Understanding jet substrucutre at the LHC

Lais Schunk

université PARIS-SACLAY

NNT : 2017SACLS284

Thèse de doctorat

UNIVERSITÉ PARIS SUD
Comprendre le monde, construire l'avenir

Understanding jet substructure at the LHC

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris Sud
(Institut de Physique Théorique, CEA-Saclay)

École doctorale n°564 Physique en Ile de France (EDPIF)
Spécialité de doctorat: Physique

Thèse présentée et soutenue à Gif-sur-Yvette, le 21 septembre 2017, par

Lais SCHUNK

Composition du Jury :

Étienne Augé Professeur des universités, Université Paris-Sud (LAL)	Président
Matteo Cacciari Professeur des universités, Université Paris Diderot (LPTHE)	Rapporteur
Giulia Zanderighi Professeur, University of Oxford	Rapporteur
Michael Spannowsky Professeur associé, Durham University	Examinateur
Grégory Soyez Chargé de recherche, Institut de Physique Théorique	Directeur de thèse

When people thought the earth was flat, they were wrong. When people thought the earth was spherical, they were wrong. But if you think that thinking the earth is spherical is just as wrong as thinking the earth is flat, then your view is wronger than both of them put together.

ISAAC ASIMOV – THE RELATIVITY OF WRONG

Acknowledgments

First of all, I would like to thank my advisor Grégory Soyez, who guided me through this PhD. I consider myself privileged for working with someone whom I admire as a scientist, who always has patience to answer my questions and has such a remarkable enthusiasm for physics. I thank him for all the valuable advice, that I will surely carry with me in the future.

I thank the referees Giulia Zanderighi and Matteo Cacciari for accepting the task of going through these pages. I realize it is a time consuming task, and I am very grateful they took the time to do this. I also thank Étienne Augé and Michael Spannowsky for accepting the invitation to be part of my jury.

I thank my parents, Lucia and Roberto. Their guidance and hard work are reflected in all my achievements, and this thesis is not different. I also thank my brother, Caio, and my extended family, without their support I would not be here today.

I thank Ricardo, for being by my side all these years, through the good and the bad. I cannot imagine how I would conclude this project without him. I am grateful as well for the encouragement I received from all my friends, both in France and in Brazil.

I also thank all collaborators I worked with during this thesis, I learned a lot from all of them. Finally, I thank the administrative and IT staff of IPhT for the support during this three years, and my colleagues in the laboratory.

Abstract

In this thesis we study jet substructure techniques, used to explore the internal dynamics of jets in boosted regimes (i.e. jets with transverse momentum much larger than their mass). We focus on techniques for two-pronged jets, meant to identify boosted W/Z/H bosons. We propose an analytical approach using all-order resummation techniques, in perturbative QCD. In the beginning of this document, we lay down the basic ideas of resummation and introduce the ingredients (basic building blocks) used for our calculations.

Our first study explores the Y-splitter tagger and how its performance is affected by combining it to different grooming techniques : the modified MassDrop Tagger (mMDT), trimming and SoftDrop. It is known that this combination increases the Y-splitter performance, and we studied the origin of this behavior from a first principle approach. We also explore the impact of non-perturbative effects and propose some variations for the original Y-splitter.

Then, we investigate the use jet shapes as discriminating variables between two-pronged hadronic decays of electroweak bosons (W/Z/H) and QCD jets background. We study three shapes: N-subjettiness, energy correlation functions and MassDrop parameter. We carry out analytical calculations for the efficiencies of signal and QCD jets with cuts on these variables. We also compare our results to Monte Carlo generators and study the impact of non-perturbative effects.

Next, we show how the knowladge accumulated in the previous studies can be used to explore the interplay between grooming/tagging techniques and the N-subjettiness. We use the ratio τ_2/τ_1 as a discriminating variable for two-pronged jets. In this work, we propose the dichroic N-subjettiness ratio, where we use a large jet for calculating τ_2 and a smaller, tagged subjet for τ_1. The resulting dichroic ratio gives enhanced performance compared to the original version of the jet shape, while keeping non-perturbative effect under control.

Finally, we perform a phenomenological study of the jet mass distribution with mMDT. Our theoretical predictions account for the resummation of the leading logarithm of the ratio of the jet mass over the jet transverse momentum and are matched to fixed-order matrix elements computed at next-to-leading order accuracy. We consider both the jet transverse momentum measured before (preferred) and after (not collinear safe) the mMDT procedure. Our predictions reproduce the recent measurement by the CMS collaboration.

Résumé

Dans cette thèse on étudie les techniques de sous-structure des jets, utilisées pour explorer la dynamique interne des jets dans les régimes boostés (i.e. jets avec une impulsion transverse beaucoup plus grande que leur masse). On se concentre sur les techniques pour les jets à deux cœurs, pour identifier les bosons W/Z/H boostés. On propose une approche analytique, utilisant des techniques de resommation à tous les ordres en QCD perturbative. Dans la première partie de ce document, on présente les idées basiques concernant la resommation et on introduit les ingrédients (basic building blocks) utilisés dans nos calculs.

Notre première étude explore le Y-splitter tagger et comment sa performance est affectée par la combinaison avec une variété de techniques de grooming : le MassDrop Tagger (mMDT), trimming et SoftDrop. Selon des études Monte Carlo, cette combinaison augmente la performance du Y-splitter, on étudie l'origine de ce comportement avec des calculs théoriques. On explore aussi l'impact des effets non-perturbatives et propose des variantes améliorées de la méthode Y-splitter originale.

Ensuite, on étudie l'utilisation des jet shapes comment une variable discriminante entre les désintégrations hadroniques à deux cœurs des bosons électrofaibles et le bruit de fond des jets QCD. On considère trois shapes couramment utilisées : N-subjettiness, energy correlation functions et le paramètre MassDrop. On calcule analytiquement les efficacités pour des jets QCD et signal avec une coupure sur la variable jet shape. On compare également nos résultats aux générateurs de Monte Carlo et on étudie l'impact des effets non-perturbatifs.

Ensuite, on montre comment le savoir-faire accumulé dans les études antérieures peut être utilisé pour explorer la combinaison des techniques de prong-finder/grooming avec le N-subjettiness. On utilise le rapport τ_2/τ_1 comment une variable discriminante pour les jets à deux cœurs. On propose le rapport dichroïque de N-subjettiness, où on utilise un gros jet (avec ou sans pre-grooming) pour calculer τ_2 et un jet plus petit, avec un prong finder pour τ_1. Cette version donne une performance améliorée par rapport aux versions utilisées actuellement par les expériences, tout en maintenant les effets non-perturbatifs sous contrôle.

Enfin, on effectue une étude phénoménologique de la distribution de masse des jets avec mMDT. Nos prédictions théoriques prennent en compte les logarithmes dominants du rapport de la masse de jet sur l'impulsion transverse et on fait le « matching » avec les éléments de matrice à ordre fixe calculés au NLO. On discute deux options possibles, selon que les distributions sont mesurées dans des bins de l'impulsion transverse avant (version préférée) ou après le mMDT (version collinear unsafe). Nos prédictions reproduisent des mesures faites récemment par la collaboration CMS.

Preface

This thesis in based on work appearing in the following publications

- M. Dasgupta, A. Powling, L. Schunk and G. Soyez, "Improved jet substructure methods: Y-splitter and variants with grooming," JHEP **1612**, 079 (2016) [arXiv:1609.07149 [hep-ph]].
- M. Dasgupta, L. Schunk and G. Soyez, "Jet shapes for boosted jet two-prong decays from first-principles," JHEP **1604**, 166 (2016) [arXiv:1512.00516 [hep-ph]].
- G. P. Salam, L. Schunk and G. Soyez, "Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging," JHEP **1703**, 022 (2017) [arXiv:1612.03917 [hep-ph]].
- S. Marzani, L. Schunk and G. Soyez, "A study of jet mass distributions with grooming," arXiv:1704.02210 [hep-ph].

Résumé de thèse en français

Actuellement, le LHC joue un rôle très important dans la physique aux hautes énergies, pas seulement grâce à la découverte récente du boson de Higgs, mais aussi pour explorer à fond des questions ouvertes dans la physique, comme la nature de la matière noire ou le problème de la hiérarchie. Il fonctionne pour le moment avec une énergie au centre de masse de 13 TeV, en atteignant des énergies bien au-dessus de l'échelle électrofaible. Les collisionneurs futurs vont atteindre des énergies encore plus importantes, comme par exemple, le futur collisionneur circulaire opérant à 100 TeV [1, 2, 3].

Dans les expériences de physique des particules, partons (quarks et gluons) produits aux hautes énergies ne peuvent pas être observés directement à cause des désintégrations colinéaires de la QCD. Ce qui est observé à la place, c'est qu'ils vont désintégrer en plusieurs partons, en produisant des structures collimatées complexes, appelées *jets*. Ces structures sont toujours présentes dans la phénoménologie des collisionneurs de particules et elles sont étudiées depuis plusieurs années. Les particules plus lourdes, comme les bosons de Higgs, Z et W, et le quark top, ne peuvent pas être observées directement non plus, car elles se désintègrent en particules plus légères. Par exemple, un boson W qui se désintègre hadroniquement comme $W \to \bar{q}q$, en principe doit produire deux jets et un quark top doit produire trois jets.

Quand on explore des régimes à hautes énergies cependant, on est confronté avec une situation particulière: la productions d'une quantité importante de particules boostées, c'est-à-dire, particules avec un moment transverse beaucoup plus important que leurs masse $p_t \gg m$. Dans ce régime, le produit de la désintégration hadronique d'une particule lourde va être très collimaté et il peut finir par être groupé dans un seul jet. De plus, jets QCD du bruit de fond peuvent acquérir une masse à cause de leur radiation, fréquemment de la même ordre de grandeur que la masse des particules lourdes discutées précédemment. De telle façon, la question qui émerge naturellement est comment on peut faire la différence entre les jets du signal (boson massifs ou quark top) et les jets du bruit de fond (quarks et gluons légers).

Devant ces challenges, des techniques de *sous-structure des jets* ont été développées pour examiner la dynamique interne des jets. Dans cette thèse, on se concentre sur les désintégrations W/Z/H. On explore le fait que les boson électrofaibles n'ont pas de préférence pour les émissions molles et ils présentent une partage d'énergie plus symétrique, donc les jets du signal ont une structure interne caractéristique à deux cœurs d'énergie. De l'autre côté, les émissions QCD ont une probabilité avec une divergence infrarouge et donc un jet QCD typiquement possédé un seul cœur d'énergie.

Plusieurs outils de sous-structure des jets ont été développés dans les dernières années, ils sont fréquemment divisés dans trois catégories principales:

Prong finders qui imposent que le jet doit contenir deux cœurs durs d'énergie (ou trois pour le quark top), cette situation est plus fréquente pour les jets du signal que pour les jets du bruit de fond QCD (par exemple, MassDrop Tagger [11], le modified MassDrop Tagger [12] et le Y-splitter [13, 14, 15]).

Groomers qui nettoient les jets de la radiation molle et à grand angle, normalement dominée par l'événement sous-jacent, de façon à améliorer la résolution de masse (quelques exemples sont trimming [16] et SoftDrop [17]).

Jet shapes qui contraignent la radiation molle et à grand angle dans le jet, normalement ces observables présentent des valeurs plus importantes pour les jets du bruit de fond QCD que pour les jets du signal (par exemple N-subjettiness [18, 19] et energy correlation functions [18, 19, 21]).

On note que groomers et prong finders fréquemment se comportent de manière similaire et donc la distinction entre les deux n'est pas toujours claire.

Les techniques de sous-structure de jets ont été assimilées par la communauté expérimentale rapidement. Ce sont des outils importants pour le tagging de quarks top et bosons lourds [22, 23], pour la reconstruction de la masse des jets [22, 23] et aussi pour faire la différence entre quarks et gluons [23, 25]. Ces techniques sont utilisées dans plusieurs mesures et recherches, ici on présente quelques exemples (sans avoir une liste complète):

- Mesures de la sections efficace de quarks top boostés [26, 27] ;
- Production de paires de bosons vectorielles WW/WZ associés à un dijet massif [28, 29];
- Étude de certains modes de désintégration du boson de Higgs, notamment $H \to b\bar{b}$ [30, 31];
- Mesure de l'impulsion de subjets dans collisions pp et PbPb, pour mesurer les fonctions de splitting [32];
- Excès temporaire de diboson vers la fin du Run-I du LHC [33, 34].

Une partie considérable de la recherche en sous-structure des jets est basée sur des simulations Monte Carlo, avec des outils comme Pythia [35], Herwig [36] et Sherpa [37]. Ces outils sont très importants, mais ils peuvent être très coûteux du point de vue numérique et ils ne vont pas nécessairement apporter l'information qu'on cherche. Dans cette thèse on propose une approche analytique, qui va nous permettre de comprendre les particularités de chaque méthode. Le premier effort dans la compréhension analytique de la sous-structure des jets a été présenté dans [12], qui a calculé la distribution de la masse des jets pour différentes techniques de groomers et prong finders.

De plus, avoir les expressions analytiques pour les observables combinés avec méthodes de sous-structure de jets nous permet de développer ces techniques dans plusieurs directions. On peut utiliser le savoir-faire acquis pendant ces études pour créer des nouveaux

outils ou pour combiner des outils existants de façon optimale. Aussi, on peut utiliser ces expressions pour optimiser les paramètres de chaque méthode. Finalement, l'approche analytique nous permet de faire des calculs avec une incertitude théorique robuste, de telle façon qu'on peut comparer nos calculs avec les résultats expérimentaux.

Dans cette thèse, on se concentre sur l'approche "traditionnelle" de QCD perturbative, qui explore les propriétés de factorisation et d'exponentiation des éléments de matrices QCD et les caractéristiques des espaces de phase disponibles pour les émissions. On note que, pour mesurer la performance des méthodes de sous-structure des jets, on prend en compte deux aspects. Premièrement, le pouvoir de discrimination (la capacité de distinguer le signal du bruit de fond) et deuxièmement l'insensibilité à l'événement sous-jacent. Fréquemment on observe un compromis entre ces deux propriétés. Une approche analytique nous permet de mieux identifier ces effets et de trouver les moyens pour améliorer les outils de sous-structure de jets dans les deux directions.

Les premiers chapitres de cette thèse sont une introduction générale au domaine où on introduit des concepts importants pour la suite, définit les observables de jet pertinents pour le reste de la thèse et établit la notation utilisée dans la suite.

Dans le chapitre 2, on présente les aspects fondamentaux de la QCD, spécialement les concepts importants pour la suite de la thèse. En particulier, on discute quelques caractéristiques des interactions fortes, comme le principe du confinement de couleur, la liberté asymptotique et la sécurité colinéaire et infra-rouge ("infrared and collinear safety" ou IRC).

Dans le chapitre 3, on présente une définition formelle du concept de jet. Dans la première section 3.1 on discute comment on peut combiner des particules pour former des jets et on présente l'accord de Snowmass – un ensemble de règles à suivre pour avoir une définition cohérente de jet. Ensuite, on présente quelques exemples de définitions de jets. Ils sont divisés en deux "familles": les algorithmes de recombinaison séquentielle (section 3.2) et les algorithmes de cône (section 3.3).

Dans le chapitre 4 on présente quelques techniques de sous-structure de jets (prong finders, groomers et jet shapes) utilisées pour discriminer entre jets du signal et jets du bruit de fond QCD. On se concentre sur les techniques pour l'identification de jets à deux cœurs d'énergie, pour identifier bosons massifs comment le W, Z et Higgs, spécialement celles qui vont être utilisées dans la suite de la thèse.

Dans le chapitre 5 on présente les techniques de resommation qui sont utilisées dans la suite de la thèse. On commence en motivant la nécessité de resommation pour les observables de sous-structure de jet.On étudie un observable relativement simple, la masse d'un jet. Après le calcul, on démontre que la série à l'ordre fixé en α_s ne converge pas dans la limité des environnement boostés. Pour éviter ce problème il faut faire la resommation à tous les ordres $\mathcal{O}(\alpha_s^n)$. De plus, on introduite les diagrammes de Lund, un outil graphique qui illustre les aspects cinématiques des émissions et qui sert de guide pour les discussions physiques. On introduit aussi les notations utilisées dans la thèse, en particulier les "basic building blocks", qui seront utilisés dans la plupart de nos calculs pour la suite.

Les chapitres suivants contiennent des travaux originaux, avec des résultats pour observable de sous-structure des jets. Comme mentionné avant, les premiers résultats ana-

lytiques dans ce domaine ont été obtenus par [12] pour la distribution de masse d'un jet avec différentes techniques de grooming: trimming, pruning et MassDrop Tagger. Dans cette thèse, on a étendu cette compréhension dans deux directions: premièrement pour prong finders dans le chapitre 6 et pour jet shapes dans le chapitre 7.

Dans le chapitre 6 on explore comment la performance du tagger Y-splitter peut être améliorée en combinant celui-ci avec une variété de techniques de grooming, à savoir, modified MassDrop Tagger, trimming ou SoftDrop. Des études récentes [81] ont montre que la combinaison de Y-splitter avec trimming peut surpasser en performance les techniques standards de tagging, même si le Y-splitter tout seul a une performance relativement basse. Notre objectif est d'étudier l'origine de cette différence de performance d'un point de vue analytique. Pour faire cela on présente de calculs pour la distribution de masse des jets QCD avec une coupure $y_{\rm cut}$ sur le Y-splitter à l'ordre fixe et aussi resommé à l'ordre logarithmique dominant en $\log\rho$ et $\log y_{\rm cut}$ (avec $\rho = m^2/R^2p_t^2 \ll 1$ pour les régimes boostés). On présente les même résultats pour le Y-splitter combiné avec un groomer (soit trimming ou mMDT). Finalement, on propose des variations pour le Y-splitter et on étudie l'impact des effets non-perturbatives dans nos prédictions.

Dans le chapitre 7, on étudie les jet shapes comme variables discriminantes entre les jets de signal à deux cœurs et les jets QCD. Les jets shapes imposent des contraintes sur la radiation de gluons dans un jet. On s'attend à ce qu'ils aient un bon pouvoir de discrimination parce que la radiation d'un boson neutre en couleur est différent de la radiation d'un parton QCD coloré. Les observables étudies dans ce chapitre sont N-subjettiness $\tau_{21} = \tau_2/\tau_1$, les energy correlation functions C_2 et une variation du paramètre μ^2 du tagger MDT (voir eq. (4.3)).

On a calculé la distribution de la masse des jets QCD et de jets de signal avec une coupure sur les observables $v < v_{\rm max}$, avec $v = \tau_{21}, C_2$ et μ^2. Cette coupure peut être fixe ou récursive. On a travaillé dans la limite de jets boostées et supposé $v_{max} \ll 1$, qui est la régiondont on a naturellement besoin pour faire la séparation entre les structures à deux cœurs d'énergie et le bruit de fond. Nous visons uniquement à capturer le comportement à l'ordre du logarithme dominant, ce qui est assez pour comprendre les différences entre les jets shapes, même si on a discuté des sources de corrections à un ordre plus élevé. On a testé nos résultats analytiques en comparant avec des générateurs Monte Carlo, aussi utilisés pour étudier l'impact des effets non-perturbatifs.

Cette étude a montré que la différence entre la performance des différentes méthodes vient principalement de la région de radiation à grand angle, avec energy correlation functions le plus efficace. On a aussi étudié la version avec une étape additionnelle de grooming, avec SoftDrop, pour observer comment cela affecte l'efficacité de tagging. Comme le groomer élimine plutôt la radiation molle et à grand angle, cela a diminué l'écart entre les méthodes, mais sans changer l'ordonnancement des efficacités.

Avec une compréhension de chaque ingrédient individuel (prong finders, groomers et jet shapes), on a étudié l'interaction entre ces outils dans le chapitre 8, en utilisant mMDT et SoftDrop combinés avec N-subjettiness. On a introduit le concept de N-subjettiness "dichroïque". Dans cette version, on utilise des (sous-)jets différents au numérateur et au dénominateur du rapport de N-subjettiness τ_2/τ_1. Ces deux sous-jets correspondent à des

dégrées de grooming différents: on calcule τ_2 dans un gros jet (grooming moins agressif ou pas de grooming) et τ_1 dans un petit jet (grooming plus agressif). Calculer τ_2 dans un gros jet nous donne une sensibilité substantielle aux différentes structures de couleur du signal (un singlet de couleur à grandes angles) et du bruit de fond (triplet de couleur pour un quark ou octet pour un gluon). Calculer τ_1 dans un jet plus petit nous assure que cette valeur est dominée principalement par la masse invariante de la structure à deux cœurs. La N-subjettiness dichroïque donne une performance plus élevée que celle de la version originale (qui utilise même jet pour le numérateur et le dénominateur). On a encore une fois utilisé des générateurs Monte Carlo pour étudier les effets non-perturbatives.

Finalement, dans le chapitre 9, on a fait une prédiction phénoménologique pour la distribution de masse avec mMDT (ou SoftDrop avec $\beta = 0$), motivé par des mesures de la collaboration CMS. Pour comparer avec les résultats expérimentaux, on a fait une procédure de "matching" avec des prédictions à l'ordre fixe (valables pour le régime non-boosté) et on a calculé les bandes d'incertitude théoriques. On considère la distribution de masse des jets dans plusieurs classes d'impulsion transverse. Notre prédiction théorique contient les résultats resommés à l'ordre du logarithme dominant en ρ, combinée avec les éléments de matrice à l'ordre fixe au NLO ("next-to-leading order"). On a aussi considéré les effets de z_{cut} fini, qui sont déjà présents à l'ordre des logarithmes dominants.

Travailler avec z_{cut} nous permet de suivre la distinction entre l'impulsion transversale des jets avant ou après le grooming, ce qu'on a appelé $p_{t,\text{jet}}$ et $p_{t,\text{mMDT}}$, respectivement On a trouvé que $p_{t,\text{mMDT}}$ possède plusieurs désavantages théoriques par rapport à $p_{t,\text{jet}}$. Même si les deux resommations sont les mêmes dans la limite $z_{\text{cut}} \to 0$, $p_{t,\text{mMDT}}$ a une structure plus compliquée déjà à l'ordre dominant. Cette différence provient du fait que le spectre $p_{t,\text{mMDT}}$ n'est pas IRC sûr, tandis que le spectre $p_{t,\text{jet}}$ est sûr. Par contre, la version $p_{t,\text{mMDT}}$ est légèrement moins sensible à l'événement sous-jacent. On a exploré les deux versions en détail.

Depuis notre travail original, les résultats CMS on été publiés [47]. On a observé que nos prédictions sont en bon accord avec les données. En particulier, utiliser les éléments de matrice NLO pour faire la procédure de matching améliore substantiellement la concordance à grande masse.

Il y a plusieurs directions possibles à explorer à l'avenir. Premièrement, étendre nos calculs pour des ordres logarithmiques plus élevés. Cela est faisable pour les situations où les effets non-perturbatifs sont contrôlés et partialement éliminés avec grooming. Deuxièmement, on peut développer encore plus nos calculs pour les observables de sous-structure des jets, en particulier pour le N-subjettiness et les energy correlation functions, qui se révèlent efficaces pour discriminer les jets et sont largement utilisés dans les expériences. Un exemple est le calcul pour une coupure finie $v_{\text{cut}} \lesssim 1$. Ce calcul est plus compliqué mais peut être important pour faire des études phénoménologiques car les valeurs typiques pour les coupures expérimentales ne sont pas nécessairement très petites. Une autre possibilité est de calculer les jet shapes avec un exposant angulaire $\beta = 1$, ce cas présente moins de simplifications que le cas $\beta = 2$, mais est le choix par défaut dans plusieurs mesures expérimentales.

Une autre application possible pour les calculs analytiques est d'explorer l'optimisation

des paramètres. Les techniques de sous-structure de jets peuvent être composées de plusieurs outils (jet shape + grooming), habituellement avec une longue liste de paramètres. L'optimisation d'un grand nombre de paramètres avec des outils numériques est trop compliqué, mais elle est réalisable si l'on a des expressions analytiques. De plus, on peut utiliser le savoir-faire des calculs analytiques, spécialement pour les combinaisons de prong-finders/groomers et jet shapes, pour la conception de taggers "décorrélés" [133]. Ils seront capables de fournir un rejet du bruit de fond qui est indépendant de la masse du jet et donc plus simple à utiliser dans le contexte d'estimation du bruit de fond dans des analyses expérimentales.

En résumé, l'approche analytique de la sous-structure des jets peut améliorer la compréhension des techniques qui sont déjà largement utilisées dans les expériences. Nous pouvons utiliser ce savoir-faire pour développer de nouvelles techniques. La combinaison de techniques existantes peut également améliorer considérablement leur performance de manière non triviale. En outre, des calculs précis avec un matching aux éléments de matrice à ordre fixe et des bandes d'erreurs théoriques peuvent être comparés avec des résultats expérimentaux. Ce domaine compte une communauté très active, à la fois théorique et expérimentale, et son importance va surement grandir dans les années suivantes, dans le cadre d'un effort commun pour répondre à des questions ouvertes en physique des particules.

Contents

— 1 —
Introduction

The LHC plays a central role in high energy physics today, not only due to the recent Higgs discovery, but also as a tool to investigate open questions in physics, such as the nature of dark matter and the hierarchy problem. It is running at a center-of-mass energy of 13 TeV, thus reaching energies far above the electroweak scale. Future detectors will achieve even higher energies, for example, there are plans for a future circular collider operating at 100 TeV [1, 2, 3].

In particle physics experiments, partons (quarks and gluons) produced in high energy collisions cannot be directly observed due to collinear branchings and to the QCD confinement principle. Instead they decay into many other partons, producing complex collimated structures called *jets*. These structures are ubiquitous in colliders phenomenology and have been studied for decades. Similarly, heavier particles, such as the Higgs boson, W or Z bosons and the top quark, cannot be directly observed as they decay into lighter particles. For example, a W boson decaying hadronically as $W \to \bar{q}q$, in principle producing two jets, and a top quark decaying as $t \to Wb$, producing three jets.

As we probe higher energies, we face an unprecedented situation: the production of a large quantity of heavy particles in the *boosted* regime, i.e. particles with a transverse momentum much larger than their masses $p_t \gg m$. In this regime, the (hadronic) decay products of a heavy particle will be very collimated and might end up being clustered into a single jet. Additionally, QCD jets in the background may acquire a mass through radiation, often in the same mass range as the heavy particles previously discussed. So the question of how to discriminate between signal jets (from a massive boson or a top quark) and background jets (from lighter partons) naturally emerges.

In face of these challenges, *jet substructure* techniques have been developed to examine the internal dynamics of the jet. In this thesis we will be focusing on W/Z/H decays. We will explore the fact that electroweak bosons show no preference for soft splittings and present a more symmetric energy sharing, so signal jets usually have a characteristic two-pronged internal structure. By the other hand, QCD emission probabilities are infrared enhanced, favoring soft splittings, and hence a QCD jet would typically consist of a single hard prong. Jet substructure techniques can also be applied to other situations, for example, top tagging [4, 5, 6, 7, 8], where one usually has three-pronged jets and quark/gluon jet discrimination [9, 10].

Many different tools have been developed in the past few years, they can often be divided in three main categories:

Prong finders which impose that a jet contain two hard cores (or three for a top-quark), a situation more common in signal jets than in QCD jets which are dominated by soft-gluon radiation; in this category we have the MassDrop Tagger [11] and its modified version [12], and the Y-splitter [13, 14, 15].

Groomers which "clean" the jets of soft and large angle radiation, often dominated by the underlying event, hence ensuring a better mass resolution; some examples are trimming [16], SoftDrop [17] and the modified MassDrop Tagger, that may also work as a groomer in some situations.

Jet shapes or radiation constraints, which constrain soft-gluon radiation inside jets, they are expected to be larger for QCD jets than for weak-boson decays; some examples of jet shapes are N-subjettiness [18, 19]; energy correlation functions [20] and their generalizations [21].

It is worth notice that groomers and prong finders often behave similarly, so the distinction between these two categories is not always clear and depends on what they are being used for.

Jet substructure techniques were rapidly assimilated by the experimental community. They are valuable tools for top quark and heavy boson tagging [22, 23], jet mass reconstruction [24], quark and gluon discrimination [23, 25]. These techniques are used in many measurements and searches, here we present a few of these applications, without the pretension of providing an exhaustive list.

- Measurement of the cross-section of boosted top quarks [26, 27] ;
- Production of vector boson pairs WW/WZ associated with a high mass dijet [28, 29];
- Study of some decaying channels of the Higgs boson, e.g. $H \to b\bar{b}$ [30, 31];
- Measure of momentum sharing between two subjets in pp and PbPb collision, in order to measure splitting functions [32];
- Temporary excess of diboson in mass spectrum towards the end of Run-I [33, 34]

A considerable amount of the research on jet substructure has been based on parton shower Monte Carlo generators, like Pythia [35], Herwig [36] and Sherpa [37]. They are very valuable tools and combined with additional software to cluster jets, the most notable example being FastJet [38, 39], they can be used to simulate jet observables. Nevertheless, in this thesis, we propose an analytical approach which will allow us to understand more deeply the particularities of each method. The work presented in this thesis follows the pioneering work in ref. [12], which calculated jet mass distributions for a few groomers and prong finders. Having analytic expressions allows one, for example, to understand the dependence on each parameter of a given jet substructure method. Doing this type

of investigation using numerical tools can be very costly from a computational point of view, and not necessarily bring useful information. In fig. 1.1, we illustrate this point with a plot of signal efficiency ϵ_{sig} vs. background efficiency ϵ_{bkg}, obtained with a Monte Carlo generator for different jet substructure methods. Although we can see an ordering in performance, it is not obvious to understand the differences between methods using only this type of plot.

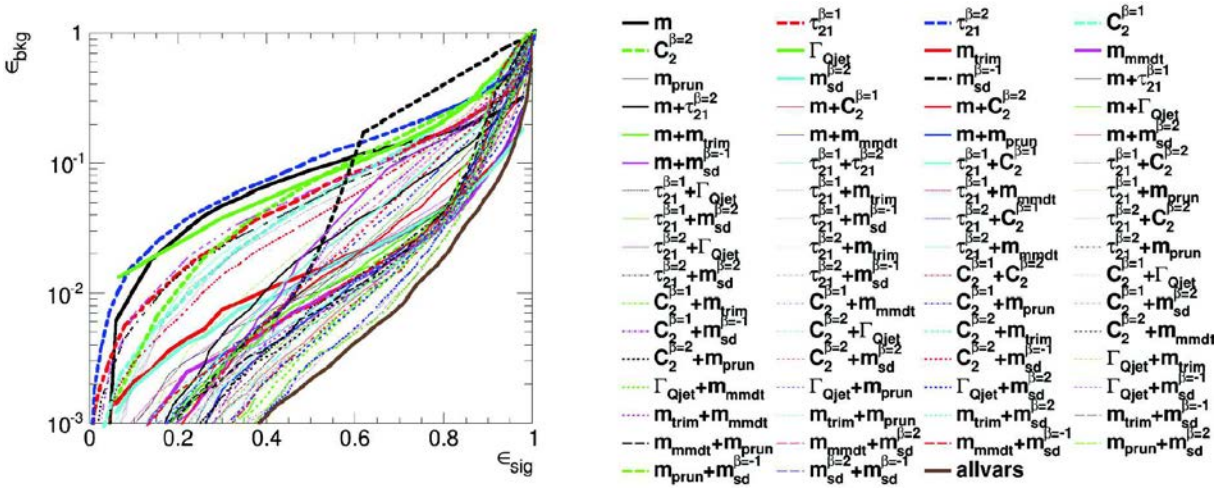


Figure 1.1 — Monte Carlo efficiency curves for several jet substructure methods.

Additionally, having analytic expressions for jet observables with substructure tools allow us to expand them in many directions. We can use the insight from calculations to develop new tools or to combine optimally different tools, through an understanding of how each individual tool constraints the available phase space for emissions. Also, we can use these expressions to optimize jet shapes and prong finders/groomers parameters. This is specially useful in situations with many parameters, where simulations with MC generators are not practical. Finally, the analytical approach also allow the computation of uncertainty bands, which are necessary if one wants to compare theoretical predictions with experiments.

In this work, we will focus on the "traditional" perturbative QCD (pQCD) approach. It explores factorizations and exponentiation properties of QCD matrix elements, as well as the characteristics of the phase space available for parton emissions. Despite being beyond the scope of this thesis, there are other possible approaches to calculate jet substructure observables. Notably, there is the Soft Collinear Effective Theory (SCET) approach [40, 41, 42, 43], an effective field theory for collinear and soft parton emissions. Most recently, there are works using neural networks and machine learning techniques [44, 45, 46] to study jet substructure problems.

Finally, we note that when discussing performance of jet substructure methods we will take into account two aspects. First the discriminating power, i.e. ability to distinguish signal from background, and second the insensitivity to the underlying event and hadronisation effects. Avoiding these effects is important as they are notably difficult to control:

from a theoretical point of view, they contain non-perturbative effects that cannot be computed using perturbative QCD techniques and, from an experimental point of view, they carry a lot of contamination from soft physics and can be dependent on particular characteristics of the detector. Often we observe a trade-off between these two properties. Jet substructure tools usually constraint the phase space available for emission. More aggressive constraints decrease non-perturbative effects, but also decrease the amount of informations we have to identify the jet. An analytical approach allow us to better identify these effects, and find ways to improve jet substructure tools in both directions.

This thesis is divided in two parts: the first one is a general introduction to the domain, where we present the notations being used and we define some important jet and jet substructure tools. The second part contains original work in jet substructure, using an analytical approach to understand existing tools and develop new ones.

The first part is divided as follows, in chapter 2 we recall some of QCD fundamentals, specially focusing on properties that will be useful in this thesis, such as Infrared and Collinear (IRC) safety. In chapter 3 we define jets and we present some selected clustering algorithms, in particular the ones used in this thesis. Chapter 4 introduces boosted jets and present some common jet substructure tools (jet shapes, groomers and prong finders) which will be used in the rest of the thesis. Chapter 5 contains an introduction to the analytical methods used in our calculations, applied to a simple example, the jet mass distribution. We will use this example to justify the need to perform calculation of boosted jet observables at all orders in the perturbative expansion. We also define Lund diagrams, a graphical tools useful to understanding the physical aspects of jet substructure calculations and the basic building blocks, which are the fundamental blocks representing integrations in emissions phase space, and are used to express the results obtained in the second part of the thesis.

The second part of the thesis is organized as follows. In chapter 6 we study how the Y-splitter tagger can be combined with a variety of grooming techniques: the modified MassDrop Tagger (mMDT), trimming and SoftDrop. It is known from Monte Carlo studies that such combination increases the Y-splitter performance, we investigate the origin of this particular behavior. We use the insight to propose improved variation of the original Y-splitter taggers.

In chapter 7 we compute analytically the jet mass distribution with an additional cut on the jet shape variables (N-subjettiness, energy correlation functions and MassDrop parameter). We investigate the performance of these jet shapes as a discriminant variable between two-pronged hadronic decays of electroweak bosons and the QCD jets background. We explain the origin of differences between those methods. We also study the impact of non-perturbative effects and discuss how adding a grooming procedure affects the results.

In chapter 8, we further investigate the interplay between jet shapes and groomers / prong finders. We use the insight gained in the previous chapter to propose the dichroic N-subjettiness ratio. This version uses a large jet (with or without a pre-grooming step) for calculating τ_2 and a smaller, tagged subjet for τ_1. This observable gives an enhanced performance compared to the variants currently used in experimental analyses, while

keeping non-perturbative effects under control.

In chapter 9, we present a phenomenological calculation for the inclusive jet mass distribution with grooming (modified MassDrop Tagger), which will be compared with upcoming LHC measures. Our theoretical predictions account for the resummation of the leading-logarithm of the ratio of the jet mass over the jet transverse momentum and are matched to fixed-order matrix elements computed at next-to-leading order accuracy. We discuss two options according to whether the distributions are measured in bins of the jet transverse momentum before or after the mMDT, and discuss the fact that the latter is not collinear safe. Our predictions agree with the CMS measurement [47], which just become public.

Finally, we conclude in chapter 10, with a final discussion of the results acquired during the thesis. We also present possibilities for future works and developments in the field of jet substructure.

— 2 —
QCD theoretical fundamentals

In this chapter we shortly present some fundamental aspects of Quantum Chromodynamics (QCD), focusing on concepts that will be relevant later in this thesis. We will discuss in particular some particularities of the strong interaction, such as the confinement principle and asymptotic freedom, and also infrared and collinear safety. We do not aim to present a introductory course to QCD, which can be found in [48, 49, 50], nor a comprehensive review, which can be found in [51, 52].

This chapter is divided as follows: section 2.1 is a brief review of QCD and in section 2.2 we introduce the concept of infrared and collinear safety and of Sudakov safety, accompanied by a discussion of the importance of these properties for jet physics.

2.1 Theoretical basis

Quantum Chromodynamics is the gauge field theory describing the strong interactions between color-charged particles, which are commonly referred to as *partons*, or more precisely *quarks* and *gluons*. It corresponds to the SU(3) sector of the Standard Model SU(3) × SU(2) × U(1). The Lagrangian of QCD is given by

$$\mathcal{L} = \sum_q \bar{\psi}_{q,a}(i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t^C_{ab} \mathcal{A}^C_\mu - m_q \delta_{ab})\psi_{q,b} - \frac{1}{4} F^A_{\mu\nu} F^{A\,\mu\nu}, \tag{2.1}$$

where γ^μ are Dirac γ-matrices, $\Psi_{q,a}$ are the quark field spinors for a given flavor q and a given color a. The $\mathcal{A}^C_\mu$ represents the gluon fields, C runs over all types of gluon, which are the adjoint representation the SU(3) group. In the Standard Model there are $N_c = 3$ colors and consequently $N_c^2 - 1 = 8$ types of gluons.

The t^C_{ab} are 3×3 matrices, they are the generators of the SU(3) group. We also have the masses of the quarks m_q and the strong coupling constant g_s, which are the fundamental parameters of QCD. For simplicity, we usually define the constant $\alpha_s = g_s^2/4\pi$ as the characteristic intensity of the strong interaction. Finally, the field tensor in Eq. (2.1) is given by

$$F^A_{\mu\nu} = \partial_\mu \mathcal{A}^A_\mu - \partial_\nu \mathcal{A}^A_\mu - g_s f_{ABC} \mathcal{A}^B_\mu \mathcal{A}^C_\nu, \qquad [t^A, t^B] = i f_{abc} t^C, \tag{2.2}$$

where f_{ABC} are the structure constants of the SU(3) group.

The Feynman rules of QCD involve a 3-gluon vertex, a quark-antiquark-gluon vertex, both of order $\sim g_s$ and a 4-gluon vertex, of order $\sim g_s^2$. There are some QCD constants which emerge from the color-algebra, which will be extensively used in the following. They are associated with specific QCD emissions (respectively $q \to q\,g$, $g \to g\,g$ and $g \to q\,\bar{q}$) and are given by

$$t^A_{ab} t^A_{bc} = C_F \delta_{ac}, \qquad C_F = (N_c^2 - 1)/(2N_c) = 4/3, \tag{2.3}$$

$$f_{ACD} f_{BCD} = C_A \delta_{AB}, \qquad C_A = N_c = 3, \tag{2.4}$$

$$t^A_{ab} t^B_{ab} = T_R \delta_{AB}, \qquad T_R = 1/2. \tag{2.5}$$

In the following I will list some properties of perturbative QCD, which are relevant to jet substructure phenomenology studied in this thesis. In this framework, the observables can be expressed in terms of the renormalized coupling $\alpha_s(\mu_R^2)$. The variable μ_R is the renormalization scale, when one takes μ_R to be of the order of the momentum Q transfered in a given process, $\alpha_s(Q^2)$ can be interpreted as the effective strength of the interaction (see fig. 2.1). The perturbative approach is valid for events that take place at high energy scales. At lower energy scales ($Q^2 \lesssim 1$ GeV), the intensity of strong interactions increases and this approach is no longer valid, this is known as the non-perturbative regime.

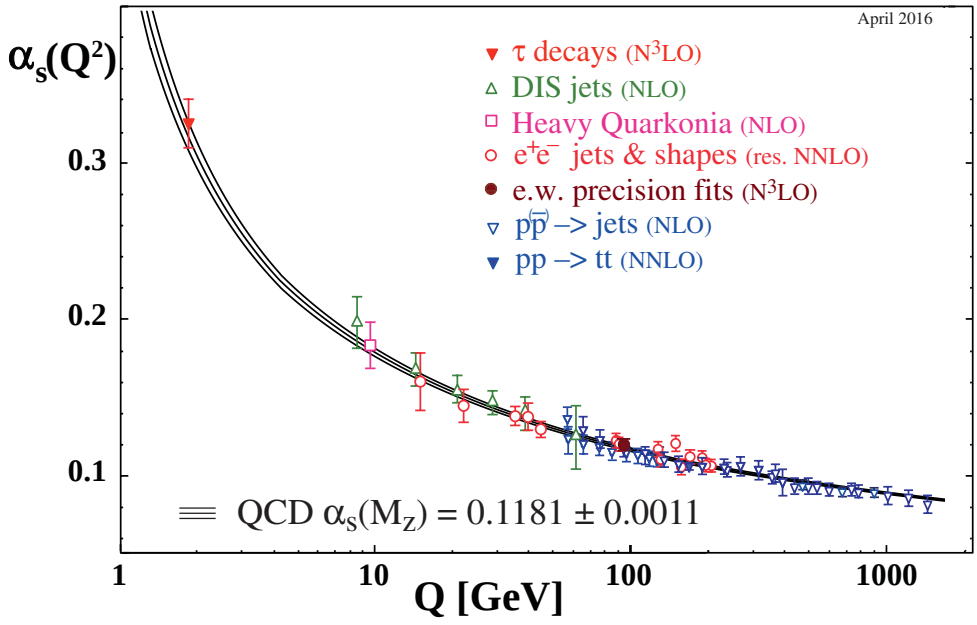


Figure 2.1 — Summary of measurements of α_s as a function of the energy scale Q (from ref. [53]).

The strong coupling satisfies the renormalization group equation

$$\mu_R^2 \frac{\mathrm{d}\alpha_s}{\mathrm{d}\mu_R^2} = -(\beta_0 \alpha_s^2 + \beta_1 \alpha_s^3 + \dots), \tag{2.6}$$

where the term β_i corresponds to the $(i+1)$-loops beta-function coefficient. Results up to the term β_3 are known in literature [54, 55], although in this thesis only the β_0 and β_1 terms are needed. If n_f denotes the number of active quark flavors at a given energy scale, they are found to be

$$\beta_0 = \frac{11C_A - 4n_f T_R}{12\pi} = \frac{33 - 12n_f}{12\pi} \tag{2.7}$$

$$\beta_1 = \frac{17C_A^2 - n_f T_R(10C_A + 6C_F)}{24\pi^2} = \frac{153 - 19n_f}{24\pi^2}. \tag{2.8}$$

Notice the minus sign in equation (2.6), it is the origin of an important QCD property called *asymptotic freedom*. It means the intensity of strong interaction between color-charged particles becomes weaker as the characteristic energy of these interactions increase (or distance between particles decreases), and respectively, interactions become stronger at low energies. For processes involving energies of order 100 GeV or higher, one has $\alpha_s \sim 0.1$.

Another fundamental property of QCD is the *confinement principle*, a well documented behavior from a phenomenological/experimental point of view, although it is not fully understood theoretically. It implies that color charged particles cannot be isolated and therefore, cannot be directly observed. They are necessarily part of bounded states called *hadrons*.

2.2 Infrared and collinear safety

One important aspect of QCD, as well as most gauge theories, is the presence of *collinear* and *infrared* divergences. A fixed-order matrix element $|\mathcal{M}(i, j, \dots)|^2$, involving the partons $i, j, \dots$ diverges in the limit where the angle of emission between two partons tends to zero or the energy carried by one of the partons tends to zero. Examples of infrared and collinear divergences in QCD can be found in literature, for example, in section 3.1 of [49], and in section 17.3 of [48], applied to gluon emissions and jet production.

In this context, one can introduce the concept of *infrared and collinear (IRC) safety* in perturbative QCD. An observable is IRC safe when it can be computed as an expansion of the strong coupling α_s, at any order, in a way that QCD divergences described above cancel.

Let us take an observable $\mathcal{O}$, depending on the 4-momentum p_i of a given number n of QCD partons. A measurement of this observable can be written as

$$\langle \mathcal{O} \rangle = \sum_{n=0}^{\infty} \int \mathrm{d}\Phi_1 \dots \mathrm{d}\Phi_n \mathcal{O}(p_1, \dots, p_n), \tag{2.9}$$

where $\mathrm{d}\Phi_i$ is the integration over the available phase space for each particle and $\mathcal{O}(p_1, \ldots, p_n)$ is the value of the observable as a function of the momenta p_i of each parton i in the event.

One can say this observable is IRC safe if it respects the conditions below. First, $\mathcal{O}$ is safe against soft radiation if adding any number of infinitely soft particles does not change the value of the observable. Which translates into

$$\lim_{E_s \to 0} \mathcal{O}(p_1, \ldots, p_n, p_s) = \mathcal{O}(p_1, \ldots, p_n), \tag{2.10}$$

where p_s is an infinitely soft parton emission, i.e. its energy E_s is infinitely small.

Similarly, an observable is safe against collinear radiation if splitting one existing parton into two collinear partons (i.e. the angle between them is infinitely small) does not change the value of the observable. This property can be written as

$$\mathcal{O}(p_1, \ldots, p_i, \ldots, p_n) = \mathcal{O}(p_1, \ldots, \lambda p_i, (1-\lambda) p_i, \ldots, p_n), \tag{2.11}$$

where λ is any real number, such that $0 < \lambda < 1$.

IRC safety is an important aspect that must be taken into account when building jet observables. From a physical point of view, it makes no sense for an observable to change radically if a very soft or very collinear emission is added to the event. For experiments, IRC safety is also important, as in practice detectors have resolution limitations, which will provide some regularization of the IRC divergences of any observable they measure.

Note that a non-IRC-safe observable, i.e. one that does not have a valid expansion in α_s, can sometimes be calculable in perturbative QCD using all-orders resummation (which will be discussed in chapter 5). This property is know as *Sudakov safety*, meaning that the perturbative Sudakov factor effectively suppresses the singular region of phase space. It is was initially observed in jet substructure objects, the ratio angularities [56]. A robust definition of this condition, based on conditional probabilities, can be found in ref. [57].

— 3 —
Jet basics

The goal of this chapter is to present a formal definition of jets. In section 3.1, we discuss how to combine particles into a jet and we present the Snowmass accord, which sets the basic requirements to be followed by any jet definition. Then, we list some jet algorithms. We do not provide an exhaustive list, but present an overview of some existing algorithms, in particular the ones which will be used in the rest of the thesis. There are two main "families" of jet algorithms: sequential recombination algorithms are covered in section 3.2 and cone algorithms in section 3.3.

3.1 What is a jet

Due to the IRC divergences in QCD, partons in a high energy collider emit a large quantity of other soft and collinear partons. This is known as the parton shower, which produces complex collimated structures called *jets*, ubiquitous in high energy colliders physics. We note that although jets are commonly described in terms of quarks and gluons, these particles are never observed as final-state particles. They will necessarily form bound states, called hadrons. This transition is known as *hadronization*, a non-perturbative phenomenon. At high energies, hadronization does not change considerably the energy flow of an event.

A formal interpretation of jet can be achieved trough a *jet definition*, a set of rules that determines how to cluster a group of particles into a jet. A definition is formed by a *jet algorithm*, a set of rules to recombine particles into a jet; a *recombination scheme*, a rule on how to define the total momentum of a jet and also an ensemble of *parameters*. There is an agreement in the scientific community that a jet definition should meet some basic requirements, this is known as "Snowmass accord" [58], proposed in 1990. The requirements are:

1. Simple to implement in an experimental analysis;
2. Simple to implement in a theoretical calculation;
3. Defined at any order of perturbation theory;

4. Yields finite cross sections at any order of perturbation theory;

5. Yields a cross section that is relatively insensitive to hadronization.

A definition makes sense from a physical standpoint if it can be applied to experiments, Monte Carlo generators and partonic calculations and is coherent across these different representations of an event. Also, it is desirable to have jets that are insensitive to hadronization, which allow us to consider jets at a partonic level. The criteria proposed above aim to fulfill this need.

Regarding the recombination schemes, the simplest example (and the most commonly used) is the 4-momentum sum, where the total momentum of a jet is given by the sum of the 4-momentum of each constituent. Another option useful for some jet substructure tools is the winner-takes-all (WTA) scheme [59, 60], where transverse momentum[1] of the jet is given by the sum of the transverse momentum of the constituents, but the direction of the momentum is the same as the hardest constituent.

There are many different jet algorithms in the literature. One can split these in two broad classes, the first one being *cone algorithms*, which have a "top-down" approach. Their goal is to identify "cone-like" structures in the event, relying on the idea that soft and collinear branching does not modify the direction of energy flow in the event. The second class is *sequential recombination algorithms*, which have a "bottom-up" approach, that iteratively recombine the closest pair of particles according to some predefined measure of distance.

3.2 Sequential recombination algorithms

Sequential recombination algorithms are well adapted to calculations as they deal well with QCD divergences by construction. Also, as they work by successively combining pair of particles, they have the advantage of assigning a clustering sequence to the jet, which is useful for some jet substructure techniques. Another advantage is their fast implementation, a comparison between different reclustering strategies can be found in Appendix A of [39]. In particular, these are the algorithms being used in this thesis.

One of the first recombination algorithm known is the *Jade* algorithm, it was proposed by the JADE collaboration in the 80's [13, 61], in the context of e^+e^- collisions. It depends on a single parameter y_{cut} and it works as follows:

1. For each pair of particles in the ensemble, computes the distance

$$y_{ij}^{\text{Jade}} = \frac{2E_iE_j(1-\cos\theta_{ij})}{Q^2}, \tag{3.1}$$

where Q is the total energy in the event, θ_{ij} is the angle between particles i and j, and E_i is the energy of a given particle i.

[1] The transverse momentum is the component of the 4-momentum transverse to the beam direction, this quantity is frequently used when studying jets from pp collisions.

2. Find the minimum y_{min} in the ensemble of all y_{ij}^{Jade},

   (a) If $y_{min} < y_{cut}$, recombine the two particles into a new one $i + j \to k$ (using a predefined recombination scheme), and replace i and j by the new particle k, then it goes back to step 1;

   (b) If $y_{min} > y_{cut}$, all remaining particles are jets and stop iteration.

The main problem with Jade is that often, in the first stages of clustering, it may happen that two very soft particles going into different directions may be recombined into a single jet due to the factor $\sim E^2$ in the distance definition, which clashes with the expectation of jets as collimated structures.

To overcome this issue, one redefines the distance measure (3.1). Following this line, the k_t *algorithm* was proposed in 1993 [13]. In this case, the distance measure is

$$y_{ij}^{k_t} = \frac{2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{Q^2}, \tag{3.2}$$

and all other steps of the iteration remain the same. We see here that the minimum function prevents the problematic behavior described before. In this case the distance between two soft particles in very different directions will always be larger than one of the soft particles and a hard particle nearby.

Notice that all distance measures introduced above are dimensionless. In the context of pp collisions, this may be a problem as it is much harder to estimate the total energy Q of the event. To avoid this issue, one may replace (3.2) by a dimensional variation [14, 15], where the variables have been changed so they are invariant under longitudinal boosts.

$$d_{ij}^{k_t} = \min(p_{ti}^2, p_{tj}^2)\frac{\Delta\theta_{ij}^2}{R^2}, \tag{3.3}$$

where p_{ti} is the tranverse momentum of a given constituent i, R is an additional parameter, and $\Delta\theta_{ij}$ is the distance between constituents i and j,

$$\Delta\theta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{3.4}$$

with y_i is the rapidity and ϕ_i is the azimuthal angle.

The second challenge comes from the fact that hadron collision are more complex, not only there are interactions between the outgoing partons themselves, but also from the incoming beam with outgoing partons, which cause "extra" emissions. In order to avoid these issues, the *inclusive-k_t algorithm* [15] was proposed. In this case we use the distance measure (3.3), with a parameter R, which works as a characteristic jet radius, and we also introduce the notion of particle-beam distance $d_{iB}^{k_t} = p_{ti}^2$. It works as follows

1. For each subjet and pair of subjets we define $d_{iB}^{k_t}$ and $d_{ij}^{k_t}$;

2. Find the minimum among all distances $d_{iB}^{k_t}$ and $d_{ij}^{k_t}$;

3. If it is a distance between subjets $d_{ij}^{k_t}$, merge i and j as a new subjet;

4. If it is a "beam distance", i is a final state jet and it is removed from the list.

5. Return to step 1 until there are no more subjets, then stop.

Note that our clustering only depends on R, and it is possible for arbitrary soft particles to form jets on their own. In practice, one can add a minimum threshold on transverse momentum to avoid this issue.

Posteriorly, the *generalized-k_t algorithm* has been introduced [62], by redefining the distance measure and adding an extra angular parameter p

$$d_{ij}^{\text{gen-}k_t} = \min\left(p_{ti}^{2p}, p_{tj}^{2p}\right)\frac{\Delta\theta_{ij}^2}{R^2}, \qquad d_{iB}^{\text{gen-}k_t} = p_{ti}^{2p}. \tag{3.5}$$

For particular values of p one can recover different algorithms, for example, for $p = 1$ one recovers the inclusive-k_t algorithm. For $p = -1$, the distance measure becomes

$$d_{ij}^{\text{anti-}k_t} = \min\left(\frac{1}{p_{ti}^2}, \frac{1}{p_{ti}^2}\right)\frac{\Delta\theta_{ij}^2}{R^2}, \qquad d_{iB}^{\text{anti-}k_t} = \frac{1}{p_{ti}^2}, \tag{3.6}$$

this one know as *anti-k_t algorithm*, proposed in [62]. Another possibility is to set $p = 0$, in this case we recover *Cambridge/Aachen algorithm (C/A)* [63], for which the distance measure is simply defined as $\Delta\theta_{ij}^2$ and one stops the iterations when all jets are separated by $\Delta\theta_{ij} > R$. This algorithm is specially useful for jet substructure methods because it preserves angular ordering between constituents.

In this thesis we will use mostly use anti-k_t to do the initial clustering and use the Cambridge/Aachen algorithm and the generalized-k_t algorithm for substructure methods that require a declustering step.

3.3 Cone algorithms

In this category we find the first jet algorithm, proposed by Sterman and Weinberg in the 70s [64] in the context of e^+e^- collisions. It states that an event can be classified as having two jets if more than a fraction $1-\epsilon$ of the total event's energy is contained in two cones of opening angle δ, where δ and ϵ are free parameters. However, this algorithm cannot be generalized to pp colliders straightforwardly. First, the total energy of the event is not necessarily known; and second, because one needs to consider events with more than two jets.

Cone algorithms for pp collisions still maintain the idea from Sterman and Weinberg algorithm of defining an angular cone around the direction of dominant energy flow. For that one has to introduce the concept of *stable cones*. A cone in this context is a circle of a fixed radius R, a free parameter of the algorithm, in the plan of the variables $y \times \phi$, i.e. rapidity vs. azimuthal angle. One can say that a cone is stable when the sum of 4-momenta of its constituents has the same direction as the center of the cone. A cone algorithm seeks to identify all stable cones in an event.

Most of cone algorithms use *seeds* to start this process, which are trial cone directions we start with. Then for each seed we establish a trial cone, evaluate the sum of the 4-momenta of its constituents, and use the resulting 4-momentum as a new trial direction. This procedure is iterated until the cone direction no longer changes, i.e. the cone is stable. There are several variations between methods, namely on how to treat the overlap between cones, i.e. particles that are in multiple stable cones. There are two main approaches to this issue:

Progressive removal we select the particle with the highest momentum as the first seed. Once we found the corresponding stable cone, we remove all particles in this cone from the list and repeat the process with the remaining particles until no particle (above an optional threshold) is left or no other stable cones are found.

Split-merge we selected all particles (above an optional threshold) to be seeds, and find all stable cones corresponding to this initial set of seed. Then we run a split-merge procedure, for each pair of cones that share particles: if more than a given fraction f of of the softer cone's transverse momentum belongs to particles shared with the harder cone they merge; otherwise, they split and each particle goes to the nearest cone.

A more detailed description of these methods, together with a list of possible variations is presented in section 2 of [65].

The seeded approaches present a known issue, they are not IR safe. This boils down to the fact that the algorithms fail to find all possible stable cones, and a infinitely soft or collinear emission can change the number of cones they find. In order to remedy this issue a seedless split-merge approach[2] was proposed in [67]. The Seedless Infrared-Safe Cone (SISCone) explores the geometrical observation that any enclosure in the $y \times \phi$ plane can be moved without changing its contents until it touches two points. As a consequence, one can browse all pair of particles in the event and find their respective cones and then check the stability of said cone. It has a complexity of $\mathcal{O}\left(N^2 \log N\right)$, N being the number of particles in the event, which is faster than other split-merge alternatives.

An infra-red safe alternative to cone algorithms using the progressive removal approach is the anti-k_t algorithm. Despite it being a sequential recombination algorithm, it tends to recombine soft particles with hard ones, before recombining soft particles between themselves. It produces soft-resilient jets with regular boundaries, e.g. jets with that are similar to the ones obtained with cone algorithms [62].

[2]Seedless approaches were proposed before [66], but they were not viable due to the time of computation.

— 4 —
Boosted jets

As discussed in the introduction (chapter 1), at the LHC one has the production of a large quantity of boosted heavy particles. We observe that in boosted environments the hadronic decay products of these particles are very collimated, and they tend to be clustered in a single jet. In fact, one can estimate the opening angle of a heavy boson decaying hadronically, it will be proportinal to $\theta \sim m/p_t$, where m is the mass of the decaying boson and p_t its traverse momentum, as illustrated in fig. 4.1.

In this situation, one has to use jet substructure techniques to separate H/W/Z jets from QCD jets. In this chapter, we present some of these techniques, focusing on the ones that will be used later on the thesis. In section 4.1 we present groomers and prong finders, and in section 4.2 we present jet shapes. The shapes listed here focus on the identification of 2-pronged signal jets (Z, W and Higgs bosons). It is possible to generalize them to other cases, in particular to top tagging, where one has typically a 3-pronged structure, but it is beyond the scope of this work.

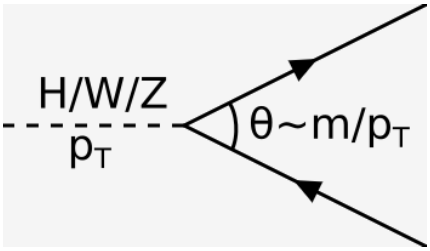


Figure 4.1 — Hadronically decaying boson in a boosted environment.

4.1 Groomer and prong finders

Jet groomers are procedures conceived to "clean" part of the soft and large-angle radiation present in a jet. This is the region of the phase space where one has a considerable contribution of the underlying event, i.e. what is detected in a event that is not coming purely from the primary hard scattering process. Prong finders on the other hand try to identify hard prongs in the jet, exploring the fact that QCD jets usually have one prong and signal jets have multiple prongs. The distinction between groomers and prong finder is not always clear, and one same procedure can be considered a prong finder or a groomer depending on many factors, such as how they are combined with other methods or even the choice of parameters.

The **Y-splitter** tagger [13, 14, 15] clusters the jet j with the k_t algorithm, then undo the last step of the clustering, dividing the jet into constituents j_1 and j_2, with a correponding k_t distance, defined as

$$d_{12} = \min\left(p_{t1}^2, p_{t2}^2\right)\theta_{12}^2, \tag{4.1}$$

where p_{t1} and p_{t2} are the transverse momenta of the two subjets and θ_{12}^2 is their angular separation in the rapidity-azimuth plane.[1] Then one examines the value of this distance with respect to a given mass m, usually taken to be the mass of the jet or the mass of the boson we are tagging

$$y = \frac{d_{12}}{m_{\text{jet}}^2}. \tag{4.2}$$

We retain jets that satisfy $y > y_{\text{cut}}$. This cut is designed to retain more symmetric signal splittings (i.e. a genuine two-pronged structure). We study the Y-splitter in detail in chapter 6 and introduce variants of this original definition.

Trimming [16] is a grooming tool that depends on two parameters R_{sub} and z_{cut}. We can apply it to a given jet j, with a characteristic radius R, as follows

1. We take its constituents and recluster them into subjets of radius $R_{sub} < R$ with a jet algorithm of choice, usually the Cambridge/Aachen algorithm or the k_t algorithm;
2. For each subjet we compare $p_{t,\text{sub}} > z_{\text{cut}} p_{t,\text{jet}}$, if the subjet fails this test it is eliminated;
3. Assemble remaining subjets into the trimmed jet.

A first analytic study of trimming was achieved in [12]. Here, we will use it briefly in chapter 6.

The **MassDrop Tagger (MDT)** [11] was initially designed to be used with jets found by the Cambridge/Aachen algorithm. It involves two parameters y_{cut} and μ and, for an initial jet labeled j, proceeds as follows:

1. Break the jet j into two subjets by undoing its last stage of clustering. Label the two subjets j_1, j_2 such that $m_{j1} > m_{j2}$;
2. If there was a significant mass drop,

$$m_{j1} < \mu m_{j2}, \tag{4.3}$$

and the splitting is not too asymmetric,

$$y = \frac{\min\left(p_{t,j1}^2, p_{t,j2}^2\right)\theta_{12}^2}{m_{\text{jet}}^2} > y_{\text{cut}}, \tag{4.4}$$

then we consider j to be the tagged jet.

[1] For simplicity, we will use the simplified notation θ_{12} instead of $\Delta\theta_{12}$.

3. Otherwise redefine j to be equal to j_1 and go back to step 1 (unless j consists of just a single particle, in which case the original jet is deemed untagged).

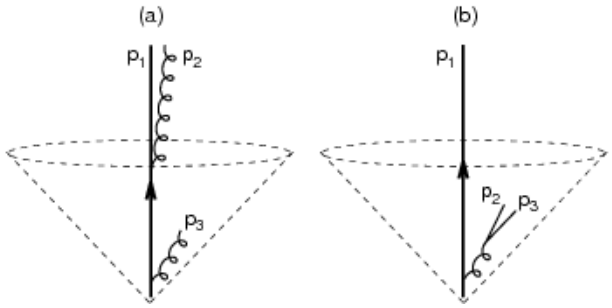


Figure 4.2 — Three particle configurations $\mathcal{O}(\alpha_s^2)$ at tree-level (from ref. [12]).

This algorithm is intended to find hard substructure in the jet, even in the presence of soft substructure strongly influencing the total jet mass. Lets observe a simple 3-particle configuration in fig. 4.2 (a) to better understand it. The parton p_1 emits p_2 and p_3, such that $\theta_{13} \ll \theta_{12}$ and the soft emission p_3 dominates the mass $m_{\text{jet}} \ll m_{12}$ (m_{12} is the mass of the subjet formed by partons $p1$ and p_2). The algorithm C/A will split the jet into subjets j_{12} and j_3, if it passes MDT the jet is tagged, if not it follows the subjet j_{12}, which contains the hard structure.

An unintended behavior for MDT was pointed out in [12], the problematic configurations is represented in fig. 4.2 (b). Here we have $\theta_{23} \ll \theta_{12} \sim \theta_{13}$, so the C/A declustering will separate j_1 and j_{23} as subjets. If the parent gluon of j_{23} is too soft, the jet will fail the symmetry condition and the prong finder will follow the most massive subjet j_{23}, and drop j_1, which happens to contain the hard substructure of the jet.

In order to solve this "wrong branch issue" an alternative version for MDT was proposed, the **modified MassDrop Tagger** (mMDT) [12]. It is similar to the previous version, except that it follows the subjet with the largest transverse momentum, instead of the most massive one. It follows same procedure as the original version, except that one replaces the step 3 of the recursion with

3. Otherwise redefine j to be that of j_1 and j_2 with the larger transverse momentum $p_{t,i}$ and go back to step 1 (unless j consists of just a single particle, in which case the original jet is deemed untagged).

Other variations can also be considered: following the subjet with largest transverse mass $\sqrt{m^2 + p_t^2}$; changing the y cut by a $z = \min(p_{t1}, p_{t2})/p_{t,\text{jet}}$ cut; or dropping the mass drop condition (4.3). In this document, we will use by default the largest p_t, the z_{cut} option and neglect the condition (4.3). We study the mMDT combined with the Y-splitter tagger in chapter 6, combined with N-subjettiness in chapter 8, and as a groomer for a phenomenological study of the jet mass distribution in chapter 9.

Finally, the **SoftDrop** grooming procedure [17] starts by reclustering the jet with the Cambridge/Aachen algorithm. Then it recursively undoes the last step of the clustering,

splitting the current jet j into two subjets j_1 and j_2. The procedures then stop if the splitting is symmetric enough, i.e.

$$z > z_{\text{cut}} \left(\frac{\theta_{12}}{R}\right)^{\beta}, \qquad z \equiv \frac{\min(p_{t1}, p_{t2})}{p_{t,\text{jet}}}. \tag{4.5}$$

If the symmetry condition is not met, the procedure is recursively applied to the subjet with the largest p_t.

Note that eq. (4.5) with $\beta = 0$ corresponds to the mMDT procedure, if one does not consider the mass drop condition. The SoftDrop generalizes it to the case of $\beta \neq 0$. Typically SD is considered a groomer for $\beta > 0$ and a prong finder for $\beta \leq 0$. The largest the value of β, the less aggressive will be the grooming procedure (i.e. the less particles are going to be groomed from the jet). In this document we use SD mostly as a groomer in chapters 6, 7 and 8.

4.2 Jet shapes

Jet shapes are functions of the constituents of the jet. They explore the fact that different type of jets have different patterns of soft-gluon radiation and are conceived in such a way that they have larger values for QCD jets and smaller values for signal jets. One way of discriminating signal jets from QCD background is imposing a cut on these observables. Below, we only define the jet shapes stidied in this thesis, mainly in chapters 7 and 8.

***N*-subjettiness** [18, 19] is defined as follows: for a given jet, one finds a set of N axes $a_1, \ldots, a_N$ (more details on that below) and introduces

$$\tau_N^{\beta_\tau} = \sum_{i \in \text{jet}} p_{ti} \min(\theta_{ia_1}^{\beta_\tau}, \ldots, \theta_{ia_N}^{\beta_\tau}), \tag{4.6}$$

where the sum runs over all the constituents of the jet, of transverse momentum p_{ti} and with an angular distance $\theta_{ia_j} = \sqrt{\Delta y_{ia_j}^2 + \Delta\phi_{ia_j}^2}$ to the axis a_j. The standard choice for angular parameter is $\beta_\tau = 1$, but for this thesis we consider the case $\beta_\tau = 2$. This choice has shown good performance in Monte-Carlo numerical simulations [68] and considerably simplifies analytical calculation, as we will see in chapter 7. A comparison between $\beta_\tau = 1$ and $\beta_\tau = 2$ will also be presented in chapter 8.

In order to have a complete definition of N-subjettiness one needs to determine how to choose the axis $a_1, \ldots, a_n$. Common options include:

- the optimal axes, which minimize τ_N ;
- the k_t axes, obtained by declustering the jet successively with the k_t algorithm until we are left with N subjets;
- the generalized-k_t axes with $p = 1/2$, that we introduced in [69] (see chapter 7), obtained by declustering the jet successively with the generalized-k_t algorithm until we are left with N subjets;

The generalized-k_t option has the advantage of leading to a similar performance to the optimal axes at much smaller computational cost, for the case $\beta_\tau = 2$. More generally, for $\tau_N^{\beta_\tau}$ with a generic β_τ, we would expect the generalized-k_t axes with $p = 1/\beta_\tau$ to give a close-to-optimal result. Note that for $\beta_\tau \leqslant 1$, this choice should be used with the winner-take-all recombination scheme (defined in the previous chapter) in order to avoid recoil effects.

In what follows, we will consider the ratio $\tau_{21}^{\beta_\tau} = \tau_2^{\beta_\tau}/\tau_1^{\beta_\tau}$ to discriminate 2-pronged signal jets. Notice that, for our choice of $\beta_\tau = 2$, the 1-subjettiness $\tau_1^{(2)}$ is equivalent to the squared jet mass at NNLL accuracy. This is true whenever axis defining τ_1 aligns with the jet axis, in particular for the axes choices described above.

The **MassDrop parameter** μ^2 [69], a variation of the MassDrop Tagger [11]. For a given jet, we uncluster it into two subjets $j \to j_1 + j_2$, and define

$$\mu^2 = \frac{\max\left(m_{j1}^2, m_{j2}^2\right)}{m_{\text{jet}}^2}. \tag{4.7}$$

We use the generalized-k_t (3.5), with a given parameter p, to decluster the jet. The typical choice is $p = 1/2$. The main difference from the original formulation of the MassDrop Tagger is that in this observable we do not apply any recursion, i.e. we simply discard the jet if the condition (7.12) is not satisfied for the first declustering.

The **energy correlation functions** [20] are defined as following

$$e_2^{\beta_C} = \frac{1}{p_{t,\text{jet}}^2 R^{\beta_C}} \sum_{i<j\in\text{jet}} p_{t,i} p_{t,j} \theta_{ij}^{\beta_C}, \tag{4.8}$$

$$e_3^{\beta_C} = \frac{1}{p_{t,\text{jet}}^3 R^{3\beta_C}} \sum_{i<j<k\in\text{jet}} p_{t,i} p_{t,j} p_{t,k} \theta_{ij}^{\beta_C} \theta_{ik}^{\beta_C} \theta_{jk}^{\beta_C}, \tag{4.9}$$

$$\tag{4.10}$$

where we summed over all pairs/triplets of constituents of a jet j of radius R and $p_{t,i}$ is the transverse momentum of a particle i, and $\theta_{ij}^2 = \Delta y_{ij}^2 + \Delta\phi_{ij}^2$ is the distance between particles i and j in the plan of rapidity vs. azimuthal angle.

For 2-pronged jet reconstruction, one generally uses the observable $C_2^{\beta_C} = e_3^{\beta_C}/(e_2^{\beta_C})^2$ as a discriminating variable. Alternatively, one can use $D_2^{\beta_C} = e_3^{\beta_C}/(e_2^{\beta_C})^3$ [70]. Once again, we will concentrate on the cases $\beta_C = 2$, for which $e_2^{(2)}$ is equal to the squared jet mass, for massless partons and small emission angles.

There are more general versions of the energy correlation functions, proposed recently in [21] and not studied in this thesis. This version depends on the angular parameter β_C, and on two parameters n and v. It is written as

$${}_v e_n^{\beta_C} = \sum_{1<i_1<\cdots<i_n<n_j} z_{i_1} \ldots z_{i_n} \prod_{m=1}^{v} \min_{s<t\in\{i_1\ldots i_n\}}^{m} \{\theta_{st}^{\beta_C}\}, \tag{4.11}$$

where are $\min^m$ denotes the m-th smallest number in the sequence.

There are new substructure discriminants, based on this energy correlation functions, notably:

- The $M_i = {}_1e_{i+1}^{\beta_C} / {}_1e_i^{\beta_C}$ series, constructed to identify i hard prongs;
- The $N_i = {}_2e_{i+1}^{\beta_C} / ({}_1e_i^{\beta_C})^2$ series, also suited to identify i hard prongs;
- The $U_i = {}_2e_{i+1}^{\beta_C}$ series, designed for quark/gluons discrimination.

— 5 —
Bases of resummation techniques

In this chapter we present the techniques used to do analytic calculation in the rest of this thesis. We start by motivating the need of resummation for calculating jet substructure observables. We study a relatively simple observable, the jet mass, we first do this calculation at fixed order in α_s. We demonstrate that, for the boosted regime, this series does not converge. In order to solve this situation, we introduce all-order resummation, where we take into account terms from all orders $\mathcal{O}(\alpha_s^n)$. Additionally, we introduce Lund diagrams, a helpful graphic tool for guiding physical discussion. We also introduce the notation used in the rest of the thesis, in particular the basic building blocks. They are objects that will be extensively used in all substructure calculations in the next chapters.

The chapter is organized as follows, in section 5.1, we explain Lund diagrams. We present a first example of calculation in section 5.2, the jet mass distribution, first at fixed order (leading order and next-to-leading order), and then the all-order resummed calculation. Finally in section 5.3, we will present the basic building blocks.

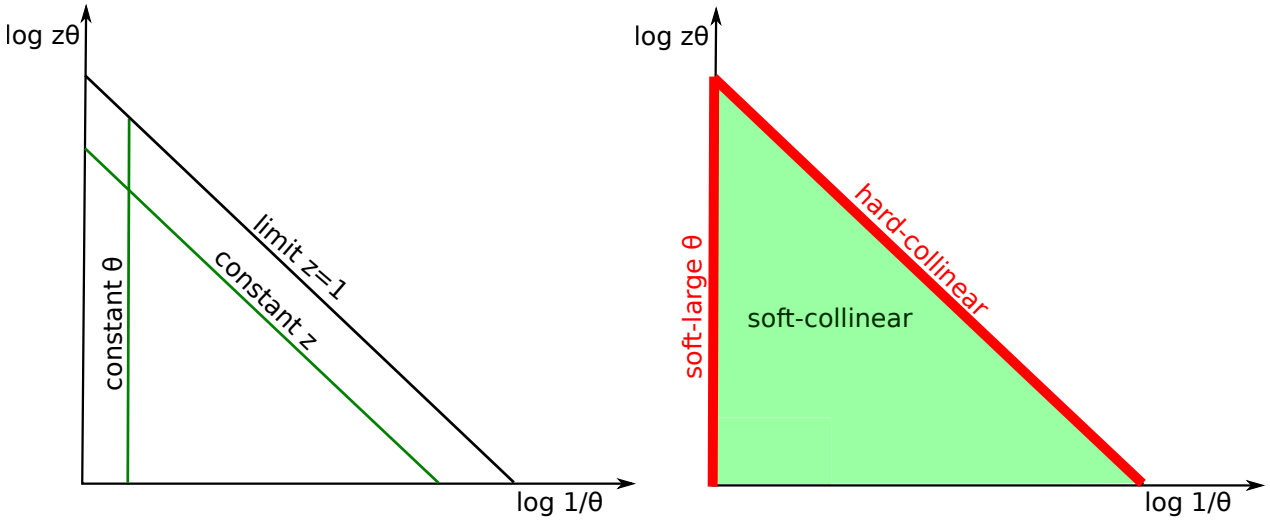
5.1 Lund Diagram

Lund diagrams [71] are a graphic tool that illustrate the kinematic aspects of particles emissions. A given emission carrying a fraction z of the total transverse momentum of the parton that initiated the jet and with an emission angle θ, will be represented as a point on the diagram. The diagram has $\log(z\theta)$ (proportional to the transverse momentum of the emission k_t) along the vertical axes and $\log(1/\theta)$ along the horizontal axes. The kinematic limit $z = 1$ is represented by the diagonal black line, and it is impossible to have emissions above this line.

