UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS FÍSICAS
Instituo de Física
Física de Altas Energías, Física Nuclear,
Gravitación y Física Matemática

TÍTULO DEL TRABAJO

"ttWW Production at NLO QCD plus Parton Showers in the POWHEG-BOX"

TESIS
QUE PARA OPTAR POR EL GRADO DE:

Maestría en Ciencias (Física)

PRESENTA:

LUZ SILVANA RODRÍGUEZ MELÉNDEZ

TUTOR/A O TUTORES PRINCIPALES
Dr. Manfred Kraus, IF
MIEMBROS DEL COMITÉ TUTOR
Dra. Myriam Mondragón Ceballos, IF
Dr. Alexis Armando Aguilar Arévalo, ICN

Ciudad de México, México, abril del año 2026

«It's the questions we can't answer that teach us the most. They teach us how to think. If you give a man an answer, all he gains is a little fact. But give him a question and he'll look for his own answers.»

Patrick Rothfuss

Acknowledgments

A mi familia... mi refugio. A mis padres, Yolanda Meléndez Peñaflor y César Rodríguez Gutiérrez, que formaron un hogar donde priorizaron el respeto e inculcaron la importancia del pensamiento crítico y constructivo, pero también del amor, el cariño y la unión; a mis hermanas, Aranza Irán y Aline Valeria, con quienes he forjado una alianza inquebrantable cuando se trata de luchar por lo que realmente vale la pena en esta vida,a quienes les auguro un futuro brillante. A mi tía, Ángeles Meléndez, quien nos procuró un cariño grande y sincero, y un apoyo incondicional a nuestros sueños. A este gran equipo, firme y fuerte en las grandes batallas.

A cada uno de mis colegas, compañeros, maestros y amigos.

Agradezco y reconozco a Emmanuel Velázquez Hermenegildo, mi primer colega en esta profesión, al amigo que se convierte en familia. Te agradezco los años de trabajo y aprendizaje; las horas de estudio y de colaboración, pero también el espacio para la discusión y construcción ideológica de otra índole, y a quien siempre agradeceré el hecho de estar y de ser.

A mi asesor, el Dr. Manfred Kraus, uno de los científicos más jóvenes y talentosos que conozco, que me brindó todo el soporte para iniciar mi profesionalización hacia el hermoso camino de la física de altas energías y la cromodiámica cuántica, y de quien he tenido la dicha y el honor de ser estudiante.

A Juan Felipe Jiménez Román, quien se ha convertido en una de las personas más importantes en este camino, mi compañero entrañable de grandes vivencias; En ti, de tu ejemplo, de tu capacidad de cuestionar y de la nobleza de tus acciones he encontrado una forma más genuina y espontánea de vivir la vida. A Javier Enrique Molina Ariza, mi admiración por tu gran potencial académico y profesional, pero también por tu gran calidad humana y de pensamiento. Y mi gratitud por los lazos de colaboración que formamos, pero también por la candidez y el amparo que me has brindado.

A Luis Eduardo y a Abner Alejandro, mi primer equipo, mis primeros amigos en esta maravillosa institución, y que hoy también recorren sus caminos de forma excepcional. A Gustavo Becerril, por convertirte poco a poco en un amigo entrañable y llegar a ser uno de los miembros más importantes de esta gran familia. Valoro enormemente tu amistad. A Emiliano Garrido, a Ricardo Orozco, a Antonio Samaniego, grandes personas a quienes tengo la dicha de llamar amigos.

A mi querida Universidad Nacional Autónoma de México, por adoptarme, hacerme uno más de tus jóvenes orgullosos. Te prometo reconocer en mis logros, tu nombre.
Y a mi patria querida, México, a su lucha histórica por la dignidad y la soberanía de sus pueblos. Me enorgullezco de ser hija de tu educación pública; de ser heredera de la escuela libre, laica y gratuita. En mis oportunidades de hoy, reconozco el legado y el esfuerzo de quienes lucharon por un país más justo ayer (y de quienes lo siguen haciendo). A ti, querido pueblo, te prometo corresponder con mis mejores esfuerzos para contribuir a la construcción de una patria más justa, donde la ciencia sea una herramienta para edificar, avanzar y hermanar.

A todos los que han contribuido a la construcción de la física.
A SECIHTI (CBF-2025-I-615) y PAPIIT (IA102224,IA103126) por el apoyo en este trabajo.

Compendio

La presente tesis ha tenido como objetivo principal el estudio del proceso $pp \to ttWW$ a NLO en QCD. Este proceso constituye el fondo ("background") dominante del tipo $t\bar{t}VV$ para producción de cuatro tops $t\bar{t}t\bar{t}$, proceso que es de particular interés no sólo en la física dentro del Modelo Estándar, sino de aquella Más Allá del Modelo Estándar en el sector pesado, ya que se espera que posibles señales de nueva física puedan manifestarse a través de este canal. Dado que la producción de $t\bar{t}t\bar{t}$ es un proceso extremadamente raro en experimentos como el LHC, resulta fundamental contar con un control preciso de sus procesos de fondo.

En la literatura ya existen estudios previos de los procesos de fondo del tipo $t\bar{t}V$, como lo son $t\bar{t}Z$ y $t\bar{t}W$ a nivel next-to-leading order (NLO), con radiación tipo cromodinámica (QCD) y/o electrodébil (EW) con algoritmos de Parton Shower (PS), que han demostrado que resulta fundamental y absolutamente complementario integrar $t\bar{t}VV$, donde V representa bosones vectoriales del Modelo Estándar, al compendio de estudios de fondo, por lo que este trabajo es el primero en abordar este proceso a un nivel profundo.

El objetivo ha sido proporcionar una predicción teórica de $pp \to t\bar{t}WW$ en QCD mediante un marco que permita acoplar PS, siendo este específicamente el método POWHEG (*Positive Weight Hardest Emission Generator*). Este enfoque permite la generación de eventos mediante técnicas Monte Carlo con pesos positivos, al mismo tiempo que trata la primera emisión radiativa a partir de los elementos de matriz exactos, lo cual es crucial para evitar el doble conteo de emisiones más duras al realizar el matching con los algoritmos PS.

La implementación de este cálculo se lleva a cabo dentro del framework POWHEG-BOX, donde los elementos de matriz necesarios son generados mediante OpenLoops. La información del evento se almacena en formato Les Houches.

Con ello, POWHEG permite usar un documento que tenga el formato correcto para alimentar posteriormente al generador del shower (PYTHIA) encargado de simular la evolución completa de la cascada partónica y la hadronización.

A partir de esto, se obtienen las predicciones para observables inclusivas y exclusivas

del evento, incluyendo variables hadrónicas, bosónicas, aquellas asociadas al quark top, y al proceso completo.
Debido a las particularidades del esquema de **matching** implementado en POWHEG, resulta pertinente realizar la simulación en MG5_aMC@NLO, que emplea un esquema NLO+PS distinto y que no requiere de la introducción de funciones de amortiguamiento (*damping functions*). Esta comparación permite validar e identificar la dependencia de resultados sobre los esquemas utilizados en las predicciones realizadas.
El desarrollo teórico y numérico de este proceso requiere de la comparación de diversos aspectos fundamentales, los cuales se abordan en esta tesis:

- Física del quark top y su relación con el bosón de Higgs.
- Colisiones hadrónicas y partónicas, por cálculos a orden fijo y por parton showers.
- Estructura y tratamiento de divergencias infrarrojas.
- Combinación de cálculos a NLO con Parton Showers, en particular, el método POWHEG.
- Implementación computacional del proceso $t\bar{t}WW$.
- Resultados fenomenológicos y discusión

Finalmente, cabe recalcar que el valor de este trabajo no se limita al estudio de $t\bar{t}WW$, sino que también sirve para establecer los fundamentos teóricos y computacionales necesarios para extender el análisis a los procesos secundarios de fondo, que en particular son $t\bar{t}WZ$ y $t\bar{t}ZZ$, que serán objeto de futuros proyectos.

Resumen

Este trabajo tiene como objetivo estudiar las contribuciones a NLO QCD en el proceso de producción de $t\bar{t}WW$ en colisiones hadrónicas, emparejando los cálculos a orden fijo a "*next-to-leading order*" con efectos de cascadas de partículas (*Parton Showers*). El emparejamiento de ambos marcos se realiza por medio del método POWHEG, dentro del *framework* POWHEG-BOX, buscando evitar el doble conteo en la región de radiaciones más duras y generando pesos positivos en los procesos Monte Carlo. Los cálculos también se realizan en MG5_aMC@NLO, que emplea un modo distinto de emparejamiento NLO+PS y se comparan los resultados de las distintas observables generadas en ambas fuentes. Al mismo tiempo, este trabajo es base para la continuación de cálculos tipo $t\bar{t}tVV$, donde V denota bosones vectoriales, imprescindibles para el estudio del proceso de gran interés fenomenológico, producción de $t\bar{t}t\bar{t}$.

Abstract

This work aims to study the NLO QCD contributions to the $t\bar{t}WW$ production process in hadron collisions, matching fixed-order calculations at *next-to-leading order* with parton shower effects (*Parton Showers*). The matching of both frameworks is performed by means of the POWHEG method, within the POWHEG-BOX framework, seeking to avoid double counting in the region of hardest emissions and generating positive weights in Monte Carlo processes. The calculations are also performed in MG5_aMC@NLO, which employs a different matching scheme, and results are compared for different generated observables. At the same time, this work serves as a basis for the continuation of calculations of the type $t\bar{t}tV$, where V denotes vector bosons, essential for the study of the process of great phenomenological interest, the production of $t\bar{t}t\bar{t}$.

Contents

1 Introduction

1.1 History of the top quark

In 1964, CP violation in the neutral kaon system was first observed [1]. This implied that the electroweak model had to be able to incorporate a mechanism capable of explaining such a phenomenon while preserving the consistency of what had been described up to that point.

In 1973, Makoto Kobayashi and Toshihide Maskawa [2] demonstrated that CP violation can exist in a physical way if there are at least three quark generations. The inclusion of a third generation allows for the emergence of a complex phase that cannot be removed by the redefinition of the quark fields, as is the case when only two generations are considered. The presence of this irreducible complex phase causes the quark mixing matrix to behave differently under CP conjugation, giving rise to observable differences between a process and its CP-conjugate counterpart.

After the experimental discovery of the bottom quark in 1977 [3], it was just a matter of time for the top quark to be discovered, and major efforts were made to continue its search. But finding it was not an easy task, due to its large mass, comparable to that of a gold atom (larger than previously thought) which required high energy collisions for its production.

The experimental discovery of the top quark was finally reached in 1995 at Fermilab, through the CDF and D0 experiments [4,5], which operated at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV. Through these experiments, single and top-quark pair production events were produced and observed, allowing for the first experimental studies on its properties, including the first measurements of its mass.
After its discovery, experiments have focused on improving the precision with which its physical properties are known, since the top quark plays a special role in the calculation of quantities within the Standard Model (SM). For example, the correction

to the W and Z propagators receive dominant contributions from the top-quark mass through the ρ parameter (when considering $m_b \ll m_t$) [6, 7] of the form

$$\Delta\rho_t \approx \frac{3\mathrm{G}_F m_t^2}{8\sqrt{2}\pi^2} \,. \tag{1.1}$$

Here, the quadratic dependence on its mass makes the top quark a fundamental piece in electroweak precision fits. Similarly, phenomena associated with physics beyond the Standard Model may manifest through deviations of the top-quark and Higgs-boson observables from the SM predictions. In this regard, the operational start of the Large Hadron Collider (LHC) in 2009 marked the beginning of a new experimental era.

The first runs, globally known as Run 1 (2010-2012), were performed at center-of-mass energies of 7 and 8 TeV. Subsequently, the Run 2 stage (2015-2018) operated at 13 TeV. After a period of upgrades in the accelerator (Long Shutdown 2), Run 3 began in 2022 and continues to operate at 13.6 TeV. This substantial increase in the energy with respect to the Tevatron allowed for a much more abundant production of massive particles, such as the top quarks and Higgs bosons, and enabled the observation of phenomena favored by higher energies. At the LHC, top quarks are predominantly produced as $t\bar{t}$-pairs and to a less extent in association with Z, W, γ and H bosons. As these processes, as well as other rare ones such as $t\bar{t}t\bar{t}$ production, are crucially important to study the top-quark couplings to SM particles and possibly to physics Beyond the SM. This large statistical dataset has enabled an increase in the precision of the determination of masses, decay widths, couplings, and also the study of rare processes such as $t\bar{t}t\bar{t}$ production, as well as processes associated with Higgs boson production [8].

Nowadays, the ATLAS and CMS experiments at the LHC continue to explore the properties of the heavy sector of the SM in order to search for anomalies that may signal the presence of new physics.

1.2 A brief overview of Top-quark physics

1.2.1 The Top-Quark in the Standard Model

The electroweak sector of the Standard Model is based on the gauge group $SU(2)_L \times U(1)_Y$. Based on experimental data, left-handed fermions are postulated as doublets under $SU(2)_L$, while right-handed fermions are singlets. For example, the third-generation quark doublet is written as [9],

$$Q_L = \begin{pmatrix} t_L \\ b_L \end{pmatrix} \tag{1.2}$$

where the two components have weak isospin third component $T_3 = +1/2$ and $T_3 = -1/2$, respectively. Here, T_3 denotes the eigenvalue of the diagonal generator of $SU(2)_L$. Right-handed quarks are postulated to be an $SU(2)_L$ singlet, where $T_3 = 0$, which means that there is no interaction with the W boson.

The charged weak interactions connect the two components of the doublet through the exchange of the charged gauge bosons $W^{\pm}$, giving rise to charged currents such as $t \to bW^+$. These interactions are first written in the weak (flavor) basis as

$$\mathcal{L}_W \supset \frac{g}{\sqrt{2}} \bar{t}_L \gamma^\mu b_L W^+ + \frac{g}{\sqrt{2}} \bar{b}_L \gamma^\mu t_L W^-. \tag{1.3}$$

On the other hand, after electroweak symmetry braking, the Yukawa interactions generate mass terms for the quarks, with non-diagonal mass matrices,

$$\mathcal{L}_Y = -\bar{Q}_L Y_d \Phi d_R - \bar{Q}_L Y_u \tilde{\Phi} u_R + \text{h.c.}. \tag{1.4}$$

To obtain the physical quark states, unitary transformations must be performed on the up-type and down-type quark fields independently to diagonalize the mass matrices.

Because the rotations in the up and down-type sectors are different, the charged-current interactions acquires a mixing matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb}. \end{pmatrix} \tag{1.5}$$

This also changes the form of the $\mathcal{L}_{\mathcal{W}}$. Now, in the mass basis, the weak interaction

is

$$\mathcal{L}_W \supset \frac{g}{\sqrt{2}}\bar{t}_L\gamma^\mu(V_{tb}b_L + V_{ts}s_L + V_{td}d_L)W^+_\mu + \text{h.c.} \tag{1.6}$$

Moreover, due to the hierarchy $V_{td} \ll V_{ts} \ll V_{tb} \approx 1$, the top quark decays almost exclusively via $t \to bW$, with a branching ratio close to 100 %. Its large mass implies a very short lifetime, shorter than the typical hadronization timescale. Consequently, the top quark decays before forming bound states, allowing direct access to its spin and polarization properties through its decay products [10].
From this, a distinctive feature is the large Yukawa coupling of the top quark,

$$y_t = \frac{\sqrt{2}m_t}{v} \approx 0.995 \ , \tag{1.7}$$

which is the largest of the SM, and makes it strongly coupled to the Higgs sector. This implies that it could play a special role in the electroweak symmetry breaking mechanism.

1.2.2 Mass

In the Standard Model, fermion masses, m_f, arise from their Yukawa couplings, y_f, to the Higgs field after Spontaneous Symmetry Breaking (SSB). The relation is

$$m_f = \frac{v}{\sqrt{2}}y_f \ , \tag{1.8}$$

where v is the Higgs-field vacuum expectation value.

Among all quarks in the SM, the top quark is the heaviest and therefore possesses the largest Yukawa coupling, $y_t \approx 0.995$, implying a particularly strong coupling to the Higgs boson. Due to this, the top quark plays an important role in electroweak radiative corrections and in the scalar potential evolution at high energies.

Since the top quark carries color charge, it cannot be a physically observable free particle due to the color confinement property in Quantum Chromodynamics (QCD). For this reason, the top quark mass does not directly correspond to an observable quantity in the strict sense of an on-shell stable particle, but it must be determined through its decay products and carefully defined within a specific theoretical scheme. The most precise determination of the top-quark mass comes from *direct measurements*, where the kinematics of its decay products is reconstructed.
These measurements compare kinematic distributions that are sensitive to the mass with simulations generated using Monte Carlo event generators. At the LHC, this

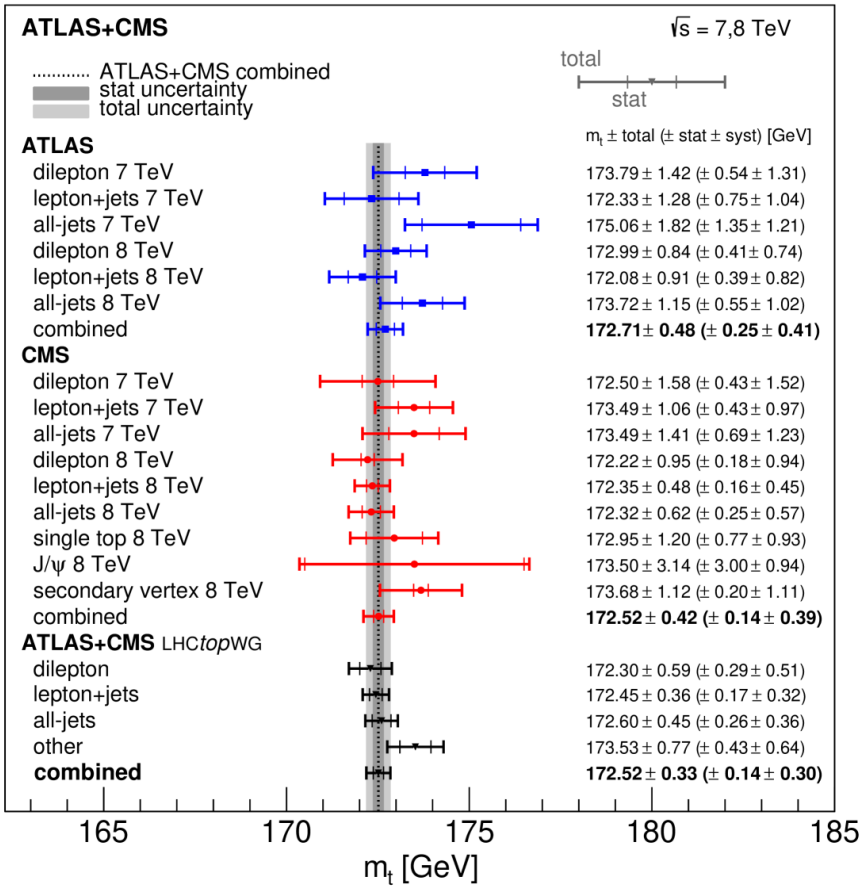
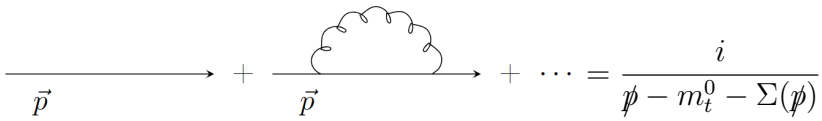


Figure 1.1: Comparison of ATLAS and CMS data at $\sqrt{s} = 8$ TeV for the m_t mass meassured thorugh different channels, and its combinations [11].

approach has enabled precision at the level of a few hundred MeV. The combined results from the ATLAS and CMS experiments at 8 TeV, Fig. (1.1) have led to the determination of the top-quark mass of [11]

$$m_t^{LHC} = 172.52 \pm 0.14(\text{stat}) \pm 0.30(\text{syst}). \tag{1.9}$$

In quantum field theory, the definition of particle masses requires a *renormalization* procedure to deal with the unphysical ultraviolet divergences, UV (referring to large momenta or energies) arising from radiative corrections. Depending on the renormalization scheme adopted, for the top quark different mass definitions can be introduced, [12]. For instance the *pole mass* is identified with the position of the pole in the renormalized fermion propagator. Pictorially



$$= \frac{i}{\not{p} - m_t^0 - \Sigma(\not{p})} \tag{1.10}$$

Figure 1.2: Self-energy for the top-quark at higher orders.

Here, $\Sigma(\not{p})$ is the fermion self energy that arises from loop corrections to the propagator. It shifts the position of the pole in the propagator and consequently modify the mass of the fermion.

$$\Sigma(\not{p}, m_t^0, \mu) = m_t^0 \left(\frac{\alpha_s(\mu)}{\pi}\right) \left[\frac{1}{\epsilon} + \ln\left(4\pi \mathrm{e}^{-\gamma_E}\right) + A^{fin}\left(\frac{m_t^0}{\mu}\right)\right]. \tag{1.11}$$

Here, μ is the renormalization constant, m_t^0 the mass of the propagator at tree level, $1/\epsilon$ the pole after regularization in $d = 4 - 2\epsilon$ dimensions, and A^{fin} is the finite contribution. The full propagator is obtained by resumming self-energy insertions to all orders. The pole mass is then defined by the condition

$$\det[\not{p} - m_t^0 - \Sigma(\not{p})] = 0. \tag{1.12}$$

where $\Sigma(\not{p}_{pole}) = \Sigma^{(1)} + \Sigma^{(2)} + \cdots$,.
Solving equation (1.12) for $\not{p}$ at first order in α_s,

$$\not{p}_{pole} \approx m_t^0 + \Sigma^{(1)}(\not{p}_{pole}), \tag{1.13}$$

and substituting in $\Sigma(\not{p}_{pole})$,

$$\Sigma(\not{p}_{pole}) = \Sigma(m_t^0 + \Sigma^{(1)}) = \Sigma(m_t^0) + \Sigma^{(1)}\Sigma'(m_t^0) + \cdots, \tag{1.14}$$

which at first order in α_s gives us

$$\Sigma(\not{p}_{pole}) \simeq \Sigma(m_t^0). \tag{1.15}$$

Then,

$$\not{p}_{pole} = m_t^{pole} \approx m_t^0 + \Sigma^{(1)}(m_t^0) \tag{1.16}$$

This equation means that the repeated absorption and emission of virtual particles modify the fermion propagator, shifting the position of its pole and consequently the physical mass. After substituting equation (1.11) in (1.16),

$$m_t^{pole} = m_t^0 \left[1 + \left(\frac{\alpha_s(\mu)}{\pi}\left[\frac{1}{\epsilon} + \ln(4\pi/e^{\gamma_E}) + A^{fin}(m_t^0/\mu)\right]\right)\right] + ..., \tag{1.17}$$

This is the mass that within the perturbative theory defines the notion of rest mass of the quark as an on-shell particle. And it is strictly defined within dimensional regularization. It depends linearly on the renormalization scheme implemented to deal with the infrared momenta.

In particular, in QCD, the one-loop self-energy contribution $\Sigma^{(1)}$ can be generalized by inserting a series of vacuum polarization bubbles into the gluon propagator. The resulting series diverges factorially at large orders. To deal with this, a Borel transformation is introduced, which translates the divergencies to singularities in the Borel plane. Due to that the contour defined in the inverse Borel transform to avoid singularities is not unique, the pole mass on the top quark acquires an intrinsic ambiguity of the order of Λ_{QCD}. [13]

Other renormalization schemes implemented to define the mass are

- the *Modified Minimal Subtraction* $\overline{\text{MS}}$ scheme, in which the mass absorbs only the UV divergences of the quark self-energy. This mass is explicitly scale dependent and evolves according to the group equations. Its running with the renormalization scales makes it the preferred definition for high-energy processes and renormalization group analyses.

- the *mass With Subtracted Renormalon* MSR scheme that can be regarded as an extension of the $\overline{\text{MS}}$ mass to scales below m_t. It absorbs quantum corrections above a chosen scale R, while preserving the cancellation of low-energy contributions with real radiation effects. This makes it a short-distance mass free from renormalon ambiguities and particularly suitable for observables whose mass sensitivity is governed by scales much smaller than the top-quark mass

and the Monte Carlo mass, m_t^{MC}, which is a parameter defined within a Monte Carlo generator.

The pole mass is not measured directly. It is extracted by comparing experimental observables with theoretical predictions expressed in terms of the pole mass. Establishing a precise relation between the mass definition by means of kinematic reconstruction and Monte Carlo methods, and the definition of the pole mass or the mass calculated by other renormalization schemes is still an open area of research.

1.2.3 Charge

As a member of the weak-isospin doublet, see equation (1.2), its electric charge follows from the Gell-Mann–Nishijima relation

$$Q = T_3 + \frac{Y}{2}, \tag{1.18}$$

where $T_3 = +1/2$ and $Y = 1/3$ for up-type quarks. Therefore,

$$Q_t = +\frac{2}{3}e, \tag{1.19}$$

being e the elementary electric charge.
The charge has been measured through experiments such as those performed in the ATLAS collaboration, yielding a value of $q_t = 0.64 \pm 0.02(\text{stat}) \pm 0.08(\text{syst})$ [14].

Due to the confinement, quarks will form hadrons, in which case electric charge always appears in integer quantities. The top quark is an exceptional case, as it decays before hadronizing. Information about its charge is reconstructed directly from its decay products.

1.2.4 Decays and Spin

Decays

The top quark is a 1/2 spin particle. At hadron colliders such as the LHC, the tops are produced unpolarized. Since weak interactions couple only left-handed fermions, the top quark decay is governed by the $V - A$ interaction.

