3 Fixed-order calculations for jet substructure

3.1 Introduction ........................................... 89
3.2 Plain jet mass ........................................ 90
3.3 Mass drop tagger ..................................... 92
  3.3.1 Definition ......................................... 92
  3.3.2 Leading order ..................................... 93
  3.3.3 Next-to-leading order: The wrong branch issue .. 96
3.4 Modified mass drop tagger ......................... 98
  3.4.1 Definition ......................................... 98
  3.4.2 Next-to-leading order: Independent emission .... 98
  3.4.3 Next-to-leading order: Flavour changing contributions. 105
  3.4.4 Non-global logarithms ......................... 106
  3.4.5 Phenomenology .................................. 109
3.5 Pruning .................................................. 111
  3.5.1 Definition ......................................... 111
  3.5.2 Leading order ..................................... 112
  3.5.3 Next-to-leading order: Independent emission .... 113
  3.5.4 Next-to-leading order: Non-Abelian terms ....... 117
  3.5.5 Y-pruning ......................................... 120
  3.5.6 Pruning in the region $\rho > z_{\text{cut}}^2 \Delta R$ .... 122
  3.5.7 Phenomenology .................................. 124
3.6 Trimming ................................................. 126
  3.6.1 Definition ......................................... 126
  3.6.2 Leading order ..................................... 126
  3.6.3 Next-to-leading order: Independent emission .... 127
  3.6.4 Next-to-leading order: Non-Abelian terms ....... 131
  3.6.5 Phenomenology .................................. 132
3.7 Conclusions ............................................ 134
<table>
<thead>
<tr>
<th>4</th>
<th>Jet substructure methods for signal jets</th>
<th>139</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>139</td>
</tr>
<tr>
<td>4.2</td>
<td>Results for plain jet mass</td>
<td>141</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Initial state radiation</td>
<td>143</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Final state radiation</td>
<td>147</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Non-perturbative contributions</td>
<td>148</td>
</tr>
<tr>
<td>4.3</td>
<td>Trimming</td>
<td>149</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Lowest order result</td>
<td>149</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Initial state radiation</td>
<td>151</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Final state radiation</td>
<td>154</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Non-perturbative contributions</td>
<td>158</td>
</tr>
<tr>
<td>4.4</td>
<td>Pruning and mMDT</td>
<td>159</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Pruning</td>
<td>159</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Modified mass drop tagger</td>
<td>161</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Non-perturbative effects and MC results</td>
<td>163</td>
</tr>
<tr>
<td>4.5</td>
<td>Y-pruning and Y-splitter</td>
<td>164</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Y-pruning</td>
<td>165</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Y-splitter</td>
<td>167</td>
</tr>
<tr>
<td>4.6</td>
<td>Y-splitter with trimming</td>
<td>171</td>
</tr>
<tr>
<td>4.7</td>
<td>Optimal parameter values</td>
<td>173</td>
</tr>
<tr>
<td>4.7.1</td>
<td>mMDT</td>
<td>175</td>
</tr>
<tr>
<td>4.7.2</td>
<td>Pruning and Y-pruning</td>
<td>178</td>
</tr>
<tr>
<td>4.7.3</td>
<td>Trimming</td>
<td>181</td>
</tr>
<tr>
<td>4.8</td>
<td>Conclusions</td>
<td>184</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5</th>
<th>Y-splitter with trimming</th>
<th>189</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>189</td>
</tr>
<tr>
<td>5.2</td>
<td>Y-splitter</td>
<td>191</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Definition</td>
<td>191</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Leading order</td>
<td>192</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Next-to-leading order</td>
<td>194</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Resummation</td>
<td>200</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Y-splitter: The anomalous configuration</td>
<td>204</td>
</tr>
<tr>
<td>5.3</td>
<td>Y-splitter with trimming</td>
<td>206</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Next-to-leading order</td>
<td>207</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Resummation</td>
<td>210</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Hadronisation and UE</td>
<td>216</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Optimal parameter values</td>
<td>217</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusions</td>
<td>219</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1.1 Feynman diagrams contributing to the 1-loop running coupling. The dotted line represents ghost propagators. 44

1.2 A schematic representation of a hard scattering process in the parton model description. In both panels, a pair of partons from two hadrons (circles) participate in a partonic hard scattering (square) with a given longitudinal momentum fraction $x_i$. The left panel shows the tree-level picture. The right panel is identical to the left, but with an additional QCD correction corresponding to an emission in the initial state; note that this emission changes the momentum fraction of the parton obtained from the hadron. The additional parallel lines emerging from the hadrons represent the remnants of the hadron after scattering. 46

1.3 Feynman diagrams showing the production of a final state quark-antiquark pair with momentum $p_1$ and $p_2$ respectively via an arbitrary colour neutral decay, in that one leg emits a real gluon with momentum $k$. 50

1.4 Feynman diagram showing the production of a final state quark-antiquark pair with momentum $p_1$ and $p_2$ respectively via an arbitrary colour neutral decay, in that a virtual gluon is transferred with momentum $k$. 53

2.1 An example of a parton level event that has been clustered into jets using different jet algorithms. The event consists of a set of hard particles generated using Herwig [86] that occupy cells in rapidity and azimuthal angle $(y, \phi)$ space with $p_T$ given by the vertical axis. A large number of soft, randomly distributed “ghost” particles are added to the event before clustering. If a cell contained a particle that is clustered into a hard jet, then it is coloured accordingly. This gives an indication of the relative behaviour and catchment area of different jet algorithms. This figure is taken directly from Ref. [83]. 67
2.2 Higgs decay in its rest frame (left) and in the boosted frame (right).
   The boost axis is represented by the dashed line and $\beta$ is the velocity
   coefficient in the Lorentz factor. For highly boosted Higgs decay, this
   axis is approximately perpendicular to the beam direction. The quantity
   $x$ is the energy fraction carried by the Higgs decay products in the
   boosted frame. ...................................................... 69

2.3 The normalised differential distribution of the minimum subjet $p_T$
   fraction for quark, gluon and Higgs jets. Events were generated using
   HERWIG++ 2.7.0 [75] with 14 TeV $pp$ collisions, generator level cuts $p_T^{\text{jet}}>3$
   TeV and $M_{\text{jet}} = M_H \pm 25$ GeV with hadronisation and underlying event.
   The Higgs jets were generated using the process $pp \rightarrow ZH$ with the
   $Z$ and Higgs boson decaying leptonically and hadronically respectively.
   The C/A algorithm as implemented by FASTJET [106] and RIVET [107],
   was used with $R = 1$ to cluster the jets and the hardest hadronic jet in
   each event was selected. The constituents of this hardest jet were sub-
   sequently clustered using an inclusive $k_t$ algorithm and then the final
   merging step was undone to obtain exactly two subjets, from which $z$
   was calculated. ...................................................... 72

2.4 The normalised differential jet mass distribution for quark (left) and
   Higgs (right) jets before and after application of the MDT with filtering
   algorithm. This was generated using HERWIG++ 2.7.0 [75] with gener-
   ator level cut $p_T^{\text{jet}}>3$ TeV including hadronisation and underlying
   event. For comparison purposes, we assumed a b-tagging efficiency of
   100% and considered all final-state hadrons to be stable. We used the
   C/A algorithm with $R = 1$ to cluster the jets and only considered the
   hardest jet in each event. Parameters used for MDT with filtering are
   $y_{\text{cut}} = 0.11$, $\mu = 0.67$, $R_{\text{filter}} = \min(0.3, R_{\text{tag}}/2)$ and $n_{\text{filter}} = 3$. ........ 76

2.5 Comparison of fixed-order LO and NLO calculations and resummed re-
   sults for differential distribution of total broadening observable at parton
   level to experimental data. [167][168] ........................................ 83
2.6 Schematic diagram that shows the perturbative contributions to an observable that result from an $\alpha_s$ expansion of a resummed result at each order. Each circle indicates the existence of a contribution of the form $\alpha_s^m L^n$. The colour of each circle shows which terms in the expansion are resummed by an exponential with LL accuracy (grey), NLL accuracy in the expansion (grey+red) and NLL accuracy in the exponent with the first hard, wide-angle correction (grey+red+green). The partially filled circles indicate that some (but not all) terms of this order are captured by the resummed result. The contributions with white circles require greater accuracy in the exponent and/or additional hard wide-angle corrections to the observable to be included.

2.7 Examples of kinematical configurations that give rise to non-global logarithms (left) and clustering logarithms (right) for a jet defined with radius $R$. In the left-hand diagram, soft emission $k_2$ is emitted with opening angle less than $R$ with respect to $p$. If clustered to $p$, $k_2$ contributes to the jet, introducing non-global logarithms. In the right-hand diagram, soft emission $k_2$ is emitted with opening angle greater than $R$ with respect to $p$. However, if initially clustered to $k_1$, $k_2$ will contribute to the jet, introducing clustering logarithms.

2.8 Resummed results for the differential distribution of the total broadening observable with and without hadronic corrections, compared to experimental data. [167,168]

3.1 The differential distribution in $\rho$ for a range of jets tagged with different jet substructure techniques using Monte Carlo. The results were obtained from $qq \rightarrow qq$ Monte Carlo simulation with Pythia 6.425 [74] in the DW tune [189] (virtuality-ordered shower), with a minimum $p_T$ cut in the generation of 3 TeV, for 14 TeV pp collisions, at parton level, including initial and final-state showering, but without the underlying event (multiple interactions). The tagger has been applied to the two leading C/A jets with $R = 1$. This figure and caption has been adapted directly from Ref. [116].

3.2 Comparison of EVENT2 at LO with the analytic expression Eq. (3.9) for a range of $y_{\text{cut}}$ values. The red points tend to a constant value in the small mass limit, which corresponds to the presence of a single logarithm in the integrated distribution. The green points show the difference between the analytic expression and the EVENT2 result in the region $\rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta_R$, which tends to zero.
3.3 NLO wrong branch configuration in the $C_F C_A$ channel. To obtain the $C_F n_f$ contribution, we replace $k_1$ and $k_2$ with a quark-antiquark pair. .... 96

3.4 Comparison of Event2 in the $C_F C_A$ and $C_F n_f$ channels at NLO with the analytic expression Eq. (3.13) on the left and Eq. (3.14) on the right, for a range of $y_{\text{cut}}$ values. The red points demonstrate an extra logarithm in both the $C_F C_A$ and $C_F n_f$ channels. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < \frac{y_{\text{cut}}}{\rho_{\text{cut}} + y_{\text{cut}}} \Delta_R$, which demonstrates that we control all terms up to $O(\alpha_s^2)$ in the differential distribution. .... 99

3.5 All real and virtual contributions to the jet mass distribution at NLO in the $C_F^2$ channel. .......................................................... 100

3.6 Comparison of Event2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.24), for a range of $y_{\text{cut}}$ values. The red points show a straight line for small $\rho$, which indicates single logarithmic behaviour in the integrated distribution. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < \frac{y_{\text{cut}}}{\rho_{\text{cut}} + y_{\text{cut}}} R^2$, which demonstrates that we control all terms up to $O(\alpha_s^2)$ in the differential distribution, for small $\rho$. .... 104

3.7 Comparison of Event2 in the $C_F C_A$ and $C_F n_f$ channels at NLO with the analytic expression Eq. (3.28), Eq. (3.30) on the left and Eq. (3.28), Eq. (3.31) on the right, for a range of $y_{\text{cut}}$ values. The red points indicate a straight line at small $\rho$, this corresponds to a leading single logarithmic result for the integrated distribution. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < \frac{y_{\text{cut}}}{\rho_{\text{cut}} + y_{\text{cut}}} \Delta_R$, which demonstrates that we control all terms up to $O(\alpha_s^2)$ in the differential distribution. .... 108

3.8 Comparison of the differential distribution in $\rho$ for mMDT ($\mu = 0.67$) tagged jets using Monte Carlo (left) and full resummed analytical results (right). Details of MC generation are given in Fig. 3.1. This figure and caption has been adapted directly from Ref. [116]. .... 109

3.9 Comparison of Event2 at LO with the analytic expression Eq. (3.39) for a range of $z_{\text{cut}}$ values. The red points indicate a constant in the small $\rho$ limit, which corresponds to a leftover $O(\alpha_s)$ term in the differential distribution. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < z_{\text{cut}} \Delta_R$, which tend to zero for each $z_{\text{cut}}$ value. .... 113
3.10 Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.48) for a range of $z_{cut}$ values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut}\Delta_R$, which gives a straight line for small $\rho$, indicating we control all terms up to $O(\alpha_s^2 L)$ in the differential distribution.  

3.11 Comparison of EVENT2 in the $C_FC_A$ channel at NLO with the analytic expression Eq. (3.50) for a range of $z_{cut}$ values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut}\Delta_R$, which gives a straight line for small $\rho$, indicating we control all terms up to $O(\alpha_s^2 L)$ in the differential distribution.  

3.12 Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.57) for a range of $z_{cut}$ values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut}\Delta_R$, which gives a straight line for small $\rho$, indicating we control all terms up to $O(\alpha_s^2 L)$ in the differential distribution.  

3.13 Comparison of the differential distribution in $\rho$ for pruning and Y-pruning tagged jets using Monte Carlo (left) and full resummed analytical results (right). Details of MC generation are given in Fig. 3.1. The arrows denote the expected position of the transition points, as derived by the analytical calculations in this section, i.e. $\rho \simeq z_{cut}^2\Delta_R$ (black) and $\rho \simeq z_{cut}\Delta_R$ (grey). The I-pruning distribution corresponds to all contributions that do not satisfy the Y-pruning constraints, hence pruning is the sum of Y- and I-pruning. This figure and caption has been adapted directly from Ref. [116].  

3.14 Comparison of EVENT2 at LO with the analytic expression Eq. (3.67) in the region $\rho < f_{cut}\Delta_R$ for a range of $f_{cut}$ values. The red points demonstrate linear behaviour for small $\rho$, indicating the presence of double logarithms. This transitions into a flat, single logarithmic, region for intermediate jet mass. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < f_{cut}\Delta_R$, which tend to zero for small $\rho$. This demonstrates that we correctly control all logarithmically enhanced terms at LO.
3.15 Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.77) in the region $\rho < f_{\text{cut}} \Delta_{R_{\text{trim}}}$ for a range of $f_{\text{cut}}$ values. The red points indicate the presence of double logarithms. The green points demonstrate the difference between the analytic expression and EVENT2 result in the region $\rho < f_{\text{cut}} \Delta_{R_{\text{trim}}}$. The straight line at small $\rho$, corresponds to leftover single logarithm $O(\alpha_s^2 L)$ in the differential distribution. .......................................................... 131

3.16 Comparison of EVENT2 in the $C_F C_A$ channel (left) and $C_F n_f$ (right) with the analytic expression Eq. (3.78) in the region $\rho < f_{\text{cut}} \Delta_{R_{\text{trim}}}$ for two $f_{\text{cut}}$ values. The green points show the difference between the analytic expression and EVENT2, which is a straight line at small $\rho$, corresponding to a leftover single logarithm $O(\alpha_s^2 L)$ in the differential distribution. .......................................................... 132

3.17 Comparison of the differential distribution in $\rho$ for trimming tagged jets using Monte Carlo (left) and full resummed analytical results (right). Details of MC generation are given in Fig. 3.1. The arrows denote the expected position of the transition points, as derived by the analytical calculations in this section, i.e. $\rho \simeq f_{\text{cut}} \Delta_{R_{\text{trim}}}$ (black) and $\rho \simeq f_{\text{cut}} \Delta_{R}$ (grey). This figure and caption has been adapted directly from Ref. [116]. .......................................................... 133

4.1 Comparison of MC (left) and analytic (right) Eq. (4.12) tagging efficiencies for a range of mass windows as a function of a generator level cut on minimum jet transverse momentum $p_T$. This result has been generated using HERWIG++ 2.7.0 for $pp \rightarrow ZH$ at 14 TeV with the $Z$ decaying leptonically and $H \rightarrow b\bar{b}$, setting $M_H = 125$ GeV. We have tagged the signal jet as the highest $p_T$ Cambridge/Aachen jet with $R = 1$. In this figure we have generated events at parton level with ISR only and divided out the contribution due to the lowest order result in both panels for clarity. .......................................................... 147

4.2 An MC study of the impact of hadronisation and underlying event on the signal efficiency for a plain jet mass cut as a function of the minimum jet transverse momentum. One can see the sizeable impact of both hadronisation and especially underlying event on the signal efficiency for the window $\delta M = 16$ GeV. Details of the generation are as in Fig. 4.1 but now we also include FSR at parton level. .......................................................... 148
4.3 Comparison of MC (left) with the analytic result Eq. (4.16) (right) with \( R_{\text{trim}} = 0.3 \) for the tagging efficiency for two values of \( f_{\text{cut}} \) as a function of a minimum generator level cut on jet transverse momentum. This result has been generated using HERWIG++ 2.7.0 at parton level with no additional radiation for \( H \to bb \) jets. We note that the location of the transition points are reproduced by MC.

4.4 Comparison of MC (left) and analytic Eq. (4.18) (right) \( R_{\text{trim}} = 0.3 \) trimming \( f_{\text{cut}} \) tagging efficiencies for a range of \( f_{\text{cut}} \) values as a function of a generator level cut on jet transverse momentum. This result has been generated using HERWIG++ 2.7.0 [75] at parton level with ISR only for \( H \to bb \) jets, setting \( M_H = 125 \) GeV and \( \delta M = 16 \) GeV. The transition points correspond to the change from plain jet mass like behaviour to a \( \ln f_{\text{cut}} \) term, discussed in the main text.

4.5 A configuration that contributes to the FSR correction to trimming signal efficiency in the region \( R_{\text{trim}} \ll \theta_{b\bar{b}} \). Here gluon \( k \) is emitted collinear to the \( b \) quark with momentum fraction \( x \) outside a cone with radius \( R_{\text{trim}} \).

4.6 An MC study of the impact of hadronisation and underlying event (UE) on the signal efficiency using the trimmed jet \( f_{\text{cut}} = 0.1 \) as a function of the minimum jet transverse momentum for two different values of \( R_{\text{trim}} \). One can see the impact of hadronisation and underlying event on the signal efficiency in the window \( \delta M = 16 \) GeV. Details of generation given in Fig. 4.2.

4.7 An MC study of the impact of hadronisation and underlying event (UE) on the signal efficiency for pruning (left) and mMDT (right) \( (z_{\text{cut}}, y_{\text{cut}} = 0.1) \) as a function of jet transverse momentum with \( \delta M = 16 \) GeV. Details of generation are given in Fig. 4.2.

4.8 Comparison of MC (left) and analytic (right) \( Y \)-pruning signal tagging efficiencies for a range of \( z_{\text{cut}} \) values as a function of a generator level cut on jet transverse momentum. This result has been generated using HERWIG++ 2.7.0 [75] at parton level with ISR only for \( H \to bb \) jets, setting \( M_H = 125 \) GeV and \( \delta M = 16 \) GeV. We have divided out the contribution due to the born configuration in both panels for clarity.

4.9 An MC study of the impact of hadronisation and underlying event (UE) on the signal efficiency using the \( Y \)-pruned jet \( (z_{\text{cut}} = 0.1) \) as a function of the minimum jet transverse momentum. One can see negligible impact resulting from hadronisation corrections but some degradation coming from underlying event contamination on the signal efficiency in the window \( \delta M = 16 \) GeV. Details of generation given in Fig. 4.2.
4.10 A MC study of the impact of ISR on the signal efficiency for various tags, \( (y_{\text{cut}}, z_{\text{cut}}, f_{\text{cut}} = 0.1, R_{\text{trim}} = 0.3) \) as a function of the minimum jet transverse momentum. One can note the similarity of mMDT, pruning and trimming while Y-splitter and Y-pruning are different, with Y-splitter in particular virtually indistinguishable from the plain jet mass. Details of generation given in Fig. 4.1. . . . . . . . . . . . . . . . . . . 170

4.11 Signal efficiency for tagging hadronic \( H \) jets using \textsc{Herwig++} 2.7.0 with underlying event and hadronisation as a function of a generator level cut \( p_T \) on transverse jet momentum. . . . . . . . . . . . . . . . . . . 171

4.12 MC results for the differential jet mass distribution using Y-splitter with trimming \( (y_{\text{cut}} = 0.1, R_{\text{trim}} = 0.3) \), Y-splitter \( (y_{\text{cut}} = 0.1) \) and Y-pruning \( (z_{\text{cut}} = 0.1) \). This result has been generated using \textsc{Herwig++} 2.7.0 at parton level. A minimum \( p_T \) cut on generation of the hard process \( qq \rightarrow qq \) was made at 3 TeV for 14 TeV pp collisions. We take the mass of the two hardest jets found using the C/A algorithm with \( R = 1.0 \) and apply the algorithms with a fixed \( z_{\text{cut}}/y_{\text{cut}} \). . . . . . . . . . . . . . . . . . . . . . . . . . 172

4.13 Signal significance for tagging hadronic \( H \) jets with quark (left panel) and gluon (right panel) backgrounds using \textsc{Herwig++} 2.7.0 with underlying event and hadronisation as a function of a generator level cut \( p_T \) on transverse jet momentum. We compare the signal significance for different algorithms to Y-splitter+trimming and find that the latter outperforms the others at high \( p_T \). Here we have used \( R_{\text{trim}} = 0.1 \) for pure trimming because, in contrast to Y-splitter+trimming, we expect this value to be closer to optimal than \( R_{\text{trim}} = 0.3 \) at high transverse momenta (see Fig. 4.18) . . . . . . . . . . . . . . . . . . . . . . . . . . 173

4.14 Signal significance for tagging hadronic W jets with quark (left panel) and gluon (right panel) backgrounds using \textsc{Herwig++} 2.7.0 with underlying event and hadronisation as a function of a generator level cut \( p_T \) on transverse jet momentum. We deem a jet tagged if it has a final mass in the window 64 – 96 GeV. We compare the signal significance for different algorithms to Y-splitter+trimming and find that the latter outperforms the others at high \( p_T \). In this plot, we use tagger parameters that match those used for W tagging in Ref. [116] for ease of comparison. 174
4.15 mMDT analytical signal significance from tree level signal and resummed background as a function of $y_{\text{cut}}$ (top left) compared to HERWIG++ 2.7.0 at parton level (top right) and with hadronisation and MPI (bottom left). The signal process used is $pp \rightarrow ZH$ where we require the Higgs and Z to decay hadronically and leptonically respectively with quark backgrounds. We place a generator level cut on the Higgs transverse momentum $p_T$ of 1, 2 and 3 TeV. Jets are tagged around the Higgs mass with a mass window $\delta M = 16$ GeV. The bottom right panel shows the analytic optimal $y_{\text{cut}}$ values as a function of $p_T$ (red line) with a 2% variation in signal significance about the peak (pink shaded area). We overlay the optimal results for $y_{\text{cut}}$ obtained using HERWIG++ 2.7.0 with hadronisation and underlying event at 1, 2 and 3 TeV, with an equivalent 2% variation about the peak signal significance (blue bars) and at parton level (black bars).

4.16 Pruning analytical signal significance from tree level signal and resummed background as a function of $z_{\text{cut}}$ compared to HERWIG++ 2.7.0 at parton level and with hadronisation and MPI. Details of generation given in Fig. 4.15.

4.17 Y-pruning analytical signal significance from tree level signal and resummed background as a function of $z_{\text{cut}}$ compared to HERWIG++ 2.7.0 at parton level and with hadronisation and MPI. Details of generation given in Fig. 4.15.

4.18 2D density plot showing the trimming signal significance as a function of $R_{\text{trim}}$ and $f_{\text{cut}}$ using HERWIG++ 2.7.0 for $H \rightarrow bb$ jets with quark backgrounds with a minimum jet transverse momentum cut. The top panels are generated at parton level with transverse momenta 2 TeV and 3 TeV left and right and the bottom panels include hadronisation and underlying event. The area inside the black contour represents the FSR improved analytic prediction for optimal values within 2% of the analytic peak signal significance.

4.19 Trimming signal significance for $R_{\text{trim}} \sim \sqrt{\Delta} \sim 0.06$ as a function of $f_{\text{cut}}$ generated at parton level and with hadronisation and MPI using HERWIG++ 2.7.0.

5.1 Comparison of LO EVENT2 and Eq. (5.9) in the region $\rho < y_{\text{cut}}\Delta R/(1 + y_{\text{cut}})$ for a range of $y_{\text{cut}}$ values. We have plotted the coefficient of normalised differential jet mass distribution as a function of ln $\rho$ at LO. After subtracting the analytical result from the EVENT2 output, the result vanishes in the limit $\rho \rightarrow 0$. 

15
5.2 Comparison of NLO EVENT2 and Eq. (5.20) in the $C_F^2$ channel for a range of $y_{\text{cut}}$ values. After subtracting the analytical result from the EVENT2 output, we get a straight line as $\rho \to 0$, corresponding to residual single logarithmic terms ($\alpha_s^2 L$) in the differential distribution.

5.3 Comparison of NLO EVENT2 and Eq. (5.21) in the $C_F C_A$ (left) and $C_F n_f$ (right) channels for two different $y_{\text{cut}}$ values. After subtracting the analytical result from the EVENT2 output, we get a flat line as $\rho \to 0$, corresponding to residual constant terms ($\alpha_s^2$) in the differential distribution.

5.4 Lund Diagram for the Y-splitter algorithm applied to a jet with given normalised mass $\rho$. We define the normalised $k_t$ scale $y_{\text{scale}}^2 = y_{\text{cut}}$ as a horizontal line and write that the emission that dominates the jet mass lies on the contour of constant $\rho$. For the jet to be accepted, this emission must have a normalised $k_t$ scale greater than $y_{\text{cut}}$, (this is equivalent here to $x > y_{\text{cut}}$ because the emissions with the largest $k_t$ scale also sets the jet mass). Further emissions are not checked by Y-splitter, so we need to veto on the shaded area to give the correct Sudakov suppression term. Emissions in the unshaded region have a subleading impact on the jet mass, so we ignore these. Note that the x-axis begins at $\eta \simeq -\ln R/2$, which is the minimum pseudorapidity with respect to the jet axis for an emission to be within a jet of radius $R$.

5.5 Comparison of Monte Carlo (left) and analytic differential jet mass distributions (right) for Y-splitter. This result has been generated using HERWIG++ 2.7.0 \cite{75} at parton level without underlying event. A minimum $p_T$ cut on generation of the hard process $qq \to qq$ was made at 3 TeV for 14 TeV pp collisions. We take the mass of the two hardest jets found using the C/A algorithm with $R = 1.0$ and apply the Y-splitter algorithm with several values of $y_{\text{cut}}$. In the above plots, we indicate the expected transition points with dashed lines at $\rho = y_{\text{cut}} R^2$, at this point the distribution changes from the plain jet mass to the next-to-double distribution characteristic of the Y-splitter algorithm.

5.6 Lund Diagram for the anomalous contribution to the Y-splitter algorithm applied to a jet with given normalised mass $\rho$. We define $y_{\text{scale}}^2 = y_{\text{cut}}$ as a horizontal line and write that the emission that dominates the jet mass lies on the contour of constant $\rho$. The jet mass is defined by a soft ($x < y_{\text{cut}}$) emission but a secondary emission with larger $k_t$ scale passes the Y-splitter condition. For the jet to be accepted, the secondary emission must have an energy $x_2 > y_{\text{cut}} \rho / \rho_2$ and a mass contribution $\rho > \rho_2 > y_{\text{cut}} \rho$. 
5.7 Monte Carlo results for the differential jet mass distribution using Y-splitter followed by trimming compared to trimming followed by Y-splitter. We have used the parameters $f_{\text{cut}} = y_{\text{cut}} = 0.1, R_{\text{trim}} = 0.3$. Details of generation are given in Fig. 5.3. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
5.13 2D density plot showing the Y-splitter+trimming signal significance as a function of $R_{\text{trim}}$ and $y_{\text{cut}}$ using HERWIG++ 2.7.0. The signal process used is $pp \rightarrow ZH$ where we require the Higgs and Z boson to decay hadronically and leptonically respectively with quark backgrounds. We place a generator level cut on the Higgs transverse momentum. Jets are tagged around the Higgs mass with a mass window $M_H \pm \delta M$ where $\delta M = 16$ GeV. The top panels are generated at parton level with transverse momentum cuts 2 TeV and 3 TeV left and right respectively and the bottom panels include hadronisation and underlying event.

A.1 Ratio for signal efficiency normalised to lowest-order result, with EVENT2 and SHERPA 2.0.0 at parton level, for $e^+e^-$ annihilation with virtual Z production and hadronic decay, where we consider a Z boson with a transverse boost to $p_T = 3$ TeV.

A.2 Contour plot showing the maximum percentage difference in signal efficiency between EVENT2 at order $\alpha_s$ and SHERPA 2.0.0 final state shower at parton level both normalised to the lowest order result. In the left hand panel we apply pruning with different values for $\epsilon$ and $z_{\text{cut}}$ with $p_T = 3$ TeV. In the right hand panel we apply trimming with different values for $\sqrt{\Delta}$ and $R_{\text{trim}}$ with $f_{\text{cut}} = 0.1$.

C.1 Lund Diagram for an ungroomed jet with given normalised mass $\rho$. We define the jet radius cutoff in the collinear limit $\eta = -\ln R/2$, as a vertical line and the emission that dominates the jet mass lies on the contour of constant $\rho$, coloured red. The shaded region denotes the vetoed region for real emissions within the jet radius, which results in a Sudakov suppression factor.
LIST OF TABLES

1.1 Feynman rules for the propagators and vertices in QCD (excluding ghosts). Quarks are represented as solid lines, gluons are represented as curly lines. Arrows on fermion lines denote the direction of charge flow and the momentum is also defined in this direction. Note that the momenta in the three gluon vertex are defined as incoming, i.e. $p + q + r = 0$. .................................................. 40

3.1 A table that summarises the main features of each tagger. The first two columns show the leading logarithmic terms in the limit $\rho \rightarrow 0$ in the integrated jet mass distribution at leading and next-to leading order. We have also tabulated the transition points in $\rho$ and indicated whether non-global logarithms in $\rho$ are present. ............................................. 136

4.1 A table summarising the logarithmic structure of radiative corrections to the signal efficiency for each tagger. For each tagger we show the coefficient of $-\frac{\alpha_s C_F}{\pi}$ for ISR and FSR results in the small $\Delta$ and $z_{cut}/y_{cut}/f_{cut}$ limit. We have defined $\epsilon = 2\delta M/M_H$ as in the main text. The coefficient $C_2$ for the trimming FSR logarithm is given in Eq. (4.26). .................. 186
# LIST OF ALGORITHMS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusive sequential recombination</td>
<td>64</td>
</tr>
<tr>
<td>Exclusive sequential recombination modification</td>
<td>65</td>
</tr>
<tr>
<td>Mass drop tagger</td>
<td>73</td>
</tr>
<tr>
<td>Filtering</td>
<td>75</td>
</tr>
<tr>
<td>MDT modification (mMDT)</td>
<td>98</td>
</tr>
<tr>
<td>Pruning</td>
<td>111</td>
</tr>
<tr>
<td>Y-pruning modification</td>
<td>121</td>
</tr>
<tr>
<td>Trimming</td>
<td>126</td>
</tr>
<tr>
<td>Y-splitter tagger</td>
<td>192</td>
</tr>
</tbody>
</table>
ABSTRACT

Jet substructure: An analytical approach
A thesis submitted to The University of Manchester
for the degree of doctor of philosophy (PhD)
in the faculty of Engineering and Physical Sciences

Alexander Powling
18th September 2015

In the past few years, detailed study of the internal structure of hadronic jets has become an active area of research. In particular, jet substructure information has been used to distinguish signal from QCD initiated jets, which constitute a significant background for many physics searches. Prior to the work undertaken in this thesis, theoretical research on jet substructure was largely Monte Carlo based, with limited analytical input. This work presents an analytical approach to the study of jet substructure techniques in the context of high-$p_T$ heavy resonance searches at the LHC.

In this thesis, we compute the mass distribution of QCD initiated jets after application of several jet substructure algorithms using approximate fixed-order perturbative QCD at leading and next-to-leading order. This is sufficient to extract the leading logarithmic structure for each technique, which we compare to exact fixed-order results. Using this analytical insight, we propose modifications to some of these algorithms and use our results to discuss the phenomenological impact of different parameter choices.

We also perform analytical calculations and Monte Carlo studies to examine the impact of QCD radiation on jets that arise from boosted Higgs decay after application of several jet substructure algorithms. Understanding the action on signal jets is important when two techniques perform similarly on background jets. An example studied here is the Y-splitter and Y-pruning techniques, which both perform well at rejecting background; however, the former retains signal jets which are subject to significant radiative and non-perturbative corrections. We demonstrate that the combination of Y-splitter with trimming ameliorates the poor signal tagging efficiency of Y-splitter whilst retaining effective background rejection. Consequently, we find that this combination outperforms the other techniques studied here, at high $p_T$. We use our analytical expressions to perform an approximate optimisation of parameters for each algorithm and compare our results to Monte Carlo simulation.

Finally, we undertake an analytical fixed-order and resummed study of the mass distribution for QCD jets for the combination of Y-splitter with trimming. We demonstrate that the trimming has a numerically subleading effect on the Y-splitter distribution for typical parameter choices and discuss why such techniques can prove to be superior when compared to the currently proposed individual methods.
DECLARATION

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: [Signature]  Date: 18.09.15

Alexander Powling
COPYRIGHT STATEMENT

- The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

- Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

- The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

- Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
ACKNOWLEDGEMENTS

I would first like to thank my supervisor Mrinal Dasgupta for his advice and support throughout the duration of my PhD. His input ensured the success of the work and subsequent papers we published over the past four years. I also want to acknowledge my other collaborators: Simone Marzani, Alessandro Fregoso for teaching me a great deal in my first few years and Andrzej Siodmok whom constantly assured me that we were “almost finished” despite several setbacks and delays.

I would also like to thank the denizens of the HEP theory office, both past and present for many fruitful and often humorous discussions. These include Rene Angeles Martinez, Lee Tomlinson, Daniele Teresi and Sotirios Karamitsos. In particular, I greatly appreciate the proof reading of this thesis performed by Daniel Burns and Graeme Nail, which introduced me to a world of grammar pedantry I never knew existed. I would also like to thank the rest of the HEP group at Manchester for providing me with a great working environment and reminding me of the practical considerations associated with my research.

I would like to thank my parents, who provided encouragement throughout and financial aid during the final year. Finally, I want to dedicate this thesis to my wonderful girlfriend Sivanne, who moved to Manchester in order to be with me during my PhD. Her constant support, reassurance and help has been invaluable and I am looking forward to what the future brings for both of us.
LIST OF PUBLICATIONS

The following publications are based on material discussed in this thesis:


INTRODUCTION

The theoretical framework that describes our current understanding of fundamental particles that exist in nature and their interaction at infinitesimal scales via the electromagnetic, weak and strong forces is known as the standard model (SM) of particle physics. As its name suggests, the standard model as formulated in the late 1970’s, was highly successful in describing data from early particle physics experiments. The SM also predicted additional unobserved particle content, for example, the W and Z bosons were discovered in 1983 [1–4], the top quark in 1995 [5] and most recently the Higgs boson in 2012 [6,7]. These discoveries, as well as precision measurements of many SM predictions have provided extensive verification and demonstrated the robustness of this theory.

Despite the success of the standard model, it is understood that it does not provide a theory for all fundamental interactions. Notable absences in the SM include a consistent theory of gravitational interactions, dark matter candidates [8], matter anti-matter asymmetry [9] and neutrino oscillations [10]. Additionally, by computing radiative corrections to the Higgs mass using the standard model framework, one finds large quantum corrections that must be reconciled with the observed mass at 125 GeV [11]. These arguments motivate the formulation of additional beyond standard model (BSM) theories to provide some improvements. Some BSM theories, notably several supersymmetric models [12], predict fundamental particles with masses around the TeV scale. Hence, it becomes of importance to begin to experimentally verify which (if any) BSM model is realised in nature by observing these particles in experiments located at the large hadron collider (LHC) [13].

Heavy particles, BSM or otherwise, will have short lifetimes and tend to decay before they reach the detectors surrounding the interaction point at a particle collider, hence one must instead consider the decay products of these particles. Given that these decay products are sufficiently long lived and interact strongly with the matter within the particle detector, one can use these final-state particles to reconstruct features of the heavy particle resonance. However, for a typical proton-proton collision at LHC, such a “signal” will appear alongside large numbers of SM “background” processes.
These may have very similar signatures in a detector; therefore, in order to distinguish the signal, one must precisely understand the features of the radiation originating from both the background and signal processes.

In terms of extracting signal event candidates, a particularly challenging set of decay products for a signal process are those produced via strong interactions. These often create large multiplicities of particles known as hadrons, which are often observed as collimated groups called jets. The number and rate of SM processes associated with this class of decay products is extremely large in proton-proton colliders. Hence, in this hadronic channel, one must devise a way of enhancing these heavy particle processes relative to the SM background, in order to decrease the amount of statistics required for a discovery. One approach is to examine jets with large transverse momentum with respect to the beam axis; in this regime, most of the decay products of the heavy resonance will be contained within a single jet. By examining this subset of the particle activity in a given event and using the constituent hadron substructure associated with the jet, one can employ algorithmic techniques to distinguish between signal and background candidates.

This thesis focusses on the search for heavy resonances at the LHC and some of the algorithmic jet substructure techniques used to distinguish signal from background initiated decay products in the high transverse momentum regime. In Chapter 1 we begin by giving an overview of some quantum field theoretic features of strong interactions using quantum chromodynamics (QCD). Using these features, we examine the production of jets in Chapter 2 and some of algorithms used to classify them in a collider environment. We also demonstrate the substructure differences between heavy resonance and background initiated jets and show how one can discriminate between the two using jet substructure algorithms. Additionally, we outline some of the theoretical aspects of observables based on measurements of jets. In Chapter 3 we analytically calculate the jet mass observable for background jets after application of a range of jet substructure algorithms and demonstrate how each technique determines the analytic structure of this observable. The material presented in this chapter is based on Ref. [14]. In Chapter 4 we examine the jet mass observable for signal jets after application of several jet substructure techniques. After examining the analytic structure of the jet mass in this context, we identify a promising combination of complementary substructure techniques that outperforms the others studied here, at high transverse momentum. Ref. [15] is based on the material presented in this chapter. Given the observed effectiveness of this combination, in Chapter 5 we embark on one of the first analytical studies of a combination of jet substructure algorithms. The material presented in this chapter forms the basis of an upcoming publication. Finally, we summarise the work and the conclusions of this thesis and discuss prospects for further investigation.
CHAPTER ONE

QCD PHENOMENOLOGY

1.1 Introduction

In the 1950’s and 60’s, significant effort was made to identify and classify the myriad of hadrons, hadronic mass spectra and hadronic interactions created in early particle scattering experiments at Brookhaven \[16\] and the Bevatron at Berkeley \[17\]. Organisation of hadrons into flavour octets and decuplets \[18\] strongly suggested that hadrons were made up of fundamental building blocks named quarks \[19,20\]. However, the quark model of hadrons was not able to explain the apparent violation of Fermi-Dirac statistics observed in some spin-3/2 baryons such as $\Delta^{++}$ (containing three u quarks) \[21\], as well as discrepancies between theoretical predictions and experimental measurements of the total cross-section for $e^+e^- \rightarrow \text{hadrons}$ \[22\]. As a solution, it was proposed that quarks must have a new quantum number named colour \[23\]. We now understand that colour is a manifestation of the SU(3) symmetry group that forms the basis of a non-Abelian gauge field theory \[24\]. It was shown that quantum field theories with this symmetry can satisfy the requirement of asymptotic freedom \[25–27\] and were renormalisable \[28\]. Thus, a consistent theory of quark dynamics was established in the early 1970’s and named *quantum chromodynamics* (QCD) \[26,29\].

QCD is a theory describing strong interactions between coloured particles, mediated by a non-Abelian gauge field named the gluon. A key difference, in contrast to the photon in *quantum electrodynamics* (QED), is that the gluon is itself coloured, leading to self interactions of gauge bosons that is crucial for asymptotic freedom of the theory at high energies. This means that the interaction strength between coloured particles becomes increasingly weak at short distances. Therefore, at high energies we may use a perturbative approach to quantum field theory when describing coloured reactions. This is known as perturbative quantum chromodynamics (pQCD).

This chapter is outlined as follows: in Section 1.2 we examine how the free quark
Lagrangian transforms under gauge transformations of the quark field and introduce the gluon gauge field to ensure invariance under this transformation. We then write down the QCD Lagrangian density and discuss some important results from the SU($N_c$) symmetry group that will be used throughout this thesis in Section 1.3. In Section 1.4, we present the Feynman rules for QCD, which are required to build up invariant squared matrix elements at each order in perturbation theory. We then discuss the renormalisation procedure applied to QCD in Section 1.5 and derive the running of the strong coupling, $\alpha_s$, which exhibits the property of asymptotic freedom. In Section 1.6, we consider hadronic interactions and discuss the factorisation of initial-state collinear QCD emissions into parton distribution functions. In Section 1.7, we explicitly compute the cancellation of infrared divergences for QCD emissions in the final state and derive the angular ordering property of soft emissions in Section 1.8. Finally in Section 1.9, we discuss how to build up large multiplicities of collinear final-state emissions via an iterative process called the parton shower.

1.2 Gauge invariance and the QCD Lagrangian

Let us now consider the Dirac Lagrangian density for non-interacting fermion (quark) fields [30]:

$$\mathcal{L}_{\text{Dirac}} = \sum_f \bar{\psi}_f (i\gamma^\sigma \partial_\sigma - m_f) \psi_f,$$

(1.1)

where $\psi_f, \bar{\psi}_f$ are the fermion, anti-fermion fields respectively with flavour $f$, mass $m_f$ and we have summed over all flavours. Dirac gamma matrices are written as $\gamma^\alpha$ where $\alpha$ is a Lorentz vector index; these matrices satisfy the Clifford algebra relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{I}$, where $\mathbb{I}$ is the $4 \times 4$ identity matrix and the metric is defined $g^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. In order to construct a locally invariant theory for QCD, we require that the Lagrangian density in Eq. (1.1) is invariant under gauge transformations of the form [31]

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x),$$

(1.2)

where the local rotation matrix $U(x)$, at space-time coordinate $x$, is defined by the special unitary group, SU($N_c$), where $N_c$ is the number of quark colours. This rotation matrix can be written as the sum:

$$U(x) = e^{i\theta_a(x)t^a},$$

(1.3)

where $\theta_a(x)$ represents a local (position dependent) angle of rotation and $t^a$ are the generators of the SU($N_c$) group in the fundamental representation indexed by parameter
a = \{1, \ldots, N_c^2 - 1\} and are traceless hermitian matrices.

In order to obtain a gauge invariant Lagrangian density, we must introduce an additional vector field, known as the gluon, to the theory. We achieve this gauge invariance by replacing the partial derivative \( \partial_\sigma \) in Eq. (1.1) with a covariant derivative \( D_\sigma \), defined as

\[
D_\sigma = \partial_\sigma \mathbb{I} - ig_s t^a A^a_\sigma, \tag{1.4}
\]

where \( g_s \) is the strong gauge coupling constant for QCD interactions and \( A^a_\sigma \) is a component of the gluon field. The integer \( a \) is a colour index \( \{1, \ldots, N_c^2 - 1\} \) of the gauge boson field. The covariant derivative must transform as \( D_\sigma \to D'_\sigma = U(x)D_\sigma U^{-1}(x) \) to preserve gauge invariance, which implies that the gluon field must transform as:

\[
A_\sigma \to A'_\sigma = U(x) \left[ A_\sigma + \frac{i}{g_s} \partial_\sigma \right] U^{-1}(x), \tag{1.5}
\]

where we define \( A_\sigma \equiv t^a A^a_\sigma \). The presence of the derivative term in the transformation Eq. (1.5) implies that the inclusion of a gluon mass term proportional to \( A^a_\sigma A^a_\sigma \) would break gauge invariance, hence gluons are massless.

Using these results, the QCD Lagrangian density can be written as

\[
\mathcal{L}_{\text{QCD}} = \sum_f^{n_f} \bar{q}_f (i \gamma^\sigma D_\sigma - m_f) q_f - \frac{1}{4} F_{a \mu \nu} F^{a \mu \nu}, \tag{1.6}
\]

where \( q_f, \bar{q}_f \) are explicitly the quark and anti-quark fields. \( i, j \) are colour indices for \( N_c \) quark colours and we have summed over quark flavours, \( f \), up to the total number of flavours \( n_f \). The first term represents the propagation of free quarks and their interactions with gluons. The second term in Eq. (1.6) is a kinetic term for the gluon fields, which gives rise to propagation and self interaction. This term satisfies gauge invariance and preserves renormalisability of the theory. The non-Abelian field strength tensor \( F_{a \mu \nu} \), is defined by the commutation relation \([D_\mu, D_\nu] = -ig_s t^a F^{a \mu \nu} \), which gives

\[
F_{a \mu \nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_s f^{abc} A^b_\mu A^c_\nu, \tag{1.7}
\]

where \( f^{abc} \) are the SU\((N_c)\) structure constants\(^1\). One should note that the last term in Eq. (1.7) is a direct result of the non-Abelian nature of the theory and gives rise to gluon self-interactions.

In order to construct a perturbative description of QCD, we require the existence of a gluon propagator\(^2\). Defining this propagator as an inverse of the operator associated

---

\(^1\)For an Abelian theory, such as QED, the structure constants are all equal to zero.

\(^2\)The Feynman rule associated with this propagator is defined explicitly in Table 1.1.
with terms bilinear in the action, we first must eliminate additional gauge degrees of freedom from the QCD Lagrangian. This is achieved via an additional gauge fixing term (this modification is also present in Abelian gauge theories), the form of which is given by

\[ \mathcal{L}_{G,F} = -\frac{1}{2\lambda} (\partial^\mu A^a_\mu)^2, \]  

(1.8)

where \( \lambda \) is a gauge fixing parameter. However, due to the non-Abelian nature of QCD, this term must be accompanied by a Faddeev-Popov “ghost” term [34]. This acts by removing all remaining longitudinal degrees of polarisation from the gluon propagator. By introducing this correction, one introduces a complex-scalar “ghost” field, \( \chi_a \), which only couples to gluons inside loop diagrams. With exception to a brief mention in the context of renormalisation of QCD in Section 1.5, we do not consider the action of ghost fields further in this thesis.

1.3 \text{SU}(N_c)

In this section, we examine some important features of the symmetry group \text{SU}(N_c) in the context of QCD and derive some relevant results. The generators of \text{SU}(N_c) form a Lie algebra that obey the commutation relation and Jacobi identity [35]:

\[ [t^a, t^b] = if^{abc} t^c, \]  

(1.9)

\[ [t^a, [t^b, t^c]] + [t^b, [t^c, t^a]] + [t^c, [t^a, t^b]] = 0, \]  

(1.10)

where the generators \( t^a \) are expressed in the fundamental representation as \( N_c \times N_c \) matrices\(^3\) \( f^{abc} \) are structure constants that are totally antisymmetric under interchange of indices. We can define additional matrices from the structure constants themselves \( T^a \), where \( (T^a)^{bc} \equiv if^{abc} \). These \( (N_c^2 - 1) \times (N_c^2 - 1) \) matrices also satisfy the commutation relations in Eqs. (1.9,1.10), and form the adjoint representation. In both representations, the number of generators is equal to the number of degrees of freedom of the \text{SU}(N_c) group, which is equal to \( N_c^2 - 1 \) (see Eq. (1.3)).

It is conventional to normalise the \text{SU}(N_c) matrices in the following way:

\[ \text{Tr} (T^a_R T^b_R) = \mathcal{N}_R \delta^{ab}, \]  

(1.11)

where \( T_R \) are the generators of \text{SU}(N_c) in the representation \( R \) and \( \mathcal{N}_R \) is a constant that depends on the representation. In the fundamental representation \( \mathcal{N}_F = \frac{1}{2} \) and adjoint \( \mathcal{N}_A = N_c \). For each representation, we construct an operator that is the sum of the squares of generators, named the Casimir. The Casimir commutes with each

\(^3\)In \text{SU}(3), these generators are given by \( t^a = \frac{1}{2} \lambda^a \), where \( \lambda^a \) are the Gell-Mann matrices, the form of which can be found in e.g. Ref. [32].
1.4. FEYNMAN RULES

generator and is therefore proportional to the identity via Schur’s first lemma \[35\]:

\[
\sum_{a,c} (T_R^a)^{bc} (T_R^a)^{cd} = C_R \delta^{bd}
\]  

(1.12)

where \(C_R\) is a constant that depends on the representation. Combining Eq. (1.11) and Eq. (1.12) we obtain the Casimirs for SU\((N_c)\):

\[
C_F = \frac{N_c^2 - 1}{2N_c},
\]

\[
C_A = N_c,
\]

which we call colour factors in the fundamental and adjoint representation respectively. Henceforth, we adopt the commonly used notation for the normalisation constant in the fundamental representation \(T_R \equiv N_F = \frac{1}{2}\) and we set \(N_c = 3\), to give the SU\((3)\) colour factors:

\[
C_F = \frac{4}{3},
\]

\[
C_A = 3.
\]  

(1.14)

1.4 Feynman rules

We now want to compute particle scattering amplitudes by evolving an initial state \(|i\rangle\), with definite momenta at time \(-t\) to time \(t\) using the Hamiltonian of the system. We can calculate the overlap of the initial state with a defined final state \(|f\rangle\) and write as an element of a unitary operator called the S-matrix, \(S_{fi}\), defined in the limit \(t \to \infty\).

This is given by the time ordered product of the exponentiated interaction part of the (QCD) Lagrangian density integrated over all spacetime \[30,36\]:

\[
S_{fi} = \langle f | T \left\{ \exp \left[ i \int d^4x L_{\text{int}}(x) \right] \right\} | i \rangle,
\]  

(1.15)

where \(L_{\text{int}}\) is the interaction part of the Lagrangian density after performing the decomposition into a combination of “free” and interacting terms: \(L = L_0 + L_{\text{int}}\). \(T\{\ldots\}\) indicates a time ordered product, such that operators contained within act in chronological order on the initial state.

It is not known (currently) how to fully calculate \(S_{fi}\) in the exponentiated form given in Eq. (1.15). However, one solution is to perturbatively expand the exponential in a region where the exponent is small\[4\]. This allows computation of the S-matrix as

\[\text{For QCD, we expand the coupling constant, } g_s, \text{ around zero. This expansion is convergent in the high energy limit due to the observed phenomena of asymptotic freedom, as we will demonstrate in Subsection 1.5.1.} \]
\[ \delta_{ij} \frac{i (p + m)}{p^2 - m^2 + i\epsilon} \]

\[ g^{ab} \frac{i}{(p^2 + i\epsilon)} \left[ -g_{\mu\nu} + (1 - \lambda) \frac{p_{\mu}p_{\nu}}{(p^2 + i\epsilon)} \right] \]

\[ -ig_{s}\gamma^{\mu}t_{ij}^{a} \]

\[ -g_{s}f_{abc}^{a}(p - q)^{\sigma}g^{\mu\nu} \]

\[ + (q - r)^{\mu}g^{\nu\sigma} + (r - p)^{\nu}g^{\sigma\mu} \]

\[ -ig_{s}^{2}\left[ f^{x ba}f^{x cd}(g_{\mu x}g_{\nu\sigma} - g_{\nu e}g_{\sigma\mu}) \right. \]

\[ + f^{x bd}f^{x ca}(g_{\mu x}g_{\nu\sigma} - g_{\nu e}g_{\sigma\mu}) \]

\[ + f^{x bc}f^{x ad}(g_{\nu x}g_{\mu\sigma} - g_{\nu e}g_{\sigma\mu}) \]

Table 1.1: Feynman rules for the propagators and vertices in QCD (excluding ghosts). Quarks are represented as solid lines, gluons are represented as curly lines. Arrows on fermion lines denote the direction of charge flow and the momentum is also defined in this direction. Note that the momenta in the three gluon vertex are defined as incoming, i.e. \( p + q + r = 0 \).
successive corrections to the non-interaction amplitude, $\delta f_i$:

$$S_{fi} = \delta f_i + i \int d^4 x \langle f | \mathcal{L}_{\text{int}}(x) | i \rangle - \frac{1}{2} \int d^4 x d^4 y \langle f | T \{ \mathcal{L}_{\text{int}}(x) \mathcal{L}_{\text{int}}(y) \} | i \rangle + \ldots .$$

(1.16)

Defining a transition matrix, $T$, such that $S = I + iT$, to remove the trivial non-interaction term, one can write

$$iT_{fi} = (2\pi)^4 \delta^{(4)} \left( \sum p_{\text{in}} - \sum p_{\text{out}} \right) iM_{fi},$$

(1.17)

where $p_{\text{in(out)}}$ is the 4-momenta of an incoming (outgoing) stable particle and we have factored out the 4-momentum conservation constraint via a 4-dimensional delta function. The remaining terms define $M_{fi}$, which is the invariant matrix element for a given transition $i \rightarrow f$.

We use the interaction terms (those proportional to $g_s$) in the QCD Lagrangian density, Eq. (1.6), to derive a set of Feynman rules that can be used to write down $M_{fi}$ at each order in perturbation theory. This is achieved by visualising the scattering process as a set of vertices, internal propagators and external particles, via Feynman diagrams. We show the relevant Feynman diagrams and rules in Table 1.1 for QCD

1.5 Renormalisation and asymptotic freedom

Renormalisation is a procedure in which one relates the unmeasurable, “bare”, parameters of a theory to observable, renormalised, parameters. In a general sense, this is achieved by factorising all short-distance (UV) behaviour of the theory above a given scale into redefinitions of the parameters in the Lagrangian density. Perturbative calculations of QCD observables in general contain divergences in the UV limit; therefore it is essential to perform renormalisation in order to make theoretical predictions. QCD is a renormalisable theory, which explicitly means that UV divergences present at all orders of perturbation theory can be absorbed by suitable redefinitions of a finite number of fields, masses and coupling constants (for an overview of renormalisation in the context of QCD see Refs. [33,37,38]).

In the QCD Lagrangian, we have a number of bare fields $\psi_f, A_\mu^a, \chi_a$ and bare parameters $g_s, m_f, \lambda$ that we redefine to give the renormalised quantities:

$$\psi_f = \sqrt{Z_2} \psi_{f,R}, \quad g_s = Z_g g_{s,R},$$

$$A_\mu^a = \sqrt{Z_3} A_{\mu,R}^a, \quad m_f = Z_{m,f} m_{f,R},$$

$$\chi_a = \sqrt{Z_3} \chi_{a,R}, \quad \lambda = Z_3 \lambda_R.$$
where subscript $R$ denotes a renormalised parameter and $Z_i$ are the renormalisation constants associated with them. We can write $Z_i = 1 - \delta Z_i$ where $\delta Z_i$ is a counterterm introduced to be equal and opposite to all UV divergent terms, calculated to all orders in perturbation theory. In order to generate the counterterms, one has to first regularise the divergent limits of the loop integrals. This regularisation procedure introduces a scale dependence to the finite terms generated alongside the divergent structures. Additionally, one also has an arbitrary choice concerning which scale independent constants to include in the counterterm; this introduces further renormalisation scheme dependence into the renormalisation procedure.

In order to make theoretical predictions for observable quantities, we require input from experimental measurements at a given energy scale to provide a fixed value for the renormalised parameters. From these measurements, we can use the Callan-Symanzik equation [39,40] to make predictions about the value of each renormalised parameter at one scale when compared to another. In the context of QCD, we can use this equation to compute the running of the strong coupling with renormalisation scale.

### 1.5.1 The running coupling constant

As discussed in the previous subsection, we now calculate explicitly the scale dependence of the renormalised strong coupling constant, $g_{s,R}$. One can use dimensional regularization [41] to parametrise the divergences in Feynman diagram calculations at each order in perturbation theory. The loop integrals are regularised by evaluating them in $d = 4 - 2\epsilon$ dimensions such that the result is finite for an infinitesimal $\epsilon$ (see, for example Refs. [36,38,42,43] for an overview of these techniques). By absorbing all terms which are divergent in the limit $\epsilon \to 0$ into the counterterms defined by $Z_i$, one can write the bare coupling, $g_s$, in terms of a dimensionful renormalisation scale, $\mu$. Explicitly, this can be written as

$$g_s = \mu^\epsilon Z_g(g_{s,R}(\mu)) g_{s,R}(\mu),$$  

(1.18)

where the $\mu^\epsilon$ term is a direct result of the dimensional regularization procedure and is introduced to ensure that the coupling remains dimensionless. For brevity, we will suppress the explicit dependence of each parameter on $\mu$ henceforth. The renormalisation constant $Z_g$ can be written in the minimal subtraction scheme (MS) [44,45] as a sum of divergent terms:

$$Z_g = 1 + \sum_{n=1}^{\infty} \frac{A_n}{\epsilon^n},$$  

(1.19)

\footnote{This renormalisation scheme defines the counterterms such that they only contain the divergent, pole terms.}

42
where $A_n$ are finite coefficients of poles of order $\epsilon^n$. We define the beta function, $\beta (g_{s,R})$, which describes how the renormalised strong coupling varies with $\ln (\mu)$, as

$$\beta (g_{s,R}) \equiv \mu \frac{\partial g_{s,R}}{\partial \mu}.$$  \hfill (1.20)

All bare quantities are independent of the renormalisation procedure, hence can not depend on the scale $\mu$. Differentiating Eq. (1.18) with respect to $\mu$ and insisting $\mu \frac{\partial g_s}{\partial \mu} = 0$, one can write the beta function in terms of $Z_g$ [43]:

$$\beta (g_{s,R}) = -\epsilon g_{s,R} - g_{s,R} \frac{\mu}{Z_g} \frac{\partial Z_g}{\partial \mu}.$$  \hfill (1.21)

If we substitute Eq. (1.19) into the beta function expression Eq. (1.21) and impose the constraint that $\beta (g_{s,R})$ is finite, one can show that the beta function depends only on the single pole coefficient, $A_1$, in the expansion of $Z_g$ [46]. This gives an equivalent expression:

$$\beta (g_{s,R}) = g_{s,R}^2 \frac{\partial A_1}{\partial g_{s,R}}.$$  \hfill (1.22)

$A_1$ has contributions from each order in perturbation theory of the form, $A_1 = 1 + A_1^{(1)} g_{s,R}^2 + A_1^{(2)} g_{s,R}^4 + ...$, where the superscript denotes the loop order. We can write this perturbative expansion in the usual notation [32]:

$$\beta (g_{s,R}) = -b'_0 g_{s,R}^3 + O (g_{s,R}^5),$$  \hfill (1.23)

where $b'_0$ is the finite 1-loop coefficient of the beta function, which can be identified as $b'_0 = -A_1^{(1)}/2$. To obtain this coefficient for QCD, we use the 1-loop expression for $Z_g$ defined in MS scheme, which can be obtained from a sum of the counterterms derived from the Feynman diagrams in Fig. 1.1 [33,38,43]:

$$Z_g^{(1)} = 1 - \frac{g_{s,R}^2}{(4\pi)^2} \frac{1}{\epsilon} \left[ \frac{1}{6} (11C_A - 4T_R n_f) \right],$$  \hfill (1.24)

to give the lowest order result:

$$b'_0 = \frac{1}{(4\pi)^2} \left[ \frac{1}{3} (11C_A - 4T_R n_f) \right].$$  \hfill (1.25)

We insert Eq. (1.25) into Eq. (1.23), and use the definition Eq. (1.20) to integrate between two arbitrary scales $\mu_0$ and $\mu_1$. This gives the following the 1-loop running
coupling relation:

$$\alpha_s(\mu_1^2) = \frac{\alpha_s(\mu_0^2)}{1 + \beta_0 \alpha_s(\mu_0^2) \ln \frac{\mu_1^2}{\mu_0^2}}, \quad (1.26)$$

where we have defined $\beta_0 = 4\pi b'_0$ and $\alpha_s(\mu^2) = g^2_{s,R}(\mu)/(4\pi)$. Typically a value of $\alpha_s$ will be fixed at a particular scale by an experimental measurement (commonly at $\mu_0 = M_Z$) and Eq. (1.26) can be used to evolve the strong coupling to a different scale $\mu_1$.

We observe that the number of active quark flavours $n_f$, is always less than the constant value $11C_A/(4T_R) \simeq 17$, hence $\beta_0$ is positive. This means that, in the limit $\mu_1 \gg \mu_0$, the value of the strong coupling approaches zero. Therefore, the strength associated with QCD interactions diminishes as one increases the energy scale. This phenomena is known as asymptotic freedom [25-27] and facilitates the use of perturbation theory for QCD in the high energy regime. Conversely, in the region $\mu_1 < \mu_0$, there exists a scale, $\Lambda_{QCD}$, whereby the coupling $\alpha_s$ diverges. At 1-loop, this scale is the value of $\mu_1$ at which the denominator of Eq. (1.26) is equal to zero. As we approach this scale, we expect a breakdown of the validity of pQCD, whereby the perturbative series becomes increasingly non-convergent. One can use Eq. (1.26) to estimate the scale at which this transition from perturbative to non-perturbative behaviour occurs. However, this determination is not strictly well defined because we are using a truncated perturbative expression for the running of $\alpha_s$ in a region that
1.6 The parton model and factorisation

At the large hadron collider (LHC), the colliding particles are hadrons (proton-proton). When these hadrons are probed at scales much larger than \( \Lambda_{\text{QCD}} \), the parton model describes high energy hadrons as a set of point-like, weakly interacting, coloured partons (quarks and gluons). Hence, one can think of high-energy hadronic interactions occurring between the constituent partons of each hadron, rather than the hadrons themselves. Quantum mechanically, this can be explained by noting that the internal dynamics of the hadron and the hard scattering process occur at disparate distance scales. Therefore, the two processes are approximately incoherent and can be factorised into a product of long-distance, non-perturbative, distributions and short-range, perturbative, partonic subprocesses.

The distribution of parton momenta within a hadron is an inherently non-perturbative quantity and can not (currently) be calculated theoretically from first principles; however, these distributions can be obtained experimentally via the measurement of universal parton distribution functions (PDFs), \( f_i(x_i) \). These functions parametrise the probability of instantaneously finding a parton \( i \) with longitudinal momentum fraction \( x_i \) in hadron of type \( f \).

Convolving PDFs with a perturbative, partonic cross-section, we can write the lowest-order (tree-level), total cross-section for a hard scattering process between two hadrons in a factorised form:

\[
\sigma(P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1) f_j(x_2) \hat{\sigma}_{ij}(x_1 P_1, x_2 P_2, \mu), \tag{1.27}
\]

where we have denoted the momenta of the incoming hadrons \( P_1 \) and \( P_2 \), neglected \( \mathcal{O}(1/Q^2) \) corrections, where \( Q \) is a hard scale of the process, and summed over all possible parton types \( i, j \) and partonic momentum fractions \( x_1, x_2 \). We have defined the cross-section of a tree-level hard perturbative subprocess as \( \hat{\sigma}_{ij} \), which depends on the UV renormalisation scale, \( \mu \) (see Fig. 1.2a).
CHAPTER 1. QCD PHENOMENOLOGY

Figure 1.2: A schematic representation of a hard scattering process in the parton model description. In both panels, a pair of partons from two hadrons (circles) participate in a partonic hard scattering (square) with a given longitudinal momentum fraction $x_i$. The left panel shows the tree-level picture. The right panel is identical to the left, but with an additional QCD correction corresponding to an emission in the initial state; note that this emission changes the momentum fraction of the parton obtained from the hadron. The additional parallel lines emerging from the hadrons represent the remnants of the hadron after scattering.

However, this “naive” parton model has divergent contributions in the massless limit when we compute the first QCD correction to the total cross-section (see e.g Refs. [32, 36, 53] and references therein). One shows this result by considering the tree-level process with a real or virtual gluon emitted by a parton (either a quark or gluon) prior to, or after, the hard interaction. Performing the perturbative calculation for real emission using the Feynman rules in Table 1.1, one finds that divergent structures emerge in the soft and/or collinear limits (we demonstrate this explicitly in the final state for $e^+e^-$ annihilation in Section 1.7). Additionally, this divergent behaviour does not cancel in the hard-collinear limit for real and virtual initial-state emissions. The basic interpretation of this result is that the emission of a real hard-collinear gluon from an initial state leg, changes the momentum of the parton obtained from the PDF. We must integrate over the full phase space of the additional emission, which can be recast as an integral over all possible initial momentum fractions $z$, of the parton from the PDF for a given momentum fraction entering the hard process $x_i$, (see Fig. 1.2b). The divergent behaviour for a real emission contributes to the $x_i$ integral, for all $z > x_i$, whereas the virtual piece only contributes at $z = x_i$. Therefore, these
divergences only fully cancel in the soft limit $z = x_i$, leaving a real, hard-collinear divergence behind. One can interpret this partial miscancellation as a manifestation of over-counting between the PDFs for a given $x_i$ (which should contain all the collinear dynamics of the hadron) and collinear perturbative evolution via emission of hard radiation.

