



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

**„Top Quark Mass Calibration for Monte Carlo Event  
Generators“**

verfasst von / submitted by

Lizhi Oliver Jin, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 066 876

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Physik UG2002

Betreut von / Supervisor:

Univ.-Prof. Dr. André H. Hoang

## Abstract

We tested and improved a top quark mass calibration framework for Monte Carlo (MC) event generators, that was originally used to calibrate the top mass parameter  $m_{\text{PYTHIA}}^{\text{MC}}$  of PYTHIA, using hadron level  $e^+e^- \rightarrow t\bar{t}$  2-Jettiness predictions, to field theoretically well-defined renormalized Lagrangian masses, the MSR mass  $m^{\text{MSR}}(R)$  and the pole mass  $m^{\text{pole}}$ . The current most precise top mass measurements specifically determine  $m^{\text{MC}}$ , even though its relation to the masses in the Standard Model Lagrangian is in general not fully understood. The theoretical predictions employ two effective field theories, SCET and bHQET, at  $\text{N}^2\text{LL} + \text{NLO}$  accuracy. The  $\mathcal{O}(\Lambda_{\text{QCD}})$  pole mass and soft function renormalon ambiguities are canceled by using the MSR mass and the R-gap scheme respectively. In this work we added two additional observables, *sum of jet mass* and *modified jet mass*, two additional gap schemes and two other major MC generators, HERWIG and SHERPA. We achieved universal results for these observables and gap schemes, by using a more reliable nonperturbative shape function parametrization and by including the leading kinematic mass-dependent power corrections in the bHQET measurement function. Remaining power corrections were reduced and estimated by parametrizing partial absorption of power corrections from the nonsingular into the singular distribution. Our calibration shows that the MC masses using standard tunes and the MSR mass at  $R = 1 \text{ GeV}$  agree with each other within uncertainties of 200 MeV, but the nonperturbative parameters for our observables are measured to be vastly different for each MC.

## Zusammenfassung

Wir haben ein Top-Quark-Massen-Kalibrierungs-Framework, basierend auf theoretischen  $e^+e^- \rightarrow t\bar{t}$  2-Jettiness Berechnungen für hadronisierte Endzustände, für Monte-Carlo-Event-Generatoren (MC) getestet und verbessert. Dieses System wurde ursprünglich verwendet, um die Beziehungen zwischen dem Top-Massen-Parameter  $m_{\text{PYTHIA}}^{\text{MC}}$  von PYTHIA und klar definierten renormierten Lagrangian-Massen, der MSR-Masse  $m^{\text{MSR}}(R)$  und der Polmasse  $m^{\text{pole}}$ , zu bestimmen. Die zurzeit genauesten Messungen der Top-Masse bestimmen spezifisch  $m^{\text{MC}}$ , obwohl die Relation zwischen  $m^{\text{MC}}$  und feldtheoretisch renormierten Massen-Schemata im Standardmodell-Lagrangian im Allgemeinen nicht vollständig verstanden ist. Die theoretischen Voraussagen verwenden zwei effektive Feldtheorien, SCET und bHQET, mit  $\text{N}^2\text{LL} + \text{NLO}$  Genauigkeit. Die  $\mathcal{O}(\Lambda_{\text{QCD}})$  Polmasse- und Soft-Funktion-Renormalon-Ambiguitäten werden mit der MSR-Masse und dem R-Gap-Schema entfernt. In dieser Arbeit haben wir zwei zusätzliche Observablen, *sum of jet mass* und *modified jet mass*, zwei zusätzliche Gap-Schemata und zwei andere bedeutende MC Generatoren, HERWIG und SHERPA, implementiert. Nach der Einführung einer verlässlicheren nichtperturbativen Shape-Funktion-Parametrisierung und der Inklusion von führenden kinematischen massen-abhängigen Power-Korrekturen in der bHQET Messfunktion erhielten wir universelle Resultate für diese Observablen und Gap-Schemata. Die restlichen Power-Korrekturen wurden durch eine Parametrisierung von partieller Absorption von Power-Korrekturen in die singuläre Verteilung reduziert und abgeschätzt. Unsere Kalibrierungs-Resultate zeigen, dass die MC-Massen für die Standard-Tunes und die MSR-Masse bei  $R = 1 \text{ GeV}$  innerhalb der Unsicherheiten von 200 MeV miteinander übereinstimmen. Bei den gemessenen nichtperturbativen Parametern für unsere Observablen wurden jedoch große Unterschiede zwischen den MC Generatoren festgestellt.

## Acknowledgements

I am extremely grateful to my supervisor Prof. André Hoang for giving me the opportunity to work on this very interesting project in the particle physics group at the University of Vienna. His invaluable feedback and guidance were key for the success of this project and his deep understanding and insights about the intricacies of this subject allowed me to effectively analyze and solve all problems that we encountered along the way. These experiences were very educational for me and they have given me a clear view about the path that I want to pursue in the future.

This endeavor would also not have been possible without Vicent Mateu, who was involved in all aspects of this project. He provided us with the original code and an improved version of the calibration framework. He also introduced me to programming in PYTHON, which clearly improved my ability to build an effective data processing pipeline and which therefore saved me and any person who will interact with this framework a lot of time. I want to thank him for taking the time to attend our regular meetings, answering questions about the details of the original calibration setup, checking some of the new calculations, and being involved in all the decisions that we made.

Next, I want to thank Prof. Simon Plätzer, Daniel Samitz and Moritz Preisser. Their deep insight into the current state of Monte Carlo event generators (MC) and the standardized tools that are available to interface with these and the code examples that they provided were instrumental in building an unified calibration and analysis framework for any MC generator. The former two also provided us with a custom HERWIG build, which fixed a problem that we encountered in the official version.

Moreover, I would like to extend my thanks to all members of the particle physics group, especially my colleagues and friends Michael Adam, Christoph Bartsch, Miguel Angel Benitez-Rathgeb, Sergio Leal Gomez, Daniel Lechner, Christopher Lepenik, Jan Lüdtke, Peter Majcen, Christoph Regner, Ines Ruffa, Daniel Samitz and Angelika Widl. The discussions with them helped me to better understand the subjects that were worked on in this group and also physics in general. It was also a pleasure to work in a relaxed, friendly and cooperative atmosphere and I enjoyed the fun conversations that we had during lunch breaks and other meetups.

Finally, I want to thank my family and especially my parents who supported me throughout this journey and who also gave me the correct advice when it was needed. This work would not have been possible without them.

# Contents

<b>1. Introduction</b>	<b>6</b>
<b>2. Resummed Cross Section</b>	<b>9</b>
2.1. Factorization Theorem in the Peak Region . . . . .	9
2.2. Renormalon Subtractions . . . . .	13
2.2.1. Mass schemes . . . . .	13
2.2.2. Soft Gap subtraction schemes . . . . .	14
<b>3. Nonsingular Corrections</b>	<b>17</b>
3.1. QCD and SCET Nonsingular Distributions . . . . .	17
3.2. QCD Massive Event-Shape Distribution . . . . .	19
3.2.1. Born Cross Section . . . . .	20
3.2.2. Phase Space and Kinematic Variables . . . . .	20
3.2.3. Observable Definitions and Properties . . . . .	21
3.2.4. Delta- and Plus-Function Coefficients . . . . .	24
3.2.5. Radiative Tail . . . . .	25
3.3. Absorption of Nonsingular Coefficients . . . . .	28
<b>4. Full Event Shape Distribution</b>	<b>30</b>
4.1. Combining Ingredients . . . . .	30
4.2. Profile Functions . . . . .	31
<b>5. Fit Procedure Improvements</b>	<b>34</b>
5.1. Fixed $\Delta_0$ Fits . . . . .	34
5.1.1. Fit Procedure . . . . .	34
5.1.2. Previous Calibration Results for PYTHIA . . . . .	37
5.1.3. Gap Dependence . . . . .	39
5.1.4. $\Delta_0$ -Dependence . . . . .	40
5.1.5. $\lambda$ -Dependence . . . . .	40
5.2. Variable $\Delta_0$ Fits . . . . .	41
5.2.1. Fit Procedure Modification . . . . .	41
5.2.2. Choosing suitable $\lambda$ for each Event-Generator . . . . .	44
5.2.3. Results of improved Fit Procedure . . . . .	49
5.2.4. First Test of Universality . . . . .	50
<b>6. Kinematic Mass Power Corrections</b>	<b>51</b>
6.1. Measurement Function . . . . .	51
6.2. Rescale Factors for Event Shapes . . . . .	52

6.3. Factorization Theorem Modifications . . . . .	55
6.4. SCET Nonsingular Modifications . . . . .	56
6.5. QCD Nonsingular Modifications . . . . .	56
<b>7. Results</b>	<b>58</b>
7.1. Analysis of Power Corrections . . . . .	58
7.1.1. Fits: Rescale Prescription . . . . .	58
7.1.2. Partial Absorption of Power Corrections . . . . .	60
7.1.3. Fits: Partial Absorption . . . . .	65
7.2. Final Results . . . . .	70
<b>8. Conclusions</b>	<b>74</b>
<b>A. Evolutions</b>	<b>76</b>
A.1. Evolution Factors and Anomalous Dimensions . . . . .	76
A.2. R-Evolution . . . . .	78
<b>B. Distributions</b>	<b>80</b>
<b>C. Additional Analysis</b>	<b>81</b>
C.1. Comparison: $R_s < \mu_S$ and $R_s = \mu_S$ . . . . .	81

# 1. Introduction

The top quark mass parameter  $m_t$  is one of most important parameters of the Standard Model (SM). Due to its large size, it plays an important role in many quantitative and conceptual aspects of the SM and beyond. The most precise determinations are obtained by so called "direct measurements" where the kinematics of the top mass decay products (jets and charged leptons) in  $t\bar{t}$  events are reconstructed and then compared to kinematic distributions obtained from Monte Carlo (MC) event generator simulations. These MCs are based on first principles QCD, but due to generality all parts of their ingredients, e.g. the parton shower and hadronization use approximations or modeling, can have systematic effects on the meaning of the top mass parameter  $m_t^{\text{MC}}$  of these MCs. The measured mass obtained by direct measurements is therefore specifically this MC mass parameter  $m_t^{\text{MC}}$ .

The current world average for direct measurements is  $m_t^{\text{MC}} = 172.69 \pm 0.30$  GeV [1], which uses, among others, the measurements  $m_t^{\text{MC}} = 172.44 \pm 0.48$  GeV (CMS combination) [2],  $m_t^{\text{MC}} = 172.69 \pm 0.48$  GeV (ATLAS combination) [3] and  $m_t^{\text{MC}} = 174.30 \pm 0.65$  GeV (Tevatron combination) [4]. Recently there has been a very precise measurement not yet included in the world average  $m_t^{\text{MC}} = 171.77 \pm 0.38$  GeV from CMS [5]. Future projections for the HL-LHC indicate, that uncertainties as small as 200 MeV for individual measurements can be reached [6].

There have also been measurements of the pole mass, extracted from inclusive or differential cross section measurements using theory prediction expressed in the pole mass scheme. These measurements however have a lower sensitivity to  $m_t$  than direct measurements. The current world average is  $m^{\text{pole}} = 172.5 \pm 0.7$  GeV [1].

The MC mass  $m_t^{\text{MC}}$  has been frequently identified with the pole mass, but the pole mass scheme is characterized by absorption of all self-energy corrections into its mass definition. These corrections contain an  $\mathcal{O}(\Lambda_{\text{QCD}})$  renormalon ambiguity, due to problematic infrared contributions at the hadronization scale  $\sim \Lambda_{\text{QCD}}$ . This ambiguity conceptually limits the precision to which the pole mass can be measured. The size of this ambiguity has been estimated by two studies to amount to 110 MeV [7] or 250 MeV [8].

However, the parton shower in a MC has to impose an infrared cutoff scale on its shower evolution. This introduces a resolution scale on the real and virtual radiation, which are left unresolved below this scale. These unresolved corrections will therefore combine and cancel the contributions that cause the renormalon ambiguity. These properties of the MC have led to the conjecture, that  $m_t^{\text{MC}}$  should be closely related to the MSR mass with its infrared subtraction scale  $R$ , which acts like a resolution scale, set to the shower cutoff scale  $Q_0$  [9, 10].

These uncertainties in the interpretation of the MC mass have led to a concerted effort of the community into devising methods to obtain a better theoretical and numerical

understanding of the relation between  $m_t^{\text{MC}}$  and any well defined renormalized Lagrangian QCD mass.

Recently there has been a theoretical determination of the HERWIG mass scheme in terms of the pole mass or the MSR mass, which quantified the corrections that arise from the shower cutoff dependence [11].

Numerical efforts of understanding this relation on the other hand rely on comparison of MC samples with state of the art hadron level QCD predictions for observables with strong mass sensitivity, at a sufficiently high perturbative order to be able to fully control the quark mass scheme dependence. Many of these strongly mass sensitive observables require the use of effective field theories (EFT) that provide factorization theorems to resum large logarithms of ratios of scales, which would otherwise spoil a fixed order (FO) analysis, to all orders of perturbation theory. The description of at least the dominant non-perturbative effects, used in these factorization theorems, can be consistently quantified in a field theoretical framework.

These predictions can not only be used as a diagnostic tool to improve MC simulations, but they also show promises for future experimental extractions of the top mass, once suitable experimental measurements of these observables are possible.

Some of these results relevant to our work are the factorized hadron-level event shape distribution for boosted top quark jets in the peak region for  $e^+e^-$  collisions [12, 13], where the resummation has recently been extended to N<sup>3</sup>LL (next-to<sup>3</sup> leading logarithmic order) [14], a automated method to obtain fixed order NLO QCD massive event-shape differential distributions [15], and for future comparison: boosted top jets with soft drop grooming in pp collisions [16].

This master thesis is an direct update of the work presented in [17]. They calibrated the PYTHIA MC top quark mass parameter using a N<sup>2</sup>LL EFT prediction of the hadron-level boosted top quark 2-Jettiness distribution in the peak region for  $e^+e^-$  colliders, which included power correction by matching to fixed order NLO QCD.

We used this framework to test the universal validity of the calibration results and the effects of power corrections by using different observables and gap schemes. This has led to the discovery of many issues that were not apparent with the original setup of [17] and to subsequent improvements of this framework. The improved analysis was then, in addition to PYTHIA, extended to the other major MC generators HERWIG and SHERPA to quantify how their masses relate to each other and the MSR mass and the pole mass.

The structure of this thesis is as follows: We describe all the ingredients of the leading power EFT description (singular cross section) in Chap. 2, which includes the factorization theorem, the description of nonperturbative effects, mass and soft gap renormalon subtraction. In Chap. 3 we review the power correction contributions (non-singulars) that arise when bHQET (boosted heavy quark effective theory) is matched to SCET (soft collinear effective theory) and subsequently to fixed order QCD and we present the expressions that are required for the description of the newly added observables. Here we will also explain a possible prescription of absorbing power corrections into the boundary conditions of the bHQET cross section. Chap. 4 shows how the ingredients are combined and we state how perturbative scale variation is parameterized in terms of so-called profile functions. In Chap. 5 we first review the fit procedure setup and the

results of the original work, then we present a necessary improvement of the fitting process of adding another fit parameter  $\Delta_0$  to the analysis. At the end of this chapter we show that the calibration framework does not give compatible results when the new observable *sum of jet mass* (SJM) is used instead of 2-Jettiness. In Chap. 6 we investigate the bHQET measurement function and discover that the leading power corrections responsible for the observable differences can be accounted for by including power correction rescaling factors of bHQET collinear and soft variables formulated directly within the measurement function. In Chap. 7 we present the final fit results using different observables, soft gap schemes, MC event generators and different power correction treatments. For the final version of our calibration procedure we had to devise an optimized “partial absorb” parametrization to estimate power correction uncertainties, since it became apparent that the global observable independent power corrections are large after the treatment of the leading observable dependent power corrections.

## 2. Resummed Cross Section

### 2.1. Factorization Theorem in the Peak Region

A factorization theorem that resums large logs in the peak region of the 2-jettiness distribution for  $e^+e^- \rightarrow t\bar{t} + X$  was derived in [12, 13] by using a sequence of EFTs.

The observables<sup>1</sup>  $\tau$  for which the factorization theorem is valid have the same definition, when first applying the power counting of SCET (soft-collinear effective theory) and then the power counting of bHQET (boosted heavy quark effective theory):

$$\tau \approx (M_t^2 + M_{\bar{t}}^2)/Q^2, \quad (2.1)$$

where  $M_t$  and  $M_{\bar{t}}$  are the invariant masses of the top and antitop hemisphere respectively and  $Q$  is the center of mass energy of the colliding electron and positron. We have, for the peak region that we are interested in, top and antitop jets that are characterized by pencil-shaped, i.e. very narrow and collimated, cones that are aligned to the thrust axis. The top and antitop hemispheres are defined by the plane perpendicular to the thrust axis. The minimum value for these observables in the case of stable top quarks has the form

$$\tau_{\min} = 2\hat{m}^2 + \mathcal{O}(\hat{m}^4), \quad (2.2)$$

with  $\hat{m} \equiv m_t/Q$ .  $m_t$  ( $\equiv m$ ) is the top quark mass parameter. In Sec. 3.2.3 we give details on the observables that we consider in this work and we explain how the hemispheres are exactly defined.

The scattering process exhibits large hierarchies of scales in the peak region of the distribution. The top and antitop are produced from the two colliding particles with a center of mass Energy  $Q$ . In the peak region the invariant masses  $M_{t,\bar{t}}$  are close to the top mass

$$\hat{s}_{t,\bar{t}} \equiv \frac{s_{t,\bar{t}}}{m} \equiv \frac{M_{t,\bar{t}}^2 - m^2}{m} \sim \Gamma \ll m, \quad (2.3)$$

where  $\Gamma$  is the top width and where we have defined the off-shellness variables  $\hat{s}_{t,\bar{t}}$  and  $s_{t,\bar{t}}$ . The quark and gluon fluctuations of the top and antitop jets that are compatible with this off-shellness region have a so-called ultra-collinear scaling of momentum components. In the top and antitop rest frames this scaling is homogeneous and has a size of  $\sim \Gamma$ . The scaling in the lab frame, which we define as the center of mass frame, is given in Sec. 6.2. The smallest ultra-collinear component is power counted as  $\sim \hat{m}\Gamma$ . This expression is used to define the homogeneous momentum scaling  $k_s^\mu \sim \Lambda \sim \hat{m}\Gamma$  of the large angle soft

---

<sup>1</sup>We will use  $e$  synonymously for  $\tau$  in later chapters. Both stand for generic thrust variables which give the same expression as 2-Jettiness in the leading power counting of SCET+bHQET.

fluctuations, that are responsible for the cross-talk between the two jets, where  $\Lambda \geq \Lambda_{\text{QCD}}$  is the size of the soft momenta.

The fixed order perturbative description of this process exhibits large logarithms of ratios of these momentum scales with  $Q \gg m \gg \Gamma > \hat{m}\Gamma$ . These logarithms invalidate the expansion in fixed orders of  $\alpha_s$ , since terms across different loop orders have to be counted to have the same size. A precise calculation of the differential cross section in the peak region therefore requires the introduction of effective field theories, which resum logarithms of the same size across all orders of perturbation theory (see Tab. 4.1 for the naming convention of the logarithmic resummation orders and which logarithms they resum).

The authors of Refs. [12, 13], which give the derivation of this factorization theorem, first apply soft-collinear effective theory (SCET) to integrate out fluctuations at the production scale  $Q$ , which leads to an expansion in  $\hat{m} = m/Q \ll 1$  and resums logs of  $\tau \sim \hat{m}^2$ . This theory can be used in the tail region where the off-shellness of the top quark jets  $s_{t,\bar{t}} = M_{t,\bar{t}}^2 - m^2$  are still of the size  $m^2$ . In the peak region the fluctuations are constrained to be much smaller, with the top width  $\Gamma$  acting as a lower bound to the rescaled off-shellness  $\Gamma \lesssim \hat{s}_{t,\bar{t}} \ll m$ . They therefore use boosted heavy quark effective theory (bHQET) to integrate out large momenta and resum logs of  $\hat{s}/m \sim \Gamma/m$ . The resulting bHQET event shape distribution reads

$$\begin{aligned} \frac{1}{\sigma_0} \frac{d\sigma_{\text{bHQET}}(\tau)}{d\tau} &= m_t Q^2 H_Q^{(6)}(Q, \mu_H) U_{H_Q}^{(6)}(Q, \mu_H, \mu_m) H_m^{(6)}(m, \varrho, \mu_m) U_v^{(5)}(\varrho, \mu_m, \mu) \\ &\times \int d\ell d\hat{s} U_B^{(5)}(\hat{s}_\tau - \varrho\ell - \hat{s}, \mu, \mu_B) B_\tau^{(5)}(\hat{s}, \Gamma_t, \delta m, \mu_B) \\ &\times \int d\ell' dk U_S^{(5)}(\ell - \ell', \mu, \mu_S) \hat{S}_\tau^{(5)}(\ell' - k, \bar{\delta}, \mu_S) F(k - 2\bar{\Delta}) \end{aligned} \quad (2.4)$$

where the superscripts (6) and (5) indicate the number of active dynamical quark flavours,

$$\varrho \equiv Q/m \quad (2.5)$$

is the leading term of the on-shell top quark boost factor in the SCET limit and the off-shellness variable

$$\hat{s}_\tau \equiv \frac{Q^2(\tau - \tau_{\min})}{m} \quad (2.6)$$

is a generic observable defined by Eq. (2.1) shifted by its threshold  $\tau_{\min}$ , which is given in Eq. (2.2). Usually, the exact expression  $\tau_{\min}$  for the specific observable considered is used, since the  $\mathcal{O}(\hat{m}^4)$  term turns out to have a large impact on the peak position. These  $\tau_{\min}$  expressions are given in Sec. 3.2.3 for observables relevant to this work. We will see in Sec. 5.2.4 that including only this threshold correction, i.e. using the exact  $\tau_{\min}$ , is not enough to make fits for different observables agree in the extracted top mass. The main ingredient in the solution of this problem is the inclusion of further mass power corrections in the measurement function, which is presented in Chap. 6.

## Hard Matching Functions

The SCET hard function  $H_Q$  is the squared Wilson coefficient obtained by matching QCD and SCET currents at leading order in  $\hat{m}$ . It contains the short distance dynamics at the scale  $Q$  that were integrated out in SCET [13]

$$H_Q(Q, \mu) = 1 + \frac{C_F \alpha_s(\mu)}{4\pi} \left\{ -2 \log^2 \left( \frac{Q^2}{\mu^2} \right) + 6 \log \left( \frac{Q^2}{\mu^2} \right) - 16 + \frac{7\pi^2}{3} \right\}. \quad (2.7)$$

Similarly,  $H_m$  is the matching coefficient between SCET and bHQET. It contains fluctuations at the top quark mass scale, that were integrated out in bHQET [13, 18]

$$H_m(m, \mu_m) = 1 + \frac{C_F \alpha_s(\mu_m)}{4\pi} \left\{ 2L_m^2 - 2L_m + 8 + \frac{\pi}{3} \right\} + \frac{C_F T_F \alpha_s^2(\mu_m)}{4\pi} \log \left( \frac{m^2}{Q^2} \right) \left\{ \frac{8}{3} L_m^2 + \frac{80}{9} L_m + \frac{224}{27} \right\}. \quad (2.8)$$

where  $L_m = \log(m^2/\mu_m^2)$ . The 2-loop term, which is enhanced by a so-called *rapidity logarithm*, is formally counted as  $\alpha_s^2 \log \hat{m}^2 \sim \mathcal{O}(\alpha_s)$  and is therefore included at N<sup>2</sup>LL. This term appears because there are two types of fluctuations at the mass scale, collinear and soft mass modes, which have the same invariant mass but different rapidities.

## The bHQET Jet Function

The bHQET jet function describes the remaining leading order dynamics of the ultra-collinear fluctuations along the top and anti-top directions

$$m^2 B(\hat{s}, \Gamma = 0, \delta_m = 0, \mu) = \delta(\hat{s}) + \frac{C_F \alpha_s(\mu)}{4\pi} \left\{ (8 - \pi^2) \delta(\hat{s}) + 16 \mathcal{L}_1^\mu(\hat{s}) - 8 \mathcal{L}_0^\mu(\hat{s}) \right\}, \quad (2.9)$$

where  $\mathcal{L}_i$  are the standard plus distributions defined in Eq. B.4.

The leading finite width effects can be expressed as a convolution of the stable jet function with a Breit-Wigner function

$$B_\tau(\hat{s}, \Gamma_t, \delta m, \mu_B) = \int \frac{d\hat{s}'}{\pi} \frac{2\Gamma_t}{(\hat{s} - \hat{s}')^2 + (2\Gamma_t)^2} B_\tau(\hat{s}', \Gamma_t = 0, \delta m, \mu_B) \quad (2.10)$$

where  $\Gamma_t$  is the top quark width.

The residual mass term  $\delta_m$  specifies the renormalization scheme for the top mass. It enters the threshold by the replacement  $\hat{s} \rightarrow \hat{s} - \frac{Q^2}{m} \frac{d\tau_{\min}}{dm} \delta m$ . The mass schemes used in this work are explained in Sec. 2.2.1.

## The Soft Function and Nonperturbative Effects

The soft function accounts for the effects of large-angle soft radiation with respect to the thrust axis. It also contains non-perturbative corrections due to hadronization effects. In the tail region of the distribution where  $\ell \sim \mu_S \gg \Lambda_{\text{QCD}}$  it is sufficient to use an operator product expansion (OPE) where the leading nonperturbative correction is given by one

parameter  $\overline{\Omega}_1$ , which describes how hadronization effects shift the perturbative cross section. In the peak region one has to use a full model to quantify the modifications of the peak shape caused by nonperturbative corrections. This can be achieved by including the nonperturbative effects in a model function (shape function)  $F(k)$  [19], which is convoluted with the leading term in the OPE, the perturbative soft function  $\hat{S}(\ell)$

$$S(\ell, \mu) = \int dk \hat{S}(\ell - k, \mu) F(k - 2\Delta) \quad (2.11)$$

$$\hat{S}(\ell, \mu) = \delta(\ell) + \frac{C_F \alpha_s(\mu)}{4\pi} \left\{ \frac{\pi^2}{3} \delta(\ell) - 16 \mathcal{L}_1^\mu(\ell) \right\} \quad (2.12)$$

The model function  $F(k)$  has support for  $k \geq 0$ , peaks at  $k \sim \Lambda_{\text{QCD}}$  and is normalized to unity. The parameter  $\Delta$  accounts for the average minimum hadronic energy deposit in each hemisphere originating from the hadronic masses and is also referred to as the ‘gap’ [19].

For our numerical studies we use the shape function parametrization developed by Ref. [20]

$$F(k; \lambda, \{c_i\}) \equiv \frac{1}{\lambda} \left[ \sum_{n=0}^N c_n f_n \left( \frac{k}{\lambda} \right) \right]^2 \quad (2.13)$$

$$f_n(z) = 8 \sqrt{\frac{2z^3(2n+1)}{3}} e^{-2z} P_n(g(z)) \quad (2.14)$$

$$g(z) = \frac{2}{3} [3 - e^{-4z} (3 + 12z + 24z^2 + 32z^3)] - 1$$

where  $P_n$  are the Legendre polynomials. The normalization is fixed by  $\sum_i c_i^2 = 1$ . We truncate the sum over basis functions  $f_n$  at  $N = 3$ , since this is sufficient to describe corrections to the peak shape due to nonperturbative effects. The first term in the series is a function with one peak. The terms with larger  $n$  have successively more oscillations and are less important for the shape of the peak of the cross section, because the details of the shape function will be smeared by the convolution. This conversely means that the first moment and the width of the truncated sum have to be set accordingly. The width can be adjusted with the parameter  $\lambda$  which will therefore be set at an appropriate value somewhat close to  $\Lambda_{\text{QCD}}$ . We will see in Chap. 5 that the first moment of the partial sum is very sensitive to the choice of the reference value of the renormalon free gap parameter  $\overline{\Delta}(R_s, \mu_S)$ , which is defined in Eq. (2.18), at the reference scale  $R_s = \mu_S = R_\Delta$ . We will demonstrate that  $\overline{\Delta}(R_\Delta, R_\Delta)$  should be determined by fitting it together with the mass parameter.

## Evolution Factors

The evolution kernels  $U_{HQ}$ ,  $U_v$ ,  $U_B$  and  $U_S$  perform the renormalization group (RG) evolutions for the hard function, the bHQET squared current

$$\mathcal{J}_\nu \equiv B_\tau \otimes S_\tau, \quad (2.15)$$

the jet function and the soft function respectively. They sum large logs between  $Q \gg m \gg \Gamma > \hat{m}\Gamma$ .

The hard function  $H_Q$  is evolved through the SCET current evolution kernel  $U_{H_Q}$  from the production scale  $\mu_H \sim Q$  to the mass scale  $\mu_m \sim m$ . The jet and soft function are evolved from their respective scales  $\mu_B \gtrsim \Gamma$  and  $\mu_S \gtrsim \hat{m}\Gamma$  to a scale  $\mu$  and this combination  $\mathcal{J}_\nu$  at scale  $\mu$  is then evolved with the bHQET current evolution kernel  $U_\nu$  to  $\mu_m$ . The cross section at any order is strictly invariant with respect to the choice of  $\mu$ .

The hard evolution factors are multiplicative and only change the normalization, while the soft and jet evolutions are given by convolutions. Their expressions and the anomalous dimensions are given in App. A.1.

## 2.2. Renormalon Subtractions

### 2.2.1. Mass schemes

It is known that the pole mass in QCD has a  $\mathcal{O}(\Lambda_{\text{QCD}})$  *renormalon*, caused by on-shell quark self energy corrections which are linearly sensitive to small momenta. These corrections are factorially divergent at large orders of the perturbation series and lead to an ambiguity of the pole mass of around 110 MeV [7] or 250 MeV [8]. The pole mass scheme has therefore been abandoned in high precision top, bottom and charm quark mass analysis in favor of *short distance* quark mass schemes that do not have this renormalon. The  $\overline{\text{MS}}$  mass is such a mass scheme and the perturbative relation to the pole mass is given by  $(\overline{m}_t \equiv \overline{m}_t^{(5+1)}(\overline{m}_t^{(5+1)}))$

$$\delta\overline{m}(\overline{m}_t) = m_t^{\text{pole}} - \overline{m}_t = \overline{m}_t \sum_{n=1} a_n^{\overline{\text{MS}}}(n_l = 5, n_h = 1) \left[ \frac{\alpha_s^{(5+1)}(\overline{m}_t)}{4\pi} \right]^n \quad (2.16)$$

in the approximation that the masses of all quarks lighter than the top quark are zero. The coefficients  $a_n^{\overline{\text{MS}}}(n_l, n_h)$  are known analytically at  $\mathcal{O}(\alpha_s, \alpha_s^2, \alpha_s^3)$  from Refs. [21–26]. Here,  $n_l$  is the number of massless quarks and  $n_h = 1$  means that the corrections include, apart from the  $n_l$  massless quarks, also the virtual loops of the heavy top quark.

The subtraction term is  $\mathcal{O}(m_t)$ , which is not compatible with the bHQET power counting. We therefore use the MSR mass  $m_t^{\text{MSR}}(R)$  [27], for which  $\delta m(R) \sim \alpha_s R$ , such that  $R$  can be adapted to be consistent with the power counting of the bHQET jet function. There are two possible  $\mathcal{O}(\Lambda_{\text{QCD}})$  renormalon free MSR mass prescriptions, which differ by how virtual top quark loops are treated in the matching with the  $\overline{\text{MS}}$  mass, called the “natural” and the “practical prescription”. The natural MSR mass does not absorb the self energy corrections that contain virtual top quark loops, while the practical MSR mass accounts for the virtual corrections from the top quark when evolving with  $R$ . We have used the practical prescription [27] in our analysis, since it was also employed in the original top mass calibration analysis of [17] and because the difference between the two prescriptions is quite small [27]. We have also tested the effect of using the natural prescription in our fits and the differences for the fitted mass were below 10 MeV. The defining series for the practical MSR mass  $m_t^{\text{MSRp}}(R)$  is obtained from the

$\overline{\text{MS}}$ -pole perturbative series of Eq. (2.16) by rewriting  $\alpha_s^{(5+1)}(\overline{m}_t)$  as a series in  $\alpha_s^{(5)}(\overline{m}_t)$  and then replacing  $\overline{m}_t$  by  $R$  [27]:

$$\delta m^{\text{MSRp}}(R) = m_t^{\text{pole}} - m_t^{\text{MSRp}}(R) = R \sum_{n=1} a_n^{\text{MSRp}}(n_l = 5) \left[ \frac{\alpha_s^{(5)}(R)}{4\pi} \right]^n \quad (2.17)$$

where the coefficients  $a_n^{\text{MSRp}}(n_l)$  and the solution of the  $R$ -evolution, that is the renormalization group evolution of  $m^{\text{MSRp}}(R)$  with respect to the renormalization scale  $R$ , are given in App. A.2. The subtraction series  $\delta m^{\text{MSRp}}(R)$  has to be expanded in powers of  $\alpha_s$  with its renormalization scale set to  $\mu_B$  of the jet function to consistently cancel the pole mass renormalon. Our profiles set the infrared subtraction scale  $R = \mu_B$  so that they automatically have a common renormalization scale.

