The scale dependence of perturbative QCD predictions for e^+e^- event shape moments and LHC observables

Katharine Elizabeth Morgan

The Institute for Particle Physics Phenomenology,
Department of Physics

A thesis submitted to Durham University
for the degree of Doctor of Philosophy

September 2012

Abstract

Perturbative QCD predictions that are truncated at fixed order have an unphysical dependence on the renormalisation procedure. We investigate two methods of avoiding scale and scheme dependence in QCD predictions of physical observables: the Effective Charges (ECH) method and the Principle of Minimal Sensitivity (PMS).

The ECH method is used to avoid the renormalisation scale and scheme dependence of fixed-order predictions of event shape moments. Values of $\alpha_s(M_Z)$ are extracted from e^+e^- data using both ECH and the physical scale method in the $\overline{\text{MS}}$ scheme. The ECH method at NLO is found to perform better than standard $\overline{\text{MS}}$ perturbation theory ($\overline{\text{MS}}$ PT) when applied to means of event shapes. However ECH at NNLO functions less well than at NLO, and the ECH method also fails to describe data for higher moments of event shapes. Padé Approximant methods are used to estimate missing higher orders in the perturbative expansions, a technique that works especially well for $\overline{\text{MS}}$ PT applied to the higher moments. We also examine the effect of adding non-perturbative power corrections to the perturbative approximations. It is found that power corrections are insufficient to counteract the undesirable behaviour of ECH at NNLO.

The PMS method is used to provide predictions of the $b\bar{b}$ and $t\bar{t}$ total cross-sections at the Tevatron and the LHC. Hadronic cross-sections depend on the factorisation scale as well as the renormalisation scale. PMS is applied by searching for stationary points on the cross-section surface in the space of the two scales. The PMS method predicts substantially larger $b\bar{b}$ cross-sections than using standard diagonal scale choices. For $t\bar{t}$ production, however, there is very little difference observed between the two methods. Both produce predictions that are in good agreement with the current experimental data.

Declaration

I declare that this thesis has not been submitted for another degree at this or any other university.

The work contained in this thesis was carried out in collaboration with Dr. C. J. Maxwell. The material in Chapter 3 is published in Ref. [1], except for that in Section 3.4 which is in preparation. The contents of Chapter 5 are also in preparation for publication.

This PhD was funded by an STFC studentship.

Acknowledgements

Firstly, I would like to thank my supervisor, Chris Maxwell, for his help and support over the course of my PhD.

I would also like to thank Peter Richardson and David Grellscheid for their assistance in using HERWIG++. My thanks go to Sebastian Sapeta and Andrew Papanastasiou for useful discussions regarding MCFM and top physics.

More generally, I thank the students and staff of the IPPP for providing a friendly and welcoming working environment.

Especial thanks go to my friends and family — for making my four years in Durham great fun, and also for supporting me through the writing up process.

Contents

Preface

The quantum field theory of Quantum Chromodynamics (QCD) has made many successful predictions regarding the behaviour of the strong force [2, 3]. QCD is a renormalisable SU(3) gauge theory which describes colour interactions between quarks and gluons. The non-Abelian nature of QCD results in the gauge bosons of the theory carrying colour, leading to the effects of asymptotic freedom and confinement [2, 4]. Since QCD possesses the property of asymptotic freedom, the strength of the coupling interaction decreases at large energies. This allows perturbation theory, with an expansion in the coupling, to be applied at sufficiently high energies.

When QCD is renormalised, an unphysical scale, $\mu_{\rm R}$, is introduced. Renormalised predictions of physical observables have an unphysical dependence on this renormalisation scale, and on the renormalisation scheme (RS) used, when the perturbative expansion is truncated. The standard method of accounting for this is to work in the $\overline{\rm MS}$ renormalisation scheme and to set $\mu_{\rm R}$ equal to some physical scale related to the observable, such as the centre-of-mass energy of e^+e^- collisions [3, 5]. The renormalisation scale is then varied by a factor of two around this choice to provide an estimate of the theoretical uncertainty [3,5,6]. The choices of $\mu_{\rm R}$ and the range of its variation are arbitrary and can lead to theoretical uncertainties that are greater than the experimental uncertainties involved. In this thesis we examine two different ways of

avoiding or reducing this unphysical scale dependence: the Effective Charges (ECH) method and the Principle of Minimal Sensitivity (PMS).

The ECH technique [7–15] starts from a β-function type equation, analagous to that for the strong coupling constant, which consists of an expansion in terms of the observable itself. The perturbative coefficients in this expansion are renormalisation scheme invariant and so predictions obtained from this equation are also scheme-independent. The β-function equation for the observable is integrated, resulting in an integration constant that is related to $\Lambda_{\rm QCD}$, the characteristic scale of QCD [2, 3, 5]. Once $\Lambda_{\rm QCD}$ is determined $\alpha_s(M_Z)$ can be extracted using the QCD β-function equation.

The PMS approach [16] is to choose the renormalisation scheme parameters such that the truncated fixed-order prediction of an observable is locally invariant to the RS. This invariance reflects the behaviour of the full all-orders result which is independent of the RS. PMS can be easily extended to cross-sections at hadronic colliders, which require parton distribution functions (PDFs) to describe the structure of the incoming hadrons. Cross-sections of hadronic collisions can be factorised into a PDF part and a partonic cross-section, and a factorisation scale, M, is introduced to define the split between the two parts. In this case, the PMS method is applied by looking for a stationary point on the cross-section surface in M and $\mu_{\rm R}$ space.

In this thesis we begin by briefly introducing QCD, renormalisation and scheme dependence in Chapter 1. The QCD β-function and its solutions are also outlined.

In Chapters 2 and 3 we derive the Effective Charges method and use it to extract $\alpha_s(M_Z)$ from e^+e^- event shape moment data. The data used come from LEP (DELPHI, OPAL and L3) and PETRA (JADE) e^+e^- collisions [17–22]. ECH is also used to provide predictions to the data. In both cases ECH is

compared with the standard scale choice in $\overline{\rm MS}$ perturbation theory ($\overline{\rm MS}$ PT). Padé Approximant methods are used to provide an estimate of missing higher orders in the perturbative expansions in Section 3.4. The effect of adding non-perturbative power corrections to ECH and $\overline{\rm MS}$ PT is also considered in Section 3.5. The PMS predictions of the event shape moments are examined in Chapter 4.

We move on, in Chapter 5, to consider heavy quark cross-sections at hadronic colliders. PMS predictions for the $b\bar{b}$ total cross-section at the Tevatron are given in Section 5.2. The $t\bar{t}$ total cross-section is also analysed, using PMS, at both the Tevatron and the LHC.

Finally, our conclusions are given in Chapter 6.

Chapter 1

An Introduction to QCD

Quantum Chromodynamics (QCD) is the quantum field theory that describes the strong interactions between quarks. It is a non-Abelian SU(3) gauge theory. The gauge bosons of QCD are gluons which, because of the non-Abelian nature of the theory, are self-interacting.

In the next section we briefly introduce the QCD Lagrangian. We then go on to discuss renormalisation in Section 1.2 and renormalisation scheme dependence in Section 1.3.

The sections in this chapter follow the material in Refs. [2, 4, 5] closely. Ref. [23] also provides a useful introduction to the subject, and Ref. [24] gives a comprehensive mathematical discussion of field theory.

1.1 QCD Lagrangian

The Lagrangian of QCD can be separated into three parts:

$$\mathcal{L}_{\rm QCD} = \mathcal{L}_{\rm class.} + \mathcal{L}_{\rm gauge} + \mathcal{L}_{\rm ghost}. \tag{1.1.1}$$

$\mathcal{L}_{\text{class.}}$, the classical Lagrangian density, contains the kinetic terms for the fields, and the mass term for the quarks:

$$\mathcal{L}_{\text{class.}} = -\frac{1}{4} F^A_{\alpha\beta} F_A^{\alpha\beta} + \sum_{a,b=1}^{N_f} \bar{q}_a (i \not{D} - m)_{ab} q_b, \tag{1.1.2}$$

where the indices (a, b) run over the quark flavours, (A, B) run over the gluon colours, and (α, β) are space-time indices. $\not{D} = D_\mu \gamma^\mu$ is the covariant derivative dotted with the gamma matrices, which have the anti-commutation relation $\{\gamma^\alpha, \gamma^\beta\} = 2g^{\alpha\beta}$. The covariant derivative is given by:

$$\begin{aligned} (D_\alpha)_{ab} &= \partial_\alpha \delta_{ab} + i g_s (t^C A^C_\alpha)_{ab} \,, \\ (D_\alpha)_{AB} &= \partial_\alpha \delta_{ab} + i g_s (T^C A^C_\alpha)_{AB} \,, \end{aligned} \tag{1.1.3}$$

where the former definition acts on the quark fields, using the t^A generators of SU(3) in the fundamental representation, and the latter on the gluon fields, with the T^A generators in the adjoint representation. The generators have the following commutation relations:

$$[t^A, t^B] = i f^{ABC} t^C \,, \tag{1.1.4}$$

where f^{ABC} are the structure constants of SU(3). The T^A generators follow the same Lie algebra, and are given by $(T^A)_{BC} = -i f^{ABC}$. By convention the generators are normalised by:

$$\text{Tr}\, t^A t^B = T_R \delta^{AB} = \frac{1}{2} \delta^{AB}. \tag{1.1.5}$$

The generators obey the following identities:

$$\sum_A t^A_{ab} t^A_{bc} = C_F \delta_{ac} = \frac{N^2 - 1}{2N} \delta_{ac} \,, \tag{1.1.6}$$

$$\mathrm{Tr}\, T^C T^D = \sum_{A,B} f^{ABC} f^{ABD} = C_A \delta^{CD} = N \delta^{CD}, \tag{1.1.7}$$

where N is the number of colours.

q_a are the quark fields and have mass m. $F^A_{\alpha\beta}$ is the field strength tensor, defined by:

$$F^A_{\alpha\beta} = \partial_\alpha A^A_\beta - \partial_\beta A^A_\alpha - g_s f^{ABC} A^B_\alpha A^C_\beta, \tag{1.1.8}$$

where the indices (A, B) run over the 8 colour degrees of freedom of the gluon fields and g_s is the strength of the coupling interaction. Note that in QCD it is frequently convenient to consider the quantity:

$$\alpha_s \equiv \frac{{g_s}^2}{4\pi} \,, \tag{1.1.9}$$

instead of g_s.

The third term in (1.1.8) is not present in quantum electrodynamics (QED) and is a consequence of the non-Abelian nature of QCD. It results in gluon self-interactions, which in turn contribute to the effects of asymptotic freedom and infra-red slavery.

$\mathcal{L}_{\text{class.}}$ is invariant under local gauge transformations, where the fields trans-

form in the following way:

$$
\begin{aligned}
&q_a(x) \to q'_a(x) = \exp(it\cdot\theta(x))_{ab} q_b(x)\,,\\
&D_{\alpha,ab} q_b(x) \to D'_{\alpha,ab} q'_b(x) \equiv \exp(it\cdot\theta(x))_{ab} D_{\alpha,bc} q_c(x)\,,\\
&t\cdot A_\alpha \to t\cdot A'_\alpha = \exp(it\cdot\theta(x))\, t\cdot A_\alpha \exp(-it\cdot\theta(x))\\
&\qquad\qquad + \frac{i}{g_s}\left(\partial_\mu \exp(it\cdot\theta(x))\right)\exp(-it\cdot\theta(x))\,,\\
&t\cdot F_{\alpha\beta} \to t\cdot F'_{\alpha\beta} = \exp(it\cdot\theta(x))\, t\cdot F_{\alpha\beta} \exp(-it\cdot\theta(x))\,,
\end{aligned}
\tag{1.1.10}
$$

where $t\cdot A_\alpha \equiv \sum_A t^A A^A_\alpha$ and the colour indices have been dropped in the last two equations for clarity.

Before reading off the Feynman rules it is necessary to fix the gauge. For covariant gauges $\partial^\alpha A_\alpha = 0$, so we can add the following term to the Lagrangian:

$$
\mathcal{L}_{\text{gauge}} = -\frac{1}{2\lambda}(\partial^\alpha A^A_\alpha)^2. \tag{1.1.11}
$$

Although $\mathcal{L}_{\text{gauge}}$ explicitly breaks gauge invariance, physical quantities will be independent of the gauge fixing parameter λ and the particular gauge chosen. The decision of which gauge to work in is usually driven by convenience of calculation.

Finally, in non-Abelian theories it is necessary to add ghosts. These are scalar fields that obey Fermi statistics. They cancel unphysical degrees of freedom that would otherwise propagate in gluon loops. The ghost term that needs to be added to the Lagrangian is:

$$
\mathcal{L}_{\text{ghost}} = \partial_\alpha \eta^{A\dagger}(D^\alpha_{AB}\eta^B). \tag{1.1.12}
$$

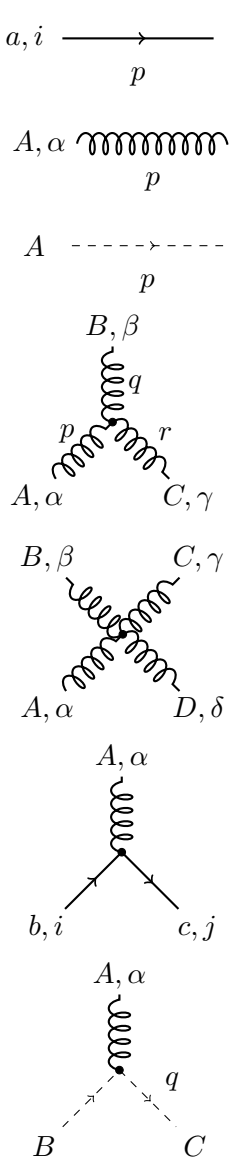
1.1.1 Feynman rules

The Feynman rules are obtained from the action: $S = i \int \mathcal{L}_{\mathrm{QCD}}\, d^4x$. It is not possible to invert for the gluon propagator without adding a gauge-fixing term to the Lagrangian. Only the gluon propagator is gauge-dependent. The Feynman rules for QCD in a covariant gauge are given in Fig. 1.1.

1.2 Renormalisation

In Fig. 1.1 we have identified g_s as the strength of the coupling between a quark and a gluon, and related m with the quark propagator. However, there will also be higher order loop corrections to these quantities, such as the ones in Fig. 1.2. There is no upper limit on the momenta which run in these loops and so they are UV divergent. Any quantities, including those which are physical observables, calculated with these loop corrections will be infinite. The parameters in the original Lagrangian can therefore not be the physical mass and coupling. In order to obtain finite predictions the parameters in the Lagrangian must be related to their physically observable counterparts. This is done by redefining, or renormalising, the quark mass, coupling, and fields to remove any infinities.

Before renormalising, we have to identify the divergences that are present. This is done by using a regularisation method. One such method is to introduce a cut-off, μ_{R}, on the momenta in the loops. Another technique, which is more generally used due to its preservation of gauge and Lorentz invariance, is dimensional regularisation [4, 24, 25]. Calculations are performed in $d = 4 - \epsilon$ dimensions. The terms in the Lagrangian have to be altered slightly when working in d dimensions: the coupling g_s has to be multiplied by $\mu_{\mathrm{R}}^{2-d/2}$, where μ_{R} is a mass scale. This keeps g_s dimensionless in d dimensions. An arbitrary



$$\delta^{ab}\frac{i}{(\not{p}-m+i\epsilon)_{ji}}$$

$$\delta^{AB}\left(-g^{\alpha\beta}+(1-\lambda)\frac{p^\alpha p^\beta}{p^2+i\epsilon}\right)\frac{i}{p^2+i\epsilon}$$

$$\delta^{AB}\frac{i}{(p^2+i\epsilon)}$$

$$-g_s f^{ABC}\left[(p-q)^\gamma g^{\alpha\beta}+(q-r)^\alpha g^{\beta\gamma}+(r-p)^\beta g^{\gamma\alpha}\right]$$

$$(\text{with } p+q+r=0)$$

$$-ig_s^2 f^{XAC}f^{XBD}\left[g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\delta}g^{\beta\gamma}\right]$$
$$-ig_s^2 f^{XAD}f^{XBC}\left[g^{\alpha\beta}g^{\gamma\delta}-g^{\alpha\gamma}g^{\beta\delta}\right]$$
$$-ig_s^2 f^{XAB}f^{XCD}\left[g^{\alpha\gamma}g^{\beta\delta}-g^{\alpha\delta}g^{\beta\gamma}\right]$$

$$-ig_s(t^A)_{cb}(\gamma^\alpha)_{ji}$$

$$g_s f^{ABC}q^\alpha$$

Figure 1.1: Feynman rules for QCD. Quarks are represented by solid lines, gluons by curly lines and ghosts by dotted lines. The gluon propagator is given in a covariant gauge with gauge fixing parameter λ.

dimensionful scale has been introduced by dimensional regularisation, as in the case of a cut-off applied to the momentum in the loop.

In this scheme, loops are found to contribute $1/\epsilon$ poles. Once the poles in ϵ have been identified they can be subtracted off or absorbed into multiplicative constants, called wave function renormalisation constants, in the renormali-

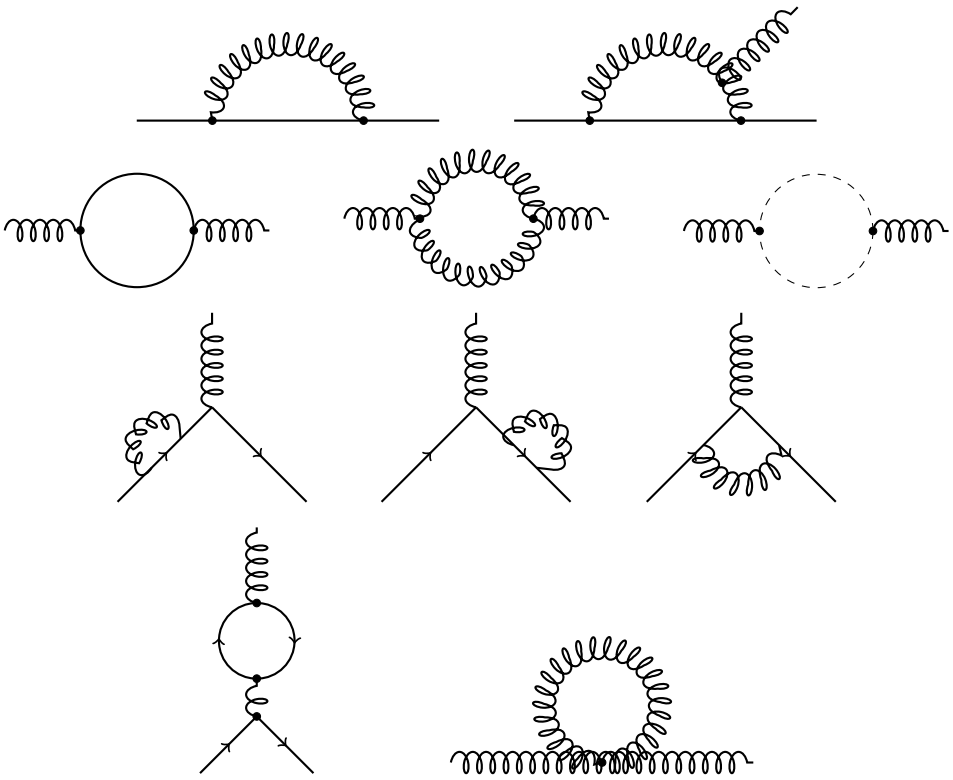


Figure 1.2: Higher order loop corrections to the quark propagator and the quark gluon vertex.

sation procedure. In addition to absorbing the infinite $1/\epsilon$ pole, some finite terms can also be removed. The exact finite terms to be included are specified by the renormalisation scheme, along with the choice of μ_{R}. Once the poles in ϵ are absorbed the limit $\epsilon \to 0$ can safely be taken, and the result is a finite calculation in $d = 4$ dimensions.

We can now identify renormalised masses, couplings and fields as those in the original 'bare' Lagrangian multiplied by combinations of the infinite wave function renormalisation constants. The infinities in the bare quantities and the renormalisation constants cancel to give finite 'physical' values.

An alternative view of renormalisation is that of the method of counter-terms. Instead of considering the original parameters to be infinite, they are

associated with the finite, physical masses and couplings. The fact that the Lagrangian does not yield finite quantities when calculating objects such as Green's functions and cross-sections is remedied by adding additional 'counter-terms' which cancel the infinite parts. The 'bare' Lagrangian, containing infinite bare parameters, consists of the original Lagrangian plus the counter-terms. These two views of renormalisation are equivalent, and both lead to a definition of renormalised quantities where UV divergences from loops are cancelled by infinite coefficients (the wave function renormalisation constants).

To illustrate the above discussion, consider the example of the self-energy of the quark propagator to one loop, as drawn in the top left hand corner of Fig. 1.2. Applying the Feynman rules from Fig. 1.1 to this diagram, and working in the Feynman gauge so that $\lambda = 1$, gives:

$$-i\Sigma^{ab}(p) = -g_s^2 \mu^{4-d} \int \frac{d^d k}{(2\pi)^d} \gamma_\mu \frac{1}{\not{p} - \not{k} - m} \gamma_\nu \frac{g^{\mu\nu}}{k^2} (t^A)^{ac} (t^A)^{cb} \,, \tag{1.2.1}$$

where we have dropped the $i\epsilon$ convention and the i, j subscripts from Fig. 1.1 for clarity. As stated above, when working in d dimensions it is necessary to introduce the extra multiplicative factor of $\mu^{2-d/2}$ to keep the coupling dimensionless. The integral is over the internal loop momentum, and the lack of an upper limit is the cause of the UV divergence. Performing the integration gives:

$$\Sigma^{ab}(p) = \left(\frac{g_s^2}{8\pi^2 \epsilon} (-\not{p} + 4m) + \text{finite} \right) (t^A t^A)^{ab} \,. \tag{1.2.2}$$

Working in $d = 4 - \epsilon$ dimensions has isolated the UV divergence in the form of the pole in ϵ. The additional non-divergent terms are represented by 'finite'.

The value of $(t^A t^A)^{ab}$ in SU(3) is given by (1.1.6), and leads to:

$$\begin{aligned}\Sigma^{ab}(p) &= \frac{4}{3}\delta^{ab}\left(\frac{g_s^2}{8\pi^2\epsilon}(-\not{p}+4m)+\text{finite}\right)\\ &= \frac{g_s^2}{6\pi^2\epsilon}(-\not{p}+4m)\delta^{ab}+\text{finite}\,.\end{aligned} \tag{1.2.3}$$

Extra counter-terms are needed in the Lagrangian to cancel the pole in ϵ, of the form:

$$\mathcal{L}_{CT} = \sum \bar{q}_a(iB\not{\partial}-A)_{ab}q_b\,, \tag{1.2.4}$$

where B and A are chosen to give a finite Σ^{ab} (at one loop). This gives the requirement:

$$\frac{g_s^2}{6\pi^2\epsilon}(-\not{p}+4m)\delta^{ab}+(A-B\not{p})\delta^{ab} = \text{finite}\,, \tag{1.2.5}$$

and so, neglecting the finite terms:

$$A = -\frac{2mg_s^2}{3\pi^2\epsilon} \qquad \text{and} \qquad B = -\frac{g_s^2}{6\pi^2\epsilon}\,. \tag{1.2.6}$$

Neglecting the finite terms and only subtracting off the poles is equivalent to working in the minimal subtraction (MS) scheme — for more information see Section 1.3.

The 'bare' Lagrangian, which consists of the original $\mathcal{L}$ and the counter-terms given in (1.2.4), is given by:

$$\begin{aligned}\mathcal{L}_B &= \sum \bar{q}_a(i(1+B)\not{\partial}-(m+A))_{ab}q_b\\ &= \sum(\overline{q_B})_a(i\not{\partial}-m_B)_{ab}(q_B)_b\,.\end{aligned} \tag{1.2.7}$$

The subscript B in these equations refers to the fact that they are bare quantities and should not be confused with an index over the colour degrees of freedom. In the second line of (1.2.7), the counter-terms have been absorbed

into a redefinition of the quark wave function and the mass. The new, renormalised, bare quark wave function is defined as:

$$(q_B)_a = \sqrt{Z_2} q_a \,, \tag{1.2.8}$$

with $Z_2 = 1 + B = 1 - g_s^2/6\pi^2\epsilon$, the quark wave function renormalisation constant.

The bare mass is defined as:

$$\begin{aligned} m_B &= \frac{1}{Z_2}(m + A) \\ &= m\left(1 - \frac{2g_s^2}{3\pi^2\epsilon}\right)\left(1 - \frac{g_s^2}{6\pi^2\epsilon}\right)^{-1} \\ &= m\left(1 - \frac{2g_s^2}{3\pi^2\epsilon}\right)\left(1 + \frac{g_s^2}{6\pi^2\epsilon} + \mathcal{O}(g_s^4)\right) \\ &= m\left(1 - \frac{g_s^2}{2\pi^2\epsilon} + \mathcal{O}(g_s^4)\right) \\ &= m + \delta m \,. \end{aligned} \tag{1.2.9}$$

The infinities that arose in calculations using the original Lagrangian are cancelled, at the one loop level, by the infinite multiplicative factor Z_2 and the δm term in the bare mass.

A similar process can be applied to the gluon propagator — which is modified by gluon, quark and ghost loops — and the quark-gluon vertex. This will result in a renormalised, bare gluon wave function and coupling. The UV divergences in physical observables, caused by the presence of loops, will be cancelled to $\mathcal{O}(g_s^4)$. This procedure can be repeated at higher orders of perturbation theory to remove divergences caused by the presence of larger numbers of loops. QCD is a renormalisable theory and so the parameters of the theory can be renormalised at all orders of perturbation theory.

1.2.1 Renormalisation Group Equation

The previous section described how to move between bare, unphysical, divergent fields, couplings, and masses and renormalised, physically observable ones through the use of multiplicative, infinite coefficients. When renormalising these quantities an arbitrary mass scale was introduced: the renormalisation scale, μ_{R}. This scale was introduced as a mathematical convenience during the process of dimensional regularisation and is unphysical. The bare quantities are independent of μ_{R}, but the renormalised quantities and the renormalisation constants do have a dependence on the unphysical scale μ_{R}. In other words, the bare quantities are invariant under transformations of the renormalisation group: $\mu_{\mathrm{R}} \to \mu'_{\mathrm{R}}$.

Predictions of more complicated quantities, such as Green's functions, can be renormalised by combining the wave function renormalisation constants in appropriate ways. This will, again, lead to a bare quantity that is equal to a renormalised prediction multiplied by some combination of renormalisation constants. For example, consider the bare n-point Green's function, $G_B^{(n)}$, for a general theory with coupling g, and fields with mass m that are renormalised by a wave function renormalisation constant $\sqrt{Z}$. $G_B^{(n)}$ can be related to the renormalised Green's function, $G^{(n)}$, in the following way:

$$G_B^{(n)}(x_i, g_B, m_B) = Z^{n/2} G^{(n)}(x_i, g(\mu_{\mathrm{R}}), m(\mu_{\mathrm{R}}))\,, \tag{1.2.10}$$

where the product of $\sqrt{Z}$ constants renormalises the wave functions contained in $G_B^{(n)}$, and i runs from 1 to n.

Bare quantities are independent of the arbitrary scale μ_{R}, so:

$$\mu_{\mathrm{R}} \frac{d}{d\mu_{\mathrm{R}}} G_B^{(n)}(x_i, g_B, m_B) = \mu_{\mathrm{R}} \frac{d}{d\mu_{\mathrm{R}}} \left(Z^{n/2} G^{(n)}(x_i, g(\mu_{\mathrm{R}}), m(\mu_{\mathrm{R}})) \right) = 0\,. \tag{1.2.11}$$

Using the chain rule this can be written as:

$$\left(\mu_{\mathrm{R}}\frac{\partial}{\partial\mu_{\mathrm{R}}}+\mu_{\mathrm{R}}\frac{\partial g}{\partial\mu_{\mathrm{R}}}\frac{\partial}{\partial g}+\mu_{\mathrm{R}}\frac{\partial m}{\partial\mu_{\mathrm{R}}}\frac{\partial}{\partial m}\right.$$
$$\left.+\frac{n}{\sqrt{Z}}\mu_{\mathrm{R}}\frac{\partial\sqrt{Z}}{\partial\mu_{\mathrm{R}}}\right)G^{(n)}(x_i,g(\mu_{\mathrm{R}}),m(\mu_{\mathrm{R}}))=0\,. \quad (1.2.12)$$

Defining the following quantities:

$$\beta(g)=\mu_{\mathrm{R}}\frac{\partial g}{\partial\mu_{\mathrm{R}}}\,, \quad (1.2.13)$$

$$\beta_m=\mu_{\mathrm{R}}\frac{\partial m}{\partial\mu_{\mathrm{R}}}\,, \quad (1.2.14)$$

$$\gamma=\frac{\mu_{\mathrm{R}}}{\sqrt{Z}}\frac{\partial\sqrt{Z}}{\partial\mu_{\mathrm{R}}}\,, \quad (1.2.15)$$

we can write (1.2.12) as:

$$\left(\mu_{\mathrm{R}}\frac{\partial}{\partial\mu_{\mathrm{R}}}+\beta(g)\frac{\partial}{\partial g}+\beta_m\frac{\partial}{\partial m}+n\gamma\right)G^{(n)}(x_i,g(\mu_{\mathrm{R}}),m(\mu_{\mathrm{R}}))=0\,. \quad (1.2.16)$$

This is known as the Callan-Symanzik equation, or the Renormalisation Group Equation (RGE). $\beta(g)$ is known as the β-function and describes the running of the coupling with the scale μ_{R}, and γ is the anomalous dimension of the renormalised fields.

For the specific case of QCD, there will be two $n\gamma$ terms: one for the gluon fields present in $G^{(n)}$ and one for the quarks. There is one other additional term that appears in the RGE for QCD, which describes the variation of the gauge fixing parameter, λ, with μ_{R}. The full QCD Callan-Symanzik equation is therefore:

$$\left(\mu_{\mathrm{R}}\frac{\partial}{\partial\mu_{\mathrm{R}}}+\beta(g_s)\frac{\partial}{\partial g_s}+\beta_m\frac{\partial}{\partial m}+n_q\gamma_q+n_g\gamma_g+\beta_\lambda\frac{\partial}{\partial\lambda}\right)G^{(n)}(x_i,g_s,m,\lambda)=0\,, \quad (1.2.17)$$

where $\beta_\lambda = \mu_\mathrm{R}\dfrac{\partial\lambda}{\partial\mu_\mathrm{R}}$, and $n_q + n_g = n$.

1.2.2 The QCD β-function

In the previous section, the function $\beta(g_s)$ was introduced in (1.2.13) to describe how the coupling varies, or runs, with the renormalisation scale. From now on we will consider the quantity α_s, as defined in (1.1.9), instead of g_s.

In QCD the perturbative expansion of $\beta(\alpha_s)$ is given by:

$$\frac{\partial\alpha_s}{\partial\ln\mu_\mathrm{R}} \equiv \beta(\alpha_s) = -\frac{b}{\pi}\alpha_s^2\left(1 + c\frac{\alpha_s}{\pi} + \sum_{n=2}^{\infty} c_n\left(\frac{\alpha_s}{\pi}\right)^n\right). \tag{1.2.18}$$

The first two coefficients:

$$b = \frac{33-2n_f}{6} \qquad \text{and} \qquad c = \frac{153-19n_f}{12b} \tag{1.2.19}$$

are universal, and n_f is the number of light flavours. The higher c_n are scheme-dependent coefficients. For example, c_2 in the $\overline{\mathrm{MS}}$ scheme is given by [2, 3]:

$$c_2^{\overline{\mathrm{MS}}} = \frac{77139 - 15099n_f + 325n_f^2}{1728b}. \tag{1.2.20}$$

The β-function coefficients are known up to c_3 in the $\overline{\mathrm{MS}}$ scheme [26, 27]:

$$c_3^{\overline{\mathrm{MS}}} = \frac{3564\zeta_3 + \frac{149753}{6} - (\frac{6508}{27}\zeta_3 + \frac{1078361}{162})n_f + (\frac{6472}{81}\zeta_3 + \frac{50065}{162})n_f^2 + \frac{1093}{729}n_f^3}{128b}, \tag{1.2.21}$$

where ζ is the Riemann zeta-function and $\zeta_3 \approx 1.202$.

For convenience, in this thesis we shall generally use the quantity $a = \alpha_s/\pi$.

Note that, since $b > 0$ for $n_f \leq 16$, the β-function for QCD is negative. This is the opposite of QED, where b_QED is negative and $\beta(g_\mathrm{QED}) > 0$. This results in the QED coupling getting larger at higher energies (larger μ_R) and

smaller at low energies. This can be explained physically by observing that the QED charge is smaller at large distances due to the shielding effect of vacuum polarisation. At higher energies, or shorter distances, this screening effect is penetrated and the observed charge increases.

In QCD the case is more complicated, since the gauge bosons also carry colour charge (unlike photons in QED which are neutral). This results in the sign of b changing, and hence the QCD coupling decreases with increasing energy. This phenomenon is known as asymptotic freedom. Since the coupling is small at sufficiently high energies, it is possible to use perturbation theory in this region.

At low energies the coupling becomes large, an effect known as infrared (IR) slavery. In regions of low energy, perturbation theory breaks down and non-perturbative effects become more important. IR slavery is manifested as the confinement of quarks and gluons — only colour singlet states can propagate over large distances. At high energies quarks and gluons behave more like free particles.

1.2.3 Solutions of the running coupling equation

At leading order (LO) the QCD β-function is:

$$\frac{\partial a^{(1)}}{\partial \ln \mu_{\mathrm{R}}} = -b(a^{(1)})^2, \tag{1.2.22}$$

where the (1) superscript on a indicates that this is the one-loop coupling constant. This can be easily integrated to give a solution for the one-loop running coupling constant:

$$a^{(1)}(\mu_{\mathrm{R}}) = \frac{1}{b \ln \dfrac{\mu_{\mathrm{R}}}{\tilde{\Lambda}}}\,. \tag{1.2.23}$$

$\tilde{\Lambda}$ is related to $\Lambda_{\rm QCD}$, the characteristic scale of QCD. $\Lambda_{\rm QCD}$ represents the energy where the coupling becomes large, and hence the region where perturbation theory breaks down. In this thesis we have considered a Lagrangian that contains a mass term for the quark fields. It would be possible to neglect these terms, however, and repeat the treatment for massless QCD. In this case renormalisation still introduces a massive scale, $\Lambda_{\rm QCD}$, despite there being no massive scales present in the original Lagrangian. This process is known as dimensional transmutation, and $\Lambda_{\rm QCD}$ is therefore called the dimensional transmutation parameter [2,5,28]. The tilde on Λ reflects the particular definition used in this thesis.[1] $\tilde{\Lambda}$ can be related to the standard definition of $\Lambda_{\rm QCD}$ using:

$$\Lambda_{\rm QCD} = \left(\frac{2c}{b}\right)^{c/b} \tilde{\Lambda}\,. \tag{1.2.24}$$

$\Lambda_{\rm QCD}$ is a scheme dependent quantity, but it is possible to move between Λ's in different schemes exactly, as long as the NLO coefficients in both schemes are known — for more details see the discussion in Section 2.1 below (2.1.19).

At next-to-leading order (NLO) the QCD β-function gains an extra term:

$$\frac{\partial a^{(2)}}{\partial \ln \mu_{\rm R}} = -b\left(a^{(2)}\right)^2\left(1 + ca^{(2)}\right), \tag{1.2.25}$$

where we are now looking at the two-loop coupling constant, hence the (2)

[1] The exact definition of Λ depends on the way the β function is integrated, and specifically how the infinite integration constant is defined — see Section 2.1 for a further discussion.

superscript. Integrating this equation gives:

$$\begin{aligned} b \ln \frac{\mu_{\mathrm{R}}}{\tilde{\Lambda}} &= \int_{\infty}^{a} dx \frac{-1}{x^2(1+cx)} \\ &= \int_{\infty}^{a} dx \left[-\frac{1}{x^2} + \frac{c}{x} - \frac{c^2}{1+cx} \right] \\ &= \left[\frac{1}{x} + c \ln x - c \ln(1+cx) \right]_{\infty}^{a} \\ &= \frac{1}{a} + c \ln \left(\frac{ca}{1+ca} \right) , \end{aligned} \tag{1.2.26}$$

where we have used asymptotic freedom as a boundary condition.[2] Equation (1.2.26) can be inverted analytically:

$$a^{(2)}(\mu_{\mathrm{R}}) = \frac{-1}{c\,[1 + W(z)]} , \tag{1.2.27}$$

where W is the Lambert W-function [29], defined by $W(z)\exp(W(z)) = z$, and z is given by:

$$z = -\exp\left[-\left(1 + \frac{b}{c} \ln \frac{\mu_{\mathrm{R}}}{\tilde{\Lambda}} \right) \right] \tag{1.2.28}$$

The W_{-1} branch should be used to ensure asymptotic freedom [30].

Solving the next-to-NLO (NNLO) QCD β-function for the three-loop coupling constant can not be done analytically. However, it is possible to relate the coupling at NNLO to an expansion in powers of $a^{(2)}$ as follows:

$$a^{(3)}(\mu_{\mathrm{R}}) = a^{(2)}(\mu_{\mathrm{R}}) + c_2 \left(a^{(2)}(\mu_{\mathrm{R}}) \right)^3 + \dots . \tag{1.2.29}$$

Using this power expansion for the higher order coupling avoids the need to solve transcendental equations. The approximation using the first two terms is

[2] The introduction of an integration constant is required to implement this boundary condition and to cancel singularities at $x = 0$ in the integrand. This treatment is shown explicitly in the context of the Effective Charges method in Section 2.1.

valid to NNLO. Higher order couplings can be approximated by adding higher order terms. For example the next-to-NNLO (N^3LO) QCD coupling can be approximated by:

$$a^{(4)}(\mu_R) = a^{(2)}(\mu_R) + c_2 \left(a^{(2)}(\mu_R)\right)^3 + \frac{c_3}{2} \left(a^{(2)}(\mu_R)\right)^4 + \dots . \tag{1.2.30}$$

As we shall see in the next section, at each new order of perturbation theory the renormalisation scheme is labelled by an extra coefficient from the QCD β-function equation. Therefore, in moving from the NNLO coupling to the N^3LO coupling a term involving c_3 appears.

1.3 Renormalisation Scheme Dependence

As discussed in Section 1.2, the process of renormalisation introduces an arbitrary scale, μ_R. There is also another choice to be made: which, if any, finite terms are to be absorbed along with the poles in ϵ by the renormalisation constants. These two choices, the selection of a μ_R and of a prescription for the exact term absorbed, together specify the renormalisation scheme (RS). The minimal subtraction (MS) scheme specifies that only the ϵ poles should be absorbed. The commonly used modified minimal subtraction ($\overline{\text{MS}}$) scheme also absorbs a finite term : $(\ln 4\pi + \gamma_E)$, where γ_E is the Euler gamma constant.

A dimensionless, physical, IR and collinear safe observable, dependent on a single energy scale Q, can be expanded in perturbation theory in terms of the coupling, $a = \alpha_s/\pi$:

$$O = a(1 + p_1 a + p_2 a^2 + \ldots + p_n a^n + \ldots) . \tag{1.3.1}$$

At n^{th} order of perturbation theory the RS dependence of $O^{(n)}$ is given by the

set of parameters $\{\tau, c_2, \ldots, c_{n-1}\}$ [16], where:

$$\tau = b \ln(\mu_{\rm R}/\tilde{\Lambda}) \,, \tag{1.3.2}$$

and $O^{(n)}$ is (1.3.1) truncated beyond the $\mathcal{O}(a^n)$ term.