In the soft and collinear limit, each emission in the diagram has the following weight

$$\mathrm{d}\omega^2 = C_R \frac{\mathrm{d}z}{z} \frac{\mathrm{d}\theta}{\theta} \frac{\alpha_s(z\theta p_t)}{\pi}, \tag{5.1}$$

where C_R is a color factor depending on the parton that originated the emission ($C_R = C_F$ for a quark, and $C_R = C_A$ for a gluon), and we included the strong coupling α_s which depends on the relative transverse momentum scale of the emission, with p_t being the



Figu e 5.1 — Left: Lund diagram with lines to represent emissions at a fixed angle θ and a fixed momentum z. Right: Lund diagram with soft/hard and collinear/large θ regions represented.

transverse momentum of the original parton. We can see that the weight (5.1) has the characteristic QCD soft and collinear divergences. It will become explicit for the jet mass distribution calculation, performed in next section, that jet observable distributions in boosted regimes can be interpreted as an integration over a Lund diagram phase space, with additional constraints imposed by the observable being considered.

An emission with a fixed emission angle θ can be anywhere on the vertical green line represented in fig. 5.1 (left), and a particle with a fixed momentum fraction z can be anywhere in the diagonal green line. One can also represent other observables in this diagram as we show later for the jet mass (see fig. 5.4). Emissions that are both collinear and soft occupy the central area of the diagram, represented by the green shades area in fig. 5.1 (right). As will be discussed in next section, this region corresponds to double-logarithmic contributions, which are usually the dominant terms in our results. Soft emissions at large angle are along the vertical axes (represented by the red line), in the $\theta \sim 1$ region. They can generate at most single-logarithmic corrections. The same is true for hard and collinear emissions, which must be along the diagonal line $z = 1$. Hard and large-angle emissions are not enhanced by any QCD divergence, they can give at most a constant contribution.

5.2 Need for resummation for jet mass distribution

5.2.1 Fixed order calculation

In this section we give a concrete example of the need for resummation for boosted jets observables. We will start by a relatively simple observable, the integrated jet mass

distribution. At leading order in α_s, we take a jet containing only one partonic emission. We will suppose the jet is originated by a quark, as shown in fig. 5.2, but the same reasoning can be applied to a gluon jet. In the collinear jet limit[1] one can write a general distributions for any given observable using the splitting functions, which corresponds to the weight of a given emission. The splitting function for a $q \to q\ g$ emission is

$$p_{gq}(z) \quad = \quad C_F \frac{1+(1-z)^2}{z}. \tag{5.2}$$

For $g \to q\bar{q}$ emissions (resp. $g \to g\ g$) the corresponding splitting function is

$$\begin{aligned} p_{qg}(z) &= T_R((1-z)^2+z^2) \\ p_{gg}(z) &= 2C_A\left(2\frac{1-z}{z}+z(1-z).\right) \end{aligned} \tag{5.3}$$

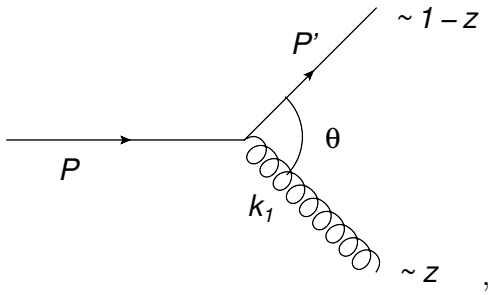


Figure 5.2 — Feynman diagram for $q \to q\ g$

.

For a jet with transverse moment p_t and a radius R, the integrated distribution is the probability that a jet observable $v(p_t, \theta, z, R)$ is smaller than a certain cut v_{cut}. We also have to take into account virtual gluons emissions, i.e. partons that are emitted and then reabsorbed. This step is fundamental to cancel out QCD divergences. Finally, the integrated mass distribution is given by [2]

$$\Sigma_{\text{LO}}(v_{\text{cut}}) = \frac{\alpha_s}{2\pi}\int_0^1 \frac{d\theta^2}{\theta^2}\int_0^1 dz\, p_{gq}(z)\left(\Theta\left[v(p_t,\theta,z,R) < v_{\text{cut}}\right]-1\right). \tag{5.4}$$

For simplicity we work in the fixed coupling approximation, where the strong coupling is considered as constant. The differential v distribution is obtained by taking the derivative of (5.4) wrt v

$$\frac{1}{\sigma}\frac{d\sigma_{\text{LO}}}{dv} = \frac{\alpha_s}{2\pi}\int_0^1 \frac{d\theta^2}{\theta^2}\int_0^1 dz\, p_{gq}(z)\delta(v - v(p_t,\theta,z,R)). \tag{5.5}$$

[1] Large-angle emissions are sligthly more complicated, and we will not consider them for this example. For other cases in this thesis, they are either eliminated thought grooming or are proven to cancel out – a discussion about large-angle emissions is presented in appendix B.2.

[2] For clarity we will write Heaviside step functions as $\Theta[a-b] \equiv \Theta[a>b]$.

Expressions (5.5) and (5.4) can be extended to a jet initiated by a gluon simply by replacing p_{gq} by

$$p_{xg}(z) = \frac{1}{2} p_{gg}(z) + n_f p_{qg}(z), \tag{5.6}$$

where n_f is the number of active quark flavors.

The mass of a jet is given by the squared sum of the (quadri-)momentum of its constituents. For the jet represented in fig. 5.2 we have

$$m^2 = \left(\sum_{i \in \text{jet}} p_i \right)^2 = 2P' \cdot k_1 \sim z(1-z)\theta^2 p_t^2 \tag{5.7}$$

where we consider the collinear limit $\theta \ll 1$ and neglect parton masses. The exact meaning of the emission angle θ depends on the type of collision being considered. For e^+e^- collisions, it is the angle between the emitted parton and the jet direction, so we have $2P' \cdot k_1 = 2z(1-z)(1-\cos\theta)p_t^2$, which gives eq. (5.7) in the collinear limit. For pp collisions it is the separation between the emission and the original parton in the rapidity vs. azimuthal angle $(y \times \phi)$, i.e. $\theta = \sqrt{\Delta\phi^2 + \Delta y^2}$. In this case $2P' \cdot k_1 = 2z(1-z)(\cosh\Delta y - \cos\Delta\phi)p_t^2$, and in the collinear limit we also recover eq. (5.7).

From expression (5.4), we see that the jet mass distribution will be

$$\Sigma_{\text{LO}}(m) = \frac{\alpha_s}{2\pi} \int_0^{R^2} \frac{\mathrm{d}\theta^2}{\theta^2} \int_0^1 dz\, p_{gq}(z) \left(\Theta\left[z(1-z)\theta^2 p_t^2 < m^2 \right] - 1 \right), \tag{5.8}$$

We can integrate over θ, and obtain an additional constraint coming from the $\theta < R$ limit,

$$\Sigma_{\text{LO}}(\rho) = C_F \frac{\alpha_s}{2\pi} \int_0^1 dz\, p_{gq}(z) \log\left(\frac{z}{\rho}\right) \Theta\left[z(1-z) > \rho \right], \tag{5.9}$$

where we introduced the dimensionless variable ρ

$$\rho = \frac{m^2}{R^2 p_t^2}. \tag{5.10}$$

In a boosted jet the transverse momentum p_t is much larger than the mass m, therefore we have $\rho \ll 1$. Under those circumstances we can simplify eq. (5.9), by neglecting constant terms and powers of ρ. In the end we find

$$\Sigma_{\text{LO}}(\rho) = C_F \frac{\alpha_s}{\pi} \left(\frac{1}{2} L_\rho^2 + B_q L_\rho + \mathcal{O}(1) \right), \tag{5.11}$$

with $L_\rho = \log(1/\rho)$ and the B_q is the term coming from hard-collinear emissions. It is given by the non-divergent part of the splitting function

$$B_q = \int_0^1 \mathrm{d}z \left(\frac{p_{qg}(z)}{2C_F} - \frac{1}{z} \right) = -\frac{3}{4}. \tag{5.12}$$

If we consider a gluon jet instead, we would have the same expression but with $C_F \to C_A$ and $B_q \to B_g$, where

$$B_g = \int_0^1 \mathrm{d}z \left(\frac{p_{xg}(z)}{2C_A} - \frac{1}{z} \right) = -\frac{11}{12} + \frac{1}{3} n_f \frac{T_R}{C_A}. \tag{5.13}$$

An important observation about these expressions is that the $(1 - z)$ terms in eq. (5.8) only give power corrections in ρ. We would obtain the same results if we ignore these terms from the beginning.

The same reasoning can be applied to next-to-leading order accuracy, where we consider two parton emissions. Again, we work in the boosted regime $\rho \ll 1$. We have to take into account both virtual gluons emissions, as shown in fig. 5.3.

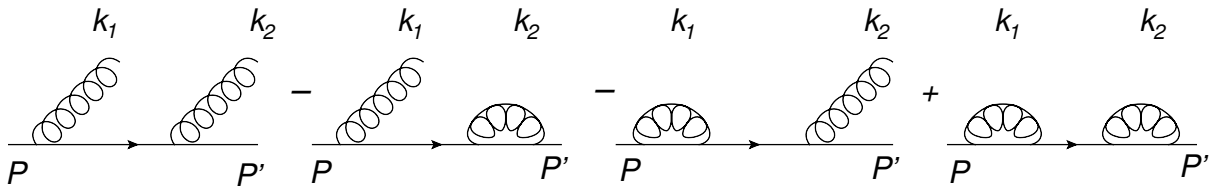


Figure 5.3 — Quark with multiple emissions, both real and virtual, at order α_s^2

In the case with two real emissions the mass is given by

$$m^2 \sim p_t^2 R^2 (x_1 + x_2), \tag{5.14}$$

where $x_i = z_i \theta_i^2$, where each parton carries a fraction of momentum z_i and has an angle of emission θ_i, which we have normalized by R. We are neglecting $(1 - z_i)$ factors, as they will give at most $\mathcal{O}(\rho)$ corrections. We can write the integrated mass distribution directly as a function of ρ

$$\begin{aligned} \Sigma_{\mathrm{NLO}}(\rho) &= \frac{\alpha_s}{2\pi} \int_0^1 \frac{\mathrm{d}\theta_1^2}{\theta_1^2} \int_0^1 \mathrm{d}z_1 \, p_{qg}(z_1) \int_0^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2} \int_0^1 \mathrm{d}z_2 \, p_{qg}(z_2) \\ &\times \left[\Theta(x_1 + x_2 < \rho) - \Theta(x_1 < \rho) - \Theta(x_2 < \rho) + 1 \right]. \end{aligned} \tag{5.15}$$

Since we are only interested in the terms maximally enhanced by logarithms, we can consider that the emissions are strongly ordered in mass. It means we can factorize the mass constraint and replace the term in the second line of (5.15) by

$$[1 - \Theta(x_1 < \rho)] \times [1 - \Theta(x_2 < \rho)], \tag{5.16}$$

and we finally have the expression

$$\Sigma_{\mathrm{NLO}}(\rho) = C_F^2 \frac{\alpha_s^2}{\pi^2} \left(\frac{1}{4} L_\rho^4 + B_q L_\rho^3 + \mathcal{O}(L_\rho^2) \right). \tag{5.17}$$

Observing eqs. (5.11) and (5.17), we can identify logarithmic contributions, in particular $(\alpha_s L_\rho^2)^n$, $\alpha_s^n L_\rho^{2n-1}$, and lower powers of L_ρ, where the exponent n depends on the

order of α_s being considered. In boosted regimes, where $\rho \ll 1$, one can expect these terms of the order of unity $\alpha_s L_\rho \sim 1$ or higher. In this case, the next-to-leading order terms, like $(\alpha_s L_\rho^2)^2$, can be the same order of magnitude (or larger) than the leading order terms $(\alpha_s L_\rho^2)$. Evidently, truncating this calculation at a given order n will not present a good predictive power.

Instead, we use an alternative method and do the resummation of all orders in α_s. For resummed calculations, the accuracy of the results is based on which logarithmic terms are being considered. For example, in (5.17) we call the leading-logarithm (LL) the terms $(\alpha_s L_\rho^2)^2$. They only contain contributions of the soft and collinear region of the Lund diagram, and our calculation is exact at this order. The next-to-leading logarithm (NLL) terms contain for example the hard-collinear region of the Lund diagram. If we relax the strongly-ordered supposition made in (5.15), there would be corrections entering at NLL accuracy (multiple emission corrections).

Note that in fig. 5.3 we do not include the possibility of secondary emissions, i.e. when the initial quark emits a gluons that decay into two partons. In fact, for the jet mass this contribution in only sub-leading[3], except for its contribution to the running coupling, as we discuss in next section.

5.2.2 Resummation for the jet mass

In this section we generalize the results obtained in the previous section to all orders in α_s. The mass of a jet with n gluons in the final state will be:

$$m^2 = p_t^2 R^2 \sum_{j=1}^{n} x_j + \mathcal{O}(z^2), \qquad x_j = z_j \theta_j^2. \tag{5.18}$$

From fixed order calculations, we know that neglecting $\mathcal{O}(z^2)$ terms does not alter the results at leading logarithm.

We again consider all possibilities for virtual emissions. Supposing that gluons are strongly ordered, we can write the mass constraint as a product of constraints over each emission. Then, the integrated distribution is written as a sum over the distributions for all possible number of emissions:

$$\Sigma(\rho) = \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^1 \frac{d\theta_1^2}{\theta_1^2} \int_0^1 dz_1 p_{gq}(z_1) \dots \int_0^1 \frac{d\theta_n^2}{\theta_n^2} \int_0^1 dz_n p_{gq}(z_n) \prod_{i=1}^{n} [1 - \Theta(x_i < \rho)]. \tag{5.19}$$

Expression (5.19) can be rewritten as an exponential

$$\Sigma(\rho) = e^{-R_{\text{plain}}(\rho)}, \tag{5.20}$$

where $R(\rho)$ is the *Sudakov exponent*, it is given by the integrated distribution for one emission with a mass limited by ρ

$$R_{\text{plain}}(\rho) = \frac{\alpha_s}{2\pi} \int_0^1 \frac{d\theta^2}{\theta^2} \int_0^1 p_{gq}(z) \Theta\left(z\theta^2 > \rho\right) = \frac{\alpha_s C_F}{\pi} \left(\frac{1}{2} L_\rho^2 + B_q L_\rho\right). \tag{5.21}$$

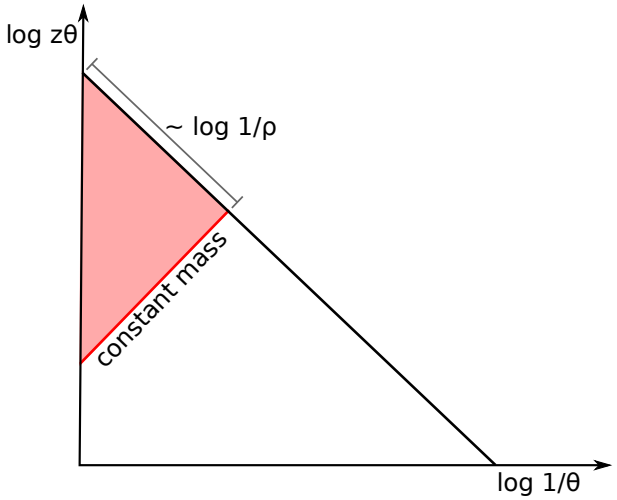


Figure 5.4 — Lund diagram for mass distribution

We can use Lund diagrams, represented in fig. 5.4, as a tools to better understand the physical aspects of eqs. (5.20) and (5.21). The red line in the diagram represents a constant mass $z\theta^2 = \rho$, and the pink shaded region in the diagram represents the large-mass region $z\theta^2 > \rho$. At leading-logarithm, where emissions are strongly-ordered, we can say the mass constraint impose that each real emission is forbidden in the pink region, while virtual emissions are not constrained, so outside the pink region real and virtual contributions cancel each other. The Sudakov exponent at LL is the integration over the pink area, so a double-logarithm. For higher order contributions, e.g. hard and collinear emissions, we know that they must be along the $z = 1$ line. Instead of integration over an area, we integrate over a line, so we lose a logarithmic power, and have a NLL contribution.

Similarly, if we consider particles which are no longer strongly ordered in mass, we would have multiple emissions along the same mass line, so a NLL contribution. The multiple emission correction can be computed by rewriting eq. (5.19) without the strong ordering supposition. This calculation is the same as the trust observable done in chapter 3.2 of ref. [72]. In the end we obtain eq. (5.20) with a multiplicative pre-factor

$$\Sigma(\rho) = \frac{e^{-\gamma_E R'_{\text{plain}}(\rho)}}{\Gamma(1 + R'_{\text{plain}}(\rho))} e^{-R_{\text{plain}}(\rho)}, \tag{5.22}$$

where $R'_{\text{plain}}(\rho)$ is the derivative of $R_{\text{plain}}(\rho)$ wrt L_ρ. Eq. (5.22) remains valid for all additive observables, like N-subjettiness and energy correlation functions.

5.2.3 General comments on resummation

The need for resummation emerges whenever one has multiple scales involved in a measurement. In the case of jet mass, these scales are the mass itself m and the transverse

[3] This is not true for all jet observables, see for example the jet shapes in chapter 7.

momentum of the jet p_t, but this technique can be applied to a variety of calculations in jet physics.

As a general rule, the cross-section for a generic observable v can be expressed as

$$\Sigma(v) = \sum_{\delta} f_{\mathrm{NGL}} \sigma_0^{(\delta)} g_0^{(\delta)}(\alpha_s) e^{\beta} \tag{5.23}$$

$$\beta = L g_1^{(\delta)}(\alpha_s L) + g_2^{(\delta)}(\alpha_s L) + \alpha_s g_3^{(\delta)}(\alpha_s L) + \dots \tag{5.24}$$

where $\sigma_0 = \sum \sigma_0^{(\delta)}$ is the corresponding Born cross-section, $L = \log(1/v)$ is the logarithm of the observable under consideration [4] and we are summing over all possible parton configurations $\delta = q, g$.

Additionally, we may have the contribution of *non-global logarithms* (NGL) [73] . They originate from hard boundaries in phase space, for example a cut in rapidity or a geometric jet boundary, like a limit in jet radius. These terms starts affecting predictions at NLL level, and are usually written as a multiplicative factor f_{NGL}. They are complicated to treat analytically, but in our case they can often be avoided with grooming techniques.

If one only considers the lowest order of resummed calculation, so exclusively the $g_1^{(\delta)}$ terms and leading order in $g_0^{(\delta)}$, it is refereed to as the leading-log (LL) approximation. If one includes the $g_i^{(\delta)}$ terms up to $k+1$ and the coefficient $g_0^{(\delta)}$ up to order α_s^{k-1} it is refereed as next^k-to-leading-order (N^{k}LL). For the moment, state of the art analytical calculations are usually at NNLL accuracy [74, 75, 76], with some cases going up to NNNLL accuracy for specific observables [77].

Finally, to have predictions that are valid all over the available phase space, it is necessary to consider simultaneously fixed-order and resummed results. In order to do that, one needs to do a *matching procedure*. For the jet mass, doing the matching ensures that the cross-section is valid both at low-p_t ($p_t \sim m$), where fixed-order dominates and high-p_t ($p_t \gg m$), where resummation dominates. There are multiple types of matching schemes, but globally, they add the two ends of the calculations and subtract the double counting terms. Commonly used schemes in jet substructure include R-matching [78], where we match the Sudakov factor directly, and the logR-matching, where one matches its logarithm.

5.3 Basic building blocks

In this section we will introduce the *basic building blocks* [69], which are helpful objects for writing several results presented in the following chapters of this thesis. We deal with many integrations over the phase space z vs. θ of parton emissions, with constraints depending on the jet observable being considered. They usually can be written as combination of objects with simpler constraints, which we refer to as basic building blocks.

[4] There are cases where another scale is added to the problem, for example when we are computing the jet mass cross section plus a cut is a jet shape, as we do in chapter 7. In these cases, there are two different logarithms that need to be resummed, so eq. (5.24) needs to be adapted accordingly.

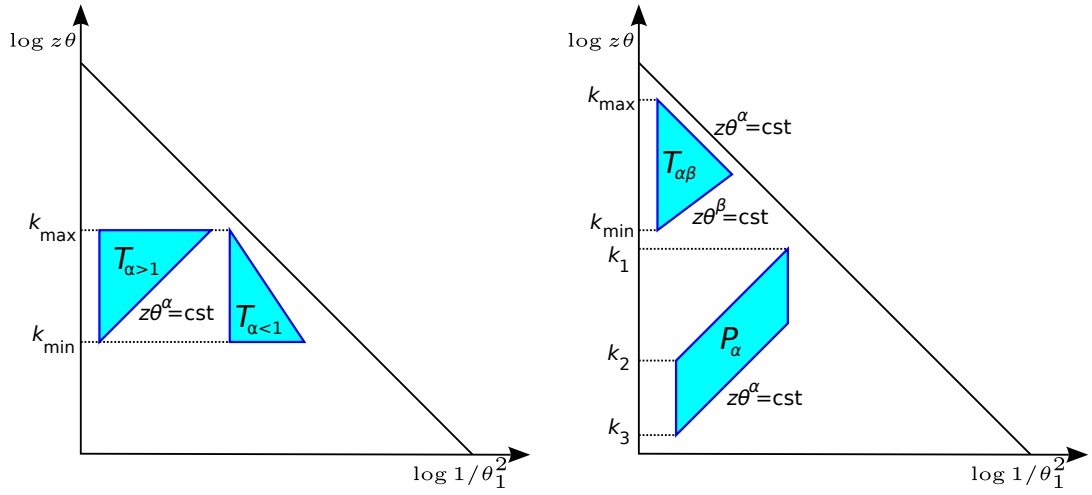


Figure 5.5 — Left: representation of the basic building block used to present our results. It appears in two different forms whether we have $\alpha < 1$ or $\alpha > 1$. Right: two additional fundamental objects built from T_α.

In this section, we provide results for these simple objects, in order to avoid lengthy calculations in the rest of the thesis.

The most basic building block we shall use is the integral over a "triangle" bounded by a maximal angle, a constant $k_t \propto z\theta$ line (upper or lower bound) and a generic line of constant $z\theta^\alpha$, as represented in fig. 5.5. The lines $k_t \sim z\theta$ correspond to a fixed transverse momentum. Expressed as a function of the minimal and maximal k_t scales, the triangle can be written as

$$
\begin{aligned}
&T_\alpha(k_{\max}, k_{\min}; C_R, B_i) \\
&\quad \overset{\alpha\leq 1}{=} \int \frac{d\theta^2}{\theta^2}\, dz\, P(z)\, \frac{\alpha(z\theta)}{2\pi}\, \Theta(\theta < k_{\max})\, \Theta(z\theta > k_{\min})\, \theta(z\theta^\alpha < k_{\max}^\alpha) \qquad (5.25) \\
&\quad \overset{\alpha\geq 1}{=} \int \frac{d\theta^2}{\theta^2}\, dz\, P(z)\, \frac{\alpha(z\theta)}{2\pi}\, \Theta(\theta < 1)\, \Theta(z\theta < k_{\max})\, \theta(z\theta^\alpha > k_{\min}) \qquad (5.26)
\end{aligned}
$$

where we take $C_R = C_F, C_A$ and $B_i = B_q, B_g$ for quark or gluon jets respectively.

We no longer consider the fixed coupling approximation, for this calculation we use the strong coupling described in eq. (2.6) up to 2-loops. We also have to take into account extra sources of logarithm corrections coming from hard hadronic processes. For example, take a hard process involving $q \to q\,g$, where the gluon can further decay collinearly (a secondary emission). This kind of contribution can be taken into account by doing an all-order resummation and it can be represented as an extra factor in the strong coupling $\alpha_s(Q^2)$. In practice, this can be done using the CMW (Bremsstrahlung) scheme [79, 80], in which these corrections are represented by a term proportional to a constant K. At

our accuracy the strong coupling can be written as

$$\begin{aligned}\alpha_s(Q^2) &= \frac{\alpha_s}{1+\alpha_s\beta_0\log\left(\frac{Q^2}{\mu_R^2}\right)} - \alpha_s^2\frac{\beta_1}{\beta_0}\frac{\log\left[1+\alpha_s\beta_0\left(\log\left(\frac{Q^2}{\mu_R^2}\right)\right)\right]}{\left(1+\alpha_s\beta_0\log\left(\frac{Q^2}{\mu_R^2}\right)\right)^2} \\ &+ \alpha_s^2\frac{K}{2\pi\beta_0}\frac{1}{\left(1+\alpha_s\beta_0\log\left(\frac{Q^2}{\mu_R^2}\right)\right)^2},\end{aligned} \tag{5.27}$$

for which,

$$K = \left(\frac{67}{18} - \frac{\pi^2}{6}\right)C_A - \frac{5}{9}n_f. \tag{5.28}$$

In practice, we may need to add a freezing scale to the running coupling, in order to avoid any possible soft divergences as we integrate over the phase space. We impose that the coupling $\alpha_s(k_t)$ is replaced by $\alpha_s(\max(k_t, \tilde{\mu}_{\rm fr}))$, so that for $k_t < \tilde{\mu}_{\rm fr}$, the strong coupling is constant. One usually sets $\tilde{\mu}_{\rm fr} = 1$ GeV. The exact expressions for these integrals depend on the positions of $k_{\rm min}$ and $k_{\rm max}$ compared to $\tilde{\mu}_{\rm fr}$. For $k_{\rm min} > \tilde{\mu}_{\rm fr}$ we find, introducing $L_{\rm min} = \log(1/k_{\rm min})$, $\lambda_{\rm min} = 2\alpha_s\beta_0 L_{\rm min}$ and similar quantities associated with $k_{\rm max}$,

$$\begin{aligned}&T_\alpha(k_{\rm max}, k_{\rm min}; C_R, B_i) \\ &\overset{\alpha\leq 1}{=} \frac{C_R}{2\pi\alpha_s\beta_0^2}\frac{1}{1-\alpha}\Bigg\{\left[(1-\lambda_{\rm max}+2\alpha_s\beta_0 B_i\Theta(\alpha=0))\log\left(\frac{1-\lambda_{\rm max}}{1-\lambda_{\rm min}}\right)+\lambda_{\rm max}-\lambda_{\rm min}\right] \\ &\quad - \frac{\alpha_s\beta_1}{\beta_0}\Bigg[\frac{1}{2}\log^2(1-\lambda_{\rm min}) - \frac{1}{2}\log^2(1-\lambda_{\rm max}) + \frac{1-\lambda_{\rm max}}{1-\lambda_{\rm min}}\log(1-\lambda_{\rm min}) - \log(1-\lambda_{\rm max}) \\ &\qquad + \frac{\lambda_{\rm min}-\lambda_{\rm max}}{1-\lambda_{\rm min}}\Bigg] + \frac{\alpha_s K}{2\pi}\left[\log\left(\frac{1-\lambda_{\rm min}}{1-\lambda_{\rm max}}\right)+\frac{\lambda_{\rm min}-\lambda_{\rm max}}{1-\lambda_{\rm min}}\right]\Bigg\} \\ &\overset{\alpha\geq 1}{=} \frac{C_R}{2\pi\alpha_s\beta_0^2}\frac{1}{\alpha-1}\Bigg\{\left[(1-\lambda_{\rm min})\log\left(\frac{1-\lambda_{\rm min}}{1-\lambda_{\rm max}}\right)+\lambda_{\rm min}-\lambda_{\rm max}\right] \\ &\quad - \frac{\alpha_s\beta_1}{\beta_0}\Bigg[\frac{1}{2}\log^2(1-\lambda_{\rm max}) - \frac{1}{2}\log^2(1-\lambda_{\rm min}) + \frac{1-\lambda_{\rm min}}{1-\lambda_{\rm max}}\log(1-\lambda_{\rm max}) - \log(1-\lambda_{\rm min}) \\ &\qquad + \frac{\lambda_{\rm max}-\lambda_{\rm min}}{1-\lambda_{\rm max}}\Bigg] + \frac{\alpha_s K}{2\pi}\left[\log\left(\frac{1-\lambda_{\rm max}}{1-\lambda_{\rm min}}\right)+\frac{\lambda_{\rm max}-\lambda_{\rm min}}{1-\lambda_{\rm max}}\right]\Bigg\},\end{aligned} \tag{5.29}$$

where the B_i term for $\alpha < 1$ only has to be included if the "triangle" upper edge corresponds to $z = 1$. The expressions above are dominate by double-logarithmic terms, known as leading-logarithm (LL). The next-to-leading logarithms (NLL) contributions are also included, from different sources: terms proportional to B_i come from hard-collinear emissions, the β_1 terms from 2-loop running coupling corrections and the K terms from the CMW scheme corrections.

For $k_{\min} < \tilde{\mu}_{\rm fr}$ but $k_{\max} > \tilde{\mu}_{\rm fr}$, one obtains

$$
\begin{aligned}
&T_\alpha(k_{\max},k_{\min};C_R,B_i) \qquad (5.30)\\
&\overset{\alpha\le 1}{=} \frac{C_R}{2\pi\alpha_s\beta_0^2}\frac{1}{1-\alpha}\Bigg\{\Bigg[(1-\lambda_{\max}+2\alpha_s\beta_0 B_i\Theta(\alpha=0))\log\left(\frac{1-\lambda_{\max}}{1-\lambda_{\rm fr}}\right)+\lambda_{\max}-\lambda_{\rm fr}\Bigg]\\
&\quad-\frac{\alpha_s\beta_1}{\beta_0}\Bigg[\frac{1}{2}\log^2(1-\lambda_{\rm fr})-\frac{1}{2}\log^2(1-\lambda_{\max})+\frac{1-\lambda_{\max}}{1-\lambda_{\rm fr}}\log(1-\lambda_{\rm fr})-\log(1-\lambda_{\max})\\
&\qquad+\frac{\lambda_{\rm fr}-\lambda_{\max}}{1-\lambda_{\rm fr}}\Bigg]+\frac{\alpha_s K}{2\pi}\Bigg[\log\left(\frac{1-\lambda_{\rm fr}}{1-\lambda_{\max}}\right)+\frac{\lambda_{\rm fr}-\lambda_{\max}}{1-\lambda_{\rm fr}}\Bigg]\Bigg\}\\
&\quad+\frac{C_R}{\pi(1-\alpha)}(L_{\min}-L_{\rm fr})\Big[\alpha_s(\tilde{\mu}_{\rm fr})(L_{\min}+L_{\rm fr}-2L_{\max})+2\alpha_{s,\text{1-loop}}(\tilde{\mu}_{\rm fr})B_i\Theta(\alpha=0)\Big]\\
&\overset{\alpha\ge 1}{=} \frac{C_R}{2\pi\alpha_s\beta_0^2}\frac{1}{\alpha-1}\Bigg\{\Bigg[(1-\lambda_{\min})\log\left(\frac{1-\lambda_{\rm fr}}{1-\lambda_{\max}}\right)+\lambda_{\rm fr}-\lambda_{\max}\Bigg]\\
&\quad-\frac{\alpha_s\beta_1}{\beta_0}\Bigg[\frac{1}{2}\log^2(1-\lambda_{\max})-\frac{1}{2}\log^2(1-\lambda_{\rm fr})+\frac{1-\lambda_{\min}}{1-\lambda_{\max}}\log(1-\lambda_{\max})-\frac{1-\lambda_{\min}}{1-\lambda_{\rm fr}}\log(1-\lambda_{\rm fr})\\
&\qquad+\frac{(\lambda_{\max}-\lambda_{\rm fr})(1-\lambda_{\min})}{(1-\lambda_{\max})(1-\lambda_{\rm fr})}\Bigg]+\frac{\alpha_s K}{2\pi}\Bigg[\log\left(\frac{1-\lambda_{\max}}{1-\lambda_{\rm fr}}\right)+\frac{(\lambda_{\max}-\lambda_{\rm fr})(1-\lambda_{\min})}{(1-\lambda_{\max})(1-\lambda_{\rm fr})}\Bigg]\Bigg\},\\
&\quad+\frac{\alpha_s(\tilde{\mu}_{\rm fr})C_R}{\pi(\alpha-1)}(L_{\min}-L_{\rm fr})^2.
\end{aligned}
$$

In that expression, we have introduced $\alpha_{s,\text{1-loop}}(k_t) = \alpha_s/(1-2\alpha_s\beta_0\log(p_tR/k_t))$, the running-coupling at 1-loop, which multiplies the contributions proportional to B_i in the frozen region. This reflects the fact that contributions proportional to $\beta_1 B_i$ and KB_i, coming from the 2-loop corrections to the running of α_s are beyond our accuracy.

And, finally, for $k_{\max} < \tilde{\mu}_{\rm fr}$, one gets

$$
\begin{aligned}
&T_\alpha(k_{\max},k_{\min};C_R,B_i) \qquad (5.31)\\
&\quad=\frac{C_R}{\pi|1-\alpha|}(L_{\min}-L_{\max})\Big[\alpha_s(\tilde{\mu}_{\rm fr})(L_{\min}-L_{\max})+2\alpha_{s,\text{1-loop}}(\tilde{\mu}_{\rm fr})B_i\Theta(\alpha=0)\Big].
\end{aligned}
$$

From this fundamental building block, we can build two derived objects which will be used to describe most of the expressions in the rest of this thesis. The first one is again a triangle bound by a maximal angle, a maximal $z\theta^\alpha$ line and a minimal $z\theta^\beta$ line, see the right plot of fig. 5.5. This can be seen as a sum of two of the above triangles. Again, we can express this new object as a function of the minimal and maximal k_t scales on the maximal-angle side of the triangle, and, assuming $\alpha < \beta$, we get

$$
\begin{aligned}
T_{\alpha\beta}(k_{\max},k_{\min};C_R,B_i) &\overset{\alpha<\beta<1}{=} T_\alpha(k_{\max},k_{\rm med};C_R,B_i)-T_\beta(k_{\min},k_{\rm med};C_R,B_i) &(5.32)\\
&\overset{\alpha<1<\beta}{=} T_\alpha(k_{\max},k_{\rm med};C_R,B_i)+T_\beta(k_{\rm med},k_{\min};C_R,B_i) &(5.33)\\
&\overset{1<\alpha<\beta}{=} T_\beta(k_{\rm med},k_{\min};C_R,B_i)-T_\alpha(k_{\rm med},k_{\max};C_R,B_i), &(5.34)
\end{aligned}
$$

with $k_{\rm med} = k_{\max}^{\frac{\beta-1}{\beta-\alpha}}k_{\min}^{\frac{1-\alpha}{\beta-\alpha}}$.[5]

[5] $T_{0\beta}(k_{\max}=1,k_{\min})$ is related to the radiator given in Appendix A.1 of Ref. [72].

The last object we shall use is a "parallelogram" bounded by a minimal and a maximal angle and two parallel lines of constant $z\theta^\alpha$, assuming here $\alpha > 1$, see again the right plot of fig. 5.5. This is expressed as a function of the maximal k_t scale k_1 (at the minimal angle) and the maximal and minimal k_t scales, k_2 and k_3 at the maximal angle. We can view this as a function of three of our basic triangles

$$P_\alpha(k_1, k_2, k_3; C_R) = T_\alpha(k_1, k_3; C_R, 0) - T_\alpha(k_1, k_2; C_R, 0) - T_\alpha(k_1, k_4; C_R, 0) \tag{5.35}$$

with $k_4 = k_1 k_3/k_2$.

Note that we will often represent the k_t scale by their logarithm, $\log(1/k_t)$ and it is worth keeping in mind that the maximal k_t would correspond to the minimal $\log(1/k_t)$.

Exemple: jet mass Coming back to the jet mass distribution in section 5.2, we can see by the Lund diagram in fig. 5.4, that the mass constraint is equivalent to a triangle limited by lines $z = 1$ and $z\theta^2 = \rho$. Using the basic building blocks formalism, we can write the corresponding Sudakov factor as

$$R_{\text{plain}}(\rho) = T_{02}(1, \rho; C_R, B_i). \tag{5.36}$$

Notice that we recover the expression (5.21) in the fixed coupling approximation, i.e. in the limit $\beta_0, \beta_1, K \to 0$.

Alternative treatment of hard and collinear emissions We would like to add a word of caution about the treatment of B_i terms. A possible drawback of expression (5.21) is that it becomes negative for $L_\rho < -2B_i$. This situation is also present in the basic building blocks, despite the transition point not being so trivial. In order to avoid this problem we propose an alternative way of treating these terms. In expression (5.21) for example, we can replace

$$\frac{L_\rho^2}{2} + B_i L_\rho \to \frac{1}{2}\left(L_\rho + B_i\right)^2, \tag{5.37}$$

and restrict ourselves to $L_\rho + B_i > 0$. At NLL accuracy it does not change our results and it avoids the problem of negative Sudakov factors. It has the disadvantage of adding extra (uncontrolled) terms at the NNLL accuracy. For the basic building blocks, it is equivalent to include a multiplicative factor e^{-B_i} in the $k_{\max}$, for the triangles that include the $z = 1$ line.

Using this alternative approach, the plain mass Sudakov in eq. (5.36) becomes

$$R_{\text{plain}}(\rho) = T_{02}(e^{-B_i}, \rho; C_R, 0). \tag{5.38}$$

This alternative approach has the advantage of having strictly positive Sudakov exponents, which eliminates undesired kinks in the plots, and is specially useful if we are doing a matching procedure. We use this approach in chapters 6, 8 and 9. The original approach, on the other hand, has the advantage of not including terms beyond our control, so it is more robust from a theoretical point of view. We use this approach in chapter 7.

— 6 —
Y-splitter and variations

One way of developing better jet substructure methods is to find interesting combinations between existing tools (prong finders, groomers and jet shapes). In this chapter we explore how the performance of the Y-splitter tagger can be improved by combining it to a variety of grooming techniques, namely modified MassDrop Tagger, trimming or SoftDrop. Recent studies [81] have shown that the particular combination of Y-splitter with trimming can outperform the standard tagging techniques, even thought the Y-splitter alone has a relatively low performance. We study the origin of this increase in performance from a theoretical point of view, using the formalism developed in the previous chapter.

In section 6.1 we present the calculation for the mass distribution of QCD jets with a Y-splitter cut y_{cut}, both fixed order and resummation results at leading logarithm accuracy in $\log \rho$ and $\log y_{\text{cut}}$. In section 6.2, we present the jet mass distribution for Y-splitter and an additional grooming (either trimming or mMDT). In section 6.3, we propose variations for the Y-splitter and in section 6.4 we explore the impact of non-perturbative effects on our predictions and discuss their consequences. Finally, in section 6.5 we present an overall discuss of our findings and conclude.

6.1 Plain Y-splitter

6.1.1 Leading-order calculation

We start by computing the leading-order result for the jet mass distribution for QCD jets that are tagged by Y-splitter, presented in (4.2). In order to generate leading logarithmic contributions it is sufficient to consider contributions from soft and collinear gluon emissions from a hard parton.

Therefore at leading order in QCD (order α_s) we have to consider a jet made up of a hard quark or gluon and a single accompanying soft and collinear quark or gluon emissions. Here we shall explicitly consider the case of quark jets to begin with, but it is trivial to obtain the corresponding results for gluon initiated jets from the ones we derive below.

Let us write the four-momenta of the particles as

$$\begin{aligned} p &= p_t\,(1,1,0,0)\,, \\ k &= \omega_t\,(\cosh y, \cos\phi, \sin\phi, \sinh y)\,, \end{aligned} \tag{6.1}$$

where p is the four-momentum of the hard quark, written in terms of its transverse momentum p_t wrt the beam and where without loss of generality we can set its rapidity wrt the beam to zero. Likewise ω_t is the transverse momentum of the emitted soft gluon, with rapidity y and azimuthal angle ϕ. In the soft and collinear limit we have $\omega_t \ll p_t$ and $\theta^2 = (y^2+\phi^2) \ll 1$.

Let us first study the jet mass distribution with a cut on d_{12}/m^2 (see eq. (4.1)), with m being the jet mass. In the soft and collinear approximation $d_{12} = \omega_t^2\theta^2$ while $m^2 = \omega_t p_t \theta^2$ so that we apply the Y-splitter cut on the quantity $x = \omega_t/p_t$ i.e. the transverse momentum fraction of the gluon, such that $x > y_{\rm cut}$. The calculation for the jet mass distribution is then simple to write down[1]

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\rm LO,soft-coll.} = \frac{C_F\alpha_s}{\pi}\int_0^1 \frac{dx}{x}\frac{d\theta^2}{\theta^2}\,\delta\left(\rho - x\theta^2\right)\Theta\left(x > y\right), \tag{6.2}$$

where, as we are working in the LL accuracy, we have taken a fixed-coupling approximation. In writing (6.2), we have implicitly normalized all angles to R so that θ runs up to 1 (instead of up to R) and all R dependence that arises at our accuracy is incorporated into our definition of $\rho = m^2/(p_t R)^2$. Note that eq. (6.2) is written for quark jets. One can easily extrapolate this, and the following formulae, to gluon jets by replacing C_F by C_A.

We can easily integrate (6.2) to obtain

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\rm LO,soft-coll.} = \frac{C_F\alpha_s}{\pi}\left(\ln\frac{1}{y}\Theta\left(y>\rho\right) + \ln\frac{1}{\rho}\Theta\left(\rho > y\right)\right). \tag{6.3}$$

The result above is identical to results obtained for the MassDrop Tagger (and the Mass-Drop Tagger (mMDT)) as well as for pruning [12]. It reflects that at this order the action of Y-splitter, in the small ρ limit, is to remove a logarithm in ρ and replace it with a (smaller) logarithm in y. This implies a reduction in the QCD background at small ρ relative to the plain jet mass result. For $\rho > y$, the cut is redundant and we return to the case of the plain QCD jet mass.

It is also straightforward to extend the soft approximation by considering hard-collinear corrections. To include these effects one simply makes the replacement $\frac{1}{x} \to \frac{1+(1-x)^2}{2x}$ i.e. includes the full p_{gq} splitting function. It is also simple to include finite y corrections in the above result by inserting the proper limits of integration that are obtained from the Y-splitter condition when one considers hard collinear rather than soft gluon emission.

[1] In order to lighten notation we will omit the subscript "cut" in $y_{\rm cut}$ when there is no room for ambiguity.

The Y-splitter condition is satisfied for $y/(1-y) < x < 1/(1+y)$ and we obtain the result, for $\rho < y/(1+y)$:

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\text{LO,coll.}} = \frac{C_F\alpha_s}{\pi}\left(\ln\frac{1}{y} - \frac{3}{4}\left(\frac{1-y}{1+y}\right)\right). \tag{6.4}$$

This result is again identical to the case of (m)MDT with the $y_{\text{cut}} > y$ condition.

6.1.2 NLO result and all-orders form

In this section we compute the next-to-leading order result in the soft and collinear limit, before extending this result to all orders in the next section.

Thus we need to consider the case of two real emissions off the primary hard parton as well as a real and a virtual emission, also treated in the soft and collinear limit. We shall work in the classical independent emission approximation which is sufficient to obtain the leading logarithmic result for jet mass distributions.

We consider a jet made up of a primary hard parton and two soft gluons with four-momenta k_1 and k_2. When the jet is declustered one requires the Y-splitter cut to be satisfied for the jet to be tagged. There are two distinct situations that arise at this order: the situation where the largest k_t gluon passes the Y-splitter cut as well as sets the mass of the jet or the situation where the largest k_t gluon passes the Y-splitter cut so the jet is accepted but the jet mass is set by a lower k_t emission.

For the one-real, one-virtual contributions the situation is the same as that for the leading-order calculation i.e. the real emission both passes the Y-splitter cut and sets the mass.

Let us assume that the jet mass is set by emission k_1 with energy fraction x_1 and which makes an angle θ_1 with the jet axis or equivalently the hard parton direction, with $x_1, \theta_1 \ll 1$. For simplicity, it is useful to introduce for every emission k_i, the respective quantities

$$\kappa_i \equiv x_i\theta_i, \qquad \rho_i \equiv x_i\theta_i^2, \tag{6.5}$$

related to the transverse momentum (k_t scale) of emission k_i wrt the jet axis and the contribution of emission k_i to the jet mass. We can then write

$$\begin{aligned}\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,soft}-\text{coll.}} = &\left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2\, \delta(\rho-\rho_1)\Big[\Theta(\kappa_1>\kappa_2)\,\Theta(x_1>y)\,\Theta(\rho_2<\rho)\\ &+\Theta(\kappa_2>\kappa_1)\,\Theta(\kappa_2>\rho y)\,\Theta(\rho_2<\rho) - \Theta(x_1>y)\Big],\end{aligned} \tag{6.6}$$

where we introduced the notation

$$d\Phi_2 \equiv \frac{dx_1}{x_1}\frac{dx_2}{x_2}\frac{d\theta_1^2}{\theta_1^2}\frac{d\theta_2^2}{\theta_2^2}, \tag{6.7}$$

for the two-gluon emission phase space in the soft-collinear limit.

The first line within the large parenthesis expresses the condition that the gluon which sets the mass has the higher k_t i.e. $\kappa_1 > \kappa_2$ as well as satisfies the Y-splitter constraint on the higher k_t gluon $\kappa_1^2/\rho_1 = x_1^2\theta_1^2/(x_1\theta_1^2) = x_1 > y$. The emission k_2 cannot dominate the jet mass by assumption, which gives rise to the veto condition $\rho_2 < \rho$. The first term on the second line within the parenthesis expresses the condition that the gluon k_1 now has lower k_t than emission k_2. Emission k_2 passes the Y-splitter cut $\kappa_2^2/\rho > y$, where ρ is the mass set by emission k_1. The final term on the last line, with negative sign, is the contribution where emission k_2 is virtual.

For the term on the first line we make the replacement $\Theta(\kappa_1 > \kappa_2) = 1 - \Theta(\kappa_2 > \kappa_1)$. These two terms can be combined with the virtual corrections and the first term of the second line, respectively, to give

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,soft}-\text{coll.}} = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \left[\int d\Phi_2\, \delta(\rho_1 - \rho)\, \Theta(x_1 > y)\, (\Theta(\rho_2 < \rho) - 1) + \int d\Phi_2\, \delta(\rho - \rho_1)\, \Theta(\kappa_2 > \kappa_1)\, \Theta(\rho_2 < \rho)\, [\Theta(\kappa_2 > y\rho) - \Theta(x_1 > y)]\right]. \quad (6.8)$$

The fundamental reason for writing the result in the above form is to separate what we expect to be the leading logarithmic contribution in the first line from sub-leading contributions which involve a higher k_t emission giving a smaller contribution to the jet mass than emission k_1. Hence we anticipate that the term in the second line in eq. (6.8) will produce results that are beyond our accuracy, in the limit of small ρ. On explicit calculation of this term one gets, for $\rho < y$,

$$\left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2\, \delta(\rho_1 - \rho)\, \Theta(\kappa_2 > \kappa_1)\, \Theta(\rho_2 < \rho)\, [\Theta(\kappa_2 > y\rho) - \Theta(x_1 > y)] = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{2\rho}\left(\ln\frac{1}{\rho}\ln^2\frac{1}{y} - \ln^3\frac{1}{y}\right) = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{2\rho}\ln\frac{y}{\rho}\ln^2\frac{1}{y}. \quad (6.9)$$

The above result implies that in the $\rho \to 0$ limit there are at best single logarithmic (in ρ) contributions to the integrated jet mass distribution from the second line of eq. (6.8). Using $\Theta(\rho_2 < \rho) - 1 = -\Theta(\rho_2 > \rho)$, the first line of eq. (6.8) gives

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,LL}} = -\left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2 \Theta(x_1 > y)\, \delta(\rho - \rho_1)\, \Theta(\rho_2 > \rho), \quad (6.10)$$

which produces the leading logarithmic (LL) corrections we require. Upon evaluation, it produces for $\rho < y$,

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,LL}} = -\left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{2}\ln\frac{1}{y}\ln^2\frac{1}{\rho}, \quad (6.11)$$

which has the structure of the leading-order result multiplied by a double logarithmic term in ρ. We note that for $\rho > y$ the Y-splitter cut becomes redundant and one returns to the result for the standard plain jet mass distribution, computed in section 5.2. We

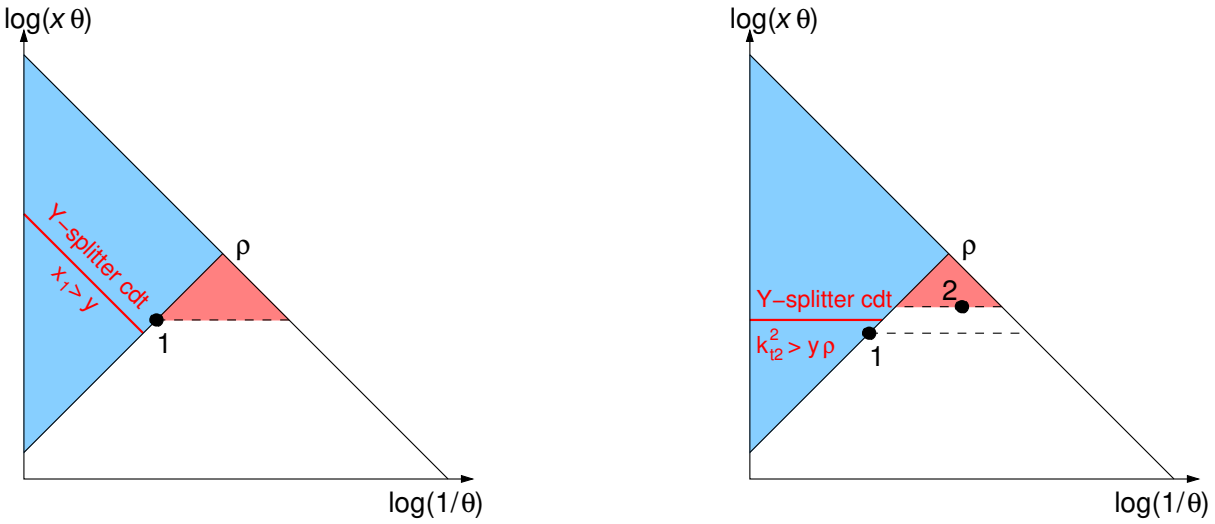


Figure 6.1 — Lund diagrams for the two possible configurations in resummed mass distribution: emission that dominates the jet mass also has the largest k_t (left) and emission with larger k_t does not dominates mass.

recall that by "leading logarithmic (LL) accuracy" we mean that we only keep the terms that are maximally enhanced in $\ln \rho$.

The result in eq. (6.11) has a simple physical interpretation. The largest k_t emission which sets the mass comes with a cut on its energy precisely as at leading order which, produces an $\alpha_s \ln \frac{1}{y}$ behavior. Emission k_2 on the other hand is subject to a veto condition such that $\rho_2 < \rho$. After cancellation against virtual corrections one obtains an $\alpha_s \ln^2 \frac{1}{\rho}$ behaviour from this emission, exactly as for the leading-order contribution to the integrated plain jet mass distribution. Based on this we can expect that at all orders, to leading-logarithmic accuracy, one ought to multiply the leading-order (LO) result by a double logarithmic Sudakov suppression factor like that for the plain jet mass. The leading-order result then appears as a single-logarithmic prefactor in front of a resummed double-logarithmic Sudakov exponent, as we shall see in the following.

Lastly we note that the full result of our calculation of eq. (6.6) can be written in the form

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,soft}-\text{coll.}} = \left(\frac{C_F \alpha_s}{\pi}\right)^2 \frac{1}{2\rho}\left(-\ln\frac{1}{y}\ln^2\frac{1}{\rho} + \ln\frac{y}{\rho}\ln^2\frac{1}{y}\right), \tag{6.12}$$

where the first term on the right side contains the leading logarithms in ρ while the second term is subleading in ρ (being purely single logarithmic), although it is enhanced by logarithmic terms in y.

6.1.3 All-orders resummation for $\log(\rho)$

Eqs. (6.10), (6.11) can be easily generalized to all orders. To LL accuracy, one can generalize the reasoning for the NLO case and one has to consider only the situation

where the highest k_t emission dominates the jet mass. A jet-mass veto then applies to all other real emissions. This situation is depicted in the Lund diagram to the left in fig. 6.1. The emission denoted with a black dot sets the jet mass i.e. satisfies $\rho_1 \equiv x_1\theta_1^2 = \rho$. The blue shaded region corresponds to emissions that give a contribution to the mass $x\theta^2 > \rho$ and hence are vetoed. Considering these emissions to be emitted according to an independent emission pattern the veto condition gives a Sudakov suppression factor represented by the blue shaded area in the figure, which is identical to the suppression factor obtained for the plain jet mass at leading-logarithmic accuracy. In addition to this, emissions with a higher transverse momentum which set a lower mass than ρ are also vetoed since we assumed that the emission which sets the mass is the highest k_t emission. This is denoted by the red shaded area in the figure but as this region produces only terms that are sub-leading in ρ we shall not consider it for the moment. Finally, we also have to consider the Y-splitter constraint which for this configuration corresponds to $x_1 > y$ where the line $x = y$ is shown in red in the figure. The all-orders fixed-coupling result from this configuration, which captures the leading double-logarithms in ρ, is

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\text{LL}} = \frac{C_F\alpha_s}{\pi}\ln\frac{1}{y}\times\exp\left[-\frac{C_F\alpha_s}{2\pi}\ln^2\frac{1}{\rho}\right], \quad (\text{for } \rho < y), \tag{6.13}$$

while for $\rho > y$ the result is that for the plain mass distribution. Eq. (6.13) corresponds to the result reported already in [81]. Note that a similar result is obtained also for the case of Y-pruning in the regime $\alpha_s \ln\frac{1}{z_{\text{cut}}}\ln\frac{1}{\rho} \ll 1$ (see eq. 5.10b of [12]).

It is simple to include running-coupling corrections both in the prefactor i.e. those associated to the emission which sets the mass as well as in the Sudakov exponent. Likewise hard-collinear emissions may be treated by using the full splitting function in the prefactor and the Sudakov exponent, yielding the modified leading logarithmic approximation. Lastly we can also include finite y corrections into the prefactor as they may be of numerical significance since they occur already at leading order (see eq. (6.4)).

The general result, for $\rho < y$ then reads

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\text{LL}} = \int_{\frac{y}{1+y}}^{\frac{1}{1+y}} dx_1\, p_i(x_1)\,\frac{\alpha_s(x_1\rho)}{2\pi}e^{-R_{\text{plain}}(\rho)}, \tag{6.14}$$

where $p_i = p_{gq}$ for a quark jet, and $p_i = p_{xg}$ for a gluon jet. The Sudakov exponent ("radiator") is the same as for the plain jet mass case, studied in previous section,

$$R_{\text{plain}}(\rho) = \int\frac{d\theta^2}{\theta^2}dx\, p_i(x)\,\frac{\alpha_s(x^2\theta^2)}{2\pi}\Theta\left(x\theta^2 > \rho\right) = T_{02}(e^{-B_i},\rho;C_R,0), \tag{6.15}$$

where we are using the basic building blocks defined in (5.32). We take $C_R = C_F$ and $B_i = B_q$ for a gluon jet, and $C_R = C_A$ and $B_i = B_g$ for a gluon jet. Note that here and henceforth we shall only specify the transition points in a small y approximation. Thus the exact transition point $\rho = y/(1+y)$ will be approximated by $\rho = y$. In the above expression and the remainder of the text, the arguments of the running coupling have to be understood as factors of $p_t^2R^2$.

6.1.4 All-orders resummation for $\log(y)$

In the present case, if y becomes small enough, we can also perform an all-order resummation of the logarithms of $1/y$. Such terms, which are formally at the level of sub-leading logarithms in ρ, were already identified in our fixed-order NLO calculation, see eq. (6.12). In order to resum them we will have to consider also situations where the highest transverse momentum emission does not set the jet mass. To write a general resummed result it is convenient to return to the Lund diagrams in fig. 6.1. The figure on the left denotes, as we stated before, the situation where the highest transverse momentum emission both passes the Y-splitter constraint and also sets the mass, with a veto on higher mass emissions. Now however we also account for the contribution from the red shaded region that corresponds to an additional veto on emissions with a higher transverse momentum than the emission which sets the mass. The figure on the right denotes a second situation where there is an emission k_2 which is the highest k_t emission i.e. $\kappa_2 > \kappa_1$. The red shaded region now denotes the additional veto on any emissions with transverse momentum greater than κ_2. The blue region as before corresponds to a veto on emissions with larger mass than $\rho = \rho_1$ and the Y-splitter condition now corresponds to $\kappa_2^2 > \rho y$ where the line $x^2\theta^2 = \rho y$ is shown in the figure.

Taking both the above described situations into account one can write the result as (for now we ignore finite y effects to which we shall return)

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\mathrm{LL+LL}_y} = \int_\rho^1 dx_1\, p_i(x_1)\,\frac{\alpha_s(\rho x_1)}{2\pi} e^{-R_{\mathrm{plain}}(\rho)} \left[\Theta(x_1 > y) e^{-R_{k_t}(\kappa_1,\rho)} + \right. \tag{6.16}$$
$$\left. + \int \frac{d\theta_2^2}{\theta_2^2} dx_2\, p_i(x_2)\,\frac{\alpha_s(\kappa_2^2)}{2\pi}\Theta\left(\rho_2 < \rho\right)\Theta\left(\kappa_2 > \rho x_1\right)\Theta\left(\kappa_2 > \rho y\right) e^{-R_{k_t}(\kappa_2,\rho)}\right],$$

where the first term in large brackets comes from the Lund diagram on the left and the second term from that on the right. Note that R_{k_t} is also a Sudakov type exponent defined as

$$R_{k_t}(\kappa,\rho) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\,\frac{\alpha_s(x^2\theta^2)}{2\pi}\Theta\left(x\theta^2 < \rho\right)\Theta\left(x\theta > \kappa\right)$$
$$= 2T_{0,2}\left(\sqrt{\rho e^{-B_i}}, k_t; C_R, 0\right). \tag{6.17}$$

which arises from a veto on transverse momentum of emissions above the scale k_t while at the same time imposing that the mass of the vetoed emissions is lower than ρ, as required for taking into account the red shaded regions in the Lund diagrams of fig. 6.1.

This expression can be simplified quite significantly: one first splits the second line into a contribution with $x_1 > y$ and a contribution with $\rho < x_1 < y$. After integration over x_2 and θ_2 and combining the contribution from $x_1 > y$ with the first line of (6.16) one can write the final result as

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\mathrm{LL+LL}_y} = e^{-R_{\mathrm{plain}}(\rho)}\left[\int_{\frac{y}{1+y}}^{\frac{1}{1+y}} dx_1\, p_i(x_1)\frac{\alpha_s(x_1\rho)}{2\pi} + \left(1 - e^{-R_{k_t}(\sqrt{\rho y},\rho)}\right)\int_\rho^y dx_1\, p_i(x_1)\frac{\alpha_s(x_1\rho)}{2\pi}\right], \tag{6.18}$$

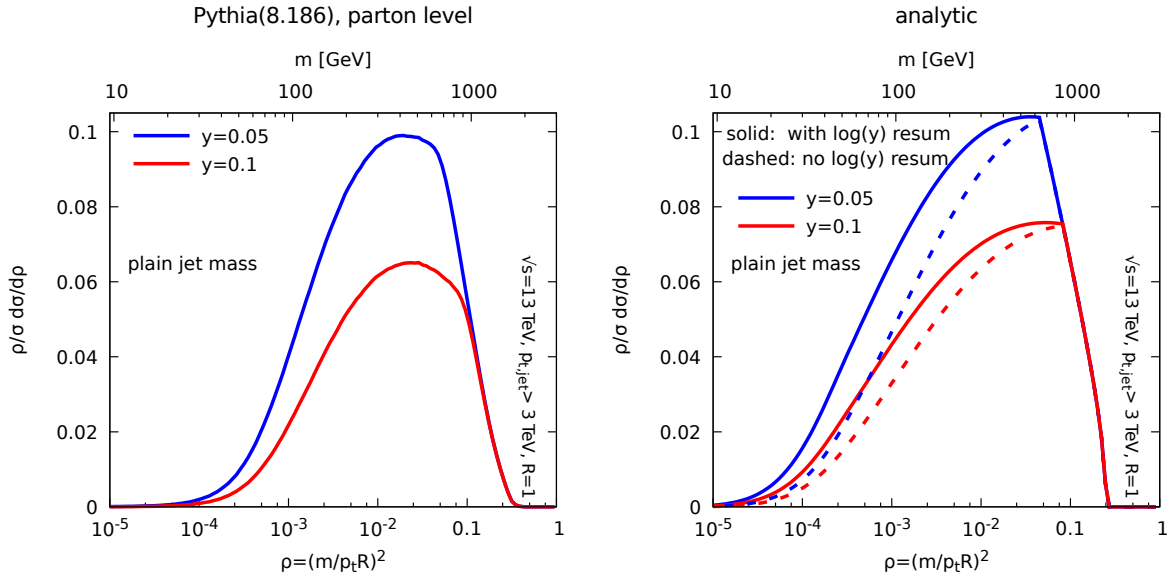


Figure 6.2 — Jet mass distribution with Y-splitter for $y = 0.05, 0.1$, Pythia simulations (left) and analytical calculation (right). For analytical results, solid lines include $\log y$ resummation and dashed lines do not.

where we have restored the finite y corrections in the leading contribution (first term). The correction term one thus obtains relative to (6.14) has a prefactor proportional to $\alpha_s \ln \frac{y}{\rho}$ multiplied by a Sudakov-like factor, starting at order α_s and resumming terms of the form $\alpha_s^n \ln^{2n} \frac{1}{y}$. This is consistent with the result obtained at NLO in eq. (6.12).

6.1.5 Comparison to Monte Carlo

In order to validate our analytic results, we have compared them to Monte-Carlo simulations. We have used Pythia (v8.186) [35] with the 4C tune [82] to generate $qq \to qq$ events at parton level with $\sqrt{s} = 13$ TeV. Jets are reconstructed with the anti-k_t algorithm with $R = 1$ as implemented in FastJet [38, 39] and we require that the jets satisfy $p_t > 3$ TeV and rapidity $|y| < 4$. Unless explicitly stated otherwise, we use this same setup to all Monte Carlo simulations involving Y-splitter tagger in this chapter.

The comparison to our analytic calculations is shown in fig. 6.2 with Pythia on the left and our results on the right. All our results include the contribution from the full splitting function including hard-collinear effects to the Sudakov exponent, and use a 1-loop approximation for the running of the strong coupling with $\alpha_s(m_Z) = 0.1383$. This value matches the one used in Pythia for the final-state shower. Furthermore, the plot with our analytic results includes both the leading logarithmic result described in eq. (6.14) (dashed curves) as well as the result augmented to include resummation of double logarithms in y, eq. (6.16) (solid curves) for two values of y. We note firstly the good overall agreement with Monte Carlo results for both variants of the analytics, which indicates that our modified leading-logarithmic results successfully explain the performance of Y-

splitter on QCD background jets. The observed differences between analytics and Monte Carlo can arise due to different treatment of next-to–leading logarithmic effects such as those due to soft emissions at large angles and initial state radiation included in the Monte Carlo studies but left out of our resummed calculations.

It is noteworthy that the $\ln y$ resummation although a visible effect, is fairly modest. The essential dependence of the results on y is already captured by the leading-logarithmic resummation of eq. (6.14).

6.2 Y-splitter with grooming

In this section we shall consider the Y-splitter method supplemented with grooming procedures, specifically the modified MassDrop Tagger (equivalently SoftDrop $\beta = 0$) and trimming. The effectiveness of applying grooming subsequent to the use of Y-splitter on a jet has been clearly demonstrated in the Monte Carlo studies carried out in [81]. It was shown that while Y-splitter alone has a very poor signal efficiency (similar to that for an ungroomed jet which is severely affected by ISR and Underlying Event), grooming makes a considerable difference to the performance of Y-splitter on signal jets. On the other hand we have already seen that on QCD background jets Y-splitter gives a double-logarithmic Sudakov type factor multiplying a single logarithmic prefactor, which implies a desirable strong suppression of background. Using Y-splitter with grooming did not significantly alter the performance of Y-splitter on background jets, in the sense that applying a grooming procedure after one imposes a Y-splitter cut does not alter the double-logarithmic Sudakov behavior for the QCD background.

This fact coupled with the great improvement seen in signal efficiency resulted in Y-splitter+grooming outperforming other standard taggers for signal significance at high p_t . Here we seek to understand from a first principles viewpoint why grooming does not appear to strongly impact the basic performance of Y-splitter on background. We start by studying Y-splitter with trimming in the next sub-section, which was the combination employed in [81].

6.2.1 Y-splitter with trimming: fixed-order results

To study the impact of trimming on Y-splitter, we shall consider taking a jet accepted by Y-splitter and then apply trimming to it (see definition in section 4.1). It is important to highlight that it is crucial to apply the Y-splitter condition on the plain jet and apply grooming afterwards. We show in appendix A that applying grooming first and then imposing the Y-splitter condition on the groomed jet leads to a smaller suppression of the QCD background.

We shall set the f_{cut} parameter of trimming to be equal to the parameter y of Y-splitter, a choice that will become clear presently[2]. We first note that, at leading order,

[2]If we keep into account finite y corrections, we should actually use $f_{\text{cut}} = y/(1+y)$, which is what we have done in practice in our Monte Carlo simulations.

for a soft emission to pass Y-splitter it must have an energy fraction $x > y$. When one applies trimming afterwards such an emission is unaffected as, with our choice of $f_{\rm cut}$ trimming removes only emissions with $x < y$. Thus at leading-order Y-splitter with trimming trivially returns the same result as Y-splitter alone.

We shall now examine the role of trimming at the NLO level. Let us consider that the mass of the final jet *after* grooming is set by an emission k_1. In other words, we first impose the Y-splitter cut on the plain jet and, if it passes, we compute the trimmed jet mass.

At order α_s^2 we have to consider both a second real emission k_2 as well as a virtual gluon contribution. Cosidering the leading logarithm limit (soft and collinear emissions), the mass distribution can be written as

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\rm NLO,soft-coll} = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2 \left(\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4\right) \tag{6.19}$$

with

$$\mathcal{I}_1 = \delta\left(\rho - \rho_1\right)\Theta\left(\kappa_1 > \kappa_2\right)\Theta\left(\frac{\kappa_1^2}{\rho_1+\rho_2} > y\right)\Theta\left(\rho_2 < \rho\right)\Theta_2^{\rm in}, \tag{6.20}$$

$$\mathcal{I}_2 = \delta\left(\rho - \rho_1\right)\Theta\left(\kappa_1 > \kappa_2\right)\Theta\left(\frac{\kappa_1^2}{\rho_1+\rho_2} > y\right)\Theta_2^{\rm out}, \tag{6.21}$$

$$\mathcal{I}_3 = \delta\left(\rho - \rho_1\right)\Theta\left(\kappa_2 > \kappa_1\right)\Theta\left(\frac{\kappa_2^2}{\rho_1+\rho_2} > y\right)\Theta\left(\rho_2 < \rho\right)\Theta_1^{\rm in}, \tag{6.22}$$

$$\mathcal{I}_4 = -\delta\left(\rho - \rho_1\right)\Theta(x_1 > y), \tag{6.23}$$

where we introduced the shorthand notations $\Theta_i^{\rm in}$ and $\Theta_i^{\rm out}$ to represent that emission k_i is respectively left in or removed by trimming. We recall the condition for an emission to be removed by trimming is

$$\Theta_i^{\rm out} = 1 - \Theta_i^{\rm in} = \Theta(x_i < y)\,\Theta(\theta_i > r), \tag{6.24}$$

with $r \equiv \frac{R_{\rm trim}}{R}$ and $R_{\rm trim}$ the trimming radius.

Let us detail the physical origin of these different contributions. The contribution $\mathcal{I}_1$ contains the conditions on $x_1, x_2, \theta_1, \theta_2$ such that k_1 sets the mass ($\rho = \rho_1$) and has the higher transverse momentum, $\kappa_1 > \kappa_2$. It also contains the condition for the Y-splitter cut to pass $\kappa_1^2/(\rho_1+\rho_2) > y$, and the condition that k_2 is left in by trimming represented by $\Theta_2^{\rm in}$. Lastly it contains the veto on the mass $\rho > \rho_2$ such that emission k_2 cannot set the mass. Likewise $\mathcal{I}_2$ contains the conditions that emerge when k_2 is removed by trimming which itself corresponds to the condition $\Theta_2^{\rm out}$. For both $\mathcal{I}_1$ and $\mathcal{I}_2$, the Y-splitter condition implies $x_1 > y$ and therefore guarantees that emission k_1 is left in by trimming. These configurations reproduce the leading-logarithmic terms of the pure Y-splitter cut, and also generate subleading contributions coming from the region where k_2 is removed by trimming and has $\rho_2 > \rho$, One can easily see this by inserting $1 = \Theta(\rho_2 > \rho) + \Theta(\rho_2 < \rho)$ in $\mathcal{I}_2$.

On the other hand, $\mathcal{I}_3$ represents the situation when k_1 is the lower transverse momentum emission and sets the mass. In this case, the Y-splitter condition implies $x_2 > y$, i.e. emission k_2 is kept by trimming, and we thus have to impose that $\rho_2 < \rho_1$. We also have to impose that emission k_1 is left in by trimming corresponding to Θ_1^{in}. Lastly $\mathcal{I}_4$ corresponds to the situation when k_2 is virtual and all that is required is for k_1 to pass the Y-splitter cut.

A comment is due about the Y-splitter condition used in the above formulae (6.20) –(6.22). In situations where emission k_1 dominates the mass even though emission k_2 is not groomed away it is possible, at leading logarithmic accuracy, to replace $\rho_1 + \rho_2$ in the denominator of the Y-splitter constraints by $\rho = \rho_1$. Specifically this applies to the $\mathcal{I}_1$ and $\mathcal{I}_3$ terms above. We have however chosen to treat the Y-splitter constraint exactly in all terms since in the term involving $\mathcal{I}_2$, where emission k_2 is groomed away, there is no condition on ρ_2 requiring it to be less than ρ. Retaining the exact Y-splitter constraint in all terms proves convenient for reorganising and combining various contributions as we shall do below, while only differing from the leading-logarithmic simplification by subleading terms which we do not control.

In order to highlight that the use of grooming techniques does not drastically modify the background rejection obtained with Y-splitter alone, it is interesting to express the calculations as grooming-induced corrections to those already carried out for Y-splitter. To this end, in the contribution involving $\mathcal{I}_1$ let us replace Θ_2^{in} with $1 - \Theta_2^{\text{out}}$ which splits the contribution from $\mathcal{I}_1$ into two pieces $\mathcal{I}_1 = \mathcal{I}_1^{\text{full}} - \mathcal{I}_1^{\text{out}}$. The contribution from $\mathcal{I}_1^{\text{full}}$, where we can use $\rho_1 + \rho_2 \approx \rho_1$ in the Y-splitter condition, is just the same as the corresponding leading term for the pure Y-splitter case. It can be combined with the virtual term $\mathcal{I}_4$ (which is also identical to the pure Y-splitter case) to produce the NLO leading-logarithmic result we reported earlier for Y-splitter, cf. (6.13) and (6.14). We can apply a similar procedure for the term $\mathcal{I}_3$ such that $\mathcal{I}_3 = \mathcal{I}_3^{\text{full}} - \mathcal{I}_3^{\text{out}}$, where $\mathcal{I}_3^{\text{full}}$ is the contribution to the pure Y-splitter case from the situation that the the highest k_t emission passes Y-splitter but does not set the jet mass. Recall that this configuration produces only terms beyond our formal leading-logarithmic accuracy (cf. the second term in eq. (6.18)). The remaining terms, all involving Θ_2^{out}, constitute the trimming-induced corrections to Y-splitter. It is then useful to write the result in the following form:

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,soft}-\text{coll}} = \frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,YS}} + \mathcal{F}^{\text{trim,a}} + \mathcal{F}^{\text{trim,b}} \tag{6.25}$$

where $\frac{1}{\sigma}\frac{d\sigma}{d\rho}^{\text{NLO,YS}}$ is the pure Y-splitter result given by (6.18), and we defined

$$\mathcal{F}^{\text{trim,a}} = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2 \delta\left(\rho - \rho_1\right) \Theta\left(\kappa_1 > \kappa_2\right) \Theta\left(\frac{\kappa_1^2}{\rho_1 + \rho_2} > y\right) \left[1 - \Theta\left(\rho_2 < \rho\right)\right] \Theta_2^{\text{out}}, \tag{6.26}$$

which arises from combining the contributions from $\mathcal{I}_2$ and $-\mathcal{I}_1^{\text{out}}$ and

$$\mathcal{F}^{\text{trim,b}} = -\left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2 \delta\left(\rho - \rho_1\right) \Theta\left(\kappa_2 > \kappa_1\right) \Theta\left(\frac{\kappa_2^2}{\rho_1 + \rho_2} > y\right) \Theta\left(\rho_2 < \rho\right) \Theta_1^{\text{out}}, \tag{6.27}$$

which arises from the $-\mathcal{I}_3^{\text{out}}$ term.