### 2.2.2. Soft Gap subtraction schemes

The partonic soft function  $\hat{S}$  also has a leading  $\mathcal{O}(\Lambda_{\text{QCD}})$  renormalon which leads to instabilities of the partonic threshold. It is independent of the pole mass renormalon, which is isolated to the jet function, only an artifact of the pole mass prescription and can be removed just by switching to a short distance mass scheme. The ambiguity in the partonic soft function on the other hand is physical and related to a non-perturbative effect. The corrections in the partonic soft function that cause the renormalon ambiguity have to be canceled by an oppositely signed subtraction series  $\bar{\delta}(R_s, \mu_S)$  contained in the gap parameter  $\Delta$  which appears in the argument of the nonperturbative model function  $F(k - 2\Delta)$ . The RGE and scheme invariant gap parameter  $\Delta$  can therefore be split into the *gap subtraction* series  $\bar{\delta}(R_s, \mu_S)$  and a renormalon-free model parameter for the hadronic threshold  $\overline{\Delta}(R_s, \mu_S)$  [19]:

$$\bar{\delta}(R_s, \mu_S) = \Delta - \overline{\Delta}(R_s, \mu_S) = R_s \sum_{i=1} d_i(R_s, \mu_S) \left[ \frac{\alpha_s^{(5)}(\mu_S)}{4\pi} \right]^i, \quad (2.18)$$

where, depending on the gap definition, the  $\mu_S$  dependence of  $d_i(R_s, \mu_S)$  either just arises from a re-expansion of  $\alpha_s(R_s)$  in terms of  $\alpha_s(\mu_S)$  or it contains a non-vanishing anomalous dimension inherited from the soft function. So, the renormalon free gap parameter  $\overline{\Delta}(R_s, \mu_S)$  depends on the infrared subtraction scale  $R_s$  and for gap scheme 1, introduced below, also on the soft function renormalization scale  $\mu_S$ . The  $R_s$  and  $\mu_S$  evolutions are given in App. A.2. The perturbative and nonperturbative thresholds are renormalon free after splitting the gap parameter  $\Delta$

$$S(\ell, \mu_S) = \int dk \hat{S}(\ell - k, \bar{\delta}(R_s, \mu_S), \mu_S) F(k - 2\overline{\Delta}(R_s, \mu_S)) \quad (2.19)$$

with the renormalon subtracted perturbative soft function

$$\hat{S}(\ell, \bar{\delta}(R_s, \mu_S), \mu_S) = \hat{S}(\ell - 2\bar{\delta}(R_s, \mu_S), \mu_S), \quad (2.20)$$

where we have indicated that, analogously to the pole mass renormalon subtraction, the subtraction series has to be re-expanded in powers of  $\alpha_s(\mu_S)$  to consistently cancel the renormalon. A parametrization for suitable subtraction schemes, collectively referred to as R-gap schemes, has been introduced in Ref. [14] by imposing a general condition on the soft function at a point in position space. The solution is given by

$$\bar{\delta}(R_s, \mu_S; A, n, \xi) \equiv \begin{cases} \frac{R_s}{2\xi} \frac{d^n}{d \log(iy)^n} \log \left[ \tilde{S}_\tau(y, \mu_S) \right]_{iy=\frac{\xi}{R_s}} & \text{if A=on} \\ \frac{R_s}{2\xi} \frac{d^n}{d \log(iy)^n} \log \left[ \tilde{S}_\tau(y, R_s) \right]_{iy=\frac{\xi}{R_s}} & \text{if A=off} \end{cases} \quad (2.21)$$

where the position space soft function is defined as

$$\tilde{S}_\tau(y, \mu) = \int_{-\infty}^{\infty} d\ell^+ e^{-iy\ell^+} \hat{S}_\tau(\ell^+, \mu) = \exp \left\{ \sum_{i=1}^{\infty} \left[ \frac{\alpha_s^{(5)}(\mu)}{4\pi} \right]^i \sum_{j=0}^{i+1} s_{ij} \log^j(i e^{\gamma_E} y \mu) \right\}, \quad (2.22)$$

The coefficients  $s_{ij}$  can be generated according to Eq. (A.17). The switch A turns the non-trivial anomalous dimension in  $\mu_S$  on or off. When A=on the scale of the strong coupling in the subtraction series is  $\mu_S$  by construction, while A=off defines a gap subtraction series that only depends on  $R_s$ . We remind the reader that the strong coupling of the subtraction series for the latter case has to be expanded in terms of  $\alpha_s(\mu_S)$ , so that the renormalon can be properly canceled. In this work we will use three different schemes to test the gap dependence of the extracted fit parameters:

$$\begin{aligned} \bar{\delta}_1(R_s, \mu_S) &\equiv \bar{\delta}(R_s, \mu_S; \text{on}, 1, e^{-\gamma_E}) \\ \bar{\delta}_2(R_s, \mu_S) &\equiv \bar{\delta}(R_s, \mu_S; \text{off}, 0, e^{5\gamma_E}) \\ \bar{\delta}_3(R_s, \mu_S) &\equiv \bar{\delta}(R_s, \mu_S; \text{off}, 0, 1) \end{aligned} \quad (2.23)$$

The gap scheme 1 [28] was used in the original PYTHIA top mass calibration of Ref. [17]

$$\bar{\delta}_1(R_s, \mu_S) = \frac{R_s e^{\gamma_E}}{2} \sum_{i=1}^{\infty} \left[ \frac{\alpha_s^{(5)}(\mu_S)}{4\pi} \right]^i \sum_{j=0}^i (j+1) s_{i,j+1} \log^j \frac{\mu_S}{R_s} \quad (2.24)$$

The choice of  $n = 1$  sets the first constant coefficient to  $s_{11} = 0$ . Thus, the infrared cutoff scale  $R_s$  in this scheme has to be strictly set below the soft renormalization scale  $\mu_S$  to achieve a useful subtraction term with the proper sign at  $\mathcal{O}(\alpha_s)$ .

Gap scheme 3 was preferred in Ref. [17], since setting  $R_s$  in  $\bar{\delta}_1$  below  $\mu_S \sim \hat{s}/\rho$  in the peak region led to instabilities at N<sup>3</sup>LL due to the larger values of  $\alpha_s(\mu_S)$ .

$$\bar{\delta}_3(R_s, \mu_S) = \frac{R_s}{2} \sum_{i=1}^{\infty} \left[ \frac{\alpha_s^{(5)}(R_s)}{4\pi} \right]^i \sum_{k=0}^{i+1} s_{ik} \gamma_E^k \quad (2.25)$$

Scheme 3 is formally independent of the soft renormalization scale  $\mu_S$  when the  $\alpha_s$  series is summed to all orders. To properly cancel the renormalon in the soft function we have

to expand the strong coupling in terms of  $\alpha_s(\mu_S)$ . In the notation of Eq. (2.18) we have

$$d_i^{\text{gap } 3}(R_s, \mu_S) = \sum_{j=0}^{i-1} d_{ij}^{\text{gap } 3} \log^j \left( \frac{\mu_S}{R_s} \right) \quad (2.26)$$

with the coefficients given by the recursion relation

$$d_{ij}^{\text{gap } 3} = \frac{2}{j} \sum_{k=j}^{i-1} k d_{k,j-1}^{\text{gap } 3} \beta_{i-k-1}^{(5)} \quad (2.27)$$

$$d_{i0}^{\text{gap } 3} = d_i^{\text{gap } 3}(R_s, R_s),$$

where the QCD beta function coefficients are given in Eq. A.11

We have noticed in our numerical studies that  $\bar{\delta}_3$  has some unphysical behavior in the transition from the peak to the tail region when paired together with the pole mass scheme and using profile functions with fast changing scales. This is caused by large  $R$ -running effects of gap scheme 3 at NLL when the scales start to grow right after the peak. The thrust variable  $\tau$  in the cross section argument always appears in combination with the mass-dependent threshold and the soft gap in the form  $\tau - 2[m^2/Q^2 + \Delta/Q]$  (power corrections are omitted here). At NLL + pole mass scheme, there are no subtraction terms and only the gap  $\bar{\Delta}(R_s(\tau))$  runs as a function of  $\tau$ . The combination  $\tau - 2\bar{\Delta}(R_s(\tau))/Q$  might therefore be constant or even decreasing as a function of  $\tau$  if  $\bar{\Delta}(R_s(\tau))$  grows too fast over some  $\tau$ -range. This behaviour leads to stretched out peaks for profiles with fast growing scales. The MSR mass running has the opposite sign so that its effect on the cross section argument combination goes into the opposite direction. Accordingly, using the MSR mass scheme together with the gap 3 scheme at NLL results in a better behaviour of the cross section argument. The problematic large running effects of gap 3 led us to introduce the scheme gap 2

$$\bar{\delta}_2(R_s, \mu_S) = \frac{R_s}{2e^{5\gamma_E}} \sum_{i=1} \left[ \frac{\alpha_s^{(5)}(R_s)}{4\pi} \right]^i \sum_{k=0}^{i+1} s_{ik} (6\gamma_E)^k, \quad (2.28)$$

which differs from gap 3 by setting  $\xi$  to  $e^{5\gamma_E}$  instead of 1. This value removes the unphysical behavior in the above stated cases and sets the numerical value of the order  $\alpha_s$  subtraction halfway between gap 1 and gap 3.

### 3. Nonsingular Corrections

The resummed cross section from the previous chapter, also referred to as the singular cross sections, contains the leading distributional terms in an expansion in  $m/Q$ ,  $\Lambda_{\text{QCD}}/Q$  and  $\Gamma/m$  and the corrections are  $\mathcal{O}(\alpha_s \hat{m})$ ,  $\mathcal{O}(\hat{m}\Gamma/Q)$ ,  $\mathcal{O}(\hat{m}^2)$  and  $\mathcal{O}(\Gamma/m)$  [12]. In Ref. [17] the resummed bHQET distribution has been matched first to SCET and then to the full QCD 2-Jettiness distribution at NLO ( $\mathcal{O}(\alpha_s)$ ) to also include these missing power corrections in the fixed order expansion. We will in this chapter briefly review the matching prescription and present the missing results required for the additional observables *sum of jet mass* (sJM, or also referred to as *hemisphere mass sum*) and “modified jet mass” (mJM).

#### 3.1. QCD and SCET Nonsingular Distributions

We will use the term “QCD nonsingular” for corrections that come from QCD with respect to SCET and “SCET nonsingular” for corrections that are obtained from SCET with respect to bHQET. We will also freely switch between the two commonly used variable names for a generic thrust variable:  $\tau$  and  $e$ . They will be treated as synonyms of each other. In our notation these letters and also the jet function variable  $\hat{s}$  can be read as general expressions that are valid regardless of which prescription we use for the rescale factors that are introduced by Eq. (6.1). The measurement function that includes the rescaling factors  $r_c$  and  $r_s$  in the first line of Eq. (6.20), with  $\hat{s} = \hat{s}_1 + \hat{s}_2$ , will always be correct for the definition of  $\hat{s}$  (and in extension also  $s = m\hat{s}$ ). The rescale factors  $r_s$  and  $r_c$  should be set to 1 before we introduce them in Chap. 6. This just means that we are using the “no rescale” prescription, i.e. we do not include power correction in the factors that multiply the collinear variable  $\hat{s}$  and the soft variable  $\ell$ . In Chap. 6 we will restate general versions of those formulas where explicit factors of  $r_s$  and  $r_c$  are missing. Setting  $r_s$  and  $r_c$  in these general formulas to 1 will always again recover the corresponding formulas for the “no rescale” treatment.

The SCET and QCD matched cross section is given by

$$\frac{d\sigma}{de} = \frac{d\sigma_{\text{bHQET}}}{de} + \frac{d\sigma_{\text{nsb}}}{de} + \frac{d\sigma_{\text{ns}}}{de}. \quad (3.1)$$

The SCET matching contribution  $d\sigma_{\text{nsb}}/de$  is defined by the terms missing in bHQET when the bHQET jet scale  $\mu_B$  and the SCET jet scale  $\mu_J$  meet at the mass scale  $\mu_m$

$$\left. \frac{d\sigma_{\text{SCET}}}{de} \right|_{\mu_J=\mu_m} = \left[ \frac{d\sigma_{\text{bHQET}}}{de} + \frac{d\sigma_{\text{nsb}}}{de} \right]_{\mu_B=\mu_m}. \quad (3.2)$$

The term  $d\sigma_{\text{nsb}}/de$  arises from the non-distributional part  $J_{\text{nsb}}(s = m\hat{s}, m, \mu)$  of the SCET jet function  $J_{\text{SCET}}(s, \mu)$  [12], which is power suppressed in bHQET.

$$J_{\text{SCET}}(s, \mu_m) = H_m B_\tau(s/m, \mu_m) + J_{\text{nsb}}(s, m, \mu_m) \quad (3.3)$$

At N<sup>2</sup>LL this correction can therefore be simply expressed as the bHQET formula of Eq. (2.4) with  $B_\tau$  replaced by  $J_{\text{nsb}}$  [13]

$$\begin{aligned} \frac{d\sigma_{\text{nsb}}}{de} &= \frac{d\sigma_{\text{bHQET}}}{de} \Big|_{B_\tau^{(5)}(\hat{s}, \Gamma_t, \delta m, \mu_B) \rightarrow J_{\text{nsb}}^{(5)}(m\hat{s}, m, \mu_B)} \\ J_{\text{nsb}}(s, m, \mu) &= \frac{C_F \alpha_s(\mu)}{4\pi} \left[ \frac{2s}{(s+m^2)^2} - \frac{8}{s} \ln \left( 1 + \frac{s}{m^2} \right) \right] \theta(s) \end{aligned} \quad (3.4)$$

The running of this SCET nonsingular jet scale has been changed compared to the original PYTHIA calibration of [17] to not freeze out at  $\mu_m$ , but to exactly follow the physically more sensible bHQET jet scale  $\mu_{\text{nsb}} = \mu_B$ . Note that the SCET nonsingular is numerically negligible small in the peak region, when compared to the distributional part of the QCD nonsingulars.

The QCD nonsingulars are determined by setting all the scales in the SCET cross section to a common scale and computing the difference to the QCD cross section normalized by the massless quark tree-level cross section, the born cross section  $\sigma_0^C$

$$\begin{aligned} \frac{1}{\sigma_0^C} \frac{d\sigma_{ns}^C}{de} &\equiv \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{QCD}}^C}{de} - \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{SCET}}^C}{de} \Big|_{\text{FO}} \\ &= R_0^{C,\text{ns}}(\hat{m}) \delta(e - e_{\min}) + \frac{C_F \alpha_s}{4\pi} \left\{ A_e^{C,\text{ns}}(\hat{m}) \delta(e - e_{\min}) \right. \\ &\quad \left. + B_{\text{plus}}^{C,\text{ns}}(\hat{m}) \left[ \frac{1}{e - e_{\min}} \right]_+ + F_e^{\text{NS},C}(e, \hat{m}) \right\} \end{aligned} \quad (3.5)$$

The born cross section  $\sigma_0^C$ , given in Eq. 3.9, depends on the current, axial-vector ( $C = A$ ) or vector current ( $C = V$ ), due to different contributions from the photon and the Z boson to the propagator. The normalized QCD cross section  $(\sigma_0^C)^{-1} d\sigma_{\text{QCD}}^C/de$  also depends on  $C$ . In a strict treatment of power corrections the normalized SCET distribution  $(\sigma_0^C)^{-1} d\sigma_{\text{SCET}}^C/de$  is independent of  $C$ , but we want to keep Eq. (3.5) general, when we start to include power corrections in the SCET cross section by absorbing them into the coefficients of the delta and plus distribution, see Sec. 3.3. This absorption has to be performed separately for vector and axialvector current and leads to a  $C$  dependence also for the normalized redefined SCET distribution. After calculating the nonsingular distribution for each current we can then obtain the final QCD nonsingular distribution by adding both contributions

$$\frac{d\sigma_{\text{ns}}}{de} = \frac{d\sigma_{\text{ns}}^V}{de} + \frac{d\sigma_{\text{ns}}^A}{de}. \quad (3.6)$$

We note that the coefficients of the distributions, i.e.  $R_0^{C,\text{ns}}$ ,  $A_e^{C,\text{ns}}$  and  $B_{\text{plus}}^{C,\text{ns}}$ , are always extracted from the standard form shown in Eq. (3.5), which restricts the delta and plus

distribution arguments to be  $e - e_{\min}$  without any other multiplicative factors. This standard form can always be obtained by applying the rescaling relation Eq. (B.6) to plus functions and the standard rescaling formula to delta functions. This procedure guarantees that all the coefficients in Eq. (3.5) are unambiguously defined even if we use a power correction prescription, see Chap. 6, that rescales the argument of the SCET cross section.

We also note that the QCD term in Eq. (3.5) contains different power corrections for each of the observables that we consider, i.e.  $\tau_2$ ,  $\tau_m$  and  $\tau_s$ , but for a fixed observable it will not change due to any prescription that we introduce in this work, since our matched cross section in Eq. (3.1) should always recover the QCD cross section in the fixed order (FO) limit, i.e. when all renormalization scales are set to the same  $\mu$ .

Conceptually, it is actually not required to explicitly calculate the nonsingular coefficients in the second and third line of Eq. (3.5). Evaluating the SCET cross section at fixed order directly and then subtracting it from the QCD distribution is the more intuitive approach. In a numerical computation, however, it is more efficient to perform calculations with the nonsingular coefficients rather than repeating each computation for QCD and SCET separately and then taking the difference.

The fixed order SCET distribution<sup>1</sup> that has to be subtracted from the QCD event-shape distribution in 3.5 is given by [29]

$$\begin{aligned} \frac{1}{\sigma_0} \left. \frac{d\sigma_{\text{SCET}}}{de} \right|_{\text{FO}} &= \delta(e - e_{\min}) + \frac{C_F \alpha_s}{4\pi} \left\{ [2\pi^2 + 4 \log(\hat{m}) + 16 \log^2(\hat{m})] \delta(e - e_{\min}) \right. \\ &\left. - 8[1 + 2 \log(\hat{m})] \left[ \frac{1}{e - e_{\min}} \right]_+ + \left[ \frac{2(e - e_{\min})}{(e - e_{\min} + \hat{m}^2)^2} - \frac{8 \log(1 + \frac{e - e_{\min}}{\hat{m}^2})}{e - e_{\min}} \right] \theta(e - e_{\min}) \right\} \end{aligned} \quad (3.7)$$

## 3.2. QCD Massive Event-Shape Distribution

The full QCD calculation is required to obtain the QCD nonsingular distribution according to Eq. (3.5). The QCD distribution has the general form

$$\begin{aligned} \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{QCD}}^C}{de} &= R_0^C(\hat{m}) \delta(e - e_{\min}) + \frac{C_F \alpha_s}{4\pi} \left\{ A_e^C(\hat{m}) \delta(e - e_{\min}) \right. \\ &\left. + B_{\text{plus}}^C(\hat{m}) \left[ \frac{1}{e - e_{\min}} \right]_+ + F_e^{\text{NS},C}(e, \hat{m}) \right\} + \mathcal{O}(\alpha_s^2). \end{aligned} \quad (3.8)$$

The following subsections will list all the relevant formulas for the coefficients in this equation and for the function  $F_e^{\text{NS},C}(e, \hat{m})$  for each observable  $e$  that we consider.

---

<sup>1</sup>This formula uses strict power counting, i.e.  $r_s = 1$  and  $r_e = 1$ , see Eq. (6.22) for the general case including mass-dependent measurement power corrections. We have also dropped the  $C$  superscript since the normalized SCET cross section only depends on  $C$  when the ‘‘absorb’’ prescription is used, see Sec. 3.3

### 3.2.1. Born Cross Section

Event-shape distributions are customary normalized to the Born cross section defined as the cross section for massless quarks at tree-level. The Born cross section can be split into vector and axial-vector contributions  $\sigma_0 = \sigma_0^V + \sigma_0^A$ . The former is mediated by both photon and Z boson exchange, while the latter only contains Z boson contributions. They are given by [29]

$$\begin{aligned}\sigma_0^V &= \frac{N_c}{3} \frac{4\pi\alpha_{\text{em}}^2}{Q^2} \left[ Q_q^2 + \frac{v_q^2(v_e^2 + a_e^2)}{(1 - \hat{m}_Z^2)^2 + \left(\frac{\Gamma_Z}{m_Z}\right)^2} + \frac{2Q_q v_e v_q (1 - \hat{m}_Z^2)}{(1 - \hat{m}_Z^2)^2 + \left(\frac{\Gamma_Z}{m_Z}\right)^2} \right] \\ \sigma_0^A &= \frac{N_c}{3} \frac{4\pi\alpha_{\text{em}}^2}{Q^2} \left[ \frac{a_q^2(v_e^2 + a_e^2)}{(1 - \hat{m}_Z^2)^2 + \left(\frac{\Gamma_Z}{m_Z}\right)^2} \right]\end{aligned}\quad (3.9)$$

with  $N_c$  the number of colors,  $\alpha_{\text{em}}$  the electromagnetic coupling,  $Q_q$  the quark electric charge,  $\hat{m}_Z = m_Z/Q$  the reduced Z boson mass,  $\Gamma_Z$  the finite width of the Z boson,  $v_i = (T_3^i - 2Q_i \sin^2 \theta_W)/\sin(2\theta_W)$ ,  $a_i = T_3^i/\sin(2\theta_W)$  the vector and axial couplings of the electron or quark to the Z boson.

The finite quark mass tree level delta-distribution coefficients then read

$$R_0^V(\hat{m}) = \frac{(3 - v^2)v}{2}, \quad R_0^A(\hat{m}) = v^3 \quad (3.10)$$

### 3.2.2. Phase Space and Kinematic Variables

For the calculation of the  $\mathcal{O}(\alpha_s)$  distribution, which consists of the virtual radiation contribution with back-to-back top and anti-top quarks and the real radiation contribution with an additional real gluon in the phase space integral, we follow the instructions of [15]. We define the convenient phase space variables  $y$  and  $z$

$$x_q = 1 - zy, \quad x_{\bar{q}} = 1 - (1 - z)y \quad (3.11)$$

with  $x_i = 2E_i/Q$  the fractional energy of particle  $i$ . Energy conservation  $x_q + x_{\bar{q}} + x_g = 2$  then fixes  $y = 2E_g/Q$ . The Dalitz region is parametrized as

$$\begin{aligned}0 &\leq y \leq y_{\text{max}}(z) \equiv 1 - \frac{\hat{m}^2}{z(1 - z)}, \quad z_- \leq z \leq z_+, \\ z_{\pm} &\equiv \frac{1 \pm v}{2}\end{aligned}\quad (3.12)$$

with  $v = \sqrt{1 - 4\hat{m}^2}$  the velocity of an on-shell top quark in the center of mass frame.

The 3-body phase space can be split into three regions, which we call the quark-, the antiquark- and the gluon-region. In the quark(antiquark)-region the thrust axis is aligned with the quark(antiquark) momentum and it is alone in one hemisphere while the antiquark(quark) and gluon momenta point into the other hemisphere. The gluon region

describes a gluon traveling alone into one hemisphere with its momentum aligned with the thrust axis while the quark and anti-quark travel into the other hemisphere. The gluon region only contributes far away from the peak of our event-shape distributions. The boundaries of these regions are given by [15]

$$\begin{aligned}
\text{quark - region :} & \quad z_- \leq z \leq \frac{1}{2}, & 0 \leq y \leq \min[y_{\max}(z), y_{\tau}(z)] \\
\text{antiquark - region :} & \quad \frac{1}{2} \leq z \leq z_+, & 0 \leq y \leq \min[y_{\tau}(1-z), y_{\max}(z)] \\
\text{gluon - region :} & \quad \hat{m} \leq z \leq 1 - \hat{m}, & \max[y_{\tau}(z), y_{\tau}(1-z)] \leq y \leq y_{\max}(z),
\end{aligned} \tag{3.13}$$

where

$$y_{\tau}(z) = \frac{\sqrt{1 - 4\hat{m}^2(1 - z^2)} - z}{1 - z^2} \tag{3.14}$$

separates the quark and gluon region. The phase space as well as matrix elements are invariant under the transformation  $z \rightarrow 1 - z$ , i.e. mirror symmetric with respect to the  $z = 1/2$  vertical line. The quark and anti-quark regions are separated by this line and this transformation interchanges the quark and anti quark regions and maps the gluon region onto itself.

Expressions relating the observable to the phase space variables  $z$  and  $y$ , also called measurement functions, can be calculated from the following expressions for the quark, antiquark and gluon momenta ( $\tilde{p}^{\mu} \equiv p^{\mu}/Q = [E, p_X, p_Y, p_Z]/Q$ ) [15]:

$$\begin{aligned}
\tilde{p}_g^{\mu} &= \frac{y}{2}[1, 0, 0, 1] \\
\tilde{p}_q^{\mu} &= \left[ \frac{1 - yz}{2}, 0, -\sqrt{(1 - y)(1 - z)z - \hat{m}^2}, \frac{2z - 1 - yz}{2} \right] \\
\tilde{p}_{\bar{q}}^{\mu} &= \left[ \frac{1 - y(1 - z)}{2}, 0, \sqrt{(1 - y)(1 - z)z - \hat{m}^2}, \frac{1 - y(1 - z) - 2z}{2} \right]
\end{aligned} \tag{3.15}$$

where we use capital letters to refer to Cartesian coordinates. The gluon three-momentum has been aligned with the  $Z$ -direction and the scattering plane has been fixed to lie in the  $X, Y$  plane. This is sufficient, because our event shapes as well as our phase space variables,  $y$  and  $z$ , are invariant with respect to rotation of the coordinate system.

### 3.2.3. Observable Definitions and Properties

For the following subsections we will need the definitions of the observables that are used in the calibration fits, the boundaries of the quark and gluon regions, and their measurement functions in these regions. Our observables depend on the thrust axis which appears in the definition of the conventional ‘‘thrust’’ event shape

$$T = \max_{\hat{t}} \frac{\sum_i |\hat{t} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|} \tag{3.16}$$

where the sum runs over all final state particles with momenta  $\vec{p}_i$ . This thrust definition has its peak region located around the maximum value of  $T$  for the stable top quark

distribution. The peak region is associated with 2-jet events. These are characterized by two jets, consisting of highly boosted collinear particles traveling either parallel or anti-parallel to the thrust axis, and large angle soft radiation in between the jet directions. The plane perpendicular to the thrust axis separates the top and antitop hemispheres. This separation plane is used to define other event shapes that depend on the invariant mass of the sum of momenta in these hemispheres. Given Hemispheres  $a$  and  $b$ , we define the dimensionless invariant masses as

$$\rho_{a,b} = \frac{1}{Q^2} \left( \sum_{i \in a,b} p_i \right)^2 \quad (3.17)$$

where the sum runs over all final state particles in either hemisphere  $a$  or  $b$  and  $Q$  is the center of mass energy.

The observables that we consider are the same at leading order power counting in SCET. Their differences lie in higher order power corrections. Our event shape variables are defined so that the peak of the stable quark event shape distribution lies at  $\tau_{\min} = 2\hat{m}^2 + \mathcal{O}(\hat{m}^4)$  and the radiative tail continuous into the positive  $\tau$  direction. Performing fits using event shapes that differ in their power corrections will allow us to estimate how important these corrections are. The  $\hat{m}^2 \sim 1/10$  corrections are numerically the largest power corrections, and it is therefore these power corrections that we examine and treat carefully.

### Heavy Jet Mass (HJM)

This observable is not directly used for our fits, but the available results for the  $\alpha_s$  fixed order distribution can be reused for other event shapes. It is given by the larger of the two invariant hemisphere masses

$$\rho = \max(\rho_a, \rho_b). \quad (3.18)$$

We need the measurement functions for each of the phase space regions defined above (3.13), that is  $\rho(y, z)$ , to later perform the phase space integrals. The measurement functions in the quark and gluon regions can be obtained by inserting Eq. (3.15) into the observable definition while taking into account which particles are together in one hemisphere. For the quark and gluon region we obtain

$$\begin{aligned} \rho_{\text{qu}}(y, z) &= \hat{m}^2 + yz \\ \rho_{\text{gl}}(y, z) &= 1 - y \end{aligned} \quad (3.19)$$

The expression for the antiquark region is not needed, because the antiquark contribution to the event shape distribution is identical to the quark contribution.

We also have to determine the region boundaries, given by Eq. (3.13), in terms of  $\rho$  values to determine which of the three phase space regions contribute to a specific  $\rho$  value

$$\begin{aligned} \hat{m}^2 &\leq \rho_{\text{qu}} \leq (2r - 1)/3 + \hat{m}^2 \\ 4\hat{m}^2 &\leq \rho_{\text{gl}} \leq (5 - 4r)/3, \end{aligned} \quad (3.20)$$

where  $r = \sqrt{1 - 3\hat{m}^2}$ . The minimum and maximum values are reached in the quark region at the phase space points  $y = 0$  and  $y_r(z = 1/2)$  respectively, in the gluon region they are reached at  $y_{max}(z = 1/2)$  and  $y_r(z = 1/2)$  respectively. The gluon region only contributes far away from the peak, so it will not affect our fit results.

### Sum of Jet Mass (SJM)

Sum of jet mass, also referred to as hemisphere mass sum, is defined as the sum of the invariant hemisphere masses

$$\tau_s = \rho_a + \rho_b. \quad (3.21)$$

The measurement functions in the quark and gluon regions are

$$\begin{aligned} \tau_{s,\text{qu}}(y, z) &= 2\hat{m}^2 + yz = \rho_{\text{qu}}(y, z) + \hat{m}^2 \\ \tau_{s,\text{gl}}(y, z) &= \rho_{\text{gl}}(y, z) \end{aligned} \quad (3.22)$$

These relations show that the contributions from the phase space regions are just shifted by a constant for the quark region and identical for the gluon region when compared to HJM. The  $\tau_s$  ranges are therefore

$$\begin{aligned} 2\hat{m}^2 &\leq \tau_{s,\text{qu}} \leq (2r - 1)/3 + 2\hat{m}^2 \\ 4\hat{m}^2 &\leq \tau_{s,\text{gl}} \leq (5 - 4r)/3 \end{aligned} \quad (3.23)$$

### Modified Jet Mass (MJM)

We define a new observable *modified jet mass*  $\tau_m$  as a quadratic function of SJM so that the leading  $\mathcal{O}(m^2/Q^2)$  power correction between  $\tau_m$  and the bHQET soft convolution variable  $\ell$  vanishes. The motivation for this definition is explained in Chap. 6.

$$\tau_m(\tau_s) = \tau_s + \frac{1}{2}\tau_s^2 \quad (3.24)$$

Measurement functions are obtained by just inserting the SJM expressions into the definition

$$\begin{aligned} \tau_{m,\text{qu}}(y, z) &= 2(\hat{m}^2 + \hat{m}^4) + (1 + 2\hat{m}^2)yz + \frac{y^2 z^2}{2} \\ \tau_{m,\text{gl}}(y, z) &= (1 - y) + \frac{1}{2}(1 - y)^2 \end{aligned} \quad (3.25)$$

The event shape ranges can be directly obtained from Eq. (3.23) by applying the observable definition on the boundaries, since Eq. (3.24) is strictly increasing within the phase space

$$\tau_{m,\text{qu/gl,min/max}} = \tau_m(\tau_s = \tau_{s,\text{qu/gl,min/max}}). \quad (3.26)$$

## 2-Jettiness (“thrust”)

2-Jettiness ( $\tau_J$  or  $\tau_2$ ) is the observable which has been used in the original MC mass calibration [17]. It is defined as  $1 - T_Q$  where  $T_Q$  is the conventional thrust variable Eq. (3.16) with the sum over 3-momenta lengths in the denominator replaced by  $Q$

$$\tau_J = \frac{1}{Q} \min_{\hat{t}} \sum_i (E_i - |\hat{t} \cdot \vec{p}_i|). \quad (3.27)$$

We will sometimes also refer to 2-Jettiness as “thrust”, since we will not use any other thrust definitions for our fits. The measurement functions are

$$\begin{aligned} \tau_{J,\text{qu}}(y, z) &= 1 - \sqrt{(1 - yz)^2 - 4\hat{m}^2} \\ &= 1 - v + \frac{yz}{v} + \mathcal{O}(y^2) \\ \tau_{J,\text{gl}}(y, z) &= \rho_{\text{gl}}(y, z) \end{aligned} \quad (3.28)$$

The  $\tau_J$  ranges are

$$\begin{aligned} 1 - v &\leq \tau_{J,\text{qu}} \leq (5 - 4r)/3 \\ 4\hat{m}^2 &\leq \tau_{J,\text{gl}} \leq (5 - 4r)/3. \end{aligned} \quad (3.29)$$

### 3.2.4. Delta- and Plus-Function Coefficients

The observables  $e$  that we use are at the threshold linear sensitive to the soft momentum and can in the soft limit  $y \rightarrow 0$  be expanded as

$$e(y, z) = e_{\text{min}} + y f_e(z) + \mathcal{O}(y^2) \quad (3.30)$$

The differences of the  $\mathcal{O}(\alpha_s)$  delta distribution coefficients  $A_e^C(\hat{m})$  between these observables are fully determined by the coefficient  $f_e(z)$  of the linear term [15]

$$\begin{aligned} A_e^V(\hat{m}) &= 4(1 + 2\hat{m}^2) \left\{ (1 - 2\hat{m}^2) \left[ \text{Li}_2 \left( -\frac{v(1+v)}{2\hat{m}^2} \right) - 3 \text{Li}_2 \left( \frac{v(1-v)}{2\hat{m}^2} \right) + 2 \log^2(\hat{m}) + \pi^2 \right. \right. \\ &\quad \left. \left. - 2 \log^2 \left( \frac{1+v}{2} \right) \right] + 2v[\log(\hat{m}) - 1] - 2I_e(\hat{m}) \right\} + 4(4 + v^2 - 16\hat{m}^4)L_v \\ A_e^A(\hat{m}) &= 4v^2 \left\{ (4 + v^2)L_v + 2v[\log(\hat{m}) - 1] - 2I_e(\hat{m}) + (1 - 2\hat{m}^2) \right. \\ &\quad \left. \times \left[ \text{Li}_2 \left( -\frac{v(1-v)}{2\hat{m}^2} \right) - 3 \text{Li}_2 \left( \frac{v(1-v)}{2\hat{m}^2} \right) + \pi^2 + 2 \log^2(\hat{m}) - 2 \log^2 \left( \frac{1+v}{2} \right) \right] \right\} \end{aligned} \quad (3.31)$$

with  $L_v \equiv \log[(1+v)/(2\hat{m})]$  and the only event shape dependent term is

$$I_e(\hat{m}) = \int_{z_-}^{1/2} dz \frac{(1-z)z - \hat{m}^2}{(1-z)^2 z^2} \log[f_e(z)] \quad (3.32)$$

This integral is formulated in terms of the quark region boundaries, so the function  $f_e(z)$  has to be extracted from the quark region measurement function. For our observables this term can be expressed as a combination of the HJM integral  $I_\rho$  and the integral

$$\mathcal{I} \equiv \int_{z_-}^{1/2} dz \frac{(1-z)z - \hat{m}^2}{(1-z)^2 z^2} = (1+v^2)L_v - v \quad (3.33)$$

HJM:

$$\begin{aligned} I_\rho(\hat{m}) = & \frac{1}{24} \left\{ -12(v^2 + 1) \text{Li}_2 \left( \frac{v+1}{2} \right) \right. \\ & + \pi^2(v^2 + 1) - 6(v(v(2 + \log^2 2) + 2 - 4 \log 2) + \log^2 2) \\ & \left. + 6 \log(1-v) \left( (v^2 + 1) \log \frac{4}{1-v} - 4v \right) + 6(v^2 - 1) \log \frac{1-v}{v+1} \right\} \end{aligned} \quad (3.34)$$

SJM:

$$I_{\tau_s}(\hat{m}) = I_\rho(\hat{m}) \quad (3.35)$$

MJM:

$$I_{\tau_m}(\hat{m}) = I_\rho(\hat{m}) + \log(1 + 2\hat{m}^2) \mathcal{I}(\hat{m}) \quad (3.36)$$

2-Jettiness:

$$I_{\tau_J}(\hat{m}) = I_\rho(\hat{m}) - \log(v) \mathcal{I}(\hat{m}) \quad (3.37)$$

The coefficient of the plus-function is universal

$$B_{\text{plus}}(\hat{m}) = 4 \begin{pmatrix} 3 - v^2 \\ 2v^2 \end{pmatrix} [(1+v^2)L_v - v], \quad (3.38)$$

where the upper (lower) element is the vector (axial-vector) current expression.