In order to see that the RS dependence is given solely by $\{\tau, c_2, \ldots, c_{n-1}\}$ we follow the argument given by Stevenson in Ref. [16]. First, consider the observable O at $n^{\rm th}$ order of perturbation theory in two different schemes, $O^{(n)}_{\rm RS1}$ and $O^{(n)}_{\rm RS2}$. These quantities have an expansion of the form given in (1.3.1) terminated at a^n, with couplings $a_{\rm RS1}$ and $a_{\rm RS2}$ respectively. Similarly the perturbative coefficients will vary between the schemes: $p_n^{\rm RS1}$ and $p_n^{\rm RS2}$. The approximations $O^{(n)}_{\rm RS1}$ and $O^{(n)}_{\rm RS2}$ agree with the full expansion O to a^{n+1} order:

$$\begin{aligned} O &= O^{(n)}_{\rm RS1} + \mathcal{O}(a^{n+1}) \\ &= O^{(n)}_{\rm RS2} + \mathcal{O}(a^{n+1}) \,. \end{aligned} \tag{1.3.3}$$

$O^{(n)}_{\rm RS1}$ and $O^{(n)}_{\rm RS2}$ therefore differ by $\mathcal{O}(a^{n+1})$. This is known as the self-consistency of perturbation theory and can be written as:

$$\frac{\partial O^{(n)}}{\partial ({\rm RS})} = \mathcal{O}(a^{n+1}) \,. \tag{1.3.4}$$

As stated above, the RS in this equation can be specified by $\{\tau, c_2, \ldots\}$ only. If this were not the case, and there was an additional parameter η that contributed to the RS and was independent of $\{\tau, c_2, \ldots\}$, then the observable $O^{(n)}$ (in either scheme) would depend on η in the following way:

$$\frac{\partial O^{(n)}}{\partial \eta} = \sum_{i=1}^{n-1} \frac{\partial p_i}{\partial \eta} (a^{(n)})^{i+1} \,. \tag{1.3.5}$$

This only agrees with the self-consistency requirement in (1.3.4) if:

$$\frac{\partial p_i}{\partial \eta} = 0 \qquad \text{for } i = 1, \ldots, n-1\,, \tag{1.3.6}$$

and so the truncated approximation $O^{(n)}$ cannot depend on η. We conclude that the RS dependence of $O^{(n)}$ is given solely by $\{\tau, c_2, \ldots, c_{n-1}\}$.

1.3.1 Physical scale choice

As seen in the previous section, a truncated expansion of a generic QCD observable depends on the RS through the parameters $\{\tau, c_2, \ldots, c_{n-1}\}$. We therefore have to choose the values of these parameters before we make a calculation in fixed order perturbation theory. A standard approach is to use the $\overline{\text{MS}}$ scheme (or some alternate RS) and to set $\mu_\text{R} = Q$ where Q is a physical scale that the observable depends on, such as the centre of mass energy in e^+e^- collisions. Equation (1.3.1) then becomes:

$$O = a(Q)(1 + p_1^{\overline{\text{MS}}}(\mu_\text{R} = Q)a(Q) + \ldots + p_n^{\overline{\text{MS}}}(\mu_\text{R} = Q)a(Q)^n + \ldots)\,. \tag{1.3.7}$$

At leading order the β function for the coupling is given by (1.2.22). In Section 1.2.3 we integrated this equation between an arbitrary μ_R and Λ_QCD, where the coupling tends to ∞. Now consider the same equation integrated between μ_R and an energy Q:

$$-\frac{1}{a(\mu_\text{R})} + \frac{1}{a(Q)} = b \ln \frac{\mu_\text{R}}{Q}\,. \tag{1.3.8}$$

Rearranging this for $a(Q)$ gives:

$$a(Q) = \frac{a(\mu_{\rm R})}{1 + b \ln\left(\frac{\mu_{\rm R}}{Q}\right) a(\mu_{\rm R})} \tag{1.3.9}$$

$$= a(\mu_{\rm R}) \left(1 - b \ln\left(\frac{\mu_{\rm R}}{Q}\right) a(\mu_{\rm R}) + \left(b \ln\left(\frac{\mu_{\rm R}}{Q}\right)\right)^2 a(\mu_{\rm R})^2 + \ldots\right) . \tag{1.3.10}$$

We see that an expansion of $a(Q)$ involves logs of $\mu_{\rm R}/Q$. Setting $\mu_{\rm R} = Q$ (or close to Q) causes the logs to disappear (or to be small). Choosing a $\mu_{\rm R}$ that is far from Q will guarantee large truncation errors, although the reverse is not necessarily true. This is the rationale behind choosing the physical scale [5]. As an estimate of the effect of RS dependence in missing higher orders $\mu_{\rm R}$ is then often varied between $0.5Q$ and $2Q$, or some similar range of values, to give a theoretical renormalisation scale error.

This procedure is highly arbitrary and can result in theoretical uncertainties that are larger than the experimental uncertainties on the physical observables involved. In the remainder of this thesis we examine two alternative methods which attempt to minimise the RS dependence of QCD predictions: the Effective Charges (ECH) method in Chapter 2, and the Principle of Minimal Sensitivity (PMS) in Chapter 4.

Chapter 2

The Effective Charges method

Following on from our discussion of RS dependence and the physical scale method we shall now look at the Effective Charges (ECH) method. This method integrates up a β-function type equation for the observable that is analogous to the running equation for the strong coupling constant. The resulting integration constant is related to $\Lambda_{\rm QCD}$, the characteristic scale of QCD [2, 3, 5]. Once determined this universal parameter of QCD can be used to evaluate $\alpha_s(M_Z)$ using the QCD β-function equation, as outlined in Section 1.2.3. As we shall see, ECH gives predictions that are RS invariant. There is no unphysical dependence on $\mu_{\rm R}$ and therefore no need to estimate a theoretical uncertainty to account for this dependence.

In Section 2.1 we introduce the ECH method and give a brief outline of its derivation. We then discuss the data on event shape moments from e^+e^- colliders in Section 2.2, which we go on to analyse using the ECH method in the next chapter.

2.1 Derivation of the ECH method

In this section we shall derive the ECH method [7–16][1] and show how it can be used to extract $\Lambda_{\overline{\rm MS}}$, $\Lambda_{\rm QCD}$ in the $\overline{\rm MS}$ scheme, and hence the strong coupling constant.[2] We shall start from the principle of dimensional analysis [28]. Consider a dimensionless QCD observable which depends on a single energy scale, $\mathcal{R}(Q)$. Since the observable is dimensionless we have to introduce another, dimensionful, quantity Λ:

$$\mathcal{R}(Q) = f\left(\frac{\Lambda}{Q}\right), \tag{2.1.1}$$

where f is some function of Λ and the physical scale Q. We will assume massless quarks; the extension to the case with non-zero quark mass has been considered in Refs [12,14]. The parameter Λ must therefore be related to $\Lambda_{\rm QCD}$, the dimensional transmutation scale of QCD, introduced in Section 1.2.3. We want to extract Λ, and to do this we must invert (2.1.1). To find the inverse function f^{-1} we shall start by looking at the derivative of $\mathcal{R}(Q)$ with respect to Q:

$$\frac{d\mathcal{R}(Q)}{dQ} \equiv \frac{\rho(\mathcal{R}(Q))}{Q}. \tag{2.1.2}$$

We have again used dimensional analysis to determine the form of (2.1.2). The expression can be rewritten as:

$$\frac{d\mathcal{R}(Q)}{d\ln Q} \equiv \rho(\mathcal{R}(Q)). \tag{2.1.3}$$

To solve (2.1.3) we need to know something about $\mathcal{R}(Q)$. We shall take the form of its perturbative expansion to be:

$$\mathcal{R}(Q) = a(1 + r_1 a + r_2 a^2 + \dots), \tag{2.1.4}$$

[1] The description of ECH in this Chapter is based closely on that given in Ref. [9]

[2] $\Lambda_{\rm QCD}$ and α_s can be related using the equations in Section 1.2.3.

where $a \equiv \alpha_s(\mu_{\rm R})/\pi$ and $\mu_{\rm R}$ is the renormalisation scale. For any observable of this type, i.e. one that depends on only a single energy scale, we can always ensure that the perturbative expansion takes the form of (2.1.4) by applying an appropriate scaling and, if necessary, raising to a power. This step therefore results in no loss of generality.

An observable with an expansion of the form of (2.1.4) is known as an effective charge [15, 31] as it has a running equation like the strong coupling does. The running coupling equation for α_s was introduced in Section 1.2.2. We can derive the analogous expression for the effective charge by setting $\mu_{\rm R} = Q$ in (2.1.4) and differentiating with respect to $\ln Q$. If we also invert (2.1.4) to get an expression for $a(\mathcal{R})$ and substitute in we find that:

$$\frac{d\mathcal{R}(Q)}{d\ln Q} \equiv \rho(\mathcal{R}(Q)) = -b\mathcal{R}^2(1 + c\mathcal{R} + \sum_{n=2}^{\infty} \rho_n \mathcal{R}^n)\,, \tag{2.1.5}$$

where b and c are the first two universal coefficients from (1.2.18). The ρ_n are specific to the observable $\mathcal{R}(Q)$; they are RS invariant and Q-independent combinations of coefficients from (2.1.4) and (1.2.18). For example:

$$\rho_2 = r_2 + c_2 - r_1 c - r_1^2. \tag{2.1.6}$$

More details of this step are given in Appendix A.

We now integrate (2.1.5) to give:

$$\ln \frac{Q}{\Lambda_{\mathcal{R}}} = \int_0^{\mathcal{R}(Q)} \frac{dx}{\rho(x)} + \kappa. \tag{2.1.7}$$

This step has introduced an integration constant that we have split into two parts: $\ln \Lambda_{\mathcal{R}}$ and κ. The constant $\Lambda_{\mathcal{R}}$ is a finite dimensionful scale specific to the observable $\mathcal{R}$. κ is an infinite constant needed to implement the boundary

condition of aymptotic freedom,[3] i.e. the requirement that $\mathcal{R}(Q) \to 0$ as $Q \to \infty$, and which must cancel the singularities in $1/\rho(x)$ at $x = 0$. Once this requirement is met there is still some freedom to choose the exact form of κ.

Consider a κ of the form:

$$\begin{aligned} \kappa &= \int_0^\Gamma \frac{dx}{bx^2(1+cx+\Delta)} \\ &\simeq \int_0^\Gamma \frac{dx}{bx^2}(1-cx-\Delta)\,. \end{aligned} \tag{2.1.8}$$

The first two terms in (2.1.8) will cancel the singularities at $x = 0$ in:

$$\begin{aligned} \frac{1}{\rho(x)} &= -\frac{1}{bx^2(1+cx+\rho_2 x^2+\ldots)} \\ &= -\left(\frac{1}{bx^2} - \frac{c}{bx} - \frac{\rho_2}{b} - \ldots\right)\,. \end{aligned} \tag{2.1.9}$$

The only restriction on Δ, therefore, is that Δ/x^2 is finite as $x \to 0$. This is achieved by setting $\Delta = 0$. A convenient choice for Γ is ∞, which gives:

$$\kappa = \int_0^\infty \frac{dx}{bx^2(1+cx)}. \tag{2.1.10}$$

Any variation in the choices of Δ and Γ, and the definition of κ, can be absorbed into the definition of $\Lambda_{\mathcal{R}}$.

Adding in this explicit form for κ we can now rewrite (2.1.7) as:

$$\begin{aligned} b\ln\frac{Q}{\Lambda_{\mathcal{R}}} &= \int_{\mathcal{R}(Q)}^\infty \frac{dx}{x^2(1+cx)} + \int_0^{\mathcal{R}(Q)} dx\left[\frac{b}{\rho(x)} + \frac{1}{x^2(1+cx)}\right] \\ &= \mathrm{F}(\mathcal{R}) + \mathrm{G}(\mathcal{R}). \end{aligned} \tag{2.1.11}$$

The integration in κ has been split into two parts: from 0 to $\mathcal{R}(Q)$ and from

[3]Note that this treatment is analogous to that required for integration of the β-function in Section 1.2.3. In order to obtain the equivalent method for the coupling, replace $\Lambda_{\mathcal{R}}$ with $\tilde{\Lambda}$, Q with μ_{R}, $\mathcal{R}(Q)$ with $\alpha_s(\mu_{\mathrm{R}})$ and $\rho(x)$ with $\beta(x)$.

$\mathcal{R}(Q)$ to ∞. Integrating up the first term on the right-hand side gives:

$$\mathrm{F}(\mathcal{R}) = \frac{1}{\mathcal{R}} + c\ln\left[\frac{c\mathcal{R}}{1+c\mathcal{R}}\right]. \tag{2.1.12}$$

The form of the second integral, $\mathrm{G}(\mathcal{R})$, depends on the order at which $\rho(x)$ is truncated. For example, at NLO we have $\rho(x) = -bx^2(1+cx)$ which cancels with the $\frac{1}{x^2(1+cx)}$ term to give $\mathrm{G}(\mathcal{R}) = 0$. At NNLO we have:

$$\begin{aligned}\mathrm{G}(\mathcal{R}) &= \int_0^{\mathcal{R}(Q)} dx \left[\frac{-1}{x^2(1+cx+\rho_2 x^2)} + \frac{1}{x^2(1+cx)}\right] \\ &= \int_0^{\mathcal{R}(Q)} dx \frac{\rho_2}{(1+cx)(1+cx+\rho_2 x^2)}.\end{aligned} \tag{2.1.13}$$

This integration has an analytic result with three different forms, depending on the relative sizes of the coefficients c and ρ_2:

$(i)\, 4\rho_2 > c^2; \qquad \Delta \equiv \sqrt{4\rho_2 - c^2},$

$$\mathrm{G}(\mathcal{R}) = \frac{c}{2}\ln\left|\frac{(1+c\mathcal{R})^2}{1+c\mathcal{R}+\rho_2\mathcal{R}^2}\right| + \frac{2\rho_2 - c^2}{\Delta}\left[\arctan\left[\frac{2\rho_2\mathcal{R}+c}{\Delta}\right] - \arctan\left[\frac{c}{\Delta}\right]\right],$$

$(ii)\, 4\rho_2 < c^2; \qquad \Delta \equiv \sqrt{c^2 - 4\rho_2},$

$$\mathrm{G}(\mathcal{R}) = \frac{c}{2}\ln\left|\frac{(1+c\mathcal{R})^2}{1+c\mathcal{R}+\rho_2\mathcal{R}^2}\right| + \frac{2\rho_2 - c^2}{2\Delta}\left[\ln\left|\frac{2\rho_2\mathcal{R}+c-\Delta}{2\rho_2\mathcal{R}+c+\Delta}\right| + \ln\left|\frac{c+\Delta}{c-\Delta}\right|\right],$$

$(iii)\, 4\rho_2 = c^2;$

$$\mathrm{G}(\mathcal{R}) = \frac{c}{2}\ln\left|\frac{(1+c\mathcal{R})^2}{1+c\mathcal{R}+\rho_2\mathcal{R}^2}\right| + (c^2 - 2\rho_2)\left[\frac{1}{2\rho_2\mathcal{R}+c} - \frac{1}{c}\right]. \tag{2.1.14}$$

We now rearrange for $\Lambda_{\mathcal{R}}$ in (2.1.11):

$$\Lambda_{\mathcal{R}} = Q\mathcal{F}(\mathcal{R})\mathcal{G}(\mathcal{R}), \tag{2.1.15}$$

where

$$\mathcal{F}(\mathcal{R}) = \exp\left(-\frac{\mathrm{F}(\mathcal{R})}{b}\right) = e^{-1/b\mathcal{R}}\left(1+\frac{1}{c\mathcal{R}}\right)^{c/b}, \tag{2.1.16}$$

and

$$\mathcal{G}(\mathcal{R}) = \exp\left(-\mathrm{G}(\mathcal{R})/b\right). \tag{2.1.17}$$

$\Lambda_{\mathcal{R}}$ is a scheme independent quantity associated with the specific effective charge in question. It can be related to the universal quantity $\tilde{\Lambda}_{\overline{\mathrm{MS}}}$ using:

$$\tilde{\Lambda}_{\overline{\mathrm{MS}}} = \Lambda_{\mathcal{R}}\, e^{-r/b}, \tag{2.1.18}$$

where $r \equiv r_1^{\overline{\mathrm{MS}}}(\mu_{\mathrm{R}} = Q)$. The tilde in (2.1.18) indicates that this definition is slightly different from the standard one for $\Lambda_{\overline{\mathrm{MS}}}$ (Λ_{QCD} in the $\overline{\mathrm{MS}}$ scheme). This is due to our choice of κ and, to regain the standard definition of $\Lambda_{\overline{\mathrm{MS}}}$ [32,33], we have to shift κ by $c\ln(b/2c)$, as in (1.2.24), resulting in the relation:

$$\begin{aligned}\Lambda_{\overline{\mathrm{MS}}} &= \left(\frac{2c}{b}\right)^{c/b}\tilde{\Lambda}_{\overline{\mathrm{MS}}} \\ &= \left(\frac{2c}{b}\right)^{c/b} Q\mathcal{F}(\mathcal{R})\mathcal{G}(\mathcal{R})e^{-r/b}.\end{aligned} \tag{2.1.19}$$

Note that all scheme dependence in $\Lambda_{\overline{\mathrm{MS}}}$ is contained in the exponential factor $e^{-r/b}$. Although $\Lambda_{\overline{\mathrm{MS}}}$ is a scheme dependent quantity it is possible to relate Λ's in different schemes exactly, providing that the NLO coefficients are known in both schemes. To move between Λ's in different schemes we note that, from (2.1.18), $\tilde{\Lambda}_{\overline{\mathrm{MS}}}\, e^{r/b} = \Lambda_{\mathcal{R}}$. Since $\Lambda_{\mathcal{R}}$ is scheme invariant we can equate the left-hand side of this expression in two different schemes. For instance, for Λ in the $\overline{\mathrm{MS}}$ scheme and Λ in the MS scheme the relation is:

$$\Lambda_{\overline{\mathrm{MS}}} = \exp\left\{\frac{r^{\mathrm{MS}} - r^{\overline{\mathrm{MS}}}}{b}\right\}\Lambda_{\mathrm{MS}}, \tag{2.1.20}$$

where $r^{\overline{\rm MS}}$ and $r^{\rm MS}$ are the NLO coefficients in the $\overline{\rm MS}$ and MS schemes respectively.

From $\Lambda_{\overline{\rm MS}}$ we can find $\alpha_s(M_Z)$ using the equations in Section 1.2.3, replicated here for ease of reference:

$$a = \frac{-1}{c\left[1 + W\left(-\exp\left[-\left(1 + \frac{b}{c}\ln\frac{M_Z}{\tilde{\Lambda}_{\overline{\rm MS}}}\right)\right]\right)\right]}, \tag{2.1.21}$$

$$\alpha_s(M_Z) = \begin{cases} \pi a & \text{for NLO} \\ \pi(a + c_2 a^3) & \text{for NNLO,} \end{cases} \tag{2.1.22}$$

where W is the Lambert W-function [29]. The coupling a in (2.1.21) is the $\overline{\rm MS}$ coupling corresponding to a two-loop (NLO) β-function solution, with the β-function coefficients c_2 and higher set to zero. This can be related to the coupling at NNLO, which has a non-zero c_2 but vanishing higher order coefficients, by an expansion in powers of a — see the bottom line of (2.1.22) and (1.2.29) in Section 1.2.3.

2.2 Event Shape Observables

Event shape observables give information about the topology of an event, and are designed to be IR- and collinear-safe. They are good observables to use to study α_s since they start at $\mathcal{O}(\alpha_s)$ in perturbation theory. In this thesis we analyse e^+e^- event shape moments. The $n^{\rm th}$ moment of an event shape observable, y, is given by [6, 34, 35]:

$$\langle y^n \rangle = \frac{1}{\sigma_{\rm had}} \int_0^{y_{\rm max}} y^n \frac{d\sigma}{dy} dy, \tag{2.2.1}$$

where $y_{\max}$ is the greatest value of the event shape that is kinematically allowed, $\frac{d\sigma}{dy}$ is the differential cross-section, and σ_{had} is the hadronic cross-section.

$\langle y^n \rangle$ has the following perturbative expansion [6, 34]:

$$\begin{aligned}\langle y^n \rangle = &\left(\frac{\alpha_s(\mu_{\text{R}})}{2\pi}\right)\overline{\mathcal{A}}_{y,n} + \left(\frac{\alpha_s(\mu_{\text{R}})}{2\pi}\right)^2\left(\overline{\mathcal{B}}_{y,n} + \overline{\mathcal{A}}_{y,n}\, b \ln\frac{\mu_{\text{R}}^2}{Q^2}\right) \\ &+ \left(\frac{\alpha_s(\mu_{\text{R}})}{2\pi}\right)^3\left(\overline{\mathcal{C}}_{y,n} + 2\overline{\mathcal{B}}_{y,n}\, b \ln\frac{\mu_{\text{R}}^2}{Q^2} + \overline{\mathcal{A}}_{y,n}\left(b^2 \ln^2\frac{\mu_{\text{R}}^2}{Q^2} + 2bc \ln\frac{\mu_{\text{R}}^2}{Q^2}\right)\right) \\ &+ \ldots, \end{aligned} \tag{2.2.2}$$

where the b and c are coefficients from the running coupling equation (1.2.18), and $\overline{\mathcal{A}}_{y,n}$, $\overline{\mathcal{B}}_{y,n}$ and $\overline{\mathcal{C}}_{y,n}$ are perturbative coefficients that are specific to the event shape moment. The NNLO perturbative coefficients ($\overline{\mathcal{C}}_{y,n}$) for several event shapes have recently been calculated [34, 36, 37].

In order to apply the ECH method to an event shape moment, we must first rearrange (2.2.2) to get it into the form of an effective charge as given in (2.1.4). To do this we divide through by $\overline{\mathcal{A}}_{y,n}$ and, remembering that $a \equiv \alpha_s(\mu_{\text{R}})/\pi$, put in appropriate factors of two:

$$\mathcal{R}(Q) = \frac{2\langle y^n \rangle}{\overline{\mathcal{A}}_{y,n}} = a\left(1 + \frac{\overline{\mathcal{B}}_{y,n}}{2\overline{\mathcal{A}}_{y,n}}a + \frac{\overline{\mathcal{C}}_{y,n}}{4\overline{\mathcal{A}}_{y,n}}a^2 + \ldots\right). \tag{2.2.3}$$

For clarity we have taken $\mu_{\text{R}} = Q$. Comparing this with (2.1.4) we see that

$$r_1 = \frac{\overline{\mathcal{B}}_{y,n}}{2\overline{\mathcal{A}}_{y,n}} \qquad \text{and} \qquad r_2 = \frac{\overline{\mathcal{C}}_{y,n}}{4\overline{\mathcal{A}}_{y,n}}. \tag{2.2.4}$$

In this thesis we examine the first three moments of five different event shape observables: one minus the thrust $(1 - T)$ [38, 39], the C-parameter (C) [40, 41], the wide and total jet broadenings (B_W and B_T) [42, 43], and the three-to-two jet transition parameter in the Durham algorithm (Y_3) [44].

In addition, we analyse the mean (1^{st} moment) of the heavy jet mass in the E-scheme ($\rho_E = (M_h^2/s)_E$) [35, 45].

Definitions of these event shape observables are as follows:

- Thrust, T:

$$T = \max \frac{\sum_i |\boldsymbol{p}_i \cdot \boldsymbol{n}_T|}{\sum_i |\boldsymbol{p}_i|}\,, \tag{2.2.5}$$

  where $\boldsymbol{p}_i$ are the final state particle momenta and $\boldsymbol{n}_T$ is a unit vector in the direction of the thrust axis. The thrust axis is determined by maximising T. Thrust varies between 1, for back-to-back events, and 1/2, for isotropic events. It is conventional to use the quantity $1 - T$, which is 0 in the two-jet limit, to be consistent with other event shape observables.

- C-parameter, C:

$$C = 3(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1)\,, \tag{2.2.6}$$

  where λ_i are eigenvalues of the linear momentum tensor:

$$\theta_{mn} = \frac{\sum_i p_i^m p_i^n/|\boldsymbol{p}_i|}{\sum_i |\boldsymbol{p}_i|}\,,$$

  with m, n running over the components of $\boldsymbol{p}_i$.

- Wide jet broadening, B_W: jet broadening is defined by splitting events into two hemispheres ($\boldsymbol{H}_a, \boldsymbol{H}_b$) using a plane perpendicular to $\boldsymbol{n}_T$. The broadening in each hemisphere is then:

$$B_x = \frac{\sum_{\boldsymbol{p}_i \in \boldsymbol{H}_x} |\boldsymbol{p}_i \times \boldsymbol{n}_T|}{2\sum_{\boldsymbol{p}_i \in \boldsymbol{H}_x} |\boldsymbol{p}_i|}\,, \tag{2.2.7}$$

  where x runs over the hemispheres a and b. B_W is the larger of the two broadenings, i.e. $B_W = \max(B_a, B_b)$.

- Total jet broadening, B_T:

$$B_T = B_a + B_b \,, \tag{2.2.8}$$

  with B_a and B_b defined as for the wide jet broadening.

- Three-to-two jet transition parameter in the Durham algorithm, Y_3: the value of the jet measure y_{ij} at which an event changes from being classified as a three-jet event to a two-jet one. In the Durham algorithm the jet measure is defined as:

$$y_{ij}^D = \frac{2\min(E_i^2, E_j^2)(1-\cos\theta_{ij})}{E_{\rm vis}^2} \,, \tag{2.2.9}$$

  where E_i are the energies of the final state hadrons, θ_{ij} is the angle between them, and $E_{\rm vis} = \sum_i E_i$.

- Heavy jet mass in the E-scheme, ρ_E: events are divided into two hemispheres, as for the broadening observables, and the invariant mass of each is calculated. The heavy jet mass is the larger of the two masses normalised to $E_{\rm vis}$:

$$\rho = \max \frac{M_x^2}{E_{\rm vis}^2} = \max \frac{1}{E_{\rm vis}^2} \left(\sum_{\boldsymbol{p}_k \in \boldsymbol{H}_x} p_k \right)^2 \,. \tag{2.2.10}$$

  The use of the E-scheme corrects for the effect of hadron masses. In the E-scheme the four-momentum $p = (\boldsymbol{p}, E)$ is replaced by $(\hat{p}E, E)$ where $\hat{p}$ is a unit vector in the $\boldsymbol{p}$-direction. The definition of ρ in the E-scheme is identical to (2.2.10) for massless hadrons.

The data come from various LEP experiments including OPAL [17, 19], DELPHI [20, 21] and L3 [22]. Some low energy JADE (PETRA) data are also

included [18, 19]. The data cover a range of energies from 14.0 to 206.6 GeV.

All observables are corrected for b and c quark decays using HERWIG++ simulations [46, 47]. Samples of 10^6 events were run with $N_f = 3$ (only light quarks) and also with $N_f = 5$ (with b and c quarks). The event shape moment was calculated in both of these cases and a ratio of the two quantities was taken to give a correction factor. This was done for each energy. The data in this thesis have been multiplied by the appropriate correction factors. The correction factors for the means ($n = 1$ moments) are shown in Fig. 2.1. Plots of the correction factors for the higher moments can be found in Appendix B. Similar trends are seen in the $n = 2$ and 3 moments, although the deviation of the correction factors from 1 tends to get larger at low energies for higher moments. These corrections have been compared with those calculated by the DELPHI collaboration in Fig. 3 of Ref. [35]. The factors in Fig. 2.1 show the same energy dependence as in Ref. [35], although our correction factors deviate further from 1 at low energies. This is because we have corrected for c decays as well as b decays. When only b decays are taken into consideration the corrections we calculated using HERWIG++ agree well with those that the DELPHI collaboration calculated using PYTHIA 6.1.

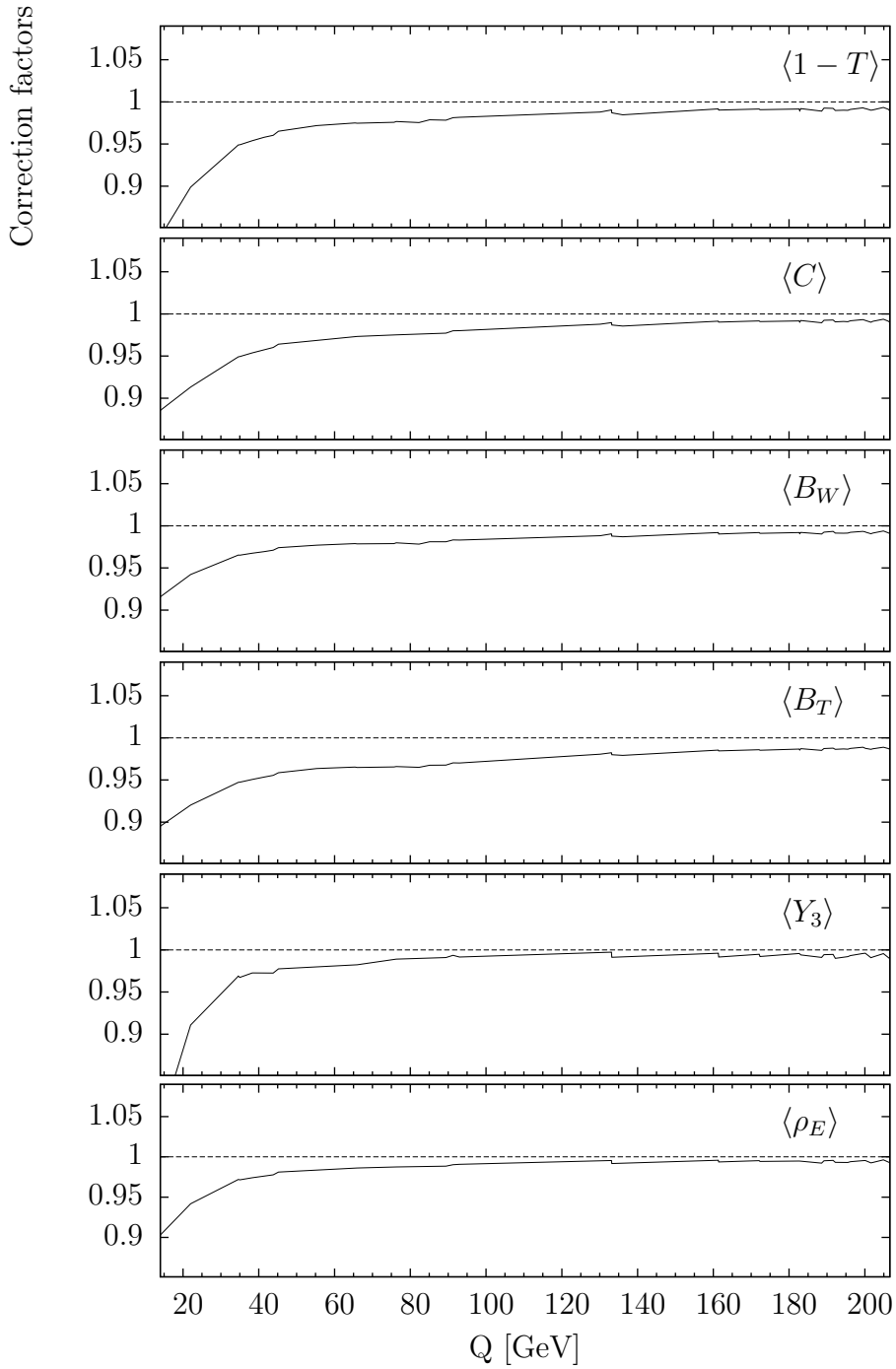


Figure 2.1: Heavy quark mass correction factors for the event shape means.

Chapter 3

Applying the ECH method to data

In the previous chapter we introduced the Effective Charges method, along with describing the e^+e^- event shape moment data. We now proceed to apply ECH to the data in order to extract α_s. The results are compared with the standard $\overline{\text{MS}}$ method in Section 3.1. In Section 3.2 we use the ECH method to give a prediction of the data. We discuss the results obtained in Section 3.3.

We move on, in the remaining sections, to look at the ECH method used in conjunction with a Padé Approximant (Section 3.4) and non-perturbative power corrections (Sections 3.5 and 3.6).

The DELPHI collaboration performed a similar study [35], applying the ECH method and standard $\overline{\text{MS}}$ perturbation theory ($\overline{\text{MS}}$ PT) at NLO to event shape means. We replicate this analysis, and also extend the work to NNLO, using the NNLO perturbative coefficients for event shape moments calculated by Gehrmann-De Ridder *et al.* in Ref. [34]. In a further extension of the DELPHI collaboration's analysis, we study some higher moments of event shape observables (see Section 2.2 for details). As in Ref. [35] we also add power cor-

rections to these perturbative models to determine whether the estimation of $\alpha_s(M_Z)$ is improved. We use the dispersive power correction model, outlined in Refs. [6, 48, 49], and also a simple power correction model [35, 50].

The results in this chapter, with the exception of those in Section 3.4, are published in Ref. [1].

3.1 Finding $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ from data

Now that we have corrected data in the form of an effective charge (see Section 2.2) we can use the ECH method to extract $\alpha_s(M_Z)$. For every event shape moment, and at each energy, a value of $\tilde{\Lambda}_{\overline{\rm MS}}$ is obtained using (2.1.15) and (2.1.18). An overall value of $\tilde{\Lambda}_{\overline{\rm MS}}$ is then found by performing a weighted average over all the energies for a specific event shape moment. The corresponding values of $\alpha_s(M_Z)$ are calculated from the averages using (2.1.21) and (2.1.22). This is done at both NLO and NNLO.

Table 3.1 shows the weighted averages of $\tilde{\Lambda}_{\overline{\rm MS}}$ obtained from the means and the corresponding values of $\alpha_s(M_Z)$. These results are also plotted in the top two panels of Fig. 3.1.

For comparison, $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ are also extracted from the event shape means using $\overline{\rm MS}$ perturbation theory ($\overline{\rm MS}$ PT), by setting $\mu_{\rm R} = Q$ in equation (2.2.2) and fitting to data. The results are shown in Table 3.2 and the bottom half of Fig. 3.1. The first uncertainty on the $\overline{\rm MS}$ PT results is the experimental uncertainty combined in quadrature with the uncertainties from the perturbative coefficients $\overline{\mathcal{B}}_{y,n}$ and $\overline{\mathcal{C}}_{y,n}$ [34]. These small uncertainties arise from the numerical integration of the perturbative coefficients associated with the differential cross-section $\frac{d\sigma}{dy}$ in equation (2.2.1). The second is a theoretical scale uncertainty obtained by varying $\mu_{\rm R}$ between $\mu_{\rm R} = 0.5Q$ and $\mu_{\rm R} = 2Q$.

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$
	NLO		NNLO	
$\langle 1-T\rangle$	262 ± 4	0.1193 ± 0.0003	319 ± 6	0.1233 ± 0.0004
$\langle C\rangle$	269 ± 4	0.1198 ± 0.0003	300 ± 5	0.1221 ± 0.0003
$\langle B_W\rangle$	326 ± 3	0.1234 ± 0.0002	312 ± 4	0.1229 ± 0.0003
$\langle B_T\rangle$	265 ± 2	0.1196 ± 0.0002	337 ± 5	0.1244 ± 0.0003
$\langle Y_3\rangle$	176 ± 5	0.1126 ± 0.0004	211 ± 7	0.1157 ± 0.0006
$\langle \rho_E\rangle$	268 ± 5	0.1198 ± 0.0003	289 ± 7	0.1214 ± 0.0004

Table 3.1: Weighted averages for $\tilde{\Lambda}_{\overline{\rm MS}}$ and the corresponding values of $\alpha_s(M_Z)$ for the means using the ECH method. The uncertainties are a combination of those from experiment and the perturbative coefficients (see Section 3.1).

The largest difference between the calculations using the central value of $\mu_{\rm R}$ and the upper and lower variations is taken as the uncertainty. The error bars in Fig. 3.1 represent the combined experimental and perturbative coefficient uncertainty and the theoretical scale uncertainty added in quadrature. In contrast, the uncertainty on the ECH results is just the combined experimental and perturbative coefficient uncertainty. Results for the $n = 2$ moments are shown in Tables 3.3–3.4 and Fig. 3.2, and for the $n = 3$ moments in Tables 3.5–3.6 and Fig. 3.3.

