

Non-global logarithms in filtered jet algorithms

Mathieu Rubin

*LPTHE, UPMC Univ. Paris 6, CNRS UMR 7589,
4 place Jussieu, 75252 Paris, France*

E-mail: rubin@lpthe.jussieu.fr

ABSTRACT: We analytically and numerically study the effect of perturbative gluons emission on the “Filtering analysis”, which is part of a subjet analysis procedure proposed two years ago to possibly identify a low-mass Higgs boson decaying into $b\bar{b}$ at the LHC. This leads us to examine the non-global structure of the resulting perturbative series in the leading single-log large- N_c approximation, including all-orders numerical results, simple analytical approximations to them and comments on the structure of their series expansion. We then use these results to semi-analytically optimize the parameters of the Filtering analysis so as to suppress as much as possible the effect of underlying event and pile-up on the Higgs mass peak reconstruction while keeping the major part of the perturbative radiation from the $b\bar{b}$ dipole.

KEYWORDS: Jets, Hadronic Colliders, QCD

ARXIV EPRINT: [1002.4557](https://arxiv.org/abs/1002.4557)

Contents

1	Introduction	2
2	Non-global structure: analytical insights	3
2.1	The filtered Higgs mass: a non-global observable	3
2.2	Some results for $n_{\text{filt}} = 2$	4
2.2.1	Primary emissions	5
2.2.2	Non-global contributions	7
2.3	Some results for $n_{\text{filt}} = 3$	9
3	Non-global structure: numerical results	11
3.1	Comparison with analytics	12
3.2	Comparison with fixed-order results	12
4	Choice of the filtering parameters	14
4.1	Study of the Higgs perturbative width	14
4.2	Study of the Higgs width due to underlying event and pile-up	18
4.3	Study of the Higgs width in presence of both UE/PU and perturbative radiation	20
4.4	Variations of the results with z and f	24
4.5	Hadronisation corrections	28
5	Conclusions	29
A	Analytical considerations on the non-global structure of the perturbative expansion	30
A.1	Primary coefficients	31
A.2	Non-global coefficients	33
A.2.1	Calculation of S_{tot}	34
A.2.2	Calculation of S_{int}	35
B	Analytical considerations on the dependence of the results on z and f	35
B.1	Dependence on z	35
B.2	Comments on the uncertainty due to the choice of f	37
C	Convergence of the non-global series	38
C.1	Case $n_{\text{filt}} = 2$	38
C.2	Case $n_{\text{filt}} = 3$	41
C.3	The Slice case	42

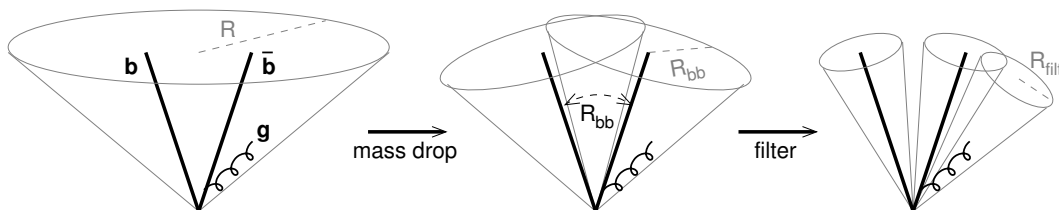


Figure 1. The Mass Drop and Filtering analysis in the procedure used to enhance the signal from a light Higgs decaying into $b\bar{b}$ at the LHC.

1 Introduction

In recent years, there has been a growing interest in jets studies in order to identify a boosted massive particle decaying hadronically, for instance the W boson [1–4], top quarks [5–8], supersymmetric particles [9, 10] and heavy resonances [11–13] (see also [14] for related work on general massive jets). Some of these studies revealed themselves to be successful in looking for a boosted light Higgs boson decaying into $b\bar{b}$ at the LHC [8, 15–17]. That of [15, 16] can be briefly summed up as follows: after having clustered the event with a radius R large enough to catch the b and \bar{b} from the Higgs decay into a single jet,¹ this jet can be analysed in 2 steps:

- A Mass Drop (MD) analysis that allows one to identify the splitting responsible for the large jet mass, i.e. separate the b and \bar{b} and thus measuring the angular distance R_{bb} between them, while suppressing as much QCD background as possible.
- A Filtering analysis where one reclusters the 2 resulting subjets with a smaller radius and takes the 3 highest- p_t subjets² obtained in order to keep the major part of the perturbative radiation while getting rid of as many underlying event (UE) and pile-up (PU) particles as possible (used also in [8, 17, 18], and a variant is proposed in [19]).

Concerning the MD analysis, the only thing we need to know is that we end up with 2 b -tagged jets, each with a radius roughly equal to R_{bb} . Notice that due to angular ordering [20–24], these 2 jets should capture the major part of the perturbative radiation from the $b\bar{b}$ dipole. The whole procedure is depicted in figure 1 (taken from [15]).

In this paper, we are going to focus on the Filtering analysis. One can generalize it with respect to its original definition using 2 parameters, n_{filt} and η_{filt} (as discussed also in [19]): after the MD analysis was carried out, one reclusters the 2 resulting subjets with a radius $R_{\text{filt}} = \eta_{\text{filt}} R_{bb}$ and takes the n_{filt} hardest jets obtained. Obviously, the larger the value of η_{filt} the more perturbative radiation we keep, but also the more the UE/PU degrades the Higgs peak. The same holds for n_{filt} . So there is a compromise to make between losing too much perturbative radiation and being contaminated by soft particles from UE/PU. In [15], the values that were found to give nice results were $n_{\text{filt}} = 3$ and $\eta_{\text{filt}} = \min(0.3/R_{bb}, 1/2)$. However, these values had been chosen on brief Monte-Carlo

¹The value chosen was $R = 1.2$.

²The value of 3 was found to work well in [15].

event generator study and one would like to gain a little more analytical control over them. One question would be for instance to understand even approximately how the optimal $(n_{\text{filt}}, \eta_{\text{filt}})$ values change when one increases the Higgs p_{t_H} cut, or when the PU becomes more and more important during the high luminosity running of the LHC. Though the MD and Filtering analysis were originally designed to identify a light Higgs boson, one should be aware that similar calculations may apply in other uses of filtering, for instance to study any boosted colorless resonance decaying hadronically, including W and Z bosons.

The article will be devoted in large part to the study of the dependence of the perturbative radiation loss with respect to the filtering parameters. As usual in this kind of work, large logarithms arise due to soft or collinear gluon emissions, and one is forced to deal with them in order to obtain reliable results in the region where the observable is sensitive to this kind of emission. We will thus compute analytically the two first orders in the leading soft logarithmic (LL) approximation when $n_{\text{filt}} = 2$ and to all-orders in the large- N_c limit³ when $n_{\text{filt}} = 2$ or 3 for small enough values of η_{filt} (section 2). With these in hand, and using a program that allows one to make all-orders leading-log calculations in the large- N_c approximation, we check the analytical results and examine if the small- η_{filt} limit and/or the truncation of the LL expansion can be trusted to estimate the loss of perturbative radiation in practice (section 3). Finally, in section 4, we will analyse the Higgs mass peak width due respectively to the loss of perturbative radiation and to the presence of UE/PU, before combining them in a simple and approximate but physically reasonable way in order to be able to conclude about the optimal parameters choices.

2 Non-global structure: analytical insights

2.1 The filtered Higgs mass: a non-global observable

It is now very well known [25–40] that soft or collinear gluons can give rise, in multiscale problems, to the appearance of large logarithms in the perturbative expansion of an observable, and more precisely in a region of phase space where it is sensitive to the soft or collinear divergences of QCD. In this article, the observable considered is $\Delta M = M_H - M_{\text{filtered jet}}$, where $M_{\text{filtered jet}}$ is the reconstructed Higgs-jet mass and M_H is its true mass. ΔM has the property that it is 0 when no gluon is emitted. We are interested in $\Sigma(\Delta M)$, the probability for the difference between the reconstructed and true Higgs masses to be less than a given ΔM . In this case, large soft logarithms have to be resummed at all-orders to obtain a reliable description of the small ΔM distribution.

For this observable, soft gluons emissions lead to powers of $\ln \frac{M_H}{\Delta M}$, whereas collinear gluons emissions leads to powers of $\ln \frac{R_{\text{bb}}}{R_{\text{filt}}}$. In this study, gluons are strongly ordered in energy (the first emitted gluon being the most energetic one, and so on), and we aim to control the $\left(\alpha_s \ln \frac{M_H}{\Delta M}\right)^k$ series, in a region where

$$\ln \frac{M_H}{\Delta M} \gg \ln \frac{R_{\text{bb}}}{R_{\text{filt}}} . \tag{2.1}$$

³ $N_c = 3$ denotes as usual the number of QCD colors.

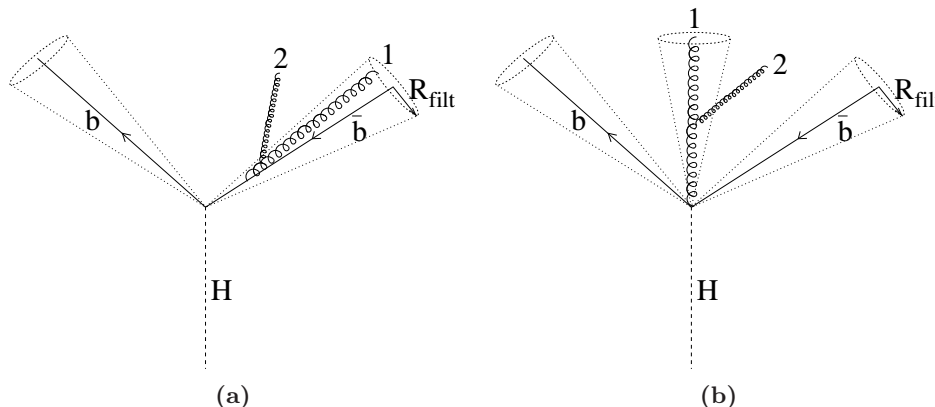


Figure 2. Configurations leading to non-global logarithms when (a) $n_{\text{filt}} = 2$ and (b) $n_{\text{filt}} = 3$. In each case, the hardest gluon 1, which is inside the filtered jet region, emits a softer gluon 2 outside the filtered jet region.

Therefore, at leading-log accuracy, one has to resum terms like

$$I_k(\Delta M) = f_k \left(\frac{R_{bb}}{R_{\text{filt}}} \right) \left(\alpha_s \ln \frac{M_H}{\Delta M} \right)^k, \quad (2.2)$$

where all the f_k are functions to be computed. We thus disregard all the subleading terms, i.e. those suppressed by at least one power of $\ln \frac{M_H}{\Delta M}$. Unfortunately, such a calculation is highly non-trivial due to the fact that the observable is *non-global*. This property, first studied in [35], means that it is sensitive to radiation in only a part of the phase space. In the case of ΔM , only emissions of gluons outside the filtered jets region contribute to the observable (cf figure 2). As a consequence of this property, one must consider soft gluons emissions not just from the $b\bar{b}$ dipole (usually called primary emissions, the only ones that would be present in QED) but also from the whole ensemble of already emitted gluons [35, 41]. As the number of gluons is increased, the geometry and the color structure of all these gluons become rapidly too complex to perform an analytical calculation. Therefore, to deal with this, one is forced to apply numerical Monte-Carlo calculations that can easily take care of the geometry. But the colour structure remains prohibitive, and one must usually also resort to the large- N_c approximation in order to go beyond the 2 first orders in perturbation theory [35, 41–43] (though some authors have derived some analytical results in special cases [44, 45] and others have examined contributions beyond the leading large- N_c approximation [46, 47]).

However, before considering numerical calculations, some results can be derived analytically at 2nd order for $n_{\text{filt}} = 2$ (where f_1 and f_2 are computed exactly) and $n_{\text{filt}} = 3$ (where only the leading behaviour of the f_k in $\ln \frac{R_{bb}}{R_{\text{filt}}}$ and N_c is looked for).

2.2 Some results for $n_{\text{filt}} = 2$

Perturbatively, one can write $\Sigma(\Delta M)$ as

$$\Sigma(\Delta M) = 1 + \sum_{k=1}^{\infty} I_k(\Delta M), \quad (2.3)$$

where $I_k(\Delta M)$ is the $\mathcal{O}(\alpha_s^k)$ contribution to the observable. To simplify the calculation, $\Sigma(\Delta M)$ will be computed using the anti- k_t algorithm [48], even if the numerical study will be done using the C/A algorithm [49, 50] to be in accordance with the choice in [15]. However, the anti- k_t algorithm is enough to catch the dominant behaviour of the leading-log series, in the sense that it does not affect the leading large collinear logarithm in the function f_k at small R_{filt} :

$$f_k \left(\frac{R_{\text{bb}}}{R_{\text{filt}}} \right) = a_k \ln^k \left(\frac{R_{\text{bb}}}{R_{\text{filt}}} \right) + \mathcal{O} \left(\ln^{k-1} \left(\frac{R_{\text{bb}}}{R_{\text{filt}}} \right) \right), \quad (2.4)$$

i.e. a_k is unchanged when moving from C/A to anti- k_t .⁴ This jet algorithm gives simpler results because the gluons outside the filtered jet region tend not to cluster with the ones inside. It is this property which ensures that the hardest jets in an event are generally perfect cones, as particles usually cluster with the hardest ones in their neighbourhood first [48].

As a first step, primary emissions are considered, defined to be those one would obtain if gluons were only emitted from the $b\bar{b}$ dipole (as for photons in QED).

2.2.1 Primary emissions

Due to the use of the anti- k_t algorithm, the result of the primary emissions can be easily shown to exponentiate, as will be roughly seen in the next section with the $\mathcal{O}(\alpha_s^2)$ analysis. Here, we just review the very well known result that the contribution to $\Sigma(\Delta M)$ from primary emissions, denoted $\Sigma^{(P)}(\Delta M)$, can be written as.⁵

$$\Sigma^{(P)}(\Delta M) = e^{I_1(\Delta M)}, \quad (2.5)$$

with:

$$I_1(\Delta M) = \int \frac{d^3\vec{k}_1}{(2\pi)^3 2|\vec{k}_1|} M(k_1) \left(\Theta(\vec{k}_1 \in J_{b\bar{b}}) + \Theta(\vec{k}_1 \notin J_{b\bar{b}}) \Theta(\Delta M - \Delta M(\vec{k}_1)) - 1 \right). \quad (2.6)$$

$M(k_1)$ is the matrix element squared for emitting one soft gluon from the $b\bar{b}$ dipole (the b quark is taken to be massless):

$$M(k_1) = 4\pi\alpha_s C_F \frac{2(p_b \cdot p_{\bar{b}})}{(p_b \cdot k_1)(k_1 \cdot p_{\bar{b}})}. \quad (2.7)$$

Concerning the notations, $\Theta(\vec{k}_1 \in J_{b\bar{b}})$ equals 1 when gluon 1 is emitted inside the jet regions around b and \bar{b} , denoted by $J_{b\bar{b}}$ (and is 0 otherwise), which, for $R_{\text{filt}} < R_{\text{bb}}$, is just 2 cones of radius R_{filt} centered on b and \bar{b} (figure 2a). Then, concerning the expression in brackets in eq. (2.6), we separate the 2 different regions where the gluon can be: either inside or outside the filtered Higgs jet. The first term $\Theta(\vec{k}_1 \in J_{b\bar{b}})$ means that the gluon

⁴When $R_{\text{filt}} \sim \frac{R_{\text{bb}}}{2}$, the discarding of the $\mathcal{O}(\ln^{k-1}(\frac{R_{\text{bb}}}{R_{\text{filt}}}))$ terms is not a priori justified, but figure 5, which compares numerical results obtained using C/A with analytical estimates using anti- k_t , supports the dominance of the leading collinear logarithms.

⁵The superfix (P) serves as a reminder that only primary emissions are being accounted for.

does not contribute to the observable (as it is kept in the Higgs jet, the reconstructed Higgs mass is the true Higgs mass: $\Delta M(k_1) = 0$). If the gluon is outside the filtered jet region (second term), then it does contribute to the observable:

$$\Delta M(k) \sim k_t \frac{M_H}{p_{tH}}, \tag{2.8}$$

up to prefactors that can be neglected in the leading-log approximation, see appendix A. Finally, the -1 stands for the virtual corrections, for which there's obviously no loss of mass for the Higgs, and whose matrix element is just the opposite of the soft real one.⁶ One thus obtains:

$$I_1(\Delta M) = - \int_{\vec{k}_1 \notin J_{b\bar{b}}} \frac{d^3 \vec{k}_1}{(2\pi)^3 2|\vec{k}_1|} M(k_1) \Theta(\Delta M(k_1) - \Delta M). \tag{2.9}$$

The computation of this integral in the boosted regime, where $p_{tH} \gg M_H$, or equivalently $R_{bb} \ll 1$, is done in appendix A. From now on, we will essentially use $\eta_{\text{filt}} = R_{\text{filt}}/R_{bb}$ instead of R_{filt} and we define $n \equiv n_{\text{filt}}$ and $\eta \equiv \eta_{\text{filt}}$ for more clarity in mathematical formulae. In order to keep in mind that it depends on the 2 parameters of the Filtering analysis, the distribution $\Sigma(\Delta M)$ is renamed $\Sigma^{(n)}(\eta, \Delta M)$. What we obtain at fixed coupling is the following:⁷

$$\Sigma^{(2),(P)}(\eta, \Delta M) = e^{-\frac{\alpha_s C_F}{\pi} J(\eta) \ln \frac{M_H}{\Delta M}}, \tag{2.10}$$

with

$$J(\eta) = 2 \ln \left(\frac{1 - \eta^2}{\eta^2} \right) \quad \text{if } \eta < \frac{1}{2}, \tag{2.11}$$

$$= \frac{8}{\pi} \int_{\eta}^{+\infty} \frac{du}{u(u^2 - 1)} \arctan \left(\frac{u - 1}{u + 1} \sqrt{\frac{2u + 1}{2u - 1}} \right) \quad \text{if } \frac{1}{2} < \eta < 1. \tag{2.12}$$

We give the value of $J(1)$, a quantity that is important to discuss some aspects of the results obtained in the following sections:

$$J(1) \simeq 0.646. \tag{2.13}$$

Notice that the case $\eta > 1$ will not be used in this study, but is mentioned in appendix A. The function $J(\eta)$ is plotted in figure 3.

Two remarks can be made:

1. The result does not depend on the energy fraction z of the Higgs splitting into $b\bar{b}$.
2. When $\eta \ll 1$, another large logarithm $\ln \frac{1}{\eta}$ appears due to collinear enhancement.

⁶Even if the result seems obvious here, this way of doing the calculation can be easily generalised to higher orders and other kinds of jet algorithms.

⁷To obtain the result at running coupling, one simply makes the replacement (see eqs. (3.1), (3.2) later in the article):

$$\alpha_s \ln \frac{M_H}{\Delta M} \rightarrow \frac{1}{2\beta_0} \ln \left(\frac{1}{1 - 2\beta_0 \alpha_s(M_H) \ln \frac{M_H}{\Delta M}} \right).$$

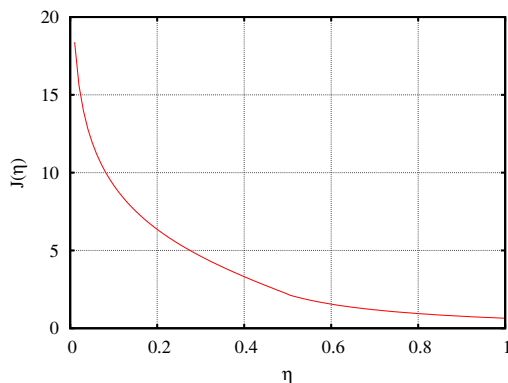


Figure 3. The coefficient $J(\eta)$ in front of the primary soft logarithm $\ln \frac{M_H}{\Delta M}$.

