

## Topical Review

# Chapter 12: Amplitudes and collider physics

**Chris D White\*** 

Department of Physics and Astronomy, Centre for Theoretical Physics,  
Queen Mary University of London, Mile End Road, London E1 4NS,  
United Kingdom

E-mail: [christopher.white@qmul.ac.uk](mailto:christopher.white@qmul.ac.uk)

Received 13 April 2022, revised 28 June 2022

Accepted for publication 2 August 2022

Published 30 November 2022

**Abstract**

We explore how various topics in modern scattering amplitudes research find application in the description of collider physics processes. After a brief review of experimentally measured quantities and how they are related to amplitudes, we summarise recent developments in perturbative QFT, and how they have impacted our ability to do precision physics with colliders. Next, we explain how the study of (next-to-)soft radiation is directly relevant to increasing theoretical precision for key processes at the LHC and related experiments. Finally, we describe the various techniques that are used to turn theoretical calculations into something more closely approaching the output of a particle accelerator.

Keywords: scattering amplitudes, collider physics, quantum chromodynamics

(Some figures may appear in colour only in the online journal)

**Contents**

1. Introduction	2
2. What is a collider, and what does it measure?	3
3. Cross-sections and amplitudes	4
4. Cross-sections at hadron colliders	7
5. Fixed-order perturbation theory	10
6. Resummation	12

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

7. From theory to experiment	18
8. Summary	22
9. Conclusion	24
Acknowledgments	24
Data availability statement	25
References	25

## 1. Introduction

As the SAGEX network has made clear, the study of scattering amplitudes has become an extraordinarily broad endeavour. Theories studied range from well-established theories of nature (e.g. the standard model, plus effective field theory corrections) right through to highly supersymmetric versions of non-abelian gauge theories, or even string theories. The latter are less relevant for current experiments, but can be useful for developing new calculational techniques, or for showing up relationships between different types of theory. Furthermore, new mathematics is often needed in order to bring new aspects of field theory to light, such that dialogue between pure mathematicians and physicists has been an increasing facet of the amplitudes community in recent years. What makes this mix so important is that new techniques used to study more formal theories are very often *the same techniques* that have been used to generate highly non-trivial results in more physically relevant theories. However, the diversity of the amplitudes community can be as much of a hindrance as it is a strength: it becomes ever-more difficult for any given person—and especially a newcomer to the field—to see how different branches of amplitudes research are related, and to know how formal ideas in one corner are applied to very practical ends in another.

The aim of this article is to redress this balance somewhat, by focussing on one of the main motivations for studying amplitudes in the first place, namely that they underlie the physics of collider experiments. The latter form our main way of testing theories of fundamental physics, which itself is at a crucial crossroads. With the discovery of the Higgs boson in 2012, the SM is experimentally complete, but its obvious deficiencies<sup>1</sup> cry out for a deeper explanation. The current flagship experiment is the Large Hadron Collider at CERN, which will be with us for at least another two decades. Replacement colliders are being actively considered, although the lack of any particularly striking new physics hints at the LHC make it rather unclear what the optimum follow-up machine will look like.

In order to make progress, we must continue to develop new types of collider, as well as our understanding of quantum field theory. On very general grounds, any new physics will look QFT-like at currently available collider energies. Furthermore, the lack of any very clear new physics means that we have to understand the old physics (i.e. the SM) extremely well. Only by knowing the old physics *very very precisely* can we be sure that tiny deviations we might see in experiments are genuine new physics, rather than poorly understood old physics. Furthermore, given that collider experiments involve scattering particles, it is the study of *scattering amplitudes* in QFT that is the most relevant thing to be doing.

<sup>1</sup> Examples include the observed matter/antimatter asymmetry in our Universe, quantum instability of the Higgs boson mass (the hierarchy problem), the absence of gravity, lack of explanation of particle masses, dark matter or dark energy, and much else besides.

Given that you are reading this article, I am assuming that you will already be familiar with what a scattering amplitude is, whether or not this is in our own apparently four-dimensional and distinctly non-conformal world. However, you may well not be familiar with what goes on inside a collider, what experimentalists do with it, and how you can start to turn a scattering amplitude into something they might be interested in. Although some of the ideas may well have already been presented to you in some region of your past light-cone, we review relevant material in the following sections.

## 2. What is a collider, and what does it measure?

A *particle accelerator* or *collider* is an experiment that accelerates one or more beams of particles so that they collide. Early versions of this idea had a single beam colliding with a fixed target, which is easier to build, but suffers from the fact that much of the energy in the initial state gets wasted as kinetic energy in the final state. Hence, for the past few decades, colliders have consisted of two beams of particles that collide in the centre-of-mass frame. Electric and magnetic fields are used to accelerate the particles and focus the beams, and thus the particles being accelerated must be (electromagnetically) charged. Recent examples include LEP ( $e^+e^-$ ), the Tevatron ( $p\bar{p}$ ), and the LHC ( $pp$ ), where  $e^\pm$  denotes the electron and positron, and  $p(\bar{p})$  the (anti-)proton. Colliders tend to be circular, which gives multiple chances to accelerate the particles each time they go around the ring. However, charged particles of energy  $E$  and mass  $m$  radiate *synchrotron radiation* at a rate  $\propto (E/m)^4 r^{-2}$ , imposing a practical limit on how much we can accelerate them. This is why the most modern collider (the LHC) uses heavy particles (protons rather than electrons), and has a very large circumference of 27 km! Protons are not ideal—they are composite, wobbly bags of quarks and gluons. This makes hadron colliders messier than  $e^-e^+$  colliders, making precision physics more difficult. Thus, future colliders may return to using electrons and positrons, at the expense of needing new technology for colliding them in a *linear* fashion.

The beams in a particle accelerator are focussed so that they collide, which in modern circular colliders happens at multiple points around the ring. Each collision of the beam particles is referred to as a (*scattering*) *event*, and the aim of each experiment is to record and analyse interesting-looking events. The collision points are surrounded by a cylindrical detector, whose aim is to collect as many of the particles emerging from a collision point as possible. The detectors contain multiple layers for capturing different types of particle, and measuring their momenta and charges. Some particles, such as weakly-interacting neutrinos or possible new physics particles, pass through the detector and thus appear as ‘missing four-momentum’ (i.e. a failure of momentum conservation) in a given scattering event. From a very naïve theorist’s viewpoint, one may thus characterise individual scattering events by a recorded list of particles, together with their measured four-momenta. We will return to how accurate this picture is in section 7, but for now let us note that in a quantum theory it is not possible to predict exactly which particles will emerge in any given event, nor what their precise momenta will be. Instead, we can only calculate *probabilities* for certain final states to appear in the detector, and then compare the measured properties of events with predicted distributions from a given QFT.

If we consider a simple classical scattering process, such as throwing tennis balls at a target, it is clear that the probability of hitting the target depends on its cross-sectional area. This is equally true of quantum scattering, for which the relevant cross-sectional area represents an

effective range of interaction for particles incident on the target. It follows that the rate of scattering events for a given pair of incoming beams can be written as

$$\frac{dN}{dt} \propto \sigma, \quad (1)$$

where  $N$  is the number of events,  $t$  the time, and  $\sigma$  has units of area. It is called the *cross-section*, and depends on the intrinsic properties of the incoming particles, thus is calculable from QFT. The quantity relating the two sides of equation (1) is usually written  $\mathcal{L}(t)$ , as it may depend on time in general. It is called the *luminosity* and, roughly speaking, measures how the incoming particles are distributed within the beams i.e. how ‘bright’ the beams are. Here we have been talking about the total number of scattering events, but we can choose to focus on a particular type of event e.g. those containing a pair of top quarks, or a single Higgs boson. Letting  $N_i$  denote the number of events for distinct processes  $\{i\}$ , we have

$$N = \sum_i N_i \Rightarrow \frac{dN_i}{dt} = \mathcal{L}(t)\sigma_i, \quad (2)$$

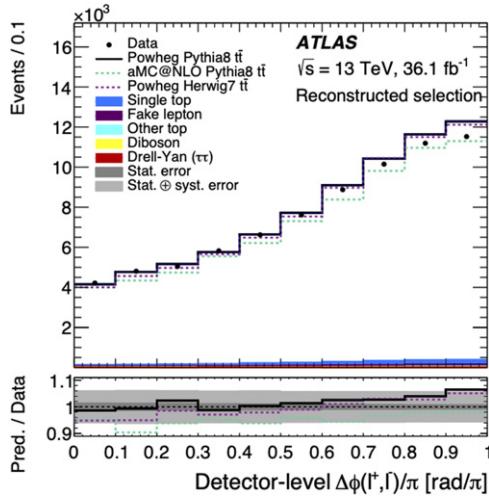
where  $\sigma_i$  is the cross-section for process  $i$ , such that the total cross-section for interaction is  $\sigma = \sum_i \sigma_i$ . It is not possible to think of each individual cross-section as a physical area, but we do not have to do this: equation (2) tells us that the cross-section for a given process is simply given by the event rate, divided by the luminosity. Experimentalists measure the luminosity  $\mathcal{L}(t)$ , so that they can then present results for cross-sections directly, for comparison with theory. For historical reasons, the conventional unit of cross-section is the *barn* (b), where one barn is defined to be  $10^{-28} \text{ m}^2$ . For context, the total  $pp$  cross-section at the LHC (at 13 TeV) is about 0.1 b. The cross-section for single Higgs boson production in its main production mode (gluon–gluon fusion) is about 57 pb (here the ‘p’ stands for *pico*, or  $10^{-12}$ ), which gives some indication of the huge range of cross-sections for interesting processes at the LHC, and how very tiny signals have to be extracted from a massive amount of other data. Given that the luminosity is related to the total event rate, it follows that the *integrated luminosity*

$$L(t) = \int_0^t dt' \mathcal{L}(t')$$

is a measure of the total number of events collected by a given collider after a time of operation  $t$ . From equation (1) and the fact that the event number  $N$  is dimensionless, the integrated luminosity has units of inverse area. As an example, the LHC delivered  $156 \text{ fb}^{-1}$  (inverse femtobarns) of integrated luminosity during run 2 at 13 TeV. Given the inverse units, a smaller prefix (nano, pico, femto) means more events have been collected.

### 3. Cross-sections and amplitudes

In the previous section, we have seen that a given scattering process  $i$  has an associated cross-section  $\sigma_i$ . In principle, QFT tells us how to calculate this. Let us first consider the case that the incoming and outgoing particles are fundamental particles. Let  $p_1$  and  $p_2$  be the incoming particle momenta, and  $\{p_3, \dots, p_n\}$  the final momenta, so that there are  $(n - 2)$  particles in the final state. Then your QFT textbook will tell you that the cross-section can be decomposed into the following form:



**Figure 1.** Distribution of the azimuthal angle between leptons originating from top quark decays. Reproduced from [1]. CC BY 4.0.

$$\sigma_i = \frac{1}{F(p_1, p_2)} \int d\Phi^{(n-2)}(p_3, \dots, p_n) \mathcal{A}_i(\{p_i\}). \quad (3)$$

Here

$$F(p_1, p_2) = 4[(p_1 \cdot p_2)^2 - m_1^2 m_2^2]^{1/2}, \quad (4)$$

where  $m_i$  is the mass of beam particle  $i$ , is called the *Lorentz-invariant flux*. It is a convenient relativistic generalisation of the usual flux describing how the beam particles are moving relative to each other. There is also an integration over the momenta of the final state particles, given by the *Lorentz-invariant phase space*

$$\int d\Phi^{(n-2)} = (2\pi)^4 \left( \prod_{i=1}^{n-2} \int \frac{d^3 \vec{p}_i}{(2\pi)^3 2E_i} \right) \delta^{(4)}(P_{\text{final}} - P_{\text{init.}}). \quad (5)$$

Here  $E_i$  is the energy of particle  $i$ , and  $\vec{p}_i$  its (relativistic) three-momentum. Also,  $P_{\text{init.}}$  ( $P_{\text{final}}$ ) are the total four-momentum in the initial (final) state respectively. We then see that the delta function in equation (5) implements total momentum conservation in the scattering process. Finally in equation (3),  $\mathcal{A}_i$  is the *scattering amplitude* for process  $i$ , which the SAGEX collaboration was set up to investigate in a myriad of ways. Equation (3) suggests how you as a theorist might compare a given QFT with experiment: for a given process, you can calculate the appropriate scattering amplitude using perturbation theory, before integrating over the phase-space, and dividing by the appropriate flux factor. The resulting number may then be compared with something produced by experimentalists, if they have isolated the relevant events. The process might be a potential new physics process, and if the number agrees you may end up with an expenses-paid trip to Stockholm. Or it might be a process that occurs in the standard model, so that your number allows an experimentalist to efficiently disentangle new from old physics, so that they might go to Stockholm.