Within the Standard Model at tree level, three decay modes for the top quark are possible,

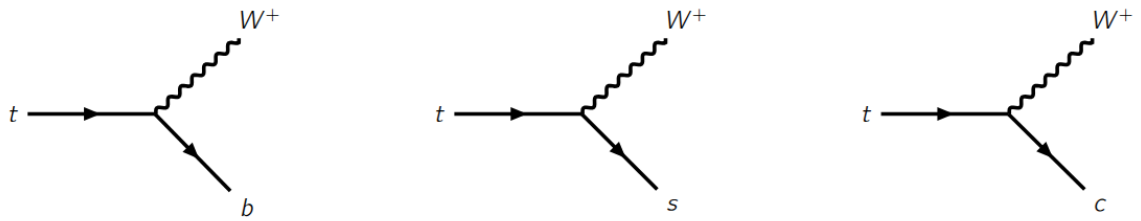


Figure 1.3: Top quark SM decays.

where the decay channel $t \to bW$ is by far the dominant one by several orders of magnitude due to the CKM matrix elements.

The decay width at tree level (Leading Order) is

$$\Gamma_{LO(t\to bW)} = \Gamma_0 \left(1 - 3\frac{M_W^4}{m_t^4} + 2\frac{M_W^6}{m_t^6}\right) \approx 1.56 \text{ GeV} \tag{1.20}$$

with $\Gamma_0 = G_F m_t^3 |V_{tb}|^2/(8\pi\sqrt{2})$ at the lowest order. By considering $m_W = 0$, $m_b = 0$, $|V_{tb}| \to 1$ and $M_W = 80.4$ GeV, $\Gamma_0 \approx 1.76$ GeV. [15].
Considering radiative corrections at second order in QCD,

$$\Gamma(t \to bW)/|V_{tb}|^2 \approx 0.807\ \Gamma_0 = 1.42 \text{ GeV}. \tag{1.21}$$

Spin

Due to non-perturbative QCD interactions, hadronization in light quarks smears the spin information of the original partons.

The case of the top quark constitutes an exception. Its large mass conditions its lifetime and hadronization time, which respectively are

$$\tau_t \sim \frac{1}{\Gamma_t} \sim 5 \times 10^{-25} s, \qquad \tau_{had} \sim \frac{1}{\Lambda_{QCD}} \sim 10^{-23} s. \tag{1.22}$$

Hence, the top quark decays before it hadronizes, allowing the spin information to be accessed by reconstructing the decay process.

By way of example, in the semileptonic decay of top quarks at tree level, the angular distribution of the decay width is directly correlated with the spin orientations through [16]

$$\frac{1}{\Gamma}\frac{\mathrm{d}\Gamma}{\mathrm{d}\cos\chi_i} = \frac{1}{2}(1 + |\vec{P}|\kappa_p \cos\chi_i), \tag{1.23}$$

where $\kappa_p \approx 1$ at tree level in the SM, χ_i is the angle between the momentum of the i-th product of the decay and the top-quark spin axis in the top-quark rest frame.

1.3 Background Processes

The experimental identification of a process at the LHC is performed through the identification and counting of events with signal-like signatures. Disentangling the underlying process responsible for each of these signatures can be challenging, especially considering the technical limitations of the detectors in identifying, for instance,

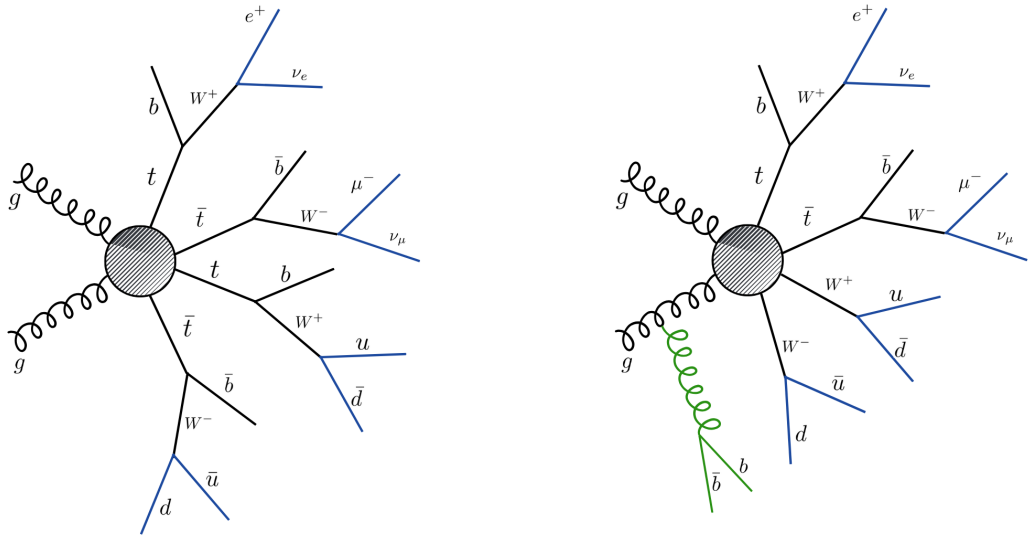


Figure 1.4: Comparison between two gluon fusion processes leading to the same experimental signature. *Left:* four top production, $gg \to t\bar{t}t\bar{t} \to e^+\nu_e\mu^+\nu_\nu b\bar{b}b\bar{b}\, jjjj$. *Right:* top-pair production in association with two W gauge bosons, and an initial state,$g \to b\bar{b}$ radiation, $gg \to t\bar{t}WW + g \to e^+\nu_e\mu^+\nu_\nu\ jjjj\ b\bar{b} + b\bar{b}$.

low-energy and collinear radiation. This becomes an even greater challenge when searching for rare processes such as $t\bar{t}t\bar{t}$ production.

For events associated with top-quark production, one has $t \to W^+b$ with $W^+ \to l^+\nu$ or $W^+ \to q\bar{q}'$. Both the b and the q partons cannot be detected directly, due to color confinement. Instead they are reconstructed as hadronic jets that are formed from successive radiated particles that come from a primary gluon or quark, followed by their hadronization, producing collimated spray of hadrons. Therefore, the final state typically contains many jets, b-jets, leptons, and missing energy. In the left hand side (l.h.s) of the figure 1.4, a possible final state for $t\bar{t}t\bar{t}$ is illustrated, where multiple leptons and several b-jets are produced. Also, a possible final state for $t\bar{t}W^+W^-$ is shown in the right hand side (r.h.s.). The latter can also yield a comparable number of leptons and b-jets. In principle, the two processes could be distinguished by the total number of jets; however, considering that the experiment may have difficulties identifying all jets, or that additional $b\bar{b}$ pairs may be produced through radiative effects, an accurate identification may become very difficult. Therefore, **the production of $t\bar{t}W^+W^-$ represents a background of $t\bar{t}t\bar{t}$.**

1.4 Single and di-quark production

Single top quark production

$$\mathcal{L}_W \supset \frac{g_s}{2\sqrt{2}} V_{tb} \left[\bar{t}\gamma^\mu(1-\gamma^5)b\right] W^+ + \text{h.c.}$$

arises from the charged weak current interaction introduced in eq. (1.6) that couples up-type and down-type quarks through the exchange of a W boson, with a strength determined by the CKM matrix.
In particular, the Wtb vertex couples top and bottom quarks and provides the mechanism for producing a single top quark through electroweak interactions.

On the other hand, top-quark pair production can be understood from the QCD interaction term (see Section 2.3)

$$\mathcal{L}_{QCD} \supset g_s \bar{t}\ \gamma^\mu t\ T^a G^a_\mu,$$

that comes from Eq. (2.22). Unlike the charged-current weak interaction, QCD conserves quark flavor and, as a result, it does not contain vertices connecting quarks of different flavor or different generations. As can be seen from Eq. (1.4), a gluon couples only to quarks of the same flavor and cannot mediate transitions such as $b \to t$, then, the production of a single top quark requires the weak interaction discussed previously.

Top quarks can nevertheless be produced through strong interactions in top-antitop pairs. At the Large Hadron Collider (LHC), $t\bar{t}$ production is dominated by gluon-gluon fusion, $gg \to t\bar{t}$, owing to the large gluon density inside the proton at the energies probed by the LHC.

Pair production can be associated with additional particles, for example

$$pp \to t\bar{t}X, \quad t\bar{t}XY, \qquad\qquad (X, Y = \gamma, Z, W, H, j).$$

These production channels, which are worth highlighting because of their greater complexity due to the larger number of particles in the final state, allow the investigation of the coupling of the top quark to other particles, such as gauge bosons and the Higgs boson.

Of course, the larger particle multiplicity leads to more complex event topologies, increasing the challenge of reconstructing and identifying the final state. These final

states require the identification of their own suitable associated experimental signatures, as well as reconstruction techniques to identify each of the final-state objects.

1.5 Four-top quark production

Since the experimental discovery of the top-quark in 1995, many events with at least one t have been produced. For instance, due to the much larger energies and luminosities in collisions at the LHC, during Run 2 from 2016 to 2018 around 300 million top quarks were detected by the CMS collaboration [17], and around the same amount by the ATLAS collaboration [18]. That enormous amount of data has allowed physicists to study the main properties of the heaviest quark in the Standard Model and its close relationship with sectors of great interest in physics, such as the previously mentioned connection to the Higgs sector and, consequently, to the mechanism of electroweak symmetry breaking.

However, not all processes involving top quarks are produced abundantly. Among the rare processes is four-top production, $t\bar{t}t\bar{t}$ production, of which only around 2000 events were observed during LHC Run 2 [17]. Despite the low amount of data the analysis performed in the CMS collaboration allowed this process to be observed with a statistical significance of 5.6 standard deviations, exceeding the 5σ threshold conventionally required for a discovery in particle physics. [17]. Owing to its sensitivity to the top-Higgs interaction, this process provides a unique window into the Higgs sector and possible effects of physics beyond the Standard Model. It can constrain the total width of the Higgs boson and can be enhanced by beyond the Standard Model scenarios where resonant states are present. So, the four top production could provide a unique window into the search of a broad class of new physics, as that related to the Higgs sector.

The $t\bar{t}t\bar{t}$ production process has been calculated at next-to-leading order (NLO) in QCD [19], and later extended to include electroweak EW contributions and complete QCD+EW predictions [20]. More recently, top-quark decays have been incorporated using different approximations [21]. These calculations have also been matched to parton showers [20]. Fig. 1.5 illustrates an example of $t\bar{t}t\bar{t}$ production through gluon-gluon fusion. In this diagram, a top-quark pair is produced together with a virtual Higgs boson, which subsequently couples to a second top-quark pair.

On the other hand, a possible Beyond-the-Standard-Model scenario that may contribute to an enhancement of the $t\bar{t}t\bar{t}$ is *Two Higgs Doublet Model* (2HDM) [22],

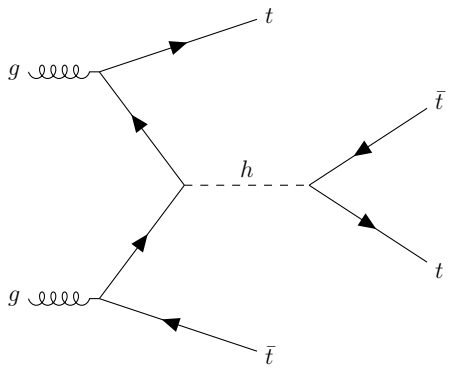


Figure 1.5: $t\bar{t}t\bar{t}$ production from gg-fusion mediated by a virtual SM Higgs.

where a $H/A \to t\bar{t}$ in association with a top-pair (fig. 1.6). Here H corresponds to a new heavy neutral scalar while A to a neutral pseudo scalar particle both decaying to $t\bar{t}$. Direct searches for heavy neutral scalars have been carried out by the ATLAS and CMS collaborations. Indirect searches, on the other hand, rely on precision measurements of Standard Model Higgs-boson production and decay rates. In this approach, possible deviations from the Standard Model predictions can provide evidence for the presence of additional scalar states. Diagram in (fig. 1.6) can avoid the destructive interference with the SM background which distorts the Breit-Wigner peak in the $t\bar{t}$ mass peak [23].

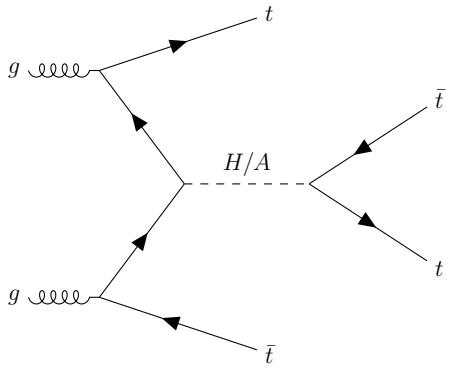


Figure 1.6: $t\bar{t}t\bar{t}$ production mediated by H/A Higgs bosons from a 2HDM.

In the following, a general discussion of the differences between various channels used to search for deviations from Standard Model predictions is presented.

Consider the production of a heavy Higgs boson through gluon fusion,

$$gg \to H \to t\bar{t}. \tag{1.24}$$

If this were the only mechanism capable of producing a $t\bar{t}$ final state, the invariant-mass distribution $m_{t\bar{t}}$ would exhibit a Breit–Wigner resonance centered at $m_{t\bar{t}} = m_H$. However, within the Standard Model there exists a significant background arising from the process

$$gg \to t\bar{t}, \tag{1.25}$$

so that the observed amplitude results from the coherent combination of both contributions. Then, the corresponding cross section is therefore determined by

$$|\mathcal{M}_{t\bar{t}}|^2 = |\mathcal{M}_{\mathrm{QCD}}|^2 + |\mathcal{M}_H|^2 + 2\,\mathrm{Re}\left(\mathcal{M}_{\mathrm{QCD}}\mathcal{M}_H^*\right), \tag{1.26}$$

where the last term corresponds to the interference between the resonant signal and the Standard Model background.
The amplitude associated with the heavy Higgs boson takes the form

$$\mathcal{M}_H \propto \frac{1}{\hat{s} - m_H^2 + i\, m_H \Gamma_H}, \tag{1.27}$$

with

$$\hat{s} = (p_g + p_g)^2 = (p_t + p_{\bar{t}})^2 = m_{t\bar{t}}^2. \tag{1.28}$$

Across the resonance, the real part of this amplitude changes sign due to the factor $(\hat{s} - m_H^2)$ and the consequence is that the interference term becomes constructive on one side of the resonance and destructive on the other. This effect distorts the expected Breit–Wigner shape and complicates the experimental search for a heavy Higgs boson, particularly because the Standard Model production of $t\bar{t}$ pairs is much more abundant than the corresponding new-physics signal.
One alternative strategy is the searching for heavy-Higgs effects in four-top production, as in fig. 1.6.

The Standard Model background for this processs is considerably smaller than which is a big difference with the $t\bar{t}$ production case. Consequently, interference effects have a reduced relative impact, making the resonant contribution of a heavy Higgs boson potentially easier to identify.

The production of $t\bar{t}t\bar{t}$ also arises in supersymmetric scenarios through the production of gluinos $\tilde{g}$ (fermionic supersymmetric partners of the gluons in supersymmetric models). Gluino pairs can subsequently decay into top quarks through virtual or on-shell top squarks $\tilde{t}_1$ (the scalar supersymmetric partners of the top quark), leading

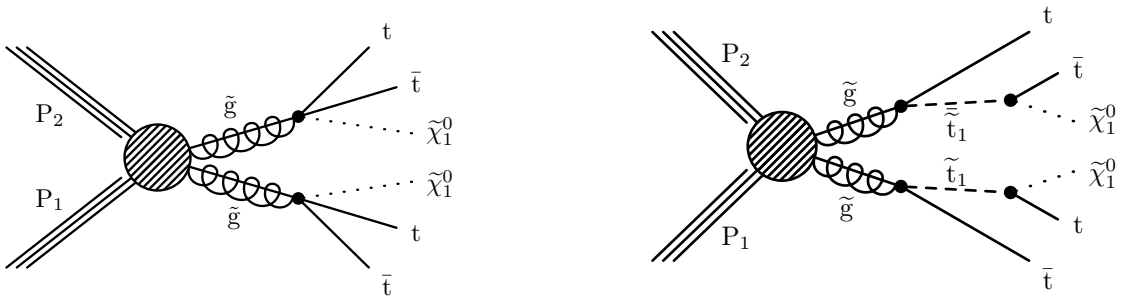


Figure 1.7: SUSY production of $t\bar{t}t\bar{t}$. Two processes producing the same experimental signature.

to final states containing four top quarks [24].

Then,

$$pp \to \tilde{g}\tilde{g} \to t\bar{t}t\bar{t} + E_T^{\text{miss}},$$

where E_T^{miss} denotes the missing transverse energy associated with particles $\tilde{\chi}_1^0$ that escape detection. Since these processes lead to signatures that closely resemble Standard Model four-top production, deviations in the observed $t\bar{t}t\bar{t}$ production rate or in kinematic distributions sensitive to missing transverse energy may provide evidence for physics beyond the Standard Model.

2 Theoretical Framework

2.1 Schematics of hadron collisions

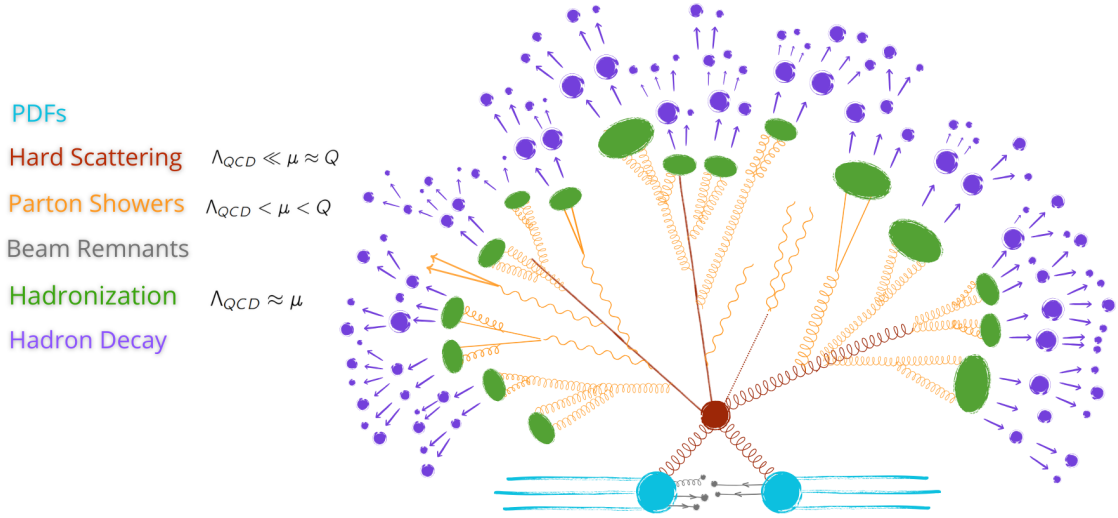


Figure 2.1: Representation of a hadron collision: Hadrons in collision (blue); the hard scattering (red); radiative processes (orange); hadronization (green); hadron decays (purple); secondary interactions (gray). Λ_{QCD} is the scale where QCD becomes a strongly coupled theory, μ is the scale of the interaction and Q, the scale of the hadron interaction.

Unlike leptonic collisions, hadron collisions involve bound states known as *hadrons*. The theoretical description of these processes is not straightforward, as it requires modeling the internal structure of hadrons in terms of their *partonic constituents* while simultaneously taking into account the effects of QCD. The confinement of those constituents and the presence of non-perturbative QCD effects, together with the technological requirements needed to achieve the high collision energies involved in the processes, make their theoretical description a major challenge.

Hadron collisions are a fundamental tool to study the elementary interactions of nature. They allow to test the Standard-Model predictions while searching for deviations

that may indicate signs of new physics. Since hadrons are bound states composed by quarks and gluons (partons), the theoretical description of their interactions is more complex than that of the elementary particles. In the following, a general description of the different stages involved in a hadron collision is presented [25].

In a hadron interaction, the internal constituents are the interacting pieces. The internal composition of the hadron is characterized by means of the **Parton Distribution Functions** (PDFs), $f_i^{\mathbf{h}}(x, \mu_F)$, that describe the probability density of finding a parton i inside a hadron **h**, carrying a fraction x of its total longitudinal momentum. These functions provide a probabilistic description of the internal structure of hadrons.

The direct interaction of partons coming from each incoming hadron is called the **hard scattering**, since it takes place at high energies $Q \gg \Lambda_{QCD}$, where Λ_{QCD} denotes the characteristic scale of QCD. Therefore, this interaction probes very short distances. This stage can be described through perturbative QCD.

During the event, more than one hard interaction may take place, however, such processes are highly suppressed and require a different theoretical treatment.

The energetic final-state partons, emerging from the hard scattering, subsequently undergo QCD radiation consisting of multiple particle emissions. This process is known as **parton shower** and occurs for energy scales $\Lambda_{QCD} < \mu < Q$. The evolution of this radiation is described by the *Parton Shower generators* (PS) which are mainly based on the implementation of the *Dokshitzer-Gribov-Lipatov-Altarelli-Parisi* (DGLAP) equations and will be discussed in further detail in a later section.

At scales $\mu \approx \Lambda_{QCD}$, $\alpha_s \sim 1$ and perturbation theory is no longer applicable. This process must be described using non-perturbative hadronization models that define how color charged partons combine to form color−neutral hadrons.

These hadrons can further decay, due to their instability. The chain of decay continues until it produces stable particles or particles long-lived enough to be detected experimentally. The proton constituents that do not participate in the hard scattering form the **beam remnants**, that carry the rest of the longitudinal momentum and quantum information of the incoming proton. Their subsequent hadronization contributes mainly to soft hadronic activity and is beyons the scope of this work. [26–28]

2.2 Lorentz Invariant Phase Space

The phase space is the space constituted by all possible kinematic states that a particle or a system can occupy that respect energy and momentum conservation.
In particle physics, it cannot depend on the reference frame.

For an n-particle final-state particles, the Lorentz invariant phase space element is defined as

$$\mathrm{dLips}(p_1, p_2, \ldots, p_n) \equiv \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i}, \tag{2.1}$$

where i runs over all final particles, $p_i = (E_i, \vec{p_i}\,)$ and $E_i = \sqrt{|\vec{p_i}\,|^2 + m_i^2}$,
The total four-momentum is defined as $p = \sum_i p_i$, and the invariant mass squared of the system is $s \equiv p^2$.
The integration of the Lorentz-invariant phase-space measure (2.1) over all kinematically allowed final state configurations yields,

$$\mathrm{d}\Phi_n \equiv \mathrm{dLips}(p_1, \ldots, p_n)\, (2\pi)^4 \delta^{(4)}\left(p - \sum_i^n p_i\right), \tag{2.2}$$

$$R_n(s) \equiv \int \mathrm{dLips}(p_1, \ldots, p_n)\, (2\pi)^4 \delta^{(4)}\left(p - \sum_i^n p_i\right). \tag{2.3}$$

The case where $n = 2$ will be useful later, when we discuss the factorization of the n-particle phase space. The corresponding $R_2(s)$ is

$$R_2(s) = (2\pi)^{-2} \int \frac{d^3 p_1}{2E_1} \frac{d^3 p_2}{2E_2}\, \delta^{(4)}(p - p_1 - p_2). \tag{2.4}$$

For a generic momentum, $p^\mu \equiv (p^0, \vec{p})$, such that $p^2 = p^0 - |\vec{p}|^2$ we can use the identity

$$\frac{1}{2E}\, \delta(p^0 - E) = \delta(p^2 - m^2)\, \Theta(p^0), \tag{2.5}$$

and express R_2 as

$$R_2(s) = (2\pi)^{-2} \int \frac{d^3 p_1}{2E_1}\, \delta\big[(p - p_1)^2 - m_2^2\big]\ \Theta\big[(p - p_1)^0\big]\,. \tag{2.6}$$

Since $R_2(s)$ is Lorentz invariant, we may evaluate it in the center-of-momentum (CM) frame, defined by

$$p_1 = (E_1, \vec{p}_{\mathrm{CM}}), \qquad p_2 = (E_2, -\vec{p}_{\mathrm{CM}}), \qquad p = (\sqrt{s},\, \vec{0}). \tag{2.7}$$

From momentum conservation $p = p_1 + p_2$, we obtain $\sqrt{s} = E_1 + E_2$, where

$$E_1^2 = |\vec{p}_{\rm CM}|^2 + m_1^2, \qquad E_2^2 = |\vec{p}_{\rm CM}|^2 + m_2^2. \tag{2.8}$$

Through these relations, we can derive

$$E_1 = \frac{s + m_1^2 - m_2^2}{2\sqrt{s}}, \tag{2.9}$$

$$E_2 = \frac{s + m_2^2 - m_1^2}{2\sqrt{s}}, \tag{2.10}$$

$$|\vec{p}_{\rm CM}| = \frac{1}{2\sqrt{s}}\,\lambda^{1/2}(s, m_1^2, m_2^2), \tag{2.11}$$

where λ is the Källén triangle function:

$$\lambda(s, m_1^2, m_2^2) = s^2 + m_1^4 + m_2^4 - 2sm_1^2 - 2sm_2^2 - 2m_1^2m_2^2 \tag{2.12}$$

$$= \left[s - (m_1 + m_2)^2\right]\left[s - (m_1 - m_2)^2\right]. \tag{2.13}$$

Using Eqs. (2.9, 2.10, and 2.11),

$$(p - p_1)^2 - m_2^2 = s - 2\sqrt{s}\,E_1 + m_1^2 - m_2^2. \tag{2.14}$$

and we obtain

$$R_2(s) = (2\pi)^{-2} \int \frac{d^3p_1}{2E_1}\,\delta\big(s - 2\sqrt{s}E_1 + m_1^2 - m_2^2\big)\;\Theta(\sqrt{s} - E_1). \tag{2.15}$$

While knowing that in the CM frame

$$d^3p_1 = p_{\rm CM}^2\,dp_{\rm CM}\,d\Omega = p_{\rm CM}\,E_1\,dE_1\,d\Omega, \tag{2.16}$$

and Eq. (2.11), we can reduce $R_2(s)$ to the following simple form:

$$R_2(s) = \frac{1}{2}(2\pi)^{-2} \int p_{\rm CM}\,dE_1\,d\Omega\,\delta\big(s - 2\sqrt{s}E_1 + m_1^2 - m_2^2\big) \tag{2.17}$$

$$= \frac{p_{\rm CM}}{16\pi^2\sqrt{s}} \int d\Omega. \tag{2.18}$$

$$= \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{32\pi^2 s} \int d\Omega. \tag{2.19}$$

The integration over the Lorentz-invariant phase-space, corresponding corresponding

with a process with n final-state particles becomes more complex as n increases. Additionally, in the integrand corresponding to the squared amplitude of the process, sharply peaked and localized structures may arise. For this reason, a sampling of points using Monte Carlo algorithms may fail to properly capture the physics in these peaks and becomes expensive, computationally speaking, in high-dimensional spaces.