3.2 Predicting event shape moments from a fixed value of $\alpha_s(M_Z)$

Another way to implement the ECH method is to work backwards from a fixed value of $\alpha_s(M_Z)$ to get an approximation to the data. In this section we fix $\alpha_s(M_Z)$ at the result extracted from next-to-NNLO ($\rm N^3LO$) calculations on Z-decays [51]: $\alpha_s(M_Z) = 0.1190 \pm 0.0026$. This value is close to the world

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	χ^2/dof
		NLO	
$\langle 1-T\rangle$	$646 \pm 9 \pm 279$	$0.1385 \pm 0.0003 \pm 0.0107$	52.3/47
$\langle C\rangle$	$597 \pm 7 \pm 247$	$0.1365 \pm 0.0003 \pm 0.0099$	45.7/41
$\langle B_W\rangle$	$336 \pm 3 \pm 105$	$0.1241 \pm 0.0002 \pm 0.0062$	39.5/47
$\langle B_T\rangle$	$429 \pm 3 \pm 147$	$0.1291 \pm 0.0001 \pm 0.0073$	69.7/47
$\langle Y_3\rangle$	$268 \pm 5 \pm 81$	$0.1197 \pm 0.0003 \pm 0.0056$	61.5/32
$\langle \rho_E\rangle$	$340 \pm 6 \pm 90$	$0.1243 \pm 0.0004 \pm 0.0053$	11.5/14
		NNLO	
$\langle 1-T\rangle$	$396 \pm 7 \pm 77$	$0.1277 \pm 0.0004 \pm 0.0041$	51.1/47
$\langle C\rangle$	$378 \pm 6 \pm 69$	$0.1267 \pm 0.0003 \pm 0.0038$	44.3/41
$\langle B_W\rangle$	$317 \pm 12 \pm 13$	$0.1232 \pm 0.0007 \pm 0.0008$	37.8/47
$\langle B_T\rangle$	$347 \pm 9 \pm 36$	$0.1250 \pm 0.0005 \pm 0.0021$	78.1/47
$\langle Y_3\rangle$	$228 \pm 5 \pm 18$	$0.1171 \pm 0.0004 \pm 0.0014$	65.1/32
$\langle \rho_E\rangle$	$295 \pm 7 \pm 21$	$0.1218 \pm 0.0005 \pm 0.0014$	11.5/14

Table 3.2: $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ for the means using $\overline{\rm MS}$ PT. The first error is experimental and the second is theoretical.

average value of $\alpha_s(M_Z) = 0.1184 \pm 0.0031$ reported in Ref. [52]. The N^{3}LO Z-decay value of $\alpha_s(M_Z)$ corresponds to $\tilde{\Lambda}_{\overline{\rm MS}} = 254$ MeV. We can then find $\Lambda_{\mathcal{R}}$ using (2.1.18) and hence $\mathcal{R}(Q)$ by inverting (2.1.15). At NLO, where $\mathcal{G}(\mathcal{R}) = 1$, this can be done analytically:

$$\mathcal{R}_{\rm NLO}(Q) = \frac{-1}{c\left[1 + W\left(-\exp\left[-\left(1 + \frac{b}{c}\ln\frac{Q}{\Lambda_{\mathcal{R}}}\right)\right]\right)\right]}. \tag{3.2.1}$$

Note that if we rewrite $\Lambda_{\mathcal{R}}$ in terms of $\tilde{\Lambda}_{\overline{\rm MS}}$ using (2.1.18) and compare it with (2.1.21) we find that $\mathcal{R}_{\rm NLO}(Q) = \alpha_s(\mu_{\rm ECH})/\pi$. The particular choice of $\overline{\rm MS}$ renormalisation scale of $\mu_{\rm R} = \mu_{\rm ECH} = e^{-r/b}\,Q$ is equivalent to the ECH

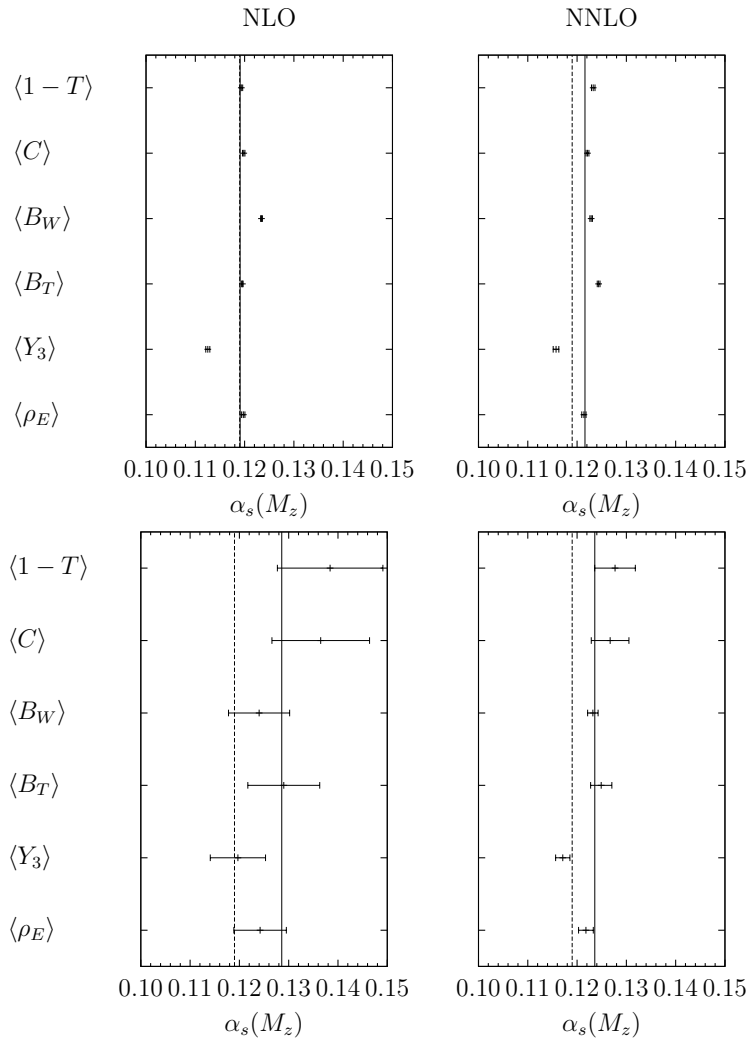


Figure 3.1: Scatter plots showing the values of $\alpha_s(M_Z)$ obtained from the means of the event shapes using the ECH method (top two panels) and $\overline{\text{MS}}$ PT (bottom two panels). The dotted lines show the value of the coupling obtained from N^{3}LO calculations on Z-decays: $\alpha_s(M_Z) = 0.1190$. The solid lines show the unweighted average for the NLO and NNLO cases.

scheme. This ECH scale corresponds to absorbing radiative corrections into the definition of the coupling.

An NLO ECH approximation to data is then given by $\langle y^n\rangle_{\text{NLO}} = \frac{\overline{A}_{y,n}}{2}\mathcal{R}_{\text{NLO}}$, using (2.2.3). At NNLO it is not possible to invert (2.1.15) analytically and so this is done numerically using Mathematica 7.0 [53].

The NLO and NNLO ECH approximations are shown against the data in

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$
	NLO		NNLO	
$\langle(1-T)^2\rangle$	130 ± 4	0.1079 ± 0.0005	438 ± 22	0.1298 ± 0.0011
$\langle C^2\rangle$	123 ± 3	0.1071 ± 0.0004	466 ± 30	0.1312 ± 0.0014
$\langle B_W^2\rangle$	166 ± 5	0.1116 ± 0.0005	203 ± 7	0.1151 ± 0.0006
$\langle B_T^2\rangle$	0.05 ± 0.02	0.0520 ± 0.0017	287 ± 34	0.1213 ± 0.0023
$\langle Y_3^2\rangle$	32 ± 13	0.0905 ± 0.0041	38 ± 16	0.0924 ± 0.0045

Table 3.3: $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ for the $n = 2$ moments using the ECH method.

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	χ^2/dof
	NLO		
$\langle(1-T)^2\rangle$	$1047 \pm 30 \pm 607$	$0.1515 \pm 0.0008 \pm 0.0172$	23.6/32
$\langle C^2\rangle$	$1048 \pm 29 \pm 609$	$0.1516 \pm 0.0008 \pm 0.0173$	28.3/26
$\langle B_W^2\rangle$	$245 \pm 9 \pm 76$	$0.1181 \pm 0.0007 \pm 0.0055$	22.5/32
$\langle B_T^2\rangle$	$775 \pm 15 \pm 490$	$0.1431 \pm 0.0005 \pm 0.0167$	26.9/32
$\langle Y_3^2\rangle$	$184 \pm 29 \pm 63$	$0.1132 \pm 0.0026 \pm 0.0056$	45.5/17
	NNLO		
$\langle(1-T)^2\rangle$	$467 \pm 13 \pm 147$	$0.1312 \pm 0.0006 \pm 0.0070$	23.1/32
$\langle C^2\rangle$	$469 \pm 13 \pm 147$	$0.1313 \pm 0.0006 \pm 0.0070$	26.4/26
$\langle B_W^2\rangle$	$207 \pm 8 \pm 19$	$0.1154 \pm 0.0006 \pm 0.0015$	22.5/32
$\langle B_T^2\rangle$	$370 \pm 7 \pm 110$	$0.1263 \pm 0.0004 \pm 0.0061$	26.2/32
$\langle Y_3^2\rangle$	$147 \pm 23 \pm 16$	$0.1099 \pm 0.0024 \pm 0.0017$	46.9/17

Table 3.4: $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ for the $n = 2$ moments using $\overline{\rm MS}$ PT. The first error is experimental and the second is theoretical.

Figs 3.4-3.9. For comparison the $\overline{\rm MS}$ PT predictions are also shown. These are obtained by running the coupling from the fixed $\alpha_s(M_Z)$ over a range of energies, using (2.1.21) and (2.1.22) with $\tilde{\Lambda}_{\overline{\rm MS}} = 254$ MeV, at NLO or NNLO as appropriate. The data are then approximated using (2.2.2).

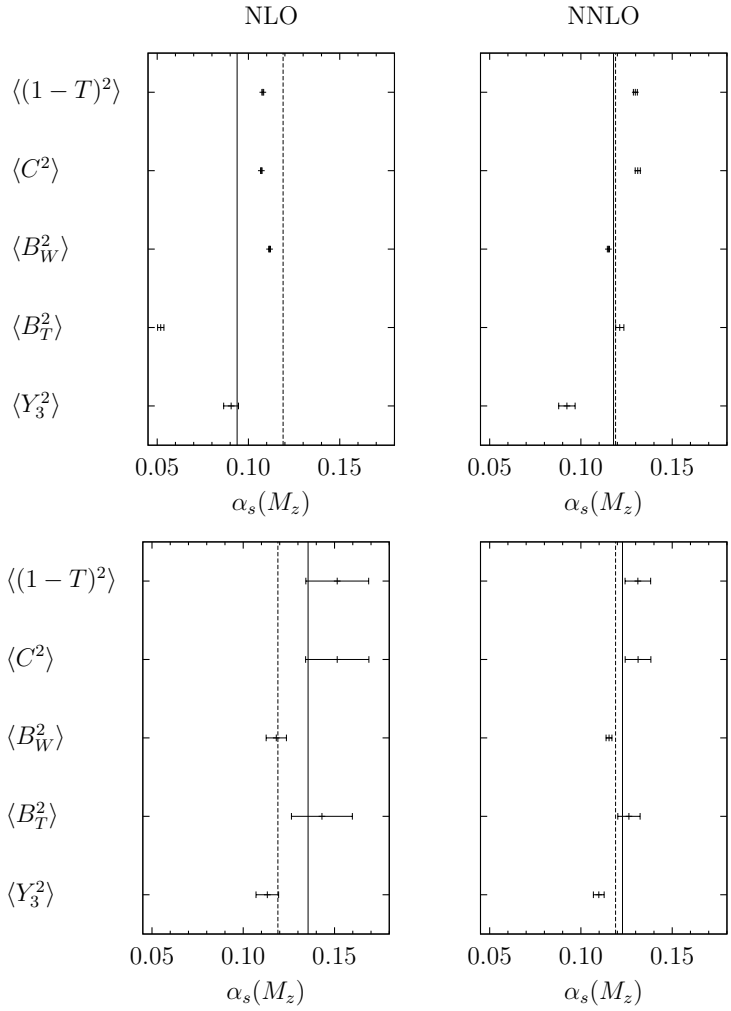


Figure 3.2: Scatter plots showing the values of $\alpha_s(M_Z)$ obtained from the $n = 2$ moments of the event shapes using the ECH method (top two panels) and $\overline{\rm MS}$ PT (bottom two panels). The solid and dotted lines are as in Fig 3.1.

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$
	NLO		NNLO	
$\langle(1-T)^3\rangle$	90 $\pm$ 5	0.1027 ± 0.0007	647 ± 37	0.1389 ± 0.0014
$\langle C^3\rangle$	73 $\pm$ 3	0.0999 ± 0.0006	344 ± 57	0.1248 ± 0.0033
$\langle B_W^3\rangle$	117 ± 6	0.1063 ± 0.0007	152 ± 7	0.1104 ± 0.0007
$\langle B_T^3\rangle$	7.5 ± 0.5	0.0776 ± 0.0005	49 $\pm$ 8	0.0953 ± 0.0018
$\langle Y_3^3\rangle$	0.6 ± 1.4	0.0620 ± 0.0122	0.6 ± 1.6	0.0624 ± 0.0131

Table 3.5: $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ for the $n = 3$ moments using the ECH method

Observable	$\tilde{\Lambda}_{\overline{\rm MS}}$ [MeV]	$\alpha_s(M_Z)$	χ^2/dof
		NLO	
$\langle(1-T)^3\rangle$	$1249 \pm 35 \pm 780$	$0.1570 \pm 0.0009 \pm 0.0200$	7.4/17
$\langle C^3\rangle$	$1316 \pm 33 \pm 844$	$0.1587 \pm 0.0008 \pm 0.0210$	9.4/17
$\langle B_W^3\rangle$	$177 \pm 10 \pm 56$	$0.1126 \pm 0.0009 \pm 0.0051$	8.2/17
$\langle B_T^3\rangle$	$976 \pm 21 \pm 692$	$0.1495 \pm 0.0006 \pm 0.0205$	9.1/17
$\langle Y_3^3\rangle$	$163 \pm 40 \pm 55$	$0.1113 \pm 0.0039 \pm 0.0053$	67.2/17
		NNLO	
$\langle(1-T)^3\rangle$	$493 \pm 14 \pm 176$	$0.1324 \pm 0.0006 \pm 0.0081$	7.2/17
$\langle C^3\rangle$	$526 \pm 14 \pm 191$	$0.1340 \pm 0.0006 \pm 0.0084$	10.0/17
$\langle B_W^3\rangle$	$151 \pm 8 \pm 13$	$0.1103 \pm 0.0008 \pm 0.0013$	8.5/17
$\langle B_T^3\rangle$	$400 \pm 9 \pm 141$	$0.1279 \pm 0.0005 \pm 0.0074$	9.7/17
$\langle Y_3^3\rangle$	$135 \pm 33 \pm 13$	$0.1086 \pm 0.0037 \pm 0.0015$	68.5/17

Table 3.6: $\tilde{\Lambda}_{\overline{\rm MS}}$ and $\alpha_s(M_Z)$ for the $n = 3$ moments using $\overline{\rm MS}$ PT. The first error is experimental and the second is theoretical.

It should be noted that when ECH is applied at NLO to the higher moments of $1-T$, C, and B_T, some imaginary values are obtained as predictions of the observables at (generally) low energies. In the case of $\langle(1-T)^2\rangle$ and $\langle C^2\rangle$ this is only for energies of less than 15 GeV. For $\langle(1-T)^3\rangle$ there are problematic results up to energies of ~ 25 GeV, and results for $\langle C^3\rangle$ are well behaved above ~ 35 GeV. B_T is the worst behaved variable, with imaginary predictions up to Q ~ 60 GeV for the $2^{\rm nd}$ moment, and across the whole energy range for $n = 3$. This behaviour is discussed further in Section 3.3.1.

3.3 Commentary on the results

The values of $\alpha_s(M_Z)$ extracted using the ECH method in Section 3.1 agree well with those found by the DELPHI collaboration in Ref. [35]. They are well

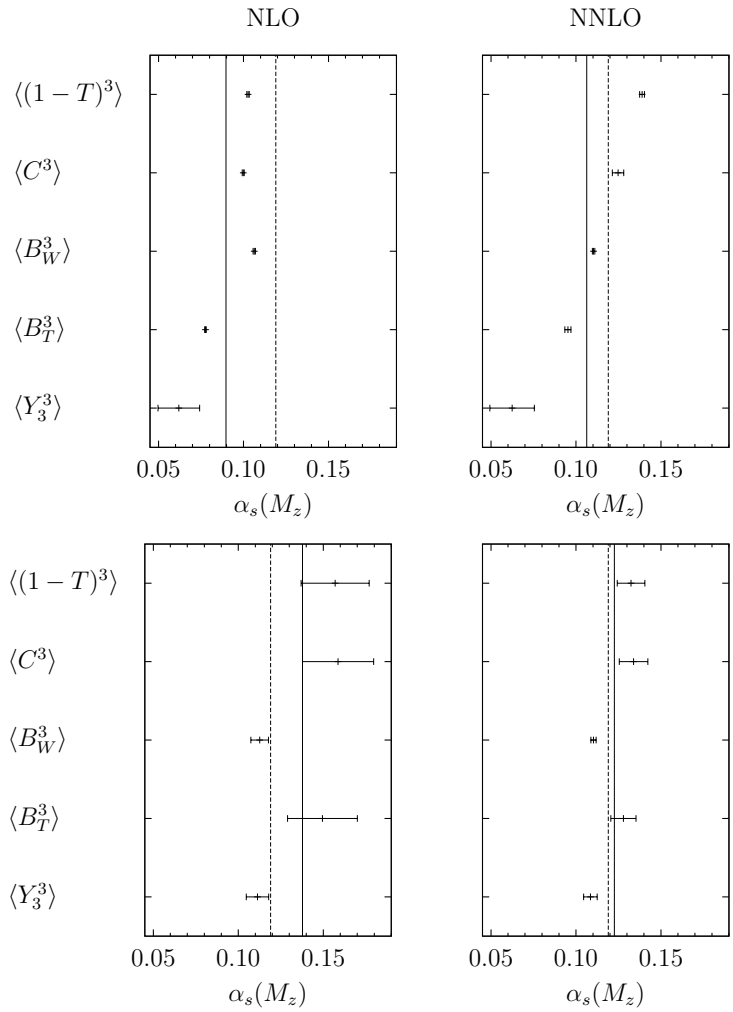


Figure 3.3: Scatter plots showing the values of $\alpha_s(M_Z)$ obtained from the $n = 3$ moments of the event shapes using the ECH method (top two panels) and $\overline{\text{MS}}$ PT (bottom two panels). The solid and dotted lines are as in Fig 3.1.

grouped, as can be seen in Fig. 3.1 (top left panel); the values for B_W and Y_3 are slight outliers. The average for NLO ECH agrees very well with the N^{3}LO Z-decay calculation value [51]. The average for the $\overline{\text{MS}}$ perturbation theory ($\overline{\text{MS}}$ PT) calculation at NLO (bottom left panel) sits further from the N^{3}LO calculation, and the scatter of the central values is larger than for ECH at NLO. The numerical values of $\alpha_s(M_Z)$ are given in Tables 3.1 and 3.2.

We also demonstrated, in Section 3.2, how the ECH method can be inverted to give an approximation to data. Figs. 3.4–3.9 show that the data for the

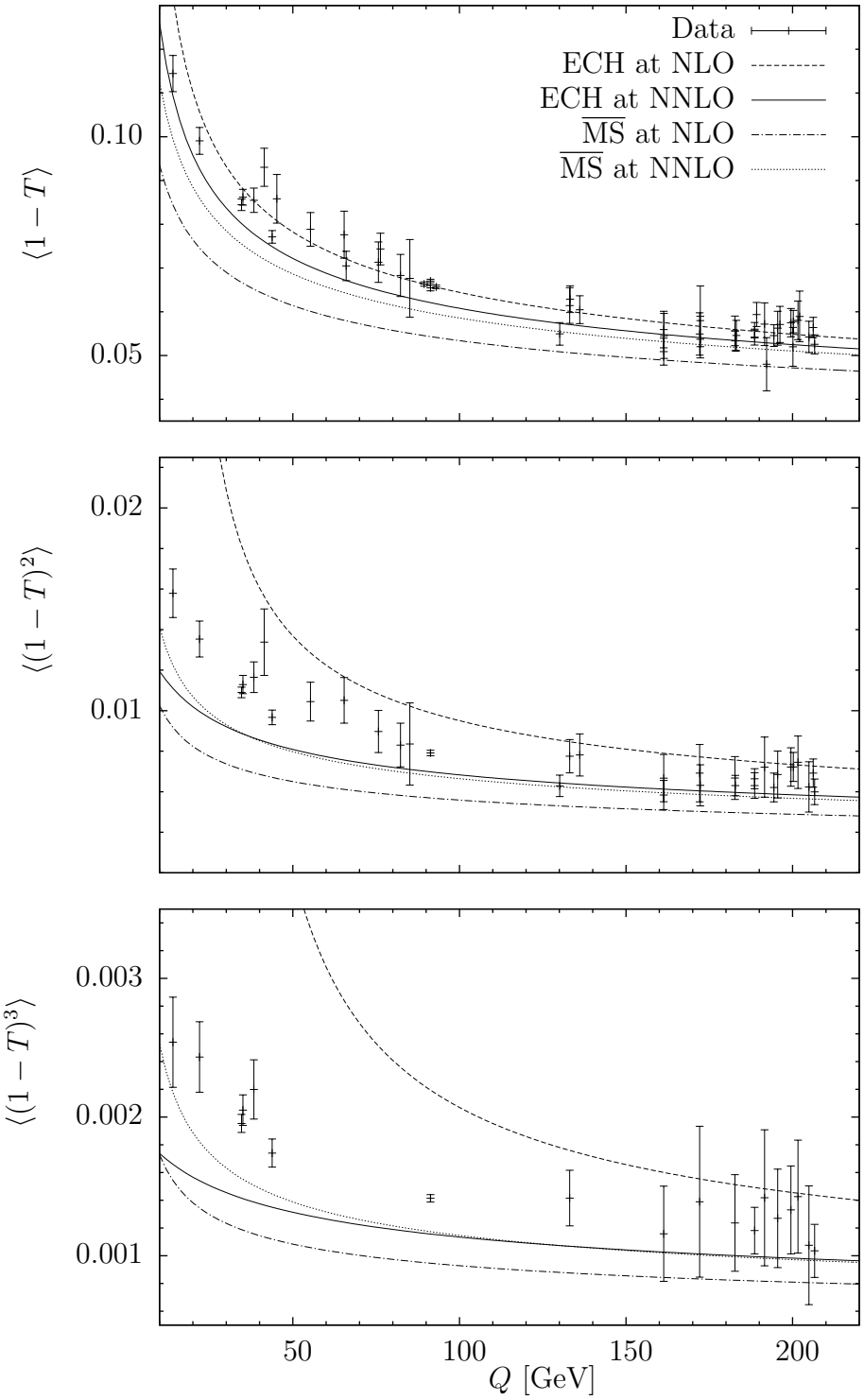


Figure 3.4: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle (1-T)^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

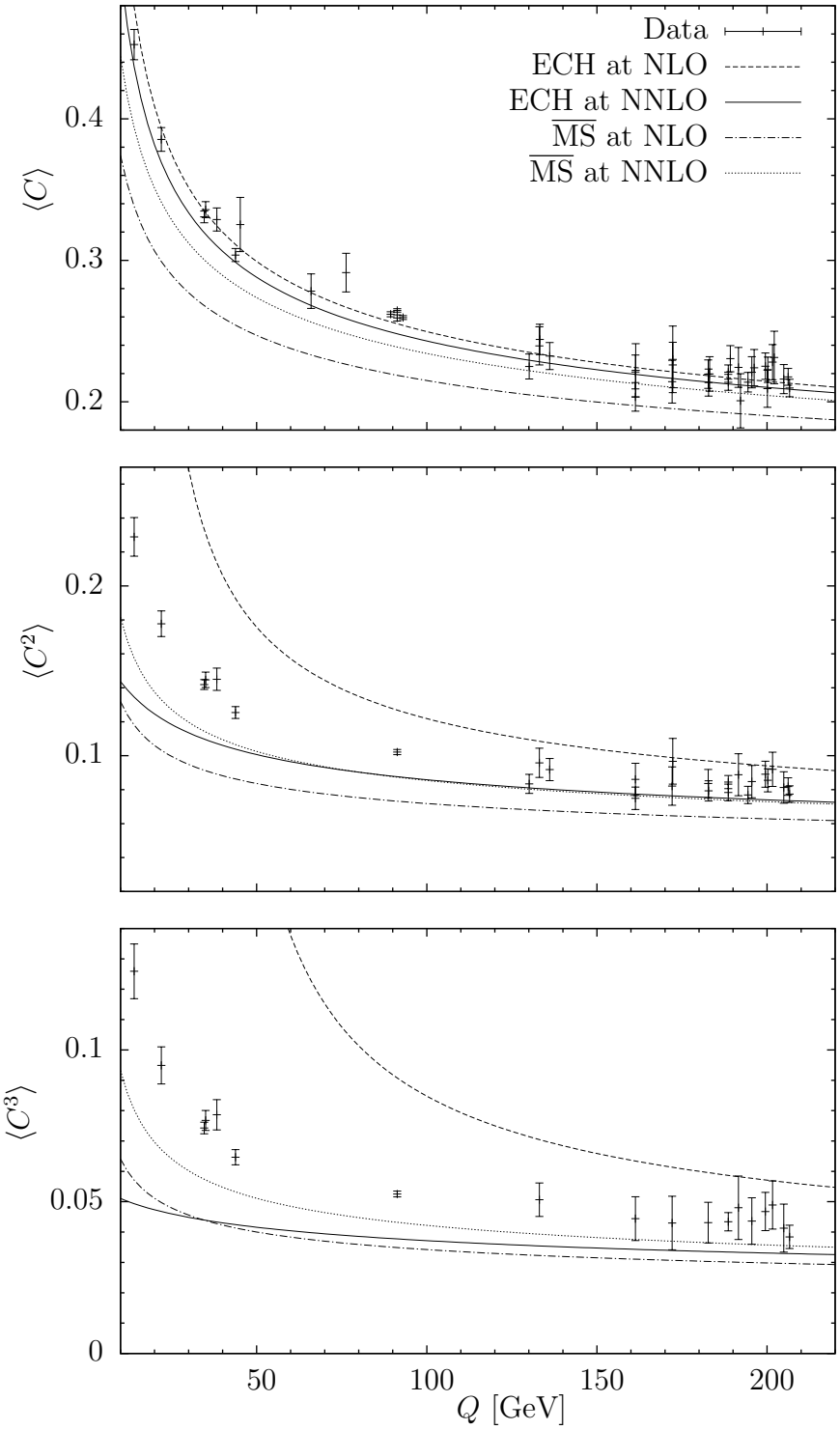


Figure 3.5: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle C^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

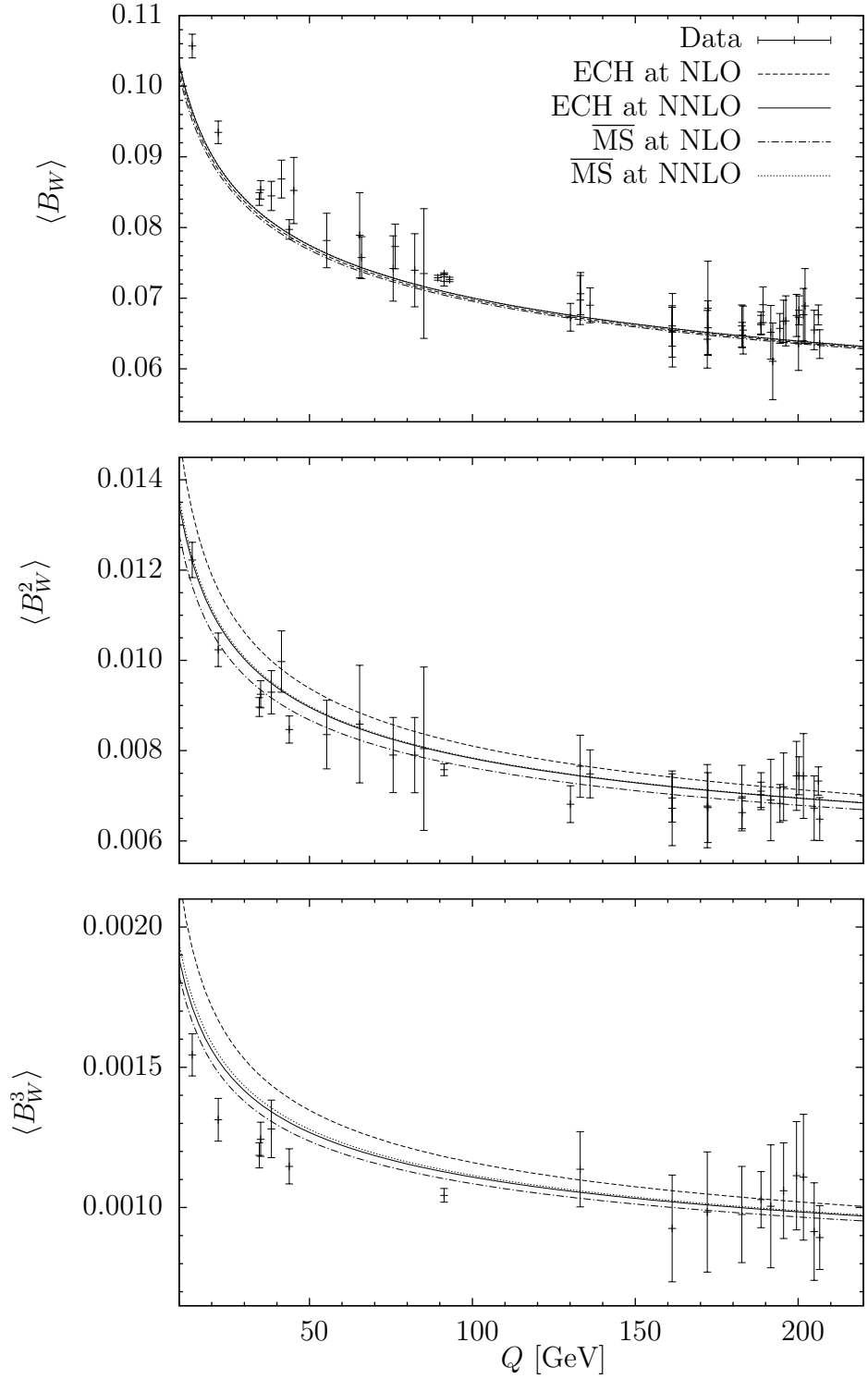


Figure 3.6: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle B_W^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

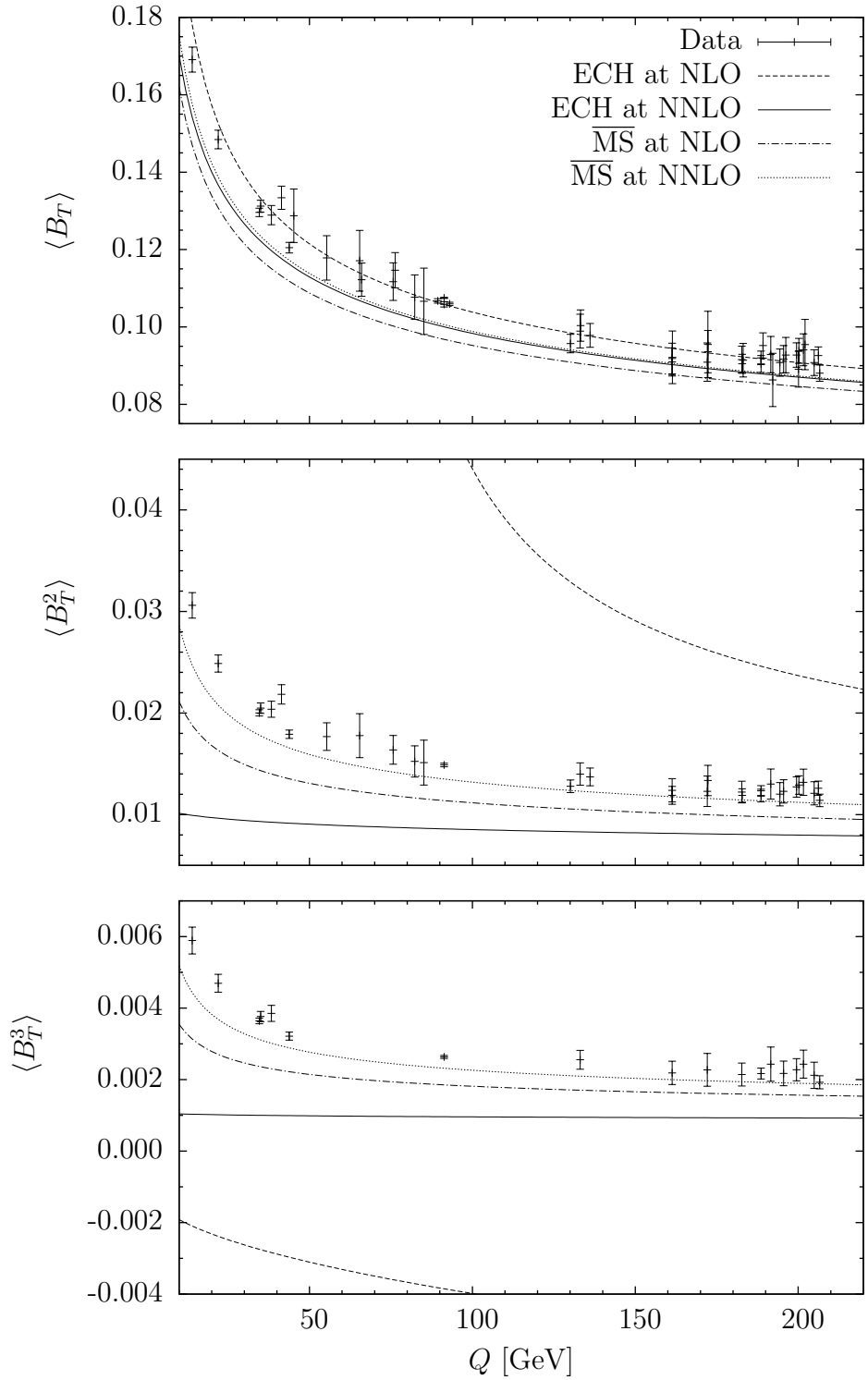


Figure 3.7: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle B_T^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

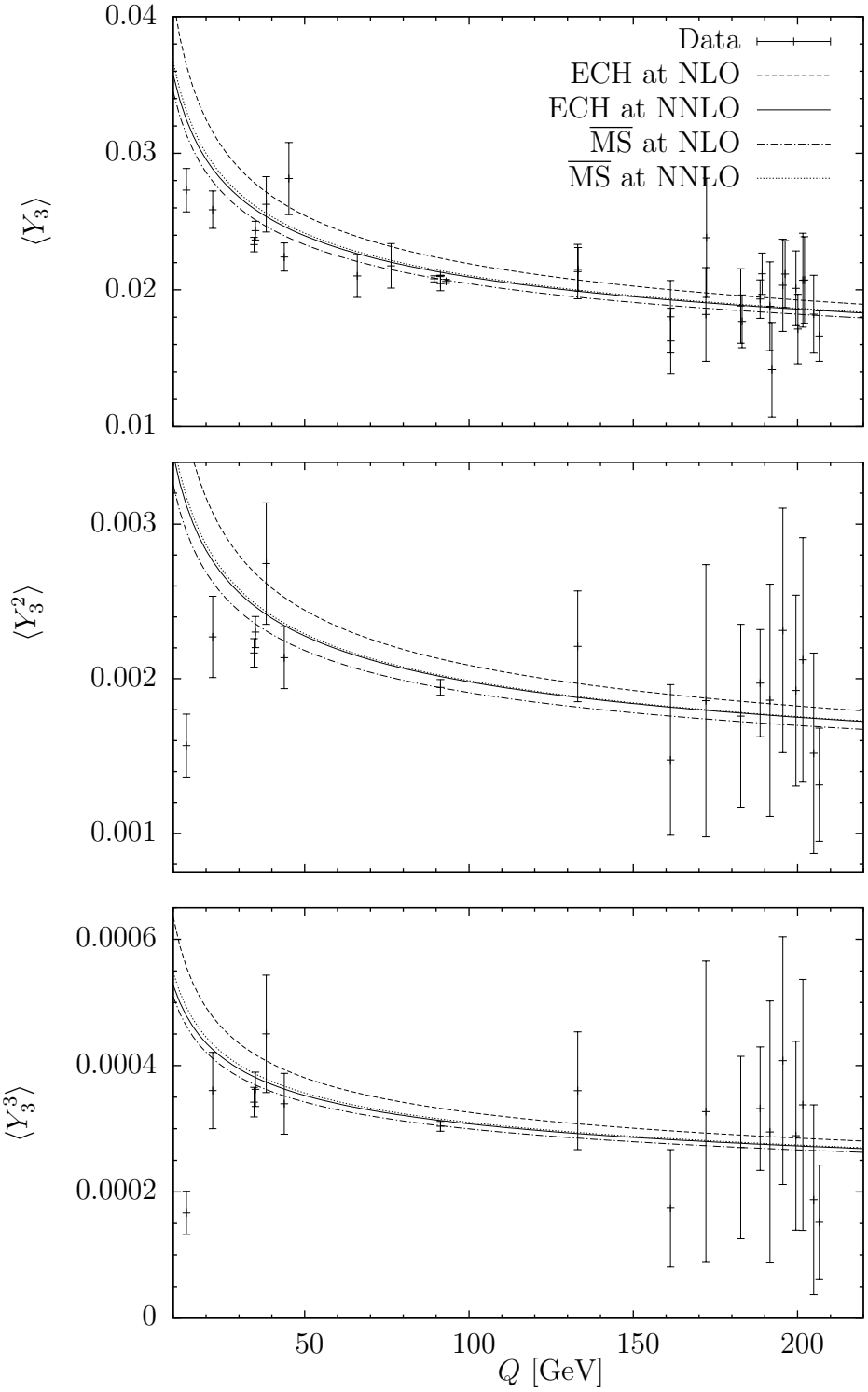


Figure 3.8: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle Y_3^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

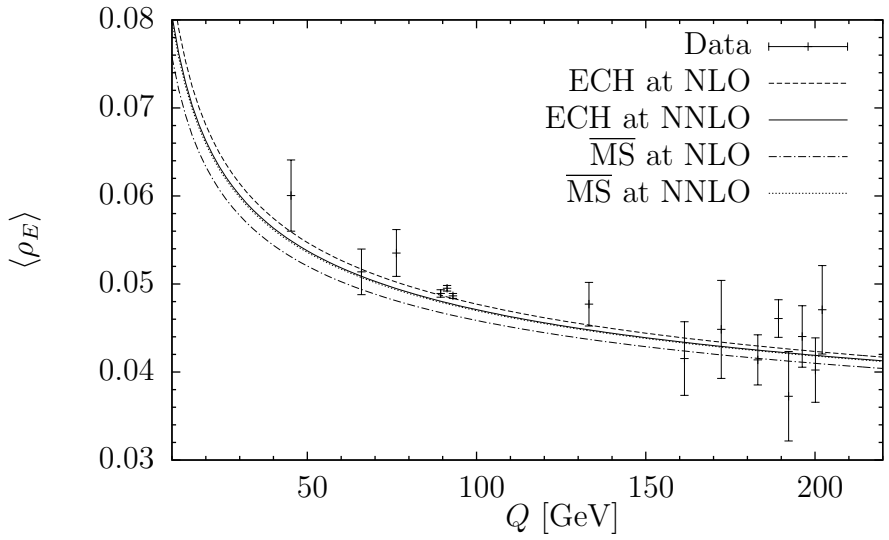


Figure 3.9: ECH and $\overline{\text{MS}}$ PT approximations to data at NLO and NNLO for $\langle \rho_E \rangle$.

means are best described by the NLO ECH curves, which lie closer to data than both the NLO and NNLO $\overline{\text{MS}}$ PT curves. The exceptions to this are $\langle B_W \rangle$, for which all the perturbative techniques give a very similar description of the data, and $\langle Y_3 \rangle$, where the $\overline{\text{MS}}$ PT at NLO curve describes the data best.

However the ECH method at NLO works less well for higher moments. This is particularly the case for $\langle (1-T)^n \rangle$, $\langle C^n \rangle$ and $\langle B_T^n \rangle$ with $n = 2, 3$, as can be seen in the bottom two panels of Figs. 3.4, 3.5 and 3.7.[1] $\overline{\text{MS}}$ PT also does not do a good job at describing the higher moment data.

ECH at NNLO describes the data less well than ECH at NLO, as shown in Section 3.2. The exception to this observation is $\langle Y_3 \rangle$, in Fig. 3.8, for which the NNLO ECH curve lies closer to the data than that for NLO ECH. Also, there is very little difference between any of the curves for $\langle B_W \rangle$. The NNLO ECH curves tend to give a description of the means that is similar to or better than the $\overline{\text{MS}}$ PT curves. For the higher moments ECH at NNLO is again often

[1]There are some energies where ECH at NLO predicts unphysical imaginary values for these observables — this is discussed in the next subsection.

very similar to the $\overline{\text{MS}}$ PT curves, although it sometimes fails to describe the behaviour of the low energy data as well as $\overline{\text{MS}}$ PT does (e.g. $\langle C^2 \rangle$ in Fig. 3.5).

In the top right panel of Fig. 3.1, we see that the values of $\alpha_s(M_Z)$ are still fairly well grouped when extracted using ECH at NNLO, with $\langle Y_3 \rangle$ remaining an outlier, but that the average has moved away from the N^{3}LO result. $\overline{\text{MS}}$ PT at NNLO, shown in the bottom right panel, is seen to have improved from the $\overline{\text{MS}}$ PT at NLO case: there are smaller errors due to a reduced scale dependence at higher order, the grouping is improved and the average is closer to the N^{3}LO calculation. In the higher moment plots (Figs. 3.2 and 3.3) the grouping in the ECH method gets worse whereas the improvement seen going from NLO to NNLO in $\overline{\text{MS}}$ PT continues to hold.