Total cross-sections are a very crude way of comparing theory with data. It is also possible to compare *differential cross-sections*, where one does not integrate over the full phase

space in equation (3). Given an observable  $\mathcal{O}$ , the quantity  $(d\sigma_i/d\mathcal{O})$  represents the distribution of  $\mathcal{O}$  as measured in events collected in scattering process  $i$ . Experimentalists often normalise this by dividing by the total cross-section, given that many uncertainties cancel out when doing this. An example distribution is shown in figure 1, which comes from an analysis of events containing (anti-)top quark pairs by the ATLAS experiment [1]. The (anti-)top quarks decay to produce leptons, and the collaboration have measured the azimuthal angle in the detector between these decay products, where the data is compared to various predictions from theorists. Interestingly, the theory curves do not agree with the data particularly well by eye. Statistically, however, there is no cause for excitement: the lower panel depicts the uncertainty on the measurement as a grey band, which indeed overlaps with the (standard model) theory.

It goes almost without saying that the above discussion is vastly oversimplified. The comparison of theory with data is, in practice, almost nothing like the simple procedure outlined here. Typical complications include the following:

- Amplitudes in QFT suffer from ultraviolet (UV) divergences at loop-level. These can be dealt with via renormalisation in some scheme, and modern-day theorists typically use *dimensional regularisation* in  $4 - 2\epsilon$  dimensions. Then cross-sections for comparison with experiment will depend on the renormalisation scale  $\mu_R$ . In particular, the QCD expansion parameter will become a *running coupling*  $\alpha_s(\mu_R)$ . The coefficients of the perturbation expansion in this parameter involve large logarithms involving ratios of  $\mu_R$  with energy scales typical of the scattering process (e.g. invariants formed from the particle momenta). Thus, dependence on  $\mu_R$  is minimised at a given order in perturbation theory by choosing  $\mu_R$  to be comparable with these energy scales.
- The incoming particles at the LHC are protons, which are not fundamental particles in QCD, and thus not fully described by standard perturbation theory. We must then generalise the formula of equation (3), given that we can calculate scattering amplitudes in perturbation theory for incoming (anti-)quarks and/or gluons, but not protons. We discuss this in section 4.
- We can typically only calculate the amplitude to low orders in perturbation theory, where the strong interaction (with expansion parameter  $\alpha_s \equiv g_s^2/(4\pi)$ , where  $g_s$  is the QCD coupling constant) is the most relevant. The state of the art for many processes of interest is next-to-leading order (NLO) in  $\alpha_s$ , although NNLO is known for some cases. Some cross-sections are known to a highly impressive N<sup>3</sup>LO! Examples include different production modes for the Higgs boson [2–5], and the so-called *Drell–Yan* production of a heavy particle [6], which we will see in more detail below. Proceeding order-by-order in the coupling in this way is known as *fixed-order perturbation theory*, and we discuss this in section 5.
- Perturbation theory can be unstable: for differential cross-sections, the coefficients of the perturbation series depend on the external particle momenta, and can diverge in certain kinematic regions (such as the *soft limit*, in which the four-momentum of emitted radiation goes to zero). One must then sum up certain contributions to all orders in perturbation theory for meaningful comparison with data, a process known as *resummation*, discussed in section 6.
- Real scattering events contain huge numbers of quarks and gluons, going way beyond what a perturbative QCD calculation can achieve. One must therefore estimate the effect of this additional radiation using well-motivated QFT-based arguments. We return to this in section 7.

- Real scattering events do not contain free (anti-)quarks and gluons, due to the confinement property of QCD. They instead contain hadrons, and we must try and estimate how. This is also discussed in section 7.

This list—which is not even complete—is enough to scare off any theorist who hopes that their research on scattering amplitudes might have something to say about experiments. Let us thus retreat to the relative safety of equation (3), and get gradually closer to what happens at the LHC.

#### 4. Cross-sections at hadron colliders

The first thing we have to deal with is the issue of incoming protons. At first glance, one might think that we could never calculate cross-sections with incoming protons, given that—with a few exceptions in special circumstances or theories—we only really know how to calculate amplitudes using perturbation theory. However, what saves us is the well-known property of *asymptotic freedom* [7, 8], according to which the QCD expansion parameter  $\alpha_s(\mu_R)$  becomes weaker with increasing renormalisation scale  $\mu_R$ . Given that this is identified with a typical energy scale  $Q$  in the scattering process of interest, we find that highly energetic protons will contain approximately free (anti-)quarks and gluons. We can indeed calculate a perturbative cross-section for the latter, and may then combine it with appropriate distribution functions measuring how the (anti-)quarks are distributed within the proton, in order to describe the total proton–proton cross-section. This was first formulated by Feynman [9, 10], who called it the *parton model*. At that time, partons were stipulated as hypothetical elementary constituents of the proton, and were only later realised to be the (anti-)quarks and gluons of QCD. The word ‘parton’, however, is still in use, as it is very convenient to have a single word that can be used to denote any (anti-)quark or gluon!

More precisely, consider partons that emerge from the incoming protons, where the latter can be taken to have momenta  $P_1$  and  $P_2$ . If the protons are very fast-moving, the emerging partons will have momenta  $\{p_i\}$  that are approximately collinear with the  $\{P_i\}$ , and which may thus be written as

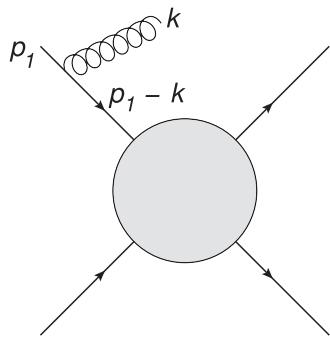
$$p_i = x_i P_i, \quad 0 \leq x_i \leq 1, \quad (6)$$

where no summation is implied on the right-hand side, and where  $x_i$  represents the (longitudinal) momentum fraction of parton  $i$ . The hadronic cross-section for a given process may then be written as<sup>2</sup>

$$\sigma = \sum_{i,j \in \{q, \bar{q}, g\}} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1) f_j(x_2) \hat{\sigma}_{ij}(\{p_i\}, \mu_R). \quad (7)$$

Here  $\hat{\sigma}_{ij}$  denotes the *partonic cross-section*, for incoming (anti-)quarks and gluons. The quantity  $f_i(x_j)$  is a *parton distribution function (PDF)*, that represents, loosely speaking, the probability to find a parton with momentum fraction  $x_j$  inside incoming proton  $i$ . We must then sum over all possible momentum fractions, and over all possible species of parton, including over all possible flavours of quark  $q$  and their anti-particles. The latter might surprise you if you are familiar only with the conventional undergraduate wisdom that protons contain two up quarks and a down quark. Needless to say, this statement is not to be taken too literally: in quantum

<sup>2</sup> To simplify the notation in equation (7), we have omitted the subscript on  $\sigma$  that we used above to denote a particular scattering process.



**Figure 2.** Gluon radiation from an incoming quark leg.

field theory, partons are constantly popping in and out of the vacuum due to virtual effects, so that the proton is some sort of localised collection of quantum partonic froth, whose *net* quantum numbers are the same as two up quarks and a down quark! While the partonic cross-section  $\hat{\sigma}_{ij}$  is calculable in perturbation theory, the parton distribution functions  $\{f_i\}$  are not. They can, however, be measured from experiment, and used to predict the results of subsequent experiments.

You may be wondering if the parton model can be rigorously justified from first-principles QFT, where the relevant field theory is QCD in this case. Indeed it can for certain scattering processes, using the operator product expansion (see e.g. reference [11] for a textbook treatment). More simply, we might ask what goes wrong if we were to ignore equation (7), and try to use perturbation theory to describe QCD scattering. At leading order (LO) in  $\alpha_s$ , everything is well-behaved. However, at NLO and beyond, problems occur due to emitting additional radiation from either incoming or outgoing partons. An example is shown in figure 2 which, using Feynman rules (or otherwise) results in an extra factor in the scattering amplitude for a given process

$$\frac{1}{(p_1 - k)^2} = \frac{1}{-2p_1 \cdot k} = \frac{1}{-2|\vec{p}_1| |\vec{k}| (1 - \cos \theta)}, \quad (8)$$

where four-momenta are labelled as in the figure, and we neglect the mass of the incoming quark, given that this is negligible at all relevant collider energies (at least for the up and down quarks). We have also introduced the three-momenta of the quark and gluon,  $\vec{p}_1$  and  $\vec{k}$ , as well as the angle  $\theta$  between them. Clearly the additional factor of equation (8) diverges in two cases: (i) the additional radiation is *soft* ( $|\vec{k}| \rightarrow 0$ ); (ii) the additional radiation is *collinear* with the emitting particle ( $\theta \rightarrow 0$ ). Furthermore, these cases are not mutually exclusive: the radiation may be soft *and* collinear, which is even more divergent! These are known as *infrared (IR)* divergences, to distinguish them from the UV divergences encountered above. Although we considered the example of real radiation here, the virtual (loop) corrections to the amplitude are also IR-divergent, and the solution to this problem is well-known. First, one must be careful to consider only those cross-sections and related observables that are *infrared-safe*, meaning that the definition is robust under the inclusion of additional soft and/or collinear radiation. For example, the cross-section for a single Higgs boson accompanied by any amount of QCD radiation is IR-safe, whereas the cross-section for a Higgs boson plus exactly three detected gluons is not. Once we have chosen an IR-safe observable, the divergences are guaranteed to cancel in QED by the *Bloch–Nordsieck theorem* [12], if we add

together all virtual and real contributions at any given order in perturbation theory. In QCD, the *KLN theorem* [13, 14] theorem states that this cancellation will only occur if we include initial states containing arbitrary numbers of incoming particles, which is inconvenient for describing scattering processes with two incoming beams. Instead requiring only two incoming particles leads to cancellation of all soft singularities, together with collinear singularities associated with the outgoing particles in the final state. However, one is left with uncancelled collinear singularities associated with the two incoming particles, which potentially pose a serious problem.

What saves us is the presence of the parton distributions  $\{f_i\}$  in equation (7). We are clearly free to redefine both the partons and the partonic cross-section such that equation (7) remains invariant. Thus, we can choose to remove collinear divergences from the partonic cross-section by absorbing them into the parton distributions. In other words, one replaces the *bare parton distributions* appearing in equation (7) with modified ones, chosen so that singularities in the partonic cross-section are removed. One way to do this is simply to dictate that all radiation below a particular *factorisation scale*  $\mu_F$  is reabsorbed into a redefined parton distribution, so that it is absent from the partonic cross-section. However, this procedure is typically not used in practical calculations, given that applying momentum cut-offs is not gauge invariant. Instead, dimensional regularisation is often used in  $d = 4 - 2\epsilon$  dimensions, in which IR divergences show up as poles in  $\epsilon$ . The factorisation scale then emerges through the additional scale  $\mu$  that is introduced to keep the coupling dimensionless. A particular means of removing divergences from the partonic cross-section is called a *factorisation scheme* and, while the partonic cross-section becomes finite, the calculated parton distributions are not. This is not a problem: we do not claim to be able to calculate the partons in perturbation theory, and so instead can simply measure them from experiment. Indeed, this is reminiscent with what happens when we remove UV singularities via renormalisation: in doing so, we lose the ability to calculate the coupling constant, which instead becomes a measurable parameter in some given scheme.

The upshot of the above discussion is that equation (9) gets replaced by the more general formula

$$\sigma = \sum_{i,j \in \{q, \bar{q}, g\}} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu_F) f_j(x_2, \mu_F) \hat{\sigma}_{ij}(\{p_i\}, \mu_R, \mu_F) + \mathcal{O}\left(\frac{\Lambda^2}{Q^2}\right). \quad (9)$$

That is, both the partons and partonic cross-section now depend on the factorisation scale  $\mu_F$ , which typically enters through dimensionless ratios involving an energy scale  $Q$  associated with the scattering process. As for the renormalisation scale  $\mu_R$ , dependence on  $\mu_F$  cancels at a given order in perturbation theory, but there can be a residual dependence associated with missing higher-order corrections, and this takes the form of logarithms of ratios of the factorisation scale with typical energy scales involved in the scattering process. As for the case of  $\mu_R$  above, one can reduce this dependence by choosing  $\mu_F$  to be such a typical energy scale  $Q$ , and there is clearly some ambiguity in this choice. Varying  $\mu_R$  and  $\mu_F$  around some default value then gives a measure of the *theoretical uncertainty* of a given cross-section prediction. We have also noted in equation (9) that the parton model formula receives formal corrections when derived from first-principles QFT, involving ratios of the QCD confinement scale  $\Lambda$  and the typical energy scale  $Q$ . Provided that  $Q$  is large enough that we trust perturbation theory in the first place, these so-called *power corrections* will be small.