To deal with this, a factorization of the phase space is introduced. For a process $2 \to n$, the differential final state phase space element is [29]

$$\mathrm{d}\Phi_n(p) = \left[\prod_{i=1}^{n} \frac{\mathrm{d}^4 p_i}{(2\pi)^4} \delta(p_i^2 - m_i^2)\Theta(p_i^0)\right] (2\pi)^4 \delta^{(4)}\left(p_a + p_b - \sum_{i=1}^{n} p_i\right), \tag{2.20}$$

where a and b denote the incoming particles, and $i = 1, ..., n$ the outgoing particles. Then, equation (2.20) can be factorized as

$$\mathrm{d}\Phi_n(p_1, p_2; p_3, ..., p_n) = \mathrm{d}\Phi_{n-m+1}(p_1, p_2; p_3, ..., p_{n-m}, P)\frac{\mathrm{d}P^2}{2\pi}\mathrm{d}\Phi_m(P; p_{n-m+1}, ..., p_n), \tag{2.21}$$

where P is the momentum for a Virtual intermediate particle that connects both factorized processes.
This means that we can split the process into two subprocesses

$$2 \to (n - m + 1) + P, \text{ then, } P \to m \text{ particles.}$$

The recursive factorization of Eq. (2.20) allows an n-body phase space to be expressed as a sequence of two-body phase-space elements connected through virtual intermediate states of momentum P. This property is particularly useful for Monte Carlo event generation, where complex multi-particle final states can be constructed through successive two-body branchings.

2.3 Quantum Chromodynamics (the core of the interaction)

The lagrangian of Quantum Chromodynamics (QCD) is

$$\mathcal{L}_{QCD} = \sum_f \bar{\psi}_f (i\gamma^\mu D_\mu - m)\psi_f - \frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}. \tag{2.22}$$

D_μ is the covariant derivative,

$$D_\mu = \partial_\mu - ig_s G^a_\mu T^a, \tag{2.23}$$

where G^a_μ is the gluon field, while T^a are the generators of $SU(3)_c$, called **color generators** and $a = 1, ..., 8$.

They satisfy

$$tr(T^a T^b) = \delta^{ab}. \tag{2.24}$$

And the relation,

$$\left[T^a, T^b\right] = if^{abc}T^c. \tag{2.25}$$

On the other hand, the field tensor is

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + ig_s f^{abc} G^b_\mu G^c_\nu, \tag{2.26}$$

that comes from the commutator

$$[D_\mu, D_\nu] = ig_s F_{\mu\nu}. \tag{2.27}$$

which measures the curvature of the space-time due to the gauge field. Here, $F_{\mu\nu} = F^a_{\mu\nu}T^a$, and f^{abc} being the *structure constants*.

The last term of eq. (2.26) appears due to the equation (2.25), which is what makes the theory **non-abelian**, unlike the QED case, where the generator of the transformation is the electric charge, which trivially satisfies [Q,Q]=0.

Because the gluon field obeys $A_\mu = G^a_\mu T^a$, then the gluon carries a charge (it differs from the photon case). Then, $F^a_{\mu\nu}F^{a\mu\nu}$ gives rise to the *gluon propagator*, but also to the *three and the four gluon vertices*.

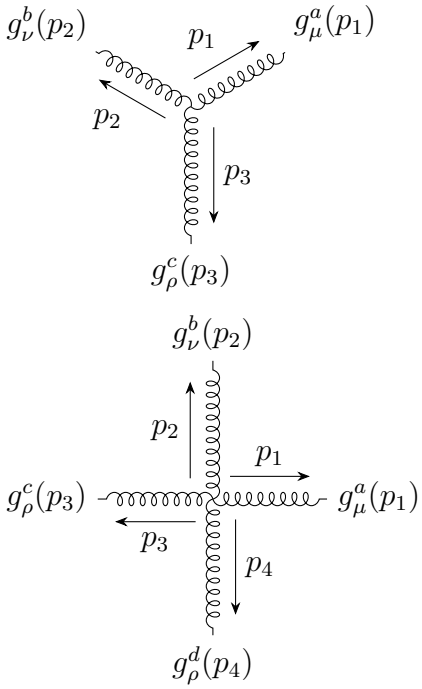
The propagators for quarks and gluons are

$$\delta^{ij}\frac{\not{k} + m}{k^2 - m^2 + i\epsilon^+} \tag{2.28}$$

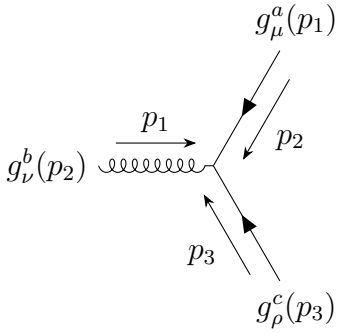
$$\frac{\delta^{ij}}{k^2 + i\epsilon^+}\left(\eta_{\mu\nu} + (\xi - 1)\frac{k_\mu k_\nu}{k^2}\right), \tag{2.29}$$

for i, j color indices.

It is important to note that (2.25), the interaction terms for gluons define the vertices, which does ot exist in QED theory.



At last, the *fermion-gluon* vertex is



2.3.1 Amplitudes in QCD

The amplitudes in QCD can be separated into a **kinematical part** and a **color structure** part.

$$|\mathcal{M}\rangle = \sum_k \mathcal{M}_k \, |c_k\rangle \,, \tag{2.30}$$

where $\mathcal{M}_k$ is the partial amplitude depending on the kinematics of the system; $|c_k\rangle$ are the basis vectors in the color space, which contain all the color information. This

means that the color and the kinematical information can be treated separately in a generic QCD amplitude.

For instance, for a $qq' \longrightarrow QQ'$ scattering process we can describe the color flow by the color operators $\mathbf{T}^a$, and how they act on the full scattering amplitude (2.30). It means, they encode how the color is exchanged between particles (quarks and gluons) and correlations between different parts of the amplitude.

In the color space, the color operators will act for quarks and gluons in the initial or final state as

$$\begin{aligned}
(\mathbf{T}_i^a)_{km} &= T_{km}^a, \text{for } i = \text{final-state quarks (or initial state antiquarks)} \\
(\mathbf{T}_i^a)_{km} &= -T_{mk}^a, \text{for } i = \text{initial-state quarks (or final state antiquarks)} \\
(\mathbf{T}_j^a)_{km} &= if_{akm}, \text{for } j = \text{for } j \text{ gluons}).
\end{aligned} \tag{2.31}$$

The Fierz identity for $SU(N_c)$ generator is

$$(T^a)_{i_3 i_4} (T^a)_{i_2 i_1} = \left(\delta_{i_3 i_1} \delta_{i_4 i_2} - \frac{1}{N_c} \delta_{i_3 i_4} \delta_{i_2 i_1} \right). \tag{2.32}$$

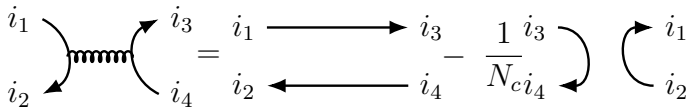


Figure 2.2: color-flow representation of the Fierz identity.

which describes the gluon exchange between two quark lines in terms of a leading-color contribution (first term) and a subleading contribution suppressed by $1/N_c$ (second term).

Then, for this particular process, the elements that define the color basis are

$$\begin{aligned}
|c_1\rangle &= \delta_{i_3 i_1} \delta_{i4 i2} \\
|c_2\rangle &= \delta_{i_3 i_4} \delta_{i_2 i_1}.
\end{aligned} \tag{2.33}$$

To illustrate the way these operators act over the color base, let us calculate $\mathbf{T}_4^a \mathbf{T}_3^a |c_1\rangle$

$$\begin{aligned}
\mathbf{T}_4^a\mathbf{T}_3^a\left|c_1\right\rangle &= T^a_{i_4j_4}T^a_{j_3i_3}\delta_{i_1j_3}\delta_{i_2j_4} = -T^a_{i_4i_2}T^a_{i_1i_3} = \delta_{i_4i_3}\delta_{i_2i_1} - \frac{1}{N_c}\delta_{i_4i_2}\delta_{i_1i_3}, \\
\mathbf{T}_4^a\mathbf{T}_3^a\left|c_1\right\rangle &= -\left|c_1\right\rangle + \frac{1}{N_c}\left|c_2\right\rangle .
\end{aligned} \tag{2.34}$$

while for $\mathbf{T}_1^a\mathbf{T}_1^a\left|c_1\right\rangle$

$$\begin{aligned}
\mathbf{T}_1^a\mathbf{T}_1^a\left|c_1\right\rangle &= (-1)^2T^a_{ji_1}T^a_{j_1j}\delta_{j_1i_3}\delta_{i_2i_4} = C_F\delta_{i_3i_1}\delta_{i_2i_4}, \\
\mathbf{T}_1^a\mathbf{T}_1^a\left|c_1\right\rangle &= C_F\left|c_1\right\rangle .
\end{aligned} \tag{2.35}$$

This identity illustrates that the color flow in QCD processes is not trivial, but instead involves non-trivial correlations between different quark and gluon lines. Even for a simple quark scattering process, the exchange of a single gluon generates multiple independent color structures, corresponding to different possible color flows. The leading contribution describes the dominant color connection between quark lines, while the subleading term, suppressed by $1/N_c$, represents singlet-like configurations and interference effects. This reflects the intrinsically non-Abelian nature of QCD, where color is exchanged dynamically among partons and gluons themselves carry color charge. As the number of external legs increases, these correlations become increasingly complex, leading to a rapidly growing number of possible color configurations. This rich color structure is one of the main reasons why QCD calculations become particularly laborious, especially for multi-leg processes and higher-order corrections.

2.3.2 Perturbative and non-perturbative Quantum Chromodynamics

At high energy scale, strong partonic interactions can be described within the framework of perturbative QCD (pQCD). In this regime, physical observables can be describe through an expansion in powers of the strong coupling constant α_s, giving rise to sufficiently precise theoretical predictions. At lower energy scales, on the other hand, α_s becomes sufficiently large to spoil the perturbative expansion. Focusing on the perturbative description, higher order corrections introduce loop diagrams that contain integrals on virtual momenta. These integrals become divergent when contributions from arbitrarily large momentum scales are included, giving rise to ultraviolet (UV) singularities. To treat those singularities, a renormalization procedure is ap-

plied, in which the parameters of the theory are redefined to absorb the divergent contributions into physical quantities. As a consequence, an unphysical parameter known as the renormalization scale μ_R is introduced. The strong coupling constants become dependent on this scale and can be expressed as [30],

$$\alpha_s(\mu_R) = \frac{4\pi}{(11 - \frac{2}{3}n_f)\ln[\mu_R^2/\Lambda_{QCD}^2]}, \tag{2.36}$$

where n_f is the number of active quark flavors. Physical observables cannot depend on the parameters μ_R. Their theoretical dependence on the renormalization scale decreases as the perturbative order considered in the calculation increases.

2.4 Fixed-Order calculations

An inclusive observable is obtained by summing over all final states that contribute to the quantity under consideration. Then, the *inclusive* partonic cross-section is

$$\hat{\sigma} = \frac{1}{2\hat{s}} \sum_f \int \mathrm{d}\Phi_f |\mathcal{M}(i \to f)|^2. \tag{2.37}$$

Thus, the final states that contribute at order α_s^m, can correspond to different multi-particle states.

The cross section can be written as the perturbative expansion can be expressed in terms of α_s^k corresponding to the leading order contribution to the *partonic cross section*,

$$\hat{\sigma}(ij \longrightarrow X) = \alpha_s^k \sum_{m=0}^{\infty} \hat{\sigma}_{ij}^{(m)} \alpha_s^m \tag{2.38}$$

α_s is the *strong coupling constant*. Here, k denotes the LO power of the strong coupling associated with the Born process, while the sum over m accounts for the higher order corrections to the cross section that arise from loop and Real diagrams.

Then, for the *Leading Order* (LO) cross-section $m = 0$, while for a *next-to-leading order* (NLO) calculation of the cross-section, m runs from 0 to 1 where the final states n and $n + 1$ are included. For instance, for the $q\bar{q} \to q\bar{q}$ process, contributions to the cross section at NLO are constructed using the terms proportional to α_s and α_s^2,

In hadronic collisions, perturbative and non-perturbative descriptions can be separated due to the *factorization theorem* that allows non-perturbative effects to be

absorbed into the Parton Distribution Functions (PDFs), while the short-distance interaction remains perturbatively calculable. This separation introduces the factorization scale μ_F.

At NLO, calculation for the *hadronic cross section* is given by [31]

$$\sigma^{NLO} = \sum_{ij} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i^{\mathbf{h}_1}(x_1, \mu_F) f_j^{\mathbf{h}_2}(x_2, \mu_F) \left[\int \mathrm{d}\hat{\sigma}_{ij}^{NLO}(x_1, x_2, Q^2, \mu_F, \mu_R) \right] \tag{2.39}$$

which means

$$\begin{aligned} \sigma^{NLO} &= \sum_{ij} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i^{\mathbf{h}_1}(x_1, \mu_F) f_j^{\mathbf{h}_2}(x_2, \mu_F) \left[\sum_{m=0}^{1} \int \mathrm{d}\hat{\sigma}_{ij}^{(m)}(x_1, x_2, Q^2, \mu_F, \mu_R) \alpha_s^{k+m}(\mu_R) \right] \\ &= \int \mathrm{d}\Phi_B [B_n(\Phi_B) + V_n(\Phi_B; \mu_F, \mu_R)] + \int \mathrm{d}\Phi_R [R(\Phi_R; \mu_F, \mu_R)]. \end{aligned} \tag{2.40}$$

The contributions appearing in the integrands consist of Born (B), Virtual (V) and Real (R) contributions. As discused in [25], B corresponds to the squared amplitude for a n-particle final state. The Virtual contribution V is an interference term that comes from combining the Born and the one-loop corrected amplitudes for a n-particle final state. And finally, R corresponds to the contribution that comes from the square amplitude for a tree level process of a $n+1$ final state particles. Figure 2.3 illustrates these terms for the $q\bar{q} \to q\bar{q}$ process in QCD.

In general, contributions at NLO can be written as

$$\begin{aligned} B(\Phi_B; \mu_F, \mu_R) &= \sum_h |\mathcal{M}_n^{(b)}(\Phi_B, h; \mu_F, \mu_R)|^2, \\ V(\Phi_B; \mu_F, \mu_R) &= 2 \sum_h \mathcal{R}e \left[\mathcal{M}_n^{(b)}(\Phi_B, h; \mu_F, \mu_R) \mathcal{M}_n^{*(b+1)}(\Phi_B, h; \mu_F, \mu_R) \right], \\ R(\Phi_R; \mu_F, \mu_R) &= 2 \sum_h |\mathcal{M}_{n+1}^{(b+1)}(\Phi_R, h; \mu_F, \mu_R)|^2 \end{aligned} \tag{2.41}$$

The corresponding phase-space at the Born and Real level is

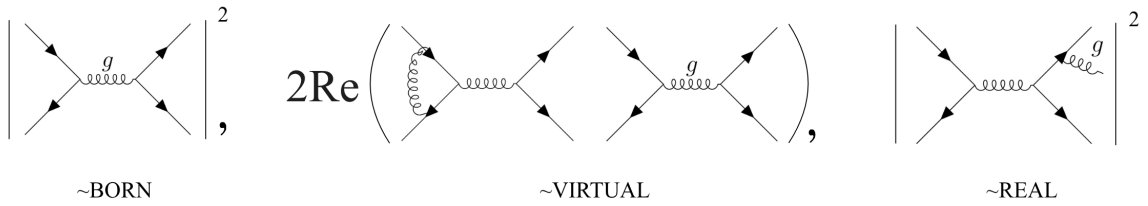


Figure 2.3: Terms that contribute to the cross-section at a NLO calculation for a process like $q\bar{q} \to q\bar{q}$. While the Born and Real contributions are obtained from squaring the amplitude matrix elements, the Virtual contribution is generated by an interference between the loop and the Born diagrams.

$$\begin{aligned} \mathrm{d}\Phi_B &= \mathrm{d}x_1 \mathrm{d}x_2 f_i^{\mathbf{h}_1}(x_2, \mu_F) f_j^{\mathbf{h}_2}(x_2, \mu_F) \frac{1}{2\hat{s}_{ab}} \mathrm{d}\Phi_n, \\ \mathrm{d}\Phi_R &= \mathrm{d}x_{1'} \mathrm{d}x_{2'} f_i^{\mathbf{h}_{1'}}(x_{2'}, \mu_F) f_j^{\mathbf{h}_{2'}}(x_{2'}, \mu_F) \frac{1}{2\hat{s}_{ab}} \mathrm{d}\Phi_{n+1}. \end{aligned} \tag{2.42}$$

The reason for the $x_{1'}$ and $x_{2'}$ is because the outgoing particle may affect the incoming particles considered.
Then, the cross section at fixed order can be written as

$$\sigma^{NLO} = \int \mathrm{d}\Phi_n [B(\Phi_n) + V_b(\Phi_n)] + \int \mathrm{d}\Phi_{n+1} R(\Phi_{n+1}). \tag{2.43}$$

Complications appear when evaluating the cross due to the presence of loops in the calculation. A loop introduces an integration variable that did not exist at Born level and it is not constrained by momentum conservation or on-shell conditions: the loop four-momentum l. As a result, l can become infinitely large, and calculations regarding d^4l/l^m for $m \leq 4$ will become infinite too. This phenomenon is known as **ultraviolet divergence** (UV). *Dimensional regularization* is often used to deal with these divergences, which consists in switching the integral dimension to $d = 4 - 2\varepsilon$, which guaranties gauge and Lorentz invariance. This allows us to express the divergence as poles in the dimensional regulator ϵ. After which the theory can be renormalized through a scheme-dependent redefinition of the Lagrangian [32]. In the following, we will assume that Virtual corrections have been renormalized

Alongside the UV divergences, infrared divergences (IR) also arise in the NLO con-

tributions, and they originate from loop momenta associated with massless particles either **soft**, i.e. with energy tending to zero, or **collinear**, i.e. emitted almost parallel to another particle. They will be discussed in more detail in Sect. (2.6).

$$\sigma^{NLO} = \int \mathrm{d}\Phi_n[B(\Phi_n) + V_b(\Phi_n)] + \int \mathrm{d}\Phi_n \mathrm{d}\Phi_r R(\Phi_n, \Phi_r), \tag{2.44}$$

after considering $\Phi_{n+1} = (\Phi_n, \Phi_r)$.

According to the Kinoshita-Lee-Nauenberg (KLN) theorem, the infrared divergences in both the Real and Virtual contributions cancel each other in sufficiently inclusive observables. The cancellation does not occur point by point since they live in different phase spaces. That is why subtraction methods, such as the Catani-Seymour method [33], which will be discussed in a later section, must be implemented. These methods introduce a counterterm that locally reproduces the singular behavior of the Real contribution and allows the cancellation before integration over the phase space.

2.5 Infrared divergences in the Virtual contribution

Beyond Leading Order (LO), theoretical predictions receive contributions from both Real and Virtual corrections. As discussed previously, these corrections introduce ultraviolet (UV) and infrared (IR) divergences that must be carefully handled to obtain physical results.

While UV divergences may be treated by regularization, where they will be translated into poles that are going to be absorbed in the redefinition of physical parameters, IR divergences that appear in the Virtual and in the Real contributions cancel each other when inclusive observables are considered.
In the Virtual corrections, due to particular regions of the *loop-momentum integration* namely when the loop momentum of the massless parton becomes soft or when the it becomes collinear to an external leg. The resulting singularities possess a universal structure and can be separated from the finite part of the amplitude.
The virtual contribution is determined by the interference term that arises in the squared matrix element,

$$|\mathcal{M}_n|^2 = |\mathcal{M}_n^{(0)}|^2 + 2\,\mathrm{Re}\left[\mathcal{M}_n^{(0)}\mathcal{M}_n^{(1)\,*}\right] + \mathcal{O}(\alpha_s^2), \tag{2.45}$$

where the full amplitude is expanded perturbatively at first order in α_s as

$$\mathcal{M}_n \approx \mathcal{M}_n^{(0)} + \mathcal{M}_n^{(1)}. \tag{2.46}$$

Here, $\mathcal{M}_n^{(0)}$ denotes the Born amplitude for a process with n final-state particles, while $\mathcal{M}_n^{(1)}$ corresponds to the one-loop correction to the same process (see Eq. (2.41)).
It has been shown in [34, 35] that the renormalized one-loop amplitude exhibits a universal factorization structure of the form

$$\mathcal{M}_n^{(b+1)} = \mathbf{I}^{(1)}\mathcal{M}_n^{(b)} + \mathcal{M}_{n,\text{fin}}^{(b+1)}, \tag{2.47}$$

The operator $\mathbf{I}^{(1)}$ contains the infrared singular structure of the renormalized one-loop amplitude. The form of $\mathbf{I}^{(1)}$ is determined by the momenta, the color charges and flavors of the external partons, while $\mathcal{M}_{n,\text{fin}}^{(1)}$ represents the finite part in the limit $\epsilon \to 0$.

$$\mathbf{I}^{(1)}(\epsilon, \mu^2; \{p\}) = \frac{1}{2}\frac{e^{-\epsilon\psi(1)}}{\Gamma(1-\epsilon)} \sum_i \frac{1}{\mathbf{T}_i^2} \mathcal{V}_i^{\text{sing}}(\epsilon) \sum_{j\neq i} \mathbf{T}_i \cdot \mathbf{T}_j \left(\frac{\mu^2 e^{-i\lambda_{ij}\pi}}{2p_i \cdot p_j}\right)^{\epsilon}. \tag{2.48}$$

The color correlations between external partons are encoded in the color-charge operators $\mathbf{T}_i \cdot \mathbf{T}_j$, while the kinematic dependence enters through $2p_i \cdot p_j$. The singular behavior is in

$$\mathcal{V}_i^{\text{sing}}(\epsilon) = \mathbf{T}_i^2 \frac{1}{\epsilon^2} + \gamma_i \frac{1}{\epsilon}, \tag{2.49}$$

which has both double and single poles on ϵ. The coefficients of these poles depend only on the flavor of the external partons,

$$\mathbf{T}_q^2 = \mathbf{T}_{\bar{q}}^2 = C_F, \qquad \mathbf{T}_g^2 = C_A, \tag{2.50}$$

and

$$\gamma_q = \gamma_{\bar{q}} = \frac{3}{2}C_F, \qquad \gamma_g = \frac{11}{6}C_A - \frac{2}{3}T_R N_f. \tag{2.51}$$

2.6 Infrared behavior in the Real contribution

To illustrate the origin of infrared divergences in the Real contribution, let us consider the emission of a gluon from a quark.

For Real contributions, we can consider a quark with four-momentum $p + k$ emitting a gluon whose four-momentum is k as illustrated in fig. 2.4.

$$q(p+k) \to g(k) + q(p). \tag{2.52}$$

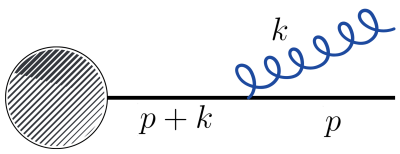


Figure 2.4: Gluon emission from a quark with four-momentum $p + k$.

The quark propagator obeys

$$\frac{i(\not{p} + \not{k} + m)}{(p+k)^2 - m^2}.$$

Where m is for the mass of the quark. For a massless emitter quark, $|\vec{p}| = E_p$ and $|\vec{k}| = E_k$ and the denominator in the propagator approximates as [30]

$$\frac{1}{(p+k)^2} = \frac{1}{2p \cdot k} = \frac{1}{2(E_p E_k - |\vec{p}||\vec{k}|)} = \frac{1}{2E_p E_k(1 - \cos\theta)}. \tag{2.53}$$

The structure of the equation (2.53) shows that there are two types of singular behavior. The first occurs in the limit $E_k \to 0$. This radiation is known as *soft radiation* and costs almost zero energy. The second arises when $\theta \to 0$, corresponding to the *collinear radiation*.

