

# Contents

<b>1</b>	<b>The Standard Model</b>	<b>3</b>
1.1	An Overview of the Theory . . . . .	3
1.2	Gauge Symmetries . . . . .	6
1.2.1	QED Lagrangian . . . . .	6
1.2.2	QCD Lagrangian . . . . .	8
1.3	Electroweak Unification . . . . .	10
1.3.1	Lagrangian for Pure Weak Interactions . . . . .	10
1.3.2	Electroweak Unification . . . . .	10
1.4	Spontaneous Symmetry Breaking . . . . .	13
1.5	Yukawa Lagrangian . . . . .	16
1.6	The Standard Model Lagrangian . . . . .	17
<b>2</b>	<b>Collisions at the LHC</b>	<b>19</b>
2.1	Factorization . . . . .	19
2.2	Parton Distribution Functions . . . . .	20
2.2.1	Deep Inelastic Scattering . . . . .	21
2.3	Fragmentation Functions . . . . .	27
2.4	Jets . . . . .	27
2.4.1	Jet Definitions . . . . .	27
2.4.2	Jet Cross Sections . . . . .	28
2.5	Higher Order Corrections . . . . .	30
2.5.1	Infrared and Collinear Divergences . . . . .	30
2.5.2	Infrared and Collinear Safety for Jets . . . . .	31
<b>3</b>	<b>Higgs Physics</b>	<b>35</b>
3.1	Higgs Production Mechanisms . . . . .	35
3.2	Higgs Decay Mechanisms . . . . .	37
3.3	Properties of the Higgs . . . . .	40
<b>4</b>	<b>A Machine Learning Study of Color Sensitive Observables</b>	<b>41</b>
4.1	The Observables . . . . .	41
4.1.1	The Pull Variables . . . . .	41
4.1.2	$D_2$ . . . . .	43
4.1.3	Color Ring . . . . .	44
4.2	Simulation . . . . .	46
4.3	Analysis . . . . .	48
4.4	Machine Learning Architecture . . . . .	49
4.5	Results . . . . .	49

4.6 Mass Bias . . . . .	54
<b>Conclusions and Outlook</b>	<b>63</b>
<b>Appendices</b>	<b>65</b>
<b>A Neural Networks</b>	<b>65</b>
<b>B Boosted Decision Trees</b>	<b>67</b>

# Chapter 1

## The Standard Model

At the basis of all physics at the Large Hadron Collider (LHC) lies the Standard Model. This theory describes all fundamental interactions involving the electroweak and strong forces, as well as the fields which partake in these interactions. In this chapter, we will briefly illustrate the main features of this theory and show how together they paint a complete picture of our current understanding of elementary particle physics.

### 1.1 An Overview of the Theory

The Standard Model of particle physics is composed of two sectors: the matter fields and gauge fields. The matter fields are fermionic fields whose excitations lead to the particles which make up ordinary matter, i.e. quarks and leptons, and their corresponding antiparticles. The gauge fields, on the other hand, are bosonic fields which arise from symmetries of the model and describe the force-carrying particles, specifically the photon ( $\gamma$ ), gluons ( $g$ ),  $W^\pm$  and  $Z$ , as well as the Higgs boson,  $H$ .

The matter fields are organized in three generations, as shown in Figure 1.1. The particles in the three generations all have the same properties, with the sole exception of mass which increases with each generation. Neutrinos might be the exception to this rule, since their masses are not yet known to a useful degree of precision.

The properties of the matter fields determine the interactions in which they participate. All fermions can interact via the weak force. The electrically charged particles participate in electromagnetic interactions, and quarks, the only particles carrying color charge, can also interact via the strong nuclear force.

The interactions with these forces are mediated by the exchange of the gauge bosons. Photons are responsible for electromagnetic interactions, gluons for strong interactions, and the  $W^\pm$  and  $Z$  for weak interactions.

The strength of the interactions are determined by the value of the coupling constants,  $\alpha$  for electromagnetic interactions,  $\alpha_W$  for weak interactions and  $\alpha_S$  for strong interactions. Despite their name, the coupling constants are known to run, i.e. they depend on the energy scale  $Q^2$  of the process. Figure 1.2 shows how the coupling constants vary with energy. As can be seen,  $\alpha$  and  $\alpha_W$  increase as  $Q^2$  increases, while  $\alpha_S$  decreases with increasing energy, a property known as asymptotic freedom.

The running of the coupling constants is determined by the  $\beta$ -function, a renormalization group equation which can be solved in perturbation theory. At one-loop, the electromagnetic

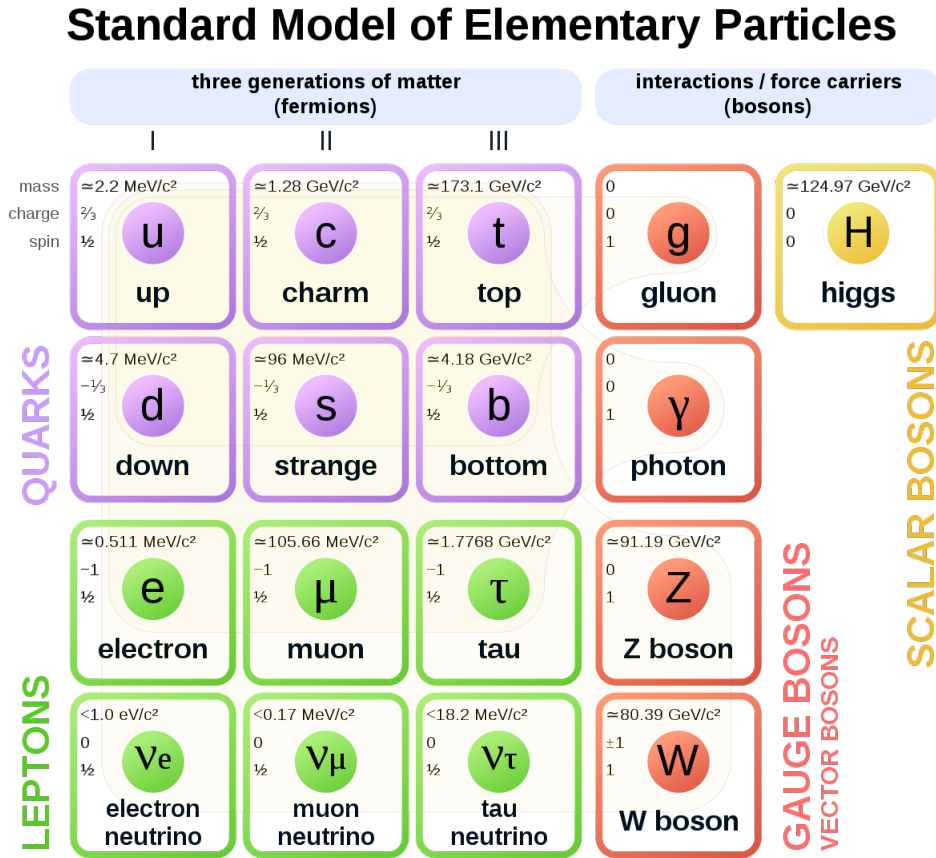


Figure 1.1: The “Periodic Table” of the Standard Model, showing the division between matter fields and gauge bosons, as well as the three different generations of the matter fields. Properties such as mass, charge and spin are also shown [1].

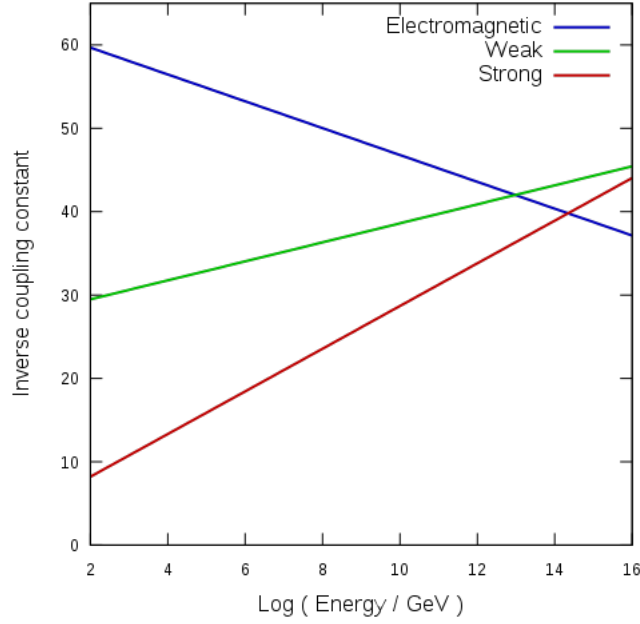


Figure 1.2: The  $\beta$ -functions for the three coupling constants included in the Standard Model [2].

$\beta$ -function is given by

$$\beta(\alpha) = \frac{2\alpha^2}{3\pi} \quad (1.1)$$

while the  $\beta$ -function for the strong interaction can be written as

$$\beta(\alpha_S) = - \left( 11 - \frac{2n_f}{3} \right) \frac{\alpha_S^2}{2\pi} \quad (1.2)$$

This inverse dependence on energy is a direct consequence of the non-Abelian nature of the strong nuclear force, as well as the number of flavors of quarks,  $n_f$  which enters the loops.

As mentioned, the gauge bosons originate from symmetries of the Standard Model. Specifically, the Standard Model Lagrangian is invariant under a local  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$  gauge symmetry. This means that, when a field  $\psi$  transforms under the action of this group, i.e. when

$$\psi \rightarrow \psi' = U(x)\psi, \quad (1.3)$$

the Lagrangian remains unchanged.

This is an example of invariance under a continuous symmetry. Per Nöther's theorem, these symmetries lead to conserved quantities, namely the conservation of color, weak isospin and hypercharge in every point of spacetime. There are a number of conserved quantities which arise from symmetries: momentum from translational symmetries, angular momentum from rotational symmetries, and electric charge from a  $U(1)_{EM}$  symmetry, just to name a few.

There are also a number of discrete symmetries present in the model. The most significant of these are *parity*, which inverts the spatial coordinates of a quantity

$$\hat{P}\psi(\vec{x}, t) = \psi(-\vec{x}, t), \quad (1.4)$$

*charge symmetry*, which exchanges all particles with the corresponding antiparticles,

$$\hat{C}\psi(x) = i\gamma^2\psi^*(x) \quad (1.5)$$

and *time-reversal symmetry* inverts the sign of the temporal coordinate of a quantity

$$\hat{T}\psi(\vec{x}, t) = \psi(\vec{x}, -t). \quad (1.6)$$

These symmetries are not individually present in all processes. For example, QED and QCD are both invariant under transformations of parity. These theories involve vector currents of the form  $j^\mu = \bar{\psi}\gamma^\mu\psi$ . These transform as

$$\begin{cases} \hat{P}j^0 = \bar{\psi}\gamma^0\gamma^0\gamma^0\psi = \bar{\psi}\gamma^0\psi \\ \hat{P}j^i = \bar{\psi}\gamma^0\gamma^i\gamma^0\psi = -\bar{\psi}\gamma^i\gamma^0\gamma^0\psi = -\bar{\psi}\gamma^i\psi = -j^i. \end{cases} \quad (1.7)$$

When we consider the matrix element  $\mathcal{M}$  for one of these interactions, which is proportional to the scalar product of two currents, we find that

$$\hat{P}j_1 \cdot \hat{P}j_2 = j_1^0 j_2^0 - \sum_i (-j_1^i)(-j_2^i) = j_1 \cdot j_2, \quad (1.8)$$

i.e. the interaction is invariant under parity transformations. On the other hand, weak interactions famously show parity violations [3]. This is due to the fact that the currents involved are of type V-A, i.e. of the form  $j^\mu \propto \bar{\psi}\gamma^\mu(1 - \gamma_5)\psi$ . This form mixes a vector current, with an axial vector current of the form  $j_A^\mu = \bar{\psi}\gamma^\mu\gamma^5\psi$ . In this case, when we consider the matrix element

$$\mathcal{M} \propto j_1 \cdot j_2 = g_V^2(j_V \cdot j_V) + g_A^2(j_A \cdot j_A) + g_V g_A(j_V \cdot j_A + j_A \cdot j_V) \quad (1.9)$$

we will find three terms which conserve parity, and one parity-violating term, since

$$\hat{P}j_V \cdot \hat{P}j_A = -j_V \cdot j_A. \quad (1.10)$$

Overall, only the combination of charge, parity, and time-reversal symmetry is expected to be conserved in all Standard Model interactions.

## 1.2 Gauge Symmetries

### 1.2.1 QED Lagrangian

We shall now begin to construct the Standard Model Lagrangian. Let us start by considering the free Lagrangian for a massive fermion field, given by the Dirac Lagrangian

$$\mathcal{L}_D = \bar{\psi}(i\cancel{D} - m)\psi, \quad (1.11)$$

where  $\psi$  is the fermion field,  $\bar{\psi}$  its Dirac adjoint, and, given the Dirac matrices  $\gamma^\mu$ ,  $\cancel{D}$  is the del operator in Feynman slash notation. It is easy to show that this Lagrangian is invariant under transformations of the type

$$\psi \rightarrow \psi' = \exp(i e \alpha) \psi \quad (1.12)$$

where  $e$  is a parameter which represents the coupling constant and  $\alpha$  is, for now, a parameter independent of the space-time coordinate  $x$ . In fact, the analogous transformation for the adjoint field  $\bar{\psi}$  is

$$\bar{\psi} \rightarrow \bar{\psi}' = [\exp(i e \alpha) \psi]^\dagger \gamma^0 = \psi^\dagger \exp(-i e \alpha) \gamma^0 = \bar{\psi} \exp(-i e \alpha) \quad (1.13)$$

since the operator  $\exp(-ie\alpha)$  commutes with  $\gamma^0$ . When applied to the whole Lagrangian, the transformation has the overall effect of leaving it unchanged:

$$\mathcal{L}_D \rightarrow \mathcal{L}'_D = \bar{\psi}'(i\cancel{\partial} - m)\psi' = \bar{\psi}(i\cancel{\partial} - m)\psi = \mathcal{L}_D. \quad (1.14)$$

Since the Lagrangian is unchanged, so too are the equations of motion. The transformed fields will therefore have the same dynamics.

We have just shown that the Dirac Lagrangian is invariant under a *global*  $U(1)$  gauge symmetry in charge space.

At this point, if we want to construct the QED Lagrangian, we must add the kinetic term describing the free photon field

$$\mathcal{L}_{kin} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (1.15)$$

where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ , as well as a term describing the interaction between the two fields. We can do this in two ways.

### Minimal Coupling

The first, more direct, prescription calls for applying the minimal coupling rule. This requires substituting the four-momentum of the fermion, which we will take to be an electron, with an expression which includes the electromagnetic potential

$$p_\mu \rightarrow p_\mu - eA_\mu \quad (1.16)$$

and the coupling constant  $e$ . Quantum mechanically, this corresponds to substituting the del operator in the Lagrangian. The Lagrangian thus becomes

$$\mathcal{L}_{QED} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\cancel{\partial} - e\cancel{A} - m)\psi = \mathcal{L}_{kin} + \mathcal{L}_D - e\bar{\psi}\gamma^\mu\psi A_\mu. \quad (1.17)$$

This Lagrangian is still invariant under the same global gauge symmetry as before.

We would like, however, to impose a more stringent symmetry requirement: a *local* gauge symmetry dependent on the space-time coordinate  $x$ . Whereas a global symmetry establishes the conservation of a conserved quantity, e.g. electric charge, in any closed system, the local symmetry imposes the same requirement in each point  $x$ .

If we promote the gauge symmetry to a local symmetry, i.e. we apply the transformation

$$\psi \rightarrow \psi' = \exp[ie\alpha(x)]\psi, \quad (1.18)$$

we find that the Lagrangian is no longer invariant under this transformation due to the action of the derivative. In fact, ignoring the terms which remain invariant, we find that

$$\mathcal{L}' = i\bar{\psi}'\cancel{\partial}\psi' = i\bar{\psi}'\exp[-ie\alpha(x)]\gamma^\mu\partial_\mu\{\exp[ie\alpha(x)]\psi\} = i\bar{\psi}\cancel{\partial}\psi - e\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha(x). \quad (1.19)$$

We can use a trick to reobtain the gauge invariance. We know that the electromagnetic tensor  $F^{\mu\nu}$  is gauge invariant. This means that if  $A_\mu$  undergoes the transformation

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu f(x) \quad (1.20)$$

where  $f(x)$  is a function such that  $\square f = 0$ , then

$$F_{\mu\nu} \rightarrow F'_{\mu\nu} = \partial_\mu(A'_\nu + \partial_\nu f) - \partial_\nu(A'_\mu + \partial_\mu f) = F_{\mu\nu}. \quad (1.21)$$

If we choose  $f$  opportunely, we can cancel out the extra term which appears in (1.19) with the last term in (1.17). Specifically, the choice  $f(x) = -\alpha(x)$  satisfies our request. Therefore, by combining the transformations (1.18) and (1.20), we can obtain an invariant Lagrangian.

### Gauge Principle

The second, more general, way of adding an interaction term to the Lagrangian is by the Gauge Principle. This principle describes a protocol through which we can obtain the dynamics of QED, or any field theory, starting from the global gauge transformation (1.12).

We start, once again, from the Lagrangian (1.11). Having identified the global gauge symmetry of the Lagrangian and having promoted it to a local symmetry, we define the covariant derivative as

$$D_\mu \doteq \partial_\mu + ieA_\mu. \quad (1.22)$$

We then require that the term  $D_\mu\psi$  transforms as the field  $\psi$  itself

$$D_\mu\psi \rightarrow (D_\mu\psi)' = D'_\mu\psi' = \{\partial_\mu + ieA'_\mu\}\psi' = \exp[ie\alpha(x)]D_\mu\psi. \quad (1.23)$$

By developing the equality, we find that  $A'_\mu$  must be given by

$$A'_\mu = A_\mu - \partial_\mu\alpha(x) \quad (1.24)$$

in order for the Lagrangian to remain invariant.

We can then use the covariant derivative to build a term which describes the free propagation of the field  $A_\mu$ . We do this by computing the commutator. With some basic algebra, we find that

$$[D_\mu, D_\nu] = ie\{\partial_\mu A_\nu - \partial_\nu A_\mu\} \equiv ieF_{\mu\nu}, \quad (1.25)$$

where  $F_{\mu\nu}$  is now a generic tensor of the field  $A_\mu$ . We thus have

$$F_{\mu\nu} = -\frac{i}{e}[D_\mu, D_\nu]. \quad (1.26)$$

We can then use the field tensor to construct a normalised, gauge invariant Lorentz scalar which will necessarily take the form (1.15).

We have thus arrived at the QED Lagrangian in a general way, without assuming any prior knowledge about the field  $A_\mu$ .

### 1.2.2 QCD Lagrangian

Armed with the gauge principle, it is now straightforward to derive the QCD Lagrangian. We must note, however, that a few complications arise from the fact that we are now dealing with a non-abelian gauge theory, i.e. a theory whose symmetry group is non-commutative. For a general Yang-Mills theory, the gauge group is  $SU(N)$ , but in QCD we will be working with  $N = 3$ .

The Dirac field for the quark can be indicated as  $q_f^\alpha$  where  $f$  is the flavour index and  $\alpha$  is the color index. We know that each flavour comes in three colours, so we can group the fields for each flavour in a three-component vector

$$q_f = \begin{bmatrix} q_f^1 \\ q_f^2 \\ q_f^3 \end{bmatrix}. \quad (1.27)$$

We can thus write the free Lagrangian for the quarks as

$$\mathcal{L}_D = \sum_f \bar{q}_f (i\not{\partial} - m_f) q_f \quad (1.28)$$



where  $m_f$  is a parameter representing the quark mass and  $(i\cancel{D} - m_f)$  is a 3-dimensional diagonal matrix. The quark mass  $m_f$  must be understood as a free parameter of the Lagrangian since it is not directly measurable, as free quarks do not exist in nature.

The Lagrangian is invariant under the following global gauge transformations in colour space:

$$q_f \rightarrow (q_f)' = \exp \left[ i\theta_a \frac{\lambda^a}{2} \right] q_f \quad (1.29)$$

where  $\theta_a$  is a parameter and  $a = 1, \dots, 8$  since, in general, the fundamental representation of  $SU(N)$  has  $N^2 - 1$  generators.  $\lambda^a$  represents the Gell-Mann matrices, which in the fundamental representation of  $SU(3)$  can be written as

$$\begin{aligned} \lambda^1 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \lambda^2 &= \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \lambda^3 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \lambda^4 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, & \lambda^5 &= \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, & \\ \lambda^6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & \lambda^7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, & \lambda^8 &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned} \quad (1.30)$$

The Gell-Mann matrices also allow us to define the structure constant of  $SU(3)$ ,  $f_{abc}$

$$\left[ \frac{\lambda_a}{2}, \frac{\lambda_b}{2} \right] = if_{abc} \frac{\lambda_c}{2}. \quad (1.31)$$

We can now proceed with the gauge principle. We define the covariant derivative as

$$D_\mu = \partial_\mu + ig_s \frac{\lambda_a}{2} G_\mu^a \quad (1.32)$$

where we have introduced the strong coupling constant  $g_s$  and 8 spin-1 vector fields  $G_\mu^a$ . These are the gluon fields. We now promote  $\theta_a$  to  $\theta_a(x)$  and require that  $D_\mu q_f$  transform as  $q_f$  so as to fix the interaction term between the quarks and the gauge bosons. We find that

$$G_\mu^a \rightarrow (G_\mu^a)' = G_\mu^a - \frac{1}{g_s} \partial_\mu \theta^a(x) - f^{abc} \partial_\mu \theta_b(x) G_{\mu c}. \quad (1.33)$$

Last but not least, using the relation

$$-\frac{i}{g_s} [D_\mu, D_\nu] = \frac{\lambda_a}{2} G_{\mu\nu}^a \quad (1.34)$$

we can define the gluon tensor field

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_{\mu b} G_{\nu c} \quad (1.35)$$

which we use to construct the gauge-invariant kinetic term with proper normalisation. We thus find that the complete Lagrangian takes the following form:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu} + \sum_f \bar{q}_f (i\cancel{D} - m_f) q_f. \quad (1.36)$$

## 1.3 Electroweak Unification

### 1.3.1 Lagrangian for Pure Weak Interactions

The gauge principle can also be applied to weak interactions. In this case, the fundamental symmetry is  $SU_L(2)$ , labelled as such because it only applies to left-handed particle states or right-handed anti-particle states. From here on we will consider only particle states, though analogous considerations hold for anti-particle states.

The symmetry acts on a weak-isospin doublet, e.g.

$$\psi_L = \begin{bmatrix} \nu_\ell \\ \ell^- \end{bmatrix}_L \quad (1.37)$$

composed of a left-handed neutrino and a lepton, or in general the left-handed states of any two fermions belonging to the same generation<sup>1</sup>. Corresponding leptonic right-handed states  $\ell_R^-$  are placed in a singlet state, and right-handed neutrino states are not considered as they have not been observed in nature [4].

As before, by applying the local gauge transformations

$$\psi_L \rightarrow \psi'_L = \exp \left[ i \frac{\tau_j}{2} \alpha^j(x) \right] \psi_L \quad (1.38)$$

$$\ell_R \rightarrow \ell'_R = \ell_R \quad (1.39)$$

where  $i=1,2,3$  and  $\tau^i$  are the generators of  $SU(2)$ , usually chosen to be the Pauli spin matrices

$$\tau^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \tau^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \tau^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (1.40)$$

we can derive the full Lagrangian describing weak interactions

$$\mathcal{L}_W = -\frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} + \bar{\psi}_L (i \not{D} - m_i) \psi_L + \ell_R (i \not{\partial} - m_i) \ell_R \quad (1.41)$$

where  $D_\mu = \partial_\mu + ig \frac{\tau_j}{2} W_\mu^j(x)$  and  $g$  is the weak coupling constant.

The fields  $W_\mu^1(x)$ ,  $W_\mu^2(x)$ , and  $W_\mu^3(x)$  in their raw form are not sufficient to describe the observed phenomenology of weak interactions. By considering

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \quad (1.42)$$

we obtain charged currents which describe the transition from upper and lower components of the weak-isospin doublet observed in nature. Naturally, one would then attempt to identify  $W_\mu^3$  with the  $Z^0$ , however  $W_\mu^3$  only couples to left-handed particles or right-handed anti-particles, in contrast to what is observed for the physical  $Z^0$ .