2.2.2 Non-global contributions

Now, we turn to the $\mathcal{O}(\alpha_s^2)$ term, and more precisely to the contribution of the non-global terms that have to be added to the primary logarithms computed in the previous section. That corresponds to the analysis of $I_2(\Delta M)$ in the perturbative expansion of $\Sigma(\Delta M)$ from eq. (2.3). The matrix element squared for 2 real gluons emission is expressed as [24, 35, 51, 52]:

$$M(k_1 \text{ real}, k_2 \text{ real}) = (4\pi\alpha_s)^2(W_1 + W_2), \tag{2.14}$$

with

$$W_1 = 4C_F^2 \frac{(p_b \cdot p_{\bar{b}})}{(p_b \cdot k_1)(k_1 \cdot p_{\bar{b}})} \frac{(p_b \cdot p_{\bar{b}})}{(p_b \cdot k_2)(k_2 \cdot p_{\bar{b}})}, \tag{2.15}$$

$$W_2 = 2C_F C_A \frac{(p_b \cdot p_{\bar{b}})}{(p_b \cdot k_1)(k_1 \cdot p_{\bar{b}})} \left(\frac{(p_b \cdot k_1)}{(p_b \cdot k_2)(k_2 \cdot k_1)} + \frac{(p_{\bar{b}} \cdot k_1)}{(p_{\bar{b}} \cdot k_2)(k_2 \cdot k_1)} - \frac{(p_b \cdot p_{\bar{b}})}{(p_b \cdot k_2)(k_2 \cdot p_{\bar{b}})} \right). \tag{2.16}$$

This expression is valid when there is a strong energy ordering between the two real gluons 1 and 2, either $E_1 \gg E_2$ or $E_2 \gg E_1$ (the formula is completely symmetric under the interchange $k_1 \leftrightarrow k_2$). For the cases with one or both gluons being virtual, the following matrix elements are obtained, valid only when $E_1 \gg E_2$ [52]:

$$\begin{aligned} M(k_1 \text{ real}, k_2 \text{ virt}) &= -(4\pi\alpha_s)^2(W_1 + W_2), \\ M(k_1 \text{ virt}, k_2 \text{ real}) &= -(4\pi\alpha_s)^2 W_1, \\ M(k_1 \text{ virt}, k_2 \text{ virt}) &= (4\pi\alpha_s)^2 W_1. \end{aligned} \tag{2.17}$$

Using these properties, separating the 4 phase space regions depending on whether the gluons are inside or outside the filtered jet region in the same way as was done for I_1 , and defining dk as:

$$dk = \frac{d^3 \vec{k}}{(2\pi)^3 2|\vec{k}|}, \tag{2.18}$$

we can then write I_2 in the following form:

$$I_2 = \int dk_1 dk_2 (4\pi\alpha_s)^2 \Theta(E_1 - E_2)$$

$$\begin{aligned}
& \times \left\{ \Theta(k_1 \in J_{b\bar{b}}) \Theta(k_2 \in J_{b\bar{b}}) ((W_1 + W_2) - (W_1 + W_2) - W_1 + W_1) \right. \\
& \quad + \Theta(k_1 \in J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) ((W_1 + W_2) \Theta(\Delta M - \Delta M(k_2)) \\
& \quad - (W_1 + W_2) - W_1 \Theta(\Delta M - \Delta M(k_2)) + W_1) \\
& \quad + \Theta(k_1 \notin J_{b\bar{b}}) \Theta(k_2 \in J_{b\bar{b}}) ((W_1 + W_2) \Theta(\Delta M - \Delta M(k_1)) \\
& \quad - (W_1 + W_2) \Theta(\Delta M - \Delta M(k_1)) - W_1 + W_1) \\
& \quad \left. + \Theta(k_1 \notin J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) ((W_1 + W_2) \Theta(\Delta M - \Delta M(k_1, k_2)) \right. \\
& \quad \left. - (W_1 + W_2) \Theta(\Delta M - \Delta M(k_1)) - W_1 \Theta(\Delta M - \Delta M(k_2)) + W_1) \right\}. \quad (2.19)
\end{aligned}$$

For each phase space region, the 4 terms $(k_1, k_2) = (\text{real}, \text{real}) - (\text{real}, \text{virt}) - (\text{virt}, \text{real}) + (\text{virt}, \text{virt})$ are considered. The strong energy ordering $E_1 \gg E_2$ implies that $\Delta M(k_1, k_2) = \Delta M(k_1)$, and one immediately gets:

$$\begin{aligned}
I_2 &= \int dk_1 dk_2 (4\pi\alpha_s)^2 \Theta(E_1 - E_2) \Theta(k_1 \notin J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) W_1 \Theta(\Delta M(k_2) - \Delta M) \\
&\quad - \int dk_1 dk_2 (4\pi\alpha_s)^2 \Theta(E_1 - E_2) \Theta(k_1 \in J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) W_2 \Theta(\Delta M(k_2) - \Delta M), \\
&= I_2^{(P)}(\Delta M) + I_2^{(NG)}(\Delta M), \quad (2.20)
\end{aligned}$$

where $I_2^{(P)}(\Delta M)$ corresponds to the first integral containing the function W_1 whereas $I_2^{(NG)}(\Delta M)$ corresponds to the second integral with the function W_2 . $I_2^{(P)}$ is just the second order contribution to the primary emissions, already computed above. To be convinced, one can notice that $(4\pi\alpha_s)^2 W_1$ can be expressed as the product of 2 one-gluon matrix elements $M(k_1)M(k_2)$ and, when $E_1 \gg E_2$,

$$\Theta(\Delta M(k_2) - \Delta M) = \Theta(\Delta M(k_2) - \Delta M) \Theta(\Delta M(k_1) - \Delta M), \quad (2.21)$$

if k_1 and k_2 belong to the same phase space region. Therefore $I_2^{(P)}$ can be written in a more symmetric way:

$$\begin{aligned}
I_2^{(P)}(\Delta M) &= \frac{1}{2} \int dk_1 dk_2 (4\pi\alpha_s)^2 \Theta(k_1 \notin J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) \\
&\quad \times W_1 \Theta(\Delta M(k_1) - \Delta M) \Theta(\Delta M(k_2) - \Delta M), \\
&= \frac{1}{2} \left(\int dk \Theta(k \notin J_{b\bar{b}}) M(k) \Theta(\Delta M(k) - \Delta M) \right)^2, \\
&= \frac{1}{2} (I_1(\Delta M))^2, \quad (2.22)
\end{aligned}$$

so that it corresponds to the second order perturbative expansion of the result eq. (2.5), obtained with primary emissions only.

The important term for this section is the one containing W_2 , denoted by $I_2^{(NG)}$. As mentioned in section 2.1, it receives a non-zero contribution when the hardest gluon 1 is emitted inside the filtered jet region whereas the softest gluon 2 is emitted outside. For the opposite configuration, there is an exact cancellation between gluon 2 being real and

virtual. Here again the computation of $I_2^{(NG)}$ is postponed to appendix A, giving directly what will help to interpret some results later. S_2 is defined such that

$$I_2^{(NG)}(\eta, \Delta M) = \frac{1}{2} C_F C_A \left(\frac{\alpha_s}{\pi} \ln \left(\frac{M_H}{\Delta M} \right) \right)^2 S_2(\eta), \quad (2.23)$$

where we explicitly introduce the dependence on η and we factorize out the soft divergence, still revealed in the large logarithm $\ln \frac{M_H}{\Delta M}$. When $\eta < 1/2$, the result for S_2 can be written as:

$$\begin{aligned} S_2(\eta) &= -\frac{\pi^2}{3} + 8 \int_0^1 \frac{du_1}{u_1} \int_0^1 \frac{du_2}{u_2} \left(\frac{1}{\sqrt{(1-\eta^2(u_1^2+u_2^2))^2 - 4\eta^4 u_1^2 u_2^2}} - \frac{1}{1-\eta^2(u_1^2+u_2^2)} \right), \\ &= -\frac{\pi^2}{3} + 4\eta^4 + 12\eta^6 + \mathcal{O}(\eta^8). \end{aligned} \quad (2.24)$$

The important point to notice in this result is the absence of collinear logarithms, which would appear as $\ln \frac{1}{\eta}$, contrary to the primary emission case (eq. (2.11)). So that the primary emissions dominate for this observable, at least for η sufficiently small.

As mentioned in previous studies [35, 53], one notices the presence of “ π^2 terms” in non-global results at second order.

2.3 Some results for $n_{\text{filt}} = 3$

The goal in this section is to have an estimate of the analytical behavior in the large N_c limit of $\Sigma^{(n)}(\eta, \Delta M)$ for $n = 3$, which is the probability of having no second gluon emission leading to a $\Delta M'$ greater than ΔM . Notice that, contrary to the previous part where we obtained the function $\Sigma^{(2)}$, only the leading behavior in $L = \ln \frac{1}{\eta}$ and N_c will be derived, so that in this context $\Sigma^{(2)}$ can be simply written:⁸

$$\Sigma^{(2)}(L, t) = e^{-4N_c L t} \quad (2.25)$$

where for further convenience we introduce the parameter $t = \frac{\alpha_s}{2\pi} \ln \frac{M_H}{\Delta M}$ and we change the arguments of Σ which becomes now a function of L and t . In this formula, $2L = 2 \ln \frac{R_{bb}}{R_{\text{filt}}}$ can be interpreted as the “logarithmic size” of the $b\bar{b}$ dipole, i.e. the allowed phase space in rapidity for an emission from this dipole (in its center of mass frame) outside the jet region. The parameter t means that this emission cannot occur with a t' between 0 and t .

Now we turn to $\Sigma^{(3)}(L, t)$. To have no second gluon emission in $[0, t]$, either there is no first gluon emission in $[0, t]$ outside the jet region (which corresponds to $\Sigma^{(2)}(L, t)$), or there is such an emission but the new dipole configuration is prohibited from emitting a second gluon in $[0, t]$ outside the jet region. This is depicted in figure 4. As the calculation is done in the large- N_c limit, after the emission of a first gluon, the second one cannot be emitted from the $b\bar{b}$ dipole, but only from the bg and $\bar{b}g$ ones. Figure 4 can be translated mathematically as:

$$\Sigma^{(3)}(L, t) \simeq \Sigma^{(2)}(L, t) + \int_0^t dt' 4N_c \Sigma^{(2)}(L, t') \int_0^L dl \Sigma^{(2)}(L, t-t') \Sigma^{(2)}(l, t-t'). \quad (2.26)$$

⁸This results simply from the combination of equations (2.10) and (2.11) with $\eta \ll 1$, and $C_F = \frac{N_c}{2}$ in the large N_c limit.

$$\Sigma^{(3)}(L, t) \simeq \text{diagram 1} + \int dldt' \text{diagram 2}$$

Figure 4. How to compute the leading behavior of $\Sigma^{(3)}(L, t)$ from $\Sigma^{(2)}(L, t)$ when $L \gg 1$ and $N_c \gg 1$. In the second term, t' is the gluon's emission scale.

Notice that L_{bg} , the logarithmic size of the bg dipole in figure 4, does not depend on l in the leading collinear log approximation.⁹ In this expression, $4N_c L \Sigma^{(2)}(L, t') dt'$ is the probability not to emit the first gluon in $[0, t']$ and to emit it only at $t \in [t', t' + dt']$. The remaining part $\frac{1}{L} \int_0^L dl \Sigma^{(2)}(L, t-t') \Sigma^{(2)}(l, t-t')$ is the probability to emit no second gluon from the bg and $\bar{b}g$ dipoles in $[t', t]$. Using eq. (2.25) for $\Sigma^{(2)}$, $\Sigma^{(3)}$ is then given by:

$$\Sigma^{(3)}(L, t) \simeq e^{-4N_c L t} \left(1 + \int_0^{4N_c L t} dt' \frac{1 - e^{-t'}}{t'} \right). \quad (2.27)$$

Two limits can be considered:

$$\Sigma^{(3)}(L, t) \simeq \begin{cases} 1 - \frac{3}{4}(4N_c L t)^2 + \mathcal{O}((4N_c L t)^3 + N_c t) & \text{if } 4N_c L t \ll 1, \\ e^{-4N_c L t} (\ln(4N_c L t) + \mathcal{O}(1)) & \text{if } 4N_c L t \gg 1. \end{cases} \quad (2.28)$$

The limit $4N_c L t \ll 1$ reveals two important aspects:

1. One can notice the absence of the $\mathcal{O}(Lt)$ term, which is indeed the goal of the filtering analysis as it was presented in its original version [15]: it is intended to catch the major part of the $\mathcal{O}(\alpha_s)$ perturbative radiation. It cannot catch *all* the $\mathcal{O}(\alpha_s)$ contribution because a hard gluon emitted at an angle $\theta > R_{bb}$ from the b and \bar{b} escapes the filtering process as it is rejected by the Mass Drop analysis. Therefore, when $4N_c L t \ll 1$, the expansion eq. (2.28) misses a term $\mathcal{O}(N_c t)$, but this is legitimate in a leading collinear log estimate. Notice that the missing term is simply $-J(1)N_c t$ where $J(\eta)$ was given in eq. (2.12).

⁹One can easily show the following relation:

$$L_{bg} = 2L + \mathcal{O}(e^{l-L}).$$

If we introduce the neglected component of L_{bg} in the $\Sigma^{(3)}$ calculation eq. (2.26), then we have to compute an integral of the form

$$\int_0^L dl \frac{1 - e^{(l + \mathcal{O}(e^{l-L}))t}}{l + \mathcal{O}(e^{l-L})}.$$

Expanding the exponential and keeping the term of order k gives

$$\int_0^L dl \left(l + \mathcal{O}(e^{l-L}) \right)^{k-1} t^k = \frac{(Lt)^k}{k} + \mathcal{O}(L^{k-2} t^k).$$

The leading $\mathcal{O}((Lt)^k)$ term is already taken into account in eq. (2.27). Therefore, including the l dependent component of L_{bg} gives rise to terms of the form $N_c^k L^{k-2} t^k$ at order k , suppressed by 2 powers of L with respect to the leading one.

ΔM (GeV)	1	2	5	10	20	50	115
t	0.141	0.108	0.075	0.054	0.036	0.016	0

Table 1. Correspondence between ΔM and t for some particular values.

2. it shows that the purely non-global result for $n = 3$ contains large collinear logarithms L , contrary to the case $n = 2$ (eq. (2.24)). Indeed, the primary result for $n = 3$ at second order can be proven to behave as¹⁰ $-32C_F^2(Lt)^2$ at order α_s^2 , so that the S_2 term for $n = 3$ should be equivalent to $-8C_FC_A(Lt)^2$ at large L .

Having understood some analytical features of the Filtering analysis, we now examine what can be learnt from a numerical calculation of the reconstructed Higgs mass observable.

3 Non-global structure: numerical results

In all that follows t is defined so as to gather all the information about the soft logarithms in a running coupling framework:

$$\begin{aligned}
 t &= \frac{1}{2\pi} \int_0^{p_{tH}} \frac{dk_t}{k_t} \alpha_s \left(k_t \frac{M_H}{p_{tH}} \right) \Theta(\Delta M(k) - \Delta M), \\
 &= \frac{1}{2\pi} \int_{p_{tH} \frac{\Delta M}{M_H}}^{p_{tH}} \frac{dk_t}{k_t} \alpha_s \left(k_t \frac{M_H}{p_{tH}} \right), \\
 &= \frac{1}{4\pi\beta_0} \ln \left(\frac{1}{1 - 2\beta_0\alpha_s(M_H) \ln \frac{M_H}{\Delta M}} \right), \tag{3.1}
 \end{aligned}$$

where the last equality holds at the one-loop level and $\beta_0 = \frac{11C_A - 2n_f}{12\pi}$. The argument of α_s was taken as the gluon's transverse momentum with respect to the Higgs boson direction, of order $k_t \frac{M_H}{p_{tH}}$, k_t being its transverse momentum with respect to the beam. In the case of a fixed coupling constant α_s , the definition for t here coincides with that of section 2.3:

$$t = \frac{\alpha_s}{2\pi} \ln \frac{M_H}{\Delta M}. \tag{3.2}$$

But from now on, and unless stated otherwise, t is given in the running coupling framework, eq. (3.1), and the function $\Sigma(\eta, \Delta M)$ is rewritten as $\Sigma(\eta, t)$.

To get an idea of the range of values covered by t , table 1 presents a few t values corresponding to a given ΔM for a Higgs mass of 115 GeV ($\alpha_s(M_H) = 0.114$). It reveals that the physical values for t are below 0.15.

To numerically investigate non-global observables, two approaches can be followed:

¹⁰In fact, one can show the following general estimate for the primary emissions in the leading soft and collinear approximations:

$$\Sigma^{(n)}(L, t) = e^{-8C_FLt} \sum_{k=0}^{n-2} \frac{(8C_FLt)^k}{k!}.$$

- an all-orders approach where one resums the leading-logs at all-orders in the large- N_c limit, the output being the function $\Sigma(t)$, i.e. the probability that the loss of perturbative emission results in a Higgs mass in the range $[M_H - \Delta M(t), M_H]$, with

$$\Delta M(t) = M_H e^{-\frac{1}{2\beta_0\alpha_s}(1-e^{-4\pi\beta_0 t})}, \tag{3.3}$$

simply obtained by inverting the relation eq. (3.1).

- a fixed-order approach where the first few coefficients from the expansion of $\Sigma(t)$ are computed in the large- N_c limit. More precisely, if $\Sigma(t) = \sum_{k=0}^{\infty} \frac{c_k}{k!} (N_c t)^k$, then the program returns the first few coefficients c_k .

From a numerical point of view, the way to write an all-orders program was explained in [35]. On the other side, a result at fixed-order may be obtained by developing a systematic approach like the one presented at second order in eq. (2.19). For the filtered Higgs jet mass observable, we used the Fastjet package [54] to perform the clustering (and mass-drop + filtering) with the C/A algorithm, consistently with the choice made in [15].

As the all-orders program gives immediately what we are looking for, which is $\Sigma(t)$, we will use it (section 4.1) to compute the perturbative Higgs width. But in order to check it and be confident with the results obtained, we compare them with the previous analytical estimates and see how well the perturbative leading log series fits them. This leads us to study the behaviour of the higher order terms and to gain a better understanding of the convergence and structure of the non-global series. Though treated in more details in appendix C, the main points are mentioned in this section.

3.1 Comparison with analytics

Using the all-orders Monte-Carlo program, a comparison between the all-orders numerical curves obtained using the C/A algorithm and their corresponding analytical estimates obtained previously with anti- k_t in eqs. (2.25), (2.27) can be done. The results are presented in figure 5 and show good agreement, at least in the region of physical t values.

Notice that the slight discrepancy between analytical estimations and numerics starts to occur at $t > 0.1$, which is at the edge of the physical region (cf table 1), beyond which ΔM would be below the perturbative scale of around 1 GeV. This agreement manifests that:

- In the physical region, the leading terms in $(\alpha_s L t)^k$, with $L = \ln \frac{1}{\eta}$ seem to completely dominate and we do not need to compute the subleading corrections.
- One can use these analytical expressions to get an accurate estimate of the reconstructed Higgs peak width.

3.2 Comparison with fixed-order results

The structure of the non-global series at fixed-order is now examined so as to independently cross-check the all-orders program and to understand if the perturbative leading-log series can be usefully truncated.

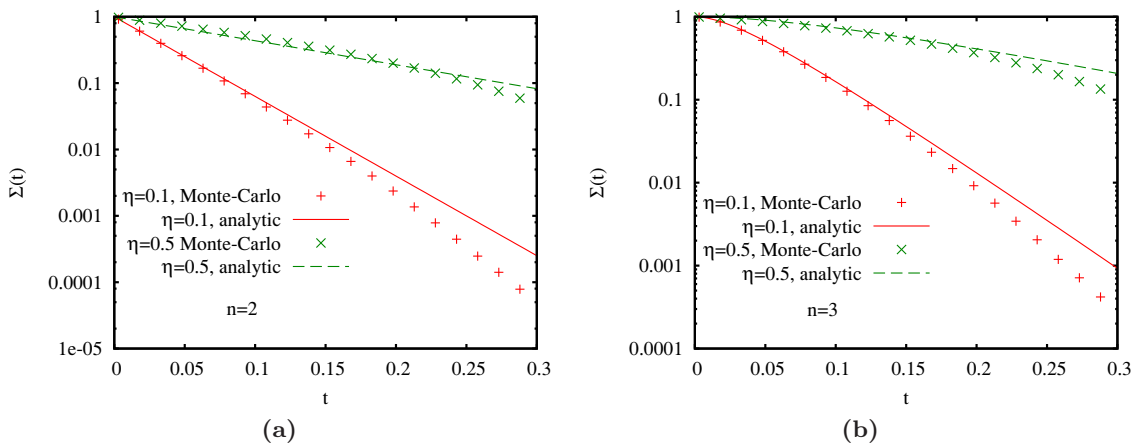


Figure 5. Comparison between numerical all-orders result (obtained using C/A algorithm) and leading collinear logarithm estimate of $\Sigma(t)$ derived with anti- k_t for (a) $n = 2$ and (b) $n = 3$ for 2 values of η : 0.1 and 0.5.

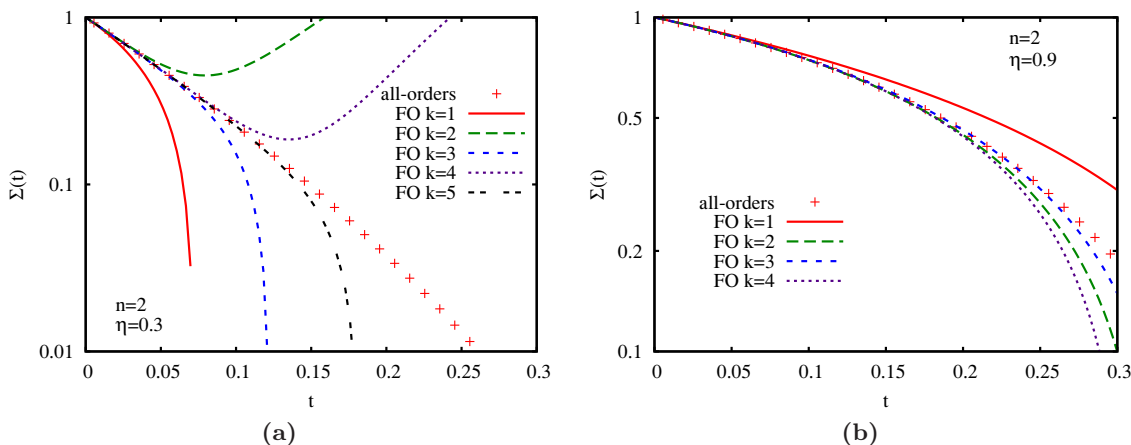


Figure 6. Comparison between fixed-order (FO) expansion and all-orders result when $n = 2$ for (a) $\eta = 0.3$ and (b) $\eta = 0.9$. Both fixed-order and all-orders results were obtained using the C/A algorithm.