At this stage, within our accuracy we can replace $\rho_1 + \rho_2$ by ρ_2 in (6.26) and by ρ_1 in (6.27). We can then express the constraints in (6.26) in the form

$$\delta(\rho - \rho_1)\,\Theta\left(\frac{\rho x_1}{x_2} > \rho_2\right)\Theta\left(\frac{\rho x_1}{y} > \rho_2\right)\left[1 - \Theta(\rho_2 < \rho)\right]\Theta_2^{\text{out}}. \tag{6.28}$$

We note that the above implies the condition $x_1 > y$ and Θ_2^{out} imposes the condition $x_2 < y$ since emission k_2 has to be removed by trimming. Thus we have that $x_1/x_2 > x_1/y$. As a consequence (6.28) can be written as

$$\delta(\rho - \rho_1)\left[\Theta\left(\rho_2 < \frac{\rho x_1}{y}\right) - \Theta(\rho_2 < \rho)\,\Theta\left(\rho_2 < \frac{\rho x_1}{y}\right)\right]\Theta_2^{\text{out}}. \tag{6.29}$$

For $x_1 < y$ this vanishes while for $x_1 > y$ the term in big square brackets gives $\Theta\left(\rho_2 < \frac{\rho x_1}{y}\right) - \Theta(\rho_2 < \rho)$. Thus one finally gets for $\mathcal{F}^{\text{trim,a}}$

$$\mathcal{F}^{\text{trim,a}} = \left(\frac{C_F\alpha_s}{\pi}\right)^2 \int d\Phi_2\,\Theta_2^{\text{out}}\delta(\rho - \rho_1)\,\Theta(x_1 > y)\left[\Theta\left(\rho_2 < \frac{\rho x_1}{y}\right) - \Theta(\rho_2 < \rho)\right]. \tag{6.30}$$

The above result has a simple interpretation. The veto on emissions that one places for the case of pure Y-splitter is modified by the action of trimming. In the region where emissions are removed by trimming, emissions are no longer subject to the direct constraint that the mass must be less than ρ, which represents the subtraction of the $\Theta(\rho_2 < \rho)$ veto condition in the Θ_2^{out} region. However emissions in this region, even though they are removed by trimming, are still subject to the constraint $k_{t1}^2/m_{\text{jet}}^2 > y$ which is the Y-splitter cut and where m_{jet}^2 is the squared invariant mass of the ungroomed jet, to which all emissions, including those removed eventually by grooming, do contribute. Thus one gets the correction to pure Y-splitter given by (6.30), from those configurations where the highestk_t emission sets the final jet mass. These, we recall, are the configurations that generate the leading logarithmic corrections for pure Y-splitter.

It is simple to calculate $\mathcal{F}^{\text{trim,a(b)}}$. The form of the result depends on the value of ρ and there are various regimes that emerge. In what follows we shall choose values such that $r^2 < y$, as is common for phenomenological purposes, although our main conclusions will be unchanged by making a different choice. One has:

- The regime $\rho < y^2 r^2$

  Here we find

$$\mathcal{F}^{\text{trim,a}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{2}\ln\frac{1}{r^2}\ln^2 y \tag{6.31}$$

$$\mathcal{F}^{\text{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{2}\ln\frac{1}{r^2}\ln^2 y \tag{6.32}$$

$$\mathcal{F}^{\text{trim,a}} + \mathcal{F}^{\text{trim,b}} = 0. \tag{6.33}$$

The above results are noteworthy since they indicate that in the small ρ limit, $\rho \to 0$, where one may regard resummation of logarithms of ρ to be most important, the overall correction to Y-splitter vanishes at our leading-logarithmic accuracy. This is also the essential reason for the fact that trimming does not appear to significantly modify the performance of Y-splitter on background jets, as the basic structure of a Sudakov form factor suppression at small ρ is left unchanged.

- The regime $y^2 r^2 < \rho < y r^2$

  One obtains

$$\mathcal{F}^{\text{trim,a}} = \frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \left(\frac{1}{2}\ln^2\frac{1}{y}\ln\frac{1}{r^2} - \frac{1}{6}\ln^3\frac{\rho}{y^2 r^2}\right), \tag{6.34}$$

  while for $\mathcal{F}^{\text{trim,b}}$ the result coincides with that quoted in (6.32). Thus we have for the full correction from trimming:

$$\mathcal{F}^{\text{trim,a}} + \mathcal{F}^{\text{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \frac{1}{6}\ln^3\frac{\rho}{y^2 r^2}. \tag{6.35}$$

  It is instructive to ex,mine the behavior of (6.35) at the transition points: for $\rho = y^2 r^2$ it vanishes and hence trivially matches (6.33), while for $\rho = y r^2$ we get

$$-\frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \frac{1}{6}\ln^3\frac{1}{y}. \tag{6.36}$$

- The regime $y^2 > \rho > y r^2$

  Here one gets

$$\mathcal{F}^{\text{trim,a}} = \frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \left(\frac{1}{2}\ln\frac{y}{\rho}\ln^2\frac{1}{y} - \frac{1}{6}\ln^3\frac{1}{y}\right). \tag{6.37}$$

  On the other hand the result for $\mathcal{F}^{\text{trim,b}}$ in this region is

$$\mathcal{F}^{\text{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \frac{1}{2}\ln\frac{y}{\rho}\ln^2\frac{1}{y}, \tag{6.38}$$

  such that

$$\mathcal{F}^{\text{trim,a}} + \mathcal{F}^{\text{trim,b}} = -\frac{1}{\rho}\left(\frac{C_F \alpha_s}{\pi}\right)^2 \frac{1}{6}\ln^3\frac{1}{y}, \tag{6.39}$$

  i.e. independent of ρ.

  Note that the above result is identical to that reported in (6.36) for $\rho = y r^2$ as one would expect.

- The regime $y > \rho > y^2$

  Here one obtains

$$\mathcal{F}^{\text{trim,a}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2\left(\frac{1}{3}\ln^3\frac{y}{\rho} + \frac{1}{2}\ln^2\frac{y}{\rho}\ln\frac{\rho}{y^2}\right). \tag{6.40}$$

  The result for $\mathcal{F}^{\text{trim,b}}$ in this region remains the same as in (6.38) so that

$$\mathcal{F}^{\text{trim,a}} + \mathcal{F}^{\text{trim,b}} = \frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2\ln\frac{y}{\rho}\left(\frac{5}{6}\ln\frac{1}{\rho}\ln\frac{1}{y} - \frac{7}{6}\ln^2\frac{1}{y} - \frac{1}{6}\ln^2\frac{1}{\rho}\right), \tag{6.41}$$

  which matches (6.39) at $\rho = y^2$ and vanishes at $\rho = y$.

For $\rho > y$ the functions $\mathcal{F}^{\text{trim,a(b)}}$ vanish and there is no correction to Y-splitter which itself coincides with the plain jet mass.

To summarise, we find that, in the formal small ρ limit, we recover the same result as for the pure Y-splitter case at this order (see the region $\rho < y^2r^2$). As we move towards larger values of ρ i.e. beyond $\rho = y^2r^2$, we find that the result becomes substantially more complicated. We find transition points at y^2r^2, yr^2, y^2 and y which arise due to the use of trimming. The result in all these regions contains logarithms of ρ along with logarithms of y (as well as $\ln r$ terms) . However in these regions logarithms of ρ cannot be considered to be dominant over other logarithms such as those in y. To get a better feeling for the size of the corrections to the pure Y-splitter case in various regions it is helpful to look at the behavior at the transition points. At $\rho = y^2r^2$ the correction due to trimming vanishes while at $\rho = yr^2$ one finds an overall correction varying as $\frac{1}{\rho}\alpha_s^2\ln^3 y$ which is formally well beyond our leading-logarithmic accuracy in ρ, although enhanced by logarithms of y. The behavior at other transition points is similarly highly subleading in ρ though containing logarithms in y. As we have already noted before resummation of $\ln y$ enhanced terms has only a modest effect and does not affect our understanding of the basic behavior of the tagger (see fig. 6.2).

The fixed-order results of this section already explain why the action of trimming following the application of Y-splitter only changes the performance of Y-splitter at a subleading level. It is simple to carry out a resummed calculation valid at the leading logarithmic level in ρ but with only an approximate treatment of subleading terms. Such a resummed calculation is in fact seen to be in qualitative agreement with Monte Carlo studies. However a feature of the result obtained with trimming, which is perhaps undesirable from a phenomenological viewpoint, is the presence of multiple transition points in the final result. While these transition points are not as visible as for the case of pure trimming itself (see [12]) it may nevertheless be desirable to think of using grooming methods which are known to have less transition points in conjunction with Y-splitter. To this end we shall first investigate the modified MassDrop Tagger (mMDT) at fixed-order before addressing the question of resummation and comparisons to Monte Carlo of Y-splitter with grooming.

6.2.2 Y-splitter with mMDT: fixed-order results

The NLO calculation for Y-splitter with mMDT proceeds similarly to the case of the Y-splitter trimming combination but with differences of detail. If one considers the correction to the pure Y-splitter case at this order, we arrive at functions $\mathcal{F}^{\text{mMDT,a(b)}}$ which can be computed exactly like $\mathcal{F}^{\text{trim,a(b)}}$ with the only difference being in the condition Θ_2^{out} for removal of emission k_2 by the mMDT as well as condition $\Theta_1^{\text{in}} = 1 - \Theta_1^{\text{out}}$ which differs from the trimming case. To be more explicit, for mMDT to remove the emission k_2 one has that $\Theta_2^{\text{out}} = \Theta(\theta_2 > \theta_1)\,\Theta(x_2 < y)$ since mMDT would not reach emission k_2 if it were at smaller angle than k_1, as k_1 passes the mMDT cut.

In contrast to trimming, the final result contains only two transition points at $\rho = y^2$ and $\rho = y$. We obtain for the correction to Y-splitter $\mathcal{F}^{\text{mMDT}} = \mathcal{F}^{\text{mMDT,a}} + \mathcal{F}^{\text{mMDT,b}}$ such that:

- For $\rho < y^2$

$$\mathcal{F}^{\text{mMDT}} = -\frac{1}{\rho}\left(\frac{C_F\alpha_s}{\pi}\right)^2 \frac{1}{6}\ln^3\frac{1}{y}. \tag{6.42}$$

  This agrees with the result for trimming at $yr^2 < \rho < y^2$, quoted in eq. (6.39).

- For $y > \rho > y^2$

  Here again the result is identical to that obtained for trimming i.e. the sum of $\mathcal{F}^{\text{trim,a}}$ and $\mathcal{F}^{\text{trim,b}}$ in the same region.

Note that one can alternatively obtain the mMDT results by taking the limit $r \to 0$ in the trimming results.

As before, for $\rho > y$ one obtains no correction from grooming or Y-splitter and the result for the plain mass is recovered, meaning once more that grooming will not substantially affect the small-ρ behavior of Y-splitter.

In summary using mMDT as a groomer produces a result that, as for the case of trimming, produces only sub-leading corrections in terms of logarithms of ρ and hence leaves the pure Y-splitter Sudakov unaltered at leading logarithmic level in the limit of small ρ. The sub leading terms carry enhancements involving logarithms of y as for trimming, but there are fewer transition points for mMDT than trimming, which is certainly a desirable feature from a phenomenological viewpoint.

6.2.3 All-orders calculation for Y-splitter with mMDT

As explicitly shown via fixed-order calculations in the previous section, the use of grooming methods subsequent to the application of Y-splitter does not modify the leading logarithmic results in a small ρ resummation. It is straightforward to see that this statement extends beyond fixed-order to all perturbative orders and is the reason why the performance of Y-splitter on background jets is not fundamentally altered by groomers.

Beyond the leading logarithmic level however the situation with Y-splitter becomes more complicated when one introduces grooming. One may therefore wonder about the

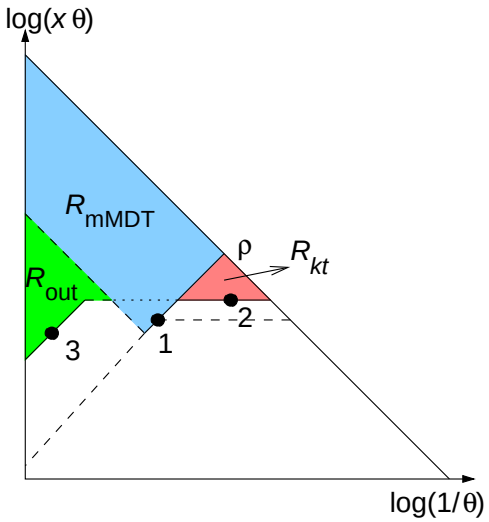


Figure 6.3 — Lund diagram representing Sudakov exponents for the resummation of the $\log y$ for Y-splitter with mMDT.

practical impact of formally sub-leading corrections on the tagger behavior. It is therefore of some interest to write down a resummed result that goes beyond leading-logarithmic accuracy in ρ and captures some of the formally sub-leading terms that emerge in the various regimes we have identified, such as those enhanced by logarithms of y. In this section we will carry out the same kind of resummation as reflected by eqs. (6.16) and (6.18) for the Y-splitter combined with mMDT case. We work in the leading logarithmic accuracy and keep both leading logarithms in ρ and y_{cut}.

At our accuracy, we can assume that the groomed mass is dominated by a single emission, say emission 1 with momentum fraction x_1 and at an angle θ_1 to the jet axis. The fact that emission 1 is kept in the groomed jet guarantees that $x_1 > y_{\text{cut}}$. We then have to consider four separate cases according to which emissions dominate the k_t and mass scales entering the Y-splitter condition. We can write

$$\begin{aligned}\frac{\sigma}{\rho}\frac{d\sigma}{d\rho} = \int_y^1 dx_1\, p_i(x_1)\frac{\alpha_s(\rho x_1)}{2\pi}e^{-R_{\text{mMDT}(\rho)}} \\ \left\{e^{-R_{k_t}(\kappa_1;\rho)}\left[e^{-R_{\text{out}}(\rho;\kappa_1)} + \int_\rho^y \frac{d\rho_3}{\rho_3}R'_{\text{out}}(\rho_3;\kappa_1)e^{-R_{\text{out}}(\rho_3;\kappa_1)}\Theta(\kappa_1^2 > y\rho_3)\right]\right. \\ + \int_{\kappa_1}^{\sqrt{\rho}}\frac{d\kappa_2}{\kappa_2}R'_{k_t}(\kappa_2;\rho)e^{-R_{k_t}(\kappa_2;\rho)}\left[e^{-R_{\text{out}}(\rho;\kappa_2)}\Theta(\kappa_2^2 > y\rho)\right. \\ \left.\left. + \int_\rho^y \frac{d\rho_3}{\rho_3}R'_{\text{out}}(\rho_3;\kappa_2)e^{-R_{\text{out}}(\rho_3;\kappa_2)}\Theta(\kappa_2^2 > y\rho_3)\right]\right\}.\end{aligned} \tag{6.43}$$

In the above expression, the two terms on the second line correspond to emission 1 also dominating the k_t scale, while the last two lines correspond to an additional emission 2 dominating the k_t scale. In both cases, the plain jet mass can either be dominated by emission 1 (the first term in each squared brackets) or by an additional emission 3 (the

second terms in each squared brackets). Different terms are weighted by different Sudakov factors:

$$R_{\text{mMDT}}(\rho) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta(x > y)\, \Theta(x\theta^2 > \rho), \tag{6.44}$$

$$R_{k_t}(\kappa_i;\rho) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta(x\theta > \kappa_i)\, \Theta(x\theta^2 < \rho), \tag{6.45}$$

$$R_{\text{out}}(\rho;\kappa_i) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta(x < y)\, \Theta(x\theta > \kappa_i \text{ or } x\theta^2 > \rho). \tag{6.46}$$

These are graphically represented in fig. 6.3. The $R'_{\text{mMDT}}(\rho)$, $R'_{k_t}(\kappa;\rho)$ and $R'_{\text{out}}(\rho;\kappa)$ are the derivatives of the above radiators wrt to the logarithm of (one over) their first argument.

Note that the intermediate transition at κ_i in R_{out} comes from the fact that an emission with $x < y$ and a k_t scale larger than κ_i would dominate both the k_t and mass scales and the Y-splitter condition would not be satisfied. This region is therefore automatically excluded.

Both integrations on ρ_3 can be performed quite straightforwardly:

$$\int_\rho^y \frac{d\rho_3}{\rho_3} R'_{\text{out}}(\rho_3;\kappa_i) e^{-R_{\text{out}}(\rho_3;\kappa_i)} \Theta(\rho_3 < \kappa_i^2/y) = e^{-R_{\text{out}}(\kappa_i^2/y)} - e^{-R_{\text{out}}(\rho)}. \tag{6.47}$$

In the above equation, we can drop the κ argument of $R_{\text{out}}(\rho;\kappa)$ for the following reason: for $\rho < \kappa^2/y$, $x\theta > \kappa$ and $x < y$ automatically imply $x\theta^2 > \rho$ so that we can replace $\Theta(x\theta > \kappa_i \text{ or } x\theta^2 > \rho)$ by $\Theta(x\theta^2 > \rho)$. We therefore have

$$R_{\text{out}}(\rho) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta(x < y)\, \Theta(x\theta^2 > \rho). \tag{6.48}$$

Using (6.47) for both squared brackets in (6.43), we obtain

$$\begin{aligned}\frac{\sigma}{\rho}\frac{d\sigma}{d\rho} = \int_y^1 dx_1\, p_i(x_1) \frac{\alpha_s(\rho x_1)}{2\pi} e^{-R_{\text{mMDT}(\rho)}} \\ \left[e^{-R_{k_t}(\kappa_1;\rho) - R_{\text{out}}(\kappa_1^2/y)} + \int_{\kappa_1}^{\sqrt{\rho}} \frac{d\kappa_2}{\kappa_2} R'_{k_t}(\kappa_2;\rho) e^{-R_{k_t}(\kappa_2;\rho) - R_{\text{out}}(\kappa_2^2/y)} \right].\end{aligned} \tag{6.49}$$

While this equation is suitable for practical purposes, specifically numerical integration over k_{t2} and z_1, it is not ideal to see the logarithmic structure of the result. For that purpose it proves to be better to factor $\exp[-R_{\text{out}}(\rho)]$, which would combine with the $\exp[-R_{\text{mMDT}}(\rho)]$ prefactor to give the plain jet mass Sudakov, leading to

$$\begin{aligned}\frac{\sigma}{\rho}\frac{d\sigma}{d\rho}^{\text{LL+LL}_y} = \int_y^1 dx_1\, p_i(x_1) \frac{\alpha_s(\rho x_1)}{2\pi} e^{-R_{\text{plain}}(\rho)} \\ \left[e^{-R_{k_t}(\kappa_1;\rho) - (R_{\text{out}}(\kappa_1^2/y) - R_{\text{out}}(\rho))} + \int_{\kappa_1}^{\sqrt{\rho}} \frac{d\kappa_2}{\kappa_2} R'_{k_t}(\kappa_2;\rho) e^{-R_{k_t}(\kappa_2;\rho) - (R_{\text{out}}(\kappa_2^2/y) - R_{\text{out}}(\rho))} \right],\end{aligned} \tag{6.50}$$

where $R_{\text{plain}}(\rho)$ and $R_{k_t}(\kappa;\rho)$ are defined in eqs. (6.15) and (6.17) respectively, and

$$R_{\text{out}}(\rho) - R_{\text{out}}(\kappa_1^2/y) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta(x<y)\,\Theta(\kappa_1^2/y > x\theta^2 > \rho). \tag{6.51}$$

One can show that the second line in (6.50) only brings subleading logarithmic contributions (in $\ln\rho$), so that the LL result is fully given by the first line in (6.50) and corresponds to the LL result for pure Y-splitter. This can be obtained from the following observations. The R_{k_t} factors, already encountered before, bring at most subleading corrections proportional to $\alpha_s \ln^2 y$. Then, since $\kappa_1^2/y = \rho x_1/y$ and $y < x_1 < 1$, $R_{\text{out}}(\rho) - R_{\text{out}}(\kappa_1^2/y)$ can at most bring single-logarithmic corrections proportional to $\alpha_s \ln\rho \ln y$. This remains valid for $R_{\text{out}}(\rho) - R_{\text{out}}(\kappa_2^2/y)$ since $\ln(\kappa_1^2/\kappa_2^2)$ can at most introduce logarithms of y.

Alternatively, it is instructive to evaluate (6.50) with a fixed-coupling approximation. Assuming, for simplicity, that $\rho < y^2$, and working in the soft-collinear approximation (for quark jets) where we can use $p_i(x) = 2C_F/x$, we have

$$R'_{k_t}(\kappa_i;\rho) = \frac{2\alpha_s C_F}{\pi} \ln\frac{\rho}{\kappa_i^2}, \tag{6.52}$$

$$R_{k_t}(\kappa_i;\rho) = \frac{\alpha_s C_F}{2\pi} \ln^2\frac{\rho}{\kappa_i^2}, \tag{6.53}$$

$$R_{\text{out}}(\rho) - R_{\text{out}}(\kappa_i^2/y) = \frac{\alpha_s C_F}{2\pi}\left(\ln^2\frac{y}{\rho} - \ln^2\frac{y^2}{\kappa_i^2}\right). \tag{6.54}$$

Substituting these expressions in eq. (6.50) one can reach after a few manipulations

$$\frac{\sigma}{\rho}\frac{d\sigma}{d\rho}^{\text{LL+LL}_y} = e^{-R_{\text{plain}}(\rho)} \int_y^1 \frac{dx}{x}\frac{\alpha_s C_F}{\pi}\left(1 + \frac{\alpha_s C_F}{\pi}\ln\frac{1}{x}\ln\frac{x}{y}\right) e^{-\frac{\alpha_s C_F}{2\pi}\left(\ln^2 x - \ln\frac{x}{y}\ln\frac{y^3}{\rho^2 x}\right)}. \tag{6.55}$$

In the above expression, the factor in front of the exponential as well as the first term in the exponential only yield terms of the form $(\alpha_s \ln^2 y)^n$, and the second term in the exponential will lead to both $(\alpha_s \ln^2 y)^n$ and $(\alpha_s \ln y \ln\rho)^n$ contributions. These are both subleading compared to the leading-logarithmic accuracy in ρ so that (6.55) will lead to the $\frac{\alpha_s C_F}{\pi}\ln\frac{1}{y} e^{-R_{\text{plain}}(\rho)}$ result plus subleading contributions as expected.

While a complete evaluation of the integral over x in (6.55) is not particularly illuminating – it gives an error function – it is interesting to expand it to second order in α_s. One obtains

$$\frac{\sigma}{\rho}\frac{d\sigma}{d\rho}^{\text{LO+NLO,soft-coll}} = \frac{\alpha_s C_F}{\pi}\ln\frac{1}{y} - \frac{1}{2}\left(\frac{\alpha_s C_F}{\pi}\right)^2 \ln\frac{1}{y}\left(\ln^2\rho - \ln\rho\ln y + \frac{4}{3}\ln^2 y\right), \tag{6.56}$$

which correctly reproduces the sum of (6.12) and (6.42).

6.2.4 Comparison to Monte Carlo

Our result eq. (6.50) shows that the leading logarithmic results obtained for Y-splitter with mMDT coincide with those for pure Y-splitter since the factor in the big square

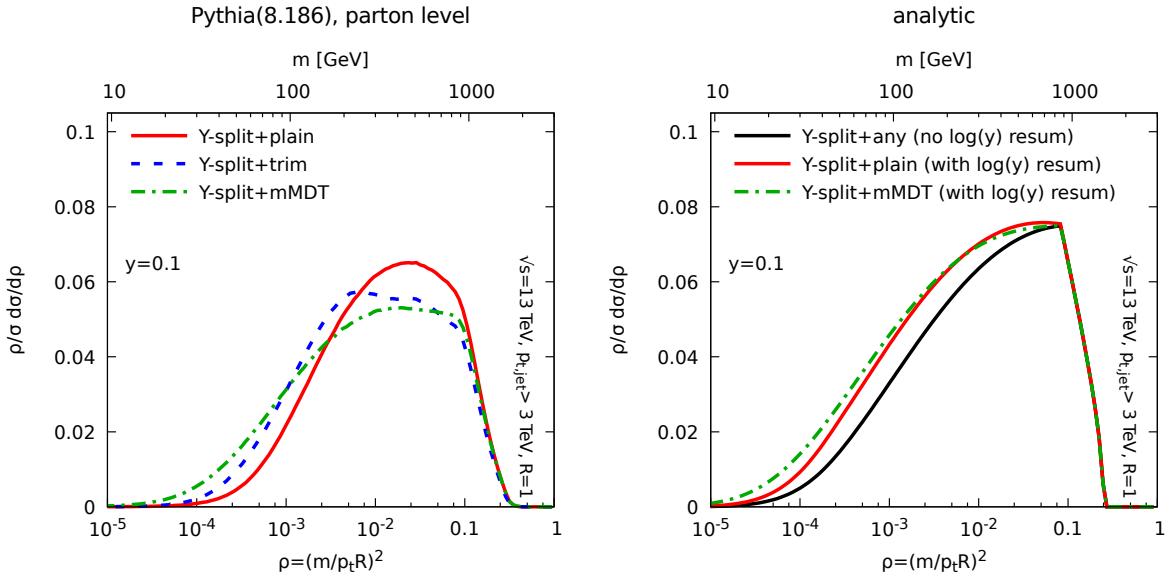


Figure 6.4 — Jet mass distribution for Y-splitter with selected groomers, Pythia simulations (left) and analytical results (right). For analytical curves, we show results with only $\log \rho$ resummation (black) and results with $\log y$ terms for pure Y-splitter (red) and Y-splitter with mMDT (green).

bracket only generates subleading corrections to the pure Y-splitter result. This result also contains the resummation of leading logarithmic terms in y, which are subleading from the point of view of $\ln \rho$ resummation. The analytic results for mMDT with $\ln y$ resummation are plotted in fig. 6.4. Also plotted for reference is the leading logarithmic resummed result, which is independent of whether we groom with mMDT or trimming, or not at all. We can see that, as also observed before for the pure Y-splitter case, resummation of $\ln y$ terms brings only modest differences compared to the leading logarithmic answer. In fig. 6.4 the plot on the left shows the results obtained with Monte Carlo studies for Y-splitter with trimming and mMDT compared to pure Y-splitter.[3] The plot reaffirms our observation that grooming does not alter the essential feature of a Sudakov suppression at small ρ. The Monte Carlo result for trimming also shows some hints of the transition in behavior induced by subleading terms and is correspondingly less smooth than the mMDT result which has fewer transition points.

We note that while we have performed a $\ln y$ resummation in order to assess their impact on the LL result we do not claim that these terms are numerically more important (for practically used values of y) than other subleading in ρ effects we have neglected, such as non-global logarithms and multiple emission effects. Non-global logarithms in particular are known to have a substantial impact on the peak height of the jet-mass spectrum [73]. However these other effects are harder to treat and hence we used the $\ln y$ resummation as a convenient method to assess the impact of some subleading terms on

[3] We used the implementation of mMDT (and SoftDrop) provided in `fjcontrib` [83].

the LL result.

6.3 Variants

6.3.1 Y-splitter with mass declustering

We have seen in the previous section that beyond the strict leading logarithmic approximation in $\ln \frac{1}{\rho}$, the behavior of the tools can be quite complex, especially when we combine Y-splitter with grooming. In this section, we discuss a small modification to the definition of Y-splitter that largely simplifies this calculation and has the fringe benefit of coming with a small performance enhancement.

Most of the complication in the calculations we have done so far comes from the fact that the emission which passes the Y-splitter cut is the highest k_t emission, which can be different from the emission that dominates the mass. Such configurations produce only terms beyond leading-logarithmic accuracy but as we have seen their structure is rather involved. The discussion and results beyond LL would clearly be simpler if the k_t scale entering Y-splitter was directly calculated based on the emission that dominates the jet mass. One can readily achieve this by *replacing the k_t declustering by a generalized-k_t declustering with $p = 1/2$* which respects the ordering in mass so that the emission that passes Y-splitter is also the emission that dominates the jet mass. If we consider a soft emission with momentum fraction x_1 at an angle θ_1, which dominates the mass, this would give a cut of the form

$$\frac{x_1^2\theta_1^2}{x_1\theta_1^2} = x_1 > y. \tag{6.57}$$

More precisely if we choose to include finite y corrections one obtains

$$\frac{\left(\min\left(x_1, 1-x_1\right)\right)^2 \theta_1^2}{x_1(1-x_1)\theta_1^2} > y \qquad \Rightarrow \qquad \frac{1}{1+y} > x_1 > \frac{y}{1+y}. \tag{6.58}$$

We denote this variant $\mathrm{Y_m}$-splitter, where the subscript m refers to the fact that we now use a mass-ordered declustering procedure. Regardless of whether we ultimately measure the jet mass without grooming or the groomed jet mass, $\mathrm{Y_m}$-splitter computed on the plain jet will always impose that the emission that dominates the plain jet mass has a momentum fraction larger than y. In the case where we measure the plain jet mass, we would therefore simply recover the result quoted in (6.14) with no $\alpha_s^2 \ln \frac{y}{\rho} \ln^2 \frac{1}{y}$ correction.

On top of that, the $\mathrm{Y_m}$-splitter condition guarantees that the emission dominating the plain mass also passes the trimming (or mMDT) condition. We would therefore also recover (6.14) for the $\mathrm{Y_m}$-splitter+grooming case, as only emissions that do not essentially affect the jet mass can be removed by grooming.

Comparisons between Monte-Carlo simulations, still using Pythia8 at parton level, and the analytic expectation (6.14) are presented in fig. 6.5. We clearly see that our analytic result captures very well the shape observed in the Monte-Carlo simulation. It also appears that differences between the ungroomed case and the two groomed cases are

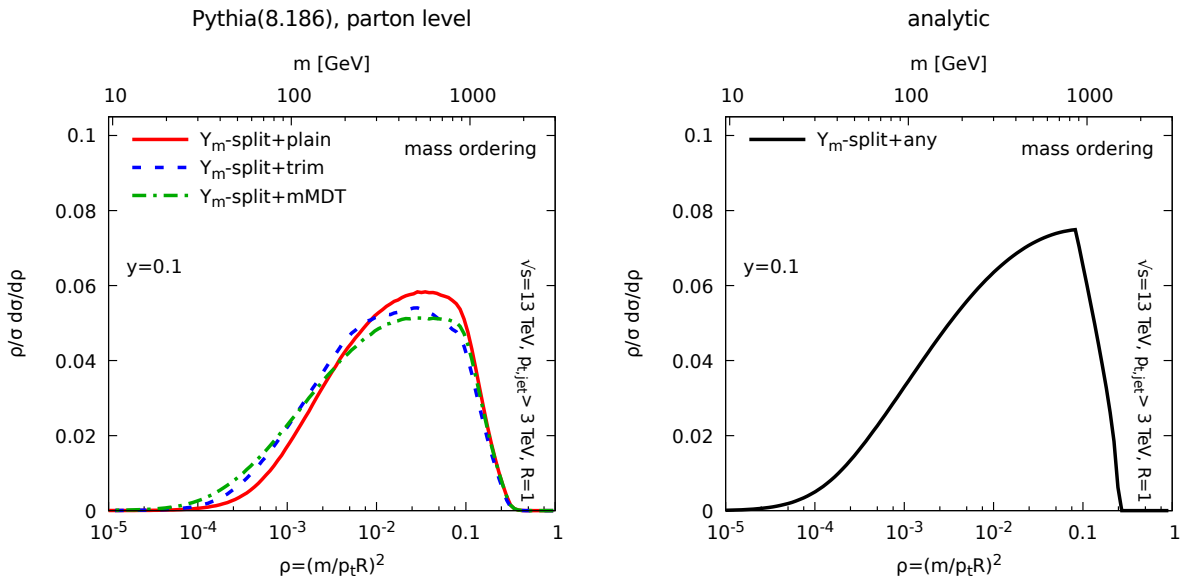


Figure 6.5 — Analytical results for jet mass distribution for Y_m-splitter (right) compared to Pythia simulations (left) combined with selected groomers.

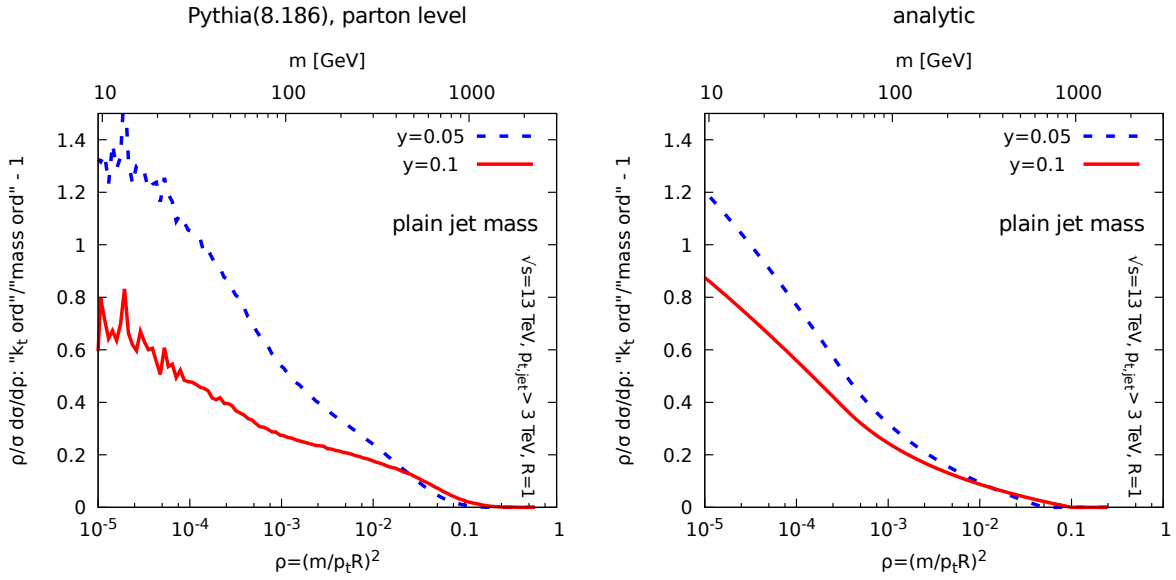


Figure 6.6 — Ratio of mass distribution for (k_t-ordered) Y-splitter divided by (mass-ordered) Y_m-splitter, Pythia simulations (left) and analytical results (right).

smaller than what was observed for the standard Y-splitter case discussed in the previous two sections (see e.g. fig. 6.4), as one would expect from the analytical viewpoint. It appears also that using $\mathrm{Y_m}$-splitter comes with a fringe benefit, namely the fact that it suppresses the mass spectrum somewhat more than Y-splitter does. As an additional test of our analytic calculations, we can compare the difference between our results for the mass-ordered case eq. (6.18) and eq. (6.14) representing our result for the usual k_t ordered Y-splitter to Monte-Carlo results. This is shown in fig. 6.6 and, bearing in mind that our analytic calculation only resums contributions maximally enhanced by $\ln \frac{1}{y}$, shows a good agreement between the two sides of the figure. fig. 6.6 also illustrates the fact that the difference between Y- and $\mathrm{Y_m}$-splitter essentially behaves like $\ln \frac{y}{\rho}$ up to running coupling corrections.

A comment is due about differences between the groomed and ungroomed jet mass after imposing the $\mathrm{Y_m}$-splitter condition. We would still expect these differences to appear at subleading logarithmic orders in ρ but they would not be enhanced by double logarithms of y. It is also interesting to notice that while most of the NLL corrections to the overall $\exp[-R_{\mathrm{plain}}(\rho)]$ Sudakov factor would be the same as for the plain jet mass, the correction due to multiple emissions would be different. This can be understood from the fact that, if several emissions, $(x_1, \theta_1), \dots (x_n, \theta_n)$ contribute significantly to the plain jet mass, only the largest, say (x_1, θ_1), will be used to compute the k_t scale leading to the $\mathrm{Y_m}$-splitter constraint

$$x_1^2 \theta_1^2 > y \sum_{i=1}^{n} x_i \theta_i^2, \tag{6.59}$$

which is no longer as simple as (6.57), albeit more constraining. One can still carry out a resummation with this exact condition but it leads to more complicated expressions which go beyond our accuracy. Note that at the same, single-logarithmic, order of accuracy, one would anyway have to include additional contributions, in particular the nontrivial contribution from non-global logarithms.

6.3.2 Y-splitter with mass declustering and a z cut

It is possible to further simplify the analytic computations by having the Y-splitter condition behave like a z_{cut} rather than a y_{cut}. As before, we first decluster the jet using the generalized k_t algorithm with $p = 1/2$ to obtain two subjets j_1 and j_2. We then impose the condition

$$z_{\mathrm{cut}} \equiv \frac{\min(p_{t1}, p_{t2})}{p_{t1} + p_{t2}} > z. \tag{6.60}$$

As for the case of a mass declustering with a y_{cut}, this would lead to (6.14) at leading logarithmic accuracy in $\ln \frac{1}{\rho}$, and be free of subleading corrections enhanced by logarithms of z. Moreover, if multiple emissions, $(x_1, \theta_1), \dots (x_n, \theta_n)$, contribute to the plain jet mass, with $x_1 \theta_1^2 \geq x_i \theta_i^2$, the $\mathrm{Y_m}$-splitter condition will give

$$z_{\mathrm{cut}} = x_1 > z, \tag{6.61}$$

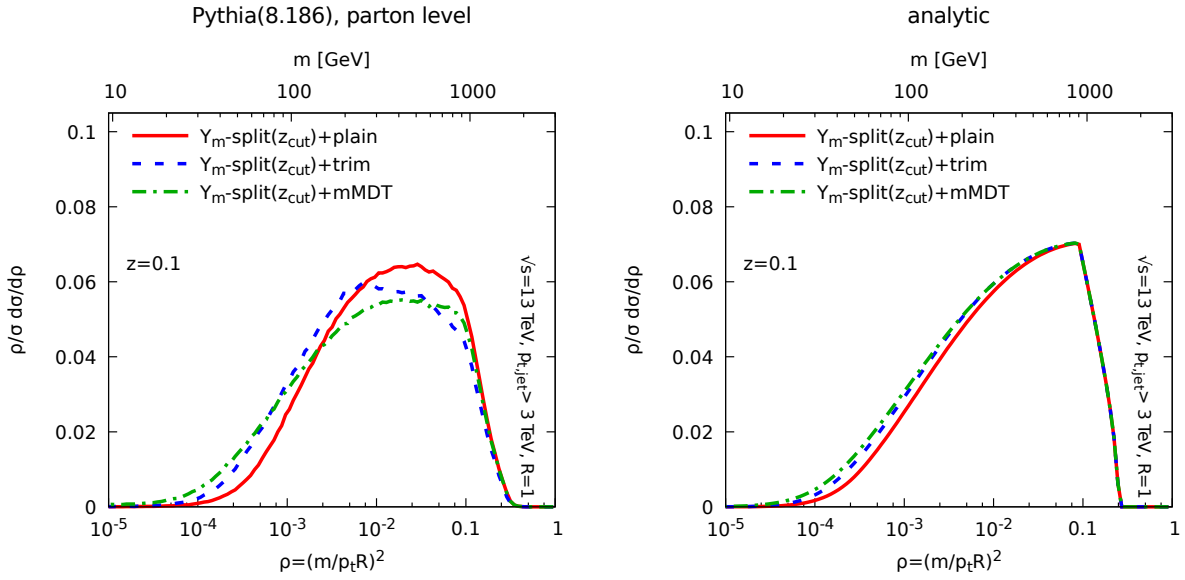


Figure 6.7 — Mass distributions for a $\mathrm{Y_m}$-splitter condition with a z_{cut} instead of a y_{cut}, followed by mMDT (green), trimming (blue) or no grooming (red), Pythia simulations (left) and analytical results (right) with multiple-emission corrections.

which is significantly simpler than the corresponding condition with a y_{cut}, eq. (6.59). This is valid independently of which mass, groomed or ungroomed, we decide to measure. However, even if we apply a grooming procedure, the $\mathrm{Y_m}$-splitter condition (6.61) guarantees that the emission (x_1, θ_1) which dominates the jet mass is kept by grooming and dominates also the groomed jet mass. The multiple-emission correction to the measured jet mass, groomed or ungroomed, will therefore be sensitive to all the emissions, including (x_1, θ_1), kept in the jet used to measure the mass. Their resummation leads to the standard form [78] for additive observables $\exp(-\gamma_E R'_{\mathrm{mass}})/\Gamma(1 + R'_{\mathrm{mass}})$, where R'_{mass} is the $\ln \frac{1}{\rho}$-derivative of the Sudakov associated with the mass we consider i.e. either the plain jet mass or the groomed jet mass Sudakov. The mass distribution is then given by

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\mathrm{LL+ME}} = \int_z^{1-z} dx_1\, p_i(x_1)\, \frac{\alpha_s(x_1\rho)}{2\pi}\, \frac{e^{-R_{\mathrm{plain}}(\rho)-\gamma_E R'_{\mathrm{mass}}(\rho)}}{\Gamma\left(1+R'_{\mathrm{mass}}(\rho)\right)}, \tag{6.62}$$

with the superscript "ME" indicating that the contribution from multiple emissions is included and

$$R'_{\mathrm{mass}}(\rho) = \int_0^1 \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi}\, \rho\delta(x\theta^2-\rho)\, \Theta_{\mathrm{in}}, \tag{6.63}$$

where the Θ_{in} imposes that the emission is kept by grooming, or is set to 1 for the plain jet mass. A comparison between (6.62) and Monte-Carlo simulations is provided in fig. 6.7.

Despite the simplicity of the analytic results, and the fact that the general shape is well reproduced by the analytic results, one should note that the Monte Carlo simulations show a slightly larger spread between the different groomers than what was observed with

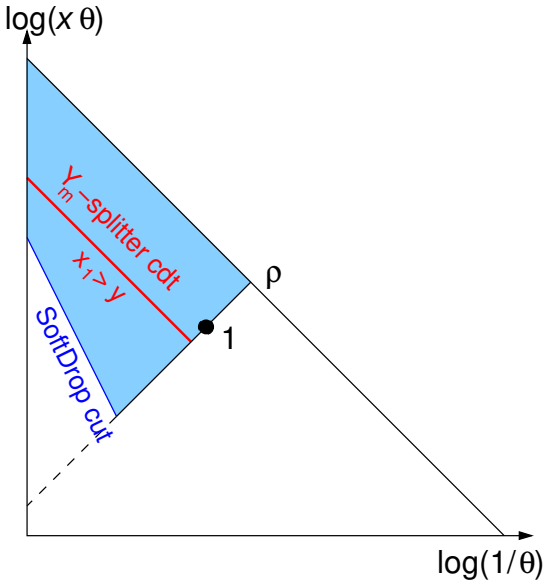


Figure 6.8 — Lund diagram corresponding to Y_m-splitter applied on a pre-groomed jet with SoftDrop. The shadowed area corresponds to the region allowed by SoftDrop and entering into the Sudakov factor. The red line corresponds to the Y_m-splitter condition.

a y_{cut} Y_m-splitter condition, indicating a larger impact of subleading terms for the z_{cut} condition.

Furthermore, the mass spectrum is slightly higher at small masses with a z_{cut} than with a y_{cut}, and we should therefore expect a slightly better tagging performance for the latter. This can be seen directly in the Monte Carlo plots in figs. 6.5 and 6.7, and ought to be apparent from an analytic calculation including multiple emissions also for the y_{cut} case. Physically, we attribute that to the fact that the Y_m-splitter condition including multiple emissions is more constraining in the case of a y_{cut}. (6.59), than with a z_{cut} (6.61).

Conversely, one should expect a z_{cut}-based Y_m-splitter to be less sensitive to non-perturbative effects than a y_{cut}-based Y_m-splitter. Which is studied with more detail in section 6.4.

6.3.3 Y-splitter with SoftDrop pre-grooming

There is one last possible adaptation of the Y-splitter method that we wish to introduce. Our original motivation to combine Y-splitter with grooming was to reduce the sensitivity of the plain jet mass to non-perturbative effects, especially important for the consequent loss of signal efficiency. We have then considered the mMDT and trimming as possible ways to solve that issue. For these situations, we have shown that it was crucial to apply the Y-splitter condition on the plain jet mass and use grooming to determine the final jet mass after applying the Y-splitter condition.

There is however an alternative, and in some sense intermediate, possibility. Instead of using the modified MassDrop Tagger or trimming we can groom the jet using SoftDrop. More precisely, one first applies a SoftDrop procedure — with parameters $\zeta_{\text{cut}} < y_{\text{cut}}$ and β — to the jet in order to reduce the non-perturbative effects and, after this pre-grooming step, we impose the Y-splitter condition on the pre-groomed jet.

In practice, this would be very similar to the case of the plain jet mass discussed in

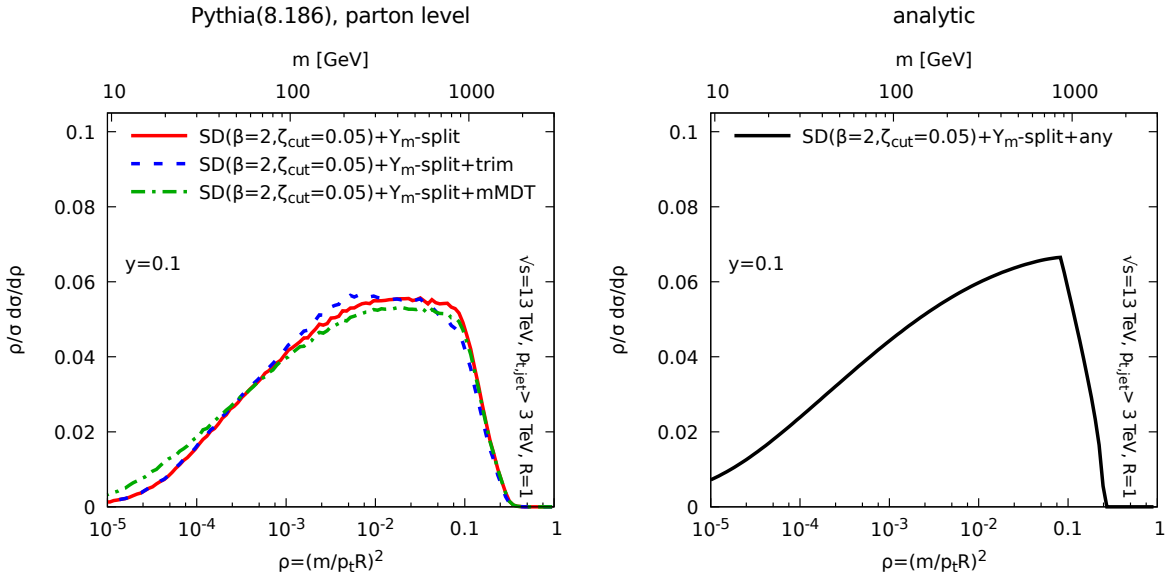


Figure 6.9 — Jet mass distribution with SoftDrop pre-grooming with $\zeta_{\rm cut} = 0.05$ and $\beta = 2$. Left, Pythia simulations with Y_m-splitter with mMDT (green), trimming (blue) or no grooming (red); right, analytical predictions.

section 6.1 except that it would apply to a SoftDropped jet in which soft and large-angle emissions have been groomed away. Focusing on the Y_m-splitter case, i.e. using a mass declustering, it is straightforward to realise that the mass distribution would be given by

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{\rm LL} = \int_{\frac{y}{1+y}}^{\frac{1}{1+y}} dx_1\, p_i(x_1)\, \frac{\alpha_s(x_1\rho)}{2\pi} e^{-R_{\rm SD}(\rho)}, \tag{6.64}$$

where the Sudakov exponent, graphically represented in fig. 6.8, now includes the effect of SoftDrop

$$\begin{aligned} R_{\rm SD}(\rho) &= \int \frac{d\theta^2}{\theta^2} dx\, P(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi}\, \Theta\left(x > \zeta_{\rm cut}\theta^\beta\right)\, \Theta\left(x\theta^2 > \rho\right) \\ &= T_{02}(e^{-B_i}, \rho; C_R, 0) - T_{-\beta,0}(\zeta_{\rm cut}, \rho; C_R, 0). \end{aligned} \tag{6.65}$$

As for the "pure" Y_m-splitter case discussed in section 6.3.1, this result captures the leading behavior, without any additional subleading logarithms of $y_{\rm cut}$ to resum. Furthermore, (6.64) is also largely unaffected by a possible mMDT or trimming one would apply after the Y_m-splitter condition since the latter guarantees that the emission that dominates the mass carries a momentum fraction larger than $y_{\rm cut}$.[4]

Compared to the pure Y-splitter case (6.14), we should expect the pre-groomed result (6.64) to show a worse performance. This is due to the fact that SoftDrop grooms away

[4]Differences between groomers would still apply due to sub-leading single logarithmic terms coming from multiple-emission contributions to the jet mass. Note also that in the case of trimming, there would be an interference between the SoftDrop and trimming conditions when the latter starts cutting angles smaller than $R_{\rm trim}$, which occurs for $\rho = \zeta_{\rm cut} R_{\rm trim}^{2+\beta}$.

a region of the phase-space that would otherwise be constrained in the ungroomed case, resulting into a smaller Sudakov suppression for the SoftDrop+Y-splitter case compared to the pure Y-splitter case.

Conversely, the region which is groomed away is also the region which is expected to be the most affected by non-perturbative effects, the underlying event in particular. We should therefore expect the pre-groomed Y-splitter to be more robust against non-perturbative effects.

Note also that, although we have advocated so far that it is important to apply the groomer after the Y-splitter condition, here we apply the grooming procedure first. This makes sense since we here apply a much gentle grooming procedure – SoftDrop with positive β – and, as a consequence, we still benefit from a large Sudakov suppression.

Finally, we have compared our analytic result (6.64) with Pythia8 Monte Carlo simulations in fig. 6.9 and we see once again that it does capture the overall behavior. We also notice in the Monte-Carlo simulations that once the pre-grooming step has been applied, the effect of an extra grooming (mMDT or trimming) has almost no effect.

6.4 Non-perturbative effects

Our discussion has so far focused on pure perturbative effects. It is nevertheless also important to assess the size of non-perturbative effects, which we would like to be as small as possible, for better theoretical control.

While for a perturbative understanding of taggers one can use methods based on first principles of QCD, for understanding the role of non-perturbative corrections this is much less straightforward. Non-perturbative corrections at hadron colliders originate both from hadronisation corrections as well as from the underlying event and can have a substantial effect on tagger performances. In this section, we use Monte Carlo event generators to estimate the magnitude of this effects.

Hence in order to estimate non-perturbative effects, we have used Pythia8 with tune 4C [82] to simulate W jets (our signal, obtained from WW events) and quark jets (our background, obtained from $qq \to qq$ Born-level events). For each event, we select the (plain) jets passing a given p_t cut that we shall vary between 250 GeV and 3 TeV and then apply one of the tagging procedures used in this paper to obtain a mass distribution for the signal and background jets. For Y-splitter, we have used a y_{cut} (or z_{cut}) of 0.1, adapting the mMDT and trimming energy cut accordingly. Finally, in order to obtain the signal and background efficiencies we have kept jets which, after the whole procedure, have a mass between 60 and 100 GeV. All efficiencies presented in this section are normalised to the total inclusive jet cross-section to obtain (W or quark) jets above the given p_t cut.

In the previous section, we have considered a large range of Y-splitter conditions (k_t or mass declustering, y_{cut} or z_{cut}) and grooming options (ungroomed jets, mMDT, trimming or pre-grooming). It is hopeless to compare all possible combinations in a human-readable plot. We have therefore selected a few representative cases to illustrate both signal-v-background performance and sensitivity to non-perturbative effects. Between Y-splitter and Y_{m}-splitter conditions, we have limited ourselves to the latter, since it has a slightly

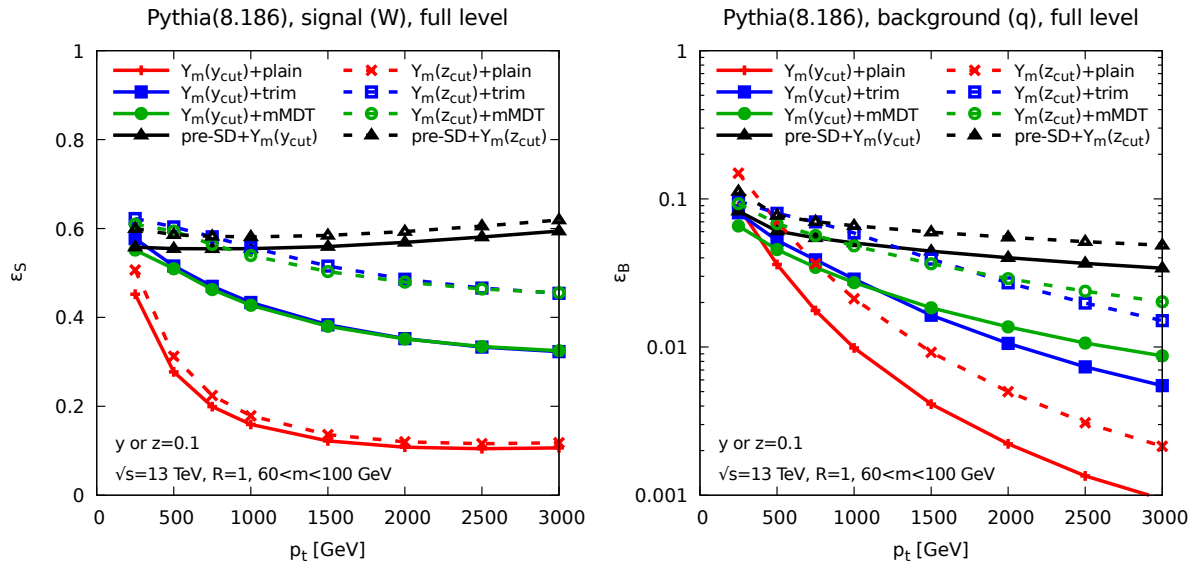


Figure 6.10 — Signal and background efficiencies at full level for a few selected tagging methods for signal (W jets) at left side and background (quark) jets at the right side. Solid (resp. dashed) lines are obtained applying a $\mathrm{Y_m}$-splitter y_{cut} (resp. z_{cut}) condition.

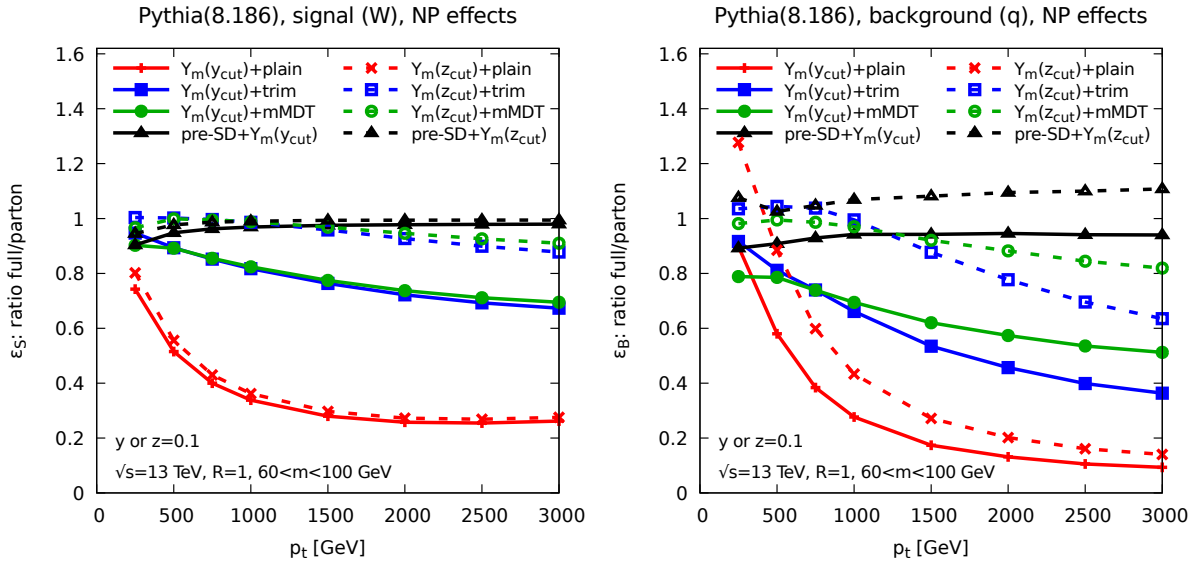


Figure 6.11 — Ratio of efficiencies obtained with-to-without non-perturbative effects for signal (left) and background (right) selected tagging methods. Solid (resp. dashed) lines are obtained applying a $\mathrm{Y_m}$-splitter y_{cut} (resp. z_{cut}) condition.

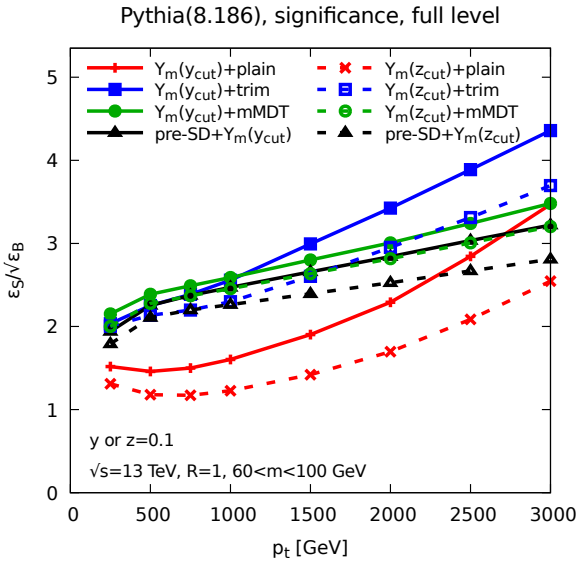


Figure 6.12 — Signal significance obtained from the efficiencies in fig. 6.10 for selected taggers. Solid (resp. dashed) lines are obtained applying a $\mathrm{Y_m}$-splitter y_{cut} (resp. z_{cut}) condition.

better performance than the former. The better performance is expected from our analytic calculations and also confirmed directly in Monte Carlo studies. We have considered both a y_{cut} and a z_{cut} type of condition, using in practice $y_{\mathrm{cut}} = z_{\mathrm{cut}} = 0.1$. We have then studied 4 grooming options: the ungroomed (or pure) case which acts as a baseline, mMDT and trimming both applied after the $\mathrm{Y_m}$-splitter condition, and SoftDrop pre-grooming for which the $\mathrm{Y_m}$-splitter condition is applied after the pre-grooming. With a y_{cut}-based $\mathrm{Y_m}$-splitter condition, the momentum fraction used in the mMDT and trimming is set to $y_{\mathrm{cut}}/(1+y_{\mathrm{cut}})$, while for a z_{cut}-based $\mathrm{Y_m}$-Splitter condition it is simply set to z_{cut}. For the SoftDrop pre-grooming, we have set $\beta = 2$ and $\zeta_{\mathrm{cut}} = 0.05$.

The signal and background efficiencies obtained from our simulations when varying the boosted jet p_t are presented in fig. 6.10 for simulations including hadronisation and the underlying event, which we also refer as "full event". For a more direct comparison of the performance of the variants of Y-splitter we have considered here, we have shown the resulting signal significance, computed as $\varepsilon_S/\sqrt{\varepsilon_B}$ in fig. 6.12.

However, differences observed in background efficiencies are usually exponential – notice the logarithmic scale on the right-hand plot of fig. 6.10 — and are therefore expected to have more impact than smaller variations in signal efficiencies. The ordering is therefore usually respected when we look at the signal significance, fig. 6.12.

These plots should be considered together with fig. 6.11 where we have plotted the ratio of the efficiencies obtained with hadronisation and the underlying event to those obtained without, as a measure of non-perturbative effects.

Additionally, to facilitate the discussion, we have plotted in fig. 6.13 two important quantities when considering the performance of a boosted-object tagging method: on the vertical axis we show the raw performance of the method, measured as usual by the signal significance. On the horizontal axis we have a measure of the method's robustness defined in terms of insensitivity to non-perturbative contributions. Here we have used a

non-perturbative correction factor defined as the ratio of the efficiencies at particle (full) and parton levels and have explicitly considered the case of quark jets, with similar trends expected for gluon jets.

Our association of robustness of a given tool with the role of non-perturbative corrections should perhaps be clarified. As we have mentioned before, non-perturbative effects cannot be fully estimated using first principles of QCD. Although one can always use event generator models for hadronisation and the underlying event there can be considerable variation in results between different event generators and also between tunes for a given generator. Given the existence of such differences and the potential dependence on event generator models and tunes it is clearly desirable to at least attempt to quantify the extent to which different taggers receive non-perturbative contributions.

Furthermore, it also becomes important to consider designing tools which give high performance without relying on large non-perturbative contributions. In that respect, there have been previous instances of developing such improved tools as for example can be seen in the much smaller non-perturbative contributions to the modified mass-drop tagger (in particular when defined with a z_{cut}), or SoftDrop, when compared to the substantial plain-mass like non-perturbative corrections for tools such as pruning and trimming [12]. Isolating and understanding these tools which systematically show limited non-perturbative contributions is important for the design of future substructure methods.

Given the above discussion, ideally we want a method with high performance and robustness, i.e. with a large signal significance and a non-perturbative correction factor close to 1.

We can then make the following generic observations:

- *Effect of grooming.* We can see from fig. 6.13 that adding grooming improves considerably both the performance and the robustness. Based on what we have discussed before, the improvement in performance comes mainly from the impact on signal efficiency. However it is crucial to impose the Y-splitter constraint on the plain jet instead of the groomed jet, otherwise one only gets a much smaller Sudakov suppression of the QCD background, as discussed in appendix A.1. We should however stress that subleading corrections sometimes come with several transition points in the mass distribution, which can be an issue for practical applications in an experimental context.

- *k_t or mass declustering?* As we have seen in our calculations, even though they lead to the same LL result, the overall analytic structure is found to be much simpler for the case of mass declustering. In particular, the groomed (trimmed or mMDT) and plain jet results are given by the LL result with no additional double-logarithmic contributions in the LL+LL$_y$ approximation. Corrections to that result would be purely single-logarithmic in the jet mass, e.g. coming from multiple emissions. Then, although it is not explicitly shown in the figure, using mass declustering comes with a small gain in performance. We traced it back to the absence of the extra terms between the LL and LL+LL$_y$ results.

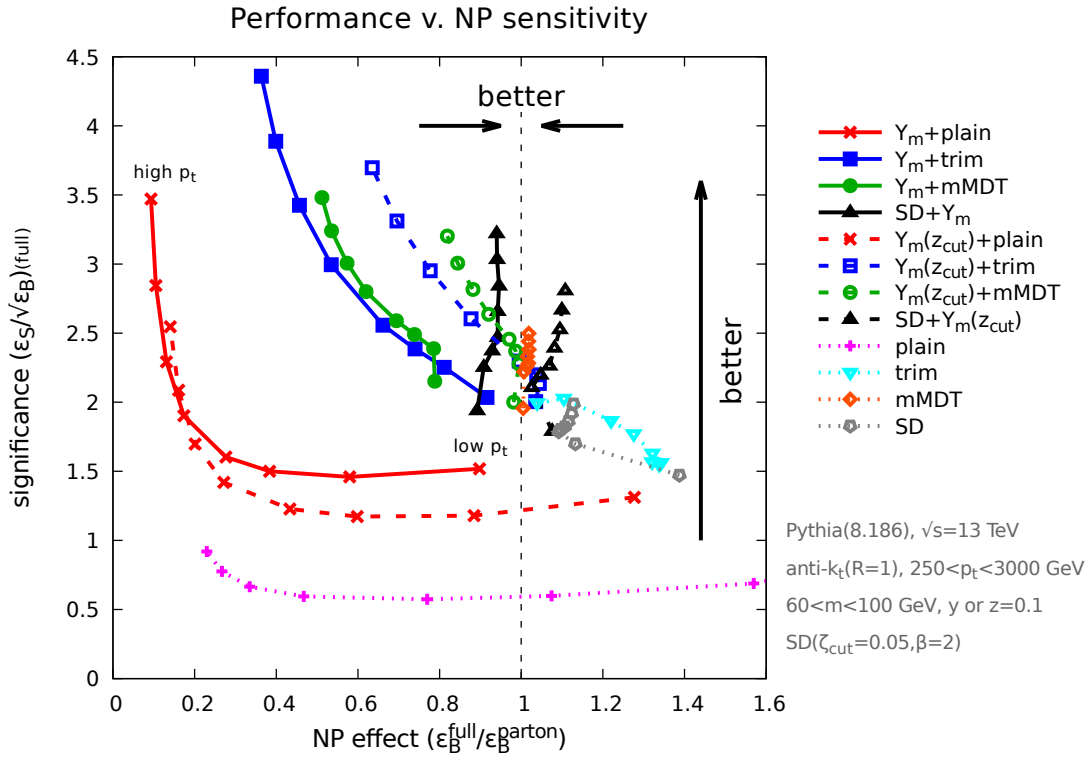


Figure 6.13 — Summary plot showing the signal efficiency, computed as $\varepsilon_S/\sqrt{\varepsilon_B}$ for events at full level, versus the ratio of the background efficiency calculated, for a quark-jet sample, at full level and at parton level. In all cases, we have required that the mass is between 60 and 100 GeV, and signal and background efficiencies are computed wrt the inclusive jet rate for each p_t cut. The different points on each curve correspond to different values of the jet p_t, from 250 GeV to 3 TeV. Each curve represents a specific method: we use either a standard y_{cut} condition (solid lines) or with a z_{cut} condition (dashed lines), with $y_{\text{cut}} = z_{\text{cut}} = 0.1$. Results are presented for a Y_m-splitter condition computed on the plain jet followed by no grooming (red), trimming (blue) or mMDT (green); or with a pre-grooming, *e.g.* SoftDrop with $\beta = 2$ and $\zeta_{\text{cut}} = 0.05$. For comparison, the results without the Y-splitter condition were added (pointed lines).

- *Trimming or mMDT?* At LL accuracy, both give the same perturbative performance. In practice, at large p_t we see that trimming tends to give a slightly better performance and is slightly less robust. It remains to be investigated whether this is generally true or a consequence of our specific choice of parameters (see "A word of caution" below). Even if it was a general observation, it is not obvious that one should prefer trimming over the mMDT. Indeed, we have seen that trimming introduces more transition points (and therefore kinks) in the mass distribution than the mMDT, although they are reduced by the use of $\mathrm{Y_m}$-splitter). These can have undesirable effects in experimental analyses, e.g. for side-band estimates of the backgrounds or if the signal lies on top of a transition point.

- y_{cut} *or* z_{cut}*?* Contrary to the case of k_t vs. mass declustering, the situation is less obvious here: the y_{cut} variant shows a better performance, in part traced back to single-logarithmic effects like multiple emissions, but at the same time the z_{cut} variant appears less sensitive to non-perturbative effects. The choice between the two is therefore again a trade-off between performance and robustness. In terms of the analytic structure of the results, we should point out that the z_{cut} variant is likely more amenable to a higher logarithmic accuracy resummation more than the y_{cut} version. In particular it gives a simple expression for the resummation of multiple emission effects.

- *Pre-grooming.* We see yet again the same trade-off between performance which is globally in favor of $\mathrm{Y_m}$-splitter+grooming, and robustness which is globally in favor of pre-grooming. The differences in performance are explicitly predicted by our analytic results, already at LL accuracy. The differences in robustness are also expected from the fact that SoftDrop cuts out soft-and-large-angle radiation. It is however interesting to notice that compared to the results obtained for mMDT, trimming and SoftDrop alone, the addition of the $\mathrm{Y_m}$-splitter condition still results in a sizeable performance gain.

- *A word of caution.* We should point out that fig. 6.13 was obtained for one specific choice of the free parameters like the jet radius, y_{cut}, z_{cut} or mass-window parameters. In practice, we do not expect to see substantial differences if we were to adopt a different setup, especially for the main features which are backed up by analytic calculations. However, some of the differences observed in fig. 6.13 go beyond our analytic accuracy and can depend on our choice of parameters. This concerns, in particular, the subleading differences observed between trimming and the mMDT, or details about the precise size of non-perturbative effects.

It is worth noting that our analytical calculations alone do not provide the final word when comparing the performance of say Y-splitter with mMDT and Y-splitter with trimming, since these methods differ by subleading perturbative corrections and sensitivity to non-perturbative effects terms we have not attempted to control. With more work subleading terms, such as those beyond leading-logarithmic accuracy, should be calculable

within our theoretical framework. However it is less simple to account for the role of non-perturbative effects, as one needs to rely on event generator tools to quantify them. At the same time a marked dependence of tagger performance on non-perturbative corrections will lead to greater theoretical uncertainty which motivates the study of taggers that are less susceptible to non-perturbative corrections, as the z_{cut} variation or of pre-grooming with Soft-Drop shown in this chapter.

6.5 Final discussion

In this chapter, we have studied analytically the effect of imposing a Y-splitter condition on boosted jets. We also have considered the combination of a Y-splitter cut together with a grooming procedure, namely trimming and mMDT. We concluded that this combination increases Y-splitter performance.

We have also considered variants of the Y-splitter condition: first the standard one defined in terms of a cut on k_t^2/m^2 (known also as a y_{cut} condition), secondly a variant called Y_m-splitter where the k_t scale is computed using a "mass declustering", i.e. by undoing the last step of a generalized-k_t clustering with $p = 1/2$, and finally replacing the standard y_{cut} condition by a z_{cut} condition, where we cut directly on the subjet momentum fractions instead of k_t^2/m^2. One important observations about these variations is that, keeping only the dominant terms enhanced by logarithms of the jet mass at all orders (LL), the same behavior is recovered for all these variants when applied to QCD background jets.

Alternatively, we can apply a more gentle SoftDrop grooming to the jet and then impose the Y-splitter condition and compute the jet mass on that pre-groomed jet. In this case we do observe a change in the leading $\log \rho$ term, which represents a small loss of discriminating power. Nevertheless, the pre-grooming procedure also decreases non-perturbative effects.

From fig. 6.13, we can observe that there is a trade-off between discriminant power and robustness against non-perturbative when comparing different variation options. But in the end, these variants still outperform the standard methods such as pure mMDT, pure trimming or pure SoftDrop at high p_t.

The good performance of the Y-splitter+grooming boosted object tagger comes from the combination of two effects. First, for the Y-splitter without grooming the QCD background is suppressed relative to the case of the plain jet mass. The exponential Sudakov factor, which is usually double-logarithmic in the jet mass, becomes a mass logarithm multiplied by a $\log y_{\text{cut}}$ factor, which is considerably smaller for boosted jets. Second, the use of grooming does not significantly affect this background suppression due to the fact that it induces only subleading corrections to the pure Y-splitter case. Due to the combination of both effects, the use of grooming considerably improves the signal efficiency relative to the pure Y-splitter case.

— 7 —
Jet Shapes

In this chapter we study jet shapes as a discriminant variable between two-pronged hadronic decays of an electroweak boson (W/Z/H) and the background formed by decaying QCD partons. These jet shapes put constraints on the gluon radiation patterns in a jet. We expect them to have a good discriminating power because gluon radiation from color-neutral bosons is different from that of colored QCD jets.

The shapes studied are the N-subjettiness ratio $\tau_{21}^{(\beta_\tau)} = \tau_2^{(\beta_\tau)}/\tau_1^{(\beta_\tau)}$, the energy correlation function $C_2^{(\beta_C)}$ and a variant of the μ^2 parameter of the MDT (see eq. (4.3)). We note that, while the MDT also has an energy cut y_{cut}, in this work we only consider the mass-drop condition. We have fixed the angular parameter $\beta_\tau = \beta_C = 2$, this choice has the advantage of imposing $\tau_1^{(2)} = e_2^{(2)} = \rho = m^2/R^2p_t^2$ and shows good performance in Monte Carlo studies.

We carry out analytical studies for the jet mass distributions for background and signal jets with cuts on shape variables $v < v_{\text{max}}$, with $v = \tau_{21}, C_2$ and μ^2. We apply this cuts in two different ways: either a fixed cut or recursive cut. We work in the limit $\rho \ll 1$ (relevant for boosted object studies) and $v_{max} \ll 1$, which is desirable to separate two-pronged structures from QCD background. We aim only at capturing the (modified) leading-logarithmic behavior, which is sufficient to capture the main differences between the three shapes, although we also commented several sources of next-to-leading logarithmic corrections. We test our analytical results by comparing to fixed-order results from EVENT2 and to results from parton shower Monte Carlo generators. We also present pure Monte Carlo studies of the impact of non-perturbative corrections.

In section 7.1, we discuss the general form of the results obtained. In section 7.2, we perform the detailed calculations for background jets for both non-recursive and recursive variants for each shape variable and we compare the expansion of our results Monte Carlo generators. In section 7.3 we do the same for signal jets. In sections 7.4, we study non-perturbative effects and discuss the impact of a combination with a grooming procedure. Finally, in section 7.5 we discuss the comparative performance of all three shapes and their variations.

7.1 Generic structure of the results

For QCD jets, there are two basic physical quantities that we will be interested in: the jet mass distribution after applying a given fixed, recursive or not, cut on one of the shapes (N-subjettiness, mMDT parameter or energy correlation); or the distribution of a jet shape for a given fixed value of the jet mass. The latter situation only applies to the non-recursive cases.

For signal jets, we are interested in jets of a fixed mass so the calculation will mostly focus on what fraction of these jets satisfy the constraint on the jet shape v, hence on the distribution of v for an object of a given mass. Jets which fail the constraint on v will be discarded.

Our calculations apply to the boosted regime, where the jet transverse momentum is much larger than its mass, so we can take the limit $\rho \ll 1$. As we focus on the leading double logarithm, soft and collinear emissions can be considered as strongly ordered and the mass of the jet is dominated by the strongest of these emissions, which we will refer as p_1. We will assume that this emission occurs at an angle[1] $R\theta_1$ and with a fraction z_1 of the jet total transverse momentum p_t. This has to satisfy the constraint $z_1(1-z_1)\theta_1^2 = \rho$, where, for QCD jets we can neglect the $(1-z_1)$ factor which would only lead to subleading power corrections in ρ.

All the shapes, v, that we consider put constraints on additional emissions. This means that we can always consider, as a starting point, a system made of two partons — the "leading parton p_0" initiating the jet and the "first, leading, emission p_1" which sets the jet mass for QCD jets, or the two prongs of a massive boson decay for signal jets — and study additional radiation from this system.

In the leading-logarithmic approximation, the constraint on radiation will always take the form of a Sudakov suppression coming on top of the mass requirement. For QCD jets, the mass distribution with a cut on v can always be written as

$$\begin{aligned}\frac{\rho}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\rho}\bigg|_{<v} &= \int_\rho^1 \frac{d\theta_1^2}{\theta_1^2}\int_\rho^1 dz_1\, p_i(z_1)\,\rho\,\delta(z_1\theta_1^2-\rho)\frac{\alpha_s(z_1\theta_1 p_t R)}{2\pi}e^{-R_{\text{plain}}(\rho)-R_v(z_1,\rho)}\\ &= \int_\rho^1 dz_1\, p_i(z_1)\,\frac{\alpha_s(\sqrt{z_1\rho}\,p_t R)}{2\pi}e^{-R_{\text{plain}}(\rho)-R_v(z_1,\rho)},\end{aligned} \tag{7.1}$$

where p_i is the corresponding splitting function, depending if it is a quarks or gluon jet. In the above $R_{\text{plain}}(\rho)$ (see eq. (5.36)) is the Sudakov resumming the leading $\log(1/\rho)$ contributions to the plain jet mass and $R_v(z_1,\rho)$ the extra contribution coming from the additional cut on v.