### 1.3.2 Electroweak Unification

A more complete description is thus required to match the theory to the physical reality. The  $Z^0$  boson is not the only neutral boson observed in nature: there is also the  $\gamma$ . We can therefore attempt to include electromagnetic interactions in our description of weak interactions and derive the  $Z^0$  and  $\gamma$  fields from two neutral fields.

---

<sup>1</sup>Quark mixing slightly complicates this.

To this aim, we start by introducing *hypercharge*, defined as

$$Y = 2(Q - I_W^{(3)}) \quad (1.43)$$

This is a quantity meant to replace electric charge, so that the  $U(1)$  local gauge symmetry of QED, becoming  $U_Y(1)$ . In this way, we have identified a quantum number capable of distinguishing the states composing the left-handed doublet from the one composing the right-handed singlet, since, for example the states considered in (1.37) both have hypercharge  $Y = -1$  according to this definition, whereas the  $SU_L(2)$  singlet state  $\ell_R^-$  has hypercharge  $Y = -2$ .

We can now consider the full gauge symmetry for electroweak interactions,  $SU_L(2) \otimes U_Y(1)$ . Under this new gauge symmetry, the transformations (1.38) become

$$\psi_L \rightarrow \psi'_L = \exp[iy_1\beta(x)] \exp\left[i\frac{\tau_j}{2}\alpha^j(x)\right] \psi_L \quad (1.44)$$

$$\ell_R \rightarrow \ell'_R = \exp[iy_2\beta(x)] \ell_R \quad (1.45)$$

where  $y_1$  and  $y_2$  are the hypercharges of the weak isospin doublet and singlet, respectively. The covariant derivatives thus are

$$D_\mu \psi_L(x) = \left[ \partial_\mu + ig\frac{\tau_j}{2}W_\mu^j(x) + ig'\frac{y_1}{2}B_\mu(x) \right] \psi_L(x) \quad (1.46)$$

$$D_\mu \ell_R = \left[ \partial_\mu + ig'\frac{y_2}{2}B_\mu(x) \right] \ell_R(x) \quad (1.47)$$

where  $g$  and  $g'$  are the two coupling constants, in general different from one another.

We now have four different gauge bosons,  $W_\mu^j(x)$  and  $B(x)$ , which must be identified with the physical gauge bosons  $W^\pm$ ,  $Z^0$  and  $\gamma$ . The physical W bosons can be identified through the relation in (1.42). The mapping from  $W_\mu^3$  and  $B_\mu$  to  $Z_\mu$  and  $A_\mu$  can be achieved through a rotation in the neutral sector the gauge bosons

$$\begin{bmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{bmatrix} \begin{bmatrix} B_\mu \\ W_\mu^3 \end{bmatrix} = \begin{bmatrix} A_\mu \\ Z_\mu \end{bmatrix}. \quad (1.48)$$

If we define the vector

$$\psi = \begin{bmatrix} \nu_{\ell L} \\ \ell_L \\ \ell_R \end{bmatrix} \quad (1.49)$$

and write out the full Lagrangian containing the neutral currents of the electroweak sector of Standard Model,

$$\begin{aligned} \mathcal{L}_{NC} = & \bar{\psi} \gamma_\mu \left\{ g \sin \theta_W \frac{\tau_3}{2} + g' \cos \theta_W \frac{Y(\psi)}{2} \right\} \psi A^\mu \\ & + \bar{\psi} \gamma_\mu \left\{ g \cos \theta_W \frac{\tau_3}{2} - g' \sin \theta_W \frac{Y(\psi)}{2} \right\} \psi Z^\mu, \end{aligned} \quad (1.50)$$

we can see that we are required to impose a condition on the coefficients in (1.50) to re-obtain the physical currents that we are familiar with. Specifically, the first part of (1.50) corresponds to the interaction term of the Lagrangian (1.17) and the second term corresponds to an interaction term involving a second neutral boson. For the sake of simplicity, we can consider the case of an electron for the purpose of this matching.

The interaction term of (1.17), when specifying the left-handed and right-handed components, corresponds to

$$\mathcal{L} = -e [\bar{e}_L \gamma_\mu e_L + \bar{e}_R \gamma_\mu e_R] A^\mu. \quad (1.51)$$

fermion	$Q$	$I_W^{(3)}$	$Y_L$	$Y_R$
$\nu_\ell$	0	$+\frac{1}{2}$	-1	0
$\ell^-$	-1	$-\frac{1}{2}$	-1	-2

Table 1.1: Values of parameters

By inspection, we find that

$$-e = g \sin \theta_W \frac{\tau_3}{2} + g' \cos \theta_W \frac{Y(\psi_e)}{2}. \quad (1.52)$$

By specifying  $\tau_3$  and  $Y(\psi_e)$  to the appropriate component of  $\psi$ , in accordance with Table (1.1), we find that

$$g \sin \theta_W = g' \cos \theta_W = e. \quad (1.53)$$

$\theta_W$  is the *weak mixing angle*, corresponding to the angle of rotation in neutral sector necessary to achieve the desired mapping. Experimentally,  $\sin^2 \theta_W$  has been measured to be  $0.22290 \pm 0.00030$ , though theoretically it can be parametrized in terms of other quantities as we shall see in the next section.

We can also see that the second neutral current in (1.50) does indeed correspond to the  $Z$  current. It is easy to show that it can be written as

$$\mathcal{L}_{NC}^Z = \bar{\psi} \gamma_\mu \frac{e}{\sin \theta_W \cos \theta_W} Q_Z \psi Z^\mu. \quad (1.54)$$

where  $Q_Z = \{\frac{\tau_3}{2} - Q \sin^2 \theta_W\}$ .  $Q_Z$  is a  $3 \times 3$  diagonal matrix, allowing for access to each of the components of  $\psi$ . It needs to be specified in order to find the full coupling constant. This is straightforward for  $\nu_{eL}$ : the coupling constant turns out to be  $\frac{e}{2 \sin \theta_W \cos \theta_W}$ .

For the electron, some additional manipulations must first be made since we must deal with the two components. The Lagrangian for the interaction can be written as

$$\mathcal{L}_{NC}^{Ze} = \frac{e}{\sin \theta_W \cos \theta_W} \{ \bar{e}_L \gamma_\mu Q_Z^L e_L + \bar{e}_R \gamma_\mu Q_Z^R e_R \} Z^\mu. \quad (1.55)$$

The projections can be obtained by considering the operators  $P_{L/R} = \frac{1}{2}(1 \mp \gamma_5)$ , i.e.

$$\begin{cases} \bar{e}_L \gamma_\mu e_L = \bar{e} \gamma_\mu \frac{1}{2} (1 - \gamma_5) e \\ \bar{e}_R \gamma_\mu e_R = \bar{e} \gamma_\mu \frac{1}{2} (1 + \gamma_5) e. \end{cases} \quad (1.56)$$

The Lagrangian thus becomes

$$\mathcal{L}_{NC}^{Ze} = \frac{e}{\sin \theta_W \cos \theta_W} \left\{ \bar{e} \gamma_\mu \frac{1}{2} (Q_Z^L + Q_Z^R) e + \bar{e} \gamma_\mu \frac{1}{2} (Q_Z^L - Q_Z^R) e \right\} Z^\mu. \quad (1.57)$$

$Q_Z^{L/R}$  can be specified directly from (1.54) using the values from Table (1.1). We can use those values to specify the couplings which appear in (1.57). The Lagrangian simplifies to

$$\mathcal{L}_{NC}^{Ze} = \frac{e}{2 \sin \theta_W \cos \theta_W} \bar{e} \gamma_\mu \{v_e - a_e \gamma_5\} e Z^\mu \quad (1.58)$$

where

$$\begin{cases} v_e = I_W^{(3)}(e_L) (1 + 4Q_e \sin^2 \theta_W) \\ a_e = I_W^{(3)}(e_L) \end{cases} \quad (1.59)$$

are the vector and axial components, respectively.

Thus, we correctly find that the interaction with the  $Z$  can involve both  $e_L$  and  $e_R$  and that the interaction is of the type V-A, with different coupling constants for the two chiral states of the electron.

The full Lagrangian for electroweak interactions can thus be written as

$$\mathcal{L}_{EW} = \bar{\psi} (i\not{D} - m) \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^j W_j^{\mu\nu}, \quad (1.60)$$

which, when combined with the rotation in (1.48), gives a correct description of the observed phenomenology.

## 1.4 Spontaneous Symmetry Breaking

So far, we have only considered Lagrangians which contain massless gauge bosons. This is for a very precise reason: mass terms vary under gauge transformations. If we consider, for example, the Proca action for a generic massive bosonic field in an abelian gauge theory

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - m^2 A_\mu A^\mu \quad (1.61)$$

it is clear that when we apply the transformation (1.20) the Lagrangian is no longer invariant. This is a significant problem since it is known that the  $W^\pm$  and  $Z$  bosons are massive.

To solve the problem of massive gauge bosons, it is necessary to introduce the *Brout-Englert-Higgs Mechanism* [5, 6], which induces the spontaneous breaking of the gauge symmetry.

This mechanism introduces a scalar field, known as the Higgs field, composed of a weak isospin doublet of two complex scalar fields, or equivalently four real scalar fields

$$\phi(x) = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(\phi_3 + i\phi_4) \\ \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \end{bmatrix} \quad (1.62)$$

governed by a complex  $\phi^4$  Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi, \phi^\dagger) \quad (1.63)$$

where the potential  $V(\phi, \phi^\dagger) = \frac{m^2}{2} \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2$ . The components of the doublet, as part of the Electroweak sector of the Standard Model, have quantum numbers

$$\begin{cases} I_W^{(3)}(\phi^+) = \frac{1}{2} \\ Y(\phi^+) = 1 \\ I_W^{(3)}(\phi^0) = \frac{1}{2} \\ Y(\phi^0) = 1 \end{cases} \quad (1.64)$$

This means that  $\phi^+$  carries electrical charge, based on (1.43).

The field  $\psi_L$  is added to the Lagrangian (1.60), leading to a Lagrangian which remains invariant under a global  $SU_L(2) \otimes U_Y(1)$  gauge symmetry, as is easily verifiable. The symmetry can be promoted to a local gauge symmetry, resulting in the following Lagrangian for the Higgs field

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi, \phi^\dagger) \quad (1.65)$$

where  $D_\mu$  is the covariant derivative defined in (1.46).



Figure 1.3: The “Mexican hat” Higgs Potential for an Abelian gauge theory.

By minimizing the potential  $V$ , we can identify the ground state of the field

$$\frac{\partial V}{\partial |\phi|} = m^2 |\phi| + \lambda |\phi|^3 \quad (1.66)$$

If  $m^2 > 0$  and  $\lambda > 0$ , the minimum occurs when  $\phi = 0$ . However, if we interpret  $m^2$  as a parameter rather than as a mass and allow  $m^2 < 0$ , we find that there is a local maximum at  $\phi = 0$  and a set of minimum value states at

$$\langle 0 | \phi^\dagger \phi | 0 \rangle \equiv (\phi^\dagger \phi)_0 = -\sqrt{\frac{m^2}{\lambda}}. \quad (1.67)$$

The quantum vacuum has thus shifted, as represented in Figure 1.3. The vacuum is *degenerate* since there are infinite values of  $\phi^\dagger \phi$  which minimize  $V$ . The gauge symmetry is said to be spontaneously broken, the choice of parameters causes the vacuum state and the Lagrangian to no longer share the same symmetry. Since  $\lambda$  is an adimensional quantity, it has the dimensions of energy.

Without loss of generality, we can choose for the vacuum states

$$(\phi_1)_0 = -\sqrt{\frac{m^2}{\lambda}} \equiv \frac{v}{\sqrt{2}}, \quad (\phi_i)_0 = 0 \quad (1.68)$$

where  $i = 2, 3, 4$  and  $v$  is the vacuum expectation value of the Higgs field. Our doublet is thus

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v \end{bmatrix}. \quad (1.69)$$

The Lagrangian (1.60) along with (1.63) remain invariant under the local gauge transformation, however the vacuum state is no longer invariant under neither the  $SU_L(2)$  nor the  $U_Y(1)$  local gauge symmetries. For example,

$$\phi_0 \rightarrow \phi'_0 = \exp \left[ ig \frac{\tau_j}{2} \alpha^j(x) \right] \phi \approx \left\{ 1 + ig \frac{\tau_j}{2} \alpha^j(x) + \dots \right\} \phi_0 \neq \phi_0 \quad (1.70)$$

Specifically, the invariance is lost due to the action of the generators  $\tau_j$ . For this reason, these generators are said to be *broken*. Likewise,  $Y$  is a broken generator.

Before continuing the discussion, we must first introduce an important theorem involving broken generators.

**Theorem 1.4.1 (Goldstone Theorem)** *For all continuous global symmetries which do not leave the vacuum state unchanged, there exist corresponding massless particles equal in number to the number of broken generators.*

The theorem holds for global symmetries, however it is also relevant when dealing with local symmetries. In this case, the massless bosons which appear cannot be interpreted as physical particles. They can be gauged away, leading to Higgs-Kibble ghosts which result in *massive* bosons and allow for a physical interpretation of the theory. These ghosts are crucial for our ultimate goal: to give mass to the  $W^\pm$  and  $Z$  while leaving  $\gamma$  massless.

In order to get to our desired result, we must make sure to have one unbroken generator. To obtain it, we can consider the following linear combinations

$$\begin{cases} Q = \frac{\tau_3}{2} - \frac{Y}{2} \\ Q' = \frac{\tau_3}{2} + \frac{Y}{2}. \end{cases} \quad (1.71)$$

It is easy to show that  $Q$  is an unbroken generator and  $Q'$  is broken. In this way we have obtained 3 broken generators ( $\tau_1, \tau_2, Q'$ ) and 1 unbroken ( $Q$ ).

We can now proceed to study the vacuum fluctuations of  $\phi$ . Naively, these can be written as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_3(x) + i\phi_4(x) \\ v + \phi_1(x) + i\phi_2(x) \end{bmatrix}, \quad (1.72)$$

though, equivalently, we can write

$$\phi(x) = \frac{1}{\sqrt{2}} \exp \left[ \frac{iT_j \xi_j(x)}{2} \right] \begin{bmatrix} 0 \\ v + H(x) \end{bmatrix} \quad (1.73)$$

where  $T_j$  are the three  $SU(2)$  generators. In the latter form, we have merely parametrized the “naive” expression, as can be seen by expanding the exponential term to first order. The fields  $\xi_j(x)$  are the ghosts, which we shall gauge away by choosing the unitary gauge

$$\phi(x) \rightarrow \phi'(x) = \exp \left[ -\frac{iT_j \xi_j(x)}{2} \right] \phi. \quad (1.74)$$

In accordance with the gauge protocol, this requires a subsequent modification of  $D_\mu$ , which leads to a modification of the gauge fields, which are said to “eat” the ghosts. After having done so, if we go on to calculate the first term in (1.63) and use (1.42), we find

$$(D_\mu \phi)^\dagger (D_\mu \phi) = \frac{g^2}{4} (v^2 + 2vH + H^2) W_\mu^+ W^{-\mu} + \frac{1}{8} (v+H)^2 (g^2 W_\mu^3 W^{3\mu} - 2gg' W_\mu^3 B^\mu + g'^2 B_\mu B^\mu) \quad (1.75)$$

which contains all the physically meaningful terms. In particular, we can see that (1.75) contains Proca mass terms such as

$$\frac{g^2}{4} v^2 W_\mu^+ W^{-\mu} \equiv M_W^2 W_\mu^+ W^{-\mu}. \quad (1.76)$$

We can then follow the same logic as in that used to derive the unified Electroweak theory and map  $B_\mu$  and  $W_\mu^3$  to  $A^\mu$  and  $Z_\mu$ , respectively. After doing so, the second term in (1.75) becomes

$$\frac{1}{8}(g^2 + g'^2)(v + H)^2 Z_\mu Z^\mu \quad (1.77)$$

and we can see that there is no mass term for  $A_\mu$ , no interaction term involving the Higgs field  $H(x)$  and  $A_\mu$ , and that the  $Z$  acquires the mass  $M_Z^2 = \frac{v^2}{4}(g^2 + g'^2)$ . We can also see that the coupling of the Higgs boson to the other bosons is proportional to the bosons' masses.

We have succeeded in giving mass to the three weak gauge bosons. The choice of parameters  $\lambda$  and  $m^2$  spontaneously breaks the gauge symmetry, and the interaction of the field  $\phi$  with the potential generates would-be Goldstone bosons which manifest as ghosts. The choice of unitary gauge allows for the gauge bosons to eat the ghosts, thus gaining mass. The energy scale at which the gauge symmetry is spontaneously broken, known as the vacuum expectation value, is given by  $v$ , which in numerical terms corresponds to 246 GeV. Above this energy, the electromagnetic force and the weak nuclear force become one unified electroweak force.

## 1.5 Yukawa Lagrangian

A similar problem occurs when considering the mass terms for fermions. In this case, the mass term appearing in the Dirac Lagrangian (1.11) does not respect the  $SU_L(2) \otimes U_Y(1)$  gauge symmetry due to the fact that the left and right-handed components of the spinor transform differently

$$-m\bar{\psi}\psi = -m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R). \quad (1.78)$$

This problem can be solved by introducing an interaction with the Higgs field. An infinitesimal  $SU(2)$  local gauge transformation has the following effect on the Higgs

$$\phi \rightarrow \phi' = \left\{1 + ig\frac{\tau_j}{2}\alpha^j(x)\right\}\phi. \quad (1.79)$$

On the other hand, the same transformation has the opposite effect on  $\bar{\psi}_L$

$$\bar{\psi}_L \rightarrow \bar{\psi}'_L = \bar{\psi}_L \left\{1 - ig\frac{\tau_j}{2}\alpha^j(x)\right\}. \quad (1.80)$$

Therefore, if we consider the combination  $\bar{\psi}_L\phi$ , we find that this is a gauge invariant quantity. The same holds true for the  $U(1)$  gauge symmetry. Since  $\ell_R$  transforms independently from  $\psi_L$ , we can add it to the combination so as to account for the right-handed component as well. Thus the Lagrangian

$$\mathcal{L}_Y = -k(\bar{\psi}_L\phi\ell_R + \bar{\ell}_R\bar{\phi}\psi_L) \quad (1.81)$$

where  $k$  is a coupling constant, is invariant under a  $SU_L(2) \otimes U_Y(1)$  local gauge transformation. We can take once again the electron as an example and specify the terms in (1.81). We find that

$$\mathcal{L}_Y = -\frac{kev}{\sqrt{2}}(\bar{e}_Le_R + \bar{e}_Re_L) - \frac{keH}{\sqrt{2}}(\bar{e}_Le_R + \bar{e}_Re_L). \quad (1.82)$$

We thus find the Dirac mass term

$$m_e = \frac{kev}{\sqrt{2}} \quad (1.83)$$

as well as a term which couples the Higgs to the fermion field. This term is proportional to the fermion mass.



In contrast to the derivation of the gauge bosons' mass, the derivation of the fermionic masses is ad-hoc. The fermionic mass terms depends on the coupling  $k$  which must be measured from experiment. There is no explanation for the observed mass hierarchy of the fermions.

discutere come i neutrini rimangono fuori

## 1.6 The Standard Model Lagrangian

We are now ready to put all the ingredients discussed together and bake the cake that is the Standard Model. The full Lagrangian for the model is given by

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{SSB} + \mathcal{L}_Y. \quad (1.84)$$

$\mathcal{L}_{QCD}$  describes all strong interactions,  $\mathcal{L}_{EW}$  describes electroweak interactions,  $\mathcal{L}_{SSB}$  gives mass to the gauge bosons and  $\mathcal{L}_Y$  gives mass to the fermions. There are a few additional complications due to quark mixing and neutrino mixing which for the sake of brevity we shall pass over.

Figure 1.4 shows all the interactions allowed by the  $\mathcal{L}_{SM}$ .

We should mention that, in this form, the Lagrangian is purely classical: it must then be quantized and renormalized [7] in order to fully describe our quantum world.

Write out full Lagrangian (found online, just need to copy & paste, or is the image of the diagrams enough?

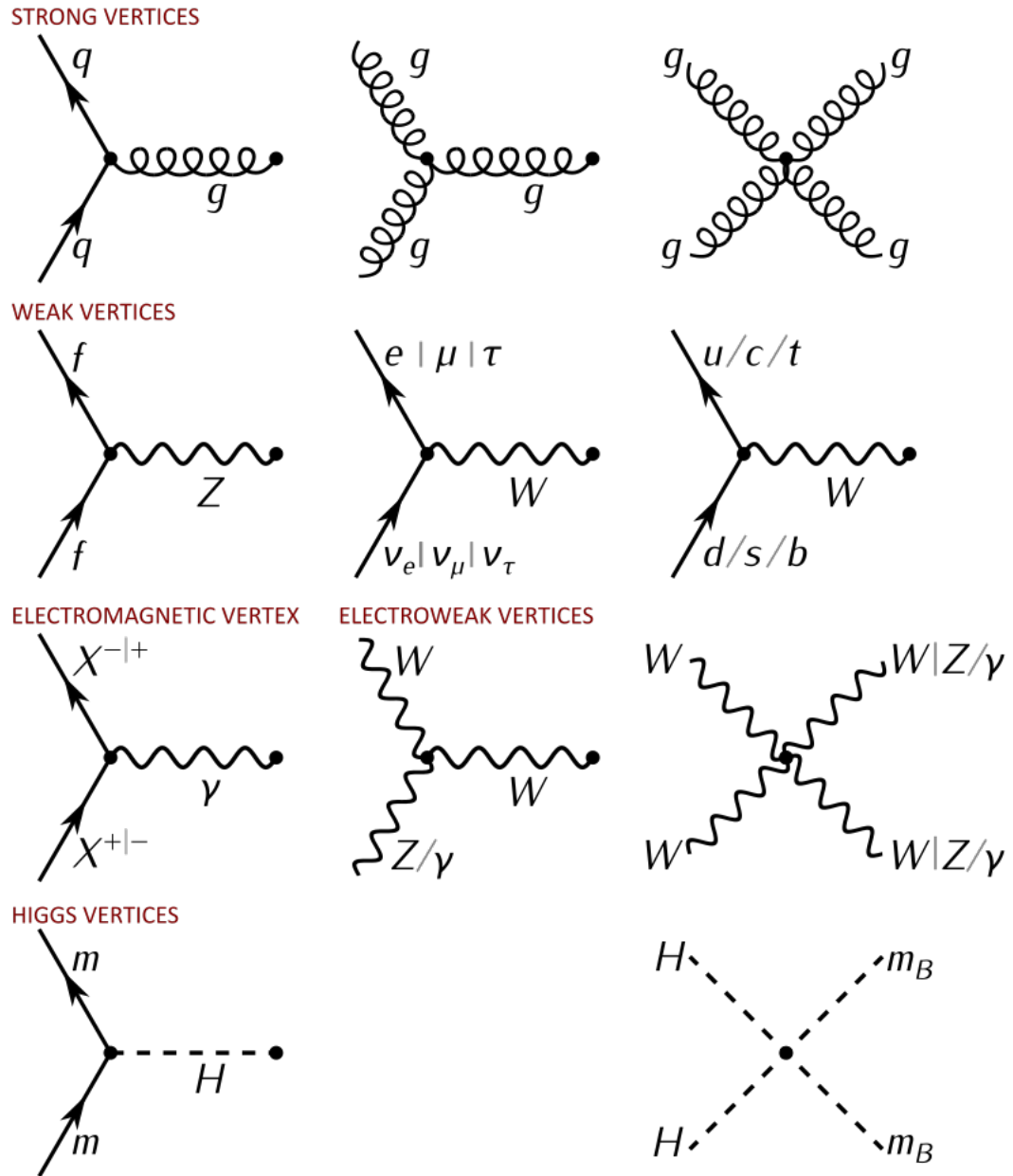


Figure 1.4: The Feynman diagrams of all interactions predicted by the Standard Model [1].  
 Change image because HHH HHH coupling is missing, for example

## Chapter 2

# Collisions at the LHC

One of the best ways to explore the Standard Model is through high-energy particle collisions. The LHC is a 27 km hadron-hadron circular collider where two protons or nucleons interact with each other in a multitude of ways, resulting in a myriad of possible final states. By identifying these final states, and selecting those which correspond to processes of interest, it is possible to study these processes in detail. A number of detectors, including ATLAS and CMS, lie on the beam pipe for this purpose.

The LHC has undergone several phases. During Run 1, the LHC ran at a center of mass energy  $\sqrt{s}$  of 7-8 TeV. The energy was increased to  $\sqrt{s} = 13$  TeV during Run 2. Run 2 delivered a total integrated luminosity of  $156 \text{ fb}^{-1}$ . During Run 3, set to begin next year, the  $\sqrt{s}$  could increase up to 14 TeV, and it is expected that the integrated luminosity will double to  $300 \text{ fb}^{-1}$ . In this chapter, we will describe the physics of hadron colliders.