As an example, figure 6 compares the all-orders result to the fixed-order ones up to α_s^5 for $n = 2$ and two different values of η (only the coefficients with an uncertainty of at most a few percent are plotted¹¹). The curves are represented up to $t = 0.3$, which is far beyond the physical region but is instructive to study the convergence of the series.

The left plot for $\eta = 0.3$ shows a nice convergence of the perturbative series eq. (2.3), as the t range for which the all-orders and fixed-order curves coincide grows with k . However, the second plot for $\eta = 0.9$ gives an unexpected result: the fourth order diverges with respect to the third one, in the sense that the point of disagreement is shifted to smaller t . The question arises whether this divergence will remain at higher orders. To answer it,

¹¹This uncertainty obviously increases with the perturbative order, but also with η because at small η , the coefficients are sensitive to the large logarithm $\ln \frac{1}{\eta}$, which is easy to compute.

one needs to go further in perturbation theory. In appendix C, a parallel is made between the filtered Higgs jet observable and the slice observable, studied for instance in [41], for which, due to computational speed, it is possible to obtain reliable coefficients up to order 6. The same effect is observed and is even enhanced at orders 5 and 6. Therefore, it seems that the fixed-order information cannot be safely used in general: one has to be aware that the leading-log large- N_c non-global series may be divergent for any value of t .

4 Choice of the filtering parameters

In the previous sections we examined the structure and convergence of the perturbative leading-log series, analytically and numerically. We could then cross-check the analytical expressions and the fixed-order approach with the all-orders program, which we are going to use throughout this part.

We would like to decide how one should choose the filtering parameters (n, η) depending on the level of UE and PU as well as the p_t of the Higgs boson. Here, we do not claim to make an exact and complete analysis, but we want to obtain some estimates. First, we consider the width of the Higgs mass distribution separately in presence of perturbative radiation (using the all-orders results) and UE/PU (using a simple model for it). Then, we try to minimize the Higgs width in presence of both of these effects. Finally, we will estimate hadronisation corrections.

In all this part, we set the Higgs mass M_H at 115 GeV, as in [15].

4.1 Study of the Higgs perturbative width

As we could see in the previous sections, even without considering additional particles from UE/PU, $\Delta M \equiv M_H - M_{\text{filtered jet}} \neq 0$ because of the loss of perturbative radiation. The Higgs boson thus acquires a perturbative width, denoted δM_{PT} . At first sight, knowing the distributions $\Sigma^{(n)}(\eta, \Delta M)$, one might simply define it as:

$$\delta M_{\text{PT}} = 2\sqrt{\langle \Delta M^2 \rangle - \langle \Delta M \rangle^2}, \tag{4.1}$$

as we do for gaussian distributions for instance. Unfortunately, if we simply take $n = 2$ as an example and if we consider the primary emission result eq. (2.10), we can deduce the following distribution for ΔM :

$$\frac{d\Sigma^{(2)}(\eta, \Delta M)}{d\Delta M} = \frac{\alpha_s C(\eta)}{M_H^{\alpha_s C(\eta)}} \frac{1}{\Delta M^{1-\alpha_s C(\eta)}}, \tag{4.2}$$

with

$$C(\eta) = \frac{C_F J(\eta)}{\pi}. \tag{4.3}$$

Computing $\langle \Delta M \rangle$ and $\langle \Delta M^2 \rangle$ implies dealing with integrals of the form

$$\int_0^{M_H} \frac{d\Delta M}{\Delta M^{1-\alpha_s C(\eta)}} \Delta M = \int_0^{M_H} d\Delta M \Delta M^{\alpha_s C(\eta)}, \tag{4.4}$$

$$\int_0^{M_H} \frac{d\Delta M}{\Delta M^{1-\alpha_s C(\eta)}} \Delta M^2 = \int_0^{M_H} d\Delta M \Delta M^{1+\alpha_s C(\eta)}. \tag{4.5}$$

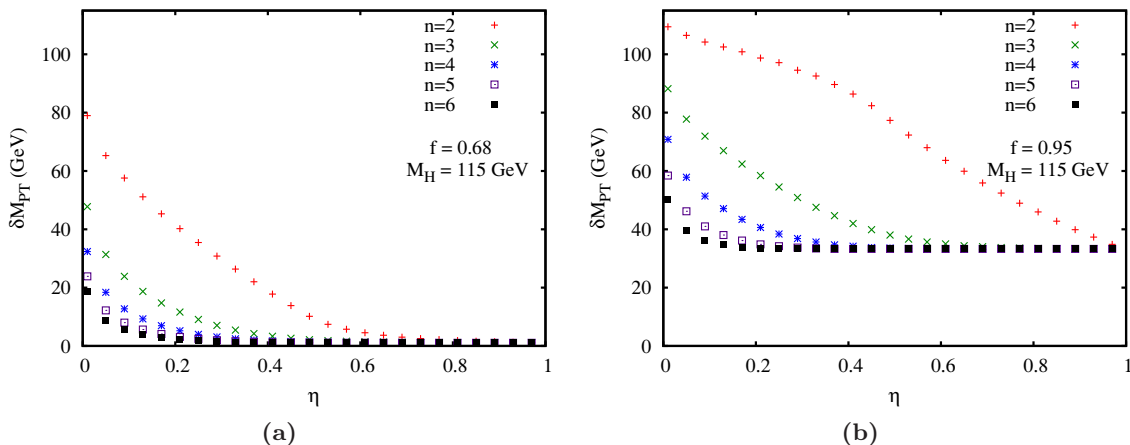


Figure 7. Perturbative width of the Higgs boson (in GeV) as a function of η for several values of n when (a) $f = 0.68$ and (b) $f = 0.95$.

Such integrals give a large importance to the $\Delta M \sim M_H/2$ region, where there should be very few events, and do not describe what happens in the neighbourhood of the peak near $\Delta M = 0$. Therefore, the definition eq. (4.1) does not seem adequate for the perturbative width. That’s why we shall adopt another definition, adapted from [55]. The Higgs perturbative width is defined as the size δM_{PT} for which a given fraction f of events satisfy $0 < \Delta M < \delta M_{PT}$. Using the all-orders function previously computed, this is equivalent to solving the equation $\Sigma^{(n)}(\eta, \Delta M) = f$. This leads to the width function $\delta M_{PT}(n, \eta, f)$.

Figure 7 shows δM_{PT} as a function of η for $n = 2 \dots 6$. When $\Delta M \sim 50$ GeV (i.e. $\sim M_H/2$), one should be aware that soft approximation loses sense and results on these plots should no longer be taken seriously. We chose the values $f = 0.68$ and $f = 0.95$, corresponding respectively to 2σ and 4σ for gaussian distributions, to show that the Higgs mass perturbative distribution is not gaussian (otherwise, going from 2σ to 4σ would have multiplied the width by a factor of 2, see also eq. (4.2)). One important thing to notice is a kind of “saturation” effect that one observes for η close enough to 1 for every fraction f . It manifests itself as a flat curve at a value $\delta M_{PT} = \delta M_{sat}(f)$, independent of n . For instance, $\delta M_{sat}(f = 0.68) \simeq 1$ GeV and $\delta M_{sat}(f = 0.95) \simeq 33$ GeV. This can be understood simply by considering that when the radius of the filtering is large enough, say $\eta > \eta_{sat}(n)$, it captures (almost) all the particles resulting from the Mass Drop analysis, i.e. all those that are within angular distance R_{bb} from b or \bar{b} , but it still fails to capture particles outside the Mass Drop region.¹² Of course, the larger n , the smaller $\eta_{sat}(n)$ as we keep more jets. This saturation property is equivalent to saying that all the functions $\Sigma^{(n)}(\eta, \Delta M)$ become independent of n and η when $\eta > \eta_{sat}(n)$.

For the rest of this analysis we keep the value $f = 0.68$, even if it is not clear which value should be chosen, and more generally what should be the relevant definition of the Higgs perturbative width. However, we will mention in section 4.4 what happens if we vary f between $f = 0.5$ and $f = 0.8$, so as to obtain a measure of the uncertainty of the calculations.

¹²The probability to emit a gluon outside the MD region in $[0, t]$ is roughly given by $1 - e^{-J(1)N_c t}$.

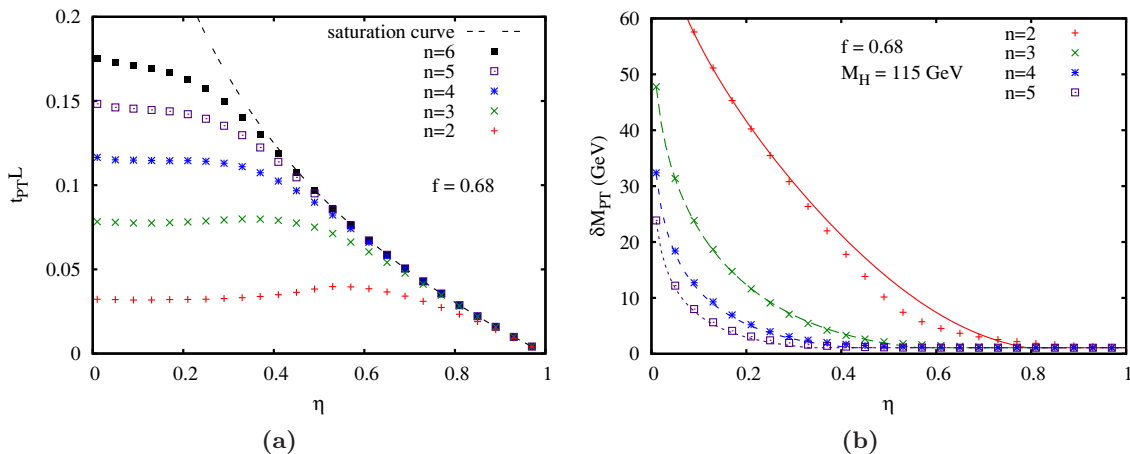


Figure 8. (a) $t_{PT}L$ as a function of η for $f = 0.68$ and different values of n . The saturation curve simply comes from the fact that all widths saturate to the same constant δM_{sat} for η large enough, and its equation is therefore given by $t_{\text{sat}}L$ (with $t_{\text{sat}}(f = 0.68) \simeq 0.136$). (b) δM_{PT} as a function of η for $f = 0.68$ and different values of n (curves with points). For each n is also represented the corresponding approximate width (lines) given by eq. (4.11).

The curves in figure 7 only give us an overview of the scales involved in the Higgs boson width. But one can go a little further. At small η , we should get a large collinear enhancement revealing itself as a large logarithm $L = \ln \frac{1}{\eta}$ multiplying t . The perturbative expansion is thus a series in $(N_c L t)^k$. As a direct consequence, at small η , the all-orders function $\Sigma^{(n)}(\eta, t)$ can be written as a function of a single variable $\Sigma^{(n)}(N_c L t)$. Solving the “width equation”

$$\Sigma^{(n)}(N_c L t) = f, \tag{4.6}$$

gives

$$t_{PT} = \frac{C_{PT}(n, f)}{L}, \tag{4.7}$$

where t_{PT} is simply related to δM_{PT} by

$$t_{PT} = \frac{1}{4\pi\beta_0} \ln \left(\frac{1}{1 - 2\beta_0\alpha_s(M_H) \ln \frac{M_H}{\delta M_{PT}}} \right), \tag{4.8}$$

and where $C_{PT}(n, f)$ is a function, independent of η , which increases with n and decreases when f increases. This is confirmed by figure 8a which shows that $t_{PT}L$ is indeed independent of η as long as η and n are not too large.

As an example, for $n = 2$, let us take the simple result $\Sigma^{(2)}(L, t) = e^{-4N_c L t}$ from eq. (2.25) in the small η limit. It was shown in section 3.1 that this result is very close to the all-orders one in the physical t region. Solving $\Sigma^{(2)}(L, t) = f$ immediately implies

$$C_{PT}(2, f) = \frac{\ln \frac{1}{f}}{4N_c}, \tag{4.9}$$

which, for $f = 0.68$, gives $C_{PT} \simeq 0.032$ in accordance with figure 8a.

n	2	3	4	5
C_{PT}	0.032	0.078	0.117	0.149
η_{sat}	0.79	0.56	0.42	0.34

Table 2. C_{PT} and η_{sat} as a function of n when $f = 0.68$.

One observes that $t_{\text{PT}}L$ is not strictly speaking a constant for higher n values. This may be due to the saturation effects discussed above. Indeed, even at large L , the perturbative expansion is not only a function of Lt but also of t for the lowest orders, as mentioned at the end of section 2.3:

$$\Sigma^{(n)}(L, t) = 1 + \sum_{k=1}^{n-2} a_k t^k + \sum_{k=n-1}^{+\infty} \left(a_k (Lt)^k + \mathcal{O}(L^{k-1}t^k) \right), \quad (4.10)$$

If we only had QED like emissions, i.e. primary ones, with the use of the anti- k_t jet algorithm, we would obtain $a_k = \frac{(-J(1)N_c)^k}{k!}$ for $k \leq n - 2$, where $J(\eta)$ was derived in section 2.2. As n increases, the term $a_1 t$ becomes more and more important with respect to $a_{n-1}(Lt)^{n-1}$, leading to larger and larger deviations from the simple law $t_{\text{PT}}L = \text{constant}$. However, until $n = 5$, assuming $t_{\text{PT}}L$ is a constant at small η seems a good approximation. Therefore, using eq. (4.8), one can model the Higgs perturbative width in the following form:

$$\delta M_{\text{PT}}(n, L, f) = \begin{cases} M_H e^{-\frac{1}{2\beta_0\alpha_s} \left(1 - e^{-4\pi\beta_0 \frac{C_{\text{PT}}(n,f)}{L}} \right)} & \text{if } \eta < \eta_{\text{sat}}(n, f), \\ \delta M_{\text{sat}}(f) & \text{if } \eta > \eta_{\text{sat}}(n, f). \end{cases} \quad (4.11)$$

$\eta_{\text{sat}}(n, f)$ is given by the intersection between the curve $t_{\text{PT}} = C_{\text{PT}}/L$ and $t_{\text{PT}} = t_{\text{sat}}$. Therefore:

$$\eta_{\text{sat}}(n, f) = e^{-\frac{C_{\text{PT}}(n,f)}{t_{\text{sat}}}}. \quad (4.12)$$

Table 2 shows C_{PT} and η_{sat} for $f = 0.68$ and different n values. Figure 8b shows the curves corresponding to the parametrisation eq. (4.11). We can see that it works rather well for all values of n except $n = 2$ in the region $\eta \sim 0.4 - 0.6$. This can be improved using the relation $J(\eta)t_{\text{PT}} = \text{constant}$, which works better for $n = 2$ because it is exact for primary emissions with anti- k_t . But implementing it would not change the main conclusions presented in sections 4.3–4.5. Therefore, for the sake of simplicity, we will not use it here: we keep eq. (4.11) as the expression for δM_{PT} for the rest of this study.

Of course, were it only for the perturbative radiation, it would be nicer to choose $\eta \geq \eta_{\text{sat}}$ in order to catch as many gluons as possible, leading to $\delta M_{\text{PT}} \rightarrow \delta M_{\text{sat}}$. But we also have to take into account Initial State Radiation (ISR) from the incoming $q\bar{q}$ pair and non-perturbative effects like PU and UE that can spoil our Higgs neighbourhood, thus increasing the jet mass.

For the purpose of this article we will only add UE and PU to the Final-State Radiation (FSR) effect studied above. We will thus ignore ISR, this partially for a question of

simplicity of the analysis, but also because the results of work such as [18, 56] suggest that for LHC processes whose hard scales are few hundred GeV, the crucial interplay is that between FSR and UE/PU. This is evident in the preference for small R values in dijet mass reconstructions in those references, where ISR is not playing a major role. Similarly we believe that the optimal values of η that we will determine here will have limited impact from ISR, though we shall not check this explicitly.

4.2 Study of the Higgs width due to underlying event and pile-up

For this simple analysis, which does not aim to give precise numbers but only an estimate of the influence of the UE/PU on the mass of the Higgs jet when we vary for instance η , n , or when the Higgs boson becomes more and more boosted, we model the UE and PU as soft particles uniformly distributed in the (y, ϕ) plane [57, 58], and with transverse momentum per unit area denoted by ρ . In order to get this estimate, we consider the simple case of a symmetric ($z = 1/2$) Higgs decay along the x axis. In the limit $M_H \ll p_{t_H}$, the Higgs momentum p_H is given by:

$$p_H = \left(p_{t_H} + \frac{M_H^2}{2p_{t_H}}, p_{t_H}, 0, 0 \right). \tag{4.13}$$

The UE/PU momentum, denoted p_{UE} ,¹³ is simply the sum of all the UE/PU particles g belonging to the filtered jet J . Still in the limit $M_H \ll p_{t_H}$, we recall the following formula [15]:

$$R_{bb} \simeq \frac{1}{\sqrt{z(1-z)}} \frac{M_H}{p_{t_H}}. \tag{4.14}$$

Throughout this section, we will apply it with $z = 1/2$. We can now write $\Delta M = M_{\text{filtered jet}} - M_H$ as:

$$\begin{aligned} \Delta M &= \frac{1}{2M_H} \left((p_H + p_{UE})^2 - M_H^2 \right), \\ &\simeq \frac{1}{M_H} \sum_{g \in J} p_{t_g} p_{t_H} \left(\frac{\theta_{gH}^2}{2} + \frac{M_H^2}{2p_{t_H}^2} \right), \\ &\simeq \frac{M_H}{p_{t_H}} \sum_{g \in J} p_{t_g}. \end{aligned} \tag{4.15}$$

In the last line we used the approximation:

$$\theta_{gH} \sim \theta_{bH} = \frac{R_{bb}}{2}, \tag{4.16}$$

which comes from the fact that the UE and PU particles tend to cluster around the perturbative radiation, which is usually close to the b and \bar{b} because of the collinear logarithmic divergence of QCD. As all the filtered UE/PU particles flow approximately in the same

¹³For brevity, we define p_{UE} to be the sum of the UE and/or PU particles' momentum but without referencing the PU dependence, which will always be implicit.

direction, the remaining sum is just the total transverse momentum of the UE which, by definition of ρ , is equal to ρA , A being the total area of the filtered jets.¹⁴ We thus obtain

$$\Delta M \simeq \frac{\rho A M_H}{p_{t_H}}, \quad (4.17)$$

with

$$\langle A \rangle \simeq n\pi\eta^2 R_{\text{bb}}^2, \quad (4.18)$$

for the C/A jet algorithm, taking into account the anomalous dimension that comes from the fact that there should be some perturbative radiation in the jets (cf figure 14 in [57]). Notice that eq. (4.18) is only true if all the jets do not overlap, so usually when η is small enough. But this is sufficient for the purpose of our study, and we shall use this formula in all the following calculations. The correction eq. (4.17) for ΔM only induces a shift towards higher masses of the Higgs mass peak. However, there are 3 sources of fluctuations that give a width to this Higgs peak:

1. ρ is not strictly uniform in the (y, ϕ) plane in a given event.
2. ρ is not the same from one event to the next.
3. The jets' area fluctuates.