On the other hand, if the emitter is massive, the propagator is proportional to

$$\frac{1}{(p+k)^2 - m^2} = \frac{1}{p^2 + 2p \cdot k + k^2 - m^2} = \frac{1}{2p \cdot k}, \tag{2.54}$$

where $p^2 = m^2$ and $k^2 = 0$. Considering $p = (E_p, \vec{p})$ and $k = (E_k, \vec{k})$, then

$$p \cdot k = E_{\vec{p}} E_k - |\vec{p}| E_k \cos\theta. \tag{2.55}$$

By defining

$$\beta = \frac{|\vec{p}|}{E_p} = \sqrt{1 - \frac{m^2}{E_p^2}}, \tag{2.56}$$

we have

$$\frac{1}{(p+k)^2 - m^2} = \frac{1}{2E_p E_k(1 - \beta\cos\theta)}. \tag{2.57}$$

It is precisely the presence of the quark mass m that controls the divergence of the propagator through β. Due to $E_p^2 = |\vec{p}| + m^2$, $\beta < 1$. This means that even if the

angle $\theta \to 0$, the propagator cannot be singular. Although the collinear singularity is regulated by the quark mass, large logarithmic contributions may still arise when integrating probabilities over the collinear phase space.

The Real matrix element exhibits a universal behavior since the structure of these divergences does not depend on the specific process and depends only on the properties of the emitted radiation. That is, the dependence on the dynamics is contained in the underlying process without the additional emission $B(\Phi_n)$, while the description of the emission and the mass logarithms is encoded in a factor known as the *universal kernel*.

In the limits of soft and collinear QCD radiation we can find an approximation to the probability of splitting of one of the particles of a process.
The amplitude in the **soft limit**, where the extra parton is emitted with a momentum $k^\mu \to 0$, is

$$\mathcal{M}_{n+1}(p,k) \longrightarrow g_s \sum_i^n \mathbf{T}_i \frac{p_i^\mu}{p_i \cdot k} \mathcal{M}_n(p) \varepsilon_\mu(k), \tag{2.58}$$

the expression above can be generalized as

$$\mathcal{M}_{n+1} \longrightarrow J \mathcal{M}_n, \tag{2.59}$$

where

$$J^\mu = \sum_i^n \mathbf{T}_i \frac{p_i^\mu}{p_i \cdot k}, \tag{2.60}$$

is the so-called *Eikonal current*, which is described in terms of the color charge operator $\mathbf{T}_i$.
Then, the squared matrix element takes the form

$$|\mathcal{M}_{n+1}|^2 \approx \mu^{2\epsilon} g_s^2 \sum_i^n \mathbf{T}_i \mathbf{T}_j \frac{p_i \cdot p_j}{(p_i \cdot k)(p_j \cdot k)} |\mathcal{M}_n|^2, \tag{2.61}$$

with $\mathbf{T}_i$ being the color charge operator for the external partons [33]. Then, the approximated amplitude for a soft emission can be described by

$$R(\Phi_{n+1}) \approx B(\Phi_n) \left[\mu^{2\epsilon} g_s^2 \sum_{ij} \mathbf{T}_i \mathbf{T}_j \frac{p_i \cdot p_j}{(p_i \cdot k)(p_j \cdot k)} \right], \tag{2.62}$$

where the expression in the square brackets is often referred to as the *eikonal factor*.

In the **collinear limit** [36] where the radiated gluon is collinear to a parton, we can approximate the Real contributions as

$$R(\Phi_{n+1}) \approx B(\Phi_n) \left[\frac{2g_s^2}{s_{ij}} P(z) \right], \tag{2.63}$$

with $P(z)$ the Altarelli Parisi splitting functions in terms of z, the *momentum fraction* carried by one of the partons, while $s_{ij} = (p_i + p_j)^2$.

We can see that in these particular limits, the Real contribution can be approximated by the Born contribution times a factor that will describe the soft/collinear contribution, known as the *splitting kernel* ($K(\Phi_r)$). This is a crucial element in Parton Shower algorithms and in the subsequent interface with NLO calculations as Realized by the POWHEG method.

As an example, let us consider the process $e^+e^- \to q\bar{q}g$ which in the **collinear emission limit** exhibits the mentioned structure: A Born contribution, and a universal factor describing the emission

$$|\mathcal{M}_{e^+e^-\to q\bar{q}g}|^2 \;\to\; |\mathcal{M}_{e^+e^-\to q\bar{q}}|^2 \, C_F g^2 \, \frac{1}{p_1 \cdot k} \, \frac{1+z^2}{1-z}. \tag{2.64}$$

A similar behavior is exhibited in the **soft limit**,

$$|\mathcal{M}_{e^+e^-\to q\bar{q}g}|^2 \;\to\; |\mathcal{M}_{e^+e^-\to q\bar{q}}|^2 \, C_F g^2 \, \frac{2p_1 \cdot p_2}{(p_1 \cdot k)(p_2 \cdot k)}. \tag{2.65}$$

In both equations, the universal structure already in the soft and collinear limits mentioned above becomes evident: *Real contribution* $\sim$ *Born* $\otimes$ *splitting kernel*. This structure will be key to the introduction of algorithms that help us describe the large number of emissions originating from these limits, which will be discussed later in Section 2.7.2.

It is also important to emphasize that, after integration over the radiative phase space, the singular behavior associated with these infrared limits gives rise to logarithmically enhanced contributions:

Using the phase-space factorization equation (2.21), with Virtual momentum $\mathcal{P}^2 = (p_i + k_j)^2 = s_{ij}$. Combining this with the collinear factorization in the equation (2.63) then,

$$|\mathcal{M}_{n+1}|^2 \approx |\mathcal{M}_n|^2 2g_s^2 P(z) \mathrm{d}\Phi_n \frac{\mathrm{d}s_{ij}}{s_{ij}} \mathrm{d}\Phi_{split}. \tag{2.66}$$

where $\mathrm{d}\Phi_{split}$ denotes the two-body phase space associated with the branching of the intermediate parton into the parent partons i and j.

Then, in terms of the integration variables of the radiated particle, the real contribution to the differential cross section is

$$\mathrm{d}\sigma_R \approx \mathrm{d}\sigma_B \, \frac{\alpha_s}{2\pi} P(z) \, \frac{\mathrm{d}q^2}{q^2} \, \mathrm{d}z \mathrm{d}\phi, \tag{2.67}$$

where q is the characteristic scale of the branching. So that the integration over a range of scales $q_0 < q < Q$, where q_0 denotes IR cutoff scale below which no further branchings are resolved and Q the hard scale of the process, produces terms of the form

$$\int_{Q_0^2}^{Q^2} \frac{\mathrm{d}q^2}{q^2} = \log\left(\frac{Q^2}{Q_0^2}\right). \tag{2.68}$$

Similar logarithmic enhancements come from the soft region of the phase space.
At higher perturbative orders, multiple unresolved emissions generate contributions proportional to

$$\alpha_s^n \log^m\left(\frac{Q^2}{q_0^2}\right), \qquad m \leq 2n. \tag{2.69}$$

2.7 Subtraction of divergences

In this section, we discuss how infrared divergences (IR) arising from Virtual and Real contributions are treated in next−to−leading order (NLO) calculations of observables. In particular we introduce a successful framework designed for this purpose known as the **Catani−Seymour** subtraction method [37].

2.7.1 Subtraction of IR divergences

As already explained, the NLO structure of a cross section calculation is

$$\sigma = \int_m \left(\mathrm{d}\sigma^B + \mathrm{d}\sigma^V\right) + \int_{m+1} \mathrm{d}\sigma^R, \tag{2.70}$$

On the other hand, IR divergences in the Real and in the Virtual contributions must compensate each other, due to the Kinoshita-Lee-Nauenberg (KLN) theorem. This theorem establishes that IR divergences must cancel once all indistinguishable contributions to a process are taken into account [38]. To make explicit this cancellation, a subtraction scheme must be applied. To this end, a term that reproduces the singular

IR behavior of the Real emission contribution is introduced to rearrange the NLO cross section,

$$\sigma^{NLO} = \int_n \mathrm{d}\Phi_n \left[B(\Phi_n) + V(\Phi_n) + \int \mathrm{d}\Phi_r A(\Phi_n \Phi_r) \right] + \int \mathrm{d}\Phi_n \Phi_r \left[R(\Phi_n \Phi_r) - A(\Phi_n \Phi_r) \right]. \tag{2.71}$$

For the Real contribution, $\mathrm{d}\Phi_n \mathrm{d}\Phi_r A(\Phi_m \mathrm{d}\Phi_r)$ behaves as a local counterterm for $\mathrm{d}\Phi_m \Phi_r R(\Phi_m \Phi_r)$. After dimensional regularization, poles can be treated by making $\epsilon \to 0$ under the integral sign corresponding to $\mathrm{d}\Phi_{n+1}$. The result is an 4-dimensional integral for the Real contribution.

The same subtracted term added in the Virtual part must be analytically integrable in the radiated−parton phase space, $\mathrm{d}\Phi_r$. This integration leads to ϵ poles under the integral sign corresponding to $\mathrm{d}\Phi_n$, that can be combined with the poles in the Virtual contribution, to cancel all the divergences. This results in a finite expression for $\epsilon \to 0$, and then in a numerically integrable contribution.

$$\sigma^{NLO} = \int \mathrm{d}\Phi_n \Big[B(\Phi_n) + V(\Phi_n) + \int \Phi_r A(\Phi_n) \Big]_{\epsilon=0} + \int \mathrm{d}\Phi_n \mathrm{d}\Phi_r \left[R(\Phi_n \Phi_r)_{\epsilon=0} - A(\Phi_n \Phi_r)_{\epsilon=0} \right]. \tag{2.72}$$

Using the known behavior of n+1 parton matrix elements in the soft and collinear limits (eq. (2.59)), where the IR singularities arise, one can implement a factorization formulae that allows the subtraction counterterm to be written as

$$A(\Phi_n) = \sum_{dipoles} |\mathcal{M}_n|^2 \otimes V_{dipole}, \tag{2.73}$$

.

This is known as the dipole factorization formula. Here, V_{dipole} are the *universal dipole kernels*, which depend on the emitter–spectator configuration, and reproduce the singular behavior of the Real contribution. Their form is universal, making it independent of the details of the underlying process.

$$\int \mathrm{d}\Phi_n \mathrm{d}\Phi_r A(\Phi_n) = \sum_{dipoles} \int \mathrm{d}\Phi_n |\mathcal{M}_B|^2 \otimes \int \mathrm{d}\Phi_r V_{dipole} = \int \mathrm{d}\Phi_n |M_B|^2 \otimes \mathbf{I}, \tag{2.74}$$

where $\mathbf{I}$ corresponds to the phase-space integral of the radiated particle in $d = 4 - 2\epsilon$

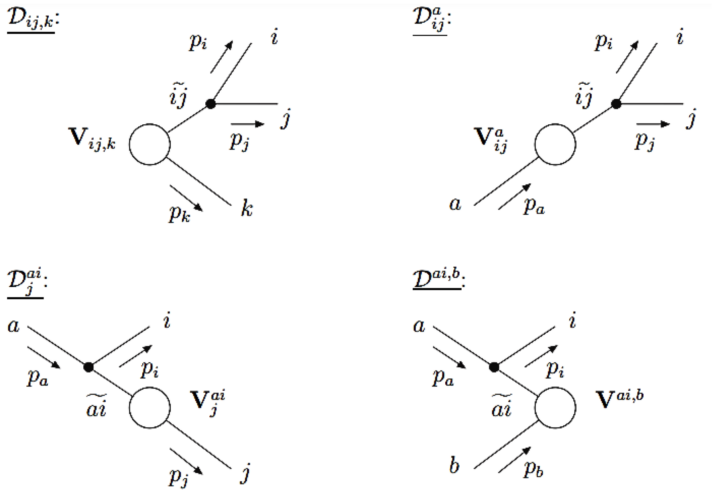


Figure 2.5: Diagrams for different emitter-spectator cases defining the dipoles $D_{ij,k}$.

dimensions,

$$\mathbf{I} = \sum_{dipoles} \int_r \mathrm{d}V_{dipole}. \tag{2.75}$$

Eq. (2.73) can be expressed in terms of the *dipoles* $D_{ij,k}$ as

$$\sum_{dipoles} D_{ij,k} = \sum_{dipoles} |\mathcal{M}_n|^2 \otimes V_{dipole}. \tag{2.76}$$

And the cross section,

$$\sigma^{NLO} = \int_n \mathrm{d}\Phi_n \left[B(\Phi_n) + V(\Phi_n) + I(\Phi_n)\right] + \int_{n+1} \mathrm{d}\Phi_{n+1} \left[R(\Phi_{n+1}) - \sum_{dipoles} D_{ij,k}\right]. \tag{2.77}$$

Then

$$D_{ij,k} = -\frac{1}{2p_i p_j} \langle \mathcal{M}_n | \frac{\mathbf{T}_k \cdot \mathbf{T}_{ij} V_{ij,k}}{\mathbf{T}^2_{ij}} | M_n \rangle , \tag{2.78}$$

which in the soft limit becomes [39],

$$D_{ij,k} = -8\pi\mu^{2\epsilon}\alpha_s \frac{1}{2p_i p_j} \langle \mathcal{M}_n | \frac{\mathbf{T}_k \cdot \mathbf{T}_{ij} P_{ij,k}}{\mathbf{T}^2_{ij}} | M_n \rangle , \tag{2.79}$$

being $V_{ij,k}$ the *universal kernel*, spin-dependent functions that can reproduced the

soft and collinear limits.
It is important to note that

$$D_{ij,k} \sim \frac{1}{p_i p_j}, \tag{2.80}$$

which reproduces the form of IR contributions in the soft and collinear limits (see e.g. Eq. (2.64) and Eq. (2.65).

The explicit form of the $V_{ij,k}$ depends on the subtraction scheme used. For instance, the Catani–Seymour method defines these dipole terms by constructing specific expressions for the $V_{ij,k}$ corresponding to the different kinematic configurations of the system.

- $q \to g(p_i) + q(p_j)$,
- $g \to q(p_i) + \bar{q}(p_j)$,
- $g \to g(p_i) + q(p_j)$.

that can be found in [37, 39].
Then the factorization of the $n+1$-particle final state can be approximated by implementing the dipole formalism times a splitting kernel

$$\sum_{\{\tilde{i},\tilde{k}\}\in\{\tilde{f}\}} \sum_{f_i} d\Phi_{n+1}\, D^{ij,k}_{n+1} = d\sigma^B_n \sum_{\{\tilde{i},\tilde{k}\}\in\{\tilde{f}\}} \sum_{f_i} dt\, dz\, \frac{1}{2p_i \cdot p_j} \frac{\alpha_s}{2\pi} P_{ij}(z). \tag{2.81}$$

where $\mathrm{d}\sigma^B_n$ where n-particle final state, where b_{ij} denotes the ratio of symmetry factors associated with the Born and real-emission configurations, while the remaining factors represent a differential emission kernel which, in the soft/collinear limit, reduces to a universal form governed by the DGLAP equations $P(z)$. Here, the differential probability of emission is,

$$\mathrm{dP} \sim \frac{\mathrm{d}t}{t} \mathrm{d}z \frac{\alpha_s}{2\pi} P(z). \tag{2.82}$$

Then, equation 2.81 provides an approximation of the contribution of $n+1$-particle configurations in the region where the radiated particle is soft and collinear.
Adding higher-order contributions by iterating this formula would violate unitarity, since each emission contribution is positive and increases the cross section of the process. To deal with this, it is necessary to include approximate Virtual contributions and unresolved emissions, that leads us to the introduction of the *Sudakov form factors*, which will be discussed in the next section.

2.7.2 Parton Shower

To obtain a more accurate description of kinematic distributions, it is necessary to account for additional QCD radiation that can modify the shape of the distributions of inclusive or exclusive observables, such as

$$\begin{aligned} &\frac{\mathrm{d}\sigma}{\mathrm{d}y_t}, \frac{\mathrm{d}\sigma}{\mathrm{d}p_T(t)}, \text{etc.} \\ &\frac{\mathrm{d}\sigma}{\mathrm{d}p_T(j_1)}, \frac{\mathrm{d}\sigma}{\mathrm{d}N_{jets}}, \text{etc.} \end{aligned} \tag{2.83}$$

A **Parton Shower Event Generator** (PS) produces, in a probabilistic way, a cascade of successive emissions of partons that are radiated from a highly energetic particle in the final state of a hard process in a simulated collision. This evolution takes place towards particles that are gradually softer and more collinear. This process is controlled through a variable that generates the successive emission. The process continues until an energy scale of the order of Λ_{QCD} is reached, scale at which α_s is not small anymore, the theory becomes non-perturbative, and confinement physics sets in.

Parton Shower event generators are designed to describe contributions in regions of phase space where convergence of calculations at fixed order is spoiled due to large logarithms that appear from integration over soft and collinear regions (see Sect. 2.5) of the Real emission. It does so by using the already known *universal structure* for the approximation of collinear and soft contributions and by **resumming** (see Apendix A.1) the logarithms associated with such contributions.
Parton Shower are therefore responsible for simulating the soft and collinear radiation by resumming the corresponding logarithmically enhanced contributions.

For this purpose, an ordering variable or scale t is used to guide the evolution of the Parton Shower. Transverse momentum, p_T, is often used as t, whose small values correspond to either soft or collinear emissions. The first emission generated by implementing this variable will be the hardest one, after which the system evolves towards emissions with progressively smaller p_T.

In this work, we will use the PYTHIA parton shower event generator [40] based on *transverse momentum ordering*.

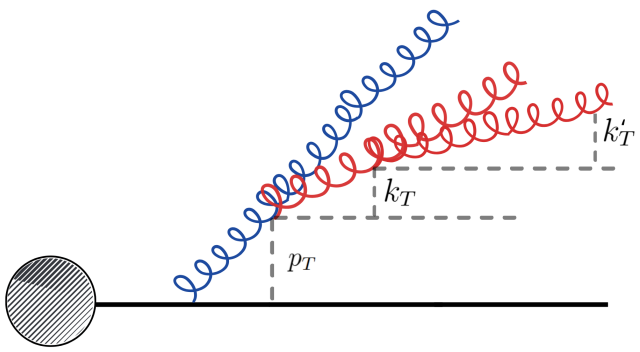


Figure 2.6: Ordered emission of gluons in a radiative process in a PS algorithm. p_T is for the transverse momentum of the hardest emission, while subsequent emissions must satisfy $k_T < p_T$.

In the following, a brief discussion on the impact of considering radiation to calculate the total cross section is presented.

For LO calculation, the cross section is

$$\sigma^{LO} = \int \mathrm{d}\Phi_n B(\Phi_n). \tag{2.84}$$

Within the PS framework, the cross section σ_{PS} is defined in terms of the shower *generating functional* $S(\mu_Q^2, \mathcal{O})$, following a procedure analogous to that described in [29]

$$\sigma_{PS} = \int \mathrm{d}\Phi_n B(\Phi_n) S_n(\mu_Q^2, \Phi_n, \mathcal{O}), \tag{2.85}$$

S is constructed by a recursive procedure that starts by first defining $\mathrm{P}(no{-}emision) \equiv \Delta(t_i, t)$ as the probability of no-emission during the evolution of the variable t in the interval (t_i, t), where t_i denotes the initial value of the shower evolution variable, typically of the order of Q^2, the characteristic hard squared scale of the process, until t. We also have $\Delta(t_i, t_f)$ corresponding to the probability of no radiation from t_i to the final evolution scale t_f at which the perturbative description ceases to be reliable. in the following, we briefly discuss the form of $\Delta(t_i, t)$ as in [41].
The probability density for the **first emission** to occur at the scale t is given by

$$\mathrm{dP} = \Delta(t_i, t) P(z) \frac{\alpha_s}{2\pi} \mathrm{d}t\, \mathrm{d}z, \tag{2.86}$$

i.e. given by the probability density of not having emitted between t_i and t, times the probability of emitting at a scale t given by the splitting kernel, where z specifies the momentum fraction carried by the daughter parton after the splitting. Therefore,

the probability for at least one emission to occur during the interval (t_i, t_f) is

$$\mathrm{P(emission)} \equiv \int_{t_f}^{t_i} \mathrm{d}t \int_0^1 \mathrm{d}z, \Delta(t_i, t) \frac{\alpha_s}{2\pi} P(z, t). \tag{2.87}$$

Then, by the unitarity of the total probability during the interval (t_i, t_f),

$$\begin{aligned} 1 &= \mathrm{P}(no - emission) + \mathrm{P}(emission), \\ 1 &= \Delta(t_i, t_f) + \int_{t_f}^{t_0} \mathrm{d}t \int_0^1 \mathrm{d}z \Delta(t_i, t) \frac{\alpha_s}{2\pi} P(z, t). \end{aligned} \tag{2.88}$$

From this equation, we obtain

$$\begin{aligned} \Delta(t_i, t_f) =& 1 - \int_{t_1}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \Delta(t_i, t_1) \frac{\alpha_s}{2\pi} P(z_1, t_1) \\ =& 1 - \int_{t_f}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \left[1 - \int_{t_1}^{t_i} \mathrm{d}t_2 \int_0^1 \mathrm{d}z_2 \Delta(t_i, t_2) \frac{\alpha_s}{2\pi} P(z_2, t_2) \right] \frac{\alpha_s}{2\pi} P(z_1, t_1) \\ =& 1 - \int_{t_f}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \Bigg[1 - \int_{t_1}^{t_i} \mathrm{d}t_2 \int_0^1 \mathrm{d}z_2 \left[1 - \int_{t_2}^{t_i} \mathrm{d}t_3 \int_0^1 \mathrm{d}z_3 \Delta(t_i, t_3) \frac{\alpha_s}{2\pi} P(z_3, t_3) \right] \\ & \times \frac{\alpha_s}{2\pi} P(z_2, t_2) \Bigg] \frac{\alpha_s}{2\pi} P(z_1, t_1). \end{aligned} \tag{2.89}$$

We can continue iterating on this process,

$$\begin{aligned} \Delta(t_i, t_f) = & 1 - \int_{t_f}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \frac{\alpha_s}{2\pi} P(z_1, t_1) \\ & + \int_{t_f}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \int_{t_1}^{t_i} \mathrm{d}t_2 \int_0^1 \mathrm{d}z_2 \left(\frac{\alpha_s}{2\pi} P(z_2, t_2) \frac{\alpha_s}{2\pi} P(z_1, t_1) \right) \\ & - \int_{t_f}^{t_i} \mathrm{d}t_1 \int_0^1 \mathrm{d}z_1 \int_{t_1}^{t_i} \mathrm{d}t_2 \int_0^1 \mathrm{d}z_2 \int_{t_2}^{t_i} \mathrm{d}t_3 \int_0^1 \mathrm{d}z_3 \left(\frac{\alpha_s}{2\pi} P(z_3, t_3) 2\pi P(z_2, t_2) \frac{\alpha_s}{2\pi} P(z_1, t_1) \right) \\ & + \dots . \end{aligned} \tag{2.90}$$

And by using simplex integration

$$\int_{t_f}^{t_i}\int_{t_1}^{t_i}\int_{t_2}^{t_i}...\int_{t_n}^{t_i} = \frac{1}{n!}\left(\int_{t_i}^{t_f} \mathrm{d}t\right)^n, \tag{2.91}$$

we can express eq. (2.90) as

$$\Delta(t_i,t_f) = 1 - \sum_1^\infty \left[\frac{1}{n!}\left(-\int_{t_f}^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P(z,t)\right)^n\right]$$

$$\Delta(t_i,t_f) = \sum_0^\infty \left[\frac{1}{n!}\left(-\int_{t_f}^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P(z,t)\right)^n\right]. \tag{2.92}$$

For $\Delta(t_i,t)$ in exponential form,

$$\Delta(t_i,t) = \exp\left(-\int_t^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P(z,t)\right), \tag{2.93}$$

which is known as the **Sudakov Form Factor**. If p_T labels the transverse momentum of the first emission, and k_T the transverse momentum of subsequent emissions, the contributions may be separated in regions where $p_T > k_T$ and $p_T < k_T$ by introducing a heaviside function, $1 = \Theta(k_T(z,t) - p_T) + \Theta(p_T - k_T(z,t))$.

$$\begin{aligned}\Delta(t_i,t) = &\exp\left(-\int_t^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P(z,t)\Theta[k_T(z,t)-p_T]\right)\\ &\times\exp\left(-\int_t^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P(z,t)\Theta[p_T-k_T(z,t)]\right).\end{aligned} \tag{2.94}$$

The first factor in Eq. (2.94) corresponds to emissions with transverse momentum satisfying $p_T < k_T$, while the second emissions satisfying $k_T > p_T$. The second line in equation (2.94) represent the probability of no emission harder than p_T.

$$\Delta(t_i,t) = \exp\left(-\int_t^{t_i}\mathrm{d}t\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}\mathrm{P}(z,t)\Theta[p_T-k_T(z,t)]\right). \tag{2.95}$$

This is the Sudakov form factor implemented in parton-shower algorithms based on transverse-momentum ordering.

The generating functional of the function (2.85) is

$$S_n(t_i) = \Delta(t_i,t) + \sum_{i=1}^N \int_t^{t_i}\mathrm{d}t'\int_0^1 \mathrm{d}z\frac{\alpha_s}{2\pi}P_i(z,t')\Delta(t_i,t')\times S_{n+1}(t'). \tag{2.96}$$

Then, the equation (2.85) can be rewritten as

$$\sigma^{PS} = \int \mathrm{d}\Phi_n B(\Phi_n) \left[\Delta(t_0, t_f) + \sum_{i=1}^{N} \int_t^{t_0} \mathrm{d}t \int_0^1 \mathrm{d}z \frac{\alpha_s}{2\pi} P_i(z,t) \Delta(t_i, t) S_{n+1} \right]. \quad (2.97)$$

which describes the parton shower.