### 3.2.5. Radiative Tail

The event shape distribution away from the threshold ( $e > e_{\min}$ ) can be written as a phase space integral over the matrix element squared in four dimensions [15]

$$F_e^C(e, \hat{m}) \equiv \frac{B_{\text{plus}}(\hat{m})}{e - e_{\min}} + F_e^{\text{NS},C}(e, \hat{m}) = \int dz dy \frac{M_C(y, z)}{y} \delta[e - \hat{e}(y, z)] \quad (3.39)$$

with

$$\begin{aligned} M_C(y, z) &= M_C^0(z) + y M_C^2(z) + y^2 M_C^3(z) \\ M_C^0(z) &= 4 \begin{pmatrix} 1 + 2\hat{m}^2 \\ v^2 \end{pmatrix} \frac{(1-z)z - \hat{m}^2}{(1-z)^2 z^2} \\ M_C^2(z) &= -4 \begin{pmatrix} 1 + 2\hat{m}^2 \\ v^2 \end{pmatrix} \frac{1}{z(1-z)} \\ M_C^3(z) &= 4 \left[ \begin{pmatrix} 1 \\ 1 + 2\hat{m}^2 \end{pmatrix} \frac{1}{2z(1-z)} - 1 \right], \end{aligned} \quad (3.40)$$

where again upper (lower) elements correspond to vector (axial-vector) current expressions. The delta function selects the constant event shape lines in the  $y$ - $z$  plane. For the observables used in this work one can write the result as a sum of contributions from the quark and gluon regions

$$F_e^C(e, \hat{m}) = 2 \sum_{i \in \{\text{qu, gl}\}} \int_{z_{\min}^i}^{\frac{1}{2}} dz \left| \frac{M_C(y, z)}{y \frac{d\hat{e}(y, z)}{dy}} \right|_{y=y_i(e, z)}. \quad (3.41)$$

$$\equiv 2F_{e, \text{qu}}^C(e, \hat{m}) + F_{e, \text{gl}}^C(e, \hat{m})$$

The lower integration limit  $z_{\min}^i$  either coincides with the phase space boundary of Eq. (3.12) or the quark-gluon-region boundary Eq. (3.14). The upper limit is  $1/2$  due to the mirror symmetry around  $z = 1/2$ . The factor 2 in the first line of the equation above therefore accounts for the anti-quark region ( $i = \text{qu}$ ) and the second half of the gluon region ( $i = \text{gl}$ ).

### HJM (Heavy Jet Mass)

The results for HJM are [15]

$$F_{\rho, \text{qu}}^V(\rho, \hat{m}) = \frac{(2 - 4z)[\rho(1 - z)(\rho - 4z) - 2\hat{m}^2(\rho + 2(1 - 2\rho)z^2 + 3\rho z) + \hat{m}^4(1 - z - 8z^2)]}{(1 - z)z^2(\rho - \hat{m}^2)}$$

$$+ 4 \left( \rho - 2 - 5\hat{m}^2 + 2 \frac{1 - 4\hat{m}^4}{\rho - \hat{m}^2} \right) \log \left( \frac{1 - z}{z} \right)$$

$$F_{\rho, \text{qu}}^A(\rho, \hat{m}) = 4[4 - 8\hat{m}^2(2 + \rho) + 8\hat{m}^4] + \frac{2}{\rho - \hat{m}^2} \left\{ \frac{(1 + 2\hat{m}^2)(\rho - \hat{m}^2)^2}{z^2} + \frac{4(1 - 4\hat{m}^2)\hat{m}^2}{1 - z} \right.$$

$$\left. - \frac{2\rho(2 + \rho) - 2\hat{m}^2\rho(5 + \rho) + \hat{m}^4(1 + 4\rho) - 2\hat{m}^6}{z} \right.$$

$$\left. + 2[2 - (2 - \rho)\rho + 2\hat{m}^2(\rho(\rho + 3) - 5) + \hat{m}^4(9 - 4\rho) + 2\hat{m}^6] \log \left( \frac{1 - z}{z} \right) \right\}$$

$$F_{\rho, \text{gl}}^V(\rho, \hat{m}) = \frac{4}{t} \left\{ [2 - (2 - t)t - 4\hat{m}^2t - 8\hat{m}^4] \log \left( \frac{1}{z} - 1 \right) - \frac{(1 - 2z)[(1 - z)zt^2 + 2\hat{m}^2 + 4\hat{m}^4]}{(1 - z)z} \right\}$$

$$F_{\rho, \text{gl}}^A(\rho, \hat{m}) = \frac{4[2 - 2t + t^2 + 2\hat{m}^2(t^2 + 4t - 6) + 16\hat{m}^4] \log \left( \frac{1}{z} - 1 \right) - \frac{(4 - 8z)[t^2(1 - z)z + 2\hat{m}^2 - 8\hat{m}^4]}{(1 - z)z}}{t} \quad (3.42)$$

given in terms of the lower integration limit  $z = z_{\min}^i$

$$z_{\min}^{\text{qu}} = \begin{cases} (t_\rho - \xi_\rho)/2 & m^2 < \rho < \hat{m}(1 - \hat{m} - \hat{m}^2)/(1 - \hat{m}) \\ (t_\rho - 1)/\sqrt{(1 - \rho)^2 - 2\hat{m}^2(1 + \rho) + \hat{m}^4} & \text{larger values} \end{cases}$$

$$z_{\min}^{\text{gl}} = \begin{cases} [1 - \sqrt{1 - 4\hat{m}^2/\rho}]/2 & 4\hat{m}^2 < \rho < \hat{m}/(1 - \hat{m}) \\ [1 - \sqrt{(1 - \rho)^2 + 4\hat{m}^2}]/(1 - \rho) & \text{larger values,} \end{cases} \quad (3.43)$$

where we defined  $t \equiv 1 - \rho$ ,  $t_\rho \equiv 1 + \rho - \hat{m}^2$ ,  $\xi_\rho \equiv \sqrt{t_\rho^2 - 4\rho}$ . Here, “larger values” refers to the rest of the event shape values up to the maximum value for either quark or gluon region, which are given in Sec. 3.2.3

### sJM (sum of Jet Mass) and mJM (modified Jet Mass)

The measurement functions of sJM and mJM, restricted to either quark or gluon region, can be related to the HJM measurement function by a bijective map

$$e'(y, z; \hat{m}) = e'(e(y, z; \hat{m}); \hat{m}). \quad (3.44)$$

The differential distribution can therefore be obtained from the HJM results by means of variable transformations

$$\int de F_{e,i}^C(e, \hat{m}) = \int de' \left| \frac{de}{de'} \right| F_{e,i}^C[e(e'; \hat{m}), \hat{m}] = \int de' F_{e',i}^C(e', \hat{m}) \quad (3.45)$$

We can for example obtain the sJM quark region contribution by inserting the inverse of the measurement function Eq. (3.22) into the result for HJM in Eq. (3.42), if we set  $e' = \tau_{s,\text{qu}}$  and  $e = \rho_{\text{qu}}$

$$F_{\tau_s,\text{qu}}^C(\tau_s, \hat{m}) = |1| F_{\rho,\text{qu}}^C[\rho = \tau_s - \hat{m}^2, \hat{m}] \quad (3.46)$$

### 2-Jettiness

The 2-Jettiness distribution has been calculated in [29] and reads in our notation

$$\begin{aligned} F_{\tau_J,\text{qu}}^V(\tau, \hat{m}) &= \frac{2t_\tau}{(z-1)z^2(\xi - t_\tau^2 - 4m^2)} \{4(z-1)z^2 \tanh^{-1}(1-2z) \\ &\quad \times (4m^2\xi - 8m^4 + (\tau-2)\tau + 2) - (2z-1)[4m^2(2z((\xi-1)z + 2 - \xi) - 1) \\ &\quad + 8m^4z + (z-1)((\tau-2)\tau + 4(\xi-1)z + 2 - 2\xi)]\} \\ F_{\tau_J,\text{qu}}^A(\tau, \hat{m}) &= \frac{2t_\tau}{(z-1)z^2(\xi - t_\tau^2 - 4m^2)} \{(2z-1)[2m^2((\tau-2)\tau \\ &\quad + 4z^2(4\xi - (\tau-2)\tau - 4) + z(3(\tau-2)\tau + 10 - 14\xi) + 4 - 2\xi) \\ &\quad + 8m^4(z(5-4z) + 1) - (z-1)((\tau-2)\tau + 4(\xi-1)z + 2 - 2\xi)] \\ &\quad + 4(z-1)z^2 \tanh^{-1}(1-2z)[2m^2((\tau-2)\tau + 2 - 6\xi) + 24m^4 + (\tau-2)\tau + 2]\} \\ F_{\tau_J,\text{gl}}^C(\tau, \hat{m}) &= F_{\rho,\text{gl}}^C(\tau, \hat{m}) \end{aligned} \quad (3.47)$$

with  $t_\tau \equiv 1 - \tau$ ,  $\xi \equiv \sqrt{t_\tau^2 + 4\hat{m}^2}$  and  $z = z_{\min}$ :

$$z_{\min}^{\text{qu}} = \begin{cases} (1 + \tau - \xi)/2 & 1 - v < \tau < \hat{m}/(1 - \hat{m}) \\ (1 - \xi)/t_\tau & \text{larger values} \end{cases} \quad (3.48)$$

### 3.3. Absorption of Nonsingular Coefficients

Instead of using strict power counting for the coefficients in the bHQET boundary conditions  $H_Q$ ,  $H_m$  and  $B_\tau$ , that is Eq. (2.7), (2.8) and (2.9), which we will call the “no absorb” prescription, we can absorb the coefficients from the QCD nonsingular distribution into these functions to assess how resummation of power suppressed corrections might affect the results. In this sections we will describe the prescription of fully absorbing all distributional (delta and plus function) power corrections. In section 7.1.2 we will generalize this prescription to partial absorption. Hence we demand that the “full absorb” or in short “abs” prescription sets the QCD nonsingular to

$$\frac{1}{\sigma_0^C} \frac{d\sigma_{\text{abs,ns}}^C}{de} \stackrel{!}{=} \frac{C_F \alpha_s}{4\pi} F_e^{\text{NS,C}}(e, \hat{m}) + \mathcal{O}(\alpha_s^2) \quad (3.49)$$

where  $F_e^{\text{NS,C}}(e, \hat{m})$  is the nondistributional part of the QCD nonsingular, see Eq. (3.8), which we do not absorb since it is numerically insignificant. According to the definition of the QCD nonsingular in the first line of Eq. (3.5) this then means, that our singular bHQET distribution has to be modified, so that the sum of singular and nonsingulars at fixed order still equals the QCD differential cross section. We can first split up the delta and plus function coefficients as

$$R_0^{\text{QCD}} = 1 + R_0^{\text{abs}} \equiv R_0^{\text{newsing}}, \quad A_e^{\text{QCD}} = A_e^{\text{base}} + A_e^{\text{abs}}, \quad B_{\text{plus}}^{\text{QCD}} = B_{\text{plus}}^{\text{base}} + B_{\text{plus}}^{\text{abs}}, \quad (3.50)$$

where we have dropped the current superscript  $C$  from terms with “QCD”, “newsing” or “abs” superscript. Here the “QCD” superscript terms are identical to the corresponding coefficients in the fixed order QCD differential cross section Eq. (3.8). These terms obviously do not depend on our treatment of power corrections. The “abs” superscript terms are defined to be identical to the coefficients in the nonsingular distribution Eq. (3.5) for the “no absorb” case. The “base” superscript terms are then identical to the coefficients of the “no absorb” SCET distribution at fixed order Eq. (3.7). The bHQET boundary conditions  $H_Q$ ,  $H_m$  and  $B_\tau$  in Eq. (2.4) at N<sup>2</sup>LL are then replaced by corresponding functions, here with a tilde, which contain power corrections [29]

$$\begin{aligned} \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{bHQET}}^C}{de} \Big|_{\text{withAbs}} &= R_0^{\text{newsing}} \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{bHQET}}^C}{de} \Big|_{H_Q \rightarrow \tilde{H}_Q, H_m \rightarrow \tilde{H}_m, B \rightarrow \tilde{B}} \\ \tilde{H}_Q(\mu_H) &= H_Q(\mu_H) + \frac{C_F \alpha_s(\mu_H)}{4\pi} (1 - \xi_J - \xi_B) h^{\text{corr}} \\ \tilde{H}_m(\mu_m) &= H_m(\mu_m) + \frac{C_F \alpha_s(\mu_m)}{4\pi} \xi_J h^{\text{corr}} \\ m^2 \tilde{B}(\hat{s}, \mu_B) &= m^2 B(\hat{s}, \mu_B) + \frac{C_F \alpha_s(\mu_B)}{4\pi} \left\{ \xi_B h^{\text{corr}} \delta(\hat{s}) + b^{\text{corr}} \frac{1}{m} \left[ \frac{1}{\hat{s}/m} \right]_+ \right\}. \end{aligned} \quad (3.51)$$

where we have now also dropped the current superscript  $C$  from  $h^{\text{corr}}$  and  $b^{\text{corr}}$ . The  $\mathcal{O}(\alpha_s)$  delta function correction  $h^{\text{corr}}$  and the plus function correction  $b^{\text{corr}}$  read

$$\begin{aligned} R_0^{\text{newsing}} h^{\text{corr}} &= A_e^{\text{abs}} + \log(\hat{m}^2 r_c) R_0^{\text{newsing}} b^{\text{corr}} - R_0^{\text{abs}} A_e^{\text{base}} \\ R_0^{\text{newsing}} b^{\text{corr}} &= B_{\text{plus}}^{\text{abs}} - R_0^{\text{abs}} B_{\text{plus}}^{\text{base}}. \end{aligned} \quad (3.52)$$

where the logarithm comes from rescaling of the plus function argument using Eq. (B.6), because all  $B_{\text{plus}}$  coefficients are extracted from formulas where all plus functions appear with the standard argument  $e - e_{\text{min}}$ , see the discussions below Eq. (3.6), while we used  $\hat{s}/m = (e - e_{\text{min}})/(\hat{m}^2 r_c)$  in the plus function argument<sup>2</sup> in Eq. (3.51). This conversion of plus function arguments generates terms in the delta function coefficient. The last terms of the equations in Eq. (3.52) subtract  $\mathcal{O}(\alpha_s)$  mixed terms generated by the global multiplicative correction  $R_0^{\text{newsing}} = 1 + R_0^{\text{abs}}$  in the first line of Eq. (3.51).

The soft function  $S_\tau$  will not be modified, because  $\hat{m}$  corrections should factorize out of the soft function, since fluctuations at the mass scale were already integrated out. Our focus is only on the  $m^2/Q^2$  power corrections since these are numerically the largest ones. Other possible power corrections  $\sim \Lambda_{\text{QCD}}/Q$  or  $\sim \Lambda_{\text{QCD}}/m$  will not be treated separately, since they are much smaller compared to the  $m^2/Q^2$  power corrections.

The delta function correction  $h^{\text{corr}}$  cannot be uniquely assigned and has to be split between the hard matching coefficients and the jet function. The fractional contributions to  $\tilde{H}_m$  and  $\tilde{B}$  are governed by the parameters  $\xi_J$  and  $\xi_B$  respectively. These parameters are in principle not restricted to the range  $[0,1]$  since the term  $1 - \xi_J - \xi_B$  in  $\tilde{H}_Q$  will automatically compensate so that all contributions add to 1, but we will use the sensible restriction  $0 < \xi_i < 1$  and  $\xi_J + \xi_B < 1$  to estimate this uncertainty in our numerical analysis.

---

<sup>2</sup>The rescaling factor  $r_c$  only appears when including additional power correction rescalings in the bHQET measurement function as shown in Eq. (6.22) and Eq. (6.23). The factor  $r_c$  is set to 1 if the “rescale” prescription discussed in Chap. 6 is turned off. Our description here has been formulated to be fully compatible with the “rescale” prescription. The formulas are general and all changes happen implicitly. The origin of all differences for the rescale prescription, besides the  $r_c$  in the  $\hat{s}$ , lie in the coefficients  $A_e^{\text{base}}$ ,  $A_e^{\text{abs}}$ ,  $B_{\text{plus}}^{\text{base}}$  and  $B_{\text{plus}}^{\text{abs}}$ , which will contain additional terms  $\log(r_s)$  and  $\log(r_c)$ , which are 0 in the “no rescale” treatment. In the “rescale” case  $A_e^{\text{base}}$  and therefore also  $A_e^{\text{abs}}$  and  $h^{\text{corr}}$  will acquire an explicit  $\log(\mu/Q)$  dependence in a next to leading power contribution (see Eq. (6.22)). These terms will then be evaluated at the same scale as the strong coupling constant.

## 4. Full Event Shape Distribution

### 4.1. Combining Ingredients

The full event shape distribution, that we use for the fits, is a sum of the bHQET cross section  $\sigma_s$ , the SCET matching contribution  $\sigma_{\text{nsb}}$  and the QCD matching contribution  $\sigma_{\text{ns}}$ , presented in Chap. 2 and Chap. 3:

$$\begin{aligned} & \frac{1}{\sigma_0} \frac{d\sigma(e)}{de} \\ &= \frac{1}{\sigma_0} \int d\hat{s} dk \frac{d(\hat{\sigma}_s + \hat{\sigma}_{\text{nsb}} + \hat{\sigma}_{\text{ns}})}{de} \left( e - e_{\text{min}} - \frac{de_{\text{min}}}{dm} \delta m - \frac{m\hat{s}}{Q^2} - \frac{k + 2\bar{\delta}}{Q} \right) \\ & \quad \times G(\hat{s}, \Gamma) F(k - 2\bar{\Delta}), \end{aligned} \quad (4.1)$$

where it is understood that, other than explicit appearances in the arguments shown,  $e_{\text{min}} = \delta_m = \bar{\delta} = \Gamma = 0$  and that the mass appearing in the argument shown and the denominator of  $\hat{s}_\tau = Q^2(e - e_{\text{min}} - \dots)/m$  refers to  $m_{\text{MSR}}(R)$  in the MSR mass scheme and to  $m_{\text{pole}}$  in the pole mass scheme. Interpreting the mass in the denominator of  $\hat{s}_\tau$  as  $m_{\text{pole}} = m_{\text{MSR}} + \delta m$  in the MSR mass scheme would in any way lead to power suppressed contributions after expanding in the  $\delta m \sim \hat{s} \sim \Gamma_t \ll m_t$  bHQET power counting, which are small compared to the more problematic  $m^2/Q^2$  power corrections. The MSR mass subtraction and the gap subtraction have to be expanded strictly in  $\alpha_s$  together with the hard, jet and soft function to guarantee a correct order-by-order cancellation of the renormalons in the jet and soft function. The same applies to the nonsingular parts. The correct way to cancel the pole renormalon in pure bHQET would be to drop all power corrections in  $de_{\text{min}}/dm$ , but since we are using the exact threshold  $e_{\text{min}}$  for the singular distribution and since we are matching to full QCD we also have to include the power corrections originating from the exact threshold in the subtraction series. The mass parameter in the hard function  $H_m$  can be either expressed in the pole or  $\overline{\text{MS}}$  mass and the choice of mass scheme for the boost factor  $\varrho = Q/m$  in  $H_m$  and the running factor  $U_v$  is ambiguous with respect to bHQET power corrections. The differences of these choices are found to be numerically small [18], so we have used the unexpanded pole mass for these factors, see discussion in [14].

The function  $G(\hat{s}, \Gamma)$  is the Breit Wigner function appearing in Eq. (2.10) and  $F(k - 2\bar{\Delta})$  is the nonperturbative shape function, defined in Eq. (2.13), with the renormalon free gap parameter. These two corrections have to be also applied to the nonsingulars. The terms  $d\sigma_{\text{nsb}}/de$  and  $F_e^{\text{NS},C}$ , defined in Eq. (3.39), are regular functions and numerically insignificant at the threshold. We have therefore dropped the convolution of  $G(\hat{s}, \Gamma)$  over them.

order	log terms	cusps	non-cusps	matching	$\beta[\alpha_s]$	$\gamma_R$	$\delta$
LL	$\alpha_s^n L^{n+1}$	1	-	tree	1	-	-
NLL	$\alpha_s^n L^n$	2	1	tree	2	1	-
N <sup>2</sup> LL	$\alpha_s^n L^{n-1}$	3	2	1	3	2	1

Table 4.1.: Required loop orders for resummation of  $\alpha_s^n L^{n+1-k} = \mathcal{O}(\alpha_s^{k-1})$  logarithmic terms and same order N<sup>k-1</sup>LO boundary conditions [30]. Cusp and non-cusp anomalous dimensions and beta function coefficients are given in App. A.1. Matching refers to the fixed order contributions, i.e. the bHQET and SCET-nonsingular boundary conditions and the QCD-nonsingular given in Sec. 2.1 and Chap. 3. The R-anomalous dimension  $\gamma_R$  and the renormalon subtraction series  $\delta$  refer to both the soft gap and pole mass renormalon.

For our fits we use next-to-leading-logarithmic resummation (NLL) + LO boundary conditions and N<sup>2</sup>LL + NLO. The required loop orders for the ingredients are listed in Tab. 4.1. Note that in our code the  $\alpha_s$  and the MSR mass evolutions are fixed to 4-loop and 3-loop respectively.

## 4.2. Profile Functions

The resummed cross section is strictly independent on the scale  $\mu$ . It can therefore be set arbitrarily. But, the dependence on the other renormalization scales  $\mu_i$  ( $i = Q, m, B, S$ ) and  $R_j$  ( $j = m, S$ ) only cancels order-by-order in  $\alpha_s$ . To estimate the uncertainty induced by the choice of scales and to guarantee proper resummation of large logarithms, we will use so-called profile functions. These are in general  $\tau$ -dependent<sup>1</sup> renormalization scales, that are determined by the characteristic scales of the kinematics and dynamics in each sector of the factorization theorem across different regions of the event shape distribution. We adopt the parametrization for massive profile functions constructed in Ref. [29] and their value ranges.

The hard function and mass matching scale are constant and their variation, given by the rescaling parameter  $e_H$ , is correlated to retain correct hierarchies between scales.

$$\begin{aligned}\mu_H &= e_H Q \\ \mu_m &= \sqrt{e_H} m.\end{aligned}\tag{4.2}$$

All mass parameters in the profile functions use the profile mass, which we fix to the MSR mass at  $R = 5$  GeV. The  $\tau$ -dependent jet and soft function scales are given by piecewise functions, which describe the *nonperturbative* region ( $\tau < t_0$ ), the *resummation* region ( $t_1 < \tau < t_2$ ) and the *fixed order* region ( $\tau > t_s$ ). In the nonperturbative region the scales are frozen at a low but still perturbative value. In the resummation region the scales grow steadily and in the fixed order region the scales are merged with the hard function scale  $\mu_H$ . These three regions are connected by transition regions which allow the piecewise

<sup>1</sup>We use  $\tau$  and  $e$  interchangeably for a generic event shape variable.

functions  $F(\tau < t_0)$  and  $G(\tau > t_1)$  for  $t_0 < t_1$  to be smoothly joined (i.e. the function and their first derivative are continuous) by double quadratic functions

$$\begin{aligned} & \zeta [\{F := F(t_0), f := F'(t_0)\}, \{G := G(t_1), g := G'(t_1)\}, t_0, t_1, \tau] \\ &= \begin{cases} F + f(\tau - t_0) + \frac{4(G-F) + (3f+g)(t_0-t_1)}{2(t_0-t_1)}(\tau - t_0)^2 & t_0 \leq \tau < \frac{t_1+t_0}{2} \\ G + g(\tau - t_1) - \frac{4(G-F) + (f+3g)(t_0-t_1)}{2(t_0-t_1)}(\tau - t_1)^2 & \frac{t_1+t_0}{2} \leq \tau \leq t_1 \end{cases} \end{aligned} \quad (4.3)$$

The region boundaries are given by

$$\begin{aligned} t_0 &= \frac{2}{Q/1 \text{ GeV}} + \frac{d_0}{(Q/1 \text{ GeV})^{0.5}} + \tau_{\min} \\ t_1 &= \frac{2.25}{(Q/1 \text{ GeV})^{0.75}} + \frac{d_1}{(Q/1 \text{ GeV})^{0.5}} + \tau_{\min} \\ t_b &= \tau_{\min} + \frac{\hat{m}^2}{r_{\text{slope}} e_H} \\ t_2 &= n_2 + \hat{m} \\ t_s &= n_s + \hat{m}, \end{aligned} \quad (4.4)$$

where  $t_b$  is the transition point for switching from bHQET to the SCET description. We only need the distribution in the nonperturbative region and the transition region to the resummation region, since we fit only in the peak region. It is therefore sufficient to specify the profile functions up to  $t_b$ .

The soft function scale is given by

$$\mu_S(\tau) = \begin{cases} \mu_0 & \tau < t_0 \\ \zeta [\mu_S(\tau < t_0), \mu_S(\tau > t_1), t_0, t_1, \tau] & t_0 \leq \tau \leq t_1 \\ \left(1 + \frac{n_s e_S}{n_s + \hat{m} - \tau_{\min}}\right) r_{\text{slope}} \mu_H(\tau - \tau_{\min}) & t_1 < \tau < t_b. \end{cases} \quad (4.5)$$

The bHQET jet function scale reads

$$\mu_J(\tau) = \begin{cases} 1 + \tilde{e}_J(t_s - t_0)^2 \tilde{\mu}_J(t_0) & \tau < t_0 \\ \zeta [\mu_J(\tau < t_0), \mu_J(\tau > t_1), t_0, t_1, \tau] & t_0 \leq \tau \leq t_1 \\ 1 + \tilde{e}_J(t_s - \tau)^2 \tilde{\mu}_J(\tau) & t_1 < \tau < t_b, \end{cases} \quad (4.6)$$

where  $\tilde{\mu}_J(\tau) = \sqrt{e_H} \mu_S(\tau) / \hat{m}$  and

$$\tilde{e}_J = e_J \left[ \frac{n_s - (t_0 - \tau_{\min})}{t_s - t_0} \right]^2 \quad (4.7)$$

The renormalization scale of the MSR mass  $m_t^{\text{MSR}}(R)$  is set to the jet scale to avoid large logarithms

$$R_m(\tau) = \mu_J(\tau). \quad (4.8)$$

$n_2$	0.225 to 0.275
$n_s$	0.375 to 0.425
$d_0$	-0.05 to 0.05
$d_1$	-0.05 to 0.05
$\mu_0$	2.5 to 3.5
$r_{\text{slope}}$	2
$e_H$	0.5 to 2
$e_S$	$1/1.13 - 1$ to $1.13 - 1$
$e_J$	-3 to 0
$n_{\text{ns}}$	-1, 0, 1

Table 4.2.: Parameter ranges for scale variation

The subtraction scale of the soft gap subtraction should in general also be set close to the soft scale to avoid generating large logarithms. But it has been observed in Refs. [29, 30] that for gap scheme 1 a nonvanishing logarithm at order  $\alpha_s$  is favorable for the soft renormalon cancellation in the peak region, since the constant term of the one loop subtraction for this scheme is zero. In this case the subtraction scale has to be set strictly below the soft scale to obtain a nonvanishing and correctly signed order  $\alpha_s$  subtraction term

$$R_S(\tau) = \begin{cases} 0.75\mu_0 & \tau < t_0 \\ \zeta [R_S(\tau < t_0), R_S(\tau > t_1), t_0, t_1, \tau] & t_0 \leq \tau \leq t_1 \\ \mu_S(\tau) & t_1 < \tau < t_b \end{cases} \quad (4.9)$$

The other two schemes have finite constant terms at order  $\alpha_s$ , hence a more natural choice for them would be to set  $R_S(\tau)$  exactly to the soft scale  $\mu_S(\tau)$ . We have for simplicity adopted the scheme 1 choice also for the other schemes. Fig. C.1 (in App. C.1) shows a comparison of our final results for the default parameter settings presented here and with the modification  $R_s = \mu_S$ . The differences are discussed in App. C.1. The summary is, that the fitted mass values are not affected by the choice. The fitted  $\Omega_1$  values for gap 2 and gap 3 schemes also show insignificant differences, while setting  $R_s = \mu_S$  for gap 1 scheme results in significant changes of the fitted  $\Omega_1$ .

The QCD nonsingular scale is set to a weighted average between hard and jet scale

$$\mu_{\text{ns}}(\tau) = \mu_H + \frac{n_{\text{ns}}}{2}(\mu_J(\tau) - \mu_H). \quad (4.10)$$

The 501 profiles used in the fits are obtained by random sampling parameters in the ranges given in Tab. 4.2. Each profile contains one value for each parameter.

# 5. Fit Procedure Improvements

## 5.1. Fixed $\Delta_0$ Fits

### 5.1.1. Fit Procedure

We follow a standard  $\chi^2$  fit for the top quark mass  $m_t$  and the nonperturbative model parameters  $c_0 - c_3$  employed in Ref. [17], which we outline in the following. The  $c_i$  are restricted by  $c_i^2 = 1$ , so the actual fit parameters are three euclidean angles  $\{a\} = (a_0, a_1, a_2)$ . The renormalon free gap parameter

$$\Delta_0 \equiv \bar{\Delta}(R_\Delta, \mu_\Delta) \quad (5.1)$$

at the reference scales  $R_\Delta = \mu_\Delta = 2 \text{ GeV}$  is fixed at  $0.05 \text{ GeV}$  and  $\lambda = 0.5$ , unless otherwise stated. The results are presented in the variables  $m_t$ , the mean  $\Omega_1(R_\Delta, \mu_\Delta) = \int dk \frac{k}{2} F(k - 2\Delta_0)$  and the variance  $\Omega'_2 = \int dk (\frac{k}{2} - \Omega_1(R_\Delta, \mu_\Delta))^2 F(k - 2\Delta_0)$  of the nonperturbative model function (also called first and second cumulants) at the reference scales. The factors of  $1/2$  adjust these results to single jet parameters. The values of  $\Omega_1$  in the plots will always be shown converted to the one loop gap 1 scheme by adding the difference of the one loop subtraction to provide better comparability of the results:

$$\Omega_{1,\text{plots}} = \Omega_{1,\text{gap } i}(R_\Delta, \mu_\Delta) + [\bar{\delta}_{\text{gap } i}(R_\Delta, \mu_\Delta) - \bar{\delta}_{\text{gap } 1}(R_\Delta, \mu_\Delta)]_{\text{one loop}} \quad (5.2)$$

The red lines in the plots are given by the inverse of this conversion for a reference value of  $\Omega_{1,\text{plots}} = 0.5 \text{ GeV}$  and should serve as a visual indication of how large the conversion correction in Eq. (5.2) is:

$$\Omega_{1,\text{redline,gap } i} = 0.5 \text{ GeV} - [\bar{\delta}_{\text{gap } i}(R_\Delta, \mu_\Delta) - \bar{\delta}_{\text{gap } 1}(R_\Delta, \mu_\Delta)]_{\text{one loop}} \quad (5.3)$$

The reference data were binned distributions of either 2-Jettiness, sJM or mJM, which we will here refer to as  $\tau$ , obtained from the Monte Carlo generator for a  $e^+e^- \rightarrow t\bar{t}$  process, each histogram contains  $10^7$  events. We used three different fit ranges around the peak of the distribution. The fit range can be written as  $(x, y)$ , with the minimum and maximum value  $\tau_{\text{min}}$  and  $\tau_{\text{max}}$  defined as the position where the distribution drops to a value given by a fraction of the value at the peak:

$$\frac{d\sigma(\tau_{\text{min}})}{d\tau} = x \frac{d\sigma(\tau_{\text{peak}})}{d\tau}, \quad \frac{d\sigma(\tau_{\text{max}})}{d\tau} = y \frac{d\sigma(\tau_{\text{peak}})}{d\tau} \quad (5.4)$$

The three ranges used are  $(0.6, 0.8)$ ,  $(0.7, 0.8)$  and  $(0.8, 0.8)$ .

Due to the functional relation between the top quark mass  $m_t$  and the soft convolution variable  $\ell$ , it is necessary to simultaneously include distributions at multiple center of mass

energies  $Q$  in the fit to break the degeneracy between  $m_t$  and  $\Omega_1$ . We used five different sets of  $Q$  values. They read in units of GeV: (700, 1000, 1400), (800, 1000, 1400), (700 – 1400), (600, 1000, 1400) and (600 – 1400). The ranges are in steps of 100 GeV. This gives 3 (ranges around peak)  $\times$  5 ( $Q$  sets) = 15 different “fit settings” of bins included in the fit and the spread of these fit results estimates the “incompatibility” between theory and Monte Carlo event generator description. To estimate the perturbative scale uncertainty we use 501 sets of profile functions  $\mu_i(\tau)$  with parameters randomly varied in the ranges stated in Tab. 4.2.

We followed these steps to obtain a central value and uncertainties for  $m_t$  (and the same for the cumulants):

1. For one fit setting  $i$  remove 1.5% of the upper and 1.5% of the lower values of the 501 best fit values from the variation over the profiles to remove potential outliers. Let us call this cleaned up set  $\{m\}_i$ .
2. Then take the central value  $m_{\text{set},i} = [\max(\{m\}_i) + \min(\{m\}_i)]/2$  as the result for this fit setting and half the range as the scale uncertainty  $\Delta m_{\text{set},i} = [\max(\{m\}_i) - \min(\{m\}_i)]/2$ .
3. Take the central value of the results for the 15 fit settings as the final result for  $m_t$ :  $m_{\text{fit}} = [\max_i(m_{\text{set},i}) + \min_i(m_{\text{set},i})]/2$ .
4. Take the mean of the scale uncertainties for the 15 fit settings as the final *perturbative uncertainty*:  $\Delta m_{\text{pert}} = [\sum_i \Delta m_{\text{set},i}]/15$ .
5. Take half the range of the individual results for the 15 fit settings as *incompatibility uncertainty*:  $m_{\text{incomp}} = [\max_i(m_{\text{set},i}) - \min_i(m_{\text{set},i})]/2$ .

A best fit value for one profile and one fit setting is obtained by minimizing the  $\chi^2$  function with respect to its arguments using the program MINUIT, with

$$\chi^2(m_t, \{a\}) = \sum_Q \sum_{\tau_{\min} \leq \tau_i < \tau_{\max}} \frac{[f_{Q,i}^{\text{theo}}(m_t, \{a\}) - f_{Q,i}^{\text{mc}}]^2}{\sigma_{Q,i}^2} \quad (5.5)$$

The theory bin  $f_{Q,i}^{\text{theo}}(m_t, \{a\})$  at observable value  $\tau_i$  is defined as the difference of the cumulant of the differential cross section at points  $\tau_i$  and  $\tau_{i+1}$ , which we call  $\hat{f}_{Q,i}^{\text{theo}}(m_t, \{a\})$ , divided by the norm  $\sum_i \hat{f}_{Q,i}^{\text{theo}}(m_t, \{a\})$ . The Monte Carlo generator bin  $f_{Q,i}^{\text{mc}}$  is defined as the sum of events with  $\tau_i < \tau < \tau_{i+1}$ , which we call  $\hat{f}_{Q,i}^{\text{mc}}$ , divided by the norm  $\mathcal{N}_Q^{\text{mc}} = \sum_i \hat{f}_{Q,i}^{\text{mc}}$ . Both histograms are normalized to 1 across the fit range  $(\tau_{\min}, \tau_{\max})$ . The error  $\sigma_{Q,i}$  is the statistical error of the event generator bin  $f_{Q,i}^{\text{mc}}$  obtained by naively dividing the bin errors  $\Delta \hat{f}_{Q,i}^{\text{mc}}$  of the unnormalized bins  $\hat{f}_{Q,i}^{\text{mc}}$  by the norm  $\mathcal{N}_Q^{\text{mc}}$ . This “naive” bin error  $\sigma_{Q,i}$  ignores correlations between bins that are introduced through the division of the original statistical independent errors  $\Delta \hat{f}_{Q,i}^{\text{mc}}$  by the norm  $\mathcal{N}_Q^{\text{mc}}$ .

