two cross sections tends to 1, also as they must.

From this discussion it should also be apparent that we do not expect the correlations obtained after resummation to be the same as in the pure fixed-order calculation. Indeed, including the resummation the perturbative uncertainties in the logarithmic series induced by the jet binning are significantly reduced compared to in the fixed-order case. This means the correlation between the uncertainties in σ_0 and $\sigma_{\geq 1}$ should be more negative at fixed order. This is indeed what happens when using the method of Ref. [12], for which at pure NNLO we find $\rho(\sigma_0, \sigma_{\geq 1})$ rises from -0.7 to -0.2 over the $p_T^{\rm cut}$ range shown in Fig. 10. The added advantage of the resummation framework used here is that it automatically provides theory based handles to estimate both the correlated contributions C_μ and anticorrelated contributions in $C_{\rm resum}$ without having to make an assumption about the correlation between any two quantities. As a final cautionary note, we remark that one should recall that the magnitude of the correlation coefficients does

not indicate the relative importance of their entries in determining the final uncertainties since the size of the corresponding diagonal uncertainties is also required.

V. CONCLUSIONS

In this paper we have presented results for Higgs production via gluon fusion with a jet veto. Jets are identified with a k_T-type clustering algorithm (which includes the experimentally used anti-k_T algorithm) with jet radius R, and are vetoed via the requirement $p_T^{\rm jet} < p_T^{\rm cut}$. The logarithms of $p_T^{\rm cut}/m_H$ are resummed to NNLL$'$ and the resummation is matched to the full fixed NNLO cross section. Our analysis is based around the small R limit, where the cross section can be factorized into hard, beam, and soft functions. To achieve NNLL$'$ order we computed the relevant soft function to $\mathcal{O}(\alpha_s^2)$ and computed the full $\alpha_s^2 \ln R^2$ term for the beam function, determining the remaining $p_T^{\rm cut}$ independent $\mathcal{O}(\alpha_s^2)$ terms in the beam function numerically. Our resummation results also include π^2 summation in the hard corrections through imaginary scale setting. To consistently incorporate the full NNLO result we made use of profile functions that properly handle both the small and large $p_T^{\rm cut}$ regions, and in particular the experimentally relevant transition region in between. We also included a precise numerical determination of the $\mathcal{O}(\alpha_s^2)$ nonsingular terms. Our results include predictions for the exclusive 0-jet cross section, the 0-jet efficiency, and the inclusive 1-jet cross section.

A key aspect of our numerical analysis is the robust estimation of perturbative uncertainties. The uncertainty comes from two independent components: overall yield uncertainties (which are correlated between jet bins) and resummation uncertainties (related to predicting the migration between jet bins as we vary $p_T^{\rm cut}$). Each of these can be estimated through the variation of various invariant mass and rapidity scales in the factorization theorem. The uncertainty framework discussed in Sec. III allows us to construct the complete covariance matrix for the total, exclusive 0-jet, and inclusive 1-jet cross sections.

In Sec. IV, we presented results for the 0-jet cross section, the inclusive 1-jet cross section, and the 0-jet efficiency. Our numerical results for several phenomenological points of interest ($p_T^{\rm cut} = 25, 30\,\text{GeV}$ and $R = 0.4, 0.5, 0.7$) are given in Table II. The precision of the predictions increases significantly as the resummation and matching is improved, from NLL_{p_T} to NLL'_{p_T}+NLO to NNLL'_{p_T}+NNLO. For the most precise predictions, the uncertainties are significantly smaller than the fixed-order NNLO uncertainties, which are currently the nominal benchmark uncertainties for the experimental $H \to WW$ and $H \to \tau\tau$ analyses. Our results add a few additional ingredients on top of the NNLL results in Ref. [9], in particular: by including π^2 summation [35], by including the complete NNLO singular terms in the fixed-order matching for soft and beam functions at $\mathcal{O}(\alpha_s^2)$, and because our factorization based framework also makes predictions for both correlated and anticorrelated contributions to the theory uncertainty correlation matrix between different jet bins. We observe a corresponding modest improvement in the size of the uncertainties, where details can be found in Sec. IV A and Sec. IV C.

Our results are part of an ongoing effort to more completely understand jet vetoes for Higgs production and their associated uncertainties. The H + 0-jet cross section is an excellent testing ground for the new methods being developed to improve the theoretical predictions. Currently, the fixed-order perturbative uncertainties due to the jet binning in the $H \to WW$ analysis are the dominant systematic uncertainties. Our results can be directly applied to provide improved theory predictions with substantially reduced perturbative uncertainties.