One can solve this over-counting by introducing a scale that divides initial-state emissions into two classes: those that can be considered part of the hard perturbative evolution and those that are part of the non-perturbative parton distribution functions. We start by noting that collinear emission is long-range with respect to the hard process by virtue of the vanishingly small virtuality of the parent propagator in this limit and is therefore non-perturbative. Taking inspiration from renormalisation in the UV limit, discussed in Section 1.5, we absorb these remaining non-perturbative collinear singularities (to all orders) into redefinitions of the PDFs. In doing so, one introduces a scale dependence to the new, renormalised PDFs, which we call the factorisation scale, $\mu_F$. This scale denotes an artificial threshold at which one separates dynamics which are attributed to the hard process from the internal dynamics included in the PDF. We write this factorisation of the collinear 1-loop correction to the hard process as a sum of a renormalised PDF and a hard correction term in a minimal subtraction scheme as

$$f^F_i(x, \mu_F) = f^R_i(x, \mu_F) + \sum_j \int_x^1 \frac{dz}{z} f^R_j(z, \mu_F) \frac{\alpha_s}{2\pi} \left( P_{ij}^{(0)} \left( \frac{x}{z} \right) \ln \frac{Q^2}{\mu_F^2} + C_j \left( \frac{x}{z} \right) \right)$$

$$+ \mathcal{O} (\alpha_s^2),$$

where the sum over $j$ denotes all possible emissions from parton $i$ (e.g. if $i = g$ then $j = q, \bar{q}, g$), $C_j$ are (finite) process-dependent coefficient functions and the superscript $F$ denotes a factorised quantity. Note that the $\mu_F$ dependence of $f^F_i(x, \mu_F)$ enters at the $\alpha_s^2$ level and higher because we have only included the 1-loop divergent terms in the redefinition. The renormalised PDF $f^R_i(x, \mu_F)$, is defined as the sum of an unobservable bare PDF $f_i$, and a collinear pole regulated by an infrared (IR) cutoff, $\kappa$:

$$f^R_i(x, \mu_F) = f_i(x) + \sum_j \int_x^1 \frac{dz}{z} f_j(z) \frac{\alpha_s}{2\pi} P_{ij}^{(0)} \left( \frac{x}{z} \right) \ln \frac{\mu_F^2}{\kappa^2} + \mathcal{O} (\alpha_s^2).$$

The coefficients of the collinear singularities in Eqs. (1.28,1.29) are given by the Altarelli-Parisi (AP) splitting functions. We are working at $\mathcal{O} (\alpha_s)$, hence $P_{ij}^{(0)}$
are the regularised, massless, tree-level AP splitting functions\(^6\) which are defined as:

\[
P^{(0)}(z)_{qq} = C_F \left[ \frac{1 + z^2}{(1 - z)_+} + \frac{3}{2} \delta (1 - z) \right],
\]

\[
P^{(0)}(z)_{gg} = T_R \left[ z^2 + (1 - z)^2 \right],
\]

\[
P^{(0)}(z)_{gq} = C_F \left[ 1 + (1 - z)^2 \right],
\]

\[
P^{(0)}(z)_{gq} = 2C_A \left[ \frac{z}{(1 - z)_+} + \frac{1 - z}{z} + z (1 - z) \right] + \delta (1 - z) \left( \frac{11C_A - 4TRn_f}{6} \right),
\]

where “+” denotes the plus distribution:

\[
\int_0^1 \frac{dz}{(1 - z)_+} = \int_0^1 \frac{dz}{1 - z},
\]

where \(g(z)\) is a smooth distribution that is finite at \(z = 1\). This prescription acts by cancelling the real divergences in the soft limit \(z \to 1\) by including the equal and opposite virtual contributions at the point \(z = 1\). The integrals over \(z\) in Eqs. (1.28,1.29) are therefore finite in the interval \(x\) to 1.

We can rewrite the tree-level cross-section in Eq. (1.27) in a form that does not contain collinear divergences at 1-loop using this factorised parton model:

\[
\sigma (P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 f_i^F (x_1, \mu_F) f_j^F (x_2, \mu_F) \hat{\sigma}_{ij} (x_1 P_1, x_2 P_2, \mu)
\]

\[
= \sum_{i,j} \int dx_1 dx_2 f_i^R (x_1, \mu_F) f_j^R (x_2, \mu_F) \hat{\sigma}_{ij}^{(1)} (x_1 P_1, x_2 P_2, \mu, \mu_F),
\]

where the first line is obtained by replacing the bare PDFs in Eq. (1.27) with the factorised expression Eq. (1.28). In the second line, we have identified the second term in Eq. (1.28) as a single emission correction to the partonic subprocess and redefined \(\hat{\sigma}_{ij}\) as \(\hat{\sigma}_{ij}^{(1)}\). Note that the renormalised PDF and the hard process with a single initial-state emission are now functions of the factorisation scale, \(\mu_F\). In order to minimise the numerical impact of this scale in fixed order calculations, we note that for each order in perturbation theory, the factorisation scale cutoff introduces leading terms of the form \(\alpha_s^n \ln^n Q^2 / \mu_F^2\) (see Eq. (1.28)). Therefore, given that one is free to choose \(\mu_F\), a natural choice is \(\mu_F \sim Q\) such that these logarithmic contributions do not spoil fixed order convergence of the perturbation series.

---

\(^6\)An additional splitting function for \(g \to \bar{q}q\) is given simply as \(P^{(0)}_{\bar{q}g} (z) = P^{(0)}_{qg} (z)\) by symmetry.
\section*{1.6.1 DGLAP evolution}

Observables such as the total cross-section must be independent of the (arbitrary) factorisation scale when computed to all orders in perturbation theory, this implies that

\[ \mu_F \frac{\partial \sigma}{\partial \mu_F} = 0. \tag{1.33} \]

However, because we are only using \( \mathcal{O}(\alpha_s) \) counterterms for the factorised PDFs shown in Eq. (1.29), the \( \mu_F \) dependence of the cross-section in Eq. (1.32) begins at \( \mathcal{O}(\alpha_s^2) \). Nevertheless, we can compute the perturbative evolution of the PDF at 1-loop between two factorisation scales given they are both much larger than \( \Lambda_{\text{QCD}} \). If we insert the first line of Eq. (1.32) into Eq. (1.33), one notes that each \( f_i^F(x, \mu_F) \), must be independent of the factorisation scale at \( \mathcal{O}(\alpha_s) \). Therefore, differentiating the 1-loop expression given in Eq. (1.28) with respect to \( \mu_F \) gives

\[ \mu_F \frac{\partial f_i^R(x, \mu_F)}{\partial \mu_F} = \frac{\alpha_s}{2\pi} \sum_j \int_x^1 \frac{dz}{z} P_{ij}^{(0)} \left( \frac{x}{z} \right) f_j^R(z, \mu_F) + \mathcal{O}(\alpha_s^2), \tag{1.34} \]

which is the well known \textit{Dokshitzer-Gribov-Lipatov-Altarelli-Parisi} (DGLAP) evolution equation at 1-loop \cite{54,57}. The DGLAP equation is analogous to the beta function for the running coupling equation Eq. (1.20) and describes how the PDF evolves perturbatively with the factorisation scale \( \mu_F \) for a given momentum fraction \( x \). This can be generalised to higher orders by computing higher order corrections to the running of \( \alpha_s \) and loop corrections to the AP splitting functions \cite{58}. The form of Eq. (1.34) indicates that the DGLAP equation is a \((2n_f+1)\)-dimensional matrix equation for the scale evolution of quark, anti-quark and gluon PDFs. Unlike the beta function for the running coupling, this coupled integro-differential equation can not be solved easily, even at 1-loop order. Hence, the running of the PDF from one scale to another is often performed numerically \cite{59,61}.

\section*{1.7 Infrared divergences of QCD in the final state}

We have dealt with UV divergences present in QCD via the renormalisation process, redefining parameters of the Lagrangian, such as the coupling constant, to include divergent perturbative corrections. By considering radiative corrections of partons that originated in hadrons, we encountered additional collinear divergences that we factorised into a redefinition of universal parton distribution functions associated with the non-perturbative structure of the colliding hadron. We now want to examine the perturbative structure of QCD emissions from outgoing partons, which we anticipate to also contain divergences in the soft and/or collinear limits.
1.7.1 Real contribution

Consider an emission of a real gluon with momentum $k$, polarisation $\lambda$ and colour $c$ from a $q\bar{q}$ pair created by a colour neutral decay in Fig. 1.3. The amplitude for this process can be derived from the Feynman rules given in Table 1.1:

$$M_1 = \bar{u}_a(p_1, s_1) \left( -ig_s t^c_{ab} \epsilon^*(k, \lambda) \right) \times \left[ \gamma^\mu \frac{i(p_1 + \not{k} + m_1)}{(p_1 + k)^2 - m_1^2 + i\varepsilon} V + V \frac{-i(p_2 + \not{k} - m_2)}{(p_2 + k)^2 - m_2^2 + i\varepsilon} \gamma^\mu \right] v_b(p_2, s_2), \quad (1.35)$$

where $\bar{u}$ and $v$ are spinors corresponding to a fermion and anti-fermion respectively in the final state. $\epsilon^*$ is the final-state gluon polarisation vector and the quark and anti-quark have masses $m_1$ and $m_2$ respectively. The 4-momenta are labelled with $p_i$, with spin $s_i$ ($i = \pm 1$) and quark colours by $a, b$. The object $V$ contains all the information about the colour neutral vertex and the process that created the final-state $q\bar{q}$ pair, which we leave unspecified. The factor $i\varepsilon$ acts by displacing the on-shell poles of the propagators from the real axis, however for this calculation they are not required so we drop them henceforth.

We are interested in the divergent structure of this amplitude, hence we examine the soft limit of Eq. (1.35), which corresponds to terms that are singular as $k \to 0$. To find this limit, one scales the 4-vector associated with soft emission $k^\mu \to \lambda k^\mu$ in the amplitude and only retain the leading terms as $\lambda \to 0$. Neglecting all subleading terms in this limit (that are associated with recoil) is known as the eikonal approximation \[62\]. Additionally, we work in the high energy limit whereby the momenta of the quarks are much greater than their masses; in this region we make the approximation $m_i \simeq 0$. 

Figure 1.3: Feynman diagrams showing the production of a final state quark-antiquark pair with momentum $p_1$ and $p_2$ respectively via an arbitrary colour neutral decay, in that one leg emits a real gluon with momentum $k$. 


This gives the massless eikonal approximation for the amplitude:

\[ \mathcal{M}_1 \simeq g_s \bar{u}_a(p_1, s) t^c_{ab} \epsilon^c_\mu(k, \lambda) \left[ \gamma^\mu \frac{p_1}{2p_1 \cdot k} V - \frac{p_2}{2p_2 \cdot k} \gamma^\mu \right] v_b(p_2, s') \]

\[ = g_s t^c_{ab} \epsilon^c_\mu(k, \lambda) \left[ \frac{p_1^\mu}{p_1 \cdot k} - \frac{p_2^\mu}{p_2 \cdot k} \right] \bar{u}_a(p_1, s) V v_b(p_2, s') \]

\[ = \epsilon^c_\mu(k, \lambda) J^{\mu,c} M_0, \] \hspace{1cm} (1.36)

where we have used the Clifford algebra relation \( \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}I \) and the result \( \bar{u}(p) \bar{p} = p v(p) = 0 \) in the second line. In the third line, we have written this process in a factorised form at amplitude level \([63]\) of a soft current term \( J^{\mu,c} = g_s t^c_{ab} \left[ \frac{p_1^\mu}{p_1 \cdot k} - \frac{p_2^\mu}{p_2 \cdot k} \right] \)

and the amplitude without the gluon emission \( M_0 = \bar{u}_a(p_1, s) V v_b(p_2, s') \), known as the Born amplitude. We square \( \mathcal{M}_1 \) in Eq. (1.36) and sum over final-state spins, flavours, colours and polarisations to get

\[ |M_1|^2 \equiv \sum \mathcal{M}_1 \mathcal{M}_1^\dagger = J^{\mu,c} J^{\nu,d} M_0 | ^2 \sum_\lambda \epsilon^c_\mu(k) \epsilon^d_\nu(k), \] \hspace{1cm} (1.37)

where \( |M_0|^2 \equiv \sum M_0 M_0^\dagger \) and we perform the polarisation sum in the light-cone gauge \([64]\)

\[ \sum_\lambda \epsilon^c_\mu(k, \lambda) \epsilon^d_\nu(k, \lambda) = \delta^{cd} \left( -g^\mu_\nu + \frac{k^\mu n_\nu + k^\nu n_\mu}{k \cdot n} \right), \] \hspace{1cm} (1.38)

where \( n \) is a light-like vector that satisfies \( k \cdot n \neq 0 \). However, we note that the soft current \( J^{\mu,c} \) satisfies the constraint\([7]\) \( k_\mu J^{\mu,c} = 0 \), therefore the second term in Eq. (1.38) does not contribute to the square amplitude. This gives

\[ |M_1|^2 = -J^{\mu,c} J^{\nu,d} M_0 | ^2 \]

\[ = g_s^2 C_F \left[ \frac{2p_1 \cdot p_2}{(p_1 \cdot k) (p_2 \cdot k)} \right] |M_0|^2, \] \hspace{1cm} (1.39)

where we have summed over colour generators in the fundamental representation via the identity in Eq. (1.12).

One notices that this square amplitude is divergent in three regions of phase space:

1. The gluon is collinear to the quark: \( p_1 \cdot k \to 0 \),

2. The gluon is collinear to the anti-quark: \( p_2 \cdot k \to 0 \),

3. The gluon momentum approaches zero: \( k \to 0 \).

\[ ^7 \text{This is a direct result of colour conservation and can be generalised to an arbitrary number of coloured final state legs} \] \hspace{1cm} (65, 67)
To show this explicitly for an observable quantity, we write down the one real emission correction to the leading-order pair-production process by integrating over the phase space of the quarks and gluon using the squared amplitude given in Eq. (1.39):

$$\sigma_{q\bar{q}g} = g_s^2 C_F \int \frac{d^3k}{(2\pi)^3} \frac{2p_1 \cdot p_2}{2k_0 (p_1 \cdot k) (p_2 \cdot k)} d\sigma_B,$$

where $d\sigma_B$ is the differential Born cross-section for creating the quark anti-quark pair. We specify back-to-back kinematics for the quark-antiquark pair each with energy $Q/2$:

$$p_1 = \frac{Q}{2} (1, 0, 0, 1),$$
$$p_2 = \frac{Q}{2} (1, 0, 0, -1),$$
$$k = k_0 (1, \sin \theta, 0, \cos \theta),$$

where $\theta$ is the angle between the emitted gluon and the quark. We have neglected recoil in the soft and/or collinear limit for the kinematics defined in Eq. (1.41). Writing the integral over the gluon momenta in terms of these kinematics, we obtain:

$$\sigma_{q\bar{q}g} = 4 \frac{\alpha_s}{2\pi} C_F \int \frac{dk_0}{k_0} \frac{d\cos \theta}{(1 - \cos \theta) (1 + \cos \theta)} \frac{d\phi}{2\pi} d\sigma_B.\tag{1.42}$$

When we integrate over the full phase space of the gluon we find that the real $O(\alpha_s)$ contribution results in a divergent observable as we approach the collinear ($\theta \to 0, \pi$) and/or soft ($k_0 \to 0$) limit. However, we still have one more diagram that can contribute at this order, the 1-loop virtual correction to the Born amplitude.

### 1.7.2 Virtual contribution

We consider the emission of a virtual gluon that is emitted between two final state legs in Fig. 1.4. This configuration constitutes the only non-zero contribution to the 1-loop correction to the Born amplitude in the massless limit and is denoted by the superscript (1). This amplitude in the massless limit is given by

$$\mathcal{M}_{0}^{(1)} = \int \frac{d^4k}{(2\pi)^4} \bar{u}_a(p_1, s_1) (-ig_s t_{ab})$$
$$\times \gamma^\mu \frac{i(p_1 + k)}{(p_1 + k)^2 + i\varepsilon} V \frac{-i(p_2 - k)}{(p_2 - k)^2 + i\varepsilon} (-ig_s t_{ba}) \gamma^\nu v_b(p_2, s_2) \frac{-ig_{\mu\nu}}{k^2 + i\varepsilon},\tag{1.43}$$

where we explicitly have an integral over the loop 4-momentum, $k$. We again take the soft limit in the numerator, but retain $k^2$ terms in the denominator because the gluon
Figure 1.4: Feynman diagram showing the production of a final state quark-antiquark pair with momentum $p_1$ and $p_2$ respectively via an arbitrary colour neutral decay, in that a virtual gluon is transferred with momentum $k$. 

is off-shell:

\[ M_0^{(1)} \simeq -ig_s^2 C_F \int \frac{d^4k}{(2\pi)^4} \bar{u}_a(p_1, s_1) \frac{2p_1^\mu}{k^2 + 2p_1 \cdot k + i\varepsilon} V \frac{-2p_{2,\mu}}{k^2 - 2p_2 \cdot k + i\varepsilon} u_b(p_2, s_2) \frac{1}{k^2 + i\varepsilon} \]

\[ = 4ig_s^2 C_F \left[ \int \frac{d^4k}{(2\pi)^4} \frac{p_1 \cdot p_2}{[k^2 + 2p_1 \cdot k + i\varepsilon][k^2 - 2p_2 \cdot k + i\varepsilon][k^2 + i\varepsilon]} \right] M_0 \tag{1.44} \]

where we have a written a factorised form of the gluon loop, at amplitude level, from the Born amplitude. We again take the leading terms as $\lambda \to 0$ in the soft limit $k^\mu \to \lambda k^\mu$ to simplify the denominator:

\[ M_0^{(1)} \simeq -ig_s^2 C_F \left[ \int \frac{d^4k}{(2\pi)^4} \frac{p_1 \cdot p_2}{(p_1 \cdot k + i\varepsilon)(p_2 \cdot k - i\varepsilon)[k^2 + i\varepsilon]} \right] M_0 \tag{1.45} \]

The leading singularity as $\lambda \to 0$ is obtained in the limit that the gluon approaches the on-shell condition $k^2 \to 0$, given by the factor coming from the gluon propagator. Inspired by the form of Eq. (1.40), we will now evaluate the poles of the gluon propagator via the decomposition:

\[ \int dk_0 \frac{1}{[k^2 + i\varepsilon]} = \int dk_0 \frac{1}{(k_0 + |k| - i\varepsilon)(k_0 - |k| + i\varepsilon)}. \tag{1.46} \]

Choosing to integrate anticlockwise around the $k_0 = -|k| + i\varepsilon$ pole in the complex plane and using Cauchy’s residue theorem, we get

\[ \int dk_0 \frac{1}{[k^2 + i\varepsilon]} = 2\pi i \times \frac{1}{-2|k|}. \tag{1.47} \]

\[ ^8 \text{The contour integral of Eq. (1.45) contains an additional pole that emerges when each quark propagator satisfies the on-shell condition. This contribution is purely imaginary and corresponds to a Coulomb exchange (see e.g. Ref. [68]), these virtual interactions do not contribute at the level of precision required in this thesis and are neglected henceforth.} \]
which gives an expression for the 1-loop Born amplitude:

$$\mathcal{M}_0^{(1)} \simeq -g_s^2 C_F \left[ \int \frac{d^3 k}{(2\pi)^3 2k_0} \frac{p_1 \cdot p_2}{(p_1 \cdot k) (p_2 \cdot k)} \right] \mathcal{M}_0.$$  \hspace{1cm} (1.48)

To generate the $\mathcal{O}(\alpha_s)$ squared amplitude, we multiply Eq. (1.48) and its complex conjugate by the Born amplitude:

$$\sum \left( \mathcal{M}_0^{(1)} \mathcal{M}_0^\dagger + \mathcal{M}_0^{(1)} \mathcal{M}_0 \right) = -g_s^2 C_F \left[ \int \frac{d^3 k}{(2\pi)^3 2k_0} \frac{2p_1 \cdot p_2}{(p_1 \cdot k) (p_2 \cdot k)} \right] |\mathcal{M}_0|^2.$$ \hspace{1cm} (1.49)

As anticipated, this expression has the same divergent structure as the squared amplitude in the real emission case, as shown in Eq. (1.39), but with a relative minus sign. This indicates that the sum of real and virtual square amplitudes at $\mathcal{O}(\alpha_s)$ is finite due to complete cancellation of equal and opposite divergences in the soft and/or collinear limit. Therefore, the virtual contribution at $\mathcal{O}(\alpha_s)$ to the leading-order process is given by

$$\sigma_{q\bar{q}(g)} = -g_s^2 C_F \int \frac{d^3 k}{(2\pi)^3 2k_0} \frac{2p_1 \cdot p_2}{(p_1 \cdot k) (p_2 \cdot k)} d\sigma_B,$$ \hspace{1cm} (1.50)

which gives the full $\mathcal{O}(\alpha_s)$ correction to the leading-order pair-production process in the eikonal limit:

$$\sigma_{q\bar{g}} + \sigma_{q\bar{q}(g)} = 0.$$ \hspace{1cm} (1.51)

This observable receives no corrections at $\mathcal{O}(\alpha_s)$ from eikonal emission due to a complete cancellation of the singular and finite terms. This result is the 1-loop consequence of the Kinoshita-Lee-Nauenberg theorem [69,70], which states that a theory consisting of massless fields is free of infrared divergences at each order in perturbation theory if we carry out summation over the initial and final degenerate states. Degenerate in this context implies that a final state consisting of a massless quark and an arbitrary number of collinear and/or soft gluons, is indistinguishable from the quark alone.

We have shown that for a fully inclusive observable (one that integrates over all particle phase space), the $\mathcal{O}(\alpha_s)$ contribution from the addition of a real eikonal gluon is completely cancelled by the corresponding virtual diagram. However, if one defines an infrared and collinear safe observable in which the real emissions are integrated over a limited region of phase space, the corresponding cancellation of non-singular terms is incomplete, giving rise to logarithmic corrections. We will examine the consequences

---

9In general, one can also compute terms previously discarded in the soft and/or collinear limit; the full result is finite but non-zero. For example, the corresponding result for the process $e^+e^- \rightarrow q\bar{q}$ is

$$\sigma_{q\bar{q}} + \sigma_{q\bar{q}(g)} = \frac{\alpha_s}{\pi} \frac{2\sigma_B}{\pi}$$ \hspace{1cm} [32]
of this miscancellation in Chapter 2 in the context of jet observables.

## 1.8 Angular ordering

We now want to derive the phenomenon of angular ordering of soft emissions from pairs of hard coloured partons.\(^{10}\) We showed in Eq. (1.39) that terms associated with the emission of a real eikonal gluon from a quark pair factorise from the Born process at the square amplitude level. One can show that an eikonal splitting of a gluon produces an identical expression for the square matrix element given in Eq. (1.39), but with a different colour factor \(C_A\) or \(T_R\), depending on whether the splitting is into gluons or quarks respectively.\(^{53}\) Consequently, we can write down a general differential cross-section \(d\sigma_{n+1}\) for a process with \(n\) external hard partons plus one additional eikonal emission in terms of \(d\sigma_n\) (see Eq. (1.40) for \(n = 2\)) as\(^{32}\)

\[
d\sigma_{n+1} = d\sigma_n \frac{dk_0}{k_0} \frac{d\Omega}{2\pi} \frac{\alpha_s}{2\pi} \sum_{ij} C_{ij} W_{ij},
\]

(1.52)

where \(d\Omega\) is the solid angle element for the emitted parton with 4-momentum \(k\) and we have summed over all possible pairs of \(n\) coloured hard partons labelled \(i, j\). \(C_{ij}\) is the relevant colour factor associated with coloured pair \(i, j\) and \(W_{ij}\) is an angular radiation function, given by

\[
W_{ij} = k_0^2 \frac{p_i \cdot p_j}{(p_i \cdot k)(p_i \cdot k)} = \frac{1 - \cos \theta_{ij}}{(1 - \cos \theta_{ik})(1 - \cos \theta_{jk})},
\]

(1.53)

where \(\theta_{ab}\) is defined as the angle between parton \(a\) and parton \(b\).

One can rewrite this function as a sum of two terms, each containing a single collinear singularity by adding and subtracting pole-like counterterms to give

\[
W_{ij} = W_{ij}^{(i)} + W_{ij}^{(j)},
\]

(1.54)

where

\[
W_{ij}^{(i)} = \frac{1}{2} \left( W_{ij} + \frac{1}{1 - \cos \theta_{ik}} - \frac{1}{1 - \cos \theta_{jk}} \right),
\]

(1.55)

and \(W_{ij}^{(j)}\) is obtained by switching \(i \leftrightarrow j\). We substitute the sum Eq. (1.54) into Eq. (1.52) and perform the azimuthal integration of the term \(W_{ij}^{(i)}\) around parton direction \(i\), by writing the solid angle element as \(d\Omega = d\phi_i d\cos \theta_{ik}\). Similarly, we separately perform the azimuthal integration around direction \(j\) for the term \(W_{ij}^{(j)}\) and

\(^{10}\) This property is a feature of all gauge theories, in QED it is called the Chudakov effect.\(^{71}\)
sum the terms to give:

\[ d\sigma_{n+1} \simeq d\sigma_n \frac{dk_0}{k_0} \frac{\alpha_s}{2\pi} \sum_{i \neq j} C_{ij} \left[ \frac{d \cos \theta_{ik}}{1 - \cos \theta_{ik}} \Theta(\theta_{ij} - \theta_{ik}) + \frac{d \cos \theta_{jk}}{1 - \cos \theta_{jk}} \Theta(\theta_{ij} - \theta_{jk}) \right], \]

(1.56)

where one can see from the Heaviside step functions \( \Theta(x) \), that eikonal radiation pattern from each hard leg \( i \) and \( j \), is confined to a cone with angle less than the pairwise opening angle, \( \theta_{ij} \). This property is known as angular ordering and is a direct result of colour coherence (see Ref. [72] and references therein). If we explicitly consider an emission \( k \) produced from a coloured pair at an angle much larger than \( \theta_{ij} \), the probability of emission is proportional to the total colour charge of the pair. This is because the gluon does not have sufficient resolving power to differentiate the hard partons, so it interacts with the coherent sum of the individual partons \( i \) and \( j \). Hence, in the wide angle limit, one may consider that the emission \( k \) is instead emitted with a probability proportional to the colour charge of the parent, rather than one of the daughter partons in a given branching [32].

One should note that the angular ordering property in Eq. (1.56) is only approximate after separation of the total phase space into two “independent” parton directions \( i \) and \( j \). One should perform the azimuthal integration over the entire function \( W_{ij} \) for a single (arbitrary) direction. This gives a non-zero probability of a soft emission far outside the angular ordered regime, but these contributions are suppressed relative to emissions close to the hard pair. Specifically, the angular ordered emissions are distributed as \( d\theta^2 / \theta^2 \) (see Eq. (1.56)) whereas wide, non-angular ordered emissions are distributed as \( d\theta^2 / \theta^4 \) in the small opening angle limit, where \( \theta \simeq \theta_{ik}, \theta_{jk} \) is the angle between the emission and the coloured pair.

So far, we have considered a single eikonal emission from a coloured QCD pair produced in a hard process. In doing so, we found that we can factorise the emission phase space as a process-independent angular-ordered sum over all coloured hard pairs. One can exploit this simple structure to iteratively generate an arbitrary number of final state emissions given an initial set of hard coloured particles. This process is known as the parton shower.

1.9 The parton shower

We have shown that soft-collinear emissions factorise from the hard process, however one finds that hard-collinear emissions also factorise and can therefore be incorporated into a parton shower. This is achieved by taking a single leg in Eq. (1.56) and replacing the emission energy factor \( 1/k_0 \) with the real emission part of the appropriate azimuthally-averaged tree-level AP splitting function \( p_{ij}(z) \) [32,73]. This gives the
following differential distribution for a collinear splitting of a parton $j \rightarrow i + k$:

$$d\sigma_{n+1} = d\sigma_n \frac{\alpha_s \pi}{\xi} C_{ij} p_{ij}(z) \, dz,$$  

(1.57)

where $C_{ij}$ is the appropriate colour factor for the branching (see Eq. (1.30)), $z$ is the energy fraction $k_0/E_j$ and $\xi$ is an ordering variable proportional\(^{11}\) to $1 - \cos \theta_{jk}$. We require that subsequent emissions are strongly ordered in this variable, so that each emission can be treated as an independent process. To leading logarithmic accuracy, one is free to parametrise the collinear limit with any suitable ordering variable\(^{[73]}\); specifically, one can order emissions in virtuality (as implemented by PYTHIA 6\(^{[74]}\)), emission angle (HERWIG++\(^{[75]}\)), or transverse momentum (SHERPA\(^{[76]}\), PYTHIA 8\(^{[77]}\)).

The unregularised, reduced splitting functions, $p_{ij}(z)$, used in Eq. (1.57) are obtained from the regularised splitting functions Eq. (1.30) by removing terms associated with virtual corrections and stripping the colour factors\(^{12}\):

\[
\begin{align*}
p_{qq}(z) &= \frac{1 + z^2}{2 (1 - z)}, \\
p_{qg}(z) &= \frac{z^2 + (1 - z)^2}{2}, \\
p_{gq}(z) &= \frac{1 + (1 - z)^2}{2z}, \\
p_{gg}(z) &= \frac{z}{(1 - z)} + \frac{1 - z}{z} + z (1 - z). 
\end{align*}
\]

(1.58)

Using Eq. (1.58) in the expression in Eq. (1.57), we now notice that any integral over the $n + 1$ differential cross-section would contain divergences in the ordering variable $\xi$ and potentially the energy fraction $z$. A physically motivated solution, and one that can be implemented easily in a numerical algorithm, is introduce a resolution cut-off $Q_0$. For example, one can impose that a resolvable branching must have some minimum relative transverse momentum with respect to the parent parton. This ensures that arbitrarily soft and/or collinear emissions are not generated, and can be thought of as a non-perturbative scale whereby the parton shower terminates and hadronisation begins.

Using a single cutoff $z (1 - z) \xi > Q_0$ to regulate the integral, where $Q_0$ is a small parameter, we use Eq. (1.57) to write down the probability of parton $j$ branching in

\(^{11}\)We remove the negative sign introduced by this change of variables, by inverting the order of integration such that the result is positive.

\(^{12}\)The reader will notice that a factor of $1/2$ has been moved into the definition of the splitting functions. This is of course arbitrary but is performed for consistency with the remainder of this thesis.
the interval \( \xi \) and \( \xi + d\xi \) as

\[
P(\xi) = \frac{1}{\xi} \int_{Q_0/\xi}^{1-Q_0/\xi} \frac{\alpha_s}{\pi} C_{ij} p_{ij}(z) \, dz.
\] (1.59)

One can use this expression to write a differential equation for the probability of parton \( j \) having no resolvable branching \( \Delta_j \), above the cutoff, as \[74\]

\[
d\Delta_j(\xi_H, \xi) = \Delta_j(\xi_H, \xi) \, P(\xi),
\] (1.60)

where the rate of change of \( \Delta_j \) at \( \xi \) is equal to the probability of a decay at \( \xi \) given that no emissions occurred between \( \xi_H \) and \( \xi \), where \( \xi_H \) is interpreted as a maximum starting scale corresponding to the hard process. Solving Eq. (1.60) gives the exponential function

\[
\Delta_j(\xi_H, \xi) = \exp\left[ - \int_{\xi}^{\xi_H} \frac{d\xi'}{\xi'} \int_{Q_0/\xi'}^{1-Q_0/\xi'} \frac{\alpha_s}{\pi} C_{ij} p_{ij}(z) \, dz \right],
\] (1.61)

which is known as the Sudakov form factor \[32\]. This factor sums virtual contributions and unresolvable real emissions below cutoff \( Q_0 \), to all orders.

A numerical implementation of the parton shower, implemented in Monte Carlo generators, uses a random number to determine the scale at which a QCD emission occurs and then uses this new scale as an iterative starting point for further emissions. Specifically, we employ a Sudakov form factor, Eq. (1.61), as the probability of no emission between two scales \( \xi_H \) and \( \xi \). We then solve the equation \( \Delta_j(\xi_H, \xi) = R \) for scale \( \xi \), where \( R \) is a random number between 0 and 1. If \( \xi \) is greater than a resolution cutoff \( Q_0 \), we generate a branching at the scale \( \xi \) and set \( \xi_H = \xi \), otherwise the shower terminates. This Markovian algorithm implicitly includes the all-orders resummation of real and virtual emissions below \( Q_0 \) via the Sudakov form factor, hence correctly includes all the leading logarithms in \( Q_0 \). For this reason, this algorithm is called the leading collinear logarithmic parton shower algorithm (see Refs. \[32,73\] and references therein).

To see the leading, double logarithmic structure of the Sudakov form factor, we can evaluate Eq. (1.61) in the soft limit for no resolvable gluon emissions between a hard scale \( Q \) and the resolution limit, for fixed \( \alpha_s \), as \[13\]

\[
\Delta(Q, Q_0) \sim \exp\left[ - \frac{\alpha_s}{2\pi} C_I \ln^2 \frac{Q}{Q_0} \right].
\] (1.62)

\[13\]The strong coupling \( \alpha_s \), is usually evaluated at the transverse momentum of the emission with respect to the parent emitter \[78\]. This means that the coupling is generally a function of \( z \) and \( \xi \); this point is discussed further in Subsection 2.4.3.
where we have retained only leading, double logarithmic terms in the small $Q_0$ limit and $C_I = C_F, C_A$ for an initial quark or gluon respectively. If we perturbatively expand Eq. (1.62), we notice the appearance of a tower of logarithms at each order in $\alpha_s$ associated with non-cancellation of finite real and virtual contributions between the two scales $Q$ and $Q_0$. We will explore these logarithmic contributions in the next chapter in the context of jet observables.
CHAPTER TWO

JET PHENOMENOLOGY

2.1 Introduction

A typical hard event in the LHC may be divided into several, approximately factorisable stages, which evolve through a range of energy scales. As discussed in Section 1.6, this factorisation results from the quantum incoherence of dynamical processes that occur at disparate scales and is complete up to power corrections in the ratio of these scales. In the context of a hard event, these corrections are typically of the form \((\Lambda_{\text{QCD}}/Q)^p\), where \(Q\) is the hard scale of the process and \(p\) is a positive integer [53]. Hence, when \(Q \gg \Lambda_{\text{QCD}}\), we can effectively factorise the internal dynamics of the colliding protons in the initial state, which are characterised by small momentum transfers \(\mathcal{O}(\Lambda_{\text{QCD}})\), from the hard process. This factorisation of the hard, partonic interaction from the non-perturbative parton distribution functions was shown explicitly in Eq. (1.32). The hard interaction is characterised by a short range, large momentum transfer process; in this stage, we can use the perturbative techniques outlined in Section 1.4 to calculate transition probabilities. In the final stage, partons from the hard interaction are observed as colour-neutral hadronic particles due to colour confinement. The formation and subsequent evolution of these hadrons occurs at scales \(\mathcal{O}(\Lambda_{\text{QCD}})\); hence, this non-perturbative evolution can also be effectively factorised from the hard process [53].

Considering just the hard interaction stage of the event for now, we discuss some perturbative simplifications used in the calculation of multi-scale observables that are sensitive to multiple soft and/or collinear emissions of coloured partons. In Section 1.7 we demonstrated that a single-scale observable like the total cross section was completely insensitive to soft and/or collinear emission of gluons. Hence, one can effectively compute this observable using pQCD using the real and virtual contributions at each order in perturbation theory and one finds that each higher order term is suppressed...
by an additional factor $\alpha_s$. However, when computing multi-scale observables that constrain the phase space of real emissions, one finds additional logarithmic terms at each order of the perturbative expansion. Explicitly, for an observable with a second scale $\lambda < Q$ we find that the perturbative expansion can contain terms as singular as $\alpha_s^n \ln^{2n}(\lambda/Q)$ (see e.g. Eq. (1.62)). Depending on the magnitude of $\lambda/Q$, these terms may degrade (or even spoil) the convergence of a perturbative calculation (this is discussed extensively in Section 2.4). Even for a convergent perturbative series, the presence of such terms requires calculation to higher orders in $\alpha_s$ compared to single-scale observable to obtain similar perturbative accuracy. This calculation quickly becomes cumbersome beyond the first few orders in $\alpha_s$; this is, in part, due to the difficulty of evaluating higher order virtual loop diagrams whilst ensuring cancellation of infrared divergences associated with QCD radiation present at each order. Instead, for calculation of these observables and Monte Carlo simulation, we adopt the simplified picture that the hard interaction only occurs between a perturbatively calculable, small number of hard “external” partons. These partons can subsequently radiate additional soft and/or collinear partons, which in turn branch via further radiation. As discussed in Section 1.9 this perturbative process is known as the parton shower and this procedure builds up a large multiplicity of final-state particles, thereby approximating the soft and/or collinear contributions to all-orders, which result from the hard interaction.

First, we consider the initial-state radiation (ISR) shower, which proceeds by evolving the relevant partonic constituent of each incoming proton between a factorisation scale and a hard interaction scale via the emission of perturbative radiation. The initial distribution of the partonic momenta within the proton are parametrised by the PDFs, which are measured with an associated factorisation scale. Initial-state evolution of incoming partons above this cutoff scale creates additional, final-state, external particles that do not participate in the hard interaction.

The final state radiation (FSR) stage evolves final-state hard partons (and those originating from ISR) via further perturbative radiation. This process starts at the hard interaction scale and terminates, for a given branch, when an emission occurs below a cutoff in the shower evolution scale, which is typically of $\mathcal{O}(\Lambda_{\text{QCD}})$. At this point, one transitions from the perturbative to non-perturbative regime, whereby the partons hadronise into colour-singlet hadrons. This hadronisation process is characterised by small momentum transfers between confined particles, hence the distribution of hadrons approximately conserves the momenta and topology of the underlying partonic structure of a hard event [79]. Therefore, for hard interactions involving coloured partons, rather than observing well separated particles in the final state, we often instead see collimated sprays of energetic, colour-neutral hadrons that we call jets [80][81].

Jets are not fundamental objects defined by pQCD; we must therefore find a way of relating these collections of collimated hadrons to their underlying partonic ele-
2.1. INTRODUCTION

ments, such that one can apply perturbative techniques. To this end, one can use jet algorithms to provide a precise definition of observable quantities such as jet total momentum and/orientation, which can be mapped back to its partonic origins. Additionally, by combining collimated emissions into jets, we implicitly do not resolve collinear emissions, which ensures that jet-based observables are insensitive to the IR divergent nature of the QCD matrix elements.

A jet algorithm is a defined set of rules or procedures that select and group subsets of collimated hadrons in an event to form jets. In order to construct a jet algorithm, we must define a distance measure between each possible pair of particles in the event. This quantity parametrises how the kinematics of each individual particle is correlated with the rest of the event. We also define a maximum distance measure (a jet radius cutoff) in order to impose a limit to which a particle can be associated with another to form a single jet and to cut off IR divergences. This ensures that widely separated, collimated groups (clusters) of particles form distinct jets and often provides an endpoint for the jet finding algorithm. Additionally, when two or more particles are deemed part of the same jet, we must define a way of combining the particle momenta into a combined jet momentum, which is called the recombination scheme. In this way, we also define how the individual momenta of the hadron ensemble can be combined to define a total jet momentum.

There exists two main classes of jet algorithms, cone algorithms and sequential recombination algorithms. Cone algorithms start with an initial jet axis and then act by summing over all the particle momenta within a cone of fixed radius from this axis. The direction of this momentum sum forms the new jet axis and the algorithm iterates until a stable axis is found. However, the choice of initial seed is often sensitive to IR emissions and ambiguities exist with how to deal with particles that lie within two separate jet cones (for discussion see Ref. [81] and references therein). In this thesis, we will only consider sequential recombination algorithms, which are simpler to define and are not sensitive to IR emissions.

This chapter is outlined as follows: in Section 2.2 we examine how sequential recombination algorithms form jets and explicitly consider the behaviour of the generalised $k_t$ algorithm. In Section 2.3 we discuss how jet substructure information can be used to identify hard substructure and remove uncorrelated soft contamination from high $p_T$ jets. In Section 2.4 we discuss global and non-global event shapes and demonstrate the presence of logarithmically enhanced contributions to jet observables at each order of perturbation theory. We then examine several sources of these logarithms and

---

1 For soft, wide-angle emissions, an additional cut must be placed on the minimum energy of a jet.
2 In this thesis we use the E-scheme [82], whereby the combination of $i,j$ results in new object $k$ with 4-momentum equal to the 4-vector sum of $i$ and $j$. This scheme allows combinations of massless particles to define massive jets.
outline the resummation procedure required to restore analytical predictivity in the large logarithmic region. Lastly, we discuss some non-perturbative contributions to jet observables.

### 2.2 Sequential recombination jet algorithms

In most modern analyses of particle collider data, jet formation is provided by sequential jet algorithms; these operate by recursively combining (clustering) pairs of particles. Specifically, for a given particle ensemble, the algorithm selects and clusters particle pairs that are closest in a pairwise distance measure. This is iterated on the new collection of particles until no pairs remain, or the algorithm reaches a termination condition.

#### Algorithm 1: Inclusive sequential recombination.

**Input**: Set of objects with energy and momenta considered for recombination.

1. Define a pairwise distance measure, $\Delta_{ij}$, beam distance, $\Delta_{iB}$ and cutoff distance $\Delta_{\text{cut}}$;
2. Define a combination procedure, Combine, using a chosen recombination scheme;
3. while size(objects) > 0 do
   4. for all $i,j$ in objects: Calculate $\Delta_{ij}$ and $\Delta_{iB}$;
   5. Find the smallest distance, $\Delta_{\text{min}}$, from all $\{\Delta_{ij}, \Delta_{iB}\}$;
   6. if $\Delta_{\text{min}}$ is a pairwise distance, $\Delta_{ij}$, then
      7. Combine $i$ and $j$ into a new object, $k$;
      8. Add $k$ to objects and Remove $i$ and $j$ from objects;
   else if $\Delta_{\text{min}}$ is a beam distance $\Delta_{iB}$ then
      10. if $\Delta_{\text{min}} > \Delta_{\text{cut}}$ then
          11. Label $i$ a final-state jet and Remove $i$ from objects;
      else
          12. Remove $i$ from objects;

The distance measure is defined specifically to reflect the probability of underlying soft and/or collinear dynamics expressed by the QCD matrix elements, and must always satisfy the requirement of infrared and collinear (IRC) safety. This means that, given a set of hard jets, the addition of arbitrary numbers of soft and/or collinear radiation to the initial particle ensemble will not change observable quantities such as jet multiplicity or momentum (for an in-depth discussion see Ref. [78]). By imposing this constraint, we ensure that the IR divergent nature of the QCD matrix elements are not manifest...
2.2. SEQUENTIAL RECOMBINATION JET ALGORITHMS

in calculating observable quantities, at any order in perturbation theory. Specifically, we require complete cancellation of all divergent contributions to the observable, which are associated with real and virtual emissions in the soft and/or collinear limit. Given a distance measure that is IRC safe, a generic inclusive sequential jet algorithm used in a hadron-hadron collider will proceed as defined in Algorithm 1 where \( \Delta_{\text{cut}} \) is a transverse momentum cutoff that ensures that well-separated arbitrarily-soft emissions cannot form jets by themselves.

One crucial feature of sequential recombination jet algorithms, for this thesis, is that each jet has a well defined and accessible recombination (or “clustering”) history. This history can provide an approximate picture of the sequential, shower-like, branching evolution of the hard parton that initiated the jet. Using this information, one can apply cuts based on the clustering history of the jet to the kinematics of the jet constituents; in doing so one can potentially separate signal from background jets, and remove some external contamination coming from the rest of the collider environment. This can be a powerful tool, and we will return to it in Section 2.3.

One can also define an exclusive version of Algorithm 1 that is typically used in \( e^+e^- \) colliders. In this context we do not consider the distance to a colour-neutral beam, so we simply remove \( \Delta_{iB} \) and \( \Delta_{\text{cut}} \) from line 1 in Algorithm 1 and replace them with a single jet resolution measure instead, \( \Delta_{\text{RES}} \). Furthermore, we replace lines 9 - 13 with the modification defined in Algorithm 2.

Algorithm 2: Exclusive sequential recombination modification.

```
else if \( \Delta_{\text{min}} \) is the resolution distance \( \Delta_{\text{RES}} \) then
    Label all remaining objects final-state jets and exit the loop.
```

If one varies \( \Delta_{\text{RES}} \), the final multiplicity of jets in a given event will be affected. For example, increasing the resolution parameter will tend to decrease the total number of jets, by merging a greater number of pairs before terminating. The behaviour of each algorithm will depend on the definition of the distance measures \( \Delta_{ij} \) and \( \Delta_{iB} \) (or \( \Delta_{\text{RES}} \)); in the next subsection, we consider a specific definition called the generalised \( k_t \) algorithm.

2.2.1 The generalised \( k_t \) algorithm

We want to construct an IRC safe definition of the pairwise distance measure \( \Delta_{ij} \), and beam distance \( \Delta_{iB} \). In a hadron collider, it is useful to consider longitudinally boost invariant variables such as particle transverse momentum \( p_T \) and squared angular distance between particles, \( \Delta R^2_{ij} \). The generalised inclusive \( k_t \) algorithm \([82, 85]\) uses
these variables to define the following distance measures:

$$\Delta_{ij} = \min \left( p_{T_i}^2, p_{T_j}^2 \right) \frac{\Delta R_{ij}^2}{R^2},$$

$$\Delta_{iB} = p_{T_i}^2,$$

(2.1)

where $\Delta R_{ij}^2 = \Delta \eta_{ij}^2 + \Delta \phi_{ij}^2$, and $\Delta \eta_{ij}, \Delta \phi_{ij}$ are the separation of particles $i$ and $j$ in pseudorapidity and azimuthal angle with respect to the incoming beam axis respectively. $R$ is a jet “radius” parameter that ensures that clustering does not occur for pairwise distances $\Delta R_{ij}^2 > R^2$. The distance measures in Eq. (2.1) are IRC safe because the addition of an arbitrarily soft and/or collinear branching will always be clustered to another parton before we declare the final-state jet (i.e. before $\Delta_{iB}$ is the smallest distance) given that $\Delta R_{ij}^2 < R^2$ is satisfied. The remaining parameter, $p$, crucially determines both the order of clustering and the shape of the resulting jets in $(\eta, \phi)$ space. Three typical choices are given by [82–85]:

$$p = \begin{cases} 
1, & k_t \text{ algorithm}, \\
0, & \text{Cambridge/Aachen (C/A) algorithm}, \\
-1, & \text{anti-}k_t \text{ algorithm}.
\end{cases}$$

(2.2)

An example of jet clustering with each parameter choice, for a given set of hard partons, is visualised in Fig. 2.1. The coloured regions depict the extent of the jets that have been obtained from a set of hard partons in the presence of diffuse soft emissions. The differences in shape and area of the jets in rapidity-$\phi$ space is characteristic of each distance measure employed. When one uses the $k_t$ algorithm ($p = 1$), $\Delta_{ij}$ is typically small for pairs of particles containing a particle with small transverse momentum, hence the clustering will tend to recursively combine the smallest $p_T$ particle with its nearest neighbour. In this way, the clustering history for a $k_t$ jet will tend to consist of initial soft combinations that get progressively harder towards the end of the clustering sequence. In contrast, the Cambridge/Aachen algorithm ($p = 0$) is defined to be insensitive to the transverse momentum of each particle, and combines pairs purely based on the relative distance $\Delta R_{ij}^2$. Both of these algorithms approximate the tree-like partonic structure that results from the QCD splitting functions, with a preference for reconstructing collinear (and soft) pairs of particles. This produces an underlying jet clustering history (or jet substructure) that we can associate with the theoretical picture of the perturbative final-state parton shower.

The anti-$k_t$ algorithm ($p = -1$) behaves differently, the smallest $\Delta_{ij}$ often cor-

\footnote{For soft-wide angle emissions $\Delta R_{ij}^2 > R^2$, we make use of the additional cut $\Delta_{cut}$ to restore IRC safety.}
Figure 2.1: An example of a parton level event that has been clustered into jets using different jet algorithms. The event consists of a set of hard particles generated using Herwig [86] that occupy cells in rapidity and azimuthal angle \((y, \phi)\) space with \(p_T\) given by the vertical axis. A large number of soft, randomly distributed “ghost” particles are added to the event before clustering. If a cell contained a particle that is clustered into a hard jet, then it is coloured accordingly. This gives an indication of the relative behaviour and catchment area of different jet algorithms. This figure is taken directly from Ref. [83].

responds to pairs that contain a particle with a large transverse momentum. These (relatively) hard particles will therefore repeatedly combine with their nearest soft neighbours, forming a small number of approximately static hard “cores”. Hence, all soft radiation tends to be clustered to these cores up to the maximum invariant distance cutoff, \(R\). Unlike \(k_t\) or C/A clustering, this method often produces uniformly circular jets (see bottom panel of Fig. 2.1), but one sacrifices some correlation between the underlying sequential branching picture and the anti-\(k_t\) clustering history. For this reason, one must take care when using a jet substructure analysis on anti-\(k_t\) jets.

Given that a collection of particles have been combined to form one or more jets, one can associate measurable quantities such as jet multiplicity or momentum with the hard partons involved in the hard process. However, in practice, this simple mapping is often insufficient. Jets may be “contaminated” by perturbative radiation from the incoming
partons (ISR) or particles originating from additional soft interactions in the event, these include underlying event (UE) and pileup (PU). These effects are proportional to the reach\(^4\) of the jet, hence can become significant for large \(R\). Additionally, at small \(R\), contributions from soft, non-perturbative corrections that result from the hadronisation process and perturbative losses from final state radiation outside the reach of the algorithm can impact the measurement of jet observables \(^87\). We will discuss these effects in more detail in Subsection 2.4.5.

Depending on the physics search, one or more of these contamination effects may contribute significantly to jet observables. One possible solution is to introduce a radius dependent counterterm that can subtract (event by event) the average “noise” from a particular jet, which is approximated by a uniform background of soft particles \(^88\)–\(^90\). An alternative, jet substructure approach to this problem involves using the clustering history of a jet to selectively remove particles likely to be contamination. In the next subsection, we will examine how jet substructure information can be used to tag signal jets and remove spurious particle contributions.

2.3 Jet substructure

With the LHC probing multi-TeV interactions, it is now feasible to analyse EW bosons and top quarks that have transverse momenta much greater than their mass. Hadronic decay products of these boosted “signal” particles tend to be collimated into single, “fat” jets\(^5\). Following the pioneering work by Seymour \(^91\) and more recently Butterworth, Davison, Rubin and Salam (BDRS) \(^92\), an active research field has emerged to investigate the discriminatory power of jet substructure in the context of boosted heavy-particle decay with respect to jets initiated by light QCD partons. In addition, substructure information associated with a boosted jet can be used to identify and remove emissions that exhibit kinematics uncorrelated with the perturbative, collinear evolution of the hard initiator of the fat jet. This means that some of the contamination from soft radiation can be removed from the jet, improving signal jet mass resolution whilst reducing the mass of light QCD initiated background jets (see Refs. \(^93\)–\(^95\) for an overview). In the next two subsections, we will explore the discriminatory action and soft contamination removal of jet substructure algorithms applied to boosted fat jets.

2.3.1 Tagging

An important objective for any particle physics analysis is to distinguish between (or “tag”) signal (heavy-resonance initiated jet) and background (boosted QCD jet) events.

\(^4\)This is given by the total phase space in which an algorithm combines soft particles into the jet, i.e. the coloured areas in Fig. 2.1, for further discussion see Ref. \(^81\).

\(^5\)Fat jets are typically defined as \(R \sim O(1)\) boosted jets with non-trivial substructure, i.e. those which potentially contain the tree-level decay products of a boosted heavy resonance.
Information about the substructure of a jet, as described by the clustering history of the jet algorithm, can be used to discriminate between the different distributions of radiation inside signal and background jets.

For example, the hadronic decay of a boosted colourless heavy resonance such as $H \to b\bar{b}$ creates a collimated quark, anti-quark pair. The resulting QCD parton shower will be angular ordered with respect to the colour dipole defined by the quark pair. Therefore, the collinear enhanced radiation pattern in the fat jet will be concentrated into two narrow subjets with momenta corresponding to the initial momenta of the quark, anti-quark pair. Crucially, each subjet will tend to contain a significant fraction of the total jet transverse momentum compared to QCD initiated jets, which favour subjets with asymmetric $p_T$ fractions. In order to see this result, we now compute the distribution of the transverse momentum fraction between the Higgs decay products in the highly boosted frame. Starting in the rest frame of a scalar boson, such as the Higgs, one always obtains a back-to-back isotropic decay into particles with equal energy (see left hand panel of Fig. 2.2). We require a boosted Higgs decay, hence we perform a Lorentz boost of this system from the frame in which the Higgs is at rest and one of its decay products forms an angle $\theta'$ with respect to the boost axis, to frame in which the entire system acquires a transverse momentum $p_T$ with respect to the beam axis (see right hand panel of Fig. 2.2); we also set the azimuthal angle $\phi' = 0$ for simplicity. We can write the particle kinematics in the boosted frame, in terms of rest
frame quantities, as:

\[
p_H = \gamma M_H (1, 0, 0, \beta), \\
p_b = \frac{\gamma M_H}{2} \left( 1 + \beta \cos \theta', 0, \frac{1}{\gamma} \sin \theta', \cos \theta' + \beta \right), \\
p_\bar{b} = \frac{\gamma M_H}{2} \left( 1 - \beta \cos \theta', 0, -\frac{1}{\gamma} \sin \theta', -\cos \theta' + \beta \right),
\]

(2.3)

where the z-axis is defined as the boost direction, \( \beta = 1/\sqrt{1 + \Delta} \), \( \Delta = M_H^2/p_T^2 \) and the Lorentz factor is \( \gamma = 1/\sqrt{1 - \beta^2} \). In the highly boosted regime \( p_T \gg M_H \), the boost axis is approximately perpendicular to the beam direction and the minimum transverse momentum fraction \( z \) and energy fraction \( x \), carried by \( b \) or \( \bar{b} \) are equal up to correction terms of \( \mathcal{O}(\Delta) \). Hence, neglecting these corrections, we can write \( z \) as the ratio of the energies in the boosted frame:

\[
z \simeq \min \left( \frac{p_0^b, p_0^{\bar{b}}}{p_0^H} \right) = \frac{1}{2} \left( 1 - \frac{1}{\sqrt{1 + \Delta}} |\cos \theta'| \right).
\]

(2.4)

The square matrix element for the decay of the scalar Higgs boson to fermions is independent of the angle \( \theta' \) (see e.g. Ref. [30]), hence the \( b\bar{b} \) pair are uniformly distributed\(^6\) in the rest frame both azimuthally and in \( \cos \theta' \). Hence, we see from Eq. (2.4) that the distribution of \( z \) in the highly boosted frame is also uniform in the region\(^7\) \( 1/2 > z \gtrsim \Delta/4 \). This means that if one samples the minimum subjet transverse momentum fraction of a highly boosted Higgs jet, one is equally likely to observe any value of \( z \).

This may not appear to be an effective discriminant for boosted Higgs jets, until one considers the contrasting substructure of a light QCD initiated jet. A jet initiated by a light quark or gluon will be colour connected to external partons in the hard interaction, hence the resulting QCD shower will tend to be single subjet like and exhibit a relatively diffuse (wide angle) radiation pattern compared to the signal jet. Additionally, the \( p_T \) fraction carried by the hardest two subjets reflect the distribution in \( z \) as defined by the relevant AP splitting function given in Eq. (1.58); hence, the distribution in \( z \) is strongly peaked in the region \( z \to 0 \). It is therefore more probable for a QCD jet to contain subjets that are asymmetric in total jet transverse momentum fraction compared to an electroweak boson signal jet. We show this explicitly in Fig. 2.3 after some further discussion.

\(^6\)For W/Z bosons, the decay in the rest frame will not be isotropic due to spin correlations. This translates into a non-trivial (but non-divergent) distribution in \( z \) after boosting. However, the result is still sufficiently different from the QCD splitting function such that \( z \) is still an effective variable for discrimination.

\(^7\)Strictly speaking, in the context of boosted fat jets, the lower limit on \( z \) will be defined by the requirement that both decay products are within the same boosted jet, i.e. \( z \gtrsim \Delta/R^2 \), when \( R < 2 \).
Differentiating between signal and background jets based on subjet $p_T$ asymmetry began with the proposal of the Y-splitter technique \[96\]. This algorithm places a simple cut on the $k_t$ scale of the final clustering within a jet performed using the inclusive $k_t$ algorithm, as defined in Algorithm 4. Equivalently, one can obtain this scale by undoing (or “declustering”) the final merging step of a given $k_t$ jet to obtain two subjets$^8$ which we label $i$ and $j$. For a $1 \rightarrow 2$ decay in the collinear limit with subjet $p_T$ fractions $z$ and $1 - z$, the $k_t$ scale between $i$ and $j$ is given by

$$d_{ij} \simeq \min \left( z^2, (1 - z)^2 \right) p_T^2 \Delta R_{ij}^2,$$

where the Y-splitter scale $d_{ij}$, corresponds to $\Delta_{ij}$ in Eq. (2.1) after setting $R = 1$ and $p = 1$. Noting that the angle between two collinear massless decay products for a boosted two-prong decay that define a final jet mass, $M_{\text{jet}}$, is given by

$$\Delta R_{ij}^2 \simeq \frac{M_{\text{jet}}^2}{z (1 - z) p_T^2},$$

we obtain, given $z < 1/2$:

$$d_{ij} \simeq \frac{z}{1 - z} M_{\text{jet}}^2.$$  

We know that the distribution in $z$ is approximately uniform for EW boson initiated jets from Eq. (2.4), therefore we expect to obtain $d_{ij} \approx M_{\text{EW}}^2$ for signal jets. QCD initiated jets, on the other hand, tend to exhibit an asymmetric $p_T$ fraction between the two subjets as previously discussed, leading to smaller values of $d_{ij}$. The Y-splitter algorithm removes jets that exhibit asymmetry in $z$ via a cut on $d_{ij} \sim \mathcal{O}(M_{\text{EW}}^2)$. This reduces the probability of mistagging a background jet as signal, whilst retaining a significant fraction of signal jets. Recently, several new jet shapes have been proposed to directly measure the “subjet-ness” of a jet, mainly N-subjettiness \[97\], template based methods \[98\] and others \[99–103\].

One can modify the original Y-splitter technique by instead cutting on the ratio $y = d_{ij}/M_{\text{jet}}^2$ (see Refs. \[81,104\]). By rearranging Eq. (2.7), we see that this ratio directly probes the $p_T$ fraction of the $k_t$ subjets when we make the replacement $z = y/(1 + y)$ at leading order. We show this explicitly in Fig. 2.3 by plotting the normalised differential distribution of $z$, where $z$ is the minimum $p_T$ fraction of two subjets $i$ and $j$ produced by declustering the final merging step of an inclusive $k_t$ jet such that we always obtain

\[\footnote{If one was to examine the $k_t$ scale of the final clustering within a jet defined with the C/A or anti-$k_t$ algorithm, one would often not resolve any hard substructure (if present) inside the jet. This is because the distance measure Eq. (2.1) with $p = 0$ or $p = -1$ does not favour hard pairwise recombinations at the end of the clustering procedure.}

71
CHAPTER 2. JET PHENOMENOLOGY

Figure 2.3: The normalised differential distribution of the minimum subjet $p_T$ fraction for quark, gluon and Higgs jets. Events were generated using HERWIG++ 2.7.0 with 14 TeV $pp$ collisions, generator level cuts $p_T^{\text{jet}} > 3$ TeV and $M_{\text{jet}} = M_H \pm 25$ GeV with hadronisation and underlying event. The Higgs jets were generated using the process $pp \rightarrow ZH$ with the $Z$ and Higgs boson decaying leptonically and hadronically respectively. The C/A algorithm as implemented by FASTJET and RIVET, was used with $R = 1$ to cluster the jets and the hardest hadronic jet in each event was selected. The constituents of this hardest jet were subsequently clustered using an inclusive $k_T$ algorithm and then the final merging step was undone to obtain exactly two subjets, from which $z$ was calculated.

One observes that the shape of the $z$ distribution for quark/gluon and Higgs jets are significantly different. Specifically, the signal jet exhibits an approximately flat distribution in $z$ whilst QCD jets favour small (asymmetric) values of $z$. One can place a cut on the subjet $p_T$ fraction, $z > z_{\text{cut}}$, such that we only retain jets containing sufficiently symmetric subjets, thereby, on average, rejecting a larger fraction of background than signal jets. For example, Fig. 2.3 indicates that for $z_{\text{cut}} = 0.1$, with the given cuts, one retains approximately 84% of Higgs jets, 68% of gluon jets and 56% of quark jets.

9This observable is not strictly collinear safe. However, this ratio satisfies an extended IRC safety after introducing a generalised fragmentation function in divergent regions of phase space. For an extensive discussion see Ref. 105.

10One would also have to include the $b$-tagging efficiency for $H \rightarrow bb$ jets for a more complete picture of signal tagging efficiency.
2.3. JET SUBSTRUCTURE

In this subsection, we have demonstrated how jet substructure information alone can be used as a tool to discriminate between Higgs and light QCD jets. These techniques are commonly used in conjunction with a mass window in the vicinity of the signal peak to further enhance the signal to background discrimination. However, to ensure that soft contamination does not cause a significant number of signal jets to fall outside this jet mass window, we can employ some “grooming” techniques to the jet.

2.3.2 Grooming

In order to improve signal to background discrimination, one must ensure that the influence of spurious radiation on jet observables, prevalent in the high-luminosity environment of the LHC, can be reduced in order to obtain good signal resolution. This cleaning up, or “grooming”, of jets can be combined with a tagging procedure via the introduction of multi-purpose tools known as jet substructure algorithms. These work, in a general sense, by attempting to remove some (or all) constituents of a jet such that one retains only the dominant, perturbative radiation from the signal decay products. In this way, jet substructure algorithms perform two complimentary actions; firstly, they improve the resolution of signal jet observables by removing contamination and secondly, attempt to remove emissions in QCD background jets that have large contributions to a jet observable, like the mass, such that it can be used as a discriminant.

Algorithm 3: Mass drop tagger.

External parameters: \( \mu, y_{\text{cut}} \).

Input: Jet, \( j \) with radius \( R \) and constituents defined by a C/A clustering history.

1. while \( \text{size(constructs)} > 0 \) do
2.   Undo the last clustering step of \( j \) to give subjets \( j_1 \) and \( j_2 \) such that \( M_{j_1} > M_{j_2} \);
3.   if \( \frac{\min(p_{T1}^2, p_{T2}^2) \Delta R_{12}^2}{M_{j}^2} > y_{\text{cut}} \) and \( M_{j_1} < \mu M_{j} \) then
4.     Label \( j \) the mass drop tagged jet and exit the loop;
5.   else
6.     Remove \( j_2 \) from constituents and Relabel \( j = j_1 \)

Output: Mass drop tagged jet.
Specific examples include mass drop tagger (MDT) \cite{92}, pruning \cite{108,109}, trimming \cite{110}, however there are many others \cite{14,104,111–116}. In this thesis, we will concentrate on the named three; all of which contain a pairwise asymmetry check, as seen in Y-splitter, but extend this to a multi-step procedure and include additional angular and mass dependent constraints on the jet constituents. We will now consider the specific example of the mass drop tagger with filtering and its action on a boosted fat jet.

The mass drop tagger was defined in the pioneering BDRS paper \cite{92} as a way of separating hadronically decaying boosted EW boson jets from the extensive QCD and top backgrounds at the LHC. It is defined in Algorithm 3, where the constraints on line 3 impose that the two subjets are sufficiently symmetric (left inequality) and exhibit a mass drop (right inequality). The external parameters of the algorithm, $y_{\text{cut}}$ and $\mu$, are dimensionless and can be tuned as required for each analysis (for instance $y_{\text{cut}} = 0.09$ and $\mu = 0.67$ were used in Ref. \cite{92}).

Similarities exist between this algorithm and the Y-splitter technique, namely the mass drop tagger uses a cut on $d_{ij}$ (see Eq. (2.5)) that is proportional to the squared jet mass. As before, this eliminates the mass dependence of the $k_t$ measure such that the asymmetry check is dimensionless (a $p_T$ fraction). This ensures that the same cut is applied at each step of the algorithm. However, in MDT, the asymmetry criteria is imposed iteratively to remove particles from the C/A clustering history, rather than using it exclusively as a tagging step, as was the case in Y-splitter. This iteration selectively removes (grooms) the wide-angle, $p_T$ asymmetric contributions that are commonly associated with uncorrelated contamination or wide-angle QCD emissions. This grooming proceeds until one reaches an angular scale that exhibits the symmetry and mass drop characteristics of signal-like hard substructure and then terminates. All radiation inside the jet at smaller angles with respect to the jet axis are not considered for grooming, which, in principle, corresponds to the upper limit of angular ordered FSR emitted by the signal prongs. This algorithm contains both a grooming and tagging step, however it was found that contamination entering the jet at smaller angular scales than the typical splitting angle for a two prong decay, defined in Eq. (2.6), still had a significant effect on jet mass resolution for moderate jet $p_T \sim 300$ GeV \cite{92}. In response, a pure grooming technique called filtering was introduced and applied to the jets tagged by MDT. It is defined in Algorithm 4, where $R_{\text{filter}}$ is chosen such that it is smaller than the angular scale at which the mass drop tagger algorithm tagged the jet $R_{\text{tag}}$, and $n_{\text{filter}}$ is an integer.
Algorithm 4: Filtering.