The first term in the last equation describes the LO Cross Section times the probability of no-emitting any particle (described by the $\Delta(t_0, t)$)in the interval $[t_0, t]$. That is, it is the probabilistically regulated Born term. The second term in the cross section describes the probability that an emission with p_T has occurred after $t < t_i$. In addition, $\frac{\alpha_s}{2\pi} P_i$ is the splitting kernel of the process, which describes the probability of emission precisely at that point in the shower [9, 42]. They are given by

$$\begin{aligned} P_{qq}(z) &= C_F \left[\frac{1+z^2}{[1-z]} + \frac{3}{2} \right], \\ P_{gq}(z) &= C_F \left[\frac{1+(1-z)^2}{z} \right], \\ P_{qg}(z) &= T_F \left[z^2 + (1-z)^2 \right], \\ P_{gg}(z) &= 2C_A \left[\frac{z}{[1-z]} + \frac{1-z}{z} + z(1-z) \right], \end{aligned} \quad (2.98)$$

where the constant values appearing in the equations are $C_F = 4/3$, $T_F = 1/2$, $C_A = 3$

2.8 Next-to-leading order calculations plus Parton Shower matching

As discussed earlier, **Parton Shower algorithms** are implemented to model the contribution to observables in collision-physics from a radiative cascade originating from a hard event. Parton shower algorithms are based on approximations that provide a reliable description of emissions in the soft and collinear regions of phase space. In an ordered sequence of emissions, the contribution associated with the hardest radiation is less accurate than its estimation through fixed order calculations. This motivates the search for a technique that allows the improvement of the contribution of at least the first emission, for example, by evaluating the first radiation not through parton shower approximation, but entirely by a fixed-order calculation at NLO, while

allowing to combine the results with the subsequent radiation estimations through the Parton Shower.

In perturbation theory, fixed-order calculations describe well regions for the distribution of observables related to high-transverse momentum and large emission angles, such as the jet spectrum and hard radiation. That is, they accurately describe results in hard regions of the phase space.

But in the soft and collinear regime of radiation, fixed-order calculations fail to give reliable results due to the logarithmic contributions, $\alpha_s^n \log^{2n}(Q^2/p_T^2)$ which become very large and spoil the convergence of the series. In fact, these contributions appear precisely because fixed-order calculations are a truncation of the perturbative expansion.
On the other hand, as we already discussed in Sect. (2.7.2), Parton Shower algorithms and resummation techniques deal with such divergences. These algorithms resumm (reorganize) the expansion in order to treat these logarithmic contributions to all orders. So, Parton Shower algorithms are necessary to describe the behavior of soft and collinear emissions and to simulate the evolution of the cascade and hadronization.

Due to the complementary advantages of both frameworks, combining next-to-leading order calculations and Parton Shower into a unified approach becomes fundamental for achieving an improved simulation of QCD results at the LHC, capable of accurately describing both the soft and hard regions of phase space.

Therefore, with the continuous refinement and improvement of results in particle physics, the consistent combination of matrix elements with resummation effects from Parton Shower has become a fundamental objective in recent years, particularly for improving the description of observables that are sensitive to soft and collinear regions of phase space.

But achieving this combination is non-trivial, because one must ensure that the accuracy of the fixed-order contribution is preserved in the hardest regions, while complementing the calculation through the implementation of the Sudakov form factor, the core of the Parton Shower algorithms [25].

Special care must be taken when matching a Parton Shower algorithms to a next-to-leading order calculation. If a PS is naively combined with a NLO calculation, a problem of double counting will arise, particularly at the level of the first emission. This is because, at NLO, the Real contribution is already included (in both the hard and soft/collinear emissions). On the other hand, when a PS shower is later applied

it will generate its own first approximated emission. This process is illustrated in Figure 2.7. The consequence is that the Real emission-phase space in the soft and collinear regions will be populated twice.
In this context, matching schemes were developed to combine fixed-order calculations with parton showers while avoiding double counting:

Matching: The expression of the Parton Shower evaluated at fixed order is computed and subtracted from higher order calculations in order to eliminate double counting. The resulting expression is later processed by the Parton Shower [29].

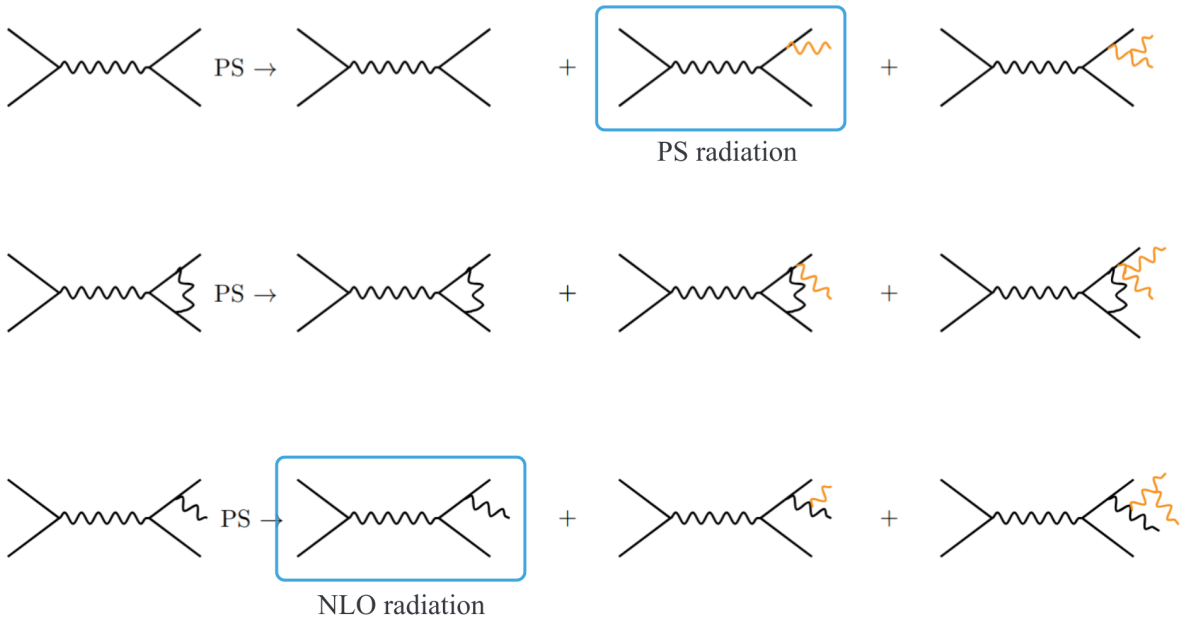


Figure 2.7: Schematic representation of the combination of diagrams involved in a next-to-leading order calculation with a Parton Shower. The fixed-order computation already includes an exact Real emission contribution (radiated particles in black). The Parton Shower generates an approximated radiative emission (radiation marked in yellow). A simultaneous inclusion of both contributions (highlighted in the blue boxes) leads to a double counting effect, which spoils the accuracy of the calculations. Then, a special matching scheme is required to deal with this overlap effect

.

To achieve the generation of positive weights, a procedure is defined that will be discussed in the following section.

For this purpose, the approximation of the splitting kernel, $\mathcal{K}(\Phi_1)$, in the parton shower algorithms with its exact counterpart is implemented

$$R(\Phi_{n+1}) \approx B(\Phi_n) \otimes \mathcal{K}(\Phi_1). \tag{2.99}$$

Recalling that

$$\hat{\sigma}^{NLO} = \int \mathrm{d}\Phi_n \left[B(\Phi_n) + V_b(\Phi_n)\right] + \int \mathrm{d}\Phi_{n+1} R(\Phi_{n+1}),$$

where $V_b(\Phi_n)$ denotes the regularized virtual contribution. From now on, this finite quantity will be denoted by $V(\Phi_n)$. Using the fact that the $n+1$ Lorentz invariant phase space, Φ_{n+1}, can be parametrized in terms of the Born phase space, Φ_n, and the phase space of the radiated particle, Φ_r, we have

$$\int \mathrm{d}\Phi_{n+1} = \int \mathrm{d}\Phi_n \int \mathrm{d}\Phi_r. \tag{2.100}$$

After considering that process, we have

$$\lim_{p_T \to 0} |M_{n+1}|^2 \approx \frac{\alpha_s}{2\pi} \mathrm{P}_i |M_n|^2. \tag{2.101}$$

We can therefore approximate this in terms of the differential cross section of the Real emission and Born contributions as

$$\frac{\alpha_s}{2\pi} \mathrm{P}_i \to \frac{R(\Phi_n + 1)}{B(\Phi_n)}. \tag{2.102}$$

Applying eq. (2.100) and eq. (2.102) to (2.97) we obtain *Modified Sudakov Form Factor*,

$$\Delta(t_0, t) \longrightarrow \Delta_M(\Phi_n, p_T) = \exp\left(-\int \mathrm{d}\Phi_r \frac{R(\Phi_{n+1})}{B(\Phi_n)} \Theta(k_T - p_T)\right). \tag{2.103}$$

In this way, by incorporating matrix elements from NLO calculations, a first way to correct parton shower is obtained

$$\sigma^{PS(NLO)} = \int \mathrm{d}\Phi_n B(\Phi_n) \left(\Delta_M(\Phi_n, p_T^{min}) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{n+1})}{B(\Phi_n)} \Delta_M(\Phi_n, p_T)\right). \tag{2.104}$$

The implementation of this approach has practical limitations, particularly the fact that events can be generated with negative weights. To counteract this effect, it is necessary to generate a larger number of events. This will be discussed in greater detail in the following section.

2.9 POWHEG

Positive Weight Hardest Emission Generator
Is a matching scheme in which the hardest radiation is generated first using matrix elements computed at NLO accuracy.

This framework aims to avoid negative weights by adding the Virtual and Real corrections as follows:

$$\overline{B}(\Phi_n) = B(\Phi_n) + V(\Phi_n) + \int \mathrm{d}\Phi_r R(\Phi_n, \Phi_r). \tag{2.105}$$

Then, the cross section is

$$\sigma_{PW} = \int \mathrm{d}\Phi_n \overline{B}(\Phi_n) \left(\Delta_M(\Phi_n, p_T^{min}) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{n+1})}{B(\Phi_n)} \Delta_M(\Phi_n, p_T) \right), \tag{2.106}$$

and in general, for an observable $\mathcal{O}_{PW}$

$$\langle \mathcal{O}_{PW} \rangle = \int \mathrm{d}\Phi_m \bar{B}(\Phi_m) \left[\Delta_M(\Phi_m, p_T^{min}) \mathcal{O}_m(\Phi_m) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m, p_T) \mathcal{O}_{m+1}(\Phi_{m+1}) \right], \tag{2.107}$$

In the POWHEG method, the first emission is generated implementing the *Modified Sudakov* constructed using R/B, which guaranties a correct description of the collinear and soft regions of the phase space. However it is important to notice that the naive use of the Sudakov in all regions of the radiation phase space introduces resummation effects that are suitable for handling divergences in the soft region but that are not justified in the hard region, and actually these effects may lead to undesired effects in, for example, emissions of high-transverse momentum, which can distort the predictions in those regions. Solution to this issue is discussed in the next section, through the introduction of a *Damping Function*.

2.10 Damping function in POWHEG

As mentioned before, it is necessary to split the Real contributions in R_s (singular, which contains IR divergences) and R_h (hard, which contains the finite contributions). This is because the Sudakov form factor, through resummation, correctly describes only the singular (soft/collinear) regions of the phase space, rendering finite and accurate predictions due to the implementation of R/B. However, in the semi-singular

region, its application is not completely justified, and in the hard region,it may even introduce distortions to the predictions.

The split of R can be done by following the procedure mentioned in [43] and in [44], i.e. by writing

$$\sigma_{PW} = \int \mathrm{d}\Phi_n \overline{B}_s(\Phi_n) \left(\Delta_{M_s}(\Phi_n, p_T^{min}) + \int \mathrm{d}\Phi_r \frac{R_s(\Phi_{n+1})}{B(\Phi_n)} \Delta_{M_s}(\Phi_n, p_T) \right) + R_h(\Phi_{n+1}) \mathrm{d}\Phi_{n+1}, \tag{2.108}$$

where now in terms of the singular part

$$\overline{B}_s(\Phi_n) = B(\Phi_n) + V(\Phi_n) + \int \mathrm{d}\Phi_r R_s(\Phi_n, \Phi_r) \tag{2.109}$$

and

$$\Delta_{M_s}(\Phi_n, p_T) = \exp\left(- \int \mathrm{d}\Phi_r \frac{R_s(\Phi_{n+1})}{B(\Phi_n)} \Theta(k_T - p_T) \right). \tag{2.110}$$

The split of the contributions may be done by defining a *Splitting Function* F such that

$$R_s(\Phi_{n+1}) \equiv F(\Phi_{n+1}) R(\Phi_{n+1}),$$

$$R_h(\Phi_{n+1}) \equiv [1 - F(\Phi_{n+1})] R(\Phi_{n+1}),$$

$$R = R_s + R_h. \tag{2.111}$$

The value of the function F varies across the phase space and determines how the Real contribution is separated in its singular and hard components. By default, the POWHEG-BOX sets $F = 1$, corresponding to $R_s = R$ and $R_h = 0$, which means that the entire real part is treated as singular contribution. If a non-trivial form for F is chosen, allowing it to take values in the interval $[0, 1]$, it allows the Real matrix element to be separated into a singular contribution and a finite part.

$$\begin{aligned} F(\Phi_{n+1}) &\to 1 \quad \text{singular regions,} \\ F(\Phi_{n+1}) &\to 0 \quad \text{hardest regions,} \\ F(\Phi_{n+1}) &\in (0, 1) \quad \text{other regions.} \end{aligned}$$

The introduction of this continuous function, guarantees a smooth split of the Real contributions in eq. (2.108).

In POWHEG-BOX, $F(\Phi_{n+1})$ can be decomposed as the product of two functions,

$$F(\Phi_{n+1}) = F_{Bornzero} F_{damp}, \tag{2.112}$$

where,

$$F_{damp} = \frac{h_{damp}^2}{h_{damp}^2 + p_T^2}, \tag{2.113}$$

and

$$F_{Bornzero}(\Phi_{n+1}) = \Theta\left(h_{Bornzero} - \frac{R(\Phi_{n+1})}{P_{ij}(\Phi) \otimes B(\Phi_n)}\right). \tag{2.114}$$

Here, a dynamical definition of this function has been introduced by making it dependent on the transverse momentum of the radiated particle. Here, h_{damp} is an arbitrary energy scale, and $h_{bornzero}$ is an arbitrary factor.

One way to see this is that F is composed of two filters. The first is given by F_{damp} which will act as a purely kinematical filter in terms of p_T. Then, when the radiation is $p_T \to 0$, $F_{damp} \approx 1$. But if $p_T \gg h_{damp}$, then $F_{damp} \approx 0$, and this will push $F \approx 0$. On the other hand, $F_{Bornzero}$ is designed to compare Real contribution with IR contribution, where $P_{ij} \otimes B \approx R$ which is the expected behavior if the splitting is within QCD. In that case, $F_{bornzero} = 1$. In that sense, $h_{bornzero}$ controls how far outside the limits of strictly soft/collinear radiation we can associate the calculations $F_{bornzero}$ blocks regions where $R(\Phi_{n+1})$ is anomalously big in which case $F_{bornzero} = 0$ and the event is not introduced to the Sudakov.

2.11 Accuracy of POWHEG

The formula in the POWHEG eq. (2.107) has an accuracy NLO. This can be demonstrated in a simple way by expanding the exponential of the Sudakov form factor following the procedure presented in [45]

$$\langle \mathcal{O}_{PWG} \rangle = \int \mathrm{d}\Phi_m \bar{B}(\Phi_m) \left[\Delta_M(\Phi_m, p_T^{min}) \mathcal{O}_m(\Phi_m) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m, p_T) \mathcal{O}_{m+1}(\Phi_{m+1})\right], \tag{2.115}$$

adding a convenient zero,

$$\begin{aligned}\langle \mathcal{O}_{PWG} \rangle = &\int \mathrm{d}\Phi_m \bar{B}(\Phi_m) \left[\Delta_M(\Phi_m, p_T^{min}) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m,\, p_T) \right] \mathcal{O}_m(\Phi_m) \\ &+ \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m, p_T) \left[\mathcal{O}_{m+1}(\Phi_{m+1}) - \mathcal{O}_m(\Phi_m) \right],\end{aligned} \tag{2.116}$$

and taking into account that $k_T > p_T^{min}$, we can introduce a Dirac delta function in the second term of eq. (2.116)

$$\int_{p_T^{min}}^{\infty} \mathrm{d}k_T \delta(p_T - k_T) \;=\; 1, \tag{2.117}$$

and obtain

$$\int_{p_T^{min}}^{\infty} \mathrm{d}k_T \delta(p_T - k_T) \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m,\, p_T). \tag{2.118}$$

Let us recall that the Sudakov form factor has the form

$$\Delta_M(\Phi_m,\, p_T) = -\exp\left(\int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Theta(k_T - p_T) \right),$$

then,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}k_T} \Delta_M(\Phi_m, p_T) &= \Delta_M(\Phi_m, p_T) \left(\frac{\mathrm{d}}{\mathrm{d}k_T} \left[-\int \mathrm{d}\Phi_r \frac{R}{B} \Theta(k_T - p_T) \right] \right) \\ &= \Delta_M(\Phi_m, p_T) \left(\left[-\int \mathrm{d}\Phi_r \frac{R}{B} \delta(k_T - p_T) \right] \right).\end{aligned} \tag{2.119}$$

Substituting this into eq. (2.118), we obtain

$$\int_{p_T^{min}}^{\infty} \mathrm{d}k_T \frac{\mathrm{d}}{\mathrm{d}k_T} \Delta_M(\Phi_m, p_T) = \Delta_M(\Phi_m, p_T)|_{p_T \to \infty} - \Delta_M(\Phi_m, p_T)|_{p_T \to p_T^{min}}. \tag{2.120}$$

and

$$\int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m,\, p_T) = 1 - \Delta_M(\Phi_m, p_T^{min}), \tag{2.121}$$

such that the eq. (2.116) takes the form

$$\langle \mathcal{O}_{PWG} \rangle = \int \mathrm{d}\Phi_m \bar{B}(\Phi_m) \left[\mathcal{O}_m(\Phi_m) + \int \mathrm{d}\Phi_r \frac{R(\Phi_{m+1})}{B(\Phi_m)} \Delta_M(\Phi_m, p_T) \left[\mathcal{O}_{m+1}(\Phi_{m+1}) - \mathcal{O}_m(\Phi_m) \right] \right]. \tag{2.122}$$

From the expansion

$$\Delta_M(\Phi_m, p_T) = 1 + \mathcal{O}(\alpha_s),$$

and recalling

$$\bar{B}(\Phi_m) = B(\Phi_m) + V(\Phi_m) + \int \mathrm{d}\Phi_r R(\Phi_{m+1}),$$

the eq. (2.122) will contain terms proportional to different powers of α_s, namely,

- $B\frac{R}{B} \sim \alpha_s$,
- $V\frac{R}{B} \sim \alpha_s^2$,
- $R\frac{R}{B} \sim \alpha_s^2$,
- $B\frac{R}{B}\mathcal{O}(\alpha_s) > \alpha_s^2$,
- $V\frac{R}{B}\mathcal{O}(\alpha_s) > \alpha^3$,
- $R\frac{R}{B}\mathcal{O}(\alpha_s) > \alpha_s^3$.

Retaining terms up to order α_s and taking into account the collinear approximation of the phase space

$$\langle \mathcal{O}_{PWG} \rangle \approx \int \mathrm{d}\Phi_m \left[B(\Phi_m) + V(\Phi_m) \right] \mathcal{O}_m(\Phi_m) + \int \mathrm{d}\Phi_{m+1} R(\Phi_{m+1}) \mathcal{O}_{m+1}(\Phi_{m+1}). \tag{2.123}$$

It means that 2.116 has the accuracy of a NLO calculation.

2.12 Event generation

2.12.1 Monte Carlo

In particle physics, computational methods are often implemented to solve, for example, complicated integrals. Some of those methods are based on the generation of random numbers distributed according to a given probability density [46].

In probability theory, the *expectation value* of a function $f(X)$ of a random variable X is obtained by averaging f over all possible outcomes weighted by the probability associated to the random variable. This can be expressed as

$$E[f(X)] = \int f(x)\mathrm{d}P(x). \tag{2.124}$$

Here, the differential *probability measure* associated to the random variable X is $\mathrm{d}P(x) = p(x)\mathrm{d}x$ where $p(x)$ is the corresponding probability density function. Applying this result to the integral

$$I = \int_\Omega f(x)\mathrm{d}x = \int_\Omega \frac{f(x)}{p(x)}p(x)\mathrm{d}x, \tag{2.125}$$

where $f(x) \geq 0$.
From eq. (2.124) and eq.(2.125)

$$E\left[\frac{f(X)}{p(X)}\right] = \int_\Omega \frac{f(x)}{p(x)}p(x)\mathrm{d}x. \tag{2.126}$$

In the **hit-and-miss** Monte Carlo method, an integral can be estimated by generating random points over the integration domain and counting the fraction of points that fall below the integrand.
In that sense, for an integral,

$$I = \int_a^b f(x)\,\mathrm{d}x, \qquad\qquad 0 \leq f(x) \leq M.$$

where M is a bound for $f(x)$, we generate a uniform distribution of points X_i and Y_i in the region bounded by

$$\begin{array}{ll} (a, b) & x \text{ coordinate,} \\ (0, M) & y \text{ coordinate.} \end{array}$$

We can consider a point as a *hit* if Y_i of the coordinate (X_i, Y_i) is under the curve generated by $f(X_i)$, and a *miss* otherwise.
After several trials, the fraction of accepted points estimates the integral in the region

$$I = (b-a)M \cdot P(Y \leq f(X)). \tag{2.127}$$

2.12.2 Veto Algorithm

In the process of generating a parton shower, the *Sudakov* is interpreted as the probability of no radiation to occur above a certain scale, which in POWHEG is set as

the *transverse momentum* p_T of the radiated particles. To generate the first emission, an algorithm known as *the Sudakov Veto Algorithm* is constructed based on this probability and the density of the subsequent branching.

In the POWHEG method, this emission is generated at NLO accuracy by modifying, among other aspects in the calculation, the Sudakov used in the Parton Shower. Once the first radiation is generated, the rest of the radiation is left to the shower generator. POWHEG-BOX first generates Φ_B with a probability proportional to $\bar{B}$, and then the hardest radiation variables Φ_r via the veto, where a limit that corresponds to the hardness of the emission is set by this algorithm. The generated event is stored in the Les Houches Event (LHE) format, a standardized file format used to communicate parton-level events between matrix-element generators and parton-shower programs [47]. The evolution of the next radiations generated by the shower generator must meet this limit in order to avoid a double counting of the hardest radiation.

In the following, we provide a general overview of the veto algorithm, following the formulation presented in [48]. For a more detailed discussion, we refer the reader to Refs. [41, 48, 49].
Random points x (which can be associated to points Φ_r in the radiation phase space) can be generated according to

$$\begin{aligned} \mathcal{P} &= k(x)\Delta(h(x)) \\ &= k(x)\exp\left(-\int \mathrm{d}^{\mathrm{d}}x' k(x')\Theta(h(x')-h(x))\right), \end{aligned} \tag{2.128}$$

here, $h(x)$ denotes the ordering variable associated with that point (in the context of Parton Shower algorithms it corresponds to p_T), $k(x)$ plays the role of the branching kernel, it corresponds to the probability distribution that the first emission occurs at the hard scale $h(x)$. while $\Delta(h)$ would be the probability of evolving down to the scale h without producing a harder emission (associated to Parton shower algorithms, it corresponds to the Sudakov form factor).

Then, in POWHEG,

$$\mathcal{P} = \frac{R}{B}\exp\left(-\int \mathrm{d}\Phi_r \frac{R}{B}\Theta(k_T'-k_T)\right). \tag{2.129}$$

It is the *probability distribution for the first emission to occur in the phase-space point* x.

To generate events distributed according to eq. (2.129), it is necessary to solve the integral appearing in the exponent of the exponential. One of the problems is that the integral appearing in the exponent cannot be evaluated analytically; moreover, direct sampling from the probability density of Eq. (2.129) is impractica. To overcome this problem, the Sudakov form factor can be used to construct an equivalent sampling procedure. The key observation is that the probability density for the first emission to occur at a hardness scale h is obtained by differentiating the no-emission probability $\Delta(h)$. This leads to

$$\frac{\mathrm{d}\Delta(h(x))}{\mathrm{d}h} = \int \mathrm{d}^{\mathrm{d}}x k(x)\delta(h(x)-h)\exp\left(-\int \mathrm{d}^{\mathrm{d}}x' k(x')\Theta(h(x')-h)\right). \tag{2.130}$$

The transformed variable scale $r = \Delta(h)$ is uniformly distributed in $[0,1]$. To obtain the hardness H a random number r is generated, and $\Delta(H) = r$ is solved for H. Once H is fixed, variables Φ_r are distributed accordingly to

$$k(x)\delta(h(x)-H)\exp\left(-\int \mathrm{d}^{\mathrm{d}}x' k(x')\Theta(h(x')-H)\right), \tag{2.131}$$

where $\exp\left(-\int \mathrm{d}^{\mathrm{d}}x' k(x')\Theta(h(x')-H)\right)$ now depends on H that has been determined, and there is no dependence on x. It is important to notice that x is now just conditioned by the factor $k(x)\delta(h(x)-H)$.