## 5. Fixed-order perturbation theory

So far the job of a theorist seems clear enough: to compare theory to data at a hadron collider such as the LHC, we must calculate amplitudes involving incoming partons [i.e. (anti-)quarks and gluons], turn these into partonic cross-sections using equation (3), and then combine these with parton distributions using one of the many versions of these which are available (see e.g. references [15–18] for recent analyses). However, anyone who has tried to calculate scattering amplitudes for relevant scattering processes knows that this becomes extraordinarily difficult as the number of loops or number of legs increases. Traditionally, QCD amplitudes were calculated using Feynman rules and diagrams, which represent the  $L$ -loop amplitude as a sum of integrals of the form

$$\mathcal{A}^{(L)}(\{p_i\}) = \sum_{\text{diags } j} \left( \prod_{k=1}^L \int \frac{d^d l_k}{(2\pi)^d} \right) \frac{c_j \mathcal{N}_j(\{p_i\}, \{l_k\})}{\mathcal{D}_j(\{p_i\}, \{l_k\})}. \quad (10)$$

Here each individual diagram  $j$  has a colour factor  $c_j$ , and integrals over the  $L$  loop momenta  $\{l_k\}$ . There is then a kinematic numerator  $\mathcal{N}_j$  and denominator  $\mathcal{D}_j$ , each of which depends on both the external momenta  $\{p_i\}$  and loop momenta  $\{l_k\}$  in general. Terms in the integrand with zero, one or more than one powers of loop momenta in the numerator are referred to as *scalar*, *vector* or *tensor* integrals respectively, and it is in fact possible to apply algebraic identities to reduce all higher-rank integrals to scalar ones. This may be formalised in a process known as *Passarino–Veltman reduction* [19], which may be automated in principle, albeit with some subtle caveats for higher-leg processes. Then, all  $L$ -loop amplitudes can in principle be straightforwardly calculated once a suitable basis for all  $L$ -loop scalar integrals is known, a problem which is entirely solved for  $L = 1$  (see e.g. reference [20]). Roughly speaking, most QCD calculations up until the mid-2000s were calculated using this approach, which ultimately proved to be unsustainable, not least due to the fact that the number of Feynman diagrams grows factorially as the order in perturbation theory increases.

More modern methods for calculating amplitudes rely on the observation that, once a basis of  $L$ -loop scalar integrals is known, one can simply expand a given amplitude in terms of this basis, and then develop clever methods for fixing the coefficients that bypass Feynman diagrams altogether. A heavily-used method is that of *generalised unitarity* [21–26], which involves first classifying the list of scalar integrals according to how many internal lines (or *propagators*) they have. Next, one may perform *unitarity cuts*, consisting of placing one or more propagators on-shell. If one takes the maximal number of such cuts at a given loop order, only the integrals with the maximal number of internal lines survive, so that one may straightforwardly fix their coefficients. One may then iterate the procedure to fix the integrals with one less than the maximal number of internal lines, and so on. The so-called *cut-constructible* part of the amplitude thus obtained must then be supplemented by an additional *rational piece*, for which various methods exist [27–34, 34, 35] (see e.g. reference [36] for a comprehensive review). Ultimately, the effect of such tools is to reduce the computational complexity of multiloop calculations from being factorial in the perturbative order, to being merely polynomial. Indeed, the calculation of one-loop processes in QCD has now been automated, such that general purpose computer programs exist that allow users to specify a given initial and final state (see e.g. references [37–39], or references [40–42] for their latest incarnations). One then presses a button, waits for a reasonable (but not too inconvenient) length of time, and receives results for (differential) cross-sections!

At two-loop level and beyond, a full basis of scalar integrals is not known, and discussion is still ongoing regarding the best way to systematise calculations. One can of course focus

on specific scattering processes rather than trying to make general statements. Then a reduced set of integrals appear, but the difficult problem remains of carrying out the integrals themselves, in terms of known functions. In recent years, this has generated a fascinating dialogue between pure mathematicians and theoretical physicists, and more details can be found in references [43–45] (chapters 3–5 of this review [46]). What is clear, however, is that there is a clear direction of travel of new computational techniques on average. They tend to be developed first in a formal `hep-th` context, typically due to the fact that it is easier to probe new structures in highly symmetric theories such as  $\mathcal{N} = 4$  super-Yang–Mills. However, these same techniques then filter down to more immediately applicable theories such as QCD, and become a standard part of the `hep-ph` toolkit. Very often, individual researchers will be working in both subfields, providing a vibrant counterpoint to the once-prevalent notion that ‘theory’ and ‘phenomenology’ are distinct activities, whose practitioners are not able—or unwilling—to communicate with each other.

The calculation of a given scattering process at higher orders may proceed completely analytically, or (partially) numerically (e.g. numerical integration of scalar integrals may be used). This may seem distasteful to those formal theorists who are used to being able to fully probe the analytic structure of their results. But numerical results are ultimately what is needed anyway to compare with experiment, and numerical methods have allowed us to carry out calculations that otherwise would simply not have been possible. However, the use of numerical methods raises a significant issue: as discussed above, the calculation of physical observables involves infrared singularities at intermediate stages. Although these cancel in final results, and may be regulated in separate parts of a calculation (e.g. cross-section contributions involving different numbers of loops), one cannot combine numerical results for individually IR-divergent parts of an observable, due to the large numerical uncertainties that result. One must thus devise a suitable *infrared-subtraction scheme* for organising the perturbative calculation, such that large numerical cancellations never appear. This problem was solved a long time ago at one-loop (for popular schemes, see references [47, 48]), and the Frontier is how to systematically accomplish this at two-loop level and beyond [49–67].

The motivation of a phenomenologist can be very different to a more formal theorist: the choice of which amplitude or scattering process to consider next is much more likely to be guided by experiment. More specifically, members of the two large general-purpose search experiments at the LHC (ATLAS and CMS) have often produced wishlists of scattering processes for which they would like higher-order corrections. A recent example can be found in reference [68], which also reviews state-of-the-art calculational tools. Various processes are discussed, and requests made for both QCD and electro-weak (EW) corrections. The interplay between including both QCD and EW perturbative information needs some careful accountancy (see reference [69] for a review). Furthermore, differential cross-sections are needed as well as total cross-sections. These typically lag behind total cross-sections in terms of the available precision: the presence of more resolved momenta in the final state means that there are more energy scales in the problem, which complicates the computation of relevant loop integrals, as well as the general book-keeping of the calculation itself.

The absolute cutting-edge in fixed-order perturbation theory is  $N^3\text{LO}$  [2–6], and the vast majority of interesting processes are still known only at NLO. For total cross-sections, only a few orders in perturbation theory are typically needed in order to start to approach a theoretical uncertainty of sub-percent level (as estimated by variation of the factorisation and renormalisation scales, whose residual dependence tells us about missing higher-order corrections). However, for differential quantities serious problems can occur, in that the perturbation expansion becomes unstable. We explore this in the following section.

## 6. Resummation

For a classic example of how perturbation theory can go awry, let us consider the *Drell–Yan process*, in which an off-shell photon is produced that eventually decays to a lepton pair. The LO Feynman diagram is shown in figure 3(a), where we label four-momenta as shown. It is then conventional to define the variable

$$z = \frac{Q^2}{s}, \quad (11)$$

which may be loosely interpreted as the fraction of the partonic centre of mass energy  $s$  that is carried by the photon (of virtuality  $Q^2$ ).

The LO differential partonic cross-section in this variable then has the form

$$\frac{d\hat{\sigma}}{dz} = \sigma_0 \delta(1 - z), \quad (12)$$

for some  $\sigma_0$ , where the delta function reflects the fact that the photon is carrying all of the energy in the final state at LO, so that we must have  $z = 1$ . In computing high-order corrections, we must define the total cross-section for Drell–Yan production to include any amount of additional QCD radiation, in line with our earlier comments regarding infrared safety. We must then include both virtual and real corrections, examples of which are shown in figures 3(b) and (c). These are individually infrared divergent, such that any formal singularity cancels upon combining all real and virtual graphs (bar those initial-state collinear singularities that must be absorbed into the parton distributions). However, upon combining all contributions, the NLO differential cross-section for the  $q\bar{q}$  initial state turns out to have the following form:

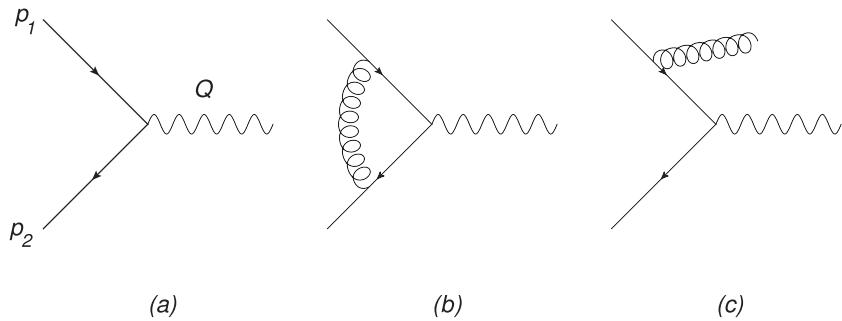
$$\begin{aligned} \frac{d\hat{\sigma}_{q\bar{q}}^{(1)}}{dz} = \frac{\alpha_s C_F}{2\pi} & \left[ 4(1 + z^2) \left( \frac{\log(1 - z)}{1 - z} \right)_+ - 2 \frac{1 + z^2}{1 - z} \log(z) \right. \\ & \left. + \delta(1 - z) \left( \frac{2\pi^2}{3} - 8 \right) \right], \end{aligned} \quad (13)$$

where  $C_F$  is a colour-dependent constant, and the  $+$  notation denotes a so-called *plus distribution*, defined by its action on a test function  $g(z)$  as:

$$\int_0^1 dz f_+(z) g(z) \equiv \int_0^1 dz f(z) [g(z) - g(1)]. \quad (14)$$

We now have a non-trivial dependence on the variable  $0 \leq z \leq 1$ , reflecting the fact that additional radiation may carry away some of the energy in the final state. However, we also see that equation (13) contains a term that is highly divergent as  $z \rightarrow 1$ , involving a logarithm of  $\xi \equiv (1 - z)$  divided by  $\xi$  itself. Finiteness of the total cross-section is guaranteed by equation (14), but it is still the case that the logarithmic term becomes extremely large near  $z \rightarrow 1$ . This threatens the validity of perturbation theory: expansion in the coupling  $\alpha_s$  only makes sense if the coefficients of the expansion are small. At the very least, we should look at higher orders in perturbation theory and see how the coefficients behave as  $z \rightarrow 1$ . Unfortunately, however, the problem gets even worse! At NNLO, for example, we have

$$\frac{d\hat{\sigma}_{q\bar{q}}^{(2)}}{dz} = \left( \frac{\alpha_s C_F}{2\pi} \right)^2 \left[ 128 \left( \frac{\log^3(1 - z)}{1 - z} \right)_+ - 256 \left( \frac{\log(1 - z)}{1 - z} \right)_+ + \dots \right], \quad (15)$$



**Figure 3.** (a) LO Feynman diagram for the production of an off-shell vector boson by a quark–antiquark pair; (b) a virtual correction at NLO; (c) a real correction at NLO.

where the ellipsis denotes terms that are suppressed by a power of  $(1 - z)$ . The highest power of the log in  $(1 - z)$  has gotten larger, and one may indeed show that this pattern recurs at higher orders, so that perturbation theory is indeed breaking down.