**External parameters:** $R_{\text{filter}}$, $n_{\text{filter}}$.

**Input:** Jet, $j$ with constituents.

1. **Recluster** the constituents of $j$ with a generalised $k_t$ algorithm using radius $R_{\text{filter}}$ into subjets $i$;
2. **Combine** the hardest $n_{\text{filter}}$ subjets $i$ and **Relabel** as $j$;
3. **Label** $j$ as the filtered jet;

**Output:** Filtered jet.

In the BDRS paper, the parameter values $R_{\text{filter}} = \min(0.3, R_{\text{tag}}/2)$ and $n_{\text{filter}} = 3$ were found to be effective for Higgs decay. By retaining only radiation within $R_{\text{filter}}$ of each prong (and also the hardest perturbative emission for $n_{\text{filter}} = 3$), we have effectively reduced the impact of contamination (which is proportional to $R^4$ for the squared plain jet mass [87]) in the signal jet by reducing the “ungroomed” jet area. This corresponds directly to the phase space that is never checked by any asymmetry (or mass drop) condition.

Pruning and trimming will be described in depth in Chapter 3, but for now it suffices to note that whilst these techniques differ in implementation to MDT, they share the same basic features. Specifically, each algorithm includes an iterative step that removes uncorrelated emissions via an asymmetry cut on the $p_T$ fractions of subjets. Additionally, each applies grooming up to a minimum angular scale that is less than the jet radius; this is designed to retain perturbative FSR from signal subjets.

### 2.3.3 Jet substructure in practice

Given a description of how to groom and tag jets using jet substructure algorithms, we now require a quantitative assessment of the effectiveness of these techniques when applied to appropriate physics searches. In order to do so, we must examine how jet observables are affected by the application of substructure techniques across a wide range of parameter space. By comparing different substructure algorithms, we can gain important information about the relative efficiency when applied to either QCD background or signal jets, as well as the dependence on each external parameter.

One possible way to assess the action of jet substructure techniques is to make use of the multitude of numerical tools available. Monte Carlo generators, such as **Herwig++** [75], **Sherpa** [76] and **Pythia** 6/8 [74,77], provide a way of generating signal and background QCD events with a set of final-state particle momenta. Each generator simulates large multiplicities of final-state particles via a parton shower algorithm, as discussed in Section 1.9. One can also include additional perturbative and non-perturbative effects such as ISR, UE and hadronisation for study. Clustering
Figure 2.4: The normalised differential jet mass distribution for quark (left) and Higgs (right) jets before and after application of the MDT with filtering algorithm. This was generated using Herwig++ 2.7.0 \[75\] with generator level cut $p_T > 3$ TeV including hadronisation and underlying event. For comparison purposes, we assumed a b-tagging efficiency of 100% and considered all final-state hadrons to be stable\[11\]. We used the C/A algorithm with $R = 1$ to cluster the jets and only considered the hardest jet in each event. Parameters used for MDT with filtering are $y_{\text{cut}} = 0.11$, $\mu = 0.67$, $R_{\text{filter}} = \min(0.3, R_{\text{tag}}/2)$ and $n_{\text{filter}} = 3$.

and analysis of these jets can be performed easily using RIVET \[107\] combined with the FASTJET package \[106\], which contains implementations of most jet substructure tools. General purpose jet substructure techniques have been adapted and applied using Monte Carlo techniques to a range of specific boosted object searches \[117–137\]. Indeed, many of these have already been applied at the LHC to QCD jets \[138–141\] as well as specific experimental searches \[142–145\]. Additionally, assessment and comparison of the effectiveness of different substructure techniques for physics searches have also been performed using Monte Carlo generators \[93–95, 146–152\] in order to give an indication of the relative signal to background performance of each algorithm.

Adopting this numerical approach in Fig. 2.4, we show the normalised differential jet mass distribution for boosted quark (left) and Higgs (right) jets at $p_T = 3$ TeV before and after application of the MDT with filtering algorithm\[12\] as defined in Algorithm 3 and Algorithm 4 respectively. We first note that the plain mass distribution for

\[11\] At resonance level, there is the additional complication of hadrons containing $b$ quarks decaying via an electroweak boson to invisible particles such as neutrinos. This has the effect of degrading the signal peak equally for both groomed and ungroomed jets. For comparison purposes, it suffices to ignore this contribution.

\[12\] One may ask if experiments at the LHC can effectively measure multi-TeV jets with EW scale masses jets due to angular limits on calorimeter resolution (see Eq. (2.6)). However, work in Refs. \[153, 154\] indicated that by using the superior spatial resolution offered by a tracking detector, alongside the hadronic calorimeter, one can improve signal mass resolution for highly boosted jets.
the Higgs jet is smeared from a peak at $M_H = 125$ GeV to larger jet masses due to contamination from ISR and underlying event. After the application of the jet substructure algorithm, which removes soft, wide-angle contamination of the jet, we see that the signal peak is considerably sharper. In contrast, the QCD jet mass is instead dominated by soft, wide-angle emissions, and so the MDT algorithm recursively removes contributions to the jet such that $M_{\text{jet}} \to 0$. Therefore, we observe a reduction in the proportion of QCD jets in the vicinity of the Higgs mass peak, further increasing the discriminative power of a mass window in separating signal from background jets.

An alternative, non-numerical approach, and the main subject of this thesis, is to use analytical techniques to study the action of jet substructure algorithms on QCD and signal jets. In part pioneered by the work within this thesis, there have been several jet substructure studies that have implemented analytic tools such as resummed calculations in perturbative QCD \cite{14,116,155,159} and soft-collinear effective theory \cite{160,161}. Studies of this type are able to unambiguously parametrise the features of each algorithm; in particular, the interplay between the external parameters and the kinematic regions in which the cuts are employed. Whilst one is able to observe these features via Monte Carlo, it is often difficult to disentangle the precise dependence of the algorithmic behaviour with respect to each parameter. This behaviour can also often be obscured by additional physics effects such as initial/final-state radiation, hadronisation and underlying event. Additionally, the numerical impact of these effects on the jet observable may be Monte Carlo/tune dependent (see, for example Refs. \cite{95,138,141,162}). Clearly, analytical results lead to a broader and more coherent understanding of jet substructure algorithms. In particular, these enable the identification of redundant parameters and insight into how each algorithm removes soft contributions to the jet and the kinematic regions in which they are active. It is also instructive to consider how jet substructure algorithms analytically affect the logarithmic structure of jet observables, such as the jet mass, and the resulting convergence of the perturbative series. In the next section we discuss some simple jet observables and use pQCD to describe their associated analytical structure.

\section{Jet shapes and resummation}

In this section we define observables that can be used to probe the QCD structure of a hard interaction by measuring the distribution of hadronic momenta in an event. We then calculate the leading logarithmic structure of these observables and discuss the contribution of non-perturbative effects.

\subsection{Global and non-global event shapes}

Global event-shapes are defined as observables that parametrise the distribution of all hadronic momenta in an event. For example, one can define a variable parametrising
the degree of collimation of an event; this is known as the thrust parameter \[163,164\] and for an event containing \(m\) final-state particles, is defined as:

\[
T_m(p_1, \ldots, p_m) = \max_n \frac{\sum_{i=1}^m |\vec{p}_i \cdot \hat{n}|}{\sum_{i=1}^m |\vec{p}_i|},
\]

where \(\vec{p}_i\) is the 3-vector of particle \(i\) and the unit vector \(\hat{n}\) is the thrust axis, this is taken to be the direction that maximises \(T_m\). One can see that \(T_m \to 1\) as the momentum of each particle in the event tends to (anti-)collinearity with the thrust axis.

In order to ensure that event shapes are theoretically well defined we require that they satisfy the IRC safety requirements, as discussed in Section 2.2. In the context of event shapes, this is the requirement that the value of the observable is unchanged if an arbitrary number of final-state particles are replaced with collinear pairs and/or one adds soft emissions to the ensemble. For example, if we introduce a collinear splitting to a particle \(p^*_m \to z p^*_m + (1-z) p^*_m\), where \(z\) is a momentum fraction, the thrust measure becomes:

\[
T_{m+1}(\vec{p}_1, \ldots, \vec{p}_{m-1}, z\vec{p}_m, (1-z)\vec{p}_m) = \max_{\hat{n}} \frac{\left(\sum_{i=1}^{m-1} |\vec{p}_i \cdot \hat{n}| + z|\vec{p}_m \cdot \hat{n}| + (1-z)|\vec{p}_m|\right)}{\left(\sum_{i=1}^{m-1} |\vec{p}_i| + z|\vec{p}_m| + (1-z)|\vec{p}_m|\right)} = T_m(\vec{p}_1, \ldots, \vec{p}_m).
\]

This property demonstrates that the value of the thrust observable is invariant under a collinear splitting. It is also trivial to show that a soft emission in the limit \(z \to 0\) contributes to thrust with zero weight, which implies that Eq. (2.10) is also true for soft emission. Fulfilment of IRC safety constraints depends on the precise definition of the event shape and it is possible to define event shapes that do not satisfy these requirements. However, these unsafe observables usually have theoretical issues such as non-convergence of perturbative calculations and correspondingly large non-perturbative corrections. Additionally, some unsafe observables can not be resummed to all orders (for example jet fractions using the JADE algorithm \[165\]).

Non-global event shape variables are defined as observables that only parametrise the distribution of a subset of hadronic momenta in an event. Relevant examples include jet shapes, which only consider the constituent particles of a given jet that have been obtained via application of a jet finding algorithm. These observables are inherently non-global because they only include particles within a limited phase space region around the jet axis and discard all others. Jet shapes are more sensitive to the substructure of jets than global observables, hence are more suitable for probing the underlying QCD structure of a jet. One example of a non-global jet observable, used
extensively in this thesis, is the normalised squared jet mass:

\[ \rho = \frac{M_{\text{jet}}^2}{Q^2} = \frac{1}{Q^2} \left( \sum_{i \in \text{jet}} p_i \right)^2, \]  

(2.11)

where the jet mass \( M_{\text{jet}} \), is normalised to a hard scale \( Q \), and we have summed over the 4-momentum of all particles within the jet. One can easily show that this jet mass definition is IRC safe using the same arguments presented for the thrust observable.

### 2.4.2 Jet shape distributions and large logarithms

We can use these observables to build up event shape distributions that can be compared to theoretical predictions. To this end, we can start by writing down the integrated cross section for an \( n \)-jet event configuration \( \Delta_n \), which has a global observable \( \tilde{G} \) less than a value \( \lambda \), in the form \[ \Delta_n (\lambda) = \sum_{k=n}^{\infty} \int d\sigma_k J_n (p_1, ..., p_k) \Theta (\lambda - \tilde{G} (p_1, ..., p_k)), \]  

(2.12)

where \( J_n \) is a set of constraints that are equal to 1 for exactly \( n \) jets or 0 otherwise and \( d\sigma_k \) is the differential cross section for producing \( k \) final-state partons. In this thesis, we are interested in non-global jet observables in the small \( \lambda \) limit; therefore, using the form of Eq. (2.12), we now derive an expression for an \( n \)-jet integrated cross section \( \Sigma_n (\lambda) \) with a non-global jet observable \( \tilde{O} \) less than an arbitrary value \( \lambda \).

Just considering the leading terms in the small \( \lambda \) limit, we specialise to a “Born” configuration, defined as \( n \) particles forming \( n \) hard, final-state jets\footnote{We neglect configurations in which more than one hard particle is within a single jet. These contributions to the integrated cross section are not logarithmically enhanced and are therefore subleading in the small \( \lambda \) limit.} and take the measured jets to also contain an arbitrary number of eikonal emissions. Firstly, we use the result for the factorisation of eikonal emissions given in Eq. (1.57), to write

\[ d\sigma_k \simeq d\sigma_n \frac{1}{(k-n)!} \prod_{i=n+1}^{k} d\sigma_{eik}^i, \]  

(2.13)

where the combinatoric factor \( (k-n)! \) arises when considering identical emissions, \( d\sigma_n \) is the differential cross section for producing \( n \) hard partons and \( d\sigma_{eik}^i \) is the differential cross section for eikonal emission \( i \). Now, inserting Eq. (2.13) into the global observable expression Eq. (2.12) and replacing \( \tilde{G} \) with a non-global observable \( \tilde{O} \), we can write
the leading $n$-jet integrated cross section in a factorised form:

$$
\Sigma_n (\lambda) \simeq \sum_{k=n}^{\infty} \int d\sigma_n J_n (p_1, ..., p_n) \left( \frac{1}{(k-n)!} \prod_{i=n+1}^{k} d\sigma_{i}^{\text{eik}} \hat{R} (p_i) \right) \times \Theta \left( \lambda - \hat{O} (p_J) \right),
$$

(2.14)

here we have defined the set of particle 4-momenta within the jets as $p_{J_x} = \{ p_{j_x} | x \in \{1, ..., n\} \}$ where $p_{j_x}$ is the set of 4-momenta of the particles within jet $j_x$ and $\hat{R} (p_i)$ is a Heaviside step function that constrains particle $i$ to be within one of the jets. The observable $\hat{O}$ receives contributions from all possible combinations of real, eikonal particles within the jets and the corresponding virtual emissions. We define the differential eikonal cross section in the soft and collinear limit (see Eq. (1.57)) as:

$$
d\sigma_{i}^{\text{eik}} = C_I \frac{\alpha_s}{\pi} \frac{d\theta_i^2 dz_i}{\theta_i^2 z_i} = C_I \frac{\alpha_s}{\pi} \frac{d\rho_i dz_i}{\rho_i z_i},
$$

(2.15)

where $C_I$ is a colour factor that corresponds to the hard initiator of the jet, a quark ($C_F$) or gluon ($C_A$) and we use a fixed $\alpha_s$, which is sufficient for leading logarithmic accuracy. We define $\theta_i$ and $z_i$ to be the angle and energy fraction of particle $i$ with respect to the parent emitter. In the last equality we have chosen to parametrise the collinear limit with a kinematical variable $\rho_i$ that scales linearly with $\theta_i^2$.

Ordering the eikonal emissions with respect to their contribution to $\rho$ introduces an additional combinatoric factor $(k-n)!$ and we can replace $\hat{R} (p_i) = \Theta (R^2 - \theta_i^2)$ for small $R$ jets clustered using the C/A algorithm. Inserting Eq. (2.15), we obtain the expression

$$
\Sigma_n (\lambda) = \sigma_n \sum_{k'=0}^{\infty} k'! \times \left( \frac{1}{k'!} \prod_{i=1}^{k'} C_I \frac{\alpha_s}{\pi} \int \frac{d\rho_i dz_i}{\rho_i z_i} \Theta \left( R^2 - \theta_i^2 \right) \Theta \left( \rho_{(i-1)} - \rho_i \right) \right) \times \Theta \left( \lambda - \hat{O} (p_J) \right),
$$

(2.16)

where we have redefined the summation variable $k' \equiv k - n$, integrated over the $n$-jet Born configuration to give the Born cross section $\sigma_n$, and imposed ordering of eikonal emissions $\rho_{(i-1)} > \rho_i$.

In order to simplify the above expression, we must consider the form of the phase space constraints imposed by the observable $\hat{O}$ on the real emissions and the contribution of the corresponding virtual emissions. We can write the contribution of all possible combinations of $k-n$ emissions, which may be real or virtual, to the observable
in the following form:
\[
\Theta \left( \lambda - \tilde{O} (p_J) \right) = \sum_{\mathcal{R} \subseteq E} \left[ (-1)^{|E| - |\mathcal{R}|} \times \Theta \left( \lambda - \sum_{i \in \mathcal{R}} \rho_i \right) \right],
\]
(2.17)

where \( E \) is the set of emissions \( E = \{ n+1, \ldots, k \} \) and \( \mathcal{R} \) is the set of real emissions, which is defined as \( E \) and all its subsets. For a given subset \( \mathcal{R} \), the Heaviside step function on the right hand side imposes that the sum of the individual contributions to \( \tilde{O} \) from each real emission, \( \rho_i \), is less than \( \lambda \). Additionally, each set \( \mathcal{R} \) contains contributions from \( |E| - |\mathcal{R}| \) virtual emissions, where \( |E|, |\mathcal{R}| \) is the number of elements in set \( E \) and \( \mathcal{R} \) respectively; this introduces an overall factor \( -1 \) if the number of virtual emissions is odd. We are interested in the leading logarithmic form of this observable in the small \( \lambda \) limit, hence we now impose strong ordering of emissions such that the jet observable is defined only by a single contribution \( i \). This approximation simplifies Eq. (2.17) to:
\[
\sum_{\mathcal{R} \subseteq E} \left[ (-1)^{|E| - |\mathcal{R}|} \times \Theta \left( \lambda - \sum_{i \in \mathcal{R}} \rho_i \right) \right] \approx \sum_{\mathcal{R} \subseteq E} \left[ (-1)^{|E| - |\mathcal{R}|} \times \prod_{i \in \mathcal{R}} \Theta \left( \lambda - \rho_i \right) \right] = \prod_{i \in E} \left[ \Theta \left( \lambda - \rho_i \right) - 1 \right],
\]
(2.18)

where we have used the leading logarithmic approximation\(^{14}\)\(^{166}\) in the first line to rewrite the sum over \( \rho_i \) as a product of Heaviside step functions. Expanding this expression, one can easily show that it can be re-written in the simple form shown in line 2. Inserting Eq. (2.18) into Eq. (2.16) we obtain the expression
\[
\Sigma_n (\lambda) = \sigma_n \sum_{k'=0}^{\infty} \left( \prod_{i=1}^{k'} C_i \frac{\alpha_s}{\pi} \int \frac{d\rho_i \, dz_i}{\rho_i \, z_i} \Theta \left( R^2 - \theta_i^2 \right) \Theta \left( \rho_i - \rho_i \right) \left[ \Theta \left( \lambda - \rho_i \right) - 1 \right] \right),
\]
(2.20)

where now each real emission individually satisfies \( \lambda > \rho_i \) and the \( "-1" \) term in the square brackets represents the corresponding virtual correction. Henceforth, as an example, we explicitly use the normalised jet mass (or equivalently the virtuality of the off-shell emitter) defined in Eq. (2.11): for one emission in the soft and collinear limit, this is given by \( \rho_i \approx z_i \theta_i^2 \). Performing the virtual cancellation, such that the

\(^{14}\)To see this, one can use the decomposition for \( \rho_i > 0 \):
\[
\Theta \left( \lambda - \sum_{i} \rho_i \right) = \left[ 1 - \Theta \left( \sum_{i} \rho_i - \lambda \right) \right] \prod_{i} \Theta \left( \lambda - \rho_i \right),
\]
(2.19)

to show that if any one contribution to the observable is much greater than the rest, i.e. \( \sum_{i} \rho_i \approx \rho_x \), where \( x \in i \), the second term in the square brackets will contradict a constraint in the product. Hence, one can neglect this term in the strongly ordered approximation.
integral over $\rho_i$ is finite, we can now explicitly write the phase space constraints on each emission $i$:

$$
\Sigma_n (\lambda) = \sigma_n \sum_{k'=0}^{\infty} \left( \prod_{i=1}^{k'} - C I \frac{\alpha_s}{\pi} \int_{R^2}^{\rho_i} \frac{d\rho_i}{\rho_i} \int_{\rho_i/R^2}^{1} \frac{dz_i}{z_i} \right) = \sigma_n \sum_{k'=0}^{\infty} \frac{1}{k'!} \left( - C I \frac{\alpha_s}{\pi} \int_{R^2}^{\rho_0} \frac{d\rho'}{\rho'} \int_{\rho'/R^2}^{1} \frac{dz'}{z'} \right)^{k'},
$$

(2.21)

where we have defined the upper limit on emission $\rho_1$ as $\rho_0 \equiv R^2$. This can be rewritten as an exponential of the single emission result and evaluated to give a leading logarithmic expression:

$$
\Sigma_n (\lambda) = \sigma_n \times \exp \left[ - C I \frac{\alpha_s}{\pi} \int_{R^2}^{\rho_0} \frac{d\rho'}{\rho'} \int_{\rho'/R^2}^{1} \frac{dz'}{z'} \right] = \sigma_n \times \exp \left[ - C I \frac{\alpha_s}{\pi} \frac{1}{2} \ln^2 \frac{R^2}{\lambda} \right].
$$

(2.22)

The exponential term can be interpreted as the fraction of events, for a given Born process, which satisfy the phase space constraints imposed by the value of the observable $^{15}$ The form of Eq. (2.22) indicates that, at each order in $\alpha_s$, there are logarithmic contributions to $\Sigma_n (\lambda)$. This is a direct result of the constraints imposed on the real emission phase space via the jet shape observable. Whilst the divergent terms of real and virtual contributions cancel for a non-global IRC safe observable,miscancellation in finite regions of phase space manifest as logarithms in the collinear and/or soft emission limit.

One can write down an expansion of the exponential in Eq. (2.22) and truncate the series in $\alpha_s$ to obtain an approximate fixed-order description of $\Sigma_n (\lambda)$. However, due to the presence of the double logarithmic term, care must be taken to ensure that this series is sufficiently convergent; such an expansion is only strictly valid in the region $\ln (R^2/\lambda) \gg 1/\sqrt{\alpha_s}$, i.e. not when $\lambda \ll R^2$. Otherwise, one must use an all-orders, or resummed, result such as the one given in Eq. (2.22). As an explicit example, fixed-order and resummed calculations for total broadening $^{16}$ $B_T$, are illustrated in Fig. 2.5. The fixed-order LO and NLO results in the small $B_T$ limit diverge due to the presence of large logarithms in the truncated expansion. Inclusion of the exponentiated next-to leading logarithmic contribution reproduces the peak observed in the OPAL experimental data $^{169}$.

$^{15}$This fraction is equivalent to the probability of no emissions, $i$, in jet $j$ that satisfy $\rho_i > \lambda$, hence one can see the correspondence between this result and the Sudakov form factor defined in Eq. (1.62).

$^{16}$This global observable is related to thrust Eq. (2.9) and defined as $B_T = \frac{\sum |p_i \times \hat{n}|}{\sum |p_i|}$, where $\hat{n}$ is the thrust axis.
2.4. JET SHAPES AND RESUMMATION

Figure 2.5: Comparison of fixed-order LO and NLO calculations and resummed results for differential distribution of total broadening observable at parton level to experimental data. \[167,168\]

2.4.3 Resummation

In the last subsection, we generated an all-orders result for a simple jet observable to a leading approximation in the small $\lambda$ limit. This exponentiated structure extends to many IRC safe observables and can be written in a general form for small $\lambda$ \[32,78\]:

$$
\frac{\Sigma_n(\lambda)}{\sigma_n} = \left(1 + \sum_{i=1}^{\infty} \left(\frac{\alpha_s}{2\pi}\right)^i A_i\right) \exp\left[L g_1(\alpha_s L) + g_2(\alpha_s L) + \alpha_s g_3(\alpha_s L) + \ldots\right] + B(\alpha_s, \lambda), \tag{2.23}
$$

where $L \equiv \ln (1/\lambda)$. The function $L g_1(\alpha_s L)$ contains all the leading logarithmic (LL) contributions of the form $\alpha_s^n L^{n+1}$, $g_2(\alpha_s L)$ contains all the next-to leading logarithmic (NLL) contributions of the form $\alpha_s^n L^n$ and $\alpha_s g_3(\alpha_s L)$ contains all the next-to-next-to leading logarithmic (NNLL) contributions of the form $\alpha_s^{n+1} L^n$. $B(\alpha_s, \lambda)$ represents all power corrections that vanish in the limit $\lambda \to 0$. $A_i$ are the coefficients of non-exponentiating, non-logarithmic correction terms.

In order to obtain the full expression for $g_1$ (rather than just the double-logarithmic leading term given in Eq. (2.22)), we must compute additional subleading logarithmic corrections to the single emission integral given in Eq. (2.21). Crucially, one has to properly treat the running of $\alpha_s$, which is evaluated at the $k_t$ of the emission relative to the parent \[78\], rather than using a fixed coupling. To this end, we use the result for the 1-loop running coupling given in Eq. (1.26), to write

$$
\alpha_s(k_t^2) = \frac{\alpha_s(M_Z^2)}{1 + \beta_0 \alpha_s(M_Z^2) \ln \frac{M_Z^2}{k_t^2}} = \alpha_s(M_Z^2) \left[1 - \beta_0 \alpha_s(M_Z^2) \ln \frac{M_Z^2}{k_t^2} + \ldots\right], \tag{2.24}
$$
where we evolve from a measurement of $\alpha_s$ at a fixed scale (commonly the $Z^0$ boson mass, $M_Z$) and the ellipsis denotes higher order terms from the expansion around small $\alpha_s$. In the collinear limit, the $k_t$ of the emission with respect to the parent is proportional to the emission angle, which is cutoff by the $\lambda$ constraint on our IRC safe observable, so this correction generates additional terms in the exponent $\alpha_s^n L^{n+1}$ starting at $n = 2$. Inclusion of the running coupling at 1-loop to the single emission integral produces a full expression for the LL term $g_1$ in Eq. (2.23).

### 2.4.4 NLL contributions

Given that we can compute all the LL contributions in the exponent for a given observable, one would like to improve this result by computing the NLL term, $g_2$. This is less straightforward and includes logarithms originating from multiple, non-strongly ordered emissions, hard-collinear effects and a 2-loop running coupling expression in the single emission integral.

To include logarithmic contributions from hard-collinear effects, one can replace the soft splitting $1/z$ in Eq. (2.21) with the (unregularised) splitting function, $p_{gg}(z)$ or $p_{qg}(z)$, given in Eq. (1.58). This replacement introduces additional logarithms of the form $\alpha_s^n L^n$ to the exponent starting at $n = 1$, as well as power corrections that vanish in the small $\lambda$ limit. By including these hard-collinear corrections, along with the 1-loop running coupling in the single emission integral, one obtains an expression that is NLL accurate in the expansion. This means that we control all terms $\alpha_s^n L^{2n}$ and $\alpha_s^n L^{2n-1}$ after perturbative expansion of the resummed result, as depicted in Fig. 2.6.

For NLL accuracy in the exponent, one must also consider single logarithmic corrections associated with two or more non-strongly ordered emissions. This accuracy implies that we now control all terms up to and including $\alpha_s^n L^n$ contained within the exponent after perturbative expansion of the resummed result (this does not imply that we control all terms $\alpha_s^n L^n$ in Eq. (2.23), see Fig. 2.6), hence one is less sensitive to uncalculated, logarithmically subleading contributions. Firstly, it is possible for a combination of two emissions, which are not strongly ordered, to jointly contribute to the precise value of an observable, which introduces single logarithmic corrections to the exponent (see Ref. 78 for a treatment of these effects in global observables).

Next, one must also include logarithmic contributions present in non-global observables, known as non-global logarithms (NGL). These contribute when an emission outside the observable phase space radiates a soft parton into the observable region or vice versa. This secondary soft emission contributes to the non-global observable, whilst the corresponding virtual correction does not. These logarithms are purely soft in origin due to the wide angle kinematics of each emission in the non-global configuration (see left panel in Fig. 2.7). Currently, analytical resummation of NGLs to all orders is prohibitively difficult due to the colour structure and geometry associated
2.4. JET SHAPES AND RESUMMATION

Figure 2.6: Schematic diagram that shows the perturbative contributions to an observable that result from an $\alpha_s$ expansion of a resummed result at each order. Each circle indicates the existence of a contribution of the form $\alpha_n^s L^m$. The colour of each circle shows which terms in the expansion are resummed by an exponential with LL accuracy (grey), NLL accuracy in the expansion (grey+red) and NLL accuracy in the exponent with the first hard, wide-angle correction (grey+red+green). The partially filled circles indicate that some (but not all) terms of this order are captured by the resummed result. The contributions with white circles require greater accuracy in the exponent and/or additional hard wide-angle corrections to the observable to be included.

with large numbers of wide-angle gluons. However, the resummation can be undertaken numerically in the large $N_c$ approximation [170]. After performing this resummation, these contributions introduce additional single logarithmic terms of the form $\alpha_n^s L^m$ in the exponent, starting at $n = 2$.

Finally, emissions close to the limit of the observable phase space can be clustered with other emissions via a clustering algorithm such that real contributions are added or removed from the non-global observable. For example, if we have a (semi-)hard emission with opening angle approximately, but less than, jet radius $R$ and a soft emission outside the jet boundary, as shown in the right panel Fig. 2.7, the 4-vector sum of the clustered gluon pair can be inside the jet. A miscancellation between the soft gluon $k_2$, which contributes to the jet, occurs with its corresponding virtual correction. Hence, we also obtain a tower of clustering logarithms (CL) of the form $\alpha_n^s L^m$ starting at $n = 2$. Like NGLs, these contributions have been resummed to all orders numerically [172].

---

This correction is required to capture all terms given by the filled green circles. Matching and merging techniques are used in Monte Carlo generators to incorporate some NLO hard corrections into the parton shower. This enables a better description of hard, wide-angle emissions, which are not reproduced well by the collinear parton shower (for an overview of these techniques see Refs. [73][171] and references therein).
Figure 2.7: Examples of kinematical configurations that give rise to non-global logarithms (left) and clustering logarithms (right) for a jet defined with radius $R$. In the left-hand diagram, soft emission $k_2$ is emitted with opening angle less than $R$ with respect to $p$. If clustered to $p$, $k_2$ contributes to the jet, introducing non-global logarithms. In the right-hand diagram, soft emission $k_2$ is emitted with opening angle greater than $R$ with respect to $p$. However, if initially clustered to $k_1$, $k_2$ will contribute to the jet, introducing clustering logarithms.

The relative magnitude of clustering and non-global logarithms depends on the jet algorithm and recombination scheme employed. This is because kinematical boundaries defined by the non-global jet observable are influenced by self clustering of emissions. For example, observables derived from jets defined using the anti-$k_t$ algorithm have a larger correction from NGLs than clustering logarithms due to a smaller phase space available for self clustering of emissions. The $k_t$ algorithm, on the other hand, favours soft clustering, hence the phase space for the non-global configuration is reduced, increasing the overall coefficient of the CL contributions.

2.4.5 Non-perturbative effects

So far we have only considered the partonic description of final-state particles. In order to relate this description with measurements at particle collider experiments, we must account for the conversion of partons to hadronic objects, which occurs at a non-perturbative scale $\sim \Lambda_{\text{QCD}}$. The hadronisation process can be modelled in several ways. One simple example is the tube model, whereby each coloured parton pair hadronises into a jet of light hadrons occupying a tube in $(\eta, \phi)$ space. If we assume that the hadronisation process produces a jet of hadrons that are distributed uniformly in $p_T$ with respect to a coloured parton pair, one can write the total energy and longitudinal momentum carried by the hadrons as

$$\langle E \rangle_{\text{had}} = \lambda \int_0^{\eta_{\text{max}}} \cosh \eta \, d\eta = \lambda \sinh \eta_{\text{max}},$$

For anti-$k_t$ jets, the phase space for self-clustering is not logarithmically enhanced, hence CL are entirely absent.\(^\text{172}\)
2.4. JET SHAPES AND RESUMMATION

Figure 2.8: Resummed results for the differential distribution of the total broadening observable with and without hadronic corrections, compared to experimental data. \[167, 168\]

\[
\langle P \rangle_{\text{had}} = \lambda \int_0^{\eta_{\text{max}}} \sinh \eta \, d\eta = \lambda (\cosh \eta_{\text{max}} - 1), \tag{2.25}
\]

where \(\eta_{\text{max}}\) is an arbitrary collinear cutoff for hadron creation and the mean transverse momentum of each hadron \(\lambda\), is derived from the corresponding Fermi motion of the bound partons, \(\lambda \sim 1\) GeV. For \(\eta_{\text{max}} \gg 1\), one can use Eq. (2.25) to write \(\langle P \rangle_{\text{had}} \sim \langle E \rangle_{\text{had}} - \lambda \) \[174\] and therefore estimate the mean shift in the normalised squared jet mass \(\rho\), as

\[
\left\langle \frac{M_{\text{jet}}^2}{Q^2} \right\rangle_{\text{had}} = \frac{\langle E \rangle_{\text{had}}^2 - \langle P \rangle_{\text{had}}^2}{Q^2} \sim \frac{\lambda}{Q}, \tag{2.26}
\]

where \(Q\) is the hard scale and we have neglected \(\mathcal{O}(\lambda^2/Q^2)\) corrections. Therefore, we find that hadronisation corrections to a normalised differential event shape distribution causes a modification that scales like \(1/Q\). Specifically, hadronisation corrections to the normalised jet mass distribution for boosted jets is only comparable in magnitude to the analytic resummed result when \(\rho \sim \frac{\lambda}{Q}\), i.e. when the jet mass is dominated by contributions resulting from emissions with \(k_t \sim \Lambda_{\text{QCD}}\). In Fig. 2.8, we show how a correction in the total broadening observable at parton level (see also Fig. 2.5) of \(\mathcal{O}(1/Q)\) improves the agreement of the theoretical prediction with the observed experimental result.

The tube model provides an estimate of the scaling behaviour of the corrections associated with hadronisation. Several relatively sophisticated models have been developed to model the hadronisation process for an arbitrary number of final-state coloured particles. An example is the string model \[175, 178\], which extends the tube model by
defining “strings” between each colour connected, final-state parton. As the particles move apart, the colour strings fragment into $q\bar{q}$ pairs that eventually create the final-state hadrons. Additionally, gluons are colour connected to two partons and produce kinks in the strings. This produces angular distributions of hadrons that is in good agreement with experiment [179]. Another example is the cluster model [180,181], whereby colour singlet clusters are constructed from the final-state partons, which decay into hadrons. This can be achieved by forcing non-perturbative decays of $g \rightarrow q\bar{q}$ and combining neighbouring colour singlet pairs of quarks into clusters, which decay isotropically in their rest frame into hadrons.

When considering jet observables, we must also account for underlying event (UE) effects. As mentioned previously in Section 1.6, colliding protons interact in a hard event via a single pair of partons. However, we neglected the possibility that scattering occurs between multiple pairs of partons coming from each proton. Given that one scattering results in a large momentum transfer, it is unlikely that an additional interaction also participates in a comparably hard scattering. Instead, these interactions tend to be soft, diffractive QCD events, which produce soft particles with energy of order a few GeV. These scattering events are highly non-perturbative and have to be modelled and tuned to experimental data for use in Monte Carlo generators [182–185].

Additionally, we must consider pileup (PU), which occurs when multiple proton pairs interact within the same bunch crossing. Like UE, the probability of multiple hard scattering events is small, therefore these additional interactions tend to result in further, non-perturbative, soft emissions. In the high-luminosity environment of the LHC, at $\sqrt{s} = 8$ TeV, the average number of pileup vertices in ATLAS was measured to be $\sim 21$ interactions per bunch crossing [186], hence it is important to model PU effects in MC generators via minimum bias simulation (see for example Ref. [74] for discussion).

UE and PU act as major sources of soft hadronic noise in detectors at the LHC and these particles can be included in jets when using a jet clustering algorithm. This uncorrelated hadronic contamination can significantly distort the measurement of many jet observables, hence it becomes of importance to either subtract or remove this contamination. As mentioned in Subsection 2.3.2 jet substructure algorithms offer a mechanism to selectively remove soft emissions within a jet, thereby eliminating some of this noise from measurement of jet observables.

In the next chapter, we explicitly calculate some of the leading logarithmic terms in the collinear limit at $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ in the perturbative expansion of the differential quantity $\frac{1}{\sigma_0} \frac{d\Sigma(\lambda)}{d\lambda}$ derived from the integrated expression in Eq. (2.23), where $\lambda$ is the normalised differential jet mass after application of several jet substructure techniques and $\sigma_0$ is the two-jet Born cross section.
3.1 Introduction

In Chapter 2 we argued that an analytical understanding of jet substructure algorithms would provide several advantages over numerical simulations using MC generators. Analytical expressions for jet observables would be independent of the event generator/parton shower/tune employed for a particular study and provide precise parametrisation of the features of the algorithm observed in MC. Knowing the parametric dependence of jet observables on parameters such as jet transverse momentum $p_T$, and radius $R$, as well as tagger parameters such as $y_{\text{cut}}$ and $\mu$, (see Algorithm 3) facilitates unambiguous evaluation of tagger behaviour. For example, one may be able to better understand the role of each parameter and define precisely the kinematic regions in which the taggers remove soft radiation.

The action of jet substructure algorithms on the logarithmically enhanced, perturbative structure of the plain jet mass distribution is also of interest. It is known that the integrated plain mass distribution contains large double logarithms at each order of the coupling of the form $\alpha_s^n \ln^2 n (M_j/p_T)$ where $M_j$ is the jet mass \[159,166,187,188\]. These logarithms are large in the region $M_j \ll p_T$, i.e. in the boosted regime, and must therefore be resummed to all orders to restore predictivity of perturbative results. We anticipate that, by placing cuts on the energy of soft emissions, the application of substructure algorithms will partially eliminate these logarithms in some regions of emission phase space. The resulting perturbative structure will dictate which methods provide the best theoretical description of such jet observables. This may include resummation, fixed-order calculations or matched combinations of the two. For this study it suffices to consider QCD jets produced in $e^+e^-$ collisions and therefore neglect
CHAPTER 3. FIXED-ORDER CALCULATIONS FOR JET SUBSTRUCTURE

initial state radiation present in hadron colliders. These subleading contributions are non-collinear in the boosted regime and would not change the conclusions reached in this chapter (see Chapter 4 for an analytical discussion of ISR contamination in signal jets).

This chapter can be outlined as follows: in Section 3.2, we present the leading logarithmic structure for the plain jet mass distribution for ease of comparison with subsequent calculations of the taggers. Following this, we explore the logarithmic structure at leading order (LO) and next-to-leading order (NLO), for the mass drop tagger and propose a modification that we call the modified mass drop tagger (mMDT) [116] in Section 3.3. We then proceed by calculating the NLO structure of the mMDT algorithm in Section 3.4, including non-global logarithms and compare our predictions to fixed-order estimates using Event2 [66]. Next, we consider the analytical structure of pruning and Y-pruning [116] at LO and NLO and test our results against Event2 in Section 3.5. We repeat this investigation for trimming in Section 3.6 and conclude by noting the differences in the fixed-order structure calculated for each of the taggers studied here.

3.2 Plain jet mass

We start by considering the differential distribution \( \frac{d\sigma}{\sigma d\rho} \), where \( \rho = \frac{M^2_j}{E^2_j} \) is the jet mass normalised to the jet energy squared. In the small \( \rho \) limit, we can write this distribution in pQCD as

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \approx \frac{\alpha_s}{\pi} (a_{12}L + a_{11}) + \left( \frac{\alpha_s}{\pi} \right)^2 \left( a_{24}L^3 + a_{23}L^2 + a_{22}L \right) + \mathcal{O}(\alpha^3_s),
\]

(3.1)

where we have neglected power corrections of \( \mathcal{O}(\rho) \) and defined \( L \equiv \ln \left( \frac{R^2}{\rho} \right) \) in the limit \( R \ll 1 \). The terms with coefficients \( a_{12i} \) are referred to as double logarithms, \( \mathcal{O}(\alpha^2_s L^{2n}) \), in the integrated distribution) and originate from emissions in the soft and collinear limit. The single logarithmic terms in Eq. (3.1) have coefficients \( a_{1i} \). In order to determine \( a_{11} \) and \( a_{23} \), we need to consider hard collinear emissions and running coupling effects. We do not compute \( a_{22} \) because this contribution to the plain jet mass contains many additional sources of logarithms including multiple emission effects [78], non-global logs [170], clustering logs [172], hard-wide angle \( \mathcal{O}(\alpha_s) \) emissions and 2-loop running coupling effects. One should note that the fixed-order expression in Eq. (3.1) is consistent with the normalised, resummed, integrated cross section for a non-global observable presented in Eq. (2.23) after one perturbatively expands Eq. (2.23) around small \( \alpha_s \) and differentiates the expression with respect to \( L \).

For the plain jet mass distribution the coefficients, \( a_{ij} \) in Eq. (3.1) have the following
values, for quark-initiated jets\cite{166}:  
\begin{align*}
a_{12} &= C_F, \\
a_{11} &= -\frac{3C_F}{4}, \\
a_{24} &= -\frac{C_F^2}{2}, \\
a_{23} &= \frac{3}{8}C_F(3C_F + 4b_0),
\end{align*}
(3.2)
where $b_0 = \pi\beta_0 = \frac{1}{12}(11C_A - 4T_Rn_f)$. This term comes from the 1-loop running coupling correction Eq. (1.26) to the leading order result. These coefficients indicate a double logarithmic distribution at LO and NLO and also suggest an exponentiated form. Indeed, these fixed-order results can be written as an expansion of the exponential (see Subsection 2.4.2 for an in-depth derivation of the exponentiation of logarithmic contributions to the plain jet mass):
\begin{align*}\frac{\rho}{\sigma} \frac{d\sigma^{(\text{plain})}}{d\rho} &\approx \frac{d}{dL} \exp \left[ -C_F \frac{\alpha_s}{\pi} \int_{\rho}^{R^2} \frac{d\rho'}{\rho'} \int_{\rho'/R^2}^{1} p_{gg}(x) \, dx \right] \\
&\approx C_F \frac{\alpha_s}{\pi} \left( L - \frac{3}{4} \right) \exp \left[ -C_F \frac{\alpha_s}{\pi} \frac{L}{2} \left( L - \frac{3}{2} \right) \right],
\end{align*}
(3.3)
where we have taken a fixed-coupling approximation, i.e. ignored running coupling effects. The first line is simply a logarithmic derivative acting on the probability of observing no emissions with a jet mass contribution greater than $\rho$. As discussed previously, the expression on the first line of Eq. (3.3) can be obtained by exponentiating the single emission probability using the full splitting function $p_{gg}(x)$, where $x$ is the energy fraction of the emitted parton. The form of Eq. (3.3) indicates that the differential distribution initially rises as one decreases $\rho$, however this growth is eventually suppressed by the exponential factor, which approaches zero when $\rho \to 0$. This produces the familiar Sudakov peak for the differential plain jet mass distribution (see Fig. 3.1).
In order to further motivate the following calculations in this chapter, we provide the MC simulation taken from Ref. [116] in Fig. 3.1. This plot shows the differential distribution in $\rho$ for highly boosted 3 TeV jets after application of a range of taggers at parton level. One can immediately observe significant differences between the plain jet result and that of each tagger. Notably, one observes one or more transition points in the distribution and, for the case of trimming and pruning, Sudakov peaks that

\textsuperscript{1}In this chapter we always compute results for quark initiated jets for ease of comparison with the fixed-order code EVENT2 [66]. One can obtain equivalent results for gluon jets by re-computing the integrals under replacement of the colour factor $C_F$ and splitting function $p_{gg}(x)$ with $C_A$ and $p_{gg}(x)$ respectively.
are shifted relative to the plain mass result. For a given tagger, this rich structure originates from the existence of distinct regions in which the removal (or non-removal) of perturbative radiation results in modifications to the logarithmic structure of the plain jet mass observable. The position of each transition point from one behaviour to another, will in general, depend on the choice of tagger parameters. In order to describe the overall shape of each distribution we should therefore start by calculating the logarithmic structure of the jet mass observable after application of each substructure algorithm. Hence, we now seek to determine expressions for the perturbative coefficients of the post-tagging jet mass distribution, like those for the plain jet mass in Eq. (3.1), starting with the mass drop tagger.

3.3 Mass drop tagger

3.3.1 Definition

The mass drop tagger (MDT) \[92\] algorithm starts with a fixed number of hard jets \( j \), with radius \( R \) that have been clustered using the C/A algorithm \[84\] and is defined in Algorithm 3. We remind the reader that the MDT algorithm acts on a jet by firstly undoing the last stage of the C/A clustering into two subjets such that \( M_{j_1} > M_{j_2} \). If these subjets exhibit a significant mass drop: \( \mu M_j > M_{j_1} \), and the splitting is not too \( p_T \) asymmetric: \( y = \min(\frac{p_{T,j_1}}{M_{j_1}^2}, \frac{p_{T,j_2}}{M_{j_2}^2}) \Delta R_{j_1,j_2}^2 > y_{cut} \), the jet is tagged as an MDT jet. Else, the subjet with the smallest mass is removed and this process is iterated for the remaining subjet until tagging occurs or the subjet has no constituents.

MDT contains two dimensionless external parameters \( \mu \) and \( y_{cut} \), which one is free to define and optimise, as required by a particular physics analysis. For this study, we shall use an \( e^+e^- \) adaptation of Algorithm 3 whereby we simply replace transverse...
momentum $p_{T,j}$ with the relevant energy $E_j$ and rewrite $\Delta R_{j_1,j_2}^2 \to 2(1 - \cos \theta_{j_1,j_2})$ where $\theta_{j_1,j_2}$ is the angle between $j_1$ and $j_2$. After this replacement we can write $M_j^2 \simeq 2E_{j_1}E_{j_2}(1 - \cos \theta_{j_1,j_2})$ and simplify the asymmetry measure, $y$:

$$y = \frac{\min(E_{j_1}, E_{j_2})}{\max(E_{j_1}, E_{j_2})} \simeq \frac{\min(x, 1-x)}{\max(x, 1-x)} \quad (3.4)$$

where we have used the collinear limit of the splitting $j \to j_1, j_2$ to write the energy of each parton as a fraction, $x$, of the parent parton energy.

One should note, in the original paper MDT includes an additional step, whereby one applies filtering \[92]. For this study we shall ignore the action of the filtering step on the mass distribution. This algorithm has little effect for high $p_T$ QCD jets as shown in Ref. \[116]. However, in Chapter 4 we discuss how filtering can be important in the context of moderately boosted signal jets after application of MDT.

### 3.3.2 Leading order

In this section we carry out a leading order (LO) calculation of the differential normalised jet mass distribution $\frac{1}{\sigma} \frac{d\sigma}{d\rho}$ after application of the MDT algorithm.

We start by considering a back-to-back quark-antiquark pair produced in a $e^+e^-$ collision and introduce a leading order correction to the system via the emission of a single soft gluon with momentum $k$. We consider the quark to be massless and have a momentum $p$ that lies along the z-axis:

$$p = E_q (1, 0, 0, 1) ,$$

$$k = E_g (1, 0, \sin \theta, \cos \theta) , \quad (3.5)$$

where $\theta$ is the angle between the quark and gluon, which have energies $E_q$ and $E_g$ respectively. Note that we have neglected recoil of the parent parton, which introduces correction terms subleading to our required accuracy.

We now redefine the energies in terms of fractions of the total jet energy, $E_j$; we write this as $E_q = (1-x) E_j$ and $E_g = xE_j$. We require that the gluon is combined with the quark when clustered using the C/A algorithm, which occurs when $\Delta_\theta < \Delta_R$, where

$$\Delta_\lambda \equiv 2(1 - \cos \lambda) , \quad (3.6)$$

which in the collinear limit, is simply $\theta^2 < R^2$.

We now consider the constraints imposed on the real emission by the MDT algorithm. The mass drop constraint is trivially satisfied because each subjet contains only one massless parton. To pass the asymmetry cut however, we require that the emission
energy fraction lies in the region
\[
\frac{y_{\text{cut}}}{1 + y_{\text{cut}}} < x < \frac{1}{1 + y_{\text{cut}}},
\] (3.7)
such that neither the quark nor the gluon is removed.

The leading order differential distribution is given by the following integral:
\[
\frac{1}{\sigma} \frac{d\sigma}{d\rho} \bigg|_{\text{MDT,LO}} = 2\alpha_s C_F \pi \int \frac{d\cos \theta}{1 - \cos^2 \theta} \int_{\frac{1}{1+y_{\text{cut}}}}^{1+y_{\text{cut}}} dx p_{gq}(x) \delta \left( \rho - x \Delta \phi \right) \Theta \left( \Delta_R - \Delta \phi \right),
\] (3.8)
where \(p_{gq}(x)\) is the unregularised AP splitting function given in Eq. (1.58). We are working in the soft limit, i.e. \(x \ll 1\), so we have written the normalised jet mass\(^2\) as \(\rho = 2x(1 - x)(1 - \cos \theta) \approx x \Delta \phi\).

Evaluating Eq. (3.8) and neglecting terms of \(\mathcal{O}(\rho)\), we arrive at the result:
\[
\rho \frac{d\sigma}{\sigma} \bigg|_{\text{MDT,LO}} = \frac{\alpha_s C_F}{\pi} \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{1}{1+y_{\text{cut}}} \right)} \right) \Theta \left( \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta_R - \rho \right) + \frac{\alpha_s C_F}{\pi} \ln \left( \frac{\frac{4 \tan^2 R}{\rho} e^{-\frac{3}{4}}}{\rho \left(1 + y_{\text{cut}}\right)} \right) \Theta \left( \rho - \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta_R \right),
\] (3.9)
where we have neglected corrections of \(\mathcal{O}(y_{\text{cut}})\) in the second line. In contrast to the plain jet mass distribution Eq. (3.1), one notices that for small masses, \(\rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta_R\), the soft logarithm in \(\rho\) has been replaced with one in \(y_{\text{cut}}\). At small \(\rho\), the magnitude of the logarithmic contribution to the differential distribution at this order, relative to the plain jet mass, is reduced, given one chooses \(y_{\text{cut}} \sim \mathcal{O}(1)\). For large masses \(\rho > \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta_R\), we obtain the double logarithmic plain jet mass distribution\(^3\) up to subleading corrections in \(y_{\text{cut}}\). Indeed, one can (and should) recover the leading order plain jet mass result for finite \(R\) by taking the limit \(y_{\text{cut}} \to 0\) of Eq. (3.9):
\[
\rho \frac{d\sigma}{\sigma} \bigg|_{\text{Plain,LO}} = \frac{\alpha_s C_F}{\pi} \ln \left( \frac{\frac{4 \tan^2 R}{\rho} e^{-\frac{3}{4}}}{\rho} \right),
\] (3.10)

The result in Eq. (3.9) indicates that, at leading order, MDT successfully removes sufficiently soft emissions from de-clustered parton pairs. In doing so, it modifies
\footnote{\footnotesize{2 We have also neglected contributions to the jet mass bilinear in \(x\), which correspond to longitudinal recoil. This replacement does not change the logarithmic result at the accuracy required in this chapter, but will introduce corrections of \(\mathcal{O}(y^2_{\text{cut}})\) to the location of the transition points. The reader should be aware that the transition points derived in this chapter correspond to the small \(y_{\text{cut}}/z_{\text{cut}}/f_{\text{cut}}\) limit.}}
\footnote{\footnotesize{3This is only true at double logarithmic level, if one treats the kinematics more carefully the result can be extended past this accuracy. Specifically, one could symmetrise the integral in Eq. (3.8) by imposing \(x < 1/2\), replacing \(p_{gq}(x) \to p_{gq}(x) + p_{gq}(1-x)\) and using \(\rho = x(1-x)\Delta \phi\).}}
the analytical perturbative logarithmic structure of the observable. After choosing an appropriate value for $y_{\text{cut}}$ we can expect that the fixed-order perturbative expansion for MDT is more convergent than the plain jet mass in the small $\rho$ limit. This would increase the domain of validity of fixed-order perturbative estimates used to calculate the MDT jet mass.

We now want to compare the result in Eq. (3.9), to the fixed-order program EVENT2, which we show in Fig. 3.2. EVENT2 [66] is a fixed-order Monte Carlo program that calculates next-to leading order corrections to two- and three-jet observables in $e^+e^-$ annihilation. It implements Catani-Seymour dipole subtraction [62] to numerically handle the cancellation of divergences present in the real and virtual contributions to an IRC safe observable. The EVENT2 numerical results are obtained at a centre of mass energy $Q = 1$ TeV, and we take the two hardest jets defined using the C/A algorithm ($R = 0.8$) and apply the MDT algorithm to each jet. We then measure the average mass of the two jets, which is equivalent to the analytical single jet calculation. One can see that the difference between the EVENT2 and the analytic result vanishes

---

3.3. MASS DROP TAGGER

![Graph showing comparison of Event2 at LO with the analytic expression Eq. (3.9) for a range of $y_{\text{cut}}$ values. The red points tend to a constant value in the small mass limit, which corresponds to the presence of a single logarithm in the integrated distribution. The green points show the difference between the analytic expression and the EVENT2 result in the region $\rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta R$, which tends to zero.](image_url)
in the limit $\rho \to 0$ for all values of $y_{\text{cut}}$, this demonstrates that we correctly reproduce the LO behaviour of the algorithm in this region.

Now that we have described the action of MDT at LO and observed a single logarithmic distribution in $\rho$, it is important to verify to what extent the replacement of soft logarithms occurs at next-to-leading order. In the next section, we calculate the leading logarithmic structure of MDT at NLO.

### 3.3.3 Next-to-leading order: The wrong branch issue

In this section, we want to establish whether MDT exhibits only single logarithms at next-to-leading order (NLO). Unfortunately, as we will demonstrate, there is an extra effect that appears at NLO in the behaviour of MDT that we call the “wrong branch” effect. This additional contribution means that the simple picture of single logarithms at each order is spoiled at NLO by an $O(\alpha_s^3 L^3)$ contribution to the integrated jet mass distribution.

In order to show this explicitly, we consider the configuration shown in Fig. 3.3. In this figure, we depict a gluon, $k$, emitted from a hard quark that subsequently splits into two partons $k_1$ and $k_2$. In the leading collinear regime, the angle between the emitted partons, $\theta_{12}$, is the smallest. In this limit, the C/A algorithm first clusters the two soft gluons together, then clusters this subjet to the hard quark, to finally form the fat jet.

We now apply the MDT algorithm to this jet, which starts by undoing the last clustering between the quark and double-gluon subjet. This produces two subjets: one massive subjet containing the soft parton pair, $j_1$, and another massless subjet containing the hard quark, $j_2$. In the soft and collinear limit, the mass of the fat jet is always much greater than the mass of $j_1$, which only contains soft partons. Therefore, the mass drop condition is always satisfied in the leading approximation and we move onto the asymmetry check. If this condition is satisfied, we return the mass defined by all three particles, however we want to examine the leading result when the asymmetry
condition fails. In this situation, MDT removes the subject with the smallest mass, which explicitly means that the hard quark is discarded and the algorithm recurses to the soft parton branch.

In order to evaluate the logarithmic contribution of this effect, we first denote the soft parton jet, \( j_1 \), to contain a total energy fraction \( x \) and note that the first asymmetry check will fail if

\[
x < \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \quad \text{or} \quad x > \frac{1}{1 + y_{\text{cut}}}. \tag{3.11}
\]

We are interested in the leading contribution, so we examine the phase space region in which the gluon subject is soft and discard the rightmost inequality.

We now follow the soft gluon subject and apply MDT cuts to the branching within \( j_1 \). We define partons \( k_1 \) and \( k_2 \) such that they carry momentum fractions \( 1 - z \) and \( z \) of the parent gluon, \( k \), respectively. The mass drop condition is always satisfied for the two massless subjets, \( k_1 \) and \( k_2 \), so we only require \( \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} < z < \frac{1}{1 + y_{\text{cut}}} \) for this subject to pass the asymmetry cut. We can write the resulting integral in the soft and collinear limit as

\[
C_F \left( \frac{\alpha_s}{\pi} \right)^2 \int \frac{d\theta^2}{\theta^2} \int \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \frac{dx}{x} \times \int \frac{d\theta_{12}^2}{\theta_{12}^2} \int \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \frac{dz}{z} \left( \frac{C_A}{2} p_{gg}(z) + T_{Rn_f} p_{qg}(z) \right) \delta \left( \rho - z (1 - z) x^2 \theta_{12}^2 \right), \tag{3.12}
\]

where the normalised mass of the soft parton jet is given by \( M_{j_1}^2 / E_{j_1}^2 = z (1 - z) x^2 \theta_{12}^2 \).

In Eq. \((3.12)\), we include the probability of the soft gluon, \( k \), splitting into a pair of gluons via the splitting function \( p_{gg}(z) \) and to a quark-antiquark pair via the splitting function \( p_{qg}(z) \), as defined in Eq. \((1.58)\). We evaluate Eq. \((3.12)\) in the small \( \rho \) limit, dividing the result into the \( C_F C_A \) and \( C_F n_f \) contributions respectively:

\[
\frac{\rho}{d\sigma} \left( \frac{M_{j_1}}{C_F C_A} \right) = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \left( \frac{1}{4} \ln \left( \frac{1}{y_{\text{cut}}} \right) + \frac{11 y_{\text{cut}}^3 + 9 y_{\text{cut}}^2 - 9 y_{\text{cut}} - 11}{12 (1 + y_{\text{cut}})^3} \right) \times \ln \frac{1}{\rho} \Theta \left( \frac{R^2 y_{\text{cut}}^3}{1 + y_{\text{cut}}^3} - \rho \right), \tag{3.13}
\]

\[
\frac{\rho}{d\sigma} \left( \frac{M_{j_1}}{C_F n_f} \right) = C_F n_f \left( \frac{\alpha_s}{\pi} \right)^2 \left( \frac{1}{4} \ln \left( \frac{1}{y_{\text{cut}}} \right) + \frac{11 y_{\text{cut}}^3 + 9 y_{\text{cut}}^2 - 9 y_{\text{cut}} - 11}{12 (1 + y_{\text{cut}})^3} \right) \ln \frac{1}{\rho} \Theta \left( \frac{R^2 y_{\text{cut}}^3}{1 + y_{\text{cut}}^3} - \rho \right), \tag{3.14}
\]

where we have neglected terms \( O(\alpha_s^2 L) \) in the differential distribution. The domain of validity of these expressions, given by the step function in Eqs. \((3.13, 3.14)\), is for smaller
than calculated in the LO result Eq. (3.9). However, crucially in the limit $\rho \to 0$, the wrong branch effect introduces terms $O(\alpha_s^2 L^3)$ to the integrated distribution. Whilst these are still less singular than the double logarithms present in the plain jet mass, such behaviour would likely require resummation to all orders to address the issue of large logarithms. It is not clear how a resummation that included these wrong branch effects would be written for MDT, hence it is desirable to eliminate them.

Before proposing a solution to the wrong branch effect, we first check the analytic expressions Eq. (3.13) and Eq. (3.14) against Event2, shown in Fig. 3.4. One can see that MDT contains leading contributions that are larger than single logarithms in both channels. Considering the difference between the Event2 and analytic result, we see that the difference is given by a straight line, corresponding to leftover terms $O(\alpha_s^2 L)$ in the differential distribution. This indicates that we successfully control the leading logarithms, which are a result of the wrong branch effect.

One solution to the wrong branch effect, is to modify the definition of MDT. The problem stems from the feature that MDT has the potential to follow a soft branch inside a jet when either the mass drop or energy asymmetry cut fails. It is clear this behaviour is undesirable because the intended action of MDT is to identify hard jet substructure. The easiest modification is to impose that MDT always recurses to the hardest subjet rather than the most massive. In the next section we explicitly write down the proposed modification to MDT, which we henceforth refer to as the modified mass drop tagger (mMDT).

### 3.4 Modified mass drop tagger

#### 3.4.1 Definition

The proposed modification of the mass drop tagger is to replace line 6 in the definition of MDT, defined in Algorithm [3], with Algorithm [5]:

**Algorithm 5:** MDT modification (mMDT).

Remove the softer of $(j_1, j_2)$ from constituents and Relabel $j$ as the remainder.

where softer means smaller transverse mass, $M^2 + p_T^2$, or transverse momentum, $p_T$, at hadron colliders or least energetic at $e^+e^-$ colliders.

It is clear that the leading order calculation for mMDT is identical to MDT, so we move on to the NLO calculation.

#### 3.4.2 Next-to-leading order: Independent emission

In this section, we will carry out a next-to-leading order calculation of the normalised differential mMDT jet mass distribution, concentrating on the leading loga-
3.4. MODIFIED MASS DROP TAGGER

![Graphs showing coefficient of $C_F C_A$ and $C_F n_f$ for MDT $R=0.8$, $y_{cut}=0.1$, 0.2, 0.4]}

Figure 3.4: Comparison of EVENT2 in the $C_F C_A$ and $C_F n_f$ channels at NLO with the analytic expression Eq. (3.13) on the left and Eq. (3.14) on the right, for a range of $y_{cut}$ values. The red points demonstrate an extra logarithm in both the $C_F C_A$ and $C_F n_f$ channels. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < \frac{y_{cut}}{(1+y_{cut})^3} \Delta R$, which demonstrates that we control all terms up to $O(\alpha_s^2 L)$ in the differential distribution.

We start by considering the independent emission of two real gluons, $k_1$ and $k_2$, from a quark (or antiquark), $p$, and the corresponding virtual corrections in the collinear limit. These are shown diagrammatically in Fig. 3.5. We assume each gluon is combined into a fat jet defined with radius $R$ using the C/A algorithm. We have to consider (a): the double-real (2R) and (b)+(c): one-real one-virtual (1R1V) contributions to the jet
mass distribution. One can ignore (d): the double virtual diagram because it does not contribute to the jet mass.

We start by writing the momenta of the two gluons $k_1$ and $k_2$:

\[
\begin{align*}
  k_1 &= x_1 E_j (1, 0, \sin \theta_1, \cos \theta_1), \\
  k_2 &= x_2 E_j (1, \sin \theta_2 \sin \phi, \sin \theta_2 \cos \phi, \cos \theta_2),
\end{align*}
\]

where we have defined $x_1$ and $x_2$ as the energy fractions of the fat jet energy, $E_j$, carried by $k_1$ and $k_2$ respectively. We consider the region $\Delta_{\theta_1} < \Delta_{\theta_2} < \Delta_R$ and $\Delta_{\theta_1} < \Delta_{\theta_{12}}$ where $\theta_{12}$ denotes the angle between $k_1$ and $k_2$. For this configuration, the C/A algorithm will first cluster $k_1$ to the hard parton $p$ and then cluster $k_2$ to the combination $(k_1 + p)$. We consider the action of mMDT algorithm when applied to this fat jet. The first step unclusters $k_2$ from the massive subjet $j_1$ that contains $p$ and $k_1$, and checks the pairwise energy asymmetry and mass drop conditions.
The mass drop condition is satisfied if $M_{j_1} < \mu M_j$, which translates into

$$(p + k_1)^2 < \mu (p + k_1 + k_2)^2.$$  

(3.16)

We can neglect terms that are bilinear in the soft parton momenta, $k_1$ and $k_2$, and write the constraint in Eq. (3.16) in terms of energy fractions and angles:

$$f x_1 \Delta \theta_1 < x_2 \Delta \theta_2,$$

(3.17)

where we have defined $f \equiv 1 - \mu$. We require that the splitting is not too energy asymmetric, which in the soft limit, is satisfied in the region $\frac{y_{\text{cut}}}{1 + y_{\text{cut}}} < x_2 < \frac{1}{1 + y_{\text{cut}}}$. Therefore, for all three particles to be accepted with given normalised jet mass, $\rho$, we impose the constraints:

$$\Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_2 \right) \Theta \left( x_2 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \Theta \left( x_2 \Delta \theta_2 - f x_1 \Delta \theta_1 \right) \delta \left( \rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2 \right),$$

(3.18)

where we have again neglected terms bilinear in soft parton momenta when defining the three-particle jet mass.

We now consider the possibility that either the mass drop or energy asymmetry cut fails. In this situation, mMDT removes the softer branch, $k_2$, and repeats the procedure for the subjet $j_1$. At this point, we essentially have the two particle LO configuration, therefore we check $k_1$ and $p$ for the energy asymmetry condition only. One should note that there is a possibility that $k_2$ is harder than subjet $j_1$, in this case mMDT removes $j_1$ and we obtain a zero mass jet. In order to avoid this configuration, we only examine the region in which $x_2 < \frac{1}{2}$. Putting this together, along with Eq. (3.18), we get the constraints on the double-real configuration:

$$\Theta^{2R} = \Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_2 \right) \Theta \left( x_2 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \Theta \left( x_2 \Delta \theta_2 - f x_1 \Delta \theta_1 \right) \delta \left( \rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2 \right) + \Theta \left( x_2 \Delta \theta_2 - f x_1 \Delta \theta_1 \right) \Theta \left( \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} - x_2 \right) + \Theta \left( f x_1 \Delta \theta_1 - x_2 \Delta \theta_2 \right) \Theta \left( \frac{1}{2} - x_2 \right) \times \Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_1 \right) \Theta \left( x_1 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \delta \left( \rho - x_1 \Delta \theta_1 \right),$$

(3.19)

where the top line corresponds to a final jet that contains three particles. The left and right terms on the second line correspond to emission $k_2$ failing the asymmetry cut and mass drop cut respectively. The last line constrains emission $k_1$ to pass the asymmetry cut, which is required for a non-zero final jet mass.

Now we consider the case of one-real one-virtual emission, and notice that we have the same kinematic configuration as the LO calculation. In order to obtain a non-zero
jet mass, we therefore have the following constraints on the real emission:

\[
\Theta^{\text{IR} \text{IV}} = -\Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_1 \right) \Theta \left( x_1 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \delta (\rho - x_1 \Delta \theta_1 ) \\
- \Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_2 \right) \Theta \left( x_2 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \delta (\rho - x_2 \Delta \theta_2 ),
\]

(3.20)

where the overall minus sign comes from the emission probability for a virtual gluon, which is otherwise identical to the corresponding real emission probability in the eikonal limit.