### 2.1 Factorization

The proton is a dynamic system. In a simplistic view, it is composed of three valence quarks,  $u, u, d$  but these are bound together by gluons. The gluons interact both with the valence quarks and themselves, leading to a “sea” composed of gluons as well as quarks and anti-quarks of all flavors. This sea is dominant at low energies, and suppressed at higher energies.

When two hadrons collide at high energies, the resulting interaction does not directly involve the hadrons as a whole but the *partons* which constitute the hadron. Thanks to asymptotic freedom, these partons are quasi-free. This means that, even when the interaction with the hadron as a whole is deeply inelastic, the parton-parton interaction which occurs is instead elastic.

Relativistic considerations also allow us to deduce that the time-scale is such that only interactions with one parton per hadron are possible. Indeed, in the rest frame of the proton, the time-scale of the interactions holding the proton together are of the order  $1/m_p$ . In the laboratory frame of the collision, this is boosted by a factor  $\gamma = \sqrt{s}/2m_p$ . Since the energies available at the LHC place us firmly in the ultrarelativistic limit, interactions with a virtual particle of energy  $Q^2 \gg m_p^2$  occur on a time-scale much shorter than  $\gamma/m_p$ , the parton probed has no time to communicate with the other partons.

In the case of deep inelastic scattering, after the constituent parton has been struck, the virtual particles emitted by the constituent as part of normal interactions within the hadron can no longer be reabsorbed. This effect is exacerbated at higher  $Q$ . The end result is a perturbative evolution of the final state particles from the interaction, together with these liberated virtual

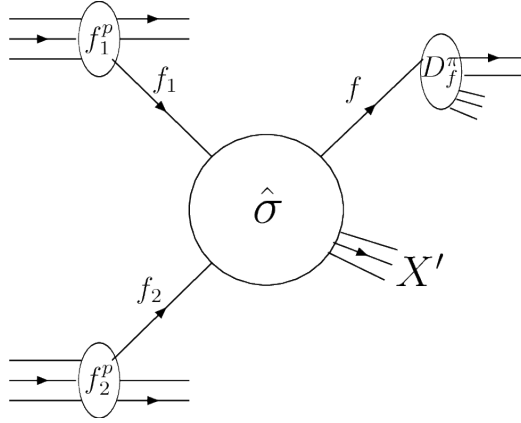


Figure 2.1: A schematic representation of factorization in a process which results in the production of pion in a proton-proton collision [12].

particles, down to energies of the order of the Landau pole of QCD,  $\Lambda_{QCD}$ . At these energies, due to the running of the coupling constant  $\alpha_S$ , it is no longer possible to describe QCD using perturbative physics. What follows is the hadronization process, where final state particles undergo non-perturbative interactions which transform them into relatively long-lived hadrons. Since the time-scale for hadronization is much longer when compared to the elementary process, the cross section for the collision is said to be *factorized* into a hard process described by an elementary cross section and functions describing the non-perturbative physics involved in the hadron before the interaction as well as the hadronization process.

This discussion can be neatly summarized in a formula:

$$\frac{d\sigma}{dX} = \sum_{j,k} \int dx_1 dx_2 f_j(x_1, Q^2) f_k(x_2, Q^2) \frac{d\hat{\sigma}_{jk}(x_1 P_1, x_2 P_2, Q^2, \mu_F^2)}{d\hat{X}} F(\hat{X} \rightarrow X; Q^2, \mu_F^2). \quad (2.1)$$

Here, we are stating that the differential cross section with respect to a hadronic observable  $X$  can be written in terms of the parton-level cross section  $d\hat{\sigma}/d\hat{X}$ , differential in the parton-level observable  $\hat{X}$ .  $f_j(x_1 P_1, \mu_F)$  and  $f_k(x_2 P_2, Q^2)$  are known as Parton Distribution Functions (PDFs) and, at the lowest order in perturbation theory, describe the probability of extracting a parton of type  $j$  or  $k$  with momentum fraction  $x_1$  or  $x_2$ , respectively from the colliding hadrons with momenta  $P_1$  and  $P_2$  when probed at energy  $Q^2$ . We must sum over all possible partons, and integrate over all momentum fractions. Finally, the function  $F(\hat{X} \rightarrow X; Q^2, \mu_F^2)$  describes the (non-perturbative) transition from partonic states to hadronic states and indicates how to relate any partonic observable  $\hat{X}$  with the measured hadronic observable  $X$ . This can be, for example, a Fragmentation Function, used when a single final-state hadron is observed, or a Jet Function, used instead when describing an aggregate transition to hadrons. We will describe all of the involved functions in more detail in subsequent sections. Figure 2.1 is a schematic representation of factorization in a proton-proton collision.

## 2.2 Parton Distribution Functions

As previously stated, Parton Distribution Functions contain information regarding the constituents of a hadron. To better understand them, we shall study in detail the process known as Deep Inelastic Scattering (DIS).

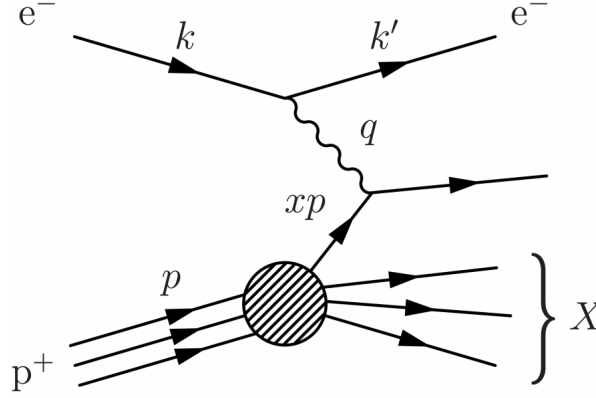


Figure 2.2: A schematic representaiton of the DIS process.

### 2.2.1 Deep Inelastic Scattering

DIS is a process which probes the insides of hadrons using leptonic probes. In this section we will focus on high-energy electron-proton inelastic scattering. In this process, the incoming electron exchanges a virtual photon (if below the production threshold of the massive gauge bosons) with the proton. Specifically, we will look at the proccess

$$e^- p \rightarrow e^- X \quad (2.2)$$

where X represents an undetermined hadronic final state, mediated by a virtual photon.

#### Kinematics

A number of kinematic variables are needed to fully describe the process. Without loss of generality, we will work in the targert rest frame (TRF), where the struck proton is at rest. With reference to Figure 2.2, we have

$$\begin{cases} k^\mu = (E, 0, 0, E) \\ p^\mu = (m_p, 0, 0, 0) \\ k^{\mu'} = (E', E' \sin \theta, 0, E' \cos \theta) \\ q^\mu = k^\mu - k^{\mu'} \end{cases} \quad (2.3)$$

where  $k^\mu$  and  $k^{\mu'}$  refer to the four-momenta of the initial and final state electron, respectively,  $p^\mu$  the four-momentum of the initial state proton and  $q^\mu$  to the four momentum of the virtual photon. The quantities  $E$  and  $E'$  refer to the respective energies of the electron in the initial and final state, and  $\theta$  is the angle at which the electron is scattered with respect to its initial momentum  $\vec{k}$ .

The first kinematic variable to consider is the energy of the virtual probe, known also as the scale of the process. Since DIS is a space-like process, the four-momentum of the virtual probe  $q^2 < 0$ , therefore we take

$$Q^2 = -q^2. \quad (2.4)$$

Next we introduce a series of Lorentz-invariant variables. The first variable,  $y$ , is defined as

$$y = \frac{p \cdot q}{p \cdot k} \stackrel{\text{TRF}}{=} 1 - \frac{E'}{E}. \quad (2.5)$$

As is clear, in the TRF,  $y$  represents the fraction of energy lost during the inelastic process. For this reason,  $0 \leq y \leq 1$ . Next, we introduce a similar variable

$$\nu = \frac{p \cdot q}{m_p} \stackrel{\text{TRF}}{=} E - E' \quad (2.6)$$

which in the TRF represents the the energy lost by the electron in the scattering process. Finally, we introduce the variable Bjorken  $x$ , defined as

$$x = \frac{Q^2}{2p \cdot q} \stackrel{\text{TRF}}{=} \frac{Q^2}{2m_p \nu} \quad (2.7)$$

If we consider the definition of the invariant mass of the system,

$$W^2 = (p + q)^2 = p^2 + 2p \cdot q + q^2 = m_p^2 + Q^2 \left( \frac{1}{x} - 1 \right) \geq m_p^2 \quad (2.8)$$

it becomes clear that  $x$  represents the “elasticity” of the process:  $x$  is limited to the range  $0 \leq x \leq 1$ , and  $x = 1$  corresponds to a perfectly elastic collision, whereas  $x = 0$  corresponds to a perfectly inelastic one. By definition, the Deep Inelastic limit of this process is the limit in which  $Q^2 \rightarrow \infty$  while  $x$  is held constant.

### Cross Section

It is possible to calculate the cross section for the inelastic scattering process described above. The most general Lorentz-invariant cross section for the interaction considered is

$$\frac{d^2\sigma}{dx dQ^2} = \frac{4\pi\alpha^2}{Q^4} \left[ \left( 1 - y - \frac{m_p^2 y^2}{Q^2} \right) \frac{F_2(x, Q^2)}{x} + y^2 F_1(x, Q^2) \right]. \quad (2.9)$$

In the deep inelastic limit, this simplifies to

$$\frac{d^2\sigma}{dx dQ^2} \approx \frac{4\pi\alpha^2}{Q^4} \left[ (1 - y) \frac{F_2(x, Q^2)}{x} + y^2 F_1(x, Q^2) \right]. \quad (2.10)$$

$F_1$  and  $F_2$  are *structure functions*, which describe the internal structure of the proton.

### Bjorken Scaling and the Callan-Gross Relation

An important observation involving the aforementioned structure functions is *Bjorken scaling*. This is the observation that, to first order,  $F_1$  and  $F_2$  are independent of  $Q^2$ , as shown in Figure 2.3. In addition to this scaling, experimental observations also found that, at sufficiently large  $Q^2$

$$F_2(x) = 2xF_1(x). \quad (2.11)$$

This relation is known as the Callan-Gross relation and is shown in Figure 2.4. We would expect to see scaling if the scattering occurred against point-like particles, giving evidence to the composite nature of the proton. In addition to this, the Callan-Gross relation tells us that the constituent partons carry spin-1/2. We therefore have experimental evidence to support the statement that the proton is composed of point-like spin-1/2 particles, namely quarks<sup>1</sup>. The

---

<sup>1</sup>The proton also has a significant gluon component, as will be discussed in a later section

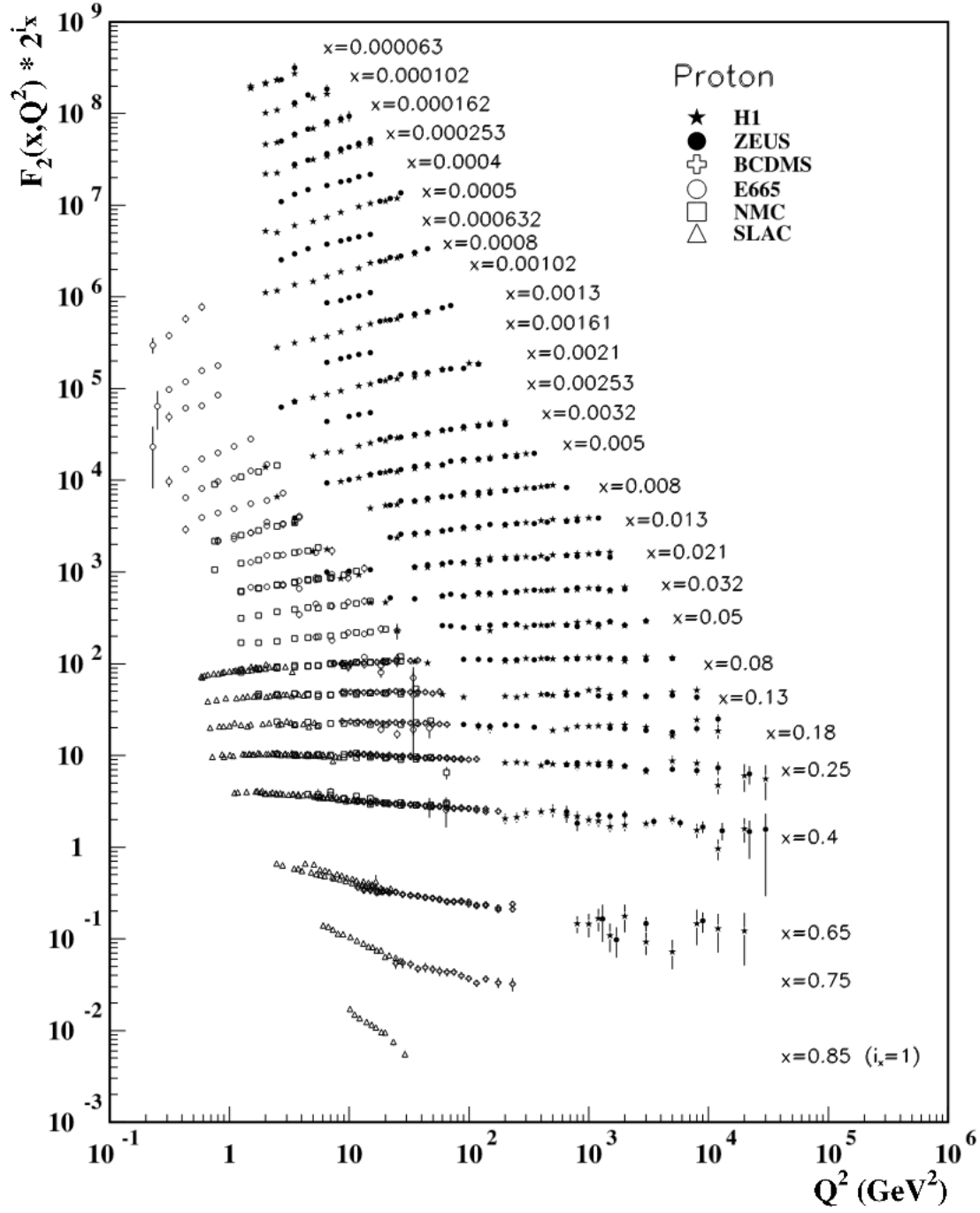


Figure 2.3: The observation of approximate scaling at various experiments. The observed violations are caused by higher-order contributions [13].

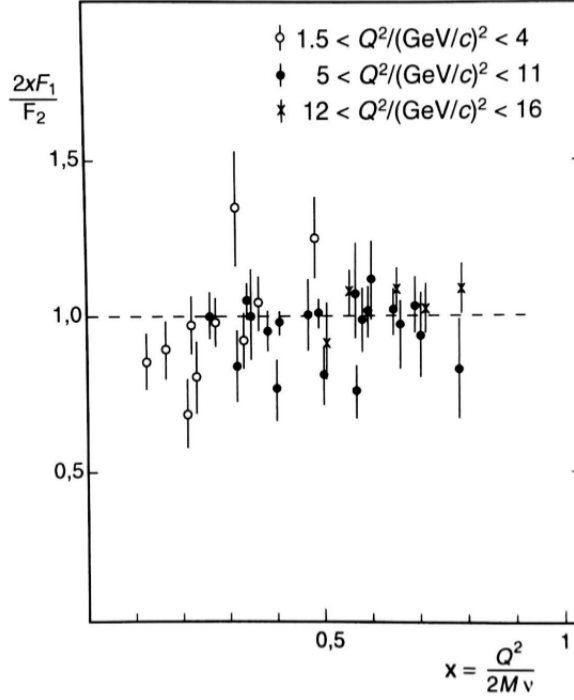


Figure 2.4: An experimental observation of the Callan-Gross relation - [Needs citation](#)

incoming electron elastically scatters against these constituents, explaining these observations and justifying the formula (2.1).

If we consider the DIS process in the infinite momentum frame, defined as the frame in which the energy of the proton in the initial state  $E_p \gg m_p$ , i.e.  $p^\mu = (E_p, 0, 0, E_p)$ , we can deduce that the four-momentum of the struck quark can be written as

$$p_q^\mu = (\xi E_p, 0, 0, \xi E_p) \quad (2.12)$$

where  $\xi$  represents the fraction of the proton's momentum carried by the quark. The quark in the final state of the electron-quark scattering will have four-momentum

$$(\xi p + q)^2 = \xi^2 p^2 + 2\xi p \cdot q + q^2 = m_q^2. \quad (2.13)$$

For this relation to hold, we must have  $2\xi p \cdot q + q^2 = 0$ , which implies that

$$\xi = \frac{-q^2}{2p \cdot q} = \frac{Q^2}{2p \cdot q} = x. \quad (2.14)$$

We can therefore identify  $x$  Bjorken with the momentum fraction of the struck constituent quark.

### Determining PDFs

We can now proceed in actually determining the constituents of the proton. Experimentally, this is done using formula (2.1), which in the case of DIS becomes

$$\frac{d\sigma}{dE' d\Omega} = \sum_f \int_0^1 dx \frac{d\hat{\sigma}}{dE' d\Omega}(xP, q) \phi_f(x). \quad (2.15)$$



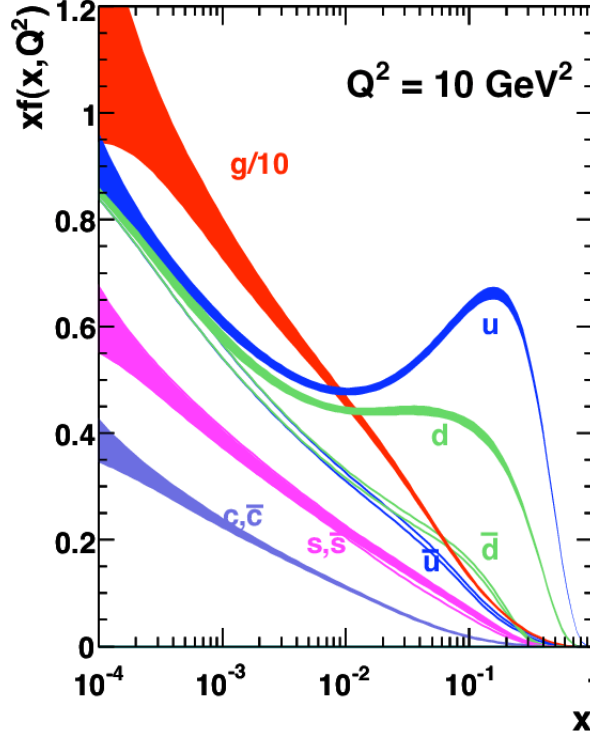


Figure 2.5: PDFs of the proton measured at various values of  $x$  at  $Q^2 = 10 \text{ GeV}^2$  [14].

The partonic cross section is easily calculable in QED

$$\frac{d\hat{\sigma}}{dE'd\Omega} = \frac{4\alpha^2}{Q^4} E'^2 \cos^2 \frac{\theta}{2} e_f^2 \frac{2m_f x}{Q^2} \delta(x' - x) \left[ 1 + \frac{Q^2}{2m_f^2} \tan^2 \frac{\theta}{2} \right]. \quad (2.16)$$

By inserting (2.10) and (2.16) in (2.15), we find that

$$\begin{cases} F_1(x) = \frac{1}{2} \sum_f e_f^2 \phi_f(x) \\ F_2(x) = x \sum_f e_f^2 \phi_f(x). \end{cases} \quad (2.17)$$

We have managed to describe the structure functions in terms of the PDFs, and in turn found the Callan-Gross relation!

By experimentally measuring the cross-section (2.15), we can deconvolve the PDF contribution. Figure 2.5 shows the PDFs of various partons within the proton, measured by ... at  $Q^2 = 10 \text{ GeV}^2$ .

PDFs have the notable property of being *universal*. This means that, regardless of what is used to probe them, they will always be the same since they are an intrinsic property of a given hadron. Given this property, determining the proton PDFs is fundamental for making accurate predictions at the LHC.

### Gluon PDF

The discussion up to now has focused exclusively on quarks, since, within the original framework of the so-called *Quark Parton Model*, did not include gluons. However, Figure 2.5 shows that, particularly at small  $x$ , a significant fraction of the proton's constituents are gluons.

The gluon PDF can be determined in a manner analogous to the quark PDFs. With DIS, we must, however, consider higher-order QCD corrections to be able to study elementary processes such as  $\gamma g \rightarrow q\bar{q}$ . This is not ideal since the process of interest is suppressed. Alternatively, we can look at hadron-hadron collisions, where gluon interactions do appear at leading order for some processes.

### PDF Evolution

In general, PDFs are functions both of the momentum fraction  $x$  and the scale of the hard process  $Q^2$ . To first order, this dependence is negligible, but the inclusion of the gluon leads to higher-order QCD corrections which lead to the scaling violations observed.

PDFs cannot be calculated from first principles; they must be measured from experiment. Thankfully, it is not necessary to perform measurements at different values of  $x$  and  $Q^2$  to fully determine PDFs.

The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) Equations [15, 16] allow us to calculate the evolution of a PDF from the scale  $Q^2$  to a different scale  $Q'^2$ . This is a notable achievement, as the PDFs can be measured at a certain energy, but re-utilized in a theoretical prediction for a process occurring at a different energy scale.

For the sake of brevity, we shall limit ourselves to citing this incredible result, rather than deriving the whole equation.

The evolution equation for the parton density is:

$$\frac{dq_f(x, Q^2)}{d \log Q^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} \left[ q_f(y, Q^2) P_{qq} \left( \frac{x}{y} \right) + g(y, Q^2) P_{gq} \left( \frac{x}{y} \right) \right]. \quad (2.18)$$

This equation states that, to calculate the PDF for a parton of flavor  $f$  at a given  $x$  and  $Q^2$  we need only to integrate the parton and gluon PDFs at  $y$  and  $Q^2$ , along with the splitting functions  $P_{qq}$  and  $P_{gq}$ .  $P_{qq}$  gives the probability of finding a real quark  $q$  with a certain momentum fraction after the emission of a virtual gluon, and  $P_{gq}$  is the analogous function for the emission of a real quark after a gluon splitting.

$$\begin{cases} P_{qq}(z) = C_F \frac{1+z^2}{(1-z)_+} \\ P_{gq}(z) = T_R [z^2 + (1-z)^2] \end{cases} \quad (2.19)$$

where  $C_F = (N_c^2 - 1)/2N_c$  is the Casimir invariant of  $SU(3)$ ,  $T_R$  is the trace of the Gell-Mann matrices and the plus-prescription, defined as

$$[f(x)]_+ = f(x) - \delta(1-x) \int_0^1 f(z) dz, \quad (2.20)$$

has been used to regularize the divergent integral.

An analogous function exists for the evolution of the gluon PDF

$$\frac{dg(x, Q^2)}{d \log Q^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} \left[ \sum_f q_f(y, Q^2) P_{qg} \left( \frac{x}{y} \right) + g(y, Q^2) P_{gg} \left( \frac{x}{y} \right) \right] \quad (2.21)$$

where this time

$$\begin{cases} P_{gq}(z) = C_F \left[ \frac{1+(1-z)^2}{z} \right] \\ P_{gg}(z) = 2C_A \left[ \frac{1-z}{z} + \frac{z}{(1-z)_+} + z(1-z) \right] \end{cases} \quad (2.22)$$

where  $C_A$  is the Casimir operator for the adjoint representation of  $SU(3)$ .

## 2.3 Fragmentation Functions

Fragmentation functions play a role similar to that of PDFs, but rather than describing the possible initial states of the interaction, they describe the possible final states that may be observed.

To understand Fragmentation Functions, it is the *semi-inclusive* DIS process, where we observe one hadron in the final state. The kinematic variables are the same as in (2.3), though we must now include the four-momentum of the final-state hadron

$$P_h^\mu = (E_h, \vec{P}_h) \quad (2.23)$$

and one additional Lorentz invariant variable

$$z_h = \frac{P \cdot P_h}{P \cdot q} \stackrel{\text{TRF}}{=} \frac{E_h}{\nu} \quad (2.24)$$

which in the TRF represents the fraction of energy lost in the hadronization process of the observed particle.