To understand the origin of the problem, note that  $z \rightarrow 1$  amounts to the photon carrying all the final state energy, so that any additional radiation is forced to be soft. Although the formal soft divergences cancelled when we added together real and virtual contributions, the large logarithms are ‘echoes’ of the fact that these singularities were present in the first place. Here we have considered the case where additional radiation is forced to be soft. In other scattering processes, it may also be the case that the radiation is forced to be collinear in some kinematic regions, and this also ends up leading to large logs. Indeed, the generic nature of this explanation suggests the following general remarks. For processes involving heavy particles being produced near threshold (i.e. with only just enough four-momentum), we can define a *threshold variable*  $\xi$ , such that  $\xi \rightarrow 0$  near threshold (in our above example, we had  $\xi = (1 - z)$ ). Then, the general structure of the partonic differential cross-section in  $\xi$  can be shown to be

$$\frac{d\hat{\sigma}}{d\xi} = \sum_{n=0}^{\infty} \alpha_s^n \sum_{m=0}^{2n-1} \left[ c_{nm}^{(0)} \left( \frac{\log^m \xi}{\xi} \right)_+ + c_{nm}^{(1)} \log^m \xi + \dots \right]. \quad (16)$$

The first set of terms on the right-hand side generalise the large terms we have already seen in equations (13) and (15). They are called *leading power (LP)* threshold logs, and correspond to the emission of purely soft and/or collinear radiation. The remaining terms constitute a systematic expansion in  $\xi$ , such that the second set of terms comprises the *next-to-leading power (NLP)* contributions. The ellipsis then denotes terms which are NNLP and beyond. It is then genuinely true that we cannot trust perturbation theory as  $\xi \rightarrow 0$ . Furthermore, this is not just a problem in principle, but very much one in practice: there are many observables at the LHC where the instability of perturbation theory becomes a problem.

The solution sounds impossible at first. We must somehow work out what the large logarithms are to all orders in  $\alpha_s$ , and sum them up to get a function of  $\alpha_s$  that is much better behaved than any fixed-order perturbation expansion. You will be familiar with this idea from your undergraduate days. Consider, for example, the toy function

$$e^{-\alpha_s x} = \sum_{n=0}^{\infty} \frac{\alpha_s^n (-x)^n}{n!}. \quad (17)$$

Each term on the right-hand side diverges as  $x \rightarrow \infty$ , but the left-hand side is perfectly well-behaved. What's more, this example is not quite as trivial as it might first appear: when the threshold logarithms in equation (16) are summed up, they do indeed tend to exponentiate. What makes this possible is the fact that we can understand the large logs as being related to infrared divergences. The latter, as is well-known, can be classified to all orders in perturbation theory (see e.g. reference [70] for a review of this subject), so that summing threshold logs in perturbation theory amounts to classifying the structure of infrared singularities.

The process of summing up the large logs is known as *resummation*, and typically proceeds as follows. First, one can take the most divergent logs at each order in  $\alpha_s$ , which at LP go as  $\alpha_s^n \log^{2n-1} \xi/\xi$ . These are referred to as *leading logarithmic (LL)* terms, and are the easiest to resum. Next, one may consider the *next-to-leading logarithmic (NLL)* terms  $\sim \alpha_s^n \log^{2n-2} \xi/\xi$  to all orders, and so on. This expansion implies a reordering of perturbation theory, and it is this that the prefix ‘re-’ in ‘resummation’ is meant to signify. There are by now many different approaches to resumming logs in QCD, involving diagrammatic arguments [71–77], use of Wilson lines [78, 79], renormalisation group arguments [80], and effective field theory [81–84] (see references [85–87] for pedagogical reviews). Common to all of them, however, is the fact that soft and collinear radiation essentially factorises from the underlying hard scattering process. That is, a general  $n$ -point amplitude dressed by soft and collinear radiation has the following form (see e.g. reference [88]):

$$\mathcal{A}_n = \mathcal{H}_n \mathcal{S} \frac{\prod_{i=1}^n J_i}{\prod_{i=1}^n \mathcal{J}_i}. \quad (18)$$

Here  $\mathcal{H}_n$  is a so-called *hard function*. It depends on the particular scattering process, but is infrared finite. The *soft function*  $\mathcal{S}$  collects all singularities associated with purely soft radiation, and the *jet functions*  $J_i$  collect collinear singularities associated with external leg  $i$ . However, radiation can be both soft and collinear, which means it has been counted twice by being included in both the soft and jet functions. Thus, one must divide by *eikonal jet functions*  $\mathcal{J}_i$  to remove the double-counting. It turns out that the soft and jet functions have universal definitions, which are independent of the particular scattering process being considered. The physics of this is indeed straightforward: soft radiation has vanishing momentum, and collinear radiation has vanishing momentum transverse to the relevant particle direction. In both cases, the emitted radiation thus has an infinite Compton wavelength, and cannot resolve the details of the underlying hard scattering process. We then expect it to factorise, leading to something like equation (18).

Studying the soft function in more detail allows us to make contact with more formal amplitudes literature. In QCD, one may define it as a vacuum expectation value of Wilson line operators. That is, if  $\{\beta_i\}$  are the four-velocities of the coloured external particles, one may write

$$\mathcal{S}(\{\beta_i\}) = \langle 0 | \Phi_1 \Phi_2 \dots \Phi_n | 0 \rangle, \quad (19)$$

where

$$\Phi_i = \mathcal{P} \exp \left[ ig_s \int_{x_i=\tau_i \beta_i} dx_i^\mu A_\mu^a \mathbf{T}_i^a \right] \quad (20)$$

is the usual gauge theory Wilson line consisting of the integral of the gauge field along a curve, which has been taken in each case to be the classical straight-line trajectory  $x_i = \tau_i \beta_i$  ( $0 \leq \tau_i \leq \infty$ ) associated with the four-velocity  $\beta_i$ . Also  $\mathbf{T}_i^a$  is a colour generator associated

with external line  $i$ , and  $\mathcal{P}$  denotes path-ordering of the colour generators along the contour. One way of seeing why equation (19) is correct is to note that each of the Wilson line factors contains a coupling of the gauge field to each coloured external particle. By Fourier transforming the exponent of equation (20) to momentum space, we can then interpret the Feynman rule for this coupling. To this end, we may rewrite the exponent of equation (20) as

$$ig_s \mathbf{T}_i^a \int dx_i^\mu \int \frac{d^d k}{(2\pi)^d} \tilde{A}_\mu^a(k) e^{ik \cdot x_i} = \int \frac{d^d k}{(2\pi)^d} \tilde{A}_\mu^a(k) \left[ ig_s \mathbf{T}_i^a \beta_i^\mu \int_0^\infty d\tau_i e^{i(k \cdot \beta_i) \tau_i} \right].$$

It is straightforward to carry out the integral over the parameter  $\tau_i$ , and we obtain<sup>3</sup>

$$\int \frac{d^d k}{(2\pi)^d} \tilde{A}_\mu^a(k) \left[ -g_s \mathbf{T}_i^a \frac{\beta_i^\mu}{\beta_i \cdot k} \right]. \quad (21)$$

The square bracket contains the *eikonal Feynman rule* for the emission of soft radiation, that enters the well-known *soft theorem* in gauge theory [89], describing the emission of soft radiation. The above argument is somewhat technical, but there is a more physical way to see that the soft function should be given by equation (19). If the coloured external particles are emitting purely soft radiation, they cannot recoil against anything, and so by definition must be following classical straight-line trajectories. Then, the only quantum behaviour they are allowed to have is to experience a phase change which, if it is to have the right gauge-covariance properties to form part of a scattering amplitude, must be described by a Wilson line operator, which is known to transform covariantly.

We see, then, that classifying the structure of infrared singularities in scattering amplitudes amounts to studying vacuum expectation values of Wilson lines. Computations involving Wilson lines occur widely throughout the formal amplitudes literature, not least due to the well-known duality between scattering amplitudes in  $\mathcal{N} = 4$  super-Yang–Mills theory, and certain Wilson loops formed from the particle momenta [90]. A key property of Wilson lines that we need for our present purposes is that equation (19) is subject to UV singularities, associated with the cusp at which the Wilson lines meet. These then correspond to the IR singularities of the original amplitude. To see how this works, consider some external particles that are emitting virtual radiation, as in e.g. figure 3(b). Relative to leading order, there will be an additional propagator on the upper line

$$\sim \frac{1}{(p_1 - k)^2} = \frac{1}{-2p_1 \cdot k + k^2},$$

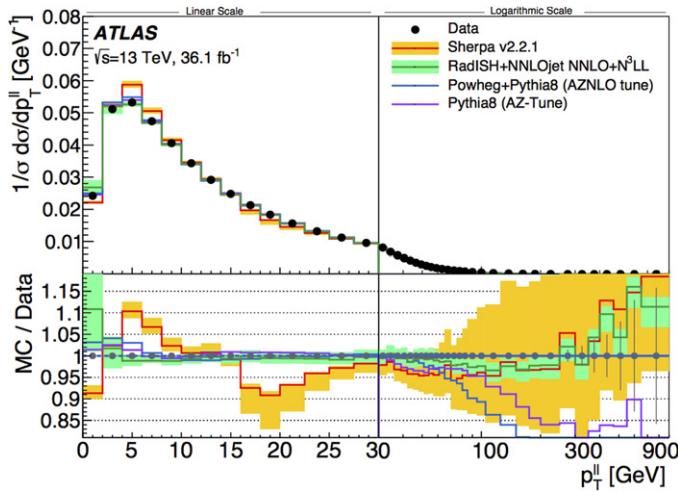
where  $k$  is the four-momentum of the exchanged gluon. In the soft limit, one linearises the propagator by neglecting the term of  $\mathcal{O}(k^2)$ , leading to the eikonal Feynman rule dependence of equation (21). However,  $k$  is a loop momentum, and thus replacing the denominator in this fashion changes the behaviour of the integrand as  $k$  becomes large. Taking also the lower line into account, one modifies the denominators in the loop integral over  $k$  as follows:

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2(-2p_1 \cdot k + k^2)(2p_2 \cdot k + k^2)} \rightarrow \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2(-2p_1 \cdot k)(2p_2 \cdot k)}. \quad (22)$$

<sup>3</sup> The upper limit of the integral in  $\tau_i$  vanishes once the Feynman  $i\varepsilon$  prescription is properly taken into account.

The original integral is UV-finite, but the modified integral is logarithmically divergent. Hence, taking the soft approximation has introduced a *spurious* UV divergence. Furthermore, the second integral turns out to be scaleless in dimensional regularisation, and thus formally vanishes. Thus, the spurious UV singularity precisely cancels the original IR singularity we are interested in. It follows that the UV singularity of a Wilson line integral (involving the linear propagators) matches the IR singularity of the amplitude, and this property in fact generalises to all orders in the soft function.

The UV singularities of VEVs of Wilson lines are controlled by a quantity called the *soft anomalous dimension*, otherwise known as the *cusp anomalous dimension* when only two Wilson lines are involved (again, see reference [70] for a review). Its specific form depends upon the particular theory being considered, and whether there are only two Wilson lines, or many. From the above discussion, it follows that knowledge of the soft anomalous dimension is a key ingredient in being able to perform resummation: (i) the soft anomalous dimension determines the UV singularities of Wilson lines; (ii) these in turn are directly related to the IR singularities of amplitudes; (iii) classifying the latter allows us to resum large logarithms to all-orders in perturbation theory, thus getting sensible results for certain observables, that we can compare with experiments. Further anomalous dimensions control the behaviour of collinear singularities, but we stress the soft anomalous dimension here given that it is a widely studied quantity in a variety of formal contexts, where the people involved may not have realised its role in helping us to get the most out of LHC data: each successive order in the soft anomalous dimension allows us to sum up a further tower of large logs (NLL, NNLL, NNNLL...) in cross-sections of interest. The state of the art in QCD is three-loop order for processes involving only two coloured particles at LO [91–99] (four loops in QED [100]). For many Wilson lines, the soft anomalous dimension is known at three-loop order [48, 88, 101–124] if the Wilson lines are lightlike, as would be relevant for the scattering of massless particles. For non-lightlike Wilson lines (which would be relevant for e.g. top quark pair production), the state-of-the-art is two-loop order [84, 125–134], although progress towards a full three-loop result has been reported [133, 135–141]. Interestingly, the three-loop soft anomalous dimension for massless particles has a remarkably simple analytic form. It was subsequently shown that it could be obtained without explicit calculation, by using a *bootstrap approach*, in which one expands it in a basis of known functions, before applying known constraints from collinear and high energy limits to fix the coefficients [142]. This strongly suggests that further techniques from formal amplitudes research—Involving both the theory of special functions, and insights into which functions can appear at which loop orders—may prove to be highly useful in extending our ability to resum perturbation theory. To illustrate the importance of this goal, figure 4 shows the measured distribution of the transverse momentum (relative to the incoming beams) of the virtual particle produced in Drell–Yan production, and which decays to a pair of leptons. At LO in perturbation theory, the Feynman diagram of figure 3(a) tells us that the transverse momentum must be exactly zero, as there is nothing for the virtual photon to recoil against. At NLO, it may recoil against an emitted gluon, but the collinear singularity of this emission means that the distribution would diverge as the transverse momentum goes to zero, due to a large logarithm involving the transverse momentum. Only by resumming large logs to all orders does the theory prediction match the data. This is one of probably hundreds of individual observables at the LHC where resummation is important. Also, desired improvements in the order of the logarithms summed in specific processes form part of experimental wishlists [68].