The complete NLO constraints are given by: \( \Theta^{\text{NLO}} = \Theta^{2\text{R}} + \Theta^{\text{IR} \text{IV}} \). The first contribution in Eq. (3.19) and the last in Eq. (3.20) cancel in the limit \( x_1 \) or \( \theta_1 \to 0 \), therefore this combination can only contribute non-logarithmically enhanced terms to the differential jet mass distribution. We can ignore these terms and write

\[
\Theta^{\text{NLO}} = -\Theta \left( \frac{1}{1 + y_{\text{cut}}} - x_1 \right) \Theta \left( x_1 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \delta (\rho - x_1 \Delta \theta_1 ) \\
\times \left[ \Theta \left( x_2 - \frac{y_{\text{cut}}}{1 + y_{\text{cut}}} \right) \Theta (x_2 \Delta \theta_2 - f x_1 \Delta \theta_1 ) + \Theta \left( x_2 - \frac{1}{2} \right) \Theta (f x_1 \Delta \theta_1 - x_2 \Delta \theta_2 ) \right].
\]

(3.21)

Computing the second term in square brackets with the constraint \( \Delta \theta_2 > \Delta \theta_1 \), one can show that this contribution is not logarithmically enhanced in \( \rho \), so we neglect it in further calculation. We now write an integral using the constraints in Eq. (3.21) and the two-gluon independent emission probability over the required phase space:

\[
\frac{1}{\alpha_s C_F} d\sigma = 4 \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int \frac{d \cos \theta_1}{1 - \cos^2 \theta_1} \frac{d \cos \theta_2}{1 - \cos^2 \theta_2} \frac{d \phi}{2\pi} dx_1 p_{gq}(x_1) dx_2 p_{gq}(x_2) \Theta^{\text{NLO}} \\
\times \Theta (\Delta \theta_2 - \Delta \theta_1 ) \Theta (\Delta \theta_{12} - \Delta \theta_1 ).
\]

(3.22)

As previously mentioned, we specify a kinematical configuration in which \( \theta_1 \) is the smallest angle, such that \( k_1 \) is clustered first by the C/A algorithm. The constraints on the second line originate from this angular restriction and we have inserted a factor of two to account for the configuration when \( \theta_2 \) is the smallest angle, which is otherwise identical by symmetry. We do not indicate explicitly, but we still impose the constraint that each gluon is emitted at an angle, \( \theta_i < R \) with respect to the hard quark.

We can simplify the integral in Eq. (3.22) by writing the self-clustering constraint, \( \Theta (\Delta \theta_{12} - \Delta \theta_1 ) \), in the collinear limit as \( \Theta (\theta_{12}^2 - \theta_1^2 ) = \Theta (\theta_{12}^2 - 4 \theta_1^2 \cos^2 \phi ) \), where we have used \( \theta_{12}^2 \approx \theta_1^2 + \theta_2^2 - 2 \theta_1 \theta_2 \cos \phi \). When we compare this constraint with the second angular requirement, \( \theta_2 > \theta_1 \), we find that we have two azimuthal regions to compute: \( \cos \phi > \frac{1}{2} \) and \( \cos \phi < \frac{1}{2} \). However, because the azimuthal integral is non-
3.4. MODIFIED MASS DROP TAGGER

singular, the overall result is that the limits on \( \theta_1 \) with respect to \( \theta_2 \) in each region are simply modified by a numerical factor, \( \mathcal{O}(1) \), after integration over \( \phi \). Therefore, the logarithmic structure of the final result is identical in each azimuthal region up to corrections below our level of accuracy. Therefore, it is equivalent to ignore this constraint and integrate over the full range of \( \phi \), which becomes trivial. Additionally, the configuration in which \( \theta_{12} \) is the smallest, leads to self-clustering of gluons, which does not contribute to the mMDT distribution to the accuracy we consider here. Note that this was not the case for MDT, where one could obtain leading contribution from following a massive, soft (wrong) branch.

We are interested in the leading terms of the NLO distribution, so we take the collinear limit of Eq. (3.22) via the replacement \( \cos \theta_i \simeq 1 - \theta_i^2/2 \), and compute the angular integrals:

\[
\frac{\rho}{\sigma} \frac{d\sigma^{(mMDT,C_2^F)}}{d\rho} = - \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int_{\frac{y_{\text{cut}}}{1+y_{\text{cut}}}}^{1} dx_2 p_{gq}(x_2) \int_{\frac{y_{\text{cut}}}{1+y_{\text{cut}}}}^{1} dx_1 p_{gq}(x_1) \times \left[ \Theta(x_2 - f x_1) \Theta \left( R^2 - \frac{\rho}{x_1} \right) \ln \frac{x_1 R^2}{\rho} + \Theta(f x_1 - x_2) \Theta \left( R^2 - \frac{\rho f}{x_2} \right) \ln \frac{x_2 R^2}{\rho f} \right]. \tag{3.23}
\]

We now evaluate the energy integrals of Eq. (3.23) in the region \( \rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} R^2 \):

\[
\frac{\rho}{\sigma} \frac{d\sigma^{(mMDT,C_2^F)}}{d\rho} = - \left( \frac{\alpha_s C_F}{\pi} \right)^2 \ln \left( \frac{1}{\frac{y_{\text{cut}}}{1+y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \right)^2}} \right) \times \left[ \ln \left( \frac{1 + y_{\text{cut}}}{y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \right)^2} \right) - \frac{y_{\text{cut}} (3 y_{\text{cut}} + 2)}{4 (1 + y_{\text{cut}})^2} \right] \ln \frac{1}{\rho}, \tag{3.24}
\]

where we have neglected subleading terms \( \mathcal{O}(\alpha_s^2) \) in the differential distribution. This result demonstrates that the mMDT distribution is single logarithmic in the \( C_2^F \) channel and gives a clear picture of the underlying physics. Like the leading order result, the logarithms from the collinear regions remain, but the integrals over each energy fraction are bounded in the soft limit by a cutoff in \( y_{\text{cut}} \). This acts by replacing the soft logarithms in the jet mass, \( \rho \), by ones in \( y_{\text{cut}} \). Therefore, if we lift the collinear approximation, contributions from wide-angle emissions provide terms that are subleading with respect to the single logarithmic accuracy we aim to calculate here. One should also note the absence of any dependence on the mass drop parameter, \( f \equiv \frac{1-\mu}{\mu} \), in Eq. (3.24). This indicates that the influence of the mass drop condition is subleading with respect to that of the asymmetry cut at this order. In Ref. [116], the authors use Monte Carlo techniques to demonstrate that the mass drop condition has little
Figure 3.6: Comparison of Event2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.24), for a range of $y_{\text{cut}}$ values. The red points show a straight line for small $\rho$, which indicates single logarithmic behaviour in the integrated distribution. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < \frac{y_{\text{cut}}}{(1+y_{\text{cut}})} R^2$, which demonstrates that we control all terms up to $\mathcal{O}(\alpha_s^2)$ in the differential distribution, for small $\rho$.

numerical effect on the mMDT jet mass distribution for $\mu \gtrsim 0.4$. This leads to the proposal of a simpler variant of mMDT in which one simply removes the mass drop condition.

We compare the analytical equation Eq. (3.24) to Event2, $C_F^2$ channel in Fig. 3.6. One can see from Event2 that mMDT contains leading single logarithms, and that the difference between this result and the analytic expression is constant for small $\rho$, which corresponds to leftover terms $\mathcal{O}(\alpha_s^2)$ in the differential distribution.

In summary, we have computed fixed-order results for mMDT at LO and NLO given by Eq. (3.9) and Eq. (3.24) respectively. We can take the small $y_{\text{cut}}$ limit of each result to give:

$$\frac{\rho d\sigma}{d\rho}^{(\text{mMDT,LO})} = \frac{\alpha_s C_F}{\pi} \ln \left( \frac{1}{y_{\text{cut}} e^{-\frac{3}{4}}} \right),$$

$$\frac{\rho d\sigma}{d\rho}^{(\text{mMDT,} C_F^2)} = -\left( \frac{\alpha_s C_F}{\pi} \right)^2 \ln \left( \frac{1}{y_{\text{cut}} e^{-\frac{3}{4}}} \right) \ln \frac{1}{\rho},$$

(3.25)
which gives a NLO expression that is consistent with exponentiation of the leading order result. We will return to this point in the summary of this section. We now consider the $C_F C_A$ and $C_F n_f$ contributions, which produce additional single logarithms.

3.4.3 Next-to-leading order: Flavour changing contributions

In this section, we will carry out a calculation of mMDT in the $C_F C_A$ and $C_F n_f$ channels. The first contribution comes from the 1-loop running coupling corrections to the LO result. We obtain this result by recalculating Eq. (3.8) with $\alpha_s$ evaluated at the transverse momentum, $k_t$, of the soft emission with respect to the hard emitter. Performing the angular integration and writing $k_t^2 = \rho x E_j^2$ in the collinear limit, we get the integral:

$$
\frac{\rho \, d\sigma}{\sigma \, d\rho} = \frac{C_F}{\pi} \int_{1+y_{\text{cut}}}^{1+y_{\text{cut}}} dx \, \rho \, p_{gq} \left( x \right) \alpha_s \left( \rho x E_j^2 \right).
$$

We can evaluate Eq. (3.26) using the 1-loop running coupling expression given in Eq. (1.26), to get the leading expression

$$
\frac{\rho \, d\sigma}{\sigma \, d\rho} = \frac{C_F \alpha_s}{\pi} \left( E_j^2 R^2 \right) \frac{1}{1 - \lambda} \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{1-y_{\text{cut}}}{1+y_{\text{cut}}} \right)} \right),
$$

where $\lambda = b_0 \frac{\alpha_s \left( E_j^2 R^2 \right)}{\pi} \ln \frac{1}{\rho}$. This gives a NLO contribution of the form

$$
\frac{\rho \, d\sigma}{\sigma \, d\rho} = C_F b_0 \left( \frac{\alpha_s}{\pi} \right)^2 \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{1-y_{\text{cut}}}{1+y_{\text{cut}}} \right)} \right) \ln \frac{1}{\rho},
$$

A second source of single logarithms arises when we have a kinematic configuration of the type depicted in the wrong branch configuration discussed in Subsection 3.3.3, shown in Fig. 3.3 whereby the partons $k_1$ and $k_2$ are the closest in angle. However, this time we consider the case where gluon $k$ is emitted with an energy fraction that is sufficiently hard, $x > \frac{1}{1+y_{\text{cut}}}$, for the quark to fail the asymmetry cut. mMDT follows the harder, massive branch containing the offspring of gluon $k$, rather than the quark, which leads to a flavour changing contribution. This is given by

$$
C_F \left( \frac{\alpha_s}{\pi} \right)^2 \int R^2 d\theta^2 \int_{1+y_{\text{cut}}}^{1+y_{\text{cut}}} dx \left( \frac{C_A}{2} p_{gg} \left( z \right) + T_R n_f p_{gq} \left( z \right) \right) \delta \left( \rho - z \left( 1 - z \right) x^2 \theta_{12}^2 \right),
$$

where we have taken the collinear limit, which is sufficient to capture the leading single
logarithmic terms we require. Evaluating the integral in the region $\rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} R^2$, for each channel, we get:

$$\frac{\rho \, d\sigma}{\sigma \, d\rho} (^{(\text{mMDT},C_FC_A)}_{}) = C_FC_A \left( \frac{\alpha_s}{\pi} \right)^2 \left[ \ln \left( \frac{1}{y_{\text{cut}}} \right) + \frac{11y_{\text{cut}}^3 + 9y_{\text{cut}}^2 - 9y_{\text{cut}} - 11}{12 \left( 1 + y_{\text{cut}} \right)^3} \right] \times \left[ \ln (1 + y_{\text{cut}}) + \frac{y_{\text{cut}} (2 + 3y_{\text{cut}})}{4 \left( 1 + y_{\text{cut}} \right)^2} \right] \ln \frac{1}{\rho},$$

(3.30)

$$\frac{\rho \, d\sigma}{\sigma \, d\rho} (^{(\text{mMDT},C_Fn_f)}_{}) = C_Fn_f \left( \frac{\alpha_s}{\pi} \right)^2 \frac{1 - y_{\text{cut}}^3}{6 \left( 1 + y_{\text{cut}} \right)^3} \left[ \ln (1 + y_{\text{cut}}) + \frac{y_{\text{cut}} (2 + 3y_{\text{cut}})}{4 \left( 1 + y_{\text{cut}} \right)^2} \right] \ln \frac{1}{\rho}.\quad (3.31)$$

Therefore, we conclude that mMDT is also single logarithmic in the $C_FC_A$ and $C_Fn_f$ channels unlike MDT, which contained an additional logarithm (see Eqs. $[3.13, 3.14]$). This is because the wrong branch effect is only initiated in mMDT by the branching of a hard emission, meaning the soft logarithm that was present in MDT has been replaced by one in $y_{\text{cut}}$. Additionally, the available phase space for this configuration vanishes in the limit $y_{\text{cut}} \to 0$, due to constraints on producing a sufficiently hard gluon to fail the energy asymmetry condition. For these reasons, combined with the common choice of $y_{\text{cut}} \sim 0.1$ in phenomenological studies, we expect the contributions to the mMDT jet mass from flavour changing contributions to be modest. These terms are explicitly included in the resummation of mMDT in Ref. $[116]$.

### 3.4.4 Non-global logarithms

We now examine the role of non-global logarithms $[166,170]$ in mMDT. As we have already seen, all leading logarithms in the jet mass at LO and NLO are purely collinear in origin. Non-global contributions only arise when we consider contributions from multiple, soft-wide angle emissions. Specifically, for plain jet mass, non-global logarithms originate from integration over the soft and wide-angle regions of phase space, whereby an emission outside the jet radius emits a soft parton into the jet.

One may anticipate that the mMDT mass distribution is free from non-global logarithms in $\rho$ because the integration over each soft energy fraction are cut off by the parameter $y_{\text{cut}}$. In order to show this explicitly, we consider the classic non-global configuration shown in Fig. $[3.3]$ where $k_1$ is emitted outside the jet and a soft gluon, $k_2$, is emitted inside the jet. We will only seek to compute the upper bound on the coefficient of the non-global logarithms because we do not explicitly consider the phase space available for self-clustering of gluons $k_1$ and $k_2$, when using the C/A algorithm. The inclusion of gluon self-clustering has been shown to reduce the coefficient of non-global logarithms in the $C_FC_A$ channel $[190, 191]$. However, because we seek only to
disprove the existence of logarithmically enhanced terms originating from non-global
configurations in this channel, this approximation will suffice. We therefore have the
following non-global constraints:

$$\Theta_{NG} = \Theta (x_1 - x_2) \Theta (\Delta \theta_1 - \Delta R) \Theta (\Delta R - \Delta \theta_2) \Theta \left( \frac{1}{1 + y_{cut}} - x_2 \right) \Theta \left( x_2 - \frac{y_{cut}}{1 + y_{cut}} \right),$$

(3.32)

where we have imposed (from left to right) that $k_1$ is more energetic than $k_2$, $k_1$ is
outside the jet, $k_2$ is within the jet and that $k_2$ survives the energy asymmetry cut
respectively.

We consider the constraints in Eq. (3.32) with the $C_F C_A$ correlated emission term
of the squared matrix element for two gluon emission. This gives the integral

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \int d\cos \theta_1 d\cos \theta_2 \frac{dx_1}{x_1} \frac{dx_2}{x_2} \Theta_{NG} \Omega_2 \delta (\rho - x_2 \Delta \theta_2),$$

(3.33)

where $\Omega_2$ is the angular function [170]:

$$\Omega_2 = \frac{2}{(1 - \cos \theta_1)(1 + \cos \theta_2) |\cos \theta_1 - \cos \theta_2|}. \quad (3.34)$$

Evaluating the integral in the region $\rho < \frac{y_{cut}}{1 + y_{cut}} \Delta R$, we obtain the result

$$\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \cot^2 \left( \frac{R}{2} \right) \frac{1 + y_{cut}}{4y_{cut}} \times \left( \ln \frac{1}{y_{cut}} - (1 - y_{cut}) (1 - \ln (1 + y_{cut})) \right). \quad (3.35)$$

As anticipated, the mMDT differential jet mass distribution has no non-global loga-
ithms in $\rho$ at $O(\alpha_s^2)$. However, if one was to perform a resummation of logarithms in
$\alpha_s y_{cut}$, these non-global effects would have to be included. We also expect that Abelian
clustering logarithms in the $C_F^2$ channel, which require at least two wide angle emissions
near the jet radius, to have no large logarithmic enhancements at this order.

We have explicitly shown that mMDT has no non-global logarithms at NLO, but
we believe that this behaviour extends to all orders due to the action of the asymmetry
cut on soft emissions. MDT is also free of non-global logarithms at $O(\alpha_s^3)$ because
the corresponding calculation is identical to mMDT. However, if we consider a subse-
quent soft-collinear branching of $k_2$ within the jet radius we obtain a wrong-branch-like
configuration. MDT can follow this massive, soft branch, introducing a non-global con-
tribution starting at 3 emissions of the form $O(\alpha_s^3 L^3)$, generalising to higher orders.
Figure 3.7: Comparison of EVENT2 in the $C_F C_A$ and $C_F n_f$ channels at NLO with the analytic expression Eq. (3.28), Eq. (3.30) on the left and Eq. (3.28), Eq. (3.31) on the right, for a range of $y_{\text{cut}}$ values. The red points indicate a straight line at small $\rho$, this corresponds to a leading single logarithmic result for the integrated distribution. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < \frac{y_{\text{cut}}}{1+y_{\text{cut}}} \Delta R$, which demonstrates that we control all terms up to $O(\alpha_s^2)$ in the differential distribution.

This is another reason to adopt mMDT over MDT.

We compare the analytical expressions Eq. (3.28), Eq. (3.30) and Eq. (3.31) to the EVENT2 program in the $C_F C_A$ and $C_F n_f$ channels in Fig. 3.7. One can see from the EVENT2 result that the leading behaviour of mMDT in the $C_F C_A$ and $C_F n_f$ channels is single logarithmic. The difference between this result and the corresponding analytic
3.4. MODIFIED MASS DROP TAGGER

expression is constant for small $\rho$, which corresponds to leftover terms $O(\alpha_s^2)$ in the differential distribution.

3.4.5 Phenomenology

In this section, we calculated that the mMDT jet mass distribution contains only single logarithmic terms at LO and NLO. This result is in contrast to the double logarithmic terms present in the plain jet mass distribution. By removing these large double logarithms, we have a distribution that does not exhibit a Sudakov peak in the background jet mass distribution. This feature may be beneficial, because peaks in the QCD jet mass distribution can be phenomenologically undesirable for background rejection.

Neglecting terms $O(y_{\text{cut}})$, we can write the coefficients as defined in Eq. (3.1), for mMDT in the region $\rho < y_{\text{cut}} \Delta_R$, as:

$$a_{11}^{\text{mMDT}} = C_F \ln \frac{e^{-3/4}}{y_{\text{cut}}},$$

$$a_{22}^{\text{mMDT}} = -C_F^2 \ln^2 \frac{e^{-3/4}}{y_{\text{cut}}} + C_F b_0 \ln \frac{e^{-3/4}}{y_{\text{cut}}},$$

where $a_{12}^{\text{mMDT}} = a_{24}^{\text{mMDT}} = a_{23}^{\text{mMDT}} = 0$.

The form of the coefficients in Eq. (3.36) suggests an exponentiation of the leading order result with running coupling effects included. For a detailed discussion, we refer...
the reader to Ref. [116], which builds on the results in this chapter to perform a resummation of each tagger to all orders, including finite \( y_{\text{cut}} \) effects. The fixed-order results in Eq. (3.36) are consistent with Ref. [116], in which the authors compute an all orders result in the small \( y_{\text{cut}} \) limit, which in a fixed-coupling approximation assumes the form

\[
\frac{1}{\sigma} \frac{d\sigma}{d\rho} (\text{mMDT,all-orders}) = \frac{d}{d\rho} \exp \left[ -\frac{C_F \alpha_s (E^2 R^2)}{\pi} \ln \left( \frac{1}{y_{\text{cut}} e^{-\frac{3}{4}}} \right) \frac{1}{\ln \frac{1}{\rho}} \right], \tag{3.37}
\]

where one has set the scale of the coupling proportional to the jet energy, in accordance with Ref. [166].

In Fig. 3.8 we provide the comparison of the full analytic resummed result 4 and the corresponding MC simulation from Ref. [116]. These show the differential distribution in \( \rho \) for highly boosted jets tagged with mMDT at parton level. From this figure, one can immediately see that all the features of the fixed-order calculation are manifest in the differential jet mass distribution. Specifically, the transition point at \( \rho \simeq y_{\text{cut}} \Delta R \) obtained at LO Eq. (3.9) and NLO Eq. (3.24) can be seen as a kink in the resummed and MC results. Additionally, the fixed-order calculation of single logarithmic behaviour in the region \( \rho < y_{\text{cut}} \Delta R \) translates into a resummed distribution without a Sudakov peak. For a logarithmic scale in \( \rho \) (as shown), this generates a simple straight line, the gradient of which depends on the value of \( y_{\text{cut}} \). Lastly, for \( \rho > y_{\text{cut}} \Delta R \) one observes an increase in the distribution that coincides with the behaviour of plain jet, also predicted by the fixed-order results.

The sum in the second line of Eq. (3.36) indicates that it is possible to tune the value of \( y_{\text{cut}} \) such that \( a_{\text{MDT}}^{2\text{MDT}} = 0 \). For \( n_f = 5 \), this corresponds to \( y_{\text{cut}} \simeq 0.11 \). At this value, the mMDT jet mass distribution could be approximated by only the LO result, leading to a flat distribution in \( \frac{\sigma}{\sigma} \frac{d\sigma}{d\rho} \) (see Fig. 3.8). Phenomenologically, this tuning would greatly simplify data driven background estimates if used in experimental studies. In other words, one can perform a background estimate by measuring the background jet mass distribution either side of an expected signal resonance and use this data to predict the background inside the signal peak region. Such estimates implicitly assume smoothness in the background distribution, hence taggers that produce a rich background structure complicate the determination of the interpolation. In some cases, non-trivial structure in the background may even be missed if it resides within the signal peak region, leading to poor estimates of the number of signal events. Hence, a tagger that exhibits a featureless background jet mass distribution with no transition point

4The analytic results, taken from Ref. [116], contain resummed running coupling and hard collinear contributions, unlike the simplified expression in Eq. (3.37). We show how to perform a more precise resummation, which includes these effects in Chapter 5. However, in this chapter, it suffices to compare the equivalent fixed-coupling resummed expressions, i.e. Eq. (3.37), with our fixed-order results.
kinks, greatly simplifies such estimates and can be reliably interpolated into the signal jet mass region.

Additionally, mMDT exhibits the remarkable and unique property that non-global and clustering logarithms are absent in the jet mass distribution for a single jet due to removal of all soft logarithms. The pure collinear nature of the logarithmic structure of mMDT means that we do not have to consider non-trivial, soft wide-angle colour structure when resumming this observable to all orders. Currently resummation of non-global logarithms is subject to corrections from the required large $N_c$ approximation \[170\], hence it is advantageous to eliminate these non-global contributions in precision calculations. In the next section, we consider a different algorithm: pruning.

3.5 Pruning

3.5.1 Definition

Algorithm 6: Pruning.

External parameters: $R_{\text{fact}}$, $z_{\text{cut}}$.

Input: Jet, $j$, with radius $R$ and constituents.

1. Define $R_{\text{prune}} = R_{\text{fact}} \frac{2M_j}{p_T}$;

2. Define a pairwise combination procedure PruningCombine as:
   - Label the pair of constituents considered for combination, $j_1$ and $j_2$, such that $p_{T_1} > p_{T_2}$;
   - if
     \[
     \frac{\min(p_{T_1}, p_{T_2})}{|p_{T_1} + p_{T_2}|} < z_{\text{cut}} \quad \text{and} \quad \Delta_{\theta_{12}} > R_{\text{prune}}^2
     \]
     then
     Remove $j_2$ from constituents;
   else
   Combine $j_1$ and $j_2$;

3. Recluster the constituents of jet $j$ using a generalised-$k_t$ algorithm but replace Combine with PruningCombine;

Output: Pruned jet.

The pruning algorithm \[108\] \[109\] begins with a hard jet $j$, with radius $R$ that has been clustered using the C/A algorithm \[84\] and is defined in Algorithm 6. Line 1 defines the dynamic pruning radius, which is determined according to the mass and transverse momentum of the fat jet and we set $R_{\text{fact}} = \frac{1}{2}$ henceforth. In line 2, we
redefine the pairwise recombination procedure used in the generalised-$k_t$ algorithm to ensure that recombination does not occur between sufficiently $p_T$ asymmetric (left inequality) and wide angle (right inequality) pairs. For this study, we shall use the C/A algorithm for reclustering and adapt Algorithm 6 for $e^+e^-$, replacing transverse momentum, $p_T$, with the relevant energy, $E_i$.

3.5.2 Leading order

The leading order calculation for pruning is fairly straightforward. We consider the fat jet to be comprised of a quark-gluon pair that arises from a $q \rightarrow qg$ branching, such that the gluon and quark carry energy fractions $x$ and $1 - x$ respectively. At this order, the pruning radius is given by $R_{\text{prune}}^2 = x(1 - x)\Delta_{\theta_{qg}}^2$, which is always less than $\Delta_{\theta_{qg}}$. As a result, the angular condition $\Delta_{\theta_{qg}} > R_{\text{prune}}^2$ is always satisfied at leading order. Therefore, we only check the energy asymmetry condition, which is satisfied in the region $z_{\text{cut}} < x < 1 - z_{\text{cut}}$.

The LO pruning differential jet mass distribution is therefore given by the following integral:

$$
\frac{1}{\sigma} \frac{d\sigma}{d\rho} = \frac{2\alpha_s C_F}{\pi} \int \frac{d \cos \theta}{1 - \cos^2 \theta} \int_{z_{\text{cut}}}^{1-z_{\text{cut}}} dx \, p_{gq}(x) \, \delta (\rho - x\Delta_{\theta}) \Theta (\Delta_R - \Delta_{\theta}) ,
$$

(3.38)

which, neglecting contributions of $O(\rho)$, evaluates to

$$
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} = \frac{\alpha_s C_F}{\pi} \ln \left( \frac{1 - z_{\text{cut}}}{z_{\text{cut}}} e^{-\frac{3}{4}(1-2z_{\text{cut}})} \right) \Theta (z_{\text{cut}}\Delta_R - \rho)
$$

$$
+ \frac{\alpha_s C_F}{\pi} \ln \left( (1 - z_{\text{cut}}) \frac{4\tan^2 \frac{R}{2}}{\rho} \right) \Theta (\rho - z_{\text{cut}}\Delta_R) .
$$

(3.39)

This result is similar to the LO (m)MDT expression given in Eq. (3.9), insofar that we have a single logarithmic distribution after replacement of a soft logarithm in $\rho$ by a logarithm in $z_{\text{cut}}$. Like $y_{\text{cut}}$, if $z_{\text{cut}} \sim O(1)$ these logarithms are not considered large for small $\rho$. One may notice that this result is identical to the (m)MDT expression under the replacement $z_{\text{cut}} \rightarrow \frac{y_{\text{cut}}}{1 + y_{\text{cut}}}$. This correspondence originates from different definitions of the asymmetry cut in each algorithm.

We check the analytic expression Eq. (3.39) against EVENT2 at LO, shown in Fig. 3.9. The difference between EVENT2 and the analytic result vanishes in the limit $\rho \rightarrow 0$, for all values of $z_{\text{cut}}$. This demonstrates that the analytical expression correctly reproduces the single logarithmic behaviour of the algorithm at this order.
Figure 3.9: Comparison of EVENT2 at LO with the analytic expression Eq. (3.39) for a range of $z_{cut}$ values. The red points indicate a constant in the small $\rho$ limit, which corresponds to a leftover $\mathcal{O}(\alpha_s)$ term in the differential distribution. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut}\Delta_R$, which tend to zero for each $z_{cut}$ value.

### 3.5.3 Next-to-leading order: Independent emission

In this section, we carry out a next-to-leading order calculation of the normalised differential pruning jet mass distribution in the Abelian $C_F^2$ channel. We concentrate on the logarithmic structure and drop contributions that are not enhanced in the limit $\rho \to 0$. Proceeding identically to the mass drop calculation, we consider the real and virtual emissions of gluons $k_1$ and $k_2$ in the independent emission approximation (see Fig. 3.5), which yields the leading singular behaviour in the small $\rho$ limit.

The definition of the pruning radius for double-real emission is given by

$$R^2_{prune} = \frac{M^2_{\text{jet}}}{{\Delta}_{\theta_j}} \approx x_1 \Delta_{\theta_1} + x_2 \Delta_{\theta_2}, \quad (3.40)$$

where $x_1$ and $x_2$ are the jet energy fractions carried by $k_1$ and $k_2$ respectively, and we again have neglected recoil of the parent parton. We find it convenient to divide the calculation into three separate phase space regions, each related to the pruning radius:
1. Both emissions are emitted with angle relative to the quark larger than the pruning radius: $\Theta(\Delta_{\theta_1} - R_{\text{prune}}^2) \Theta(\Delta_{\theta_2} - R_{\text{prune}}^2)$.

2. One emission is emitted with angle relative to the quark larger than the pruning radius: $\Theta(\Delta_{\theta_1} - R_{\text{prune}}^2) \Theta(R_{\text{prune}}^2 - \Delta_{\theta_2}) + \Theta(\Delta_{\theta_1} - R_{\text{prune}}^2) \Theta(R_{\text{prune}}^2 - \Delta_{\theta_2})$.

3. Neither emission is emitted with angle relative to the quark larger than the pruning radius: $\Theta(R_{\text{prune}}^2 - \Delta_{\theta_1}) \Theta(R_{\text{prune}}^2 - \Delta_{\theta_2})$.

Starting with the one-real one-virtual diagrams, we note that these configurations are kinematically identical to the LO case. Specifically, the value of $R_{\text{prune}}^2$ in each case is always less than the real emission angle $\Delta_{\theta_i}$. Hence, we only check the asymmetry condition for these contributions and there is no constraint on emission angle (other than $\Delta_R > \Delta_{\theta_i}$). Therefore, we can divide the angular phase space of these one-real one-virtual diagrams in precisely the same way as the double-real case, i.e. using the definition for $R_{\text{prune}}$ in Eq. (3.40), along with the angular regions defined above. This means that we can easily combine real and virtual contributions together in each region to cancel divergences.

Combining real and virtual contributions when both emissions are within the $R_{\text{prune}}$ radius (region 3 above), we get no logarithmic enhancements in $\rho$. This complete cancellation between real and virtual contributions is a direct consequence of the dynamical nature ($\rho$ dependence) of the pruning radius. When both emissions are outside the $R_{\text{prune}}$ radius (region 1 above), we again find no logarithmic enhancements in the small $\rho$ limit. We will comment on this region for non-vanishing $\rho$ in Subsection 3.5.6, but for now it suffices to discard these contributions for vanishing $\rho$.

The only remaining region to compute is a single emission outside the pruning radius (region 2 above). We consider the explicit case of $k_2$ emitted within the pruning radius and $k_1$ emitted at an angle greater than $R_{\text{prune}}$. Combining the real and virtual diagrams in this region we get

$$\Theta^{\text{NLO}} = \left[ \Theta(x_1 - z_{\text{cut}}) \delta(\rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2}) + \Theta(z_{\text{cut}} - x_1) \delta(\rho - x_2 \Delta_{\theta_2}) - \Theta(x_2 - z_{\text{cut}}) \delta(\rho - x_2 \Delta_{\theta_2}) - \Theta(x_1 - z_{\text{cut}}) \delta(\rho - x_1 \Delta_{\theta_1}) \right]$$

$$\times \Theta(\Delta_{\theta_1} - R_{\text{prune}}^2) \Theta(R_{\text{prune}}^2 - \Delta_{\theta_2}), \quad (3.41)$$

where the first line corresponds to the double-real emission contribution. Specifically, the left hand term corresponds to $k_1$ passing the asymmetry cut, in this situation both real gluons contribute to the jet mass. The right hand term arises when $k_1$ fails the energy cut and is removed by the algorithm; however, $k_2$ is retained because it is emitted within $R_{\text{prune}}$ and it alone contributes to the jet mass. The second line corresponds to
3.5. PRUNING

the contribution from the one-real one-virtual configuration. We require that the real emission satisfies $x_{1,2} > z_{\text{cut}}$ to obtain a non-zero jet mass, and the virtual emission provides an overall negative sign.

The first and last terms within the square brackets in Eq. (3.41) cancel in the limit $x_2$ or $\Delta\theta_2 \to 0$. Therefore, this combination gives only subleading contributions to the NLO integrated jet mass distribution. Eliminating these terms, we can rewrite the NLO constraints as

$$\Theta_{\text{NLO}} \simeq [\Theta (z_{\text{cut}} - x_1) - \Theta (x_2 - z_{\text{cut}})] \delta (\rho - x_2 \Delta\theta_2) \Theta (\Delta\theta_1 - R^2_{\text{prune}}) \Theta (R^2_{\text{prune}} - \Delta\theta_2),$$

we therefore have the following integral to compute:

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} (\text{Pruning}, C^2_F) = 4 \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int \frac{d\cos \theta_1}{1 - \cos^2 \theta_1} \frac{d\cos \theta_2}{1 - \cos^2 \theta_2} dx_1 p_{gq}(x_1) dx_2 p_{gq}(x_2) \Theta_{\text{NLO}}.$$

In order to simplify the calculation we perform the decomposition:

$$\frac{1}{1 - \cos^2 \theta_i} = \frac{1}{2} \left( \frac{1}{1 - \cos \theta_i} + \frac{1}{1 + \cos \theta_i} \right),$$

where the first term contains the leading collinear singularity and the second term is non-singular as $\theta_i \to 0$, which corresponds to wide-angle emissions. We only calculate leading, $O(\alpha_s^2 L^4)$, and next-to leading, $O(\alpha_s^2 L^3)$, contributions to the integrated jet mass distribution, so we need to evaluate three integrals, denoted with $I_i$. The leading contribution will be in the region $z_{\text{cut}} > x_1$:

$$I_1 = \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int \frac{d\cos \theta_1}{1 - \cos \theta_1} \frac{d\cos \theta_2}{1 - \cos \theta_2} dx_1 p_{gq}(x_1) dx_2 p_{gq}(x_2) \times \Theta (z_{\text{cut}} - x_1) \Theta (\Delta\theta_1 - R^2_{\text{prune}}) \Theta (R^2_{\text{prune}} - \Delta\theta_2) \delta (\rho - x_2 \Delta\theta_2),$$

and the wide-angle $k_1$ contribution will contain next-to leading terms:

$$I_2 = \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int \frac{d\cos \theta_1}{1 + \cos \theta_1} \frac{d\cos \theta_2}{1 - \cos \theta_2} dx_1 x_1 dx_2 p_{gq}(x_2) \times \Theta (z_{\text{cut}} - x_1) \Theta (\Delta\theta_1 - R^2_{\text{prune}}) \Theta (R^2_{\text{prune}} - \Delta\theta_2) \delta (\rho - x_2 \Delta\theta_2).$$

We have used only the soft pole of the splitting function ($1/x_1$) for the wide angle emission of $k_1$, which is sufficient for the accuracy we seek. Lastly, we have a next-to
leading contribution from the region $x_2 > z_{\text{cut}}$:

$$I_3 = \left( \frac{\alpha_s C_F}{\pi} \right)^2 \int \frac{d \cos \theta_1}{1 - \cos \theta_1} \frac{d \cos \theta_2}{1 - \cos \theta_2} dx_1 p_{gq} (x_1) dx_2 p_{gq} (x_2) \
\times \Theta (x_2 - z_{\text{cut}}) \Theta (\Delta \theta_1 - R_{\text{prune}}^2) \Theta (R_{\text{prune}}^2 - \Delta \theta_2) \delta (\rho - x_2 \Delta \theta_2).$$

(3.47)

Evaluating these integrals, the sum $I_1 + I_2 + I_3$ gives the final result in the region $\rho < z_{\text{cut}} \Delta R$:

$$\frac{\rho}{\sigma} \frac{d \sigma}{d \rho}^{(\text{Pruning}, C_F^2)} = \left( \frac{\alpha_s C_F}{\pi} \right)^2 \left( \frac{1}{6} \ln^3 \frac{1}{\rho} \
+ \left[ \frac{1}{2} \ln \frac{z_{\text{cut}}^2}{1 - z_{\text{cut}}} - \frac{5}{4} z_{\text{cut}} + \frac{1}{8} z_{\text{cut}}^2 + \frac{1}{2} \ln \left( 4 \tan^2 \frac{R}{2} \right) \right] \ln^2 \frac{1}{\rho} \right),$$

(3.48)

where we have neglected terms $O (\alpha_s^2 L)$ in the differential mass distribution.

This result is remarkable, we obtain double logarithmic behaviour at NLO, which is in contrast to the LO pruning calculation, Eq. (3.39), and the results obtained for (m)MDT. This result implies that, starting at NLO, the pruning algorithm does not remove all soft gluon contributions as was possibly intended. We instead find that pruning is as singular as the plain jet mass distribution (albeit with a smaller coefficient) in the small $\rho$ limit. This leading term corresponds explicitly to the configuration in which a soft gluon dominates the jet mass, but is pruned away by the asymmetry condition, leaving a pruned jet with mass dominated by soft-collinear substructure that is never checked by the asymmetry condition (denoted part of the “I-pruned” class in Ref. [116]). Additionally, one would expect the presence of non-global and Abelian clustering logarithms in the pruned jet mass distribution, further complicating calculation of pruning at single logarithmic accuracy. We explicitly calculate some of these single logarithmic terms due to non-global effects in the next subsection.

We now check Eq. (3.48) against results from EVENT2 in the $C_F^2$ channel, shown in Fig. 3.10. As predicted, pruning contains terms that are more divergent than single logarithms at NLO in this channel. By considering the difference between the analytical result and EVENT2, we conclude that we successfully control all leading and next to leading terms at NLO. The leftover straight line corresponds to uncalculated single logarithmic contributions to the differential distribution. In the next section, we explicitly calculate the leading contributions in the $C_F C_A$ channel and demonstrate the presence of non-global terms.
3.5. PRUNING

Figure 3.10: Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.48) for a range of $z_{cut}$ values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut} \Delta_R$, which gives a straight line for small $\rho$, indicating we control all terms up to $O(\alpha_s^2 L)$ in the differential distribution.

3.5.4 Next-to-leading order: Non-Abelian terms

We now consider contributions to the pruning mass distribution in the $C_F C_A$ channel that result from secondary gluon splitting. These include running coupling effects (which dress the LO result), non-global logarithms and configurations in which one examines the mass of the gluon pair after removal of the hard quark (analogous to the wrong branch effect seen in (m)MDT). The latter will occur when the energy carried away by the primary gluon is sufficient to cause the hard quark energy fraction to drop below $z_{cut}$, such that the pruned jet mass is defined only by the gluon jet. This will lead to contributions of the form $O(\alpha_s^2 L^3)$, in the integrated distribution because we only require the first emission is sufficiently hard. Running coupling and non-global effects will contribute at the $O(\alpha_s^2 L^2)$ level to the integrated distribution. Hence, we do not explicitly treat the running coupling effects here, but, as before, we compute the upper bound on the non-global logarithms in order to demonstrate their existence and contrast them with the (m)MDT results.
Starting by computing the $O(\alpha_s^2 L^3)$ terms, we consider the configuration given in Fig. 3.3, where gluon $k$ branches into collinear gluon pair $k_1$ and $k_2$. In order to remove the quark, the primary gluon has to be sufficiently hard to fail the asymmetry condition, this is satisfied in the region $x > 1 - z_{\text{cut}}$, where $x$ is the energy fraction carried by the parent gluon, $k$. In this kinematic configuration, the pruning algorithm discards the quark and we study the mass distribution of the gluon jet. Working in the collinear limit, we can write the leading integral for this wrong branch contribution:

$$
\frac{1}{\sigma} \frac{d\sigma}{d\rho}^{(\text{Pruning},C_F C_A)} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \int \frac{d\theta^2}{\theta^2} \int_{1-z_{\text{cut}}}^{R^2} dx \ p_{gg}(x) \times \int \frac{d\theta_{12}^2}{\theta_{12}^2} \int_1^{\rho} \frac{dz}{z} \delta \left( \rho - z \frac{x^2 \theta_{12}^2}{z_{\text{cut}}} \right),
$$

(3.49)

where we have assumed that the gluon splitting is soft, $z \ll 1$. We have imposed that gluons $k_1$ and $k_2$ are clustered first by the C/A algorithm by placing the constraint: $\theta_{12}^2 < \theta^2$, where $\theta$ is the angle of gluon $k$ with respect to the quark. We evaluate this integral in the small $\rho$ limit to get

$$
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(\text{Pruning},C_F C_A)} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \frac{1}{2} \left( \ln \frac{1}{1-z_{\text{cut}}} - \frac{z_{\text{cut}}}{4} (z_{\text{cut}} + 2) \right) \ln \frac{1}{\rho},
$$

(3.50)

where we have neglected subleading terms, $O(\alpha_s^2 L)$, in the differential distribution.

We check Eq. (3.50) against EVENT2 in the $C_F C_A$ channel in Fig. 3.11. We see from the EVENT2 result that pruning contains terms more divergent than single logarithms, and the relative magnitude of these contributions decrease with $z_{\text{cut}}$. This is because the available phase space for the quark to fail the asymmetry cut diminishes in the limit $z_{\text{cut}} \to 0$. Subtraction of the analytical result from EVENT2 reveals that we successfully control up to $O(\alpha_s^2 L)$ in the differential distribution, and we are left with a straight line corresponding to leftover single logarithmic contributions.

We now examine the non-global contributions to pruning. Firstly, we reiterate that pruning does not check for asymmetry of particle pairs that have an opening angle less than $R_{\text{prune}}$; for this reason, we expect non-global logarithms to be present in pruning because one does not have a soft cutoff in some collinear regions of phase space. Explicitly, we consider the configuration given in Fig. 3.3, whereby gluon $k_3$ lies outside $R_{\text{prune}}$. If this gluon has an energy fraction less than $z_{\text{cut}}$, it is removed by pruning and does not contribute to the jet mass. However, gluon $k$ can radiate a soft gluon $k_2$ with an angle relative to the quark less than $R_{\text{prune}}$, i.e., into the core of the pruned jet. This is analogous to the non-global contribution to the plain jet mass under the replacement $R \to R_{\text{prune}}$. The soft emission, $k_2$, is always retained because it is within the pruning radius and can therefore contribute large logarithms to the jet mass.
3.5. PRUNING

![Graphs showing coefficient of C_F C_A for pruning](image)

Figure 3.11: Comparison of Event2 in the C_F C_A channel at NLO with the analytic expression Eq. (3.50) for a range of z_{cut} values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the Event2 result in the region ρ < z_{cut}∆_R, which gives a straight line for small ρ, indicating we control all terms up to O(α_s^2 L) in the differential distribution.

mass distribution. In the region, θ_{12} < θ, self-clustering of gluons k_1 and k_2 using the C/A algorithm will eliminate this non-global contribution [172, 191, 192]. Hence, we require θ_{12} > θ such that k_2 clusters to the hard parton first. We can therefore write the phase space that contributes non-global logarithms as

\[ \Theta^{NG} = \Theta (x_1 - x_2) \Theta (\Delta_\theta_1 - R_{prune}^2) \Theta (R_{prune}^2 - \Delta_\theta_2) \Theta (z_{cut} - x_1) \Theta (\Delta_{\theta_{12}} - \Delta_\theta_2), \]

(3.51)

where R_{prune}^2 is defined as Eq. (3.40).

We now compare the phase space limits imposed by the no self-clustering constraint, \( \theta_{12} > \theta_2 \), and the requirement \( R_{prune} > \theta_2 \) in Eq. (3.51). Using the expression for \( R_{prune} \) given in Eq. (3.40) in the small angle limit, and the collinear approximation \( \theta_{12}^2 \approx \theta_1^2 + \theta_2^2 - 2\theta_1\theta_2 \cos \phi \), where \( \phi \) is the azimuthal angle between \( k_1 \) and \( k_2 \) with
respect to the hard axis, we find that self-clustering is absent in the region
\[ x_1 < \frac{1 - x_2}{4 \cos^2 \phi}. \] (3.52)

Taking the soft limit, \( x_2 \to 0 \), and imposing that \( x_1 < z_{\text{cut}} \), we can see that self-clustering is always absent when \( z_{\text{cut}} < \frac{1}{4} \) is satisfied. In this region, we can ignore the \( \phi \) dependence in what follows and integrate the squared matrix element for correlated two gluon emission:

\[
\frac{1}{\sigma} \frac{d\sigma^{(\text{Pruning,NG})}}{d\rho} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \int d\cos \theta_1 \, d\cos \theta_2 \frac{dx_1}{x_1} \frac{dx_2}{x_2} \Theta_{\text{NG}} \Omega_2 \delta (\rho - x_2 \Delta \theta_2),
\] (3.53)

where \( \Omega_2 \) is an angular function defined in Eq. (3.34). As discussed, we ignore the no self-clustering constraint \( \Delta \theta_{12} > \Delta \theta_2 \) in \( \Theta_{\text{NG}} \) and take the small angle limit, which is sufficient to capture the leading behaviour, we evaluate the integral in the region \( \rho < z_{\text{cut}}^2 R^2 \) to get the leading result:

\[
\frac{\rho}{\sigma} \frac{d\sigma^{(\text{Pruning,NG})}}{d\rho} = C_F C_A \left( \frac{\alpha_s}{\pi} \right)^2 \text{Li}_2 (z_{\text{cut}}) \ln \frac{1}{\rho},
\] (3.54)

where \( \text{Li}_2 (x) \) is the dilogarithm function

\[
\text{Li}_2 (x) \equiv -\int_0^x \frac{\ln (1 - y)}{y} dy.
\] (3.55)

Hence, we have shown that pruning exhibits single logarithmic, non-global contributions to the jet mass distribution at NLO. However, the dilogarithmic coefficient \( (\text{Li}_2 (z_{\text{cut}}) \simeq z_{\text{cut}} \text{ for small } z_{\text{cut}}) \) means that the magnitude of this contribution vanishes in the limit \( z_{\text{cut}} \to 0 \). When we compute the coefficient for a typical value \( z_{\text{cut}} \sim 0.1 \), we find that the size of the non-global contribution is roughly 10% of the plain jet mass non-global coefficient, \( \pi^2/3 \), reported in Ref. [166]. One should note that the estimation of the non-global logarithms in Eq. (3.54) applies at NLO only, to get a full picture, one would have to consider higher order contributions to assess the overall impact of non-global logarithms on the pruned jet mass distribution.

### 3.5.5 Y-pruning

We now calculate the NLO fixed-order expression for a variant of pruning, Y-pruning, proposed in Ref. [116]. This modification eliminates the double logarithmic terms present in the \( C_F^2 \) channel, calculated in Eq. (3.48). This is achieved by introducing an additional condition, Algorithm 7, to the end of the pruning procedure defined in Algorithm 6.
3.5. PRUNING

Algorithm 7: Y-pruning modification.

4 if the clustering history of \( j \) contains a PruningCombine in which \( \Delta \theta_{12} > R_{\text{prune}}^2 \) and \( \min(\frac{p_{T1}, p_{T2}}{|p_{T1} + p_{T2}|}) > z_{\text{cut}} \) are satisfied then
   Label jet \( j \) a Y-pruned jet;
else
   Discard jet \( j \);
Output: Y-pruned jet.

This condition ensures that the fat jet mass was dominated by hard radiation, such that the pruning radius, \( R_{\text{prune}} \), was set appropriately. This hard substructure “two-prong” requirement discards pruned jets with masses that are dominated by arbitrarily soft and collinear radiation, thereby removing the double logarithmic contribution.

To demonstrate this explicitly, we go back to the \( C_F^2 \) pruning constraints, Eq. (3.42), for two emissions: \( k_1 \) outside and \( k_2 \) within the pruning radius. For the double-real emission case, we require that \( k_1 \) is sufficiently hard, \( x_1 > z_{\text{cut}} \), to satisfy the Y-pruning requirement. This contradicts the first step function, \( \Theta (z_{\text{cut}} - x_1) \), and we are only left with the one-real one-virtual contribution:

\[
\Theta^{\text{NLO,Y-pruning}} \sim -\Theta (x_2 - z_{\text{cut}}) \delta (\rho - x_2 \Delta \theta_2) \Theta \left( \Delta \theta_1 - R_{\text{prune}}^2 \right) \Theta \left( R_{\text{prune}}^2 - \Delta \theta_2 \right),
\]

(3.56)

where \( k_2 \) provides the necessary hard substructure to pass the Y-pruning condition. This contribution corresponds to the integral \( I_3 \) alone, given in Eq. (3.47). We evaluate this integral explicitly in the region \( \rho < z_{\text{cut}} \Delta R \) to get

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(\text{Y-Pruning}, C_F^2)} = -\left( \frac{\alpha_s C_F}{\pi} \right)^2 \left( \frac{1}{2} \ln \frac{1 - z_{\text{cut}}}{z_{\text{cut}}} - \frac{3}{4} z_{\text{cut}} + \frac{3}{8} \right) \ln^2 \frac{1}{\rho},
\]

(3.57)

where we have neglected terms \( \mathcal{O}(\alpha_s^3 L) \) in the differential distribution. This shows that Y-pruning has one less logarithm in \( \rho \) at NLO when compared to pruning or the plain jet mass, i.e. the leading behaviour of Y-pruning is \( \mathcal{O}(\alpha_s^3 L^2) \) in the integrated jet mass distribution.

We compare the analytical expression for Y-pruning, Eq. (3.57), with EVENT2 in the \( C_F^2 \) channel at NLO in Fig. 3.12. We observe that the EVENT2 result contains terms that are more divergent than single logarithms. After subtracting the analytical result, we find that we control the leading divergence in the differential distribution, \( \mathcal{O}(\alpha_s^2 L^2) \), and are left with a straight line corresponding to leftover \( \mathcal{O}(\alpha_s^2 L) \) terms.
Figure 3.12: Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.57) for a range of $z_{cut}$ values. The red points demonstrate the presence of terms at NLO that are more divergent than single logarithms. The green points demonstrate the difference between the analytic expression and the EVENT2 result in the region $\rho < z_{cut}\Delta R$, which gives a straight line for small $\rho$, indicating we control terms up to $O(\alpha_s^2 L)$ in the differential distribution.

3.5.6 Pruning in the region $\rho > z_{cut}^2 \Delta_R$

So far we have been working exclusively in the small $\rho$ limit and established double logarithmic behaviour for pruning in the $C_F^2$ channel in the region $\rho < z_{cut}\Delta_R$ (see Eq. (3.48)). However, this leading behaviour turns out to have a smaller range of validity than anticipated if we consider the jet mass distribution for non-vanishing values of $\rho$. In the range, $z_{cut}^2 \Delta_R < \rho < z_{cut}\Delta_R$, we calculate some additional logarithmic contributions from the region where both gluons are outside the pruning radius that cancel the leading double logarithmic term.

Explicitly, we consider the double-real and one-real one-virtual contributions in the
3.5. PRUNING

region $\Delta_{\theta_1}, \Delta_{\theta_2} > R_{\text{prune}}^2$. In this region we have the following constraints:

$$\Theta^{\text{NLO,large}} \rho = \Theta (x_1 - z_{\text{cut}}) \Theta (x_2 - z_{\text{cut}}) \delta (\rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2})$$

$$+ \Theta (x_1 - z_{\text{cut}}) \Theta (z_{\text{cut}} - x_2) \delta (\rho - x_1 \Delta_{\theta_1})$$

$$+ \Theta (z_{\text{cut}} - x_1) \Theta (x_2 - z_{\text{cut}}) \delta (\rho - x_2 \Delta_{\theta_2})$$

$$- \Theta (x_1 - z_{\text{cut}}) \delta (\rho - x_1 \Delta_{\theta_1}) - \Theta (x_2 - z_{\text{cut}}) \delta (\rho - x_2 \Delta_{\theta_2}), \quad (3.58)$$

where the first three lines correspond to the double-real emission contribution; in this configuration, whenever we prune away an emission with energy fraction $x_i < z_{\text{cut}}$, the final jet mass is set by the remaining emission. The last line corresponds to the one-real one-virtual contributions to the pruned jet mass, in this line we only require that the real emission, $i$, has an energy fraction $x_i > z_{\text{cut}}$ to obtain a non-zero jet mass. After some rearranging this simplifies to

$$\Theta^{\text{NLO,large}} \rho = \Theta (x_1 - z_{\text{cut}}) \Theta (x_2 - z_{\text{cut}})$$

$$\times \left[ \delta (\rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2}) - \delta (\rho - x_1 \Delta_{\theta_1}) - \delta (\rho - x_2 \Delta_{\theta_2}) \right]. \quad (3.59)$$

We are only interested in the leading term, so we take the soft and collinear limit and write the integral, denoted $I_4$, as

$$I_4 = \left( \frac{\alpha_s C_F}{\pi} \right)^2 \frac{1}{2} \int \frac{dx_1}{x_1} \frac{dx_2}{x_2} d\theta_1^2 d\theta_2^2 \Theta (\theta_1^2 - R_{\text{prune}}^2) \Theta (\theta_2^2 - R_{\text{prune}}^2) \Theta^{\text{NLO,large}} \rho, \quad (3.60)$$

where in the collinear limit: $R_{\text{prune}}^2 = x_1 \theta_1^2 + x_2 \theta_2^2$. Evaluation of this integral is straightforward; there are no logarithmically enhanced terms for $\rho < z_{\text{cut}}^2 \Delta R$, but in the region $z_{\text{cut}}^2 \Delta R < \rho < z_{\text{cut}} \Delta R$ we get the following contribution:

$$I_4 = \frac{1}{\rho} \left( \frac{\alpha_s C_F}{\pi} \right)^2 \left[ \frac{1}{6} \ln^3 \frac{R_{\text{cut}}^2}{\rho} + \mathcal{O} \left( \ln \frac{1}{\rho} \right) \right], \quad (3.61)$$

which cancels the leading double logarithms in Eq. (3.48). If we include the next-to-leading contributions, we find that the final sum, $I_1 + I_2 + I_3 + I_4$, in this region is single logarithmic:

$$\frac{\rho d\sigma}{\sigma d\rho}^{\text{(Pruning,C}_{\text{F}}^2)} = - \left( \frac{\alpha_s C_F}{\pi} \right)^2 \ln^2 \frac{1}{z_{\text{cut}}} \ln \frac{1}{\rho}, \quad z_{\text{cut}}^2 \Delta R < \rho < z_{\text{cut}} \Delta R. \quad (3.62)$$

This result coincides with the expression for mMDT in the soft and collinear limit, Eq. (3.25), and indicates that mMDT and pruning have the same leading single logarithmic behaviour in this region.
3.5.7 Phenomenology

In this section, we examined the LO and NLO behaviour of the pruning algorithm and revealed logarithmic structure that is more complex than mMDT. At LO we found that pruning is single logarithmic and is equal to mMDT under a change of variables. At NLO we discovered that pruning contains double logarithmic contributions, $O(\alpha_s^2 L^4)$, to the integrated jet mass distribution, which arises from the dynamical nature of the pruning radius. Specifically, this occurs when a soft gluon dominates the fat jet mass, setting $R_{\text{prune}}$, but is removed by the pruning algorithm.

We now write down the logarithmic coefficients for pruning in the small $z_{\text{cut}}$ limit and in the region $\rho < z_{\text{cut}} \Delta_R$, as defined in Eq. (3.1):

\begin{align}
a_{11}^{\text{Pruning}} &= C_F \ln \frac{e^{-3/4}}{z_{\text{cut}}}, \\
a_{24}^{\text{Pruning}} &= \frac{1}{6} C_F^2 \Theta \left( z_{\text{cut}}^2 \Delta_R - \rho \right), \\
a_{23}^{\text{Pruning}} &= -\frac{1}{2} C_F^2 \ln \frac{1}{z_{\text{cut}}} \Theta \left( z_{\text{cut}}^2 \Delta_R - \rho \right),
\end{align}

where $a_{12}^{\text{Pruning}} = 0$ and we have not reported $a_{22}^{\text{Pruning}}$, but leave these NLO single logarithmic contributions to further work. We re-emphasise that in the region $z_{\text{cut}}^2 \Delta_R < \rho < z_{\text{cut}} \Delta_R$, the coefficients $a_{24}^{\text{Pruning}} = a_{23}^{\text{Pruning}} = 0$ and we are left with a purely single logarithmic result up to NLO.

We also calculated fixed-order expressions for Y-pruning. We showed that by retaining a subset of pruned jets that contain at least one sufficiently hard, wide-angle emission, we can eliminate the double logarithmic term. The coefficients for Y-pruning in the small $z_{\text{cut}}$ limit therefore read:

\begin{align}
a_{11}^{Y\text{-Pruning}} &= C_F \ln \frac{e^{-3/4}}{z_{\text{cut}}}, \\
a_{24}^{Y\text{-Pruning}} &= 0, \\
a_{23}^{Y\text{-Pruning}} &= -\frac{1}{2} C_F^2 \ln \frac{e^{3/4}}{z_{\text{cut}}},
\end{align}

The fixed-order results in Eq. (3.64) are consistent with Ref. [116], in which the authors compute an all orders result in the small $z_{\text{cut}}$ limit of the form

\begin{equation}
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y\text{-pruning,all-orders})} \simeq \frac{\alpha_s C_F}{\pi} \ln \frac{1}{z_{\text{cut}}} \times \exp \left[ -\frac{\alpha_s C_F}{2\pi} \ln^2 \frac{1}{\rho} \right].
\end{equation}

The presence of double logarithms in pruning will give rise to a Sudakov peak in the resummed expression for the differential jet mass distribution (see Fig. 3.13). The transition from single logarithmic to double logarithmic behaviour occurs at $\rho \simeq z_{\text{cut}}^2 \Delta_R$.
3.5. PRUNING

Figure 3.13: Comparison of the differential distribution in $\rho$ for pruning and Y-pruning tagged jets using Monte Carlo (left) and full resummed analytical results (right). Details of MC generation are given in Fig. 3.1. The arrows denote the expected position of the transition points, as derived by the analytical calculations in this section, i.e. $\rho \simeq z_{cut}^2 \Delta R$ (black) and $\rho \simeq z_{cut} \Delta R$ (grey). The I-pruning distribution corresponds to all contributions that do not satisfy the Y-pruning constraints, hence pruning is the sum of Y- and I-pruning. This figure and caption has been adapted directly from Ref. [116].

(see Eq. (3.62)); for high $p_T$ jets, this transition in the background may coincide with signal jets that have masses around the electroweak scale. For example, the transition occurs in the vicinity of the $Z$ boson mass for 1 TeV jets with $R = 1$ and $z_{cut} = 0.1$ (at $M_j \simeq 96$ GeV). As discussed previously, this non-smoothness in the background jet mass distribution can be undesirable for phenomenological studies that use data driven background estimates.

In order to see this explicitly, in Fig. 3.13 we provide the comparison of the full analytic resummed result and the corresponding MC simulation from Ref. [116]. These show the differential distribution in $\rho$ for highly boosted jets tagged with pruning and Y-pruning at parton level. One observes that the features obtained with fixed-order calculations emerge in the differential jet mass distribution. For pruning, the NLO result in Eq. (3.48), combined with Eq. (3.61), predicted a Sudakov peak from the double logarithmic behaviour in the region $\rho < z_{cut}^2 \Delta R$. In the MC results (left panel), the transition point at $\rho \simeq z_{cut}^2 \Delta R$ is labelled with a black arrow and can be clearly seen as the onset of a shifted Sudakov peak towards small $\rho$. The result in Eq. (3.62) predicted a flat, single logarithmic distribution in the region $z_{cut}^2 \Delta R < \rho < z_{cut} \Delta R$, which is also manifest in the MC results. Finally, we calculated a transition point at $\rho \simeq z_{cut} \Delta R$ to the plain jet mass distribution at large $\rho$, this is labelled with a grey
Y-pruning was shown to be the subset of the pruning contribution that does not contain the double logarithmic contributions, hence for Y-pruning we do not predict a transition point at $\rho \simeq z_{cut}^2 \Delta R$ or a Sudakov peak. These analytic features can be seen in Fig. 3.13 as a single transition and a suppression for small $\rho$, as suggested by Eq. (3.65). By removing the double logarithmic contribution but not the next-to double logarithms, i.e. $O(\alpha_s^2 L ^{2n-1})$ in the perturbative expansion as seen in Eq. (3.57), this algorithm exhibits favourable background rejection rates compared to other taggers studied here. This important feature will be revisited in the next chapter.

In the next section, we consider a different algorithm: trimming.

### 3.6 Trimming

#### 3.6.1 Definition

The trimming algorithm is applied to a hard jet, $j$, with radius $R$ and is defined in Algorithm 8.

**Algorithm 8:** Trimming.

<table>
<thead>
<tr>
<th>External parameters: $R_{\text{trim}}$, $f_{\text{cut}}$.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> Jet, $j$ with constituents.</td>
</tr>
<tr>
<td>1 <strong>Recluster</strong> the constituents of $j$ with a generalised-$k_t$ algorithm using radius $R_{\text{trim}}$ into subjets $i$;</td>
</tr>
<tr>
<td>2 <strong>Combine</strong> all subjets that satisfy $p_{T,i} &gt; f_{\text{cut}} p_{T,j}$ and <strong>Relabel</strong> the combination as $j$;</td>
</tr>
<tr>
<td>3 <strong>Label</strong> jet $j$ as the trimmed jet;</td>
</tr>
</tbody>
</table>

**Output:** Trimmed jet.

The trimming algorithm reclusters the jet into subjets using trimming radius $R_{\text{trim}} < R$ in line 1 and retains those that carry at least a fraction $f_{\text{cut}}$ of the total jet transverse momentum in line 2. One can use any clustering algorithm to find and/or cluster jet $j$ or subjets $i$ in line 1, but for this chapter we will use C/A throughout. For this study, we shall use an $e^+e^-$ adaptation of Algorithm 8 whereby we replace transverse momentum $p_{T,i}$ with the relevant energy $E_i$.

#### 3.6.2 Leading order

At LO, we consider a fat jet comprised of a quark-gluon pair resulting from a $q \rightarrow gg$ branching, such that the gluon and quark carry energy fractions $x$ and $1-x$ respectively. For emission angle $\Delta \theta < \Delta R_{\text{trim}}$, the reclustering procedure produces a single subjet that trivially passes the $p_T$ condition. However, if $\Delta \theta > \Delta R_{\text{trim}}$, we obtain two massless subjets that only pass the $p_T$ condition in the region $f_{\text{cut}} < x < 1 - f_{\text{cut}}$. 

126
This gives the following integral for the trimming LO jet mass distribution:

\[
\frac{1}{\sigma} \frac{d\sigma}{d\rho} \left( \text{Trimming,LO} \right) = 2 \frac{\alpha_s C_F}{\pi} \int \frac{d\cos \theta}{1 - \cos^2 \theta} \int dx p_{yy} (x) \delta (\rho - x(1 - x)\Delta_\theta) \Theta (\Delta_R - \Delta_\theta) 
\times \left[ \Theta (\Delta_{R_{\text{trim}}} - \Delta_\theta) + \Theta (\Delta_\theta - \Delta_{R_{\text{trim}}}) \Theta (x - f_{\text{cut}}) \Theta (1 - f_{\text{cut}} - x) \right].
\]

(3.66)

We can compute the integral in Eq. (3.66), using the angular decomposition given in Eq. (3.44) and neglecting contributions of \( O(\rho) \) to get

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \left( \text{Trimming,LO} \right) = \frac{\alpha_s C_F}{\pi} \left[ \ln \left( \frac{4 \tan^2 \frac{R_{\text{trim}}}{2}}{\rho e^{-\frac{3}{4}}} \right) \Theta (f_{\text{cut}} \Delta_{R_{\text{trim}}} - \rho) 
+ \ln \left( \frac{1 - f_{\text{cut}}}{f_{\text{cut}}} e^{-\frac{3}{4}(1 - 2f_{\text{cut}})} \right) \Theta (\rho - f_{\text{cut}} \Delta_{R_{\text{trim}}}) \Theta (f_{\text{cut}} \Delta_R - \rho) 
+ \ln \left( \frac{4 \tan^2 \frac{R}{2}}{\rho} e^{-\frac{3}{4}} \right) \Theta (\rho - f_{\text{cut}} \Delta_R) \right].
\]

(3.67)

This result reveals some interesting structure, namely that for \( \rho < f_{\text{cut}} \Delta_{R_{\text{trim}}} \), behaviour of trimming is double logarithmic at LO. This region is identical in structure to the LO plain jet mass distribution with reduced radius \( R \rightarrow R_{\text{trim}} \). For intermediate values of \( \rho \), we observe a transition to single logarithmic behaviour in the region \( f_{\text{cut}} \Delta_{R_{\text{trim}}} < \rho < f_{\text{cut}} \Delta_R \), analogous to mMDT, and pruning in the region \( z_{\text{cut}}^2 \Delta_R < \rho < z_{\text{cut}} \Delta_R \). As observed for the other taggers, we obtain the double logarithmic like plain jet mass distribution for large values of \( \rho \).