Through a function $K(x)$ that overestimates $k(x)$, and ensuring it is chosen to be simpler than k and that it can be invertible and integrated, its corresponding Sudakov is constructed,

$$\Delta_K(h(x)) = \exp\left(-\int \mathrm{d}^{\mathrm{d}}x' K(x')\Theta(h(x')-H)\right), \tag{2.132}$$

such that its probability distribution is

$$\mathcal{P}_K(x) = K(x)\Delta_K(h(x)). \tag{2.133}$$

In this way, H_{max} is established such that $\Delta_K(H_{max}) = 1$ which fixes the maximum scale that the process can take,

$$r = \frac{\Delta_K(H)}{\Delta_K(H_{max})} = \Delta_K(H). \tag{2.134}$$

Moreover, a random number q is generated $q \in [0,1]$ and it is accepted if $qK(x) < k(x)$. If it is accepted, the event is stored and $x \to \Phi_r$ is generated by $x \sim K(x)\delta(H-$

$h(x)$). If the event is not accepted, H is set as the new H_{max} y and the process is repeated.

2.12.3 Event generation in POWHEG

Event generation using the POWHEG method can be schematically described as follows [48]:

First step: Total cross section

To this end, we must compute the quantity

$$\bar{B}(\Phi_B) = B(\Phi_B) + V(\Phi_B) + \int \mathrm{d}\Phi_r R(\Phi_B, \Phi_r), \tag{2.135}$$

which depends only on the Born kinematics, Φ_B.

Among the three terms that contribute to $\bar{B}(\Phi_B)$:

- $B(\Phi_B)$ corresponds to the Born term and is evaluated analytically using Born kinematics.
- $V(\Phi_B)$ represents the Virtual one-loop contribution. This term is also evaluated analytically in Born kinematics.
- The last term contains the integral

  $$\int \mathrm{d}\Phi_r R(\Phi_m, \Phi_r),$$

  where the integration is performed over the radiation phase space $\Phi_{rad} = (k_T, z, \phi)$. This is the contribution that is evaluated numerically within the POWHEG algorithm.

In this way, we have defined $\bar{B}(\Phi_m)$. Using this quantity, the total cross section can be estimated through Monte Carlo integration

$$\sigma_{\bar{m}} = \int d\Phi_m \, \bar{B}(\Phi_m) = \int d\Phi_m \, g(\Phi_m) \, \frac{\bar{B}(\Phi_m)}{g(\Phi_m)} \approx \frac{1}{N} \sum_{i=1}^{N} \frac{\bar{B}(\Phi_m^i)}{g(\Phi_m^i)}, \tag{2.136}$$

where, in the last line, the points Φ_m^i are generated according to an importance function $g(\Phi_m)$.

This function is introduced in Monte Carlo methods to improve the efficiency of the integration estimate by replacing the uniform generation of points in phase space with one weighted by g. This function should approximate the shape of $\bar{B}(\Phi_m)$. This

method increases the sampling in regions of phase space that give larger contributions to the cross-section calculation, while at the same time reducing the statistical variance of our estimates.

Second step: Event generation at born level

Once the total cross section has been estimated, events distributed according to the differential cross section $\bar{B}(\Phi_B)$ are generated. To achieve this, an acceptance–rejection algorithm is used. First, the maximum value $\bar{B}_{max}$ of the function $\bar{B}(\Phi_B)$ is determined. Then candidate points Φ_m are generated in the Born phase space together with a random number r uniformly distributed in the interval $[0, 1]$.
The point Φ_m is accepted if the following condition is satisfied

$$r\bar{B}_{max} < \bar{B}(\Phi_m). \tag{2.137}$$

Geometrically, the generated point is accpeted whenever $r\bar{B}^{max}$ falls below the $\bar{B}(\Phi_m)$ distribution (see figure 2.8). If the condition is not satisfied, the point is rejected and a new candidate is generated.

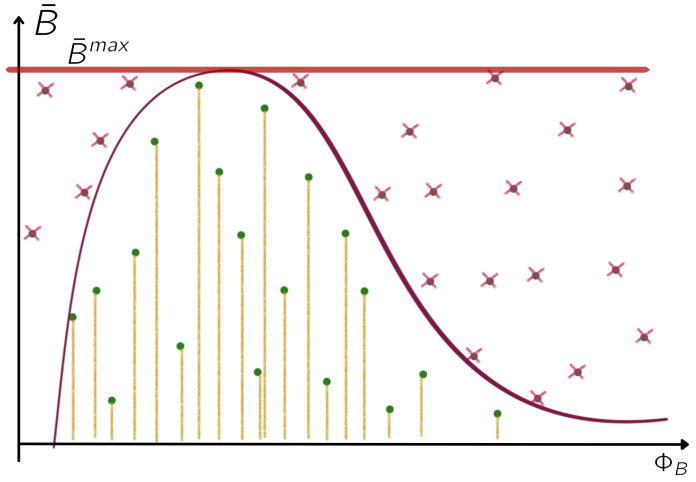


Figure 2.8: Schematic illustration of the acceptance-regection procedure implemented to generate phase space points Φ_m. For a given random r, if $r\bar{B}(\Phi_m)$ lies below the curve $\bar{B}(\Phi_m)$, the corresponding phase-space point Φ_m is accpeted; otherwise, rejected.

The accepted points correspond to events that now obey the form of the Born distribution $\bar{B}(\Phi_B)$.

Third step: Event with radiation

In the last step, Born vents Φ_m have been already generated. Now, the next step consists in determining whether an additional radiation is generated. In POWHEG, the probability distribution for the hardest emission is given by

$$\Delta_M(\Phi_m, p_T^{min}) + \int \mathrm{d}\Phi_r \frac{R(\Phi_m \Phi_r)}{B(\Phi_m)} \Delta_M(\Phi_m, p_T^{min}) \tag{2.138}$$

Let us recall that within the Parton Shower formalism, the Sudakov form factor in the first term represents the probability of evolving from a given hard scale dow to a cutoff p_T^{min} without any resolvable emission, while the second term represents the probability of of generating the hardest emission occured at p_T.
To generate the radiation variables corresponding to the radiated particle, a Monte Carlo method is implemented again. For a given Φ_m, a random number $r \in [0, 1]$ is generated, and

$$\Delta_M(\Phi_m, p_T^{min}) = r \tag{2.139}$$

must be solved for p_T. Once the transverse momentum of the emission has been determined, the remaining radiation variables are generated according to the corresponding phase-space parametrization.
Among all candidate emissions, the one with the largest transverse momentum is selected according to the Highest-Bid procedure. If no resolvable emission is generated, corresponding to $p_T = 0$, the event remains at Born level.

3 Details of the calculation

We investigate the parton shower effects in the production of $t\bar{t}WW$ at the LHC, with the goal of improving the theoretical description of this process as a relevant background to four-top quark production. We aim to study the impact of the parton shower modeling on kinematic observables. We compare different Monte Carlo event generators for the fully inclusive signatures and at the fiducial level.

This study provides insight into the interaction between fixed-order calculations and parton shower modeling, and the associated theoretical uncertainties.

3.1 Computational Setup

For the computational setup, we follow a similar procedure as in Ref. [44].
We compute next-to-leading order (NLO) QCD corrections at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV (at the LHC) for the process $pp \to t\bar{t}W^+W^-$. The input parameter of the Standard Model are chosen according to Ref. [50] to be

$$\begin{aligned} M_W &= 80.3692\,\text{GeV}\,, & M_Z &= 91.1880\,\text{GeV}\,, & M_H &= 125.2\,\text{GeV}\,, \\ m_t &= 172.57\,\text{GeV}\,, & m_b &= 4.78\,\text{GeV}\,, & & \\ \Gamma_Z &= 2.4955\,\text{GeV} & \Gamma_H &= 3.7 \cdot 10^{-4}\,\text{GeV}\,. & & \end{aligned} \tag{3.1}$$

While the Fermi-constant is given by

$$G_F = 1.663788 \cdot 10^{-5}\,\text{GeV}^{-2}\,, \tag{3.2}$$

which determines the electroweak coupling to be

$$\alpha = \frac{\sqrt{2}}{\pi}\, G_F\, M_W^2 \left(1 - \frac{M_W^2}{M_Z^2}\right) = \frac{1}{132.1}\,. \tag{3.3}$$

For the parton distribution functions and the one-loop running of strong coupling, we are employing the `NNPDF40_nlo_as_01180_nf_4` (334700) PDF set [51], which is

interfaced via LHAPDF library [52].

3.2 Implementation

For the implementation of a new process in the POWHEG-BOX framework, it is necessary to provide some components that allow us to construct a NLO calculation such as the Born, Virtual and Real contributions through the corresponding matrix elements, as well as a parametrization of the phase space at born level. In this work, the matrix elements have been obtained from OpenLoops [53–55], which are then given to the POWHEG-BOX framework [43,48,56] that is responsible for generating the hardest emission in the process primarily through the Modified Sudakov. Once this is obtained, the event is written in Les Houches format and subsequently passed to the parton shower (PYTHIA) which is responsible for generating the subsequent emissions.

In order to validate the implementation in POWHEG, a comparison between LO+PS and NLO+PS predictions is performed, which allows a direct analysis of the impact of higher-order corrections. Subsequently, to validate the NLO+PS implementation, the results obtained are compared with predictions generated using MG5_aMC@NLO [20,57], which implements the MC@NLO matching scheme [45,58]. This comparison allows, among other things, the confrontation of uncertainties associated with the matching between fixed-order calculations and the parton shower.

In LO+PS vs NLO+PS and POWHEG vs MG5 comparisons, the theoretical uncertainties associated with the renormalization and factorization scales are evaluated by implementing variations within the seven-point scheme. This allows the quantification of the residual dependence of the perturbative calculation.

Using matrix elements from OpenLoops [53–55] we have developed an event generator in the POWHEG-BOX framework [43,48,56]. We compare this to predictions obtained using aMC@NLO_MG5 [20,57], which employs the MC@NLO matching [45,58].

3.3 Diagrams

Some diagrams corresponding to the $t\bar{t}WW$ production are shown below in Figures (Figs. 3.1, 3.2, and 3.3) to illustrate the types of topologies corresponding to the Born, Virtual, and Real contributions. These diagrams are generated using MG5_aMC@NLO.

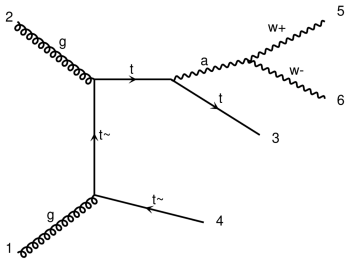


Figure 3.1: One of the Feynman diagrams contributing to $t\bar{t}WW$ production at the Born level is shown. It corresponds to a gluon fusion process involving two QED-type vertices and two QCD-type vertices.

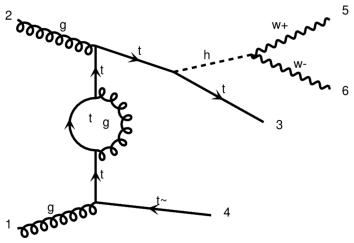


Figure 3.2: One of the Feynman diagrams, at virtual level, contributing to $t\bar{t}WW$. It results from a gluon-fusion process involving two QED-type vertices and four QCD-type vertices.

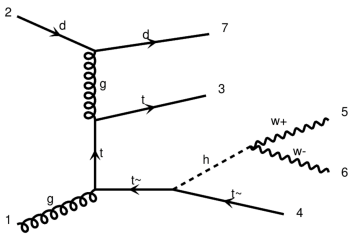


Figure 3.3: Above, a Feynman diagram contributing to the real level production of o $t\bar{t}WW$. This diagram in particular results from gd interaction, involving two QED-type vertices and three QCD-type vertices. The radiated particle is a quark which carries a color charge.

3.3.1 Processes and subprocesses

The subprocesses contributing to the NLO QCD calculations to $pp \to t\bar{t}WW$ are shown in Table 3.1. The calculation is performed under the four-flavor scheme (4FS) [59], where u, d, s, c quarks and gluons contribute as initial state partons.

Directory (partonic process)	Type	# Diagrams	# Subproc.	Subprocesses (partonic channels)
$d\bar{d}$ and $s\bar{s}$ ($d\bar{d},\ s\bar{s} \to t\bar{t}W^+W^-$)	born	16	2	$d\bar{d} \to t\bar{t}W^+W^-,\quad s\bar{s} \to t\bar{t}W^+W^-$
	virt	305	2	$d\bar{d} \to t\bar{t}W^+W^-,\quad s\bar{s} \to t\bar{t}W^+W^-$
	real	102	2	$dg \to t\bar{t}W^+W^-d,\quad sg \to t\bar{t}W^+W^-s$
	real	102	2	$d\bar{d} \to t\bar{t}W^+W^-g,\quad s\bar{s} \to t\bar{t}W^+W^-g$
	real	102	2	$gd \to t\bar{t}W^+W^-d,\quad gs \to t\bar{t}W^+W^-\bar{s}$
$\bar{d}d$ and $\bar{s}s$ ($\bar{d}d,\ \bar{s}s \to t\bar{t}W^+W^-$)	born	16	2	$d\bar{d} \to t\bar{t}W^+W^-,\quad \bar{s}s \to t\bar{t}W^+W^-$
	virt	305	2	$d\bar{d} \to t\bar{t}W^+W^-,\quad \bar{s}s \to t\bar{t}W^+W^-$
	real	102	2	$dg \to t\bar{t}W^+W^-d,\quad \bar{s}g \to t\bar{t}W^+W^-\bar{s}$
	real	102	2	$d\bar{d} \to t\bar{t}W^+W^-g,\quad \bar{s}s \to t\bar{t}W^+W^-g$
	real	102	2	$gd \to t\bar{t}W^+W^-d,\quad gs \to t\bar{t}W^+W^-s$
gg ($gg \to t\bar{t}W^+W^-$)	born	39	1	$gg \to t\bar{t}W^+W^-$
	virt	969	1	$gg \to t\bar{t}W^+W^-$
	real	102	2	$dg \to t\bar{t}W^+W^-d,\quad sg \to t\bar{t}W^+W^-s$
	real	102	2	$ug \to t\bar{t}W^+W^-u,\quad cg \to t\bar{t}W^+W^-c$
	real	258	1	$gg \to t\bar{t}W^+W^-g$
	real	102	2	$\bar{d}g \to t\bar{t}W^+W^-\bar{d},\quad \bar{s}g \to t\bar{t}W^+W^-\bar{s}$
	real	102	2	$gd \to t\bar{t}W^+W^-d,\quad g\bar{s} \to t\bar{t}W^+W^-\bar{s}$
	real	102	2	$gd \to t\bar{t}W^+W^-d,\quad gs \to t\bar{t}W^+W^-s$
	real	102	2	$g\bar{u} \to t\bar{t}W^+W^-\bar{u},\quad g\bar{c} \to t\bar{t}W^+W^-\bar{c}$
	real	102	2	$gu \to t\bar{t}W^+W^-u,\quad gc \to t\bar{t}W^+W^-c$
	real	102	2	$\bar{u}g \to t\bar{t}W^+W^-\bar{u},\quad \bar{c}g \to t\bar{t}W^+W^-\bar{c}$
$u\bar{u}$ and $c\bar{c}$ ($u\bar{u},\ c\bar{c} \to t\bar{t}W^+W^-$)	born	16	2	$u\bar{u} \to t\bar{t}W^+W^-,\quad c\bar{c} \to t\bar{t}W^+W^-$
	virt	305	2	$u\bar{u} \to t\bar{t}W^+W^-,\quad c\bar{c} \to t\bar{t}W^+W^-$
	real	102	2	$ug \to t\bar{t}W^+W^-u,\quad cg \to t\bar{t}W^+W^-c$
	real	102	2	$u\bar{u} \to t\bar{t}W^+W^-g,\quad c\bar{c} \to t\bar{t}W^+W^-g$
	real	102	2	$g\bar{u} \to t\bar{t}W^+W^-\bar{u},\quad g\bar{c} \to t\bar{t}W^+W^-\bar{c}$
$\bar{u}u$ and $\bar{c}c$ ($\bar{u}u,\ \bar{c}c \to t\bar{t}W^+W^-$)	born	16	2	$\bar{u}u \to t\bar{t}W^+W^-,\quad \bar{c}c \to t\bar{t}W^+W^-$
	virt	305	2	$\bar{u}u \to t\bar{t}W^+W^-,\quad \bar{c}c \to t\bar{t}W^+W^-$
	real	102	2	$\bar{u}g \to t\bar{t}W^+W^-\bar{u},\quad \bar{c}g \to t\bar{t}W^+W^-\bar{c}$
	real	102	2	$\bar{u}u \to t\bar{t}W^+W^-g,\quad \bar{c}c \to t\bar{t}W^+W^-g$
	real	102	2	$gu \to t\bar{t}W^+W^-u,\quad gc \to t\bar{t}W^+W^-c$

Table 3.1: Born, virtual and real subprocesses contributing to $pp \to t\bar{t}W^+W^-$ in the POWHEG-BOX implementation.

A total of 4590 diagrams are considered for this process, which correspond to 103 Born, 2189 Virtual, and 2298 Real diagrams.

POWHEG and MadGraph weights

Table 3.2 shows the fraction of positive- and negative-weight events obtained in the generation of $t\bar{t}WW$ events with POWHEG and MadGraph5_aMC@NLO.

Generator	**Positive weights**	**Negative weights**
POWHEG	99.995%	0.005%
MadGraph5_aMC@NLO	77.63%	22.37%

Table 3.2: Comparison of positive and negative event weights in POWHEG and MadGraph5_aMC@NLO.

POWHEG is designed to generate almost entirely positive weights, unlike MC@NLO, which produces a considerable fraction of negative weights. For the $t\bar{t}WW$ process, the POWHEG sample contains (99.995%) positive weights, whereas the MadGraph5_aMC@NLO sample contains (77.630%) positive weights. Since the total cross section is obtained through the sum of the weights of all generated events, both positive and negative weights must be taken into account.

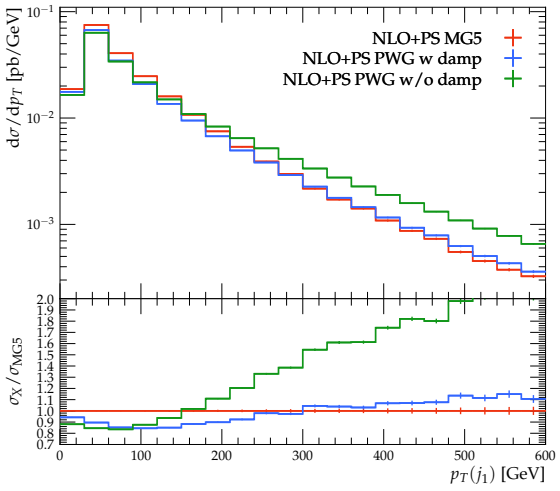


Figure 3.4: An example of results when the damping function is activated in POWHEG

3.3.2 Total cross sections and differential distributions

As a first step, we examine two different choices of renormalization and factorization scales,

$$\mu_{\rm fix} = m_t + m_W \ , \qquad \mu_{\rm dyn} = \frac{H_T}{2} \equiv \frac{1}{2} \sum_{i \in \text{ final state}} E_T(i) \ , \tag{3.4}$$

where the transverse energy is given by $E_T(i) = \sqrt{p_T^2(i) + m_i^2}$. First we study the scale dependence of the total cross section. For this, we perform computations for $\mu_R = \mu_F = \xi\ \mu_0$, where μ_0 is either the fixed or the dynamical scale.

Next, we proceed to study the differential distributions and their residual theoretical uncertainties. To this end, we adopt the dynamical scale choice, $\mu_R = \mu_F = \mu_0 = \mu_{\rm dyn}$. Theoretical uncertainties associated with missing higher-oreder corrections are estimated using the 7-point envelope

$$\left(\frac{\mu_R}{\mu_0}, \frac{\mu_F}{\mu_0}\right) = \Big\{(0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (1, 2), (2, 1), (2, 2)\Big\} \tag{3.5}$$

The damping function reduces deviation of the POWHEG calculations. In the POWHEG-BOX framework we use the following damping parameters

$$h_{\rm damp} = \frac{H_T}{2} \ , \qquad h_{\rm bornzero} = 5 \ , \tag{3.6}$$

the impact of the damping functions can be visualized through the Fig. (3.4).

and study their related theoretical uncertainties through the following variations

$$(h_{\text{damp}}, h_{\text{bornzero}}) = \left\{ \left(\frac{H_T}{2}, 5\right), \left(\frac{H_T}{2}, 2\right), \left(\frac{H_T}{2}, 10\right), \left(\frac{H_T}{4}, 5\right), (H_T, 5) \right\} . \quad (3.7)$$

For aMC@NLO_MG5 the equivalent parameter would be the initial shower scale μ_Q, which we choose to be $\mu_Q = H_T/2$. Uncertainties related to this choice are obtained by considering the envelope of

$$\mu_Q = \left\{ \frac{H_T}{4}, \frac{H_T}{2}, H_T \right\} . \quad (3.8)$$

We study the distributions for

- Top-quark observables,
- W-boson observables,
- Hadronic distributions

3.3.3 Jet formation

After the hard event is generated by POWHEG and the parton shower is simulated with PYTHIA, the resulting stable final-state particles are clustered into jets using the algorithm of anti-k_T, which combines particles that are close in the pseudorapidity-azimutal angle space, and clustering soft particles around hard ones. To define a fiducial region (a region of phase space defined by kinematic cuts to match the detector acceptance and the experimentally accessible phase space) of analysis, only jets satisfying the kinematic requirements $p_T^j > p_{T,\text{min}}^j, |\eta_j| < \eta_{\text{max}}^j$ are used.
Similar requirements are imposed on the charged leptons in the final state, ensuring that all reconstructed objects lie within the fiducial acceptance of the analysis.

4 Results

Results at fixed-order matched to parton showers (PS) are presented here using the POWHEG-BOX framework for the process $pp \to t\bar{t}WW$ with a center of mass energy $\sqrt{s} = 13.6$ TeV at the LHC, and are compared to predictions made with MG5_aMC@NLO. The differential cross-section distributions are presented for several observables (inclusive and exclusive).

4.1 Top-Quark Observables

4.1.1 Transverse momentum distribution for the hardest top quark in the events ($p_T(t_1)$)

Fig. (4.1) shows the transverse momentum distribution of the hardest top quark in the event. This observable is considered inclusive since it depends only on the kinematics of the hardest top quark and does not require a fixed number of extra particles in the final state.

Left panel

For the fixed-order calculation + PS within the POWHEG-BOX the LO+PS (red band) and NLO + PS (blue band) were generated in order to compare the contributions arising from the inclusion of one additional order. The differential cross section $\frac{\mathrm{d}\sigma}{\mathrm{d}p_T}$ (upper panel) represents cross section as a function of the transverse momentum of the hardest top quark, $p_T(t_1)$. The vertical axis is presented on a logarithmic scale. It is observed that the inclusion of one additional perturbative order leads to an enhancement of the distribution, with the NLO+PS prediction lying above the LO+PS result over the entire transverse momentum range.
Both distributions decrease along the p_T axis, as expected since high-p_T emissions are less likely to occur due to the large momentum transfers required.

The lower panel shows the ration between NLO+PS and LO+PS as a function of

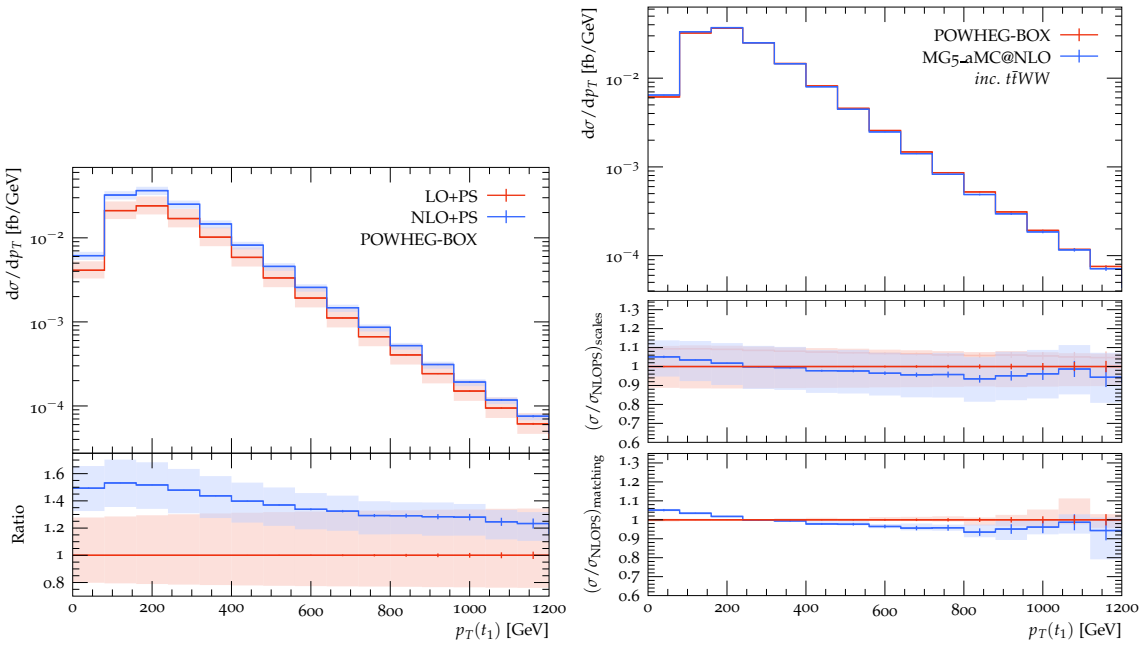


Figure 4.1: Inclusive $t\bar{t}WW$ QCD production as a function of the transverse momentum of the hardest top-quark. In the left-hand panel (l.h.p.), a comparison between calculations an LO+PS and NLO+PS in POWHEG are shown, including their respective theoretical uncertainties. In the right-hand panel (r.h.p) a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

$p_T(t_1)$. We see that an increase between $\sim$ 20% and $\sim$ 50% is introduced by the NLO+PS which is more pronounced in the low-p_T region. It means that the NLO contribution, does not only correct the underestimated LO prediction but also redistributes events across the p_T space.