Acknowledgments

We thank Robert Schabinger for assistance in converting our numerical values of the 2-loop non-cusp anomalous dimensions into analytic expressions. The authors thank each other's institutions and the Erwin Schrödinger Institute program "Jets and Quantum Fields for LHC and Future Colliders" for hospitality while portions of this work were completed.

This work was supported in part by the Director, Office of Science, Offices of Nuclear Physics and High Energy Physics of the U.S. Department of Energy under the Grant No. DE-FG02-94ER40818 and the Contract No. DE-AC02-05CH11231, the DFG Emmy-Noether grant TA 867/1-1, and the US National Science Foundation, grant NSF-PHY-0705682, the LHC Theory Initiative.

This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Note Added:

While finalizing this paper Ref. [60] appeared, which also makes predictions for the H+0-jet cross section including contributions beyond the NNLL results of Ref. [9]. In this note added we compare their theoretical ingredients with ours.

Regarding the derivation of factorization for terms of $\mathcal{O}(R^2)$ we believe the discussion of rapidity scaling in our footnote 2 still applies to Ref. [60].

A common goal of both our work and Ref. [60] is the inclusion of fixed-order corrections from the low-energy matrix elements (corresponding to beam and/or soft functions) that are needed as ingredients in a calculation at N^{3}LL order. In our analysis we have fully calculated the $\mathcal{O}(\alpha_s^2)$ soft function, including the R-dependent anomalous dimension and finite corrections that depend on $\ln R^2$. In addition, we have calculated the finite corrections in the beam function that depend on $\ln R^2$. Thus, the dominant R dependence has been fully determined analytically, and the only numerical ingredient is the re-

maining contribution in the beam function. In contrast, in Ref. [60] an analytic calculation is done for the anomalous dimension terms, but a numerical extraction is done for the combined finite soft + beam contributions including their R dependence. We make use of the rapidity renormalization group in our analysis, including rapidity scale variations in our uncertainties to estimate the size of higher-order rapidity logarithms, while Ref. [60] accounts for these contributions using the "collinear anomaly" formalism without variations of the rapidity scales. A resummation of π^2 contributions through imaginary scale setting is used in both our work and their work.

Ref. [60] refers to the accuracy of their resummation as "$\mathrm{N^3LL_p}$", where "p" stands for partial, which can be contrasted with our NNLL′. As far as perturbative ingredients that have been either calculated analytically or extracted numerically, both our results include the same theoretical ingredients. Ref. [60] makes an additional ansatz about the anomalous dimensions required for $\mathrm{N^3LL}$ resummation, since none of the required coefficients are currently known. Their method of estimating and varying the size of these coefficients in some range is another method for estimating uncertainties from unknown higher-order perturbative corrections. It does not however improve the perturbative accuracy of the resummation beyond NNLL′ order.

In our analysis we have used profile scales to properly describe the transition between the resummation and fixed-order regimes, which ensures that we have canonical scales in the small p_T^{cut} region and also reproduce the fixed-order cross section in the large p_T^{cut} limit. In contrast, Ref. [60] limit themselves to using canonical scales, which can only be used to properly describe the cross section in the small p_T^{cut} region below the transition region. As we have seen in our analysis, for phenomenologically relevant values of p_T^{cut}, the cross section and its uncertainties are influenced by the transition region. The connection to the fixed-order cross section also provides an important constraint when predicting correlations (which are not considered in Ref. [60]). Overall, our method of calculating perturbative uncertainties by varying all scales appearing in the RGE is therefore quite different from Ref. [60]. Numerically, the resummed perturbation theory as organized in Ref. [60] show a slower convergence (as shown, e.g., in their Figs. 8 and 11) compared to our results shown in Fig. 7.

Appendix A: Results for Real μ_H

For completeness and to demonstrate the benefit of the imaginary scale setting for μ_H, in this Appendix we give predictions for the real scale setting $\mu_H = \mu_{\mathrm{FO}}$, which excludes the large π^2 terms from the resummation in the hard function.