These results confirm the success of NLO ECH in providing an accurate description of event shape means, as was noted in the analysis of Ref. [35]. There is generally no need to add a non-perturbative power correction. However, for $\overline{\text{MS}}$ PT at NLO additional power corrections are often necessary; these can be obtained using a dispersive model of power corrections [6, 48, 49]. In Section 3.6, we will discuss the results of fits using this dispersive model. A new feature of these fits is that the ECH method is used in conjunction with the dispersive power corrections.

Before exploring power correction models, we attempt to explain why the ECH method at NNLO does not provide a good description of the data, even for event shape means. We also try to motivate why NLO ECH does not give an accurate description of the higher moments. We do this by studying the relative sizes of terms in the asymptotic series expansion of $\rho(\mathcal{R})$, and examining the size of $\Lambda_{\mathcal{R}}$ for the variables, as shown in the next section.

3.3.1 Terms in the expansion of $\rho(\mathcal{R})$ and the size of $\Lambda_{\mathcal{R}}$ values

In order to attempt to explain the results in Sections 3.1 and 3.2, we now look at the expansions $\rho(\mathcal{R})$. From Section 2.1 we have:

$$\rho(\mathcal{R}) \equiv \frac{d\mathcal{R}(Q)}{d\ln Q} = -b\mathcal{R}^2(1 + c\mathcal{R} + \rho_2\mathcal{R}^2 + \dots). \tag{3.3.1}$$

Perturbative expansions in QCD and other quantum field theories are expected to be asymptotic series, with an $n!$ growth of coefficients in n^{th} order perturbation theory. This divergent large-order behaviour arises from renormalon diagrams involving chains of vacuum polarisation bubbles in QED (for a review see Ref. [54]). The large-order behaviour of the $\rho(\mathcal{R})$ function for the inclusive QCD $R_{e^+e^-}$ ratio has been investigated using renormalon techniques in Ref. [55], and has been shown to be an asymptotic series.

A series that is asymptotic to a function, $F(a)$, has the following property (see Ref. [56] for a discussion of asymptotic series):

$$F(a) \approx \sum_{n=0}^{\infty} f_n a^n \qquad \text{if} \qquad \left| F(a) - \sum_{n=0}^{N} f_n a^n \right| \leq \left| f_{N+1} a^{N+1} \right|, \tag{3.3.2}$$

where $\approx$ indicates that the series $\sum f_n a^n$ is asymptotic to the function $F(a)$. An important feature of (3.3.2) is that the error associated with terminating the series at a particular point (after N terms) is bounded by the size of the first term neglected (the $(N+1)^{\text{th}}$ term in this case). To get the best approximation to $F(a)$ we therefore want to terminate the series just before the term of smallest absolute magnitude. This is a very important property of asymptotic series; although they are divergent, with zero radius of convergence, one can still obtain accurate approximations by truncating them after a suitable

Observable	n	$c\mathcal{R}$ (NLO)	$\rho_2\mathcal{R}^2$ (NNLO)	Order at which series should be terminated
$\langle(1-T)^n\rangle$	1	0.078	−0.049	NLO
	2	0.105	−0.323	LO
	3	0.119	−0.575	LO
$\langle C^n\rangle$	1	0.076	−0.030	NLO
	2	0.106	−0.376	LO
	3	0.123	−0.912	LO
$\langle B_W^n\rangle$	1	0.045	0.004	NLO
	2	0.057	−0.031	NLO
	3	0.055	−0.039	NLO
$\langle B_T^n\rangle$	1	0.066	−0.060	NLO
	2	0.111	−1.803	LO
	3	0.140	−5.715	LO
$\langle Y_3^n\rangle$	1	0.058	−0.042	NLO
	2	0.060	−0.049	NLO
	3	0.060	−0.053	NLO
$\langle \rho_E^n\rangle$	1	0.059	−0.015	NLO

Table 3.7: Size of terms in $\frac{\rho(\mathcal{R})}{-b\mathcal{R}^2}$

number of terms.

In Table 3.7 we show the relative sizes of the known terms in $\frac{\rho(\mathcal{R})}{-b\mathcal{R}^2}$ for the event shape moments: $c\mathcal{R}$, the NLO term, and $\rho_2\mathcal{R}^2$, the NNLO term. These numbers are calculated using data from the M_Z peak. The leading order (LO) term in $\frac{\rho(\mathcal{R})}{-b\mathcal{R}^2}$, 1, is universal. The last column indicates the order at which the series should be terminated based on the relative magnitudes of the currently known terms. Note that, where $\rho_2\mathcal{R}^2$ is the smallest term, it is not possible to say definitively whether the series should be terminated at NLO or NNLO without knowing the size of the N^{3}LO term.

This information can help explain the results discussed in Section 3.3. For example, the expansions for the higher moments of $1-T$, C and B_T all suggest that it is best to terminate the series at leading order. The energy dependence of $\mathcal{R}(Q)$ is controlled by the RS invariant dimensionful constant $\Lambda_{\mathcal{R}}$ of (2.1.18) which involves the NLO coefficient $r_1^{\overline{\rm MS}}(\mu_{\rm R} = Q)$. This implies that LO truncation is not physically meaningful. Therefore using perturbation theory with this type of expansion, where the observable itself is used as the expansion parameter, is inappropriate for these event shape moments. As discussed previously, it can be seen that applying the ECH method to the higher moments of these observables does not work well. However, in some instances (e.g. $\langle(1-T)^2\rangle$ in Fig. 3.4) the ECH method at NLO appears to be converging with the data at high energies. This implies that we are examining data below the perturbative regime of the method for these variables, but that if we go to sufficiently high energies the ECH method may be found to describe the data again. As Q increases $\mathcal{R}(Q)$ decreases; for a suitably high energy the NLO truncation will therefore represent the optimal number of terms, and will successfully describe the data. Of course this is of theoretical rather than practical interest, since we do not have data at indefinitely large energies.

The expansions for all other variables in Table 3.7 suggest that it is preferable to terminate at NLO. As discussed in Section 3.3, it appears that the ECH method generally works better at NLO than at NNLO, although a firm conclusion cannot be made in these cases since we have no information on the size of the N^{3}LO term. We can also see that the higher moments of B_W and Y_3 do not display the same undesirable behaviour as higher moments of the other variables: see the bottom two panels of Figs. 3.6 and 3.8.

We also examine the size of $\Lambda_{\mathcal{R}}$ for each variable. The values of $\Lambda_{\mathcal{R}}$, shown

in Table 3.8, are calculated from a fixed value of $\alpha_s(M_Z) = 0.1190$ [2] and (2.1.18), which when inverted gives $\Lambda_{\mathcal{R}} = e^{-r/b}\,\tilde{\Lambda}_{\overline{\rm MS}}$. From (2.1.11) we see that $\mathcal{R}(Q)$ has a Landau pole when $Q = \Lambda_{\mathcal{R}}$, and will formally diverge there [2]. From Table 3.8, the $\Lambda_{\mathcal{R}}$ values for the $n = 1$ moments are a few GeV. Therefore the data we are using are at energies considerably higher than the Landau pole and the ECH method should work well. For the higher moments, the values of $\Lambda_{\mathcal{R}}$ increase. We are therefore analysing data in a region of Q which contains the Landau pole, and the perturbative ECH method will not function properly until we have data at much larger energies. This is the same conclusion that we reached by studying the magnitudes of terms in the $\rho(\mathcal{R})$ series. The $\langle B_W^n \rangle$ and $\langle Y_3^n \rangle$ moments for $n = 2, 3$ are an exception, with small values of $\Lambda_{\mathcal{R}}$. As stated above, the ECH method performs better for these higher moments compared with the other observables.

These values of $\Lambda_{\mathcal{R}}$ also explain the unphysical imaginary predictions seen at some energies for the higher moments of $1-T$, C and B_T in Section 3.2. In ECH at NLO imaginary predictions are produced for energies that are lower than the Landau pole — for example, at $Q < 15$ GeV for $\langle (1-T)^2 \rangle$, where $\Lambda_{\mathcal{R}}$ is ~ 15 GeV, and across the whole energy region studied for $\langle B_T^3 \rangle$, where $\Lambda_{\mathcal{R}}$ is almost 400 GeV. This is because $Q/\Lambda_{\mathcal{R}}$ is less than 1, and therefore the argument of the Lambert-W function in (3.2.1) is less than $-1/e$. Below this value of z, $W_{-1}(z)$ gives imaginary values. This behaviour is not seen in ECH at NNLO, although this method still does not give good predictions in the region outside of the perturbative regime.

[2] or $\tilde{\Lambda}_{\overline{\rm MS}} = 254$ MeV,

n	$\langle(1-T)^n\rangle$	$\langle C^n\rangle$	$\langle B_W^n\rangle$	$\langle B_T^n\rangle$	$\langle Y_3^n\rangle$	$\langle\rho_E^n\rangle$
1	3	3	0.1	2	1	1
2	14	15	1.1	58	2	-
3	27	34	1.2	384	1	-

Table 3.8: Values of $\Lambda_{\mathcal{R}}$ in GeV

3.4 Padé approximants

In this section we apply the ECH method in conjunction with Padé approximant methods. We start by introducing Padé approximants and describing how they can be used, and then use the methods together to give a prediction of event shape moment data in Section 3.4.2.

A Padé Approximant consists of a ratio of two polynomials [29, 57–60]. A power series, S:

$$S = S_0 + S_1 x + S_2 x^2 + \ldots + S_{N+M} x^{N+M} + \ldots \tag{3.4.1}$$

has a Padé approximant defined as:

$$[N/M] = \frac{a_0 + a_1 x + \ldots + a_N x^N}{1 + b_1 x + \ldots + b_M x^M}, \tag{3.4.2}$$

where the coefficients a_n and b_n are chosen such that $[N/M]$ approximates the series S up to a term of order x^{N+M+1}:

$$[N/M] = S + \mathcal{O}(x^{N+M+1}). \tag{3.4.3}$$

The Padé approximant (PA) can either be used to estimate the coefficient of the next term in the power series, S_{N+M+1}, or as an estimate of the total sum

of S. We describe this latter case as using the Padé Sum (PS).

We briefly note two properties of PAs:

1. The [N / N] diagonal Padé approximants will converge if the series they represent is a Stieltjes series — for a proof and more details of Stieltjes series see Ref. [56].

2. Diagonal PAs to $\mathcal{R}(Q)$ are RS invariant in the 't Hooft scheme $c_2 = c_3 = \ldots = 0$ and with the coefficient c also set to zero [11].

3.4.1 Using Padé Approximants to estimate ρ_3

In this section we show how to get an estimate for ρ_3 in (3.3.1). We start by rewriting the expression for $\rho(\mathcal{R})$ in the form of S in (3.4.1):

$$-\frac{\rho(\mathcal{R})}{b\mathcal{R}^2} = 1 + c\mathcal{R} + \rho_2\mathcal{R}^2 + \rho_3\mathcal{R}^3 + \ldots. \tag{3.4.4}$$

We know the coefficients in the series up to ρ_2 and we want to use the PA methods outlined above to get an estimate of ρ_3. Comparing with (3.4.1) we see that $N + M = 2$. We can therefore construct two PAs: [0/2] and [1/1]. The PA [2/0] will be equivalent to $-\frac{\rho(\mathcal{R})}{b\mathcal{R}^2}$ and will not give any additional information. The [0/2] PA is:

$$[0/2] = \frac{a_0}{1 + b_1\mathcal{R} + b_2\mathcal{R}^2}. \tag{3.4.5}$$

When we set this equal to the right hand side of (3.4.4) and multiply up by $1 + b_1\mathcal{R} + b_2\mathcal{R}^2$ we get:

$$a_0 = 1 + (b_1 + c)\mathcal{R} + (b_2 + \rho_2 + b_1 c)\mathcal{R}^2 + (b_1\rho_2 + cb_2 + \rho_3)\mathcal{R}^3 + \mathcal{O}(\mathcal{R}^4). \tag{3.4.6}$$

Comparing coefficients we find that $a_0 = 1$, $b_1 = -c$ and $b_2 = c^2 - \rho_2$, which means that the full expression for the PA is:

$$[0/2] = \frac{1}{1 - c\mathcal{R} + (c^2 - \rho_2)\mathcal{R}^2}. \tag{3.4.7}$$

We also see from the $\mathcal{R}^3$ term in (3.4.6) that $\rho_3 = 2c\rho_2 - c^3$.

Constructing the [1/1] PA and proceeding in a similar fashion we find that:

$$[1/1] = \frac{a_0 + a_1\mathcal{R}}{1 + b_1\mathcal{R}} = \frac{1 + \left(\frac{c^2-\rho_2}{c}\right)\mathcal{R}}{1 - \frac{\rho_2}{c}\mathcal{R}} \tag{3.4.8}$$

and that $\rho_3 = \rho_2^2/c$.

We now have two estimates of ρ_3 from the two PAs [0/2] and [1/1]. The $\rho_3\mathcal{R}^3$ terms as calculated from the PA estimations of ρ_3 are shown in Table 3.9. The [1/1] PA displays better behaviour than [0/2] in the sense that for those variables where $\rho_2\mathcal{R}^2 > c\mathcal{R}$ it predicts $\rho_3\mathcal{R}^3 > \rho_2\mathcal{R}^2$. The [0/2] PA, on the other hand, predicts that the $\rho_3\mathcal{R}^3$ term gets smaller again, which is not behaviour that would be expected from an asymptotic series. For this reason we shall concentrate on the [1/1] PA from now on.

3.4.2 Using the Padé Sum

In this subsection we use the [1/1] PA to approximate the entire series $-\rho/b\mathcal{R}^2$. Instead of using the PA to predict the value of one higher order coefficient, ρ_3, we use it as a Padé Sum (PS) to provide an approximation to the resummed quantity. We extend the analysis performed in Section 3.2 to using the ECH method with the [1/1] PS for $\rho(\mathcal{R})$, as given in (3.4.8). As before $\alpha_s(M_Z)$ is fixed at the result extracted from N^{3}LO calculations on Z-decays [51], $\alpha_s(M_Z) = 0.1190 \pm 0.0026$, corresponding to $\tilde{\Lambda}_{\overline{\rm MS}} = 254$ MeV.

Observable	n	$c\mathcal{R}$ (NLO)	$\rho_2\mathcal{R}^2$ (NNLO)	$\rho_3\mathcal{R}^3$ (N^3LO approx. from [0/2] PA)	$\rho_3\mathcal{R}^3$ (N^3LO approx. from [1/1] PA)
$\langle(1-T)^n\rangle$	1	0.078	−0.049	−0.008	0.031
	2	0.105	−0.323	−0.069	0.992
	3	0.119	−0.575	−0.139	2.766
$\langle C^n\rangle$	1	0.076	−0.030	−0.005	0.012
	2	0.106	−0.376	−0.081	1.338
	3	0.123	−0.912	−0.226	6.769
$\langle B_W^n\rangle$	1	0.045	0.004	0.0003	0.0003
	2	0.057	−0.031	−0.004	0.017
	3	0.055	−0.039	−0.004	0.027
$\langle B_T^n\rangle$	1	0.066	−0.060	−0.008	0.056
	2	0.111	−1.803	−0.403	29.182
	3	0.140	−5.715	−1.599	233.842
$\langle Y_3^n\rangle$	1	0.058	−0.042	−0.005	0.030
	2	0.060	−0.049	−0.006	0.039
	3	0.060	−0.053	−0.007	0.048
$\langle \rho_E^n\rangle$	1	0.059	−0.015	−0.002	0.004

Table 3.9: Size of terms in $\frac{\rho(\mathcal{R})}{-b\mathcal{R}^2}$, including estimates of $\rho_3\mathcal{R}^3$ obtained from [0/2] and [1/1] PAs.

Again, we want to invert (2.1.15) to get an approximation to the data. To do this we need to find a new expression for $\mathrm{G}(\mathcal{R})$ which replaces $\rho(\mathcal{R})$ with the

[1/1] PS:

$$\begin{aligned}
\mathrm{G}(\mathcal{R}) &= \int_0^{\mathcal{R}(Q)} dx \left[\frac{b}{\rho(x)} + \frac{1}{x^2(1+cx)}\right] \\
&= \int_0^{\mathcal{R}(Q)} dx \left[-\frac{1-\frac{\rho_2}{c}x}{x^2\left(1+\left(\frac{c^2-\rho_2}{c}\right)x\right)} + \frac{1}{x^2(1+cx)}\right] \\
&= \int_0^{\mathcal{R}(Q)} dx \, \frac{\rho_2}{(1+cx)\left(1+\left(\frac{c^2-\rho_2}{c}\right)x\right)} \\
&= c\ln\left[\frac{c\,(1+c\mathcal{R})}{c+(c^{\,2}-\rho_2)\,\mathcal{R}}\right]. \qquad (3.4.9)
\end{aligned}$$

Putting this together with the expression for $\mathrm{F}(\mathcal{R})$ gives:

$$\begin{aligned}
b\ln\frac{Q}{\Lambda_{\mathcal{R}}} &= \frac{1}{\mathcal{R}} + c\ln\left[\frac{c\mathcal{R}}{1+c\mathcal{R}}\right] + c\ln\left[\frac{c\,(1+c\mathcal{R})}{c+(c^{\,2}-\rho_2)\,\mathcal{R}}\right] \\
&= \frac{1}{\mathcal{R}} + c\ln\left[\frac{c\mathcal{R}}{1+(c-\rho_2/c)\mathcal{R}}\right]. \qquad (3.4.10)
\end{aligned}$$

This expression is inverted numerically to give $\mathcal{R}$, and an approximation to data is obtained using $\langle y^n\rangle = \frac{\overline{\mathcal{A}}_{y,n}}{2}\mathcal{R}$, from (2.2.3).

It is also possible to apply Padé Approximant methods to standard $\overline{\mathrm{MS}}$ perturbation theory. We can approximate the series in (2.2.2), truncated at NNLO and with μ_{R} set equal to Q:

$$\langle y^n\rangle = \left(\frac{\alpha_s(Q)}{2\pi}\right)\overline{\mathcal{A}}_{y,n} + \left(\frac{\alpha_s(Q)}{2\pi}\right)^2\overline{\mathcal{B}}_{y,n} + \left(\frac{\alpha_s(Q)}{2\pi}\right)^3\overline{\mathcal{C}}_{y,n}, \qquad (3.4.11)$$

with its [1/1] PS:

$$\begin{aligned}
\langle y^n\rangle &= \left(\frac{\alpha_s(Q)}{2\pi}\right)[1/1] + \mathcal{O}(\alpha_s(Q)^4) \\
&= \left(\frac{\alpha_s(Q)}{2\pi}\right)\frac{\overline{\mathcal{A}}_{y,n} + \frac{1}{\overline{\mathcal{B}}_{y,n}}\left(\overline{\mathcal{B}}^2_{y,n} - \overline{\mathcal{C}}_{y,n}\overline{\mathcal{A}}_{y,n}\right)\frac{\alpha_s(Q)}{2\pi}}{1-\frac{\overline{\mathcal{C}}_{y,n}}{\overline{\mathcal{B}}_{y,n}}\frac{\alpha_s(Q)}{2\pi}} + \mathcal{O}(\alpha_s(Q)^4). \qquad (3.4.12)
\end{aligned}$$

(3.4.12) gives a second PS approximation to the event shape data. We plot it, alongside the PS ECH approximation and the NLO ECH and NNLO $\overline{\text{MS}}$ predictions to data, in Figs. 3.10-3.15. The NLO ECH and NNLO $\overline{\text{MS}}$ predictions were chosen as they were found in Section 3.2 and Ref. [1] to be the best approximations to data for the respective methods.

From Figs. 3.10-3.15, we can see that ECH at NLO still generally gives the best approximation to the data for the means (with the exceptions noted in Section 3.3). For the higher moments both ECH and $\overline{\text{MS}}$ give good predictions when combined with the PS, often substantially better than when using ECH or $\overline{\text{MS}}$ alone. $\overline{\text{MS}}$ with the PS generally slightly outperforms the ECH method with the PS. This is especially the case for the higher moments of $\langle B_T \rangle$.

As in Section 3.2, for some higher moments there are imaginary predictions at certain energies. However, the situation is vastly improved when using the PS. For both $1-T$ and C the higher moment $n = 2, 3$ are free of unphysical imaginary predictions across the energy range studied. $\langle B_T^2 \rangle$ only has imaginary values at the very lowest energies of $Q \sim 10$ GeV, and the 3^{rd} moment of B_T gives real predictions above ~ 60 GeV. Equation (3.4.10) has a very similar structure to the fundamental equation for ECH at NLO which is why these imaginary values can still occur. Using ECH with the PS appears to increase the region where perturbation theory can be applied successfully.

When a PS is added to standard $\overline{\text{MS}}$ PT the scale uncertainty, which is calculated by varying μ_{R} between $0.5Q$ and $2Q$, is found to be considerably smaller than that at NNLO. To illustrate this we plot $\overline{\text{MS}}$ PT predictions, along with their scale variations, at NLO, NNLO and with a PS for $\langle (1-T)^n \rangle$ in Fig. 3.16. This behaviour suggests that the Padé Sum is estimating the missing higher orders effectively, thus removing most of the residual RS dependence.

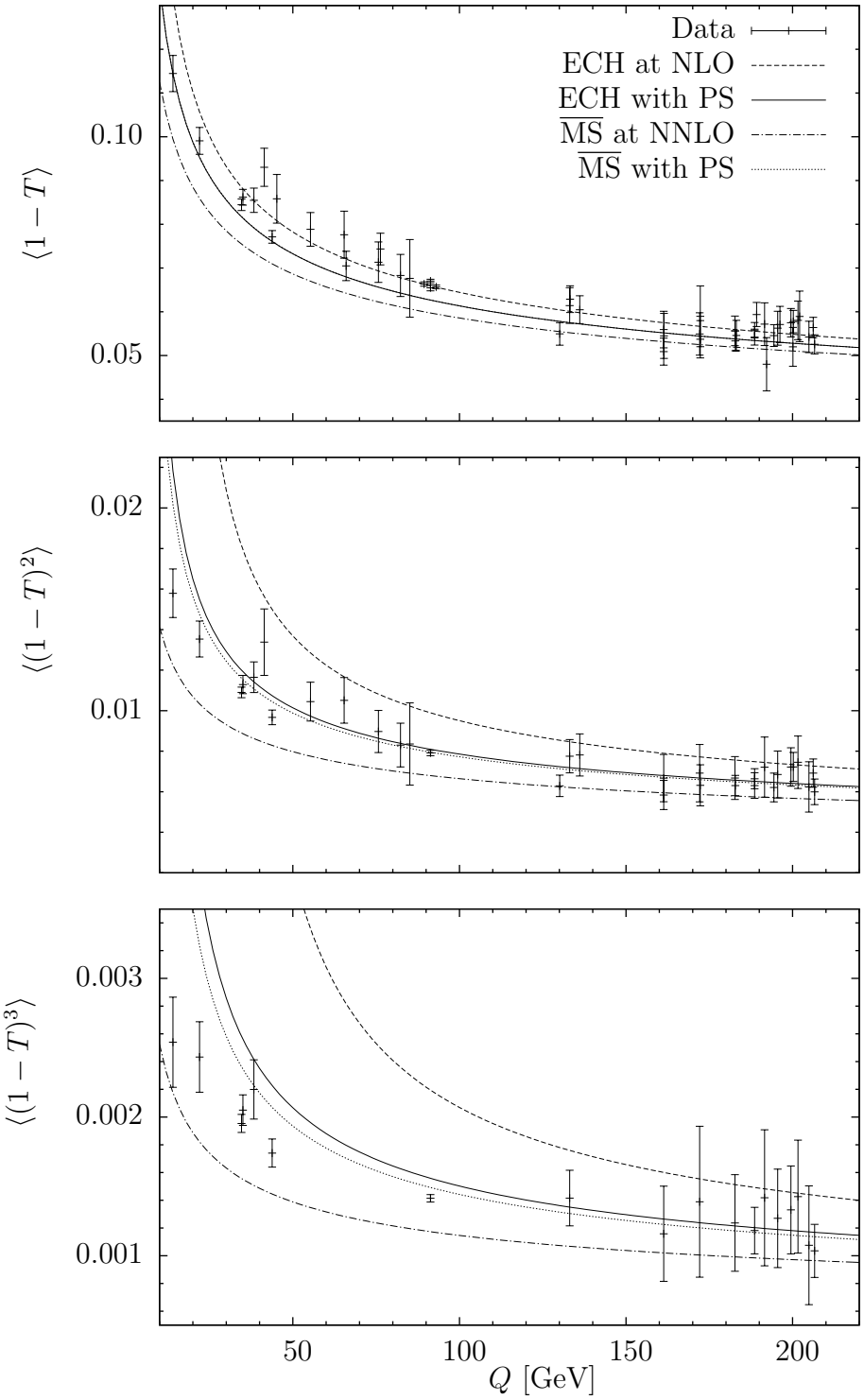


Figure 3.10: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximation to data at NNLO and with PS for $\langle (1-T)^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

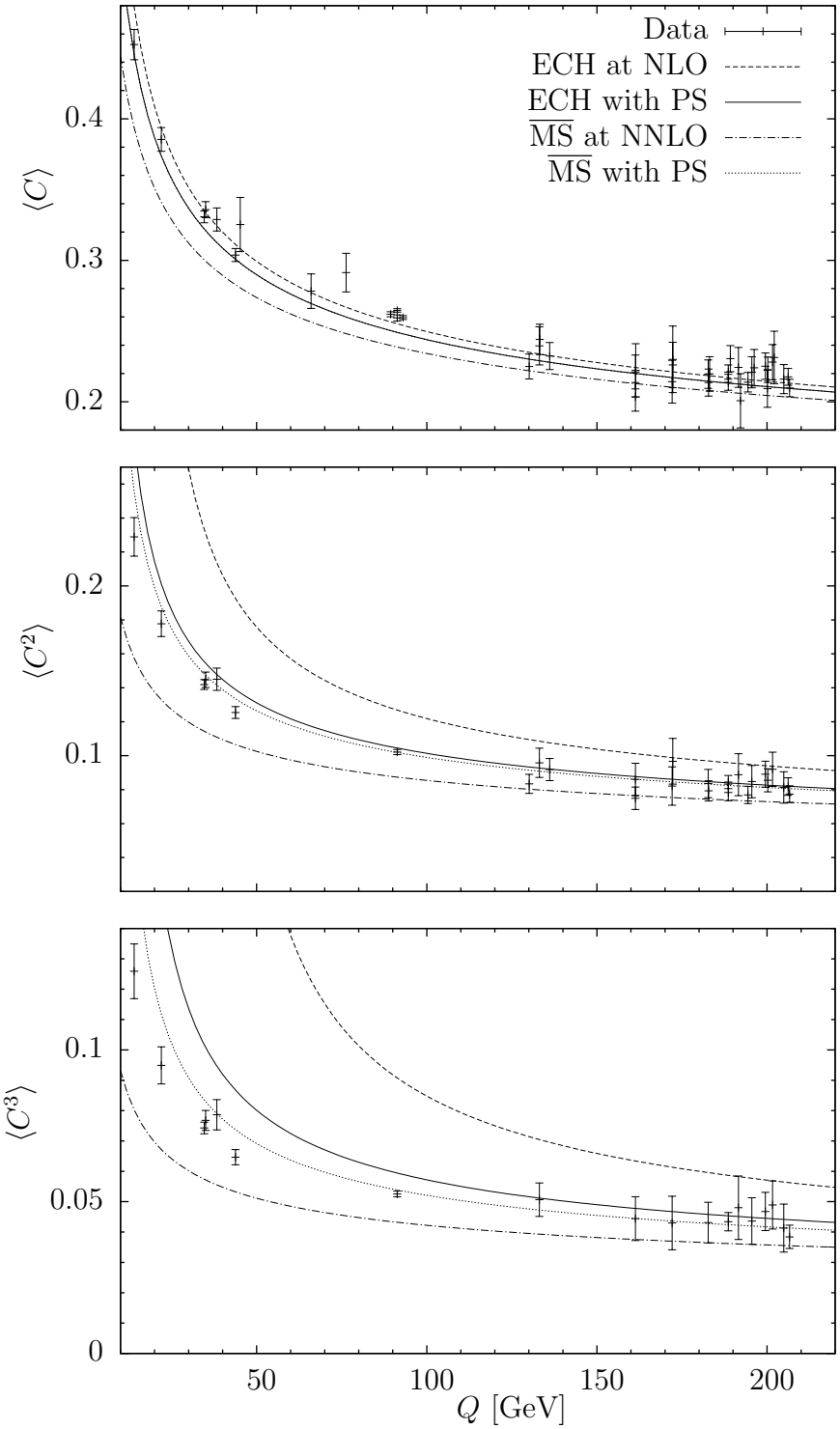


Figure 3.11: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximation to data at NNLO and with PS for $\langle C^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

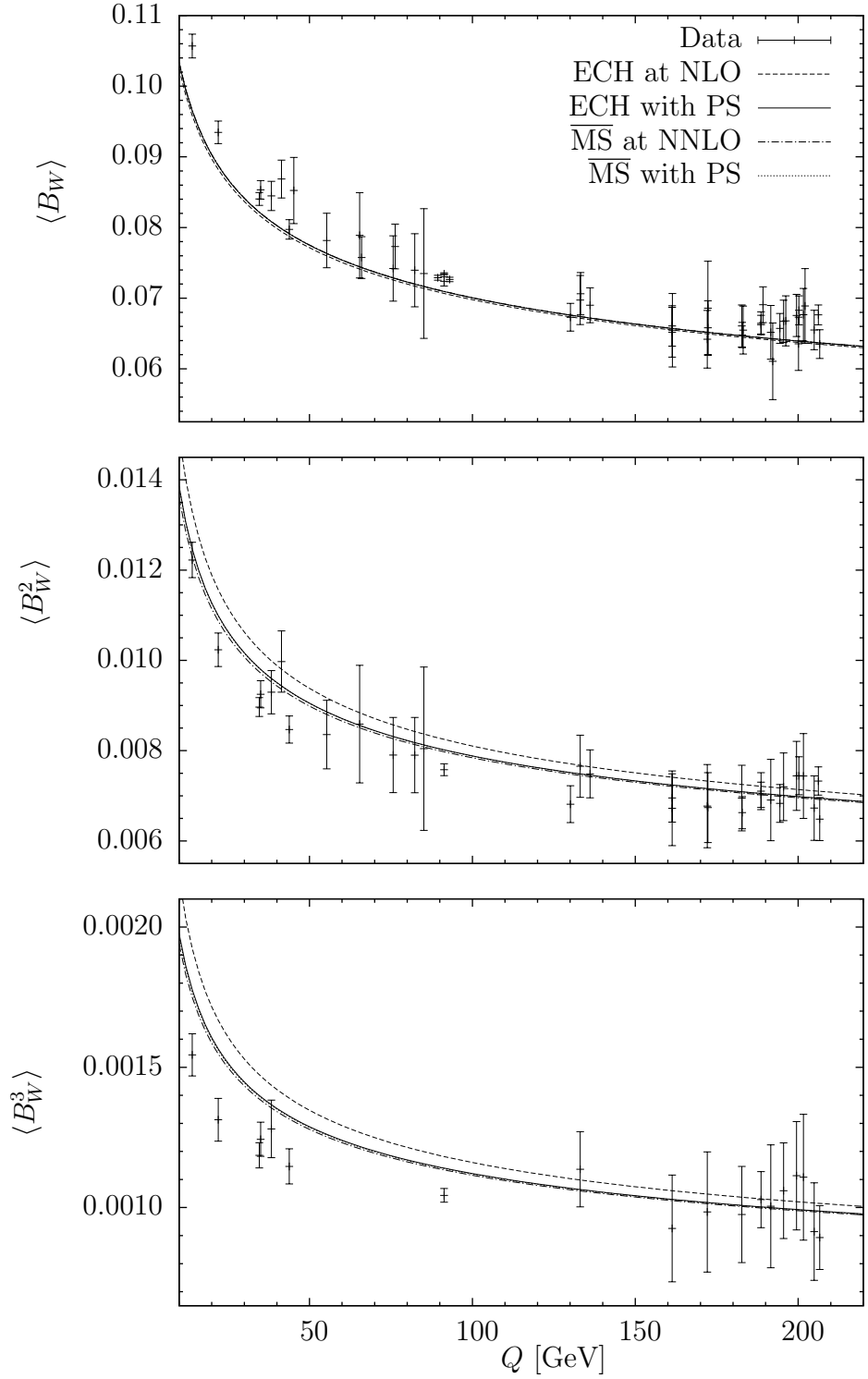


Figure 3.12: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximation to data at NNLO and with PS for $\langle B_W^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

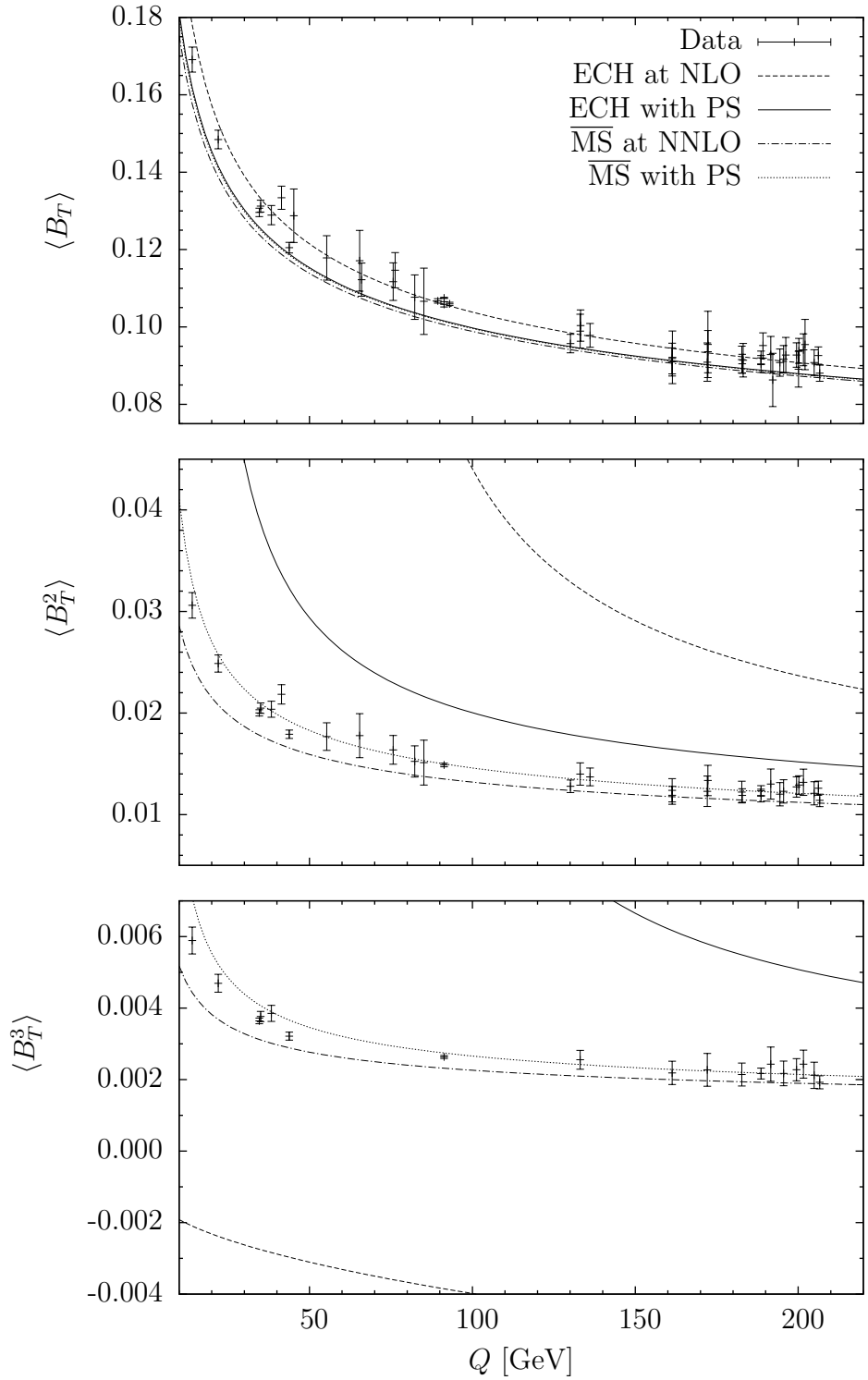


Figure 3.13: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximation to data at NNLO and with PS for $\langle B_T^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

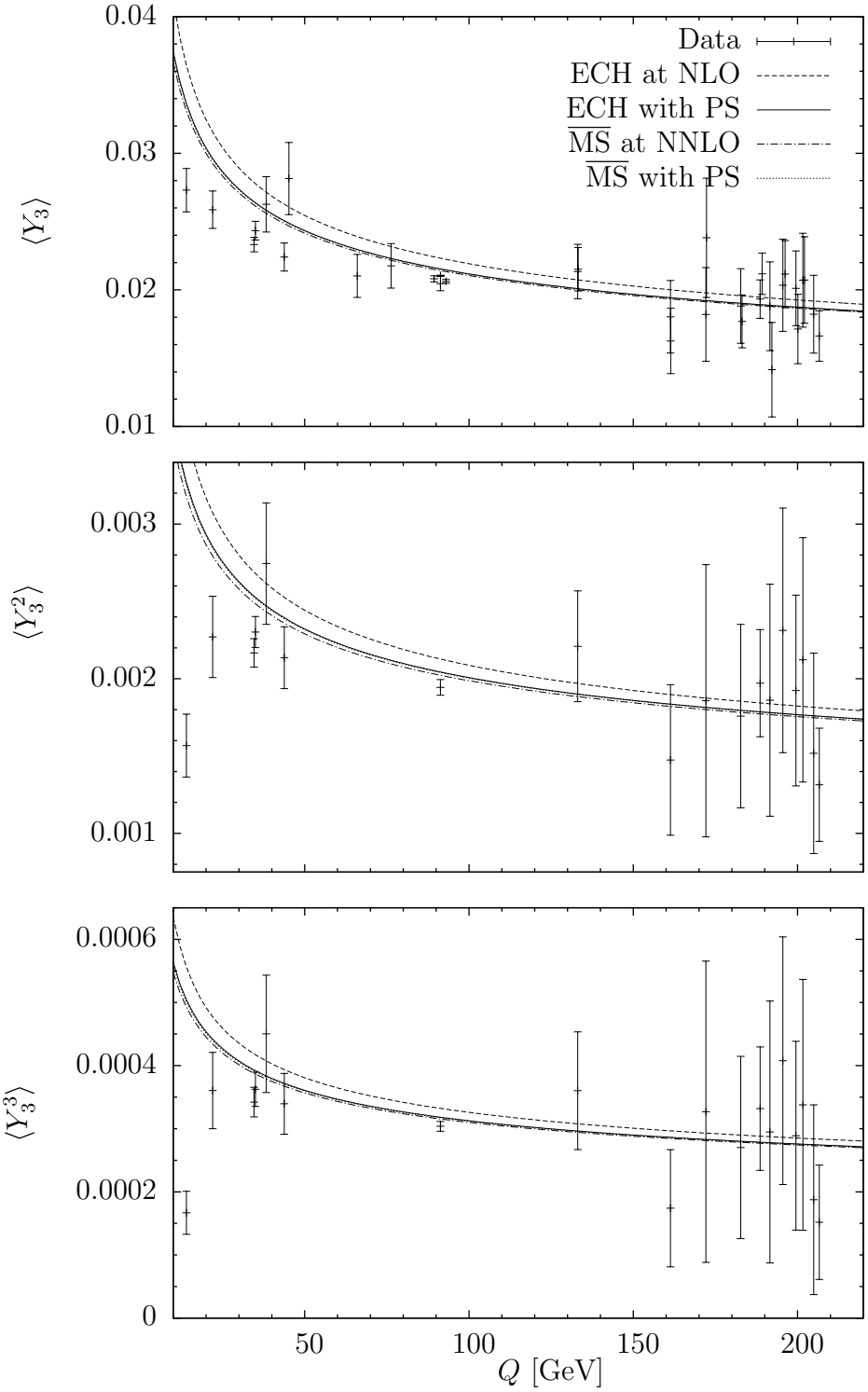


Figure 3.14: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximation to data at NNLO and with PS for $\langle Y_3^n \rangle$, with $n = 1, 2, 3$ from top to bottom.