We also tested the strict statistical treatment of performing the fits with the  $\chi^2$  values obtained by using the full covariance matrix for the normalized bins. The differences to the naive treatment of Eq. (5.5) for the fitted mass were at the sub-MeV level for individual profile fits. In light of the negligible differences, we concluded that using the naive bin errors for calculating the  $\chi^2$  is sufficient.

The error bars shown in the results are obtained by quadratically adding the perturbative and incompatibility uncertainty. The  $\chi^2/\text{dof}$  value quoted is an average over all profiles and fit settings. The MSR mass results are quoted at the scale  $R = 1 \text{ GeV}$ .

$\alpha_s(m_Z)$  on the theory side was fixed at the world average of 0.1181(13). Variations within its errors induces an additional uncertainty in  $m_{\text{fit}}$  of  $\simeq 20 \text{ MeV}$ , which is not shown in the results due to its negligible size compared to the other uncertainties.

The parameters of the Monte Carlo event generators were set to  $m_t^{\text{MC}} = 173 \text{ GeV}$ . The top quark width was fixed for theory and event generator to  $\Gamma_t = 1.4 \text{ GeV}$ . The generators use a tree level  $e^+e^- \rightarrow t\bar{t}$  matrix element, which goes through their respective internal standard decayer, parton shower and hadronization model. Initial state radiation has been turned off. For PYTHIA 8.305 [31] and SHERPA 2.2.11 [32] we use the default settings and the default tune. The standard HERWIG 7.2 [33] exhibited some unusual flattening of its peak for smaller center of mass energies  $Q \simeq 600 - 1000 \text{ GeV}$ . We therefore received an unpublished preliminary version from its developers [34], that fixes this issue, and has its parameters re-tuned. That version is referred to as ‘‘FullTune’’ in our results.

## Details on Data Processing

The input files for the event generators were modified according to the settings described above, all other settings were kept at their standard values. We use the program RIVET [35] paired with a custom written analysis to convert per event kinematic information into histograms in the format YODA for our observables. This workflow works with all leading Monte Carlo event generators, that support RIVET directly or the event record format HEPMC [36]. The MC produces events across the full event shape range and the choice of the bin specification has no impact on the MC runtime. It is therefore safer to keep a large range and use narrow bins, since wider bins can always be produced by merging smaller bins. For the histograms we used 10000 evenly spaced bins between 0.0 and 0.5 for each of our observables and all  $Q$  values. This was also the width of the bins we used in Eq. (5.5), since there was no obvious problem with this choice.

The theory cross section is generated in a Fortran program CALIPER [37], which we modified if necessary with our own code. The following steps have to be repeated for each fit setting and each profile. We will in the following assume that the profile parameters have been fixed. One can write the cross section as  $\sum_{kl} c_k c_l f_{kl}(\tau; m_t, Q)$  using the double sum in the shape function, as defined in Eq. (2.13), over the basis functions. We can therefore save  $f_{kl}(\tau; m_t, Q)$  evaluated on a grid in  $m_t, Q, \tau, (k, l)$  (with  $f_{kl} = f_{lk}$ ), and evaluate the sum over the  $c_i$  analytically during the calculation of the  $\chi^2$  fit function. Recall that the shape function fit parameters are actually the euclidean angles  $a_j$  in  $c_i(\{a_j\})$  due to the normalization condition of the  $c_i$ 's. Hence, the cross section will depend on sines and cosines of these angles, which means that one has to sample multiple

starting values to reliably find the true minimum of the fit function.

The  $\tau$  values of the grid lie in a range between 0 and  $t_1(m_t = 177 \text{ GeV}, Q, d_1 = 0.25)$  which is defined in Eq. (4.4).  $t_1$  with the given arguments will always be greater than the boundary of our fit ranges  $\tau_{\max}$ , which is given below Eq. (5.4). The range of our MC histogram has also been chosen so that  $t_1$  will always lie inside it. The  $\tau$  values of the theory grid do not have to coincide with the MC histogram bin boundaries, because we will interpolate over the  $\tau$  values and then obtain a interpolation of the cumulant from which the final theory bins, with matching boundaries, are calculated. To determine appropriate  $\tau$  values for our grid we first find the peak  $t_{\text{peak}}$  of the  $f_{00}$  distribution with the Fortran routine `COMPASS_SEARCH` [38] using the on-shell top threshold  $\tau_{\min}$  as starting value. Then we generate 15 evenly spaced points left of  $t_{\text{peak}} - 0.4(t_1 - t_{\text{peak}})$ , the next interval up to  $t_{\text{peak}} + 0.4(t_1 - t_{\text{peak}})$  is filled with 75 evenly spaced points, and the rest up to  $t_1$  has 10 evenly spaced data points (each point contains  $\tau$ , and the  $f_{kl}$ 's). We checked that this provided sufficient points around the peak to offer a good interpolation quality for the fits.

The other dimensions of the grid were the  $Q$  values as given above and the mass (either  $m_{\text{pole}}$  or  $m_{\text{MSR}}$ ) in steps of 0.25 GeV between 170 GeV and 175 GeV.

The fits themselves were performed using a python script. At the beginning of the fit one determines from the MC histograms the range of bins that satisfy the ‘‘above  $(x, y)$  fraction of peak height’’ condition in Eq. (5.4) for each  $(x, y)$  and  $Q$  combination. The  $f_{kl}(\tau, m, Q)$  are then turned into bins  $f_{kli}(m, Q)$  with the same specifications as the MC bins by using an integral over a spline interpolation. Before the interpolation over  $\tau$  we clean numerical outliers from the data, if there are any, but this might not be necessary since outliers are also removed after the fits and one has to be careful for  $l \neq k$ , because those functions are oscillatory. Finally, these bins are spline interpolated over  $m_t$ . The  $\chi^2$  in Eq. (5.5) with

$$f_{Q,i}^{\text{theo}}(m_t, \{a\}) = \sum_{kl} c_k(\{a\}) c_l(\{a\}) f_{kli}(m_t, Q) \quad (5.6)$$

can then be smoothly sampled by MINUIT, without having to calculate the theory bins from scratch for each evaluation.

### 5.1.2. Previous Calibration Results for PYTHIA

The original calibration of the PYTHIA MC top mass was carried out in Ref. [17]. The same methodology as described in the previous section was applied to extract the top mass. They used the observable 2-Jettiness and the gap 1 scheme defined in Eq. (2.24) for the soft renormalon subtraction.

Their results are quoted in the left half of Tab. 5.1. They showed that  $m_{t,\text{PYTHIA}}^{\text{MC}}$  is indeed close to the MSR mass at  $R = 1 \text{ GeV}$  and not a tree level or one-loop pole mass. The perturbative uncertainties decreased when using NNLL instead of NLL theory, but the incompatibility uncertainties remained comparable, as was expected. The pole mass exhibited a large correction between orders, which is associated with the  $\mathcal{O}(\Lambda_{\text{QCD}})$

	order	central	perturb.	incomp.	central	perturb.	incomp.
$m_{t,1\text{ GeV}}^{\text{MSR}}$	N <sup>2</sup> LL	172.82	0.19	0.11	172.82	0.17	0.10
$m_{t,1\text{ GeV}}^{\text{MSR}}$	NLL	172.80	0.26	0.14	172.83	0.29	0.12
$m_t^{\text{pole}}$	N <sup>2</sup> LL	172.43	0.18	0.22	172.40	0.18	0.20
$m_t^{\text{pole}}$	NLL	172.10	0.34	0.16	172.06	0.34	0.16
$\Omega_1$	N <sup>2</sup> LL	0.42	0.07	0.03	0.43	0.06	0.03
$\Omega_1$	NLL	0.41	0.07	0.02	0.42	0.07	0.03

Table 5.1.: Calibration results from Ref. [17] for MSR( $R = 1\text{ GeV}$ ) mass, pole mass and  $\Omega_1$ ( $R = 2\text{ GeV}$ , from simultaneous MSR mass fit) for PYTHIA 8.205 with  $m_t^{\text{MC}} = 173\text{ GeV}$  (left) and our new fits for PYTHIA 8.305 (right). Shown are central values, perturbative and incompatibility uncertainties, all in GeV.

renormalon, while the MSR mass allowed a stable mass parameter extraction since it does not suffer from these infrared corrections.

The right half of Tab. 5.1 shows our new fits for the current PYTHIA version 8.305, which are up to rounding errors identical to the values that we obtained for an updated fit, using our current setup, for the old version 8.205. The new results and those from Ref. [17] are practically identical, which means that the changes made between both fits only had a marginal effect. The differences, apart from the newer PYTHIA version were:

1. As noted below Eq. (3.4), the renormalization scale of nonsingular terms in the jet function present in SCET, but not in bHQET, are now run down to the same scale  $\mu_B$  as the distributional terms in bHQET, rather than being frozen at the mass mode matching scale  $\mu_m$ . This effect is small since the contribution of this nonsingular is small.
2. The interpolation over  $m_t$  is now happening at the bin level and all parameters are simultaneously minimized. The method used before was to first minimize with respect to the shape function parameters at fixed  $m_t$  giving  $\chi^2(m_t, \{a_{\min}(m_t)\})$ , then interpolating this over  $m_t$  and finding the minimum with respect to  $m_t$ . Both methods are equivalent if the grid in  $m_t$  is fine enough.
3. The fit of [17] included 2 additional  $Q$  sets: (600 – 900) and (700 – 1000). We will see in Sec. 5.1.4 that the shape function was not given enough freedom in its fit parametrization to vary its  $\Omega_1$ . After fixing this issue these 2  $Q$  sets which are very restricted in their  $Q$  range will not be able to break the degeneracy between  $m_t$  and  $\Omega_1$  well enough and lead to an increased spread of the individual best fit results and an inaccurate estimate of the final uncertainties. For the old fit parametrization on the other hand, removing these  $Q$  sets has minimal effects on the end results because the inflexibility of  $\Omega_1$  made the fits insensitive to the  $Q$  dependence of the thrust peak position.

The PYTHIA 8.305 results from the fit of Ref. [17] will be shown as reference values in the majority of plots in later parts of this thesis. They are labeled as “pythia8305”.

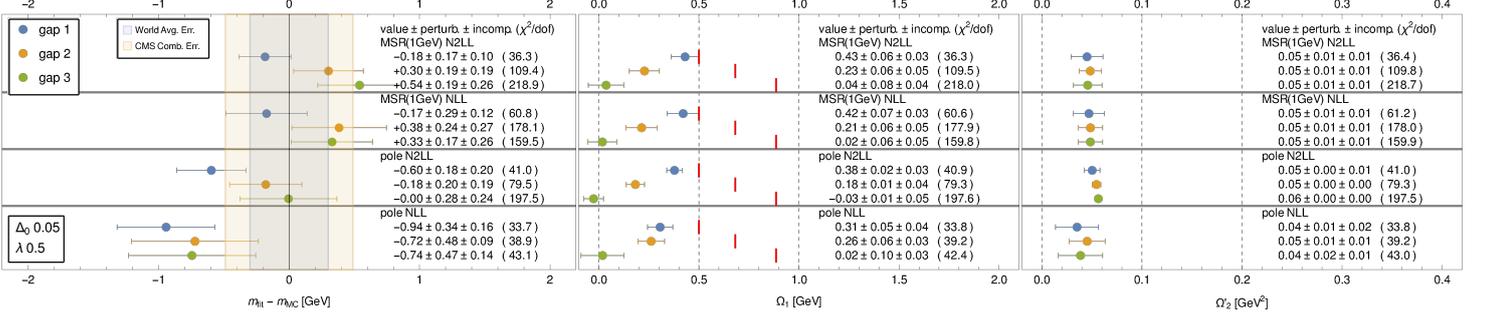


Figure 5.1.: Gap dependence of fitted top mass  $m_{\text{fit}}$  (either  $m_t^{\text{MSR}}(R = 1 \text{ GeV})$  or  $m_t^{\text{pole}}$ ), the first moment  $\Omega_1$  converted to the 1-loop gap 1 scheme, see Eq. (5.2), and the variance  $\Omega_2'$  of the shape function for N<sup>2</sup>LL and NLL orders. The corrected  $\Omega_1$  values for the same mass scheme and order don't agree with each other, because the shape function failed to adapt to the change in gap scheme. Shifts in  $m_{\text{fit}}$  partly compensate for the unchanging shape function, but due to the different functional dependence on  $m_{\text{fit}}$  and  $\Omega_1$  the  $\chi^2$  successively worsens.

### 5.1.3. Gap Dependence

The first thing that has to be addressed is the insufficient flexibility of the original fit parametrization of the non-perturbative shape function. For that we repeated the original calibration shown in Tab. 5.1, that was using gap 1, also for gap 2 and gap 3. The fitted parameters are plotted in Fig. 5.1. All fits in this chapter use the 2-Jettiness observable, unless explicitly stated otherwise. The  $\Omega_1$  plotted is converted to the 1-loop gap 1 scheme as explained in Eq. (5.2), so that at least the N<sup>2</sup>LL  $\Omega_1$  values for a fixed mass scheme should agree with each other. But what we found is that the shape function has not adapted at all to changes of the gap scheme, so the converted  $\Omega_1$  exhibits shifts that are exactly opposite to the expected change of the original values. The position of the red vertical lines in these plots as defined in Eq. (5.3) give the  $\Omega_1$  value in the original scheme for a reference converted  $\Omega_{1,\text{gap1}} = 0.5 \text{ GeV}$ . The inability of the shape function to change its first moment forces the mass parameter  $m_{\text{fit}}$  to compensate for the change in gap scheme, which makes the fitted mass values meaningless. One can also see that the obtained  $\chi^2$  values worsens quite dramatically for gap schemes that are more different. That is because the functional dependence of the cross section on the mass and the soft convolution variable is given in a leading approximation by the combination  $2\hat{m}^2 + \ell/Q$ , i.e.  $\delta m \hat{\simeq} -\frac{Q}{2\hat{m}} \delta \Omega_1$ . Thus,  $m_{\text{fit}}$  can only compensate in an average fashion for bins with different  $Q$  values.

### 5.1.4. $\Delta_0$ -Dependence

A parameter that is naturally connected to the first moment of the shape function and the effects of the gap scheme is the renormalon free gap parameter  $\overline{\Delta}(R_s, \mu_\delta)$  defined in Eq. (2.18) and we define its value at the reference scale as  $\Delta_0 \equiv \overline{\Delta}(R_\Delta = \mu_\Delta = 2 \text{ GeV})$ . If we take the limit of using infinite basis functions of the nonperturbative model functions, that is letting  $N \rightarrow \infty$  in Eq. (2.13), then all moments of the shape functions can be accurately modeled without the need to explicitly insert a reference gap, that is  $F[k - 2(\overline{\Delta}(R_s, \mu_\delta) - \Delta_0) - 2\Delta_0; \lambda, c_1, \dots, c_\infty] \simeq F[k - 2(\overline{\Delta}(R_s, \mu_\delta) - \Delta_0); \lambda, c'_1, \dots, c'_\infty]$  where  $\overline{\Delta}(R_s, \mu_\delta) - \Delta_0$  is just the gap running difference. In this limit  $\Delta_0$  in

$$\Omega_1(R_\Delta, \mu_\Delta) = \Delta_0 + \lambda(0.5c_0^2 + 0.47c_0c_1 + \dots) \quad (5.7)$$

is a superfluous parameter degenerate to a combination of  $c_i$ 's. It would only be required to at least be smaller than the first moment. Since we are only using 4 non-zero  $c_i$ s, which amounts to 3 free parameters due to the normalization, it is possible that they are unable to model the correct peak shape and peak position at the same time.

Fig. 5.2 shows the capability of the fitted  $c_i$ 's to compensate for changes in  $\Delta_0$ . What we see is that the  $c_i$  dependent part of Eq. (5.7) is not adapting at all to changes of  $\Delta_0$ . This means that the  $c_i$ 's are already determined by higher order moments of the shape function and they effectively behave completely orthogonal to  $\Delta_0$ . The only fitted parameter that reacted to the  $\Delta_0$  change is the mass parameter  $m_{\text{fit}}$ . As explained at the end of the last subsection, a change of the first moment can be compensated by a  $Q$  dependent shift of the mass  $\delta m \simeq -\frac{Q}{2m} \delta \Omega_1$ . Since we are fitting simultaneously to multiple  $Q$  values, the shift can only be done for an intermediate  $Q$  value, which explains the clear increase of  $\chi^2$  the farther  $\Omega_1$  and  $m_{\text{fit}}$  are from their optimal values. The changes of  $m_{\text{fit}}$  are relatively large even for small changes of  $\Delta_0$ . We therefore conclude, that  $\Delta_0$  has to be treated as a fit parameter to remove any bias in the fitted mass values.

### 5.1.5. $\lambda$ -Dependence

We have also tested the dependence of the fitted parameters on the  $\lambda$  parameter of the shape function in Eq. (2.13) at fixed  $\Delta_0 = 0.05 \text{ GeV}$ .

The results for gap 1 are shown in Fig. 5.3. For the MSR mass fits  $\lambda = 0.5 \text{ GeV}$  turns out to be too small. For larger values of  $\lambda$  the fitted MSR mass is stable over a large range with small increases of the uncertainty in the negative direction.

The pole mass fits are apparently very correlated with the choice of  $\lambda$ . The  $\Delta_0$ -fit data point shows that the problem lies in the value of the fixed  $\Delta_0$ , which in combination with restrictions from higher order moments effectively sets a lower bound for  $\Omega_1$ . Increasing  $\lambda$  and thus the width of the basis functions of the shape function only worsens this problem. Further tests in Sec. 5.2.2 on the  $\lambda$  dependence in combination with simultaneous  $\Delta_0$  fits will show that  $\lambda$  can be chosen in a reasonably large range with negligible effects on  $m_{\text{fit}}$ .

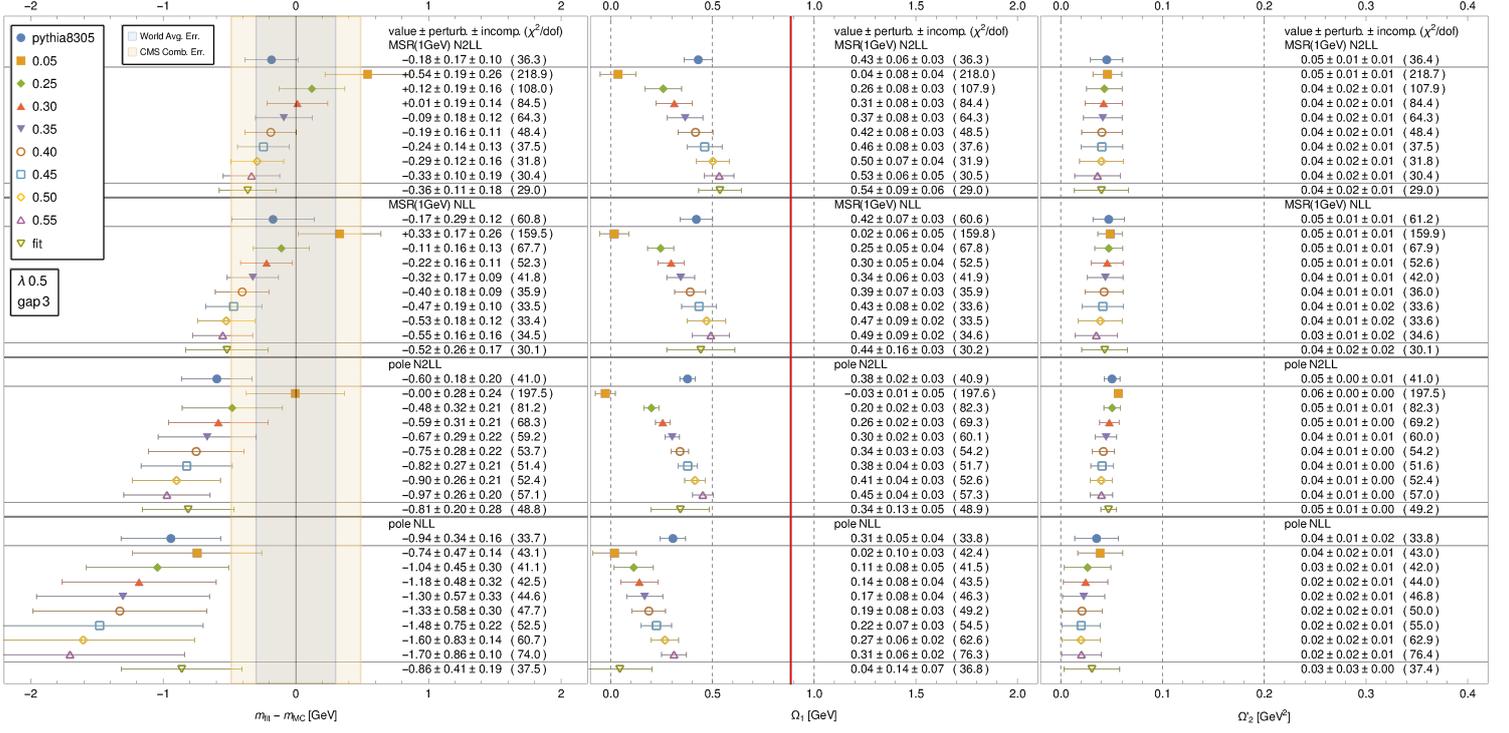


Figure 5.2.: Dependence of fitted parameters on  $\Delta_0$  values in the range 0.05 – 0.55 GeV using gap 3 scheme. “pythia8305” is the gap 1 scheme result with  $\Delta_0 = 0.05$  GeV listed on the right half of tab. 5.1 and is shown as reference. The “fit” data point is obtained by also fitting for  $\Delta_0$ . We see that the other fit parameters of the shape function are unable to compensate for changes of  $\Delta_0$ . In contrast, the fitted mass is strongly correlated with the value of  $\Delta_0$ . The optimal values can clearly be determined by minimizing the  $\chi^2$ .

## 5.2. Variable $\Delta_0$ Fits

### 5.2.1. Fit Procedure Modification

We have demonstrated in the previous section, that the gap parameter  $\Delta_0$  at the reference scales as defined in Eq. (5.1) also has to be fitted for, so that the shape and the position of the peak can be correctly described. We do this by adding this additional fit parameter to the simultaneous fits for  $m_{\text{fit}}$  and the shape function coefficient angles  $a$ . This modifies

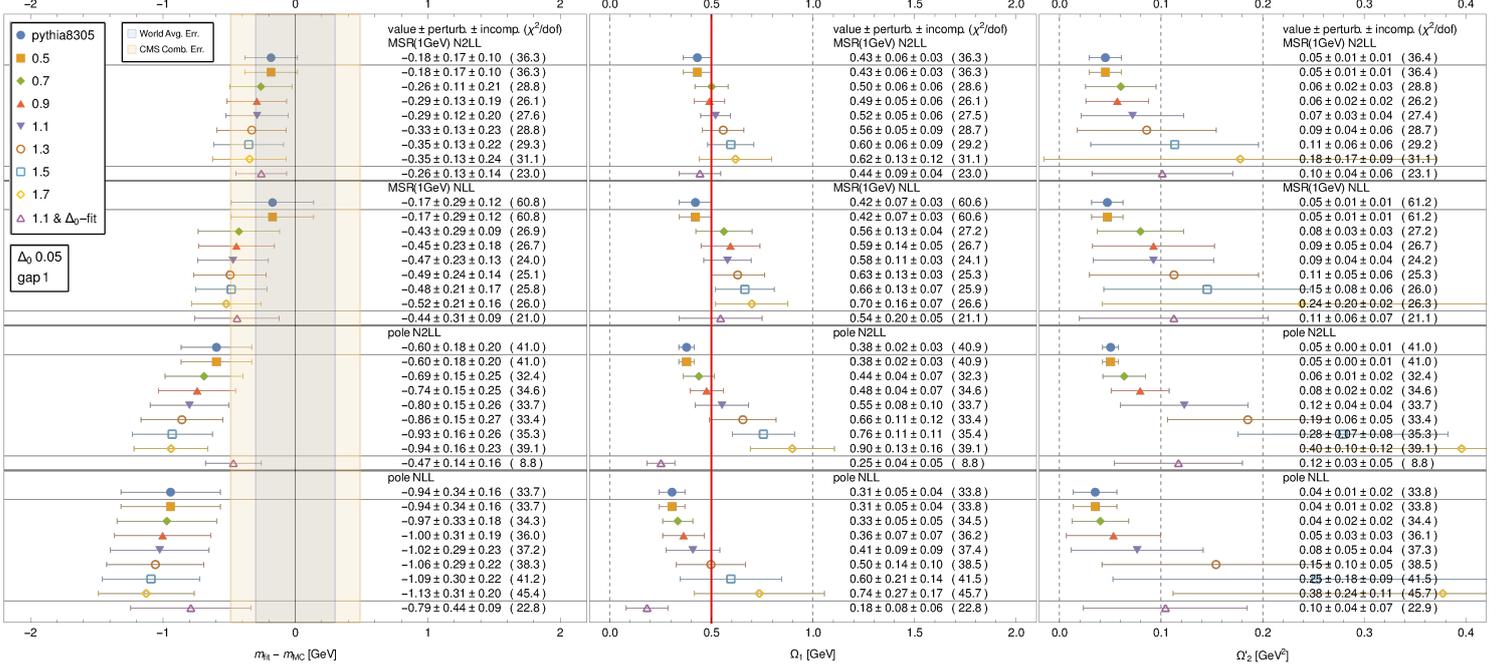


Figure 5.3.: Dependence of fitted parameters on  $\lambda$  values in the range 0.5 – 1.7 GeV for fixed  $\Delta_0 = 0.05$  GeV and gap 1. The fitted MSR masses are stable over a large range, if  $\lambda$  is not too small. The pole mass fits are apparently already too constrained by the fixed  $\Delta_0$  value (compare to the  $\Delta_0$ -fit data point), so that larger  $\lambda$  values push the  $\Omega_1$  and  $m_{\text{fit}}$  even farther away.

the previous  $\chi^2$  fit function introduced in Eq. (5.5) to:

$$\chi^2(m_t, \Delta_0, \{a\}) = \sum_Q \sum_{\tau_{\min} \leq \tau_i < \tau_{\max}} \frac{[f_{Q,i}^{\text{theo}}(m_t, \Delta_0, \{a\}) - f_{Q,i}^{\text{mc}}]^2}{\sigma_{Q,i}^2}. \quad (5.8)$$

We have two methods to deal with the additional  $\Delta_0$  fit parameter. The first method just uses the old  $\Delta_0$  independent profiles. The second method is computationally much faster and uses  $\Delta_0$  dependent profiles.

### Fits with $\Delta_0$ -independent profiles

In this version the only change to the theory cross section is a reference gap  $\Delta_0 = \overline{\Delta}(R_\Delta, \mu_\Delta)$  that can be fitted freely. To efficiently perform the fits we add an additional dimension to the grid, that we had before. This is handled the same way as the mass grid dimension. We used steps of the size  $\delta\Delta_0 = 0.05$  GeV. A range of

$\Delta_0 \in [-1.00 \text{ GeV}, 1.90 \text{ GeV}]$  should be generally safe for all gaps, mass schemes and orders, but due to the sheer size of the grid it might be more efficient to have smaller customized ranges. The one-dimensional spline interpolation that was applied to the mass dimension before is now replaced by a two-dimensional spline interpolation in the mass and  $\Delta_0$  direction. The equivalent of Eq. (5.6) now reads

$$f_{Q,i}^{\text{theo}}(m_t, \Delta_0, \{a\}) = \sum_{kl} c_k(\{a\}) c_l(\{a\}) f_{kli}(m_t, \Delta_0, Q) \quad (5.9)$$

where MINUIT is now able to smoothly sample in  $\{a\}$ ,  $m_t$ , and  $\Delta_0$ .

### Fits with $\Delta_0$ -dependent profiles

The second method relies on the fact that  $\Delta_0$  always appears as a  $Q$  dependent shift of the event shape  $e$ . Hence, we can use a grid with *one* fixed  $\Delta_{0,\text{grid}}$  value and generate any other  $\Delta_0$  cross section by sampling shifted points<sup>1</sup>:

$$\frac{d\sigma}{de}(e, \Delta'_0) = \frac{d\sigma}{de'}(e', \Delta_{0,\text{grid}}) \Big|_{e'=e-r_s \frac{\Delta'_0 - \Delta_{0,\text{grid}}}{Q}}. \quad (5.10)$$

But, this can only work if we also define the profile functions to behave the same way

$$\mu_i(e, \Delta'_0) \equiv \mu_i\left(e - r_s \frac{\Delta'_0 - \Delta_{0,\text{grid}}}{Q}, \Delta_{0,\text{grid}}\right). \quad (5.11)$$

The transition points in Eq. (4.4) of the profiles already depend on a fitted quantity, the profile mass  $m_t^{\text{MSR}}(R=5)$ , which is obtained from the fitted mass parameter. This allows the profiles to correctly adapt to changes of the peak position caused by changes of the mass. Since Monte Carlo event generators don't exactly behave like nature and their hadronization models can be very different compared to each other, it would also make sense to make the transition points depend on the fitted  $\Omega_1$ , so that the profiles also adapt to changes of peak position caused by different  $\Omega_1$ 's. The problem with  $\Omega_1$  is that its value as given in Eq. 5.7 is a function of soft shape function parameters, hence one cannot in contrast to the mass, attribute  $\Omega_1$  values to the grid, i.e. the profiles cannot depend on a fitted  $\Omega_1$  when using our method. The next best alternative is to use a  $\Delta_0$  dependence as described above. This is equivalent with  $\Omega_1$  dependence if we assume that the other shape function parameters  $c_i$  are constant. We define these new profiles to be identical to the old profiles when  $\Delta_0$  is at its old default value  $\Delta_{0,\text{gap } 1}^{\text{match}} = 0.05 \text{ GeV}$  for the gap 1 scheme. For other gap schemes the old profiles are matched at the one loop converted  $\Delta_0$  value, i.e. Eq. (5.2) with the replacements  $\{\Omega_{1,\text{redline,gap } i} \rightarrow \Delta_{0,\text{gap } i}^{\text{match}}, 0.5 \rightarrow 0.05\}$ .

In practice we first save a grid with the same specifications as for the fixed  $\Delta_0$  fits with  $\Delta_0 = \Delta_{0,\text{grid}}$ . The cumulant over the event shape variable  $\tau$  is evaluated at the bin edges.

<sup>1</sup>Here, we have already displayed a power correction  $r_s = 1 + \#\hat{m}^2$ , that will be used when we consider rescaling corrections starting with Chap. 6.

Then we do a 2-D spline interpolation of the cumulant over  $m_t$  and  $\tau$ , which we define as  $F_{kl}(\tau, m_t, Q)$ . The bin values can still be defined as in Eq. (5.9) with the shape function pieces now obtained by evaluating

$$f_{kli}(m_t, \Delta_0, Q) = F_{kl}(\tau_{i+1} - r_s \frac{\Delta_0 - \Delta_{0,\text{grid}}}{Q}, m_t, Q) - F_{kl}(\tau_i - r_s \frac{\Delta_0 - \Delta_{0,\text{grid}}}{Q}, m_t, Q) \quad (5.12)$$

which is much faster than before, since we can use an out of the box interpolator, the PYTHON class `scipy.interpolate.RectBivariateSpline`, that vectorizes the evaluation of  $F_{kl}$  along the  $\tau$  direction. The creation of the grids is obviously also faster and produces smaller files due to the removal of the  $\Delta_0$  grid dimension and this additionally reduces the time required to distribute the grids to each node of the cluster during fits.

### 5.2.2. Choosing suitable $\lambda$ for each Event-Generator

We still have to determine which  $\lambda$  should be used for the shape function. Fig. 5.4 shows the  $\lambda$  dependence of fitted parameters if we use the old  $\Delta_0$ -independent profiles and PYTHIA. It tells us, that as long as we chose  $\lambda$  large enough and at least within the range plotted, we will obtain stable mass and  $\Omega_1$  results. The  $\Omega'_2$  value is less under control, and it has little correlation with the other parameters, as long as, again,  $\lambda$  is large enough. We see that  $\lambda = 0.5$  is restricting the  $\Omega'_2$  too much, so that the fitted distribution has a too narrow width, which is clearly reflected in the much higher  $\chi^2$  value.

This fitting method of using  $\Delta_0$ -independent profiles is very slow. In the following we will show that using the  $\Delta_0$ -dependent profiles produces results that are compatible with the results of the old method. The  $\lambda$  dependence for this new method and PYTHIA is shown in Fig. 5.5. With this method the values have shifted slightly and the uncertainty in some cases has increased. This might be explained by the circumstance that the profiles already contained a variation for the transition points  $t_i$ . Now that the profiles also change with the fitted value of  $\Delta_0$  we have introduced an additional variation. We also noticed that the  $\chi^2$  has decreased slightly overall. This is probably caused by the fit trying to optimize the profile transition point through the value of  $\Delta_0$ . The fit results of the new and the old method shown in Fig. 5.4 and Fig. 5.5 respectively are fully compatible within errors, so we will from now on exclusively use  $\Delta_0$ -dependent profiles, since they improve the computation time by an order of magnitude. The new method will also lead to profiles that automatically adapt to event generators with larger  $\Omega_1$  values.

The fitted parameters for PYTHIA in Fig. 5.5 stabilize around  $\lambda = 1.1$  GeV. We take this value as the default for PYTHIA since all other results for  $\lambda \in [0.7 \text{ GeV}, 1.5 \text{ GeV}]$  are compatible with this choice. As before,  $\lambda = 0.5$  GeV is clearly too limiting for the width of the shape function.

We repeated the same test for HERWIG and SHERPA, the results are plotted in Fig. 5.6 and Fig. 5.7 respectively. The behaviour of HERWIG with respect to  $\lambda$  variation is similar to PYTHIA. Here we also chose  $\lambda = 1.1$  GeV as default. In contrast, SHERPA's distributions turn out to be wider when using equivalent generator settings. As a consequence, the compatible  $\lambda$  values are larger and we have to chose  $\lambda = 1.3$  GeV as default for SHERPA.