Following [57], we can write the total UE/PU transverse momentum contributing to the Higgs p_t as

$$p_{t_{\text{UE}}} = \rho A \pm \left(\sqrt{A}\sigma + A\delta\rho + \rho\Sigma \right), \quad (4.19)$$

where

$$\sigma = \sqrt{\langle \rho^2 \rangle - \langle \rho \rangle^2} \quad \text{with } \langle \dots \rangle \text{ a spatial average in a given event,} \quad (4.20)$$

$$\delta\rho = \sqrt{\langle \rho^2 \rangle - \langle \rho \rangle^2} \quad \text{with } \langle \dots \rangle \text{ an average over events,} \quad (4.21)$$

$$\Sigma = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} \quad \text{with } \langle A \rangle \text{ the average over events of the filtered jets' area.} \quad (4.22)$$

For pure UE events, i.e. without PU, these terms can be estimated [57, 58]:

$$\rho_{\text{UE}} \simeq 2 - 3 \text{ GeV/area,} \quad (4.23)$$

$$\sigma_{\text{UE}} \simeq 0.6\rho_{\text{UE}}, \quad (4.24)$$

$$\delta\rho_{\text{UE}} \simeq 0.8\rho_{\text{UE}}, \quad (4.25)$$

$$\Sigma \simeq 0.26\sqrt{n}\pi\eta^2 R_{\text{bb}}^2. \quad (4.26)$$

Though ρ_{UE} seems to be around 2 GeV/area, the tuning used in [15] was closer to 3 GeV/area, the value that we choose here. In presence of PU, i.e. when there is more than 1 pp collision per bunch crossing at the LHC (thus leading to the emission of other soft particles), ρ , σ and $\delta\rho$ have to be modified. We define N_{PU} to be the number of pp

¹⁴In the active sense, see [57].

collisions in a bunch crossing except the one at the origin of the hard interaction. We use a simple model to write the parameters of the UE/PU as:

$$\rho \simeq \left(1 + \frac{N_{\text{PU}}}{4}\right) \rho_{\text{UE}}, \quad (4.27)$$

$$\sigma \simeq \sqrt{1 + \frac{N_{\text{PU}}}{4}} \sigma_{\text{UE}}, \quad (4.28)$$

$$\delta\rho \simeq \sqrt{1 + \frac{N_{\text{PU}}}{4}} \delta\rho_{\text{UE}}. \quad (4.29)$$

Some comments are needed: since ρ measures the level of noise, it should grow like N_{PU} . In the expression $1 + N_{\text{PU}}/4$, the 1 corresponds to the pp collision that leads to the UE and to the hard interaction, whereas the $N_{\text{PU}}/4$ term simply corresponds to the other pp interactions and could be derived from the numbers given in [59]. The intra and inter events fluctuations of ρ are modelled as growing like $\sqrt{\rho}$: we thus just give σ and $\delta\rho$ the factor $\sqrt{1 + N_{\text{PU}}/4}$, though further studies might be of value to parametrize these terms in a more adequate manner. Notice that the value given for $\delta\rho$ ignores the fluctuations in the number of PU events from one bunch crossing to the next, but this is beyond the accuracy of our model here. At high luminosity at LHC, N_{PU} is expected to be ~ 20 , which implies $\rho \sim 10 - 20 \text{ GeV}$ [59–61].

Assuming gaussian distributions for these three kinds of fluctuations, one can deduce the Higgs width due to the presence of UE/PU,¹⁵ $\delta M_{\text{UE}} = 2\sqrt{\langle\Delta M^2\rangle - \langle\Delta M\rangle^2}$:

$$\delta M_{\text{UE}} = 2\sqrt{A\sigma^2 + A^2\delta\rho^2 + \rho^2\Sigma^2} \frac{M_H}{p_{tH}}. \quad (4.30)$$

For a gaussian peak, defining a 2σ width means that we keep roughly 68% of the events around the average, which is in correspondence with the value $f = 0.68$ chosen for the perturbative calculation.

We now have all the important results in hand to consider both UE/PU and FSR simultaneously.

4.3 Study of the Higgs width in presence of both UE/PU and perturbative radiation

The purpose of this part is to give an estimate of how one should choose the couple of filtering parameters (n, η) . For that, one has to convolute the effects of UE/PU and perturbative radiation and compute the resulting reconstructed Higgs peak width, and then minimize it with respect to the filtering parameters. This is highly non trivial to do analytically and we leave it for future work. The simple choice made here is to say that, for a given n , the optimal η , denoted η_{opt} , is the one for which the two widths are equal. This is obviously not true in general, but seems reasonable to obtain an estimate (figure 9) and to understand how η_{opt} changes when we vary p_{tH} and N_{PU} . Notice that, using this method,

¹⁵Here again, for brevity, we define δM_{UE} to be the Higgs width in presence of UE and/or PU without referencing the PU dependence. Actually it serves only to distinguish the width due to UE/PU from the perturbative width δM_{PT} .

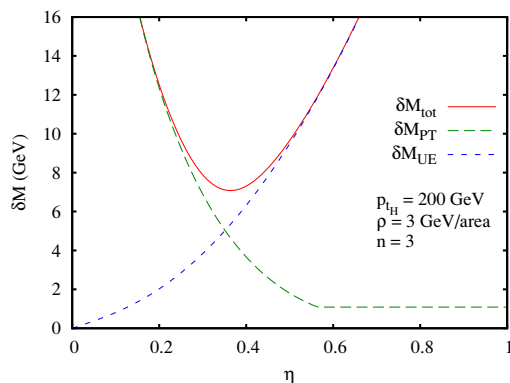


Figure 9. The Higgs width due to UE/PU and loss of perturbative radiation, combined as if the 2 distributions were gaussians, i.e. $\delta M_{\text{tot}} = \sqrt{\delta M_{\text{PT}}^2 + \delta M_{\text{UE}}^2}$ when $n = 3$. In this case, η_{opt} , though slightly larger, is approximately given by the intersection of the 2 curves, at least as long as η is not in the saturation region.

we have to impose $\eta_{\text{opt}} < \eta_{\text{sat}}$ where η_{sat} is the saturation point (eq. (4.12)), because beyond η_{sat} , increasing η makes δM_{UE} larger without decreasing δM_{PT} , thus solving the equation $\delta M_{\text{PT}} = \delta M_{\text{UE}}$ has no sense in this region. Finally, we numerically minimize $\sqrt{\delta M_{\text{PT}}^2 + \delta M_{\text{UE}}^2}$, calculated at $\eta = \eta_{\text{opt}}(n)$, with respect to n in order to find n_{opt} .

First, we would like to understand how η_{opt} evolves with respect to the physical parameters. The equality $\delta M_{\text{PT}} = \delta M_{\text{UE}}$ gives an equation in $L = \ln \frac{1}{\eta}$:

$$M_H e^{-\frac{1}{2\beta_0\alpha_s} \left(1 - e^{-4\pi\beta_0 \frac{c_{\text{PT}}}{L}}\right)} = 2\sqrt{c_\sigma^2 e^{-2L} + c_{\delta\rho}^2 e^{-4L} + c_\Sigma^2 e^{-4L}} \rho_{\text{UE}} \frac{M_H}{p_{t_H}}, \quad (4.31)$$

where the coefficients c_σ , $c_{\delta\rho}$ and c_Σ can be easily calculated using eqs. (4.18), (4.23)–(4.29), (4.30):

$$c_\sigma(n, N_{\text{PU}}, R_{\text{bb}}) \simeq 0.6\sqrt{\pi}\sqrt{n}R_{\text{bb}}\sqrt{1 + \frac{N_{\text{PU}}}{4}}, \quad (4.32)$$

$$c_{\delta\rho}(n, N_{\text{PU}}, R_{\text{bb}}) \simeq 0.8\pi n R_{\text{bb}}^2 \sqrt{1 + \frac{N_{\text{PU}}}{4}}, \quad (4.33)$$

$$c_\Sigma(n, N_{\text{PU}}, R_{\text{bb}}) \simeq 0.26\pi\sqrt{n}R_{\text{bb}}^2 \left(1 + \frac{N_{\text{PU}}}{4}\right). \quad (4.34)$$

If the solution of eq. (4.31) for a given n is found to be above $\eta_{\text{sat}}(n, f)$, then $\eta_{\text{opt}} = \eta_{\text{sat}}(n, f)$ in order to take the saturation of δM_{PT} into account. We start by solving this equation numerically. In figure 10 we show η_{opt} as a function of p_{t_H} and N_{PU} for different values of n . As it should, η_{opt} increases with p_{t_H} at fixed N_{PU} . Indeed, if p_{t_H} grows at fixed η , R_{bb} decreases and so does the effect of UE/PU, whereas the perturbative radiation is kept fixed (no dependence on R_{bb}). Notice also, for $n = 3$, that the values obtained for η_{opt} are roughly consistent with the choice in [15] where we had $\eta = \min(0.3/R_{\text{bb}}, 1/2)$. The saturation comes into effect at relatively low p_{t_H} , around 400 – 500 GeV. Above this value, the total width is small and hadronisation corrections start to become relevant, so that

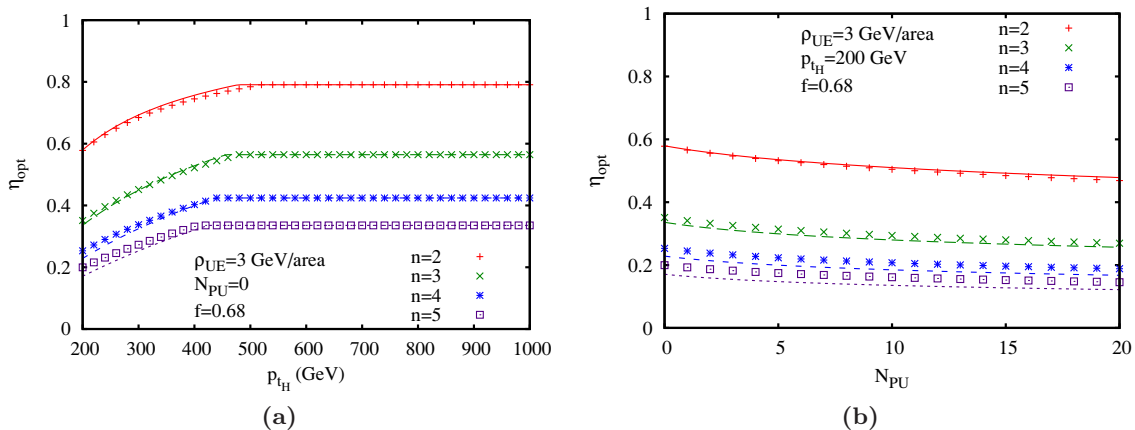


Figure 10. The numerical solutions (points) of eq. (4.31) shown for different values of n : (a) as a function of p_{t_H} when $N_{PU} = 0$, and (b) as a function of N_{PU} when $p_{t_H} = 200$ GeV. We also show the corresponding approximate analytical solutions (lines) derived in eq. (4.42).

the results presented on these plots become not very reliable. However, for $p_t > \sim 500$ GeV and $\eta > \eta_{\text{sat}}$, the Higgs width due to perturbative radiation and UE/PU vary slowly with η and we thus believe that the precise value chosen for η is not so important: one can take any value above η_{sat} without changing the result too much. The decrease of η_{opt} with N_{PU} seems to be weaker than one might have expected *a priori*. However, in figure 9, we can see that the negative slope of the perturbative width is very large, and therefore increasing the noise from PU will not change too much the η_{opt} value.

It would be interesting to understand analytically the evolution of η_{opt} with respect to the physical parameters p_{t_H} and N_{PU} . Unfortunately, eq. (4.31) cannot be easily dealt with. That's why we have to make an approximation: in this equation, one of the 3 terms under the square root may be dominant when $\eta = \eta_{\text{opt}}$. At first sight, one would expect that at low η_{opt} , the $c_\sigma^2 e^{-2L}$ term, which scales like η^2 , should be the largest, whereas at large N_{PU} , it should be the $c_\Sigma^2 e^{-4L}$ term that is the largest one as it scales like N_{PU}^2 . But figure 11 for $n = 3$ reveals that the $c_{\delta\rho}$ term surprisingly brings the largest contribution to δM_{UE} for physical values of the parameters (the same holds for other values of n). Therefore, to simplify things a little, one can consider eq. (4.31) and put $c_\sigma = c_\Sigma = 0$. However, to be more general, and to consider the possible situation where one of the other terms might be dominant,¹⁶ we rewrite eq. (4.31) in the following approximate form:

$$M_H e^{-\frac{1}{2\beta_0\alpha_s} \left(1 - e^{-4\pi\beta_0 \frac{C_{PT}}{L}}\right)} = C_{UE} \rho_{UE} e^{-pL} R_{bb}^p \frac{M_H}{p_{t_H}}, \quad (4.35)$$

¹⁶The subtraction procedure proposed in [59] seems to eliminate most of the fluctuations from the $c_{\delta\rho}$ and c_Σ terms, so that the remaining c_σ term would be dominant in this case.

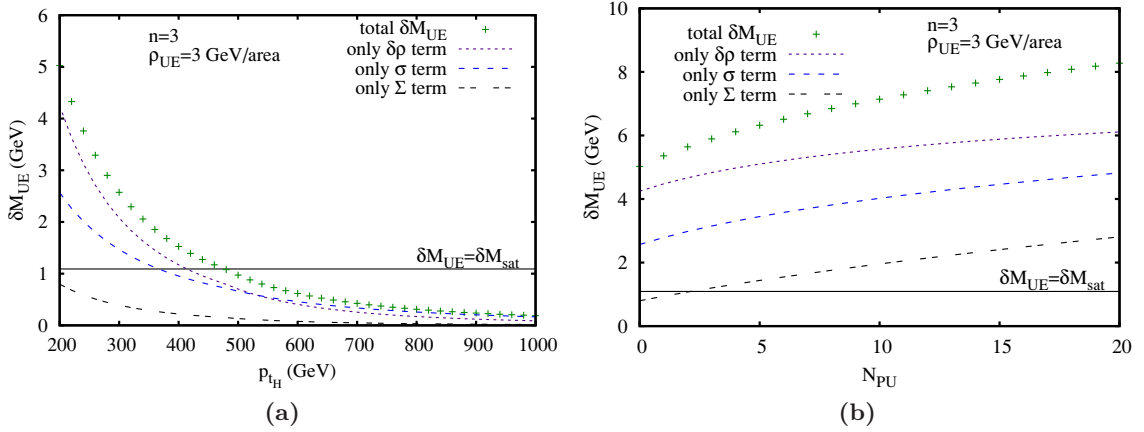


Figure 11. δM_{UE} computed at $\eta = \eta_{\text{opt}}$ with respect to (a) p_{tH} when $N_{\text{PU}} = 0$ and (b) N_{PU} when $p_{tH} = 200$ GeV. On these plots is also represented the contribution to δM_{UE} of each term separately. When the UE/PU width falls below the saturation line $\delta M_{\text{UE}} = \delta M_{\text{sat}}$, then $\eta_{\text{opt}} = \eta_{\text{sat}}$.

where $p = 1$ if the c_σ term dominates and $p = 2$ otherwise. Moreover:

$$C_{\text{UE}}(n, N_{\text{PU}}) = \begin{cases} 1.2\sqrt{\pi}\sqrt{n}\sqrt{1 + \frac{N_{\text{PU}}}{4}}, & \text{if the } c_\sigma \text{ term is dominant,} \\ 1.6\pi n\sqrt{1 + \frac{N_{\text{PU}}}{4}}, & \text{if the } c_{\delta\rho} \text{ term is dominant,} \\ 0.52\pi\sqrt{n}\left(1 + \frac{N_{\text{PU}}}{4}\right), & \text{if the } c_\Sigma \text{ term is dominant.} \end{cases} \quad (4.36)$$

Eq. (4.35) can be written in a slightly different way:

$$\frac{B_{\text{PT}}}{L} = \ln\left(\frac{1}{B_{\text{UE}} - 2\beta_0\alpha_s p L}\right), \quad (4.37)$$

with:

$$B_{\text{PT}} = 4\pi\beta_0 C_{\text{PT}}, \quad (4.38)$$

$$B_{\text{UE}} = 1 - 2\beta_0\alpha_s \ln\left(\frac{p_{tH}}{C_{\text{UE}}\rho_{\text{UE}}R_{\text{bb}}^p}\right). \quad (4.39)$$

Despite its simpler form, eq. (4.37) for L cannot be solved analytically. Here comes the second approximation, which is to make a perturbative expansion:

$$\frac{B_{\text{PT}}}{L} = \ln\frac{1}{B_{\text{UE}}} + \frac{2\beta_0\alpha_s p}{B_{\text{UE}}}L + \mathcal{O}((\alpha_s L)^2). \quad (4.40)$$

Neglecting the $\mathcal{O}((\alpha_s L)^2)$ term, the resulting quadratic equation immediately implies

$$L_{\text{opt}} = \frac{-B_{\text{UE}} \ln \frac{1}{B_{\text{UE}}} + \sqrt{B_{\text{UE}}^2 \ln^2 \frac{1}{B_{\text{UE}}} + 8\beta_0\alpha_s p B_{\text{UE}} B_{\text{PT}}}}{4\beta_0\alpha_s p}. \quad (4.41)$$

Taking into account the saturation effect, η_{opt} is then given by:

$$\eta_{\text{opt}} = \begin{cases} e^{-L_{\text{opt}}}, & \text{if } L_{\text{opt}} > -\ln \eta_{\text{sat}}, \\ \eta_{\text{sat}}, & \text{otherwise.} \end{cases} \quad (4.42)$$

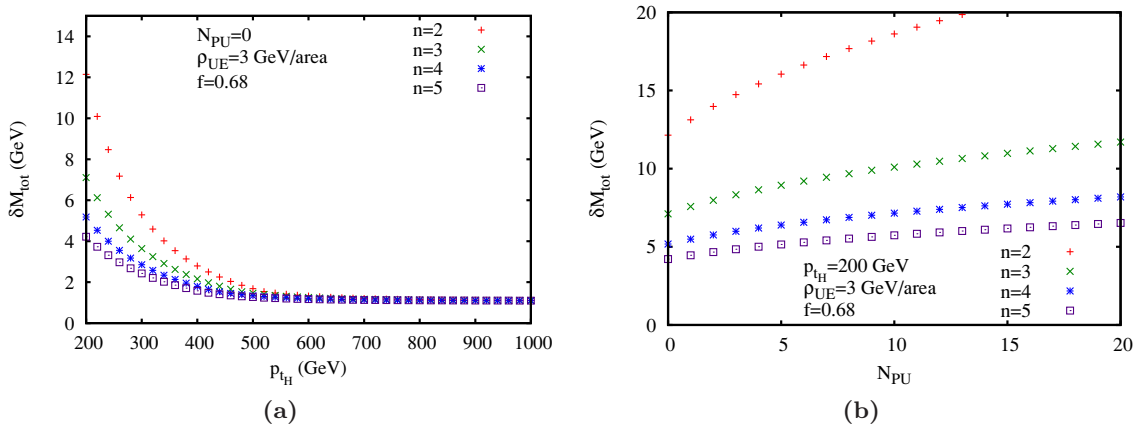


Figure 12. δM_{tot} computed at $\eta = \eta_{\text{opt}}$ as a function of (a) p_{t_H} and (b) N_{PU} for different values of n .

We used this expression with C_{UE} corresponding to the $\delta\rho$ term in eq. (4.36) and $p = 2$ in order to plot the approximate solutions in figure 10. This reveals that the above relation for η_{opt} (eq. (4.42)) works rather well, within a few %.

As a second step, we would like to find the optimal n , denoted n_{opt} . This also should depend on the way UE/PU and perturbative radiation are combined. However, as a simple approximation, one can combine them as if they were both gaussian distributions. Therefore, one should minimize

$$\delta M_{\text{tot}}(n) = \sqrt{\delta M_{\text{PT}}^2(n) + \delta M_{\text{UE}}^2(n)}, \tag{4.43}$$

computed at $\eta = \eta_{\text{opt}}(n)$ for a given p_{t_H} and N_{PU} .

The results are plotted in figure 12. We can notice that the larger n , the narrower the peak, and thus the better the result. However, one should keep in mind that when n increases, the optimal $R_{\text{filt}} = \eta R_{\text{bb}}$ becomes small, and we have to deal with hadronisation corrections that grow as $1/R_{\text{filt}}$ [56] as well as detector resolution and granularity $\delta\eta \times \delta\phi = 0.1 \times 0.1$ that both start to have an important impact on the reconstructed Higgs width, and thus degrade the results presented here. In section 4.5 we will examine what happens when we include a very rough estimate for hadronisation corrections. However, at first sight, it seems that one should definitely not take $n = 2$. The value $n = 3$ chosen in [15] is good, but it may be possible to do better with $n = 4$. Beyond this value, the optimal R_{filt} falls below ~ 0.2 (cf figure 10), which is too small for this study to be fully reliable, as we shall see in section 4.5.

4.4 Variations of the results with z and f

Until now, we have only presented some results for $f = 0.68$ and $z = 1/2$, z being defined as

$$z = \min\left(\frac{E_b}{E_H}, \frac{E_{\bar{b}}}{E_H}\right), \tag{4.44}$$

with E_i the energy of particle i in the Higgs splitting into $b\bar{b}$. What happens if we change these values?