In the approximation we shall be working at, instead of the usual splitting functions $P(z_1)$ (5.2) it is sufficient to consider its leading logarithmic contribution from its $2C_R/z_1$ term and a subleading hard collinear contribution $2C_R B_i\delta(z_1-1)$, where C_R is the color charge of a jet initiated by a parton of flavor i and B_i is the integral of the non-singular part of the corresponding splitting function, already defined in section 5.3. Finally, eq. (7.1)

[1] Practically, it is easier to normalize all angles to the jet radius R.

can therefore be replaced by

$$\frac{\rho}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\rho}\bigg|_{<v} = \int_\rho^1 \frac{dz_1}{z_1}\frac{\alpha_s(\sqrt{z_1\rho}\,p_t R)C_R}{\pi} e^{-R_{\text{plain}}(\rho)-R_v(z_1,\rho)} + \frac{\alpha_s(\sqrt{\rho}\,p_t R)C_R}{\pi} B_i\, e^{-R_{\text{plain}}(\rho)-R_v(z_1=1,\rho)}. \tag{7.2}$$

Note however that keeping the full integration over the splitting function is sometimes useful in comparing background and signal efficiencies and can lead to potentially large subleading corrections.[2] For all the analytic plots in this chapter, where the integration over z_1 is done numerically, we have decided to keep the exact $p_i(z_1)$ splitting function and use eq. (7.1).

If instead we want to obtain the probability to satisfy the cut on the shape v for a jet of a given mass one get (for the non-recursive versions):

$$\Sigma_v(v) = \left[R'_{\text{plain}}(\rho)e^{-R_{\text{plain}}}\right]^{-1} \frac{\rho}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\rho}\bigg|_{<v}, \tag{7.3}$$

with R'_{plain} being the derivative of R_{plain} wrt $\log(1/\rho)$. Note that the shapes we consider all require at least three particles in the jet to be non-zero, meaning that the distribution $\mathrm{d}\sigma/\mathrm{d}\rho|_{<v}$ – or, equivalently, the double-differential distribution in both the mass and the shape, $\mathrm{d}^2\sigma/\mathrm{d}\rho\mathrm{d}v$ – starts at order α_s^2. Conversely, $\Sigma(v)$ will start at order α_s, since it is normalized to the jet mass which itself starts at order α_s.

For the following calculation, we treat logarithms of the shape and the jet mass on an equal footing. Hence, by leading logarithms, we mean, for fixed coupling, double logarithms of any kind, i.e. in either the shape or the jet mass or both. For the figures and the comparisons to Monte Carlo simulations, we will also include the (leading order) running-coupling contributions as well as a few relevant NLL effects. A list of NLL contributions and a more detailed discussion in included in 7.4.

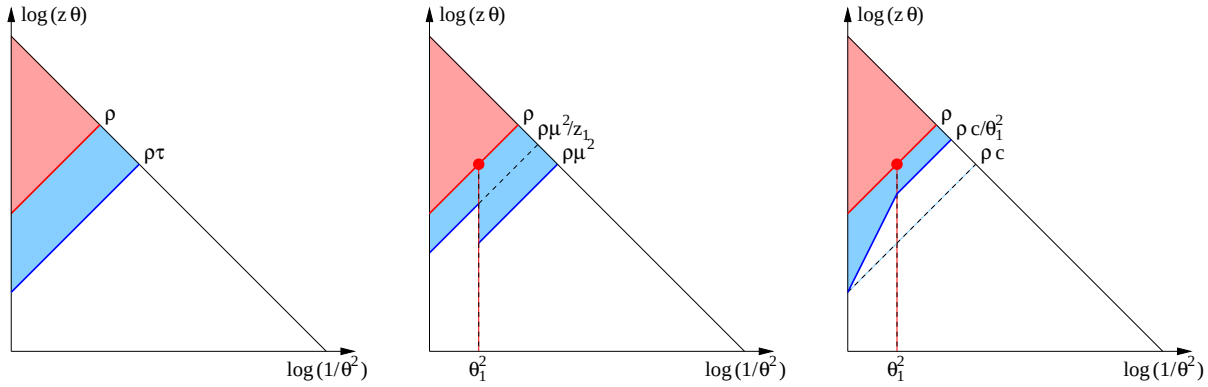
For signal jets, we will directly be interested in the efficiency, i.e. in the fraction of jets (of the original jet mass) that will satisfy the constraint on v. This can be written as

$$\Sigma_{\text{sig}}(v) = \int_\rho^1 dz_1\, p_{\text{sig}}(z_1) e^{-R_{v,\text{sig}}(z_1,\rho)} \tag{7.4}$$

where the signal "splitting function" $p_{\text{sig}}(z_1)$ depends on the process being studied and is assumed to be normalised to unity. Again, we can either decide to keep the full integration over z_1 or, at our level of accuracy, keep only the dominant part without any z_1 dependence and the first $\log(1/z_1)$ and $\log(1/(1-z_1))$ corrections. Note that here z_1 can no longer be neglected in the constraint on the jet mass, $\rho = z_1(1-z_1)\theta_1^2$.

Given these basic expressions, our main task is to compute the Sudakov factors R_v for all the shapes under consideration. We do that in the next two sections.

[2] See also the discussion in section 7.2.7.



(a) N-subjettiness **(b)** Mass-drop **(c)** Energy correlation function

Figure 7.1 — Plots of the phase-space constraints on emissions setting the mass (in red) and the jet shape (in blue).

7.2 Calculations for the QCD background

The results below give the generic expression for the Sudakov form factor assuming one works in the (modified) leading-log approximation , using the basic building blocks formalism. It is helpful to clarify the notations once and for all:

$$\begin{aligned} L_\rho &= \log(1/\rho) = \log(p_t^2 R^2/m^2), & L_\tau &= \log(1/\tau_{21}), \\ L_1 &= \log(1/z_1), & L_\mu &= \log(1/\mu^2), \\ L_v &= \log(1/[\tau_{21},\, \mu^2 \text{ or } C_2]), & L_e &= \log(1/C_2). \end{aligned} \tag{7.5}$$

We assume, as stated before, that the angles are normalized to the jet radius R. For a fixed mass ρ and momentum fraction z_1, we have $\theta_1^2 = \rho/z_1$. We also recall the constants being used, for quark jets, we have $C_R = C_F$ and $B_i = B_q = -3/4$ while for gluon jets we have $C_R = C_A$ and $B_i = B_g = -(11C_A - 4n_f T_R)/(12C_A)$.

7.2.1 τ_{21} cut (pure N-subjettiness cut)

We first consider the case where we impose a cut $\tau_{21} < \tau_{\text{cut}}$ on the N-subjettiness of a jet of a given mass ρ. We are interested in the limit $\tau_{\text{cut}} \ll 1$.[3]

The first step is to find an expression for τ_{21} in the limit where emissions are strongly ordered in angle and transverse momentum fraction. Let us assume there is a second leading emission (in mass) occurs at an angle θ_2, wrt the leading parton p_0, (initiating the jet) and carries a transverse momentum fraction z_2 of the leading parton.[4]

[3] In order to keep the notation as light as possible, we shall drop the "cut" subscript when no confusions are possible, we will also ignore the subscript β in the N-subjettiness and energy correlation function, as they are already fixed.

[4] Note that for N-subjettiness with $\beta = 2$, we do not have to worry about recoil effects and we can focus on the simpler E-scheme recombinations, which uses 4-momentum sum of the particles.

For τ_2, three different situations are possible, depending on how these partons are combined with the N-subjettiness axes:

- one axis coincides with p_0, the other with $p_1 + p_2$, giving $\tau_2^{(0,12)} = z_1 z_2/(z_1 + z_2)\theta_{12}^2$,
- one axis coincides with p_1, the other with $p_0 + p_2$, giving $\tau_2^{(1,02)} = z_2\theta_2^2$,
- one axis coincides with p_2, the other with $p_0 + p_1$, giving $\tau_2^{(2,01)} = z_1\theta_1^2$,

where we have again neglected subleading large-z_i contributions, and θ_{12} is the angle between the first and second emissions.

Since the emission p_1 dominates the mass, we have always $\tau_2^{(2,01)} \gg \tau_2^{(1,02)}$. By the other hand, the ordering between $\tau_2^{(0,12)}$ and $\tau_2^{(1,02)}$ is less clear. When $\theta_2 \gg \theta_1$, $z_2\theta_2^2 \ll z_1\theta_1^2$ imposes $z_2 \ll z_1$; we can then approximate $\theta_{12} \approx \theta_2$ and get $\tau_2^{(0,12)} \approx z_2\theta_2^2$, i.e. both choices $\tau_2^{(0,12)}$ and $\tau_2^{(1,02)}$ are equivalent. In the opposite case, when $\theta_2 \ll \theta_1$, $\theta_{12} \approx \theta_1$ and $\tau_2^{(0,12)} \approx z_1 z_2/(z_1 + z_2)\theta_1^2$. For $z_1 \ll z_2$, we get $\tau_2^{(0,12)} \approx z_1\theta_1^2 \gg z_2\theta_2^2$, while for $z_1 \gg z_2$, we get $\tau_2^{(0,12)} \approx z_2\theta_1^2 \gg z_2\theta_2^2$. [5]

The value of τ_{21} will depend on how we choose out axes for the N-subjettiness. For the cases described in section 4.2, we have

- the *optimal axes* should minimize τ_2 and hence give $\tau_2 = z_2\theta_2^2$.
- for the k_t *axes*, we should therefore find the minimum of $d_{01}^{(k_t)} = z_1\theta_1$, $d_{02}^{(k_t)} = z_2\theta_2$, and $d_{12}^{(k_t)} = \min(z_1, z_2)\theta_{12}$. In that case, we also will find $\tau_2 \approx z_2\theta_2^2$ *except* in a region $z_2\theta_2^2 \ll z_1\theta_1^2$, $z_2\theta_2 \gg z_1\theta_1$, i.e. the region where the emission p_2 has smaller mass but larger k_t than the emission p_1, and where we get $\tau_2 \approx z_1\theta_1^2$.
- for the *gen-k_t(1/2) axes*, we should find the pair that minimises the distance $d_{ij}^{(1/2)} = \min(z_i, z_j)\theta_{ij}^2$. In this case, the minimum will always be d_{02} or d_{12} and yield $\tau_2 = z_2\theta_2^2$.

In the end, the case of k_t *axes* is clearly more complex. In what follows we shall therefore focus on the two other axes choices. Based on considerations similar to the ones above, one can show that the gen-k_t(1/2) axes will agree with the minimal axes up to NNLL corrections (mostly occurring when two angles become comparable or when there is a hard splitting). From a numerical point of view, computing the optimal axes can be an expensive step and we can view the gen-k_t(1/2) option as a simpler alternative reproducing essentially the same performance, in what follows, we shall concentrate on this choice. Both choices lead to following jet shape

$$\tau_{21} = \frac{z_2\theta_2^2}{z_1\theta_1^2}, \tag{7.6}$$

up to corrections which are beyond the LL accuracy we aim for here.

[5] Note that if we target single logarithmic accuracy, we should also worry about the situation where $\theta_2 \approx \theta_1$. In that case, $z_2 \ll z_1$ and $\tau_2^{(0,12)} \approx z_2\theta_{12}^2$. This would give at most a constant-factor correction to τ_{21} and hence only contribute at a NNLL compared to the approximation $\tau_2 \approx z_2\theta_2^2$.

Furthermore, we also have to consider secondary emissions, where the radiation is emitted from the gluon (z_1, θ_1^2) itself. If z_2 denotes the fraction of the (first emitted) gluon energy carried by the extra emission at an angle θ_{12}, with $\theta_{12} < \theta_1$ due to angular ordering, we find

$$\tau_{21}^{\text{sec}} = z_2 \frac{\theta_{12}^2}{\theta_1^2}, \tag{7.7}$$

where the different normalization wrt eq. (7.6) is purely due to z_2 being normalized to the gluon energy fraction z_1.

In the limit of small τ_{21}, additional emissions at smaller mass do not affect the result. The one-gluon emission will thus exponentiate according to eq. (7.1) and we get

$$\begin{aligned} R_\tau(z_1) = &\int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \, \Theta(\rho > z_2\theta_2^2 > \rho\tau) \\ &+ \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_1 z_2 \theta_{12})}{2\pi} \, p_g(z_2) \, \Theta(z_2\theta_{12}^2/\theta_1^2 > \tau), \end{aligned} \tag{7.8}$$

where the first line takes into account emissions from the leading parton p_0 while the second accounts for secondary gluon emissions from the first emitted gluon p_1. The arguments of the strong coupling are given as factors multiplying the "natural" scale of the problem, $p_t R$. The phase-space corresponding to the primary emissions is represented in fig. 7.1a.

This is the most simple result because the phase-space just corresponds to a triangle for the primary emissions and another one for the secondary emissions:

$$\begin{aligned} R_\tau(z_1) = &\, T_{02}(0, L_\rho + L_v; C_R, B_i) - T_{02}(0, L_\rho; C_R, B_i) \\ &+ T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_A, B_g\Big), \end{aligned} \tag{7.9}$$

where the negative term subtracts the Sudakov factor for the plain jet mass which has been factored out in our expressions.

For greater clarity and comparison proposes, we shall also quote results with a fixed coupling approximation. In this case, the final exponent does not depend on z_1 and we find

$$R_\tau^{(\text{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi} \left[L_\tau^2/2 + L_\rho L_\tau + B_i L_\tau\right] + \frac{\alpha_s C_A}{\pi} \left[L_\tau^2/2 + B_g L_\tau\right]. \tag{7.10}$$

7.2.2 μ^2 cut

As for the case of N-subjettiness, we first have to find the value of the MassDrop parameter μ^2. Since μ^2 is defined by undoing the last clustering step, it will depend on the jet algorithm we use to (re-)cluster the jet. The C/A algorithm is a common choice but does not work here. Indeed, undoing the last step of a C/A clustering would separate the emission at the largest angle from the rest of the jet, regardless of the transverse

momentum of that emission. This is not infrared safe, we further discuss infrared-safety in appendix B.1.

Instead, we shall define μ^2 by undoing the last step of a generalised-k_t clustering with $p = 1/2$. The motivation for this is the same as the motivation for the axes choice in the previous section: the generalised-k_t algorithm with $p = 1/2$ follows closely the ordering in mass. To keep things unambiguous, we shall denote by μ_p^2 the mass-drop parameter obtained by undoing the last step of a generalised-k_t clustering with parameter p. The (infrared-unsafe) case of a C/A clustering would correspond to μ_0^2 while we will be interested in $\mu_{1/2}^2$, although the calculation can be performed for any positive p. To avoid any possible confusion, we must stress that this argument only applies to the non-recursive version of the μ^2 parameter and that the recursive application of a μ_p^2 cut is infrared-safe for any p.

In order to find the two subjets, we need to find the minimal distance amongst the gen-$k_t(1/2)$ distances d_{01}, d_{02} and d_{12} which gives the two subjets and $\mu_{1/2}^2$ will be given by the mass of the two particles which have been clustered divided by the total mass of the jet.

Again, we consider the leading parton p_0 and two emissions $p_1(\theta_1, z_1)$ and $p_2(\theta_2, z_2)$ with $z_1\theta_1^2 \gg z_2\theta_2^2$. The smallest distance is either $d_{02} = z_2\theta_2^2$ or $d_{12} = \min(z_1, z_2)\theta_{12}^2$. For $\theta_2 \ll \theta_1$, $\theta_{12} \approx \theta_1$ and $d_{12} \geq z_2\theta_1^2 \gg z_2\theta_2^2$, so that the hard subjet mass is $z_2\theta_2^2$. The opposite case, $\theta_2 \gg \theta_1$ (implying $z_2 \ll z_1$), is more subtle: one has to compare the pairwise clustering distances $d_{02} = z_2\theta_2^2$ with $d_{12} = z_2\theta_{12}^2$, where we have used $\theta_{12} \approx \theta_2$. If we remember that each emission comes with an additional angle, φ_i around the jet axis, the minimum depends on $\varphi_2 - \varphi_1$. In half the cases this will cluster 0 and 1 and giving a subjet mass $z_2\theta_2^2$, in the other half, it will cluster 1 and 2, giving a subjet mass of $z_1 z_2\theta_2^2$. Similar considerations allow one to show that the secondary emissions also have an extra factor z_1 compared to the N-subjettiness case.

In the end we find

$$(z_1\theta_1^2)\mu_{1/2}^2 \approx \begin{cases} \dfrac{z_2\theta_2^2}{z_1\theta_1^2} & \text{for } \theta_2 < \theta_1 \text{ or } (\theta_2 > \theta_1 \text{ and } \theta_2 < \theta_{12}), \\ \dfrac{z_2\theta_2^2}{\theta_1^2} & \text{for } (\theta_2 > \theta_1 \text{ and } \theta_2 > \theta_{12}), \\ \dfrac{z_1 z_2\theta_{12}^2}{\theta_1^2} & \text{for secondary emissions.} \end{cases} \tag{7.11}$$

There is a crucial difference between MassDrop and N-subjettiness: the latter can be seen as $(1/p_t)\sum_{\text{j}\in\text{subjets}} m_j^2/p_{t,j}$ which has an extra $1/p_{t,j}$ compared to $\mu_{1/2}^2$. This leads to different expressions whenever the jet with the largest mass is not the one with the largest p_t. The secondary emissions and large-angle radiations will therefore give additional suppressions for N-subjettiness compared to the MassDrop.

With similar arguments, it is easy to realise that additional emissions with smaller masses will not affect this calculation, so that, at leading-logarithmic accuracy, the lowest order simply exponentiates according to eq. (7.1). The vetoed phase-space for emissions

is represented in fig. 7.1b and we get

$$R_{\mu^2_{1/2}}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \Big\{ \Theta(\theta_2^2 < \theta_1^2)\, \Theta(\rho > z_2\theta_2^2 > \rho\mu^2) \\ + \Theta(\theta_2^2 > \theta_1^2) \Big[\frac{1}{2}\Theta(\rho > z_2\theta_2^2 > \rho\mu^2) + \frac{1}{2}\Theta(\rho > z_2\theta_2^2 > \theta_1^2\mu^2) \Big] \Big\} \\ + \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_1 z_2 \theta_{12})}{2\pi} \, p_g(z_2)\, \Theta(z_1 z_2 \theta_{12}^2/\theta_1^2 > \mu^2). \tag{7.12}$$

In order to present the results in a more clear manner, we will divide the Sudakov exponent into two contributions: R_0 clustered with the main parton and R_1 clustered with the emission setting the mass.

$$R_{\mu^2_{1/2},0}(z_1) = T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_R, B_i\Big) - T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2}; C_R, B_i\Big) \\ + \frac{1}{2} P_2\Big(\frac{L_\rho + L_1}{2}, L_\rho, L_\rho + L_v; C_R, B_i\Big) \\ R_{\mu^2_{1/2},1}(z_1) = \Big[\frac{1}{2} P_2\Big(\frac{L_\rho + L_1}{2}, L_\rho, L_\rho - L_1 + L_v; C_R, B_i\Big) \\ + T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{L_\rho - L_1}{2} + L_v; C_A, B_g\Big)\Big]\, \Theta(L_v > L_1) \tag{7.13}$$

The total Sudakov $R_{\mu^2_{1/2}}$ is the sum of these two contributions. For a fixed coupling approximation, we find

$$R^{(\text{fixed})}_{\mu^2_{1/2}}(z_1) = \frac{\alpha_s C_R}{\pi}\Big[(L_\rho + L_1 + L_\mu)L_\mu/2 + \frac{1}{2}(L_\rho - L_1)(L_\mu - L_1)\Theta(L_\mu > L_1) + B_i L_\mu\Big] \\ + \frac{\alpha_s C_A}{\pi}\Big[(L_\mu - L_1)^2/2 + B_g(L_\mu - L_1)\Big]\, \Theta(L_\mu > L_1). \tag{7.14}$$

7.2.3 C_2 cut

For two strongly-ordered emissions $p_1(z_1, \theta_1)$ and $p_2(z_2, \theta_2)$, such that $z_1\theta_1^2 \gg z_2\theta_2^2$, one finds, for primary emissions,

$$C_2 = \frac{1}{z_1^2\theta_1^4} z_1 z_2 (1 - z_1 - z_2)\theta_1^2\theta_2^2\theta_{12}^2 \simeq \frac{z_2\theta_2^2}{z_1\theta_1^2} \max(\theta_1^2, \theta_2^2) \tag{7.15}$$

which is the same result as the one we obtained in the N-subjettiness case with an extra factor $\max(\theta_1^2, \theta_2^2)$.[6] For secondary emissions, $\theta_{12} \ll \theta_1$, hence $\theta_2 \simeq \theta_1$ and we have (with z_2 measuring the momentum fraction wrt emission 1)

$$C_2 \simeq z_2 \frac{\theta_{12}^2}{\theta_1^2}\theta_1^2 = z_2\theta_{12}^2. \tag{7.16}$$

[6] Contrary to what we have for $\mu^2_{1/2}$ (see Appendix. B.2), eq. (7.15) is continuous for $\theta_1 = \theta_2$. Using the exact expression for θ_{12} in the region $\theta_2 \approx \theta_1$ will therefore not lead to (single) logarithmically enhanced terms.

The corresponding phase-space is represented in fig. 7.1c and gives

$$R_{C_2}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \, \Theta(\rho > z_2\theta_2^2) \\ \Big[\Theta(\theta_2^2 < \theta_1^2) \, \Theta(z_2\theta_2^2\theta_1^2 > \rho C) + \Theta(\theta_2^2 > \theta_1^2) \, \Theta(z_2\theta_2^4 > \rho C)\Big] \\ + \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_1 z_2 \theta_{12})}{2\pi} \, p_g(z_2) \, \Theta(z_2\theta_{12}^2 > C). \tag{7.17}$$

After the integration, we observe the need to disentangle two cases depending on whether we have a contribution from emissions at small angles or not:

$$\begin{aligned} R_{C_2}(z_1) &\overset{L_v<L_\rho-L_1}{=} T_{24}(L_\rho, L_\rho + L_v; C_R, B_i) \\ &\overset{L_v>L_\rho-L_1}{=} T_{02}(0, L_\rho - L_1 + L_v; C_R, B_i) - T_{02}(0, L_\rho; C_R, B_i) \\ &\quad + T_{24}(L_1 + L_v, L_\rho + L_v; C_R, B_i) + T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{3L_1 - L_\rho}{2} + L_v; C_A, B_g\Big). \end{aligned} \tag{7.18}$$

For a fixed coupling approximation, one finds [7]

$$R_{C_2}^{(\text{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi}\Big[L_e^2/2 + (L_e - L_\rho + L_1)(L_1 + B_i)\Theta(L_e > L_\rho - L_1)\Big] \\ + \frac{\alpha_s C_A}{\pi}\Big[(L_e - L_\rho + L_1)^2/2 + B_g(L_e - L_\rho + L_1)\Big]\Theta(L_e > L_\rho - L_1). \tag{7.19}$$

7.2.4 Recursive τ_{21} cut

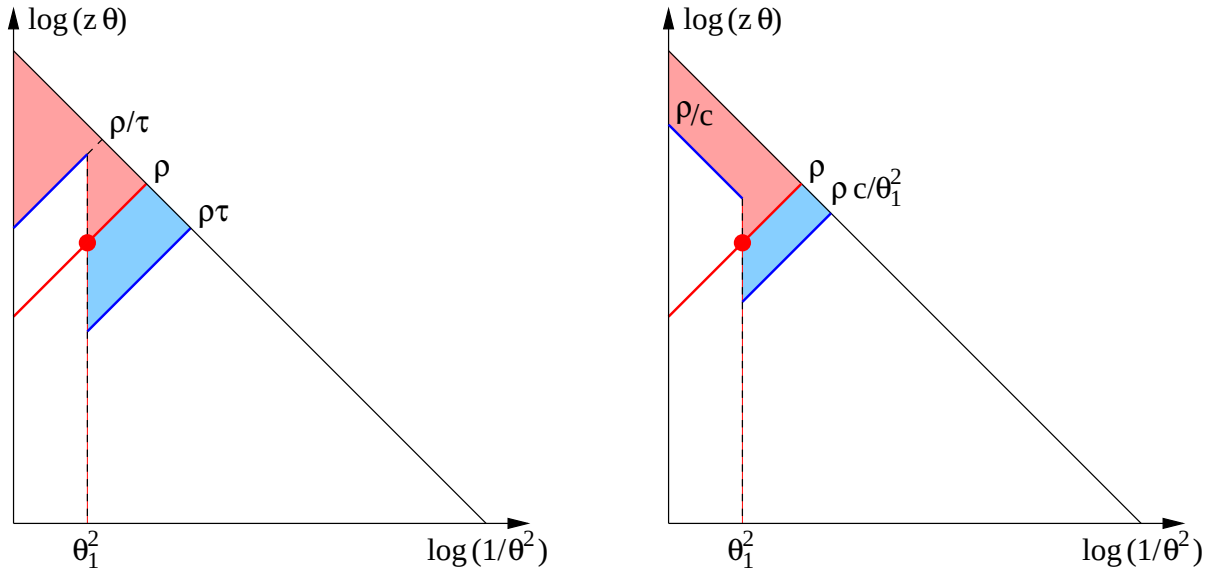
We now move to the same calculations as above but apply the cut recursively. We will undo one step of clustering, with the C/A algorithm. Then we check if the jet pass the cut v_{cut}, is it passes we stop the procedure; if not, we eliminate the softer subjet and repeat.

The calculation of the shapes mostly remains unchanged but the recursion will affect the allowed phase-space for emissions. As before, let us assume that $p_1(\theta_1, z_1)$ is the emission that dominates the mass *after* the recursion procedure has been applied and see what constraints on the phase-space the cut imposes on additional emissions $p_2(\theta_2, z_2)$.

For emissions at angles θ_2 smaller than θ_1, the de-clustering will reach p_1 before p_2, which corresponds to the same situation as for the non-recursive case. In fact it remains true for all the three shape variables considered.

Differences occur for emissions at angles larger than θ_1. The physical reason for that comes from emissions at angles larger than θ_1 and which would dominate the mass, i.e. for which $z_2\theta_2^2 > z_1\theta_1^2$. In the non-recursive case, these emissions are forbidden by our constraint on the jet mass and this is included in the Sudakov suppression for the jet mass $R_{\text{plain}}(\rho)$ in eq. (7.1), which imposes that the mass of the jet is truly dominated by the (z_1, θ_1^2) emission. In the situation where the cut on the shape is applied recursively,

[7] Eqss (7.18) and (7.19) can be trivially expressed as a result of D_2 [70] by replacing L_v by $L_v - L_\rho$.



(a) N-subjettiness or MassDrop **(b)** Energy correlation function

Figure 7.2 — Same as fig. 7.1 but this time for cases where the cut is applied recursively.

some extra care is needed since some of these emissions – that are vetoed in the non-recursive case because they would lead to a larger jet mass – can be simply discarded by the recursive procedure. In such a case, they should no longer be forbidden.

For the large-angle region $\theta_2 > \theta_1$, we have to separate 4 different regions:

- For $z_2\theta_2^2 < \rho\tau$, we have $\tau_{21} \approx z_2\theta_2^2/z_1\theta_1^2 = z_2\theta_2^2/\rho < \tau$, meaning that the constraint is satisfied, this region is therefore allowed.

- For $\rho\tau < z_2\theta_2^2 < \rho$, we have $\tau_{21} \approx z_2\theta_2^2/z_1\theta_1^2 = z_2\theta_2^2/\rho$ as in the previous case, but this time it does not satisfy the condition $\tau_{21} < \tau$. The emission (z_2, θ_2^2) will thus be discarded, meaning that this region is again allowed.

- For $\rho < z_2\theta_2^2 < \rho/\tau$, we now have $\tau_{21} \approx z_1\theta_1^2/z_2\theta_2^2 = \rho/z_2\theta_2^2$, i.e. $\tau_{21} > \tau$. The condition is once again not satisfied and the region is allowed.

- For $z_2\theta_2^2 > \rho/\tau$, we find similarly $\tau_{21} \approx z_1\theta_1^2/z_2\theta_2^2 = \rho/z_2\theta_2^2 < \tau$. The condition on τ_{21} would be met, leaving a jet with a mass $z_2\theta_2^2 > \rho$. This region is therefore forbidden.

Compared to the non-recursive case, the vetoed region at large angle is reduced. Notice that, because the recursive version of the jet shapes only alter the phase space at angles larger than θ_1, contributions coming from secondary emissions are left unchanged.

In the above discussion, we tacitly assumed that we were working with the gen-$k_t(1/2)$ axes or with the optimal axes, but the argument is more general. We could also define τ_{21} using the exclusive C/A axes, automatically available from the declustering procedure.

Indeed, in that case, all emissions with $z_2\theta_2^2 < \rho/\tau$ would fail the cut on τ_{21} and be discarded. We will come back to that point later on.

Again, the lowest order result simply exponentiates and the Sudakov suppression, depicted in fig. 7.2a is

$$\begin{aligned} R_{\tau,\mathrm{rec}}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \Big[&\Theta(\theta_2^2 > \theta_1^2)\, \Theta(z_2\theta_2^2 > \rho/\tau) \\ &+ \Theta(\theta_2^2 < \theta_1^2)\, \Theta(z_2\theta_2^2 > \rho\tau) \Big] \\ + \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_1 z_2 \theta_2)}{2\pi} \, p_g(z_2)\, &\Theta(z_2\theta_{12}^2/\theta_1^2 > \tau) - R_{\mathrm{plain}}(\rho), \end{aligned} \tag{7.20}$$

where we have subtracted $R_{\mathrm{plain}}(\rho)$ which has already been included in (7.1).

The phase-space constraints can take three different forms, depending on the angle of the mass-dominant emission. Remember also that we do subtract the Sudakov factor corresponding to the plain jet mass.

$$\begin{aligned} R_{\tau,\mathrm{rec}}(z_1) \overset{L_v \le L_1}{=}\; & T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_R, B_i\Big) - T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2}; C_R, B_i\Big) \\ & - P_2\Big(\frac{L_\rho - L_1}{2} - L_v, L_\rho - L_v, L_\rho; C_R, B_i\Big) \\ & + T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_A, B_g\Big) \\ \overset{L_1 < L_v < L_\rho}{=}\; & T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_R, B_i\Big) + T_{02}(0, L_v - L_\rho; C_R, B_i) \\ & - T_{02}(0, L_\rho; C_R, B_i) + T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_A, B_g\Big) \\ \overset{L_v \ge L_\rho}{=}\; & T_{02}\Big(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_R, B_i\Big) - T_{02}(0, L_\rho; C_R, B_i) \\ & + T_{02}\Big(\frac{L_\rho + L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_A, B_g\Big). \end{aligned} \tag{7.21}$$

For a fixed coupling approximation, this gives

$$\begin{aligned} R_{\tau,\mathrm{rec}}^{(\mathrm{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi} \Big\{ &\Big[L_\tau^2/2 - L_\rho L_\tau + 2L_1 L_\tau + B_i L_\tau\Big]\, \Theta(L_\tau < L_1) \\ &+ \Big[L_\tau^2 - L_\rho L_\tau + L_1 L_\tau + L_1^2/2 + B_i L_1\Big]\, \Theta(L_1 < L_\tau < L_\rho) \\ &+ \Big[\frac{1}{2}(L_\rho + L_1 + L_\tau + 2B_i)(L_\tau + L_1 - L_\rho)\Big]\, \Theta(L_\rho < L_\tau) \Big\} \\ + \frac{\alpha_s C_A}{\pi} &\Big[L_\tau^2/2 + B_g L_\tau\Big]. \end{aligned} \tag{7.22}$$

7.2.5 Recursive μ^2 cut (pure MassDrop Tagger)

The situation is mostly the same as for the recursive τ_{21} cut. Here, the use of a recursive criterion allows to use either the subjets naturally given by the C/A declustering or

the gen-$k_t(1/2)$ subjets. The results presented in this section are valid for both μ_0^2 and $\mu_{1/2}^2$, although, as we will see in the next paragraph, different axes choice yield the same answer for the mass distribution in different ways, and would give different answers for other observables.

As before, for θ_2 smaller than θ_1, the declustering has no effect and the results are as obtained in sec. 7.2.2. The complication related to the clustering distance for $\theta_2 \gg \theta_1$ is absent here because of the declustering, and only emissions with $z_2\theta_2^2 > \rho/\mu^2$ have to be vetoed. In all other cases, either the MassDrop condition fails and the emission is simply discarded, or the MassDrop condition is satisfied but the mass of the jet remains $z_1\theta_1^2$.[8] For the natural choice, μ_0^2, all emissions in the region $z_2\theta_2^2 < \rho/\mu_0^2$ will fail the condition and be discarded before the recursion continues. That said, the only remaining difference between a recursive μ^2 cut and a recursive τ_{21} cut will be in the extra factor z_1 in the secondary emissions and we find

$$\begin{aligned} R_{\mu^2,\text{rec}}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \Big[&\Theta(\theta_2^2 > \theta_1^2)\, \Theta(z_2\theta_2^2 > \rho/\mu^2) \\ &+ \Theta(\theta_2^2 < \theta_1^2)\, \Theta(z_2\theta_2^2 > \rho\mu^2) \Big] \\ + \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_1 z_2\theta_2)}{2\pi} \, p_g(z_2)\, &\Theta(z_1 z_2 \theta_{12}^2/\theta_1^2 > \mu^2) - R_{\text{plain}}(\rho). \end{aligned} \tag{7.23}$$

The result of this integration is the same as for the recursive N-subjettiness cut (7.21), except that the second argument of the C_A term should be $\frac{L_\rho - L_1}{2} + L_v$ instead of $\frac{L_\rho + L_1}{2} + L_v$ and that term comes with a $\Theta(L_v > L_1)$.

For a fixed coupling approximation, we get

$$\begin{aligned} R^{(\text{fixed})}_{\mu^2,\text{rec}}(z_1) = \frac{\alpha_s C_R}{\pi} \Big\{ &\Big[L_\mu^2/2 - L_\mu L_\rho + 2L_\mu L_1 + B_i L_\mu \Big]\, \Theta(L_\mu < L_1) \\ &+ \Big[L_\mu^2 - L_\mu L_\rho + L_\mu L_1 + L_1^2/2 + B_i L_1 \Big]\, \Theta(L_1 < L_\mu < L_\rho) \\ &+ \Big[\frac{1}{2}(L_\rho + L_1 + L_\mu + 2B_i)(L_\mu + L_1 - L_\rho) \Big]\, \Theta(L_\rho < L_\mu) \Big\} \\ + \frac{\alpha_s C_A}{\pi} &\Big[(L_\mu - L_1)^2/2 + B_g(L_\mu - L_1) \Big]\, \Theta(L_\mu > L_1), \end{aligned} \tag{7.24}$$

where the C_R contribution is the same as for the recursive τ_{21} cut and the C_A contribution is the same as for the non-recursive $\mu_{1/2}^2$ cut.

7.2.6 Recursive C_2 cut

Again, the calculation unfolds as for the two recursive cases above with a contribution from "failed" conditions for $\theta_2 > \theta_1$ and a standard constraint for $\theta_2 < \theta_1$. In the first

[8] As for the axes choice in N-subjettiness, these regions will differ for μ_0^2 and $\mu_{1/2}^2$.

case, e_2 (resp. e_3) is set by emission p_2 (resp. p_1) and $\theta_{12} \approx \theta_2$. In the second case, e_2 (resp. e_3) is set by emission p_1 (resp. p_2) and $\theta_{12} \approx \theta_1$, yielding

$$C_2 = \frac{z_1\theta_1^2}{z_2}\Theta(\theta_2 > \theta_1) + \frac{z_2\theta_2^2}{z_1}\Theta(\theta_2 < \theta_1). \tag{7.25}$$

The Sudakov exponent will ultimately be given by

$$\begin{aligned} R_{C,\mathrm{rec}}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2}\int_0^1 dz_2\,\frac{\alpha_s(z_2\theta_2)}{2\pi}\,p_i(z_2)\Big[&\Theta(\theta_2^2>\theta_1^2)\,\Theta(z_2\theta_2^2>z_1\theta_1^2)\,\Theta(z_2>\rho/C)\\ &+\Theta(\theta_2^2<\theta_1^2)\,\Theta(z_2\theta_2^2>z_1\theta_1^2)\\ &+\Theta(\theta_2^2<\theta_1^2)\,\Theta(z_2\theta_2^2<z_1\theta_1^2)\,\Theta(z_2\theta_2^2>\rho C/\theta_1^2)\Big]\\ +\int_0^{\theta_1^2}\frac{d\theta_{12}^2}{\theta_{12}^2}\int_0^1 dz_2\,\frac{\alpha_s(z_1z_2\theta_2)}{2\pi}\,p_g(z_2)\,\Theta(z_2\theta_{12}^2>C) - R_{\mathrm{plain}}(\rho). \end{aligned} \tag{7.26}$$

Again, we have three different cases [9]

$$\begin{aligned} R_{C_2,\mathrm{rec}}(z_1) &\overset{L_v<L_\rho-L_1}{=} -T_{02}(L_\rho-L_v,L_\rho;C_R,B_i)\Big)\\ &\overset{L_v\le L_\rho}{=} T_{02}\Big(\frac{L_\rho-L_1}{2},\frac{3L_1-L_\rho}{2}+L_v;C_R,B_i\Big) - T_{02}\Big(\frac{L_\rho-L_1}{2},\frac{L_\rho+L_1}{2};C_R,B_i\Big)\\ &\quad - T_{02}(L_\rho-L_v,L_\rho;C_R,B_i) + T_{02}\Big(\frac{3L_\rho-L_1}{2}-L_v,\frac{L_\rho+L_1}{2};C_R,B_i\Big)\\ &\quad + T_{02}\Big(\frac{L_\rho+L_1}{2},\frac{3L_1-L_\rho}{2}+L_v;C_A,B_g\Big)\\ &\overset{L_v>L_\rho}{=} T_{02}\Big(\frac{L_\rho-L_1}{2},\frac{3L_1-L_\rho}{2}+L_v;C_R,B_i\Big) - T_{02}(0,L_\rho;C_R,B_i)\\ &\quad + T_{02}\Big(\frac{L_\rho+L_1}{2},\frac{3L_1-L_\rho}{2}+L_v;C_A,B_g\Big) \end{aligned} \tag{7.27}$$

For a fixed coupling approximation, we obtain

$$\begin{aligned} R^{(\mathrm{fixed})}_{C,\mathrm{rec}}(z_1) = \frac{\alpha_s C_R}{\pi}\Big\{&\Big[-L_e^2/2\Big]\,\Theta(L_e<L_\rho-L_1)\\ &+\Big[(L_v+L_1-L_\rho)(L_v+2L_1-L_\rho+B_i)-L_e^2/2\Big]\,\Theta(0<L_\rho-L_e<L_1)\\ &+\big[(L_e+2L_1-2L_\rho)(L_e+2L_1)/2+B_i(L_e-2L_\rho+2L_1)\big]\,\Theta(L_e>L_\rho)\Big\}\\ +\frac{\alpha_s C_A}{\pi}&\Big[(L_e+L_1-L_\rho)^2/2+B_g(L_e+L_1-L_\rho)\Big]\,\Theta(L_e>L_\rho-L_1). \end{aligned} \tag{7.28}$$

[9] Again, (7.27) can be expressed as a result for D_2 replacing L_v by $L_v - L_\rho$.

7.2.7 Towards NLL accuracy

In the previous chapters we were only aiming to achieve only a (modified) leading-logarithmic description of the shape variables we study here. This level of approximation is enough to capture the main physical features of various jet tagging and grooming tools

Nevertheless it is interesting to extend the scope of our calculations. As the studied jet shapes have some broad similarities, so in order to highlight the differences between these tools it would be helpful to increase the accuracy of the analytical predictions, so that differences that may arise beyond LL effects are effectively highlighted. We would also expect such differences to show up in the Monte Carlo event generator studies, like those carried out below, since event generators would partially capture many sources of subleading corrections.

In this section we discuss several extra ingredients that are required to reach NLL accuracy: soft-and-large-angle contributions, multiple emissions, the two-loop β function for α_s, finite z_1 corrections and non-global logarithms [73]. For the plots where we compare to Monte Carlo simulations, we will include some of these effects: multiple emission effects, two-loop running coupling corrections, and finite z_1 corrections.

We will not include contributions which are power-suppressed in the jet radius R. Although they would be relevant for a full phenomenological prediction, and can be substantial at the peak of the distributions (see section 5 of [84]), these are expected to have little impact when comparing the discriminative power of different jet shapes.

Soft-and-large-angle radiation. A source of single-logarithmic corrections comes from radiating soft gluons at large angles. This would correspond to all the limits beyond the strict collinear ordering that we have adopted until now i.e. it can come from either $\theta_1 \sim R$, or $\theta_2 \sim R$, or $\theta_1 \sim \theta_2$.

The first two regions would give single-logarithmic corrections proportional to R^2. In the small-R approximation we have adopted so far, these would further be suppressed. At the same order of accuracy, one would also have to include contributions coming from initial-state radiation and potential color-correlation with the recoiling partonic system [84]. Taking these into account would also add single-logarithmic contributions to the mass distributions. This significantly complicates the discussion, especially for signal jets, where the mass would no longer be identical to the boosted heavy-boson mass and we would have to impose a certain window around the signal mass. In practice, therefore, one usually applies these techniques together with some grooming procedure which would drastically change this discussion. Some results have already been obtained in [81] for grooming techniques. We present results for jet shapes with SoftDrop grooming in appendix B.5.

The situation for $\theta_1 \sim \theta_2$ is a bit more involved and we show in Appendix B.2 that it would only contribute to single-logarithmic corrections suppressed by θ_1^2. These contributions are also at most proportional to R^2, although since radiation constraints tend to take most of their discriminative power from the large-angle region $\theta_2 > \theta_1$, it makes sense to consider a region $\theta_1 \ll R$. In that case, the contribution from the $\theta_1 \sim \theta_2$ region would be even further suppressed.

Multiple emissions. Multiple gluon emissions also bring single-logarithmic corrections to our results and we briefly discuss below how to account for them for the non-recursive variants of the shapes.

They correspond to cases where several gluon emissions, $(z_2, \theta_2), \ldots, (z_n, \theta_n)$, are only strongly ordered in angle and give similar contributions to the shape v, i.e. when $v(z_2, \theta_2^2; z_1, \theta_1^2) \sim \cdots \sim v(z_n, \theta_n^2; z_1, \theta_1^2)$. This will come with a single-logarithmic correction $\alpha_s^{n-1} L_v^{n-1}$ to the resummed Sudakov exponent.

It is important to realise that we will keep working in the $v \ll 1$ limit and so neglect the contribution where all the $z_i\theta_i^2$, $i \geq 2$, are of the same order as $z_1\theta_1^2$. This would also give a single logarithmic correction of the form $\alpha_s^n L_\rho^n f_n(v)$. Up to power corrections, we can take f_n constant and this correction would therefore simply be equivalent to the multiple-emission correction to the plain jet mass, cancelling against the corresponding normalisation in the spectrum of v.[10] So, from now on, we focus on the region where all the $z_i\theta_i^2$, $i \geq 2$, are much smaller than $z_1\theta_1^2$ and compute the corresponding correction to $R_v(z_1)$ for a fixed z_1.

The case of N-subjettiness and energy correlation functions are mostly straightforward. In the kinematical configurations under consideration, the (optimal or gen-k_t) N-subjettiness axes will still align with the jet axis and with the emission (z_1, θ_1) setting the mass. At a given z_1, both τ_{21} and C_2 will therefore be additive and the correction to $R_v(z_1)$ will be $\gamma_E R_v'(z_1) + \log[\Gamma(1 + R_v'(z_1))]$ where γ_E is the Euler constant and $R_v'(z_1)$ is the derivative of $R_v(z_1)$ wrt L_v.

The situation is a bit more involved for the MassDrop parameter. Had we defined μ^2 as $(m_{j_1}^2 + m_{j_2}^2)/m^2$, μ^2 would have been additive and the similar conclusion as for τ_{21} and C_2 would have been reached. Since μ^2 is defined as a maximum over the two subjets rather than a sum, we should instead use the fact that the condition $\mu^2 < \mu_{\text{cut}}^2$ will be satisfied if both $m_{j_1}^2 < \mu_{\text{cut}}^2 m^2$ and $m_{j_2}^2 < \mu_{\text{cut}}^2 m^2$.

In practice, the emissions will either be clustered with the original hard parton or with the emission setting the mass. How exactly the particles in the jet are sifted in these two sets can depend non-trivially on the details of the clustering. If we take as an approximation, the assumption that particles behave independently, they will be clustered with the hard parton or the emission setting the mass according to which is geometrically closer, in a way similar to the heavy-jet mass in e^+e^- collisions [85]. If we split $R_{\mu_{1/2}^2}(z_1)$ in two contributions according to whether the emissions are clustered with one or the other of the subjets,

$$R_{\mu_{1/2}^2,0}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \Big[\Theta(\theta_2^2 < \theta_1^2) \, \Theta(\rho > z_2\theta_2^2 > \rho\mu^2) + \frac{1}{2} \Theta(\theta_2^2 > \theta_1^2) \Theta(\rho > z_2\theta_2^2 > \rho\mu^2) \Big] \tag{7.29}$$

and

$$R_{\mu_{1/2}^2,1}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, p_i(z_2) \Theta(\theta_2^2 > \theta_1^2) \frac{1}{2} \Theta(\rho > z_2\theta_2^2 > \theta_1^2\mu^2)$$

[10] These type of corrections may however be crucial in trying to obtain the spectrum of v at finite v.

$$+\int_0^{\theta_1^2}\frac{d\theta_{12}^2}{\theta_{12}^2}\int_0^1 dz_2\,\frac{\alpha_s(z_1z_2\theta_{12})}{2\pi}\,p_g(z_2)\,\Theta(z_1z_2\theta_{12}^2/\theta_1^2>\mu^2), \tag{7.30}$$

each of these two parts become additive and we obtain the following correction to $R_{\mu_{1/2}^2}$

$$\gamma_E R'_{\mu_{1/2}^2}(z_1)+\log[\Gamma(1+R'_{\mu_{1/2}^2,0}(z_1))]+\log[\Gamma(1+R'_{\mu_{1/2}^2,1}(z_1))]. \tag{7.31}$$

This is however only an approximation, but even at this stage, it can also be seen as the fact that, compared to N-subjettiness and energy correlation functions, the MassDrop parameter is more delicate to tackle analytically.

Before going to comparisons with Monte Carlo simulations, we can observe that the two axes of 2-subjettiness can be viewed as partitioning the jet in two subjets, one with the jet constituents closer to the hard parton, one with those closer to the emission setting the mass. If instead of summing over all particles in the jet we were summing independently over the contributions of each of the two subjets and defining a modified 2-subjettiness as the maximum of these two contributions, the resummation of multiple emissions for that observable would follow eq. (7.31). However, since $\Gamma(1+R'_0)\Gamma(1+R'_1)/\Gamma(1+R'_0+R'_1)<1$ we should expect this variant of 2-subjettiness to perform worse than its original definition. Conversely, defining the MassDrop parameter as $(m_{j_1}^2+m_{j_2}^2)/m_j^2$ would not only make its analytic behaviour simpler but could also translate into a slightly more efficient tool.

Two-loop running coupling. The inclusion of the two-loop β function is purely a technical complication and they are already included in the basic building blocks defined in chapter 5.3. For simplicity and in order to keep a uniform notation along the thesis the were already directly included in the previous sections.

Non-global logarithms. Non-global logarithms are known to be difficult contributions to handle, especially if we want to go beyond the large-N_c approximation, where a general treatment is still lacking. We will not provide an explicit calculation of their contribution in this paper. We note however that it might be beneficial to apply grooming techniques such as SoftDrop which are known to eliminate the contributions from non-global logarithms, this option is presented in appendix B.5.

Finite z_1 corrections. Finite z_1 corrections would typically give contributions to $R(z_1)$ like $\alpha_s\log(1/v)\log(1/z_1)$ or $\alpha_s\log(1/v)\log(1/(1-z_1))$. The first of these two terms, integrated over the $1/z_1$ part of the splitting function corresponding to the first emission, will give a double-logarithmic contribution that we already have included. The second term, as well as the first term integrated over the non-singular contributions to the $p(z_1)$ splitting function will become important at NLL accuracy. Indeed, after integration over z_1, they would give corrections proportional to $\alpha_s L_v$ which contribute at the single-logarithm accuracy. To properly include these corrections, it is sufficient to integrate over the full $P(z_i)$ splitting function (rather than just including the finite piece as a B_i term) and to keep the full z_1 dependence when we calculate the shapes in order to get single-logarithmic corrections to $R(z_1)$.

The corresponding results are presented in appendix B.3. It is interesting to note that their calculation allows for a nice physical discussion of similarities and differences between background and signal jets. Unless explicitly mentioned, these results will be used for the plots in this paper.

7.2.8 Comparison with fixed-order Monte Carlo

As a partial cross-check of our results, the expressions obtained above can be expanded in a series in α_s and compared to EVENT2 [86, 87] simulations. Here we compare the (non-recursive) τ_{21}, $\mu^2_{1/2}$ and C_2 distributions at order α_s.

Note that since we are using the N-subjettiness implementation from FastJet contrib, we have to use pp coordinates (transverse momentum, rapidity and azimuth) rather than e^+e^- ones (energy and polar coordinates).[11] To maximise the efficiency and provide quark jets with a monochromatic p_t, events are rotated so that their original $2 \to 2$ scattering gives 2 jets at $y = 0$.[12] After that rotation, jets are reconstructed with the standard (pp) anti-k_t algorithm [62] with $R = 0.4$.

On the analytic side, we take the fixed-order results, as running coupling corrections would only enter at order α_s^2, expand (7.3) to first order in α_s, and perform the z_1 integration.

For N-subjettiness, starting from (7.10) we get

$$\tau \frac{d\Sigma(\tau)}{d\tau} = \frac{\alpha_s C_F}{\pi}(L_\rho + L_\tau + B_q) + \frac{\alpha_s C_A}{\pi}(L_\tau + B_g). \tag{7.32}$$

For the mass-drop parameter, we use (7.14) and reach

$$\begin{aligned}
\mu^2 \frac{d\Sigma(\mu^2)}{d\mu^2} &\overset{L_\mu<L_\rho}{=} \frac{1}{L_\rho + B_q}\left[\frac{\alpha_s C_F}{4\pi}\left(3L_\rho^2 + 6L_\rho L_\mu - L_\mu^2 + 4B_q(2L_\rho + L_\mu) + 4B_q^2\right)\right. \\
&\qquad \left. + \frac{\alpha_s C_A}{2\pi}\left(L_\mu^2 + 2B_q L_\mu + 2B_g(L_\mu + B_q)\right)\right] \\
&\overset{L_\mu>L_\rho}{=} \frac{1}{L_\rho + B_q}\left[\frac{\alpha_s C_F}{\pi}\left(L_\rho^2 + L_\rho L_\mu + B_q(2L_\rho + L_\mu) + B_q^2\right)\right. \\
&\qquad \left. + \frac{\alpha_s C_A}{2\pi}\left(2L_\mu L_\rho - L_\rho^2 + 2B_q L_\mu + 2B_g(L_\rho + B_q)\right)\right].
\end{aligned} \tag{7.33}$$

Finally, for the energy correlation function, we start from (7.19) and obtain

$$C_2 \frac{d\Sigma(C_2)}{dC_2} \overset{L_e<L_\rho}{=} \frac{1}{L_\rho + B_q}\left[\frac{\alpha_s C_F}{2\pi} L_e(4L_\rho - L_e + 4B_q) + \frac{\alpha_s C_A}{2\pi} L_e(L_e + 2B_g)\right] \tag{7.34}$$

[11] Alternatively, we could have used an e^+e^- implementation of the jet shapes (and clustering) together with unmodified e^+e^- events. Such an implementation is already readily available in the fastjet-contrib implementation of energy correlation functions. This would however give the same logarithms as in our pp study so we decided to stay with a single coordinate system throughout this paper.

[12] Given the block structure of EVENT2 events, each event can be uniquely associated with a corresponding event with 2 partons in the final state. The latter can be used to define the event rotation. Another approach would be to rotate the event so as to align its thrust axis at $y = 0$.

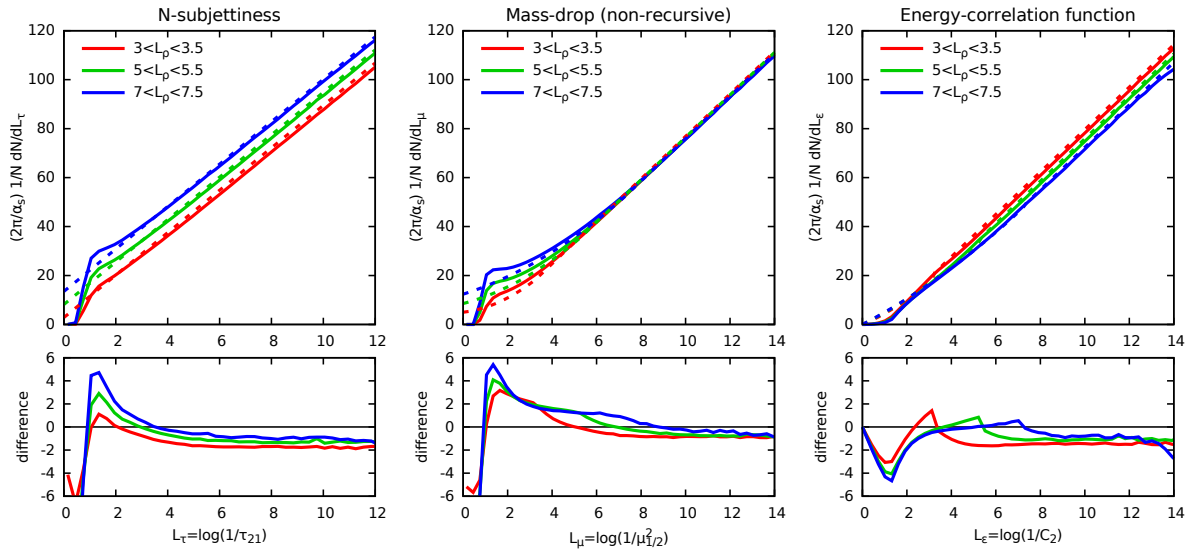


Figure 7.3 — Distributions for the (non-recursive) shapes at order α_s for a few specific bins in the jet mass. A constant factor $\alpha_s/(2\pi)$ has been factored out of the cross-section. The top row shows the distributions themselves, with solid lines corresponding to EVENT2 simulations and dashed lines to our analytic calculation. The bottom row show the difference between the two.

$$\overset{L_e>L_\rho}{=} \frac{\alpha_s C_F}{2\pi}\left(2L_e + L_\rho + B_q\frac{L_\rho+2B_q}{L_\rho+B_q}\right) + \frac{\alpha_s C_A}{2\pi}\left(2L_e - L_\rho + 2B_g - B_q\frac{L_\rho}{L_\rho+B_q}\right).$$

The comparison with EVENT2 is presented in fig. 7.3 where we have plotted the shape distributions at order α_s together with our analytic prediction. In these plots, a constant factor $\alpha_s/(2\pi)$ has been factored out. From fig. 7.3, we see that this difference goes at least to a constant at large L_v, meaning that we do control the leading logarithmic behavior.

In principle, one can also wonder if the constant term can be obtained from an analytic calculation, which is, strictly speaking, beyond our leading-logarithmic accuracy. For example, we have included in equations (7.32)-(7.34) corrections coming from the hard part of the splitting function. However, we have neglected large-angle contributions proportional to R^2 and expected to be small for $R = 0.4$, as well as possible finite z_1 corrections. It is unclear from fig. 7.3 whether or not this fully accounts from the apparent constant value observed at large L_v. In this respect, it is also interesting to note that, contrary to the jet mass where besides the logarithmic and constant terms we would only have power corrections, the constant term in the L_v expansion has some corrections proportional to $1/L_\rho$, coming from the normalisation of the shape distributions by the jet mass cross-section (see eq. (7.3)). These terms can make the convergence slower.

To extract more precise information, we have fitted, in each bin of the jet mass, the coefficient of L_v and the constant term. This has been done in each color channel and reported in fig. 7.4. Again, we see a good agreement for the linear rise with L_v as well as for the constant terms proportional to C_A and N_f. The slow convergence of the C_F term is related to the above discussion.

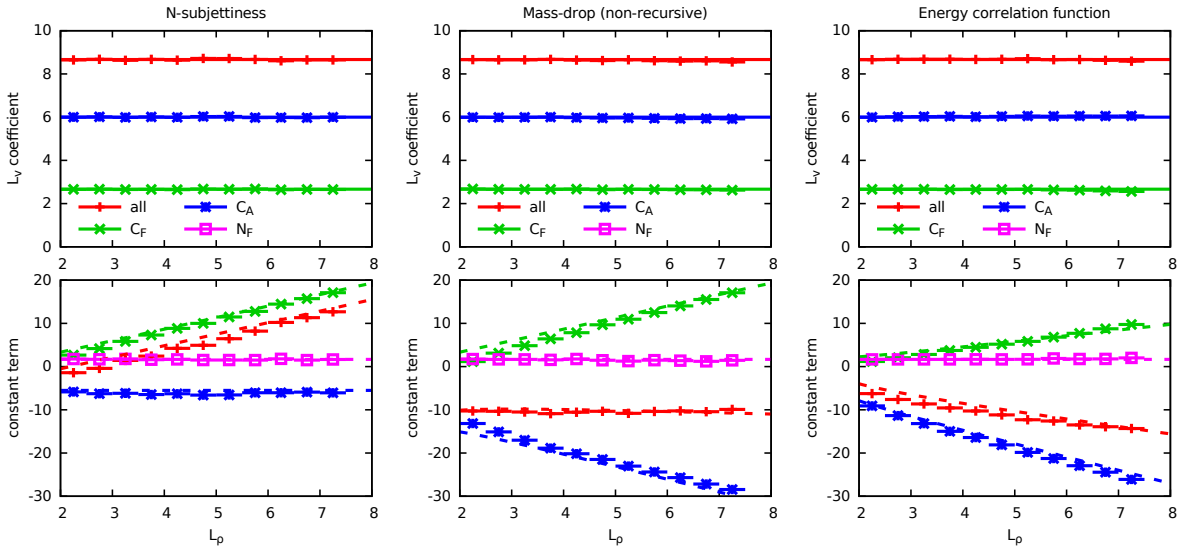


Figure 7.4 — Coefficients of the L_v (top row) and constant (bottom row) terms extracted from the distributions in different bins of the jet mass. For each distribution, we have separated the results in the different color channels. In all cases, a factor $\alpha_s/(2\pi)$ has been factored out of the numbers that are shown.

More precise statements would require going to larger values of L_v and L_ρ. This is difficult to explore due to limited machine precision.

7.2.9 Comparison with parton-shower Monte Carlo

Our resummed analytic results can be directly compared to parton-shower Monte Carlo event generators such as Pythia [35] or Herwig [36]. To do this, we have generated QCD dijet events in 14 TeV pp collisions simulated with Pythia. We have selected anti-k_t (R=1) jets with a transverse momentum of at least 3 TeV.

For our analytical predictions, unless explicitly mentioned otherwise, we have included all the computed global NLL corrections discussed in section 7.2.7. We have fixed $\alpha_s(m_z) = 0.1185$ with $n_f = 5$ and frozen the coupling at $\mu_{fr} = 1$ GeV. Note that Pythia uses a different prescription for the strong coupling, with $\alpha_s(m_z) = 0.1383$ and a 1-loop running. However, our analytic results use the 2-loop β function. It is discussed in appendix E of [69] that it does not alter the results.

In fig. 7.5, we compare the analytic results obtained for the distribution of N-subjettiness, the MassDrop parameter and the energy correlation functions, at a given jet mass, with the same distributions obtained with Pythia at parton-level, including only final-state radiation. First of all, if we look at the large L_v region, where our analytic description is valid, we see that it does reproduce nicely the Pythia simulations. However, at smaller L_v, Pythia tends to produce more peaked distributions than what we obtain analytically.[13] In

[13] Using the prescription from [88] we can replace $R(v)$ by $R(v/(1-v))$ and impose an endpoint, e.g. at $v = 1/2$, which would be the case for N-subjettiness at the order α_s. That would produce distributions

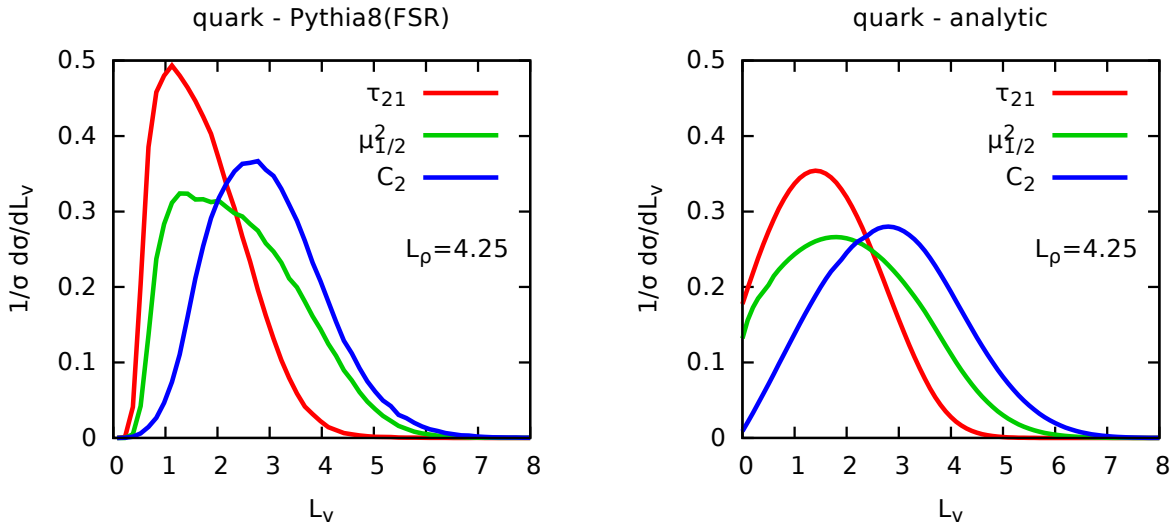


Figure 7.5 — Distributions obtained from quark jets for each of the three shapes studies. Left: results obtained with Pythia including only final-state radiation (we used $p_{t,\text{jet}} > 3$ TeV, and $4 < L_\rho < 4.5$); right: results of our analytic calculations (for $p_t = 3$ TeV and $L_\rho = 4.25$).

any case, the main message that one has to take from this comparison is that the generic ordering between the different shapes is well captured by our analytic calculations.

Instead of plotting the distributions themselves, we can instead look at the mass distributions. This has the advantage that we can also consider the recursive versions of the cuts on the shapes. In fig. 7.6, we plotted the ratio of the mass distribution obtained after a given cut, $L_v > 2.4$, applied recursively (dashed lines) or not (solid lines) on our three shapes, divided by the jet mass distribution without applying any cut. Globally, our analytic calculations tends to reproduce the main features of the Monte Carlo simulations, although they show longer tails at small masses. Note that for these plots, we have used D_2 instead of C_2 since the former peaks at values of L_v closer to the other two shapes. Furthermore, since we have not computed multiple-emission corrections for the recursive versions of the shape constraints, we have also left aside the multiple-emission corrections to the non-recursive versions for the analytic results plotted in fig. 7.6. We notice that including the multiple-emission corrections for the non-recursive shapes tends to reduce the tails towards small mass, bringing more resemblance to the Pythia results. We could expect a similar behaviour for the corresponding recursive versions.

Finally, we want to investigate how the three shapes we have considered are affected by initial-state radiation (ISR) and non-perturbative effects such as hadronisation and the underlying event (UE). To get an insight about the importance of these effects, we

which look much closer to Pythia, although a more detailed resummation of subleading logarithms of ρ (and L_v when if becomes small), and potentially fixed-order corrections (e.g. for secondary emissions) would be needed to draw stronger conclusions.

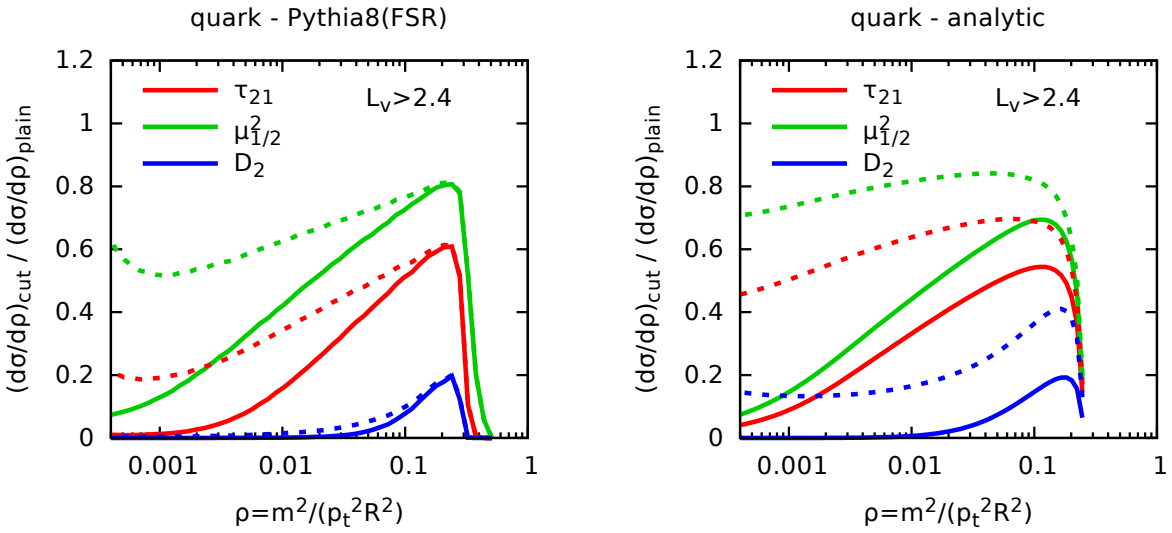


Figure 7.6 — Ratio of the mass spectrum obtained with a cut on one of the shapes, divided by the plain jet mass spectrum. The solid lines are obtained imposing a fixed cut on the jet, while the dashed lines are obtained by imposing the cut recursively. Left: results obtained with Pythia including only final-state radiation (we used $p_{t,\text{jet}} > 3$ TeV, and $L_v > 2.4$ corresponding to $v < 0.09$); right: results of our analytic calculations (for $p_t = 3$ TeV). Note that multiple emissions are not included in these expressions since they have not been computed for the recursive versions.

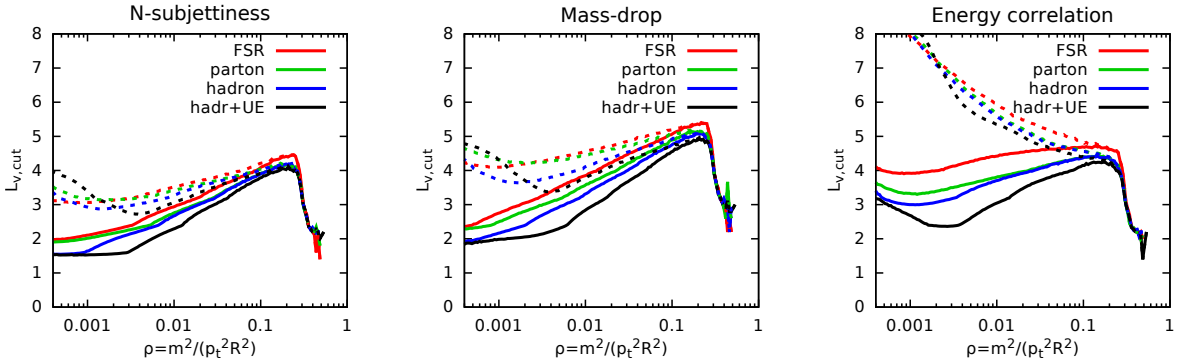


Figure 7.7 — As a function of the jet mass, value of the cut on a given shape, $\log(1/v_{\text{cut}})$ which would correspond to a 25% tagging rate. Results correspond to dijet events obtained with Pythia with $p_{t,\text{jet}} > 3$ TeV. The various curves correspond to different levels of the simulations. The three plots, from left to right, correspond to N-subjettiness, the MassDrop parameter and the energy correlation function.

have looked, for each jet mass, at the cut on L_v that has to be applied to obtain a 25% tagging rate compared to the plain jet mass. This is plotted in fig. 7.7 where we see that, as expected, the cuts are quite sensitive to ISR and the UE, with hadronisation effects

remaining relatively small.

We attribute this behavior to the sensitivity of the shapes to soft and large-angle radiation. We also see that the energy correlation function tends to be more sensitive to these effects than N-subjettiness and the mass-drop parameter. In practice, one would rarely use such a cut without some additional grooming of the jet, limiting the non-perturbative effects at least on the reconstruction of the jet mass. We will come back to this point later, in section 7.4.

7.3 Calculations for the signal

We now turn to the case of signal jets, i.e. jets coming from boosted colorless objects that decay into a $q\bar{q}$ pair (or a pair of gluons), like a W, Z or Higgs boson, or a photon.

As already briefly discussed in sec. 7.1, the splitting of such a boosted object X into a $q\bar{q}$ pair differs from a QCD gluon emission in the sense that it does not diverge as $1/z$ at small transverse-momentum fraction. This means that, although we are still in the regime $\rho \ll 1$ and we shall still consider the limit of small v for all jet shapes v we study in this paper, now $L_1 = \log(1/z_1)$ is no longer large. As for the case of QCD jets, we shall write the results as a function of z_1, see eq. (7.4), but now we will keep the correction in z_1 and $1 - z_1$. These finite z_1 corrections would generate single-logarithmic terms under the form of contributions with one logarithm of z_1 or $1 - z_1$ and one logarithm of ρ or v. It is illustrative to expand out results in series of $\log(1/\rho)$ and $\log(1/v)$ to see explicitly how these terms appear.

Besides the careful inclusion of the z_1 and $1 - z_1$ dependence, the calculation follows the same logic as what has been done above and mostly consists of two copies of the contribution from "secondary emissions" in the QCD case, one for each of the decay products of the boosted colorless object. The contributions from each parton will just differ by the replacement $z_1 \leftrightarrow (1 - z_1)$. For simplicity, we still use $L_1 = \log(1/z_1)$ and additionally introduce $L_- = \log(1/(1 - z_1))$.

Finally, as was already seen to be the case for the secondary emission contributions for QCD jets, the results presented in this section apply invariantly for the recursive or non-recursive versions of the shapes.

7.3.1 τ_{21} cut

Following the same construction as in section 7.2.1, we find that for an emission off the parton carrying a momentum $(1 - z_1)p_t$, we have

$$\tau_{21} = \frac{z_2\theta_2^2}{z_1\theta_1^2}. \tag{7.35}$$

This leads to

$$R_\tau(z_1) = \int_0^{\theta_1^2} \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 \, \frac{\alpha_s(z_2\theta_2)}{2\pi} \, P_q(z_2) \, \Theta(z_2\theta_2^2/\theta_1^2 > z_1\tau_{21}) + [z_1 \leftrightarrow (1 - z_1)], \tag{7.36}$$

where $\theta_1^2 = \rho/[z_1(1-z_1)]$.

After doing this integration one finds

$$R_\tau(z_1) = T_{02}\Big(\frac{L_\rho + L_- - L_1}{2}, \frac{L_\rho + L_- + L_1}{2} + L_v; C_R, B_i\Big) + T_{02}\Big(\frac{L_\rho + L_1 - L_-}{2}, \frac{L_\rho + L_1 + L_-}{2} + L_v; C_R, B_i\Big). \tag{7.37}$$

In order to better highlight the physics behind our calculation, we present the fixed coupling approximation. Keeping only the first non-trivial terms in L_1 and L_-, we find

$$R_\tau^{(\text{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi}\left[L_\tau^2 + (L_1 + L_- + 2B_i)L_\tau\right]. \tag{7.38}$$

7.3.2 μ^2 cut

As for the case of QCD jets discussed in Section 7.2.2, expressions for μ^2 differ from the N-subjettiness ones due to the fact that the p_t normalisations are different.

For an emission off the parton carrying a momentum $(1-z_1)p_t$, we have

$$\mu_{1/2}^2 = \frac{(1-z_1)z_2\theta_2^2}{z_1\theta_1^2}. \tag{7.39}$$

This leads to

$$R_{\mu_{1/2}^2}(z_1) = \int_0^{\theta_1^2}\frac{d\theta_2^2}{\theta_2^2}\int_0^1 dz_2\,\frac{\alpha_s(z_2\theta_2)}{2\pi}\,P_q(z_2)\,\Theta(z_2\theta_2^2/\theta_1^2 > z_1/(1-z_1)\mu_{1/2}^2) + [z_1 \leftrightarrow (1-z_1)] \tag{7.40}$$

Note that formally the Θ constraint above will result in the condition $\Theta(\mu^2 < (1-z_1)/z_1)$ but this will only lead to power corrections in μ^2 and can hence be neglected.

The only difference between N-subjettiness and a $\mu_{1/2}^2$ cut lies in the z_1 and $1-z_1$ corrections. We find

$$R_{\mu_{1/2}^2}(z_1) = T_{02}\Big(\frac{L_\rho + L_- - L_1}{2}, \frac{L_\rho - L_- + L_1}{2} + L_v; C_R, B_i\Big)\,\Theta(L_v > L_- - L_1) + T_{02}\Big(\frac{L_\rho + L_1 - L_-}{2}, \frac{L_\rho - L_1 + L_-}{2} + L_v; C_R, B_i\Big)\,\Theta(L_v > L_1 - L_-) \tag{7.41}$$

For a fixed coupling approximation the extra contributions from the two legs thus cancel, giving

$$R_{\mu_{1/2}^2}^{(\text{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi}\left[L_\mu^2 + 2B_i L_\mu\right]. \tag{7.42}$$

Note that in the case of the signal, the calculation for μ_0^2 would lead to the same result. However, other effects like soft and large-angle gluon emissions that we have neglected here would appear at the same order and lead to an infrared divergence for μ_0^2.