We compare the analytical results, Eq. (3.67), in the region \( \rho < f_{\text{cut}} \Delta_R \) to LO EVENT2 in Fig. 3.14 for \( R_{\text{trim}} = 0.2 \) and two different values of \( f_{\text{cut}} \). Each EVENT2 result exhibits three distinct regions, as predicted by the analytic result. We observe double logarithmic behaviour in the limit \( \rho \rightarrow 0 \), which transitions into a single logarithm for intermediate values of \( \rho \). The difference between EVENT2 and the analytic result for each value of \( f_{\text{cut}} \) vanishes for small \( \rho \). This indicates that we correctly control the logarithmic structure of the trimming differential jet mass distribution at LO.

We now examine the next-to-leading order behaviour of trimming.

3.6.3 Next-to-leading order: Independent emission

In this section, we carry out a next-to-leading order calculation of the normalised differential trimming jet mass distribution in the Abelian \( C_F \) channel. We again consider the real and virtual emission of gluons \( k_1 \) and \( k_2 \) from a hard quark in the independent emission approximation (see Fig. 3.5). Like the pruning calculation, we find it convenient to separate the angular phase space into three regions. We define each region by the number of subjets obtained after reclustering the double-real emission
Figure 3.14: Comparison of Event2 at LO with the analytic expression Eq. (3.67) in the region $\rho < f_{\text{cut}}\Delta R$ for a range of $f_{\text{cut}}$ values. The red points demonstrate linear behaviour for small $\rho$, indicating the presence of double logarithms. This transitions into a flat, single logarithmic, region for intermediate jet mass. The green points demonstrate the difference between the analytic expression and the Event2 result in the region $\rho < f_{\text{cut}}\Delta R$, which tend to zero for small $\rho$. This demonstrates that we correctly control all logarithmically enhanced terms at LO.

contribution with radius $R_{\text{trim}}$:

1. One subjet: $\Theta (\Delta R_{\text{trim}} - \Delta \theta_1) \Theta (\Delta R_{\text{trim}} - \Delta \theta_2)$
2. Two subjets: $\Theta (\Delta \theta_1 - \Delta R_{\text{trim}}) \Theta (\Delta R_{\text{trim}} - \Delta \theta_2) + \Theta (\Delta R_{\text{trim}} - \Delta \theta_1) \Theta (\Delta \theta_2 - \Delta R_{\text{trim}})$
3. Three subjets: $\Theta (\Delta \theta_1 - \Delta R_{\text{trim}}) \Theta (\Delta \theta_2 - \Delta R_{\text{trim}})$

where we neglect subleading contributions associated with self-clustering of gluons. We can write the differential distribution for trimming in the $C_F^2$ channel as the integral over the sum of the contributions from each angular region:

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} \left(\text{Trimming}, C_F^2\right) = 4 \left(\frac{\alpha_s C_F}{\pi}\right)^2 \int \frac{d \cos \theta_1}{1 - \cos^2 \theta_1} \frac{d \cos \theta_2}{1 - \cos^2 \theta_2} dx_1 p_{gq} (x_1) dx_2 p_{gq} (x_2) \times \left[\Theta^{(1)} + \Theta^{(2)} + \Theta^{(3)}\right],$$

where we denote the contribution associated with each angular region as $\Theta^{(s)}$, where $s$ is the number of subjets. Writing a shorthand for the energy constraints required for an emission $i$ to pass the asymmetry check as:

$$\mathcal{P}_i = \Theta (1 - f_{\text{cut}} - x_i) \Theta (x_i - f_{\text{cut}}),$$
we combine the double-real and one-real one-virtual contributions in each region to get the contributions:

\[
\Theta^{(3)} = \left[ P_1 P_2 \delta (\rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2) + (1 - P_1) P_2 \delta (\rho - x_2 \Delta \theta_2) + (1 - P_2) P_1 \delta (\rho - x_1 \Delta \theta_1) - P_1 \delta (\rho - x_1 \Delta \theta_1) - P_2 \delta (\rho - x_2 \Delta \theta_2) \right] \Theta (\Delta \theta_1 - \Delta_{R_{\text{trim}}}) \Theta (\Delta \theta_2 - \Delta_{R_{\text{trim}}}) , \\
\Theta^{(2)} = \left\{ \left[ P_1 \delta (\rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2) + (1 - P_1) \delta (\rho - x_2 \Delta \theta_2) - P_1 \delta (\rho - x_1 \Delta \theta_1) - \delta (\rho - x_2 \Delta \theta_2) \right] \Theta (\Delta \theta_1 - \Delta_{R_{\text{trim}}}) \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_2) \right\} + \{1 \leftrightarrow 2\}, \\
\Theta^{(1)} = \left[ \delta (\rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2) - \delta (\rho - x_1 \Delta \theta_1) - \delta (\rho - x_2 \Delta \theta_2) \right] \times \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_1) \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_2) ,
\]

(3.70)

where the first and second lines in \( \Theta^{(3)} \) corresponds to the double-real configuration and the last line contains the one-real one-virtual contributions, which have a minus sign.

The quantity \( (1 - P_i) \) is the condition for emission \( i \) to fail the asymmetry cut, hence the terms in the second line correspond to the removal of subjets 1 and 2 respectively.

The first line of \( \Theta^{(2)} \) is the double-real configuration and the second line corresponds to the one-real one-virtual contributions, however we only check the asymmetry of subjet 1 because gluon \( k_2 \) forms a subjet with the quark. The third line accounts for gluon \( k_2 \) forming the second subjet, which is identical to the first two lines under the label exchange \( \{1 \leftrightarrow 2\} \). There are no asymmetry conditions applied to the single subjet region \( \Theta^{(1)} \). The constraints defined in Eq. (3.70) simplify to

\[
\Theta^{(3)} = \delta_\rho \left[ P_1 P_2 \Theta (\Delta \theta_1 - \Delta_{R_{\text{trim}}}) \Theta (\Delta \theta_2 - \Delta_{R_{\text{trim}}}) \right] , \\
\Theta^{(2)} = \delta_\rho \left[ P_1 \Theta (\Delta \theta_1 - \Delta_{R_{\text{trim}}}) \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_2) \right] + \{1 \leftrightarrow 2\}, \\
\Theta^{(1)} = \delta_\rho \left[ \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_1) \Theta (\Delta_{R_{\text{trim}}} - \Delta \theta_2) \right] ,
\]

(3.71)

where the sum of the relevant double-real and one-real one-virtual contributions to \( \rho \) is written as:

\[
\delta_\rho = \delta (\rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2) - \delta (\rho - x_1 \Delta \theta_1) - \delta (\rho - x_2 \Delta \theta_2) .
\]

(3.72)

The result in Eq. (3.71) implies that, for each region, complete real and virtual cancellation for each emission \( i \) occurs in the region \( f_{\text{cut}} > x_i \) and \( 1 - f_{\text{cut}} < x_i \) given that \( \Delta \theta_i > \Delta_{R_{\text{trim}}} \). The remaining phase space contributes logarithmic terms to the trimming jet mass distribution.
We can simplify the integral by factorising the sum of the constraints into a product of two terms that have the same form as the leading order constraints in Eq. (3.66):

\[
\sum_{i=1}^{3} \Theta^{(i)} = \delta_{\rho} \prod_{i=1}^{2} \left[ \Theta (1 - f_{\text{cut}} - x_{i}) \Theta (x_{i} - f_{\text{cut}}) \Theta (\Delta_{\theta_{i}} - \Delta_{R_{\text{trim}}}) + \Theta (\Delta_{R_{\text{trim}}} - \Delta_{\theta_{i}}) \right].
\] (3.73)

Furthermore, we are interested in the leading, \( O(\alpha_s^2 L^4) \), and next-to leading terms, \( O(\alpha_s^2 L^3) \), in the integrated jet mass distribution. Therefore, we replace the delta functions defined in Eq. (3.72) using \( \delta(f(\rho)) = \frac{\partial}{\partial \rho} \Theta(f(\rho)), \) where \( f \) is a function of \( \rho \), and take the “leading log” approximation [166]:

\[
\Theta (\rho - x_{1}\Delta_{\theta_{1}} - x_{2}\Delta_{\theta_{2}}) \rightarrow \Theta (\rho - x_{1}\Delta_{\theta_{1}}) \Theta (\rho - x_{2}\Delta_{\theta_{2}}), \tag{3.74}
\]

substituting this into Eq. (3.72) gives

\[
\delta_{\rho} \approx \frac{\partial}{\partial \rho} \left[ \Theta (x_{1}\Delta_{\theta_{1}} - \rho) \Theta (x_{2}\Delta_{\theta_{2}} - \rho) \right]. \tag{3.75}
\]

The kinematical constraints for each emission, given in Eq. (3.73), are independent of \( \rho \), hence we can completely factorise the integral in Eq. (3.68) up to single logarithmic level into the product of two integrals. Explicitly, this corresponds to

\[
\frac{1}{\sigma} \frac{d\sigma}{d\rho} (\text{Trimming}, C_{F}^{2}) = 4 \left( \frac{\alpha_{s} C_{F}}{\pi} \right)^{2} \frac{1}{2} \left[ \int \frac{d \cos \theta_{i}}{1 - \cos^{2} \theta_{i}} \int dx_{i} p_{gq} (x_{i}) \Theta (x_{i} \Delta_{\theta_{i}} - \rho) \Theta (\Delta_{R} - \Delta_{\theta_{i}}) \right.
\]

\[
\times \left[ \Theta (\Delta_{\theta_{i}} - \Delta_{R_{\text{trim}}}) \Theta (x_{i} - f_{\text{cut}}) \Theta (1 - f_{\text{cut}} - x_{i}) \right], \tag{3.76}
\]

which is consistent with an all-orders exponentiation of the LO result given in Eq. (3.66). Evaluating Eq. (3.76) in the region \( \rho < f_{\text{cut}} \Delta_{R_{\text{trim}}} \), we get the result:

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} (\text{Trimming}, C_{F}^{2}) = - \left( \frac{\alpha_{s} C_{F}}{\pi} \right)^{2} \left( \frac{1}{2} \ln^{4} \frac{1}{\rho} + \left[ 3 \ln \left( 4 \tan^{2} \frac{R_{\text{trim}}}{2} \right) - \frac{9}{8} \right] \ln^{2} \frac{1}{\rho} \right), \tag{3.77}
\]

where we have neglected terms \( O(\alpha_s^2 L) \) in the differential distribution. This result is double logarithmic in the small \( \rho \) limit and has an identical logarithmic structure to the differential plain mass distribution for a jet with radius \( R_{\text{trim}} \).
Figure 3.15: Comparison of EVENT2 in the $C_F^2$ channel at NLO with the analytic expression Eq. (3.77) in the region $\rho < f_{\text{cut}}\Delta R_{\text{trim}}$ for a range of $f_{\text{cut}}$ values. The red points indicate the presence of double logarithms. The green points demonstrate the difference between the analytic expression and EVENT2 result in the region $\rho < f_{\text{cut}}\Delta R_{\text{trim}}$. The straight line at small $\rho$, corresponds to leftover single logarithm $O(\alpha_s^2 L)$ in the differential distribution.

We check this analytical result, Eq. (3.77), against EVENT2 in the $C_F^2$ channel in Fig. 3.15. The difference between the analytic and EVENT2 result indicates that we correctly reproduce the leading and next-to-leading logarithms in the differential jet mass distribution for trimming in the $C_F^2$ channel. The remaining linear behaviour is consistent with leftover $O(\alpha_s^2 L)$ terms in the differential distribution. Note we have subtracted the analytical result in the region $\rho < f_{\text{cut}}\Delta R_{\text{trim}}$ because we are only interested in capturing the large logarithmic contributions as $\rho$ vanishes.

3.6.4 Next-to-leading order: Non-Abelian terms

Up to, but not including, single logarithmic accuracy, the only relevant contributions in the $C_F C_A$ or $C_F n_f$ channels come from the running of the strong coupling. This produces an additional contribution that is identical to the plain jet mass running coupling correction in the small $\rho$ limit:

$$
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(\text{Trimming,NLO,rc})} = \left( \frac{\alpha_s}{\pi} \right)^2 C_F b_0 \frac{3}{2} \ln^2 \frac{1}{\rho}.
$$

We check this result against EVENT2 in the $C_F C_A$ and $C_F n_f$ channels in Fig. 3.16. We find that we correctly control the $O(\alpha_s^2 L^2)$ terms contributing to the differential jet mass distribution at NLO. As usual, we are left with a straight line corresponding to leftover single logarithms.

Non-global logarithms are not explicitly calculated here, but it is fairly obvious they contribute to the trimmed jet mass distribution. One requires that a soft gluon is emitted at an angle $\theta > R_{\text{trim}}$, which radiates another soft gluon into the core of the
Figure 3.16: Comparison of EVENT2 in the $C_F C_A$ channel (left) and $C_F n_f$ (right) with the analytic expression Eq. (3.78) in the region $\rho < f_{cut} \Delta_{R_{trim}}$ for two $f_{cut}$ values. The green points show the difference between the analytic expression and EVENT2, which is a straight line at small $\rho$, corresponding to a leftover single logarithm $O(\alpha_s^2 L)$ in the differential distribution.

The primary gluon is removed if its energy fraction $x < f_{cut}$, but the secondary emission contributes to the jet mass, resulting in non-global logarithmic contributions.

3.6.5 Phenomenology

In this section we computed the logarithmic structure of the differential jet mass distribution after application of the trimming algorithm at leading and next-to leading order. We found in the small $\rho$ limit that this distribution contains double logarithms at LO and NLO. If we redefine $L \equiv \ln \left( \frac{R_{2,trim}^2}{\rho} \right)$, we can write the coefficients for the series defined in Eq. (3.1), in the region $f_{cut} \Delta_R > \rho$ as:

$$a_{12}^{Trimming} = C_F \Theta \left( f_{cut} \Delta_{R_{trim}} - \rho \right)$$
$$a_{11}^{Trimming} = -\frac{3C_F}{4} \Theta \left( f_{cut} \Delta_{R_{trim}} - \rho \right) + C_F \ln \frac{e^{-3/4}}{f_{cut}} \Theta \left( \rho - f_{cut} \Delta_{R_{trim}} \right)$$
$$a_{24}^{Trimming} = -\frac{C_F}{2} \Theta \left( f_{cut} \Delta_{R_{trim}} - \rho \right)$$
Figure 3.17: Comparison of the differential distribution in $\rho$ for trimming tagged jets using Monte Carlo (left) and full resummed analytical results (right). Details of MC generation are given in Fig. 3.1. The arrows denote the expected position of the transition points, as derived by the analytical calculations in this section, i.e. $\rho \simeq f_{\text{cut}} \Delta R_{\text{trim}}$ (black) and $\rho \simeq f_{\text{cut}} \Delta R$ (grey). This figure and caption has been adapted directly from Ref. [116].

$$a_{23}^{\text{Trimming}} = \frac{3}{8} (3C_F + 4b_0) \Theta (f_{\text{cut}} \Delta R_{\text{trim}} - \rho), \quad (3.79)$$

which are identical to the plain jet mass results given in Eq. (3.2) in the limit $\rho \to 0$. We re-emphasise that in the region $f_{\text{cut}} \Delta R_{\text{trim}} < \rho < f_{\text{cut}} \Delta R$, the coefficients $a_{12}^{\text{Trimming}} = a_{24}^{\text{Trimming}} = a_{23}^{\text{Trimming}} = 0$ and we are left with a purely single logarithmic result up to NLO.

The form of the NLO contribution, Eq. (3.68), suggests that the all-orders expression for trimming in the small $\rho$ limit is a simple exponentiation. In the fixed coupling and small $f_{\text{cut}}$ limit, the exponent is given by the integrated result for single gluon emission defined in Eq. (3.67). It follows that exponentiation of the double logarithmic terms will lead to a trimming jet mass distribution that contains a Sudakov peak, like plain jet mass (see Fig. 3.17). The Sudakov peak in the background distribution will be situated at jet masses below the transition point $\rho \simeq f_{\text{cut}} \Delta R_{\text{trim}}$. For boosted jets with $p_T \simeq 3$ TeV and typical trimming parameters $f_{\text{cut}} \simeq 0.03$, $R_{\text{trim}} \simeq 0.2$, this corresponds to a transition point close to electroweak scale jet masses. As with pruning, this may impact signal to background rejection rates and may be undesirable for data driven background estimates.

In order to see this explicitly, in Fig. 3.17 we provide the comparison of the full analytic resummed result and the corresponding MC simulation from Ref. [116]. These
show the differential distribution in $\rho$ for highly boosted jets tagged with trimming at parton level. From this figure, one can immediately see that all the features of the fixed-order calculation are manifest in the differential jet mass distribution. Specifically, in the LO Eq. (3.67) and NLO calculations Eq. (3.77), we calculated a double logarithmic distribution in the region $\rho < f_{\text{cut}} \Delta_{R_{\text{trim}}}$, which implies the existence of a Sudakov peak for small $\rho$. This transition point is labelled with a black arrow and can be seen in MC as the onset of the Sudakov peak with decreasing $\rho$. Like pruning, trimming was calculated to have a single logarithmic distribution in an intermediate region $f_{\text{cut}} \Delta_{R_{\text{trim}}} < \rho < f_{\text{cut}} \Delta_{R}$ that transitions to the plain result for $\rho > f_{\text{cut}} \Delta_{R}$, which can be seen in the MC results.

### 3.7 Conclusions

In this chapter we studied the jet masses of quark jets after application of several substructure algorithms, specifically mass drop, pruning and trimming boosted-object techniques. The novel feature of this study is that we have employed analytical techniques to describe the perturbative structure and behaviour of each tagger. This is in contrast to studies which use Monte Carlo event generators to simulate the features of each tagger, which is standard practice in most substructure analyses. This enabled us to calculate the leading and next-to-leading fixed-order perturbative structure of each tagger in the eikonal approximation, extended to treat hard-collinear emissions. Explicitly, we considered the jet mass of boosted quark jet pairs as an observable in the context of $e^+e^-$ collisions. The results obtained in this chapter can be applied in the context of hadron-hadron colliders provided one considers jets in the small $R$ approximation. In order to go beyond this approximation, one must compute contributions from initial state radiation. However, contributions to the boosted jet mass from emissions that are associated with the initial state beam will be subleading due to a lack of collinear enhancement, i.e. at most single logarithmic.

All of the substructure techniques studied in this chapter provide cuts on radiation within a tagged jet, which are designed to reduce uncorrelated soft contamination and enable use of the jet mass as an effective discriminant between signal and QCD jets. It is natural to expect that these soft radiation cuts affect the logarithmic structure of the observable, and one can ask whether they remove some of the large logarithms in $M_j/p_T$ that contribute at each order to the jet mass observable. By using analytical methods, one can unambiguously determine the perturbative structure of the observable expanded in the strong coupling. Crucially, we can write this logarithmically enhanced, analytical structure as a function of the algorithmic parameters (e.g. $y_{\text{cut}}$, $\mu$) and jet variables (e.g. $R$, $p_T$, $M_j$). Using this information, we can examine which theoretical methods can be used to most accurately compute such substructure observables. Depending on the specific parameter choices, it may be possible to employ
pure fixed-order calculations to accurately compute these observables, in regions where logarithms in $M_j/p_{T_j}$ are not too large. Otherwise, it may be necessary to perform a resummation on any large leftover logarithms to accurately describe these substructure observables.

In this chapter, we started by examining the mass drop tagger algorithm and found a single logarithmic expression for the jet mass distribution at LO. However, this description changed at NLO due to a “wrong branch” effect (following a massive branch consisting of soft emissions), giving rise to terms $\mathcal{O}(\alpha_s^2 L^3)$ in the integrated distribution. By locating the origin of this contribution, we proposed a modification of MDT, named mMDT, which replaces the recursion choice after branch rejection from the most massive branch to the hardest one. We showed that this modification eliminated the leading wrong-branch terms, such that the mMDT jet mass consists only of single logarithms at NLO. At NLO we calculated that all soft logarithms are removed by mMDT and demonstrated explicitly that non-global logarithms, which are purely soft in origin, are not present in the mMDT jet mass distribution, making it a unique single jet observable. We demonstrated that these features, as obtained using fixed-order techniques, are manifest in the MC and analytic resummed results. Specifically, approaching small $\rho$, we observe a transition from the plain jet mass result into a single logarithmic region, which results in a linear distribution in $\ln \rho$ for the differential jet mass.

Next we examined pruning, which at leading order was calculated to be single logarithmic and equivalent to the LO (m)MDT expression under a change of variables. However, at NLO, we encountered a double logarithmic result, $\mathcal{O}(\alpha_s^2 L^4)$ in the integrated jet mass distribution. This is due to the dynamic nature of the pruning radius, whereby a soft emission can dominate the fat jet mass and set $R_{\text{prune}}$. Removal of this emission can lead to a pruned jet that is dominated by soft and collinear emissions inside $R_{\text{prune}}$, all of which are never checked for the asymmetry condition. Additionally, we showed the presence of non-global logarithms in the pruning jet mass distribution, in contrast to mMDT. NLO calculations at non-vanishing values of $\rho$ showed that pruning mass distribution has two transition points and in the region $z_{\text{cut}}^2 \Delta_R < \rho < z_{\text{cut}} \Delta_R$ we obtained a single logarithmic distribution, which is equivalent to mMDT under a change of variables. We also proposed and considered the logarithmic structure of Y-pruning at NLO and demonstrated that this modification removes the double logarithmic contributions by rejecting jets in which the pruning radius was set by a soft emission. The resulting distribution at NLO contained leading next-to-double logarithmic contributions to the integrated jet mass of the form $\mathcal{O}(\alpha_s^2 L^3)$. We then demonstrated that the distinct regions and transition points obtained using fixed-order techniques can be observed in the MC and resummed analytical results. Using this information, we argued that the transition points in the highly boosted pruned jet mass distribution
Table 3.1: A table that summarises the main features of each tagger. The first two columns show the leading logarithmic terms in the limit $\rho \rightarrow 0$ in the integrated jet mass distribution at leading and next-to-leading order. We have also tabulated the transition points in $\rho$ and indicated whether non-global logarithms in $\rho$ are present.

<table>
<thead>
<tr>
<th>Tagger</th>
<th>LO</th>
<th>NLO</th>
<th>Transition point(s)</th>
<th>Non-global logarithms?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>$\alpha_s L^2$</td>
<td>$\alpha_s^2 L^4$</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td>MDT</td>
<td>$\alpha_s L$</td>
<td>$\alpha_s^2 L^3$</td>
<td>$\frac{f_{\text{cut}} \Delta_R}{1+y_{\text{cut}}}$</td>
<td>Yes</td>
</tr>
<tr>
<td>mMDT</td>
<td>$\alpha_s L$</td>
<td>$\alpha_s^2 L^2$</td>
<td>$\frac{f_{\text{cut}} \Delta_R}{1+y_{\text{cut}}}$</td>
<td>No</td>
</tr>
<tr>
<td>Pruning</td>
<td>$\alpha_s L$</td>
<td>$\alpha_s^2 L^4$</td>
<td>$z_{\text{cut}}^2 \Delta_R, z_{\text{cut}} \Delta_R$</td>
<td>Yes</td>
</tr>
<tr>
<td>Y-pruning</td>
<td>$\alpha_s L$</td>
<td>$\alpha_s^2 L^3$</td>
<td>$z_{\text{cut}} \Delta_R$</td>
<td>Yes</td>
</tr>
<tr>
<td>Trimming</td>
<td>$\alpha_s L^2$</td>
<td>$\alpha_s^2 L^4$</td>
<td>$f_{\text{cut}} \Delta_{R_{\text{trim}}}, f_{\text{cut}} \Delta_R$</td>
<td>Yes</td>
</tr>
<tr>
<td>Y-splitter</td>
<td>$\alpha_s L$</td>
<td>$\alpha_s^2 L^3$</td>
<td>$\frac{f_{\text{cut}} \Delta_{R_{\text{trim}}}}{1+y_{\text{cut}}}$</td>
<td>No</td>
</tr>
</tbody>
</table>

can arise around the electroweak mass scale. Therefore, this combination of non-trivial structure in the background jet mass distribution, combined with double logarithmic behaviour, may potentially make pruning undesirable for use in conjunction with data driven background estimates in some regions of phase space.

Finally, we calculated the jet mass distribution after application of trimming at LO and NLO. We found double logarithmic contributions to the integrated jet mass distribution for $\rho < f_{\text{cut}} \Delta_{R_{\text{trim}}}$ of the form $O(\alpha_s L^2)$ and $O(\alpha_s^2 L^4)$. Analogous to the pruning result, we calculated two transition points in the trimming jet mass distribution that, for boosted jets, can also arise close to the electroweak scale. We also calculated that single logarithmic behaviour exists in the region $f_{\text{cut}} \Delta_{R_{\text{trim}}} < \rho < f_{\text{cut}} \Delta_R$, as was shown for mMDT and pruning (for intermediate values of $\rho$). Additionally, the form of the leading NLO contribution in the small $\rho$ limit suggested exponentiation of the leading-order result. We finally noted that the logarithmic structure for trimming in the small $\rho$ limit is similar to that obtained for the plain jet mass, under the replacement $R \rightarrow R_{\text{trim}}$. These results were compared against MC and analytical resummed results, and we discussed how the transition points and logarithmic behaviour in each region emerge in the differential jet mass distribution. A summary of all the results are given in Table 3.1; note that we have included the jet substructure algorithm Y-splitter, the logarithmic structure of which we will calculate explicitly in Chapter 5.

The analytic results derived in this chapter can be used to understand the precise dependence of jet observables, like the mass, on parameters used in jet clustering and substructure algorithms and the interplay between them. One can use this information to optimise these parameters for QCD jet background rejection in the context of phe-
nomenological studies that use substructure techniques in a discovery context. We also anticipate that the location of transition points and Sudakov peaks in the background jet mass distribution will be useful for assessing data driven background estimates. For a complete picture, one would need an analytical study of these techniques applied to signal jets (see Chapter 4) in order to assess performance in terms of signal to background ratios. Our results concerning the leading logarithmic structure of each tagger, will also help in determining the suitability of fixed-order or resummed computations of jet observables in different regions of parameter space.
CHAPTER
FOUR

JET SUBSTRUCTURE METHODS FOR SIGNAL JETS

4.1 Introduction

One of the main aims of Chapter 3 was to better understand how aspects of jet substructure algorithm definition and design may interplay with QCD dynamics to dictate the performance of jet substructure algorithms (that we shall collectively refer to as “taggers”), as reflected by their action on background jets. The improved analytical understanding that was achieved in the previous chapter led to a better appreciation of the role of tagger parameters (including the discovery of the apparent redundancy of the mass-drop parameter $\mu$ in the mass-drop tagger). These analytical studies also paved the way for improvement of theoretical properties of taggers. Examples of improvements that were suggested or made in the last chapter and Ref. [116] included the design of taggers with a perturbative expansion more amenable to resummation as for the modified mass-drop tagger (mMDT) as well as removing undesirable tagger features as for the case of pruning via the Y-pruning modification.

In the previous chapter, our focus was exclusively on taggers applied to light QCD background jets. We now conduct an analytical investigation of the impact of radiative effects for signal jets after application of a range of substructure methods. The reader should be aware that limited analytical calculations have already been performed to study the action of filtering for $H \rightarrow b\bar{b}$ [155] and for N-subjettiness [158].

We first observe that it is common to study high $p_T$ signal jets in some relatively narrow mass window of width $\sim \delta M$ around the mass, $M$ of some boosted decaying heavy resonance of interest, this mass cut being a step in tagging signal jets. One then has a situation where there are various disparate scales involved in the problem such as the (potentially) TeV scale transverse momenta of the fat jets, the mass $M$
of the resonance (which for our studies we can consider to be around the electroweak scale) and the width of the window $\delta M$, which for most purposes we can consider as a parameter $\sim 10$ GeV. These scales are in addition, of course, to the various parameters corresponding to angular distances and energy cuts introduced by tagging and jet finding.

In such multi-scale problems, radiative corrections have the potential to produce large logarithms involving ratios of disparate scales. In the example of filtering studied in Ref. [155], large logarithms in the Higgs mass to window ratio $M_H/\delta M$ arose from considering soft emissions, which were accompanied by collinear logarithmic enhancements in $R_{\bar{b}b}/R_{\text{filt}}$, i.e. the ratio of the $\bar{b}b$ opening angle to the filtering radius. On the other hand, Ref. [116] observed via MC studies of the signal that for the taggers studied there (mMDT, pruning, trimming, Y-pruning) the tagger performance was primarily driven by the action of taggers on QCD background, with signals not appearing to display very sizeable radiative corrections for the chosen default parameters.

In order to better understand these apparently contrasting observations, it is desirable to acquire a higher level of analytical insight into the action of taggers on signal jets. When comparing the performance of taggers one may also meet a situation where two taggers shall act similarly on background jets and hence their action on signal becomes of critical significance. An explicit example of this situation is provided later in Section 4.6. It is also of importance to understand and assess the impact of QCD radiative corrections and non-perturbative effects on tagger efficiency for signals, to ascertain what theoretical tools (fixed-order calculations, MC methods, resummed calculations or combinations thereof), should ideally be employed to get the most reliable picture for the signal efficiency for a given tagger.

With all the above aims in mind, here we embark on a more detailed study dedicated to signal jets. We shall focus our attention on the case of a jet arising from a boosted Higgs boson. Specifically, for Higgs production in association with a vector boson $pp \to W/Z, H$, with $H \to \bar{b}b$, where we will work in a narrow width approximation. This means that we explicitly neglect the decay width of the mass resonance peak of the Higgs ($\Gamma \sim 1$ GeV [21]) with respect to its central mass value, $M_H$. For the approximate analytical calculations we perform in this chapter, we can safely ignore the decay width for a sufficiently large tagging mass window $\Gamma \ll \delta M$. For our analytical approximations we shall also typically consider highly boosted configurations, i.e. those where the Higgs has transverse momentum $p_T \gg M_H$ and shall further take a fat jet with radius $R \gg M_H/p_T$. We shall work in a small-angle approximation throughout\footnote{This approximation, for practical purposes, is valid up to reasonably large values of $R \sim \mathcal{O}(1)$ due to significant constant suppression terms in the coefficients of the soft and wide-angle contributions proportional to $R$ [187].} though we will often consider $R \sim 1$. 

140
We stress here that we do not intend to provide precise high-order calculations for radiative corrections to any given process but seek mainly to understand and compare the behaviour of taggers via a combination of approximate analytics and MC cross-checks. For an example of exact fixed-order calculations involving jet substructure and signal processes we refer the reader to Ref. [193].

We start in Section 4.2 by analysing the case of a plain jet mass cut, focussing on a mass window around the signal mass and deeming a jet to be tagged as signal if the jet mass falls within this window. We consider the impact of initial state radiation (ISR), final state radiative corrections (FSR) both analytically and in MC studies on the signal efficiency. We also study the impact of non-perturbative effects (NP) using MC. The results so obtained can then provide a point of reference and comparison to judge the improvements that are offered by use of substructure taggers, which impose further kinematical requirements in addition to a simple cut on mass. In Section 4.3 we study trimming at lowest order and also with ISR and FSR corrections. We investigate the logarithmic structure that emerges in $M$, $\delta M$ and $p_T$ as well as in tagger parameters and compare to MC results where appropriate. We also study the impact of non-perturbative corrections, though purely with MC results. In Section 4.4 we analyse pruning and the modified mass-drop tagger (mMDT) along similar lines. FSR is analysed further for these taggers in Appendices A.1 and A.2 where we also compare parton shower results to those from full leading-order calculations, i.e. those that go beyond the soft/collinear approximation. In Section 4.5 we study Y-pruning and the Y-splitter tagger [96]. We observe that while the action of Y-splitter on QCD background is similar to Y-pruning, the signal jet with Y-splitter is subject to severe loss of mass resolution due to ISR and underlying event (UE) effects. In Section 4.6 we show with MC studies that combining Y-splitter with trimming dramatically improves the signal behaviour while leaving the background largely unmodified. As a consequence, we show that Y-splitter with trimming outperforms the other taggers we study here, especially at high $p_T$. This example further illustrates how even a relatively basic analytical understanding of all aspects of taggers (for both signal and background) can be exploited to achieve important performance gains. Finally, in Section 4.7 we carry out analytical studies of optimal values for tagger parameters, obtained by maximising signal significance, and compare to MC results. We conclude with a summary and mention prospects for future work.

4.2 Results for plain jet mass

Here we shall consider the plain jet-mass distribution for fat signal jets without the application of substructure methods, other than the imposition of a mass window $\delta M$, as previously stated. As also mentioned before, we shall consider the case of Higgs boson production in association with an electroweak vector boson $pp \rightarrow W/Z, H$
with Higgs decay to a $b\bar{b}$ quark pair and shall work in a narrow width approximation throughout. For the purposes of examining the jet substructure, we shall not need to write down matrix elements for the production of the high $p_T$ Higgs boson. Instead, we shall be concerned purely with the details of the Higgs decay and the resulting fat jet, as well as the impact of ISR, FSR and non-perturbative effects.

Let us take a boosted Higgs boson produced with transverse momentum $p_T \gg M_H$ and, purely for convenience, set it to be at zero rapidity with respect to the beam direction, such that the corresponding energy is $\sqrt{p_T^2 + M_H^2}$. We further consider that the Higgs decays into a $b\bar{b}$ pair, such that, in terms of four-momenta, one has $p_H = p_b + p_{\bar{b}}$. Thus the invariant mass of the Higgs can be expressed as

$$p_H^2 = M_H^2 = 2p_b \cdot p_{\bar{b}} = 2z(1-z)(p_T^2 + M_H^2)(1-\cos\theta_{b\bar{b}}),$$  \hspace{1cm} (4.1)

where $z$ and $1-z$ are the energy fractions of the decay products, $\theta_{b\bar{b}}$ is the $b\bar{b}$ opening angle and we have neglected the $b$ quark masses. Furthermore, we shall consider the highly boosted regime by taking $\Delta = \frac{M_H^2}{p_T^2} \ll R^2$ and systematically neglect power corrections in $\Delta$. Then from Eq. (4.1), taking a small-angle approximation, we obtain the standard result:

$$\theta_{b\bar{b}}^2 \approx \frac{\Delta}{z(1-z)}. $$  \hspace{1cm} (4.2)

We require that the Higgs decay products are contained in a single fat jet, $\theta_{b\bar{b}}^2 < R^2$, which translates into a constraint on $z$:

$$z(1-z) > \frac{\Delta}{R^2}. $$  \hspace{1cm} (4.3)

Let us start by providing the results for the signal efficiency $\varepsilon_S^{(0)}$, to lowest order, i.e. taking just the $H \rightarrow b\bar{b}$ decay without any radiative corrections. This can be considered as the fraction of decays that are reconstructed inside a fat jet of radius $R$. Here one has to consider the relevant Feynman amplitude for $H \rightarrow b\bar{b}$ and the full decay phase-space with an integral over the final state parton momenta. However, for the fraction of decays inside the boosted fat jet, $\varepsilon_S^{(0)}$, we simply obtain

$$\varepsilon_S^{(0)} = \int_0^1 dz \Theta \left( R^2 - \frac{\Delta}{z(1-z)} \right) = \sqrt{1 - \frac{4\Delta}{R^2} \Theta (R^2 - 4\Delta)} \approx 1 - \frac{2\Delta}{R^2} + O \left( \frac{\Delta^2}{R^4} \right),$$  \hspace{1cm} (4.4)

which is trivially in good agreement with corresponding MC event generator results when all ISR, FSR and non-perturbative effects are turned off. The result in Eq. (4.4) simply suggests, as one can easily anticipate, that with increasing transverse boosts, i.e. smaller $\Delta$, the efficiency of reconstruction inside a single jet increases. At this lowest-
order there is no role for the mass-window $\delta M$ because the jet mass $M_j$ coincides with
the Higgs mass $M_H$.

4.2.1 Initial state radiation

Let us now account for the impact of initial state radiation on the jet mass distribution. We can anticipate that the impact of soft radiation may be significant here because we require the invariant mass of the fat jet to be within $\delta M$ of the Higgs mass, with $\delta M \ll p_T$. This requirement imposes a constraint on the phase space of real emissions arising from ISR that enter the jet because these contribute directly to the deviation of the jet mass from $M_H$. Hence, one can expect that large logarithmic corrections arise as a consequence. In order to understand the structure of these corrections, we consider the process $pp \rightarrow ZH$ with the additional production of soft gluons radiated by the incoming hard partons (here we consider a $q\bar{q}$ pair). Let us start by taking a single ISR gluon that is soft, i.e. has energy $\omega \ll p_T$. In the soft limit we can work with the eikonal approximation in which production of the ISR factorises from the Born-level hard process $pp \rightarrow ZH$. To compute the signal efficiency we shall require the jet invariant mass to be within a relatively narrow mass-window $\delta M$ of the Higgs mass:

$$M_H - \delta M < M_j < M_H + \delta M.$$  \hfill (4.5)

At lowest order (Born level), this inequality is always true because $M_j = M_H$, however, with ISR it amounts to a constraint on the ISR gluon energy. Neglecting corrections of relative order $\delta M/M_H$ we can write the mass window constraint as:

$$M_j^2 - M_H^2 = 2p_H \cdot k < 2M_H\delta M,$$  \hfill (4.6)

where $p_H$ is the four-momentum of the Higgs (or equivalently the sum of the four-momenta of its decay products) and $k$ that of the ISR particle. Defining $\theta$ as the angle between the soft emission and the Higgs direction we can write:

$$2p_H \cdot k = 2\omega \left( \sqrt{p_T^2 + M_H^2} - p_T \cos \theta \right).$$  \hfill (4.7)

Taking $\Delta = M_H^2/p_T^2 \ll 1$ and using the small $\theta$ approximation, we can expand in small quantities and write Eq. (4.6) as the following constraint on gluon energy:

$$\omega < \frac{2M_H\delta M}{p_T (\theta^2 + \Delta)}.$$  \hfill (4.8)

One can express this equation in terms of standard hadron collider variables $k_t$, $\eta$ and $\phi$ defined with respect to the beam direction by noting simply that $\omega = k_t \cosh \eta$ and $\theta^2 \approx \eta^2 + \phi^2$.  

143
CHAPTER 4. JET SUBSTRUCTURE METHODS FOR SIGNAL JETS

We wish to examine only the leading logarithmic structure that arises from soft ISR emissions, starting with a single emission, i.e. to leading order in the strong coupling. Because we are considering an emission that enters the high $p_T$ fat jet we are concerned with large-angle radiation from the incoming hard partons. This in turn implies that there are no collinear enhancements associated with such radiation and the resulting leading logarithmic structure ought to be single-logarithmic, arising purely from the infrared singularities in the gluon emission probability.\(^2\) The gluon emission probability is given by the standard two particle antenna for soft emissions associated with the incoming quark/anti-quark:

$$W_{ij} = 2C_F \frac{\alpha_s}{\pi} \frac{(p_i \cdot p_j)}{(p_i \cdot k)(p_j \cdot k)}.$$  \hspace{1cm} (4.9)

where $C_F = 4/3$. Analogous to the case of jet mass distributions for hadron collider QCD jets (see, for example, calculations in Ref. [187]) the ISR contribution can be written by integrating Eq. (4.9) over the gluon emission phase space. The result can be expressed in terms of $k_t$, $\eta$ and $\phi$ and reads:

$$
\varepsilon_{S,\text{ISR}}^{(1)} = \int dz \Theta \left( z(1-z) - \frac{\Delta}{R^2} \right) \times 
\int \frac{dk_t}{k_t} d\eta \frac{d\phi \alpha_s(k_t^2)}{2\pi} \left( \Theta \left( \frac{2M_H\delta M}{p_T(\theta^2 + \Delta)} - k_t \cosh \eta \right) - 1 \right) \Theta_{\text{jet}}. \hspace{1cm} (4.10)
$$

This equation contains an integral over the energy fraction of the $b$ quark involved in the Higgs decay (i.e. that over $z$), with a constraint that is identical to the zeroth order requirement that the hard quarks be contained in the fat jet; this constraint is unmodified by the presence of soft ISR at leading logarithmic level, i.e. in the limit $\omega = k_t \cosh \eta \ll p_T$. The step function involving a restriction on the transverse momentum $k_t$ on the second line follows directly from Eq. (4.8) and the subsequent arguments. Virtual corrections are incorporated via unitarity through the $-1$ term also in parenthesis. Lastly, we have a factor $\Theta_{\text{jet}}$ that is the condition that the soft ISR is within the fat jet.

The clustering condition $\Theta_{\text{jet}}$ is in principle quite complicated because it involves recombination of three particles within the fat jet, namely the $b$, $\bar{b}$ and the ISR gluon. In our approximation of $\Delta \ll R^2$, i.e. in the limit of large transverse boosts, we are considering a highly collimated quark pair, relative to the radius of the fat jet. One can thus ignore the effect of the finite $b\bar{b}$ opening angle; these effects contribute only terms that are relatively suppressed by powers of $\Delta$ compared to the leading term we

\(^2\)This expectation is challenged by the discovery of superleading logs [68]. Because these appear at order $\alpha_s^4$ and are suppressed as $1/N_c^2$ we do not expect them to make a significant impact on the essential arguments we make here.

144
compute. Then one only has to consider that the soft ISR gluon is in the interior of
the fat jet, which amounts to the condition \( \Theta_{\text{jet}} = \Theta(R^2 - (\eta^2 + \phi^2)) \), because we have
taken the Higgs rapidity to be zero.

Within the context of the current purely order \( \alpha_s \) calculation, we shall also consider
the coupling as fixed at scale \( p_T \) and ignore its running. Running coupling effects are
of course important to include for leading logarithmic resummation and we shall do
so for our final answers. Defining \( \varepsilon_{S,\text{ISR}} = \varepsilon_S^{(0)} + \varepsilon_{S,\text{ISR}}^{(1)} \) and carrying out the relevant
integrations with fixed coupling in Eq. (4.10), we arrive at the leading logarithmic
result:
\[
\frac{\varepsilon_{S,\text{ISR}}^{(0)}}{\varepsilon_S^{(0)}} \approx 1 - \frac{C_F \alpha_s}{\pi} R^2 \ln \left( \frac{p_T^2 R^2}{2M_H \delta M} \right). \tag{4.11}
\]
In order to obtain this result, starting from Eq. (4.10), one first integrates over \( k_t \),
discarding the accompanying \( \cosh \eta \) factor as this will only generate subleading terms.
This integral over \( k_t \) produces the large logarithm we seek. One can then express the
\( \eta, \phi \) integral as one with respect to \( \theta^2 = \eta^2 + \phi^2 \) and integrate over \( \theta^2 \) with the condition
\( \theta^2 < R^2 \). Neglecting subleading terms, including those that vanish with \( \Delta \), one then
obtains the result reported in Eq. (4.11) for the quantity \( \varepsilon_{S,\text{ISR}}/\varepsilon_S^{(0)} \). We have chosen
to retain the formally subleading logarithmic dependence on \( R \), which can become
important at smaller values of \( R \).

Given the large single logarithms that emerge from the approximate fixed-order
calculation in Eq. (4.11), it is natural to wish to attempt to resum at least the leading
logarithms to all perturbative orders. This is far from a straightforward exercise.
One of the main obstacles to performing a soft, single logarithmic resummation, in
the present context, is the presence of non-global logarithms, associated clustering
logarithms \([170,190,194,195]\) as well as superleading logarithms \([68]\), as referred to
previously. Such logarithmic contributions pose a serious challenge to the current state
of the art and are beyond the scope of this work.

In the absence of a complete resummed calculation, one can still obtain a working
estimate that can be compared to MC results, simply by exponentiating the order \( \alpha_s \)
result obtained from Eq. (4.10) and including running coupling effects. The exponen-
tiated result including the running of the QCD coupling is given by
\[
\frac{\varepsilon_{S,\text{ISR}}^P}{\varepsilon_S^{(0)}} \approx \exp \left( -2C_F R^2 t \right) \tag{4.12}
\]
where we defined the single-log evolution variable

$$ t = \frac{1}{2\pi} \int_{\frac{2M_H^2 \delta M}{p_T^2 R^2}}^{p_T^2} \frac{dk_t}{k_t} \alpha_s(k_t^2) $$

$$ = \frac{1}{4\pi \beta_0} \ln \frac{1}{1 - 2\lambda}, \quad \lambda = \beta_0 \alpha_s(p_T^2 R^2) \ln \frac{p_T^2 R^2}{2M_H \delta M}, \quad (4.13) $$

where $\beta_0 = \frac{1}{12\pi} (11C_A - 2n_f)$, and we shall use $n_f = 5$. Note that we have indicated the exponentiated result by the superfix $P$, which indicates the resummed contribution from primary emissions alone, i.e. excluding secondary emissions that lead to non-global logarithms. We observe here that the perturbative calculations break down at $\lambda = 1/2$, which corresponds to an ISR emission with $k_t \sim \Lambda_{\text{QCD}}$. Using the lower limit in Eq. (4.12), one can see this translates into a value of $\delta M$ that is

$$ \delta M_{\text{NP}} = \frac{\Lambda_{\text{QCD}} p_T R^2}{2M_H} \quad (4.14) $$

where $\delta M_{\text{NP}}$ is the point of breakdown for perturbative calculations. Taking a value of $\Lambda_{\text{QCD}} = 1$ GeV for $p_T = 3$ TeV and $R = 1$ we can deduce that we should not use perturbative results below $\delta M \sim 12$ GeV. One should note that the validity of perturbative results can be extended to smaller values of $\delta M$ if one correspondingly reduces the jet radius $R$ at high $p_T$.

Although we have emphasised that our estimates of the ISR corrections to the signal efficiency are incomplete, even to leading logarithmic accuracy, it is nevertheless of interest to compare these results to MC event generators. This is at least, in part, because MC generators themselves do not attain full single logarithmic accuracy. They do however contain a number of effects that would be formally subleading from the viewpoint of our calculation but could be of non-negligible significance numerically. Hence, while we do not intend to make a detailed quantitative comparison, we do expect to find qualitative similarities with MC results.

To make this comparison, we generate $pp \rightarrow ZH$ events at 14 TeV using Herwig++ 2.7.0 with the UE-EE-5-MRST tune [183] and constrain the Higgs and Z boson to decay hadronically and leptonically respectively. Each generated event is directly handed over to the Rivet package [107], which implements our analyses. We tag the signal jet as the highest $p_T$ Cambridge/Aachen [84, 196] jet with $R = 1$ as implemented in FastJet package [106] and plot the fraction of jets that lie in mass in the window $M_H \pm \delta M$ as a function of a generator level cut on jet transverse momentum for three separate values of $\delta M$. To make a comparison with our ISR results we omit FSR and non-perturbative corrections including hadronisation and underlying event (UE), switching them off for the MC results. The resulting comparison is shown in Fig. 4.1 where results are displayed for the ratio of the signal efficiency to the lowest order
4.2. RESULTS FOR PLAIN JET MASS

Figure 4.1: Comparison of MC (left) and analytic (right) Eq. (4.12) tagging efficiencies for a range of mass windows as a function of a generator level cut on minimum jet transverse momentum $p_T$. This result has been generated using HERWIG++ 2.7.0 for $pp \rightarrow ZH$ at 14 TeV with the $Z$ decaying leptonically and $H \rightarrow b\bar{b}$, setting $M_H = 125$ GeV. We have tagged the signal jet as the highest $p_T$ Cambridge/Aachen jet with $R = 1$. In this figure we have generated events at parton level with ISR only and divided out the contribution due to the lowest order result in both panels for clarity.

result. We observe that the MC signal efficiency decreases with transverse momentum for all values of $\delta M$, as indicated by our exponentiated analytical result in Eq. (4.12). As one makes the mass window wider, i.e. choosing a larger $\delta M$, one obtains a smaller Sudakov suppression; hence, we observe a larger signal efficiency but begin to lose the association with a well defined signal peak.

4.2.2 Final state radiation

For the case of plain jet-mass we would expect that the correction due to final-state radiation can be neglected in our region of interest: $p_T \gg M_H$ or equivalently $\Delta \ll 1$. Physically, FSR is associated with the $b\bar{b}$ dipole originating from Higgs decay. It is captured within the fat jet as long as the FSR gluons are not radiated at angles beyond those corresponding to the jet radius $R$. However, due to angular ordering, we would expect that most of the FSR radiation from the $b$ quarks is emitted at angles smaller than $\theta_{bb}^2$. In this limit, the final state emission is always recombined inside the fat jet. To be more precise, large-angle radiation beyond the jet-radius $R$ is cut off by the ratio of the dipole size ($b\bar{b}$ opening angle, given by $\Delta/(z(1-z))$, to the jet radius squared.

Upon integration over $z$, such corrections translate into terms varying at most as $\Delta \ln \Delta$, which we shall neglect as they vanish with $\Delta$. We have verified this with MC and find that for sufficiently large transverse momenta, the correction due to final state
radiation is of negligible magnitude ($\mathcal{O}(0.5\%)$) when compared to ISR ($\mathcal{O}(20\%)$) for $R = 1$ and $p_T \gtrsim 500$ GeV. We shall need to consider FSR more carefully when it comes to analysing the taggers in future sections.

### 4.2.3 Non-perturbative contributions

In order to get a complete picture of the physical effects that dictate the signal efficiency we also need to study how the signal efficiency changes after including non-perturbative effects such as hadronisation and underlying event (UE). In order to estimate these effects we used HERWIG++ 2.7.0 with improved modelling of underlying event $^{197}$ and the most recent UE-EE-5-MRST tune $^{183}$, which is able to describe the double-parton scattering cross section $^{198}$ and underlying event data from $\sqrt{s} = 300$ GeV to $\sqrt{s} = 7$ TeV. It can readily be anticipated that the underlying event contribution, in particular, will significantly degrade the signal mass peak and hence lead to a loss of signal efficiency.

For this study, we consider all final state hadrons to be stable, therefore we switch off the decay handler module in HERWIG++. In doing so, we eliminate the possibility of $b$ flavour hadrons decaying into invisible particles such as neutrinos. If one includes hadronic decay via invisible particles, one notices a universal reduction in signal efficiency for each tagger due to a loss of signal mass resolution. This is particularly important for the jets formed from the decay $H \rightarrow bb$ as compared to $W/Z$ jets because these electroweak bosons have stronger couplings to light quarks. For further information on experimental techniques to mitigate the impact of these particular sources of missing transverse energy, see for example Ref. $^{199}$. We also assume a $b$-tagging efficiency of 100%, which is sufficient for a relative comparison of tagger performance and behaviour. The reader is referred to Ref. $^{92}$ for a discussion on the impact of $b$-tagging efficiency on signal efficiency.

In Fig. 4.2 we see how non-perturbative effects such as hadronisation and underlying
event affect the signal efficiency when using plain jets tagged with $\delta M = 16$ GeV. One immediately notices that whilst hadronisation has a more moderate effect on the signal efficiency, which however increases with $p_T$ (more precisely like $\sqrt{p_T}$, see Ref. [87]), the dominant contribution comes from underlying event contamination. UE reduces the efficiency at $p_T = 3$ TeV from about 60 percent to around 20 percent. This implies simply that one needs to consider removal of the UE for efficient tagging, which we shall discuss later in the context of boosted object taggers. We have also presented results here for $R = 1$, whilst the averaged UE contribution to the squared jet mass varies as $R^4$ [87]; thus, working with smaller $R$ jets, one may expect this contribution to be less significant. One should of course also consider the presence of considerable pile-up contamination, which we do not treat in this chapter (see [88–90] for discussion of pileup subtraction techniques), but to which the plain jet mass will also be very susceptible.

4.3 Trimming

Trimming [110], as defined in Algorithm [8], takes all the particles in a jet defined with radius $R$ and reclusters them into subjets using a new jet definition with radius $R_{\text{trim}} < R$. It retains only the subjets that carry a minimum fraction $f_{\text{cut}}$ of the original jet transverse momentum $p_T^{(\text{subjet})} > f_{\text{cut}} \times p_T^{(\text{jet})}$ and discards the others. The final subjets are merged to form the trimmed jet.

It is standard to use the Cambridge-Aachen (C/A) jet algorithm [84,196] for substructure studies with trimming (and other taggers) and this is what we shall employ here.

4.3.1 Lowest order result

Compared to the plain jet mass, trimming already has a more interesting structure even without considering any additional radiation. If the opening angle between the $b\bar{b}$ pair is less then $R_{\text{trim}}$ then the trimming constraints are inactive. However, if this angle is greater than $R_{\text{trim}}$, one removes the softer particle when its energy fraction is below $f_{\text{cut}}$. The result for signal efficiency is therefore given by an integral over $z$ that can be expressed as

$$\epsilon_S^{(0)} = \int_0^1 dz \left( 1 - \Theta \left( f_{\text{cut}} - \min \left[ z, 1 - z \right] \right) \Theta \left( \frac{\Delta}{R_{\text{trim}}^2} - z \left( 1 - z \right) \right) \right). \quad (4.15)$$

Strictly we should also have written the condition for the hard prongs to be inside the fat jet as we did for the plain jet case. However, because this condition only results in terms varying as $\Delta/R^2$ we shall neglect it here, consistent with our approximation $\Delta/R^2 \ll 1$.

The subtracted term in Eq. (4.15) represents the removal of any prong that has
energy fraction below $f_{\text{cut}}$, in the region where trimming is active. Evaluating the integral in Eq. (4.15) gives the result

$$
\varepsilon_S^{(0)} = (1 - 2f_{\text{cut}}) \Theta(1 - 2f_{\text{cut}}) + \sqrt{1 - \frac{4\Delta}{R_{\text{trim}}^2}} \Theta\left(\frac{1}{4} - \frac{\Delta}{R_{\text{trim}}^2}\right) \Theta\left(f_{\text{cut}} - \frac{1}{2}\right)
$$

$$
+ \left(2f_{\text{cut}} - 1 + \sqrt{1 - \frac{4\Delta}{R_{\text{trim}}^2}}\right)
\times \Theta\left(\frac{1}{4} - \frac{\Delta}{R_{\text{trim}}^2}\right) \Theta\left(\frac{1}{2} - f_{\text{cut}}\right) \Theta\left(f_{\text{cut}} - \frac{1}{2}\right) \left(1 - \sqrt{1 - \frac{4\Delta}{R_{\text{trim}}^2}}\right)
$$

(4.16)

For now we shall consider values of $f_{\text{cut}}$ that are standard in trimming analyses and therefore are considerably smaller than $1/2$. For such choices of $f_{\text{cut}}$ the second term on the first line in Eq. (4.16), which requires $f_{\text{cut}} > 1/2$, clearly does not contribute.

While the result in Eq. (4.16) is general, let us, for illustrative purposes, consider the remaining terms for values of $R_{\text{trim}}$ not too small, such that $\Delta/R_{\text{trim}}^2 \ll 1$. Then Eq. (4.16) implies a transition point at $\Delta \simeq f_{\text{cut}} R_{\text{trim}}^2$, which translates to a transition point at $p_T \simeq M_H/(\sqrt{f_{\text{cut}} R_{\text{trim}}})$. In the previous chapter we noted that the mass distribution for background jets also had transition points at $M_j^2/p_T^2 \simeq f_{\text{cut}} R^2$ and $M_j^2/p_T^2 \simeq f_{\text{cut}} R_{\text{trim}}^2$ [14,116]. The latter transition point is coincident with that reported above for the signal and corresponds to the minimal jet mass that can be obtained with trimming for a splitting with opening angle $R_{\text{trim}}$. As one increases the $p_T$ beyond $M_j/(\sqrt{f_{\text{cut}} R_{\text{trim}}})$ the background distribution starts to grow due to the onset of double logarithmic behaviour, thereby increasing the background mistag rate. Below this value of $p_T$, the signal efficiency is $p_T$ independent and equal to $1 - 2f_{\text{cut}}$. In contrast, at a $p_T$ above this value the signal efficiency is given by $\sqrt{1 - \frac{4\Delta}{R_{\text{trim}}^2}}$ and hence acquires a $p_T$ dependence. We remind the reader that these results apply specifically to Higgs decay and for processes involving $W/Z$ tagging different results will be obtained due to the different splitting functions involved in hadronic $W/Z$ decay.

In Fig. 4.3 we compare the signal efficiency using HERWIG++ 2.7.0 for trimming applied to boosted Higgs jets with no ISR, FSR or non-perturbative effects to the analytical calculation in Eq. (4.16). We generate the tagging efficiency with two different $f_{\text{cut}}$ values and for $R_{\text{trim}} = 0.3$ as a function of $p_T$ for both MC (left) and analytics (right). One observes, as we would expect, that the MC clearly reproduces the analytic behaviour of the tagger at lowest order and, for our choice of parameters, the expected transition points at around 1320 GeV and 1860 GeV for $f_{\text{cut}} = 0.1$ and $f_{\text{cut}} = 0.05$ respectively.
4.3. TRIMMING

4.3.2 Initial state radiation

Let us consider the action of trimming on ISR and compare to the case of the plain jet. For the plain jet we found a large logarithmic term that results in loss of signal with increasing $p_T$. On the other hand, we would expect trimming to substantially remove ISR radiation and hence wish to check the impact on the logarithmically enhanced terms that emerge from considering soft ISR. The key difference with the plain jet case is that when the angle between the ISR gluon and the jet axis exceeds $R_{\text{trim}}$, the soft gluon is retained only if it has $k_t/p_T$ greater than $f_{\text{cut}}$, where $k_t$ is the transverse momentum of the soft gluon. If the $k_t$ fraction is below $f_{\text{cut}}$ the ISR emission is removed by trimming, thus not contributing to the jet mass. Hence, in this region there is a complete cancellation with virtual corrections. Alternatively, if the ISR falls into the trimming radius, we always retain the emission, much like the plain jet case. These constraints on real emission can be expressed as:

$$
\Theta_{\text{ISR,trim}} = \Theta (\theta^2 - R_{\text{trim}}^2) \left( \Theta (x - f_{\text{cut}}) \Theta \left( \frac{2M_H\delta M}{p_T^2 (\theta^2 + \Delta)} - x \right) + \Theta (f_{\text{cut}} - x) \right) \\
+ \Theta (R_{\text{trim}}^2 - \theta^2) \Theta \left( \frac{2M_H\delta M}{p_T^2 (\theta^2 + \Delta)} - x \right),
$$

where we defined $x$ as $k_t/p_T$ and $\theta^2 = \eta^2 + \phi^2$ is the angle between the ISR gluon and the fat jet axis.

One can then repeat the calculation carried out for the plain jet mass in the previous
CHAPTER 4. JET SUBSTRUCTURE METHODS FOR SIGNAL JETS

section using the constraint in Eq. (4.17). Taking the ISR emission probability in the
eikonal approximation as before, and incorporating virtual corrections we get (in a
fixed-coupling approximation)

\[ \frac{\varepsilon_{S,ISR}}{\varepsilon_{S}^{(0)}} = 1 + C_F \frac{\alpha_s}{\pi} \int_0^1 \frac{dx}{x} d\theta^2 \left[ \Theta_{ISR,\text{trim}} - 1 \right]. \quad (4.18) \]

We can evaluate this integral straightforwardly; the result obtained, after discarding
terms that are power suppressed in \( \Delta \), has two distinct regimes: for \( f_{\text{cut}} > \frac{2 M_H \delta M}{p_T^2} \) one
gets an answer of the form

\[ \frac{\varepsilon_{S,ISR}}{\varepsilon_{S}^{(0)}} \approx 1 - C_F \frac{\alpha_s}{\pi} \left( R^2 \ln \frac{1}{f_{\text{cut}}} + R_{\text{trim}}^2 \ln \left( \frac{f_{\text{cut}} p_T^2 R_{\text{trim}}^2}{2 M_H \delta M} \right) \Theta \left( f_{\text{cut}} - \frac{2 M_H \delta M}{p_T^2 R_{\text{trim}}^2} \right) \right), \quad (4.19) \]

while for \( f_{\text{cut}} < \frac{2 M_H \delta M}{R^2 p_T^2} \), we get

\[ \frac{\varepsilon_{S,ISR}}{\varepsilon_{S}^{(0)}} \approx 1 - C_F \frac{\alpha_s}{\pi} R^2 \ln \frac{R^2 p_T^2}{2 M_H \delta M} \cdot \quad (4.20) \]

Eqs. (4.19) and (4.20) tell us that for sufficiently large \( f_{\text{cut}} \), i.e. above \( \frac{2 M_H \delta M}{R^2 p_T^2} \) one
eliminates the logarithm we obtained for the plain jet mass replacing it by a smaller
\( \ln 1/f_{\text{cut}} \), provided one chooses \( f_{\text{cut}} \) not too small. On the other hand, for smaller \( f_{\text{cut}} \) we see a transition to the logarithmic dependence seen for the plain mass.

The second term in equation Eq. (4.19) represents the region of integration with
\( \theta^2 < R_{\text{trim}}^2 \). This term vanishes as \( R_{\text{trim}} \rightarrow 0 \) and suggests that choosing smaller \( R_{\text{trim}} \)
values will result in less contamination from ISR, as one may readily expect. However,
when studying FSR radiative corrections, we will see that we can not choose \( R_{\text{trim}}^2 \) too small, i.e. \( R_{\text{trim}}^2 \ll \Delta \), due to degradation of the jet due to FSR losses. If one
chooses \( R_{\text{trim}}^2 \ll 1 \) but of order \( \Delta \), then within our small \( \Delta \) approximation we can simply ignore this term. If, on the other hand, one chooses \( R_{\text{trim}} \) not too small then
at very high \( p_T \) one should also consider the presence of this term, which appears only
for \( f_{\text{cut}} > \frac{2 M_H \delta M}{R^2 p_T^2} \). For most practical purposes, with commonly used parameter values,
this term can safely be ignored. For example with \( f_{\text{cut}} = 0.1 \), \( R_{\text{trim}} = 0.3 \) and \( \delta M = 16 \)
GeV, even at \( p_T = 3 \) TeV it only contributes order 10 percent corrections relative to
the main \( \ln 1/f_{\text{cut}} \) piece.

In principle, we should resum the logarithms of \( f_{\text{cut}} \) that are obtained with trim-
ing. However, such a resummation is beset by additional non-global and clustering
logarithms and therefore highly involved. Moreover, the \( \ln 1/f_{\text{cut}} \) terms play only a
modest role numerically, for typical choices of \( f_{\text{cut}} \sim 0.1 \), thus their resummation is
not particularly motivated on phenomenological grounds. We note that \( \ln 1/f_{\text{cut}} \) en-
hanced terms are also produced in the corresponding calculations for QCD background
4.3. TRIMMING

Figure 4.4: Comparison of MC (left) and analytic Eq. (4.18) (right) trimming \( R_{\text{trim}} = 0.3 \) tagging efficiencies for a range of \( f_{\text{cut}} \) values as a function of a generator level cut on jet transverse momentum. This result has been generated using HERWIG++ 2.7.0 at parton level with ISR only for \( H \rightarrow b\bar{b} \) jets, setting \( M_H = 125 \) GeV and \( \delta M = 16 \) GeV. The transition points correspond to the change from plain jet mass like behaviour to a \( \ln f_{\text{cut}} \) term, discussed in the main text.

Let us then compare the main features of our simple analytical NLO approximation, augmented to include running coupling effects, to what is seen in MC event generators. In Fig. 4.4 we again compare our analytical approximations, with running coupling effects as in the plain mass case Eq. (4.13), to HERWIG++ 2.7.0. For the MC studies we turn on ISR effects with boosted \( H \rightarrow b\bar{b} \) jets for a range of \( f_{\text{cut}} \) values, as a function of jet \( p_T \), keeping \( \delta M \) fixed at 16 GeV. Plotting the ratio of the ISR corrected signal efficiency to the lowest order result, we can see that the approximate NLO analytic result reproduces the MC trends reasonably well. For values of \( f_{\text{cut}} = 0.1 \) and 0.05 we do not obtain a transition over the range of \( p_T \) values shown (transition points are expected at roughly 200 GeV and 280 GeV, which are beyond the range shown) and none is seen in the MC plots. The behaviour over the entire plotted \( p_T \) range is quite flat with \( p_T \) because it depends mainly on \( \ln 1/f_{\text{cut}} \), with running coupling and uncalculated subleading effects (in the case of the MC results) providing the weak \( p_T \) dependence that is observed. Instead, for \( f_{\text{cut}} = 0.005 \), we would anticipate a plain mass like degradation of the signal efficiency until the transition at about 890 GeV and then for higher \( p_T \) a flatter behaviour with \( p_T \), which is consistent with MC results. This relative flatness over a large range of \( p_T \) is of course in contrast to the pure mass cut case.
4.3.3 Final state radiation

Let us consider the response of trimming to final state radiation. In principle there are a number of parameters to be considered, in particular $f_{\text{cut}}$, $\Delta$ and $R_{\text{trim}}$ as well as the mass window $\delta M$ and transverse momentum $p_T$. Final state radiation, when not recombined into the fat jet, results in a shift in mass that can cause the resulting jet to fall outside the mass window $\delta M$. Imposing a veto on soft FSR that degrades the jet mass results in the appearance of large logarithms, the structure of which we examine here. Additionally, a relatively hard FSR gluon can also result in one of the primary $b$ quarks energy fraction falling below the asymmetry cuts that are used in taggers, resulting in a loss of signal efficiency. Such hard configurations can still come with collinear enhancements and so their role should also be considered.