In Fig. 11, we plot the analog of Fig. 7 for $\sigma_0(p_T^{\mathrm{cut}})$ but using real μ_H. Comparing these two figures, it is clear that including the π^2 terms in the resummation significantly improves the convergence and precision of the 0-jet predictions at small p_T^{cut}. This improvement also translates into an improved convergence and reduced uncertainties at larger values of p_T^{cut}. In Table IV, we give the analogous values without π^2 summation to those in Table II. For $p_T^{\mathrm{cut}} = 25\,\mathrm{GeV}$, $R = 0.4$ and $p_T^{\mathrm{cut}} = 30\,\mathrm{GeV}$, $R = 0.5$, the corresponding components of the uncertainty are

	$p_T^{\mathrm{cut}} = 25\,\mathrm{GeV}$ $R = 0.4$	$p_T^{\mathrm{cut}} = 30\,\mathrm{GeV}$ $R = 0.5$
Δ_{tot} :	1.91	1.91
Δ_{resum} :	1.08	0.95
$\Delta_{\mu 0}$:	1.16	1.32
$\Delta_{\mu\geq 1}$:	0.75	0.59

(A1)

Both the resummed and fixed-order uncertainties for the 0-jet cross section are larger when the π^2 terms are excluded from the resummation, indicating that these large π^2 terms have an effect on the shape as well as the normalization of the cross section. This is also reflected in Fig. 11.

[1] ATLAS Collaboration, G. Aad *et al.*, Phys. Lett. B **716**, 1 (2012), [arXiv:1207.7214].

[2] CMS Collaboration, S. Chatrchyan *et al.*, Phys. Lett. B **716**, 30 (2012), [arXiv:1207.7235].

[3] C. Anastasiou, K. Melnikov, and F. Petriello, Phys. Rev. Lett. **93**, 262002 (2004), [hep-ph/0409088].

[4] C. Anastasiou, K. Melnikov, and F. Petriello, Nucl. Phys. **B724**, 197 (2005), [hep-ph/0501130].

[5] S. Catani and M. Grazzini, Phys. Rev. Lett. **98**, 222002 (2007), [hep-ph/0703012].

[6] M. Grazzini, JHEP **02**, 043 (2008), [arXiv:0801.3232].

[7] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **04**, 092 (2011), [arXiv:1012.4480].

[8] T. Becher and M. Neubert, JHEP **1207**, 108 (2012), [arXiv:1205.3806].

[9] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, Phys. Rev. Lett. **109**, 202001 (2012), [arXiv:1206.4998].

[10] A. Banfi, G. P. Salam, and G. Zanderighi, JHEP **06**, 159 (2012), [arXiv:1203.5773].

[11] F. J. Tackmann, J. R. Walsh, and S. Zuberi, Phys. Rev. D **86**, 053011 (2012), [arXiv:1206.4312].

[12] I. W. Stewart and F. J. Tackmann, Phys. Rev. D **85**, 034011 (2012), [arXiv:1107.2117].

[13] X. Liu and F. Petriello, Phys. Rev. D **87**, 014018 (2013), [arXiv:1210.1906].

[14] X. Liu and F. Petriello, Phys. Rev. D **87**, 094027 (2013), [arXiv:1303.4405].