The cross section can again be calculated without difficulty:

$$\frac{d\sigma}{dx dy dz} = \frac{4\pi\alpha_S^2}{Q^4} \left( \frac{y}{2} + 1 - y \right) x \sum_f e_f^2 \phi_f(x) D_f(z) \quad (2.25)$$

where the sum over all flavors includes the anti-quarks. In this way, the Fragmentation Function  $D_f(z)$  represents the probability of finding a hadron with fraction  $z$  of the available energy. The cross section remains factorized, allowing for information on the Fragmentation Functions by comparison with data, as neither they are calculable in QCD. Finally, it should be said that the Fragmentation Functions depend on the quark from which they originate, hence the index.

**Togliere questa parte sulle Fragmentation functions? Non e' molto pertinente ma faccio un accenno piu' avanti**

## 2.4 Jets

Due to the non-perturbative nature of the hadronization process and the myriad of particles produced, it is not possible to calculate cross-sections for all of the different possible hadronic final states. Instead, *jets* are considered in their place. Jets are, as the name suggests, groups of collimated particles. They can be treated as singular objects, allowing for a notable simplification of the calculations.

### 2.4.1 Jet Definitions

In order to be able to precisely calculate jet cross-sections and confront these with experimental data, it is necessary to unambiguously *define* what a jet is. There are multiple possible definitions of jets, and we will briefly give some examples of these.

### Cone Algorithms

Historically, cone algorithms were the first class of jet algorithms introduced. The very first algorithm was used to classify jets in  $e^+e^-$  collisions and depended on two arbitrary parameters,  $\delta$  and  $\epsilon$ . If an event had a fraction of energy of at least  $1 - \epsilon$  concentrated within two cones of half-angle  $\delta$ , then that event was said to contain two jets [17]. The fact that the definition relied on arbitrary parameters meant that these had to be specified so that the predictions could be compared to data. This remains a general feature of jet algorithms to this day.

Cone algorithms have progressed in the years since their inception. Today, two of the most widely used code algorithms are iterative and fixed cone algorithms. In iterative cone algorithms, the direction of the jet is initially set by a particle  $i$ . All particles within distance  $R_{ij}$

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 < R_{ij}^2 \quad (2.26)$$

in the rapidity/azimuthal angle plane are then taken as part of the jet, and their momenta are summed. The result of this summation is then used as the new seed, and the process is iterated until the jet cone is stable. To fully specify the algorithm, one must choose how to take the seed, and what to do when jets overlap.

Fixed cone algorithms function similarly. In this case, rather than iterating the cone direction, a cone is fixed around a seed, and that cone is called a jet. The particles within the radius  $R_{ij}$  are assigned to the jet, and removed from the event record. The algorithm proceeds until all possible jets have been identified.

### $k_t$ -Algorithms

$k_t$  algorithms are part of a family of algorithms known as sequential recombination jet algorithms. The  $kt$  algorithms use a momentum-weighted distance

$$\begin{cases} d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \\ d_{iB} = k_{ti}^{2p} \end{cases} \quad (2.27)$$

to establish which particles  $j$  lie closest to the particle  $i$ . If  $d_{ij}$  is less than the distance between  $i$  and the beam  $d_{iB}$ ,  $i$  and  $j$  are combined into a jet. This procedure iterates over all particles.

These definitions also depend on two parameters:  $p$  and  $R$ . There exist three noteworthy cases:  $p = 1$  is known as the  $k_t$  algorithm, and weighs the distance  $d_{ij}$  using the square of the transverse momentum of the softer particle;  $p = 0$  is known as the Cambridge-Aachen algorithm, and features no weighting;  $p = -1$  is known as the anti- $k_t$  algorithm, and weights the distance using the inverse of the square of the transverse momentum of the harder particle.

Due to the distance used, the anti- $k_t$  algorithm tends to cluster soft particles together with hard particles. This is a useful property as it tends to lead to jets centered about hard particles, and correctly recombines the soft radiation emitted from the hard seed together with that seed. If the hard particles are well-separated, this also leads to conical jets, as shown in Figure 2.6.

### 2.4.2 Jet Cross Sections

Once we have defined our jets, we can go on to calculate jet cross sections. Without going into too much detail, we will just illustrate a general feature which highlights the usefulness of jets.

If we calculate the two-hadron semi-inclusive cross section for  $e^+e^-$  into hadrons, we find that

$$\frac{d\sigma}{dydz_1dz_2} = N_c \frac{\pi\alpha^2}{Q^2} (1 + \cos\theta) \sum_f e_f^2 D_f(z_1) D_f(z_2), \quad (2.28)$$

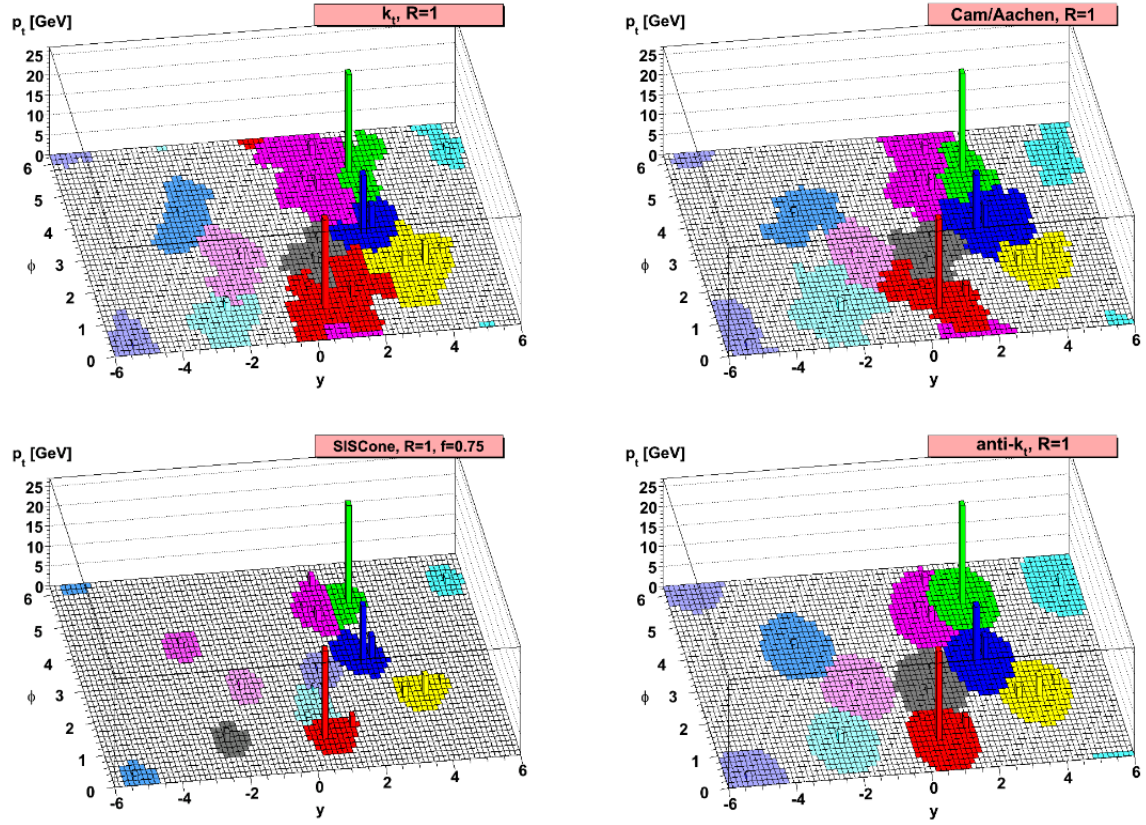


Figure 2.6: A representation of the jets formed from the same event using the three main  $k_t$  algorithms as well as SIScone, a commonly used cone algorithm. The jets clustered using anti- $k_t$  are conical, and centered around the hard particles [18].

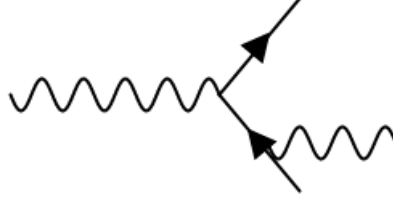


Figure 2.7: A Feynman diagram depicting a real emission in QED.

where again the sum over the flavors  $f$  runs over both quarks and antiquarks.  $N_c$  stands for the number of colors that can be produced, and  $\theta$  is the angle between the momentum of the quarks produced and the momentum of the colliding leptons.

If rather than observing the two final-state hadrons we observe the jets surrounding them, the Fragmentation Functions are replaced by  $\delta$ -functions  $\delta(1 - z_1)$  and  $\delta(1 - z_2)$ . After integrating, we find that

$$\frac{d\sigma^{jets}}{dy} = N_c \frac{\pi\alpha^2}{Q^2} (1 + \cos\theta) \sum_f e_f^2. \quad (2.29)$$

*This is exactly the QED cross section for  $e^+e^-$  annihilation!* The use of jets allows us to notably simplify calculations and make accurate predictions, since we neither have to consider fully inclusive cross sections, nor a fully exclusive one.

## 2.5 Higher Order Corrections

A number of higher order corrections to the elementary cross section are possible. These include real and virtual corrections in all possible combinations. In QCD, we currently know how to treat corrections up to  $\mathcal{O}(\alpha_s^3)$  for a few select processes [citation](#). In this section we will briefly discuss the importance of higher order corrections, as well as some challenges which arise during their calculation.

### 2.5.1 Infrared and Collinear Divergences

The matrix element for a radiative correction in QED and QCD is easily calculable using the Feynman rules. For example, for the diagram shown in Figure 2.7, the matrix element for all possible real emissions and all possible photon polarizations is found to be

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_0 \frac{\alpha}{2\pi} \left[ 2 \frac{p_- \cdot p_+}{p_- \cdot k p_+ \cdot k} - \frac{m_e^2}{(p_- \cdot k)^2} - \frac{m_e^2}{(p_+ \cdot k)^2} \right] \frac{d^3k}{\omega} \quad (2.30)$$

where  $\left(\frac{d\sigma}{d\Omega}\right)_0$  represents the tree-level cross section. In the center-of-mass frame, where the kinematic variables can be expressed as

$$\begin{cases} p_-^\mu = (E, 0, 0, \beta E) \\ p_+^\mu = (E, 0, 0, -\beta E) \\ k^\mu = (\omega, \omega \sin\theta \cos\phi, \omega \sin\theta \sin\phi, \omega \cos\theta) \end{cases} \quad (2.31)$$

where  $\beta = \sqrt{1 - \frac{m_e^2}{E^2}}$ , (2.30) becomes

$$\frac{d\sigma}{d\Omega} = \left( \frac{d\sigma}{d\Omega} \right)_0 \frac{\alpha}{2\pi} \frac{1}{\omega} \frac{\beta^2 \sin^2 \theta}{(1 - \beta^2 \cos^2 \theta)^2} d\omega d\cos\theta \quad (2.32)$$

where we have already integrated in  $d\phi$ . This expression is divergent in two cases: when the radiated particle is either very soft ( $E \ll 1$ ) or when it is collinear to the lepton from which it is emitted ( $\theta \ll 1$ ). These cases are known as *infrared* and *collinear* divergences, respectively, and they are problematic as they prohibit the calculation of the cross section at next-to-leading order.

Integrating in  $d\cos\theta$  resolves the collinear divergence

$$\frac{d\sigma}{d\Omega} = \left( \frac{d\sigma}{d\Omega} \right)_0 \frac{1}{\omega} \frac{2\alpha}{\pi} \left[ \frac{\beta^2 + 1}{2\beta} \ln \left( \frac{1 + \beta}{1 - \beta} \right) + 1 \right] d\omega \quad (2.33)$$

which, when we consider that  $s \gg m^2$  simplifies to

$$\frac{d\sigma}{d\Omega} = \left( \frac{d\sigma}{d\Omega} \right)_0 \frac{1}{\omega} \frac{2\alpha}{\pi} \left[ \ln \frac{s}{m_e^2} - 1 \right] d\omega. \quad (2.34)$$

Lastly, we must integrate over all energies  $d\omega$ . However, because of finite detector resolution, the experimental cross section can only be sensitive to those photons radiated over a threshold  $\Delta E$ , the experimental cross section contains two parts

$$\left( \frac{d\sigma}{d\Omega} \right)_{exp} = \left( \frac{d\sigma}{d\Omega} \right)_{elastic} + \left( \frac{d\sigma}{d\Omega} \right)_{\omega < \Delta E} \quad (2.35)$$

one corresponding to the tree-level cross section, and the other with a radiative correction with  $\omega < \Delta E$ . When we integrate over  $d\omega$ , the second term on the right-hand side diverges as can be seen from (2.34).

Thankfully, a brilliant solution to this problem exists. If we assign a mass  $\lambda$  to the radiated photon, the (2.34) becomes

$$\left( \frac{d\sigma}{d\Omega} \right)_{exp} = \left( \frac{d\sigma}{d\Omega} \right)_{elastic} + \left( \frac{d\sigma}{d\Omega} \right)_{\omega < \Delta E} \ln \frac{\Delta E}{\lambda}. \quad (2.36)$$

If we also consider the virtual corrections, such as in Figure 2.8 and calculate the relative cross section, which also contributes to (2.35), we find

$$\frac{d\sigma}{d\Omega} = \left( \frac{d\sigma}{d\Omega} \right)_0 \frac{1}{\omega} \frac{2\alpha}{\pi} \left[ \ln \frac{s}{m_e^2} - 1 \right] \ln \frac{\lambda}{E} d\omega \quad (2.37)$$

We can sum the real and virtual corrections, and find a total cross section independent from  $\lambda$ ! We can now safely take the limit of  $\lambda \rightarrow 0$  and integrate over  $\omega$ .

### 2.5.2 Infrared and Collinear Safety for Jets

As mentioned in the previous section, infrared and collinear (IRC) divergences will, in theory, exactly cancel. In practice, however, this is not always the case. Depending on the type of jet algorithm used, it may happen that the cancellation breaks and an infinite cross section is calculated. Obviously, this is a problem since the measured cross section will always be finite. It is therefore important from a theoretical standpoint to define in a way that is IRC safe.

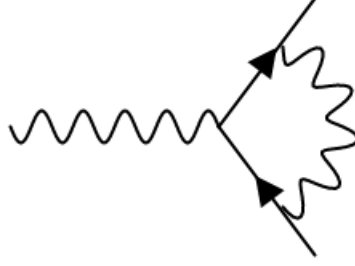


Figure 2.8: A Feynman diagram depicting a virtual correction in QED.

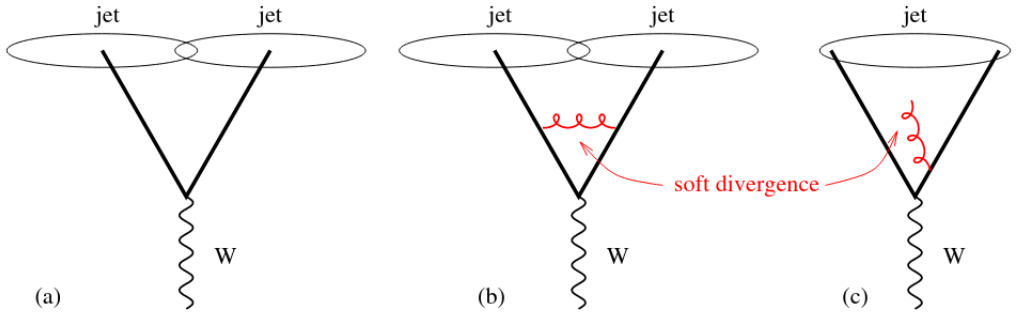


Figure 2.9: A schematic representation of an IR unsafe jet. The emission of a soft gluon changes the event from a two-jet event to a one-jet event due to its effect on the clustering algorithm [19].

An observable  $\mathcal{O}$  is said to be IRC safe, respectively, when the following properties are satisfied:

$$\mathcal{O}(X; p_1, \dots, p_n, p_{n+1} \rightarrow 0) \rightarrow \mathcal{O}(X; p_1, \dots, p_n) \quad (2.38)$$

$$\mathcal{O}(X; p_1, \dots, p_n \parallel p_{n+1}) \rightarrow \mathcal{O}(X; p_1, \dots, p_n), \quad (2.39)$$

i.e. when the observable reduces to that of  $n$  particle case, in absence of the soft or collinear emission. An example of an infrared safe and unsafe jet is shown in Figure 2.9, while Figure 2.10 illustrates an example of a collinear safe and unsafe jet.

The anti- $k_t$  algorithm is an example of a jet algorithm which is IRC safe. Due to its weighting, when a soft particle is emitted, this will tend to cluster together with the hard center of the jet. On the other hand, when a particle is emitted collinear to another, the distance  $\Delta_{ij}$  will be very small, and it will again be clustered together with its parent particle.



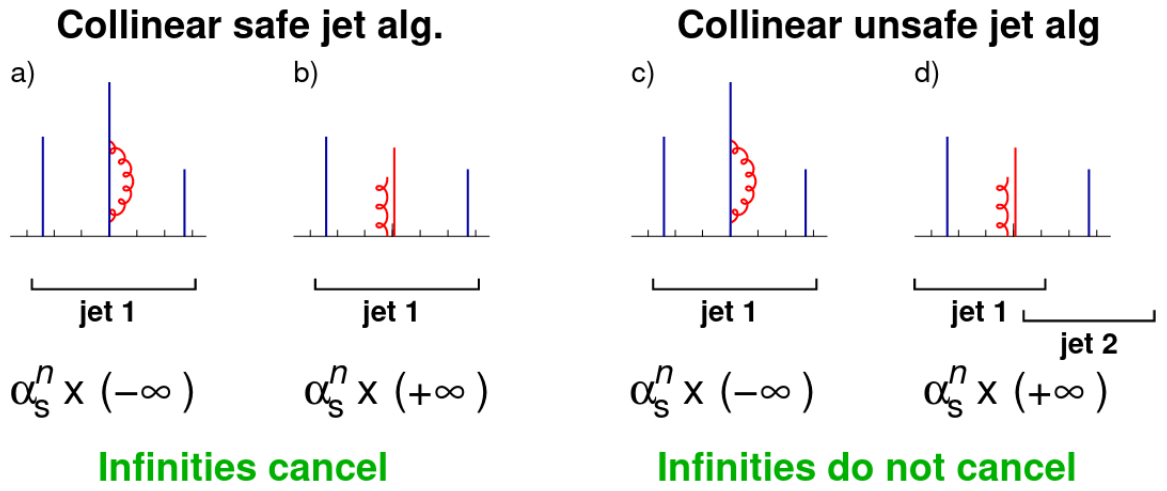


Figure 2.10: A schematic representation of a collinear safe and unsafe jet. The blue lines represent particles, and the length corresponds to the particle's transverse momentum. The horizontal axis represents rapidity. If the emission of a collinear gluon leads to the formation of a second jet, the cancellation is broken and the cross section diverges, as shown on the right. If, on the other hand, it and its parents are treated as a single particle by, for example, summing their momenta, the emission satisfies the definition 2.38 [19].



## Chapter 3

# Higgs Physics

The discovery of the Higgs Boson in 2012 [20] ushered in a new era of physics. Since then, one of the main goals of the ATLAS and CMS programs has been to robustly test the Standard Model by studying the properties of the Higgs. This requires knowledge on the production and decay mechanisms of the Higgs.

These mechanisms could be studied as a function of the Higgs mass  $m_H$  even prior to the discovery of the Higgs. Now that  $m_H$  is known, the cross sections depend exclusively on  $\sqrt{s}$ .

### 3.1 Higgs Production Mechanisms

At the LHC, there are four different processes which can produce the Higgs. Their cross sections at  $\sqrt{s} = 13$  TeV as a function of mass are shown in Figure 3.1.

#### Gluon-Gluon Fusion

Gluon-gluon fusion is the process which dominates Higgs production at the LHC. In the Standard Model, there is no direct coupling between the Higgs and the gluon, though an indirect coupling is possible through a virtual top quark loop, as shown in Figure 3.2. The top quark loop is favored is due the high mass of the top, though there is a small contribution from the bottom.

#### Vector Boson Fusion

Vector boson fusion is the second most important cross section, responsible for about 10% of the cross section.

Figure 3.3 shows the Feynman diagram for the process. The initial-state quarks deviate only slightly from their initial direction, leading to two jets near the beam axis in opposite regions of the detector. This provides an experimental signature of the process.

#### Higgstrahlung

Higgstrahlung, named after bremsstrahlung, is another process which relies on the Higgs coupling to two vector bosons. In this case, the Higgs radiates off a virtual  $Z$  or  $W$ . Experimentally, this channel can be identified by the decay of vector boson into leptons. Figure 3.4 shows the Feynman diagram for this process.

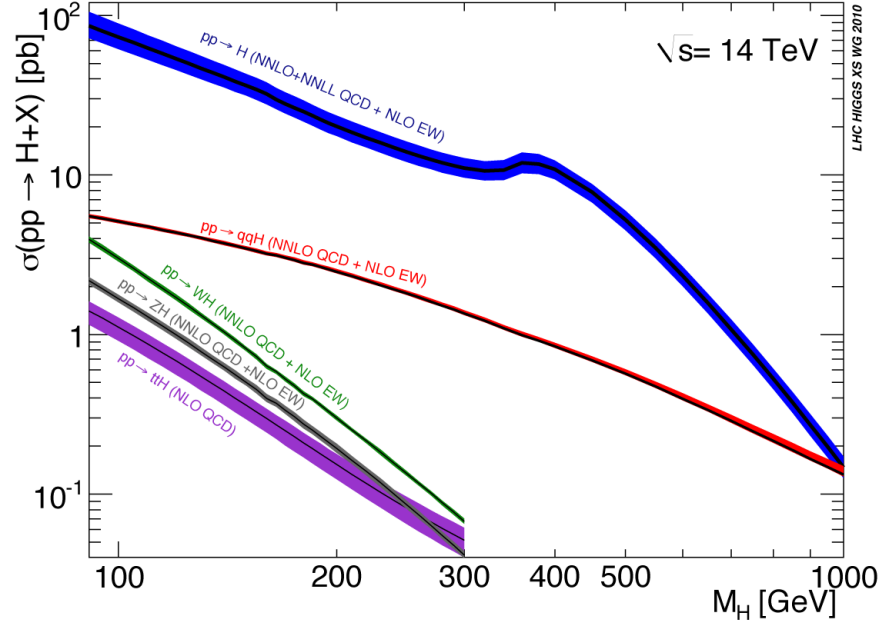


Figure 3.1: The cross sections of various Higgs production mechanisms as a function of  $m_H$  at LHC energies [cite twiki](#). The thickness of the lines represent various theoretical uncertainties.

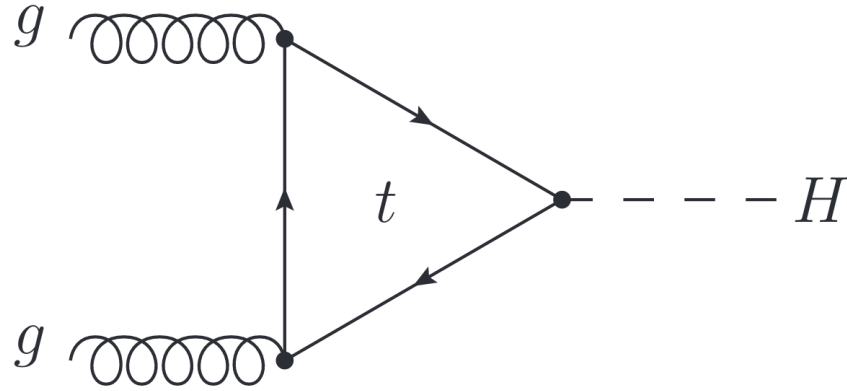


Figure 3.2: The Feynman diagram for gluon-gluon fusion.

### Heavy Quarks Associated Production

The last production channel of the Higgs is  $q\bar{q} \rightarrow q\bar{q}H$ , as shown in Figure 3.5. Since again the Higgs coupling favors heavy masses, the top quark dominates this channel. However, the coupling is also possible for other heavy quarks, such as the bottom. This channel provides a direct way to measure the Yukawa coupling.

[unify diagrams into single figure?](#)

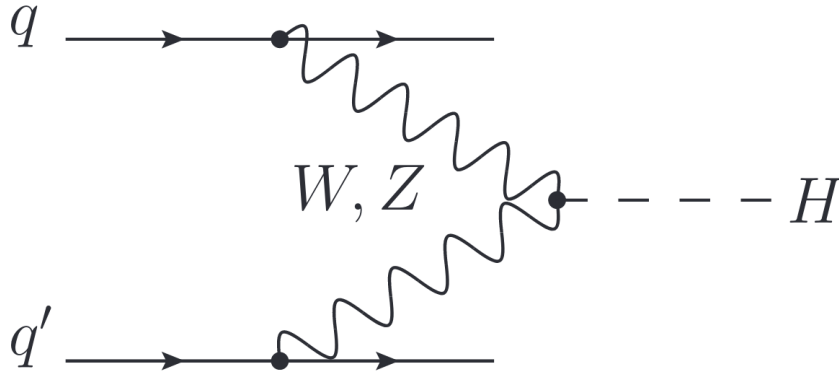


Figure 3.3: A Feynman diagram depicting vector boson fusion.