7.3.3 C_2 cut

This time for emissions off the parton carrying a momentum $(1-z_1)p_t$, we find

$$C_2 = \frac{\rho}{z_1^2(1-z_1)}\, z_2\, \frac{\theta_2^2}{\theta_1^2}. \tag{7.43}$$

This leads to

$$R_{C_2}(z_1) = \int_0^{\theta_1^2} \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2\, \frac{\alpha_s(z_2\theta_2)}{2\pi}\, P_q(z_2)\, \Theta\!\left(\frac{z_2\theta_2^2}{\theta_1^2} > \frac{z_1^2(1-z_1)C_2}{\rho}\right) + [z_1 \leftrightarrow (1-z_1)] \tag{7.44}$$

Again, the expression for C_2 looks very similar, except for the logarithms involving z_1. We find [14]

$$R_{C_2}(z_1) = T_{02}\!\left(\frac{L_\rho + L_- - L_1}{2}, \frac{3L_- + 3L_1 - L_\rho}{2} + L_v; C_R, B_i\right) \Theta(L_v > L_\rho - L_- - 2L_1) \tag{7.45}$$

$$+ T_{02}\!\left(\frac{L_\rho + L_1 - L_-}{2}, \frac{3L_1 + 3L_- - L_\rho}{2} + L_v; C_R, B_i\right) \Theta(L_v > L_\rho - L_1 - 2L_-)$$

For a fixed coupling approximation, we get

$$R_{C_2}^{(\text{fixed})}(z_1) = \frac{\alpha_s C_R}{\pi}\left[(L_e - L_\rho)^2 + (3L_1 + 3L_- + 2B_i)(L_e - L_\rho)\right]\, \Theta(L_e > L_\rho). \tag{7.46}$$

Again, formally the extra factor $z_1^2(1-z_1)$ will enter in the $\Theta(L_e > L_\rho)$ condition but its effect is only power corrections and then can be neglected.

7.3.4 Integration over the z_1 splitting

For most of the splitting relevant for phenomenological studies, the splitting function in terms of z_1 is expressed as $z_1^k(1-z_1)^k$ or as a linear combination of such terms (typically, only $k=0$ and $k=1$ are needed for $W/Z/H$ or photon signals).

Introducing $B_2(x) = B(x,x) = \Gamma^2(x)/\Gamma(2x)$, the integration over z_1 can be performed in the fixed-coupling approximation, using

$$\int_0^1 dz_1\, z_1^k(1-z_1)^k \exp\left(-\frac{\alpha_s C_R}{\pi}\, p\, L_v\, (L_1 + L_-)\right) = B_2\left(1 + k + \frac{\alpha_s C_R}{\pi}\, pL_v\right), \tag{7.47}$$

with p a number varying from one shape to another.

7.3.5 Comparison with fixed-order Monte Carlo

Similarly to what was presented in section 7.2.8 for QCD jets, we can compare our results with EVENT2 simulations. In this case, we boost the event along the z axis and rotate it to obtain boosted photons decaying to a jet at $y = 0$.[15]

[14] Expressions (7.45) and (7.46) can be trivially expressed as a result for D_2 replacing L_v by $L_v - L_\rho$.

[15] It appears that the exact outcome depends on the value used for the EVENT2 parameter `metype`, referring to the matrix elements. Set to 1, our default here, we recover the expected situation of a boosted photon. Set to 0, it behaves like a boosted scalar particle, i.e. with a z-independent splitting function.

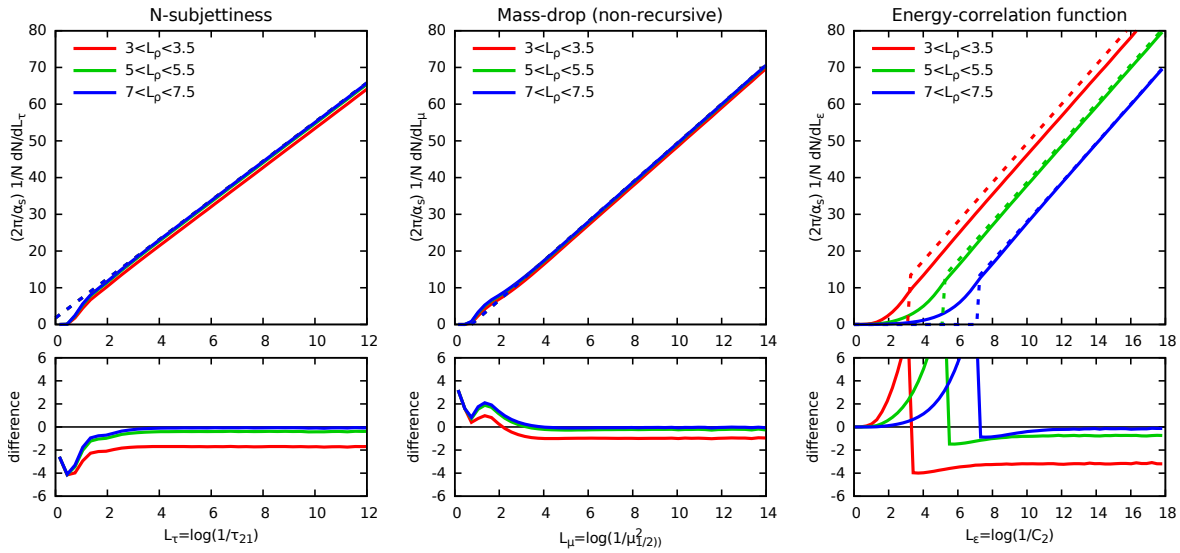


Figure 7.8 — Distributions for the (non-recursive) shapes at order α_s for a few specific bins in the jet mass for the hadronic decay of a Z boson. A constant factor $\alpha_s/(2\pi)$ has been factored out of the cross-section. The top row shows the distributions themselves, with solid lines corresponding to EVENT2 simulations and dashed lines to our analytic calculation. The bottom row shows the difference between the two.

The expansion of the above results to first order in α_s gives, after integration over z_1

$$\tau \frac{d\Sigma(\tau)}{d\tau} = \frac{\alpha_s C_F}{\pi}(2L_\tau + 2B_q + a_\gamma), \tag{7.48}$$

$$\mu^2 \frac{d\Sigma(\mu^2)}{d\mu^2} = \frac{\alpha_s C_F}{\pi}(2L_\mu + 2B_q), \tag{7.49}$$

$$C_2 \frac{d\Sigma(C_2)}{dC_2} = \frac{\alpha_s C_F}{\pi}\Big(2(L_e - L_\rho) + 2B_q + 3a_\gamma\Big)\Theta(L_e > L_\rho). \tag{7.50}$$

In the above expressions, $a_\gamma = \frac{3}{2}a_0 - \frac{1}{2}a_1 = \frac{13}{6}$ with $a_0 = 2$ and $a_1 = \frac{5}{3}$.

The comparison of these analytic results with EVENT2 simulations is presented in fig. 7.8 and shows a good agreement. It is also interesting to notice that the convergence seems faster than it was for QCD jets, probably due to the fact that here the jet mass is fixed.

7.3.6 Comparison with parton-shower Monte Carlo

As for the case of the QCD background jets, we want to compare our analytic calculations to parton-shower Monte Carlo simulations. This time, we used Pythia to generate ZZ events with both Z bosons decaying to hadrons. To match the jet selection of section 7.2.9 in the case of QCD jets, we have selected anti-$k_t(R = 1)$ jets with $p_t \geq 3$ TeV and artificially varied the mass of the Z boson to scan over the ρ range.

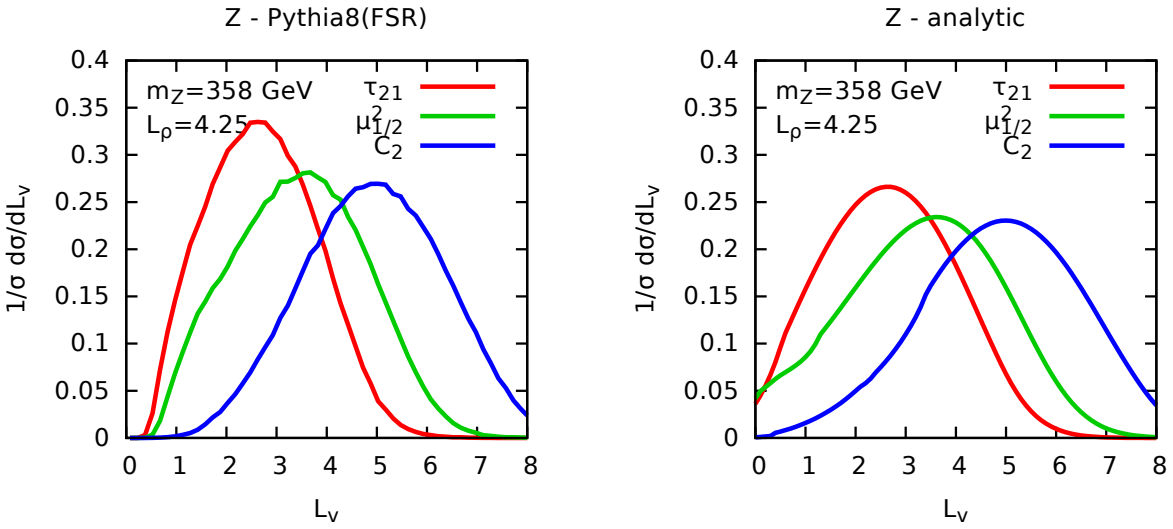


Figure 7.9 — Distributions obtained from $Z \to q\bar{q}$ jets for each of the three shapes studies. Left: results obtained with Pythia including only final-state radiation (for $4 < L_\rho < 4.5$); right: results of our analytic calculations (for $L_\rho = 4.25$).

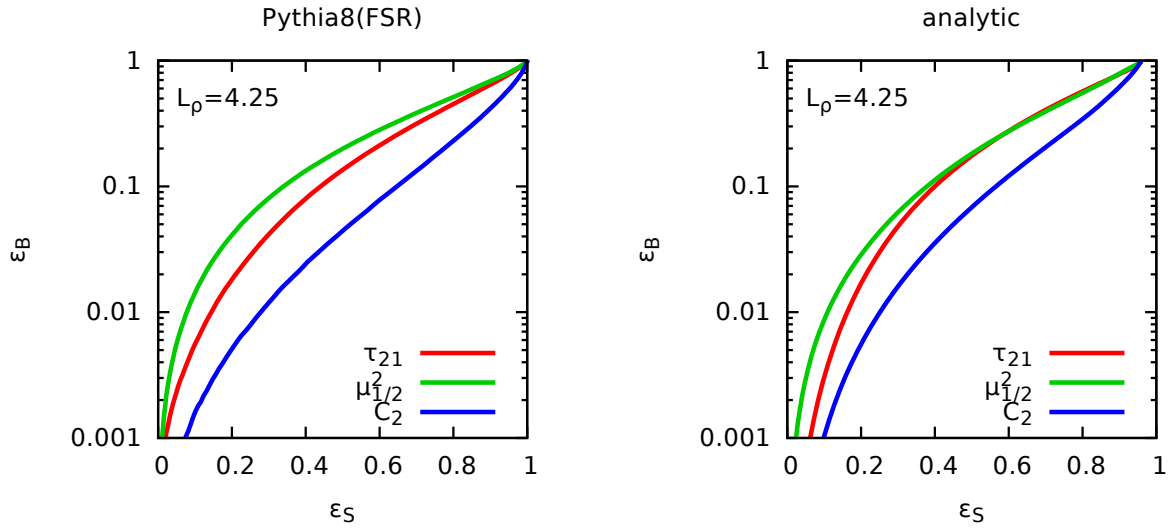


Figure 7.10 — ROC curves showing the background fake rate as a function of the signal efficiency obtained from $Z \to q\bar{q}$ jets for each of the three shapes studies. Left: results obtained with Pythia including only final-state radiation (for $4 < L_\rho < 4.5$); right: results of our analytic calculations (for $L_\rho = 4.25$).

The distributions obtained for the shapes are plotted on fig. 7.9 for Z bosons decaying hadronically. As for the case of QCD jets, we see a good overall description of the features of the distributions and of the differences between the three shapes, particularly in the

large L_v region which is targeted by our calculation.

Based on the results for both the signal and the QCD background, we have plotted a set of ROC curves on fig. 7.10 obtained by varying the cut on the three shapes for a given value of the jet mass. Note that here, the signal and background efficiencies are normalised to the sample of jets that are within the mass window under investigation. The main result here is that a cut on the energy correlation function is more efficient at rejecting the QCD background than a cut on N-subjettiness, itself performing a bit better than a cut on the MassDrop parameter. This behavior is clearly seen in both the Pythia simulations and our analytic calculations.[16] We leave a detailed discussion of this comparison for section 7.5.

7.4 Non-perturbative effects and combination with grooming

We have already seen in section 7.2.9 and in fig. 7.7 that initial-state radiation and non-perturbative effects can have a large impact on the shapes we have studied. One difficulty in trying to assess these effects is that they do not only affect the different shapes we are interested in but also the jet mass and hence our selection of a sample of jets with a mass lying within a given window.

To make a physically meaningful comparison, we have to adapt our normalisation of the background and signal efficiencies compared to what we used to produce fig. 7.10. Instead, we shall now compute the efficiencies as the fraction of the jets passing the initial p_t cut which satisfy both the constraint on the mass and the constraint on the shape. In such a case, as the cut on the shape increases, the signal and background efficiencies progressively increase to ultimately reach an endpoint, common to all shapes, where just the cut on the mass is effective.

As before, we work with anti-k_t jets with $R = 1$ and impose a p_t cut of 3 TeV. For the signal, we used a massive Z' boson with a mass of 217 GeV and impose the constraint on the mass that $5 < \log(p_t^2 R^2/m^2) < 5.5$.[17] Here the background is taken as quark-only to match with the results presented in the previous sections.

The top row of fig. 7.11 shows the ROC curves obtained for our three shapes starting from events including only final-state radiation effects at parton level (in red) and adding successively initial-state radiation (in green), hadronisation effects (in blue) and the underlying event (in black). We clearly see large deviations from what we observe for pure FSR results, noticeably when adding initial-state radiation and the underlying event. Concentrating on the endpoint of these curves, where the cut on the shapes has no effect, we see that these effects are already present when applying the initial mass cut.

In practice, when working with large-R jets, one usually first applies a grooming procedure in order to obtain, at the very least, a good resolution on the jet mass. The

[16] We show in appendix B.4 that this remains valid for less boosted jets, e.g. with $p_t = 500$ GeV.

[17] Working with the nominal Z mass would bring us yet closer to the non-perturbative region and increase even further the effects observed here.

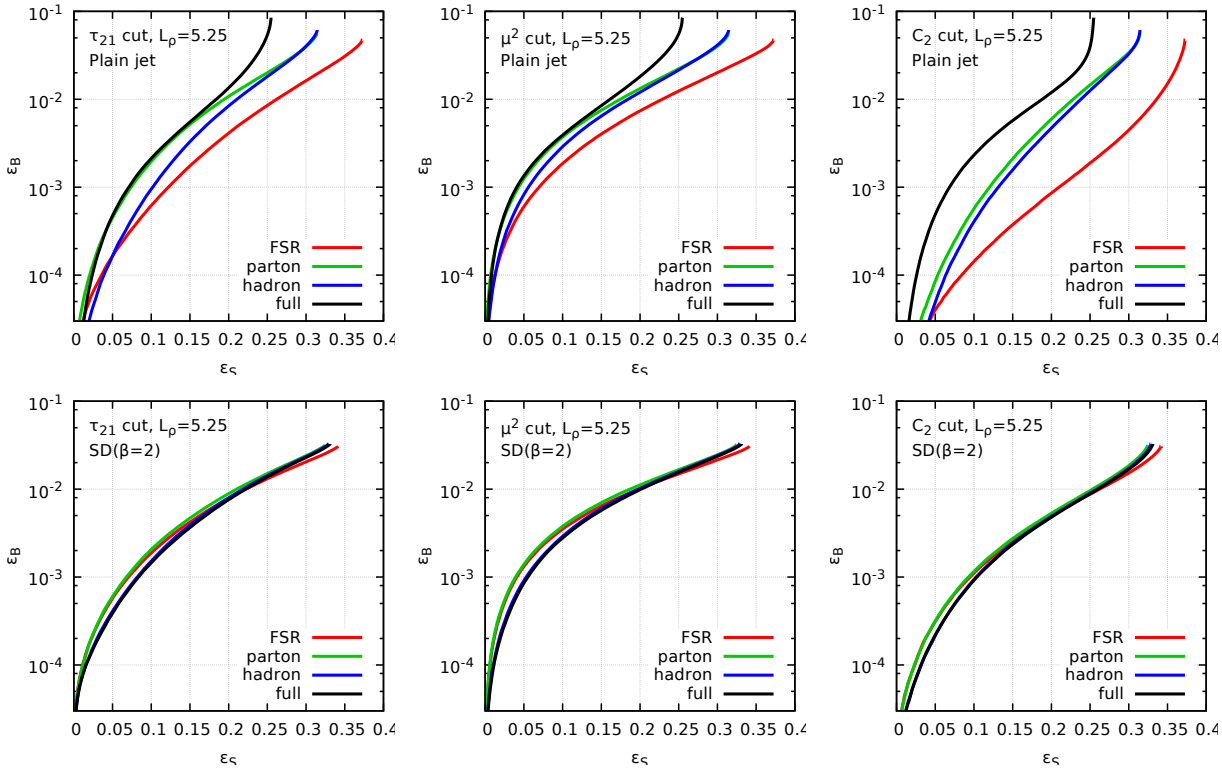


Figure 7.11 — Effects of the initial-state radiation (green), hadronisation (blue) and underlying event (black) on the ROC curves, compared to pure final-state radiation (red). In all cases, we impose that $5 < \log(p_t^2 R^2/m^2) < 5.5$. The left, central and right columns correspond to τ_{21}, $\mu^2_{1/2}$ and C_2, respectively. For the top row, the mass and shape constraints are imposed on the plain, ungroomed, jet. For the plots on the bottom row, we have first applied a SoftDrop procedure with $\beta = 2$ and $z_{\text{cut}} = 0.1$ before imposing the mass and shape constraints.

bottom row of fig. 7.11 shows the same plot as on the top row, now obtained by first grooming the jet with the SoftDrop procedure [17], using $z_{\text{cut}} = 0.1$ and $\beta = 2$, before imposing the cut on the mass and on the shapes. Although this reduces the performance observed on events with pure final-state radiation, this has two positive effects: (i) it stabilises remarkably the ROC curves against initial-state radiation and non-perturbative effects, and (ii) at full parton level it even gives better performance than without the grooming procedure. Again, the ordering between the three shapes remains the same, albeit with strongly reduced differences compared to the plain jet case.

Despite the fact that the results in this section only depend on numerical simulations, as the goal is to study non-perturbative effects, it is rather straightforward to compute LL distributions with SoftDrop using the basic building blocks formalism. In appendix B.5 we present the Sudakov factors exponents in Section 7.2 for QCD jets with a SoftDrop grooming. The results for the signal jets remain the same, which is explained with more detail in the same appendix.

7.5 Discussion

In this chapter, we have provided a first-principles comparison of the performance of three common jet-shapes : N-subjettiness, the MassDrop parameter and energy correlation functions. In order to ensure infrared safety, we have defined the mass-drop parameter based on the subjets obtained via a clustering with the generalized k_t with $p = 1/2$. Similarly, for N-subjettiness, we find that using the exclusive gen-$k_t(p = 1/2)$ algorithm is an efficient alternative to the more complicated optimal axes. The usage of the gen-k_t algorithm is closely connected to the fact that it respects the ordering in mass, which is helpful in our situation where we work at a fixed jet mass and study shapes that have a mass-like behavior.

The main observation from our analytical results and simulations involving only final-state radiation is that there appears to be a clear ordering in the discriminating power of the shapes we have studied: the energycorrelation function ratio is more powerful than the N-subjettiness ratio which, in turn, is more powerful than the μ^2 parameter.

Our results indicate a Sudakov suppression of both the signal and the background for $v \ll 1$. This suppression is however more powerful for the background for two major reasons. Recall that, since we work at a fixed jet mass, both the QCD jets and the signal jets can be seen as two-pronged objects.[18] A cut on the shape thus constrains additional radiation from that system. Given that, discrimination power comes from constraints on radiation at angles smaller and larger than the opening angle between the two prongs. For large angles, the cut on the shape only affects the background due to the color-singlet nature of the signal. At small angles, the radiation from each of the two prongs is proportional to their color factors, which tend to be larger for QCD jets, involving

[18] Strictly speaking, this is only true in the strongly-ordered limit, relevant in the small v context considered in this paper (up to NLL in L_v). For more generic situations, one would also have to consider multi-pronged QCD jets.

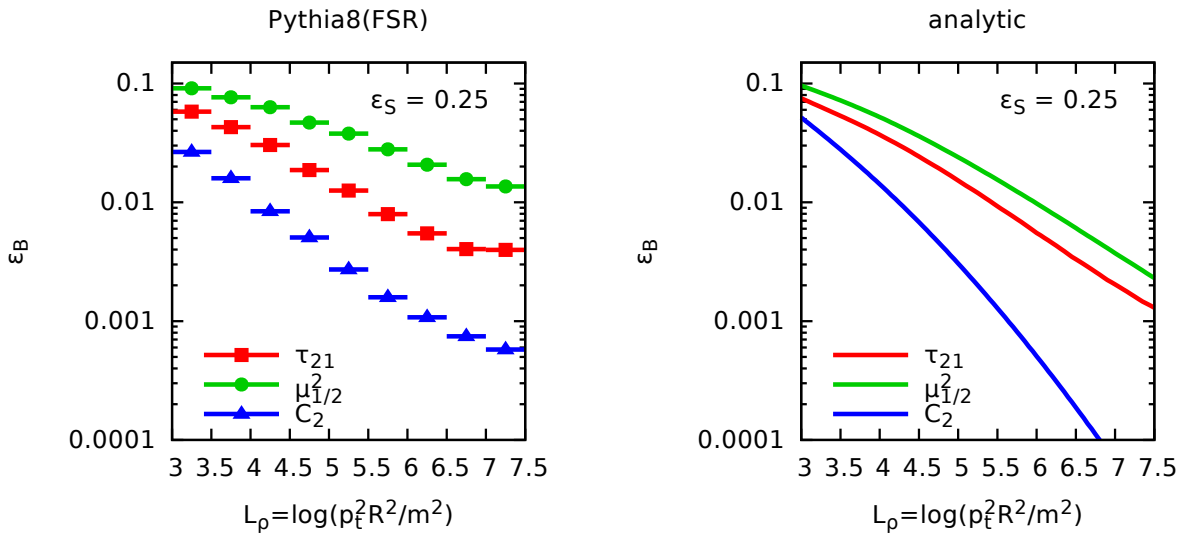


Figure 7.12 — Background fake rate for a 25% signal efficiency as a function of the jet mass. As above, we used $R = 1$ and $p_{t,\text{jet}} > 3$ TeV for the Pythia simulation (left plot) and $p_t = 3$ TeV, for the analytic calculation (right plot).

gluons in their two-prong decay, than for resonances mostly decaying to quarks.[19] Since quark-gluon discrimination exploiting differences in color factors only lead to moderate discrimination power[20, 89, 90, 9], we expect that the large-angle effect would be the main source of difference in tagging two-body decays.

The ordering in discrimination power between the different shapes can also be understood from a different viewpoint. Say we work at a given signal efficiency. The corresponding cut on the shape would determine the constraints on small-angle radiation for both the signal and the background (up to color-factor effects discussed above). Once this is fixed, one has to look at the constraint put on the large-angle radiation for QCD jets. In that region, it is clear from our results, that the radiation veto imposed by a cut on C_2 is more constraining than that imposed by a cut on τ_{21}, itself more constraining than a cut on μ^2. This effect can be directly observed in fig. 7.1, comparing the large angle region of each Lund diagram.

This statement can be made more quantitative from our analytic results. First, the difference between τ_{21} and μ^2 mostly comes from the large-angle region where gluon emissions are clustered with the gluon setting the mass. The extra z_1 factor in the expression for μ^2 compared to τ_{21}, see eq. (7.6) vs. (7.11), results in a smaller vetoed region for μ^2. Parametrically, this region scales like $\alpha_s \log(1/\theta_1^2) \log(1/v) \propto \alpha_s \log(1/\rho) \log(1/v)$. This can be deduced algebraically from our results by fixing the signal efficiency and computing the background for the corresponding cut (with additional $\alpha_s \log^2(1/v)$ terms also coming from the small-angle region).

In the case of C_2, the constraint at large angle now becomes proportional to θ_2^4, see

[19] This argument would be reversed for resonances decaying to gluons.

eq. (7.15), and this translates into an additional vetoed region compared to τ_{21}, which is proportional to $\alpha_s \log^2(1/\theta_1^2) \propto \alpha_s \log^2(1/\rho)$. In conclusion, we expect the ordering between the shapes to be more visible when increasing the boost of the jet. This difference should also grow faster with p_t/m when comparing C_2 and τ_{21} than for τ_{21} and μ^2. This is indeed what is observed from both pure-FSR Monte Carlo studies and from our analytic calculations, as seen in fig. 7.12, where we have plotted the background rejection rate for a 25% signal efficiency as a function of $\log(1/\rho) = \log(p_t^2 R^2/m^2)$.[20]

The next important observation is that, without grooming, the shapes are significantly affected by ISR and non-perturbative effects, UE in particular. These model-dependent effects can be substantial enough to wash out or even invert the differences between the shapes observed from pure FSR and analytic studies (see e.g. the top row of fig. 7.11). This is due to the impact of these effects on both the mass resolution for the jet — mostly for signal jets — and the sensitivity of the shapes themselves. Since ISR and UE mostly affect the soft-and-large-angle region, we expect C_2 to be more affected than τ_{21}, itself more affected than μ^2 (see the discussion above) and this is indeed what we observe from Monte Carlo studies.

The arguments above can be applied when comparing the recursive and non-recursive versions of the shapes: the recursive versions have a smaller vetoed region at large angle while retaining the same small-angle region as their corresponding non-recursive version. Thus, although the recursive versions have the advantage of being less sensitive to ISR and non-perturbative effects, they have a smaller discriminating power.

Furthermore, we have seen that applying a grooming procedure on the jet before computing its mass and values of the shapes largely improves the robustness against ISR and non-perturbative effects, also restoring the ordering between the shapes observed with pure FSR. Again, this can be interpreted as grooming cutting away a part of the soft-and-large-angle region. This increased robustness however comes at a price in that reducing the soft-and-large-angle region using grooming also reduces the discriminating power of the shape cuts. In practice, there will be a trade-off between sheer efficiency and robustness against model-dependent effects.

In addition, note that working at a fixed jet mass ensures that our results are infrared- and collinear safe because it fixes automatically the value of τ_1 and e_2. If we were to impose a cut on the shapes without fixing the jet mass, our results would still be finite after integration of (7.1) over ρ because the infrared region is killed by the plain mass Sudakov. This is an example of Sudakov-safe observables [56, 57].

Another key aspect of our results is that a cut on the shapes leads to an exponential suppression of the signal efficiency. This has to be contrasted with two-prong taggers like the MassDrop Tagger, trimming or pruning which would only give a linear suppression [81]. This means that although it initially seems natural to work in the small v limit, in practice one will not be able to take the cut on v too small. In practice, computing corrections for finite v could then become relevant for this discussion.

[20] We used the same samples as sections 7.2.9 and 7.3.6, using a 3 TeV cut on the jet p_t and varying its mass.

— 8 —
Dichroic N-subjettiness

In this chapter, we further explore the interplay between grooming/tagging techniques and jet shapes, in particular N-subjettiness. We introduce the concept of "dichroic" N-subjettiness ratios for improved radiation constraints. Starting from an object in which two hard prongs have been identified ("tagged"), the dichroic variant of N-subjettiness differs from standard subjettiness ratios in that it uses different (sub)jets for the numerator and denominator of the τ_2/τ_1 ratio. These two (sub)jets will generally overlap and correspond to different degrees of tagging/grooming. The reason for calling this "dichroic" is that the radiation patterns in the two different (sub)jets are driven by distinct color flows.

In this work we will use a large jet for calculating τ_2 and a smaller, tagged subjet for τ_1. Calculating τ_2 on the large jet provides substantial sensitivity to the different color structures of signal (color singlet when viewed at large angles) and background (color triplet for a quark-jet or octet for a gluon-jet). Calculating τ_1 on the tagged subjet ensures that it is not substantially affected by the overall color flow of the large jet, but rather is governed essentially by the invariant mass of the two-prong structure found by the tagger. The resulting dichroic τ_2/τ_1 ratio gives enhanced performance compared to the original version of N-subjettiness, which adopts the same (sub)jet for numerator and denominator.

Section 8.1 presents the dichroic N-subjettiness. In section 8.2 we study the performance of this observable using Monte Carlo simulations. Finally, in section 8.3 we present analytical calculations for this observable using the basic bulding blocks computed in chapter 5 and compare our predictions to Monte Carlo generators.

8.1 Dichroic subjettiness ratios

8.1.1 Combining mMDT/SD with N-subjettiness

In this chapter we will present the dichroic combination of a tagger with a radiation constraint. The discussion below assumes that we use SoftDrop or the modified MassDrop Tagger as our tagger and a cut on τ_{21} as a radiation constraint, but we believe that the

core argument can also be applied to other shapes, for example to energy correlation functions.

Let us consider a high-p_t large-radius ($R \simeq 1$) jet on which we have applied an mMDT (or SD) tagger. The original large-radius jet will be called the full jet. The part of the jet that remains after the mMDT/SD tagging procedure will be called the tagged jet, and has an angular size comparable to the angle between the two hard prongs identified by the tagger. The N-subjettiness variables τ_1 and τ_2 can be evaluated either on the full or the tagged jet and there are three combinations of interest:

$$\tau_{21}^{\text{tagged}} \equiv \frac{\tau_2^{\text{tagged}}}{\tau_1^{\text{tagged}}}, \tag{8.1a}$$

$$\tau_{21}^{\text{full}} \equiv \frac{\tau_2^{\text{full}}}{\tau_1^{\text{full}}}, \tag{8.1b}$$

$$\tau_{21}^{\text{dichroic}} \equiv \frac{\tau_2^{\text{full}}}{\tau_1^{\text{tagged}}}. \tag{8.1c}$$

One can be tempted to also consider a fourth option where τ_1 is computed on the full jet and τ_2 on the tagged jet. It is straightforward to show, following the same arguments as below, that this is not the best combination, as one might expect intuitively.

To understand how these different variants work, we will take two approaches. First we will consider what values of τ_{21} arise for different kinematic configurations involving three particles in the jet, i.e. two emissions in the case of QCD jets, and the original two prongs plus one additional emission in the case of signal jets. Then we will use this information to understand how a cut on τ_{21} constrains the radiation inside the jet.

During this discussion it will be useful to keep in mind the core difference between signal and background jets. In the case of the background jets, the whole Lund plane and the "leaf", where secondary emissions take place, can contain emissions. In the case of signal jets, emissions are mostly limited to the region where angles are smaller than the decay opening angle and transverse momenta smaller than the mass. The leaves in the two cases have different color factors, however we will neglect this aspect in our discussion. [1] Rather we will concentrate on the differences that arise at large angle, i.e. from the different coherent radiation patterns of colored versus net color-neutral objects.

We consider the situation where, after the tagger has been applied, the tagged jet mass is dominated by emission "a", i.e. $\rho \approx z_a\theta_a^2$ (in the case of the signal jet this is the softer of the two prongs). The Lund plane phase space can then be separated into 3 regions depicted in fig. 8.1. Region A (in red) is the region that is constrained to be free of radiation by the fact that the tagger has triggered on emission a. This corresponds to the region where both $z\theta^2 > z_a\theta_a^2$ and the SoftDrop condition (4.5) are satisfied.

[1] At low p_t a significant part of τ_{21}'s discriminating power is arguably associated with the leaf and, for gluon-initiated background jets, with the part of the main Lund plane that is at small angles compared to the decay opening. This is mostly equivalent to quark–gluon discrimination, which is known to be only moderately effective [89, 10, 20, 91] and not to improve significantly at high-p_t. These effects are included in the analytic calculations of section 8.3.

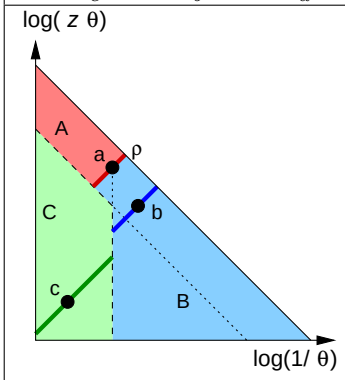
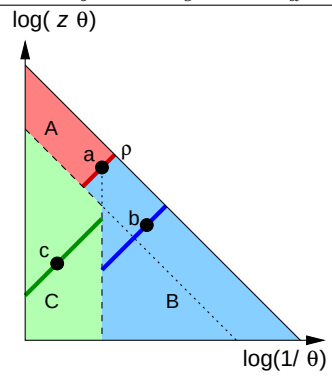
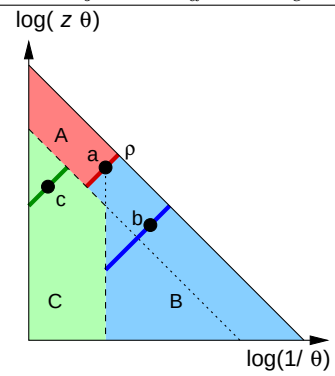
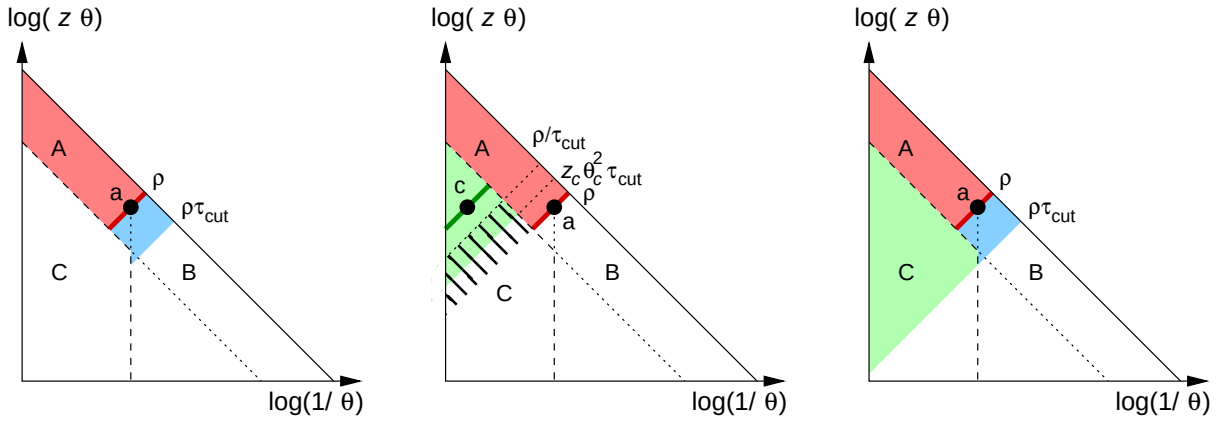
	Case 1		Case 2		Case 3	
	$z_c\theta_c^2 \ll z_b\theta_b^2 \ll z_a\theta_a^2$		$z_b\theta_b^2 \ll z_c\theta_c^2 \ll z_a\theta_a^2$		$z_b\theta_b^2 \ll z_a\theta_a^2 \ll z_c\theta_c^2$	
						
	bkg	sig	bkg	sig	bkg	sig
$\tau_{21}^{\text{tagged}}$	**b/a**	b/a	b/a	b/a	b/a	b/a
τ_{21}^{full}	**b/a**	b/a	**c/a**	b/a	a/c	b/a
$\tau_{21}^{\text{dichroic}}$	**b/a**	b/a	**c/a**	b/a	**a/a**	b/a

Figure 8.1 — Schematic representation of three possible kinematic configurations for the combination of τ_{21} with mMDT/SD (shown specifically for mMDT or SD with $\beta = 0$). In each Lund diagram, emission "a" corresponds to the emission that dominates the mMDT/SD jet mass. This defines three regions: region A (red) is vetoed by mMDT, region B (blue) contains the constituents of the mMDT/SD jet and region C (blue) is the difference between the mMDT/SD jet and the full jet. Emissions "b" and "c" are respectively in regions B and C, and the three plots correspond to three different orderings of $z_c\theta_c^2$ compared to $z_a\theta_a^2$ and $z_b\theta_b^2$. The table below the plots shows the corresponding value of τ_{21} for both the QCD background (where all three regions have to be included) and the signal (where only regions A and B are present). For simplicity, "b/a" stands for $(z_b\theta_b^2)/(z_a\theta_a^2)$, and so forth.

Of the remaining phase space, region B (blue) corresponds to emissions that are contained inside the tagged jet. It is populated in both signal and background cases. It contains not only emissions that satisfy the mMDT/SD condition ($z > z_{\text{cut}}$ in the case of mMDT), but also emissions with $z\theta^2 < z_a\theta_a^2$ and $\theta < \theta_a$, due to the Cambridge/Aachen declustering used by mMDT/SD. Region C (green) corresponds to emissions that are in the original full jet, but not in the tagged jet. It is uniformly populated in the background case, while in the signal case it is mostly empty of radiation, except at the left-hand edge (initial-state radiation) and the right-hand edge (leakage of radiation from the color-singlet $q\bar{q}$ decay). The emission with the largest $z\theta^2$ in each of regions B and C will respectively be labelled b and c and we will assume strong ordering between emissions.

There are three kinematic cases to consider for the relative $z\theta^2$ ordering of emissions a, b and c, cf. fig. 8.1. In each case, fig. 8.1 gives the result for each of the τ_{21} variants, for both background and signal. The signal case simply assumes that there are no emissions in region C, which is appropriate in a double-logarithmic approximation. The results are



(a) $\tau_{21}^{\text{tagged}}$ **(b)** τ_{21}^{full} **(c)** $\tau_{21}^{\text{dichroic}}$ and τ_{21}^{full}

Figure 8.2 — Regions where real emissions are vetoed when combining a mMDT/SD tagger with a cut on τ_{21}. See text for details.

expressed as a shorthand, $i/j \equiv z_i\theta_i^2/z_j\theta_j^2$.

The case of the signal is particularly simple: since $z_b\theta_b^2 < z_a\theta_a^2$ and there is nothing in region C, all variants give $\tau_{21} = z_b\theta_b^2/z_a\theta_a^2$. Given that the signal result is always the same, the performance of the signal/background discrimination will be best for the method that gives the largest background τ_{21} result (recall that one enhances signal relative to background by requiring $\tau_{21} < \tau_{\text{cut}}$).

Let us examine the background separately for each of the three kinematic cases shown in fig. 8.1:

1. For $z_a\theta_a^2 \gg z_b\theta_b^2 \gg z_c\theta_c^2$, all three τ_{21} variants give the same result as for the signal, $z_b\theta_b^2/z_a\theta_a^2$.[2]

2. For $z_a^2\theta_a^2 \gg z_c^2\theta_c^2 \gg z_b^2\theta_b^2$, $\tau_{21}^{\text{tagged}}$ is still still given by $z_b\theta_b^2/z_a\theta_a^2$, but τ_{21}^{full} and $\tau_{21}^{\text{dichroic}}$ now both take the larger value of $z_c\theta_c^2/z_a\theta_a^2$. They should therefore perform better in this case.

3. Finally, for $z_c^2\theta_c^2 \gg z_a^2\theta_a^2 \gg z_b^2\theta_b^2$, $\tau_{21}^{\text{tagged}}$ is again given by $z_b\theta_b^2/z_a\theta_a^2$; τ_{21}^{full} is given by $z_a\theta_a^2/z_c\theta_c^2$, since τ_1 is dominated by emission c, while τ_2 is dominated by emission a. Depending on the exact configuration, τ_{21}^{full} may be larger or smaller than $z_b\theta_b^2/z_a\theta_a^2$ and so may or may not be advantageous. $\tau_{21}^{\text{dichroic}}$ has a value of $z_a\theta_a^2/z_a\theta_a^2 = 1$, which is always larger than the signal and larger than the other two variants.

Overall therefore, $\tau_{21}^{\text{dichroic}}$ is expected to be the best of the three variants.

Alternatively, we can also see the benefit of the dichroic combination by examining directly how emissions are constrained when one applies a given cut on the τ_{21} ratio. We

[2]Even if the signal and background have the same value, the different color factor of the leaf, discussed earlier, still ensures discriminating power, because $z_b\theta_b^2/z_a\theta_a^2$ tends to be smaller for C_F color factors (signal) than for C_A color factors (background).

have represented the Lund diagrams relevant for our discussion in fig. 8.2, where we have used the same regions A, B and C as in the above discussion.

We start by considering a jet for which we already have applied the mMDT/SD procedure, resulting in a (mMDT/SD) mass ρ dominated by emission "a". This automatically comes with a mMDT/SD prefactor and Sudakov suppression represented by the solid red line and shaded light red area (region A) in fig. 8.2, guaranteeing that there are no emissions at larger mass kept by the mMDT/SD.

For $\tau_{21}^{\text{tagged}}$, emissions in region B are vetoed down to a mass scale $\rho\tau_{\text{cut}}$ while emissions in region C, i.e. outside the mMDT/SD tagged jet, are left unconstrained. This results in the (additional) Sudakov suppression given by the blue area (region B) in fig. 8.2(a).

The situation for τ_{21}^{full} is a bit more involved and we have three cases to consider. The first case is when there is (at least) one emission in region C with $z\theta^2 > z_a\theta_a^2/\tau_{\text{cut}} \equiv \rho/\tau_{\text{cut}}$ and is represented in fig. 8.2(b). Let us then call emission "c" the emission in region C with the largest $z\theta^2$, which thus comes with a Sudakov suppression imposing that there are no other emissions in region C with $z\theta^2 > z_c\theta_c^2$. Emission "c" will dominate τ_1 so that the cut on τ_{21} will come with an extra suppression factor in region C extending from $z_c\theta_c^2$ down to $z\theta^2 = z_c\theta_c^2\tau_{\text{cut}}$. Consequently, all emissions down to $z_c\theta_c^2\tau_{\text{cut}}$ are vetoed as depicted in fig. 8.2(b). The second case is when the emission in region C with the largest $z\theta^2$ satisfies $z_a\theta_a^2 \equiv \rho < z_c\theta_c^2 < \rho/\tau_{\text{cut}}$. This region, represented by the hatched area in fig. 8.2(b), is entirely forbidden because it would give a value of $\tau_{21} \geq z_a\theta_a^2/z_c\theta_c^2$ which is always larger than τ_{cut}. The third case is when there are no emissions in region C with $z\theta^2 > \rho$. This directly comes with a Sudakov suppression in region C vetoing emission down to $z\theta^2 = \rho$. In this case, τ_1 is dominated by emission "a" and the constraint on τ_{21} further vetoes emissions with $\rho\tau_{\text{cut}} < z\theta^2 < \rho$ in both regions B and C. These two vetoes combine to vetoing all emission down to $\rho\tau_{\text{cut}}$ as represented in fig. 8.2(c).

If instead we use our new $\tau_{21}^{\text{dichroic}}$ variable, we are always in the situation of fig. 8.2(c), where we veto all emissions down to a mass scale $\rho\tau_{\text{cut}}$ in both regions B and C. This new version therefore comes with the strongest Sudakov suppression, i.e. of the three τ_{21} variables it is the one that, for background jets, is least likely to have a small τ_{21} value. Given that the three τ_{21} variants behave similarly to each other for signal, the signal-to-background discrimination should be improved for the dichroic variant.

With our dichroic method, we actually recover the same overall Sudakov suppression as the one we had when measuring the full jet mass and cutting on the full N-subjettiness. The gain of our new method (8.1c) compared to this full N-subjettiness case comes from the fact that the prefactor associated with the jet mass is now subject to the constraint imposed by the tagger. If we take for example the case of the mMDT, this prefactor would be largely suppressed for the background — going from $\sim \alpha_s \log(1/\rho)$ for full N-subjettiness to $\sim \alpha_s \log(1/z_{\text{cut}})$ for the dichroic method — while the signal would only be suppressed by a much smaller factor $\sim 1 - 2z_{\text{cut}}$. Additionally, measuring the tagged jet mass instead of the full jet mass significantly reduces ISR and non-perturbative effects which would otherwise affect the resolution of the signal mass peak.

Finally, we note that the gain in performance is expected to increase for larger boosts due to region C getting bigger (double-logarithmically in ρ).

8.1.2 Dichroic subjettiness with SoftDrop (pre-)grooming

Since $\tau_{21}^{\text{dichroic}}$ uses τ_2 computed on the full jet, including all the soft radiation at large angles, we can expect this observable to be quite sensitive to poorly-controlled non-perturbative effects – hadronisation and the underlying event – and to pileup.

The standard strategy to mitigate these effects is to kill two birds with one stone and to use mMDT (or SD) both as a two-prong tagger and as a groomer, and impose the τ_{21} constraint on the result. This is equivalent to the $\tau_{21}^{\text{tagged}}$ variant discussed (fig. 8.2(a)), with the drawback and loss of performance described in the previous section.

We show here how we can achieve a background rejection that is larger than for $\tau_{21}^{\text{tagged}}$ and more robust with respect to non-perturbative effects than $\tau_{21}^{\text{dichroic}}$. Conceptually, the idea is that the tagger and groomer achieve two different tasks: the tagger selects a two-prong structure in the jet, imposing a rather hard constraint on the soft radiation in order to do so, leading to a small R' prefactor for the jet mass. This is not quite what we want from a groomer, which should get rid of the soft-and-large-angle radiation while retaining enough of the jet substructure to have some discriminating power when using radiation constraints.

This suggests the following picture: we first apply a "gentle" grooming procedure to the jet, like a SoftDrop procedure with a positive value of β. This is meant to clean the jet of the unwanted soft effects [3] while retaining as much as possible the information about the perturbative radiation in the jet. We can then carry on with the dichroic method presented in the previous section, i.e. use a more aggressive tagger, like mMDT,[4] to compute the jet mass and τ_1 and compute τ_2 on the SD (pre-)groomed jet:

$$\tau_{21,\text{groomed}}^{\text{dichroic}} = \frac{\tau_2(\text{SD jet})}{\tau_1(\text{mMDT jet})}. \tag{8.2}$$

This is depicted in fig. 8.3, where regions A and B are the same as in the previous section, but now region C indicates the region where emissions are kept by the groomer but rejected by the tagger. Similarly, we can introduce

$$\tau_{21,\text{groomed}}^{\text{full}} = \frac{\tau_2(\text{SD jet})}{\tau_1(\text{SD jet})}. \tag{8.3}$$

Note that we will always choose our mMDT-tagging and SD-grooming parameters such that the tagged jet is the same whether tagging is performed before or after grooming. For mMDT-tagging with parameter z_{cut} and SD-grooming with parameters ζ_{cut} and β, this implies $\zeta_{\text{cut}} \leq z_{\text{cut}}$ and $\beta \geq 0$.

Using the same arguments as in section 8.1.1, we can show straightforwardly that this method will have a larger rejection than with the other two variants where one would be

[3] In the presence of pileup, one should still apply a pileup subtraction procedure [92], like area–median subtraction [93, 94], charged-track-based techniques [95, 96, 97], the constituent subtractor [98], SoftKiller [99] or PUPPI [100]. This can be done straightforwardly with SoftDrop and mMDT.

[4] Or SD with a smaller value of β than used in the grooming.

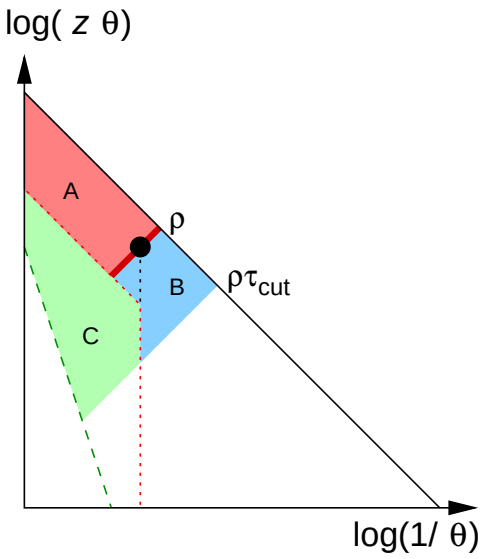


Figure 8.3 — Phase space constraints on QCD jets obtained from our new combination including grooming: we first groom the jet, e.g. with SoftDrop. We then compute both the jet mass and τ_1 on the tagged jet (here using the mMDT), yielding the solid red line prefactor and the shaded red region (A) for the Sudakov exponent. We then impose a cut on the τ_{21} ratio with τ_2 computed on the SD jet, leading to the extra shaded blue and green regions (B and C) for the Sudakov exponent.

computing the jet mass on the mMDT-tagged jet and the τ_{21} ratio either on the mMDT-tagged jet, $\tau_{21,\text{groomed}}^{\text{tagged}} \equiv \tau_{21}^{\text{tagged}}$, or on the SD-groomed jet, $\tau_{21,\text{groomed}}^{\text{full}}$, owing to a larger Sudakov suppression of the background, for a similar signal efficiency.

Compared to the other possible situation where both the jet mass and the τ_{21} ratio are computed on the SD-groomed jet, the dichroic variant would have a smaller R' prefactor, associated with mMDT instead of SD. This again leads to a larger background rejection.

Because of the initial grooming step, the groomed dichroic subjettiness ration is expected to be less discriminating than the ungroomed version introduced in Section 8.1.1. Indeed, the associated Sudakov exponent is smaller since we have amputated part of the soft-large-angle region. One should however expect that this groomed variant will be less sensitive to non-perturbative effects. Overall, there is therefore a trade-off between effectiveness, in terms of achieving the largest suppression of the QCD background for a given signal efficiency, and perturbative robustness, in terms of limiting the sensitivity to poorly-controlled non-perturbative effects.

8.2 Performance in Monte Carlo simulations

Let us now investigate the effectiveness and robustness of dichroic subjettiness ratios in Monte Carlo simulations, using Pythia 8.186 [35], at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Our signal sample consists of WW events, while for the background we use dijet events. Jets are reconstructed with the anti-k_t algorithm with $R = 1$ and in determining signal and background efficiencies we keep all jets above a given p_t cut.[5] We use the modified MassDrop Tagger with $z_{\text{cut}} = 0.1$ for the 2-prong tagging and vary the cut on the τ_{21} ratio. Whenever a SoftDrop grooming procedure is included, we use $\zeta_{\text{cut}} = 0.05$ and

[5] All jets in the signal sample above that cut are considered to be signal-like, even if they came from initial-state radiation; however such initial-state jets will have been relatively rare in our sample and so should not affect our final conclusions.

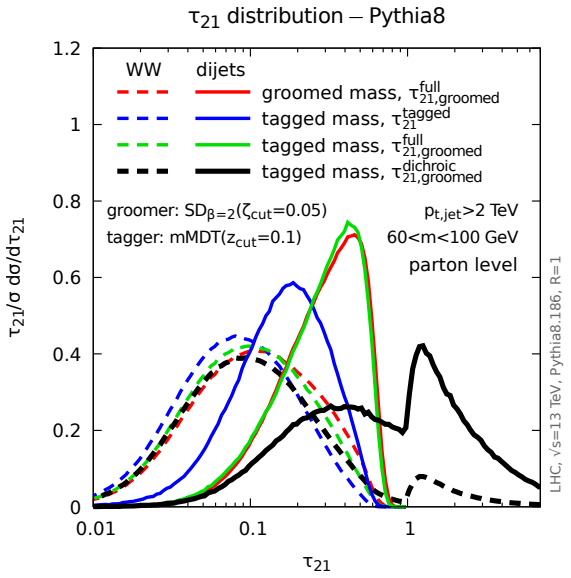


Figure 8.4 — τ_{21} distributions for jets in dijet (solid lines) and WW (dashed lines) events again imposing $p_t > 2$ TeV and including SoftDrop grooming. Different colors correspond to different combinations of jets used for the computation of the jet mass, τ_1 and τ_2 as indicated in the legend, our new dichroic combination being plotted in black. We have selected jets with a mass is between 60 and 100 GeV. The cross-section used for normalisation, σ, is defined after the jet p_t and mass cut, so that all curves integrate to one.

$\beta = 2$ as illustrative parameter choices, for clarity we use separate symbols z_{cut} and ζ_{cut} respectively for the parameters of mMDT and SD. Jet reconstruction and manipulation are performed with FastJet 3.2.0 [38, 39] and fjcontrib 1.024 [83].

8.2.1 N-subjettiness and mass distributions with various τ_{21} ratios ($\beta_\tau = 2$)

We start by examining the τ_{21} distribution. This is plotted in fig. 8.4 for both QCD jets (solid lines) in dijet events and W jets (dashed lines) in WW events. We select jets above 2 TeV and always apply SoftDrop grooming. In practice, we use parton-level events, and impose a cut on the reconstructed jet mass (SD-groomed or mMDT-tagged) $60 < m < 100$ GeV. We consider four cases: the $\tau_{21,\text{groomed}}^{\text{full}}$ distribution when we cut on the SD-groomed mass and the $\tau_{21}^{\text{tagged}}$, $\tau_{21,\text{groomed}}^{\text{full}}$ and $\tau_{21,\text{groomed}}^{\text{dichroic}}$ distributions when we cut on the mMDT-tagged mass. As expected, the distributions for signal (W) jets are peaked at smaller values of τ_{21} than the corresponding distribution for background (QCD) jets. Fig. 8.4 shows that all the signal distributions, and in particular the three options where one measures the mMDT-tagged jet mass, are very similar. This is in agreement with our discussion in the previous section. Comparatively the background distributions look rather different. The case where everything is computed from the mMDT-tagged jet (the solid blue curve) peaks at smaller values of τ_{21} as expected from its smaller Sudakov suppression, related to the fact that this combination puts no constraints on large-angle emissions (region C in the previous section). Furthermore, the dichroic combination, the solid black curve in fig. 8.4, is expected to have the largest suppression and is indeed peaked at larger τ_{21} values, translating into a larger discrimination against signal jets.

Note that the τ_{21} distribution for the dichroic combination also shows a peak for $\tau_{12} > 1$ that we have not discussed in our earlier argumentation. This comes from events with multiple emissions in region C, as discussed in our analytic calculations in section 8.3.

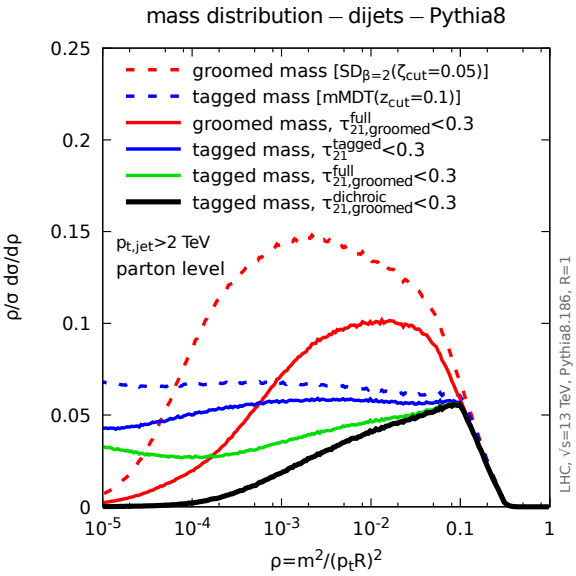


Figure 8.5 — Mass distribution for QCD jets with $p_t > 2$ TeV (anti-k_t, $R = 1$) at parton level, including SoftDrop grooming. The dashed lines, in red for the SD-groomed jet and in blue for the mMDT-tagged jet, are the mass distributions with no constraint on N-subjettiness. The solid lines have an additional cut $\tau_{21} < 0.3$ with different combinations of jets used for the computation of the jet mass, τ_1 and τ_2 as indicated in the legend, our dichroic combination being plotted using a solid black line. The cross section used for normalisation, σ is that for jets above the p_t cut.

Results for the mass distribution obtained for background (QCD dijets) jets at parton level (without UE) are presented in fig. 8.5. As in fig. 8.4, SoftDrop grooming has always been applied prior to any additional tagging or N-subjettiness cut. Again, we can identify most of the features discussed in section 8.1. First of all, if we compare the mMDT-tagged mass (dashed blue curve) to the SD-groomed jet mass (dashed red curve) we see that the latter is smaller than the former at small masses, owing to the larger Sudakov factor $R_{\rm SD} > R_{\rm mMDT}$, but larger at intermediate masses, due to the larger prefactor $R'_{\rm SD} > R'_{\rm mMDT}$.

Then, we can consider the effect of the additional constraint on the τ_{21} ratio, taken here as $\tau_{21} < 0.3$ for illustrative purpose. If we compute τ_{21} on the same jet as for the mass ($\tau^{\rm full}_{21,\rm groomed}$ in solid red and $\tau^{\rm tagged}_{21}$ in solid blue for the SD-groomed and mMDT-tagged jets respectively), we see that the cut reduces the background, that the reduction increases for smaller masses and that the reduction is larger for the SD-groomed jet than for the mMDT-tagged jet. This last point is a reflection of the fact, that the Sudakov suppression associated with the N-subjettiness cut is larger when both the mass and τ_{21} are computed on the SD-groomed jet than when both the mass and τ_{21} are computed on the mMDT-tagged jet (fig. 8.2(left)). Then, when measuring the mMDT-tagged jet mass, one sees that computing τ_{21} on the SD-groomed jet ($\tau^{\rm full}_{21,\rm groomed}$, the solid green curve in fig. 8.5) shows a larger suppression than computing τ_{21} on the mMDT-tagged jet, although the difference is reduced at very small masses. Finally, if we consider our new, dichroic case, eq. (8.2) ($\tau^{\rm dichroic}_{21,\rm groomed}$, the solid black curve), we see a larger suppression than in all other cases, as expected from our earlier arguments.

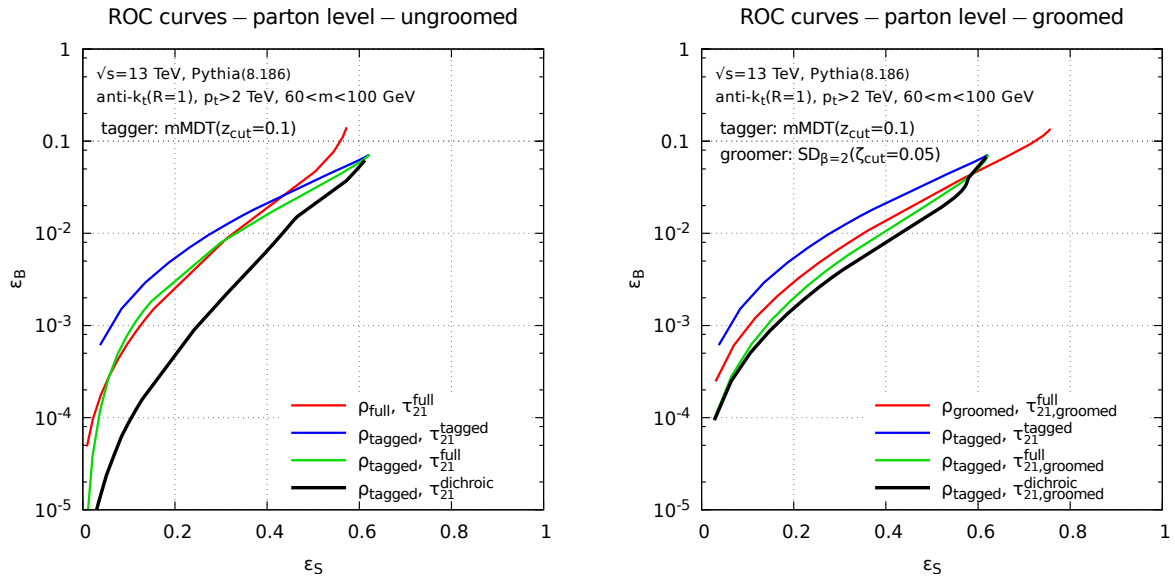


Figure 8.6 — ROC curves for various τ_{21} combinations, i.e. background versus signal efficiency, at parton level. The left plot is obtained starting from the full jet, while for the right plot, a SoftDrop (pre-)grooming step has been applied. The ROC curves are obtained by varying the cut on the τ_{21} ratio. In all cases, we considered anti-$k_t(R = 1)$ jets with $p_t > 2$ TeV.

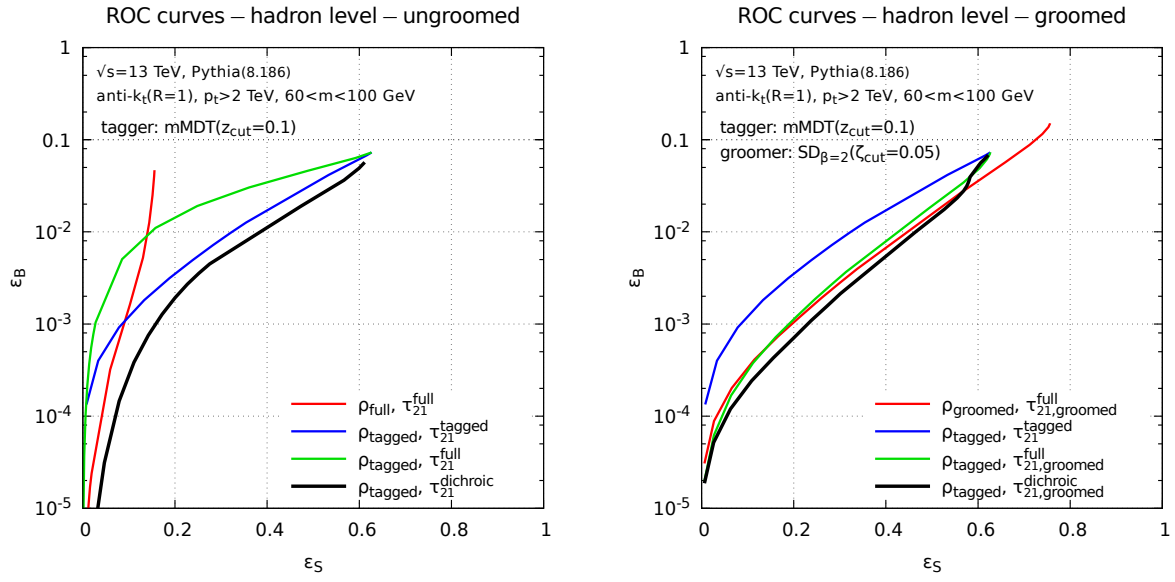


Figure 8.7 — Same as figure as 8.6, now for hadron level (including the underlying event).

8.2.2 Signal v. background discrimination and other performance measures

To further test the performance of our new method, we have also studied ROC (receiver operating characteristic) curves, shown in fig. 8.6 for parton-level simulations and in fig. 8.7 for hadron-level events including hadronisation and the underlying event. In all cases, we impose the constraint that the (full, tagged or groomed) mass is between 60 and 100 GeV. Efficiencies are given relative to the inclusive cross-section for having jets above our p_t cut.

Let us first discuss the result of parton-level simulations, fig. 8.6, where the dichroic ratio is again represented by the black curves. Without grooming (the left-hand plot in the figure), our method shows a substantial improvement compared to all other combinations considered, outperforming them by almost 30% in background rejection at a signal efficiency of 50% and by more than a factor of 2 at a signal efficiency of 40%. After Soft-Drop grooming (right-hand plot), the dichroic method, i.e. computing the jet mass and τ_1 on the mMDT-tagged jet and τ_2 on the SD-groomed jet, still shows an improvement, albeit less impressive than what is observed using the full jet to compute τ_2.

If instead we consider the results at hadron level, including both the perturbative parton shower as well as non-perturbative effects, in fig. 8.7, we see that the dichroic subjettiness ratio still does a better job than the other variants but the gain is smaller. For example, measuring the mMDT-tagged mass with a cut on the groomed dichroic ratio, $\tau^{\text{dichroic}}_{21,\text{groomed}}$, the optimal choice in fig. 8.7, is only slightly better than the next best choice where one measures the SD-groomed mass and imposes a constraint on $\tau^{\text{full}}_{21,\text{groomed}}$. This is because in going from parton to hadron level, the ρ_{groomed} vs. $\tau^{\text{full}}_{21,\text{groomed}}$ curve has moved down more than the ρ_{tagged}–$\tau^{\text{dichroic}}_{21,\text{groomed}}$ curve, i.e. the former is getting a significantly larger boost in its discriminating power from non-perturbative effects.[6] This is potentially problematic, because one does not necessarily want signal-to-background discrimination power for a multi-TeV object to be substantially driven by the physics that takes place at a scale of 1GeV, physics that cannot, with today's techniques, be predicted from first principles. Additionally, phenomena happening on a scale of 1GeV are difficult to measure reliably.

In evaluating the overall performance of different τ_{21} combinations we will consider both the signal significance and the size of non-perturbative effects, as we done in fig. 6.13 in chapter 6. For a given method and p_t cut, we first determine the τ_{21} cut required to obtain a desired signal efficiency (at hadron level). For that value of the τ_{21} cut, we can compute the signal significance, defined as $\varepsilon_S/\sqrt{\varepsilon_B}$ (computed at hadron level) which is a measure of the discriminating power of the method; we then estimate non-perturbative effects as the ratio between the background efficiency at hadron level divided by the background efficiency at parton level, which is a measure of robustness against non-perturbative effects. We will show results for a range of different signal-efficiency choices

[6] That there should be larger non-perturbative effects in the ρ_{groomed}–$\tau^{\text{full}}_{21,\text{groomed}}$ can be understood as follows: because ρ_{groomed} accepts a larger fraction of signal events in a given mass window than ρ_{tagged}, to reach the same final efficiency the τ_{21} cut must be pushed closer to the non-perturbative region.

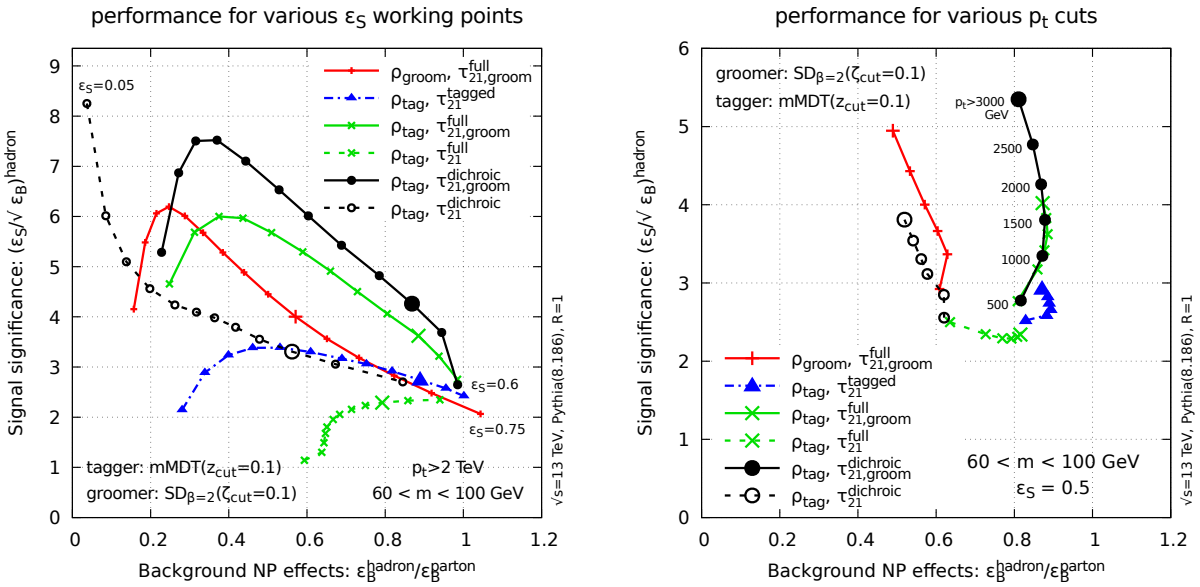


Figure 8.8 — Signal significance plotted versus the non-perturbative effects for the QCD background (defined as the ratio between the background "fake" tagging rate at hadron and parton level). Different curves correspond to different combinations indicated in the legend. For the solid curves, a SoftDrop ($\beta = 2$ and $\zeta_{\text{cut}} = 0.05$) grooming is applied, while no grooming is applied for the dashed curves. In the left-hand plot, we impose a 2 TeV p_t cut on the initial jet. The symbols on each curve then correspond to a signal efficiency (computed at hadron level) ranging from 0.05 upwards in steps of 0.05, with the large symbol on each line corresponding to $\varepsilon_S = 0.5$ and the efficiency at the right-hand extremity explicitly labelled. In the right-hand plot, the signal efficiency (computed at hadron level) is fixed to be 0.5 and the p_t cut on the jet is varied between 500 GeV and 3 TeV (in steps of 500 GeV, labelled explicitly for the groomed dichroic ratio), with the large symbol on each line corresponding to a 3 TeV cut.

and jet p_t cuts.

In fig. 8.8, which highlights the key performance features of the dichroic method, we plot the signal significance versus the non-perturbative effects for different methods. In the left-hand panel, the curves correspond to a range of τ_{21} cuts for jets with $p_t > 2\,\text{TeV}$. The points on the curves correspond to different signal efficiencies (starting from 0.05, in steps of 0.05, and with $\varepsilon_S = 0.5$ indicated by a bigger point). In the right-hand panel, the points on the curves correspond to different p_t cuts, with the τ_{21} cut adjusted (as a function of p_t) so as to ensure a constant signal efficiency of 0.5. In both plots, the τ_{21} cut is determined so as to achieve the expected signal efficiency at hadron level and the same cut is used for parton-level results. To avoid the proliferation of curves, the result for the ungroomed ρ_{full}–τ_{21}^{full} is not shown since it is obvious from the ROC curves in figs. 8.6 and 8.7(left) that it is extremely sensitive to non-perturbative effects.

In both plots, we see that the dichroic method comes with larger discriminating power

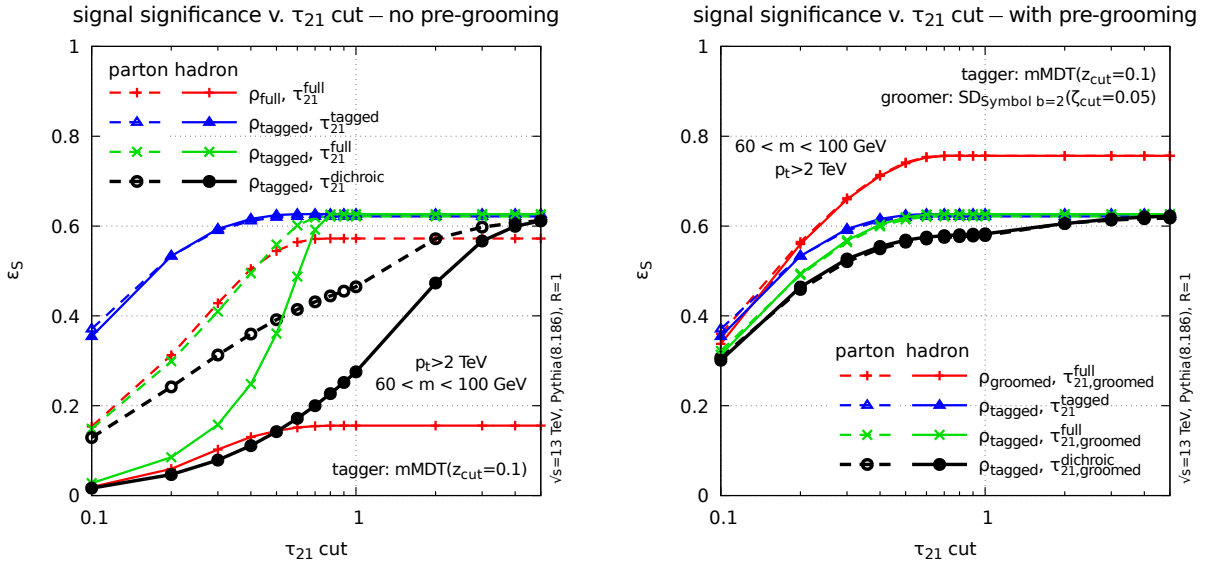


Figure 8.9 — Signal efficiency plotted as a function of the cut τ_{cut} on τ_{21} for all the combinations considered in figs. 8.6 and 8.7. Solid curves correspond to hadron-level results while dashed curves are obtained at parton level. The left plot is obtained starting from the full jet, while for the right plot, a SoftDrop grooming has been applied.

with a relatively limited sensitivity to non-perturbative effects, provided one first applies a grooming step. Without the grooming step, one observes a much larger sensitivity to non-perturbative effects, as one might expect.[7] It also appears that the performance gain increases when the boost, i.e. the jet p_t, increases. This was also expected from our arguments in section 8.1. Finally, compared to the common setups in the literature, namely with modified MassDrop tagging with a cut on τ_{21} applied either on the mMDT (ρ_{tag}–$\tau_{21}^{\text{tagged}}$, the dot-dashed blue curve) or on the full jet (ρ_{tag}–τ_{21}^{full}, the dashed green curve), our dichroic method with grooming (solid black) gives up to a factor of two improvement in signal significance, with comparable non-perturbative effects. Considering other combinations that have not been widely used experimentally, $\tau_{21,\text{groomed}}^{\text{full}}$ with either a groomed (ρ_{groom}, solid red) or a tagged (ρ_{tag}), solid green) jet mass both perform well, however $\tau_{21,\text{groomed}}^{\text{dichroic}}$ still remains the best, with an optimal significance that is about 25% larger, and smaller non-perturbative corrections for any given signal significance.

As a final check, we have studied the dependence of the signal efficiency on the τ_{21} cut, as shown in fig. 8.9. Comparing the left and right-hand plots, it appears clearly that applying SoftDrop grooming helps to reduce non-perturbative effects which otherwise significantly lower the signal efficiency. It is also interesting to notice that without grooming, the signal efficiency obtained with our dichroic method (the dashed black curve on the left plot of fig. 8.9) only reaches its plateau for cuts on τ_{21} larger than 1 already at parton level. This can likely be attributed to initial-state radiation in the jet at angles larger

[7] As seen in chapter 7, grooming largely reduces the impact of initial-state radiation as well.

than the decay angle of the W boson. These effects are strongly reduced by SoftDrop grooming (see also the discussion in Section 8.3).