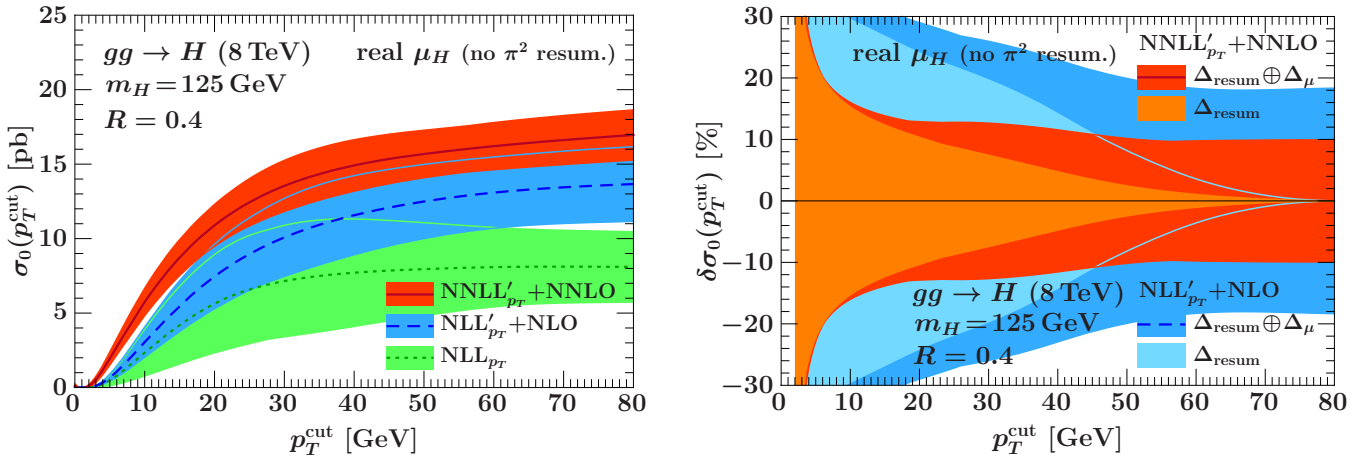

FIG. 11: The 0-jet cross section for $m_H = 125\,\text{GeV}$ and $R = 0.4$ using the real scale setting $\mu_H = \mu_{\rm FO}$, which excludes the π^2 resummation. The poor convergence of the hard function results in larger uncertainties and a poorer convergence of the cross section at all values of $p_T^{\rm cut}$.

	$\sigma_{\geq 0}$ [pb]	$\sigma_0(p_T^{\rm cut})$ [pb]	$\sigma_{\geq 1}(p_T^{\rm cut})$ [pb]	$\epsilon_0(p_T^{\rm cut})$
NLL$'_{p_T}$+NLO				
$p_T^{\rm cut} = 25\,\text{GeV}$	14.57 ± 2.91 (20.0%)	8.96 ± 2.44 (27.2%)	5.61 ± 2.44 (43.5%)	0.615 ± 0.136 (22.1%)
$p_T^{\rm cut} = 30\,\text{GeV}$	14.57 ± 2.91 (20.0%)	10.08 ± 2.62 (26.0%)	4.49 ± 2.32 (51.7%)	0.692 ± 0.138 (19.9%)
NNLL$'_{p_T}$+NNLO ($R = 0.4$)				
$p_T^{\rm cut} = 25\,\text{GeV}$	18.38 ± 1.91 (10.4%)	12.44 ± 1.59 (12.8%)	5.94 ± 1.32 (22.2%)	0.677 ± 0.059 (8.8%)
$p_T^{\rm cut} = 30\,\text{GeV}$	18.38 ± 1.91 (10.4%)	13.54 ± 1.71 (12.6%)	4.84 ± 1.13 (23.4%)	0.737 ± 0.055 (7.4%)
NNLL$'_{p_T}$+NNLO ($R = 0.5$)				
$p_T^{\rm cut} = 25\,\text{GeV}$	18.38 ± 1.91 (10.4%)	12.14 ± 1.50 (12.4%)	6.24 ± 1.29 (20.7%)	0.661 ± 0.056 (8.4%)
$p_T^{\rm cut} = 30\,\text{GeV}$	18.38 ± 1.91 (10.4%)	13.29 ± 1.63 (12.2%)	5.09 ± 1.12 (21.9%)	0.723 ± 0.052 (7.2%)
NNLL$'_{p_T}$+NNLO ($R = 0.7$)				
$p_T^{\rm cut} = 25\,\text{GeV}$	18.38 ± 1.91 (10.4%)	11.69 ± 1.41 (12.1%)	6.68 ± 1.33 (19.9%)	0.636 ± 0.055 (8.6%)
$p_T^{\rm cut} = 30\,\text{GeV}$	18.38 ± 1.91 (10.4%)	12.91 ± 1.54 (11.9%)	5.47 ± 1.18 (21.6%)	0.703 ± 0.052 (7.5%)

TABLE IV: Predictions for various cross sections with real scale setting $\mu_H = \mu_{\rm FO}$ and $\mu_{\rm FO} = m_H$ as central scale.