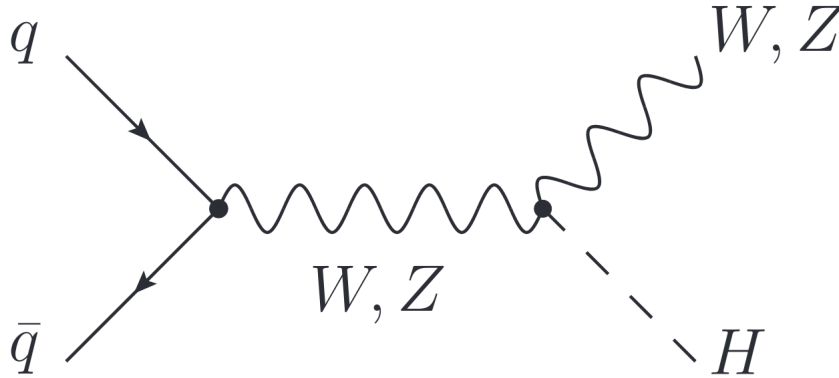


Figure 3.4: A Feynman diagram depicting the Higgstrahlung process.

## 3.2 Higgs Decay Mechanisms

As opposed to the Higgs production mechanisms, which depend in part on the structure of the hadrons used in collisions, Higgs decay mechanisms depend exclusively on the properties of the Higgs. The branching ratios of the various decay channels, again as a function of  $m_H$  are shown in Figure 3.6. Since they are many, we will only focus on a few key processes.

$$H \rightarrow b\bar{b}$$

The decay of the Higgs to  $b\bar{b}$  is by far the most important decay channel. Despite the fact that the Higgs couples more strongly to the top, the decay is not possible because of conservation of energy.

The extremely large branching ratio means that, experimentally, it is of utmost importance to be able to observe these decays if we want to study the properties of the Higgs in detail. To this aim, flavor tagging algorithms are a fundamental tool in determining whether a jet originates from a b-quark, a c-quark,  $\tau$ , or a light-quark or gluon.

A jet originating from a heavy quark can be identified based on some unique properties which

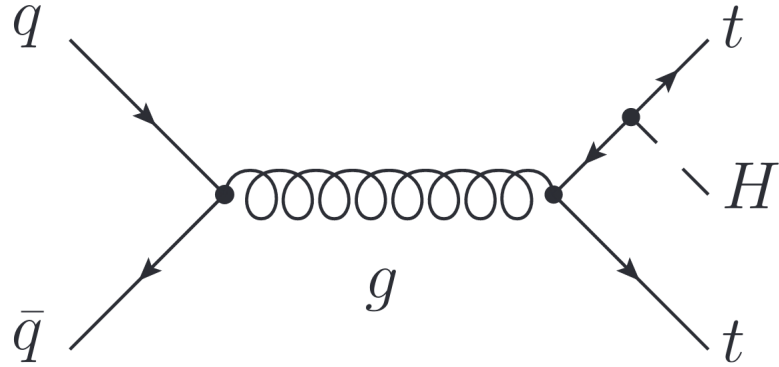


Figure 3.5: Higgs production in association with heavy quarks *sistemare linea fermionica*.

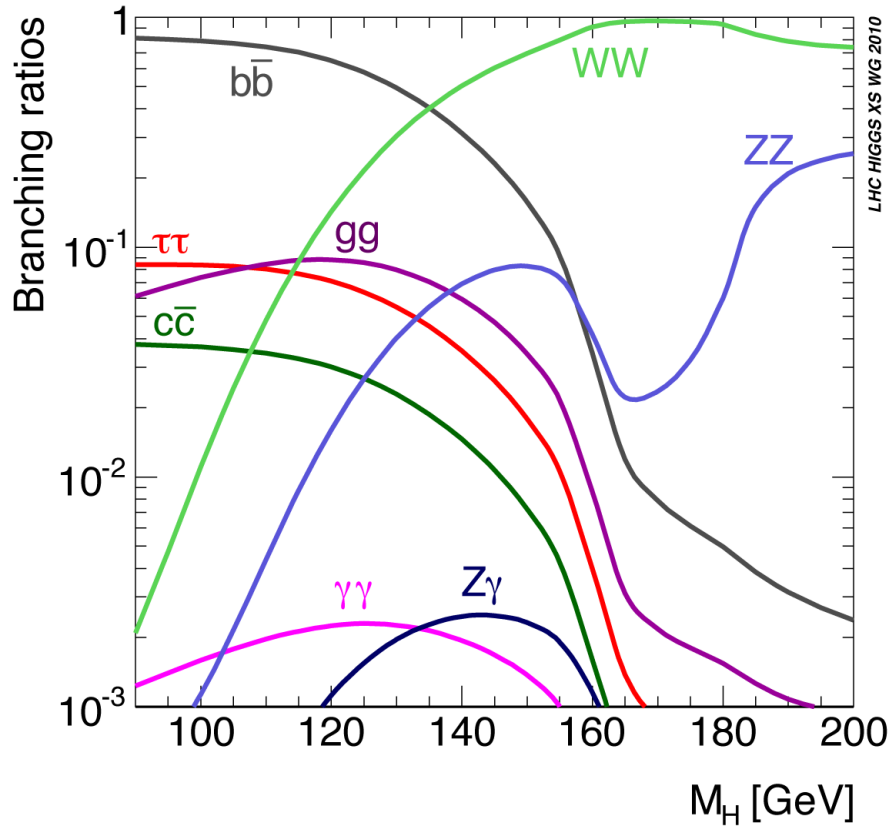


Figure 3.6: The branching ratios of the various decay channels of the Higgs boson as a function of  $m_H$  *cite Dani*.

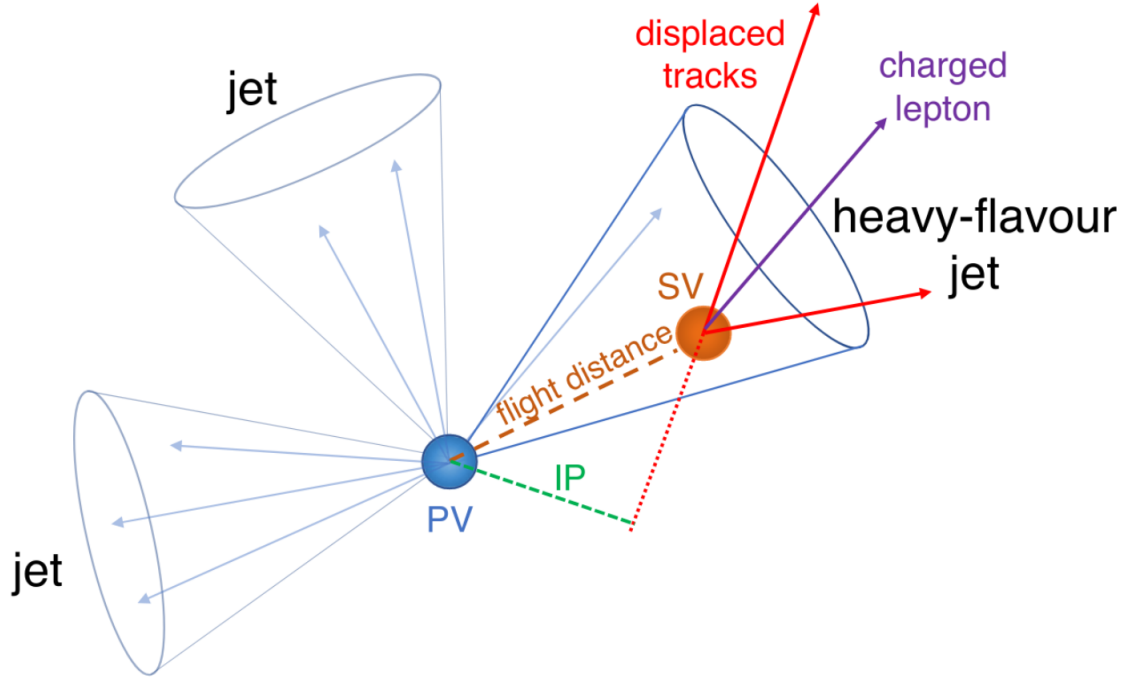


Figure 3.7: A caption [21]

stem from the high mass of the quark. For example, the relatively long lifetime of b-hadrons, of the order of 1.5 ps, combined with the high energies involved result in the formation of a secondary vertex, which can distance anywhere from a few hundred  $\mu\text{m}$  up to  $\sim 1$  cm from the primary vertex. This secondary vertex gives rise to displaced tracks with respect to the primary vertex, from which the secondary vertex can be recognized. It is also possible to measure the “mass” of the secondary vertex, which will be related to the mass of the b-hadron. The decay products of a b-hadron are characterized by a large transverse momentum with respect to jet axis when compared to other jet constituents. In addition to this, in about 20% of cases, the decay of the b-hadron at the secondary vertex results in an soft electron or muon, whose properties allow for the selection of a pure sample of b-jets. The combination of this information is used as input to algorithms, which can either be classical or make use of machine learning techniques, in order to determine with what likelihood a given jet is a b-jet. Figure 3.7 illustrates the geometry of these secondary vertices.

Because these tagging techniques are general, similar considerations hold for  $H \rightarrow \tau^+\tau^-$  and  $H \rightarrow c\bar{c}$ . Although  $H \rightarrow gg$  also leads to two jets, these cannot be effectively discriminated from the QCD background.

$$H \rightarrow WW, H \rightarrow ZZ, H \rightarrow \gamma\gamma$$

The large QCD background is the foremost obstacle that physicists must contend with when make discoveries at the LHC. For this reason, the decay channels  $H \rightarrow WW$ ,  $H \rightarrow ZZ$  and  $H \rightarrow \gamma\gamma$  were crucial in the discovery of the Higgs boson. These channels are characterized by a clean experimental signal, thanks to the possibility of leptonic decays and the ease with each photons can be identified. Specifically, the channels  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ ,  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow WW^{(*)} \rightarrow e\nu\mu\nu$ ,

when combined with data from the channels  $H \rightarrow b\bar{b}$  and  $H \rightarrow \tau^+\tau^-$  allowed for the discovery of the Higgs in 2012 by the ATLAS and CMS collaborations. The infamous plots are shown in Figure ... **quanti e quali mettere? Ogni canale ha il proprio plot** .

### 3.3 Properties of the Higgs

Measured mass, width, ecc.



## Chapter 4

# A Machine Learning Study of Color Sensitive Observables

As discussed previously, the  $H \rightarrow b\bar{b}$  decay is fundamental to unveiling the properties of the Higgs due to its large branching ratio. In this chapter, we will describe a method that we have developed to discriminate b-jets originating from a color singlet, such as the Higgs boson, from those originating from a color octet.

We have used simulated data to train two (three? Ma forse e' meglio lasciare Keras a Luca) different machine learning algorithms on high-level, color-sensitive variables which have been introduced in literature. We will first introduce these variables, and then go on to discuss the details of the simulation, the analysis of the simulated events, and our findings.

### 4.1 The Observables

In total, this method makes use of 8 variables. These will be described below.

#### 4.1.1 The Pull Variables

The first variables we will consider are related to the pull vector. With reference to Figure 4.1, consider an event in which two jets,  $J_a$  and  $J_b$ , with centers  $(y_a, \phi_a)$  and  $(y_b, \phi_b)$ , respectively, are emitted. The pull vector  $\vec{t}$  relative to jet  $J_a$  is defined as

$$\vec{t} = \frac{1}{p_{ta}} \sum_{i \in J_a} p_{ti} |\vec{r}_i|^2 \hat{r}_i \quad (4.1)$$

where  $p_{ta}$  is the transverse momentum of the jet, and the sum runs over all the the jet constituents.  $y$  and  $\phi$  again represent rapidity and azimuthal angle, and  $\vec{r}_i$  is the distance vector between the jet and its  $i$ -th constituent in the  $y$ - $\phi$  plane

$$\vec{r}_i = (y_i - y_a, \phi_i - \phi_a). \quad (4.2)$$

In particular, we would like to consider the projections of  $\vec{t}$  along two lines: one, generated by the unit vector which points from the center of  $J_a$  to the center of  $J_b$  of the two jets

$$\hat{n}_{\parallel} = \frac{1}{\sqrt{\Delta y^2 + \Delta \phi^2}} (\Delta y, \Delta \phi), \quad (4.3)$$

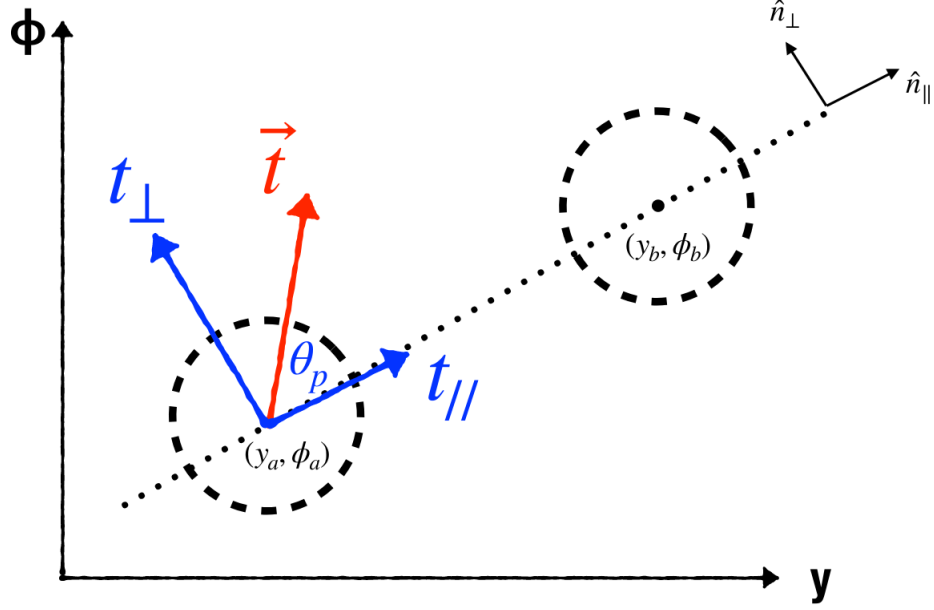


Figure 4.1: A caption

and the other generated by the unit vector perpendicular to  $\hat{n}_{\parallel}$

$$\hat{n}_{\perp} = \frac{1}{\sqrt{\Delta y^2 + \Delta \phi^2}} (-\Delta \phi, \Delta y), \quad (4.4)$$

i.e.

$$t_{\parallel} = \vec{t} \cdot \hat{n}_{\parallel} \quad (4.5)$$

$$t_{\perp} = \vec{t} \cdot \hat{n}_{\perp} \quad (4.6)$$

We would also like to consider  $\theta_p$ , known as the pull angle, defined as

$$\theta_p = \arccos \frac{t_{\parallel}}{|\vec{t}|}. \quad (4.7)$$

The pull vector is sensitive to the different color connections which are present in the decay of a color singlet and the decay of a color octet. An illustration of this feature is shown in Figure 4.2. Because of the different color connections, the radiation from jets originating from a singlet decay will tend to be emitted between the two jets, causing the pull vector of  $J_a$  to point in the direction of  $J_b$  and viceversa. In the octet case, the pull vectors will instead point in opposite (random?) directions. Of all the variables considered, the pull angle has been shown to be the most effective discriminator of the two different color configurations, [22]. However, there was a noticeable mismatch between theoretical predictions for  $\theta_p$  and experimental measurements [23].

The theoretical difficulties in calculating  $\theta_p$  stem from the fact that it is not an IRC safe variable. However,  $t_{\parallel}$  and  $t_{\perp}$  are IRC safe observables.

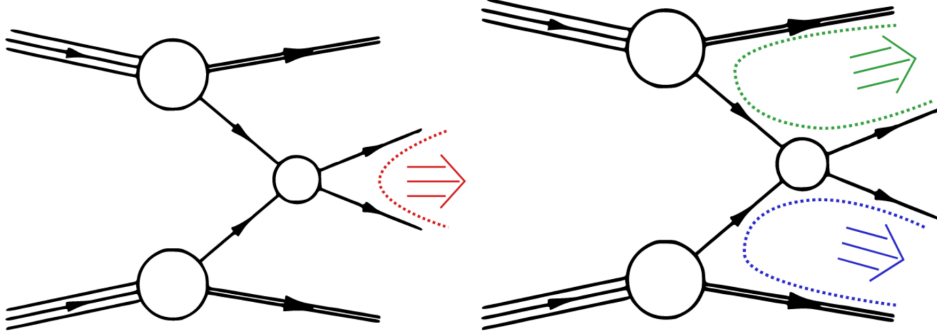


Figure 4.2: A representation of the color flow for the decay of a color singlet (left) vs. a color octet (right) at a collider experiment. In the case of a color singlet, the two colored particles stemming from the decay are color-connected only to each other, while for the case of an octet, the particles are color-connected to the rest of the proton, as it is from there that their color originates [22].

Our method includes these two variables for this purpose.  $\theta_p$  is also included since it can be measured at experiments, and its discrimination potential is still good. We consider these variables for both jets  $J_a$  and  $J_b$ .

#### 4.1.2 $D_2$

To introduce the observable  $D_2$ , we must first familiarize the reader with energy correlator functions (ECF) [24]. These are a class of observables sensitive to jet substructure. Specifically, they are sensitive to a jet's *prongedness*, or how many distinct subjets compose a given jet. This is particularly useful in the boosted regime, defined by  $p_T \gg 1$ , kinematics forces jets which would otherwise be separated to come together, as shown in Figure 4.3. The  $(N + 1)$ -point correlator is used to determine if a jet has  $N$  prongs.

The definition of the  $N$ -point ECF is based on the the  $p_T$  and angular distance between the components of a jet:

$$ECF(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left( \prod_{a=1}^N p_{Ti_a} \right) \left( \prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta, \quad (4.8)$$

where  $\beta$  is an arbitrary parameter. If  $\beta > 0$ , the observable is IRC safe. For the sake of clarity, let us write out the first few ECFs:

$$\begin{aligned} ECF(0, \beta) &= 1 \\ ECF(1, \beta) &= \sum_{i \in J} p_{Ti} \\ ECF(2, \beta) &= \sum_{i < j \in J} p_{Ti} p_{Tj} R_{ij}^\beta \\ ECF(3, \beta) &= \sum_{i < j < k \in J} p_{Ti} p_{Tj} p_{Tk} (R_{ij} R_{ik} R_{jk})^\beta. \end{aligned}$$

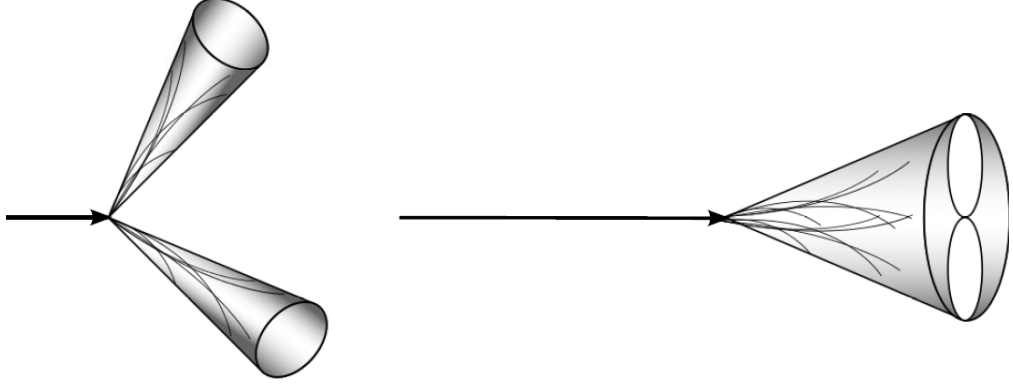


Figure 4.3: An illustration of a 2-prong jet caused by boosted kinematics [cite https://slidetodoc.com/download.php?id=7569163](https://slidetodoc.com/download.php?id=7569163).

There exists also a normalized  $N$ -point ECF,  $e_N^{(\beta)}$ , which differs from the definition (4.8) by a factor proportional to the  $p_T$  of the jet

$$e_N^{(\beta)} = \left( \frac{1}{p_{TJ}} \right)^N ECF(N, \beta) \quad (4.9)$$

Clearly,  $ECF(N, \beta) = 0$  when it is computed on a jet with fewer than  $N$  constituents. In addition to this,  $ECF(N+1, \beta)$  will be significantly smaller than  $ECF(N, \beta)$  if calculated on a jet containing only  $N$  subjets. This naturally leads to the consideration of the following ratio:

$$r_N^\beta = \frac{ECF(N+1, \beta)}{ECF(N, \beta)}. \quad (4.10)$$

The ratio  $r_N^\beta$  is useful to discriminate  $N$ -prong jets from  $N+1$ -prong jets. The  $D_2$  variable [25] is a variation on this ratio, specified to the case of 1-prong and 2-prong jets, and calculated using the normalized ECFs

$$D_2(\beta) = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3} \quad (4.11)$$

Due to the different color configurations, a large jet originating from the decay of a color singlet in the boosted regime will tend to exhibit a 2-prong substructure, and thus will tend to have a larger value of  $D_2$ . On the other hand, in the case of a color octet, the same decay will tend to have a 1-prong structure, leading to smaller values for the observable.

For the purpose of our study, we have chosen the value  $\beta = 2$ .

### 4.1.3 Color Ring

The final variable we have considered is known as the color ring [26]. To derive the observable, we must consider the boosted decay of a color singlet into two jets, which, at the parton level, correspond to (anti)quarks or gluons. We will consider the decay into two hard partons and the subsequent emission of a low-energy gluon.

The matrix element describing this transition is as follows:

$$|\mathcal{M}_S|^2 = -\mathbf{T}_\alpha \cdot \mathbf{T}_\beta \frac{n_a \cdot n_b}{(n_a \cdot k)(n_b \cdot k)}. \quad (4.12)$$

In this expression,  $n_a$  and  $n_b$  are the light-like vectors parallel to the hard jets,  $k$  is the impulse of the soft emission, and  $\mathbf{T}_i$  is the color operator. The dependence on  $\alpha_S$  has been neglected as it is not pertinent to the discussion. Latin indices are used to refer to kinematics while Greek indices refer to color.

We know that color must be conserved, as since we are dealing with a singlet decay, we must have

$$\mathbf{T}_\alpha + \mathbf{T}_\beta = 0. \quad (4.13)$$

This necessarily implies that  $\beta = \bar{\alpha}$ , and so

$$-\mathbf{T}_\alpha \cdot \mathbf{T}_\beta = \mathbf{T}_\alpha^2 = C_S \mathbf{1} \quad (4.14)$$

where  $C_S$  is Casimir operator for either the fundamental or adjoint representation of  $SU(3)$ , depending on the decay, and we have explicitly written the identity matrix.

Let us now consider the same matrix element for a background process, in which a color octet decays into two hard partons, with a subsequent soft emission in the boosted regime. Thanks to the boosted kinematics of the event, we can safely assume that the two hard partons are closer in angle to each other than to any other colored object, allowing for the factorization of the matrix element into two elements. The first describes the dipole formed by the hard partons and the emission of the soft radiation. The second contains additional contributions from the initial state of the event and from extra jets, if present. This term will again be neglected since it will lead to a constant. We can thus write

$$|\mathcal{M}_B|^2 = \sum_{\alpha, \beta} \left[ -\mathbf{T}_\alpha \cdot \mathbf{T}_\beta \frac{n_a \cdot n_b}{(n_a \cdot k)(n_b \cdot k)} - \mathbf{T}_\alpha \cdot \mathbf{T}_\gamma \frac{n_a \cdot \bar{n}}{(n_a \cdot k)(\bar{n} \cdot k)} - \mathbf{T}_\beta \cdot \mathbf{T}_\gamma \frac{n_b \cdot \bar{n}}{(n_b \cdot k)(\bar{n} \cdot k)} \right] \quad (4.15)$$

where  $\bar{n}$  is a light-like vector antiparallel to the system.  $\bar{n}$  has overall color  $\gamma$  which is fixed by the relation

$$\mathbf{T}_\alpha + \mathbf{T}_\beta + \mathbf{T}_\gamma = 0. \quad (4.16)$$

We can use this to simplify the expression (4.15), which becomes

$$|\mathcal{M}_B|^2 = \sum_{\alpha, \beta} \left[ -\mathbf{T}_\alpha \cdot \mathbf{T}_\beta \frac{n_a \cdot n_b}{(n_a \cdot k)(n_b \cdot k)} + (\mathbf{T}_\alpha^2 + \mathbf{T}_\alpha \cdot \mathbf{T}_\beta) \left( \frac{n_a \cdot \bar{n}}{(n_a \cdot k)(\bar{n} \cdot k)} + \frac{n_b \cdot \bar{n}}{(n_b \cdot k)(\bar{n} \cdot k)} \right) \right]. \quad (4.17)$$

The background matrix element has been written in terms of two color factors:

$$\begin{cases} C_B = -\mathbf{T}_\alpha \cdot \mathbf{T}_\beta \\ \tilde{C}_B = \mathbf{T}_\alpha^2 + \mathbf{T}_\alpha \cdot \mathbf{T}_\beta. \end{cases} \quad (4.18)$$

Again, the exact color factor depends on the process considered. We are interested in  $H \rightarrow b\bar{b}$  as our signal process, which fixes  $C_S = C_F$  and  $C_B = C_F - C_A/2$  and  $\tilde{C}_B = C_A/2$ . A decay into gluons, e.g.  $H \rightarrow gg$  would fix these constants differently.