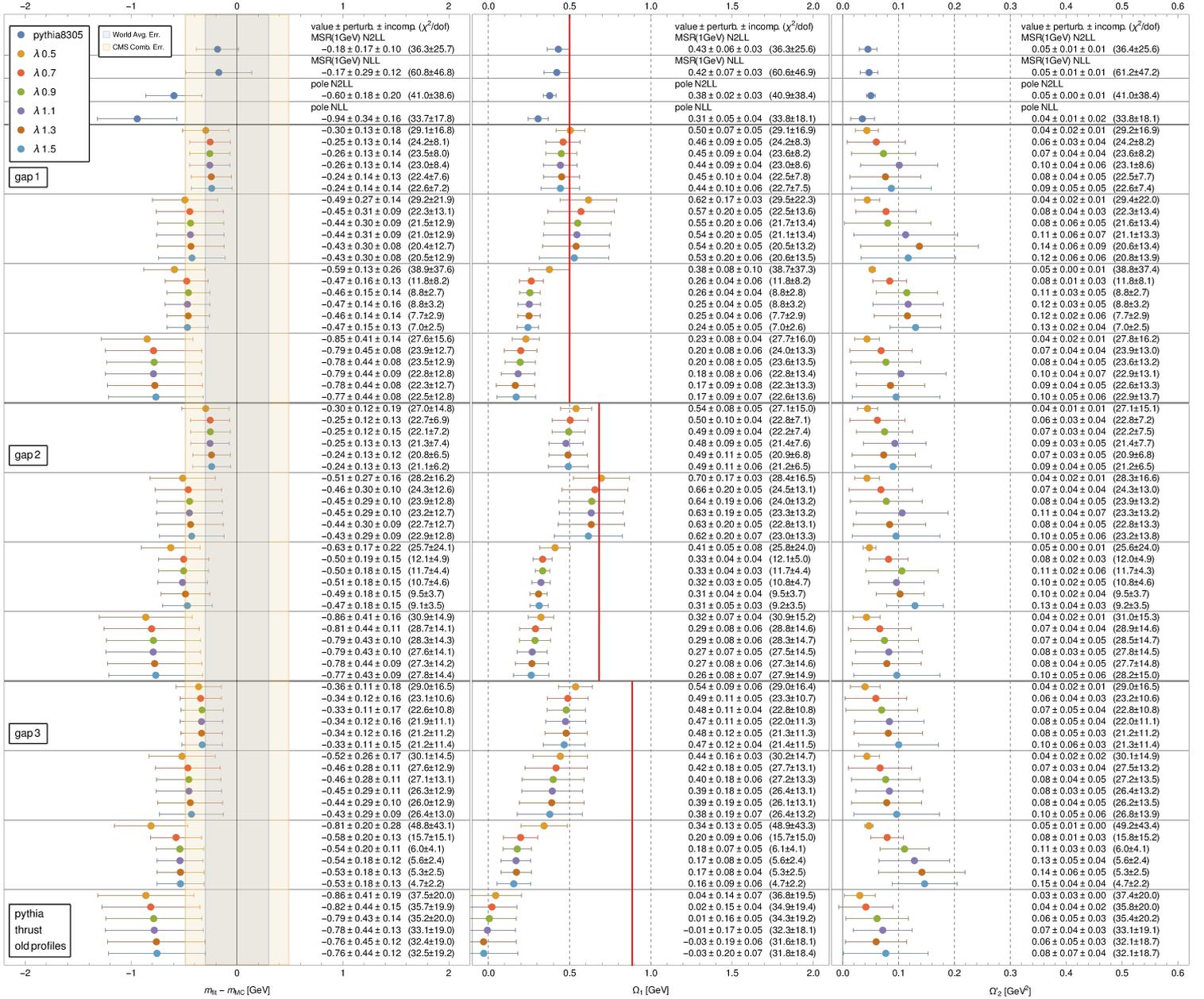


Figure 5.4.: Dependence of fitted parameters on  $\lambda$  values in the range  $0.5 - 1.5$  GeV using old  $\Delta_0$ -independent profiles and PYTHIA.  $m_{\text{fit}}$  and  $\Omega_1$  are very stable starting from  $\lambda = 0.7$  GeV. (Thin horizontal lines separate different combinations of mass scheme and perturbative order, the ordering is the same as for the “pythia8305” data points.)

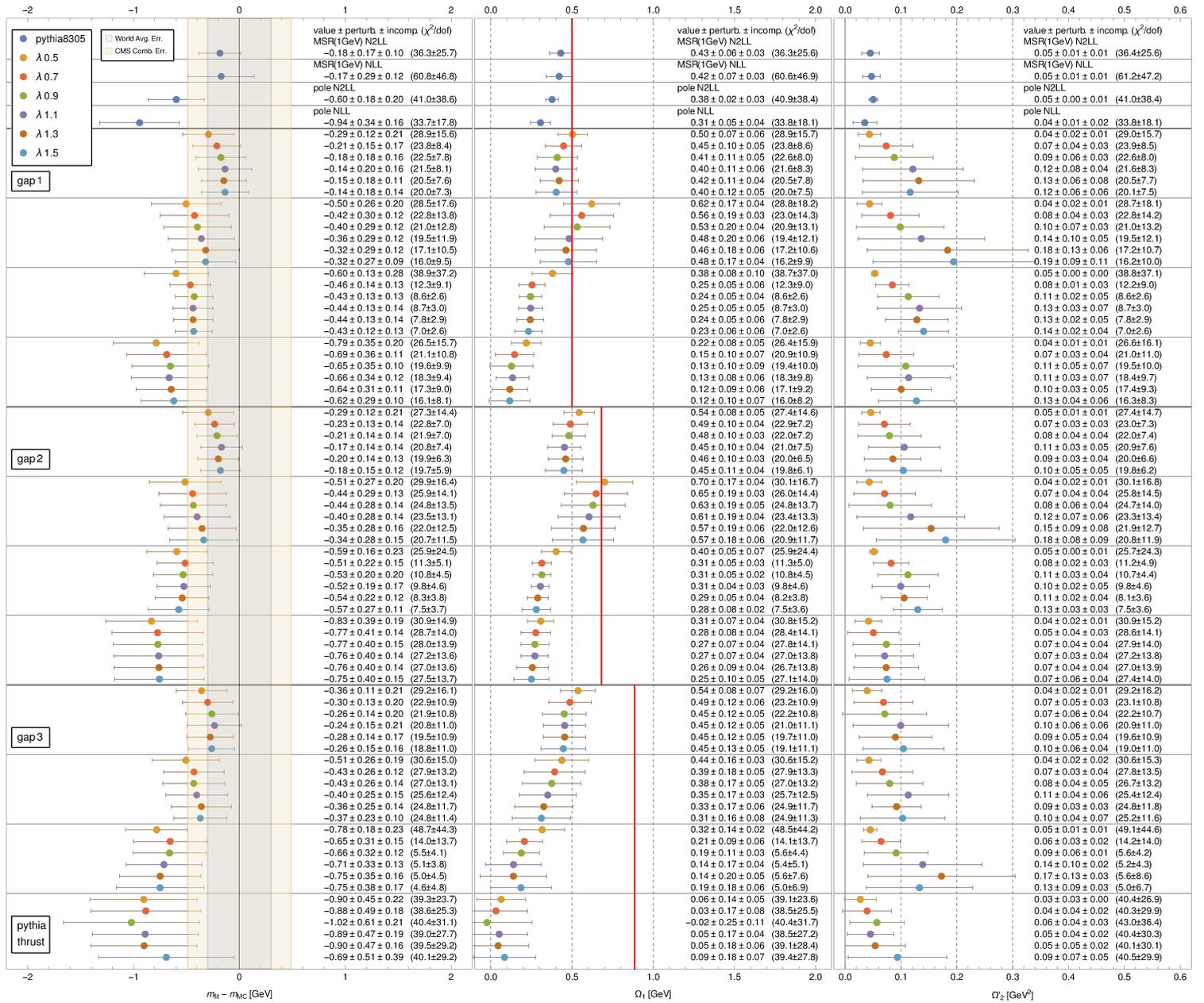


Figure 5.5.: Dependence of fitted parameters on  $\lambda$  values in the range 0.5 – 1.5 GeV using  $\Delta_0$ -dependent profiles for PYTHIA. We chose  $\lambda = 1.1$  GeV as default for PYTHIA.

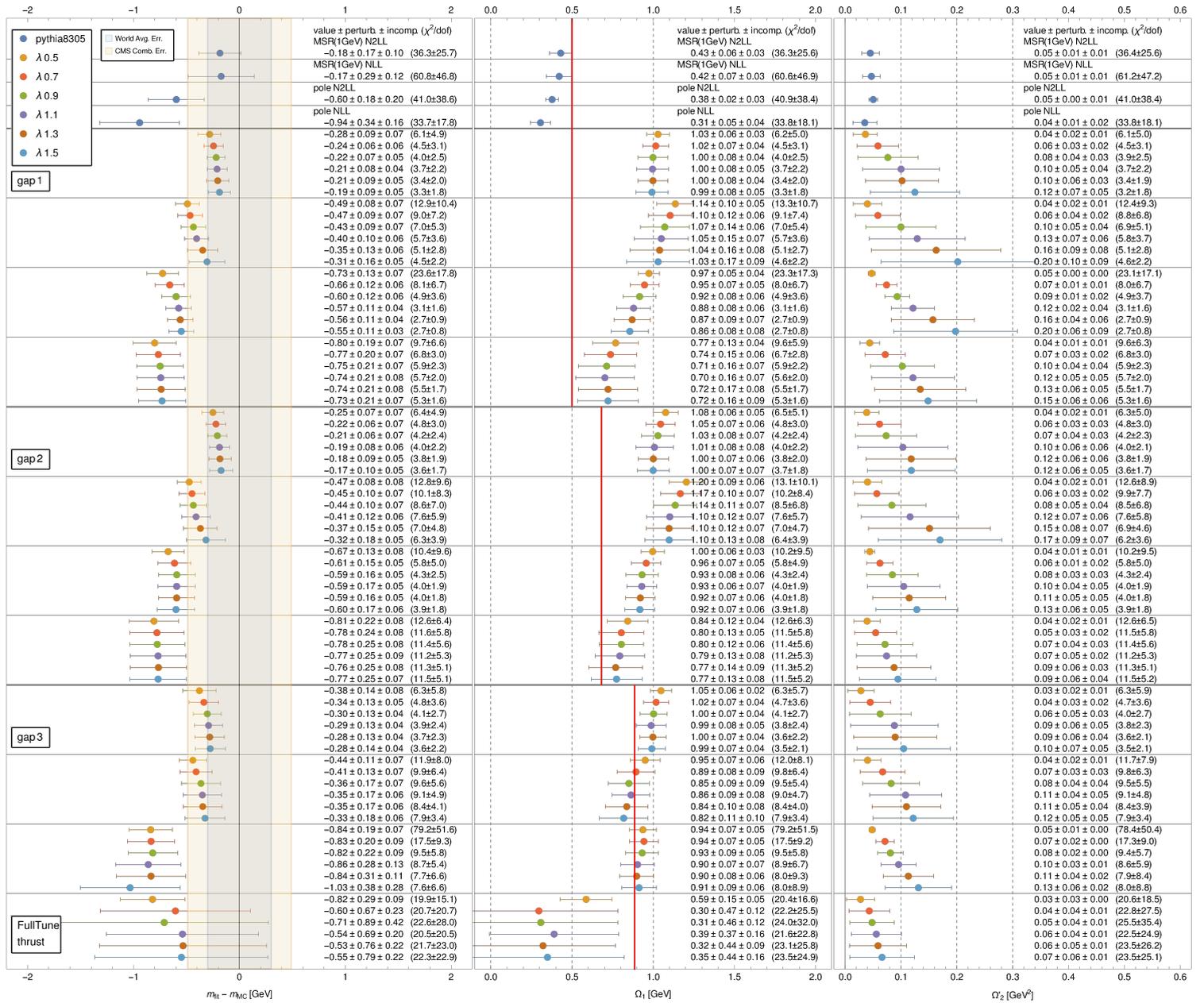


Figure 5.6.: Dependence of fitted parameters on  $\lambda$  values in the range 0.5 – 1.5 GeV using  $\Delta_0$ -dependent profiles for HERWIG. We chose  $\lambda = 1.1$  GeV as default for HERWIG.

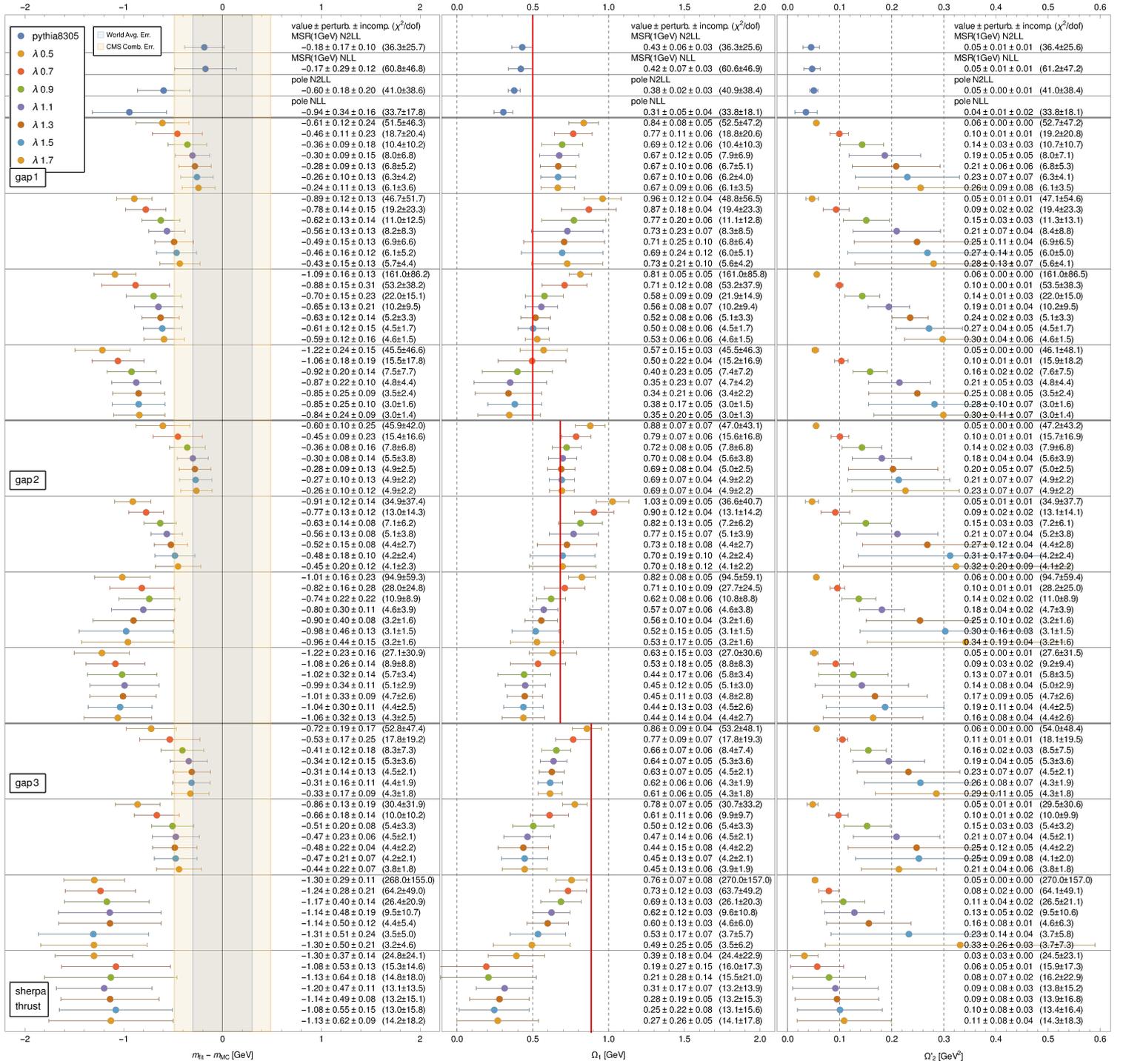


Figure 5.7.: Dependence of fitted parameters on  $\lambda$  values in the range 0.5 – 1.7 GeV using  $\Delta_0$ -dependent profiles for SHERPA. We chose  $\lambda = 1.3$  GeV as default for SHERPA.

### 5.2.3. Results of improved Fit Procedure

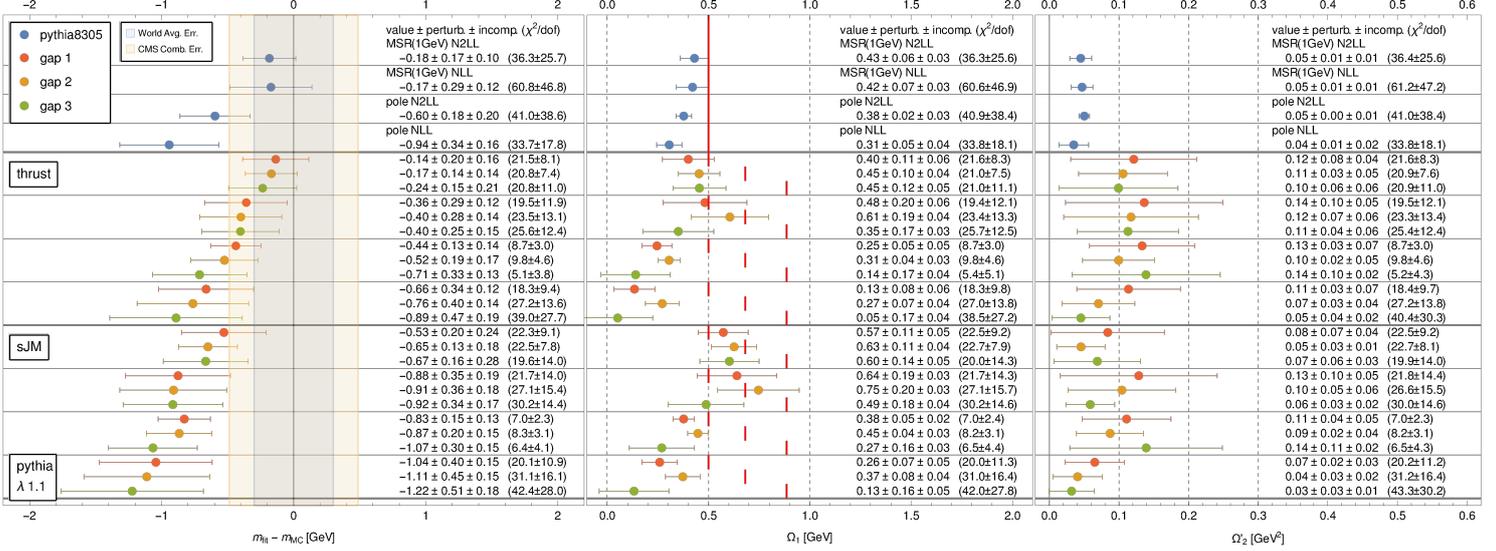


Figure 5.8.: Dependence of fitted parameters on different gap schemes using variable  $\Delta_0$  fit method for 2-Jettiness (referred to as thrust) and sJM (sum of jet mass). The fitted values for  $\Omega_1$  now agree with each other at N<sup>2</sup>LL for different gap schemes. The mass can now also be correctly extracted without interference from the shape function. The combinations of pole mass and gap 3 result in large scale uncertainties due to the extreme running of gap 3. Note that 2-Jettiness and sJM fits exhibit large differences in extracted mass and  $\Omega_1$

After determining how to best perform the variable  $\Delta_0$  fits in the last subsection, we are now able to test if our fitting procedure has improved compared to the old method. For that, we repeat the gap scheme dependence fit using the new fit method. The results for 2-Jettiness (denoted as thrust in these plots) are plotted in the upper half of Fig. 5.8. It is very clear that, in contrast with the old method shown in Fig. 5.1, the new method provides very stable extracted masses, regardless of the gap scheme used. The reason is that the fits automatically adjust  $\Delta_0$  so that the difference in the gap subtraction is compensated for. This is evident in the agreement of the fitted N<sup>2</sup>LL values for the  $\Omega_1$  parameter. Remember that all plots show  $\Omega_1$  corrected to the one-loop gap 1 scheme. The combination of pole mass and gap 3 at N<sup>2</sup>LL exhibits large uncertainties due to large running effects specific to gap 3. This problem does not appear for the MSR mass, since MSR mass and gap 3 have opposite signed subtractions, thus the overall subtraction does not cause perturbative problems. The validity of this new approach is also reflected in the dramatic improvement of the  $\chi^2/\text{d.o.f}$  values for measurements that rely on the two new gap schemes, while the fits using the old gap scheme have also achieved significantly

lower  $\chi^2/\text{d.o.f}$  values.

#### 5.2.4. First Test of Universality

Given that our shape function parametrization now works reliably, we can turn our attention to study the universality of our “bHQET + nonsingulars” approach with respect to different observables. The fit results for the sJM (sum of jet mass) observable is plotted in the lower half of Fig. 5.8. The extracted mass values are about 400 MeV smaller than the corresponding masses obtained from the 2-Jettiness distributions. This difference is a systematic effect and turns out to be related to the difference in the power corrections for these two observables and its size is between one to two times the uncertainty that we obtained from perturbative scale and fit range variations.

We already know, that 2-Jettiness has large power corrections in its threshold  $\tau_{\min} = 1 - \sqrt{1 - 4\hat{m}^2} = 2\hat{m}^2 + \mathcal{O}(\hat{m}^4)$ , whereas sJM does not. So one could speculate that sJM is better suited for the SCET power counting and produces more reliable fit results. But this would make the obtained result even more perplexing, because the extracted sJM mass values are actually much lower than the theoretical expectation of the Monte Carlo generator mass being close to the MSR mass at  $R \sim 1$  GeV [9, 10]. Also recall that we are already using a prescription of including the exact threshold in the bHQET cross section, so the disagreement observed should not originate from the power corrections in the threshold.

Apart from the disagreement in the mass values, we also observe a systematic increase of the fitted  $\Omega_1$  values of about 150 MeV. We will see in the next chapter that relative power corrections between the threshold  $\tau_{\min}$  and the soft function convolution variable  $\ell$  explain these shifts.

# 6. Kinematic Mass Power Corrections

## 6.1. Measurement Function

In this chapter we will identify the  $\hat{m}^2 \sim 1/10$  power corrections in the bHQET measurement function. We will focus specifically on these corrections since all the other possible power corrections that we have mentioned in the intro of Chap. 3 are in the peak region parametrically significantly smaller. The original factorization theorem in [12] can be read as a differential distribution in the variables  $k_n^+$ ,  $k_{\bar{n}}^-$  and  $k_s^+$ ,  $k_{\bar{s}}^-$ , where  $\ell = k_s^+ + k_{\bar{s}}^-$  is the argument of the thrust soft function  $S_\tau(\ell)$  and  $\hat{s} = \hat{s}_1 + \hat{s}_2 = k_n^+ + k_{\bar{n}}^-$  (see Eq. (6.8)) is the argument of the thrust jet function  $B_\tau(\hat{s})$ . This distribution generates the event shape distribution by the measurement function<sup>1</sup>

$$Q[e - e_{\min}(\hat{m})] = r_c(\hat{m})\hat{m}(k_n^+ + k_{\bar{n}}^-) + r_s(\hat{m})(k_s^+ + k_{\bar{s}}^-). \quad (6.1)$$

The leading expressions for the power counting in  $\hat{m}$  are  $e_{\min} = 2\hat{m}^2$ ,  $r_c = 1$  and  $r_s = 1$ . The  $\hat{m}^4$  correction in the threshold  $e_{\min}$  is equivalent to a  $\mathcal{O}(m_t\hat{m}^2) \approx$  several GeV shift of  $m_t$  in the resummed cross section and attempting to formally compensate this correction in the nonsingular distribution as defined in the first line of Eq. 3.5 would essentially lead to a large positive (the original QCD) and a large negative peak in the matching correction. This would undermine the purpose of using resummation. The original calibration of PYTHIA using the 2-Jettiness event shape therefore included the exact perturbative threshold at  $e_{\min}(\hat{m})$  in the singular factorization-theorem.

The collinear rescaling factor  $r_c$  mainly modifies the width of the jet function (see Eq. 6.20), since it is peaked close to the origin.  $r_c$  should therefore only have a small impact on  $m_t$ .

The soft rescaling factor  $r_s$  in principle represents corrections of the same power as  $r_c$ , but the combination with the first moment of the soft shape function  $\Omega_1$ , which contains the hadronic threshold, effectively causes an  $\mathcal{O}(\hat{m}\Omega_1)$  shift of the fitted calibration result for the top mass due to a shift of the peak position<sup>2</sup>. Matching to perturbative QCD also cannot account for this shift since it mainly originates from a non-perturbative correction. This problem can only be fixed by applying the measurement correction consistently to singular and nonsingular parts.

<sup>1</sup>The variables  $k_n^+$  and  $k_{\bar{n}}^-$  are the components of the residual collinear momenta in the rest frames of the top and anti-top respectively. The factor that boosts them into the lab frame is contained in the prefactor  $r_c\hat{m}$ .

<sup>2</sup>An  $\mathcal{O}(\hat{m}^2\Lambda)$  correction in the measurement function can be related to an  $\mathcal{O}(\hat{m}\Lambda)$  shift in  $m_t$ , since a mass shift in the threshold has the form  $(m + \delta m)^2/Q^2 \approx \hat{m}^2 + \hat{m}(2\delta m)/Q$ , while the rescaling corrections are power counted as  $\hat{m}(\hat{m}\Lambda)/Q$ .

In the following sections we will show the definitions of the collinear and soft variables and their power counting, calculate the rescale factors for relevant observables and show how the factorization theorem and the nonsingular corrections have to be modified.

## 6.2. Rescale Factors for Event Shapes

We set the light cone four vectors parallel to the thrust axis

$$n^\mu = (1, \vec{n}), \quad \bar{n}^\mu = (1, -\vec{n}) \quad (6.2)$$

with the unit vector  $\vec{n}$  pointing into the top quark hemisphere. Four vectors can then be decomposed as

$$p^\mu = p^+ \frac{\bar{n}^\mu}{2} + p^- \frac{n^\mu}{2} + p_\perp^\mu \quad (6.3)$$

which defines the light cone coordinates  $(p^+, p^-, p_\perp) = (n \cdot p, \bar{n} \cdot p, p_\perp)$ . The inner product then reads  $p^\mu q_\mu = (p^+ q^- + p^- q^+)/2 + p_\perp^\mu q_{\perp\mu}$ . The  $n$  and  $\bar{n}$  hemispheres are separated by the plane perpendicular to  $\vec{n}$ . The total momentum going into the  $n$  hemisphere can be split into the momentum of an on-shell top quark traveling parallel to the thrust axis, the sum of ultra-collinear momenta  $k_n = \sum k_{n,i}$ , and the sum of soft momenta  $k_s = \sum k_{s,i}$ . The  $\bar{n}$  momenta are split analogously

$$\begin{aligned} p_n &= m v_+ + k_{n+} + k_s \\ p_{\bar{n}} &= m v_- + k_{\bar{n}-} + k_{\bar{s}}. \end{aligned} \quad (6.4)$$

where this is a component-wise equation in the lab frame (which we will define to be the center of mass frame). The  $+$ ( $-$ ) subscript indicates a boost of the components from the on-shell top (anti-top) “rest” frame to the lab frame.

The four-velocity of the on-shell top and anti-top in their respective zero three momentum frames is  $v = (1, 1, 0_\perp)$ . The transformation to the lab frame reads  $v_+ = (\lambda, \lambda^{-1}, 0_\perp)$  and  $v_- = (\lambda^{-1}, \lambda, 0_\perp)$  with the boost factor given by

$$\lambda \equiv \frac{1 - \beta}{2\hat{m}} \approx \hat{m} \quad (6.5)$$

where we will in this section use the letter  $\beta = \sqrt{(Q/2)^2 - m^2}/(Q/2) = \sqrt{1 - 4\hat{m}^2}$  as the velocity of the on-shell top quark in the lab frame.

The soft momenta  $k_{s,i}$  belong to large angle gluon radiation in the  $n$  hemisphere and the sum of these parton momenta is defined as  $k_s = (k_s^+, k_s^-, k_s^\perp) \simeq (\Lambda, \Lambda, \Lambda)$ . The soft momentum  $k_{\bar{s}}$  in the  $\bar{n}$  hemisphere has the same scaling. These soft momenta are in the peak region parametrically smaller than the top width  $\Lambda \gtrsim \hat{m}\Gamma$ .

The ultra-collinear momenta correspond to quark and gluon fluctuations that are small in the top rest frame  $k_{n,i} \simeq \Lambda/\hat{m} \gtrsim \Gamma$ . The sum of these ultra-collinear parton momenta is defined in the top rest frame as  $k_n = (k_n^+, k_n^-, k_n^\perp) \simeq (\Lambda/\hat{m}, \Lambda/\hat{m}, \Lambda/\hat{m})$ . In the lab frame the components are given by  $k_{n+} = (\lambda k_n^+, \lambda^{-1} k_n^-, k_n^\perp) \simeq (\Lambda, \Lambda/\hat{m}^2, \Lambda/\hat{m})$ . The momenta  $k_{\bar{n}}$  and  $k_{\bar{n}-}$  are defined analogously for the  $\bar{n}$  hemisphere.

The definitions above will be used to calculate the measurement function for the observables that we consider. The collinear momenta in the final results will be expressed in terms of  $k_n^+$  and  $k_{\bar{n}}^-$ , i.e. components defined in the top and antitop rest frame respectively. The thrust jet function  $B_\tau(\hat{s})$ , however, is defined with an off-shellness argument  $\hat{s} = \hat{s}_1 + \hat{s}_2$ , where  $\hat{s}_1$  and  $\hat{s}_2$  are the off-shellnesses  $\hat{s}_{1/2} = (p_{1/2}^2 - m^2)/m$  of the collinear top and antitop momenta respectively,  $p_1 = p_n - k_s$  and  $p_2 = p_{\bar{n}} - k_{\bar{s}}$ . To relate these two expressions we have to first expand the off-shellness in the SCET power counting and then insert our bHQET expressions. The SCET momentum decomposition is given by [12]

$$p_1 = (p_1^+, p_1^-, p_1^\perp) = (p_1^+, \tilde{p}_1 + k_{n,\text{SCET}}^-, p_1^\perp) \quad (6.6)$$

where the label momentum  $\tilde{p}_1 \sim Q$  is large and the other momenta are residual and are power counted to have the same size. By first applying this power counting and then identifying the label and residual momenta with bHQET momenta we obtain the invariant mass

$$p_1^2 \approx \tilde{p}_1 p_1^+ = (mv_+^-)(mv_+^+ + k_{n,+}^+) = m^2 + mk_n^+. \quad (6.7)$$

The expression for the  $\bar{n}$  hemisphere is calculated analogously and we finally obtain the expressions for the off-shellness variables

$$\begin{aligned} \hat{s}_1 &= (p_1^2 - m^2)/m \approx k_n^+ \\ \hat{s}_2 &= (p_2^2 - m^2)/m \approx k_{\bar{n}}^- \end{aligned} \quad (6.8)$$

## 2-Jettiness

The definition of 2-Jettiness is

$$Q\tau_J = \min_{\hat{t}} \sum_i (E_i - |\hat{t} \cdot \vec{p}_i|). \quad (6.9)$$

We replace the thrust axis  $\hat{t}$  with the light-cone vector  $\vec{n}$  that was chosen to lie along the thrust axis, so we can assign the momenta to their hemispheres.

$$\begin{aligned} Q\tau_J &= \sum_{\vec{n} \cdot \vec{p}_a > 0} (E_a - \vec{n} \cdot \vec{p}_a) + \sum_{\vec{n} \cdot \vec{p}_b < 0} (E_b + \vec{n} \cdot \vec{p}_b) \\ &= p_n^+ + p_{\bar{n}}^- \end{aligned} \quad (6.10)$$

This turns out to be exactly the total plus(minus) momenta in the  $n(\bar{n})$  hemisphere. Now we just plug in Eq. (6.4)

$$Q\tau_J = Q(1 - \beta) + \frac{1 - \beta}{2\hat{m}^2} \hat{m} (k_n^+ + k_{\bar{n}}^-) + (k_s^+ + k_{\bar{s}}^-). \quad (6.11)$$

The extracted corrections read

$$\begin{aligned} e_{\min} &= 1 - \beta & r_c &= \frac{1 - \beta}{2\hat{m}^2} & r_s &= 1 \\ &= 2\hat{m}^2 + 2\hat{m}^4 + 4\hat{m}^6 + \mathcal{O}(\hat{m}^8) & &= 1 + \hat{m}^2 + \mathcal{O}(\hat{m}^4) \end{aligned} \quad (6.12)$$

Note that 2-Jettiness has  $r_s = 1$ , i.e. there is no power correction for the soft rescaling factor.

## SJM

sJM is the sum of hemisphere invariant masses

$$Q\tau_s = \frac{1}{Q}(p_n^2 + p_{\bar{n}}^2) \approx \frac{1}{Q}(p_n^+ p_n^- + p_{\bar{n}}^- p_{\bar{n}}^+) \quad (6.13)$$

We can replace the  $p_n^- (p_{\bar{n}}^+)$  components with  $p_n^+ (p_{\bar{n}}^-)$  using the relations

$$\begin{aligned} p_n^- + p_n^+ &= Q + 2\Delta E \\ p_{\bar{n}}^- + p_{\bar{n}}^+ &= Q - 2\Delta E. \end{aligned} \quad (6.14)$$

$\Delta E(-\Delta E)$  is the sum of the residual energy in the  $n(\bar{n})$  hemisphere and inherits a power counting of  $\Delta E = [(k_{n+}^- + k_s^-) + (k_{n+}^+ + k_s^+)]/2 \simeq \Lambda/\hat{m}^2$ . Inserting everything into the definition

$$Q\tau_s = 2Q\hat{m}^2 + \beta \frac{1-\beta}{2\hat{m}^2} \hat{m} (k_n^+ + k_{\bar{n}}^-) + \beta (k_s^+ + k_{\bar{s}}^-) + \mathcal{O}\left(\frac{\Lambda^2}{Q\hat{m}^2}\right) \quad (6.15)$$

gives us the corrections for sJM

$$\begin{aligned} e_{\min} = 2\hat{m}^2 \quad r_c = \beta \frac{1-\beta}{2\hat{m}^2} \quad r_s = \beta \\ = 1 - \hat{m}^2 + \mathcal{O}(\hat{m}^4) \quad = 1 - 2\hat{m}^2 + \mathcal{O}(\hat{m}^4) \end{aligned} \quad (6.16)$$

## How to include these corrections

We see that 2-Jettiness and sJM have a collinear rescaling, which is opposite in sign. 2-Jettiness has no power corrections for the soft rescaling factor while the sJM rescaling factor contains power corrections. This suggests that sJM will suffer from threshold shifting effects due to kinematic power correction interacting with nonperturbative corrections.

There are now two possible ways to test the impact of these corrections:

1. Modify the measurement function in the original factorization theorem to incorporate the power corrections. This is presented in the next section.
2. Define a new observable as a function of sJM that is cured from the leading soft rescaling. This new observable can be tested with the standard measurement function or a corrected one. This will allow us to test if the power corrections to  $r_s$  are the main issue to address.