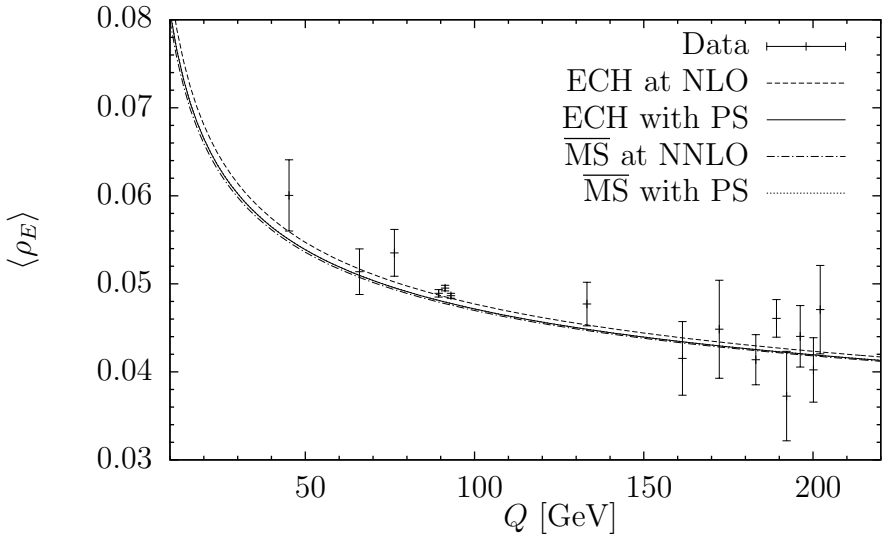


Figure 3.15: ECH at NLO and with PS and $\overline{\text{MS}}$ PT approximations to data at NNLO and with PS for $\langle \rho_E \rangle$.

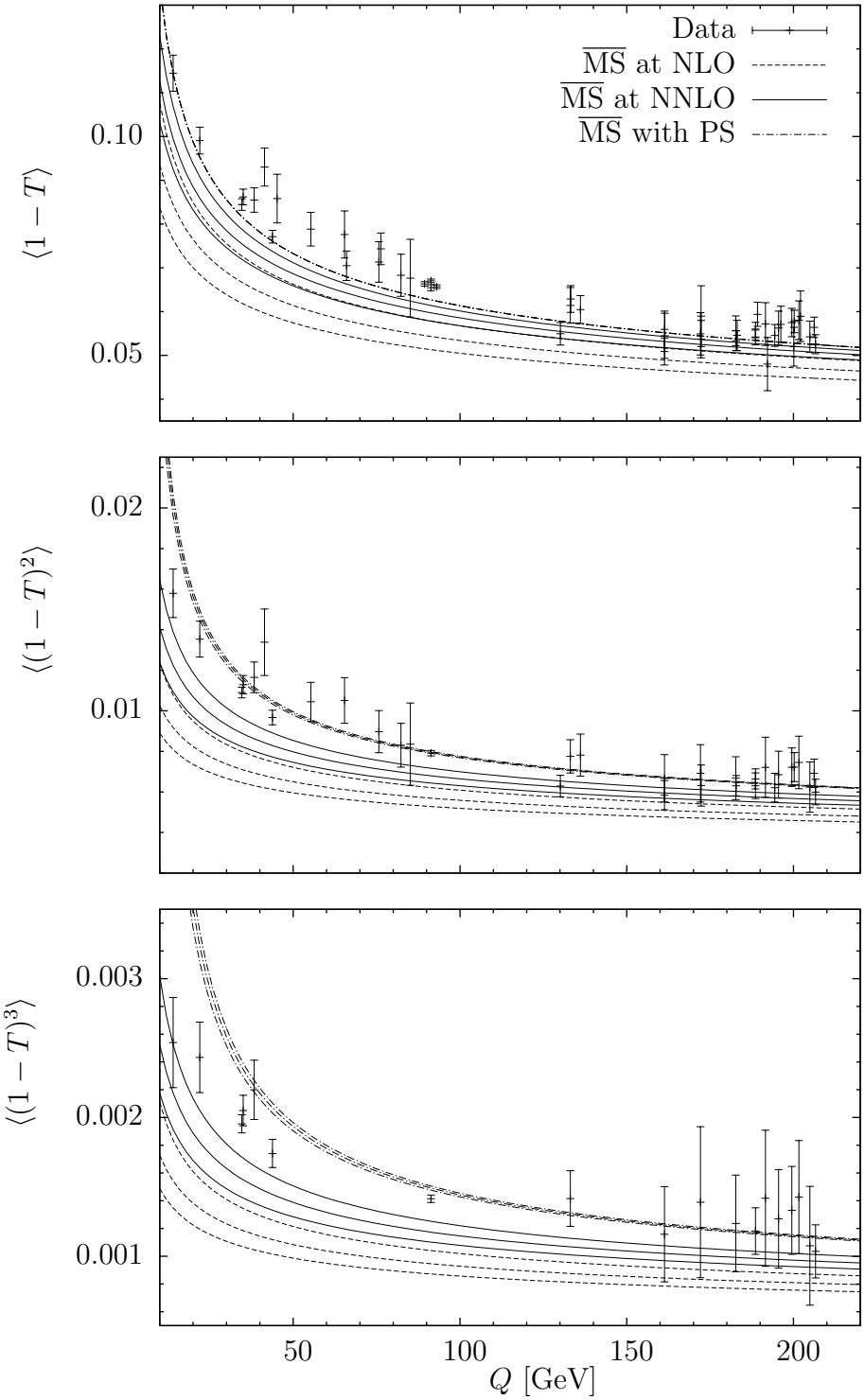


Figure 3.16: $\overline{\rm MS}$ PT approximations to data at NLO, NNLO and with PS, along with the scale variation in each case, for $\langle (1-T)^n\rangle$, with $n = 1, 2, 3$ from top to bottom.

3.5 Non-perturbative power corrections

In this section we use the ECH and standard $\overline{\text{MS}}$ methods in conjunction with two different non-perturbative power correction models. The first is a dispersive model, which is outlined in Subsection 3.5.1, and the second is a simple power correction model, described in Subsection 3.5.2.

3.5.1 Dispersive power corrections

In this subsection we use the ECH method in combination with a dispersive model for power corrections [6, 48, 49]. Hadronisation corrections for event shape moments are expected to be additive [6, 34]:

$$\langle y^n \rangle = \langle y^n \rangle_{pt} + \langle y^n \rangle_{np}, \tag{3.5.1}$$

where $\langle y^n \rangle_{pt}$ is the perturbative part of the event shape moment and $\langle y^n \rangle_{np}$ is the non-perturbative part.

The dispersive model for power corrections accounts for non-perturbative behaviour at low energies by replacing the strong coupling constant with an effective coupling, α_{eff}, below an IR cutoff scale, μ_{I}. This is done in such a way that the integral of the effective coupling up to μ_{I} is finite:

$$\frac{1}{\mu_{\text{I}}} \int_0^{\mu_{\text{I}}} dQ\, \alpha_{\text{eff}}(Q^2) = \alpha_0(\mu_{\text{I}}). \tag{3.5.2}$$

The effective coupling is related to the strong coupling through a dispersion relation [61, 62]:

$$\alpha_s(k) = k^2 \int_0^{\infty} dm^2 \frac{\alpha_{\text{eff}}(m^2)}{(m^2 + k^2)^2}\,. \tag{3.5.3}$$

Introducing an effective coupling in this way leads to a shift in the distributions

of the event shapes [6]:

$$\frac{d\sigma}{dy}(y) = \frac{d\sigma_{\rm pt}}{dy}(y - a_y P)\,, \tag{3.5.4}$$

where the a_y are numerical factors that depend on the event shape in question [6]. Their values are given in Table 3.10. This in turn leads to the following contribution to the event shape moments [6]:

$$\begin{aligned}\langle y^n \rangle &= \int_{-a_y P}^{y_{\rm max}-a_y P} dy\,(y + a_y P)^n \frac{1}{\sigma_{\rm had}} \frac{d\sigma_{\rm pt}}{dy}(y) \\ &\approx \int_0^{y_{\rm max}} dy\,(y + a_y P)^n \frac{1}{\sigma_{\rm had}} \frac{d\sigma_{\rm pt}}{dy}(y)\,,\end{aligned} \tag{3.5.5}$$

and for the event shape means the non-perturbative contribution is therefore of the form:

$$\langle y \rangle_{np} = a_y P. \tag{3.5.6}$$

Note that since $a_{Y_3} = 0$ there are no dispersive power corrections for $\langle Y_3^n \rangle$.

The form of P at NNLO [6, 63] is:

$$\begin{aligned}P = \frac{4C_F}{\pi^2}\mathcal{M}\Bigg(\alpha_0 - \Bigg[\alpha_s(\mu_{\rm R}) + \frac{b}{\pi}\left(1 + \ln\frac{\mu_{\rm R}}{\mu_{\rm I}} + \frac{K}{2b}\right)\alpha_s^2(\mu_{\rm R}) \\ + \left(4bc\left(1 + \ln\frac{\mu_{\rm R}}{\mu_{\rm I}} + \frac{L}{4bc}\right) + 8b^2\left(1 + \ln\frac{\mu_{\rm R}}{\mu_{\rm I}} + \frac{K}{2b}\right)\right. \\ \left. + 4b^2 \ln\frac{\mu_{\rm R}}{\mu_{\rm I}}\left(\ln\frac{\mu_{\rm R}}{\mu_{\rm I}} + \frac{K}{b}\right)\right)\frac{\alpha_s^3(\mu_{\rm R})}{4\pi^2}\Bigg]\Bigg)\frac{\mu_{\rm I}}{Q},\end{aligned} \tag{3.5.7}$$

where $C_F = \frac{N^2-1}{2N}$ (with $N = 3$ being the number of colours), and:

$$K = \left(\frac{67}{18} - \frac{\pi^2}{6}\right) C_A - \frac{5}{9} N_F,$$
$$L = C_A^2 \left(\frac{245}{24} - \frac{67}{9}\frac{\pi^2}{6} + \frac{11}{6}\zeta_3 + \frac{11}{5}\left(\frac{\pi^2}{6}\right)^2\right) + C_F N_F \left(-\frac{55}{24} + 2\zeta_3\right) + C_A N_F \left(-\frac{209}{108} + \frac{10}{9}\frac{\pi^2}{6} - \frac{7}{3}\zeta_3\right) - \frac{1}{27} N_F^2,$$

where $C_A = N$ and N_F is the number of quark flavours. The first term in P is the contribution to the low energy region from the effective coupling, and is proportional to $1/Q$. The terms in the square brackets in P subtract off the perturbative contribution in the region below μ_I to avoid double counting [63]. $\mathcal{M}$ is the Milan factor, a two-loop enhancement factor [64]. $\mathcal{M}$ is universal for observables with this type of $1/Q$ power correction. Given the energy range of the data used in this thesis, the number of active quark flavours is 5. The form of P at NLO ($P^{(NLO)}$) is the same as P in (3.5.7) except that the α_s^3 term in the square brackets is omitted [64].

For B_W and B_T there are additional corrections to P, due to quark recoil complications: there is a small mismatch between the quark axis and the thrust axis, caused by perturbative gluon radiation, that these observables are sensitive to [61]. At NLO this results in additional contributions to the non-perturbative parts of B_W and B_T of [61]:

$$P_{B_W}^{(NLO)} = P^{(NLO)} \left(\frac{\pi}{\sqrt{8 C_F \hat{\alpha}_s \left(1 + \frac{K\hat{\alpha}_s}{2\pi}\right)}} + \frac{3}{4} - \frac{b}{6C_F} + \eta_0\right), \tag{3.5.8}$$

$$P_{B_T}^{(NLO)} = P^{(NLO)} \left(\frac{\pi}{\sqrt{4 C_F \hat{\alpha}_s \left(1 + \frac{K\hat{\alpha}_s}{2\pi}\right)}} + \frac{3}{4} - \frac{b}{3C_F} + \eta_0\right), \tag{3.5.9}$$

where $\hat{\alpha}_s(\mu_\mathrm{R}) \equiv \alpha_s(e^{-3/4}\mu_\mathrm{R})$ and $\eta_0 = -0.6137$. The corrections to P for B_W

and B_T at NNLO have not been explicitly calculated yet but the potentially dominant terms are approximated by [6]:

$$P_{B_W} = P\left(\frac{\pi}{\sqrt{8C_F\hat{\alpha}_s\left(1+\frac{K\hat{\alpha}_s}{2\pi}+\frac{L\hat{\alpha}_s^2}{4\pi^2}\right)}}+\frac{3}{4}-\frac{b}{6C_F}+\eta_0\right), \tag{3.5.10}$$

$$P_{B_T} = P\left(\frac{\pi}{\sqrt{4C_F\hat{\alpha}_s\left(1+\frac{K\hat{\alpha}_s}{2\pi}+\frac{L\hat{\alpha}_s^2}{4\pi^2}\right)}}+\frac{3}{4}-\frac{b}{3C_F}+\eta_0\right). \tag{3.5.11}$$

In this section we use the forms in (3.5.1) and (3.5.6) to predict the form of the data and then perform simultaneous fits for $\alpha_s(M_Z)$ and α_0. The parameter α_0 is expected to be universal. For $\langle y\rangle_{pt}$ we use the ECH prediction at NLO and NNLO (see Section 3.2). In order to facilitate inversion of (2.1.15) at NNLO we use Padé Approximant (PA) methods [29, 57–60], which were described in Section 3.4. We do this by writing (2.1.5) at NNLO as x^2 multiplied by a [1/1] PA:

$$\rho(x) = -bx^2\left(\frac{1+(c-\frac{\rho_2}{c})x}{1-(\frac{\rho_2}{c})x}+\mathcal{O}(x^3)\right). \tag{3.5.12}$$

We can then integrate and invert analytically to get:

$$\mathcal{R}_{\mathrm{NNLO}}(Q) = \frac{-1}{c\left[1-\frac{\rho_2}{c^2}+W\left(-\exp\left[-\left(1-\frac{\rho_2}{c^2}+\frac{b}{c}\ln\frac{Q}{\Lambda_{\mathcal{R}}}\right)\right]\right)\right]}. \tag{3.5.13}$$

In the expressions for the power correction term, μ_{R} is set to $\mu_{\mathrm{ECH}} = e^{-r/b}\,Q$ (see the discussion below (3.2.1)) when the ECH perturbative part is used.

The IR cutoff should be chosen such that $\Lambda_{\mathrm{QCD}} \ll \mu_{\mathrm{I}} \ll Q$. We therefore take μ_{I} to be 2 GeV. Since the fits for α_s and α_0 have a dependency on the choice of cutoff, we vary μ_{I} between 1 and 3 GeV to assess the extent of the dependency. The variation in the fits is used to assign a theoretical uncertainty.

Event shape	$1-T$	C	B_W	B_T	Y_3	ρ_E
a_y	2	3π	1/2	1	0	1

Table 3.10: The a_y coefficients for different event shapes

	Central value	Upper value	Lower value
$\mu_{\rm I}$ [GeV]	2	3	1
$\mathcal{M}$	1.49	1.788 (+20 %)	1.192 (−20 %)
$\mu_{\rm R}$ [GeV]	Q	$2Q$	$Q/2$

Table 3.11: Variation of $\mu_{\rm I}$, $\mathcal{M}$ and $\mu_{\rm R}$

This is in keeping with the method used in Ref. [6]. The Milan factor is known at two-loop order to be $\mathcal{M} = 1.49 \pm 20\%$, so $\mathcal{M}$ is also varied around this central value. The central, upper and lower values of $\mu_{\rm I}$ and $\mathcal{M}$ are shown in Table 3.11. When the effect of the upward and downward variation result in uneven differences the larger of the two is taken to contribute towards the theoretical uncertainty. The individual differences from the variation of $\mu_{\rm I}$ and $\mathcal{M}$ are added together in quadrature to give the theoretical uncertainty.

Since it was found in the previous sections that generally higher moments work less well than the $n = 1$ moments, we confine ourselves to analysing the means in this part of the thesis. The results of the dual fits for means at NLO and NNLO are shown in Table 3.12. The first uncertainty is the combined experimental and perturbative coefficient uncertainty, and the second is the theoretical uncertainty (estimated by varying $\mu_{\rm I}$ and $\mathcal{M}$). These results are plotted in Fig. 3.17. We also perform fits using the $\overline{\rm MS}$ PT prediction at NLO and NNLO. These results are shown in Table 3.13 and Fig. 3.18. In these cases $\mu_{\rm R}$ is varied (around a central value of $\mu_{\rm R} = Q$), as well as $\mu_{\rm I}$ and $\mathcal{M}$, to find the theoretical uncertainty (see Table 3.11).

Observable	$\alpha_s(M_Z)$	α_0	χ^2/dof
	NLO		
$\langle 1-T\rangle$	$0.1144 \pm 0.0007 \pm 0.0028$	$0.436 \pm 0.011 \pm 0.006$	52.6/46
$\langle C\rangle$	$0.1133 \pm 0.0006 \pm 0.0031$	$0.423 \pm 0.006 \pm 0.010$	58.5/40
$\langle B_W\rangle$	$0.1192 \pm 0.0006 \pm 0.0014$	$0.307 \pm 0.014 \pm 0.011$	41.4/46
$\langle B_T\rangle$	$0.1188 \pm 0.0010 \pm 0.0038$	$0.312 \pm 0.027 \pm 0.053$	59.8/46
$\langle \rho_E\rangle$	$0.1163 \pm 0.0055 \pm 0.0010$	$0.342 \pm 0.136 \pm 0.012$	11.6/13
	NNLO		
$\langle 1-T\rangle$	$0.1142 \pm 0.0006 \pm 0.0017$	$0.584 \pm 0.010 \pm 0.027$	58.3/46
$\langle C\rangle$	$0.1130 \pm 0.0005 \pm 0.0019$	$0.539 \pm 0.008 \pm 0.026$	67.0/40
$\langle B_W\rangle$	$0.1184 \pm 0.0007 \pm 0.0014$	$0.392 \pm 0.014 \pm 0.013$	43.1/46
$\langle B_T\rangle$	$0.1184 \pm 0.0008 \pm 0.0028$	$0.537 \pm 0.020 \pm 0.008$	67.2/46
$\langle \rho_E\rangle$	$0.1168 \pm 0.0054 \pm 0.0010$	$0.449 \pm 0.114 \pm 0.017$	11.6/13

Table 3.12: Fits for $\alpha_s(M_Z)$ and α_0 using ECH and dispersive power corrections at NLO and NNLO.

In analogy to Section 3.2, we also perform fits for α_0 while keeping $\alpha_s(M_Z)$ fixed at 0.1190. This is done for both ECH and standard $\overline{\text{MS}}$ PT at NLO and NNLO, and the results are shown in Tables 3.14–3.15 and Figs. 3.19–3.20.

3.5.2 Simple power corrections

For comparison we now use a second model for non-perturbative corrections: a simple power correction [35,50]. For the ECH method this is done by adding a c_1/Q power correction (with c_1 a constant) to the perturbative expansion of $\mathcal{R}$ given in (2.1.4). This results in an altered form for the $\rho(\mathcal{R})$ function as follows:

$$\rho(\mathcal{R}) = -b\mathcal{R}^2(1 + c\mathcal{R} + \rho_2\mathcal{R}^2 + \dots) + \kappa_0\mathcal{R}^{-c/b}e^{-1/(b\mathcal{R})}. \tag{3.5.14}$$

Observable	$\alpha_s(M_Z)$	α_0	χ^2/dof
	NLO		
$\langle 1-T\rangle$	$0.1302 \pm 0.0010 \pm 0.0088$	$0.382 \pm 0.012 \pm 0.012$	49.7/46
$\langle C\rangle$	$0.1274 \pm 0.0008 \pm 0.0085$	$0.364 \pm 0.007 \pm 0.013$	48.5/40
$\langle B_W\rangle$	$0.1191 \pm 0.0007 \pm 0.0021$	$0.346 \pm 0.015 \pm 0.126$	42.7/46
$\langle B_T\rangle$	$0.1252 \pm 0.0008 \pm 0.0060$	$0.312 \pm 0.013 \pm 0.005$	64.3/46
$\langle \rho_E\rangle$	$0.1200 \pm 0.0060 \pm 0.0042$	$0.323 \pm 0.138 \pm 0.018$	11.5/13
	NNLO		
$\langle 1-T\rangle$	$0.1216 \pm 0.0009 \pm 0.0038$	$0.427 \pm 0.011 \pm 0.020$	49.5/46
$\langle C\rangle$	$0.1197 \pm 0.0007 \pm 0.0039$	$0.412 \pm 0.007 \pm 0.022$	47.9/40
$\langle B_W\rangle$	$0.1179 \pm 0.0007 \pm 0.0017$	$0.429 \pm 0.015 \pm 0.045$	44.3/46
$\langle B_T\rangle$	$0.1212 \pm 0.0008 \pm 0.0029$	$0.385 \pm 0.012 \pm 0.038$	65.1/46
$\langle \rho_E\rangle$	$0.1178 \pm 0.0056 \pm 0.0014$	$0.389 \pm 0.127 \pm 0.033$	11.6/13

Table 3.13: Fits for $\alpha_s(M_Z)$ and α_0 using $\overline{\text{MS}}$ PT and dispersive power corrections at NLO and NNLO.

Observable	α_0			
	NLO	χ^2 / dof	NNLO	χ^2/dof
$\langle 1-T\rangle$	$0.377 \pm 0.007 \pm 0.003$	101 / 47	$0.557 \pm 0.013 \pm 0.019$	166/47
$\langle C\rangle$	$0.375 \pm 0.007 \pm 0.004$	209 / 41	$0.524 \pm 0.013 \pm 0.013$	373/41
$\langle B_W\rangle$	$0.318 \pm 0.005 \pm 0.028$	42.0 / 47	$0.382 \pm 0.014 \pm 0.025$	43.9/47
$\langle B_T\rangle$	$0.314 \pm 0.003 \pm 0.007$	59.8 / 47	$0.526 \pm 0.017 \pm 0.036$	67.9/47
$\langle \rho_E\rangle$	$0.282 \pm 0.010 \pm 0.008$	11.8 / 14	$0.404 \pm 0.015 \pm 0.017$	11.8/14

Table 3.14: Single fits for α_0, with $\alpha_s(M_Z) = 0.1190$ held fixed, for ECH and dispersive power corrections at NLO and NNLO.

This modified form of $\rho(\mathcal{R})$ is then substituted into (2.1.11), with the perturbative part truncated at NLO or NNLO, and a fit to data is performed to extract $\Lambda_{\mathcal{R}}$ (and hence $\alpha_s(M_Z)$) and κ_0.

A corresponding power correction can be added to an NLO or NNLO $\overline{\text{MS}}$

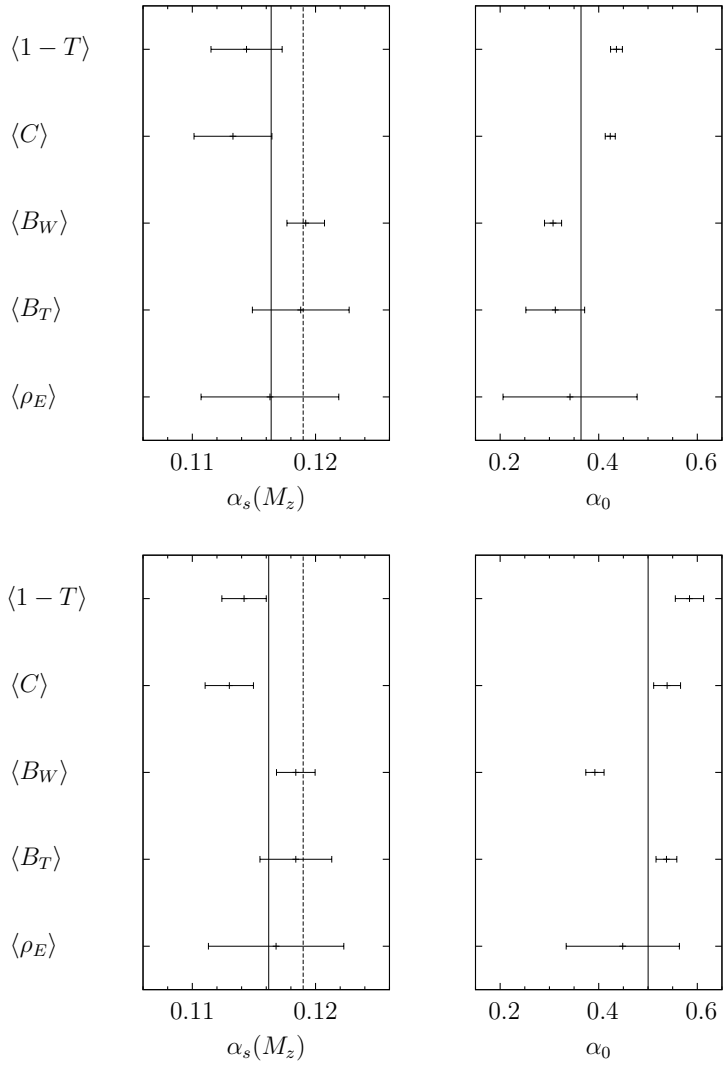


Figure 3.17: Scatter plots of $\alpha_s(M_Z)$ and α_0 from fits of ECH and dispersive power corrections for the means at NLO (top two panels) and NNLO (bottom two panels). The dotted lines on the $\alpha_s(M_Z)$ plots show the value of the coupling obtained from N^{3}LO calculations. The solid lines show the unweighted averages.

perturbative model using:

$$\langle y^n\rangle = \left(\frac{\alpha_s(\mu_{\rm R})}{2\pi}\right)\overline{\mathcal{A}}_{y,n} + \left(\frac{\alpha_s(\mu_{\rm R})}{2\pi}\right)^2\overline{\mathcal{B}}_{y,n} + \left(\frac{\alpha_s(\mu_{\rm R})}{2\pi}\right)^3\overline{\mathcal{C}}_{y,n} + \ldots - \kappa_0 e^{r/b}\left(\frac{b}{2}\right)^{\frac{c}{b}}\frac{\Lambda_{\overline{\rm MS}}}{Q}. \qquad (3.5.15)$$

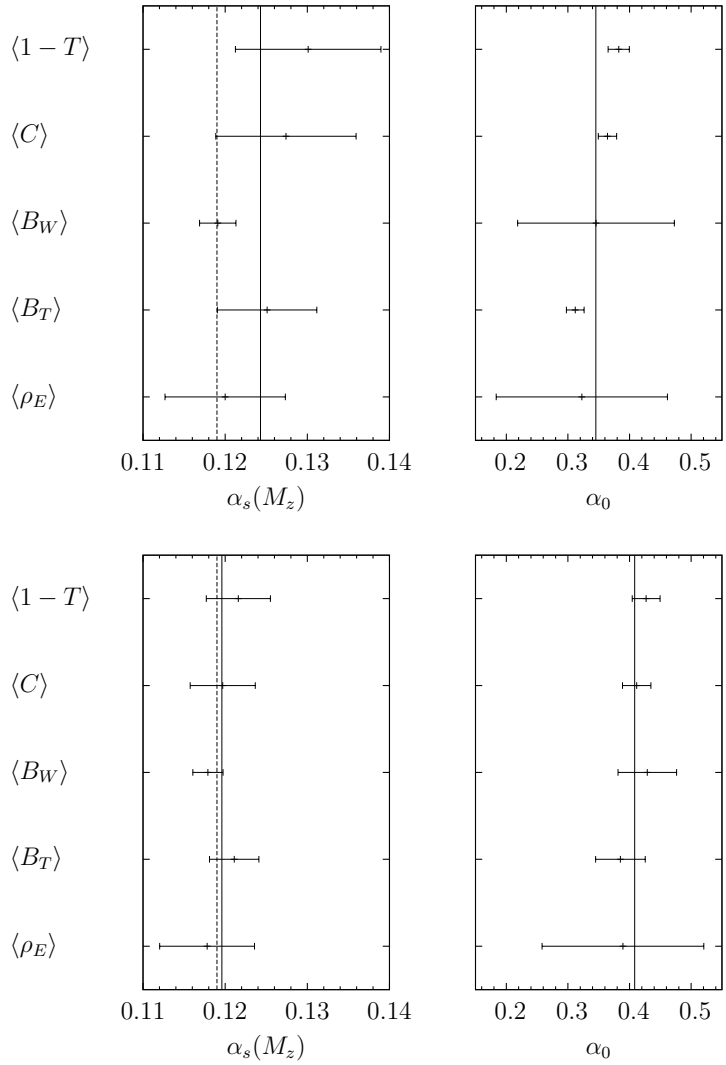


Figure 3.18: Scatter plots of $\alpha_s(M_Z)$ and α_0 from fits of $\overline{\rm MS}$ PT and dispersive power corrections at NLO (top two panels) and NNLO (bottom two panels). The solid and dotted lines are as in Fig. 3.17.

In this model it is not expected that the fits will yield a universal value of κ_0. Each particular observable will have a different value of κ_0, unlike the dispersive model where the power correction parameter α_0 is expected to be universal. The results are shown in Tables 3.16–3.17 and Figs. 3.21–3.22.

Observable	α_0			
	NLO	χ^2 / dof	NNLO	χ^2/dof
$\langle 1-T\rangle$	$0.505 \pm 0.010 \pm 0.088$	186 / 47	$0.457 \pm 0.006 \pm 0.043$	59.9/47
$\langle C\rangle$	$0.433 \pm 0.007 \pm 0.063$	199 / 41	$0.418 \pm 0.005 \pm 0.028$	49.6/41
$\langle B_W\rangle$	$0.354 \pm 0.005 \pm 0.154$	43.0 / 47	$0.409 \pm 0.016 \pm 0.041$	46.8/47
$\langle B_T\rangle$	$0.413 \pm 0.004 \pm 0.071$	168 / 47	$0.419 \pm 0.010 \pm 0.039$	79.9/47
$\langle \rho_E\rangle$	$0.352 \pm 0.010 \pm 0.089$	11.6 / 14	$0.363 \pm 0.012 \pm 0.025$	11.6/14

Table 3.15: Single fits for α_0, with $\alpha_s(M_Z) = 0.1190$ held fixed, for $\overline{\text{MS}}$ PT and dispersive power corrections at NLO and NNLO.

Observable	$\alpha_s(M_Z)$	κ_0	χ^2/dof
	NLO		
$\langle 1-T\rangle$	0.1230 ± 0.0008	-0.039 ± 0.008	51.3/46
$\langle C\rangle$	0.1230 ± 0.0006	-0.036 ± 0.006	43.7/40
$\langle B_W\rangle$	0.1225 ± 0.0005	0.084 ± 0.038	34.2/46
$\langle B_T\rangle$	0.1225 ± 0.0004	-0.054 ± 0.008	60.7/46
$\langle Y_3\rangle$	0.1222 ± 0.0021	-0.237 ± 0.045	31.2/31
$\langle \rho_E\rangle$	0.1184 ± 0.0052	0.047 ± 0.181	11.4/13
	NNLO		
$\langle 1-T\rangle$	0.1253 ± 0.0009	-0.016 ± 0.009	50.7/46
$\langle C\rangle$	0.1247 ± 0.0008	-0.022 ± 0.008	43.7/40
$\langle B_W\rangle$	0.1224 ± 0.0006	0.077 ± 0.038	34.3/46
$\langle B_T\rangle$	0.1269 ± 0.0010	-0.020 ± 0.010	61.5/46
$\langle Y_3\rangle$	0.1261 ± 0.0024	-0.204 ± 0.045	31.2/31
$\langle \rho_E\rangle$	0.1196 ± 0.0053	0.063 ± 0.179	11.4/13

Table 3.16: Fits for $\alpha_s(M_Z)$ and κ_0 using ECH and simple power corrections at NLO and NNLO.

3.6 Commentary on power correction fits

In Section 3.5 we examine the effect of adding a dispersive power correction to the perturbative model. In Fig. 3.17 it is seen that the values of $\alpha_s(M_Z)$ and

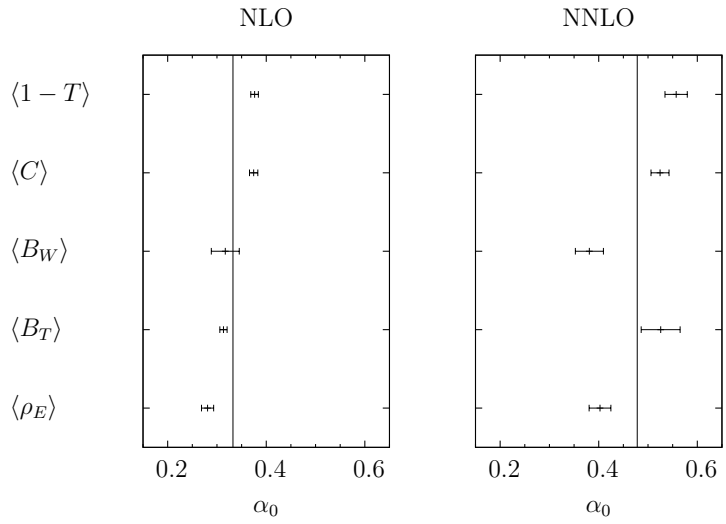


Figure 3.19: Scatter plots for single α_0 fits for ECH and dispersive power corrections at NLO and NNLO.

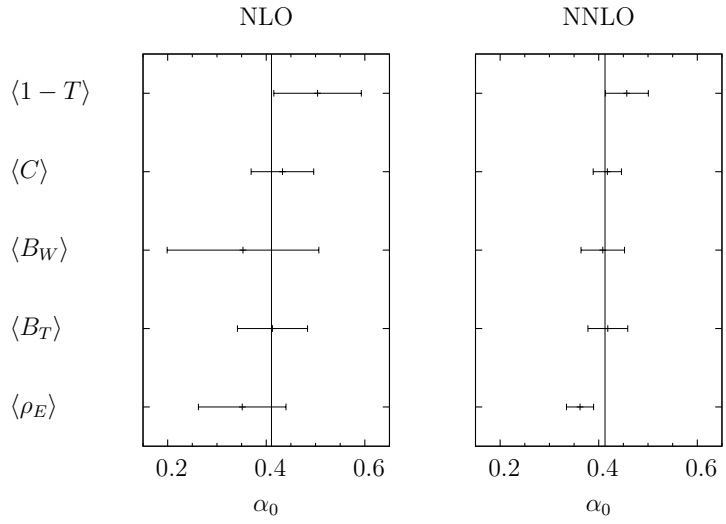


Figure 3.20: Scatter plots for single α_0 fits for $\overline{\text{MS}}$ PT and dispersive power corrections at NLO and NNLO. The solid lines are unweighted averages.

α_0 extracted from $\langle B_W\rangle$, $\langle B_T\rangle$ and $\langle \rho_E\rangle$ using ECH plus a dispersive power correction at NLO agree well with each other. The results for $\langle 1-T\rangle$ and $\langle C\rangle$ are somewhat discrepant. The values of $\alpha_s(M_Z)$ for the other three event shapes are not considerably changed from those extracted using pure ECH. This implies that only very small power corrections are required for ECH at

Observable	$\alpha_s(M_Z)$	κ_0	χ^2/dof
	NLO		
$\langle 1-T\rangle$	$0.1374 \pm 0.0013 \pm 0.0085$	$-0.006 \pm 0.008 \pm 0.012$	51.5/46
$\langle C\rangle$	$0.1357 \pm 0.0009 \pm 0.0079$	$-0.026 \pm 0.028 \pm 0.057$	44.7/40
$\langle B_W\rangle$	$0.1228 \pm 0.0005 \pm 0.0035$	$-0.286 \pm 0.119 \pm 0.314$	34.6/46
$\langle B_T\rangle$	$0.1305 \pm 0.0006 \pm 0.0063$	$0.042 \pm 0.016 \pm 0.051$	61.5/46
$\langle Y_3\rangle$	$0.1260 \pm 0.0012 \pm 0.0054$	$0.060 \pm 0.008 \pm 0.022$	31.6/31
$\langle \rho_E\rangle$	$0.1218 \pm 0.0062 \pm 0.0041$	$-0.097 \pm 0.276 \pm 0.016$	11.4/13
	NNLO		
$\langle 1-T\rangle$	$0.1279 \pm 0.0011 \pm 0.0025$	$0.002 \pm 0.012 \pm 0.020$	51.1/46
$\langle C\rangle$	$0.1270 \pm 0.0008 \pm 0.0023$	$0.018 \pm 0.042 \pm 0.088$	44.1/40
$\langle B_W\rangle$	$0.1223 \pm 0.0007 \pm 0.0003$	$-0.238 \pm 0.129 \pm 0.122$	34.7/46
$\langle B_T\rangle$	$0.1269 \pm 0.0007 \pm 0.0013$	$0.075 \pm 0.020 \pm 0.033$	61.1/46
$\langle Y_3\rangle$	$0.1235 \pm 0.0012 \pm 0.0011$	$0.074 \pm 0.009 \pm 0.010$	31.6/31
$\langle \rho_E\rangle$	$0.1199 \pm 0.0060 \pm 0.0008$	$-0.089 \pm 0.307 \pm 0.016$	11.4/13

Table 3.17: Fits for $\alpha_s(M_Z)$ and κ_0 using $\overline{\rm MS}$ PT and simple power corrections at NLO and NNLO.

NLO for these observables. At NNLO the extracted values of $\alpha_s(M_Z)$ do not change substantially, but the fitted values of α_0 become larger. The fits for $\langle 1-T\rangle$ and $\langle C\rangle$ are still not consistent with those from the other observables. The reduced χ^2 values[3] for the fits generally get further from 1 when going from NLO to NNLO, indicating that the quality of the fits decreases.

For NLO $\overline{\rm MS}$ PT plus a dispersive power correction, shown in Fig. 3.18, the error bars are generally larger than for the corresponding ECH fits. This is because it is necessary to vary $\mu_{\rm R}$, in addition to the Milan factor and the IR cutoff scale, when calculating the theoretical uncertainties. At NNLO these uncertainties get smaller, as expected, and furthermore the agreement between values of $\alpha_s(M_Z)$ and α_0 improves. There is no substantial change in the χ^2 values between the NLO and NNLO fits for $\overline{\rm MS}$ PT. The description of the

[3]Reduced χ^2 values are χ^2 divided by the number of degrees of freedom.