Let us start with z . Though the Higgs splitting into $b\bar{b}$ is more often symmetric than in QCD events (and this is what was used in [15] to distinguish it from pure QCD splittings), it still has a distribution in z that is uniform in the range:

$$\frac{1}{2} \left(1 - \frac{1}{\sqrt{1 + \frac{M_H^2}{p_{tH}^2}}} \right) < z < \frac{1}{2}, \quad (4.45)$$

which follows from simple kinematics in the limit $m_b = 0$. But in order to reduce the large QCD background, one usually cuts on small z , so that

$$z_{\text{cut}} < z < \frac{1}{2}, \quad (4.46)$$

with $z_{\text{cut}} \sim 0.1$. As an example, assume the b quark carries the fraction z of the Higgs splitting. In such a case, b and \bar{b} are not equidistant from the Higgs direction: they are respectively at an angular distance $(1-z)R_{bb}$ and zR_{bb} from H (see for instance figure 16 in appendix A). Therefore, as UE/PU particles tend to cluster around the perturbative radiation, eq. (4.16) has to be modified:

$$\theta_{gH} \sim zR_{bb} \text{ or } \theta_{gH} \sim (1-z)R_{bb}, \quad (4.47)$$

for a given UE/PU particle g in the filtered jet. This leads to the modification of eq. (4.15) according to g being relatively close to b (region called “ J_1 ”) or \bar{b} (region called “ J_2 ”):

$$\begin{aligned} \Delta M &\simeq \frac{p_{tH}}{M_H} \left(\sum_{g \in J_1} p_{t_g} \left(\frac{(1-z)^2 R_{bb}^2}{2} + \frac{M_H^2}{2p_{tH}^2} \right) + \sum_{g \in J_2} p_{t_g} \left(\frac{z^2 R_{bb}^2}{2} + \frac{M_H^2}{2p_{tH}^2} \right) \right), \\ &\simeq \frac{M_H}{2p_{tH}} \left(\frac{1}{z} \sum_{g \in J_1} p_{t_g} + \frac{1}{1-z} \sum_{g \in J_2} p_{t_g} \right), \\ &= \frac{M_H}{2p_{tH}} \left(\frac{1}{z} \rho A_1 + \frac{1}{1-z} \rho A_2 \right). \end{aligned} \quad (4.48)$$

In this calculation we used eq. (4.14). To compute the dependence of the fluctuations on z , we take the simplest case $n = 2$. For the σ and Σ fluctuations, the terms ρA_1 and ρA_2 vary independently, leading to the following contribution to δM_{UE} :¹⁷

$$\delta M_{UE, \sigma, \Sigma}^2 = 4 \left(\frac{M_H \rho_{\text{UE}}}{2p_{tH}} \right)^2 \left(\frac{1}{z^2} \delta_{1, \sigma, \Sigma}^2 + \frac{1}{(1-z)^2} \delta_{2, \sigma, \Sigma}^2 \right), \quad (4.49)$$

where

$$\delta_{1, \sigma, \Sigma}^2 = \delta_{2, \sigma, \Sigma}^2 = c_\sigma^2 e^{-2L} + c_\Sigma^2 e^{-4L}, \quad (4.50)$$

¹⁷The factor of 4 comes from the fact that we compute the width at 2σ .

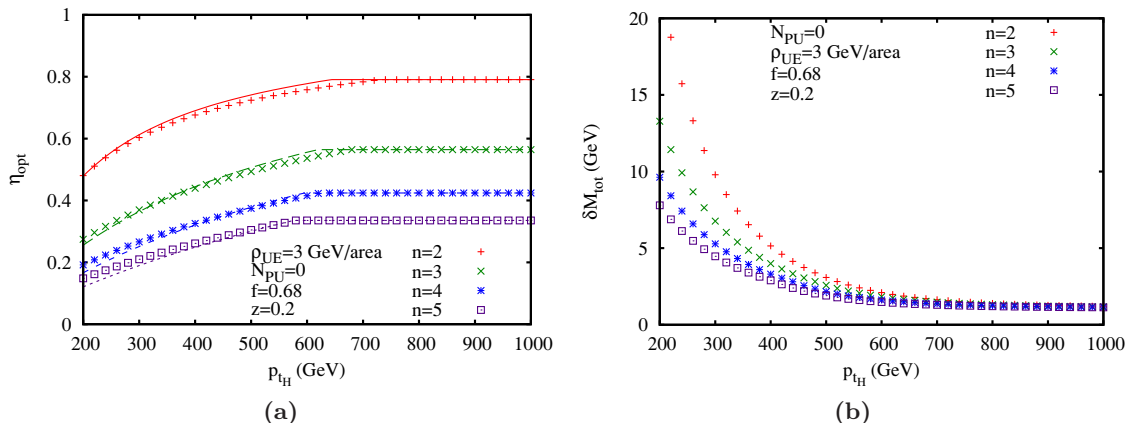


Figure 13. (a) η_{opt} as a function of p_{tH} when $f = 0.68$ and $z = 0.2$ for different values of n . The points correspond to the numerical determination of η_{opt} , found solving eq. (4.31) whose z dependence is derived in appendix B, whereas the curves correspond to its approximate analytical solutions. (b) $\delta M_{\text{tot}} = \sqrt{\delta M_{\text{PT}}^2 + \delta M_{\text{UE}}^2}$ computed at $\eta = \eta_{\text{opt}}$ as a function of p_{tH} for $f = 0.68$ and $z = 0.2$.

with c_σ and c_Σ given by eqs. (4.32), (4.34) for $n = 1$. Concerning the $\delta\rho$ fluctuations, the 2 terms ρA_1 and ρA_2 vary the same way from one event to another. Therefore, if it were only for the $\delta\rho$ term, we would write $\rho A_1 = \rho A_2$ leading to:

$$\delta M_{\text{UE},\delta\rho}^2 = 4 \left(\frac{M_H \rho_{\text{UE}}}{2p_{tH}} \right)^2 \frac{1}{z^2(1-z)^2} \delta_{\delta\rho}^2, \quad (4.51)$$

where

$$\delta_{\delta\rho}^2 = c_{\delta\rho}^2 e^{-4L}, \quad (4.52)$$

with $c_{\delta\rho}$ given by eq. (4.33) for $n = 1$. Adding all these contributions,

$$\delta M_{\text{UE}}^2 = \delta M_{\text{UE},\sigma,\Sigma}^2 + \delta M_{\text{UE},\delta\rho}^2, \quad (4.53)$$

this apparently leads to an enhancement of the width by a factor of $1/z$. But we have to take into account that the coefficients $c_{\delta\rho}$ and c_Σ also contain a factor R_{bb}^2 (eqs. (4.32), (4.33)) leading to another factor $1/z$, and thus an enhancement $1/z^2$ at small z .¹⁸ Therefore, we can conclude that the effect of $z \neq 1/2$ is to broaden the reconstructed Higgs peak. Such a factor may partly explain the width of ~ 14 GeV that was observed in [62], to be compared with the various widths found in the previous subsection (see for instance figure 12), and should also lead to decreasing η_{opt} . This is illustrated in figure 13, which was obtained with the results derived in appendix B, where we carry out the above analysis for a general n .

Now, we turn to the f value. As we explained in section 4.1, the choice $f = 0.68$ was made to correspond to a 2σ gaussian width, as we did for δM_{UE} , which is somewhat arbitrary. We would like to estimate how the results change when f is modified. We

¹⁸This is valid when $p_{tH} > \frac{1}{\sqrt{z(1-z)}} \frac{M_H}{R_0}$, with R_0 the radius of the initial clustering of the event, in order for the b and \bar{b} to be clustered together. For lower p_{tH} , there is a kinematic cut on z and the enhancement is less strong.

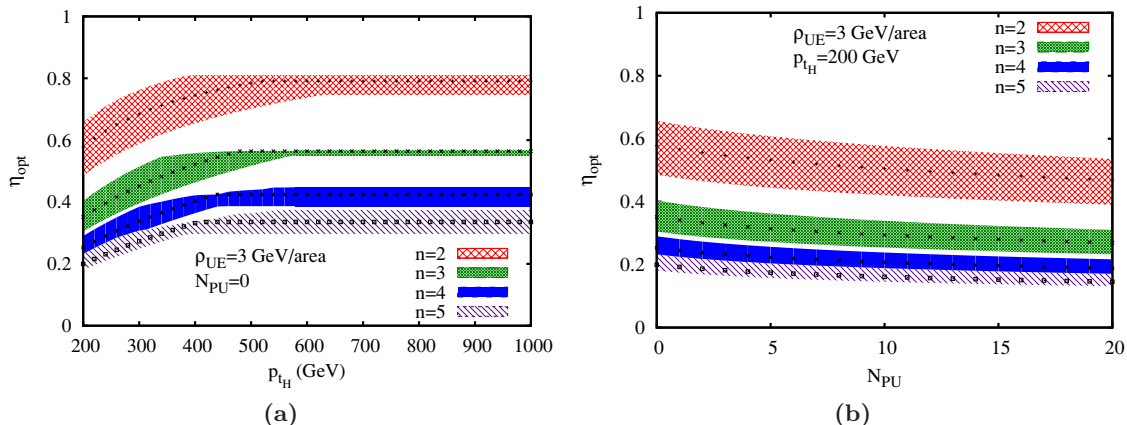


Figure 14. Uncertainty on η_{opt} when f varies from 0.5 to 0.8, for different values of n as a function of (a) p_{t_H} when $N_{\text{PU}} = 0$ and (b) N_{PU} when $p_{t_H} = 200 \text{ GeV}$. The results for $f = 0.68$ are also plotted as a reference.

thus also consider a range of values for f between 0.5 and 0.8. In this case the $C_{\text{PT}}(n, f)$ constants characterizing δM_{PT} are changed (see for instance eq. (4.9)), and δM_{UE} is also changed, i.e. eqs. (4.30), (4.31), (4.36) have to be slightly modified:

$$\delta M_{\text{UE}} = 2\sqrt{2} \operatorname{erf}^{-1}(f) \sqrt{A\sigma^2 + A^2\delta\rho^2 + \rho^2\Sigma^2} \frac{M_H}{p_{t_H}}, \quad (4.54)$$

$$= 2\sqrt{2} \operatorname{erf}^{-1}(f) \sqrt{c_\sigma^2 e^{-2L} + c_{\delta\rho}^2 e^{-4L} + c_\Sigma^2 e^{-4L}} \rho_{\text{UE}} \frac{M_H}{p_{t_H}}, \quad (4.55)$$

where $\operatorname{erf}(x)$ is the usual error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du. \quad (4.56)$$

Notice that the constants c_σ , $c_{\delta\rho}$ and c_Σ are left unchanged with this convention. However C_{UE} becomes:

$$C_{\text{UE}}(n, f, N_{\text{PU}}) = \begin{cases} 2\sqrt{2} \operatorname{erf}^{-1}(f) 0.6\sqrt{\pi}\sqrt{n}\sqrt{1 + \frac{N_{\text{PU}}}{4}}, & \text{if the } c_\sigma \text{ term is dominant,} \\ 2\sqrt{2} \operatorname{erf}^{-1}(f) 0.8\pi n \sqrt{1 + \frac{N_{\text{PU}}}{4}}, & \text{if the } c_{\delta\rho} \text{ term is dominant,} \\ 2\sqrt{2} \operatorname{erf}^{-1}(f) 0.26\pi\sqrt{n} \left(1 + \frac{N_{\text{PU}}}{4}\right), & \text{if the } c_\Sigma \text{ term is dominant.} \end{cases} \quad (4.57)$$

The bands corresponding to the uncertainties on η_{opt} that we obtain including these modifications are presented in figure 14. The uncertainty that we get, $\sim 20 - 30\%$, is not larger than the precision of the whole study of this paper, which limits itself to a large- N_c leading-log calculation. Notice that the variation with N_{PU} remains small.

One finally observes that $\eta_{\text{sat}}(n, f)$ is almost independent of f for $n = 3$. In appendix B, we will show that it can be approximately written as:

$$\eta_{\text{sat}} \simeq e^{-0.58} \left(1 + 0.044 \left(f - \frac{1}{2} \right) + \mathcal{O} \left(\left(f - \frac{1}{2} \right)^2 \right) \right). \quad (4.58)$$

Because of the small coefficient of its first order correction, $\eta_{\text{sat}} = e^{-0.58}$ is a good approximation within less than 1% for a large range of f values. But this seems to be a coincidence with no deep physical reason.

4.5 Hadronisation corrections

It is difficult to calculate what happens during the process of hadronisation, though some analytical results can be found concerning jet studies for instance [56, 63, 64]. In particular, it was shown in [56] that such non-perturbative corrections lead to a p_t shift for QCD jets equals on average $\sim -\Lambda/R_{\text{filt}}C_i$ where $\Lambda = 0.4 \text{ GeV}$ and $C_i = C_F$ or C_A depending on whether it is a quark jet or a gluon jet. This can be translated in our study by the following averaged p_t shift for the filtered jet:

$$\begin{aligned} \langle \delta p_t \rangle_{\text{had}} &= -(2C_F + (n-2)C_A) \frac{\Lambda}{R_{\text{filt}}}, \\ &\simeq -\frac{(n-1)N_c\Lambda}{R_{\text{filt}}}, \end{aligned} \tag{4.59}$$

where the second equality holds in the large N_c limit. Unfortunately, there is no result concerning the dispersion of the p_t distribution, which is the relevant quantity to compute in our case. Therefore, we are going to assume that the spread is of the same order of magnitude as the shift. This is in principle a crude approximation, but the only aim here is to illustrate the consequences of including hadronisation corrections in order to emphasize the fact that taking n too large is certainly not a good choice. Therefore, we use eq. (4.15) to estimate very roughly the hadronisation corrections to the reconstructed Higgs mass peak width:

$$\delta M_{\text{had}} \sim \frac{(n-1)N_c\Lambda}{R_{\text{filt}}} \frac{M_H}{p_{tH}} = \frac{(n-1)N_c\Lambda}{2\eta}, \tag{4.60}$$

when $z = 1/2$. As before, one should know how to combine perturbative radiation with UE/PU and hadronisation corrections in order to minimize the resulting combined width. However, we simply choose to minimize the quantity

$$\delta M_{\text{tot}} = \sqrt{\delta M_{\text{PT}}^2 + \delta M_{\text{UE}}^2 + \delta M_{\text{had}}^2}, \tag{4.61}$$

with respect to η and plot the resulting minimal δM_{tot} for different values of n (figure 15).

The first thing one can notice on these plots is that increasing n also increases the hadronisation corrections. For $n = 5$ they become so important that it is now clearly not an optimal filtering parameter contrary to what could be deduced from figure 12. The relevant p_t region in our study is roughly 200–400 GeV where we find the major part of the Higgs cross-section above 200 GeV and where our results are more reliable (see section 4.3). In this region $n = 3$ gives the best result. And at high PU, $n = 3$ and $n = 4$ both seem optimal, whereas $n = 2$ is far from being as good.

To conclude, our estimates seem to indicate within the accuracy of our calculations that $n = 2$ is not a good choice, nor is $n \geq 5$. Taking $n = 3$ or $n = 4$ gives equally good

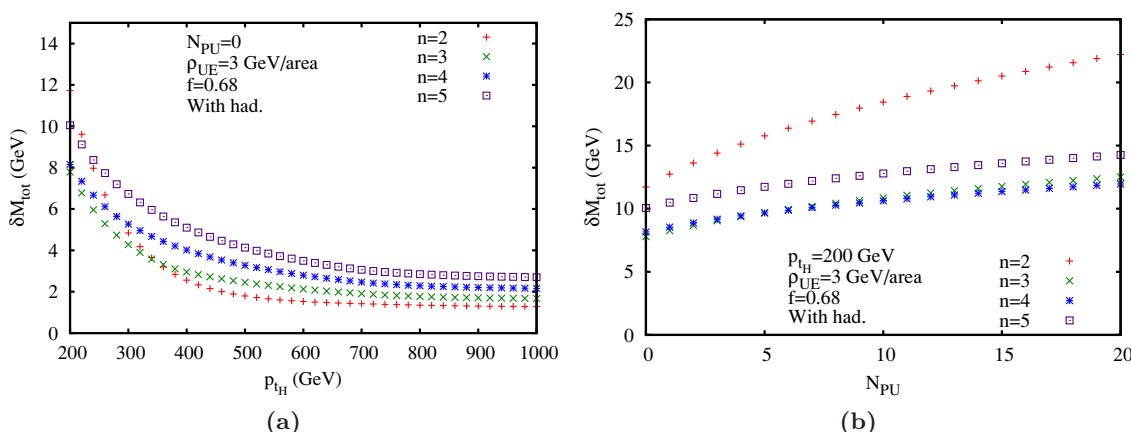


Figure 15. δM_{tot} including hadronisation corrections computed at $\eta = \eta_{\text{opt}}$ as a function of (a) p_{t_H} and (b) N_{PU} for different values of n .

quality to the mass peak. Increasing the hadronisation effects with respect to eq. (4.60), would lead to $n_{\text{opt}} = 3$, whereas if we lower them, we would find $n_{\text{opt}} = 4$. The only thing we can say is that $n = 3$ and $n = 4$ both seem to work rather well.

One way to go beyond these results would be to use event generators like Herwig [65, 66] or Pythia [67], to compute directly the Higgs width in presence of UE/PU, perturbative radiation, ISR and hadronisation, and to find for which value of the couple (n, η) the reconstructed Higgs mass peak width δM_H becomes minimal (that would still depend on p_{t_H} and the level of UE/PU). But our study here aimed to understand as much as possible the physical aspects behind such an optimisation, the price to pay being larger uncertainties on the result because of the necessary simplifications that were made.

5 Conclusions

This work has investigated the effect of QCD radiation on the reconstruction of hadronically decaying boosted heavy particles, motivated in part by the proposal of [15] to use a boosted search channel for the $H \rightarrow b\bar{b}$ decay. Though this article took the Higgs boson as an example, all the results presented here can be applied to the W and Z bosons, as well as any new colorless resonance decaying hadronically that might be observed at the LHC. The main effect of the QCD radiation is to distort and spread out the boosted heavy resonance shape well beyond the intrinsic width of the resonance. The aim therefore is to calculate the resulting resonance lineshape. This is a function of the parameters of the reconstruction method, notably of the “filtering” procedure, which aims to limit contamination from underlying event and pile-up, but which causes more perturbative radiation to be lost than would otherwise be the case.

Calculations were performed in a leading (single) logarithmic and leading colour approximation, which is the state of the art for this kind of problem. Analytic results were provided up to $\alpha_s^2 \ln^2 \frac{M_H}{\Delta M}$ for $n = 2$, and all-orders analytic results for the cases $n = 2$ and $n = 3$ were given for the terms that dominate in the small η limit. Numerical fixed-order

results up to $\alpha_s^5 \ln^5 \frac{M_H}{\Delta M}$ and all-orders resummed results were also given and are treated in more details in appendix C for a range of n and η . For the $n = 2$ and $n = 3$ cases there is quite acceptable agreement between the small- η analytic results and the full numerical results, even for values of $\eta \simeq 0.5$.

One unexpected feature that was observed was the behaviour of order-by-order expansion as compared to the resummed result: indeed there are indications that the series in $\alpha_s \ln \frac{M}{\Delta M}$ has a radius of convergence that is zero (but in a way that is unrelated to the renormalon divergence of the perturbative QCD series). This seems to be a general feature of the non-global logarithm series. Its practical impact seems to be greater for large η , or equivalently when the coefficients of the “primary” logarithms are small.

With these results in hand, it was then possible to examine how the perturbative width of the resonance peak depends on the parameters of the filtering. Though this was accessible only numerically for the full range of filtering parameters, figure 8a lends itself to a simple parametrisation for practically interesting parameter-ranges.

This parametrisation was then used in section 4.3 together with a parametrisation for the effect of UE and PU, so as to examine how to minimize the overall resonance width as a function of the filtering parameters and of the physical parameters of the problem such as the resonance p_t and the level of UE/PU. The approximations used might be described as overly simple, yet they do suggest interesting relations between optimal choices of the filtering parameters and the physical parameters of the problem. Though it is beyond the scope of this article to test these relations in full Monte Carlo simulation, we believe that investigation of their applicability in realistic conditions would be an interesting subject for future work. It should also be noticed that the methods used in this paper may be adapted to other reconstruction procedures like jet pruning [14] and jet trimming [19] as well as filtering as applied to jets without explicit substructure [18].

Acknowledgments

I am very grateful to Gavin Salam for suggesting this work and for helpful discussions while it was being carried out. I also wish to thank him as well as Sebastian Sapeta for comments on the manuscript. This work was supported in part by the French ANR under contract ANR-09-BLAN-0060.

A Analytical considerations on the non-global structure of the perturbative expansion

We derive here all the results presented in section 2. The Higgs boson is taken to move along the x direction and the angular coordinates (θ, ϕ) are defined with respect to the Higgs direction (so that $\theta = 0$ corresponds to the Higgs direction and $\phi = 0$ to the y axis for instance). As it is very boosted, the b and \bar{b} resulting from its decay should be close to the Higgs, i.e. with $\theta_{b,\bar{b}} \ll 1$. Due to angular ordering, so will be the major part of the perturbative radiation from $b\bar{b}$. To take this property into account, the angular coordinates

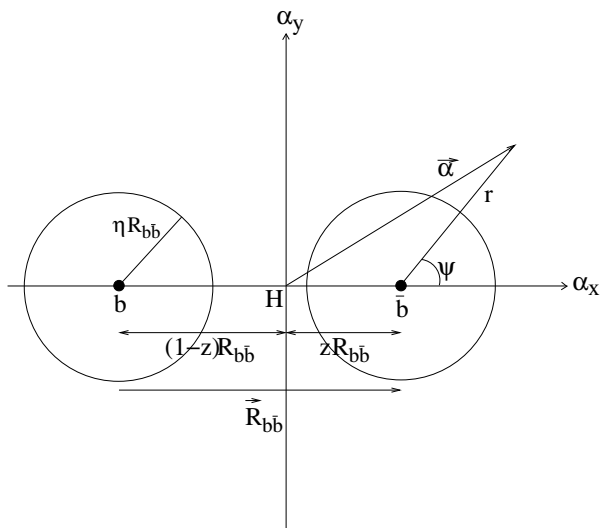


Figure 16. The $\vec{\alpha}$ plane, with various variables used in the calculation and all along this study. In this figure, the b quark is supposed to carry a fraction z of the Higgs energy, and the center of the frame coincides with the direction of the Higgs boson momentum.