**Figure 4.** The distribution of the transverse momentum of a lepton pair in Drell–Yan production. The curve labelled with ‘NNLO +  $N^3LL$ ’ includes threshold logarithms resummed to all orders in perturbation theory.

As well as the order of the logs, another potential improvement is to include the terms that we have so far neglected in equation (16). In particular, while a lot is known about the LP terms, much less is known about the NLP terms. If the LP terms are dictated by the emission of soft and collinear radiation, the NLP terms are in principle described by *next-to-soft* (or *next-to-collinear*) radiation. Coincidentally, this is also a topic that has been widely studied on `hep-th` in recent years, due in particular to the fact that next-to-soft properties of amplitudes have been linked to symmetries at asymptotic infinity [143–145], thus starting an ongoing programme known as *celestial holography*, discussed in chapter 11 of this review [146]. As for the use of Wilson lines discussed above, practitioners in this field may be entirely unaware that the need to understand next-to-soft physics has a highly practical application at the LHC. The need to describe and potentially resum the NLP terms in equation (16) has been emphasised in e.g. references [147–151]. Indeed, the study of such corrections has a remarkably long history, beginning with the classic work of references [152, 153] for massive external particles, which was extended to massless particles in [154]. In QCD, next-to-soft effects have been investigated using a variety of techniques, including diagrammatic approaches [155–165] that aim to extend the factorisation formula of equation (18) to NLP order and/or resum NLP logs, insights from fixed-order perturbation theory [166–177], and effective field theory [178–201]. We are just starting to learn how to resum NLP effects for certain specific processes, and the coming years are likely to see a rapid increasing of our understanding of their importance. Furthermore, extensions to the factorisation formula of equation (18) involve new universal objects in field theory, that may be of interest beyond QCD, for those wishing to classify new structures in e.g.  $\mathcal{N} = 4$  SYM theory.

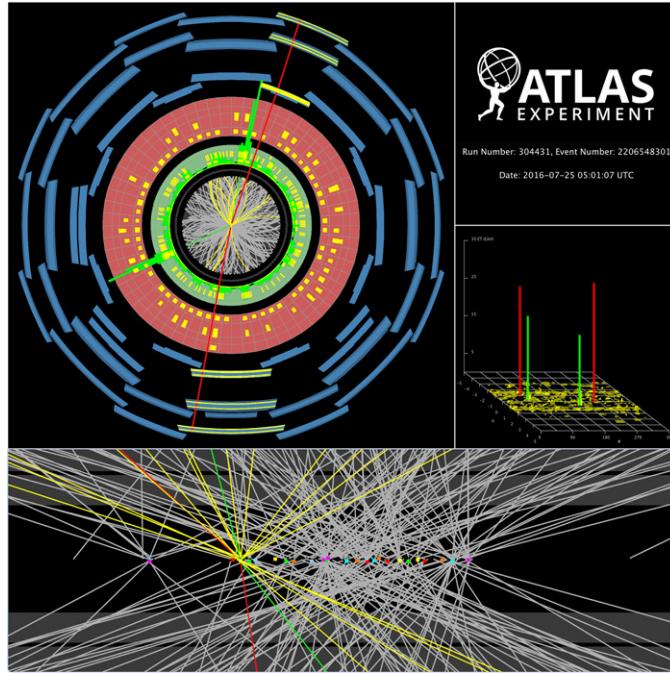
We have here seen some basic ideas from threshold resummation, which involves the inclusion of arbitrary amounts of soft/collinear radiation. However, the idea of summing up corrections to all orders in perturbation theory clearly generalises. Another kinematic limit that has been widely studied is the *Regge limit*, in which the centre of mass energy is much larger than the momentum transfer. This limit becomes experimentally more relevant as the energy of particle accelerators increases, and there is significant evidence that the inclusion of

enhanced effects in this limit is needed to better describe scattering data (see e.g. references [202–212]). This makes direct contact with much earlier work on S-matrix theory (see the classic texts of references [213, 214], and reference [215] for a more modern review), ideas from which (e.g. the bootstrap) seem to be coming back into fashion in contemporary **hep-th** physics. Of course, with any resummation, only a subset of the full information at each order in perturbation theory is included at higher orders. Resummation thus proceeds in tandem with fixed-order perturbation theory to provide a two-pronged attack on QFT: given any observable we want to calculate, we must include as much information as possible. Low-order information in perturbation theory can be calculated exactly, and then supplemented with higher-order results from resummation, being careful that no contributions have been counted twice. There are then two frontiers in perturbation theory, namely the inclusion of subleading terms in both the coupling (NLO, NNLO, ...) and logarithmic (NLL, NNLL, ...) expansions. Each of these requires clever thinking and new techniques, and also a joined-up approach. For example, new methods for obtaining fixed-order results can often be recycled for calculating higher-order Wilson line correlators, which are needed for resummation. There is significant scope for more formal amplitudists to contribute to both areas.

## 7. From theory to experiment

In the previous two sections, we have sketched the status of modern-day efforts to calculate perturbative QFT observables for collider physics. Anyone working in this area knows how extraordinarily difficult it can be to squeeze new results out of a non-abelian gauge theory. However, as we already hinted at in section 3, the output of even the most intricate QFT calculation looks almost nothing like what comes out of a particle accelerator! To illustrate this, figure 5 shows an event that was measured by the ATLAS detector in 2016, and which was believed to be a Higgs boson decaying to two  $Z$  bosons, which themselves then decay to a pair of leptons each. The upper-left panel shows a cross-sectional slice through the cylindrical detector, where the red and green lines constitute the best guess for what the leptons did. However, there is an enormous number of additional particles. The yellow lines denote extra charged particles that accompanied the Higgs boson event. Many of these will be charged hadrons, which arose from additional quark and gluon radiation. The sheer number of these goes way beyond what we can reliably calculate in fixed-order perturbation theory. What's more, only a tiny fraction of these charged particles have been kept in the event display—those with sufficiently small momentum transverse to the beam direction are thrown away so that we can even see what is going on! The grey lines demonstrate an additional complication: the beams at the LHC do not consist of single protons, but bunches of many protons (with over 100 billion protons per bunch in fact). This means that a large number of collisions happen simultaneously, so that any event we want to look at is swamped not just by its own mess, but by that of many other independent collisions! This is called *pile-up*, and can be corrected for by carefully ascertaining that the various particle tracks originate from different scattering vertices, as is shown in the lower panel of the figure. However, it is clear that many theorist's idea of what 'comparing theory to data' means, is very far indeed from what actually happens in practice. Our aim here is to provide a brief review on how QFT can nevertheless be meaningfully applied to the analysis of scattering events, and to point out some of the open issues.

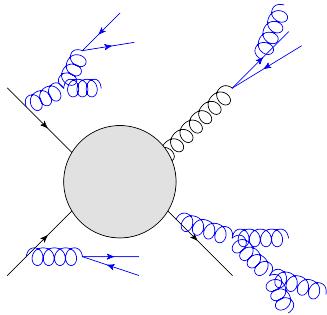
Theory calculations have two uses. Either we are calculating a signal (e.g. a new physics process), in which case we want to compare our prediction for a particular observable (e.g. a cross-section) with something measured. Or we might be considering a *background*, namely a standard model process that is in principle different to the signal, but which may contribute a proportion of events that happen to look similar. If the background process is much more



**Figure 5.** Event display from the ATLAS experiment, for a candidate event in which a Higgs boson decays to two  $Z$  bosons. The event was used in the analysis of reference [216]. Reproduced with permission from ATLAS Experiment © 2022 CERN.

probable than the signal, we will get lots of events that mimic the signal and thus act as ‘fake news’. Experimentalists apply very stringent statistical techniques to make sure that any signal they find has negligible probability of having been caused by a background. However, to do this, they need to know the background processes *very very precisely*, and this is then the job of theorists. But how do the latter turn their calculations into something approaching what is seen in the collider?

The first problem is to try to estimate the effect of large amounts of additional quark and gluon radiation, and there is a very well-established way to do this known as a *parton shower* (see e.g. references [87, 217] for reviews). It relies on the observation made above, that collinear radiation is enhanced, and also factorises off from an underlying hard scattering process in a universal (process-independent) manner. The radiation is in fact described by known *splitting functions* which, roughly speaking, give the probability that a parton splits into two other partons, each carrying a certain momentum fraction of the parent. The factorisation property means that different collinear splittings are uncorrelated and independent, and this allows for the construction of an algorithm for the generation of arbitrary amounts of collinear radiation, modelled as a Markov chain process. Essentially, one can generate a set of four-momenta for the additional particles, whose probability is given by a known distribution, which becomes exact in the limit in which all the radiation is strictly collinear. It includes both real and virtual QFT corrections, so that all probability weights are infrared finite, and applies this distribution even for radiation that is not collinear. That this turns out to be a reasonable approximation follows from the fact that collinear radiation is anyway enhanced. A schematic view of the action of a parton shower is shown in figure 6.



**Figure 6.** A *parton shower* algorithm dresses a given hard scattering process (whose particles are shown here in black) with additional quark and gluon radiation (shown in blue). The generated probability for such a process is approximate, but exact in the limit in which the emitted radiation is collinear with the outgoing hard particles. The approximation can be improved by including higher-order information in the hard scattering process. Reproduced with permission from ATLAS Experiment © 2022 CERN.

The energy of modern colliders such as the LHC is such that many particles are produced, which are widely separated in the detector and thus not necessarily collinear. Then the parton shower approximation is insufficient, and must be made more accurate. One way to do this is to include also full higher-order tree-level amplitudes, which can be done up to a few extra legs. One can then invent a suitable *matching prescription*, that guarantees that the most widely separated particles in any event are described by the tree-level amplitudes (where these are most accurate), but the remaining particles with less separation are described by the parton shower. Several prescriptions exist (see e.g. references [218, 219] for some of the first), and they must carefully make sure that no radiation is double-counted, by being included in both the higher-order tree-level matrix elements *and* the parton shower. From a QFT point of view, including only tree-level amplitudes is not ideal, given that the formal accuracy of the total cross-section then remains at LO only. Thus, it is desirable to start with full NLO amplitudes, and to match these with a parton shower. The matching procedure is now even more delicate than the tree-level case discussed above: the NLO matrix elements contain both real and virtual corrections, both of which potentially overlap with what the parton shower is doing. Two NLO matching schemes are in widespread use [220, 221], and their predictions for a given process are often compared as a means of estimating the reliability of the matching procedure. The state-of-the art for many LHC analyses is that either tree-level amplitudes matched to a parton shower are used, or the NLO approach (which only includes one additional emission in the amplitude). The choice of which approach to use is dictated by what is deemed to be more accurate for the observable of interest. However, one can clearly go further than this, by combining as many  $N^n$ LO matrix elements as possible (including higher-order tree-level matrix elements), and matching the whole lot to a parton shower. The subtleties in doing this are many, and the computational expense of doing this means that more understanding is needed of where such corrections are genuinely important. This is a topic that will evolve a great deal in the coming years [222].

After a parton shower has been applied, our QFT results start to look a lot more like those in figure 5. But they still contain free (anti-)quarks and gluons in the final state, rather than the colour-singlet hadrons that we now are observed in real experiments, due to the confinement property of QCD. For many observables this is not a problem. If all we want to do is to estimate a total cross-section, for example, every parton will end up in *some* hadron, so the observables

calculated assuming final-state partons, or final-state hadrons must be the same. But if we want to more realistically model scattering events for use in experimental analyses, we do indeed want to estimate how the process of *hadronisation* changes the final state. In practice, this is done using a variety of phenomenological models, containing free parameters that can be tuned to data. What helps is that the quantitative effect of hadronisation on a given observable is known to be suppressed by powers of the QCD confinement scale  $\Lambda$  i.e. it constitutes a power correction as appears in equation (9).

After parton showering and hadronisation, our scattering events are still not fully realistic: we have to include the fact that the incoming partons were only part of the incoming protons. The rest of the latter will somehow be distributed throughout a given scattering event, and their colour information may be non-trivially entangled with the rest of the event. Various models exist for describing this mess, which is loosely referred to as the *underlying event*. New ideas are always needed, as the uncertainties involved in such models can have dramatic effects. For example, the mass of the top quark is currently measured to within about 0.5 GeV, which is very similar to the estimated theory uncertainty due to underlying event effects (in particular how the colour of the top quark interacts with the colour of the beam remnants). This may not sound like a problem, but the top quark mass enters the expression for quantum corrections to the potential energy of the Higgs field. This can become negative for a certain range of top mass values, which results in the vacuum of our Universe becoming unstable. Given the measured Higgs mass value, the mass of the top quark we currently observe is perilously close to the unstable range. We cannot quite say how close, without knowing the top quark mass more precisely, and thus our ability to settle our collective fate rests on better modelling of non-perturbative QCD!