In order for an FSR gluon to be removed by trimming it has to be emitted at an angle larger than $R_{\text{trim}}$ with respect to both the hard primary partons. In addition, its energy, expressed as a fraction of the fat jet energy, must fall below the $f_{\text{cut}}$ cutoff. Lastly, for the resulting jet to be retained, the consequent loss in mass must be less than $\delta M$. One can therefore write the following result for real emission contributions, valid in the soft limit where the gluon energy $\omega \ll p_T$:

$$
\varepsilon^{(1)}_{\text{S,FSR,REAL}} = C_F \frac{\alpha_s}{\pi} \int_{f_{\text{cut}}}^{1-f_{\text{cut}}} dz \int \frac{d\omega}{2\pi} \frac{d\Omega}{\omega} (b\bar{b}) (bk) \Theta_{\text{FSR}} \Theta \left( f_{\text{cut}} - \frac{\omega}{p_T} \right) \Theta (\delta M - (M_H - M_j)), \quad (4.21)
$$

where $d\Omega$ is the solid angle element for the emitted gluon and the spatial distribution of radiation has been expressed in terms of the standard antenna pattern with the notation $(ij) = 1 - \cos \theta_{ij}$. The condition $\Theta_{\text{FSR}}$ simply represents the constraint that the angular integration should be carried out over the region where trimming is active, i.e. when the emitted gluon makes an angle larger than $R_{\text{trim}}$ with both $b$ and $\bar{b}$. Moreover, there is an additional step function that constrains the gluon energy fraction to be below $f_{\text{cut}}$ and the factor $\Theta (\delta M - (M_H - M_j))$ represents the constraint on real emissions due to the mass window $\delta M$.

There are three distinct regimes one can consider according to the value of $R_{\text{trim}}$. Firstly, when $R_{\text{trim}} \ll \theta_{b\bar{b}}$ one can expect a collinear enhancement with a logarithm in $R_{\text{trim}}$, which accompanies a soft logarithm arising out of the $\delta M$ constraint. This is the most singular contribution one obtains for trimming so we analyse it in more detail below. On the other hand, in the region where $R_{\text{trim}} \sim \theta_{b\bar{b}}$, there will be no collinear enhancement and one obtains a purely soft, single logarithm. In this region trimming is similar to pruning and the mMDT as far as FSR is concerned, and we shall comment on these results in somewhat more detail in the next section. Finally, in the region
4.3. TRIMMING

Figure 4.5: A configuration that contributes to the FSR correction to trimming signal efficiency in the region $R_{\text{trim}} \ll \theta_{b\bar{b}}$. Here gluon $k$ is emitted collinear to the $b$ quark with momentum fraction $x$ outside a cone with radius $R_{\text{trim}}$.

where $R_{\text{trim}} \gg \theta_{b\bar{b}}$, the FSR correction for trimming becomes more like the plain jet whereby large angle corrections are strongly suppressed.

For the soft and collinear enhanced region $R_{\text{trim}} \ll \theta_{b\bar{b}}$, let us perform a more detailed calculation. Examining the loss in jet mass by removing a FSR emission $k$, in more detail, one has the following expression:

$$M_H^2 - M_j^2 = 2 (p_b \cdot k + p_{\bar{b}} \cdot k) = \omega \left( E_b \theta_{bk} + E_{\bar{b}} \theta_{\bar{b}k} \right).$$  \hspace{1cm} (4.22)

Consider first that the gluon $k$ is emitted collinear to the $b$ quark with momentum fraction $x$, as shown in Fig. 4.5. In this limit, one can neglect the $E_b \theta_{bk}^2$ contribution in Eq. (4.22) and set $\theta_{bk}^2 \approx \theta_{\bar{b}k}^2 = \frac{\Delta}{z(1-z)}$. Requiring $M_H - M_j < \delta M$ and neglecting correction terms of order $\delta M/M_H$ relative to $M_H \delta M$ gives the mass window constraint $2M_H \delta M > \omega p_T \Delta / z$. Next, defining $x$ as the energy fraction of the soft gluon with respect to the energy of the hard emitting prong, i.e. $x = \omega / (zp_T)$, one can write the mass window constraint as the condition $x < 2\delta M/M_H$, where we have used the definition $\Delta = M_H^2 / p_T^2$. Likewise, the $f_{\text{cut}}$ cut expressed as a condition on $x$ just gives $f_{\text{cut}} / z > x$.

In the collinear limit, one can also simplify the angular integration in Eq. (4.21), by assuming a simple $d\theta^2 / \theta^2$ form, resulting in a logarithmic contribution, $\ln \frac{\theta_{bk}^2}{R_{\text{trim}}}$. It is also possible to perform the angular integration exactly, i.e. beyond the collinear limit, to account for less singular soft, large-angle contributions. More details of the derivation and results on the angular integration are provided in Appendix A.1.
We can therefore express the soft-collinear contribution to the FSR corrections as:

\[ \varepsilon^{(1)}_{S,\text{FSR}} = 2C_F \frac{\alpha_s}{\pi} \int_{f_{\text{cut}}}^{1-f_{\text{cut}}} dz \ln \frac{\theta_{bb}^2}{R_{\text{trim}}^2} \int dx \ln \frac{f_{\text{cut}}/z - x}{x} \Theta (f_{\text{cut}}/z - x) \left[ \Theta \left( \frac{2\delta M}{M_H} - x \right) - 1 \right] , \] (4.23)

where a factor of 2 has been inserted to account for an identical result from the region where \( k \) is collinear to \( \bar{b} \) rather than \( b \) and the corresponding virtual corrections have been introduced via the “−1” term in square brackets.

Now let us write \( \theta_{bb}^2 = \Delta / (z(1-z)) \) and carry out the integration over \( x \), which gives

\[ \varepsilon^{(1)}_{S,\text{FSR}} = -2C_F \frac{\alpha_s}{\pi} \int_{f_{\text{cut}}}^{1-f_{\text{cut}}} dz \left( \ln \frac{\Delta}{R_{\text{trim}}^2} - \ln(z(1-z)) \right) \ln \frac{f_{\text{cut}}}{z\epsilon} \Theta (f_{\text{cut}} - z\epsilon) , \] (4.24)

where we have introduced \( \epsilon = \frac{2\delta M}{M_H} \). The structure of this result is, in essence, a double logarithmic form with a soft divergence in the limit \( \delta M \to 0 \) and an accompanying collinear divergence when \( R_{\text{trim}} \to 0 \). Let us now take \( R_{\text{trim}}^2 \ll \Delta^3 \) for simplicity ignore the accompanying \( \ln(z(1-z)) \) term, and integrate over \( z \) to generate the following result:

\[ \varepsilon^{(1)}_{S,\text{FSR}} = -2C_F \frac{\alpha_s}{\pi} \ln \frac{\Delta}{R_{\text{trim}}^2} \times \left[ C_1 (f_{\text{cut}}, \epsilon) \Theta \left( f_{\text{cut}} - \frac{\epsilon}{1+\epsilon} \right) + C_2 (f_{\text{cut}}, \epsilon) \Theta \left( \frac{\epsilon}{1+\epsilon} - f_{\text{cut}} \right) \right] , \] (4.25)

where

\[ C_1 = (1 - 2f_{\text{cut}}) + (1 - 2f_{\text{cut}}) \ln \frac{f_{\text{cut}}}{\epsilon} + f_{\text{cut}} \ln f_{\text{cut}} - (1 - f_{\text{cut}}) \ln (1 - f_{\text{cut}}) , \] (4.26)

\[ C_2 = \frac{f_{\text{cut}}}{\epsilon} - f_{\text{cut}} - f_{\text{cut}} \ln \frac{1}{\epsilon} . \]

We note that for values of \( \epsilon \ll f_{\text{cut}} \) the signal efficiency will be dominated by a \( \ln \frac{f_{\text{cut}}}{\epsilon} \) term in the coefficient \( C_1 \). However, the presence of the \( f_{\text{cut}} \) constraint means that, in practice, such logarithms make only modest or negligible contributions for a wide range of values of \( \epsilon \), or equivalently \( \delta M \). This can be contrasted to the case of filtering, computed in Ref. [155], which has an identical collinear divergence as trimming calculated in Eq. (4.25), but the additional absence of an \( f_{\text{cut}} \) condition leads to a much stronger \( \ln \frac{1}{\epsilon} \) enhancement, which needs to be treated with resummation. It is also straightforward to include the effects of hard collinear radiation by considering the full \( p_{\text{gg}} \) splitting function rather than just its divergent \( 1/x \) piece. In this region, it is possible for the quark to fall below the \( f_{\text{cut}} \) threshold and to therefore be removed by

\footnote{In this limit, the leading order result Eq. (4.16) is simply \( 1 - 2f_{\text{cut}} \), i.e. the two subjets never form a single subjet and are always subject to the asymmetry condition \( x > f_{\text{cut}} \).}
4.3. TRIMMING

Such corrections do not come with soft enhancements and produce terms that vanish with either $f_{\text{cut}}$ or $\epsilon$, hence do not have a sizeable numerical effect that would require resummation. For this reason, we do not calculate these terms explicitly, continuing to work in the soft and collinear limit.

To make the above statements more explicit, let us consider the situation at $p_T = 300$ GeV and choose $f_{\text{cut}} = 0.1$. The zeroth order result for signal efficiency is then $\varepsilon_S^{(0)} = 1 - 2f_{\text{cut}} = 0.8$. If one chooses a value of $R_{\text{trim}} = 0.1$ then $\ln (\Delta/R_{\text{trim}}^2) \sim \ln 17$ and one may expect significant (collinear enhanced) radiative corrections. Choosing a larger $R_{\text{trim}} = 0.3$ one can instead reduce $\ln \Delta/R_{\text{trim}}^2$ to $\sim \ln 2$, which is not enhanced and does not require resummation, implying a much more modest FSR contribution. However, we should also examine the effect of this increased $R_{\text{trim}}$ on ISR and UE contributions. For our choice of parameters it is evident from MC studies that we do not pay a significant price in signal efficiency for the increased $R_{\text{trim}}$ value in terms of the ISR contribution. At the same value of $p_T$ the UE contribution for $R_{\text{trim}} = 0.3$ is also small (see Fig. 4.6). If one moves to higher $p_T$, say 3 TeV, one should correspondingly lower $R_{\text{trim}}$. Here one has $\Delta = 0.0017$, hence choosing a value of $R_{\text{trim}} \sim 0.1$ would ensure a smaller FSR contribution as well as reduce the impact of ISR and the underlying event. This illustrates that, by making an appropriate choice of $R_{\text{trim}}$, one can negate large radiative losses due to FSR, without necessarily suffering from large ISR/UE effects. In general, the optimal value of $R_{\text{trim}}$ will involve a trade-off between FSR radiative corrections and ISR/UE effects. We shall return to this point in Section 4.7.

We should also examine the role of soft divergences that are formally important in the $\epsilon \rightarrow 0$ limit. Taking a value of $\delta M = 2$ GeV leads to $\epsilon = 0.032$. For our choice of $f_{\text{cut}} = 0.1$, we have $\ln f_{\text{cut}}/\epsilon \sim \ln 3$, which is not a genuinely large logarithm. The overall coefficient of $-2C_F \alpha_s \pi \ln \Delta/R_{\text{trim}}^2$, which is given by the $C_1$ and $C_2$ terms in Eq. (4.26), is for $\delta M = 2$ GeV, approximately 1.58 and for $\delta M = 10$ GeV approximately 0.34, thus indicating that resummation of soft logarithms is not a necessity. Expressed as a percentage of the tree level result $1 - 2f_{\text{cut}}$, the FSR corrections, as computed in Eq. (4.25), constitute a roughly two percent to ten percent effect for $\delta M$ ranging from 2 GeV to 10 GeV, if one chooses $R_{\text{trim}}^2 \approx \Delta/2$. Hence, we find that even the leading soft-collinear enhanced contribution makes only modest contributions to the signal efficiency, at best comparable to pure order $\alpha_s$ corrections. The main implication of this finding is that full fixed-order calculations or combinations of fixed-order results with parton showers (see Refs. [73,162] for a review of the latter methods), would give a better description of the signal efficiency than pure parton showers. We explore, in somewhat more detail, the role of fixed-order calculations to describe the signal efficiency in Appendix A.2.

---

4As we shall see later, this constitutes a somewhat non-optimal choice for $R_{\text{trim}}$. It is made here for purely illustrative purposes in order to estimate the magnitude of soft, non-collinear enhanced effects.
Figure 4.6: An MC study of the impact of hadronisation and underlying event (UE) on the signal efficiency using the trimmed jet \( f_{\text{cut}} = 0.1 \) as a function of the minimum jet transverse momentum for two different values of \( R_{\text{trim}} \). One can see the impact of hadronisation and underlying event on the signal efficiency in the window \( \delta M = 16 \text{ GeV} \). Details of generation given in Fig. 4.2.

The FSR radiative correction given in Eq. \((4.25)\) is intended to address the formal limit \( R_{\text{trim}}^2 \ll \Delta \). It indicates that choosing extremely small values of \( R_{\text{trim}} \) is problematic due to degradation of the jet from FSR loss. In the opposite limit, i.e. \( R_{\text{trim}}^2 \gg \Delta \), the FSR correction will be negligible, as seen for the case of the plain jet mass. On the other hand, ISR corrections calculated in Eq. \((4.19)\) indicate that large choices of \( R_{\text{trim}} \) may not be optimal due to increased ISR (and UE) contamination. One is thus led to consider the region \( R_{\text{trim}}^2 \sim \Delta \). This is reminiscent of the choice made in pruning for \( R_{\text{prune}}^2 \sim M_j^2/p_T^2 \). In this limit, the behaviour of FSR corrections for trimming is therefore expected to be similar to that for pruning, for which a detailed calculation is carried out in Subsection 4.4.1. We simply note here that, in the region \( R_{\text{trim}}^2 \sim \Delta \), FSR corrections are relatively modest and can be thought of as pure order \( \alpha_s \) corrections, rather than carrying significant logarithmic enhancements.

### 4.3.4 Non-perturbative contributions

Let us now study the impact of non-perturbative corrections to the signal efficiency using trimming on boosted Higgs jets.

In Fig. 4.6 we show the signal efficiency for a boosted Higgs signal jet after application of trimming with parameters \( R_{\text{trim}} = 0.3 \) (left), \( R_{\text{trim}} = 0.1 \) (right) with \( f_{\text{cut}} = 0.1 \), as a function of the minimum jet transverse momentum. One can see that hadronisation has little effect on the tagging rate of signal jets, due to the action of trimming on contributions that are soft and wide angle in the jet. UE has a larger impact on the
signal efficiency due to soft contamination that is not checked for energy asymmetry. In other words, inside the trimming radius the algorithm is inactive, and we automatically include all contamination coming from UE, which inside this region would contribute, on average, to a change in the jet mass squared varying as $R_{\text{trim}}^4$. The UE contribution could thus be substantially reduced by choosing a smaller $R_{\text{trim}}$. Reduction of $R_{\text{trim}}$ is, in particular, required at higher $p_T$, as evident from Fig. 4.6. Finally, in contrast to the plain jet result in Fig. 4.2, one notes a significant reduction in sensitivity to non-perturbative effects when tagging signal jets using trimming.

4.4 Pruning and mMDT

In this section we shall study pruning and the modified mass drop tagger. We describe these methods together because, unlike for the case of the QCD background studied in detail in Chapter 3 and Ref. [116] whereby the taggers can exhibit substantial differences, for the signal one finds similar behaviour.

4.4.1 Pruning

Pruning, as defined in Algorithm 6, uses the initial jet to calculate a pruning radius that is dependent on the mass of the jet and its transverse momentum $R_{\text{prune}} = R_{\text{fact}} \times \frac{2M_j}{p_T}$ where $R_{\text{fact}}$ is a parameter of the tagger. It proceeds by reclustering the jet, at each step checking if both the angle between the two objects $i$ and $j$ is greater than the pruning radius $\Delta R_{ij} > R_{\text{prune}}$ and the splitting is $p_T$ asymmetric, i.e. $\min(p_{T_i}, p_{T_j}) < z_{\text{cut}} \times p_T(i+j)$. If these conditions are both true, pruning discards the softer of $i$ and $j$, else $i,j$ are combined as usual. This is repeated for each clustering step until all particles are either discarded or combined into the final pruned jet. For this study we use the default value $R_{\text{fact}} = \frac{1}{2}$ [108] and again use the C/A algorithm to both find and recluster the jets.

At zeroth order the two signal prongs are always at an angle larger than $R_{\text{prune}}$ and so the result is simply $1 - 2z_{\text{cut}}$. For initial state radiation one can consider pruning to be similar to trimming with $R_{\text{trim}}$ replaced by $R_{\text{prune}}$. The pruning radius is given by

$$R_{\text{prune}}^2 = \frac{(p_1 + p_2 + k)^2}{p_T^2} \approx \Delta + \frac{2p_H \cdot k}{p_T^2} \approx \Delta + x\theta^2,$$

(4.27)

where $\theta$ is the angle between the soft gluon and the Higgs direction (or equivalently, with neglect of recoil against soft ISR, the fat jet axis).

One then ends up comparing the gluon angle $\theta^2$ to $x\theta^2 + \Delta$ and thus for sufficiently soft emissions, i.e. in the limit $x \to 0$, responsible for logarithmic corrections, one can just replace $R_{\text{prune}}^2$ by $\Delta$. The situation is therefore identical to trimming but with $R_{\text{trim}}^2$ replaced by $\Delta$. Because we work in the limit $\Delta \ll 1$ we can neglect corrections varying as powers of $\Delta$ under this replacement, i.e. we can neglect the second term in
The result should then be identical to trimming in that one should obtain a $\ln \frac{1}{z_{\text{cut}}}$ dependence with a transition to the plain mass behaviour visible for smaller $z_{\text{cut}}$ values as in Fig. 4.4. We have verified that this is indeed the case with MC and that the efficiencies for pruning and trimming look essentially identical in terms of the response to ISR. An MC plot comparing ISR for all taggers is shown in the next section (Fig. 4.10), after we discuss the cases of mMDT, Y-pruning and Y-splitter.

Next we discuss briefly the situation with regard to FSR corrections, in the context of pruning. Again one can employ the insight we gained in the previous section for the FSR calculation of trimming. For the case of pruning there is no collinear enhancement because radiation that is lost is emitted at an angle (with respect to both hard prongs) larger than $R_{\text{prune}}^2 \sim \Delta = z(1-z)\theta_{bb}^2$, i.e. essentially of order $\theta_{bb}^2$. This angular integration produces a finite $O(1)$ coefficient and we will thus obtain only a single logarithmic enhancement, that results from the loss of soft radiation at relatively large angles, i.e. those comparable to the $b\bar{b}$ dipole size. The corresponding loss in mass can be expressed as

$$M_H^2 - M_j^2 = (p_1 + p_2 + k)^2 - (p_1 + p_2)^2 = 2k \cdot (p_1 + p_2) \approx \omega p_T \left( \theta^2 + \Delta \right), \quad (4.28)$$

where $\theta$ is the angle between the gluon and the jet axis. Noting that $\theta^2$ is at most of order $\Delta$ (contributions from the region where $\theta^2 \gg \Delta$ are negligible due to angular ordering of soft radiation) one can replace $M_H^2 - M_j^2$ by $\omega p_T \Delta$, ignoring any multiplicative factors of order one, that lead to only subleading logarithmic corrections. The condition on gluon energy due to the mass window constraint is then $\frac{2\delta M}{M_H} > \frac{\omega}{p_T}$. One also requires a constraint on the gluon energy such that it fails the $z_{\text{cut}}$ condition.

Denoting $\frac{\omega}{p_T}$ as $x$, and accounting for virtual corrections as for the case of trimming, we have the expression for FSR corrections to pruning:

$$\varepsilon_{S,\text{FSR}}^{(1)} \simeq -C_F \frac{\alpha_s}{\pi} \int_{z_{\text{cut}}}^{1-z_{\text{cut}}} \frac{dz}{z} \int \frac{dx}{x} \Theta(z_{\text{cut}} - x) \Theta(x - \epsilon) \int \frac{(bb)}{(bk)(\bar{b}k)} \frac{d\Omega}{2\pi} \Theta_{\text{FSR}}^{\text{prune}}, \quad (4.29)$$

where the condition $\Theta_{\text{FSR}}^{\text{prune}}$ is simply the condition that the gluon is emitted outside an angle $R_{\text{prune}}$ with respect to both hard prongs. The angular integration and $z$ integration is performed in Appendix A.1, hence carrying out the integrals over $x$ we obtain the result:

$$\varepsilon_{S,\text{FSR}}^{(1)} = -C_F \frac{\alpha_s}{\pi} \frac{2\pi}{\sqrt{3}} \ln \frac{z_{\text{cut}}}{\epsilon}, \quad z_{\text{cut}} > \epsilon. \quad (4.30)$$

This result suggest that logarithmic enhancements for pruning are, in principle, present

---

5Strictly, with the precise definition of the $z_{\text{cut}}$ condition we would have to consider a cut on $\omega$ normalised to the energy of the corresponding declustered prong, i.e. $z p_T$ or $(1-z) p_T$ but at single-logarithmic accuracy one can just always take a cut on $\omega/p_T$. 160
4.4. PRUNING AND MMDT

For $\epsilon \ll z_{\text{cut}}$. However, even with a choice of $\delta M$ as low as 2 GeV, with $z_{\text{cut}} = 0.1$ we obtain a modest logarithm $\sim \ln 3$, implying that soft enhanced effects can be neglected. To therefore assess FSR corrections in more detail, as we found for trimming, it is necessary to go beyond the soft approximation and study hard corrections, which we explore further in Appendix A.2. However, it should be apparent that radiative corrections due to FSR corrections represent essentially order $\alpha_s$ corrections without significant log enhancements over a wide range of values of $\delta M$, $p_T$ and $z_{\text{cut}}$. We can exploit this stability against radiative corrections in optimising the tagger parameters, which we do in Section 4.7.

To obtain a complete picture of pruning we also need to account for non-perturbative corrections arising from hadronisation and UE corrections. We shall comment on MC results for these aspects together with the mMDT results in Subsection 4.4.3.

4.4.2 Modified mass drop tagger

Here we shall consider the modified mass drop tagger along similar lines. We start by recalling the definition first of the regular mass drop tagger, defined in Algorithm 3: The mass drop tagger (MDT) [92] is intended for use with jets clustered using the C/A algorithm. For each jet $j$ one applies the algorithm:

1. If jet $j$ contains subjets, split $j$ into two subjets $j_1$ and $j_2$ by undoing the last stage of clustering such that $M_{j_1} > M_{j_2}$.

2. If there is a significant mass drop $\mu \times M_j > M_{j_1}$ and the splitting is not too asymmetric $\min(p_{T,j_1}, p_{T,j_2}) \Delta R_{j_1,j_2} > y_{\text{cut}}$ deem $j$ to be a “tagged” jet and exit the loop.

3. Else relabel $j_1 = j$ and repeat from step 1.

The modified mass drop tagger (mMDT) corrected a flaw in the mass drop tagger so that in the event that the mass drop and asymmetry conditions are not satisfied one follows the more energetic (higher $p_T$ branch) rather than the heavier branch $j_1$, as advocated above. This is not only physically relevant (as one ensures that one identifies hard substructure rather than, for a small fraction of events, following soft massive jets) but also significantly ameliorates the logarithmic dependence of the perturbative structure for calculations related to QCD background jets, rendering for instance the QCD jet mass distribution purely single logarithmic and free from non-global logarithms.

One other observation that was made in the last chapter and in Ref. [116] concerned the role of the mass drop parameter $\mu$ itself. There it was noted that the mass drop condition had a negligible impact on the result obtained for the jet mass distribution for QCD background jets and hence that it was possible to entirely ignore the mass
drop requirement. In this chapter we shall consider this variant of the mMDT, where we do not impose the mass-drop condition but just the asymmetry requirement via a $y_{\text{cut}}$ cut-off.

At zeroth order we obtain a signal efficiency $\varepsilon_S^{(0)} = 1 - 2y_{\text{cut}}$ coming from the asymmetry cut, which is the same result as for pruning. As far as the response to ISR is concerned, one can straightforwardly see that the general behaviour will be similar to the taggers we have considered before. Consider a fat jet consisting of a $b\bar{b}$ pair and an ISR gluon. If the gluon makes an angle less than $\theta_{b\bar{b}}$ with either of the hard prongs of the fat jet then, during the declustering step, we will break the jet into a massless prong and a prong with a small mass consisting of a quark and the soft gluon. In the soft limit the asymmetry condition will pass if the hard prongs are sufficiently energetic, i.e. exactly as at zeroth order, and the soft ISR will contaminate the jet. Such corrections will vanish with $\Delta$ just like for the case of pruning and hence we can ignore them here, at high $p_T$. If the soft gluon forms an angle with respect one of the hard prongs greater than $\theta_{b\bar{b}}$, it emerges first on declustering the jet. If it fails the asymmetry condition it will be removed and its effects cancel against virtual corrections. If it passes the asymmetry cut, one obtains, in the small $\Delta$ limit, essentially a logarithm in $y_{\text{cut}}$ as calculated for pruning and trimming. One should note that, if one uses the asymmetry condition, with the $y_{\text{cut}}$ measure exactly as defined above, the condition for the gluon to pass the asymmetry cut can be expressed as $x^2\theta^2/(\Delta + x\theta^2) > y_{\text{cut}}$, where $x = \omega/p_T$. Combining this with the condition $\theta^2 < R^2$, on the angular integration, we get the constraint:

$$x > \frac{y_{\text{cut}}}{2} \left( 1 + \frac{4\Delta}{y_{\text{cut}}R^2} \right), \quad (4.31)$$

which for $y_{\text{cut}} \gg \Delta/R^2$ reduces to the same constraint as for the case of pruning and trimming, i.e. $x > y_{\text{cut}}$. The main effect of this slightly different relationship between the gluon energy and $y_{\text{cut}}$ manifests itself, only for rather small $y_{\text{cut}}$ values and at low $p_T$, as a change in the transition point to the plain jet mass like behaviour. The position of the transition can be computed as before by setting the right hand side of Eq. (4.31) equal to $2M_H\delta M/p_T^2$, recalling that we take the fat jet radius $R = 1$. Let us consider $f_{\text{cut}} = 0.005$ and $\delta M = 16$ GeV, for which one obtains a transition point with trimming at $p_T \sim 900$ GeV, below which one sees a plain jet like behaviour. For the same value $y_{\text{cut}} = 0.005$ the corresponding transition point for mMDT occurs at roughly $p_T \sim 400$ GeV, i.e. is absent over the range of $p_T$ considered here. In any case, very small values of $y_{\text{cut}}$ would mean that logarithms in $y_{\text{cut}}$ become large and hence should, in general, be avoided. For reasonable values of $y_{\text{cut}} \sim 0.1$, mMDT behaves essentially identical to pruning and trimming. We have verified all of the above points with MC studies and shall provide a plot comparing the response of taggers to ISR in
the next section, after discussing Y-pruning and Y-splitter (see Fig. 4.10). Lastly, we shall mention that for calculation of FSR corrections in the soft approximation, we do not observe any significant differences between mMDT and pruning. To understand this it is enough to realise that a soft FSR gluon emitted at an angle smaller than \( \theta_{b\bar{b}} \) with respect to either the \( b \) or \( \bar{b} \) direction is not examined for the asymmetry condition and hence does not contribute to a loss in jet mass, implying also the absence of any collinear enhancements. Emissions at an angle larger than \( \theta_{b\bar{b}} = \frac{\Delta}{z(1-z)} \) contribute to a loss in mass and give a soft, single logarithmic contribution identical to that for pruning. The result obtained is identical to Eq. (4.30) with the angular coefficient \( 2\pi/\sqrt{3} \approx 3.63 \) replaced by a coefficient that we have determined numerically. For \( p_T = 3 \) TeV the coefficient we obtain is \( \approx 0.646 \).\(^6\) The key point however, is that no large logarithmic corrections arise due to soft FSR emissions, owing to the presence of the \( y_{\text{cut}} \) cut-off. Once again it would therefore be of interest to study the role of genuinely hard radiative corrections beyond the eikonal approximation, a study we carry out in Appendix A.2.

4.4.3 Non-perturbative effects and MC results

We have analysed the effects of ISR and FSR for both pruning and mMDT and concluded that the taggers exhibit essentially similar behaviour for the case of signal jets. However, our studies have focussed thus far on the perturbative component and hence it is prudent to examine non-perturbative effects before reaching any final conclusions.

In Fig. 4.7 we show the MC results for pruning and mMDT. One observes that the signal efficiency has only a weak dependence on \( p_T \) and that, relative to the lowest order expectation of \( \varepsilon_S^{(0)} = 0.8 \), at high \( p_T \), one sees a roughly 10 percent difference for the full parton level result with radiative corrections. One also observes a remarkable similarity between the two taggers over the entire \( p_T \) range as far as parton level results and those including hadronisation are concerned. However, the effect of UE contamination is more clearly visible in the mMDT case towards lower \( p_T \) values, which owes to the larger effective radius \( \theta_{b\bar{b}} = \frac{1}{\sqrt{z(1-z)}} \frac{M_H}{p_T} \) as compared to \( R_{\text{prune}} \approx \frac{M_H}{p_T} \) for pruning, as well as differences in the definitions of the asymmetry parameters \( y_{\text{cut}} \) vs \( z_{\text{cut}} \)\(^7\). Therefore, at lower \( p_T \) it has been standard practice to use the mass drop tagger in conjunction with filtering as suggested in the original paper Ref. [92]. One should also bear in mind the contrasting results of Ref. [116], whereby the impact of non-perturbative effects on the QCD background jet mass was much more pronounced for pruning than for mMDT. In the final analysis one expects the impact on the background to dictate the ultimate

---

\(^6\)This is the same result, \( J(1) \), found by Rubin in Ref. [155] for the coefficient of the FSR soft logarithm for filtering when the filtering radius \( R_{\text{filter}} = \theta_{b\bar{b}} \).

\(^7\)It is of course possible to use mMDT with a \( z_{\text{cut}} \) constraint defined as for pruning instead of \( y_{\text{cut}} \), as was studied in Ref. [116]. This choice would further enhance the similarity we observe for signal jets and is the default choice currently implemented in FastJet 3.1.2 [106].
performance of the taggers, rather than the comparatively small corrections one sees here for the signal, over most of the $p_T$ range studied.

A final point to make about Fig. 4.7 concerns the contrast between the FSR corrections observed for mMDT and pruning to those seen in Fig. 4.6 for trimming. In order to make this comparison we firstly note that for the right hand panel of Fig. 4.6 we have chosen $r_{cut} = 0.1$ and $R_{trim} = 0.1$. Within the $p_T$ range presented, the zeroth order result for trimming is the same as for mMDT and pruning, i.e. $\varepsilon_S = 1 - 2f_{cut}$. Secondly, it is evident from Figs. 4.6 and 4.7 that, while the FSR results for mMDT and pruning show hardly any dependence on $p_T$ over the range studied, the corresponding results for trimming show a more pronounced $p_T$ dependence. This feature already emerges from our simplified fixed-coupling analytics, namely the FSR corrections for trimming depend on $p_T$ via a dependence on $\ln \Delta/R_{trim}^2$, whilst for pruning and mMDT we have shown that the FSR corrections are $p_T$ independent (see e.g. Eq. (4.30) for pruning). We shall return to this point in Section 4.7.

4.5 Y-pruning and Y-splitter

In this section we shall study the Y-pruning modification of pruning suggested in the previous chapter and in Ref. [116] and defined in Algorithm 7, along with the older Y-splitter method [96]. We shall study these two methods together because they have a remarkably similar action on QCD background jets and are particularly effective in removing QCD background in the vicinity of signal peaks for boosted Higgs and electroweak gauge bosons, which makes them potentially valuable tools. However,
they differ significantly from each other and from other taggers in their response to initial state radiation (and even more significantly to UE), for reasons we highlight below, and that were also mentioned in Ref. [116] for the case of Y-pruning. In the next subsection we shall use the insight we gain in the present section to suggest improvements to Y-splitter in particular.

4.5.1 Y-pruning

We begin by examining the case of Y-pruning. We remind the reader that this is a modification of pruning whereby one requires that at least one clustering is explicitly checked for and passes the pruning criteria, else one discards the jet. In this way, one removes spurious configurations whereby all emissions that are left after application of pruning are within an angular distance \( R_{\text{prune}} \) of one another. This configuration implies that these emissions never get examined for an asymmetry condition, resulting in the tagging of jet substructure with only a single hard prong.

A known issue with Y-pruning for the case of signal jets, already discovered in Ref. [116], concerns its response to soft, wide-angle emissions from ISR or UE. Here one can have a situation whereby a soft ISR emission contributes to setting the pruning radius but is itself removed by pruning. If the pruning radius set by the ISR emission is larger than \( \theta_{b\bar{b}} \) then one would discard the resulting jet as it does not satisfy the Y-pruning condition, causing a loss of signal. In the same kinematic region, virtual corrections would lead to a jet that is accepted by Y-pruning (assuming the hard prongs arising from Higgs decay satisfy the asymmetry criterion), and hence contribute to the signal tagging efficiency, as we shall demonstrate below.

We first consider a soft ISR emission with energy (or \( k_t \)) fraction \( x \ll 1 \) that makes an angle \( \theta \) with the fat jet axis (again we neglect recoil against soft ISR); in this configuration we again have that \( R_{\text{prune}}^2 = M_j^2/p_T^2 \approx \Delta + x\theta^2 \). We wish to compute the virtual contribution in the kinematic region where the real ISR is removed, i.e. in the region that satisfies \( \theta^2 > R_{\text{prune}}^2 \) and \( x < z_{\text{cut}} \). Moreover, we require that \( R_{\text{prune}}^2 > \theta_{b\bar{b}}^2 \) so that the hard prongs are within the pruning radius. Thus we have, for the contribution of uncancelled virtual gluons, the extra contribution for Y-pruning, relative to pruning:

\[
\Delta \varepsilon_{\text{S}} = - \frac{2}{\pi} \int_{z_{\text{cut}}}^{1-x_{\text{cut}}} dz \int dx \frac{dx}{x} d\theta^2 \Theta(\theta^2 - \Delta) \Theta(\theta^2 - \frac{\Delta}{x} f(z)) \Theta(R^2 - \theta^2)\Theta(z_{\text{cut}} - x),
\]

where we have defined the function \( f(z) = \frac{1}{z(1-z)} - 1 \) and the overall minus sign indicates that we are considering just the virtual contribution. The first step function in Eq. (4.32) comes from the requirement that the ISR emission lies at a larger angle than \( R_{\text{prune}} \) relative to the jet axis (with \( x \ll 1 \)). The second step function expresses the constraint that the \( b\bar{b} \) opening angle is less than \( R_{\text{prune}} \). Lastly, we require that the ISR radiation is within the fat jet radius \( R^2 > \theta^2 \) and the real contribution is removed.
via the energy condition $x < z_{\text{cut}}$. The integration over $x$ produces the logarithmically enhanced term we seek, here one obtains a logarithm in the ratio of $z_{\text{cut}}$ to $\Delta$. For high $p_T$, one may expect that this logarithm can become large, hence it should have a visible effect for values of $z_{\text{cut}}$ that are not too small. The final result, discarding all other terms that are less singular in the high $p_T$ limit (e.g. those that vanish with $\Delta$), is

$$
\Delta \varepsilon_S \approx -C_F \frac{\alpha_s}{\pi} R^2 \ln \frac{z_{\text{cut}} R^2}{\Delta} \left( \frac{1}{1 + \beta} - z_{\text{cut}} (1 - z_{\text{cut}}) \right) + (1 - 2z_{\text{cut}}) \Theta \left( z_{\text{cut}} (1 - z_{\text{cut}}) - \frac{1}{1 + \beta} \right),
$$

(4.33)

where we defined $\beta = \frac{z_{\text{cut}} R^2}{\Delta}$. Therefore, at high $p_T$, $\Delta \varepsilon_S$ can potentially dominate the normal logarithmic dependence of pruning on $z_{\text{cut}}$. Hence, we can distinguish $Y$-pruning from other taggers by looking at the transverse momentum dependence of the signal response to ISR in the high $p_T$ limit.

In Fig. 4.8 we compare the sum of the analytic result from pruning and the additional contribution described in Eq. (4.33) to MC with ISR for a range of $z_{\text{cut}}$ values. One immediately notices, in both analytical and MC plots, that the $p_T$ dependence of $Y$-pruning is significantly different from that of pruning for the commonly used value

---

Figure 4.8: Comparison of MC (left) and analytic (right) $Y$-pruning signal tagging efficiencies for a range of $z_{\text{cut}}$ values as a function of a generator level cut on jet transverse momentum. This result has been generated using Herwig++ 2.7.0 at parton level with ISR only for $H \rightarrow b \bar{b}$ jets, setting $M_H = 125$ GeV and $\delta M = 16$ GeV. We have divided out the contribution due to the born configuration in both panels for clarity.
4.5. Y-PRUNING AND Y-SPLITTER

Figure 4.9: An MC study of the impact of hadronisation and underlying event (UE) on the signal efficiency using the the Y-pruned jet ($z_{\text{cut}} = 0.1$) as a function of the minimum jet transverse momentum. One can see negligible impact resulting from hadronisation corrections but some degradation coming from underlying event contamination on the signal efficiency in the window $\delta M = 16$ GeV. Details of generation given in Fig. 4.2.

$z_{\text{cut}} = 0.1$, (see Fig. 4.10). In agreement with our expectations, the signal efficiency as given by MC in Fig. 4.8 first increases with $p_T$ (like pruning) and then decreases beyond a transition point, which is the onset of the logarithmic behaviour we have computed in Eq. (4.33). Our calculation indicates that the onset of the logarithm in $z_{\text{cut}}/\Delta$ is for $\beta > 3$, which for $z_{\text{cut}} = 0.1$ and $z_{\text{cut}} = 0.05$ corresponds to approximately 680 GeV and 970 GeV respectively. This is consistent with what is seen in MC, although the transitions are not as sharp as in the analytical result.

As far as final state radiation is concerned, there are no significant differences between Y-pruning and pruning. The soft, large-angle contributions that are responsible for the loss of signal we saw for ISR are strongly suppressed for the case of FSR, due to the colour singlet nature of the parent Higgs particle and angular ordering. We conclude by presenting Fig. 4.9, an MC plot for Y-pruning at both parton level (including both FSR and FSR+ISR) and with non-perturbative corrections. As expected, there is significant loss of signal due to UE contributions for precisely the same reasons as for the case of ISR, also observed in Ref. [116]. In spite of this deficiency, it was also shown in Ref. [116] that, due to its strong suppression of background jets, Y-pruning produced a signal significance that was at least comparable and, at high $p_T$, exceeded that from the other taggers studied (mMDT, pruning and trimming), especially for gluon jet backgrounds. In the next subsection we shall study an older method, Y-splitter, that has a similar action to Y-pruning on background jets, which makes its action on signal worth exploring further.

4.5.2 Y-splitter

The Y-splitter technique was first introduced in Ref. [96] in the context of W boson tagging. The main observation was that the $k_t$ distance measure (as employed in the $k_t$ algorithm [85]) between the two partonic prongs of a $W$ decay tended to be close to the $W$ mass, which is a consequence of a typically symmetric energy sharing between
the two prongs; this is in contrast to the case of QCD background whereby the energy sharing is typically asymmetric (see Section 2.3). To exploit this fact, one takes a fat jet constructed with the $k_t$ algorithm and undoes the last step of the clustering. This produces two prongs that have energy fractions $z$ and $1 - z$. The $k_t$ distance $d_{ij}$ is given, at small opening angles, by the square of the transverse momentum of the softer prong with respect to the direction of the harder prong:

$$d_{ij} = \min(z^2, (1 - z)^2)p_T^2\theta_{ij}^2.$$  

One can either cut directly on this distance by requiring it to be of order $M^2/W$ ($M^2$ in our case) or cut on the ratio of $d_{ij}$ to the jet invariant mass squared $M^2_j = p_T^2 z (1 - z) \theta_{ij}^2$ (see e.g. Ref. [81,104]). In the present case, we shall choose the latter option and hence demand that

$$\frac{d_{ij}}{M^2_j} = \min(\frac{z, 1 - z}{\max(z, 1 - z)}) > y_{\text{cut}}.$$  

If this condition is satisfied, one tags the jet, else, it is discarded. Taking for instance $z < 1/2$ this cut amounts to requiring that $z > y_{\text{cut}}/(1 + y_{\text{cut}}) = y_{\text{cut}} + O(y_{\text{cut}}^2)$. Likewise for $z > 1/2$ one obtains $z < 1 - y_{\text{cut}} + O(y_{\text{cut}}^2)$.

The Y-splitter method has not been as widely used in recent times as some of the other methods we have studied here, though one relatively recent application has been for the purposes of top tagging [200]. Also, a detailed comparison of Y-splitter with N-subjettiness was carried out in Ref. [97].

Let us first briefly consider the action of Y-splitter on QCD background. If one considers a quark jet with an additional soft-collinear gluon emission, then the $y_{\text{cut}}$ condition is active on the gluon energy, which means it regulates the soft divergence associated to gluon emission. The usual double logarithmic structure of the QCD jet mass gives way to a single logarithmic answer precisely as for the mMDT, pruning and Y-pruning methods in the previous chapter:

$$\rho \frac{d\sigma}{d\rho}^{(\text{Y-splitter,LO})} \simeq C_F \frac{\alpha_s}{\pi} \left( \ln \frac{1}{y_{\text{cut}}} - \frac{3}{4} \right), \quad \rho < y_{\text{cut}} R^2,$$

where $\rho = \frac{M^2_j}{p_T^2}$ and we have neglected terms varying as powers of $y_{\text{cut}}$. For $\rho > y_{\text{cut}} R^2$ one obtains a transition to the normal jet mass result. At all orders, the result for Y-splitter can be derived using methods similar to those in Ref. [116]. Because the derivation of this result takes us away from our current focus on signals, we shall not provide it here, but shall do so in the next chapter. The basic fixed coupling result, for small $\rho$, can be expressed in the form:

$$\rho \frac{d\sigma}{d\rho}^{(\text{Y-splitter})} \simeq C_F \frac{\alpha_s}{\pi} \left( \ln \frac{1}{y_{\text{cut}}} - \frac{3}{4} \right) \exp \left[ -\frac{C_F \alpha_s}{2\pi} \ln^2 \frac{1}{\rho} \right],$$

where $\rho = \frac{M^2_j}{p_T^2}$.
which represents a Sudakov suppression of the leading order result. The form of this result is identical to that derived for Y-pruning in the region $\rho < z_{\text{cut}}^2$ and $\alpha_s \ln \frac{1}{z_{\text{cut}}} \ln \frac{1}{\rho} \ll 1$ (see Eq. (5.10b) of Ref. [116]), though subleading logarithmic terms will differ. One can verify this similarity of Y-splitter to Y-pruning, for the case of QCD jets, by examining the results produced by MC and we shall do so in the next subsection.

We shall now study the response of Y-splitter to signal jets, for our case of Higgs decay. At zeroth order the result is similar to that for mMDT and pruning, hence neglecting terms of order $y_{\text{cut}}^2$ one simply gets $\varepsilon_S^{(0)} = 1 - 2y_{\text{cut}}$, which is, as usual, a consequence of the uniform $z$ distribution and the asymmetry cuts on $z$. Beyond zeroth order one should expect very significant differences between Y-splitter and the other taggers. These originate from the response to ISR (and UE/pile-up).

In order to understand the ISR response let us consider our usual configuration of a $b\bar{b}$ pair and a large-angle soft ISR gluon with $\theta \sim 1 \gg \theta_{b\bar{b}}$. For Y-splitter, there are two configurations of interest. Firstly, when the distance $d_{ij}$ between the ISR gluon and either hard prong is larger than that of the $b\bar{b}$ quark pair. In this situation, the gluon gets declustered first and, for the jet to be retained, one requires $k^2_t/M^2_j > y_{\text{cut}}$ where $k_t$ is the transverse momentum of the gluon with respect to the jet axis. Given the jet passes this Y-splitter cut, we also require that the mass window constraint is satisfied as for the plain jet mass. On the other hand, when $d_{ij}$ between the gluon and the $b\bar{b}$ pair is smaller than that between the $b$ and $\bar{b}$ prongs, the gluon is simply clustered into the jet. Here, like at zeroth order, one just declusters the two hard prongs and imposes the $y_{\text{cut}}$ condition. The gluon thus contaminates the jet and one again has to impose the mass window constraint.

One can argue that the configuration whereby the gluon has a larger $k_t$ distance and is not too energetic so as to comply with the mass window limits, is confined to a small corner of phase space that vanishes with $\delta M/p_T$. To see this most straightforwardly, one uses Eq. (4.34) to note that the gluon has a $k_t$ distance with respect to the $b\bar{b}$ pair (or equivalently in our soft large-angle approximation from either the $b$ or $\bar{b}$) that is approximately given by $d_{ij} = x^2 p_T^2 \theta^2 \simeq x^2 p_T^2$ where $x = \omega/p_T$. For convenience suppose that $z < 1/2$, the $k_t$ distance between the $b$ and $\bar{b}$ is $z^2 p_T^2 \theta_{b\bar{b}}^2$. Hence, the the gluon is declustered first when $x^2 p_T^2 > z^2 p_T^2 \theta_{b\bar{b}}^2$. The mass window condition for the contamination of an ISR gluon for $\theta^2 \sim 1 \gg \Delta$ gives $x < \frac{2M_{\delta M}}{p_T^2}$ (see Eq. (4.8)). These $k_t$ and mass window conditions are only simultaneously satisfied if $z/(1 - z) < (2\delta M/p_T)^2$, which corresponds to a negligibly small region of phase space and, given the uniform distribution in $z$, can be safely ignored. Hence, we are left with the situation that, up to small corrections, the ISR always gluon contaminates the jet and gives a result that is essentially like the plain jet mass. This implies considerable degradation in mass due to ISR, UE and, of course in the final analysis, pile-up.
Figure 4.10: A MC study of the impact of ISR on the signal efficiency for various taggers (\(y_{\text{cut}}, z_{\text{cut}}, f_{\text{cut}} = 0.1, R_{\text{trim}} = 0.3\)) as a function of the minimum jet transverse momentum. One can note the similarity of mMDT, pruning and trimming while Y-splitter and Y-pruning are different, with Y-splitter in particular virtually indistinguishable from the plain jet mass. Details of generation given in Fig. 4.1.

Let us compare Y-splitter with the other taggers and the plain jet mass using MC. In Fig. 4.10 we show the response of all taggers studied thus far to ISR, as a function of jet \(p_T\). One can immediately see that Y-splitter and the plain jet mass are essentially identical. One also notes that mMDT, trimming and pruning have a very similar behaviour to one another as expected from our analytical estimates. In contrast, the Y-pruning algorithm suffers at high \(p_T\) as already observed, however still remains far better than Y-splitter.

As far as FSR is concerned, one does not expect any significant corrections with Y-splitter. In a similar way to ISR, a soft FSR gluon will nearly always be clustered with the hard emitting partons, as a consequence of the soft limit and angular ordering, and so end up as part of the fat jet, thus not contributing to a loss in mass. Its effects will cancel against soft virtual corrections leaving us to study genuinely hard, non-collinear configurations that ought to have a relatively modest impact at the level of pure order \(\alpha_s\) corrections. Studies of the non-perturbative effects for Y-splitter have been carried out with MC for hadronisation and UE. As expected, the findings here are that the effects are comparable in size to the plain jet mass.

Thus in the final analysis it appears that Y-splitter may not be as useful as the other methods studied here, particularly Y-pruning, even though it shares a very similar suppression of the QCD background. While Y-splitter appears effective at identifying hard substructure and removing background, it is not effective in grooming away soft contamination, a feature inbuilt, to varying extents, in the other methods we have studied. This suggests using Y-splitter along with another effective grooming method may alleviate some of the issues we see with the signal. Therefore, in the next subsection we shall consider its combination with trimming, which we find has some noteworthy features and produces interesting results.
4.6 Y-splitter with trimming

Here we shall study the combination of Y-splitter with trimming, in view of the lack of any effective grooming element in Y-splitter, as mentioned previously. We do not have to necessarily choose trimming in this respect and it is possible to study a combination of Y-splitter with other methods with a grooming element such as mMDT and the recently introduced soft drop method [115]. Indeed it has been known for some time that combinations of substructure tools can often produce better results than the individual tools themselves [152] and thus one may hope to improve the performance of Y-splitter using a suitable complementary tool.

We first study the impact of applying trimming on signal jets that are tagged by Y-splitter. For the combination of Y-splitter with trimming we choose $f_{\text{cut}} = y_{\text{cut}} = 0.1$ and $R_{\text{trim}} = 0.3$. An MC analysis is shown in Fig. 4.11, which demonstrates that the use of trimming substantially ameliorates the loss of signal we saw with Y-splitter. It is evident from the same figure that, while Y-splitter with trimming still does not reach the signal efficiency of some other methods, the difference is much less pronounced than before. In fact one observes that the signal efficiency for Y-splitter with trimming bears a qualitative similarity to Y-pruning. The reason for this is that the use of trimming turns the plain mass like behaviour of Y-splitter into the behaviour for trimming, except for configurations that have been rejected by Y-splitter, on which subsequent trimming does not act. This corresponds to the Y-splitter rejection region whereby an ISR gluon has the largest $k_t$ distance but fails the $y_{\text{cut}}$ requirement. This kinematic configuration is reminiscent of that which resulted in the extra $\Delta\varepsilon_S$ correction for Y-pruning. In the present case, uncancelled virtual corrections integrated over the Y-splitter rejection region produce a term $\sim -C_F \frac{2\pi}{\sigma} R^2 \ln \frac{y_{\text{cut}}}{\Delta}$, which corrects the simple trimming result in Eq. (4.19). However, the overall efficiency remains considerably higher than the Y-splitter or plain mass result, due to elimination of the plain jet mass like logarithm.

![Figure 4.11: Signal efficiency for tagging hadronic $H$ jets using HERWIG++ 2.7.0 with underlying event and hadronisation as a function of a generator level cut $p_T$ on transverse jet momentum.](image-url)
Next, one should study the impact of using trimming in conjunction with Y-splitter on the QCD background. In this chapter, given our focus on signal jets, we shall not attempt to provide a detailed analytical study of the background case, which, together with the derivation of the basic Y-splitter formula Eq. (4.37), we shall carry out in Chapter 5. Here we shall instead confine ourselves to MC studies, the results of which are shown in Fig. 4.12. Once again we study the action of trimming on jets that have been tagged by Y-splitter and we choose $f_{\text{cut}} = y_{\text{cut}} = 0.1$ for both Y-splitter and trimming and $R_{\text{trim}} = 0.3$. In this plot, we remind the reader that $\rho = \frac{M_j^2}{p_T^2}$, i.e. the normalised jet mass as defined and used in the previous chapter.

One notes from Fig. 4.12 that Y-pruning, Y-splitter and Y-splitter+trimming all have a similar action on background jets and provide, for our choice of parameters, a significant suppression of background around the signal mass-peak $\sim 100$ GeV. These results are for quark backgrounds but similar results are obtained for gluon jets. It is important to note that the action of trimming for the chosen parameters, appears only to have an apparently subleading effect and hence the desirable property of Y-splitter, that of reducing background via a Sudakov suppression term (see Eq. (4.37)), is largely unaffected. Such findings are certainly worthy of an analytical follow-up for general choices of parameters, which we shall provide in the forthcoming chapter.

Given the improvement in signal efficiency that we have achieved using Y-splitter with trimming, and the fact that the backgrounds are comparably (and in fact apparently somewhat more) suppressed compared to Y-pruning in the mass region of interest, it is worth examining the signal significances (i.e. the ratio $\varepsilon_S/\sqrt{\varepsilon_B}$ of signal efficiency to the square root of background efficiency) that can be achieved with the various taggers, as a function of jet transverse momentum. These are shown in Fig. 4.13 for quark and gluon backgrounds. One observes that the Y-splitter with trimming method outperforms the other taggers discussed here, particularly at high $p_T$. In this figure,
4.7. OPTIMAL PARAMETER VALUES

Figure 4.13: Signal significance for tagging hadronic $H$ jets with quark (left panel) and gluon (right panel) backgrounds using Herwig++ 2.7.0 with underlying event and hadronisation as a function of a generator level cut $p_T$ on transverse jet momentum. We compare the signal significance for different algorithms to Y-splitter+trimming and find that the latter outperforms the others at high $p_T$. Here we have used $R_{\text{trim}} = 0.1$ for pure trimming because, in contrast to Y-splitter+trimming, we expect this value to be closer to optimal than $R_{\text{trim}} = 0.3$ at high transverse momenta (see Fig. 4.18).

One should note that we have used $R_{\text{trim}} = 0.3$ for Y-splitter with trimming; this represents a non-optimal choice at high $p_T$ for standalone trimming (see Fig. 4.18 later), hence we have chosen to present our results for trimming with $R_{\text{trim}} = 0.1$. A detailed study of optimal parameters for Y-splitter+trimming will be presented in the following chapter.

The results shown in Fig. 4.13 are for our standard process, $pp \rightarrow ZH$, but similar results are also obtained for $W$ tagging as shown in Fig. 4.14. Here we observe that Y-splitter with trimming now consistently outperforms the other taggers discussed over a range of $p_T$. This emerges from the different mass window of the $W$ boson ($64-96$ GeV) compared to the Higgs ($109-141$ GeV). In the window $M_W \pm 16$ GeV, the background mass distribution of Y-splitter+trimming is smaller relative to Y-pruning than the window around $M_H$ (see Fig. 4.12). Hence, we observe a greater signal significance for tagging $W$ rather than Higgs relative to the other taggers at large $p_T$.

### 4.7 Optimal parameter values

In this section we shall use analytical expressions to derive values of parameters that maximise the signal significance $\frac{\varepsilon_S}{\sqrt{\varepsilon_B}}$ for the different taggers. We do not expect the

---

8We have performed preliminary studies for other possible combinations such as Y-splitter+mMDT/pruning/soft drop. These all have a similar qualitative effect on both the background and signal jet mass distribution as Y-splitter+trimming. Hence, one observes a comparable gain in signal significance over Y-splitter for all of these combinations. However, we find that Y-splitter+trimming has the best signal significance for tagging $W$ bosons over background in the high $p_T$ limit.
values so derived to really be optimal in the sense that they will not take into account non-perturbative effects. Indeed we should emphasise that optimal parameter values have already been extracted using full MC studies for all methods considered in the original papers and also examined in subsequent studies such as in Ref. [95]. Analytical studies of optimal parameters have also been carried out by Rubin in Ref. [155] in the context of a filtering analysis, which we do not consider here.

Nevertheless, we can regard it as one of the tests of the robustness of these methods that the values derived here with analytical formulae as inputs should be reasonable approximations to what one obtains in complete MC studies. This is because one ideally wants to have substructure methods where statements about performance are largely independent of our detailed knowledge about non-perturbative corrections. We are also interested in examining to what extent general trends that emerge with analytics, such as the dependence of optimal parameters on $p_T$, are replicated in full MC studies. For the following studies we confine ourselves to quark backgrounds as we have no reason to believe that gluon backgrounds will differ significantly in terms of the conclusions we reach here.

Having observed in this chapter the relatively small radiative corrections, both for ISR and FSR, that emerge for signal processes over a broad range of parameter values, one feels encouraged in a first approximation to turn off these effects and treat the signal in a tree-level approximation, except for the case of trimming as we discuss below in more detail. In other words, we anticipate that the signal significance ought
to primarily be driven by the tree-level results for signal while for the background we shall use the resummed formulae first derived in Ref. [116]. For self-consistency, one should also verify that, for the optimal values one derives, the radiative corrections to signal efficiency can indeed be considered small relative to the tree level result.

### 4.7.1 mMDT

Let us follow the previously described procedure for the mMDT and extract the optimal value of $y_{\text{cut}}$. First, one needs to study the background mistag rate in the window $M_H - \delta M < M_j < M_H + \delta M$. Using $R = 1$ henceforth in this section, this corresponds to a range in the normalised jet mass $\rho$ of the form $\rho_H - \delta \rho < \rho < \rho_H + \delta \rho$, with $\delta \rho \approx 2M_H\delta M/p_T^2$. Using the tree-level signal efficiency, we have the following expression for the signal significance:

$$
\frac{\varepsilon_S}{\sqrt{\varepsilon_B}} = \sqrt{\frac{1 - 2y_{\text{cut}}}{\Sigma (y_{\text{cut}}, \rho_H + \delta \rho) - \Sigma (y_{\text{cut}}, \rho_H - \delta \rho)}},
$$

where $\rho_H = \frac{M_H}{p_T^2}$ and $\Sigma(y_{\text{cut}}, \rho)$ is the integrated mMDT background jet mass distribution calculated in Ref. [116]. Thus the quantity $\Sigma (y_{\text{cut}}, \rho_H + \delta \rho) - \Sigma (y_{\text{cut}}, \rho_H - \delta \rho)$ represents the integral of the background jet-mass distribution over the mass window corresponding to signal tagging with mMDT. We can find the value of $y_{\text{cut}}$ that maximises this signal significance $y_{\text{max}}$, by taking the derivative of Eq. (4.38) with respect to $y_{\text{cut}}$ and setting it to zero, which gives:

$$
-4\frac{1 - 2y_{\text{cut}}}{1 - 2y_{\text{max}}} = \frac{\Sigma' (y_{\text{cut}}, \rho_H + \delta \rho) - \Sigma' (y_{\text{cut}}, \rho_H - \delta \rho)}{\Sigma (y_{\text{cut}}, \rho_H + \delta \rho) - \Sigma (y_{\text{cut}}, \rho_H - \delta \rho)} |_{y_{\text{cut}} = y_{\text{max}}},
$$

where $\Sigma'$ denotes a derivative with respect to $y_{\text{cut}}$. Neglecting higher order corrections in $\delta \rho$, the optimal value for $y_{\text{cut}}$ satisfies

$$
-4\frac{1 - 2y_{\text{max}}}{1 - 2y_{\text{max}}} = \left. \frac{d}{dy_{\text{cut}}} \left( \frac{\partial \Sigma}{\partial \rho} \right) \right|_{\rho = \rho_H} \left. \frac{\partial \Sigma}{\partial \rho} \right|_{\rho = \rho_H} |_{y_{\text{cut}} = y_{\text{max}}},
$$

where $\Sigma(y_{\text{cut}}, \rho)$ is the fixed-coupling mMDT result for $\rho < y_{\text{cut}}$ derived in Ref. [116]. The fixed-coupling mMDT result for $\rho < y_{\text{cut}}$ reads:

$$
\Sigma(y_{\text{cut}}, \rho) = \exp \left[ -\frac{C_F\alpha_s}{\pi} \left( \ln \frac{1}{y_{\text{cut}}} + \ln \frac{1}{\rho} - \frac{3}{4} \ln \frac{1}{\rho} + \frac{1}{2} \ln^2 \frac{1}{y_{\text{cut}}} \right) \right].
$$
We can use this expression in Eq. (4.40), assuming that the optimal value lies in the region $\rho < y_{\text{cut}}$.[9] This gives us an implicit equation for optimal $y_{\text{cut}}$:

$$
\frac{-4y_{\text{max}}}{1 - 2y_{\text{max}}} = C_F \frac{\alpha_s}{\pi} \ln \frac{y_{\text{max}}}{\rho_H} + \frac{4}{3 + 4 \ln y_{\text{max}}}.
$$

(4.42)

One can numerically solve this equation, which contains the essential information about the optimal $y_{\text{cut}}$ and its dependence on $p_T$. The values we obtain for $y_{\text{max}}$ at $p_T = 1, 2, 3$ TeV with $\alpha_s = 0.1$ are approximately 0.124, 0.102 and 0.088 respectively. Whilst we have not included running coupling effects in the above derivation, one finds it is straightforward to do so. Using the full calculation of Ref. [116] for the background, i.e. including running coupling effects and a transition to the plain mass like behaviour for $\rho > y_{\text{cut}}$, we compute the analytical signal significance and plot in Fig. 4.15 as a function of $y_{\text{cut}}$.

From Fig. 4.15 we note firstly that the peak position of the analytical signal significance is approximately in agreement with the numbers we quoted in the last paragraph for the fixed-coupling calculation. A kink can be seen in the analytical result for $p_T = 1$ TeV, the origin of which is the transition from a single-logarithmic dependence on $\rho$, which is only valid at $\rho < y_{\text{cut}}$, to the usual double logarithmic result for the plain mass when $\rho > y_{\text{cut}}$. We have also shown, in the same figure, results from HERWIG++ 2.7.0 at both parton level and at full hadron level including UE. We find the HERWIG++ 2.7.0 results at parton level are in reasonable agreement with the simple analytical estimates we have made, for both the peak positions and the evolution of optimal $y_{\text{cut}}$ with $p_T$, though the values of the signal significance at each peak differ somewhat. It is noteworthy also that hadronisation and UE do not change the picture significantly at the $p_T$ values we have studied here. One other feature that emerges from both analytical and MC studies is that the peaks themselves are fairly broad such that choosing a slightly non-optimal $y_{\text{cut}}$ does not greatly impact the tagger performance. In Fig. 4.15 we have provided a direct comparison between optimal values from HERWIG++ 2.7.0 and the analytical estimates. We show the results for the range of $y_{\text{cut}}$ values (denoted by the pink shaded region) that correspond to a ±2% variation in the signal significance around the peak. For HERWIG++ 2.7.0 instead we overlay the equivalent range of $y_{\text{cut}}$ values as blue/black bars. We find a good degree of overlap within this tolerance band between full HERWIG++ 2.7.0 results and analytical estimates.

One can draw at least a couple of inferences from our observations. Firstly, as we have argued, radiative corrections to the signal for mMDT are clearly of minor significance to the tagger performance. The fact that the analytics are generally in

---

[9] This is reasonable at high $p_T$, because $\rho_H(1\text{TeV}) \approx 0.015$, i.e. much smaller than typical $y_{\text{cut}}$ values $\sim 0.1$ quoted in the literature.
Figure 4.15: mMDT analytical signal significance from tree level signal and resummed background as a function of $y_{\text{cut}}$ (top left) compared to HERWIG++ 2.7.0 \cite{75} at parton level (top right) and with hadronisation and MPI (bottom left). The signal process used is $pp \rightarrow ZH$ where we require the Higgs and Z to decay hadronically and leptonically respectively with quark backgrounds. We place a generator level cut on the Higgs transverse momentum $p_T$ of 1, 2 and 3 TeV. Jets are tagged around the Higgs mass with a mass window $\delta M = 16$ GeV. The bottom right panel shows the analytic optimal $y_{\text{cut}}$ values as a function of $p_T$ (red line) with a 2\% variation in signal significance about the peak (pink shaded area). We overlay the optimal results for $y_{\text{cut}}$ obtained using HERWIG++ 2.7.0 with hadronisation and underlying event at 1, 2 and 3 TeV, with an equivalent 2\% variation about the peak signal significance (blue bars) and at parton level (black bars).

Good agreement with HERWIG++ 2.7.0 points to the importance of the background contribution, in the context of the signal significance, and the success of analytical approaches in describing this background jet mass distribution \cite{116}. The fact that non-perturbative effects play an evidently minor role at the values of $p_T$ studied is also
reassuring from the point of view of a robust understanding of tagger performance.

We end with a caveat. If one considers lower $p_T$ values than the ones studied here, then one has to reconsider some of the previous arguments. In the low $p_T$ region e.g. 200 – 300 GeV, one would have a situation where we expect $\rho > y_{cut}$ to be satisfied, hence Eq. (4.42) for the background does not directly apply. Apart from this, perhaps more significantly, one can expect UE contamination to play a larger role due to the larger effective radius $\sim \frac{M_j}{p_T}$ whereby UE particles accumulate without being removed by the asymmetry cut. Here one ought to consider the use of mMDT with filtering and optimise the parameters of both methods together as in the original analysis [92].

### 4.7.2 Pruning and Y-pruning

Here we carry out a similar analysis for the case of pruning. The resummed expression for pruning, for QCD jets, is considerably more complicated compared to mMDT. The result essentially has two pieces that in Ref. [116] were dubbed the Y and I components. In this chapter we have already dealt with Y-pruning in some detail in the context of signal jets. For the background, as we have also discussed in a previous section, Y-pruning, for small jet masses $\rho < z_{cut}$, consists of a suppression of the leading order single-logarithmic result by a Sudakov-like form factor. At high $p_T$, this gives rise to a desirable suppression of the background in the signal region. The I-pruning contribution, on the other hand, starts at order $\alpha_s^2$ and is as singular as the plain jet mass, i.e. double logarithmic. For the sum of Y and I components, i.e. for pruning as a whole, one observes three distinct regions: 1) for $\rho > z_{cut}$ the behaviour is like the plain mass, 2) for $z_{cut}^2 < \rho < z_{cut}$ there is a single log behaviour as seen in the leading-order result (as for mMDT) and 3) for $\rho < z_{cut}^2$ we observe the I-pruning contribution start to become important. Specifically, this piece causes growth of the background and the appearance of a second Sudakov peak for quark jets and a shoulder-like structure for gluon jets (see Fig. 3.13).