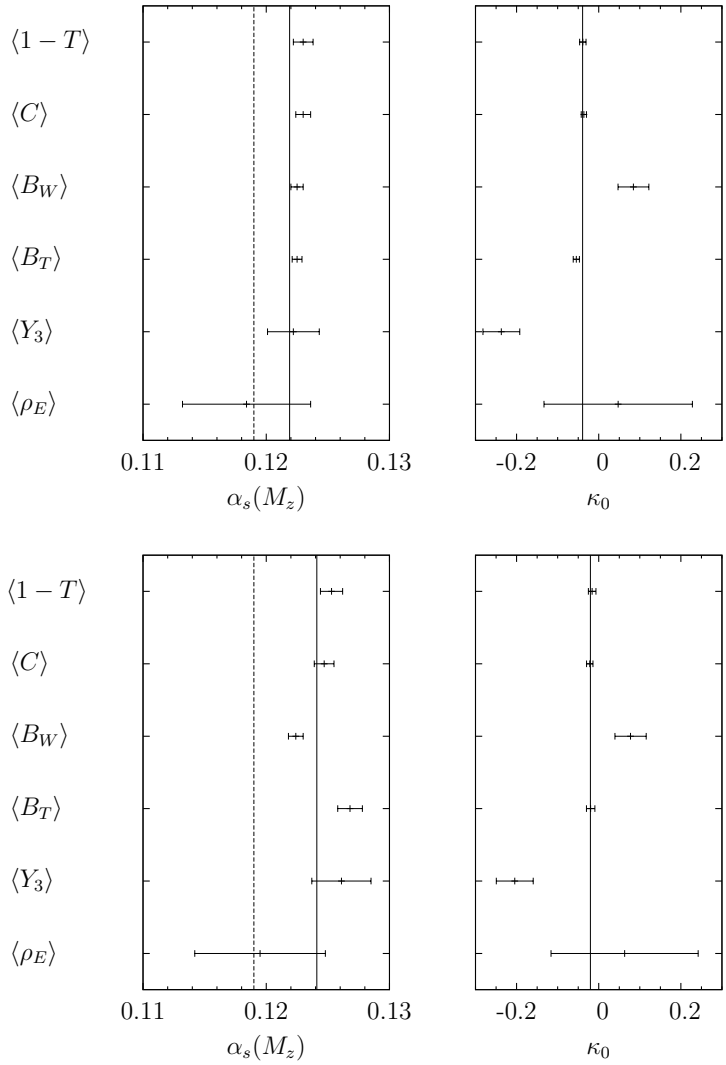


Figure 3.21: Scatter plots of $\alpha_s(M_Z)$ and κ_0 from fits of ECH and simple power corrections for the means at NLO (top two panels) and NNLO (bottom two panels). The dotted lines on the $\alpha_s(M_Z)$ plots show the value of the coupling obtained from N^{3}LO calculations. The solid lines are unweighted averages.

data by $\overline{\rm MS}$ PT plus dispersive power corrections is generally better than the corresponding ECH fits, as can be seen from the χ^2 values in Tables 3.12 and 3.13.

In Fig. 3.19, when $\alpha_s(M_Z)$ is held fixed and only α_0 fitted for, good agreement is found for ECH at NLO. There is now less discrepancy between $\langle 1-T\rangle$ and $\langle C\rangle$ and the other three observables. However the χ^2 values for these two

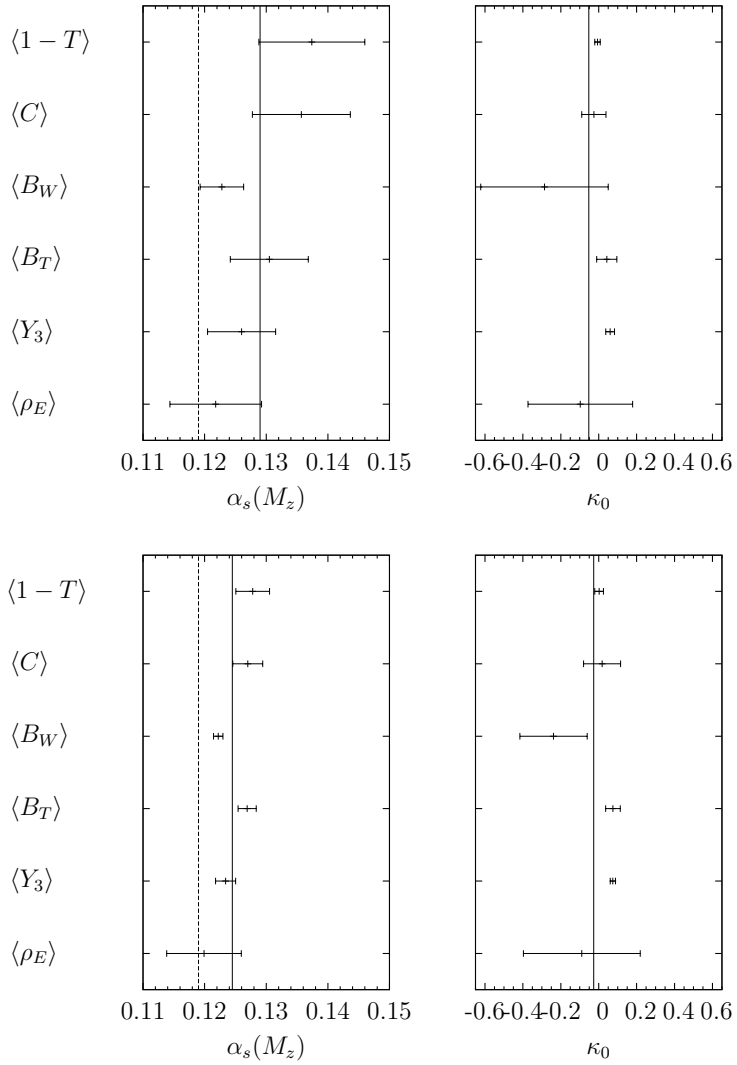


Figure 3.22: Scatter plots of $\alpha_s(M_Z)$ and κ_0 from fits of $\overline{\rm MS}$ PT and simple power corrections for the means at NLO (top two panels) and NNLO (bottom two panels). The dotted and solid lines are as in Fig. 3.21.

fits are very large compared with the number of degrees of freedom, indicating a poor quality of fit. At NNLO the results are more scattered and the values of α_0 required are larger. Again, the reduced χ^2 values for the NNLO fits are generally worse than those for the NLO fits.

For NLO $\overline{\rm MS}$ PT, in Fig. 3.20, there is good agreement in α_0 but the error bars are large. As before the uncertainties decrease at NNLO, as do the reduced

χ^2 values. From the χ^2 values in Tables 3.14 and 3.15 we see that at NLO the quality of the fits using the ECH method is comparable with or better than those using $\overline{\text{MS}}$ PT. At NNLO $\overline{\text{MS}}$ PT describes the data for $\langle 1-T \rangle$ and $\langle C \rangle$ much better than the ECH method.

In Fig. 3.21 in Section 3.5.2, when the simple power correction model is added to ECH at NLO, very good agreement is found between the extracted values of $\alpha_s(M_Z)$. The fits for κ_0 are generally close to zero, again showing that NLO ECH does not require large power corrections; this is an observation that was also made in the DELPHI analysis [35]. Slight differences in our fit results from the DELPHI collaboration's paper are most probably due to differences in the data analysed and the fitting programs used. Compared with the results for pure ECH from Fig. 3.1 we see that adding a simple power correction brings the values of $\alpha_s(M_Z)$ extracted from the outliers $\langle B_W \rangle$ and $\langle Y_3 \rangle$ into better agreement, but also increases the unweighted average of $\alpha_s(M_Z)$. At NNLO the grouping is less good and the average of $\alpha_s(M_Z)$ is increased even further.

The fitted values of κ_0 when using $\overline{\text{MS}}$ PT, shown in Fig. 3.22, are also generally close to zero. The corresponding values of $\alpha_s(M_Z)$ have therefore not altered substantially from when pure $\overline{\text{MS}}$ PT (see Fig. 3.1) is used. The values of χ^2 for the ECH method and $\overline{\text{MS}}$ PT are very similar, implying that the descriptions of each observable given by the two methods are comparable.

Chapter 4

The Principle of Minimal Sensitivity

We now turn our attention to another method that attempts to deal with unphysical RS dependence: the Principle of Minimal Sensitivity (PMS) [16]. This method tries to find values of $\{\tau, c_2, \ldots\}$[1] where the variable being studied is locally insensitive (or minimally sensitive) to the renormalisation scheme. The minimal sensitivity of this point replicates the behaviour at all orders, as a complete description of a physical observable should be RS-independent. We describe the PMS method in Section 4.1, and apply it to give predictions of e^+e^- event shape moment data in Section 4.2. The results are compared with those obtained from the ECH method and data at M_Z. In the next chapter we go on to apply the Principle of Minimal Sensitivity numerically to hadronic observables.

[1] These parameters label the renormalisation scheme — see Section 1.3.

4.1 The PMS method

In this section we describe the Principle of Minimal Sensitivity and how it can be applied. The description of PMS is based closely on that given in Ref. [16]. As stated before, a perturbative expansion that is truncated at a given order will generally be renormalisation scheme dependent. The rationale behind PMS states that, at any given order, one should find an optimum expansion that is locally insensitive to the parameters of the renormalisation scheme. This replicates the behaviour of the all orders result, which should not vary with the renormalisation scheme (RS) since this is merely an arbitrary mathematical tool: physical observables should be scheme independent quantities. The requirement of independence to the RS can be written symbolically as:

$$\left.\frac{\partial O}{\partial(\mathrm{RS})}\right|_{\text{Opt. RS}} = 0, \tag{4.1.1}$$

where O is an observable and Opt. RS denotes the optimum renormalisation scheme that gives no scheme dependence in the truncated perturbative expansion of O [16].

The RS can be labelled by $\{\tau, c_2, c_3, \ldots\}$ [16], where $\tau = b \ln \frac{\mu}{\tilde{\Lambda}}$ and $c_2, c_3, \ldots$ are higher coefficients in the β-function equation (1.2.18), as demonstrated in Section 1.3. The number of c_n coefficients required to define the RS depends on the order of perturbation theory. At NLO the RS is defined solely by τ. At $\mathcal{O}(a^{n+1})$ the RS is labelled by $\{\tau, c_2, \ldots, c_n\}$.

The self-consistency of perturbation theory states that approximations of O, to i^{th} order, in different renormalisation schemes must agree to $\mathcal{O}(a^{i+1})$. This can be written as:

$$\frac{\partial O^{(i)}}{\partial(\mathrm{RS})} = \mathcal{O}(a^{i+1}), \tag{4.1.2}$$

where the (i) denotes that the observable is expanded up to orders of a^i.

In the next two sub-sections we outline the PMS method at NLO and NNLO.

4.1.1 PMS at NLO

Consider an observable with an expansion like that of the effective charges considered in Section 2.1. At second order:

$$\mathcal{R}^{(2)} = a(\tau)(1 + r_1 a(\tau)), \tag{4.1.3}$$

where the superscript (2) denotes that we are considering the expansion of $\mathcal{R}$ up to $\mathcal{O}(a^2)$. We want to find $\mathcal{R}^{(2)}_{\text{opt}} = \mathcal{R}^{(2)}(\tau_{\text{opt}})$ such that:

$$\left.\frac{\partial \mathcal{R}^{(2)}}{\partial \tau}\right|_{\tau=\tau_{\text{opt}}} = 0. \tag{4.1.4}$$

Differentiating (4.1.3) with respect to τ gives:

$$\frac{\partial \mathcal{R}^{(2)}}{\partial \tau} = -a^2(1 + ca)(1 + 2r_1 a) + \frac{\partial r_1}{\partial \tau} a^2, \tag{4.1.5}$$

where we have substituted in the NLO β-function, from (1.2.18) in Section 2.1, noting that $\dfrac{\partial a}{\partial \tau} = \dfrac{1}{b}\dfrac{\partial a}{\partial \ln \mu}$.

Using the self consistency of perturbation theory, as shown in (4.1.2), the terms of $\mathcal{O}(a^2)$ in the above equation must cancel. This gives us:

$$\frac{\partial r_1}{\partial \tau} = 1. \tag{4.1.6}$$

Integrating this equation and rearranging for the constant of integration, which

we shall call X_1, gives:

$$X_1 = \tau - r_1. \tag{4.1.7}$$

X_1 is a scheme invariant, Q-dependent, quantity. The Q-dependence of X_1 is obtained from that of r_1 [9][2]:

$$r_1 = b \ln \frac{\mu_{\rm R}}{\tilde{\Lambda}} - b \ln \frac{Q}{\Lambda_{\mathcal{R}}}\,, \tag{4.1.8}$$

and the definition of τ in (1.3.2). Putting them together gives:

$$X_1 = b \ln \frac{Q}{\Lambda_{\mathcal{R}}}\,, \tag{4.1.9}$$

and so this scheme invariant quantity is related to that of the ECH method, $\Lambda_{\mathcal{R}}$.

We want to find the value of τ for which the right hand side of (4.1.5) is zero:

$$2\,c\,r_1(\tau_{\rm opt})\,a(\tau_{\rm opt}) + 2r_1(\tau_{\rm opt}) + c = 0 \tag{4.1.10}$$

To condense expressions such as this we now adopt the notation of Ref. [16], where $\tau_{\rm opt} = \bar{\tau}$, $r_1(\bar{\tau}) = \bar{r}_1$ and $a(\bar{\tau}) = \bar{a}$. Rewriting (4.1.10) in the new notation, and rearranging for $\bar{r}_1$, gives:

$$\bar{r}_1 = \frac{-c}{2(1 + c\bar{a})}. \tag{4.1.11}$$

We now just need to find $\bar{a}$. We can do this by substituting $\bar{r}_1 = \bar{\tau} - X_1$ in to (4.1.10):

$$2(\bar{\tau} - X_1)(c\bar{a} + 1) + c = 0, \tag{4.1.12}$$

[2]This equation is obtained by rearranging (2.1.18), and then generalising to the case where $\mu_{\rm R} \neq Q$.

and using:

$$\tau = \frac{1}{a} + c \ln\left(\frac{ca}{1+ca}\right). \tag{4.1.13}$$

This equation is obtained by integrating the NLO β-function equation, as described in Section 1.2.3. Putting this expression for τ into (4.1.12) gives:

$$\frac{1}{\bar{a}} + c \ln\left(\frac{c\bar{a}}{1+c\bar{a}}\right) + \frac{c}{2(1+c\bar{a})} = X_1, \tag{4.1.14}$$

which can be solved numerically for $\bar{a}$.

Now that we have $\bar{a}$ and $\bar{r}_1$ (from (4.1.11)) we can put them together to get an optimum, renormalisation scale independent, formulation of $\mathcal{R}$:

$$\mathcal{R}^{(2)}_{\text{opt}} = \bar{a}\,(1 + \bar{r}_1 \bar{a})\,. \tag{4.1.15}$$

4.1.2 PMS at NNLO

We now consider PMS at NNLO, with the observable $\mathcal{R}$ terminated at $\mathcal{O}(a^3)$:

$$\mathcal{R}^{(3)} = a(1 + r_1 a + r_2 a^2). \tag{4.1.16}$$

At NNLO the RS is defined by τ and c_2. The PMS point is therefore determined by the following two equations:

$$\left.\frac{\partial \mathcal{R}^{(3)}}{\partial \tau}\right|_{\tau=\bar{\tau}} = 0 \qquad \text{and} \qquad \left.\frac{\partial \mathcal{R}^{(3)}}{dc_2}\right|_{c_2=\bar{c}_2} = 0. \tag{4.1.17}$$

To derive PMS at NNLO we will need the β-function equation to NNLO:

$$\frac{\partial a}{\partial \tau} = -a^2(1 + ca + c_2 a^2), \tag{4.1.18}$$

and the equivalent expression for c_2:

$$\frac{\partial a}{\partial c_2} = a^2(1 + ca + c_2a^2) \int_0^a \frac{dx}{(1 + cx + c_2x^2)^2}. \tag{4.1.19}$$

Differentiating $\mathcal{R}^{(3)}$ with respect to τ and c_2 and substituting in the equations above gives:

$$\frac{\partial \mathcal{R}^{(3)}}{\partial \tau} = -a^2(1 + ca + c_2a^2)(1 + 2r_1a + 3r_2a^2) + a^2\frac{\partial r_1}{\partial \tau} + a^3\frac{\partial r_2}{\partial \tau}, \tag{4.1.20}$$

$$\frac{\partial \mathcal{R}^{(3)}}{\partial c_2} = a^2(1 + ca + c_2a^2) \int_0^a \frac{dx}{(1 + cx + c_2x^2)^2}(1 + 2r_1a + 3r_2a^2) + a^2\frac{\partial r_1}{\partial c_2} + a^3\frac{\partial r_2}{\partial c_2}. \tag{4.1.21}$$

We can again use the requirement of self-consistency of perturbation theory to cancel the $\mathcal{O}(a^2)$ and $\mathcal{O}(a^3)$ terms. From (4.1.20) we have:

$$\frac{\partial r_1}{\partial \tau} = 1, \qquad \text{and} \qquad \frac{\partial r_2}{\partial \tau} = c + 2r_1. \tag{4.1.22}$$

For (4.1.21) we first have to Taylor expand the integrand:

$$\int_0^a \frac{1}{(1 + cx + c_2x^2)^2} \approx \int_0^a (1 - 2cx + \ldots) = a - ca^2 + \ldots. \tag{4.1.23}$$

The self-consistency of perturbation theory therefore requires that:

$$\frac{\partial r_1}{\partial c_2} = 0, \qquad \text{and} \qquad \frac{\partial r_2}{\partial c_2} = -1. \tag{4.1.24}$$

The two equations referring to r_1 integrate up to give (4.1.7) again. The two involving r_2 can also be integrated to give a second RS invariant quantity.

First we rewrite $\dfrac{\partial r_2}{\partial \tau}$ as:

$$\frac{\partial r_2}{\partial \tau} = \frac{\partial r_1}{\partial \tau}\frac{\partial r_2}{\partial r_1} = \frac{\partial r_2}{\partial r_1} = c + 2r_1\,. \tag{4.1.25}$$

This then integrates up to:

$$r_2 = cr_1 + r_1^2 + f(c_2) + \text{const.}\,, \tag{4.1.26}$$

where const. is the integration constant. f is a function of c_2 and its form can be determined from the other partial differential equation, $\dfrac{\partial r_2}{\partial c_2}$. Putting this together, and calling the RS invariant integration constant X_2, gives:

$$X_2 = r_2 + c_2 - r_1^2 - r_1 c. \tag{4.1.27}$$

Note that X_2 is exactly the same as the coefficient ρ_2 from the Effective Charges Method (see (2.1.6)). At higher orders the X_n quantities remain closely related to the ρ_n coefficients of the ECH method, although they are only identical for $n = 2$. For example, $X_3 = \rho_3/2$ and $X_4 = \rho_4/3 + c\rho_3/6 + 2\rho_2$ [65].

At NNLO we want to find the quantites $\bar{\tau}$ and $\bar{c}_2$ for which (4.1.20) and (4.1.21) are zero:

$$(\bar{c}_2 + 3\bar{r}_2 + 2c\bar{r}_1) + (2\bar{c}_2\bar{r}_1 + 3c\bar{r}_2)\bar{a} + 3\bar{c}_2\bar{r}_2\bar{a}^2 = 0, \tag{4.1.28}$$

$$\begin{aligned}\int_0^{\bar{a}} \frac{dx}{(1 + cx + \bar{c}_2 x^2)^2} &= \frac{\bar{a}}{(1 + (c + 2\bar{r}_1)\bar{a} + (\bar{c}_2 + 3\bar{r}_2 + 2c\bar{r}_1)\bar{a}^2 + (2\bar{c}_2\bar{r}_1 + 3c\bar{r}_2)\bar{a}^3 + 3\bar{c}_2\bar{r}_2\bar{a}^4)} \\ &= \frac{\bar{a}}{1 + (c + 2\bar{r}_1)\bar{a}}.\end{aligned} \tag{4.1.29}$$

where we have used (4.1.28) to simplify the denominator in (4.1.29). We can

obtain expressions for $\bar{r}_1$ and $\bar{r}_2$ from (4.1.7) and (4.1.27):

$$\begin{aligned}\bar{r}_1 &= \bar{\tau} - X_1 \\ &= \frac{1}{\bar{a}} + c\ln\left(\frac{c\bar{a}}{1+c\bar{a}}\right) + \bar{c}_2\int_0^{\bar{a}} \frac{dx}{(1+cx)(1+cx+\bar{c}_2x^2)} - X_1, \qquad (4.1.30)\end{aligned}$$

$$\bar{r}_2 = X_2 - \bar{c}_2 + (\bar{\tau} - X_1)^2 + (\bar{\tau} - X_1)c, \qquad (4.1.31)$$

where the NNLO version of (4.1.13) is used for $\bar{\tau}$.[3]

Now that we have expressions for $\bar{r}_1$ and $\bar{r}_2$ we can solve the simultaneous equations (4.1.28) and (4.1.29) for $\bar{c}_2$ and $\bar{a}$ numerically. Substituting these into (4.1.30) and (4.1.31) the optimum observable at NNLO can be constructed [16]:

$$\mathcal{R}^{(3)}_{\rm opt} = \bar{a}(1 + \bar{r_1}\bar{a} + \bar{r_2}\bar{a}^2). \qquad (4.1.32)$$

4.2 PMS applied to event shape moments

In this section we use the PMS method in order to predict to event shape moments from e^+e^- collisions. PMS is implemented at both NLO and NNLO and the results are shown in Tables 4.1 and 4.2. As in Section 3.2 the predictions are obtained from a fixed value of $\alpha_s(M_Z) = 0.1190$. From the corresponding value of $\tilde{\Lambda}_{\overline{\rm MS}}$, $\tau^{(\overline{\rm MS})}$ is calculated as:

$$\tau^{(\overline{\rm MS})} = b\ln\frac{M_Z}{\tilde{\Lambda}_{\overline{\rm MS}}}, \qquad (4.2.1)$$

where $\mu_{\rm R} = Q$ since we are working in the $\overline{\rm MS}$ scheme, and we consider the energy M_Z.[4] The RS independent quantity X_1 is then calculated from τ and

[3] See (2.1.11) in Section 2.1 for analogous expressions containing $\mathcal{R}$ and ρ_2 instead of a and c_2.

[4] τ, and hence X_1, is Q-dependent, but we restrict ourselves to examining $Q = M_Z$ in this analysis. PMS could easily be extended to give predictions at other energies.

r_1 in the $\overline{\text{MS}}$ scheme:

$$X_1 = \tau^{(\overline{\text{MS}})} - r_1^{(\overline{\text{MS}})} . \tag{4.2.2}$$

Similarly, for the NNLO case, X_2 can be found from the $\overline{\text{MS}}$ coefficients r_1, r_2 and c_2 from (4.1.27).[5]

For PMS at NLO, once X_1 has been found, $\bar{a}$ can be extracted numerically from (4.1.14). The optimal prediction for the event shape can then be constructed from (4.1.11) and (4.1.15).

At NNLO, $\bar{a}$ and $\overline{c}_2$ from (4.1.28) and (4.1.29) are simultaneously extracted. These values are used along with (4.1.32) to give a PMS prediction of the event shape at $Q = M_Z$.

For comparison, experimental data at 91.3 GeV and the ECH prediction of the event shape at M_Z are also shown in both tables. It can be seen that the PMS and ECH predictions are very similar, and that at NLO they are also close to the experimental data. The predictions are generally not as close to the data at NNLO.

For $\langle B_T^3 \rangle$ at NLO, unlike the other observables, it is not possible to extract $\bar{a}$ at $Q = M_Z$. $\bar{a}$ can be extracted at a higher energy, but when the coupling is run down to M_Z an imaginary value for $\bar{a}$ is obtained. This is because the value for $\Lambda_{\overline{\text{MS}}}$ that corresponds to the $\bar{a}$ extracted is larger than M_Z. Trying to run the coupling down to M_Z involves moving out of the perturbative regime, and hence results in an unrealistic value. A PMS prediction for $\langle B_T^3 \rangle$ is therefore not included in Table 4.1. This is also the case for some variables and energies for the ECH method at NLO, as discussed in Sections 3.2 and 3.3.1.

[5]The two invariants X_1 and X_2 could be calculated from coefficients in any scheme, but it is convenient to use the $\overline{\text{MS}}$ scheme since many calculations are performed in this RS.

Observable	n	$\overline{\alpha}_s(M_Z)$	$\overline{\langle y \rangle} \times 10^2$	$\langle y \rangle_{\text{exp}} \times 10^2$	$\langle y \rangle_{\text{ECH}} \times 10^2$
$\langle (1-T)^n \rangle$	1	0.202	6.51	6.55 ± 0.07	6.50
	2	0.340	0.97	0.79 ± 0.01	0.96
	3	0.488	0.21	0.141 ± 0.003	0.21
$\langle C^n \rangle$	1	0.193	25.5	25.9 ± 0.2	25.5
	2	0.351	12.7	10.2 ± 0.2	12.7
	3	0.590	9.16	5.3 ± 0.1	9.07
$\langle B_W^n \rangle$	1	0.112	7.07	7.24 ± 0.07	7.07
	2	0.159	0.83	0.76 ± 0.01	0.82
	3	0.161	0.12	0.104 ± 0.002	0.12
$\langle B_T^n \rangle$	1	0.169	10.6	10.58 ± 0.07	10.6
	2	1.14	5.17	1.49 ± 0.02	5.02
	3	-	-	0.263 ± 0.004	-
$\langle Y_3^n \rangle$	1	0.162	2.23	2.05 ± 0.05	2.23
	2	0.170	0.21	0.194 ± 0.005	0.21
	3	0.168	0.03	0.030 ± 0.001	0.03
$\langle \rho_E^n \rangle$	1	0.149	4.86	4.95 ± 0.03	4.85

Table 4.1: Results for applying PMS at NLO to e^+e^- event shape moments. Values of α_s at the stationary point are shown alongside a prediction for the effective charge at M_Z, $\overline{\langle y \rangle}$. Experimental values of the effective charge at M_Z and predictions using ECH at NLO are also shown for comparison.

Observable	n	$\overline{\alpha}_s(M_Z)$	$\overline{c}_2$	$\overline{\langle y \rangle} \times 10^2$	$\langle y \rangle_{\rm exp} \times 10^2$	$\langle y \rangle_{\rm ECH} \times 10^2$
$\langle (1-T)^n \rangle$	1	0.185	-19.6	6.20	6.55 ± 0.07	6.21
	2	0.226	-72.3	0.69	0.79 ± 0.01	0.70
	3	0.232	-100	0.11	0.141 ± 0.003	0.12
$\langle C^n \rangle$	1	0.179	-12.8	24.8	25.9 ± 0.2	24.8
	2	0.223	-82.8	8.64	10.2 ± 0.2	8.76
	3	0.232	-149	4.43	5.3 ± 0.1	3.77
$\langle B_W^n \rangle$	1	0.110	4.58	7.09	7.24 ± 0.07	7.09
	2	0.148	-23.8	0.80	0.76 ± 0.01	0.80
	3	0.148	-31.0	0.11	0.104 ± 0.002	0.11
$\langle B_T^n \rangle$	1	0.154	-34.1	10.0	10.58 ± 0.07	10.0
	2	0.145	-360	0.81	1.49 ± 0.02	0.86
	3	0.110	-736	0.09	0.263 ± 0.004	0.10
$\langle Y_3^n \rangle$	1	0.151	-30.5	2.13	2.05 ± 0.05	2.13
	2	0.154	-32.6	0.20	0.194 ± 0.005	0.20
	3	0.154	-36.5	0.03	0.030 ± 0.001	0.03
$\langle \rho_E^n \rangle$	1	0.145	-10.1	4.79	4.95 ± 0.03	4.79

Table 4.2: Results for applying PMS at NNLO to e^+e^- event shape moments. Values of α_s and c_2 at the stationary point are shown alongside a prediction for the effective charge at M_Z, $\overline{\langle y \rangle}$. Experimental values of the effective charge at M_Z and predictions using ECH at NNLO are also shown for comparison.

Chapter 5

Applying PMS to hadronic observables

In the previous chapter the Principle of Minimal Sensitivity was introduced, and applied analytically to produce predictions for event shape moments. It was found to give results that are very similar to those obtained from the ECH method, despite the two methods being motivated by different philosophies and resulting in different master equations.

We now proceed to apply PMS to hadronic observables. Physical observables at hadron colliders are generally more complicated than those measured at e^+e^- colliders, since parton distribution functions (PDFs) are needed to describe the incoming hadrons. This results in an additional scale, the factorisation scale M, that needs to be taken into account in addition to the renormalisation scale and scheme. PMS can no longer be applied analytically to observables of this type, but it is possible to look for a stationary point on the surface of the observable in M–$\mu_{\rm R}$ space. A stationary point on this surface is locally invariant with respect to M and $\mu_{\rm R}$, which satisfies the PMS philosophy.

In Section 5.1 we describe in more detail the application of PMS to hadronic observables. The results obtained from applying PMS to heavy quark cross-sections at the Tevatron and the LHC are given in Sections 5.2 and 5.3. Section 5.4 contains a discussion of the results from this chapter and those from Section 4.2 from the end of Chapter 4.

5.1 PMS for processes involving hadrons

We have so far considered applying PMS to observables that depend only on the renormalisation scheme. An example of this type is the event shape moments measured at e^+e^- colliders. In this part of the thesis we apply PMS to various Tevatron and LHC cross-sections, including the $b\bar{b}$ total cross-section at the Tevatron, and the $t\bar{t}$ total cross-section at the Tevatron and the LHC.

Predictions of observables with hadrons in the initial state are more complicated than those from e^+e^- collisions as they depend on parton distribution functions (PDFs) [2, 3]. The PDFs are needed to describe the internal structure of the incoming hadrons and are essentially non-perturbative in nature. Long- and short-distance contributions are separated out into the PDFs and the hard partonic cross-section respectively. PDFs are not calculable in perturbative QCD, although their evolution with M is, and have to be extracted from experimental data.

There are low energy divergences that arise when a quark or gluon is emitted colinearly to an incoming parton. These initial state IR singularities are dealt with in a similar fashion to the way in which we renormalised UV divergences in Section 1.2 — they are absorbed into a redefinition of the PDFs [2]. This procedure introduces an arbitrary, unphysical factorisation scale, M, which is analogous to μ_{R}. A factorisation scheme is also needed to specify

the exact contributions that are absorbed by the PDFs. The $\overline{\text{MS}}$ scheme is frequently used.

A generic total cross-section from a hadronic collision can be written as:

$$\sigma(S) = \sum_{i,j} \int dx_1 dx_2 F_i^A(x_1, M) F_j^B(x_2, M)\, \hat{\sigma}(\hat{s}, M), \tag{5.1.1}$$

where $\sqrt{S}$ is the centre of mass energy, F_i^A is the PDF for parton i from hadron A, x_1 and x_2 are the fractions of momentum carried by each parton, and M is the factorisation scale. $\hat{\sigma}$ is the partonic cross-section which depends on the fraction of the centre of mass energy $\hat{s} = x_1 x_2 S$. Since $\hat{\sigma}$ contains only short distance physics after factorisation, and at short distances (or high energies) the strong coupling constant will be small, perturbation theory can be used to expand the partonic cross-section. At fixed order $\hat{\sigma}$ will also have a dependence on μ_{R}. The M dependence of the PDFs will cancel with that of $\hat{\sigma}$ at all orders.

The scales μ_{R} and M are often set equal to each other when performing calculations [2], but their physical origins are very different. M is related to singularities at low energy, and μ_{R} regulates high energy divergences. Setting them equal to each other is therefore not physically motivated. When we apply PMS to hadronic cross-sections we keep M and μ_{R} separate.

Applying PMS to observables of this type involves numerically solving for the values of M and μ_{R} at which the cross-section, $\sigma(M, \mu_{\text{R}})$, is locally invariant with respect to these two observables. This corresponds to a stationary point of the surface of σ in M and μ_{R} space. It is not possible to find the stationary points analytically as we do not have an explicit expression for the PDFs.

As stated above, the procedure of factorisation is to some extent arbitrary and a scheme is required to specify what finite pieces are absorbed along with the singularity. Beyond LO, results are dependent on the factorisation scheme

used and in order to perform a full PMS analysis of the cross-section we should take this into account too. However, finding a stationary point while varying three different parameters would be considerably harder, so in this analysis we restrict ourselves to considering the two scales M and μ_{R}.

There is no guarantee that there is a unique stationary point, or indeed any stationary point at all, on a given surface [16]. At leading order (LO) the cross-section is generally monotonic in μ_{R} and there is therefore no stationary point [66]. At NLO however there is often a saddle point observed at some values of M and μ_{R}, as shown for the $b\bar{b}$ total cross-section in Ref. [66].

In this thesis we have looked for stationary points with respect to M and μ_{R} for two different observables:

1. $b\bar{b}$ total cross-section at the Tevatron,

2. $t\bar{t}$ total cross-section at the Tevatron and the LHC.

In both these cases MCFM [67–69] was used to calculate the cross-sections. For the total $t\bar{t}$ cross-section, the Top++ program [70] was also used. The saddle points were found by minimising the function:[1]

$$f = \frac{1}{2}\left(\frac{\partial\sigma}{\partial \ln M}\right)^2 + \frac{1}{2}\left(\frac{\partial\sigma}{\partial \ln \mu_{\mathrm{R}}}\right)^2 . \tag{5.1.2}$$

Note that, in this thesis, the terms stationary points and saddle points are often used interchangeably as in general the stationary points on the cross-section surface are saddle points. However the gradients in f are squared so that all stationary points will be found by minimising this function, including minima and maxima.

[1] In practice the gradients in this equation were often normalised to the cross-section — see Section C.3 in Appendix C.

5.1.1 Algorithms for finding stationary points

MCFM uses a Monte-Carlo integrator which requires several thousand event generations to produce cross-sections with sufficiently low statistical errors. We generally run with 4000 events for total cross-sections. This takes about 4 minutes per MCFM calculation and it is clearly desirable to make as few function calls to MCFM as possible. We therefore wrote our own algorithm, implemented in Python [71], based on a secant Levenberg-Marquadt algorithm with BFGS updates [72, 73].

The Levenberg-Marquadt algorithm is a hybrid method combining the method of steepest descent and the Gauss-Newton method [72, 73]. There is a damping parameter that governs how much of each method to use. When the damping parameter is large a small step is taken in the direction of steepest descent. As the stationary point is approached the damping parameter becomes smaller and the step taken approaches the Gauss-Newton direction. This, in general, provides better convergence close to the stationary point.

The advantage of using a secant method is that the Jacobian does not have to be found at every step. The Jacobian for the initial starting point, or an approximation to it, is provided at the start and this is then updated during each iteration. There are various ways of doing this: we chose to use a BFGS (Broyen-Fletcher-Goldfarb-Shanno) update as this preserves any initial symmetry in the matrix [73]. This property is useful since our Jacobian will contain second order derivatives, given that f in (5.1.2) is already a function of derivatives. The commutativity of partial derivatives means that the Jacobian will be symmetric.

The details of the algorithm and the BFGS update are given in Appendix C. Occasionally a built-in Python algorithm was used as a cross-check on the results obtained from our program.

5.2 Results for the total $b\bar{b}$ cross-section

An initial study, used to develop and check our algorithm, was performed to identify the stationary points of the total $b\bar{b}$ cross-section surface. The cross-sections were calculated at NLO for $p\bar{p}$ collisions at energies from 62 GeV to 1.96 TeV. These energies are used by Chýla in Ref. [66], and this reference serves as a cross-check for our results. We have used more recent PDFs than Chýla however, and it is therefore interesting to determine whether the cross-sections obtained vary significantly as a result of this.

In Fig. 5.1 the total cross-section surface at 630 GeV is plotted against M and μ_{R} at both LO and NLO. It can be seen that at LO there is no stationary point in the range of M and μ_{R} studied. At NLO, however, there is a saddle point at $M \sim 15$ GeV and $\mu_{\mathrm{R}} \sim 1$ GeV.

Fig. 5.2 demonstrates how the saddle point moves with $\sqrt{S}$. The optimal values of M and μ_{R} according to the PMS method are generally far from the standard diagonal choice of $M = \mu_{\mathrm{R}} = m_b$. For comparison, the cross-section calculated using the PMS choices of scale is plotted against the standard diagonal choice in Fig. 5.3. We also plot the ratio of the PMS cross-section (i.e. the cross-section at the stationary point) to the diagonal cross-section in Fig. 5.4. R_{PMS} is defined as:

$$R_{\mathrm{PMS}} = \frac{\sigma(\overline{M}, \overline{\mu}_R)}{\sigma(M = \mu_{\mathrm{R}} = m_b)}, \tag{5.2.1}$$

where $\overline{M}$ and $\overline{\mu}_R$ are the values of M and μ_{R} at the stationary point. $R_{1/2}$ and R_2 are also shown, as in Fig. 7 of Ref. [66], where these quantities are defined

as:

$$R_{1/2} = \frac{\sigma(M = \mu_{\rm R} = 0.5 m_b)}{\sigma(M = \mu_{\rm R} = m_b)}, \tag{5.2.2}$$

$$R_2 = \frac{\sigma(M = \mu_{\rm R} = 2 m_b)}{\sigma(M = \mu_{\rm R} = m_b)}. \tag{5.2.3}$$

The data for these plots have been calculated using $m_b = 4.75$ GeV and the 2008 set of NLO MSTW PDFs [74]. The value of m_b was chosen to match that used in the MSTW PDF set. We include Figs. 5.2–5.4 for ease of comparison with Chýla's results in Ref. [66]. Generally we find that our results are similar to those in Ref. [66]. There are some differences however, which are probably due to our use of more recent PDFs and value of m_b. The differences include generally larger values of the NLO cross-section and a larger ratio $R_{\rm PMS}$.

Fig. 5.2 shows that the PMS scales move in a highly non-diagonal fashion. As the collision energy increases, the factorisation scale tends to higher values, while the renormalisation scale remains around 1-2 GeV. The PMS cross-section is consistently larger than that predicted by fixed order perturbation theory using diagonal scales, as seen in Fig. 5.3.

From Fig. 5.4 we see that at the Tevatron energy 1.96 TeV the cross-section as calculated using PMS is almost a factor of two larger than that from standard $\overline{\rm MS}$ perturbation theory using diagonal scale choices: 129 μb compared with 72 μb. There have been discrepancies between QCD predictions and experimental data for $b\bar{b}$ differential cross-sections at the Tevatron [75,76]. While the latest theoretical predictions do agree well with Run II CDF data [77, 78], it is necessary to use a next-to-leading log resummation in addition to a fixed-order NLO calculation, and improved b-fragmentation models. There are also still inconsistencies between the experimental measurements [76, 79]. As mentioned in Section 5.1, there is no physical motivation for setting $M = \mu_{\rm R}$

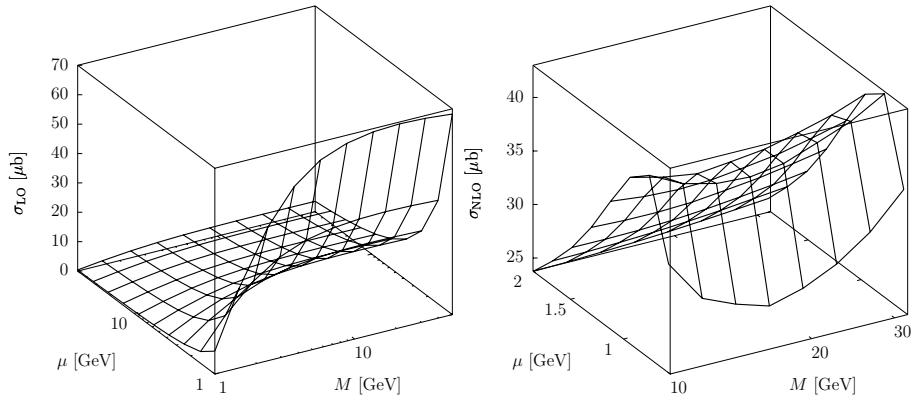


Figure 5.1: Total $b\bar{b}$ cross-section surface plotted against factorisation and renormalisation scale, calculated at 630 GeV at the Tevatron. The LO cross-section is shown on the left and the NLO cross-section on the right.