The above steps are computationally technical, and clearly very complicated for the uninitiated. However, various general purpose computer programs exist that allow users to simulate scattering events for a given process, including higher-order amplitudes, parton showers, hadronisation, underlying event modelling, and more besides. They are usually referred to as *Monte Carlo event generators*, and popular programs include [40–42, 223]. The output of such programs consists of a set of simulated scattering events, comprising a list of particles (e.g. leptons, photons, hadrons) and their four-momenta. One may then write additional code for analysing these events, e.g. to select those events that look interesting, and then plot distributions of various measurable quantities, to mimic what is done in an actual collider experiment.

Monte Carlo event generators are widely used both by experimentalists, but also phenomenologists who are trying to find new and better things to measure. They can also be interfaced with further simulations, that mimic the behaviour of the real ATLAS and CMS detectors, for an even more realistic characterisation of how actual scattering events are likely to behave. This then gives various levels at which theorists may compare their results with data, and examples include:

- *Parton level*: in this case, experimentalists will try to correct for non-perturbative effects, decays of particles etc, and present results for (differential) cross-sections containing final-state partons, vector bosons, top quarks etc. Theorists can then calculate simple final states involving quarks, gluons and other particles, using their favourite methods for calculating scattering amplitudes, for direct comparison with the ‘data’. This approach is going out of fashion for differential observables, due to the many assumptions that go in to *unfolding* the raw data back to parton level. If these assumptions turn out to be incorrect or superceded years after an experiment finishes, it can be impossible to replace them with a more correct analysis, especially if the raw data is no longer available.

- *Particle/hadron level*: here the theory calculation will include a parton shower, hadronisation, underlying event etc, as is obtained as the output of a Monte Carlo event generator. For some very inclusive observables (e.g. total cross-sections), this is not necessary, given that the additional steps needed to get to particle level will not change the total cross-section. However, the output of a particle-level calculation is a set of simulated scattering events that looks much closer to what happens in a real detector. In particular, this allows theorists to simulate the effect of proposed experimental analyses, including how ‘interesting’ scattering events will be selected, and what the distributions of various measured quantities will look like.
- *Detector level*: this is similar to particle-level, but includes an additional detector simulation, which would be important for theorists if they believed that realistic detector effects (e.g. finite energy/momentum resolution, gaps in where particles can be recorded) may be affecting their predictions.

For differential cross-sections, particle level has become a standard way of comparing theory with data. Even in that case, however, further steps are usually needed to make experimental events look similar to particle-level predictions. Both experimental scattering events and simulated particle-level ones will contain hundreds of particles, most of which will be hadrons, due to the large amounts of strongly-coupled quark and gluon radiation. In order to simplify each event, we can cluster these particles into *jets*, using a suitable *jet algorithm*. We can then select interesting events based on how many jets they have, and what the distribution of these jets looks like in the detector. A surprising amount of care is needed to make sure that a given definition of how to cluster particles into jets is well-defined in perturbation theory (see reference [224] for an excellent review).

The detailed comparison of theory with data requires a constant dialogue between theorists and experimentalists. At the heart of all event generators that are used in this process lie our beloved scattering amplitudes, and thus new techniques from formal theory can clearly contribute to the ongoing vast international efforts to understand what the Universe is trying to tell us in our detectors.

## 8. Summary

In the previous sections, we have seen a large number of steps in between the calculation of scattering amplitudes, and the direct comparison of theory with quantities measured by experimentalists. It is thus useful to have a quick summary of these ideas, with some further context of how they are applied in practice:

- Squared amplitudes must be renormalised to remove UV singularities, and converted into cross-sections by integrating over the phase space of any final-state particles, and dividing by the Lorentz-invariant flux factor.
- At hadron colliders, one must combine cross-sections for incoming *partons* [i.e. (anti-)quarks and gluons] with *parton distribution* functions for the incoming beam particles (e.g. protons).
- Theoretical quantities [e.g. (differential) cross-sections] should be such that they are *infrared-safe*: final states with different numbers of particles must be combined so that there are no IR singularities in the final results. One should then include as many perturbative orders in each amplitude as possible.
- If the coefficients of the perturbation expansion are unstable, it may be necessary to include additional logarithmically-enhanced terms to all orders in perturbation theory

(‘resummation’), and to avoid double-counting contributions which have already been included in the fixed-order expansion.

- (e) One can approximate the emission of additional QCD radiation by applying a *parton shower* algorithm to the theoretical prediction for a given final state. Care is needed if combining higher-order amplitudes with the parton shower, to avoid double-counting contributions.
- (f) After the parton shower, one may include further algorithms that simulate the combination of outgoing partons into colour-singlet hadrons, interactions with the beam remnants etc.

Depending on what data we are comparing to, it is not necessary to include all of these steps. To clarify this, let us consider one of the simplest types of observable that an experiment might measure, namely the cross-section for a given process. Experiments sometimes present a best estimate for the total cross-section, having attempted to reconstruct any intermediate unstable particles, and corrected for the finite detector volume etc. This is perhaps the easiest type of data for a theorist to compare to. If they wish, they can use fixed-order perturbation theory to obtain the total cross-section (including the parton distributions if it is a hadron collider) at some order, and then directly compare this number with the data. The theory result will have some uncertainty estimated by varying the renormalisation and factorisation scales. Likewise, the experimental result will also have an uncertainty, which may be split further into estimated *systematic* and *statistical errors*. The first of these is due to possible biases in the measurement, whereas the second measures the limited power of a finite data set. If the theory and data numbers agree within their respective uncertainties, then we would say that the theory matches the data. If not, we can statistically quantify the level of disagreement. This corresponds to the *parton level* comparison outlined above.

Even for total cross-sections, there are more sophisticated approaches. For example, we may wish to resum contributions from certain kinematic regions, the results of which are typically available in dedicated public codes for given processes. It is routine, for example, to include resummed contributions for various top quark and Higgs boson cross-sections. The effect of including these resummed contributions may be to change the central value of our theory prediction, or modify its uncertainty. However, it is still a single number which we can compare with the measured result. An even more complicated approach would be to dress our theoretical prediction with a parton shower algorithm, hadronisation etc. This simulates final states with large numbers of particles, and we can then apply similar selection criteria to those used in the experiment to our simulated events, in order to reproduce the steps they used in performing their own measurement. This is particularly useful if the experiment presents a measured cross-section that does not correspond to the total cross-section one would calculate in QFT, but has some exclusion criteria already applied (e.g. requirements that the particles be in certain regions of the detector only).

Other common measurements are differential cross-sections involving numbers of particles or jets, or their kinematic properties (e.g. angles, transverse momenta etc). For low numbers of particles, it may be possible to rely purely on fixed-order perturbation theory, provided there are enough particles in the final state of the relevant amplitudes to coincide with all the particles one wants for a given kinematic distribution. However, a much more realistic estimate will be obtained by using parton-shower programs to simulate realistic events containing many particles, and then clustering the particles into jets using the same algorithms that the experiment has used. In order to do this, a theorist would use one of the publicly available Monte Carlo event generators mentioned above.

Concrete examples of processes at the LHC for which the above remarks apply are perhaps too numerous to list explicitly. Furthermore, modern-day Monte Carlo event generators essentially allow the user to specify *any* desired process, and then to automatically generate simulated events which can be analysed to produce kinematic distributions for comparison with real data. A particularly convenient repository for experimental data in particle physics is the HEPData database [225], which also provides links to the relevant publications.

Finally, it is worth commenting on what theory input is missing for forthcoming data analyses and/or new collider experiments beyond the currently running LHC. Broadly speaking, the answer is that we must learn to compute ever-higher orders in perturbation theory, and to be able to match amplitudes calculated at these orders to state-of-the-art parton-shower algorithms, where the latter can also be improved. Given the large amount of effort involved, it is necessary to prioritise those processes that experiments are particularly focussed on. To this end, wishlists such as those in reference [68] are highly useful, as they distil the opinions of large numbers of experimentalists into a coherent request! Another area requiring improvement is the determination of the parton distribution functions, as the uncertainty on extracting these from data can limit the theoretical precision of a given collider prediction, if the relevant cross-section is known to a very high perturbative order. It is currently not definitively known what the next collider after the LHC will be. However, it is highly likely to be a lepton-based machine (e.g. colliding electrons and positrons). Many high-precision cross-sections are already known for such colliders, given that the calculations are actually somewhat simpler than those needed for hadron machines, due to the lack of coloured particles in the initial state. But it is always useful to have higher orders in perturbation theory, as for hadron colliders.

## 9. Conclusion

It is well-known that scattering amplitudes in perturbative QFT are relevant for collider experiments. But the precise way in which they are used, including the huge number of steps that are needed to meaningfully compare theory with data, are never dealt with in textbooks. It is perfectly possible nowadays for formal theorists to live out an entire career without comparing anything to data, despite a desire to do so. It is also very often true that the topics being discussed on **hep-th** [e.g. Wilson lines, (next-to)-soft radiation] are the *same* topics being discussed on **hep-ph**, albeit in completely different notation, and for entirely different reasons. Hence, there is significant scope for more formal theorists to contribute to the challenging and rewarding world of collider physics, should they wish to do so. Recent years have seen remarkable revolutions in how we think about scattering amplitudes in different theories, and how those different theories are related. My hope—which is also the very ethos of the SAGEX network—is that the coming years will see similar revolutions in how we apply new techniques to extract the most we can from current and forthcoming collider experiments.

## Acknowledgments

This work was supported by the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 764850 ‘SAGEX’. We also acknowledge support from the Science and Technology Facilities Council (STFC) Consolidated Grant ST/T000686/1 ‘*Amplitudes, strings & duality*’.

## Data availability statement

No new data were created or analysed in this study.

## ORCID iDs

Chris D White  <https://orcid.org/0000-0002-6752-2394>

## References

- [1] Aaboud M *et al* (ATLAS) 2020 *Eur. Phys. J. C* **80** 754
- [2] Anastasiou C, Duhr C, Dulat F, Herzog F and Mistlberger B 2015 *Phys. Rev. Lett.* **114** 212001
- [3] Anastasiou C, Duhr C, Dulat F, Furlan E, Gehrmann T, Herzog F, Lazopoulos A and Mistlberger B 2016 *J. High Energy Phys.* **JHEP05(2016)058**
- [4] Dreyer F A and Karlberg A 2016 *Phys. Rev. Lett.* **117** 072001
- [5] Dreyer F A and Karlberg A 2018 *Phys. Rev. D* **98** 114016
- [6] Duhr C, Dulat F and Mistlberger B 2020 *Phys. Rev. Lett.* **125** 172001
- [7] Gross D J and Wilczek F 1973 *Phys. Rev. Lett.* **30** 1343–6
- [8] Politzer H D 1973 *Phys. Rev. Lett.* **30** 1346–9
- [9] Feynman R P 1969 *Conf. Proc. C* **690905** 237–58
- [10] Feynman R P 1969 *Phys. Rev. Lett.* **23** 1415–7
- [11] Collins J C 1986 *Renormalization: An Introduction to Renormalization, The Renormalization Group, and The Operator Product Expansion* (Cambridge Monographs on Mathematical Physics vol 26) (Cambridge: Cambridge University Press)
- [12] Bloch F and Nordsieck A 1937 *Phys. Rev.* **52** 54–9
- [13] Kinoshita T 1962 *J. Math. Phys.* **3** 650–77
- [14] Lee T D and Nauenberg M 1964 *Phys. Rev.* **133** B1549–62
- [15] Bailey S, Cridge T, Harland-Lang L A, Martin A D and Thorne R S 2021 *Eur. Phys. J. C* **81** 341
- [16] Abdul Khalek R *et al* (NNPDF) 2019 *Eur. Phys. J. C* **79** 931
- [17] Alekhin S, Blümlein J and Moch S 2018 *Eur. Phys. J. C* **78** 477
- [18] Hou T J *et al* 2021 *Phys. Rev. D* **103** 014013
- [19] Passarino G and Veltman M J G 1979 *Nucl. Phys. B* **160** 151–207
- [20] Ellis R K and Zanderighi G 2008 *J. High Energy Phys.* **JHEP02(2008)002**
- [21] Bern Z, Dixon L J, Dunbar D C and Kosower D A 1994 *Nucl. Phys. B* **425** 217–60
- [22] Bern Z, Dixon L J, Dunbar D C and Kosower D A 1995 *Nucl. Phys. B* **435** 59–101
- [23] Bern Z and Morgan A G 1996 *Nucl. Phys. B* **467** 479–509
- [24] Bern Z, Dixon L J and Kosower D A 1998 *Nucl. Phys. B* **513** 3–86
- [25] Britto R, Cachazo F and Feng B 2005 *Nucl. Phys. B* **725** 275–305
- [26] Brandhuber A, McNamara S, Spence B J and Travaglini G 2005 *J. High Energy Phys.* **JHEP10(2005)011**
- [27] Britto R, Cachazo F and Feng B 2005 *Nucl. Phys. B* **715** 499–522
- [28] Bern Z, Dixon L J and Kosower D A 2005 *Phys. Rev. D* **71** 105013
- [29] Bern Z, Dixon L J and Kosower D A 2005 *Phys. Rev. D* **72** 125003
- [30] Badger S D, Glover E W N, Khoze V V and Svrcek P 2005 *J. High Energy Phys.* **JHEP07(2005)025**
- [31] Badger S D, Glover E W N and Khoze V V 2006 *J. High Energy Phys.* **JHEP01(2006)066**
- [32] Anastasiou C, Britto R, Feng B, Kunszt Z and Mastrolia P 2007 *Phys. Lett. B* **645** 213–6
- [33] Anastasiou C, Britto R, Feng B, Kunszt Z and Mastrolia P 2007 *J. High Energy Phys.* **JHEP03(2007)111**
- [34] Giele W T, Kunszt Z and Melnikov K 2008 *J. High Energy Phys.* **JHEP04(2008)049**
- [35] Huang Y t and McGady D 2014 *Phys. Rev. Lett.* **112** 241601
- [36] Ellis R K, Kunszt Z, Melnikov K and Zanderighi G 2012 *Phys. Rep.* **518** 141–250
- [37] Gleisberg T and Krauss F 2008 *Eur. Phys. J. C* **53** 501–23