[15] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. D **81**, 094035 (2010), [arXiv:0910.0467].
[16] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev.Lett. **105**, 092002 (2010), [arXiv:1004.2489].
[17] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys. Rev. Lett. **106**, 032001 (2011), [arXiv:1005.4060].
[18] A. Papaefstathiou, J. M. Smillie, and B. R. Webber, JHEP **04**, 084 (2010), [arXiv:1002.4375].
[19] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), [arXiv:0802.1189].
[20] C. W. Bauer, S. Fleming, and M. E. Luke, Phys. Rev. D **63**, 014006 (2000), [hep-ph/0005275].
[21] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, Phys. Rev. D **63**, 114020 (2001), [hep-ph/0011336].
[22] C. W. Bauer and I. W. Stewart, Phys. Lett. B **516**, 134 (2001), [hep-ph/0107001].
[23] C. W. Bauer, D. Pirjol, and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002), [hep-ph/0109045].
[24] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, Phys. Rev. D **66**, 014017 (2002), [hep-ph/0202088].
[25] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, Phys. Rev. Lett. **108**, 151601 (2012), [arXiv:1104.0881].
[26] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, JHEP **05**, 084 (2012), [arXiv:1202.0814].
[27] V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, Eur. Phys. J. C **62**, 333 (2009), [arXiv:0809.4283].
[28] S. Mantry and F. Petriello, Phys. Rev. D **81**, 093007 (2010), [arXiv:0911.4135].
[29] A. Idilbi, X.-d. Ji, and F. Yuan, Phys. Lett. B **625**, 253 (2005), [hep-ph/0507196].
[30] A. Idilbi, X.-d. Ji, J.-P. Ma, and F. Yuan, Phys. Rev. D **73**, 077501 (2006), [hep-ph/0509294].
[31] A. Vogt, S. Moch, and J. Vermaseren, Nucl. Phys. B **691**, 129 (2004), [hep-ph/0404111].
[32] G. Parisi, Phys. Lett. B **90**, 295 (1980).
[33] G. F. Sterman, Nucl. Phys. **B281**, 310 (1987).
[34] L. Magnea and G. F. Sterman, Phys. Rev. D **42**, 4222 (1990).
[35] V. Ahrens, T. Becher, M. Neubert, and L. L. Yang, Phys. Rev. D **79**, 033013 (2009), [arXiv:0808.3008].
[36] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and

B. R. Webber, Phys. Lett. B **269**, 432 (1991).
[37] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
[38] S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993), [hep-ph/9305266].
[39] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997), [hep-ph/9707323].
[40] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, JHEP **09**, 005 (2010), [arXiv:1002.2213].
[41] T. Gehrmann, T. Lubbert, and L. L. Yang, Phys. Rev. Lett. **109**, 242003 (2012), [arXiv:1209.0682].
[42] A. V. Manohar and I. W. Stewart, Phys. Rev. D **76**, 074002 (2007), [hep-ph/0605001].
[43] J. M. Campbell and R. K. Ellis, Phys. Rev. D **65**, 113007 (2002), [hep-ph/0202176].
[44] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002), [hep-ph/0201206].
[45] C. Anastasiou and K. Melnikov, Nucl. Phys. B **646**, 220 (2002), [hep-ph/0207004].
[46] V. Ravindran, J. Smith, and W. L. van Neerven, Nucl. Phys. B **665**, 325 (2003), [hep-ph/0302135].
[47] S. Gangal and F. J. Tackmann, Phys. Rev. D **87**, 093008 (2013), [arXiv:1302.5437].
[48] Z. Ligeti, I. W. Stewart, and F. J. Tackmann, Phys. Rev. D **78**, 114014 (2008), [arXiv:0807.1926].
[49] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, Phys. Rev. D **83**, 074021 (2011), [arXiv:1006.3080].
[50] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, arXiv:1302.0846.
[51] C. W. Bauer, F. J. Tackmann, J. R. Walsh, and S. Zuberi, Phys. Rev. D **85**, 074006 (2012), [arXiv:1106.6047].
[52] S. Alioli *et al.*, arXiv:1211.7049.
[53] D. Kang, C. Lee, and I. W. Stewart, arXiv:1303.6952.
[54] A. Jain, M. Procura, B. Shotwell, and W. J. Waalewijn, Phys. Rev. D **87**, 074013 (2012), [arXiv:1207.4788].
[55] H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, arXiv:1306.6630.
[56] S. Gritschacher, A. H. Hoang, I. Jemos, and P. Pietrulewicz, arXiv:1302.4743.
[57] R. D. Ball, M. Bonvini, S. Forte, S. Marzani, and G. Ridolfi, arXiv:1303.3590.
[58] A. Martin, W. Stirling, R. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009), [arXiv:0901.0002].
[59] R. Boughezal, F. Caola, K. Melnikov, F. Petriello, and M. Schulze, arXiv:1302.6216.
[60] T. Becher, M. Neubert, and L. Rothen, arXiv:1307.0025.