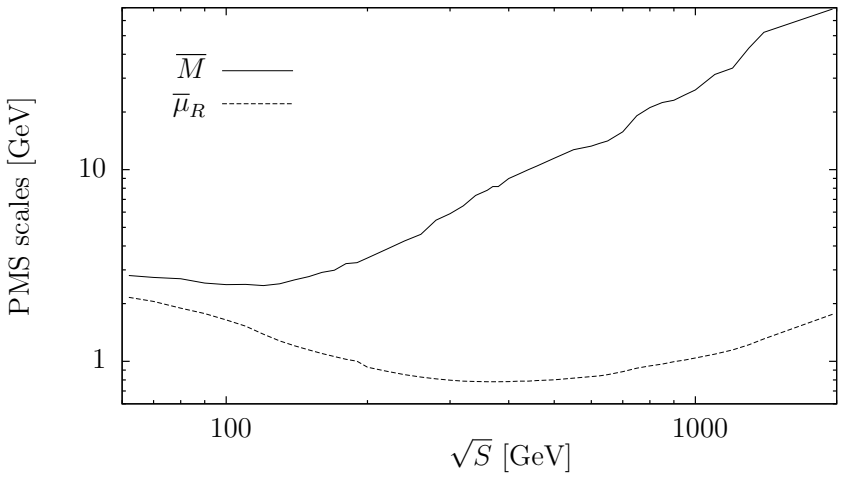


Figure 5.2: $\overline{M}$ and $\overline{\mu}_R$ at the saddle point are plotted over a range of collision energies.

and alternative scale setting procedures should be considered along with the other improvements to the QCD predictions discussed in e.g. Ref. [75].

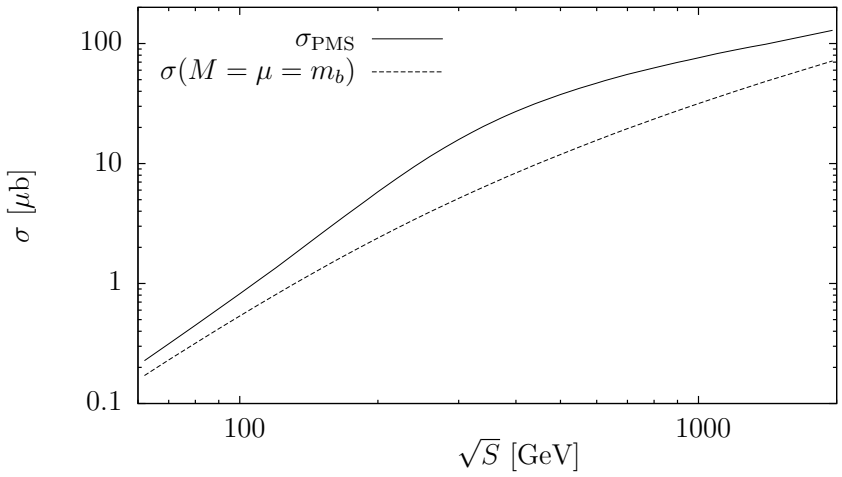


Figure 5.3: The total NLO cross-section at the saddle point ($\sigma_{\rm PMS}$) is shown, along with the cross-section for the conventional diagonal choice of $M = \mu_{\rm R} = m_b$.

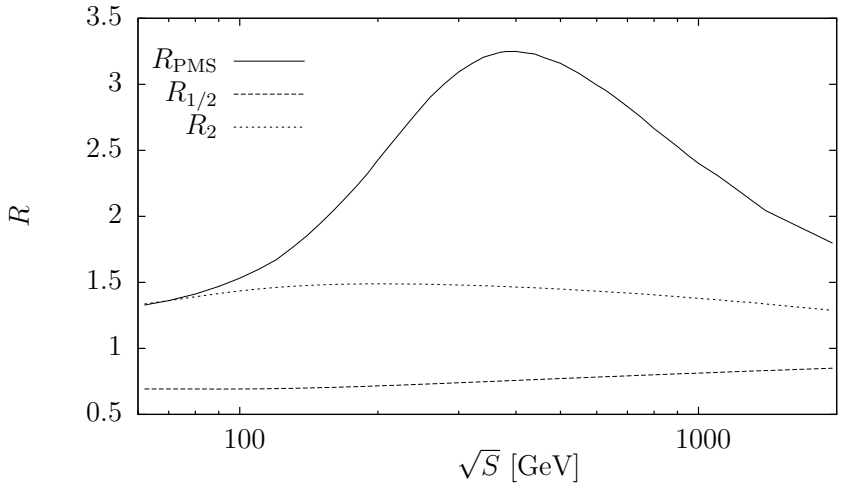


Figure 5.4: Ratio of PMS cross-section to standard $\overline{\rm MS}$ cross-section plotted against energy. Also shown is the scale variation in the standard cross-section when varying M and $\mu_{\rm R}$ by a factor of two.

5.3 Results for the total $t\bar{t}$ cross-section

In this section the analysis of $b\bar{b}$ pair production is extended to look at the total top quark pair production cross-section. Some example surfaces, for the NLO $t\bar{t}$ total cross-section at the Tevatron at 1.96 TeV and for the LHC at

7 TeV, are shown in the left-hand panels of Figs. 5.5 and 5.6. As for the $b\bar{b}$ cross-section, the secant Levenberg-Marquadt algorithm is used to find the stationary point on these surfaces, with the cross-sections provided by MCFM. In addition, the program Top++ [70] is used to calculate the cross-section. The stationary point for the NLO surface was again found and compared with the MCFM results. Top++ is also used to look at the preliminary NNLO results calculated by Czakon and Mitov [80]. For more details see Section 5.3.2.

5.3.1 MCFM results

At the Tevatron energy 1.96 TeV a saddle point was found at $M \sim 110$ GeV and $\mu_{\mathrm{R}} \sim 84$ GeV. This corresponds to a cross-section of 7.20 ± 0.01 pb, where the error is statistical and comes from the numerical integration in MCFM. The standard cross-section, calculated by setting $M = \mu_{\mathrm{R}} = m_t$, is $6.86 \pm 0.01 \pm 0.75$ pb. The first error is statistical and the second is obtained by varying the scales between $2m_t$ and $m_t/2$, with the larger difference from the central value taken as the theoretical error. Using the PMS method therefore gives a cross-section that is about 5% larger than the standard method does, which is well within the theoretical error on the latter value.

A saddle point for the surface created by cross-sections at 7 TeV at the LHC was found at $M \sim 63$ GeV and $\mu_{\mathrm{R}} \sim 41$ GeV, at a cross-section of 176.5 ± 0.3 pb. Again, the error is statistical. The cross-section using diagonal scales is $150.5 \pm 0.2 \pm 19.6$ pb. In this case, therefore, using the PMS method results in an increase of $\sim$17% which is larger than the upper scale variation of the diagonal scale.

For 14 TeV, a stationary point was found at $M \sim 80$ GeV and $\mu_{\mathrm{R}} \sim 27$ GeV. This gives a cross-section of 1024 ± 2 pb, compared with the diagonal value of $842 \pm 1 \pm 98$ pb. Using PMS for this energy gives an increase of 22% compared

with the diagonal scale choice, almost a 2σ discrepancy.

For all three energies, the PMS scales are found to be far from the diagonal choice of $M = \mu_{\mathrm{R}} = m_t$. The increase in cross-section predicted by PMS method for the LHC energies is larger than the scale variation on the standard cross-section prediction.

All these results are calculated using CT10 PDFs [81] and a top mass of 172.9 GeV.

These cross-sections can be compared to experimental results from the Tevatron and the LHC. At 1.96 TeV, D0 measures top quark pair production to be $7.56^{+0.63}_{-0.56}$pb, and CDF obtain 7.50 ± 0.48 pb [82]. These results assume a top mass of 172.5 GeV. However, the dependence on the mass is found to be less than the theoretical errors involved [82]. At 7 TeV at the LHC the total $t\bar{t}$ cross-section is measured by ATLAS as 179 ± 12 pb and by CMS as 164 ± 14 pb [82]. Both the standard diagonal scale choice and the PMS cross-section agree with these experimental results. However, the PMS cross-sections are closer to the central values of the experimental results than the standard scale choice.

5.3.2 Top++ results

In addition to using MCFM to calculate the $t\bar{t}$ total cross-section, we also used Top++ [70]. This program uses analytical methods rather than a Monte Carlo approach. It also includes some preliminary NNLO results; an exact fixed order NNLO calculation for the $q\bar{q} \to t\bar{t} + X$, $qq \to t\bar{t} + X$, $q\bar{q}' \to t\bar{t} + X$ and $qq' \to t\bar{t} + X$ channels [80], and an approximation to the NNLO result for $gg \to t\bar{t} + X$ [83]. The $q\bar{q}$ channel is dominant at the Tevatron. At the time of writing, the authors of Top++ are working on calculating the remaining channels exactly.

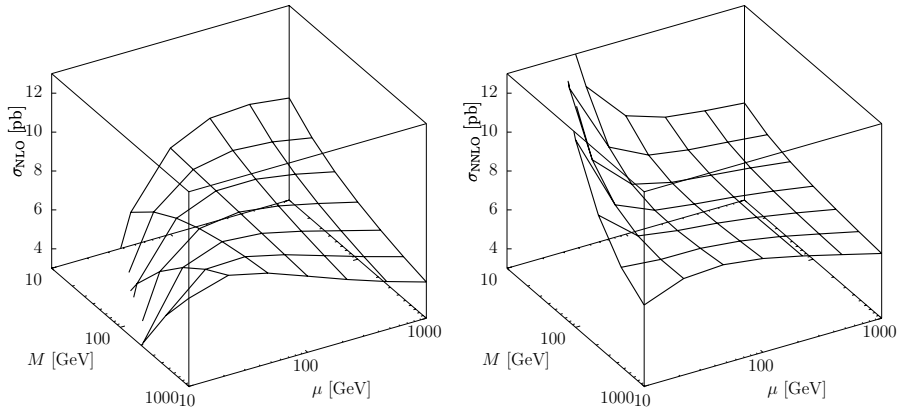


Figure 5.5: Total $t\bar{t}$ cross-section surface, calculated at 1.96 TeV at the Tevatron, plotted against factorisation and renormalisation scale. The NLO cross-section is shown on the left and NNLO on the right.

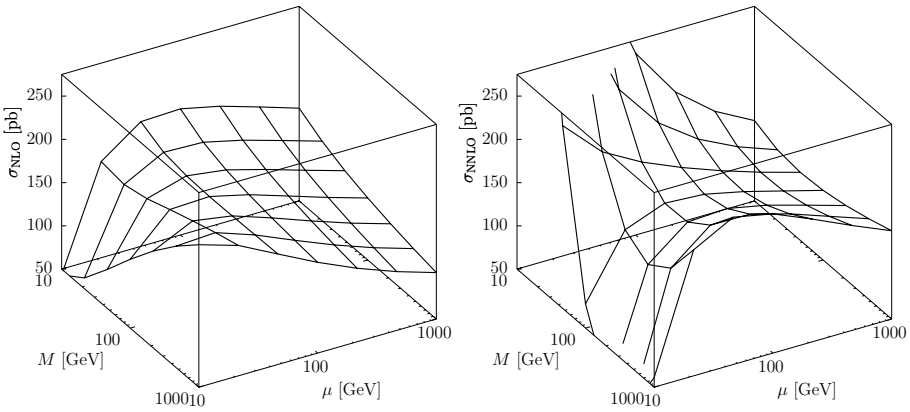


Figure 5.6: Total $t\bar{t}$ cross-section surface, calculated at 7 TeV at the LHC, plotted against factorisation and renormalisation scale. The NLO cross-section is shown on the left and NNLO on the right.

We used Top++ to provide a fixed order NLO calculation and compared the saddle points found with those obtained using MCFM. Since Top++ does not employ Monte Carlo methods each individual calculation takes less time to perform than when using MCFM. It was therefore possible to use a Levenberg-Marquadt algorithm, where the Jacobian is calculated explicity for each step of

the iteration, instead of using the secant approach as outlined in Appendix C. A comparison of the results obtained from MCFM and Top++ is given in Table 5.1. As expected both methods of calculation give very similar results.

A built-in Python function using a Nelder-Mead simplex algorithm [84] was also used in conjunction with Top++ as a further validation of the results, and found stationary points in similar regions as the Levenberg-Marquadt algorithm. The Nelder-Mead algorithm was not generally used with MCFM as it tends to require a larger number of function calls, rendering it inefficient given the greater run time for MCFM. However, when used in conjunction with Top++ it was found to be less sensitive to starting guesses for M and μ_{R} than the Levenberg-Marquadt algorithm, and has therefore been used for the remainder of this section.

In order to assess the effect of the choice of PDF set on our results we repeated the process of finding the stationary point using MSTW 2008 NLO PDFs. The stationary points are shown in Table 5.2, alongside the original CT10 results for ease of comparison. The saddle points were generally found to be in the same region, with the exception of the LHC energy 7 TeV where M increased by about 20 GeV and μ_{R} decreased by 5 GeV. However, even in this case the cross-section at the saddle point only changes by 7%. For the Tevatron at 1.96 TeV, the PMS cross-sections given by the two different PDF sets agree to within 1%. Although the stationary points at 14 TeV agree to within 1 GeV, the cross-section differs again by about 7%. The effect of changing these PDF sets on the NLO results is therefore seen to be negligible in the case of the Tevatron energy, and less than 10% for the LHC energies.

We also used Top++ to provide approximate NNLO results. At NNLO the RS can be defined by $\{\tau, c_2\}$, where τ is a function of μ_{R}, as discussed in Section 1.3. However, for hadronic observables it is also necessary to consider

	MCFM			Top++		
$\sqrt{S}$ [TeV]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]
1.96	110	84	7.20	110	84	7.18
7	63	41	177	63	41	177
14	80	27	1020	85	27	1020

Table 5.1: Comparison of stationary points found using MCFM and Top++ at NLO. CT10 PDFs are used in both cases, with $m_t = 172.9$ GeV.

	CT10			MSTW		
$\sqrt{S}$ [TeV]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]
1.96	110	84	7.18	110	83	7.12
7	63	41	177	64	42	188
14	85	27	1020	86	28	1090

Table 5.2: Comparison of stationary points found using CT10 and MSTW 2008 NLO PDFs with Top++ at NLO and $m_t = 172.9$ GeV.

the factorisation scale. In this analysis we therefore only examine the variation of the cross-section with M and μ_{R}, as in the NLO case.

Example surfaces are shown in the right-hand panels of Figs. 5.5 and 5.6. The NNLO surfaces are generally a little flatter than at NLO, especially for large M and μ_{R}, as would be expected from adding an extra order of perturbation theory. The stationary points of these surfaces are given in Table 5.3. The results were calculated using both CT10 and MSTW 2008 NNLO PDF sets, with $m_t = 172.9$ GeV. As with the NLO results, the effect on the stationary point of the PDF set is generally small. The largest change in this case is for the Tevatron energy, with the cross-section using CT10 PDFs about 5% larger than that using MSTW PDFs.

Comparing the NLO and NNLO results, we see that at 1.96 TeV the stationary point moves to a much higher M and lower μ_{R}. This results in a

	CT10			MSTW		
$\sqrt{S}$ [TeV]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]
1.96	440	26	7.94	420	24	7.59
7	120	64	168	120	64	168
14	130	64	933	130	64	936

Table 5.3: Comparison of stationary points found using CT10 and MSTW 2008 NNLO PDFs with Top++ at NNLO and $m_t = 172.9$ GeV.

$\sim 10\%$ increase on the cross-section prediction at NLO. At the LHC at 7 TeV the saddle point moves to higher M and μ_{R}, producing a 10% decrease in PMS cross-section compared with NLO. A similar trend is seen at 14 TeV, with a slightly larger decrease of $\sim 15\%$.

In addition to studying the effect of PDF set on the PMS results the variation with respect to top quark mass was also considered. The analysis was repeated for MSTW PDFs using $m_t = 172.5$ GeV and $m_t = 174.5$ GeV. These values were chosen as they are the upper and lower limits of the PDG value of the top mass as obtained from direct measurements: $m_t = 173.5 \pm 1.0$ GeV [85]. The results are shown in Table 5.4. The stationary points are generally found to be in the same place for both values of the top mass, with the exception of NNLO PMS at the Tevatron and NLO PMS at 7 TeV. The cross-sections at the PMS point for $m_t = 172.5$ GeV are 1–2% larger than 172.9 GeV, and 5–7% larger than with $m_t = 174.5$ GeV. Even for NNLO PMS at 1.96 TeV, where $\overline{M}$ is almost 100 GeV greater at the lower top mass than for 172.9 GeV, the cross-section at the saddle point is still only $\sim 7\%$ larger. This implies that the saddle point covers a moderately large range of the factorisation scale. This is confirmed by calculating $\sigma_{\mathrm{NNLO}}(M = 510, \mu_{\mathrm{R}} = 26)$ with $m_t = 172.9$ GeV at 1.96 TeV, which is found to be the same as $\sigma_{\mathrm{NNLO}}(M = 420, \mu_{\mathrm{R}} = 24)$ to the precision quoted. Similarly, $\sigma_{\mathrm{NNLO}}(M = 510, \mu_{\mathrm{R}} = 26)$ with $m_t = 174.5$ GeV

	$m_t = 174.5$ GeV			$m_t = 172.5$ GeV		
$\sqrt{S}$ [TeV]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]	$\overline{M}$ [GeV]	$\overline{\mu}_R$ [GeV]	$\sigma(\overline{M}, \overline{\mu}_R)$ [pb]
			NLO			
1.96	110	84	6.77	110	82	7.21
7	85	37	180	64	42	191
14	85	28	1040	86	28	1100
			NNLO			
1.96	460	26	7.21	510	26	7.69
7	120	65	160	120	64	170
14	130	64	897	130	64	946

Table 5.4: Comparison of stationary points found using $m_t = 174.5$ GeV and $m_t = 172.5$ GeV with Top++ at NLO and NNLO with MSTW PDFs.

agrees with that quoted for $M = 460$, $\mu_{\rm R} = 26$. The same conclusions can be drawn for the NLO case at the LHC energy 7 TeV. Changing the top mass is therefore found to affect the cross-section only, and not the position of the saddle point. Altering m_t by 2 GeV alters the cross-sections by less than 7 %, comparable with the PDF variation between the MSTW and CT10 sets.

The NNLO PMS results can be compared with the cross-sections obtained using the standard diagonal scale choice of $M = \mu_{\rm R} = m_t = 172.9$ GeV. For the Tevatron at 1.96 TeV this gives $\sigma(M = \mu_{\rm R} = m_t) = 7.09 \pm 0.31$ pb. For the LHC at 7 and 14 TeV, the standard cross-sections are 163 ± 4 pb and 902 ± 31 pb respectively. The errors are obtained from varying the scale between 0.5 and 2 m_t. These calculations use MSTW NNLO PDFs. The PMS cross-sections vary from these standard cross-sections by $\sim 7\%$ in the case of the Tevatron, which is slightly larger than the scale variation on the standard cross-section. The difference decreases to only 3–4% at the LHC. This is less than the variation between using CT10 and MSTW PDFs. The slight differences between the PMS method and the standard scale choices is

	CT10	MSTW	CT10	MSTW
	σ_{diag} [pb]		$(\sigma_{\text{PMS}} - \sigma_{\text{diag}})/\sigma_{\text{diag}} \cdot 100$ [%]	
$\sqrt{S}$ [TeV]	NLO			
1.96	6.83 ± 0.75	6.77 ± 0.75	5.1	5.2
7	150 ± 20	160 ± 21	18	18
14	842 ± 98	894 ± 108	21	22
	NNLO			
1.96	7.42 ± 0.33	7.09 ± 0.31	7.0	7.1
7	163 ± 4	163 ± 4	3.1	3.1
14	898 ± 31	902 ± 31	3.9	3.8

Table 5.5: Comparison of cross-sections obtained using the diagonal scale choices of $M = \mu_{\text{R}} = m_t$ found using $m_t = 172.9$ GeV with Top++ at NLO and NNLO with CT10 and MSTW PDFs. At NLO, MSTW NLO PDFs were used and at NNLO, MSTW NNLO PDFs were used. In the two right-hand columns we show the percentage difference between the PMS and diagonal scale choice cross-sections.

therefore seen to lessen at NNLO.

In order to allow the reader to fully compare the diagonal scale choices against PMS at NLO and NNLO with both PDF sets we have summarised the cross-sections obtained at $M = \mu_{\text{R}} = 172.9$ GeV in Table 5.5. These results were all calculated using Top++. On the right-hand side of Table 5.5 are the percentage differences between the cross-sections obtained using the PMS and physical scale methods. The two PDF sets give consistently larger cross-sections in the PMS method. The difference between the two methods at the LHC is considerably smaller at NNLO than at NLO.

Finally, as a visual summary, in Figs. 5.7–5.9 we have plotted the cross-sections given by the PMS method and the standard $\overline{\text{MS}}$ method against experimental data where it is available. Cross-section predictions at NLO and NNLO are both shown. For each of the four theoretical cases, three points are

plotted: one for the MSTW PDFs, one for the CT10 PDFs, and one an average of the two. The experimental data point is an average of the two values at each energy — 7.53 ± 0.63 pb for the Tevatron and 172 ± 14 pb for the LHC — with the larger uncertainty taken in each case. The uncertainties on the PMS cross-sections are formed of a "PDF" uncertainty and a top mass uncertainty combined in quadrature. The PDF uncertainty is simply the difference between the MSTW and CT10 results in each case, and it should be noted that this may well be an underestimation since only two PDF sets were examined. The top mass uncertainty is calculated by subtracting $\sigma(m_t = 174.5\text{GeV})$ and $\sigma(m_t = 172.5\text{GeV})$ from $\sigma(m_t = 172.9\text{GeV})$, with the larger difference taken as the uncertainty.[2] The error bars on the diagonal scale choice cross-sections include the PDF and top mass uncertainties, along with the scale variation uncertainty.

5.4 Commentary on PMS results

In the previous few sections we applied the PMS method in two different ways. In Section 4.2, PMS was applied analytically at NLO and NNLO to e^+e^- event shape moments. The resulting predictions for the moments at M_Z agree very closely with those given by the ECH method, as calculated in Section 3.2. At NLO both these predictions are close to the experimental data for the means at M_Z. For the higher moments the predictions are not as close, a trend that was noted in Chapter 2. At NNLO the PMS and ECH methods again agree very closely, and the predictions produced are not as good as those at NLO.

The similarity between PMS and the Fastest Apparent Convergence (FAC)

[2]In practice this means taking $\sigma(m_t = 174.5\text{GeV}) - \sigma(m_t = 172.9\text{GeV})$ as the uncertainty. Since the mass was varied only in conjunction with MSTW PDFs for the PMS method (see Table 5.4), the mass uncertainty calculated from these results was applied to both the MSTW and CT10 predictions.

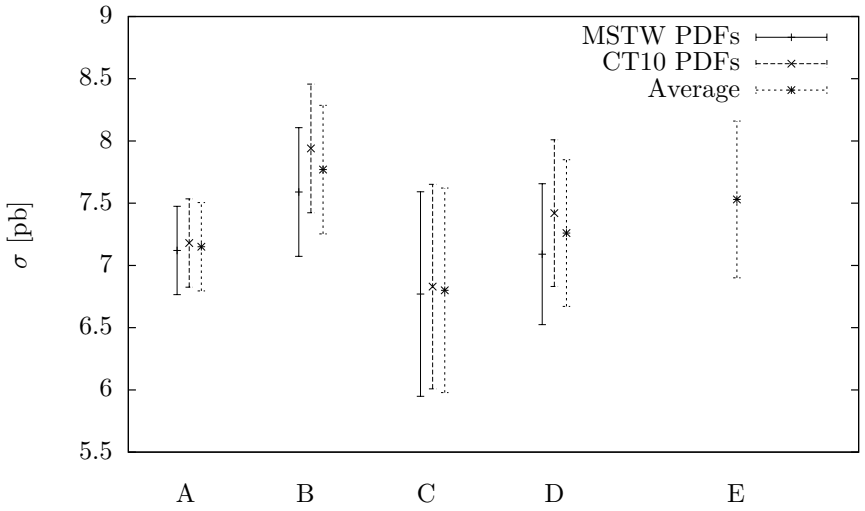


Figure 5.7: A comparison of the cross-sections predicted by the PMS method and the diagonal scale choice for the Tevatron at 1.96 TeV. The cluster of points at A show the PMS at NLO results with MSTW NLO and CT10 PDFs and the average of the two. B contains the same for PMS at NNLO, using MSTW NNLO PDFs. C and D are standard $\overline{\text{MS}}$ results at NLO and NNLO respectively. An average of the experimental data is shown at E.

method, which is equivalent to ECH at NLO, has been noted in Refs. [57, 58]. The PMS point was also observed to be close to the predictions of ECH by Chýla in Ref. [66]. Note that at NLO, where $c_2 = c_3 = \ldots = 0$, and with c set equal to zero in addition the ECH and PMS methods are exactly the same. This can be seen by setting $c = 0$ in (4.1.14) and (2.1.12), and remembering that $\mathcal{R}_{\text{NLO}} = \alpha_s(\mu_{\text{ECH}})/\pi$ as discussed in Section 3.2.

This good agreement between the PMS and ECH methods was motivation for applying PMS to hadronic observables. Because these quantities involve two scales — the renormalisation scale, μ_{R}, and the factorisation scale, M — it is not possible to use the ECH method. There are logs involving both scales and it is not (currently) known how to separate out the μ_{R} and M dependency in order to achieve an expansion such as in (2.1.5).

The first hadronic observable studied was the total $b\bar{b}$ cross-section at the Tevatron. PMS was applied numerically at NLO to find the values of M and

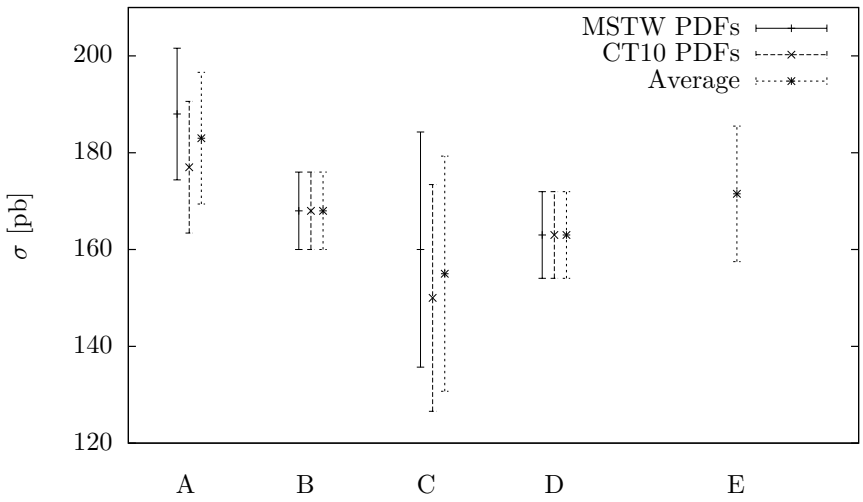


Figure 5.8: A comparison of the cross-sections predicted by the PMS method and the diagonal scale choice for the LHC at 7 TeV. The labels A to E are defined in Fig. 5.7.

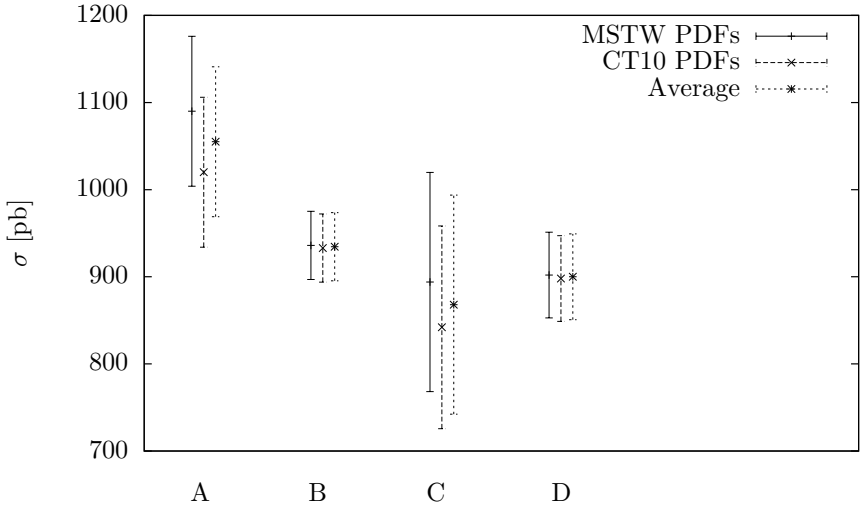


Figure 5.9: A comparison of the cross-sections predicted by the PMS method and the diagonal scale choice for the LHC at 14 TeV. The labels A to D are defined in Fig. 5.7. There are no experimental data for this energy.

μ_{R} at the stationary point on the cross-section surface. The M and μ_{R} values that mark the saddle points are seen to move in a highly non-diagonal way in Fig. 5.2. The PMS predictions for the cross-section are generally much larger than those given by the physical scale choice. For example, at the Tevatron

energy of 1.96 TeV the PMS cross-section is almost a factor of two larger than when using the standard diagonal scale choice. In Fig. 5.4 it can be seen that the PMS prediction generally lies outside of the scale variation on the standard $\overline{\text{MS}}$ result. At $\sim$ 400 GeV the PMS cross-section is over 3 times larger than the central $\overline{\text{MS}}$ prediction.

We also examined the $t\bar{t}$ total cross-section at both the LHC and the Tevatron. MCFM was used to study the NLO cross-section. At the Tevatron at 1.96 TeV the PMS cross-section and the physical scale cross-section agree to within the scale variation on the latter value. Both predictions agree with the experimental data. At the LHC energies σ_{PMS} is a little larger than the upper scale variation given by $M = \mu_{\text{R}} = m_t$. The predictions at 7 TeV again both agree with the experimental data to within the errors. At the energies where experimental data is available σ_{PMS} is slightly closer to the central values of the data than the diagonal cross-section when using CT10 PDFs.

These results were replicated using cross-sections calculated with Top++ in Section 5.3.2. The effect of using another PDF set, MSTW 2008 at NLO and NNLO, was examined. Similar results were found as for the CT10 PDFs, with the cross-sections generally being around 7% larger for both methods at LHC energies. This results in the PMS cross-section no longer being closer to the central experimental value than the standard $\overline{\text{MS}}$ result.

Similar results were found when using different methods of calculating the cross-section (MCFM and Top++), and when using different algorithms (Secant Levenberg-Marquadt and a Nelder-Mead algorithm) at a variety of energies at the Tevatron and the LHC, indicating robustness of the PMS method. The variation when changing between the CT10 and MSTW PDF sets was found to be small, although since the predictions from the PMS and diagonal scale choice methods were similar the PDF uncertainty is large enough to make

it difficult to distinguish between the two sets of results. Similar effects were seen when varying the top mass, although interestingly only the cross-section seems to vary with m_t, in general, and not the position of the saddle point. The two methods are therefore found to give equally good predictions given the uncertainties involved.

Top++ was also used to give preliminary NNLO cross-sections. As in the NLO case, the PMS and diagonal scale predictions are very similar, and close to the experimental data. Both methods are therefore found to perform well for this observable. One interesting point, however, is that the percentage difference between the two methods at LHC energies is significantly lower at NNLO than at NLO. The opposite trend is seen at 1.96 TeV but the differences in this case are smaller. One explanation for the difference between the two methods decreasing at NNLO could be that Top++ only includes an approximation to the gg channel at NNLO, which is the dominant channel at the LHC. However, it could also just be a reflection of the decreased scale uncertainty moving from NLO to NNLO. The agreement between the PMS and standard $\overline{\text{MS}}$ cross-sections, despite σ_{PMS} being calculated at non-diagonal scales, suggests that the surface is fairly flat and therefore that the scale uncertainty is small.

One possible extension to this work is to apply the PMS method to the top forward-backward asymmetry at the Tevatron, where there is currently a 2–3σ discrepancy between QCD predictions and experimental data [86–88]. Preliminary work carried out by the author suggests that there is no saddle-point in the asymmetry, as calculated from NLO QCD cross-sections, however. This may be because the effect of the top forward-backward asymmetry only appears at NLO in the cross-section, so it is in fact a LO result for the asymmetry.

To see this we define the top asymmetry as:

$$A_{FB} = \frac{\sigma^{\mathrm{NLO}}(y_t^{t\bar{t}} > 0) - \sigma^{\mathrm{NLO}}(y_t^{t\bar{t}} < 0)}{\sigma^{\mathrm{NLO}}(y_t^{t\bar{t}} > 0) + \sigma^{\mathrm{NLO}}(y_t^{t\bar{t}} < 0)}\,, \tag{5.4.1}$$

where $y_t^{t\bar{t}}$ is the rapidity of the top-quark in the $t\bar{t}$ rest frame. Expanding this in terms of α_s gives [87]:

$$A_{FB} = \frac{\alpha_s^3 N_1 + \alpha_s^4 N_2 + \mathcal{O}(\alpha_s^5)}{\alpha_s^2 D_0 + \alpha_s^3 D_1 + \alpha_s^4 D_2 + \mathcal{O}(\alpha_s^5)} \tag{5.4.2}$$

$$= \frac{\alpha_s}{D_0}\left(N_1 + \left(N_2 - \frac{N_1 D_1}{D_0}\right)\alpha_s + \ldots\right)\,. \tag{5.4.3}$$

The coefficients D_0, N_1 and D_1 are known at NLO, but N_2 is not. We do not fully know the coefficient at $\mathcal{O}(\alpha_s^2)$ in A_{FB}, and therefore do not have a full NLO calculation for the asymmetry. As can be seen in the left hand side of Fig. 5.1, LO surfaces for the heavy quark cross-section appear to not have a stationary point. The incomplete NLO result for A_{FB} might also not have a saddle point, or have a saddle point in a misleading position. In this case, the NNLO cross-sections will be necessary to apply PMS to the asymmetry surface. This cannot be implemented using Top++ as it gives only total cross-sections and not the differential cross-sections that are necessary to calculate the asymmetry. If NNLO results for the differential cross-section in bins of rapidity become available, and there is still a discrepancy between theory and experiment, it would be interesting to apply PMS to the asymmetry surface.

In theory the PMS method can be applied numerically to any observable with a scale dependence. Practically, however, it would be challenging to apply the method to more than two scales. In addition, as finding stationary points is a reasonably time consuming procedure, it would be sensible to only consider PMS in cases where there is a large discrepancy between experiment and

perturbative QCD predictions, or where scale variation from missing higher orders appears to be a problem.

Chapter 6

Conclusions

In this thesis we have examined two methods of combating the unphysical renormalisation scheme dependence seen in truncated perturbative expansions of physical observables: the Effective Charges (ECH) method and the Principle of Minimal Sensitivity (PMS). The ECH technique is to integrate up a β-function type equation that has an expansion in terms of the observable itself, resulting in scheme independent predictions. The PMS approach uses scales at which the observable is locally invariant to the renormalisation scheme.

In Chapter 3 the ECH method was used to extract values of $\alpha_s(M_Z)$ from e^+e^- event shape moment data, and also to provide predictions of the data. At NLO and applied to the event shape means, ECH was found to work very well. It extracts values of $\alpha_s(M_Z)$ that are in good agreement between the different event shapes, as also observed in Ref. [35]. The predictions of ECH at NLO for data of event shape means also show good agreement with experimental data over a wide range of energies.

We had anticipated that applying the ECH method at NLO to higher moments of event shapes would result in similarly good agreement with data, and that extending the analysis to NNLO would improve on the NLO results.

However, for some higher moments it does not provide such good predictions and the previously good agreement with data for the means deteriorates when the NNLO corrections are added to the ECH method. An explanation of these observations was given in Section 3.3.1. By examining the $\Lambda_{\mathcal{R}}$ parameter for the observables it is seen that the data on higher moments are generally at energies below the perturbative regime. There is a Landau pole at $Q = \Lambda_{\mathcal{R}}$ at which the integrated β-function equation diverges and so it is not appropriate to apply the ECH method to these moments at the energies studied. The terms in the asymptotic series $\rho(\mathcal{R})$ suggest that for the higher moments the optimum expansion would involve truncating the series at LO, again indicating that the ECH method will not be effective for these variables. The energy dependence of $\mathcal{R}(Q)$ is controlled by the RS invariant dimensionful constant $\Lambda_{\mathcal{R}}$ of (2.1.18) which involves the NLO perturbative coefficient $r_1^{\overline{\rm MS}}(\mu_{\rm R} = Q)$. This implies that leading order truncation is not physically meaningful. The asymptotic series also indicate that NLO may be the optimum order of truncation for the means, although without calculating the next term in the expansion this cannot be definitively stated. However, the good agreement of the NLO ECH predictions with the experimental data for the means supports this interpretation.

In conclusion, the ECH method is very successful for certain physical observables. There is an optimum order of expansion at which the method works well, and this is seen to be surprisingly small for e^+e^- event shape moment data. Although the RS independent predictions of ECH can be very effective, care needs to be taken that the observable in question is within the perturbative regime for the ECH method.

Using a Padé Sum (PS) of the $\rho(\mathcal{R})$ expansion in the Effective Charges method is found to improve the predictions of the higher moments that were

previously problematic, although it does not perform as well as pure NLO ECH for the means. The $\overline{\rm MS}$ prediction at NNLO is improved by the use of a PS, and some of the predictions for the $n = 2, 3$ moments are particularly good. $\overline{\rm MS}$ with a PS generally outperforms the equivalent ECH combination for the higher moments. The scale uncertainty for $\overline{\rm MS}$ PT with the PS is very small, as shown in Fig. 3.16, and considerably less than that for $\overline{\rm MS}$ at NNLO. Approximating the missing higher orders in $\rho(\mathcal{R})$ of the $\overline{\rm MS}$ perturbative expansion with a PS is therefore seen to be highly effective, especially for the $n = 2$, 3 moments of $\langle(1-T)^n\rangle$, $\langle C^n\rangle$ and $\langle B_T^n\rangle$.

The success of adding a PS to standard $\overline{\rm MS}$ perturbation theory indicates that this is a useful method which should be considered for use in other appropriate situations. The potential for the application of Padé Approximant methods to QCD has also been noted in e.g. Refs. [58, 59, 89]. The addition of a PS to $\overline{\rm MS}$ PT could be applied to hadronic cross-sections. The predictions might, as for the event shape moments, be less dependent on $\mu_{\rm R}$ than $\overline{\rm MS}$ PT at NNLO. However, there would be residual M-dependence, so it is not clear that $\overline{\rm MS}$ with a PS would be as effective in these cases.

Possible non-perturbative effects were investigated using the dispersive model of Refs. [6, 48, 49], which has been widely used in other analyses of event shapes, and a simple power correction model used in Ref [35]. When performing fits for NLO ECH plus dispersive power corrections, we find consistent values of $\alpha_s(M_Z)$ and α_0, except for the observables $\langle 1-T\rangle$ and $\langle C\rangle$. This is an improvement on the fit results gained from NLO $\overline{\rm MS}$ PT plus power corrections. However ECH at NNLO again performs less well than the $\overline{\rm MS}$ PT counterpart. ECH at NLO gives good agreement in $\alpha_s(M_Z)$ when simple power corrections were added. Generally only small values of κ_0 are required, as expected since pure ECH at NLO gives good predictions across the energy

range studied. We conclude that adding power corrections does not improve the performance of the ECH method at NNLO, and that non-perturbative power corrections are not required by ECH to describe the means at NLO.