[38] Hirschi V, Frederix R, Frixione S, Garzelli M V, Maltoni F and Pittau R 2011 *J. High Energy Phys.* **JHEP05(2011)044**

[39] Bellm J *et al* 2016 *Eur. Phys. J. C* **76** 196

[40] Bothmann E *et al* (Sherpa) 2019 *SciPost Phys.* **7** 034

[41] Alwall J *et al* 2014 *J. High Energy Phys.* **JHEP07(2014)079**

[42] Bellm J *et al* 2020 *Eur. Phys. J. C* **80** 452

[43] Abreu S, Britto R and Duhr C 2022 arXiv:2203.13014

[44] Blümlein J and Schneider C 2022 arXiv:2203.13015

[45] Papathanasiou G 2022 arXiv:2203.13016

[46] Travaglini G *et al* 2022 arXiv:2203.13011

[47] Frixione S, Kunszt Z and Signer A 1996 *Nucl. Phys. B* **467** 399–442

[48] Catani S and Seymour M H 1997 *Nucl. Phys. B* **485** 291–419  
Catani S and Seymour M H 1998 *Nucl. Phys. B* **510** 503–4 (erratum)

[49] Gehrmann-De Ridder A, Gehrmann T and Glover E W N 2005 *J. High Energy Phys.* **JHEP09(2005)056**

[50] Currie J, Glover E W N and Wells S 2013 *J. High Energy Phys.* **JHEP04(2013)066**

[51] Czakon M 2010 *Phys. Lett. B* **693** 259–68

[52] Czakon M 2011 *Nucl. Phys. B* **849** 250–95

[53] Boughezal R, Melnikov K and Petriello F 2012 *Phys. Rev. D* **85** 034025

[54] Anastasiou C, Melnikov K and Petriello F 2004 *Phys. Rev. D* **69** 076010

[55] Binoth T and Heinrich G 2004 *Nucl. Phys. B* **693** 134–48

[56] Catani S and Grazzini M 2007 *Phys. Rev. Lett.* **98** 222002

[57] Boughezal R, Liu X and Petriello F 2015 *Phys. Rev. D* **91** 094035

[58] Gaunt J, Stahlhofen M, Tackmann F J and Walsh J R 2015 *J. High Energy Phys.* **JHEP09(2015)058**

[59] Del Duca V, Duhr C, Somogyi G, Tramontano F and Trócsányi Z 2015 *J. High Energy Phys.* **JHEP04(2015)036**

[60] Caola F, Melnikov K and Röntsch R 2017 *Eur. Phys. J. C* **77** 248

[61] Delto M and Melnikov K 2019 *J. High Energy Phys.* **JHEP05(2019)148**

[62] Caola F, Melnikov K and Röntsch R 2019 *Eur. Phys. J. C* **79** 386

[63] Magnea L, Maina E, Pelliccioli G, Signorile-Signorile C, Torrielli P and Uccirati S 2018 *J. High Energy Phys.* **JHEP12(2018)107**  
Magnea L, Maina E, Pelliccioli G, Signorile-Signorile C, Torrielli P and Uccirati S 2019 *J. High Energy Phys.* **JHEP06(2019)013** (erratum)

[64] Magnea L, Maina E, Pelliccioli G, Signorile-Signorile C, Torrielli P and Uccirati S 2018 *J. High Energy Phys.* **JHEP12(2018)062**

[65] Caola F, Melnikov K and Röntsch R 2019 *Eur. Phys. J. C* **79** 1013

[66] Asteriadis K, Caola F, Melnikov K and Röntsch R 2020 *Eur. Phys. J. C* **80** 8

[67] Cacciari M, Dreyer F A, Karlberg A, Salam G P and Zanderighi G 2015 *Phys. Rev. Lett.* **115** 082002  
Cacciari M, Dreyer F A, Karlberg A, Salam G P and Zanderighi G 2018 *Phys. Rev. Lett.* **120** 139901 (erratum)

[68] Huss A, Huston J, Jones S and Pellen M 2022 arXiv:2207.02122

[69] Denner A and Dittmaier S 2020 *Phys. Rep.* **864** 1–163

[70] Agarwal N, Magnea L, Signorile-Signorile C and Tripathi A 2021 arXiv:2112.07099

[71] Parisi G 1980 *Phys. Lett. B* **90** 295–6

[72] Curci G and Greco M 1980 *Phys. Lett. B* **92** 175–8

[73] Sterman G F 1987 *Nucl. Phys. B* **281** 310–64

[74] Catani S and Trentadue L 1989 *Nucl. Phys. B* **327** 323–52

[75] Gatheral J G M 1983 *Phys. Lett. B* **133** 90–4

[76] Frenkel J and Taylor J C 1984 *Nucl. Phys. B* **246** 231–45

[77] Sterman G F 1981 *AIP Conf. Proc.* **74** 22–40

[78] Korchemsky G P and Marchesini G 1993 *Nucl. Phys. B* **406** 225–58

[79] Korchemsky G P and Marchesini G 1993 *Phys. Lett. B* **313** 433–40

[80] Forte S and Ridolfi G 2003 *Nucl. Phys. B* **650** 229–70

[81] Becher T and Neubert M 2006 *Phys. Rev. Lett.* **97** 082001

[82] Schwartz M D 2008 *Phys. Rev. D* **77** 014026

[83] Bauer C W, Fleming S P, Lee C and Sterman G F 2008 *Phys. Rev. D* **78** 034027

[84] Chiu J y, Fuhrer A, Kelley R and Manohar A V 2009 *Phys. Rev. D* **80** 094013

[85] Luisoni G and Marzani S 2015 *J. Phys. G: Nucl. Part. Phys.* **42** 103101

[86] Becher T, Broggio A and Ferroglio A 2015 *Introduction to Soft-Collinear Effective Theory* vol 896 (Berlin: Springer) (arXiv:1410.1892)

[87] Campbell J, Huston J and Krauss F 2017 *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era* (Oxford: Oxford University Press)

[88] Dixon L J, Magnea L and Sterman G F 2008 *J. High Energy Phys.* **JHEP08(2008)022**

[89] Weinberg S 1965 *Phys. Rev.* **140** B516–24

[90] Alday L F and Maldacena J M 2007 *J. High Energy Phys.* **JHEP06(2007)064**

[91] Polyakov A M 1980 *Nucl. Phys. B* **164** 171–88

[92] Arefeva I Y 1980 *Phys. Lett. B* **93** 347–53

[93] Dotsenko V S and Vergeles S N 1980 *Nucl. Phys. B* **169** 527–46

[94] Brandt R A, Neri F and Sato M a 1981 *Phys. Rev. D* **24** 879

[95] Korchemsky G P and Radyushkin A V 1986 *Phys. Lett. B* **171** 459–67

[96] Korchemsky G P and Radyushkin A V 1986 *Sov. J. Nucl. Phys.* **44** 877

[97] Korchemsky G P and Radyushkin A V 1987 *Nucl. Phys. B* **283** 342–64

[98] Grozin A, Henn J M, Korchemsky G P and Marquard P 2016 *J. High Energy Phys.* **JHEP01(2016)140**

[99] Grozin A, Henn J M, Korchemsky G P and Marquard P 2015 *Phys. Rev. Lett.* **114** 062006

[100] Brüser R, Dlapa C, Henn J M and Yan K 2021 *Phys. Rev. Lett.* **126** 021601

[101] Korchemsky G P 1994 *Phys. Lett. B* **325** 459–66

[102] Korchemskaya I A and Korchemsky G P 1996 *Phys. Lett. B* **387** 346–54

[103] Korchemskaya I A and Korchemsky G P 1995 *Nucl. Phys. B* **437** 127–62

[104] Catani S 1998 *Phys. Lett. B* **427** 161–71

[105] Sterman G F and Tejeda-Yeomans M E 2003 *Phys. Lett. B* **552** 48–56

[106] Kidonakis N, Oderda G and Sterman G F 1998 *Nucl. Phys. B* **531** 365–402

[107] Bonciani R, Catani S, Mangano M L and Nason P 2003 *Phys. Lett. B* **575** 268–78

[108] Dokshitzer Y L and Marchesini G 2006 *J. High Energy Phys.* **JHEP01(2006)007**

[109] Aybat S M, Dixon L J and Sterman G F 2006 *Phys. Rev. D* **74** 074004

[110] Gardi E and Magnea L 2009 *J. High Energy Phys.* **JHEP03(2009)079**

[111] Becher T and Neubert M 2009 *Phys. Rev. Lett.* **102** 162001  
Becher T and Neubert M 2013 *Phys. Rev. Lett.* **111** 199905 (erratum)

[112] Becher T and Neubert M 2009 *J. High Energy Phys.* **JHEP06(2009)081**  
Becher T and Neubert M 2013 *J. High Energy Phys.* **JHEP11(2013)024** (erratum)

[113] Gardi E and Magnea L 2009 *Nuovo Cimento C* **32N5-6** 137–57

[114] Dixon L J 2009 *Phys. Rev. D* **79** 091501

[115] Dixon L J, Gardi E and Magnea L 2010 *J. High Energy Phys.* **JHEP02(2010)081**

[116] Del Duca V, Duhr C, Gardi E, Magnea L and White C D 2011 *J. High Energy Phys.* **JHEP12(2011)021**

[117] Caron-Huot S 2015 *J. High Energy Phys.* **JHEP05(2015)093**

[118] Ahrens V, Neubert M and Vernazza L 2012 *J. High Energy Phys.* **JHEP09(2012)138**

[119] Naclilch S G, Nastase H and Schnitzer H J 2013 *J. High Energy Phys.* **JHEP04(2013)114**

[120] Erdogan O and Sterman G 2015 *Phys. Rev. D* **91** 065033

[121] Gehrmann T, Glover E W N, Huber T, Ikizlerli N and Studerus C 2010 *J. High Energy Phys.* **JHEP06(2010)094**

[122] Agarwal B, von Manteuffel A, Panzer E and Schabinger R M 2021 *Phys. Lett. B* **820** 136503

[123] Almelid O, Duhr C and Gardi E 2016 *Phys. Rev. Lett.* **117** 172002

[124] Gardi E, Almelid O and Duhr C 2016 *PoS* **LL2016** 058

[125] Kidonakis N 2009 *Phys. Rev. Lett.* **102** 232003

[126] Mitov A, Sterman G F and Sung I 2009 *Phys. Rev. D* **79** 094015

[127] Becher T and Neubert M 2009 *Phys. Rev. D* **79** 125004  
Becher T and Neubert M 2009 *Phys. Rev. D* **80** 109901 (erratum)