8.2.3 Brief comparison with other tools

To complete our Monte Carlo studies, in fig. 8.10 we compare the performance of $\tau_{21,\text{groomed}}^{\text{dichroic}}$ with various other tools: mMDT tagging alone, SoftDrop grooming alone ($\beta = 2$ as above), and also the Y_m-variant of Y-splitter, combined with SoftDrop (pre-)grooming or with trimming, as described in detail in ch. 6. (see also [81]). Whereas in the analogous fig. 8.8, all curves involved the same signal efficiency, here this is no longer the case. Accordingly efficiencies are reported versus p_t in table 8.1.

Let us start by examining the pure mMDT result: as known already from [12] it provides mild tagging, it has small non-perturbative corrections and only modest dependence on p_t. SoftDrop ($\beta = 2$), when used alone, has slightly lower significance and larger non-perturbative corrections.[8] These two tools have the highest signal efficiencies, of about 63% and 76% respectively at 2 TeV.

Next we examine combinations that involve Y_m-splitter. This cut is similar in its effect to z_{cut} in mMDT. When used in conjunction with SD (pre-)grooming, the highest-mass emission that passes the SD cut is also the one that is unclustered by Y_m-splitter and so it is required to pass the y_{cut} condition. As a result, the constraint in the Lund plane turns out, at the leading-log level, to be identical to that obtained with $\tau_{21,\text{groomed}}^{\text{dichroic}}$ and the condition $\tau_{\text{cut}} = 1$, with a Sudakov suppression vetoing all emission down to a mass scale ρ in the SD-groomed jet, and a small prefactor $\sim \alpha_s \ln(1/y_{\text{cut}})$. This is reflected in fig. 8.10, where one sees that the $\tau_{21,\text{groomed}}^{\text{dichroic}} < 1$ curve (black open diamonds) is remarkably similar to the SD+Y_m-splitter curve (red open squares). Where the $\tau_{21,\text{groomed}}^{\text{dichroic}}$ variable has an advantage is that one can now further adjust the choice τ_{cut}, whereas with SD+Y_m-splitter that freedom is not available.

Of the various possible combinations involving Y-splitter, we are considering the Y_m with trimming, as it gives the best signal-to-background discrimination. It is shown as red solid squares in fig. 8.10. Overall it performs less well than the mMDT plus $\tau_{21,\text{groomed}}^{\text{dichroic}}$ combination with ε_S fixed to 0.4, even though is has a broadly similar signal efficiency.

Another point to discuss concerns the choice of β_τ in the N-subjettiness definition, eq. (4.6). Many experimental uses of N-subjettiness ratios have concentrated on the choice $\beta_\tau = 1$, while throughout this article we have used $\beta_\tau = 2$. A discussion of the $\beta_\tau = 1$ case is given in appendix C.1, including comparisons of dichroic and normal variants. Dichroic always perform best also for $\beta_\tau = 1$, and so in the brief summary that we give here we will only show dichroic results.

An argument often given for the choice of $\beta_\tau = 1$ is that it is less sensitive to non-perturbative effects. Fig. 8.10 (right) shows groomed (filled symbols, solid lines) and ungroomed (open symbols, dashed lines) results for $\beta_\tau = 1$ (squares and triangles) and

[8] The performance of SD can be somewhat improved for a specific m/p_t value by taking a negative value for β and adjusting z_{cut} such that one effectively removes branchings with $z < 0.1$ at that m/p_t scale (see section 7 of [17]).

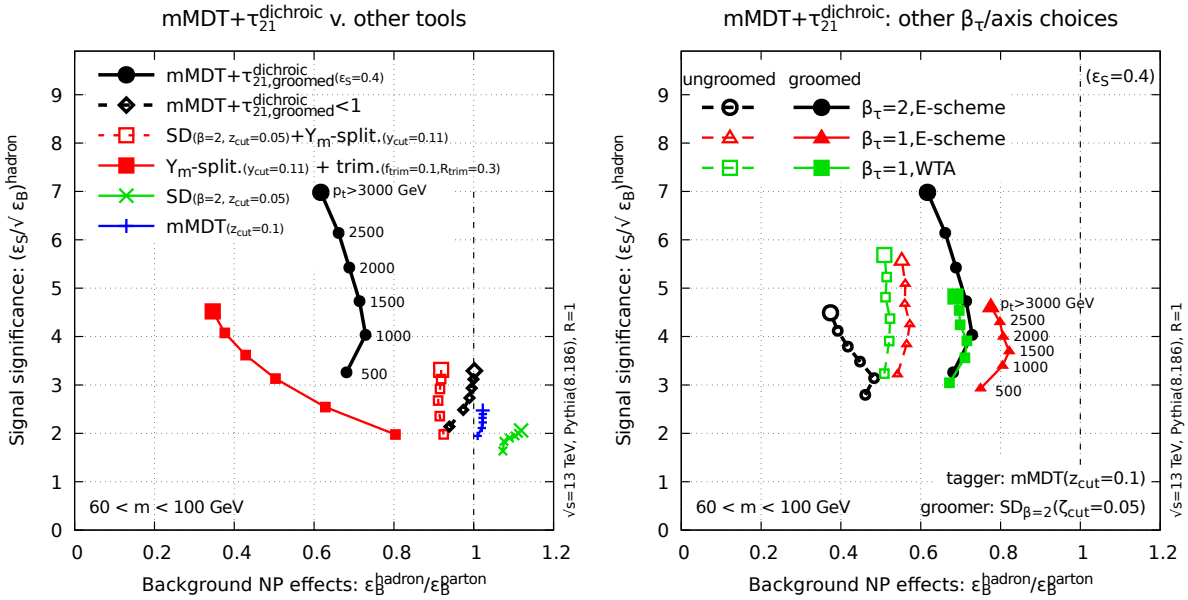


Figure 8.10 — Signal significance and non-perturbative effects for background, for jet p_t cuts ranging from 500GeV to 3 TeV in steps of 500GeV, as in fig. 8.8(right). The 3 TeV point is always labelled with a larger symbol. The plots compare $\tau_{21,\text{groomed}}^{\text{dichroic}}$ ($\beta_\tau = 2$) with a range of other tools, including Y_m-splitter (left) and $\beta_\tau = 1$ dichroic subjettiness ratios (right). Where the β_τ value is not explicitly labelled, it is equal to 2. Note that the default signal-efficiency working point for the dichroic subjettiness ratios is 0.4 here rather than the 0.5 chosen in fig. 8.8. The signal efficiencies for other cases are given in table 8.1.

	jet p_t cut [GeV]					
method	500	1000	1500	2000	2500	3000
mMDT	0.63	0.62	0.62	0.63	0.64	0.65
SoftDrop	0.74	0.74	0.75	0.76	0.77	0.79
Y_m-splitter+trimming	0.49	0.41	0.36	0.33	0.31	0.30
SoftDrop+Y_m-splitter	0.56	0.55	0.55	0.55	0.57	0.58
mMDT + $\tau_{21,\text{groomed}}^{\text{dichroic}} < 1$	0.60	0.57	0.58	0.58	0.59	0.61
all other variants	0.4	0.4	0.4	0.4	0.4	0.4

Table 8.1 — Signal efficiencies for the various tools shown in fig. 8.10.

$\beta_\tau = 2$ (circles). For the $\beta_\tau = 1$ case, we have considered either exclusive-k_t axes with the standard E-scheme four-vector recombination (triangles), or the exclusive-k_t axes with the winner-takes-all (WTA) recombination scheme (squares). In both the SD-groomed and ungroomed cases, the non-perturbative corrections are somewhat smaller for $\beta_\tau = 1$ (except in the WTA groomed case). In the ungroomed case, $\beta_\tau = 1$ also leads to better signal-discrimination. However once SD-grooming is included the signal discrimination is best for the $\beta_\tau = 2$ case. If one is concerned about the slightly larger non-perturbative effects for the SD-groomed $\beta_\tau = 2$ case, then one can slightly increase the τ_{cut} choice: in fig. 8.8(right) where τ_{cut} was chosen so as to obtain a higher signal efficiency of $\varepsilon_S = 0.5$ the $\tau^{\text{dichroic}}_{21,\text{groomed}}(\beta_\tau = 2)$ performance is very similar to the $\tau^{\text{dichroic}}_{21,\text{groomed}}(\beta_\tau = 1, \varepsilon_S = 0.4)$ performance in fig. 8.10(right). Therefore, it is the SD-groomed, $\beta_\tau = 2$, dichroic ratio that appears to give the best overall performance.

8.3 Analytic calculations

In this section, we consider brief analytic calculations relating to the observables we have presented so far. Our main goal here is to illustrate that the discussion from section 8.1 — where we used Lund diagrams to motivate dichroic subjettiness ratios — does indeed capture the qualitative picture observed in Monte Carlo simulations. To that aim, it is sufficient to use leading-logarithmic accuracy, where we control double logarithms, i.e. $\alpha_s^n \ln^j\rho \ln^k\tau_{\text{cut}} \ln^\ell z_{\text{cut}} \ln^m\zeta_{\text{cut}}$ with $j + k + \ell + m = 2n$, assuming $\rho, \tau_{\text{cut}}, z_{\text{cut}}, \zeta_{\text{cut}} \ll 1$.

For the QCD background, we find, for $\tau_{\text{cut}} < 1$: [9]

$$\rho_{\text{full}},\ \tau_{21}^{\text{full}}: \quad \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_\rho^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \exp\Big[-R_{\text{full}}(\rho,\tau,z)\Big], \tag{8.4a}$$

$$\rho_{\text{mMDT}},\ \tau_{21}^{\text{tagged}}: \quad \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_{z_{\text{cut}}}^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \exp\Big[-R_{\text{mMDT}}(\rho,\tau,z)\Big], \tag{8.4b}$$

$$\rho_{\text{SD}},\ \tau_{21,\text{groomed}}^{\text{full}}: \quad \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_{z_{\text{SD}}(\rho)}^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \exp\Big[-R_{\text{SD}}(\rho,\tau,z)\Big], \tag{8.4c}$$

$$\rho_{\text{mMDT}},\ \tau_{21}^{\text{dichroic}}: \quad \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_{z_{\text{cut}}}^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \exp\Big[-R_{\text{full}}(\rho,\tau,z)\Big], \tag{8.4d}$$

$$\rho_{\text{mMDT}},\ \tau_{21,\text{groomed}}^{\text{dichroic}}: \quad \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_{z_{\text{cut}}}^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \exp\Big[-R_{\text{SD}}(\rho,\tau,z)\Big], \tag{8.4e}$$

where $z_{\text{SD}}(\rho) = \max\Big((\rho^\beta \zeta_{\text{cut}}^2)^{1/(2+\beta)}, \rho\Big)$ and

$$\begin{aligned} R_{\text{full}}(\rho,\tau,z) = &\int_0^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2} \int_0^1 \mathrm{d}z p_i(z_2) \frac{\alpha_s(z_2\theta_2)}{2\pi} \Theta(z_2\theta_2^2 > \tau\rho), \\ &+ \int_0^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2} \int_0^1 \mathrm{d}z p_{xg}(z_2) \frac{\alpha_s(\sqrt{z\rho}z_2\theta_2)}{2\pi} \Theta(z_2\theta_2^2 > \tau), \end{aligned} \tag{8.5}$$

[9] In order to simplify the notation we will drop the subscription "cut" in τ_{cut}, when no confusion is possible.

$$R_{\mathrm{SD}}(\rho,\tau,z) = \int_0^1 \frac{\mathrm{d}\theta^2}{\theta^2}\int_0^1 \mathrm{d}z p_i(z)\,\frac{\alpha_s(z_2\theta_2)}{2\pi}\Theta(z\theta^2>\tau\rho)\Theta(z>z_{\mathrm{cut}}\theta^\beta) \\ + \int_0^1 \frac{\mathrm{d}\theta_2^2}{\theta_2^2}\int_0^1 \mathrm{d}z p_{xg}(z_2)\,\frac{\alpha_s(\sqrt{z\rho}z_2\theta_2)}{2\pi}\Theta(z_2\theta_2^2>\tau), \tag{8.6}$$

where we supposed that the emission angle θ is already normalized to the jet radius R and the choice of $p_i(z)$ depends on the flavor of the initial parton (quark or gluon). We can obtain R_{mMDT} by imposing $\beta \to 0$ in the expression R_{SD}. Notice that these expression (8.5) in particular is equivalent to eq. (7.8), obtained in chapter 7.

Using the basic building blocks, this exponents can be written as

$$R_{\mathrm{full}}(\rho,\tau,z) = T_{02}(\tau\rho, b_i; C_R) + T_{02}(\sqrt{z\rho}\tau, \sqrt{z\rho}b_g; C_A), \tag{8.7a}$$

$$R_{\mathrm{mMDT}}(\rho,\tau,z) = R_{\mathrm{full}}(\rho,\tau,z) - T_{02}(\tau\rho, z_{\mathrm{cut}}; C_R) + T_{02}(\sqrt{z\rho}\tau, z_{\mathrm{cut}}\sqrt{\rho/z}; C_R), \tag{8.7b}$$

$$R_{\mathrm{SD}}(\rho,\tau,z) = R_{\mathrm{full}}(\rho,\tau,z) - T_{-\beta 0}(\tau\rho, \zeta_{\mathrm{cut}}; C_R) + T_{-\beta 0}(\sqrt{z\rho}\tau, \zeta_{\mathrm{cut}}(\rho/z)^{\frac{(\beta+1)}{2}}; C_R). \tag{8.7c}$$

Note that the full and mMDT jet mass Sudakov introduced respectively in eq. (8.5) and eq. (8.6) can be written as

$$R_{\mathrm{full}}(\rho) = R_{\mathrm{full}}(\rho, 1, \text{“any } z\text{”})\,, \tag{8.8a}$$

$$R_{\mathrm{mMDT}}(\rho) = R_{\mathrm{mMDT}}(\rho, 1, \text{“any } z\text{”})\,. \tag{8.8b}$$

In the above expressions, z corresponds to the momentum fraction of the emission dominating the jet mass (emission “a” in figs. 8.1 and 8.2). We keep the z integration explicit since the secondary emissions, the C_A terms, depend explicitly on z. In all cases, the integration over z runs over the region kinematically allowed by the tagger defining the jet mass. The Sudakov exponent in these expressions is then essentially given by the jet on which we compute τ_2.

While we only target leading-logarithmic accuracy, our results also include the single-logarithmic contributions coming from hard collinear splittings, which are often phenomenologically important. They appear as the b_i factors in eqs. (8.4) and (8.7), where we have introduced $b_i = \exp(B_i)$.[10] These contributions can effectively be taken into account by limiting all z integrations to b_i for primary emissions and b_g for secondary emissions.

Finally, as expected, if one takes the limit $\beta \to \infty$ of the SD results, one recovers the full results. Also, the limit $\beta \to 0$ of (8.4c), reduces to (8.4b).

So far, we have not yet discussed the case where ρ is computed from the mMDT-tagged jet and τ_{21} from the full jet. This is more involved due to the two separate kinematic configurations involved (see fig. 8.1(b-c)). In the end, we find (assuming $\rho < z_{\mathrm{cut}}$)

$$\rho_{\mathrm{mMDT}},\ \tau_{21}^{\mathrm{full}}:\ \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} = \int_{z_{\mathrm{cut}}}^{b_i}\frac{dz}{z}\frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi}\exp\Big[-R_{\mathrm{full}}(\rho,\tau,z)\Big]$$

[10] We are adopting the alternative treatment to hard and collinear emissions, as discussed in sec. 5.3

$$+\,\Theta\left(z_{\text{cut}} > \frac{\rho}{\tau}\right)\int_{z_{\text{cut}}}^{b_i}\frac{dz}{z}\frac{\alpha_s(\sqrt{z\rho}p_tR)C_R}{\pi}\exp\Big[-R_{\text{mMDT}}(\rho)\Big]\times$$
$$\times\int_{\rho/\tau}^{z_{\text{cut}}}\frac{d\rho_c}{\rho_c}\int_{\rho_c}^{z_{\text{cut}}}\frac{dz_c}{z_c}\frac{\alpha_s(\sqrt{z_c\rho_c}p_tR)C_R}{\pi}\exp\Big[-R_{\text{out,full}}(\rho_c,\tau,z_c)\Big], \quad (8.9)$$

and a similar expression with "full" replaced by "SD" for the case where τ_{21} is calculated on the SD jet. In the above expression, we have used $\rho_c = z_c\theta_c^2$ and

$$R_{\text{out,full}}(\rho_c,\tau,z_c) = T_{02}(\rho_c\tau, z_{\text{cut}}; C_R) + T_{02}(\sqrt{\rho_c z_c}\tau, \sqrt{\rho_c z_c}b_g; C_A)\,, \quad (8.10a)$$
$$R_{\text{out,SD}}(\rho_c,\tau,z_c) = R_{\text{out,full}}(\rho_c,\tau,z_c)$$
$$- T_{-\beta 2}(\rho_c\tau,\zeta_{\text{cut}};C_R) + T_{-\beta 2}(\sqrt{\rho_c z_c}\tau,\zeta_{\text{cut}}(\rho_c/z_c)^{(\beta+1)/2};C_R). \quad (8.10b)$$

The configurations contributing to the last two lines of eq. (8.9) come from jets with at least one emission in region C (discarded by mMDT) with $\rho_c \equiv z_c\theta_c^2 > \rho/\tau_{\text{cut}}$. They result in an extra contribution to the mass distribution, which would then be larger than what we obtain with our dichroic combination (eq. (8.4d) or, equivalently, the first line of eq. (8.9)). When using the dichroic combination, these configurations would all have $\tau_{21} \geq 1$ (up to $\tau_{21} = z_{\text{cut}}/\rho$). In particular, for a cut $\tau_{21} < \tau_{\text{cut}}$ with $\tau_{\text{cut}} > 1$, the dichroic combination leads to:

$$\rho_{\text{mMDT}},\ \tau_{21}^{\text{dichroic}}:\ \frac{\rho}{\sigma}\frac{d\sigma}{d\rho}\bigg|_{<\tau} \overset{\tau\geq 1}{=} \int_{z_{\text{cut}}}^{b_i}\frac{dz}{z}\frac{\alpha_s(\sqrt{z\rho}p_tR)C_R}{\pi}e^{-R_{\text{mMDT}}(\rho)} \quad (8.11)$$
$$\left(e^{-R_{\text{out,full}}(\rho\tau)} + \int_{\rho\tau}^{z_{\text{cut}}}\frac{d\rho_c}{\rho_c}\int_{\rho_c}^{z_{\text{cut}}}\frac{dz_c}{z_c}\frac{\alpha_s(\sqrt{z_c\rho_c}p_tR)C_R}{\pi}e^{-R_{\text{out,full}}(\rho_c,\rho\tau/\rho_c,z_c)}\right)$$

with

$$R_{\text{out,full}}(\rho\tau) = T_{02}(\rho\tau, z_{\text{cut}}; C_R) \quad (8.12a)$$
$$R_{\text{out,SD}}(\rho\tau) = R_{\text{out,full}}(\rho\tau) - T_{-\beta 2}(\rho\tau,\zeta_{\text{cut}};C_R). \quad (8.12b)$$

This result splits into 2 contributions corresponding to the two terms in the round bracket on the second line of (8.11): the first term comes from configurations where there is no emission in region C with $z\theta^2 > \rho\tau_{\text{cut}}$, and it corresponds to values of $\tau_{21}^{\text{dichroic}} < 1$ (this is manifest, because in eq. (8.11), given for $\tau_{\text{cut}} > 1$, it has no dependence on τ_{cut}). For the second contribution, the part corresponding to values of $\tau_{21}^{\text{dichroic}} \geq 1$, there is an emission "c" with $z_c\theta_c^2 > \rho\tau_{\text{cut}}$. To guarantee $\tau_{21} < \tau_{\text{cut}}$, we then need to veto emissions (both primary and secondary) with $z\theta^2 > \rho\tau_{\text{cut}}$.[11] Note that this second contribution itself includes two sub-contributions: the case where emission "c" is the only emission in region C with $z\theta^2 > \rho$, yielding a contribution to the τ_{21} distribution proportional to $\delta(\tau_{21}-1)$ (recall that τ_2^{full} is set by the second hardest emission overall, which makes it equal to τ_1^{tagged}); and a second sub-contribution where, in addition to emission "c", there is at least one additional emission with $\rho\tau_{\text{cut}} > z\theta^2 > \rho$, yielding a continuum with $\tau_{21} > 1$ in the

[11] Note that the difference between the Sudakov suppression in the two contributions comes from secondary emissions, i.e. we have $R_{\text{out,full}}(\rho_c,\rho\tau_{\text{cut}}/\rho_c,z_c) = R_{\text{out,full}}(\rho\tau_{\text{cut}}) + T_0(\sqrt{\rho_c z_c}\tau_{\text{cut}}, \sqrt{\rho_c z_c}b_g; C_A)$.

τ_{21} distribution (see fig. 8.4 as well as the right plot of fig. 8.11 below). One can calculate the $\delta(\tau_{21}-1)$ contribution to the τ_{21} distribution by taking the difference between (8.11) and (8.4d) for $\tau_{\text{cut}} \to 1$ which gives

$$\int_{z_{\text{cut}}}^{b_i} \frac{dz}{z} \frac{\alpha_s(\sqrt{z\rho}p_t R)C_R}{\pi} \int_{\rho}^{z_{\text{cut}}} \frac{d\rho_c}{\rho_c} \int_{\rho_c}^{z_{\text{cut}}} \frac{dz_c}{z_c} \frac{\alpha_s(\sqrt{z_c\rho_c}p_t R)C_R}{\pi} e^{-R_{\text{full}}(\rho)-R_{C_A}(\rho_c,z_c,\rho)}, \quad (8.13)$$

where $R_{\text{full}}(\rho)$ is the full jet mass Sudakov given in eq. (8.8a), and $R_{C_A}(\rho_c, z_c, \rho) = T_0(\sqrt{z_c/\rho_c}\rho, \sqrt{\rho_c z_c}b_g; C_A)$. eq. (8.13) is equal to the $\tau_{\text{cut}} \to 1$ limit of the second term in round brackets in eq. (8.11). In practice the δ-function contribution gets smeared out to values of $\tau_{12} > 1$ through the effect of multiple emissions.

Note that it is relatively straightforward to check that the limit $\tau_{\text{cut}} \to 1$ in eq. (8.9), or the limit $\tau_{\text{cut}} \to z_{\text{cut}}/\rho$ in eq. (8.11) both tend to the mMDT jet mass distribution.

From the equations above, the τ_{21} distribution, for a given jet mass, can be obtained by taking the derivative with respect to τ_{cut} and normalising by the jet mass distribution without any cut on τ_{21}. Background efficiencies can also be obtained straightforwardly by integrating any of the above mass distributions over the allowed mass window.

For signal jets, we assume that if the jet mass is not within some reasonable window around the boson mass, then the jet is discarded. We then find the following signal efficiency

$$\varepsilon_S = f_{\text{ISR}} \int_{z_{\text{min}}}^{1-z_{\text{min}}} dz\, p_{\text{sig}}(z) \exp\Big[-R_{\text{sig}}(\rho, \tau_{\text{cut}}, z)\Big], \quad (8.14)$$

with $z_{\text{min}} = \rho$, $z_{\text{SD}}(\rho)$ or z_{cut} depending on whether the mass is computed on the full jet, the SD-groomed jet or the mMDT-tagged jet, respectively. The τ_{21} distribution for a given jet mass can be obtained by taking the derivative of ε_S with respect to τ_{cut} (and normalising appropriately). In eq. (8.14) the Sudakov exponent is given by

$$\begin{aligned} R_{\text{sig}}(\rho,\tau,z) = &\Big[T_{02}(\sqrt{z(1-z)\rho}\tau; \sqrt{(1-z)\rho/z}b_i; C_R) \\ &- T_{02}(\sqrt{z\rho/(1-z)}; \sqrt{(1-z)\rho/z}b_i; C_R)\Big] \\ &+ \Big[T_{02}(\sqrt{z(1-z)\rho}\tau; \sqrt{z\rho/(1-z)}b_i; C_R) \\ &- T_{02}(\sqrt{(1-z)\rho/z}; \sqrt{z\rho/(1-z)}b_i; C_R)\Big], \end{aligned} \quad (8.15)$$

valid for small τ. Here we target double-logarithmic accuracy, $\alpha_s^n \ln^{2n}\tau$, though we also include a set of finite-z and hard-splitting corrections that were found to be numerically important (see discussion in the end of chapter 7). These represent only a subset of next-to-leading logarithmic terms. Note that for $z \ll 1$ ($1-z \ll 1$) the term on the fourth (second) line is zero, while the term on the third (first) line corresponds to each one of the (symmetric) branches in the signal emission. For simplicity, in our numerical results we will use $p_{\text{sig}}(z) = 1$ in eq. (8.14).[12]

[12]For the WW process under consideration, correlations between the incoming quarks and the final

Eq. (8.14) also includes a factor f_{ISR} that accounts for the effect of initial-state radiation (ISR). Such effects are present both for signal and background jets and are generically single-logarithmic. As such they are subleading compared to the double-logarithms that we resum.

Nevertheless, if we consider signal jets and examine the limit of large p_t with M, τ_{cut}, etc. all fixed, then because of the absence of double logarithms of ρ, single-logarithmic ISR effects $(\alpha_s \ln \rho)^n$ can be numerically dominant [81]. Physically, they are associated with the requirement that ISR should not substantially modify the mass of the signal jet. The correction involves $(\alpha_s \ln \rho)^n$ terms, only when the mass is determined on the full jet and the factor f_{ISR} then takes the form

$$f_{\text{ISR}} = \exp\left[-\frac{C_R}{2\pi\beta_0} R^2 \log\frac{1}{1-2\lambda}\right], \tag{8.16}$$

$$\lambda = \beta_0 \alpha_s(p_t)\left[\log\frac{1}{\rho} + \mathcal{O}\left(\log\frac{M}{\delta M}, \log\frac{1}{\tau_{\text{cut}}}, \cdots\right)\right], \tag{8.17}$$

where a non-global contribution (formally of the same logarithmic order) is ignored for simplicity. In the above formula, δM is size of the mass window in which signal jets are accepted, and a full treatment of all single-logarithmic corrections would need to account also for logarithms of $\delta M/M$. A more complete treatment of f_{ISR} would be relevant for precise phenomenological applications. The finite $\mathcal{O}(\alpha_s)$ component associated with high-p_t emissions could be obtained e.g. using POWHEG [102, 103, 104], aMC@NLO [105] or at NNLO using MATRIX [106] or MCFM [107, 108, 109].

We can now compare our analytic predictions with the Monte Carlo results from the previous Section. We use $\alpha_s(m_Z) = 0.1383$, as in the Pythia 8 simulations presented in the previous Section, and freeze the coupling for scales below $\mu_{\text{fr}} = \tilde{\mu}_{\text{fr}} p_t R$, which we set to 1 GeV. We start with the QCD mass distributions, shown on the left plot of fig. 8.11, to be compared to the Monte Carlo results presented in fig. 8.5. Globally, we see that our analytic calculation captures correctly the main patterns discussed earlier. We note however that the analytic distributions, especially those involving the full jet mass, are less peaked than the Monte Carlo ones. This is likely due to subleading logarithmic corrections, like multiple-emission corrections which would effectively increase the Sudakov exponent.

The τ_{21} distributions for both QCD jets and signal (W) jets are shown in the right plot of fig. 8.11, to be compared with fig. 8.4. The ordering between the different curves is well captured by our analytic expressions. Differences related to the over-simplicity of our leading-logarithmic approximation are larger than what was seen for the mass distribution. First, our analytic calculations are non-zero when $\tau_{21} \to 1$. This region is however not under control within our strongly-ordered approximation. Similarly, the kink observed for $\tau_{21} \sim 0.5$ is not physical. It comes from the onset of the secondary-emission

quarks after the decay of the two W bosons have been calculated in [101] and could in principle be used to compute $p_{\text{sig}}(z)$. This would however be specific to the WW process considered here just as an example. We therefore use the "splitting function" of an unpolarised W boson. This simplification does not affect significantly any of the results presented here.

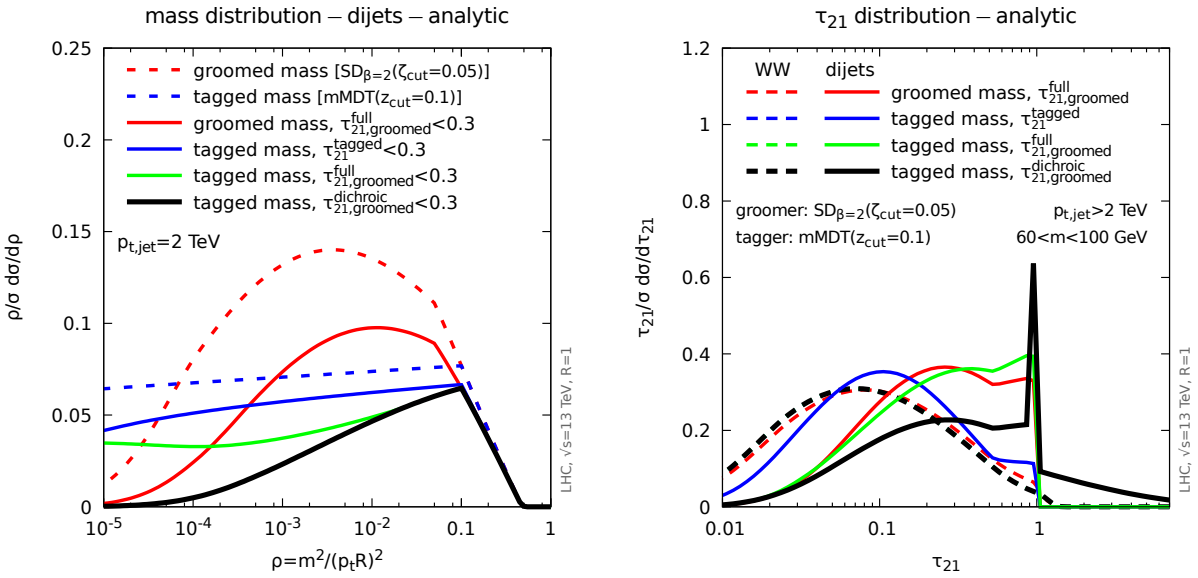


Figure 8.11 — Same as figure as 8.5 and 8.4 now obtained from our analytic calculation instead of Monte Carlo simulations. In the right-hand plot, for clarity, the δ-function that appears at $\tau_{21,\text{groomed}}^{\text{dichroic}} = 1$ (dijets) has been represented with finite width and scaled down by a factor of 5.

contribution which starts, in our formulas, at $\tau_{21} = b_g$. The analytic calculation for our dichroic combination is given by the black curves in the right plot of fig. 8.11. The dijet case clearly has a contribution proportional to $\delta(\tau_{21} - 1)$ (cf. eq. (8.13)) (scaled down by a factor of 5 for clarity), which is not observed in the Monte Carlo results. In practice, additional emissions at smaller $z\theta^2$ would also contribute to τ_{21}, and they would transform the $\delta(\tau_{21} - 1)$ contribution into a Sudakov peak at $\tau_{21} \gtrsim 1$, which is visible on the Monte Carlo simulations.

Finally, let us turn to the ROC curves, plotted in fig. 8.12. We again see that they reproduce the main qualitative features observed in section 8.2. There are however quantitative differences between our analytic results and the Monte Carlo simulations. For example, our calculation over-estimates the signal efficiencies. A more quantitative description would require a more precise analytic treatment including subleading corrections, beyond the strong-ordering approximation, and fixed-order corrections for signal efficiencies.

8.4 Final considerations

In this chapter we have examined the interplay between boosted-object tagging algorithms, mMDT or SoftDrop, and radiation constraints, notably as imposed through N-subjettiness cuts. The analysis points to a new N-subjettiness ratio, $\tau_{21}^{\text{dichroic}} = \tau_2^{\text{full}}/\tau_1^{\text{tagged}}$, where the numerator is evaluated on the full jet, while the denominator is evaluated on

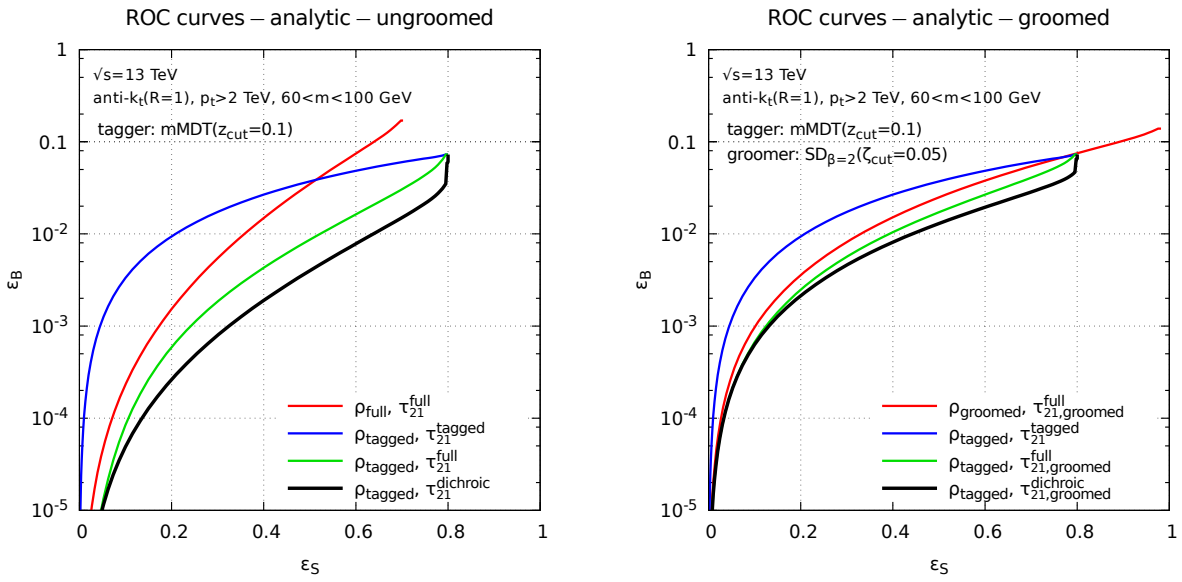


Figure 8.12 — Same as figure as 8.6 now obtained from our analytic calculation instead of Monte Carlo simulations.

the set of constituents left after the tagging stage. The name "dichroic" comes from the fact that the large-angle color flow, present in backgrounds but not signals, gets directed exclusively to the numerator and not the denominator. It is this feature that leads to an enhanced significance in distinguishing (color-singlet) signals from (color-triplet or octet) backgrounds, notably compared to current widely used N-subjettiness ratios. We have found that the combination of $\tau_{21}^{\text{dichroic}}$ with a light grooming step based on SoftDrop ($\beta = 2$), $\tau_{21,\text{groomed}}^{\text{dichroic}} = \tau_2^{\text{SD}}/\tau_1^{\text{tagged}}$ is effective in maintaining good signal-to-background significance while substantially limiting non-perturbative effects.

The overall behavior of our dichroic τ_{21} variable with grooming (see fig. 8.4), is that the τ_{21} distribution for signal jets is left largely unmodified by the change to a dichroic variant, whereas the distribution for background jets is shifted to substantially higher values of τ_{21}, increasing the ability to distinguish signal and background. Figs. 8.8 and 8.10 provide a summary of the signal-significance and non-perturbative corrections for a range of boosted-object identification methods. One sees that $\tau_{21,\text{groomed}}^{\text{dichroic}}$ with $\beta_\tau = 2$ provides the best signal significance of any of the methods and that, for a given signal significance, it tends to limit the size of non-perturbative effects relative to other methods.

— 9 —
Groomed jet mass distribution

In this chapter we perform a phenomenological study of the jet mass distribution with mMDT (or SoftDrop with $\beta = 0$) motivated by an upcoming CMS measurement [47].[1] In order to compare with experiments, we matched resummed and fixed-order predictions and computed theoretical uncertainty bands.

We consider jet mass distributions in several transverse momentum bins. Our theoretical prediction accounts for the resummation of the leading large logarithms of ρ and is matched to fixed-order matrix elements computed at next-to-leading order (NLO), we also consider finite $z_{\rm cut}$ effects, which enter already at LL accuracy. Crucially, working at finite $z_{\rm cut}$ allows us to keep track of the distinction between the jet transverse momentum before or after grooming, henceforth $p_{t,\rm jet}$ and $p_{t,\rm mMDT}$, respectively. We find that the use of $p_{t,\rm mMDT}$ has several theoretical disadvantages with respect to $p_{t,\rm jet}$. While the two resummations coincides for $z_{\rm cut} \to 0$, the former has a more involved perturbative structure already at leading order. This difference stems from a basic fact, namely while the ungroomed $p_{t,\rm jet}$ spectrum is an Infra-Red and Collinear (IRC) safe quantity, the jet $p_{t,\rm mMDT}$ spectrum (with no additional cuts) is Sudakov safe but not IRC safe. Conversely, the $p_{t,\rm mMDT}$ spectrum is slightly less sensitive to the underlying event than $p_{t,\rm jet}$ one and, arguably, more resilient to pile-up. It is therefore interesting to explore both options in more details.

This chapter is organized as follows. Resummation and matching of the mass distribution with $p_{t,\rm jet}$ are done in section 9.1, followed by the case of $p_{t,\rm mMDT}$ in section 9.2. A Monte Carlo study of non-perturbative corrections is presented in section 9.3 and we collect our final phenomenological predictions and discuss them in section 9.4.

9.1 Jet mass distributions with mMDT

In this chapter, we focus on the invariant mass of a mMDT jet produced in proton-proton collisions with a centre-of-mass energy of 13 TeV. Our selection cuts closely follow the ones of the upcoming CMS measurement [47]: jets are defined with the anti-k_t algorithm

[1] The CMS results have recently become public and are is good agreement with our calculation.

with jet radius $R = 0.8$. Next, we select the two hardest jets, j_a and j_b, of the event and impose the following conditions:

1. both jets must have $p_{t,\text{jet}} > 200\text{GeV}$ and central rapidity, namely $|y| < 2.4$;
2. the transverse momenta of the jets must satisfy $|p_{ta} - p_{tb}| < 0.3(p_{ta} + p_{tb})$ in order to select symmetric configurations;
3. the jets should be well-separated in azimuth, i.e. $\Delta\phi_{j_a,j_b} > \pi/2$.

In practice, these cuts are intended to select dijet events. We note however that the transverse momentum cut on the second jet results in large perturbative corrections for the dijet cross-section which render the mass distribution unstable in the first transverse momentum bin. Imposing only a p_t cut on the leading jet and the symmetry condition would have been similarly efficient at selecting dijet events, and would have improved the perturbative convergence.

For every jet that passes the above cuts, we apply the mMDT procedure with $z_{\text{cut}} = 0.1$. We compute the (groomed) jet mass squared $m^2 = \left(\sum_i p_i\right)^2$, where the sum runs over all particles in the groomed jet. We continue using the dimensionless variable $\rho = m^2/(p_{t,\text{jet}}^2 R^2)$.

We calculate the ρ distribution in a given transverse momentum bin $p_{t1} < p_{t,\text{jet}} < p_{t2}$:

$$\frac{d\sigma}{d\rho}\left(\rho; z_{\text{cut}}, p_{t1}, p_{t2}\right) = \int_{p_{t1}}^{p_{t2}} dp_{t,\text{jet}} \frac{d^2\sigma}{dp_{t,\text{jet}} d\rho}. \tag{9.1}$$

We also define the normalised distribution as

$$\frac{d\tilde{\sigma}}{d\rho}\left(\rho; z_{\text{cut}}, p_{t1}, p_{t2}\right) = \frac{1}{\sigma_{\text{bin}}(p_{t1}, p_{t2})} \frac{d\sigma}{d\rho}\left(\rho; z_{\text{cut}}, p_{t1}, p_{t2}\right), \tag{9.2}$$

where σ_{bin} is the jet cross section in the transverse momentum bin under consideration. We also explicitly consider the jet mass distribution

$$\frac{d\sigma}{dm}\left(m; z_{\text{cut}}, p_{t1}, p_{t2}\right) = \int_{p_{t1}}^{p_{t2}} dp_{t,\text{jet}} \frac{d^2\sigma}{dp_{t,\text{jet}} dm} = \int_{p_{t1}}^{p_{t2}} dp_{t,\text{jet}} \frac{2m}{p_{t,\text{jet}}^2 R^2} \frac{d^2\sigma}{dp_{t,\text{jet}} d\rho}, \tag{9.3}$$

and the corresponding normalised version. Furthermore, the quantity that is measured experimentally is the mass distribution integrated over a set of mass bins $m_i < m < m_{i+1}$, which is the observable we are going to explicitly show in our plots. Note that in eq. (9.1) $p_{t,\text{jet}}$ is the jet transverse momentum *before* grooming. We will consider the alternative choice, namely the groomed transverse momentum $p_{t,\text{mMDT}}$ in section 9.2.

The leading logarithmic resummation of mMDT jet masses has been performed in [12] and resummation for SoftDrop observables, i.e. for generic β, was performed to NLL accuracy in [17] and to NNLL accuracy in [74, 75]. All the logarithmic contributions in SoftDrop observables are of collinear origin, while soft-emission at large angle can at most contribute with logarithms of z_{cut}. Thanks to this observation, the resummed calculation can be done in the collinear limit and the resulting structure is much simpler than the one

that we encounter in the resummation of the jet mass distributions without grooming, see for instance [76, 84, 110]. In particular, soft radiation at large angle, which would give rise to a nontrivial matrix structure in color space, is groomed away: only dipoles involving the measured jet are logarithmically enhanced and require resummation, while initial-state radiation does not contribute. For the same reason, these observables are free of non-global logarithms.

At this stage, a word of caution about our counting of the logarithmic accuracy is in order. While for a generic (non-zero) β, the SoftDrop mass distribution is dominated by double logarithms — with LL accuracy resumming those double logarithms, NLL accuracy including single-logarithms as well, etc... — these double logarithms are absent for mMDT (i.e. SoftDrop with $\beta = 0$) in the region $\rho < z_{\rm cut}$:

$$\rho \frac{d\tilde{\sigma}}{d\rho}(\rho; z_{\rm cut}) = \left[\sum_{n=1}^{\infty} \sum_{m=1}^{n} c_{n,m}(z_{\rm cut})\, \alpha_s^n \log^{m-1}\left(\frac{1}{\rho}\right) + \mathcal{O}(\rho) \right], \tag{9.4}$$

where the dependence on the transverse momentum bin is understood. Single logarithmic terms in the jet mass are therefore formally the leading contribution and will be referred to as LL in what follows. Note that this counting is different from the (modified) leading-logarithm we used in previous sections, where our accuracy was double logarithms (both of the jet shape and of ρ).

Also, the logarithmic counting of refs. [74, 75] differs from ours because it refers to the accuracy of the objects that appear in the factorization theorem. These functions are separately double-logarithmic, even for $\beta = 0$, and the cancellation of the double logarithms only happens when they are combined.[2] In our counting, the NLL [17] and NNLL [74, 75] results obtained for a generic β, actually correspond respectively to LL and NLL accuracy, in the small $z_{\rm cut}$ limit, for mMDT. Thus, the state-of-the art evaluation of eq. (9.4) accounts for all the coefficients $\tilde{c}_{n,n}(z_{\rm cut})$ and $\tilde{c}_{n,n-1}(z_{\rm cut})$, where

$$\lim_{z_{\rm cut}\to 0} c_{n,m}(z_{\rm cut}) = \tilde{c}_{n,m}(z_{\rm cut}) + \mathcal{O}(z_{\rm cut}). \tag{9.5}$$

For phenomenology, one typically uses $z_{\rm cut} \simeq 0.1$, so it is important to investigate the size of finite $z_{\rm cut}$ corrections. In this study we restrict ourselves to LL accuracy, while maintaining for the full $z_{\rm cut}$ dependence, i.e. we fully account for all coefficients $c_{n,n}(z_{\rm cut})$.

Finally, in the region $\rho > z_{\rm cut}$ grooming is not active and we recover the traditional jet mass result [12]. In this region we are going to perform a less sophisticated calculation which resums the double logarithms and those single logarithmic contributions of collinear origin. We find this procedure acceptable because in this region $\rho \sim z_{\rm cut}$ and we expect these contributions to be less important than the fixed-order corrections, which we include at NLO.

9.1.1 Resummation at finite $z_{\rm cut}$

The major complication with respect to the small-$z_{\rm cut}$ limit has to do with the flavor structure. Let us consider for instance a $q \to qg$ splitting which does not satisfy the

[2] We would like to thank Andrew Larkoski for clarifying this point.

mMDT condition. There is an $\mathcal{O}(z_{\text{cut}})$ probability for the gluon to be harder than the quark. In such a case, the declustering sequence would follow the gluon branch rather than the quark, resulting into a nontrivial mixing between quarks and gluons. The resummed distribution therefore acquires a matrix structure in flavor space [12] [3]

$$\rho\frac{d^2\sigma}{dp_{t,\text{jet}}d\rho} = (R'_q \;\; R'_g)\exp\begin{pmatrix} -R_q - R_{q\to g} & R_{g\to q} \\ R_{q\to g} & -R_g - R_{g\to q}\end{pmatrix}\begin{pmatrix}\sigma_q \\ \sigma_g\end{pmatrix}, \tag{9.6}$$

where $\sigma_{q(g)}$ is Born-level cross section for a quark (gluon) with transverse momentum $p_{t,\text{jet}}$ and $R'_{q(g)} = \partial_L R_{q(g)}$, with $L = \log(1/\rho)$. As previously discussed, because we are dealing with a SoftDrop observable, the radiators R_i can be computed in the collinear limit. Denoting by θ the emission angle (in units of the jet radius R) with respect to the hard momentum and with z the momentum fraction, we have

$$R_q = \int_0^1\frac{d\theta^2}{\theta^2}\int_0^1 dz\; p_{gq}(z)\frac{\alpha_s(z\theta p_{t,\text{jet}}R)}{2\pi}\Theta\left(z_{\text{cut}} < z < 1 - z_{\text{cut}}\right)\Theta(z\theta^2 > \rho), \tag{9.7a}$$

$$R_g = \int_0^1\frac{d\theta^2}{\theta^2}\int_0^1 dz\; p_{xg}(z)\frac{\alpha_s(z\theta p_{t,\text{jet}}R)}{2\pi}\Theta\left(z_{\text{cut}} < z < 1 - z_{\text{cut}}\right)\Theta(z\theta^2 > \rho), \tag{9.7b}$$

$$R_{q\to g} = \int_0^1\frac{d\theta^2}{\theta^2}\int_0^1 dz\; p_{gq}(z)\frac{\alpha_s(z\theta p_{t,\text{jet}}R)}{2\pi}\Theta\left(1 - z < z_{\text{cut}}\right)\Theta(z\theta^2 > \rho), \tag{9.7c}$$

$$R_{g\to q} = \int_0^1\frac{d\theta^2}{\theta^2}\int_0^1 dz\; p_{qg}(z)\frac{\alpha_s(z\theta p_{t,\text{jet}}R)}{2\pi}\left[\Theta\left(1 - z < z_{\text{cut}}\right) + \Theta\left(z < z_{\text{cut}}\right)\right]\Theta(z\theta^2 > \rho), \tag{9.7d}$$

recalling that we are using the QCD constants defined in chapter 2 and $p_{ab}(z)$ are the splitting functions given in chapter 5.

At the LL accuracy we are working at, the above expressions can be further simplified. Besides the strict leading-logarithmic terms in ρ, it is trivial to also include the double-logarithmic terms in z_{cut} and this allows for a more transparent treatment of the transition point at $\rho = z_{\text{cut}}$. In that context, it is helpful to separate eq. (9.7) in a contribution $\mathcal{R}_i$, coming from the $1/z$ part of the splitting function that includes the logarithmic and constant terms in z_{cut}, and a remainder which contains the corrections power-suppressed in z_{cut}. Later, this will make it easy to study the size of the finite-z_{cut} corrections. For these contributions, we neglect the z factor in the argument of α_s and in the constraint $z\theta^2 > \rho$. The details of our calculation are given in appendix D.1.1 and, our final result reads

$$R_q = C_F\mathcal{R}_q(\rho; z_{\text{cut}})\,\Theta(\rho < e^{B_q}) + C_F\,\mathcal{I}(\rho; z_{\text{cut}})\,\pi_q(z_{\text{cut}})\,\Theta(\rho < z_{\text{cut}}), \tag{9.8a}$$

$$R_g = C_A\mathcal{R}_g(\rho; z_{\text{cut}})\,\Theta(\rho < e^{B_g}) + C_A\,\mathcal{I}(\rho; z_{\text{cut}})\,\pi_g(z_{\text{cut}})\,\Theta(\rho < z_{\text{cut}}), \tag{9.8b}$$

$$R_{q\to g} = C_F\,\mathcal{I}(\rho; z_{\text{cut}})\,\pi_{q\to g}(z_{\text{cut}})\,\Theta(\rho < z_{\text{cut}}), \tag{9.8c}$$

$$R_{q\to g} = n_f T_R\,\mathcal{I}(\rho; z_{\text{cut}})\,\pi_{g\to q}(z_{\text{cut}})\,\Theta(\rho < z_{\text{cut}}), \tag{9.8d}$$

[3] More precisely, the resummation of ref. [12] was performed in case of a y_{cut}, but its modification to a z_{cut} is straightforward.

where we have introduced

$$\mathcal{R}_i(\rho; z_{\rm cut}) = \frac{1}{2\pi\alpha_s\beta_0^2}\Big[W\Big(1+2\alpha_s\beta_0 B_i\Big) - W\Big(1+2\alpha_s\beta_0\log(z_m)\Big) + 2W\Big(1+\alpha_s\beta_0\log(\rho z_m)\Big) - 2W\Big(1+\alpha_s\beta_0(\log(\rho)+B_i)\Big)\Big], \tag{9.9a}$$

$$\mathcal{I}(\rho; z_{\rm cut}) = \int_\rho^{z_{\rm cut}} \frac{dx}{x}\frac{\alpha_s(xp_tR)}{\pi} = \frac{1}{\pi\beta_0}\log\left(\frac{1+\alpha_s\beta_0\log(z_{\rm cut})}{1+\alpha_s\beta_0\log(\rho)}\right), \tag{9.9b}$$

with $W(x) = x\log(x)$, $z_m = \max(z_{\rm cut}, \rho)$, B_q and B_g given in (5.12) and (5.13),

$$\pi_q(z_{\rm cut}) = \log(1-z_{\rm cut}) + \frac{3z_{\rm cut}}{2}, \tag{9.10a}$$

$$\pi_g(z_{\rm cut}) = \log(1-z_{\rm cut}) + 2z_{\rm cut} - \frac{z_{\rm cut}^2}{2} + \frac{z_{\rm cut}^3}{3} - \frac{n_f T_R}{C_A}\left(z_{\rm cut} - z_{\rm cut}^2 + \frac{2z_{\rm cut}^3}{3}\right), \tag{9.10b}$$

$$\pi_{q\to g}(z_{\rm cut}) = -\log(1-z_{\rm cut}) - \frac{z_{\rm cut}}{2} - \frac{z_{\rm cut}^2}{4}, \tag{9.10c}$$

$$\pi_{g\to q}(z_{\rm cut}) = z_{\rm cut} - z_{\rm cut}^2 + \frac{2z_{\rm cut}^3}{3}. \tag{9.10d}$$

We note that the diagonal radiators vanish for $\rho = \exp(B_i)$ and, since B_q is (slightly) larger than B_g, this produces distributions with an end-point at $\rho = \exp(B_q)$. Furthermore, the appearance of $z_m = \max(z_{\rm cut}, \rho)$ reproduces the transition point at $\rho = z_{\rm cut}$, when the mMDT becomes active. We show explicitly below that it corresponds to a transition between a plain jet mass behaviour at large mass and a single-logarithmic behaviour at low mass.

To gain some insight in this direction, it is helpful to consider the limit of these expressions in a fixer-order approximation, where we find

$$\mathcal{R}_i^{(\rm f.c.)}(\rho; z_{\rm cut}) = \frac{\alpha_s}{2\pi}\Big[\Big(\log(\rho) - B_i\Big)^2 - \log^2(z_m/\rho)\Big], \tag{9.11a}$$

$$\mathcal{I}^{(\rm f.c.)}(\rho; z_{\rm cut}) = \int_\rho^{z_{\rm cut}} \frac{dx}{x}\frac{\alpha_s(xp_tR)}{\pi} = \frac{\alpha_s}{\pi}\log\left(\frac{z_{\rm cut}}{\rho}\right). \tag{9.11b}$$

This clearly shows that the distribution is double-logarithmic for $\rho > z_{\rm cut}$ (where $z_m = \rho$), where we recover the ungroomed jet mass distribution calculated in chapter 5, and becomes single-logarithmic for $\rho < z_{\rm cut}$ (where $z_m = z_{\rm cut}$). In the latter case, we also see that the finite-$z_{\rm cut}$ corrections, proportional to $\mathcal{I}$ are entering at the same order as the small-$z_{\rm cut}$ contributions, that is at the leading-logarithmic accuracy. Thus, these contributions must be included to formally obtain the full LL result.

In order to assess perturbative uncertainties we follow a standard procedure. We vary the factorisation scale (in the Born-level cross-sections σ_q and σ_g) and the renormalisation scale (both in the resummation formula and in the Born-level cross-sections) by a factor of two around the hard scale $p_{t,\rm jet}R$, keeping the ratio of scales never larger than 2 or smaller than 1/2, i.e. we employ a canonical 7-point scale variation [111]. We also introduce a

resummation scale μ_Q, which we use to rescale the argument of the logarithms we are resumming $L = \log \frac{p_{t,\text{jet}} R}{\mu_Q \rho}$. We use variations of μ_Q by a factor of 2 around the hard scale $p_{t,\text{jet}} R$ to assess the size of logarithmic contributions beyond our accuracy.

9.1.2 Fixed-order calculations and matching prescription

The resummed jet mass spectrum discussed in the previous section is reliable in the $\rho \ll 1$ region, where the distribution is dominated by collinear splittings. In order to accurately describe the $\rho \sim 1$ region we have to resort to fixed-order computations. Ultimately, we will match the two calculations yielding theoretical predictions which are accurate at both small and large ρ, as discussed in the following.

All our fixed-order predictions are obtained using the public code NLOJet++ [86, 112] together with the parton distribution set CT14 [113] at NLO. Jets are then clustered with the anti-k_t algorithm as implemented in FastJet [38, 39] and we use the implementation of mMDT in fjcontrib [83]. Jet mass distributions are obtained by considering $2 \to 3$ partonic processes at LO and NLO. Moreover, we also use NLOJet++ to calculate the bin cross section σ_{bin}, see eq. (9.2), and the quark and gluon cross sections, σ_q and σ_g respectively. In order to estimate the theoretical uncertainty, we vary renormalisation and factorisation scales around the central value $\mu_R = \mu_F = p_{t,\text{jet}} R$, with the 7-point method.

We are now ready to match the resummed and the fixed-order calculations. Before discussing different matching schemes, we address the issue of the end-point of the distribution at large ρ. It is not difficult to show, see e.g. [84], that the LO distribution has an end-point at $\rho_{\text{max,LO}} = \frac{1}{4} + \mathcal{O}(R^2)$. At NLO up to three partons can be reconstructed in a single jet, leading to $\rho_{\text{max,NLO}} = \frac{25}{64} + \mathcal{O}(R^2)$ (see appendix D.2 for details). On the other hand, our resummed calculation has an end-point at $\rho = \exp(B_q)$, see eq. (9.8). It is desirable to match curves with the same end-point, therefore we modify the argument of the logarithms in the resummation in such a way that the resummed distribution has the same end-point as the fixed-order it is matched to (see e.g. [78])

$$\log\left(\frac{1}{\rho}\right) \to \log\left(\frac{1}{\rho} - \frac{1}{\rho_{\text{max},i}} + e^{-B_q}\right), \tag{9.12}$$

where for $R = 0.8$ the end-points are found to be $\rho_{\text{max,LO}} = 0.279303$ and $\rho_{\text{max,NLO}} = 0.44974$ (see appendix D.2).

The combination of resummed and fixed-order results comes with a certain degree of ambiguity. Different matching schemes must produce resummed and matched distributions, LO+LL and NLO+LL, at the quoted accuracy but they can differ for terms that are subleading in both logarithmic and fixed-order counting. The simplest matching scheme is the additive one, which consists of adding the two results while removing double counting. This scheme suffers from two issues. Firstly, when matching to NLO fixed-order results, our LL calculation only includes the leading $\alpha_s^2 \log(1/\rho)$ contribution and misses the constant α_s^2 term, so an additive matching would tend to a constant at small ρ which is not physically correct. Secondly, even at LO, matching with our LL calculation requires a precise numerical calculation of the small-ρ tail, which can be delicate to reach in the

fixed-order calculation. Therefore, we have decided to employ an alternative matching scheme, namely multiplicative matching. We discuss it in some detail for the NLO+LL case and then recover from it the simpler LO+LL. Naively, multiplicative matching can be defined as

$$\sigma^{(m)}_{\text{NLO+LL,naive}} = \frac{\sigma^{(m)}_{\text{LL}}\,\sigma^{(m)}_{\text{NLO}}}{\sigma^{(m)}_{\text{LL,NLO}}}, \tag{9.13}$$

where, to keep the notation compact, $\sigma^{(m)}_X$ indicates the jet mass differential distribution computed at accuracy X, i.e. $\sigma^{(m)}_X \equiv \frac{d\sigma_X}{dm}$. This construction applies both to the normalised and unnormalised distributions.

Equation (9.13) is however not ideal either because at NLO accuracy, the fixed-order cross-section turns negative at small mass. Asymptotically both $\sigma^{(m)}_{\text{NLO}}$ and $\sigma^{(m)}_{\text{LL,NLO}}$ would be negative and their ratio would tend to 1 but there is a region where they would be close to zero and where eq. (9.13) would therefore be unreliable. To fix this issue, we can write the fixed-order distribution explicitly as

$$\sigma^{(m)}_{\text{NLO}} = \sigma^{(m)}_{\text{LO}} + \alpha_s \delta^{(m)}_{\text{NLO}}, \tag{9.14}$$

while the expansion of the resummation to second order is

$$\sigma^{(m)}_{\text{LL,NLO}} = \sigma^{(m)}_{\text{LL,LO}} + \alpha_s \delta^{(m)}_{\text{LL,NLO}}. \tag{9.15}$$

We can then substitute eq. (9.14) and (9.15) into eq. (9.13) and expand to the desired accuracy, to obtain

$$\sigma^{(m)}_{\text{NLO+LL}} = \sigma^{(m)}_{\text{LL}} \left[\frac{\sigma^{(m)}_{\text{LO}}}{\sigma^{(m)}_{\text{LL,LO}}} + \alpha_s \left(\frac{\delta^{(m)}_{\text{NLO}}}{\sigma^{(m)}_{\text{LL,LO}}} - \sigma^{(m)}_{\text{LO}} \frac{\delta^{(m)}_{\text{LL,NLO}}}{{\sigma^{(m)}_{\text{LL,LO}}}^2} \right) \right]. \tag{9.16}$$

This is the expression we use in order to obtain our matched results. The LO+LL results can be easily deduced from the above expression by simply dropping the $\mathcal{O}(\alpha_s)$ correction in brackets, in which case the expression corresponds to what would have been obtained with a naive multiplicative matching. We can also define alternative matching schemes. For instance, we can work with cumulative distributions

$$\Sigma_X(m) = \int_0^m dm' \frac{d\tilde{\sigma}_X}{dm'} = 1 + \alpha_s \Sigma^{(1)}_X + \alpha_s^2 \Sigma^{(2)}_X + \mathcal{O}\left(\alpha_s^3\right), \tag{9.17}$$

and employ the so-called log-R matching [78], which combines together the logarithm of the cumulative distributions. This results in

$$\Sigma^{\text{log-}R}_{\text{NLO+LL}} = \Sigma_{\text{LL}} \exp\left[\alpha_s \left(\Sigma^{(1)} - \Sigma^{(1)}_{\text{LL}} \right) + \alpha_s^2 \left(\Sigma^{(2)} - \Sigma^{(2)}_{\text{LL}} \right) - \frac{\alpha_s^2}{2} \left({\Sigma^{(1)}}^2 - {\Sigma^{(1)}_{\text{LL}}}^2 \right) \right]. \tag{9.18}$$

A comparison between the different matching schemes will be discussed in the following.

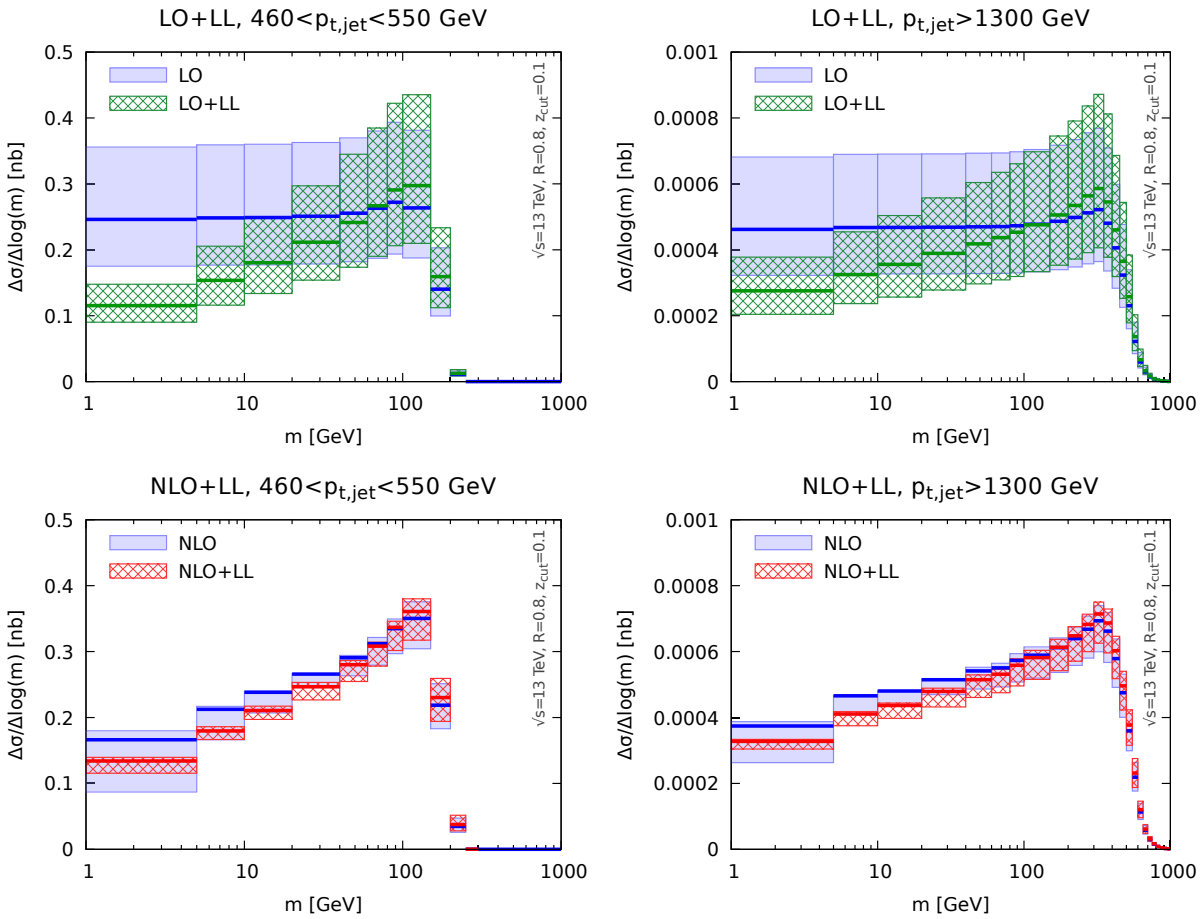


Figure 9.1 — In this figure we show the resummed and matched jet mass distribution in the $460 < p_{t,\text{jet}} < 550$ GeV transverse momentum bin (on the left), and in the $p_{t,\text{jet}} > 1300$ GeV bin (on the right). The top panels show LO+LL, while the bottom panels NLO+LL.

9.1.3 Perturbative results

We now present our results for the resummed and matched jet mass distribution. We pick two representative bins in transverse momentum, namely $460 < p_{t,\text{jet}} < 550$ GeV and $p_{t,\text{jet}} > 1300$ GeV. In fig. 9.1 we show the mass distribution in logarithmic bins of the mass:[4]

$$\frac{\Delta\sigma}{\Delta \log m} \equiv \frac{m_{i+1} - m_i}{\log\left(m_{i+1}/m_i\right)} \frac{\Delta\sigma}{\Delta m}, \tag{9.19}$$

where m_{i+1} and m_i are, respectively, the upper and lower edge of each mass bin. Blue lines with a solid band represent distributions obtained with fixed-order calculations and their uncertainty, while green or red curves with a hatched band are for resummed and matched results obtained using eq. (9.16). We estimate the theoretical uncertainty on the matched result by taking the envelope of all the curves obtained by varying the arbitrary scales $(\mu_\text{R}, \mu_\text{F}, \mu_\text{Q})$ which enter the fixed-order and resummed calculations, as previously detailed. At the top we compare leading order distributions to LO+LL results, while at the bottom we show the NLO curve compared to NLO+LL. The plots on the left are for the lower-$p_{t,\text{jet}}$ bin, while the ones on the right for the boosted bin. We can see that the normalisation uncertainty is rather large especially when we consider LO distributions. Therefore, it is also interesting to look at normalised distribution, with the normalisation taken to be the jet cross-section in the relevant transverse momentum bin calculated at LO and NLO, respectively for the LO(+LL) and NLO(+LL) results. We show our results for the normalised distributions in fig. 9.2.

In order to estimate the importance of finite z_cut we compared in fig. 9.3 the resummed and matched NLO+LL normalised distribution, in red, to an approximation in which the resummation is performed in the $z_\text{cut} \to 0$ limit, in grey, for two different transverse momentum bins. From the top plots we can already see that, for $z_\text{cut} = 0.1$, these effects are small and the two curves fall well within each other's uncertainties bands. Looking at the bottom plots we can see that these effects at at most a couple of percent at NLO+LL (red curves). For comparison, we also show, in green, the same ratio in the case of the LO+LL result. Note that the bands in the ratio plots represent the uncertainty on the effect, not the overall uncertainty which is of the order of 10%, as can be seen from the top plots. These findings justify the approximation of refs. [74, 75], which achieved higher-logarithmic accuracy but in the small-z_cut limit. We will see in the next section that the situation radically changes when consider bins in $p_{t,\text{mMDT}}$.

Finally, in fig. 9.4 we compare two different matching schemes. In particular, we plot the ratio between the NLO+LL distribution obtained with log-R matching eq. (9.18) to the one obtained with multiplicative matching eq. (9.16), with their respective perturbative uncertainties. We see that the two results are in good agreement and they fall within each other's scale variation bands.

[4] The binned distribution is computed using eq. (9.3). For a given $p_{t,\text{jet}}$ we thus need to integrate $\rho d^2\sigma/(dp_{t,\text{jet}} d\rho)$ over a range in ρ. In practice, this can be written as a difference between the cumulative ρ distribution taken at the bin edges, which, for the resummed results, is obtained by removing the $(R'_q \;\; R'_g)$ pre-factor in eq. (9.6).

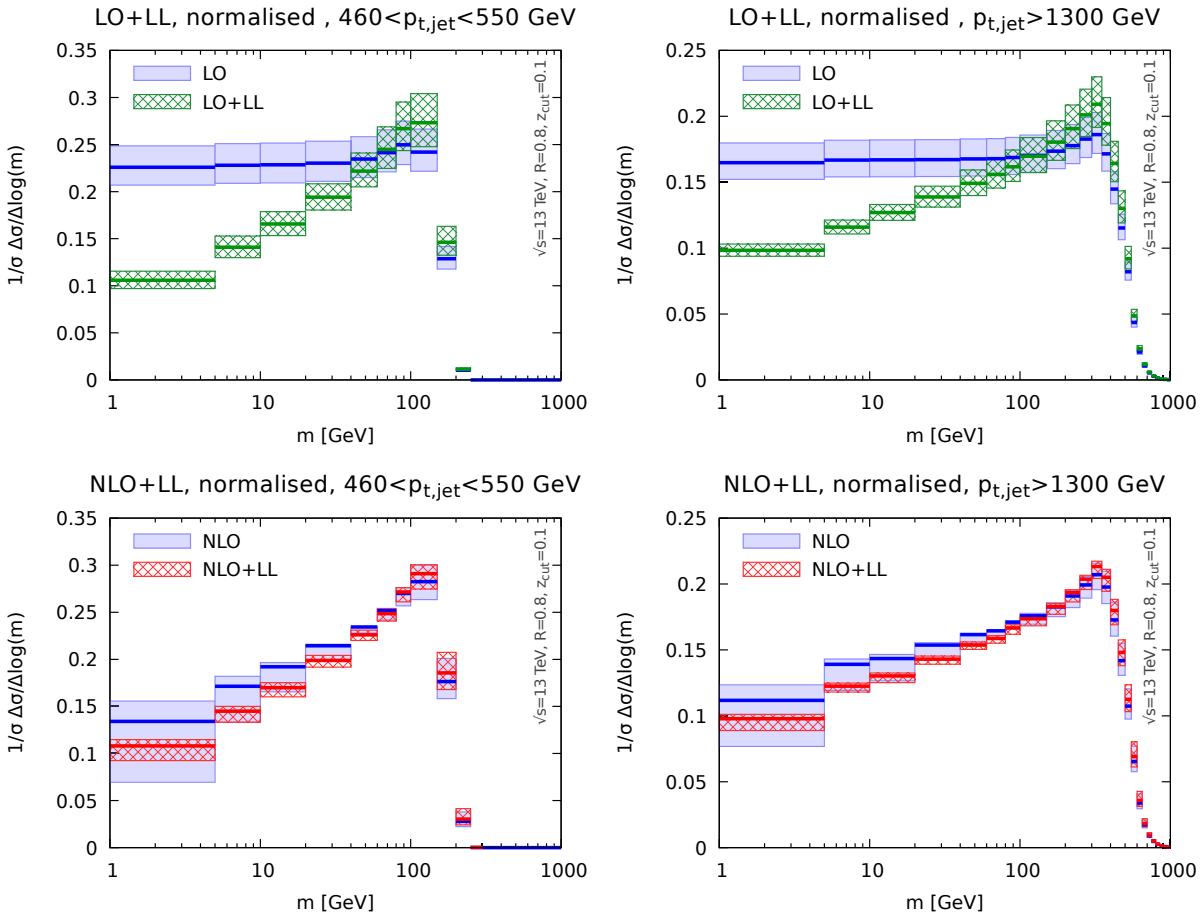


Figure 9.2 — Same as in fig. 9.1 but for the normalised distribution.

9.2 Jet mass distributions with mMDT using $p_{t,\mathrm{mMDT}}$

We now consider the alternative option where the mMDT jet mass is measured in bins of $p_{t,\mathrm{mMDT}}$ rather than $p_{t,\mathrm{jet}}$. We begin our discussion pointing out a known but perhaps under-appreciated fact: the transverse momentum distribution $\frac{d\sigma}{dp_{t,\mathrm{mMDT}}}$ is not IRC safe, see e.g. [17]. We then proceed, as before, by discussing our calculation for the jet mass distribution in bins of $p_{t,\mathrm{mMDT}}$.

9.2.1 Collinear unsafety (but Sudakov safety) of $p_{t,\mathrm{mMDT}}$

The mMDT groomer only imposes a cut on the transverse momentum fraction z. Therefore, real emission emissions below z_{cut} are groomed away without any constraint on the emission angle, resulting in collinear singularities that do not cancel against the corresponding virtual corrections. Thus, the $p_{t,\mathrm{mMDT}}$ distribution is IRC unsafe and it cannot be computed order-by-order in the strong coupling α_s, producing a divergence even at the level of the first emission. However, this observable still enjoys the property of Sudakov

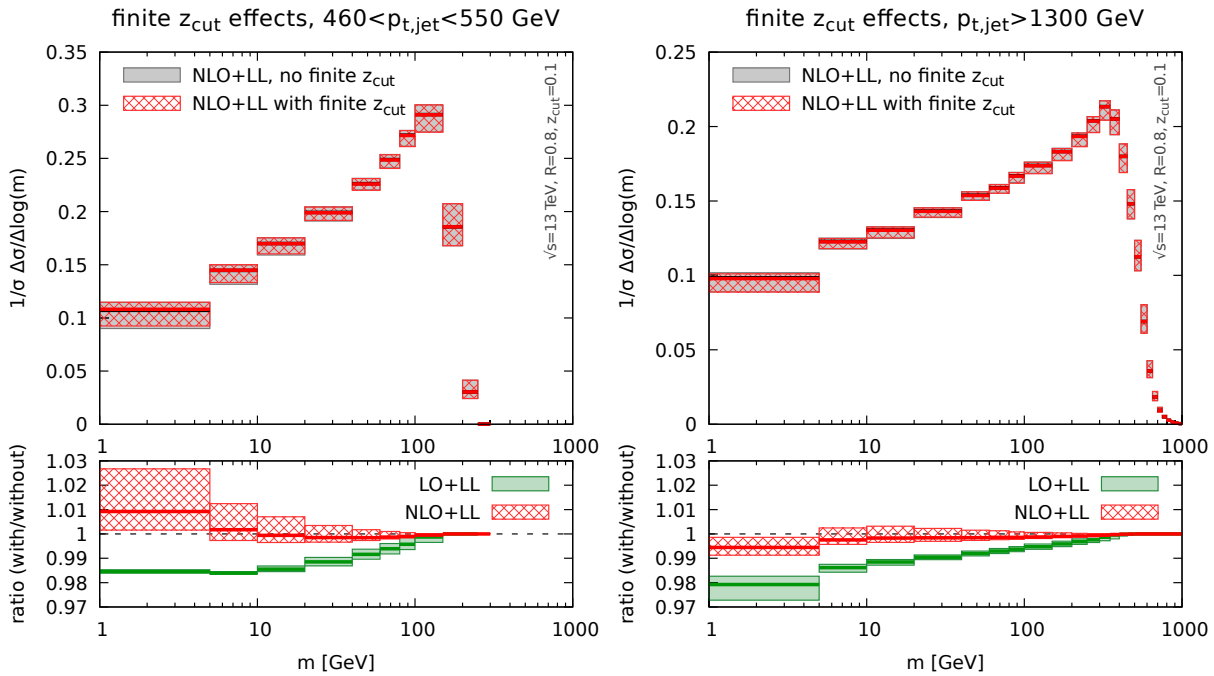


Figure 9.3 — Comparison between the resummed and matched calculation with finite z_{cut} (red) and the result with the resummation computed in the $z_{\mathrm{cut}} \to 0$ limit. The ratio plots at the bottom show that for $z_{\mathrm{cut}} = 0.1$ these type of corrections are very small.

safety and it is therefore calculable provided we perform an all-order computation. We note that the situation is instead different if one considers SoftDrop with $\beta > 0$, which does regulate the collinear region.