It is a well-established fact that the optimal variable to discriminate these two configurations is monotonic in their ratio [27]. The ratio turns out to be

$$\frac{|\mathcal{M}_S|^2}{|\mathcal{M}_B|^2} \simeq \frac{(n_a \cdot \bar{n})(n_b \cdot k)}{(n_a \cdot n_b)(\bar{n} \cdot k)} + \frac{(n_b \cdot \bar{n})(n_a \cdot k)}{(n_a \cdot n_b)(\bar{n} \cdot k)} \quad (4.19)$$

where we have left out the color factors as they are just constants. If we assume the collinear limit and utilize the small-angle approximation, this expression simplifies to

$$\frac{|\mathcal{M}_S|^2}{|\mathcal{M}_B|^2} \simeq \frac{\theta_{ak}^2 + \theta_{bk}^2}{\theta_{ab}^2} \equiv \mathcal{O}. \quad (4.20)$$

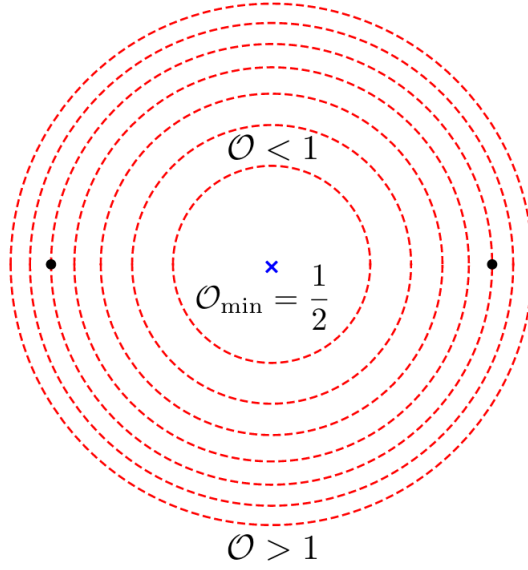


Figure 4.4: A geometric interpretation of the color ring. The two black dots represent the two hard jets, and the contour passing through them has value  $\mathcal{O} = 1$ . Depending on the direction of  $k$ , the  $\mathcal{O}$  can assume different values. If  $k$  is collinear to either  $a$  or  $b$ ,  $\mathcal{O}$  assumes its minimum value [26].

$\mathcal{O}$  is the definition of the jet color ring. This variable gets its name from its geometric interpretation, shown in Figure 4.4. Radiation from color singlets will tend to fall between the two jets, leading to values of  $\mathcal{O} < 1$ , while in the case of color octets, we will tend to have  $\mathcal{O} > 1$ .

Unfortunately, without any sort of momentum-weighting, (4.20) cannot be IRC safe. However, if we assume that the emission of the soft gluon leads to a distinct subjet within a larger jet, then IRC safety is recovered.

## 4.2 Simulation

The aforementioned variables were produced via Monte Carlo simulation. Using **MadGraph5\_aMC** v2.8.3.2 [28], we generated two hard processes:

$$pp \rightarrow Z(\nu_\ell \bar{\nu}_\ell) H(b\bar{b}) \quad (4.21)$$

$$pp \rightarrow b\bar{b}\nu_\ell \bar{\nu}_\ell. \quad (4.22)$$

The first process, hereby referred to as the signal, represents the most likely decay of the Higgs, a color singlet. The second process, or the background, contains all QCD diagrams which lead to the same final state as the signal. Figure 4.5 shows the Feynman diagrams constituting signal and the background. The outputs of the parton-level simulations were saved as Les Houches Event Files [29].

These hard processes were subsequently showered using two programs, **PYTHIA8** v8.305 [30] and **Herwig** v7.2.2 [31, 32]. These programs simulate the perturbative evolution of the event (e.g. radiation) as well as hadronization to produce particle-level events. The simulations included

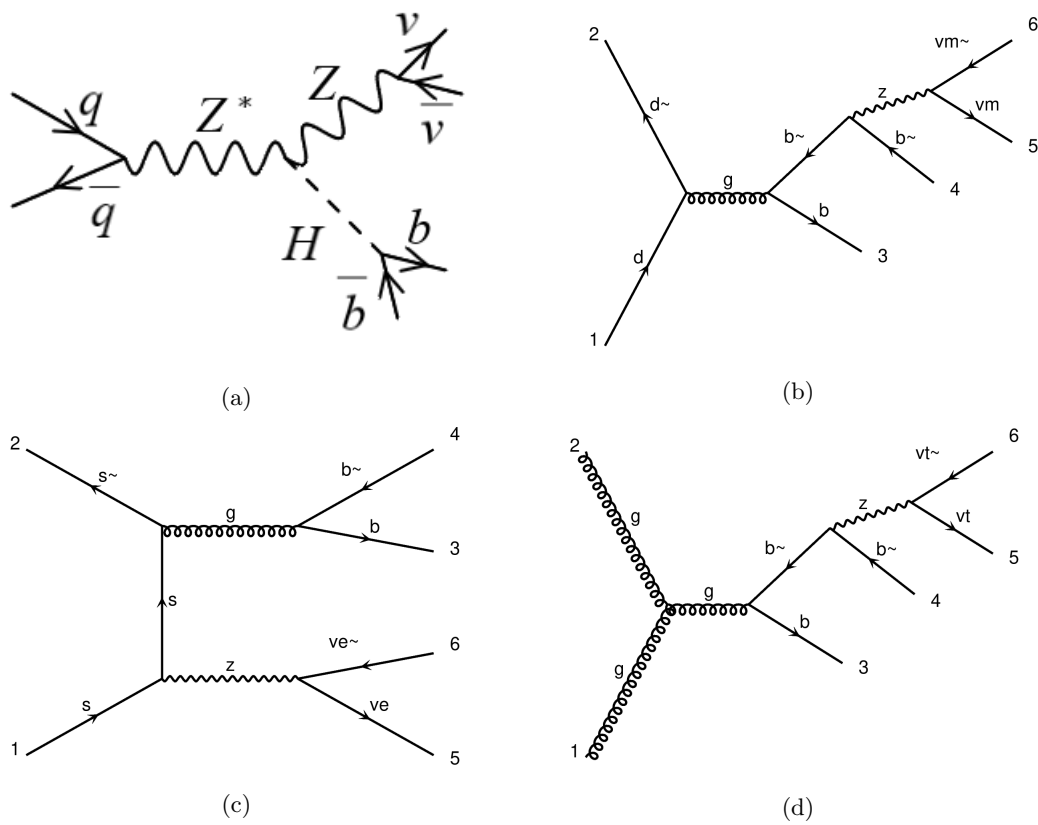


Figure 4.5: The Feynman diagrams constituting the signal (a) and background (b),(c),(d). mettere immagine migliore.

both Multi Parton Interactions and underlying events. The output of the `PYTHIA8` shower was saved as a `HepMC3` file, and the output of the `Herwig` shower was saved as a `HepMC2` file [33, 34]. `HepMC3` v3.2.2 and `HepMC2` v2.06.11 were used.

Since every program which simulates the parton shower necessarily includes its own systematic uncertainties, the shower in `Herwig` served as an independent check. We wanted to ensure that the results found were independent of the choice of parton shower, and thus that our results were reliable.

Finally, rather than simulating an entire detector, `DELPHES` v3.5.0 was used to perform a fast detector simulation [35, 36]. This allowed us to understand how our method would perform in reality without having to run the computationally expensive full simulation. From `DELPHES`, we extracted both the Monte Carlo truth of the event, containing the particle level information, as well as the reconstructed events including the detector effects.

The `DELPHES` simulation was run using a version of the ATLAS card modified to fit our analysis standards. The output was stored as a `ROOT` file, and `ROOT` v6.22/06 was used [37].

In total, we simulated 1 M signal and background events.

### 4.3 Analysis

We analyzed the simulated data in both the truth case and the reconstructed case. These cases differ in just one aspect: for the truth case, all stable particles with  $p_T > 0.5$  GeV were clustered into jets. In the reconstructed case, the jets were clustered using the simulated calorimeter towers and tracks. All Electromagnetic Calorimeter towers with energy  $E > 0.5$  GeV and Significance  $S > 2.0$  and all Hadronic Calorimeter towers with energy  $E > 1.0$  GeV and Significance  $S > 2.0$  are considered. Tracks are required to have  $p_T > 0.5$  GeV. The jets are clustered using `FastJet` v3.3.4 [38]. In a few cases, tracks are also included as jet constituents. The tracks are

We can now discuss the details of the analysis. First, the constituents are clustered into Large Jets with radius  $R = 1.0$ . For each event, we choose the Large Jet with the highest  $p_T$ . We only accept the event if the jet has  $p_T > 250$  GeV,  $|y| < 1.5$ , and invariant mass  $m_J \in [0, 500]$  GeV (in final analysis remove upper limit?). The restriction on  $y$  comes from experimental considerations: in practice, it is only possible to tag b-jets within this rapidity range.

We also cluster the constituents into smaller jets with radius  $R = 0.2$ . We identify those jets which distance  $\Delta R < 0.8$  from the selected Large Jet, and call these subjets. We then proceed to identify the b-subjets which originate from the b-partons, through a process known as b-labelling. We do this by first identifying the b-partons originating from the hard scattering in the event record. For each b-parton, we find the closest subjet. If the b-parton has  $p_T > 5.0$  GeV and the subjet is within  $\Delta R = 0.2$  from the b-parton, we label the subjet as a b-subjet.

For the event to be accepted, we require that there be two unique b-labelled subjets with  $p_T > 10$  GeV. The pull variables are calculated on these two b-labelled subjets, and  $D_2$  is calculated on the Large Jet. For the color ring to be defined, there must also be a third non-b subjet within  $\Delta R = 0.8$  from the Large Jet. In a majority of cases, this third jet is not present. To avoid discarding too many events, in these cases we assign a default value of  $\mathcal{O} = -1$  to the color ring. This allows for higher statistics, but also useful information to the machine learning algorithms, since these are events for which only the color ring is not defined.

Table 4.1 shows the percentage of events which passed the selections in all cases considered. As is clear, the cuts were more severe on the background events than the signal events in both the truth and reconstructed cases. Two cuts are significantly more important than the rest: the  $p_T$  cut on the Large Jet, responsible for about 60% of all cuts, and the rapidity cut on the Large Jet, responsible for another 10%.



Truth		Reco	
	Events Passed		Events Passed
Signal	20%	Signal	17%
Background	1.6%	Background	1.3%

Table 4.1: The approximate efficiencies of the analysis after all cuts were applied.

	AUC - Test Sample	
	Truth	Reco
BDT	0.818	0.777
NN - TMVA	0.817	0.773

Table 4.2: The areas under the ROC curves found.

Figure 4.6 shows the distributions of the 8 variables for the truth case. In Figure 4.7, the same distributions are shown in the reconstructed case. In both cases, clear differences between signal and background allow for discrimination of the two. Figures 4.8 and 4.9 show the comparison between the signal and background distributions of the eight variables in the truth case obtained after a shower in `PYTHIA8` and `Herwig`. There is little to no dependence on the parton shower. *devo ancora girare Herwig*

*Mettere anche sgn true vs sgn reco e bkg true vs bkg reco per mostrare effetto della simulazione del rivelatore? Potrei commentare che hanno gli stessi “general features”, pero’ e’ un’informazione che puo’ essere dedotta anche guardando le prime due figure*

## 4.4 Machine Learning Architecture

The output variables were used to train two different machine learning algorithms: a Neural Network (NN) and Boosted Decision Tree (BDT) from the `ROOT` Toolkit for Multivariate Analysis (TMVA) library [39]. In this section, we will limit ourselves to describing the architectures of these algorithms. For a more detailed discussion on how they work, see the Appendix [add hyperlink](#).

The NN is made of one hidden layer containing nine neurons. All signal and background events are passed to the network in one batch. The 60% of the data is used as a training sample and the remaining 40% as a test sample. The number of epochs is set to 600, and a Rectified Linear Unit activation function is used. The learning rate is set to 0.02.

The BDT is made up of 50 trees, with a maximum depth of 5 and minimum node size set to 2.5% *of the total number of events?* Adaptive boosting is used as a boosting model and the number of cuts is set to 80. Also in this case, we use a 60/40 train/test sample, and there is no downsampling.

*E’ forse piu’ chiaro riassumere tutto in una tabella anziche’ a parole?*

## 4.5 Results

We are now ready to present our findings. In Figure 4.10, the ROC curves from both the NN and BDT are shown for the truth and reconstructed cases. The areas under the ROC curves (AUCs) are reported in Table 4.2.

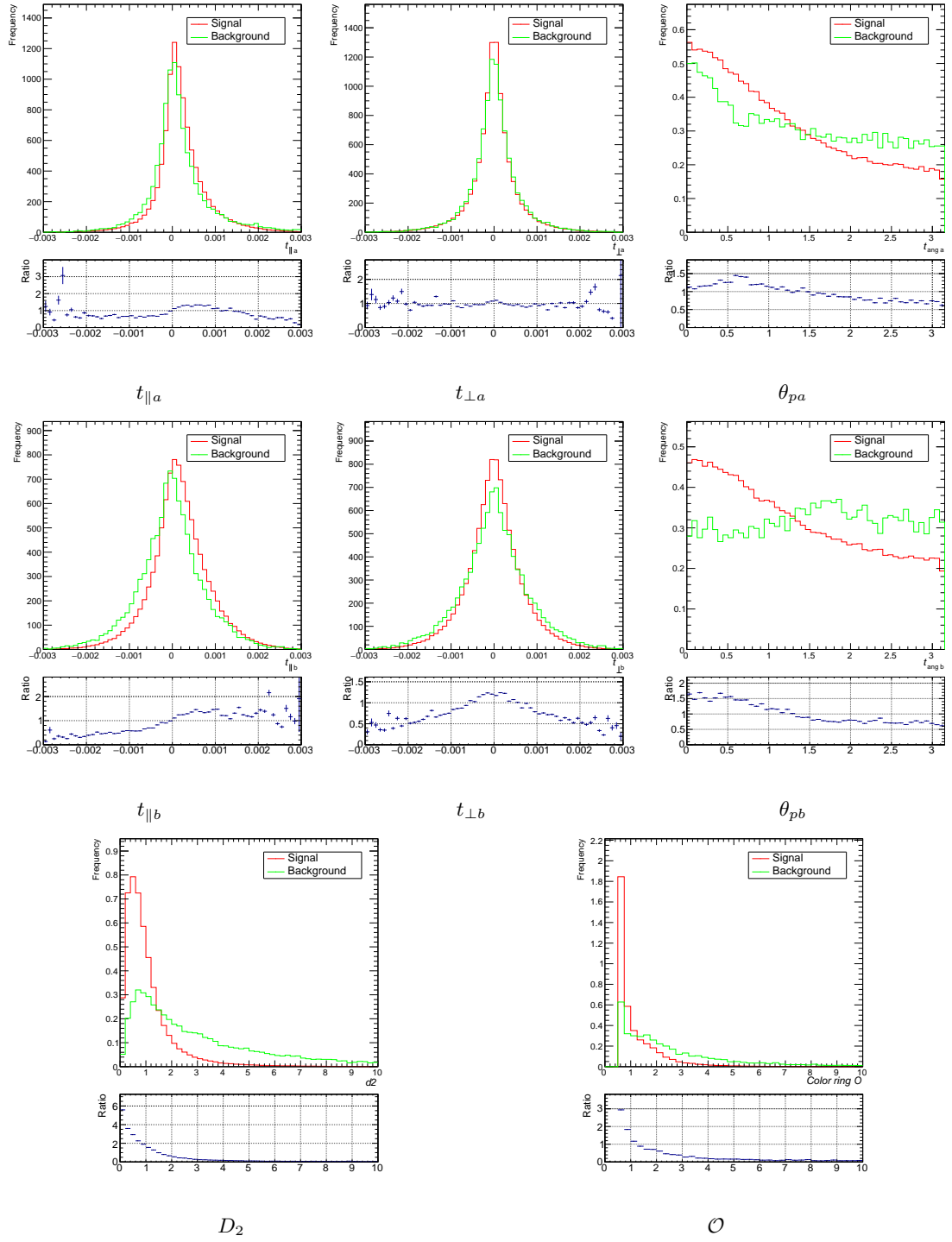


Figure 4.6: The distributions of the 8 variables after all selection cuts in the truth case together with the ratio between signal and background events.

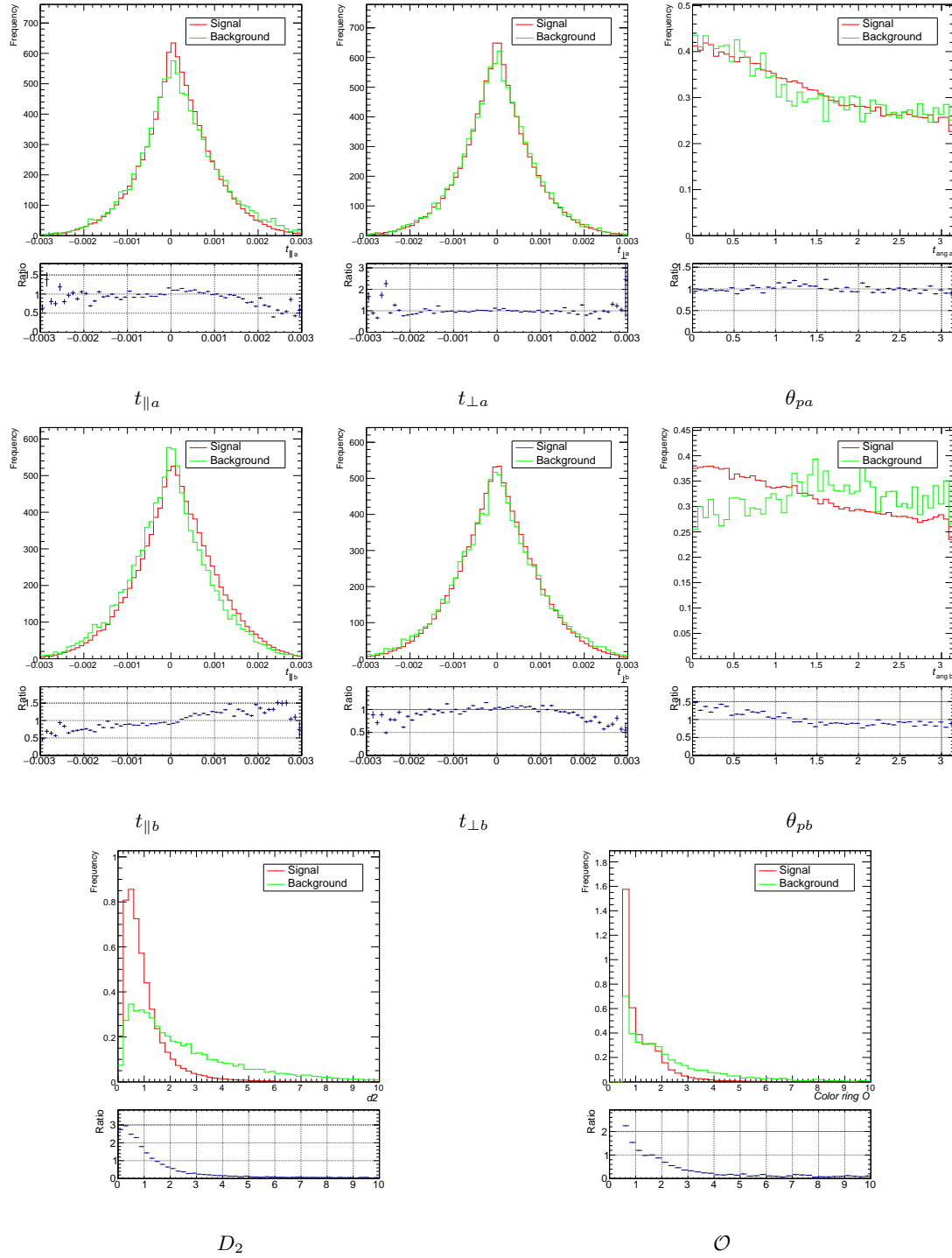


Figure 4.7: The distributions of the 8 variables after all selection cuts in the reconstructed case together with the ratio between signal and background events.

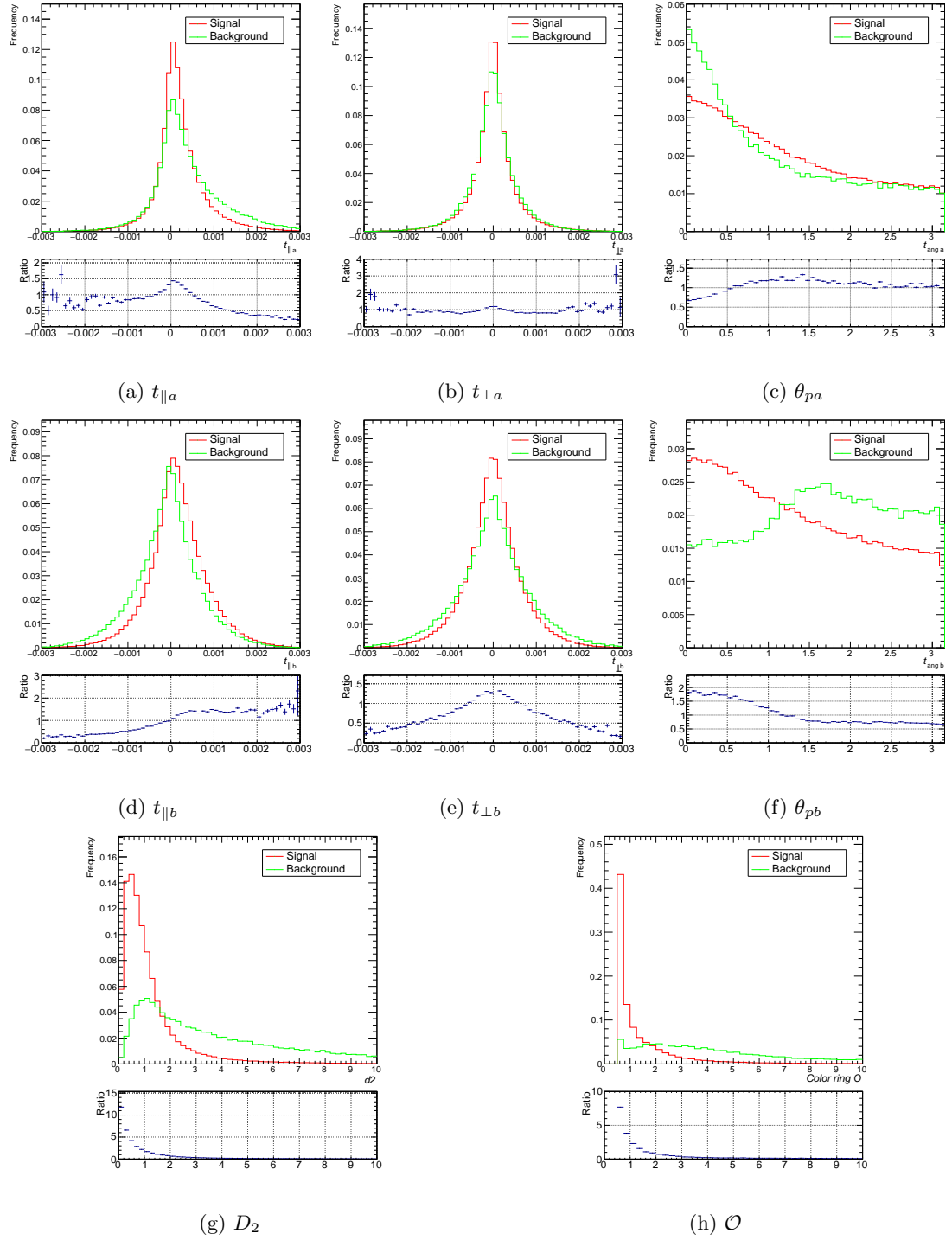


Figure 4.8: The signal distributions of the 8 variables in the truth case showered in PYTHIA8 and Herwig. Generare plot, togliere lettere dalla caption.