## MJM

This new observable which we will call *modified jet mass* (mJM) can be defined as a polynomial in  $\tau_s$ . A quadratic function is sufficient to remove the leading soft rescaling

$$\tau_m = \tau_s + \frac{1}{2}\tau_s^2. \quad (6.17)$$

Inserting the measurement function for  $\tau_s$  into this definition results in

$$Q\tau_m = 2Q(\hat{m}^2 + \hat{m}^4) + (1 + 2\hat{m}^2)\beta \frac{1-\beta}{2\hat{m}^2} \hat{m} (k_n^+ + k_n^-) + (1 + 2\hat{m}^2)\beta (k_s^+ + k_s^-) + \mathcal{O}\left(\frac{\Lambda^2}{Q\hat{m}^2}\right) \quad (6.18)$$

and gives us the corrections for mJM

$$\begin{aligned} e_{\min} &= 2\hat{m}^2 + 2\hat{m}^4 & r_c &= (1 + 2\hat{m}^2)\beta \frac{1-\beta}{2\hat{m}^2} & r_s &= (1 + 2\hat{m}^2)\beta & (6.19) \\ & & &= 1 + \hat{m}^2 + \mathcal{O}(\hat{m}^4) & &= 1 - 6\hat{m}^4 + \mathcal{O}(\hat{m}^6). \end{aligned}$$

This observable has the same leading  $\hat{m}^2$  corrections as 2-Jettiness. That means that mJM and 2-Jettiness distributions will be very similar, but the experimental data will still be based on the sJM data, which contains large corrections. This will accordingly serve as check for how well the leading kinematic power corrections can describe the differences between these two observables.

### 6.3. Factorization Theorem Modifications

We will schematically show how the bHQET factorization theorem is modified due to the corrected measurement function. The functions here are already RGE evolved to a common scale  $\mu$  and the hard matching factors are combined into one function  $H$ . We use the more commonly used symbols  $\hat{s}$  and  $\ell$  instead of  $k_n^+$  and  $k_s^+$ .

$$\begin{aligned} & \frac{1}{\sigma_0} \left. \frac{d\sigma_{\text{bHQET}}}{de} \right|_{\text{resc.}} \\ &= Q^2 H(Q, m, \mu) \int d\ell d\hat{s}_1 d\hat{s}_2 [mB_n(\hat{s}_1, \Gamma, \mu)][mB_{\bar{n}}(\hat{s}_2, \Gamma, \mu)] S_\tau(\ell, \mu) \\ & \quad \times \delta[Q^2(e - e_{\min}) - r_c m(\hat{s}_1 + \hat{s}_2) - r_s Q\ell] \\ &= \frac{Q^2}{r_c} H(Q, m, \mu) \int d\ell d\hat{s} B_n \left[ \frac{Q^2(e - e_{\min}) - r_s Q\ell}{r_c m} - \hat{s}, \Gamma, \mu \right] [mB_{\bar{n}}(\hat{s}, \Gamma, \mu)] S_\tau(\ell, \mu) \\ &= \frac{mQ^2}{r_c} H(Q, m, \mu) \int d\ell B_\tau \left[ \frac{Q^2(e - e_{\min}) - r_s Q\ell}{r_c m}, \Gamma, \mu \right] S_\tau(\ell, \mu) \\ &= \frac{mQ^2}{r_c} H(Q, m, \mu) \int d\ell d\hat{s} dk B_\tau \left[ \frac{Q^2(e - e_{\min}) - r_s Q\ell}{r_c m} - \hat{s}, \mu \right] G(\hat{s}, \Gamma) \\ & \quad \times \hat{S}_\tau(\ell - k, \bar{\delta}, \mu) F(k - 2\bar{\Delta}) \\ &= \frac{mQ^2}{r_c} H(Q, m, \mu) \int d\ell d\hat{s} dk B_\tau \left[ \frac{Q^2(e - e_{\min}) - r_s Q(\ell + k)}{r_c m} - \hat{s}, \mu \right] \hat{S}_\tau(\ell, \bar{\delta}, \mu) \\ & \quad \times G(\hat{s}, \Gamma) F(k - 2\bar{\Delta}) \\ &= \frac{1}{\sigma_0} \int d\hat{s} dk \frac{d\hat{\sigma}_s}{de} \left( e - r_c \frac{m\hat{s}}{Q^2} - r_s \frac{k}{Q}, e_{\min}, r_c, r_s, \Gamma = 0 \right) G(\hat{s}, \Gamma) F(k - 2\bar{\Delta}) \end{aligned} \quad (6.20)$$

In the second line we replace  $s_1$  by using the delta function. The third line uses the definition of the thrust jet function as a convolution of the hemisphere jet functions  $B_\tau = B_n \otimes B_{\bar{n}}$ . In the fourth line we write the jet function as a convolution with a Breit Wigner function  $G(\hat{s}, \Gamma)$  and the soft function as a convolution of perturbative soft function and nonperturbative model function  $F(k - 2\Delta)$ . In the second to last line we just redefine the soft convolution variable so that the convolution variable of the nonperturbative model function appears together with the observable. This form shows how the convolutions over the Breit Wigner function the soft perturbative function and the nonperturbative model functions have to be modified. In the last line we combine all non-model-dependent parts into a ‘‘perturbative’’ cross section, which should be seen as the counterpart to the perturbative QCD cross section with stable top quark. So the remaining convolutions over the model functions should be implemented the same way for the nonsingular cross section as shown in Eq. 6.21.

Note, that the  $\rho$  in  $U_v$  in A.1 has to be replaced now by by the exact boost factor  $\lambda^{-1} = r_s/(r_c \hat{m})$  to make the running  $\mu$  independent. The pole mass renormalon cancellation is implemented as before by the replacement rule  $e_{\min} \rightarrow e_{\min} - \frac{de_{\min}}{dm} \delta m$  as explained below Eq. (4.1).

## 6.4. SCET Nonsingular Modifications

Recall that the exact threshold has already been included in our singular distribution by the prescription  $e - 2\hat{m}^2 \rightarrow e - e_{\min}$ . We have also seen that the nonsingular contribution from SCET to the full cross section in the peak region has been numerically much smaller than the bHQET singular and the QCD nonsingular contributions. We will therefore adopt a minimal modification prescription to keep the SCET nonsingular small by using the same rescaled measurement function as in the bHQET case. This will exactly match the bHQET jet function onto the distributional parts of the SCET jet function and what remains is the nondistributional part. That is, we replace  $B_\tau[\hat{s}, \mu]$  in the rescaled cross section in the second to last line of Eq. (6.20) with  $J_{\text{nsb}}(s = m\hat{s}, m, \mu)$  which is defined in Eq. (3.4) to obtain the rescaled SCET nonsingular (To be concrete,  $\hat{s}$  is exactly the first argument of  $B_\tau[\hat{s}, \mu]$  in the line mentioned, which contains all the the rescale factors. Also note that we have not smeared nondistributional parts with the Breit Wigner function in the numerical code).

## 6.5. QCD Nonsingular Modifications

At the end of Sec. 6.3 we have already mentioned that the convolution variables over the model-dependent functions  $G(\hat{s}, \Gamma)$  and  $F(\ell)$  should be modified the same way as for the bHQET cross section. This is done for consistency, but it might only have a negligible numerical effect, since it is a subleading power correction.

$$\frac{1}{\sigma_0} \frac{d\sigma_{\text{ns}}}{de} = \frac{1}{\sigma_0} \int d\hat{s} dk \frac{d\hat{\sigma}_{\text{ns}}}{de} \left( e - r_c \frac{m\hat{s}}{Q^2} - r_s \frac{k}{Q}, \Gamma = 0 \right) G(\hat{s}, \Gamma) F(k - 2\bar{\Delta}) \quad (6.21)$$

The perturbative nonsingular cross section is still defined as in Eq. (3.5), but we now have to restate the fixed order SCET cross section result with explicit factors of  $r_s$  and  $r_c$  that have been set to 1 in Eq. (3.7):

$$\begin{aligned}
\frac{1}{\sigma_0} \frac{d\sigma_{\text{SCET}}}{de} \Big|_{\text{resc.,FO}} &= \delta(e - e_{\min}) \\
&+ \frac{C_F \alpha_s(\mu)}{4\pi} \left\{ [2\pi^2 + 4 \log(\hat{m}) + 16 \log^2(\hat{m})] \delta(e - e_{\min}) - 8[1 + 2 \log(\hat{m})] \left[ \frac{1}{e - e_{\min}} \right]_+ \right\} \\
&+ \frac{C_F \alpha_s(\mu)}{4\pi} \{ 8[(1 + L_c + 2L_m)L_c - L_s^2 + 2(L_c - L_s)L_\mu] \} \delta(e - e_{\min}) \\
&+ \frac{C_F \alpha_s(\mu)}{4\pi} \{ 16(L_s - L_c) \} \left[ \frac{1}{e - e_{\min}} \right]_+ \\
&+ \frac{C_F \alpha_s(\mu)}{4\pi} \left\{ 2 \left[ \frac{e - e_{\min}}{(e - e_{\min} + r_c \hat{m}^2)^2} - \frac{4 \log\left(1 + \frac{e - e_{\min}}{r_c \hat{m}^2}\right)}{e - e_{\min}} \right] \theta(e - e_{\min}) \right\} + \mathcal{O}(\alpha_s^2).
\end{aligned} \tag{6.22}$$

The sum over all lines except the last one is identical to

$$\frac{mQ^2}{r_c} H(Q, m, \mu_H = \mu, \mu_m = \mu) \int d\ell B_\tau \left[ \frac{Q^2(e - e_{\min}) - r_s Q \ell}{r_c m}, \Gamma = 0, \mu_B = \mu \right] \hat{S}_\tau(\ell, \mu_s = \mu). \tag{6.23}$$

The first two lines of Eq. (6.22) are just the distributional part with  $r_s$  and  $r_c$  set to 1, this is identical to the distributional part of Eq. (3.7). The additional distributional terms that are produced by the measurement function rescalings are gathered in the third and fourth line. The logarithms

$$L_c = \log r_c, \quad L_s = \log r_s, \quad L_m = \log \hat{m}, \quad L_\mu = \log \frac{\mu}{Q} \tag{6.24}$$

appear after using the plus function rescaling relation Eq. (B.6). Note that there is an explicit logarithm of  $\mu$  which should be canceled by corrections in the hard functions in a complete treatment of mass power corrections. For our observables the nonsingular plus distribution coefficient  $B_{\text{plus}}^{\text{ns}}$  is still observable independent after rescaling and the differences between observables for the order  $\alpha_s$  nonsingular delta function coefficient  $A_e^{\text{ns}}$  start at  $\hat{m}^4$  after rescaling ( $\hat{m}^2$  before rescaling). This suggests, that the leading fixed order observable differences at  $\mathcal{O}(\hat{m}^2)$  are completely explained by argument rescalings of the leading power distributions (and obviously the shift of the threshold). The last line, finally, is the contribution from the nondistributional part, which is numerically insignificant in the peak region.

# 7. Results

## 7.1. Analysis of Power Corrections

### 7.1.1. Fits: Rescale Prescription

We have tested the effects of the measurement mass power corrections derived in chapter 6 and different prescriptions (“absorb”, “no absorb”) of including the remaining fixed order power corrections, which are described in Sec. 3.3. The results for PYTHIA using gap 2 are compiled in Fig. 7.1. Other combinations of gap and MC generator do not contain any qualitative differences relevant to the discussion of power corrections.

The first part labeled “default” contains the standard “no absorb” and “no rescaling factors” results. The new observable mJM (modified jet mass), that was introduced at the end of Sec. 6.2 as a quadratic function of sJM, was designed to remove the leading soft rescaling correction appearing in the sJM measurement function (this has automatically set both rescale factors  $r_c$  and  $r_s$  and the threshold  $e_{\min}$  of mJM to have the same  $\hat{m}^2$  corrections as 2-Jettiness). This data point is therefore based on sJM event generator data and the agreement with the 2-Jettiness data point signifies, that the problematic relative power corrections between sJM and 2-Jettiness are well described by the kinematic power corrections in the measurement function.

The “abs” labeled points are obtained by absorbing the coefficients of delta and plus distributions from the QCD nonsingulars into the corresponding coefficients of the bHQET cross section. The remaining regular QCD nonsingulars and SCET nonsingulars are numerically tiny in the peak region, thus all perturbative fixed order power corrections at the given order are essentially absorbed into the resummed cross section. The improved agreement between NLL and N<sup>2</sup>LL extracted MSR masses for measurements of the same observable, i.e. when only comparing points with the same color, is the consequence of resummation also acting on sub power terms. According to the rescale identity Eq. (B.6) a rescaling of the plus distribution is equivalent to changing coefficients of distributions. That means that absorbing all fixed order QCD coefficients should automatically also contain all corrections that would be added by applying the rescale prescription to the perturbative distribution. The rescale corrections that cannot be accounted for in this way are the rescalings of the Breit Wigner (probably small effect) and soft shape function convolution variables. These missing corrections are responsible for the clearly still incompatible sJM data point

The results of only applying the rescale factors  $r_s$  and  $r_c$  in the measurement function to the cross section are plotted in the “rescale” part. We see good agreement between all observables for both mass and  $\Omega_1$  values. The factor  $r_s$  has successfully repaired the relation between the event shape variable and the soft function convolution variable. The

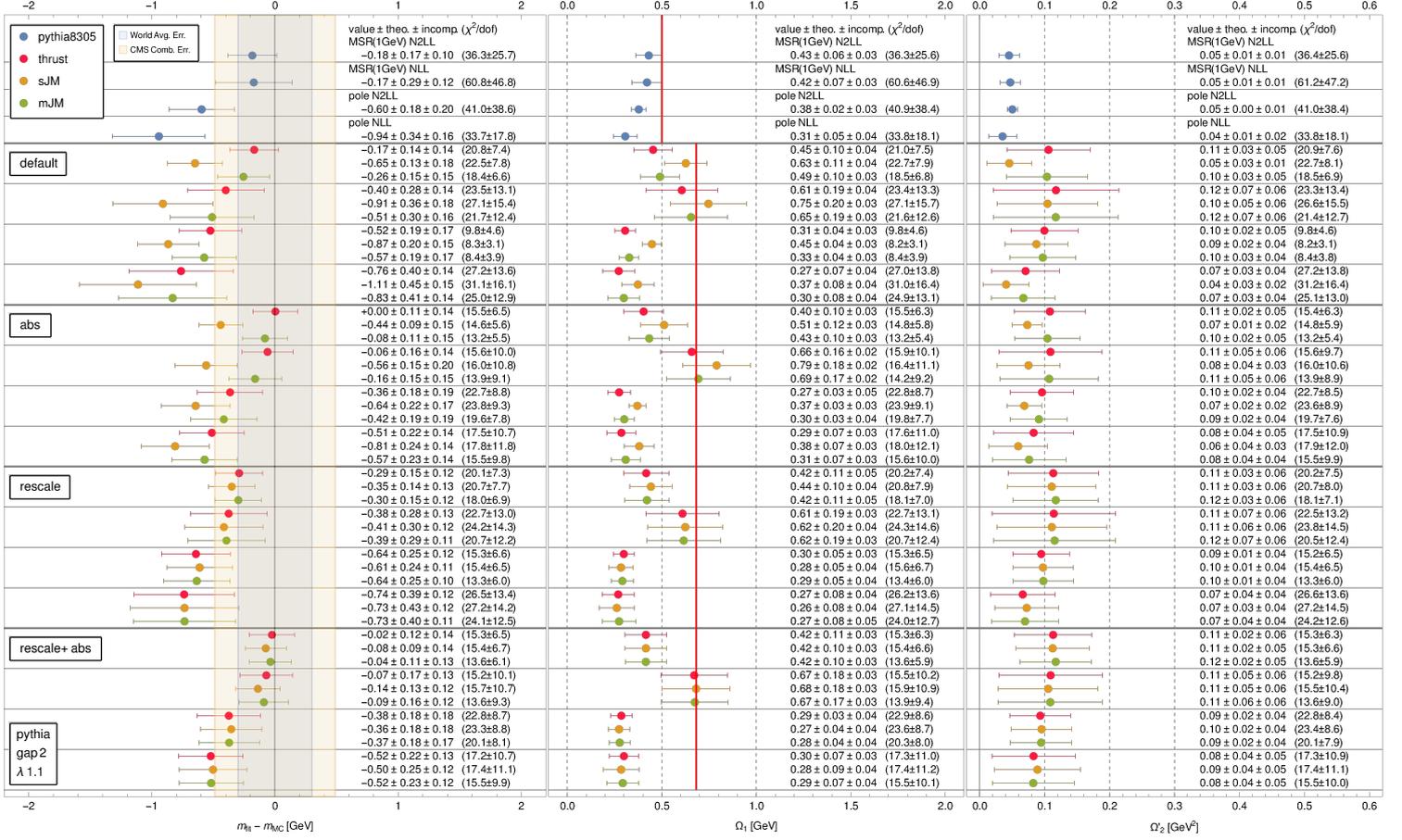


Figure 7.1.: Fit results for PYTHIA using gap 2. The different sections are: “default” has no rescaling and no absorb, “abs” absorbs coefficients of distributions from the nonsingular cross section into the resummed cross section but has no rescaling, “rescale” contains both soft and collinear rescalings  $r_s$  and  $r_c$  but no absorb, “rescale + abs” has absorption and contains both rescale factors.

rescale prescription reduces the observable differences of the nonsingular coefficients by one power of  $\hat{m}^2$  as we have stated at the end of Sec. 6.5, but it turns out that the now largely observable equivalent (in the sense that they have the same  $\hat{m}^2$  corrections apart from those in  $e_{\min}$ )  $N^2\text{LL}$  nonsingulars have significantly increased in size when compared to the “default” prescription in the first panel. This explains the rather large shift of the  $N^2\text{LL}$  results in the wrong direction.

The final prescription for including power corrections is to apply both rescale factors and then to absorb the the remaining power corrections. In the figure this version is referred

to as “rescale + abs”. The main correction that is missing from the “abs” prescription, i.e. the rescaling of the soft shape function convolution variable, is now also included, which leads to 2-Jettiness and mJM results that are comparable to the “abs” results and a repaired sJM result.

We can conclude, that the rescale prescription is mandatory for consistent results. But, we have also discovered that rescaling leads to relative large global (i.e. universal for  $\tau_2$ ,  $\tau_m$ ,  $\tau_s$ ) power corrections and subsequently a large sensitivity to the “absorb” prescription. In the next sections we will devise a procedure to obtain a reasonable estimate for these power correction uncertainties.

### 7.1.2. Partial Absorption of Power Corrections

In section Sec. 3.3 we described the prescription of fully absorbing all distributional QCD power corrections into bHQET, so that only the non-distributional term, which is numerically negligible in the peak regions, remains in the nonsingular, see Eq. (3.49). We will now use a parametrization that will allow us to exactly specify how much of the distributional contributions should remain in the nonsingular. In this partial absorption prescription we set the nonsingular to

$$\begin{aligned} \frac{1}{\sigma_0^C} \frac{d\sigma_{\text{parAbs,ns}}^C}{de} & \stackrel{!}{=} R_0^{C,\text{ns}}(\xi_{A0}; \hat{m}) \delta(e - e_{\min}) + \frac{C_{F\alpha_s}}{4\pi} \left\{ A_e^{C,\text{ns}}(\xi_{R0}, \xi_{A1}; \hat{m}) \delta(e - e_{\min}) \right. \\ & \left. + B_{\text{plus}}^{C,\text{ns}}(\xi_{R0}, \xi_{B1}; \hat{m}) \left[ \frac{1}{e - e_{\min}} \right]_+ + F_e^{\text{NS,C}}(e, \hat{m}) \right\}, \end{aligned} \quad (7.1)$$

where the coefficients are parametrized by (we will now for readability drop the current superscript  $C$  and the  $\hat{m}$  dependence in the argument)

$$\begin{aligned} R_0^{\text{ns}}(\xi_{A0}) &= (1 - \xi_{A0}) R_0^{\text{ns}}(0) \\ A_e^{\text{ns}}(\xi_{R0}, \xi_{A1}) &= (1 - \xi_{A1}) A_e^{\text{ns}}(\xi_{R0}, 0) \\ B_{\text{plus}}^{\text{ns}}(\xi_{R0}, \xi_{B1}) &= (1 - \xi_{B1}) B_{\text{plus}}^{\text{ns}}(\xi_{R0}, 0) \end{aligned} \quad (7.2)$$

and the baseline for the  $\xi_{A1}$ ,  $\xi_{B1}$  parametrization is given by

$$\begin{aligned} A_e^{\text{ns}}(\xi_{R0}, 0) &= A_e^{\text{ns}}(0, 0) - \xi_{R0} R_0^{\text{ns}}(0) A_e^{\text{s}}(0, 0) \\ B_{\text{plus}}^{\text{ns}}(\xi_{R0}, 0) &= B_{\text{plus}}^{\text{ns}}(0, 0) - \xi_{R0} R_0^{\text{ns}}(0) B_{\text{plus}}^{\text{s}}(0, 0) \end{aligned} \quad (7.3)$$

The fixed order (FO) contribution of the singular cross section after absorption is then given by the FO QCD coefficients defined in Eq. (3.8) minus the contribution, that

remains in the nonsingular

$$\begin{aligned}
R_0^{\text{QCD}} - R_0^{\text{ns}}(\xi_{A0}) &= 1 + \underbrace{R_0^{\text{ns}}(0) - R_0^{\text{ns}}(\xi_{A0})}_{R_0^{\text{abs}}} \equiv R_0^{\text{newsing}} \\
A_e^{\text{QCD}} - A_e^{\text{ns}}(\xi_{R0}, \xi_{A1}) &= \underbrace{A_e^{\text{s}}(0, 0)}_{A_e^{\text{base}}} + \underbrace{A_e^{\text{ns}}(0, 0) - A_e^{\text{ns}}(\xi_{R0}, \xi_{A1})}_{A_e^{\text{abs}}} \\
B_{\text{plus}}^{\text{QCD}} - B_{\text{plus}}^{\text{ns}}(\xi_{R0}, \xi_{B1}) &= \underbrace{B_{\text{plus}}^{\text{s}}(0, 0)}_{B_{\text{plus}}^{\text{base}}} + \underbrace{B_{\text{plus}}^{\text{ns}}(0, 0) - B_{\text{plus}}^{\text{ns}}(\xi_{R0}, \xi_{B1})}_{B_{\text{plus}}^{\text{abs}}},
\end{aligned} \tag{7.4}$$

where  $R_0^{\text{s}}(\xi_{A0} = 0)$ ,  $A_e^{\text{s}}(\xi_{R0} = 0, \xi_{A1} = 0)$ ,  $B_{\text{plus}}^{\text{s}}(\xi_{R0} = 0, \xi_{B1} = 0)$  are the delta and plus distribution coefficients of the FO singular distribution before any absorption, given in Eq. (3.7) and (6.22), and  $R_0^{\text{ns}}(0)$ ,  $A_e^{\text{ns}}(0, 0)$ ,  $B_{\text{plus}}^{\text{ns}}(0, 0)$  are the nonsingular coefficients before any absorption defined in Eq. 3.5. The right hand sides of Eq. (7.4) have the form of Eq. (3.50) where the “base” superscript terms are identical and the “abs” superscript terms are modified due to the remaining nonsingulars. We can therefore directly use Eq. 3.51 with these definitions of “base”, “abs” and “newsing” terms to obtain the correct modification of the bHQET cross section. It is clear from the equations above, that the sum of the resulting singular and nonsingular distributions will automatically add up to the QCD distribution at FO for any choice of the parameters  $\xi_{A0}$ ,  $\xi_{R0}$ ,  $\xi_{A1}$  and  $\xi_{B1}$ .

Let us now explain the meaning of these parameters.  $\xi_{A0}$  is the fractional amount of the original tree level delta function nonsingular coefficient, that is absorbed into the bHQET distribution. Hence,  $1 - \xi_{A0}$  is the fractional amount that remains in the nonsingular. At NLL,  $\xi_{A0} = 0$  therefore corresponds to the “no absorb” prescription and  $\xi_{A0} = 1$  to the “full absorb” prescription.

$\xi_{A1}$  and  $\xi_{B1}$  do the same for the order  $\alpha_s$  delta function and plus distribution coefficients respectively.  $\xi_{A1} = 1$  or  $\xi_{B1} = 1$  again correspond to full absorption, i.e. the nonsingular coefficient vanishes. However the baseline for the parametrization, that is the nonsingular coefficients for  $\xi_{A1} = 0$  or  $\xi_{B1} = 0$ , can be taken either as the nonsingular coefficients before any absorption  $A_e^{\text{ns}}(0, 0)$ ,  $B_{\text{plus}}^{\text{ns}}(0, 0)$  or as the recalculated nonsingular coefficients after absorption of the tree level coefficient, treated as part of a global multiplicative factor, as it is illustrated by the first line of Eq. (3.51). This global factor contains the tree level power corrections, but it will also add power corrections to the order  $\alpha_s$  singular FO coefficient due to mixed terms of the form (power correction  $\times \alpha_s$  leading power). The parameter  $\xi_{R0}$  then governs how much of these mixed terms we consider to be part of the baseline for the  $\xi_{A1}$  and  $\xi_{B1}$  parametrization.  $\xi_{R0} = 0$  sets the baseline to the order  $\alpha_s$  nonsingular coefficients before any absorption and  $\xi_{R0} = \xi_{A0}$  sets the baseline to the recalculated nonsingular coefficients after absorption of the tree level coefficient as a global factor.

For example, the choice  $\xi_{A0} = 1$ ,  $\xi_{R0} = \xi_{A0}$ ,  $\xi_{A1} = \xi_{B1} = 0$  means that the tree level coefficient is fully absorbed, the order  $\alpha_s$  nonsingular baseline is recalculated and we absorb nothing from this recalculated  $\alpha_s$  nonsingular. For this choice the new bHQET hard and jet functions in Eq. 3.51 are identical to the old ones before any absorption:

$\tilde{H}_Q = H_Q$ ,  $\tilde{H}_m = H_m$ ,  $\tilde{B} = B$ . The only change is the global factor  $R_0^{\text{newsing}} = 1 + R_0^{\text{abs}}$  in front of everything.

For the second example we take the choice  $\xi_{A0} = 1$ ,  $\xi_{R0} = 0$ ,  $\xi_{A1} = \xi_{B1} = 0$ . The tree level coefficient is still fully absorbed, but now the order  $\alpha_s$  nonsingular baseline does not change due to the absorbed tree level coefficient, we then proceed to absorb nothing measured from this baseline. The resulting order  $\alpha_s$  nonsingular coefficients are therefore exactly the same as in the “no absorb” case. However, the new bHQET hard and jet functions are now different than the old ones before any absorption:  $\tilde{H}_Q \neq H_Q$ ,  $\tilde{H}_m \neq H_m$ ,  $\tilde{B} \neq B$ , since the mixed terms produced by the global factor  $R_0^{\text{newsing}}$  have to be canceled explicitly at one loop, so that the global factor is effectively removed from the order  $\alpha_s$  singular distribution. This choice  $\xi_{R0} = 0$  is useful to quantify how much each coefficient of the “no absorb” nonsingular distribution separately affects the fit result. Each of the nonsingular coefficients in this case interpolates independently between the “no absorb” prescription value and 0 (“full absorb”). This version will however provide unrealistic and over-conservative estimates of power correction uncertainties, if the global  $R_0^{\text{newsing}}$  factor in the singular distribution generates the correct power corrections at order  $\alpha_s$  by itself. To avoid a subtraction of these correct terms in the singular distribution we would therefore use  $\xi_{R0} = \xi_{A0}$ , so that the order  $\alpha_s$  nonsingular coefficients now interpolate between the recalculated nonsingular coefficients, that have a smaller absolute value than before, and 0. For our analytic FO expressions the tree level coefficient seems to indeed factorize to a large extent.

The QCD tree level coefficient absorbed into the singular as a global multiplicative factor  $R_0^{\text{newsing}} = R_0^{\text{QCD}} = 1 + R_0^{\text{ns}}(0)$  produces by itself all next to leading  $\hat{m}^2$  power corrections of the plus distribution coefficient  $B_{\text{plus}}^{\text{QCD}}$ . Note, that this only happens in the full “rescale” prescription case, because kinematic power corrections would otherwise be missing. The first interesting thing that happens is that  $R_0^{\text{V,QCD}} \neq R_0^{\text{A,QCD}}$  and  $B_{\text{plus}}^{\text{V,QCD}} \neq B_{\text{plus}}^{\text{A,QCD}}$ , but

$$\frac{B_{\text{plus}}^{\text{V,QCD}}}{R_0^{\text{V,QCD}}} = \frac{B_{\text{plus}}^{\text{A,QCD}}}{R_0^{\text{A,QCD}}}, \quad (7.5)$$

i.e. the tree level coefficients are different for vector and axial-vector currents and the same applies to the plus distribution coefficients, but the ratios are exactly the same. This supports the treatment of the tree level coefficient as global factor. Furthermore, we can now subtract the singular contribution from the tree level subtracted plus function coefficient

$$\frac{B_{\text{plus}}^{\text{QCD}}}{R_0^{\text{QCD}}} - B_{\text{plus}}^{\text{s}}(0, 0) = -32 \log(\hat{m}) \hat{m}^4 + \mathcal{O}(\hat{m}^6). \quad (7.6)$$

This shows, that we only need the assumption that  $R_0^{\text{QCD}}$  is a global factor, to account for all next to leading power  $\hat{m}^2$  corrections in  $B_{\text{plus}}^{\text{QCD}}$ . It should be clear, that the kinematic corrections from the rescale prescription, given in Eq. (6.22) have been taken into account in  $B_{\text{plus}}^{\text{s}}(0, 0)$  to obtain the expression above. It should also be mentioned, that the corrections in  $R_0^{\text{V,QCD}}$  start at  $\hat{m}^4$ , while those in  $R_0^{\text{A,QCD}}$  start at  $\hat{m}^2$ . The leading  $\hat{m}^2$

corrections of  $B_{\text{plus}}^{\text{V,QCD}}$  are therefore accounted for by rescale corrections alone. All the objects in the equation above are, at least for the observables that we use, observable independent.

The same treatment applied to the order  $\alpha_s$  delta function coefficient cannot account for all leading power  $\hat{m}^2$  corrections, which is not surprising, because it is very probable that these power corrections will also appear in the hard functions. We nevertheless observe that the corresponding nonsingulars are significantly reduced.

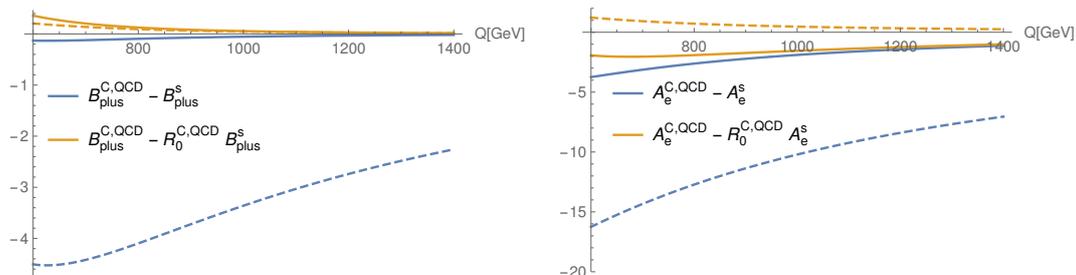


Figure 7.2.: Order  $\alpha_s$  nonsingular plus distribution (left) and delta distribution (right) coefficients with (yellow) and without (blue) absorption of tree level delta function coefficient as global factor. Shown separately for vector (solid) and axial-vector (dashed) current. The top mass is set to 173 GeV.

Fig. 7.2 shows the nonsingular coefficients with and without absorbing the tree level coefficient as a global factor for vector and axial-vector contributions. These values have to be multiplied with the vector and axial-vector born cross section, given in Eq. (3.9), respectively to obtain their relative contribution to the total nonsingular coefficients. The range of the born cross section ratios within our  $Q$ -range is  $\sigma_0^V/\sigma_0^A \in [3.67, 3.80]$ . In these plots we have included all rescale prescription corrections in the singular coefficients, except for the  $\log \frac{\mu}{Q}$  correction shown in Eq. (6.22), which can be treated separately due to its  $\mu$  dependence and its observable independence.

Below Eq. (6.22) we stated that the differences between our observables for the order  $\alpha_s$  nonsingular coefficient  $A_e^{\text{ns}}$  start at power  $\hat{m}^4$  after the rescale prescription ( $\hat{m}^2$  before rescale prescription). This difference is further reduced after absorbing the tree level coefficient  $R_0^{\text{QCD}}$  into the singular as global factor, the difference now starts at  $\hat{m}^6$ . The reason is that a  $\hat{m}^{2n}$  term in the observable independent  $B_{\text{plus}}^{\text{ns}}$  coefficient leads to observable dependent rescaling corrections starting at  $\hat{m}^{(2+1)n}$ . These observable differences then finally end up in  $A_e^{\text{ns}}$  after using the rescaling identity (B.6) to bring the arguments into the standard form shown in Eq. (3.5). Since the absorption of  $R_0^{\text{QCD}}$  as a global factor moves all  $\hat{m}^2$  terms of  $B_{\text{plus}}^{\text{QCD}}$  and its associated  $\hat{m}^4$  rescale corrections into the singular distribution, the remaining nonsingular coefficient  $B_{\text{plus}}^{\text{ns}}$  starting at  $\hat{m}^4$  can only lead to  $\hat{m}^6$  rescaling corrections in  $A_e^{\text{ns}}$ . The  $A_e^{\text{ns}}$  plot in Fig. 7.2 is therefore representative for any of our observables ( $B_{\text{plus}}^{\text{ns}}$  plot is observable independent).

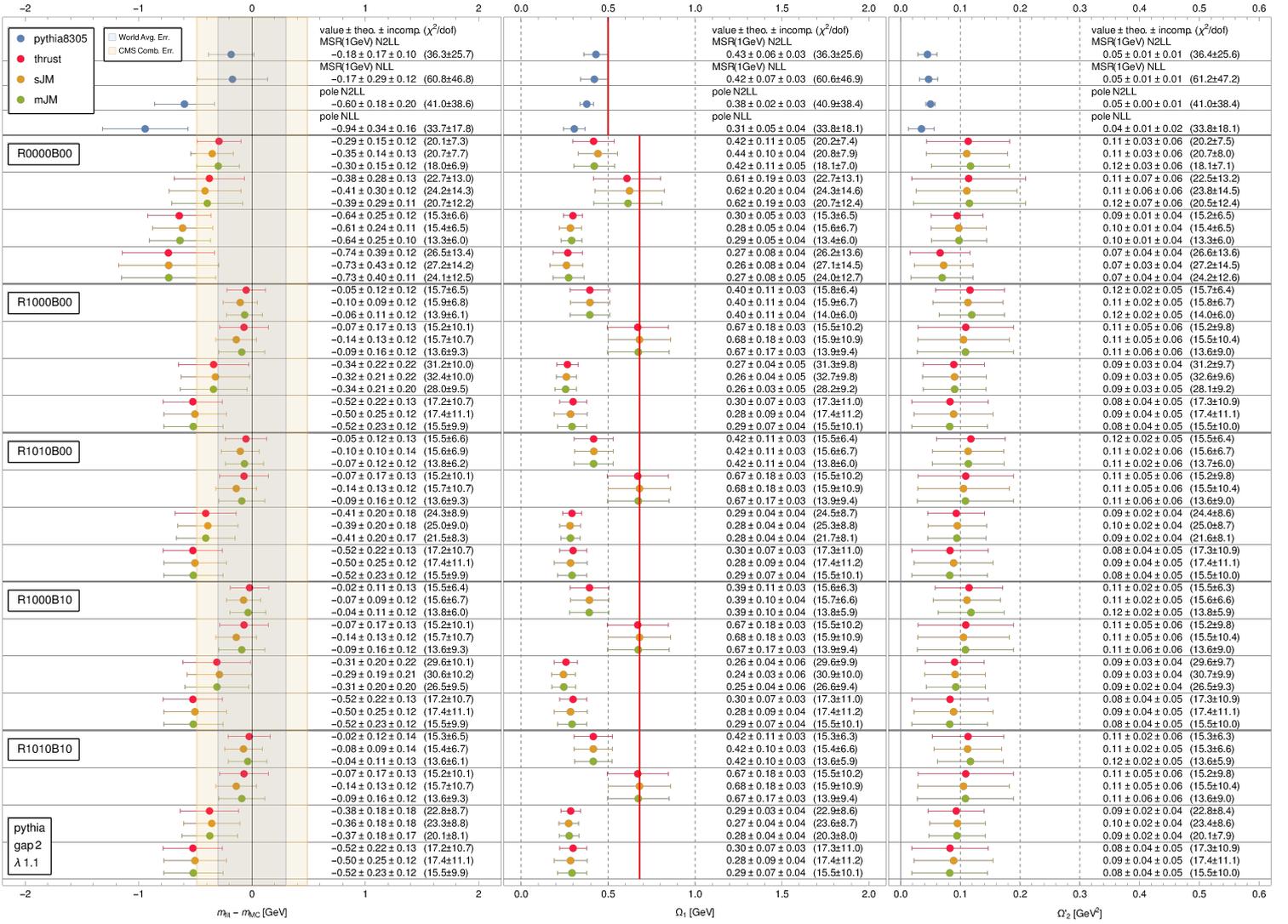


Figure 7.3.: Fit results using partial absorption parametrization. R0000B00: no absorb, R1000B00: tree level coefficient absorbed as global factor, R1010B00: both delta coefficients absorbed, R1000B10: tree level and plus distribution absorbed, R1010B10: all distributional nonsingulars absorbed.