(θ, ϕ) are slightly changed into a two-dimensional vector $\vec{\alpha}$:

$$\begin{aligned}\vec{\alpha} &= \theta(\cos \phi, \sin \phi), \\ d^2 \vec{\alpha} &= \theta d\theta d\phi.\end{aligned}\tag{A.1}$$

This vector is useful because of the property that, at small θ , one can express the angle θ_{ij} between 2 particles i and j as:

$$\theta_{ij} = |\vec{\alpha}_i - \vec{\alpha}_j|.\tag{A.2}$$

$\vec{R}_{b\bar{b}}$ is defined to be the vector from b to \bar{b} in the $\vec{\alpha}$ plane (see figure 16). Even if $|\vec{\alpha}| < \mathcal{O}(1)$, the integrations over α_x and α_y are extended to ∞ . By doing so, the error made on a $\mathcal{O}(1)$ result is of order $\mathcal{O}(R_{bb}^2)$.

A.1 Primary coefficients

Let us go back to the integral of eq. (2.9) and write it in the $\vec{\alpha}$ plane:

$$\begin{aligned}I(\Delta M) &= - \int_{\vec{k} \notin J_{b\bar{b}}} \frac{d^3 \vec{k}}{(2\pi)^3 2|\vec{k}|} |M(k)|^2 \Theta(\Delta M(k) - \Delta M), \\ &= - \frac{\alpha_s C_F}{2\pi^2} \int_0^{p_{tH}} \frac{dk_t}{k_t} \int_{\vec{\alpha} \notin J_{b\bar{b}}} d^2 \vec{\alpha} \frac{2R_{bb}^2}{|\vec{\alpha} - z\vec{R}_{bb}|^2 |\vec{\alpha} + (1-z)\vec{R}_{bb}|^2} \Theta(\Delta M(k) - \Delta M).\end{aligned}\tag{A.3}$$

For $\Delta M(k) = M_H - M$ one easily finds

$$\Delta M(k) = k_t \frac{M_H}{p_{tH}} A(\vec{\alpha}),\tag{A.4}$$

with

$$A(\vec{\alpha}) = \frac{|\vec{\alpha}|^2 p_{tH}^2}{2M_H^2} + \frac{1}{2}, \tag{A.5}$$

$$= \frac{|\vec{\alpha}|^2}{2z(1-z)R_{bb}^2} + \frac{1}{2}, \tag{A.6}$$

where we used eq. (4.14). With this expression, the integration over k_t is straightforward:

$$\begin{aligned} \int_0^{p_{tH}} \frac{dk_t}{k_t} \Theta(\Delta M(k) - \Delta M) &= \int_{\frac{p_{tH} \Delta M}{M_H A(\vec{\alpha})}}^{p_{tH}} \frac{dk_t}{k_t}, \\ &= \ln \frac{M_H}{\Delta M} + \ln A(\vec{\alpha}). \end{aligned} \tag{A.7}$$

Notice that $|\vec{\alpha}| \sim (1-z)R_{bb}$ or $|\vec{\alpha}| \sim zR_{bb}$, depending on whether the perturbative gluon emission is relatively close to b or \bar{b} (due to the collinear divergence of QCD). Thus, given that a Higgs splitting is most of the time roughly symmetric ($z \sim 1/2$), this leads to $A(\vec{\alpha}) = \mathcal{O}(1)$. Therefore:

$$\ln A(\vec{\alpha}) \ll \ln \frac{M_H}{\Delta M}, \tag{A.8}$$

and the $\ln A(\vec{\alpha})$ term can be neglected in a leading-log calculation. One thus obtains:

$$I(\Delta M) = -\frac{\alpha_s C_F}{\pi} \ln \left(\frac{M_H}{\Delta M} \right) J(\eta), \tag{A.9}$$

with $J(\eta)$ the remaining angular integral. Introducing the coordinates (r, ψ) defined in figure 16, $J(\eta)$ can be rewritten in the following form:¹⁹

$$\begin{aligned} J(\eta) = J_0(\eta) &= \frac{1}{\pi} \int_{\eta R_{bb}}^{+\infty} \frac{dr}{r} \int_{-\arccos(-\frac{R_{bb}}{2r})}^{\arccos(-\frac{R_{bb}}{2r})} d\psi \frac{2R_{bb}^2}{r^2 + 2rR_{bb} \cos \psi + R_{bb}^2} \quad \text{if } \frac{1}{2} < \eta < 1, \\ &= J_0\left(\frac{1}{2}\right) + \frac{1}{\pi} \int_{\eta R_{bb}}^{\frac{R_{bb}}{2}} \frac{dr}{r} \int_{-\pi}^{\pi} d\psi \frac{2R_{bb}^2}{r^2 + 2rR_{bb} \cos \psi + R_{bb}^2} \quad \text{if } \eta < \frac{1}{2}. \end{aligned} \tag{A.10}$$

Performing the ψ integration and the r one (for $\eta < 1/2$), one arrives at the formulae (2.11), (2.12), where all the R_{bb} dependence is cancelled.

One remark: if $\eta > 1$, the b and \bar{b} quarks cluster together, and the result then depends on the z fraction of the splitting. For instance, if $\eta > 2$, $J(\eta)$ can now be written:²⁰

$$J(\eta, z) = \ln \left(\frac{(\eta^2 + z(1-z))^2}{(\eta^2 - z^2)(\eta^2 - (1-z)^2)} \right). \tag{A.11}$$

¹⁹A simple shift of the $\vec{\alpha}$ coordinates gets rid of the z dependence, for instance $\vec{\alpha}' = \vec{\alpha} - (z - \frac{1}{2}) \vec{R}_{b\bar{b}}$

²⁰The intermediate case $1 < \eta < 2$ has a more complicated phase space integration and is not presented here.

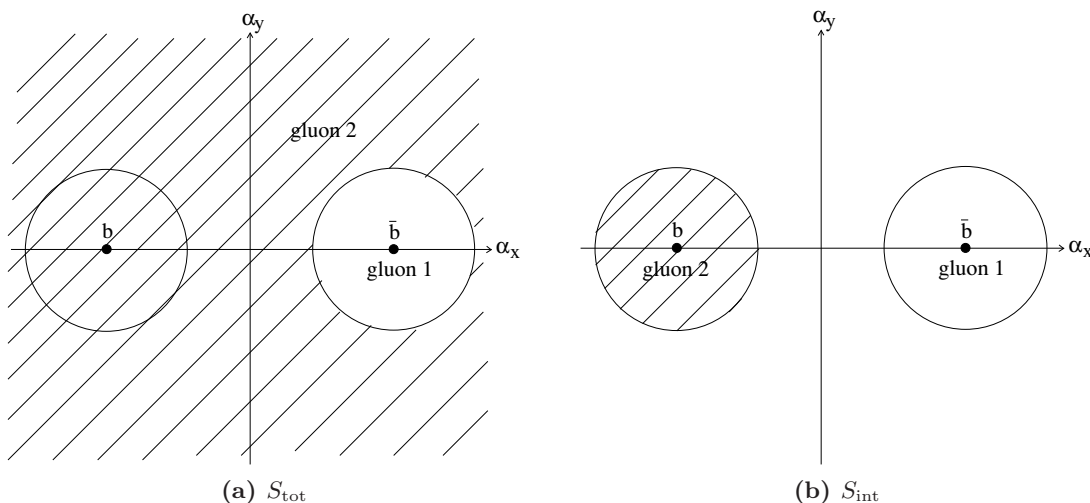


Figure 17. Regions of integration (dashed region) for gluon 2 when computing S_{tot} (left) and S_{int} (right).

A.2 Non-global coefficients

The starting point here is eq. (2.20) for I_2^{NG} :

$$I_2^{\text{NG}}(\Delta M) = - \int dk_1 dk_2 (4\pi\alpha_s)^2 \Theta(E_1 - E_2) \Theta(k_1 \in J_{b\bar{b}}) \Theta(k_2 \notin J_{b\bar{b}}) W_2 \Theta(\Delta M(k_2) - \Delta M).$$

In the same way as the primary case, using the $\vec{\alpha}$ plane and integrating over the energies of the 2 gluons, one arrives at:

$$I_2^{\text{NG}}(\Delta M) = \frac{1}{2} C_F C_A \left(\frac{\alpha_s}{\pi} \log \left(\frac{M_H}{\Delta M} \right) \right)^2 S_2(\eta),$$

with $S_2(\eta) = -\frac{R_{b\bar{b}}^2}{2\pi^2} \int_{\vec{\alpha}_1 \in J_{b\bar{b}}} d^2 \vec{\alpha}_1 \int_{\vec{\alpha}_2 \notin J_{b\bar{b}}} d^2 \vec{\alpha}_2$

$$\times \left(-\frac{R_{b\bar{b}}^2}{|\vec{\alpha}_1 - z\vec{R}_{b\bar{b}}|^2 |\vec{\alpha}_1 + (1-z)\vec{R}_{b\bar{b}}|^2 |\vec{\alpha}_2 - z\vec{R}_{b\bar{b}}|^2 |\vec{\alpha}_2 + (1-z)\vec{R}_{b\bar{b}}|^2} + \frac{1}{|\vec{\alpha}_1 + (1-z)\vec{R}_{b\bar{b}}|^2 |\vec{\alpha}_1 - \vec{\alpha}_2|^2 |\vec{\alpha}_2 - z\vec{R}_{b\bar{b}}|^2} + \frac{1}{|\vec{\alpha}_1 - z\vec{R}_{b\bar{b}}|^2 |\vec{\alpha}_1 - \vec{\alpha}_2|^2 |\vec{\alpha}_2 + (1-z)\vec{R}_{b\bar{b}}|^2} \right). \quad (\text{A.12})$$

In all this part, $\eta < 1/2$ is assumed. To deal with this integral, the frame is centered around \bar{b} for instance, and 2 quantities are computed: S_{tot} where gluon 1 is in the jet region $J_{\bar{b}}$ around \bar{b} and gluon 2 covers the whole phase space except $J_{\bar{b}}$, from which is subtracted S_{int} where gluon 2 covers J_b , the jet region around b (figure 17). Therefore:

$$S_2(\eta) = 2(S_{\text{tot}}(\eta) - S_{\text{int}}(\eta)), \quad (\text{A.13})$$

where the factor 2 is for the symmetric case (gluon 1 in J_b).

A.2.1 Calculation of S_{tot}

Using the variables $(u = \frac{r}{\eta R_{\text{bb}}}, \psi)$ (figure 16), S_{tot} can be written

$$\begin{aligned}
 S_{\text{tot}}(\eta) = & -\frac{1}{2\pi^2} \int_0^1 u_1 du_1 \int_1^{+\infty} u_2 du_2 \int_0^{2\pi} d\psi_1 \int_0^{2\pi} d\psi_2 \\
 & \times \left(-\frac{1}{u_1^2 u_2^2 (1 + 2\eta u_1 \cos \psi_1 + \eta^2 u_1^2)(1 + 2\eta u_2 \cos \psi_2 + \eta^2 u_2^2)} \right. \\
 & + \frac{1}{u_1^2 (1 + 2\eta u_2 \cos \psi_2 + \eta^2 u_2^2) (u_1^2 - 2u_1 u_2 \cos(\psi_1 - \psi_2) + u_2^2)} \\
 & \left. + \frac{1}{u_2^2 (1 + 2\eta u_1 \cos \psi_1 + \eta^2 u_1^2) (u_1^2 - 2u_1 u_2 \cos(\psi_1 - \psi_2) + u_2^2)} \right). \quad (\text{A.14})
 \end{aligned}$$

Doing the angular integrations, one arrives at:

$$S_{\text{tot}}(\eta) = -4 \int_0^1 du_1 \frac{u_1}{1 - \eta^2 u_1^2} \int_1^{\frac{1}{\eta}} \frac{du_2}{u_2} \frac{1}{u_2^2 - u_1^2}. \quad (\text{A.15})$$

2 remarks about this result:

1. The first and second terms have divergences in $u_1 = 0$ and $u_2 = \frac{1}{\eta}$, so respectively when gluon 1 is collinear to \bar{b} and gluon 2 is collinear to b , but they cancel when adding these terms.
2. The part of the integral corresponding to $u_2 > \frac{1}{\eta}$, i.e. $r_2 > R_{\text{bb}}$, is null, which can be simply interpreted as a manifestation of the angular ordering.

Performing this integration with Maple for instance gives the following result:

$$\begin{aligned}
 S_{\text{tot}}(\eta) = & \frac{\pi^2}{2} - \frac{3}{2} \ln^2(1 - \eta) + \ln(2\eta) \ln(1 - \eta) - \ln 2 \ln(1 + \eta) - \\
 & - 2\Re \left(\text{dilog} \left(\frac{\eta}{\eta - 1} \right) \right) - 2\text{Li}_2(1 - \eta) - \text{Li}_2 \left(\frac{1}{1 + \eta} \right) \\
 & + \text{Li}_2(-\eta) + \text{Li}_2 \left(\frac{1 - \eta}{1 + \eta} \right) - \text{Li}_2 \left(\frac{-2\eta}{1 - \eta} \right), \quad (\text{A.16})
 \end{aligned}$$

where

$$\text{dilog}(x) = \int_1^x dt \frac{\ln t}{1 - t}, \quad (\text{A.17})$$

$$\text{Li}_2(x) = \int_x^0 dt \frac{\ln(1 - t)}{t}. \quad (\text{A.18})$$

This is rather complicated, but that expression can be greatly simplified using the relations:

$$\forall x > 0, \quad \text{dilog}(x) = \text{Li}_2(1 - x), \quad (\text{A.19})$$

$$\text{dilog}(-x) = \frac{\pi^2}{3} - \frac{1}{2} \ln^2(1 + x) - \text{Li}_2 \left(\frac{1}{1 + x} \right) - i\pi \ln(1 + x), \quad (\text{A.20})$$

$$\forall x \text{ with } 0 < x < 1, \quad \text{Li}_2(x) + \text{Li}_2(1-x) = \frac{\pi^2}{6} - \ln(x) \ln(1-x), \quad (\text{A.21})$$

$$\text{Li}_2(1-x) + \text{Li}_2(1-1/x) = -\frac{1}{2} \ln^2 x. \quad (\text{A.22})$$

The final answer is then:

$$S_{\text{tot}}(\eta) = -\frac{\pi^2}{6}. \quad (\text{A.23})$$

The remarkable point to notice is of course that S_{tot} does not depend on η . But no simple explanation was found to interpret this result.

A.2.2 Calculation of S_{int}

S_{int} must now be subtracted from S_{tot} . The computation is similar to the previous one and is not detailed here. However, contrary to S_{tot} , a simple analytical result was not obtained, only the following expansion:

$$\begin{aligned} S_{\text{int}}(\eta) &= -4 \int_0^1 \frac{du_1}{u_1} \int_0^1 \frac{du_2}{u_2} \left(\frac{1}{\sqrt{(1-\eta^2(u_1^2+u_2^2))^2 - 4\eta^4 u_1^2 u_2^2}} - \frac{1}{1-\eta^2(u_1^2+u_2^2)} \right), \quad (\text{A.24}) \\ &= -2\eta^4 - 6\eta^6 - \frac{31}{2}\eta^8 - 40\eta^{10} - \frac{1921}{18}\eta^{12} - \frac{889}{3}\eta^{14} - \frac{20589}{24}\eta^{16} - \frac{7643}{3}\eta^{18} + \mathcal{O}(\eta^{20}). \end{aligned} \quad (\text{A.25})$$

Using eq. (A.13) with the 2 previous results, one arrives at the expression eq. (2.24).

B Analytical considerations on the dependence of the results on z and f

B.1 Dependence on z

It is interesting to understand analytically how η_{opt} evolves when the decay of the Higgs boson occurs with a z fraction different from $1/2$. In fact, we obtain the same result as eqs. (4.41), (4.42) up to a modification of the constant C_{UE} (more generally, we have to modify the coefficients c_σ , $c_{\delta\rho}$ and c_Σ , see eqs. (4.31)–(4.34)). The starting point is eq. (4.48) generalized to any value of n :

$$\Delta M \simeq \frac{M_H}{2p_{tH}} \sum_{i=1}^n a_i(z) \rho A_i, \quad (\text{B.1})$$

where $a_i(z)$ is either $\frac{1}{z}$ or $\frac{1}{1-z}$ depending on whether subject i is in the J_1 region (around the b quark) or in the J_2 region (around the \bar{b} quark). This is because the UE/PU particles tend to cluster around the perturbative radiation, which itself is emitted close to b and \bar{b} . We call a “configuration” the set of all the coefficients $a_i(z)$.

The result on the fluctuations depends on which ones are considered. Let us start with the fluctuations originating from the σ and Σ terms. In this case the ρA_i terms vary independently, thus leading to a contribution to δM_{UE} similar to that of eq. (4.49) for $n = 2$:

$$\delta M_{\text{UE},\sigma,\Sigma,\{a_i\}}^2 = 4 \left(\frac{M_H \rho_{\text{UE}}}{2p_{tH}} \right)^2 \sum_{i=1}^n a_i(z)^2 \delta_{i,\sigma,\Sigma}^2, \quad (\text{B.2})$$

with (see eq. (4.50)):

$$\begin{aligned}\delta_{i,\sigma,\Sigma}^2 &= c_\sigma^2 e^{-2L} + c_\Sigma^2 e^{-4L}, \\ &\equiv \delta_{\sigma,\Sigma}^2.\end{aligned}\tag{B.3}$$

The coefficients c_σ and c_Σ are still computed for $n = 1$ in this formula. But eq. (B.2) is only valid for a given configuration $\{a_i\}$. We thus have to average over all the 2^{n-2} possible configurations (the b and \bar{b} subjects are fixed). As the perturbative radiation pattern does not depend on z (for z not too small), each configuration arises with the same probability. Therefore, if k is the number of subjects in the J_1 region and $n - 2 - k$ the number of subjects in the J_2 region apart from the b and \bar{b} subjects, we obtain:

$$\delta M_{UE,\sigma,\Sigma}^2 = 4 \left(\frac{M_H \rho_{UE}}{2p_{t_H}} \right)^2 \frac{1}{2^{n-2}} \sum_{k=0}^{n-2} \binom{n-2}{k} \left(\frac{k+1}{z^2} + \frac{n-1-k}{(1-z)^2} \right) \delta_{\sigma,\Sigma}^2,\tag{B.4}$$

$$= 4 \left(\frac{M_H \rho_{UE}}{p_{t_H}} \right)^2 \frac{n}{8} \frac{1-2z(1-z)}{z^2(1-z)^2} \delta_{\sigma,\Sigma}^2.\tag{B.5}$$

We can follow a similar reasoning for the $\delta\rho$ fluctuations, except that the ρA_i terms in eq. (B.1) vary the same way from one event to the next. Therefore, for a given configuration $\{a_i\}$, one can deduce the following contribution to δM_{UE} :

$$\delta M_{UE,\delta\rho,\{a_i\}}^2 = 4 \left(\frac{M_H \rho_{UE}}{2p_{t_H}} \right)^2 \left(\sum_{i=1}^n a_i(z) \right)^2 \delta_{\delta\rho}^2,\tag{B.6}$$

with $\delta_{\delta\rho}$ given by eq. (4.52). As before, we have to average this result over all the 2^{n-2} possible configurations, leading to:

$$\delta M_{UE,\delta\rho}^2 = 4 \left(\frac{M_H \rho_{UE}}{2p_{t_H}} \right)^2 \frac{1}{2^{n-2}} \sum_{k=0}^{n-2} \binom{n-2}{k} \left(\frac{k+1}{z} + \frac{n-1-k}{1-z} \right)^2 \delta_{\delta\rho}^2,\tag{B.7}$$

$$= 4 \left(\frac{M_H \rho_{UE}}{p_{t_H}} \right)^2 \frac{n^2 + (n-2)(1-2z)^2}{16z^2(1-z)^2} \delta_{\delta\rho}^2.\tag{B.8}$$

One can absorb all the dependence of the resulting δM_{UE}^2 in n and z into the coefficients c_σ , $c_{\delta\rho}$ and c_Σ , and define new coefficients c' such that

$$\delta M_{UE} = 2 \sqrt{c_\sigma'^2 \eta^2 + c_{\delta\rho}'^2 \eta^4 + c_\Sigma'^2 \eta^4} \frac{M_H \rho_{UE}}{p_{t_H}},\tag{B.9}$$

with

$$c_\sigma'(n, N_{PU}, R_{bb}, z) \simeq 0.6 \sqrt{\pi} \sqrt{n} \frac{\sqrt{1-2z(1-z)}}{2\sqrt{2}z(1-z)} R_{bb} \sqrt{1 + \frac{N_{PU}}{4}},\tag{B.10}$$

$$c_{\delta\rho}'(n, N_{PU}, R_{bb}, z) \simeq 0.8 \pi \frac{\sqrt{n^2 + (n-2)(1-2z)^2}}{4z(1-z)} R_{bb}^2 \sqrt{1 + \frac{N_{PU}}{4}},\tag{B.11}$$

$$c_\Sigma'(n, N_{PU}, R_{bb}, z) \simeq 0.26 \pi \sqrt{n} \frac{\sqrt{1-2z(1-z)}}{2\sqrt{2}z(1-z)} R_{bb}^2 \left(1 + \frac{N_{PU}}{4} \right).\tag{B.12}$$

With these results in hand, we can easily generalize eq. (4.41) for any value z of the Higgs splitting. One only has to modify the value of C_{UE} . For instance, the curves in figure 13a were obtained using eqs. (4.38)–(4.42) with:

$$C_{\text{UE}}(n, N_{\text{PU}}, z) = 1.6\pi \frac{\sqrt{n^2 + (n-2)(1-2z)^2}}{4z(1-z)} \sqrt{1 + \frac{N_{\text{PU}}}{4}}, \quad (\text{B.13})$$

where we just use the dominant $c'_{\delta\rho}$ term.