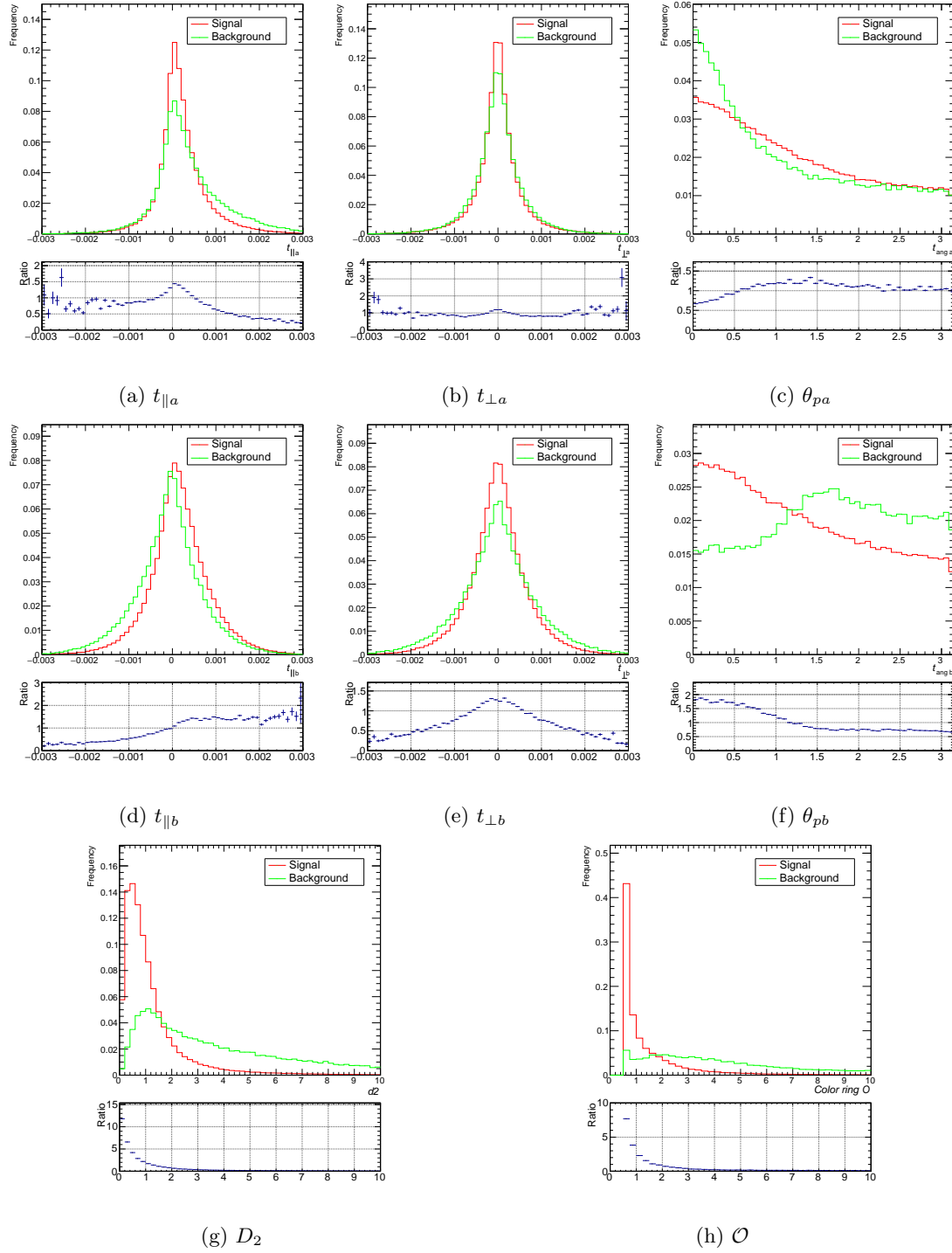


Figure 4.9: The background distributions of the 8 variables in the truth case showered in PYTHIA8 and Herwig. **Generare plot, togliere lettere dalla caption.**

Variable Ranking - Truth			Variable Ranking - Reco		
Rank	Var.	Importance	Rank	Var.	Importance
1	$D_2$	$3.743 \times 10^{-1}$	1	$D_2$	$3.536 \times 10^{-1}$
2	$\theta_{pb}$	$1.083 \times 10^{-1}$	2	$\mathcal{O}$	$1.116 \times 10^{-1}$
3	$t_{\parallel a}$	$9.508 \times 10^{-2}$	3	$\theta_{pa}$	$1.020 \times 10^{-1}$
4	$\theta_{pa}$	$9.506 \times 10^{-2}$	4	$t_{\perp b}$	$9.347 \times 10^{-2}$
5	$t_{\parallel b}$	$9.729 \times 10^{-2}$	5	$t_{\perp a}$	$9.137 \times 10^{-2}$
6	$\mathcal{O}$	$8.637 \times 10^{-2}$	6	$t_{\parallel a}$	$8.795 \times 10^{-2}$
7	$t_{\perp b}$	$8.151 \times 10^{-2}$	7	$\theta_{pb}$	$8.136 \times 10^{-2}$
8	$t_{\perp a}$	$7.206 \times 10^{-2}$	8	$t_{\parallel b}$	$7.862 \times 10^{-2}$

Table 4.3: The importance of the variables found by the BDT using the data from the truth case and the reconstructed case.

The AUCs show that in the truth case there both machine learning algorithms exhibit good signal/background discrimination. The results are unsurprisingly slightly worse for the reconstructed case due to the introduction of detector effects, but still quite good. This method is thus promising for application to real data.

We can also rank the variables based on their importance. The ranking for the BDT is reported in Table 4.3.  $D_2$  is the most discriminant variable, and the others are all of similar importance. For the reconstructed case,  $\mathcal{O}$  gains additional importance with respect to the pull variables. **Posso anche mettere ranking della NN**

Figure 4.11 shows the correlation between the variables for the signal and for the background in the reconstructed case **(add truth?)**. Unsurprisingly,  $t_{\parallel a}$  is correlated with  $\theta_{pa}$  and  $t_{\parallel b}$  is correlated with  $\theta_{pb}$ , as they are related to each other by definition. There is also some correlation present between  $D_2$  and  $\mathcal{O}$  in the background **pensare al motivo**.

Finally, Figures 4.12 4.13 shows the result of a Kolmogorov-Smirnov test [40] performed on the NN and the BDT to check for overfitting. Since the test probability  $\rho_{KS} > 0.01$  in all cases, no overtraining has occurred. **ancora da aggiornare**  
**aggiungere altri plot/test?**

## 4.6 Mass Bias

Since our goal is to develop a tagger purely sensitive to the color configuration of the decaying particle, ideally our machine learning algorithms would be insensitive to the invariant mass of system. This however is not the case.

In Figure 4.14a we show how different cuts on the output of the BDT depend on the invariant mass of the system in the truth case. We found that the sensitivity to the invariant mass was introduced mainly through the  $D_2$  observable. Figure 4.14b shows how the dependence of the output of the invariant mass is eliminated when the BDT is trained without this observable. The same effect was found in the reconstructed case. **show reco plots?**

Given the fact that the  $D_2$  observable is ranked as most important, the removal of this variable comes at the price of the efficiency of the algorithm **show ROC curves?**. Focusing again on the truth case, Table ... shows how AUCs of the NN and BDT change when the algorithms are trained without the  $D_2$  observable.

In an attempt to reduce the mass bias while still training the algorithms with  $D_2$ , we generated 100k new events for the signal and artificially varied the mass of the Higgs. Specifically, we generated 20k events with mass  $m_H = 25$  GeV,  $m_H = 50$  GeV,  $m_H = 75$ ,  $m_H = 100$  GeV and

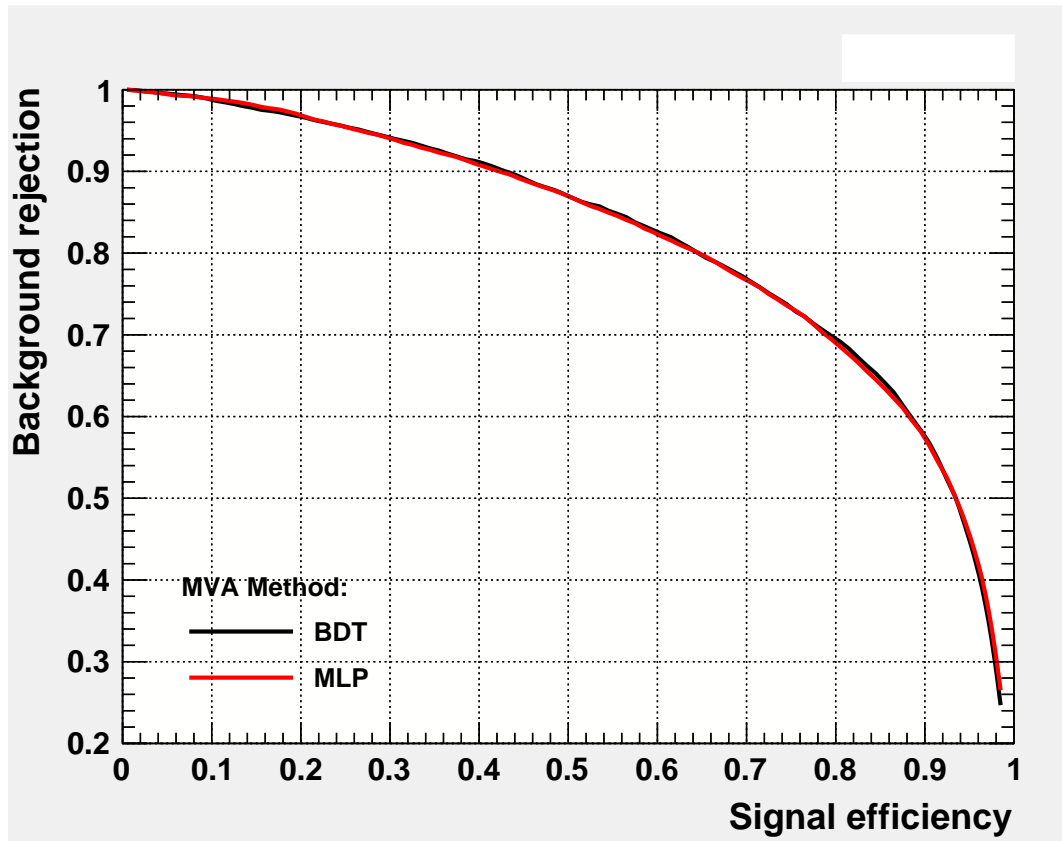
	AUC w/o $D_2$
BDT	0.775
NN - TMVA	0.773

$m_H = 125$  GeV. After showering these events in PYTHIA8 and running the analysis, we used the output to train the machine learning algorithms. For the training, 100k background events were also generated in the same manner as described in 4.2.

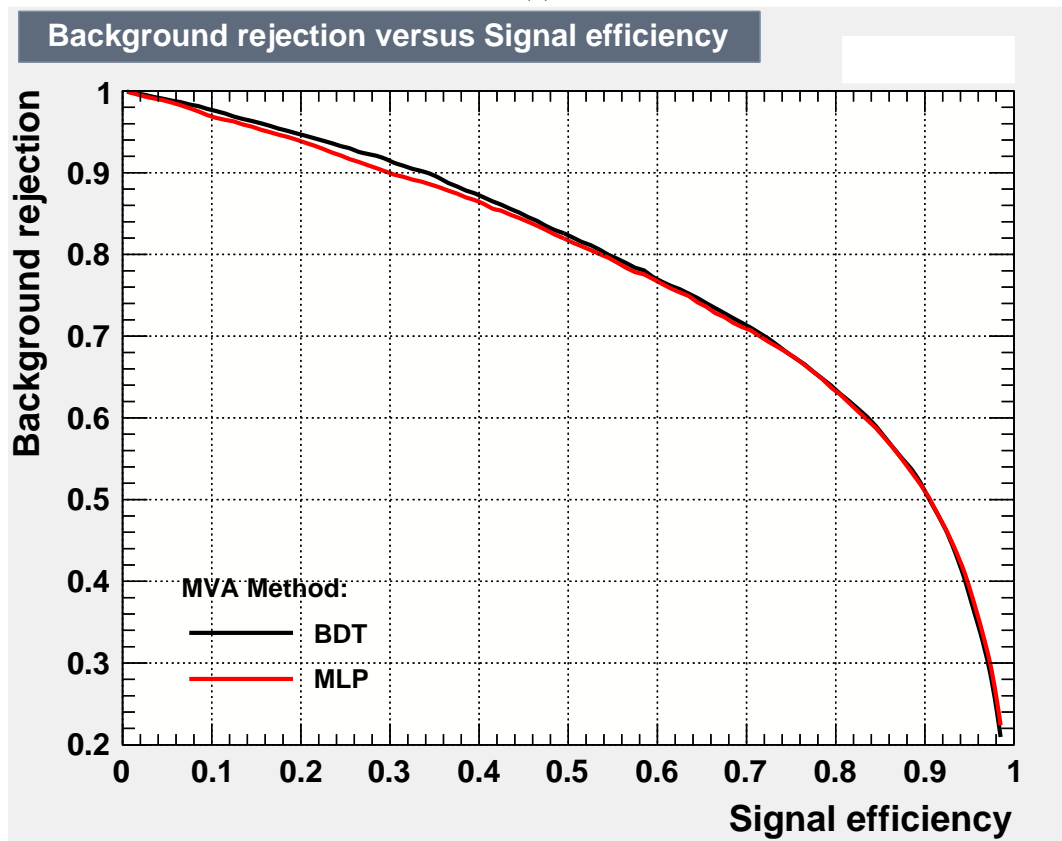
In Figure 4.15 we again show how the output of the BDT depends on the invariant mass of the system. The dependence on the mass is noticeably reduced. However, with the reduction of the mass bias came a reduction of efficiency. The AUC of the BDT reduced to **inserire cifra**, despite the fact that  $D_2$  was still included.

We have been able to find the reason for the reduction of the AUC. Figure 4.16 shows the distribution of the  $D_2$  variable in the truth case when generating events with and without varying the mass of the Higgs. When the mass of the Higgs is varied, the  $D_2$  distribution of the signal noticeably shifts towards that of the background. Effectively, the discrimination potential of the  $D_2$  observable is reduced, explaining the reduction of the AUC.

**migliorare immagine**



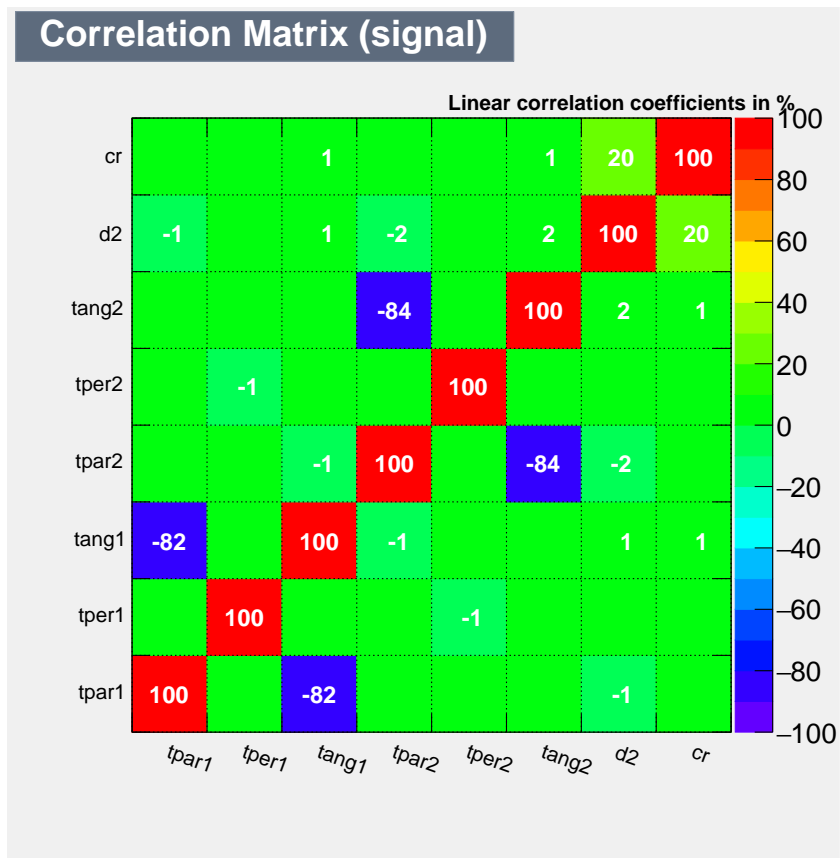
(a)



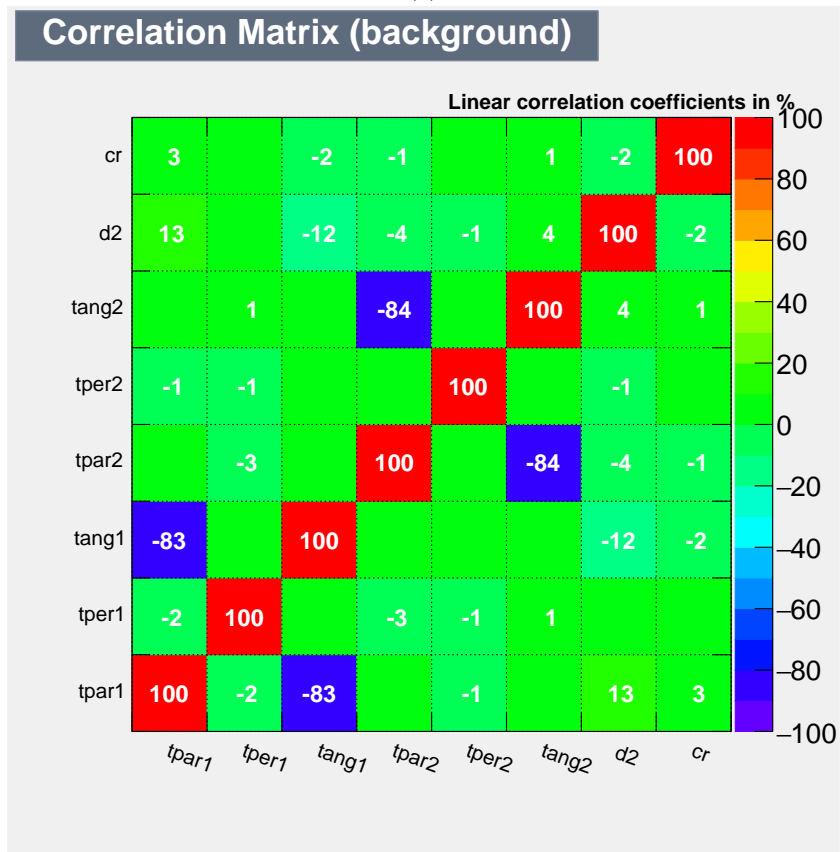
(b)

Figure 4.10: The ROC curves showing background rejection as a function of signal efficiency for the truth case (a) and for the reconstructed case (b) for both the BDT and the NN.



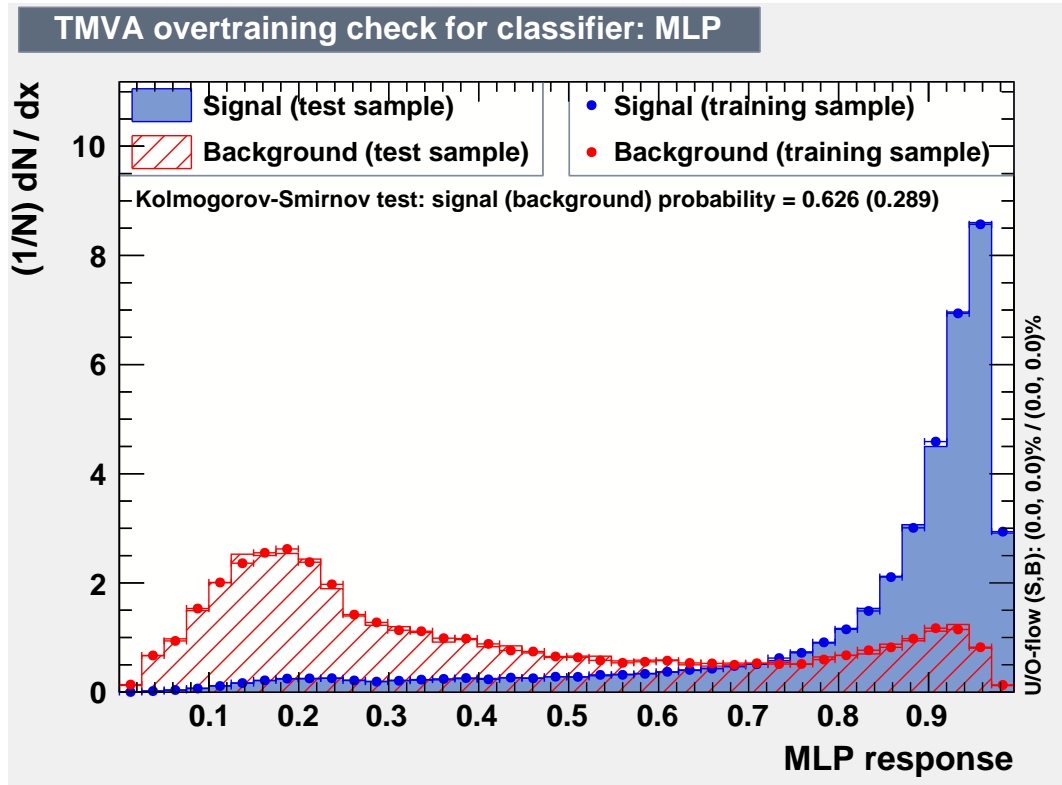


(a)

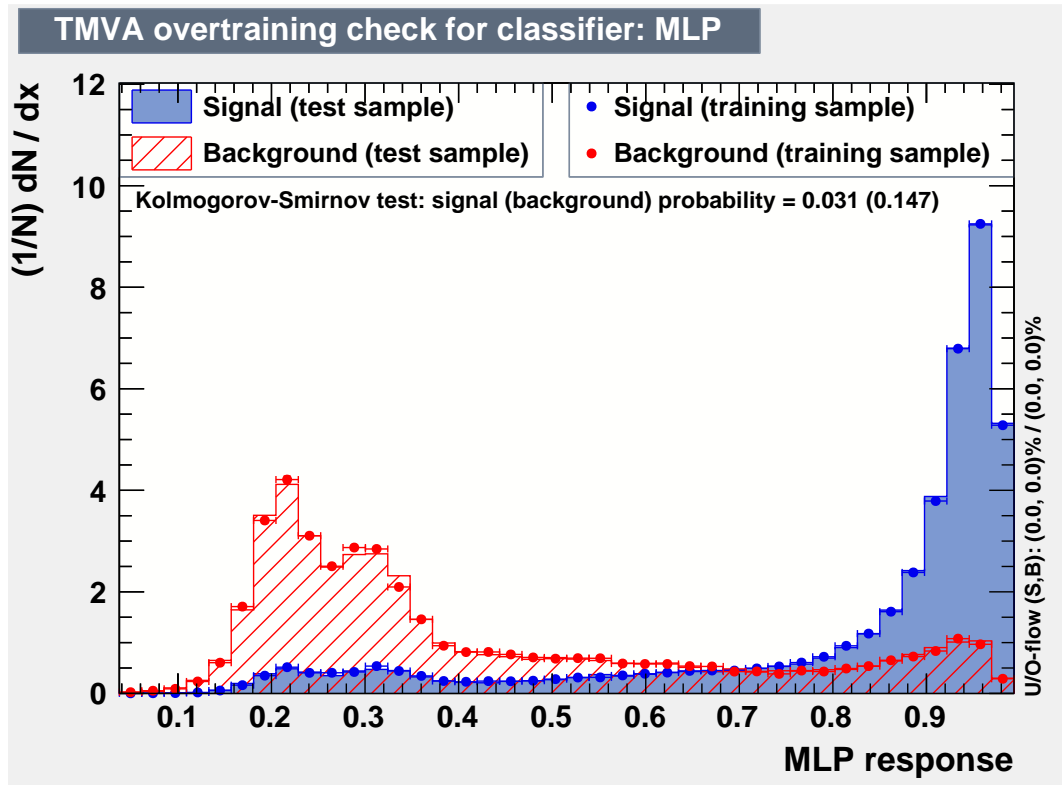


(b)

Figure 4.11: The correlation matrix found by the NN in the reconstructed case for the signal (a) and for the background (b).