### 7.1.3. Fits: Partial Absorption

#### Effects of Nonsingular Coefficients and Uncertainty Estimate

Fig. 7.3 shows fit results with different partial absorb parameters. The kinematic rescaling factors are from now on always included as default. Our naming scheme is given by

$$R \underbrace{xx}_{10 \times \xi_{A0}} \overbrace{(m)yy}^{(-)10 \times \xi_{A1}} B \underbrace{(m)zz}_{(-)10 \times \xi_{B1}}, \quad (7.7)$$

where the  $\xi$ 's have been defined and explained in Sec. 7.1.2. We will always set  $\xi_{R0} = \xi_{A0}$ , so that  $\xi_{A1}$  and  $\xi_{B1}$  correspond to fractional absorption of the remaining nonsingulars after absorbing the tree level coefficient  $1 + R_0^{\text{abs}}(\xi_{A0})$  as global factor.

We saw in the last section, that absorbing the tree level coefficient as global factor also leads to large analytical cancellations of terms in the order  $\alpha_s$  nonsingular coefficients. This is also reflected in the fit results. All tree level absorbed versions with different combinations of absorption of the remaining order  $\alpha_s$  nonsingular coefficients, between no absorb and full absorb, are very close to the treatment, where all distributional nonsingulars are fully absorbed.

In the next Fig. 7.4 we show fits with a larger variation of the remaining order  $\alpha_s$  nonsingular coefficients. We let  $\alpha_s$  delta function and plus distribution parameters change in a correlated way, to avoid numerical cancellations of these nonsingulars due to their signs. The labels are explicitly explained in the figure caption. R10m20Bm20 sets the remaining order  $\alpha_s$  nonsingular to 3 times the standard R1000B00 nonsingular. We see in this extreme case, i.e. R10m20Bm20, that the N<sup>2</sup>LL MSR fits are still very insensitive to these nonsingulars, whereas the N<sup>2</sup>LL pole fits show a relative large increase of their uncertainties. This means that the nonsingulars are relative large again (about 3/5 of “no absorb” nonsingular but with slightly shifted peak position) and that this choice would lead to an overestimate of the uncertainties. Also note, that the large uncertainty in the fitted N<sup>2</sup>LL MSR  $\Omega_1$  value for R10m20Bm20 is due to some very small peak like features at the side of the actual peak caused by negative nonsingulars, that are miss-recognized for some few profiles, due to the normalization applied over the fit range. This problem could be fixed by carefully restricting the fit range of  $\Delta_0$ . In general, the fit parameters show a very mild dependence on variations of the order  $\alpha_s$  absorption parameters.

Based on these results and on the analytical properties of these coefficients, presented in Sec. 7.1.2, we decided for the uncertainty estimate of these power corrections to always fully absorb the tree level coefficient as global factor, that is  $\xi_{A0} = \xi_{R0} = 1$ , and to include the order  $\alpha_s$  parameters as independent random parameters in our 501 profiles. We use the range  $-1 < \xi_{A1} < 1$  and  $-1 < \xi_{B1} < 1$ . This corresponds to a variation between full absorb R1010B10 and R10m10Bm10, where the latter has 2x of the nonsingulars of R1000B00. This should provide a reasonably conservative estimate of the remaining power corrections given the analytical results that are currently available to us. This final version is already included in Fig. 7.4 with the label “parAbs” and it is per construction compatible with all other reasonably parametrized fit results.

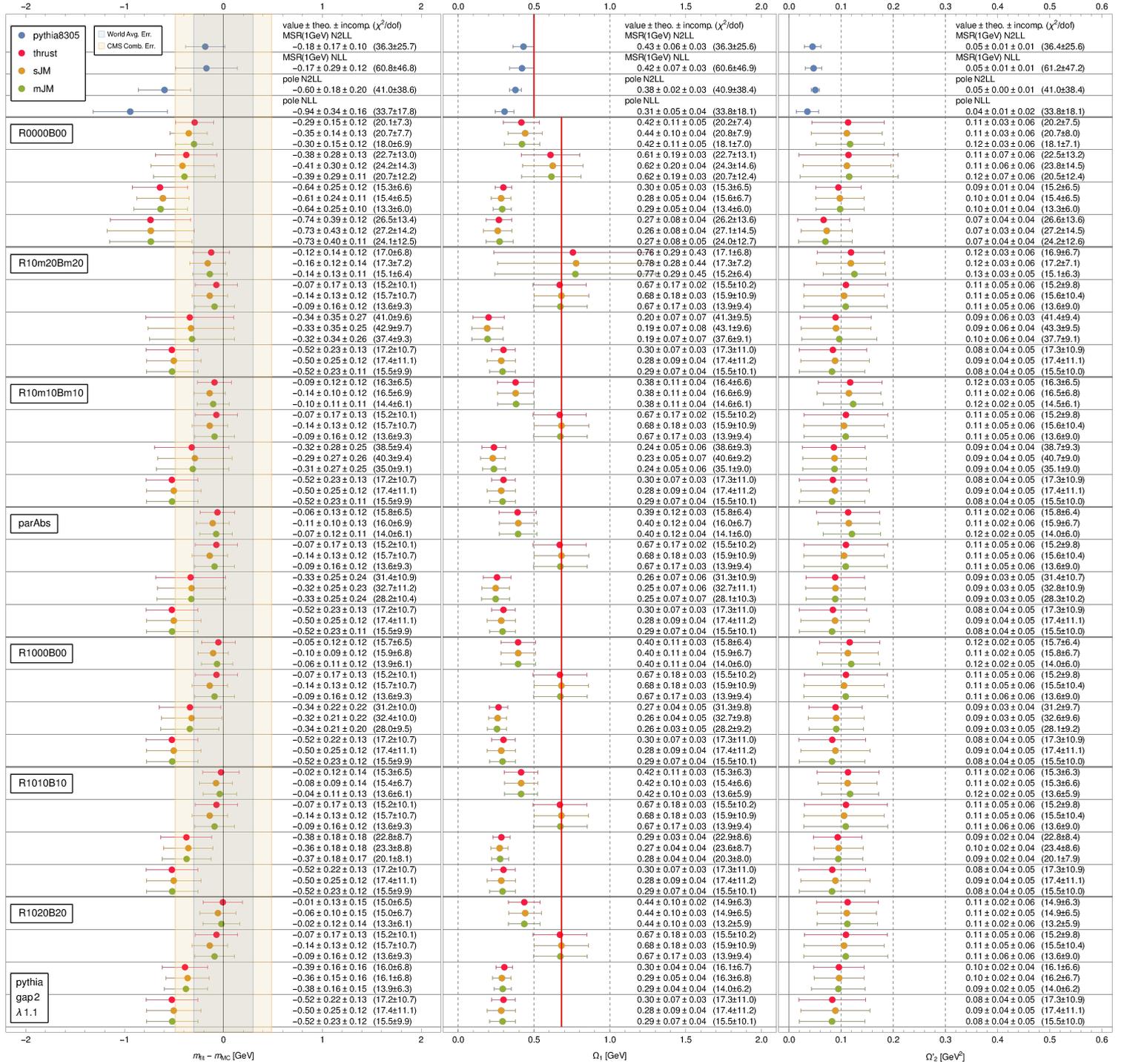


Figure 7.4.: Fit results using partial absorption parametrization. R0000B00: no absorb, R1000B00: tree level coefficient absorbed as global factor. Taking the nonsingulars in R1000B00 as  $x$ , the nonsingulars in the following versions are R10m20Bm20:  $3x$ , R10m10Bm10:  $2x$ , R1010B10:  $0x$ , R1020B20:  $-x$ . parAbs: independent random profile variation of  $\xi_{A1}$  and  $\xi_{B1}$  between R10m10Bm10 and R1010B10.

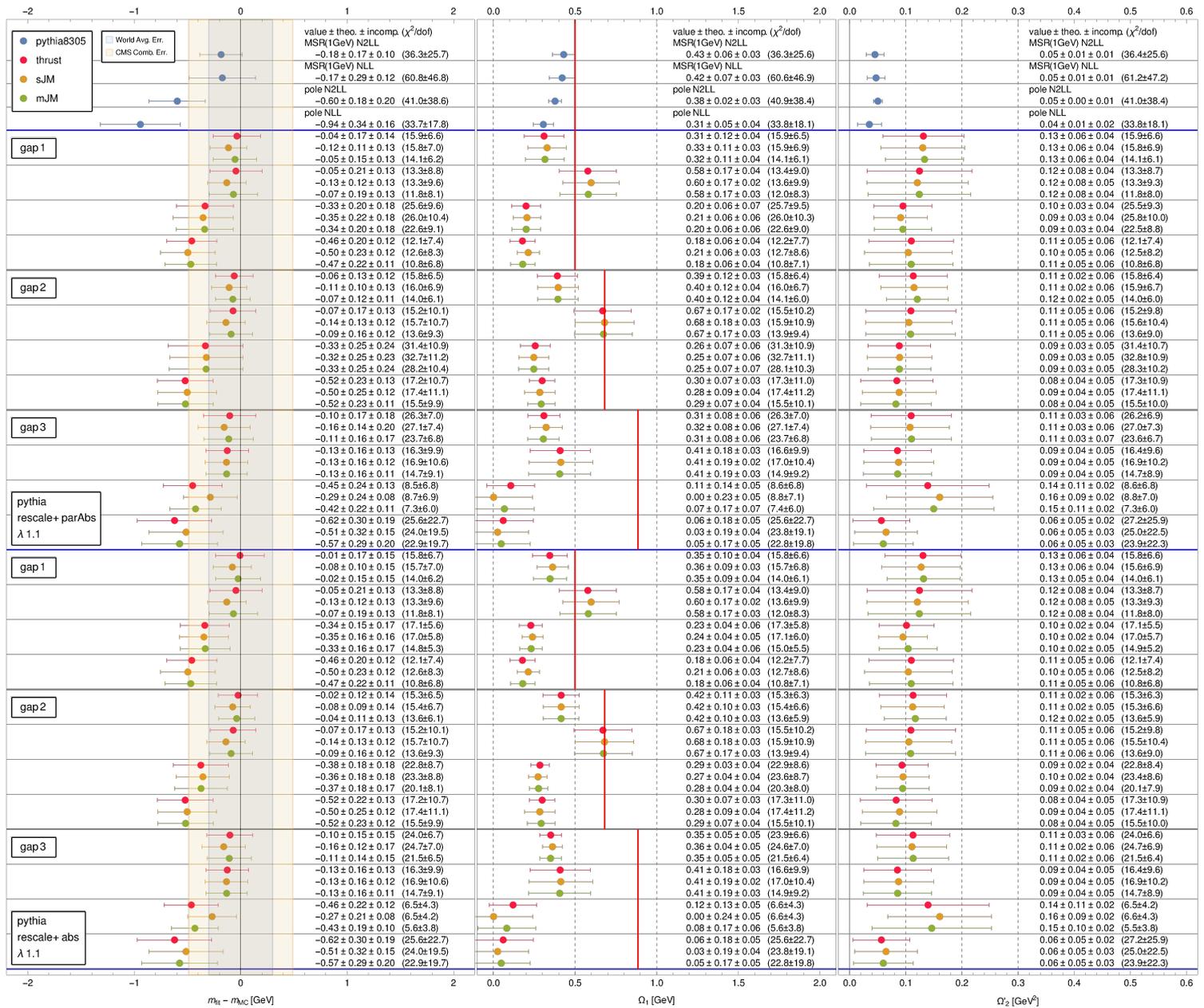


Figure 7.5.: Gap scheme and partial vs. full absorb comparison for PYTHIA

### Soft Gap Scheme Comparison for different MCs

In Fig. 7.5, 7.6 and 7.7 we show gap comparisons and “partial absorb” (“parAbs”) versus “full absorb” (“abs”) for PYTHIA, HERWIG, and SHERPA respectively. First, we can conclude based on these figures, that using partial absorption to provide an uncertainty estimate of power corrections does not change the qualitative behavior of the fits compared

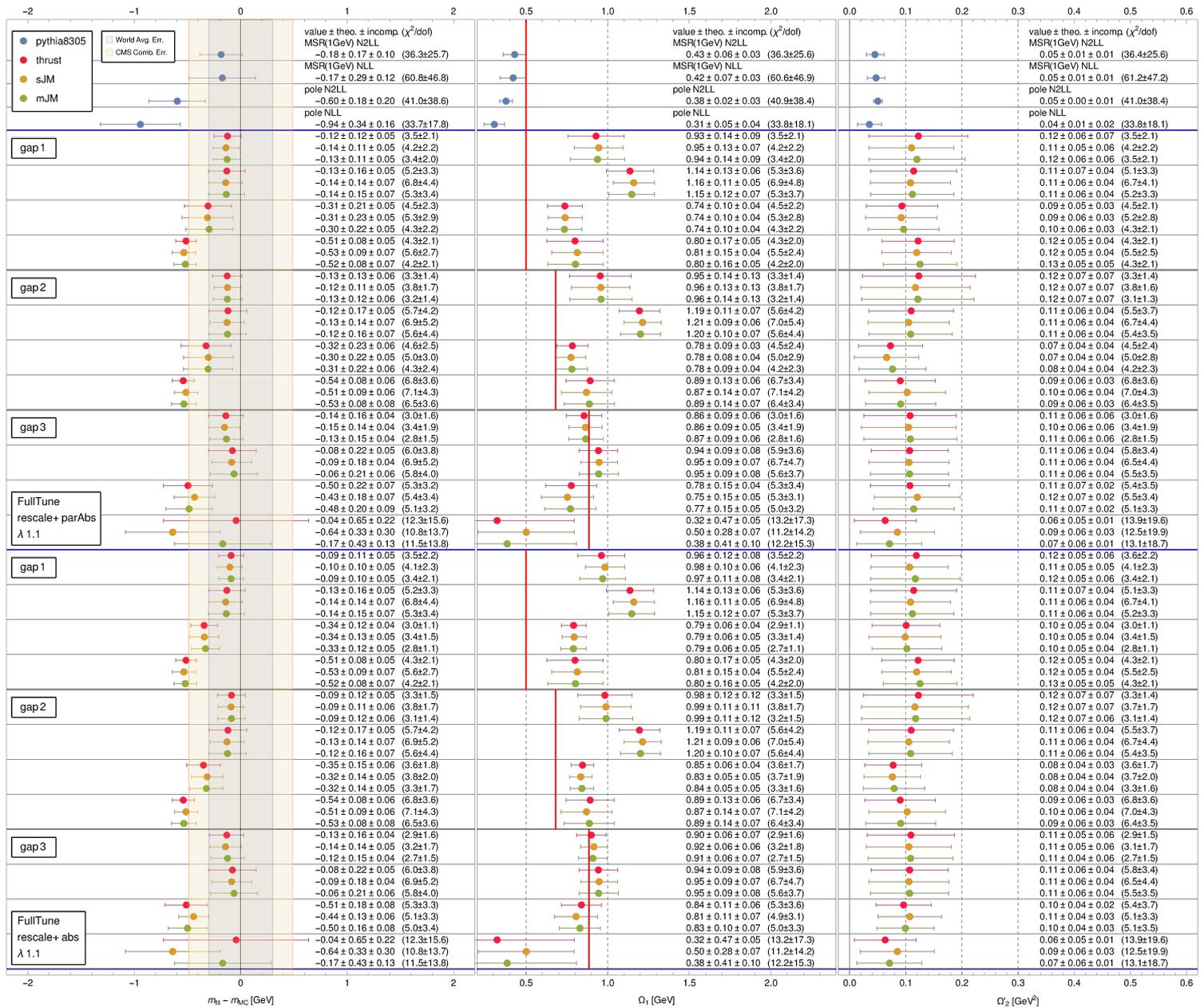


Figure 7.6.: Gap scheme and partial vs. full absorb comparison for HERWIG

to a fully absorbed treatment. This applies to every single fit and also to those that exhibit irregularities. We can therefore restrict ourselves to the more conservative “parAbs” version, to discuss gap related behaviour, since this is also the version that will be used to present the final results.

Gap 1 and gap 2 show very similar results overall. The fitted MSR mass values are in general compatible between all gaps. The gap 3 has better convergence for the MSR

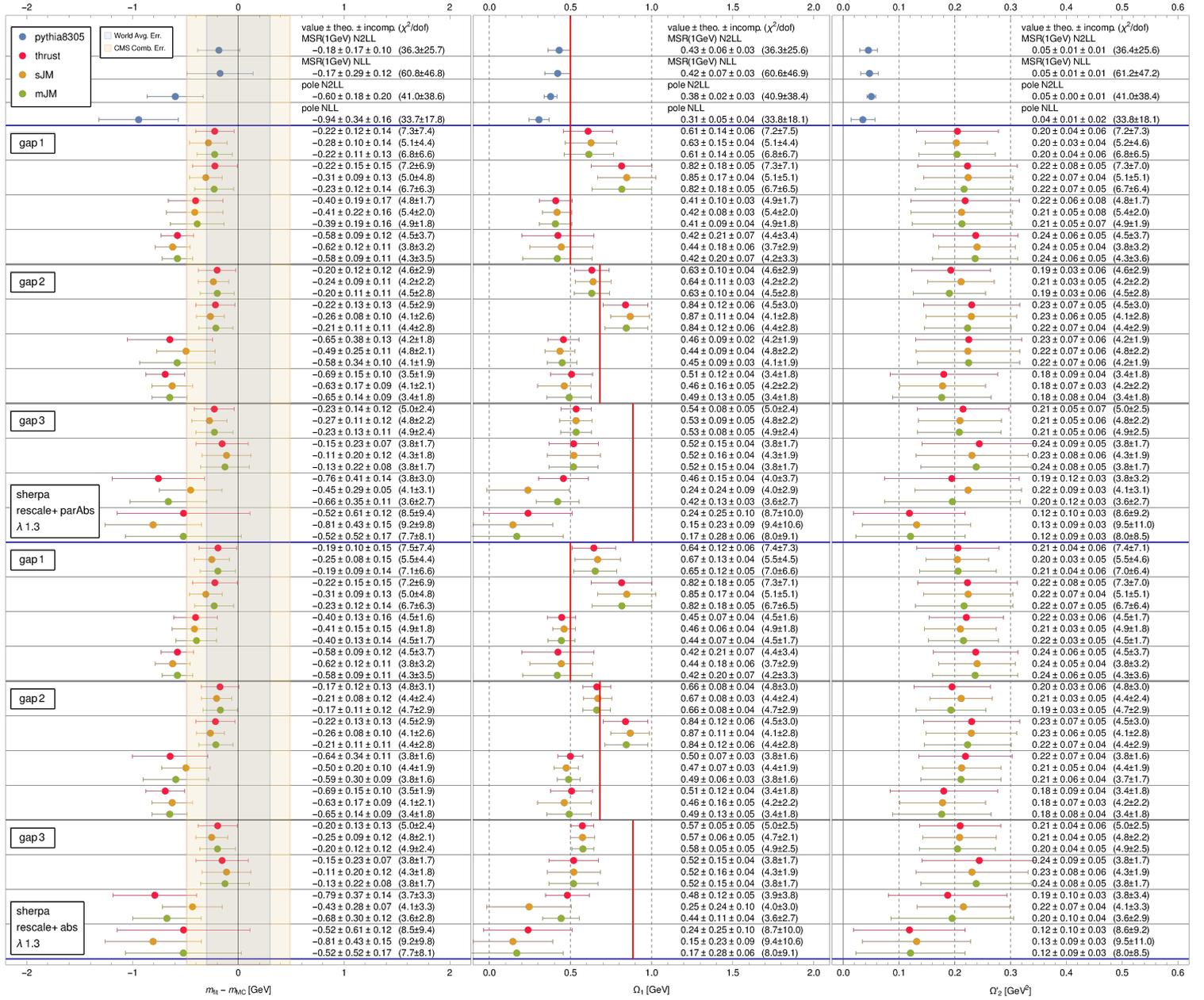


Figure 7.7.: Gap scheme and partial vs. full absorb comparison for SHERPA

$\Omega_1$  values between orders. But gap 3 has problems, when used to fit the pole mass: Gap 3 exhibits large scale dependence. For the NLL cross section this means, that the combination of event shape variable and gap parameter  $\tau - \tau_{\min}(m) - r_s 2\Delta(R_s(\tau))/Q$  in the argument of the factorization theorem is close to stationary or not monotonously increasing as a function of  $\tau$  right after the peak for some profiles with fast changing scales. This leads to a stretched peak shape or a double peak, which are responsible for

the large errors seen in the fits. In the NLL MSR case this problem does not happen, since the scale dependence of the mass  $m_{\text{MSR}}(R(\tau))$  in the threshold  $\tau_{\text{min}}(m_{\text{MSR}})$  has the opposite sign, which partially cancels the gap running numerically. Gap 3 also has some issues when used for N<sup>2</sup>LL pole fits, due to the large gap subtraction  $\delta_3(R_s)$ , which leads to worse perturbative convergence. This problem also does not occur for N<sup>2</sup>LL MSR fits, because the mass subtraction term  $\delta m^{\text{MSR}}(R)$  combines with the gap subtraction into a small total subtraction, since they contribute with opposite signs. Hence, gap 3 pole NLL and N<sup>2</sup>LL results are in general less reliable than those using the other gaps.

Gap 2 provides overall MSR mass fits with the smallest uncertainties. The gap 2 pole fits are also in general very similar to the gap 1 results, but for SHERPA we can see some irregularities, which do not occur for the other MCs and which might therefore be caused by some peculiar combination of factors.

## 7.2. Final Results

Our final results for the three event generators PYTHIA, HERWIG and SHERPA are shown in Fig. 7.8, 7.9 and 7.10 for the three gap schemes respectively. Details on the generator settings and versions are given in the last paragraph below Eq. (5.5). The errors of these results contain an estimate of remaining power correction uncertainties obtained by profile variations of nonsingular partial absorb parameters described in the last sections. We also explained in the last section why for pole mass fits the gap 3 results are less reliable than those using other gaps.

The Monte Carlo event generator mass parameters  $m_{\text{MC}}$  using standard tunes agree with each other and the MSR( $R = 1$  GeV) mass within our uncertainties of about 200 MeV. For HERWIG it is known that its mass parameter depends on the shower cutoff  $Q_0$  [11], where a change of  $Q_0$  by about 1 GeV is related to a  $\sim 200$  MeV change of the meaning of the mass. These similar fitted masses therefore suggest, that the theoretical differences between different MC masses are also of about the same size.

In contrast, the parameters for the soft shape function that we obtain are vastly different for different MCs. These differences are also reflected in the Monte Carlo distributions, that are used for these fits, which look completely different and do not agree with each other at all. The differences of the fitted  $\Omega_1$ 's are equivalent to differences, which increase with  $Q$ , of the peak positions of about up to 3 GeV in units of the jet mass, defined as  $m_J = Q\sqrt{\tau/2}$ , for which the peak position value is close to the top quark mass  $m_{J,\text{peak}} \sim m_t$ . The large  $\Omega_2$  value of SHERPA reflects an  $\sim 1.5$  times wider peak compared to the other MCs. This suggests, that the MC generator description of nonperturbative soft effects using their standard tunes does not work well for our observables. But it seems, that by simultaneous fits of the shape parameters, we can successfully separate the problematic soft behaviour from the collinear physics and therefore obtain the correct mass dependencies of these MC event generators.

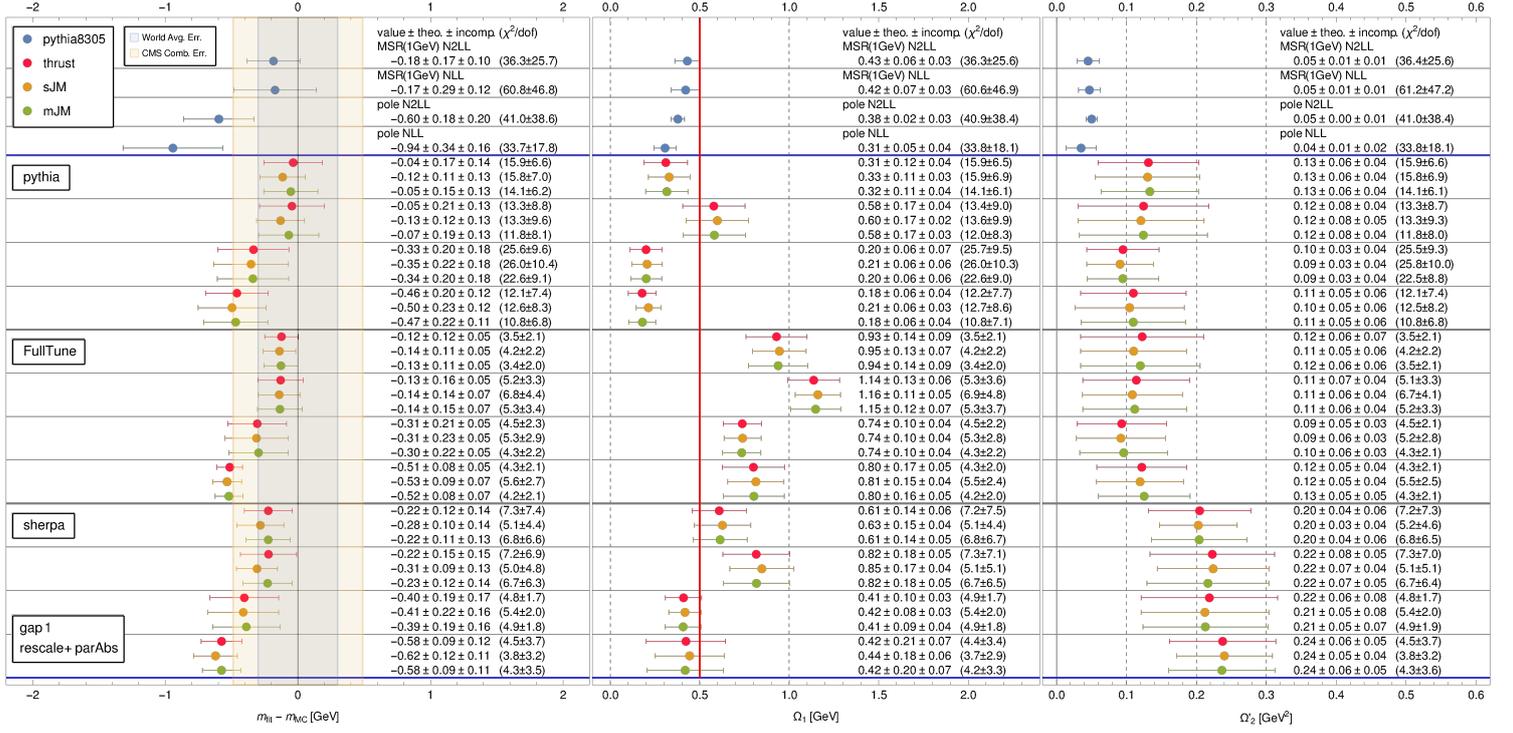


Figure 7.8.: Final results for PYTHIA, HERWIG (FullTune) and SHERPA using gap 1. The “pythia8305” reference points use the old method from [17]. The improved results are obtained from the event shapes 2-Jettiness (thrust), sum of jet mass (sJM) and modified jet mass (mJM). The ordering of the sections is: MSR N<sup>2</sup>LL, MSR NLL, pole N<sup>2</sup>LL and pole NLL.

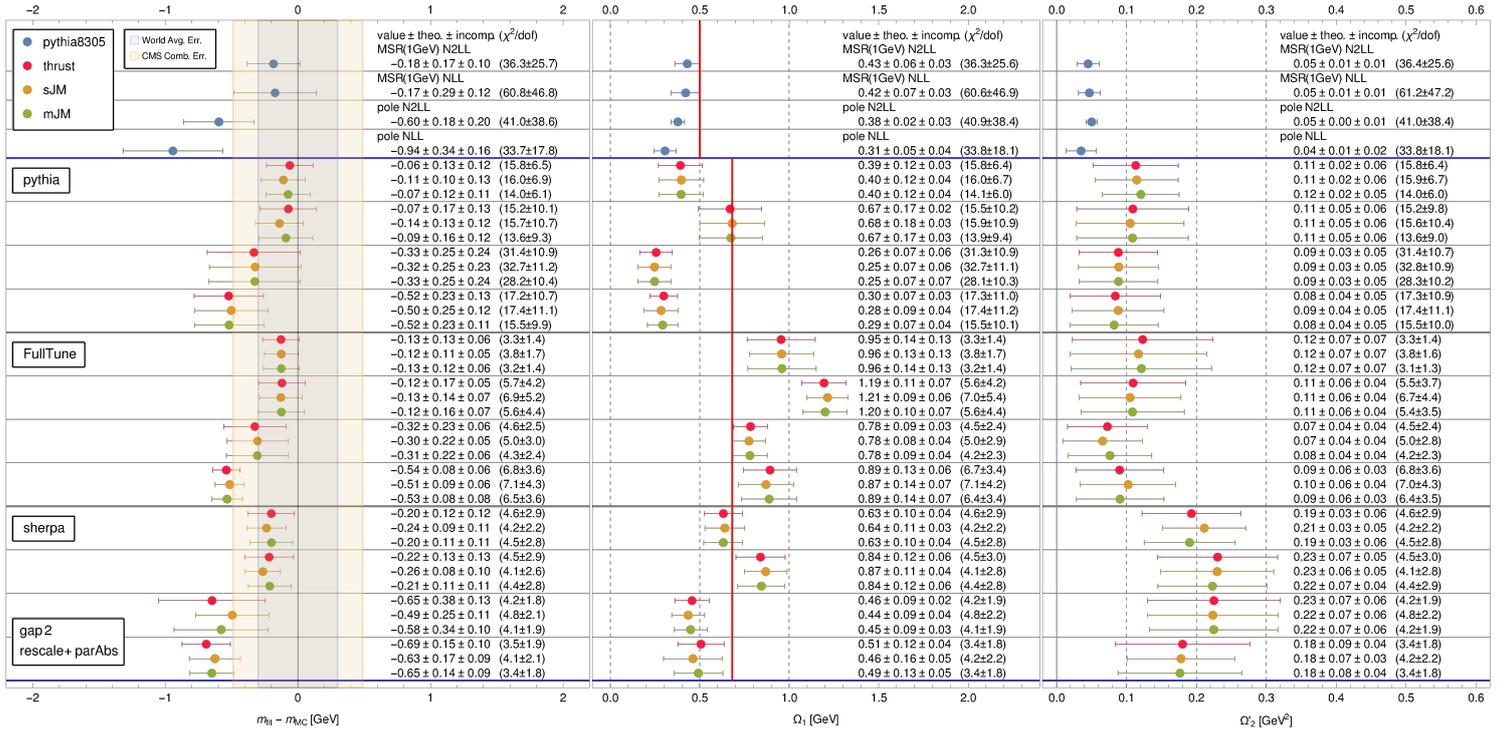


Figure 7.9.: Final results for PYTHIA, HERWIG and SHERPA using gap 2.

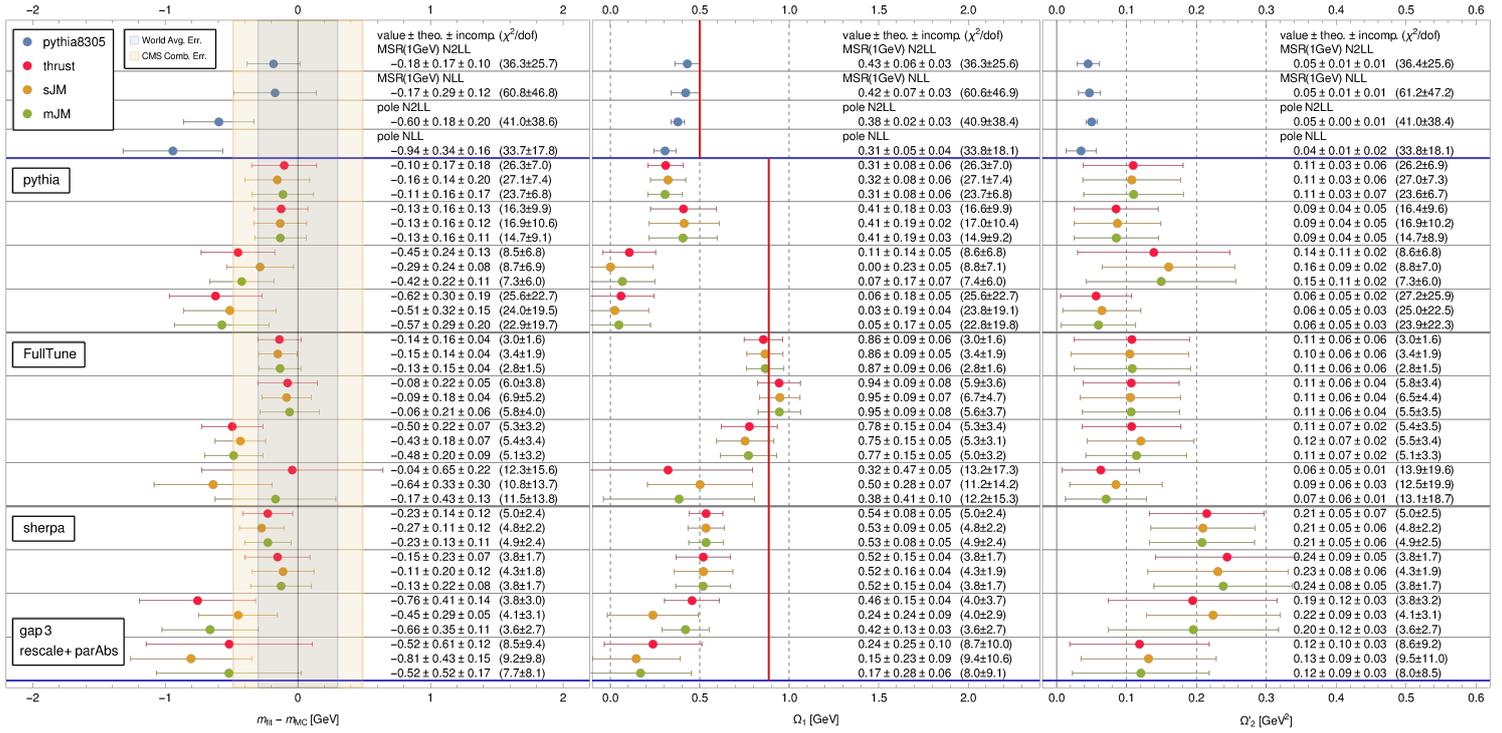


Figure 7.10.: Final results for PYTHIA, HERWIG and SHERPA using gap 3.