There is also another source of fluctuations that we haven't accounted for so far. Indeed, even if the values of ρ and of the jets area were constant, we should consider the fact that the filtered subjects can be either in the J_1 or in the J_2 region. The calculation of its effect is similar to that of the fluctuations in σ , $\delta\rho$ and Σ . Its contribution on δM_{UE}^2 can be written:

$$\delta M_{\text{UE,other}}^2 = 4 \left(\frac{M_H \rho_{\text{UE}}}{p_{t_H}} \right)^2 c_{\text{other}}'^2 \eta^4, \quad (\text{B.14})$$

with

$$c_{\text{other}}'(n, N_{\text{PU}}, R_{\text{bb}}, z) = \pi \frac{\sqrt{n-2}|1-2z|}{4z(1-z)} R_{\text{bb}}^2 \left(1 + \frac{N_{\text{PU}}}{4} \right). \quad (\text{B.15})$$

But we checked that its effect on δM_{UE}^2 is negligible compared to that from the dominant $\delta\rho$ fluctuation,²¹ and therefore we did not include it.

B.2 Comments on the uncertainty due to the choice of f

Let us return to the observation of section 4.4 that $\eta_{\text{sat}}(n, f)$ is almost independent of f for $n = 3$. To understand why, we have to estimate $t_{\text{sat}}(f)$ and $C_{\text{PT}}(3, f)$ (cf eq. (4.12)). t_{sat} can be deduced from the equation:

$$\Sigma^{(n)}(\eta = 1, t_{\text{sat}}) = f, \quad (\text{B.16})$$

for any value of n , as t_{sat} only depends on f (see for instance figures (7), (8a)). To be simple, we can use the function $\Sigma^{(2)}$ which was widely studied in this paper. Unfortunately we cannot take the primary emission result eq. (2.10) because the non-global part becomes important when $\eta = 1$. However, as an approximation, we can numerically compute the second order coefficient $a_2 \simeq -3$ of $\Sigma^{(2)}(\eta = 1, t)$ and solve:

$$1 - J(1)N_c t_{\text{sat}} + a_2 t_{\text{sat}}^2 \simeq f, \quad (\text{B.17})$$

which simply leads to

$$t_{\text{sat}}(f) \simeq \frac{J(1)N_c - \sqrt{J(1)^2 N_c^2 - 4a_2(1-f)}}{2a_2}. \quad (\text{B.18})$$

²¹This effect is strictly null when $z = 1/2$. When $z = 0.2$, we obtain:

$$\delta M_{\text{UE},\delta\rho}^2 \simeq 16.6 \delta M_{\text{UE,other}}^2,$$

for $p_{t_H} = 200 \text{ GeV}$ and $N_{\text{PU}} = 0$.

This expression can be shown numerically to give $t_{\text{sat}}(f)$ with a precision better than 1% for $f \in [0.5, 0.8]$. $C_{\text{PT}}(3, f)$ is harder to evaluate. One must solve

$$\Sigma^{(3)}(L, t) = f, \tag{B.19}$$

in the limit of large L . Using eq. (2.27), which is valid in this limit, and defining the function h such that $h(4N_c L t) = \Sigma^{(3)}(L, t) - f$, we can use Newton's method with only one iteration on h , evaluated numerically for $4N_c L t = 1$,²² to solve approximately eq. (B.19):

$$\begin{aligned} 4N_c L t &\simeq 1 - \frac{h(1)}{h'(1)}, \\ &\simeq 2.543 - 2.334f, \\ &\equiv 4N_c C_{\text{PT}}(3, f), \end{aligned} \tag{B.20}$$

which gives $C_{\text{PT}}(3, f)$ within a few %. Therefore, eq. (4.12) leads to

$$\begin{aligned} \eta_{\text{sat}}(3, f) &= e^{-\frac{C_{\text{PT}}(3, f)}{t_{\text{sat}}}}, \\ &\simeq e^{-0.58} \left(1 + 0.044 \left(f - \frac{1}{2} \right) + \mathcal{O} \left(\left(f - \frac{1}{2} \right)^2 \right) \right), \end{aligned} \tag{B.21}$$

which is eq. (4.58).

To be complete, the same analysis can be done when $n = 2$, using eq. (4.9) for $C_{\text{PT}}(2, f)$, to obtain $\eta_{\text{sat}}(2, f)$. We will just mention that

$$\lim_{f \rightarrow 1} \eta_{\text{sat}}(2, f) = e^{-\frac{J(1)}{4}} \simeq 0.85. \tag{B.22}$$

C Convergence of the non-global series

In this appendix, we go back to the study of the convergence of the non-global series that was already briefly examined in section 3.2, trying to understand a little more what may be behind the observed behaviour.

C.1 Case $n_{\text{fit}} = 2$

First, we start by studying the convergence of the non-global perturbative series when $n = 2$ for η small, where the primary coefficients are known to be enhanced with respect to the purely non-global ones because of the presence of large collinear logarithms (cf section 2.2). Figure 18 compares the fixed-order results to the all-orders one up to α_s^5 . On these 2 plots one can notice that the series seems to converge, as was already shown in section 3.2. In other respects, the convergence looks better when η is larger. This may be understood using the following simple explanation: if we make an expansion up to order k , then the series starts to diverge from the exact result when the term of order $k + 1$ becomes roughly

²²Starting from $4N_c L t \sim 1$ quickly converges to the true solution.

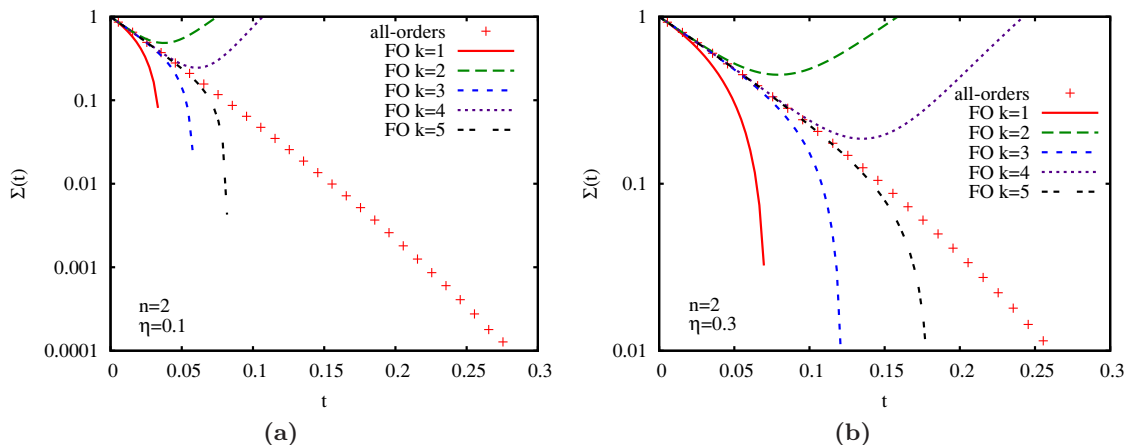


Figure 18. Comparison between fixed-order (FO) expansion and all-orders result when $n = 2$ for (a) $\eta = 0.1$ and (b) $\eta = 0.3$.

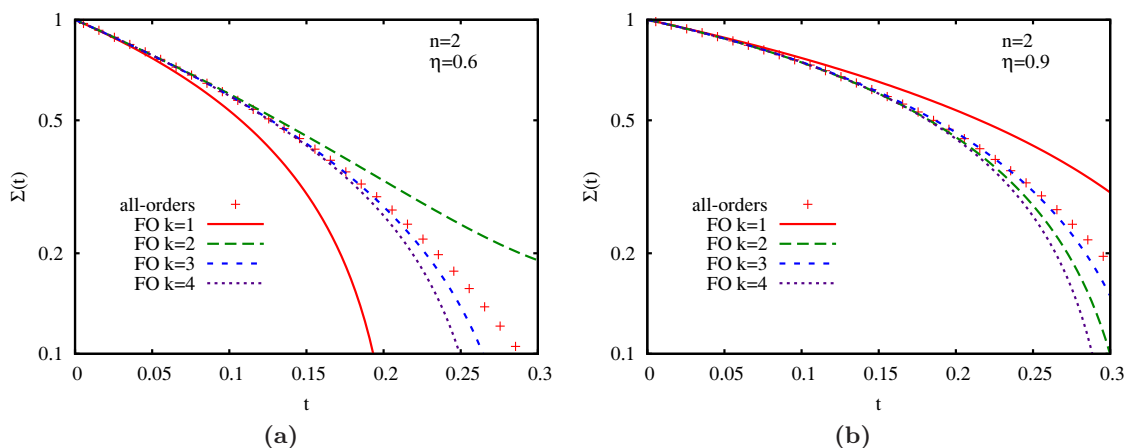


Figure 19. Comparison between fixed-order (FO) expansion and all-orders result when $n = 2$ for (a) $\eta = 0.6$ and (b) $\eta = 0.9$.

of the same size as the function itself. In the $\Sigma^{(2)}(L, t)$ case, using the analytic expression (eq. (2.25)), this can be translated to

$$\frac{(4N_c L t)^{k+1}}{(k+1)!} \sim e^{-4N_c L t}, \tag{C.1}$$

with $L = \ln \frac{1}{\eta}$. At large k ,²³ the solution gives

$$t \sim \frac{k+1}{4N_c L}. \tag{C.2}$$

So the convergence is better when k and η increase.

Let us go to larger η , where there is no large collinear logarithm anymore, and check what happens. This is done for $\eta = 0.6$ and $\eta = 0.9$ on figure 19. One striking feature

²³The derivation was done at large k , but the result seems to be reasonable even for $k = 2$.

of these plots is that the convergence seems acceptable up to the third order, but the 4th order does not give as good a result, and the situation becomes worse as η increases. Said another way, the perturbative expansion can be trusted until the third order, but then it starts to diverge. The fact that the convergence looks better for small η may come from the dominant behavior of the primary series, which converges very well to a nice exponential function. However, if one could go to sufficiently high orders, it might be possible to observe the same divergence as in figure 19, when the purely non-global coefficients become of the same order of magnitude as the primary ones.

To get an idea of these coefficients, the series of the plots are explicitly written below:

$$\Sigma^{(2)}(\eta = 0.1, t) = 1 - 27.57t + 368.8t^2 - 3195t^3 + 20200t^4 - 99300t^5 + \mathcal{O}(t^6), \quad (C.3)$$

$$\Sigma^{(2)}(\eta = 0.3, t) = 1 - 13.88t + 87.49t^2 - 334t^3 + 860t^4 - 1500t^5 + \mathcal{O}(t^6), \quad (C.4)$$

$$\Sigma^{(2)}(\eta = 0.6, t) = 1 - 4.656t + 6.53t^2 - 6.7t^3 - 10t^4 + \mathcal{O}(t^5), \quad (C.5)$$

$$\Sigma^{(2)}(\eta = 0.9, t) = 1 - 2.320t - 2.27t^2 + 1.9t^3 - 12t^4 + \mathcal{O}(t^5). \quad (C.6)$$

For η small, the growth of these coefficients comes from the powers of the large collinear logarithm L . For η near 1, the small growth that one can start to observe at the 4th order essentially comes from the purely non-global part. Indeed, at the 4th order, the primary coefficient is positive, and it is not the case for $\eta > 0.6$. This non-global growth will be confirmed at higher orders for the slice observable in section C.3.

Let us now see what happens if the perturbative series is exponentiated. This means that instead of plotting $g(t) = 1 + \sum_{i=1}^k a_i t^i$, which is the perturbative series, one plots $e^{f(t)}$ where

$$f(t) = \sum_{i=1}^k c_i t^i \text{ and } e^{f(t)} = g(t) + \mathcal{O}(t^{k+1}), \quad (C.7)$$

so that

$$c_1 = a_1, \quad (C.8)$$

$$c_2 = a_2 - \frac{a_1^2}{2}, \quad (C.9)$$

...

Notice that the exponentiated first order corresponds to the analytical estimate for $n = 2$. One can observe on figure 20 a nice convergence for $\eta = 0.1$ until the 4th order. Concerning the order 5, it seems that it diverges a little, but if the coefficients of the series are varied within their respective errors, then it can coincide with the all-orders curve. However, one can guess that it should not converge at the end because the all-orders function is not strictly a simple exponential. This will be confirmed later with the slice. When $\eta = 0.9$, the exponentiated 4th order surprisingly almost fits the all-orders result. Even more surprising, if the 5th order coefficient (not shown here) is varied within its error band, it seems that the exponentiated fit is still improved. Is it accidental? Here again, the slice example and good sense lead us to answer yes but we cannot be completely sure.

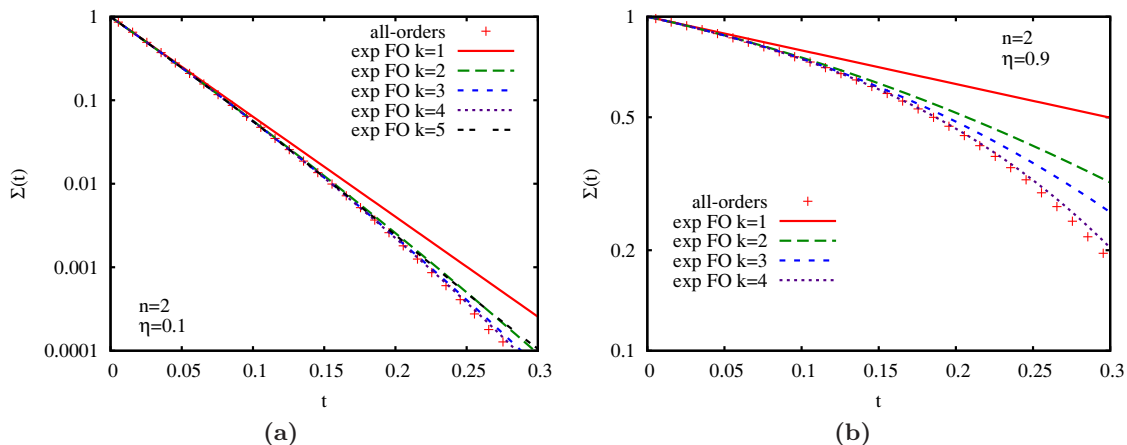


Figure 20. Comparison between exponentiated fixed-order (exp FO) expansion and all-orders result when $n = 2$ for (a) $\eta = 0.1$ and (b) $\eta = 0.9$.

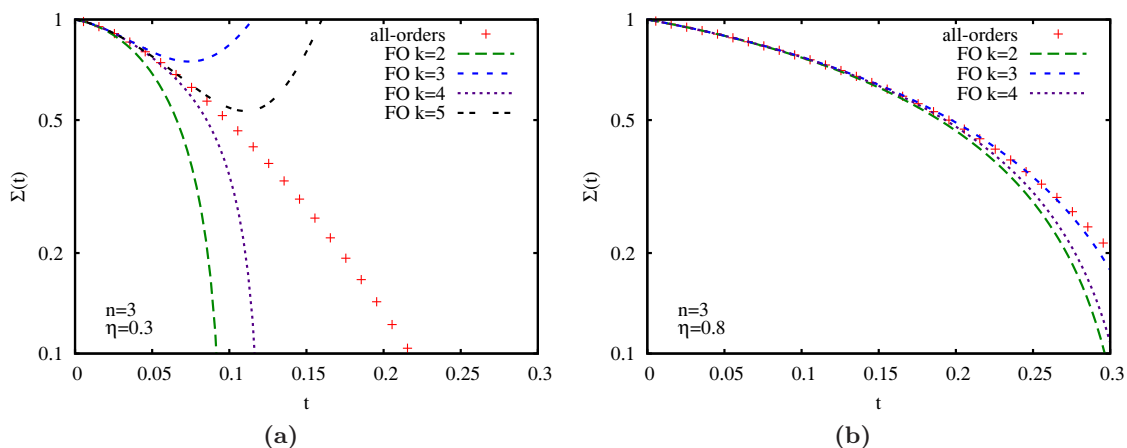


Figure 21. Comparison between fixed-order expansion and all-orders result when $n = 3$ for (a) $\eta = 0.3$ and (b) $\eta = 0.8$.

C.2 Case $n_{\text{fl}t} = 3$

For small enough η , it was found analytically in section 2.3 that the all-orders result for the Higgs ΔM distribution can be expressed as a simple function of t (eq. (2.27)), which is of the form $g(4N_c L t)$ with:

$$g(t) = e^{-t} \left(1 + \int_0^t dt' \frac{1 - e^{t'}}{t'} \right). \quad (\text{C.10})$$

The radius of convergence of the Taylor series for g is infinite as the coefficients a_k resulting from its expansion can be shown to be bounded for large k by:

$$|a_k| < \frac{1}{k!} (2^k + \mathcal{O}(1)). \quad (\text{C.11})$$

As the expansion of this function converges, one would expect the same to occur for the curves obtained numerically. However, the expansion does not converge as fast as for the

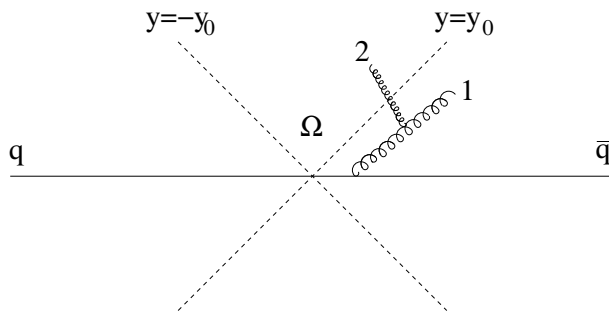


Figure 22. The region Ω between the 2 dashed lines, and the gluon’s configuration leading to the appearance of non-global logarithms.

exponential, so that the t window for which there is a convergence should be smaller than what was obtained for $n = 2$. This is indeed what is observed in figure 21a for $\eta = 0.3$, if it is compared for instance with the plot 18b.

In other respects, the curves on the plot 21b for η near 1 behave similarly as for $n = 2$, i.e. a perturbative series that converges until the 3rd order, and that starts to diverge from the 4th order. Notice that the same behaviour is observed for $n > 3$.

C.3 The Slice case

The Slice observable, studied for instance in [41], gives an interesting example of the strange behaviour of the non-global leading-log series. It is simply defined by the sum of all the particles’ energy flowing into a region Ω of the phase space corresponding to a rapidity $y \in [-y_0, y_0]$, with y_0 a parameter of the observable (figure 22). Here, we work in the $q\bar{q}$ center of mass frame and the quarks are assumed to move along the z axis. This observable is non-global as shown in figure 22, and it is interesting in 2 ways:

- A simple change of frame²⁴ shows that it is more or less equivalent in the leading-log approximation to the ΔM observable in the filtering analysis using anti- k_t with $n = 2$ and $\eta = e^{-y_0}$ when $\eta \ll 1$, while being faster to compute numerically due to the absence of clustering.
- When $e^{-y_0} \sim 1$, which should very approximately correspond to $\eta = \mathcal{O}(1)$, the strange behaviour of the non-global series, observed in section 3, is clearly confirmed with the addition of the 5th and 6th orders.

The series are represented in figure 23 for 2 different values of y_0 . The plot for $y_0 = 2.3$ is here for a comparison with figure 18a, i.e. when $n = 2$ and $\eta = 0.1$. One can notice that until order 5 the behaviours of the 2 series are very similar. The slight difference comes from the fact that the C/A algorithm was used there instead of anti- k_t . This can also be seen on the expansions eqs. (C.3), (C.12). However, a remarkable effect is the 6th order curve which does not improve the fit with the all-orders result anymore (it even makes it

²⁴Or more simply comparison between primary results, eq. (2.5) from [41] and eq. (2.10) from this paper.