This behavior reflects the impact of higher-order corrections, which not only modify the normalization of the cross section but also improve its perturbative accuracy.

The uncertainty bands associated with renormalization and factorization scale variations are shown as shaded regions in the plots. The inclusion of NLO corrections reduces the residual scale dependence, leading to an improvement in the perturbative stability of the prediction. The corresponding uncertainty is reduced from approximately 20% at LO+PS to about 10% at NLO+PS.

Right panel

The comparison between the POWHEG (red curve) and MG5_aMC@NLO (blue curve), both at NLO+PS, is shown.

In addition, the middle and low panels show the ratio of $\sigma_{\text{MG5}}/\sigma_{\text{POWHEG}}$ with the corresponding theoretical uncertainty bands arising from scale variations (middle panel),

and from the matching scheme (bottom panel), computed using the seven-point envelope prescription.
A good agreement between both frameworks is observed in the distribution of events as a function of the transverse momentum of the hardest top quark. This can be interpreted as this observable is lowly sensitive to the specific matching scheme employed (POWHEG vs MC@NLO).

4.1.2 Pseudorapidity distribution of the hardest top quark, $\eta(t_1)$

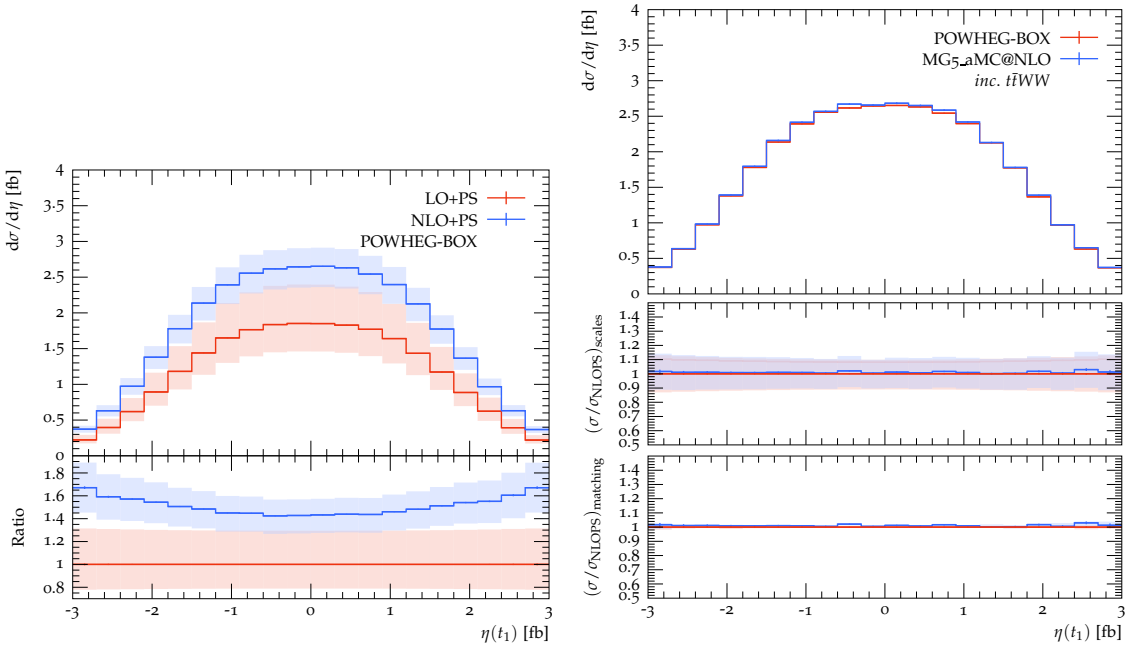


Figure 4.2: Inclusive $t\bar{t}WW$ QCD production as a function of the rapidity of the hardest top-quark. In the left-hand panel (l.h.p.), a comparison between calculations an LO+PS and NLO+PS in POWHEG are shown, including their respective theoretical uncertainties. In the right-hand panel (r.h.p) a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Results for $\eta(t_1)$ are shown in Figure (4.2). Once again the left panel presents a comparison between the LO+PS and the NLO+PS distributions both generated in POWHEG, while the right panel shows a comparison between the predictions at NLO+PS generated with POWHEG and MG5_aMC@NLO.

Left panel

The LO+PS (in red) and NLO+PS (in blue) distributions for the rapidity of the hardest top are both generated with POWHEG. In the upper panel, it is observed that the distribution is predominantly concentrated in the central region, around $\eta(t_1) = 0$, which indicates that the top quarks are mostly produced with small longitudinal boosts along the beam axis (z) . The distribution decreases systematically for rapidities away from zero. In order to produce highly forward top quarks, a strong

asymmetry between the initial-state momentum fractions x_1 and x_2 is required, which is suppressed by the parton distribution functions.
The NLO prediction is systematically larger than the LO one, which indicates that the LO+PS calculation underestimates the distribution. The ratio generated for both results, where the LO prediction is normalized, reveals a non-uniform K-factor (see Apendix A.5) over the range of $\eta(t_1)$.
It can also be seen that the uncertainty bands are reduced at NLO to about 10%, (smaller than the LO bands of around 20%), meaning that for this observable, the perturbative uncertainty decreases when higher-order corrections are included, as expected.

Right panel
In the upper panel, the NLO+PS predictions obtained with POWHEG and MG5_aMC@NLO are overlaid. Both distributions are found to be in good agreement over the entire rapidity range with the expected results. The middle and lower panels display the ratio of the two predictions, normalized with respect to the POWHEG result. In the middle panel, the theoretical uncertainty bands associated with scale variations are shown for both frameworks, amounting to approximately 12% and being nearly constant across the full range, with a significant overlap between them. In the lower panel, the uncertainty bands associated with the matching scheme are displayed, which are found to be negligible, at the level of approximately 0%. This indicates that this observable is almost no sensitive to the matching for both POWHEG and MG5_aMC@NLO

4.1.3 Invariant mass of the top-quark pair, $M_{t\bar{t}}$

Fig. (4.3) presents the invariant-mass distribution of the top-quark pair. Since this observable is directly related to the kinematics of the hard scattering process, it provides a useful probe of the impact of higher-order corrections on both the normalization and shape of the distribution.

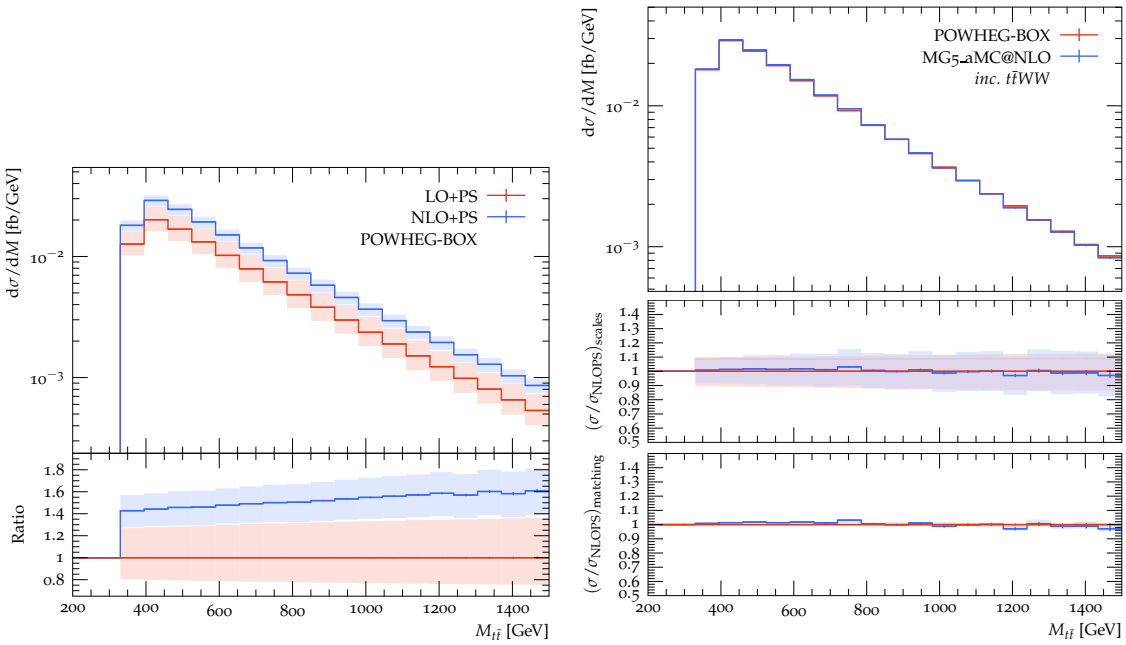


Figure 4.3: Inclusive $t\bar{t}WW$ QCD production as a function of the invariant mass of the top-quark pair. In the l.h.p., a comparison between distributions estimated at LO+PS and NLO+PS in POWHEG is shown, including their respective theoretical uncertainties. In the r.h.p a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Left panel

The LO+PS and NLO+PS results obtained with POWHEG are shown now for the invariant mass distribution of the $t\bar{t}$ pair. The distributions are shown in a logarithmic scale.

In the distribution there are no events contributing to the distributions below approximately 345 GeV, since it is not possible to produce a top-pair lighter than this value. The threshold starts at $M_{t\bar{t}} \approx 2m_t \approx 345$ GeV. Here, the tops are produced almost at rest with respect to their center of mass frame.

For $M_{t\bar{t}} \approx 2m_t \approx 345$ GeV the tops are produced far from relative rest and the system requires more energy in the center of mass to produce the event. The higher

energy required to produce the pair, the fewer events are generated. Therefore the distribution decreases as the invariant mass grows.

The NLO distribution is larger than the LO one over the entire invariant mass range, with differences that go from approximately 40% to 60%. It means that the NLO corrections enhance the LO prediction, and also changes the shape of the distribution. In the ratio, the NLO+PS contribution shows an increasing behavior over the whole range, which means that the correction becomes more relevant in the hard region of the $t\bar{t}$ invariant mass.
The shaded bands corresponds once again to the theoretical uncertainties estimated through the seven point envelope scheme.
At higher order in the calculation, its dependence on the renormalization and the factorization scales decreases, so the uncertainty band at NLO is reduced with respect to the LO one.

Right panel

In the right panel distributions for this observable generated both at NLO in the POWHEG-BOX (red line) and in MG5_aMC@NLO (blue line) are shown. They show a very good agreement, indicating that both provide consistent perturbative predictions. The corresponding scale variation ratio normalized to POWHEG remains close the unity over the entire $M_{t\bar{t}}$ range. This indicates that the dependence on the normalization and factorization scales is largely driven by the fixed-order calculation rather than by the matching scheme.

The matching ratio between POWHEG and MG5_aMC@NLO remains consistent with unity over the full range, indicating that this observable is weakly dependent on the NLO+PS matching scheme.

4.2 W-boson Observables

4.2.1 Transverse momentum of the hardest W-boson, $p_T(W_1)$

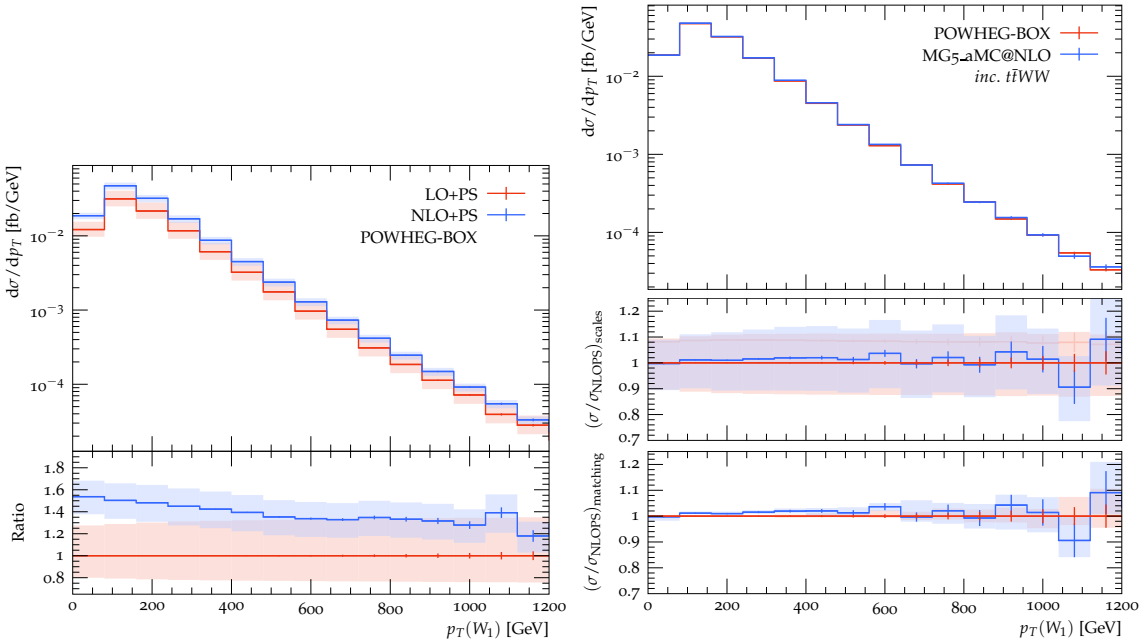


Figure 4.4: Inclusive $t\bar{t}WW$ QCD production as a function of the transverse momentum of the hardest W-boson. In the l.h.p., a comparison between distributions estimated at LO+PS and NLO+PS in POWHEG is shown, including their respective theoretical uncertainties. In the r.h.p a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Here, the transverse momentum distribution for the hardest W-boson in Fig. (4.4) is analyzed.

Left panel

Distributions for the transverse momentum $p_T(W_1)$ of the hardest W are shown for both LO+PS (red) and NLO+PS (blue) calculations in the POWHEG framework. The highest differences between the LO and the NLO predictions are of nearly 55%, while the lowest are of nearly 20%.
That means that the NLO calculations not only correct the underestimation of events made by the LO prediction, but also introduce a K-factor dependent on the transverse momentum that modifies the shape of the distribution.

The $p_T(W_1)$ distribution falls as the transverse momentum increases, since events at this scale require a higher energy configuration in the hard scattering, which are statistically suppressed.

Right panel

Distributions at NLO+PS in POWHEG and in MG5_aMC@NLO are shown. In the top panel a direct comparison is displayed, while the middle panel shows the ratio between MG5_aMC@NLO predictions compared to POWHEG calculations. The solid lines exhibit an excellent agreement between the two frameworks across the entire range.
In the middle panel, uncertainties associated to the scale variation, μ_R and μ_F are shown, which for POWHEG and MG5_aMC@NLO are consistent. This observable is weakly dependent on the scale variations employed in both frameworks.
Finally, in the bottom panel the ratio is shown, with information related to the matching of the frameworks. The uncertainty bands increase slightly as the transverse momentum increases, both for POWHEG and for MG5_aMC@NLO.
In addition, the corresponding uncertainty bands overlap throughout the spectrum. This indicates that the underlying physics of this observable is consistently described by both frameworks.

4.3 Hadronic Observables

4.3.1 Hadronic transverse energy, H_T

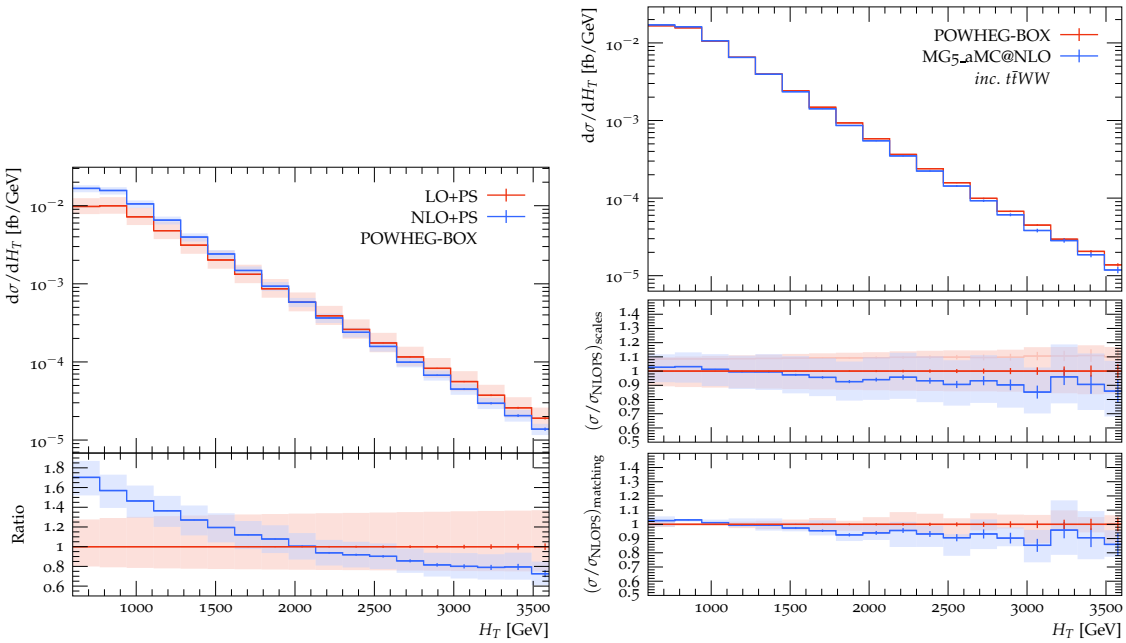


Figure 4.5: Inclusive $t\bar{t}WW$ QCD distribution in terms of H_T. In the l.h.p., a comparison between calculations an LO+PS and NLO+PS in POWHEG are shown, including their respective theoretical uncertainties. In the r.h.p a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Distributions for H_T of the system are shown in Figure (4.5).

Left panel

In this panel, NLO+PS (red) and NLO+PS (blue) predictions are displayed, both obtained with POWHEG. As expected, both distributions exhibit a decreasing behavior at larger H_T, reflecting that the events with large transverse *activity* require increasingly energetic configurations, which are kinematically suppresses.
Significant difference between the two predictions are observed. The ratio indicates that the K-factor is not constant over the H_T domain. In the region of high H_T, noticeable differences can be seen. However, the corresponding uncertainty band overlap, indicating a good agreement between the two perturbative orders within uncertainties.

On the contrary, in the low-H_T region, the LO+PS predictions underestimates the NLO+PS results and the uncertainty bands do not fully overlap, highlighting the importance of higher-order corrections in this region.

It is important to notice that the uncertainty band at NLO is reduced with respect the the LO one. This indicates an improvement in the theoretical predictions due to the inclusion of higher order effects.

Overall, this observable shows a strong dependence on the higher-order corrections suggesting that NLO calculations are essential to improve predictions and further improvements can be achieved by considering NNLO corrections.

Right panel

The comparison between POWHEG (in red) and MG5_aMC@NLO (blue) predictions at NLO+PS is shown. A good overall agreement is observed, although small differences appear in the domain of the high H_T. Also theoretical band are enhanced in the same region and indicate a moderate sensitivity to renormalization and factorization scales.

4.3.2 Number of inclusive jets, N_{jets}^{incl}

We can define the number of inclusive jets as the number of reconstructed jets in an event that contribute to the observable. The jet multiplicity is computed by counting all reconstructed jets that satisfy the analysis selection criteria for fiducial studies (see section (3.3.3)), and the inclusive jet distribution is built as the cumulative distribution of N or more jets. Each bin includes the contribution of events with N or more jets to the cross section.

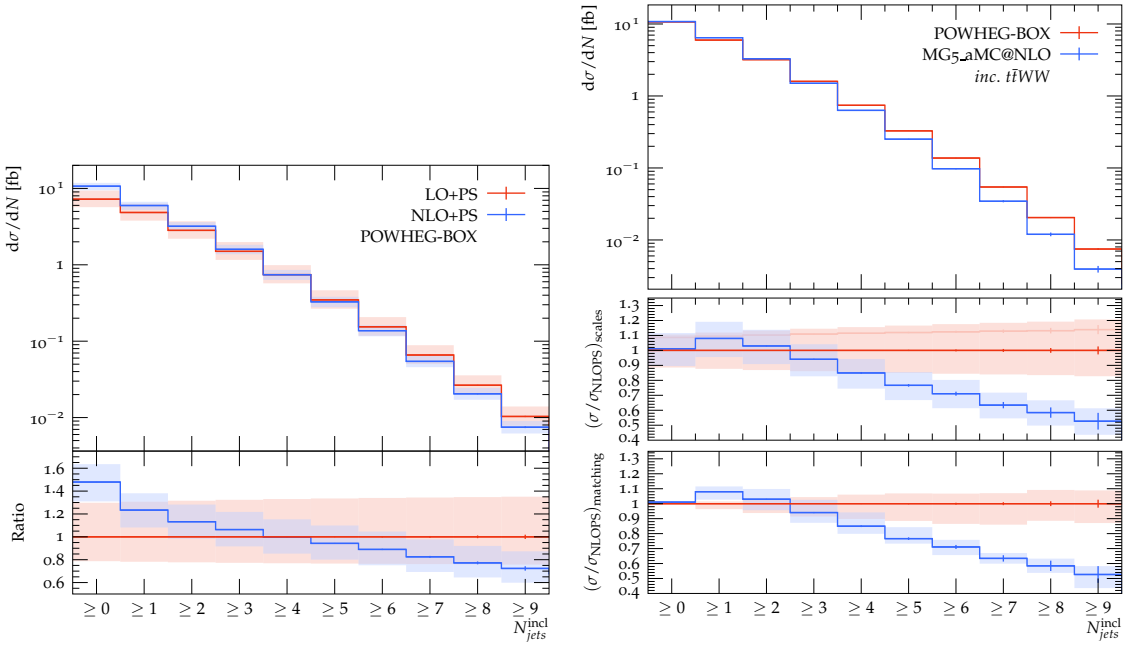


Figure 4.6: Inclusive $t\bar{t}WW$ QCD distribution in terms of the number of inclusive jets (i.e. jets that meet $N \geq N_{incl}$). In the left-hand panel (l.h.p.), a comparison between calculations an LO+PS and NLO+PS in POWHEG are shown, including their respective theoretical uncertainties. In the right-hand panel (r.h.p) a comparison between $t\bar{t}WW$ distribution in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Left panel

The distributions of the inclusive jet multiplicity N_{incl} for LO+PS and NLO+PS are shown in Fig. (4.6). It is observed that the predictions differ in shape. In the first bins, the NLO+PS estimate is above LO+PS, while for bins of high multiplicities the opposite exclusively occurs. The first bin, corresponding to $N \geq 0$ jets, includes all generated events, independently of the number of jets, and therefore it mainly contains information about the total normalization of the process. The difference reflected here

between NLO and LO is due to the inclusion of virtual and real corrections in the calculation, which modify the way in which the cross section is estimated.

For $N \geq 1$ the distribution begins to consider radiation because all events include at least one resolvabled radiation, which is generated by POWHEG at NLO+PS, while subsequent emissions, if any, are generated by the parton shower. In the same way, this explains that the region of high multiplicities is increasingly dominated by the shower dynamics. On the other hand, the first radiation at LO+PS level is generated by the parton shower. These differences may explain why there is a difference in the high multiplicity region in the estimate.

The lower panel displays the ratio of the NLO+PS prediction to the LO+PS result. The deviation from unity reflects the effect of higher-order corrections on the jet-multiplicity distribution. As discussed above, NLO corrections enhance the low-multiplicity region, while a suppression is observed at larger jet multiplicities
The uncertainty bands, obtained using the seven-point scale-variation prescription, are visibly reduced at NLO+PS with respect to LO+PS. This behavior reflects the reduced dependence of the prediction on the renormalization and factorization scales when higher-order corrections are included.
On the other hand, we can observe the ratio which is calculated by normalizing with respect to the LO+PS calculation. The theoretical uncertainties estimated following the seven-point envelope are reduced at NLO+PS with respect to LO+PS. This again reflects that calculations at higher orders reduce the dependence on the scale.
Despite the difference between the solid lines, the uncertainties overlap, except for the first bin, therefore the theoretical estimates are consistent within theoretical uncertainties.

Right panel

Distributions for the number of inclusive jets estimated implementing POWHEG and MG5_aMC@NLO are shown. In both cases, first emission is generated by these frameworks, while subsequent emissions are generated by the shower (PYTHIA). It is this first emission that determines the evolution of the rest of the radiation.
Small differences between POWHEG and MG5_aMC@NLO can be observed at the first bins in the distribution. At higher multiplicities, these differences increase. For low jet multiplicities, POWHEG and MG5_aMC@NLO show good agreement. However, as the jet multiplicity increases, the differences between the two frameworks become more pronounced, particularly in the high-multiplicity region. This behavior reflects the modeling uncertainty associated with the simulation of additional QCD

radiation beyond the first NLO emission. In particular, we observe that the differences are not accounted by the uncertainty estimates of either predictions. Therefore, the difference can be taken as an uncertainty of the matching procedure itself. This observation is especially relevant for experimental searches for four-top-quark production, whose signal regions are typically characterized by large jet multiplicities.