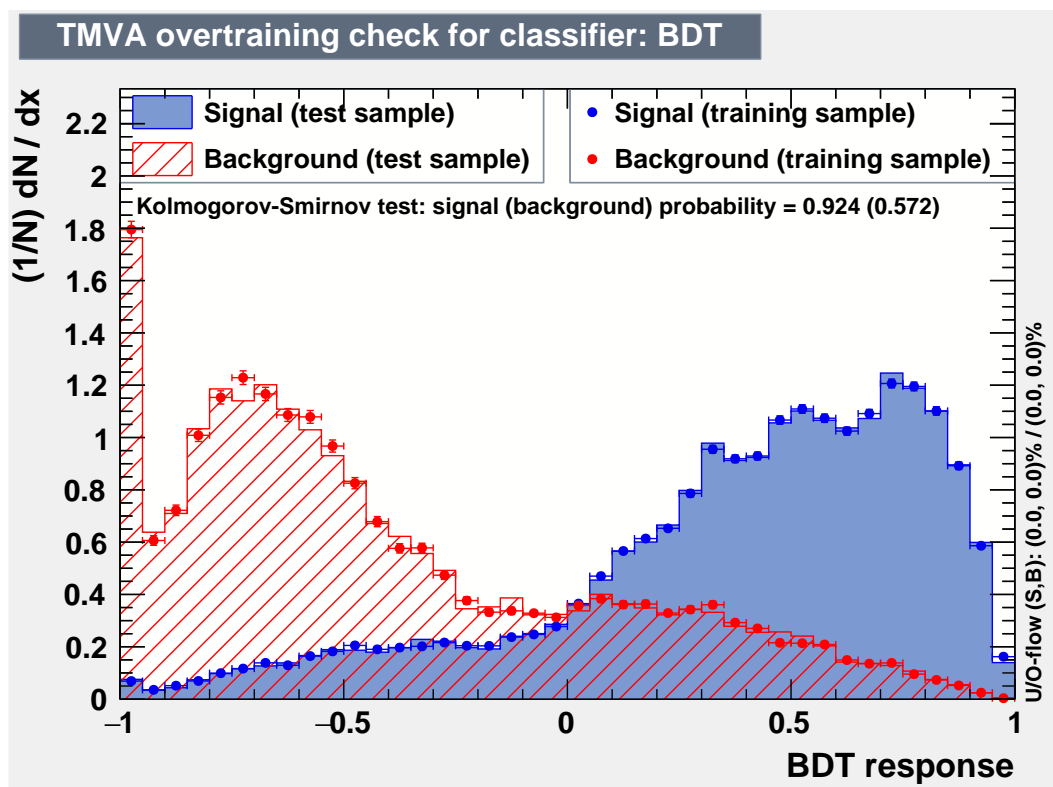


(a)

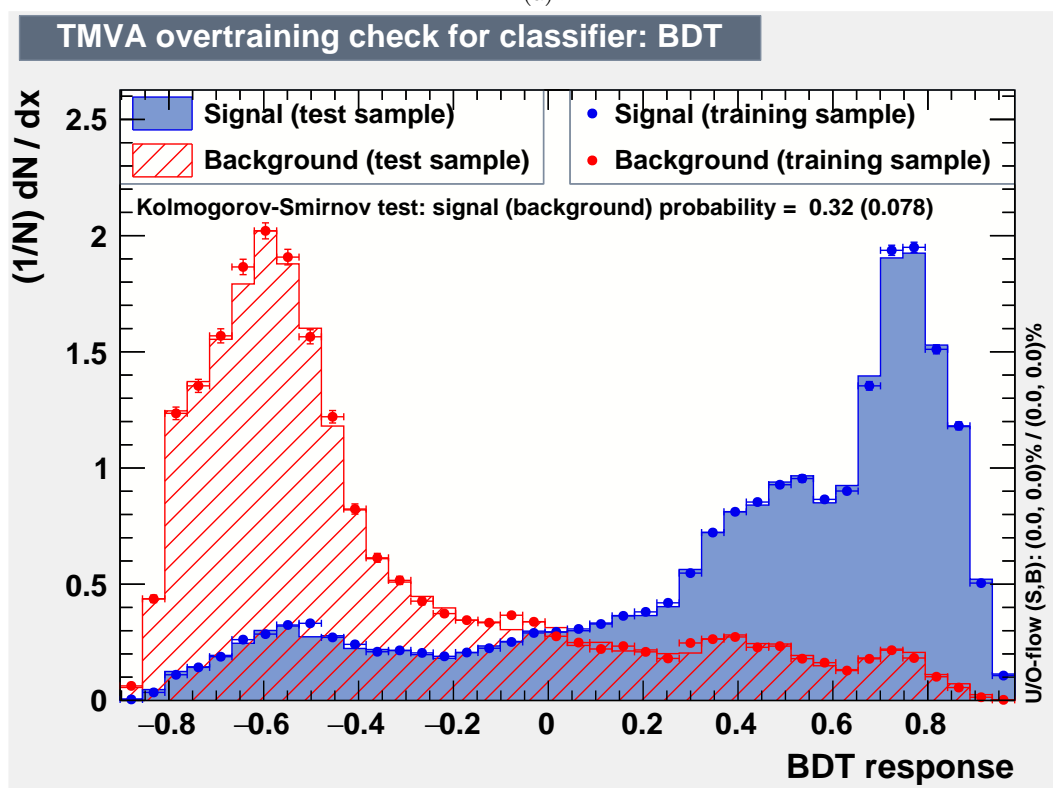


(b)

Figure 4.12: The result of a Kolmogorov-Smirnov test performed on the NN in the truth case (a) and the reconstructed case (b).

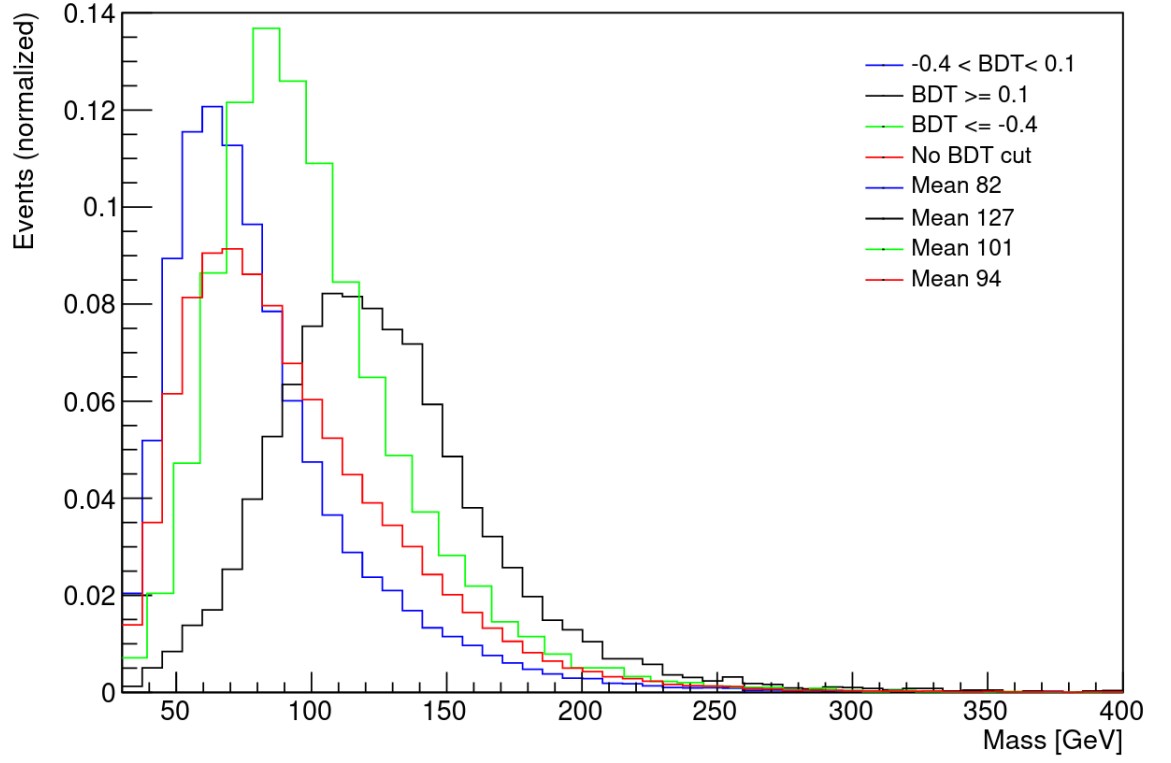


(a)

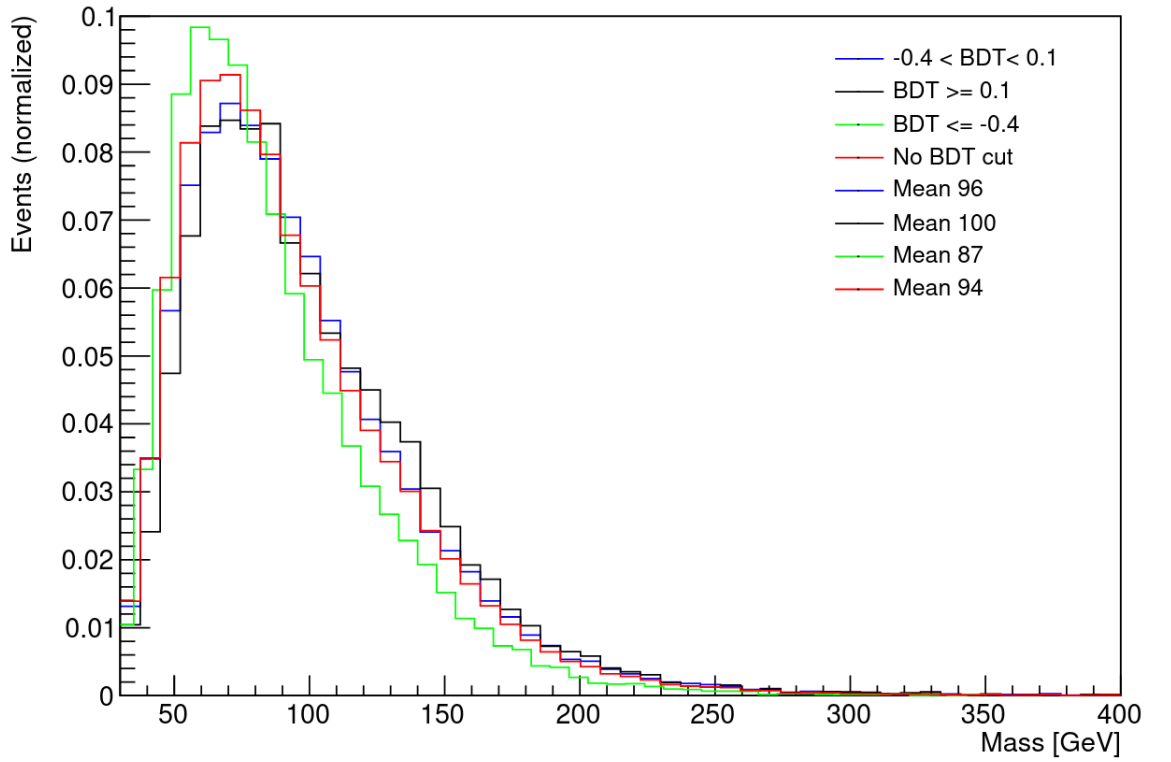


(b)

Figure 4.13: The result of a Kolmogorov-Smirnov test performed on the BDT in the truth case (a) and the reconstructed case (b).



(a)



(b)

Figure 4.14: The distribution of different cuts of the output of the BDT as a function of invariant mass of the system with the inclusion of  $D_2$  (a) and without the inclusion of  $D_2$  (b)

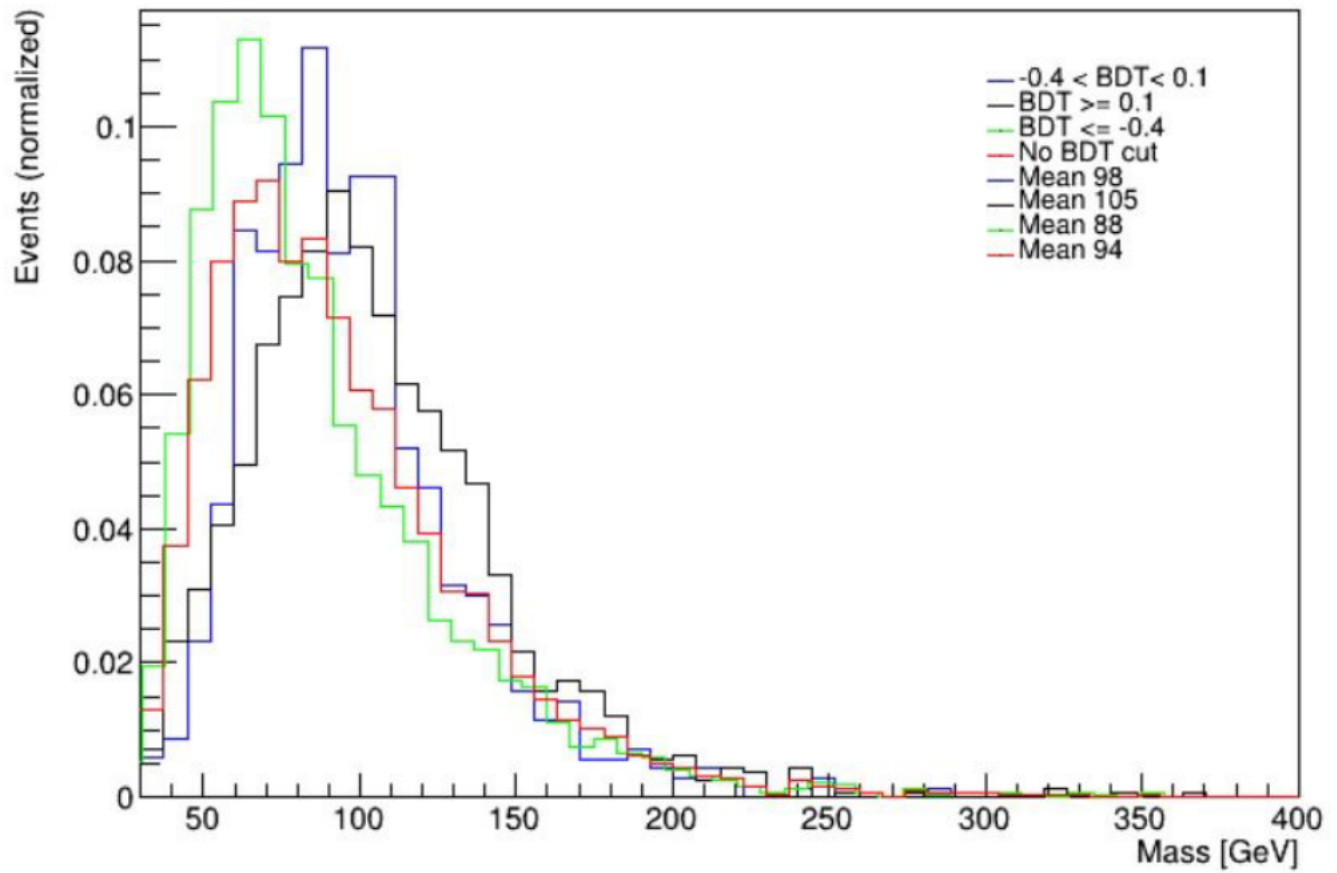


Figure 4.15: The distribution of different cuts of the output of the BDT as a function of invariant mass of the system when this has been varied when generating hard scattering events.

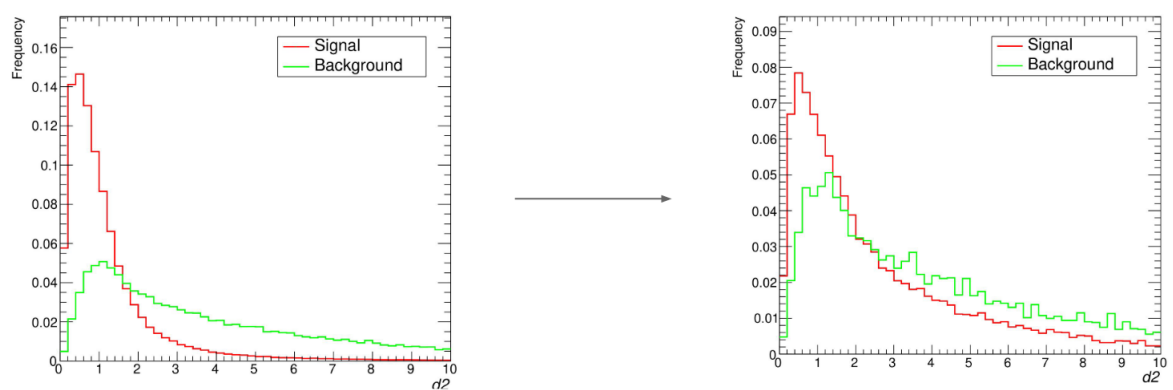


Figure 4.16: The effect of varying the invariant mass of the Higgs on the signal distribution of  $D_2$ .



# Conclusions and Outlook

In this thesis, we have illustrated a machine learning based method capable of distinguishing decays originating from color singlets from those originating from color octets. Given the high value of the AUC found when using simulated particle-level data and simulated data including detector effects, we can conclude that the method is not only effective in theory, but also shows promising prospects for application to experiment.

The ATLAS Xbb Tagger group has already expressed interest in this work. In the future, we plan to run a full simulation of the ATLAS detector, to then test the performance of the method using real data. At that point, the method can be integrated into the existing b-tagging algorithms used at ATLAS, and can provide additional information regarding the partonic origin of jets.

We also hope to improve this method by introducing other variables. In particular, we would like to include the Lund Jet Plane [41], a phase-space representation of jet substructure which has already been successfully measured at ATLAS [42].

In the coming years, the increased statistics as a result of the LHC luminosity upgrade promise to provide us with more insight on the properties of the Higgs. With luck, precision studies led by both theoretical and experimental efforts will lead to new discoveries.





## Appendix A

# Neural Networks



## Appendix B

# Boosted Decision Trees



# Bibliography

- [1] Wikipedia contributors. Standard model — Wikipedia, the free encyclopedia. [https://en.wikipedia.org/w/index.php?title=Standard\\_Model&oldid=1039528522](https://en.wikipedia.org/w/index.php?title=Standard_Model&oldid=1039528522), 2021. [Online; accessed 19-August-2021].
- [2] Wikimedia Commons. File:running coupling constants.svg — wikimedia commons, the free media repository, 2020. [Online; accessed 19-August-2021].
- [3] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.
- [4] M. Goldhaber, L. Grodzins, and A. W. Sunyar. Helicity of Neutrinos. *Phys. Rev.*, 109:1015–1017, Feb 1958.
- [5] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [6] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [7] Gerard 't Hooft. Renormalizable Lagrangians for Massive Yang-Mills Fields. *Nucl. Phys. B*, 35:167–188, 1971.
- [8] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, New York, 2013.
- [9] L. H. Ryder. *Quantum Field Theory*. Cambridge University Press, 6 1996.
- [10] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, Reading, USA, 1995.
- [11] Antonio Pich. The Standard Model of Electroweak Interactions. In *2010 European School of High Energy Physics*, 1 2012.
- [12] Gerry Bunce, Naohito Saito, Jacques Soffer, and Werner Vogelsang. Prospects for spin physics at rhic. *Annual Review of Nuclear and Particle Science*, 50, 08 2000.
- [13] Debasish Banerjee, Jajati Nayak, and Raju Venugopalan. Two introductory lectures on high-energy qcd and heavy-ion collisions. *Lecture Notes in Physics*, 785, 10 2008.
- [14] Jessica Metcalfe, Simone Seidel, Rouzbeh Allahverdi, Alice Bean, and Igor Gorelov. Development of planar and 3d silicon sensor technologies for the atlas experiment upgrades and measurements of heavy quark production fractions with fully reconstructed d-star mesons with atlas. 08 2021.

- [15] V.N. Gribov and L.N. Lipatov. Deep inelastic electron scattering in perturbation theory. *Physics Letters B*, 37(1):78–80, 1971.
- [16] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl. Phys. B*, 126:298–318, 1977.
- [17] George F. Sterman and Steven Weinberg. Jets from Quantum Chromodynamics. *Phys. Rev. Lett.*, 39:1436, 1977.
- [18] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- $k_t$  jet clustering algorithm. *JHEP*, 04:063, 2008.
- [19] Gavin P. Salam. Towards Jetography. *Eur. Phys. J. C*, 67:637–686, 2010.
- [20] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.
- [21] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13(05):P05011, 2018.
- [22] Jason Gallicchio and Matthew D. Schwartz. Seeing in Color: Jet Superstructure. *Phys. Rev. Lett.*, 105:022001, 2010.
- [23] Andrew J. Larkoski, Simone Marzani, and Chang Wu. Theory Predictions for the Pull Angle. *Phys. Rev. D*, 99(9):091502, 2019.
- [24] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. Energy Correlation Functions for Jet Substructure. *JHEP*, 06:108, 2013.
- [25] Andrew J. Larkoski, Ian Moult, and Duff Neill. Power Counting to Better Jet Observables. *JHEP*, 12:009, 2014.
- [26] Andy Buckley, Giuseppe Callea, Andrew J. Larkoski, and Simone Marzani. An Optimal Observable for Color Singlet Identification. *SciPost Phys.*, 9:026, 2020.
- [27] J. Neyman and E. S. Pearson. *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, pages 73–108. Springer New York, New York, NY, 1992.
- [28] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.
- [29] Johan Alwall et al. A Standard format for Les Houches event files. *Comput. Phys. Commun.*, 176:300–304, 2007.
- [30] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [31] Manuel Bähr, Stefan Gieseke, Martyn A. Gigg, David Grellscheid, Keith Hamilton, Oluseyi Latunde-Dada, Simon Plätzer, Peter Richardson, Michael H. Seymour, Alexander Sherstnev, and et al. Herwig++ physics and manual. *The European Physical Journal C*, 58(4):639–707, Nov 2008.

- [32] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76(4):196, 2016.
- [33] Andy Buckley, Philip Ilten, Dmitri Konstantinov, Leif Lönnblad, James Monk, Witold Pokorski, Tomasz Przedzinski, and Andrii Verbytskyi. The hepmc3 event record library for monte carlo event generators. *Computer Physics Communications*, 260:107310, 2021.
- [34] Matt Dobbs and Jorgen Beck Hansen. The HepMC C++ Monte Carlo event record for High Energy Physics. *Comput. Phys. Commun.*, 134:41–46, 2001.
- [35] S. Ovin, X. Rouby, and V. Lemaitre. DELPHES, a framework for fast simulation of a generic collider experiment. 3 2009.
- [36] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. *JHEP*, 02:057, 2014.
- [37] Fons Rademakers, Philippe Canal, Axel Naumann, Olivier Couet, Lorenzo Moneta, Vassil Vassilev, Sergey Linev, Danilo Piparo, Gerardo GANIS, Bertrand Bellenot, Enrico Guiraud, Guilherme Amadio, wverkerke, Pere Mato, TimurP, Matevž Tadel, wlav, Enric Tejedor, Jakob Blomer, Andrei Gheata, Stephan Hageboeck, Stefan Roiser, marsupial, Stefan Wunsch, Oksana Shadura, Anirudha Bose, CristinaCristescu, Xavier Valls, Raphael Isemann, and Kim Albertsson. root-project/root: v6.20/06, June 2020.
- [38] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [39] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS, ACAT*:040, 2007.
- [40] Frank J. Massey. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.
- [41] Frédéric A. Dreyer, Gavin P. Salam, and Grégory Soyez. The Lund Jet Plane. *JHEP*, 12:064, 2018.
- [42] Georges Aad et al. Measurement of the Lund Jet Plane Using Charged Particles in 13 TeV Proton-Proton Collisions with the ATLAS Detector. *Phys. Rev. Lett.*, 124(22):222002, 2020.