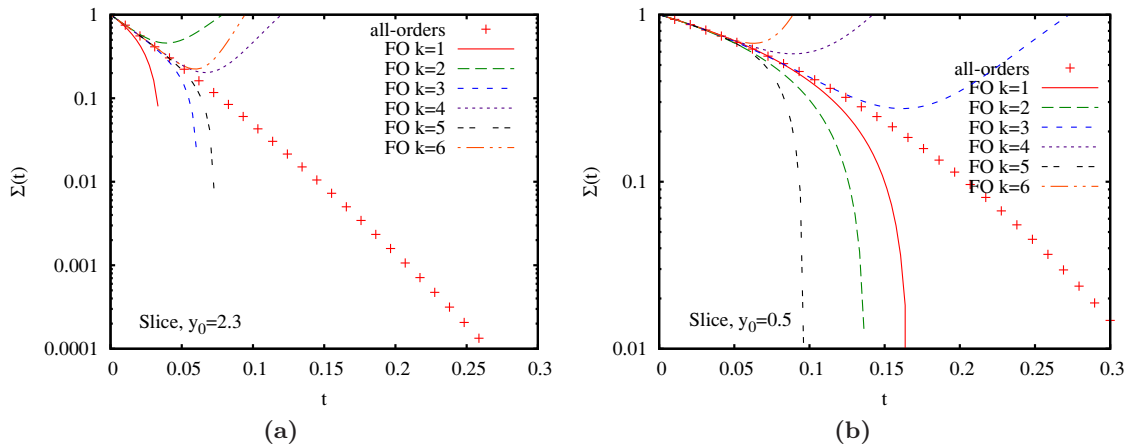


Figure 23. Comparison between all-orders and fixed-order results for the Slice when (a) $y_0 = 2.3$ and (b) $y_0 = 0.5$.

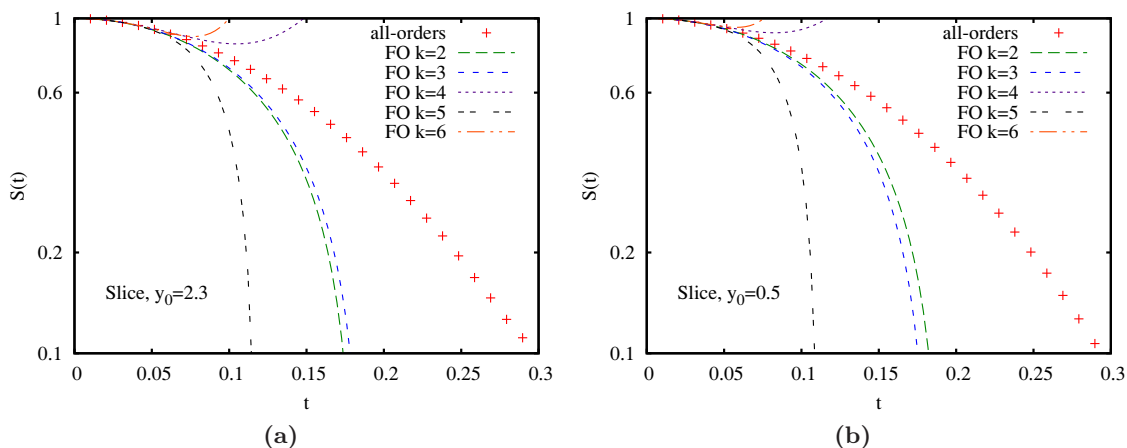


Figure 24. Comparison between all-orders and fixed-order results for the purely non-global part $S(t)$ of the Slice when (a) $y_0 = 2.3$ and (b) $y_0 = 0.5$.

worse). When $y_0 = 0.5$ (figure 23b), this effect is enhanced: indeed, one notices that the 3rd order gives the best result, with the 2nd order even worse than the 1st one. And here again, adding more orders shifts point of disagreement to smaller t .

To get an idea of the coefficients in this case, the series of the plots are given below:

$$\Sigma(y_0 = 2.3, t) = 1 - 27.6t + 351.1t^2 - 2673t^3 + 13900t^4 - 99500t^5 + 210^6t^6 + \mathcal{O}(t^7), \quad (\text{C.12})$$

$$\Sigma(y_0 = 0.5, t) = 1 - 6t - 9.108t^2 + 114t^3 + 1740t^4 - 68400t^5 + 1.610^6t^6 + \mathcal{O}(t^7). \quad (\text{C.13})$$

The growth of the coefficients for $y_0 = 2.3$ until the 5th order essentially comes from the powers of $4N_c y_0$ when expanding the primary result, which is the expression of the collinear divergence near q and \bar{q} , whereas, for $y_0 = 0.5$, it essentially takes its origin from the purely non-global part. There is no large enhancement due to collinear divergence which can explain it. As an example, we also show the function $S(y_0, t)$, which is defined

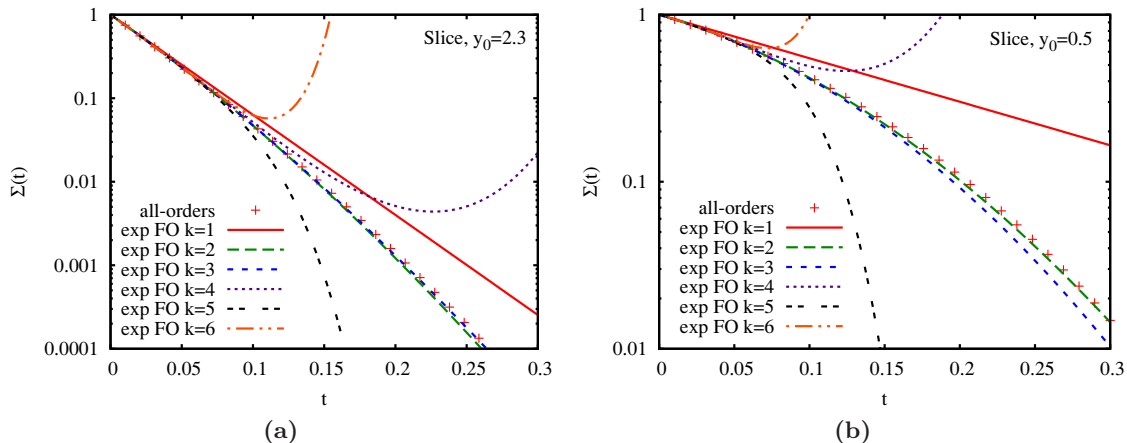


Figure 25. Comparison between all-orders and exponentiated fixed-order results for the Slice when (a) $y_0 = 2.3$ and (b) $y_0 = 0.5$.

as in [35, 41] to contain the purely non-global part of the result:

$$S(y_0, t) \equiv \frac{\Sigma(y_0, t)}{\Sigma^{(P)}(y_0, t)}, \tag{C.14}$$

where, for the slice, the primary contribution $\Sigma^{(P)}$ can be written as:

$$\Sigma^{(P)}(y_0, t) = e^{-4N_c y_0 t}, \tag{C.15}$$

in the large- N_c limit. The plots for $S(y_0, t)$ are shown in figure 24. One observes the saturation property noticed in [41] which leads to very similar plots for $y_0 = 0.5$ and $y_0 = 2.3$. There is clearly no convergence of the perturbative series.

Therefore, the non-global leading-log large- N_c series seems to behave badly at high orders. Does this mean that it is an asymptotic series like the Standard Model is known to be [68]?²⁵ This study cannot answer such a question but, at least, one should be aware of the strange behaviour of the non-global series.

To finish, let us mention the exponentiated results represented in figure 25. They show that the convergence observed in the case $n = 2$ is clearly only illusory. One point to notice is that the case $y_0 = 2.3$ is slightly different from its $\eta = 0.1$ counterpart as the divergence starts to be visible at order 4 instead of 5. This may be because using the C/A algorithm in the filtered Higgs mass observable reduces the impact of the non-global logarithms [42, 43] and the primaries still dominate at order 4, whereas it is not the case anymore for the slice. Notice also the very nice fit given by the exponentiated 2nd order curve when $y_0 = 0.5$.

Open Access. This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

²⁵Though the origin would be different if it were the case.

References

- [1] M.H. Seymour, *Searches for new particles using cone and cluster jet algorithms: a comparative study*, *Z. Phys. C* **62** (1994) 127 [[SPIRES](#)].
- [2] J.M. Butterworth, B.E. Cox and J.R. Forshaw, *WW scattering at the CERN LHC*, *Phys. Rev. D* **65** (2002) 096014 [[hep-ph/0201098](#)] [[SPIRES](#)].
- [3] W. Skiba and D. Tucker-Smith, *Using jet mass to discover vector quarks at the LHC*, *Phys. Rev. D* **75** (2007) 115010 [[hep-ph/0701247](#)] [[SPIRES](#)].
- [4] B. Holdom, *t' at the LHC: the physics of discovery*, *JHEP* **03** (2007) 063 [[hep-ph/0702037](#)] [[SPIRES](#)].
- [5] L.G. Almeida, S.J. Lee, G. Perez, I. Sung and J. Virzi, *Top jets at the LHC*, *Phys. Rev. D* **79** (2009) 074012 [[arXiv:0810.0934](#)] [[SPIRES](#)].
- [6] D.E. Kaplan, K. Rehermann, M.D. Schwartz and B. Tweedie, *Top tagging: a method for identifying boosted hadronically decaying top quarks*, *Phys. Rev. Lett.* **101** (2008) 142001 [[arXiv:0806.0848](#)] [[SPIRES](#)].
- [7] J. Thaler and L.-T. Wang, *Strategies to identify boosted tops*, *JHEP* **07** (2008) 092 [[arXiv:0806.0023](#)] [[SPIRES](#)].
- [8] T. Plehn, G.P. Salam and M. Spannowsky, *Fat jets for a light Higgs*, *Phys. Rev. Lett.* **104** (2010) 111801 [[arXiv:0910.5472](#)] [[SPIRES](#)].
- [9] J.M. Butterworth, J.R. Ellis and A.R. Raklev, *Reconstructing sparticle mass spectra using hadronic decays*, *JHEP* **05** (2007) 033 [[hep-ph/0702150](#)] [[SPIRES](#)].
- [10] J.M. Butterworth, J.R. Ellis, A.R. Raklev and G.P. Salam, *Discovering baryon-number violating neutralino decays at the LHC*, *Phys. Rev. Lett.* **103** (2009) 241803 [[arXiv:0906.0728](#)] [[SPIRES](#)].
- [11] U. Baur and L.H. Orr, *Searching for $t\bar{t}$ resonances at the Large Hadron Collider*, *Phys. Rev. D* **77** (2008) 114001 [[arXiv:0803.1160](#)] [[SPIRES](#)].
- [12] P. Fileviez Perez, R. Gavin, T. McElmurry and F. Petriello, *Grand unification and light color-octet scalars at the LHC*, *Phys. Rev. D* **78** (2008) 115017 [[arXiv:0809.2106](#)] [[SPIRES](#)].
- [13] Y. Bai and Z. Han, *Top-antitop and top-top resonances in the dilepton channel at the CERN LHC*, *JHEP* **04** (2009) 056 [[arXiv:0809.4487](#)] [[SPIRES](#)].
- [14] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination algorithms and jet substructure: pruning as a tool for heavy particle searches*, [arXiv:0912.0033](#) [[SPIRES](#)].
- [15] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001 [[arXiv:0802.2470](#)] [[SPIRES](#)].
- [16] ATLAS collaboration, *ATLAS sensitivity to the standard model Higgs in the HW and HZ channels at high transverse momenta*, Technical Report ATL-PHYS-PUB-2009-088 [ATL-COM-PHYS-2009-345], CERN, Geneva Switzerland (2009).
- [17] G.D. Kribs, A. Martin, T.S. Roy and M. Spannowsky, *Discovering the Higgs boson in new physics events using jet substructure*, [arXiv:0912.4731](#) [[SPIRES](#)].
- [18] M. Cacciari, J. Rojo, G.P. Salam and G. Soyez, *Quantifying the performance of jet definitions*

- for kinematic reconstruction at the LHC, *JHEP* **12** (2008) 032 [[arXiv:0810.1304](#)] [[SPIRES](#)].
- [19] D. Krohn, J. Thaler and L.-T. Wang, *Jet trimming*, *JHEP* **02** (2010) 084 [[arXiv:0912.1342](#)] [[SPIRES](#)].
- [20] V.S. Fadin, *Double logarithmic asymptotics of the cross-sections of e^+e^- annihilation into quarks and gluons* (in Russian), *Yad. Fiz.* **37** (1983) 408 [[SPIRES](#)].
- [21] B.I. Ermolaev and V.S. Fadin, *Log-log asymptotic form of exclusive cross-sections in quantum chromodynamics*, *JETP Lett.* **33** (1981) 269 [*Pisma Zh. Eksp. Teor. Fiz.* **33** (1981) 285] [[SPIRES](#)].
- [22] A.H. Mueller, *On the multiplicity of hadrons in QCD jets*, *Phys. Lett. B* **104** (1981) 161 [[SPIRES](#)].
- [23] Y.L. Dokshitzer, V.S. Fadin and V.A. Khoze, *Double logs of perturbative QCD for parton jets and soft hadron spectra*, *Z. Phys. C* **15** (1982) 325 [[SPIRES](#)].
- [24] A. Bassetto, M. Ciafaloni and G. Marchesini, *Jet structure and infrared sensitive quantities in perturbative QCD*, *Phys. Rept.* **100** (1983) 201 [[SPIRES](#)].
- [25] S. Catani, G. Turnock, B.R. Webber and L. Trentadue, *Thrust distribution in e^+e^- annihilation*, *Phys. Lett. B* **263** (1991) 491 [[SPIRES](#)].
- [26] S. Catani, Y.L. Dokshitzer, F. Fiorani and B.R. Webber, *Average number of jets in e^+e^- annihilation*, *Nucl. Phys. B* **377** (1992) 445 [[SPIRES](#)].
- [27] S. Catani, L. Trentadue, G. Turnock and B.R. Webber, *Resummation of large logarithms in e^+e^- event shape distributions*, *Nucl. Phys. B* **407** (1993) 3 [[SPIRES](#)].
- [28] S. Catani, G. Turnock and B.R. Webber, *Jet broadening measures in e^+e^- annihilation*, *Phys. Lett. B* **295** (1992) 269 [[SPIRES](#)].
- [29] S. Catani and B.R. Webber, *Resummed C -parameter distribution in e^+e^- annihilation*, *Phys. Lett. B* **427** (1998) 377 [[hep-ph/9801350](#)] [[SPIRES](#)].
- [30] Y.L. Dokshitzer, A. Lucenti, G. Marchesini and G.P. Salam, *On the QCD analysis of jet broadening*, *JHEP* **01** (1998) 011 [[hep-ph/9801324](#)] [[SPIRES](#)].
- [31] V. Antonelli, M. Dasgupta and G.P. Salam, *Resummation of thrust distributions in DIS*, *JHEP* **02** (2000) 001 [[hep-ph/9912488](#)] [[SPIRES](#)].
- [32] S.J. Burby, *The four-jet rate in e^+e^- annihilation*, *Phys. Lett. B* **453** (1999) 54 [[hep-ph/9902305](#)] [[SPIRES](#)].
- [33] S.J. Burby and E.W.N. Glover, *Resumming the light hemisphere mass and narrow jet broadening distributions in e^+e^- annihilation*, *JHEP* **04** (2001) 029 [[hep-ph/0101226](#)] [[SPIRES](#)].
- [34] A. Banfi, G. Marchesini, Y.L. Dokshitzer and G. Zanderighi, *QCD analysis of near-to-planar 3-jet events*, *JHEP* **07** (2000) 002 [[hep-ph/0004027](#)] [[SPIRES](#)].
- [35] M. Dasgupta and G.P. Salam, *Resummation of non-global QCD observables*, *Phys. Lett. B* **512** (2001) 323 [[hep-ph/0104277](#)] [[SPIRES](#)].
- [36] A. Banfi and M. Dasgupta, *Dijet rates with symmetric E_t cuts*, *JHEP* **01** (2004) 027 [[hep-ph/0312108](#)] [[SPIRES](#)].
- [37] A. Banfi, G.P. Salam and G. Zanderighi, *Resummed event shapes at hadron-hadron colliders*, *JHEP* **08** (2004) 062 [[hep-ph/0407287](#)] [[SPIRES](#)].

- [38] A. Banfi, G.P. Salam and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, *JHEP* **03** (2005) 073 [[hep-ph/0407286](#)] [[SPIRES](#)].
- [39] T. Becher and M.D. Schwartz, *A precise determination of α_s from LEP thrust data using effective field theory*, *JHEP* **07** (2008) 034 [[arXiv:0803.0342](#)] [[SPIRES](#)].
- [40] G. Dissertori et al., *Determination of the strong coupling constant using matched NNLO+NLLA predictions for hadronic event shapes in e^+e^- annihilations*, *JHEP* **08** (2009) 036 [[arXiv:0906.3436](#)] [[SPIRES](#)].
- [41] M. Dasgupta and G.P. Salam, *Accounting for coherence in interjet E_t flow: a case study*, *JHEP* **03** (2002) 017 [[hep-ph/0203009](#)] [[SPIRES](#)].
- [42] Y. Delenda, R. Appleby, M. Dasgupta and A. Banfi, *On QCD resummation with k_t clustering*, *JHEP* **12** (2006) 044 [[hep-ph/0610242](#)] [[SPIRES](#)].
- [43] R.B. Appleby and M.H. Seymour, *Non-global logarithms in inter-jet energy flow with k_t clustering requirement*, *JHEP* **12** (2002) 063 [[hep-ph/0211426](#)] [[SPIRES](#)].
- [44] A. Banfi, G. Marchesini and G. Smye, *Away-from-jet energy flow*, *JHEP* **08** (2002) 006 [[hep-ph/0206076](#)] [[SPIRES](#)].
- [45] Y. Hatta and T. Ueda, *Jet energy flow at the LHC*, *Phys. Rev. D* **80** (2009) 074018 [[arXiv:0909.0056](#)] [[SPIRES](#)].
- [46] H. Weigert, *Non-global jet evolution at finite N_c* , *Nucl. Phys. B* **685** (2004) 321 [[hep-ph/0312050](#)] [[SPIRES](#)].
- [47] J.R. Forshaw, A. Kyrielleis and M.H. Seymour, *Super-leading logarithms in non-global observables in QCD*, *JHEP* **08** (2006) 059 [[hep-ph/0604094](#)] [[SPIRES](#)].
- [48] M. Cacciari, G.P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [[arXiv:0802.1189](#)] [[SPIRES](#)].
- [49] Y.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001 [[hep-ph/9707323](#)] [[SPIRES](#)].
- [50] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep-inelastic scattering*, [hep-ph/9907280](#) [[SPIRES](#)].
- [51] F. Fiorani, G. Marchesini and L. Reina, *Soft gluon factorization and multi-gluon amplitude*, *Nucl. Phys. B* **309** (1988) 439 [[SPIRES](#)].
- [52] Y.L. Dokshitzer, G. Marchesini and G. Oriani, *Measuring color flows in hard processes: beyond leading order*, *Nucl. Phys. B* **387** (1992) 675 [[SPIRES](#)].
- [53] Y. Dokshitzer, V. Khoze, A. Mueller and S. Troyan, *Basics of perturbative QCD*, Editions Frontières, Gif-sur-Yvette, France (1991) [[SPIRES](#)].
- [54] M. Cacciari and G.P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett. B* **641** (2006) 57 [[hep-ph/0512210](#)] [[SPIRES](#)].
- [55] C. Buttar et al., *Standard model handles and candles working group: tools and jets summary report*, [arXiv:0803.0678](#) [[SPIRES](#)].
- [56] M. Dasgupta, L. Magnea and G.P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, *JHEP* **02** (2008) 055 [[arXiv:0712.3014](#)] [[SPIRES](#)].
- [57] M. Cacciari, G.P. Salam and G. Soyez, *The catchment area of jets*, *JHEP* **04** (2008) 005 [[arXiv:0802.1188](#)] [[SPIRES](#)].

- [58] M. Cacciari, G.P. Salam and S. Sapeta, *On the characterisation of the underlying event*, [arXiv:0912.4926](#) [[SPIRES](#)].
- [59] M. Cacciari and G.P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett. B* **659** (2008) 119 [[arXiv:0707.1378](#)] [[SPIRES](#)].
- [60] T. Sjöstrand et al., *High-energy physics event generation with PYTHIA 6.1*, *Comput. Phys. Commun.* **135** (2001) 238 [[hep-ph/0010017](#)] [[SPIRES](#)].
- [61] T. Sjöstrand, L. Lönnblad, S. Mrenna and P.Z. Skands, *PYTHIA 6.3 physics and manual*, [hep-ph/0308153](#) [[SPIRES](#)].
- [62] M. Rubin, *Light Higgs searches at the LHC using jet substructure*, [arXiv:0905.2124](#) [[SPIRES](#)].
- [63] G.P. Korchemsky and G. Sterman, *Nonperturbative corrections in resummed cross-sections*, *Nucl. Phys. B* **437** (1995) 415 [[hep-ph/9411211](#)] [[SPIRES](#)].
- [64] M. Dasgupta and Y. Delenda, *On the universality of hadronisation corrections to QCD jets*, *JHEP* **07** (2009) 004 [[arXiv:0903.2187](#)] [[SPIRES](#)].
- [65] G. Corcella et al., *HERWIG 6.5: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, *JHEP* **01** (2001) 010 [[hep-ph/0011363](#)] [[SPIRES](#)].
- [66] G. Corcella et al., *HERWIG 6.5 release note*, [hep-ph/0210213](#) [[SPIRES](#)].
- [67] T. Sjöstrand, S. Mrenna and P.Z. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026 [[hep-ph/0603175](#)] [[SPIRES](#)].
- [68] M. Beneke, *Renormalons*, *Phys. Rept.* **317** (1999) 1 [[hep-ph/9807443](#)] [[SPIRES](#)].