## 8. Conclusions

We have tested the original Monte Carlo event generator top quark mass calibration framework of [17] using different observables and soft gap schemes. We discovered that the original fit parametrization required an additional fit parameter to correctly work with these changes and to guarantee an unbiased determination when using the original setup. This improvement is not only necessary to correctly capture the effects of different gaps, but also to reliably obtain non-perturbative shape function parameters and to quantify inconsistencies arising from insufficient treatment of power corrections.

These inconsistencies of the original treatment of power corrections occurred for observables, that have  $(m/Q)^2$  power correction in the relation between the event shape variable and soft momenta. Our new “rescale” prescription, which includes these power corrections in the bHQET measurement function, cures the leading power corrections between observables, both collinear and soft, which leads to very consistent fit results between different observables.

The inclusion of these kinematic measurement power corrections made apparent, that global observable independent power corrections are also sizable and that there is a relative large difference between including these fixed order QCD matched power corrections as a fixed order correction (nonsingular) or fully absorbing them into the boundary conditions of bHQET, which was called “absorb” prescription in Ref. [29].

We analyzed the nonsingular coefficients and showed that absorbing only the scale independent tree level coefficient as global factor multiplying the bHQET cross section also completely accounted for the leading power corrections of the order  $\alpha_s$  plus distribution coefficient and a large part of the order  $\alpha_s$  delta function coefficient. Based on these facts we were able to provide a reasonable uncertainty estimate of these remaining power corrections in our final version of the mass calibration procedure, with only minimal increases in the total uncertainty.

This optimized procedure was finally used to calibrate the top quark masses of the major Monte Carlo event generators on the market, PYTHIA, HERWIG and SHERPA, to the MSR and pole mass using 3 observables and 3 soft gap schemes. The results are plotted in Fig. 7.8, 7.9, and 7.10.

Both observable and soft gap dependencies are well controlled due to the improvements that we have implemented. We have shown that the differences between the Monte Carlo mass schemes using their standard tunes are small and within  $\sim 200$  MeV for the central values and  $\sim 500$  MeV using the full uncertainty ranges.

The non-perturbative soft behaviour, which we successfully separated from the mass dependence, was in contrast vastly different between MC generators for the type of observables that we used. These differences are reflected in a clear incompatibility of the distributions produced by these MC generators and they are quantified by the soft

shape function parameters that we simultaneously fitted together with the mass parameter.

This calibration framework can be used to confirm analytical results for the MC mass scheme dependence, if they are available (e.g. HERWIG [11]). It can also be used to troubleshoot and improve Monte Carlo event generators and their models for this kinematic phase space region.

Our theoretical description can be improved by including the known N<sup>3</sup>LL bHQET ingredients. Ref. [14] has already tested the convergence behaviour of this additional order for bHQET alone, but the nonsingular contribution has to be still worked out for a complete description.

Further theoretical improvements can be achieved by investigating the leading  $(m/Q)^2$  power corrections in a strict EFT framework to exactly determine the correct resummation contributions for these power corrections. We know that compared to the strict leading power bHQET description, these corrections will be large. If this strictly derived theory result falls within the reasonable estimates of our power correction improved description, then we expect a relatively small improvement of our MSR mass fit results and a significant reduction of our pole mass fit N<sup>2</sup>LL uncertainties, since those were rather sensitive to our uncertainty estimate parametrization.

# A. Evolutions

## A.1. Evolution Factors and Anomalous Dimensions

We follow the notation and convention of Ref. [14], except our  $\Gamma^{\text{cusp}}$  has  $C_F$  absorbed. The RGE evolution factors read [13]

$$\begin{aligned}
 U_{H_Q}(Q, \mu_0, \mu_1) &= e^{K_{H_Q}} \left( \frac{\mu_0}{Q} \right)^{\omega_{H_Q}} \\
 U_v(\rho, \mu_1, \mu_0) &= e^{K_v} \rho^{-\omega_v} \\
 U_F(t, \mu_1, \mu_0) &= \frac{e^{K_F} (e^{\gamma_E})^{\omega_F}}{\Gamma(-\omega_F)} \mathcal{L}_{0, \omega_F}^{\mu_0}(t)
 \end{aligned}
 \tag{A.1}$$

where  $F \in \{B_\tau, S_\tau\}$  and  $\mathcal{L}_{0, \omega_F}^\mu$  is the fractional plus distribution defined in App. B. They are the solutions to the renormalization group equations

$$\begin{aligned}
 \mu \frac{d}{d\mu} H_Q(Q, \mu) &= \left[ \Gamma_H[\alpha_s] \ln \left( \frac{\mu}{Q} \right) + \gamma_{H_Q}[\alpha_s] \right] H_Q(Q, \mu) \\
 \mu \frac{d}{d\mu} \mathcal{J}_\nu(\tau, \mu) &= (\Gamma_v[\alpha_s] \ln \rho^{-1} + \gamma_v[\alpha_s]) \mathcal{J}_\nu(\tau, \mu) \\
 \mu \frac{d}{d\mu} F(t, \mu) &= \int_{-\infty}^{+\infty} dt' \gamma_F(t-t', \mu) F(t', \mu) \\
 \gamma_F(t-t', \mu) &= -\Gamma_F[\alpha_s] \mathcal{L}_{0, \omega_F}^{\mu_F}(t-t') + \gamma_F[\alpha_s] \delta(t-t')
 \end{aligned}
 \tag{A.2}$$

with  $\mathcal{J}_\nu$  defined as the squared bHQET current  $B_\tau \otimes S_\tau$ , the running of this combination is not a convolution anymore.  $\Gamma_F$  is the cusp anomalous dimension and  $\gamma_F$  is the non-cusp anomalous dimension. Note, that the implementation of power corrections in the measurement function discussed in Chap. 6 rescales the boost between the soft and ultra-collinear momenta, as a consequence  $\rho$  has to be replaced by the exact on-shell top quark boost factor  $r_s/(r_c \hat{m})$  to obtain consistent running, that is independent of the arbitrary starting scale of  $U_\nu$ .

The evolution kernels are given by  $[\alpha_i \equiv \alpha_s(\mu_i)]$ :

$$\begin{aligned}
 \omega(\Gamma; \mu_1, \mu_0) &= \int_{\alpha_0}^{\alpha_1} \frac{d\alpha}{\beta(\alpha)} \Gamma[\alpha] \\
 K(\Gamma, \gamma, j; \mu_1, \mu_0) - \omega(\gamma; \mu_1, \mu_0) &= j \int_{\alpha_0}^{\alpha_1} \frac{d\alpha}{\beta(\alpha)} \Gamma[\alpha] \int_{\alpha_0}^{\alpha} \frac{d\alpha'}{\beta(\alpha')}
 \end{aligned}
 \tag{A.3}$$

where  $j$  is the mass dimension of the variable in the logarithm of the cusp piece. Given our notation in Eq. A.2 all  $j = 1$ , except for  $\mathcal{J}_\nu$  for which  $j = 0$ . The results at N<sup>2</sup>LL

read [30]

$$\begin{aligned}
\omega^{\text{N}^3\text{LL}}(\Gamma; \mu_1, \mu_0) &= -\frac{\Gamma_0}{2\beta_0} \left\{ \log r + \frac{\alpha_0}{4\pi} \left( \frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) (r-1) \right. \\
&+ \frac{1}{2} \left( \frac{\alpha_0}{4\pi} \right)^2 \left( \frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} + \frac{\Gamma_2}{\Gamma_0} - \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} \right) (r^2-1) \\
&\left. + \frac{1}{3} \left( \frac{\alpha_0}{4\pi} \right)^3 \left[ \frac{\Gamma_3}{\Gamma_0} - \frac{\beta_3}{\beta_0} + \frac{\Gamma_1}{\Gamma_0} \left( \frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} \right) - \frac{\beta_1}{\beta_0} \left( \frac{\beta_1^2}{\beta_0^2} - 2\frac{\beta_2}{\beta_0} + \frac{\Gamma_2}{\Gamma_0} \right) \right] (r^3-1) \right\}
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
K^{\text{N}^2\text{LL}}(\Gamma, \gamma, j; \mu_1, \mu_0) &= \frac{j\Gamma_0}{4\beta_0^2} \left\{ \frac{4\pi}{r\alpha_0} (r \log r + 1 - r) + \left( \frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right) (r-1 - \log r) \right. \\
&- \frac{\beta_1}{2\beta_0} \log^2 r + \frac{\alpha_0}{4\pi} \left[ \left( \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} - \frac{\beta_1^2}{\beta_0^2} \right) (r-1 - r \log r) - B_2 \log r \right. \\
&\left. \left. + \left( \frac{\Gamma_2}{\Gamma_0} - \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} + B_2 \right) \frac{r^2-1}{2} + \left( \frac{\Gamma_1\beta_1}{\Gamma_0\beta_0} - \frac{\Gamma_2}{\Gamma_0} \right) (r-1) \right] \right\} + \omega^{\text{NLL}}(\gamma; \mu_1, \mu_0),
\end{aligned} \tag{A.5}$$

where  $r = \alpha_1/\alpha_0$  depends on the 4-loop running coupling and  $B_2 = \beta_1^2/\beta_0^2 - \beta_2/\beta_0$ . The QCD beta function and the cusp and non-cusp anomalous dimensions are given by the series

$$\frac{d\alpha_s(\mu)}{d \log \mu} = \beta[\alpha_s] = -2\alpha_s \sum_{n=0}^{\infty} \beta_n \left( \frac{\alpha_s}{4\pi} \right)^{n+1}, \quad \Gamma[\alpha_s] = \sum_{n=0}^{\infty} \Gamma_n \left( \frac{\alpha_s}{4\pi} \right)^{n+1}, \tag{A.6}$$

where  $\Gamma$  stands for either  $\Gamma_F, \gamma_F$  or the QCD cusp anomalous dimension  $\Gamma^{\text{cusp}}$ . The cusp anomalous dimensions are proportional to  $\Gamma^{\text{cusp}}$  [13, 18, 39–42]

$$\begin{aligned}
\Gamma_{B_\tau}[\alpha_s] &= \Gamma_\nu[\alpha_s] = -\Gamma_{S_\tau}[\alpha_s] = 4\Gamma^{\text{cusp}(5)}[\alpha_s] \\
&- \Gamma_{HQ}[\alpha_s] = 4\Gamma^{\text{cusp}(6)}[\alpha_s]
\end{aligned} \tag{A.7}$$

with the universal cusp anomalous dimension coefficients given by

$$\begin{aligned}
\Gamma_0^{\text{cusp}} &= \frac{16}{3}, \quad \Gamma_1^{\text{cusp}} = \frac{1072}{9} - \frac{16}{3}\pi^2 - \frac{160}{27}n_f, \\
\Gamma_2^{\text{cusp}} &= 1960 - \frac{2144}{9}\pi^2 + \frac{176}{15}\pi^4 + 352\zeta_3 + \left( -\frac{5104}{27} + \frac{320}{27}\pi^2 - \frac{832}{9}\zeta_3 \right) n_f - \frac{64}{81}n_f^2.
\end{aligned} \tag{A.8}$$

Consistency in the running gives the relation

$$\gamma_\nu[\alpha_s] = \gamma_{B_\tau}[\alpha_s] + \gamma_{S_\tau}[\alpha_s] \tag{A.9}$$

and the expressions for the noncusp anomalous dimensions read

$$\begin{aligned}
\gamma_0^{HQ} &= -16, \quad \gamma_1^{HQ} = -\frac{7976}{27} - \frac{136}{9}\pi^2 + \frac{736}{3}\zeta_3 + \left( \frac{1040}{81} + \frac{16}{9}\pi^2 \right) n_f, \\
\gamma_0^{B_\tau} &= \frac{32}{3}, \quad \gamma_1^{B_\tau} = \frac{11168}{27} - \frac{184}{9}\pi^2 - 160\zeta_3 + \left( -\frac{1856}{81} + \frac{16}{27}\pi^2 \right) n_f, \\
\gamma_0^{S_\tau} &= 0, \quad \gamma_1^{S_\tau} = -\frac{6464}{27} - \frac{88}{9}\pi^2 + 224\zeta_3 + \left( \frac{896}{81} - \frac{16}{27}\pi^2 \right) n_f.
\end{aligned} \tag{A.10}$$

The beta function coefficients are given by [43]

$$\begin{aligned}\beta_0 &= 11 - \frac{2}{3}n_f, & \beta_1 &= 102 - \frac{38}{3}n_f, & \beta_2 &= \frac{2857}{2} - \frac{5033}{18}n_f + \frac{325}{54}n_f^2, \\ \beta_3 &= \frac{149753}{6} + 3564\zeta_3 + \left(-\frac{1078361}{162} - \frac{6508}{27}\zeta_3\right)n_f + \left(\frac{50065}{162} + \frac{6472}{81}\zeta_3\right)n_f^2 + \frac{1093}{729}n_f^3.\end{aligned}\tag{A.11}$$

## A.2. R-Evolution

The MSR mass and soft gap scheme R-RGE can be determined from the fact that  $m^{\text{pole}} = m^{\text{MSR}}(R) + \delta m(R)$  and  $\Delta = \bar{\Delta}(R, R) + \bar{\delta}(R, R)$  are scale independent. Given a perturbative series of the form

$$f(R) = \text{const.} - R \sum_{n=1}^{\infty} \left[ \frac{\alpha_s(R)}{4\pi} \right]^n f_n \tag{A.12}$$

the R-RGE can be written as

$$\frac{df(R)}{d \ln R} = R \frac{df(R)}{dR} = -R \sum_{n=0}^{\infty} \gamma_n^{f,R} \left[ \frac{\alpha_s(R)}{4\pi} \right]^{n+1} \tag{A.13}$$

with the anomalous dimension coefficients reading

$$\begin{aligned}\gamma_n^{f,R} &= f_{n+1} - 2 \sum_{j=0}^{n-1} (n-j)\beta_j f_{n-j}, & (n \geq 1) \\ \gamma_0^{f,R} &= f_1\end{aligned}\tag{A.14}$$

The solution for the evolution is therefore

$$f(R_1) - f(R_0) = - \sum_{n=0}^{\infty} \gamma_n^{f,R} \int_{R_0}^{R_1} dR \left[ \frac{\alpha_s(R)}{4\pi} \right]^{n+1}. \tag{A.15}$$

For  $f(R) = m^{\text{MSRp}}(R)$  we have  $f_n = a_n^{\text{MSRp}}(n_l)$  as defined in Eq. (2.17) and given by [27]

$$\begin{aligned}a_1^{\text{MSRp}}(n_l) &= \frac{16}{3}, & a_2^{\text{MSRp}}(n_l) &= \frac{307}{2} + \frac{16}{3}\pi^2 + \frac{16}{9}\pi^2 \log 2 - \frac{8}{3}\zeta_3 + n_l \left( -\frac{71}{9} - \frac{8}{9}\pi^2 \right), \\ a_3^{\text{MSRp}}(n_l) &= 12185. - 1705.93n_l + 41.7722n_l^2.\end{aligned}\tag{A.16}$$

For  $f(R) = \bar{\Delta}(R, R)$  we use  $f_n = d_i(R, R)$  as defined in Eq. (2.18). These in turn depend on the coefficients in the exponent of the position space soft function as defined in Eq. 2.22, which can be generated by [14]

$$s_{mn} = s_{mn}^{[0]}[\beta] + s_{mn}^{[1]}[\beta] + s_{mn}^{[2]}[\beta], \tag{A.17}$$

where each term follows a recursion relation for  $m > 1$  and  $1 \leq n - k \leq m - 1$

$$s_{mn}^{[k]}[\beta] = \frac{2}{n} \sum_{i=n-k}^{m-1} i s_{i(n-1)}[\beta] \beta_{m-i-1}, \quad (\text{A.18})$$

and the independent starting values (with  $m \geq 1$ ) read

$$s_{m0}^{[0]}[\beta] = s_{m0}, \quad s_{m1}^{[1]}[\beta] = \gamma_{m-1}^{S_\tau}, \quad s_{m2}^{[2]}[\beta] = \frac{1}{2} \Gamma_{m-1}^{S_\tau}. \quad (\text{A.19})$$

The anomalous dimensions are listed in the previous section and the relevant non-logarithmic terms are given by [13, 44]

$$s_{10} = -\frac{4\pi^2}{3} \quad s_{20} = \frac{2}{405} (25740\zeta_3 + 756\pi^4 - 13065\pi^2 - 21400) + \frac{4}{243} (-468\zeta_3 + 231\pi^2 + 40) n_f. \quad (\text{A.20})$$

Gap 2 and 3 are  $\mu$  independent, but gap 1 inherits a non-trivial  $\mu$  anomalous dimension from the soft function and hence requires an additional  $\mu$  evolution. This  $\mu$ -RGE reads [29]

$$\mu \frac{d}{d\mu} \bar{\Delta}_1(R, \mu) = -\mu \frac{d}{d\mu} \bar{\delta}_1(R, \mu) = 2Re^{\gamma_E} \Gamma^{\text{cusp}}[\alpha_s], \quad (\text{A.21})$$

which follows from the gap definition 2.23 and the soft function RGE in position space, and where the cusp anomalous dimension are given in Eq. (A.8). The solution is

$$\bar{\Delta}_1(R, \mu) - \bar{\Delta}_1(R, \mu_0) = 2Re^{\gamma_E} \omega(\Gamma^{\text{cusp}}, \mu, \mu_0), \quad (\text{A.22})$$

with the evolution kernel as defined in Eq. (A.3).

## B. Distributions

The plus function with a fractional exponent  $1 + \omega$  and  $\omega < 1$  is defined by [13]

$$\left[ \frac{\Theta(x)}{(x)^{1+\omega}} \right]_+ \equiv \lim_{\beta \rightarrow 0} \left[ \frac{\theta(x - \beta)}{(x)^{1+\omega}} - \delta(x - \beta) \frac{\beta^{-\omega}}{\omega} \right]. \quad (\text{B.1})$$

Expanding this equation for small  $\omega$  gives the definition for the  $\log^n(x)/x$  plus distribution for  $n \geq 0$ :

$$\left[ \frac{\Theta(x) \log^n x}{x} \right]_+ \equiv \lim_{\beta \rightarrow 0} \left[ \frac{\theta(x - \beta) \log^n x}{x} - \delta(x - \beta) \frac{\log^{n+1} x}{n+1} \right]. \quad (\text{B.2})$$

Integrating with this distribution over a test function  $f(x)$  gives

$$\int_0^\Delta dx \left[ \frac{\Theta(x) \log^n x}{x} \right]_+ f(x) = \int_0^\Delta dx \frac{f(x) - f(0)}{x} \log^n x + f(0) \frac{\log^{n+1} \Delta}{n+1} \quad (\text{B.3})$$

These plus distributions appear in the jet and soft function and their evolutions. We will use a shorthand notation for them:

$$\begin{aligned} \mathcal{L}_{0,\omega}^\mu(\ell) &\equiv \frac{1}{\mu^j} \left[ \frac{\Theta(\ell)}{(\ell/\mu^j)^{1+\omega}} \right]_+ \\ \mathcal{L}_n^\mu(\ell) &\equiv \frac{1}{\mu^j} \left[ \frac{\Theta(\ell) \log^n(\ell/\mu^j)}{\ell/\mu^j} \right]_+ \end{aligned} \quad (\text{B.4})$$

where the exponent  $j$  is the mass dimension of the variable  $\ell$ . In the case of an dimensionless argument we will also use the notation

$$\left[ \frac{1}{e} \right]_+ \equiv \left[ \frac{\Theta(e)}{e} \right]_+ \quad (\text{B.5})$$

An useful identity to rescale the plus function argument for e.g. matching purposes is given by [13]

$$\kappa \left[ \frac{\theta(x) \log^n(\kappa x)}{\kappa x} \right]_+ = \frac{\log^{n+1}(\kappa)}{n+1} \delta(x) + \sum_{k=0}^n \frac{n!}{(n-k)!k!} \log^{n-k}(\kappa) \left[ \frac{\theta(x) \log^k(x)}{x} \right]_+. \quad (\text{B.6})$$

## C. Additional Analysis

### C.1. Comparison: $R_s < \mu_S$ and $R_s = \mu_S$

Fig. C.1 shows how modifying the analysis for the final result for PYTHIA shown in Fig. 7.5 by setting the soft gap subtraction scale  $R_s$  to the soft scale  $\mu_S$  affects the fitted results. The default profiles set  $R_s$  below  $\mu_S$ , see Eq. (4.9). The fitted mass values are not affected by the different choices of  $R_s$ . The fitted  $\Omega_1$  values of gap scheme 2 and 3 show insignificant differences. The fitted  $\Omega_1$  values for gap scheme 1 received large shifts which was expected. The gap parameter and also  $\Omega_1(R_s, \mu_S)$ , which is a function of the gap parameter, are extracted from the cross section which uses a dynamical scale  $\mu_S(\tau)$ , that is higher than the reference scales, which are used to quote the result for the fitted  $\Omega_1(R_\Delta = 2 \text{ GeV}, \mu_\Delta = 2 \text{ GeV})$  values. For  $R_s = \mu_S$  the scheme 1 gap parameter  $\bar{\Delta}(R_s, R_s)$  does not run with the scale  $R_s$  at NLL, since the first R-anomalous dimension is zero. When  $R_s$  is below  $\mu_S$ , however, it is possible for the scheme 1 gap parameter to run down to the reference scales. This explains the difference at NLL. For  $R_s = \mu_S$  at N<sup>2</sup>LL, the scheme 1 gap parameter is able run with the two-loop R-anomalous dimension, but the gap cannot subtract  $\mathcal{O}(\alpha_s)$  corrections, since the non-log one loop term in the subtraction series is zero. When  $R_s < \mu_S$  at N<sup>2</sup>LL, the finite logarithmic term of the series in Eq. (2.24) can be used as an effective subtraction term. These factors contribute to the differences at N<sup>2</sup>LL.

The consistency of the fitted mass values for all gap schemes indicates, that our calibration method correctly absorbs soft gap related differences into the fitted  $\Omega_1$  values.

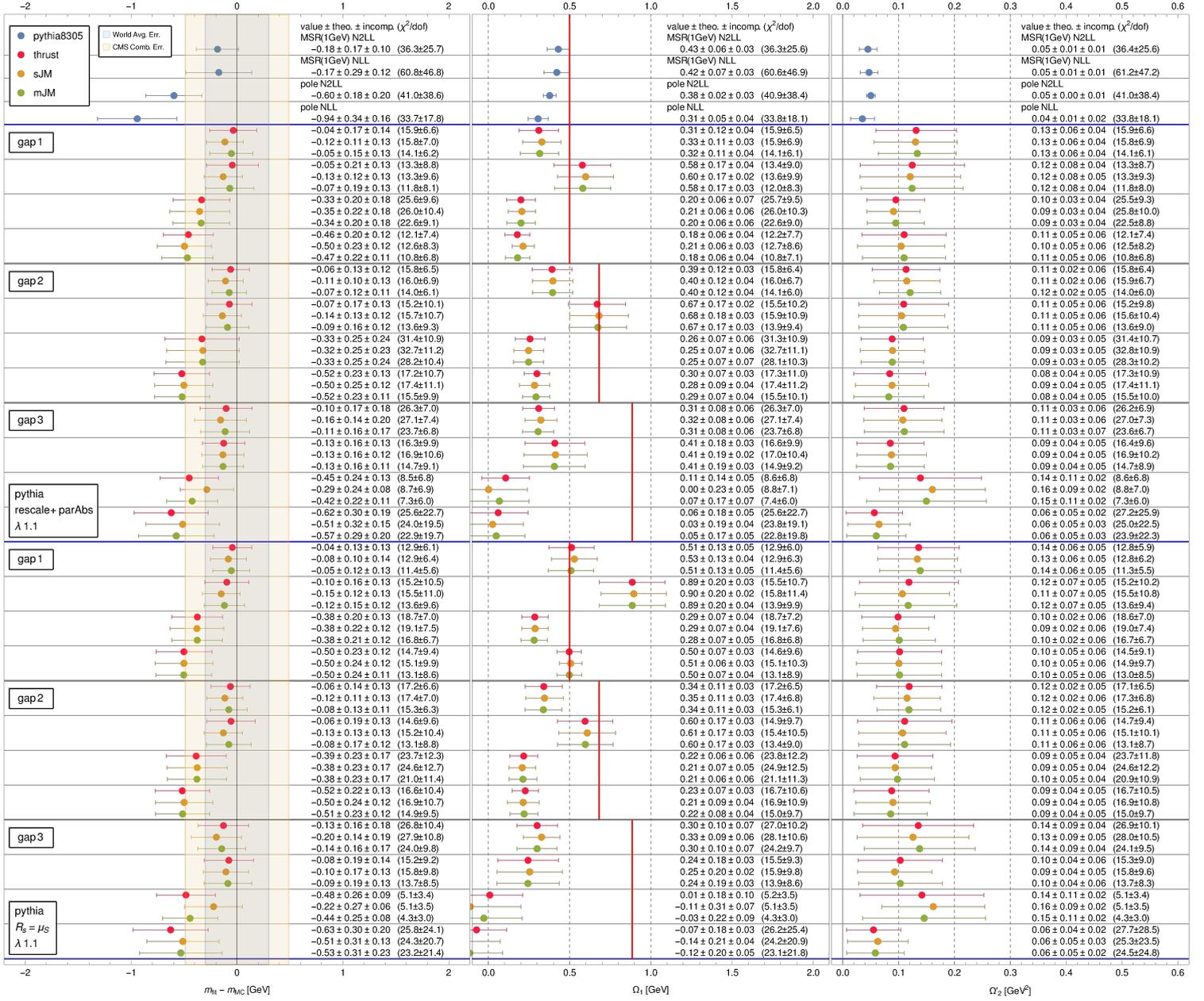


Figure C.1.:  $R_s < \mu_S$  vs.  $R_s = \mu_S$  comparison for PYTHIA. The upper half, with  $R_s < \mu_S$ , is identical to Fig. 7.5. The lower half sets  $R_s = \mu_S$ .

# Bibliography

- [1] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.
- [2] CMS collaboration, *Measurement of the top quark mass using proton-proton data at  $\sqrt{s} = 7$  and 8 TeV*, *Phys. Rev. D* **93** (2016) 072004 [1509.04044].
- [3] ATLAS collaboration, *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  lepton+jets channel from  $\sqrt{s} = 8$  TeV ATLAS data and combination with previous results*, *Eur. Phys. J. C* **79** (2019) 290 [1810.01772].
- [4] CDF, D0 collaboration, *Combination of CDF and D0 results on the mass of the top quark using up  $9.7 \text{ fb}^{-1}$  at the Tevatron*, 1608.01881.
- [5] CMS collaboration, *A profile likelihood approach to measure the top quark mass in the lepton+jets channel at  $\sqrt{s} = 13$  TeV*, .
- [6] P. Azzi et al., *Report from Working Group 1: Standard Model Physics at the HL-LHC and HE-LHC*, *CERN Yellow Rep. Monogr.* **7** (2019) 1 [1902.04070].
- [7] M. Beneke, P. Marquard, P. Nason and M. Steinhauser, *On the ultimate uncertainty of the top quark pole mass*, *Phys. Lett. B* **775** (2017) 63 [1605.03609].
- [8] A.H. Hoang, C. Lepenik and M. Preisser, *On the Light Massive Flavor Dependence of the Large Order Asymptotic Behavior and the Ambiguity of the Pole Mass*, *JHEP* **09** (2017) 099 [1706.08526].
- [9] A.H. Hoang and I.W. Stewart, *Top Mass Measurements from Jets and the Tevatron Top-Quark Mass*, *Nucl. Phys. B Proc. Suppl.* **185** (2008) 220 [0808.0222].
- [10] A.H. Hoang, *The Top Mass: Interpretation and Theoretical Uncertainties*, in *7th International Workshop on Top Quark Physics*, 12, 2014 [1412.3649].
- [11] A.H. Hoang, S. Plätzer and D. Samitz, *On the Cutoff Dependence of the Quark Mass Parameter in Angular Ordered Parton Showers*, *JHEP* **10** (2018) 200 [1807.06617].
- [12] S. Fleming, A.H. Hoang, S. Mantry and I.W. Stewart, *Jets from massive unstable particles: Top-mass determination*, *Phys. Rev. D* **77** (2008) 074010 [hep-ph/0703207].
- [13] S. Fleming, A.H. Hoang, S. Mantry and I.W. Stewart, *Top Jets in the Peak Region: Factorization Analysis with NLL Resummation*, *Phys. Rev. D* **77** (2008) 114003 [0711.2079].

- [14] B. Bachu, A.H. Hoang, V. Mateu, A. Pathak and I.W. Stewart, *Boosted top quarks in the peak region with  $NL3L$  resummation*, *Phys. Rev. D* **104** (2021) 014026 [2012.12304].
- [15] C. Lepenik and V. Mateu, *NLO Massive Event-Shape Differential and Cumulative Distributions*, *JHEP* **03** (2020) 024 [1912.08211].
- [16] A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, *Extracting a Short Distance Top Mass with Light Grooming*, *Phys. Rev. D* **100** (2019) 074021 [1708.02586].
- [17] M. Butenschoen, B. Dehnadi, A.H. Hoang, V. Mateu, M. Preisser and I.W. Stewart, *Top Quark Mass Calibration for Monte Carlo Event Generators*, *Phys. Rev. Lett.* **117** (2016) 232001 [1608.01318].
- [18] A.H. Hoang, A. Pathak, P. Pietrulewicz and I.W. Stewart, *Hard Matching for Boosted Tops at Two Loops*, *JHEP* **12** (2015) 059 [1508.04137].
- [19] A.H. Hoang and I.W. Stewart, *Designing gapped soft functions for jet production*, *Phys. Lett. B* **660** (2008) 483 [0709.3519].
- [20] Z. Ligeti, I.W. Stewart and F.J. Tackmann, *Treating the  $b$  quark distribution function with reliable uncertainties*, *Phys. Rev. D* **78** (2008) 114014 [0807.1926].
- [21] R. Tarrach, *The Pole Mass in Perturbative QCD*, *Nucl. Phys. B* **183** (1981) 384.
- [22] N. Gray, D.J. Broadhurst, W. Grafe and K. Schilcher, *Three Loop Relation of Quark (Modified)  $M_s$  and Pole Masses*, *Z. Phys. C* **48** (1990) 673.
- [23] K. Melnikov and T.v. Ritbergen, *The Three loop relation between the  $\overline{MS}$ -bar and the pole quark masses*, *Phys. Lett. B* **482** (2000) 99 [hep-ph/9912391].
- [24] K.G. Chetyrkin and M. Steinhauser, *Short distance mass of a heavy quark at order  $\alpha_s^3$* , *Phys. Rev. Lett.* **83** (1999) 4001 [hep-ph/9907509].
- [25] K.G. Chetyrkin and M. Steinhauser, *The Relation between the  $\overline{MS}$ -bar and the on-shell quark mass at order  $\alpha(s)^{**3}$* , *Nucl. Phys. B* **573** (2000) 617 [hep-ph/9911434].
- [26] P. Marquard, L. Mihaila, J.H. Piclum and M. Steinhauser, *Relation between the pole and the minimally subtracted mass in dimensional regularization and dimensional reduction to three-loop order*, *Nucl. Phys. B* **773** (2007) 1 [hep-ph/0702185].
- [27] A.H. Hoang, A. Jain, C. Lepenik, V. Mateu, M. Preisser, I. Scimemi et al., *The  $MSR$  mass and the  $\mathcal{O}(\Lambda_{QCD})$  renormalon sum rule*, *JHEP* **04** (2018) 003 [1704.01580].
- [28] A.H. Hoang and S. Kluth, *Hemisphere Soft Function at  $O(\alpha(s)^{**2})$  for Dijet Production in  $e^+ e^-$  Annihilation*, 0806.3852.
- [29] B. Dehnadi, *Heavy quark mass determinations with sum rules and jets*, Ph.D. thesis, Vienna U., 2016. 10.25365/thesis.42936.

- [30] 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  $\alpha_s(m_Z)$* , *Phys. Rev. D* **83** (2011) 074021 [1006.3080].
- [31] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, 2203.11601.
- [32] SHERPA collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].
- [33] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [1512.01178].
- [34] S. Plätzer and D. Samitz. Personal communication.
- [35] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, *SciPost Phys.* **8** (2020) 026 [1912.05451].
- [36] A. Buckley, P. Ilten, D. Konstantinov, L. Lönnblad, J. Monk, W. Pokorski et al., *The HepMC3 event record library for Monte Carlo event generators*, *Comput. Phys. Commun.* **260** (2021) 107310 [1912.08005].
- [37] V. Mateu. Personal communication.
- [38] T.G. Kolda, R.M. Lewis and V. Torczon, *Optimization by direct search: New perspectives on some classical and modern methods*, *SIAM Review* **45** (2003) 385 [<https://doi.org/10.1137/S003614450242889>].
- [39] S. Moch, J.A.M. Vermaseren and A. Vogt, *The Three loop splitting functions in QCD: The Nonsinglet case*, *Nucl. Phys. B* **688** (2004) 101 [hep-ph/0403192].
- [40] S. Moch, J.A.M. Vermaseren and A. Vogt, *The Quark form-factor at higher orders*, *JHEP* **08** (2005) 049 [hep-ph/0507039].
- [41] A. Jain, I. Scimemi and I.W. Stewart, *Two-loop Jet-Function and Jet-Mass for Top Quarks*, *Phys. Rev. D* **77** (2008) 094008 [0801.0743].
- [42] T. Becher, M. Neubert and B.D. Pecjak, *Factorization and Momentum-Space Resummation in Deep-Inelastic Scattering*, *JHEP* **01** (2007) 076 [hep-ph/0607228].
- [43] K.G. Chetyrkin, J.H. Kuhn and M. Steinhauser, *RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses*, *Comput. Phys. Commun.* **133** (2000) 43 [hep-ph/0004189].
- [44] P.F. Monni, T. Gehrmann and G. Luisoni, *Two-Loop Soft Corrections and Resummation of the Thrust Distribution in the Dijet Region*, *JHEP* **08** (2011) 010 [1105.4560].