One way to explicitly show the IRC unsafety of the $p_{t,\mathrm{mMDT}}$ distribution is to study fixed-order distributions in e^+e^- collisions using the program EVENT2 [114, 86], for which we can easily control the infrared cutoff scale. In practice, we simulate events at Born level and at $\mathcal{O}(\alpha_s)$, including both real emissions and virtual corrections. We cluster the full event with the e^+e^- version of the anti-k_t algorithm with radius $R = 0.4$ and select jets with an energy larger than $0.95\sqrt{s}/2$, with $\sqrt{s} = 1$ TeV. We note that, at this order in perturbation theory, jets have either one or two constituents. We then run the following e^+e^- version of mMDT: jets with one constituent are kept untouched, and for jets with two constituents we either keep them intact if $\min(E_1, E_2) > z_{\mathrm{cut}} E_{\mathrm{jet}}$, or only keep the most energetic particle otherwise. We use $z_{\mathrm{cut}} = 0.1$. We consider the jet cross section for $E > 0.95\sqrt{s}/2$ before and after applying mMDT. At Born level, jets after the mMDT procedure are identical to the ungroomed jets. At $\mathcal{O}(\alpha_s)$, for an initial jet with an energy above the cutoff, the mMDT jet energy can drop below the cutoff because of a collinear real emission inside the jet that does not pass the mMDT condition. This cannot happen for virtual corrections and so we do expect a leftover singularity.

In numerical codes, both the real and the virtual terms are simulated down to an infrared cutoff so that the numerical result is always finite. When lowering the infrared cutoff parameter the cross section for the ungroomed case is expected to remain constant

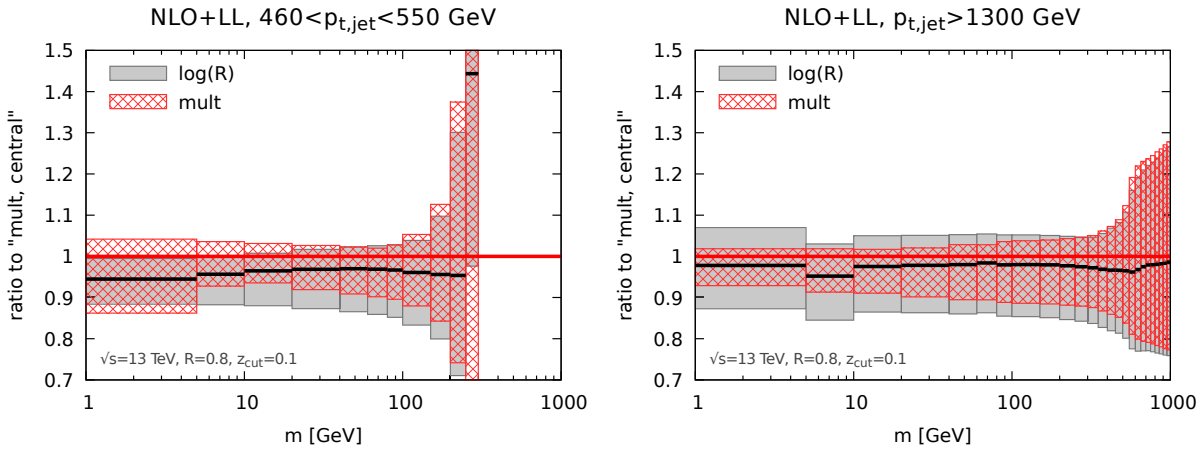


Figure 9.4 — Comparison of the jet mass distribution in two different matching schemes, the multiplicative one eq. (9.16) and the log-R one eq. (9.18).

(modulo small power corrections), while the cross section for mMDT jets is expected to have a residual logarithmic dependence on the cut-off as a consequence of the IRC unsafety. Fig. 9.5 shows the results of our simulations when varying the infra-red cut-off used in EVENT2 . We indeed clearly see a constant behaviour for the (IRC safe) inclusive cross-section and a logarithmically diverging behaviour for the (IRC unsafe) cross-section after the mMDT procedure.

Moving back to pp collisions, we study how the nature of the observable, IRC safety for $p_{t,\mathrm{jet}}$ and Sudakov safety for $p_{t,\mathrm{mMDT}}$, correlates with the size of non-perturbative corrections due to the hadronisation process and to multiple parton interactions, i.e. the underlying event (UE). To achieve this we use a phenomenological approach based on Monte Carlo parton showers simulations. In order to minimise potential bias due to a particular non-perturbative model, we use a variety of parton showers with different tunes, namely the AUET2 [115] tune of Herwig 6.521 [116, 117], the Z2 [118] and Perugia 2011 [119, 120] tunes of Pythia 6.428 [121], the 4C [82] and the Monash 13 tune [122] of Pythia 8.223 [35].

The results of this study are presented in fig. 9.6, where the plot on the left shows the ungroomed $p_{t,\mathrm{jet}}$ spectrum, while the one on the right the $p_{t,\mathrm{mMDT}}$ distribution. In each plot, we show two sets of curves. The first set (labelled "hadronisation" on the plots) represents, for each Monte Carlo, the ratio between hadron-level and parton-level results, without UE. The second set (labelled "UE") instead shows the ratio of hadron-level results with and without the UE contribution. The $p_{t,\mathrm{jet}}$ plot shows all the features we would expect from an IRC safe observable. Non-perturbative corrections are suppressed by negative powers of the jet transverse momentum. Moreover, since we are dealing with high-p_t jet with a fairly large radius ($R = 0.8$) hadronisation corrections are rather small [123]. The Sudakov-safe $p_{t,\mathrm{mMDT}}$ distribution instead exhibits larger hadronisation corrections, which do not appear to be power suppressed [57]. On the other hand, as perhaps expected in the presence a groomer, we note that $p_{t,\mathrm{mMDT}}$ is less sensitive to

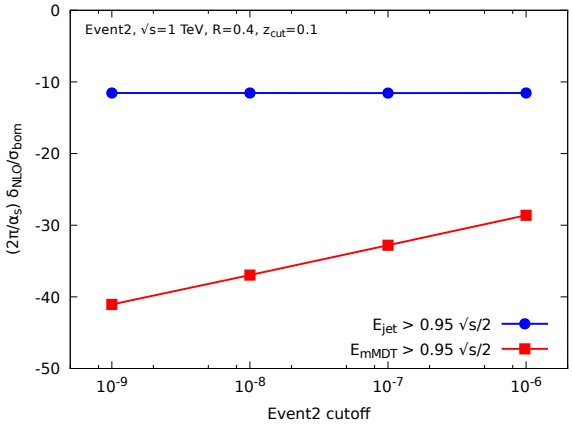


Figure 9.5 — Dependence of the jet cross-section before and after applying mMDT, as a function of the infrared cut-off used in EVENT2 . The cross-section before grooming is stable but the one after grooming diverges logarithmically, thus making the IRC unsafety apparent.

the UE contribution than $p_{t,\mathrm{jet}}$, especially at moderate transverse momentum. We can therefore expect that $p_{t,\mathrm{mMDT}}$ will be more resilient against pile-up (not considered here), which has a structure similar to the UE.

In this study we are primarily interested in jet mass distributions, while we only use the jet cross section for normalisation purposes. Measuring a non-vanishing mMDT mass resolves a two-prong structure within the jet, thus acting as an angular cut-off and regulating the collinear divergence. This means that the unnormalised distribution

$$\frac{d\sigma}{d\rho}(\rho; z_{\mathrm{cut}}, p_{t1}, p_{t2}) = \int_{p_{t1}}^{p_{t2}} dp_{t,\mathrm{mMDT}} \frac{d^2\sigma}{dp_{t,\mathrm{mMDT}} d\rho}, \quad \text{with} \quad \rho = \left(\frac{m}{p_{t,\mathrm{mMDT}} R}\right)^2, \tag{9.20}$$

is IRC safe. However, as we shall see in the following section, the resulting all-order structure is different compared to the one previously described and rather cumbersome. We also note that, because the difference between $p_{t,\mathrm{jet}}$ and $p_{t,\mathrm{mMDT}}$ is $\mathcal{O}(z_{\mathrm{cut}})$, if we choose to use $p_{t,\mathrm{mMDT}}$ we are forced to work at finite z_{cut}.

As a final note, we point out that despite its issues related to IRC safety, $p_{t,\mathrm{mMDT}}$ shows some interesting properties in perturbative QCD. For example, it is directly related to the "energy loss" distribution computed in ref. [17] in the small z_{cut} limit. Modulo small corrections induced by the running of the coupling, the energy loss distribution — i.e. the $p_{t,\mathrm{mMDT}}$ distribution at fixed $p_{t,\mathrm{jet}}$ — is independent of α_s and of the color factor of the parton initiating the jet. We discuss this briefly in the context of the $p_{t,\mathrm{mMDT}}$ jet cross-section in appendix D.3.

9.2.2 Fixed-order structure of the mass distribution

In order to better understand the structure of the mass distribution with $p_{t,\mathrm{mMDT}}$ we analytically calculate eq. (9.20) to LO and NLO, in the collinear limit. We start with

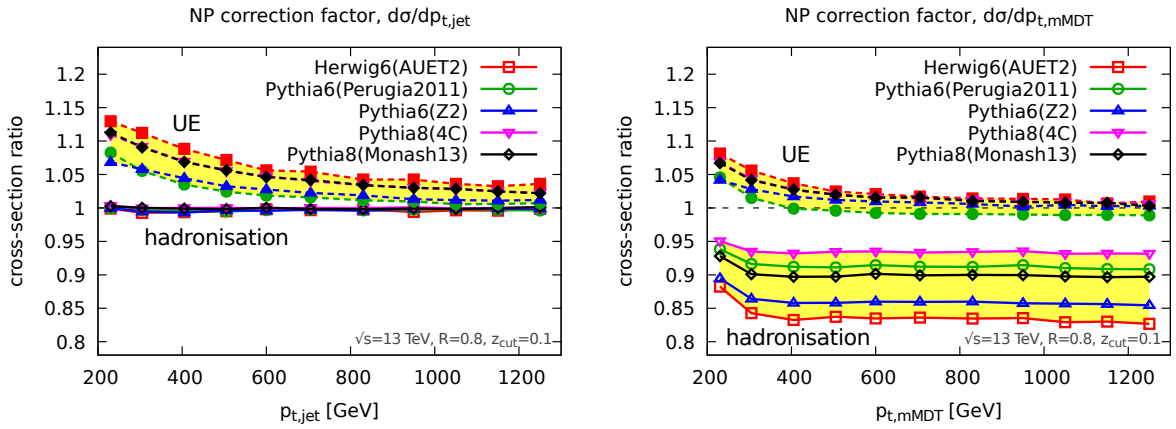


Figure 9.6 — Monte Carlo study of the impact of hadronisation and underlying event (UE) on the ungroomed $p_{t,\mathrm{jet}}$ distribution (left) and on the $p_{t,\mathrm{mMDT}}$ distribution (right).

a jet of momentum $p_{t,\mathrm{jet}}$. At $\mathcal{O}(\alpha_s)$ the jet is made of at most two partons. If one of them is groomed away by mMDT, then the resulting groomed jet is massless. Thus, in order to have a non-vanishing mass, the emission must pass the z_{cut} condition, leading to $p_{t,\mathrm{mMDT}} = p_{t,\mathrm{jet}}$. Therefore, the LL distribution at LO is the same for the two transverse momentum choices and it reads (see also ref. [124])

$$\rho \frac{d\sigma^{\mathrm{LL,LO}}}{d\rho}(\rho; z_{\mathrm{cut}}, p_{t1}, p_{t2}) = \int_{p_{t1}}^{p_{t2}} dp_{t,\mathrm{jet}} \left[\sigma_q(p_{t,\mathrm{jet}}) R'_q + \sigma_g(p_{t,\mathrm{jet}}) R'_g\right]. \tag{9.21}$$

The situation changes when we move to NLO. We consider the sum of the double real emission diagrams and the real-virtual contributions, while the double virtual only gives vanishing masses. At NLO we have di erent color structures. For convenience, we explicitly consider the C_F^2 contribution, which originates from the independent emission of two collinear gluons 1 and 2 o a quark leg. Analogous results can be obtained for the other color structures. Because we are interested in the LL contribution, we can order the two emissions in angle, i.e. $\theta_1 \gg \theta_2, \theta_{12}$. The relevant contributions correspond to the situation where gluon 2 is real (and dominates the mMDT jet mass) and the large-angle gluon 1 is either real and groomed away, or virtual. The only di erence with respect to our calculation in the $p_{t,\mathrm{jet}}$ case is that here we further have to make sure that the measured $p_{t,\mathrm{mMDT}}$ falls in the transverse momentum bin under consideration, say $p_{t1} < p_{t,\mathrm{mMDT}} < p_{t2}$. Assuming for the moment that $p_{t1} < p_{t,\mathrm{jet}} < p_{t2}$, we therefore have the additional constraint on the double-real emission contribution that $p_{t,\mathrm{mMDT}} = (1-z_1)p_{t,\mathrm{jet}}$ still falls in the same transverse momentum bin. We thus have

$$\begin{aligned}\rho \frac{d\sigma^{\mathrm{LL,NLO},C_F^2 a}}{d\rho} = \left(\frac{\alpha_s}{2\pi}\right)^2 \rho \int_{p_{t1}}^{p_{t2}} dp_{t,\mathrm{jet}}\, \sigma_q(p_{t,\mathrm{jet}}) \\ \cdot \int_0^1 \frac{d\theta_1^2}{\theta_1^2} \int_0^1 dz_1\, p_{gq}(z_1) \Big[\Theta\left(z_{\mathrm{cut}} > z_1\right) \Theta\left((1-z_1)p_{t,\mathrm{jet}} > p_{t1}\right) - 1\Big]\end{aligned} \tag{9.22}$$

$$\cdot \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2\, p_{gq}(z_2)\Theta\left(z_2 > z_{\text{cut}}\right)\Theta\left(1-z_2 > z_{\text{cut}}\right)\Theta\left(\theta_1^2 > \theta_2^2\right)\delta\left(\rho - z_2\theta_2^2\right).$$

After some algebra, the distribution in the $\rho < z_{\text{cut}}$ region can be written in terms of the R_i functions previously defined

$$\begin{aligned}\rho\frac{d\sigma^{\text{LL,NLO},C_F^2 a}}{d\rho} &= \int_{p_{t1}}^{p_{t2}} dp_{t,\text{jet}}\,\sigma_q(p_{t,\text{jet}})R_q'\Big[-R_q - R_{q\to g}\Big] \\ &\quad - \int_{p_{t1}}^{\min\left[p_{t2},\frac{p_{t1}}{1-z_{\text{cut}}}\right]} dp_{t,\text{jet}}\,\sigma_q(p_{t,\text{jet}})\; R_q'\,\frac{\alpha_s}{2\pi}\log\frac{1}{\rho}\int_{1-\frac{p_{t1}}{p_{t,\text{jet}}}}^{z_{\text{cut}}} dz_1\, p_{gq}(z_1).\end{aligned} \tag{9.23}$$

We note that the first contribution coincides with the expansion of the resummation formula eq. (9.6) to second order. However, the second term, proportional to $\alpha_s^2\log(1/\rho)$, is a new LL contribution that signals the different all-order structure of the mass distribution with $p_{t,\text{mMDT}}$. Note that we have put a label a in eqs. (9.22) and (9.23) because there is actually a second configuration that contributes, namely when the ungroomed jet has $p_{t,\text{jet}} > p_{t2}$. If the first emission is groomed away, we may end up with $p_{t,\text{mMDT}} < p_{t2}$, so that this contribution has now leaked into the lower bin. For a quark-initiated jet with two gluon emissions. this results into an additional LL piece:

$$\rho\frac{d\sigma^{\text{LL,NLO},C_F^2 b}}{d\rho} = \int_{p_{t2}}^{\frac{p_{t2}}{1-z_{\text{cut}}}} dp_{t,\text{jet}}\,\sigma_q(p_{t,\text{jet}})\; R_q'\,\frac{\alpha_s}{2\pi}\log\frac{1}{\rho}\int_{1-\frac{p_{t2}}{p_{t,\text{jet}}}}^{z_{\text{cut}}} dz_1\, p_{gq}(z_1). \tag{9.24}$$

9.2.3 Resummation

In order to resum the groomed jet mass spectrum in the case of the $p_{t,\text{mMDT}}$ selection we have to generalise the calculation described in the previous section to all orders. Clearly, the situation is much more complicated than the $p_{t,\text{jet}}$ case chiefly because the value of $p_{t,\text{mMDT}}$ is determined by all the emissions that fail the mMDT condition and therefore our calculation must keep track of them. Because of this complication we are not able to find simple analytic expressions that capture the all-order behaviour, nevertheless we can achieve LL accuracy in the groomed mass distribution using an approach based on generating functionals [125, 126] and, in particular, the application of this formalism to the description of the angular evolution of jets with small radius [127, 128].

We start by defining an evolution variable which is closely related to the angular scale θ at which we resolve a jet

$$t = \int_{\theta^2}^1 \frac{d\theta'^2}{\theta'^2}\frac{\alpha_s(\theta' p_{t,\text{jet}}R)}{2\pi} = \frac{1}{2\pi\beta_0}\log\left(\frac{1}{1+2\alpha_s\beta_0\log(\theta)}\right) = \frac{\alpha_s}{2\pi}\log\left(\frac{1}{\theta^2}\right) + \mathcal{O}(\alpha_s^2), \tag{9.25}$$

with, as before, $\alpha_s = \alpha_s(p_t R)$. This definition of t includes leading collinear logarithms induced by the running of the QCD coupling when going to small angles. When mMDT (and more generically SoftDrop) recurses to smaller and smaller angular scales, the corresponding value of evolution variable t increases until it reaches a non-perturbative value

$t_{\max}$. Thus, by considering successive $1 \to 2$ angular-ordered splittings, we can write down LL evolution equations for a generating functional associated to a quark, $Q(x,t)$, or to a gluon $G(x,t)$, where x is the momentum fraction. The relevant evolution equations were derived in ref. [127]. The only difference here is that after each splitting we follow the branch with the highest transverse momentum, as it is appropriate for the mMDT algorithm. We obtain

$$\frac{d}{dt}Q(x,t) = \int_0^1 dz\, p_{gq}(z) \left[Q\left((1-z)x\right) \Theta\left(z < \frac{1}{2}\right) + G\left(zx\right) \Theta\left(z > \frac{1}{2}\right) - Q(x,t) \right], \tag{9.26a}$$

$$\frac{d}{dt}G(x,t) = \int_0^1 dz \left[\frac{1}{2} p_{gg}(z) G\Big(\max(z,1-z)x, t\Big) + p_{qg}(z) Q\Big(\max(z,1-z)x, t\Big) - p_{xg}(z) G(x,t) \right]. \tag{9.26b}$$

These equations can be implemented numerically under the form of a Monte Carlo generator producing angular-ordered (from large angles to small ones) parton branchings. Compared to the implementation used in [127], the only difference is that the successive branchings follow the hardest of the two partons obtained at the previous step of the showering. We record the angle θ and momentum fraction z of all the emissions.

In order to obtain the mMDT mass spectrum, two extra ingredients are needed: firstly, we need to impose the mMDT condition and, secondly, we should impose an ordering in invariant mass rather than an ordering in angle. Since mMDT proceeds by declustering a C/A tree, imposing the mMDT condition on our angular-ordered events is trivial: we simply search for the first emission that satisfies $z_{\mathrm{cut}} < z < 1 - z_{\mathrm{cut}}$. From the momentum fractions of all the previous emissions, i.e. those at larger angles, we can then reconstruct the momentum fraction groomed away by the mMDT procedure and thus $p_{t,\mathrm{mMDT}}$. Then, once we have reached an emission that passes the mMDT condition, we investigate all the emissions to find the one that dominates the mass. If these emissions have angles θ_i, obtained by inverting eq. (9.25), and momentum fractions z_i, we take, to LL accuracy, $\rho = \max_i[\min(z_i, 1-z_i)\theta_i^2]$. In particular, it is worth pointing out that we can use the momentum fraction z_i, relative to each branching, instead of the actual momentum of each parton with respect to the initial jet. This is simply because the difference between the two does not generate any logarithmic enhancement.[5] Finally, since the resummation is obtained from a Mote Carlo event generator, it can directly be interfaced with NLOJet++ at Born-level to obtain predictions for the jet mass cross-section.

Before we present matched results, we note that, compared to the resummation done in the previous section for the $p_{t,\mathrm{jet}}$ case, the use of eq. (9.25) implies that we are neglecting

[5] Similarly, we can wonder why, once we have an emission satisfying the mMDT condition and the de-clustering procedure stops, we keep generating branchings only on the hardest branch. This is simply because further branchings on a soft branch would never dominate the jet mass and can therefore be neglected. This would not be valid for observables sensitive to secondary emissions, like N-subjettiness with $N > 1$, for which all branchings should be included at angles smaller than the first branching which passes the mMDT condition.

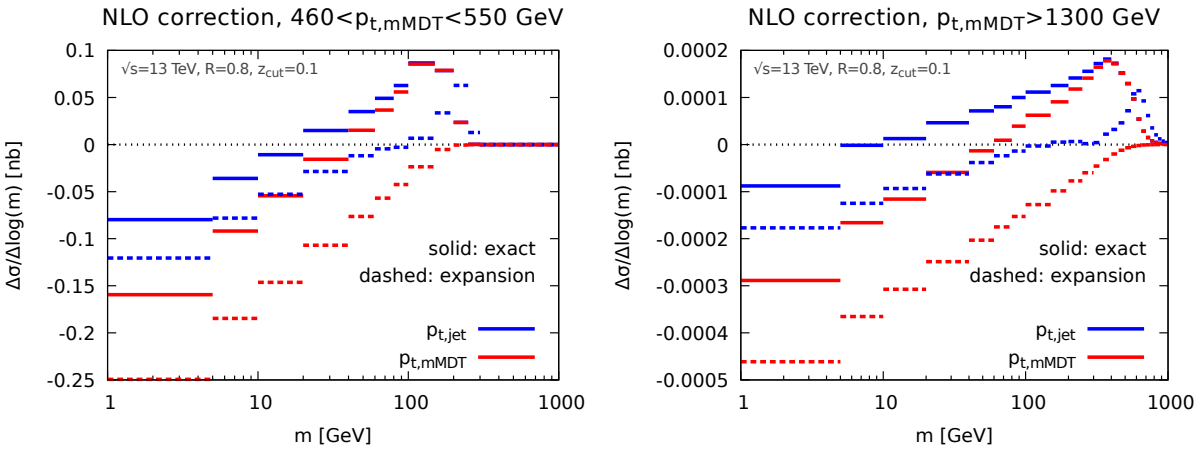


Figure 9.7 — Comparison between the full NLO correction (solid) to the jet mass distribution to the $\mathcal{O}(\alpha_s^2)$ expansion of the LL resummation (dashed) for both $p_{t,\mathrm{jet}}$ (blue) and $p_{t,\mathrm{mMDT}}$ (red) in two different transverse momentum bins.

a factor z in our choice of the scale of the running coupling. This means that we are not including running-coupling effects in the double-logarithmic small-z_{cut} contributions. This approximation can be explicitly studied in the context of a selection on $p_{t,\mathrm{mMDT}}$ and we show in appendix D.1.2 that this only have a modest impact on the final results.

9.2.4 Matching and perturbative results

As for the case of the ungroomed $p_{t,\mathrm{jet}}$, an accurate description valid both in the $\rho \ll 1$ region and in the $\rho \sim 1$ region requires the matching of our LL resummation to a fixed-order calculation. As before, the latter is obtained using NLOJet++. We note that at LO, the results are identical to the ones obtained in the $p_{t,\mathrm{jet}}$ case in section 9.1.3.

In order to match fixed-order and resummed calculations we have to work out the expansion of the resummed cross-sections to LO and NLO. This can be obtained by expanding eq. (9.26) to first and second order in α_s. In practice, we have found more convenient to reuse here the same code as in ref. [127], with minor modifications to include additional information about the successive branching angles and momentum fractions as well as simplifications related to the fact that we do not have to include splittings in the soft branch. For fixed p_t, we have checked our numerical results against an explicit analytic calculation. Note that at NLO, i.e. at $\mathcal{O}(\alpha_s^2)$, we should include both a contribution coming from two emissions (see also the earlier discussion in section 9.2.2) as well as a running-coupling correction coming from the expansion of eq. (9.25) to $\mathcal{O}(\alpha_s^2)$.

We compare the expansion of the LL resummation to $\mathcal{O}(\alpha_s^2)$ against the exact NLOJet++ NLO correction in fig. 9.7, for both $p_{t,\mathrm{jet}}$ (blue) and $p_{t,\mathrm{mMDT}}$ (red) and for two different transverse momentum bins. We first note that at small mass the expansion of the resummed distribution has the same slope of the corresponding fixed-order, meaning that we do indeed control the $\mathcal{O}(\alpha_s^2 \log(1/\rho))$ contribution, as expected from our LL cal-

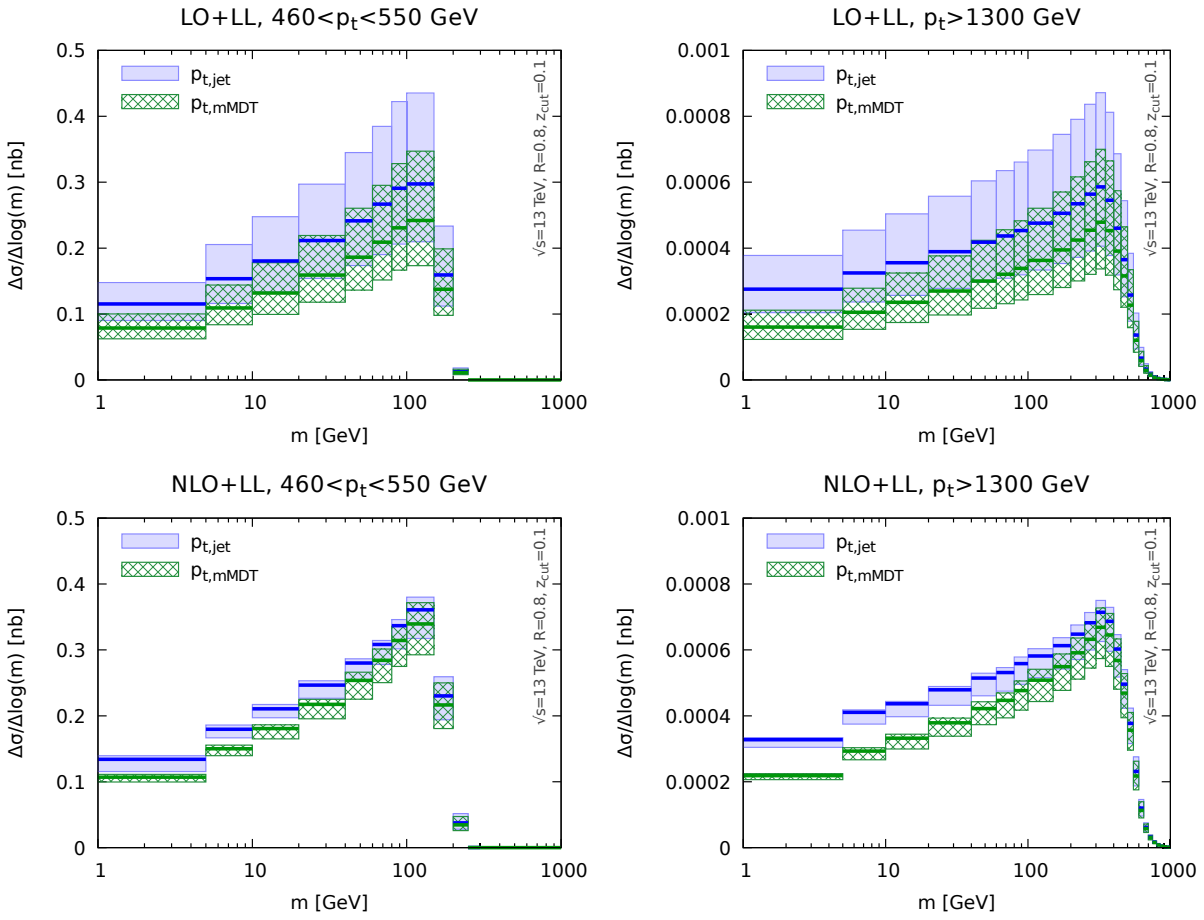


Figure 9.8 — In this figure we compare resummed and matched jet mass distributions in the case of ungroomed $p_{t,\mathrm{jet}}$ selection (blue) or groomed $p_{t,\mathrm{mMDT}}$ selection (green).

culation. More interestingly, fig. 9.7 shows explicitly that the mass distribution obtained in the $p_{t,\mathrm{mMDT}}$ and $p_{t,\mathrm{jet}}$ cases differ already at the LL accuracy. This is manifest from the fact that the $p_{t,\mathrm{mMDT}}$ and $p_{t,\mathrm{jet}}$ curves have different slopes at small mass. The difference in slope is captured by our analytic calculation and is due to the effects already discussed in section 9.2.2.

We are now ready to discuss the matching itself. We adopt the multiplicative matching scheme introduced in eq. (9.16). Our results are shown in fig. 9.8 for the (unnormalised) jet-mass cross-section. The hatched (green) curves are the results obtained for the $p_{t,\mathrm{mMDT}}$ case and they are compared to the results already obtained in section 9.1.3 shown in shaded blue. The plots on the top are for LO+LL, while the ones at the bottom for NLO+LL. We pick the same representative bins in transverse momentum as before, namely $460 < p_t < 550$ GeV and $p_t > 1300$ GeV, with p_t being either $p_{t,\mathrm{mMDT}}$ or $p_{t,\mathrm{jet}}$. The cross-sections are significantly smaller for the $p_{t,\mathrm{mMDT}}$ case than for the $p_{t,\mathrm{jet}}$ case, mostly because the overall jet cross-section is smaller. This is related to the loss of transverse momentum due to the mMDT procedure, which is also discussed in appendix D.3. We also see, in particular on

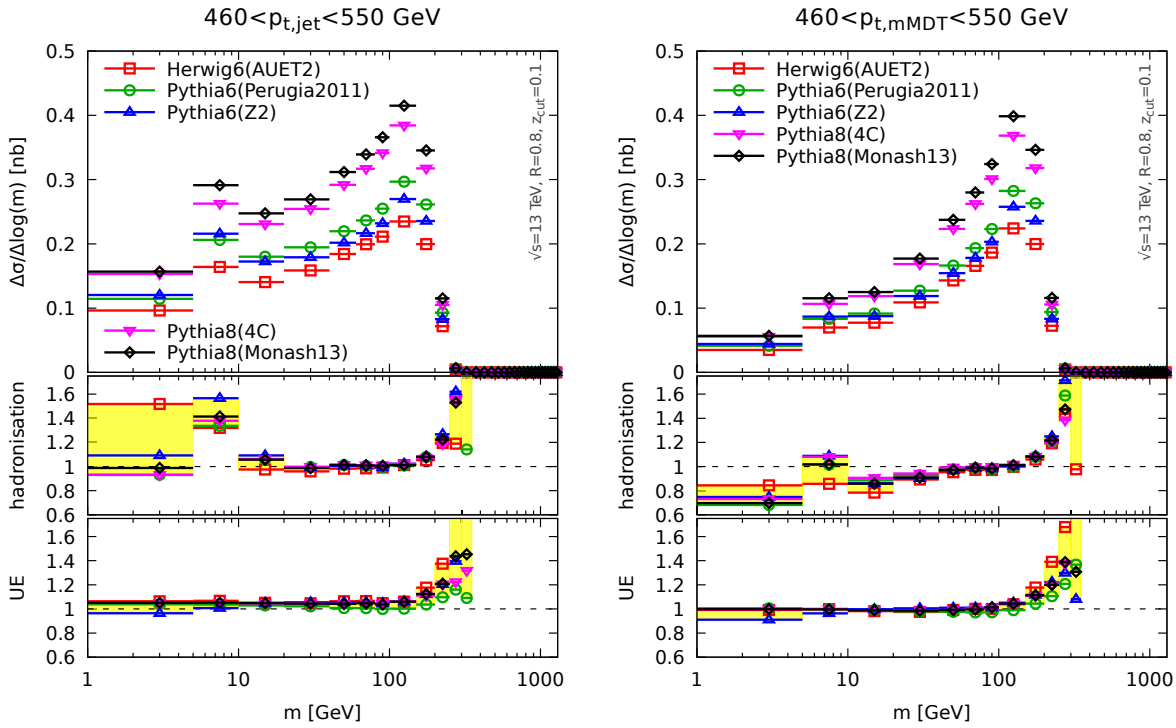


Figure 9.9 — The top plots show the groomed jet mass distribution for $460 < p_t < 550$ GeV, with hadronisation and the underlying event, for different Monte Carlo parton showers. The plot on the left is for the ungroomed $p_{t,\text{jet}}$, while the one of the right for $p_{t,\text{mMDT}}$. The bottom plots show the ratios hadron-level to parton-level and with-to-without the underlying event.

the NLO+LL results for the high-p_t bin, that the $p_{t,\text{mMDT}}$ distributions decrease slightly faster than the $p_{t,\text{jet}}$ ones at small mass. This feature was already observed in fig. 9.7.

We note that due to the IRC unsafety of the $p_{t,\text{mMDT}}$ jet cross-section, the normalisation of the fixed-order jet mass distribution is ill-defined. The resummed and matched cross-sections could simply be normalised to unity but we found that this procedure tends to clearly underestimate the size of the perturbative uncertainty and is potentially dangerous as it relies on the computation of the resummed cross-section down to very small masses where non-perturbative effects are dominant. We have therefore decided to present only predictions for the unnormalised distributions.

9.3 Non-perturbative corrections

In this section we perform a Monte Carlo study of non-perturbative contributions considering effects coming from the hadronisation process as well as from the underlying event. In order to study non-perturbative corrections to the jet mass distribution we consider the same set of Monte Carlo tunes used for studying the p_t spectra in section 9.2. As usual, we consider two representative transverse momentum bins. In fig. 9.9 we consider

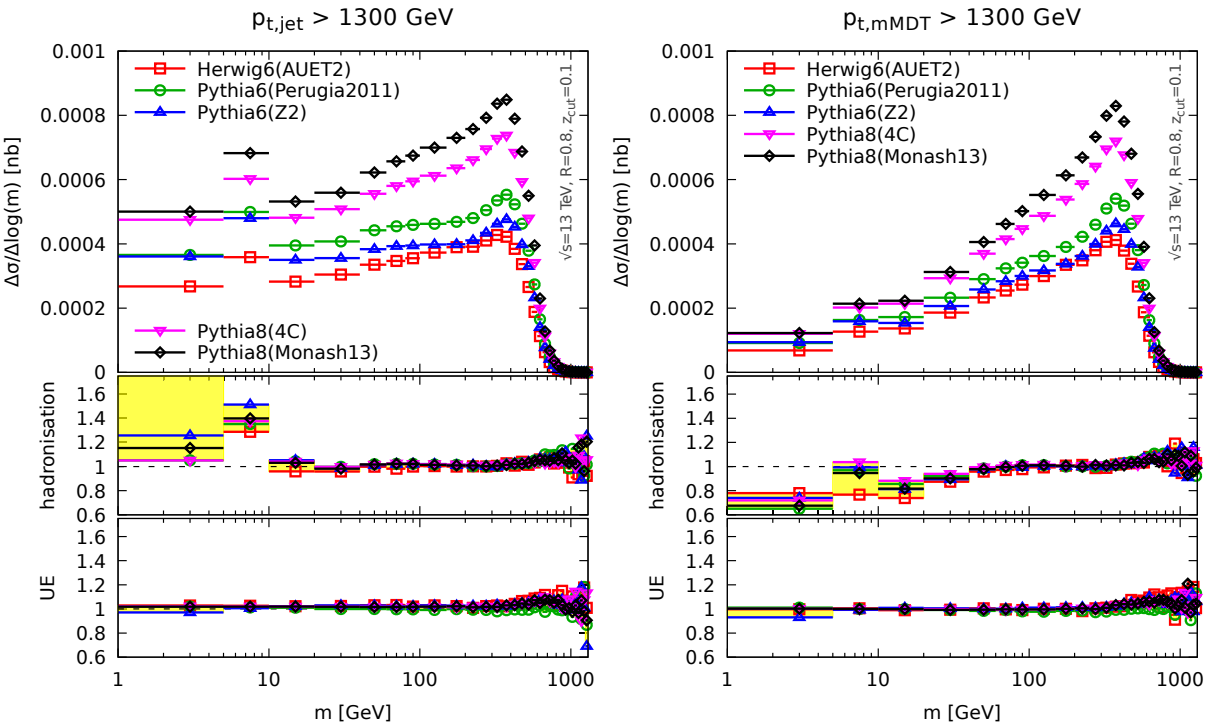


Figure 9.10 — Same as fig. 9.9 but for the bin $p_t > 1300$ GeV.

$460 < p_t < 550$ GeV, while in fig. 9.10 we consider $p_t > 1300$ GeV. In both cases, the plots on the left refer to the ungroomed $p_{t,\mathrm{jet}}$ selection, while the ones on the right refer to the $p_{t,\mathrm{mMDT}}$ case.

In the top plots we show the (unnormalised) jet mass distributions as obtained from each Monte Carlo program. The striking feature is the huge discrepancy between these results, even at large masses. In particular, the predictions obtained with the most recent Pythia 8 tunes appear to be a factor of 2 larger than the other tunes in the region of interest for this study. This performance of standard parton shower tools, worrisome at first glance, should be put in parallel with our LO+LL results (see e.g. fig. 9.8) which exhibit a similar uncertainty band. This indicates the need to match the parton shower with NLO fixed-order matrix elements.

In the bottom plots of figs. 9.9 and 9.10 we instead show, for each Monte Carlo, the ratio of hadron-to-parton level results (labelled "hadronisation") and the ratio with-to-without the underlying event contribution (labelled "UE"). We first note that in both the $p_{t,\mathrm{jet}}$ and $p_{t,\mathrm{mMDT}}$ selection cases, the groomed mass distribution has very small sensitivity to the underlying event, as we expect from mMDT being an (aggressive) groomer. This contribution becomes more sizeable at large masses essentially because the effective jet radius becomes larger. Moreover, this effect is more visible in the moderate p_t bin since the power-suppression in the hard scale of the process becomes weaker. Hadronisation corrections have instead a different shape for the $p_{t,\mathrm{jet}}$ and $p_{t,\mathrm{mMDT}}$ selections, most likely stemming from the different properties of the underlying transverse momentum distribu-

tion. For the $p_{t,\text{jet}}$ case, hadronisation corrections are sizeable in the low mass bins, with a peculiar peak in the 5-10 GeV bin, and at very large masses, close to the end-point region. For both small and large masses, this also comes with a larger spread of the hadronisation corrections across the generators and tunes. However, there exists a rather large region in mass, increasing in size as $p_{t,\text{jet}}$ grows, where these contributions are genuinely small. Hadronisation corrections appear rather different in the $p_{t,\text{mMDT}}$ selection case. They come with opposite sign at small masses and appear to be non-negligible in a wider region of the mass distribution, similarly to what was already noticed in section 9.2.1 for the jet cross-section.

Given the large kinematic range over which the non-perturbative corrections appear to be small, upcoming LHC data could bring valuable constraints on the perturbative aspects of parton showers. Additionally, the behaviour at low mass, with very little sensitivity to the underlying event, could help constraining hadronisation models. For example, measurements on both quark and gluon-enriched jet samples would be complementary to the quark-dominated LEP data currently used to tune hadronisation models [91, 129].

In practice, for this study, we use the above Monte Carlo results to estimate the size and the uncertainty of non-perturbative corrections on the groomed mass distribution. For each Monte Carlo generator and tune we construct the ratio particle-level, i.e. hadronisation with UE, to parton-level, in each mass and transverse momentum bin. We take the average value of this ratio as a correction factor to apply to the perturbative NLO+LL results obtained in the previous sections. We take the envelope of the corrections across different generators and tunes as an estimate of the non-perturbative uncertainty, which we add in quadrature to the perturbative uncertainty. We consider this solution an acceptable and rather conservative estimate of non-perturbative contributions.

9.4 Final results

We can now present our final results for the groomed jet mass distribution for both the $p_{t,\text{jet}}$ and $p_{t,\text{mMDT}}$ selection. Our perturbative results, which are accurate to NLO+LL, are multiplied by a bin-by-bin (in both mass and transverse momentum) non-perturbative correction factor obtained from Monte Carlo parton showers. The total uncertainty is taken as the sum in quadrature of the perturbative and non-perturbative uncertainties. The former is obtained by varying renormalisation, factorisation, and resummation scales as described in section 9.1 and taking the envelope of the result; the latter by considering the envelope of the five different Monte Carlo generators and tunes.

Fig. 9.11 and fig. 9.12 show the results (in black, with grey uncertainty bands) for the ungroomed $p_{t,\text{jet}}$ selection in the two representative transverse momentum bins: $460 < p_{t,\text{jet}} < 550$ GeV and $p_{t,\text{jet}} > 1300$ GeV. The former is the jet mass distribution, while the latter is normalised to the NLO jet cross-section in the appropriate transverse momentum bin. Similarly, in fig. 9.13 we show our final results for the $p_{t,\text{mMDT}}$ selection. As discussed in the paper, the NLO jet cross section is not well-defined in this case, so we only present unnormalised distributions. For comparison, we also show in red the purely perturbative NLO+LL results with their uncertainties. As previously noted, non-perturbative correc-

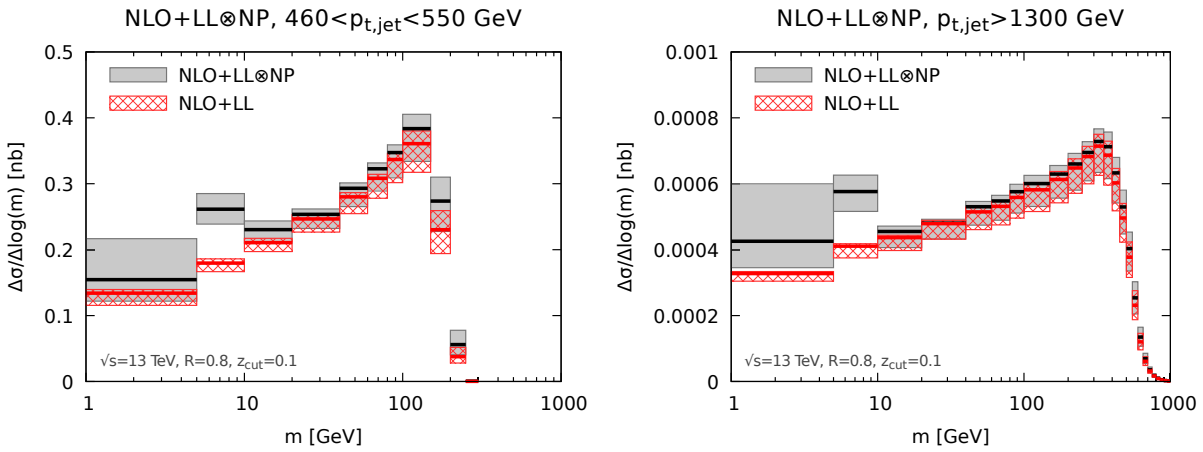


Figure 9.11 — Final results for the jet mass distribution in the case of the ungroomed $p_{t,\text{jet}}$ selection. The perturbative calculation is performed at NLO+LL and non-perturbative corrections are included as a multiplicative factor obtained from Monte Carlo parton showers. Perturbative uncertainties are obtained varying renormalisation, factorisation and resummation scales as detailed in section 9.1. Non-perturbative uncertainties are obtained considering the spread of five different Monte Carlo tunes, as detailed in section 9.3. Perturbative and non-perturbative uncertainties are added in quadrature.

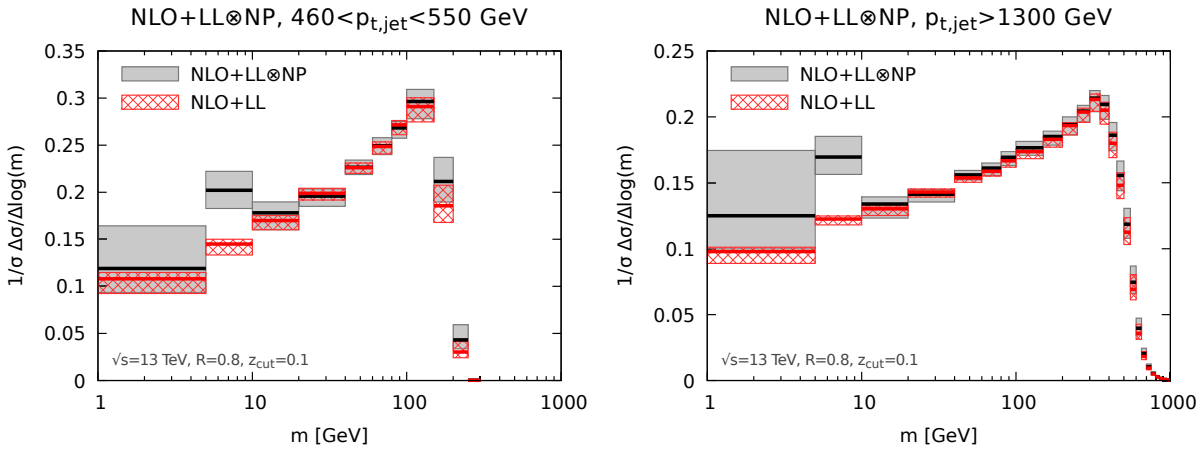


Figure 9.12 — Final results at NLO+LL, with non-perturbative corrections, for the normalised jet mass distribution, in the case of the ungroomed $p_{t,\text{jet}}$ selection.

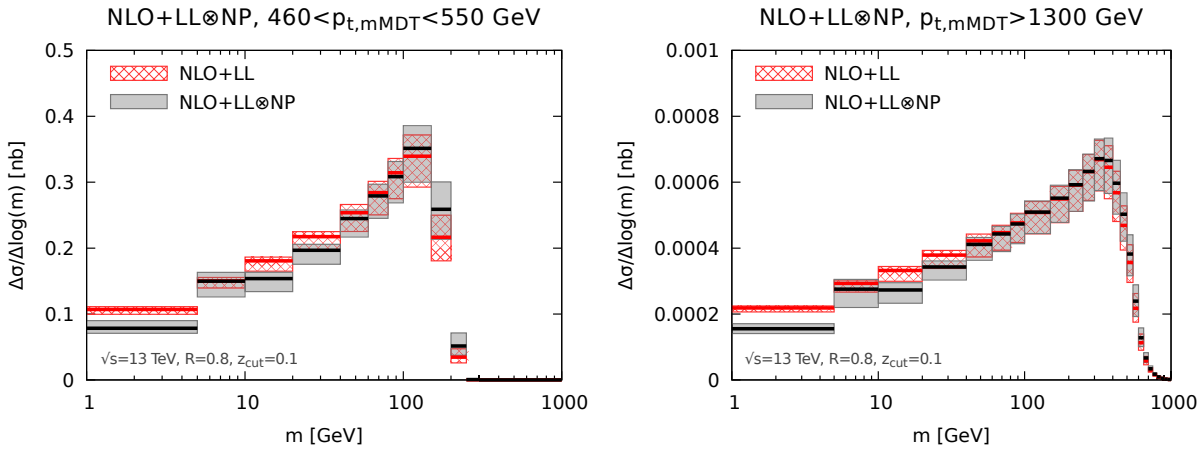


Figure 9.13 — Final results at NLO+LL, with non-perturbative corrections, for the $p_{t,\mathrm{mMDT}}$ selection.

tions are sizeable (with large uncertainties) in the first few mass bins ($m \lesssim 10$ GeV) and at very large masses, close to the end-point region. Nevertheless, there exists a region in mass, which increases in size as $p_{t,\mathrm{jet}}$ grows, where non-perturbative effects are genuinely small and a meaningful comparisons between experiments and perturbation theory can be performed. However, we have found that, when we consider normalised distributions in fig. 9.12, the uncertainty related to these non-perturbative contributions is, at best, of the same order as the NLO+LL perturbative calculation.

The above results clearly demonstrate the value of jet substructure algorithms to perform phenomenological studies in QCD. In particular, the region in mass where non-perturbative contributions are genuinely small offers an opportunity to test the modeling of perturbative radiation in analytic resummations and parton showers. In that respect, one could even consider the possibility to use experimental data in this mass region for a novel measurement of the strong coupling. On the other hand, the lower mass bins, which are sensitive to hadronisation but have small UE contaminations, can be used to test (and tune) the hadronisation models of Monte Carlo event generators.

Finally, we have found that the p_t selection is better suited for theoretical calculations and the resulting resummation has a relatively simple form that can be, in principle, extended to higher-logarithmic accuracy. Moreover, for the typical choice $z_{\mathrm{cut}} = 0.1$, finite z_{cut} corrections, although formally entering already at LL accuracy, appear to be very small. This justifies the small-z_{cut} approximation used in refs. [74, 75] to achieve higher logarithmic accuracy. However, the finite z_{cut} corrections would inevitably increase for larger values of z_{cut}. Also, it would be interesting to achieve a complete picture at NLL accuracy, including the finite z_{cut} corrections, even though our findings in this paper suggest that the latter would be small. We have also found that logarithms of z_{cut} give a non-negligible contribution, thus indicating the necessity of their resummation. We have also studied the perturbative uncertainty of our calculation, observing that matching to NLO greatly reduces the theoretical uncertainty especially in the case of unnormalised

distributions. Non-perturbative effects are reduced compared to the ungroomed jet mass and only remain sizeable at low mass, where hadronisation dominates, or at very large masses, close to the end-point of the distribution.

The $p_{t,\mathrm{mMDT}}$ selection has instead more theoretical issues but it can also present some advantages from a phenomenological viewpoint. The main theoretical complication stems from the fact that the $p_{t,\mathrm{mMDT}}$ jet spectrum is not IRC safe, but only Sudakov safe. The jet mass distribution is itself safe, with the mass acting as a regulator for collinear emissions, but the inclusive $p_{t,\mathrm{mMDT}}$ cross-section is only Sudakov safe. Due to the complicated flavor structure of the all-order resummation, we were only able to arrive at a numerical resummation of the LL contributions. A possible extension of our results to a higher logarithmic accuracy is therefore expected to be difficult, even in principle. From a phenomenological viewpoint, it would be interesting to see whether the slightly smaller sensitivity to the underlying event of the $p_{t,\mathrm{mMDT}}$ choice implies a smaller sensitivity to pileup. More generally, an understanding of hadronisation corrections for Sudakov-safe observables would also be interesting.

To summarise, in this work we have derived theoretical predictions for the invariant mass distribution of jets groomed with mMDT, including a study of the perturbative and non-perturbative theoretical uncertainties. The situation where distributions are computed in bins of the initial (ungroomed) jet p_t exhibit a simpler analytic structure, compared to the case where the binning is done using the groomed jet p_t. This means that the former is more likely to be amenable to a theoretical calculation with higher logarithmic accuracy.

— 10 —
Conclusion

As underlined in the introduction of this thesis, now that the Run II of the LHC is achieving energies well above the electroweak scale, jet substructure tools play an important role in particle physics. This importance will only grow in the next runs and, eventually, with the construction of future accelerators with even higher energy scales. More generally, jets are ubiquitous in collider phenomenology and as they are complex structures, substructure tools are helpful to understand their internal dynamics, specially at boosted regimes.

There is an active community developing and researching jet substructure tools, both on the theoretical and experimental aspects. Numerical tools play a major role in this domain, notably parton shower Monte Carlo generators as Pythia, Herwig and Sherpa. They can be used to simulate jet substructure observables when combined with the appropriated softwares like FastJet, which does the clustering of particles into jets and also with `fjcontrib`, which contains a multitude of extensions for different jet observables. Nevertheless, in this thesis we argued that an analytical approach can, in conjunction with these tools, bring valuable insights to the field. Although analytic calculations do not completely replace Monte Carlo simulations, the analytical approach is capable of understanding the source of the differences between methods. Also, it allow for precise calculations with uncertainties that can be compared to experimental measures.

In this thesis, we focused on methods for two-pronged jet tagging, i.e. methods developed to identify jets originating from heavy bosons like the Higgs, W and Z. We used the "traditional" perturbative QCD approach, which explores factorizations and exponentiation properties of QCD matrices to achieve all-orders resummation in α_s, which is necessary to have a good predictive power in boosted regimes. Other approaches to jet substructure are used in literature, notably the Soft Collinear Effective Theory which uses factorization theorems to achieve all-orders resummation. Some examples of applications if this method can be found in refs. [74, 75, 130, 131]. Other approach that gained attention recently is the use of deep learning techniques to identify jets.

The first chapters of this thesis are introductory, essentially meant to define relevant jet observables and lay down notations. In chapter 2, we recalled the basis of QCD, discussing the concepts of infra-red and collinear safety and Sudakov safety, that are relevant in the context of jet substructure. In particular, the Sudakov safety property was first observed in the study of ratio angularities [56], a jet substructure observable.

Then in chapter 3, we presented the formal definition of jets and some jet algorithms that are relevant to this work. It is important to stress that jets are known for a long time in particle physics, before we achieved the energy scales necessary to have a boosted regimes. The first jet algorithms were proposed in the context of e^+e^- collisions in the 70's. In chapter 4, we presented the particularities of jets in boosted regimes. We also presented some jet substructure tools (prong finders, groomers and jet shapes), focusing on the ones that are relevant for the rest of this thesis. Finally, in chapter 5, we motivated the need for resummed calculations using a simple example, namely the jet mass distribution. We presented the Lund diagrams, a graphic tool that we used extensively to illustrate physical discussions. We also introduced the basic building blocks, which allowed us to expresses our results in a compact form, without the need to repeat the same lengthy calculation multiple times.

The next chapters of this thesis presented original results for calculations of jet substructure observables. The first analytic results in this field have been obtained by ref. [12] for the jet mass with different groomer: trimming, pruning and MassDrop Tagger. In this work we extended this understanding in two directions, first for the combination of prong finder with groomers in chapter 6, then for jet shapes in chapter 7.

In chapter 6, we studied the Y-splitter tagger combined with three different grooming techniques: the modified MassDrop Tagger (mMDT), trimming and SoftDrop. This work followed previous Monte Carlo studies which showed that such combinations increase the tagger performance. We studied the origin of this behavior and also introduced improved variants of the original Y-splitter method. We used Monte Carlo generators to study the impact of non-perturbative effects and, comparing the different variants, we observed a trade-off between robustness against NP effects and the tagging performance. In terms of sheer performance the $\mathrm{Y_m}$-splitter+trimming or $\mathrm{Y_m}$-splitter+mMDT combinations with a standard y_{cut} should be preferred. If instead we want maximum robustness, e.g. to reduce uncertainties, $\mathrm{Y_m}$-splitter+mMDT with a z_{cut} condition or SoftDrop pre-grooming (with either a y_{cut} or a z_{cut} condition) appear at the same time both efficient and robust.

In chapter 7, we studied three jet shapes, N-subjettiness, energy correlation functions and MassDrop parameter, used as discriminating variables between two-pronged hadronic decays of electroweak bosons and QCD jets from the background. For each case, we computed the jet mass distribution with an additional cut on the jet shape variable. This study showed that the difference in performance between methods comes mostly from the large angle emissions, with energy correlation functions being the most efficient. We used Monte Carlo generators to study non-perturbative effects and validate our results. We also added a SoftDrop grooming to the jet and observed how it affects the tagging efficiencies. As this groomer removes mostly at the soft and large angle emissions, it diminishes the differences in discriminating power between methods, but the ordering remains the same.

Given the understanding of the individual ingredients (prong finders, groomers and jet shapes) we studied the interplay between these tools in chapter 8, using mMDT and SoftDrop combined with N-subjettiness. We proposed the dichroic N-subjettiness ratio, where we use a large jet (with or without a SoftDrop pre-grooming step) for calculating τ_2 and a smaller, tagged subjet for τ_1. We observed an enhanced performance compared to

the variants currently used in experimental analyses, while keeping non-perturbative effects under control. Comparing the versions with and without pre-grooming, we observed again a trade-off between performance and robustness against non-perturbative effects.

Finally, in chapter 9, we performed a phenomenological study of the jet mass distribution after applying the mMDT. Our goal it to compare our results with current LHC predictions. We accounted for the resummation of the leading logarithm of the ratio of the jet mass over the jet transverse momentum and matched it to fixed-order matrix elements computed at next-to-leading order. We discussed two options according to whether the distributions are measured in bins of the jet transverse momentum before ($p_{t,\text{jet}}$) or after the mMDT ($p_{t,\text{mMDT}}$). We call attention to the fact that the distribution differential in the transverse momentum after grooming is not collinear safe, only Sudakov safe. This generates many theoretical complications, which make a higher order computation extremely complicated, so we encourage the use of the $p_{t,\text{jet}}$ variant. Importantly, since our original work in chapter 9, the CMS results are now available [47]. It turns out that our predictions agree nicely with the data. In particular, using NLO matrix elements shows a substantial improvement at large jet mass.

There are many different directions to be explored in future works, the first one being expand our resummed calculations to higher orders. This is feasible for situations where the non-pertubative effects are relatively under control and non-global logarithms are eliminated thought grooming (as they are very difficult to control analytically). One example of observable that fulfills these properties and that we intend to compute at NLL is the groomed jet mass distribution, using the $p_{t,\text{jet}}$ variation as mentioned above. We are currently doing this calculation for the SoftDrop groomer with a general β_{SD} parameter at NLL accuracy [132]. There are results available in literature at NNLL+LO accuracy, obtained with SCET approach [74, 75], but our results take into account finite z_{cut} effects and are matched to NLO matrix elements.

There is also room for improvement in our calculations of jet shapes, in particular N-subjettiness and energy correlation functions, which are shown to be efficient to discriminate jets and are largely used in experiments. We intend to do calculation for finite jet shape cuts $v_{\text{cut}} \lesssim 1$. The calculations are more complicated, but it is important if one intends to do a phenomenological study, as the typical cuts chosen in experiments are usually not so small. Other possibility is to do calculation for jet shapes with the angular exponent $\beta = 1$, this case does not have many of the simplifications of the $\beta = 2$ case, but is the default choice in many experimental measurements.

Another application of analytic calculations we intend to explore is the optimization of parameters. Composite jet substructure methods (e.g. jet shape + grooming) usually involve a long list of parameters – for example, a cut in the jet shape v_{cut}, a cut from the groomer y_{cut} or z_{cut}, possibly angular parameters β, the jet radius R. Optimizing such a large array of parameters with numerical tools would be too complicated, but it is feasible if one has analytic expressions for observables efficiencies. In future studies, we would like to apply this procedure to some of the methods studied in the thesis (in particular the ones more resilient to NP effects, e.g. dichroic jet shapes with pre-grooming and Y_{m}-splitter+mMDT/trimming variations), to verify if it can bring improvements to

existing methods.

We also intend to apply the insight acquired from the analytic calculations, specially for combinations of taggers and dichroic radiation constraints, for the future design of "decorrelated" taggers [133]. These would provide a background rejection that is independent of the tagged jet mass and thus straightforward to use in the context of data-driven background estimates.

To summarize, the analytical approach to jet substructure can bring a better understanding of techniques that are already largely used in experiments. We can use this information to improve them or develop new techniques. Combining existing techniques can also improve considerably their performance in a nontrivial manner. Additionally, precise calculations matched with fixed order matrix elements and with theoretical error bands can be compared with experimental results. This domain has a very active community, both theoretical and experimental, and its importance will only grow in the following years, as part of a bigger effort to answer open questions in particle physics.

— A —
Further details on Y-splitter

A.1 Why not use the groomed mass in the Y-splitter condition?

We have argued in section 6.2 that we should first impose the Y-splitter condition on the plain jet and, if the condition is satisfied, measure the groomed jet mass. One might be tempted to also use the groomed jet mass in the definition of the Y-splitter condition. We show in the following that this does not lead to an efficient tagger.

For simplicity, let us use the modified MassDrop Tagger (trimming would yield similar results, albeit a bit more complex and involving additional transition points) and assume that emission 1 dominates the groomed mass. We still have two ways to proceed: we can either decluster the groomed jet or the plain jet to get the k_t scale entering the $\mathrm{Y_m}$-splitter condition. The situation where we use the groomed jet is almost trivial: the declustering will either select emission 1 or an emission, say 2, at smaller mass and larger k_t. In both cases, the resulting Y-splitter condition is trivially satisfied, since, e.g. in the second case, $k_{t2}^2 > k_{t1}^2 = x_1\rho > y\rho$. Hence, neither the grooming procedure nor the Y-splitter condition place any constraint on radiation at larger mass in the groomed-away region, meaning that we would get

$$\rho\frac{d\sigma}{d\rho} = \int_{\frac{y}{1+y}}^{\frac{1}{1+y}} dx_1\, p_i(x_1)\, \frac{\alpha_s(x_1\rho)}{2\pi} e^{-R_{\mathrm{mMDT}}(\rho)}. \tag{A.1}$$

This has to be compared to eq. (6.14) for the situation(s), considered in the main text, where we use the plain jet mass in the $\mathrm{Y_m}$-splitter condition. The result in (A.1) is significantly less efficient since it comes with a much weaker Sudakov suppression.

Let us assume instead that we decluster the plain jet in order to define the Y-splitter k_t scale. In the groomed-away region, emission with k_t smaller than k_{t1} will be unconstrained. Emission with k_t larger than k_{t1} will also be allowed since the resulting Y-splitter condition $k_{t2}^2 > \rho y$ is always met due to $k_{t2}^2 > k_{t1}^2 > \rho y$. We would therefore again recover (A.1).

Finally, let us briefly discuss the case of $\mathrm{Y_m}$-splitter, with mass declustering applied to the plain jet. This is slightly different because now there could be an emission, say

emission 2, in the groomed-away region, with a mass larger than ρ and a k_t smaller than k_{t1}. In that case the $\mathrm{Y_m}$-splitter condition would impose $k_{t2}^2 > \rho y$, yielding an additional suppression compared to (A.1)

$$\rho\frac{d\sigma}{d\rho} = \int_{\frac{y}{1+y}}^{\frac{1}{1+y}} dx_1\, p_i(x_1)\, \frac{\alpha_s(x_1\rho)}{2\pi} e^{-R_{\mathrm{mMDT}}(\rho)-R_{\mathrm{out,low}-k_t}(\rho)}, \tag{A.2}$$

with

$$R_{\mathrm{out,low}-k_t}(\rho) = \int \frac{d\theta^2}{\theta^2} dx\, p_i(x)\, \frac{\alpha_s(x^2\theta^2)}{2\pi} \Theta\left(x\theta^2 > \rho\right)\, \Theta(x^2\theta^2 < \rho y). \tag{A.3}$$

This is better than (A.1) but still remains less efficient than (6.14) by double logarithms of ρ.

In the end, it is not our recommendation to use the groomed jet mass in the Y- or $\mathrm{Y_m}$-splitter condition.

— B —
Further details on jet shapes

B.1 Infrared (un)safety of C/A de-clustering

In this appendix, we provide a few additional details regarding the infrared unsafety of the μ^2 parameter with Cambridge/Aachen de-clustering. To avoid any possible confusion, we must stress that the discussion below only applies to the non-recursive version of the μ^2 parameter and that the recursive application of a μ_p^2 cut is infrared-safe for any p.

That said, let us consider a jet with three particles: a hard parton, a first emission with momentum fraction z_1 at an angle θ_1 and a second emission with momentum fraction z_2 at an angle θ_2, with $z_1\theta_1^2 > z_2\theta_2^2$ and $\theta_2 \ll \theta_1$. This corresponds to the leading-order ($\mathcal{O}(\alpha_s^2)$) configuration for a jet with $m^2 = (z_1\theta_1^2+z_2\theta_2^2)p_t^2$ and with a generic $\mu^2 = z_2\theta_2^2/(z_1\theta_1^2+z_2\theta_2^2)$. At the next order of the perturbation theory, one would have to include real emissions of gluons with momentum fraction z_3 and angle θ_3 as well as the corresponding virtual corrections and the soft divergence $z_3 \to 0$ is supposed to cancel between the real and virtual contributions. However, for $\theta_3 \gg \theta_1$ and $z_3 \to 0$, the virtual contribution would give $\mu^2_{\text{virt}} = z_2\theta_2^2/(z_1\theta_1^2 + z_2\theta_2^2)$ as for the 2-particle configuration, but the real emissions would give $\mu^2_{\text{real}} = 1$ because of the C/A de-clustering. This would lead to an infrared unsafety at μ^2_{virt}. This situation can happen at any value of μ, depending on the original three-particle configuration.

Although we have not made an explicit calculation, one might expect that the Sudakov $R_{\mu_p^2}$ would receive a contribution proportional to $(\alpha_s/p)\log^2(1/\theta_1^2)$, with $\theta_1^2 = \rho/z_1$, which diverges in the limit $p \to 0$.

B.2 Soft and large-angle emissions

In all the calculations we have performed so far, we have included hard collinear splittings which correspond to the terms proportional to B_i and B_g in our results. At the same order we could also have single-logarithmic contributions coming from soft and large-angle emissions. In practice, keeping the same notations as above, this means working in the approximation $z_2 \ll z_1$ without assuming any specific ordering between θ_1 and θ_2.

This can affect the calculations above at various levels: either through changes in the approximation used for the shape, where so far we have assumed a strong ordering, or through modifications of the matrix element for soft gluons at large angles.[1]

Let us first discuss the first effect. Since the expressions we have used so far are correct when $\theta_2 \ll \theta_2$ or when $\theta_2 \gg \theta_1$ we only have to worry about the region $\theta_2 \sim \theta_1$.

For N-subjettiness and the energy correlation functions, the correct expression in that region will only differ from the asymptotic one used so far by a constant, not enhanced by any parametrically large quantities. As a consequence, if we compute the difference to what has already been included, the integration over z_2 will at most bring a constant. Then, the angular integration over $\theta_2 \sim \theta_1$ will also at most bring a constant giving an overall NNLL subleading correction.

The situation is potentially a bit more tricky for μ^2 since the expression at $\theta_2 \sim \theta_1$ can vary between $z_2\theta_2^2/\rho$ and $z_2\theta_2^2/\theta_1^2$ potentially introducing a correction enhanced by $\log(1/z_1)$. Not making any assumption about angular ordering, eq. (7.12) becomes

$$\begin{aligned} R_{\mu^2_{1/2}}(z_1) = \int_0^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 \frac{dz_2}{z_2} \frac{\alpha_s C_R}{\pi} \Theta(z_2\theta_2^2 < \rho) \Bigg\{ & \Theta(\theta_2^2 < \theta_1^2/4)\, \Theta(z_2\theta_2^2 > \rho\mu^2) \\ & + \Theta(\theta_2^2 > \theta_1^2/4) \int_0^{2\pi} \frac{d\phi}{2\pi} \Big[\Theta(\theta_{12}^2 > \theta_2^2)\, \Theta(z_2\theta_2^2 > \rho\mu^2) \\ & + \Theta(\theta_{12}^2 < \theta_2^2)\, \Theta(z_2\theta_2^2 > \theta_1^2\mu^2) \Big] \Bigg\} \end{aligned} \tag{B.1}$$

where we have only considered primary emissions, worked with a fixed coupling approximation, and noticed that, for the sake of our calculation, we can safely replace $p_i(z_2)$ by $2C_R/z_2$. The angle ϕ that we have introduced is the angle between the two emissions, measured from the jet axis. This means that we have $\theta_{12}^2 = \theta_1^2 + \theta_2^2 - 2\theta_1\theta_2\cos(\phi)$. After computing the above integration, we find that all single-logarithmic terms cancel, leaving the same result as what we have obtained in section 7.2.2.

We are therefore left with potential single logarithms coming from the matrix element for the emission of soft and large-angle gluons. Taking the case of a quark jet, we therefore have to compute the following generic expression:

$$R = \int \frac{dz_2}{z_2}\, d^2\theta_2\, \frac{\alpha_s}{\pi^2} \left[\frac{C_A}{2} \frac{1}{\theta_{12}^2} + \frac{C_A}{2} \frac{\theta_{01}^2}{\theta_{02}^2\theta_{12}^2} + \left(C_F - \frac{C_A}{2} \right) \frac{1}{\theta_{02}^2} \right] \Theta(z_2\theta_2^2 < \rho)\, \Theta(v(z_i,\theta_i) > v). \tag{B.2}$$

If we focus on the single-logarithmic contribution, we can subtract the double-logarithmic piece, $C_F/\theta_{02}^2 + C_A/\theta_{12}^2\, \Theta(\theta_{12} < \theta_{01})$, and set $v(z_i,\theta_i) = (z_2\theta_2^2)/(z_1\theta_1^2)$ in what remains so that the z_2 integration yields a $\log(1/v)$. This gives

$$R_{\text{SL}} = \frac{\alpha_s C_A}{2\pi^2} \log(1/v) \int d^2\theta_2\, \frac{\theta_{01}^2}{\theta_{02}^2\theta_{12}^2} + \frac{2}{\theta_{12}^2}\, \Theta(\theta_{12}^2 < \theta_{01}^2), \tag{B.3}$$

where we have used the fact that $\int d^2\theta_2/\theta_{02}^2 = \int d^2\theta_2/\theta_{12}^2$. Up to subleading corrections, we can extend the θ_2 integration to infinity and show, e.g. using dimensional regularisation,

[1] In this discussion, we neglect additional effects from non-global logarithms.

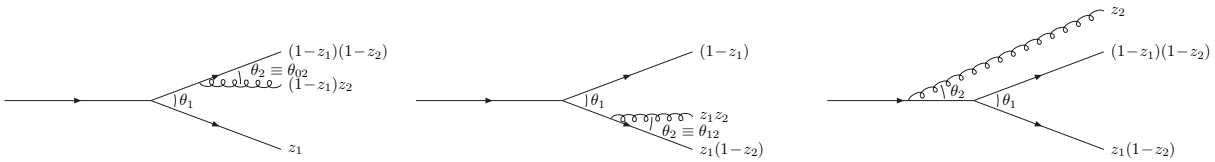


Figure B.1 — Three topologies potentially contributing to the emission of the gluon dominating the value of the shape, starting with a massive two-pronged object. Left: small-angle emission from the prong carrying a fraction $1-z_1$ of the jet p_t ("prong 1"), centre: small-angle emission from the prong carrying a fraction z_1 of the jet p_t ("prong 2"), right: large-angle emission from the parent object ("parent").

that it vanishes. In the end, there are no soft and large-angle single-logarithmic corrections to what we have computed earlier in the text.

B.3 Including finite z_1 corrections: QCD (background) and signal jets

We have argued in section 7.2.7 that if we wish to achieve NLL accuracy it is mandatory to include all finite z_1 and $1-z_1$ factors in our expressions for the shapes, with z_1 the fraction of the jet transverse momentum carried by the emission that dominates the mass of the jet. The main reason behind that is that they can be raised to powers of order $\alpha_s \log(1/v)$ which would give single-logarithmic corrections after integration over z_1.

In this appendix, our main goal is to discuss these extra source of NLL terms. As a fringe benefit of this discussion, we will at the same time provide a unified description of the signal and background distributions, allowing for interesting interpretations of the results obtained in this paper.

If we want to properly include the finite z_1 corrections we first need to carefully identify the origin of the gluon emissions. In the collinear limit, sufficient to capture all the finite z_1 corrections, colour coherence indicates that we can encounter three situations, represented in fig. B.1. The first two situations correspond to gluon emissions at small angle $\theta_2 \ll \theta_1$ from the splitting of either the hardest or the softest of the two prongs (carrying respectively a fraction $1-z_1$ and z_1 of the jet transverse momentum). These are the first two plots of fig. B.1 and will be referred to as the "prong 1" and "prong 2" topologies respectively for the $1-z_1$ and z_1 case. The third option corresponds to gluons emitted at large angle $\theta_2 \gg \theta_1$ from the parent parton in the jet. This is represented on the rightmost plot of fig. B.1 and will be called the "parent" topology in what follows. In that approach, the distribution for QCD jets will receive contributions from all three topologies — the first and third weighted by C_R and the second, corresponding to secondary emissions, weighted by C_A – while signal jets coming from the decay of colour-neutral bosons would only receive contributions from the first two topologies, both weighted by C_R.

For each of the three topologies, one then has to find the expression for the shape in

the soft and collinear limit for the gluon emission,[2] and impose that the first emission (z_1, θ_1) dominates the mass. The Sudakov factors for a given mass ρ, splitting momentum fraction z_1 and shape cut v, would then take the following form for each topology:

$$\begin{aligned}
R_{\text{prong1}} &= \int_0^{\theta_1^2} \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 p_{\text{prong1}}(z_2)\, \frac{\alpha_s}{2\pi}\, \Theta(v_{\text{prong1}}(z_1,\rho;z_2,\theta_2) > v)\Theta((1-z_1)^2 z_2\theta_2^2 < \rho),\\
R_{\text{prong2}} &= \int_0^{\theta_1^2} \frac{d\theta_{12}^2}{\theta_{12}^2} \int_0^1 dz_2 p_{\text{prong2}}(z_2)\, \frac{\alpha_s}{2\pi}\, \Theta(v_{\text{prong2}}(z_1,\rho;z_2,\theta_{12}) > v)\Theta(z_1^2 z_2\theta_{12}^2 < \rho),\\
R_{\text{parent}} &= \int_{\theta_1^2}^1 \frac{d\theta_2^2}{\theta_2^2} \int_0^1 dz_2 p_{\text{parent}}(z_2)\, \frac{\alpha_s}{2\pi}\, \Theta(v_{\text{parent}}(z_1,\rho;z_2,\theta_2) > v)\Theta(z_2\theta_2^2 < \rho),
\end{aligned} \tag{B.4}$$

where the splitting function would be the one of a quark, a gluon, or simply 0 for emissions from a color-neutral object, and $\rho = z_1(1-z_1)\theta_1^2$.