We do not, for brevity, present here the resummed results for pruning for QCD jets, referring the reader instead to Ref. [116]. Here we simply plot the analytical signal significance for pruning, as for mMDT, with neglect of radiative corrections to the signal efficiency, but with the full resummed calculation for QCD background, which we take to be quark jets alone. The resulting signal significance is displayed in Fig. 4.16 along with MC results at parton and hadronic+UE level. One would expect the optimal $z_{cut}$ to lie in a region that corresponds to the mMDT-like region, i.e. such that $\rho_H > z_{cut}^2$. Choosing a larger $z_{cut}$ would push the signal window into the region where the background starts to grow due to onset of I-pruning and hence the signal significance would decrease. Hence, at 1 TeV one can expect an optimal value of $z_{cut}$ to be below $\sqrt{\rho_H} \approx 0.125$ while for 3 TeV one may expect a value closer to 0.04. These expectations are roughly consistent with what one observes from the analytical and
Figure 4.16: Pruning analytical signal significance from tree level signal and resummed background as a function of $z_{\text{cut}}$ compared to HERWIG++ 2.7.0 at parton level and with hadronisation and MPI. Details of generation given in Fig. 4.15.

MC results shown. Once again, we observe that the introduction of non-perturbative effects do not change the essential picture one obtains from analytics and have only a limited impact on the signal significance relative to parton level.

The pruning results have clear qualitative differences relative to mMDT. In particular, at high $p_T$, we have to be more precise about the optimal choice of $z_{\text{cut}}$ due to the somewhat narrower peak in the signal significance. We can compare, as for mMDT, analytical results to those from HERWIG++ 2.7.0, once again with a ±2% tolerance band, shown in the bottom right panel of Fig. 4.16. We observe that, within this small tolerance band, the results are compatible. Although, at higher $p_T$ with non-perturbative effects, perhaps less so compared to mMDT.

In the original pruning paper [109], the authors conclude that the optimal $z_{\text{cut}}$ value
for pruning is $\sim 0.1$ when using the C/A algorithm to cluster the initial jet, as we do here. This optimisation was performed at a relatively moderate transverse momenta ($100 – 500$ GeV for $W$ bosons) compared to the work in this chapter, however our results are consistent as we approach this region. For larger boosts, we observe that the optimal value choice for $z_{\text{cut}}$ tends to slightly smaller values ($z_{\text{cut}} \sim 0.075$).

In Fig. 4.17 we also present results for the signal significance of Y-pruning, again taking quark jets as background. Here we note firstly that the analytics are broadly in agreement with MC results for the shape of the signal significance as a function of $z_{\text{cut}}$. Secondly, the peaks are quite broad and so choosing a somewhat non-optimal value of $z_{\text{cut}}$ does not critically affect the significance. Furthermore, the optimal $z_{\text{cut}}$ does not depend strongly on $p_T$ and is virtually constant over the $p_T$ range studied.

---

**Figure 4.17:** Y-pruning analytical signal significance from tree level signal and resummed background as a function of $z_{\text{cut}}$ compared to HERWIG++ 2.7.0 at parton level and with hadronisation and MPI. Details of generation given in Fig. 4.15.
Lastly, within a ±2 % tolerance band, there is good agreement between partonic/full MC results and simple analytics on the optimal values of $z_{\text{cut}}$.

Hence, for mMDT, pruning and Y-pruning we find that, over the $p_T$ values we studied here, analytical results based on resummed calculations for QCD background and lowest order results for signals, with neglect of non-perturbative effects, capture the essential features of tagger performance, as reflected in the signal significance. An extension of our studies to lower $p_T$ values would be of interest in order to ascertain the further validity of the simple picture we have used for our analytical results and probe in more detail the role of radiative corrections to the signal and that of non-perturbative contributions. We shall next examine the more involved case of optimal parameters for trimming.

### 4.7.3 Trimming

Here we carry out a similar analysis for trimming, but one now has to optimise two parameters, $R_{\text{trim}}$ and $f_{\text{cut}}$. As performed for the mMDT/pruning analysis, we use the analytic resummed expression for QCD jets given in Ref. [116]. The result for this background jet mass distribution consists of a region with single logarithmic behaviour (equivalent in structure to mMDT) for $f_{\text{cut}}R^2_{\text{trim}} < \rho < f_{\text{cut}}$ that transitions into Sudakov double logarithmic growth in the background distribution for $\rho < f_{\text{cut}}R^2_{\text{trim}}$ (see Fig. 3.17). However, in contrast to mMDT and (Y-)pruning, FSR radiative corrections to the signal efficiency are crucial for optimisation.

If one naively uses the tree level result given in Eq. (4.16), it follows that the optimum value for $R_{\text{trim}}$ tends to zero. This is because one can ensure that signal mass window is within the single logarithmic background region by simply pushing the location of the double logarithmic transition ($\rho \simeq f_{\text{cut}}R^2_{\text{trim}}$) to small values of the QCD jet mass, thereby avoiding the Sudakov peak. However, as shown in this chapter, when taking the limit $R_{\text{trim}} \to 0$, one encounters large logarithmic corrections to the signal efficiency associated with final state radiation from the signal jet (see Eq. (4.25)). This puts a limit on how small one can reduce the trimming radius whilst maintaining adequate signal mass resolution. Hence, in our approximate analytics we now include FSR radiative corrections to the signal efficiency by integrating the the expression given in Eq. (4.24) over $z$ and adding this term to the Born level result Eq. (4.16). Including this radiative correction, along with the resummed QCD background, we can obtain analytical estimates for the signal significance.

In Fig. 4.18 we show a 2D density plot for the signal significance with trimming over a range of $R_{\text{trim}}$ and $f_{\text{cut}}$ values using Monte Carlo at parton level (top) and with full hadronisation and underlying event (bottom) with a transverse momentum cut at 2 and 3 TeV left and right respectively. We overlay a black analytical contour representing the region in which the analytical signal significance is no more than ±2% away from the
analytically derived peak value for $R_{\text{trim}}$ and $f_{\text{cut}}$. One can see that we have reasonable agreement between the simple analytical estimates and the HERWIG++ 2.7.0 results at parton level. However, when one includes non-perturbative effects, we observe that contamination from underlying event significantly reduces the signal significance in the large $R_{\text{trim}}$ region.

We can use our simple analytical estimates to comment on the optimal values we observe from MC. Firstly, for optimal values of $f_{\text{cut}}$ and $R_{\text{trim}}$, one would expect
4.7. OPTIMAL PARAMETER VALUES

Figure 4.19: Trimming signal significance for $R_{\text{trim}} \sim \sqrt{\Delta} \sim 0.06$ as a function of $f_{\text{cut}}$ generated at parton level and with hadronisation and MPI using HERWIG++. The signal mass window to reside in the single logarithmic mMDT-like region of the background, hence we anticipate that the optimal parameters satisfy the constraint $\Delta > f_{\text{cut}}R_{\text{trim}}^2$. This expectation is consistent with both the analytical contour (this constraint defines the top right edge) and MC results both at parton and full hadron level. This background driven effect is manifest as a suppression in signal significance when the product $f_{\text{cut}}R_{\text{trim}}^2$ becomes large (i.e. the top right of the contour plots). For example, at 3 TeV and fixed $f_{\text{cut}} = 0.1$, one would analytically expect an optimal value for $R_{\text{trim}} \lesssim 0.13$, whilst at $f_{\text{cut}} = 0.05$ one expects $R_{\text{trim}} \lesssim 0.19$. These numbers are in agreement with the analytical contour and MC results. Secondly, FSR corrections to the signal efficiency become significant in the region $R_{\text{trim}}^2 \ll \Delta$, hence one would expect the optimal trimming radius to reside in the region $R_{\text{trim}} \gtrsim \sqrt{\Delta}$. At 3 TeV this corresponds to $R_{\text{trim}} > 0.04$ and at 2 TeV corresponds to $R_{\text{trim}} > 0.06$. This is consistent with the analytical contour and MC, whereby we observe a reduction in signal significance in the limit $R_{\text{trim}} \ll \sqrt{\Delta}$.

We notice that, like mMDT and pruning, the signal significance is fairly insensitive to variations in $f_{\text{cut}}$ provided we choose $R_{\text{trim}}$ such that $\Delta/f_{\text{cut}} > R_{\text{trim}}$. However, the signal significance is subject to non-perturbative corrections that increase with $R_{\text{trim}}$, and consequently one should favour the small $R_{\text{trim}}$ limit of the analytical optimal contour region $R_{\text{trim}} \approx \sqrt{\Delta}$ to minimise both signal FSR and non-perturbative corrections to the signal significance. It is thus of interest to choose $R_{\text{trim}}^2 \sim \Delta$ and study the dependence of the signal significance on the choice of $f_{\text{cut}}$ as shown in Figs. 4.15, 4.16, 4.17 for mMDT, pruning and Y-pruning respectively. These results can be found in Fig. 4.19 for $R_{\text{trim}} = 0.06$, which is almost identical to $\Delta$ at 2 TeV and of order $\Delta$ for the other $p_T$ values. With this given choice of $R_{\text{trim}}$ (reminiscent of the pruning radius) it is
natural to compare the results to those for pruning reported in Fig. 4.16. One immediately notes that, even with a similar choice of radius, there are differences between the two techniques. While for pruning the optimal $z_{\text{cut}}$ decreases with increasing $p_T$, the optimal value for trimming stays more constant. However, the peak signal significance increases with $p_T$ in both cases. Additionally, for a given $p_T$, the tail behaviour as a function of $f_{\text{cut}}$ is also different, especially at larger $f_{\text{cut}}$. These differences originate from several sources: the difference in FSR corrections resulting from contrasting definitions of $f_{\text{cut}}$ and $z_{\text{cut}}$ (and their resulting $p_T$ dependence) and differences arising from the QCD background jet distribution with pruning and trimming (see Ref. [116]).

In order to better understand the role, for example, of FSR effects, in the above context, we again note that for pruning one can simply replace the signal efficiency by $1 - 2z_{\text{cut}}$ as we have done for our analytical studies of optimal parameter values. If one was to similarly use a signal efficiency $1 - 2f_{\text{cut}}$ to compute the signal significance for trimming, one would observe that the MC result with $R^2_{\text{trim}} \sim \Delta$ is closer to that for pruning and consequently optimal $f_{\text{cut}}$ values show a similar trend with $p_T$. However, it is evident that the $p_T$ dependent FSR corrections in trimming cannot be neglected, especially at low $p_T$ (see Fig. 4.6), and play an important role in pushing the optimal $f_{\text{cut}}$ to smaller values than would be obtained after turning off FSR effects. This is the main reason behind the observed relative insensitivity of optimal $f_{\text{cut}}$ values in Fig. 4.19 over the $p_T$ range studied.

### 4.8 Conclusions

In this chapter we have studied both perturbative radiative corrections and non-perturbative effects for the case of signal jets, specifically for boosted Higgs production with $H \rightarrow b\bar{b}$, after application of jet substructure taggers. For the former we have carried out relatively simple analytical calculations to assess the impact of ISR and FSR as well as to study its dependence on various parameters, such as a mass window of width $\delta M$ on either side of the signal mass, the fat jet $p_T$, the mass of the resonance $M_H$ and the parameters of the various taggers. To examine non-perturbative effects we have confined ourselves to MC studies.

Our study was motivated by relatively recent calculations dedicated to the case of QCD background jets and, in particular, work presented in the last chapter and Ref. [116]. There it was noted that, while taggers should, in principle, discriminate against jets from QCD background, the degree to which this happened and the impact on the background jet mass distribution was not always as desired. While taggers such as pruning, mMDT and trimming were essentially identical over a limited range in the normalised square jet mass $\rho = M_j^2/p_T^2$, significant differences in performance and behaviour were observed at small values of $\rho$, which, especially at high $p_T$, could correspond to background jet masses in the signal mass region of interest. Likewise,
taggers should, in principle, not significantly affect signal jets and retain them as far as possible. Additionally, most taggers have a grooming element (via the $f_{\text{cut}}/y_{\text{cut}}/z_{\text{cut}}$ criteria) that is responsible for the removal of jet contamination from ISR/UE, thereby helping in the reconstruction of sharper mass peaks. Here our aim was to carry out analytical and MC studies to investigate in detail the impact of taggers on signal jets, especially with regard to the interplay between tagger parameters and kinematic cuts such as jet $p_T$, masses and mass windows.

Our findings on the whole indicate that tagger performance is more robust for the case of signal jets than was apparent for QCD background. Most taggers are similar in their response to ISR and, in general, significantly ameliorate the loss of the signal efficiency observed for the plain jet mass cuts without application of substructure techniques. An exception to this situation was the case of Y-splitter, whereby the ISR and UE contamination resulted in a loss of signal efficiency identical to that seen for plain jets.

Likewise for FSR, the radiative losses that one sees are generally modest for a reasonably wide range of tagger parameters. Here an interesting question opens up about the potential role for fixed-order calculations in the context of jet substructure studies. This is because one observes an absence of genuine logarithmic enhancements for sensibly chosen tagger parameter values. The signal efficiency, for the taggers studied here, ought then to be better described by exact calculations that incorporate hard gluon radiation or by combinations of matrix element corrections and parton showers than by the soft/collinear emissions encoded in pure parton showers. In Appendix A.2 we carried out a comparison between an MC description of the signal efficiency and exact order $\alpha_s$ results for various taggers. We find that we can reasonably adjust parameters such as the size of the mass window $\delta M$ to obtain good agreement between the two descriptions. Such observations may also be useful beyond the immediate context of our work, in situations where differences in tagger performance could come from regions of phase space that are not under the control of a soft, eikonal approximation. In these situations, one would ideally want to combine resummed calculations, where necessary, with fixed-order calculations, i.e. carry out matched resummed calculations. A summary of the results presented in this chapter for the logarithmic structure of radiative corrections to the signal efficiency for each tagger are given in Table 4.1.

A development we have made here is the introduction of a combination of Y-splitter with trimming in an attempt to improve the response of Y-splitter to ISR/UE contamination. The motivation behind this combination originated from the observation that Y-splitter was very effective at suppressing QCD background in the signal region. The resulting improvement of signal efficiency from the extra grooming stage, coupled with the fact that the background suppression from Y-splitter remained essentially intact after the use of trimming, meant that the combination of Y-splitter with trimming
Table 4.1: A table summarising the logarithmic structure of radiative corrections to the signal efficiency for each tagger. For each tagger we show the coefficient of $-\frac{\alpha_s C_F}{\pi}$ for ISR and FSR results in the small $\Delta$ and $z_{\text{cut}}/y_{\text{cut}}/f_{\text{cut}}$ limit. We have defined $\epsilon = 2\delta M/M_H$ as in the main text. The coefficient $C_2$ for the trimming FSR logarithm is given in Eq. (4.26).

- **Tagger** | **ISR** | **FSR**

| Plain | $R^2 \ln \frac{R^2}{\Delta}$ | $\Delta \ln \Delta$ |
| Trimming | $R^2 \ln \frac{1}{f_{\text{cut}}}$ | $2 \ln \frac{\Delta}{R_{\text{trim}}} C_2(f_{\text{cut}}, \epsilon)$ |
| Pruning | $R^2 \ln \frac{1}{z_{\text{cut}}}$ | $\frac{2\pi}{\sqrt{3}} \ln \frac{z_{\text{cut}}}{\epsilon}$ |
| Y-pruning | $R^2 \ln \frac{z_{\text{cut}} R^2}{\Delta}$ | $\frac{2\pi}{\sqrt{3}} \ln \frac{z_{\text{cut}}}{\epsilon}$ |
| mMDT | $R^2 \ln \frac{1}{y_{\text{cut}}}$ | $0.646 \ln \frac{y_{\text{cut}}}{\epsilon}$ |
| Y-splitter | $R^2 \ln \frac{R^2}{\Delta}$ | $\mathcal{O}(y_{\text{cut}})$ |

Actually outperforms other taggers studied here, particularly at high $p_T$. Our observation is in keeping with the general idea that suitably chosen tagger combinations may prove to be superior discovery tools compared to the currently proposed individual methods [152]. In fact, it is now becoming increasingly common to use combinations of techniques such as N-subjettiness [97] with for instance mMDT in an effort to maximise tagger performance. (see e.g. Ref. [201]). There is also much effort aimed at better understanding tagger correlations [202], the analytical calculations in the next chapter for the case of Y-splitter with trimming will shed further light on some of these issues.

Lastly, we have carried out an analytical study of optimal parameter values for various taggers. Having observed modest radiative corrections to the signal we neglected these effects and found that analytical estimates, based on lowest order results for the signal and resummed calculations for QCD background, generally provide a good indicator of the dependence of signal significance on the tagger parameters. Despite not including non-perturbative effects, the analytical formulae give rise to optimal values that are fairly compatible with those produced by full MC studies. This is encouraging from the point of view of robustness of the various methods considered because a dependence of optimal values on MC features (hadronisation models or MC tunes) are potentially not ideal.

Finally, we note that for other methods such as N-subjettiness, for example, there will also be a suppression of signal jets due to the fact that such observables directly
restrict radiation from the signal prongs. Thus, in those cases, radiative corrections arising from soft/collinear emissions by signal prongs are highly significant as can be noted from Ref. [158]. We hope that our work taken together with studies of such observables will enable a more complete understanding of features of signal jets in the context of jet substructure studies and provide yet stronger foundations for future developments.
5.1 Introduction

In the last chapter, we observed that the action of the Y-splitter tagger [96] on QCD background jets was qualitatively similar to Y-pruning [14,116]. However, a signal jet tagged with the Y-splitter algorithm was subject to large losses in signal efficiency from contamination originating from ISR and underlying event. Using Monte Carlo techniques, we showed that the combination of Y-splitter with a subsequent trimming step significantly improved the signal tagging efficiency whilst leaving the background jet mass distribution largely intact. This resulted in a tagger combination that outperformed the other taggers studied here in signal to square root background efficiency, especially at high $p_T$. However, analytical understanding of the Y-splitter background jet mass distribution and the effect of trimming on this distribution was left largely unexplored. In particular, the MC observation of an apparently small difference between Y-splitter and Y-splitter with trimming background jet mass distributions certainly calls for a higher level of analytical understanding. Using information from such analytical expressions, one can gain greater insight into how particular combinations of tagging algorithms affect the background mistag rate. Specifically, one can obtain the parametric dependence of these observables on each tagger parameter and begin to find the optimal choices for these parameters for a given physics analysis.

Combining complimentary jet substructure algorithms may be an important step in designing more effective techniques to discriminate boosted heavy-resonance decay jets from the QCD background. For example, there have been several Monte Carlo studies suggesting that some combinations of substructure techniques can offer increased discrimination power when compared to each technique alone [148,152,202]. Due to the large number of substructure techniques available and hence the extensive number of possible combinations, it becomes of greater importance to gain an analytical
understanding of each individual tagger (as well as combinations), rather than relying purely on numerical techniques. Using the analytical insight gained in the previous chapters, one can now systematically identify candidate taggers for combination and in which order they should be applied to candidate boosted jets. Explicitly, in the case of Y-splitter with trimming, we identified that Y-splitter performed well for background rejection at high $p_T$ due to a Sudakov suppression term in Y-splitter (see Eq. (4.37)). In contrast, trimming performs relatively poorly for background rejection at high $p_T$ due to double logarithmic enhancement at small jet masses (see Eq. (3.79)). On the other hand, in the previous chapter we used analytical techniques to demonstrate that trimming effectively removes contamination from signal jets, whilst Y-splitter does not. Hence, one is led to the combination whereby one firstly applies the Y-splitter algorithm to remove a significant fraction of background jets and then one applies trimming to the remaining candidate signal jets to improve the mass resolution. One should note that the work in this chapter does not claim that Y-splitter with trimming is the optimal combination of jet substructure techniques, but rather provides an analytical understanding of the features that make this combination a promising candidate. One hopes that this work will form the first of many analytical studies of tagger combinations, paving the way for the design of better techniques that incorporate some of the desirable features previously found in separate taggers.

In this chapter, we continue the work of the previous chapter by providing a thorough exploration of the analytical structure of Y-splitter and we conduct one of the first full analytical calculations for the background jet mass distribution with a combination of substructure techniques: Y-splitter with trimming. In Section 5.2, we begin by calculating leading and next-to-leading order analytical expressions for the differential jet mass distribution of quark jets after application of the Y-splitter jet substructure algorithm. These results are compared against the fixed-order estimates using EVENT2 [66]. We then use Lund kinematic maps [78,203] to resum the leading large logarithmic contributions to the Y-splitter jet mass distribution to all orders in Subsection 5.2.4, which we compare to the Monte Carlo program HERWIG++ 2.7.0 [75]. We then consider a subleading “anomalous” contribution to the Y-splitter mass distribution; this arises when an emission that passes the Y-splitter cut is accompanied by a soft wide-angle emission (with lower $k_t$), which sets the jet mass. Whilst not necessary to capture the leading all-orders behaviour of Y-splitter, this term is crucial in order to understand the action of trimming on jets that pass Y-splitter. To this end, we calculate the leading and next-to-leading correction to the Y-splitter mass distribution after applying a subsequent trimming step in Section 5.3 and compare these results to EVENT2. In Subsection 5.3.2 we generate an all-orders result for Y-splitter with trimming using Lund diagrams and compare the result to Monte Carlo. After comparing these analytic resummed results to those for Y-pruning and standalone Y-splitter, along with a
Monte Carlo comparison, we briefly explore the impact of non-perturbative effects on the Y-splitter and Y-splitter+trimming mass distributions using HERWIG++ 2.7.0. Finally, in Subsection 5.3.4 we carry an approximate Monte Carlo optimisation of the Y-splitter with trimming parameters by maximising signal significance. We conclude by summarising this work and suggest possible studies that could further advance the understanding of jet substructure techniques and combinations.

5.2 Y-splitter

Y-splitter was initially introduced in the context of separating hadronic WW decay from QCD background [96]. In its initial form, a fixed cut was placed on the $k_t$ distance scale, $y_{\text{split}}$, at which a jet decomposes into two jets when reclustered using an exclusive $k_t$ algorithm. The expectation, in this context, is that $y_{\text{split}} \approx O(M^2_W)$ for signal jets whereas highly contaminated or QCD jets tend to have an asymmetric final merging step and therefore smaller $k_t$ scale (see Eq. (2.6) and associated discussion). However, further work using Y-splitter to identify boosted top decays by Thaler and Wang [104] showed that it is more advantageous to cut on the $k_t$ distance scale normalised to the jet mass\(^1\) (see also Ref. [81]):

$$y_{\text{cut}} \equiv \frac{y_{\text{split}}}{M^2_j}. \tag{5.1}$$

This will ensure that we do not simply reject all jets below a fixed $k_t$ scale but instead check for subjet energy symmetry in the last clustering, an important signature of a signal process. Additionally, one can now use the Y-splitter technique to reject background even if the mass of the heavy resonance is not known a priori. We will now analytically calculate the background differential jet mass distribution for this version of the Y-splitter algorithm.

5.2.1 Definition

The Y-splitter algorithm [96,104] has one external parameter, $y_{\text{cut}}$, which we need to optimise to separate signal from background. We start with a number of hard jets $j$ with radius $R$ defined using the $k_t$ algorithm [85] and apply Algorithm 9 to each jet\(^2\). For this study, we shall use an $e^+e^-$ adaptation of Algorithm 9 whereby we simply replace transverse momentum $p_{T,j}$ with the relevant energy $E_j$, rewrite $\Delta R^2_{j_1,j_2} \rightarrow 2(1 - \cos \theta_{j_1,j_2})$ where $\theta_{j_1,j_2}$ is the angle between $j_1$ and $j_2$ and $R^2 \rightarrow 2(1 - \cos R)$. When using the $k_t$ algorithm the final objects clustered have the largest $k_t$ distance,

\(^1\)In this reference, the authors instead cut on an energy sharing variable $z_{\text{cut}} \equiv \frac{y_{\text{split}}}{y_{\text{cut}} + M^2_j}$. This is identical to the definition in Eq. (5.1) under the replacement $y_{\text{cut}} \rightarrow \frac{z_{\text{cut}}}{1-z_{\text{cut}}}$.\(^2\)One can think of this algorithm as a non-recursive variant of the mass drop algorithm [92] with $\mu = \infty$ using the $k_t$ algorithm.
therefore $y_{\text{split}}^2$ will correspond to the scale of the final merging step.

**Algorithm 9: Y-splitter tagger.**

- **External parameters:** $y_{\text{cut}}$.
- **Input:** Jet, $j$ with radius $R$ and *constituents* defined with a $k_t$ clustering history.

1. Undo the last clustering step of $j$ to give subjets $j_1$ and $j_2$;
2. if 
   $$y_{\text{split}}^2 = \min \left( p_{T,j_1}^2, p_{T,j_2}^2 \right) \Delta R_{j_1,j_2}^2 > y_{\text{cut}} \times M_j^2$$
   then
3. Label $j$ the Y-splitter tagged jet;
4. else
5. Discard jet $j$;
- **Output:** Y-splitter tagged jet.

### 5.2.2 Leading order

In this section we will conduct a fixed-order, analytical study of the Y-splitter algorithm applied to quark jets. The calculations that follow are similar to those carried out in Chapter 3 and we will repeat some previously stated formulae for convenience.

In this subsection we will carry out a leading order (LO) calculation of the differential distribution $\frac{1}{\sigma} \frac{d\sigma}{d\rho}$, where $\rho = \frac{M_j^2}{E_j^2}$ is the squared jet mass normalised to the jet energy after application of the Y-splitter algorithm. We start by considering a back-to-back quark, anti-quark pair produced in an $e^+e^-$ collision.

We introduce a leading order correction to the system via the emission of a single soft gluon with momentum $k$, for ease we define the quark momentum to be aligned along the $z$-axis:

$$p = E_q \left( 1, 0, 0, 1 \right),$$

$$k = E_g \left( 1, 0, \sin \theta, \cos \theta \right),$$

(5.2)

where we have neglected recoil and $\theta$ is the angle between the quark and gluon, which have energies $E_q$ and $E_g$ respectively. The $k_t$ scale for this jet configuration to decompose into two subjets is

$$y_{\text{split}}^2 = \min(E_q^2, E_g^2) \Delta \theta,$$

(5.3)

where we have defined the quantity $\Delta \lambda \equiv 2(1 - \cos \lambda)$. Redefining the energies in
terms of fractions of the total jet energy $E_j$: $E_q = (1 - x)E_j$ and $E_g = xE_j$, the normalised square jet mass can be written:

$$\rho = \frac{M_j^2}{E_j} = x(1 - x)\Delta_\theta.$$  \hspace{1cm} (5.4)

The Y-splitter cut in Algorithm 9 is defined by the inequality $y_{\text{split}}^2 > y_{\text{cut}} \times M_j^2$, which, for brevity, we rewrite in terms of a $k_t$ scale normalised to the square mass of the jet, $y_{\text{scale}}^2$:

$$y_{\text{scale}}^2 = \frac{y_{\text{split}}^2}{M_j^2} > y_{\text{cut}}.$$  \hspace{1cm} (5.5)

For soft gluon emission $x < 1/2$, we use Eq. (5.3) to re-express the inequality in Eq. (5.5) as $x^2\Delta_\theta / \rho > y_{\text{cut}}$. Hence, using the expression Eq. (5.4), one finds that this condition is satisfied for a jet with a single emission in the region

$$\frac{1}{1 + y_{\text{cut}}} > x > \frac{y_{\text{cut}}}{1 + y_{\text{cut}}},$$  \hspace{1cm} (5.6)

which is identical to the LO (m)MDT constraint on emission energy, Eq. (3.8). We require that the gluon is clustered to the quark within a jet of radius $R$, which is satisfied in the region $\Delta_\theta < \Delta_R$. The leading order distribution is therefore given by

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} (Y-\text{splitter,LO}) = 2\frac{\alpha_s C_F}{\pi} \int \frac{d\cos \theta}{1 - \cos^2 \theta} \int_{\frac{1}{1+y_{\text{cut}}}}^{\frac{1}{1+y_{\text{cut}}}} dx p_{gq}(x) \delta \left( \rho - x\Delta_\theta \right) \Theta \left( \Delta_R - \Delta_\theta \right),$$  \hspace{1cm} (5.7)

where $p_{gq}(x)$ is the unregularised AP splitting function given in Eq. (1.58) and we have neglected longitudinal recoil in the definition of the jet mass, i.e. $\rho \approx x\Delta_\theta$. This only changes the final result at the level of non-singular terms. One can evaluate the integral in Eq. (5.7) to arrive at the result:

$$\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} (Y-\text{splitter,LO}) = \frac{\alpha_s C_F}{\pi} \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3}{4} \frac{1-y_{\text{cut}}}{1+y_{\text{cut}}}} \right) \Theta \left( \frac{y_{\text{cut}}\Delta_R}{1+y_{\text{cut}}} - \rho \right)$$

$$+ \frac{\alpha_s C_F}{\pi} \ln \left( \frac{4 \tan^2 \frac{R}{2} e^{-\frac{3}{4}}}{\rho \left(1+y_{\text{cut}}\right)} \right) \Theta \left( \rho - \frac{y_{\text{cut}}\Delta_R}{1+y_{\text{cut}}} \right).$$  \hspace{1cm} (5.9)

\footnote{One again finds it convenient to separate the integral into “collinear” and “wide angle” parts via the decomposition:

$$\frac{1}{1 - \cos^2 \theta} = \frac{1}{2} \left( \frac{\cos \theta}{1 + \cos \theta} + \frac{1}{1 + \cos \theta} \right),$$  \hspace{1cm} (5.8)

where the first term will capture the leading logarithms and the second will produce subleading corrections due to the absence of a collinear divergence as $\theta \to 0$.}
We see that for small masses, $\rho < \frac{y_{\text{cut}} \Delta R}{1 + y_{\text{cut}}}$, the distribution is single logarithmic in $\rho$; this indicates that Y-splitter has replaced the soft logarithm in the plain jet mass with a logarithm in $y_{\text{cut}}$. The fixed external parameter $y_{\text{cut}}$ can be chosen to be $O(1)$, thereby giving moderately sized logarithms in this region. For this reason, one may expect a wider region of fixed-order convergence in $\rho$ for the Y-splitter algorithm when compared to the plain jet mass. However, we first need to verify that this logarithmic replacement occurs at next-to-leading order and then at all orders before we can assess the full impact of the Y-splitter algorithm on the logarithmic structure.

For the large mass region, $\rho > \frac{y_{\text{cut}} \Delta R}{1 + y_{\text{cut}}}$, we get a double logarithmic expression that is the same as the plain jet mass up to $O(y_{\text{cut}})$ corrections. In this region of phase space, the emission is always sufficiently hard and wide angle enough to satisfy the Y-splitter condition and consequently the jet is never rejected.

One should notice that the expression in Eq. (5.9) is identical to the LO analytical structure of the (m)MDT algorithm, presented in Eq. (3.9). This result follows because Y-splitter uses the same asymmetry cut present in the mass drop tagger algorithm (see Algorithm 3), hence at LO the two algorithms have identical action.

We compare the analytical expression Eq. (5.9) with numerical results from EVENT2 for a range of $y_{\text{cut}}$ values generated at centre of mass energy 1 TeV by taking the average mass of the two hardest jets in each event. We also show the result after subtraction of our analytical calculation Eq. (5.9) in order to establish to what extent this expression captures the behaviour of the algorithm at LO. One can see in Fig. 5.1 that we have full control of the constant terms in the small $\rho$ limit, and that our LO calculation correctly reproduces the distribution at this order. The next step is to repeat this analysis at NLO, and see if the Y-splitter algorithm continues to remove logarithms associated with soft emission.

### 5.2.3 Next-to-leading order

In this section we will carry out a next-to-leading order calculation (NLO) of the normalised jet mass distribution after application of the Y-splitter algorithm. This is performed in a very similar way to the LO calculation, but we have the added complexity of two gluons emitted within the jet radius $R$.

We start by considering all the possible combinations of real and virtual diagrams for the emission of two independent soft gluons $k_1$ and $k_2$ from quark $p$ in the Abelian $C_F^2$ channel, shown in Fig. 3.5. We neglect the double virtual diagram (d), as this configuration does not contribute to the jet mass distribution. Defining the 4-momentum of $p$ as before in Eq. (5.2), one can write the momentum of each emitted gluon:

$$
k_1 = E_1 \left(1, 0, \sin \theta_1, \cos \theta_1\right),$$

$$
k_2 = E_2 \left(1, \sin \theta_2 \sin \phi_2, \sin \theta_2 \cos \phi_2, \cos \theta_2\right),$$

(5.10)
5.2. Y-SPLITTER

Figure 5.1: Comparison of LO EVENT2 and Eq. (5.9) in the region \( \rho < y_{\text{cut}} \Delta R / (1 + y_{\text{cut}}) \) for a range of \( y_{\text{cut}} \) values. We have plotted the coefficient of normalised differential jet mass distribution as a function of \( \ln \rho \) at LO. After subtracting the analytical result from the EVENT2 output, the result vanishes in the limit \( \rho \to 0 \).

where \( E_i \) and \( \theta_i \) are the energy and emission angle of gluon \( i \) respectively. We are free to choose the origin of the azimuthal axis and therefore we have set \( \phi_1 = 0 \) for emission \( k_1 \). Furthermore, we note that the jet mass is independent of the azimuthal angle (if one neglects contributions bilinear in the soft gluon momenta), therefore the leading logarithmic structure is unaffected by the value of \( \phi_2 \). Therefore, for simplicity, we henceforth simply choose \( \phi_2 = 0 \) for \( k_2 \). Again, we constrain emissions to be within the jet \( \Delta \theta_1, \Delta \theta_2 < \Delta R \), neglect recoil and self-clustering of gluons, which change the result below our level of accuracy.

In a similar fashion to the calculation performed at LO, let us firstly assign a normalised \( k_t \) scale, to each gluon \( i \) that we denote \( y^2_i \):

\[
y^2_i \equiv \frac{x_i^2 \Delta \theta_i}{\rho}.
\]  

(5.11)

In order to determine whether this jet passes the Y-splitter constraint, we also need the normalised scale at which the jet declusters from \( 1 \to 2 \) jets, \( y^2_{\text{scale}} \). For a jet clustering history defined with a \( k_t \) algorithm, this is simply provided by the first unclustering,
i.e. the emission with the largest $x^2 \Delta \theta_i$, which is equivalent to $y_{scale}^2 = \max(y_1^2, y_2^2)$. Hence, the Y-splitter algorithm will accept the whole jet if the largest normalised $k_t$ scale is greater than $y_{cut}$. Therefore, we get the following contribution from the double real (2R) diagram (a) in Fig. 3.5:

$$\Theta^{2R} = \left[ \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( y_1^2 - y_{cut} \right) + \Theta \left( y_2^2 - y_1^2 \right) \Theta \left( y_2^2 - y_{cut} \right) \right] \delta \left( \rho - x_1 \Delta \theta_1 - x_2 \Delta \theta_2 \right),$$

(5.12)

where the each term in the square brackets corresponds to different $k_t$ scale ordering of the two emissions. If the emission passes the Y-splitter cut, we accept the whole jet and the final normalised mass given by the sum of the contributions from both emissions, as shown by the delta function in Eq. (5.12).

Next, we consider the one-real one-virtual (1R1V) diagrams given by (b) and (c) in Fig. 3.5, the corresponding constraints are given by

$$\Theta^{1R1V} = -\Theta \left( y_1^2 - y_{cut} \right) \delta \left( \rho - x_1 \Delta \theta_1 \right) - \Theta \left( y_2^2 - y_{cut} \right) \delta \left( \rho - x_2 \Delta \theta_2 \right),$$

(5.13)

where the overall “−” sign in front of each term comes from the matrix element of the virtual gluon loop that is otherwise equal to the real emission in the eikonal limit (see Eq. (1.51)). We only check the real contribution for the Y-splitter condition, as it is the only emission that contributes to the jet mass distribution. The double virtual configuration can never contribute to the observable at this order in $\alpha_s$, so we do not consider it further.

The total contribution at NLO is therefore the sum of $\Theta^{2R}$ and $\Theta^{1R1V}$. In order to simplify this sum, one firstly makes the leading logarithmic approximation\footnote{For independent variables $\lambda_1$ and $\lambda_2$, we can write this as:}

$$\delta (\rho - \lambda_1 - \lambda_2) = \frac{\partial}{\partial \rho} [\Theta (\rho - \lambda_1) \Theta (\rho - \lambda_2)] \approx \frac{\partial}{\partial \rho} [\Theta (\rho - \lambda_1) \Theta (\rho - \lambda_2)].$$

(5.14)

in Eq. (5.12) to obtain:

$$\Theta^{2R} \approx 2 \times \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( y_1^2 - y_{cut} \right) \times \left[ \delta \left( \rho - x_1 \Delta \theta_1 \right) \Theta \left( \rho - x_2 \Delta \theta_2 \right) + \delta \left( \rho - x_2 \Delta \theta_2 \right) \Theta \left( \rho - x_1 \Delta \theta_1 \right) \right],$$

(5.15)

where we have inserted a factor of 2, to account for symmetry under the label interchange $1 \leftrightarrow 2$. Summing Eq. (5.15) and Eq. (5.13), we arrive at the result:

$$\Theta^{2R} + \Theta^{1R1V} \approx 2 \times \Theta \left( y_1^2 - y_{cut} \right) \delta \left( \rho - x_1 \Delta \theta_1 \right) \left[ \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( \rho - x_2 \Delta \theta_2 \right) - 1 \right]$$

$$+ 2 \times \Theta \left( y_2^2 - y_{cut} \right) \delta \left( \rho - x_2 \Delta \theta_2 \right) \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( \rho - x_1 \Delta \theta_1 \right),$$

(5.16)
where the first line is the leading contribution and represents the configuration whereby the emission with the largest $k_t$ measure, $k_1$, also dominates the jet mass. Additionally, the “−1” term in the square brackets denotes the corresponding virtual contribution. The second line contains the “anomalous” configuration in which $k_2$ dominates the jet mass, but has a smaller $k_t$ measure than emission $k_1$. In this case, $k_1$ must still satisfy the Y-splitter constraints but is required to have a smaller contribution to the jet mass relative to $k_2$. This combination of constraints means that this contribution is subleading; explicit computation yields terms $O\left(\alpha_s^2 L^2\right)$ in the integrated distribution, where $L \equiv \ln \frac{1}{\rho}$. In order to obtain the leading behaviour, it suffices to ignore this contribution, however we will return to terms arising from these “anomalous” configurations in Subsection 5.2.5.

Discarding the second line in Eq. (5.16) and performing the virtual cancellation in the first line we obtain the constraints:

$$\Theta^{2R} + \Theta^{1R1V} \simeq -2 \times \Theta \left(y_1^2 - y_{\text{cut}}\right) \delta \left(\rho - x_1 \Delta \theta_1\right) \Theta \left(x_2 \Delta \theta_2 - \rho\right) - 2 \times \Theta \left(y_1^2 - y_{\text{cut}}\right) \Theta \left(y_2^2 - y_1^2\right) \delta \left(\rho - x_1 \Delta \theta_1\right) \Theta \left(\rho - x_2 \Delta \theta_2\right),$$ (5.17)

where the first line is the leading contribution and the second line is subleading in the small $\rho$ limit; this is due to the constraint that $y_2^2 > y_1^2 > y_{\text{cut}}$, indeed explicit calculation of this term simply yields contributions of $O\left(\alpha_s^2 \ln^3 \frac{1}{y_{\text{cut}}} \times \tilde{L}\right)$ to the integrated distribution. Thus we have one integral to evaluate:

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} \left(Y\text{-splitter,NLO}\right) = -4 \left(\frac{\alpha_s C_F}{\pi}\right)^2 \int \frac{d\cos \theta_1}{1 - \cos^2 \theta_1} \frac{d\cos \theta_2}{1 - \cos^2 \theta_2} dx_1 p_{gq} (x_1) dx_2 p_{gq} (x_2) \times \frac{d\phi}{2\pi} \Theta \left(\Delta R - \Delta \theta_1\right) \Theta \left(\Delta R - \Delta \theta_2\right) \times \Theta \left(y_1^2 - y_{\text{cut}}\right) \delta \left(\rho - x_1 \Delta \theta_1\right) \Theta \left(x_2 \Delta \theta_2 - \rho\right).$$ (5.18)

where the factor of 2 in Eq. (5.17) has been cancelled against a combinatoric factor $1/2!$ associated with emission of two identical gluons. The evaluation of Eq. (5.18) is straightforward, when one observes that this integral can be factorised into separate contributions from each gluon. Hence, performing the trivial integration over $\phi$ and separating the integral into a factorised form gives:

$$\frac{1}{\sigma} \frac{d\sigma}{d\rho} \left(Y\text{-splitter,NLO}\right) = -4 \left(\frac{\alpha_s C_F}{\pi}\right)^2 \times \int \frac{d\cos \theta_1}{1 - \cos^2 \theta_1} dx_1 p_{gq} (x_1) \Theta \left(\Delta R - \Delta \theta_1\right) \Theta \left(y_1^2 - y_{\text{cut}}\right) \delta \left(\rho - x_1 \Delta \theta_1\right) \times \int \frac{d\cos \theta_2}{1 - \cos^2 \theta_2} dx_2 p_{gq} (x_2) \Theta \left(\Delta R - \Delta \theta_2\right) \Theta \left(x_2 \Delta \theta_2 - \rho\right).$$ (5.19)
We see that the second line is identical to the LO integral in Eq. (5.7). Hence, we can write a leading expression for the NLO distribution by evaluating the third line for small $\rho$ and multiplying by the LO result Eq. (5.9) in the small $\rho$ limit:

$$\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} \left( \text{Y-splitter, NLO} \right) = -\left( \frac{\alpha_s C_F}{\pi} \right)^2 \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3(1-y_{\text{cut}})}{4(1+y_{\text{cut}})}} \right) \frac{1}{\ln^2 \frac{1}{\rho}} \Theta \left( \frac{y_{\text{cut}} \Delta R}{1 + y_{\text{cut}}} - \rho \right) + \mathcal{O} \left( \alpha_s^2 \ln \frac{1}{\rho} \right).$$

(5.20)

One observes that the integrated distribution in the small $\rho$ region at NLO is of the form $\mathcal{O} \left( \alpha_s^2 L^3 \right)$, in contrast to the leading-order single-logarithmic distribution $\mathcal{O} \left( \alpha_s L \right)$ given in Eq. (5.9). This indicates that, starting at NLO, the Y-splitter algorithm does not remove any additional soft divergences from the jet mass distribution. We anticipate that this behaviour persists to all orders in $\alpha_s$, because Y-splitter only checks the single emission that declusters first, hence the leading term in the integrated distribution at order $n$ will be of the form $\alpha_s^n L^{2n-1}$. Whilst this perturbative expansion
is more convergent than the double logarithmic plain jet mass, it is less convergent than the single logarithmic distribution one sees at leading order. As we will see later, after considering the resummed expression, this feature is phenomenologically favourable in terms of suppressing QCD background jets in the small $\rho$ limit.

In Fig. 5.2 we check this analytical expression against the fixed-order code Event2 at NLO in the $C_F C_A$ channel for a range of $y_{\text{cut}}$ values. The subtraction of Eq. (5.20) from the numerical result leaves a straight line in the small $\rho$ limit, which indicates that we successfully describe the differential distribution at NLO up to single logarithmic corrections $O(\alpha_s^2 \ln \rho)$.

Finally, we end by noting that the leading contribution to Y-splitter in the $C_F C_A$ and $C_F n_f$ channels will arise purely from the 1-loop running coupling corrections to the LO result. Analogous to the mMDT result, we expect that the cut on soft emissions will mean that non-global contributions will not be enhanced by logarithms in $\rho$ at this order and will instead contribute logarithms in $y_{\text{cut}}$ (see Eq. (3.35)). Given that the LO result for Y-splitter is identical to (m)MDT, we re-use the running coupling result
for mMDT from Eq. (3.28):
\[
\frac{\rho \, d\sigma}{\sigma \, d\rho}^{(Y\text{-splitter, NLO rc})} = C_F b_0 \left( \frac{\alpha_s}{\pi} \right)^2 \ln \left( \frac{1}{y_{\text{cut}}} e^{-\frac{3}{4} \left( \frac{1-y_{\text{cut}}}{1+y_{\text{cut}}} \right)} \right) \ln \frac{1}{\rho},
\]
(5.21)
where \( b_0 = \frac{1}{12} (11C_A - 4T_R n_f) \) and compare it to EVENT2 in Fig. 5.3. Here one observes that the leading behaviour of Y-splitter in the \( C_F C_A \) and \( C_F n_f \) channels is single logarithmic. The difference between this result and the corresponding analytical expression in Eq. (5.21) is constant for small \( \rho \), which corresponds to leftover terms \( \mathcal{O}(\alpha_s^2) \) in the differential distribution. This confirms that non-global logarithms do not contribute single logarithmic terms in \( \rho \) to the distribution at NLO.

Given that we now have analytic LO and NLO expressions for the Y-splitter normalised jet mass distribution, we can approximate the structure of the tagger to all orders. We achieve this via resummation of large logarithms into exponential form, as discussed extensively in the literature (e.g. Refs. [78, 204, 205]) and Section 2.4.

### 5.2.4 Resummation

In this section we will consider the effect of an arbitrary number of eikonal emissions on the normalised jet mass distribution after applying the Y-splitter algorithm. In Chapter 3 we presented resummed results for each tagger taken from Ref. [116], however, in this paper, the authors did not consider the action of the Y-splitter algorithm on background jets. Hence, we now perform a full resummation of the Y-splitter jet mass distribution with running coupling and hard collinear effects. We assume that each emission is emitted independently and ignore subsequent splitting of those emissions, apart from those accounted for by inclusion of running coupling effects. This is an approximation that reflects the essential result of independent strongly-ordered emissions in either angle or \( p_T \) and is sufficient for leading log (LL) accuracy; additionally this approximation includes a subset of the next-to-leading terms in the exponent via the running of \( \alpha_s \) (see discussion in Section 2.4). Notably, we neglect non-global terms associated with the tagger, which have a subleading effect on the differential jet mass distribution. Additionally, will be working exclusively in the collinear, \( R \ll 1 \), and small \( y_{\text{cut}} \) region, which will capture the leading contributions to the resummed distribution.

We firstly make the leading assumption that the emission that sets the jet mass, also has the highest \( k_t \) scale and therefore the Y-splitter constraint applies only to that emission. One should note that this assumption is inconsistent with configurations whereby an “anomalous” wide-angle and soft emission sets the jet mass while a second emission with a larger \( k_t \) scale enables the jet to pass the Y-splitter check. As mentioned in the previous subsection, this configuration contributes an \( \mathcal{O}(\alpha_s^2 L^2) \) term to the integrated distribution starting at NLO, which is below our current level of accuracy.
5.2. Y-SPLITTER

Figure 5.4: Lund Diagram for the Y-splitter algorithm applied to a jet with given normalised mass $\rho$. We define the normalised $k_t$ scale $y_{\text{scale}}^2 = y_{\text{cut}}$ as a horizontal line and write that the emission that dominates the jet mass lies on the contour of constant $\rho$. For the jet to be accepted, this emission must have a normalised $k_t$ scale greater than $y_{\text{cut}}$, (this is equivalent here to $x > y_{\text{cut}}$ because the emissions with the largest $k_t$ scale also sets the jet mass). Further emissions are not checked by Y-splitter, so we need to veto on the shaded area to give the correct Sudakov suppression term. Emissions in the unshaded region have a subleading impact on the jet mass, so we ignore these. Note that the x-axis begins at $\eta \approx -\ln R/2$, which is the minimum pseudorapidity with respect to the jet axis for an emission to be within a jet of radius $R$.

(see Subsection 5.2.5 for explicit calculation). However, this configuration is crucial in understanding the differences in background distribution between Y-splitter and the combination of Y-splitter with trimming, as we show shortly in Subsection 5.3.2.

The resummed result, in the small $y_{\text{cut}}$ limit, is illustrated by the Lund diagram in Fig. 5.4 (See Appendix. C for more details on Lund diagrams). In this figure, we show Y-splitter applied to a jet with normalised square mass $\rho$ (thick red line) such that the emission that sets the jet mass has the largest $k_t$ scale and satisfies the constraint $y_{\text{scale}}^2 > y_{\text{cut}}$. We must veto on all real emissions $i$ that have an individual contribution to normalised jet mass $\rho_i > \rho$; this veto is represented by the blue shaded region of phase space. We can therefore write a resummed expression for the differential jet
mass distribution as an exponential of the vetoed region times an integral over the
phase space of having an emission with an energy fraction that sets the jet mass \( \rho \) (see Eq. (C.4)):

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y\text{-splitter})} = e^{-D(\rho)} \int_{\text{max}(y_{\text{cut}},\rho/R^2)}^{1} dx p_{gq}(x) \frac{\alpha_s(xp_T^2) C_F}{\pi},
\]

(5.22)

where the constraint \( y_{\text{scale}}^2 > y_{\text{cut}} \) is equivalent to a limit on the energy of the emission
that sets the mass, i.e. \( x > y_{\text{cut}} \). Additionally, we have defined the double logarithmic
function

\[
D(\rho) = \int_{\rho}^{R^2} \frac{d\rho'}{\rho'} \int_{\rho'/R^2}^{1} dx p_{gq}(x) \frac{\alpha_s(xp_T^2) C_F}{\pi},
\]

(5.23)

where we have evaluated \( \alpha_s \) at the \( k_t \) of the emission with respect to the jet axis in
Eqs. (5.22,5.23) and used the full \( q \rightarrow gq \) splitting function. This double logarithmic
function is identical to the argument of the plain jet mass exponent and corresponds
to the sum of virtual emissions within the shaded veto region of the Lund diagram.

We note that, unlike plain jet mass, the expression Eq. (5.22) does not have a simple
exponentiated structure, insofar that a single emission (the one with the highest \( k_t \)
scale) has different phase space constraints compared to the remaining real emissions.
The resummed result for Y-splitter presented in Eq. (5.22) resums all terms \( \alpha^n_s L^{2n-1} \)
and some terms \( \alpha^n_s L^{2n-2} \) in the expansion of the integrated distribution up to finite
corrections proportional to \( y_{\text{cut}} \).

It is interesting to compare this expression to the analytical, resummed integrated
distribution for Y-pruning presented in Ref. [116], which has identical leading \( \alpha^n_s L^{2n-1} \)
structure in the region \( \rho < y_{\text{cut}}^2 R^2 \) when \( \alpha_s \ln \frac{1}{y_{\text{cut}}} \ln \frac{R^2}{\rho} \ll 1 \) but differs at a subleading
level. Numerically, these extra subleading terms present in Y-pruning, when evaluated,
appear to act by enhancing the background mass distribution relative to Y-splitter at
small \( \rho \). This suggests that Y-splitter has the potential to offer better background
rejection of QCD jets in some regions of phase space (we later show this explicitly
using Monte Carlo simulation in Fig. 5.11).

It is informative to evaluate Eq. (5.22) in the region \( \rho < y_{\text{cut}} R^2 \) for fixed-coupling
in the soft limit:

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y\text{-splitter})} \simeq \frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} \exp \left[ -\frac{\alpha_s C_F}{\pi} \frac{1}{2} \ln^2 \frac{R^2}{\rho} \right],
\]

(5.24)

which, to leading logarithmic accuracy, is the same result quoted in the previous chapter
Eq. (4.37). One can expand Eq. (5.24) for small values of the exponent to give the
same leading LO and NLO terms we previously calculated in Eq. (5.9) and Eq. (5.20).
for small $\rho$, up to corrections $O(y_{\text{cut}})$. For $\rho > y_{\text{cut}} R^2$ we just obtain the plain jet mass result.

We are now able to compare the resummed expression for the differential normalised jet mass Eq. (5.22) against MC results. At this level of accuracy, we expect our analytical expression to describe the shape and transition points of the differential distribution. Hence, in Fig. 5.5 we check the shape of the analytical distribution against the corresponding parton level Monte Carlo results for quark jets with minimum $p_T = 3$ TeV. The left panel shows the Y-splitter mass distribution from HERWIG++ 2.7.0 at parton level without underlying event. A minimum $p_T$ cut on generation of the hard process $qq \rightarrow qq$ was made at 3 TeV for 14 TeV pp collisions. We take the mass of the two hardest jets found using the C/A algorithm with $R = 1.0$ and apply the Y-splitter algorithm with several values of $y_{\text{cut}}$. In the above plots, we indicate the expected transition points with dashed lines at $\rho = y_{\text{cut}} R^2$, at this point the distribution changes from the plain jet mass to the next-to-double distribution characteristic of the Y-splitter algorithm.

---

5When evaluating the running coupling in the infrared region at 1-loop, we can encounter the non-perturbative region as $\rho$ or $x$ becomes small. In order to deal with this, we introduce a “freezing scale” that fixes the value of $\alpha_s$ below a non-perturbative scale $\mu_{\text{NP}}$. This corresponds to the replacement

$$\alpha_s (k_t^2) = \alpha_s^{1\text{-loop}} (k_t^2) \Theta (k_t^2 - \mu_{\text{NP}}^2) + \alpha_s^{1\text{-loop}} (\mu_{\text{NP}}^2) \Theta (\mu_{\text{NP}}^2 - k_t^2)$$

where $k_t^2 = p x p_T^2$, $\mu_{\text{NP}} = 1$ GeV and we have used the usual 1-loop running coupling expression:

$$\alpha_s^{1\text{-loop}} (\lambda^2) = \frac{\alpha_s (M_Z^2)}{1 + \beta_0 \alpha_s (M_Z^2) \ln \frac{\lambda^2}{M_Z^2}}.$$
differential jet mass produced by Monte Carlo, including the relative dependence of the shape and transition points on the parameter $y_{\text{cut}}$. We note that for small masses, the distribution exhibits a Sudakov suppression of background as shown in the analytic result Eq. (5.22). One also notices that the analytically derived transition point to the plain double logarithmic distribution at $\rho \simeq y_{\text{cut}} R^2$, indicated by the dashed lines on each plot, is in good agreement with MC.

5.2.5 Y-splitter: The anomalous configuration

When we calculated the NLO and resummed expressions for Y-splitter, we discounted the possibility that a soft and wide-angle emission sets the jet mass whilst another hard emission with a larger $k_t$ scale causes the jet to pass the Y-splitter condition. As mentioned previously, this is a subleading contribution, however in order to compare Y-splitter and Y-splitter with trimming later on, we must briefly consider this effect. We can again use a Lund diagram, shown in Fig. 5.6 to write down this correction. We start with a soft, wide-angle emission that dominates the jet mass but has an energy $x < y_{\text{cut}}$ (thick red line), such that it would fail the Y-splitter cut if checked. However, if this jet contains a secondary emission, $k_2$, that does not dominate the jet mass, but has a sufficiently large $k_t$ scale to satisfy the Y-splitter condition as defined by the mass contribution of the first emission (thick orange line), the jet will be accepted. These constraints define the following integral:

$$
\frac{\rho \, d\sigma}{\sigma \, d\rho}^{(Y-\text{splitter,Anomalous})} = e^{-D(\rho)} \left[ \int_{y_{\text{cut}}}^{\rho} \frac{dx}{x} \alpha_s \left( \frac{\rho x p_T^2}{\pi} \right) C_F \right] \times \left[ \int_{y_{\text{cut}} \rho / \rho_2}^{\rho} \frac{d\rho_2}{\rho_2} \int_{y_{\text{cut}} \rho / \rho_2}^{1} \frac{dx_2 p_{gq} (x_2)}{\pi} \alpha_s \left( \frac{\rho_2 x_2 p_T^2}{\pi} \right) C_F \right],
$$

(5.25)

where the second line denotes an integral over the secondary emission that passes the Y-splitter condition, whilst the jet mass is defined by the emission on the first line. One can evaluate this in the fixed-coupling approximation for $\rho < y_{\text{cut}} R^2$ to give the expression:

$$
\frac{\rho \, d\sigma}{\sigma \, d\rho}^{(Y-\text{splitter,Anomalous})} = \left[ \left( \frac{\alpha_s C_F}{\pi} \right)^2 \frac{1}{2 \ln^2} \frac{1}{y_{\text{cut}}} \ln \frac{y_{\text{cut}} R^2}{\rho} \right] e^{-D(\rho)},
$$

(5.26)

which corresponds to a single logarithmic correction starting at $\alpha_s^2$ to the integrated distribution for Y-splitter of the form $O_2$. Combining the resummed result

---

6 Here we neglect the configuration whereby the emission that sets the jet mass would pass the Y-splitter condition if it had the highest $k_t$ scale, i.e. $x > y_{\text{cut}}$, but is accompanied by a larger $k_t$ emission. This configuration is subleading and, if included, would contribute terms of order $\alpha_s^2 \ln^3 \frac{1}{y_{\text{cut}}}$ to the differential distribution starting at NLO.

204
Figure 5.6: Lund Diagram for the anomalous contribution to the Y-splitter algorithm applied to a jet with given normalised mass $\rho$. We define $y_{\text{scale}}^2 = y_{\text{cut}}$ as a horizontal line and write that the emission that dominates the jet mass lies on the contour of constant $\rho$. The jet mass is defined by a soft ($x < y_{\text{cut}}$) emission but a secondary emission with larger $k_t$ scale passes the Y-splitter condition. For the jet to be accepted, the secondary emission must have an energy $x_2 > y_{\text{cut}} \rho / \rho_2$ and a mass contribution $\rho > \rho_2 > y_{\text{cut}} \rho$.

Eq. (5.22) with the integral in Eq. (5.25), one now has an expression that resums all terms $\alpha_s^n L^{2n-1}$ and $\alpha_s^n L^{2n-2}$ terms in the expansion of the integrated distribution. For fixed-coupling and $\rho < y_{\text{cut}} R^2$ we get the result:

$$
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y-\text{splitter})} = \left[ \frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} + \left( \frac{\alpha_s C_F}{\pi} \right)^2 \frac{1}{2} \ln^2 \frac{1}{y_{\text{cut}}} \ln \frac{y_{\text{cut}} R^2}{\rho} \right] e^{-D(\rho)}. \quad (5.27)
$$

We reiterate that the Y-splitter distribution remains identical to the plain result in the region $\rho > y_{\text{cut}} R^2$.

Now that we understand the leading analytical behaviour of Y-splitter, we can begin to associate the double logarithmic Sudakov suppression term $e^{-D(\rho)}$ with the small background mistag rate at large $p_T$. However, results in the previous chapter suggested that Y-splitter performs poorly in signal efficiency under non-perturbative corrections (see Fig. 4.10). Consequently, Y-splitter was combined with trimming in
order to remove soft contamination and improve the signal efficiency. In the next section we consider the combination of Y-splitter with trimming, specifically the effect trimming has on the analytical structure of Y-splitter when applied to background jets.

### 5.3 Y-splitter with trimming

In the last chapter we identified a potential flaw in the signal resolution of Y-splitter under UE and ISR contamination, which was significantly improved by introducing a subsequent trimming step. It now becomes of importance to examine how the analytic structure of Y-splitter with trimming applied to QCD background differs from Y-splitter alone. The order in which one applies these techniques is, in general, important; this is because some algorithms modify the substructure of the jet upon which the next algorithm acts. To take a specific example, applying trimming first to a QCD jet will tend to remove energy asymmetric contributions to the jet. Hence, passing these trimmed jets to Y-splitter will degrade the effectiveness of the Y-splitter asymmetry check to further reject background jets. In contrast, Y-splitter does not modify the substructure of a jet that is tagged, leaving trimming to act on all its constituents.

In Fig. 5.7 we present MC results for the differential jet mass distribution for Y-splitter followed by trimming and trimming followed by Y-splitter for $f_{\text{cut}} = y_{\text{cut}}$. One immediately notices that the order in which one applies each algorithm is important for background rejection. For Y-splitter followed by trimming, the distribution is similar to that observed in Fig. 5.5 for standalone Y-splitter. In contrast, by applying Y-splitter to trimmed jets, one obtains a distribution that has inherited some of the features of standalone trimming, seen in Fig. 3.17. A basic explanation for the latter result is that the trimming step removes contributions to the jet that would fail the Y-splitter step, provided they are emitted at an angle greater than $R_{\text{trim}}$. Hence, for jets with masses in the region $\rho > f_{\text{cut}}\Delta R_{\text{trim}}$ the Y-splitter constraints are always satisfied and the distribution is unmodified from standalone trimming. However, in the region $\rho < f_{\text{cut}}\Delta R_{\text{trim}}$ the mass of the jet is dominated by the mass of a single trimming subjet, i.e. the soft emission that sets the mass is emitted at angle less than $R_{\text{trim}}$. These configurations can fail the Y-splitter constraints, hence one observes a suppression in this region relative to standalone trimming.

In this section, we always apply Y-splitter as an initial tagging step before employing the trimming procedure because Fig. 5.7 indicates that it largely retains the favourable background rejection rates observed for standalone Y-splitter. Additionally, we set $f_{\text{cut}} = y_{\text{cut}}$ in the small $y_{\text{cut}}$ limit. We first note that the leading order modification to Y-splitter due to trimming is trivial. At this order, we require that the single emission

---

7One is free to choose different values of $y_{\text{cut}}$ or $f_{\text{cut}}$ for each tagger. However, by doing so, the analytic expressions gain further structure that may be important for the background distribution. We leave this for future study.
5.3. **Y-SPLITTER WITH TRIMMING**

Figure 5.7: Monte Carlo results for the differential jet mass distribution using Y-splitter followed by trimming compared to trimming followed by Y-splitter. We have used the parameters $f_{\text{cut}} = y_{\text{cut}} = 0.1$, $R_{\text{trim}} = 0.3$. Details of generation are given in Fig. 5.5.

has energy $y_{\text{scale}}^2 > y_{\text{cut}}$ to satisfy the Y-splitter constraint. This reduces to a constraint on energy $x > y_{\text{cut}}$, which always satisfies the trimming conditions; consequently, the jet is left unmodified by the trimming step. Therefore, at leading order we get the same result as Y-splitter given in Eq. (5.9) for Y-splitter with trimming.

### 5.3.1 Next-to-leading order

At next-to-leading order, given that a jet passes the Y-splitter step, the trimming algorithm proceeds by removing soft $x < y_{\text{cut}}$ and wide angle $\theta > R_{\text{trim}}$ emissions. This means that only the emission with the highest $k_t$ scale, that satisfied the Y-splitter constraints, is guaranteed to pass this grooming step, which, as we have seen, is not necessarily the emission that sets the jet mass. Hence, one now has the possibility that trimming can change the mass of the jet by removing a soft, wide-angle emission that dominates the jet mass.

We start by reiterating that in the leading configuration, whereby the emission that defines the jet mass also has the highest $k_t$ scale, the trimming algorithm does not, at leading logarithmic level, change the final jet mass. We therefore obtain the same $O(\alpha_s^2 L^3)$ leading integrated NLO expression as standalone Y-splitter given in Eq. (5.20). However, we must consider the anomalous contribution to Y-splitter, whereby an emission with $y_{\text{scale}}^2 < y_{\text{cut}}$ sets the jet mass but is accompanied by another emission with a $k_t$ scale that passes the Y-splitter cut. The action of the trimming algorithm on the jet will result in the removal of the soft, wide-angle emission that dominates the jet mass, given that it is emitted at an angle greater than $R_{\text{trim}}$. This will alter the final mass of the jet, and has the potential to introduce a modification, relative to Y-splitter, to the logarithmic structure starting at NLO.