In the central panel the ratio between POWHEG and MG5_aMC@NLO is shown together with the uncertainty bands related to μ_f and μ_R variations, while the bottom ratio shows the uncertainty bands for matching scale variation. Theoretical uncertainties are in general smaller than the MG5_aMC@NLO ones. There exist a partial overlap of those bands, so differences between frameworks can be considered as compatible with theoretical uncertainties in certain region of the range of N_{incl}. In the region of higher multiplicities, these differences cannot be reconciled. This may indicate differences in the construction and structure of each framework.

4.3.3 Transverse momentum of the hardest jet, $p_T(j_1)$

In the following, we present the description of the results for the observable transverse momentum of the hardest jet, p_T shown in Fig. (4.7). Once again, the right panel shows a comparison between LO and NLO results, both within the POWHEG framework, while the left panel presents a comparison between POWHEG and MG_aMC@NLO.

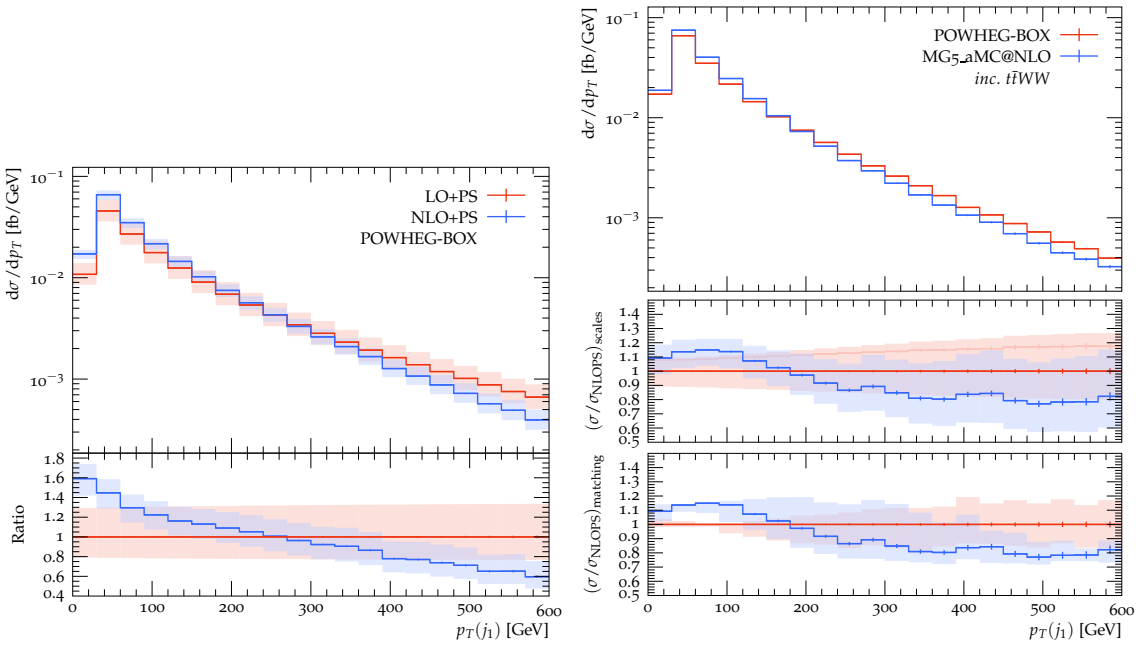


Figure 4.7: Inclusive $t\bar{t}WW$ QCD differential distribution of the transverse momentum of the hardest jet. In the l.h.p, a comparison between calculations an LO+PS and NLO+PS in POWHEG are shown, including their associated theoretical uncertainties. In the r.h.p, a comparison between $t\bar{t}WW$ production in POWHEG and in MG5_aMC@NLO is displayed, considering uncertainties related to the μ_F and μ_R scales and to the matching scheme.

Left panel

It is important to emphasize that at LO, the underlying process does not include radiative contributions. Therefore, within the LO+PS scheme, all radiation—particularly the first emission—originates from the parton shower.

It should be recalled that the shower, through the Sudakov factor, provides a reliable description of the soft and collinear regions of phase space. However, this approximation loses accuracy as one approaches the hard region of phase space. Thus, the first emission, which corresponds to the hardest one, is modeled with less accuracy.

Since POWHEG is based on p_T ordering, this first emission plays a fundamental role in determining all subsequent radiation, as it conditions, through the Sudakov factor, the structure of the subsequent emissions via a hierarchy in p_T.

Meanwhile, in the NLO+PS scheme, the first emission is generated using the matrix elements that contribute to the real corrections, and is therefore exact over the entire phase space. Once again, since the shower is conditioned by the hardest emission, this results in a more accurate description of the subsequent radiative cascade.

Right panel

In the left panel, the comparison between the POWHEG-BOX and MG_a_MC@NLO frameworks for the cross section distribution with respect to the transverse momentum of the hardest jet is shown, and both implement NLO+PS calculations.

In the upper panel it is visible that in both cases the distribution increases in the low transverse momentum region of the jet, and then begins to decrease as it increases. The distribution shows that the production of soft jets is more probable than that of hard jets, since the production of high-p_T jets requires energetic emissions, which are favored if the colliding partons carried a large momentum fraction of the parent hadron. And this probability is suppressed by the PDFs.

On the other hand, the middle panel shows the ratio between both predictions. This comparison is performed by means of POWHEG normalized $\sigma_{PW}/\sigma_{PW} = 1$ bin by bin. The blue line is σ_{MG}/σ_{PW}, which allows one to quantify the differences between both generators.
The solid lines correspond to the choice of the renormalization scale μ_R and the factorization scale μ_F, and the bands represent once again the theoretical uncertainty estimated through the variation of these scales by means of the 7-point envelope scheme.

We can see that in the high transverse momentum region, the characteristic scale of the process is large, so that $\alpha_s = \alpha_s(\mu_R)$ takes small values, such that small variations in the scale produce noticeable changes in α_s. Therefore, the prediction becomes more dependent on the scale in this regime, which produces a widening of the theoretical uncertainty bands in this region.

5 Discussion

In this work, we studied the $pp \to t\bar{t}WW$ production at NLO QCD for the events at a center of mass energy of 13.6 TeV, within the POWHEG-BOX framework. The interest in this process lies in the fact that it represents one of the most important backgrounds for $t\bar{t}t\bar{t}$ production, the latter being a rare process, and at the same time of great importance in the exploration of signals of new physics, especially related to the Higgs sector.

Being the latter a rare process, having greater control over the background processes that may appear is of special importance. Detailed studies on significant background processes such as $t\bar{t}W$, $t\bar{t}Z$ have been carried out in recent years, at different theoretical and technical levels. In this work, a detailed study of $t\bar{t}WW$ production at NLO (QCD) in hadronic collisions using the POWHEG framework is presented for the first time.

From this, differential distributions of the event for different observables of the process are studied. The objective has been to observe the predictions that this generator provides, and the theoretical uncertainties associated with its construction, among which those associated with the renormalization and factorization scales of the theory and with the matching procedure.

Due to the different matching procedures implemented by POWHEG and MadGraph5_aMC@NLO to combine fixed-order NLO calculations with parton showers, as well as their different treatments of the theoretical scales entering the simulation and the resulting fraction of positive and negative weight events, it was decided to perform a comparison of the simulation results between both frameworks.

The results were that fully inclusive observables of the event were generated consistently between both generators, resulting in distributions being generated with little difference or such that the uncertainty bands always overlapped. In the case of POWHEG it is necesarry to implement a consistent form of the damping function, to successfully separate the hardest real contributions which are treated at NLO order level.

It was not the same case for the distributions of exclusive observables, that are par-

ticularly sensitive to additional QCD radiation, such as the number of inclusive jets, the hadronic transverse energy, H_T and the transverse momentum of the hardest jet, $p_T(j_1)$, where the differences in the predictions between the different frameworks became more noticeable, despite the implementation of the damping function. This reflects the impact of the matching method used, as well as the treatment that each one makes with respect to the constants μ_R and μ_F.
Having, for the first time, access to the kinematic properties of the $t\bar{t}WW$ process provides a valuable tool for the interpretation of experimental results and contributes to the search for signals of new physics. In particular, a precise characterization of this process is important because it constitutes a relevant background to four-top-quark production. One of the most powerful observables in this context is the number of inclusive jets. The multiplicity of inclusive jets is expected to be larger in $t\bar{t}t\bar{t}$ production than in processes such as $t\bar{t}WW$, owing to the presence of four top quarks in the final state. Their decays, $t \to bW$, give rise to multiple b-jets, while additional jets can originate from the subsequent hadronic decays of the W bosons. As a consequence, $t\bar{t}t\bar{t}$ events tend to populate regions of high jet multiplicity

Moreover, part of the relevance of this work also lies in the fact that it involves the development of a new generator in the POWHEG framework for the $t\bar{t}WW$ process, which belongs to the class of $t\bar{t}VV$ processes, where V denotes a vector particle. This advancement, therefore, lays the foundations for the development of generators associated with the $t\bar{t}ZZ$ and $t\bar{t}WZ$ production processes, which constitute secondary contributions of this type to the $t\bar{t}t\bar{t}$ background, and whose inclusion complements the study of $t\bar{t}VV$ processes. In this sense, this work leads to a project whose objective is to integrate the study of these processes and to culminate in the publication of a paper that contributes to the field.

A Appendix

A.1 Resummation

Resummation is a procedure that allows one to deal with the appearance of large logarithmic terms in a perturbative series, which can spoil its convergence.
In perturbative QCD [60], a cross section can be written as an expansion in the strong coupling constant α_s :

$$\sigma = \sigma_0 + \alpha_s\sigma_1 + \alpha_s^2\sigma_2 + \cdots = \sum_{n=0}^{\infty} \alpha_s^n \sigma_n. \tag{A.1}$$

However, in certain regions of phase space, large logarithms L appear, for example:

$$\ln(1-z), \qquad \ln\left(\frac{Q^2}{\mu^2}\right), \qquad \text{etc.} \tag{A.2}$$

In these situations, the coefficients σ_n contain powers of these logarithms, such that the series takes the form

$$\sigma = \sigma_0 \left[1 + \alpha_s(a_1L^2 + b_1L + c_1) + \alpha_s^2(a_2L^4 + b_2L^3 + c_2L^2 + \cdots) + \cdots\right]. \tag{A.3}$$

In general, the dominant terms at each order are of the form $\alpha_s^n L^{2n}$, known as *Leading Logarithms* (LL). The terms $\alpha_s^n L^{2n-1}$ correspond to *Next-to-Leading Logarithms* (NLL), and so on.
When L is large, the terms $\alpha_s^n L^{2n}$ can be of order unity even if $\alpha_s \ll 1$, which breaks the convergence of the perturbative expansion. The idea of resummation is to reorganize the series and sum these dominant contributions to all orders.
At LL accuracy, this sum typically leads to an exponential:

$$\sigma^{LL} = \sigma_0 \exp\left[-A\alpha_s L^2\right], \tag{A.4}$$

whose expansion reproduces the dominant terms:

$$\sigma^{LL} = \sigma_0 \left(1 - A\alpha_s L^2 + \frac{1}{2}(A\alpha_s L^2)^2 + \cdots\right). \tag{A.5}$$

This type of exponentiation is characteristic of the resummation of soft and collinear logarithms.

A particularly important case in high-energy physics is the Sudakov factor.
The Sudakov factor $\Delta(t)$ represents the probability of not emitting resolvable radiation above a scale t (typically a p_T or a virtuality). In QCD, it can be written as:

$$\Delta(t) = \exp\left(-\int_t^{Q^2} \mathrm{d}\Phi_{\text{emission}} \frac{\alpha_s}{2\pi} P(z)\right), \tag{A.6}$$

where $P(z)$ is the splitting function.

LL approximation
In the soft and collinear limit, the emission phase space can be parametrized as:

$$\mathrm{d}\Phi_{\text{emission}} \sim \frac{\mathrm{d}k_T^2}{k_T^2}\,\mathrm{d}z\,\frac{\mathrm{d}\phi}{2\pi}. \tag{A.7}$$

The integral over k_T^2 generates a logarithm, while the integration over z, through the splitting function, produces another logarithm. As a result, double logarithmic terms appear in the exponent:

$$\mathrm{d}P \sim \frac{\alpha_s}{\pi} P(z) \frac{\mathrm{d}k_T^2}{k_T^2} \mathrm{d}z. \tag{A.8}$$

This explains the origin of the $\alpha_s^n L^{2n}$ terms that dominate in the LL approximation.

A.2 H_T

The quantity H_T is defined as

$$H_T = \sum_i p_T^i, \tag{A.9}$$

where i represents all final-state particles. It is not a vector quantity.
Due to the fact that

$$\hat{s} = x_1 x_2 s, \tag{A.10}$$

where $\hat{s}$ represents the energy scale of the hadrons in the collision, while x_1 and x_2 represent the partonic momentum fractions, it follows that, for massless particles,

$$E_1 = E_2 = \frac{\sqrt{\hat{s}}}{2}, \tag{A.11}$$

thus the maximum transverse momentum for the products could be obtained when $\sin\theta = 1$, so that $p_T^{max} = \frac{\sqrt{\hat{s}}}{2}$.
From the inversion of this relation one obtains

$$\hat{s} \geq 4p_T^2, \tag{A.12}$$

where obviously $p_T^2 \geq 0$. The above allows us to interpret that, for the production of jets with high transverse momentum, a large transfer of partonic momentum is required.
Taking this into account, H_T measures the activity in the transverse plane with respect to the collision axis.
That is, in LO+PS, the parton shower is responsible for contributing radiation, from the hardest to the softer ones. All of them contribute.

A.3 Invariant mass of the $t\bar{t}$ system

$$M_{t\bar{t}}^2 = (p_t + p_{\bar{t}})^2 \tag{A.13}$$

which in the center of mass frame is,

$$M_{t\bar{t}}^2 = (E_t + E_{\bar{t}})^2. \tag{A.14}$$

Then, $M_{t\bar{t}}$ represents the total energy of the pair $t\bar{t}$ in its center of momentum.
In the center of mass $t\bar{t}$

$$\vec{p_t} + \vec{p_{\bar{t}}} = 0 \implies \vec{p_t} = -\vec{p_{\bar{t}}}.$$

Then, the invariant mass

$$M_{t\bar{t}} = E_t + E_{\bar{t}} = 2\sqrt{m_t^2 + |\vec{p}|^2}.$$

If $|\vec{p}| \to 0$, then there is almost zero kinetic energy.

$$M_{t\bar{t}} \approx 2m_t \tag{A.15}$$

This means that the energy $M_{t\bar{t}}$ at the center of mass frame is approximately equal to the sum of the top quark masses with the tops being produced nearly at rest with respect to each other. [61]

A.4 Rapidity and pseudorapidity of particles in the final state

Due to the internal structure of the proton, the colliding partons carry a fraction of the proton momentum. Those momentum fractions are generally different for the two incoming partons, and the consequence is that the center-of-mass frame of the partonic interaction is boosted with respect to the laboratory frame along the z axis. Then, to characterize the the longitudinal motion of a resulting particle in the final state, the *rapidity* is defined as [62]

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{A.16}$$

where E and p_z are the particle energy and longitudinal momentum, respectively. A particle traveling close to the $+z$ beam direction has $p_z > 0$ and therefore $y > 0$, while a particle moving close to the $-z$ direction has $p_z < 0$ and hence $y < 0$. If the particle is emitted perpendicular to the beam axis, $p_z = 0$ and consequently $y = 0$.

The form of the equation of rapidity is useful because under a longitudinal Lorentz boost it transforms as

$$y' = y + y_{\text{boost}}, \tag{A.17}$$

where y_{boost} depends only on the boost. Therefore, rapidity differences between final-state particles in the same event,

$$\Delta y = y_1 - y_2, \tag{A.18}$$

remain invariant under longitudinal boosts, making them especially suitable for the study of hadron-collider events.

For highly relativistic particles, for which $m \ll |\vec{p}|$then $E \approx |\vec{p}|$, and consequently the rapidity approaches the pseudorapidity,

$$y \approx -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] := \eta, \tag{A.19}$$

where θ is the polar angle measured with respect to the beam axis. Since pseudorapidity depends only on the particle direction and can be reconstructed directly from detector information, it is widely used in experimental analyses.

A.5 The K-factor

A commonly used quantity in collider phenomenology is the K-factor, which measures the impact of higher-order perturbative corrections relative to a lower-order prediction. For an observable whose total cross section is known at different perturbative orders, the K-factor is defined as

$$K \equiv \frac{\sigma_{\mathrm{HO}}}{\sigma_{\mathrm{LO}}}, \tag{A.20}$$

where σ_{HO} denotes the cross section computed at a higher perturbative order (e.g. NLO or NNLO), and σ_{LO} is the corresponding leading-order prediction.
The K-factor provides a measure of the size of perturbative corrections. Values close to unity indicate that higher-order effects are relatively small, while larger deviations from unity signal substantial corrections. In hadronic collisions, sizeable K-factors often arise due to the opening of new partonic channels at higher orders or from large virtual corrections.
For differential distributions, it is useful to introduce a local or differential K-factor, defined as

$$\frac{\mathrm{d}\sigma_{\mathrm{HO}}/\mathrm{d}\mathcal{O}}{\mathrm{d}\sigma_{\mathrm{LO}}/\mathrm{d}\mathcal{O}}, \tag{A.21}$$

where $\mathcal{O}$ denotes the observable under consideration. In contrast to the total K-factor, the differential K-factor generally depends on the kinematic region. A constant differential K-factor indicates that higher-order corrections mainly affect the overall normalization of the distribution, whereas a non-uniform behavior reveals shape distortions induced by higher-order contributions.
Although the K-factor is a useful diagnostic tool, it should not be regarded as a physical observable. Its numerical value depends on the perturbative order considered, the renormalization and factorization scales, the choice of parton distribution functions, and the definition of the observable itself. Therefore, the K-factor should be interpreted only as an indicator of the relative importance of higher-order corrections within a given theoretical setup.

Bibliography

[1] J.H. Christenson, J.W. Cronin, V.L. Fitch and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, *Phys. Rev. Lett.* **13** (1964) 138.

[2] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.

[3] E288 collaboration, *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252.

[4] CDF collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [hep-ex/9503002].

[5] D0 collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [hep-ex/9503003].

[6] M. Faisst, J.H. Kuhn and O. Veretin, *Pole versus MS mass definitions in the electroweak theory*, *Phys. Lett. B* **589** (2004) 35 [hep-ph/0403026].

[7] G. Bhattacharyya, *A Pedagogical Review of Electroweak Symmetry Breaking Scenarios*, *Rept. Prog. Phys.* **74** (2011) 026201 [0910.5095].

[8] CMS, ATLAS collaboration, *Searches for Beyond Standard Model Physics at the LHC: Run1 Summary and Run2 Prospects*, *PoS* **FPCP2015** (2015) 024 [1507.08427].

[9] M.D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press (3, 2014), 10.1017/9781139540940.

[10] Particle Data Group collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.

[11] ATLAS collaboration, *Combination of measurements of the top quark mass from data collected by the ATLAS and CMS experiments at $\sqrt{s} = 7$ and 8 TeV*, .

[12] A.H. Hoang, *What is the Top Quark Mass?*, *Ann. Rev. Nucl. Part. Sci.* **70** (2020) 225 [2004.12915].

[13] M.C. Smith and S.S. Willenbrock, *Top quark pole mass*, *Phys. Rev. Lett.* **79** (1997) 3825 [hep-ph/9612329].

[14] ATLAS Collaboration, *Measurement of the top quark electric charge in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **11** (2013) 031 [1307.4568].

[15] M. Beneke et al., *Top quark physics*, in *Workshop on Standard Model Physics (and more) at the LHC (First Plenary Meeting)*, pp. 419–529, 3, 2000, DOI [hep-ph/0003033].

[16] G. Mahlon and S.J. Parke, *Spin Correlation Effects in Top Quark Pair Production at the LHC*, *Phys. Rev. D* **81** (2010) 074024 [1001.3422].

[17] CMS Collaboration, "CMS observes four-top quark production." CMS Experiment at CERN, March, 2023.

[18] A. Hoecker, "A transformative leap in physics: ATLAS results from LHC Run 2." ATLAS Experiment at CERN, April, 2025.

[19] G. Bevilacqua and M. Worek, *Constraining BSM Physics at the LHC: Four top final states with NLO accuracy in perturbative QCD*, *JHEP* **07** (2012) 111 [1206.3064].

[20] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, *JHEP* **07** (2018) 185 [1804.10017].

[21] R. Frederix and I. Tsinikos, *Subleading EW corrections and spin-correlation effects in $t\bar{t}W$ multi-lepton signatures*, *Eur. Phys. J. C* **80** (2020) 803 [2004.09552].

[22] ATLAS collaboration, *Search for $t\bar{t}H/A \to t\bar{t}t\bar{t}$ production in the multilepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, .

[23] B. Hespel, F. Maltoni and E. Vryonidou, *Signal background interference effects in heavy scalar production and decay to a top-anti-top pair*, *JHEP* **10** (2016) 016 [1606.04149].

[24] CMS collaboration, *Search for supersymmetry in pp collisions at $\sqrt{s} = 13$ TeV in the single-lepton final state using the sum of masses of large-radius jets*, *JHEP* **08** (2016) 122 [1605.04608].

[25] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics : a Primer for the LHC Era*, Oxford University Press (2018), 10.1093/oso/9780199652747.001.0001.

[26] M.L. Mangano, *Qcd and the physics of hadronic collisions*, in *Proceedings of the 2017 CERN–Latin-American School of High-Energy Physics (CLASHEP2017)*, vol. 4 of *CERN Yellow Reports: School Proceedings*, (Geneva), pp. 27–62, CERN, 2018, DOI.

[27] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and collider physics*, vol. 8, Cambridge University Press (2, 2011), 10.1017/CBO9780511628788.

[28] M. Markovych and A. Tandogan, *Analytic Evolution of DGLAP Equations*, 2304.10458.

[29] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 235–295, 2015, DOI [1411.4085].

[30] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, Boulder, CO (1995).

[31] Y.M. Shabelski, *The Role of incident parton transverse momenta in heavy quark hadroproduction*, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 474–483, 4, 1998 [hep-ph/9904492].

[32] G. 't Hooft and M.J.G. Veltman, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys. B* **44** (1972) 189.

[33] S. Catani and M.H. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, *Phys. Lett. B* **378** (1996) 287 [hep-ph/9602277].

[34] S. Catani, *The Singular behavior of QCD amplitudes at two loop order*, *Phys. Lett. B* **427** (1998) 161 [hep-ph/9802439].

[35] A. Mitov and S. Moch, *The Singular behavior of massive QCD amplitudes*, *JHEP* **05** (2007) 001 [hep-ph/0612149].

[36] A. Bassetto, M. Ciafaloni and G. Marchesini, *Jet Structure and Infrared Sensitive Quantities in Perturbative QCD*, *Phys. Rept.* **100** (1983) 201.

[37] S. Catani, S. Dittmaier, M.H. Seymour and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189 [hep-ph/0201036].

[38] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650.

[39] S. Catani and M.H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323].

[40] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012].

[41] L. Lönnblad, *Fooling Around with the Sudakov Veto Algorithm*, *Eur. Phys. J. C* **73** (2013) 2350 [1211.7204].

[42] L. Durand and W. Putikka, *Probabilistic Derivation of Parton Splitting Functions*, *Phys. Rev. D* **36** (1987) 2840.

[43] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].

[44] F. Febres Cordero, M. Kraus and L. Reina, *Top-quark pair production in association with a $W^{\pm}$ gauge boson in the POWHEG-BOX*, *Phys. Rev. D* **103** (2021) 094014 [2101.11808].

[45] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244].

[46] S. Weinzierl, *Introduction to Monte Carlo methods*, hep-ph/0006269.

[47] J. Alwall et al., *A Standard format for Les Houches event files*, *Comput. Phys. Commun.* **176** (2007) 300 [hep-ph/0609017].

[48] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092].

[49] S. Platzer and M. Sjodahl, *The Sudakov Veto Algorithm Reloaded*, *Eur. Phys. J. Plus* **127** (2012) 26 [1108.6180].

[50] Particle Data Group collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001.

[51] NNPDF collaboration, *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022) 428 [2109.02653].

[52] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [1412.7420].

[53] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [1111.5206].

[54] F. Buccioni, S. Pozzorini and M. Zoller, *On-the-fly reduction of open loops*, *Eur. Phys. J. C* **78** (2018) 70 [1710.11452].

[55] F. Buccioni, J.-N. Lang, J.M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866 [1907.13071].

[56] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[57] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[58] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavor production*, *JHEP* **08** (2003) 007 [hep-ph/0305252].

[59] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, *JHEP* **09** (2012) 130 [1207.5391].

[60] R. Bonciani, S. Catani, M.L. Mangano and P. Nason, *Sudakov resummation of multiparton QCD cross-sections*, *Phys. Lett. B* **575** (2003) 268 [hep-ph/0307035].

[61] W.-L. Ju, G. Wang, X. Wang, X. Xu, Y. Xu and L.L. Yang, *Invariant-mass distribution of top-quark pairs and top-quark mass determination*, *Chin. Phys. C* **44** (2020) 091001 [1908.02179].

[62] F.-H. Liu, Y.-H. Chen, Y.-Q. Gao and E.-Q. Wang, *On Current Conversion between Particle Rapidity and Pseudorapidity Distributions in High Energy Collisions*, *Adv. High Energy Phys.* **2013** (2013) 710534.