[128] Beneke M, Falgari P and Schwinn C 2010 *Nucl. Phys. B* **828** 69–101

[129] Czakon M, Mitov A and Sterman G F 2009 *Phys. Rev. D* **80** 074017

[130] Ferroglio A, Neubert M, Pecjak B D and Yang L L 2009 *J. High Energy Phys.* **JHEP11(2009)062**

[131] Mitov A, Sterman G F and Sung I 2010 *Phys. Rev. D* **82** 034020

[132] Gardi E 2014 *J. High Energy Phys.* **JHEP04(2014)044**

[133] Falcioni G, Gardi E, Harley M, Magnea L and White C D 2014 *J. High Energy Phys.* **JHEP10(2014)010**

[134] Henn J M, Smirnov A V and Smirnov V A 2013 *J. High Energy Phys.* **JHEP07(2013)128**

[135] Gardi E, Laenen E, Stavenga G and White C D 2010 *J. High Energy Phys.* **JHEP11(2010)155**

[136] Gardi E and White C D 2011 *J. High Energy Phys.* **JHEP03(2011)079**

[137] Dukes M, Gardi E, Steinbergsson E and White C D 2013 *J. Comb. Theory A* **120** 1012–37

[138] Gardi E, Smillie J M and White C D 2013 *J. High Energy Phys.* **JHEP06(2013)088**

[139] Gardi E, Harley M, Lodin R, Palusa M, Smillie J M, White C D and Yeomans S 2021 arXiv:[2110.01685](https://arxiv.org/abs/2110.01685)

[140] Mitov A, Sterman G and Sung I 2010 *Phys. Rev. D* **82** 096010

[141] Vladimirov A A 2015 *J. High Energy Phys.* **JHEP06(2015)120**

[142] Almelid O, Duhr C, Gardi E, McLeod A and White C D 2017 *J. High Energy Phys.* **JHEP09(2017)073**

[143] Cachazo F and Strominger A 2014 arXiv:[1404.4091](https://arxiv.org/abs/1404.4091)

[144] Casali E 2014 *J. High Energy Phys.* **JHEP08(2014)077**

[145] Strominger A 2017 arXiv:[1703.05448](https://arxiv.org/abs/1703.05448)

[146] McLoughlin T, Puhm A and Raclariu A M 2022 arXiv:[2203.13022](https://arxiv.org/abs/2203.13022)

[147] Kramer M, Laenen E and Spira M 1998 *Nucl. Phys. B* **511** 523–49

[148] Ball R D, Bonvini M, Forte S, Marzani S and Ridolfi G 2013 *Nucl. Phys. B* **874** 746–72

[149] Bonvini M, Forte S, Ridolfi G and Rottoli L 2015 *J. High Energy Phys.* **JHEP01(2015)046**

[150] van Beekveld M, Beenakker W, Basu R, Laenen E, Misra A and Motylinski P 2019 *Phys. Rev. D* **100** 056009

[151] van Beekveld M, Laenen E, Sinnenhe Damsté J and Vernazza L 2021 *J. High Energy Phys.* **JHEP05(2021)114**

[152] Low F E 1958 *Phys. Rev.* **110** 974–7

[153] Burnett T H and Kroll N M 1968 *Phys. Rev. Lett.* **20** 86

[154] Del Duca V 1990 *Nucl. Phys. B* **345** 369–88

[155] Laenen E, Stavenga G and White C D 2009 *J. High Energy Phys.* **JHEP03(2009)054**

[156] Laenen E, Magnea L, Stavenga G and White C D 2011 *J. High Energy Phys.* **JHEP01(2011)141**

[157] Bonocore D, Laenen E, Magnea L, Melville S, Vernazza L and White C D 2015 *J. High Energy Phys.* **JHEP06(2015)008**

[158] Bonocore D, Laenen E, Magnea L, Vernazza L and White C D 2016 *J. High Energy Phys.* **JHEP12(2016)121**

[159] Bonocore D 2021 *J. High Energy Phys.* **JHEP02(2021)007**

[160] Gervais H 2017 *Phys. Rev. D* **95** 125009

[161] Gervais H 2017 *Phys. Rev. D* **96** 065007

[162] Gervais H 2017 Soft radiation theorems at all loop order in quantum field theory *PhD Thesis* State University of New York at Stony Brook

[163] Laenen E, Sinnenhe Damsté J, Vernazza L, Waalewijn W and Zoppi L 2021 *Phys. Rev. D* **103** 034022

[164] Bahjat-Abbas N, Bonocore D, Sinnenhe Damsté J, Laenen E, Magnea L, Vernazza L and White C D 2019 *J. High Energy Phys.* **JHEP11(2019)002**

[165] van Beekveld M, Vernazza L and White C D 2021 arXiv:[2109.09752](https://arxiv.org/abs/2109.09752)

[166] Soar G, Moch S, Vermaseren J A M and Vogt A 2010 *Nucl. Phys. B* **832** 152–227

[167] Moch S and Vogt A 2009 *J. High Energy Phys.* **JHEP11(2009)099**

[168] Moch S and Vogt A 2009 *J. High Energy Phys.* **JHEP04(2009)081**

[169] de Florian D, Mazzitelli J, Moch S and Vogt A 2014 *J. High Energy Phys.* **JHEP10(2014)176**

[170] Lo Presti N A, Almasy A A and Vogt A 2014 *Phys. Lett. B* **737** 120–3

[171] Del Duca V, Laenen E, Magnea L, Vernazza L and White C D 2017 *J. High Energy Phys.* **JHEP11(2017)057**

[172] van Beekveld M, Beenakker W, Laenen E and White C D 2020 *J. High Energy Phys.* **JHEP03(2020)106**

[173] Bonocore D, Laenen E, Magnea L, Vernazza L and White C D 2015 *Phys. Lett. B* **742** 375–82

[174] Bahjat-Abbas N, Sinnenhe Damsté J, Vernazza L and White C D 2018 *J. High Energy Phys.* **JHEP10(2018)144**

[175] Ebert M A, Moult I, Stewart I W, Tackmann F J, Vita G and Zhu H X 2018 *J. High Energy Phys.* **JHEP12(2018)084**

[176] Boughezal R, Isgrò A and Petriello F 2018 *Phys. Rev. D* **97** 076006

[177] Boughezal R, Isgrò A and Petriello F 2020 *Phys. Rev. D* **101** 016005

[178] Kolodrubetz D W, Moult I and Stewart I W 2016 *J. High Energy Phys.* **JHEP05**(2016)139

[179] Moult I, Rothen L, Stewart I W, Tackmann F J and Zhu H X 2017 *Phys. Rev. D* **95** 074023

[180] Feige I, Kolodrubetz D W, Moult I and Stewart I W 2017 *J. High Energy Phys.* **JHEP11**(2017)142

[181] Beneke M, Garny M, Szafron R and Wang J 2018 *J. High Energy Phys.* **JHEP03**(2018)001

[182] Beneke M, Garny M, Szafron R and Wang J 2018 *J. High Energy Phys.* **JHEP11**(2018)112

[183] Bhattacharya A, Moult I, Stewart I W and Vita G 2019 *J. High Energy Phys.* **JHEP05**(2019)192

[184] Beneke M, Garny M, Szafron R and Wang J 2019 *J. High Energy Phys.* **JHEP09**(2019)101

[185] Bodwin G T, Ee J H, Lee J and Wang X P 2021 arXiv:[2107.07941](https://arxiv.org/abs/2107.07941)

[186] Moult I, Stewart I W and Vita G 2019 *J. High Energy Phys.* **JHEP11**(2019)153

[187] Beneke M, Broggio A, Jaskiewicz S and Vernazza L 2020 *J. High Energy Phys.* **JHEP07**(2020)078

[188] Liu Z L and Neubert M 2020 *J. High Energy Phys.* **JHEP04**(2020)033

[189] Liu Z L, Mecaj B, Neubert M and Wang X 2021 *Phys. Rev. D* **104** 014004

[190] Boughezal R, Liu X and Petriello F 2017 *J. High Energy Phys.* **JHEP03**(2017)160

[191] Moult I, Stewart I W and Vita G 2017 *J. High Energy Phys.* **JHEP07**(2017)067

[192] Chang C H, Stewart I W and Vita G 2018 *J. High Energy Phys.* **JHEP04**(2018)041

[193] Moult I, Stewart I W, Vita G and Zhu H X 2018 *J. High Energy Phys.* **JHEP08**(2018)013

[194] Beneke M, Broggio A, Garny M, Jaskiewicz S, Szafron R, Vernazza L and Wang J 2019 *J. High Energy Phys.* **JHEP03**(2019)043

[195] Ebert M A, Moult I, Stewart I W, Tackmann F J, Vita G and Zhu H X 2019 *J. High Energy Phys.* **JHEP04**(2019)123

[196] Beneke M, Garny M, Jaskiewicz S, Szafron R, Vernazza L and Wang J 2020 *J. High Energy Phys.* **JHEP01**(2020)094

[197] Moult I, Stewart I W, Vita G and Zhu H X 2020 *J. High Energy Phys.* **JHEP05**(2020)089

[198] Liu Z L and Neubert M 2020 *J. High Energy Phys.* **JHEP06**(2020)060

[199] Liu Z L, Mecaj B, Neubert M, Wang X and Fleming S 2020 *J. High Energy Phys.* **JHEP07**(2020)104

[200] Wang J 2019 arXiv:[1912.09920](https://arxiv.org/abs/1912.09920)

[201] Beneke M, Garny M, Jaskiewicz S, Szafron R, Vernazza L and Wang J 2020 *J. High Energy Phys.* **JHEP10**(2020)196

[202] White C D and Thorne R S 2007 *Phys. Rev. D* **75** 034005

[203] Ball R D, Bertone V, Bonvini M, Marzani S, Rojo J and Rottoli L 2018 *Eur. Phys. J. C* **78** 321

[204] Andersen J R, Black J A, Brooks H M, Byrne E P, Maier A and Smillie J M 2021 *J. High Energy Phys.* **JHEP04**(2021)105

[205] Andersen J R, Hapola T, Heil M, Maier A and Smillie J 2019 arXiv:[1902.08430](https://arxiv.org/abs/1902.08430)

[206] Andersen J R, Cockburn J D, Heil M, Maier A and Smillie J M 2019 *J. High Energy Phys.* **JHEP04**(2019)127

[207] Andersen J R, Hapola T, Heil M, Maier A and Smillie J M 2018 *J. High Energy Phys.* **JHEP08**(2018)090

[208] Andersen J R, Hapola T, Maier A and Smillie J M 2017 *J. High Energy Phys.* **JHEP09**(2017)065

[209] Andersen J R, Medley J J and Smillie J M 2016 *J. High Energy Phys.* **JHEP05**(2016)136

[210] Andersen J R, Hapola T and Smillie J M 2012 *J. High Energy Phys.* **JHEP09**(2012)047

[211] Andersen J R and Smillie J M 2011 *J. High Energy Phys.* **JHEP06**(2011)010

[212] Andersen J R and Smillie J M 2010 *J. High Energy Phys.* **JHEP01**(2010)039

[213] Eden R J, Landshoff P V, Olive D I and Polkinghorne J C 1966 *The Analytic S-Matrix* (Cambridge: Cambridge University Press)

[214] Collins P D B 2009 *An Introduction to Regge Theory and High-Energy Physics (Cambridge Monographs on Mathematical Physics)* (Cambridge: Cambridge University Press)

[215] White C D 2020 *SciPost Phys. Lect. Notes* **13** 1

[216] Aaboud M *et al.* (ATLAS) 2017 *J. High Energy Phys.* **JHEP10**(2017)132

[217] Ellis R K, Stirling W J and Webber B R 2011 *QCD and Collider Physics* vol 8 (Cambridge: Cambridge University Press)

[218] Caravaglios F, Mangano M L, Moretti M and Pittau R 1999 *Nucl. Phys. B* **539** 215–32

[219] Catani S, Krauss F, Kuhn R and Webber B R 2001 *J. High Energy Phys.* **JHEP11**(2001)063

- [220] Frixione S and Webber B R 2002 *J. High Energy Phys.* **JHEP06(2002)029**
- [221] Frixione S, Nason P and Webber B R 2003 *J. High Energy Phys.* **JHEP08(2003)007**
- [222] Re E 2021 Recent progress on high order calculations and matching to parton showers *9th Large Hadron Collider Physics Conf.*
- [223] Sjöstrand T 2020 *Comput. Phys. Commun.* **246** 106910
- [224] Salam G P 2010 *Eur. Phys. J. C* **67** 637–86
- [225] HEPData: repository for publication-related high-energy physics data <https://hepdata.net> (accessed 27 June 2022)