In this subsection, we explicitly compute the difference between Y-splitter and Y-splitter with trimming at NLO. We start by comparing the Y-splitter constraints given
by the sum of Eq. (5.12) and Eq. (5.13):

$$\Theta^{\text{Y-split}} = \left[ \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( y_1^2 - y_{\text{cut}} \right) + \Theta \left( y_2^2 - y_1^2 \right) \Theta \left( y_2^2 - y_{\text{cut}} \right) \right] \delta \left( \rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2} \right)$$

$$\Theta \left( y_1^2 - y_{\text{cut}} \right) \delta \left( \rho - x_1 \Delta_{\theta_1} \right) - \Theta \left( y_2^2 - y_{\text{cut}} \right) \delta \left( \rho - x_2 \Delta_{\theta_2} \right), \quad (5.28)$$

to the corresponding constraints for Y-splitter with trimming:

$$\Theta^{\text{Y+trim}} \Leftarrow \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( y_1^2 - y_{\text{cut}} \right)$$

$$\times \left[ \left[ 1 - \Theta_{\text{trim}} \right] \delta \left( \rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2} \right) + \Theta_{\text{trim}} \delta \left( \rho - x_1 \Delta_{\theta_1} \right) \right]$$

$$\times \left[ \left[ 1 - \Theta_{\text{trim}} \right] \delta \left( \rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2} \right) + \Theta_{\text{trim}} \delta \left( \rho - x_2 \Delta_{\theta_2} \right) \right]$$

$$\Theta \left( y_2^2 - y_1^2 \right) \Theta \left( y_2^2 - y_{\text{cut}} \right)$$

$$\times \left[ \left[ 1 - \Theta_{\text{trim}} \right] \delta \left( \rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2} \right) + \Theta_{\text{trim}} \delta \left( \rho - x_2 \Delta_{\theta_2} \right) \right]$$

$$- \Theta \left( y_1^2 - y_{\text{cut}} \right) \delta \left( \rho - x_1 \Delta_{\theta_1} \right) - \Theta \left( y_2^2 - y_{\text{cut}} \right) \delta \left( \rho - x_2 \Delta_{\theta_2} \right), \quad (5.29)$$

where $\Theta_{i} \equiv \Theta \left( y_{i} - x_i \right) \Theta \left( \Delta_{\theta_i} - \Delta_{R_{\text{trim}}} \right)$ are the constraints imposed by the trimming algorithm, such that emission $i$ is removed from the jet. The emission that passed Y-splitter with the largest $k_t$ always survives the trimming step because it is always sufficiently hard such that it can not be removed, hence we only consider the removal of the emission with the smaller $k_t$. When trimming removes an emission, it changes the groomed jet mass $\rho$, as shown by the relevant delta functions for the double real configuration in the first four lines. The final line represents the one-real one-virtual contributions to jet mass. It is important to note that the normalised $k_t$ distances, $y_i^2$, used in Eq. (5.29) are always normalised to the fat (pre-trimmed) jet mass, which may be different from the groomed jet mass, i.e. at NLO these are always given by $y_i^2 = x_i^2 \Delta_{\theta_i} / \left( x_1 \Delta_{\theta_1} + x_2 \Delta_{\theta_2} \right)$.

We now want to compute the difference; we note that the one-real, one-virtual contributions are identical, hence Eq. (5.28) and Eq. (5.29) only differ when the trimming algorithm removes a single emission from the double real configuration. Hence, we subtract one from the other and arrive at the result:

$$\Theta^{\text{Y-split}} - \Theta^{\text{Y+trim}} = 2 \times \Theta \left( y_1^2 - y_2^2 \right) \Theta \left( y_1^2 - y_{\text{cut}} \right) \Theta_{2}^{\text{trim}}$$

$$\times \left[ \delta \left( \rho - x_1 \Delta_{\theta_1} - x_2 \Delta_{\theta_2} \right) - \delta \left( \rho - x_1 \Delta_{\theta_1} \right) \right], \quad (5.30)$$

where we have accounted for the contribution that arises under the label exchange $1 \leftrightarrow 2$ by an overall factor of two. Using the leading logarithmic approximation to
5.3. **Y-SPLITTER WITH TRIMMING**

tsimplify the terms in the square brackets, we obtain:

\[
\Theta^{\text{Y-split}} - \Theta^{\text{Y+trim}} \simeq 2 \times \Theta (y_1^2 - y_2^2) \Theta (y_1^2 - y_{\text{cut}}) \Theta_2 \times \left[ \delta (\rho - x_2 \Delta \theta_2) \Theta (\rho - x_1 \Delta \theta_1) - \delta (\rho - x_1 \Delta \theta_1) \Theta (x_2 \Delta \theta_2 - \rho) \right].
\]

(5.31)

Thus, we have one integral to evaluate for the correction due to trimming:

\[
\frac{1}{\sigma} \frac{d\sigma}{d\rho} (\text{TrimCorr}, \text{NLO}) = \left( \frac{\alpha_s C_F \pi}{\pi} \right)^2 \frac{1}{2!} \int \frac{d\cos \theta_1}{1 - \cos \theta_1} \frac{d\cos \theta_2}{1 - \cos \theta_2} \frac{dx_1}{x_1} \frac{dx_2}{x_2} \times \frac{d\phi}{2\pi} \Theta (\Delta_R - \Delta \theta_1) \Theta (\Delta_R - \Delta \theta_2) \left[ \Theta^{\text{Y-split}} - \Theta^{\text{Y+trim}} \right],
\]

(5.32)

where we have only taken the double soft and collinear limit of the matrix element for gluon emission, which is sufficient to capture the leading logarithmic behaviour of this correction term in \( \rho \). Evaluating Eq. (5.32) gives the leading result:

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho} (\text{TrimCorr}, \text{NLO}) = -\frac{1}{6} \left( \frac{\alpha_s C_F}{\pi} \right)^2 \times \left[ 0 \times \Theta (y_{\text{cut}}^2 \Delta R_{\text{trim}} - \rho) 
+ \ln^3 \frac{y_{\text{cut}}^2 \Delta R_{\text{trim}}}{\rho} \Theta (\rho - y_{\text{cut}}^2 \Delta R_{\text{trim}}) \Theta (y_{\text{cut}}^2 \Delta R_{\text{trim}} - \rho) 
- \ln^3 \frac{1}{y_{\text{cut}}} \Theta (\rho - y_{\text{cut}} \Delta R_{\text{trim}}) \Theta (y_{\text{cut}} \Delta R - \rho) 
- \left( \ln^3 \frac{y_{\text{cut}}^2 \Delta R}{\rho} + \ln^3 \frac{1}{y_{\text{cut}}} \right) \Theta (\rho - y_{\text{cut}}^2 \Delta R) \Theta (y_{\text{cut}} \Delta R - \rho) \right].
\]

(5.33)

where we have neglected terms of \( \mathcal{O}(y_{\text{cut}}) \) and \( \mathcal{O}(R_{\text{trim}}) \). This leading expression only parametrises the logarithmic terms in \( \rho \), which is sufficient to obtain an understanding of the logarithmic structure of this correction term. We have neglected some constant terms and logarithms of \( y_{\text{cut}} \) and \( R_{\text{trim}} \) in order to obtain this simplified expression. For the full result, we refer the reader to Appendix B. Here one finds that the neglected terms further reduce the magnitude of the result in Eq. (5.33).

One notices that for small \( \rho < y_{\text{cut}}^2 \Delta R_{\text{trim}} \), the correction to Y-splitter is zero at NLO. This result is remarkable, for sufficiently small jet masses the addition of trimming does not change the Y-splitter background jet mass distribution at NLO. The transition boundary denotes the minimum jet mass that can be obtained after removing an emission, given that this removed emission satisfies the constraints that it has the smallest \( k_t \) distance, dominates the pre-trimmed mass and is emitted at an angle greater than \( R_{\text{trim}} \). Below this value, either all emissions lie within a single trimming
CHAPTER 5. Y-SPLITTER WITH TRIMMING

Figure 5.8: Comparison of Y-splitter and Y-splitter+trimming using EVENT2 in the $C_F^2$ channel (right) for two different $R_{\text{trim}}$ values. After subtracting Y-splitter+trimming from Y-splitter at NLO, we see that the difference is consistent with a flat line at zero in the $\rho \to 0$ limit, hence the two results are identical for two emissions in this region.

subjet or the emission that passes Y-splitter also sets the final jet mass. Hence, in this region, the addition of trimming always leaves the jet mass (and Y-splitter distribution) unmodified. In the region $y_{\text{cut}}^2 \Delta R_{\text{trim}} > \rho > y_{\text{cut}} \Delta R_{\text{trim}}$, one finds double logarithmic contributions in $\rho$. However, the magnitude of these corrections can only be as large as logarithms in $y_{\text{cut}}$ due to the upper and lower limits on $\rho$ imposed by the Heaviside step functions. Hence, given that $y_{\text{cut}}$ is not sufficiently small to motivate resummation (whilst being small enough to neglect power corrections in $y_{\text{cut}}$), these terms constitute only $O(\alpha_s^2)$ corrections to the Y-splitter result, which are below our level of accuracy. In the regions $y_{\text{cut}} \Delta R_{\text{trim}} > \rho > y_{\text{cut}}^2 \Delta R$ and $y_{\text{cut}}^2 \Delta R > \rho > y_{\text{cut}} \Delta R$, one observes a constant correction and double logarithmic correction respectively. These terms are also numerically small in these non-vanishing $\rho$ regions and also constitute $O(\alpha_s^2)$ corrections to the Y-splitter distribution. We conclude that the addition of trimming to Y-splitter should not have a significant numerical impact on the Y-splitter jet mass distribution at NLO for typical parameter choices $y_{\text{cut}} \sim 0.1$.

In Eq. (5.33), we calculated the NLO expression for the difference between Y-splitter with trimming for fixed $\alpha_s$. In order to check the form of this analytical expression, we use the fixed-order code EVENT2 at NLO to compare standalone Y-splitter with Y-splitter+trimming. In Fig. 5.8 we compare Y-splitter and Y-splitter+trimming at NLO in the $C_F^2$ channel using EVENT2 for two values of $R_{\text{trim}}$. The difference between Y-splitter and Y-splitter with trimming is consistent with the analytical result in Eq. (5.33) insofar that it vanishes in the region $\rho < y_{\text{cut}}^2 \Delta R_{\text{trim}}$.

5.3.2 Resummation

In this subsection, we compute a resummed expression for Y-splitter with trimming; in doing so, we will obtain a better understanding of the impact of trimming beyond
next-to-leading order. Specifically, we anticipate this correction to manifest as modi-
fication to the plain mass-like Sudakov exponent present in Y-splitter. By performing
this resummation, one can also approximately assess the numerical impact of the dou-
ble logarithmic terms observed for non-vanishing $\rho$ in Eq. (5.33) and investigate if these
terms are likely to result in discontinuities in the differential jet mass distribution for
a given choice of parameters.

We start by using the Lund diagrams for Y-splitter Fig. [5.4] and Fig. [5.6] to write the
differential jet mass distribution for Y-splitter+trimming when one retains the emission
that sets the fat jet mass. The only difference is that we require the anomalous emission
to be emitted at an angle less than $R_{\text{trim}}$, which translates into the additional constraint
\[ x > \rho/R_{\text{trim}}. \] Hence, this contribution is given by the sum:

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y+\text{trim},\text{Retain})} = e^{-D(\rho)} \int_{\max(y_{\text{cut}},\rho/R^2)}^1 \frac{dx}{x} \frac{p_g(x)}{\pi} \frac{\alpha_s(\rho x p_T^2)}{C_F} + e^{-D(\rho)} \int_{\rho/R_{\text{trim}}^2}^{y_{\text{cut}}} \frac{dx}{x} \frac{\alpha_s(\rho x p_T^2)}{\pi} C_F \times \int_{y_{\text{cut}}/\rho}^{\rho} d\rho_2 \int_{y_{\text{cut}}/\rho}^{1} dx_2 p_g(x_2) \frac{\alpha_s(\rho_2 x_2 p_T^2)}{\pi},
\]

(5.34)

where the first term arises when the jet is dominated by the highest $k_t$ scale emission
and the second term is the anomalous configuration when the mass is set by a soft
particle inside the trimming radius, i.e. $\theta < R_{\text{trim}}$ (see Eq. (5.25)).

Evaluating this
expression for fixed $\alpha_s$ and $\rho < y_{\text{cut}} R^2$ we get

\[
\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y+\text{trim},\text{Retain})} \simeq \left[ \frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} + \left( \frac{\alpha_s C_F}{\pi} \right)^2 \frac{1}{2} \ln^2 \frac{1}{y_{\text{cut}}} \ln \frac{y_{\text{cut}} R_{\text{trim}}^2}{\rho} \Theta \left( R_{\text{trim}}^2 y_{\text{cut}} - \rho \right) \right] \times e^{-D(\rho)}.
\]

(5.35)

This leading expression is identical in logarithmic structure to the resummed Y-splitter
expression in Eq. (5.27) under the replacement $R \rightarrow R_{\text{trim}}$ and introduces terms of the
form $\mathcal{O} \left( \alpha_s^2 \ln^2 \frac{1}{y_{\text{cut}}} \ln R_{\text{trim}}^2 \times e^{-D(\rho)} \right)$ in the small $\rho$ limit. These emerge due to the
additional constraint that the anomalous emission is emitted at an angle less than the
trimming radius.

In order to get a full picture, we must also consider all possible configurations
that result in the trimming algorithm removing the emission that dominates the jet

\textsuperscript{8}As in Section 5.2.5 when considering the plain Y-splitter anomalous contribution, we neglect
configuration in which the emission that sets the jet mass is retained by trimming but does not
have the largest $k_t$ distance. We have verified that this contribution has a negligible impact on the
resummed jet mass distribution (a uniform fraction of a percentage at $y_{\text{cut}} = 0.1$). Additionally, this
correction is also present in the plain Y-splitter distribution, hence can be neglected for comparative
purposes.
Figure 5.9: Lund Diagram for the Y-splitter+trimming algorithm applied to a jet with given normalised mass $\rho$. We define the $k_t$ scale cut as a horizontal line and assume that the emission that dominates the jet mass lies on the contour of constant $\rho$ represented by a thick (red) line. We now have the additional consideration that the trimming step will remove emissions that have an angle $\theta > R_{\text{trim}}$ and energy $x < y_{\text{cut}}$. One such emission (shown as a thick orange line) may dominate the mass of the fat jet, $\rho_{\text{fat}}$, and defines the normalisation of the Y-splitter scale $y_{\text{scale}}^2 = y_{\text{cut}}$. The top diagram shows the region $\rho > y_{\text{cut}} R_{\text{trim}}^2$ and the bottom shows $\rho < y_{\text{cut}} R_{\text{trim}}^2$ for two different values of $\rho_{\text{fat}}$. These Lund phase space diagrams for Y-splitter+trimming are shown in Fig. 5.9 for several distinct phase space regions. In this figure, we depict a soft, wide-angle emission (thick orange line) that dominates the mass of the (pre-trimmed) fat jet, $\rho_{\text{fat}}$, and defines the normalisation of the Y-splitter scale $y_{\text{scale}}^2 = y_{\text{cut}}$. The contribution is accompanied by an additional emission (thick red line) that has a normalised $k_t$ scale greater than $y_{\text{cut}}$, (i.e.
5.3. Y-SPLITTER WITH TRIMMING

$x > y_{\text{cut}}(\rho_{\text{fat}}/\rho)$, which is checked by the Y-splitter algorithm and passes the asymmetry cut.

This jet is then subjected to a trimming step that removes all contributions to the jet in the region $x < y_{\text{cut}}$ and $\theta > R_{\text{trim}}$. Therefore, if the emission that sets the fat jet mass lies in this region, it will be removed by the trimming algorithm, changing the mass of the jet such that $\rho \neq \rho_{\text{fat}}$. We must veto on all real emissions $i$ that would have a mass contribution $\rho_i > \rho$, unless they would be removed by the trimming algorithm, i.e. $\theta_i < R_{\text{trim}}$ and $x_i < y_{\text{cut}}$ (both limits are shown as dashed lines on the Lund diagram). This vetoed region is represented by the blue shaded area in each Lund diagram and accounts for Sudakov suppression of real emissions. One can easily see that this Sudakov term has additional structure that differs from the simple plain jet mass suppression for Y-splitter seen in Fig. 5.4.

The topmost panel in Fig. 5.9 shows a jet with initial mass $\rho_{\text{fat}}$ in the region $\rho > y_{\text{cut}}R_{\text{trim}}^2$. In this kinematic region, the emission that sets the fat jet mass is constrained to have an energy $x < y_{\text{cut}}$ in order to be removed by the trimming step. The bottom panels show two possible values for $\rho_{\text{fat}}$ in the region $\rho < y_{\text{cut}}R_{\text{trim}}^2$, and the associated Sudakov suppressed region. If $\rho_{\text{fat}} > y_{\text{cut}}R_{\text{trim}}^2$ (bottom left panel), we require that the emission that sets the fat jet mass to have energy $x < y_{\text{cut}}$, in order to ensure it is removed. If $\rho_{\text{fat}} < y_{\text{cut}}R_{\text{trim}}^2$ (bottom right panel), the energy fraction of the emission that sets the fat jet mass is instead constrained by the trimming radius for a given $\rho_{\text{fat}}$, i.e. $x < \rho_{\text{fat}}/R_{\text{trim}}^2$. One should note that, relative to the top panel, there is an additional triangular shaded Sudakov suppressed region for soft emissions $(x < y_{\text{cut}})$, $\rho_i > \rho$ within the trimming radius; these contributions are not resolved by trimming and will contribute to the jet mass. With these considerations, one can express the contribution of all three diagrams in a compact form:

$$\frac{\rho}{\sigma} \frac{d\sigma}{d\rho}^{(Y_{\text{trim}}, \text{Remove})} = \int_{\rho} \min\left(\rho/y_{\text{cut}}, y_{\text{cut}}R_{\text{trim}}^2\right) \frac{d\rho_{\text{fat}}}{\rho_{\text{fat}}} \times e^{-I} \times \int_{\rho_{\text{fat}}/R_{\text{trim}}^2} \min\left(\rho_{\text{fat}}/R_{\text{trim}}^2, y_{\text{cut}}\right) \frac{dx}{x} \frac{\alpha_s(\rho_{\text{fat}}x p_T^2)C_F}{\pi} \times \int_{y_{\text{cut}}\rho_{\text{fat}}/\rho}^{1} dx_2 p_{gq}(x_2) \frac{\alpha_s(\rho x_2 p_T^2)C_F}{\pi},$$

where the integral on the first line is over all possible values of $\rho_{\text{fat}}$, the second line is over the energy fraction of the emission that defines $\rho_{\text{fat}}$ and the third line is over the energy fraction of the secondary emission that satisfies the Y-splitter cut and defines the final jet mass. The exponent of the Sudakov suppressed region $I$ (shaded blue on

---

9This is analogous to the pruning calculation discussed extensively in Ref. [116] whereby one has a jet mass dependent pruning radius.
Figure 5.10: Comparison of Monte Carlo (left) and analytic differential jet mass distributions (right) for Y-splitter with trimming. In the above plots, we indicate the expected transition points \( \rho = y_{\text{cut}} R^2 \) at which the distribution transitions from the plain jet mass distribution to the next-to-double distribution characteristic of the Y-splitter (+trimming) algorithm. Details of generation are given in Fig. 5.5.
5.3. Y-SPLITTER WITH TRIMMING

corrections starting at $\mathcal{O}(\alpha_s^3)$ arise in all regions of phase space from the form of the exponent defined in Eq. (5.37), which is then integrated over $\rho_{\text{cut}}$. In particular, in the limit $\rho \to 0$ and for typical choices $R_{\text{trim}} \sim 0.3$ and $y_{\text{cut}} \sim 0.1$, the correction due to trimming can be written for $\rho < y_{\text{cut}}^2 R_{\text{trim}}^2$ as

$$
\rho \frac{d\sigma}{d\rho}^{(Y+\text{Trim})} \approx \left[ \frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} + \left( \frac{\alpha_s C_F}{\pi} \right)^2 \frac{1}{2} \ln^2 \frac{1}{y_{\text{cut}}} \ln \frac{y_{\text{cut}} R_{\text{trim}}^2}{\rho} + \mathcal{O}(\alpha_s^3) \right] e^{-D(\rho)},
$$

(5.39)

which is identical to the resummed expression for Y-splitter up to $\mathcal{O}(\alpha_s^3)$ corrections (see Eq. (5.27) for an equivalent fixed-coupling expression). Hence, we expect that the impact of trimming on the Y-splitter mass distribution to be small in this region.

In order to verify the numerical impact of the trimming correction also for non-vanishing $\rho$, we now check the full resummed expression Eq. (5.38) with running coupling effects against Monte Carlo at parton level, which we show in Fig. 5.10. The left panel shows the normalised mass distribution of quark jets generated using HERWIG++ 2.7.0 with a minimum cut on jet $p_T = 3$ TeV, for two different values of $y_{\text{cut}}$ after applying the Y-splitter+trimming algorithm and the right panel is the corresponding resummed analytical distribution. One can see that the shape is dominated by the next-to-double logarithmic distribution characteristic of the standalone Y-splitter algorithm shown in Fig. 5.5. We note that the transition points at $\rho \simeq y_{\text{cut}} R_{\text{trim}}^2$ (indicated by dashed lines) are the same as derived for Y-splitter and in good agreement with the analytical prediction. One also notices a weak discontinuity in the MC and analytical distributions at $\rho \simeq y_{\text{cut}} R_{\text{trim}}^2$ (this corresponds to $\ln \rho \simeq -4.7$ and $\ln \rho \simeq -5.4$ for $y_{\text{cut}} = 0.1$ and $y_{\text{cut}} = 0.05$ respectively), which is a result of the transition in logarithmic behaviour present in the Y-splitter+trimming exponent Eq. (5.37) and the corrections at NLO given by Eq. (5.33).

In order to understand the differences in background mistag rates between Y-splitter with and without trimming (and Y-pruning), we need to compare the mass distributions of each algorithm. In Fig. 5.11 we show the Monte Carlo results (left panel) and corresponding analytical distributions (right panel) for Y-splitter without trimming Eq. (5.22), with trimming Eq. (5.38) and Y-pruning [116]. One can see the similarity between the shape of each distribution; this arises because the logarithmic structure of each algorithm is dominated by a leading term proportional to $\alpha_s \ln \frac{1}{y_{\text{cut}}} e^{-D(\rho)}$ in the differential distribution. However, additional subleading terms vary between each algorithm, and one can see that the analytical results broadly account for the differ-

---

\textsuperscript{10}Explicitly, the correction due to trimming in the region $\rho < y_{\text{cut}}^2 R_{\text{trim}}^2$ for fixed $\alpha_s$ is given by

$$
\frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} \times \frac{1}{x} \left( 1 - x + \frac{x^2}{2} - e^{-x} \right) \times e^{-D(\rho)},
$$

where $x = \frac{\alpha_s C_F}{\pi} \ln \frac{1}{y_{\text{cut}}} \ln R_{\text{trim}}^2$. For typical choices of $y_{\text{cut}} \sim 0.1$ and $R_{\text{trim}} \sim 0.3$, we can consider this expression as a pure $\mathcal{O}(\alpha_s^3)$ correction multiplying the plain Sudakov exponential.
CHAPTER 5. Y-SPLITTER WITH TRIMMING

5.3.3 Hadronisation and UE

It is instructive to provide a Monte Carlo study of the background jet mass distribution for each tagger under the influence of both underlying event (UE) and hadronisation. By including these effects, one can infer the normalised jet mass at which non-perturbative effects become significant and see how they change the distribution. For a normalised jet mass distribution like Y-splitter (+trimming) that always decreases as $\rho \to 0$ (i.e. no Sudakov peaks), one expects that these effects act to suppress background in the small mass region by shifting $\rho$ to higher values when compared to parton level (see discussion in Section 2.4.5). In contrast, as observed in Ref. [116] for the analytic study of trimming and pruning, non-perturbative effects in double logarithmic regions can result in an enhancement for some values of $\rho$, which can be detrimental for background rejection. Additionally, an understanding of the sensitivity to UE contamination is crucial to predict the eventual performance of the tagger when applied to experimental studies.

In Fig. 5.12 we show the differential jet mass distribution produced by Monte Carlo for Y-splitter and Y-splitter+trimming at parton level, with hadronisation and hadronisation + UE. One observes that each non-perturbative effect suppresses the low-mass tail of the distribution in Y-splitter and Y-splitter+trimming. In each case,
5.3. Y-SPLITTER WITH TRIMMING

Figure 5.12: Differential jet mass distribution using Monte Carlo for Y-splitter (left) and Y-splitter+trimming (right) before and after hadronisation and with underlying event (UE). For this plot, we have used $y_{\text{cut}} = f_{\text{cut}} = 0.1$ and $R_{\text{trim}} = 0.3$. Details of generation are given in Fig. 5.5.

we find the corrections to be qualitatively similar for each algorithm and that we find that non-perturbative corrections can actually aid background rejection rates provided that the jet has a sufficiently high transverse momentum for a given signal window. Crucially, the addition of trimming does not significantly affect the response of Y-splitter to non-perturbative effects, hence one again finds that the trimming step is largely passive in terms of background jet rejection rates. An understanding of the impact of non-perturbative effects is useful when considering the optimal parameters for Y-splitter+trimming in the next subsection.

5.3.4 Optimal parameter values

In this subsection we shall use a Monte Carlo study to extract the approximate values of the parameters $y_{\text{cut}}$, $R_{\text{trim}}$ that maximise the signal significance $\varepsilon_S/\sqrt{\varepsilon_B}$ for Y-splitter with trimming. Following the same methodology used in the previous chapter, Section 4.7 for the trimming optimal parameters, we use HERWIG+ 2.7.0 to generate signal events $pp \rightarrow ZH$ and constrain the Higgs and Z boson to decay hadronically and leptonically respectively with a generator level cut on the Higgs $p_T$. We apply the Y-splitter+trimming algorithm to the hardest jet and tag jets with masses in the window $M_H \pm 16$ GeV. The fraction of tagged jets gives the signal tagging efficiency $\varepsilon_S$. We also generate background events $pp \rightarrow q\bar{q}$, with the given cuts and calculate the fraction of jets that fall in the same window, this gives the background mistag rate $\varepsilon_B$. In Fig. 5.13 we show a 2D density plot for the signal significance with Y-
splitter+trimming over a range of $y_{\text{cut}}$ and $R_{\text{trim}}$ values using Monte Carlo at parton level (top) and with full hadronisation and underlying event (bottom) with a minimum $p_T$ cut at 2 and 3 TeV left and right respectively. At parton level, the performance of Y-splitter+trimming is largely insensitive to $y_{\text{cut}}$ and $R_{\text{trim}}$ given that each parameter is not taken too small. Specifically, if $R_{\text{trim}} \ll \frac{M_H}{p_T}$, one encounters signal resolution losses from perturbative, final state radiation during the trimming step, which reduces the signal tagging efficiency. Additionally, the logarithms of $y_{\text{cut}}$ in the background distribution become large in the small $y_{\text{cut}}$ limit (see Eq. (5.39)), increasing the number of background jets within the signal mass window.

The optimal value of $y_{\text{cut}}$ is broadly similar to other taggers such as trimming or Y-pruning and emerges from the trade-off between maximising background rejection from the asymmetry cut in the large $y_{\text{cut}}$ limit (by minimising $\ln \frac{1}{y_{\text{cut}}} e^{-D(\rho)}$) and maximising the tree-level signal efficiency $\varepsilon_S = 1 - 2y_{\text{cut}}$, which favours a small $y_{\text{cut}}$. The optimal value of $R_{\text{trim}}$ will arise from a balance between minimising trimming corrections to Y-splitter for background rejection (large $R_{\text{trim}}$) and ensuring that trimming has sufficient grooming power to improve signal jet resolution ($R_{\text{trim}} \sim \frac{M_H}{p_T}$). For
this optimisation, after including non-perturbative effects, we find that Y-splitter with trimming favours relatively large trimming radius $R_{trim} \sim 0.1 - 0.5$ at high $p_T$ when compared to standalone trimming (see Fig. 4.18); however, in this range one observes that the signal significance is fairly insensitive to variations in $R_{trim}$.

From the fixed-order Eq. (5.33) and resummed Fig. 5.11 analytic calculations, we noted that the correction due to trimming on the Y-splitter background distribution was not significant for non-vanishing $y_{cut}$ and $R_{trim}$ values. This behaviour contributes to the observation that the transition points in the logarithmic behaviour have little numerical impact on the resummed background distribution. However, the impact of non-perturbative corrections are dependent on the value of $R_{trim}$. Explicitly, for large $R_{trim}$, one observes a reduction in signal efficiency, which results from a degradation in signal jet mass resolution due to contamination, i.e. the algorithm becomes more like plain Y-splitter. Conversely, we observed in Fig. 5.12 that non-perturbative corrections improve the background rejection rate by suppressing the tail of the background jet mass distribution. The magnitude of this improvement is maximised for large $R_{trim}$, hence the dependence of the optimal values on $R_{trim}$ of the bottom panels in Fig. 5.13 emerges from the balance of these two contributions. Ultimately, one observes that the dependence on $R_{trim}$ is broadly similar to the partonic picture (top) albeit with a small shift away from the combination of large $R_{trim}$ and small $y_{cut}$ values due to both non-perturbative losses in the signal efficiency and poor background rejection respectively.

In conclusion, we have shown that the addition of trimming to Y-splitter does not significantly change the background jet mass distribution both perturbatively and with respect to non-perturbative corrections. On the other hand, trimming greatly improves the signal efficiency compared to plain Y-splitter; these effects translate in the optimal parameters study as an additional weak dependence on $R_{trim}$ that originates in non-perturbative effects. For a standard choice of $y_{cut} \sim 0.15$, one is therefore able to use any moderate value of $R_{trim} \sim 0.1 - 0.5$ whilst achieving good signal significance of tagging signal over background jets for a range of $p_T$.

5.4 Conclusions

In this chapter we performed an analytical study of the jet substructure technique Y-splitter and its combination with trimming. As observed in the previous chapter, the latter combination left the Y-splitter differential jet mass distribution for background jets largely intact whilst improving the signal efficiency. In order to understand this observation via analytical means, we first embarked on a thorough study of the analytical

\footnote{Specifically, an approximately uniform $\mathcal{O}(20\%)$ increase in the differential distribution in the vicinity of $\rho \approx M_{H}^{2}/p_{T}^{2}$ at $p_T = 3$ TeV using $y_{cut} = 0.1$ and $R_{trim} = 0.3$ can be seen in the MC and analytical results in Fig. 5.11}
structure of Y-splitter when applied to quark jets. We started by calculating leading and next-to-leading fixed-order contributions to the differential jet mass distribution in the eikonal approximation and included soft wide-angle and hard collinear effects. At leading order, we found that Y-splitter exhibited single logarithmic structure and contained a transition point in the differential jet mass distribution, corresponding to the limits of the active region of phase space whereby jets can be rejected by the Y-splitter algorithm. We compared these expressions against the fixed-order code EVENT2 and found that the analytical results correctly reproduce the numerical leading order results, including the transition point. Additionally, we calculated analytical NLO expressions that contain terms of $O(\alpha_s^2 L^3)$ in the integrated jet mass distribution. These results were compared to EVENT2 and were found to be in good agreement up to subleading terms. Hence, the analytical structure of Y-splitter jet mass distribution appeared not to be a simple exponentiation of the LO result, therefore resummation was performed with the aid of Lund diagrams and compared to Monte Carlo results. The shape and transition point of the resummed expression was in agreement with MC, and due to the presence of a double logarithmic Sudakov suppression term, one can show that the Y-splitter algorithm exhibits the best background rejection efficiency compared to the others studied in this chapter (and this thesis).

After analytically demonstrating that Y-splitter performs well at rejecting background jets, it was clear from the last chapter that this technique would benefit from an additional grooming step to remove soft contamination from signal jets, thereby improving signal mass resolution. In the next section, we combined Y-splitter with a subsequent trimming step and calculated the resulting NLO fixed-order correction to the QCD jet mass distribution previously calculated for plain Y-splitter. We found that the result exhibited multiple transition points, but the correction was exactly zero in the small $\rho$ limit. For non-vanishing $\rho$ these corrections, whilst double logarithmic in some regions, were found to be numerically small for typical values of $y_{\text{cut}}$; hence, we anticipated that that the underlying Y-splitter distribution is the dominant contributor to the background distribution. In order to verify this to all-orders and check the magnitude of the modification to the Y-splitter Sudakov exponent, we then calculated an approximate analytic resummed expression using Lund diagrams for the QCD jet mass distribution for the combination of Y-splitter with trimming. We found that the shape of the background distribution was largely unchanged from the plain Y-splitter result and that the correction due to trimming was numerically small for typical parameter choices.

We then carried out a Monte Carlo study of the optimal parameters for Y-splitter with trimming by studying the signal significance of tagging Higgs jets over quark backgrounds. As indicated by the analytical result, the impact of the trimming radius $R_{\text{trim}}$ on the background jet mass was minimal for non-vanishing $R_{\text{trim}}$, hence the
performance of the Y-splitter with trimming algorithm was largely uncorrelated to the choice of $R_{\text{trim}}$ at parton level. After inclusion of non-perturbative effects including hadronisation and underlying event, one notices a weak dependence on $R_{\text{trim}}$ due to signal resolution losses and increased background rejection at large $R_{\text{trim}}$. Despite this, one can conclude that the Y-splitter with trimming optimal parameter values are driven by the $y_{\text{cut}}$ values dictated by the plain Y-splitter background distribution and the trimming step can be considered as an approximately independent improvement to the signal efficiency over a wide range of $R_{\text{trim}}$ values. Hence, for a given standard $y_{\text{cut}} \sim 0.15$, one is able to choose any moderate $R_{\text{trim}} \sim 0.1 - 0.5$ to obtain good signal significance over a range of $p_T$ values.

This work forms a part of a wider effort to analytically understand the action of jet substructure techniques when applied to both QCD and signal jets. In doing so we have obtained analytic expressions for Y-splitter and performed one of the first analytic studies of a combination of two jet substructure techniques: Y-splitter with trimming. In obtaining these results one hopes to gain a further analytical understanding of combinations of jet substructure algorithms, paving the way for the designing of better techniques to separate signal from background in boosted hadronic channels at the LHC. Possible avenues of further research may include the analytic structure of different combinations of jet substructure algorithms that include a tagging and grooming step when applied to signal and QCD jets and their effect on different jet observables.
CONCLUSIONS

In this thesis we conducted a study of some of the algorithmic jet substructure techniques used to distinguish boosted heavy resonance decay from the abundant QCD background at the large hadron collider. In particular, we emphasised an analytical approach to the calculation of jet observables after application of jet substructure taggers and examined the impact of these techniques on the logarithmic structure of the jet mass. This work has made headway into elucidating the action of these jet substructure techniques when applied to QCD and signal jets by providing both analytical expressions and numerical based Monte Carlo results.

In Chapter 3 we computed eikonal, fixed-order perturbative expressions for the normalised jet mass observable in order to examine the impact of substructure algorithms on the logarithmic structure at each order in $\alpha_s$ when applied to QCD jets. These results enabled a deeper analytical understanding of each algorithm and provided an approximate description of the dependence of the jet mass and associated transition points in the differential distribution on each tagger parameter. We identified logarithmic terms as singular as seen in the plain jet result for pruning and trimming algorithms as the jet mass approaches zero, which implies that these algorithms do not remove all soft contributions to the jet, as was possibly intended. For the mass drop tagger, we observed single logarithmic behaviour at LO, but this simple picture changed at NLO due to the possibility of following a massive branch consisting of entirely soft emissions. By identifying the source of these potentially undesirable behaviours in the mass drop tagger and pruning algorithm for QCD jets, one was able to propose some simple modifications in the definitions of each algorithm. In doing so, we showed that the resulting jet mass distribution of the modified mass drop tagger is purely single logarithmic and consequently free from non-global logarithms, making it a unique single jet observable. We also proposed a modification of the pruning technique, Y-pruning, which rejects pruned jets with no identified two-prong hard substructure. This eliminated the leading double logarithms in pruning and produced a resummed QCD jet mass distribution that was suppressed by a double logarithmic exponential; hence, Y-pruning offered favourable background rejection rates compared to the other taggers.
studied here, at high $p_T$. We also discussed how the features from the fixed-order calculation are manifest in the MC and analytical resummed results and the corresponding dependence on tagger parameters.

In order to provide a description of the overall behaviour and performance of each algorithm, we performed a similar study in Chapter 4 for the action of each tagger on the jet mass distribution for signal jets. Using analytical techniques, we demonstrated that radiative ISR and FSR corrections generally did not contain large logarithmic corrections; this observation strengthens the argument for the use of fixed-order calculations for signal efficiencies. One exception was Y-pruning, which exhibited a loss in signal efficiency due to an increase in large logarithmic ISR corrections at high $p_T$. Additionally, the Y-splitter algorithm was found to be susceptible to significant non-perturbative and ISR corrections to the signal efficiency, very similar to that observed for the ungroomed jet. This was partially ameliorated via the inclusion of a subsequent grooming step to Y-splitter, which improved the signal efficiency under non-perturbative and ISR contamination of the signal jet. Specifically, the combination of Y-splitter with trimming was shown to surpass the other taggers studied in this thesis for signal significance, particularly at high $p_T$. Using the result that radiative corrections to the signal efficiency were small for typical parameter choices, we used resummed expressions for the background and fixed-order expressions for the signal and computed the signal significance for tagging Higgs jets over quark backgrounds with a range of substructure methods. These expressions were used to perform an approximate optimisation of tagger parameters and examine the variation of the signal significance on different parameter choices. We found that, for the taggers studied here, the analytically derived optimal values were compatible with the MC results, which indicates robustness against non-perturbative effects in these regions. For each tagger, it was also observed that signal significance was often close to maximal for a fairly wide range of parameter choices.

In Chapter 5 we undertook a thorough analytical investigation of the action of Y-splitter on QCD jets and derived fixed-order and resummed expressions for the corresponding differential jet mass distribution. These results were similar to Y-pruning insofar they produced a single logarithmic leading result suppressed by a plain-like Sudakov exponential. Additional subleading corrections to this result indicated that Y-splitter had the best background rejection rates at high $p_T$, compared to the other taggers considered here. However, studies of the signal in the previous chapter suggested that Y-splitter had to be combined with a grooming step to obtain an acceptable level of signal efficiency. Using MC results we demonstrated that the combination of Y-splitter and trimming preserves the dominant shape of the Y-splitter QCD jet mass distribution. This discovery motivated one of the first analytical studies of a combination of jet substructure techniques and we demonstrated that the addition of trimming
to Y-splitter has only a numerically subleading effect on the QCD jet distribution for typical parameter choices. After deriving a fixed-order and analytic resummed result for the normalised jet mass distribution after Y-splitter with trimming, we show that this combination outperforms Y-pruning at high $p_T$. We also performed an MC optimisation of Y-splitter with trimming and found that the signal significance was fairly insensitive to the trimming radius, given it is not chosen to be much smaller than $M_j/p_T$.

This work indicates that suitably chosen combinations of jet substructure techniques may prove to be superior tools when compared to the currently proposed individual methods. Hence, one hopes that the analytical understanding provided in this thesis will form a basis for choosing combinations of complementary taggers and provide a reference for future analytical studies of such algorithmic combinations. We anticipate that Y-splitter with trimming is only the beginning of this effort and will provide further understanding that is useful for the design of new, better algorithms for use at the LHC and beyond.

Recently, a search for high-mass diboson resonances at ATLAS in Ref. [206] used the mass-drop and filtering algorithm to groom and tag highly boosted boson-tagged jets. This substructure technique enabled the collaboration to observe a $2.5 \sigma$ deviation from the expected background at $\sim 2$ TeV in the di-boson invariant mass distribution. Whilst this is not enough for a discovery, it is encouraging that these techniques are proving effective in the search for BSM heavy resonances. It would be of interest to repeat this analysis with the Y-splitter with trimming technique, which we have shown, in this thesis, has the potential to be superior to (m)MDT for heavy-resonance signal to background rejection at high $p_T$.

For a closing remark we note that, at a centre of mass energy of 13 TeV, the second run at the LHC will be an increasingly noisy environment for extraction of signal events. The resulting escalation of pileup and underlying event may render some hadronic based signal searches intractable without some method of hadronic noise removal. In the boosted regime, jet substructure techniques offer one such solution, hence one anticipates that the experimental implementation of these taggers will only increase in the near future. Therefore, it becomes of importance to design better algorithms and/or find combinations of algorithms that enable better signal to background discrimination in the noisy environment of the LHC. By providing an analytical insight into the action of existing techniques, as well as providing improvements and suggesting effective combinations, this work forms an important step in this endeavour.
FSR CORRECTIONS FOR SIGNAL EFFICIENCY

A.1 Angular integration for FSR

To work out the coefficient of the soft FSR contribution we need to perform the angular and $z$ integrals for the antenna pattern in Eq. (4.21) for trimming and likewise for all taggers. Generally, for a single gluon emission, one has to evaluate the contribution from the FSR emission outside two cones of radius $r$ centred on the $b$ and $\bar{b}$ quarks. The choice of $r$ depends on the tagger in question, so after carrying out the angular integration, one can set $r^2$ as $R_{\text{trim}}^2$ for trimming, $\Delta$ for pruning and $\theta_{b\bar{b}}^2 = \Delta/(z(1-z))$ for mMDT and then integrate over $z$. Explicitly, one has to evaluate the angular integral

$$I = \int \frac{d\Omega}{2\pi} \frac{1 - \cos \theta_{b\bar{b}}}{(1 - \cos \theta_{bk})(1 - \cos \theta_{\bar{b}k})} \Theta \left( \theta_{bk}^2 - r^2 \right) \left( \theta_{\bar{b}k}^2 - r^2 \right), \quad (A.1)$$

where we have written the conditions for the gluon to be at an angle $\theta^2 > r^2$ with respect to both hard partons as Heaviside step functions\(^1\). The simplest way to evaluate the integral in Eq. (A.1) is to first consider an integration over the entire solid angle and then to remove the contribution from inside the two cones around the hard parton directions. We shall assume that the cones do not overlap, i.e. we only consider $r < \theta_{b\bar{b}}/2$. For larger $r$, as appropriate for mMDT where $r = \theta_{b\bar{b}}$, we perform a numerical calculation and find that our results agree with those of Rubin in Ref. [155]. Therefore, we write

$$I = I_{\text{all}} - I_{C_b} - I_{C_{\bar{b}}} \quad (A.2)$$

\(^1\)While we have retained, at this stage, the full angular antenna pattern for ease of comparison to standard formulae, we shall later take the small angle approximation to compute the final answer.
where $I_{\text{all}}$ is the integration over the full solid angle and $I_{C_b, \bar{b}}$ are the integrals inside the region corresponding to cones around $b$ and $\bar{b}$ directions respectively. $I_{\text{all}}$ can be evaluated by standard techniques and yields, after azimuthal averaging, the result in Eq. (1.56) corresponding to angular ordering of soft emission:

$$I_{\text{all}} = \int d(\cos \theta_{bk}) \frac{\Theta (\theta_{bb}^2 - \theta_{bk}^2)}{1 - \cos \theta_{bk}} + \int d(\cos \theta_{bk}) \frac{\Theta (\theta_{bk}^2 - \theta_{bb}^2)}{1 - \cos \theta_{bk}}. \quad (A.3)$$

The contribution inside the cone around $b$, $I_{C_b}$, can be evaluated as follows. Taking the $b$ direction as the ”$z$” axis we define the parton directions by the unit vectors:

$$\vec{n}_b = (0, 0, 1), \quad \vec{n}_{\bar{b}} = (0, \sin \theta_{\bar{b}b}, \cos \theta_{\bar{b}b}), \quad \vec{n}_k = (\sin \theta_{bk} \sin \phi, \sin \theta_{bk} \cos \phi, \cos \theta_{bk}). \quad (A.4)$$

The in-cone subtraction term for $C_b$ can therefore be written as

$$I_{C_b} = \int d\phi \frac{d(\cos \theta_{bk})}{2\pi} \frac{1 - \cos \theta_{\bar{b}b}}{(1 - \cos \theta_{bk}) (1 - \cos \theta_{\bar{b}b} \cos \theta_{bk} - \sin \theta_{bk} \sin \theta_{\bar{b}b} \cos \phi)} \Theta (r^2 - \theta_{bk}^2). \quad (A.5)$$

Integrating over the azimuthal angle $\phi$ gives

$$I_{C_b} = \int d(\cos \theta_{bk}) \frac{1 - \cos \theta_{\bar{b}b}}{(1 - \cos \theta_{bk}) |\cos \theta_{bk} - \cos \theta_{\bar{b}b}|} \Theta (r^2 - \theta_{bk}^2). \quad (A.6)$$

This term can be combined with the corresponding contribution (the first term) in Eq. (A.3), and taking the small-angle approximation for $\cos \theta \approx 1 - \theta^2/2!$ one obtains

$$I = \int_{0}^{r^2} d\theta_{bk}^2 \left( \frac{1}{\theta_{bk}^2} - \frac{\theta_{bb}^2}{\theta_{bk}^2} \left( \frac{\theta_{bb}^2 - \theta_{bk}^2}{\theta_{bk}^2} \right) \right) + \int_{r^2}^{\theta_{bk}^2} \frac{d\theta_{bk}^2}{\theta_{bk}^2} + \{b \leftrightarrow \bar{b}\}, \quad (A.7)$$

where we have included the contribution from $I_{C_b}$ via the interchange $b \leftrightarrow \bar{b}$. The collinear divergence along each hard parton direction is cancelled by the in-cone contributions, leaving only a wide-angle contribution. Carrying out the angular integrations we get

$$I = 2 \log \left( \frac{\theta_{bb}^2 - r^2}{r^2} \right), \quad (A.8)$$

which agrees with the result found by Rubin in Ref. [155] when written in terms of the variable $\eta = \frac{r}{\theta_{bb}}$, for $\eta < \frac{1}{2}$. In the collinear limit, $r \ll \theta_{bb}$, we get the result for trimming quoted in the main text and used in Eq. (4.24). To obtain the result for pruning we substitute $\theta_{bb}^2 = \Delta/(z(1 - z))$ and $r^2 = \Delta$, then carry out the $z$ integral.
A.2. FIXED-ORDER RESULTS VS PARTON SHOWERS FOR FSR CORRECTIONS

which gives:

\[ I = 2 \int_{z_{\text{cut}}}^{1-z_{\text{cut}}} dz \ln \left( \frac{1 - z(1-z)}{z(1-z)} \right) = \frac{2\pi}{\sqrt{3}} + O(z_{\text{cut}}), \]

(A.9)

which corresponds to the result quoted for pruning in Eq. (4.30).

For mMDT where \( r = \theta_{b\bar{b}} \) our calculation above, which assumed non-overlapping cones around the \( b \) and \( \bar{b} \), does not apply. For this purpose, we have evaluated the angular integration in Eq. (A.1) numerically and for \( M_H/p_T \ll 1 \) i.e. when one can use the small-angle approximation, the result is \( I \approx 0.646 \) as found by Rubin for the corresponding quantity \( J(1) \).

A.2 Fixed-order results vs parton showers for FSR corrections

We have noted that FSR corrections computed using the soft approximation gives numerically small corrections to the leading-order results, for sensible choices of the mass window \( \delta M \), and the tagger parameters \( y_{\text{cut}}, z_{\text{cut}}, f_{\text{cut}} \) and \( R_{\text{trim}} \). This of course means that such calculations are not necessarily a good guide to the actual tagger performance in terms of the signal efficiency, because they do not produce genuine logarithmic enhancements. Instead, one can expect that fixed-order calculations, with correct treatment of hard, non-collinear radiation at order \( \alpha_s \) and beyond, will provide a better picture of the behaviour of the tagger signal efficiency. Given that resummation effects are not likely to be significant, it becomes of interest to compare signal efficiencies obtained with pure fixed-order calculations to those from MC generators. One may anticipate that precise order \( \alpha_s \) calculations give quite similar results to full MC parton showers, owing to the dominance of hard radiation and the consequent lack of importance of multiple soft/collinear emissions.

To test this, we ideally need to carry out an exact order \( \alpha_s \) calculation for the process \( H \to b\bar{b}g \). Such a calculation can be straightforwardly performed by taking the exact \( H \to b\bar{b}g \) matrix element and integrating over the available phase space after application of cuts corresponding to jet finding and tagging in various algorithms. While straightforward, this exercise proves cumbersome and has, in any case, to be carried out with numerical integration. One may instead try to obtain the same information more economically by exploiting existing fixed-order codes.

One of the most reliable and long-standing fixed-order programs available to us, and used throughout this thesis, is the code EVENT2 [66] for \( e^+e^- \) annihilation. We can exploit this program by considering the process \( e^+e^- \to Z \to q\bar{q} \) at lowest order and at order \( \alpha_s \) via the production of an extra gluon emission. One can then perform a boost such that the \( Z \) is produced with a large momentum along a given direction; in
APPENDIX A. FSR CORRECTIONS FOR SIGNAL EFFICIENCY

Figure A.1: Ratio for signal efficiency normalised to lowest-order result, with EVENT2 and SHERPA 2.0.0 at parton level, for $e^+e^-$ annihilation with virtual Z production and hadronic decay, where we consider a Z boson with a transverse boost to $p_T = 3$ TeV.

In this regime, the Z decay products will, a significant fraction of the time, form a single fat jet. One can then apply the boosted object taggers to tag the Z boson, imposing a mass window requirement $\delta M$ around $M_Z$, as we have done throughout this paper for the Higgs boson. The situation is similar, but not identical, to Higgs decay we have thus far considered, due to the polarisation of the Z boson. Hence, the matrix element for Z decay to quarks differs from the simple Higgs case and efficiencies at tree-level and beyond are affected, giving, for example, a different dependence on $z_{cut}/y_{cut}/f_{cut}$ at lowest-order. Nevertheless, all of our conclusions about radiative corrections apply to this case as well, including our findings about the logarithmic structure of FSR contributions. This follows because these results originate from the radiation of a gluon from the $q\bar{q}$ pair, which is given by a process independent antenna pattern, which factorises from the process dependent, lowest order decay of a scalar (i.e. Higgs) or a Z boson.

In order to test our basic notion that fixed-order calculations should give a comparable FSR contribution to tagging efficiency when compared to MC event generators, it should therefore suffice to compare results for boosted Z bosons from EVENT2 and MC. In order to minimise any process dependence one should choose precisely the same hard process for both studies, hence we choose to compare EVENT2 against the virtual Z boson contribution in $e^+e^- \rightarrow q\bar{q}$ events using the MC generator SHERPA 2.0.0 shower [76] at parton level. In each case, we boost the Z boson to 3 TeV and impose that it decays hadronically.

We first study the signal efficiencies, normalised to the lowest order result, that are obtained with EVENT2 and SHERPA 2.0.0 for mMDT and pruning for $y_{cut} = z_{cut} = 0.1$.  

A.2. FIXED-ORDER RESULTS VS PARTON SHOWERS FOR FSR CORRECTIONS

These are shown in Fig. A.1 as a function of the mass window, where $\epsilon = 2\delta M/M_Z$ as in the main text.

A first observation is that there is a reasonable degree of qualitative and quantitative similarity between the LO and shower estimates over a wide range of mass windows, which establishes further our point about the essential perturbative stability of taggers against FSR corrections. The difference between the normalised signal efficiencies for SHERPA 2.0.0 and EVENT2 are 2% or less when $\delta M$ is greater than $\sim 8$ GeV or $\sim 13$ GeV for mMDT and pruning respectively. One should not, in any case, consider mass windows significantly lower than these values at high $p_T$, in order to minimise NP hadronisation corrections from ISR (see Eq. (4.14) and subsequent discussion). Differences become more marked for extremely small mass windows (particularly for pruning) signalling the need for resummation and hadronisation corrections. We have also verified that hadronisation corrections have a minimal impact on the shower above the $\delta M$ values stated above, hence basically preserving the picture one obtains at leading-order.

One can similarly study trimming; here the additional choice of $R_{\text{trim}}$ is crucial to ensure that radiative corrections are minimised such that signal efficiency is not significantly affected by FSR emissions. Another way of comparing fixed-order and parton shower results is provided in Fig. A.2 where we show the difference between EVENT2 and SHERPA 2.0.0 signal efficiencies (normalised to the lowest order result) as a function of $z_{\text{cut}}$ and $\epsilon$ for pruning and as a function of $R_{\text{trim}}$ and $\sqrt{\Delta}$ for trimming. The values of $\delta M$, corresponding to the $\epsilon = 2\delta M/M_Z$ values, are shown on the upper
axis for pruning and values of $p_T$ corresponding to $\Delta$ are shown for trimming. The blue shaded region in each case represents parameter values whereby the difference between the normalised signal efficiencies for SHERPA 2.0.0 and EVENT2 are less than two percent, while the green and pink regions correspond to less than five and ten percent respectively. From the pruning plot (left) one notes that there is a correlation between values of $\epsilon$ and $z_{cut}$ needed to minimise radiative corrections. As one goes up in $z_{cut}$, to stay within the five percent zone, for example, one has to correspondingly increase the size of the window. This is in accordance with expectations from our simple analytics, which indicate that one should expect large radiative corrections for $\epsilon \ll z_{cut}$. For trimming, one may expect a correlation between the value of $\sqrt{\Delta}$ and the value of $R_{\text{trim}}$, as required to minimise radiative degradation of mass. This is reflected in the right hand panel of Fig. A.2 where once again the blue, green and pink shaded regions represent differences of 2, 5 and 10 percent respectively, for the normalised signal efficiencies. At $p_T = 600$ GeV, for example, choosing $R_{\text{trim}} \approx 0.13$ or larger gives less than 5 percent difference between leading order and shower descriptions. As one lowers $R_{\text{trim}}$ radiative losses from the shower get progressively larger compared to the fixed-order result.
B

Y-SPLITTER WITH TRIMMING AT NLO

In this section we present the full result for the correction of trimming to Y-splitter at next-to-leading order. One obtains the following result after evaluation of the integral in Eq. (5.32):

$$\frac{\rho \, d\sigma}{\sigma \, d\rho}^{\text{(TrimCorr,NLO)}} = -\frac{1}{6} \left( \frac{\alpha_s C_F}{\pi} \right)^2 \times \left[ 0 \times \Theta \left( y^2 \Delta_{R_{\text{trim}}} - \rho \right) + A_1 \times \Theta \left( \rho - y^2 \Delta_{R_{\text{trim}}} \right) \Theta \left( y \Delta_{R_{\text{trim}}} - \rho \right) + A_2 \times \Theta \left( \rho - y^2 \Delta_{R_{\text{trim}}} \right) \Theta \left( y^2 \Delta_{R} - \rho \right) + A_3 \times \Theta \left( \rho - y^2 \Delta_{R} \right) \Theta \left( y \Delta_{R} - \rho \right) \right], \quad (B.1)$$

where one still obtains zero in the region $y^2 \Delta_{R_{\text{trim}}} > \rho$. The other coefficients in each region are given by

$$A_1 = \pi^2 \ln \Delta_{R_{\text{trim}}} - 3 \ln^2 \Delta_{R_{\text{trim}}} \ln y_{\text{cut}} - 9 \ln \Delta_{R_{\text{trim}}} \ln^2 y_{\text{cut}} - 12 \ln^3 y_{\text{cut}}$$
$$- 6 \ln y_{\text{cut}} \ln \frac{1}{\rho} \ln \frac{y^2 \Delta_{R_{\text{trim}}}}{\sqrt{\rho}} - 6 \ln \frac{y^2}{\rho} \text{Li}_2 \left( -\frac{1}{y_{\text{cut}}} \right) + 6 \ln \Delta_{R_{\text{trim}}} \text{Li}_2 \left( -y_{\text{cut}} \right)$$
$$- 6 \text{Li}_3 \left( -\frac{1}{y_{\text{cut}}} \right) + 6 \text{Li}_3 \left( -\frac{y^2 \Delta_{R_{\text{trim}}}}{\rho} \right), \quad (B.2)$$

$$A_2 = \pi^2 \ln y_{\text{cut}} + 6 \ln y_{\text{cut}} \text{Li}_2 \left( -y_{\text{cut}} \right) - 6 \text{Li}_3 \left( -\frac{1}{y_{\text{cut}}} \right) - \frac{18 \zeta(3)}{4}, \quad (B.3)$$

$$A_3 = -\left( \pi^2 - 3 \ln^2 y_{\text{cut}} \right) \ln \frac{y_{\text{cut}} \Delta_{R}}{\rho} + 3 \ln y_{\text{cut}} \ln^2 \frac{y_{\text{cut}} \Delta_{R}}{\rho} - 6 \ln \frac{y_{\text{cut}} \Delta_{R}}{\rho} \text{Li}_2 \left( -y_{\text{cut}} \right)$$
$$- 6 \text{Li}_3 \left( -\frac{y_{\text{cut}} \Delta_{R}}{\rho} \right) - \frac{18 \zeta(3)}{4}, \quad (B.4)$$

233
where we have used the polylogarithmic function of order $n$:

$$\text{Li}_n(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^n}, \quad (B.5)$$

and the Riemann zeta function:

$$\zeta(x) = \sum_{k=1}^{\infty} \frac{1}{k^x}, \quad (B.6)$$

noting that $\zeta(3) \simeq 1.20$. The expression in Eq. (B.1) contains the dominant $\rho$ dependence of the simplified expression in Eq. (5.33) and hence has the same qualitative behaviour with $\rho$. Consequently, consideration of this result does not affect any of the arguments in the main text about the logarithmic structure of this correction term. However, by including the constant terms and the logarithms of $y_{cut}$ and $R_{trim}$ that were neglected, one has a better understanding of the magnitude of such a correction term in each region. Most importantly, the intermediate region $A_2$ will dictate the maximum size of the correction because $A_1$ and $A_3$ are constrained to numerically match this expression at the boundaries whilst equalling zero in the limits $\rho \to y_{cut}^2 \Delta R_{trim}$ and $\rho \to y_{cut} \Delta R$. Specifically, for typical $y_{cut} = 0.1$ the coefficient of $(\alpha_s C_F / \pi)^2$ in the region $y_{cut}^2 \Delta R_{trim} < \rho < y_{cut} \Delta R_{trim}$ is given by $-A_2/6 \simeq -1.46$ i.e. a genuinely $\mathcal{O} (\alpha_s^2)$ correction.
LUND DIAGRAMS

In order to visualise the sum of emissions to all orders, for observables with complex analytical structure, it convenient to use a kinematic map called a Lund diagram \cite{lund1, lund2}. In Fig. \ref{fig:lund1} we provide a simple Lund diagram, which corresponds to the all-orders plain jet differential mass distribution. Each emission can be represented by a point on the map, with x and y coordinates given by the pseudorapidity $\eta$ and logarithm of the normalised emission $k_t$ with respect to the jet axis respectively. Throughout this thesis we use the emission angle $\theta_i$, normalised energy fraction $x_i$ and mass contribution $\rho_i$ of an emission $i$ as kinematic variables. In order to see how these correspond to the Lund diagram, one notes that in the soft-collinear limit $\eta_i \simeq -\ln (\theta_i/2)$ and $k_t/Q \simeq x_i \theta_i/2$, hence

$$
\ln \frac{k_t}{Q} = -\eta_i + \ln x_i = \eta_i + \ln \frac{\rho_i}{4}
$$

where we have used the soft and collinear contribution of emission $i$ to the normalised jet mass $\rho_i = x_i \theta_i^2$. By inspection of Eq. \eqref{eq:lund1}, one can see that contours of constant $x$ and $\rho$ are straight lines on the Lund diagram with gradient $-1$ and $+1$ respectively. Hence, the red line in Fig. \ref{fig:lund1} depicts all combinations of $\theta_i$ and $x_i$ that an emission can have in order to contribute exactly $\rho$ to the normalised jet mass. The vertical line is the jet radius cutoff at $\theta_i = R$, which places a limit on the minimum energy fraction an emission that contributes $\rho$ to the jet mass whilst remaining inside the jet\footnote{One should note that the Lund diagrams in Fig. \ref{fig:lund4}, Fig. \ref{fig:lund6} and Fig. \ref{fig:lund9} have been drawn such that the rapidity axis starts at the minimum rapidity defined by the jet radius, $\eta = -\ln (R/2)$.}, i.e. $x > \rho/R^2$.

All that remains is the shaded region, which represents a resummed exponential of virtual contributions to the observable to all-orders. In order to understand this, we first note that in the strongly ordered approximation, we can assume that the overall

\[\text{APPENDIX} \]
value for the jet mass is defined exclusively by the most massive emission that lies on the corresponding line of constant mass, shown in red. For a given final value of the normalised jet mass $\rho$, we require that all contributions to the jet mass from other real emissions satisfy $\rho_i \ll \rho$ in such a way that they can be ignored when measuring the jet mass. We also have corresponding virtual emissions that occur for all values of $\rho$. The contribution of these virtual terms to the observable cancel with the real terms completely in the $\rho_i < \rho$ region, but not in the rest of the phase space restricted by the observable (shaded in blue). In this region, the resulting phase space miscancellation causes a tower of logarithms to appear, which can be exponentiated to all-orders. This is best shown by considering the value of the differential distribution in the limit we have an infinite number of strongly ordered real and corresponding virtual emissions.
\[
\frac{1}{\sigma} \frac{d\Sigma(\rho)}{d\rho} = \frac{\partial}{\partial \rho} \sum_{n=0}^{\infty} \frac{1}{n!} \left( C_I \int \frac{dp^T}{p^T} dx \frac{P_{ij}(x)}{\lambda_s} \frac{\alpha_s(p^T)}{\pi} [\Theta(\rho - p^T - 1)]^n \right), \tag{C.2}
\]

where the “−1” comes from the virtual contributions, \(p_{ij}(x)\) is the relevant unregularised Altarelli Parisi splitting function and \(C_I\) is the relevant colour factor (\(C_F\) or \(C_A\) for a jet initiated by a quark or gluon respectively). We have also written the argument of \(\alpha_s\) to be evaluated at a scale that corresponds to the \(k_t\) of the emission with respect to the jet axis. Performing the virtual cancellation we get:

\[
\frac{1}{\sigma} \frac{d\Sigma(\rho)}{d\rho} = \frac{\partial}{\partial \rho} \sum_{n=0}^{\infty} \frac{1}{n!} \left( -C_I \int \frac{R^2}{R^2} d\rho' \int_{\rho'/R^2}^{1} dx \frac{P_{ij}(x)}{\lambda_s} \frac{\alpha_s(p^T)}{\pi} \Theta(\rho' - \rho) \right)^n, \tag{C.3}
\]

where we have imposed that each emission is within the jet radius via the constraint \(x > \rho' / R^2\). We can write this as an exponential and differentiate out the emission that sets the jet mass (which we label with a 1):

\[
\frac{1}{\sigma} \frac{d\Sigma(\rho)}{d\rho} = \int_{\rho/ R^2}^{1} dx_1 \frac{P_{gq}(x_1)}{\lambda_s} \frac{\alpha_s(\rho_1 x_1 p^T)}{\pi} \delta(\rho_1 - \rho) \times \exp \left[ - \int_{\rho}^{R^2} \frac{dp^T}{p^T} \int_{\rho'/R^2}^{1} dx \frac{P_{gq}(x)}{\lambda_s} \frac{\alpha_s(p^T)}{\pi} \Theta(\rho' - \rho) \right]. \tag{C.4}
\]

It is now clearer to see how this expression corresponds to the Lund diagram; the emission that sets the jet mass \(\rho_1\) defines the border of Sudakov suppressed region and resides on the red line, so the non-exponentiated piece on the first line is equivalent to a line integral over all energies \(x_1\) for a given \(\rho\). In the second line one can see an exponential of an integral over the negative of the shaded area to give the Sudakov suppression associated with the veto of emissions with a mass contribution greater than \(\rho\). As one approaches zero jet mass, the shaded area gets larger and we consequently have a larger suppression term in the exponent associated with the vanishing probability of no resolvable emissions from the parton that initiates the jet. Finally, using Eq. \((C.4)\) we can integrate over \(\rho_1\) to give an expression for the resummed plain jet mass in Fig. \(C.1\) as:

\[
\frac{\rho}{\sigma} \frac{d\Sigma(\rho)}{d\rho} = \int_{\rho/R^2}^{1} dx_1 \frac{P_{gq}(x_1)}{\lambda_s} \frac{\alpha_s(\rho x_1 p^T)}{\pi} \frac{C_I}{\pi} \times \exp \left[ - \int_{\rho}^{R^2} \frac{dp^T}{p^T} \int_{\rho'/R^2}^{1} dx \frac{P_{gq}(x)}{\lambda_s} \frac{\alpha_s(p^T)}{\pi} \Theta(\rho' - \rho) \right], \tag{C.5}
\]

which can be evaluated in the fixed coupling approximation to give the differential of
the familiar result in Eq. (2.22):

\[
\frac{\rho \ d\Sigma(\rho)}{\sigma \ d\rho} = C_i \frac{\alpha_s}{\pi} \ln \frac{R^2}{\rho} \times \exp \left[ -C_i \frac{\alpha_s}{\pi} \frac{1}{2} \ln^2 \frac{R^2}{\rho} \right].
\]  

(C.6)


[22] V. V. Ezhela, S. B. Lugovsky, and O. V. Zenin, *Hadronic Part of the Muon g-2 Estimated on the $\sigma_{\text{tot}}^{2003}$ ($e^+ e^- \rightarrow$ hadrons) Evaluated Data Compilation*, [hep-ph/0312114].


[144] S. Chatrchyan et al. [CMS Collaboration], Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state, JHEP 1209 (2012) 29, [arXiv:1204.2488].


BIBLIOGRAPHY


[185] S. Chatrchyan et al. [CMS Collaboration], *Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV*, JHEP 2011 (2011) 109, [arXiv:1107.0330].


M. Bähr, M. Myska, M. H. Seymour et al., _Extracting \(\sigma_{\text{eff}}\) from the CDF \(\gamma + 3\) jets measurement_, JHEP 2013 (2013) 129, [arXiv:1302.4325].

T. Aaltonen, a. Buzatu, B. Kilminster et al., _Improved b-jet Energy Correction for \(H \rightarrow b\bar{b}\) Searches at CDF_, [arXiv:1107.3026].

G. Brooijmans, _High \(p_T\) Hadronic Top Quark Identification_, .


ATLAS Collaboration, _Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at \(\sqrt{s} = 8\) TeV with the ATLAS detector_, [arXiv:1506.0096].