In Chapter 4 it is seen that the PMS method gives very similar predictions for event shape moment data to those obtained using ECH. This applies at both NLO and NNLO. Both methods are motivated by the aim of avoiding RS dependence in predictions of physical observables, but the derivations are different so it is interesting that they produce such similar results. The similarity between PMS and ECH has been noted before in e.g. Refs. [57, 58, 66].

When applied to $b\bar{b}$ production, PMS is found to predict substantially larger total cross-sections than using the standard diagonal scale choices of $M = \mu_{\rm R} = m_b$. At the Tevatron at 1.96 TeV the PMS prediction is larger by a factor of two: $129\,\mu$b compared with $72\,\mu$b. For $t\bar{t}$ production there is very little difference observed between the two methods. Both produce predictions that are in good agreement with the experimental data to within the errors. However, even for $t\bar{t}$ production the PMS scales are found to be far from the diagonal scale choice. The standard way of setting the scales could, therefore, be giving a larger theoretical uncertainty than is necessary. When making fixed order QCD predictions the scales are frequently clamped together and then varied by a factor of two, to provide an estimate of the RS dependence of the missing higher orders. However, simply using a different scale choice could dramatically reduce this error and change the prediction obtained.

We conclude that, although the ECH method is not suitable for all variables and the PMS method predicts only a small difference for the $t\bar{t}$ total cross-section, it is still interesting to look at methods which improve fixed order perturbation theory and decrease RS dependence without the need for non-perturbative corrections.

Appendix A

More details of the ECH method

In this appendix we give more details of the derivation of (2.1.5) in Section 2.1. We start by differentiating the definition of an effective charge:

$$\mathcal{R}(Q) = a(1 + r_1 a + r_2 a^2 + \dots)\,, \tag{A.0.1}$$

with respect to $\ln Q$, while setting $\mu_{\rm R} = Q$, which gives:

$$\begin{aligned}\frac{d\mathcal{R}(Q)}{d\ln Q} = \rho(\mathcal{R}(Q)) &= \frac{\partial a}{\partial \ln Q}(1 + 2r_1 a + 3r_2 a^2 + \dots) \\ &= \beta(a)(1 + 2r_1 a + 3r_2 a^2 + \dots)\,,\end{aligned} \tag{A.0.2}$$

$\beta(a)$ is given in (1.2.18) and can be substituted in to give:

$$\begin{aligned}\rho(\mathcal{R}(Q)) &= -ba^2(1 + ca + c_2 a^2 + \dots)(1 + 2r_1 a + 3r_2 a^2 + \dots) \\ &= -ba^2(1 + (c + 2r_1)a + (c_2 + 3r_2 + 2cr_1)a^2 + \dots)\,.\end{aligned} \tag{A.0.3}$$

The next step is to invert (A.0.1) to get a perturbative expansion for a in

terms of $\mathcal{R}(Q)$. We assume that the expansion is of the form:

$$a(\mathcal{R}) = \mathcal{R}(1 + A_1\mathcal{R} + A_2\mathcal{R}^2 + \ldots)\,, \tag{A.0.4}$$

and then substitute in for $\mathcal{R}$ and compare terms in orders of a to find the coefficients A_n:

$$\begin{aligned} a(\mathcal{R}) &= a(1 + r_1 a + r_2 a^2 + \ldots)(1 + A_1 a(1 + r_1 a + \ldots) + A_2 a^2(1 + \ldots) + \ldots) \\ &= a(1 + r_1 a + r_2 a^2 + \ldots)(1 + A_1 a + (A_1 r_1 + A_2)a^2 + \ldots) \\ &= a(1 + (A_1 + r_1)a + (r_2 + 2A_1 r_1 + A_2)a^2 + \ldots)\,. \end{aligned} \tag{A.0.5}$$

Comparing coefficients gives:

$$\begin{aligned} A_1 &= -r_1 \\ A_2 &= -r_2 - 2A_1 r_1 \\ &= 2r_1^2 - r_2\,, \end{aligned} \tag{A.0.6}$$

and so the expansion of a in terms of $\mathcal{R}(Q)$ is:

$$a(\mathcal{R}) = \mathcal{R}(1 - r_1\mathcal{R} + (2r_1^2 - r_2)\mathcal{R}^2 + \ldots)\,. \tag{A.0.7}$$

Substituting (A.0.7) into (A.0.3), and keeping the coefficients in (A.0.7) as

A_n until the last line for clarity, gives:

$$\begin{aligned}
\rho(\mathcal{R}(Q)) &= -b\mathcal{R}^2[1+2A_1\mathcal{R}+(2A_2+A_1^2)\mathcal{R}^2+\ldots] \\
&\quad\times[1+(c+2r_1)\mathcal{R}(1+A_1\mathcal{R}+\ldots)+(c_2+3r_2+2cr_1)\mathcal{R}^2(1+\ldots)+\ldots] \\
&= -b\mathcal{R}^2[1+2A_1\mathcal{R}+(2A_2+A_1^2)\mathcal{R}^2+\ldots] \\
&\quad\times[1+(c+2r_1)\mathcal{R}+(cA_1+2A_1r_1+c_2+3r_2+2cr_1)\mathcal{R}^2+\ldots] \\
&= -b\mathcal{R}^2[1+(2A_1+c+2r_1)\mathcal{R}+(2A_2+A_1^2+cA_1+2A_1r_1+c_2+3r_2 \\
&\quad+2cr_1+2cA_1+4A_1r_1)\mathcal{R}^2+\ldots] \\
&= -b\mathcal{R}^2[1+c\mathcal{R}+(r_2+c_2-r_1c-r_1^2)\mathcal{R}^2+\ldots]\,. \qquad \text{(A.0.8)}
\end{aligned}$$

Comparing this with (2.1.5) we see that:

$$\rho_2 = r_2+c_2-r_1c-r_1^2\,. \qquad \text{(A.0.9)}$$

If higher orders of $\mathcal{R}$ are included the form of the higher coefficients in $\rho(\mathcal{R}(Q))$ can be found. For example:

$$\rho_3 = -6r_1r_2+4r_1^3+2r_3+c_3-2r_1c_2+r_1^2c\,. \qquad \text{(A.0.10)}$$

The ρ_n coefficients are manifestly RS invariant since both $d\mathcal{R}/d\ln Q$ and $\mathcal{R}(Q)$ are physical observables. We can explicitly show the RS-independence of ρ_2 by differentiating with respect to τ and c_2, the two parameters that define the RS at NNLO. Differentiating with respect to τ gives:

$$\frac{\partial\rho_2}{\partial\tau} = \frac{\partial r_2}{\partial\tau}-c\frac{\partial r_1}{\partial\tau}-2r_1\frac{\partial r_1}{\partial\tau}\,. \qquad \text{(A.0.11)}$$

In Section 4.1 we derive expressions for $\frac{\partial r_2}{\partial \mathrm{RS}}$ and $\frac{\partial r_1}{\partial \mathrm{RS}}$ using the self-consistency

of perturbation theory. We substitute them in here to give:

$$\frac{\partial \rho_2}{\partial \tau} = (c + 2r_1) - c - 2r_1 = 0\,. \tag{A.0.12}$$

Similarly, differentiating with respect to c_2 and using the self-consistency conditions gives:

$$\frac{\partial \rho_2}{\partial c_2} = \frac{\partial r_2}{\partial c_2} + 1 - c\frac{\partial r_1}{\partial c_2} - 2r_1\frac{\partial r_1}{\partial c_2} \tag{A.0.13}$$

$$= -1 + 1 = 0\,, \tag{A.0.14}$$

and we see that ρ_2 is indeed RS-invariant.

Appendix B

Correction factors for $n = 2, 3$ event shape moments

The data of event shape moments used in this thesis are corrected for b and c quark decays. The correction factors were calculated using HERWIG++ simulations [46, 47]. Samples of 10^6 events were run with $N_f = 3$ (only light quarks) and also with $N_f = 5$ (with b and c quarks). The event shape moment was calculated in both of these cases and a ratio of the two quantities was taken to give a correction factor. This was done for each energy. The data were then multiplied by the appropriate correction factors. The correction factors for the means are shown in the main body of the text in Fig. 2.1. The correction factors for the higher moments ($n = 2$ and 3) are shown in Figs. B.1 and B.2.

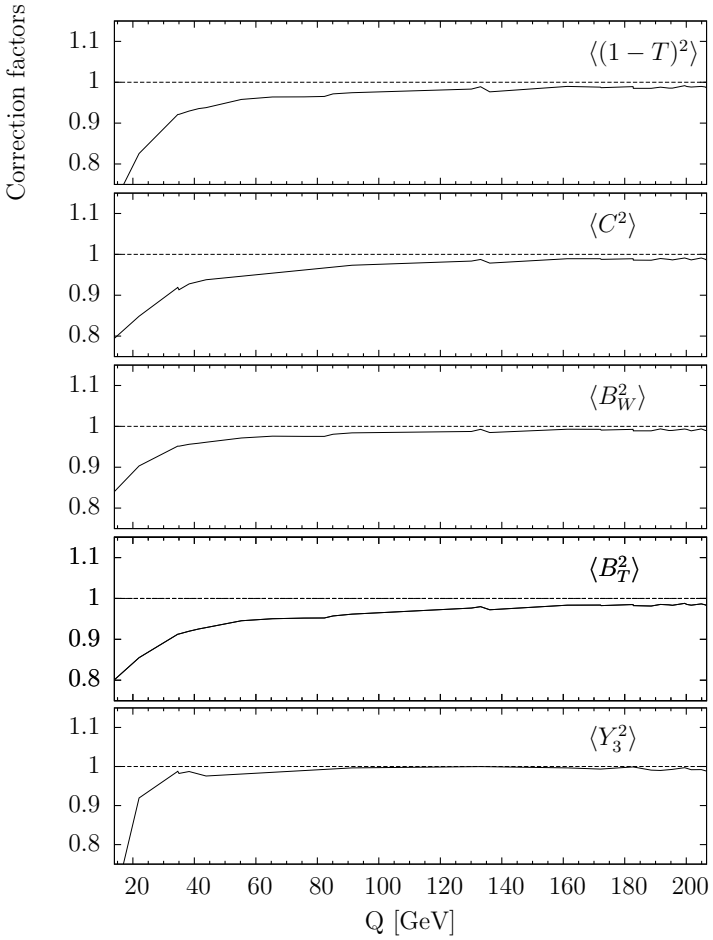


Figure B.1: Heavy quark mass correction factors for the $n = 2$ event shape moments.

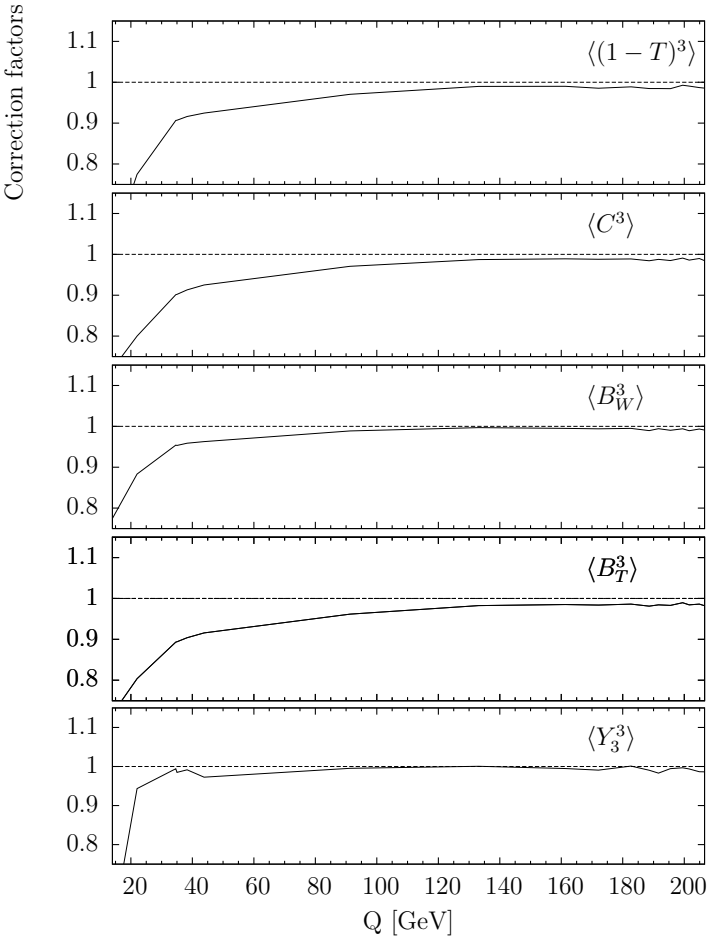


Figure B.2: Heavy quark mass correction factors for the $n = 3$ event shape moments.

Appendix C

Saddle-point finding algorithm

This appendix describes the algorithm that was used in Chapter 4 to find the stationary points of surfaces. We start by describing the Levenberg-Marquadt algorithm in Section C.1, and then explain how this is adapted to a Secant Levenberg-Marquadt approach which uses a BFGS update in Section C.2. More details of the parameters used in the algorithm are given in Section C.3. Further details of the algorithms which were used as the basis of our program can be found in Refs. [72, 73].

C.1 Levenberg-Marquadt algorithm

The Levenberg-Marquadt algorithm uses a damping method to move in a direction that is a mixture of a Gauss-Newton and a steepest descent step. It can be used to find a minimum of a function in a non-linear least squares problem. Consider a function, F, of the form:

$$F(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{m}(f_i(\mathbf{x}))^2 = \frac{1}{2}\mathbf{f}(\mathbf{x})^{\mathbf{T}}\mathbf{f}(\mathbf{x})\,, \tag{C.1.1}$$

where $\mathbf{x}$ is an n-dimensional vector. In the case of finding a saddle point on a 2D surface, $\mathbf{x}$ has length 2 and contains the co-ordinates of a point on the surface. The aim of the algorithm is to find the position of the minimum of F, $\mathbf{x}^*$.

The principal equation to determine the step taken in the Levenberg-Marquadt algorithm is:

$$(\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{J}(\mathbf{x}) + \mu\mathbf{I})\mathbf{h}_{\mathrm{LM}} = -\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{f}(\mathbf{x})\,, \tag{C.1.2}$$

where $\mathbf{J}$ is the Jacobian, $\mu \geq 0$ is the damping parameter which varies throughout the algorithm (see below), and $\mathbf{h}_{\mathrm{LM}}$ is the Levenberg-Marquadt step. The Jacobian is defined as:

$$(\mathbf{J}(\mathbf{x}))_{ij} = \frac{\partial f_i}{\partial x_j}\,. \tag{C.1.3}$$

When the damping parameter is large, and $\mathbf{x}$ is far from the minimum, a short step is taken in the steepest descent direction:

$$\mathbf{h}_{\mathrm{LM}} \simeq -\frac{\mathbf{J}^{\mathbf{T}}\mathbf{f}}{\mu} = -\frac{\nabla F(\mathbf{x})}{\mu}\,. \tag{C.1.4}$$

When the algorithm is close to $\mathbf{x}^*$, μ decreases and the step approaches the Gauss-Newton direction:

$$\mathbf{h}_{\mathrm{LM}} \simeq -\frac{\mathbf{J}^{\mathbf{T}}\mathbf{f}}{\mathbf{J}^{\mathbf{T}}\mathbf{J}}\,. \tag{C.1.5}$$

This is also a descent direction but improves the convergence from linear to nearly quadratic, provided that $F(\mathbf{x}^*)$ is zero or nearly zero.

Initially μ is chosen to be:

$$\mu = \tau \cdot \max_i\big((\mathbf{J}(\mathbf{x_0})^{\mathbf{T}}\mathbf{J}(\mathbf{x_0}))_{ii}\big)\,, \tag{C.1.6}$$

where $\mathbf{x_0}$ is the starting point for the algorithm, and τ is chosen by the user. Generally τ should be chosen to be small ($\sim 10^{-6}$) if $\mathbf{x_0}$ is thought to be close to $\mathbf{x}^*$, and larger (10^{-3} or 1) otherwise [72]. At the beginning of the algorithm a parameter ν, which is used in the updating mechanism for μ, is assigned a starting value of 2.

For each iteration of the algorithm the step is determined by solving (C.1.2). A new $\mathbf{x}$ is then defined as:

$$\mathbf{x}_{\text{new}} = \mathbf{x} + \mathbf{h}_{\text{LM}} \,. \tag{C.1.7}$$

The function F is then evaluated at $\mathbf{x}_{\text{new}}$. If the new value, $F(\mathbf{x}_{\text{new}})$, is smaller than $F(\mathbf{x})$ the step is accepted, i.e. $\mathbf{x}$ is now assigned the value of $\mathbf{x}_{\text{new}}$ and the Jacobian and $\mathbf{f}(\mathbf{x})$ are recalculated.

The value of μ is also updated and in order to do this the following quantity, known as the 'gain ratio', is calculated:

$$\rho = \frac{F(\mathbf{x}) - F(\mathbf{x}_{\text{new}})}{L(\mathbf{0}) - L(\mathbf{h}_{\text{LM}})} \,, \tag{C.1.8}$$

where L is a linear model of F, where $\mathbf{f}(\mathbf{x} + \mathbf{h}_{\text{LM}})$ is Taylor expanded to first order:

$$\begin{aligned} F(\mathbf{x} + \mathbf{h}_{\text{LM}}) &= \frac{1}{2}\mathbf{f}(\mathbf{x} + \mathbf{h}_{\text{LM}})^{\mathbf{T}}\mathbf{f}(\mathbf{x} + \mathbf{h}_{\text{LM}}) \\ &\simeq L(\mathbf{h}_{\text{LM}}) = \frac{1}{2}(\mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\mathbf{h}_{\text{LM}})^{\mathbf{T}}(\mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x})\mathbf{h}_{\text{LM}}) \\ &= F(\mathbf{x}) + \mathbf{h}_{\text{LM}}{}^{\mathbf{T}}\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{f}(\mathbf{x}) + \frac{1}{2}\mathbf{h}_{\text{LM}}{}^{\mathbf{T}}\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{J}(\mathbf{x})\mathbf{h}_{\text{LM}} \,. \end{aligned} \tag{C.1.9}$$

From this we see that $L(\mathbf{0}) = F(\mathbf{x})$. The denominator of (C.1.8) is therefore:

$$\begin{aligned} L(\mathbf{0}) - L(\mathbf{h}_{\mathrm{LM}}) &= -\mathbf{h}_{\mathrm{LM}}{}^{\mathbf{T}}\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{f}(\mathbf{x}) - \frac{1}{2}\mathbf{h}_{\mathrm{LM}}{}^{\mathbf{T}}\mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{J}(\mathbf{x})\mathbf{h}_{\mathrm{LM}} \\ &= \frac{1}{2}\mathbf{h}_{\mathrm{LM}}{}^{\mathbf{T}}(\mu\mathbf{h}_{\mathrm{LM}} - \mathbf{J}(\mathbf{x})^{\mathbf{T}}\mathbf{f}(\mathbf{x}))\,, \end{aligned} \tag{C.1.10}$$

where we have used (C.1.2) in the final step.

The gain ratio compares the difference in F over the step taken to the difference in the linear model of F. If ρ is large $L(\mathbf{h}_{\mathrm{LM}})$ is likely to be a good approximation to $F(\mathbf{x} + \mathbf{h}_{\mathrm{LM}})$. The Gauss-Newton method minimises the linear model, so that:

$$\mathbf{L}'(\mathbf{h}_{\mathrm{GN}}) = \mathbf{J}^{\mathbf{T}}\mathbf{f} + \mathbf{J}^{\mathbf{T}}\mathbf{J}\mathbf{h}_{\mathrm{GN}} = 0\,. \tag{C.1.11}$$

Therefore in the case of ρ being large μ should be decreased to give a next step in the Levenberg-Marquadt algorithm that is close to $\mathbf{h}_{\mathrm{GN}}$. Alternatively, if ρ is small then the linear model is a poor approximation and μ should be increased. The next step $\mathbf{h}_{\mathrm{LM}}$ will then be closer to that of steepest descent and the step length will also be decreased. μ is therefore updated in the following way:

$$\mu = \mu \cdot \max\left(\frac{1}{3}, 1 - (2\rho - 1)^3\right)\,. \tag{C.1.12}$$

If $F(\mathbf{x}_{\mathrm{new}}) > F(\mathbf{x})$, $\mathbf{x}$ is not updated to $\mathbf{x}_{\mathrm{new}}$ and no new Jacobian or $\mathbf{f}(\mathbf{x})$ is calculated. μ is increased to $\mu \cdot \nu$ so that the next step will be closer to that of steepest descent. ν is also increased to 2ν. The next time a step is accepted ν is set back to 2.

The algorithm is terminated by either one of two criteria. The first is if $F(\mathbf{x})$ falls below a certain value, F_{min}, in which case the minimum is deemed to have been found and the current value of $\mathbf{x}$ marks its position. The exact

value that $F(\mathbf{x})$ has to reach is decided upon by the user and will depend on the function to be minimised. The second criterion terminates the algorithm, without having found the minimum of F, if the step length falls below a small value:

$$||\mathbf{h}_{\mathrm{LM}}|| \leq \epsilon(||\mathbf{x}|| + \epsilon)\,, \tag{C.1.13}$$

where ϵ is again chosen by the user.

C.2 Secant Levenberg-Marquadt algorithm with a BFGS update

We now explain how the Levenberg-Marquadt algorithm, outlined in the previous section, is adapted to become a secant method. As we are using MCFM (or Top++) to obtain cross-sections it is not possible to calculate the gradients with respect to M and μ_{R} analytically. Instead, we replace the Jacobian with a matrix $\mathbf{B}$ whose elements contain numerical derivatives. In this thesis we generally use central difference gradients:

$$\frac{\partial f}{\partial x} \simeq \frac{f(x+\delta) - f(x-\delta)}{2\delta}\,, \tag{C.2.1}$$

rather than forward difference gradients. Although this requires more function calls — 4 evaluations of f for gradients in 2 directions rather than 3 for forward difference derivatives — it results in a more accurate estimation of the gradient.

Calculating numerical gradients requires many function calls for each $\mathbf{x}$ so it is desirable to find a way to update $\mathbf{B}$ for a new value of $\mathbf{x}$ instead of explicitly calculating it every time. A generalised secant method satisfies the condition:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_{\mathrm{new}}) + \mathbf{B}_{\mathrm{new}}(\mathbf{x} - \mathbf{x}_{\mathrm{new}})\,. \tag{C.2.2}$$

This set of equations is underdetermined, and so there is some freedom to choose a method of updating $\mathbf{B}$ in each iteration. In this thesis we have chosen to use a BFGS (Broyden, Fletcher, Goldfarb, Shanno) update, which preserves symmetry and positive definiteness in the original matrix $\mathbf{B}_0$. The BFGS update states that $\mathbf{B}$ at $\mathbf{x}_{\text{new}} = \mathbf{x} + \mathbf{h}$ is given by:

$$\mathbf{B}(\mathbf{x}_{\text{new}}) = \mathbf{B}(\mathbf{x}) + \frac{\mathbf{y}\mathbf{y}^{\mathbf{T}}}{\mathbf{y}^{\mathbf{T}}\mathbf{h}} - \frac{(\mathbf{B}(\mathbf{x})\mathbf{h})(\mathbf{B}(\mathbf{x})\mathbf{h})^{\mathbf{T}}}{\mathbf{h}^{\mathbf{T}}\mathbf{B}(\mathbf{x})\mathbf{h}}\,, \tag{C.2.3}$$

where $\mathbf{y} = \mathbf{f}(\mathbf{x}_{\text{new}}) - \mathbf{f}(\mathbf{x})$.

The BFGS update is usually used to update a Hessian matrix in quasi-Newton methods.[1] However, when looking for saddle points the function we minimise is:

$$F(\mathbf{x}) = \frac{1}{2}\left(\left(\frac{\partial\sigma}{\partial \ln M}\right)^2 + \left(\frac{\partial\sigma}{\partial \ln \mu_{\text{R}}}\right)^2\right)\,, \tag{C.2.4}$$

where $\mathbf{x} = (\ln M, \ln \mu_{\text{R}})^{\mathbf{T}}$. In the notation of (C.1.1):

$$\mathbf{f}(\mathbf{x}) = \left(\frac{\partial\sigma}{\partial \ln M}, \frac{\partial\sigma}{\partial \ln \mu_{\text{R}}}\right)^{\mathbf{T}}. \tag{C.2.5}$$

Therefore the Jacobian in (C.1.3) is:

$$\mathbf{J}(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial^2\sigma}{\partial(\ln M)^2} & \dfrac{\partial^2\sigma}{\partial \ln \mu_{\text{R}} \partial \ln M} \\ \dfrac{\partial^2\sigma}{\partial \ln M \partial \ln \mu_{\text{R}}} & \dfrac{\partial^2\sigma}{\partial(\ln \mu_{\text{R}})^2} \end{pmatrix}\,, \tag{C.2.6}$$

which has the form of a Hessian. It is symmetric due to the commutativity of partial derivatives and the BFGS update will preserve this.

In summary, the BFGS update reduces the number of function calls we

[1] Quasi-Newton algorithms look for solutions of $\mathbf{H}(\mathbf{x})\mathbf{h} = -\mathbf{F}'(\mathbf{x})$, with $\mathbf{H}$ replaced with a numerical approximation that is updated by a secant method.

have to make per iteration. $\mathbf{f}(\mathbf{x})$ has to be calculated at each step, but the approximation to $\mathbf{J}(\mathbf{x})$, given by $\mathbf{B}$, only has to be found once at the start of the algorithm. After this $\mathbf{B}$ is updated using (C.2.3).

C.3 Parameters used in the algorithm

In this section we give details of the parameters we used in our algorithm. When calculating the central difference gradients, as defined in (C.2.1), a δ of 0.05 was generally used. Occasionally different values of δ were used but 0.05 was found to work most consistently.

For the $b\bar{b}$ total cross-section the algorithm was run at a large range of energies. It was therefore found convenient to consider derivatives with respect to $\ln M$ and $\ln \mu_{\mathrm{R}}$ as defined in (C.2.4), and in addition to normalise the gradients to a cross-section, i.e.:

$$F(\mathbf{x}) = \frac{1}{2}\left(\left(\frac{1}{\sigma(M,\mu_{\mathrm{R}})}\frac{\partial\sigma}{\partial \ln M}\right)^2 + \left(\frac{1}{\sigma(M,\mu_{\mathrm{R}})}\frac{\partial\sigma}{\partial \ln \mu_{\mathrm{R}}}\right)^2\right) . \qquad \text{(C.3.1)}$$

This enables a value of F_{min}, below which a stationary point is deemed to have been found, to be chosen that can be consistently applied across a range of energies where the cross-section varies by orders of magnitude. For the $b\bar{b}$ analysis a value of $F_{\mathrm{min}} = 10^{-5}$ was chosen. For a few energies a smaller value of F_{min} was also applied to check that the stationary points found in both cases agreed.

Because the cross-sections calculated by MCFM have a statistical uncertainty associated with them the surface generated is uneven. The algorithm was therefore found to be fairly sensitive to the starting values of M and μ_{R}. At the lowest energy for the $b\bar{b}$ total cross-section starting values of $M = \mu_{\mathrm{R}} = 2$ GeV and ($M = 3$, $\mu_{\mathrm{R}} = 2$) GeV were both found to give a stationary point.

For the next energy up, $\overline{M}$ and $\overline{\mu}_R$ from the previous energy were generally used as the starting values. At higher energies, where the differences between the energies considered increased, it was sometimes necessary to choose different starting points before a stationary point was located. It was also noted that if a stationary point was not located for a particular starting choice it could sometimes be found by changing the value of τ used.

For the $t\bar{t}$ total cross-section the algorithm was found to be much less sensitive to the starting values chosen. For example, at 1.96 TeV starting values of ($M = 150$, $\mu_{\mathrm{R}} = 50$) GeV, ($M = m_t$, $\mu_{\mathrm{R}} = 60$) GeV, ($M = 1100$, $\mu_{\mathrm{R}} = 30$) GeV and $M = \mu_{\mathrm{R}} = m_t$ were all found to give a stationary point at $\sim$ (110, 84) GeV. Since fewer energies were examined for this observable, F_{min} was chosen to be 10^{-6}. When using Top++ there is no statistical uncertainty and so the stationary points were generally found, using a built-in Nelder Mead algorithm in Python, to have an F even smaller than this. As for the $b\bar{b}$ cross-section the gradients with respect to $\ln M$ and $\ln \mu_{\mathrm{R}}$ were normalised to a cross-section.

Bibliography

[1] C. J. Maxwell and K. E. Morgan, Nucl. Phys. B **858**, 405 (2012), arXiv:1108.6204v2 [hep-ph].

[2] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, 2003), ISBN: 0-521-54589-7.

[3] Particle Data Group, J. Beringer *et al.*, Phys. Rev. D **86**, 010001 (2012), Review: "Quantum Chromodynamics".

[4] L. H. Ryder, *Quantum Field Theory (2nd ed.)* (Cambridge University Press, 1996).

[5] M. H. Seymour, Quantum chromodynamics, in *European School of High-Energy Physics*, 2004, arXiv:hep-ph/0505192v2. eConf C040530

[6] T. Gehrmann, M. Jaquier, and G. Luisoni, Eur. Phys. J. C **67**, 57 (2010), arXiv:0911.2422v3 [hep-ph].

[7] A. Dhar and V. Gupta, Phys. Rev. D **29**, 2822 (1984).

[8] G. Grunberg, Phys. Rev. D **29**, 2315 (1984).

[9] S. J. Burby and C. J. Maxwell, Nucl. Phys. B **609**, 193 (2001), arXiv:hep-ph/0011203v1.

[10] G. Grunberg, Phys. Lett. B **95**, 70 (1980).

[11] C. J. Maxwell, Phys. Lett. B **409**, 450 (1997), arXiv:hep-ph/9706365v1.

[12] M. J. Dinsdale and C. J. Maxwell, arXiv:hep-ph/0408114v1.

[13] A. Dhar and V. Gupta, Pramana **21**, 207 (1983).

[14] V. Gupta, D. V. Shirkov, and O. V. Tarasov, Int. J. Mod. Phys. A **6**, 3381 (1991).

[15] C. J. Maxwell, Effective Charges, Event Shapes and Power Corrections, in *FRIF Workshop on First Principles Non-Perturbative QCD of Hadron Jets*, 2006, arXiv:hep-ph/0607039v1. eConf C0601121

[16] P. M. Stevenson, Phys. Rev. D **23**, 2916 (1981).

[17] OPAL collaboration, G. Abbiendi *et al.*, Eur. Phys. J. C **40**, 287 (2005), arXiv:hep-ex/0503051v1.

[18] JADE Collaboration, P. A. Movilla Fernandez *et al.*, Eur. Phys. J. C **1**, 461 (1998), arXiv:hep-ex/9708034v2.

[19] C. J. Pahl, PhD thesis, TU Munich, 2007.

[20] DELPHI Collaboration, P. Abreu *et al.*, Phys. Lett. B **456**, 322 (1999).

[21] R. Reinhardt, PhD thesis, BU Wuppertal, 2001.

[22] L3 collaboration, P. Achard *et al.*, Phys. Rept. **399**, 71 (2004), arXiv:hep-ex/0406049v1.

[23] F. Halzen and A. D. Martin, *Quarks and leptons: an introductory course in modern particle physics* (Wiley, 1984).

[24] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (ABP, 1995).

[25] G. 't Hooft and M. Veltman, Nucl. Phys. B **44**, 189 (1972).

[26] M. Czakon, Nucl. Phys. B **710**, 485 (2005), arXiv:hep-ph/0411261v2.

[27] T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, Phys. Lett. B **400**, 379 (1997), arXiv:hep-ph/9701390v1.

[28] P. M. Stevenson, Annals Phys. **132**, 383 (1981).

[29] E. Gardi, G. Grunberg, and M. Karliner, JHEP **9807**, 007 (1998), arXiv:hep-ph/9806462v1.

[30] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, Adv. Comput. Math. **5**, 329 (1996).

[31] K. Hamacher, Effective Charges in Practice, in *FRIF Workshop on First Principles Non-Perturbative QCD of Hadron Jets*, 2006, arXiv:hep-ex/0605123v2. eConf C0601121

[32] A. J. Buras, E. G. Floratos, D. A. Ross, and C. T. Sachrajda, Nucl. Phys. B **131**, 308 (1977).

[33] W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, Phys. Rev. D **18**, 3998 (1978).

[34] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, JHEP **0905**, 106 (2009), arXiv:0903.4658v2 [hep-ph].

[35] DELPHI Collaboration, J. Abdallah *et al.*, Eur. Phys. J. C **29**, 285 (2003), arXiv:hep-ex/0307048v1.

[36] S. Weinzierl, JHEP **0906**, 041 (2009), arXiv:0904.1077 [hep-ph].

[37] S. Weinzierl, Phys. Rev. Lett. **101**, 162001 (2008), arXiv:0807.3241 [hep-ph].

[38] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).

[39] E. Farhi, Phys. Rev. Lett. **39**, 1587 (1977).

[40] G. Parisi, Phys. Lett. B **74**, 65 (1978).

[41] J. F. Donoghue, F. E. Low, and S. Y. Pi, Phys. Rev. D **20**, 2759 (1979).

[42] S. Catani, G. Turnock, and B. R. Webber, Phys. Lett. B **295**, 269 (1992).

[43] P. E. L. Rakow and B. R. Webber, Nucl. Phys. B **191**, 63 (1981).

[44] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. B **269**, 432 (1991).

[45] L. Clavelli, Phys. Lett. B **85**, 111 (1979).

[46] Herwig++, M. Bahr *et al.*, Eur. Phys. J. C **58**, 639 (2008), `arXiv:0803.0883 [hep-ph]`.

[47] Herwig++, S. Gieseke *et al.*, `arXiv:1102.1672v1 [hep-ph]`.

[48] Y. L. Dokshitzer and B. R. Webber, Phys. Lett. B **352**, 451 (1995), `arXiv:hep-ph/9504219v1`.

[49] G. P. Salam and D. Wicke, JHEP **0105**, 061 (2001), `arXiv:hep-ph/0102343v2`.

[50] J. M. Campbell, E. W. N. Glover, and C. J. Maxwell, Phys. Rev. Lett. **81**, 1568 (1998), `arXiv:hep-ph/9803254v2`.

[51] P. A. Baikov, K. G. Chetyrkin, and J. H. Kühn, Phys. Rev. Lett. **101**, 012002 (2008), `arXiv:0801.1821v2 [hep-ph]`.

[52] S. Bethke, Eur. Phys. J. C **64**, 689 (2009), arXiv:0908.1135v2 [hep-ph].

[53] Wolfram Research, Inc., *Mathematica* (Wolfram Research, Inc., Champaign, IL, 2008), Version 7.0.

[54] M. Beneke, Phys. Rept. **317**, 1 (1999), arXiv:hep-ph/9807443v2.

[55] C. J. Maxwell and D. G. Tonge, Nucl. Phys. B **535**, 19 (1998), arXiv:hep-ph/9705314v1.

[56] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory* (Springer, 1999), ISBN: 0-387-98931-5.

[57] P. N. Burrows, Determination of α_s from hadronic event shape observables in e+e- annihilation, in *28th International Conference on High-Energy Physics (ICHEP 96)*, 1996, hep-ex/9612008. eConf C960725

[58] P. N. Burrows *et al.*, Phys. Lett. B **392**, 223 (1997), arXiv:hep-ph/9609513v1.

[59] M. A. Samuel, J. R. Ellis, and M. Karliner, Phys. Rev. Lett. **74**, 4380 (1995), arXiv:hep-ph/9503411v2.

[60] J. R. Ellis, E. Gardi, M. Karliner, and M. A. Samuel, Phys. Lett. B **366**, 268 (1996), arXiv:hep-ph/9509312v1.

[61] Y. Dokshitzer, G. Marchesini, and G. Salam, Eur. Phys. J. direct **C1**, 3 (1999), arXiv:hep-ph/9812487.

[62] Y. Dokshitzer, G. Marchesini, and B. Webber, Nucl. Phys. B **469**, 93 (1996), arXiv:hep-ph/9512336.

[63] R. Davison and B. Webber, Eur.Phys.J. **C59**, 13 (2009), arXiv:0809.3326 [hep-ph].

[64] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. P. Salam, JHEP **9805**, 003 (1998), arXiv:hep-ph/9802381v4.

[65] C. J. Maxwell, arXiv:hep-ph/9809270v1.

[66] J. Chyla, JHEP **0303**, 042 (2003), arXiv:hep-ph/0303179v1.

[67] J. M. Campbell and R. K. Ellis, (2012), arXiv:1204.1513 [hep-ph], FERMILAB-PUB-12-078-T.

[68] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, Phys. Rev. Lett. **102**, 182003 (2009), arXiv:0903.0005 [hep-ph].

[69] J. Campbell, R. K. Ellis, and F. Tramontano, Phys. Rev. D **70**, 094012 (2004), arXiv:hep-ph/0408158.

[70] M. Czakon and A. Mitov, (2011), arXiv:1112.5675 [hep-ph].

[71] http://www.python.org/.

[72] K. Madsen, H. B. Nielsen, and O. Tingleff, *Methods for non-linear least squares problems (2nd ed.)* (Technical University of Denmark, 2004), Lecture notes.

[73] C. T. Kelley, *Iterative Methods for Optimization (Frontiers in Applied Mathematics: Book 18)* (Society for Industrial and Applied Mathematics, 1999), http://www.ec-securehost.com/SIAM/FR18.html.

[74] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009), arXiv:0901.0002v3 [hep-ph].

[75] M. L. Mangano, AIP Conf.Proc **753**, 247 (2005), arXiv:hep-ph/0411020.

[76] Particle Data Group, J. Beringer *et al.*, Phys. Rev. D **86**, 010001 (2012), Review: "Production and decay of b-flavoured hadrons".

[77] M. Cacciari, S. Frixione, M. Mangano, P. Nason, and G. Ridolfi, JHEP **0407**, 033 (2004), arXiv:hep-ph/0312132.

[78] CDF collaboration, A. Abulencia *et al.*, Phys. Rev. D **75**, 012010 (2007), arXiv:hep-ex/0612015.

[79] F. Happacher, P. Giromini, and F. Ptohos, Phys. Rev. D **73**, 014026 (2006), arXiv:hep-ph/0509348.

[80] M. Czakon and A. Mitov, arXiv:1207.0236 [hep-ph].

[81] H.-L. Lai *et al.*, Phys. Rev. D **82**, 074024 (2010).

[82] Particle Data Group, J. Beringer *et al.*, Phys. Rev. D **86**, 010001 (2012), Review: "The top quark".

[83] P. Baernreuther, M. Czakon, and A. Mitov, arXiv:1204.5201 [hep-ph].

[84] http://docs.scipy.org/doc/scipy-0.7.x/reference/generated/scipy.optimize.fmin.html, Details of the built-in SciPy function "fmin" which uses a Nelder Mead algorithm.

[85] Particle Data Group, J. Beringer *et al.*, Phys. Rev. D **86**, 010001 (2012), Summary table: "Quarks".

[86] The CDF Collaboration, "Study of the Top Quark Production Asymmetry and Its Mass and Rapidity Dependence in the Full Run II Tevatron Dataset", CDF Note 10807.

[87] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **85**, 114040 (2012), arXiv:1205.1232 [hep-ph].

[88] J. M. Campbell and R. K. Ellis, arXiv:1204.1513 [hep-ph].

[89] J. R. Ellis, E. Gardi, M. Karliner, and M. A. Samuel, Phys. Rev. D **54**, 6986 (1996), arXiv:hep-ph/9607404.

List of Figures

List of Tables