In practice, the two "prong" contributions are the same as the ones that will be computed in the case of signal jets, up to the constraint that the (z_1, θ_1) emission dominates the mass. This last term is irrelevant for signal jets as it would only contribute to a constant. For QCD jets it is however crucial to impose it for the emissions from the hard prong since, there, the $z_1 \ll 1$ region can give rise to large logarithms.

Strictly speaking, the finite z_1 corrections should only be kept in the expression for the shapes and the mass constraint in the emission from the soft prong is subleading for both the signal and the background. However, keeping these contributions makes the expressions more symmetric.

To fully specify our results, we just have to find the expressions of the three shapes we consider in each of the three topologies above. Following the same considerations as in the main text, it is easy to obtain

$$\tau_{21,\text{prong1}} = \frac{z_2}{z_1}\frac{\theta_2^2}{\theta_1^2} \qquad \tau_{21,\text{prong2}} = \frac{z_2}{1-z_1}\frac{\theta_{12}^2}{\theta_1^2} \qquad \tau_{21,\text{parent}} = \frac{z_2\theta_2^2}{\rho} \tag{B.5}$$

$$\mu^2_{\text{prong1}} = \frac{(1-z_1)z_2}{z_1}\frac{\theta_2^2}{\theta_1^2} \qquad \mu^2_{\text{prong2}} = \frac{z_1 z_2}{1-z_1}\frac{\theta_{12}^2}{\theta_1^2} \qquad \mu^2_{\text{parent}} \overset{\theta_2>\theta_{12}}{=} \frac{z_1 z_2\theta_2^2}{\rho} \overset{\theta_{12}>\theta_2}{=} \frac{(1-z_1)z_2\theta_2^2}{\rho} \tag{B.6}$$

$$C_{2,\text{prong1}} = \frac{z_2}{z_1}\theta_2^2 \qquad C_{2,\text{prong2}} = \frac{z_2}{1-z_1}\theta_{12}^2 \qquad C_{2,\text{parent}} = \frac{z_2\theta_2^4}{\rho}. \tag{B.7}$$

For parent emissions, we again had to separate two cases for the mass-drop parameter corresponding to the clustering of the second emission with one of the two prongs, with θ_2 being the angle wrt "prong 1" and θ_{12} the angle to "prong 2".

With these expressions we can compute the Sudakov form factors. It is convenient to introduce $C_{R,1}$, $C_{R,2}$ and $C_{R,p}$ respectively as the colour factors associated with the "prong 1", "prong 2" and "parent" topologies. Similarly, we denote B_1, B_2 the hard-splitting coefficient associated with the two "prong" configurations, realising that the

[2] Meaning in particular that one can discard the $1 - z_2$ factors.

large-angle topology will not receive a hard-splitting correction. Note that in the case of a boson decay, we can simply set $C_{R,p} = 0$.

The results for the emissions collinear to the $1 - z_1$ branch ("prong 1") are as follows:

$$\begin{aligned} R_{\tau,\text{prong1}}(z_1) =& \Big[T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{L_\rho + L_1 + L_-}{2} + L_v; C_{R,1}, B_1\Big) \\ & - T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{L_\rho + L_1 - L_-}{2}; C_{R,1}, B_1\Big)\Theta(L_1 > L_-)\Big] \\ & \Theta(L_v + L_1 > 0)\,\Theta(L_v + L_- > 0) \end{aligned} \tag{B.8}$$

$$\begin{aligned} R_{\mu^2,\text{prong1}}(z_1) =& \Big[T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{L_\rho + L_1 - L_-}{2} + L_v; C_{R,1}, B_1\Big) \\ & - T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{L_\rho + L_1 - L_-}{2}; C_{R,1}, B_1\Big)\Theta(L_1 > L_-)\Big] \\ & \Theta(L_v > L_- - L_1)\,\Theta(L_v > 0) \end{aligned} \tag{B.9}$$

$$\begin{aligned} R_{C_2,\text{prong1}}(z_1) =& \Big[T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{3L_- + 3L_1 - L_\rho}{2} + L_v; C_R, B_i\Big) \\ & - T_{02}\Big(\frac{L_\rho - L_1 + L_-}{2}, \frac{L_\rho + L_1 - L_-}{2}; C_{R,1}, B_1\Big)\Theta(L_1 > L_-)\Big] \\ & \Theta(L_v > L_\rho - L_- - 2L_1)\,\Theta(L_v > L_\rho - L_1 - 2L_-), \end{aligned} \tag{B.10}$$

where the last two Θ constraints come from the fact that the first term has to be positive and larger than the second term. Note that the second term in each of these three expressions is the same and come from the kinematic constraint than the second emission (z_2, θ_2) does not dominate the mass.

The results for the "prong 2" topology have not been given explicitly but can be directly obtained from the "prong 1" topology by inverting L_1 and L_- which corresponds to inverting z_1 and $1 - z_1$.

For the emissions from the parent object, we find in a similar way

$$R_{\tau,\text{parent}}(z_1) = P_2\Big(\frac{L_\rho + L_1 + L_-}{2}, L_\rho, L_\rho + L_v; C_{R,p}, 0\Big)\,\Theta(L_v > 0) \tag{B.11}$$

$$\begin{aligned} R_{\mu^2,\text{parent}}(z_1) =& \frac{1}{2}P_2\Big(\frac{L_\rho + L_1 + L_-}{2}, L_\rho, L_\rho - L_- + L_v; C_{R,p}, 0\Big)\,\Theta(L_v > L_-) \\ & + \frac{1}{2}P_2\Big(\frac{L_\rho + L_1 + L_-}{2}, L_\rho, L_\rho - L_1 + L_v; C_{R,p}, 0\Big)\,\Theta(L_v > L_1) \end{aligned} \tag{B.12}$$

$$\begin{aligned} R_{C_2,\text{parent}}(z_1) =& \Big[P_2\Big(\frac{L_\rho + L_1 + L_-}{2}, L_\rho, L_1 + L_- + L_v; C_{R,p}, 0\Big) \\ & + T_{24}(L_1 + L_- + L_v, L_\rho + L_v; C_{R,p}, 0)\Big]\,\Theta(L_v > L_\rho - L_1 - L_-) \\ & + T_{24}(L_\rho, L_\rho + L_v; C_{R,p}, 0)\,\Theta(0 < L_v < L_\rho - L_1 - L_-) \end{aligned} \tag{B.13}$$

B.4 Further comparisons

In this appendix, we provide a few additional comparisons between our analytic predictions and Monte Carlo simulations.

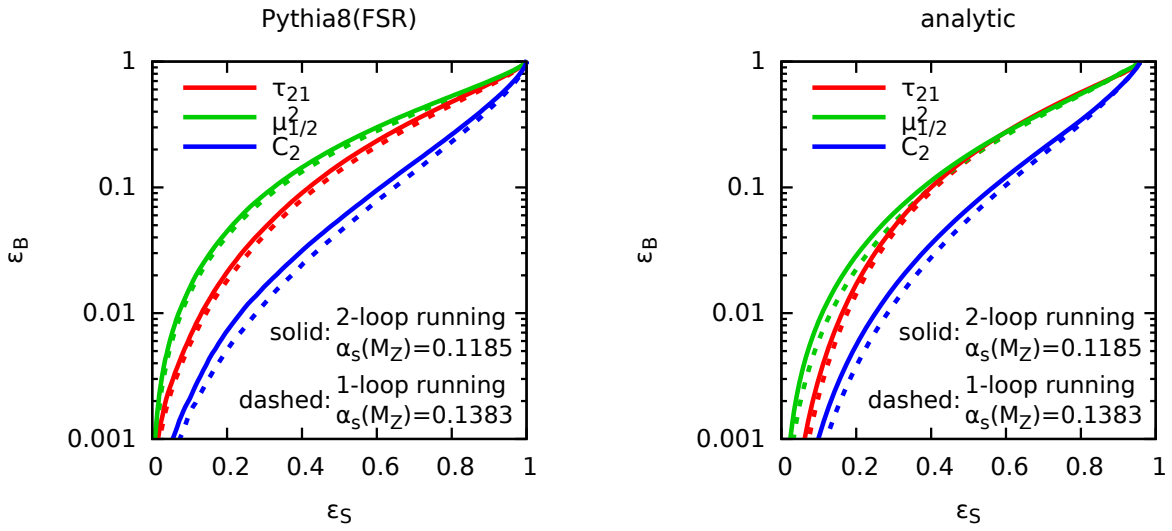


Figure B.2 — Similar plot as in fig. 7.10 where we show Pythia results (left) and analytic calculations (right) of the signal and background efficiencies for two different running-coupling prescriptions: a one-loop running with $\alpha_s(m_Z) = 0.1383$ (dashed, our default for Pythia in the main text) and a two-loop running with $\alpha_s(m_Z) = 0.1185$ (solid, our default for analytic results in the main text).

One-loop vs. two-loop running coupling. First, in sections 7.2.9 and 7.3.6, we have used a one-loop running of α_s, with $\alpha_s(m_Z) = 0.1383$, for Pythia simulations, and compared that to analytic calculations including two-loop corrections and using $\alpha_s(m_Z) = 0.1185$. In the case of our analytic calculation, this choice is motivated by the fact that two-loop corrections are easily included and we then used the world-average value [53] at the Z-boson mass. For the Pythia simulation, we simply kept the default which is a one-loop running.

We could also have run Pythia with a two-loop running of the coupling and impose $\alpha_s(m_Z) = 0.1185$. We did not do that in the main text because that can only safely be done with a retuning of other parameters in Pythia (mostly for the non-perturbative effects). It is however interesting to check that this difference in the treatment of the running of the strong coupling does not come with large effects. The result is presented in fig. B.2, where we see that this is indeed a small effect which does not alter in any way the conclusions of this paper. We also see from that figure that the size of the effect is similar in Monte Carlo simulations and in our analytic predictions.

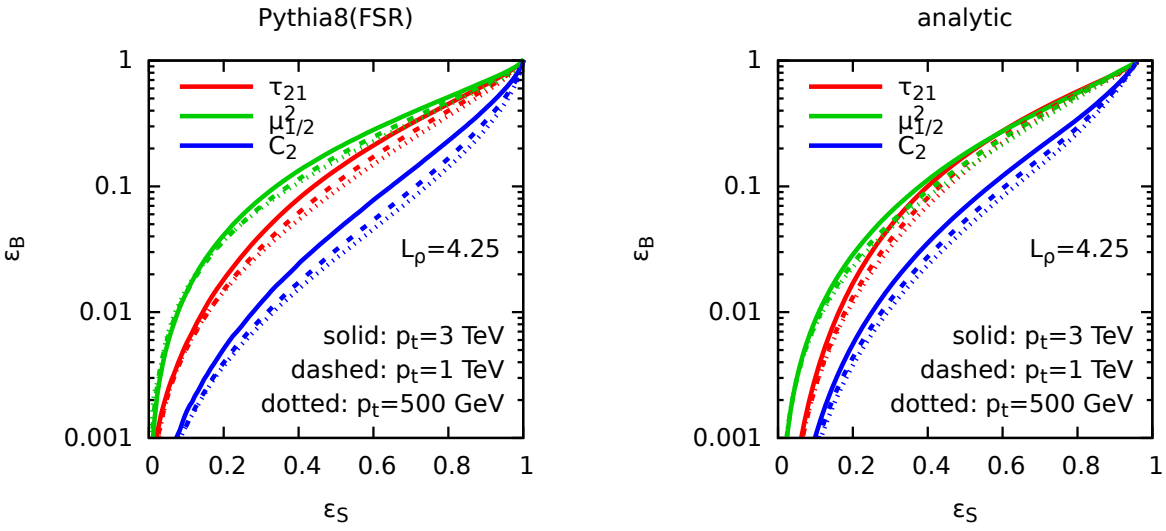


Figure B.3 — Similar plot as in fig. 7.10 where we show Pythia results (left) and analytic calculations (right) of the signal and background efficiencies for two different running-coupling prescriptions obtained for different jet transverse momenta, keeping L_ρ fixed to 4.25 (or, in the 4-4.5 range for Pythia simulations).

Note also that another interesting check of our results is to compare our fixed-order results with Pythia simulations also done with a fixed coupling. Although we do not show explicit plots here, this comparison shows similar features as the ones observed with a running-coupling prescription.

Dependence on the jet transverse momentum. Throughout this paper, we have shown results for jets with a large transverse momentum of 3 TeV. Here, we briefly show that our calculations remain valid for less boosted jets, closer to those used in today's phenomenological analyses.

In fig. B.3, we show ROC curves obtained from Pythia simulations and our analytic calculations, for three different jet transverse momenta: 3 TeV, 1 TeV and 500 GeV. For this comparison, we have kept the ratio m/p_t fixed, i.e. considered a mass of 358, 120 and 60 GeV respectively for each of the three p_t scales. We see that the dependence on the jet p_t is mild, which is expected since the result only depend on p_t through the $p_t R$ scale entering in α_s. Our conclusions are therefore also valid for jets of more moderate transverse momentum. Note that the small differences observed in Pythia simulations between different jet p_t are well reproduced by our analytic calculation.

B.5 Results for jet shapes with grooming

In this appendix we present the Sudakov factors corresponding to the (non-recursive) jet shapes studied in chapter 7 with the addition of a SoftDrop grooming step, with parameters z_{cut} and β. These results are expressed using the basic building blocks. The results for N-subjettiness are not present in this appendix, as they are already included in chapter 8, where we discussed the dichroic N-subjettiness ratio.

In this case, the jet mass distribution for QCD jets can be written in a similar manner as eq. (7.1), with some adaptations

$$\frac{\rho}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\rho}\bigg|^{\text{SD}}_{<v} = \int_{z_{\text{cut}}}^{1} dz_1\, p_i(z_1)\, \frac{\alpha_s(\sqrt{z_1}\rho)}{2\pi} e^{-R^{\text{SD}}_{\text{plain}}(\rho)-R^{\text{SD}}_{v}(z_1,\rho)}. \tag{B.14}$$

The SoftDrop mass distributions Sudakov is given by

$$R^{\text{SD}}_{\text{plain}}(\rho) = T_{02}(0, L_\rho; C_R, B_i) - T_{-\beta 2}(L_z, L_\rho; C_R, 0), \tag{B.15}$$

where $L_z = \log(1/z_{\text{cut}})$.

We note that the addition of SoftDrop does not alter the secondary emissions. This is a consequence the strong ordering between emission angles, so secondary emissions are never selected by the C/A de-clustering procedure. For a more concrete example, suppose the mass of our jet is set by an emission p_1, with a certain emission angle θ_1. Then, p_1 emits another parton p_2 with an emission angle θ_2. At LL accuracy, we have $\theta_2 \ll \theta_1$. As a consequence, p_1 is always going to be selected by the SoftDrop procedure first, it cannot be eliminated because it is the emission that sets the mass of the jet, so it passes the SoftDrop condition and stops the procedure. So there are no additional constraints in the secondary emission and they are going to be given by the same expressions as for the ungroomed case. The calculations for signal emissions are not affected for the same reason, as electroweak boson decays have no collinear divergences.

B.5.1 Mass Drop

As stated in the main text, the Sudakov will have two different contributions, depending on how partons are clustered. The first contribution has an a new transition point $L_{\text{tr},0} = -\beta L_\rho/2 + (\beta+2)L_1/2 + L_v$. It is given by

$$\begin{aligned} R^{\text{SD}}_{\mu^2_{1/2},0}(z_1) &\overset{L_z<L_{\text{tr},0}}{=} T_{-\beta 2}\left(\frac{L_\rho+L_1}{2}, L_z + \frac{\beta+1}{2}(L_\rho - L_1); C_R, 0\right) \\ &\quad + T_{02}\left(\frac{L_\rho-L_1}{2}, \frac{L_\rho+L_1}{2}+L_v; C_R, B_i\right) - T_{02}\left(\frac{L_\rho-L_1}{2}, \frac{L_\rho+L_1}{2}; C_R, B_i\right) \\ &\overset{L_z>L_{\text{tr},0}}{=} T_{-\beta 2}\left(\frac{(1+\beta)L_\rho - L_v + L_z}{2+\beta}, \frac{(1+\beta)(L_\rho+L_v)+L_z}{2+\beta}; C_R, 0\right) \\ &\quad + P_2\left(\frac{L_\rho+L_1}{2}, \frac{(1+\beta)L_\rho - L_v + L_z}{2+\beta}, \frac{(1+\beta)(L_\rho+L_v)+L_z}{2+\beta}; C_R, 0\right) \end{aligned} \tag{B.16}$$

$$+ T_{02}\left(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2} + L_v; C_R, B_i\right) - T_{02}\left(\frac{L_\rho - L_1}{2}, \frac{L_\rho + L_1}{2}; C_R, B_i\right).$$

The second contribution is non-zero only for the case $L_v > L_1$. It has an additional transition point at $L_{\text{tr},1} = -\beta(L_\rho - L_1)/2 + L_v$ and it is given by

$$\begin{aligned} R^{\text{SD}}_{\mu^2_{1/2},1}(z_1) &\overset{L_z<L_{\text{tr},1}}{=} T_{-\beta 2}\left(\frac{L_\rho + L_1}{2}, L_z + \frac{\beta+1}{2}(L_\rho - L_1); C_R, 0\right) \\ &\quad + T_{02}\left(\frac{L_\rho + L_1}{2}, \frac{L_\rho - L_1}{2} + L_v; C_A, B_g\right) \\ &\overset{L_z>L_{\text{tr},1}}{=} T_{-\beta 2}\left(\frac{(1+\beta)L_\rho + L_1 + L_z - L_v}{2+\beta}, \frac{(1+\beta)(L_\rho + L_v - L_1) + L_z}{2+\beta}; C_R, 0\right) \\ &\quad + P_2\left(\frac{L_\rho + L_1}{2}, \frac{(1+\beta)L_\rho + L_1 + L_z - L_v}{2+\beta}, \frac{(1+\beta)(L_\rho + L_v - L_1) + L_z}{2+\beta}; C_R, 0\right) \\ &\quad + T_{02}\left(\frac{L_\rho + L_1}{2}, \frac{L_\rho - L_1}{2} + L_v; C_A, B_g\right). \end{aligned} \tag{B.17}$$

The total Sudakov $R^{\text{SD}}_{\mu^2_{1/2}}(z_1)$ is the sum of these two contributions.

B.5.2 Energy Correlation

We have a transition point at $c = \rho/z_1$, the same as the case without SoftDrop (see eq. (7.18)). For the first region, $L_v < L_\rho - L_1$ the Sudakov is

$$R^{\text{SD}}_{C_2}(z_1) = \Theta\left(L_z > L_\rho - \frac{2+\beta}{2}L_v\right)\left[T_{-\beta 2}\left(\frac{(1+\beta)L_\rho + L_Z}{2+\beta}, L_y; C_R, 0\right) - T_{24}\left(L_\rho - \frac{L_v}{2}, L_y; C_R, 0\right)\right], \tag{B.18}$$

with,

$$L_y = \frac{1+\beta}{4+\beta}(L_\rho + L_v) + \frac{3}{4+\beta}L_z. \tag{B.19}$$

For the region where $L_v < L_\rho - L_1$, we have an additional transition point at $L_{\text{tr,C}} = L_v + (4+\beta)L_1/2 - (2+\beta)L_\rho/2$,

$$\begin{aligned} R^{\text{SD}}_{C_2}(z_1) &\overset{L_z>L_{\text{tr,C}}}{=} T_{02}(0, L_v + L_1; C_R, B_i) - T_{-\beta 2}(L_z, L_v + L_1; C_R, 0) - R^{\text{SD}}_{\text{plain}}(\rho) \\ &\quad + T_{-\beta 4}\left(\frac{(1+\beta)(L_v + L_1) + L_z}{2+\beta}, L_y; C_R, 0\right) + T_{24}\left(L_v + \frac{3L_1 - L_\rho}{2}, L_y; C_R, 0\right) \\ &\quad + T_{02}\left(\frac{L_\rho + L_1}{2}, \frac{3L_1 - L_\rho}{2} + L_v; C_A, B_g\right) \\ &\overset{L_z<L_{\text{tr,C}}}{=} T_{02}(0, L_v + L_1; C_R, B_i) - T_{-\beta 2}(L_z, L_v + L_1; C_R, 0) - R^{\text{SD}}_{\text{plain}}(\rho) \\ &\quad + T_{-\beta 2}\left(L_z + \frac{\beta+1}{2}(L_\rho - L_1), L_v - \frac{3L_1 - L_\rho}{2}; C_R, 0\right) \\ &\quad + T_{02}\left(\frac{L_\rho + L_1}{2}, \frac{3L_1 - L_\rho}{2} + L_v; C_A, B_g\right). \end{aligned} \tag{B.20}$$

— C —
Further details on dichroic subjettiness ratios

C.1 Dichroic subjettiness ratios for $\beta_\tau = 1$

In section 8.1, we have argued in favor of the dichroic subjettiness ratios using N-subjettiness with $\beta_\tau = 2$. In this appendix, we briefly discuss the case $\beta_\tau = 1$, for which the dichroic variant can also be considered. Note that for $\beta_\tau = 1$, we have defined the N-subjettiness axes through an exclusive-k_t declustering. This can be done either using the standard E-scheme four-vector recombination or the winner-takes-all (WTA) recombination scheme. For simplicity, we will focus on E-scheme results here. A brief comparison between the two axis choices is shown in fig. 8.10(right).

Fig. C.1 shows ROC curves similar to those presented in Figs. 8.6 and 8.7, this time including results for $\beta_\tau = 1$ as dashed lines.

We can make several observations based on these plots. First, as for $\beta_\tau = 2$, we see that the dichroic ratio also outperforms the other combination for $\beta_\tau = 1$. The performance gain is however smaller, especially with SD grooming.

In terms of the sensitivity to non-perturbative effects, we see that N-subjettiness ratios with $\beta_\tau = 1$ are rather stable even without any SD grooming step. This small sensitivity to non-perturbative effects might have been anticipated since the corresponding k_t cut is less affected by soft-and-large-angle emissions than for $\beta_\tau = 2$. A consequence of this observation is that grooming is less critical when using a cut on N-subjettiness ratios with $\beta_\tau = 1$, and without SD grooming the dichroic combination shows a more sizeable performance gain compared to the other approaches, cf. the bottom-left plot of Fig. C.1.

Finally, we can argue that $\beta_\tau = 2$ gives somewhat better performance than $\beta_\tau = 1$. To be fair, the comparison should be made between $\tau_{21,\text{groomed}}^{\text{dichroic}}$ for $\beta_\tau = 2$ (the solid black line on the bottom-right plot of fig. C.1) and $\tau_{21}^{\text{dichroic}}$ for $\beta_\tau = 1$ (the dashed black line on the bottom-left plot) which both show good signal significance and limited non-perturbative corrections. This comparison shows a somewhat larger background rejection in the $\beta_\tau = 2$ case for typical signal efficiencies in the $0.2-0.6$ range, as also seen in fig. 8.10.

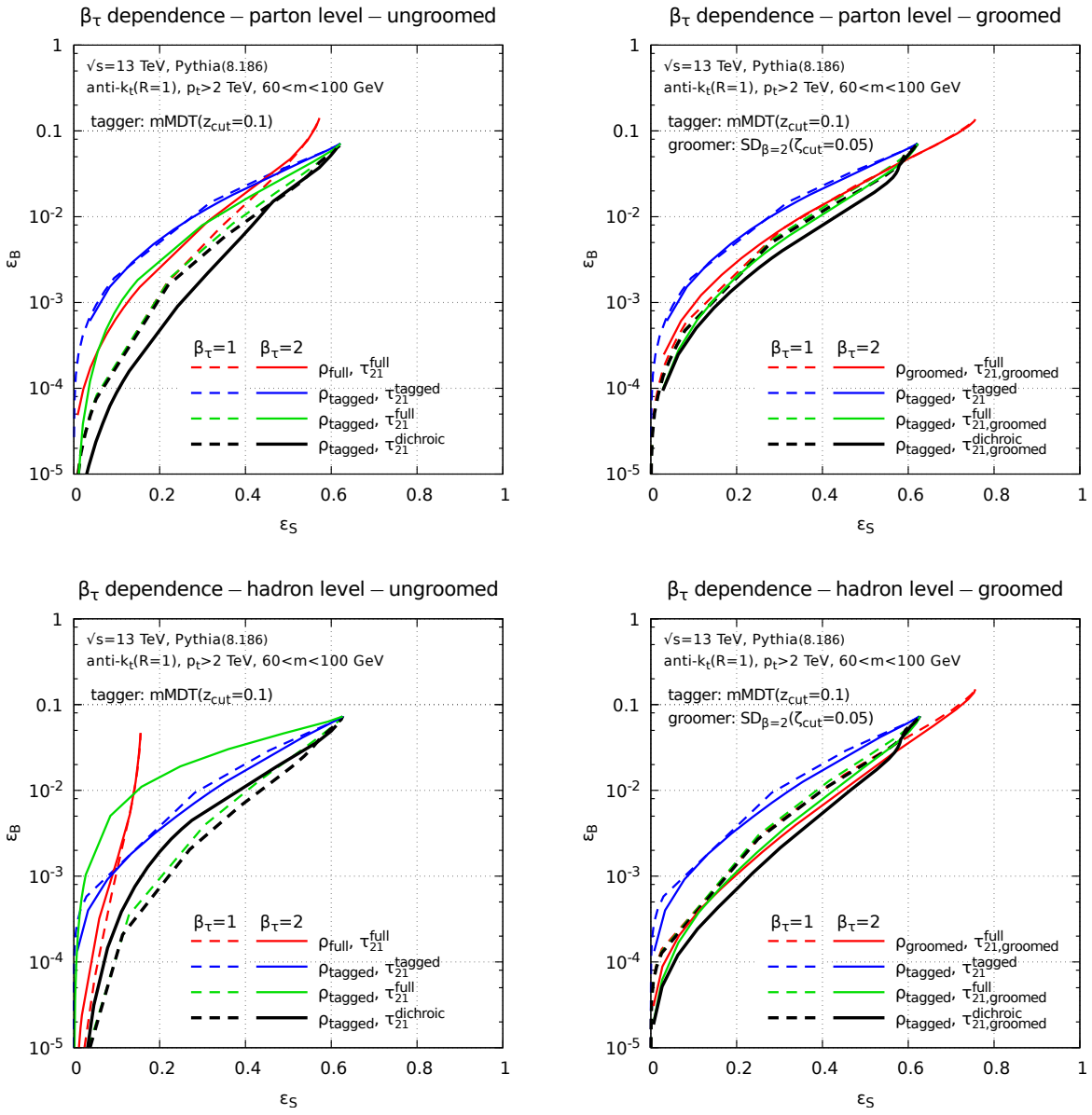


Figure C.1 — ROC curves providing a comparison between different N-subjettiness ratios for $\beta_\tau = 1$ (dashed lines) and $\beta_\tau = 2$ (solid lines). The same 4 variants as in Figs. 8.6 and 8.7 are included. The left (right) column corresponds to full (SD-groomed) jets. The top (bottom) row corresponds to parton-level (hadron-level) events.

C.2 Example code for dichroic subjettiness ratios

In this last appendix, we briefly indicate how dichroic subjettiness ratios can be implemented using tools available in `FastJet` and `fjcontrib`. In particular, we make use of the `RecursiveTools` contrib (for `ModifiedMassDropTagger` and `SoftDrop`) and of the `Nsubjettiness` contrib.

First, besides standard `FastJet` headers needed for jet clustering, one needs to include the following headers:

```
#include <fastjet/contrib/ModifiedMassDropTagger.hh>
// mMDT tagger
#include <fastjet/contrib/SoftDrop.hh>
// optional SD grooming
#include <fastjet/contrib/Nsubjettiness.hh>
// tau1 and tau2
```

Then, one should declare the basic objects needed for tagging, computing τ_1 and τ_2, and, optionally, grooming:

```
// the tagger [here mMDT with a z cut]
// Note: by default, this automatically reclusters with C/A
double zcut = 0.1;
fastjet::contrib::ModifiedMassDropTagger mmdt_tagger(zcut);

// (optional) groomer [here SoftDrop]
// Note: by default, this automatically reclusters with C/A
double beta    = 2.0;
double zetacut = 0.05;
fastjet::contrib::SoftDrop sd_pre_groomer(beta, zetacut);

// N-subjettiness with beta_tau=2 and gen-kt axes
// (for theoretical reasons it is preferred to use
// an unnormalised measure)
double beta_tau = 2.0;
fastjet::contrib::UnnormalizedMeasure measure(beta_tau);
fastjet::contrib::GenKT_Axes axes_gkt(1.0/beta_tau);
fastjet::contrib::Nsubjettiness tau1(1, axes_gkt, measure);
fastjet::contrib::Nsubjettiness tau2(2, axes_gkt, measure);
```

Note that all parameters here are given as examples and have not been optimised. Also, when used with events contaminated by pileup, a proper pileup mitigation technique should be implemented. This can for example be done by passing a `fastjet::Subtractor` to the mMDT and SD via the `set_subtractor` method, and using a `GenericSubtractor` [134] or a `ConstituentSubtractor` [98] for the N-subjettiness variables. Alternatively one can use methods that carry out event-wide pileup-suppression such as PUPPI [100] or SoftKiller [99].

Finally, for a given jet (`jet` in the example below), one can compute the dichroic subjettiness ratio using

```
fastjet::PseudoJet jet; // given jet

fastjet::PseudoJet pre_groomed_jet = sd_pre_groomer(jet);
// grooming
fastjet::PseudoJet tagged_jet = mmdt_tagger(pre_groomed_jet);
// tagging

double tau1_tagged    = tau1(tagged_jet);
// $\tau_1^{\text{tagged}}$
double tau2_groomed   = tau2(pre_groomed_jet);
// $\tau_2^{\text{groomed}}$

double tagged_mass = tagged_jet.m();
// tagged mass
double tau21_dichroic = tau2_groomed/tau1_tagged;
// $\tauDG$
```

— D —
Further details on groomed mass distribution

D.1 Details of the analytic calculation

D.1.1 Resummed exponents

Our results are expressed in terms of $\alpha_s = \alpha_s(R\,p_t)$, evolved from $\alpha_s(m_Z) = 0.118$ with a two-loop approximation ($n_f = 5$).[1] Note that for the minimal jet mass of 1 GeV that we consider in this paper and the variations of the renormalisation and resummation scales, μ_R and μ_Q, our perturbative results always remain above the Landau pole. We could decide to freeze the coupling at a scale $\mu_{\rm NP}$ that we would vary around 1 GeV, and hence obtain an uncertainty associated to using perturbative QCD in a region sensitive to non-perturbative effects. However this effect should be included already in our estimate of the non-perturbative effects via the Monte Carlo simulations discussed in section 9.3.

To obtain the results presented in the main text, we have written the splitting functions entering the flavor-diagonal contributions as a sum of two different contributions:

$$p_{gq}(z) = \frac{C_F}{z}\Theta(z < e^{B_q}) + p_{gq}^{(\rm finite)}(z), \tag{D.1a}$$

$$p_{xg}(z) = \frac{C_A}{z}\Theta(z < e^{B_g}) + p_{xg}^{(\rm finite)}(z). \tag{D.1b}$$

The cut-off at $z = e^{B_i}$ is such that the leftover finite part only generates power corrections in $z_{\rm cut}$ while the $\log(1/z_{\rm cut})$ and constant terms are included in the first terms proportional to $1/z$. Note that this will naturally produce distributions with an end-point at $\rho = e^{B_i}$. That said, the contribution from the first term can be integrated straightforwardly and gives the $\mathcal{R}_i$ function given in eq. (9.8).

[1] Our use of the two-loop running coupling to compute α_s at the hard scale comes from the fact that we ultimately match our resummed calculation to a NLO fixed-order calculation which itself uses the two-loop running coupling as obtained from the NLO CT14 PDF set [113].

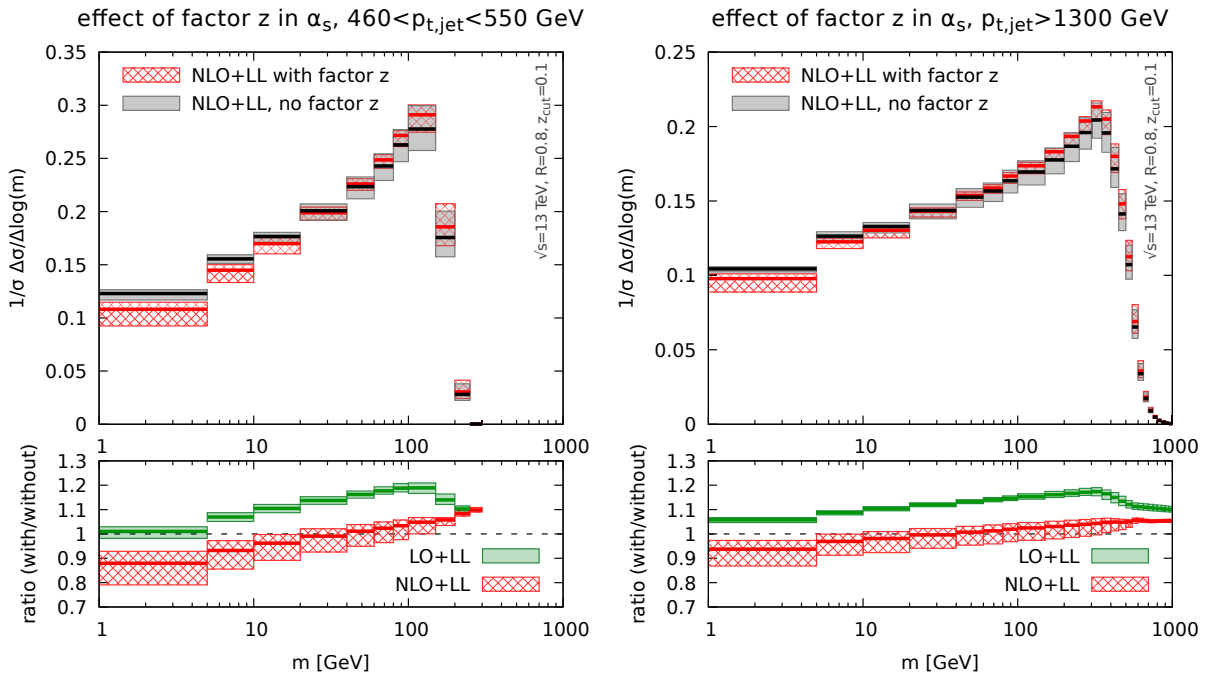


Figure D.1 — Comparison of the jet mass distribution with and without the resummation of logarithmic corrections in z_{cut} originating from the running of the strong coupling. We note that these effects are sizeable, although still within the theoretical uncertainty.

Next, we consider the contributions coming from the second term in eq. (D.1), as well as from the flavor-changing contributions, which will be power-suppressed in z_{cut}. For these, we can safely ignore the factor z in both the argument of α_s and the constraint $\Theta(z\theta^2 > \rho)$. The z and θ^2 integration then factorise to give

$$\text{finite part:} \quad \int_{\rho}^{z_{\text{cut}}} \frac{d\theta^2}{\theta^2} \frac{\alpha_s(\theta p_t R)}{\pi} \int_{z_{\text{min}}}^{z_{\text{max}}} dz\, p_{ij}^{(\text{finite})}(z), \tag{D.2}$$

where the integration boundaries z_{min} and z_{max} depend on which matrix element we consider and should match those imposed by the mMDT conditions in eq. (9.7). Once again, to our accuracy, there is some freedom in the choice of the upper integration boundary of the θ^2 integration. Setting it to z_{cut} ensures that there are no corrections beyond the transition point $\rho = z_{\text{cut}}$. Note that neglecting the finite z_{cut} effects is equivalent to keeping only the contribution from $\mathcal{R}_i$ while neglecting the contribution from eq. (D.2).

D.1.2 Impact of the z factor in the scale of the running coupling

If the parameter z_{cut} is chosen to be rather small, finite-z_{cut} corrections are negligible but logarithmic corrections can become relevant. The resummation of the leading-logarithmic corrections in z_{cut} is relatively straightforward and it was discussed in Ref. [12] (see also Refs. [74, 75]). Firstly, successive gluon emissions must be ordered in mass rather than in angle. Secondly, the argument of the QCD running coupling should be taken as $k_t = z\theta p_t R$

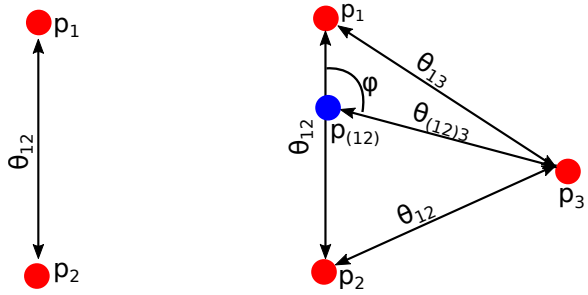


Figure D.2 — Configurations with maximal mass for LO (left) and NLO (right).

(at least for the calculation of $\mathcal{R}_i$). Both effects are included in our calculation. However, to LL accuracy (in ρ) the argument of the running coupling could more simply be chosen as $\theta p_t R$. This choice leads to simpler analytic expressions and is what we naturally obtain when we consider bins of $p_{t,\text{mMDT}}$, see eq. (9.25). It is therefore of some interest to investigate how neglecting the factor z in the argument of the running coupling affects our findings. In this case, the $\mathcal{R}_i$ functions in eq. (9.9) become

$$\widetilde{\mathcal{R}}_i = \frac{1}{\pi\alpha_s\beta_0^2}\Big[W\Big(1+\alpha_s\beta_0(\log(\rho)-B_i)\Big) - W\Big(1+\alpha_s\beta_0\log(\rho/z_m)\Big) - \alpha_s\beta_0(\log(z_m)-B_i)\Big]. \tag{D.3}$$

In fig. D.1 we show the impact of these corrections on the normalised matched distributions. Remembering that the uncertainty on the lower panels is the actual uncertainty on the ratio, we see that the effects are genuinely present. However, they remain within our overall theoretical uncertainties shown on the mass distribution (upper plots).

D.2 End-point of the ρ distribution

As discussed in section 9.1.2, we have modified the argument $\log(1/\rho)$ to take into account end-point effects i.e. the fact that ρ has a maximum value $\rho_{\max}$ for a jet with transverse momentum p_t and radius R. In this appendix, we give the details of the computation of $\rho_{\max}$ at LO and NLO.

At LO, where we have two partons p_1 and p_2 in the jet, the calculation is straightforward. The mass of the jet, and therefore ρ, will be maximal when the final partons are as distant as possible, but are still clustered into a single jet. Let us first work in the small-angle limit. Then, the angular distance between the two partons is $\theta_{12} = R$, as shown in the left plot of fig. D.2. If the two partons carry a transverse momentum $p_{t1} = xp_t$ and $p_{t2} = (1-x)p_t$, respectively, the jet mass is given by

$$m^2 = p_t^2 R^2 x(1-x). \tag{D.4}$$

This is maximal when the momentum is equally distributed between the two partons, $x = 1/2$, for which we have $\rho_{\text{max,LO}}^{(\text{small-}R)} = 1/4$. If we relax our small-angle approximation, we should take into account that the mass of the system of two partons separated by a

distance R will depend on their orientation in the rapidity-azimuthal angle plane. It is straightforward to include this in the above analytic calculation and we find that ρ is maximal when the two partons have the same rapidity, leading to $\rho_{\text{max,LO}} = R^{-2}\tan^2\frac{R}{2}$ [84]. For our choice of $R = 0.8$, this gives $\rho_{\text{max,LO}} = 0.279303$.

At NLO, the same reasoning applies but is complicated by the presence of one more parton in the jet. We start again by considering the small-R limit. Remembering that the three partons must be clustered into a single anti-k_t jet of radius R, we can assume, without loss of generality, that p_1 and p_2 are the first pair of partons to be clustered into a subjet with momentum p_{12}, with p_{12} then clustered with parton p_3. In order to have all 3 partons clustered into a single jet, we must must have $\theta_{12} \leq R$ and $\theta_{(12)3} \leq R$. We define φ as being the angle between θ_{12} and $\theta_{(12)3}$, as shown in the right plot of fig. D.2, and we parametrise the momentum fractions of the partons as

$$z_1 = xt, \qquad z_2 = x(1-t), \qquad z_3 = 1 - x. \tag{D.5}$$

Since $\theta_{(12)1} = (1-t)\theta_{12}$ and $\theta_{(12)2} = t\theta_{12}$, we have

$$\theta_{13}^2 = (1-t)^2\theta_{12}^2 + \theta_{(12)3}^2 + 2(1-t)\theta_{12}\theta_{(12)3}\cos\varphi, \tag{D.6}$$

$$\theta_{23}^2 = t^2\theta_{12}^2 + \theta_{(12)3}^2 + 2t\theta_{12}\theta_{(12)3}\cos\varphi. \tag{D.7}$$

The jet mass is then found to be

$$m^2 = p_t^2(z_1z_2\theta_{12}^2 + z_1z_3\theta_{13}^2 + z_2z_3\theta_{23}^2) = p_t^2xt(1-t)\theta_{12}^2 + p_t^2x(1-x)\theta_{(12)3}^2. \tag{D.8}$$

This is maximal for $\theta_{12} = \theta_{(12)3} = R$ and momentum equally distributed between p_1 and p_2, *i.e.* $t = 1/2$, in which case we have

$$m^2 = p_t^2R^2x\left(\frac{5}{4} - x\right). \tag{D.9}$$

The maximum jet mass is thus reached for $x = 5/8$, which corresponds to $\rho_{\text{max,NLO}}^{(\text{small-}R)} = 25/64$. If we lift the small-$R$ approximation, the situation becomes more complex since the mass now depends explicitly on the angle φ as well as on an additional overall rotation angle ψ of the 3-parton system. One can write analytic expressions for the jet mass and transverse-momentum conservation and, for given values of φ and ψ we can maximise the mass. The maximisation over φ and ψ has been done numerically — imposing that $\Delta R_{12} < R$ and $\Delta R_{(12)3} < R$ as required by the clustering — and we find is $\rho_{\text{max,NLO}} = 0.44974$ for $R = 0.8$.

D.3 LL predictions for the $p_{t,\text{mMDT}}$ jet cross-section

Before investigating in detail the double-differential cross-section $d^2\sigma/(dp_{t,\text{mMDT}}\,dm)$, one might be tempted to study the jet cross-section, $d\sigma/dp_{t,\text{mMDT}}$. Despite looking simpler, the latter is actually plagued with the issue of IRC unsafety, while for the former, the

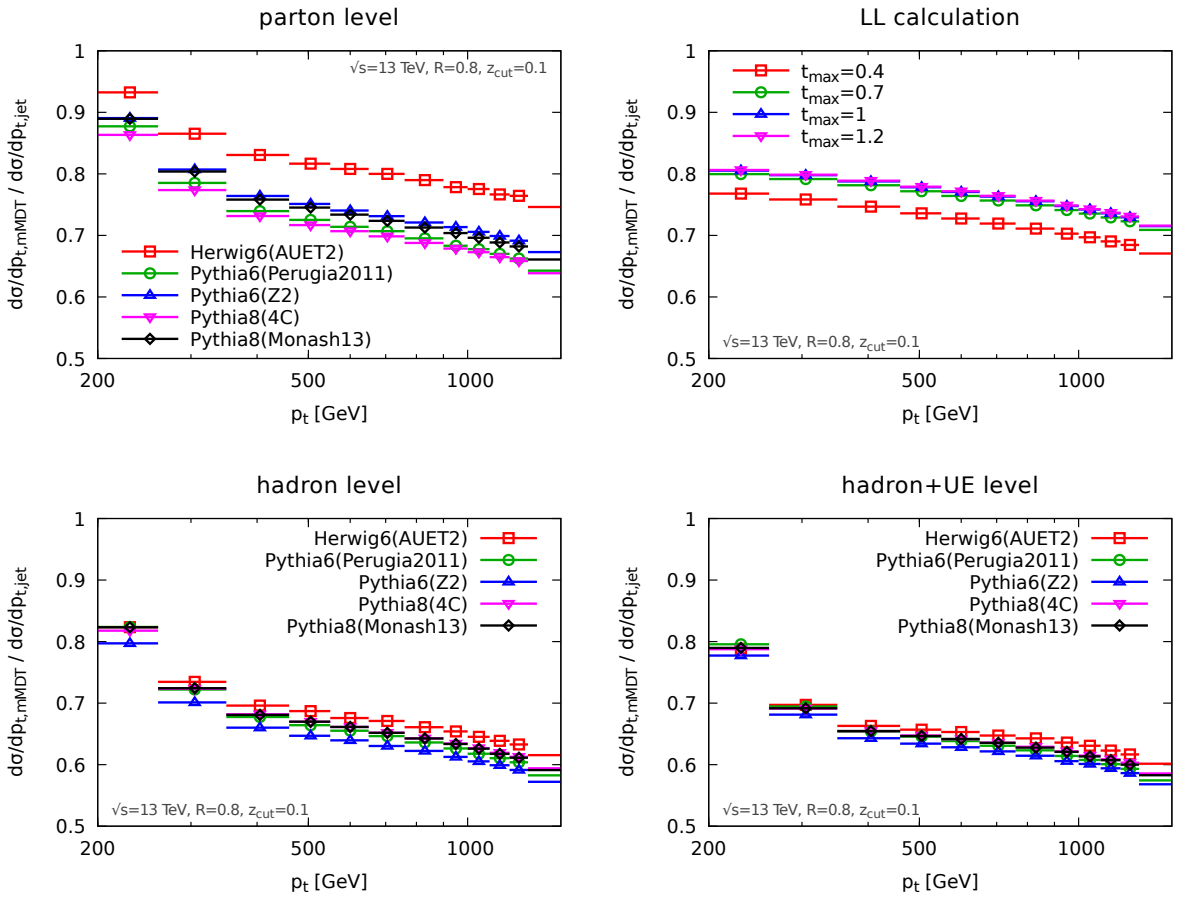


Figure D.3 — Ratio of the jet cross-section $d\sigma/dp_{t,\mathrm{mMDT}}$, binned in the groomed jet p_t, to the standard jet cross-section $d\sigma/dp_t$. The results of Monte Carlo simulations performed with different generators and tunes are shown in the top-left, bottom-left and bottom-right plots, respectively for simulations at parton-level, hadron-level without UE, and hadron-level including UE. The top-right plot instead shows our LL analytic results.

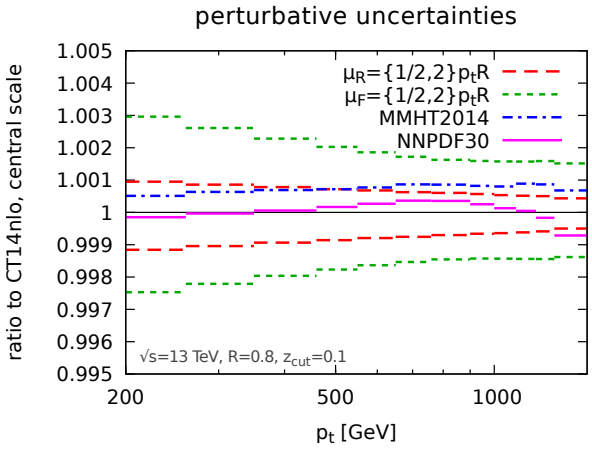


Figure D.4 — Theoretical uncertainties on the ratio $(d\sigma/dp_{t,\mathrm{mMDT}})/(d\sigma/dp_t)$. Uncertainties associated with the choice of the renormalisation and factorisation scales as well as with the choice of PDF are shown relative to the ratio obtained for the central scale choice and our default CT14nlo PDF set.

measured jet mass acts as a regulator of the collinear divergence. In this appendix, we therefore briefly depart from our study of the double-differential mass distribution to concentrate instead on the Sudakov-safe $d\sigma/dp_{t,\mathrm{mMDT}}$.

The results of both our LL calculation and of Monte Carlo simulations at different levels are presented in fig. D.3, for the ratio $(d\sigma/dp_{t,\mathrm{mMDT}})/(d\sigma/dp_t)$. We can make two main observations: firstly, our LL calculation provides a good description of what is observed at parton level. Secondly, as already noticed in fig. 9.6, hadronisation effects are sizeable while UE correction are more modest. Additionally, fig. D.3 shows the dependence of our LL calculation when varying the value $t_{\max}$ of t at which we stop parton branchings. For all the results presented in the main body of the paper, we have adopted $t_{\max} = 1.2$ which shows stable results in fig. D.3.

From a theoretical viewpoint, $d\sigma/dp_{t,\mathrm{mMDT}}$ can be viewed as the convolution of the jet spectrum $d^2\sigma/dp_t$ with the "jet energy drop", $1/\sigma\, d\sigma/d\Delta_E$ distribution, computed in the original Soft Drop paper [17] at LL accuracy in Δ_E, neglecting finite z_{cut} corrections. For the specific case of mMDT, i.e. the limit $\beta \to 0$ of Soft Drop, we found the remarkable property that, modulo running-coupling corrections, the energy-drop spectrum is independent of α_s and of the flavor of the parton initiating the jet.[2] It is therefore interesting to study the theoretical uncertainty of our LL calculation of $d\sigma/dp_{t,\mathrm{mMDT}}$, as measured from scale variation. This is shown in fig. D.4. The observed theoretical uncertainty is indeed very small, well below 1%. This should be contrasted with the much larger spread of the parton-level results from our Monte Carlo simulations, the top-left panel of fig. D.3. This could be related to subleading effects not captured by scale variation, or to effects of finite shower cut-off, seen also in our LL calculation when varying $t_{\max}$. The question

[2] See eq. (5.9) of Ref. [17].

of the power corrections to the $p_{t,\mathrm{mMDT}}$ cross-section, and to Sudakov-safe observables in general, is therefore interesting both from the point of view of Monte Carlo simulations and all-order calculations.

Bibliography

[1] M. Mangano, "Physics at the FCC-hh, a 100 TeV pp collider," 2017.

[2] E. Bothmann, P. Ferrarese, F. Krauss, S. Kuttimalai, S. Schumann, and J. Thompson, "Aspects of perturbative QCD at a 100 TeV future hadron collider," *Phys. Rev.*, vol. D94, no. 3, p. 034007, 2016.

[3] N. Arkani-Hamed, T. Han, M. Mangano, and L.-T. Wang, "Physics opportunities of a 100 TeV proton–proton collider," *Phys. Rept.*, vol. 652, pp. 1–49, 2016.

[4] G. Kasieczka, T. Plehn, T. Schell, T. Strebler, and G. P. Salam, "Resonance Searches with an Updated Top Tagger," *JHEP*, vol. 06, p. 203, 2015.

[5] C. Anders, C. Bernaciak, G. Kasieczka, T. Plehn, and T. Schell, "Benchmarking an even better top tagger algorithm," *Phys. Rev.*, vol. D89, no. 7, p. 074047, 2014.

[6] S. Schaetzel and M. Spannowsky, "Tagging highly boosted top quarks," *Phys. Rev.*, vol. D89, no. 1, p. 014007, 2014.

[7] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, "Stop Reconstruction with Tagged Tops," *JHEP*, vol. 10, p. 078, 2010.

[8] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, "Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks," *Phys. Rev. Lett.*, vol. 101, p. 142001, 2008.

[9] G. Aad *et al.*, "Light-quark and gluon jet discrimination in *pp* collisions at $\sqrt{s} =$ 7 TeV with the ATLAS detector," *Eur. Phys. J.*, vol. C74, no. 8, p. 3023, 2014.

[10] J. Gallicchio and M. D. Schwartz, "Quark and Gluon Jet Substructure," *JHEP*, vol. 04, p. 090, 2013.

[11] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC," *Phys. Rev. Lett.*, vol. 100, p. 242001, 2008.

[12] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure," *JHEP*, vol. 09, p. 029, 2013.

[13] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, "New clustering algorithm for multi - jet cross-sections in e+ e- annihilation," *Phys. Lett.*, vol. B269, pp. 432–438, 1991.

[14] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys.*, vol. B406, pp. 187–224, 1993.

[15] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys. Rev.*, vol. D48, pp. 3160–3166, 1993.

[16] D. Krohn, J. Thaler, and L.-T. Wang, "Jet Trimming," *JHEP*, vol. 02, p. 084, 2010.

[17] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop," *JHEP*, vol. 05, p. 146, 2014.

[18] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness," *JHEP*, vol. 03, p. 015, 2011.

[19] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "N-Jettiness: An Inclusive Event Shape to Veto Jets," *Phys. Rev. Lett.*, vol. 105, p. 092002, 2010.

[20] A. J. Larkoski, G. P. Salam, and J. Thaler, "Energy Correlation Functions for Jet Substructure," *JHEP*, vol. 06, p. 108, 2013.

[21] I. Moult, L. Necib, and J. Thaler, "New Angles on Energy Correlation Functions," *JHEP*, vol. 12, p. 153, 2016.

[22] "Performance of Boosted W Boson Identification with the ATLAS Detector," Tech. Rep. ATL-PHYS-PUB-2014-004, CERN, Geneva, Mar 2014.

[23] C. Collaboration, "Jet algorithms performance in 13 TeV data," 2017.

[24] T. A. collaboration, "Jet mass reconstruction with the ATLAS Detector in early Run 2 data," 2016.

[25] T. A. collaboration, "Discrimination of Light Quark and Gluon Jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector," 2016.

[26] G. Aad *et al.*, "Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector," *Phys. Rev.*, vol. D93, no. 3, p. 032009, 2016.

[27] V. Khachatryan *et al.*, "Measurement of the integrated and differential $t\bar{t}$ production cross sections for high-p_t top quarks in pp collisions at $\sqrt{s} = 8$ TeV," *Phys. Rev.*, vol. D94, no. 7, p. 072002, 2016.

[28] M. Aaboud *et al.*, "Search for anomalous electroweak production of WW/WZ in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys. Rev.*, vol. D95, no. 3, p. 032001, 2017.

[29] A. M. Sirunyan *et al.*, "Search for anomalous couplings in boosted WW/WZ $\to \ell\nu\mathrm{q}\bar{\mathrm{q}}$ production in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Phys. Lett.*, vol. B772, pp. 21–42, 2017.

[30] T. A. collaboration, "Boosted Higgs ($\to b\bar{b}$) Boson Identification with the ATLAS Detector at $\sqrt{s} = 13$ TeV," 2016.

[31] "Search for $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment," Tech. Rep. CMS-PAS-HIG-16-004, CERN, Geneva, 2016.

[32] C. Collaboration, "Splitting function in pp and PbPb collisions at 5.02 TeV," 2016.

[33] M. Aaboud *et al.*, "Searches for heavy diboson resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 09, p. 173, 2016.

[34] C. Collaboration, "Search for massive resonances decaying into pairs of boosted W and Z bosons at $\sqrt{s} = 13$ TeV," 2015.

[35] T. Sjostrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.178:852-867,2008*, Oct. 2007.

[36] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.*, vol. C58, pp. 639–707, 2008.

[37] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," *JHEP*, vol. 02, p. 007, 2009.

[38] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," *Phys. Lett.*, vol. B641, pp. 57–61, 2006.

[39] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur. Phys. J.*, vol. C72, p. 1896, 2012.

[40] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, "An Effective field theory for collinear and soft gluons: Heavy to light decays," *Phys. Rev.*, vol. D63, p. 114020, 2001.

[41] C. W. Bauer and I. W. Stewart, "Invariant operators in collinear effective theory," *Phys. Lett.*, vol. B516, pp. 134–142, 2001.

[42] C. W. Bauer, D. Pirjol, and I. W. Stewart, "Soft collinear factorization in effective field theory," *Phys. Rev.*, vol. D65, p. 054022, 2002.

[43] C. W. Bauer, F. J. Tackmann, J. R. Walsh, and S. Zuberi, "Factorization and Resummation for Dijet Invariant Mass Spectra," *Phys. Rev.*, vol. D85, p. 074006, 2012.

[44] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, "Jet-images — deep learning edition," *JHEP*, vol. 07, p. 069, 2016.

[45] T. Lapsien, R. Kogler, and J. Haller, "A new tagger for hadronically decaying heavy particles at the LHC," *Eur. Phys. J.*, vol. C76, no. 11, p. 600, 2016.

[46] P. Baldi, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson, "Jet Substructure Classification in High-Energy Physics with Deep Neural Networks," *Phys. Rev.*, vol. D93, no. 9, p. 094034, 2016.

[47] "Measurement of the differential jet production cross section with respect to jet mass and transverse momentum in dijet events from pp collisions at $\sqrt{s} = 13$ TeV," Tech. Rep. CMS-PAS-SMP-16-010, CERN, Geneva, 2017.

[48] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory; 1995 ed.* Boulder, CO: Westview, 1995. Includes exercises.

[49] R. K. Ellis, *QCD and collider physics.* Cambridge: Cambridge University Press, 1996.

[50] Q. Kim, *Elementary particles and their interactions : concepts and phenomena.* Berlin New York: Springer, 1998.

[51] J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys. Rev.*, vol. D86, p. 010001, 2012.

[52] G. P. Salam, "Elements of QCD for hadron colliders," in *High-energy physics. Proceedings, 17th European School, ESHEP 2009, Bautzen, Germany, June 14-27, 2009*, 2010.

[53] K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.*, vol. C38, p. 090001, 2014.

[54] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, "The Four loop beta function in quantum chromodynamics," *Phys. Lett.*, vol. B400, pp. 379–384, 1997.

[55] M. Czakon, "The Four-loop QCD beta-function and anomalous dimensions," *Nucl. Phys.*, vol. B710, pp. 485–498, 2005.

[56] A. J. Larkoski and J. Thaler, "Unsafe but Calculable: Ratios of Angularities in Perturbative QCD," *JHEP*, vol. 09, p. 137, 2013.

[57] A. J. Larkoski, S. Marzani, and J. Thaler, "Sudakov Safety in Perturbative QCD," *Phys. Rev.*, vol. D91, no. 11, p. 111501, 2015.

[58] J. E. Huth *et al.*, "Toward a standardization of jet definitions," in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90) Snowmass, Colorado, June 25-July 13, 1990*, pp. 0134–136, 1990.

[59] D. Bertolini, T. Chan, and J. Thaler, "Jet Observables Without Jet Algorithms," *JHEP*, vol. 1404, p. 013, 2014.

[60] A. J. Larkoski and J. Thaler, "Aspects of jets at 100 TeV," *Phys. Rev.*, vol. D90, no. 3, p. 034010, 2014.

[61] S. Bethke *et al.*, "Experimental Investigation of the Energy Dependence of the Strong Coupling Strength," *Phys. Lett.*, vol. B213, pp. 235–241, 1988.

[62] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[63] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms," *JHEP*, vol. 08, p. 001, 1997.

[64] G. F. Sterman and S. Weinberg, "Jets from Quantum Chromodynamics," *Phys. Rev. Lett.*, vol. 39, p. 1436, 1977.

[65] G. P. Salam, "Towards Jetography," *Eur. Phys. J.*, vol. C67, pp. 637–686, 2010.

[66] G. C. Blazey *et al.*, "Run II jet physics," in *QCD and weak boson physics in Run II. Proceedings, Batavia, USA, March 4-6, June 3-4, November 4-6, 1999*, pp. 47–77, 2000.

[67] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone jet algorithm," *JHEP*, vol. 05, p. 086, 2007.

[68] G. P. Salam, L. Schunk, and G. Soyez, "Dichroic subjettiness ratios to distinguish colour flows in boosted boson tagging," *JHEP*, vol. 03, p. 022, 2017.

[69] M. Dasgupta, L. Schunk, and G. Soyez, "Jet shapes for boosted jet two-prong decays from first-principles," *JHEP*, vol. 04, p. 166, 2016.

[70] A. J. Larkoski, I. Moult, and D. Neill, "Power Counting to Better Jet Observables," *JHEP*, vol. 12, p. 009, 2014.

[71] B. Andersson, G. Gustafson, L. Lonnblad, and U. Pettersson, "Coherence Effects in Deep Inelastic Scattering," *Z. Phys.*, vol. C43, p. 625, 1989.

[72] A. Banfi, G. P. Salam, and G. Zanderighi, "Principles of general final-state resummation and automated implementation," *JHEP*, vol. 03, p. 073, 2005.

[73] M. Dasgupta and G. P. Salam, "Resummation of nonglobal QCD observables," *Phys. Lett.*, vol. B512, pp. 323–330, 2001.

[74] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, "Precision physics with pile-up insensitive observables," 2016.

[75] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, "Factorization for groomed jet substructure beyond the next-to-leading logarithm," *JHEP*, vol. 07, p. 064, 2016.

[76] Y.-T. Chien, R. Kelley, M. D. Schwartz, and H. X. Zhu, "Resummation of Jet Mass at Hadron Colliders," *Phys.Rev.*, vol. D87, p. 014010, 2013.

[77] Y.-T. Chien and M. D. Schwartz, "Resummation of heavy jet mass and comparison to LEP data," *JHEP*, vol. 08, p. 058, 2010.

[78] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, "Resummation of large logarithms in e+ e- event shape distributions," *Nucl. Phys.*, vol. B407, pp. 3–42, 1993.

[79] Y. L. Dokshitzer, V. A. Khoze, and S. I. Troian, "Specific features of heavy quark production. LPHD approach to heavy particle spectra," *Phys. Rev.*, vol. D53, pp. 89–119, 1996.

[80] S. Catani, B. R. Webber, and G. Marchesini, "QCD coherent branching and semi-inclusive processes at large x," *Nucl. Phys.*, vol. B349, pp. 635–654, 1991.

[81] M. Dasgupta, A. Powling, and A. Siodmok, "On jet substructure methods for signal jets," *JHEP*, vol. 08, p. 079, 2015.

[82] R. Corke and T. Sjostrand, "Interleaved Parton Showers and Tuning Prospects," *JHEP*, vol. 03, p. 032, 2011.

[83] "Fastjet contrib." http://fastjet.hepforge.org/contrib.

[84] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, "On jet mass distributions in Z+jet and dijet processes at the LHC," *JHEP*, vol. 10, p. 126, 2012.

[85] S. Catani, G. Turnock, and B. R. Webber, "Heavy jet mass distribution in e+ e- annihilation," *Phys. Lett.*, vol. B272, pp. 368–372, 1991.

[86] S. Catani and M. H. Seymour, "A General algorithm for calculating jet cross-sections in NLO QCD," *Nucl. Phys.*, vol. B485, pp. 291–419, 1997. [Erratum: Nucl. Phys.B510,503(1998)].

[87] M. Seymour, "Event2." http://www.hep.manchester.ac.uk/u/seymour/nlo/.

[88] R. W. L. Jones, M. Ford, G. P. Salam, H. Stenzel, and D. Wicke, "Theoretical uncertainties on alpha(s) from event shape variables in e+ e- annihilations," *JHEP*, vol. 12, p. 007, 2003.

[89] J. Gallicchio and M. D. Schwartz, "Quark and Gluon Tagging at the LHC," *Phys. Rev. Lett.*, vol. 107, p. 172001, 2011.

[90] A. J. Larkoski, J. Thaler, and W. J. Waalewijn, "Gaining (Mutual) Information about Quark/Gluon Discrimination," *JHEP*, vol. 11, p. 129, 2014.

[91] J. R. Andersen *et al.*, "Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report," in *9th Les Houches Workshop on Physics at TeV Colliders (PhysTeV 2015) Les Houches, France, June 1-19, 2015*, 2016.

[92] A. Altheimer *et al.*, "Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd-27th of July 2012," *Eur. Phys. J.*, vol. C74, no. 3, p. 2792, 2014.

[93] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys.Lett.*, vol. B659, pp. 119–126, 2008.

[94] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets," *JHEP*, vol. 0804, p. 005, 2008.

[95] "Tagging and suppression of pileup jets," Tech. Rep. ATL-PHYS-PUB-2014-001, CERN, Geneva, Jan 2014.

[96] M. Cacciari, G. P. Salam, and G. Soyez, "Use of charged-track information to subtract neutral pileup," *Phys. Rev.*, vol. D92, no. 1, p. 014003, 2015.

[97] D. Krohn, M. D. Schwartz, M. Low, and L.-T. Wang, "Jet Cleansing: Pileup Removal at High Luminosity," *Phys. Rev.*, vol. D90, no. 6, p. 065020, 2014.

[98] P. Berta, M. Spousta, D. W. Miller, and R. Leitner, "Particle-level pileup subtraction for jets and jet shapes," *JHEP*, vol. 06, p. 092, 2014.

[99] M. Cacciari, G. P. Salam, and G. Soyez, "SoftKiller, a particle-level pileup removal method," *Eur. Phys. J.*, vol. C75, no. 2, p. 59, 2015.

[100] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup Per Particle Identification," *JHEP*, vol. 10, p. 059, 2014.

[101] J. F. Gunion and Z. Kunszt, "Lepton Correlations in Gauge Boson Pair Production and Decay," *Phys. Rev.*, vol. D33, p. 665, 1986.

[102] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP*, vol. 11, p. 040, 2004.

[103] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[104] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010.

[105] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[106] D. de Florian, M. Grazzini, C. Hanga, S. Kallweit, J. M. Lindert, P. Maierhöfer, J. Mazzitelli, and D. Rathlev, "Differential Higgs Boson Pair Production at Next-to-Next-to-Leading Order in QCD," *JHEP*, vol. 09, p. 151, 2016.

[107] J. M. Campbell and R. K. Ellis, "An Update on vector boson pair production at hadron colliders," *Phys. Rev.*, vol. D60, p. 113006, 1999.

[108] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC," *JHEP*, vol. 07, p. 018, 2011.

[109] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, and C. Williams, "Color singlet production at NNLO in MCFM," *Eur. Phys. J.*, vol. C77, no. 1, p. 7, 2017.

[110] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, "Jet Mass Spectra in Higgs + One Jet at NNLL," *Phys. Rev. D*, vol. 88, p. 054031, 2013.

[111] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, "The t anti-t cross-section at 1.8-TeV and 1.96-TeV: A Study of the systematics due to parton densities and scale dependence," *JHEP*, vol. 04, p. 068, 2004.

[112] Z. Nagy, "Next-to-leading order calculation of three jet observables in hadron hadron collision," *Phys. Rev.*, vol. D68, p. 094002, 2003.

[113] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.*, vol. D93, no. 3, p. 033006, 2016.

[114] S. Catani and M. H. Seymour, "The Dipole Formalism for the Calculation of QCD Jet Cross Sections at Next-to-Leading Order," *Phys. Lett.*, vol. B378, pp. 287–301, 1996.

[115] "New ATLAS event generator tunes to 2010 data," Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011.

[116] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)," *JHEP*, vol. 01, p. 010, 2001.

[117] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, "HERWIG 6.5 release note," 2002.

[118] R. Field, "Early LHC Underlying Event Data - Findings and Surprises," in *Hadron collider physics. Proceedings, 22nd Conference, HCP 2010, Toronto, Canada, August 23-27, 2010*, 2010.

[119] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes," *Phys. Rev.*, vol. D82, p. 074018, 2010.

[120] B. Cooper, J. Katzy, M. L. Mangano, A. Messina, L. Mijovic, and P. Skands, "Importance of a consistent choice of alpha(s) in the matching of AlpGen and Pythia," *Eur. Phys. J.*, vol. C72, p. 2078, 2012.

[121] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.

[122] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J.*, vol. C74, no. 8, p. 3024, 2014.

[123] M. Dasgupta, L. Magnea, and G. P. Salam, "Non-perturbative QCD effects in jets at hadron colliders," *JHEP*, vol. 02, p. 055, 2008.

[124] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, "Jet substructure with analytical methods," *Eur. Phys. J.*, vol. C73, no. 11, p. 2623, 2013.

[125] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and collider physics," *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 8, pp. 1–435, 1996.

[126] Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian, *Basics of perturbative QCD*. 1991.

[127] M. Dasgupta, F. Dreyer, G. P. Salam, and G. Soyez, "Small-radius jets to all orders in QCD," *JHEP*, vol. 04, p. 039, 2015.

[128] M. Dasgupta, F. A. Dreyer, G. P. Salam, and G. Soyez, "Inclusive jet spectrum for small-radius jets," *JHEP*, vol. 06, p. 057, 2016.

[129] P. Gras, S. Hoeche, D. Kar, A. Larkoski, L. Lönnblad, S. Plätzer, S. Prestel, A. Siódmok, P. Skands, G. Soyez, and J. Thaler, "*Systematics of quark/gluon tagging*." In preparation.

[130] A. J. Larkoski, I. Moult, and D. Neill, "Analytic Boosted Boson Discrimination," *JHEP*, vol. 05, p. 117, 2016.

[131] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, "Thrust at N^3LL with Power Corrections and a Precision Global Fit for alphas(mZ)," *Phys. Rev.*, vol. D83, p. 074021, 2011.

[132] L. Schunk, S. Marzani, and G. Soyez, "A study of jet mass distributions with grooming." Talk given at BOOST2017 Workshop, https://indico.cern.ch/event/579660/contributions/2582133/.

[133] J. Dolen, P. Harris, S. Marzani, S. Rappoccio, and N. Tran, "Thinking outside the ROCs: Designing Decorrelated Taggers (DDT) for jet substructure," *JHEP*, vol. 05, p. 156, 2016.

[134] G. Soyez, G. P. Salam, J. Kim, S. Dutta, and M. Cacciari, "Pileup subtraction for jet shapes," *Phys.Rev.Lett.*, vol. 110, no. 16, p. 162001, 2013.

université PARIS-SACLAY | ÉCOLE DOCTORALE Physique en Île-de-France (EDPIF)

Titre : Comprendre le sous-structure des jets au LHC

Mots clés : jets, sous-structure, physique de particules, LHC, phenomenologie

Résumé : Dans cette thèse on étudie les techniques de sous-structure des jets, utilisées pour explorer la dynamique interne des jets dans les régimes boostés (jets avec une impulsion transverse beaucoup plus grande que leur masse). On se concentre sur les techniques pour les jets à deux cœurs, pour identifier les bosons W/Z/H boostés. On propose une approche analytique, utilisant des techniques de resommation à tous les ordres en QCD perturbative. Dans la première partie de ce document, on présente les idées basiques concernant la resommation et on introduit les ingrédients (basic building blocks) utilisés dans nos calculs.

Notre première étude explore le Y-splitter tagger et comment sa performance est affectée par la combinaison avec une variété de techniques de grooming : le MassDrop Tagger (mMDT), trimming et SoftDrop. Selon des études Monte Carlo, cette combinaison augmente la performance du Y-splitter, on étudie l'origine de ce comportement avec des calculs théoriques. On explore aussi l'impact des effets non-perturbatives et propose des variantes améliorées de la méthode Y-splitter originale.

Ensuite, on étudie l'utilisation des jet shapes comment une variable discriminante entre les désintégrations hadroniques à deux cœurs des bosons électrofaibles et le bruit de fond des jets QCD. On considère trois shapes couramment utilisées : N-subjettiness, energy correlation functions et le paramètre MassDrop. On calcule analytiquement les efficacités pour des jets QCD et signal avec une coupure sur la variable jet shape. On compare également nos résultats aux générateurs de Monte Carlo et on étudie l'impact des effets non-perturbatifs.

Ensuite, on montre comment le savoir-faire accumulé dans les études antérieures peut être utilisé pour explorer la combinaison des techniques de prong-finder/grooming avec le N-subjettiness. On utilise le rapport $\tau_2 \setminus \tau_1$ comment une variable discriminante pour les jets à deux cœurs. On propose le rapport dichroïque de N-subjettiness, où on utilise un gros jet (avec ou sans pre-grooming) pour calculer τ_2 et un jet plus petit, avec un prong finder pour τ_1 . Cette version donne une performance améliorée par rapport aux versions utilisées actuellement par les expériences, tout en maintenant les effets non-perturbatifs sous contrôle.

Enfin, on effectue une étude phénoménologique de la distribution de masse des jets avec mMDT. Nos prédictions théoriques prennent en compte les logarithmes dominants du rapport de la masse de jet sur l'impulsion transverse et on fait le « matching » avec les éléments de matrice à ordre fixe calculés au NLO. On discute deux options possibles, selon que les distributions sont mesurées dans des bins de l'impulsion transverse avant (version préférée) ou après le mMDT (version collinear unsafe). Nos prédictions reproduisent des mesures faites récemment par la collaboration CMS.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

université PARIS-SACLAY | ÉCOLE DOCTORALE Physique en Île-de-France (EDPIF)

Title : Understanding jet substructure at the LHC

Keywords : jets, substructure, particle physics, LHC, phenomenology

Abstract : In this thesis we study jet substructure techniques, used to explore the internal dynamics of jets in boosted regimes (jets with transverse momentum much larger than their mass). We focus on techniques for two-pronged jets, meant to identify boosted W/Z/H bosons. We propose an analytical approach using all-order resummation techniques, in perturbative QCD. In the beginning of this document, we lay down the basic ideas of resummation and introduce the ingredients (basic building blocks) used for our calculations.

Our first study explores the Y-splitter tagger and how its performance is affected by combining it to different grooming techniques : the modified MassDrop Tagger (mMDT), trimming and SoftDrop. It is known that this combination increases the Y-splitter performance, and we studied the origin of this behavior from a first principle approach. We also explore the impact of non-perturbative effects and propose some variations for the original Y-splitter.

Then, we investigate the use jet shapes as discriminating variables between two-pronged hadronic decays of electroweak bosons (W/Z/H) and QCD jets background. We study three shapes: N-subjettiness, energy correlation functions and MassDrop parameter. We carry out analytical calculations for the efficiencies of signal and QCD jets with cuts on these variables. We also compare our results to Monte Carlo generators and study the impact of non-perturbative effects.

Next, we show how the knowladge accumulated in the previous studies can be used to explore the interplay between grooming/tagging techniques and the N-subjettiness. We use the ratio $\tau_2 \setminus \tau_1$ as a discriminating variable for two-pronged jets. In this work, we propose the dichroic N-subjettiness ratio, where we use a large jet for calculating τ_2 and a smaller, tagged subjet for τ_1 . The resulting dichroic ratio gives enhanced performance compared to the original version of the jet shape, while keeping non-perturbative effect under control.

Finally, we perform a phenomenological study of the jet mass distribution with mMDT. Our theoretical predictions account for the resummation of the leading logarithm of the ratio of the jet mass over the jet transverse momentum and are matched to fixed-order matrix elements computed at next-to-leading order accuracy. We consider both the jet transverse momentum measured before (preferred) and after (not collinear safe) the mMDT procedure. Our predictions reproduce the recent measurement by the CMS collaboration.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France